CALIFORNIA INSTITUTE OF TECHNOLOGY

EARTHQUAKE ENGINEERING RESEARCH LABORATORY

NEW BAYESIAN UPDATING METHODOLOGY FOR MODEL VALIDATION AND ROBUST PREDICTIONS BASED ON DATA FROM HIERARCHICAL SUBSYSTEM TESTS

BY

SAI HUNG CHEUNG AND JAMES L. BECK

REPORT NO. EERL-2008-04

PASADENA, CALIFORNIA

AUGUST 2008

ACKNOWLEDGMENTS

The first author gratefully acknowledges the financial support of Caltech's George W. Housner Fellowship

ABSTRACT

In many engineering applications, it is a formidable task to construct a mathematical model that is expected to produce accurate predictions of the behavior of a system of interest. During the construction of such predictive models, errors due to imperfect modeling and uncertainties due to incomplete information about the system and its input always exist and can be accounted for appropriately by using probability logic. Often one has to decide which proposed candidate models are acceptable for prediction of the target system behavior. In recent years, the problem of developing an effective model validation methodology has attracted attention in many different fields of engineering and applied science. Here, we consider the problem where a series of experiments are conducted that involve collecting data from successively more complex subsystems and these data are to be used to predict the response of a related more complex system. A novel methodology based on Bayesian updating of hierarchical stochastic system model classes using such experimental data is proposed for uncertainty quantification and propagation, model validation, and robust prediction of the response of the target system. After each test stage, we use all the available data to calculate the posterior probability of each stochastic system model along with the quality of its robust prediction. The proposed methodology is applied to the 2006 Sandia static-frame validation challenge problem to illustrate our approach for model validation and robust prediction of the system response. Recently-developed stochastic simulation methods are used to solve the computational problems involved.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

In recent years, the problem of model validation for a system has attracted the attention of many researchers (e.g. Babuška and Oden, 2004; Oberkampf et al. 2004; Babuška et al. 2006) because of the desire to provide a measure of confidence in the predictions of a system model. In particular, in May 2006, the Sandia Model Validation Challenge Workshop brought together a group of researchers to present various approaches to model validation (Hills et al. 2008). The participants could choose to work on any of three problems; one in heat transfer (Dowding et al. 2008), one in structural dynamics (Red-Horse and Paez 2008) and one in structural statics (Babuška et al. 2008). The difficult issue of how to validate a model is, however, still not settled; indeed, it is clear that a model that has given good predictions in tests so far might perform poorly under different circumstances, such as an excitation with different characteristics.

Our philosophy when predicting the behavior of a system of interest is that one should develop candidate sets of probabilistic predictive input-output models to give robust predictions that explicitly address errors due to imperfect models and uncertainties due to incomplete information. For model validation, it is then desirable to check based on system test data whether any of the proposed candidate model sets are highly probable and whether they provide high quality predictions of the system behavior of interest.

Sometimes the full system cannot be readily tested because it is too expensive or too large, or due to other limitations, but some of its subsystems may be tested. Here we introduce the concept of hierarchical stochastic system model classes and then propose a Bayesian methodology using them to treat modeling and input uncertainties in model validation, uncertainty propagation and robust predictions of the response of the full system. The Sandia

static-frame validation problem is used to illustrate the proposed methodology. The results of other researchers' studies of this problem are presented in a special issue of the journal Computer Methods in Applied Mechanics and Engineering (Chleboun 2008; Babuška et al. 2008; Grigoriu and Field 2008; Pradlwarter and Schuëller 2008; Rebba and Cafeo 2008).

2 STOCHASTIC SYSTEM MODEL CLASSES AND THEIR POSTERIOR PROBABILITIES

There always exist modeling errors and other uncertainties associated with the process of constructing a mathematical model of a system. A fully probabilistic Bayesian model updating approach provides a robust and rigorous framework for these applications due to its ability to characterize modeling uncertainties associated with the system and to its exclusive foundation on the probability axioms. In our applications of the Bayesian approach, we use the Cox-Jaynes interpretation of probability (Cox 1961; Jaynes 2003) as an extension of binary Boolean logic to a multi-valued logic of plausible inference where the relative plausibility of each model within a class of models is quantified by its probability.

A key concept in our approach is a *stochastic system model class* $\mathcal{M}$ which consists of a *set* of probabilistic predictive input-output models for a system together with a probability distribution, the *prior*, over this set that quantifies the initial relative plausibility of each predictive model. For simpler presentation, we will usually abbreviate the term "stochastic system model class" to "model class". Based on $\mathcal{M}$, one can use data $\mathcal{D}$ to compute the updated relative plausibility of each predictive model in the set defined by $\mathcal{M}$. This is quantified by the *posterior* PDF $p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M})$ for the uncertain model parameters $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^D$ which specify a particular model within $\mathcal{M}$. By Bayes' theorem, this posterior PDF is given by:

$$p(\boldsymbol{\theta} \mid \mathcal{D}, \mathcal{M}) = c^{-1} p(\mathcal{D} \mid \boldsymbol{\theta}, \mathcal{M}) p(\boldsymbol{\theta} \mid \mathcal{M}) \tag{2.1}$$

where $c = p(\mathcal{D}|\mathcal{M}) = \int p(\mathcal{D}|\boldsymbol{\theta},\mathcal{M})p(\boldsymbol{\theta}|\mathcal{M})d\boldsymbol{\theta}$ is the normalizing constant which makes the probability volume under the posterior PDF equal to unity; $p(\mathcal{D}|\boldsymbol{\theta},\mathcal{M})$ is the *likelihood function* which expresses the probability of getting data $\mathcal{D}$ based on the predictive PDF for the response given by model $\boldsymbol{\theta}$ within $\mathcal{M}$; and $p(\boldsymbol{\theta}|\mathcal{M})$ is the prior PDF for $\mathcal{M}$ which one can freely choose to quantify the initial plausibility of each model defined by the value of the parameters $\boldsymbol{\theta}$. For example, through the use of prior information that is not readily built into the predictive PDF that produces the likelihood function, the prior can be chosen to provide regularization of ill-conditioned inverse problems (Bishop 2006). As emphasized by Jaynes (2003), probability models represent a quantification of the state of knowledge about real phenomena conditional on the available information and should not be imagined to be a property inherent in these phenomena, as often believed by those who ascribe to the common interpretation that probability is the relative frequency of "inherently random" events in the "long run".

Based on the topology of $p(\mathcal{D}|\boldsymbol{\theta},\mathcal{M})$ in the parameter space, and, in particular, the set $\{\boldsymbol{\theta} \in \Theta : \boldsymbol{\theta}=\arg\max p(\mathcal{D}|\boldsymbol{\theta},\mathcal{M})\}$ of MLEs (*maximum likelihood estimates*), a model class $\mathcal{M}$ can be classified into 3 different categories (Beck & Katafygiotis 1991, 1998; Katafygiotis & Beck 1998): *globally identifiable* (unique MLE), *locally identifiable* (discrete set of MLEs) and *unidentifiable* (a continuum of MLEs) based on the available data $\mathcal{D}$. Full Bayesian updating can treat all these cases (Yuen et al. 2004).

Model class comparison is a rigorous Bayesian updating procedure that judges the plausibility of different candidate model classes, based on their posterior probability (that is, their probability conditional on the data from the system). Its application to system identification of dynamic systems that are globally identifiable or unidentifiable was studied in

Beck & Yuen (2004) and Muto & Beck (2008), respectively. In these publications, a model class is referred to as a *Bayesian model class*.

Given a set of candidate model classes $M=\{\mathcal{M}_j: j=1,2,\ldots N_M\}$, we calculate the posterior probability $P(\mathcal{M}_j \mid \mathcal{D}, M)$ of each model class based on system data $\mathcal{D}$ by using Bayes' Theorem:

$$P(\mathcal{M}_j \mid \mathcal{D}, M) = \frac{p(\mathcal{D} \mid \mathcal{M}_j) P(\mathcal{M}_j \mid M)}{p(\mathcal{D} \mid M)} \tag{2.2}$$

where $P(\mathcal{M}_j|M)$ is the prior probability of each $\mathcal{M}_j$ and can be taken to be $1/N_M$ if one considers all N_M model classes as being equally plausible a priori; $p(\mathcal{D}|\mathcal{M}_j)$ expresses the probability of getting the data $\mathcal{D}$ based on $\mathcal{M}_j$ and is called the *evidence* (or sometimes marginal likelihood) for $\mathcal{M}_j$ provided by the data $\mathcal{D}$ and it is given by the Theorem of Total Probability:

$$p(\mathcal{D} \mid \mathcal{M}_j) = \int p(\mathcal{D} \mid \boldsymbol{\theta}, \mathcal{M}_j) p(\boldsymbol{\theta} \mid \mathcal{M}_j) d\boldsymbol{\theta} \tag{2.3}$$

Although $\boldsymbol{\theta}$ corresponds to different sets of parameters and can be of different dimension for different $\mathcal{M}_j$, for simpler presentation a subscript j on $\boldsymbol{\theta}$ is not used since explicit conditioning on $\mathcal{M}_j$ indicates which parameter vector $\boldsymbol{\theta}$ is involved.

Notice that (2.3) can be interpreted as follows: the evidence gives the probability of the data according to $\mathcal{M}_j$ (if (2.3) is multiplied by an elemental volume in the data space) and it is equal to a weighted average of the probability of the data according to each model specified by $\mathcal{M}_j$, where the weights are given by the probability $p(\boldsymbol{\theta}|\mathcal{M}_j)d\boldsymbol{\theta}$ of the parameter values corresponding to each model. The evidence therefore corresponds to a type of integrated global

sensitivity analysis where the prediction $p(\mathcal{D}|\boldsymbol{\theta},\mathcal{M}_j)$ of each model specified by $\boldsymbol{\theta} \in \Theta$ is considered but it is weighted by the relative plausibility of the corresponding model.

The computation of the multi-dimensional integral in (2.3) is nontrivial. Laplace's method of asymptotic approximation (e.g. Beck & Katafygiotis 1991, 1998) has been proposed for its evaluation (e.g. Mackay 1992; Beck & Yuen 2004), which, in effect, utilizes a weighted sum of Gaussian PDFs centered on each MLE as an approximation to the posterior PDF. However, the accuracy of such an approximation is questionable when (i) the amount of data is small, or (ii) the chosen class of models turns out to be unidentifiable based on the available data. Under these circumstances, only stochastic simulation methods are practical and the Hybrid Gibbs TMCMC method presented in Appendix B is used later in the illustrative example.

It is worth noting that from (2.3), the log evidence can be expressed as the difference of two terms (Ching et al. 2005; Muto & Beck 2008):

$$\ln[p(\mathcal{D} \mid \mathcal{M}_j)] = E[\ln(p(\mathcal{D} \mid \boldsymbol{\theta}, \mathcal{M}_j)] - E[\ln \frac{p(\boldsymbol{\theta} \mid \mathcal{D}, \mathcal{M}_j)}{p(\boldsymbol{\theta} \mid \mathcal{M}_j)}] \tag{2.4}$$

where the expectation is with respect to the posterior $p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M}_j)$. The first term is the posterior mean of the log likelihood function, which gives a measure of the goodness of the fit of the model class $\mathcal{M}_j$ to the data, and the second term is the Kullback-Leibler divergence, or relative entropy (Cover & Thomas 2006), which is a measure of the information gain about $\mathcal{M}_j$ from the data $\mathcal{D}$ and is always non-negative. The importance of (2.4) is that it shows that the log evidence for $\mathcal{M}_j$, which controls the posterior probability of this model class according to (2.2), explicitly

builds in a trade-off between the data-fit of the model class and its “complexity” (how much information it takes from the data).

3 ROBUST PREDICTIVE ANALYSIS USING STOCHASTIC SYSTEM MODEL CLASSES

One of the most useful applications of Bayesian model updating is to make robust predictions about future events based on past observations. Let $\mathcal{D}$ denote data from available measurements on a system. Based on a candidate model class $\mathcal{M}_j$, all the probabilistic information for the prediction of a vector of future responses $\mathbf{X}$ is contained in the *posterior robust* predictive PDF for $\mathcal{M}_j$ given by the Theorem of Total Probability (Papadimitriou et al. 2001):

$$p(\mathbf{X} \mid \mathcal{D}, \mathcal{M}_j) = \int p(\mathbf{X} \mid \boldsymbol{\theta}, \mathcal{D}, \mathcal{M}_j) p(\boldsymbol{\theta} \mid \mathcal{D}, \mathcal{M}_j) d\boldsymbol{\theta} \tag{3.1}$$

The interpretation of (3.1) is similar to that given for (2.3) except now the prediction $p(\mathbf{X}|\boldsymbol{\theta},\mathcal{D},\mathcal{M}_j)$ of each model specified by $\boldsymbol{\theta} \in \Theta$ is weighted by its posterior probability $p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M}_j)d\boldsymbol{\theta}$ because of the conditioning on the data $\mathcal{D}$. If this conditioning on $\mathcal{D}$ in (3.1) is dropped so, for example, the prior $p(\boldsymbol{\theta}|\mathcal{M}_j)$ is used in place of the posterior $p(\boldsymbol{\theta}|\mathcal{D}, \mathcal{M}_j)$, the result $p(\mathbf{X}|\mathcal{M}_j)$ of the integration is the *prior robust* predictive PDF.

Many system performance measures can be expressed as the expectation of some function $\mathbf{g}(\mathbf{X})$ with respect to the posterior robust predictive PDF in (3.1) as follows:

$$E[\mathbf{g}(\mathbf{X}) \mid \mathcal{D}, \mathcal{M}_j] = \int \mathbf{g}(\mathbf{X}) p(\mathbf{X} \mid \mathcal{D}, \mathcal{M}_j) d\mathbf{X} \tag{3.2}$$

Some examples of important special cases are:

1) $\mathbf{g}(\mathbf{X})=I_F(\mathbf{X})$, which is equal to 1 if $\mathbf{X} \in F$ and 0 otherwise, where F is a region in the response space that corresponds to unsatisfactory system performance, then the integral in (3.2) is equal to the robust "failure" probability $P(F|\mathcal{D}, \mathcal{M}_j)$;

2) $\mathbf{g}(\mathbf{X})=\mathbf{X}$, then the integral in (3.2) becomes the robust mean response;

3) $\mathbf{g}(\mathbf{X})=(\mathbf{X}-E[\mathbf{X}|\mathcal{D}, \mathcal{M}_j])(\mathbf{X}-E[\mathbf{X}|\mathcal{D}, \mathcal{M}_j])^T$, then the integral in (3.2) is equal to the robust covariance matrix of $\mathbf{X}$.

The Bayesian approach to robust predictive analysis requires the evaluation of multi-dimensional integrals, such as in (3.1), and this usually cannot be done analytically. Laplace's method of asymptotic approximation has been used in the past (e.g. Beck and Katafygiotis 1998; Papadimitriou et al. 2001), which utilizes a Gaussian sum approximation to the posterior PDF, as mentioned before for (2.3). Such an approximation requires a non-convex optimization in what is usually a high-dimensional parameter space, which is computationally challenging, especially when the model class is not globally identifiable and so there may be multiple global maximizing points (Katafygiotis & Lam 2002). Thus, in recent years, focus has shifted from asymptotic approximations to using stochastic simulation methods in which samples are generated from the posterior PDF $p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M}_j)$. There are several difficulties related to this sampling: (i) the normalizing constant c in Bayes' Theorem in (2.1), which is actually the evidence in (2.3), is usually unknown a priori and its evaluation requires a high-dimensional integration over the uncertain parameter space; and (ii) the high probability content of $p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M}_j)$ occupies a much smaller volume than that of the prior PDF, so samples in the high probability region of $p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M}_j)$ cannot be generated efficiently by sampling from the prior PDF using direct Monte Carlo simulation.

To tackle the aforementioned difficulties, stochastic simulation methods, such as multi-level Metropolis-Hastings (e.g. Beck & Au 2002; Ching & Cheng 2007), Gibbs sampling (e.g. Ching et al. 2006), and Hybrid Monte Carlo simulation (e.g. Cheung & Beck 2007, 2008a) have been used recently to perform Bayesian model updating for dynamic systems. For the illustrative example later, we use a variant that we call Hybrid Gibbs TMCMC method (see Appendix B). In these methods, all the probabilistic information encapsulated in $p(\boldsymbol{\theta}|\mathcal{D}, \mathcal{M}_j)$ is characterized by posterior samples $\boldsymbol{\theta}^{(k)}$, k=1,2,...,K, and the integral in (3.1) can be approximated by:

$$p(\mathbf{X} \mid \mathcal{D}, \mathcal{M}_j) \approx \frac{1}{K} \sum_{k=1}^{K} p(\mathbf{X} \mid \boldsymbol{\theta}^{(k)}, \mathcal{D}, \mathcal{M}_j) \tag{3.3}$$

Samples of $\mathbf{X}$ can then be generated from each of the $p(\mathbf{X} \mid \boldsymbol{\theta}^{(k)}, \mathcal{D}, \mathcal{M}_j)$ with equal probability. The probabilistic information encapsulated in $p(\mathbf{X} \mid \mathcal{D}, \mathcal{M}_j)$ is characterized by these samples of $\mathbf{X}$.

If a set of candidate model classes $M=\{\mathcal{M}_j: j=1,2,\ldots N_M\}$ is being considered for a system, all the probabilistic information for the prediction of future responses $\mathbf{X}$ is contained in the *hyper-robust* predictive PDF for M given by the Theorem of Total Probability:

$$p(\mathbf{X} \mid \mathcal{D}, M) = \sum_{j=1}^{N_M} p(\mathbf{X} \mid \mathcal{D}, \mathcal{M}_j) P(\mathcal{M}_j \mid \mathcal{D}, M) \tag{3.4}$$

where the robust predictive PDF for each model class $\mathcal{M}_j$ is weighted by its posterior probability $P(\mathcal{M}_j|\mathcal{D}, M)$ from (2.2). Equation (3.4) is also called *posterior model averaging* in the Bayesian statistics literature (Raftery et al. 1997, Hoeting et al. 1999).

4 HIERARCHICAL STOCHASTIC SYSTEM MODEL CLASSES AND MODEL VALIDATION

In this section, a novel model validation methodology based on a new concept of hierarchical stochastic system model classes is proposed (building on the theoretical foundations presented in the previous sections) so that a rational decision can be made regarding which proposed model classes should be used for predicting the response of a target system. The proposed methodology is based on using full Bayesian updating to investigate multiple important aspects of the performance of the candidate model classes, including their quality of prediction, their posterior probabilities and their contribution to response predictions of the final system. We do not make a binary reject/accept step but instead provide the decision maker with information about these important aspects, which can be combined with other considerations when making a decision related to the target system; for example, should the current target system design be accepted or modified?

Suppose during construction of the system, a series of I experiments are conducted where data $\mathcal{D}_i$, i=1,…, I, are collected from each of I similarly complex, or successively more complex, subsystems and these data are to be used to predict the response of the more complex target system. The i-th level subsystem is either a standalone subsystem (especially in lower levels) or one comprised of a combination of some (or all) tested subsystems from the previous levels, together, possibly, with new untested subsystems.

4.1 Analysis and full Bayesian updating of i-th subsystem

The presentation in this subsection is very general and the reader may find it helpful to look at the example illustrating the hierarchical concepts in the last subsection of this section. We assume that a set $M_i = \{\mathcal{M}_j^{(i)}: j=1,2,\ldots N_i\}$ of model classes is proposed for the i-th subsystem which are either newly defined or built-up by extending the model classes for some (or all) tested subsystems in the previous levels. In the latter case, a model class for the i-th subsystem is built-up by extending at most one model class for each relevant lower-level subsystem since candidate model classes for each such subsystem are supposed to be competing. Denote uncertain model parameters for the model class $\mathcal{M}_j^{(i)}$ by $\boldsymbol{\theta}^{(i,j)}=[\boldsymbol{\varphi}^{(i,j)}, \boldsymbol{\xi}^{(i,j)}]$ where $\boldsymbol{\varphi}^{(i,j)}$, if any, are the new uncertain model parameters and $\boldsymbol{\xi}^{(i,j)}$, if any, are the uncertain model parameters corresponding to a model class for some subsystems in the previous levels, that is, these parameters of $\mathcal{M}_j^{(i)}$ are also in model classes of subsystems of the ith subsystem. In the proposed hierarchical approach, the model class $\mathcal{M}_j^{(i)}$ is based on the "prior" (prior to the i^{th} subsystem test but posterior to all previous tests):

$$p(\boldsymbol{\theta}^{(i,j)} \mid \mathcal{D}_1,\ldots,\mathcal{D}_{i-1},\mathcal{M}_j^{(i)}) = p(\boldsymbol{\varphi}^{(i,j)} \mid \mathcal{M}_j^{(i)})\, p(\boldsymbol{\xi}^{(i,j)} \mid \mathcal{D}_1,\ldots,\mathcal{D}_{i-1}) \tag{4.1}$$

where $p(\boldsymbol{\varphi}^{(i,j)}|\mathcal{M}_j^{(i)})$ quantifies the prior uncertainties in the new parameters $\boldsymbol{\varphi}^{(i,j)}$ in model class $\mathcal{M}_j^{(i)}$ and $p(\boldsymbol{\xi}^{(i,j)}|\mathcal{D}_1,\ldots,\mathcal{D}_{i-1})$ is the most updated PDF of $\boldsymbol{\xi}^{(i,j)}$ given data collected from all subsystems in the previous levels. For simplicity, the conditioning of $p(\boldsymbol{\xi}^{(i,j)}|\mathcal{D}_1,\ldots,\mathcal{D}_{i-1})$ on the model classes previously considered which contain components of $\boldsymbol{\xi}^{(i,j)}$ are left implicit. For i=1, $p(\boldsymbol{\theta}^{(i,j)}|\mathcal{D}_1,\ldots,\mathcal{D}_{i-1},\mathcal{M}_j^{(i)}) = p(\boldsymbol{\theta}^{(1,j)}|\mathcal{M}_j^{(1)})$.

At the end of the experiments on the i-th subsystem where data $\mathcal{D}_i$ are collected, the following procedure is used to check the prediction quality of each candidate model class being considered for the i-th subsystem. For each model class $\mathcal{M}_j^{(i)}$ in M_i and for each measured quantity in $\mathcal{D}_i$, the *consistency* of the predicted response is first investigated by calculating the difference of the measured quantity in $\mathcal{D}_i$ and the mean of the corresponding prior robust predicted response. The robust predicted response given by $\mathcal{M}_j^{(i)}$ is consistent if this difference is no more than a certain number of standard deviations (e.g., no more than 2 to 3 standard deviations). An alternative way of investigating the consistency is to check whether each measured quantity in $\mathcal{D}_i$ is within q percentile and (100-q) percentile of the robust predicted response (e.g., q can be 1). The mean and standard deviation of the prior robust predicted response can be calculated using (3.2) and (3.3) but with samples drawn from the prior in (4.1).

Next, the *accuracy* of the prediction is investigated by calculating the probability that the prior robust predicted response using $\mathcal{M}_j^{(i)}$ (again based on $p(\boldsymbol{\theta}^{(i,j)}|\mathcal{D}_1,\ldots, \mathcal{D}_{i-1},\mathcal{M}_j^{(i)})$ in (4.1)) is within a certain b% (e.g. 10%) of the measured quantity using (3.2) and (3.3). This probability is related to the prediction error of each model class for the i-th level subsystem and reflects the predictability of these models before being updated using data $\mathcal{D}_i$. Note that a model class may give consistent predictions but not accurate ones because, for example, it has a relatively large standard deviation.

Next, for each model class $\mathcal{M}_j^{(i)}$ in M_i, the uncertainties in the model parameters $\boldsymbol{\theta}^{(i,j)}$ are updated using all the available data, as quantified by $p(\boldsymbol{\theta}^{(i,j)}|\mathcal{D}_1,\ldots, \mathcal{D}_i, \mathcal{M}_j^{(i)})$ through Bayes' Theorem:

$$p(\boldsymbol{\theta}^{(i,j)} \mid \mathcal{D}_1,...,\mathcal{D}_i,\mathcal{M}_j^{(i)}) = c_{i,j}^{-1}\, p(\mathcal{D}_i \mid \boldsymbol{\theta}^{(i,j)},\mathcal{M}_j^{(i)})\, p(\boldsymbol{\theta}^{(i,j)} \mid \mathcal{D}_1,...,\mathcal{D}_{i-1},\mathcal{M}_j^{(i)}) \tag{4.2}$$

where the data $\mathcal{D}_1,\ldots,\mathcal{D}_{i\text{-}1}$ are modeled as irrelevant to the probability of getting $\mathcal{D}_i$ when $\boldsymbol{\theta}^{(i,j)}$ is given since this parameter vector defines the predictive probability model for the model class $\mathcal{M}_j^{(i)}$. Recall that $\boldsymbol{\xi}^{(i,j)}$ are the uncertain model parameters corresponding to some model classes of subsystems already considered in the previous levels. A subtle point to be noted is that sometimes uncertainties for some other model parameters $\boldsymbol{\Phi}^{(i,j)}$ corresponding to the model classes containing components of $\boldsymbol{\xi}^{(i,j)}$ will also be updated when updating uncertainties in $\boldsymbol{\xi}^{(i,j)}$ using $\mathcal{D}_1,\ldots,\mathcal{D}_{i\text{-}1}$. Since $\boldsymbol{\Phi}^{(i,j)}$ and $\boldsymbol{\xi}^{(i,j)}$ are not stochastically independent given $\mathcal{D}_1,\ldots,\mathcal{D}_i$, the uncertainties in both $\boldsymbol{\theta}^{(i,j)}$ and $\boldsymbol{\Phi}^{(i,j)}$ need to be updated together from Bayes' Theorem:

$$p(\boldsymbol{\theta}^{(i,j)},\Phi^{(i,j)} \mid \mathcal{D}_1,...,\mathcal{D}_i,\mathcal{M}_j^{(i)}) = \tilde{c}_{i,j}^{-1}\, p(\mathcal{D}_i \mid \boldsymbol{\theta}^{(i,j)},\mathcal{M}_j^{(i)})\, p(\boldsymbol{\xi}^{(i,j)},\boldsymbol{\Phi}^{(i,j)} \mid \mathcal{D}_1,...,\mathcal{D}_{i-1})\, p(\boldsymbol{\varphi}^{(i,j)} \mid \mathcal{M}_j^{(i)}) \tag{4.3}$$

where $\boldsymbol{\theta}^{(i,j)}=[\boldsymbol{\varphi}^{(i,j)}, \boldsymbol{\xi}^{(i,j)}]$ and the data $\mathcal{D}_1,\ldots,\mathcal{D}_{i\text{-}1}$ are modeled as irrelevant to the probability of getting $\mathcal{D}_i$ given $\boldsymbol{\theta}^{(i,j)}$, as before. Finally, $p(\boldsymbol{\theta}^{(i,j)}|\mathcal{D}_1,\ldots,\mathcal{D}_i,\mathcal{M}_j^{(i)})$ can be obtained as the marginal PDF of $p(\boldsymbol{\theta}^{(i,j)}, \boldsymbol{\Phi}^{(i,j)}|\mathcal{D}_1,\ldots,\mathcal{D}_i,\mathcal{M}_j^{(i)})$.

The posterior probability $P(\mathcal{M}_j^{(i)}|\mathcal{D}_1,\ldots,\mathcal{D}_i,M_i)$ of each model class in M_i can be calculated as follows to evaluate the relative plausibility of each model class. If a model class $\mathcal{M}_j^{(i)}$ is built-up by extending or using model classes which have been updated using data from subsystems in the previous levels $k_1, k_2,\ldots, k_m$ where $k_1< k_2<\ldots< k_m$ and $1\leq m<i$, $P(\mathcal{M}_j^{(i)}|\mathcal{D}_1,\ldots,\mathcal{D}_i,M_i)$ is equal to $P(\mathcal{M}_j^{(i)}|\mathcal{D}_{k_1},...,\mathcal{D}_{k_m},\mathcal{D}_i,M_i)$. The most up-to-date evidence $p(\mathcal{D}_{k_1},...,\mathcal{D}_{k_m},\mathcal{D}_i|\mathcal{M}_j^{(i)})$ for $\mathcal{M}_j^{(i)}$ that is provided by the data $\mathcal{D}_{k_1},...,\mathcal{D}_{k_m},\mathcal{D}_i$, and which is required for calculating $P(\mathcal{M}_j^{(i)}|\mathcal{D}_{k_1},...,\mathcal{D}_{k_m},\mathcal{D}_i,M_i)$, is given by:

$$p(\mathcal{D}_{k_1},...,\mathcal{D}_{k_m},\mathcal{D}_i \mid \mathcal{M}_j^{(i)}) = p(\mathcal{D}_{k_1},...,\mathcal{D}_{k_m} \mid \mathcal{M}_j^{(i)})\, p(\mathcal{D}_i \mid \mathcal{D}_{k_1},...,\mathcal{D}_{k_m},\mathcal{M}_j^{(i)}) \tag{4.4}$$

In this equation, $p(\mathcal{D}_i|\, \mathcal{D}_{k_1},...,\mathcal{D}_{k_m},\mathcal{M}_j^{(i)})$ is given by:

$$p(\mathcal{D}_i \mid \mathcal{D}_{k_1},...,\mathcal{D}_{k_m},\mathcal{M}_j^{(i)}) = \int p(\mathcal{D}_i \mid \boldsymbol{\theta}^{(i,j)},\mathcal{M}_j^{(i)})\, p(\boldsymbol{\theta}^{(i,j)} \mid \mathcal{D}_1,...,\mathcal{D}_{i-1},\mathcal{M}_j^{(i)})\, d\boldsymbol{\theta}^{(i,j)} \tag{4.5}$$

which can be determined using a stochastic simulation method, such as the Hybrid Gibbs TMCMC method presented in Appendix B. The other factor in (4.4), $p(\mathcal{D}_{k_1},...,\mathcal{D}_{k_m} \mid \mathcal{M}_j^{(i)})$, is given by a product of the evidences which have already been determined at the end of previous experiments. This point will be more clear in the example illustrating the hierarchical concepts in the last subsection of this section. Based on (4.4), $P(\mathcal{M}_j^{(i)}|\mathcal{D}_1,\ldots,\mathcal{D}_i, M_i) = P(\mathcal{M}_j^{(i)}|\, \mathcal{D}_{k_1},...,\mathcal{D}_{k_m}, \mathcal{D}_i, M_i)$ can be calculated using (2.2) with $\mathcal{M}_j$ replaced by $\mathcal{M}_j^{(i)}$, M replaced by M_i and $\mathcal{D}$ by $\mathcal{D}_{k_1},...,\mathcal{D}_{k_m}, \mathcal{D}_i$.

In the special case that $\mathcal{M}_j^{(i)}$ is newly defined, i.e., not built-up by extending any model classes for subsystems in the previous levels, the posterior probability $P(\mathcal{M}_j^{(i)}|\mathcal{D}_1,\ldots,\mathcal{D}_i, M_i)$ is given by $P(\mathcal{M}_j^{(i)}|\mathcal{D}_i, M_i)$, which can be calculated using (2.2) with $\mathcal{M}_j$ replaced by $\mathcal{M}_j^{(i)}$, M replaced by M_i and $\mathcal{D}$ by $\mathcal{D}_i$ where the evidence $p(\mathcal{D}_i|\mathcal{M}_j^{(i)})$ for $\mathcal{M}_j^{(i)}$ is given by:

$$p(\mathcal{D}_i \mid \mathcal{M}_j^{(i)}) = \int p(\mathcal{D}_i \mid \boldsymbol{\theta}^{(i,j)},\mathcal{M}_j^{(i)})\, p(\boldsymbol{\theta}^{(i,j)} \mid \mathcal{M}_j^{(i)})\, d\boldsymbol{\theta}^{(i,j)} \tag{4.6}$$

which can be determined using a stochastic simulation method.

Based on all the data, $\mathcal{D}_1,\ldots, \mathcal{D}_i$, so far, the posterior robust prediction of the response vector $\mathbf{X}$ for the target system can be calculated using (3.3) and (3.4). If a model class $\mathcal{M}_j^{(i)}$ is

very improbable compared to the others in M_i, so that its contribution to the hyper-robust response prediction of the target system is negligible in (3.4), it can be neglected when building the candidate model classes for higher level subsystems in order to save computations. Note that (3.4) allows calculation of the most robust predictions for the i-th subsystem based on all the available information and viable model classes.

For each model class $\mathcal{M}_j^{(i)}$ in M_i and for each measured quantity in $\mathcal{D}_i$, the consistency of the predicted response is again investigated by examining the difference of the measured quantity in $\mathcal{D}_i$ and the mean of the corresponding posterior robust predicted response (again judged in terms of the number of standard deviations of the posterior robust predicted response). The robust predicted response is based on the "posterior" $p(\boldsymbol{\theta}^{(i,j)}|\mathcal{D}_1,\ldots, \mathcal{D}_i, \mathcal{M}_j^{(i)})$ given by (4.2) or (4.3) and its mean and standard deviation are calculated using (3.2) and (3.3). One can also check whether each measured quantity in $\mathcal{D}_i$ is within q percentile and (100-q) percentile of the posterior robust predicted response. Next, the accuracy of the prediction is investigated by calculating the probability that the robust predicted response (again based on $p(\boldsymbol{\theta}^{(i,j)}|\mathcal{D}_1,\ldots, \mathcal{D}_i, \mathcal{M}_j^{(i)})$) is within a certain b% (e.g. 10%) of the measured quantity using (3.2) and (3.3).

4.2 Example to illustrate hierarchical model classes

The following example is presented to illustrate the above theory on how to propagate uncertainties in parameters and calculate the posterior probability for a hierarchical stochastic system model class. Figure 4.1 shows the hierarchical structure of some of the model classes for the illustrative example. The ellipses show the subsystems for different levels; a black dot inside

an ellipse shows a candidate model class corresponding to that subsystem; the lower end of an arrow points to a model class which is used to build another model class pointed to by the top end of the same arrow. Shown next to an arrow is the set of data used to update the lower level model classes, along with the posterior PDF for the previous model class and the evidence required for calculating the posterior probability of this model class.

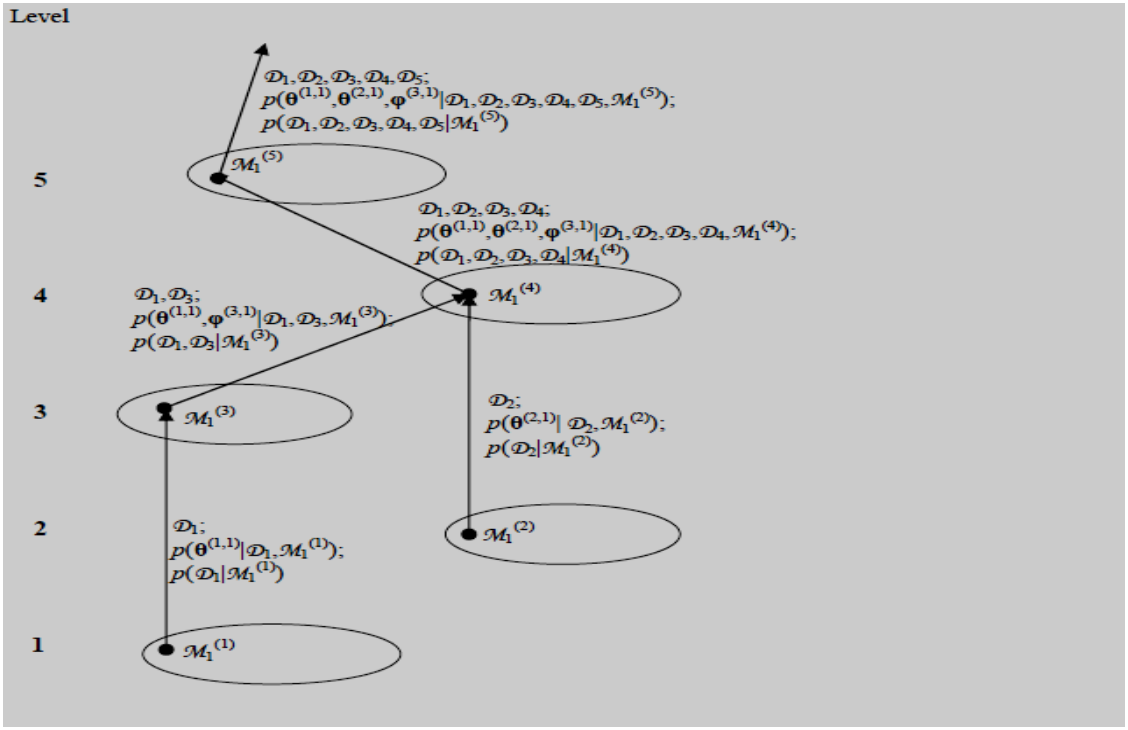

Figure 4.1: Schematic plot for an illustrative example of hierarchical model classes

Recall that $\mathcal{M}_1^{(1)}$ in M_1 is the first candidate model class with uncertain parameters $\boldsymbol{\theta}^{(1,1)}$ for the first level subsystem from which data $\mathcal{D}_1$ is collected. The posterior PDF $p(\boldsymbol{\theta}^{(1,1)}|\mathcal{D}_1, \mathcal{M}_1^{(1)})$ for $\mathcal{M}_1^{(1)}$ is given by (4.2) with the chosen prior PDF $p(\boldsymbol{\theta}^{(1,1)}|\mathcal{M}_1^{(1)})$. The evidence $p(\mathcal{D}_1|\mathcal{M}_1^{(1)})$, which is required for calculating the posterior probability $P(\mathcal{M}_1^{(1)}|\mathcal{D}_1, M_1)$ for $\mathcal{M}_1^{(1)}$, is given by (4.6) with i=1 and j=1.

Suppose that $\mathcal{D}_2$ is collected from a second level subsystem that is independent of the first level subsystem and $\mathcal{M}_1^{(2)}$ in M_2 is a newly defined candidate model class with new uncertain parameters $\boldsymbol{\theta}^{(2,1)}$. The posterior PDF $p(\boldsymbol{\theta}^{(2,1)}|\mathcal{D}_1, \mathcal{D}_2, \mathcal{M}_1^{(2)})=p(\boldsymbol{\theta}^{(2,1)}|\mathcal{D}_2, \mathcal{M}_1^{(2)})$ for $\mathcal{M}_1^{(2)}$ is given by (4.2) with the chosen prior PDF $p(\boldsymbol{\theta}^{(2,1)}|\mathcal{M}_1^{(2)})$. The evidence $p(\mathcal{D}_2|\mathcal{M}_1^{(2)})$, which is required for calculating the posterior probability $P(\mathcal{M}_1^{(2)}|\mathcal{D}_1,\mathcal{D}_2, M_2) = P(\mathcal{M}_1^{(2)}|\mathcal{D}_2, M_2)$ for $\mathcal{M}_1^{(2)}$, is given by (4.6) with i=2 and j=1.

Suppose that the third level subsystem contains the first level subsystem but not the second level subsystem. Assume that the first candidate model class $\mathcal{M}_1^{(3)}$ in M_3, with uncertain parameters $\boldsymbol{\theta}^{(3,1)}$ for the third level subsystem from which $\mathcal{D}_3$ is collected, is built-up by extending the model class $\mathcal{M}_1^{(1)}$ (i.e., existing parameters $\boldsymbol{\xi}^{(3,1)} = \boldsymbol{\theta}^{(1,1)}$) and $\boldsymbol{\varphi}(3, 1)$ are the new uncertain model parameters, so $\boldsymbol{\theta}^{(3,1)} = [\boldsymbol{\theta}^{(1,1)}, \boldsymbol{\varphi}^{(3,1)}]$. The posterior PDF $p(\boldsymbol{\theta}^{(3,1)}|\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{M}_1^{(3)})$ for $\mathcal{M}_1^{(3)}$ is given by (4.2) with the prior PDF $p(\boldsymbol{\theta}^{(3,1)}|\mathcal{D}_1, \mathcal{D}_2, \mathcal{M}_1^{(3)})=p(\boldsymbol{\theta}^{(1,1)}|\mathcal{D}_1, \mathcal{M}_1^{(1)})\ p(\boldsymbol{\varphi}^{(3,1)}|\mathcal{M}_1^{(3)})$ and so this posterior is independent of $\mathcal{D}_2$, as expected. The evidence $p(\mathcal{D}_1,\mathcal{D}_3|\mathcal{M}_1^{(3)})$, which is required for calculating the posterior probability $P(\mathcal{M}_1^{(3)}|\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, M_3) = P(\mathcal{M}_1^{(3)}|\mathcal{D}_1, \mathcal{D}_3, M_3)$ for $\mathcal{M}_1^{(3)}$, is equal to $p(\mathcal{D}_1|\mathcal{M}_1^{(3)})\ p(\mathcal{D}_3|\mathcal{D}_1,\mathcal{M}_1^{(3)})$ by (4.4) where $p(\mathcal{D}_3|\mathcal{D}_1,\mathcal{M}_1^{(3)})$ is given by (4.5) which becomes here:

$$p(\mathcal{D}_3 \mid \mathcal{D}_1, \mathcal{M}_1^{(3)}) = \int p(\mathcal{D}_3 \mid \boldsymbol{\theta}^{(3,1)}, \mathcal{M}_1^{(3)})\, p(\boldsymbol{\theta}^{(3,1)} \mid \mathcal{D}_1, \mathcal{D}_2, \mathcal{M}_1^{(3)})\, d\boldsymbol{\theta}^{(3,1)} \tag{4.7}$$

and $p(\mathcal{D}_1|\mathcal{M}_1^{(3)}) = p(\mathcal{D}_1|\mathcal{M}_1^{(1)})$, since 1) $\mathcal{M}_1^{(3)}$ is built-up by extending $\mathcal{M}_1^{(1)}$; 2) prior to the collection of $\mathcal{D}_3$, $\mathcal{D}_1$ is used to update $\mathcal{M}_1^{(1)}$. Recall that $p(\mathcal{D}_1|\mathcal{M}_1^{(1)})$ has already been determined.

Suppose that the fourth level subsystem is a combination of the first and second level subsystems but not the third one. Assume that the first candidate model class $\mathcal{M}_1^{(4)}$ in M_4, with uncertain parameters $\boldsymbol{\theta}^{(4,1)}$ for the fourth level subsystem from which $\mathcal{D}_4$ is collected, is built-up by using the model classes $\mathcal{M}_1^{(3)}$ and $\mathcal{M}_1^{(2)}$ (i.e., $\boldsymbol{\xi}^{(4,1)} = [\boldsymbol{\theta}^{(1,1)}, \boldsymbol{\theta}^{(2,1)}]$) and there are no new uncertain model parameters. Thus $\boldsymbol{\theta}^{(4,1)} = \boldsymbol{\xi}^{(4,1)} = [\boldsymbol{\theta}^{(1,1)}, \boldsymbol{\theta}^{(2,1)}]$ and $\boldsymbol{\Phi}^{(4,1)} = \boldsymbol{\varphi}^{(3,1)}$ since when updating $\mathcal{M}_1^{(3)}$, $\boldsymbol{\varphi}^{(3,1)}$ and $\boldsymbol{\theta}^{(1,1)}$ are both updated and $\mathcal{D}_1$ and $\mathcal{D}_3$ are used to update both of them. The posterior PDF $p(\boldsymbol{\theta}^{(4,1)}, \boldsymbol{\Phi}^{(4,1)}|\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{M}_1^{(4)}) = p(\boldsymbol{\theta}^{(1,1)}, \boldsymbol{\theta}^{(2,1)}, \boldsymbol{\varphi}^{(3,1)}|\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{M}_1^{(4)})$ for $\mathcal{M}_1^{(4)}$ is given by (4.3) with the prior PDF $p(\boldsymbol{\theta}^{(4,1)}, \boldsymbol{\Phi}^{(4,1)}|\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{M}_1^{(4)}) = p(\boldsymbol{\theta}^{(1,1)}, \boldsymbol{\varphi}^{(3,1)}|\mathcal{D}_1, \mathcal{D}_3, \mathcal{M}_1^{(3)})\, p(\boldsymbol{\theta}^{(2,1)}|\mathcal{D}_2, \mathcal{M}_1^{(2)})$. The evidence $p(\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4|\mathcal{M}_1^{(4)})$, which is required for calculating the posterior probability $P(\mathcal{M}_1^{(4)}|\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, M_4)$ for $\mathcal{M}_1^{(4)}$, is equal to $p(\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3|\mathcal{M}_1^{(4)})\, p(\mathcal{D}_4|\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{M}_1^{(4)})$ by (4.4) where $p(\mathcal{D}_4|\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{M}_1^{(4)})$ is given by (4.5) which becomes here:

$$p(\mathcal{D}_4 \mid \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{M}_1^{(4)}) = \int p(\mathcal{D}_4 \mid \boldsymbol{\theta}^{(4,1)}, \mathcal{M}_1^{(4)})\, p(\boldsymbol{\theta}^{(4,1)} \mid \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{M}_1^{(4)})\, d\boldsymbol{\theta}^{(4,1)} \tag{4.8}$$

where $p(\boldsymbol{\theta}^{(4,1)}|\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{M}_1^{(4)}) = p(\boldsymbol{\theta}^{(1,1)}|\mathcal{D}_1, \mathcal{D}_3, \mathcal{M}_1^{(3)})\, p(\boldsymbol{\theta}^{(2,1)}|\mathcal{D}_2, \mathcal{M}_1^{(2)})$ and $p(\boldsymbol{\theta}^{(1,1)}|\mathcal{D}_1, \mathcal{D}_3, \mathcal{M}_1^{(3)})$ is the marginal PDF of the posterior PDF $p(\boldsymbol{\theta}^{(3,1)}|\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{M}_1^{(3)})$ for $\mathcal{M}_1^{(3)}$ while $p(\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3|\mathcal{M}_1^{(4)}) = p(\mathcal{D}_1, \mathcal{D}_3|\mathcal{M}_1^{(3)})p(\mathcal{D}_2|\mathcal{M}_1^{(2)})$, since 1) $\mathcal{M}_1^{(4)}$ is built-up by using $\mathcal{M}_1^{(3)}$ and $\mathcal{M}_1^{(2)}$; 2) prior to the collection of $\mathcal{D}_4$, $\mathcal{D}_1$ and $\mathcal{D}_3$ are used to update $\mathcal{M}_1^{(3)}$ and $\mathcal{D}_2$ is used to update $\mathcal{M}_1^{(2)}$. Recall that $p(\mathcal{D}_1, \mathcal{D}_3|\mathcal{M}_1^{(3)})$ and $p(\mathcal{D}_2|\mathcal{M}_1^{(2)})$ have already been determined.

Suppose that the fifth level subsystem contains third and fourth level subsystems. Assume that the first candidate model class $\mathcal{M}_1^{(5)}$ in M_5, with uncertain parameters $\boldsymbol{\theta}^{(5,1)}$ for the fifth

level subsystem from which $\mathcal{D}_5$ is collected, is built-up by using the model class $\mathcal{M}_1^{(4)}$ with no new uncertain model parameters. Thus, $\boldsymbol{\theta}^{(5,1)} = \boldsymbol{\xi}^{(5,1)} = [\boldsymbol{\theta}^{(1,1)}, \boldsymbol{\theta}^{(2,1)}, \boldsymbol{\varphi}^{(3,1)}]$ since when updating $\mathcal{M}_1^{(4)}$, $\boldsymbol{\theta}^{(1,1)}$, $\boldsymbol{\theta}^{(2,1)}$ and $\boldsymbol{\varphi}^{(3,1)}$ are updated and $\mathcal{D}_1$, $\mathcal{D}_2$, $\mathcal{D}_3$ and $\mathcal{D}_4$ are used to update them. The posterior PDF $p(\boldsymbol{\theta}^{(5,1)} | \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{D}_5, \mathcal{M}_1^{(5)}) = p(\boldsymbol{\theta}^{(1,1)}, \boldsymbol{\theta}^{(2,1)}, \boldsymbol{\varphi}^{(3,1)} | \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{D}_5, \mathcal{M}_1^{(5)})$ for $\mathcal{M}_1^{(5)}$ is given by (4.2) with the prior PDF $p(\boldsymbol{\theta}^{(5,1)} | \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{M}_1^{(5)}) = p(\boldsymbol{\theta}^{(1,1)}, \boldsymbol{\theta}^{(2,1)}, \boldsymbol{\varphi}^{(3,1)} | \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{M}_1^{(4)})$. The evidence $p(\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{D}_5 | \mathcal{M}_1^{(5)})$, which is required for calculating the posterior model probability $P(\mathcal{M}_1^{(5)} | \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{D}_5, M_5)$ for $\mathcal{M}_1^{(5)}$, is equal to $p(\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4 | \mathcal{M}_1^{(5)})\, p(\mathcal{D}_5 | \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{M}_1^{(5)})$ by (4.4) where $p(\mathcal{D}_5 | \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{M}_1^{(5)})$ is given by (4.5) which becomes here:

$$p(\mathcal{D}_5 \mid \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{M}_1^{(5)}) = \int p(\mathcal{D}_5 \mid \boldsymbol{\theta}^{(5,1)}, \mathcal{M}_1^{(5)})\, p(\boldsymbol{\theta}^{(5,1)} \mid \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{M}_1^{(5)})\, d\boldsymbol{\theta}^{(5,1)} \qquad (4.9)$$

where $p(\boldsymbol{\theta}^{(5,1)} | \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{M}_1^{(5)}) = p(\boldsymbol{\theta}^{(1,1)}, \boldsymbol{\theta}^{(2,1)}, \boldsymbol{\varphi}^{(3,1)} | \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{M}_1^{(4)})$ while $p(\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4 | \mathcal{M}_1^{(5)}) = p(\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4 | \mathcal{M}_1^{(4)})$, since 1) $\mathcal{M}_1^{(5)}$ is built-up by using $\mathcal{M}_1^{(4)}$; 2) prior to the collection of $\mathcal{D}_5$, $\mathcal{D}_1$, $\mathcal{D}_2$, $\mathcal{D}_3$ and $\mathcal{D}_4$ are used to update $\mathcal{M}_1^{(4)}$. Recall that $p(\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4 | \mathcal{M}_1^{(4)})$ has already been determined.

5 ILLUSTRATIVE EXAMPLE BASED ON A VALIDATION CHALLENGE PROBLEM

For illustration, the static-frame validation challenge problem (Babuška et al. 2008) is considered. It is one of the problems presented at the Validation Challenge Workshop at Sandia National Laboratory on May 27-29, 2006. The purpose of this particular challenge problem is to predict the probability of the event F (*regulatory assessment*): $|w_p| \geq 3$mm, where w_p is the vertical displacement of the midpoint P of beam 4 of the frame structure (our target system) shown in Figure 1 of Babuška et al (2008) and Figure 5.1 in this report. The structure is subjected to a uniform load q = 6kN/m on beam 4. Information regarding the geometry of the frame structure is shown in Table 1 of Babuška et al (2008) and in Tables 5.1 and 5.2 in this report. Also, in the definition of the challenge problem, the structure is given to be linear elastic with a one-dimensional tension model for each of the rods and a one-dimensional Bernoulli beam model for the bending of the beam. The coupling of bending and compression is given to be negligible for beam 4. It is given that all the bars are made of the same inhomogeneous material but come from independent sources and so can have variable material properties; in fact, the only uncertainty considered in this challenge problem is Young's modulus E (or compliance $S=1/E$) along each of the bars. Given Young's modulus variation along each of the bars, w_p can be predicted using the equations in Babuška et al (2008) and in Appendix A in this report.

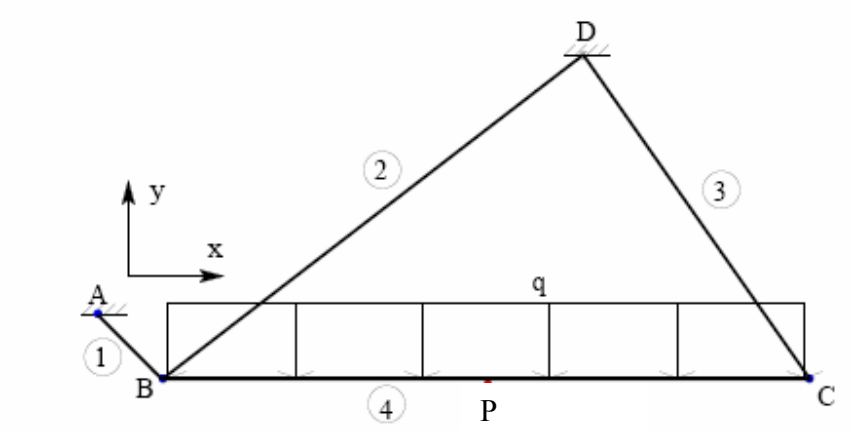

Figure 5.1: The frame structure for the prediction case

Table 5.1 Joint coordinates of the frame structure (prediction case) (from Babuska et al. 2008)

Point	x(m)	y(m)
A	0	0.2
B	0.2	0
C	2.2	0
D	1.5	1.0

Table 5.2: Geometry information of the bars, tensile force and bending moment along the bars of the frame structure (prediction case) (from Babuska et al. 2008)

Bar	A(cm^2)	I(cm^4)	Tensile force (kN)	Moment (kNm)
1	16	Not required	2.214	
2	16	Not required	7.274	
3	16	Not required	7.324	
4	80	5333	-4.200	3(2.2-x)(x-0.2)

The simulated experiments are set up to resemble a typical situation in which data are collected from a hierarchy of successively more complex subsystems that become "closer" to

the final system and the amount of data reduces in the higher levels of the hierarchy. Data from three experiments which involve systems of increasing complexity are presented as part of the challenge problem:

1. The first experiment is referred to as the *calibration* experiment. It involves N_c bars where each bar has a cross section area A_c =4.0cm^2 and length L_c= 20 cm, is fixed rigidly at one end and is loaded by a tensile axial force F_c=1.2kN at the other end. The available data $\mathcal{D}_1$ from this experiment are the elongation $\delta L_c^{(i)}$, i=1,2…, N_c, of the bars from the initial length and the Young's modulus $E_c^{(i)}(L_c/2)$ at the midpoint of the bars.
2. The second experiment is referred to as the *validation* experiment. The set-up is similar to the first experiment. The only difference is that the bars have longer length L_v= 80cm and only the total elongation $\delta L_v^{(i)}$, i=1,2…, N_v, is measured. Let $\mathcal{D}_2$ denote the data in this case.
3. The third experiment is referred to as the *accreditation* experiment. It involves a frame structure (Figure 4 in Babuška et al (2008) and Figure 5.2 in this report) subject to a point load F_a=6kN at the midpoint Q of bar 1. The available data $\mathcal{D}_3$ are the vertical displacement $w_a^{(i)}$, i=1,…, N_a, of the point Q. Information regarding the geometry of the frame is shown in Table 3 in Babuška et al (2008) and Tables 5.3 and 5.4 in this report. Notice that the system here is not a subsystem of the target system.

Table 5.3 Joint coordinates of the frame structure in the third experiment (from Babuska et al. 2008)

Point	x(m)	y(m)
A	0	0.5
B	0	0
C	0.5	0
D	0.5	0.5

Table 5.4 Geometry information of the bars of the frame structure in the third experiment (from Babuska et al. 2008)

Bar	A(cm^2)	I(cm^4)	Tensile force (kN)	Moment (kNm)
1	16	333.3	-3.000	$3x$, $0<x<0.25$ $3(0.5-x)$,$0.25<x<0.5$
2	16	Not required	4.243	
3	16	Not required	0.000	
4	20	Not required	4.243	

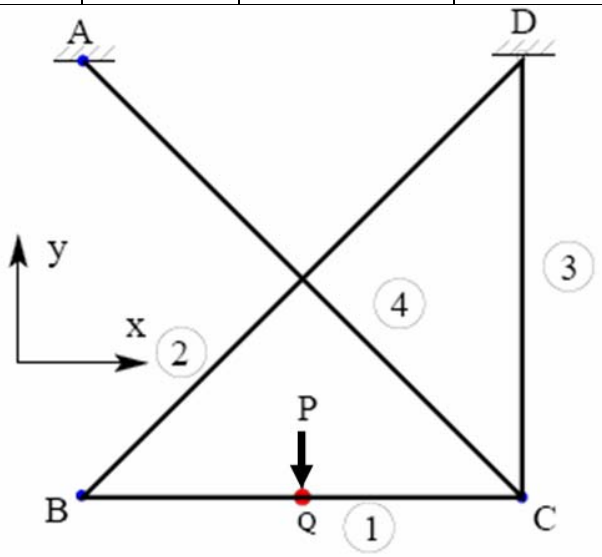

Figure 5.2: The frame structure in the accreditation experiment

Data collected from the above three experiments are shown in Tables 5.5, 5.6 and 5.7 respectively. Three cases of N_c, N_v and N_a, as shown in Table 5.8, are considered. For instance, for case 1, $N_c = 5$, $N_v = 2$ and $N_a = 1$ correspond to the first five, the first two and the first of the measurements listed in Tables 5.5, 5.6 and 5.7 respectively. A superscript is added to $\mathcal{D}_i$ to denote different data cases. For instance, $\mathcal{D}_1^{(1)}$ denotes data collected from the calibration experiment with $N_c = 5$, $\mathcal{D}_2^{(1)}$ denotes data collected from the validation experiment with $N_v = 2$ and $\mathcal{D}_3^{(1)}$ denotes data collected from the accreditation experiment with $N_a = 1$. Given Young's modulus of each of the bars, the elongation of the bars in the first and second experiment and the vertical displacement in the third experiment can be predicted using the equations in Babuška et al (2008) and in Appendix A in this report. For convenience, the superscripts in $\boldsymbol{\theta}^{(i,j)}$ are omitted in this section.

Table 5.5 Data $\mathcal{D}_1$ for the calibration experiment (from Babuska et al. 2008)

Sample i	$\delta L_c^{(i)}$ ($\times 10^{-2}$mm)	$E(L_c/2)$ (GPa)
1	5.15	13.26
2	5.35	10.86
3	5.24	14.77
4	5.51	10.94
5	5.14	11.05
6	5.38	11.06
7	4.97	11.97
8	5.41	11.66
9	4.95	12.09
10	5.42	11.30
11	5.47	10.98
12	5.74	11.92
13	5.36	11.12

14	5.42	12.00
15	5.34	10.98
16	5.60	10.71
17	5.06	10.91
18	4.99	11.89
19	5.22	11.43
20	5.57	10.87
21	5.28	11.75
22	5.10	13.47
23	5.48	11.44
24	5.35	12.44
25	4.92	12.13
26	5.51	11.38
27	5.27	10.75
28	5.14	11.92
29	5.61	10.82
30	5.56	11.04

Table 5.6 Data $\mathcal{D}_2$ for the validation experiment (from Babuska et al. 2008)

Sample i	$\delta L_v^{(i)}$ ($\times 10^{-1}$mm)
1	2.01
2	2.06
3	2.01
4	2.08
5	2.04
6	2.01
7	2.06
8	2.11
9	1.98
10	2.08

Table 5.7 Data $\mathcal{D}_3$ for the accreditation experiment (from Babuska et al. 2008)

Sample i	$u^{(i)}$ ($\times 10^{-1}$mm)
1	-6.50
2	-6.73

Table 5.8 Number of samples for different cases

Case	N_c	N_v	N_a
1	5	2	1
2	20	4	1
3	30	10	2

5.1 Using data $\mathcal{D}_1$ from the calibration experiment

For the quantification of the uncertainties in Young's modulus $E(x)$, $0 \le x \le L$, of a bar of length L using data $\mathcal{D}_1$ from the calibration experiment, a set M_1 of four candidate model classes $\mathcal{M}_j^{(1)}$, j=1,2,3,4, is considered as follows:

Model class $\mathcal{M}_1^{(1)}$: The compliance $S(x)=S=1/E$ is constant along a bar and the value for each bar is assumed to be a sample from a Gaussian distribution with mean μ_s and variance σ_s^2. The elongation δL_c of a bar of length L_c is given by $\delta L_c = F_c L_c S/A_c + \varepsilon_c$ where ε_c is the prediction error, assumed to follow a Gaussian distribution with mean zero and variance σ_ε^2. The term ε_c is needed since from $\mathcal{D}_1$, it can be seen that δL_c is obviously not proportional to S. The prior PDF for $\boldsymbol{\theta} = [\mu_s \; \sigma_s^2 \; \sigma_\varepsilon^2]^T$ is chosen as three independent probability distributions: μ_s follows a truncated Gaussian distribution (constrained to be positive) which is proportional to a Gaussian distribution with mean equal to the sample mean of measurements of the mid-point compliance $S_c(L_c/2)$ and c.o.v. (coefficient of variation) of 1.0; σ_s^2 follows an inverse gamma distribution with mean μ equal to the sample variance of measurements of $S_c(L_c/2)$ and c.o.v. δ =1.0, i.e.,

$p(\sigma_s^2) \propto (\sigma_s^2)^{-\alpha-1}\exp(-\beta/\sigma_s^2)$ where $\alpha=\delta^{-2}+2, \beta=\mu(\alpha-1)$; l_s follows an inverse gamma distribution with mean equal to 10^{-11} m^2 (slightly more than the mean-square of the elongation measurements) and c.o.v. equal to 1.0. The prior c.o.v. of all of the uncertain parameters is chosen to be 1.0 to reflect a large uncertainty in the values of these parameters. If the type of material of the bars had been known in advance, the prior mean for μ_s could have been chosen to be the nominal value of the compliance obtained from previous tests performed on such material and the prior mean for σ_s^2 could have been chosen to be the prior mean for μ_s multiplied by a coefficient of variation chosen to reflect previously observed variability in the material compliance.

Model class $\mathcal{M}_2^{(1)}$: The compliance $S(x)$ is assumed to follow a stationary Gaussian random field with mean μ_s and correlation function $\text{Cov}(S(x_1),S(x_2)|\sigma_s^2, l_s, r)= \sigma_s^2\exp(-(|x_1-x_2|/l_s)^r)$ where r is equal to 1. The prior PDF for $\boldsymbol{\theta} =[\mu_s\ \sigma_s^2\ l_s]^T$ is chosen as three independent distributions: the prior PDFs for the mean μ_s and the variance σ_s^2 follow the same distributions as in $\mathcal{M}_1^{(1)}$; the correlation length l_s follows a uniform distribution on the interval $[10^{-5}L, L]$ where we choose L=0.5m to give a reasonable range.

Model class $\mathcal{M}_3^{(1)}$: Everything is the same as $\mathcal{M}_2^{(1)}$ except r is equal to 2.

Model class $\mathcal{M}_4^{(1)}$: Everything is the same as $\mathcal{M}_2^{(1)}$ and $\mathcal{M}_3^{(1)}$ except that r is uncertain. The prior PDF for $\boldsymbol{\theta} =[\mu_s\ \sigma_s^2\ l_s\ r]^T$ is chosen as four independent distributions: μ_s, σ_s^2, l_s follow the same distributions as in $\mathcal{M}_2^{(1)}$ and $\mathcal{M}_3^{(1)}$ and r follows a uniform distribution on [0.5, 3].

Babuška et al. (2008) and Grigoriu and Field (2008) also study the static-frame challenge problem using Bayesian updating. The perfectly-correlated Gaussian model for the compliance in $\mathcal{M}_1^{(1)}$ and the partially-correlated stationary Gaussian random field model for the compliance

in $\mathcal{M}_3^{(1)}$ are also considered in Babuška et al. (2008). The partially-correlated Gaussian random field model for the compliance in $\mathcal{M}_2^{(1)}$ is considered in Grigoriu and Field (2008). $\mathcal{M}_2^{(1)}$ and $\mathcal{M}_3^{(1)}$ are included here for comparison purposes only. In practice, when r is uncertain, only $\mathcal{M}_4^{(1)}$ needs to be considered. For r=0, the correlation coefficient between the compliance at one position on the bar and that at another position is always equal to e^{-1}. This model is thought to be unreasonably constrained and so it is not considered. This is why the lower bound of r is taken to be positive.

Babuška et al. (2008) find point estimates of μ_s and σ_s^2 in $\mathcal{M}_1^{(1)}$ by matching the first two sample moments of the compliance data $S_c^{(i)}(L_c/2)$, i=1,2…, N_c, and l_s in $\mathcal{M}_3^{(1)}$ by matching the sample variance of the elongation data $\delta L_c^{(i)}$, i=1,2…, N_c, and the sample covariance of $\delta L_c^{(i)}$ and $S_c^{(i)}(L_c/2)$, i=1,2…, N_c. Grigoriu and Field (2008) approximate the uncertain parameters by point estimates by matching the sample moments similar to Babuška et al (2008) except that they do not consider the sample covariance of $\delta L_c^{(i)}$ and $S_c^{(i)}(L_c/2)$, i=1,2…, N_c. In Grigoriu and Field (2008), the uncertainties in the model parameters μ_s, σ_s^2 and l_s are not considered and not directly propagated into the predictions so probabilistic information in these parameters is not subsequently characterized. Babuška et al. (2008) quantify the uncertainties by using kernel density estimation to reconstruct the joint PDF of δL_c and $S_c(L_c/2)$ from the data for $\delta L_c^{(i)}$ and $S_c^{(i)}(L_c/2)$ and then using the bootstrapping method to generate additional "data".

Appropriate quantification of uncertainties in the parameters (i.e. obtaining complete probabilistic information in terms of the posterior PDF for each model class) is desirable since it significantly affects the effectiveness and robustness of model class updating, comparison and validation, as well as the prediction of the responses and the failure probability of the target

structure. Here we use the challenge problem to illustrate how the uncertainties can be quantified appropriately and effectively by exploiting the *full* power of Bayesian analysis using the proposed concept of hierarchical stochastic system model classes and recently-developed computational tools. Later, when we present the analysis results, it will be clear that given the calibration data, the uncertainty in μ_s is quite small but the uncertainties in other parameters and data-induced correlation between the parameters are not negligible; the complete probabilistic information is, however, encapsulated in the samples from the posterior.

To quantify the uncertainties of $\boldsymbol{\theta}$ using Bayesian analysis and $\mathcal{D}_1^{(l)}$, the elongation data $\delta L_c^{(i)}$ and the compliance data $S_c^{(i)}(L_c/2)$, i=1,2,…, N_c should be considered simultaneously since they are correlated to each other given $\boldsymbol{\theta}$ and the proposed model classes.

The posterior PDF for model class $\mathcal{M}_j^{(1)}$, for j=1,2,3,4, is given by Bayes' Theorem: $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{M}_j^{(1)}) = p(\mathcal{D}_1^{(l)}|\boldsymbol{\theta},\mathcal{M}_j^{(1)})p(\boldsymbol{\theta}|\mathcal{M}_j^{(1)})/p(\mathcal{D}_1^{(l)}|\mathcal{M}_j^{(1)})$ where the prior PDF $p(\boldsymbol{\theta}|\mathcal{M}_j^{(1)})$ is described above and the likelihood function $p(\mathcal{D}_1^{(l)}|\boldsymbol{\theta},\mathcal{M}_j^{(1)})$ is given by the following. The likelihood function for $\mathcal{M}_1^{(1)}$ is:

$$p(\mathcal{D}_1^{(l)} \mid \boldsymbol{\theta}, \mathcal{M}_1^{(1)}) = \frac{1}{[2\pi \mid \mathbf{C}(\sigma_s^{\,2}, \sigma_\varepsilon^{\,2}) \mid^{1/2}]^{N_c}} \exp\left(-\frac{1}{2}\sum_{i=1}^{N_c}[\mathbf{y}^{(i)} - \boldsymbol{\mu}(\mu_s)]^T \mathbf{C}^{-1}(\sigma_s^{\,2}, \sigma_\varepsilon^{\,2})[\mathbf{y}^{(i)} - \boldsymbol{\mu}(\mu_s)]\right) \tag{5.1}$$

where

$$\mathbf{y}^{(i)} = \begin{bmatrix} \delta L_c^{(i)} \\ S_c^{(i)}(L_c/2) \end{bmatrix} \tag{5.2}$$

$$\boldsymbol{\mu}(\mu_s) = \begin{bmatrix} \frac{F_c L_c}{A_c} \\ 1 \end{bmatrix} \mu_s \tag{5.3}$$

$$\mathbf{C}(\sigma_s{}^2, \sigma_\varepsilon{}^2) = \begin{bmatrix} (\frac{F_c L_c}{A_c})^2 \sigma_s{}^2 + \sigma_\varepsilon{}^2 & \frac{F_c L_c}{A_c} \sigma_s{}^2 \\ \frac{F_c L_c}{A_c} \sigma_s{}^2 & \sigma_s{}^2 \end{bmatrix} \tag{5.4}$$

For $\mathcal{M}_2^{(1)}$ and $\mathcal{M}_3^{(1)}$, the likelihood function is the same as that for $\mathcal{M}_4^{(1)}$ with r=1 and 2, respectively. The likelihood function for $\mathcal{M}_4^{(1)}$ is given by:

$$p(\mathcal{D}_1^{(l)} \mid \boldsymbol{\theta}, \mathcal{M}_4^{(1)}) = \frac{1}{[2\pi\sigma_s{}^2 \mid \mathbf{C}(l_s, r) \mid^{1/2}]^{N_c}} \exp(-\frac{1}{2\sigma_s{}^2} \sum_{i=1}^{N_c} [\mathbf{y}^{(i)} - \boldsymbol{\mu}(\mu_s)]^T \mathbf{C}^{-1}(l_s, r) [\mathbf{y}^{(i)} - \boldsymbol{\mu}(\mu_s)]) \tag{5.5}$$

where $\mathbf{y}^{(i)}$ and $\boldsymbol{\mu}(\theta_1)$ are given by (5.2) and (5.3) and $\mathbf{C}(l_s, r)$ is given by:

$$\mathbf{C}(l_s, r) = \begin{bmatrix} C_{11}(l_s, r) & C_{12}(l_s, r) \\ C_{12}(l_s, r) & 1 \end{bmatrix} \tag{5.6}$$

where the entries C_{11} and C_{12} of $\mathbf{C}$ are given by:

$$\begin{aligned} C_{11}(l_s, r) &= (\frac{F_c}{A_c})^2 \int_0^{L_c} \int_0^{L_c} \frac{\mathrm{Cov}(S(x_1), S(x_2) \mid \sigma_s{}^2, l_s, r)}{\sigma_s{}^2} dx_1 dx_2 \\ &= (\frac{F_c}{A_c})^2 \int_0^{L_c} \int_0^{L_c} \exp(-(\frac{\mid x_1 - x_2 \mid}{l_s})^r) dx_1 dx_2 \\ &= 2(\frac{F_c}{A_c})^2 \int_0^{L_c} (L_c - x) \exp(-(\frac{x}{l_s})^r) dx \end{aligned} \tag{5.7}$$

$$C_{12}(l_s,r)=\frac{F_c}{A_c}\int_0^{L_c}\frac{\mathrm{Cov}(S(x),S(L_c/2)\,|\,\sigma_s^{\,2},l_s,r)}{\sigma_s^{\,2}}dx=\frac{F_c}{A_c}\int_0^{L_c}\exp(-(\frac{|x-L_c/2|}{l_s})^r)dx \tag{5.8}$$

For $\mathcal{M}_2^{(1)}$, r is equal to 1 and thus the above integrals can be evaluated analytically to give:

$$C_{11}(l_s,1)=2(\frac{F_c}{A_c})^2 l_s(L_c-l_s+l_s\exp(-\frac{L_c}{l_s})) \tag{5.9}$$

$$C_{12}(l_s,1)=2(\frac{F_c}{A_c})l_s(1-\exp(-\frac{L_c}{2l_s})) \tag{5.10}$$

For $\mathcal{M}_3^{(1)}$, r is equal to 2 and thus the above integrals can be expressed in terms of the error function to give:

$$C_{11}(l_s,2)=(\frac{F_c}{A_c})^2[L_c l_s\sqrt{\pi}\,\mathrm{erf}(\frac{L_c}{l_s})-l_s^{\,2}(1-\exp(-(\frac{L_c}{l_s})^2))] \tag{5.11}$$

$$C_{12}(l_s,2)=\frac{F_c}{A_c}l_s\sqrt{\pi}\,\mathrm{erf}(\frac{L_c}{2l_s}) \tag{5.12}$$

Since the computer always has a precision limit in representing numbers, when performing the analysis, we make sure l_s is such that $\mathbf{C}(l_s,r)$ is positive definite, i.e., $C_{11}(l_s,r)$ and $|\mathbf{C}(l_s,r)|=C_{11}(l_s,r)-C_{12}^{\,2}(l_s,r)$ are both positive. The interval of l_s for its prior PDF in $\mathcal{M}_2^{(1)}$, $\mathcal{M}_3^{(1)}$ and $\mathcal{M}_4^{(1)}$ satisfies this constraint.

Before updating the above model classes, it is instructive to first study the effect of using only the compliance measurements $S_c^{(i)}(L_c/2)$, i=1,…, N_c (denote this data as $\mathcal{D}_0$ which is a subset of $\mathcal{D}_1$) on the quantification of the uncertainties of the mean μ_s and variance σ_s^2 of the

stationary Gaussian random field models for the compliance. Let $\mathcal{M}_0$ represent the model class for the mid-point compliance which uses a Gaussian model with mean μ_s and variance σ_s^2 where the prior PDF for μ_s and $\ln\sigma_s^2$ is chosen as the product of independent noninformative priors that are constant over a broad range. The likelihood function $p(\mathcal{D}_0|\, \mu_s, \sigma_s^2, \mathcal{M}_0)$ is given by:

$$p(\mathcal{D}_0 \mid \mu_s, \sigma_s, \mathcal{M}_0) = \frac{1}{(\sqrt{2\pi\sigma_s^{\,2}})^{N_c}} \exp\left(-\frac{1}{2\sigma_s^{\,2}} \sum_{i=1}^{N_c} (S_c^{\,(i)} - \mu_s)^2\right) \tag{5.13}$$

The MAP (maximum a posteriori) estimates $\hat{\mu}_s$, $\hat{\sigma}_s^{\,2}$ (i.e., the values of μ_s, σ_s^2 that globally maximize $p(\mu_s, \sigma_s^2|\, \mathcal{D}_0, \mathcal{M}_0)$) are equal to the sample mean and very close to the sample variance of the $S_c^{(i)}$, i=1,…, N_c, respectively:

$$\begin{bmatrix} \hat{\mu}_s \\ \hat{\sigma}_s^{\,2} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{N_c} \displaystyle\sum_{i=1}^{N_c} S_c^{(i)} \\ \dfrac{1}{(N_c+1)} \displaystyle\sum_{i=1}^{N_c} (S_c^{(i)} - \hat{\mu}_s)^2 \end{bmatrix} \tag{5.14}$$

It can be shown that given $\mathcal{M}_0$ and $\mathcal{D}_0$, $(\mu_s - \hat{\mu}_s)\sqrt{N_c}\,/v_s$ follows Student's t distribution with N_c-1 degrees of freedom and σ_s follows a distribution as follows (Box & Tiao 1973):

$$p(\sigma_s \mid \mathcal{D}_0, \mathcal{M}_0) \propto \sigma_s^{\,-N_c} \exp\left(-\frac{\displaystyle\sum_{i=1}^{N_c} (S_c^{(i)} - \hat{\mu}_s)^2}{2\sigma_s^{\,2}}\right) \tag{5.15}$$

Thus μ_s has a mean equal to the sample mean $\hat{\mu}_s$ and c.o.v. equal to δ_1 as follows (which is very close to the sample c.o.v. of $S_c^{(i)}$ divided by $\sqrt{N_c}$ for sufficiently large N_c):

$$\delta_1 = \frac{1}{\sqrt{N_c}}\left(\sqrt{\frac{1}{N_c - 1}\sum_{i=1}^{N_c}(S_c^{(i)} - \hat{\mu}_s)^2} \,/\, \hat{\mu}_s\right)\sqrt{\frac{N_c - 1}{N_c - 3}} \tag{5.16}$$

From (5.15), it can be shown that given $\mathcal{M}_0$ and $\mathcal{D}_0$, σ_s^2 follows an inverse gamma distribution:

$$p(\sigma_s^2 \mid \mathcal{D}_0, \mathcal{M}_0) \propto (\sigma_s^2)^{\frac{-(N_c-1)}{2}-1} \exp\left(-\frac{\sum_{i=1}^{N_c}(S_c^{(i)} - \hat{\mu}_s)^2}{2\sigma_s^2}\right) \tag{5.17}$$

Thus, given $\mathcal{M}_0$ and $\mathcal{D}_0$, σ_s^2 has a c.o.v. equal to δ_2 given as follows for N_c >5:

$$\delta_2 = \frac{\sqrt{2}}{\sqrt{N_c - 5}} \tag{5.18}$$

For large N_c, the c.o.v. of both μ_s and σ_s^2 decrease at the rate of approximately $\sqrt{N_c}$. Table 5.9 shows the results for the MAP estimates $\hat{\mu}_s$, $\hat{\sigma}_s^2$ based on data $\mathcal{D}_0$ along with the coefficients of variation δ_1 and δ_2 for μ_s and σ_s^2 respectively, from their posterior PDFs.

Table 5.9 Statistical results obtained using data $\mathcal{D}_0$

Case	$\hat{\mu}_s$	$\hat{\sigma}_s^2$	δ_1	δ_2
1	8.34×10^{-11}	1.00×10^{-22}	5.4%	63.2%
2	8.68×10^{-11}	3.85×10^{-23}	1.6%	31.6%
3	8.64×10^{-11}	3.60×10^{-23}	1.3%	25.8%

Tables 5.10, 5.11 and 5.12 shows the statistical results using the calibration data $\mathcal{D}_1^{(1)}$, $\mathcal{D}_1^{(2)}$ and $\mathcal{D}_1^{(3)}$ for three cases of N_c = 5, 20 and 30, respectively. The (j+1)-th column gives the results

obtained using a full Bayesian analysis for model class $\mathcal{M}_j^{(1)}$, j=1,2,3,4. Usually stochastic simulation methods are to be preferred because they are applicable for any case and are more robust but the results obtained by Laplace's asymptotic approximation are also given this report because of their common use in past work (e.g., Beck & Katafygiotis 1991, 1998, Beck & Yuen 2004). For the stochastic simulation methods, we used the Hybrid Gibbs TMCMC algorithm presented in Appendix B for simulating samples from the posterior $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{M}_j^{(1)})$ and for calculating the evidence $p(\mathcal{D}_1^{(l)}|\mathcal{M}_j^{(1)})$ which is required for the calculation of the probability $P(\mathcal{M}_j^{(1)}|\mathcal{D}_1^{(l)},M)$ of each model class conditioned on the data $\mathcal{D}_1^{(l)}$. This algorithm is used for simulating samples from the posterior $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{M}_j^{(1)})$ because of its ability to handle the case where we do not know apriori whether there may be several separated neighborhoods of high probability regions of $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{M}_j^{(1)})$ between which the transition using a Markov chain of samples is not efficient. This algorithm is in general more efficient than the common Metropolis-Hastings algorithm for problems in higher dimensions. For the rest of this report, for the tables where both the results obtained using Laplace's asymptotic approximation and stochastic simulation methods are shown, "AA" is used to refer the results obtained using the former and "SS" is refer to denote the results obtained using the latter. The results are obtained using stochastic simulation methods if neither "AA" nor "SS" appears in the tables.

The second row of the tables gives the MAP (maximum a posteriori) estimate $\boldsymbol{\theta}_{\mathbf{MAP}}$ (that is, $\boldsymbol{\theta}$ that globally maximizes $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{M}_j^{(1)})$). The results of the third, sixth to ninth, and eleventh and twelveth rows are obtained using stochastic simulation methods. The third row gives the mean (the number before the semicolon), c.o.v. (the number after the semicolon) and the correlation coefficient matrix R from the posterior samples for $\boldsymbol{\theta}$ where the (i,j) entry of R is the

correlation coefficient between θ_i and θ_j. Only the upper diagonal entries of R are presented since it is symmetric. Compared with the prior uncertainty in the parameters, the posterior (updated) uncertainty is reduced since the data provide information about these parameters. For all three data cases and four model classes, μ_s has a lot smaller uncertainty than the other parameters which have significant uncertainties. It can be seen that the posterior mean of σ_s^2 given data $\mathcal{D}_1^{(l)}$ is quite different from the sample variance of the compliance measurements $S_c^{(i)}$, i=1,…, N_c since the elongation data $\delta L_c^{(i)}$ in $\mathcal{D}_1^{(l)}$ give extra information about this parameter. Because the challenge problem assumes an exact theory for the deformation analysis, prediction errors for each model class are accounted for by the modeling parameters such as σ_s^2. In general, prediction errors can be explicitly accounted for by adding them to the output equation (Beck and Katafygiotis 1998), as done in $\mathcal{M}_1^{(1)}$.

It can be seen from the correlation coefficient matrix that there is only weak correlation between pairs of parameters, although one must be careful since a small correlation coefficient between two uncertain parameters only implies weak *linear* dependence and does not necessarily imply weak dependence between them unless the parameters are jointly Gaussian. A simple example for this is $W=Z^2$ and a standard normal variable Z which are uncorrelated but strongly dependent. To investigate dependence between different pairs of parameters, sample plots of some pairs of the components of $\boldsymbol{\theta}$ from the posterior $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$ obtained using stochastic simulation methods are shown in Figures 5.3-5.5 (for j=2), Figures 5.6-5.8 (for j=3) and Figures 5.9-5.11 (j=4). Each axis corresponds to an uncertain parameter θ_i divided by its posterior mean μ_i given $\mathcal{D}_1^{(l)}$ and a specific model class $\mathcal{M}_j^{(1)}$, which can be estimated as follows:

$$\mu_i = E[\theta_i \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)}] \approx \frac{1}{K}\sum_{k=1}^{K} \theta_i^{(k)} \tag{5.19}$$

where $[\theta_i^{(1)},...,\theta_i^{(K)}]$ are K posterior samples for θ_i from $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$. All the other parameters have significantly larger uncertainties than θ_1. The reduction of parameter uncertainties with increasing amount of data is also obvious from these figures. It can be seen that $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_2^{(1)})$ and $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_3^{(1)})$ are not close to a multivariate Gaussian PDF, especially when the amount of data is very small for $\mathcal{M}_2^{(1)}$ and $\mathcal{M}_3^{(1)}$ (e.g., case 1 where N_c=5); $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_4^{(1)})$ departs substantially from a multivariate Gaussian PDF and is of a very complex shape. For $\mathcal{M}_4^{(1)}$, the samples for r show truncation due to the choice of truncated uniform priors for r.

Figures 5.12, 5.13 and 5.14 give the histograms of posterior samples for r from $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_4^{(1)})$. These figures suggest that $p(r|\mathcal{D}_1^{(l)}, \mathcal{M}_4^{(1)})$ is multi-modal and every value of r is of non-negligible plausibility. The above results exhibit the strength of the stochastic simulation method in capturing the full characteristics of the complex posterior PDF $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$ represented by the generated posterior samples.

The stochastic simulation estimate for log evidence, posterior mean of the log likelihood function (a datafit measure), expected information gain and the probability $P(\mathcal{M}_j^{(1)}|\mathcal{D}_1^{(l)}, M_1)$ of the model classes are shown in the sixth through ninth rows, respectively, of Tables 5.10, 5.11 and 5.12. Based on the calibration data, for all 3 data cases, $\mathcal{M}_1^{(1)}$ is very improbable compared with the other model classes $\mathcal{M}_2^{(1)}$, $\mathcal{M}_3^{(1)}$ and $\mathcal{M}_4^{(1)}$ which have similar posterior probabilities and have essentially the same posterior mean of the log likelihood function which shows that they give a similar fit to the data on average and they also have similar expected information gains.

The fourth row of Tables 5.10-5.12 gives the estimate for the log evidence $\ln p(\mathcal{D}_1^{(l)}|\mathcal{M}_j^{(1)})$ obtained by Laplace's asymptotic approximation assuming there is only one global optimum for $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$. Although this assumption is not true for $\mathcal{M}_4^{(1)}$, as shown by Figures 12-14, all these estimates agree well with the corresponding stochastic simulation estimates. Using these asymptotic log evidence estimates, the probability $P(\mathcal{M}_j^{(1)}|\mathcal{D}_1^{(l)},M_1)$ of each model class conditioned on the data $\mathcal{D}_1$ is computed as shown in the fifth row. Compared with the results obtained using stochastic simulation methods, as expected, the posterior probability estimate obtained by Laplace's asymptotic approximation for model class $\mathcal{M}_4^{(1)}$ is not accurate since $\mathcal{M}_4^{(1)}$ is almost unidentifiable. It can be seen that the accuracy of the estimates obtained by Laplace's asymptotic approximation for the other model classes increases as the amount of data increases. Once again, the results show that stochastic simulation methods are preferable.

Grigoriu and Field (2008) perform model selection by calculating the posterior model probabilities of the MLE (maximum likelihood estimate) models (rather than the posterior probability for the whole model class) in which the modeling parameters are obtained by matching the moments calculated from the data. Such an approach considers the magnitude of the likelihood functions of the MLE models and no uncertainties in the parameters are considered when performing model selection. The fact that there exists many plausible models in a model class is not considered, in contrast to our full Bayesian treatment. In particular, when the evidence for the model class is not employed, there is no automatic downgrading of more "complex" models that extract more information from the data, so this can lead to what is commonly called "data overfitting" (Bishop 2006). Note that one cannot simply count the number of uncertain parameters in a model class to judge reliably its complexity; for example,

one should use the evidence for the model class and not the simplified version known as BIC (Bayesian information criterion) for model selection (Beck and Yuen 2004, Muto and Beck 2008).

For each of the four model classes $\mathcal{M}_j^{(1)}$, given $\boldsymbol{\theta}$, it can be shown that the response w_p of interest for the target frame structure follows a Gaussian distribution with mean $\mu_p = K_p \mu_s$ and variance $\sigma_p^2 = \sigma_s^2 V_{p,1}$ for $\mathcal{M}_1^{(1)}$ and $\sigma_s^2 V_{p,j}(l_s)$ for $\mathcal{M}_j^{(1)}$, j=2,3 and $\sigma_s^2 V_{p,j}(l_s, r)$ for $\mathcal{M}_4^{(1)}$ where K_p and $V_{p,j}$ are given as follows (the values of the force $F_{p,i}$, the cross-sectional area $A_{p,i}$ and the length $L_{p,i}$ of rod i and the cross-sectional moment of inertia I_p of beam 4 are given in Table 5.2):

$$K_p = \mathbf{a}^T \begin{bmatrix} \frac{F_{p,1}L_{p,1}}{A_{p,1}} \\ \frac{F_{p,2}L_{p,2}}{A_{p,2}} \\ \frac{F_{p,3}L_{p,3}}{A_{p,3}} \\ \frac{F_{p,4}L_{p,4}}{A_{p,4}} \end{bmatrix} - \frac{5qL_{p,4}^4}{384I_p} \tag{5.20}$$

where the vector $\mathbf{a}$ is given by:

$$\mathbf{a}^T = [a_i] = \frac{1}{2}\begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix} T_p^{-1} \tag{5.21}$$

$$V_{p,1} = \sum_{i=1}^{4} a_i \left(\frac{F_{p,i}L_{p,i}}{A_{p,i}}\right)^2 + \frac{5L_{p,4}^4 q}{192I_p}\left(\frac{q}{4I_p} - a_4 \frac{F_{p,4}L_{p,4}}{A_{p,4}}\right) \tag{5.22}$$

where for j=2 and 3,

$$V_{p,j}(l_s)=\sum_{i=1}^{4}a_i(\frac{F_{p,i}L_{p,i}}{A_{p,i}})^2 b_{i,j}(l_s)+(\frac{q}{2I_p})^2 Q_{p,j}{}^{(1)}(l_s)-a_4\frac{qF_{p,4}}{I_p A_{p,4}}Q_{p,j}{}^{(2)}(l_s) \tag{5.23}$$

where
$$b_{i,2}(l_s)=2l_s(L_{p,i}-l_s+l_s\exp(-\frac{L_{p,i}}{l_s})) \tag{5.24}$$

$$b_{i,3}(l_s)=L_{p,i}l_s\sqrt{\pi}\,\mathrm{erf}(\frac{L_{p,i}}{l_s})-l_s{}^2(1-\exp(-(\frac{L_{p,i}}{l_s})^2)) \tag{5.25}$$

For j=4,

$$V_{p,4}(l_s,r)=\sum_{i=1}^{4}a_i(\frac{F_{p,i}L_{p,i}}{A_{p,i}})^2 b_{i,4}(l_s,r)+(\frac{q}{2I_p})^2 Q_{p,j}{}^{(1)}(l_s,r)-a_4\frac{qF_{p,4}}{I_p A_{p,4}}Q_{p,j}{}^{(2)}(l_s,r) \tag{5.26}$$

where

$$b_{i,4}(l_s,r)=2\int_0^{L_{p,i}}(L_{p,i}-x)\exp(-(\frac{x}{l_s})^r)dx \tag{5.27}$$

$$Q_{p,j}{}^{(1)}=\int_0^{L_{p,4}}\int_0^{L_{p,4}}(L_{p,4}-x_1)(L_{p,4}-x_2)x_1x_2\varphi_p(x_1)\varphi_p(x_2)\exp(-(\frac{|x_1-x_2|}{l_s})^r)dx_1dx_2 \tag{5.28}$$

$$Q_{p,j}{}^{(2)}=\int_0^{L_{p,4}}\int_0^{L_{p,4}}(L_{p,4}-x_2)x_2\varphi_p(x_2)\exp(-(\frac{|x_1-x_2|}{l_s})^r)dx_1dx_2 \tag{5.29}$$

where r=1 for j=2, r=2 for j=3. It should be stressed that w_p is not Gaussian (in this case, it follows a distribution which a weighted infinite sum of Gaussian PDFs) and it is Gaussian only when given $\boldsymbol{\theta}$.

The tenth row in Tables 5.10-5.12 gives the failure probability $P(F|\boldsymbol{\theta}_{\mathbf{MAP}}, \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$ of the target frame structure with $\boldsymbol{\theta}=\boldsymbol{\theta}_{\mathbf{MAP}}$ based on the calibration data $\mathcal{D}_1^{(l)}$ and each model class, which can be expressed in terms of the CDF of a standard Gaussian random variable $\Phi(\mathbf{z})$:

$$P(F \mid \boldsymbol{\theta}_{\mathbf{MAP}}, \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)}) = \Phi\left(\frac{-0.003-\mu_p(\boldsymbol{\theta}_{\mathbf{MAP}})}{\sigma_p(\boldsymbol{\theta}_{\mathbf{MAP}})}\right)+1-\Phi\left(\frac{0.003-\mu_p(\boldsymbol{\theta}_{\mathbf{MAP}})}{\sigma_p(\boldsymbol{\theta}_{\mathbf{MAP}})}\right) \tag{5.30}$$

The eleventh row gives the predicted robust failure probability $P(F|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$ (the number outside the parenthesis) of the target frame structure with the uncertainty in $\boldsymbol{\theta}$ taken into account for each model class, and it is calculated using:

$$\begin{aligned} P(F \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)}) &= \int P(F \mid \boldsymbol{\theta}, \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)}) p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)}) d\boldsymbol{\theta} \\ &\approx 1+\frac{1}{K}\sum_{k=1}^{K}\left[\Phi\left(\frac{-0.003-\mu_p(\boldsymbol{\theta}^{(k)})}{\sigma_p(\boldsymbol{\theta}^{(k)})}\right)-\Phi\left(\frac{0.003-\mu_p(\boldsymbol{\theta}^{(k)})}{\sigma_p(\boldsymbol{\theta}^{(k)})}\right)\right] \end{aligned} \tag{5.31}$$

where $\boldsymbol{\theta}^{(k)}$, k=1,2,...,K, are posterior samples from $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$. The number inside the parenthesis gives the estimate of the coefficient of variation (c.o.v.) of the above predicted robust failure probability estimate. It can be seen that $P(F|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$ is orders of magnitude different from $P(F|\boldsymbol{\theta}_{\mathbf{MAP}}, \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$ showing that the effects of the uncertainties in the parameters on the failure probabilities is substantial. In fact, ignoring the uncertainty in $\boldsymbol{\theta}$ would be disastrous since $P(F|\boldsymbol{\theta}_{\mathbf{MAP}}, \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$ greatly underestimates the failure probability for all model classes and it varies greatly from one model class to another, in contrast with the robust case $P(F|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$. Figure 5.15 shows $P(F|\boldsymbol{\theta}^{(k)}, \mathcal{D}_1^{(3)}, \mathcal{M}_j^{(1)})$ corresponding to each posterior sample model $\boldsymbol{\theta}^{(k)}$, sorted in increasing order. Figure 5.16 shows the CDF of $P(F|\boldsymbol{\theta}, \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$

estimated using posterior samples from $p(\boldsymbol{\theta}|\mathcal{D}_1^{(3)}, \mathcal{M}_j^{(1)})$. Figures 5.15 and 5.16 confirm that there is a large variability in $P(F|\boldsymbol{\theta}, \mathcal{D}_1^{(3)}, \mathcal{M}_j^{(1)})$ due to the uncertainties in $\boldsymbol{\theta}$.

Posterior model averaging can be carried out to obtain the predicted hyper-robust failure probability $P(F|\mathcal{D}_1^{(l)}, M_1)$ given the set of candidate model classes M_1 as shown in the last row of Tables 5.10-5.12:

$$P(F \mid \mathcal{D}_1^{(l)}, M_1) = \sum_{j=1}^{4} P(F \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)}) P(\mathcal{M}_j^{(1)} \mid \mathcal{D}_1^{(l)}, M_1) \tag{5.32}$$

$P(F|\mathcal{D}_1^{(1)},M_1)$, $P(F|\mathcal{D}_1^{(2)},M_1)$ and $P(F|\mathcal{D}_1^{(3)}, M_1)$ are estimated to be 1.13×10^{-2}, 6.68×10^{-4}, 2.48×10^{-4}, respectively, showing that the predicted failure probability of the target system depends on the uncertainties in the model parameters which in turn depends on the amount of data and the model classes under consideration.

Figures 5.17-5.19 show the CDFs of the predicted vertical displacement w_p at point P in the target frame structure corresponding to each sample $\boldsymbol{\theta}^{(k)}$, k=1,2,…,4000, from $p(\boldsymbol{\theta}|\mathcal{D}_1^{(3)}, \mathcal{M}_j^{(1)})$. The robust posterior CDF of the response w_p of interest for the target frame structure can be obtained using the Theorem of Total Probability, as the previous section. Samples of w_p can be obtained as follows: For each $\boldsymbol{\theta}^{(k)}$, k=1,2,...,K, which are the posterior samples from $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$, generate a sample $w_p^{(k)}$ for w_p from a Gaussian distribution with mean $\mu_p(\boldsymbol{\theta}^{(k)})$ and variance $\sigma_p^2(\boldsymbol{\theta}^{(k)})$. These samples can also be used to find the probability in (5.31) by approximating the robust failure probability $P(F|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$ as the proportion $\tilde{p}_j$ of failure samples out of the K samples:

$$\tilde{p}_j = \int I(w_p)\, p(w_p \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})\, dw_p\, d\boldsymbol{\theta} \approx \frac{1}{K}\sum_{k=1}^{K} I(w_p^{(k)}) \tag{5.33}$$

where $w_p^{(k)} \sim N(\mu_p(\boldsymbol{\theta}^{(k)}), \sigma_p^2(\boldsymbol{\theta}^{(k)}))$ and $\boldsymbol{\theta}^{(k)}$, k=1,2,...,K, are posterior samples from $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})$. It can be shown that the estimator in (5.31), which implicitly integrates over w_p, is always of a smaller c.o.v. and thus is more accurate than $\tilde{p}_j$. A very efficient stochastic simulation method called Subset Simulation (Au and Beck 2001) can also be used for the robust posterior CDF. Here, the CDF for w_p is obtained based on the samples for w_p generated as described above for each model class. The results for $\mathcal{M}_2^{(1)}$, $\mathcal{M}_3^{(1)}$ and $\mathcal{M}_4^{(1)}$ are shown in Figures 5.20-5.25 for the three data cases. Figures 20-22 show that the CDFs for the three model classes are very close to each other in the high probability region but differ somewhat in the tails so the predicted failure probability is quite different (though still within the same order of magnitude), as shown in Tables 5.10-5.12. Figures 5.23-5.25 show that the CDFs for data cases 2 and 3 for each of these three model classes are very close to each other in the high probability region.

From the results in Tables 5.10-5.12, it can be seen that $P(F|\mathcal{D}_1^{(l)}, \mathcal{M}_1^{(1)})P(\mathcal{M}_1^{(1)}|\mathcal{D}_1^{(l)}, M_1)$ is negligible compared to $P(F|\mathcal{D}_1^{(l)}, M_1)$ for data case 3 and so the contribution of $\mathcal{M}_1^{(1)}$ is negligible to the prediction of interest, the failure probability of the target frame structure. Also, having a posterior model class probability $P(\mathcal{M}_1^{(1)}|\mathcal{D}_1^{(l)},M_1)$ that is several orders of magnitude smaller than those for the other model classes implies $\mathcal{M}_1^{(1)}$ is relatively improbable conditioned on the data $\mathcal{D}_1^{(l)}$. Thus, $\mathcal{M}_1^{(1)}$ is dropped in the subsequent analyses.

Note that the posterior probability $P(\mathcal{M}_j^{(1)}|\mathcal{D}_1^{(l)},M_1)$ for each model class conditioned on the data $\mathcal{D}_1^{(l)}$ gives the plausibility of each $\mathcal{M}_j^{(1)}$ given the set of candidate model classes $M_1=\{\mathcal{M}_j^{(1)}$,

j=1,2,3,4} and $P(F|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(1)})P(\mathcal{M}_j^{(1)}|\mathcal{D}_1^{(l)}, M_1)$ gives the contribution of each model class to the desired response prediction. These probabilities do not give information regarding the predictability of each model class for the response of other systems, including the target system. It is shown in the following sections how the data from the validation and accreditation experiments are used to evaluate the prediction consistency and accuracy of the calibrated model classes.

Table 5.10 Statistical results using data $\mathcal{D}_1^{(1)}$ from the calibration experiment

		$\mathcal{M}_1^{(1)}$	$\mathcal{M}_2^{(1)}$	$\mathcal{M}_3^{(1)}$	$\mathcal{M}_4^{(1)}$
MAP	μ_s (Pa^{-1})	8.34×10^{-11}	8.80×10^{-11}	8.80×10^{-11}	8.80×10^{-11}
	σ_s^2(Pa^{-2})	7.69×10^{-23}	8.24×10^{-23}	8.24×10^{-23}	8.26×10^{-23}
	σ_ε^2*(m^2);l_s(m)	1.76×10^{-11}*	0.00652	0.00727	0.0035
	r				0.5
Parameter Statistics (SS)	μ_s (Pa^{-1})	8.36×10^{-11};5.9%	8.81×10^{-11};2.2%	8.80×10^{-11};2.2%	8.80×10^{-11};2.4%
	σ_s^2(Pa^{-2})	1.24×10^{-22};43.0%	1.06×10^{-22};48.8%	1.13×10^{-22};50.6%	1.11×10^{-22};40.8%
	σ_ε^2*(m^2);l_s(m)	2.07×10^{-11}*;34.1%	0.0217;97.8%	0.0175;66.4%	0.0203;144.2%
	r				1.40;52.4%
	R	$\begin{bmatrix} 1 & -0.05 & 0.06 \\ & 1 & -0.15 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.01 & 0.17 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0.03 & 0.12 \\ & 1 & -0.10 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.15 & 0.07 & 0.22 \\ & 1 & 0.34 & -0.13 \\ & & 1 & -0.03 \\ & & & 1 \end{bmatrix}$
Log evidence (AA)		166.45	170.92	171.02	171.83
$P(\mathcal{M}_j^{(1)}\|\mathcal{D}_1^{(1)},M_1)$ (AA)		2.49×10^{-3}	0.217	0.240	0.540
Log evidence (SS)		167.13	172.06	172.01	172.34
$E[\ln p(\mathcal{D}_1^{(1)}\|\boldsymbol{\theta},\mathcal{M}_j^{(1)})]$		170.70	177.66	177.83	177.76
Expected information gain		3.57	5.60	5.82	5.41
$P(\mathcal{M}_j^{(1)}\|\mathcal{D}_1^{(1)},M_1)$ (SS)		2.20×10^{-3}	0.305	0.290	0.403
$P(F\|\boldsymbol{\theta}_{\mathbf{MAP}},\mathcal{D}_1^{(1)},\mathcal{M}_j^{(1)})$		5.32×10^{-2}	0	0	0
$P(F\|\mathcal{D}_1^{(1)},\mathcal{M}_j^{(1)})$		0.253(2.1%)	1.48×10^{-2}(12.6%)	6.58×10^{-3}(16.5%)	1.07×10^{-2}(14.1%)
$P(F\|\mathcal{D}_1^{(1)},M_1)$		1.13×10^{-2}			

Table 5.11 Statistical results using data $\mathcal{D}_1^{(2)}$ from the calibration experiment

		$\mathcal{M}_1^{(1)}$	$\mathcal{M}_2^{(1)}$	$\mathcal{M}_3^{(1)}$	$\mathcal{M}_4^{(1)}$
MAP	μ_s (Pa^{-1})	8.68×10^{-11}	8.86×10^{-11}	8.86×10^{-11}	8.86×10^{-11}
	σ_s^2(Pa^{-2})	3.33×10^{-23}	5.11×10^{-23}	4.98×10^{-23}	4.94×10^{-23}
	σ_ε^2*(m^2);l_s(m)	1.17×10^{-11}*	0.0263	0.0287	0.0286
	r				3
Parameter Statistics (SS)	μ_s (Pa^{-1})	8.69×10^{-11};1.6%	8.87×10^{-11};0.90%	8.87×10^{-11};0.97%	8.86×10^{-11};1.0%
	σ_s^2(Pa^{-2})	3.88×10^{-23};26.9%	5.61×10^{-23};24.7%	5.63×10^{-23};23.5%	5.78×10^{-23};23.8%
	σ_ε^2*(m^2);l_s(m)	1.32×10^{-11}*;25.4%	0.0336;34.6%	0.0318;27.3%	0.032;38.6%
	r				1.47;47.5%
	R	$\begin{bmatrix} 1 & 0.12 & 0.003 \\ & 1 & -0.03 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.001 & 0.06 \\ & 1 & -0.09 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.08 & 0.03 \\ & 1 & -0.12 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0.14 & 0.07 & 0.05 \\ & 1 & -0.13 & -0.01 \\ & & 1 & 0.10 \\ & & & 1 \end{bmatrix}$
Log evidence (AA)		702.43	710.04	710.18	712.03
$P(\mathcal{M}_j^{(1)}\|\mathcal{D}_1^{(2)},M_1)$ (AA)		5.23×10^{-5}	0.106	0.122	0.773
Log evidence (SS)		702.95	709.20	710.03	710.40
$E[\ln p(\mathcal{D}_1^{(2)}\|\boldsymbol{\theta},\mathcal{M}_j^{(1)})]$		707.77	717.85	718.18	718.02
Expected information gain		4.82	8.65	8.15	7.62
$P(\mathcal{M}_j^{(1)}\|\mathcal{D}_1^{(2)},M_1)$ (SS)		2.92×10^{-4}	0.151	0.347	0.502
$P(F\|\boldsymbol{\theta}_{\mathbf{MAP}},\mathcal{D}_1^{(2)},\mathcal{M}_j^{(1)})$		4.52×10^{-2}	1.77×10^{-7}	6.91×10^{-10}	3.87×10^{-13}
$P(F\|\mathcal{D}_1^{(2)},\mathcal{M}_j^{(1)})$		0.137(1.9%)	5.84×10^{-4}(13.5%)	2.58×10^{-4}(26.7%)	8.98×10^{-4}(12.6%)
$P(F\|\mathcal{D}_1^{(2)},M_1)$		6.68×10^{-4}			

Table 5.12 Statistical results using data $\mathcal{D}_1^{(3)}$ from the calibration experiment

		$\mathcal{M}_1^{(1)}$	$\mathcal{M}_2^{(1)}$	$\mathcal{M}_3^{(1)}$	$\mathcal{M}_4^{(1)}$
MAP	μ_s (Pa^{-1})	8.64×10^{-11}	8.87×10^{-11}	8.87×10^{-11}	8.87×10^{-11}
	σ_s^2(Pa^{-2})	3.24×10^{-23}	4.87×10^{-23}	4.76×10^{-23}	4.72×10^{-23}
	σ_ε^2*(m^2);l_s(m)	1.11×10^{-11}*	0.0284	0.0307	0.0305
	r				3
Parameter Statistics (SS)	μ_s (Pa^{-1})	8.64×10^{-11};1.2%	8.88×10^{-11};0.83%	8.87×10^{-11};0.69%	8.88×10^{-11};0.8%
	σ_s^2(Pa^{-2})	3.69×10^{-23};26.0%	5.19×10^{-23};19.5%	5.37×10^{-23};20.4%	5.20×10^{-23};19.9%
	σ_ε^2*(m^2);l_s(m)	1.24×10^{-11}*;23.7%	0.0319;27.5%	0.0327;23.6%	0.0328;27.8%
	r				1.79;40.5%
	R	$\begin{bmatrix} 1 & 0.09 & -0.11 \\ & 1 & -0.09 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.05 & 0.20 \\ & 1 & -0.10 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.10 & 0.04 \\ & 1 & -0.21 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0.01 & 0.15 & 0.07 \\ & 1 & -0.05 & -0.14 \\ & & 1 & -0.01 \\ & & & 1 \end{bmatrix}$
Log evidence (AA)		1059.27	1071.33	1071.47	1073.22
$P(\mathcal{M}_j^{(1)}\|\mathcal{D}_1^{(3)},M_1)$ (AA)		6.60×10^{-7}	0.114	0.131	0.755
Log evidence (SS)		1059.63	1071.34	1071.66	1071.87
$E[\ln p(\mathcal{D}_1^{(3)}\|\boldsymbol{\theta},\mathcal{M}_j^{(1)})]$		1064.89	1079.75	1080.15	1079.82
Expected information gain		5.27	8.41	8.49	7.95
$P(\mathcal{M}_j^{(1)}\|\mathcal{D}_1^{(3)},M_1)$ (SS)		2.01×10^{-6}	0.245	0.338	0.416
$P(F\|\boldsymbol{\theta}_{\mathbf{MAP}},\mathcal{D}_1^{(3)},\mathcal{M}_j^{(1)})$		3.61×10^{-2}	3.56×10^{-7}	3.19×10^{-9}	6.70×10^{-12}
$P(F\|\mathcal{D}_1^{(3)},\mathcal{M}_j^{(1)})$		9.81×10^{-2}(1.9%)	3.58×10^{-4}(16.1%)	1.30×10^{-4}(26.1%)	2.79×10^{-4}(16.5%)
$P(F\|\mathcal{D}_1^{(3)},M_1)$		2.48×10^{-4}			

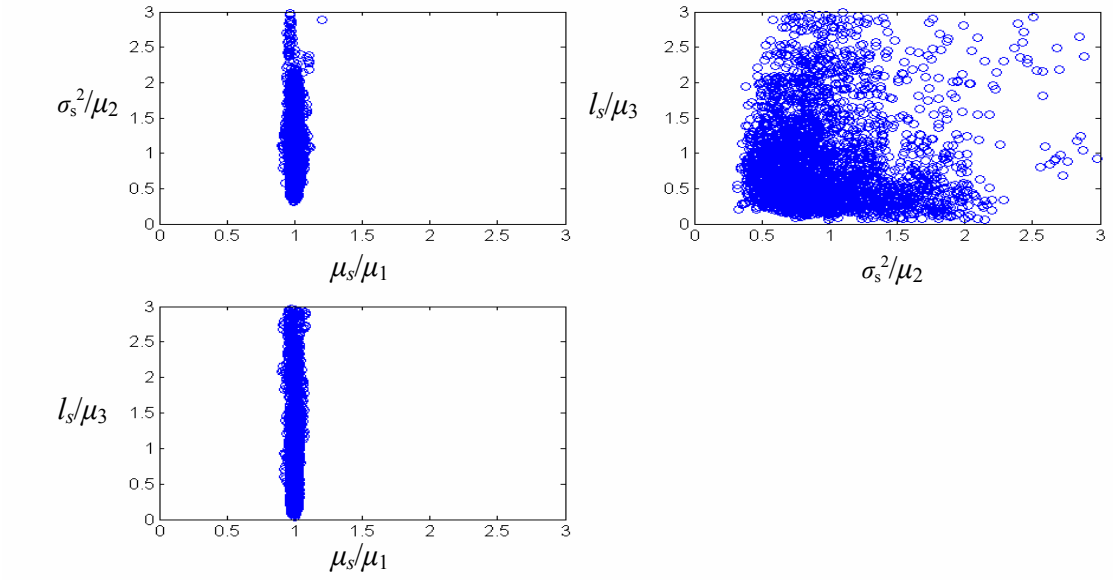

Figure 5.3: Pairwise sample plots of posterior samples for $p(\theta|\mathcal{D}_1^{(1)}, \mathcal{M}_2^{(1)})$ normalized by posterior mean

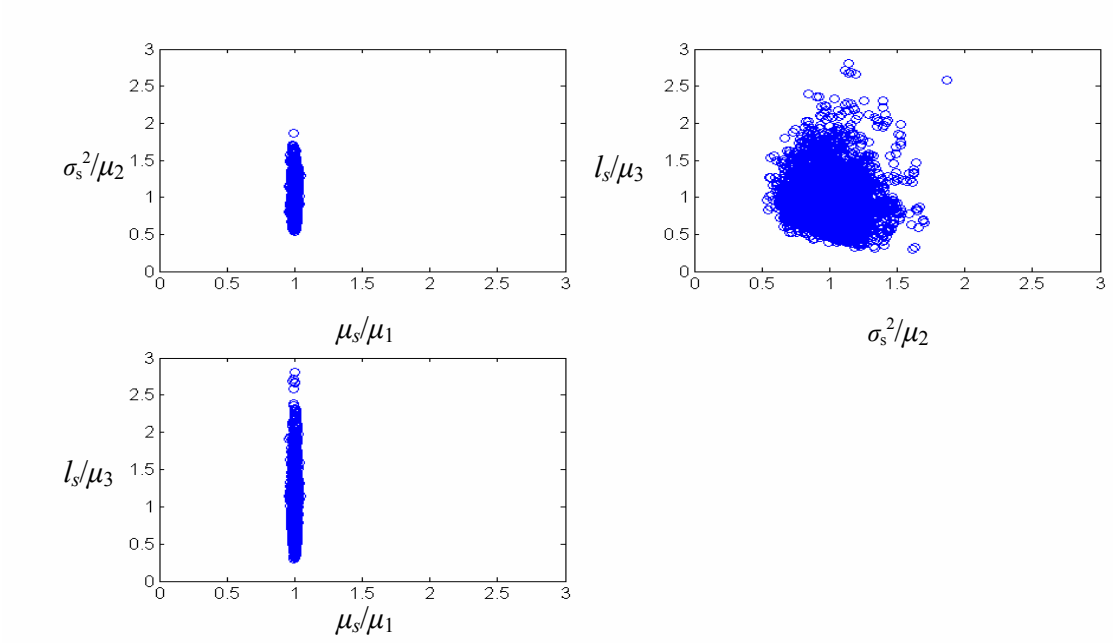

Figure 5.4: Pairwise sample plots of posterior samples for $p(\theta|\mathcal{D}_1^{(2)}, \mathcal{M}_2^{(1)})$ normalized by posterior mean

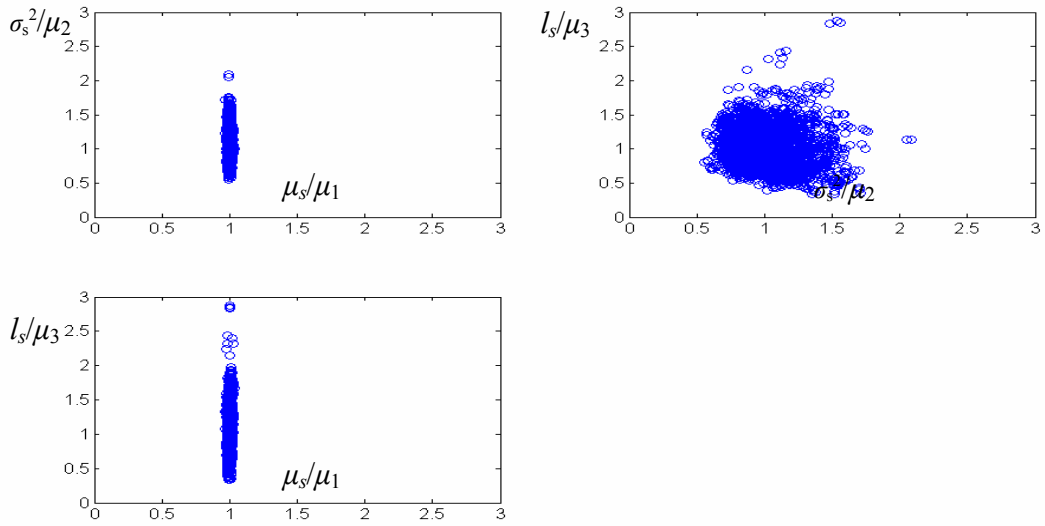

Figure 5.5: Pairwise sample plots of posterior samples for $p(\theta|\mathcal{D}_1^{(3)}, \mathcal{M}_2^{(1)})$ normalized by posterior mean

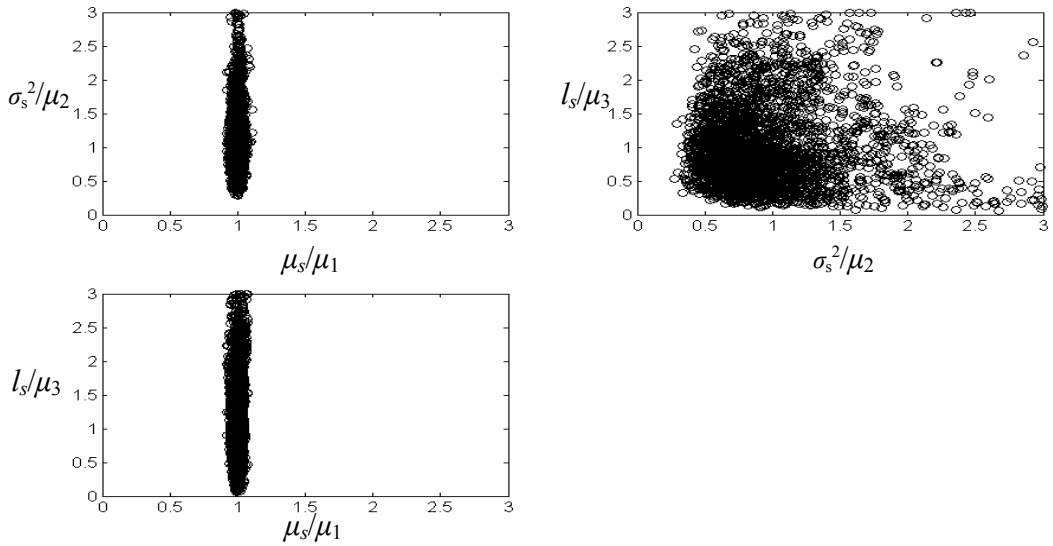

Figure 5.6: Pairwise sample plots of posterior samples for $p(\theta|\mathcal{D}_1^{(1)}, \mathcal{M}_3^{(1)})$ normalized by posterior mean

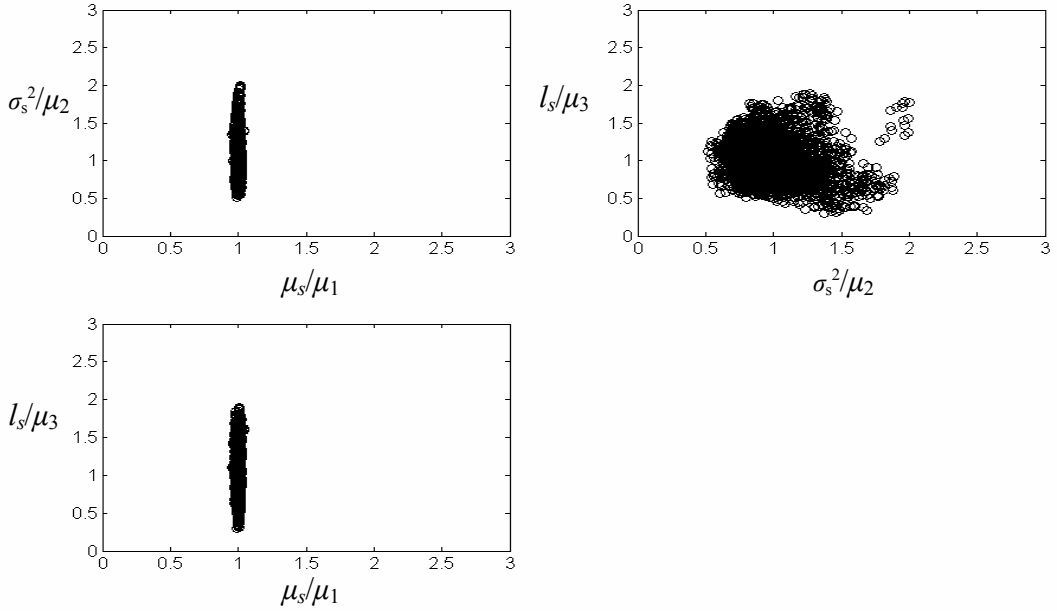

Figure 5.7: Pairwise sample plots of posterior samples for $p(\theta|\mathcal{D}_1^{(2)}, \mathcal{M}_3^{(1)})$ normalized by posterior mean

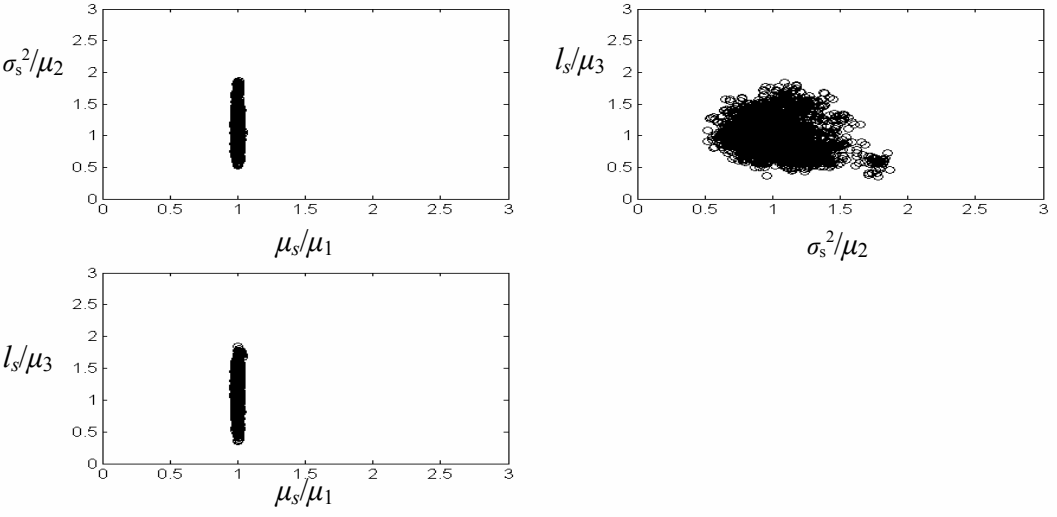

Figure 5.8: Pairwise sample plots of posterior samples for $p(\theta|\mathcal{D}_1^{(3)}, \mathcal{M}_3^{(1)})$ normalized by posterior mean

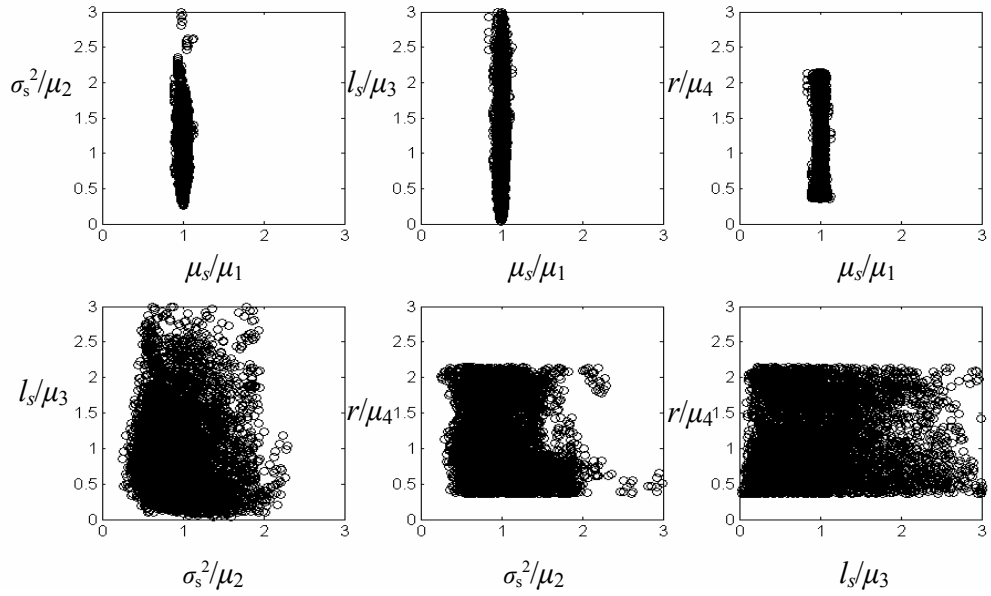

Figure 5.9: Pairwise sample plots of posterior samples for $p(\theta|\mathcal{D}_1^{(1)}, \mathcal{M}_4^{(1)})$ normalized by posterior mean

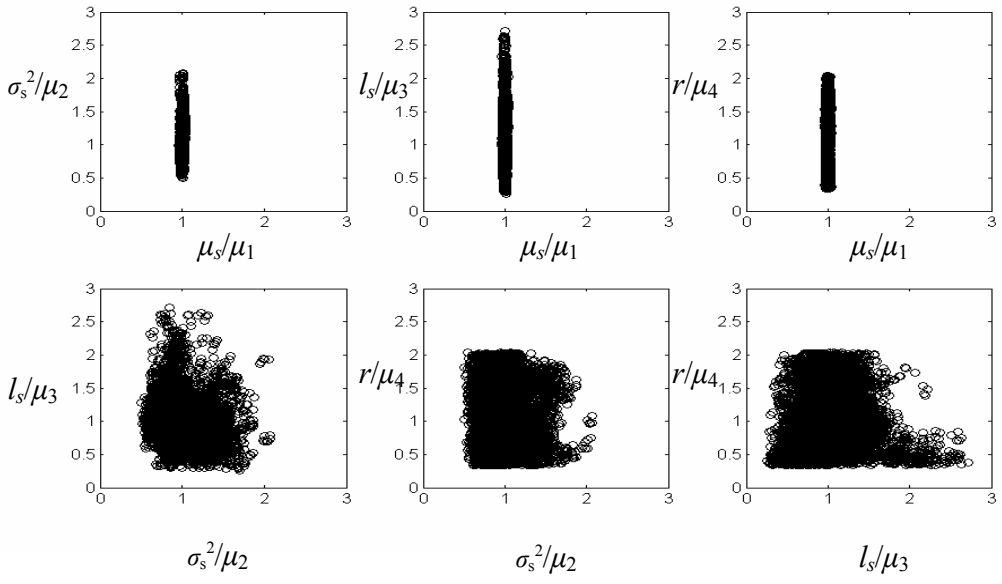

Figure 5.10: Pairwise sample plots of posterior samples for $p(\theta|\mathcal{D}_1^{(2)}, \mathcal{M}_4^{(1)})$ normalized by posterior mean

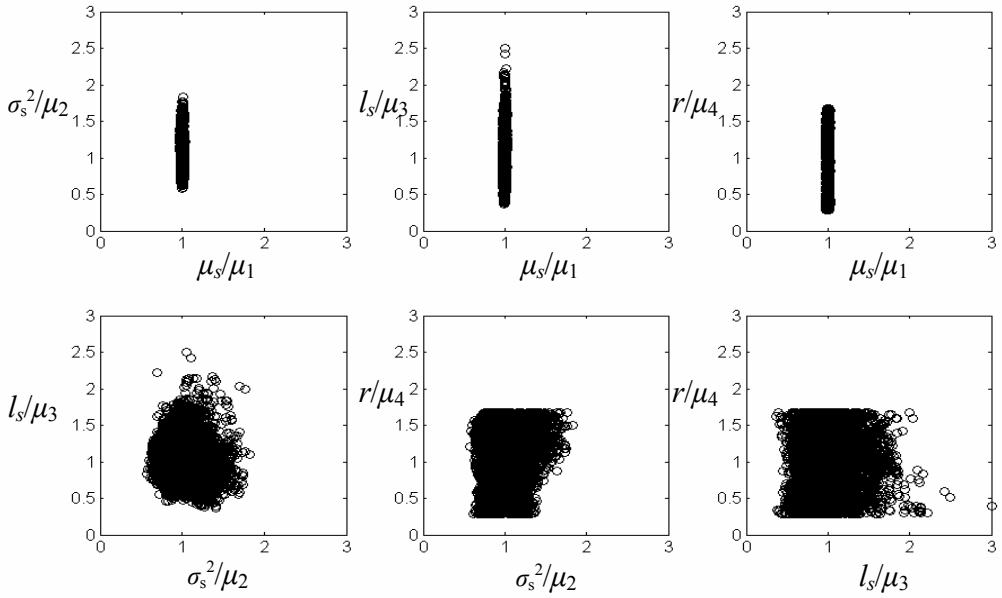

Figure 5.11: Pairwise sample plots of posterior samples for $p(\theta|\mathcal{D}_1^{(3)}, \mathcal{M}_4^{(1)})$ normalized by posterior mean

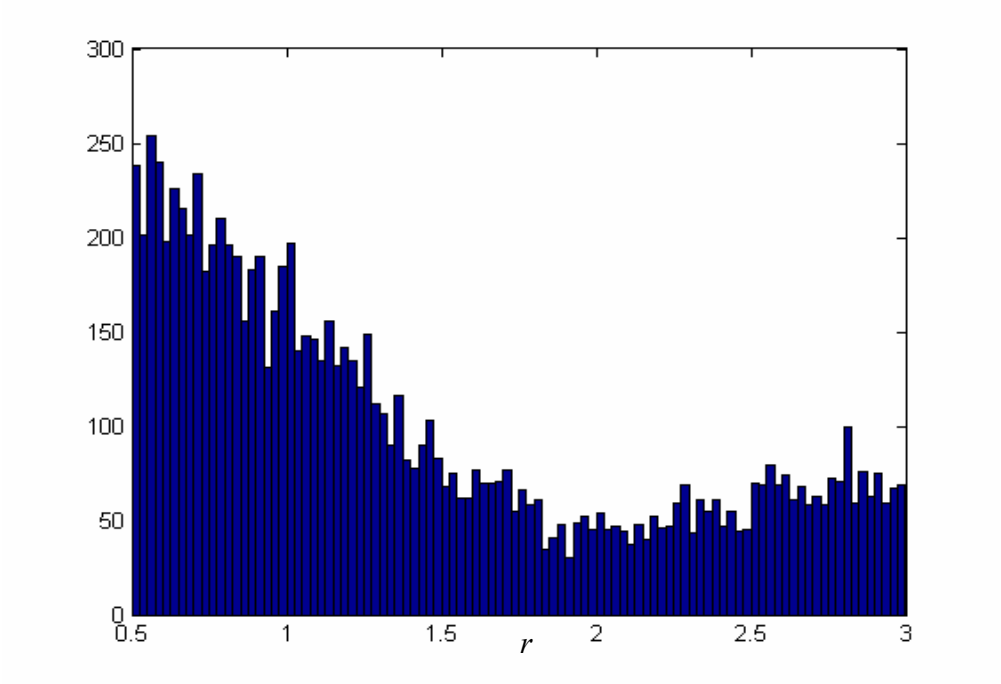

Figure 5.12: Histogram for posterior samples for $p(r|\mathcal{D}_1^{(1)}, \mathcal{M}_4^{(1)})$

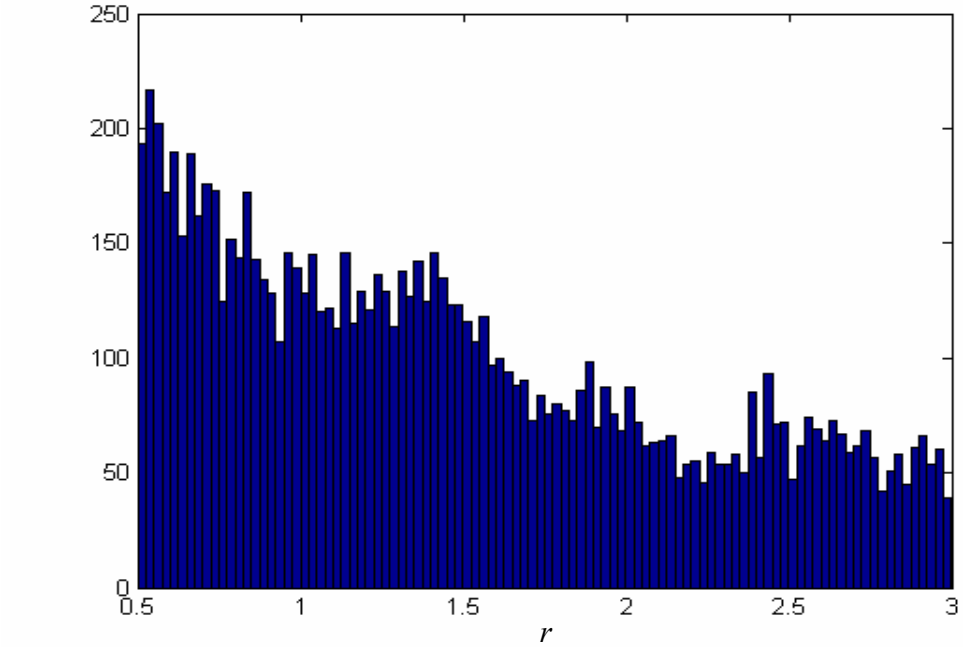

Figure 5.13: Histogram for posterior samples for $p(r|\mathcal{D}_1^{(2)}, \mathcal{M}_4^{(1)})$

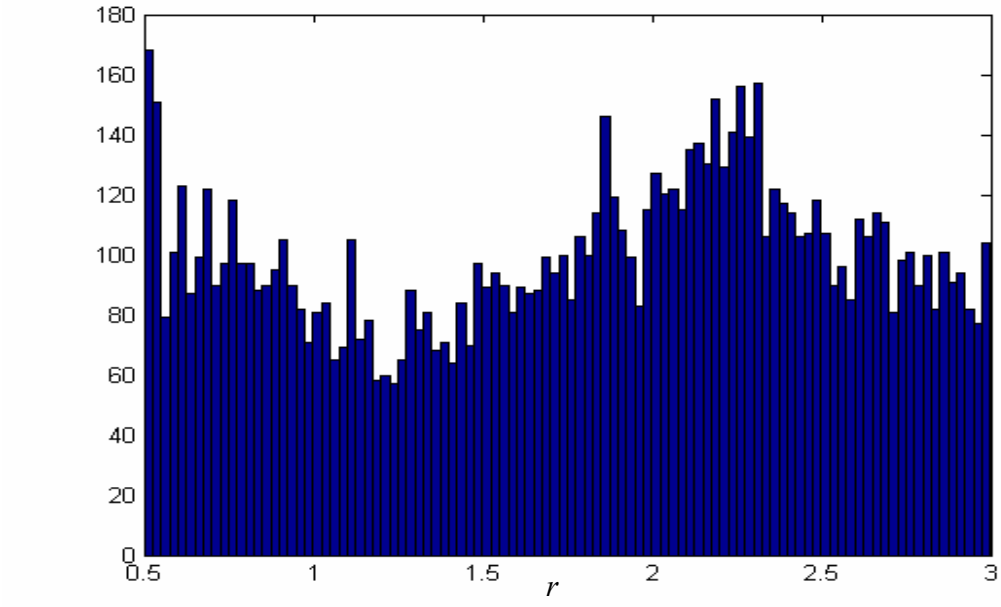

Figure 5.14: Histogram for posterior samples for $p(r|\mathcal{D}_1^{(3)}, \mathcal{M}_4^{(1)})$

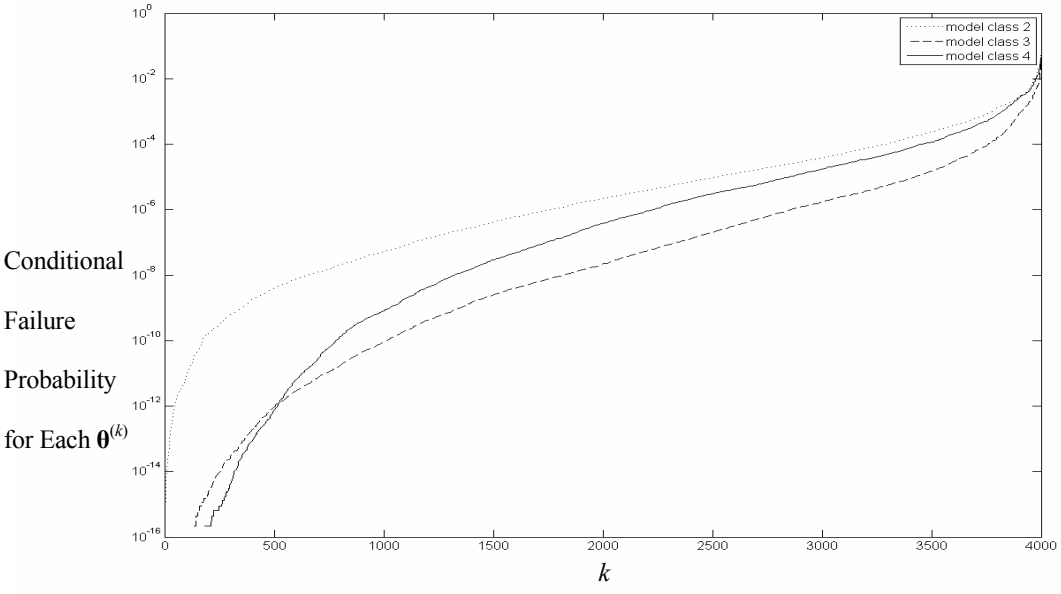

Figure 5.15: The failure probability (sorted in increasing order) conditioned on each posterior sample $\boldsymbol{\theta}^{(k)}$ for model class $\mathcal{M}_j^{(1)}$, i.e. $P(F|\boldsymbol{\theta}^{(k)},\mathcal{D}_1^{(3)}, \mathcal{M}_j^{(1)})$, for j=2,3,4

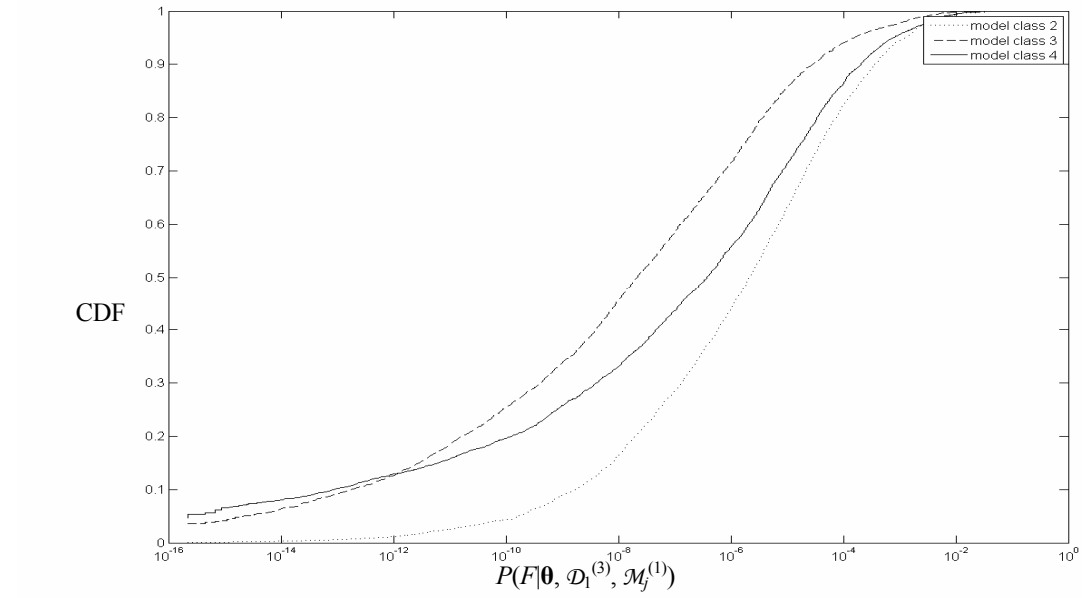

Figure 5.16: CDF of failure probability $P(F|\boldsymbol{\theta},\mathcal{D}_1^{(3)},\mathcal{M}_j^{(1)})$, j=2,3,4, estimated using posterior samples for model class $\mathcal{M}_j^{(1)}$

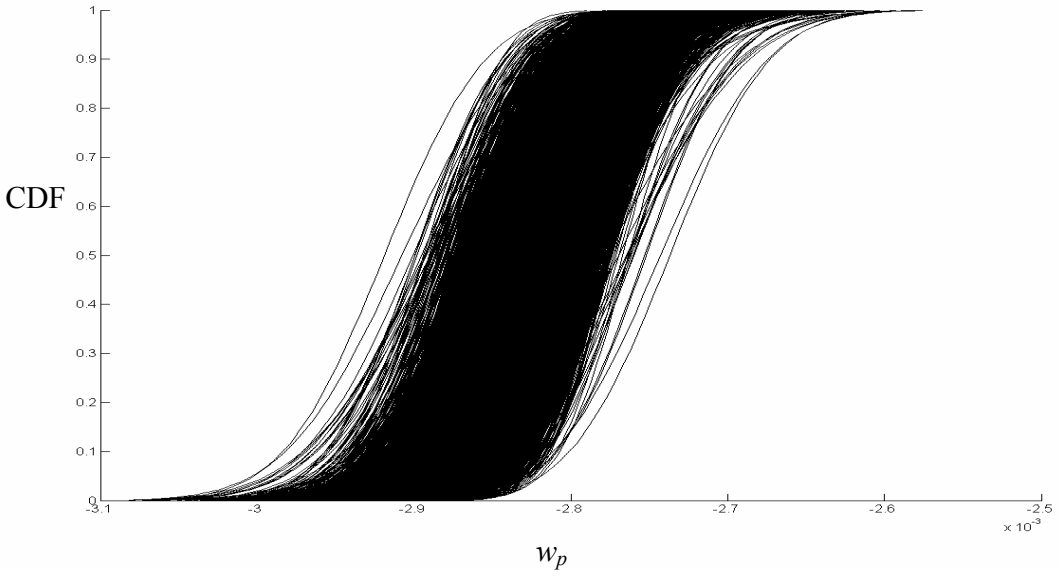

Figure 5.17: CDF of predicted vertical displacement w_p at point P in the target frame structure conditioned on each sample from $p(\theta|\mathcal{D}_1^{(3)}, \mathcal{M}_2^{(1)})$

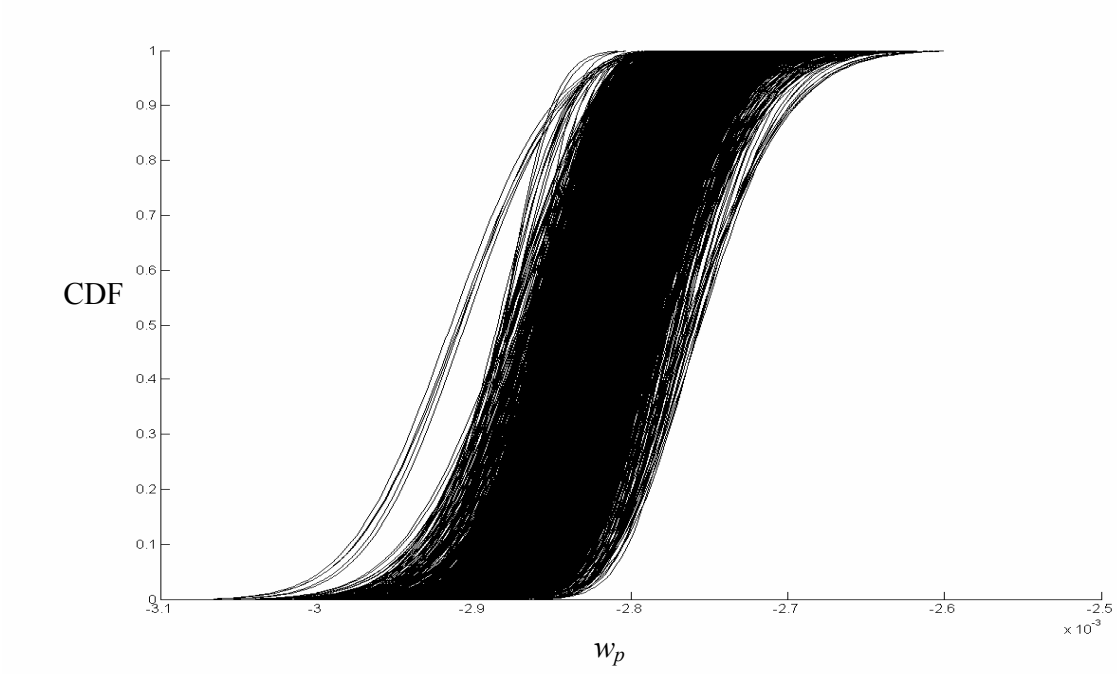

Figure 5.18: CDF of predicted vertical displacement w_p at point P in the target frame structure conditioned on each sample from $p(\theta|\mathcal{D}_1^{(3)}, \mathcal{M}_3^{(1)})$

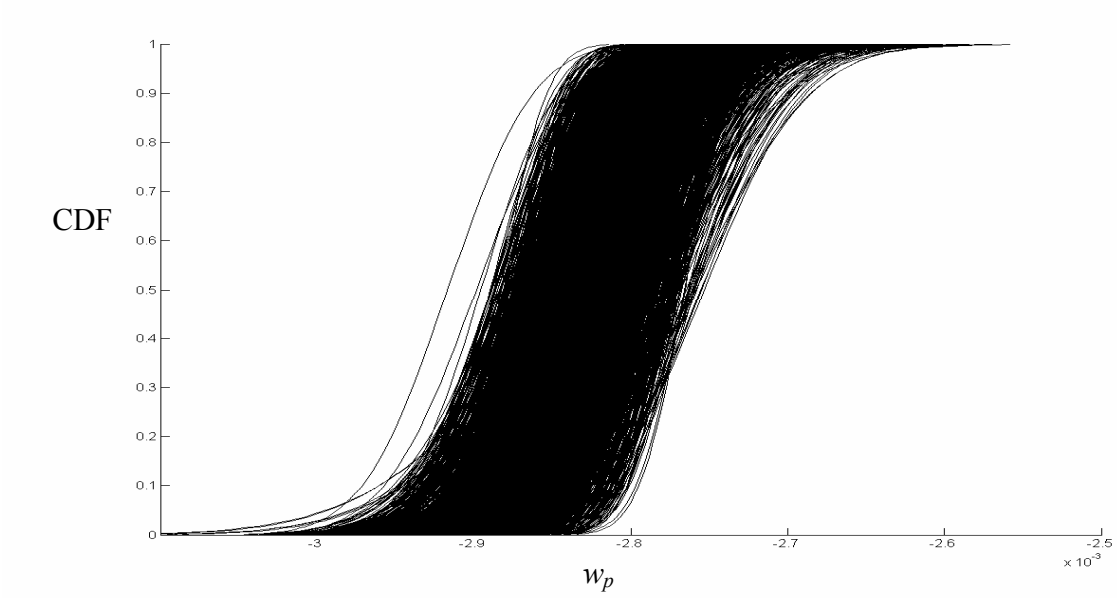

Figure 5.19: CDF of predicted vertical displacement w_p at point P in the target frame structure conditioned on each sample from $p(\theta|\mathcal{D}_1^{(3)}, \mathcal{M}_4^{(1)})$

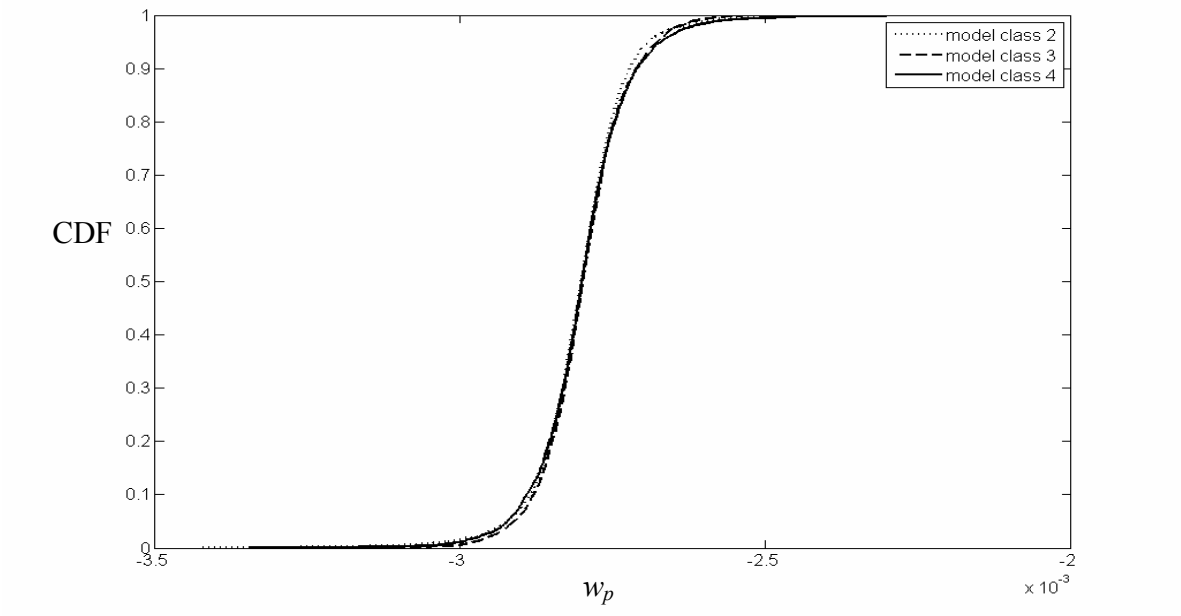

Figure 5.20: Robust posterior CDF of predicted vertical displacement w_p at point P in the target frame structure using posterior samples for $p(\theta|\mathcal{D}_1^{(1)}, \mathcal{M}_j^{(1)})$, j=2,3,4

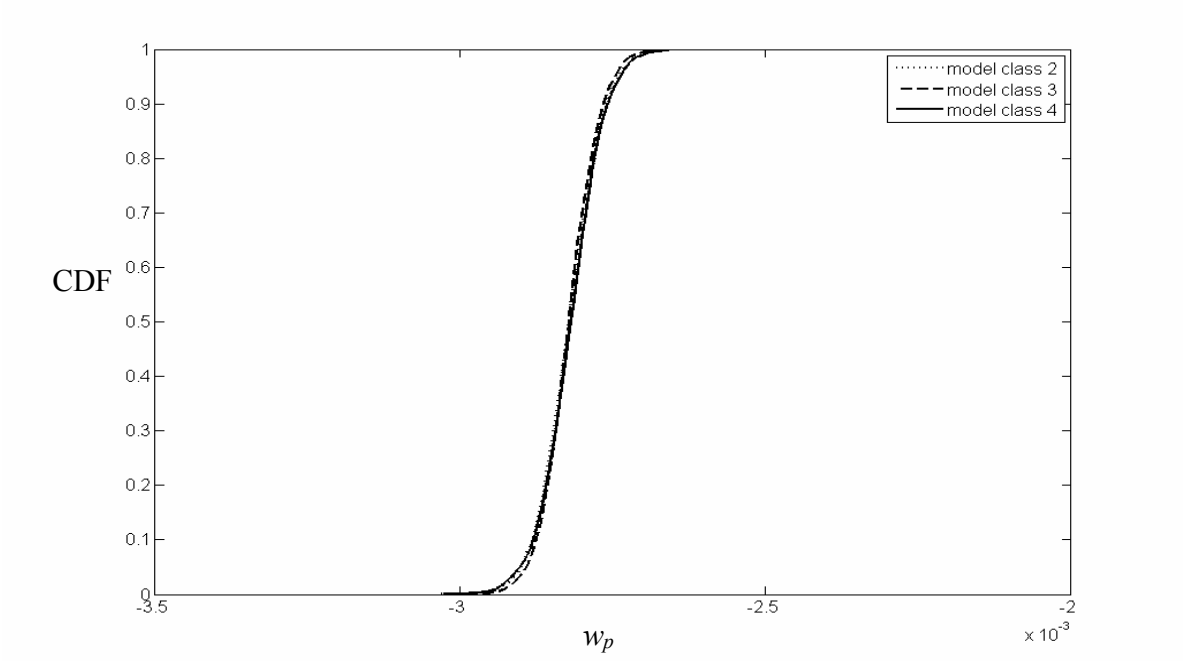

Figure 5.21: Robust posterior CDF of predicted vertical displacement w_p at point P in the target frame structure using posterior samples for $p(\theta|\mathcal{D}_1^{(2)}, \mathcal{M}_j^{(1)})$, j=2,3,4

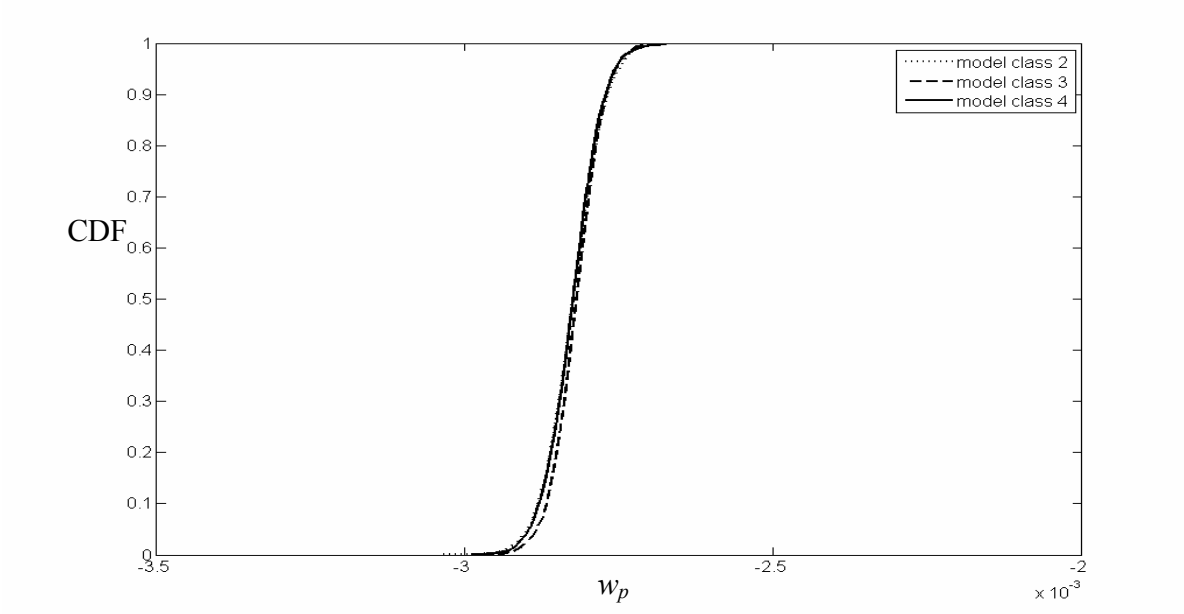

Figure 5.22: Robust posterior CDF of predicted vertical displacement w_p at point P in the target frame structure using posterior samples for $p(\theta|\mathcal{D}_1^{(3)}, \mathcal{M}_j^{(1)})$, j=2,3,4

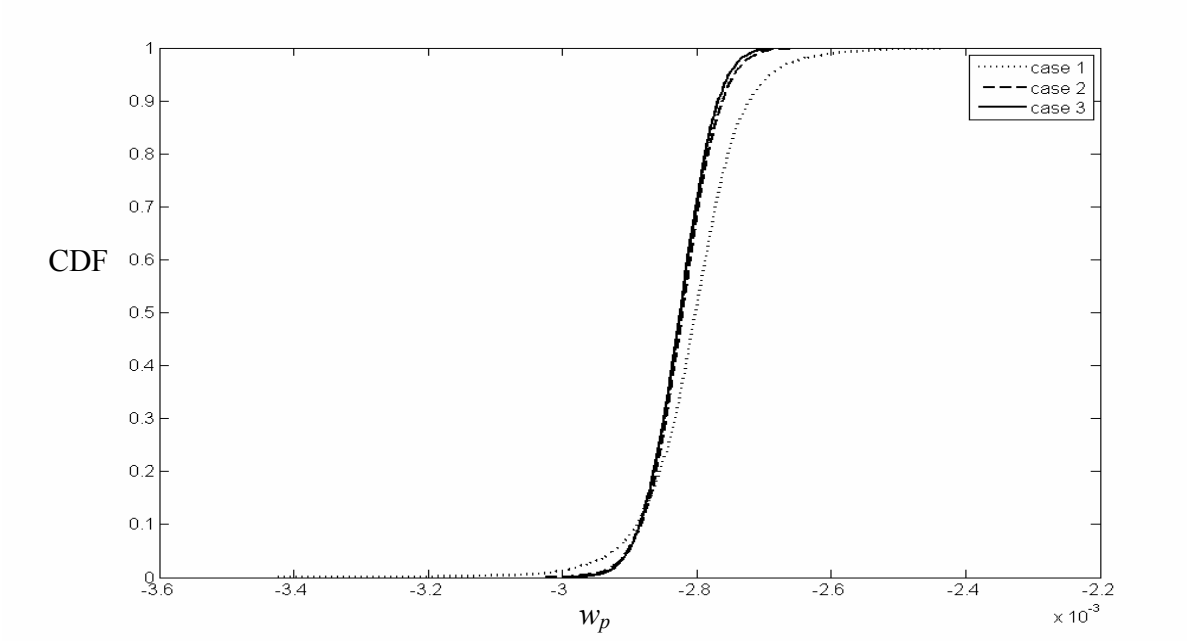

Figure 5.23: Robust posterior CDF of predicted vertical displacement w_p at point P in the target frame structure using posterior samples for $p(\theta|\mathcal{D}_1^{(J)}, \mathcal{M}_2^{(1)})$ for 3 different data cases

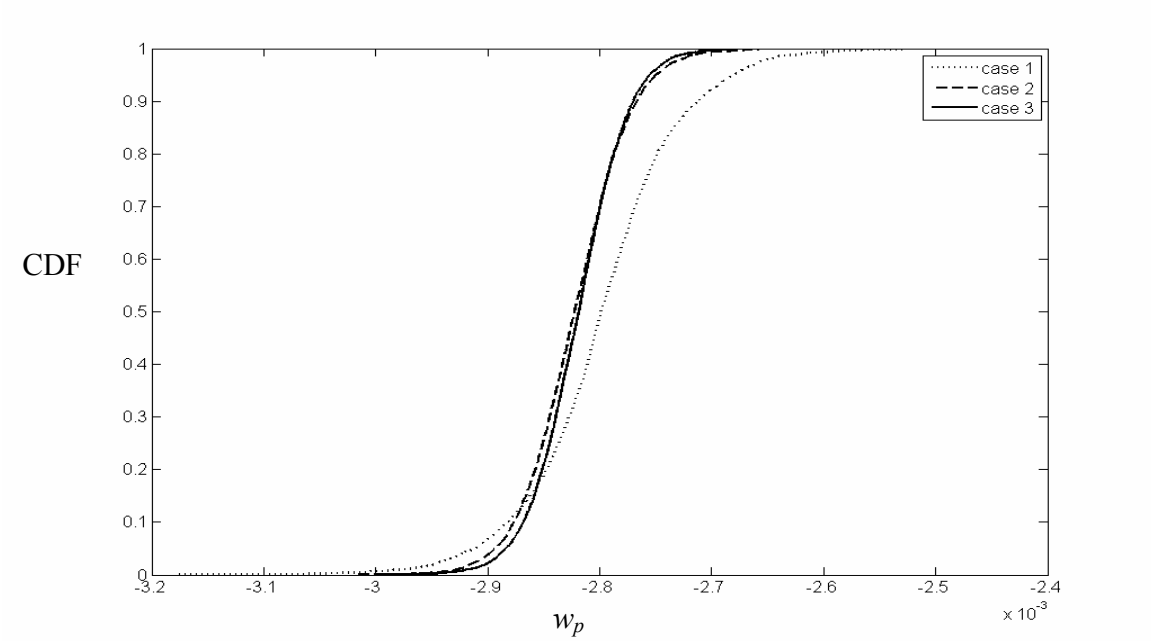

Figure 5.24: Robust posterior CDF of predicted vertical displacement w_p at point P in the target frame structure using posterior samples for $p(\theta|\mathcal{D}_1^{(J)}, \mathcal{M}_3^{(1)})$ for 3 different data cases

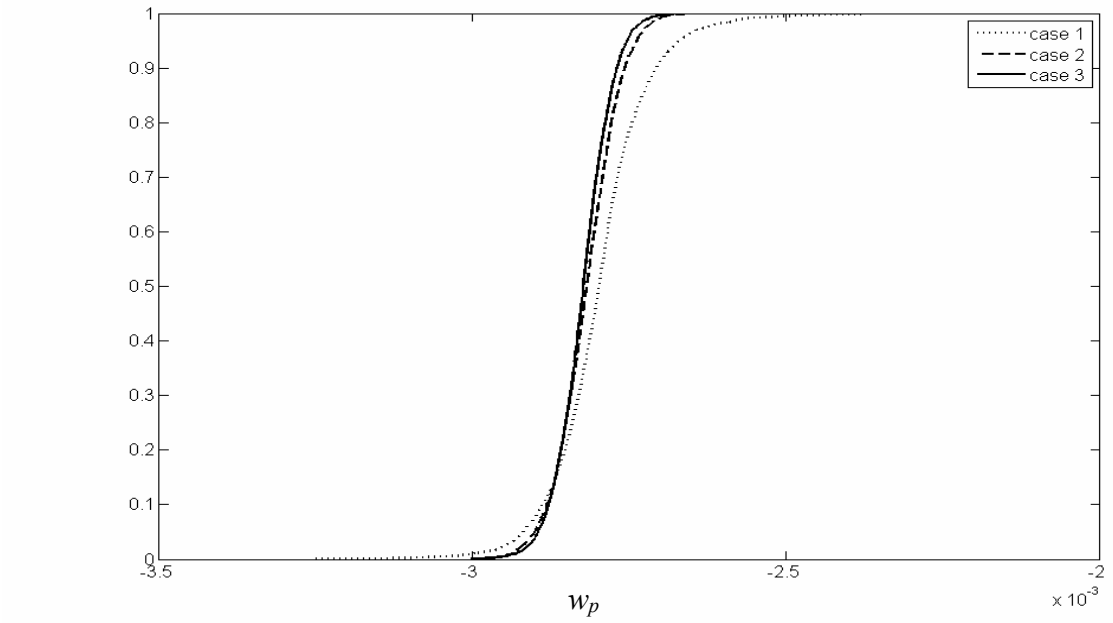

Figure 5.25: Robust posterior CDF of predicted vertical displacement w_p at point P in the target frame structure using posterior samples for $p(\theta|\mathcal{D}_1^{(J)}, \mathcal{M}_4^{(1)})$ for 3 different data cases

5.2. Using data $\mathcal{D}_2$ from the validation experiment

Candidate model classes for the subsystem in the validation experiment are $\mathcal{M}_j^{(2)}$, j=1,2,3. The only difference between the subsystem here and that in the previous experiment is the longer beam length. The uncertain parameters $\boldsymbol{\theta}^{(2,j)}$ for $\mathcal{M}_j^{(2)}$ are the same as $\boldsymbol{\theta}^{(1,j+1)}$ for $\mathcal{M}_{j+1}^{(1)}$. The "prior" PDF $p(\boldsymbol{\theta}^{(2,j)}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})$ for $\mathcal{M}_j^{(2)}$ is given by the "posterior" PDF $p(\boldsymbol{\theta}^{(1,j+1)}|\mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)})$ for $\mathcal{M}_{j+1}^{(1)}$. Data $\mathcal{D}_2^{(l)} = \{\delta L_v^{(i)}, i=1,2\ldots, N_v\}$ from the validation experiment are used to investigate the predictive performance, including the prediction consistency and accuracy of the model classes.

To evaluate *prediction accuracy*, we compute the probability that the response $\delta L_{v,p}$, which is the elongation of the bar in the validation experiment, predicted using the model classes updated by data from the previous experiment (i.e. data $\mathcal{D}_1^{(l)}$ from the calibration experiment), is within a certain b% (b=5 and 10) of the measured quantity $\delta L_v^{(i)}$ in the validation experiment. This probability is given by the following updated robust predictive PDF conditioned on $\mathcal{D}_1^{(l)}$:

$$\begin{aligned} P(e_{v,p}^{(i)} \le b\% \,|\, \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}) &= \int P(e_{v,p}^{(i)} \le b\% \,|\, \boldsymbol{\theta}^{(2,j)}, \mathcal{M}_j^{(2)})\, p(\boldsymbol{\theta}^{(2,j)} \,|\, \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})\, d\boldsymbol{\theta}^{(2,j)} \\ &= \int P(e_{v,p}^{(i)} \le b\% \,|\, \boldsymbol{\theta}^{(1,j+1)}, \mathcal{M}_j^{(2)})\, p(\boldsymbol{\theta}^{(1,j+1)} \,|\, \mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)})\, d\boldsymbol{\theta}^{(1,j+1)} \end{aligned} \tag{5.34}$$

where

$$e_{v,p}^{(i)} = \left| \frac{\delta L_{v,p} - \delta L_v^{(i)}}{\delta L_v^{(i)}} \right| \tag{5.35}$$

For convenience, the superscripts in $\boldsymbol{\theta}^{(i,j)}$ will now be omitted. For the model class $\mathcal{M}_j^{(2)}$, j=1,2,3, given $\boldsymbol{\theta}$, it can be shown that the response $\delta L_{v,p}$ follows a Gaussian distribution with mean $\mu_v = K_v \mu_s$ and variance $\sigma_{v,j}^2 = \sigma_s^2 s_{v,j}(l_s, r)$ where K_v and $s_{v,j}$ are given as follows:

$$K_v = \frac{F_v L_v}{A_v} \tag{5.36}$$

$$s_{v,j}(l_s, r) = 2\left(\frac{F_v}{A_v}\right)^2 \int_0^{L_v} (L_v - x) \exp\left(-\left(\frac{x}{l_s}\right)^r\right) dx \tag{5.37}$$

For j=1, r is equal to 1, and $s_{v,j}(l_s,r)$ is given by (5.9) with subscript 'c' replaced by 'v' and for j=2, r is equal to 2, and $s_{v,j}(l_s,r)$ is given by (5.11) with subscript 'c' replaced by 'v'. Thus, the probability $P(e_{v,p}^{(i)} \le b\% | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})$ in (5.34) becomes:

$$\begin{aligned} & P(e_{v,p}^{(i)} \le b\% \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}) \\ &= \int \left[\Phi\left(\frac{\left(1+\frac{b}{100}\right)\delta L_v^{(i)} - \mu_v(\boldsymbol{\theta})}{\sigma_{v,j}(\boldsymbol{\theta})}\right) - \Phi\left(\frac{\left(1-\frac{b}{100}\right)\delta L_v^{(i)} - \mu_v(\boldsymbol{\theta})}{\sigma_{v,j}(\boldsymbol{\theta})}\right)\right] p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)}) d\boldsymbol{\theta} \\ &\approx \frac{1}{K} \sum_{k=1}^{K} \left[\Phi\left(\frac{\left(1+\frac{b}{100}\right)\delta L_v^{(i)} - \mu_v(\boldsymbol{\theta}^{(k)})}{\sigma_{v,j}(\boldsymbol{\theta}^{(k)})}\right) - \Phi\left(\frac{\left(1-\frac{b}{100}\right)\delta L_v^{(i)} - \mu_v(\boldsymbol{\theta}^{(k)})}{\sigma_{v,j}(\boldsymbol{\theta}^{(k)})}\right)\right] \end{aligned} \tag{5.38}$$

where $\boldsymbol{\theta}^{(k)}$, k=1,2,...,K, are posterior samples from $p(\boldsymbol{\theta}^{(1,\, j+1)} | \mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)})$. Similar to before, samples of $\delta L_{v,p}$ can be obtained as follows: For each $\boldsymbol{\theta}^{(k)}$, k=1,2,...,K, which are the posterior samples from $p(\boldsymbol{\theta} | \mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)})$, generate a sample $\delta L_{v,p}^{(k)}$ for $\delta L_{v,p}$ from a Gaussian distribution with mean $\mu_v(\boldsymbol{\theta})$ and variance $\sigma_{v,j}^2(\boldsymbol{\theta})$. These samples can also be used to find the above probability by approximating it as the proportion of samples that satisfies the condition

$e_{v,p}^{(i)} \leq b\%$ out of the K samples. It can be shown that the estimator in (5.38) is always of a smaller c.o.v. and thus more accurate than the latter approximation.

The average prediction error probability, denoted $P(e_{v,p} \leq b\% | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})$, for a model class updated using data $\mathcal{D}_1^{(l)}$ can be obtained by taking the arithmetic mean of $P(e_{v,p}^{(i)} \leq b\% | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})$, i=1, 2…, N_v. Tables 5.13, 5.14 and 5.15 show the results for $P(e_{v,p}^{(i)} \leq b\% | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})$ (the numbers outside the parenthesis) and their average $P(e_{v,p} \leq b\% | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})$ (the numbers inside the parenthesis) and for j=1, 2, 3, and b=5 and 10 using $\mathcal{D}_1^{(1)}$, $\mathcal{D}_1^{(2)}$ and $\mathcal{D}_1^{(3)}$, respectively. It can be seen from these tables that the model classes $\mathcal{M}_j^{(2)}$ (and so $\mathcal{M}_{j+1}^{(1)}$ updated using $\mathcal{D}_1^{(l)}$), for j=1, 2, 3, are sufficiently accurate. It is noted that the averages $P(e_{v,p} \leq 5\% | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})$ for each j=1, 2, 3, are larger than 0.5 implying that it is more likely than not for the response prediction by the model classes to be accurate within 5% of the actual response. The averages $P(e_{v,p} \leq 10\% | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})$ are all very close to 1, showing that it is very probable that the prediction errors for each model class are less than 10%.

To evaluate *prediction consistency*, we calculate the difference of the measured quantity $\delta L_v^{(i)}$ and the posterior mean $E[\delta L_{v,p} | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}]$ of the robust predicted response (measured in terms of the number of posterior standard deviations $\sqrt{Var[\delta L_{v,p} | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}]}$) as follows:

$$c_{v,j}^{(i)} = \frac{\delta L_v^{(i)} - E[\delta L_{v,p} | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}]}{\sqrt{Var[\delta L_{v,p} | \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}]}} \tag{5.39}$$

where

$$E[\delta L_{v,p} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}] = \int E[\delta L_{v,p} \mid \boldsymbol{\theta}, \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}] p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}) d\boldsymbol{\theta}$$
$$= \int \mu_v(\boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)}) d\boldsymbol{\theta} = K_v \int \mu_s p(\mu_s \mid \mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)}) d\mu_s \approx \frac{K_v}{K} \sum_{k=1}^{K} \mu_s^{(k)} \tag{5.40}$$

where $\mu_s^{(k)}$ is the first component of $\boldsymbol{\theta}^{(k)}$, where $\boldsymbol{\theta}^{(k)}$, k=1,2,...,K, are posterior samples from $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)})$. $Var[\delta L_{v,p} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}]$ is given by:

$$Var[\delta L_{v,p} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}] = E[\delta L_{v,p}{}^2 \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}] - E^2[\delta L_{v,p} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}] \tag{5.41}$$

where

$$E[\delta L_{v,p}{}^2 \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}] = \int E[\delta L_{v,p}{}^2 \mid \boldsymbol{\theta}, \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}] p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}) d\boldsymbol{\theta}$$
$$= \int (\mu_v{}^2(\boldsymbol{\theta}) + \sigma_{v,j}{}^2(\boldsymbol{\theta})) p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)}) d\boldsymbol{\theta} \approx \frac{1}{K} \sum_{k=1}^{K} [\mu_v{}^2(\boldsymbol{\theta}^{(k)}) + \sigma_{v,j}{}^2(\boldsymbol{\theta}^{(k)})] \tag{5.42}$$

where $\boldsymbol{\theta}^{(k)}$, k=1,2,...,K, are posterior samples from $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)})$. The last rows of Tables 5.13, 5.14 and 5.15 show the results for $c_{v,j}^{(i)}$, for j=1, 2, 3 using $\mathcal{D}_1^{(1)}$, $\mathcal{D}_1^{(2)}$ and $\mathcal{D}_1^{(3)}$, respectively. It can be seen from these tables that for all three data cases, the model classes $\mathcal{M}_j^{(2)}$ (and also $\mathcal{M}_{j+1}^{(1)}$) updated just using data $\mathcal{D}_1^{(l)}$, j=1, 2, 3, are sufficiently consistent since the results are all within about 3 standard deviations.

Using data $\mathcal{D}_2^{(l)}$, which is modeled as stochastically independent of $\mathcal{D}_1^{(l)}$ given $\boldsymbol{\theta}$, one can update uncertainties in $\boldsymbol{\theta}$ for all surviving model classes using Bayes' Theorem with $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})$ as the prior (recall that in this case, $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}) = p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)})$):

$$p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)}) = c_2^{-1} p(\mathcal{D}_2^{(l)} \mid \boldsymbol{\theta}, \mathcal{M}_j^{(2)}) p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}) \tag{5.43}$$

where the likelihood function is given by:

$$p(\mathcal{D}_2^{(l)} \mid \boldsymbol{\theta}, \mathcal{M}_j^{(2)}) = \frac{1}{(2\pi\sigma_{v,j}(\boldsymbol{\theta})^2)^{N_v/2}} \exp(-\frac{1}{2\sigma_{v,j}{}^2(\boldsymbol{\theta})} \sum_{i=1}^{N_v} (\delta L_v{}^{(i)} - \mu_v(\boldsymbol{\theta}))^2 \tag{5.44}$$

and the evidence $p(\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)}|\mathcal{M}_j^{(2)})$ for model class $\mathcal{M}_j^{(2)}$ provided by the data $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$ is given by:

$$p(\mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)} \mid \mathcal{M}_j^{(2)}) = p(\mathcal{D}_1^{(l)} \mid \mathcal{M}_j^{(2)}) p(\mathcal{D}_2^{(l)} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}) \tag{5.45}$$

where $p(\mathcal{D}_1^{(l)}|\mathcal{M}_j^{(2)})$ is equal to $p(\mathcal{D}_1^{(l)}|\mathcal{M}_{j+1}^{(1)})$, which has already been determined from previous analyses, while $p(\mathcal{D}_2^{(l)}|\mathcal{D}_1^{(l)},\mathcal{M}_j^{(2)})$ is given by:

$$p(\mathcal{D}_2^{(l)} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}) = \int p(\mathcal{D}_2^{(l)} \mid \boldsymbol{\theta}, \mathcal{M}_j^{(2)}) p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)}) d\boldsymbol{\theta} \tag{5.46}$$

which is determined using the stochastic simulation method in Appendix B as before. The samples from the prior $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_j^{(2)})$ (calibration test posterior $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)}, \mathcal{M}_{j+1}^{(1)})$) obtained from the previous analyses, are used.

Tables 5.16, 5.17 and 5.18 show the statistical results using data $\mathcal{D}_2^{(l)}$ in addition to $\mathcal{D}_1^{(l)}$ for the three validation cases of N_v = 2, 4 and 10 datapoints, respectively. Compared to Tables 5.10, 5.11 and 5.12, it can be seen that the posterior c.o.v. of the parameters updated using additional data $\mathcal{D}_2^{(l)}$ is reduced somewhat for σ_s^2, l_s () for data cases 1 and 3 but for a somewhat lesser amount for data case 2. For data case 2, $\mathcal{D}_2^{(l)}$ provides only 20% additional data while for data cases 1 and 3, $\mathcal{D}_2^{(l)}$ provides 40% and 33% additional data respectively (see Table 5.8). For all data cases, the posterior means of the parameters σ_s^2 and l_s using $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$ are significantly

higher than the means using only $\mathcal{D}_1^{(l)}$. There are several possible reasons: 1) additional information is provided by the additional data $\mathcal{D}_2^{(l)}$; and 2) uncertainties of the estimators due to a finite number of samples used in the stochastic simulation. Similar to before, it can be seen from the posterior correlation coefficient matrix that there is only weak correlation between most pairs of parameters. The posterior means of r in $\mathcal{M}_3^{(2)}$ are very close for all 3 data cases: 1.81, 1.83 and 1.79 but the corresponding uncertainty in r is still significant. The results show that given both $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$, for all 3 data cases, $\mathcal{M}_1^{(2)}$, $\mathcal{M}_2^{(2)}$ and $\mathcal{M}_3^{(2)}$ are significantly probable. Thus, based on the calibration data and validation data, all the model classes $\mathcal{M}_1^{(2)}$, $\mathcal{M}_2^{(2)}$ and $\mathcal{M}_3^{(2)}$ are considered in subsequent analyses.

It can also be seen that the predicted robust failure probability $P(F|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{M}_2^{(2)})$ of the target frame structure using model class $\mathcal{M}_2^{(2)}$ is smaller than that using model classes $\mathcal{M}_1^{(2)}$ and $\mathcal{M}_3^{(2)}$. For data cases 1, 2 and 3, the predicted hyper-robust failure probabilities $P(F|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},M_2)$ are estimated to be 1.44×10^{-3}, 2.25×10^{-4} and 1.25×10^{-5}, respectively, showing that the predicted failure probability of the target system depends on the uncertainties in the model parameters, which in turn depends on the amount of data and the model classes under consideration. By comparing Tables 5.10-5.12 and Tables 5.16-5.18, it can be seen that the predicted hyper-robust failure probability is significantly smaller than that based on only data $\mathcal{D}_1^{(l)}$ for all data cases.

Table 5.19 shows the results for checking, using the following index, the consistency of the model classes $\mathcal{M}_j^{(2)}$, j =1, 2, 3, in predicting the response δL_v using data $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$:

$$\frac{\delta L_v^{(i)} - E[\delta L_{v,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)}]}{\sqrt{Var[\delta L_{v,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)}]}} \tag{5.47}$$

where $E[\delta L_{v,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)}]$ and $Var[\delta L_{v,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)}]$ can be determined using (5.40), (5.41) and (5.42) except that the samples from the most recently updated posterior PDF $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{M}_j^{(2)})$ are used instead of $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{M}_j^{(2)})$. By comparing Tables 5.13-5.15 and Table 5.19, it can be seen that the consistency of the model classes improves over the case without data $\mathcal{D}_2^{(l)}$.

The accuracy of the model classes $\mathcal{M}_j^{(2)}$, j =1, 2, 3, in predicting δL_v using data $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$ can be assessed, similar to the case without data $\mathcal{D}_2^{(l)}$, by evaluating i) $P(e_{v,p}^{(i)} \leq b\% | \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)})$, i=1, 2…, N_v, which can be determined using (5.38) except that the samples from the most recently updated posterior PDF $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{M}_j^{(2)})$ are used instead, and ii) the average prediction error probability $P(e_{v,p} \leq b\% | \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)})$ of a model class updated using data $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$ which can be obtained by taking the arithmetic mean of $P(e_{v,p}^{(i)} \leq b\% | \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)})$, i=1, 2…, N_v. The corresponding results are not shown here for brevity but they show high probability that the prediction errors for each model class will be less than 5%, with even higher probabilities for 10%..

Since the system involved in this experiment is just a longer bar subjected to the same load with the same boundary conditions and other geometrical properties, it is reasonable to use the data collected in the most recent experiment to update the uncertainties in the model parameters considered in the previous experiment. However, if the system in the validation experiment was very different from that in the previous experiment, additional parameters may have to be introduced to take into account the additional uncertainties involved.

Table 5.13 Results of predicting δL_v using data $\mathcal{D}_1^{(1)}$ from the calibration experiment

	$\mathcal{M}_1^{(2)}$	$\mathcal{M}_2^{(2)}$	$\mathcal{M}_3^{(2)}$
$P(\lvert\delta L_{v,p}-\delta L_v^{(i)}\rvert/\lvert\delta L_v^{(i)}\rvert\le 5\%\lvert\mathcal{D}_1^{(1)},\mathcal{M}_j^{(2)})$	0.469,0.775 (0.622)	0.481,0.778 (0.629)	0.470, 0.769 (0.620)
$P(\lvert\delta L_{v,p}-\delta L_v^{(i)}\rvert/\lvert\delta L_v^{(i)}\rvert\le 10\%\lvert\mathcal{D}_1^{(1)},\mathcal{M}_j^{(2)})$	0.928,0.968 (0.948)	0.930;0.981 (0.955)	0.924,0.976 (0.950)
5 percentile/95percentile of $\delta L_{v,p}$	2.00×10^{-4},2.22×10^{-4}	2.00×10^{-4},2.23×10^{-4}	1.99×10^{-4},2.23×10^{-4}
$\frac{\delta L_v^{(i)} - E[\delta L_{v,p}\lvert\mathcal{D}_1^{(1)},\mathcal{M}_j^{(2)}]}{\sqrt{Var[\delta L_{v,p}\lvert\mathcal{D}_1^{(1)},\mathcal{M}_j^{(2)}]}}$	-1.40,-0.73	-1.55,-0.82	-1.30,-0.66

Table 5.14 Results of predicting δL_v using data $\mathcal{D}_1^{(2)}$ from the calibration experiment

	$\mathcal{M}_1^{(2)}$	$\mathcal{M}_2^{(2)}$	$\mathcal{M}_3^{(2)}$
$P(\lvert\delta L_{v,p}-\delta L_v^{(i)}\rvert/\lvert\delta L_v^{(i)}\rvert\le 5\%\lvert\mathcal{D}_1^{(2)},\mathcal{M}_j^{(2)})$	0.372,0.729, 0.372,0.835 (0.577)	0.347,0.752, 0.347,0.862 (0.577)	0.374,0.744, 0.374,0.849 (0.585)
$P(\lvert\delta L_{v,p}-\delta L_v^{(i)}\rvert/\lvert\delta L_v^{(i)}\rvert\le 10\%\lvert\mathcal{D}_1^{(2)},\mathcal{M}_j^{(2)})$	0.931,0.992, 0.931,0.995 (0.920)	0.944,0.995, 0.944,0.998 (0.970)	0.932,0.992, 0.932,0.997 (0.963)
5 percentile/95percentile of $\delta L_{v,p}$	2.04×10^{-4},2.22×10^{-4}	2.04×10^{-4},2.21×10^{-4}	2.04×10^{-4},2.22×10^{-4}
$\frac{\delta L_v^{(i)} - E[\delta L_{v,p}\lvert\mathcal{D}_1^{(2)},\mathcal{M}_j^{(2)}]}{\sqrt{Var[\delta L_{v,p}\lvert\mathcal{D}_1^{(2)},\mathcal{M}_j^{(2)}]}}$	-2.17,-1.26, -2.17,-0.90	-2.31,-1.34, -2.31,-0.95	-2.31,-1.22, -2.13,-0.86

Table 5.15 Results of predicting δL_v using data $\mathcal{D}_1^{(3)}$ from the calibration experiment

	$\mathcal{M}_1^{(2)}$	$\mathcal{M}_2^{(2)}$	$\mathcal{M}_3^{(2)}$
$P(\lvert\delta L_{v,p}-\delta L_v^{(i)}\rvert/\lvert\delta L_v^{(i)}\rvert\le 5\%\lvert\mathcal{D}_1^{(3)},\mathcal{M}_j^{(2)})$	0.325,0.732, 0.325,0.844, 0.579,0.325, 0.732,0.943, 0.149,0.844 (0.579)	0.368,0.774, 0.368,0.882, 0.624,0.368, 0.774,0.956, 0.160,0.882 (0.615)	0.327,0.730, 0.327,0.846, 0.579,0.327, 0.730,0.944, 0.137,0.846 (0.579)
$P(\lvert\delta L_{v,p}-\delta L_v^{(i)}\rvert/\lvert\delta L_v^{(i)}\rvert\le 10\%\lvert\mathcal{D}_1^{(3)},\mathcal{M}_j^{(2)})$	0.940,0.994, 0.940,0.997, 0.984,0.940, 0.994,0.999, 0.815,0.997 (0.960)	0.956,0.997, 0.956,0.998, 0.988,0.956, 0.997,0.999, 0.854,0.998 (0.970)	0.943,0.993, 0.943,0.999, 0.984,0.943, 0.993,0.999, 0.817,0.999 (0.961)
5 percentile/95percentile of $\delta L_{v,p}$	2.05×10^{-4},2.22×10^{-4}	2.05×10^{-4},2.21×10^{-4}	2.05×10^{-4},2.21×10^{-4}
$\frac{\delta L_v^{(i)} - E[\delta L_{v,p}\mid\mathcal{D}_1^{(3)},\mathcal{M}_j^{(2)}]}{\sqrt{Var[\delta L_{v,p}\mid\mathcal{D}_1^{(3)},\mathcal{M}_j^{(2)}]}}$	-2.40,-1.42 -2.40,-1.02 -1.81,-2.40 -1.42,-0.43 -2.99,-1.02	-2.40,-1.38 -2.40,-0.97 -1.79,-2.40 -1.38,-0.35 -3.01,-0.97	-2.41,-1.42 -2.40,-1.03 -1.82,-2.41 -1.42,-0.44 -3.00,-1.03

Table 5.16 Statistical results using data $\mathcal{D}_2^{(1)}$ from the validation experiment in addition to $\mathcal{D}_1^{(1)}$

		$\mathcal{M}_1^{(2)}$	$\mathcal{M}_2^{(2)}$	$\mathcal{M}_3^{(2)}$
Parameter Statistics	μ_s (Pa^{-1})	8.60×10^{-11};1.7%	8.60×10^{-11};1.5%	8.60×10^{-11};1.4%
	σ_s^2(Pa^{-2})	1.11×10^{-22};39.8%	1.14×10^{-22};44.5%	1.08×10^{-22};37.1%
	l_s(m)	0.0238;82.0%	0.0229;51.2%	0.0218;66.7%
	r			1.81;44.1%
	R	$\begin{bmatrix} 1 & 0.03 & 0.22 \\ & 1 & 0.09 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.01 & 0.02 \\ & 1 & -0.10 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.10 & 0.03 & 0.10 \\ & 1 & 0.08 & 0.02 \\ & & 1 & 0.04 \\ & & & 1 \end{bmatrix}$
Log evidence		192.08	192.07	192.81
$E[\ln p(\mathcal{D}_1^{(1)},\mathcal{D}_2^{(1)}\|\boldsymbol{\theta},\mathcal{M}_j^{(2)})]$		198.58	198.67	198.77
Expected Information gain		6.51	6.60	5.96
$P(\mathcal{M}_j^{(2)}\|\mathcal{D}_1^{(1)},\mathcal{D}_2^{(1)},M_2)$		0.246	0.244	0.510
$P(F\|\mathcal{D}_1^{(1)},\mathcal{D}_2^{(1)},\mathcal{M}_j^{(2)})$		3.47×10^{-3}(14.4%)	4.42×10^{-4}(17.7%)	9.32×10^{-4}(17.7%)
$P(F\|\mathcal{D}_1^{(1)},\mathcal{D}_2^{(1)},M_2)$		1.44×10^{-3}		

Table 5.17 Statistical results using data $\mathcal{D}_2^{(2)}$ from the validation experiment in addition to $\mathcal{D}_1^{(2)}$

		$\mathcal{M}_1^{(2)}$	$\mathcal{M}_2^{(2)}$	$\mathcal{M}_3^{(2)}$
Parameter Statistics	μ_s (Pa^{-1})	8.72×10^{-11};0.90%	8.71×10^{-11};0.84%	8.74×10^{-11};0.89%
	σ_s^2(Pa^{-2})	6.04×10^{-23};23.5%	6.40×10^{-23};22.8%	6.17×10^{-23};22.6%
	l_s(m)	0.0386;33.3%	0.0341;25.2%	0.036;35.5%
	r			1.83;45.2 %
	R	$\begin{bmatrix} 1 & -0.15 & 0.11 \\ & 1 & -0.03 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.08 & 0.22 \\ & 1 & -0.29 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.005 & 0.15 & -0.25 \\ & 1 & -0.15 & -0.14 \\ & & 1 & -0.16 \\ & & & 1 \end{bmatrix}$
Log evidence		749.95	750.84	751.41
E[ln$p(\mathcal{D}_1^{(2)},\mathcal{D}_2^{(2)}\|\boldsymbol{\theta},\mathcal{M}_j^{(2)})$]		758.44	758.92	758.73
Expected Information gain		8.48	8.08	7.32
$P(\mathcal{M}_j^{(2)}\|\mathcal{D}_1^{(2)},\mathcal{D}_2^{(2)},M_2)$		0.129	0.315	0.556
$P(F\|\mathcal{D}_1^{(2)},\mathcal{D}_2^{(2)},\mathcal{M}_j^{(2)})$		1.82×10^{-4}(16.5%)	1.84×10^{-5}(24.8%)	3.51×10^{-4}(15.5%)
$P(F\|\mathcal{D}_1^{(2)},\mathcal{D}_2^{(2)},M_2)$		2.25×10^{-4}		

Table 5.18 Statistical results using data $\mathcal{D}_2^{(3)}$ from the validation experiment in addition to $\mathcal{D}_1^{(3)}$

		$\mathcal{M}_1^{(2)}$	$\mathcal{M}_2^{(2)}$	$\mathcal{M}_3^{(2)}$
Parameter Statistics	μ_s (Pa^{-1})	8.70×10^{-11};0.62%	8.68×10^{-11};0.63%	8.68×10^{-11};0.6%
	σ_s^2(Pa^{-2})	6.00×10^{-23};17.2%	5.80×10^{-23};17.8%	5.70×10^{-23};20.1%
	l_s(m)	0.0383;25.4%	0.0384;19.0%	0.0398;25.6%
	r			1.79;39.6%
	R	$\begin{bmatrix} 1 & 0.02 & 0.08 \\ & 1 & -0.14 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.04 & 0.10 \\ & 1 & -0.10 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.02 & 0.10 & -0.13 \\ & 1 & -0.28 & 0.26 \\ & & 1 & -0.41 \\ & & & 1 \end{bmatrix}$
Log evidence		1174.56	1173.82	1173.83
$E[\ln p(\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)}\|\boldsymbol{\theta},\mathcal{M}_j^{(2)})]$		1182.70	1182.83	1182.72
Expected Information gain		8.14	9.01	8.90
$P(\mathcal{M}_j^{(2)}\|\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)},M_2)$		0.510	0.244	0.246
$P(F\|\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)},\mathcal{M}_j^{(2)})$		1.32×10^{-5}(20.6%)	3.43×10^{-6}(32.1%)	1.99×10^{-5}(22.2%)
$P(F\|\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)},M_2)$		1.25×10^{-5}		

Table 5.19 Consistency assessment of model classes in predicting δL_v using data $\mathcal{D}_2^{(l)}$ from the validation experiment in addition to $\mathcal{D}_1^{(l)}$ from the calibration experiment

	$\mathcal{M}_1^{(2)}$	$\mathcal{M}_2^{(2)}$	$\mathcal{M}_3^{(2)}$
Data case 1, l=1	-0.79,-0.06	-0.82,-0.05	-0.87,-0.08
Data case 2, l=2	-1.39,-0.55, -1.39,-0.21	-1.49,-0.57, -1.49,-0.20	-1.50,-0.64, -1.50,-0.29
Data case 3, l=3	-1.30,-0.43, -1.30,-0.08, -0.78,-1.30, -0.43,0.44, -1.83,-0.08	-1.34,-0.42, -1.34,-0.05, -0.79,-1.34, -0.42,0.50, -1.90,-0.05	-1.33,-0.43, -1.33,-0.07, -0.79,-1.33, -0.43,0.47, -1.87,-0.07

5.3 Using data $\mathcal{D}_3^{(l)}$ from the accreditation experiment

Candidate model classes for the subsystem in the accreditation experiment are $\mathcal{M}_j^{(3)}$, j=1,2,3. The uncertain parameters $\boldsymbol{\theta}^{(3,j)}$ for $\mathcal{M}_j^{(3)}$ are the same as $\boldsymbol{\theta}^{(2,j)}$ for $\mathcal{M}_j^{(2)}$. The "prior" PDF $p(\boldsymbol{\theta}^{(3,j)}|\mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)})$ for $\mathcal{M}_j^{(3)}$ is given by the "posterior" PDF $p(\boldsymbol{\theta}^{(2,j)}|\mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)})$ for $\mathcal{M}_j^{(2)}$. Similar analyses to before are carried out as follows. Data $\mathcal{D}_3^{(l)} = \{w_a^{(i)}, i=1,\ldots, N_a\}$ from the accreditation experiment are used to investigate the predictive performance of the model classes. The probability that the response $w_{a,p}$ (the vertical displacement of point Q of the frame structure in the accreditation experiment) predicted using the model classes updated by data from the previous two experiments is within a certain b% of the measured quantity $w_a^{(i)}$ is given by the following updated robust predictive PDF conditioned on $\mathcal{D}_1^{(l)}$and $\mathcal{D}_2^{(l)}$:

$$\begin{aligned} P(e_{a,p}{}^{(i)} \le b\% \,|\, \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}) &= \int P(e_{a,p}{}^{(i)} \le b\% \,|\, \boldsymbol{\theta}, \mathcal{M}_j^{(3)})\, p(\boldsymbol{\theta} \,|\, \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)})\, d\boldsymbol{\theta} \\ &= \int P(e_{a,p}{}^{(i)} \le b\% \,|\, \boldsymbol{\theta}, \mathcal{M}_j^{(3)})\, p(\boldsymbol{\theta} \,|\, \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)})\, d\boldsymbol{\theta} \end{aligned} \tag{5.48}$$

where

$$e_{a,p}{}^{(i)} = \left| \frac{w_{a,p} - w_a^{(i)}}{w_a^{(i)}} \right| \tag{5.49}$$

For the model class $\mathcal{M}_j^{(3)}$, j=1, 2, 3, given $\boldsymbol{\theta}$, it can be shown that the response $w_{a,p}$ follows a Gaussian distribution with mean $\mu_a = K_a \mu_s$ and variance $\sigma_{a,j}{}^2 = \sigma_s^2 s_{a,j}(l_s, r)$ where K_a is given as follows:

$$K_a = \frac{1}{2}[\frac{F_1 L_1}{A_1} - \sqrt{2}(\frac{F_2 L_2}{A_2} + \frac{F_4 L_4}{A_4})] - \frac{F_a L_1^3}{48I} \tag{5.50}$$

The expression for $s_{a,j}$ is given in the Appendix C. Thus,

$$\begin{aligned}
&P(e_{a,p}{}^{(i)} \le b\% \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}) \\
&= \mathrm{sgn}(w_a^{(i)}) \int [\Phi(\frac{(1+\frac{b}{100})w_a^{(i)} - \mu_a(\boldsymbol{\theta})}{\sigma_{a,j}(\boldsymbol{\theta})}) - \Phi(\frac{(1-\frac{b}{100})w_a^{(i)} - \mu_a(\boldsymbol{\theta})}{\sigma_{a,j}(\boldsymbol{\theta})})] p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)}) d\boldsymbol{\theta} \\
&\approx \frac{\mathrm{sgn}(w_a^{(i)})}{K} \sum_{k=1}^{K} \Phi(\frac{(1+\frac{b}{100})w_a^{(i)} - \mu_a(\boldsymbol{\theta}^{(k)})}{\sigma_{a,j}(\boldsymbol{\theta}^{(k)})}) - \Phi(\frac{(1-\frac{b}{100})w_a^{(i)} - \mu_a(\boldsymbol{\theta}^{(k)})}{\sigma_{a,j}(\boldsymbol{\theta}^{(k)})})
\end{aligned} \tag{5.51}$$

where $\boldsymbol{\theta}^{(k)}$, k=1,2,...,K, are posterior samples from $p(\boldsymbol{\theta}|\mathcal{D}_1{}^{(l)},\mathcal{D}_2{}^{(l)},\mathcal{M}_j^{(2)})$.

Tables 5.20, 5.21 and 5.22 show the results for $P(e_{a,p}{}^{(i)} \le b\% | \mathcal{D}_1{}^{(l)},\mathcal{D}_2{}^{(l)},\mathcal{M}_j^{(3)})$ (the numbers outside the parenthesis) and the average prediction error probability $P(e_{a,p} \le b\% | \mathcal{D}_1{}^{(l)},\mathcal{D}_2{}^{(l)},\mathcal{M}_j^{(3)})$ (the numbers inside the parenthesis), for j=1, 2, 3, and b=5 and 10 using $\mathcal{D}_1{}^{(l)}$ and $\mathcal{D}_2{}^{(l)}$ data cases 1, 2 and 3, respectively. It can be seen from these tables that for all three data cases, the model classes $\mathcal{M}_j^{(3)}$ (and so $\mathcal{M}_j^{(2)}$), j=1, 2, 3, updated using $\mathcal{D}_1{}^{(l)}$ and $\mathcal{D}_2{}^{(l)}$, are sufficiently accurate. It is noted that all $P(e_{a,p} \le 5\% | \mathcal{D}_1{}^{(l)},\mathcal{D}_2{}^{(l)},\mathcal{M}_j^{(3)})$ are larger than 0.84 implying that there is a high probability for the response prediction by the model classes to be within 5% of the actual response measurements.

The difference between the measured quantity $w_a^{(i)}$ and the posterior mean $E[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}]$ of the robust predicted response (measured in terms of the number of posterior standard deviations $\sqrt{Var[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}]}$) is given by:

$$c_{a,j}^{(i)} = \frac{w_a^{(i)} - E[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}]}{\sqrt{Var[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}]}} \tag{5.52}$$

where $E[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}]$ is given by:

$$\begin{aligned} &E[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}] = \int \mu_a(\boldsymbol{\theta})\, p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}) d\boldsymbol{\theta} \\ &= \int \mu_a(\boldsymbol{\theta})\, p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)}) d\boldsymbol{\theta} \\ &= K_a \int \mu_s\, p(\mu_s \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}) d\mu_s \approx \frac{K_a}{K} \sum_{k=1}^{K} \mu_s^{(k)} \end{aligned} \tag{5.53}$$

where $\mu_s^{(k)}$ is the first component of $\boldsymbol{\theta}^{(k)}$, where $\boldsymbol{\theta}^{(k)}$, k=1,2,...,K, are posterior samples from $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{M}_j^{(2)})$. $Var[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}]$ is given by:

$$Var[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}] = E[w_{a,p}{}^2 \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}] - E^2[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}] \tag{5.54}$$

where $E[w_{a,p}{}^2 \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}]$ is given by:

$$\begin{aligned} &E[w_{a,p}{}^2 \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}] = \int [\mu_a{}^2(\boldsymbol{\theta}) + \sigma_{a,j}{}^2(\boldsymbol{\theta})]\, p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}) d\boldsymbol{\theta} \\ &= \int [\mu_a{}^2(\boldsymbol{\theta}) + \sigma_{a,j}{}^2(\boldsymbol{\theta})]\, p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(2)}) d\boldsymbol{\theta} \approx \frac{1}{K} \sum_{k=1}^{K} [\mu_a{}^2(\boldsymbol{\theta}^{(k)}) + \sigma_{a,j}{}^2(\boldsymbol{\theta}^{(k)})] \end{aligned} \tag{5.55}$$

The last rows of Tables 5.20, 5.21 and 5.22 show the results for $c_{a,j}^{(i)}$, j=1, 2, 3, using $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$ data cases 1, 2 and 3, respectively. It can be seen from these tables that for all three data cases, the model classes $\mathcal{M}_j^{(3)}$ (and so $\mathcal{M}_j^{(2)}$), j=1, 2, 3, updated using $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$ are sufficiently consistent since the results are all within a standard deviation.

Using data $\mathcal{D}_3^{(l)}$, which is modelled as stochastically independent of $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$ given $\boldsymbol{\theta}$, one can update the uncertainties in $\boldsymbol{\theta}$ for all the model classes using Bayes' Theorem with the previous posterior PDF $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{M}_j^{(2)})$ as the prior $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{M}_j^{(3)})$:

$$p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)}, \mathcal{M}_j^{(3)}) = c_3^{-1} p(\mathcal{D}_3^{(l)} \mid \boldsymbol{\theta}, \mathcal{M}_j^{(3)}) p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}) \tag{5.56}$$

where the likelihood function is given by:

$$p(\mathcal{D}_3^{(l)} \mid \boldsymbol{\theta}, \mathcal{M}_j^{(3)}) = \frac{1}{(2\pi\sigma_{a,j}(\boldsymbol{\theta})^2)^{N_a/2}} \exp(-\frac{1}{2\sigma_{a,j}^2(\boldsymbol{\theta})} \sum_{i=1}^{N_a} (w_a^{(i)} - \mu_a(\boldsymbol{\theta}))^2 \tag{5.57}$$

The evidence $p(\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{D}_3^{(l)}|\mathcal{M}_j^{(3)})$ for model class $\mathcal{M}_j^{(3)}$ that is provided by the data $\mathcal{D}_1^{(l)}$, $\mathcal{D}_2^{(l)}$ and $\mathcal{D}_3^{(l)}$ is given by:

$$p(\mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)} \mid \mathcal{M}_j^{(3)}) = p(\mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)} \mid \mathcal{M}_j^{(3)}) p(\mathcal{D}_3^{(l)} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}) \tag{5.58}$$

where $p(\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)}|\mathcal{M}_j^{(3)})$ has already been determined and $p(\mathcal{D}_3^{(l)}|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{M}_j^{(3)})$ is given by:

$$p(\mathcal{D}_3^{(l)} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}) = \int p(\mathcal{D}_3^{(l)} \mid \boldsymbol{\theta}, \mathcal{M}_j^{(3)}) p(\boldsymbol{\theta} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{M}_j^{(3)}) d\boldsymbol{\theta} \tag{5.59}$$

which is determined using the same stochastic simulation method as before. The samples from the prior $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{M}_j^{(3)})$ obtained from the previous analyses are used.

The system involved in this accreditation experiment is a lot more complicated than the one in the validation experiment. In practice, one may want to consider introducing additional parameters to take into account the additional uncertainties involved. Nonetheless, for illustration, we have kept the same number of uncertain parameters as before, which is consistent with the statement of the validation challenge problem, and use data $\mathcal{D}_3^{(l)}$ to update the uncertainties in the parameters. Tables 5.23, 5.24 and 5.25 show the statistical results using data $\mathcal{D}_3^{(l)}$ in addition to the data from the previous experiments $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$ for the three data cases of N_a = 1, 1 and 2 respectively. Compared to Tables 5.16, 5.17 and 5.18, some of the differences observed in the posterior mean, c.o.v. and correlation coefficient of parameters are due to: 1) additional information provided by the additional data $\mathcal{D}_3^{(l)}$; and 2) uncertainties of the estimators due to a finite number of samples used in stochastic simulation. Similar to before, it can be seen from the posterior correlation coefficient matrix that there is only weak correlation between most pairs of parameters. The posterior mean of r in $\mathcal{M}_3^{(3)}$ is very close for all 3 data cases: 1.77, 1.90 and 1.81 but the uncertainty in r is still significant since $\mathcal{D}_3^{(l)}$ provides only 1 or 2 additional data. The results show that given $\mathcal{D}_1^{(l)}$, $\mathcal{D}_2^{(l)}$ and $\mathcal{D}_3^{(l)}$, $\mathcal{M}_1^{(3)}$, $\mathcal{M}_2^{(3)}$ and $\mathcal{M}_3^{(3)}$ are significantly probable and the posterior probabilities are essentially unchanged from Tables 5.16, 5.17 and 5.18,. Thus, all of the model classes $\mathcal{M}_1^{(3)}$, $\mathcal{M}_2^{(3)}$ and $\mathcal{M}_3^{(3)}$ are utilized to make robust predictions.

It can also be seen from Tables 5.23, 5.24 and 5.25 that the predicted robust failure probability $P(F|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{D}_3^{(l)},\mathcal{M}_2^{(3)})$ of the target frame structure using model class $\mathcal{M}_2^{(3)}$ is again smaller than that using model classes $\mathcal{M}_1^{(3)}$ and $\mathcal{M}_3^{(3)}$, especially for data cases 2 and 3. For data cases 1, 2 and 3, the predicted hyper-robust failure probabilities $P(F|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{D}_3^{(l)},M_3)$ are

4.18×10^{-4}, 1.48×10^{-4} and 1.51×10^{-5}, respectively, showing that the predicted failure probability of the target system depends on the uncertainties in the model parameters, which in turn depends on the amount of data and the model classes under consideration. By comparing Tables 5.16-5.18 and Tables 5.23-5.25, it can be seen that the predicted hyper-robust failure probability changes little compared to that based on only data $\mathcal{D}_1^{(l)}$ and $\mathcal{D}_2^{(l)}$ for the large data cases 2 and 3. $P(F|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{D}_3^{(l)},\mathcal{M}_2^{(3)})P(\mathcal{M}_2^{(3)}|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{D}_3^{(l)},M_3)$ is small compared to $P(F|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)},\mathcal{D}_3^{(l)},M_3)$ and thus the contribution of $\mathcal{M}_2^{(3)}$ to the prediction quantity of interest is small.

Table 5.26 shows the results for checking the consistency of the model classes $\mathcal{M}_j^{(3)}$, j =1, 2, 3, in predicting the response w_a using data $\mathcal{D}_1^{(l)}$, $\mathcal{D}_2^{(l)}$ and $\mathcal{D}_3^{(l)}$:

$$\frac{w_a^{(i)} - E[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)}, \mathcal{M}_j^{(3)}]}{\sqrt{Var[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)} \mathcal{M}_j^{(3)}]}} \tag{5.60}$$

where $E[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)}, \mathcal{M}_j^{(3)}]$ and $Var[w_{a,p} \mid \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)}, \mathcal{M}_j^{(3)}]$ can be determined by using (5.53), (5.54) and (5.55) except that the samples from the most recently updated posterior PDF $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)}, \mathcal{M}_j^{(3)})$ are used instead. By comparing Tables 5.20-5.22 and Table 5.26, it can be seen that the consistency of the model classes is similar to the case without data $\mathcal{D}_3^{(l)}$ since $\mathcal{D}_3^{(l)}$ provides only one or two additional data.

The accuracy of the model classes $\mathcal{M}_j^{(3)}$, j =1, 2, 3, in predicting w_a using data $\mathcal{D}_1^{(l)}$, $\mathcal{D}_2^{(l)}$ and $\mathcal{D}_3^{(l)}$ can be assessed, similar to the case without data $\mathcal{D}_3^{(l)}$, by evaluating i) $P(e_{a,p}^{(i)} \leq b\% | \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)}, \mathcal{M}_j^{(3)})$, i=1,…, N_a, which can be determined using (5.51) except that the samples from the most recently updated posterior PDF $p(\boldsymbol{\theta}|\mathcal{D}_1^{(l)},\mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)},\mathcal{M}_j^{(3)})$ are used instead, and ii) the average prediction error probability $P(e_{a,p} \leq b\% | \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)}, \mathcal{M}_j^{(3)})$ of a model class updated

using data $\mathcal{D}_1^{(l)}$, $\mathcal{D}_2^{(l)}$ and $\mathcal{D}_3^{(l)}$ which can be obtained by taking the arithmetic mean of $P(e_{a,p}^{(i)} \le b\% | \mathcal{D}_1^{(l)}, \mathcal{D}_2^{(l)}, \mathcal{D}_3^{(l)}, \mathcal{M}_j^{(3)})$, i=1, 2…, N_v. The corresponding results are not shown here for brevity but they show high prediction accuracy (high probability of prediction errors less than 5%).

Table 5.20 Results of predicting w_a using data $\mathcal{D}_2^{(1)}$ from the validation experiment in addition to $\mathcal{D}_1^{(1)}$ from the calibration experiment

	$\mathcal{M}_1^{(3)}$	$\mathcal{M}_2^{(3)}$	$\mathcal{M}_3^{(3)}$
$P(\lvert w_{a,p}-w_a^{(i)}\rvert/\lvert w_a^{(i)}\rvert \le 5\% \lvert \mathcal{D}_1^{(1)}, \mathcal{D}_2^{(1)}, \mathcal{M}_j^{(3)})$	0.866	0.879	0.884
$P(\lvert w_{a,p}-w_a^{(i)}\rvert/\lvert w_a^{(i)}\rvert \le 10\% \lvert \mathcal{D}_1^{(1)}, \mathcal{D}_2^{(1)}, \mathcal{M}_j^{(3)})$	0.982	0.988	0.989
$\frac{w_a^{(i)} - E[w_{a,p} \mid \mathcal{D}_1^{(1)}, \mathcal{D}_2^{(1)}, \mathcal{M}_j^{(3)}]}{\sqrt{Var[w_{a,p} \mid \mathcal{D}_1^{(1)}, \mathcal{D}_2^{(1)}, \mathcal{M}_j^{(3)}]}}$	-0.05	-0.07	-0.04

Table 5.21 Results of predicting w_a using data $\mathcal{D}_2^{(2)}$ from the validation experiment in addition to $\mathcal{D}_1^{(2)}$ from the calibration experiment

	$\mathcal{M}_1^{(3)}$	$\mathcal{M}_2^{(3)}$	$\mathcal{M}_3^{(3)}$
$P(\lvert w_{a,p}-w_a^{(i)}\rvert/\lvert w_a^{(i)}\rvert \le 5\% \lvert \mathcal{D}_1^{(2)}, \mathcal{D}_2^{(2)}, \mathcal{M}_j^{(3)})$	0.865	0.895	0.863
$P(\lvert w_{a,p}-w_a^{(i)}\rvert/\lvert w_a^{(i)}\rvert \le 10\% \lvert \mathcal{D}_1^{(2)}, \mathcal{D}_2^{(2)}, \mathcal{M}_j^{(3)})$	0.996	0.998	0.995
$\frac{w_a^{(i)} - E[w_{a,p} \mid \mathcal{D}_1^{(2)}, \mathcal{D}_2^{(2)}, \mathcal{M}_j^{(3)}]}{\sqrt{Var[w_{a,p} \mid \mathcal{D}_1^{(2)}, \mathcal{D}_2^{(2)}, \mathcal{M}_j^{(3)}]}}$	0.37	0.38	0.46

Table 5.22 Results of predicting w_a using data $\mathcal{D}_2^{(3)}$ from the validation experiment in addition to $\mathcal{D}_1^{(3)}$ from the calibration experiment

	$\mathcal{M}_1^{(3)}$	$\mathcal{M}_2^{(3)}$	$\mathcal{M}_3^{(3)}$
$P(\lvert w_{a,p}-w_a^{(i)}\rvert/\lvert w_a^{(i)}\rvert\le 5\%\lvert\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)},\mathcal{M}_j^{(3)})$	0.896,0.788 (0.842)	0.907,0.782 (0.844)	0.902,0.795 (0.848)
$P(\lvert w_{a,p}-w_a^{(i)}\rvert/\lvert w_a^{(i)}\rvert\le 10\%\lvert\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)},\mathcal{M}_j^{(3)})$	0.997,0.992 (0.994)	0.999,0.995, (0.997)	0.9995,0.994 (0.997)
$\dfrac{w_a^{(i)}-E[w_{a,p}\mid\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)},\mathcal{M}_j^{(3)}]}{\sqrt{Var[w_{a,p}\mid\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)},\mathcal{M}_j^{(3)}]}}$	0.26,-0.89	0.24,-0.96	0.26,-0.94

Table 5.23 Statistical results using data $\mathcal{D}_3^{(1)}$ from the accreditation experiment in addition to $\mathcal{D}_1^{(1)}$ and $\mathcal{D}_2^{(1)}$

		$\mathcal{M}_1^{(3)}$	$\mathcal{M}_2^{(3)}$	$\mathcal{M}_3^{(3)}$
Parameter	μ_s (Pa^{-1})	8.60×10^{-11};1.2%	8.61×10^{-11};1.2%	8.61×10^{-11};1.1%
Statistics	σ_s^2(Pa^{-2})	1.08×10^{-22};39.1%	1.08×10^{-22};44.4%	1.02×10^{-22};34.9%
	l_s(m)	0.0197;73.7%	0.0190;55.0%	0.0175;60.6%
	r			1.83;42.4%
	R	$\begin{bmatrix}1 & -0.01 & 0.06\\ & 1 & -0.18\\ & & 1\end{bmatrix}$	$\begin{bmatrix}1 & 0.06 & 0.06\\ & 1 & -0.13\\ & & 1\end{bmatrix}$	$\begin{bmatrix}1 & -0.15 & 0.01 & 0.10\\ & 1 & -0.04 & 0.03\\ & & 1 & 0.02\\ & & & 1\end{bmatrix}$
Log evidence		202.03	202.07	202.84
$P(\mathcal{M}_j^{(3)}\lvert\mathcal{D}_1^{(1)},\mathcal{D}_2^{(1)},\mathcal{D}_3^{(1)},M_3)$		0.233	0.243	0.524
$P(F\lvert\mathcal{D}_1^{(1)},\mathcal{D}_2^{(1)},\mathcal{D}_3^{(1)},\mathcal{M}_j^{(3)})$		7.39×10^{-4} (26.9 %)	1.95×10^{-4}(25.9%)	3.06×10^{-4}(27.1%)
$P(F\lvert\mathcal{D}_1^{(1)},\mathcal{D}_2^{(1)},\mathcal{D}_3^{(1)},M_3)$		3.80×10^{-4}		

Table 5.24 Statistical results using data $\mathcal{D}_3^{(2)}$ from the accreditation experiment in addition to $\mathcal{D}_1^{(2)}$ and $\mathcal{D}_2^{(2)}$

		$\mathcal{M}_1^{(3)}$	$\mathcal{M}_2^{(3)}$	$\mathcal{M}_3^{(3)}$
Parameter	μ_s (Pa^{-1})	8.71×10^{-11};0.8%	8.71×10^{-11};0.82%	8.73×10^{-11};0.8%
Statistics	σ_s^2(Pa^{-2})	6.03×10^{-23};23.9%	6.25×10^{-23};22.9%	6.01×10^{-23};22.1%
	l_s(m)	0.0364;33.1%	0.0339;24.7%	0.0358;34.7%
	r			1.92;41.9%
	R	$\begin{bmatrix} 1 & -0.16 & 0.13 \\ & 1 & -0.14 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.12 & 0.19 \\ & 1 & -0.30 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.03 & 0.11 & -0.28 \\ & 1 & -0.10 & -0.13 \\ & & 1 & -0.17 \\ & & & 1 \end{bmatrix}$
Log evidence		759.79	760.73	761.24
$P(\mathcal{M}_j^{(3)}\|\mathcal{D}_1^{(2)},\mathcal{D}_2^{(2)},\mathcal{D}_3^{(2)},M_3)$		0.128	0.327	0.545
$P(F\|\mathcal{D}_1^{(2)},\mathcal{D}_2^{(2)},\mathcal{D}_3^{(2)},\mathcal{M}_j^{(3)})$		8.34×10^{-5}(21.0%)	1.21×10^{-5}(20.4%)	1.94×10^{-4}(18.0 %)
$P(F\|\mathcal{D}_1^{(2)},\mathcal{D}_2^{(2)},\mathcal{D}_3^{(2)},M_3)$		1.20×10^{-4}		

Table 5.25 Statistical results using data $\mathcal{D}_3^{(3)}$ from the accreditation experiment in addition to $\mathcal{D}_1^{(3)}$ and $\mathcal{D}_2^{(3)}$

		$\mathcal{M}_1^{(3)}$	$\mathcal{M}_2^{(3)}$	$\mathcal{M}_3^{(3)}$
Parameter Statistics	μ_s (Pa^{-1})	8.69×10^{-11};0.57%	8.69×10^{-11};0.59%	8.69×10^{-11};0.6%
	σ_s^2(Pa^{-2})	5.88×10^{-23};18.0%	5.75×10^{-23};17.5%	5.61×10^{-23};20.0%
	l_s(m)	0.0374;25.5%	0.0378;18.9%	0.0392;26.5%
	r			1.81;40.4%
	R	$\begin{bmatrix} 1 & -0.04 & 0.06 \\ & 1 & -0.17 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.06 & 0.09 \\ & 1 & -0.19 \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.06 & 0.16 & -0.18 \\ & 1 & -0.30 & 0.21 \\ & & 1 & -0.36 \\ & & & 1 \end{bmatrix}$
Log evidence		1193.94	1193.21	1193.21
$P(\mathcal{M}_j^{(3)}\|\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)},\mathcal{D}_3^{(3)},M_3)$		0.510	0.245	0.245
$P(F\|\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)},\mathcal{D}_3^{(3)},\mathcal{M}_j^{(3)})$		8.98×10^{-6}(11.8%)	1.29×10^{-6}(16.6%)	2.68×10^{-5}(20.0%)
$P(F\|\mathcal{D}_1^{(3)},\mathcal{D}_2^{(3)},\mathcal{D}_3^{(3)},M_3)$		1.14×10^{-5}		

Table 5.26 Consistency assessment of model classes in predicting w_a using data $\mathcal{D}_3^{(l)}$ from the accreditation experiment in addition to $\mathcal{D}_1^{(l)}$ from the calibration experiment and $\mathcal{D}_2^{(l)}$ from the validation experiment

	$\mathcal{M}_1^{(3)}$	$\mathcal{M}_2^{(3)}$	$\mathcal{M}_3^{(3)}$
Data case 1, l=1	-0.0745	-0.0398	-0.0316
Data case 2, l=2	0.3507	0.3534	0.4214
Data case 3, l=3	0.2975,-0.8822	0.2771,-0.9396	0.2843, -0.9177

6 CONCLUDING REMARKS

A novel methodology based on Bayesian updating of hierarchical stochastic system model classes is proposed for uncertainty quantification, model updating, model selection, model validation and robust prediction of the response of a system for which some subsystems have been separately tested. It uses full Bayesian updating of the model classes, along with model class comparison and prediction consistency and accuracy assessment. In the proposed methodology, all the results are rigorously derived from the probability axioms and all the information in the available data are considered to make predictions. The concepts and computational tools of the proposed methodology are illustrated with a previously-studied validation challenge problem, although the methodology can handle a more general process of hierarchical subsystem testing.

As shown by the illustrative example, within a model class, there are many plausible models and the predictions of response and failure probability of the final system can often vary greatly from one model to another, showing that the consequences of the uncertainties in the parameters are significant. Ignoring the uncertainty in the modeling parameters and solely relying on the MAP model (corresponding to the maximum of the posterior PDF) or the MLE model (corresponding to the maximum likelihood parameter value) for predictions can be dangerous and misleading since such predictions can greatly underestimate the failure probability and the uncertainty in the response. It is shown how more robust predictions by a model class can be obtained by taking into account the predictions from all the plausible models in the model class where the plausibilities are quantified by their respective posterior PDF values.

Multiple model classes are investigated for the illustrative example. The response and failure probability prediction vary greatly from one model class to another. Hyper-robust predictions of response and failure probability are also obtained by a weighted average of the robust predictions given by each model class where the weight is given by the posterior probability of the model class. The posterior probability of one of the candidate model classes is so small based on the calibration data that its contribution to the prediction is negligible, so it is discarded from further predictive analysis after the calibration tests.

The computational problems resulting from full Bayesian updating of hierarchical model classes, as well as model class comparison, can be challenging, especially for problems with many uncertain parameters. A number of powerful computational tools based on stochastic simulation are used to solve efficiently the computational problems involved; in particular, for the illustrative example studied, the Hybrid Gibbs TMCMC algorithm worked well.

If a model class performs well in predicting the response for the subsystems involved in all of the experiments, one can gain more confidence in its predictive performance for the final constructed system. However, it should be stressed that 1) whether the predictive performance of the model classes is acceptable or not depends on which criteria the decision maker thinks are critical, and 2) there is no guarantee that a model class which performs well enough to satisfy the selected criteria in predicting the response of the subsystems in these experiments will always predict the response of the final system well, especially in the case where some of the uncertainties in the final system which are critical to the prediction are not present in the subsystem tests (for example, there can be uncertainties in support or joint conditions in the final system, and uncertainties in input loadings, such as stronger amplitude inputs which may

be experienced by the final system that cause it to behave very differently than the subsystems during their tests).

Although it did not occur in the illustrative example, in the case where all candidate model classes give poor performance in predicting the response for subsystems involved in an experiment, one should check whether some of the uncertainties have not been adequately modeled in the failing subsystem tests and, if so, modify the candidate model classes to properly take into account these uncertainties.

To test the performance of the proposed methodology, future work should use data collected from real systems, preferably with a larger degree of complexity than the one considered in the illustrative example of this report.

APPENDIX A

For a one-dimensional linearly elastic bar with Young's modulus $E(x)$ where $0 \le x \le L$, the elongation δL of each rod of length L and area A subject to an axial force F is given by:

$$\delta L = \frac{F}{A}\int_0^L \frac{1}{E(x)}dx \tag{A.1}$$

For the accreditation experiment, the vertical displacement w_a of the beam 1 at the midpoint Q subject to a vertical force F_a is given by:

$$w_a = (\delta y_B + \delta y_C)/2 - \frac{F_a}{I}\int_0^{L_1} \frac{\varphi^2(x)}{E(x)}dx \tag{A.2}$$

where L_1 and I are the length and the cross-sectional moment of inertia of beam 1 respectively and the hinge displacements δy_B at hinge B and δy_C at hinge C are given by:

$$\begin{bmatrix} \delta y_B \\ \delta y_C \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} T_a^{-1} \begin{bmatrix} \delta L_1 \\ \delta L_2 \\ \delta L_3 \\ \delta L_4 \end{bmatrix} \tag{A.3}$$

where T_a is given by:

$$T_a = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix} \tag{A.4}$$

The function φ(x) is given by:

$$\varphi(x) = \begin{cases} x/2, 0 \le x < L_1/2 \\ (L_1 - x)/2, L_1/2 \le x \le L_1 \end{cases} \tag{A.5}$$

The elongation δL_i of rod i of length L_i, area A_i and Young's modulus $E_i(x)$ in the frame structure due to axial force F_i is given by:

$$\delta L_i = \frac{F_i}{A_i} \int_0^{L_i} \frac{1}{E_i(x)} dx \tag{A.6}$$

The values of L_i, A_i and F_i can be obtained from Tables 5.3 and 5.4.

For the prediction, the vertical displacement w_p of the beam 4 at the midpoint P subject to a uniformly distributed vertical load q is given by:

$$w_p = (\delta y_B + \delta y_C)/2 - \frac{q}{2I_p} \int_0^{L_{p,4}} (L_{p,4} - x) x \frac{\varphi_p(x)}{E_{p,4}(x)} dx \tag{A.7}$$

where $L_{p,4}$ and I_p are the length and the cross-sectional moment of inertia of beam 4 respectively and the hinge displacements δy_B at hinge B and δy_C at hinge C are given by:

$$\begin{bmatrix} \delta y_B \\ \delta y_C \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} T_p^{-1} \begin{bmatrix} \delta L_{p,1} \\ \delta L_{p,2} \\ \delta L_{p,3} \\ \delta L_{p,4} \end{bmatrix} \tag{A.8}$$

where T_p is given by:

$$T_p = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \\ -0.7926 & 0 & -0.6097 & 0 \\ 0 & 0.5735 & 0 & -0.8192 \\ -1 & 1 & 0 & 0 \end{bmatrix} \tag{A.9}$$

The function φ(x) is given by:

$$\varphi_p(x) = \begin{cases} x/2, 0 \le x < L_{p,4}/2 \\ (L_{p,4} - x)/2, L_{p,4}/2 \le x \le L_{p,4} \end{cases} \tag{A.10}$$

The elongation $\delta L_{p,i}$ of rod i of length $L_{p,i}$, area $A_{p,i}$ and Young's modulus $E_{p,i}(x)$ in the frame structure due to axial force $F_{p,i}$ is given by:

$$\delta L_{p,i} = \frac{F_{p,i}}{A_{p,i}} \int_0^{L_i} \frac{1}{E_{p,i}(x)} dx \tag{A.11}$$

The values of $L_{p,i}$, $A_{p,i}$, and $F_{p,i}$ can be obtained from Tables 5.1 and 5.2.

APPENDIX B: HYBRID GIBBS TMCMC ALGORITHM for POSTERIOR SAMPLING

Part of our methodology involves a sequential update of the posterior PDF given the data from the experiments collected from the subsystems. The following algorithm is proposed for this purpose. At the end of the experiment where data are collected from the i-th subsystem, we need to characterize $p(\boldsymbol{\theta}|D_i,\mathcal{M}_j^{(i)})$ given the data $\mathcal{D}_i$ collected from the most current subsystem experiment and all the data $D_{i\text{-}1} =\{\mathcal{D}_1,\ldots, \mathcal{D}_{i\text{-}1}\}$ collected from the previous subsystem experiments, where $D_i = D_{i\text{-}1} \cup \mathcal{D}_i$. The prior PDF corresponding to this posterior PDF is $p(\boldsymbol{\theta}|D_{i\text{-}1},\mathcal{M}_j^{(i)})$ from which samples have been previously generated and the evidences $p(D_{i\text{-}1}|\mathcal{M}_j^{(i)})$ for each model class $\mathcal{M}_j^{(i)}$ which have been obtained. Note that in the analysis below, we use the conventions $p(\boldsymbol{\theta}|D_0,\mathcal{M}_j^{(i)}) = p(\boldsymbol{\theta}|\mathcal{M}_j^{(i)})$ and $p(D_0|\mathcal{M}_j^{(i)})=1$.

For a given $\boldsymbol{\theta}$, $\mathcal{D}_1,\ldots, \mathcal{D}_i$ are modeled as stochastically independent. We propose a hybrid approach making use of the TMCMC method (Ching & Chen 2007), Metropolis Hastings algorithm and Gibbs sampling to generate samples from the posterior PDF $\pi(\boldsymbol{\theta})=p(\boldsymbol{\theta}|D_i,\mathcal{M}_j^{(i)})= p(\mathcal{D}_i|\boldsymbol{\theta},\mathcal{M}_j^{(i)})p(\boldsymbol{\theta}|D_{i\text{-}1},\mathcal{M}_j^{(i)})/p(\mathcal{D}_i|D_{i\text{-}1},\mathcal{M}_j^{(i)})$ and to calculate the evidence $p(\mathcal{D}_i|D_{i\text{-}1},\mathcal{M}_j^{(i)})$.

Consider a sequence of intermediate PDFs $\pi_l(\boldsymbol{\theta})$ for l=0,1,…, L, such that the first and last PDFs, $\pi_0(\boldsymbol{\theta})$ and $\pi_L(\boldsymbol{\theta}) = \pi(\boldsymbol{\theta})$, in the sequence are the prior $p(\boldsymbol{\theta}|D_{i\text{-}1},\mathcal{M}_j^{(i)})$ and posterior $p(\boldsymbol{\theta}|D_i,\mathcal{M}_j^{(i)})$, respectively:

$$\pi_l(\boldsymbol{\theta}) \propto p^{\tau_l}(\mathcal{D}_i \mid \boldsymbol{\theta}, \mathcal{M}_j^{(i)})\, p(\boldsymbol{\theta} \mid D_{i-1}, \mathcal{M}_j^{(i)}) \tag{B.1}$$

where $0=\tau_0<\tau_1<\ldots<\tau_L=1$. Divide $\boldsymbol{\theta}$ into B groups of components. Denote the b-th component group of $\boldsymbol{\theta}$ as $\boldsymbol{\theta}_b$.

First, N_0 samples are generated from the prior $p(\boldsymbol{\theta}|D_{i\text{-}1},\mathcal{M}_j^{(i)})$. Then do the following procedures for $l=1,\ldots,L$. At the beginning of the l-th level, we have the samples $\boldsymbol{\theta}_{l-1}^{(m)}$, $m=1,2,\ldots,N_{l\text{-}1}$, from $\pi_{l\text{-}1}(\boldsymbol{\theta})$. First, select τ_l such that the effective sample size $1/\sum_{s=1}^{N_{l-1}} \tilde{w}_s^{\,2}$ = some threshold (e.g., 0.9 $N_{l\text{-}1}$) (Cheung & Beck 2008b), where $\tilde{w}_s = w_s / \sum_{s=1}^{N_{l-1}} w_s$ and $w_s = p^{\tau_l-\tau_{l-1}}(\mathcal{D}_i \mid \boldsymbol{\theta}_{l-1}^{(s)}, \mathcal{M}_j)$, $s=1,2,\ldots,N_{l\text{-}1}$. If $\tau_l>1$, then set $L=l$ and $\tau_l=1$, then recompute w_s and $\tilde{w}_s$. Compute an estimate for the sample covariance matrix for $\pi_l(\boldsymbol{\theta})$ as follows:

$$\Sigma = \sum_{m=1}^{N_{l-1}} \tilde{w}_m (\boldsymbol{\theta}_{l-1}^{(m)} - \bar{\boldsymbol{\theta}})(\boldsymbol{\theta}_{l-1}^{(m)} - \bar{\boldsymbol{\theta}})^T, \quad \bar{\boldsymbol{\theta}} = \sum_{m=1}^{N_{l-1}} \tilde{w}_m \boldsymbol{\theta}_{l-1}^{(m)} \tag{B.2}$$

Set $E_l = \sum_{s=1}^{N_{l-1}} w_s / N_{l-1}$. Then the N_l samples $\boldsymbol{\theta}_l^{(n)}$ from $\pi_l(\boldsymbol{\theta})$ are generated by doing the following for $n=1,2,\ldots,N_l$:

1. Draw a number s' from a discrete distribution $p(S=s)=\tilde{w}_s$, $s=1,2,\ldots,N_{l\text{-}1}$.
2. Fixing the last component group of $\boldsymbol{\theta}$ at the values of $\boldsymbol{\theta}_{l-1,B}^{(s')}$, draw the samples $\boldsymbol{\theta}_{l,1}^{(n)}$, …, $\boldsymbol{\theta}_{l,B-1}^{(n)}$ for the first B-1 component groups of $\boldsymbol{\theta}$, one after another, using Gibbs sampling as described later. Set $\boldsymbol{\theta}_{l-1,b}^{(s')} = \boldsymbol{\theta}_{l,b}^{(n)}$ for $b=1,\ldots,B\text{-}1$.

3. Fixing the first B-1 component groups at the values of $\boldsymbol{\theta}_{l,1}^{(n)}$, …, $\boldsymbol{\theta}_{l,B-1}^{(n)}$, generate a sample $\boldsymbol{\theta}_{l,B}^{(n)}$ for the last component group of $\boldsymbol{\theta}$ by the Metropolis-Hastings algorithm: Generate $\boldsymbol{\theta}^*$ from a Gaussian PDF with mean $\boldsymbol{\theta}_{l-1,B}^{(s')}$ and covariance matrix $\eta\Sigma_B$ where Σ_B is the submatrix that corresponds to the last component group (i.e., the B-th component group) in the covariance matrix Σ. Compute the acceptance probability $r''=\min\{r',1\}$ where r' is given by:

$$r' = \frac{p^{\tau_i}(\mathcal{D}_i \mid \boldsymbol{\theta}_{l,1}^{(n)},...,\boldsymbol{\theta}_{l,B-1}^{(n)},\boldsymbol{\theta}^*,\mathcal{M}_j^{(i)})}{p^{\tau_i}(\mathcal{D}_i \mid \boldsymbol{\theta}_{l,1}^{(n)},...,\boldsymbol{\theta}_{l,B-1}^{(n)},\boldsymbol{\theta}_{l-1,B}^{(s')},\mathcal{M}_j^{(i)})} \times \frac{[\prod_{t=1}^{i-1} p(\mathcal{D}_t \mid \boldsymbol{\theta}_{l,1}^{(n)},...,\boldsymbol{\theta}_{l,B-1}^{(n)},\boldsymbol{\theta}^*,\mathcal{M}_j^{(i)})]p(\boldsymbol{\theta}_{l,1}^{(n)},...,\boldsymbol{\theta}_{l,B-1}^{(n)},\boldsymbol{\theta}^* \mid \mathcal{M}_j^{(i)})}{[\prod_{t=1}^{i-1} p(\mathcal{D}_t \mid \boldsymbol{\theta}_{l,1}^{(n)},...,\boldsymbol{\theta}_{l,B-1}^{(n)},\boldsymbol{\theta}_{l-1,B}^{(s')},\mathcal{M}_j^{(i)})]p(\boldsymbol{\theta}_{l,1}^{(n)},...,\boldsymbol{\theta}_{l,B-1}^{(n)},\boldsymbol{\theta}_{l-1,B}^{(s')} \mid \mathcal{M}_j^{(i)})} \quad \text{(B.3)}$$

If $r''>U(0,1)$ where $U(0,1)$ is a uniformly distributed number between 0 and 1, $\boldsymbol{\theta}_{l,B}^{(n)}=\boldsymbol{\theta}^*$, $\boldsymbol{\theta}_{l-1,B}^{(s')}=\boldsymbol{\theta}^*$. Otherwise, $\boldsymbol{\theta}_{l,B}^{(n)}=\boldsymbol{\theta}_{l-1,B}^{(s')}$.

Thus, the n-th sample for $\boldsymbol{\theta}$ with the target PDF $\pi_l(\boldsymbol{\theta})$ is given by $\boldsymbol{\theta}_l^{(n)}=[\boldsymbol{\theta}_{l,1}^{(n)}\ \boldsymbol{\theta}_{l,2}^{(n)}\ ....\boldsymbol{\theta}_{l,B}^{(n)}]$.

In step 3, η (e.g., 0.2^2) is chosen such that the average acceptance probability is larger than some threshold (e.g., 0.7). Other MCMC algorithms such as Hybrid Monte Carlo methods (Cheung and Beck 2007, 2008a) can also be used in place of the Metropolis-Hastings algorithm in step 3 for more effective sampling, as is done in Cheung and Beck (2008a, c and d). The evidence $p(\mathcal{D}_i|D_{i-1},\mathcal{M}_j^{(i)})$ for $\mathcal{M}_j^{(i)}$ given by data $\mathcal{D}_i$ can be estimated as follows:

$$p(\mathcal{D}_i \mid D_{i-1}, \mathcal{M}_j^{(i)}) \approx \prod_{l=1}^{L} E_l \tag{B.4}$$

Gibbs sampling for the posterior PDF in the illustrative example with data D_1 (i=1)

Now we describe how Gibbs sampling can be performed for the posterior PDF in the illustrative example with data D_1 (i=1). For $\mathcal{M}_1^{(1)}$ (i=1, j=1), $\boldsymbol{\theta}$ is divided into 2 component groups: $\boldsymbol{\theta}_1 = \mu_s$, $\boldsymbol{\theta}_2 = [\sigma_s^2 \ \sigma_\varepsilon^2]$. Gibbs sampling in step 2 of the above algorithm is performed on the first component group as follows: draw $\boldsymbol{\theta}_{l,1}^{(n)}$ from a truncated Gaussian PDF (constrained to be positive) which is proportional to a Gaussian distribution with mean μ and variance σ^2 given below:

$$\mu = \frac{H_{11}\frac{F_c L_c}{A_c}\sum_{i=1}^{N_c}\delta L_c^{(i)} + H_{12}\left(\frac{F_c L_c}{A_c}\sum_{k=1}^{N_c} S_c^{(k)}(L_c/2) + \sum_{k=1}^{N_c}\delta L_c^{(k)}\right) + H_{22}\sum_{k=1}^{N_c} S_c^{(k)}(L_c/2) + \frac{\mu_0}{\tau_l {\sigma_0}^2}}{N_c\left(H_{11}\left(\frac{F_c L_c}{A_c}\right)^2 + 2H_{12}\frac{F_c L_c}{A_c} + H_{22}\right) + \frac{1}{\tau_l {\sigma_0}^2}} \tag{B.5}$$

$$\sigma^2 = \frac{1}{\tau_l\left[N_c\left(H_{11}\left(\frac{F_c L_c}{A_c}\right)^2 + 2H_{12}\frac{F_c L_c}{A_c} + H_{22}\right) + \frac{1}{\tau_l {\sigma_0}^2}\right]} \tag{B.6}$$

where H_{11}, H_{12} and H_{22} are the (1,1), (1,2) and (2,2) entries of the inverse of $\mathbf{C}(\sigma_s^2, \sigma_\varepsilon^2)$ in equation (5.4) with $[\sigma_s^2 \ \sigma_\varepsilon^2] = \boldsymbol{\theta}_{l-1,2}^{(s')}$; μ_0 and σ_0^2 are the mean and variance of the prior PDF $p(\mu_s|\mathcal{M}_j^{(1)})$ of μ_s respectively

For $\mathcal{M}_4^{(1)}$ (i=1, j=4), $\boldsymbol{\theta}$ is divided into 3 component groups: $\boldsymbol{\theta}_1 = \mu_s$, $\boldsymbol{\theta}_2 = \sigma_s^2$, $\boldsymbol{\theta}_3 = [l_s^2 \ r]$. Gibbs sampling in step 2 of the proposed algorithm is performed on the first two component groups as

follows: draw $\boldsymbol{\theta}_{l,1}^{(n)}$ from a truncated Gaussian PDF (constrained to be positive) which is proportional to a Gaussian distribution with mean μ' and variance σ'^2 given below:

$$\mu' = \frac{H_{11}\frac{F_c L_c}{A_c}\sum_{k=1}^{N_c}\delta L_c^{(k)} + H_{12}\left(\frac{F_c L_c}{A_c}\sum_{k=1}^{N_c}S_c^{(k)}(L_c/2) + \sum_{k=1}^{N_c}\delta L_c^{(k)}\right) + H_{22}\sum_{k=1}^{N_c}S_c^{(k)}(L_c/2) + \frac{\sigma_s^{\,2}\mu_0}{\tau_l\sigma_0^{\,2}}}{N_c\left(H_{11}\left(\frac{F_c L_c}{A_c}\right)^2 + 2H_{12}\frac{F_c L_c}{A_c} + H_{22}\right) + \frac{\sigma_s^{\,2}}{\tau_l\sigma_0^{\,2}}} \tag{B.7}$$

$$\sigma'^2 = \frac{\sigma_s^{\,2}}{\tau_l\left[N_c\left(H_{11}\left(\frac{F_c L_c}{A_c}\right)^2 + 2H_{12}\frac{F_c L_c}{A_c} + H_{22}\right) + \frac{\sigma_s^{\,2}}{\tau_l\sigma_0^{\,2}}\right]} \tag{B.8}$$

In the above equations, $\sigma_s^2 = \boldsymbol{\theta}_{l-1,2}^{(s')}$ and H_{11}, H_{12} and H_{22} are the (1,1), (1,2) and (2,2) entries of the inverse of $\mathbf{C}(l_s, r)$ in equation (5.6) with $[l_s\ r] = \boldsymbol{\theta}_{l-1,3}^{(s')}$. Then draw $\boldsymbol{\theta}_{l,2}^{(n)}$ from an inverse gamma distribution with PDF proportional to $(\theta_2')^{-\alpha'-1}\exp(-\beta'/\theta_2')$ where $\alpha'=\alpha+\tau_l N_c$ and β' is given by:

$$\beta' = \beta + \frac{\tau_l}{2}\sum_{k=1}^{N_c}[\mathbf{y}^{(k)} - \boldsymbol{\mu}(\mu_s)]^T\mathbf{C}^{-1}(l_s, r)[\mathbf{y}^{(k)} - \boldsymbol{\mu}(\mu_s)] \tag{B.9}$$

where α and β are the parameters for the prior PDF $p(\sigma_s^2|\mathcal{M}_j^{(1)})$ of σ_s^2, the terms in the above are given by (5.2), (5.3) and (5.6) with $\mu_s = \boldsymbol{\theta}_{l,1}^{(n)}$, $[l_s\ r] = \boldsymbol{\theta}_{l-1,3}^{(s')}$. For $\mathcal{M}_2^{(1)}(i=1, j=2)$ and $\mathcal{M}_3^{(1)}(i=1, j=3)$, everything is the same as for $\mathcal{M}_4^{(1)}(i=1, j=4)$ except that r is fixed at 1 and 2 respectively.

Gibbs sampling for the posterior PDF in the illustrative example with data D_2 (i=2)

Now we describe how Gibbs sampling can be performed for the posterior PDF in the illustrative example with data $D_2=\{\mathcal{D}_1, \mathcal{D}_2\}$ (i=2), for $\mathcal{M}_3^{(2)}$ (i=2, j=3), $\boldsymbol{\theta}$ is divided into 3 component groups: $\boldsymbol{\theta}_1= \mu_s$, $\boldsymbol{\theta}_2=\sigma_s^2$, $\boldsymbol{\theta}_3=[l_s^2\ r]$. Gibbs sampling in step 2 of the proposed stochastic simulation algorithm is performed on the first two component groups as follows: draw $\boldsymbol{\theta}_{l,1}^{(n)}$ from a truncated Gaussian PDF (constrained to be positive) which is proportional to a Gaussian distribution with mean μ'' and variance σ''^2 given below:

$$\mu'' = \sigma''^2\left(\frac{\mu'}{\sigma'^2} + \frac{\tau_l K_v \sum_{k=1}^{N_v} \delta L_v^{(k)}}{\sigma_{v,j}(\sigma_s^2, l_s, r)^2}\right) \tag{B.10}$$

$$\sigma''^2 = \frac{1}{\frac{1}{\sigma'^2} + \frac{N_v K_v^2 \tau_l}{\sigma_{v,j}(\sigma_s^2, l_s, r)^2}} \tag{B.11}$$

$$\mu' = \frac{H_{11}\frac{F_c L_c}{A_c}\sum_{k=1}^{N_c}\delta L_c^{(k)} + H_{12}\left(\frac{F_c L_c}{A_c}\sum_{k=1}^{N_c} S_c^{(k)}(L_c/2) - \sum_{k=1}^{N_c}\delta L_c^{(k)}\right) + H_{22}\sum_{k=1}^{N_c} S_c^{(k)}(L_c/2) + \frac{\sigma_s^2\mu_0}{\sigma_0^2}}{N_c\left(H_{11}\left(\frac{F_c L_c}{A_c}\right)^2 + 2H_{12}\frac{F_c L_c}{A_c} + H_{22}\right) + \frac{\sigma_s^2}{\sigma_0^2}} \tag{B.12}$$

$$\sigma'^2 = \frac{\sigma_s^2}{N_c\left(H_{11}\left(\frac{F_c L_c}{A_c}\right)^2 + 2H_{12}\frac{F_c L_c}{A_c} + H_{22}\right) + \frac{\sigma_s^2}{\sigma_0^2}} \tag{B.13}$$

In the above equations, $\sigma_s^2 = \boldsymbol{\theta}_{l-1,2}^{(s')}$, $[l_s\ r] = \boldsymbol{\theta}_{l-1,3}^{(s')}$; H_{11}, H_{12} and H_{22} are the (1,1), (1,2) and (2,2) entries of the inverse of $\mathbf{C}(l_s, r)$ in (5.6); K_v is given in section 5.2; $\sigma_{v,j}(\sigma_s^2, l_s, r)^2 = \sigma_s^2 s_{v,j}(l_s, r)$

where $s_{v,j}(l_s,r)$ is given in section 5.2. Then draw $\boldsymbol{\theta}_{l,2}^{(n)}$ from an inverse gamma distribution with PDF proportional to $(\theta_2'')^{-\alpha''-1}\exp(-\beta''/\theta_2'')$ where $\alpha''=\alpha+N_c+\tau_l N_v/2$ and β'' is given by:

$$\beta'' = \beta + \frac{1}{2}\sum_{k=1}^{N_c}[\mathbf{y}^{(k)} - \boldsymbol{\mu}(\mu_s)]^T \mathbf{C}^{-1}(l_s,r)[\mathbf{y}^{(k)} - \boldsymbol{\mu}(\mu_s)] + \frac{\tau_l}{2s_v(l_s,r)}\sum_{k=1}^{N_v}(\delta L_v^{(k)} - K_v\mu_s)^2 \quad \text{(B.14)}$$

where α and β are the parameters for the PDF $p(\sigma_s^2|\mathcal{M}_{j+1}^{(1)})$ of σ_s^2, the terms in the above are given by (5.2), (5.3) and (5.6) with $\mu_s = \boldsymbol{\theta}_{l,1}^{(n)}$, $[l_s\ r] = \boldsymbol{\theta}_{l-1,3}^{(s')}$. For $\mathcal{M}_1^{(2)}(i=2, j=1)$ and $\mathcal{M}_2^{(2)}$ $(i=2, j=2)$, everything is the same as for $\mathcal{M}_3^{(2)}(i=2, j=3)$ except that r is fixed at 1 and 2 respectively.

Gibbs sampling for the posterior PDF in the illustrative example with data D_3 (i=3)

Now we describe how Gibbs sampling can be performed for the posterior PDF in the illustrative example with data $D_3=\{\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3\}$ (i=3), for $\mathcal{M}_3^{(3)}$ (i=3, j=3), $\boldsymbol{\theta}$ is divided into 3 component groups: $\boldsymbol{\theta}_1=\mu_s$, $\boldsymbol{\theta}_2=\sigma_s^2$, $\boldsymbol{\theta}_3=[l_s^2\ r]$. Gibbs sampling in step 2 of the proposed stochastic simulation algorithm is performed on the first two component groups as follows: draw $\boldsymbol{\theta}_{l,1}^{(n)}$ from a truncated Gaussian PDF (constrained to be positive) which is proportional to a Gaussian distribution with mean μ''' and variance σ'''^2 given below:

$$\mu''' = \sigma'''^2\left(\frac{\mu'}{\sigma'^2} + \frac{K_v\sum_{k=1}^{N_v}\delta L_v^{(k)}}{\sigma_{v,j}(\sigma_s^2,l_s,r)^2} + \frac{\tau_l K_a\sum_{k=1}^{N_a}w_a^{(k)}}{\sigma_{a,j}(\sigma_s^2,l_s,r)^2}\right) \quad \text{(B.15)}$$

$$\sigma'''^2 = \frac{1}{\frac{1}{\sigma'^2} + \frac{N_v K_v^2}{\sigma_{v,j}(\sigma_s^2,l_s,r)^2} + \frac{N_a K_a^2 \tau_l}{\sigma_{a,j}(\sigma_s^2,l_s,r)^2}} \quad \text{(B.16)}$$

In the above equations, $\sigma_s^2 = \boldsymbol{\theta}_{l-1,2}^{(s')}$, $[l_s \;\; r] = \boldsymbol{\theta}_{l-1,3}^{(s')}$; $\sigma_{a,j}(\sigma_s^2, l_s, r)^2 = \sigma_s^2 \; s_{a,\,j}(l_s, r)$ where $s_{a,j}(l_s,r)$ is given in Appendix C. Then draw $\boldsymbol{\theta}_{l,2}^{(n)}$ from an inverse gamma distribution with PDF proportional to $(\theta_2''')^{-\alpha'''-1}\exp(-\beta'''/\theta_2''')$ where $\alpha'''=\alpha+N_c+N_v/2+\tau_l N_a/2$ and β''' is given by:

$$\beta''' = \beta + \frac{1}{2}\sum_{k=1}^{N_c}[\mathbf{y}^{(k)} - \boldsymbol{\mu}(\mu_s)]^T \mathbf{C}^{-1}(l_s, r)[\mathbf{y}^{(k)} - \boldsymbol{\mu}(\mu_s)] + \frac{\sum_{k=1}^{N_v}(\delta L_v^{(k)} - K_v\mu_s)^2}{2s_v(l_s, r)} + \frac{\tau_l \sum_{k=1}^{N_a}(w_a^{(k)} - K_a\mu_s)^2}{2s_a(l_s, r)} \tag{B.17}$$

where $\mu_s = \boldsymbol{\theta}_{l,1}^{(n)}$, $[l_s \; r] = \boldsymbol{\theta}_{l-1,3}^{(s')}$. For $\mathcal{M}_1^{(3)}(i=3, j=1)$ and $\mathcal{M}_2^{(3)}(i=3, j=2)$, everything is the same as for $\mathcal{M}_3^{(3)}(i=3, j=3)$ except that r is fixed at 1 and 2 respectively.

Gibbs sampling in step 3 of the hybrid Gibbs TMCMC algorithm exploits the form of $p(\boldsymbol{\theta}|D_i, \mathcal{M}_j^{(i)})$ which allows direct sampling from the conditional PDF for some groups. In the case where the form of $p(\boldsymbol{\theta}|D_i, \mathcal{M}_j^{(i)})$ cannot be exploited to carry out Gibbs sampling, step 2 is skipped and $\boldsymbol{\theta}$ has only one component group which includes all the parameters and so the algorithm reduces to the original TMCMC algorithm.

APPENDIX C

$$s_{a,j}(l_s,r)=\frac{1}{4}s_{1,j}+\frac{1}{2}(s_{2,j}+s_{4,j})+(\frac{F_a}{I})^2 Q_{a,j}{}^{(1)}(l_s,r)-\frac{F_aF_1}{A_1I}Q_{a,j}{}^{(2)}(l_s,r) \tag{C.1}$$

where

$$s_{i,j}(l_s,r)=2(\frac{F_i}{A_i})^2\int_0^{L_i}(L_i-x)\exp(-(\frac{x}{l_s})^r)dx \tag{C.2}$$

$$Q_{a,j}{}^{(1)}=\int_0^{L_1}\int_0^{L_1}\varphi^2(x_1)\varphi^2(x_2)\exp(-(\frac{|x_1-x_2|}{l_s})^r)dx_1dx_2 \tag{C.3}$$

$$Q_{a,j}{}^{(2)}=\int_0^{L_1}\int_0^{L_1}\varphi^2(x_2)\exp(-(\frac{|x_1-x_2|}{l_s})^r)dx_1dx_2 \tag{C.4}$$

where r=1 for j=2, r=2 for j=3.

$$s_{i,2}(l_s,1)=2(\frac{F_i}{A_i})^2 l_s(L_i-l_s+l_s\exp(-\frac{L_i}{l_s})) \tag{C.5}$$

$$s_{i,3}(l_s,2)=(\frac{F_i}{A_i})^2[L_il_s\sqrt{\pi}\,\mathrm{erf}(\frac{L_i}{l_s})-l_s{}^2(1-\exp(-(\frac{L_i}{l_s})^2))] \tag{C.6}$$

REFERENCES

Babuška, I. and Oden, J.T. 2004. Verification and validation in computational engineering and science: basic concepts. Computer Methods in Applied Mechanics and Engineering. 193(36-38): 4057-4066.

Babuška, I., Nobile, F. and Tempone, R. 2006. Reliability of computational science. Numerical methods for partial differential equations. 23(4):753-784

Babuška, I., Nobile, F. and Tempone, R. 2008. Formulation of static frame problem. Computer Methods in Applied Mechanics and Engineering. 197(29-32): 2496-2499

Babuška, I., Nobile, F. and Tempone, R. 2008. A systematic approach to model validation based on Bayesian updates and prediction related rejection criteria. Computer Methods in Applied Mechanics and Engineering. 197(29-32): 2517-2539

Beck, J.L. and Katafygiotis, L.S. 1991. Updating of a model and its uncertainties utilizing dynamic test data. *Proceedings First International Conference on Computational Stochastic Mechanics*, 125-136, Computational Mechanics Publications, Boston, September 1991.

Beck, J.L. and Katafygiotis, L.S. 1998. Updating models and their uncertainties. I: Bayesian statistical framework. ASCE Journal of Engineering Mechanics. 124(4):455-61.

Beck, J.L. and Au, S.K. 2002. Bayesian updating of structural models and reliability using Markov Chain Monte Carlo simulation. ASCE Journal of Engineering Mechanics 128(2): 380-391.

Beck, J.L. and Yuen, K.V. 2004. Model selection using response measurements: A Bayesian probabilistic approach. ASCE Journal of Engineering Mechanics 130(2):192-203.

Berger, J. and Delampady, M. 1987. Testing precise hypotheses. Statistical Science. 3:317-352.

Cheung, S.H. and Beck, J.L. 2007. Bayesian model updating of higher-dimensional dynamic systems. 10th International Conference on Applications of Statistics and Probability in Civil Engineering, University of Tokyo, Tokyo, Japan, July 31-August 3, 2007.

Cheung, S. H. and Beck, J.L. 2008a. Bayesian model updating using Hybrid Monte Carlo simulation with application to structural dynamic models with many uncertain parameters, Journal of Engineering Mechanics, in print.

Cheung, S.H. and Beck, J.L. 2008b. Near real-time loss estimation of structures subjected to strong seismic excitation. Inaugural International Conference of the Engineering Mechanics Institute (EM08), University of Minnesota, Minneapolis, Minnesota, USA, May 18-21, 2008.

Cheung, S.H. and Beck, J.L. 2008c. On Using Posterior Samples for Model Selection for Structural Identification" Asian-Pacific Symposium on Structural Reliability and its Applications 2008 (APSSRA'08), Hong Kong University of Science and Technology, Hong Kong, China, June 18-20, 2008.

Cheung, S.H. and Beck, J.L. 2008d. Calculation of the posterior probability for Bayesian model class selection and averaging from posterior samples based on dynamic system data. Computer-Aided Civil and Infrastructure Engineering, under review.

Chleboun, J. An approach to the Sandia workshop static frame challenge problem: A combination of elementary probabilistic, fuzzy set, and worst scenario tools. Computer Methods in Applied Mechanics and Engineering. 197(29-32): 2500-2516

Ching, J., Muto, M. and Beck, J.L. 2005, Bayesian linear structural model updating using Gibbs sampler with modal data, *Proc. International Conference on Structural Safety and Reliability*, Rome, Italy, June 2005.

Ching, J., Muto, M. and Beck, J.L. 2006. Structural model updating and health monitoring with incomplete modal data using Gibbs Sampler. Computer-Aided Civil and Infrastructure Engineering 21: 242-257.

Ching, J. and Chen, Y.J. 2007. Transitional Markov Chain Monte Carlo method for Bayesian model updating, model class selection and model averaging. Journal of Engineering Mechanics: 133(7): 816-832.

Cover, T.M. and Thomas, J.A. 2006. Elements of Information Theory. Wiley Series in Telecommunications, John Wiley & Sons, Inc.

Cox, R.T. 1961. *The Algebra of Probable Inference,* Johns Hopkins University Press, Baltimore, MD.

Dowding, K.J., Pilch, M. and Hills, R.G. 2008. Formulation of the thermal problem. Computer Methods in Applied Mechanics and Engineering. 197(29-32): 2385-2389

Grigoriu, M.D. and Field Jr., R.V. A solution to the static frame validation challenge problem using Bayesian model selection. Computer Methods in Applied Mechanics and Engineering. 197(29-32): 2540-2549

Hills, R.G., Pilch, M., Dowding, K.J., Red-Horse, J., Paez, T.L., Babuška, I., and Tempone, R. 2008. Validation challenge workshop. Computer Methods in Applied Mechanics and Engineering. 197(29-32): 2375-2380

Hoeting, J.A., Madigan, D., Raftery, A.E., and Volinsky, C.T. 1999. Bayesian model averaging: a tutorial (with discussion). Statistical Science. 14(4):382-417

Jaynes, E.T. 2003. Probability Theory: The Logic of Science. Cambridge University Press.

Katafygiotis, LS. and Beck, JL. 1998. Updating models and their uncertainties: model identifiability. ASCE Journal of Engineering Mechanics 124(4): 463-467.

Katafygiotis, L.S. and Lam, H.F. 2002. Tangential-projection algorithm for manifold representation in unidentifiable model updating problems. Earthquake Engineering and Structural Dynamics 31: 791-812.

Mackay, D.J.C. 1992. Bayesian methods for adaptive methods. PhD Thesis in Computation and Neural Systems, California Institute of Technology.

Muto, M. and Beck, J.L. 2008. Bayesian updating of hysteretic structural models using stochastic simulation. Journal of Vibration and Control. 14:7-34.

Oberkampf, W.L., Helton, J.C., Joslyn, C.A., Wojtkiewicz, S.F. and Ferson, S. 2004. Challenge problems: uncertainty in system response given uncertain parameters. Reliability Engineering and System Safety. 85(1-3): 11-19

Pradlwarter, H.J. and Schuëller, G.I. The use of kernel densities and confidence intervals to cope with insufficient data in validation experiments. Computer Methods in Applied Mechanics and Engineering. 197(29-32): 2550-2560

Raftery A.E., Madigan D., and Hoeting J.A. 1997. Bayesian model averaging for linear regression models. Journal of the American Statistical Association. 92:179-191.

Rebba, R. and Cafeo, J. Probabilistic analysis of a static frame model. Computer Methods in Applied Mechanics and Engineering. 197(29-32): 2561-2571.

Red-Horse, J.R. and Paez, T.L. 2008. Sandia National Laboratories Validation Workshop: Structural dynamics application. Computer Methods in Applied Mechanics and Engineering. 197(29-32):2578-2584

Yuen, K.V., Beck, J.L. and Au, S.K. (2004). Structural damage detection and assessment by adaptive Markov chain Monte Carlo simulation. Structural Control and Health Monitoring.11: 327-347.