Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 312 (2008) 694–712
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Bayesian wavelet method for multivariate model assessment of dynamic systems

Xiaomo Jiang*, Sankaran Mahadevan

Department of Civil and Environmental Engineering, Vanderbilt University, Box 1831-B, Nashville, TN 37235, USA

Received 30 August 2006; received in revised form 31 October 2007; accepted 7 November 2007
Available online 26 December 2007

Abstract

In this paper an energy-based Bayesian wavelet method is presented for validation assessment of multivariate predictive models under uncertainty, using time-series data collected from a dynamic system. Time history data are decomposed into multiple time–frequency resolutions using a discrete wavelet packet transform method. As a signal feature, wavelet packet component energy is computed in terms of the decomposed coefficients. The effectiveness of the selected feature is assessed using both cross-correlation and cross-coherence metrics. A generalized likelihood ratio is derived as a quantitative validation metric based on Bayes' theorem and Gaussian distribution assumption of errors of the wavelet packet component energy between validation data and model prediction. The multivariate model is then assessed based on the Bayesian point and interval hypothesis testing approaches. The probability density function of the likelihood ratio is constructed using the statistics of multiple response quantities and Monte Carlo simulation. The proposed methodology is implemented in the validation of a structural dynamics problem, using multivariate time-series data sets. Sensitivity analysis is also performed to investigate the effect of parameter selection on the model validation decision.
Published by Elsevier Ltd.

1. Motivation

Within the context of model validation, model outputs are compared with experimental observations in order to quantitatively assess the validity or predictive capabilities of computational models. The uncertainties in the input variables are generally propagated either through the computational model to the model output or through the experiment to the measured output of a quantity of interest. This paper will focus on the development of quantitative approaches for multivariate model assessment under uncertainty using time-series data collected from dynamic systems.

Developing quantitative methods for model validation under uncertainty has become a problem of considerable research interest in recent years. For example, the fundamental concepts and methodologies for validation of large-scale computational models are being investigated by several government agencies such as the United States Department of Defense [1], American Institute of Aeronautics and Astronautics [2], Accelerated Strategic Computing Initiative (ASCI) program of the United Sates Department of Energy [3],

*Corresponding author. Tel.: +1 615 322 8633; fax: +1 615 322 3365.
E-mail addresses: xiaomo.jiang@vanderbilt.edu (X. Jiang), sankaran.mahadevan@vanderbilt.edu (S. Mahadevan).

0022-460X/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.jsv.2007.11.025

and American Society of Mechanical Engineers Council [4]. Detailed discussions of model verification and validation concepts can be found in many research articles (see e.g. Refs. [5–14]). Particularly, a model validation workshop was organized in 2006 by Sandia (http://www.esc.sandia.gov/VCWwebsite/vcwhome.html) consisting of researchers from the engineering, statistical, and mathematical fields to discuss and exchange ideas on quantitative model validation. Three validation challenge problems were developed in three disciplinary analyses, namely, thermal [15], structural mechanics [16], and structural dynamics [17]. Various solution techniques for these three problems were provided by researchers; these are compiled in a special issue of the *Computer Methods in Applied Mechanics and Engineering* journal ([18]). This paper uses the structural dynamics problem to illustrate the effectiveness of the proposed quantitative approach.

Recently, statistical hypothesis testing-based techniques have been proposed for assessment of structural dynamics to determine whether the model predicted output matches with the real system output (e.g. Refs. [19–21]). These methods consist of two main elements: feature extraction and quantitative assessment. The objectives of feature extraction are to reduce the dimensionality of the data used for model assessment, and to improve the efficiency and accuracy of model validation. Different features have been used in model validation of dynamic systems, such as root mean square, principal component decomposition, and frequency response functions. Refer to Hemez and Doebling [20] for details of various features extracted from time history data for model validation purpose. In this study, wavelet packet component energy is extracted from a given signal as feature for model validation of dynamic systems.

Statistical hypothesis testing is one approach to quantitative model validation under uncertainty, and both classical and Bayesian statistics have been explored (see Oberkampf and Barone [13] for a comprehensive state-of-the-art review). Classical hypothesis testing is a well-developed statistical method for accepting or rejecting a model based on an error statistic (see e.g. Refs. [13,19,22–26]). The second author and his research associates have developed Bayesian methods for determining the predictive capability of computational models [27–32]. One important difference between the Bayesian and classical hypothesis testing approaches is that, the Bayesian approach focuses on model acceptance whereas classical hypothesis testing focuses on model rejection. Particularly, Rebba [32] compared the Bayesian methods with other statistical validation metrics and approaches in terms of accuracy and adequacy requirements, ease of implementation, etc. In particular, Bayesian hypothesis testing-based methods [27] and Bayesian risk-based decision making methodology [29] have been explored for model assessment. These metrics have been investigated with various model validation problems using limited amount of experimental data. Refer to Mahadevan and Rebba [28] and Jiang and Mahadevan [29] for details of the validation metrics and their applications.

The focus of this paper is to pursue a new Bayesian wavelet method for multivariate model validation using time-series data collected from dynamic systems. The wavelet packet component energy, a wavelet packet-based feature, is explored for model validation in this paper. Wavelets provide an effective and efficient approach to obtain a multi-resolution representation of a signal. This representation provides a hierarchical framework for interpreting the information context in a signal. At different resolutions, the details of a signal characterize different physical structures of the scene. At a fine resolution, these details correspond to the transient changes which provide the signal "context". For given wavelet basis and decomposition level, the wavelet transform of a signal has been demonstrated to be unique and invariant [33–35]. Coifman and Wickerhauser [36] proposed the wavelet packet transform analysis to allow for a finer and adjustable resolution in the high frequencies (details). Compared with conventional wavelet transform methods, the wavelet packet transform method is a more effective approach to extract features from either stationary or non-stationary signals to represent the underlying dynamic systems [33,37].

The wavelet packet component energy measures the signal energy content contained in some specific frequency band. In the model validation problem, we have only one specific dynamic system to be validated; therefore the relationship between the wavelet packet component energy features and the dynamic system is unique for this particular problem. Since the wavelet packet component energy values extracted from the decomposed signals are unique, they can be used as signal features to represent system characteristic for model validation purposes. Recently, the wavelet packet component energy-based method has been demonstrated to be an effective feature representation of a signal in the context of structural system identification [37], traffic flow analysis [38], condition monitoring of dynamic systems [39], structural damage assessment [40], and

incident detection in traffic patterns [41]. The wavelet packet component energy-based approach is therefore used in this research for dynamical model validation.

In the context of feature-based model assessment, two important issues need to be addressed. The first issue is how to assess the key features. Feature extraction will inevitably result in the loss of information from the original time series. The time-series data collected from a dynamic system typically contain useful information and disturbing noise. Therefore, it is desirable to extract features that capture the characteristics of dynamic system (information) and to separate these features that represent the disturbances (noise). Various approaches (e.g. Refs. [42–48]) have been proposed to assess the selected features. For example, Chen et al. [43] used the autocorrelation function of the residuals between the original and reconstructed signals in the time domain. Cho et al. [44] used the coherence spectra of the residuals in the frequency domain. Wu and Du [47] used various assessment criteria in both time and frequency domains. Yan and Gao [48] used a distance measure-based assessment method. In this paper, both the *cross-correlation* in time domain and the *cross-coherence* in frequency domain between the original and reconstructed signals are used as the feature selection rule.

The second issue to be considered is how to assess the overall validity of multivariate time-series predictive models? A single response quantity may be predicted and observed at different spatial and temporal points, thus multiple response quantities need to be compared in the process of model validation. Multiple univariate comparisons might yield conflicting inferences for individual quantities and also do not consider correlation [32]. Therefore, another focus of this study is to explore a Bayesian wavelet-based probabilistic method for overall validation assessment of a multivariate model. A point null-based hypothesis testing Bayes factor method was derived earlier in Jiang and Mahadevan [49] for multivariate model validation. This paper derives a more practical interval-based hypothesis testing method for this purpose.

In the following sections, the discrete wavelet packet transform method is first introduced briefly, with the purpose of decomposing both measured and predicted time-series signals into different levels of wavelet coefficients. These coefficients represent the multiple time–frequency resolutions of a signal. For given different resolution levels, the wavelet component energy can then be computed, from the two data sets, in terms of the relevant wavelet and scaling function coefficients; this will be used as the signal feature. Interval Bayes factor-based validation metric is computed in terms of the difference of the wavelet energy between the experimental and predicted time-series data. The computational model is assessed using both point null and interval-based hypothesis testing approaches. The validation framework is demonstrated with a structural dynamics problem developed at Sandia National Laboratories [17].

2. Wavelet signal processing

2.1. Wavelet packet decomposition

The wavelet packets (or bases) are formed by linear combinations of wavelet functions, denoted by $\psi_{j,k}(t)$. These wavelet functions are obtained from a basis function (also known as *mother* or generating wavelet) $\varphi(t)$ by simple scaling and translation in the dyadic form as follows [34,35]:

$$\psi_{j,k}(t) = 2^{-j/2}\varphi(2^{-j}t - k), \quad j,k \in Z, \quad \psi \in L^2(R), \tag{1}$$

where t represents a continuous time variable, k and j denote the time and the frequency indices, respectively, and Z is the set of all integers. The notation $L^2(R)$ represents the square summable real number space. The resulting wavelet packets inherit properties such as orthonormality and time–frequency localization from the wavelet basis function $\varphi(t)$.

In practical applications of the discrete wavelet packet transform decomposition, a given time series with N data points, $f(t)$ ($t = t_1, t_2, \ldots, t_N$), is simultaneously decomposed into a series of scaling coefficients, $s_j(k)$, and wavelet coefficients, $w_j(k)$. The time series can then be reconstructed by the inverse wavelet transform and the discrete wavelet packet transform coefficients as follows:

$$\bar{f}(t) = \sum_{k\in Z}\sum_{j\in Z}\left[s_j(k)\psi_{j,k}(t) + w_j(k)\varphi_{j,k}(t)\right], \tag{2}$$

where $\bar{f}(t)$ is the reconstructed time series, and the double summation indicates that the scaling and wavelet subspaces are simultaneously split into second-level subspaces to provide the frequency and time breakdown of the signal. Obviously, the component signal is a superposition of wavelet functions $\varphi_{j,k}(t)$ and scaling functions $\psi_{j,k}(t)$. The j level scaling coefficients $s_j(k)$ and wavelet coefficients $w_j(k)$ are obtained by a recursive way. Refer to Burrus et al. [50] for details regarding the computation of the coefficients. Next, the wavelet packet component energy will be calculated at each node in the last decomposition level (i.e. three in this study) for model assessment purpose.

2.2. Wavelet packet component energy

The advantage of using an orthonormal basis to decompose a signal by the discrete wavelet packet transform is that the total energy of the signal can be partitioned into its various time–frequency components. The energy contribution from each component is mathematically expressed as a function of the wavelet and scaling coefficients as follows [50]:

$$E^f = \int_{-\infty}^{\infty} f^2(t)\,\mathrm{d}t = \sum_{m=1}^{2^j}\sum_{n=1}^{2^j}\int_{-\infty}^{\infty}\left|f_j^m(t)\right|\left|f_j^n(t)\right|\mathrm{d}t. \tag{3}$$

Using the orthogonal condition of wavelet functions, Eq. (3) becomes

$$E^f = \sum_{i=1}^{2^j} E_i^f = \sum_{i=1}^{2^j}\int_{-\infty}^{\infty}\left[f_j^i(t)\right]^2\mathrm{d}t, \tag{4}$$

where $f_j^i(t)$ is the discrete wavelet packet transform coefficient or component signal [$s_j(k)\psi_{j,k}(t)$ or $w_j(k)\varphi_{j,k}(t)$ in Eq. (2)], and E_i^f is the component energy stored in the component signal.

2.3. Feature assessment

We usually select the main energy components as features for model validation purpose. Now we need evaluate whether the selected components can be used to effectively represent the original signal. In this paper, the effectiveness of the selected feature packets is evaluated based on the reconstructed signal using two criteria: cross-correlation and cross-coherence.

The *cross-correlation* between the reconstructed signal $\bar{f}(t)$ (i.e., Eq. (2)) and the original signal $f(t)$ is used to measure the similarity of the two signals in the time domain. Following the definition of Wu and Du [47], the cross-correlation is computed as follows:

$$r(k) = E[\bar{f}(t)f(t-k)]\big/E[f(t)f(t-k)], \quad k = 0, 1, \ldots, N-1, \tag{5}$$

where the variable k is the time delay, the nominator is the cross-covariance between $\bar{f}(t)$ and $f(t)$, and the denominator is only the variance of $f(t)$.

The definition in Eq. (5) is slightly different from that in statistics where both the variances of $\bar{f}(t)$ and $f(t)$ are used in the denominator. Thus, the cross-correlation defined in Eq. (5) can be used to compare various constructed signals for the same original signal. If there is no cross-correlation between the two signals, then the resulting cross-correlation value is constantly equal to zero, which is obviously not desirable. On the other hand, if $\bar{f}(t)$ is equal to $f(t)$, then their cross-correlation value is constantly equal to one. This case is also not desirable since it implies that the reconstructed signal completely represent the original one, including both the useful information and the noise. The desirable result is that the cross-correlation value approaches units in the first few k, which implies that both $\bar{f}(t)$ and $f(t)$ have the same principle components.

The *cross-coherence* between $\bar{f}(t)$ and $f(t)$ is used to measure the similarity of the two signals in the frequency domain, which is defined as

$$c(w) = S_{\bar{f}f}(\omega)\Big/S_f(\omega), \tag{6}$$

where $S_{\bar{f}f}(\omega)$ is the cross-power-spectrum between $\bar{f}(t)$ and $f(t)$, and $S_f(\omega)$ is the power spectrum of $f(t)$. The well-known Welch method [51] is used to estimate the power spectra density because it provides a smoothed spectral density estimate. Similar to the cross-correlation defined in Eq. (5), the definition of cross-coherence in Eq. (6) is slightly different from that in statistics. Again, the cross-coherence being close to one but not equal to one implies that the best selected features contain the useful information on the signal but not the noise.

3. Bayesian multivariate model validation

Let $\mathbf{Y}_{\text{exp}}$ and $\mathbf{Y}_{\text{pred}}$ represent the features extracted from experimental data and model prediction of the quantity of interest in the form of multivariate time series, respectively. Within the context of binary hypothesis testing for model validation, we need to test two hypotheses H_0 and H_1, i.e., the null hypothesis (H_0: $\mathbf{Y}_{\text{exp}} = \mathbf{Y}_{\text{pred}}$) to accept the model and an alternative hypothesis (H_1: $\mathbf{Y}_{\text{exp}} \neq \mathbf{Y}_{\text{pred}}$) to reject the model. Thus, the likelihood ratio, referred to as the Bayes factor, is calculated using Bayes' theorem as [52]

$$B_{01} = \frac{f(\text{Data}|H_0)}{f(\text{Data}|H_1)}. \tag{7}$$

Since B_{01} is non-negative, the value of B_{01} may be converted into the logarithm scale for convenience of comparison over a large range of values, i.e., $b_{01} = \ln(B_{01})$, where ln(.) is a natural logarithm operator with a basis of e. Kass and Raftery [52] suggest interpreting b_{01} between 0 and 1 as weak evidence in favor of H_0, between 3 and 5 as strong evidence, and $b_{01} > 5$ as very strong evidence. Negative b_{01} of the same magnitude is said to favor H_1 by the same amount.

3.1. Point null hypothesis-based method

Recently, Jiang and Mahadevan [49] derived a generalized likelihood ratio based on Bayes' theorem, assuming that the likelihood function of the difference between the validation data and the model prediction, $\mathbf{D}$, is a Gaussian function. In the current paper, instead of directly assessing the difference between the two sets of time series, the difference of selected features (i.e., wavelet packet component energy E^f calculated using Eq. (4)) between the experimental and predicted time series is used for model assessment. Assuming that $\mathbf{D} = \begin{bmatrix} \mathbf{d}_1 & \mathbf{d}_2 & \dots & \mathbf{d}_m \end{bmatrix}^{\mathrm{T}}$ is an $m \times n$ matrix representing m variables, each $\mathbf{d}_i = [\, d_{i1} \quad d_{i2} \quad \dots \quad d_{in} \,]$ ($i = 1, 2, \dots, m$) representing the n discrepancy values of the ith variable E_i^f, i.e. $d_{ij} = E^f_{ij,\text{exp}} - E^f_{ij,\text{pred}}$. We do not know the distribution of the difference a priori, so we assume Gaussian only as an initial guess, and then do a Bayesian update. The variance of difference is estimated using the difference of energy features between the experimental data and prediction output. Usually the matrix $\mathbf{D}$ is assumed to follow a multivariate normal density $N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where the vector $\boldsymbol{\mu} = \mathbf{E}[\mathbf{d}]$ represents the corresponding m mean values and $\boldsymbol{\Sigma} = \mathbf{E}[(\mathbf{d} - \boldsymbol{\mu})(\mathbf{d} - \boldsymbol{\mu})']$ is an $m \times m$ covariance matrix of all variables.

Let $\bar{d}_i = (1/n)\sum_{j=1}^{n} d_{ij}$ ($i = 1, 2, \dots, m$) be the mean of n validation data points of the ith variable d_i. Within the context of multivariate model validation, we wish to test whether the sets of means, $\bar{\mathbf{D}} = \begin{bmatrix} \bar{d}_1 & \bar{d}_2 & \dots & \bar{d}_m \end{bmatrix}^{\mathrm{T}}$, are equal to zeros, $\mathbf{E}_0 = [\, 0 \quad 0 \quad \dots \quad 0 \,]^{\mathrm{T}}$ (an m-dimensional zero vector). Thus, the multivariate model validation problem becomes testing the two hypotheses H_0: $\boldsymbol{\mu} = \mathbf{E}_0$ versus H_1: $\boldsymbol{\mu} \neq \mathbf{E}_0$ with $\boldsymbol{\mu}|H_1 \sim N(\boldsymbol{\rho}, \boldsymbol{\Lambda})$. The Bayes factor B_M is derived for Gaussian prior and likelihood of the difference as [49]:

$$B_M = \left(\frac{n|\boldsymbol{\Lambda}| + |\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}|} \right)^{1/2} \exp\left\{ \frac{n}{2} \left[(\bar{\mathbf{D}} - \boldsymbol{\rho})'(n\boldsymbol{\Lambda} + \boldsymbol{\Sigma})^{-1}(\bar{\mathbf{D}} - \boldsymbol{\rho}) - \bar{\mathbf{D}}'\boldsymbol{\Sigma}^{-1}\bar{\mathbf{D}} \right] \right\}. \tag{8}$$

Three issues regarding the point null hypothesis testing method need to be addressed here. First, we need to select the covariance matrix $\boldsymbol{\Lambda}$ in the Bayesian hypothesis testing. Based on the Fisher information method [52], a proper selection may be made on the assumption that the amount of information in the prior is equal to that in the observation. Thus, the parameters $\boldsymbol{\rho} = \mathbf{E}_0$ and $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}$ are used in this study, following the suggestion in Migon and Gamerman [53] for the univariate case. Furthermore, sensitivity analysis will be performed in the example presented in this paper to investigate the effect of $\boldsymbol{\Lambda}$ on b_M.

Second, assume that $\pi_0 = \pi_1 = 0.5$ in the absence of prior knowledge of each hypothesis before testing. The confidence in the model based on the validation data can be obtained by [28]

$$\kappa = Pr(\mathrm{H}_0|\mathrm{Data}) = B_M/(B_M + 1). \tag{9}$$

Obviously, from Eq. (9), $B_M \to 0$ indicates 0% confidence in accepting the model, and $B_M \to \infty$ indicates 100% confidence.

Third and finally, the current work is based on the error Gaussian (or normality) assumption. In the case of non-normality, various transformation methods (see e.g., Refs. [30,54]) are available to achieve normality of error data. The transformed data can then be used in the proposed Bayesian decision methodology for model validation.

3.2. Interval hypothesis-based method

The interval-based hypothesis testing method has been demonstrated to provide more consistent model validation results than a point hypothesis testing method [32]. In Appendix A of this paper, we derive an explicit expression to calculate the Bayes factor based on interval-based hypothesis testing for *multivariate* model validation, with the purpose of facilitating the overall validation assessment of computational models.

Within the context of binary hypothesis testing for multivariate model validation, the Bayesian formulation of interval-based hypotheses is represented as H_0: $|\mathbf{D}-\mathbf{E}_0| \leqslant \boldsymbol{\varepsilon}$ versus H_1: $|\mathbf{D}-\mathbf{E}_0| > \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon}$ is a predefined threshold vector. Here we are testing whether the difference $\mathbf{D}$ is within an allowable limit ε. Assuming that the difference, $\mathbf{D}$, has a probability density function under each hypothesis, i.e., $\mathbf{D}|\mathrm{H}_0 \sim f(\mathbf{D}|\mathrm{H}_0)$ and $\mathbf{D}|\mathrm{H}_1 \sim f(\mathbf{D}|\mathrm{H}_1)$. Again, we do not know the distribution of the difference a priori, so we assume Gaussian only as an initial guess, and then do a Bayesian update. We assume that (1) the difference $\mathbf{D} = \{d_1, d_2, \ldots, d_n\}$ follows a multivariate normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with the covariance matrix $\boldsymbol{\Sigma}$ estimated from the validation data, and (2) a prior density function of $\boldsymbol{\mu}$ under both null and alternative hypotheses, denoted by $f(\boldsymbol{\mu})$, is taken to be $N(\boldsymbol{\rho}, \boldsymbol{\Lambda})$. Following the discussion in the previous section for Bayesian point null hypothesis testing, the parameters $\boldsymbol{\rho} = \mathbf{E}_0$ and $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}$ are used in the following derivation as well. Using Bayes' theorem, $f(\boldsymbol{\mu}|\mathbf{D}) \propto f(\mathbf{D}|\boldsymbol{\mu}) f(\boldsymbol{\mu})$, the Bayes factor for the multivariate case, B_{iM}, is equivalent to the volume ratio of the posterior density of $\boldsymbol{\mu}$ under two hypotheses, expressed as follows:

$$B_{iM} = \frac{\int_{-\boldsymbol{\varepsilon}+\mathbf{e}_0}^{\boldsymbol{\varepsilon}+\mathbf{e}_0} f(\boldsymbol{\mu}|\mathbf{D})\,\mathrm{d}\boldsymbol{\mu}}{\int_{-\infty}^{-\boldsymbol{\varepsilon}+\mathbf{e}_0} f(\boldsymbol{\mu}|\mathbf{D})\,\mathrm{d}\boldsymbol{\mu} + \int_{\boldsymbol{\varepsilon}+\mathbf{e}_0}^{\infty} f(\boldsymbol{\mu}|\mathbf{D})\,\mathrm{d}\boldsymbol{\mu}} = \frac{K}{1-K}, \tag{10}$$

where the multivariable integral of $K = \int_{-\boldsymbol{\varepsilon}+\mathbf{e}_0}^{\boldsymbol{\varepsilon}+\mathbf{e}_0} f(\boldsymbol{\mu}|\mathbf{D})\,\mathrm{d}\boldsymbol{\mu}$ is calculated using Eq. (A.4) in Appendix A. Note that the quantity K in Eq. (A.4) in Appendix A is dependent on the value of $\boldsymbol{\varepsilon}$. The decision maker or model user has to decide what $\boldsymbol{\varepsilon}$ are acceptable.

In this study, for the sake of comparison with the point method, the values of $\boldsymbol{\varepsilon}$ are taken to be 0.5 times of the standard deviations of the multiple variables in the numerical example. Similar to the point method, sensitivity analysis is performed in the numerical example in Section 3 to investigate the effect of the selection of $\boldsymbol{\varepsilon}$ on the Bayes factor value B_{iM}. Again, the confidence of accepting the model based on Eq. (10) can be estimated using Eq. (9) with B_{iM} in place of B_M.

3.3. Probabilistic model validation

In order to consider the uncertainties in both validation data and model prediction, Bayes factor, B_M or B_{iM}, is treated as a random variable. Given a fixed sample size n and the variable statistics of response quantity (i.e., probability density functions of $E_{i,\mathrm{exp}}$ and $E_{i,\mathrm{pred}}$), various values of $\bar{\mathbf{D}}$ are obtained by using Monte Carlo simulation technique to produce a distribution of the Bayes factor. As such, the probability of accepting the model can be estimated by finding the proportion of the Bayes factor whose values are greater than η, i.e., $\gamma = \mathbf{Pr}(B_M > \eta)$. The probabilistic method provides a quantitative measure for the overall validation assessment of the multivariate computational model.

4. Numerical application

4.1. Problem description

A structural dynamics problem provided by Sandia National Laboratories [17] is used in this study to investigate the effectiveness of the proposed Bayesian wavelet method for multivariate model validation, using time series data. This challenge problem was generated to simulate the real scenarios existing in structural dynamics, which consists of three parts: (1) a simple subsystem, (2) a more complex accreditation system, and (3) a target application. The subsystem is not fully relevant to the target application. It is used to isolate important phenomena in target application such as nonlinearity. The accreditation system is more expensive to prototype and to be tested, but it is more relevant to the target application. The target application provides both a decision context and an accompanying regulatory requirement. Refer to Red-Horse and Paez [17] for details about this challenge problem.

In this research, the experimental data virtually generated from a three-mass subsystem (Fig. 1) for input characterization and validation are used to illustrate the proposed validation methodology. This subsystem is mounted on a simply supported beam by an additional weakly nonlinear connection and only vertical motions are permitted. In addition, computational models are available to make simulation-based predictions of structural responses. All validation data used in this example have been virtually generated through solving dynamical equations of a structural system [17]. Several researchers provided different solutions to this validation challenge problem during a model validation workshop organized by Sandia in 2006 (http://www.esc.sandia.gov/VCWwebsite/vcwhome.html).

In the input characterization test, three random vibration excitations, namely, low-, medium-, and high-level, are individually applied to the mass m_1 (Fig. 1) one at a time. The acceleration responses at m_1, m_2, and m_3 in the vertical direction are obtained for every level of vibration excitation, resulting in three sets of time series data, each consisting of three acceleration responses for three masses separately. As an example, Fig. 2 shows the time series plots of (a) the medium-level random vibration input excitation and (b) the corresponding acceleration record at m_3 for the first subsystem or sample. The virtual test is repeated on 20 nominally identical subsystems (i.e., samples), resulting in 60 sets of response acceleration data. Since the three-mass subsystem has only three degrees of freedom, every set of response acceleration data is used to create 15 modal parameters via experimental modal analysis [17], namely, three modal frequencies, ω_i, three modal damping values, ζ_i, and 3×3 mode shape vector, Φ_i. Due to unit-to-unit variability within the population of components and possibly also nonlinearities in the subsystem, the created modal parameters of each subsystem are different. Therefore, the modal parameters may be considered as a 15-dimensional random vector.

In the validation test, a shock with free decay is applied to m_2 (Fig. 1) rather than the random vibration on m_1 in the input characterization test. As an example, Fig. 3 shows the time series plot of (a) the high-level

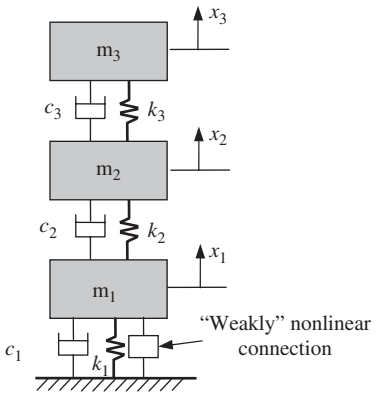

Fig. 1. Subsystem for input characterization and model validation: components with connection.

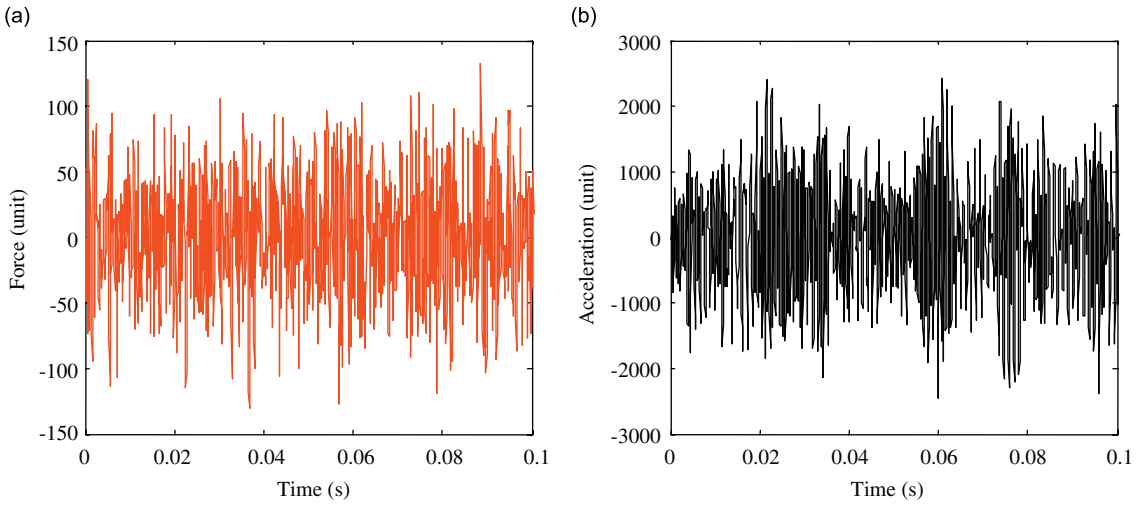

Fig. 2. Time series plots of medium-level random vibration excitation and acceleration response data of mass m_3 from 0 to 0.1 s for sample 1 (input characterization): (a) excitation and (b) acceleration response.

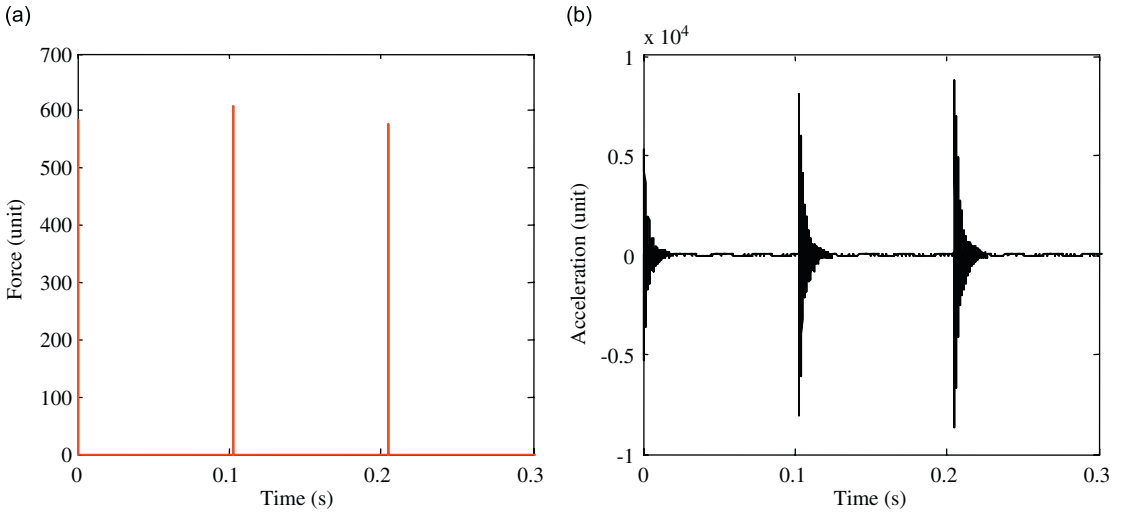

Fig. 3. Time series plots of high-level shock excitation and acceleration response data of mass m_1 from 0 to 0.3 s for sample 1 (model validation): (a) excitation and (b) acceleration response.

shock input excitation and (b) the corresponding acceleration record at m_1 for the first system or sample. Similar to the characterization case, 60 sets of response acceleration data are created, each generating 15 modal parameters. All 15 modal parameters as well as the corresponding random shock excitation are used as inputs to the computational model to predict the structural dynamic responses. For the sake of simplicity, two assumptions have been made in this problem [17]: (1) various configurations of the structural systems are completely known and therefore all uncertainties come from only the subsystem; and (2) all experimental measurements are perfect and therefore do not contain any noise.

4.2. Input characterization

The obtained 60 sets of data are explored to quantify the statistics of the 15 random input modal parameters using statistical–graphical approaches, including histogram, normal probability, and scatter plots.

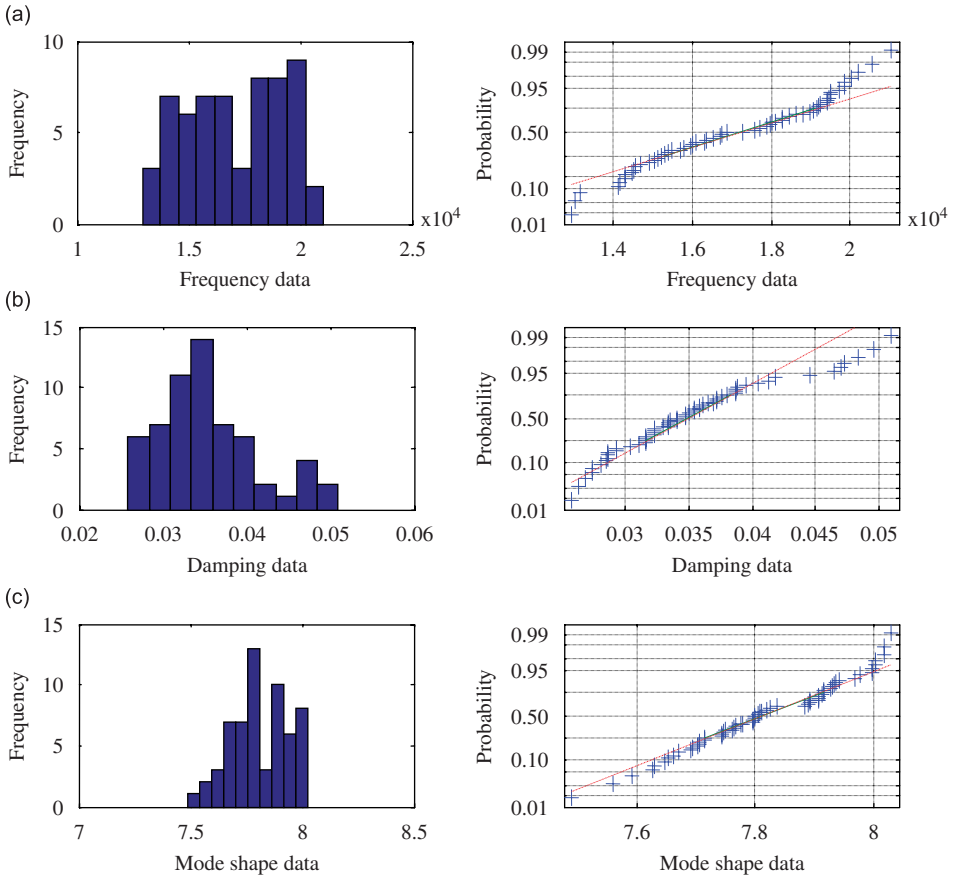

Fig. 4. Histogram and probability plots of input parameters: (a) third frequency (ω_3); (b) third damp (ζ_3); and (c) third mode shape (Φ_3).

Three phenomena are observed from the numerical results. First, not all input parameters follow a normal distribution. For demonstration purpose, Fig. 4 shows the histogram and normal probability plots of three input parameters: the third frequency (ω_3); the third damp (ζ_3); and the third mode shape (Φ_3). The normal probability plots show a reasonably linear pattern in the center of the data. However, the tails, particularly the upper tail, show departures from the fitted lines. These plots imply that a distribution other than normal distribution should be used for these data. Usually a Bayes factor-based model validation approach [27,28] requires sampling the input variables to predict model output. It is relatively difficult to sample the multiple variables with mixed distribution functions.

Second, the input variables are strongly correlated with each other. Scatter plots are drawn to find the relationship among various modal parameters. Fig. 5 shows the correlation matrix plots of three frequency parameters and two model shape parameters. Clearly the pair-wise relationships among the variables are linear. Therefore, the multiple input variables are strongly correlated, which makes sampling more complicated in the model validation.

Third and finally, the calibration and validation data may come from different populations. Fig. 6 shows the comparison of validation and calibration data for the first frequency and second mode shape parameters. The two sets of data fall in different regions, implying that their statistics may be different. As a result, it may be inaccurate to make use of the statistics of 15 modal parameters obtained from the input characterization data for model validation.

To sum up, it is difficult to apply the existing Bayes factor method directly to evaluate the predictive model. Since the uncertainties in the input variables are propagated either through the computational model or

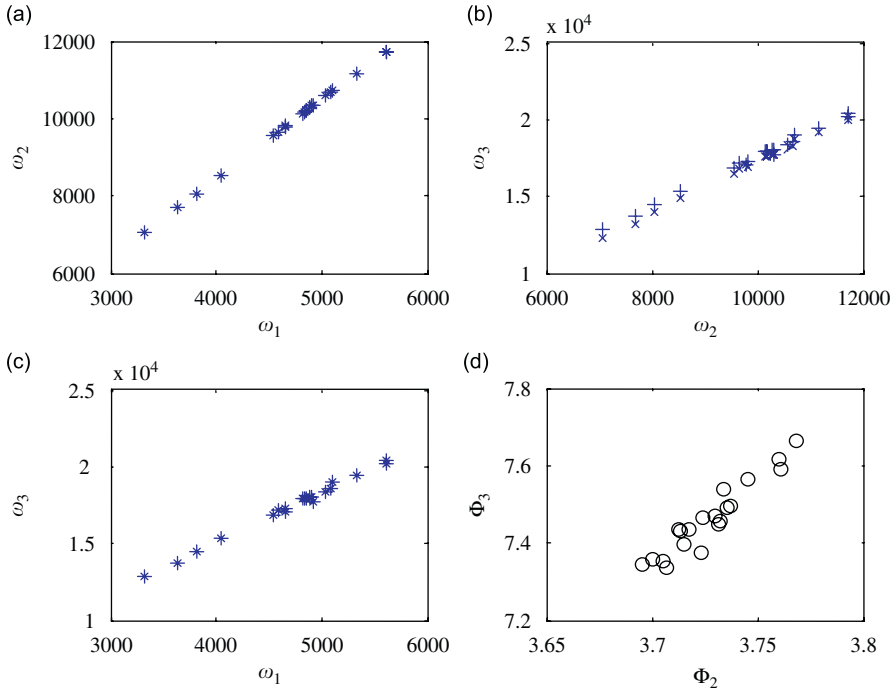

Fig. 5. Scatter plots of three frequency parameters and two mode shape parameters: (a) ω_1 vs. ω_2; (b) ω_2 vs. ω_3; (c) ω_1 vs. ω_3; and (d) Φ_2 vs. Φ_3.

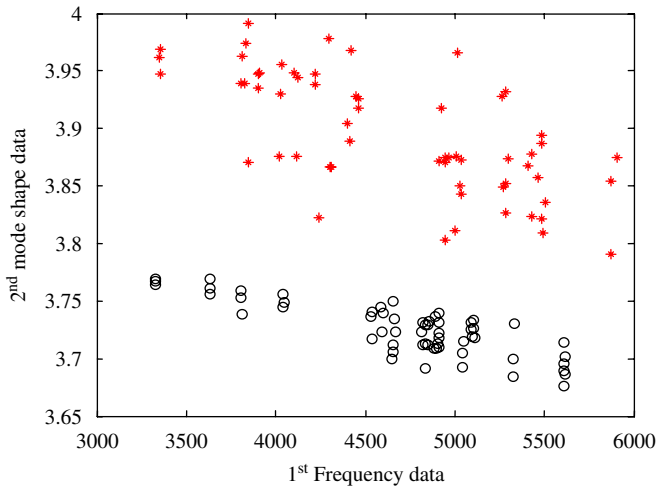

Fig. 6. Comparison of validation and calibration data for first frequency and second mode shape parameters: (∗) calibration data and (○) validation data.

through the virtual experiments to the response accelerations, the probabilistic assessment in the next section is based only on the response acceleration time histories, which is the merit of the proposed output-based Bayesian wavelet method.

4.3. Model validation

In this section, both predicted and experimental acceleration time histories data are decomposed using the three-level discrete wavelet packet transform method. Then wavelet packet component energy is computed as

the signal feature at every decomposition node at the third level. The principal energy contents which dominate the signal energy are used to construct a multivariate model validation problem. The effectiveness of the selected feature is assessed using both cross-correlation and cross-coherence metrics. This problem is handled by both point null and interval-based Bayesian hypothesis testing methods, followed by a probabilistic method for model validation.

4.3.1. Wavelet packet component energy

Using Daubechies wavelet of order 5, the three-level discrete wavelet packet transform decompositions are performed on all 60×3 sets of response acceleration time history data. Every set of data are resolved into eight series corresponding to the eight decomposition subspaces in the third level, and identified as AAA_3, which represents the third-level approximation coefficients (A_3) resulting from the second-level approximation (AA_2) to DDD_3, which represents the third-level details coefficients (D_3) resulting from the second-level details (DD_2). Thus, based on Eq. (2), the original data, $f(t)$, can be

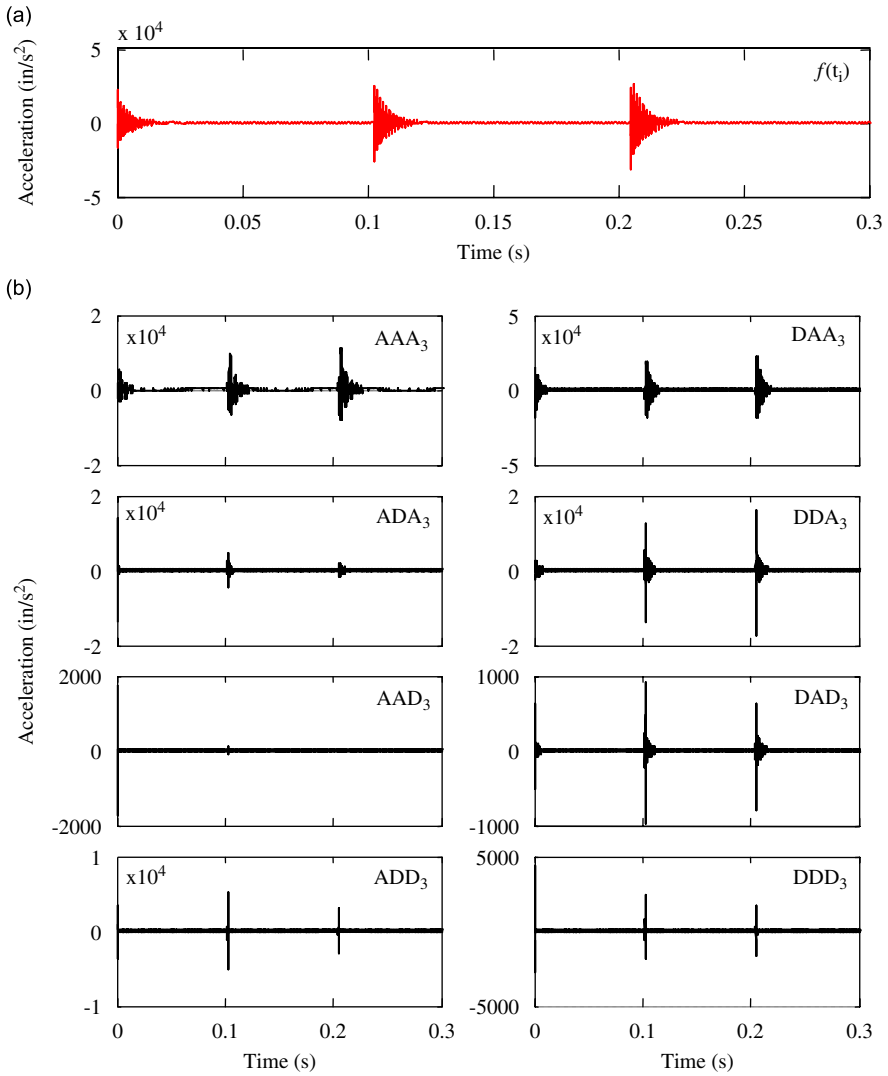

Fig. 7. Three-level discrete wavelet packet transform decomposition of experimental acceleration response data of m_2 for the first sample under high-level shock vibration using Daubechies wavelet of order 5: (a) original acceleration response; and (b) three-level decomposition of discrete wavelet packet transform.

represented mathematically by the third-level discrete wavelet packet transform decomposition coefficients as follows:

$$f(t) = AAA_3 + AAD_3 + ADA_3 + ADD_3 + DAA_3 + DAD_3 + DDA_3 + DDD_3. \quad (11)$$

Obviously, the component signal is a superposition of wavelet functions $\psi_{j,k}(t)$. As an example, Fig. 7 shows the discrete wavelet packet transform decomposition coefficients of experimental acceleration response data of m_2 under high-level shock vibration. The three shock peaks in the original response time history are clearly apparent in several decomposed coefficients. Therefore, the three-level decomposed coefficients effectively capture the features (e.g. peaks) in the original data. It should be noted that the scale of the vertical axis is chosen differently for different levels of details in Fig. 7b for the sake of visibility of the illustration.

Next, using Eq. (4), the component energy E_i^f is calculated for every set of decomposed coefficient in the last level, resulting in eight energy values from every set of response acceleration data. We sort the eight energy values in a descending way, and find that more than 99% signal energy is stored in the first four sorted wavelet packet component energy values. As an example, Fig. 8 shows the wavelet energy percentage of three-level discrete wavelet packet transform decomposition coefficients ($j = 3$) for both experimental and predicted acceleration response series of the 15th sample under low-level shock vibration.

The cross-correlation [$r(k)$ in Eq. (5)] and cross-coherence [$c(\omega)$ in Eq. (6)] between the original signal and the reconstructed signal using the first four wavelet packets are computed for both experimental data and model prediction of each mass. It is observed that both $r(k)$ and $c(\omega)$ values approach one for the first few k and ω, respectively, implying that the first four feature values (i.e., wavelet packet component energy) are enough to represent the original signal. As an example, Fig. 9 shows the assessment results of both experimental data and model prediction reconstructed using the first four wavelet packets for mass 3 of the 15th sample under low-level shock vibration. Thus only the first four feature values are used as the measure of a signal to compare the predicted and experimental time series, resulting in a 12-variable model validation problem for the three-mass subsystem.

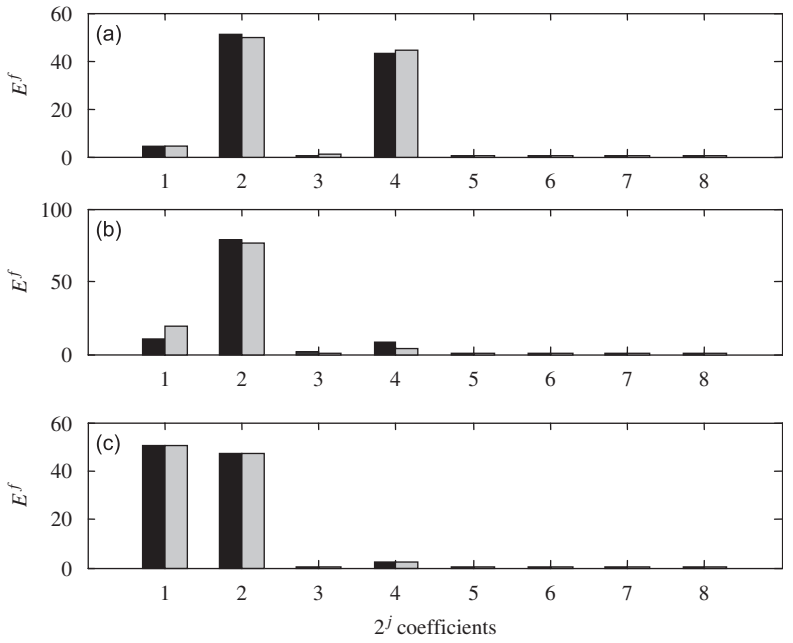

Fig. 8. Wavelet energy percentage of three-level discrete wavelet packet transform decomposition coefficients ($j = 3$) for both experimental and predicted acceleration response data of the 15th sample under low-level shock vibration using Daubechies wavelet of order 5 (Model validation): (a) mass 1: $E_4 = 99.14\%$; (b) mass 2: $E_4 = 99.56\%$; and (c) mass 3: $E_4 = 99.99\%$; ■ experimental data; and ▢ model prediction; E_4 = Sum of the first four wavelet energy values of predicted signal.

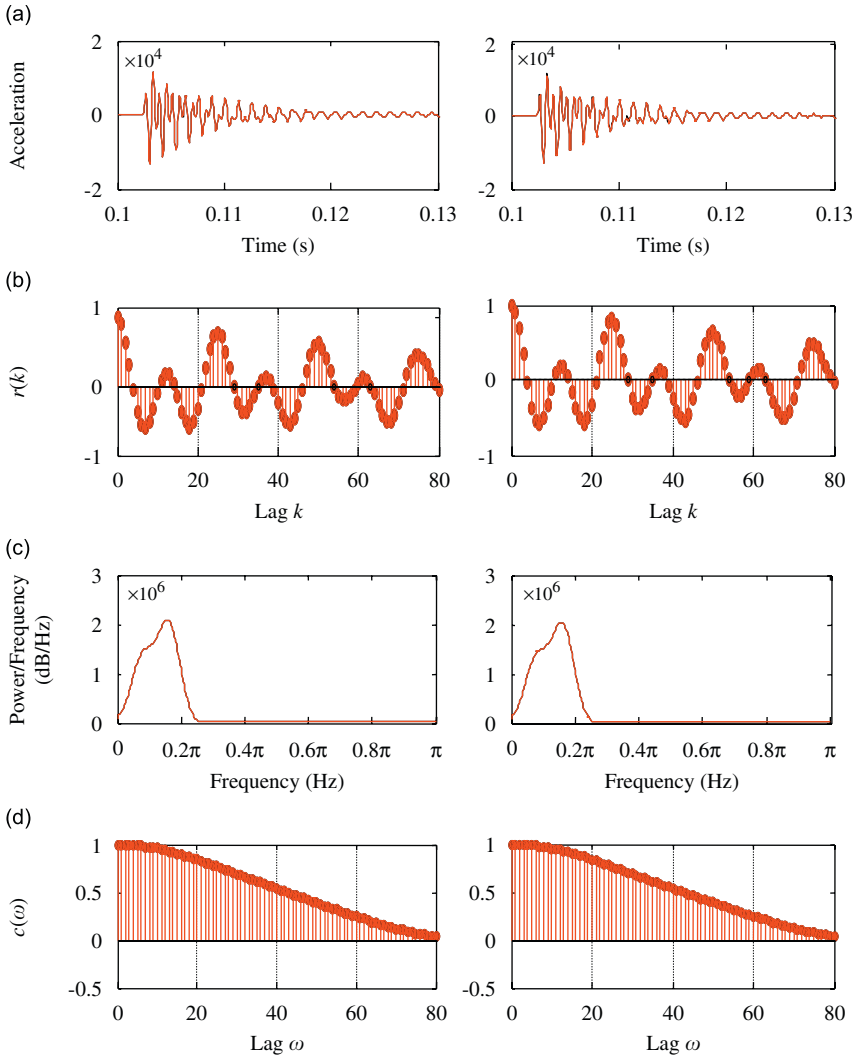

Fig. 9. Assessment of both experimental and predicted acceleration response data reconstructed using first four wavelet packets for mass 3 of the 15th sample under low-level shock vibration: (a) time series plots; (b) cross-correlation of signals; (c) power spectra density using Welch's method and (d) cross-coherence of signals.

4.3.2. *Deterministic validation*

Now we can construct a 12×60 matrix $\mathbf{D}$ representing 12 variables each having 60 values of the difference of wavelet packet component energy between the experimental and predicted response acceleration data. The 12-variable model validation problem becomes testing whether the set of means $\bar{\mathbf{D}}$ are equal to zeros.

Prior to applying the Bayesian methods described in Section 3, statistical analyses are performed on the feature data to find the parameters of priors for model assessment. Scatter plots indicate that the 12 variables are independent of each other. For example, Fig. 10 shows the scatter plots of the errors of energy variables: (a) e_1 vs. e_2, (b) e_1 vs. e_5, (c) e_5 vs. e_9, and (d) e_9 vs. e_{10}. Clearly, no significant relationship among these variables is found in these plots. Further, the probability plots are performed on the 12 sets of error data. A multivariate normal distribution is used to model these data. The obtained statistics of every variable are used for calculating the Bayes factor in Eqs. (8) and (10).

Using the statistics of 12 component energy variables for the 60 sets of energy data (i.e., the number of experiments $n = 60$), the Bayes factor values of -1566.58 and -8.51 are obtained by Eqs. (8) and (10) for the

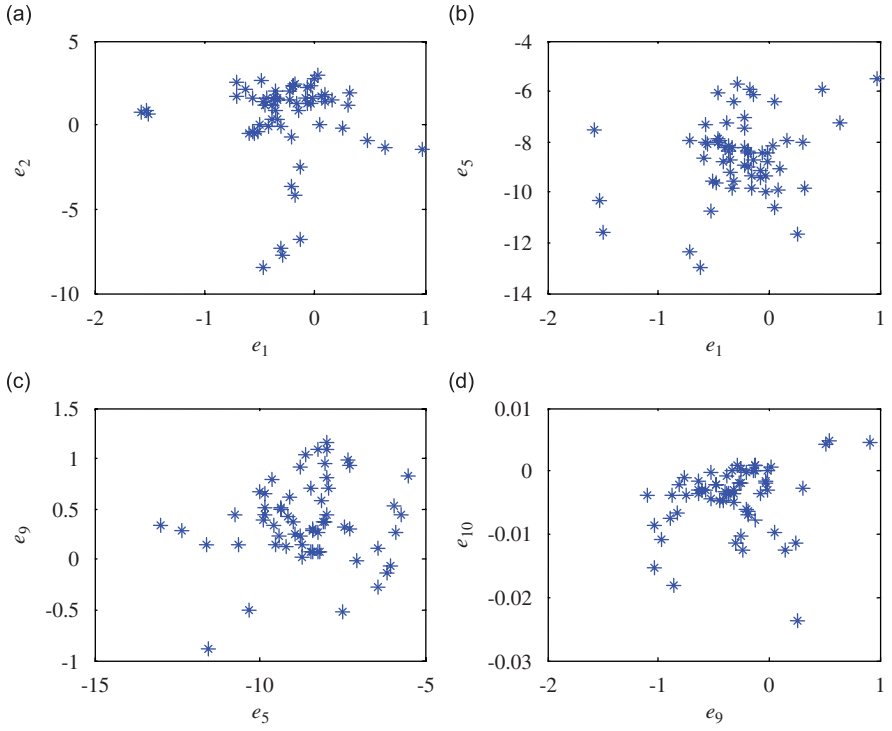

Fig. 10. Scatter plots of errors of the energy variables: (a) e_1 vs. e_2; (b) e_1 vs. e_5; (c) e_5 vs. e_9; and (d) e_9 vs. e_{10}.

Bayesian point- and interval-based methods, respectively. The acceptance confidence of 0% and 0.2% are obtained using Eq. (9) for the two methods. Obviously, the model is not suitable for practical application based on the comparison of the component energy of the response time series.

4.4. Sensitivity analyses of Bayes factor

In the context of multivariate model validation, it is important to understand the effects of the covariance matrix $\mathbf{\Lambda}$ in the point method (Eq. (8)) and the interval vector $\boldsymbol{\varepsilon}$ in the interval method (Eq. (10)) on the Bayes factor values. Sensitivity analyses of the Bayes factor are therefore performed here to investigate their effects on the model validation decision.

Fig. 11 shows the variation of Bayes factor versus various values of the two parameters in the Bayesian point and interval methods. Fig. 11a shows the variation of $b_M = \ln(B_M)$ with respect to the selection of $\mathbf{\Lambda}$ in the point method. The horizontal axes represent the multipliers of the variance matrix of error variables ($\mathbf{\Sigma}$). When $\mathbf{\Lambda}$ varies from $\mathbf{\Sigma}$ to zeros (solid line in Fig. 11a corresponding to the upper horizontal axis), the Bayes factor approaches to zero negatively. However, when $\mathbf{\Lambda}$ varies from zeros to $2\mathbf{\Sigma}$ (dashed line in Fig. 11a corresponding to the below horizontal axis), the logarithm of Bayes factor approaches a large negative value. Note that (1) the variations of the horizontal axes for the two curves in Fig. 11a are different and (2) the two curves imply different variation trends of the Bayes factor value. Obviously, the model is rejected regardless of the selection of $\mathbf{\Lambda}$. However, the value of $\mathbf{\Lambda}$ affects the Bayes factor value and further the probability of accepting the model based on Eq. (9).

Fig. 11b shows the variation of b_{iM} versus $\boldsymbol{\varepsilon}$ in the interval method. The basis vector of the interval $\boldsymbol{\varepsilon}_0$ is taken to be the standard deviation of multiple variables in this study (i.e., the squared diagonal values of the variance matrix $\mathbf{\Sigma}$). It is observed that the maximum value of Bayes factor in the interval method is -2.464. Thus, the model is always rejected regardless of the interval value. Again, the selection of various parameters

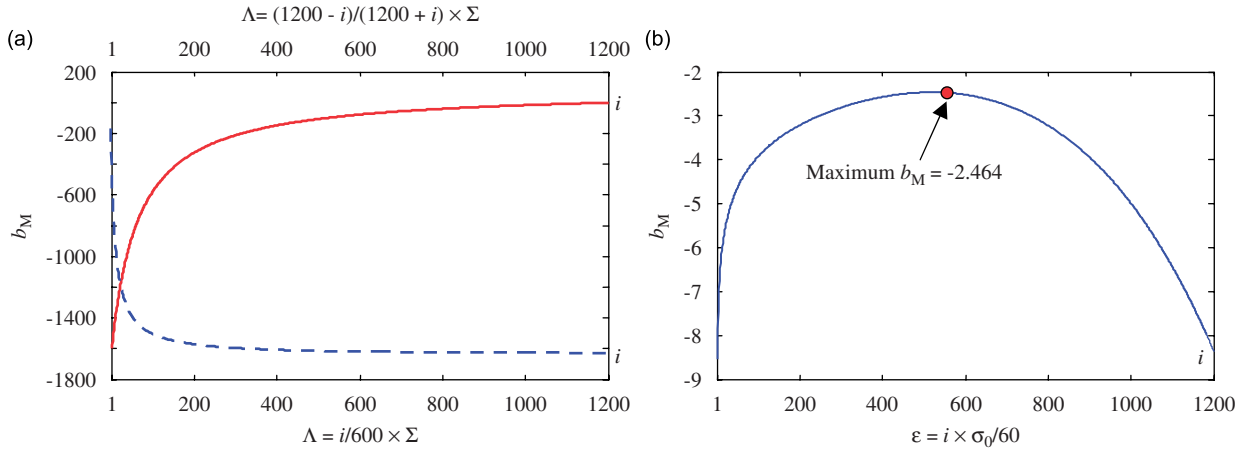

Fig. 11. Sensitivity analyses of Bayes factor: (a) variation of $\mathbf{\Lambda}$ in point method (Eq. (8)): solid line $b_M \to 0$ and dash line $b_M \to -1610$; and (b) variation of $\boldsymbol{\varepsilon}$ in interval method (Eq. (10)) (*note*: $b_M = \ln B_M$).

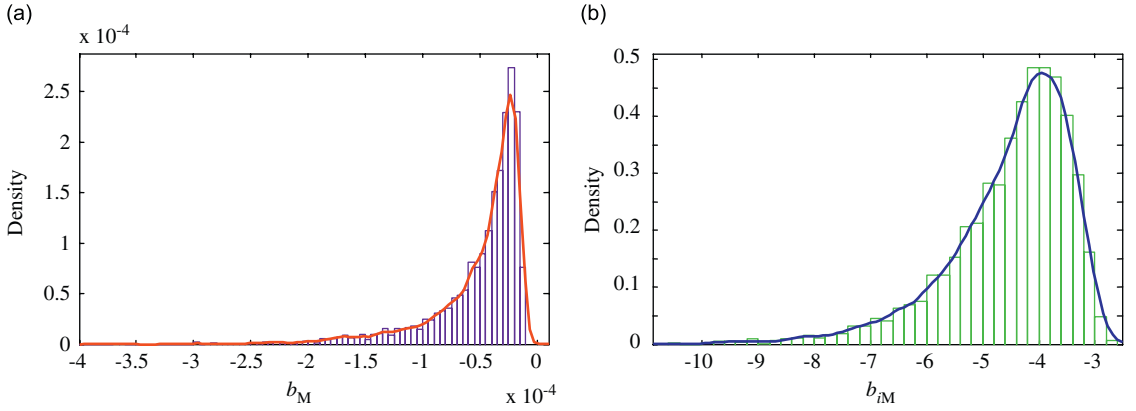

Fig. 12. Probabilistic assessment of model validity: (a) point method: $\gamma = \Pr(b_M > 0) = 0$, and (b) interval method: $\gamma = \Pr(b_{iM} > 0) = 0$.

$\mathbf{\Lambda}$ and $\boldsymbol{\varepsilon}$ will affect the magnitude of Bayes factor, but it will not affect the model validation decision in this example.

4.5. *Probabilistic assessment*

The statistics of errors of 12 energy variables are used to perform the probabilistic assessment of model validity. Fig. 12 shows the simulated results for both the point method (Fig. 12a) and the interval method (Fig. 12b). It is found that the model is rejected based on the probabilistic assessment using both Bayesian methods, which is in consistence with the results obtained from the deterministic assessment previously.

5. Concluding remarks

In this paper, a Bayesian wavelet method is explored for model validation under uncertainty, using multivariate time series data from both model output and experimental observations. Discrete wavelet packet transform is first used to decompose the signal into multiple time–frequency resolutions. The decomposed coefficients are used to compute the wavelet packet component energy stored in a signal. With the aid of both

cross-correlation and cross-coherence evaluation metrics, the feature components containing little discriminative information are discarded, resulting in a feature subset having a reduced number of parameters without compromising the model validation accuracy. A generalized likelihood ratio based on Bayesian interval hypothesis testing is derived as a quantitative validation metric based on Bayes' theorem and Gaussian distribution assumption of the difference of wavelet packet component energy between validation data and model prediction. Both the point and interval hypothesis testing-based Bayesian methods are explored for multivariate model assessment based on the wavelet packet component energy.

The proposed methodology is illustrated with a structural dynamics problem using simulated validation time history data, a validation challenge problem developed at Sandia National Laboratories. The prediction model is rejected based on both the deterministic validation metrics using the Bayesian hypothesis testing approach, and the probabilistic assessment results. Numerical results from sensitivity analyses also show that the values of covariance matrix $\mathbf{\Lambda}$ in the point method and the parameters $\boldsymbol{\varepsilon}$ in the interval method affect the magnitude of the Bayes factor metrics, but they do not affect the model validation result. As a result, the validation framework has been demonstrated to provide an effective methodology for multivariate model validation using time series data collected from dynamic systems. The results can provide objective, rational decision support for model validation under uncertainty. In future, the proposed Bayesian wavelet methodology may be extended to assess more complicated dynamic systems.

Acknowledgments

This research was supported by funds from Sandia National Laboratories, Albuquerque, New Mexico (contract no. BG-7732, project monitors: Dr. Thomas L. Paez, Dr. Laura P. Swiler, and Dr. Martin Pilch). The support is gratefully acknowledged. The authors also thank the anonymous referees for their helpful comments to improve the manuscript.

Appendix A. Interval-based hypothesis Bayes factor B_{iM}

For the multivariate case, the likelihood function of the multiple observations, $f(\mathbf{D})$, is expressed as follows:

$$f(\mathbf{D}) = \frac{|\mathbf{\Sigma}|^{-1/2}}{(2\pi)^{m/2}} \exp\left[-\frac{1}{2}(\mathbf{D}-\boldsymbol{\mu})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{D}-\boldsymbol{\mu})\right]. \tag{A.1}$$

Based on the definition of interval-based hypothesis testing for model validation, the Bayes factor in Eq. (10) can be rewritten as

$$B_M = \frac{\int_{-\boldsymbol{\varepsilon}+\mathbf{e}_0}^{\boldsymbol{\varepsilon}+\mathbf{e}_0} f(\mathbf{D}|\boldsymbol{\mu}) f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu}}{\int_{-\infty}^{-\boldsymbol{\varepsilon}+\mathbf{e}_0} f(\mathbf{D}|\boldsymbol{\mu}) f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu} + \int_{\boldsymbol{\varepsilon}+\mathbf{e}_0}^{\infty} f(\mathbf{D}|\boldsymbol{\mu}) f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu}}. \tag{A.2}$$

Substituting the distributions of $f(\mathbf{D}|\boldsymbol{\mu})$ and $f(\boldsymbol{\mu})$ into the nominator of the right-hand side of Eq. (A.2) with a few algebraic transformation yields

$$\begin{aligned} f(\mathrm{Data}|H_0) &= \int_{-\boldsymbol{\varepsilon}+\mathbf{e}_0}^{\boldsymbol{\varepsilon}+\mathbf{e}_0} f(\mathbf{D}|\boldsymbol{\mu}) f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu} \\ &= \int_{-\boldsymbol{\varepsilon}+\mathbf{e}_0}^{\boldsymbol{\varepsilon}+\mathbf{e}_0} \frac{1}{(2\pi)^{(m+n)/2}|\mathbf{\Sigma}|^{n/2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{d}_i-\boldsymbol{\mu})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{d}_i-\boldsymbol{\mu})\right] \\ &\quad \frac{1}{\sqrt{2\pi|\mathbf{\Lambda}|}} \exp\left[-\frac{1}{2}(\boldsymbol{\mu}-\boldsymbol{\rho})^{\mathrm{T}}\mathbf{\Lambda}^{-1}(\boldsymbol{\mu}-\boldsymbol{\rho})\right] \mathrm{d}\boldsymbol{\mu} \end{aligned}$$

$$
\begin{aligned}
&= \int_{-\boldsymbol{\varepsilon}+\mathbf{e}_0}^{\boldsymbol{\varepsilon}+\mathbf{e}_0} \frac{1}{(2\pi)^{(m+n)/2}|\boldsymbol{\Sigma}|^{n/2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\mathbf{d}_i-\bar{\mathbf{D}}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\mathbf{d}_i-\bar{\mathbf{D}}\right)\right] \\
&\quad \times \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\bar{\mathbf{D}}-\boldsymbol{\mu}\right)^{\mathrm{T}}\Sigma^{-1}\left(\bar{\mathbf{D}}-\boldsymbol{\mu}\right)\right]\frac{1}{\sqrt{2\pi|\boldsymbol{\Lambda}|}}\exp\left[-\frac{1}{2}(\boldsymbol{\mu}-\boldsymbol{\rho})^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}(\boldsymbol{\mu}-\boldsymbol{\rho})\right]\mathrm{d}\boldsymbol{\mu} \\
&= \frac{\exp\left(-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{d}_i-\bar{\mathbf{D}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{d}_i-\bar{\mathbf{D}})\right)}{(2\pi)^{(m+n)/2}|\boldsymbol{\Sigma}|^{n/2}}\sqrt{\frac{|\boldsymbol{\Sigma}|}{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|}} \\
&\quad \times \exp\left[-\frac{n}{2}(\bar{\mathbf{D}}-\boldsymbol{\rho})^{\mathrm{T}}(n\boldsymbol{\Lambda}+\boldsymbol{\Sigma})^{-1}(\bar{\mathbf{D}}-\boldsymbol{\rho})\right] \\
&\quad \times \int_{(-\boldsymbol{\varepsilon}+\mathbf{e}_0)\sqrt{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|}}^{(\boldsymbol{\varepsilon}+\mathbf{e}_0)\sqrt{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|}} \frac{1}{\sqrt{2\pi|\boldsymbol{\Lambda}||\boldsymbol{\Sigma}|}}\exp\left[-\frac{1}{2}(\mathbf{Z}-\mathbf{Z}_0)^{\mathrm{T}}(\boldsymbol{\Pi})^{-1}(\mathbf{Z}-\mathbf{Z}_0)\right]\mathrm{d}\mathbf{Z} \\
&= \frac{\exp\left(-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{d}_i-\bar{\mathbf{D}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{d}_i-\bar{\mathbf{D}})\right)}{(2\pi)^{(m+n)/2}|\boldsymbol{\Sigma}|^{n/2}} \\
&\quad \times \frac{K\sqrt{|\boldsymbol{\Sigma}|}}{\sqrt{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|}}\exp\left[-\frac{n}{2}(\bar{\mathbf{D}}-\boldsymbol{\rho})^{\mathrm{T}}(n\boldsymbol{\Lambda}+\boldsymbol{\Sigma})^{-1}(\bar{\mathbf{D}}-\boldsymbol{\rho})\right],
\end{aligned}
\tag{A.3}
$$

where $\mathbf{d}_i = \begin{bmatrix} d_{i1} & d_{i2} & \dots & d_{im} \end{bmatrix}^{\mathrm{T}}$ $(i = 1, 2, \dots, n)$ is the ith difference of m component energy variables, $\bar{\mathbf{D}} = \begin{bmatrix} \bar{d}_1 & \bar{d}_2 & \dots & \bar{d}_m \end{bmatrix}^{\mathrm{T}}$ is the m component average energy values with $\bar{d}_i = (1/n)\sum_{j=1}^{n} d_{ij}$ $(i = 1, 2, \dots, m)$, $|.|$ denotes the determinant of a matrix, and the parameter K can be obtained using the standard normal distribution as follows:

$$
\begin{aligned}
K &= \int_{(-\boldsymbol{\varepsilon}+\mathbf{e}_0)\sqrt{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|}}^{(\boldsymbol{\varepsilon}+\mathbf{e}_0)\sqrt{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|}} \frac{1}{\sqrt{2\pi|\boldsymbol{\Pi}|}}\exp\left[-\frac{1}{2}(\mathbf{Z}-\mathbf{Z}_0)^{\mathrm{T}}(\boldsymbol{\Pi})^{-1}(\mathbf{Z}-\mathbf{Z}_0)\right]\mathrm{d}\mathbf{Z} \\
&= \Phi(\varepsilon_2, \mathbf{Z}_0, \boldsymbol{\Pi}) - \Phi(\varepsilon_1, \mathbf{Z}_0, \boldsymbol{\Pi}),
\end{aligned}
\tag{A.4}
$$

in which the parameters

$$
\begin{aligned}
&\mathbf{Z} = \boldsymbol{\mu}\sqrt{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|}, \quad \varepsilon_1 = (-\boldsymbol{\varepsilon}+\mathbf{e}_0)\sqrt{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|}, \quad \varepsilon_2 = (\boldsymbol{\varepsilon}+\mathbf{e}_0)\sqrt{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|} \\
&\mathbf{Z}_0 = \frac{n\bar{\mathbf{Y}}|\boldsymbol{\Lambda}|+\boldsymbol{\rho}|\boldsymbol{\Sigma}|}{\sqrt{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|}} \quad \text{and} \quad \boldsymbol{\Pi} = \boldsymbol{\Sigma}|\boldsymbol{\Lambda}|
\end{aligned}
$$

and $\Phi(.)$ presents a multivariate normal cumulative distribution function, which is computed using the numerical algorithm proposed by Genz [55].

Based on the Bayes theorem, the marginal likelihood (or posterior joint function) of H_1 given $\mathbf{D}$ is

$$
\begin{aligned}
f(\mathrm{Data}|\mathrm{H}_1) &= \int_{\Theta-\Theta_0} f(\mathbf{D}, \boldsymbol{\mu}|\mathrm{H}_1)\,\mathrm{d}\boldsymbol{\mu} = \int_{\Theta-\Theta_0} f(\mathbf{D}|\boldsymbol{\mu})f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu} \\
&= \int_{\Theta} f(\mathbf{D}|\boldsymbol{\mu})f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu} - \int_{\Theta_0} f(\mathbf{D}|\boldsymbol{\mu})f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu},
\end{aligned}
\tag{A.5}
$$

where Θ and Θ_0 represent the entire space $[-\infty, \infty]$ and the interval space $[-\boldsymbol{\varepsilon}+\mathbf{e}_0, \boldsymbol{\varepsilon}+\mathbf{e}_0]$, respectively. Using the relationship of

$$
\int_{\Theta} \frac{1}{\sqrt{2\pi|\boldsymbol{\Pi}|}}\exp\left[-\frac{1}{2}(\mathbf{Z}-\mathbf{Z}_0)^{\mathrm{T}}(\boldsymbol{\Pi})^{-1}(\mathbf{Z}-\mathbf{Z}_0)\right]\mathrm{d}\mathbf{Z} = 1,
$$

similar to the derivatives of Eq. (A.3), we can easily obtain

$$\int_{\Theta} f(\mathbf{D}|\boldsymbol{\mu}) f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu} = \frac{\exp\left(-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{d}_i-\bar{\mathbf{D}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{d}_i-\bar{\mathbf{D}})\right)}{(2\pi)^{(m+n)/2}|\boldsymbol{\Sigma}|^{n/2}} \times \frac{\sqrt{|\boldsymbol{\Sigma}|}}{\sqrt{n|\boldsymbol{\Lambda}|+|\boldsymbol{\Sigma}|}} \exp\left[-\frac{n}{2}(\bar{\mathbf{D}}-\boldsymbol{\rho})^{\mathrm{T}}(n\boldsymbol{\Lambda}+\boldsymbol{\Sigma})^{-1}(\bar{\mathbf{D}}-\boldsymbol{\rho})\right]. \tag{A.6}$$

Thus, substituting Eqs. (A.3), (A.5), and (A.6) into Eq. (10), we have the likelihood ratio (or Bayes factor), B_{iM}, as follows:

$$B_{iM} = \frac{\int_{\Theta_0} f(\mathbf{D}|\boldsymbol{\mu}) f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu}}{\int_{\Theta} f(\mathbf{D}|\boldsymbol{\mu}) f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu} - \int_{\Theta_0} f(\mathbf{D}|\boldsymbol{\mu}) f(\boldsymbol{\mu})\,\mathrm{d}\boldsymbol{\mu}} = \frac{K}{1-K}, \tag{A.7}$$

where K is calculated by Eq. (A.4).

References

[1] DOD. *Verification, Validation, and Accreditation (VV&A) Recommended Practices Guide*, Department of Defense, Alexandria, VA, 1996, ⟨http://vva.dmso.mil/⟩.

[2] AIAA. *Guide for the Verification and Validation of Computational Fluid Dynamics Simulations*, American Institute of Aeronautics and Astronautics AIAA-G-077-1998, Reston, VA, 1998.

[3] DOE, *Accelerated Strategic Computing Initiative (ASCI) Program Plan*, DOE/DP-99-000010592, Department of Energy, Washington, DC, 2000.

[4] ASME. *Guide for Verification and Validation in Computational Solid Mechanics*, American Society of Mechanical Engineers, ASME V&V 10, 2006.

[5] O. Balci, Verification, validation, and accreditation of simulation models, *Proceedings of the 29th Conference on Winter Simulation*, Atlanta, GA, December 1997.

[6] P.J. Roache, *Verification and validation in computational science and engineering*, Hermosa Publishers, Albuquerque, NM, 1998.

[7] W.L. Oberkampf, T.G. Trucano, Verification and validation in computational fluid dynamics, *Progress in Aerospace Sciences* 38 (2002) 209–272.

[8] W.L. Oberkampf, T.G. Trucano, Ch. Hirsch, Verification, validation, and predictive capabilities in computational engineering and physics, Technical Report Sand. No 2003-3769, Sandia National Laboratories, Albuquerque, NM, 2003.

[9] I. Babuska, J.T. Oden, Verification and validation in computational engineering and science: basic concepts, *Computer Methods in Applied Mechanics and Engineering* 193 (2004) 4057–4066.

[10] R. Sargent, Validation and verification of simulation models, *Proceedings of the 2004 Winter Simulation Conference*, Washington, DC, March 2004.

[11] K. Campbell, A brief survey of statistical model calibration ideas, *Proceedings of International Conference on Sensitivity Analysis of Model Output*, Santa Fe, NM, 2004.

[12] W.L. Oberkampf, T.G. Trucano, Design of and comparison with verification and validation benchmarks, Technical Report Sand. No. 2006-5376C, Sandia National Laboratories, Albuquerque, NM, 2006.

[13] W.L. Oberkampf, M.F. Barone, Measures of agreement between computation and experiment: validation metrics, *Journal of Computational Physics* 217 (2006) 5–36.

[14] T.G. Trucano, R.G. Easterling, K.J. Dowding, T.L. Paez, A. Urbina, V.J. Romero, B.M. Rutherford, R.G. Hills, Description of the Sandia Validation Metrics Project, Technical Report Sand. No. 2001-1339, Sandia National Laboratories, Albuquerque, NM, 2001.

[15] K.J. Dowding, M. Pilch, R.G. Hills, Thermal validation challenge problem, Sandia National Laboratories, Albuquerque, NM, 2006, ⟨http://www.esc.sandia.gov/VCWwebsite/ThermalProblemDescrip.pdf⟩.

[16] I. Babuska, F. Nobile, R. Tempone, Model validation challenge problem: static frame problem, Sandia National Laboratories, Albuquerque, NM, 2006, ⟨http://www.esc.sandia.gov/VCWwebsite/MechanicsProblemDescrip.pdf⟩.

[17] J.R. Red-Horse, T.L. Paez, Model validation challenge problem: structural dynamics application, Sandia National Laboratories, Albuquerque, NM, 2006, ⟨http://www.esc.sandia.gov/VCWwebsite/StructuralDynProbDescr.pdf⟩

[18] J. McFarland, S. Mahadevan, Multivariate significance testing and model calibration under uncertainty, in: T.J.R. Hughes, J.T. Oden, M. Papadrakakis (Eds.), *Computer Methods in Applied Mechanics and Engineering (special issue)*, 2008.

[19] T.L. Paez, A. Urbina, Validation of mathematical models of complex structural dynamic systems, *Proceedings of the Ninth International Congress on Sound and Vibration*, Orlando, FL, 2002.

[20] F.M. Hemez, S.W. Doebling, Validation of structural dynamics models at Los Alamos national laboratory, *Proceedings of the 41st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Atlanta, GA, April 2000.

[21] S.W. Doebling, T.A. Butler, J.F. Schultze, F.M. Hemez, L.M. Moore, M.D. McKay, Validation of transient structural dynamics simulations: an example, *Proceedings of the SAMO 2001: Third International Conference on Sensitivity Analysis of Model Output*, Madrid, Spain, June 2001.

[22] R.G. Hills, T.G. Trucano, Statistical validation of engineering and scientific models: a maximum likelihood based metric, Technical Report Sand. No. 2001-1783, Sandia National Laboratories, Albuquerque, NM, 2002.
[23] R.G. Hills, I. Leslie, Statistical validation of engineering and scientific models: validation experiments to application, Technical Report Sand. No. 2003-0706, Sandia National Laboratories, Albuquerque, NM, 2003.
[24] B.M. Rutherford, K. Dowding, An approach to model validation and model-based prediction—polyurethane foam case study, Technical Report Sand. No. 2003-2336, Sandia National Laboratories, Albuquerque, NM, 2003.
[25] K.J. Dowding, R.G. Hills, I. Leslie, M. Pilch, B.M. Rutherford, M.L. Hobbs, Case study for model validation: assessing a model for thermal decomposition of polyurethane foam, Technical Report Sand. No. 2004-3632, Sandia National Laboratories, Albuquerque, NM, 2004.
[26] W. Chen, L. Baghdasaryan, T. Buranathiti, J. Cao, Model validation via uncertainty propagation and data transformation, *AIAA Journal* 42 (2004) 1406–1415.
[27] R. Zhang, S. Mahadevan, Bayesian methodology for reliability model acceptance, *Reliability Engineering and System Safety* 80 (2003) 95–103.
[28] S. Mahadevan, R. Rebba, Validation of reliability computational models using Bayes networks, *Reliability Engineering and System Safety* 87 (2005) 223–232.
[29] X. Jiang, S. Mahadevan, Bayesian risk-based decision method for model validation under uncertainty, *Reliability Engineering and System Safety* 92 (2007) 707–718.
[30] R. Rebba, S. Mahadevan, Model predictive capability assessment under uncertainty, *AIAA Journal* 44 (2006) 2376–2384.
[31] G.E.P. Box, D.R. Cox, An analysis of transformations, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 26 (1964) 211–252.
[32] R. Rebba, Model Validation and Design under Uncertainty, PhD Thesis, Vanderbilt University, 2005.
[33] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11 (1989) 674–693.
[34] I. Daubechies, Orthonormal bases of compactly supported wavelets, *Communication on Pure and Applied Mathematics* 41 (1988) 909–996.
[35] I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.
[36] P.R. Coifman, M.V. Wickerhauser, Entropy-based algorithms for best basis selection, *IEEE Transactions on Information Theory* 38 (1992) 713–718.
[37] X. Jiang, S. Mahadevan, H. Adeli, Bayesian wavelet packet denoising for structural system identification, *Structural Control and Health Monitoring* 14 (2007) 333–356.
[38] X. Jiang, H. Adeli, Wavelet packet-autocorrelation function method for traffic flow pattern analysis, *Computer-Aided Civil and Infrastructure Engineering* 19 (2004) 324–337.
[39] G.G. Yen, K.-C. Lin, Wavelet packet feature extraction for vibration monitoring, *IEEE Transactions on Industrial Electronics* 47 (2000) 650–667.
[40] Z. Sun, C.C. Chang, Structural damage assessment based on wavelet packet transform, *ASCE Journal of Structural Engineering* 128 (2002) 1354–1361.
[41] A. Karim, H. Adeli, Incident detection algorithm using wavelet energy representation of traffic pattern, *ASCE Journal of Transportation Engineering* 128 (2002) 223–242.
[42] J. Kittler, Mathematics methods of feature selection in pattern recognition, *International Journal of Man-Machine Studies* 7 (1975) 609–637.
[43] M.H. Chen, D. Lee, T. Pavlidis, Residual analysis for feature detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (1991) 30–40.
[44] Y.S. Cho, S.B. Kim, E.L. Hixson, E.J. Powers, A digital technique to estimate second-order distortion using higher order coherence spectra, *IEEE Transactions on Signal Processing* 40 (1992) 1029–1040.
[45] O. Rioul, P. Flandrin, Time-scale energy distributions: a general class extending wavelet transforms, *IEEE Transactions on Signal Processing* 40 (1992) 1746–1757.
[46] D.S. Ornstein, B. Weiss, Entropy and data compression schemes, *IEEE Transactions on Information Theory* 39 (1993) 78–83.
[47] Y. Wu, R. Du, Feature extraction and assessment using wavelet packets for monitoring of machining processes, *Mechanical System and Signal Processing* 10 (1996) 29–53.
[48] R. Yan, R.X. Gao, An efficient approach to machine health diagnosis based on harmonic wavelet packet transform, *Robotics and Computer-Integrated Manufacturing* 21 (2005) 291–301.
[49] X. Jiang, S. Mahadevan, Bayesian validation assessment of multivariate computational models, *Journal of Applied Statistics* 35 (2008) 49–65.
[50] C.S. Burrus, R.A. Gopinath, H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*, Prentice-Hall, Englewood Cliffs, NJ, 1998.
[51] P.D. Welch, The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, *IEEE Transactions on Audio Electroacoustics* AU-15 (1967) 70–73.
[52] R. Kass, A. Raftery, Bayes factors, *Journal of the American Statistical Association* 90 (1995) 773–795.
[53] H.S. Migon, D. Gamerman, *Statistical Inference—An Integrated Approach*, Arnold, a Member of the Holder Headline Group, London, UK, 1999.
[54] G.E.P. Box, D.R. Cox, An analysis of transformations, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 26 (1964) 211–252.
[55] Genz, Numerical computation of multivariate normal probabilities, *Journal of Computational and Graphical Statistics* 1 (1992) 141–149.