Probabilistic Identification of Simulated Damage on the Dowling Hall Footbridge through Bayesian FE Model Updating

Iman Behmanesh and Babak Moaveni

Dept. of Civil and Environmental Eng., Tufts University, Medford, MA, USA

ABSTRACT

This paper presents a probabilistic damage identification study on a full-scale structure, the Dowling Hall Footbridge, through a Bayesian finite element (FE) model updating. The footbridge is located at Tufts University and is equipped with a continuous monitoring system that measures its ambient acceleration response. A set of data is recorded once every hour or when triggered by large vibrations. The modal parameters of the footbridge are extracted from each set of data and are used for FE model updating. In this study, effects of physical damage are simulated by loading a small segment of the footbridge deck with concrete blocks. The footbridge deck is divided into five segments in a FE model of the test structure and the added mass on each segment is considered as an updating parameter. Overall, 72 sets of data are collected during the loading period and different subsets of these data are used to find the location and extent of the damage (added mass). The Adaptive Metropolis-Hastings algorithm with adaption on the proposal probability density function is successfully used to generate Markov Chains for sampling the posterior probability distributions of the five updating parameters. Effects of the number of data sets used in the identification process are investigated on the posterior probability distributions of the updating parameters. The probabilistic model updating framework accurately predicts the simulated damage and the level of confidence on the obtained results. The probabilistic damage identification results are found to be in good agreement with their corresponding deterministic counterparts.

KEYWORDS: Bayesian FE model updating, Adaptive Metropolis-Hastings algorithm, damage identification, uncertainty analysis, Dowling Hall Footbridge

1. INTRODUCTION

Recent structural failures have raised public attention on the need for improved infrastructure safety and maintenance [1-2]. Bridge and infrastructure owners are encouraged to evaluate and assess the true state of structural health regularly. This is currently done mostly by visual inspection. Structural health monitoring (SHM) can provide more revealing and quantitative information on structural health and is complementary to visual inspection. The ultimate goal of SHM is to determine the four levels of damage identification proposed by Rytter [3] at the earliest possible stage: (1) existence of damage; (2) location of damage; (3) severity of damage; and (4) remaining useful life of structures. Many methods have been proposed in the past two decades to address some or all the damage identification levels [4-6]. A commonly used class of
these methods is the finite element (FE) model updating methods [7-11], which have been successfully applied for damage identification of civil structures in recent years. In the FE model updating methods, model parameters are adjusted so that model predicted time histories, modal parameters, or frequency response functions best match the corresponding quantities obtained from the test data. However, the accuracy and spatial resolution of deterministic damage identification results depend on the accuracy and completeness of the measured data or the features extracted from data (e.g., modal parameters) and on the accuracy of the FE model. The extracted data features are usually estimated with significant uncertainties, mainly due to measurement noise, estimation errors, and changing environmental conditions. These uncertainties propagate through the FE model updating process yielding uncertain model updating results [12-15]. Another unavoidable source of uncertainty in the model updating process arises from modeling errors, which are due to the simplifications and idealizations involved in the modeling. In order to better quantify the accuracy of model updating results, it is necessary to account for these sources of uncertainties through a probabilistic model updating framework, where the prior probabilities of model parameters are updated to their posterior probabilities using the measured data.

The papers by Sohn and Law [16] and Beck and Katafygiotis [9] are the pioneering efforts in probabilistic FE model updating using the Bayesian inference scheme. One important advantage of Bayesian model updating methods to deterministic methods is their application in locally identifiable and unidentifiable cases of inverse problems [17]. Another important advantage of these methods is to estimate the level of confidence on the model updating results. In continuous SHM, continual accumulation of data increases the level of confidence/accuracy in model updating and therefore damage identification results by mitigating the effects of measurement noises and estimation errors as the amount of measured data increases [18]. One of the challenges in the application of Bayesian model updating methods is the case of multi-dimensional complex systems with several updating parameters. These cases involve solving multi-dimensional integrations to calculate marginal posterior probability distributions of model parameters or structural response quantities of interest. One common approach to tackle this problem is generating Markov Chain samples from the posterior joint probability distribution of the model parameters. This numerical technique has been successfully applied on several numerical and small-scale laboratory applications [19-22].

Although the deterministic FE model updating methods have been used for damage identification of several real-world, large-scale structures [23-30] including the Dowling Hall footbridge [31], there are few applications of Bayesian FE model updating methods to full-scale complex structures [32-33]. Specifically, application of Bayesian FE model updating to large-scale operational civil structures is very rare due to the existing challenges when dealing with real data and large modeling errors. Most of the probabilistic studies are applied to numerically simulated data, and in the absence of large modeling errors that can significantly affect the model updating results. In this paper, the authors investigate the challenges of implementing a Bayesian FE model updating framework for damage identification of a full-scale structure, the Dowling Hall footbridge. The available vibration data are collected during the operational condition of the structure and therefore contain realistic sources of uncertainty/variability. The main sources of uncertainties include the variability in the structural mass, stiffness and boundary conditions due to changing pedestrian traffic and environmental conditions.
Damage on the footbridge is simulated by addition of 2.29 metric tons of concrete blocks on a small segment of the bridge deck. The effect of added mass will be similar to a loss of stiffness (commonly used as damage indication) at the loaded segment of the bridge. The extracted modal parameters from ambient acceleration time histories of the damaged structure (loaded structure) are used to find the location and extent of the damage (added mass) through a Bayesian FE model updating scheme. Quantifying the value of added measured data has been recognized as one of the grand challenges in the SHM community. The value of using continuously measured data on the accuracy and resolution of model updating results has been quantified in this study. Finally, the results of probabilistic damage identifications are compared to their counterpart deterministic damage identification. Comparison of the probabilistic and deterministic FE model updating results provides insight into the benefits and limitations of the used probabilistic method over the deterministic FE model updating for structural identification.

The footbridge deck is divided into five segments and the added mass on each segment is considered as a model parameter to be calibrated in the probabilistic model updating process. An adaptive Metropolis-Hastings algorithm [34-39] is used to sample the posterior probability distributions of the updating parameters given the measured data and the model class. Effect of the amount of data used in the updating process is investigated on the accuracy of probabilistic damage identification results. This is done by comparing nine different cases of damage identification results obtained based on 1, 2, 6, 12, 24, 36, 48, 60 and 72 sets of identified modal parameters.

This paper is organized in the following order. In Section 2, the Dowling Hall Footbridge and its continuous monitoring system are introduced. Section 3 explains how damage is simulated on the footbridge through addition of mass on a segment of the bridge deck. Brief reviews of Bayesian FE model updating and the sampling technique used in this study are provided in Section 4 while the model updating results are reported in Section 5. Finally, the concluding remarks are presented in Section 6.

2. DOWLING HALL FOOTBRIDGE AND ITS CONTINUOUS MONITORING SYSTEM

The Dowling Hall Footbridge is located at the Medford, Massachusetts campus of Tufts University. Figure 1 shows the south view of the footbridge. The bridge is 3.9 m wide and consists of two 22 m spans. It connects the main campus on its western end to the Dowling Hall on its eastern end. The footbridge is composed of a reinforced concrete deck and a steel frame. More information about the structural details of the Dowling Hall Footbridge can be found in [40].

A continuous monitoring system was installed on the footbridge in November 2009 and has been providing continuous measurements since January 2010. The monitoring system consists of eight accelerometers and a data acquisition device that is connected to the Tufts wireless network. The monitoring system continuously samples the acceleration channels at a sampling rate of 2,048 Hz. A five-minute data sample is recorded at the top of every hour or when the root-mean square (RMS) value of an acceleration measurement exceeds 0.03 g. Locations of accelerometers on the footbridge are shown in Figure 2. Details about design and deployment of this continuous monitoring system can be found in [41]. The modal parameters of the first six (most excited) vibration modes identified from a preliminary test data in April 2009 are shown in
In this figure, the mode shape plots are interpolated between the locations of accelerometers using a cubic spline. The data-driven stochastic subspace identification method (SSI-Data) [42-43] is used for automatic modal identification of the footbridge.

3. SIMULATION OF STRUCTURAL DAMAGE ON THE FOOTBRIDGE

To simulate the effects of damage, a small segment of the footbridge deck was loaded with 2.29 metric tons (2,290 kg) of concrete blocks for 72 hours. The length of the loaded segment is 4.9 meters. The added mass will cause the same reduction in the natural frequency of mode one as a 35% loss of stiffness in the same segment of the bridge. Figure 4 shows the blocks on the bridge deck and safety measures that were considered for this test. 72 sets of ambient vibration measurements were collected once every hour and their corresponding modal parameters – representing model parameters of the damaged structure – were identified through an automated operational modal analysis framework. Figure 5 shows the effects of added mass on the identified natural frequencies of modes 1 to 6. The figure also shows the identified natural frequencies 24 hours before and 24 hours after the loading. The grey dots correspond to the undamaged condition while the black dots refer to the damaged (loaded) condition of the bridge. Significant drops of natural frequencies are observed for mode 1, 2, 3, and 4 while the natural frequencies of modes 5 and 6 remained almost unchanged. It is worth noting that the drops in the identified natural frequencies can be used for the detection of damage on the structure, however, they do not provide the location and extent of damage. Information about the identification success rate and statistics of modal parameters identified during the loading period are provided in Table 1. The most reliably identified modes are modes 3 and 4, while modes 5 and 6 have the largest estimation uncertainty.

4. BAYESIAN FE MODEL UPDATING AND MARKOV CHAIN SAMPLING

The first step in the FE model updating process is to calibrate an initial FE model of a structure based on its design information to a reference FE model corresponding to “as built” condition in the undamaged/baseline state of the structure. The initial FE model of the Dowling Hall Footbridge was created using FEDEASLab [44], a MATLAB-based structural analysis software, based on design drawings and visual inspections of the footbridge. The modal parameters extracted from the data recorded at 9 am of July 28, 2011 (one day before loading the footbridge) are selected as the reference modal parameters. The initial FE model is calibrated so that its modal parameters best match to reference modal parameters using a deterministic sensitivity-based FE model updating [31]. The deterministic FE model updating is performed by tuning the updating parameters of the initial model to minimize an objective function which represent the misfit between model-predicted and experimentally-identified modal parameters. The calibrated model, referred to as the reference FE model, is assumed to represent the true undamaged condition of the footbridge. In the Bayesian FE model updating performed in this study, the footbridge deck in the reference FE model is divided to five segments and the added mass of each segment is considered as an updating parameter. These five segments are shown in Figure 6. The considered segments in this study are defined based on the location of the sensors, and they are limited to five to assure having a globally identifiable problem. The instrumentation of the eastern side of the bridge, nearest to the Dowling Hall, was not possible because of the height above the ground. Low number of sensors on the eastern side of the bridge has resulted in selecting a large segment on this side of the bridge, i.e., segment 5. Using smaller segments (i.e.,
a larger number of updating parameters) on the eastern side of the bridge can therefore result in locally or globally unidentifiable problem unless more informative data (e.g., more sensors and larger number of modes) are used in the updating process. In addition, increasing the number of parameters will demand significantly higher computational efforts. Alternatively, a different number of segments and/or different segment lengths could be selected as updating parameters which would yield different updating results. FE model updating results are sensitive to the selection of updating parameters and this sensitivity can be viewed as one of the main limitations/shortcomings of this method for damage identification. Potentially, multiple model updating cases based on different sets of updating parameters (model classes) can be performed and the optimal set of updating parameters can be selected through a Bayesian model class selection technique [32].

It is also worth noting that in this study the added mass is uniformly distributed within one of the five updating segments which in turn reduces the modeling errors for damage identification. By limiting the amount of uncertainty in the distribution of damage, the performance of FE model updating with increasing the amount of measured data can be studied in the presence of modeling errors and variability in the identified modal parameters. It would also allow comparing the obtained results with the “exact” damage scenario which would not be available if the added mass was not uniformly distributed along the complete length of the substructure. In that case, the estimated damage measure will be representing the equivalent uniform change (or any predefined shape when using shape/damage functions [45]) and should be considered as another source of uncertainty (e.g., in modeling error). In real-world applications of model updating for damage identification, damage is often smeared along parts of different updating substructures, which causes increased error in damage identification results. Identification of smeared damage along different updating segments of a structure is out of scope of this study.

### 4.1. Bayesian Formulation

This section presents a summary of the Bayesian FE model updating formulation used in this study. More detailed formulation of the Bayesian model updating process can be found in seminal publications on this topic [17-21]. According to the Bayes theorem, conditional posterior probability distributions of updating parameters \( \theta \) given the measured data \( d \) (identified modal parameters in this study) and the model class \( M \) can be obtained by multiplying the likelihood function \( p(d | \theta, M) \), the conditional prior probability distribution of model parameters \( p(\theta | M) \), and a constant \( c \).

\[
p(\theta | d, M) = cp(d | \theta, M) p(\theta | M)
\]

The likelihood function is the probability of measured data given the updating model parameters and the normalization constant \( c \) is to ensure that the posterior probability density function (PDF) integrates to one. Only one model class is considered in this study, thus conditioning on the model class \( M \) will be dropped hereafter.

Data vector \( d \) contains a set of modal parameters identified from one set of ambient measurements. Modal parameters are assumed to be independently distributed from mode to mode and from natural frequencies to mode shapes. In other words, it is assumed that knowing the values of any observed modal parameters do not provide any information regarding the probability of observing other modal parameters. The likelihood function is defined as:
\[ p(d | \theta) = \prod_{m=1}^{N_m} p(\tilde{\lambda}_m | \theta) p(\tilde{\Phi}_m | \theta) \]  

(2)

where \( N_m \) is the number of identified modes, \( \tilde{\lambda}_m = \left( \frac{2\pi}{\tilde{f}_m} \right)^2 \) is the identified eigen-frequency and \( \tilde{\Phi}_m \) is the identified mode shape of mode \( m \). In the case of small modeling errors, it is reasonable to assume the estimation uncertainties of modal parameters are independent. However, when considerable modeling errors exist between the reference FE model and the real system, the assumption on independency of identified modal parameters will not be realistic. The eigen-frequency and mode shape errors are defined as:

\[ e_{\lambda_m} = \tilde{\lambda}_m - \lambda_m(\theta) \]  

(3)

\[ e_{\Phi_m} = \frac{\tilde{\Phi}_m}{\Phi_m} - a_m \frac{\Gamma_{\Phi_m}(\theta)}{\|\Gamma_{\Phi_m}(\theta)\|} \]  

(4)

The matrix \( \Gamma \) picks the observed degrees of freedom from model-calculated mode shape \( \Phi_m(\theta) \). The scaling factor \( a_m \) is defined in Equation (5).

\[ a_m = \frac{\tilde{\Phi}_m \Gamma_{\Phi_m}(\theta)}{\|\tilde{\Phi}_m\| \|\Gamma_{\Phi_m}(\theta)\|} \]  

(5)

Assuming the errors are zero-mean Gaussian distributed random variables and the mode shape covariance matrix is a diagonal matrix with all diagonal elements equal to \( \sigma^2_{\Phi_m} \), probability distributions of the identified eigen-frequency and mode shape at a given \( \theta \) become:

\[ p(\tilde{\lambda}_m | \theta) \propto \exp\left(-\frac{1}{2} \left( \frac{\tilde{\lambda}_m - \lambda_m(\theta)}{\sigma^2_{\lambda_m}} \right)^2 \right) \]  

(6)

\[ p(\tilde{\Phi}_m | \theta) \propto \exp\left[-\frac{1}{2\sigma^2_{\Phi_m}} \left( \frac{\tilde{\Phi}_m}{\Phi_m} - a_m \frac{\Gamma_{\Phi_m}(\theta)}{\|\Gamma_{\Phi_m}(\theta)\|} \right)^T \left( \frac{\tilde{\Phi}_m}{\Phi_m} - a_m \frac{\Gamma_{\Phi_m}(\theta)}{\|\Gamma_{\Phi_m}(\theta)\|} \right)\right] \]  

(7)

Inserting Equations (6) and (7) into (2) yields the following likelihood function:

\[ p(d | \theta) \propto \exp(-\frac{1}{2} J(\theta,d)) \]  

(8)

\[ J(\theta,d) = \sum_{m=1}^{N_m} \left( \frac{\tilde{\lambda}_m - \lambda_m(\theta)}{\sigma^2_{\lambda_m}} \right)^2 + \sum_{m=1}^{N_m} \frac{1}{\sigma^2_{\Phi_m}} \left( \frac{\tilde{\Phi}_m}{\Phi_m} - a_m \frac{\Gamma_{\Phi_m}(\theta)}{\|\Gamma_{\Phi_m}(\theta)\|} \right)^T \left( \frac{\tilde{\Phi}_m}{\Phi_m} - a_m \frac{\Gamma_{\Phi_m}(\theta)}{\|\Gamma_{\Phi_m}(\theta)\|} \right) \]  

(9)

In the case of having \( N_t \) independent sets of measured modal parameters, the probability \( p(d_1 : d_{N_t} | \theta) \) can be stated as:
\[ p(d_1 : d_{N_f} | \theta) = p(d_{N_f} | d_1 : d_{N_f-1}, \theta)p(d_1 : d_{N_f-1} | \theta) = \ldots = \prod_{n=1}^{N_f} p(d_n | \theta) \]
\[ = \hat{c} \prod_{n=1}^{N_f} \exp \left( -\frac{1}{2} J(\theta, d_n) \right) \]  

(10)

where it is assumed that the uncertainty in the \( n \)th data set is not affected by previous data sets, i.e., \( p(d_n | d_1 : d_{n-1}, \theta) = p(d_n | \theta) \). The posterior probability distribution of model parameters in Equation (1) is a joint PDF, and therefore a multi-dimension integration is required to obtain the marginal probability distribution of each model parameter:

\[ p(\theta_i | d) = \int_{\theta_i} p(\theta_i, \theta_{-i} | d) d\theta_{-i} \]  

(11)

where subscript ‘-i’ refers to all the updating parameters except the \( i \)th parameter. These probability distributions often cannot be calculated analytically and are usually estimated numerically. Markov Chain Monte Carlo (MCMC) methods are usually implemented to sample such complex multi-dimensional probability distributions. MCMC is a Monte Carlo sampling technique where samples are correlated through a Markov chain. In Monte Carlo method, samples are drawn independently from a target probability distribution. In some applications, generating independent and identically distributed samples from the target distribution is not feasible or computationally expensive; therefore, dependent samples can be used instead. One class of generating such dependency in a sequence of samples is Markov Chain. In Markov Chain Monte Carlo, the next sample is selected from a distribution that depends only on the current sample.

### 4.2. Sampling the Posterior Probability Distributions of Model Parameters

Among the MCMC methods, the Metropolis-Hastings (MH) algorithm is the most common method used in the application of FE model updating [19-21]. This algorithm is based on generating samples from a prescribed proposal PDF and accepting the samples with a probability of move. One intuitive choice for the proposal probability distribution is a localized, symmetric PDF at the current sample. In this case, the algorithm can be viewed as a random walk process that continues until the high probability region of the target PDF is sufficiently explored. However, direct application of the standard MH algorithm is not suited for cases where high probability regions of model parameters are concentrated in a small volume of the parameter space or in the cases with multi-modal target PDFs. To be able to address these difficulties, different adaptive Metropolis-Hastings (AMH) algorithms have been proposed. The adaption can be performed on either the proposal probability distribution or the target probability distribution. The algorithms with adaption on the target PDF, such as the adaptive algorithm used in [19] or the TMCMC algorithm [22], incorporate a number of intermediate target probability distributions that converge to the posterior PDF of Equation (1). The advantage of using these methods is their ability to sample the whole space and explore all the peaks (high probability regions). However, the main disadvantage of these methods is the high computational cost required in the case of complex structural models where the number of updating parameters increases. Another approach is to adapt the proposal PDF [37-38]. The advantage of this method is that it is fast for identifiable cases where the posterior distributions of updating parameters...
have unique clear peaks. In these cases, samples quickly converge to the highest probability region even when the region is sharply peaked.

In this study, the adaption is performed on the proposal PDF instead of the target PDF. The algorithm is briefly introduced below assuming a Gaussian distribution $N(0|\theta, \beta \Sigma)$, for the proposal probability distribution function.

- Initialize $\theta_0, \mu_0, \Sigma_0$, which are the vectors of initial parameter values, the mean vector, and the covariance matrix of the initial proposal probability distribution, respectively.
- At iteration $j+1$ given $\theta_j, \mu_j, \Sigma_j$, and $\beta_j$:
  - Generate sample $\xi_{j+1}$ from $N(\theta_j, \beta_j \Sigma_j)$ and set $\theta_{j+1} = \xi_{j+1}$ with probability of move $\alpha(\theta_j, \xi_{j+1})$. If $\xi_{j+1}$ is not accepted then set $\theta_{j+1} = \theta_j$.
  - Update the scale factor $\beta_{j+1}$, mean $\mu_{j+1}$, and covariance matrix $\Sigma_{j+1}$ of the proposal PDF as:
    \begin{align*}
    \log(\beta_{j+1}) &= \log(\beta_j) + \gamma_{j+1} \left[ \alpha(\theta_j, \xi_{j+1}) - \alpha^* \right] \\
    \mu_{j+1} &= \mu_j + \gamma_{j+1} (\theta_{j+1} - \mu_j) \\
    \Sigma_{j+1} &= \Sigma_j + \gamma_{j+1} \left[ (\theta_{j+1} - \mu_j)(\theta_{j+1} - \mu_j)^T - \Sigma_j \right]
    \end{align*}
  \label{eq:13}

The step size sequence $\gamma_j$ needs to satisfy two conditions of $\sum_{j=1}^{\infty} \gamma_j = \infty$ and $\sum_{j=1}^{\infty} \gamma_{j+\beta} < \infty$ for some $\beta > 0$, and here it is taken as $\gamma_j = 1/j^{0.1}$ which satisfies both conditions. The parameter $\alpha^*$ is the desired acceptance ratio and is taken as 44% \cite{38}. The optimal acceptance rate for one dimensional Gaussian target probability densities is 44% while the optimal rate will become 23% when the dimension of the target distribution goes to infinity \cite{39}.

5. FE MODEL UPDATING RESULTS

The first case of Bayesian FE model updating is based on the average of identified modal parameters over the 72 sets of data and the results are presented in Section 5.1. In section 5.2, nine cases of model updating are performed using different subsets (1, 2, 6, 12, 24, 36, 48, 60, and 72 sets) of available identified modal parameters. In all the FE Bayesian model updating cases, 100,000 Markov Chain samples are generated. The number of burn-in samples is 10,000. The prior distributions of segments’ added masses are considered as independent uniform distributions with the lower bound of 0 and upper bound of 6.8 tons. Values of the standard deviations for eigen-frequencies in Equation (9) are considered as:

$$\sigma_{\lambda_m} = w_m \text{COV}_{\lambda} \hat{\lambda}_m$$
\label{eq:14}
where $\text{COV}_{\lambda}$ is taken as the average of calculated eigen-frequency coefficients-of-variation (COVs) of all the six modes (the COVs of natural frequencies are reported in Table 1) and $w_m$ is a weight associated to each mode, taken as 10 for modes 2, 5, and 6 to account for their higher estimation uncertainties and one for the other modes. This factor is selected based on prior knowledge of system identification studies of this footbridge [15, 31, 41]. Standard deviations of mode shapes are computed from:

$$\sigma_{\Phi_n} = \sqrt{N_s w_m \text{COV}_{\lambda}}$$  \hspace{1cm} (15)

where $N_s$ is the number of accelerometers, or number of identified mode shape components, i.e., $N_s = 8$.

### 5.1. Bayesian Updating Results Based on Average Modal Parameters

The posterior probability distributions of the added mass at the five considered segments of the footbridge are sampled using the aforementioned sampling process. Figure 7 shows the samples versus their posterior probability densities. The burn-in samples are excluded from the graphs. The maximum a-posteriori (MAP) estimates of the updating parameters are shown by black dots and the exact values of the added mass are shown by stars. The convergence of the sampling process is confirmed by the convergence of their means, standard deviations, and the correlation between the samples. Figure 8a shows the cumulative mean values of samples generated at updating parameters 1 and 2 (i.e., added masses on segments 1 and 2). The cumulative means vary significantly when the total number of samples is small but its value becomes almost constant as the number of samples increases beyond 60,000. Figure 8b shows the cumulative standard deviations of the samples and Figure 8c plots the correlation coefficients between the samples. For both updating parameters, the correlations become negligible beyond 6,000 samples.

Table 2 provides the statistical information of the mean, standard deviation, and MAP of the five model parameters based on the MCMC samples as well as the MAP estimates when using Gaussian kernel functions to fit the posterior probability distributions of the updating parameters. The updated model parameters obtained from a deterministic FE model updating are also listed in the table. From this table and Figure 7 the following observations can be made.

- For segments 1, 3, 4 and 5, the maximum a-posteriori of added mass values are close to zero, which is expected as no mass is loaded on these segments. For segment 2, the posterior probability density of added mass is the largest in the range of 2.0 to 2.5 tons, with the MAP of 2.29 tons (which is equal to the exact value).

- The posterior probability distributions of the added mass on segments 1, 3, 4, and 5 are one-sided because of the lower bound constraint of zero for these updating parameters implemented in the prior PDFs. The posterior PDF of segment 2 is two-sided but not symmetric; it is skewed to the left. This may be due to the fact that larger values of $\theta_2$ have to be compensated by negative values of other updating factors which are not allowed. This has resulted in large bias in the samples mean for each of the model parameters.

- The estimates of parameters $\theta_1$ to $\theta_4$ have larger variances than that of $\theta_5$. This is due to the higher sensitivity of updating parameter $\theta_5$ to the considered modal data, i.e., larger observability of this parameter in the view of data. Alternatively, this observation could be
reached intuitively as this is the largest segment of the bridge deck and therefore has the largest effect on the modal parameters of the footbridge.

- The deterministic estimates of model parameters are in relatively good agreement with the MAP estimates. Theoretically, the deterministic and MAP estimates should be identical when using compatible objective/likelihood functions. However, in this application due to limited number of data sets used (i.e., one) and smaller number of samples, there is some discrepancy between these two estimates.

The MCMC samples are used to construct the posterior PDF of each model parameter instead of calculating the integral of Equation (11) analytically. Gaussian kernel functions are used to estimate the marginal posterior PDF of each model parameter. Figure 9 shows the normalized envelope of sampled marginal posterior PDFs (envelope of Figure 7) and the kernel marginal posterior PDFs of updating parameters. The envelope of samples at each point corresponds to the maximum of the joint posterior PDF for the specific value of the considered parameter. The kernel PDF estimate is based on summation of normalized Gaussian kernel functions for each of the sample bins and a weight associated with each sample. The weights of the samples are taken as the posterior probability of them and the bin widths in this section are considered as 0.08 tons for \( \theta_1 \) and \( \theta_4 \), and 0.05 tons for \( \theta_2 \), \( \theta_3 \), and \( \theta_5 \). The bin width for each updating parameter is selected based on the observed range and standard deviation of samples; therefore, different bin widths are used for different parameters. It is worth noting that the kernel probability distribution estimates fail to accurately represent the MAP and standard variation of samples. As it will be shown and discussed in the next section, the kernel PDF estimates provide much better representation of the sample probability distributions as the number of data sets used in the likelihood function increases.

### 5.2. Bayesian Updating Results Based on Subsets of Identified Modal Parameters

In this section, effects of the number of data sets used in the likelihood function on the accuracy of updating results are investigated. Nine cases of probabilistic FE model updating are performed with \( N_t = 1, 2, 6, 12, 24, 36, 48, 60, \) and 72 data sets (see Equation (10)), corresponding to the first \( N_t \) sets of identified modal parameters. As expressed in Equation (10), using \( N_t \) sets of data is equivalent to the multiplication of \( N_t \) likelihood functions, each including one set of identified modal parameters. Figure 10 shows the envelope of generated Markov Chain samples with the MAP of samples in each bin shown by black dots, the kernel marginal posterior PDFs, and the approximated Gaussian posterior PDFs for the updating model parameter \( \theta_2 \) and nine cases of model updating (\( N_t = 1, 2, \ldots, 72 \)). The light grey lines in this figure are approximated as a Gaussian distribution with mean and standard deviation of the Markov Chain samples. Alternatively, the posterior probability distribution of the updating model parameters can be shown by the cumulative distribution function (CDF) of samples. The probability that added mass at the \( i^{th} \) segment is lower than \( L \) tons is estimated using the \( N_c \) generated Markov Chain samples:

\[
P_{i}^{AM} (L \mid d) = P \{ \theta_i \leq L \mid d \} \approx \frac{1}{N_c} \sum_{j=1}^{N} H \left( L - \theta_{i,j} \right)
\]  

(16)

where \( H(z) \) is a unit step function, defined as \( H(z) = 1 \) if \( z > 0 \) and \( H(z) = 0 \) if \( z \leq 0 \). Figure 11 shows the CDFs of the added mass at segments 1, 3, 4, and 5 for the nine probabilistic FE model
updating cases. The MAP, mean, and standard deviation estimates of the posterior probabilities of all five parameters are reported in Table 3. The standard deviation and the bias between the MAP and sample means as function of data points \( (N_t) \) are plotted in Figures 12 and 13, respectively, for the five parameters. From Figures 10-13 and Table 3, the following observations can be made.

- The estimation uncertainties (i.e., standard deviations) of the updating parameters significantly decrease as more data is used in the Bayesian model updating process. However, this reduction becomes less significant as the number of data sets exceeds 36. Therefore, it is expected that additional data (more than 72 sets) would not drastically improve the estimation accuracy of updating parameters. Such information can be used to quantify the value of additional data for parameter estimation.

- The three estimates of the posterior probability distribution of each model parameter, namely the sample envelope, the kernel marginal PDF, and the approximated Gaussian PDF are very different for the case \( N_t = 1 \), but these estimates get closer to each other as more data sets are used. The three posterior PDF estimates are in good agreement for the case of \( N_t = 72 \) with MAP values of 2.48, 2.49, and 2.49 tons. It should be noted that although the sample envelope and kernel probability distribution estimate represent different information about probability distribution of an updating parameter, these two types of information become similar when the updating parameters are independent. Therefore, the closer the kernel PDFs represent the MAP and standard deviation of the samples, the smaller is the correlation between updating parameters.

- With increasing the number of data sets, the posterior PDFs become more symmetrical and therefore the discrepancies between the sample means and MAP estimates are reduced.

- The standard deviation of model parameter \( \theta_4 \) is much larger than the standard deviations of parameters \( \theta_1, \theta_3, \) and \( \theta_5 \), especially when \( N_t \geq 12 \). This can be due to the fact that this segment is located on the middle support and has smaller vibration responses and therefore, the identified modal parameters are less sensitive to this updating parameter.

- Although the identification method can successfully predict the location and extent of the added mass on the footbridge, the MAP estimate of added mass on segment 2 has some bias from the actual mass of 2.29 tons. This bias in the MAP estimation (0.19 tons for the case of \( N_t = 72 \)) is mainly due to the modeling error in the reference FE model and the variability in the identified modal parameters caused by changing environmental conditions and estimation errors.

Figure 14 shows the posterior correlation coefficients between updating parameters for the nine different cases of model updating. The correlation coefficients are high when few data sets are used, but they reduce as \( N_t \) increases. In other words, the updating parameters are sampled more independently when more data sets are used (maximum correlation coefficients is calculated as 0.17 when \( N_t \geq 36 \)). To quantify the effects of measured data on model updating results, an eigen-analysis is performed on the covariance matrices of prior and posterior joint PDFs [33, 46-47]:

\[
\Sigma_{p_{\theta}} X = \Lambda \Sigma_{p_{\theta}} X \tag{17}
\]
where the eigenvectors $X$ are orthogonal directions in the parameter space ranked based on the corresponding eigenvalues. The eigenvalues give a measure of reduction in variances from prior covariance matrix $\Sigma_{pr}$ to posterior covariance matrix $\Sigma_{po}$. Figure 15 shows the eigenvalues computed from Equation (17) for the nine model updating cases. The changes in the entropy of updating parameters from prior to posterior quantify the value of measured data on the accuracy of updated parameters and can be estimated as:

$$\Delta h = h_{pr} - h_{po} \approx -0.5 \sum_{k=1}^{N_t} \log \lambda_k$$

(18)

where $h$ refers to information entropy, and $\lambda_k$ is the $k^{th}$ eigenvalue computed from Equation (17). Each term of $-0.5 \log \lambda_k$ represents the relative contribution of the corresponding eigenvalue to the total entropy reduction from the prior to the posterior probability distribution of updating parameters. The last eigenvalue has little contribution to the total entropy reduction when $N_t$ is small, but its contributions becomes more significant as $N_t$ increases. For example, the fifth eigenvalue accounts for only 9.7% of the total entropy reduction when $N_t = 1$, but contribution of $\lambda_5$ becomes 17.7% for $N_t = 72$. Inversely, contribution of $\lambda_1$ is 29.3% when $N_t = 1$ and is reduced to 22.1% when $N_t = 72$. The information entropy reduction is more sensitive to the number of data sets when $N_t$ is small but this reduction becomes less significant for cases of using more than 36 data sets. This observation is consistent with the reduction in standard deviations and bias of model parameter estimates in Figures 12 and 13.

It is observed that using different subset of data with the same number of data sets will provide similar posterior probability distributions of model parameters but with slightly different MAP estimates. Figure 16 shows the posterior PDFs of the added mass on segment 2 using three different subsets of identified modal parameters with $N_t = 24$. The first subset corresponds to the data collected during the first 24 hours of loading, the second set contains the second 24 hours of data, and the third set is the last 24 hours of the measured data during the loading period. Although the three PDFs have almost the same standard deviations (0.044, 0.044, 0.043 tons), their MAP estimates are slightly different (2.50, 2.47, and 2.52 tons). In practice, structural damage is quantified by comparing the updating model parameters in the damaged and undamaged states of a structure. Therefore, it is important to recognize that the posterior PDFs or MAP estimates of the updating model parameters are conditional on the measured data and therefore, depend on the accuracy/uncertainty of measurements in both the damaged and undamaged conditions. It should be noted that the changes in environmental conditions such as air temperature may have significant effects on the identified modal parameters, especially when the ambient temperature drops below the freezing point [31]. These effects can consequently affect the model updating results and should be taken into account [31] when large variations in environmental conditions are observed. In the current application, the environmental conditions remained relatively consistent during the week of testing with temperatures being in the range of 17 to 33 degrees Celsius.

The obtained bias in model updating results could be reduced by including the effects of modeling errors in the Bayesian formulation. For example, Sohn and Law [16] assumed that the mean values of the natural frequency errors and mode shape errors in Equations (3) and (4) can be approximated by the corresponding mean values in the undamaged state. Beck and co-workers [48] included the effects of modeling error in their formulation by introducing system
mode shapes and system eigen-frequencies. System mode shapes are also used in [18] and [21]. In this study, the model updating results are found to be sensitive to the considered standard deviations of the identified natural frequencies and mode shapes in Equations (14) and (15), i.e., the posterior probability distributions obtained are conditional on the $\sigma_j$ and $\sigma_\Phi$ values. The model updating results are sensitive to both the relative standard deviations among different modes and the relative standard deviations between the natural frequency and mode shape of each mode. Some studies have been performed to mitigate the effects of this uncertainty [49].

Finally, to show the computational efficiency of the adaptive sampling algorithm used in this study, the generated sequence of samples for parameters 1 and 2 and the case of $N_t = 72$ are plotted in Figure 17. It can be seen that after only 3,000 iterations the samples reach the high probability region of the parameter space. The number of samples required to reach convergence was found to be much larger when using the standard MH algorithm. The acceptance ratios of all runs are close to 44%, which is the desired acceptance ratio used in Equation (13).

5.3. Deterministic FE Model Updating Results

Deterministic FE model updating is performed using the measured “damaged” data through a global optimization approach, namely the “multistart” function from the MATLAB global optimization toolbox [50]. This optimization algorithm is based on several Gauss-Newton sensitivity-based optimizations with different starting points to ensure global optimality. Equation (9) is considered as the deterministic objective function to be minimized. The maximum number of iterations for each optimization is limited to 80 and maximum number of initial points is assigned as 10. Figure 18 shows the scatter of the updating parameters for all 72 deterministic updating cases. It is observed that, except for a few outliers, the added mass on segments 1, 3, and 4 are accurately estimated as zero with no variability. In this case, the outliers are defined as identification cases with nonzero mass on either segment 1, 3, or 4. However, the added mass on segments 2 and 5 are estimated with larger variability. The main sources of estimation errors in updating parameters are incompleteness of identified modal parameters from the corresponding measured data. The modal parameters in the outlier cases are incomplete (at least missing a mode) and are identified with larger estimation errors, which yield to inaccurate model updating results. It is also observed that when the first mode is missing (i.e., could not be identified) in a data set, relatively larger added mass is predicted at segment 2. These large estimation errors are avoided in the Bayesian approach when multiple data sets are used in one updating process.

6. CONCLUSIONS

A Bayesian FE model updating approach is implemented for identification of physically simulated damage on the Dowling Hall Footbridge. The extent and location of the damage (added mass) was accurately estimated based on the continuously measured vibration data. The probability distributions of the updating parameters do not only provide an estimate for the location and extent of damage but also a measure of confidence/uncertainty of the damage identification results.

The MAP estimates of updating model parameters match the exact values and are in a good agreement with the optimum values from the deterministic FE model updating. Effects of the number of data sets used in the identification process (i.e., “value” of added data) are
investigated by using different subsets (1, 2, 6, 12, 24, 36, 48, 60, and 72) of available data. Estimation uncertainty of the updating model parameters are significantly reduced by adding more data sets to the likelihood function, which implies more accurate model updating results. However, this reduction becomes less significant as the number of data sets exceeds 36. Therefore, it is expected that additional data (more than 36 sets) would not drastically improve the estimation accuracy of updating parameters. Such information can be used to quantify the value of additional data for parameter estimation. Adding more data sets also affects the shape of the posterior PDFs of updating parameters resulting in smaller bias between the sample means and the MAP estimates of model parameters. It is also worth noting that in the application of deterministic FE model updating, addition of more data sets will not necessarily improve the model updating results. Thus, probabilistic FE model updating approaches based on multiple sets of measurements are strongly recommended for structural identification purposes to yield more accurate prediction and detection of potential structural damage.

Although the implemented FE model updating method has successfully estimated the location and extend of damage in this study, but in general the success of this method depends on the accuracy of the initial FE model, the selected updating parameters, and the considered residuals and their weights in the objective function. Sensitivity of damage identification results to these factors can be viewed as one of the main limitations of this method for implementation by the practicing engineers without experience in model updating and inverse problems. For a robust identification, however, different combination of these factors (i.e., initial models, updating parameters, and objective functions) can be considered as different model classes, and a Bayesian model class selection/averaging technique can be used to select the optimal set of these factors.

ACKNOWLEDGEMENT

The authors would like to acknowledge support of this study by the National Science Foundation Grant No. 1125624 which was awarded under the Broadening Participation Research Initiation Grants in Engineering (BRIGE) program. The authors also acknowledge Ms. Alyssa Kody for the design and performance of the bridge load test, Ms. Rachele Pesenti for her help in programming the real-time module of the data acquisition system, and Mr. Bidiak Amana for his help in setting up the wireless network connection of Dowling Hall footbridge. The opinions, findings, and conclusions expressed in the paper are those of the authors and do not necessarily reflect the views of the individuals and organizations involved in this project.

REFERENCES


Tables:

**Table 1.** Statistics of identified modal parameters during loading period

<table>
<thead>
<tr>
<th>Mode</th>
<th>$N_a$</th>
<th>$\mu^j_a$</th>
<th>$\sigma^j_a$</th>
<th>$COV^j_a$</th>
<th>$\sigma^j_{\Phi}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46</td>
<td>4.51</td>
<td>0.021</td>
<td>0.46</td>
<td>0.054</td>
</tr>
<tr>
<td>2</td>
<td>61</td>
<td>5.70</td>
<td>0.028</td>
<td>0.49</td>
<td>0.054</td>
</tr>
<tr>
<td>3</td>
<td>68</td>
<td>6.83</td>
<td>0.023</td>
<td>0.33</td>
<td>0.031</td>
</tr>
<tr>
<td>4</td>
<td>68</td>
<td>8.57</td>
<td>0.025</td>
<td>0.29</td>
<td>0.014</td>
</tr>
<tr>
<td>5</td>
<td>65</td>
<td>12.97</td>
<td>0.104</td>
<td>0.80</td>
<td>0.185</td>
</tr>
<tr>
<td>6</td>
<td>67</td>
<td>13.45</td>
<td>0.097</td>
<td>0.72</td>
<td>0.067</td>
</tr>
</tbody>
</table>

1. Total number of available modal parameters
2. Mean of available natural frequencies [Hz].
3. Standard deviation of available natural frequencies [Hz].
4. Coefficient of variations of available natural frequencies [%].
5. Standard deviation of available unit-length normalized mode shapes; defined for mode $m$ as: $\sigma^2_{\Phi_m} = \frac{1}{N_a} \sum_{j=1}^{N_a} [\Phi^j_m - \Phi_m]^2$.  

**Table 2.** Statistical information of model parameters when using average of modal parameters in the likelihood function

<table>
<thead>
<tr>
<th></th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$\theta_3$</th>
<th>$\theta_4$</th>
<th>$\theta_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAP [tons]</td>
<td>0.00</td>
<td>2.29</td>
<td>0.03</td>
<td>0.08</td>
<td>0.19</td>
</tr>
<tr>
<td>Mean of Samples [tons]</td>
<td>0.76</td>
<td>1.67</td>
<td>0.52</td>
<td>0.59</td>
<td>0.18</td>
</tr>
<tr>
<td>STD of Samples [tons]</td>
<td>0.70</td>
<td>0.46</td>
<td>0.53</td>
<td>0.51</td>
<td>0.16</td>
</tr>
<tr>
<td>MAP of kernel marginal PDF [tons]</td>
<td>0.16</td>
<td>2.00</td>
<td>0.05</td>
<td>0.16</td>
<td>0.05</td>
</tr>
<tr>
<td>Mean of kernel marginal PDF [tons]</td>
<td>0.49</td>
<td>1.89</td>
<td>0.33</td>
<td>0.49</td>
<td>0.15</td>
</tr>
<tr>
<td>STD of kernel marginal PDF [tons]</td>
<td>0.49</td>
<td>0.30</td>
<td>0.33</td>
<td>0.40</td>
<td>0.13</td>
</tr>
<tr>
<td>Optima from deterministic FE model updating [tons]</td>
<td>0.00</td>
<td>2.34</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>No. of data sets</td>
<td>$\theta_1$</td>
<td>$\theta_2$</td>
<td>$\theta_3$</td>
<td>$\theta_4$</td>
<td>$\theta_5$</td>
</tr>
<tr>
<td>------------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td>MAP  mean  STD</td>
<td>MAP  mean  STD</td>
<td>MAP  mean  STD</td>
<td>MAP  mean  STD</td>
<td>MAP  mean  STD</td>
</tr>
<tr>
<td>1</td>
<td>0.04  1.97  1.390</td>
<td>2.75  1.66  0.598</td>
<td>0.00  0.65  0.561</td>
<td>0.29  0.65  0.556</td>
<td>0.01  0.16  0.143</td>
</tr>
<tr>
<td>2</td>
<td>0.05  0.46  0.418</td>
<td>2.60  2.31  0.251</td>
<td>0.02  0.27  0.249</td>
<td>0.00  0.37  0.335</td>
<td>0.02  0.12  0.097</td>
</tr>
<tr>
<td>6</td>
<td>0.00  0.14  0.127</td>
<td>2.56  2.44  0.101</td>
<td>0.00  0.10  0.089</td>
<td>0.00  0.20  0.175</td>
<td>0.01  0.07  0.062</td>
</tr>
<tr>
<td>12</td>
<td>0.00  0.07  0.064</td>
<td>2.53  2.46  0.065</td>
<td>0.00  0.04  0.042</td>
<td>0.00  0.14  0.131</td>
<td>0.03  0.06  0.044</td>
</tr>
<tr>
<td>24</td>
<td>0.00  0.03  0.028</td>
<td>2.50  2.47  0.044</td>
<td>0.00  0.02  0.023</td>
<td>0.00  0.07  0.065</td>
<td>0.00  0.03  0.026</td>
</tr>
<tr>
<td>36</td>
<td>0.00  0.02  0.018</td>
<td>2.44  2.42  0.034</td>
<td>0.00  0.01  0.014</td>
<td>0.00  0.04  0.040</td>
<td>0.00  0.02  0.017</td>
</tr>
<tr>
<td>48</td>
<td>0.00  0.01  0.014</td>
<td>2.49  2.47  0.028</td>
<td>0.00  0.01  0.010</td>
<td>0.00  0.04  0.039</td>
<td>0.00  0.02  0.014</td>
</tr>
<tr>
<td>60</td>
<td>0.00  0.01  0.011</td>
<td>2.46  2.44  0.026</td>
<td>0.00  0.01  0.008</td>
<td>0.00  0.03  0.028</td>
<td>0.00  0.01  0.011</td>
</tr>
<tr>
<td>72</td>
<td>0.00  0.01  0.010</td>
<td>2.48  2.49  0.023</td>
<td>0.00  0.01  0.007</td>
<td>0.00  0.02  0.023</td>
<td>0.00  0.01  0.009</td>
</tr>
</tbody>
</table>
Figure 1. South view of the Dowling Hall Footbridge.
Figure 2. Layout of accelerometers on the bridge.
Figure 3. Identified modal parameters from preliminary test data.
Figure 4. Concrete blocks loaded on footbridge’s deck for three days.
Figure 5. Hourly identified natural frequencies before, during, and after loading.
Figure 6. Five segments along the footbridge deck corresponding to the five updating parameters.
Figure 7. Samples vs. their posterior probability densities using average of the 72 sets of identified modal parameters.
Figure 8. (a) Mean, (b) standard deviation, (c) and correlation coefficient between the samples of added mass at segments 1 and 2.
Figure 9. Posterior PDF of added mass at segments 1 to 5.
Figure 10. Posterior PDFs of added mass at segment 2 based on different subsets of identified modal parameters (sample envelope: black solid line, kernel marginal PDF: grey dashed line, Gaussian approximation: light grey solid line).
Figure 11. Posterior cumulative distribution functions (CDF) of segments 1, 3, 4, and 5.
Figure 12. Posterior standard deviation versus $N_t$. 
Figure 13. Difference between sample mean and MAP estimates versus $N_t$. 

*Figure 13.* Difference between sample mean and MAP estimates versus $N_t$. 

<table>
<thead>
<tr>
<th>Sample Mean − MAP</th>
<th>θ₁</th>
<th>θ₂</th>
<th>θ₃</th>
<th>θ₄</th>
<th>θ₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 14. Correlation coefficients of updating model parameters in all 9 Bayesian model updating cases.
Figure 15. Eigenvalues of the covariance matrices (information gain from prior to posterior).
Figure 16. Posterior PDFs of segment 2 with different 24-hour long subsets of measured data.
Figure 17. Distributions of samples for $\theta_1$ and $\theta_2$: (a) all samples; (b) first 5000 samples.
Figure 18. Scatter of updating parameters for the 72 cases of deterministic FE model updating, the outliers are shown by crosses and are not included in the computation of mean and standard deviation values.