Uncertainty Quantification and Stochastic Model Updating for Structural Dynamics in Engineering

Thesis submitted in accordance with the requirements
of the University of Liverpool for the degree of
Doctor in Philosophy
by

Qiong Li Wang

July 26, 2019

To my daughter, KeYu

Abstract

Uncertainty Quantification and Stochastic Model Updating for Structural Dynamics in Engineering

Qiong Li Wang

This dissertation focuses on the uncertainty quantification and model updating for structural dynamics. In deterministic model updating methods, in order to improve efficiency, the traditional sensitivity method is improved by putting the second-order sensitivity into consideration for model updating. The variability inevitably exists due to different sources. They may arise from measurement noise, or to be inherent to the test structures variability such as manufacturing and material variability in structures. For such cases, the stochastic model updating methods may be considered to represent the uncertainty in the structural parameters. Bayesian model updating, combined with computationally efficient surrogate model approaches using advanced Markov Chain Monte Carlo (MCMC) methods is promising for some complex structure updating problems, which has an advantage of alleviating the computational cost for this real-time application tasks without losing accuracy.

The computationally efficient Gaussian process emulator can be used to replace the original large finite element models since their running time may prohibit any type of finite element model analysis, which require large evaluations of the finite element model to be performed. As long as the computational cost of running the complex model is decreased, the uncertainty quantification analysis and stochastic model updating methods become tractable, which can be used to study their performance.

Model discrepancy, defined as the difference between reality and the simulator output, as a very important modeling uncertainty is considered for model updating via Gaussian process emulation. It is found that using an informative prior distriburions for the calibration parameters, which can help to obtain the true updated posterior distribution of the calibration parameters.

Another aspect of using the Gaussian process emulator is that it has the ability to model some underlying empirical relations with confidence. Because this statistical surrogate model estimates not only the expectation but also the variance. For some very important empirical relations (e.g. the empirical relations of dynamic increase factor (DIF) of steel skeleton reinforced concrete (SSRC) specimens under impact loadings), using the trained Gaussian process emulator to approximate this relationship is more convincing.

Declaration

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed …………………(Candidate)

Date…………………….

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by explicit references. A bibliography is appended.

Signed …………………(Candidate)

Date…………………….

Acknowledgments

I would like to express my deepest gratitude to my supervisors, Dr. Francisco Alejandro Diaz De la O, and Professor John E. Mottershead. Without their continued encouragement, guidance, and support, it impossible for me to complete the work.

I also want to express my deepest gratitude to their care, that kept me moving on when I once felt desperate and hopeless after heard the physical conditions of my newborn daughter.

I am grateful to have had the opportunity of studying at the Institute for Risk and Uncertainty, University of Liverpool. My thankfulness also goes to Dr. Edoardo Patelli for providing the DLR AIRMOD industrial example and data. I also would like to thank Ander K. Gray for having an opportunity to work together.

Finally, I would like to thank my wife, Wan Qi Guo. When I was doing my Ph.D. studies, she gave me a lot of encouragement and supported me emotionally. I am also grateful to my family and my friends both in Liverpool and China for their support, help, and care.

Contents

List of Figures

List of Tables

Symbols and acronyms

λ : Lagrange multiplier

Δ : Perturbation of the system matrices.

$\mathbf{Z}_m$: Measured outputs

$\mathbf{Z}_i$: Predicted outputs

$\mathbf{S}_i$: First-order sensitivity matrix

$\boldsymbol{\varepsilon}_z$: Error vector

$\mathbf{M}$: Symmetric mass matrix

$\mathbf{K}$: Symmetric stiffness matrix

λ_i : i-th eigenvalue

ϕ_i : i-th mode shapes

$\boldsymbol{\omega}_\varepsilon$: Symmetric weighting matrix

$\mathbf{H}_{ij}$: Second-order sensitivity matrix

$\mathbf{T}$: Transformation matrix

$\mathbf{D}$: Modal properties obtained from experiments

$\boldsymbol{\Sigma}_\theta$: Covariance matrix

$h(\cdot)$: Real-valued function

$\boldsymbol{\beta}$: Vector of unknown coefficients

$Z(\mathbf{x})$: Gaussian process with zero mean and covariance

σ^2 : Scalar parameter

$\mathbf{C}$: Correlation matrix

u : Parameter controls the typical amplitude

τ : Parameter controls the lengthscale of variation.

Γ : Gamma function

E : Young's modulus

A : Cross-section area of pressure bars

C_0 : Elastic wave velocity

m : Inducing variables

$\delta_{\mathbf{xx'}}$: Kronecker delta

$y^e(\mathbf{x})$: Physical experiment response

$\eta^c(\mathbf{x})$: Computer model response

$\delta(\mathbf{x})$: Discrepancy function

SD: Standard deviation

KL: Kullback-Leibler divergence

VSGP: Variational sparse Gaussian process

EM: Expectation-Maximization

SPGP: Sparse pseudo-input Gaussian process

RPC: Reactive powder concrete

GA: Genetic algorithm

MH: Metropolis-Hastings

TMCMC: Transitional Markov Chain Monte Carlo

ME: Maximum entropy

PDF: Probability density function

MOO: Multi-objective optimization

GP: Gaussian process

GPE: Gaussian process emulator

LHS: Latin hypercube sampling

FRC: Fiber reinforced concrete

SSRC: Steel skeleton reinforced concrete

SHPB: Split Hopkinson pressure bar

DIF: Dynamic increase factor

IPE: Individual prediction errors

FE: Finite element

MCMC: Markov Chain Monte Carlo

FRF: Frequency Response Functions

MAC: Modal assurance criteria

1 Introduction

1.1 Motivation and background

The dynamic analysis of the engineering structures is increasingly becoming complex and computationally intensive and it will continue to do so to meet the challenges of the world to dynamic analysis. The interest in the vibration properties of the engineering structures is growing all the time because nearly all structures are subject to the vibration of one form or another, which is always considered as undesirable under some conditions.

Today, for the dynamic design and analysis of complex structures, it is more important than ever to have reliable dynamic mathematical models. The finite element (FE) method [1] is a numerical analysis technique for obtaining approximate solutions to a wide variety of engineering problems. Due to its diversity and flexibility, it has become a well-established procedure in industry. In the FE method, the complex structures are divided into discrete areas or volumes known as 'elements' with simple and standard geometrical shapes whose dynamic behavior is known. Therefore, the original structure can be rebuilt form such 'elements' to understand its overall dynamic behavior. Under different types of loadings, the FE methods can provide predictions of the dynamic behavior of the complex structures. This can help engineering researchers to detect any deficiency in the structure from the early stage of the design process and consequently reduce the cost of design.

Although the FE method is widespread used for applications in engineering, it is always an approximate approach, and there are different sources of errors for this approximation. These errors are mainly divided into three groups, namely model structure errors, model parameter errors and model order errors. In order to reduce these errors in FE model predictions, FE model updating methods are described as follows.

FE model updating methods can be defined as the adjustment of the FE model in light of a measured vibration test. After adjustment, the updated FE model is expected to represent the dynamic behavior of the structure more accurately. FE

model updating methods [2, 3] are well established and applied to industry structure updating problems. These methods are mainly divided into three groups, namely the direct methods, the iterative methods, and the computational intelligence methods.

In the direct methods [2], a 'representational' model including the updated global mass and stiffness matrices are obtained that is capable of reproducing the measured data exactly. The major disadvantage of these methods is that the updated mass and stiffness matrices have little physical meaning and cannot be related to physical changes to the FE in the original models. In the iterative method [2], an iterative process based on sensitivity analysis is required in order to minimize an objective function which consists of the difference between predicted modal or FRF (Frequency Response Function) data and their measured counterparts. In the iterative method, the problems of convergence and ill-conditioning of the matrices [4] are needed to solve.

In fact, for the FE model updating problems, the experimental variability inevitably exists due to different sources [17]. They may arise from measurement noise, or to be inherent to the test structures variability such as manufacturing and material variability in structures. For such cases, different stochastic model updating methods may be considered to represent the uncertainty in the structural parameters. They may generally be categorized into two categories namely probabilistic stochastic model updating methods and non-probabilistic stochastic model updating methods. In non-probabilistic stochastic model updating methods, the interval model updating methods [74, 75] and fuzzy model updating methods [77-80] are the most representative methods. For probabilistic stochastic model updating method, the minimum variance method, the probabilistic perturbation method, and Bayesian model updating methods are the representative methods.

When applying the probabilistic perturbation methods to complex structures in industry, the covariance updating is generally ill-conditioned and requires regularization [4]. For Bayesian model updating, the regularization is done automatically through using the output and parameter covariance matrices (present in the likelihood function and the prior distribution). Adopting the Bayesian FE model updating to complex structures (e.g. the AIRMOD structure in chapter 5) is the main focus of this thesis. The main disadvantage of Bayesian model updating is computationally intensive. But now these methods are tractable by using

advanced Markov Chain Monte Carlo (MCMC) methods [67-70], surrogate model techniques [83-88] and parallelized system strategies [161].

Using high-fidelity surrogate models to replace the full FE model is another aspect of interest for Bayesian FE model updating in order to alleviate the computation cost without losing the accuracy. A surrogate model, also known as a metamodel [82] is an approximation to the output of the simulator, which is much less computationally expensive to use. Without generality, given n inputs, for which the simulator has been evaluated, there will be n output values. The metamodel will use this information to capture the behaviors of the original simulators. The selection of n inputs, referred to as the training samples, which is important so as to maximize the information about the simulator with a minimal computational effort. Once built, the surrogate model can be used to infer the outputs of the original simulator at input configuration not included in the training samples. There are many different types of metamodels. (e.g. Polynomial chaos expansion methods [88], Response surface methods [83], Radial basis function methods[86], and Support vector machine methods [87]). Another surrogate model is the Gaussian process emulation (GPE). This emulator is a statistical surrogate model, which has the ability to quantify the uncertainty in any predictions it makes about the overall behaviors of the simulator. For the advantages of GPEs outlined in chapter 4, the GPEs are amongst the most widespread used metamodels in engineering.

The GPE also has the ability to model some underlying empirical relations with confidence. In chapter 4, the behaviors of steel skeleton reinforced concrete (SSRC) materials are investigated by quasi-static tests and Split Hopkinson Pressure Bar (SHPB) test, respectively. The relations of the dynamic increase factor (DIF) for compressive strength and Young's modulus of SSRC specimens can be constructed using GPE, which estimates not only the expectations but also the variances. Quantifying the uncertainty of this important empirical relations is important since it can be used to estimate DIFs of SSRC in numerical prediction of SSRC structure responses to other high-rate loadings.

The inverse problem (e.g. Bayesian model updating) in uncertainty analysis, in which the experimental data are used to learn about sources of modeling uncertainties especially for calibration parameters and model discrepancy

functions (difference between reality and the simulator output) is another much challenging model updating problems, which is described in chapter 6. By an illustrative example of a cantilever beam to do the uncertainty quantification of model discrepancy, it seems that using an informative prior distribution for the calibration parameter helps to obtain the true updated posterior distribution to it.

1.2 Aim and objectives

The aim of this thesis is to model, propagate, update and identify the sources of modeling uncertainties for structural dynamics in engineering. By considering existing methodologies for model updating, various strategies are employed in terms of computational efficiency together with their level of accuracy for prediction. The application of the surrogate model, namely GPE, is the main focus of this thesis to quantify the uncertainty in any predictions to overall behavior of the complex mathematical models or underlying empirical relations. This work is an attempt:

1. To critically review the existing deterministic model updating methods, especially for the sensitivity method and find out some strategies to improve the deterministic mode updating methods' efficiency and accuracy.
2. To critically review the existing stochastic model updating methods, especially for the probabilistic perturbation method and Bayesian model updating methods and find new strategies to develop the Bayesian model updating in order to cope with industrial model updating problems.
3. To adopt surrogate modeling techniques especially GPEs to do the uncertainty quantification tasks in order to alleviate the burden of computations and model some important empirical relations (e.g. DIFs of SSRC materials under different impact loadings.).
4. To develop a Bayesian model updating framework to do stochastic model updating, which account for the computational cost of the FE models by using the computational efficient GPEs.
5. To examine the existing most comprehensive model updating formulations considering all sources of uncertainties and provide a better understanding of the issue of identifiability for calibration parameters.

1.3 Outline of the thesis

Chapter 2 gives a literature review of the deterministic model updating methods. The advantages and disadvantages of these methods are discussed. An extension to the classic sensitivity method is proposed and the performance of this method is discussed by two numerical examples.

Chapter 3 provides a literature review of the stochastic model updating methods. A detailed description of probabilistic stochastic model updating methods including probabilistic perturbation method and Bayesian model updating method is given. A brief description of non-probabilistic stochastic model updating methods including interval model updating methods and fuzzy model updating method is also described. A comparative study on the performance of probabilistic perturbation method and the Bayesian model updating method is provided.

Chapter 4 describes the background theory for the surrogate model namely Gaussian process emulator (GPE). The GPE model construction process, the selection process of covariance functions and hyperparameter estimation, the surrogate model training and diagnostics are described. An analysis of steel skeleton reinforced concrete materials under quasi-static tests and SHPB test is carried out and the DIFs relations of SSRC materials is discussed.

Chapter 5 gives a literature review of computational efficient GPEs. The Bayesian FE model updating combined with the computational efficient GPEs. and using a parallel computing strategy is carried out to the AIRMOD structure. The updating results are given and discussed.

Chapter 6 provides the modular Bayesian approach to do the model updating by considering all sources of uncertainties namely parameter uncertainty, model discrepancy, code uncertainty, and experimental uncertainty. The GPEs are used to model the simulator model and discrepancy function, respectively. An illustrative example of the cantilever beam is used to demonstrate the application of this approach. In addition, the limitation and challenges of this approach are also discussed.

Chapter 7 gives the conclusions of the research and provide suggestions for future

work.

2 Deterministic finite element model updating

2.1 Introduction

In the design of structures, making a comparison between measured datasets from prototype structures and predicted results from corresponding numerical models is very important. When the comparison process is carried out, the quality of the numerical models can be evaluated. If the results taken from numerical models agree with their experimental counterparts, the numerical models can be used for the purpose of the structural design. However, in most cases, this agreement cannot be achieved due to many kinds of errors in the numerical model, e.g. simplifications of the structure, erroneous geometrical shape assumption, discretization errors in FE mesh process and among many others. These errors can be effectively quantified through different methods in model updating, such as least squares estimators, total linear least squares, singular value decomposition and among many others in [2, 3]. The model updating methodologies [2] mainly deal with model parameter errors. In the model updating process, inaccurate parameters are selected by the analyst and corrected through valuable measured datasets. The measured datasets often include modal data such as natural frequencies, mode shapes and the Frequency Response Functions (FRFs), obtained from the experimental modal analysis [5, 6].

In order to evaluate the closeness of the experimental model and corresponding numerical counterparts, the comparison process is often carried out in the modal domain using the natural frequencies and the mode shape vectors. The modal assurance criteria (MAC) [7] is a statistical indicator that is most sensitive to large differencs and relatively insensitive to small differences in mode shapes. This yields a good statistical indicator and a degree of consistency between mode shapes. The MAC is often used to pair mode shape taken from numerical models with those obtained experimentally. The MAC indicator between a measured mode Φ_{mi} and a numerical (analytical) mode Φ_{nj} is defined as

$$MAC_{ij} = \frac{|\Phi_{mi}^T \Phi_{nj}|^2}{(\Phi_{nj}^T \Phi_{nj})(\Phi_{mi}^T \Phi_{mi})} \qquad (2.1)$$

The MAC value is between 0 and 1. The value of 1 means a perfect correlation and 0 indicates no correlation between numerical and experimental mode shapes.

In this chapter, a brief description of deterministic model updating methods including direct methods, iterative methods, and computational intelligence methods are given. A novel and efficient iterative method based on the classic sensitivity method is then proposed and its performance is evaluated in numerical examples. Finally, some useful conclusions are presented in the chapter summary.

2.2　Direct methods

The direct methods detailed described in [2] are one of the earliest methodologies employed for FE model updating. Three main direct methods defined as Lagrange multiplier method, optimal matrix method, and control theory based method are briefly described.

The Lagrange multiplier method is an optimization technique deals with minimizing a function subjected to exact constraints on the independent variables. It is implemented by minimizing a constrained objective function, where the constraints are imposed by Lagrange multipliers [8].

The optimal matrix methods use analytical rather than numerical solutions to obtain matrices from the damaged systems. They are formulated using Lagrange multipliers and perturbation matrices and the optimization problem is used to minimize

$$O(\Delta \mathbf{M}, \Delta \mathbf{C}, \Delta \mathbf{K}) + \lambda C(\Delta \mathbf{M}, \Delta \mathbf{C}, \Delta \mathbf{K}) \tag{2.2}$$

where O is the objective function, λ is the Lagrange multiplier, C is the constraint equation and Δ is the perturbation of the system matrices.

The direct method consists in using different perturbation strategies until the differences between the FE model results and the measured counterparts are minimized in model updating. Baruch and Bar [9], Kabe [10] formulated Eq. (2.2) by minimizing the Frobenius norm of the error. Zimmerman et al. [11] successfully applied this method for model updating with incomplete measured modal data.

The control theory based methods are also called eigenstructure assignment approaches based on control system theory. An updated FE model is that with eigenstructure which is obtained from measured datasets. Zimmerman and Kaouk [12] applied this approach to update a FE model of cantilever beam based on modal

properties successfully. When applying this technique, the reduction/expansion approaches (static reduction [13, 14], dynamic reduction [15]) should be considered to make sure the mode shapes and frequency response functions are expanded to the size of the FE model or the mass and stiffness matrices of the FE model are reduced to the size of the measured data.

The direct methods possess the ability to reproduce the exact experimental datasets without iteration, which makes these algorithms computational efficient. But the disadvantages are still obvious, such as the updated mass and stiffness matrices have little physical meaning and cannot be related to physical changes to the finite elements in the original models and also lack of node connectivity.

2.3 Iterative methods

Due to the drawbacks mentioned in the direct methods, the iterative methods in [2] were developed. These methods used penalty functions involving the mode shapes and eigenvalue data: often the sum of squares of the differences between the measured and estimated eigenvalues are used. A set of parameters are iteratively adjusted to minimize penalty functions, which are generally non-linear functions of the parameters. If the change in the parameters between successive iterations is small, a good estimate of the modal model is available. Widespread iterative method called sensitivity method is described in detail.

In the sensitivity method, the experimental model is treated as a perturbation in updating parameters to the initial numerical model. The perturbation equations (penalty functions) generally use a truncated Taylor series expansion after the linear term of modal data in terms of the unknown parameters may be written as follows:

$$\mathbf{Z}_m - \mathbf{Z}_i = \mathbf{S}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i) + \boldsymbol{\varepsilon}_z \tag{2.3}$$

So that linearization is carried out at $\boldsymbol{\theta} = \boldsymbol{\theta}_i$. The measured and predicted outputs are denoted as $\mathbf{Z}_m$ and $\mathbf{Z}_i$. Where $\mathbf{Z}_i = \mathbf{Z}(\boldsymbol{\theta}_i)$. These outputs can be eigenfrequencies, mode shapes or complex frequency response functions. The $\mathbf{S}_i$ is defined as the first-order sensitivity matrix, given by

$$\mathbf{S}_i = \left[\frac{\partial \mathbf{Z}_j}{\partial \theta_k} \right]_{\boldsymbol{\theta} = \boldsymbol{\theta}_i} \tag{2.4}$$

where $j = 1, 2, \ldots, q$ denotes the output data and $k = 1, 2, \ldots, p$ is the parameter index. The error vector $\boldsymbol{\varepsilon}_z$ is assumed to be small for parameters $\boldsymbol{\theta}$ in the area of $\boldsymbol{\theta}_i$.

For undamped structural eigenvalue problem, the first derivatives of eigenvalues are expressed by Wittrick [16] since considering a system can be modeled by real symmetric mass and stiffness matrices $\mathbf{M}$ and $\mathbf{K}$, the eigenvalue (natural frequency squared) λ_i and eigenvector (mode shapes) ϕ_i are given by the solution of

$$\left[\mathbf{K} - \lambda_i \mathbf{M}\right]\phi_i = \mathbf{0} \tag{2.5}$$

Differentiating Eq. (2.5) with design parameters θ_k gives

$$\left[\mathbf{K} - \lambda_i \mathbf{M}\right]\frac{\partial \phi_i}{\partial \theta_k} + \left[\frac{\partial \mathbf{K}}{\partial \theta_k} - \lambda_i \frac{\partial \mathbf{M}}{\partial \theta_k}\right]\phi_i - \frac{\partial \lambda_i}{\partial \theta_k}\mathbf{M}\phi_i = \mathbf{0} \tag{2.6}$$

By premultiplying Eq. (2.6) by ϕ_i^T and using the mass orthogonality, the eigensystem produces

$$\frac{\partial \lambda_i}{\partial \theta_k} = \phi_i^T \left[\frac{\partial \mathbf{K}}{\partial \theta_k} - \lambda_i \frac{\partial \mathbf{M}}{\partial \theta_k}\right]\phi_i \tag{2.7}$$

This expression is only requiring the $i-$th eigenvalue, $i-$th eigenvector and first-order sensitivity matrix to be calculated. The θ_k could be either element specifically defined in mass and stiffness matrices $\mathbf{M}$ and $\mathbf{K}$.

Taking parameter updating techniques by defining a residual containing the analysis and test behavior differences which could be formed by Eq. (2.3). Therefore, the single objective function is to minimize

$$J(\boldsymbol{\theta}) = \boldsymbol{\varepsilon}_z^T \boldsymbol{\omega}_\varepsilon \boldsymbol{\varepsilon}_z \tag{2.8}$$

The symmetric weighting matrix $\boldsymbol{\omega}_\varepsilon$ is considered to account for the importance of each individual term in the residual vectors. $\boldsymbol{\omega}_\varepsilon$ is difficult to estimate and Mottershead, et al. [17] suggest a reasonable choice is $\boldsymbol{\omega}_\varepsilon = \left[diag(\mathbf{Z}_m)\right]^{-2}$ to scaling and balance the effect of amplitude.

Normally, in parameter updating, the number of measurements should always be made equal or larger than the number of parameters which yields determined or overdetermined equation systems. For these cases, simply minimize $J(\boldsymbol{\theta})$ with

respect to $\Delta\boldsymbol{\theta}_i = (\boldsymbol{\theta} - \boldsymbol{\theta}_i)$ give a theoretical parameter estimate as

$$\Delta\boldsymbol{\theta}_i = \left[\mathbf{S}_i^T\boldsymbol{\omega}_\varepsilon\mathbf{S}_i\right]^{-1}\mathbf{S}_i^T\boldsymbol{\omega}_\varepsilon(\mathbf{Z}_m - \mathbf{Z}_i) \qquad (2.9)$$

A fundamental requirement to obtain a solution is $rank[\mathbf{S}_i^T\boldsymbol{\omega}_\varepsilon\mathbf{S}_i] = p$. As p is the number of parameters.

The case when fewer measurements than parameters are available in Eq. (2.3) $(q < p)$ leads to a undetermined equation systems whose solution is not unique. The treatment of ill-conditioned, noisy systems of equations is a problem to FE model updating. The classical weighted least squares method in Eq. (2.8) can be extended in case it is difficult to obtain a convergent solution because of an ill - conditioned sensitivity matrix. The objective function $J(\boldsymbol{\theta})$ is expanded for the requirement of parameter changes $\Delta\boldsymbol{\theta}_i$ should be minimized as

$$J(\boldsymbol{\theta}) = \boldsymbol{\varepsilon}_z^T\boldsymbol{\omega}_\varepsilon\boldsymbol{\varepsilon}_z + \upsilon^2\Delta\boldsymbol{\theta}_i^T\boldsymbol{\omega}_\theta\Delta\boldsymbol{\theta}_i \qquad (2.10)$$

The parameter weighting matrix $\boldsymbol{\omega}_\theta$ should be chosen to reflect the uncertainty in the initial parameter estimates and the regularisation parameter υ provides a balance between the measurement residual $\boldsymbol{\varepsilon}_z^T\boldsymbol{\omega}_\varepsilon\boldsymbol{\varepsilon}_z$ and the side constraint $\Delta\boldsymbol{\theta}_i^T\boldsymbol{\omega}_\theta\Delta\boldsymbol{\theta}_i$, which can be found in the work of Link [18] and Hansen [19]. Moreover, when ill-conditioning cannot be solved by well-thought parameterization of the model, regularization techniques can be used. A detailed overview of regularization techniques in FE model updating can be found in Titurus and Friswell [20].

2.4 Computational intelligence inverse methods

FE model updating problems are sometimes computational expensive for realistic systems. The computational intelligence inverse methods are about using modern optimization algorithms to do the FE model updating, which has advantages to manage the computational load without losing updating efficiency and accuracy. Liu et al. [21] used a fuzzy theory, while Jung and Kim [22] employed a hybrid genetic algorithm for FE model updating. Tan et al. [23] used support vector machines for FE model updating, while Zapico, et al. [24] applied artificial neural networks to update a small steel frame. Further successful application of computational intelligence methods to FE model updating include Tu and Lu [25]

who used a genetic algorithm, Feng et al. [26] applied a hybrid of genetic algorithm and simulated annealing, He et al. [27] used a hybrid of a genetic algorithm and artificial neural networks and Marwala [28] applied the particle swarm optimization technique for FE model updating.

2.5 Second-Order Sensitivities

2.5.1 Background description

As described in section 2.3 for the iterative method in FE model updating, the traditional sensitivity approach is basically using a linear approximation to represent the generally non-linear relationship between measured outputs and parameters of the model in need of correction by assuming a small change for design parameters. In reality, this assumption is sound more or less idealistic, if the variation for the design parameters is not small, only using the linear approximation approach to model the response outputs looks inadequate.

Under such conditions, it is highly demanding to put the non-linear part of Taylor series expansion of model data into consideration, specifically, obtaining Taylor series expansion truncated after the second-order nonlinear term of model data to model the generally non-linear relationship of measured response and design parameters.

$$\mathbf{Z}_{mj} - \mathbf{Z}_{ij} = \mathbf{S}_{ij}(\mathbf{\theta} - \mathbf{\theta}_i) + \frac{1}{2}(\mathbf{\theta} - \mathbf{\theta}_i)^T \mathbf{H}_{ij}(\mathbf{\theta} - \mathbf{\theta}_i) + \mathbf{\epsilon}_{zj} \tag{2.11}$$

where $j = 1, \ldots, q$ denotes the number of output datasets, $\mathbf{S}_{ij}$ is defined as the first-order sensitivity matrix expressed as $\left[\dfrac{\partial \mathbf{Z}_j}{\partial \theta_k} \right]$ at the i th iteration. The $\mathbf{H}_{ij}$ is the second-order sensitivity matrix defined as

$$\mathbf{H}_{ij} = \begin{bmatrix} \dfrac{\partial^2 \mathbf{Z}_j}{\partial \theta_k \partial \theta_k} & \cdots & \dfrac{\partial^2 \mathbf{Z}_j}{\partial \theta_k \partial \theta_u} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 \mathbf{Z}_j}{\partial \theta_u \partial \theta_k} & \cdots & \dfrac{\partial^2 \mathbf{Z}_j}{\partial \theta_u \partial \theta_u} \end{bmatrix}_{\mathbf{\theta} = \mathbf{\theta}_i} \tag{2.12}$$

where $\mathbf{\theta}_i = (\theta_k, \ldots, \theta_u)_i^T$.

The second-order derivatives could be calculated similarly to the calculation process for the first-order sensitivities. By differentiating Eq. (2.6) with possible another parameter θ_u (could be either element specifically defined in mass and stiffness matrices $\mathbf{M}$ and $\mathbf{K}$) a second time, which could be expressed as

$$\left[\frac{\partial}{\partial\theta_k}\big[\mathbf{K}-\lambda_i\mathbf{M}\big]\right]\frac{\partial\phi_i}{\partial\theta_u}+\left[\frac{\partial}{\partial\theta_u}\big[\mathbf{K}-\lambda_i\mathbf{M}\big]\right]\frac{\partial\phi_i}{\partial\theta_k}$$

$$+\left[\frac{\partial^2}{\partial\theta_k\partial\theta_u}\big[\mathbf{K}-\lambda_i\mathbf{M}\big]\right]\phi_i+\left[\big[\mathbf{K}-\lambda_i\mathbf{M}\big]\frac{\partial^2\phi_i}{\partial\theta_k\partial\theta_u}\right]=\mathbf{0} \tag{2.13}$$

Premultiplying Eq. (2.13) with ϕ_i^T, and using the mass orthogonality of the i th eigenvector (mode shapes), the second-order sensitivities of i th eigenvalue can be expressed as

$$\frac{\partial^2\lambda_i}{\partial\theta_k\partial\theta_u}=\phi_i^T\left[\frac{\partial}{\partial\theta_k}\big[\mathbf{K}-\lambda_i\mathbf{M}\big]\right]\frac{\partial\phi_i}{\partial\theta_u}+\phi_i^T\left[\frac{\partial}{\partial\theta_u}\big[\mathbf{K}-\lambda_i\mathbf{M}\big]\right]\frac{\partial\phi_i}{\partial\theta_k}$$

$$+\phi_i^T\left[\frac{\partial^2\mathbf{K}}{\partial\theta_u\partial\theta_k}-\lambda_i\frac{\partial^2\mathbf{M}}{\partial\theta_u\partial\theta_k}-\frac{\partial\lambda_i}{\partial\theta_u}\frac{\partial\mathbf{M}}{\partial\theta_k}-\frac{\partial\lambda_i}{\partial\theta_k}\frac{\partial\mathbf{M}}{\partial\theta_u}\right]\phi_i \tag{2.14}$$

In addition, if the updating parameters are assigned for only stiffness parameters in the matrix $\mathbf{K}$, Eq. (2.14) can be simplified as

$$\frac{\partial^2\lambda_i}{\partial\theta_k\partial\theta_u}=\phi_i^T\frac{\partial\mathbf{K}}{\partial\theta_k}\left(\frac{\partial\phi_i}{\partial\theta_u}\right)+\left(\frac{\partial\phi_i}{\partial\theta_u}\right)^T\frac{\partial\mathbf{K}}{\partial\theta_k}\phi_i \tag{2.15}$$

The result in Eq. (2.15) only requires the first-order derivatives for eigenvectors to calculate. It is worth noting that there are several methods that can be used to calculate the first-order derivatives of the eigenvectors (mode shapes). In this section, a method proposed by Fox and Kapoor [29] is employed to calculate first-order derivatives of the eigenvectors.

This approach is based on expanding the gradient and described as a linear combination of all eigenvectors given as

$$\frac{\partial\phi_i}{\partial\theta_u}=\sum_{h=1}^{n}\alpha_{iuh}\ ,n\le N \tag{2.16}$$

Substitution of Eq. (2.16) the derivative of the eigenvalue problems, the factor α_{iuh} is calculated in the form of

$$\alpha_{iuh}=\begin{cases}\alpha_{iuh}=\dfrac{\phi_h^T\left[-\lambda_i\dfrac{\partial\mathbf{M}}{\partial\theta_k}+\dfrac{\partial\mathbf{K}}{\partial\theta_k}\right]\phi_i}{\lambda_i-\lambda_h}\ \ i\ne h\\[2em]\alpha_{iui}=-\dfrac{1}{2}\phi_i^T\dfrac{\partial\mathbf{M}}{\partial\theta_k}\phi_i\end{cases} \tag{2.17}$$

This expansion is exact if $n=N$ modes are used. For $n\le N$, this expansion gives an approximation depending on the number of modal terms, a suitable truncation

based n is usually determined by trial-and-error. Another suggested method is proposed by Nelson [30], which has advantages that only the mode shape of interest is required.

2.5.2 A multi-objective optimization framework

Since we have obtained the non-linear approximation forms by truncated Taylor series expansion of model responses, the error vector $\boldsymbol{\varepsilon}_z$ is defined as $\boldsymbol{\varepsilon}_{zj}$, $j = 1, 2, \ldots, q$ and q is the number of output datasets. The $\boldsymbol{\varepsilon}_{zj}$ can be expressed as

$$\boldsymbol{\varepsilon}_{zj} = Z_{mj} - Z_{ij} - \left(S_{ij} \left(\boldsymbol{\theta} - \boldsymbol{\theta}_i \right) + \frac{1}{2} \left(\boldsymbol{\theta} - \boldsymbol{\theta}_i \right)^T \mathbf{H}_{ij} \left(\boldsymbol{\theta} - \boldsymbol{\theta}_i \right) \right) \tag{2.18}$$

The multi-objective function is then proposed, which can be expressed as

$$J\left(\boldsymbol{\theta} \right)_j = \boldsymbol{\varepsilon}_{zj}{}^T \boldsymbol{\omega}_{\varepsilon j} \boldsymbol{\varepsilon}_{zj}, j = 1, 2, \ldots, q \tag{2.19}$$

where the weighting value $\boldsymbol{\omega}_{\varepsilon j}$ is chosen as $\boldsymbol{\omega}_{\varepsilon j} = \left(Z_{mj} \right)^{-2}, j = 1, 2, \ldots, q$.

The multi-objective optimization problem is finally can be defined as

$$arg\ min_{\boldsymbol{\theta}}\ J\left(\boldsymbol{\theta} \right)_j = \boldsymbol{\varepsilon}_{zj}{}^T \boldsymbol{\omega}_{\varepsilon j} \boldsymbol{\varepsilon}_{zj}, j = 1, 2, \ldots, q \tag{2.20}$$

The process of optimizing systematically and simultaneously a collection of objective functions is called multi-objective optimization (MOO) or vector optimization. The general multi-objective optimization problem can be expressed as follows

$$\begin{aligned} &arg\ min_{\mathbf{X}}\ F(\mathbf{X}) = \left[F_1(\mathbf{X}), F_2(\mathbf{X}), \ldots, F_k(\mathbf{X}) \right]^T \\ &subject\ to\ g_j(\mathbf{X}) \leq 0, j = 1, 2, \ldots, m \\ &h_i(\mathbf{X}) = 0, i = 1, 2, \ldots, e \end{aligned} \tag{2.21}$$

Where k is the number of objective functions, m is the number of inequality constraints, and e is the number of equality constraints. $\mathbf{X}$ is the vector of design variables. In contrast to single-objective optimization, a solution to a multi-objective problem is more of a concept than a definition. Typically, there is no single global solution, and it is often necessary to determine a set of points that all fit a predetermined definition for an optimum.

Currently, some methods can be used to specify preferences, which may be articulated in terms of goals or the relative importance of different objectives. Most of these methods incorporate parameters, which are coefficients, exponents,

constraint limits, etc. that can be set to reflect decision-maker preferences. These optimization methods include:

(1) Weighted global criterion methods in which all objective functions are combined to form a single function referred to the work by Wierzbicki [31], Miettinen [32].

(2) Weighted sum methods in which to give weights to sum all the objective functions referred to the work by Das and Dennis [33], Messal et al. [34].

(3) The Lexicographic methods in which the objective functions are arranged in order of importance. Referred to the work by Rentmeesters et al. [35] and Weighted min-max methods in which by introducing some additional unknown parameters γ combined with weights ω to process the optimization problems.

Gembicki [36] proposed a weighted min-max method named goal attainment method for multi-objective optimization. This is concerned with the minimization of a set of objectives simultaneously, which will be used in this paper. Its definition combined with multiple-objective functions derived above can be specified by

$$
\begin{cases}
\varepsilon_{zj} = Z_{mj} - Z_{ij} - \left(S_{ij}\left(\theta - \theta_i\right) + \frac{1}{2}\left(\theta - \theta_i\right)^T \mathbf{H}_{ij}\left(\theta - \theta_i\right) \right), \ j = 1, 2, \ldots, q \\
J\left(\theta\right)_j = \varepsilon_{zj}^T \omega_{\varepsilon j} \varepsilon_{zj}, \ j = 1, 2, \ldots, q \\
arg\ min_{\theta, \gamma} \left[J\left(\theta\right)_1, J\left(\theta\right)_2, \ldots, J\left(\theta\right)_q \right]^T \leq \gamma \\
Constraint\ Conditions
\end{cases}
\tag{2.22}
$$

The weights are defined as $\omega_{\varepsilon j} = \left(Z_{mj}\right)^{-2}, \ j = 1, 2, \ldots, q$. In order to improve the computation efficiency, the sequential quadratic programming (SQR) method is incorporated, modifications are made to the linear search and Hessian matrix described by [37, 38].

2.5.3 An illustrative example of three degree of freedom mass-spring system

The numerical system, which is used to demonstrates the proposed sensitivity method is the three degree of freedom mass-spring system shown in Figure 2.1. The nominal values of parameters of the 'synthetic experimental' system are given in Table 2.1.
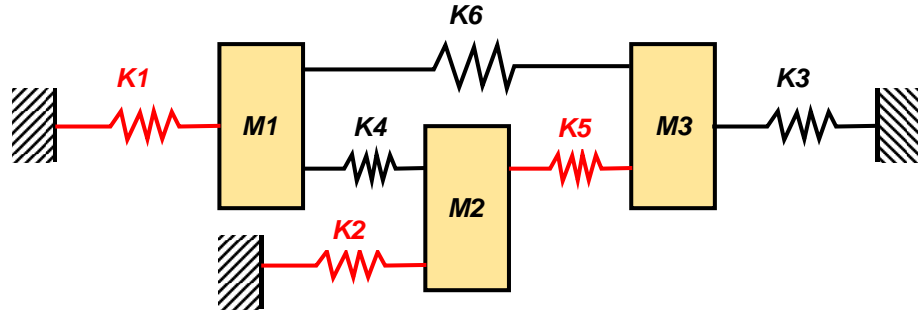
Figure 2.1 *Three degree of freedom mass-spring system.*

In this example, the three uncertain parameters, k_1, k_2, k_5 are assumed responsible for the observed variability in the three natural frequencies of the system.

Table 2.1 *Parameter values for the 'experimental' system*

Mass vector (kg)		Stiffness vector (N/m)			
m_1	1	k_1	2	k_2	2
m_2	1	k_3	1	k_4	1
m_3	1	k_5	2	k_6	3

The first case is by assuming the changes for the design uncertain parameters k_1, k_2, k_5 are relatively small compared to the parameter values in the 'experimental' system . Figure 2.2 and Figure 2.3 show the convergence of uncertain parameters using the traditional and proposed sensitivity method respectively when the initial values for the parameters are settled as $1N/m$.
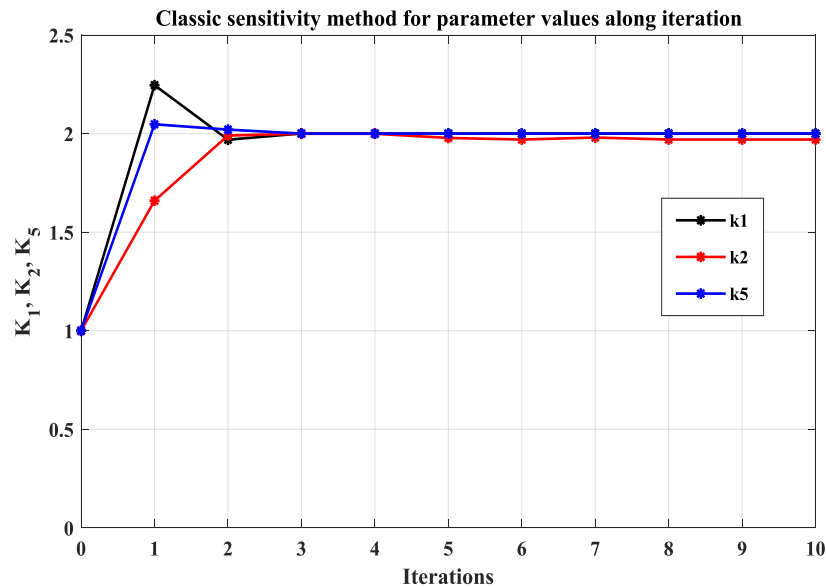
Figure 2.2 *Convergence of uncertain parameters using the traditional sensitivity method (Eq. 2.9).*

In Figure 2.2, parameters convergence is to be obtained after 6 iterations except k_2 with the final updated value $1.971N/m$, which is compared to the 'synthetic experimental' data point $2N/m$ when using traditional sensitivity method. But in Figure 2.3, all the parameters values are updated to the 'synthetic experimental' value $2N/m$ after only 2 iterations by using the proposed sensitivity method.
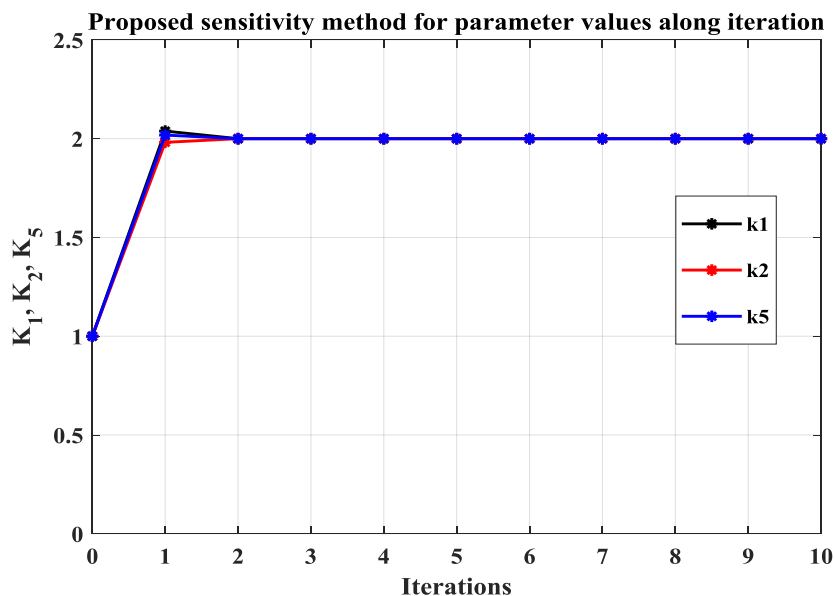
Figure 2.3 *Convergence of uncertain parameters using the proposed sensitivity method (Eq. 2.22).*
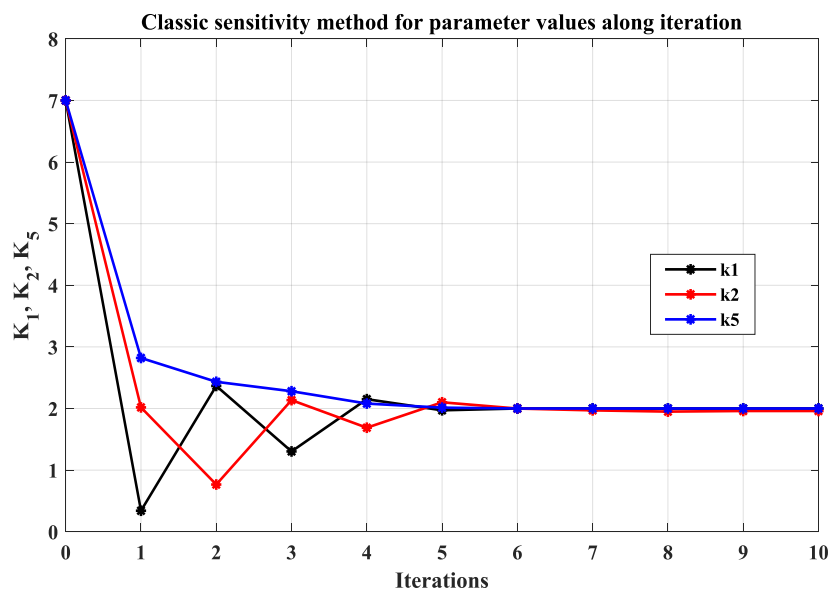
Figure 2.4 *Convergence of uncertain parameters using the traditional sensitivity method (Eq. 2.9).*
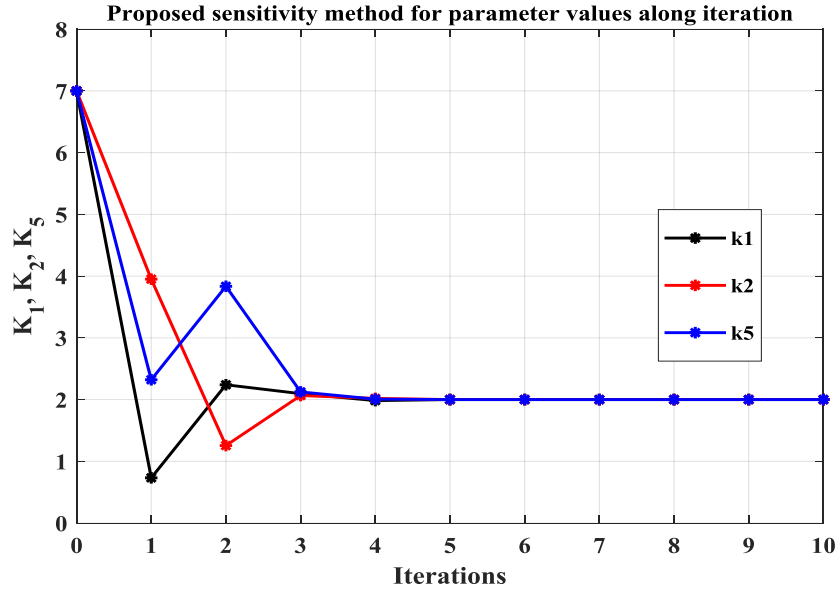
Figure 2.5 *Convergence of uncertain parameters using the proposed sensitivity method (Eq. 2.22).*

In the cases shown in Figure 2.4 and Figure 2.5, we set the initial values of the parameters to be $7N/m$, which is larger than $1N/m$. In Figure 2.4, the convergence of the parameters can be seen after 8 iterations except k_2 with the final updated value $1.987N/m$, which is compared to $2N/m$ by using the classic sensitivity method. On the other hand, using the proposed sensitivity method shown in Figure 2.5, all the updating parameters have converged to the $2N/m$ with only 4 iterations. It can be concluded that when the changes for the parameters are small, the proposed sensitivity method seems to be more efficient and accurate than the traditional method.

The second case is assuming that changes for the design uncertain parameters k_1, k_2, k_5 are relative large. The initial values for updating parameters are designed as $22N/m$. In Figure 2.6, it can be seen that all the design parameters have converged to $2N/m$ except k_2 with the final updated value $1.983N/m$ after 6 iterations by using the traditional sensitivity method. More importantly, in Figure 2.6, during the procedure of updating k_2 . Its value after the first iteration is negative -2.5848 N/m. When facing this situation, by giving constraints to the k_2 during the updating process is suggested when using the traditional sensitivity method.
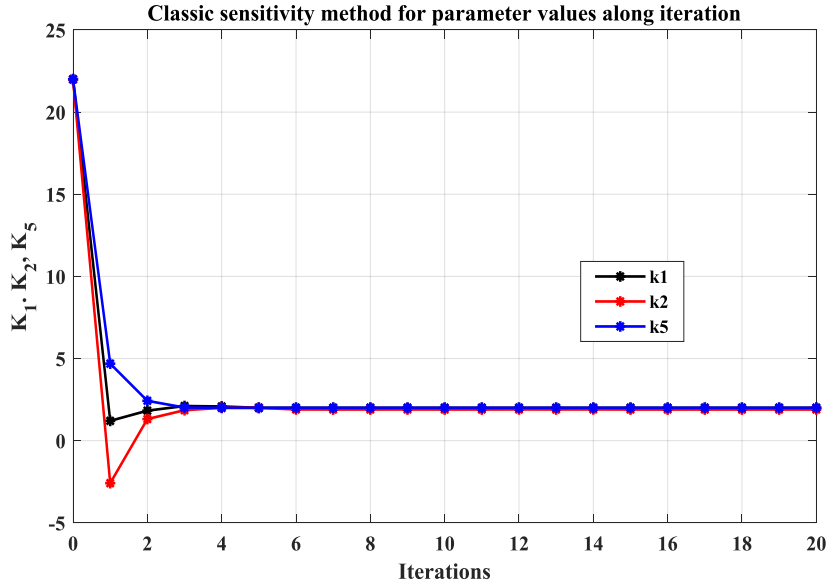
Figure 2.6 *Convergence of uncertain parameters using the traditional sensitivity method (Eq. 2.9).*
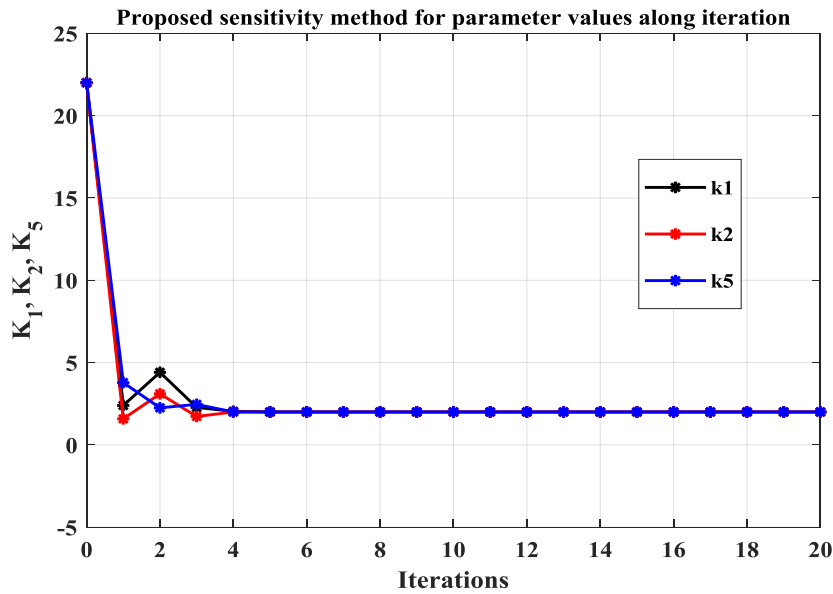
Figure 2.7 *Convergence of uncertain parameters using the proposed sensitivity method (Eq. 2.22).*

On the other hand, Figure 2.7 shows the convergence of parameters k_1, k_2, k_5 by using the proposed sensitivity method. It is clear that all the parameters are updated to $2N/m$ with just 4 iterations. In addition, there is no need to set constraints to the parameters during the updating process. For this occasion, The proposed sensitivity method shows some merits for the mode updating problems compared to using the traditional method.

2.5.4 An illustrative example of six storey bare frame

Figure 2.8 gives the numerical example under structural dynamics background. It is a six storey bare frame, which is used to demonstrate the application of the proposed sensitivity method. The nominal values of parameters of the 'synthetic experimental' system are given in Table 2.2. The corresponding measured modal properties are estimated, which is given in Table 2.3.
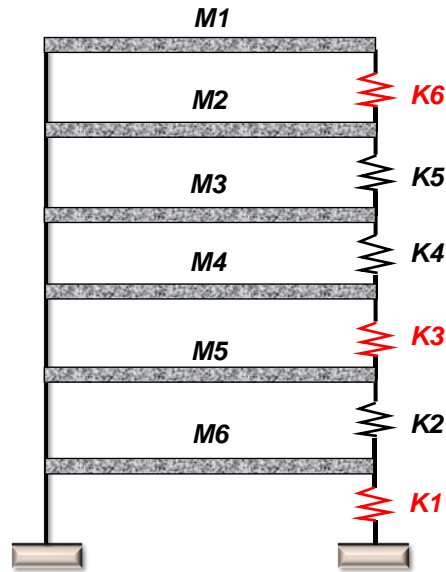
Figure 2.8 *Six storey bare frame.*

Table 2.2 *Parameter values for the 'experimental' system*

Mass vector (kg)		Stiffness vector (N/m)	
m_1	8×10^5	k_1	1.5×10^9
m_2	6×10^5	k_2	1.5×10^9
m_3	6×10^5	k_3	1.8×10^9
m_4	6×10^5	k_4	1.6×10^9
m_5	6×10^5	k_5	2.0×10^9
m_6	6×10^5	k_6	1.6×10^9

In this case, the three uncertain parameters, k_1, k_3, k_6 are assumed responsible for the observed variability in the six storey bare frame. The initial values for the design parameters are $k_1 = 1.8\times10^9 N/m$, $k_3 = 2.5\times10^9 N/m$ and $k_6 = 2\times10^9 N/m$.

Table 2.3 *The actual modal properties of the 'experimental' system*

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6
ω^2 $(rad/s)^2$	154	1327	3267	5352	8591	10894
ϕ_1	-0.00019	0.000537	-0.0007	0.000574	0.000306	-0.00011
ϕ_2	-0.00036	0.000693	-0.00018	-0.00049	-0.00079	0.0004
ϕ_3	-0.00048	0.000518	0.00045	-0.0005	0.000559	-0.00063
ϕ_4	-0.00059	0.000621	0.000609	0.000492	0.000276	0.000789
ϕ_5	-0.00065	-0.00033	0.000139	0.000498	-0.00026	-0.00065
ϕ_6	-0.00069	-0.00065	-0.00062	-0.00049	0.000298	0.000211
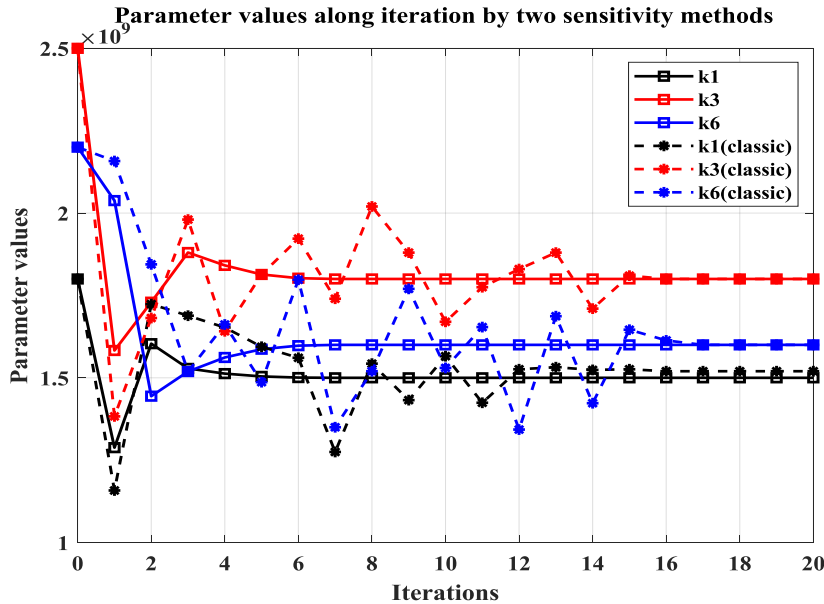
Figure 2.9 *Convergence of uncertain parameters using classic (Eq. 2.9) and proposed (Eq. 2.22) methods.*

It can be seen that in Figure 2.9, which shows the convergence of design parameters k_1, k_3, k_6 using classic and proposed sensitivity method, respectively. It is clear that all the updating parameters have converged to the initial preset values except k_1 with $1.572 \times 10^9 N/m$ using classic sensitivity method. On the other hand, using the proposed sensitivity method, all the updating parameters k_1, k_3, k_6 have converged to the initial preset values. In addition, the number of iterations obtained with the proposed sensitivity method is much smaller than the one obtained with classic sensitivity method, which means that a relatively high efficiency is also achieved.

2.6   Chapter summary

This chapter gives an overall review of the deterministic model updating approaches and the advantages and disadvantages of these methods are also discussed. The deterministic model updating methods are categories into three groups: (a) direct method, (b) iterative methods and (c) computational intelligence methods.

A novel iterative method is then proposed, which is formulated based on the classic sensitivity method. The efficiency and advantages of this method are verified by using two numerical examples, named the three degree of freedom mass-spring system and the six storey bare frame. For the numerical examples provided, the proposed sensitivity method seems to have some advantages especially when large changes for the parameters of the model in need of correction. In addition, it also enriches the theory of deterministic model updating.

The next chapter namely stochastic FE model updating will be introduced in detail.

3 Stochastic finite element model updating

3.1 Introduction

The previous chapter presented deterministic FE model updating, it was assumed that the measured data are accurate. However, variability inevitably exists in the measured data, model parameters and the uncertainty due to model discrepancy and approximations which has to be considered in the model updating procedure. In practical model updating the measured data are often imprecise, incomplete and variable. Therefore, it is very important to put statistical techniques into consideration to improve updating performance.

Variability in experimental results can be categorized into two groups: reducible and irreducible [39]. Measurement noise and the use of sensors that affect the measurement signal processing that might introduce bias are some examples of reducible uncertainty which can be minimized by gathering more information (e.g. repeating the measurement). Under such conditions, randomness arises only from the measurement noise and the updating parameters take unique values, found by iterative correction to the estimated means, whilst the variances are minimized.

On the other hand, the distributions of updating parameters become physically meaningful in the presence of irreducible uncertain measured data. The stochastic model updating problem includes not only the variability in measurement signals due to noise but also the variability that exists between nominally identical test structures, built in the same way from the same materials but with manufacturing and material variability [40]. Similar variability is known from environmental erosion, damage [41] or disassembly and reassembly of the same structure [42].

In this chapter, a detailed description of probabilistic stochastic model updating methods including minimum variance method, probabilistic perturbation method and Bayesian model updating method is given. A brief description of non-probabilistic stochastic model updating methods including interval model updating methods and fuzzy model updating method is also described. Finally, A comparative study on the performance of probabilistic perturbation method and Bayesian model updating method is carried out on a three-degree-of-freedom

mass-spring system and some useful conclusions are presented in the chapter summary.

3.2 The minimum variance method

The minimum variance method is an iterative process that takes into account the parameter variability and the uncertainties related to constructing the FE model by Collins et al. [43]. The method incorporated statistical techniques to treat the test measurement errors as well as uncertainty in the estimation of updating parameters. In other words, it is assumed that both measured data and updating parameters have errors which may be described by their variances. This technique minimizes the variance of the uncertain parameters, at each iteration.

Suppose $\mathbf{\theta}_i$ is the vector of uncertain parameters at the ith iteration of updating procedure. The variance matrix of the parameters at the ith iteration is $E(\mathbf{\theta}_i\mathbf{\theta}_i^T) = \mathbf{V}_i$. By subtracting the predicted output $\mathbf{Z}_i$ at the ith iteration from the measurement datasets $\mathbf{Z}_m$ yields:

$$\delta\mathbf{Z} = \mathbf{Z}_m - \mathbf{Z}_i = \mathbf{S}(\mathbf{\theta} - \mathbf{\theta}_i) \tag{3.1}$$

Then the approximated uncertain parameter vector at the $(i+1)$th iteration $\mathbf{\theta}_{i+1}$ could be written as

$$\mathbf{\theta}_{i+1} - \mathbf{\theta}_i = \mathbf{T}(\mathbf{Z}_m - \mathbf{Z}_i) \tag{3.2}$$

where $\mathbf{T}$ is the unknown transformation matrix. The new variance of estimated parameters $\mathbf{\theta}_{i+1}$ for the $(i+1)$th iteration is given by

$$\mathbf{V}_{i+1} = E(\mathbf{\theta}_{i+1}\mathbf{\theta}_{i+1}^T) = \mathbf{V}_i + (\mathbf{D}_i + \mathbf{V}_i\mathbf{S}_i^T)^T\mathbf{T}^T + \mathbf{T}(\mathbf{D}_i^T - \mathbf{V}_i\mathbf{S}_i) + \mathbf{T}\mathbf{V}_{\gamma i}\mathbf{T}^T \tag{3.3}$$

where $D_i = E(\mathbf{\theta}_i\mathbf{Z}_m^T)$ is the correlation between the parameter approximation and the measurement noise. The output error variance is

$$\mathbf{V}_{\gamma i} = \mathbf{S}_i\mathbf{V}_i\mathbf{S}_i^T - \mathbf{S}_i\mathbf{D}_i - \mathbf{D}_i^T\mathbf{S}_i^T + \mathbf{V}_e \tag{3.4}$$

where $\mathbf{V}_e = E(\mathbf{Z}_m\mathbf{Z}_m^T)$ and the transformation matrix is obtained by minimizing the variance at the $(i+1)$th iteration as follows

$$\mathbf{T} = (\mathbf{V}_i\mathbf{S}_i^T - \mathbf{D}_i)\mathbf{V}_{\gamma i}^{-1} \tag{3.5}$$

At last, the updated parameters $\mathbf{\theta}_{i+1}$, $\mathbf{V}_{i+1}$ and $\mathbf{D}_{i+1}$ are obtained as

$$\mathbf{\theta}_{i+1} = \mathbf{\theta}_i + (\mathbf{V}_i\mathbf{S}_i^T - \mathbf{D}_i)\mathbf{V}_{\gamma i}^{-1}(\mathbf{Z}_m - \mathbf{Z}_i) \tag{3.6}$$

$$\mathbf{V}_{i+1} = \mathbf{V}_i - (\mathbf{V}_i\mathbf{S}_i^T - \mathbf{D}_i)\mathbf{V}_{\gamma i}^{-1}(\mathbf{V}_i\mathbf{S}_i^T - \mathbf{D}_i)^T \tag{3.7}$$

$$\mathbf{D}_{i+1} = \mathbf{D}_i + (\mathbf{V}_i\mathbf{S}_i^T - \mathbf{D}_i)\mathbf{V}_{\gamma i}^{-1}(\mathbf{S}_i\mathbf{D}_i - \mathbf{V}_e) \tag{3.8}$$

3.3 The probabilistic perturbation method

The probabilistic perturbation methods applied a Tylor series expansion to extend the terms in model updating equations around a predefined data point and then to estimate the mean and variance of updated parameters [44]. Khodaparast et al. [45] also showed the computation of the Hessian matrix is not necessary for this kind of perturbation-based updating method since the correlation between the measurement and updating parameters is assumed to be not considered in model updating. Govers and Link [42] proposed analytical output covariance matrix for the stochastic model updating which was shown to be identical by Khodaparast et al. [45] and Silva et al. [46].

Let the estimated parameters $\boldsymbol{\theta}_{j+1}$ could be updated using the prior estimate $\boldsymbol{\theta}_j$ expressed as

$$\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}_j + \mathbf{T}_j(\mathbf{Z}_m - \mathbf{Z}_j) \tag{3.9}$$

where $\mathbf{Z}_j \in \mathfrak{R}^{n\times 1}$ are the estimated parameters output vectors, $\mathbf{Z}_m \in \mathfrak{R}^{n\times 1}$ is the measurement datasets vectors. $\boldsymbol{\theta} \in \mathfrak{R}^{m\times 1}$ is the system parameters vector and $\mathbf{T}_j$ is the transformation matrix. By accounting for the variabilities in measurement arising from multiple sources, the modal parameters are expressed as

$$\mathbf{Z}_m = \bar{\mathbf{Z}}_m + \Delta\mathbf{Z}_m \tag{3.10}$$

$$\mathbf{Z}_j = \bar{\mathbf{Z}}_j + \Delta\mathbf{Z}_j \tag{3.11}$$

where $\bar{\mathbf{Z}}_m$ denotes the mean value of $\mathbf{Z}_m$ and $\Delta\mathbf{Z}_m,\ \Delta\mathbf{Z}_j \in \mathfrak{R}^{n\times 1}$ are random vectors. The variabilities in physical parameters at the jth iterations are given by

$$\boldsymbol{\theta}_j = \bar{\boldsymbol{\theta}}_j + \Delta\boldsymbol{\theta}_j \tag{3.12}$$

Therefore, the stochastic model updating problems are defined as

$$\bar{\boldsymbol{\theta}}_{j+1} + \Delta\boldsymbol{\theta}_{j+1} = \bar{\boldsymbol{\theta}}_j + \Delta\boldsymbol{\theta}_j + (\bar{\mathbf{T}}_j + \Delta\mathbf{T}_j)(\bar{\mathbf{Z}}_m + \Delta\mathbf{Z}_m - \bar{\mathbf{Z}}_j - \Delta\mathbf{Z}_j) \tag{3.13}$$

where the transformation matrix is

$$\mathbf{T}_j = \bar{\mathbf{T}}_j + \Delta\mathbf{T}_j \tag{3.14}$$

$$\Delta \mathbf{T}_j = \sum_{k=1}^{n} \frac{\partial \overline{\mathbf{T}}_j}{\partial \mathbf{Z}_{mk}} \Delta \mathbf{Z}_{mk} \tag{3.15}$$

where $\overline{\mathbf{T}}_j$ denotes the transformation matrix at the parameter means, $\overline{\mathbf{T}}_j = \mathbf{T}(\overline{\boldsymbol{\theta}}_j)$ and $\Delta \mathbf{Z}_{mk}$ is the kth element of $\Delta \mathbf{Z}_m$. Separating Eq. (3.13) into zero-order and first-order term, which leads to

$$o(\Delta^0) = \overline{\boldsymbol{\theta}}_j + \overline{\mathbf{T}}_j (\overline{\mathbf{Z}}_m - \overline{\mathbf{Z}}_j)$$

$$o(\Delta^1) = \Delta \boldsymbol{\theta}_{j+1} + \overline{\mathbf{T}}_j (\Delta \mathbf{Z}_m - \Delta \mathbf{Z}_j) + ((\sum_{k=1}^{n} \frac{\partial \overline{\mathbf{T}}_j}{\partial \mathbf{Z}_{mk}} \Delta \mathbf{Z}_{mk})(\overline{\mathbf{Z}}_m - \overline{\mathbf{Z}}_j)) \tag{3.16}$$

Eq. (3.16) gives the estimates of parameter means and covariance matrix. Khodaparast et al. [45] concluded that when the correlation between the updating parameters and the measurements are omitted, the requirement to calculate the second order sensitivities is no longer necessary since there is no significant deterioration on the estimated parameters distribution. The parameter covariance matrix can be estimated at the $(j+1)$th iteration expressed as

$$Cov(\Delta \boldsymbol{\theta}_{j+1}, \Delta \boldsymbol{\theta}_{j+1}) = Cov(\Delta \boldsymbol{\theta}_j, \Delta \boldsymbol{\theta}_j) - Cov(\Delta \boldsymbol{\theta}_j, \Delta \mathbf{Z}_j) \overline{\mathbf{T}}_j^T$$

$$+ \overline{\mathbf{T}}_j Cov(\Delta \mathbf{Z}_m, \Delta \mathbf{Z}_m) \overline{\mathbf{T}}_j^T - \overline{\mathbf{T}}_\mathbf{j} Cov(\Delta \mathbf{Z}_j, \Delta \boldsymbol{\theta}_j) + \overline{\mathbf{T}}_j Cov(\Delta \mathbf{Z}_j, \Delta \mathbf{Z}_j) \overline{\mathbf{T}}_j^T \tag{3.17}$$

where $\overline{\mathbf{T}}_j$ is the transformation matrix, that is, the generalized pseudo-inverse of the sensitivity matrix.

$$\overline{\mathbf{T}}_j = (\overline{\mathbf{S}}_j^T \mathbf{W}_1 \overline{\mathbf{S}}_j + \mathbf{W}_2)^{-1} \overline{\mathbf{S}}_j^T \mathbf{W}_1 \tag{3.18}$$

where $\mathbf{W}_1$ and $\mathbf{W}_2$ are the weighting matrices to account for regularization of ill-posed sensitivity equation [47].

3.4 The Bayesian model updating method

Bayesian model updating methods are techniques based on the Bayesian inference [48], which is a statistical school of thought that quantifies the degree of belief of parameters' values. This theory has had great advancement after the significant work by Jeffreys [49] and Cox [50]. Since then, different Bayesian methods have been developed and widely applied to many different disciplines, especially in statistically physicals [51], econometrics [52], information science [53] and finance [54].

Bayesian model updating methods initially developed by Beck and Katafygiotis

[55, 56] in structural dynamics and civil engineering field. After that, a detailed up-to-date exposition for these methods is given by Yuen [57].

There are two levels of Bayesian model updating. The first level is parametric identification, in which a class of mathematical models for a particular physical phenomenon is given with unknown parameters to be identified. The second level deals with the selection of a suitable class of mathematical models for parametric identification. Using the first level of Bayesian model updating is the main purpose of the thesis. It is governed by the following Bayes' rule [58]

$$P(\mathbf{\theta}\,|\,\mathbf{D},\mathbf{M}) = \frac{P(\mathbf{D}\,|\,\mathbf{\theta},\mathbf{M})P(\mathbf{\theta}\,|\,\mathbf{M})}{P(\mathbf{D}\,|\,\mathbf{M})} \qquad (3.19)$$

where $P(\mathbf{\theta}\,|\,\mathbf{D},\mathbf{M})$ is the set of uncertain updating parameters in the presence of the data $\mathbf{D}$ and model class $\mathbf{M}$ (posterior probability). The prior probability density function (PDF) $P(\mathbf{\theta}\,|\,\mathbf{M})$ reflects the probability of updating parameters $\mathbf{\theta}$ in the absence of measured information, and it is obtained independently of current measurement results. The $P(\mathbf{D}\,|\,\mathbf{\theta},\mathbf{M})$ is the likelihood function, probability of datasets $\mathbf{D}$ in the presence of fixed updating parameters $\mathbf{\theta}$ and the model class $\mathbf{M}$ is selected. The denominator term $P(\mathbf{D}\,|\,\mathbf{M})$ is a normalization factor, known as evidence, $P(\mathbf{D}\,|\,\mathbf{M}) = \int_{\mathbf{\theta}} P(\mathbf{D}\,|\,\mathbf{\theta},\mathbf{M})P(\mathbf{\theta}\,|\,\mathbf{M})d\mathbf{\theta}$, which ensures $\int_{\mathbf{\theta}} P(\mathbf{\theta}\,|\,\mathbf{D},\mathbf{M})d\mathbf{\theta} = 1$.

Eq. (3.19) could be further simplified to be represented by the prior density and the likelihood functions. Since it is always true when the measured datasets $\mathbf{D}$ are always considered as a constant and the marginal distribution of the datasets $\mathbf{D}$ does not depend on the model parameters $\mathbf{\theta}$ and the model class $\mathbf{M}$ is assigned to only one model to be updated, which is usually omitted to simplify the notation. Therefore, the Eq. (3.19) is then written as

$$P(\mathbf{\theta}\,|\,\mathbf{D}) \propto P(\mathbf{D}\,|\,\mathbf{\theta})P(\mathbf{\theta}) \qquad (3.20)$$

During the last decades, many methods aimed at obtaining theoretically consistent prior PDFs were proposed. The review papers can be found in the work of Jaynes [59] and Berger, et al. [60]. One approach is that of conjugate priors [61], where the prior PDF is chosen in a way that the posterior PDFs belongs to the same distribution family. For example, using a Gaussian prior when the likelihood is

Gaussian will result in a Gaussian posterior PDFs, and it is also said that the family of Gaussian distributions is self-conjugate. However, the most widely used method at present choose the prior PDFs based on the principle of maximum entropy (ME) introduced by Soize [62] for Bayesian inference. The ME principle expresses the probability distribution which best represents the current state of knowledge on updating parameters $\boldsymbol{\theta}$ is the distribution, which results in the largest information entropy. For instance, the ME procedure leads to a uniform prior PDF when it is known that a parameter is non-zero over a finite range of values in a certain interval. For the multivariate situations with no information, the maximum entropy always leads to independent prior variables, which means the joint prior PFDs of $\boldsymbol{\theta}$ is constructed as the product of the marginal prior PDFs of the individual θ_i.

In general, the likelihood function $P(\mathbf{D}|\boldsymbol{\theta},\mathbf{M})$ is typically chosen as a zero mean multivariate normal distribution and covariance $\Sigma = Cov(\boldsymbol{\varepsilon}(\boldsymbol{\theta}),\boldsymbol{\varepsilon}(\boldsymbol{\theta}))$, often justified on the basis that the information entropy [63] for a given mean and covariance is maximized.

$$P(\mathbf{D}|\boldsymbol{\theta},\mathbf{M}) = \prod_{i=1}^{N} \frac{1}{(2\pi)^{(n/2)}|\Sigma|^{1/2}} \exp(-\frac{1}{2}\varepsilon_i^T(\boldsymbol{\theta})\Sigma^{-1}\varepsilon_i(\boldsymbol{\theta})) \qquad (3.21)$$

For Bayesian FE model updating in structural dynamics, the datasets $\mathbf{D}$, represents the modal properties obtained from experiments (natural frequencies $(\omega_i^2)^m$ and mode shapes ϕ_i^m), which are always chosen as eigenvalue residuals, $\varepsilon_i = (z_i^m - z_i(\boldsymbol{\theta}))$, where $z_i^m = (\omega_i^2)^m$, $z_i(\boldsymbol{\theta}) = \lambda_i(\boldsymbol{\theta})$, $i = 1,\ldots,n$, which represent the square of the ith measured natural frequency and the ith eigenvalue of FE model, respectively. When the prior PDF and likelihood function are determined, Eq. (3.20) allows for the posterior updating PDFs of the parameters based on the experimental measurements.

Although Bayes' theorem is intuitively appealing, however, in most practical applications, it requires multidimensional integration to acquire the marginal posterior probabilities of parameters which makes it impractical when the posterior distribution $P(\boldsymbol{\theta}|\mathbf{D},\mathbf{M})$ evaluation is always complex and unlikely to be evaluated analytically. Therefore, the sampling techniques named the Markov Chain Monte Carlo (MCMC) methods can be used to approximate these complex posterior PDFs. An advantage of MCMC methods is that they can provide a full characterization of the posterior uncertainty when the data is sufficient to constraint the updating

parameters [64]. Thereby, they are practical for cases where no analytical solutions are available for the posterior PDFs.

MCMC sampling methods [65, 66] form the most widespread used sampling techniques in Bayesian inference, most often using the Metropolis-Hastings (MH) algorithm described section 3.4.1 to sample the posterior PDFs $P(\boldsymbol{\theta}|\mathbf{D},\mathbf{M})$. In recent years, some methods were developed to improve the convergence speed of the classic MH algorithm. These methods include adaptive MH MCMC [67], delayed rejection and adaptive MH MCMC [68]. Another group of improved MCMC approaches involves sequential sampling strategy, which is developed by Au, et al. [69], and the transitional Markov Chain Monte Carlo (TMCMC) method [70] is considered as such a representative sampling technique for Bayesian model updating.

In this section, the classic Metropolis-Hastings (MH) approach and the transitional Markov Chain Monte Carlo (TMCMC) algorithm are described in detail, which also has been used in the thesis for Bayesian model updating in section 3.6 and chapter 5.

3.4.1 The Metropolis-Hastings Algorithm

The MH algorithm is one of the simplest MCMC methods that is able to draw samples from multivariable densities [71]. It is related to rejection and importance sampling and its general idea is to propose a PDF and then use it to generate proposed values. The proposed distribution is also used to obtain a move probability which is used to determine whether the calculated values should be accepted as the next state of the Markov chain. The move probability is defined by the product of the ratio of the target density and the ratio of the proposed density which means that a normalizing constant of the target density distribution function is not required in this algorithm.

In the MH algorithm, to sample from the posterior distribution function $p(\boldsymbol{\theta}|\mathbf{D})$, which is the target distribution function, where $\boldsymbol{\theta}=\{\theta_1,\dots\theta_n\}$ is an n-dimensional parameter vector, the proposal density distribution $Q(\boldsymbol{\theta}|\boldsymbol{\theta}_{n-1})$ is introduced in order to generate a random vector $\boldsymbol{\theta}$ given the value at the previous iteration of the algorithm, which can be summarized as follows

1. Begin the algorithm with an initial value $\boldsymbol{\theta}_0$.

2. At the iteration n, $\boldsymbol{\theta}^{\bullet}$ is drawn from the proposal probability distribution

density $Q(\theta \mid \theta_{n-1})$, where θ_{n-1} is the parameter value of the previous step.

3. Update the FE model to obtain the new analytical responses, then computer the acceptance probability, given by

$$\alpha(\theta^*, \theta_{t-1}) = \min\{1, \frac{p(\theta^* \mid D)Q(\theta_{t-1} \mid \theta^*)}{p(\theta_{t-1} \mid D)Q(\theta^* \mid \theta_{t-1})}\} \tag{3.22}$$

4. Draw u from a uniform distribution $u \sim U(0,1)$.

5. If $u \leq \alpha(\theta^*, \theta_{t-1})$ then accept state θ^*. Otherwise, reject θ^*.

6. Repeat Step 2.

Figure 3.1, Figure 3.2, Figure 3.3, and Figure 3.4 show the results of a single variable example using the MH algorithm. The proposal distribution is settled as $Q(\theta) = N(1,1)$, and the single variable target distribution is given by $p(x) = \dfrac{\exp(-x^2)(2+\sin(5x)+\sin(2x))}{\int_{-\infty}^{+\infty} \exp(-x^{'2})(2+\sin(5x^{'})+\sin(2x^{'}))dx^{'}}$, which can also be simplified as $p(x) \propto \exp(-x^2)(2+\sin(5x)+\sin(2x))$, is plotted for 500 and 5000 iterations (samples), respectively.
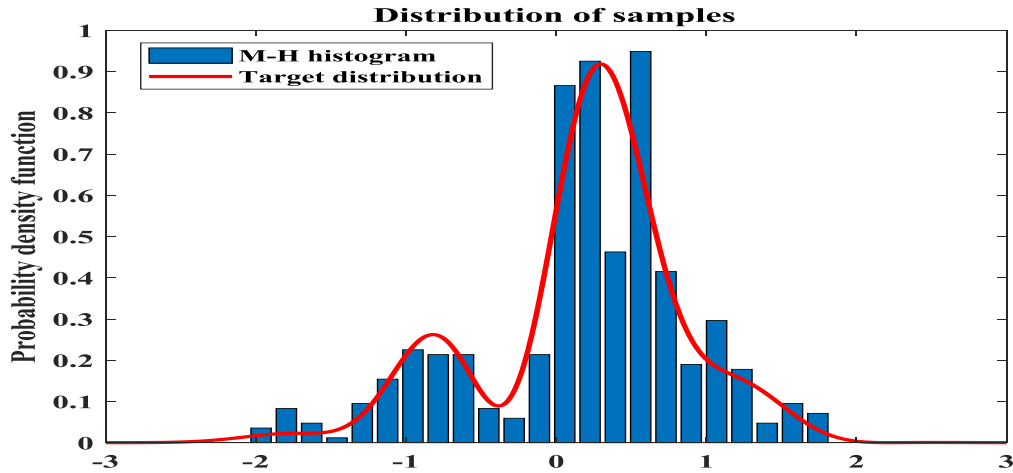
Figure 3.1 *Target distribution and histogram of the MH method with 500 iterations.*
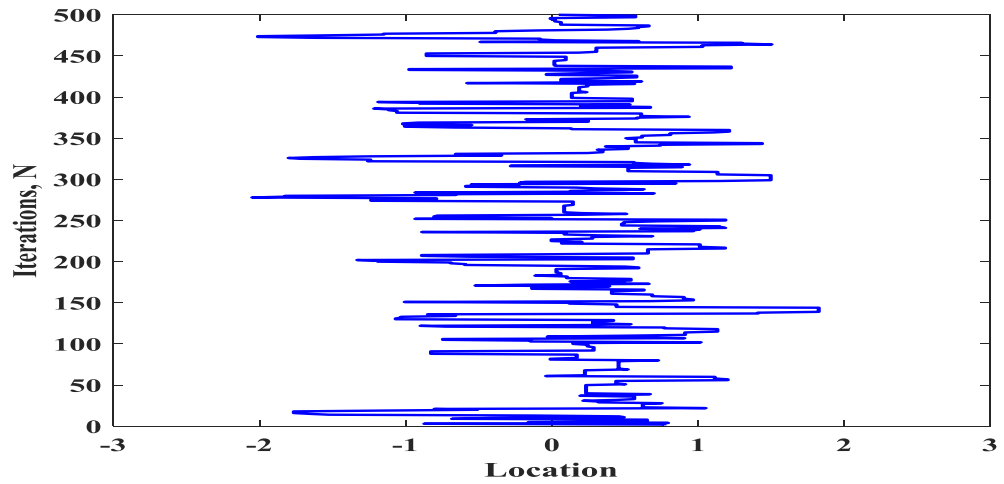
Figure 3.2 *Trace history of the MH method with 500 iterations.*
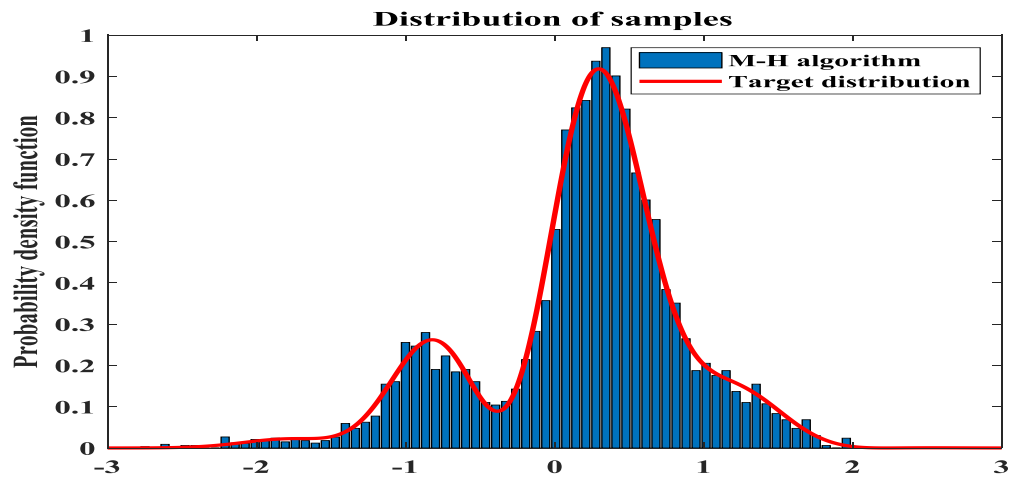
Figure 3.3 *Target distribution and histogram of the MH method with 5000 iterations.*
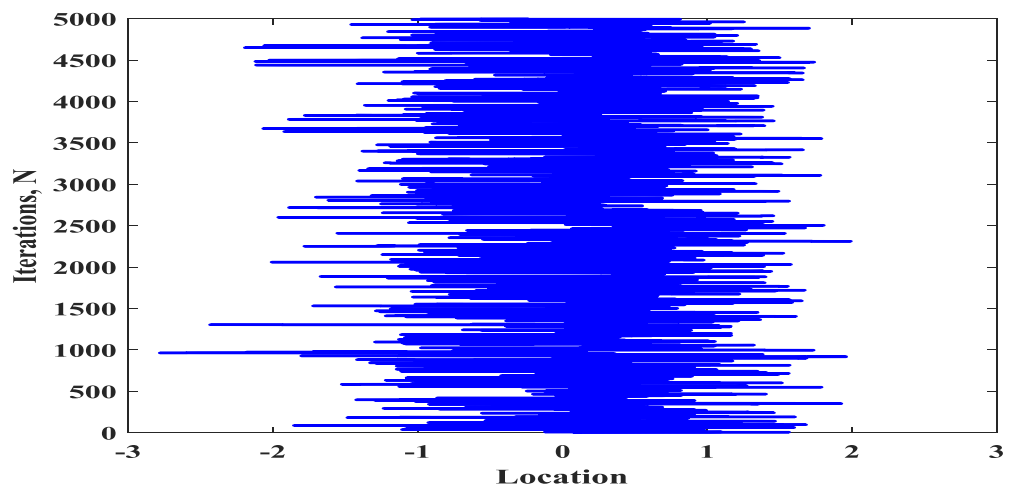
Figure 3.4 *Trace history of the MH method with 5000 iterations.*

As expected, from Figure 3.1 to Figure 3.4, the histogram of samples approximate the target distribution. However, in Figure 3.1 and Figure 3.2, 500 iterations are looked not enough to give a good representation of the target distribution given as $p(x) \propto \exp(-x^2)(2 + \sin(5x) + \sin(2x))$. This is because the acceptance rate is 48.6%, which means only less than half samples (500) is selected to approximate the target distribution and also due to some other aspects' influences like the start point, burn-in period, the lag, etc. By comparison in Figure 3.3 and Figure 3.4 through fixing aspects' influences mentioned above, increasing the number of iterations to 5000 iterations with the acceptance rate 48.9% leads to an almost excellent prediction of the form to the target distribution given as $p(x) \propto \exp(-x^2)(2 + \sin(5x) + \sin(2x))$.

3.4.2 The transitional Markov Chain Monte Carlo algorithm

The TMCMC [70], is motivated by a related adaptive MH method [72]. The fundamental basis of TMCMC is that samples are taken from a series of intermediate PDFs in an adaptive manner, expressed as

$$
\begin{aligned}
&p_i(\boldsymbol{\theta} \,|\, \mathbf{D}) \propto p(\boldsymbol{\theta}) p(\mathbf{D} \,|\, \boldsymbol{\theta})^{s_i} \\
&i = 0, \ldots, I; 0 = s_0 < s_1 < \ldots < s_I = 1
\end{aligned} \tag{3.23}
$$

where i is the stage number and s_i is the corresponding tempering parameter for the ith stage, which controls the speed of this gradual transition and is automatically computed in the process to form the intermediate PDFs. It can be seen from Eq. (3.23) that the series start with prior PDF $p(\boldsymbol{\theta})$ (when $i = 0$ and $s_0 = 0$) and convergence to the posterior PDF $p(\boldsymbol{\theta} \,|\, \mathbf{D})$ (when $i = 1$ and $s_I = 1$). Note that normalization is not necessary since only the relative probability densities are required.

Given the ith stage samples $\{\theta_{i,1}, \theta_{i,2}, \ldots, \theta_{i,n}\}$ from $p_i(\boldsymbol{\theta} \,|\, \mathbf{D})$, the $p_{i+1}(\boldsymbol{\theta} \,|\, \mathbf{D})$ samples of the $(i+1)$th stage are obtained by a resampling procedure, which can be achieved as follows [73]:

The plausibility weights of samples $\{\theta_{i,1}, \theta_{i,2}, \ldots, \theta_{i,n}\}$ with respect to $p_{i+1}(\boldsymbol{\theta} \,|\, \mathbf{D})$ are first computed according to

$$v(\theta_{i,k}) = \frac{p(\mathbf{D}|\theta_{i,k})^{s_{i+1}} p(\theta_{i,k})}{p(\mathbf{D}|\theta_{i,k})^{s_i} p(\theta_{i,k})} = p(\mathbf{D}|\theta_{i,k})^{s_{i+1}-s_i}, \ k=1,\ldots,n \qquad (3.24)$$

Having the plausibility weights, the uncertain model parameters are then resampled based on the weights such that

$$\theta_{i+1,k} = \theta_{i,t} \text{ with probability } \frac{v(\theta_{i,t})}{\sum_{t=1}^{N} v(\theta_{i,t})}, \ k=1,\ldots,\text{n}, \ t=\text{dummy index} \quad (3.25)$$

It can be seen that if N is large, samples $\{\theta_{i,1},\theta_{i,2},\ldots,\theta_{i,n}\}$ will be generated according to the intermediate PFDs $p_i(\boldsymbol{\theta}|\mathbf{D})$.

The expectation of $v(\theta_{i,k})$ can be estimated by the average of $\theta_{i,k}$ samples, $k=1,\ldots,\text{n}$

$$E(v(\theta_{i,k})) = \frac{\int p(D|\theta)^{s_{i+1}} p(\theta)d\theta}{\int p(D|\theta)^{s_i} p(\theta)d\theta} \approx \frac{1}{N} \sum_{k=1}^{N} v(\theta_{i,k}) \qquad (3.26)$$

and the evidence $p(\mathbf{D})$ can be evaluated as

$$P(\mathbf{D}) = \frac{\int p(\mathbf{D}|\theta)p(\theta)d\theta}{\int p(\theta)d\theta} = \frac{\int p(\mathbf{D}|\theta)^{s_I} p(\theta)d\theta}{\int p(\mathbf{D}|\theta)^{s_0} p(\theta)d\theta} \qquad (3.27)$$

$$= \prod_{i=1}^{I} E(v(\theta_{i,k}))$$

In chapter 5, TMCMC will be used to perform Bayesian FE model updating described on the DLR-AIRMOD structure.

3.4.3   Surrogate modelling techniques in combination

The main drawback of implementing Bayesian model updating for large-scale engineering structures for model updating problems is related to the intensive computation, e.g. carrying out Bayesian FE model updating for the DLR-ALRMOD structure described in section 5.5.

Besides using the advanced MCMC methods such as TMCMC [70] to develop the computational efficiency which allows the generation of samples to the updated posterior distribution for high dimensions, utilising the high-fidelity surrogate models described in chapter 4 and 5 to replace the full FE models are currently other aspects of interests for Bayesian FE model updating in order to alleviate the computation cost without losing accuracy.

3.4.4   Statistical summaries

The merits of using Bayesian methods for FE model updating by MCMC techniques is that the uncertainty of the updated parameters is presented and the correlations between them can be measured. Once the posterior PDFs is approximated, it can be used to obtain information on the resolution of the parameters through the Bayesian scheme. The resolution analysis typically includes comparing the mean values, variance, and covariance matrix.

The approximated mean value of updated parameters is expressed as

$$\hat{\boldsymbol{\theta}} = E(\boldsymbol{\theta}) \cong \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\theta}^i \tag{3.28}$$

where $E(\boldsymbol{\theta})$ is the mean value of $\boldsymbol{\theta}$ and N is the number of samples. The uncertainty of the estimated mean values for updated parameters can be characterized by the variance, expressed as

$$V(\hat{\boldsymbol{\theta}}) = E((\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^2) \cong \frac{1}{N} \sum_{i=1}^{N} (\boldsymbol{\theta}^i - \hat{\boldsymbol{\theta}})^2 \tag{3.29}$$

The covariance matrix $\boldsymbol{\Sigma_\theta}$ is used to identify the correlation between the updated parameters is written as

$$\boldsymbol{\Sigma_\theta} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \tag{3.30}$$

where $\quad c_{ij} = \dfrac{cov(\theta_i, \theta_j)}{\sqrt{E((\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i)^2)E((\boldsymbol{\theta}_j - \hat{\boldsymbol{\theta}}_j)^2)}} \quad , \quad cov(\theta_i, \theta_j) = E((\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i)(\boldsymbol{\theta}_j - \hat{\boldsymbol{\theta}}_j))$

and the computed values c_{ij} are bounded between -1 and 1.

3.5   Non-probabilistic stochastic model updating methods

Apart from describing the probabilistic stochastic model updating methods, a brief description of non-probabilistic stochastic model updating methods including interval model updating methods and fuzzy model updating methods is also provided.

3.5.1   The interval model updating method

The interval model updating methods stem from interval analysis [74]: the

parameter uncertainty may be modeled using a range between lower and upper bounds. The aim of interval analysis is to evaluate the range of possible outputs considering all possible combinations of the uncertain inputs within their permissible range.

If each component of the response vector is generally represented by $y_i(\mathbf{x})$, the interval analysis is a numerical procedure equivalent to solving the following equations

$$\overline{y}_i = \max(y_i(\mathbf{x})), \quad \underline{y}_i = \min(y_i(\mathbf{x})), \quad i=1,\ldots,\text{n} \tag{3.31}$$

Subject to

$$\underline{\mathbf{x}} < \mathbf{x} < \overline{\mathbf{x}} \tag{3.32}$$

The operations on intervals, can be used to solve Eq. (3.31) when the response function $y_i(\mathbf{x})$ has an analytical expression in closed-form and a global optimization procedure is always considered as the most general solution. This optimization is carried out independently on every component of the response vector $\mathbf{y}$.

The interval analysis approach has been applied to the FE model updating by Gabriele and Valente [74] and Khodaparast et al. [75] in recent years. The interval method can also be seen as a specific case of the more general convex modeling approach [76], where it is assumed that the uncertain quantities lie within a convex region. The paper for interval FE model updating is in the work of Khodaparast et al. [75], who applied the interval model updating approach for frame structures with uncertain beam positions. This interval FE model updating approach is formulated based on the Kriging predictor [89] and an iterative model updating procedure.

3.5.2 The Fuzzy model updating method

The fuzzy model updating methods stems from the fuzzy set analysis, which is the extension of interval analysis originally attributed by Zadeh [77]. The fuzzy set method carries out uncertainty modeling in situations where besides the interval bounds, confidence values regarding the uncertain quantities are also available. It is about to determine the fuzzy description of outputs when the inputs are modeled using fuzzy sets. The fuzzy set is considered as an extension of a conventional set, which discriminates between elements that belong to the set and those which do

not.

In the work of Erdogan and Bakir [78], Erdogan, et al. [79] and Khodaparast, et al. [80] proposed fuzzy model updating in structural dynamics. More precisely, Erdogan and Bakir [78] and Erdogan et al. [79] used the fuzzy models to model the uncertainty due to the measurement noise which has dealt with reducible uncertainty, and Khodaparast et al. [80] proposed a method to compute measured fuzzy membership functions of experimental data and then a fuzzy parameter is introduced to represent a set of interval parameters through the membership function. In addition, the kriging surrogate model is used to speed up the fuzzy FE model updating procedure.

3.6 Comparison of uncertainty identification with probabilistic model updating approaches

In this section, the methods introduced in section 3.3 and 3.4 namely the probabilistic perturbation method and the Bayesian model updating method are applied to the three degree of freedom mass-spring system shown in Figure 3.5 to investigate their performance for uncertainty identification.
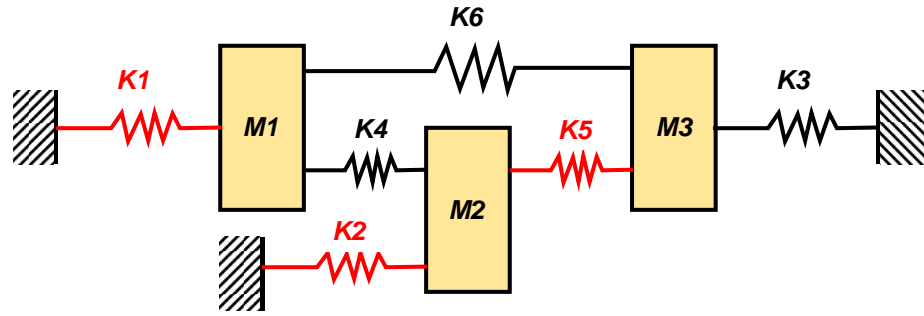
Figure 3.5 *Three degree of freedom mass-spring system.*

The nominal values of parameters to the 'synthetic experimental' system are given in Table 3.1 with $m_i = 1.0 kg (i = 1, 2, 3)$, $k_i = 1.0 N/m$ $(i = 1, \ldots, 5)$ and $k_6 = 3.0 N/m$. The updating random parameters are assumed to have Gaussian distributions with mean vector $u_{k_1} = u_{k_2} = u_{k_5} = 2.0 N/m$ and standard deviation $\sigma_{k_1} = \sigma_{k_2} = \sigma_{k_5} = 0.3 N/m$. The true mean values are the nominal values with a standard deviation $\sigma_{k_1} = \sigma_{k_2} = \sigma_{k_5} = 0.2 N/m$. The updating random parameters k_1, k_2, k_5 are assumed to be independent.

Table 3.1 *Parameter values for the 'experimental' system*

Mass vector (kg)		Stiffness vector (N/m)			
m_1	1	k_1	1	k_2	1
m_2	1	k_3	1	k_4	1
m_3	1	k_5	1	k_6	3

For this case, three uncertain updating parameters, k_1, k_2, k_5 are assumed responsible for the observed variability in the three natural frequencies of the system.
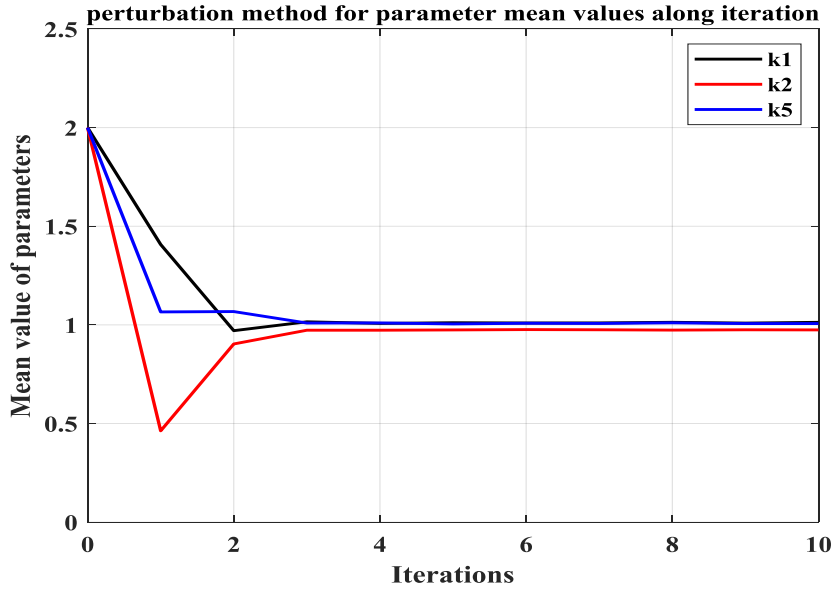
Figure 3.6 *Parameter convergence of mean values by perturbation method.*
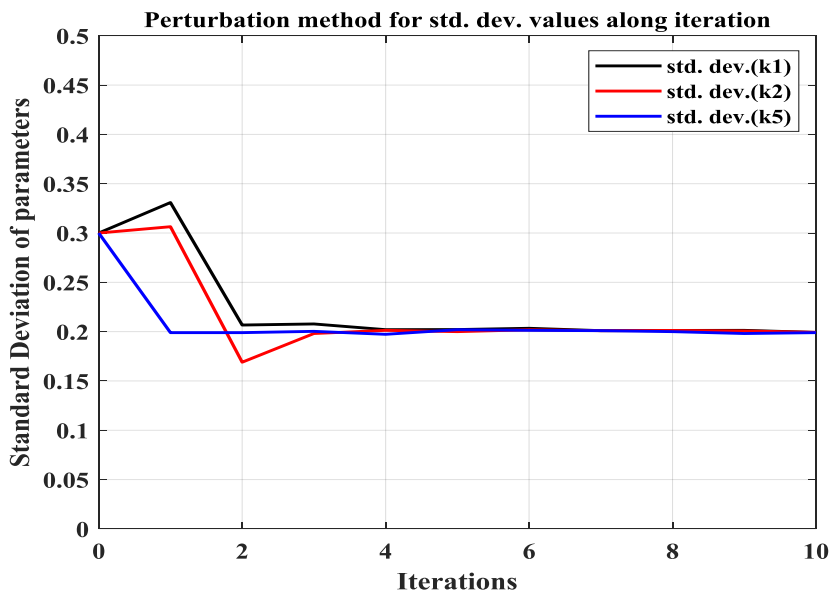
Figure 3.7 *Parameter convergence of standard deviation by perturbation method.*
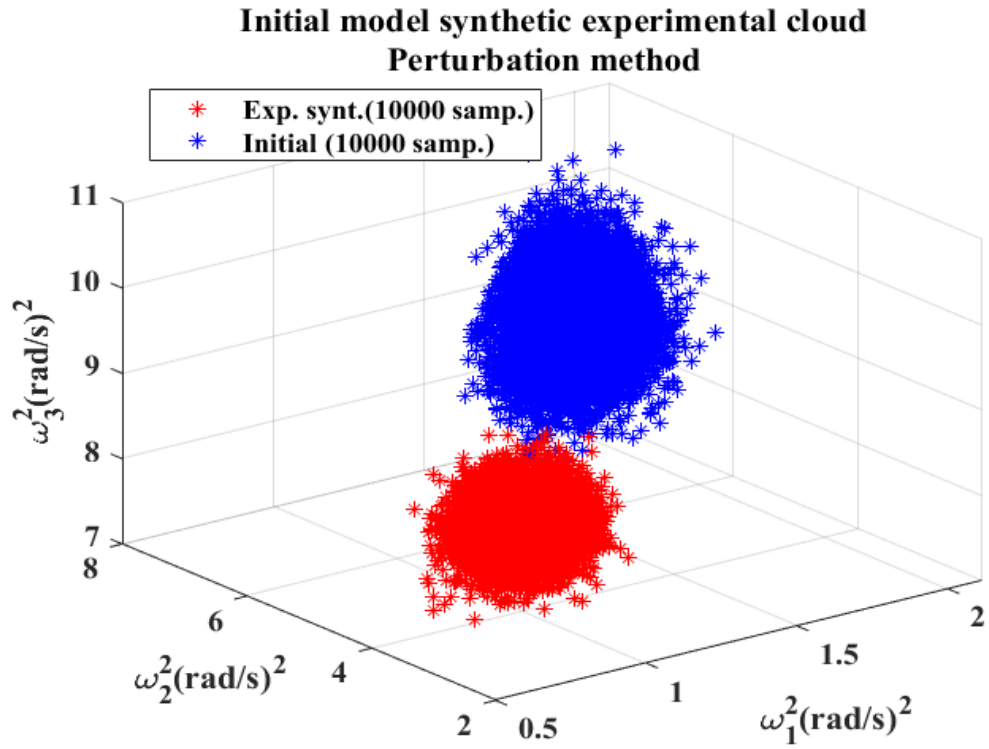
Figure 3.8 *Initial model responses cloud (10000 samples) versus Synt. Exp.*
responses cloud (10000 samples).
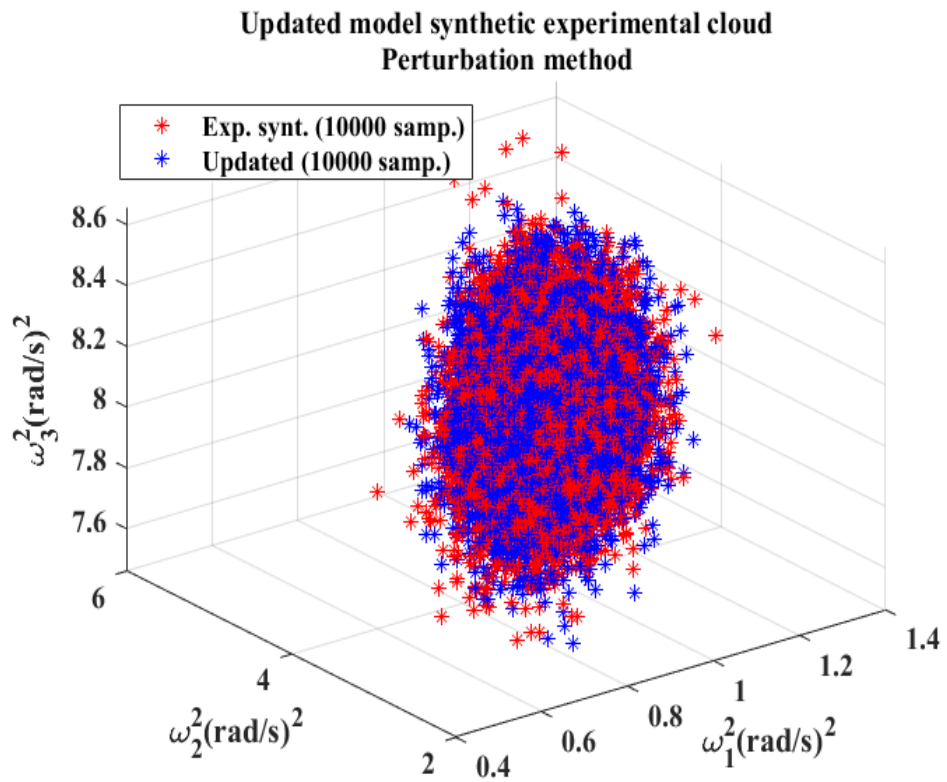
Figure 3.9 *Updated model responses cloud (10000 samples) versus Synt. Exp.*
responses cloud (10000 samples).
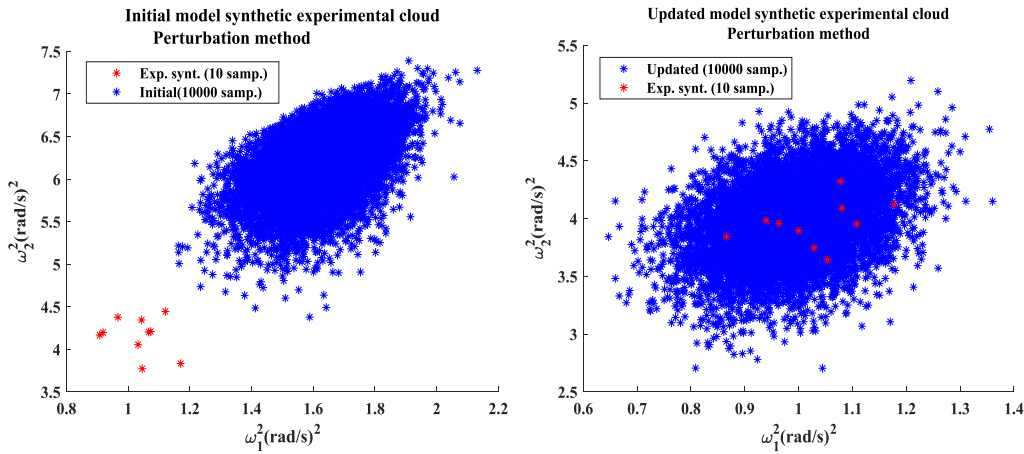
Figure 3.10 *Initial model responses cloud 10000 samples versus Synt. Exp. responses cloud 10 samples (left) and Updated model responses cloud 10000 samples versus Synt. Exp. responses cloud 10 samples (right) in $\omega_1{}^2$ and $\omega_2{}^2$ plane.*
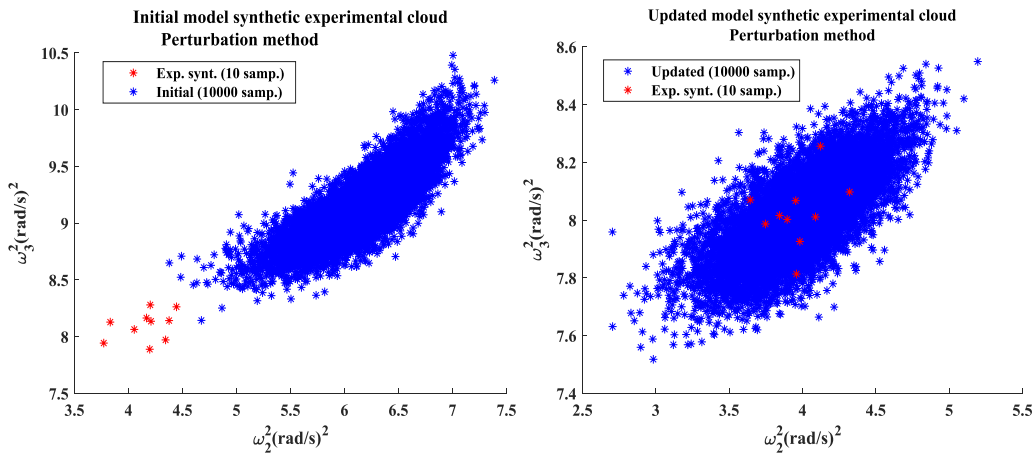
Figure 3.11 *Initial model responses cloud 10000 samples versus Synt. Exp. responses cloud 10 samples (left) and Updated model responses cloud 10000 samples versus Synt. Exp. responses cloud 10 samples (right) in $\omega_2{}^2$ and $\omega_3{}^2$ plane.*

The probabilistic perturbation method described in detail in section 3.3 is employed. The parameter convergence of mean and standard deviation is shown in Figure 3.6 and Figure 3.7 respectively, and the initial cloud of predicted round frequencies is made to converge to the cloud of 'experimental' predicted round frequencies shown from Figure 3.8 to Figure 3.11. In addition, the converged response cloud is shown to be a good agreement with the 'synthetic experimental' dataset.

The Bayesian model updating method described in detail in section 3.4 is also used

for performance comparizon. Figure 3.12, Figure 3.13, and Figure 3.14 show the updating parameter probability densities, CDFs and the trace history for k_1, k_2, k_5, which is obtained using the Bayesian model updating method with the MH algorithm described in section 3.4.1. with 15000 samples (2000 burn-in samples) in order to have a full characterization of the updated posterior distributions.

The model updating results summary obtained by the probabilistic perturbation method and the Bayesian model updating method is shown in Table 3.2. It can be concluded that the updated results by these two methods to mean values, standard deviations for parameters and predicted responses of the system are all in good agreement with the 'synthetic experimental' datasets.

More importantly, it is also worth noting that, the running time for the Bayesian model updating to the three degree of freedom mass-spring system is much longer than probabilistic perturbation method.
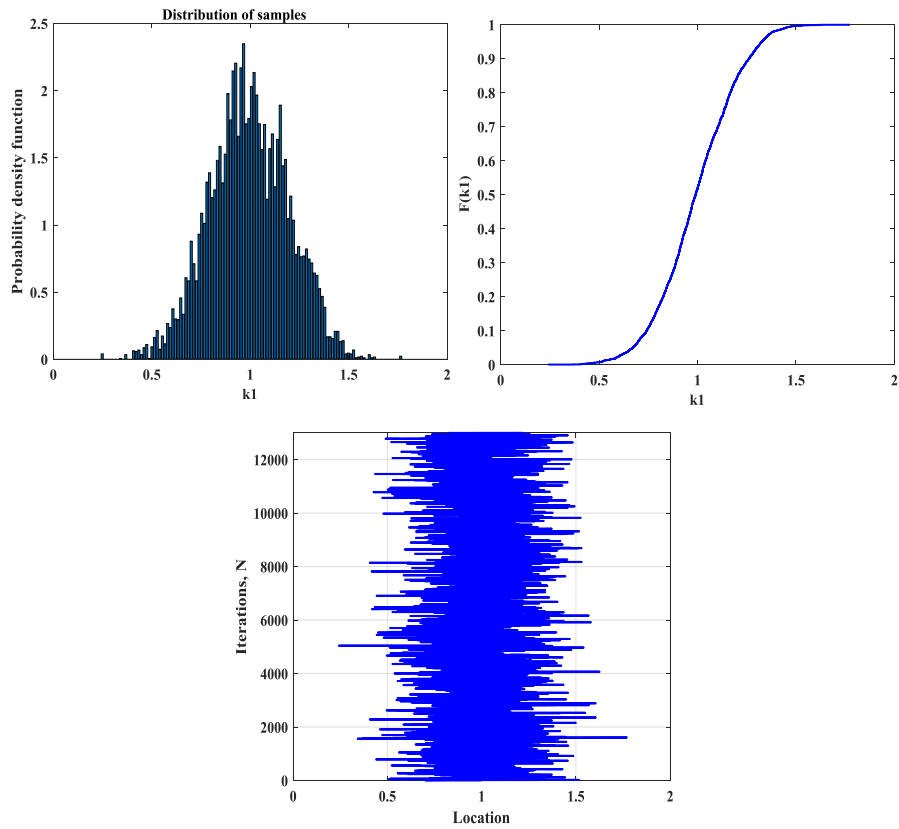
Figure 3.12 *Parameter PDFs, CDFs and Metropolis-Hastings trace history for* k_1.
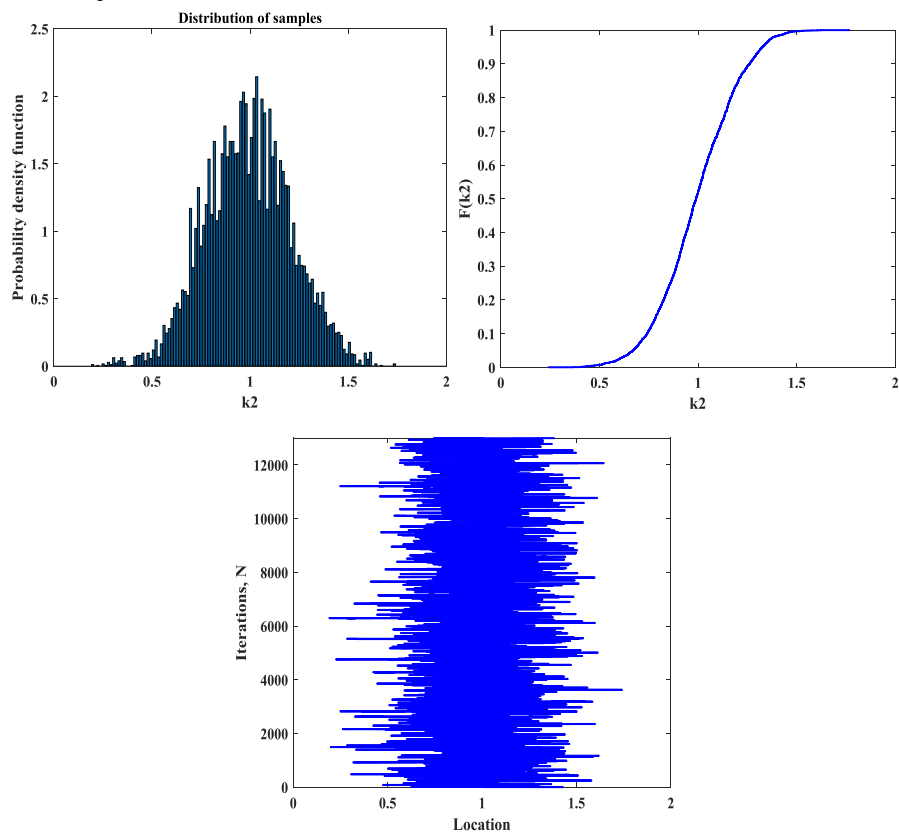
Figure 3.13 *Parameter PDFs, CDFs and Metropolis-Hastings trace history for* k_2.
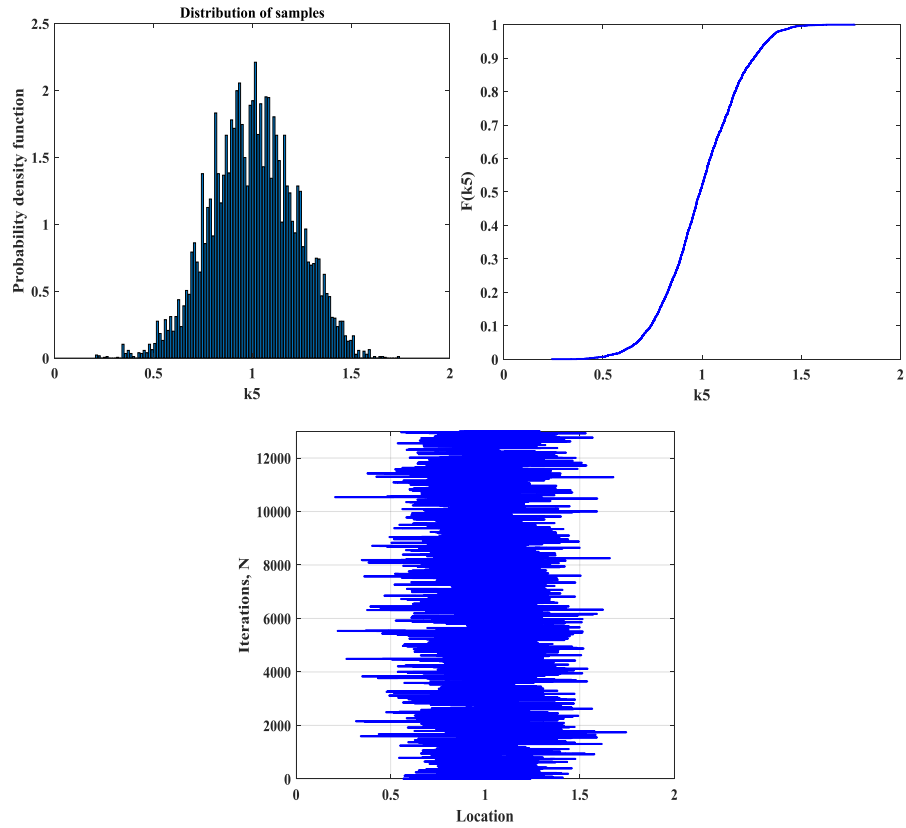
Figure 3.14 *Parameter PDFs, CDFs and Metropolis-Hastings trace history for*
k_5.

Table 3.2 *Model updating results summary*

	Value	Perturbation method	Errors (%)	Bayesian model updating method	Errors (%)
$\bar{k}_1$	1	1.011	-1.1	1.009	-0.9
$\bar{k}_2$	1	0.972	2.5	0.976	2.4
$\bar{k}_5$	1	1.003	-0.3	1.005	-0.5
σ_1	0.1997	0.1983	0.7	0.1982	0.7
σ_2	0.1997	0.1992	0.25	0.1993	0.2
σ_5	0.1997	0.2021	-1.2	0.2107	-5.5
$\bar{\omega}_1^{\,2}$ (rad/s)	0.9930	0.9925	0.5	0.9930	0
$\bar{\omega}_2^{\,2}$ (rad/s)	3.9896	3.9896	0	3.9896	0
$\bar{\omega}_3^{\,2}$ (rad/s)	8.0027	8.0026	0.001	8.0028	-0.001

3.7 Chapter summary

In this chapter, an introduction to the stochastic FE model updating has been provided. The methods are generally categorized into two groups, i.e. probabilistic methods and non-probabilistic methods. At present, the existing widespread used stochastic FE model updating methods that incorporate statistics are probabilistic perturbation method, and Bayesian model updating method, which are described in detail in this chapter. The non-probabilistic FE stochastic model updating methods including the interval model updating methods and the fuzzy model updating methods are also briefly described.

In addition, by applying the probabilistic perturbation method, and Bayesian model updating method to a three degree of freedom mass-spring system. It is found that these two methods are all applicable to the problem of model updating in the presence of uncertainty in this three degree of freedom mass-spring system with high accuracy. However, the running time of conducting Bayesian model updating when combined with the MH algorithm is much longer than the probabilistic perturbation method.

It is also important for Bayesian FE model updating methods combined with advanced MCMC methods described in section 3.4.2 and surrogate model techniques described in chapter 4 and 5 in order to solve a computationally intensive model updating task (e.g. the DLR AIRMOD structure).

4 Gaussian process emulation

4.1 Surrogate modelling

Simulators or computer codes are often used to model physical processes in science and engineering [81]. In practice, simulators which model complex physical phenomena with high-fidelity are usually computationally expensive. That is, a single input-output evaluation of the given computer codes takes a long time to prohibit statistically meaningful ensembles of experiments to be run in a practical amount of time.

In order to alleviate the cost of the expensive models, a stochastic representation of model outputs can be utilized. These approximations are called metamodels or surrogate models [82]. In the recent years, many kinds of surrogate models have been used in different research fields. These approaches include response surface methods [83], Taylor expansions [84], artificial neural networks [85], radial basis functions [86], support vector machines [87] and polynomial chaos expansions [88]. Metamodels have their own merits and disadvantages. However, none of these methodologies is statistical in its nature and is thus unable to provide an estimation of the uncertainty introduced while using the surrogate models without employing additional tools.

The Gaussian process emulator (GPE), also called Bayesian emulator or Kriging [89], in contrast, is a statistical surrogate model. It works by using the Bayesian principle: for fixed sample points in the input domain, it consists in assigning a prior distribution to the model output and updating it with more observed data, which results in a predictive posterior distribution for the model outputs. This produces not only an expectation but also a variance, which quantify the uncertainty of an emulator's approximation for the model outputs.

4.2 A brief history of Gaussian processes

The Gaussian process (GP) is a general class of probability distribution on functions. By viewing this general setting, the well known Wiener process [90] is actually a particular type of Gaussian process. Gaussian processes were first used

for time series prediction could date back to the 1940s [91]. Gaussian processes have been widely used since the 1970s in the fields of geostatistics and meteorology. In geostatistics, Gaussian processes prediction is termed Kriging, (after the South African mining engineer D. G. Krige) by Matheron [92]. The inputs to the process are only two or three dimensions.

In 1978, O'Hagan [93] first applied the GPs to slightly more general multivariate input regression problems in spatial statistics. After that, the GPs was described by Williams and Rasmussen in 1996 [94] to the machine learning community. At that time the neural networks by Bishop [95] were in vogue as general purpose function approximators, Williams and Rasmussen's work [94] were partly inspired between GPs and neural networks.

Over the past decade, the work for Gaussian processes has been actively developed and extended both in statistics and machine learning communities. In machine learning communities, the tasks are broadly divided into three fundamental classes: supervised learning, unsupervised learning, and reinforcement learning [96]. The most widely studied class is supervised learning, which concerns about learning a relationship from inputs to outputs. The supervised learning may be further subdivided into two primary tasks: classification and regression. Moreover, In classification, the outputs are discrete labels, whereas in regression the outputs are continuous variables.

The GPE is built in the form of a full probabilistic Bayesian model. By defining a probability distribution on functions $p(f)$, which is used as a prior for the regression, then Bayesian inference is used to make predictions of the function output given a dataset $\mathbf{D}$ by Bayes' theorem as:

$$p(f \mid \mathbf{D}) = \frac{p(\mathbf{D} \mid f) p(f)}{p(\mathbf{D})} \tag{4.1}$$

This is a relatively high level of description of how a GP solves the regression task described above, which allow us to have probabilistic predictions of possible interpolating functions f.

4.3 Gaussian process emulation

The GPE, also called Bayesian emulator [89] or Kriging, is a stochastic approximation that links complex simulators. This means that a GPE is capable of

quantifying the uncertainty about the underlying function which arises from the fact that the output of the simulator has not been observed at all locations of interest. In a Bayesian setting for constructing the GPE, the uncertainty about parameters is represented through probability distributions. This results in a full probabilistic model that takes into account uncertainty in the parameters and in the data generating process.

4.3.1   Surrogate model construction

By assuming the outputs are noise-free, observed from an underlying functional mapping $\eta(\mathbf{x})$, the GPE provides an interpolator, which is of the form

$$\eta(\mathbf{x}) = h(\mathbf{x})^T \boldsymbol{\beta} + Z(\mathbf{x}) \tag{4.2}$$

where $h(\cdot) \in \mathbb{R}^p$ is any real-valued function that maps $\mathbb{R}^p \to \mathbb{R}^q$. It consists in q known regression functions of $\mathbf{x}$, and is chosen to incorporate the beliefs that we have about $\eta(\mathbf{x})$. $\boldsymbol{\beta}$ is a vector of unknown coefficients and $Z(\mathbf{x})$ is a Gaussian process with zero mean and covariance $\sigma^2 c(\mathbf{x}, \mathbf{x}'; \varphi)$ where σ^2 is a scalar parameter and φ is a set of parameters describing behavior for the correlation function $c(\cdot, \cdot; \varphi)$.

Definition 2.1 (*Gaussian stochastic process*) For $x \in \mathbb{R}^d$, $\mathrm{GP}(\cdot)$ is a Gaussian stochastic process if for any $d \geq 1$ and any choice from $\mathbf{x}_1, \ldots, \mathbf{x}_d$, the vector $\left[Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_d) \right]^T$ has a multivariate Gaussian distribution.

A zero mean multivariate Gaussian distribution can be expressed as follows

$$\begin{bmatrix} Z_1(\mathbf{x}) \\ \vdots \\ Z_n(\mathbf{x}) \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} k_{11} & \cdots & k_{1n} \\ \vdots & \ddots & \vdots \\ k_{n1} & \cdots & k_{nn} \end{bmatrix} \right) \tag{4.3}$$

And the covariance matrix can be expressed as a scalar parameter σ^2 times the correlation matrix $\mathbf{C}$ expressed as

$$\sigma^2 \mathbf{C} = \begin{bmatrix} \sigma^2 & \sigma^2 c(\mathbf{x}_1, \mathbf{x}_2) & \cdots & \sigma^2 c(\mathbf{x}_1, \mathbf{x}_n) \\ \sigma^2 c(\mathbf{x}_2, \mathbf{x}_1) & \sigma^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2 c(\mathbf{x}_n, \mathbf{x}_1) & \cdots & \cdots & \sigma^2 \end{bmatrix} \tag{4.4}$$

Since any finite collection of points from a Gaussian process is distributed as a

multivariate Gaussian distribution, the distribution of the outputs at unobserved input values $\mathbf{x}^*$, given already obtained datasets must follow

$$\eta(\mathbf{x}^*) \mid \mathbf{D} \sim N(m(\mathbf{x}^*), c(\mathbf{x}^*, \mathbf{x}^*)) \tag{4.5}$$

with some kinds of mean and covariance functions.

Attention needs to be paid to

$$\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2 \sim N(\mathbf{H}\boldsymbol{\beta}, \sigma^2 \mathbf{C}) \tag{4.6}$$

where $\mathbf{H} = (h(\mathbf{x}_1), \ldots, h(\mathbf{x}_n))^T$ leads to the notation that in order to obtain a distribution for the output of simulator unconditional on the GPE parameters in Eq. (4.2), the values need to be estimated from datasets. The regression coefficients $\boldsymbol{\beta}$ and the scalar parameter, σ^2 can be obtained analytically. The parameters for the correlation function, φ, which is implied through the correlation matrix $\mathbf{C}$ can be estimated using optimization algorithms [104-106].

In Eq.(4.6), the likelihood of $\boldsymbol{\beta}$ and σ^2 can be expressed as

$$p(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2} \mid \mathbf{C} \mid^{1/2}} \exp[-\frac{(\mathbf{y} - \mathbf{H}\boldsymbol{\beta})^T \mathbf{C}^{-1}(\mathbf{y} - \mathbf{H}\boldsymbol{\beta})}{2\sigma^2}] \tag{4.7}$$

The posterior distribution of the parameters conditioned on a dataset can be obtained using the following procedure. Assuming a weak prior for the joint distribution of $\boldsymbol{\beta}$ and σ^2, which implies that for any $\mathbf{x}$, the variance of $\eta(\mathbf{x})$ is infinite [99].

$$p(\boldsymbol{\beta}, \sigma^2) \propto \frac{1}{\sigma^2} \tag{4.8}$$

This prior is known as a Jeffreys prior [97], which has a convenient feature that is invariant under reparameterization and proportional to the square root of the determinant of the Fisher information matrix [98], and in this case, it is proportional to σ^{-2}. Based on the Eq. (4.6), the posterior distribution of $\boldsymbol{\beta}$ and σ^2 are expressed as follows [99]

$$p(\boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}) \propto p(y \mid \boldsymbol{\beta}, \sigma^2) p(\boldsymbol{\beta}, \sigma^2) \tag{4.9}$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}^{-1} \mathbf{y} \tag{4.10}$$

Separating from the distribution of $\boldsymbol{\beta}$ and combining terms results in an inverse gamma posterior distribution for σ^2 with an unbiased estimator

$$\hat{\sigma}^2 = \frac{\mathbf{y}^T (\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{H}(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}^{-1})\mathbf{y}}{n - q - 2} \tag{4.11}$$

By combing Eq. (4.6), Eq. (4.10) and Eq. (4.11), which gives posterior predictive distributions for outputs conditional on the observed datasets expressed as follows

$$\eta(\mathbf{x}^* \mid D, \hat{\boldsymbol{\beta}}, \sigma^2) \sim N(m(\mathbf{x}^*), c(\mathbf{x}^*, \mathbf{x}^*)) \tag{4.12}$$

with posterior predictive mean

$$m(\mathbf{x}^\bullet) = h(\mathbf{x}^\bullet)\hat{\boldsymbol{\beta}} + \mathbf{t}(\mathbf{x}^\bullet)^T \mathbf{C}^{-1}(\mathbf{y} - \mathbf{H}\hat{\boldsymbol{\beta}}) \tag{4.13}$$

and posterior predictive covariance expressed as

$$\begin{aligned} c_1(\mathbf{x}^\bullet, \mathbf{x}^{'\bullet}) = \hat{\sigma}^2 [c_2(\mathbf{x}^\bullet, \mathbf{x}^{'\bullet}) + (h(\mathbf{x}^\bullet)^T - \mathbf{t}(\mathbf{x}^\bullet)^T \mathbf{C}^{-1}\mathbf{H}) \\ \times (\mathbf{H}^T \mathbf{C}^{-1}\mathbf{H})^{-1}(h(\mathbf{x}^{'\bullet})^T - \mathbf{t}(\mathbf{x}^\bullet)(h(\mathbf{x}^{'\bullet})^T \mathbf{C}^{-1}\mathbf{H})^T] \end{aligned} \tag{4.14}$$

With c_2 given by

$$c_2(\mathbf{x}^\bullet, \mathbf{x}^{'\bullet}) = c(\mathbf{x}^\bullet, \mathbf{x}^{'\bullet}) - \mathbf{t}(\mathbf{x}^\bullet)^T \mathbf{C}^{-1}\mathbf{t}(\mathbf{x}^{'\bullet}) \tag{4.15}$$

From Eq. (4.13) to Eq. (4.15), $\mathbf{C} \in \mathbb{R}^{n \times n}$ is such that $\mathbf{C}_{ij} = c(\mathbf{x}_i, \mathbf{x}_j; \hat{\varphi})$. $\mathbf{t}(\cdot) \in \mathbb{R}^n$ such that $\mathbf{t}(\cdot) = (c(\cdot, \mathbf{x}_1; \hat{\varphi}), \ldots, c(\cdot, \mathbf{x}_n; \hat{\varphi}))^T$.

The GPE provides both a metamodel for the function and plausible uncertainty bounds. More details on the general procedure of constructing the GPE under the assumption that outputs are noise-free observed from an underlying functional mapping $\eta(\mathbf{x})$ can be found in [99, 100] from a Bayesian standpoint.

Figure 4.1 shows five sample GP functions draw from the GP prior in Eq. (4.2), which allows all properties of Gaussian distributions to be used. Figure 4.2 gives the corresponding five sample GP functions draw from GP posterior and the GPE also provides both a surrogate for the functions and plausible uncertainty bounds after combining five noise-free observations ($\mathbf{x} = [-4, -3, -2, -1, 1]^T$, $\mathbf{y} = [0.7568, -0.1411, -0.9093, -0.8415, 0.8415]^T$).
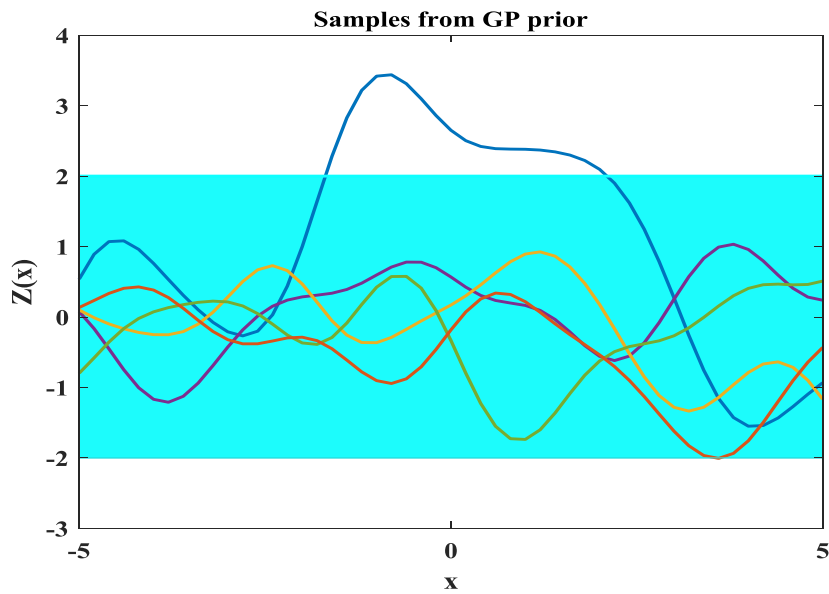
Figure 4.1 *Five sample GP functions draw from* Z(**x**)*. The light blue are indicates the 95% credible interval.*
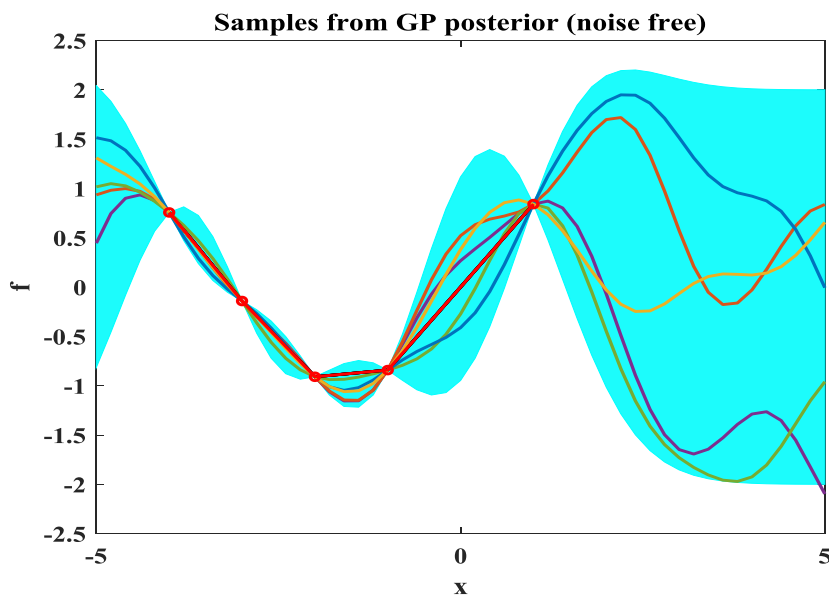
Figure 4.2 *Five sample GP functions draw from GP posterior conditioned on five observations(noise free). The light blue area indicates the 95% credible interval.*

An important assumption in the engineering literature is that the GPE is formulated on the premise that the true engineering function to be approximate is smooth and continuous. However, function evaluations are always scattered about this smooth trend due to errors in the physical experiments or the computer simulations used to calculate the functions.

If the outputs are noise polluted observed from an underlying functional mapping $\eta(\mathbf{x})$, a noise term σ^2_{noise} can be added. As a consequence, the GPE is no longer an interpolator but a regressor. In order to consider the influences of the noise, a regression constant σ^2_{noise} can be added to the leading diagonal of $\mathbf{C}$ [101]. That is, the $\mathbf{C}$ in Eq. (4.6) becomes $\mathbf{C} + \sigma^2_{noise}\mathbf{I}$ ($\mathbf{I}$ is an $n \times n$ identity matrix). This way, $\mathbf{t}(\cdot)$ is never a column of $\mathbf{C}$ (if $\mathbf{x}^* = \mathbf{x}$, the posterior predictive variance for the observed input values $\mathbf{x}$ is not zero anymore, which means the datasets are not interpolated). Using the same method of derivation as for interpolating GPE as described above, the regression GPE posterior predictive distributions for outputs conditional on the noised observed datasets expressed as follows

$$\eta_r(\mathbf{x}^* \mid D, \hat{\boldsymbol{\beta}}, \sigma^2) \sim N(m_r(\mathbf{x}^*), c_r(\mathbf{x}^*, \mathbf{x}^*)) \tag{4.16}$$

with posterior predictive mean

$$m_r(\mathbf{x}^*) = h(\mathbf{x}^*)\hat{\boldsymbol{\beta}} + \mathbf{t}(\mathbf{x}^*)^T (\mathbf{C} + \sigma^2_{noise}\mathbf{I})^{-1}(\mathbf{y} - \mathbf{H}\hat{\boldsymbol{\beta}}) \tag{4.17}$$

and posterior predictive covariance expressed as

$$\begin{aligned} c_{r1}(\mathbf{x}^*, \mathbf{x}'^*) = \hat{\sigma}^2 [c_{r2}(\mathbf{x}^*, \mathbf{x}'^*) + (h(\mathbf{x}^*)^T - \mathbf{t}(\mathbf{x}^*)^T (\mathbf{C} + \sigma^2_{noise}\mathbf{I})^{-1}\mathbf{H}) \\ \times (\mathbf{H}^T (\mathbf{C} + \sigma^2_{noise}\mathbf{I})^{-1}\mathbf{H})^{-1}(h(\mathbf{x}'^*)^T - \mathbf{t}(\mathbf{x}^*)(h(\mathbf{x}^*)^T (\mathbf{C} + \sigma^2_{noise}\mathbf{I})^{-1}\mathbf{H})^T] \end{aligned} \tag{4.18}$$

With c_{r2} given by

$$c_2(\mathbf{x}^*, \mathbf{x}'^*) = c(\mathbf{x}^*, \mathbf{x}'^*) - \mathbf{t}(\mathbf{x}^*)^T (\mathbf{C} + \sigma^2_{noise}\mathbf{I})^{-1}\mathbf{t}(\mathbf{x}'^*) \tag{4.19}$$

In Figure 4.3, through five noisy observations ($\mathbf{x} = [-4, -3, -2, -1, 1]^T$, $\mathbf{f} = [0.7568, -0.1411, -0.9093, -0.8415, 0.8415]^T$, $y = f + \varepsilon, \varepsilon \sim N(0, 0.05^2)$), which gives the corresponding 5 sample GP functions drawn from GP posterior and the GPE also provides both a surrogate for the functions and plausible uncertainty bounds. It is worth noting that the observations are not interpolated anymore.

Samples from GP posterior (noised)
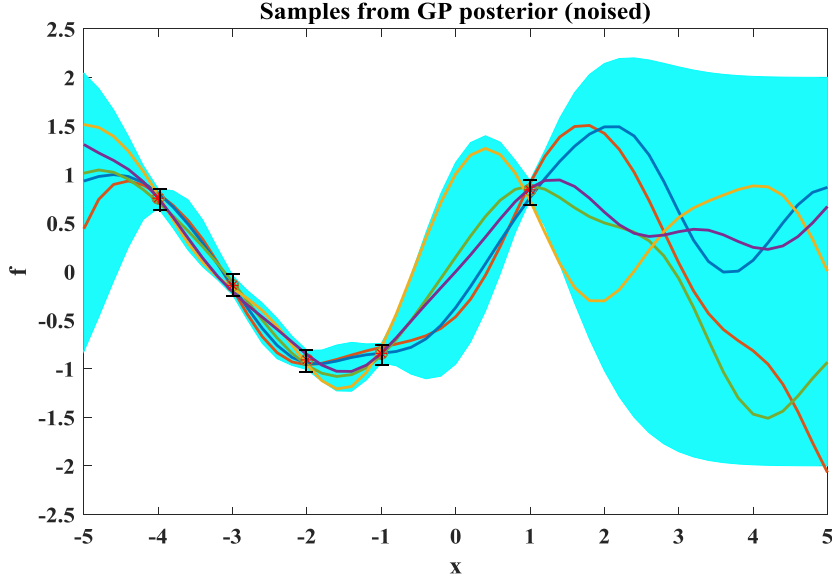
Figure 4.3 *Five sample GP functions draw from GP posterior conditioned on five observations(noised). The light blue area indicates the 95% credible interval.*

4.3.2 Covariance parameter estimation

In the procedure of specifying a particular GP prior in Eq. (4.2), the important quantity is the covariance matrix $\mathbf{C}$ with $\mathbf{C}_{ij} = c(\mathbf{x}_i, \mathbf{x}_j; \hat{\varphi})$. The particular choice of covariance function determines the properties of the sample functions drawn from the GP prior (e.g. smoothness, amplitude, etc.). Therefore, selecting an appropriate covariance function in GPE for particular uncertainty quantification problems is important. In this chapter, two well known and widely used stationary covariances will be used and discussed.

I. The squared exponential (SE) covariance

The SE covariance function is the most used function in various fields (e.g. geostatistics, machine learning, etc.). It gives very smooth sample functions, which are infinitely differentiable

$$c_{SE}(\mathbf{x}, \mathbf{x}'; \varphi) = u^2 \exp(-\frac{r^2}{2\tau^2}); \ \varphi = \{u, \tau\} \tag{4.20}$$

where $r = \|\mathbf{x} - \mathbf{x}'\|$. The two hyperparameters u and τ govern the properties of sample functions. The term u controls the typical amplitude and τ controls the lengthscale of variation.

$$x \in [-10, 10]$$
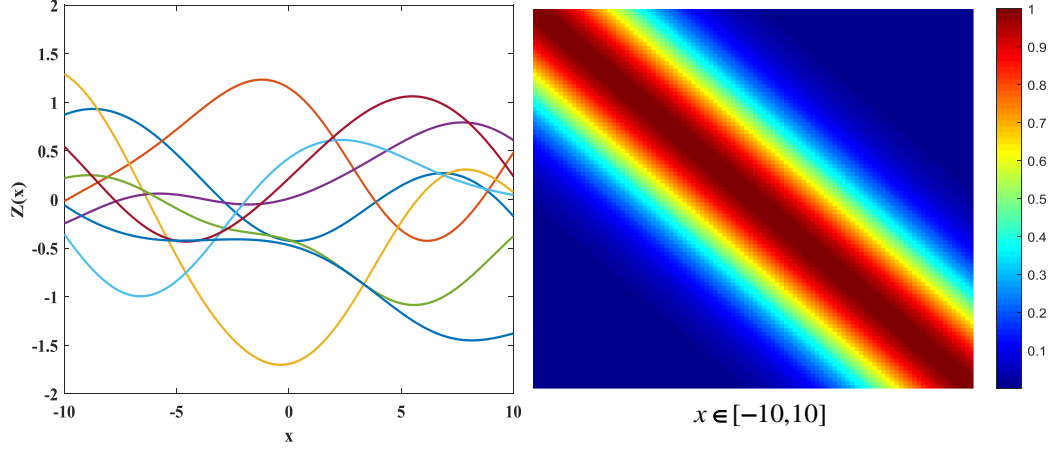
Figure 4.4 *Eight sample GP functions with SE covariance (left) and the SE prior covariance matrix for equally spaced ordered points (right).*

The left plot in Figure 4.4 shows the 8 sample GP functions with SE covariance with the hyperparameter values $u = 1$ and $\tau = 5$. It can be seen that the smoothness of the sample functions arises from the form of Eq. (4.20). Function variables close in input space are highly correlated, whereas far apart relative to the lengthscale τ are uncorrelated.

The structure of SE covariance can also be understood visually by plotting the values in the covariance matrix as colors shown on the right in Figure 4.4. The region of high covariance appears as a diagonal constant width band, which reflects the local stationary nature of the SE covariance. In addition, increasing the lengthscale τ increase the width of the diagonal band and points further away from each other become correlated.

A criticism about using SE covariance functions is that it may be unreasonable smooth for some realistic regression tasks due to its infinitely differentiable property. For this reason, many researchers prefer using Matérn class of covariance functions.

II. The Matérn class of covariance

The Matérn class of covariance functions can be expressed as follows

$$c_{Mat}(\mathbf{x}, \mathbf{x}'; \varphi) = u^2 \frac{2^{1-\nu}}{\Gamma(\nu)} (\sqrt{2\nu} \frac{r}{\tau})^{\nu} K_{\nu}(\sqrt{2\nu} \frac{r}{\tau}); \ \varphi = \{u, \tau\} \qquad (4.21)$$

where $r = \|\mathbf{x} - \mathbf{x}'\|$. ν and τ are positive parameters, Γ is gamma function, K_{ν} is the modified Bessel function of the second kind. In addition, the τ controls

the degree of smoothness .

In fact, for $v = \infty$, we regain the SE covariance function in Eq. (4.20). Especially, when $v = \dfrac{1}{2}$, we have the Matérn class of covariance functions expressed as

$$\lim_{v=1/2} \mathrm{C}_{Mat}(\mathbf{x}, \mathbf{x}'; \varphi) = u^2 \exp(-\frac{r}{\tau}) \qquad (4.22)$$

And, $v = \dfrac{3}{2}$ we have the Matérn class of covariance functions expressed as

$$\lim_{v=3/2} \mathrm{C}_{Mat}(\mathbf{x}, \mathbf{x}'; \varphi) = u^2 (1 + \frac{\sqrt{3}r}{\tau}) \exp(-\frac{\sqrt{3}r}{\tau}) \qquad (4.23)$$

The Matérn covariance functions are isotropic and again hyperparameters u and τ control the amplitude and lengthscale, respectively. v controls how smooth the sample functions are.

The special case of $v = \dfrac{1}{2}$ and $v = \dfrac{3}{2}$ for Matérn prior covariance function shown in Figure 4.5 and Figure 4.6 using the the same values in amplitude and lengthscale of SE covariance in Figure 4.4. By comparing theses figures, we can see how much rougher functions the Matérn covariance functions can produce.
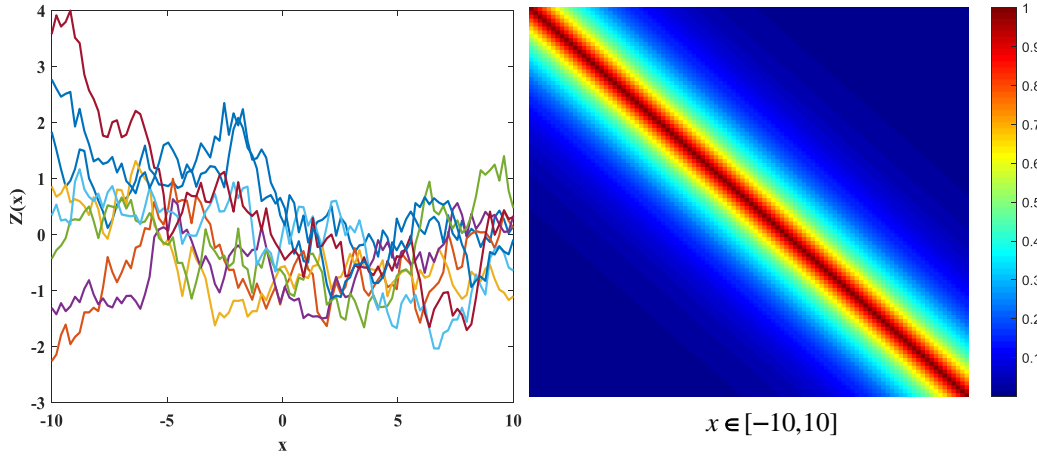
Figure 4.5 *Eight sample GP functions with Matérn* $\left(v = \frac{1}{2}\right)$ *covariance (left) and the Matérn* $\left(v = \frac{1}{2}\right)$ *prior covariance matrix for equally spaced ordered points (right).*
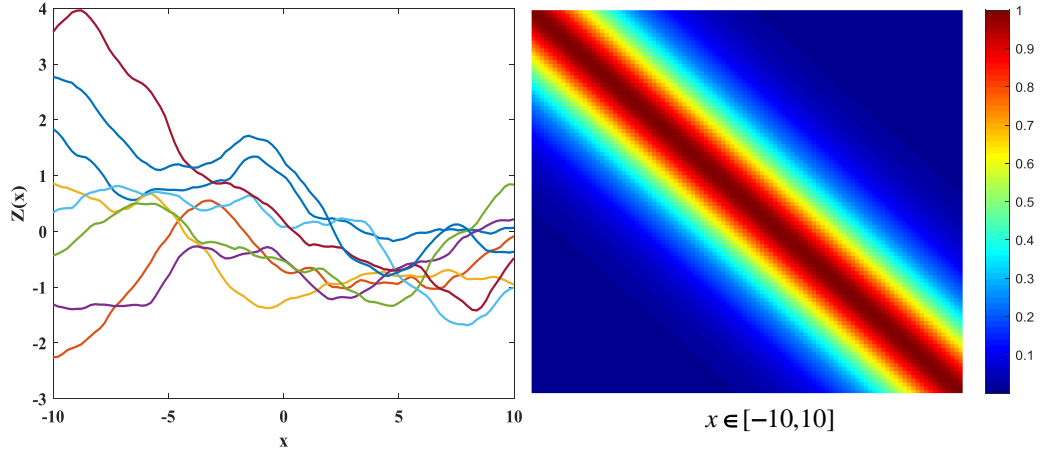
Figure 4.6 *Eight sample GP functions with Matérn* $\left(v = \frac{3}{2}\right)$ *covariance (left) and the Matérn* $\left(v = \frac{3}{2}\right)$ *prior covariance matrix for equally spaced ordered points (right).*

Due to the fact that hyperparameters φ which play an important role in the predictive capacity of the GPE, their accurate estimation is very important. In general, the optimal value φ is taken as the one that maximizes Eq. (4.7), the natural logarithm of the likelihood is used to improve the stability [102], which results in the log-likelihood function

$$\mathrm{In}\left[f(\mathbf{y}\,|\,\boldsymbol{\beta},\sigma^2)\right] = -\frac{n}{2}\mathrm{In}(2\pi) - \frac{n}{2}\mathrm{In}(\sigma^2) - \frac{n}{2}\mathrm{In}(|\,\mathbf{C}\,|)$$
$$-\frac{(\mathbf{y}-\mathbf{H}\boldsymbol{\beta})^T\mathbf{C}^{-1}(\mathbf{y}-\mathbf{H}\boldsymbol{\beta})}{2\sigma^2} \tag{4.24}$$

By substituting Eq. (4.10) and Eq. (4.11) into Eq. (4.24) yields the concentrated log-likelihood function in which constant terms are already removed

$$\mathrm{In}\left[f(\mathbf{y}\,|\,\boldsymbol{\beta},\sigma^2)\right] \approx -\frac{n}{2}\mathrm{In}(\hat{\sigma}^2) - \frac{n}{2}\mathrm{In}(|\,\mathbf{C}\,|) \tag{4.25}$$

Eq. (4.25) is used in computational implementations of MLE to increase the efficiency of the procedure [103]. In order to maximize the log-likelihood function in Eq. (4.25), some direct search algorithms can be employed to search the optimal values of φ such as Nelder-Mead simplex method [104], simulated annealing algorithm [105], genetic algorithm (GA) [106] and among others. The GA, which is an algorithm commonly applied in complex optimization problems, is employed in this implementation.

4.3.3 Surrogate model training

Design and analysis of computer experiments [107] is about exploring optimal strategies in selecting a combination of predictive values, to determine the variation in the response of a dependent variable. Since the cost of running a simulator is likely to restrict the number of available evaluations, n, the training samples should be chosen such that it explores the input space uniformly to make the training design to be space-filling in all dimensions. More details on the use of space-filling sampling plan can be found in [108].

Latin hypercube sampling (LHS) [109] is a statistical method for generating a near-ranodm samples of parameter values from a multidimensional distribution, This method is applied in the thesis to efficiently explore the input space due to its space-filling property. This is because a good space-filling LHS produces samples, which are evenly spread on every dimension to make the corresponding projections approximately follow uniform distributions.
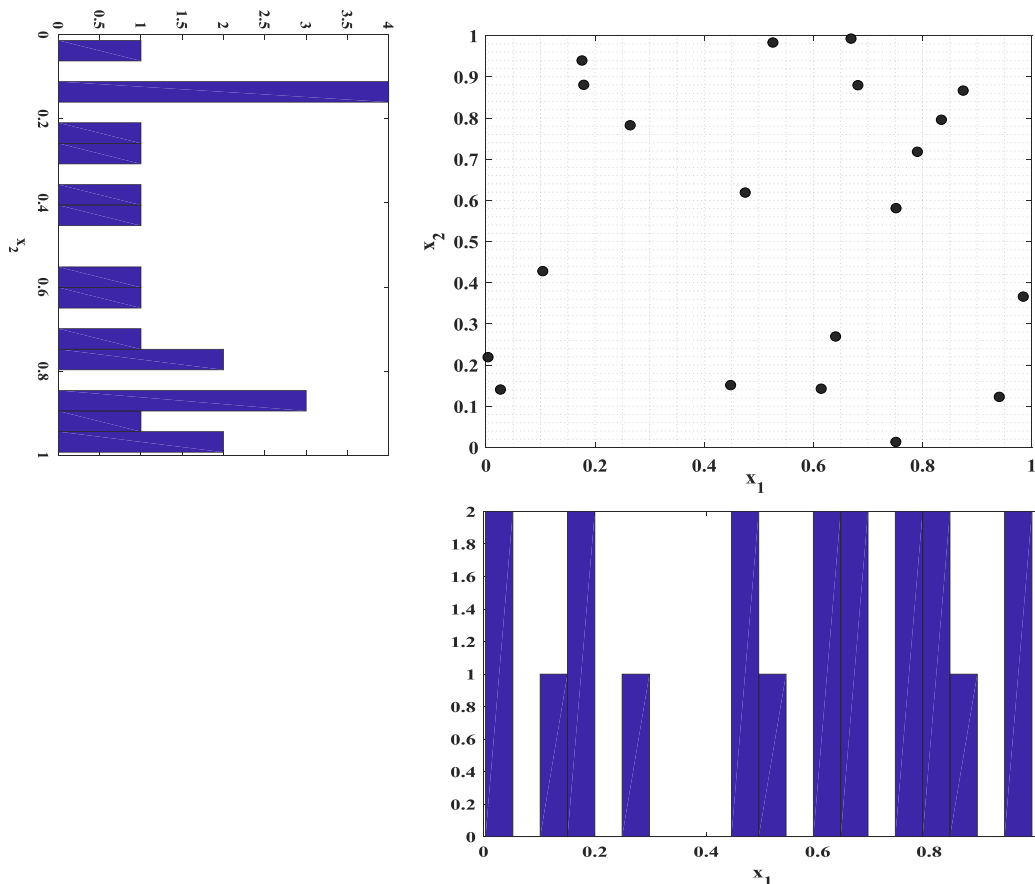
Figure 4.7 *The uniform random sampling technique with the number of data points* n = 20.
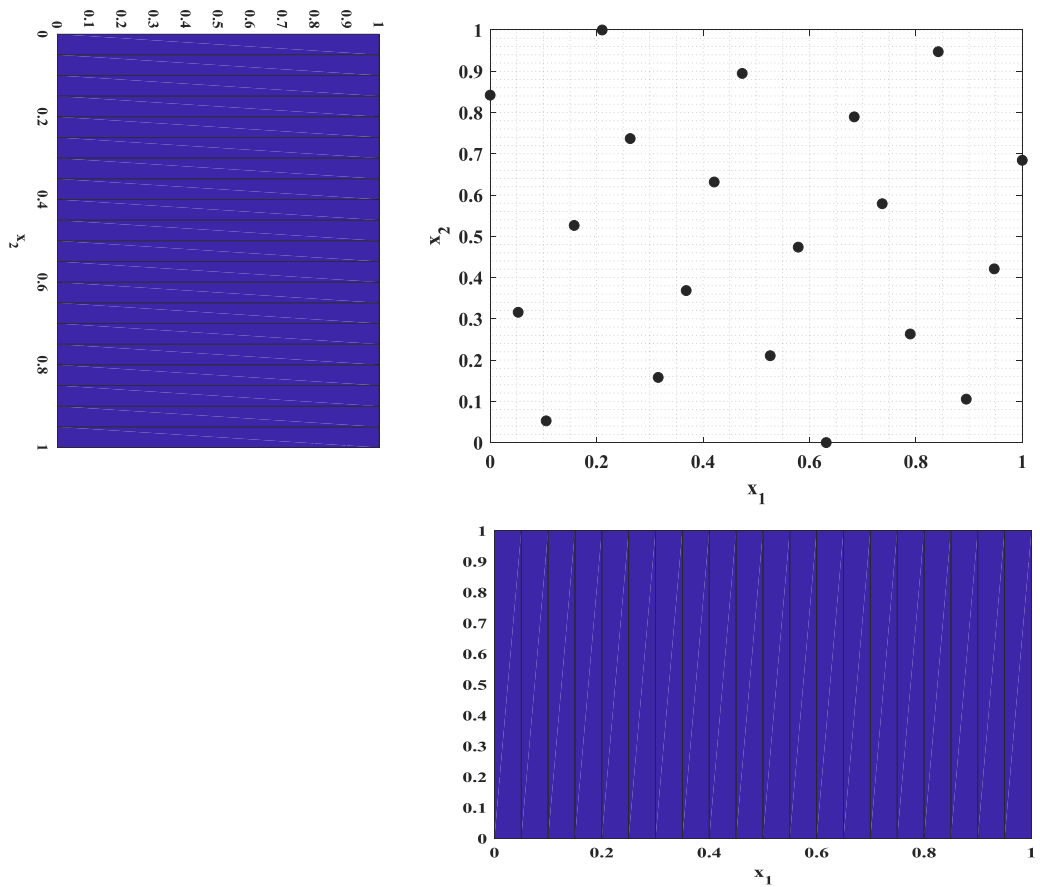
Figure 4.8 *The LHS technique with the number of data points* n = 20.

Figure 4.7 and Figure 4.8 show the uniform random sampling and the LHS strategies in terms of investigating individual variable ranges and samples distributions, which demonstrates the ability of the LHS to distribute the samples evenly along each dimension.

Choosing the size of the training samples is another aspect of GPE training. Loeppky et al. [110] provided a criterion for the number of training samples, n is related to the dimensionality of the problem, d as $n = 10d$. In the thesis, the training sample size is selected not only considering this kind of rule of thumb but also depending on the complexity of model construction.

4.3.4 Diagnostics

Due to some assumptions and additional uncertainties (e.g. inappropriate choice of estimation of parameters, inappropriate choice of correlation functions, etc.), the surrogate models might fail to construct an approximation to the output of the

simulator. Therefore, an emulator needs to be tested and validated.

There exist some methods to provide complete diagnostics to help with assessing the quality of the GPE predictions [111]. A diagnostic used in the thesis, which utilizes complete information provided by the GPE is the set of individual prediction errors (IPE) expressed as

$$I_i = \frac{y_i - \mathbb{E}\left[\eta(\mathbf{x}_i^*) \mid \mathbf{y}\right]}{\sqrt{\mathbb{V}\left[\eta(\mathbf{x}_i^*) \mid \mathbf{y}\right]}} \tag{4.26}$$

where $\mathbb{E}\left[\eta(\mathbf{x}_i^*) \mid \mathbf{y}\right]$ and $\mathbb{V}\left[\eta(\mathbf{x}_i^*) \mid \mathbf{y}\right]$ are the expectation and variance of the posterior predictive distributions of the surrogate model, respectively. For a reliable GP emulator, the IPE [111] should follow a Student-t distribution and 95% of validated points located in the interval $\left[-2, 2\right]$. Using Eq. (4.26), the IPE distribution can be visualized and studied to detect problems with emulators and ensure that, the quality of the GPE is acceptable.

In the following section, we are going to present an application of GPEs to a case of engineering reference. That is using GPEs to model the dynamic behaviors of steel skeleton reinforced concrete materials in the Split Hopkinson Pressure Bar test.

4.4 Analysis of steel skeleton reinforced concrete materials

4.4.1 Background description

In recent decades, increasingly severe challenges have arisen for modern concrete materials due to the construction of large-scale and complicated infrastructures, extreme service environments and multi-factor coupling actions [112, 113]. These facilities not only have to bear normal design loads (e.g. quasi-static loads, creep loads and among many others) but also may be exposed to short strong dynamic loads such as explosions and impacts under some extreme conditions. Therefore, investigations of the dynamic mechanical behavior of concrete are important and essential to provide valuable information about the structural designs.

It is well known that dynamic concrete material properties are different from their quasi-static conditions. In the early stage of research for concrete, the effect of strain rate on its compressive strength is at first drawn researchers' attention. Abrams [114] conducted concrete compression experiments at a strain rate $\dot{\varepsilon} = 2 \times 10^{-4} / s$ and $\dot{\varepsilon} = 8 \times 10^{-6} / s$. The results suggested the compressive strength of concrete has strain rate sensitivity property. Bresler and Bertero [115] used a hydraulic testing system to load the concrete specimens, the strain rate reached at $\dot{\varepsilon} = 1 / s$. Hughes and Gregory [116] used drop weight impactor and its strain rate reached at $\dot{\varepsilon} = 10 / s$. Watstein [117] conducted drop weight impactor tests to the cylindrical concrete specimen and found that in the range of strain rate at $\dot{\varepsilon} = 10 / s$, the dynamic compressive strength went up about 70%. Malvern, et al. [118] conducted $\varnothing = 76mm$ SHPB tests to concrete specimens and results suggest the dynamic compressive strength at a strain rate $\dot{\varepsilon} = 118 / s$, which is 2.2 times larger to the value of static strength. A classical review of concrete material on dynamic behavior under high stain rate through various testing methods is given by Bischoff and Perry [119]. After that, the research into putting steel fibers as dispersed reinforcement for concrete (FRC) became popular since Roumaldi and Batson [120] first proposed the idea of combining steel fibers and concrete together. A classical review of the strain rate effects on FRC material properties by impact and impulsive loading is given by [121].

Most of the research on dynamic materials properties of FRC is based on experimental tests. The testing instruments include drop weight impactor [122],

Split Hopkinson Pressure Bar (SHPB) [123] and detonation of explosives. FRC was found to be more strain rate sensitive than concrete and made with higher fibers content supposed to show larger strain rate sensitivity. Brandt [124] studied many types of fibers including high modulus steel, glass, carbon and asbestos of different shapes and in general, through fibers addition, the compressive strength was increased. Xu, et al. [125] conducted drop weight impactor tests on concrete specimens with seven different types of fibers and conclude that the FRC with spiral-shaped steel fibers is performed better than any other six fiber types in terms of ultimate compressive strength, post failure strength and energy absorption capacity. The energy absorption ability of concrete is also affected by steel fiber volume fraction as the strain rate is higher than $100 / s$. The experimental results of Sun et al. [126] indicated that the strain-rate strengthening effect of concrete is reduced with the increase of steel fiber volume fraction. Hou, et al. [127] and Li, et al. [128] investigated the dynamic compressive behavior of reactive powder concrete (RPC) with 2% and 5% steel fiber at the strain rate range from $75 / s$ to $274 / s$. It was concluded that the dynamic elastic modulus of concrete is sensitive to both steel fiber volume fraction and strain rate. Su, et al. [129] observed that the dynamic behavior of ultra-high-performance concrete containing 3% nanomaterial by weight has a strong dependency on the steel fiber. Soufeiani, et al. [130] summarized previous researches about the effect of steel fiber shape and volume fraction on the dynamic behavior of concrete.

Despite the general literature review above, however, the research of steel skeleton reinforced concrete (SSRC), which is also considered as a very important engineering composite material, under impact loadings, is limited. In this section, a series of experimental tests are therefore processed to investigate mechanical properties of SSRC materials. Quasi-static and dynamic compression tests are carried out by using WWJ-2008B material testing machine and split Hopkinson pressure bar (SHPB) system, respectively.

Three different longitudinal steel arrangements are considered: 0, 3×3, 3×4, which have a volume fraction of steel 0%, 1.91% and 2.94% named Concrete, SSRC33, and SSRC34, respectively. For the dynamic compression tests, the SHPB tests achieve strain rate from $30 / s$ to $100 / s$. The stress/strain equilibrium is examined. In addition, the strain rate sensitivity on the compressive strength and

Young's modulus is discussed. The dynamic increase factor (DIF) for compressive strength and Young's modulus relations for SSRC specimens are constructed probabilistically through Gaussian process emulation (GPE) in Bayesian framework.

4.4.2 Experimental program

All specimens are designed to have a dimension of $\varnothing = 75 \times 35mm$ (L/D is 0.5) which is proposed by Bertholf and Karnes [131] to eliminate the effects of the axial inertia in high-speed impact tests. Moreover, in all tests, grease is spread at both sides of specimens in order to minimize the end friction confinement due to the interaction between the specimen and instruments. For the arrangement of steel skeletons, the longitudinal steel is all $25mm$ in length and arranged in a stratified array, the transverse steel adopts two layers of the orthogonal arrangement, the space between the layers is $15mm$, the diameter of steel is $2mm$, which is shown in Figure 4.9.
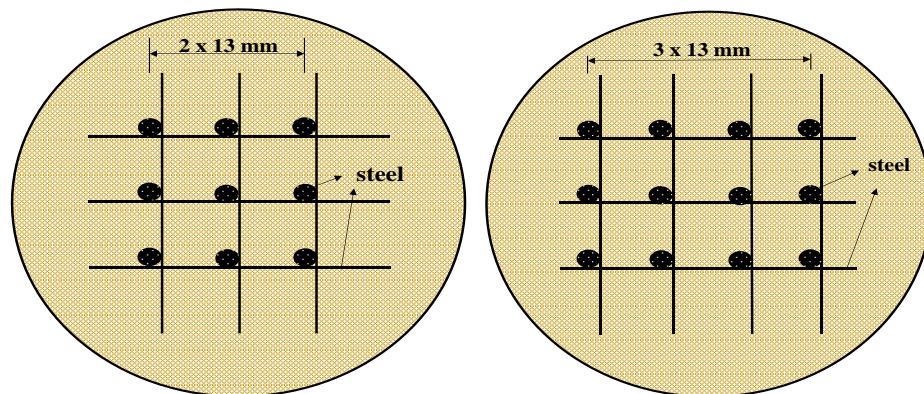
Figure 4.9 *Steel skeleton arrangement with SSRC33 (left) and SSRC34(right).*

The cement used to prepare the specimens is the ordinary Portland cement, the maximum coarse aggregate is $10mm$, the detailed mixtures for concrete are given in Table 4.1. All specimens are designed to have a dimension of $\varnothing = 75 \times 35mm$ (L/D is 0.5) and mixed into steel molds and ensure all the steel skeletons are not exposed to air. The specimens are being carefully cured for 24 hours after demolded and then placed into normal temperature water for 28 days curing. After 28 days, both end surfaces of specimens were smoothed, dried and tested. A total number of 60 specimens were prepared for quasi-state and SHPB tests,

with 20 specimens for Concrete, SSRC33, SSRC34, respectively. The prepared specimens and a typical steel skeleton of 3×3 are shown in Figure 4.10. The relations of longitudinal steel array and volume fraction of steel are given in Table 4.2.

Figure 4.10 *Steel skeleton (left) and the specimens (right).*

Table 4.1 *Mix proportions of specimens*

Mixes (%)	W/C	V_f (%)	Mix proportions (kg/m^3)					
			Water	Cement	10mm	7mm	5mm	sand
Concrete	0.36	0	200	556.2	567.1	425.25	141.95	509.6
SSRC33	0.36	1.91	200	556.2	567.1	425.25	141.95	509.6
SSRC34	0.36	2.94	200	556.2	567.1	425.25	141.95	509.6

W/C: Water to cement ratio.

V_f : Volume fraction of steel.

Table 4.2 *Longitudinal steel ratio*

Longitudinal steel array	V_f (%)
Concrete	0
SSRC33	1.91
SSRC34	2.94

V_f : Volume fraction of steel.

The quasi-static compression tests were conducted by using WWJ-2000B material test machine at national high-speed railway construction technology laboratory in the Central South University (CSU) in China in 2015 shown in Figure 4.11. The WWJ-2000B material test machine used a closed circuit electro-hydraulic system to make sure the load could be compensated quickly to guarantee the whole testing

process stable. The vertical loads and displacement versus time were recorded by the data processing system. In each test, the grease was spread on the interfaces between specimens and instruments to minimize the influence of friction.
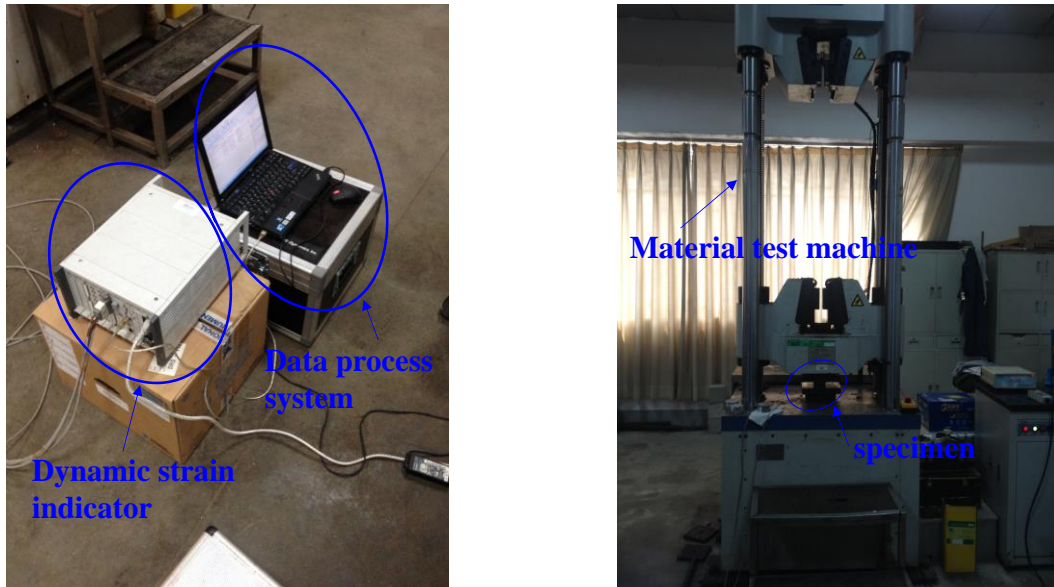
Figure 4.11 *Quasi-static compression experiment instruments.*

The dynamic impact compression tests were conducted by using the SHPB test system at the Hefei University of Technology in China in 2015. Its setup is shown in Figure 4.12.

The SHPB system was a variable cross-section type test system with $\varnothing = 74mm$. The striker bar had a diameter of $\varnothing = 37mm$ and length of $800mm$, the incident bar was a variable cross-section bar, the large side had a diameter of $\varnothing = 74mm$ and the small side had a diameter of $\varnothing = 37mm$. The transition length from small side to large side was $420mm$, the overall length of the incident bar was $3061mm$, the length of the transmission bar was $1797mm$ with a diameter of $\varnothing = 74mm$.
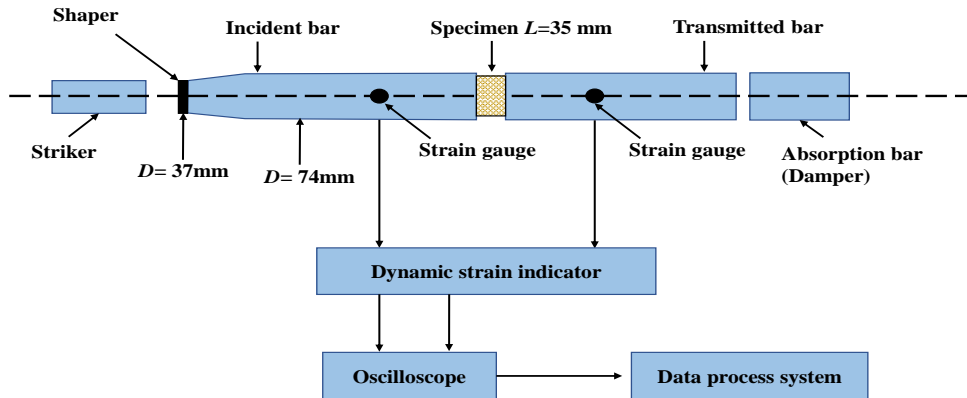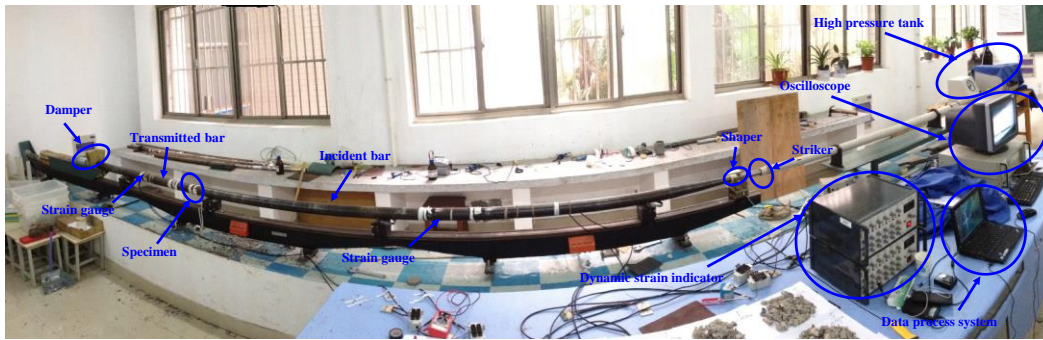
Figure 4.12 *SHPB dynamic compression experiment instruments.*

Before doing SHPB tests, a waveform shaper was used, which had the power to initiate a half-sine loading waveform which could eliminate violent oscillation and dispersion is illustrated in Figure 4.13. The strain gauges were placed at the middle surface of incident and transmission bars. The bars were made of stainless steel with density $7850kg/m^3$, Young's modulus $210GPa$ and Poison's ratio 0.25. The elastic wave velocity were $5172m/s$. In order to eliminate errors as much as possible, each specimen was carefully prepared by smoothing and grinding to guaranty both two sides were parallel. Moreover, the roughness on surfaces was controlled to be less than $0.02mm$. A typical signal from SHPB tests can be seen in Figure 4.13. In the tests, even with the help of waveform shaper, it is hard to find a relative stability region used as an average strain rate when the peak value in stress history is used. This is caused by different sources of uncertainty such as tested materials, geometrical properties et al [133]. Under these conditions, the value of strain rate was chosen by the peak stress (compressive strength).
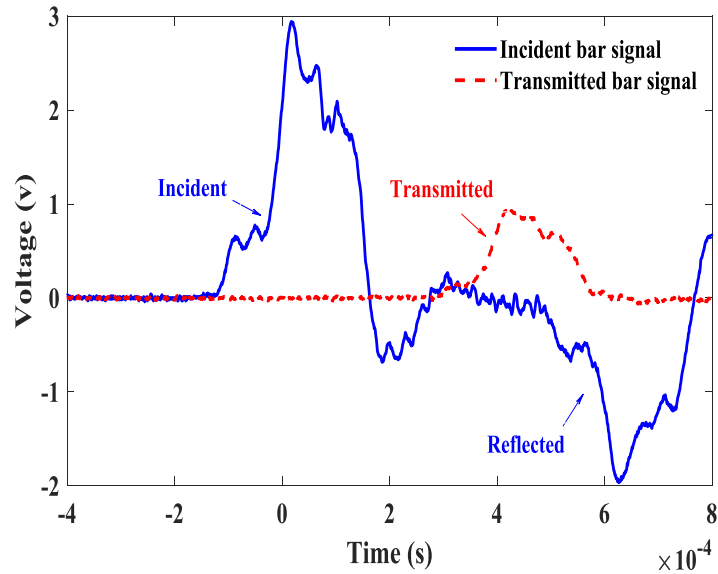
Figure 4.13 *The representative signals of SSRC34 from SHPB experiment measured at the strain gauge in Figure 4.12.*

4.4.3　Experiment results analysis

(1) Quasi-static compression tests

The typical results of Concrete, SSRC33, and SSRC34 under quasi-static loadings are given in Table 4.3. It is obvious that the compressive strength is improved when the steel skeletons are being inserted into the plane concrete specimens.

Table 4.3 *Material properties in Quasi-static condition*

	Dry density (kg/m^3)	Compression strength (MPa)	Young's modulus (GPa)
Concrete	2250	47.52	31.44
SSRC33	2279	54.68	32.35
SSRC34	2293	56.03	32.76

(2) SHPB compression tests

The SHPB compression tests were conducted for the specimens: Concrete, SSRC33, and SSRC34. The computation is based on one-dimensional stress wave propagation theory. Eq. (4.27) was used to calculate the stress, strain and strain ratio of specimens [133].

$$
\begin{cases}
\sigma_s(t) = \dfrac{EA}{A_s}\varepsilon_t(t) \\[2mm]
\varepsilon_s(t) = \dfrac{2C_0}{l_s}\int_0^t \big[\varepsilon_i(t) - \varepsilon_t(t)\big]dt \\[2mm]
\dot{\varepsilon}_s(t) = \dfrac{2C_0}{l_s}\big[\varepsilon_i(t) - \varepsilon_t(t)\big]
\end{cases}
\tag{4.27}
$$

where E, A are Young's modulus and cross-section area of pressure bars. A_s and l_s are the cross-section area and length of test specimens. $\varepsilon_i(t)$ and $\varepsilon_t(t)$ are the measured incident and transmitted strain data. C_0 is the elastic wave velocity of pressure bars.

During the SHPB tests, the longitudinal strain/stress equilibrium has to be achieved [134]. Eq. (4.28) is used to check the dynamic strain/stress equilibrium. In addition, Figure 4.14 gives an example of signals taken from SSRC33 for strain/stress equilibrium checkout. The time lags have already been moved for clear comparison. For each test, the strain/stress equilibrium status must be achieved which means the results of SHPB impact compression tests are proved to be validated.

$$
\varepsilon_i(t) + \varepsilon_r(t) = \varepsilon_t(t)
\tag{4.28}
$$
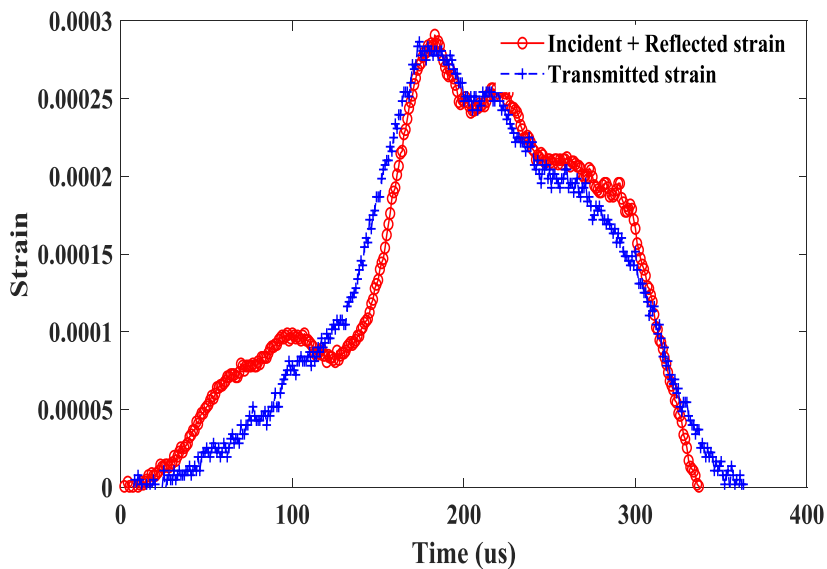
Figure 4.14 *Strain/stress equilibrium verification of SSRC33.*

Figure 4.15 shows the typical stress-strain curve of all specimens at $\dot{\varepsilon} = 35/s$. Generally, the maximum compressive stress is largely enhanced by the insertion of specific steel skeletons.
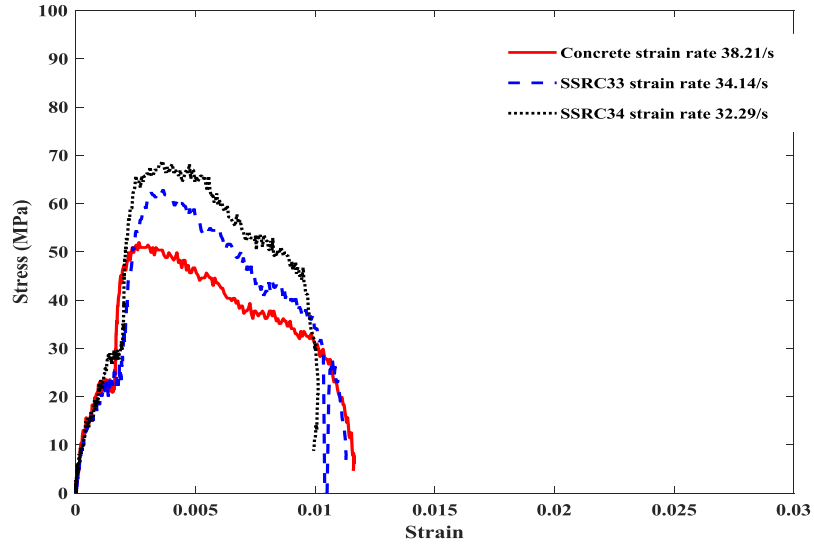
Figure 4.15 *Typical stress-strain curves at* $\dot{\varepsilon} = 35 / s$ *for Concrete, SSRC33 and SSRC34.*

4.4.4 Strain rate effects and GP emulation of DIFs on compressive strength and Young's modulus

The effect of strain rate on compressive strengths of all specimens is illustrated in Figure 4.16 The dynamic increase factor (DIF) for compressive strength is calculated by the dynamic compressive strength normalized by the average static compressive strength of the same type of specimen listed in Table 4.3.
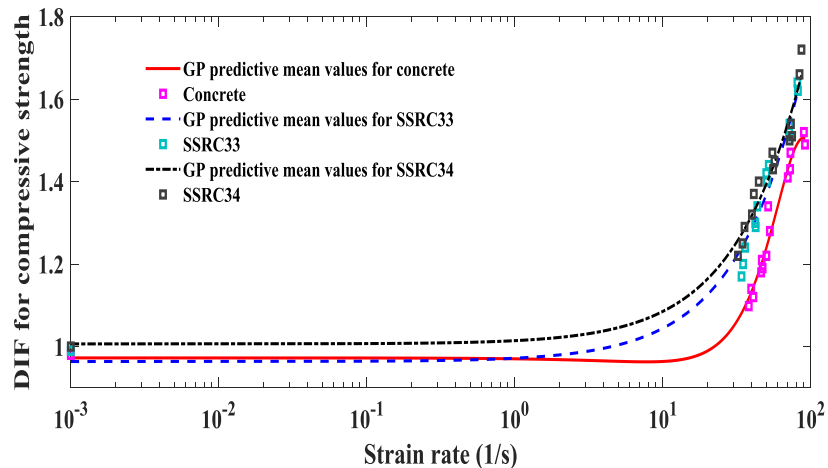
Figure 4.16 *DIFs for compression strength (Concrete-15 points (specimens), SSRC33- 15 points (specimens) and SSRC34-15 points (specimens) taken from Table 4.4-4.6).*

Moreover, the DIF results of all specimens for compressive strength and Young's modulus are listed from Table 4.4 to Table 4.6. In Figure 4.16, it is obvious that the compressive strength of all specimens is sensitive to strain rate. Compared with

the scatter values of DIF on Concrete specimens, the SSRC33, SSRC34 show a similar strain rate sensitivity trend and indicate more strain rate sensitive than the Concrete under SHPB tests.

Table 4.4 *DIFs summary for Concrete*

	Strain rate $(1/s)$	DIFs for compressive strength	DIFs for Young's modulus
1	1×10^{-3}	0.98	0.99
2	38.21	1.10	1.03
3	39.66	1.14	0.98
4	41.05	1.12	1.03
5	46.32	1.18	1.01
6	46.78	1.21	1.00
7	47.40	1.19	1.02
8	50.27	1.22	1.04
9	51.66	1.34	1.06
10	53.07	1.28	1.04
11	70.11	1.41	1.08
12	72.67	1.43	1.07
13	73.06	1.47	1.09
14	90.12	1.52	1.11
15	91.67	1.49	1.10

The effect of strain rate on Young's Modulus of all specimens is illustrated in Figure 4.17. The dynamic increase factor (DIF) for Young's Modulus is calculated by Young's Modulus under dynamic compression normalized by the average static Young's Modulus of the same type of specimen listed in Table 4.3. In Figure 4.17, the DIFs for the SSRC33 and SSRC34 are higher than Concrete and seem to follow a similar trend compared to Concrete.
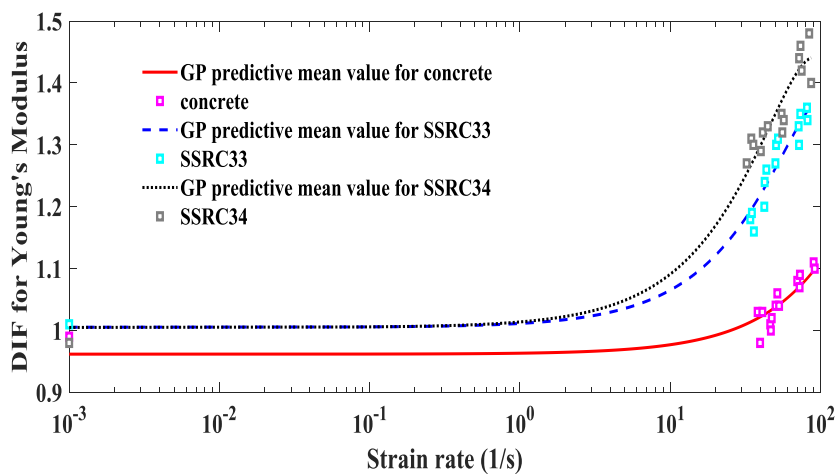
Figure 4.17 *DIFs for Young's modulus(Concrete-15 points (specimens), SSRC33-15 points (specimens) and SSRC34-15 points (specimens) taken from Table 4.4-4.6).*

Finding the best-fit DIF relations for the compressive strength and Young's modulus of SSRC specimens with different steel skeletons is important. Because these relations can be used to estimate DIFs of SSRC materials in numerical prediction of SSRC structure responses to other untested high-rate loadings. Researchers such as Xu Z, et al. [136], Y. Hao, et al. [137] used deterministically polynomial fitting techniques to do the curve fitting. However, these methodologies are limited in the sense that, being deterministic, are unable to account for the uncertainty present in the experiment and due to the fact that it is expensive to run such experiments.

Table 4.5 *DIFs summary for SSRC33*

	Strain rate $(1/s)$	DIFs for compressive strength	DIFs for Young's modulus
1	1×10^{-3}	0.99	1.01
2	34.14	1.17	1.18
3	34.99	1.20	1.19
4	36.07	1.24	1.16
5	42.31	1.30	1.20
6	42.54	1.29	1.24
7	43.66	1.34	1.26
8	50.18	1.42	1.27
9	50.66	1.40	1.30
10	52.17	1.44	1.31
11	71.62	1.52	1.33
12	72.03	1.54	1.30
13	73.44	1.51	1.35
14	81.41	1.64	1.36
15	82.07	1.62	1.34

Table 4.6 *DIFs summary for SSRC34*

	Strain rate $(1/s)$	DIFs for compressive strength	DIFs for Young's modulus
1	1×10^{-3}	1.00	0.98
2	32.99	1.22	1.27
3	34.66	1.25	1.31
4	35.77	1.29	1.30
5	40.12	1.32	1.29
6	41.33	1.37	1.32
7	44.62	1.40	1.33
8	55.16	1.47	1.35
9	55.77	1.43	1.32
10	57.03	1.45	1.34
11	72.11	1.50	1.44

12	73.66	1.54	1.46
13	75.03	1.51	1.42
14	84.19	1.66	1.48
15	86.75	1.72	1.49

GPE described in detail in section 4.3 is a modelling process based on statistical learning theory which has drawn a growing interest in solving uncertainty quantification problems, computationally intensive problems in various fields such as climate prediction [138], deterministic structural dynamics [139], stochastic structural dynamics [140], test crash modelling [141] and among many others.

Since the training process is done not by running a numerical model, but by observing experiments datasets, then a noise term has to be added to the covariance matrix $\mathbf{C}$ in Eq. (4.6), which described in detail in section 4.3, The GPE is then constructed based on the measurements to train with the SE covariance functions. In addition, in order to train the GP emulator, 10 measured points were selected across the input domain according to J. Loeppky, et al. [110] who justify that the number of measured training points, n , that is required related to the dimensionality of the problem d and a practical rule is that n should be about $n = 10d$.

The GPEs for DIF relations to the compressive strength and Young's modulus of the plane Concrete, SSRC33 and SSRC34 are constructed described in section 4.3. The GP predictive mean curves and corresponding predictive 95% confidence intervals of all specimens for compressive strength and Young's modulus are shown in Figure 4.18, Figure 4.20, and Figure 4.22.

For a reliable GP emulator, the IPE [111] should follow a Student-t distribution and 95% of validated points located in the interval $[-2, 2]$. The IPE is used to validate the GPE. Figure 4.19, Figure 4.21 and Figure 4.23 show IPEs of all specimens for compressive strength and Young's modulus with only 5 points to validate since making such experiments is expensive. It can be observed that the validated points lie in the desired region, which illustrates the well-trained GPE could be used to estimate DIFs of SSRC in numerical prediction of SSRC structure responses to other different impact loadings. It is also worth noting that obtaining more validation points means conducting more experiments, which is expensive.
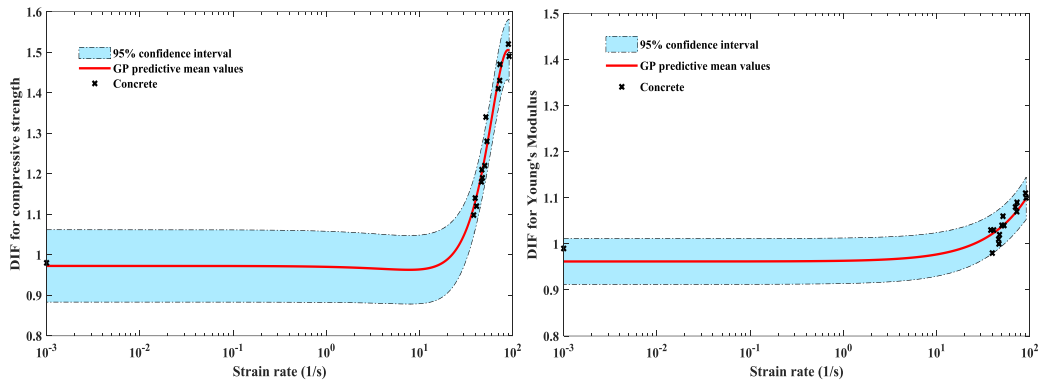
Figure 4.18 *GP emulation of DIFs for compressive strength (left) and Young's modulus (right) in Concrete (10 specimens used for training).*
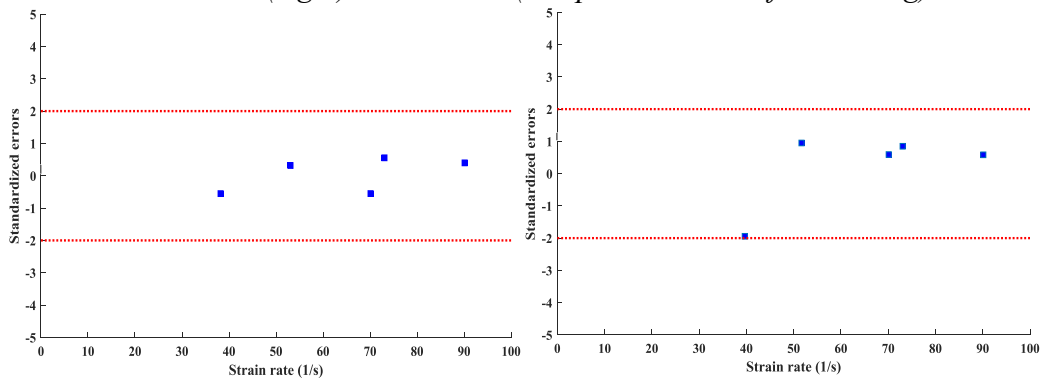
Figure 4.19 *IPE of Concrete for compression strength(left) and Young's modulus (right) (5 validation specimens).*
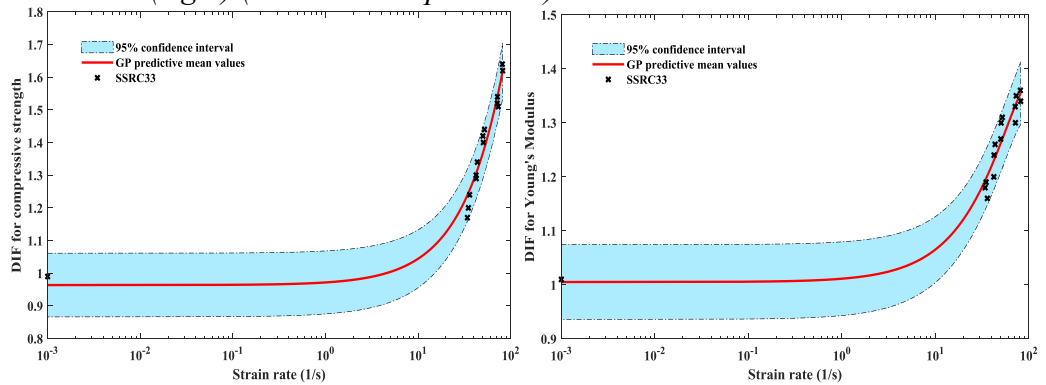
Figure 4.20 *GP emulation of DIFs for compressive strength (left) and Young's modulus (right) in SSRC33 (10 specimens used for training).*
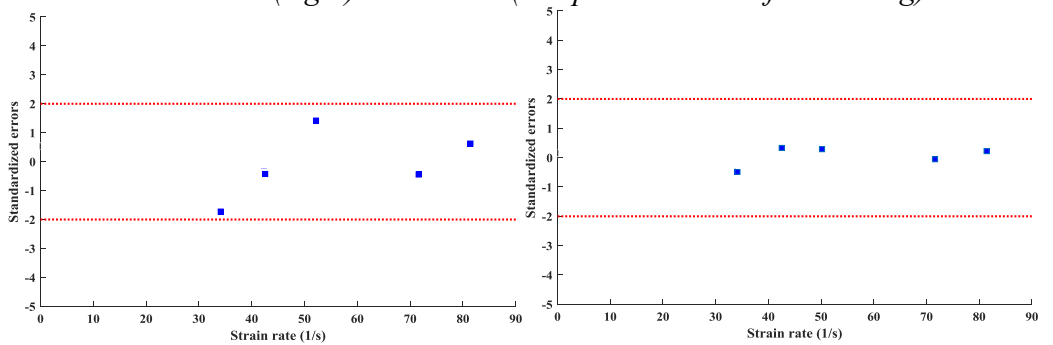
Figure 4.21 *IPE of SSRC33 for compression strength (left) and Young's modulus (right) (5 validation specimens).*

73
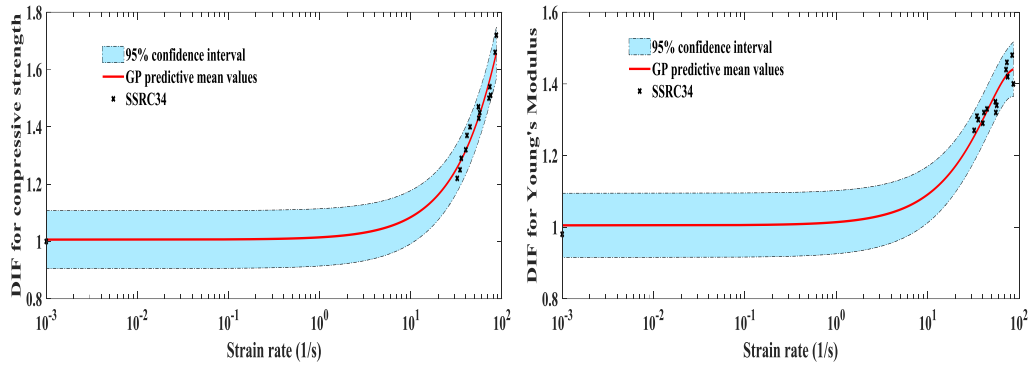
Figure 4.22 *GP emulation of DIFs for compressive strength (left) and Young's modulus (right) in SSRC34 (10 specimens used for training).*
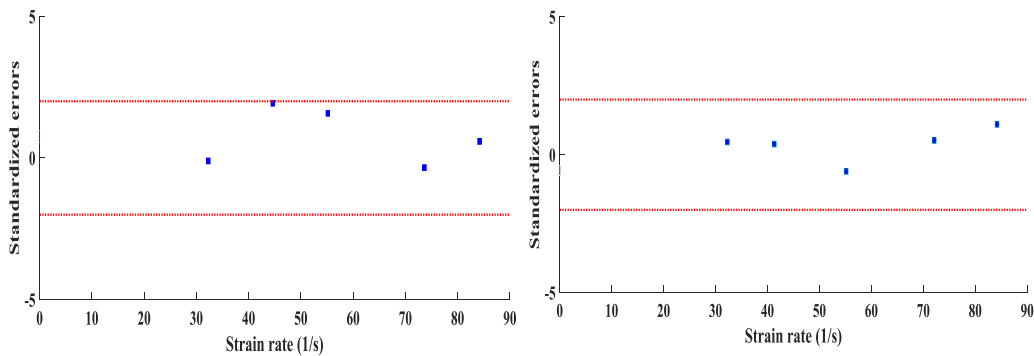
Figure 4.23 *IPE of SSRC34 for compression strength(left) and Young's modulus (right) (5 validation specimens).*

4.4.5   Discussion

Section4.4 presents some useful results from quasi-static and SHPB tests for three types of specimens: Concrete, SSRC33, and SSRC34. The plain Concrete specimens are regarded as control samples for comparison. It is observed that under quasi-static tests, the compressive strength is improved from Concrete to SSRC34. For SHPB tests, a typical stress-strain curves of all specimens at certain strain rate indicates that the maximum compressive stress are largely enhanced by the insertion of steel skeletons compared to the plain concrete. The GPE is employed to obtain the relations of DIFs for compressive strength and Young's modulus from different strain rates. The well-trained GPE is validated by individual prediction errors (IPE) diagnostics. In addition, DIFs of the compressive strength and Young's modulus show increasing rate sensitivity to strain rate from Concrete to SSRC34. More experimental tests are suggested to conduct in the future to have more validation points to validate the trained GPE and considering other kinds of arrangements of steel skeletons in order to have a more

comprehensive understanding of the dynamic mechanical behaviors of this composite material is also recommended.

4.5 Chapter summary

This chapter gives an introduction of surrogate modeling techniques (emulation). The theory behind GPE and the relevant notations are discussed in detail. The procedure of construct the GPE considering the noise-free observations and noised observations, respectively are given and discussed.

The GPEs introduced in this chapter, are the most important ingredient in the uncertainty quantification analysis performed in this thesis. The analysis of SSRC materials under quasi-static tests and SHPB tests is discussed. The GPEs are employed to obtain the relations of DIFs for compressive strength and Young's modulus from different strain rates under impact loadings. In the following chapter 5, the computational efficient Gaussian processes are built based on the standard GPE described in this chapter. In addition, the GPEs are also used in chapter 6 to carry out model updating using the modular Bayesian approach.

5 Sparse Gaussian process emulation

5.1 Computational efficiency Gaussian processes

It is a problem particularly apparent in the structural dynamics and engineering field where relatively large datasets are available. When processing the large datasets, applying the full GPE can be computationally challenging. This is due to the following. The full procedure of GPE construction consists of two stages which called training and prediction. For the hyperparameters' estimation using Eq. (4.24) in the training stage, each search computation requires the inversion of the covariance matrix $\left[\mathbf{C} + \sigma^2_{noise} \mathbf{I} \right]$ and hence cost will be scaled as $\mathcal{O}(N^3)$ where N is the number of data points. In addition, in the prediction stage for test points, the prediction cost will be scaled as $\mathcal{O}(N^2)$.

In fact, the reason for both the GPE training and prediction stage to be prohibitively expensive is due to the poor scaling when N is large. The poor scaling is caused by the nonparametric nature of the Gaussian process itself. In the past decades, some research was done that attempts to address the computational problem by developing efficient approximations to the full GPs. An overview of this is given in this section.

5.2 Sparse Gaussian process approaches

5.2.1 Subset of data

When facing big datasets, the most desirable way to reduce the complexity may be to design a method or criterion in order to discard the data we cannot afford to process. When data is redundant, most of the available information can be obtained in some selected subset.

In fact, if all the data points are equally informative, this kind of subset can be selected randomly, but there is usually an advantage to more sophisticated selection techniques. Several greedy selection criteria have been proposed, such as informative vector machine [142] and information Gain [143]. In both cases, samples added into the subset are chosen to maximize some measure of the

information held into it, e.g. a differential entropy score for subset information. The sample selection does not increase the overall computational complexity of sparse regression due to the relevant matrices are incrementally grown using low-rank updates.

5.2.2 The Nystrom method

Using the Nyström method for efficient Gaussian process regression was proposed by Williams and Seeger [144]. A few columns of the covariance matrix are used to approximately obtain its eigenvectors and eigenvalues, which are in turn used to reconstruct the full matrix. The low-rank approximation to the covariance matrix is expressed as

$$\mathbf{C}_{\eta\eta} \approx \mathbf{L}_{\eta\eta} = \mathbf{C}_{\eta m} \mathbf{C}_{mm}^{-1} \mathbf{C}_{m\eta}^{T} \tag{5.1}$$

where m are called inducing variables. For the Nyström method, m is a subset of η with size $M \ll N$. The set of inputs $\bar{\mathbf{X}} \subset \mathbf{X}$ corresponding to m is called the active set. The $\mathbf{C}_{\eta m}$ refers to the covariance matrix between the elements from η and m.

Replacing $\mathbf{C}_{\eta\eta}$ with low-rank matrix $\mathbf{L}_{\eta\eta}$ in the Eq. (4.17) and Eq. (4.18), makes the training be performed as scaled as $\mathcal{O}(M^2 N)$ and predictions for new test samples are scaled as $\mathcal{O}(N)$ for the mean and $\mathcal{O}(MN)$ for the variance. But it is worth noting that this is a numerical approximation with no probabilistic foundation and can produce meaningless results. The effective joint prior for training and test dataset is

$$p(\eta, \eta_* \mid \mathbf{x}, \mathbf{x}_*) = N(\begin{bmatrix} \eta \\ \eta_* \end{bmatrix} \mid 0, \begin{bmatrix} \mathbf{L}_{\eta\eta} & \mathbf{C}_{\eta*} \\ \mathbf{C}_{*\eta} & \mathbf{C}_{**} \end{bmatrix}) \tag{5.2}$$

where the exact prior covariance for the training dataset has been replaced with the Nyström method, whereas the remaining prior covariances are kept exact. Doing so yields an inconsistent joint probability distribution whose prior covariance is not even guaranteed to be positive definite. The approximate posterior cannot be regarded as GP and absurd results such as negative predictive variances can occur.

5.2.3 Sparse pseudo-input Gaussian processes

The sparse pseudo-input Gaussian processes (SPGP) model was introduced by

Snelson and Ghahramani [145]. One of the novelties of this model is the constraint that the samples of an active set, which are also called pseudo-inputs, must be selected among the training datasets is relaxed. This allows both pseudo-inputs and hyperparameters to be selected in a joint continuous optimization and increase flexibility.

The key idea of SPGP is to augment the existing training datasets with a noiseless pseudo datasets $\mathbf{D}_m = \{\mathbf{X}, m\}$ with pseudo inputs $\mathbf{X} = \left[\overline{\mathbf{x}}_1, \ldots, \overline{\mathbf{x}}_M \right]$ and pseudo outputs $m = \left[m_1, \ldots, m_M \right]$ ($M \ll N$) and assume that all the latent variables are conditionally independent given the pseudo datasets. By performing a gradient-based optimization, the covariance matrix of the full Gaussian process regression model is parameterized by the locations of M pseudo input points. This kind of sparse approximation is equivalent to a standard GP with covariance function expressed as

$$c_{SPGP} = c_{SPGP}(\mathbf{x}, \mathbf{x}') = \mathbf{t}_m(\mathbf{x})^T C_{mm}^{-1} \mathbf{t}_m(\mathbf{x}')(1 - \delta_{\mathbf{x}\mathbf{x}'}) + c(\mathbf{x}, \mathbf{x}')\delta_{\mathbf{x}\mathbf{x}'} \qquad (5.3)$$

where Kronecker delta $\delta_{\mathbf{x}\mathbf{x}'}$ equals one if $\mathbf{x} = \mathbf{x}'$ and zero otherwise.

The prediction equations can be expressed as

$$
\begin{aligned}
p_{SPGP}(\eta_* \mid x_*, \mathbf{D}) &= N(\eta_* \mid u_{SPGP}, \sigma_{SPGP}^2) \\
u_{SPGP} &= \mathbf{t}_*^T (\mathbf{C}_{mm} + \mathbf{C}_{\eta m}^T \Lambda_{\mathbf{y}}^{-1} \mathbf{C}_{\eta m})^{-1} \mathbf{C}_{\eta m}^T \Lambda_{\mathbf{y}}^{-1} \mathbf{y} \\
\sigma_{SPGP}^2 &= c_{**} + \mathbf{t}_*^T ((\mathbf{C}_{mm} + \mathbf{C}_{\eta m}^T \Lambda_{\mathbf{y}}^{-1} \mathbf{C}_{\eta m})^{-1} - \mathbf{C}_{mm}^{-1}) \mathbf{t}_*
\end{aligned}
\qquad (5.4)
$$

where $\Lambda_{\mathbf{y}} = \mathrm{diag}(\mathbf{C}_{\eta\eta} - \mathbf{C}_{\eta m}^T \Lambda_{\mathbf{y}}^{-1} \mathbf{C}_{\eta m}) + \sigma^2 \mathbf{I}$.

The computational complexity for the training stage is scaled as $\mathcal{O}(NM^2)$ and prediction stage scaled as $\mathcal{O}(M^2)$ for SPGP. This approach turns the inducing inputs into additional hyperparameters, which can increase flexibility when we fit the data, but it can also lead to overfitting since the approximate marginal likelihood is not a lower bound when we optimize with respect to all unknown hyperparameters.

5.3 Variational sparse Gaussian process emulation

The variational sparse Gaussian process (VSGP) model is used and implemented in the thesis, which is built on the notion of VSGP method proposed by Titsias [146]. This method is also closely related to the techniques described in [147, 148].

The key property of this method is that the inducing inputs are defined to be variational parameters which are selected by minimizing the Kullback-Leibler (KL) divergence, as detailed in Appendix B, between the variational distribution and the exact posterior distribution over the latent function values.

Defining $\mathbf{f}_m$ as the vector of true function values at the sparse points in $\mathbf{X}_m$, a prior $p(\mathbf{f},\mathbf{f}_m)$ and likelihood $p(\mathbf{y}|\mathbf{f},\mathbf{f}_m)$ gives the posterior expressed as

$$p(\mathbf{f},\mathbf{f}_m|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{f},\mathbf{f}_m)p(\mathbf{f},\mathbf{f}_m)}{p(\mathbf{y})} \tag{5.5}$$

Utilizing the optimum choice of $\mathbf{f}_m$ lead to $p(\mathbf{y}|\mathbf{f},\mathbf{f}_m) = p(\mathbf{y}|\mathbf{f}_m)$, Eq. (5.5) is also expressed as

$$p(\mathbf{f},\mathbf{f}_m|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{f}_m)p(\mathbf{f},\mathbf{f}_m)}{p(\mathbf{y})} \tag{5.6}$$

$$= \frac{p(\mathbf{y}|\mathbf{f}_m)p(\mathbf{f}_m|\mathbf{f})p(\mathbf{f})}{p(\mathbf{y})} \times \frac{p(\mathbf{f}_m)}{p(\mathbf{f}_m)} \tag{5.7}$$

$$= \frac{p(\mathbf{f}_m|\mathbf{f})p(\mathbf{f})}{p(\mathbf{f}_m)} \frac{p(\mathbf{y}|\mathbf{f}_m)p(\mathbf{f}_m)}{p(\mathbf{y})} \tag{5.8}$$

$$= p(\mathbf{f}|\mathbf{f}_m)p(\mathbf{f}_m|\mathbf{y}) \tag{5.9}$$

Using variational inference, as detailed in Appendix C, to select a variational distribution $Q(\mathbf{f},\mathbf{f}_m)$ we try to minimize the KL divergence between $Q(\mathbf{f},\mathbf{f}_m)$ and the posterior $p(\mathbf{f},\mathbf{f}_m|\mathbf{y})$.

Noting that minimizing

$$KL(Q(\mathbf{f},\mathbf{f}_m) \| p(\mathbf{f},\mathbf{f}_m|\mathbf{y})) \tag{5.10}$$

is equivalent to maximizing

$$L = \iint Q(\mathbf{f},\mathbf{f}_m)\log\left[\frac{p(\mathbf{f},\mathbf{f}_m,\mathbf{y})}{Q(\mathbf{f},\mathbf{f}_m)}\right]d\mathbf{f}d\mathbf{f}_m \tag{5.11}$$

Then enforcing the factorization property of $Q(\mathbf{f},\mathbf{f}_m) = \phi(\mathbf{f}_m)p(\mathbf{f}|\mathbf{f}_m)$, we have

$$L = \int \phi(\mathbf{f}_m)\left[\int p(\mathbf{f}|\mathbf{f}_m)\log p(\mathbf{y}|\mathbf{f})d\mathbf{f}\right] + \log(\frac{p(\mathbf{f}_m)}{\phi(\mathbf{f}_m)})d\mathbf{f}_m \tag{5.12}$$

$$\log G = \int p(\mathbf{f}|\mathbf{f}_m)\log p(\mathbf{y}|\mathbf{f})d\mathbf{f}$$

Using Gaussian distribution identities, as detailed in Appendix A, the $p(\mathbf{f}|\mathbf{f}_m)$ is given by

$$p(\mathbf{f}|\mathbf{f}_m) = N(\mathbf{f};\mathbf{C}_{nm}\mathbf{C}_{mm}^{-1}\mathbf{f}_m,\mathbf{C}_{nn} - \mathbf{C}_{nm}\mathbf{C}_{mm}^{-1}\mathbf{C}_{mn}) \tag{5.13}$$

and $\log G$ can be expressed as [8]

$$\log G = \log(N(\mathbf{y}; \mathbf{C}_{nm}\mathbf{C}_{mm}^{-1}\mathbf{f}_m, \mathbf{I}\sigma^2)) - \frac{1}{2\sigma^2}Tr(\mathbf{C}_{nn} - \mathbf{C}_{nm}\mathbf{C}_{mm}^{-1}\mathbf{C}_{mn}) \qquad (5.14)$$

Eq. (5.12) can also be expressed as

$$L = \int \phi(\mathbf{f}_m)\left[\log(G(\mathbf{f}_m, \mathbf{y})\frac{p(\mathbf{f}_m)}{\phi(\mathbf{f}_m)})\right]d\mathbf{f}_m \qquad (5.15)$$

Then reverse Jensen's inequality, as detailed in Appendix D, to maximize with respect to $\phi(\mathbf{f}_m)$, we can finally calculate a low bound on L

$$L_{lb} = \log\left[N(\mathbf{y}; 0, \mathbf{I}\sigma^2 + \mathbf{C}_{nm}\mathbf{C}_{mm}^{-1}\mathbf{C}_{mn})\right] - \frac{1}{2\sigma^2}Tr(\mathbf{C}_{nn} - \mathbf{C}_{nm}\mathbf{C}_{mm}^{-1}\mathbf{C}_{mn}) \qquad (5.16)$$

The quantity in Eq. (5.16) is computed as scaled as $\mathcal{O}(NM^2)$ for training and $\mathcal{O}(M^2)$ for prediction. The inducing inputs are defined by minimizing KL divergence between a variational Gaussian process and the true posterior Gaussian process, which allows avoiding overfitting and to rigorously approximate the exact GP by minimizing a distance between the sparse model and the exact one.

The selection of the inducing inputs and hyperparameters can be achieved using a variational EM (Expectation-Maximization) algorithm [149] where at the E step we greedily select the inducing inputs from the training data and at the M step we update the hyperparameters.

5.4 Benchmark studies

In this section, two numerical examples are provided to illustrate the application of VSGP emulation in this research.

5.4.1 An illustrative example of one dimension

In this numerical example, we suppose that we have a training dataset $(x_i, y_i)_{i=1}^n$ where $n = 100$, which has the noisy realizations of the latent function $f = x\sin(x)$. Therefore, each scalar value y_i can be expressed by adding a Gaussian noise term to $f(x)$ at input x_i (i.e. $y_i = f_i + \varepsilon_i$ where $\varepsilon_i \sim N(0, 0.2^2)$).

Figure 5.1 presents the standard (full) GPE of the latent function $f = x \sin(x)$ by the training datasets shown as red dots. The gray area is the 95% predictive confidence interval and the black line indicates the predictive mean values of the GPE.
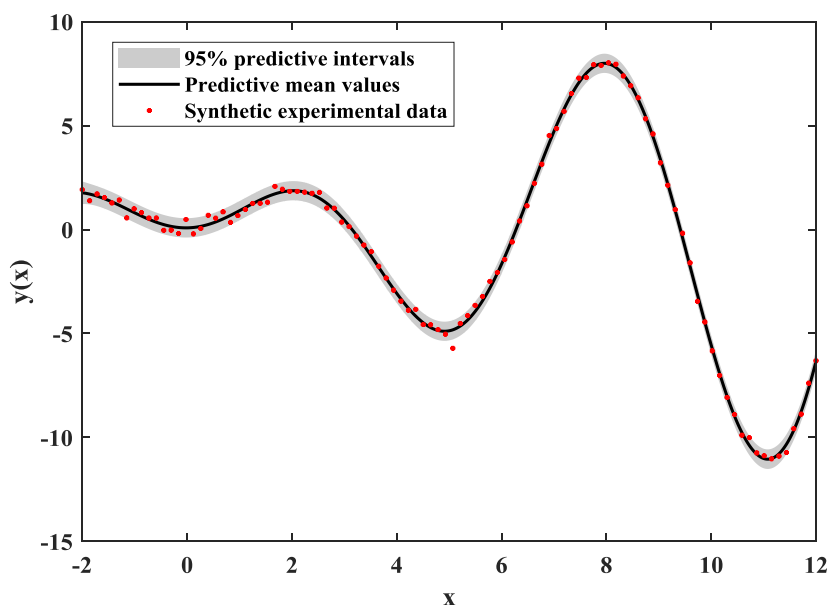
Figure 5.1 *The standard Gaussian process emulation of latent function with* $f = x \sin(x)$.

As described in detail in the section $5.2.3$ for the VSGP emulation, we try to select inducing variables $(m \leq n)$ between $[-2, 9]$ to construct the GP emulator, which has the ability to replace the standard GP emulator with approximately the same behaviors for predictions at untrained data points. In Figure 5.2, as a result, only 10 sparse inducing variables are selected to construct the VSGP emulator shown in green dots. The 95% predictive confidence intervals and predictive mean values of standard GP emulation and VSGP emulations are represented by color with black and blue, respectively.
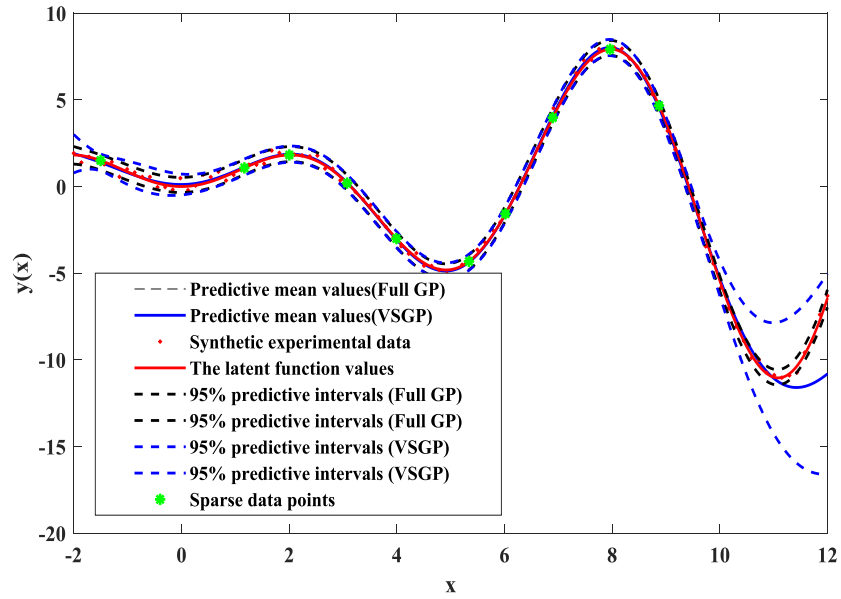
Figure 5.2 *The variational sparse Gaussian process emulation versus standard Gaussian process emulation of latent function with* $f = x \sin(x)$.

In addition, in order to explore the degree of approximation between the standard GP emulator and VSGP emulator, one quantitative performance measure called the predictive standard deviation diagnostics is used. In Figure 5.3, it is obvious that the values of predictive standard deviation for the standard GP emulator and VSGP emulator in $[-2, 9]$ are getting the closest. But this value is overestimated outside the support of inducing variables.
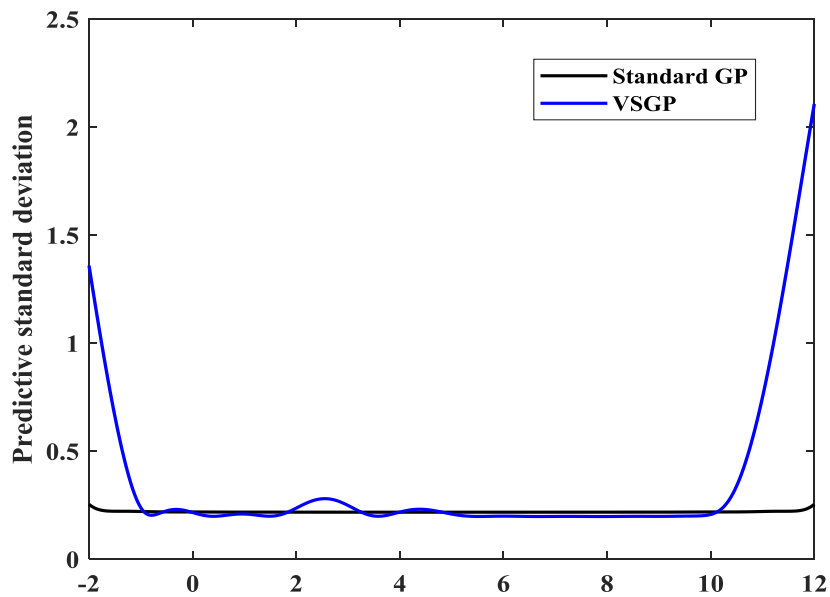
Figure 5.3 *The predictive standard deviation diagnostics for VSGPE (Blue) and standard Gaussian process emulation (Black) (both models are obtained through training and prediction).*

5.4.2 An illustrative example of two dimensions

In the following example, we assume that the explicit form of input-output mapping is not known and it is considered as a 'Black-box' model. In fact, Black-box models are ubiquitous in engineering (e.g. FE model solvers, computational fluid dynamics tools, climate models and among many others). In this case, only the mapping of inputs $\mathbf{X} \in \mathfrak{R}^{75 \times 2}$, and outputs $\mathbf{Y} \in \mathfrak{R}^{75 \times 1}$ are available. Figure 5.4 shows the standard GPE of the 'Black-box' model with the mapping relations of inputs and outputs for the response surface. In addition, the corresponding contour figure with 75 training points and the predictive standard deviation diagnostics are given in Figure 5.5. The value of negative log marginal likelihood (nlml) to training datasets, which is employed as another quantitative performance measure, is 91.5294.
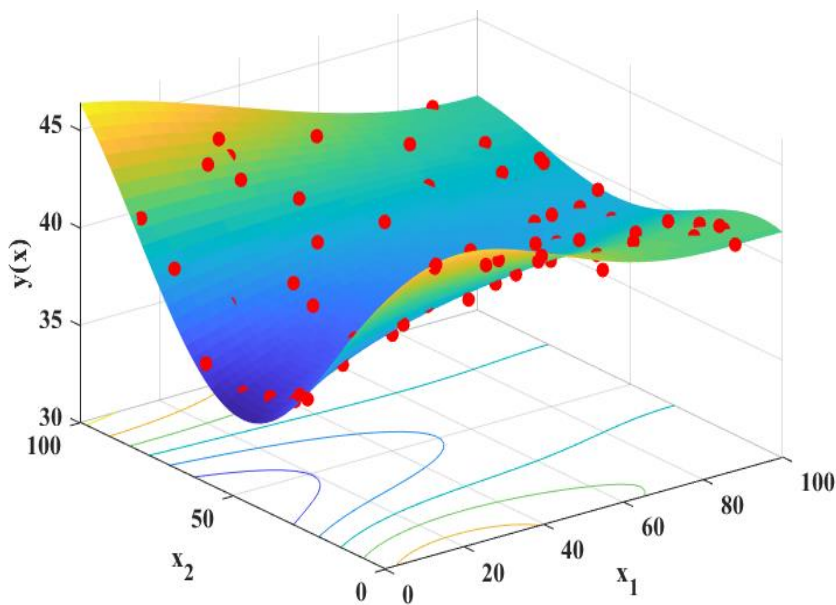
Figure 5.4 *The standard Gaussian process emulation of underlying function treated as 'Black-box' model for response surface.*
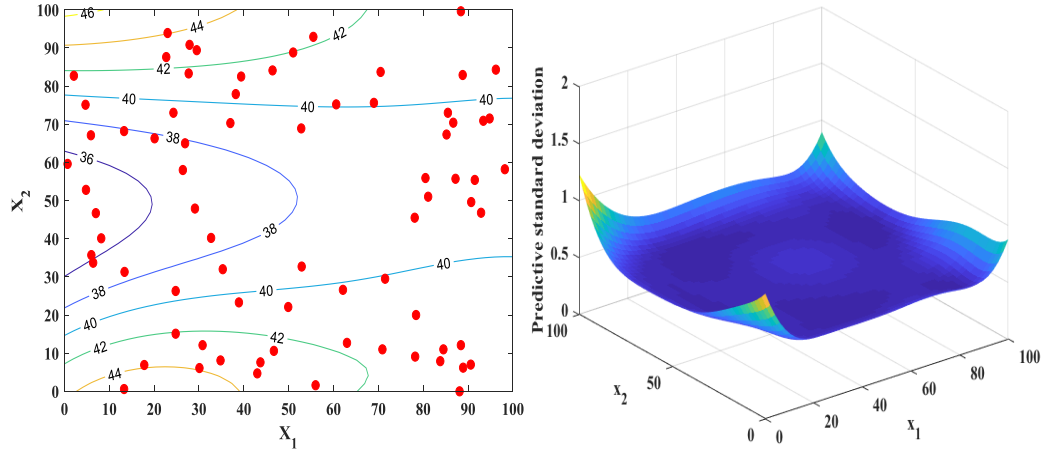
Figure 5.5 *Contour display (left) and predictive standard deviation diagnostics (right) for standard GP emulation.*

In Figure 5.6, the response surface of the 'Black-box' model is constructed using the VSGP emulation with 15 sparse inducing variables. The contour display is given on the left in Figure 5.7 with 15 sparse inducing points in blue dots. it seems that there are not many differences for the response surface constructed by the standard GP model and VSGP model with 15 sparse inducing variables, respectively. But if we compare the predictive standard deviation diagnostics figure shown on the right in Figure 5.7 to the same index by standard GP emulation shown on the right in Figure 5.5, they do have big differences, which indicates a bad VSGPE to be built. In addition, the value of nlml is 93.1725, a relatively large deviation from the value by standard GP emulation. As a consequence, the constructed VSGP emulator by 15 sparse inducing variables cannot be used to replace the standard GP emulator for the response surface of the 'Black-box' model.
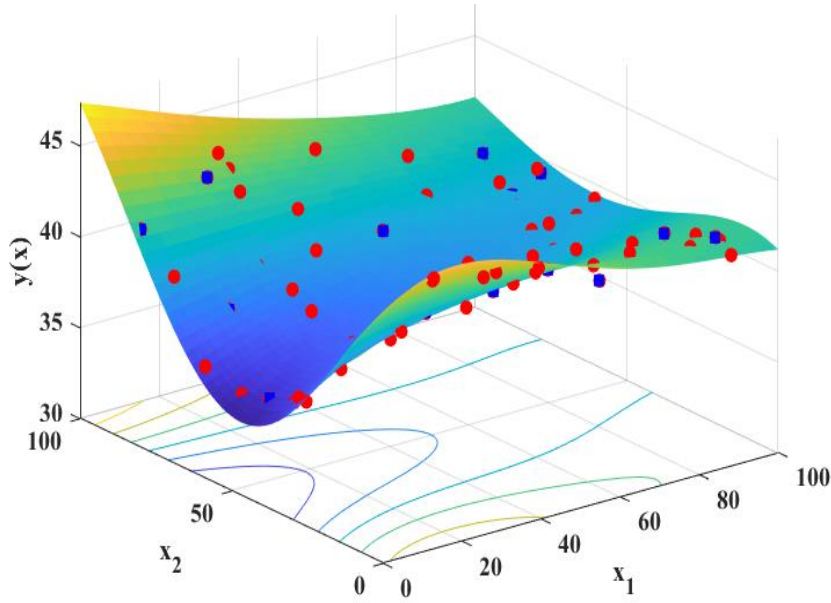
Figure 5.6 *The VSGP emulation of underlying function treated as 'Black-box' model for response surface with 15 sparse inducing variables.*
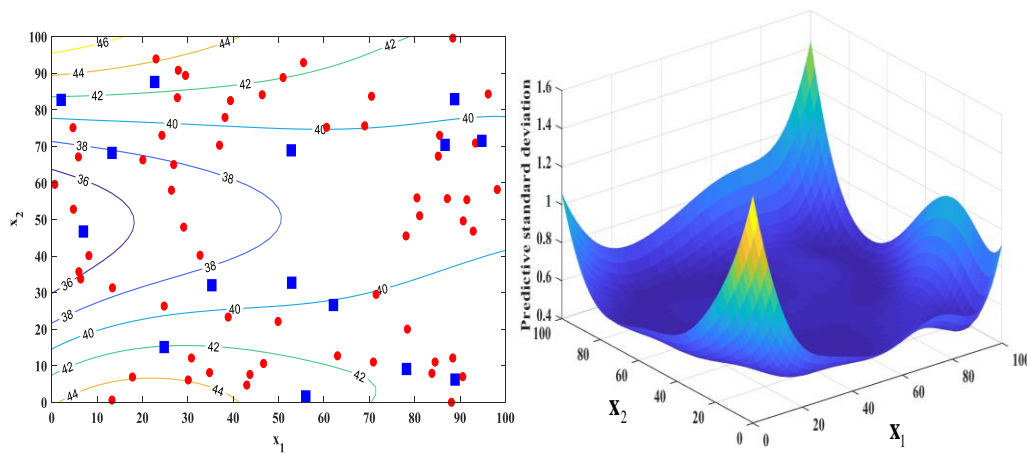
Figure 5.7 *Contour display (left) and predictive standard deviation diagnostics (right) for VSGP emulation with 15 sparse inducing variables.*

As described in detail in section 5.3, therefore, we increase the inducing inputs begins from 15 sparse points in the training data to build VSGPE. In Figure 5.8, the response surface of the 'Black-box' model is constructed using the VSGP emulation with 19 sparse inducing variables. The contour display is given on the left in Figure 5.9 with 19 sparse inducing points in blue dots. it seems that the response surface constructed by the standard GP model and VSGP model with 19 sparse inducing variables respectively are almost the same. In addition, there are also not big differences for figures of predictive standard deviation diagnostics by the standard GP emulator and VSGP emulator with 19 sparse inducing variables,

respectively. The value of nlml is 91.5677 compared with 91.5294 by standard GP emulator. As a result, this well constructed VSGP emulator by 19 sparse inducing variables can be used to represent the standard GP emulator for the response surface prediction of the 'Black-box' model.
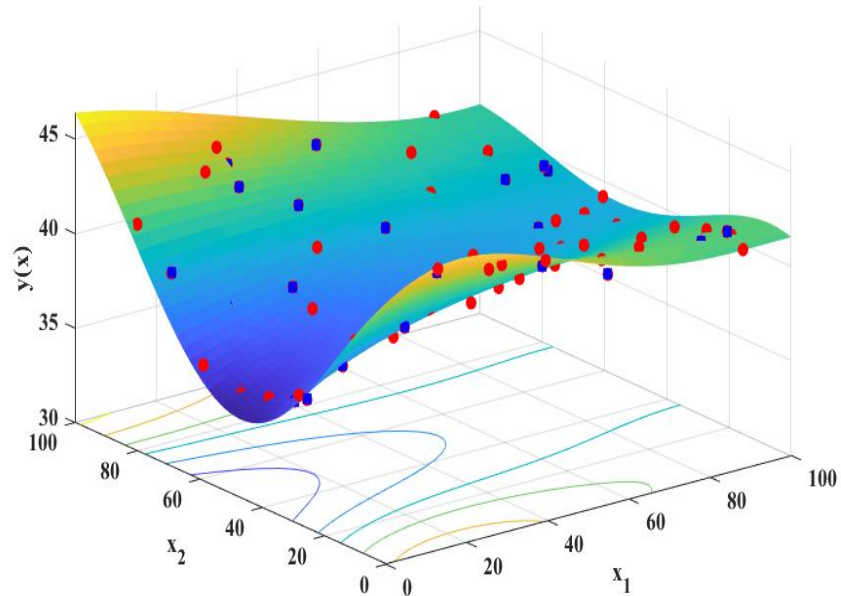
Figure 5.8 *The VSGP emulation of underlying function treated as 'Black-box' model for response surface with 19 sparse inducing variables.*
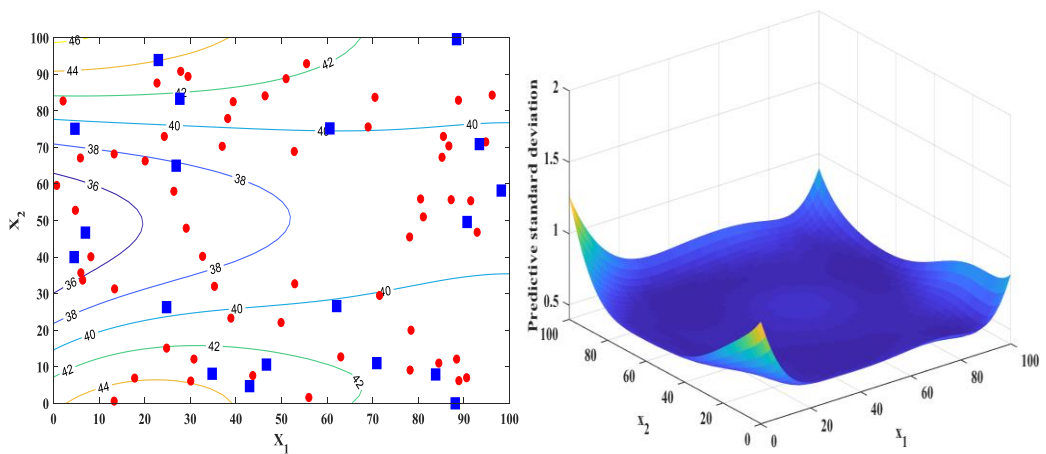
Figure 5.9 *Contour display (left) and predictive standard deviation diagnostics (right) for VSGP emulation with 19 sparse inducing variables.*

5.5 The DLR AIRMOD structure

5.5.1 Background description

Applying Bayesian FE model updating framework, as detailed in chapter 3, combined with the surrogate model techniques, as detailed in chapter 4 and 5, to the DLR-AIRMOD experimental structure is analyzed in this section. The DLR-AIRMOD structure is a replica of GARTEUR SM-AG19 benchmark, proposed by Balmes, et al. [150] and described in detail in references [151, 152]. This structure consists of 6 beam like components which are connected to each other with bolted joints shown in Figure 5.10.
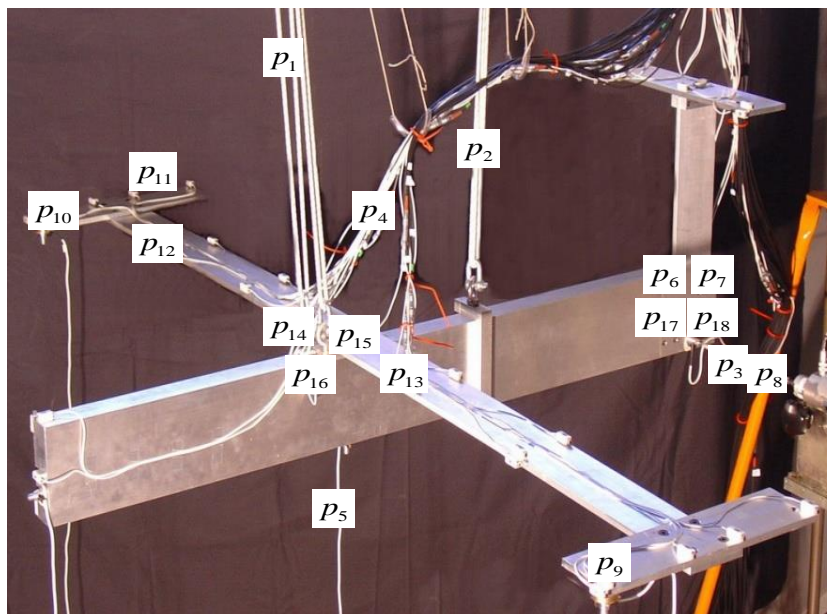
(*a*)
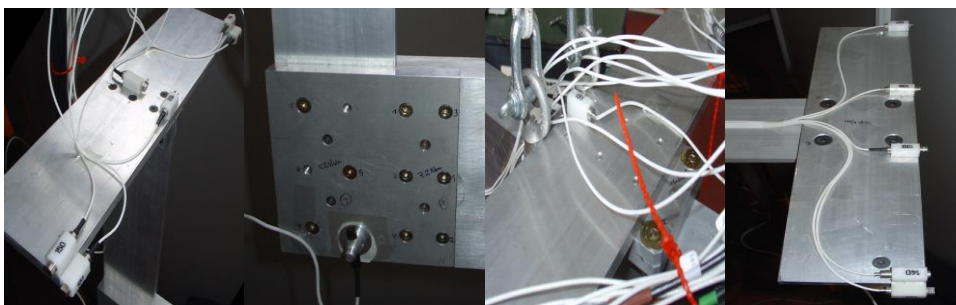
(*b*)   (*c*)   (*d*)   (*e*)

Figure 5.10 *DLR-AIRMOD structure (a) and beam joints (b)-(e) [152].*

The modes used in model updating were the lower modes with simpler shapes shown in Figure 5.11. In order to assess the variability in the measured data, a

series of modal tests are carried out with the bolted joints disassembled and reassembled 130 times, producing a maximum of 260 modal test datasets from single point excitation at two locations at $626z$ and $627y$ shown in Figure 5.12.

Mode 1: RBM Yaw   Mode 2: RBM Roll   Mode 3: RBM Pitch

Mode 4: RBM Heave   Mode 5: 2n Wing
Bending   Mode 6: 2n Wing
Bending

Mode 7: Anti-symm
Wing Torsion   Mode 8: Symm Wing
Torsion   Mode 10: 4n Wing
Bending

Mode 11: 1n Wing
Force-Aft   Mode 12: 2n Wing
Force-Aft   Mode 14: VTP Torsion

Mode 19: HTP Bending   Mode 20: HTP Force-Aft
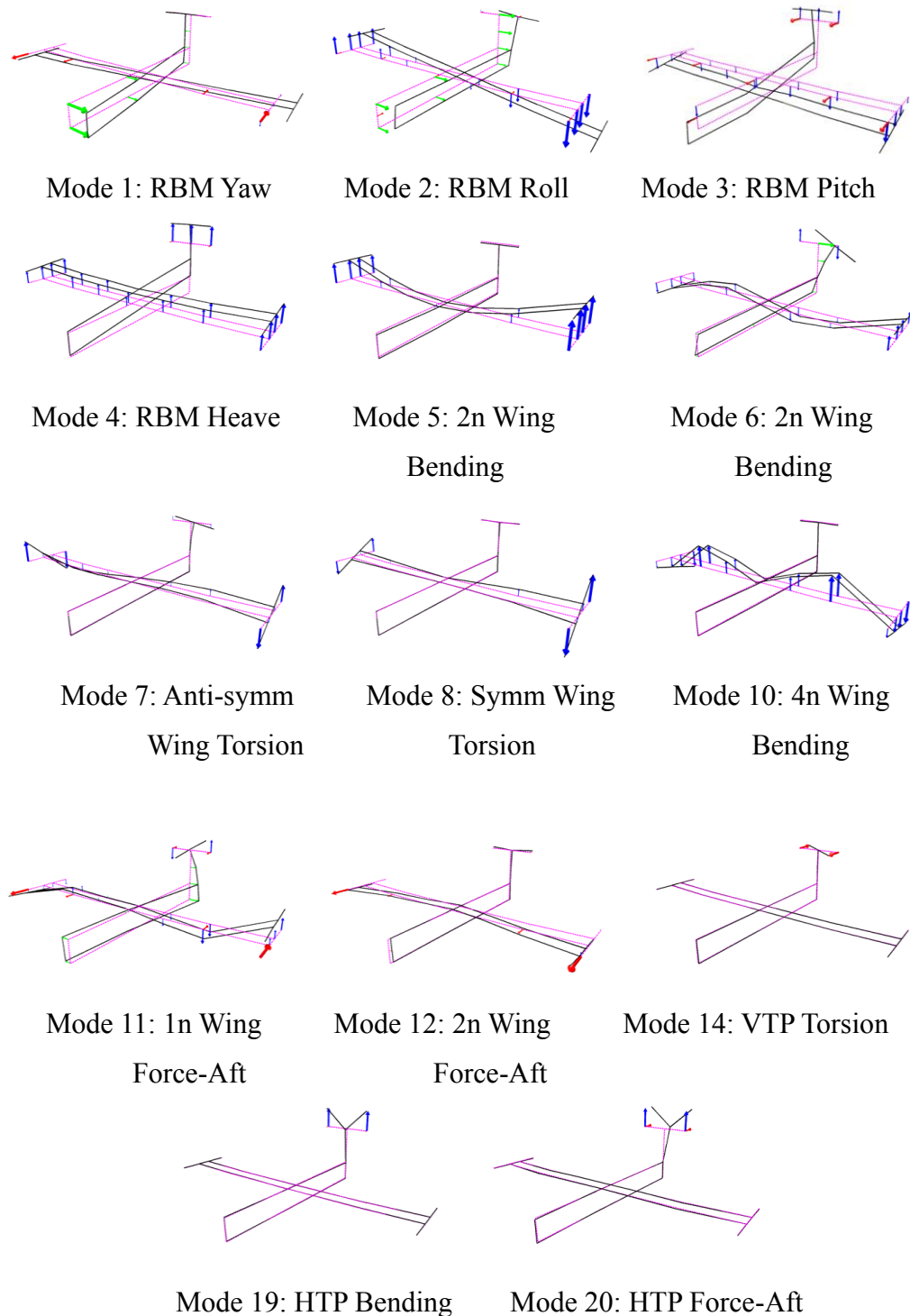
Figure 5.11 *Mode shapes used in model updating [152].*

<div align="center">(<i>a</i>)   (<i>b</i>)</div>
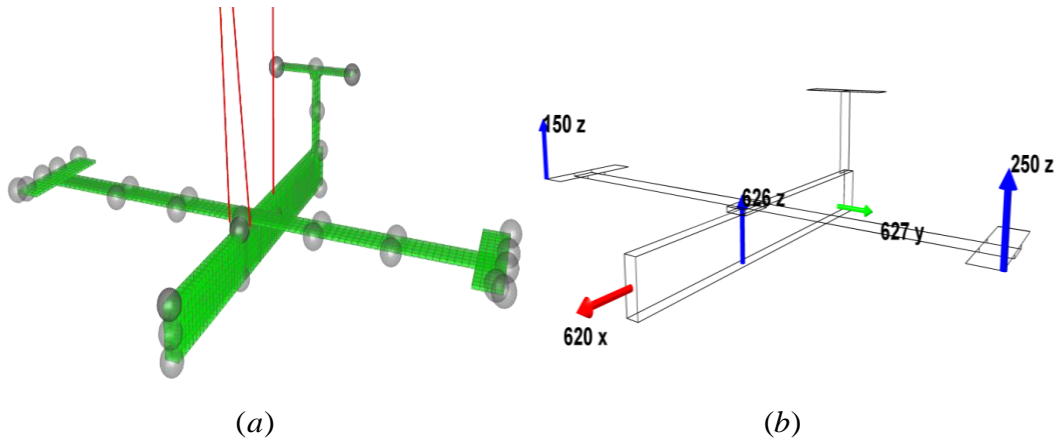
Figure 5.12 *The finite element model and exciter locations of the AIRMOD structure [152].*

The modal test results in statistics used for model updating with lower modes are shown in Table 5.1. From 130 samples, sample 82 was identified as an outlier which could be detected by a frequency drop for all extracted modes of that sample. Furthermore a system change took place between sample 43 and 44 resulting in a frequency shift of the first bending mode. The exact reason for this effect could not be determined because the data analysis was not conducted immediately after acquiring the time data. Therefore, 86 samples starting with sample 44 to 130 could be used. Nevertheless, some modes have been identified less than 86 times which is due to the automatic identification and correlation process. Govers, et al. [153] demonstrate the manufacturing variability is not reducible and should be categorized to the aleatory uncertainty for model updating. This subject has also attracted great attention described by Simoen, et al. [154] for damage assessment.

Table 5.1 *Statistics of identified frequency*

Mode	Description	$f_{mean}(Hz)$	$f_{std}(Hz)$	$f_{CoV}(\%)$	Samples
1	RBM Yam	0.23	0.006	2.41	41
2	RBM Roll	0.65	0.019	2.89	81
3	RBM Pitch	0.83	0.017	1.99	83
4	RBM Heave	2.17	0.024	1.11	86
5	2n Wing Bending	5.50	0.004	0.07	86
6	3n Wing Bending	14.91	0.017	0.12	86
7	Wing torsion (anti)	31.96	0.020	0.06	86

8	Wing torsion (symm)	32.33	0.017	0.05	86
10	4n Wing Bending	43.89	0.015	0.03	86
11	1n Wing force-aft	46.71	0.149	0.32	86
12	2n Wing force-aft	51.88	0.012	0.02	86
14	VTP Torsion	65.93	0.274	0.42	86
19	2n HTP bending	205.89	1.023	0.50	86
20	HTP force-aft	219.07	1.663	0.76	86

The FE model of AIRMOD shown in Figure 5.12 is built in MSC/NASTRAN consists of a variety of element types described in detail in Table 5.2. The selected parameters, both mass, and stiffness are listed in Table 5.3. The mass parameters (p_8 to p_{13}) are represented to the position of cable bundles, screws, and glue after each reassembly of the AIRMOD structure. The stiffness parameters are divided into two groups, namely support stiffness (p_1 to p_5) and joint stiffness (p_6, p_7 and p_{14} to p_{18}). These selected updating parameters are chosen after a detailed sensitivity analysis in [152].

Table 5.2 *Details of Finite element model*

Element type	Number
CHEXA	1440
CPENTA	6
CELAS1	561
SMASS1	55
CONM2	18
CROD	3

Table 5.3 *List of updating parameters*

	Type	Location	Description	Init. Val	Unit
p_1	Stiffness	Front bungee cord	Elastic modulus	1.80E+03	N/m^2
p_2	Stiffness	Rear bungee cord	Elastic modulus	7.50E+03	N/m^2

p_3	Stiffness	VTP/HTP joint	Sensor cable arness-y dir^n	1.30E+02	N/m
p_4	Stiffness	Wing/fuselage joint	Sensor cable arness-y dir^n (top)	7.00E+01	N/m
p_5	Stiffness	Wing/fuselage joint	Sensor cable arness-y dir^n (bott^m)	7.00E+01	N/m
p_6	Stiffness	VTP/HTP joint	Joint stiffness-x,y directions	1.00E+07	N/m
p_7	Stiffness	VTP/HTP joint	Joint stiffness-z direction	1.00E+09	N/m
p_8	Mass	VTP/HTP joint	Sensor cables	2.00E-01	kg
p_9	Mass	Wingtip right wing	Screws and glue	1.86E-01	kg
p_{10}	Mass	Wingtip left wing	Screws and glue	1.86E-01	kg
p_{11}	Mass	Wingtip left wing	Sensor cable on wings	1.50E-02	kg
p_{12}	Mass	Out^r wing left/right	Sensor cable on wings	1.50E-02	kg
p_{13}	Mass	Inn^r wing left/right	Sensor cable on wings	1.50E-02	kg
p_{14}	Stiffness	Wing/fuselage joint	Joint stiffness-x direction	2.00E+07	N/m
p_{15}	Stiffness	Wing/fuselage joint	Joint stiffness-y direction	2.00E+07	N/m
p_{16}	Stiffness	Wing/fuselage joint	Joint stiffness-z direction	7.00E+06	N/m
p_{17}	Stiffness	VTP/ fuselage joint	Joint stiffness-x direction	5.00E+07	N/m
p_{18}	Stiffness	VTP/ fuselage joint	Joint stiffness-y direction	1.00E+07	N/m

5.5.2 Stochastic finite element model updating

Constructing the validated surrogate models to represent the computationally expensive FE models when carrying out the Bayesian FE model updating procedure described in chapter 3, which should be constructed beforehand. Since the parameter dimensionality increases to 18 and the big differences in value magnitude between them so that the available data will become sparse. This sparsity is problematic for a method that requires statistical significance. Therefore, all the training datasets are designed to be normalized [155] to guarantee the numerical stability of surrogate modeling to the AIRMOD structure in this section. Through running the full FE model 2000 times, we have 1800 training points and 200 separately validation points by Latin Hypercube Sampling (LHS).

In order to explore the influence of different covariance functions on the predictive responses of surrogate models, three different covariance functions (kernels) (i.e. SE, Matérn $\left(v = \frac{1}{2} \right)$, and Matérn $\left(v = \frac{3}{2} \right)$) are considered in this section. The final validation results in 200 separately validation points indicate that the standard GPE with SE and Matérn $\left(v = \frac{3}{2} \right)$ covariance functions are well trained and validated through IPE diagnostics. However, the standard GP emulators with Matérn $\left(v = \frac{1}{2} \right)$ covariance functions result in 8 of 14 surrogate models failed to validated by IPE diagnostics to frequency responses. The failure validation for f_4 and f_{11} by IPE diagnostics are selected to display in Figure 5.13. In Figure 5.13, a large number of individual values, with absolute values larger than 2 which indicate a conflict between the simulator and the emulator. As a result, the other two covariance functions (SE and Matérn $\left(v = \frac{3}{2} \right)$) are chosen to construct the VSGPE.
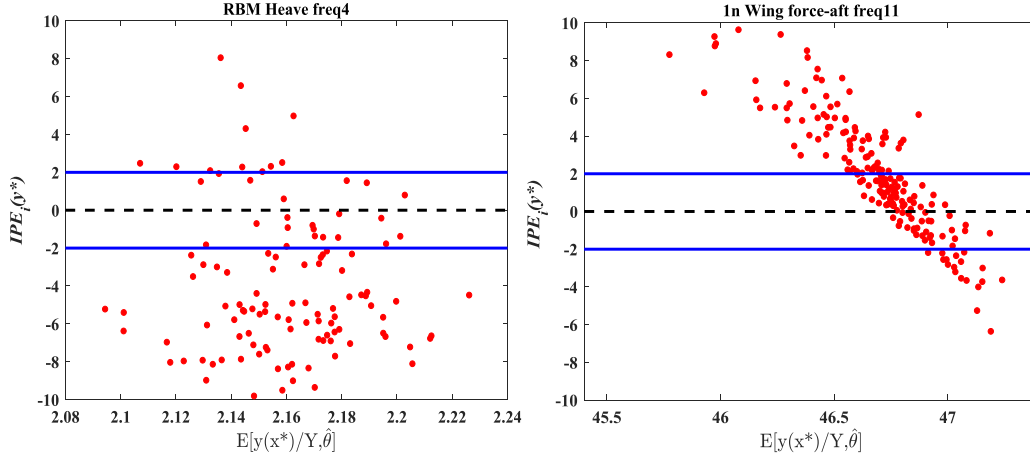
Figure 5.13 *The IPE of standard GP emulation with Matérn $\left(\text{v} = \frac{1}{2}\right)$ kernel for*
   f_4 *and* f_{11}.

In order to construct the surrogate models-VSGPE described in detail in section 5.3, the performance measure called Negative Log Marginal Likelihood (nlml), also refers to the negative log probability of the training data described in section 5.4, is employed. If the trained VSGPE could represent the corresponding full standard GP models with certain inducing variables, the values of nlml between them have to be infinitely close or even identical. In addition, The VSGPE construction procedure indicates that the number of inducing variables for VSGP metamodels with SE covariance function to 14 natural frequency responses is between 350 and 600 instead of original 1800 training points and the number of inducing variables for of VSGP metamodels with Matérn $\left(v = \frac{3}{2}\right)$ covariance function to 14 natural frequency responses is between 350 and 450 instead of original 1800 training points. Figure 5.14 and Figure 5.15 show VSGPE for natural frequency responses with SE and Matérn $\left(v = \frac{3}{2}\right)$ kernels with selected certain inducing variables respectively, which are assessed by performance measure (nlml). All the VSGP metamodels with certain inducing variables to 14 frequency responses by SE and Matérn $\left(v = \frac{3}{2}\right)$ kernels are constructed and a further validation process namely IPE, which is supposed to carry out.
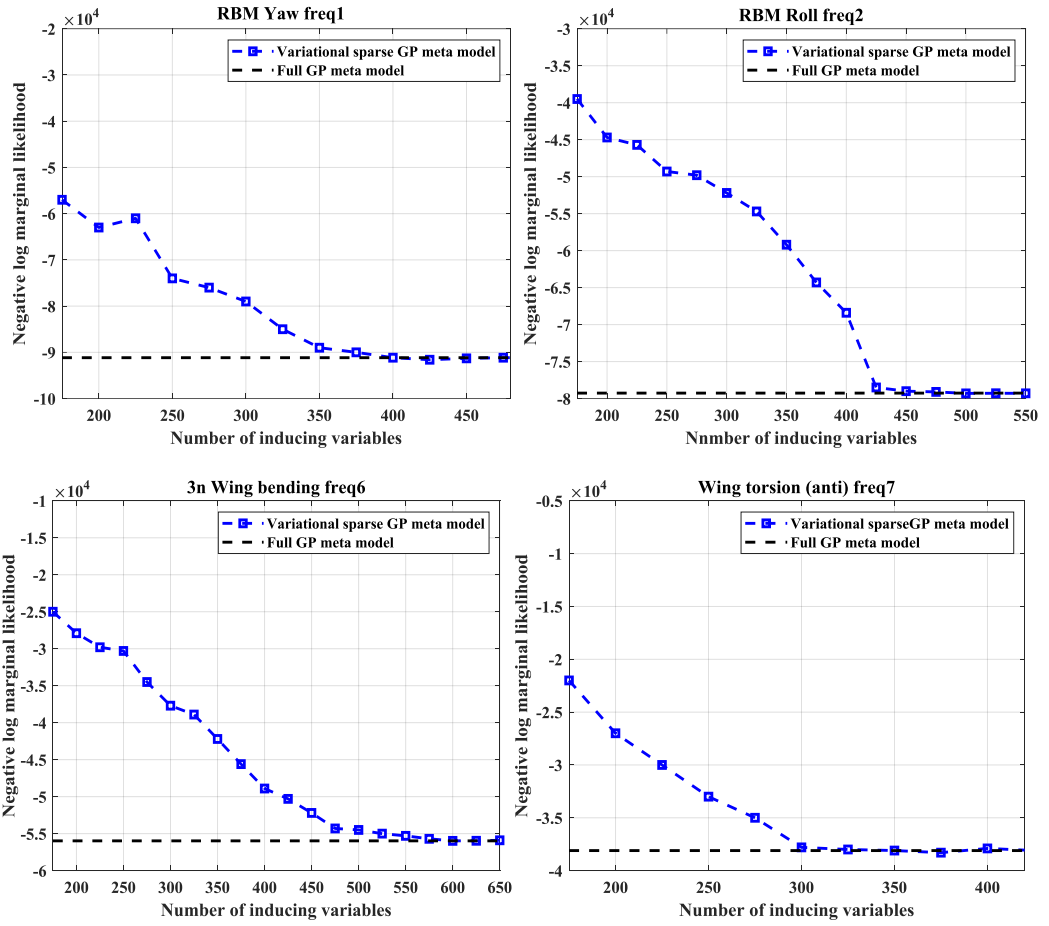
Figure 5.14 *The performance measure of VSGP emulation with SE kernel for* f_1
f_2 f_6 *and* f_7.
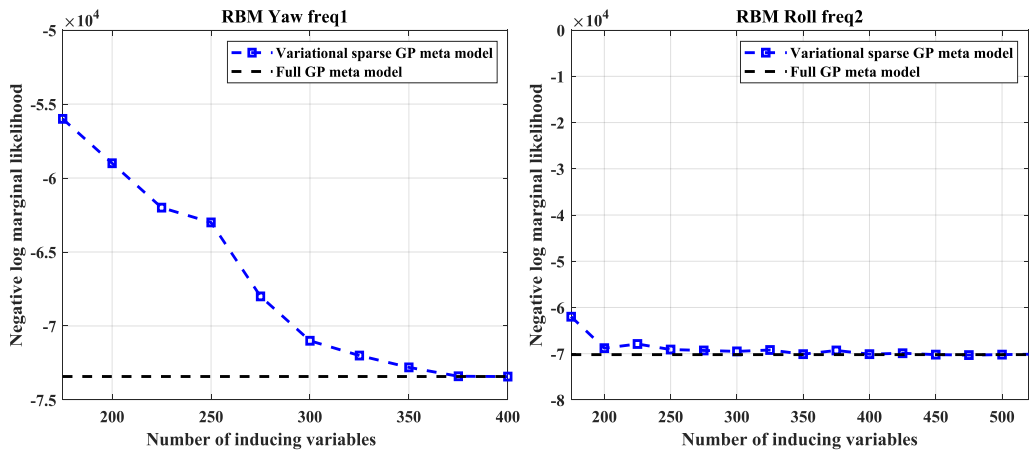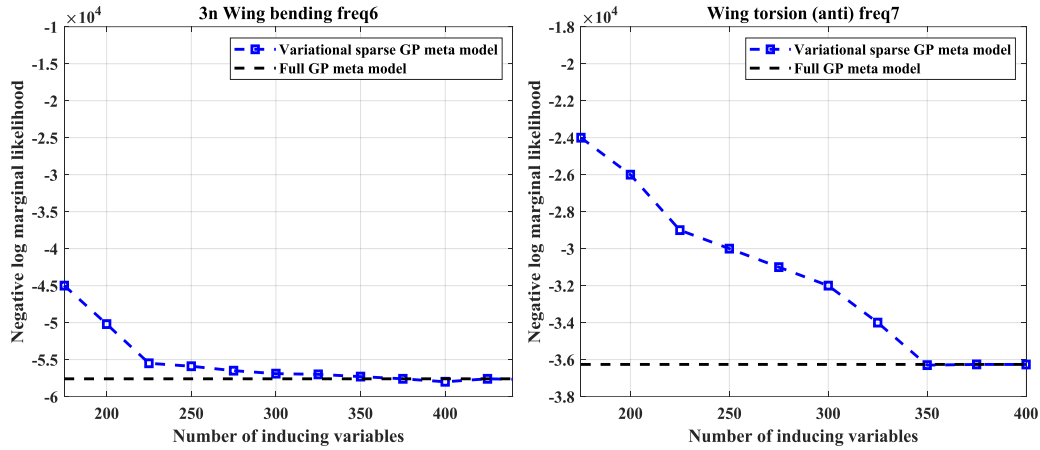
Figure 5.15 *The performance measure of VSGP emulation with Matérn* $\left(v = \frac{3}{2}\right)$ *kernel for* f_1 f_2 f_6 *and* f_7 .

When these two types of VSGP emulators are validated to represent the full FE model by IPE, Bayesian model updating procedure is processed to the AIRMOD structure. Figure 5.16 and Figure 5.17 present the IPE to VSGPE with SE and Matérn $\left(v = \frac{3}{2}\right)$ kernels for f_1, f_6, f_7 and f_{19}, respectively. In addition, it can be seen from Figure 5.16 and Figure 5.17, the validation performance of Matérn $\left(v = \frac{3}{2}\right)$ kernel under IPE diagnostics for certain natural frequency performances is better than SE kernel as fewer validation points are outside the region $[-2, 2]$.
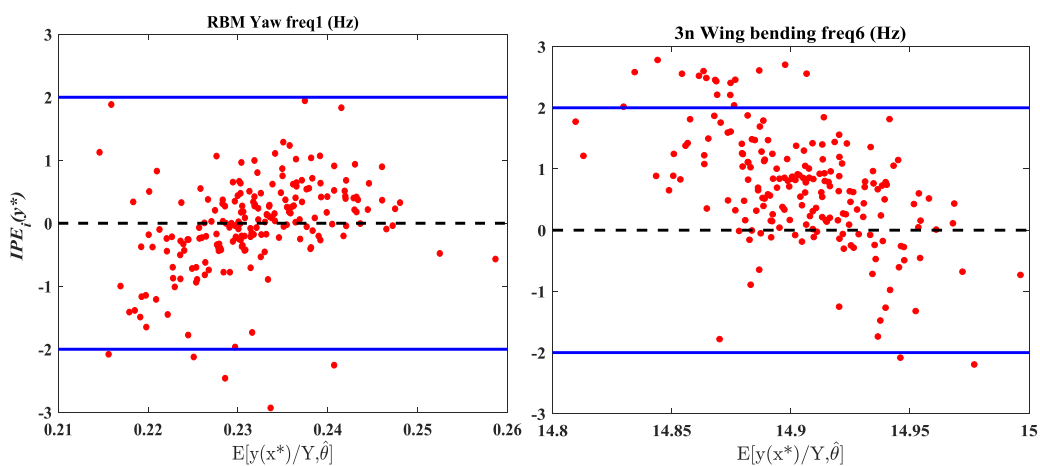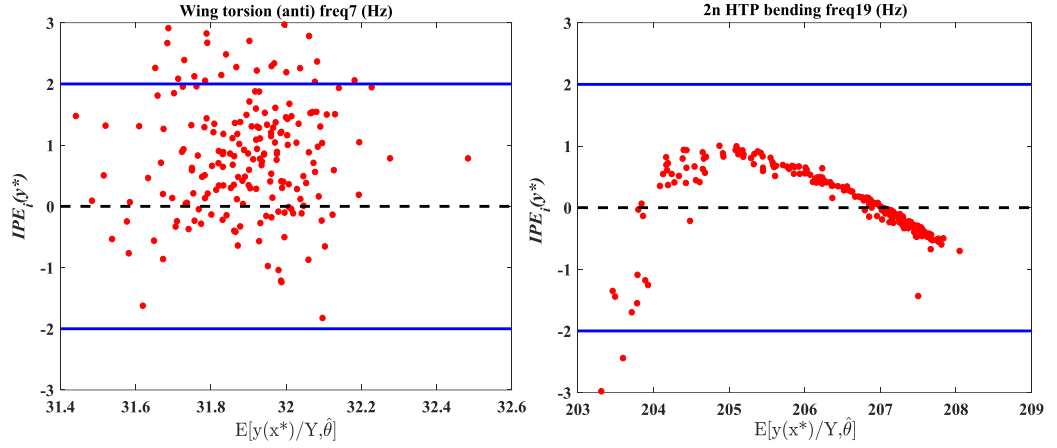
Figure 5.16 *The IPE of VSGP emulation with SE kernel for* f_1, f_6, f_7 *and* f_{19}.
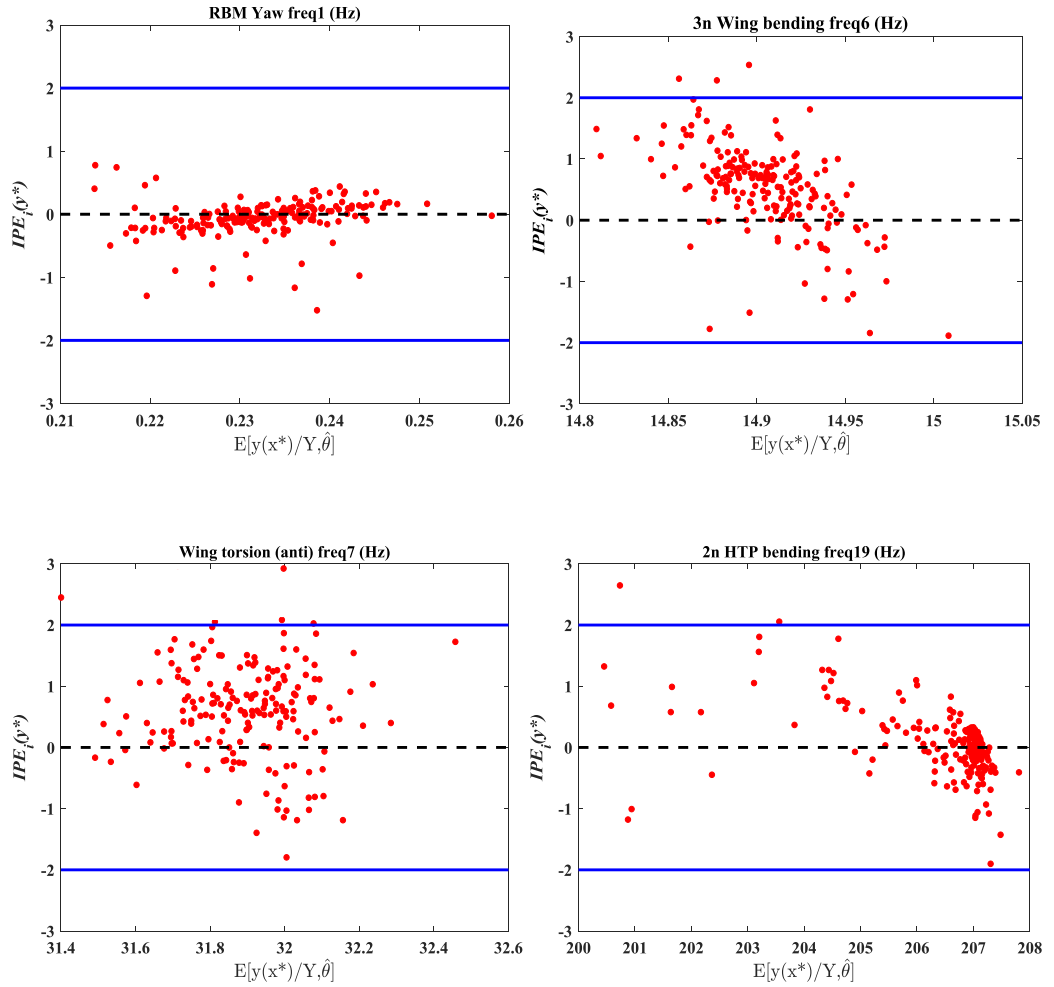
Figure 5.17 *The IPE of VSGP emulation with Matérn* $\left(v = \frac{3}{2} \right)$ *kernel for* f_1, f_6 f_7 *and* f_{19}.

After the two types of well-trained VSGPE with SE and Matérn $\left(v = \frac{3}{2} \right)$ covariance functions respectively, which are built, the Bayesian FE model

updating procedure using TMCMC described in chapter 3 are then conducted. For both situations, a uniform prior is used within the range of $5\% - 200\%$ for the initial parameter values, to indicate virtually no knowledge of the AIRMOD physics [156], and the updated results are in principle entirely dependent upon the test data. All the updated values of 18 parameters are shown in Table 5.4 and the posterior distribution of selected updated parameters through Bayesian FE model updating is estimated by kernel density estimation technique [157] shown in Figure 5.18. In Figure 5.18, the red line represents posterior distribution for certain updating parameters with Matérn $\left(v = \dfrac{3}{2} \right)$ covariance function, the blue line represents posterior distribution for certain updating parameters with SE covariance function, and the black dotted line is initial parameter value. In addition, the red and blue dotted line indicates mean values of the posterior distribution for certain parameters with Matérn $\left(v = \dfrac{3}{2} \right)$ and SE kernels, respectively.
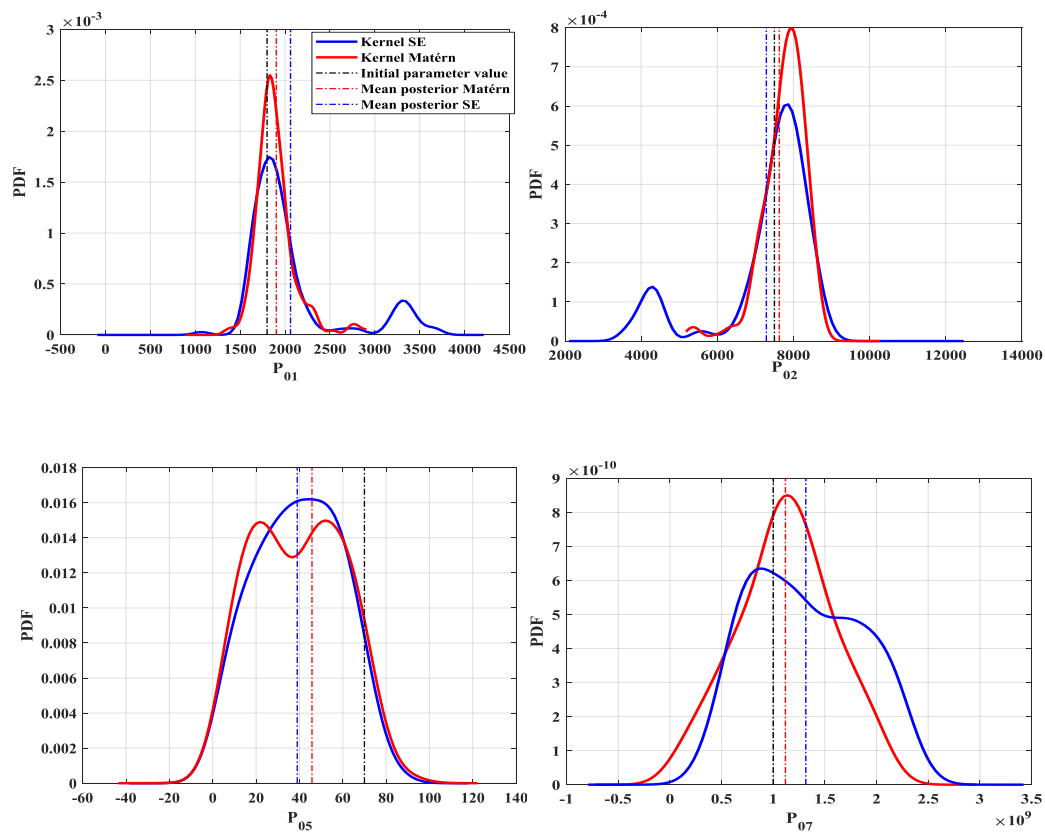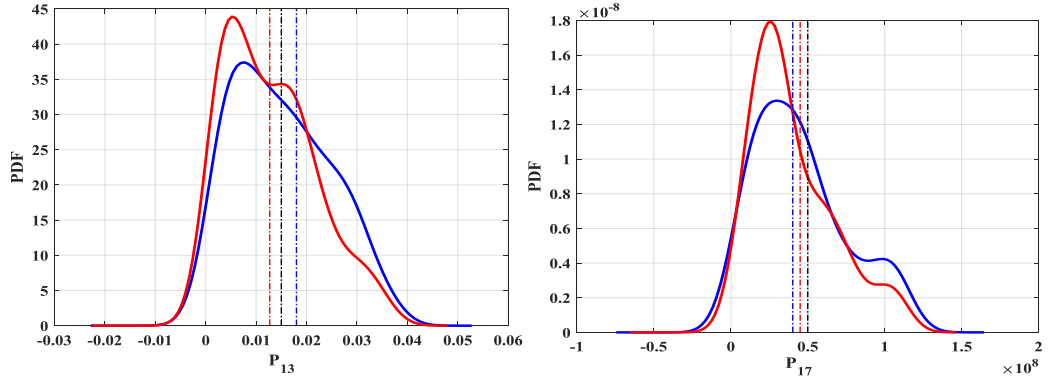
Figure 5.18 *Results of selected updating parameters of AIRMOD structure.*

The detailed results for all the updated parameters are shown in Table 5.4. In Table 5.4, the mean value u_{mean}, the 5th and the 95th percentiles $\left[p_5, p_{95} \right]$ for posterior distributions of updated parameters are given. The KL divergence is used to compare two distributions of updating parameters for SE kernel and Matérn $\left(v = \dfrac{3}{2} \right)$ kernel, respectively. In Table 5.4, it is concluded that the relative large difference for KL divergence is found in $p_6, p_7, p_{12}, p_{15}, p_{17}$ and p_{18}.

In addition, the overall number of selected inducing variables to construct the VSGPE with Matérn $\left(v = \dfrac{3}{2} \right)$ kernel is between $\left[300, 450 \right]$ for each of 14 frequency responses, that is fewer than the number of selected inducing variables to construct the VSGPE with SE kernel between $\left[350, 600 \right]$. It indicates a relative less computational cost is required for Matérn $\left(v = \dfrac{3}{2} \right)$ kernel.

Table 5.4 *List of updating parameters*

	Init. Val	Unit	SE kernel (updated)		Matérn $\left(v = \dfrac{3}{2} \right)$ kernel (updated)		KL Div
			u_{mean}	$\left[p_5, p_{95} \right]$	u_{mean}	$\left[p_5, p_{95} \right]$	
p_1	1.80E+03	N/m^2	2.06E+03	$\left[1.61, 3.36 \right]$ E+03	1.90E+03	$\left[1.62, 2.32 \right]$ E+03	4.48
p_2	7.50E+03	N/m^2	7.28E+03	$\left[4.17, 8.56 \right]$ E+03	7.73E+03	$\left[6.72, 8.51 \right]$ E+03	4.02
p_3	1.30E+02	N/m	1.51E+02	$\left[1.22, 1.63 \right]$ E+02	1.43E+02	$\left[1.21, 1.80 \right]$ E+02	4.40

p_4	7.00E+01	N/m	4.61E+01	$[0.91, 8.34]$ E+01	4.56E+01	$[0.60, 8.03]$ E+01	3.39
p_5	7.00E+01	N/m	3.96E+01	$[0.73, 6.98]$ E+01	4.58E+01	$[1.30, 8.47]$ E+01	4.76
p_6	1.00E+07	N/m	1.36E+06	$[1.07, 1.87]$ E+06	1.55E+06	$[1.10, 2.89]$ E+06	24.9
p_7	1.00E+09	N/m	1.11E+09	$[0.31, 1.92]$ E+09	1.32E+09	$[0.58, 2.19]$ E+09	49.8
p_8	2.00E-01	kg	2.30E-01	$[1.90, 2.53]$ E-01	2.20E-01	$[1.86, 2.46]$ E-01	0.18
p_9	1.86E-01	kg	2.04E-01	$[1.81, 2.23]$ E-01	2.00E-01	$[1.71, 2.21]$ E-01	0.86
p_{10}	1.86E-01	kg	1.99E-01	$[1.81, 2.23]$ E-01	2.00E-01	$[1.77, 2.29]$ E-01	0.21
p_{11}	1.50E-02	kg	1.50E-02	$[1.23, 2.04]$ E-02	1.47E-02	$[1.17, 1.71]$ E-02	1.71
p_{12}	1.50E-02	kg	1.09E-02	$[0.41, 1.70]$ E-02	1.17E-02	$[0.36, 1.76]$ E-02	17.1
p_{13}	1.50E-02	kg	1.80E-02	$[0.20, 3.07]$ E-02	1.27E-02	$[0.20, 3.03]$ E-02	5.11
p_{14}	2.00E+07	N/m	3.59E+07	$[2.87, 4.09]$ E+07	3.38E+07	$[1.88, 4.04]$ E+07	3.91
p_{15}	2.00E+07	N/m	1.53E+07	$[1.23, 1.87]$ E+07	1.71E+07	$[1.37, 2.27]$ E+07	28.9
p_{16}	7.00E+06	N/m	2.97E+06	$[2.44, 3.43]$ E+06	2.92E+06	$[2.55, 3.23]$ E+06	1.91
p_{17}	5.00E+07	N/m	4.51E+07	$[0.56, 10.5]$ E+07	4.02E+07	$[0.92, 9.97]$ E+07	30.2
p_{18}	1.00E+07	N/m	6.70E+07	$[1.40, 11.0]$ E+07	5.73E+07	$[0.69, 10.9]$ E+07	18.6

The correlation matrix, which is calculated based on the Pearson correlation coefficient definded as $\rho_{XY} = \dfrac{cov(X,Y)}{\sigma_X \sigma_Y}$ for updating parameters with Matérn $\left(v = \dfrac{3}{2} \right)$, which is given in Figure 5.19. It is concluded that the relative strong correlations are always to be found between parameters located at the same joint. For example, p_4 and p_5 which are the sensor harness stiffness at the wing-fuselage connection indicate the strongest correlation. Mode 19 (2n HTP bending) is seen in Table 5.1 to have large coefficient of variation (CoV) and is sensitive to

parameter p_6 (joint stiffness-x,y directions), p_7 (joint stiffness-z direction), p_{17} (joint stiffness-x direction), p_{18} (joint stiffness-y direction), which are shown to have large percentile $\left[p_5, p_{95} \right]$ in Table 5.4. Similarly, mode 20 (HTP force-aft) is strongly affected by parameter p_{17} (joint stiffness-x direction). It is seen in Table 5.4 that p_{17} has large percentile $\left[p_5, p_{95} \right]$ and the the mode with highest variability in Table 5.1 is mode 20.
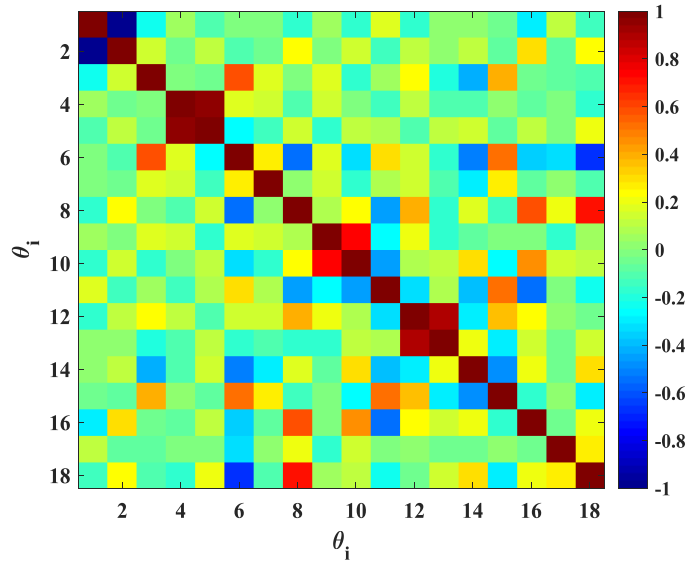
Figure 5.19 *Updated parameter correlation matrix with Matérn* $\left(v = \frac{3}{2} \right)$ kernel.

Figure 5.20 and Figure 5.21 show frequency distributions under Bayesian FE model updating using SE and Matérn $\left(v = \frac{3}{2} \right)$ covariance functions, respectively. It can be seen from the Figure 5.20 and Figure 5.21 that the general performance of frequency distributions by VSGP emulators using Matérn $\left(v = \frac{3}{2} \right)$ kernels is better than the corresponding SE kernels since the 95% confidence region (ellipse) of updated frequency responses marked by blue get closer to the test datasets marked by red. In addition, it is also worth noting that, in Figure 5.20 and Figure 5.21 with higher modes, especially f_{14}, f_{19}, f_{20}, the 95% confidence region (ellipse) of the frequency distributions from the Bayesian model updating are not replicated closely from the test datasets. But, the performance of VSGP emulators

with Matérn $\left(v = \dfrac{3}{2} \right)$ covariance functions is still better than the SE covariance functions.

The reasons to cause discrepancies for higher modes (f_{14}, f_{19} and f_{20}) come from many aspects. For example, the uniform priors for updating parameters maybe not appropriate which should be incorporated with some strong information. Model discrepancy [158], as another very important kind of modeling uncertainty, is not being considered in this application. However, it is an important topic in uncertainty quantification and some insight will be given in chapter 6.
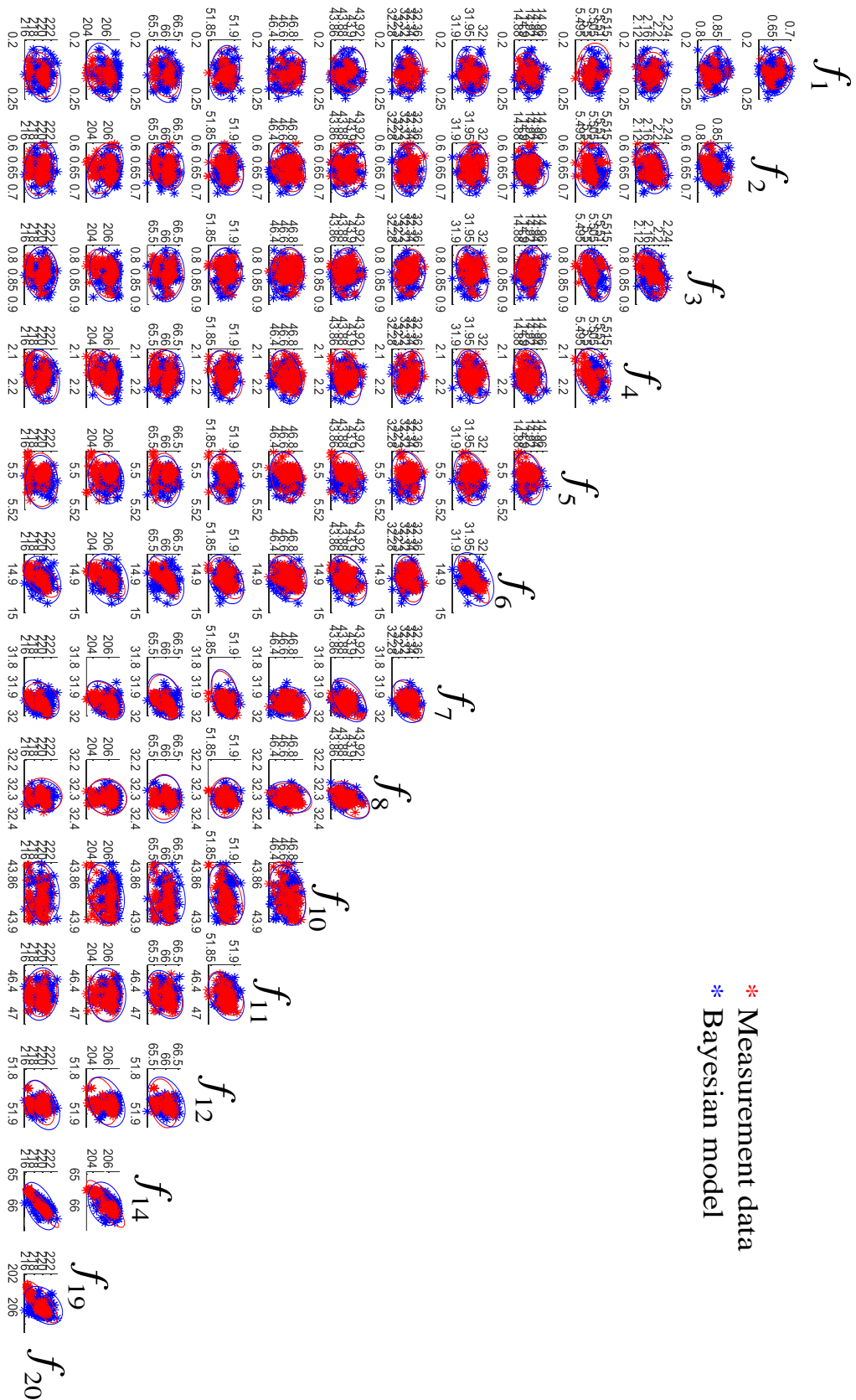
Figure 5.20 *Frequency distributions using VSGP emulation with* Matérn $\left(v = \dfrac{3}{2}\right)$ *kernel versus measurement data.*
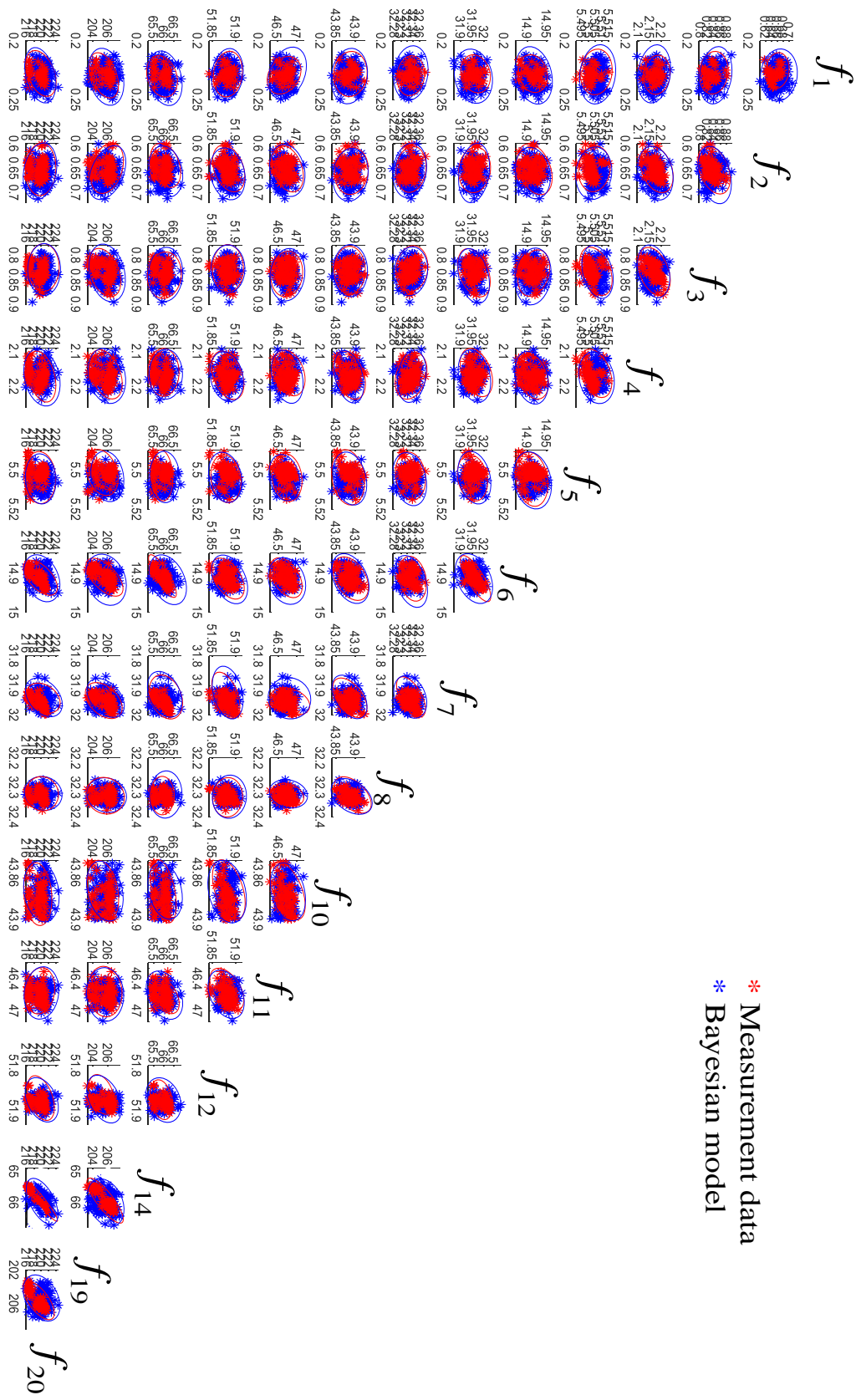
Figure 5.21 *Frequency distributions using VSGP emulation with* SE *kernel versus measurement data.*

The training time for the single independent output from VSGP metamodel of the full FE model (excluding the cost of running the full FE model 2000 times) is around 39 seconds of CPU time which compared to 157 seconds for standard full GP metamodels training and 172 seconds for ANNs [159]. The well-trained metamodel (VSGP) with relatively small inducing variables at around 375 $\left(375 \ll 1800\right)$ that allows the reduction of the training cost scaled from $\mathcal{O}(1800^3)$ to $\mathcal{O}(1800 \times 375^2)$ and prediction cost from $\mathcal{O}(1800^2)$ to $\mathcal{O}(375^2)$ for each test case. As the results obtained so far, it is concluded that applying Bayesian FE model updating combined VSGP surrogate models with Matérn $\left(v = \frac{3}{2}\right)$ covariance functions using TMCMC [70] is an efficient model updating method to the DLR-AIRMOD structure.

5.5.3 Results discussion

In this section, Bayesian FE model updating framework combined with high-fidelity, less time-consuming surrogate models called VSGP metamodels are presented. Two case studies were carried out. The first case was two numerical examples to illustrate the formulation of VSGPE. The second case is putting this kind of metamodels into a Bayesian model updating framework to do model updating to a complex experimental structure DLR-AIRMOD. The computational cost for this real-time application of Bayesian model updating is a hard requirement and the problem has been solved effectively using the well-constructed VSGP surrogate models and parallel computing strategy (20 cores of a multi-AMD Opteron processor 6168 system). In addition, the presented Bayesian model updating procedure is implemented in the open source software Open-COSSAN [160, 161].

Three different kinds of covariance functions are employed and analyzed in order to obtain high-fidelity surrogate models. By conducting updating parameter results comparison for the AIRMOD structure between two different covariance functions (Matérn $\left(v = \frac{3}{2}\right)$ and SE kernels), it is concluded that using Matérn $\left(v = \frac{3}{2}\right)$ covariance functions to build the metamodels is more favorable than the SE covariance function. The proposed Bayesian FE model updating framework

combined with VSGP surrogate models and parallel computing strategy, which was proved to be effective and far less computationally intensive compared to using other standard surrogate modeling techniques to the DLR-AIRMOD structure. More importantly, this Bayesian model updating framework also enriches the theory of stochastic model updating to large scale complex engineering structures.

5.6 Chapter summary

In this chapter, a brief literature review of the computationally efficient GPEs was described. The Bayesian FE model updating was then carried out to the DLR AIRMOD structure. This Bayesian model updating method combined surrogate modeling techniques named the VSGPE, which is built based on the notion of standard GPE, the TMCMC method described in chapter 3 and parallel computing strategy to improve the updating efficiency. Two numerical examples were provided to illustrate the application of VSGPE. In addition, the SE covariance function and Matern $\left(v = \dfrac{3}{2} \right)$ covariance functions are employed and analyzed in order to obtain high-fidelity surrogate models to the AIRMOD structure updating. It was found that using Matern $\left(v = \dfrac{3}{2} \right)$ covariance functions to build the VSGPE is more favorable than the SE covariance function to do the Bayesian model updating to AIRMOD structure.

6 The importance of model discrepancy

6.1 Introduction

In chapter 5, the model discrepancy was briefly mentioned and this chapter will give some insight on how this can be used in model updating.

In recent years, there has been a lot of research in design under uncertainty, which has, in fact, drawn attention in engineering design field [162, 163]. The inverse problem in uncertainty analysis is considered as a hard task, in which experimental datasets are used to learn about modeling of uncertainty sources such as calibration parameters of a computer model and computer model discrepancy, defined as the difference between reality and the simulator output. This topic is receiving increased attention in recent years because quantifying these modeling uncertainties and the resulting system response prediction is essential for reliable design decision making.

Carrying out the uncertainty quantification analysis for inverse problems, which is typically implemented in a model updating process. Kennedy and O'Hagan [164] defined the calibration parameter as any physical parameter that can be specified as an input to the computer model and that is unknown or not measurable when conducting the physical experiment.

Within the existing model updating techniques, the concept of calibration is used to adjust or tune the calibration parameters $\boldsymbol{\theta}$, and an additive discrepancy function $\delta(\mathbf{x})$ is used to present the model discrepancy. The majority of the approaches for model updating consider either calibration only [165, 166] or discrepancy function only [167] and few consider both calibration parameters and discrepancy functions together [164, 168].

The model updating formulation in [164] is one such approach, which incorporates both the calibration parameters and discrepancy function, and it is considered as the most applicable method to tackle the inverse problem analysis under uncertainty in diverse fields, such as ecology [169], health science [170] and climate modelling [171].

6.2 Sources of uncertainty

Following the important work by Kennedy and O'Hagan [164], several different sources of uncertainty are considered for model updating.

(1) **Parameter uncertainty.**

The parameter uncertainty derives from model calibration parameters that are inputs to a computer model but that are unknown and cannot be measured in physical experiments. In addition, the set of calibration parameters can be selected properly based on prior expert knowledge and parameter screening techniques [172].

(2) **Model uncertainty**

The model uncertainty (commonly termed model discrepancy), which results from underlying missing physics, numerical approximations and other inaccuracies of computer models (simulators) even if all the parameters are known.

(3) **Code uncertainty**

The code uncertainty [173] results from having only a finite set of simulations and experiment datasets collected at discrete input settings, in which case, one can make interpolation and regression analysis to predict the response at other input settings.

(4) **Experimental uncertainty**

The experimental uncertainty (commonly termed experimental measurement error) comes from a physical experiment. This represents any variability in the experimental response that would happen if the experiment were repeated many times, even using the same settings for all of the design variables.

Before illustrating the model updating formulation proposed in [164], the other two formulations considered various uncertainties had been reviewed.

The first model updating formulation considers a discrepancy function with no calibration parameters [174], which can be expressed as

$$y^e(\mathbf{x}) = \eta^c(\mathbf{x}) + \delta(\mathbf{x}) + \varepsilon \tag{6.1}$$

where $y^e(\mathbf{x})$ is the physical experiment response, $\eta^c(\mathbf{x})$ is the computer model response, $\delta(\mathbf{x})$ is the discrepancy function and ε accounts for the experimental uncertainty and typically assumed to follow a normal distribution with mean 0

and variance σ^2.

The advantage of only considering the discrepancy function is that a Bayesian closed form solution exists. The obvious limitation of the Eq. (6.1) is that it cannot account for the effects of unknown calibration parameters.

The second model updating formulation incorporates the calibration parameters with no discrepancy functions [175], which is expressed as

$$y^e(\mathbf{x}) = \eta^c(\mathbf{x}, \boldsymbol{\theta}^\bullet) + \varepsilon \tag{6.2}$$

where $\eta^c(\mathbf{x}, \boldsymbol{\theta}^\bullet)$ is the computer model response, $\boldsymbol{\theta}^\bullet$ is the true values of the unknown calibration parameters. Many instances of this approach [166, 175] only consider the calibration parameters and omit $\mathbf{x}$ when fitting the model.

Kennedy and O'Hagan [164] has combined the above two approaches into one model updating formulation, which is expressed as

$$y^e(\mathbf{x}) = \eta^c(\mathbf{x}, \boldsymbol{\theta}^\bullet) + \delta(\mathbf{x}) + \varepsilon \tag{6.3}$$

The discrepancy function $\delta(\mathbf{x})$ in Eq. (6.3) is a function of $\mathbf{x}$. The design inputs $\mathbf{x}$ are variables that can be controlled and set by the user during the experiments. The discrepancy function is defined as the difference between the experimental response $y^e(\mathbf{x})$ and the computer model response $\eta^c(\mathbf{x}, \boldsymbol{\theta}^\bullet)$ at the true values of $\boldsymbol{\theta}^\bullet$.

The code uncertainty and model discrepancy is quantified by the GPE discussed in section 4.3. Because the GPE can give a reliable estimate of their own uncertainty and model an infinite set of random variables with the property that any finite subset has a multivariate normal distribution it be used to model the errors (can be positive or negative values). In order to quantify the uncertainty in the calibration parameters and discrepancy function, the modular Bayesian model updating stages are provided in section 6.3 since the Eq. (6.1) accounts for these uncertainties which is viewed as a comprehensive and applicable model updating approach.

6.3   The modular Bayesian approach

In this section, the modular Bayesian model updating approach in [164] is discussed. The GPE models discussed in section 4.3 are created for both the computer model and the discrepancy function to interpolate the limited datasets from the computer models and physical models.

In Figure 6.1, the GPE formulating process is visualized. It depicts the changes of posterior distributions of the responses with assigned hyperparameters after collecting 4 observations (b), 8 observations (c) and 15 observations (d). This example illustrated an attractive characteristic of the GPE model, which has the ability to provide a reasonable predicted response that interpolates the observed data points (shrinks to 0 at the observed data points).
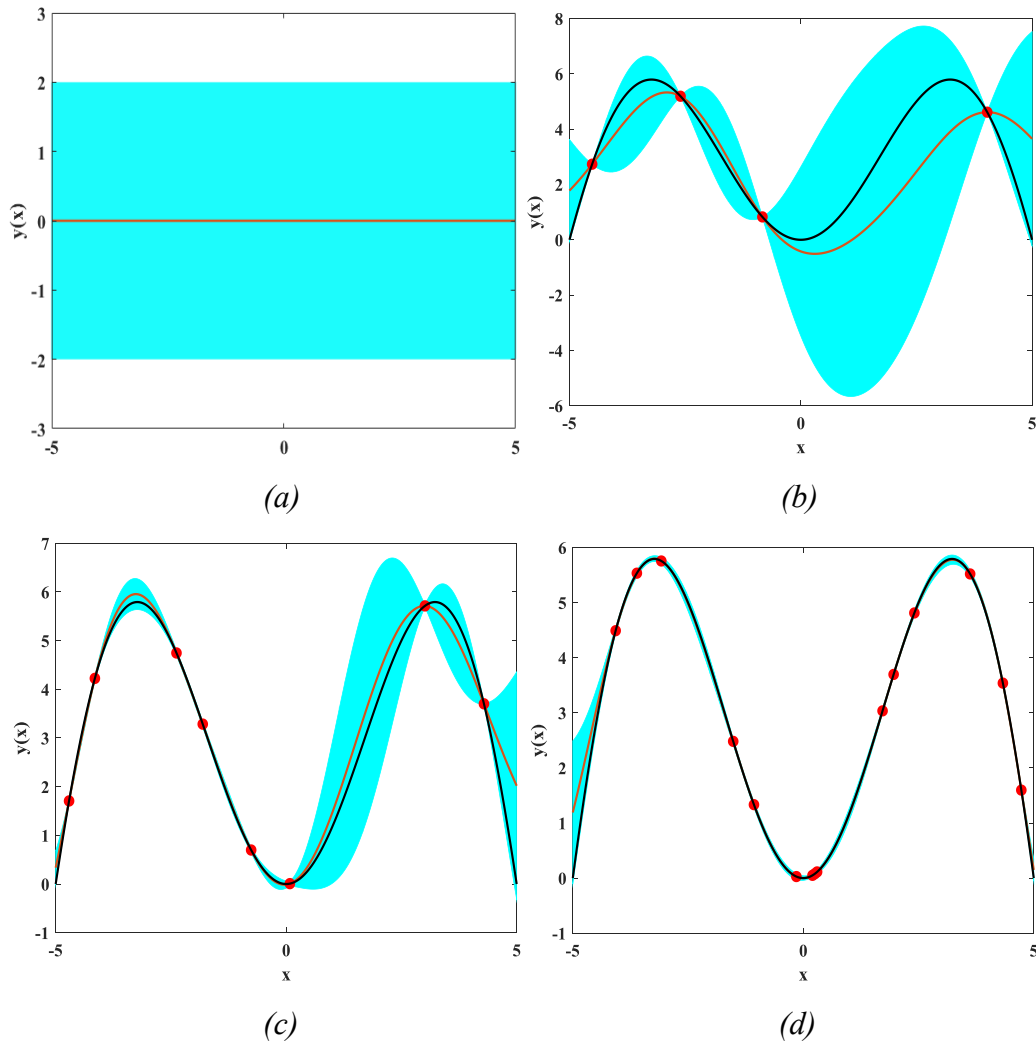
Figure 6.1 *Description by (a) the prior distribution of a GP model with constant mean* 0 *and constant variance. (b), (c) and (d) the predictive posterior distribution for the same GP model after collecting response observations (red dots) with 4,8, and 15. The deep orange solid line is the predicted mean values and the sky blue shaded region indicates a 95% confidence interval.*

The uncertainty in the prediction of $y(\mathbf{x})$, expressed in the form of 95% confidence interval is given as the sky blue shaded regions. The black solid line in

Figure 6.1 is expressed with the objective function $y = 2\sin(2\pi x/10)x$. It is obvious that as the limited observations increased from 4 to 15, the GPE tends to approximate exactly the objective function.

In Figure 6.2, a modular Bayesian approach is given. It separates the estimation of the GP model hyperparameters for both computer model and the discrepancy function (modulus 1 and 2 in Figure 6.2) and the posterior distribution of the calibration parameters, the discrepancy function is calculated from the modular Bayesian approach (modulus 3 in Figure 6.2). the details of each modulus are expressed in the following sections.
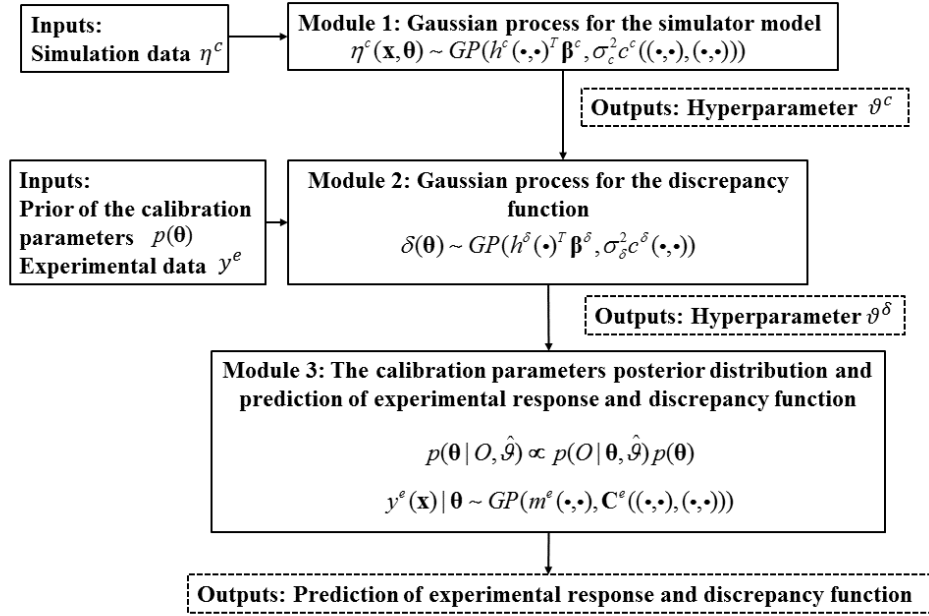
Figure 6.2 *The modular Bayesian approach flowchart.*

6.3.1 Gaussian process for the simulator model (module 1)

A GP is used to infer the computer model output by using the simulation data. The prior of the GP model is defined as

$$\eta^c(\bullet,\bullet) \sim GP(h^c(\bullet,\bullet)^T \boldsymbol{\beta}^c, \sigma_c^2 c^c((\bullet,\bullet),(\bullet,\bullet))) \tag{6.4}$$

where the prior mean function comprised the unknown coefficients $\boldsymbol{\beta}^c$ and the known regression function $h^c(\mathbf{x},\boldsymbol{\theta})^T$. The prior covariance function is the product of an unknown constant σ_c^2 and a correlation function $c^c((\mathbf{x},\boldsymbol{\theta}),(\mathbf{x}',\boldsymbol{\theta}'))$, where

$(\mathbf{x}, \boldsymbol{\theta})$ and $(\mathbf{x}^{'}, \boldsymbol{\theta}^{'})$ denotes two sets of computer model inputs, and the chosen Gaussian correlation function is used. The hyperparameters $\vartheta^c = [\boldsymbol{\beta}^c, \sigma_c^2, \varphi^c]$ can be estimated by maximizing the multivariate normal log-likelihood function for the simulation datasets η^c collected at a limited set of input sites $\mathbf{x}^c$ and $\boldsymbol{\theta}^c$. After obtained the estimates of the hyperparameters of the Gaussian process model for the computer model, the next stage is to estimate the hyperparameters of the Gaussian process model for the discrepancy function.

6.3.2   Gaussian process for the discrepancy function (module 2)

A GP model is adopted for the discrepancy function. The prior is expressed as

$$\delta(\bullet) \sim GP(h^\delta(\bullet)^T \boldsymbol{\beta}^\delta, \sigma_\delta^2 c^\delta(\bullet, \bullet)) \tag{6.5}$$

where the mean function comprised the unknown coefficients $\boldsymbol{\beta}^\delta$ and the known regression function $h^\delta(\mathbf{x})^T$, The covariance function is the product of an unknown constant σ_δ^2 and a correlation function $c^\delta(\mathbf{x}, \mathbf{x}^{'})$. Following the work of Kennedy and O'Hagan [164], we assume the computer model, the discrepancy function and the experimental uncertainty are prior statistically independent. This kind of assumptions simplifies many of the calculations required in the modular Bayesian approach.

As a consequence, the priors for computer model in Eq. (6.4) and the discrepancy function in Eq. (6.5) are combined to form the prior for the experimental response GP, which is expressed as

$$\begin{aligned} y^e(\bullet) \,|\, \boldsymbol{\theta} &\sim GP(m^e(\bullet, \bullet), \mathbf{C}^e((\bullet, \bullet), (\bullet, \bullet))) \\ m^e(\mathbf{x}, \boldsymbol{\theta}) &= h^c(\mathbf{x}, \boldsymbol{\theta})^T \boldsymbol{\beta}^c + h^\delta(\mathbf{x})^T \boldsymbol{\beta}^\delta \\ \mathbf{C}^e((\mathbf{x}, \boldsymbol{\theta}), (\mathbf{x}^{'}, \boldsymbol{\theta}^{'})) &= \sigma_c^2 c^c((\mathbf{x}, \boldsymbol{\theta}), (\mathbf{x}^{'}, \boldsymbol{\theta}^{'})) + \sigma_\delta^2 c^\delta(\mathbf{x}, \mathbf{x}^{'}) + \sigma^2 \end{aligned} \tag{6.6}$$

Since Eq. (6.6) depends on the calibration parameters $\boldsymbol{\theta}$, the problem now is to estimate the hyperparameters of the discrepancy function Gaussian process without knowing the true value of $\boldsymbol{\theta}$.

Kennedy and O'Hagan [164] developed a procedure to obtain the MLEs of the hyperparameters ϑ^δ using only the prior distribution for the calibration parameters. In addition, Kennedy and O'Hagan in [176] also provide a closed form solution for the likelihood function under the conditions of Gaussian correlation functions, constant regression functions.

6.3.3   The calibration parameters posterior distribution (module 3)

This module calculates the posterior distribution of the calibration parameters and predictions of experimental responses and discrepancy function based on the simulations, the experimental datasets, and the estimated hyperparameters from module 1 and 2. The posterior distribution of the calibration parameters can be expressed as

$$p(\boldsymbol{\theta} \mid O, \hat{\vartheta}) \propto p(O \mid \boldsymbol{\theta}, \hat{\vartheta}) p(\boldsymbol{\theta}) \tag{6.7}$$

where $\hat{\vartheta}$ are the MLEs of $\vartheta = [\vartheta^c, \vartheta^\delta]$, O denotes all response observations and $p(\boldsymbol{\theta})$ is the prior of the calibration parameters. The likelihood function $p(O \mid \boldsymbol{\theta}, \hat{\vartheta})$ is multivariate normals. Section 4.4 of [164] gives detailed equations for the likelihood function.

After obtained the simulation and experimental datasets and estimating the hyperparameters in module 1 and 2, given a specific value of $\boldsymbol{\theta}$, the conditional posterior distribution of experimental response can be calculated at any location $\mathbf{x}$, with mean and covariance similar to Eq. (4.13) and Eq. (4.14) (Section 4.6 of [164] for detailed equations). The unconditional distribution of the experimental response is then obtained by marginalizing the conditional distribution with respect to the posterior distribution of the calibration parameters. As a result, the marginalized posterior distribution of the experimental response accounts for parameter uncertainty, model discrepancy, code uncertainty, and experimental uncertainty.

6.4 An illustrative example of a cantilever beam

An illustrative example is presented to conduct the calibration parameter identification by putting the model discrepancy into consideration using the modular Bayesian approach described in section 6.3. The example here is a cantilever beam system shown in Figure 6.3. The 'true' beam structure is presented in Figure 6.3 (b) with a semi-rigid cantilever connection modeled using a rotational spring, which has a stiffness parameter K. The Young's modulus E^* along this beam has a value of $70 \times 10^3 \text{MPa}$. The vertical force applied to its end is $F = 5 \times 10^3 N$.

The idealized cantilever beam

(a)

The true cantilever beam

(b)
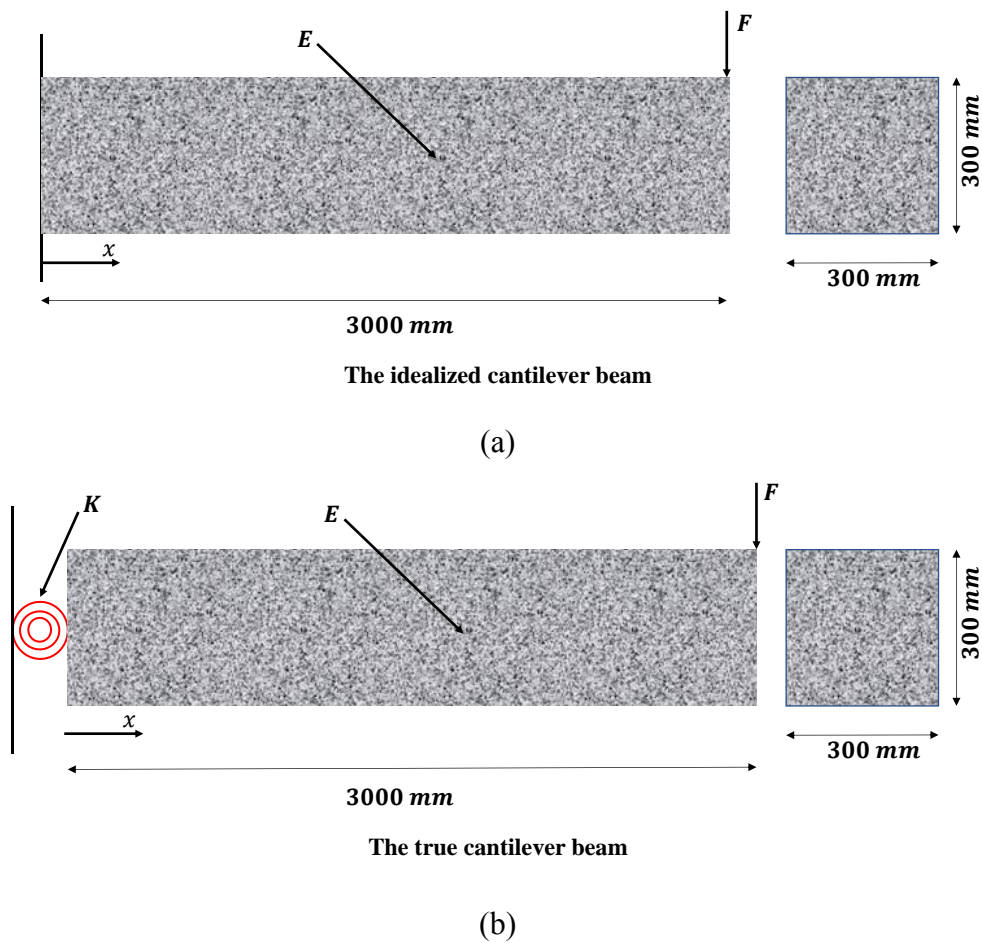
Figure 6.3 *The idealized and true cantilever beams. The calibration parameter to be identified is Young's modulus (E).*

In order to be representative of the full-scale situation where it is not possible to capture reality in a model. An idealized beam is shown in Figure 6.3 (a), which is used to model the 'true' beam system shown in Figure 6.3 (b). This idealized structure does not include the partial rigidity of the cantilever beam connection.

The calibration parameter E, Young's modulus, assigned as the calibration parameter which has a possible range of values $[35,140] \times 10^3 \text{MPa}$ corresponding to the $5\% - 200\%$ of the initial parameter value. To be consistent with the notation used in section 6.3. the parameter to be identified is denoted as $\theta = E$.

The beam is $3000mm$ long and has a square cross-section of $300 \times 300 mm^2$. Its inertia I is $6.75 \times 10^8 mm^4$. The vertical displacement $D(x)$ of the cantilever beam at any location $x \in [0,l] (l = 3000mm)$ is given by

$$D(x) = \frac{Fx^2(3l-x)}{6EI} \tag{6.8}$$

For any location x, the error introduced by the idealized beam model is $\gamma(x) = -\frac{Flx}{K}$. The simulated measured values $y(x)$ are obtained according to

$$y(x) = D^*(x) - \gamma(x) + \varepsilon \tag{6.9}$$

where $D^*(x)$ is the displacement computed with the correct parameter value E^* and ε is a realization of $\varepsilon \sim N(u_e, \sigma_e^2)$, a Gaussian random variable indicates the experimental uncertainty.

The number of measurement n_m is selected starting by $n = 10d$ according to Loeppky, et al. [30], and finally, 25 measurements are chosen in order to have a good interpolation for GPE. For each value of n_m, the displacement is evaluated at $x = x_{start} + i(l - x_{start})/n_m, \forall i \in \{1, \ldots, n_m\}$. By using Eq. (6.9), the minimal distance from the cantilever support designed to have the value of $x_{srart} = 500mm$. In this example, the magnitude of the distance is chosen as the design variable **x** and refer to Table 6.1 for the notation of this example.

Table 6.1 *Notation of the cantilever beam example*

Variable	Description
x	the magnitude of the distance along the beam $[500,3000] mm$
θ	Young's modulus E (10^3MPa) on the range of $[35,140] \times 10^3 \text{MPa}$, $\theta^* = 70 \times 10^3 \text{MPa}$ is the true value
$\eta^c(x,\theta)$	$D(x,\theta)$ calculated using Figure 6.3 (a)
$y^e(x)$	$y(x)$ calculated using Figure 6.3 (b)

ε	Experimental uncertainty with $\varepsilon \sim N(0, 0.02^2)$

The first prior distribution for the calibration parameter is a uniform prior over the range $[35,140] \times 10^3$ MPa. The prior distribution is chosen to be wide in order to minimize the effect of the prior on the posterior distributions. To infer the calibration parameter and the discrepancy function, we use the modular Bayesian approach described in section 6.3. Figure 6.4 (a) and Figure 6.5 (a) show the posterior distribution of discrepancy function and calibration parameter, respectively. In addition, collecting additional 'measured' observations helps little to the performance for the posterior distribution of discrepancy function and calibration parameter.

(a)

(b)

(c)
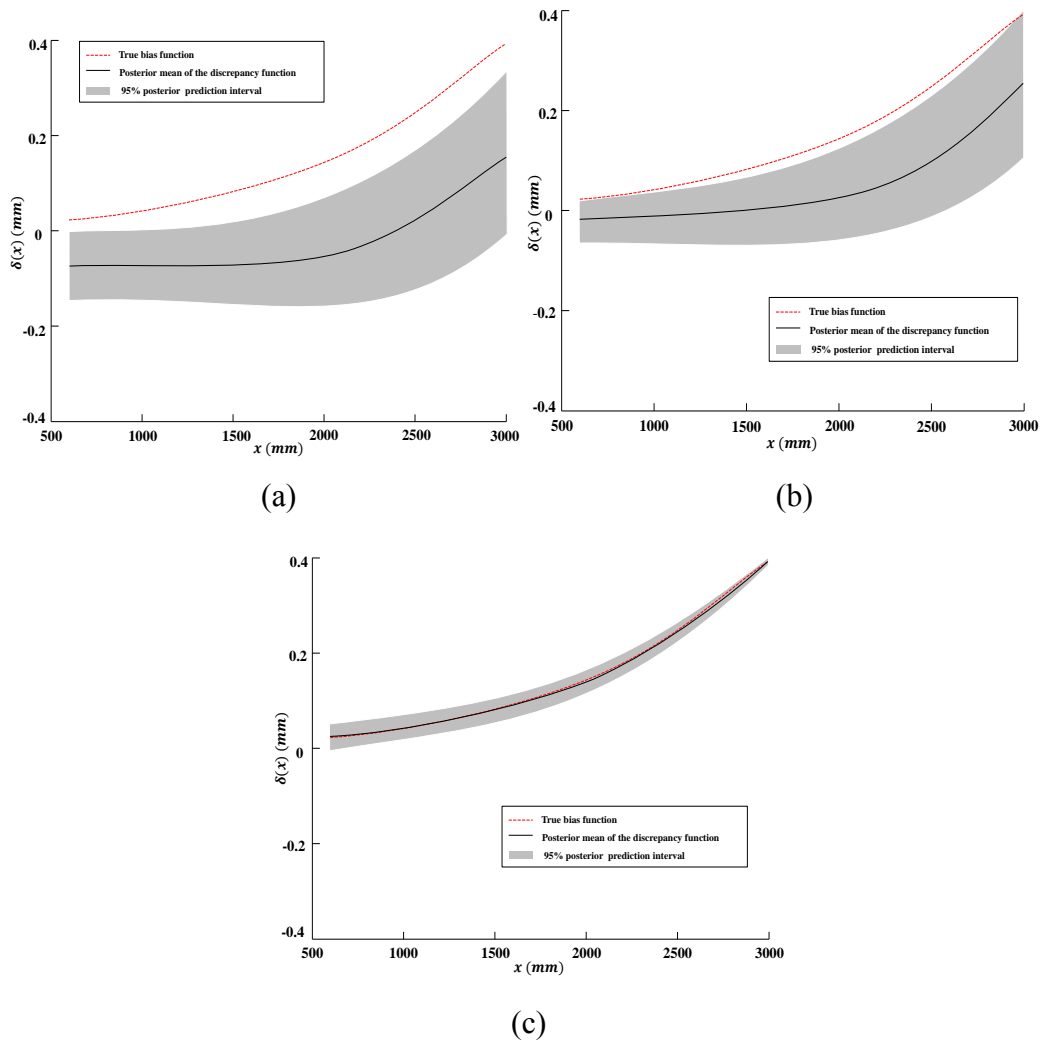
Figure 6.4 *The posterior distribution for the discrepancy function for the three prior distributions of the calibration parameter (a)uniform (b)N(105*10³, 10.33²) (c)N(85*10³, 3.5²).*
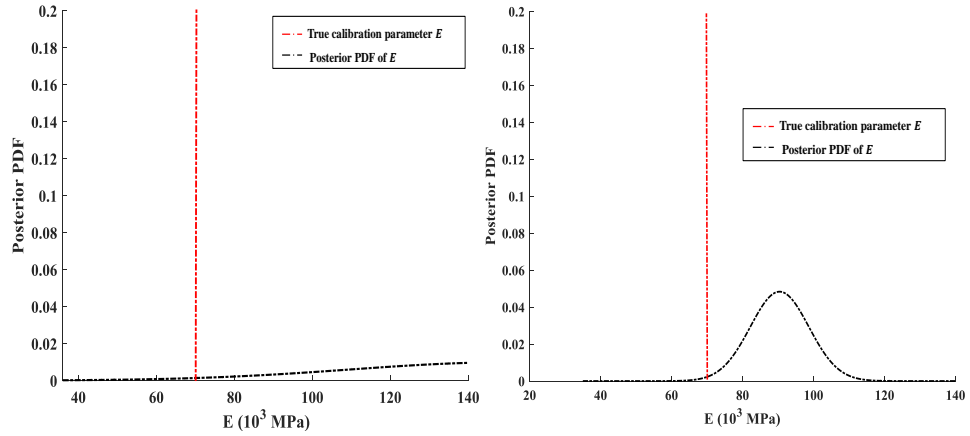
(a) *Posterior distribution*
N(149.78*10³, 39.95²).

(b) *Posterior distribution*
N(90.43*10³, 8.25²).
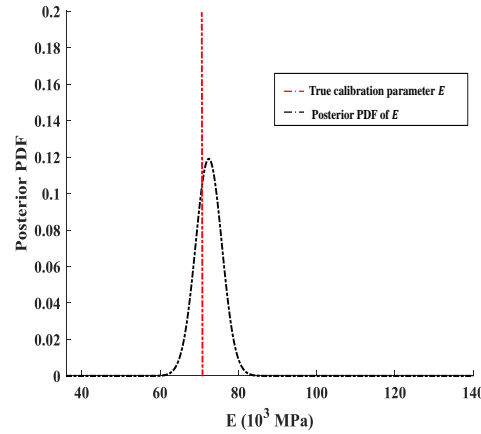
(c) *Posterior distribution*
N(72.05*10³, 3.35²).

Figure 6.5 *Posterior distribution for the calibration parameter* E *for the three prior distributions (a)uniform (b)N(105*10³, 10.33²) (c)N(85*10³, 3.5²).*

In order to improve the identifiability, Liu, et al. [16] suggested using an informative prior distribution for the calibration parameters, or the discrepancy function. To illustrate how this assumption improves the identifiability, more 2 cases are conducted using 2 different normal prior distributions for θ, each with different mean and variance.

The second prior distribution for the calibration parameter assumes a less informative prior distribution with prior mean $105 \times 10^3 \, \text{MPa}$ and prior standard deviation (SD) 10.33. Figure 6.4 (b) and Figure 6.5 (b) show the posterior distribution for the discrepancy function and calibration parameter. The posterior distribution for the calibration parameter E with a mean value $90.43 \times 10^3 \, \text{MPa}$

and SD 8.25. In this case, the posterior distribution for the calibration parameter is not precise, which indicates a lack of accuracy.

The last prior distribution for the calibration parameter is therefore assigned an informative prior distribution (small SD) with prior mean $85 \times 10^3 \text{MPa}$ and prior SD 3.5. Figure 6.4 (c) and Figure 6.5 (c) show the posterior distribution for the discrepancy function and calibration parameter. The posterior distribution for the calibration parameter E with a mean value $72.05 \times 10^3 \text{MPa}$ and SD 3.35. In this case, it is evident that the posterior distribution for the calibration parameter E is more precisely relative to the previous cases, results in an identifiable system. In conclusion, despite the difficulty of quantifying the model discrepancy, choosing informative priors does help to obtain the true posterior distribution of the calibration parameter E in this example.

6.5 Chapter summary

This chapter gives an introduction of model updating framework considering several different sources of uncertainty namely parameter uncertainty, model discrepancy, code uncertainty, and experimental uncertainty, which is proposed by Kennedy and O' Hagan. The GPEs are used to model the simulator model and discrepancy function, respectively.

In order to have a better understanding of this model updating procedure to a cantilever beam. We first show that the calibration parameters E and the discrepancy function are not be identified when given a uniform prior distribution to the calibration parameter E. After that, two other prior distributions to the calibration parameter E are applied, it is found that using an informative prior distribution for the calibration parameter, which can improve the identifiability.

The challenge with incorporating model discrepancy in statistical model updating is being confounded with calibration parameters, which will be solved with informative priors. For our cantilever beam example, the model discrepancy is modeled via GPE, and demonstrate that using the informative priors on the calibration parameters, which helps to obtain the true posterior distribution of the calibration parameter E.

7 Conclusion and future work

7.1 Summary of completed work

This dissertation has presented the work, which focuses on the uncertainty quantification and stochastic model updating for structural dynamics.

In chapter 1, the motivation and background to the scope of the work were provided. In Chapter 2, a literature review of deterministic FE model updating methods including direct methods, iterative methods, and computational intelligence methods were described. The advantages and disadvantages of these methods were also discussed.

In chapter 3, the notion of stochastic FE model updating methods was introduced as the variability inevitably exists in the measured data and had to be considered in the model updating procedure. In practical model updating the measured data were often imprecise, incomplete and variable. Therefore, it was very important to put statistical techniques into consideration to improve updating performance. The stochastic FE model updating methods were generally categorized into two categories: probabilistic methods and non-probabilistic methods. The representative existing stochastic model updating methods: probabilistic perturbation method and the Bayesian model updating method were described in detail. The non-probabilistic stochastic model updating including the interval model updating methods and the fuzzy model updating methods were also briefly discussed. The performance of these two stochastic model updating methods was discussed with illustrative examples.

In chapter 4, the concept of GPE was introduced, together with relevant supporting theory and methodologies. Aspects concern about the GPE construction process, optimizing GPEs and validate their qualities were described and discussed. In addition, the GPEs were employed to model the relations of DIFs in SSRC specimens for compressive strength and Young's modulus from different strain rates in SHPB tests.

Chapter 5 described the importance of employing the computational efficient GPEs to do the stochastic model updating. Several sparse GP approaches were described.

The VSGPE was given in detail as this modeling technique had been used to do the Bayesian model updating to the AIRMOD structure. The Bayesian model updating method combined surrogate modeling techniques of VSGPE, which was built based on the notion of standard GPE, TMCMC method described in chapter 3 and parallel computing strategy to improve the updating process efficiency. In addition, by using two different covariance functions namely SE kernel and Matérn $\left(v = \dfrac{3}{2} \right)$ kernel to do the model updating performance comparison of the AIRMOD structure, which was discussed in detail.

In chapter 6, an introduction of model updating framework considering several different sources of uncertainty namely parameter uncertainty, model discrepancy, code uncertainty, and experimental uncertainty was provided. The comprehensive model updating formulation under uncertainty was expressed and applied to a cantilever beam example using modular Bayesian approach. Some useful conclusions and challenges were provided and discussed.

7.2   Summary of contributions

The main contributions of this dissertation are the following:

For deterministic FE model updating, A novel iterative method formulated based on the classic sensitivity method is proposed. The efficiency and advantages of this method are verified by using two numerical examples, named the three degree of freedom mass-spring system and the six storey bare frame. It is concluded that the proposed iterative method is more accurate and efficient than the traditional sensitivity method especially when the changes for the parameters of the model in need of correction are relatively large. This improved iterative method also enriches the theory of deterministic model updating. In addition, the proposed method is applicable to determined or overdetermined mode updating problems and need to improve to solve undetermined model updating problems (Chapter 2).

The probabilistic perturbation method and Bayesian model updating method are implemented to a three degree of freedom mass-spring system to carry out a performance comparison analysis. It is found that these two methods are all applicable to the problem of model updating in the presence of uncertainty with high accuracy. It is found that the time of conducting Bayesian model updating

when combined with the MH algorithm is much longer than the probabilistic perturbation method (Chapter 3).

The theory and procedure of constructing the GPE considering the noise-free observations and noisy observations respectively are systematically described. Some useful conclusions from quasi-static and SHPB tests for three types of specimens: Concrete, SSRC33, and SSRC34 are given. It is observed that under quasi-static tests, the compressive strength is improved from Concrete to SSRC34. For SHPB tests,a typical stress-strain curves of all specimens under certain strain rate indicates that the maximum compressive stress are largely enhanced by the insertion of steel skeletons compared to the plain Concrete. The GPEs is employed to obtain the relations of DIFs for compressive strength and Young's modulus from different strain rates. The well-trained GPE is validated by individual prediction errors (IPE) diagnostics. In addition, DIFs of the compressive strength and Young's modulus show increasing rate sensitivity to strain rate from Concrete to SSRC34. (Chapter 4)

A Bayesian FE model updating framework combined with high-fidelity, less time-consuming surrogate models called VSGP metamodels is proposed to a complex experimental structure DLR-AIRMOD for model updating. The computational cost for this real-time application of Bayesian model updating is a hard requirement and the problem has been solved effectively by using a parallel computing strategy and the well-constructed VSGP surrogate models together. In addition, the SE covariance function and Matern $\left(v = \dfrac{3}{2} \right)$ covariance functions are employed to implement performance comparison analysis in order to obtain high-fidelity surrogate models to the AIRMOD structure updating. It is found that using Matern $\left(v = \dfrac{3}{2} \right)$ covariance functions to build the VSGPE is more favorable than the SE covariance function (Chapter 5).

By examining the existing most comprehensive model updating formulations considering all sources of uncertainties namely parameter uncertainty, model discrepancy, code uncertainty, and experimental uncertainty to an illustrative

example of a cantilever beam and using GPEs to model the simulator model and discrepancy function, respectively. It is found that using an informative prior distribution for the calibration parameter E, which helps to obtain the true posterior distribution of the calibration parameter E (Chapter 6).

7.3 Research outlook

The work conducted in this dissertation leads to several research directions, which can be considered in the future.

Firstly, implementing the proposed iterative method to do model updating for industrial complex structures to assess its performance since the examples used to demonstrate in the thesis are simplified numerical models.

Secondly, considering other kinds of arrangements of steel skeletons in order to have a more comprehensive understanding of the dynamic mechanical behaviors of this composite material under uncertainties. More experimental tests are suggested to conduct in the future to have more validation points to validate the trained GPE to do uncertainty quantification analysis in SHPB tests.

Thirdly, the Bayesian FE model updating framework is employed to do the stochastic model updating to the AIRMOD structure in success. However, the frequency distributions especially f_{14}, f_{19} and f_{21} from the test is still not replicated closely enough by the results of Bayesian model updating. The reason to cause this discrepancy is that by constructing VSGPE with multiple inputs and only one output is not appropriate, especially for f_{14}, f_{19} and f_{21}. Considering these three frequency responses correlated to build the VSGPE for stochastic model updating is suggested in the future.

Finally, research and develop the model updating formulation in chapter 6 to put all sources of uncertainty into consideration to do model updating is important since the analysis of uncertainty in the subject of structural dynamics in model updating is a challenging problem. As long as we have the developed model updating formulation, exploring the uncertainty of structural parameters (calibration parameters) under model discrepancy in order to have better parametric identifiability can be investigated and analyzed.

7.4 List of publications

Journal article:

- Q.L. Wang, F.A. DiazDelaO, Experimental investigation and Gaussian process emulation of steel skeleton reinforced concrete behaviors in split Hopkinson pressure bar tests, International Journal of Structural and Civil Engineering Research. (Accepted)

- Q.L. Wang, F.A. DiazDelaO, A.K. Gray, J.E. Mottershead, E. Patelli, Stochastic finite element model updating with variational sparse Gaussian processes using the DLR AIRMOD data, Mechanical Systems and Signal Processing. (Under preparation)

Conference paper:

- Q.L. Wang, F.A. DiazDelaO, A.K. Gray, J.E. Mottershead, E. Patelli, Stochastic finite element model updating with variational sparse Gaussian processes, *Proceedings of the 3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering, UNCECOMP 2019, 24-26 June, Crete, Greece.*

Appendix A

Gaussian identities

A.1 Multivariate Gaussian distribution

A random vector $\mathbf{x}$ is said to have a normal multivariate distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$ if and only if

$$p(\mathbf{x}) = N(\mathbf{x} \mid \boldsymbol{\mu}, \mathbf{C}) = \mid 2\pi\mathbf{C} \mid^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$$

A.2 Marginal and conditional distributions

Given a joint Gaussian multivariate distribution

$$p(\mathbf{x}_1, \mathbf{x}_2) = N(\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \mid \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{B} \end{bmatrix})$$

the marginal distribution of $\mathbf{x}_1$ is

$$p(\mathbf{x}_1) = N(\mathbf{x}_1 \mid \boldsymbol{\mu}_1, \mathbf{A})$$

and the distribution of $\mathbf{x}_1$ conditioned on $\mathbf{x}_2$ is

$$p(\mathbf{x}_1 \mid \mathbf{x}_2) = N(\mathbf{x}_1 \mid \boldsymbol{\mu}_1 + \mathbf{C}\mathbf{B}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \mathbf{A} - \mathbf{C}\mathbf{B}^{-1}\mathbf{C}^{\mathbf{T}})$$

A.3 Integral of the product of two Gaussians

The following identity

$$\int N(\mathbf{x} \mid \mathbf{a}, \mathbf{A})N(\mathbf{a} \mid \mathbf{b}, \mathbf{B})d\mathbf{a} = N(\mathbf{x} \mid \mathbf{b}, \mathbf{A} + \mathbf{b})$$

is useful to compute marginal likelihood $p(\mathbf{x})$ from likelihood $p(\mathbf{x}|\mathbf{a})$ and prior $p(\mathbf{a})$, when both are Gaussian distributions.

A.4 Gaussian likelihood with linear parameter

If the mean of a Gaussian depends linearly on some parameter with Gaussian prior, it is possible to obtain and posterior probabilities in closed form. Assume the

following parameterization for the likelihood and the prior

$$p(\mathbf{y} \mid \mathbf{w}, \mathbf{x}) = N(\mathbf{y} \mid \mathbf{xw}, \mathbf{C}_y)$$
$$p(\mathbf{w}) = N(\mathbf{w} \mid \mathbf{a}, \mathbf{C}_w)$$

The joint distribution is then $p(\mathbf{y}, \mathbf{w} \mid \mathbf{x}) = p(\mathbf{y} \mid \mathbf{w}, \mathbf{x}) p(\mathbf{w})$. This product can be conveniently arranged as

$$p(\mathbf{y}, \mathbf{w} \mid \mathbf{x}) = N\left(\begin{bmatrix} \mathbf{y} \\ \mathbf{w} \end{bmatrix} \middle| \begin{bmatrix} \mathbf{xa} \\ \mathbf{a} \end{bmatrix}, \begin{bmatrix} \mathbf{C}_y + \mathbf{x}\mathbf{C}_w\mathbf{x}^{\mathrm{T}} & \mathbf{x}\mathbf{C}_w \\ \mathbf{C}_w\mathbf{x}^{\mathrm{T}} & \mathbf{C}_w \end{bmatrix} \right)$$

The marginal likelihood of $\mathbf{y}$ is contained in the previous expression due to the marginalization property of Gaussian distribution (A. 2). The posterior distribution over $\mathbf{w}$ given $\mathbf{y}$ and $\mathbf{x}$ is obtained applying (A. 3) to (A. 5)

$$p(\mathbf{w} \mid \mathbf{y}, \mathbf{x}) = N(\mathbf{w} \mid \boldsymbol{\mu}_{w|y}, \mathbf{C}_{w|y})$$
$$\boldsymbol{\mu}_{w|y} = \mathbf{a} + \mathbf{C}_w\mathbf{x}^{\mathrm{T}}(\mathbf{C}_y + \mathbf{x}\mathbf{C}_w\mathbf{x}^{\mathrm{T}})^{-1}(\mathbf{y} - \mathbf{xa})$$
$$\mathbf{C}_{w|y} = (\mathbf{C}_w^{-1} + \mathbf{x}^{\mathrm{T}}\mathbf{C}_y^{-1}\mathbf{x})^{-1}$$

A.5 Linear transformations

If $\mathbf{x}$ is a random vector following a multivariate Gaussian distribution

$$p(\mathbf{x}) = N(\mathbf{x} \mid \boldsymbol{\mu}, \mathbf{C})$$

The linear transformation $\mathbf{y} = \mathbf{R}^{\mathrm{T}}\mathbf{x} + \mathbf{r}$ is also Gaussian distribution

$$p(\mathbf{y}) = N(\mathbf{y} \mid \mathbf{R}^{\mathrm{T}}\boldsymbol{\mu} + \mathbf{r}, \mathbf{R}^{\mathrm{T}}\mathbf{C}\mathbf{R})$$

Appendix B

Kullback-Leibler Divergence

The Kullback-Leibler (KL) divergence between two distributions $p(\mathbf{x})$ and $q(\mathbf{x})$ is defined as

$$KL(p \| q) = \int p(\mathbf{x}) \log \frac{p(\mathbf{x})}{q(\mathbf{x})} d\mathbf{x}$$

It is easy to show that $KL(p \| q) \geq 0$, with equality if $p = q$.

Appendix C

Variational inference

In variational inference, the posterior distribution over a set of unobserved variables $\mathbf{M} = \{M_1, \ldots, M_n\}$ given some data $\mathbf{x}$ is approximated by a variational distribution $Q(\mathbf{M})$

$$p(\mathbf{M} \mid \mathbf{x}) \approx Q(\mathbf{M})$$

The distribution $Q(\mathbf{M})$ is restricted to belong to a family of distributions of simpler form than $p(\mathbf{M} \mid \mathbf{x})$, selected with the intention of making $Q(\mathbf{M})$ similar to the true posterior $p(\mathbf{M} \mid \mathbf{x})$.

The lack of similarity is measured in terms of a dissimilarity function $d(Q; p)$ and hence inference is performed by selecting the distribution $Q(\mathbf{M})$ that minimizes $d(Q; p)$.

Appendix D

Jensen's inequality

Jensen's inequality, named after the Danish mathematician Johan Jensen, relates the value of a convex function of an integral to the integral of the convex function. When making the form involving a probability density function, suppose Ω is a measurable subset of the real line and $f(x)$ is a non-negative function such that

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

f is a probabilistic density function.

Jensen's inequality becomes the following statements:

If g is any real-valued measurable function and φ is convex over the range of g, then

$$\varphi(\int_{-\infty}^{\infty} g(x)f(x)dx) \leq \int_{-\infty}^{\infty} \varphi(g(x))f(x)dx$$

If $g(x) = x$, then this form of inequality reduced to a commonly used special case:

$$\varphi(\int_{-\infty}^{\infty} xf(x)dx) \leq \int_{-\infty}^{\infty} \varphi(x)f(x)dx$$

Bibliography

[1] O.C. Zienkiewicz and R.L. Taylor. The finite element method, Volume 1, the basis. Butterworth Heinemann, 2000.

[2] M.I. Friswell and J.E. Mottershead. Finite element model updating in structural dynamics. Kluwer Academic Press, 1995.

[3] J. E. Mottershead and M. I. Friswell. Model updating in structural dynamics: a survey. Journal of Sound and Vibration, 167(2):347–375, 1993.

[4] H. Ahmadian, J.E. Mottershead, and M.I. Friswell. Regularisation methods for finite element model updating. Mechanical Systems and Signal Processing, 12:47–64, 1998.

[5] D.J. Ewins. Modal testing: theory and practice. Research Studies Press LTD, Letchworth, Hertfordshire, England, 1984.

[6] W. Heylen, S. Lammens, and P. Sas. Modal analysis theory and testing. KUL Press, Leuven, 1997.

[7] R.J. Allemang and D.L. Brown. A correlation coefficient for modal vector analysis. In Proceedings of the 1st International Modal Analysis Conference (IMAC I), Orlando, Florida, U.S.A, 1982.

[8] W. Heylen, P. Sas (1987). Review of model optimization techniques. Proceedings of the 5th International Modal Analysis Conferences, Schenectady, NY: Union College, pp. 1177-1182.

[9] M. Baruch, IY Itzhack Bar (1978). Optimum weighted orthogonalization of measured modes. American Institute of Aeronautics and Astronautics Journal 16: 1208-1210.

[10] A. M. Kabe (1985). Stiffness matrix adjustment using mode data. American Institute of Aeronautics and Astronautics Journal 23: 1431-1436.

[11] D. C. Zimmerman, M. Kaouk, T. Simmermacher (1995). Structural damage using frequency response functions. Proceedings of the 13th International Modal Analysis Conference, Bethel, CT: SPIE, pp. 179-184.

[12] D. C. Zimmerman, M. Kaouk (1992). Eigenstructure assignment approach for structural damage detection. American Institute of Aeronautics and Astronautics Journal 30: 1848-1855.

[13] R. J. Guyan (1965). Reduction of stiffness and mass matrices. American Institute of Aeronautics and Astronautics Journal 3: 380.

[14] H. Gysin (1990). Comparison of expansion methods for FE model localization. Proceedings of the 8th International Modal Analysis Conferences, Schenectady, NY: Union College, pp. 195-204.

[15] M. Paz (1984). Dynamic condensation. American Institute of Aeronautics and Astronautics Journal 22: 724-727.

[16] Wittrick, W.H. 1962. "Rates of Change of Eigenvalues, with Reference to Buckling and Vibration Problems" Journal of the Royal Aeronautical Society, 66, 590-591.

[17] J.E. Mottershead, M. Link, M.I. Friswell, The sensitivity method in finite element model updating: A tutorial. Mechanical Systems and Signal Processing. 25(7), 2275-2296 (2011).

[18] M. Link, Updating of analytical models-procedures and experience, in: J.L, Wearing (Ed). Proceedings of the Conference on Modern Practice in Stress and Vibration Analysis, Sheffield Academic Press. (1993) 35-52.

[19] P.C. Hansen, Regularization tools: A MATLAB package for analysis and solution of discrete ill-posed problems, Numerical Algorithms 6 (1994) 1-35.

[20] B. Titurus, M.I. Friswell, Regularization in model updating, International Journal for Numerical Methods in Engineering. 75 (2008) 440-478.

[21] Y. Liu, Z. Duan, H. Liu, Updating of finite element model in considering mode errors with fuzzy theory. Key Engineering Materials, (2006) 413-414: 785-792.

[22] D.S. Jung, C.Y. Kim, FE model updating based on hybrid genetic algorithm and its verification on numerical bridge model. Structural Engineering and Mechanics (2009) 32: 667-683.

[23] D. Tan, W. Qu, J Wang, The finite element model updating of structure based on wavelet package analysis and support vector machines, Journal of Huazhong University of Science and Technology (Natural Science Edition) (2009) 37: 104-107.

[24] J.L. Zapico, A. Gonzlez-Buelga, M.P. Gonzlez, R. Alonso, Finite element model updating of a small steel frame using neural networks. Smart Materials and Structures (2008) 17, no. 045016.

[25] Z. Tu, Y. Lu, FE model updating using artificial boundary conditions with genetic algorithms. Computers and Structures (2008) 86: 714-727.

[26] F.Z. Feng, Y.H. Kim, B.S. Yang, Application of hybrid optimization techniques for model updating of rotor shafts. Structural and Multidisciplinary Optimisation (2006) 32: 65-75.

[27] H.X. He, W.M. Yan, Z. Wang, Stepwise model updating method based on substructures and GA-ANN. Engineering Mechanics (2008) 25: 99-105

[28] T. Marwala Finite Element Model Updating Using Computational Intelligence Techniques. (2010) London: Spring-Verlag.

[29] R. Fox, M. Kooper, Rate of change of eigenvalues and eigenvectors, AIAA Journal, 14(9) (1976) 1201-1205.

[30] R. Nelson, Simplified calculation of eigenvector derivatives, AIAA Journal, 14(9) (1976) 1201-1205.

[31] A.P. Wierzbicki, 1986: A methodological approach to comparing parametric characteristics of efficient solutions. Large-Scale Modeling and Interactive Decision Analysis, Lecture Notes in Economics and Mathematical Systems, No. 273, pp. 27-45. Berlin: Spring-Verlag.

[32] K. Miettinen, 1999, Nonlinear Multiobjective Optimization. Boston: Kluwer Academic Publishers.

[33] I. Das, J.E. Dennis, 1997, A closer look at drawbacks of minimization weighted sums of objectives for Pareto set generation in multicriteria optimization problems. Struct. Optim. 14, 63-69.

[34] A. Messac, C.P. Sukam, Melachrinoudis, E 2000a, Aggregate objective functions and Pareto frontiers: required relationships and practical implications. Optim. Eng. 1, 171-188.

[35] M.J. Rentmeesters, W.K. Tsai, K.J. Lin, 1996, A theory of lexicographic multi-criteria optimization. In: Second IEEE International Conference on Engineering of Complex Computer Systems, pp 76-79.

[36] Gembicki, F.W., "Vector Optimization for control with Performance and Parameter Sensitivity Indices" Ph.D. Dissertation, Case Western Reserve Univ., Cleveland, OH, 1974.

[37] R.K. Brayton, S. W. Director, G. D. Hachtel, and L. Vidigal, "A New algorithm for statistical Circuit Design Based on Quasi-Newton Methods and Function Splitting," IEEE Transactions on Circuits and systems, Vol, CAS-26, pp 784-794, Sep. 1979.

[38] A.C.W. Grace, "Computer-Aided Control System Design Using Optimization Techniques," Ph.D Thesis, University of Wales, Bangor, Gwynedd, UK, 1989.

[39] F.O. Hoffman, J.S. Hammonds, Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Anal 1994; 14:707-12.

[40] C. Mares, J.E. Mottershead, and M.I. Friswell. Stochastic model updating: Part 1: Theory and simulated example. Mechanical Systems and Signal Processing, 20(7):1674–1695, 2006.

[41] X.G. Hua, Y.Q. Ni, Z.Q. Chen, and J.M. Ko. An improved perturbation method for stochastic finite element model updating. International Journal for Numerical Methods in Engineering, 73(13):1845–1864, 2008.

[42] Y. Govers and M. Link. Stochastic model updating: Covariance matrix adjustment from uncertain experimental modal data. Mechanical Systems and Signal Processing, 24(3):696–706, 2010.

[43] J.D. Collins, G.C. Hart, T.K. Hasselman, and B. Kennedy. Statistical identification of structures. AIAA Journal, 12(2):185–190, 1974.

[44] X.G. Hua, Y.Q. Ni, Z.Q. Chen, J.M. Ko, An improved perturbation method for stochastic finite element model updating, International Journal for Numerical Methods in Engineering, 73 (13) (2008) 1845-1864.

[45] H. H. Khodaparast, J.E. Mottershead, M.I. Friswell, Perturbation methods for the estimation of parameter variability in stochastic model updating, Mechanical Systems and Signal Processing, 22 (8) (2008) 1751-1773.

[46] T. AN Silva, M.M.M Nuno, M. Link, J.E. Mottershead, Parameter selection and covariance updating, Mechanical Systems and Signal Processing, 70 (2016) 269-283.

[47] J.E. Mottershead, M. Link, M.I. Friswell, The sensitivity method in finite element model updating: A tutorial, Mechanical Systems and Signal Processing, 25(7):2275 - 2296, 2011.

[48] T. Bayes, R. Price, An Essay towards solving a Problem in the Doctrine of Chance. By the 489 late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. 490 (PDF). Philos. Trans. R. Soc. Lond. 53:370-418 (1763).

[49] H. Jeffreys, Theory of Probability (3rd Edition). Oxford Clarendon Press, Oxford, UK, 1961.

[50] R.T. Cox, The Algebra of Probable Inference, John Hopkins Press, Baltimore, MA, 1961.

[51] E.T. Jaynes. Bayesian methods: General background. In Maximum Entropy and Bayesian Methods in Applied Statistics, Cambridge University Press, Cambridge, UK (1986), pp. 1-25.

[52] A. Zellner, An introduction to Bayesian Inference in Econometrics. John Wiley & Sons Ltd, Inc., New York, NY, 1996.

[53] H. Jiang, and L. Deng, A Bayesian approach to the verification problem: applications to the speaker verification. IEEE Transactions on Speech and Audio Processing. 9(8) (2001), 874-884.

[54] S.T. Rachev, J.S.J. Hsu, B.S. Bagasheva, F.J. Fabozzi. Bayesian Methods in Finance. John Wiley & Sons, Inc., New York, NY, 2008.

[55] J. Beck, L. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, Journal of Engineering Mechanics, 124 (4) (1998) 455-461.

[56] L. Katafygiotis, J. Beck, Updating models and their uncertainties. II: model identifiability, Journal of Engineering Mechanics, 124 (4) (1998) 463-467.

[57] K. -V. Yuen, Bayesian Methods for Structural Dynamics and Civil Engineering. Wiley, Singapore (2010).

[58] I. Boulkaibet, T. Marwala, L. Mthembu, M.I. Friswell, S. Adhikari (2015) Finite element model updating using the shadow hybrid Monte Carlo technique. Mechanical Systems and Signal Processing, 52: 115-132.

[59] E. Jaynes, Prior probabilities, IEEE Transactions on Systems Scicence and Cybernetics, sec-4 (3) (1968) 227-241.

[60] J.O. Berger, J.M. Bernardo, D. Sun, The formal definition of references priors, Annals of Statistics, 37 (2), 2009, 905-938.

[61] P. Diaconis, D. Ylvisaker, Conjugate priors for exponential families, Annals of Statistics, 7 (2), 1979, 269-281

[62] .C. Soize, Construction of probability distributions in high dimensions using the maximum entropy principle: applications to stochastic processes, random fields and random matrices, Internatianal Journal for Numerical Methods in Engineering, 75 (2008) 1583-1611.

[63] C. Soize, Maximum entropy approach of modeling random uncertainties in a transient elastodynamics. The Journal of the Acoustical Society of America, 109 (5), 1979-1996 (2001).

[64] J.L. Beck, Bayesian system identification based on probability logic. Structural Control Health Monitoring. 2010, 17, 825–847. (22)

[65] D. Gamerman, Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Chapman & Hall, London, 1997.

[66] K. Mosegaard, A. Tarantola, Monte Carlo sampling of solutions to inverse problems, Journal of Geophysical Research, 100, 1995, 12431-12447.

[67] H. Haario, E. Saksman, J. Tamminen, An adaptive Metropolis algorithm, Bernouilli 7 (2), 2001, 223-242.

[68] H. Haario, M. Laine, A. Mira, E. Saksman, DRAM: efficient adaptive MCMC, Statistics and Computing. 16 (4), 2006, 339-354.

[69] S. Au, C. Papadimitriou, J. Beck, Reliability of uncertain dynamical systems with multiple design points, Structural Safety, 21, 1999, 113-133.

[70] J Ching, Y.-C. Chen, Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class Selection, and Model Averaging. Journal of Engineering Mechanics, ASCE 2007, 133, 816–832.

[71] B. Calderhead, A general construction for parallelizing Metropolis-Hastings algorithms. Proceedings of the National Academy of Science of the United States of America, 2014, 111: 17408-17413.

[72] J.L. Beck, S.-K. Au, Bayesian Updating of Structural Models and Reliability Using Markov Chain Monte Carlo Simulation. Journal of Engineering Mechanics, ASCE 2002, 128, 380–391.

[73] T.G. Chondros. Distinguished Figures in Mechanism and Machine Science. Springer Netherlands, 2007.

[74] S. Gabriele, C. Valente, An interval-based technique for FE model updating, International Journal of Reliability, Quality and Safety Engineering. 3, 2009, 79-103

[75] H. H. Khodaparast, J.E. Mottershead, K.J. Badcock, Interval model updating with irreducible uncertainty using Kriging predictor, Mechanical Systems and Signal Processing, 25 (4) (2011) 1204-1226

[76] Y. Ben-Haim, I. Elishakoff, Convex models of uncertainty in applied mechanics, Elsevier Science, Amsterdam, 1990.

[77] L.A. Zadeh, Fuzzy sets, Information and Control, 8 (3), 1965, 338-353.

[78] Y.S. Erdogan, P.G. Bakir, Inverse propagation of uncertainties in finite element model 477 updating through use of fuzzy arithmetic, Engineering Applications of Artificial Intelligence, 26 (1) (2013) 357-367.

[79] Y.S. Erdogan, M. Gul, F.N. Catbas, P.G. Bakir, Investigation of uncertainty changes in model outputs for finite-element model updating using structural health monitoring data, Journal of Structural Engineering, 140 (11) (2014) 04014078.

[80] H. H. Khodaparast, Y. Govers, I. Dayyani, S. Adhikari, M, Link, M.I. Friswell, J.E. Mottershead, J. Sienz, Fuzzy finite element model updating of the DLR AIRMOD test structure, Applied Mathematical Modelling, 52 (2017) 512–526.

[81] A. O'Hagan, Bayesian analysis of computer code outputs: A tutorial, Reliability Engineering and System Safety 91 (10-11) (2006) 1290-1300.

[82] A. Forrester, A. Sobester, A. Keane, Engineering Design via Surrogate Modelling: A Practical Guide, 1st Edition, John Wiley & Sons, 2008.

[83] G. E. P. Box, K. B. Wilson, On the experimental attainment of optimum conditions, Journal of the Royal Statistical Society. Series B 13 (1) (1951) 1-45.

[84] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables, Applied mathematics series, Dover Publications, 1965.

[85] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural networks 2 (5) (1989) 359-366.

[86] M. J. D. Powell, in: J. C. Mason, M. G. Cox (Eds.), Algorithms for Approximation, Clarendon Press, 1987, Ch. Radial Basis Functions for Multivariable Interpolation: A Review, pp. 143-167.

[87] V. Vapnik, Statistical Learning Theory, John Wiley & Sons, 1998.

[88] A. O'Hagan, Polynomial Chaos: A Tutorial and Critique from a Statisticianis Perspective (2011).

[89] J. E. Oakley and A. O'Hagan. Probabilistic sensitivity analysis of complex models: a Bayesian approach. Journal of the Royal Statistical Society. Series B (Statistical Methodology), (3):751-770, 2004.

[90] A. Papoulis. Probability, Random Variables, and Stochastic Processes. McGraw-Hill, New York, third edition, 1991.

[91] N. Wiener. Extrapolation, Interpolation and Smoothing of Stationary Time Series. MIT Press, 1949.

[92] G. Matheron. The intrinsic random functions and their applications. Advances in Applied Probability, 5:439–468, 1973.

[93] A. O'Hagan. Curve fitting and optimal design for prediction. Journal of the Royal Statistical Society B, 40:1–42, 1978.

[94] C. K. I. Williams and C. E. Rasmussen. Gaussian processes for regression. In Advances in Neural Information Processing Systems 8. MIT Press, 1996.

[95] C. M. Bishop. Neural Networks for Pattern Recognition. Clarendon Press, Oxford, 1995.

[96] D. J. C. MacKay. Information Theory, Inference, and Learning Algorithms. Cambridge University Press, 2003.

[97] H. Jeffreys. An invariant form for the prior probability in estimation problems. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, pages 453{461, 1946.

[98] I. J. Myung and D. J. Navarro. Information matrix. Encyclopedia of Statistics in Behavioral Science, 2005.

[99] J. Oakley, Bayesian uncertainty analysis for complex computer codes, Ph.D. thesis, University of Sheffield (1999).

[100] W. Becker, Uncertainty Propagation through Large Nonlinear Models, Ph.D. thesis, University of Sheffield (2011).

[101] A.E. Hoerl and R.W. Kennard (1970), Ridge regression: biased estimation for nonorthogonal problems. Technometrics, 12(1), 55-67, February.

[102] C. M. Bishop, Pattern recognition and machine learning., Information science and statistics, New York: Springer, 2006.

[103] A. Forrester, A. Sobester, A. Keane, Engineering Design via Surrogate Modelling: A Practical Guide, 1st Edition, John Wiley & Sons, 2008.

[104] J. A. Nelder, R. Mead, A simplex method for function minimization, The Computer Journal 7 (4) (1965) 308-313.

[105] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, Optimization by simulated annealing, Science 220 (4598) (1983) 671-680.

[106] D. Whitley, A genetic algorithm tutorial, Statistics and Computing 4 (2).

[107] T. Santner, B. Williams, W. Notz, The Design and Analysis of Computer Experiments, Springer Series in Statistics, Springer New York, 2010.

[108] L. Pronzato, W. G. Muller, Design of computer experiments: space filling and beyond, Statistics and Computing 22 (3) (2012) 681-701.

[109] M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, (2): 239-245, 1979.

[110] J. L. Loeppky, J. Sacks, and W. J. Welch. Choosing the sample size of a computer experiment: a practical guide. Technometrics, (4):366-378, 2009.

[111] L. S. Bastos, A. O'Hagan, Diagnostics for Gaussian Process Emulators, Technometrics 51 (4) (2009) 425-438.

[112] B.G. Han, X. Yu, J.P. Ou, in: Self-Sensing Concrete in Smart Structures, Publisher: Elsevier, 2014, p. 385.

[113] B.G. Han, L.Q. Zhang, J.P. Ou, Smart and Multifunctional Concrete Toward Sustainable Infrastructures, Publisher: Springer, 2017, p. 400.

[114] D.A. Abrams, Effect of rate of application of load on the compressive strength of concrete [J]. Journal of ASTM International. 1917, 17(2): 364-377.

[115] B. Bresler, V.V. Bertero, Influence of high strain rate and cyclic loading of unconfined and confined concrete in compression. Proceedings of Second Canadian Conference on Earthquake Engineering. Hamilton, Ontario, 1975. P, 1-13.

[116] B.P. Hughe, R. Gregory, Concrete subjected to high rates of loading in compression [J]. Acta Mechanic, 1972, 24(78): 25-36.

[117] D. Watstein, Effect of straining rate on the compressive concrete [J]. ACI Journal. 1953, 49(8): 729-744.

[118] L.E. Malvern, D.A. Jenkins, T.X. Tang, et al. Dynamic compressive testing of concrete [A]. Proceedings of 2nd Symposium on the Interaction of Non-nuclear Munitions with Structures [C]. Florida：Us Department of Defense, 1985, 194-199.

[119] P.H. Bischoff, S. H. Perry. Compressive behavior of concrete at high strain rates [J]. Materials and Structures. 1991, 144(24): 425-450.

[120] J.P. Roumaldi, G.B. Batson. Mechanics of crack arrest in concrete. Journal of the Engineering Mechanics Division, 1963;89(EM3):147-68.

[121] W. Suaris, S.P. Shah, Strain-rate effects in fiber-reinforced concrete subjected to impact and impulsive loading. Composites 1982:153-9.

[122] C.V. Srinivasa, A. Arifulla, N. Goutham, T. Santhosh, H.J. Jaeethendra, R.B. Ravikumar, et al. Static bending and impact behavior of areca fibers composites. Materials and Design, 2011; 32:2469-75.

[123] T.S. Lok, P.J. Zhao, G. Lu. Using the split Hopkinson pressure bar to investigate the dynamic behavior of SFRC. Magazine of Concrete Research 2003:183-191.

[124] A.M. Brandt. Fiber reinforced cement-based (FRC) composites after over 40 years of development in building and civil engineering. Composite Structures 2008; 86:3-9.

[125] Z. Xu, H. Hao, H.N. Li. Experimental study of dynamic compressive properties of fibre reinforced concrete material with different fibres. Materials and Design 2012; 33:42-55.

[126] X. Sun, K. Zhao, Y. Li, R. Huang, Z. Ye, Y. Zhang, J. Ma, A study of strain-rate effect and fiber reinforcement effect on dynamic behavior of steel fiber reinforced concrete, Construction and Building Materials, 158 (2018) 657–669.

[127] X. Hou, S. Cao, Q. Rong, W. Zheng, G. Li, Effects of steel fiber and strain rate on the dynamic compressive stress-strain relationship in reactive powder concrete, Construction and Building Materials, 170 (2018) 570–581.

[128] Q. Li, X. Zhao, S. Xu, X. Gao, Influence of steel fiber on dynamic compressive behavior of hybrid fiber ultra-high toughness cementitious composites at different strain rates, Construction and Building Materials, 125 (2016) 490–500.

[129] Y. Su, J. Li, C. Wu, P. Wu, Z. Li, Influences of nano-particles on dynamic strength of ultra-high-performance concrete, Composites Part B: Engineering. 91 (2016) 595–609.

[130] L. Soufeiani, S.N. Raman, M.Z.B. Jumaat, U.J. Alengaram, G. Ghadyani, and P. Mendis. Influences of the volume fraction and shape of steel fibers on fiber-reinforced concrete subjected to dynamic loading – A review. Engineering Structures, 124:405–417, 2016

[131] L.D. Bertholf, C.H. Karnes. Two-dimensional analysis of the split Hopkinson pressure bar system. Journal of Mechanics and Physics of Solids, 1975; 23:1-19.

[132] J.B. Mander, M.J.N. Priestley, R. Park, Theoretical stress-strain model for confined concrete. Journal of Structural Engineering, ASCE 1988; 114:1804–26.

[133] U.S. Lindholm, Some experiments with the split Hopkinson pressure bar. Journal of Mechanics and Physics of Solids, 1964; 12:317–35.

[134] D.J. Frew, M.J. Forrestal, W. Chen. Pulse shaping techniques for testing brittle materials with a split Hopkinson pressure bar. Experimental Mechanics 2002;42(1):93–106.

[135] F. Dai, K. Xia, L. Tang. Rate dependence of the flexural tensile strength of Lauren tian granite [J]. International Journal of Rock Mechanics and Mining Sciences, 2010, 47 (3): 469-475.

[136] Z. Xu, H. Hao, H.N. Li. Experimental study of dynamic compressive properties of fibre reinforced concrete material with different fibres. Materials and Design, 2012; 33:42-55.

[137] Y. Hao, H. Hao, Dynamic compressive behaviour of spiral steel fibre reinforced concrete in split Hopkinson pressure bar tests. Construction and Building Materials 2013, 48, 521-532.

[138] P. Challenor, R. Hankin, R. Marsh, Avoiding dangerous climate change. Cambridge, UK: Cambridge University Press; 2006.

[139] F.A. DiazDelaO, S. Adhikari. Structural dynamic analysis using Gaussian process emulators. Engineering Computations 2010;27(5):580-605.

[140] T.E. Fricker, J.E. Oakley, N.D. Sims, K. Worden. Probabilistic uncertainty analysis of an FRF of a structure using a Gaussian process emulator. Mechanical Systems and Signal Processing, 2011, 25(8):2962-2975.

[141] Y. Xiong, W. Chen, D. Appley, X. Ding. A non-stationary covariance-based kriging method for meta modelling in engineering design, International Journal of Numerical Methods in Engineering, 2006, 71: 733-756.

[142] N. D. Lawrence, M. Seeger, and R. Herbrich, (2003). Fast sparse Gaussian process methods: The informative vector machine. In Advances in Neural Information Processing Systems 15, pages 609–616. MIT Press.

[143] M. Seeger, C. K. I. Williams, and N. D. Lawrence, (2003). Fast forward selection to speed up sparse Gaussian process regression. In Proceedings of the 9th International Workshop on AI Stats.

[144] C. K. I. Williams, and M. Seeger,(2001). Using the Nyström method to speed up kernel machines. In Advances in Neural Information Processing Systems 13, pages 682–688. MIT Press.

[145] E. Snelson and Z. Ghahramani, (2006). Sparse Gaussian processes using pseudoinputs. In Advances in Neural Information Processing Systems 18, pages 1259–1266. MIT Press.

[146] M. K. Titsias, Variational Learning of Inducing Variables in Sparse Gaussian Processes. Artificial Intelligence and Statistics, 5:567-574, 2009.

[147] L. Csato and M. Opper, (2002). Sparse online Gaussian processes. Neural Computation, 14:641-668.

[148] M. Seeger, 2003., Bayesian Gaussian Process Models: PAC-Bayesian Generalisation Error Bounds and Sparse Approximations. PhD thesis, University of Edinburgh.

[149] M.K. Titsias (2009), Variational Model Selection for Sparse Gaussian Process Regression. Technical report, School of Computer Science, University of Manchester.

[150] Etienne Balmes, J.R. Wright, GARTEUR group on ground vibration testing. Results from the test of a single structure by 12 laboratories in Europe, OFFICE NATIONAL D ETUDES ET DE RECHERCHES AEROSPATIALE ONERA-PUBLICATION-TP (1997).

[151] Y. Govers, M. Link, Stochastic model updating of an aircraft like structure by parameter covariance matrix adjustment. In: Proceedings of International Conference on Noise and Vibration Engineering, ISMA/USD2010, Leuven, Belgium, pp. 2639-2656 (2010).

[152] Y. Govers, M. Link, Using stochastic experimental modal data for identifying stochastic finite element parameters of the AIRMOD benchmark structure. In: Proceedings of International Conference on Noise and Vibration Engineering, ISMA/USD2010, Leuven, Belgium, pp. 4697-4715 (2012).

[153] Y. Govers, M. Böswald, U. Füllekrug, D. Göge, M. Link, Analysis of sources and quantification of uncertainty in experimental modal data, in: Proceedings of the International Seminar on Modal Analysis, 2006, pp. 4161-4173.

[154] E. Simoen, G. Roeck De, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: a review, Mech, Syst. Signal Process 56 (2015) 123-149.

[155] H.P. Wan, W.X. Ren, A residual-based Gaussian process model framework for finite element model updating, Computers & Structures, 156 (2015), 149-159.

[156] S. Gull, Bayesian inductive inference and maximum entropy, Maximum-Entropy Bayesian Methods in Science and Engineering, 1 (1988) 53-74.

[157] X. Xu, Z. Yan, S. Xu, 2015, Estimating wind speed probability distribution by diffusion-based kernel density method, Electric Power Systems Research, 121: 28–37.

[158] Marc C. Kennedy, A. O' Hagan, Bayesian calibration of computer models, Journal of the royal statistical society. Series B (statistical methodology), Vol 63, No. 3 (2001), pp. 425-464.

[159] E. Patelli, Y. Govers, M. Broggi, H. M.Gomes, M. Link, J. E. Mottershead, Sensitivity or Bayesian model updating: a comparison of techniques using the dlr airmod test data, Archive of Applied Mechanics 87 (2017) 905-925.

[160] E. Patelli, COSSAN: A Multidisciplinary Software Suite for Uncertainty Quantification and Risk Management, in: Springer International Publishing, Cham, 2016, pp. 1-69. Doi: 10.1007/978-3-319-11259-6-59-1.

[161] E. Patelli, M. Broggi, D. Marco, M. Beer, OpenCossan: An efficient open tool for dealing with epistemic and aleatory uncertainties, in: Vulnerability, Uncertainty, and Risk, Ameriacan Society of Civil Engineers, 2014, pp. 2564-2573. URL: doi:10.1061/9780784413609.258.

[162] W. Chen, Y. Xiong, K.L. Tsui and S. Wang, 2008, A Design-Driven Validation Approach Using Bayesian Prediction Models, ASME Journal of Mechanical Design, 130(2), p.021101.

[163] Z. Wang, H.-Z. Huang and Y. Liu, 2010, A Uniform Framework for Integrated Optimization Under Uncertainty, ASME Journal of Mechanical Design, 132(5), p051008.

[164] M. C. Kennedy and A. O'Hagan, Bayesian calibration of computer models, Journal of the Royal Statistical Society B, 63 (2001), pp. 425-464.

[165] L.-E. Lindgren, H. Alberg and K. Domkin, Constitutive Modelling and Parameter Optimization. Proceeding 7th International Conference on Computational Plasticity, Barcelona.

[166] J. McFarland, J. Mahadevan, V.S. Romero and L. Swiler, 2008. Calibration and Uncertainty Analysis for Computer Simulation with Multivariate Output, AIAA Journal, 46(5), pp. 1253-1265.

[167] S. Wang, K.L. Tsui and W. Chen, 2009, Bayesian Validation of Computer Models, Technometrics, 51(4), pp. 439-451.

[168] Y. Xiong, W. Chen, K.L. Tsui and D.W. Apley, 2009, A Better Understanding of Model Updating Strategies in Validation Engineering Models. Computer Methods in Applied Mechanics and Engineering. 198(15-16), pp. 1327-1337.

[169] G. B. Arhonditsis, D. Papantou, W. Zhang, G. Perhar, E. Massos, and M. Shi, Bayesian calibration of mechanistic aquatic biogeochemical models and bene_ts for environmental management, Journal of Marine Systems, 73 (2008), pp. 8-30.

[170] M. Strong, J. E. Oakley, and J. Chilcott, Managing structural uncertainty in health economic decision models: a discrepancy approach, Journal of the Royal Statistical Society C, 61 (2012), pp. 25-45.

[171] J. M. Murphy, B. B. B. Booth, M. Collins, G. R. Harris, D. M. H. Sexton, and M. J.Webb, A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles, Philos Transact A Math Phys Eng Sci, 365 (2007), pp. 1993-2028.

[172] M.J. Bayarri, J. O. Berger, R. Paulo, J. Sacks, J.A. Cafeo, J. Cavendish, C.H. Lin and J. Tu, 2007, A Framework for Validation of Computer Models, Technometrics, 49(2), pp. 138-154.

[173] Y. Liu, W. Chen, P. Arendt and H.-Z. Huang, 2011, Toward a Better Understanding of Model Validation Metrics, ASME Journal of Mechanical Design, 133(7), p071005.

[174] R.G. Easterling and J.O. Berger, 2002, Statistical Foundations for the Validation of Computer Models, Presented at Computer Model Verification and Validation in the 21st Century Workshop, John Hopkins University.

[175] D. Drignei, 2009, A Kriging Approach to the Analysis of Climate Model Experiments, Journal of Agriculture, Biological and Environmental Statistics, 14(1), pp. 99-114.

[176] M. C. Kennedy and A. O'Hagan, 2000, Supplementary details on Bayesian Calibration of Computer Models, available at http://www2.stat.duke.edu/~fei/samsi/Oct_09/bayesian_calibration_of_computer_models_supplement.pdf

[177] F. Liu, M.J. Bayarri and J.O. Berger, 2009, Modularization in Bayesian Analysis, With Emphasis on Analysis of Computer Models, Bayesian Analysis, 4(1), pp. 119-150.