Proc. of the 6th Intl. Symposium on Reliability Engineering and Risk Management (6ISRERM)
31 May – 1 June 2018, Singapore
Editor(s) Xudong Qian, Sze Dai Pang, Ghim Ping Raymond Ong, Kok-Kwang Phoon
Copyright © 2018 Author(s). All rights reserved.

BAYESIAN MODEL SELECTION IN FINITE-ELEMENT MODEL UPDATING OF A BOLT-CONNECTED STEEL FRAME

YIN TAO[1] and ZHU HONG-PING[2]

[1] *School of Civil Engineering, Wuhan University, Wuhan 430072, P.R. China.*
E-mail: tyin@whu.edu.cn
[2] *School of Civil Engineering & Mechanics, Huazhong University of Science and Technology, Wuhan 430074, P.R. China.*
E-mail: hpzhu@mail.hust.edu.cn

This paper is intended to investigate the issues of finite-element (FE) model-class selection for the purpose of choosing suitable parameterized models to update FE model of a bolt-connected steel frame based on Bayesian evidence inference method together with Markov Chain Monte Carlo (MCMC) technique. By employing the concept of information divergence, the amount of information needed to be extracted from the measured data is explicitly quantified during the FE model updating procedure. Then, for achieving a trade-off between the complexity of a prescribed FE model class and that of its corresponding information-theoretic interpretation, such information is utilized for penalizing the complexity of FE parameterization to ensure that a relatively simple parameterization scheme can be obtained for keeping the similar model-data matching and avoiding the over-fitting problem arisen from excessive modeling parameters. Through numerical case studies conducted for a two-story bolt-connected steel frame structure, the feasibility and validity of proposed methodology is demonstrated.

Keywords: Model updating, Bayesian evidence, information divergence, MCMC.

1 Introduction

The finite element (FE) approach has been extensively used in various engineering fields over the past few decades. However, due to assumption and uncertainty arisen from the theoretical hypothesis, boundary condition, and geometric and material properties, there is an unavoidable mismatching between the model-predicted and measured responses. By utilizing the measured lower order modes to update the initial FE model, the FE model updating technique is an effective means intending to improve the model accuracy in order to achieve a refined FE model matching well with the field measurement (Friswell and Mottershead 1995).

In order to obtain higher accuracy for structural analysis, the FE model of the target structure tends to be fine enough to approximate the structural details, leading to the increase of the model complexity. This, however, is not beneficial for the procedure of FE model updating, which is typically an inverse problem of structural dynamics. Since the repeated solution of the large eigenvalue problem is usually required in the FE model updating, the computational cost of this procedure eventually resulted unaffordable when dealing with complex models. In addition, the measured data is incomplete due to the problems of limited number of sensors, measurement noise and truncation error of higher order modes, etc., which cause the inability to capture the full dynamics of the structure, rendering the inverse problem of FE model updating uncertain and ill-posed (Datta 2002, Mthembu et al. 2011). Moreover, there usually exist multitudinous FE models with varying level of complexity that can be developed from the engineering judgment. The inverse problem may be non-uniquely solvable for FE models with higher parameterized complexity due to the large number of uncertain parameters to be identified as compared to the limited measurement information available. Thus, it is particularly important to choose the FE model with the suitable complexity for the model updating purpose.

In this paper, the problem of choosing suitable class of parameterized models for the FE model updating is addressed following the Bayesian evidence inference method. Within the concept of information theory, the amount of information needed to be extracted from the measured data for the specified parameterized model class is explicitly quantified during the FE model updating procedure. Then, this information is utilized for penalizing the parameterization complexity of FE model for ensuring a FE parameterization scheme with suitable complexity. It also avoids the over-fitting problem arisen from excessive modeling parameters, guaranteeing a trade-off between the complexity of a prescribed class of parameterized models and that of its corresponding information-theoretic interpretation. The validity of the proposed methodology is verified by the numerical cases studies of a two-story steel frame with bolted connection.

2 Theoretical background

Let $\mathcal{D}$ denote the dynamic data measured from the target structural system. The goal is to use $\mathcal{D}$ to select the most plausible class of models representing the system out of N_M given classes of parameterized models $\mathcal{M}_1$, $\mathcal{M}_2, \ldots, \mathcal{M}_{N_M}$, where $\mathcal{M}_j$ specifies not only a class of deterministic dynamic models but also the probability descriptions for the prediction error. By following the Bayes' theorem, the probability of model class $\mathcal{M}_j$ conditional on the data $\mathcal{D}$ is obtained as

$$p(\mathcal{M}_j|\mathcal{D},\mathcal{U}) = p(\mathcal{D}|\mathcal{M}_j,\mathcal{U})p(\mathcal{M}_j|\mathcal{U})/p(\mathcal{D}|\mathcal{U}) \tag{1}$$

where $p(\mathcal{D}|\mathcal{U})$ is given by the law of total probability. $\mathcal{U}$ represents the user's judgment on the initial plausibility of the model classes expressed as the prior plausibility $p(\mathcal{M}_j|\mathcal{U})$ on the model class $\mathcal{M}_j$, $j = 1,2,\cdots,N_M$. Gaussian prior is assumed in this paper. The prior plausibilities are normalized as $\sum_{j=1}^{N_M} p(\mathcal{M}_j|\mathcal{U}) = 1$, and it's simply assumed that each class of models has the same initial plausibility in this paper.

The factor $p(\mathcal{D}|\mathcal{M}_j,\mathcal{U})$, expressing how likely the data $\mathcal{D}$ are obtained with the specified model class $\mathcal{M}_j$, is the evidence for this class of models provided by the data. By assuming further that $\mathcal{M}_j$ alone specifies the probability density function (PDF) for the data, the user's preference $\mathcal{U}$ can thus be dropped from the notation $p(\mathcal{D}|\mathcal{M}_j,\mathcal{U})$ hereinafter. It is shown in Eq. (1) that the most plausible class of models is the one that maximizes $p(\mathcal{M}_j|\mathcal{D})$ which is equivalent to maximize the evidence $p(\mathcal{D}|\mathcal{M}_j)$ with respect to j. The evidence for $\mathcal{M}_j$ provided by the data $\mathcal{D}$ is given by the law of total probability as

$$p(\mathcal{D}|\mathcal{M}_j) = \int_{\boldsymbol{\Theta}_j} p(\mathcal{D}|\boldsymbol{\theta}_j,\mathcal{M}_j)p(\boldsymbol{\theta}_j|\mathcal{M}_j)\mathrm{d}\boldsymbol{\theta}_j \tag{2}$$

where $\boldsymbol{\theta}_j \in \mathbb{R}^{N_j\times 1}$ is the parameter vector of the jth class of models $\mathcal{M}_j$ in the parameter space $\boldsymbol{\Theta}_j$. It's noted that the parameter vector $\boldsymbol{\theta}_j$ depends on the model class $\mathcal{M}_j$ even though it is not explicitly reflected in the symbol for simplicity. $p(\boldsymbol{\theta}_j|\mathcal{M}_j)$ is the prior PDF specified by the user.

The likelihood function $p(\mathcal{D}|\boldsymbol{\theta}_j,\mathcal{M}_j)$ is given by

$$p(\mathcal{D}|\boldsymbol{\theta}_j,\mathcal{M}_j) = (2\pi\sigma^2)^{-\frac{N_s N_m}{2}} \exp\left[-\frac{N_s N_m}{2\sigma^2}\mathcal{J}(\boldsymbol{\theta}_j;\mathcal{D},\mathcal{M}_j)\right] \tag{3}$$

where $\mathcal{J}(\boldsymbol{\theta}_j;\mathcal{D},\mathcal{M}_j)$ is the measure-of-fit function defining the differences between the measured and model-predicted modal parameters as

$$\mathcal{J}(\boldsymbol{\theta}_j;\mathcal{D},\mathcal{M}_j) = \frac{1}{N_s N_m}\sum_{s=1}^{N_s}\sum_{m=1}^{N_m}\left[\left(\frac{f_m(\boldsymbol{\theta}_j) - \hat{f}_m^{(s)}}{\hat{f}_m^{(s)}}\right)^2 + \frac{\boldsymbol{\phi}_m(\boldsymbol{\theta}_j)^{\mathrm{T}}\left(\mathrm{I} - \widehat{\boldsymbol{\phi}}_m^{(s)}\widehat{\boldsymbol{\phi}}_m^{(s)\mathrm{T}}\right)\boldsymbol{\phi}_m(\boldsymbol{\theta}_j)}{\alpha\left\|\boldsymbol{\phi}_m(\boldsymbol{\theta}_j)\right\|^2}\right] \tag{4}$$

and N_m denotes the number of measured natural frequencies $\hat{f}_m$ and mode shapes vector $\widehat{\boldsymbol{\phi}}_m \in \mathbb{R}^{N_O}$, and N_O is the number of observed degrees of freedom (DOF), and N_s is the number of measured data sets. $f_m(\boldsymbol{\theta}_j)$ and $\boldsymbol{\phi}_m(\boldsymbol{\theta}_j)$ represent the mth predicted natural frequencies and mode shape vectors with model class $\mathcal{M}_j$ parameterized by $\boldsymbol{\theta}_j$. α is the ratio between the prediction-error variances of mode shape vectors and modal frequencies. It is assumed here that the measured mode shapes are normalized so that its Euclidean norm $\left\|\widehat{\boldsymbol{\phi}}_m^{(s)}\right\| = 1$.

By employing the Bayesian statistical identification framework (Beck and Katafygiotis 1998), the posterior PDF of the parameter vector $\boldsymbol{\theta}_j$ can be given by

$$p(\boldsymbol{\theta}_j|\mathcal{D},\mathcal{M}_j) = p(\mathcal{D}|\boldsymbol{\theta}_j,\mathcal{M}_j)p(\boldsymbol{\theta}_j|\mathcal{M}_j)/p(\mathcal{D}|\mathcal{M}_j) \tag{5}$$

In globally identifiable cases, the posterior PDF $p(\boldsymbol{\theta}_j|\mathcal{D},\mathcal{M}_j)$ in Eq. (5) given a large amount of data $\mathcal{D}$ may be approximated accurately by a Gaussian distribution, so the evidence $p(\mathcal{D}|\mathcal{M}_j)$ can be approximated by using Laplace's method for asymptotic expansion (Papadimitriou 1997). However, this asymptotic expansion is not valid for the general case where the posterior PDF may not be approximated by the Gaussian distribution. In the circumstances, by using the Bayes' theorem, the logarithm of evidence can be expressed as (Ching et al. 2006):

$$\begin{aligned}\ln\left[p(\mathcal{D}|\mathcal{M}_j)\right] &= \ln\left[\int_{\boldsymbol{\Theta}_j} p(\mathcal{D}|\boldsymbol{\theta}_j,\mathcal{M}_j)p(\boldsymbol{\theta}_j|\mathcal{M}_j)\mathrm{d}\boldsymbol{\theta}_j\right] = \\ &= \int_{\boldsymbol{\Theta}_j}\ln\left[p(\mathcal{D}|\boldsymbol{\theta}_j,\mathcal{M}_j)\right]p(\boldsymbol{\theta}_j|\mathcal{D},\mathcal{M}_j)\mathrm{d}\boldsymbol{\theta}_j - \int_{\boldsymbol{\Theta}_j}\ln\left[\frac{p(\boldsymbol{\theta}_j|\mathcal{D},\mathcal{M}_j)}{p(\boldsymbol{\theta}_j|\mathcal{M}_j)}\right]p(\boldsymbol{\theta}_j|\mathcal{D},\mathcal{M}_j)\mathrm{d}\boldsymbol{\theta}_j\end{aligned} \tag{6}$$

The first term in Eq. (6) is a measure of the average data-fit of the model class $\mathcal{M}_j$, accounting for the log-goodness of fit for different combinations of the parameters weighted by the posterior PDF. The second term is the Kullback–Leibler information, which is a non-negative measure of the information gain about $\mathcal{M}_j$ from the data $\mathcal{D}$. If the selection of a model class is solely determined by the data-fit term, then more complex models will usually be preferred over simpler ones. This often leads to over-fitting of the data and the updated model depending too much on the details of the specific data will be unreliable. The combination of these two factors in the log evidence for $\mathcal{M}_j$ provides a mathematically rigorous and robust way to builds in a trade-off between the data-fit of the model class and its information-theoretic complexity. The MCMC technique is employed to calculate the integration in Eq. (6).

3 Numerical case studies

A two-storey plane frame with standard 10# I-steel columns and beams is employed for verification. The span of the frame is 1.6 m and the height is 1.15 m for each storey. The sectional and material properties are: Young's modulus $E = 2.01\times10^{11}$ N/m^2, cross-sectional area $A = 1.57 \times 10^{-3}$ m^2, moment of inertia $I = 2.21 \times 10^{-6}$ m^4, and material density $\rho = 7.58 \times 10^3$ kg/m^3. The column-base and beam-column connections of the frame are treated as semi-rigid with rotational stiffness. The frame is discretized by plane beam elements into a FE model of 30 elements and 30 nodes with 84 DOFs. It's noted that these semi-rigid connections are simulated by beam elements of very short length with smaller flexural rigidity as compared to the regular beam and column components. The rotational stiffness of the semi-rigid connection is quantified by the flexural stiffness of the short beam element, and it is considered as uncertain modeling parameter to be identified.

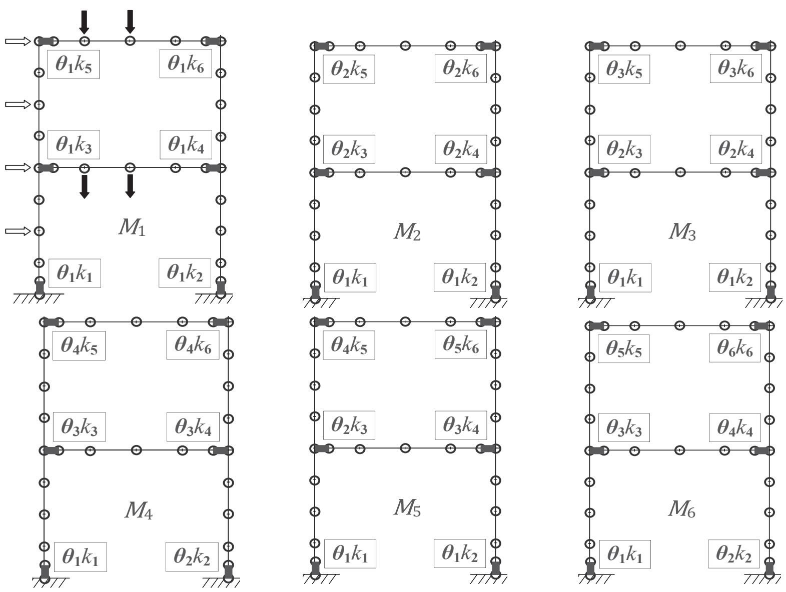

Figure 1. Different classes of FE parameterized models considered.

Table 1. Cases considered for numerical simulations.

	Case description
Case 1	$k_1 = k_2 = k_{cb}$, $k_3 = k_4 = k_5 = k_6 = k_{bc}$
Case 2	$k_1 = k_2 = k_{cb}$, $k_4 = k_5 = k_6 = k_{bc}$, $k_3 = 0.7k_{bc}$
Case 3	$k_2 = k_{cb}$, $k_3 = k_4 = k_5 = k_6 = k_{bc}$, $k_1 = 0.7k_{bc}$

There are six classes of models considered for this portal frame as shown in Figure 1. The complexity of model parameterization increases gradually from $\mathcal{M}_1$ to $\mathcal{M}_6$. For instance, $\mathcal{M}_1$ with one parameter θ_1 to scale all six rotational stiffness has the lowest degree of complexity among all considered model classes. $\mathcal{M}_2$ is a little bit more complex than $\mathcal{M}_1$, and it has two parameters θ_1 and θ_2 to scale the rotational stiffness of all the column-base and beam-column connections, respectively. $\mathcal{M}_6$ is the class of models with the most complex parameterization among all six model classes, and there are six scalar parameters in this class of models to update all column-base and beam-column connections separately.

Table 2. Results of model class selection for each numerical case.

		$p(\mathcal{M}_j \mid \mathcal{D})$	$\ln[p(\mathcal{D} \mid \mathcal{M}_j)]$	Data matching	Information gain
Case 1	$\boldsymbol{\mathcal{M}_1}$	**72.01%**	**83.45**	**86.01**	**2.56**
	$\mathcal{M}_2$	13.26%	81.75	85.63	3.88
	$\mathcal{M}_3$	3.02%	80.27	85.04	4.77
	$\mathcal{M}_4$	3.94%	80.54	85.16	4.62
	$\mathcal{M}_5$	3.73%	80.49	84.90	4.41
	$\mathcal{M}_6$	4.04%	80.56	84.92	4.36
Case 2	$\mathcal{M}_1$	0.05%	72.60	75.43	2.83
	$\mathcal{M}_2$	0.34%	74.48	78.46	3.98
	$\boldsymbol{\mathcal{M}_3}$	**34.63%**	**79.11**	**84.56**	**5.45**
	$\mathcal{M}_4$	26.43%	78.84	84.37	5.53
	$\mathcal{M}_5$	15.99%	78.34	84.57	6.23
	$\mathcal{M}_6$	22.56%	78.68	84.30	5.62
Case 3	$\mathcal{M}_1$	2.84%	77.84	80.72	2.88
	$\boldsymbol{\mathcal{M}_2}$	**49.43%**	**80.70**	**85.01**	**4.31**
	$\mathcal{M}_3$	9.26%	79.02	84.55	5.53
	$\mathcal{M}_4$	11.16%	79.21	84.66	5.45
	$\mathcal{M}_5$	13.81%	79.42	84.32	4.90
	$\mathcal{M}_6$	13.50%	79.40	84.45	5.05

There are eight sensors utilized for this frame as shown in Figure 1(a). The first four sensors (denoted as hollow arrow) are used to measure the horizontal vibration of the left column, while the latter four (denoted as solid arrow) are employed to monitor the vertical motion of the two beams. The measurement includes the natural frequencies and partial mode shapes at the measured DOFs of the first four modes. The considered noise level is 1% and 10% for the eigenvalues and partial mode shapes, respectively. It's assumed that there are 30 sets of repeated data obtained. There are three cases considered as showed in Table 1. k_{cb} and k_{bc} denote the nominal or baseline values of rotational stiffness of column-base and beam-column connections, and they are 0.03 and 0.06 times of the flexural rigidity of the beam and column elements, respectively. Case 1 is the nominal case, where all the modeling parameters are taken as their baseline values. In Case 2, the rotational stiffness of left beam-column connection of the first storey is set to be 0.7 times of original values. Based on Case 2, the semi-rigid stiffness of the left column-base connection is further reduced by 0.3 in Case 3.

A total of $1{\times}10^4$ samples are obtained by the MCMC sampling algorithm in which the samples within the 'burn-in' period are further removed before using Eq. (6). The obtained results are listed in Table 2, where the results corresponding to the optimal model class are highlighted in bold. It's noted that the probability of models $\mathcal{M}_j$ conditional on the data $\mathcal{D}$, i.e., $p(\mathcal{M}_j|\mathcal{D})$ is obtained by the method in Yuen (2010). As for Case 1, it's very clear that the most plausible model class is $\mathcal{M}_1$, the parameterization scheme of which is the most simple one. This implies that in the nominal status, only one parameter is enough for updating the FE model of the frame with the present data. For Case 2, it is found from Table 2 that the most plausible class of models is $\mathcal{M}_3$. This is not surprised since it's well anticipated that additional parameters are needed to characterize the local reduction of the rotation stiffness of the left beam-column connection. In addition, it's shown that, with the increase of parameterization complexity of the model class, there is more information needed to be extracted from the available data for updating these excessive parameters. It's reflected by the relatively larger value of information gain, which, conversely, penalizes the complexity of model parameterization. For instance, in Case 2, the model complexities of $\mathcal{M}_4$ and $\mathcal{M}_5$ are relatively higher than that of $\mathcal{M}_3$, rendering the value of information gain to be larger. Furthermore, it should be pointed out that since there is no difference between the rotational stiffness of beam-column connections in the first and second storey in the model class $\mathcal{M}_2$, the most plausible model class $\mathcal{M}_3$ matches the data better, even though the degree of complexity of $\mathcal{M}_2$ is less than that of $\mathcal{M}_3$. The similar phenomenon can also be observed from Case 3, where the most plausible class of modes with suitable parameterization complexity is chosen.

4 Conclusions

This paper addresses the issues of choosing suitable parameterized models for FE model updating through the model-class selection procedure following the Bayesian evidence inference. The concept of information divergence is employed to quantify the amount of information needed to be extracted from the measured data for achieving a trade-off between the complexity of a prescribed model class and that of its corresponding information-theoretic interpretation. Numerical simulations of a two-story bolt-connected steel frame structure are utilized for demonstration. The obtained results show that the combination of data matching and information gain indexes provides an efficient and mathematically rigorous way to select the most plausible class of FE models with a relatively simple parameterization scheme being suitable for model updating.

Acknowledgments

The authors gratefully acknowledge the financial support provided by the National Natural Science Foundation of China (Grants No. 51778506, 51578260 and 51629801).

References

Beck, J. L. and Katafygiotis, L. S., Updating Models and Their Uncertainties. I: Bayesian Statistical Framework, *J. Eng. Mech.-ASCE*, 124(4), 455–461, 1998.

Ching, J., Muto, M. and Beck, J. L., Structural Model Updating and Health Monitoring with Incomplete Modal Data Using Gibbs Sampler, *Comput.-Aided Civ. Inf.*, 21(4), 242–257, 2006.

Datta, B. N., Finite Element Model Updating, Eigenstructure Assignment and Eigenvalue Embedding Techniques for Vibrating Systems, *Mech. Syst. Signal Pr.*, 16(1), 83-96, 2002.

Friswell, M. I. and Mottershead, J. E., *Finite element model updating in structural dynamics*, Kluwer Academic Publishers, 1995.

Mthembu, L., Marwala, T., Friswell, M. I. and Adhikari, S., Model Selection in Finite Element Model Updating Using the Bayesian Evidence Statistic, *Mech. Syst. Signal Pr.*, 25(7), 2399-2412, 2011.

Papadimitriou, C., Beck, J. L. and Katafygiotis, L. S., Asymptotic Expansions for Reliability and Moments of Uncertain Systems, *J. Eng. Mech.-ASCE*, 123(12), 1219-1229, 1997.

Yuen, K. V., *Bayesian Methods for Structural Dynamics and Civil Engineering*, John Wiley & Sons, Ltd, 2010.