Conference Proceedings of the Society for Experimental Mechanics Series

T. Simmermacher · S. Cogan · L.G. Horta · R. Barthorpe *Editors*

Topics in Model Validation and Uncertainty Quantification, Volume 4

Proceedings of the 30th IMAC, A Conference on Structural Dynamics, 2012

SEM

Springer

Conference Proceedings of the Society for Experimental Mechanics Series

Series Editor
Tom Proulx
Society for Experimental Mechanics, Inc.,
Bethel, CT, USA

For further volumes:
http://www.springer.com/series/8922

T. Simmermacher • S. Cogan • L.G. Horta • R. Barthorpe
Editors

Topics in Model Validation and Uncertainty Quantification, Volume 4

Proceedings of the 30th IMAC, A Conference on Structural Dynamics, 2012

Springer

Editors
T. Simmermacher
Sandia National Laboratories
Albuquerque, NM, USA

S. Cogan
University of Franche-Comte
Besançon, France

L.G. Horta
NASA Langley Research Center
Hampton, VA, USA

R. Barthorpe
University of Sheffield
Sheffield, UK

ISSN 2191-5644 e-ISSN 2191-5652
ISBN 978-1-4614-2430-7 e-ISBN 978-1-4614-2431-4
DOI 10.1007/978-1-4614-2431-4
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2012934958

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Topics in Model Validation and Uncertainty Quantification represents one of six volumes of technical papers presented at the 30th IMAC, A Conference and Exposition on Structural Dynamics, 2012 organized by the Society for Experimental Mechanics, and held in Jacksonville, Florida, January 30–February 2, 2012. The full proceedings also include volumes on Dynamics of Civil Structures; Substructuring and Wind Turbine Dynamics; Nonlinear Dynamics; and Modal Analysis, I & II.

Each collection presents early findings from experimental and computational investigations on an important area within Structural Dynamics. This volume focuses on the importance of Model Validation and Uncertainty Quantification. In keeping with the goals of encouraging advancement and application of model Verification and Validation (V&V) and uncertainty quantification methods in Structural Dynamics, this volume includes tutorials on the use of the Bayesian and Markov Chain Monte Carlo Methods. The Bayesian paradigm is of great interest to the engineering community as it provides a principled framework for handling uncertainties in complex systems. The Markov Chain Monte Carlo method is, in turn, a useful tool for practical implementation of the Bayesian paradigm.

The organizers would like to thank the authors, presenters, session organizers, and session chairs for their participation in this track.

Albuquerque, NM, USA	T. Simmermacher
Besançon, France	S. Cogan
Hampton, VA, USA	L.G. Horta
Sheffield, UK	R. Barthorpe

Contents

Chapter 1
On Assessing the Robustness of Structural Health Monitoring Technologies

Christopher J. Stull, François M. Hemez, and Charles R. Farrar

Abstract As *Structural Health Monitoring* (SHM) continues to gain popularity, both as an area of research and as a tool for use in industrial applications, the number of technologies associated with SHM will also continue to grow. As a result, the engineer tasked with developing a SHM system is faced with myriad hardware and software technologies from which to choose, often adopting an *ad hoc* qualitative approach based on physical intuition or past experience to making such decisions, and offering little in the way of justification for a particular decision. The present paper offers a framework that aims to provide the engineer with a qualitative approach for choosing from among a suite of candidate SHM technologies. The framework is outlined for the general case, where a supervised learning approach to SHM is adopted, and is then demonstrated on a problem commonly encountered when developing SHM systems: selection of a damage classifier, where the engineer must select from among a suite of candidate classifiers, the one most appropriate for the task at hand. The data employed for these problems are taken from a preliminary study that examined the feasibility of applying SHM technologies to the RAPid Telescopes for Optical Response observatory network. ***(Approved for unlimited public release on September 20, 2011, LA-UR 11-05398, Unclassified)***

Keywords Structural health monitoring • Info-gap decision theory • Decision making • Model selection • Uncertainty • Autoregressive model

1.1 Introduction

At the most general level, *Structural Health Monitoring* (SHM) often employs a model to approximate the changes in behavior of a physical system that result from damage. An increased interest in SHM by the civil, aerospace and mechanical engineering communities has led to a relative explosion of technologies proposed to both improve upon and extend the application of these research fields. As these technologies are brought online, the engineer tasked with developing a SHM system is not only faced with the difficulty of choosing from among these technologies, but must also do so defensibly.

Offering such defensible approaches, two recent research efforts [1, 2] adopted Bayesian model selection frameworks to select from among a family of models, the model that most probably represented a system for which data exist. Whereas [1] focused more generally on the problem of selecting a model that approximates a physical system, [2] focused more specifically on the determining the most probable order of an autoregressive model with exogenous inputs (ARX). In general, these approaches employ Bayes' theorem as [3]

$$P(\mathcal{M}_i|\mathcal{D}) \propto P(\mathcal{D}|\mathcal{M}_i)\,P(\mathcal{M}_i). \tag{1.1}$$

C.J. Stull (✉) • C.R. Farrar
Engineering Institute, Los Alamos National Laboratory, PO Box 1663, Mail Stop T001, Los Alamos, NM 87545, USA
e-mail: stull@lanl.gov

F.M. Hemez
X-Theoretical Design Division, Los Alamos National Laboratory, PO Box 1663, Mail Stop T087, Los Alamos, NM 87545, USA

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_1,

Here, $P(\mathcal{M}_i)$ is the *prior probability* assigned to the model $\mathcal{M}_i$, or the probability that the model $\mathcal{M}_i$ is the true model *before* observing the data. $P(\mathcal{D}|\mathcal{M}_i)$ is the *model evidence*, or the probability that the data $\mathcal{D}$ will be observed given the model $\mathcal{M}_i$. $P(\mathcal{D}|\mathcal{M}_i)$ is also referred to as the *marginal likelihood*, as it is essentially a likelihood function associated with model $\mathcal{M}_i$, where the effects of the model parameters are removed by randomly sampling from the prior probabilities associated with the model parameters [3]. Lastly, the term $P(\mathcal{M}_i|\mathcal{D})$ is the *posterior probability* of the model $\mathcal{M}_i$, or the probability that model $\mathcal{M}_i$ is the true model *after* observing the data $\mathcal{D}$. With respect to the work of [1], the $\mathcal{M}_i$ in (1.1) corresponded to different model classes employed to approximate the dynamic behavior of structures subjected to seismic excitations. For [2], the $\mathcal{M}_i$ corresponded to the parameters that define different orders of an ARX model.

When solving Bayesian model comparison problems, models are typically assigned *equal* prior probabilities to represent the analyst's ignorance as to which model is the true model for the data at hand. Once the posterior probabilities associated with the models are estimated, these estimations may be used to compute information criteria (e.g. Akaike, Schwartz [3]) as a means of evaluating the relative merit of each model. While this framework offers a principled treatment of the analyst's uncertainty about the model selection problem, it is conditioned on data $\mathcal{D}$. As such, concern arises when it is admitted that these data are subject to variability (aleatoric and epistemic), as such changes have the potential to introduce additional uncertainties into the model selection problem.

1.1.1 Overview of Proposed Framework

The present paper offers an alternative framework with which to address the model selection problem within the context of SHM. Anchored in *Information-Gap Decision Theory* (IGDT) [4], this framework assumes a *non-probabilistic* description of the engineer's lack of knowledge about the data. Using this as a point of departure, the proposed framework aims to inform the engineer as to the *robustness*, or insensitivity to this ignorance, of a particular decision (i.e. SHM technology).

IGDT has been applied in a number of research topics, and the reader is directed to [5] for an exhaustive list of references. A recent application of IGDT to SHM is found in [6] where the authors assessed the robustness of an artificial neural network's ability to detect damage, given variability in the in the network inputs. While a similar strategy is adopted herein, an attempt has been made by the authors to generalize the approach in order to tackle broader classes of problems within SHM. To this end, the four-part statistical pattern recognition paradigm discussed in [7] is presented:

(i) Operational Evaluation;
(ii) Data Acquisition, Normalization, and Cleansing;
(iii) Feature Selection and Information Condensation; and
(iv) Statistical Model Development for Feature Discrimination.

The focus herein will be on decisions related to Part (iv) of this paradigm, and as such, it will be assumed for the purposes of this paper that the technologies, data, etc. associated with Parts (i), (ii), and (iii) have been defined and are fixed at this point in the SHM system design process. It is noted however, that the proposed framework is general, in that it could also be applied to decisions associated with Parts (i), (ii), and (iii).

1.1.2 Paper Organization

The remainder of this paper is organized around two main sections. Section 1.2 is devoted to formulating the problem of how to assess SHM technologies, in support of designing SHM systems that are robust to variability in the data and hence, the damage-sensitive features extracted from these data. First, the problem is cast within the context of IGDT. Then, important quantities relevant to this application of IGDT are defined, which feed into the definitions of the info-gap models of uncertainty and the robustness function. Section 1.3 then demonstrates the application of IGDT to an aspect of an SHM design, related to Part (iv) of the paradigm outlined above. The paper concludes with a summary of the work performed, improvements to the SHM development process, and ideas for future directions of research.

1.2 Problem Formulation

1.2.1 Overview

IGDT aims to facilitate decision-making in the face of uncertainty, particularly when there is little information available about this uncertainty *a priori*, rendering a probabilistic definition difficult to employ. To begin the discussion of the problem formulation, three questions that are central to any IGDT analysis are asked: (1) "What is the decision that needs to be made?" (2) "What is the uncertainty that affects the quality of this decision?" (3) "How is the quality of this decision quantified?" Coincidentally, these questions also arise when adopting a supervised learning approach to SHM. While there are many decisions that need to be made when adopting such an approach, the focus herein relates to choosing an appropriate classifier that employs these features for damage detection. Engineers faced with this question typically offer little in the way of objective reasoning for choosing one classifier over another, relying instead on a "tried-and-true" classifier. This statement is not meant to imply that expert opinion is invalid. Indeed, reliance on such tried-and-true solutions is often a *robust* decision, as the engineer does not expect to be surprised by the results associated with such a decision (or else that decision would not be made). However, the proposed methodology goes beyond this subjective approach and offers a principled and hence, defensible manner of assessing the robustness of decisions (tried-and-true and otherwise) to uncertainties affecting those decisions.

As for the second question, an important concern related to the supervised learning approach to SHM is addressed: data that initially represent the undamaged (or damaged) structural system can and often do change, with the potential to do so in an unexpected manner. That is, the structural condition information offered by damage-sensitive features may not be consistent over a long time-frame, potentially rendering the damage classifier obsolete. An adaptive SHM system offers a means of addressing this concern, but such approaches can be difficult to develop, implement, and maintain. The proposed methodology offers an alternative by assuming little about the future form of the damage-sensitive features, other than that it will vary due to environmental and/or operational variability. The robustness of the damage classifier then relates to its sensitivity to these variations.

The focus of the third question lies in the quantification of this robustness. While generally the most subjective question of the three, for the particular case of SHM, the answer to this question should relate to the effectiveness of the damage classifier to accurately assess the state of the structure. That is, as the future form of the damage-sensitive features undergoes more and more variability, a damage classifier will be assessed as robust to this variability if it is able to meet a minimum performance criterion related to the accuracy of its assessments.

1.2.2 Definitions

Having answered the above questions, the problem may now be formulated in mathematical terms. Prior to continuing, the reader is reminded that the focus of the present paper is on Part (iv) of the four-part statistical pattern recognition paradigm outlined above. Thus, the formulation begins with the assumption that there exist two sets of time-histories containing the measured responses of a structural system in its undamaged and damaged configurations. *(This assumption implies the completion of Parts (i) and (ii) of the paradigm.)* These sets are denoted $\mathcal{X}^u = \left\{\mathcal{X}_i^u : \forall i = 1, 2, \ldots, R\right\}$ and $\mathcal{X}^d = \left\{\mathcal{X}_i^d : \forall i = 1, 2, \ldots, S\right\}$, where the right superscripts u and d refer to the undamaged and damaged configurations, respectively, and R and S refer to the number of time-histories collected for the undamaged and damaged configurations, respectively. For typical applications of the supervised learning approach to SHM, R will be much greater than S, resulting in a well-characterized definition of the undamaged configuration. It would then be assumed that data exhibiting characteristics that deviate from this well-characterized, undamaged configuration would indicate the presence of damage in the structural system. It is further assumed that damage-sensitive features have been selected from each of the time-histories within $\mathcal{X}^u$ and $\mathcal{X}^d$, and collected in the sets $\Phi^u = \left\{\boldsymbol{\varphi}_i^u : \forall i = 1, 2, \ldots, R\right\}$ and $\Phi^d = \left\{\boldsymbol{\varphi}_i^d : \forall i = 1, 2, \ldots, S\right\}$. *(This assumption implies the completion of Part (iii) of the paradigm.)* In this study, the damage-sensitive features are the coefficients of the autoregressive models employed to approximate the time-histories within $\mathcal{X}^u$ and $\mathcal{X}^d$.

With these definitions, the means of the features associated with the undamaged and damaged configurations may be computed as in (1.2) and (1.3), respectively, where it is noted that the dimensions of these quantities are consistent with those of the features. That is, for scalar-based representations, the means are also scalars, and likewise for vector-based representations.

$$\boldsymbol{\mu}^u(\mathcal{P}) = \frac{1}{R}\sum_{i=1}^{R}\boldsymbol{\varphi}_i^u(\mathcal{P}) \tag{1.2}$$

$$\boldsymbol{\mu}^d(\mathcal{P}) = \frac{1}{S}\sum_{i=1}^{S}\boldsymbol{\varphi}_i^d(\mathcal{P}) \tag{1.3}$$

Note that for the discussion below, the terms “undamaged features” and “damaged features” are used. This terminology implies that the features *extracted from the data* are acquired from the structural system in its undamaged or damaged configuration, respectively, and not that the features themselves are undamaged or damaged.

1.2.3 Info-Gap Models of Uncertainty and the Robustness Function

Using the definitions provided in Sect. 1.2.2, the info-gap models of uncertainty $\mathcal{B}^u$ and $\mathcal{B}^d$., are now formulated. These are given as

$$\mathcal{B}^u(\alpha, \mu^u) = \left\{ \boldsymbol{b}^u : \left| \frac{b_i^u - \mu_i^u}{\mu_i^u} \right| \le \alpha, \forall i \right\}, \alpha \ge 0 \tag{1.4}$$

and

$$\mathcal{B}^d\left(\alpha, \mu^d\right) = \left\{ \boldsymbol{b}^d : \left| \frac{b_i^d - \mu_i^d}{\mu_i^d} \right| \le \alpha, \forall i \right\}, \alpha \ge 0 \tag{1.5}$$

for the structural system in its undamaged and damaged configurations, respectively, where it is noted that the functional dependencies on $\mathcal{P}$ have been omitted for clarity. Referred to as an “envelope-bound” info-gap model [4], (1.4) represents the *set* of all undamaged features b_i^u having elements centered about means μ_i^u, and uniformly distributed throughout the ranges $\pm\left(\alpha \times \mu_i^u\right)$. Analogously, (1.5) represents the set of all damaged features b_i^d having elements centered about means μ_i^d, and uniformly distributed throughout the ranges $\pm\left(\alpha \times \mu_i^d\right)$. α is referred to as the “horizon of uncertainty” and for the present discussion, represents the engineer’s lack of knowledge about the damage-sensitive features associated with the data. Thus, as α grows, the assumption that the features are clustered about the mean values μ_i^u and μ_i^d, becomes more invalid. Note that the formulation outlined above takes into account variability associated with the characterization of the undamaged configuration *and* the damaged configuration.

As stated previously, the robustness of a damage classifier is measured as its insensitivity to variability in the damage-sensitive features, while still providing adequate damage detection performance. Thus, a metric of performance must be defined against which the robustness of each damage classifier will be measured. While the performance of a damage classifier may be measured in variety of ways, a common metric involves the *Receiver Operating Characteristic* (ROC) [8], or more specifically, the ROC *curve*. ROC curves are constructed by computing the ordered pair $(TPR(t), FPR(t))$, where $TPR(t)$ and $FPR(t)$ are the *true positive* and *false positive rates*, respectively, for a given decision threshold t. In words, the $TPR(t)$ is the probability that a damage classifier correctly detects damage in the system, and correspondingly, the $FPR(t)$ is the probability that a damage classifier produces a false alarm or Type I error. Given the preference for a high $TPR(t)$ paired with a low $FPR(t)$, the *Area Under the ROC Curve* (AUC) has been proposed and employed as a global measure of performance for diagnostic tests [9, 10]. It is this measure, paired with the qualitative guidelines given in [11] that will be employed to assess the robustness curves in Sect. 1.3.2.

With a metric of performance against which to measure the robustness of damage classifiers, the corresponding robustness function is defined as

$$\hat{\alpha}(\mathcal{C}) = \max_{\alpha}\left\{ \min_{\mathbf{B}^u \subseteq \mathcal{B}^u(\alpha, \boldsymbol{\mu}^u), \mathbf{B}^d \subseteq \mathcal{B}^d(\alpha, \boldsymbol{\mu}^d)} AUC > AUC_{critical} \right\}. \tag{1.6}$$

The functional dependency $\mathcal{C}$ in (1.6) refers to a particular damage classifier, which is chosen by the analyst. Literally translated, the robustness function seeks the maximum horizon of uncertainty α, for which the minimum AUC or worst-case performance, is greater than the minimum performance requirement $AUC_{critical}$. Equation 1.6 is thus evaluated for various $\mathcal{C}$'s, providing the analyst with a principled way in which to compare and contrast the relative merits of the chosen $\mathcal{C}$'s.

Before presenting a conceptual example, the practical problem of computing the inner minimum of (1.6) is discussed. Oftentimes, there exists an analytical expression for the performance of a model that permits the derivation of an explicit expression for the robustness function. Consider for example, a cantilevered beam under the influence of uncertain loading with respect to the assumed info-gap model of uncertainty, whose performance is measured by the magnitude of the tip deflection. In this situation, an expression for the tip deflection may be derived that takes into account the uncertainty in the loading, and the robustness function may subsequently be derived using elementary calculus. In cases for which such closed-form solutions are intractable, as in the present case, the info-gap model(s) of uncertainty may instead be sampled. For each sample, the performance is computed and for the case of (1.6), the minimum performance from among the samples is stored. If the info-gap model(s) of uncertainty are sufficiently sampled for a fixed value of α, it is reasonable to assume that the robustness curves are well-approximated, and the IGDT analysis may proceed. Particular to robustness function presented in (1.6), the performance metric is derived from a ROC curve, the construction of which requires *multiple* classified instances of the uncertain feature vector. Thus, the robustness actually depends upon a sampled *subset* of the info-gap models given in (1.4) and (1.5); these subsets are denoted $\mathbf{B}^u$ and $\mathbf{B}^d$ in (1.6). As above, the performance of each sampled subset is computed and the minimum performance from among these subsets is stored. Alternately, an optimization scheme (e.g. genetic algorithm) may be employed to determine the minimum performance. (While not a focus of the present paper, this issue is noted here for the sake of completeness.)

1.3 Demonstration Problem

1.3.1 Experimental Data and Analysis Procedure

The data employed for the following demonstration problems are taken from a preliminary study that examined the feasibility of applying SHM technologies to the *RAPid Telescopes for Optical Response* (RAPTOR) observatory network [12]. The RAPTOR observatory network consists of several autonomous, robotic, astronomical observatories designed to search for astrophysical transients, with specific focus on gamma ray bursts [13]. The high duty cycles and remote locations of these observatories present significant difficulties with respect to planning maintenance schedules, making them ideal candidates for SHM.

At issue in this study is the wear of a component of the drive mechanism that controls the right ascension axis of the telescope (see Fig. 1.1). This component, referred to as a "capstan," is a stainless steel rod with a urethane coating that provides the friction interface between the motor that drives the telescope mount and the drive wheel. The urethane coating is known to experience wear with use, resulting in irregular travel of the drive wheel and the eventual inability to actuate the drive wheel. Figure 1.2 provides examples of both undamaged and damaged capstans, where these qualitative assessments were provided by an expert familiar with the RAPTOR observatories.

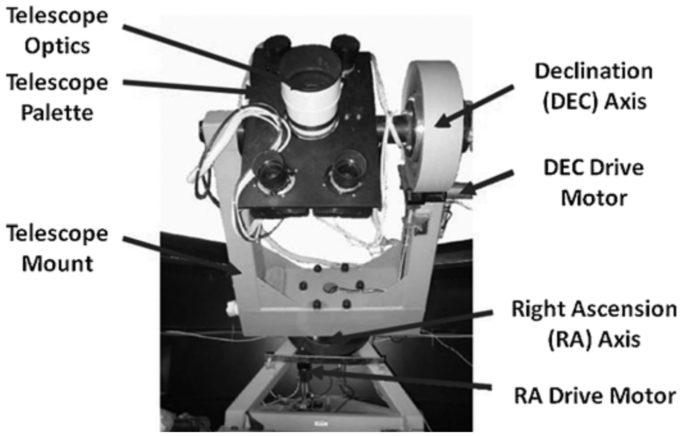


Fig. 1.1 Telescope system with mount and drive mechanisms

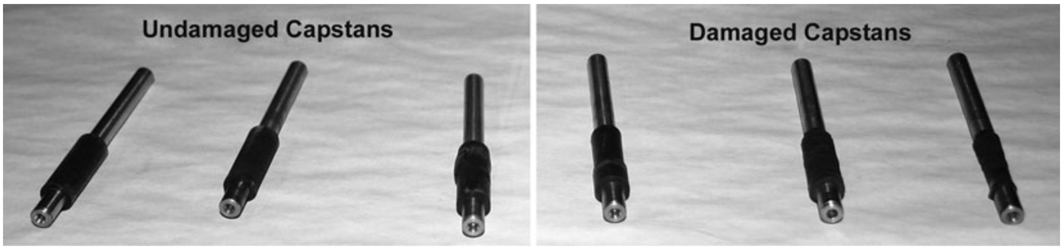


Fig. 1.2 Capstans for which the sets of acceleration time-histories, $\mathcal{X}^u$ and $\mathcal{X}^d$, were obtained

Data for the study are collected using a PCB Piezotronics model 352A24 accelerometer with a data sampling rate of 640 Hz over a time of 6.4 s, giving a total record length of 4,096 samples, for each of the ten tests executed, and for each of the six capstans illustrated in Fig. 1.2. This gives a total of 60 acceleration time-histories that were then divided into equal-size records, each having a length of 256 time points. Referencing the sets $\mathcal{X}^u = \{\boldsymbol{x}_i^u : \forall i = 1, 2, \ldots, R\}$ and $\mathcal{X}^d = \{\boldsymbol{x}_i^d : \forall i = 1, 2, \ldots, S\}$, $\mathcal{X}_i^u$ and $\mathcal{X}_i^d$ contain 256 time-points for all i, and $R = S = 480$. The reader is reminded that typical SHM applications present situations where $R \gg S$. However, in this application, access to both undamaged *and* damaged capstans and an out-of-service RAPTOR observatory with which to test the capstans was readily available, presenting a unique case where $R = S$.

Relating this problem to the robustness function in (1.6), the sets Φ^u and Φ^d are composed of vectors of the AR model coefficients associated with the structural system in its undamaged and damaged configurations, respectively. AR models provide a relatively straightforward means of accurately describing time-history data, and because the AR model coefficients represent the poles of the structural response, they are sensitive to changes in the dynamic response of the structure. The use of AR models is widespread in the SHM literature (see e.g. [14, 15]). However, a survey of the literature also indicates that the problem of selecting an appropriate AR model order remains a topic of importance in the SHM community [2, 16].

For the time-histories considered herein, having discretely sampled components $x_1, x_2, \ldots, x_{256}$, the corresponding AR models are computed as

$$x_i = \sum_{j=1}^{p} \varphi_j x_{i-j} + e_i, \tag{1.7}$$

where p is the order of the AR model, φ_j .s the jth AR model coefficient, and e_i is a white-noise random error at time i. The coefficients φ_j are computed using a least-squares technique or the Yule-Walker equations [17], with the number of coefficients being equal to the order of the AR model: a number that is typically several orders of magnitude smaller than the number of samples within the time-history.

To conduct the info-gap analysis, AR model coefficients are computed for each of the acceleration time-histories in the sets $\mathcal{X}^u$ and $\mathcal{X}^d$, and their respective means are computed as in (1.2) and (1.3). A damage classifier is then "trained" using the features contained in the set Φ^u. Thereafter, multiple subsets of samples (i.e. multiple $\mathbf{B}^u$'s and $\mathbf{B}^d$'s) are assembled from the info-gap models in (1.4) and (1.5) at monotonically increasing values of the horizon of uncertainty α. The samples within these subsets are classified as undamaged or damaged, using the aforementioned damage classifier, and ROC curves are constructed for each of the subsets. Then, for each value of α, the worst-case performance (i.e. minimum AUC) of the damage classifier is stored. This process is repeated for each of three damage classifiers $\mathcal{C} = \{$Mahalanobis Squared Distance, Principal Component Analysis, Singular Value Decomposition$\}$ (a discussion of these damage classifiers follows).

1.3.2 Robust Selection of Damage Classifier

This demonstration problem is centered on Part (iv) of the SHM paradigm presented in Sect. 1.1.1, Statistical Model Development for Feature Discrimination. The model employed to discern the state of the structural system can rely on a purely mathematical representation of the data as in [18] or instead rely on a high-fidelity, physics-based model as in [19]. For the present case, three mathematical representations employed as damage classifiers are examined: Mahalanobis Squared Distance, Principal Component Analysis, and Singular Value Decomposition. Brief explanations of each are given in the following sections, followed by a presentation of the results from the info-gap analyses of each.

For reference in the following sections, (1.8) defines the $(p \times R)$ matrix $\mathbf{Y}^u$ that contains the standardized AR model coefficients associated with the undamaged system state, where $\boldsymbol{\mu}^u$ is as defined in (1.2). Equation 1.9 is then the $(p \times p)$ covariance matrix of $\mathbf{Y}^u$, where it is noted that $\mathrm{E}[\cdot]$ is the expected value operator. Lastly, p is the selected order of the AR model employed to represent the time-history data (see (1.7)).

$$\mathbf{Y}^u = \left(\boldsymbol{\varphi}_1^u - \boldsymbol{\mu}^u, \ldots, \boldsymbol{\varphi}_R^u - \boldsymbol{\mu}^u\right) \tag{1.8}$$

$$\boldsymbol{\Sigma}^u = \mathrm{E}\left[(\mathbf{Y}^u)(\mathbf{Y}^u)^{\mathrm{T}}\right] \tag{1.9}$$

1.3.2.1 Mahalanobis Squared Distance

The most straightforward of the four damage classifiers employs the *Mahalanobis Squared Distance* (MSD) metric to quantify the similarity between a training data set and a data set of features obtained from an unknown system state. The formula for computing this metric is given as

$$D_M = \left(\boldsymbol{\varphi}^{uknown} - \boldsymbol{\mu}^u\right)^{\mathrm{T}} (\boldsymbol{\Sigma}^u)^{-1} \left(\boldsymbol{\varphi}^{uknown} - \boldsymbol{\mu}^u\right), \tag{1.10}$$

where $\boldsymbol{\varphi}^{uknown}$ is the feature vector being classified and $\boldsymbol{\mu}^u$ and $\boldsymbol{\Sigma}^u$ are as defined in (1.2) and (1.9), respectively. In this way, the classification problem reduces to one of scoring individual feature vectors, where the scores represent the deviation of each feature vector from the mean behavior of the set Φ^u. Thus, higher MSDs are assumed to be associated with the damaged system state.

1.3.2.2 Principal Component Analysis

Principal Component Analysis (PCA) operates by projecting multi-dimensional feature vectors onto a lower-dimensional subspace. This is accomplished by operating on $\mathbf{Y}^u$ with an $(p \times n)$ transformation matrix T as

$$X = T^{\mathrm{T}} \mathbf{Y}^u, \tag{1.11}$$

where n is less than p, which yields a reduced set of R vectors in n-dimensional feature space. If T is orthogonal (i.e. if $T^{\mathrm{T}} T = \mathbf{I}$), then $\mathbf{Y}^u$. can be reconstructed as

$$\hat{Y} = TX = TT^{\mathrm{T}} \mathbf{Y}^u. \tag{1.12}$$

The error in this reconstruction can be expressed as the Euclidian norm $\mathbf{Y}^u - \hat{Y}_2$, and the columns of T that minimize this error for a projection onto a subspace of dimension n, are the eigenvectors corresponding to the largest $\mathbf{n}$ eigenvalues of the covariance matrix of $\mathbf{Y}^u$ [20]. The classification then employs the transformation matrix T to operate on $\boldsymbol{\varphi}^{uknown}$ as in (1.12), where large residual error in the reconstruction indicates that $\boldsymbol{\varphi}^{uknown}$ is associated with the damaged system state.

1.3.2.3 Singular Value Decomposition

Singular Value Decomposition (SVD) is a method that has been proposed to discern whether a feature vector from an unknown system state should be assessed as damaged or undamaged [21]. To begin, consider the matrix

$$\mathbf{Z} = \left(\boldsymbol{\varphi}_1^u, \ldots, \boldsymbol{\varphi}_R^u, \boldsymbol{\varphi}^{unknown}\right), \tag{1.13}$$

where the quantities are as previously defined and linear independence among the $\boldsymbol{\varphi}_i^u$ is assumed. If it is assumed that linear independence between the $\boldsymbol{\varphi}_i^u$ and $\boldsymbol{\varphi}^{unknown}$ is indicative of damage, then the state from which $\boldsymbol{\varphi}^{unknown}$ emanates may be assessed by simply computing $rank(\mathbf{Z})$. Thus, if $rank(\mathbf{Z}) = R + 1$, then $\varphi^{unknown}$ is likely to be associated with the damaged condition and the converse is true for $rank(\mathbf{Z}) = R$. However, due for example to noise-contaminated data, assessing $\varphi^{unknown}$ is rarely as simple as computing $rank(\mathbf{Z})$. This difficulty is typically overcome by comparing the singular values of the matrices $\mathbf{Y}^u$ and $\mathbf{Z}$. up to a certain threshold, where large residual differences between the singular values of $\mathbf{Y}^u$ and $\mathbf{Z}$ are assumed to indicate damage.

1.4 Results

The robustness curves given in the left panel of Fig. 1.3 present the worst-case performances associated with the robust-optimal AR model orders, for each of the three damage classifiers. For each damage classifier, the robust-optimal AR model order is chosen by first examining the performance levels of all AR model orders at a horizon of uncertainty $\alpha = 1.0$. The AR model order that exhibits the most robust performance at this horizon of uncertainty (i.e. the highest AUC value at $\alpha = 1.0$) is then selected as the robust-optimal AR model order. *(For brevity, the details of this analysis are not included in the present document.)* For the MSD, PCA, and SVD damage classifiers, the robust-optimal AR model order is 20. Shifting focus to the right panel of Fig. 1.3, a limit of $AUC_{critical} = 0.9$ has been set (see (1.6)). Imposition of this limit means that when the performance of any damage classifier falls *below* $AUC = AUC_{critical}$, the corresponding value of its robustness becomes zero. Put another way, there is no permissible horizon of uncertainty α, for which this damage classifier is able to meet the minimum performance requirement $AUC_{critical}$. Conversely, when a damage classifier meets this minimum performance requirement, it is said to *satisfice* the minimum performance requirement [4].

It is noted that the value assigned to $AUC_{critical}$ uld ultimately need to take into account many aspects of the SHM system design (e.g. life-safety concerns, equipment costs). In the case of the RAPTOR observatories, there are no life-safety concerns, and equipment failures that could damage high-cost equipment are readily detected. According to the quantitative assessment provided by [11], $AUC = 0.9$ corresponds to a "highly accurate" diagnostic test, which is assumed herein to be adequate for this application.

For the present case, the most robust of the four algorithms is clearly the PCA damage classifier. Examination of Fig. 1.3 also highlights two central properties or axioms of IGDT. The first relates to the nesting property common to info-gap models of uncertainty in that as the horizon of uncertainty α, increases, the sets defined by info-gap models become more inclusive. It is noted that this property requires that the constructed info-gap models be convex in nature, which is the case for (1.4) and (1.5), but it is also a property typical to the types of uncertainties handled in IGDT [4]. As an example, consider Fig. 1.4 where it is seen that at $\alpha = \alpha_2$, the range of values that b_i^u may assume also includes that defined at $\alpha = \alpha_1$. . Likewise, the range of values that b_i^u may assume at $\alpha = \alpha_3$, includes that defined at $\alpha = \alpha_2$ (which also includes that defined at $\alpha = \alpha_1$). The implication of this property is that as the horizon of uncertainty increases, the worst-case performance of a model, which emanates from this nested set of model parameters, does not and moreover, *cannot* improve (but it can stay the same), producing the monotonically increasing robustness curves given in Fig. 1.3.

The second property is that of contraction [4], in that as α approaches a value of zero, the info-gap models reduce to a single set of parameters that describe "nominal models;" these nominal models have zero robustness to uncertainty. For the present case, the nominal models of the undamaged and damaged configurations are represented by the mean values of the

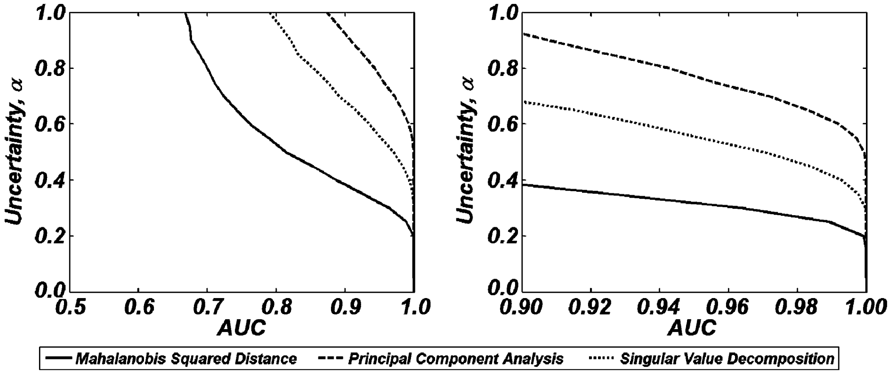


Fig. 1.3 Robustness curves for damage classifiers associated with robust-optimal AR model orders. The scale of the *right* figure is magnified 5X over that of the *left*, in the x-axis direction

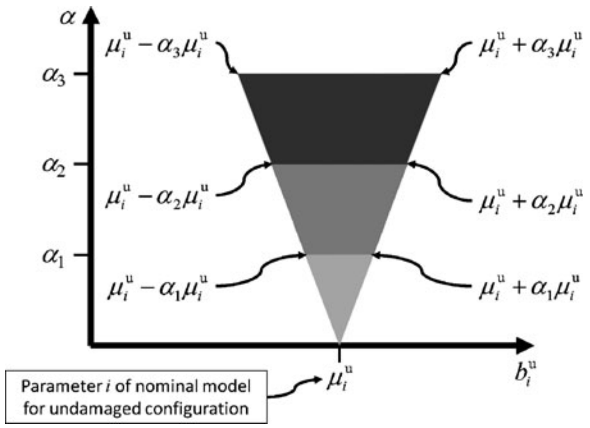


Fig. 1.4 Nesting property of info-gap models of uncertainty

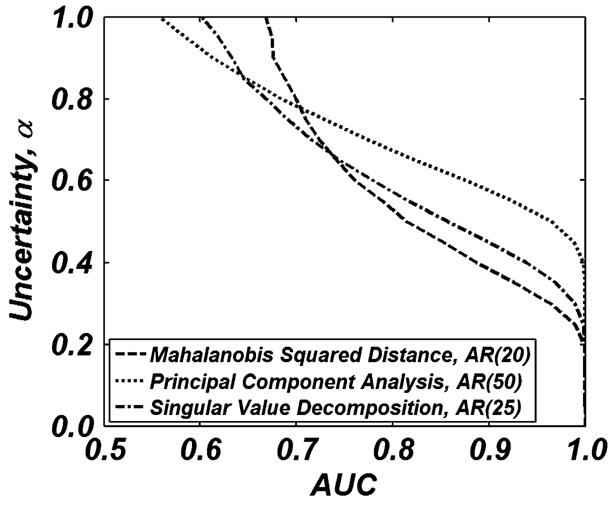


Fig. 1.5 Robustness curves for damage classifiers associated with hypothetically chosen AR model orders (for illustration purposes)

AR model coefficients $\boldsymbol{\mu}^u$ and $\boldsymbol{\mu}^d$, built from the time-histories contained in $\mathcal{X}^u$ and $\mathcal{X}^d$, respectively (given in Sect. 1.3.1). The contraction property is demonstrated in Fig. 1.3, in that the robustness curves reduce to a constant value of $AUC = 1.0$ at a value of $\alpha = 0$. While representing the best information available to the engineer, these nominal models also represent a precariously optimistic point-of-view in that *no* variability in the data may be tolerated without the potential of negatively affecting the quality of the model, represented as a reduction in the value of AUC.

1.4.1 *Hypothetical Example*

As a benefit to the readers, the authors propose the following hypothetical example derived from the info-gap analyses conducted in Sect. 1.3.2. An engineer is presented with the time-history data described in Sect. 1.3.1 and asked to assess the robustness of the MSD, PCA, and SVD damage classifiers against variability in the data. The engineer conducts a similar analysis as given in Sect. 1.3.2, but instead analyzes a subset of randomly selected AR model orders. A plot of the robustness curves from one of these subsets is given in Fig. 1.5, where the randomly selected AR model orders are $\mathcal{P} = \{20, 50, 25\}$ for $\mathcal{C} = \{\text{MSD,PCA,SVD}\}$, respectively.

While the authors are quick to admit that this is a contrived example, the results of this example demonstrate an important characteristic common to info-gap analyses, namely that of preference reversal. Upon examination of Fig. 1.5, it can be seen that the robustness curves associated with three damage classifiers cross at various points throughout the space. The crossing of these curves represents a point where the preference for one algorithm over another *changes* due to the engineer's

Table 1.1 Damage classifier preferences for robustness curves given in Fig. 1.5

Horizon of uncertainty range	Damage classifier preference[a]
$\alpha<0.2$	No preference
$0.2<\alpha<0.67$	PCA, SVD, MSD
$0.67<\alpha<0.78$	PCA, MSD, SVD
$0.78<\alpha<0.85$	MSD, PCA, SVD
$0.85<\alpha<1.0$	MSD, SVD, PCA

[a] Ordered left to right, with the more preferred approach to the left

increasing lack of knowledge about the data. In the present example, the robustness curves cross at three points within the horizon of uncertainty examined. This presents four regions in which damage classifier preferences may be specified (see Table 1.1), where it is noted that when $\alpha<0.2$, there appears to be no preference of one classifier over another. That is, when $\alpha<0.2$, the engineer may choose a classifier without worrying about negative consequences of variability in the data.

This example illustrates one of the considerable strengths of IGDT in that the engineer can, *in a defensible manner*, state his preference for one damage classifier over another by assessing the robustnesses of these damage classifiers to variability in the data. However, it must be noted that the question of what the appropriate horizon of uncertainty is remains a judgment call that is heavily influenced by the engineer's, policy-maker's, or regulatory body's aversion to risk.

An important result that can be seen in both Fig. 1.5 and Table 1.1 is that the PCA damage classifier, which starts out as the most robust algorithm (as evidenced by its high *AUC* as compared to the MSD and SVD damage classifiers), becomes the least robust algorithm as the horizon of uncertainty increases. Put another way, *the optimal damage classifier becomes sub-optimal in the face of uncertainty about the data*. This result is of paramount importance if the engineer suspects that the data may be, or may become subject to environmental and/or operational variability in the future.

1.5 Conclusions

A framework for assessing robustness of various *Structural Health Monitoring* (SHM) technologies is presented. This framework, anchored in *Information-Gap Decision Theory* (IGDT), offers a means of accounting for the engineer's lack of knowledge affecting the decision at hand, while not requiring a probabilistic treatment of said lack of knowledge. For the demonstration problems outlined above, this lack of knowledge corresponds to environmental and/or operational variability affecting data collected from a structure to be monitored by a SHM system; the engineer is tasked with designing a SHM system robust to this variability. Traditionally, accounting for this variability would require the engineer to collect enough data to formulate probabilistic descriptions (e.g. probability density functions) about these data. In the present case, the variability is treated non-probabilistically, and the worst-case performance for a given level of variability is determined. The results of Sect. 1.3.2 highlight properties typical to IGDT problems, providing the reader with a reasonable overview of IGDT and its application to SHM. While the applications demonstrated (i.e. selection of robust, damage-sensitive features and a robust damage classifier) are of a rather specific nature, the framework outlined above is general enough that its applicability could be extended to other problems within SHM.

The limitations of this framework stem from the fact that it is a tool that analyzes the robustness of one or more decisions, but does not actually make a decision. Such decisions are certainly better informed vis-à-vis IGDT analyses but ultimately rest with the engineer. Furthermore, if a closed-form, analytical expression for the robustness function cannot be derived, its evaluation either requires sampling of the info-gap models or the implementation of an optimization procedure to "search" for the worst-case performance, at *every* horizon of uncertainty desired. For the demonstration problems outlined above, this is a trivial concern as building AR models and constructing ROC-curves are relatively inexpensive operations. If instead the evaluation of the robustness function requires calls to a high-fidelity, physics-based model, this concern is no longer trivial. It can, however, be mitigated by employing surrogate or reduced-order models, keeping in mind that such models may introduced modeling inaccuracies that affect the evaluation of the robustness function.

Future research directions include applying this framework to simple laboratory structures that can be monitored in a controlled environment, where variability in the data can be introduced in a controlled manner (i.e. the horizon of uncertainty can be quantified). These laboratory tests would serve to validate the claim that a robust SHM system design should be preferred over an optimal (or nominal) SHM system design when faced with environmental and/or operational variability, paving the way for applications of this framework to real-world SHM system designs.

Acknowledgements The first author is grateful to Yakov Ben-Haim for his guidance during the initial phases of development of this work. The algorithmic development supporting this work was facilitated, in part, by the use of functions available in SHMTools, a software package developed at the Los Alamos National Laboratory//University of California San Diego Engineering Institute to aid in the construction of SHM processes. Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Los Alamos National Security, LLC, for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396.

References

1. Beck JL, Yuen K-V (2004) Model selection using response measurements: bayesian probabilistic approach. J Eng Mech 130(2):192–203
2. Saito T, Beck JL (2010) Bayesian model selection for ARX models and its application to structural health monitoring. Earthquake Eng Struct Dynam 39(15):1737–1759
3. Bishop CM (2006) Pattern recognition and machine learning. Springer, New York
4. Ben-Haim Y (2006) Info-gap decision theory: decisions under severe uncertainty, 2nd edn. Academic, Oxford, UK
5. Ben-Haim Y (2011) Info-gap decision theory: decisions under severe uncertainty. http://info-gap.com/. Retrieved 13 April 2011
6. Pierce SG, Worden K, Manson G (2006) A novel information-gap technique to assess reliability of neural network-based damage detection. J Sound Vib 293(1–2):96–111
7. Farrar CR, Worden K (2007) An introduction to structural health monitoring. Philos Trans Roy Soc A 365(1851):303–315
8. Kay SM (1998) Fundamentals of statistical signal processing, vol II, Detection theory. Prentice-Hall, Upper Saddle River
9. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1):29–36
10. Bradley AP (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recogn 30(7):1145–1159
11. Greiner M, Pfeiffer D, Smith RD (2000) Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Prev Vet Med 45(1–2):23–41
12. Stull CJ, Taylor SG, Wren J, Mascarenus DL, Farrar CR (2012) Real-time condition assessment of RAPTOR telescope systems. ASCE J Struct Eng, in press
13. Vestrand WT, Borozdin KN, Brumby SP, Casperson DE, Fenimore EE, Galassi MC et al (2002) The RAPTOR experiment: a system for monitoring the optical sky in real time. In: Proceedings of SPIE, vol 4845, Bellingham, pp 126–136
14. Farrar CR, Doebling SW, Nix DA (2001) Vibration–based structural damage identification. Philos Trans Roy Soc Math Phys Eng Sci 359(1778):131–149
15. Lynch JP, Sundararajan A, Law KH, Kiremidjian AS, Kenny T, Carryer E (2003) Embedment of structural monitoring algorithms in a wireless sensing unit. Struct Eng Mech 15(3):285–297
16. Figueiredo E, Figueiras J, Park G, Farrar CR, Worden K (2011) Influence of the autoregressive model order on damage detection. Int J Comput-Aided Civil Infrastruct Eng 26(3):225–238
17. Box GE, Jenkins GM (1976) Time series analysis: forecasting and control, Revised edn. Holden-Day, San Francisco
18. Park G, Figueiredo E, Farinholt KM, Farrar CR (2010) Time series predictive models of piezoelectric active-sensing for SHM applications. In: Proceedings of SPIE, vol 7650. Bellingham
19. Stull CJ, Nichols JM, Earls CJ (2011) Stochastic inverse identification of geometric imperfections in shell structures. Comput Method Appl Mech Eng 200(25–28):2256–2267
20. Duda RO, Hart PE, Stork DG (2001) Pattern classification, 2nd edn. Wiley, New York
21. Ruotolo R, Surace C (1999) Using SVD to detect damage in structures with different operational conditions. J Sound Vib 226(3):425–439

Chapter 2
Design of Uncertain Prestressed Space Structures: An Info-Gap Approach

Aurélien Hot, Scott Cogan, Emmanuel Foltête, Gaetan Kerschen,
Fabrice Buffe, Jérôme Buffe, and Stéphanie Behar

Abstract Uncertainty quantification is an integral part of the model validation process and is important to take into account during the design of mechanical systems. Sources of uncertainty are diverse but generally fall into two categories: aleatory uncertainties due to random processes and epistemic uncertainty resulting from a lack of knowledge or erroneous assumptions. This work focuses on the impact of uncertain levels of prestress on the behavior of solar arrays in their stowed configuration. In this context, snubbers are inserted between two adjacent panels to maintain contact and absorb vibrations during launch. However, under high excitation loads, a loss of contact between the two panels may occur. This results in impacts that can cause extensive damages to fragile elements.

In practice, the specific load configuration for which the separation of the two panels occurs is difficult to determine precisely since the exact level of prestress applied to the structure is unknown. An info-gap robustness analysis is applied to study the impact of this lack of knowledge on the prestress safety factor required to avoid loss of contact. The proposed methodology is illustrated using a simplified model of a solar array.

2.1 Introduction

In the field of structural dynamics, the mathematical model must not only be validated against tests, it must also be validated with respect to its final performance goal in the presence of uncertain parameters. Indeed, the required performance may be drastically affected even by a slight perturbation. Moreover, real environmental conditions and loads on a structure are sometimes unknown or can not be reproduced during tests. It is thus important to take these uncertainties into account in the design process.

One can distinguish two sources of uncertainty. The first class of uncertainties, called aleatory uncertainty, is due to the intrinsic randomness of the structure parameters such as variability of its physical properties, of the assembling process, of its service conditions (temperature, hydrometry, etc.), or again the variability of the measuring process. The second class pertains to epistemic uncertainties. They result from a lack of knowledge or erroneous assumptions, such as unknown values of stiffnesses, linearity assumptions, omission of contact elements or other nonlinear elements, unknown probability distributions, among others.

A. Hot • S. Cogan (✉) • E. Foltête
FEMTO-ST Institute – Applied Mechanics Department, 24 rue de l'Epitaphe, 25000 Besançon, France
e-mail: aurelien.hot@univ-fcomte.fr; scott.cogan@univ-fcomte.fr; emmanuel.foltete@univ-fcomte.fr

G. Kerschen
Department of Aerospace and Mechanical Engineering, University of Liege, Chemin des Chevreuils 1 (B52), 4000 Liége, Belgium
e-mail: g.kerschen@ulg.ac.be

F. Buffe
Centre National d'Etudes Spatiales, 18 Avenue Edouard Belin, 31401 Toulouse, Cedex 9
e-mail: fabrice.buffe@cnes.fr

J. Buffe • S. Behar
Thales Alenia Space, 100bd du Midi – BP99 – 06156, Cannes la Bocca, Cedex
e-mail: jerome.buffe@thalesaleniaspace.com; stephanie.behar@thalesaleniaspace.com

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_2,

Dealing with aleatory uncertainties, it is common practice to use probabilistic methodologies [1, 2] to quantify and propagate them through the system model. A classical approach consists in using a deterministic finite element model for the mechanical system and then adding the uncertainty through either parametric or non-parametric methods. When using the parametric approach, the uncertain physical parameters, e.g. Young's modulus or plate thickness, are characterized by their mean value, their standard deviation or by probability density functions, e.g. Gaussian or uniform distribution. However, the difficulty in quantifying the level of stochastic uncertainty in physical parameters has motivated the development of alternative non-parametric approaches. Soize proposes in [3] an overview of these methodologies, for which the uncertainties are no longer applied to physical parameters but directly introduced in the global randomized matrices through dispersion parameters.

However, using a stochastic approach for all types of problems would be too restrictive and inappropriate. Indeed, for epistemic uncertainties, where a large amount of information is missing, the probabilistic approach is unsuitable. Thus, to study these cases, one can use a second class of uncertainty models, based on non-probabilistic methods such as the interval arithmetic theory [4], the Fuzzy sets approach [5] or the Transformation method [6]. Moreover, to deal with problems subject to severe uncertainty, Ben-Haim proposed in [7] the Info-gap theory. The basic idea of this theory is to investigate the degree of lack of knowledge that can be tolerated while satisfying a critical level of performance.

In this paper, only the case of epistemic uncertainty due to lack of knowledge is tackled using the Info-gap theory. This method will be first briefly described and then applied to an uncertain prestressed structure.

2.2 The Info-Gap Theory

The Info-gap theory proposed by Ben-Haim [7] helps to quantify the robustness of model-based decisions to lack of knowledge in the system model. This approach provides a useful tool since the trade-off between performance and robustness can be investigated and different design configurations can be compared. The methodology has already been applied in many different fields. In structural dynamics analysis, studies concerning model updating [8], reliability [9], optimal sensors and exciters positioning [10] and design of civil structures in seismically active regions [11] have already been carried out. Moreover, applications of the Info-gap theory can also be found in medical [12] and management [13] applications, as well as elsewhere.

An Info-gap model quantifies the difference—the gap—between the information that is known, for example the nominal value of a parameter, and the information that has to be known in order to insure a critical level of performance. This non-probabilistic approach to uncertainty quantification is briefly described in the following sections.

An Info-gap robustness analysis is composed of four components:

- The *system model* defines the relation between the system inputs and outputs. In the case of structural dynamics, the system model consists in the equation of motion that links the inputs, the excitation force, to the outputs, displacement, velocity and acceleration, through the finite element model. The system model is denoted $S(\nu)$ and depends on various unknown parameters ν.
- The *uncertainty model* $\mathcal{U}$ represents the uncertainty in the variables ν. It consists in a collection of nested sets of uncertain events whose size is controlled by a horizon of uncertainty h. The larger this horizon, the more inclusive is the uncertainty model. A simple model could be an interval within which ν varies around a nominal value ν_0:

$$\mathcal{U}(\nu_0, h) = \{\nu \quad : \quad |\nu - \nu_0| \leq h\}, \quad h \geq 0 \tag{2.1}$$

- The *performance requirement* is a specification of a critical level, denoted P_{crit}, of the system performance $P(S(\nu))$, e.g. a maximum tolerable displacement or stress in the structure:

$$P(S(\nu)) \leq P_{crit} \tag{2.2}$$

- The *robustness to uncertainty* $\hat{h}$ is defined as the greatest value of the uncertainty parameter h for which the critical performance requirement is still insured. This can be expressed by:

$$\hat{h} = max\left\{h \quad : \quad \max_{\nu \in \mathcal{U}(\nu_0, \alpha)} P(S(\nu)) \leq P_{crit}\right\} \tag{2.3}$$

2.3 Application to a Simplified Solar Array System

2.3.1 *Presentation*

A specific difficulty encountered on solar arrays concerns the impacts that occurs between two adjacent panels. In order to limit this phenomenon, which can lead to significant damage, dedicated snubbers are inserted between the panels. The presence of the snubbers has two consequences. First, they lead to an increase in the local stiffness, thus limiting the relative displacements. Second, the induced prestress forces insures that contact is maintained between the two panels. However, for high input levels, loss of contact between the snubbers and the panels may arise and impacts still occur.

The Info-gap approach is used here to assess the robustness of the solar array system performance to epistemic uncertainties due to lack of knowledge in the induced prestress forces. The performance criterion here is the non-separation of two adjacent panels that will guarantee the integrity of the structure. The four components of an Info-gap analysis described above are presented in the following paragraphs, namely:

- System model: Simplified solar array system prestressed due to the presence of snubbers
- Performance requirement: Non-separation of the two panels in order to avoid impacts
- Uncertainty model associated with uncertain prestress values
- Robustness: What maximum level of uncertainty in the prestress can be tolerated while insuring that the panels do not impact one another?

2.3.2 *System Model*

The finite element model of the simplified solar array is depicted in Fig. 2.1a. The two plates are meshed with 3072 NASTRAN CQUAD4 elements. The stacking points are modeled using three CBAR beams that are linked to the plates through rigid elements RBE2. Only the masses of the two hinges are taken into account in this model by concentrated masses CONM2, also linked to the plates with rigid elements. The element properties are listed in Table 2.1, where E, ν, ρ, T, d and m are the Young's modulus, the density, the Poisson's ratio, the plates thickness, the beams cross section dimension and the

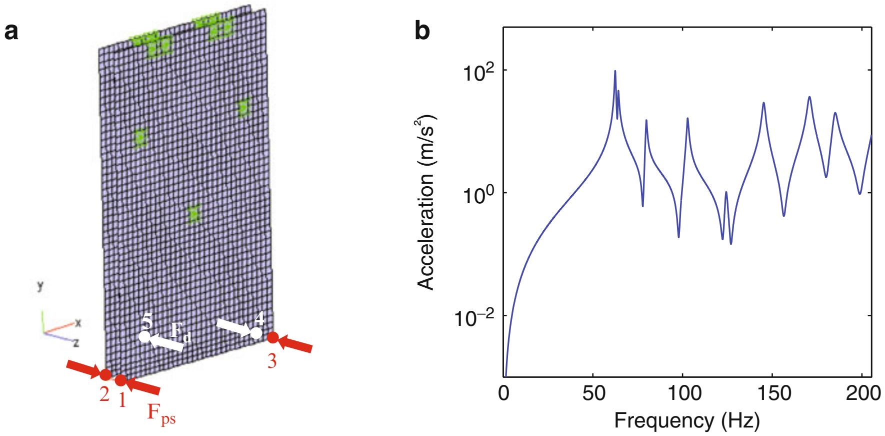


Fig. 2.1 System model. (**a**) Simplified solar array model FEM; (**b**) Frequency response of the system at node 1

Table 2.1 Element property of solar array model

CQUAD4	$E = 6.52e10$ Pa	$\rho = 2{,}761$ kg.m^{-3}	ν=0.34	T= 5 mm
CBAR	$E = 5.18e10$ Pa	$\rho = 3{,}453$ kg.m^{-3}	ν=0.34	d= 16 mm
CONM2	$m = 191$ g			

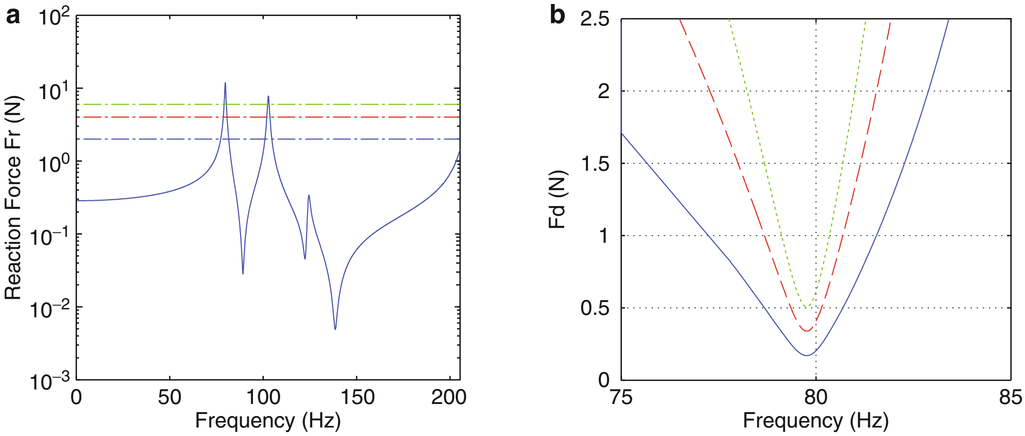


Fig. 2.2 Model performance. (**a**) Reaction force F_r at node 1 — $F_{ps} = 2$ N, - - - $F_{ps} = 4$ N, ··· $F_{ps} = 6$ N; (**b**) Separation zone $\mathcal{V}(\omega, F_d)$ for prestress values — $F_{ps} = 2$ N, - - - $F_{ps} = 4$ N, ··· $F_{ps} = 6$ N

concentrated masses value, respectively. The whole model contains 19,548 dofs. Two linear springs that model the snubbers behavior are inserted between 1 and 2 and 3 and 4 whose stiffnesses are denoted k_s. In a first step, $k_s = 1e5$ N/m.

A 1 N sine excitation F_d is applied at node 5 in the z-direction. The corresponding displacement frequency response of node 1 is plotted in Fig. 2.1b.

2.3.3 Performance Requirement

The performance requirement in this study consists in avoiding the separation of the two panels. Indeed this requirement insures that no impact can occur between the panels, so that the system integrity is preserved. This statement is expressed as a condition on local forces: separation occurs when the reaction force F_r of the snubber on the panel exceeds the applied static prestress load F_{ps}:

$$F_r \leq s_f F_{ps} \tag{2.4}$$

where s_f is a safety factor that enables to allow variations on the separation condition. A small safety factor guaranties a safe behavior but has very strict requirements, whereas a large value is easier to realize but is dangerous regarding the performance requirement. Unless otherwise mentioned, this s_f value remains equal to 1.

Figure 2.2a plots the reaction force F_r at node 1 obtained using the MPCF NASTRAN card, as well as three given threshold prestress force values. One remarks that two modes are liable to lead to the performance failure. We will focus on the one around 80 Hz, the highest amplitude one. However, the extension to a frequency band containing several modes can be readily made.

With the performance requirement defined, one can seek all realizable configurations (frequency, dynamic input force, prestress loads) that lead to failure. The set $\mathcal{V}(\omega, F_d)$, expressed in (2.5), corresponds to the points that bounds the performance. In other words, all the configuration points inside the parabola do not satisfy the performance requirement. Figure 2.2b shows these sets for the three prestress loads. For a given prestress load, the lowest dynamic force F_d for which the panels separate occurs for the resonance frequency. Indeed, with the increasing relative displacement of the two corners of the panels, the reaction force increases and exceeds the prestress threshold value. Moreover, one can also remark that, as expected, the higher the prestress load, the narrower the $\mathcal{V}(\omega, F_d)$ set and thus the smaller the opportunity of failure.

$$\mathcal{V}(\omega, F_d) = \{\{\omega, F_d\} \quad : \quad |F_r(\omega)| = s_f F_{ps}\} \tag{2.5}$$

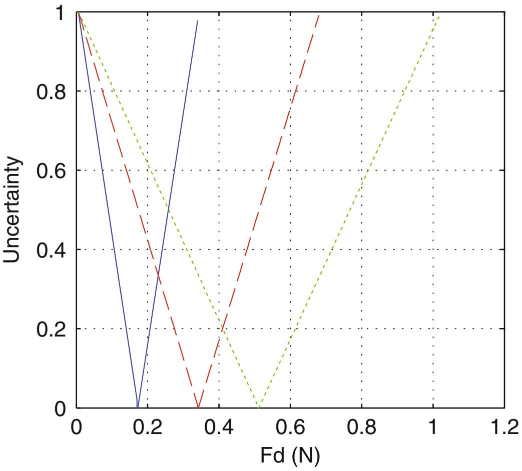


Fig. 2.3 Evolution of the critical separation point with uncertainty — $F_{ps} = 2$ N, - - - $F_{ps} = 4$ N, ··· $F_{ps} = 6$ N

2.3.4 *Uncertainty Model*

As mentioned in the presentation paragraph, the source of epistemic uncertainty in this example is the prestress load value F_{ps}. Indeed, it is assumed here that the level of prestress introduced by the snubbers is difficult to assess due to structural uncertainties (geometric defects, interface stiffness, etc).

Let $\widetilde{F_{ps}}$ be the best estimate of F_{ps}. This $\widetilde{F_{ps}}$ quantity is a known but perhaps unreliable estimate of the corresponding coefficient. In fact, we do not know by how much the estimated value is wrong. The uncertainties on F_{ps} are modeled here using an fractional-error Info-gap model:

$$\mathcal{U}(h) = \left\{ F_{ps} \quad : \quad \frac{|F_{ps} - \widetilde{F_{ps}}|}{\widetilde{F_{ps}}} \le h \right\}, \quad h \ge 0 \tag{2.6}$$

The inequality states that the fractional error of the estimate $\frac{|F_{ps} - \widetilde{F_{ps}}|}{\widetilde{F_{ps}}}$ is bounded by the horizon of uncertainty h. Since we do not know the magnitude of error, no realistic worst case is known, h is unbounded. When $h = 0$, there is no uncertainty and the estimation is correct. Whereas, as h increases the uncertainty set $\mathcal{U}(h)$ is more and more inclusive. Thus the Info-gap model is an unbounded family of sets of possible realizations of F_{ps}.

Figure 2.3 plots the evolution of the failure point in an uncertainty vs dynamic force diagram. For each curve, the first decreasing part corresponds to the worst case of performance failure, and the second increasing branch corresponds to the best case. As engineering structures are usually designed in the worst case, the study focuses on the curves left hand part.

For a null uncertainty h, the F_{ps} value is the nominal one and the failure point corresponds to the one seen in Fig. 2.2b. Regarding the nominal value of 2 N prestress load example, the separation occurs for a dynamic force of 0.17 N, which actually corresponds to the minimum of the corresponding curve in Fig. 2.2b. Then, the more the amount of uncertainty increases, the lower the dynamic force required for failure.

These curves also enable to evaluate the allowed uncertainty that still avoids failure. For example, if the system is excited with a dynamic force of 0.17 N, no uncertainty is tolerable if the prestress load is equal to 2 N. Thus, the prestressed force value must be exactly known. However, if $F_{ps} = 0.\ 4$ N, a 0.51 uncertainty value is now tolerable, and even 0.68 in the last case.

2.3.5 *Robustness*

We now define the robustness of the system model to uncertainty in the prestress load value. The robustness $\hat{h}$ is the greatest horizon of uncertainty h up to which the performance requirement is satisfied for all realizations of the uncertain quantity:

$$\hat{h} = max \left\{ h \quad : \quad \left(\max_{F_{ps} \in \mathcal{U}(h)} [F_r - s_f F_{ps}] \right) \le 0 \right\} \tag{2.7}$$

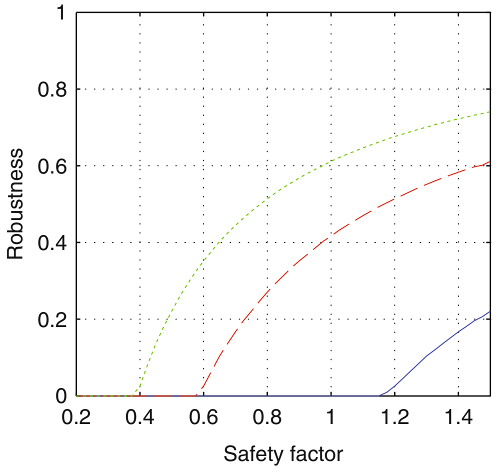


Fig. 2.4 Robustness evaluation for three different prestress loads — $F_{ps} = 2$ N, - - - $F_{ps} = 4$ N, ··· $F_{ps} = 6$ N

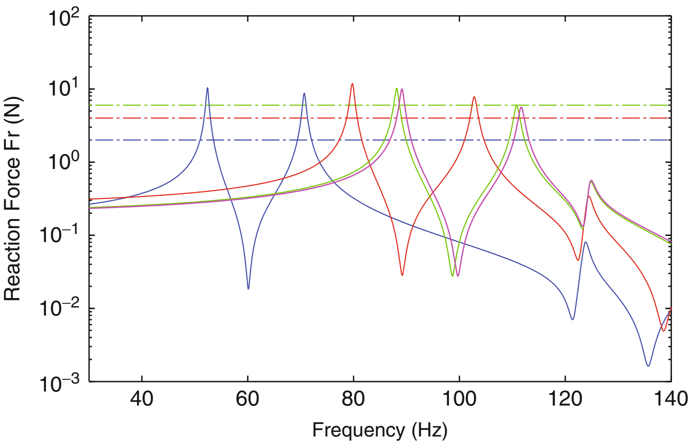


Fig. 2.5 Reaction force F_r at node 1 for different snubber stiffnesses values — $k_s = 1e4$ N/m, — $k_s = 1e5$ N/m, — $k_s = 1e6$ N/m, — $k_s = 1e7$ N/m - · - $F_{ps} = 2$ N, - · - $F_{ps} = 4$ N, - · - $F_{ps} = 6$ N

Figure 2.4 displays the robustness $\hat{h}$ vs the safety factor s_f, that remained equal to 1 up till now. Several conclusions can be drawn from these curves.

First, for low safety factors, the robustness remains null. This means that, in these cases, what ever the prestress configuration, failure of the performance requirement is certain.

Then, for the 6 N prestressed configuration, the robustness curve presents two different behaviors. For a safety factor below 0.8, the slope is quite steep, which means that for a small effort in increasing the safety factor the gain in robustness is important. Whereas, for safety factors above 0.8, the slope is now flat: a large variation of s_f only leads to a small profit of the robustness.

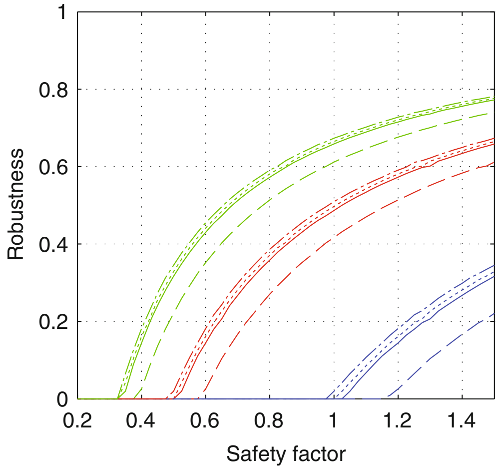
Finally comparing the three curves, one can note that, for all safety factor values, the higher the prestress load value, the greater the robustness to the sought performance, that is to say, avoiding separation between the panels.

2.3.6 *Comparison of Different Snubber Stiffnesses*

The previous paragraph compared the robustness of the system for different prestress loads and safety factor values. Now a third parameter is introduced: the snubber stiffness k_s. The previous results were obtained using a unique value: $k_s = 1e5$ N/m. Let us now assess the influence of k_s on the robustness function.

The reaction forces on the snubbers nodes are first recomputed for four different k_s values: $1e4$, $1e5$, $1e6$ and $1e7$ N/m. These results are plotted in Fig. 2.5 and one can observe, as expected, a softening or stiffening effect depending on the snubber stiffness value.

Fig. 2.6 Robustness evaluation for different snubber stiffnesses values — $k_s = 1e4$ N/m, - - - $k_s = 1e5$ N/m, ⋯ $k_s = 1e6$ N/m, - · - $k_s = 1e7$ N/m — $F_{ps} = 2$ N, — $F_{ps} = 4$ N, — $F_{ps} = 6$ N



The same procedure as above is then followed to estimate the robustness of the different designs. The corresponding results are displayed in Fig. 2.6. For all k_s values, the obtained curves present the same shape, which enables to directly compare the different robustness levels. It appears that, in this specific example, the case $k_s = 1e7$ N/m leads to the most robust design. However, the greater stiffness does not always means the more robust. Indeed, here the $k_s = 1e4$ N/m configuration is more robust than the $k_s = 1e5$ N/m one.

2.4 Conclusions

The objective of this work was to develop a model-based decision indicator for designing the prestress level required between the panels of a stowed solar generator in order to avoid damage due to impacts during launch. Given the presence of severe uncertainty in the prestress values induced by the introduction of dedicated snubbers, an info-gap approach was used to study the robustness of the critical design performance, namely the absence of separation, to lack of knowledge in the prestress states at each snubber location. Moreover, the impact of snubber stiffness on the robustness function was also investigated and the existence of an optimal value was demonstrated. The proposed methodology was illustrated using a simplified solar array model.

Before applying this methodology to an industrial structure, several limitations in the current approach must be addressed. Future work will consist in building a more accurate nonlinear model of the contact phenomena between the snubbers and the panels. Moreover, an estimation of how much prestress uncertainty must be tolerated in practice would be very useful for design purposes. This question will be examined by studying the impact of lack of knowledge in the mechanical properties of the structure on the prestress forces. Finally, an experimental validation of the methodology will be carried out on a simplified prototype structure.

Acknowledgements The work presented in this paper has been carried out with the generous support of the Centre National d'Etudes Spatiales (Toulouse, France) and Thales Alenia Space (Cannes La Bocca, France). The authors would also like to thank Professor Yakov Ben-Haim for his helpful remarks.

References

1. Schueller G (2007) On the treatment of uncertainties in structural mechanics and analysis. Comput Struct 85:235–243
2. Dempster AP (1968) Upper and lower probabilities generated by a random interval. Ann Math Stat 39(3):957–966
3. Soize C (2005) A comprehensive overview of a non-parametric probabilistic approach of model uncertainties for predictive models in structural dynamics. J Sound Vib 288:623–652
4. Moore RE (1979) Methods and applications of interval analysis. Prentice-Hall, London
5. Zadeh LA (1965) Fuzzy sets. Inf Control 8:338–353

6. Hanss M (2003) The extended transformation method for the simulation and analysis of fuzzy-parametrized models. Int J Uncertain Fuzz 11 (6):711–727
7. Ben-Haim Y (2001) Information-gap decision theory, decisions under severe uncertainty. Academic Press, Cornwall
8. Hemez F, Ben-Haim Y (2004) Info-gap robustness for the correlation of test and simulations in non-linear transient. Mech Syst Signal Process 18(6):1443–1467
9. Vinot P, Cogan S, Lallement G (2003) Approche non-probabiliste de fiabilité basée sur les modèles convexes. Mécanique Indutries 4:45–50
10. Vinot P, Cogan S, Cipolla V (2005) A robust model-based test planning procedure. J Sound Vib 288(3):571–585
11. Takewaki I, Ben-Haim Y (2005) Info-gap robust design with load and model uncertianties. J Sound Vib 288:551–570
12. Ben-Haim Y, Zacksenhouse M, Keren C, Dacso CC (2009) Do we know how to set decision thresholds for diabetes. Med Hypotheses 73:189–193
13. Ben-Haim Y, Laufer A (1998) Robust reliability of projects with activity-duration uncertainty. ASCE J Constr Eng Manag 124:125–132

Chapter 3
Robust Control Design for Uncertain Nonlinear Dynamic Systems

Sean P. Kenny, Luis G. Crespo, Lindsey Andrews, and Daniel Giesy

Abstract Robustness to parametric uncertainty is fundamental to successful control system design and as such it has been at the core of many design methods developed over the decades. Despite its prominence, most of the work on robust control design has focused on linear models and uncertainties that are non-probabilistic in nature. Recently, researchers have acknowledged this disparity and have been developing theory to address a broader class of uncertainties. This paper presents an experimental application of robust control design for a hybrid class of probabilistic and non-probabilistic parametric uncertainties. The experimental apparatus is based upon the classic inverted pendulum on a cart. The physical uncertainty is realized by a known additional lumped mass at an unknown location on the pendulum. This unknown location has the effect of substantially altering the nominal frequency and controllability of the nonlinear system, and in the limit has the capability to make the system neutrally stable and uncontrollable. Another uncertainty to be considered is a direct current motor parameter. The control design objective is to design a controller that satisfies stability, tracking error, control power, and transient behavior requirements for the largest range of parametric uncertainties. This paper presents an overview of the theory behind the robust control design methodology and the experimental results.

Keywords Robust design • Uncertainty quantification • Nonlinear control

3.1 Introduction

This paper presents an application of a control design and analysis methodology for arbitrary dynamic systems in the presence of parametric uncertainty. The specific classes of uncertainty considered here include multidimensional sets and probability density functions (PDFs). In the context of this paper, multidimensional sets are simply regions in parameter space where the stability and performance of the dynamic system are to be assessed. A bounding set in two dimensions, for example, could be a rectangle that is defined by its center and semi-diagonal. Other bounding sets are permitted, e.g., circular, and elliptical, with no restrictions on the dimensionality of the uncertain parameter space. Bounding set models are typically used when little is known about the true variability of the uncertain parameters, or when the stability and performance of the system need to be certified over a specific range of uncertainties. The other form of uncertainty used in this work is the PDF that describes the relative likelihood for a random variable to occur at a given point. Uncertainty models based upon PDFs require more information to create, but yield likelihood information of stability and performance of the controlled system. Despite the fundamental differences in these two forms of uncertainty, a unified framework [1] has been developed to accommodate both simultaneously.

S.P. Kenny (✉)
Senior Research Engineer, NASA Langley Research Center, Hampton, VA, 23681, USA
e-mail: Sean.P.Kenny@nasa.gov

L.G. Crespo
Senior Research Scientist, National Institute of Aerospace, 100 exploration way, Hampton, VA, 23666, USA

L. Andrews
Student, Mechanical Engineering Department, Old Dominion University, Norfolk, VA, 23508, USA

D. Giesy
Research Mathematician, NASA Langley Research Center, Hampton, VA, 23681, USA

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_3,

The primary objective of this paper is to apply the methods outlined in [1] to a hardware realization of the classic inverted pendulum on a cart. This objective was broken into two tasks, one for pure analysis of both stability and performance, and one that focused on controller design for the nonlinear pendulum system. The analysis task was to assess the robustness of the system to parametric uncertainty in two parameters, whereas the design task was to synthesize an optimally robust fixed-architecture controller for the nonlinear system.

3.2 Inverted Pendulum System

Figure 3.1 illustrates the experimental apparatus used in this work. The system features a cart driven by a rack and pinion mechanism using a DC motor. The cart position is measured using an optical encoder coupled to the rack via an additional pinion. A pendulum mounted on the cart is free to rotate and its angle is measured via an optical encoder. The pendulum arrangement has been modified to accommodate the needs of this research. The primary modification is the addition of a rigid coupling connected to the pivot point to allow an additional link to be added. The modified system consists of two collinear links, the original upper link with length $2L_1$, and a new lower link with length $2L_2$. The lower link has been designed to permit the addition of a movable point mass, M_p, whose distance from the pivot point, L_p, is considered as one of the uncertain parameters. A total of two uncertain parameters were considered in this study and were distinct in that one was physically realizable by changing the location of a movable mass on the pendulum, and one was a simulation model parameter corresponding to the relationship of current input to torque output of the DC motor used to actuate the system. Both uncertain parameters have a profound effect on the stability, performance, and controllability of the dynamic system. For example, the movable mass degrades controllability of the system by requiring larger control signals to maintain stability and achieve adequate performance as the mass moves away from the pivot point.

3.2.1 *Model of the Dynamic System*

The classic inverted pendulum is well-studied and the equations of motion can be found in many references, e.g., [2]. However, the configuration used in this work departs slightly from the conventional inverted pendulum, so the equations of motion are stated here for completeness.

The input force to the system is generated by applying a voltage to a DC motor to move the cart. The relationship between voltage, V , and applied force, F, can be represented as:

$$
\begin{aligned}
V &= I_m R_m + K_m K_g \omega_g = I_m R_m + K_m K_g \frac{\dot{X}_c}{r} \\
T &= K_m K_g I_m \\
F &= \frac{K_m K_g V}{R_m r} - \frac{K_m^2 K_g^2 \dot{X}_c}{R_m r^2}
\end{aligned}
\tag{3.1}
$$

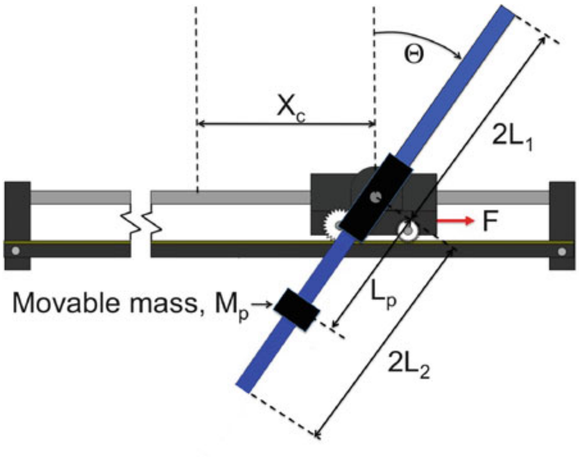


Fig. 3.1 Inverted pendulum with movable mass

where K_m is the DC motor torque constant, K_g is the planetary gearbox gear ratio, ω_g is the angular velocity of the pinion gear, R_m is the motor armature resistance, I_m is the motor armature current, r is the motor pinion diameter, and V is the applied voltage. Using the relationships given in (3.1) together with the fundamental physics results in the following nonlinear equations of motion:

$$\ddot{X}_c = \frac{VNW - \dot{X}_c NW^2 + \cos(\Theta) g \sin(\Theta) R_m P^2 + \dot{\Theta}^2 \sin(\Theta) R_m NP}{R_m \left(NM - \cos^2(\Theta) P^2\right)} \tag{3.2}$$

$$\ddot{\Theta} = -\frac{P\left(g \sin(\Theta) R_m M + \cos(\Theta) \dot{\Theta}^2 \sin(\Theta) R_m P + \cos(\Theta) V W - \cos(\Theta) \dot{X}_c W^2\right)}{R_m \left(NM - \cos^2(\Theta) P^2\right)} \tag{3.3}$$

where

$$W = \frac{K_g K_m}{r} \tag{3.4}$$

$$P = M_2 L_2 + M_p L_p - M_1 L_1 \tag{3.5}$$

$$N = \frac{4}{3}\left(M_2 {L_2}^2 + M_1 {L_1}^2\right) + M_p {L_p}^2 \tag{3.6}$$

$$M = M_c + M_p + M_1 + M_2 \tag{3.7}$$

X_c is the position of the cart and Θ is the angle of the pendulum. The angle Θ is measured counter-clockwise with zero defined when link one is pointing vertically up. Note that with this definition of Θ, the angular displacement shown in Fig. 3.1 is actually negative. The remaining parameters above are defined as follows: M_c is the mass of the cart, M_1 is the mass of link one, M_2 is the mass of link two, L_1 is the half-length of link one, L_2 is the half-length of link two, and g is the gravitational acceleration. The MKS system of units are used throughout the paper. Furthermore, it should be noted that equations given in (3.2) and (3.3) neglect damping.

3.2.2 Linear System Model

A general representation of the nonlinear equations given in (3.2) and (3.3) can be given as

$$\dot{X} = \mathscr{F}(X, U, p) \tag{3.8}$$

where $\mathscr{F}$ is a nonlinear function of the state vector X, the control input U, and the vector of *uncertain parametersp*. In this system, the vector of uncertain parameters is $p = [L_p, K_m]$.

For control design purposes, this nonlinear plant is linearized about the equilibrium point $(\bar{x}, \bar{u})$ satisfying $\mathscr{F}(\bar{x}, \bar{u}, p) = 0$. Deviations from the equilibrium values $\bar{x}(p)$ and $\bar{u}(p)$ are written as lowercase letters hereafter, e.g., $X = \bar{x} + x$ and $U = \bar{u} + u$. Linearization of (3.8) about this equilibrium point leads to the system

$$\dot{x} = A(p)x + B(p)u + \eta(x, u, p) \tag{3.9}$$

where

$$A(p) = \left.\frac{\partial \mathscr{F}}{\partial X}\right|_{\bar{x}(p), \bar{u}(p)} \qquad B(p) = \left.\frac{\partial \mathscr{F}}{\partial U}\right|_{\bar{x}(p), \bar{u}(p)} \tag{3.10}$$

and $\eta(x, u, p)$ contains higher-order terms. Note that A, B and the high-order terms depend on p. In a sufficiently small neighborhood of the equilibrium point the effect of the higher-order terms is negligible. The Linear Time Invariant (LTI) representation of the plant results from dropping the higher-order terms in (3.9). This LTI system can be written as

$$\begin{aligned}\dot{x} &= A(p)x + B(p)u \\ y &= Cx\end{aligned} \tag{3.11}$$

where $A(p)$ is the system matrix, $B(p)$ is the input matrix, C is the output matrix, x is the state vector, and u is the input. This is applied to the inverted pendulum by choosing the state vector to be $x = [x_c, \theta, \dot{x}_c, \dot{\theta}]^\top$ and the control input to be the applied voltage; i.e., $u = V$. The equilibrium point about which we linearize is $\bar{x}(p) = [0, 0, 0, 0]^\top$ and $\bar{u}(p) = 0$. With these choices, the LTI representation matrices are:

$$A(p) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{gP^2}{NM-P^2} & -\frac{NW^2}{R_m\left(NM-P^2\right)} & 0 \\ 0 & -\frac{gPM}{NM-P^2} & \frac{PW^2}{R_m\left(NM-P^2\right)} & 0 \end{bmatrix} \tag{3.12}$$

$$B(p) = \begin{bmatrix} 0 \\ 0 \\ \frac{NW}{R_m\left(NM-P^2\right)} \\ -\frac{PW}{R_m\left(NM-P^2\right)} \end{bmatrix} \tag{3.13}$$

and

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{3.14}$$

3.3 Control Architecture

There have been numerous control architectures developed and successfully implemented on similar inverted pendulum systems. The objective of this work is not to introduce a new control architecture, but instead use one that is familiar to many with control design experience and then demonstrate methods to optimally robustify this fixed architecture controller to parametric uncertainties. The mathematical framework for robustness analysis and control tuning, covered in Sect. 3.4, is capable of handling a wide array of control architectures and plant models, linear or nonlinear, with arbitrary parameter dependence. The primary objective of the control system is for the cart to track a periodic reference signal while maintaining stability and adequate performance. One period of the commanded reference signal is given in Fig. 3.2. See Sect. 3.4.3 for a detailed description of performance requirements.

3.3.1 Linear Quadratic Regulator Control with State Estimation

A Linear Quadratic Regulator (LQR) control together with state estimation constitute the control architecture. The estimator was necessary since only the two positional states (cart position and pendulum angle) of the four states were measured

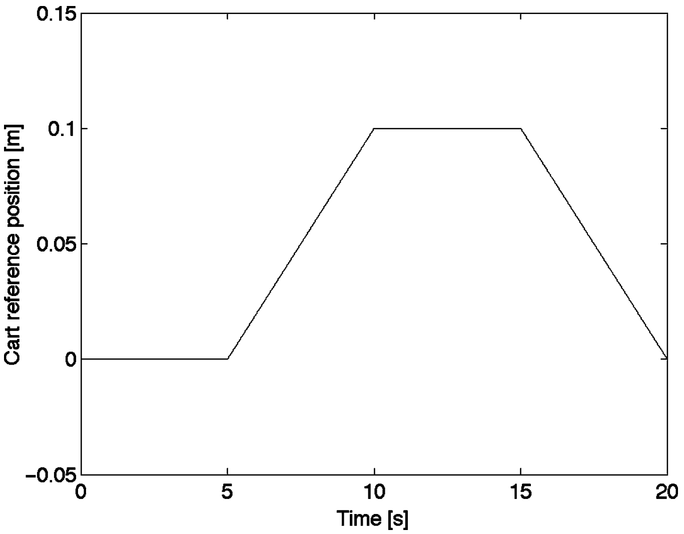


Fig. 3.2 Cart position reference signal

outputs. Conventional steady-state implementations of LQR control works by multiplying the states by a constant gain matrix, K, to produce the plant control signal. The gain matrix is generated by minimizing the following cost function:

$$J = \int_0^\infty [x^\top Qx + u^\top Ru]\; dt \tag{3.15}$$

In this cost function, weighting matrices $Q \geq 0$ and $R > 0$ are used to penalize deviation of the states from zero and penalize control energy. Choosing these weighting matrices is the fundamental task of the control designer. In the results to follow, these weighting matrices were chosen using a numerical optimization strategy.

Two additional features were added to the control architecture to achieve better tracking and steady state error performance. The first was the addition of a feed-forward control term using the known reference command, and the second was an integral error feedback term on cart position error. The feed-forward command was added to regulate the system about the reference command instead of regulating it about $x(t) = 0$. This resulted in the following control signal:

$$u(t) = -K(x - x_{ref}) = -Kx + Kx_{ref} \tag{3.16}$$

where x_{ref} is the four-element state reference vector. The second feature is an integral error feedback, which is added by augmenting the nominal state vector with an additional element. This additional element is an error between the cart's actual position and its commanded position. The total control signal is:

$$u(t) = -K(x - x_{ref}) - K_i e \tag{3.17}$$

where e is defined as:

$$e = \int_0^t (x_{c_{ref}}(\tau) - x_c(\tau))\; d\tau \tag{3.18}$$

with $x_{c_{ref}}$ defined as: $x_{c_{ref}} = [1, 0, 0, 0]x_{ref}$. Note that integral action is only added to the cart error signal and not to the full measurement vector. The scalar gain term, K_i, is the last element in the gain vector obtained by solving an LQR problem on an augmented system that includes the error state.

The remaining element of the control system is the state estimator which is a standard implementation of the Kalman filter. As is standard, the state estimate, $\hat{x}$, is used to replace x in (3.17) for implementation on the inverted pendulum hardware. In the design and analysis that follows, the state estimator parameters are fixed. This is consistent with current practices, since the true values of the parameters are not known prior to implementation. The Kalman filter gains in this work were computed according to the system performance when the uncertainties assume their nominal values.

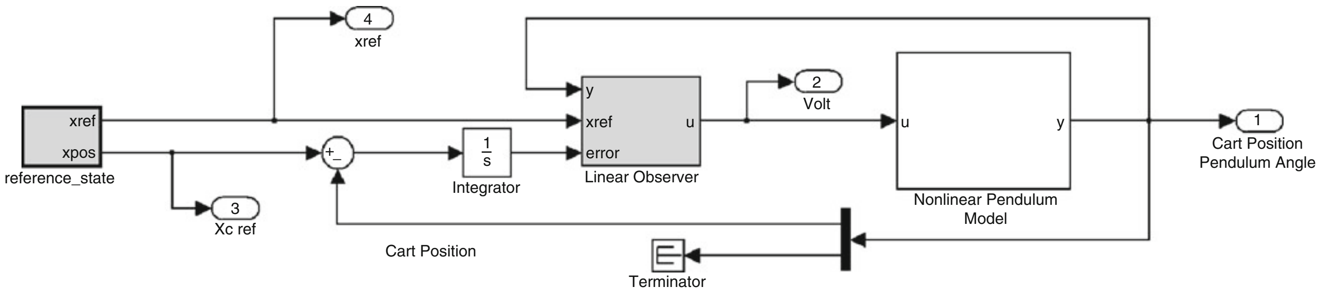


Fig. 3.3 Simulink model of the control architecture

A Simulink®; implementation of the control architecture is shown in Fig. 3.3. A linear representation of this system is as follows:

$$\begin{bmatrix} \dot{\hat{x}} \\ \dot{x} \\ \dot{e} \end{bmatrix} = \begin{bmatrix} A - LC - BK & LC & -BK_i \\ -BK & A & -BK_i \\ 0 & -C_1 & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ x \\ e \end{bmatrix} + \begin{bmatrix} BK \\ BK \\ C_1 \end{bmatrix} x_{ref} \tag{3.19}$$

where L is the Kalman gain matrix, and $C_1 = [1, 0, 0, 0]$.

3.4 Mathematical Framework for Robustness Analysis and Control Tuning

3.4.1 Robustness Analysis

A brief introduction to the methodology proposed in [1] is presented next. Consider the closed-loop dynamic system

$$\dot{X} = \mathscr{F}(X, U, p, r) \tag{3.20}$$

where r is a generic representation of a reference signal, with x_{ref} in (3.19) as a specific realization of the reference signal.

The control input can be represented as

$$U = U(X, r, K) \tag{3.21}$$

where the function U refers to a particular control structure and $K \in \mathbb{R}^{\dim(K)}$ is the set of free parameters, to be chosen by the control engineer (e.g., control gains), that fully prescribe the controller. The controller is deemed acceptable if the closed-loop response satisfies a set of requirements. These requirements, which are represented by a set of inequality constraints on selected output metrics, depend on the uncertain parameter vector p. The control system is deemed acceptable for a given realization, p, of the uncertain parameters if all inequalities are satisfied. Specifically, the satisfaction of the vector inequality

$$g(p, r, K) < 0 \tag{3.22}$$

implies the satisfaction of all the requirements,[1] and where $g \in \mathbb{R}^{\dim(g)}$. These constraints partition the uncertain parameter space into two sets, the failure domain $F = \{\, p : g_i(p) \geq 0, \text{ for some } i\}$, where at least one requirement is violated, and the safe domain $S = \{\, p : g(p) < 0\}$, where all requirements are satisfied. The term "satisfactory performance", which is to be

---

[1] For the reminder of the paper is assumed that vector inequalities hold component-wise, vector and set super-indices denote a particular realization, and vector sub-indices denote a particular component, e.g., p_i^j is the ith component of $p^j \in \mathbb{R}^{\dim(p)}$.

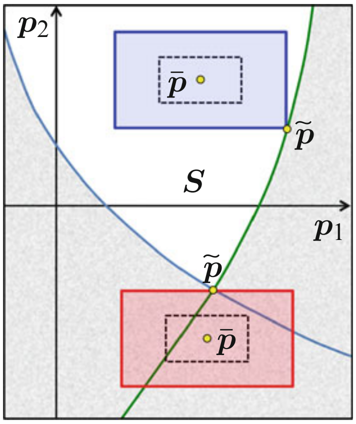


Fig. 3.4 Maximal deformations in the safe and failure domains

captured in g, is broad in scope and refers to acceptable ranges of variation in metrics of interest. These metrics may correspond to specific performance specifications as well as to basic notions of goodness. For instance, if $y(p, t)$ is the system output to a step-input and t_{as} is the admissible settling time, the corresponding requirement function is $g_{settling} = t_s - t_{as}$ where t_s is the minimal time for which $y(p, t) - y(p, \infty) \leq 0.99$ for all $t > t_s$.

The parameter realization $\bar{p} \in \mathbb{R}^{\dim(p)}$, called the nominal operating conditions point, denotes our best estimate of the actual value of p. The objective of this robustness analysis is to determine the largest deviation of p from $\bar{p}$ for which all the requirements are met. In this paper, such a deviation is prescribed as a hyper-rectangular set of fixed proportions. Note that the hyper-rectangle centered at $\bar{p}$ with semi-diagonal vector $m > 0$ is given by

$$R(\bar{p}, m) = \{p : \bar{p} - m \leq p \leq \bar{p} + m\} \tag{3.23}$$

The notions required to calculate this deviation are introduced next.

A *homothetic* deformation results from a uniform, radial expansion or contraction of the space about a fixed point. The distance from any point in the space to the fixed point changes by a factor α after the deformation. This factor is called the *similitude ratio* of the homothetic deformation. Note that if α is greater than 1, the deformation is an expansion, while if α is less than 1, the deformation is a contraction. A *reference set*, denoted as $\Omega \subset \mathbb{R}^{\dim(p)}$, is deformed with respect to its geometric center $\bar{p}$. Intuitively, one can imagine that Ω is being deformed with respect to $\bar{p}$ until its boundary just touches the failure domain boundary ∂F. This deformation is called the *maximal deformation*. The set resulting from this deformation, denoted as $M \subseteq \mathbb{R}^{\dim(p)}$, is the *maximal set*. A *critical parameter value* (CPV), denoted as $\tilde{p} \in \mathbb{R}^{\dim(p)}$, is (one of) the point(s) where the maximal set touches ∂F. This point is a worst-case uncertainty combination associated with Ω. The *critical similitude ratio* (CSR), denoted by $\tilde{\alpha} \in \mathbb{R}$, is the similitude ratio of that deformation and is a non-dimensional metric proportional to the separation between the point $\bar{p}$ and the failure domain. The *parametric safety margin* (PSM), denoted as $\rho \in \mathbb{R}$ is its dimensional equivalent. The values taken on by the CSR and the magnitude of the PSM are proportional to the size of the maximal set and measure the separation between $\bar{p}$ and ∂F. The PSM is assigned a negative value if the controller does not even satisfy the requirements at $\bar{p}$. If the PSM is zero, the controller exhibits no robustness because there is an arbitrarily small deviation from $\bar{p}$ leading to a requirement violation. If the PSM is positive, the requirements are satisfied at $\bar{p}$ and in its vicinity. The larger the PSM, the larger the Ω-shaped vicinity.

Figure 3.4 illustrates some of these variables when two requirements are present. The CPV, $\tilde{p}$, the nominal parameter point, $\bar{p}$, the reference set (region within dashed line) and the maximal set corresponding to deformations in the safe (blue) and failure (red) domains are shown. Note that by construction, the maximal set is fully contained in the safe or failure domains.

The formulation enabling sizing the maximal deformation is as follows. If $\Omega = R(\bar{p}, m)$, the CPV, CSR, PSM and maximal sets corresponding to the ith requirement are

$$\tilde{p}^i = \underset{p}{\operatorname{argmin}} \left\{ \|p - \bar{p}\|_m^\infty : \sigma_i g_i(p, r, K) \geq 0, p_{\min} \leq p \leq p_{\max} \right\} \tag{3.24}$$

$$\tilde{\alpha}_i = \frac{\|\tilde{p}^i - \bar{p}\|_m^\infty}{\|m\|} \tag{3.25}$$

$$\rho_i = \sigma_i \tilde{\alpha}_i \|m\| \tag{3.26}$$

$$M^i = R(\bar{p}, \tilde{\alpha}_i m) \tag{3.27}$$

where $\|x\|_m^\infty = \max_i\{|x_i|/m_i\}$, $p_{\min}$ and $p_{\max}$ bound the region of interest, and $\sigma_i = 1$ if $g_i(\bar{p}) < 0$, otherwise $\sigma_i = -1$. Note that σ_i can reverse the inequality sign of the constraint. This is needed since deformations centered at a nominal point in the safe domain require a different formulation than those where the nominal point is in the failure domain (See Fig. 3.4). While in the former case all points within the maximal set must satisfy all constraints, in the latter case all points must violate at least one constraint. The value of the controller's parameters K and the commands in r are set a priori and kept fixed during the analysis. While K can be prescribed according to any control design practice, r is prescribed to achieve the ultimate goal of the system.

The CPV, CSR, PSM and maximal sets associated with all requirements are given by (3.24)–(3.27) after replacing g_i with $\max_i(g_i)$. The resulting values are qualified as *aggregate* since they take all requirements into account. If $\bar{p} \in F$, these aggregate values coincide with those corresponding to the individual requirement attaining the smallest PSM, e.g., $\tilde{\alpha} = \tilde{\alpha}_j$ where $j = \text{argmin}(\rho_i)$. Therefore, for a maximal safe set, the aggregate PSM is the worst-case individual PSM.

The calculation of the maximal deformation requires solving an standard optimization problem. In this work, a gradient-based nonlinear constrained optimization was used. This problem class is, in general, non-convex when the dependency of the vector g_i of requirement functions on the uncertainty p is nonlinear. In any non-convex optimization problem there is always the possibility of convergence to a non-global optimum. When this occurs, the CPV resulting from the numerical optimization is mistaken and the corresponding maximal set intrudes into the failure region. This intrusion yields to a spuriously larger PSM. Absolute guarantees are not possible, but a variety of algorithmic safeguards can be used to deal with this deficiency. For instance, g can be evaluated at a few sample points in the maximal set and if one happens to fall into the failure domain, it can be used as an initial condition in a subsequent optimization. Requirements functions with a polynomial parameter dependency can be handled using the methods in [3]. These methods, which are based on linear matrix inequalities and Bernstein expansions, do not suffer from this potential drawback and guarantee convergence to actual maximal deformation. Notice however, that in most practical applications the explicit functional form of g is unknown. This intrinsic feature of the problem precludes the usage of convexity conditions from the outset.

3.4.2 *Control Tuning*

Ultimately, a good controller should satisfy the stability and performance requirements imposed upon the closed-loop system with sufficiently large robustness margins. The PSM and CSR introduced above are examples of such margins. A computational framework, also used in [4], enables searching for the set of controller gains that maximize these margins. This framework is introduced next.

Equation 3.22 indicates that the geometry of the failure domain, thus the CSR and the PSM, is a function of the control parameter K. Recall that this parameter is kept fixed when a robustness analysis is performed. A control design challenge is to find the value of K for which the closed-loop system exhibits satisfactory robustness characteristics. Specifically, an optimal design will be one that has the largest overall PSM for the given control structure. This entails maximizing the separation between the nominal operating condition point $\bar{p}$ and the failure domain F, so the largest maximal set is attained. Since the resulting controllers tolerate the largest deviations from nominal operating conditions before violating any requirement, they are called *optimally robust*. In the presence of conflicting design objectives, an optimally robust controller trades off the PSMs of all individual requirements to attain the largest aggregate PSM.

Two optimization-based formulations for pursuing the optimally robust design point $\tilde{K}$ are presented next. The first one evaluates robustness via the exact CSR. The second one uses an approximation to the CSR rooted in a multi-point approximation to the maximal set. The approximate nature of this formulation may not lead to controllers with the intended features. However, its relaxed computational demands may justify the potential drawbacks associated with it.

3.4.2.1 Enlarging the Maximal Set

A formulation based on the exact overall CSR is given by

$$\tilde{K} = \underset{K}{\text{argmin}} \{\tilde{\alpha}(K) : \ h(\bar{p}, K) > 0\} \tag{3.28}$$

where the vector constraint $h(K, \bar{p}){>}0$ is used to enforce additional characteristics to the controller based on standard control metrics. For instance, $h = PM(K, \bar{p}) - 60^\circ{>}0$ ensures that the feedback loop corresponding to the plant under nominal operating conditions has a phase margin of more than 60°. Examples of other performance/stability metrics are time delay, settling time, disturbance rejection, noise attenuation, reference tracking, and control energy. These metrics can be extracted from the time and/or frequency domains and may correspond to linear and nonlinear representations of the plant dynamics.

Recall that calculating the CSR $\tilde{\alpha}$ entails solving an optimization problem. Therefore, this formulation has an optimization in the outer loop and another one in the inner loop. While the outer loop searches for the robustly optimal gains $\tilde{K}$, the inner one searches for the CSR corresponding to the design point under consideration. Nested optimizations, such as (3.28) and those resulting from worst-case-based design policies, impose stringent computational demands. Such demands can be substantially mitigated by using multi-point approximations.

3.4.2.2 Enlarging an Approximation to the Maximal Set

The variables $\hat{K}$ and $\hat{\alpha}$, which are estimates of $\tilde{K}$ and $\tilde{\alpha}(\tilde{K})$ respectively, are given by

$$\langle \hat{K}, \hat{\alpha} \rangle = \underset{K,\alpha}{\operatorname{argmin}} \left\{ \alpha\colon \max_{i \le \dim(g), j \le n} \left\{ g_i\left(\bar{p} + \alpha(p^j - \bar{p}), r, K\right) \right\} {<} 0,\ \alpha \ge 0,\ h(\bar{p}, K){>}0 \right\}, \tag{3.29}$$

where $\{p^1, p^2, \ldots, p^n\}$ are parameter-points on the surface of $R(\bar{p}, m)$. These points should be uniformly distributed over the surface. Note that this formulation replaces the inner optimization loop in (3.28) with a multi-point constraint over parameter points lying on the surface of a maximal set estimate. Equation 3.29 may lead to suboptimal designs, for which $\hat{\alpha} > \tilde{\alpha}(\tilde{K})$ and $M(\hat{K}) \subset M(\tilde{K})$, because the satisfaction of the multi-point constraint does not guarantee the containment of the maximal set estimate by the safe domain. The formulation of Sect. 3.4.1 can be used to determine if this anomaly has occurred.

Since the robustness analysis and tuning procedures above identify the critical combination of uncertainties and/or failures (i.e., the CPV) the particular control architecture is more sensitive to, we expect the resulting controllers to be *safer* than those validated by brute force Monte Carlo simulations.

3.4.3 Application to the Inverted Pendulum

The components of the uncertain parameter p are the distance, L_p, of the mass M_p to the pivot point, and the motor torque constant K_m. These parameters are free to assume any non-negative value, i.e., $p_{\min} = [0,\ 0]^\top$, $p_{\max} = [0.\ 330\ \text{m},\ \infty]^\top$. The nominal operating condition point is $\bar{p} = [0.165\ \text{m},\ 0.00767\ \text{N-m/Amp}]^\top$. The robustness of the controller to uncertainty in the value of these two parameters is evaluated and improved using the methods above. The reference command is given by $r = [x_{\text{ref}}, \dot{x}_{\text{ref}}, 0, 0]^\top$, where x_{ref} is the periodic signal shown in Fig. 3.2. The design parameters K are diagonal elements of the Q and R matrices of the LQR formulation of Sect. 3.3. Denote by K_{base} the gains of a baseline controller attaining satisfactory performance at $p = \bar{p}$, and t_f the final simulation time. The closed-loop requirements defining the constraint function g are as follows:

1. **Command tracking**: the requirement $g_1 < 0$, where

$$g_1 = \int_0^{t_f} \left(x_{c_{ref}}(t) - x_c(p, r, K, t)\right)^2 dt - \gamma_1 \int_0^{t_f} \left(x_{c_{ref}}(t) - x_c(\bar{p}, r, K_{\text{base}}, t)\right)^2 dt \tag{3.30}$$

and $\gamma_1 > 1$ ensures satisfactory tracking of the reference command.

2. **Actuation**: the requirement $g_2 < 0$, where

$$g_2 = \int_0^{t_f} u^2(p, r, K, t) - \gamma_2 u^2(\bar{p}, r, K_{\text{base}}, t) dt \tag{3.31}$$

and $\gamma_2 > 1$ avoids over actuation.

3. **Peak overshoot**: the requirement $g_3 < 0$, where

$$g_3 = \max_t \left\{x_c(p,r,K,t) : t_0 \leq t \leq t_f\right\} - \gamma_3 \max_t \left\{x_c(\bar{p},r,K_{\text{base}},t) : t_0 \leq t \leq t_f\right\} \tag{3.32}$$

   and $\gamma_3 > 1$ bounds the peak overshoot of the response.
4. **Settling time**: the requirement $g_4 < 0$, where

$$g_4 = t_s(p,r,K) - \gamma_4 t_s(\bar{p},r,K_{\text{base}}) \tag{3.33}$$

$$t_s(p,r,K) = \min_t \{|x_c(p,r,K,t) - 0.1| < \varepsilon \text{ for all } t \text{ satisfying } t_s < t < 5 \text{ s}\} \tag{3.34}$$

   and $\gamma_4 > 1$ imposes an upper bound to the settling time.
5. **Local Stability**: the requirement $g_5 < 0$, where

$$g_5 = \max\{\text{Real}\{\text{Eigenvalues}\{A(p,K)\}\}\} \tag{3.35}$$

   and A is the closed-loop system matrix of the LTI system, ensures that the up-right position of the pendulum is locally stable.

The constraint functions corresponding to the first two requirements are evaluated using the L_2 norm, and therefore, they are a function of the transient response. Notice that the dependency of g on p assumes an unknown and implicit functional form. Also notice that evaluating g for a particular realization of p requires simulating the closed-loop response and calculating the five constraint functions comprising g. Further notice that while the notions of goodness supporting the requirement functions are fairly universal, their particular functional form is not.

3.5 Results

3.5.1 Baseline Controller

This controller, to be denoted as K_{base}, was designed using the structure presented in (3.17) based on the closed-loop performance at $p = \bar{p}$. The parameters determining K are the LQR matrices $Q = \text{diag}\{[40, 0, 0, 0, 50]^{\top}\}$ and $R = 10^{-3}$. This controller performed well in both the simulation and experimental settings. Figure 3.5 shows the number of requirement violations in the parameter space and the maximal set when the time responses, thus the constraint function g, are evaluated using a linear simulation. Note that the safe domain, where all requirements are satisfied, corresponds to the green region. The color scale indicates the number of constraints violated, with black corresponding to all five constraints violated. The colored lines correspond to constraint boundaries for each of the five constraints. The corresponding PSM is $\rho = 6.4 \times 10^{-3}$. This figure, as some that follow, was generated by simulating the closed-loop response at the points of an uniformly spaced grid, calculating the constraint function g at each point, and counting the number of non-negative components of each g. This large number of simulations (over 8,000), which are not required to calculate the maximal deformation, has only been made for illustrative purposes.

Figure 3.6 shows the number of requirement violations and the maximal set for K_{base} when evaluated using the nonlinear simulation. The corresponding PSM is $\rho = 1.6 \times 10^{-2}$. This larger PSM obtained from the nonlinear evaluation model is consistent with the design practice used for determining K_{base}. The process was an iterative technique of modifying Q and R in the laboratory and evaluating the closed-loop performance on hardware.

By definition, the CPV is a point on the surface of the maximal set that touches the boundary of the failure domain. Some of the figures may misleadingly show that the CPV is at the interior of the safe domain. This is a mere consequence of the

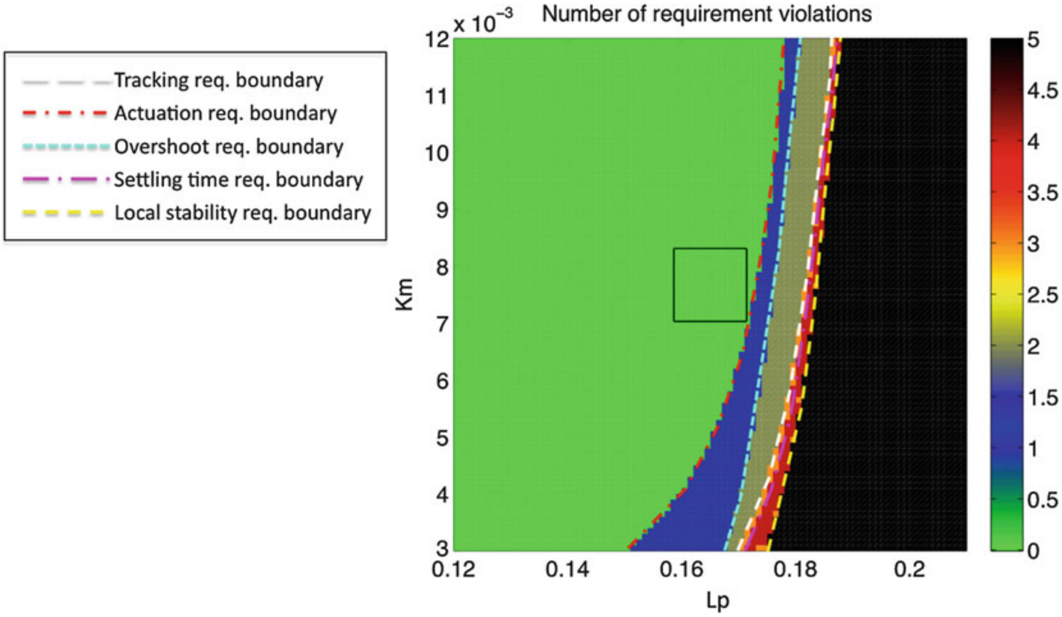


Fig. 3.5 Number of requirement violations in the parameter space and maximal set for K_{base} using linear simulation

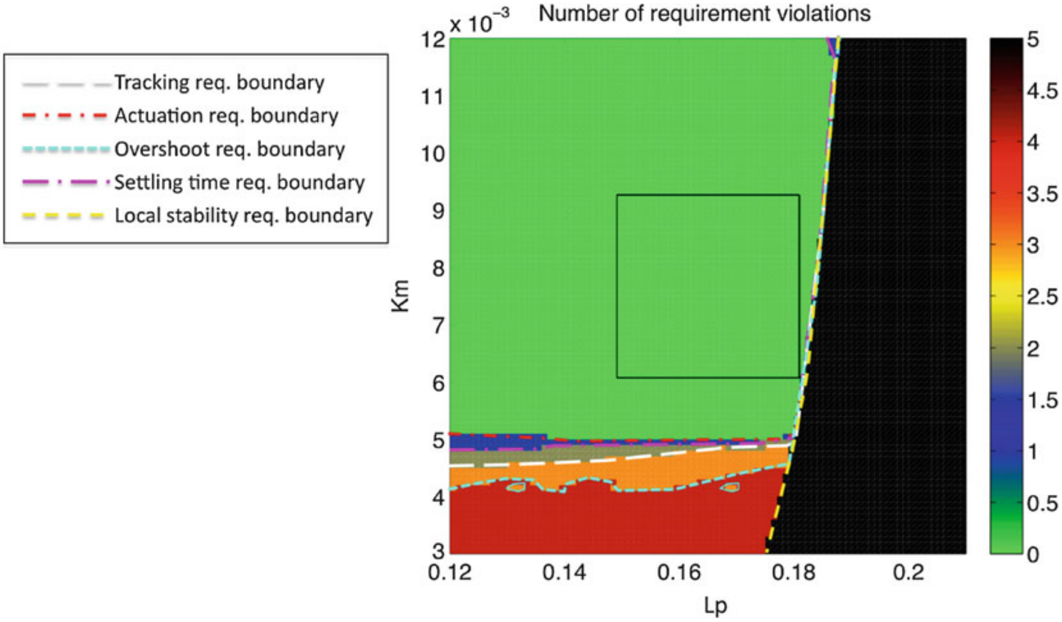


Fig. 3.6 Number of requirement violations in the parameter space and maximal set for K_{base} using nonlinear simulation

manner in which the figures are generated: they show a piecewise constant function over a uniform partition of the parameter space. The same value/color is assigned to all parameter realizations constituting an element of the partition. That value corresponds to the center of the box. The actual function, which is used when calculating the maximal deformation, may vary across any given uncertain parameter range.

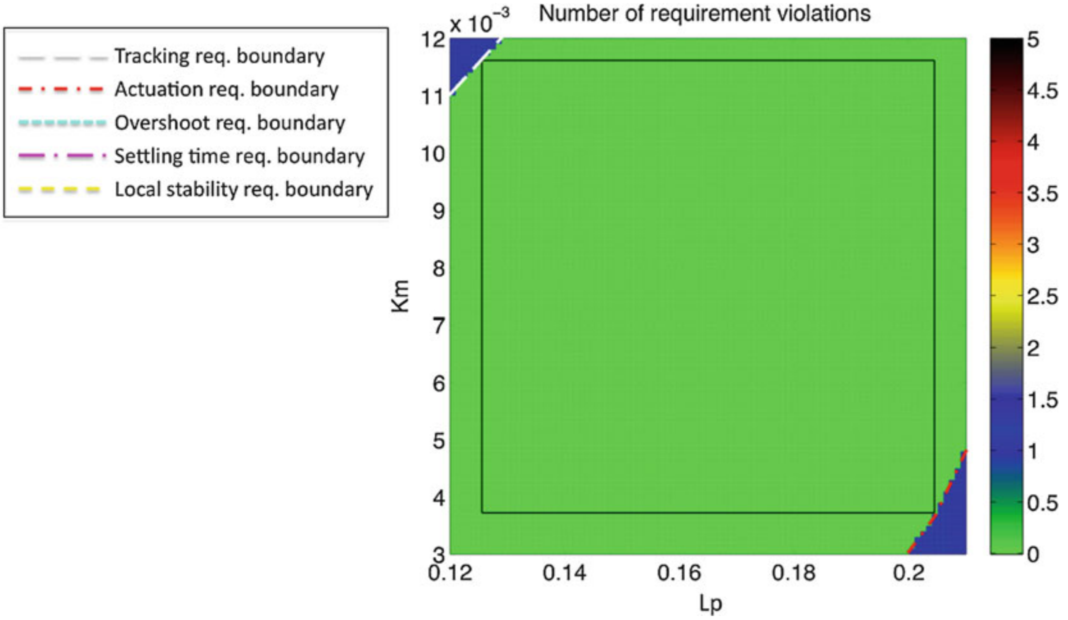


Fig. 3.7 Number of requirement violations in the parameter space for $K_{\text{tuned-1}}$ according to the linear simulation

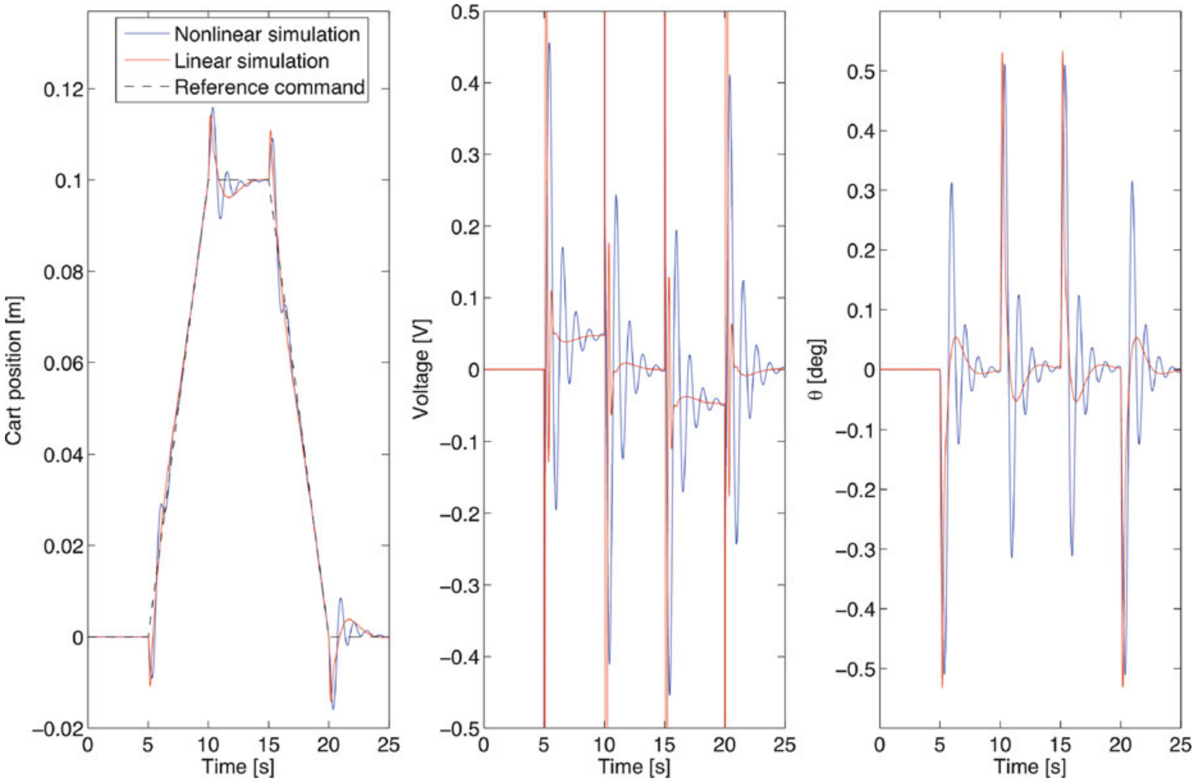


Fig. 3.8 Time responses for controller $K_{\text{tuned-1}}$ at $p = [0.2, 4 \times 10^{-3}]$

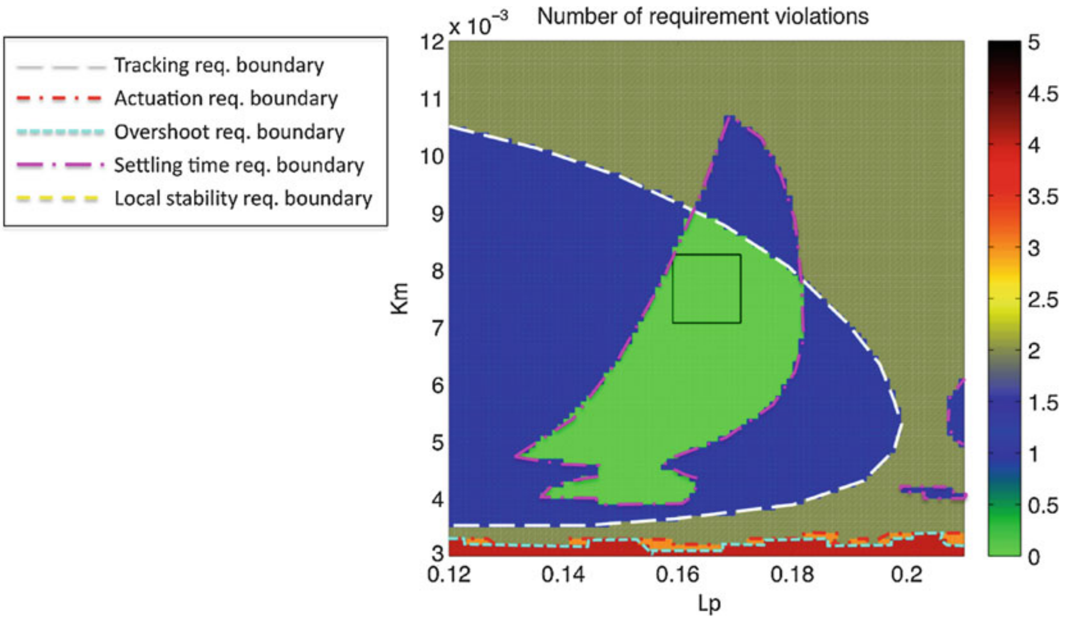


Fig. 3.9 Number of requirement violations in the parameter space and maximal set for $K_{\text{tuned-1}}$ according to the nonlinear simulation

3.5.2 Tuned Controller 1: Designed Based on Linear Simulation

The application of the control tuning procedure above using time responses of the linearized model yields controller $K_{\text{tuned-1}}$. The design parameters used throughout this study are the Q and R parameters of the LQR controller. The controller parameters for this linearized system model are $Q = diag\{[7.67, 0.28, 7.62, 1.4 \times 10^{-3}, 79.94]^{\top}\}$ and $R = 3.2 \times 10^{-3}$. Figure 3.7 shows the corresponding uncertain parameter space. The corresponding PSM, which is equal to $\rho = 3.965 \times 10^{-2}$, is more than six times larger than that of the baseline. This illustrates the effectiveness of the control tuning procedure. Figure 3.8 shows the linear and nonlinear time responses of $K_{\text{tuned-1}}$ for $p = [0.2, 4 \times 10^{-3}]^{\top}$. Note that the more oscillatory nature of the nonlinear response makes the controller unacceptable.

A robustness analysis of $K_{\text{tuned-1}}$ based on the nonlinear response leads to $\rho = 6 \times 10^{-3}$. This PSM is 6. 5 times smaller than that of the linear response. Figure 3.9 shows the corresponding uncertain parameter space. Major differences between the linear analysis of Fig. 3.7 and the nonlinear analysis of Fig. 3.9 are apparent. The requirements based on the L_2 norms, whose satisfaction depend upon the transient response, are the most affected.

3.5.3 Tuned Controller 2: Designed Based on Nonlinear Simulation

The application of the control tuning procedure above using time responses of the nonlinear model equations, as given in (3.2) and (3.3), yields controller $K_{\text{tuned-2}}$. This controller parameters are $Q = diag\{[8.20, 10.63, 0.19, 10^{-4}, 42.46]^{\top}\}$ and $R = 1.04 \times 10^{-4}$. Figure 3.10 shows the corresponding uncertain parameter space. The corresponding PSM is equal to $\rho = 2.34 \times 10^{-2}$. As expected, this controller is more robust than $K_{\text{tuned-1}}$ according to the nonlinear response. Note the transition between the safe domain (green area) and the region where the system is unstable (black area) may occur without a progressive and noticeable degradation in the transient response, i.e., the violation of $g_5 < 0$ resulting from a small drift in p may not be preceded by any other requirement violation.

The robustness analysis of $K_{\text{tuned-2}}$ according to the linear response leads to $\rho = 3.7 \times 10^{-3}$. Figure 3.11 shows the corresponding number of requirement violations in uncertain parameter space. As before, significant differences between the linear and nonlinear analyses are present. These differences, which are the result of model-form uncertainty, highlight the need for assessing the dependence of the high order terms in (3.9) on the uncertain parameters.

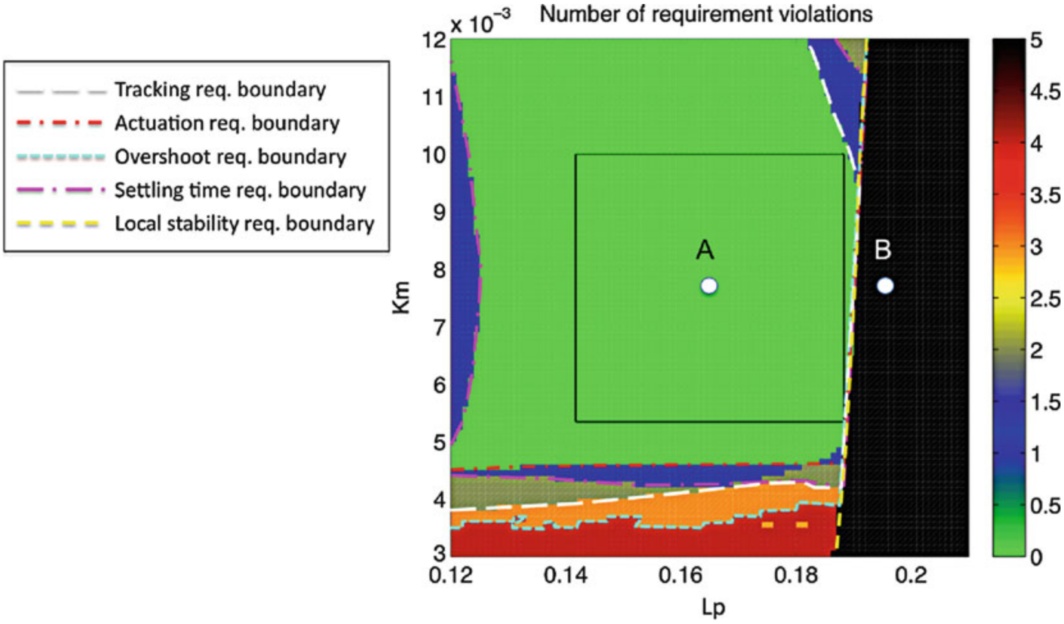


Fig. 3.10 Number of requirement violations in the parameter space and maximal set for $K_{\text{tuned-2}}$ according to the nonlinear simulation

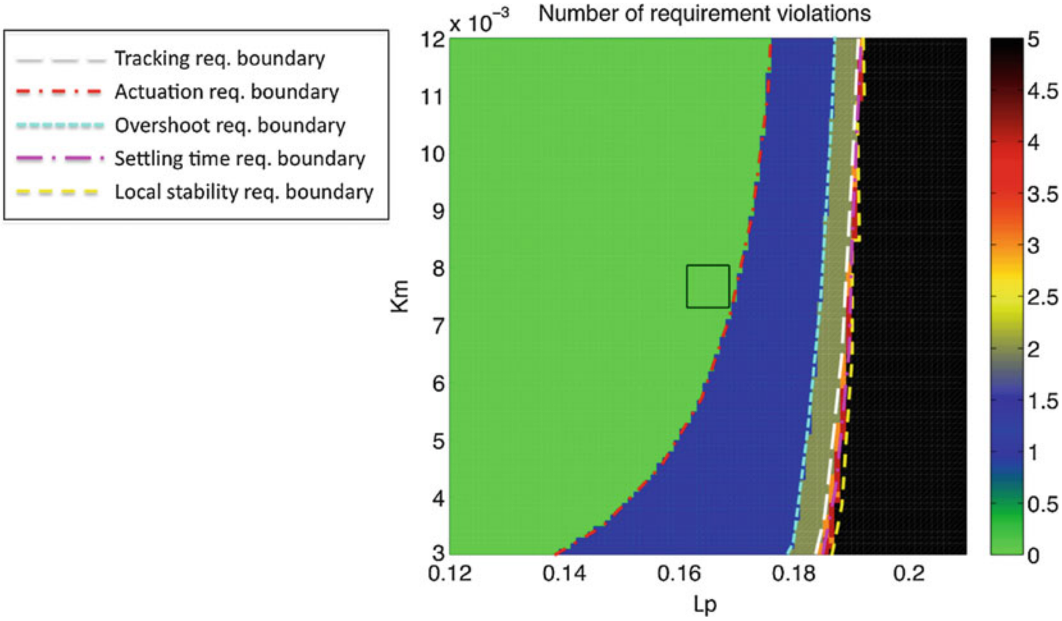


Fig. 3.11 Number of requirement violations in the parameter space and maximal set for $K_{\text{tuned-2}}$ according to the linear simulation

Experimental results for tuned controller 2 are shown in Figs 3.12 and 3.13. Figure 3.12 shows two cycles of the cart tracking response for the uncertain parameter point $p = [0.165, 7.67 \times 10^{-3}]$. This point is well within the maximal set (see point A in Fig. 3.10) and therefore we should expect a reasonable system response. Qualitatively the system response is good, however, due to additional unmodeled effects, such as: cart and pendulum damping, drive pinion eccentricity, drive rack flexibility, data and power cable frictional effects, etc., the true system response slightly differs from the simulated

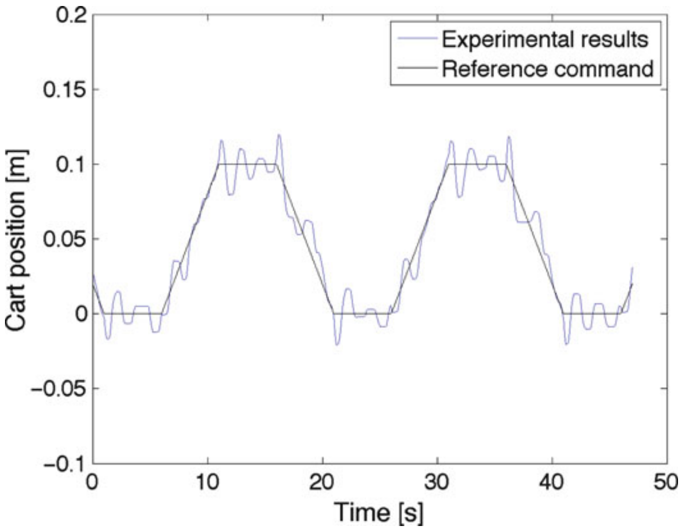


Fig. 3.12 Experimental time responses for controller $K_{\text{tuned-2}}$ at $p = [0.165, 7.67 \times 10^{-3}]$

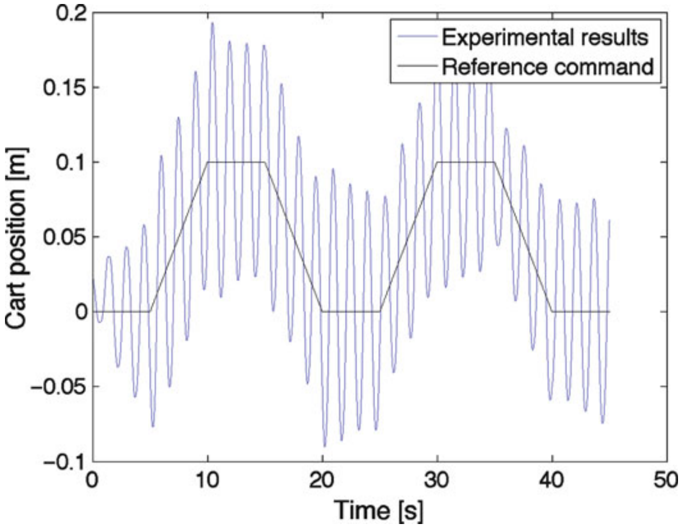


Fig. 3.13 Experimental time responses for controller $K_{\text{tuned-2}}$ at $p = [0.20, 7.67 \times 10^{-3}]$

response. Figure 3.13 shows the cart tracking response for the uncertain parameter point $p = [0.20, 7.67 \times 10^{-3}]$. This point is outside of the maximal set where all five requirements fail (see point B in Fig. 3.10) and therefore we should expect an unacceptable system response. Indeed this is the case as the system tracking response is unacceptably poor. It should be noted that if the system was perfectly linear then the response would be unstable yielding unbounded output. Instead a limit cycle oscillation is observed, which is consistent with the nonlinear system.

3.5.4 Probabilistic Analysis

This section assesses the probabilistic performance of all of the control designs previously presented. The unified framework in [1] provides explicit methods on how to use maximal sets to compute bounds to failure probabilities. To proceed with failure probability computations, one must first prescribe a PDF for the uncertain parameters. Independent generalized beta distributions were chosen to represent the uncertainty in both L_p and K_m. Generalized beta distributions are

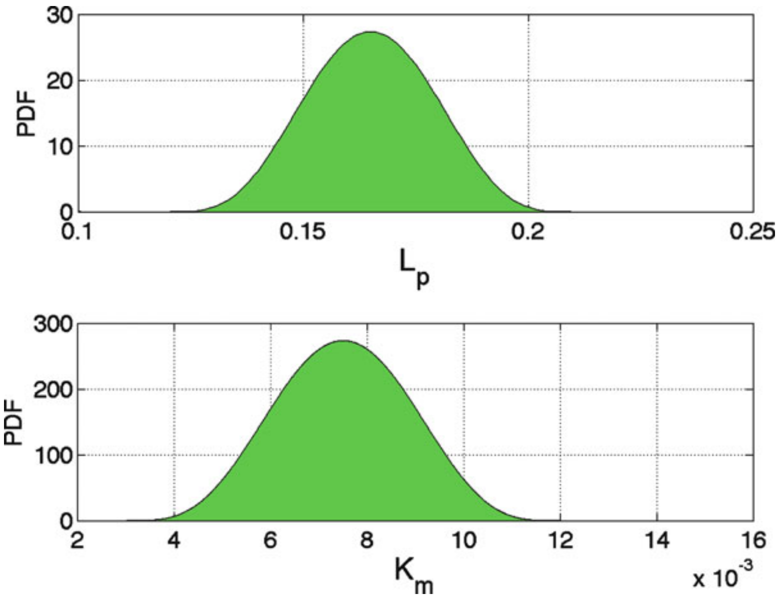


Fig. 3.14 Probability density functions

Table 3.1 Robustness assessment results

Controller	Test model	PSM	Failure probability upper bound	Failure probability using sampling
K_{base}	Linear	6.4×10^{-3}	0.8847	0.3146
K_{base}	Nonlinear	1.6×10^{-2}	0.4538	0.1236
$K_{\text{tuned}-1}$	Linear	3.9×10^{-2}	5.22×10^{-4}	1.5455×10^{-7}
$K_{\text{tuned}-1}$	Nonlinear	6.0×10^{-3}	0.8989	0.5035
$K_{\text{tuned}-2}$	Linear	3.7×10^{-3}	0.9604	0.3990
$K_{\text{tuned}-2}$	Nonlinear	2.34×10^{-2}	0.1653	0.0441

parameterized by four values, [*a*, *b*, *xmin*, *xmax*], two shaping parameters, a lower limit, and an upper limit, respectively. The specific parameters used in this paper were for L_p: [5, 5, 0. 12, 0. 21], and for K_m: [5, 5, 0. 003, 0. 012]. Plots of the probability density functions are given in Fig. 3.14. The probabilistic assessments performed in this work were based upon probability bounds computed using maximal sets, and estimates of failure probability using sampling. A robustness comparison of the three controllers is given in Table 3.1. There are several things to note from these results. In terms of robustness, the baseline controller, K_{base}, shows good robustness when evaluated on the nonlinear system, but very poor robustness when evaluated on the linear system. As stated earlier, this is consistent with the design practice used to generate K_{base}. The controller $K_{\text{tuned-1}}$ had good robustness when evaluated using the linear design model, but suffered greatly when deployed on the nonlinear model. The best choice of controller is $K_{\text{tuned-2}}$ because it has good robustness margins, as indicated by the 46.3% improvement in the PSM over K_{base}. These results reinforce two important principles. The first is that parametric uncertainty must be explicitly considered to achieve good robustness, and the second is that model nonlinearities can have a dramatic impact on system robustness. These results clearly demonstrate that controllers designed using linear models may have very poor robustness when evaluated on nonlinear models.

Most of the prior state-of-the-art techniques (e.g., μ-synthesis, H_∞, linear parameter varying, etc.) for the analysis and control design in the presence of parametric uncertainty require linear plants and a particular structure of parameter dependencies (affine dependencies, or limited polynomial dependencies). Furthermore, explicitly incorporating time domain specifications into conventional robust control design techniques can be very cumbersome. In contrast, the strategies proposed in this work do not require linear system dynamics or a specific functional form of the uncertainty and can easily accommodate a variety of time or frequency domain requirements. However, one of the greatest advantages is the ease at which the methods used here interface with arbitrary simulation models and the ability to incorporate probabilistic definitions of parametric uncertainty models and therefore represents an advancement in the state of practice for control analysis and design.

3.6 Conclusions

This paper presents a control design and analysis methodology for arbitrary dynamic systems in the presence of parametric uncertainty. A brief overview of techniques to compute maximal multidimensional bounding sets and associated probability of failure bounds for parametric uncertainties was given. The classic inverted pendulum was used to demonstrate these techniques for control system design and analysis on a realistic nonlinear system. Generic requirements in the form of command tracking, control actuation, peak overshoot, settling time, and local stability were used to demonstrate the versatility of the method. One of the key results from this work was that the methods are equally applicable to either linear or nonlinear system models, and therefore represent an advancement in the state of practice for control system design and analysis. This technique allows one to use the nonlinear model explicitly in the design process, something many conventional robust control analysis and design methods simply cannot accommodate. Using the nonlinear model in the design process was shown to result in a controller with better robustness properties when applied to the nonlinear system than a controller based on the linear model.

References

1. Crespo LG, Kenny SP, Giesy DP (2010) A computational framework to control verification and robustness analysis. NASA TP 2010-216189. NASA Langley Research Center, Hampton, pp 1–38
2. Franklin GF, Powell JD, Emami-Naeini A (2010) Feedback control of dynamic systems, 6th edn. Prentice Hall, Upper Saddle River, pp 37–38
3. Crespo LG, Munoz CA, Narkawicz A, Kenny SP, Giesy D (2011) Uncertainty analysis via failure Domain characterization: polynomial requirement functions. In: ESREL 2011, Troyes, Sept 2011
4. Crespo LG, Matsutani M, Annaswamy A, (2010) Verification and tuning of an adaptive controller for an unmaned air vehicle. In: AIAA guidance navigation and control conference, AIAA-2010-8403, Toronto, Aug 2010

Chapter 4
Bayesian Damage Localisation at Higher Frequencies with Gaussian Process Error

Christophe Lecomte, J.J. Forster, B.R. Mace, and N.S. Ferguson

Abstract This paper concerns the estimation of the location (and properties) of damage in structures using Bayesian methods and Markov Chain Monte Carlo (MCMC). It is widely recognised that the consideration of uncertainty in structural dynamic systems may be essential, for example from an economic point of view ("Does is make sense to add expensive damping if it will only affect a small proportion of the vehicles we produce?") or for critical safety purposes ("What is the risk of failure of an airplane engine due to bladed disk mistuning?"). The use of Bayesian methods appears to be a viable approach to obtain inferences about the parameters of such uncertain systems. Here we report on numerical experiments on the use of MCMC to locate a frequency dependent damage in a one-dimensional structure. Transfer function measurements subject to a Gaussian process measurement error are available. The particular structure of the resulting system matrices is then seen to have a special form which results in a semi-analytic solution method being available and a much reduced computational cost. We discuss the characteristics and efficiency of the Bayesian model and MCMC computation and highlight features in the analysis of structural dynamic systems such as higher-frequency multimodality.

Keywords Bayesian damage identification • Mid-frequency and high-frequency • Frequency dependent damage • Markov Chain Monte Carlo (MCMC) • Gaussian process • Analytic conditional posterior • Tridiagonal matrix • Tempering

4.1 Introduction

Identification of parameters of dynamic systems is a critical aspect for understanding, control and health monitoring of dynamic structures. Uncertainty is always present in this process. The system itself may be inherently random due to changing conditions such as temperature or uncertain components such as a damping material with properties varying from one sample to another. In particular, any damage occurring during the life of the structure is in essence uncertain both in terms of its location and extent. Since perfect modelling of a system is not possible, there is also uncertainty in the choice and quality of the model itself and in the value of its parameters. Finally, measurement errors add to the randomness.

C. Lecomte (✉)
Institute of Sound of Vibration Research (ISVR) and Southampton Statistical Sciences Research Institute (S3RI)
University of Southampton, Southampton SO17 1BJ, UK
e-mail: c.lecomte@soton.ac.uk; lecomtec@member.ams.org

J.J. Forster
School of Mathematics and Southampton Statistical Sciences Research Institute (S3RI), University of Southampton, Southampton SO17 1BJ, UK
e-mail: J.J.Forster@soton.ac.uk

B.R. Mace
Institute of Sound of Vibration Research (ISVR), University of Southampton, Southampton SO17 1BJ, UK

Department of Mechanical Engineering, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
e-mail: b.mace@auckland.ac.nz

N.S. Ferguson
Institute of Sound of Vibration Research (ISVR), University of Southampton, Southampton, SO17 1BJ, UK
e-mail: N.S.Ferguson@soton.ac.uk

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_4,

Bayesian analysis is a natural framework for identifying the parameters of (necessarily random) dynamic systems. It allows the consideration of as many or as few random parameters as judged useful or necessary and the consideration of engineering "common sense" as prior knowledge. It also provides estimates of full probability density functions of the parameters that can further be used to infer statistics of derived functions and properties.

Application of Bayesian methods in structural dynamics have indeed started showing success during the recent decades, for example in updating of structural models [1], model identification [2] and selection [3], or damage detection [4, 5]. While most methods are based on time domain measurements or estimation of the resonance frequencies and mode shapes [6], the work presented here and in [4, 7] makes use of measurements of whole frequency response transfer functions. The goal here is to identify a damage location and magnitude when both damage magnitude and measurement errors at discrete frequencies are a Gaussian process.

In the next section, the structure, the randomness and their parameters are introduced. The Bayesian framework is then presented and a Metropolis MCMC scheme is described. Simulations are then reported and discussed.

4.2 Response of Nominal and Damaged Structures

In this section, the studied structure is described and expressions of its nominal and damaged transfer functions are given.

4.2.1 Description of the Structure

In this paper, a regular spring-mass system is subjected to damage at a random location. This homogeneous collinear structure is a model for various damaged structures such as rods or other continuous systems as illustrated in Fig. 4.1.

A choice of normalised unit masses and springs can be made without loss of generality about the presence of damping or not, and about the value of the masses or spring stiffness (the transformation of a general homogeneous system into its normalised form is discussed in [8]). The equations of the response vector $\mathbf{x} = [x_1, \ldots, x_N]^T$ of the system to a force vector $\mathbf{f} = [f_1, \ldots, f_N]^T$ at circular frequency ω are then

$$[\mathbf{K}(1 + \mathrm{i}\eta) - \omega^2\mathbf{M} - s\mathbf{D}_k]\mathbf{x}(\omega, k, s) = \mathbf{f}. \tag{4.1}$$

The structure is free at one end and fixed at the other so that the mass matrix $\mathbf{M}$ is the identity matrix while the stiffness matrix $\mathbf{K}$ has $[1, 2, \ldots, 2]$ on its main diagonal and -1 on its two neighbouring diagonals. Structural damping can be considered through the introduction of a non-zero hysteretic loss factor, η. Damage produces a change in the dynamic stiffness matrix in the form of a damage matrix $\mathbf{D}_\mathrm{k}$ multiplied by a scalar damage magnitude s. Both $\mathbf{D}_k$ and s can in general depend on frequency. Here the damage occurs at a single k-th location so that $\mathbf{D}_K = \mathbf{e}_k\,\mathbf{e}_k^T$ is the outer product of the k-th unit vector $\mathbf{e}_k$. The damage magnitude $s(\omega)$ is random and its probability density function (pdf) $p(s)$ will be discussed later. Usually, only a transfer function $g(\omega, k, s) = \mathbf{c}^T\,\mathbf{x}(\omega, k, s)$ (i.e. part of the response vector) is of interest for some output vector $\mathbf{c}$.

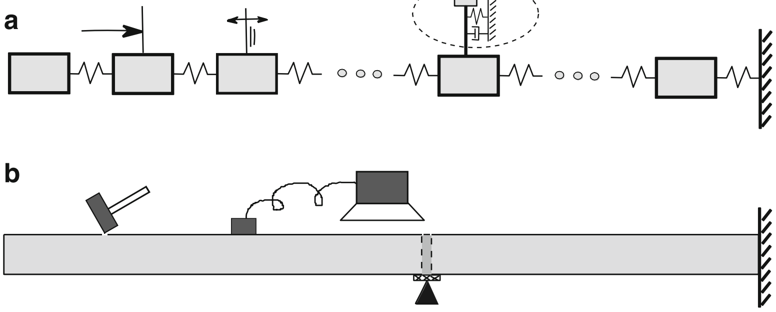


Fig. 4.1 Illustration of the damaged structure with random damage at a specific location. (**a**) Discrete case of the spring-mass system; (**b**) continuous case of a rod as a limit of $N \to \infty$

It is now shown that the transfer functions of the damaged structure can be obtained cheaply by an update of the transfer functions of the nominal (undamaged, $s = 0$) structure.

4.2.2 Transfer Functions of the Nominal and Damaged Structures

The (nominal) response of the nominal structure is $\mathbf{x}(\omega, k, 0) = (\mathbf{K}{-}\omega^2\,\mathbf{M})^{-1}\mathbf{f}$, from which the nominal transfer function $g(\omega, k, 0) = \mathbf{c}^T\,\mathbf{x}(\omega, k, 0)$ is directly available. The transfer function of the damaged structure can then be obtained through the rank-one update [9, 10],

$$g(\omega,k,s) = g(\omega,k,0) + b_1(\omega,k)\left[\frac{s_1(\omega,k)}{s_1(\omega,k)-s} - 1\right] \tag{4.2}$$

where the functions s_1 and b_1 only depend on nominal transfer functions, $s_1(\omega,k) = \left[\mathbf{e}_k{}^T(\mathbf{K}-\omega^2\mathbf{M})^{-1}\mathbf{e}_k\right]^{-1}$ and $b_1(\omega,k) = s_1(\omega,k)\mathbf{c}^T(\mathbf{K}-\omega^2\mathbf{M})^{-1}\mathbf{e}_k\mathbf{e}_k{}^T(\mathbf{K}-\omega^2\mathbf{M})^{-1}\mathbf{f}$. This update expression allows rapid and efficient evaluation of the damaged transfer functions for various values of the damage magnitude s and location k, from only four nominal transfer functions.

Several measurements with different force and output vectors, $\mathbf{f}_m$ and $\mathbf{c}_m$, can be made on the actual physical system, each of those indexed by a subscript $m = 1, \ldots, M_e$. The measured transfer functions, $y_m(\omega)$, differ from the model transfer functions, g_m, due to both error in measurements and model errors. This difference is denoted as the measurement error ε_m (which is also a random variable) so that

$$y_m(\omega) = g_m(\omega,k,s) + \varepsilon_m(\omega,k,s). \tag{4.3}$$

4.2.3 Gaussian Process Statistics of the Measurement Error

In the current analysis, we assume that the measurement errors with different force and output vectors are independent from each other and that there is correlation of the measurement errors between different frequencies. Specifically, for a given measurement y_m evaluated at discrete frequencies, $\omega_j = \omega_0 + j\Delta\omega,\ j = 1, \ldots, Q$, the true measurement error (corresponding to the actual location k and damage magnitude $\mathbf{s} = [s(\omega_1), \ldots, s(\omega_Q)]^T$) is a Gaussian process such that

$$p(\boldsymbol{\varepsilon}_m) = \frac{1}{\pi^Q\det(\mathbf{S})}\mathrm{e}^{-\boldsymbol{\varepsilon}_m^H\mathbf{S}^{-1}\boldsymbol{\varepsilon}_m} \tag{4.4}$$

Where $\det(\mathbf{S})$ denotes the determinant of $\mathbf{S}$, $\boldsymbol{\varepsilon}_m = \left[\varepsilon_m(\omega_1,k,s), \ldots, \varepsilon_m(\omega_Q,k,s)\right]^T$ is the (complex) vector of errors at different frequencies, and $\boldsymbol{\varepsilon}_m^H$ is the transpose conjugate of $\boldsymbol{\varepsilon}_m$. The matrix $\mathbf{S} = \sigma^2\ \mathbf{C}$ is the product of the standard deviation σ^2 and the correlation matrix $\mathbf{C}$ whose (q, r) coefficient is $\mathrm{e}^{-\mathrm{i}\rho|q-r|\Delta\omega}$ for $\rho > 0$. The matrix $\mathbf{C}$ thus has the form

$$\mathbf{C} = \begin{pmatrix} 1 & \mathrm{e}^{-\mathrm{i}\rho\Delta\omega} & \cdots & \mathrm{e}^{-\mathrm{i}\rho(Q-1)\Delta\omega} \\ \mathrm{e}^{-\mathrm{i}\rho\Delta\omega} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{e}^{-\mathrm{i}\rho\Delta\omega} \\ \mathrm{e}^{-\mathrm{i}\rho(Q-1)\Delta\omega} & \cdots & \mathrm{e}^{-\mathrm{i}\rho\Delta\omega} & 1 \end{pmatrix} \tag{4.5}$$

and the cross-correlations decrease as the frequency space, $\Delta\omega$, or the parameter ρ increase. Thanks to the form of this correlation matrix, the inverse and determinant of $\mathbf{S}$ can be evaluated cheaply as will be discussed in the next section. The parameters σ^2 and ρ are here assumed equal for all measurements and are a priori unknown (as are the two other parameters, i.e. the location k and damage vector $\mathbf{s}$).

4.3 Bayesian Inference

In this section, the principles providing posterior probability of the parameters based on measurements are presented.

4.3.1 General Bayes Considerations

The values of the four sets of parameters, namely the position $\theta_1 = k$ of the damage, the error measurement parameters, $\theta_2 = 1/\sigma^2$, and $\theta_3 = \rho$, and the vector of damage magnitudes, $\boldsymbol{\theta}_4 = \mathbf{s}$, can be stored in a vector $\boldsymbol{\theta} = [\theta_1, ..., \boldsymbol{\theta}_4{}^T]^T$ of length 3 + Q. The uncertainty pervading these parameters based on the measurements and prior knowledge can then be evaluated using Bayes theorem, i.e.

$$p(\boldsymbol{\theta}|\mathbf{Y}) = \frac{p(\mathbf{Y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{Y})} \tag{4.6}$$

with $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_{M_e}]$ and $\mathbf{y}_m = [y_m(\omega_1), \ldots, y_m(\omega_Q)]^T$for $m = 1, \ldots, M_e$. The likelihood $p(\mathbf{Y}|\boldsymbol{\theta})$ and prior statistics $p\ (\boldsymbol{\theta})$ will now be discussed while the denominator probability, $p(\mathbf{Y}) = \int_{D(\boldsymbol{\theta})} p(\mathbf{Y}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}$ (which can be expressed as an integral of the numerator on the domain $D(\boldsymbol{\theta})$ of the parameters) is a scaling factor that is not necessary for implicit sampling by Markov Chain Monte Carlo. Since this function is the integral of the numerator, in principle equation (4.5) gives *probabilities of all combined and marginal probabilities* of the location and extent of damage, as well as of the parameters of the measurement errors.

4.3.2 System Randomness and Likelihood

Randomness or uncertainty is propagated in the system as follows: the knowledge of the position k and damage magnitude vector $\mathbf{s}$ uniquely determines the vector of model transfer functions, $\mathbf{g}_m(k, \mathbf{s}) = [g_m(\omega_1, k, s(\omega_1)), \ldots, g_m(\omega_Q, k, s(\omega_Q))]^T$. The probability of the measurements $\mathbf{y}_m$ for given σ^2 and ρ is then equal to the probability of the corresponding error in measurements, $\boldsymbol{\varepsilon}_m = \mathbf{y}_m - \mathbf{g}_m(k, s)$. The likelihood of the measurements is therefore

$$p(\mathbf{Y}|\boldsymbol{\theta}) = \frac{1}{\pi^{(Q.M_e)}.\det(\mathbf{S})^{M_e}} \exp\Big(-\Sigma_{m=1,\ldots,M_e}[(\mathbf{y}_m - \mathbf{g}_m(k, \mathbf{s}))^H \mathbf{S}^{-1}(\mathbf{y}_m - \mathbf{g}_m(k, \mathbf{s}))]\Big) \tag{4.7}$$

$$= \frac{1}{(\pi.\sigma^2)^{(Q.M_e)}.\det(\mathbf{C})^{M_e}} \exp\left\{-\left(\frac{1}{\sigma^2}\right)\Sigma_{m=1,\ldots,M_e}\left[(\mathbf{y}_m - \mathbf{g}_m(k, \mathbf{s}))^H \mathbf{C}^{-1}(\mathbf{y}_m - \mathbf{g}_m(k, \mathbf{s}))\right]\right\}. \tag{4.8}$$

As already mentioned, the determinant and inverse of S can be evaluated efficiently.

4.3.3 Prior Engineering Knowledge and Probabilities

Any prior knowledge and resulting prior probability of the parameters can be incorporated in the statistical model. The particular choices made here are now described. No a priori information on the location of damage is known. A uniform prior is therefore selected such that all locations, $k = 1, \ldots, N$, are equiprobable, $p(\theta_1 = k) = 1/N$. For efficiency, a Gamma distribution (with parameters α and β) is chosen for the measurement error standard deviation. The prior is then such that $p\left(\theta_2 = \frac{1}{\sigma^2}; \alpha, \beta\right) = \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\alpha-1} \mathrm{e}^{-\beta\frac{1}{\sigma^2}}$ where $\Gamma(\alpha) = \int_0^\infty t^{(\alpha-1)}\mathrm{e}^{-t}\mathrm{d}t$ is the Gamma function. It will be shown that with this choice, analytical steps can be taken in the MCMC process. Since the a priori knowledge of the correlation factor ρ can vary over a large range, a lognormal prior probability is selected such that the probability of ρ is assumed to be a normal variable with mean $\mu_1 = \log(\rho_0)$ and standard deviation τ_1. Therefore, $p(\theta_3 = \rho; \mu_1, \tau_1) = \frac{1}{\sqrt{2\pi\tau_s^2}} \exp\left(-\left[\frac{(\log(\rho)-\mu_1)^2}{2(\tau_1)^2}\right]\right)$. A (complex) Gaussian process damage vector is considered in the examples presented in this paper. The same Gaussian process structure, with parameters σ_s^2 and ρ_s, is used as prior, so that $p(\boldsymbol{\theta}_4 = \mathbf{s}; \sigma_s, \rho_s) = \frac{1}{(\pi.\sigma_s^2)^Q \det(\mathbf{C_s})} \mathrm{e}^{-\left(\frac{1}{\sigma_s^2}\right)\mathbf{s}^H \mathbf{C_s}^{-1}\mathbf{s}}$ where the (q, r) coefficients of $\mathbf{C_s}$ are $\mathrm{e}^{-\mathrm{i}\rho_s|q-r|\Delta\omega}$.

4.4 Markov Chain Monte Carlo

In this section, a MCMC metropolis walk is described to estimate the posterior probability of the system parameters.

4.4.1 General MCMC Considerations

Samples can be explicitly drawn from the posterior $p(\boldsymbol{\theta}|\mathbf{Y})$ without the need to recur to the evaluation of the integral $p(\mathbf{Y})$, by using a Metropolis MCMC scheme [11]. At each step, the new sampled values $\boldsymbol{\theta}^{(j+1)}$ of the parameters depend only on the previous sample $\boldsymbol{\theta}$ and, after a burn-in period, the distribution of samples $\boldsymbol{\theta}^{(1)}$, $\boldsymbol{\theta}^{(2)}$, … tends to the equilibrium posterior probability, $p(\boldsymbol{\theta}^{(\mathbf{j})}) \to p(\boldsymbol{\theta}|\mathbf{Y})$ [11, Chap. 10, Sect. 10.3]. New random trial samples $\boldsymbol{\theta}^{(*)}$are drawn from a jump probability distribution $J(\boldsymbol{\theta}^{(*)}; \boldsymbol{\theta}^{(j)})$ by updating one or more of the parameters of $\boldsymbol{\theta}^{(j)}$. The trial samples are then accepted as $\boldsymbol{\theta}^{(j+1)} = \boldsymbol{\theta}^{(*)}$ with probability $r = \min\left(1, \frac{p(\boldsymbol{\theta}^{(*)}|\mathbf{Y})}{p(\boldsymbol{\theta}^{(j)}|\mathbf{Y})}\right)$, and otherwise $\boldsymbol{\theta}^{(j+1)} = \boldsymbol{\theta}^{(j)}$. Each of the trial samples are generated by updating one or more of the individual components of $\boldsymbol{\theta}^{(j)}$ by a random linear or logarithmic step. Here we report experiments in which a single component of the four components of the parameter vector is updated when drawing each new trial sample. Note that the trial steps must be chosen such that the probability of stepping from one to the other parameter step is symmetric, $J(\boldsymbol{\theta}^{(*)}; \boldsymbol{\theta}^{(j)}) = J(\boldsymbol{\theta}^{(j)}; \boldsymbol{\theta}^{(*)})$, in classic Metropolis MCMC.

4.4.2 Choice of Trial Steps

The four components of the parameter vector are updated one after the other (only modifying θ_1 at step j, θ_2 at step $j + 1$,…) in the following way. The trial locations $\theta_1 = k$ are generated by discrete normal steps, $\theta_1^{(*)} - \theta_1^{(j)} \sim \text{round}[N(0, t)]$, that are the rounded version (closest integer) of a real normal variable $x \sim N(0, t), p(x) = \frac{1}{\sqrt{2\pi t^2}} \mathrm{e}^{\frac{-x^2}{2t^2}}$. Trials that fall outside the range of possible values of k are automatically refused. The steps for the standard deviation, $\theta_2 = 1/\sigma^2$, as well as for the error correlation ρ are log-normal $\theta_2^{(*)} \sim \theta_2^{(j)}\mathrm{e}^{N(0,\eta_2)}$, and $\theta_3^{(*)} \sim \theta_3^{(j)}\mathrm{e}^{N(0,\eta_3)}$. Finally, the steps for the vector of the disturbance magnitude at all frequencies are generated from a multi variable (complex) normal variable with parameters σ_p and ρ_p (instead of σ and ρ).

4.4.3 Improvement of the Walks

Two improvements of the Bayesian MCMC walk are presented here: direct sampling from an analytical probability density function, and use of parallel tempered chains. While the main purpose of these improvements is to accelerate the convergence of the samples to the posterior distribution, the use of tempered chains may be critical in practical situations to avoid the samples being limited to a single region of relatively high probability of the posterior.

4.4.3.1 Faster Convergence by Using Analytic Steps

The length of the burn-in is, in principle, reduced by sampling the standard deviation, $\theta_2 = \frac{1}{\sigma^2}$ from its exact conditional posterior distribution. Thanks to the choice of prior and likelihood function, the conditional posterior (as a function of θ_2) is indeed a Gamma distribution (with parameters $\alpha + Q.M_e$ and $\beta + \Sigma_{m=1,\ldots,M_e}[(\mathbf{y}_m - \mathbf{g}_m(k,\mathbf{s}))^{\mathrm{H}}\boldsymbol{C}^{-\mathbf{1}}(\mathbf{y}_m - \mathbf{g}_m(k,\mathbf{s}))]$) since it is proportional to

$$p\left(\theta_2 = \frac{1}{\sigma^2}\middle|\mathbf{Y}, \boldsymbol{\theta}_{\setminus 2}\right) \propto p(\mathbf{Y}|\boldsymbol{\theta})p(\theta_2; \alpha, \beta) \propto$$

$$\frac{1}{(\sigma^2)^{(\alpha+Q.M_e-1)}} \mathrm{e}^{-\left(\frac{1}{\sigma^2}\right)\left\{\beta+\Sigma_{m=1,\ldots,M_e}\left[(\mathbf{y}_m-\mathbf{g}_m(k,\mathbf{s}))^{\boldsymbol{H}}\mathbf{C}^{-1}(\mathbf{y}_m-\mathbf{g}_m(k,\mathbf{s}))\right]\right\}} \tag{4.9}$$

where $\boldsymbol{\theta}_{\setminus 2}$ indicates the parameters (the part of the parameter vector $\boldsymbol{\theta}$) complementary to θ_2. Samples from this parameter can therefore be drawn directly from this Gamma distribution in the MCMC walk when all the other parameters are kept constant.

4.4.3.2 Treatment of Multimodality by Tempering

Statistical multimodality (several separated regions of parameters with high probability) can be such that a single MCMC chain can stagnate in a single region of relatively higher probability. If this is the case, some regions of the parameter space with high probability, even that with the highest probability, may not be reached during a feasible finite number of samples. For the current model, such multimodality exists at high frequency. A solution chosen here is to run several chains $i = 0, 1, 2, \ldots, I$ in parallel. Besides the main original chain, the other chains correspond to a tempered posterior p_i such that the likelihood is raised to a given power, t_i, with $t_0 = 1$ and $t_i < t_{i-1}$,

$$p_i(\boldsymbol{\theta}|\mathbf{Y}) \propto [p(\mathbf{Y}|\boldsymbol{\theta})]^{t_i} p(\boldsymbol{\theta}). \tag{4.10}$$

While these tempered posteriors are not identical to the original posterior and the walks therefore do not share the same equilibrium distribution, they allow for increased mixing of the samples. At each step, either ‘regular’ trial samples are drawn and accepted or refused in each of the chains independently of the other chains, or an exchange of samples is proposed between two neighbouring (indices i and $i-1$) chains. If the full vectors of samples are indexed by their chain number such that $\boldsymbol{\theta}_i = \left[\theta_{i,1}, \ldots, \boldsymbol{\theta}_{i,4}{}^T\right]^T$, then trial exchange samples at a step $j + 1$ are $\boldsymbol{\theta}_i^{(*)} = \boldsymbol{\theta}_{i-1}^{(j)}$ and $\boldsymbol{\theta}_{i-1}^{(*)} = \boldsymbol{\theta}_i^{(j)}$. This exchange is accepted with probability $r = \min\left(1, \frac{p_i\left(\boldsymbol{\theta}_i^{(*)}|\mathbf{Y}\right)p_{i-1}\left(\boldsymbol{\theta}_{i-1}^{(*)}|\mathbf{Y}\right)}{p_i\left(\boldsymbol{\theta}_i^{(j)}|\mathbf{Y}\right)p_{i-1}\left(\boldsymbol{\theta}_{i-1}^{(j)}|\mathbf{Y}\right.}\right)$, in which case $\boldsymbol{\theta}_i^{(j+1)} = \boldsymbol{\theta}_i^{(*)}$ and $\boldsymbol{\theta}_{i-1}^{(j+1)} = \boldsymbol{\theta}_{i-1}^{(*)}$. Otherwise, $\boldsymbol{\theta}_i^{(j+1)} = \boldsymbol{\theta}_i^{(j)}$and $\boldsymbol{\theta}_{i-1}^{(j+1)} = \boldsymbol{\theta}_{i-1}^{(j)}$. The exchange steps occur every other N_{full} +1 steps and the two neighbouring chains (i and $i-1$) are chosen by drawing the index i uniformly within the range 1, …, I.

4.4.3.3 Efficient Evaluation of the Gaussian Process Likelihood

At each step of the MCMC walk, the inverse of (or solution with) the matrix $\mathbf{C}$ is in principle required to evaluate terms of the form $\boldsymbol{\varepsilon}_m^H \mathbf{C}^{-1} \boldsymbol{\varepsilon}_m$ as seen from (4.8) or (4.9). Since $\mathbf{C}$ depends on the correlation parameter, ρ, its inverse cannot be evaluated once for all steps and the cost of the solutions could therefore be prohibitive. The very particular form of this Toeplitz matrix however allows for very efficient evaluation at each step. Details of this will be presented elsewhere.

4.5 Simulations

Results of simulations are now reported and discussed. The parameters of the system, likelihood, priors, and MCMC steps are first presented.

Table 4.1 Parameters of the system and MCMC walk

System	System dimension	$N = 300$
	Loss factor	$\eta = 0.01$
	Damage location	$k = 171$
	Damage magnitude	$\mathbf{s} \sim GP_{\mathrm{C}}(\rho_s = 0.1, \sigma_s = 0.5)$
Measurements	# of measurements	$M_e = 5$
	Input locations	$\mathbf{l}_f = [240\ 150\ 133\ 270\ 172]$
	Output locations	$\mathbf{l}_{\mathrm{c}} = [240\ 270\ 72\ 240\ 245]$
	Meas. error	$\boldsymbol{\varepsilon}_{\mathrm{m}} \sim GP_{\mathrm{C}}(\rho = 400, \sigma = 0.1)$
	Frequency step	$\Delta\omega = 0.0005$
Prior	Prior θ_2	Gamma $p\left(\theta_2 = \frac{1}{\sigma^2}; \alpha = 3, \beta = 20\right)$
	Prior θ_3	Lognormal, $p(\theta_3 = \rho; \mu_1 = 200, \tau_1 = 1)$
	Prior $\boldsymbol{\theta}_4$	Gaussian process, $p(\boldsymbol{\theta}_4 = \mathbf{s}; \sigma_s = 0.5, \rho_s = 1)$
MCMC walk	Maximum # of steps	$M_{iter} = 30000$
	Walk order	one component at a time: θ_1, θ_2, θ_3, $\boldsymbol{\theta}_4$
	Step θ_1	$\theta_1^{(j)} + \mathrm{round}[N(0, t_i)]$
	Step θ_2	$\theta_2^{(j)} \mathrm{e}^{N(0,\eta_{2,i})}$
	Step θ_3	$\theta_3^{(j)} \mathrm{e}^{N(0,\eta_{3,i})}$
	Step $\boldsymbol{\theta}_4$	$\theta_4^{(j)} + GP_{\mathrm{C}}(\rho_{p,i}, \sigma_{p,i})$
	Initial values:	$[k_i^{(0)}, 0.011, 151.25, \mathbf{0}^T]^T$
Tempering parameters	# of tempered levels	$I = 7$
	Regular/exchange	$N_{full} = 3$
	Tempering powers	$\mathbf{t} = [1,0.500,0.333,0.167,0.100,0.050,0.033,0.020]$

4.5.1 Parameters

Results are presented for simulations in low and higher frequency ranges. The parameters of the system and MCMC simulations in the higher frequency range are summarised in Table 4.1.

The system is made of $N = 300$ masses and is lightly damped (loss factor $\eta = 0.01$). The actual location of damage is at $k = 171$. Measurements are made at Q discrete frequencies (with frequency step $\Delta\omega = 0.0005$) at several locations $\mathbf{l}_c = [240\ 270\ 72\ 240\ 245]$ for unit forces at mass locations $\mathbf{l}_f = [240\ 150\ 133\ 270\ 172]$ respectively. The actual value of the damage magnitude is drawn from a complex Gaussian process, $\mathbf{s} \sim GP_{\mathrm{C}}(\rho_s = 0.1, \sigma_s = 0.5)$, and the measurement error is similarly drawn (independently at each measurement point) from the likelihood Gaussian process distribution $\boldsymbol{\varepsilon}_{\mathrm{m}} \sim GP_{\mathrm{C}}(\rho = 400, \sigma = 0.1)$. The prior knowledge is uniform for the location of damage and the following Gamma and lognormal prior distributions are considered for the standard deviation and correlation of the measurements, $p\left(\theta_2 = \frac{1}{\sigma^2}; \alpha = 1.9, \beta = \frac{1}{2.9}\right)$ and $p(\theta_3 = \rho; \mu_1 = 50, \tau_1 = 1)$. The prior for the damage magnitude vector is similar to the Gaussian process from which actual damage magnitude was drawn, $p(\boldsymbol{\theta}_4 = \mathbf{s}; \sigma_s = 0.5, \rho_s = 1)$.

The parameters of the MCMC walk to implicitly sample the posterior are a maximum number of $M_{iter} = 30{,}000$ steps, and a sequential consideration of the components θ_1, θ_2, θ_3, $\boldsymbol{\theta}_4$. Every $N_{full} = 3$ times a full sequence of the four parameter components have been updated, an exchange between two neighbour tempered levels is considered. Seven tempered chains are considered (additional to the regular chain) with powers $t_i = 0.500,0.333,0.167,0.100,0.050,0.033,0.020$. The starting location of the main chain has been chosen relatively far from the actual damage location $k_0^{(0)} = 60$ and the starting locations of the other chains have been randomly drawn. The other starting values were identical for all tempered levels, $\theta_2^{(0)} = 33.33$, $\theta_3^{(0)} = 0.05$, $\boldsymbol{\theta}_4^{(0)} = \mathbf{0}$. The parameters of the sampling process, $t_i, \eta_{2,i}, \eta_{3,i}, \rho_{p,i}$, and $\sigma_{p,i}$ have been assigned values that allow appropriate mixing of the chains and similar parameters were used for the low frequency range simulations.

4.5.2 Results

Results of the MCMC walks are presented for two frequency configurations: In the first set of results, a walk is considered at each single frequency ($Q = 1$) independently of the measurements at other frequencies while in the second set, all $Q = 150$ frequencies are considered at the same time in the same Gaussian process likelihood. The same range of frequencies $\Delta\omega$, $2\Delta\omega$, …, 0.075 is covered in both cases.

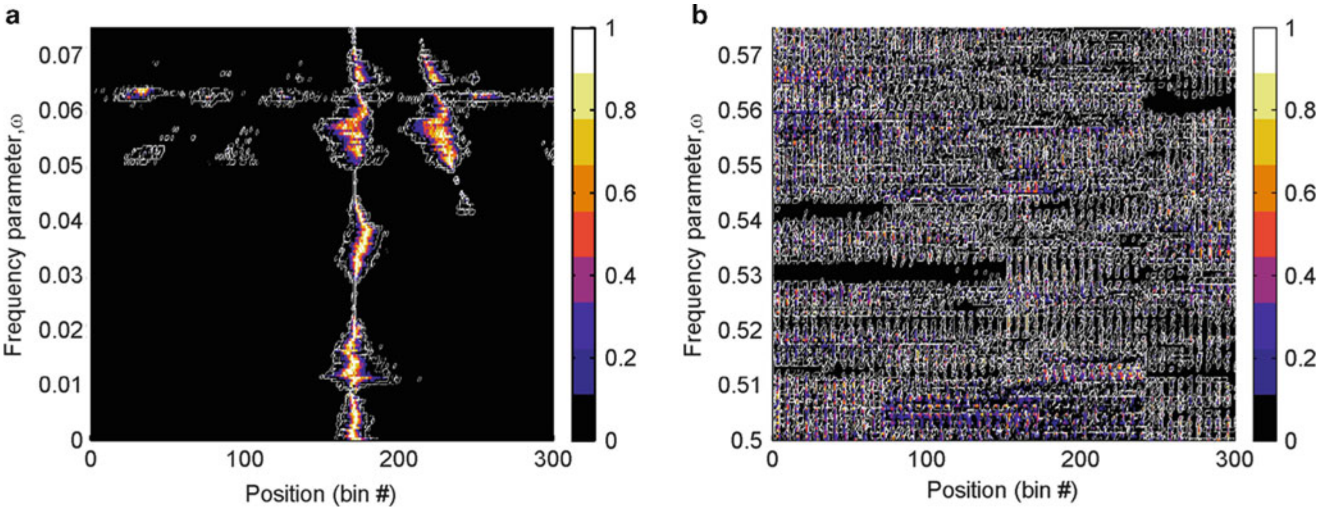


Fig. 4.2 Scaled histogram map of the damage ∈ location considering one frequency (horizontal line) at a time. Actual location of the damage is $k = 171$. (**a**) Low frequency range $\omega \in [0, 0.075]$. (**b**) Higher frequency range $\omega \in [0.5, 0.575]$

4.5.2.1 Individual Frequency Simulations

The case of frequencies considered individually is similar to the analysis presented in [4] with additional recourse here to tempering. The (left) part a) of Fig. 4.2 is a map of histograms of the second half of samples ($j = 15000 \ldots 30001$). Each horizontal line is therefore the (scaled) posterior probability of damage location estimated from measurements at a single frequency. Patterns can be clearly distinguished. At very low frequency, the (significant) probability of damage is located in a single region and when the frequency increases, other modes (in a statistical sense) of probability appear. The actual damage location is the only location at which significant posterior probability exists at all frequencies. Furthermore, there are frequencies at which the location is perfectly well identified (it is the only one with significant posterior probability). Similar patterns were already noted in the evaluation of analytical posteriors in [7] and could be explained by the form of the likelihood function. These patterns are also present at higher frequency as is illustrated in the (right) part b) of Fig. 4.2. Although they are in theory qualitatively identical, the high (statistical) multimodality, quality of the estimation (identical walk parameters were used at all frequencies), and (structural dynamic) modal characteristics make for a very different story. Visually, at first sight, all locations appear to be almost equally likely and it is hard to distinguish the exact location amongst these. Also, despite the presence, again, of discrete frequencies at which the exact location is better identified by the (single frequency) posterior, the situation is very different than in the low frequency case, since small variation from one given frequency appears to possibly lead to large qualitative change in the posterior. Useful information may however be extracted from this apparent relative chaos through the use of MCMC with Gaussian process.

4.5.2.2 Gaussian Process Simulations

The maps of posterior probabilities at individual frequencies show that one should expect the location of damage to be perfectly well identified if a range of frequencies is considered at a time. This is verified by considering all the $Q = 150$ frequencies in single Gaussian processes (both for the measurement error likelihood and the damage magnitude prior). In the higher frequency range, the posterior of the damage location estimated by MCMC identifies the actual location ($k = 171$) as the only possible position of damage (*All* the location samples $\theta_1^{(j)}$ are equal to 171 for $j = 15000 \ldots 30000$). The convergence of the samples to the actual location is explained by the conditional likelihood of the location (when other exact parameters are fixed) presented in Fig. 4.3 which shows that only the exact location is likely.

A big advantage of Bayesian methods is that they provide information on all the parameters at the same time. This is demonstrated in Fig. 4.4 where additional statistical information on the damage magnitude is presented.

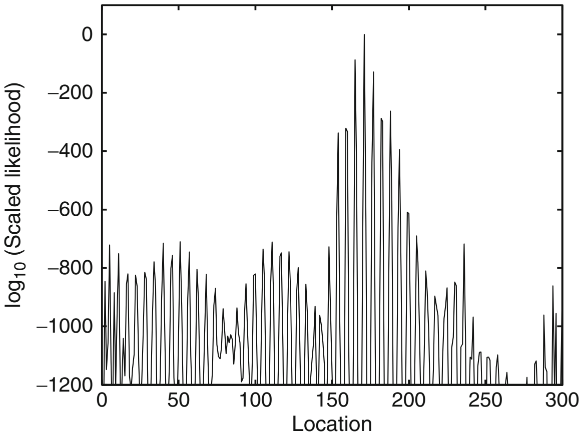


Fig. 4.3 Scaled likelihood of the damage location considering all frequencies of the higher frequency range $\omega \in [0.5, 0.575]$ together in Gaussian processes. Actual location of the damage is $k = 171$

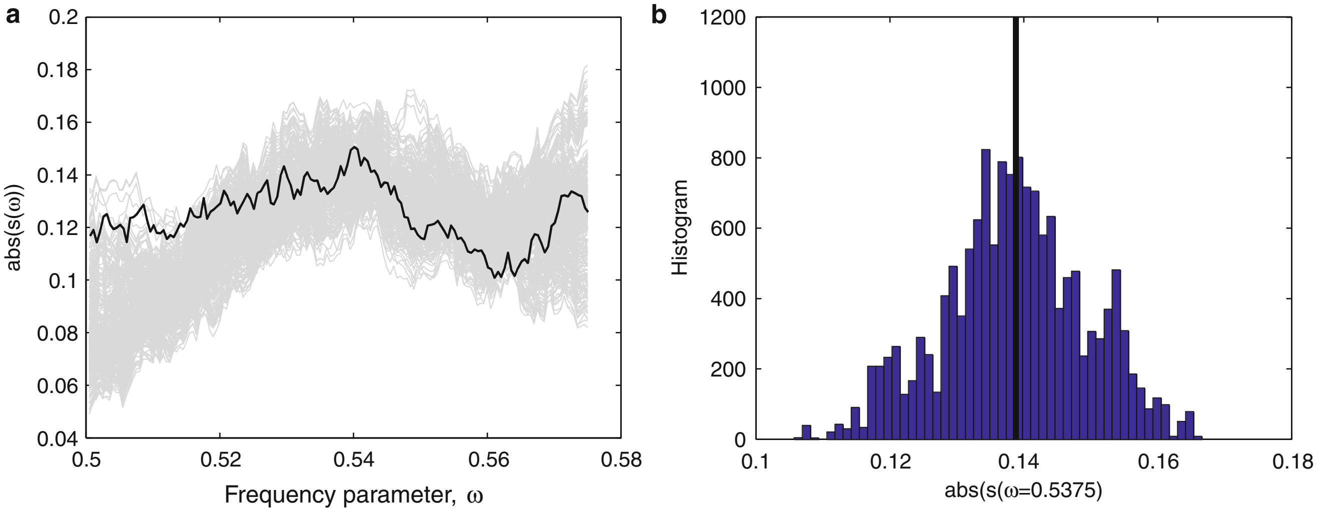


Fig. 4.4 Samples and histogram of the absolute value of the damage magnitude in the higher frequency range. (**a**) Three hundred samples of the converged MCMC chain are superposed in light gray with the exact value. (**b**) An histogram of samples together with the exact value at $\omega = 0.5375$

4.6 Conclusion

The use of Bayesian methods based on transfer function measurements of a structural dynamic system appears to be a feasible technique to identify structural damage. In the simulations reported here, a nominal one-dimensional structure is known to have been damaged at a single location. Measurements of the transfer functions are taken over a range of discrete frequencies. Whilst the value of the damage is unknown, it is assumed known that it has a Gaussian process form (as a function of frequency) with unknown scale and correlation parameters. Error in measurements has a similar form. The location and magnitude (at each frequency) of damage as well as the parameters of the measurement error Gaussian process are parameters of a Bayesian analysis. It is shown that MCMC samples can be used to estimate the posterior distribution of these parameters. The consideration of parallel multi-chain tempering allows one to avoid the main chain becoming stuck in a local region of relatively high probability. Results of Gaussian process walks are compared to results of walks run at individual frequencies. It is shown that the posteriors have a different qualitative behaviour at different individual or range of frequencies. Notably, (statistical) multimodality increases at higher frequencies. Considering several frequencies together nevertheless allows one to identify the location of damage almost deterministically for the particular system considered here. This appears to be true both at low and higher frequency ranges.

Acknowledgements The authors gratefully acknowledge the financial support provided by the Engineering and Physical Sciences Research Council under grant EP/G056765/1.

References

1. Beck JL, Au S-K (2002) Bayesian updating of structural models and reliability using Markov Chain Monte Carlo simulation. J Eng Mech 128 (4):380–391
2. Au S-K (2011) Fast Bayesian FFT method for ambient modal identification with separated modes. J Eng Mech 137(3):214–226
3. Beck JL, Yuen K-V (2004) Model selection using response measurements: Bayesian probabilistic approach. J Eng Mech 130(2):192–203
4. Lecomte C, Mace BR, Forster JJ, Ferguson NS (2010) Bayesian localisation of damage in a linear dynamic system. In: USD2010 international conference on uncertainty in structural dynamics, Leuven, pp 4953–4964
5. Nichols JM, Link WA, Murphy KD, Olson CC (2010) A Bayesian approach to identifying structural nonlinearity using free-decay response: application to damage detection in composites. J Sound Vib 329:2995–3007
6. Yuen K-V (2010) Bayesian methods for structural dynamics and civil engineering. Wiley, Singapore
7. Lecomte C, Forster JJ, Mace BR, Ferguson NS (2010) Bayesian inference for uncertain dynamic systems. In: Proceedings of the tenth international conference on recent advances in structural dynamics (RASD2010), Southampton, Paper 162, 12–14 July 2010
8. Lecomte C (2010) Zero and root loci of disturbed spring-mass systems. In: Proceedings of the tenth international conference on recent advances in structural dynamics (RASD2010), Southampton, Paper 161, 12–14 July 2010
9. Sherman J, Morrison WJ (1950) Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. Ann Math Stat 21:124–127
10. Lecomte C (2011) Vibration analysis of an ensemble of structures using a theory of stochastic linear systems. In: IUTAM symposium on the vibration analysis of structures with uncertainties, St Petersburg, July 2009, Springer, pp 301–315
11. O'Hagan A, Forster JJ (2004) Statistical methods, Kendall's advanced theory of statistics. vol 2B: Bayesian statistics, 2nd (edn). Imprint: London: Arnold, pp 480. ISBN: 0340807520

Chapter 5
Identification of Hysteretic Systems Using NARX Models, Part I: Evolutionary Identification

K. Worden and R.J. Barthorpe

Abstract Although there has been considerable work on the identification of hysteretic systems over the years, there has been comparatively little using discrete NARX or NARMAX models. One of the reasons for this may be that many of the common continuous-time models for hysteresis, like the Bouc-Wen model are nonlinear in the parameters and incorporate unmeasured states, and this makes a direct analytical discretisation somewhat opaque. Because NARX models are universal in the sense that they can model any input–output process, they can be applied directly without consideration of the hysteretic nature; however, if the polynomial form of NARX were to be used for a Bouc-Wen system, the result would be input-dependent because of the non-polynomial (indeed discontinuous) nature of the original model. The objective of the current paper is to investigate the use of NARX models for Bouc-Wen systems and to consider the use of non-polynomial basis functions as a potential means of alleviating any input-dependence. As the title suggests, the parameter estimation scheme adopted will be an evolutionary one based on Self-Adaptive Differential Evolution (SADE). The paper will present results for simulated data.

Keywords Hysteresis • NARX models • Nonlinear system identification • Self-Adaptive Differential Evolution (SADE) • The Bouc-Wen model

5.1 Introduction

This paper represents the first of two in a short sequence discussing aspects of the system identification problem for hysteretic systems i.e. systems with memory. The basis for discussion here is the extremely versatile parametric form for the modelling of hysteretic systems provided by the Bouc-Wen model [1, 2]. The model form has been widely adopted and the literature on the Bouc-Wen model is extensive. There is no intention here to survey the said literature and the reader is instead referred to the recent book [3] for a comprehensive guide. In the past, one of the main problems associated with adopting the Bouc-Wen form for a system model was the identification of the model parameters, as the problem is complicated by nonlinearity in the parameters and the presence of unmeasured states. These issues have been largely dealt with by the adoption of evolutionary identification schemes; the Differential Evolution (DE) scheme discussed in [4] proved particularly effective, as did the later self-adaptive variant [5].

Having determined an effective identification scheme for the Bouc-Wen system, it is possible to turn ones attention to the model form itself in order to see if one can develop a more convenient or effective variant. The first thing which comes to mind is that the Bouc-Wen system is described by a continuous-time differential equation of motion and it is often more convenient for system identification purposes to adopt a discrete form. In terms of discrete models of nonlinear systems, it is probably fair to say that the NARMAX (Nonlinear Auto-Regressive Moving Average with Exogenous Inputs) or NARX form is the 'standard model' for system identification [6, 7]. One issue with the NARMAX or NARX models is that one has to specify the basis of model terms in use and, although other bases have been adopted successfully, a polynomial basis is most usually adopted. A problem with this approach is that the fitted model may be input-dependent; although the polynomial basis can approximate *any* function with any desired accuracy, the approximation will depend on the level of

---

K. Worden • R.J. Barthorpe (✉)
Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK
e-mail: k.worden@sheffield.ac.uk; r.j.barthorpe@sheffield.ac.uk

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_5,

excitation if the system of interest does not truly have polynomial nonlinearities—as in the case of the Bouc-Wen system. A way around the problem of input-dependence is to include in the NARMAX or NARX basis, terms which represent the true nonlinearities of the system. This will be the approach adopted in the current paper in order to derive a NARX (noise will not be considered here so the moving average part is not estimated) class of models for Bouc-Wen systems. It will be shown that the Bouc-Wen NARX class also potentially overcomes other problems which arise in fitting polynomial NARX models to systems of Bouc-Wen type.

The layout of the paper is as follows: Sect. 5.2 briefly describes the Bouc-Wen model of hysteresis considered here and explains the evolutionary approach to system identification. Sect. 5.3 describes how the NARX Bouc-Wen model is derived and identification results based on simulated data are presented in Sect. 5.4. The paper concludes with a little discussion in Sect. 5.5.

5.2 Identification of the Bouc-Wen Hysteresis Model

The system of interest here will be the general Single-Degree-of-Freedom (SDOF) hysteretic system described in the terms of Wen [2]. The system is a continuous-time one and is represented below with $z(y, \dot{y})$ the hysteretic part of the restoring force,

$$m\ddot{y} + c\dot{y} + ky + z(y, \dot{y}) = x(t) \tag{5.1}$$

where m, c and k are the mass, damping and stiffness parameters respectively, $y(t)$ is the system displacement response and $x(t)$ is the excitation force; overdots denote differentiation with respect to time. The system in (5.1) is a restricted form of the Bouc-Wen model as it has a linear non-hysteretic internal force; in general one could have an arbitrarily nonlinear non-hysteretic component. The hysteretic component is defined by Wen [2] via an additional equation of motion,

$$\dot{z} = A\dot{y} - \alpha|\dot{y}|z^n - \beta\dot{y}|z^n| \tag{5.2}$$

for n odd, or,

$$\dot{z} = A\dot{y} - \alpha|\dot{y}|z^{n-1}|z| - \beta\dot{y}z^n \tag{5.3}$$

for n even.

The parameters α, β and n govern the shape and smoothness of the hysteresis loop. A simplification from the point of view of parameter estimation is that the stiffness term in (5.1) can be combined with the term $A\dot{y}$ in the state equation for z i.e. one can set $k = 0$. As a system identification problem, this set of equations has presented a number of difficulties in the past, foremost are: (a) the state z is not measurable and therefore it is not possible to use (5.2) or (5.3) directly in a least-squares formulation, and (b) The parameter n enters the state equations (5.2) and (5.3) in a nonlinear way; this means that a linear least-squares approach is not applicable to the estimation of the full parameter set. However, it has been shown that there are a number of identification procedures which can accommodate the difficulties peculiar to the Bouc-Wen system; as discussed in [5], the differential evolution algorithm [8], and it's adaptive variant [9, 10], offer a powerful approach to the parameter estimation problem. As in all evolutionary optimisation procedures, a population of possible solutions (here, the vector of parameter estimates), is iterated in such a way that succeeding generations of the population contain better solutions to the problem in accordance with the Darwinian principle of 'survival of the fittest'. The problem is framed as a minimisation problem with the cost function defined as a normalised mean-square error between the 'measured' data and that predicted using a given parameter estimate, i.e.

$$J(m, c, k, \alpha, \beta) = \frac{100}{N\sigma_y^2} \sum\nolimits_{i=1}^{N} (y_i - \hat{y}_i(m, c, A, \alpha, \beta, n))^2 \tag{5.4}$$

where ${\sigma_y}^2$ is the variance of the 'measured' sequence of displacements y_i and the caret denotes a predicted quantity. The identification scheme adopted in this paper is based on the Self-Adaptive Differential Evolution (SADE) algorithm mentioned above and illustrated previously on Bouc-Wen systems in [5].

Before moving on to the NARX form of the Bouc-Wen model, some description of the simulated data used here for illustration is given. The main data for the current study were generated by numerical simulation. The coupled equations (5.1) and (5.3) were integrated forward in time in Matlab [11] using a fixed-step fourth-order Runge–Kutta scheme for initial value problems [12]. The fixed-step solver was used in order to avoid a subtle issue with adaptive solvers which was raised in [5].

The parameters for the baseline system adopted here were: $m = 1, c = 20, \alpha = 1.\ 5, \beta = -1.5, A = 6680.\ 0$ and $n = 2$. The excitation was a Gaussian white noise sequence with mean zero and standard deviation 9.92. These are the same parameter values as those adopted in [5]. The step-size (or sampling interval) was initially taken as 0.004 s, corresponding to a sampling frequency of 250 Hz. As the main aim of the paper is concerned with discussing the new NARX model form, the data are not corrupted by noise, this matter will be investigated at a later date. The 'training set' or identification set used here was composed of 1,000 points corresponding to a record duration of 4 s.

5.3 The NARX Bouc-Wen Model

Without any real loss of generality here, the Bouc-Wen system will be specified by (5.1) and (5.3) for n even. In fact, to simplify matters further, the parameter n will be fixed at the value 2 in order to remove it from the identification problem as it does not affect the arguments which will follow. As a first step in deriving a discrete model for the Bouc-Wen system, a continuous-time form independent of z will be obtained. The first step in this process is simply to differentiate equation (5.1) to form,

$$m\ddot{v} + c\dot{v} + \dot{z} = \dot{x}(t) \tag{5.5}$$

with the assumption that $k = 0$ as discussed in the last section. The natural notation $v = \dot{y}$ has been adopted. Substituting for $\{v = \dot{z}\}$ from (5.3) then gives,

$$m\ddot{v} + c\dot{v} + Av - \alpha|v|z^{n-1}|z| - \beta v z^n = \dot{x}(t) \tag{5.6}$$

Finally, one substitutes,

$$z = x - cv - m\dot{v} \tag{5.7}$$

which is a trivial rearrangement of (5.1), to give,

$$m\ddot{v} + c\dot{v} + Av - \alpha|v|(x - cv - m\dot{v})^{n-1}|x - cv - m\dot{v}| - \beta v(x - cv - m\dot{v})^n = \dot{x}(t) \tag{5.8}$$

The latter equation is independent of the unmeasured state z; however, this freedom has been achieved at the expense of making the form very nonlinear in the excitation x. Discretisation of the Bouc-Wen system now proceeds in the standard manner. First, one adopts discrete approximations to the first and second derivative terms in (5.8),

$$\dot{v} \approx \frac{1}{\Delta t}(v_i - v_{i-1}) \tag{5.9}$$

$$\ddot{v} \approx \frac{1}{\Delta t^2}(v_{i+1} - 2v_i + v_{i-1}) \tag{5.10}$$

where $v_i = v(i\Delta t)$ and Δt is the sampling interval. With this prescription for the derivatives (which is by no means compulsory), one finds that,

$$x - cv - m\dot{v} \approx x_i - cv_i - \frac{m}{\Delta t}(v_i - v_{i-1}) = x_i + \gamma_1 v_i + \gamma_2 v_{i-1} \tag{5.11}$$

with the latter equation serving to define the coefficients γ_i. By a similar process, one finds the discrete form,

$$m\ddot{v} + c\dot{v} + Av \approx a_0 v_{i+1} + a_1 v_i + a_2 v_{i-1} \tag{5.12}$$

and similar for the x terms.

Putting all of this together in (5.8), redefining a_i by $-a_i/a_0$ throughout and shifting the time index yields finally,

$$v_i = \sum\nolimits_{j=1}^{2} a_j v_{i-j} + \sum\nolimits_{j=1}^{2} b_j x_{i-j} + \alpha|v_{i-1}|\left(x_{i-1} + \sum\nolimits_{j=1}^{2}\gamma_j v_{i-j}\right)\left|x_{i-1} + \sum\nolimits_{j=1}^{2}\gamma_j v_{i-j}\right|^{n-1} + \beta v_i\left(x_{i-1} + \sum\nolimits_{j=1}^{2}\gamma_j v_{i-j}\right)^n \tag{5.13}$$

on restoring a general (even) n. The parameters which require estimation here are (a_1, a_2, b_1, b_2, α, β, γ_1, γ_2) and so there are eight in total in this, the minimal NARX Bouc-Wen model. Allowing for the fact that more accurate approximations for the derivatives could have been used, one arrives at a more general NARX Bouc-Wen form,

$$v_i = \sum_{j=1}^{p} a_j v_{i-j} + \sum_{j=1}^{q} b_j x_{i-j} + \alpha |v_{i-1}| \left(x_{i-1} + \sum_{j=1}^{r} \gamma_j v_{i-j}\right) \left|x_{i-1} + \sum_{j=1}^{r} \gamma_j v_{i-j}\right|^{n-1} + \beta v_i \left(x_{i-1} + \sum_{j=1}^{r} \gamma_j v_{i-j}\right)^n \quad (5.14)$$

however, only the minimal form (5.13) will be discussed further in this paper.

As promised, (5.13) is tailored to the nonlinear terms expected from a Bouc-Wen system; however, it will offer other advantages over the polynomial form of the NARX model. First, it must be observed that the NARX Bouc-Wen system is still nonlinear-in-the-parameters. This is not an issue for the evolutionary optimisation scheme adopted here, but it is a problem for a polynomial NARX approach. The problem is avoided by making the polynomial approximation,

$$|z| \approx c_0 + c_2 z^2 \quad (5.15)$$

which renders all terms in (5.13) polynomial at the expense of making the model input-dependent as discussed in the introduction. A further problem now is that the highest-order polynomial term needed to represent the minimal NARX Bouc-Wen model (with $n = 2$) will be fifth-order. This means that the polynomial NARX structure detection algorithm must search over all candidate basis terms up to fifth order and this is likely to be prohibitively expensive.

The final advantage of the NARX Bouc-Wen form over the continuous form is shared by NARX models in general and is that forward prediction is a simple iteration rather than requiring the use of a differential equation solver. This is particularly significant in the context of the evolutionary identification scheme as predictions have to made for all candidate parameter sets in the SADE population. In all other respects, the SADE approach to identification is identical to that for the continuous-time Bouc-Wen model except for the fact that one now minimises a cost function,

$$J(a_1, a_2, b_1, b_2, \alpha, \beta, \gamma_1, \gamma_2) = \frac{100}{N\sigma_y^2} \sum_{i=1}^{N} \left(y_i - \hat{y}_i(a_1, a_2, b_1, b_2, \alpha, \beta, \gamma_1, \gamma_2)\right)^2 \quad (5.16)$$

(assuming n fixed).

5.4 Illustration on Simulated Data

The application of the NARX Bouc-Wen model structure will later be shown on data simulated from a continuous-time Bouc-Wen model; however, in order to develop the algorithm an even simpler benchmark was developed first. The parameters $m = 1$, $c = 20$, $\alpha = 1.\ 5$, $\beta = -1.5$, $A = 6680.\ 0$ and $n = 2$ discussed in Sect. 5.2 were adopted for the continuous-time Bouc-Wen system and the corresponding parameters for the NARX model were calculated. With a sampling interval of $\Delta t = 0.\ 004$, this led to $a_1 = 1.813, a_2 = -0.92, b_1 = 0.004, b_2 = -0.004, \alpha = 0.000024$, $\beta = -0.000024, \gamma_1 = -270, \gamma_2 = 250$. Adopting a Gaussian white noise sequence with zero-mean and standard deviation 9.92, the exact discrete form in (5.13) was used to generate reference data for v_i. 1,000 points of identification data were generated. In the first stage of the identification process, one usually has little prior knowledge of the parameters and this will be reflected here by assuming initial bounds on the parameters which range over an order of magnitude above and below the true values; this means the initial ranges are: [0. 1813, 18. 13] for a_1, $[-9.2, -0.092]$ for a_2 etc. On applying the SADE approach a problem immediately arose. This issue related to the stability of the NARX model; it transpired that the randomly-generated initial population of parameter vectors in SADE almost always generated unstable predictions. After some investigation, it became clear that the main source of the problem was the γ_1 parameter; when the corresponding term was removed from the model in (5.13), the stability problem was reduced to the extent that stable predictions were generated. A further source of instability was introduced by any attempt to include initial conditions in the forward prediction of the velocity data; while this bears further investigation, the problem was 'solved' here by prescribing $v_1 = v_2 = 0$. A result of this is that an initial transient appears in all predicted velocity time-series and this lower bounds the prediction accuracy which can be achieved by the identification. New simulated data were generated as above, but with $\gamma_1 = 0$.

The hyperparameters of the SADE algorithm adopted here are the same as those described in [5], which is a further confirmation of the robustness of the adaptive algorithm. The only modifications were that a population size of 200 was

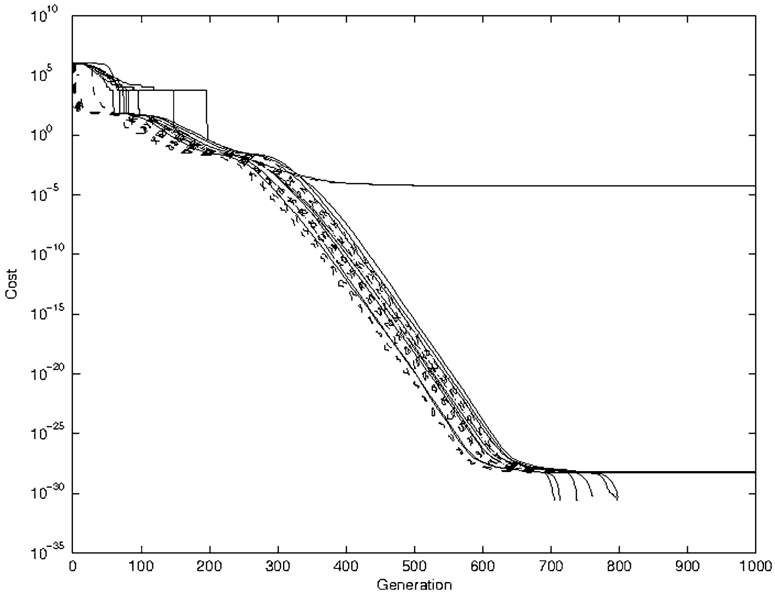


Fig. 5.1 Evolution of cost functions for identification of exact NARX data (*solid*: average cost over population; *dashed*: minimum cost over generation)

used and the maximum number of generations was set at 1,000. Because the algorithm begins from a random initial population, 10 runs of the algorithm were made with different initial populations. When the data generated by the 'exact' NARX form were identified using the SADE algorithm, all parameters were estimated to machine accuracy. The evolution of the cost functions over the 10 runs is shown in Fig. 5.1; the descent to machine precision over 600 generations or so is shown clearly. There is no purpose in showing how the model-predicted data compares with the 'true' data as they would simply appear superimposed.

Having established that the algorithm is implemented correctly (verification) it became necessary to show that it generated appropriate models in more demanding cases (validation). This was accomplished here by identifying data generated from a continuous-time Bouc-Wen process as described in Sect. 5.2. In this case, the 'true' NARX Bouc-Wen parameters were not known, so it was not immediately obvious how to set initial bounds on the parameters. However, based on the general behaviour of the discretised models e.g. $a_i = O(1)$, $b_i = O(\Delta t)$ etc. it was possible to set bounds which gave initial convergence of the predictions. These 'trial and error' bounds were: $[-2, 2]$ for a_1, a_2, b_1 and b_2; $[-0.01, 0.01]$ for α and β and $[-500, 500]$ for γ_2. With the initial ranges specified, the SADE algorithm converged to a solution with cost value 4.90% in 4 of the 10 runs; the optimal parameter set found was: $a_1 = 1.819, a_2 = -0.92, b_1 = 0.00344, b_2 = -0.00361, \alpha = -0.0000104, \beta = -0.0000222, \gamma_2 = 13.6$. The linear term parameter estimates agree with those of the algebraically-derived form of the discrete model, while the nonlinear parameters are somewhat different; this is to be expected because of the inexact nature of the discretisation. Figure 5.2 shows the model predicted data compared to the 'true' data.

Previous experience with the cost function used here as shown that a value of 5% indicates a good model fit while 1% indicates a truly excellent fit. The value of 4.9% obtained here, based on noise-free data was considered a little disappointing. There are various reasons why the model fit could be less than perfect, but the first one that was considered here was the accuracy of the discretisation. In order to investigate this, data were simulated from the same continuous-time Bouc-Wen system as before, except that 'sampling' was carried out at 1,000 Hz (i.e. $\Delta t = 0.001$s). When the identification scheme was applied to this data, an error of 1.99% was found, which showed that sampling was an issue; in practice, this could also be addressed by using more lags in the model i.e. adopting the more general form in (5.14). In contrast to the case for the 'exact' NARX data, the SADE algorithm only found a good solution in one of 10 runs (as shown in the cost evolutions in Fig. 5.3). In 4 of the runs, the algorithm failed to escape from regions of the parameter space leading to instability of the predictions. With the best parameter set, which was: $a_1 = 1.97283, a_2 = -0.9802159, b_1 = 0.000923679, b_2 = -0.000925449, \alpha = 0.000005516, \beta = -0.000008497, \gamma_2 = -116.53$, the comparison between the predicted and true data was as shown in Fig. 5.4. The prediction error in this case is dominated by the component in the initial transient which appears for reasons discussed earlier; the data later in the time series show an excellent fit of the model.

The observant reader will no doubt question why parameter estimates are being quoted to six significant figures when previously three sufficed. The reason is interesting; when only three significant figures are used in order to make the model predictions, the results in this case are rather poor as shown in Fig. 5.5. This shows that the predictions are very sensitive to

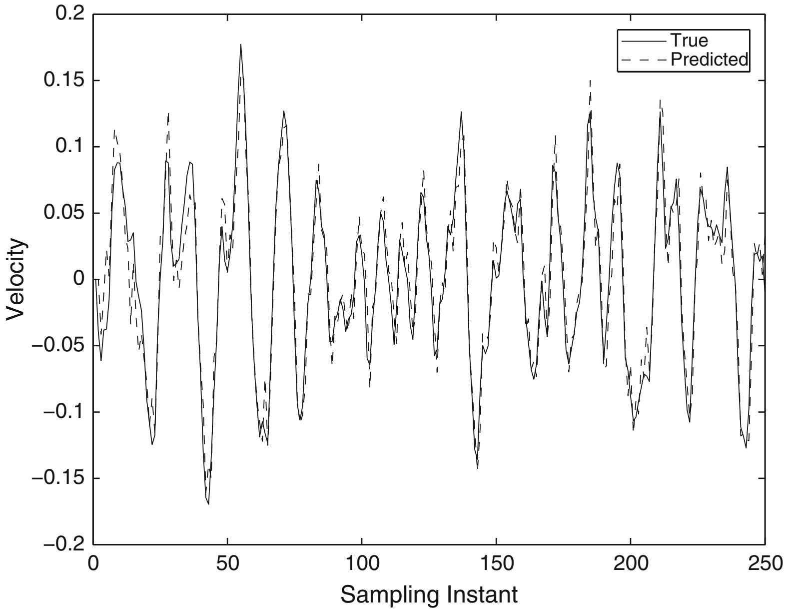


Fig. 5.2 Comparison between model-predicted and true data for NARX Bouc-Wen model fitted to simulated continuous-time data: sampling frequency 250 Hz

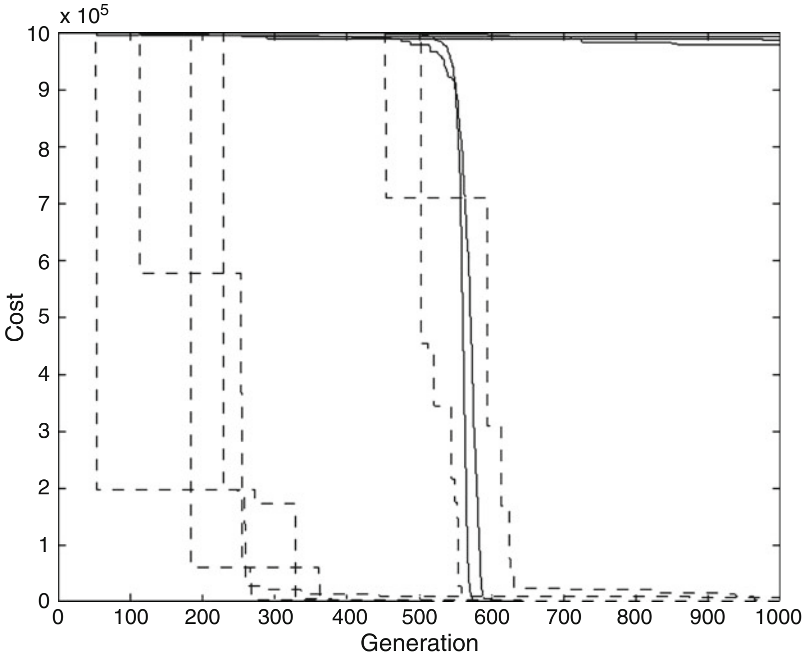


Fig. 5.3 Evolution of cost functions for identification of continuous-time NARX data sampled at 1,000 Hz (*solid*: average cost over population; *dashed*: minimum cost over generation)

the parameter estimates. The extent of this sensitivity bears further investigation and could be considered by using Bayesian sensitivity analysis as discussed in the context of the continuous-time Bouc-Wen system in [13].

As a final consideration, a linear model was fitted to the 1,000 Hz data, i.e., the parameters α, β and γ_2 were constrained to zero in the identification. This exercise gave a cost value of 3.7% showing that the nonlinear terms are needed.

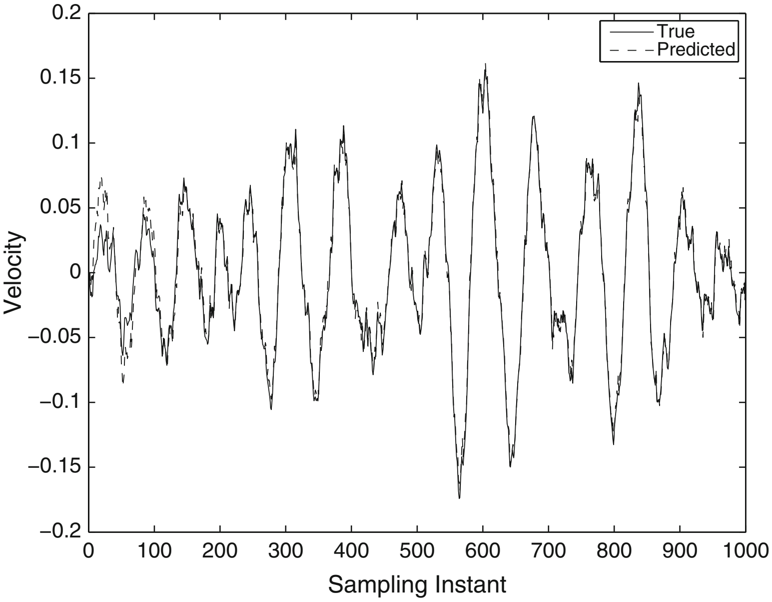


Fig. 5.4 Comparison between model-predicted and true data for NARX Bouc-Wen model fitted to simulated continuous-time data: sampling frequency 1,000 Hz

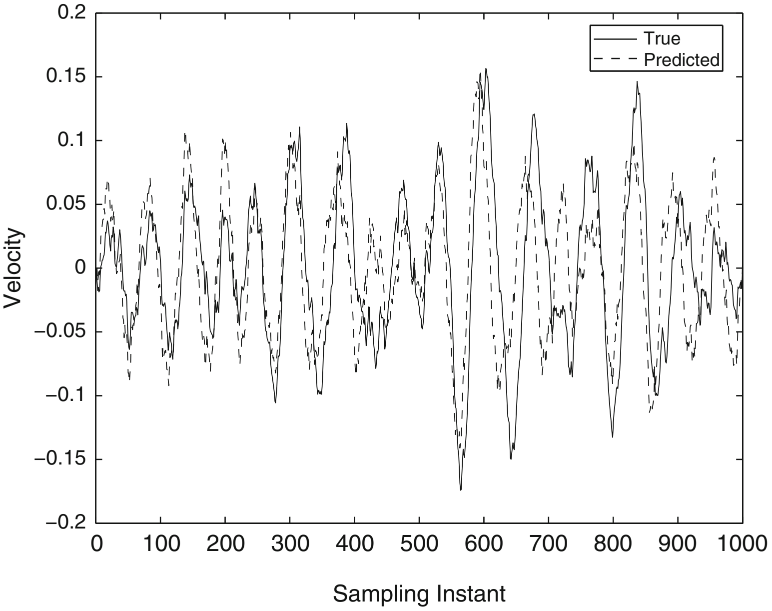


Fig. 5.5 Comparison between model-predicted and true data for NARX Bouc-Wen model fitted to simulated continuous-time data: sampling frequency 1,000 Hz. Parameters truncated to three significant figures

5.5 Conclusions

Lengthy conclusions do not appear to be warranted here as the work presented is largely of an exploratory nature and much remains to be followed up. A discrete NARX structure has been introduced tailored to hysteretic systems of Bouc-Wen type in the hope that input-independent NARX models can be found for hysteretic systems with small parameter sets. It has been shown that the SADE algorithm provides an effective means of estimating parameters for the NARX Bouc-Wen structure. Among the many issues raised here that bear further investigation are:

- It will be important in general to understand the stability properties of the NARX Bouc-Wen form, at least as far as it is necessary to generate initial populations for the SADE algorithm which do not all blow-up. At least, the influence of the a_1 parameter and the initial conditions needs further investigation.

- It will be necessary to consider more general forms for the NARX model, both in terms of the number of lags in the linear components in the model and in terms of allowing a nonlinear non-hysteretic restoring force.
- Sensitivity to the parameter uncertainty appears to be an issue, this may be critical when the identification algorithm is applied to real noisy data.
- Estimation of parameter confidence (related to but not the same as sensitivity) is an issue and can be investigated using various approaches.
- One could add basis terms for a range of n values and then determine which make a significant contribution.

As indicated in the introduction, this paper is one of a pair; the second in the series will consider a Bayesian system identification scheme based on Markov-Chain Monte Carlo and will by its nature encounter some of the issues indicated above in that separate context.

Acknowledgements The authors would like to sincerely thank Dr. Graeme Manson for providing Matlab code implementing the SADE algorithm.

References

1. Bouc R (1967) Forced vibration of mechanical system with hysteresis. In: Proceedings of 4th conference on nonlinear oscillation, Prague
2. Wen Y (1976) Method for random vibration of hysteretic systems. ASCE J Eng Mech Div 102:249–263
3. Ikhouane F, Rodellar J (2007) Systems with hysteresis: analysis, identification and control using the Bouc-Wen model. Wiley-Blackwell, Chichester/Hoboken
4. Kyprianou A, Worden K, Panet M (2001) Identification of hysteretic systems using the differential evolution algorithm. J Sound Vib 248:289–314
5. Worden K, Manson G (2011) On the identification of hysteretic systems. Part I: fitness landscapes and evolutionary identification. Mech Sys and Sig Proc (In Press). http://dx.doi.org/10.1016/j.ymssp.2012.01.004
6. Leontaritis IJ, Billings SA (1985) Input–output parametric models for nonlinear systems, part I: deterministic nonlinear systems. Int J Control 41:303–328
7. Leontaritis IJ, Billings SA (1985) Input–output parametric models for nonlinear systems, part II: stochastic nonlinear systems. Int J Control 41:329–344
8. Price K, Storn R (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 11:341–359
9. Qin AK, Suganthan PN (2005) Self-adaptive differential evolution algorithm for numerical optimization. In: Proceedings of IEEE congress on evolutionary computation (CEC 2005), Edinburgh
10. Huang VL, Qin AK, Suganthan PN (2006) Self-adaptive differential evolution algorithm for constrained real-parameter optimization. In: Proceedings of IEEC congress on evolutionary computation (CEC 2006), Vancouver, Canada, pp 17–24
11. The Mathworks Inc., (2004) MATLAB version 7 Natick, Massachusetts
12. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical recipes: the art of scientific computing, 3rd edn. Cambridge University Press, New York/Cambridge
13. Worden K, Becker WE (2011) Submitted to mechanical systems and signal processing. On the identification of hysteretic systems. Part II: Bayesian sensitivity analysis and parameter confidence. Mech Sys and Sig Proc. (In Press). http://dx.doi.org/10.1016/j.ymssp.2012.01.005

Chapter 6

Identification of Hysteretic Systems Using NARX Models, Part II: A Bayesian Approach

K. Worden, R.J. Barthorpe, and J.J. Hensman

Abstract Following on from the first part of this short sequence, this paper will investigate the use of a Bayesian methodology for the identification of Bouc-Wen hysteretic systems by NARX models. The approach—based on Markov Chain Monte Carlo—offers a number of advantages over the evolutionary approach of the first paper. Among them are the ability to sample from the probability density functions of the parameters in order to develop nonparametric estimators and the possibility of selecting model terms in a principled manner. The paper will investigate the use of the Deviance Information Criterion (DIC) as a means of selecting model terms, specifically the special basis functions developed for the Bouc-Wen system in Part I. Results for simulated data will be given.

Keywords Bayesian inference • Hysteresis • Markov Chain Monte Carlo (MCMC) • Nonlinear system identification • The Bouc-Wen model

6.1 Introduction

This paper represents the second in a two-part series. In the first part [1], the authors showed how to transform the Bouc-Wen hysteretic system [2, 3] into a discrete NARX (Nonlinear Auto-Regressive model with eXogeneous inputs) form [4, 5]. The resulting model structures are tailored to the nonlinear terms expected from a Bouc-Wen system in order to allow the possibility of establishing input-independent models rather than the input-dependent models that would be expected from polynomial NARX approximations; however, as noted in [1], they will offer other advantages over polynomial NARX. If one assumes a linear non-hysteretic force, there is no structure detection problem for the models and the system identification problem is reduced to parameter estimation. In [1], the parameter estimation algorithm was an evolutionary scheme based on Self-Adaptive Differential Evolution (SADE). This paper will present an alternative parameter estimation method based on a Bayesian Markov Chain Monte Carlo (MCMC) approach. There are many advantages of such an approach as discussed in [6]; among them is the ability to generate confidence intervals for parameters; in fact the scheme allows one to actually sample from the probability density functions of the parameters.

The layout of the paper is as follows: Sect. 6.2 describes the newly-proposed Bouc-Wen NARX model structures. In Sect. 6.3 the necessary background to the MCMC scheme is described and the results of applying the method are given in Sect. 6.4. The paper concludes with a little discussion in Sect. 6.5.

K. Worden • R.J. Barthorpe (✉)
Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK
e-mail: k.worden@sheffield.ac.uk; r.j.barthorpe@sheffield.ac.uk

J.J. Hensman
Department of Neuroscience, Sheffield Institute for Translational Neuroscience (SITraN), The University of Sheffield, 385a Glossop Road, Sheffield S1 3JD, UK

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_6,

6.2 The NARX Bouc-Wen Model

As in the first paper in this short sequence [1], the system of interest here is specified as the general Single-Degree-of-Freedom (SDOF) hysteretic system described in the terms of Wen [3] as,

$$m\ddot{y} + c\dot{y} + ky + z(y, \dot{y}) = x(t) \tag{6.1}$$

$y(t)$ is the system displacement response and $x(t)$ is the excitation force; overdots denote differentiation with respect to time. As represented above, $z(y, \dot{y})$ as the hysteretic part of the restoring force; the non-hysteretic part is taken here as linear, where m, c and k are the mass, damping and stiffness parameters respectively, The hysteretic component of the internal force is defined by Wen [3] via the additional equation of motion,

$$\dot{z} = A\dot{y} - \alpha|\dot{y}|z^n - \beta\dot{y}|z^n| \tag{6.2}$$

for n odd, or,

$$\dot{z} = A\dot{y} - \alpha|\dot{y}|z^{n-1}|z| - \beta\dot{y}z^n \tag{6.3}$$

for n even.

The parameters α, β and n govern the shape and smoothness of the hysteresis loop.

In [1], a discrete NARX (Nonlinear Auto-Regressive model with eXogeneous inputs) form for Bouc-Wen systems. For even n, the minimal Bouc-Wen NARX model is given,

$$v_i = \sum\nolimits_{j=1}^{2} a_j v_{i-j} + \sum\nolimits_{j=1}^{2} b_j x_{i-j} + \alpha|v_{i-1}|\left(x_{i-1} + \sum\nolimits_{j=1}^{2} \gamma_j v_{i-j}\right)^n + \beta v_i \left|x_{i-1} + \sum\nolimits_{j=1}^{2} \gamma_j v_{i-j}\right|^n \tag{6.4}$$

for n odd and,

$$\begin{aligned} v_i = {} & \sum\nolimits_{j=1}^{2} a_j v_{i-j} + \sum\nolimits_{j=1}^{2} b_j x_{i-j} + \alpha|v_{i-1}|\left(x_{i-1} + \sum\nolimits_{j=1}^{2} \gamma_j v_{i-j}\right)|x_{i-1} \\ & + \sum\nolimits_{j=1}^{2} \gamma_j v_{i-j}|^{n-1} + \beta v_i \left(x_{i-1} + \sum\nolimits_{j=1}^{2} \gamma_j v_{i-j}\right)^n \end{aligned} \tag{6.5}$$

for n even. Note that the primary variable of interest is now the velocity $v = dy/dt/$ The parameters which require estimation here are (a_1, a_2, b_1, b_2, α, β, γ_1, γ_2) and so there are eight in total in the minimal NARX Bouc-Wen model. Allowing for the fact that more accurate approximations for the derivatives could have been used, one arrives at the more general NARX Bouc-Wen forms,

$$v_i = \sum\nolimits_{j=1}^{p} a_j v_{i-j} + \sum\nolimits_{j=1}^{q} b_j x_{i-j} + \alpha|v_{i-1}|\left(x_{i-1} + \sum\nolimits_{j=1}^{r} \gamma_j v_{i-j}\right)^n + \beta v_i \left|x_{i-1} + \sum\nolimits_{j=1}^{r} \gamma_j v_{i-j}\right|^n \tag{6.6}$$

for n odd and,

$$\begin{aligned} v_i = {} & \sum\nolimits_{j=1}^{p} a_j v_{i-j} + \sum\nolimits_{j=1}^{q} b_j x_{i-j} + \alpha|v_{i-1}|\left(x_{i-1} + \sum\nolimits_{j=1}^{r} \gamma_j v_{i-j}\right)|x_{i-1} \\ & + \sum\nolimits_{j=1}^{r} \gamma_j v_{i-j}|^{n-1} + \beta v_i \left(x_{i-1} + \sum\nolimits_{j=1}^{r} \gamma_j v_{i-j}\right)^n \end{aligned} \tag{6.7}$$

for n even. However, only the minimal forms (6.4) and (6.5) will be discussed further in this paper. For reasons relating to the stability of predictions that are discussed in [1], the parameter γ_1 is set to zero here and there are thus seven parameters for estimation: (a_1, a_2, b_1, b_2, α, β, γ_1, γ_2)

As discussed in the introduction, if one assumes a linear non-hysteretic force, there is no structure detection problem for the models and the system identification problem is reduced to parameter estimation. The estimation strategy in this paper is a Bayesian method based on Markov Chain Monte Carlo (MCMC) and this is defined in the following section. It has to be said that, in order to make the present paper as self-contained as possible, substantial portions of earlier papers by the authors on the MCMC approach are reproduced including material from [6].

6.3 Bayesian Inference and Markov Chain Monte Carlo

6.3.1 Bayesian Inference for System Identification

The problem of system identification is simply stated: given measured data from a structure or system of interest, how does one infer the equations of motion which 'generated' the data. The main classical, almost exclusively linear, approaches to the problem are well-described in [7, 8]. Although the problem can be stated simply, mathematically it can be extremely ill-posed even if the underlying equations are assumed to be linear in the parameters of interest [9]. If the equations of motion are not linear in the parameters of interest, the difficulties multiply.

Another issue is concerned with *confidence* in any parameter estimates. This issue is a result of the fact that measurements or data from a system will, in reality, almost always be contaminated by random noise. Suppose that one has a set of data $D = \{(x_i, y_i), i = 1, \ldots, N\}$ of sampled system inputs x_i and outputs y_i. If one assumes that there is no measurement noise, then the application of an identification algorithm, will yield (subject to the conditioning of the problem), a deterministic estimate of the system parameters $\underline{w}$,

$$\underline{w} = id(D) \tag{6.8}$$

where the function *id* represents the application of the identification algorithm to the data D. In the context of the current problem (a_1, a_2, b_1, b_2, α, β, γ_1, γ_2). Now, if noise $\varepsilon(t)$ is present on the input or output data (or both), $\underline{w}$ will become a random variable conditioned on the data. The problem is no longer to find an *estimate* of $\underline{w}$, but rather to specify the belief in its value. For simplicity, it will be assumed that the noise is Gaussian with (unknown) variance σ. The parameter σ will be subsumed into $\underline{w}$, since it is to be inferred along with the model parameters. In probabilistic terms, instead of (6.8) one now has,

$$\underline{w} \sim p(\underline{w}|D, \mathcal{M}) \tag{6.9}$$

where $\mathcal{M}$ represents the choice of model (the Bouc-Wen system with Gaussian noise in the case here). In the presence of measurement noise, the most one can hope to learn from any data is the probability density function of the parameter estimates; in fact, in a probabilistic context, this is everything. The Bayesian approach to the parameter estimation problem described here was motivated by the recent work in [10, 11] where Bayesian approaches to determining the parameters of continuous-time differential equations were proposed and developed. Such an approach was applied in a structural dynamic context in [6]; the method described here differs only in that parameters are estimated for discrete-time systems.

The usual objective of system identification is to provide a predictive model i.e. a mathematical model which can estimate or predict system outputs if a different system input were provided. In the probabilistic context described above, the best that one could do is to determine a predictive distribution. Suppose a new input sequence $\underline{x}^*$ were applied to the system, one would wish to determine the density for the predicted outputs,

$$\underline{y}^* \sim p(\underline{y}^*|\underline{x}^*, \underline{w}, D, \mathcal{M}) \tag{6.10}$$

noting all the dependencies. However, the true Bayesian viewpoint on the prediction, would require one to marginalise over the parameter estimates, i.e. to derive,

$$p(\underline{y}^*|\underline{x}^*, D, \mathcal{M}) = \int p(\underline{y}^*|\underline{x}^*, \underline{w})p(\underline{w}|D, \mathcal{M})d\underline{w} \tag{6.11}$$

This is a very powerful idea: allowing for a fixed model *structure*, *one is making predictions using an entire set of parameters consistent with the training data*, with each point in the space of parameters weighted according to its likelihood given the data. In practice, there are considerable problems in implementing the full Bayesian approach. For the purposes of this paper, the main issue for discussion will be the problem of inferring the distributions for the parameters as given in (6.9).

Another potential advantage of a Bayesian approach is that it may be possible to assess the relative evidence for a number of competing model structures. This issue is discussed in more detail in [6]. One simple approach to model selection, which will be applied later is the *Deviance Information Criterion* (DIC), which is a Bayesian generalisation of the Akaike Information Criterion (AIC) [12]. This is specified in terms of the *deviance* $D(\underline{w})$, defined by,

$$D(\underline{w}) = -2\log p(D|\underline{w}) + C \tag{6.12}$$

where $p(D|\underline{w})$ is simply the likelihood of the data and C is a constant. Because the DIC will only ever be used in comparing models, the C will cancel out when a difference is taken and is therefore unimportant. The expectation of the deviation,

$$\overline{D} = E_{\underline{w}}[D(\underline{w})] \tag{6.13}$$

is a measure of how well a given model *structure* fits the data, as the parameters themselves have been marginalised. Now, the *effective number of parameters* is defined as,

$$p_D = \overline{D} - D(\overline{\underline{w}}) \tag{6.14}$$

where $\overline{\underline{w}} = E[\underline{w}]$. A little thought is enough to show that this function will be an increasing function of the number of model parameters. Finally, the DIC is defined by,

$$DIC = p_D + \overline{D} \tag{6.15}$$

The DIC is higher if more parameters are used because of the first term and lower if the model explains the data better. The first term therefore balances the second as models with more parameters will naturally fit data better. Like the AIC, the DIC penalises model complexity and rewards model fidelity.

6.3.2 Markov Chain Monte Carlo

As stated earlier, the central problem for this paper will be to determine the probability density function for parameter estimates, given data from a Bouc-Wen hysteretic system. Unfortunately, in terms of exact statistical inference this problem is already too difficult. It is not possible to analytically determine the form of the required distribution and it will be necessary to adopt numerical methods. If one were to assume a specific parametric form for the density e.g. Gaussian, one could make headway; however, this involves imposing a restrictive, and quite possibly incorrect, assumption. One ideally wishes to adopt a nonparametric approach to the problem so that the true statistics of the parameter density are obtained. An obvious step in the direction of nonparametric estimation is to sample from the parameter distribution itself and thus build a picture; however, the distribution in question is unknown. Fortunately, there exists a very powerful class of numerical methods which allow sampling from such densities—the class of *Markov Chain Monte Carlo (MCMC)* methods [13].

While many sampling methods are available for Monte Carlo estimation, one of the most versatile and powerful is MCMC; the variant used in this paper will be the *Metropolis–Hastings* (MH) method. In order to describe the method, it is sufficient to assume that one has a single random variable x with true, but unknown, density $p(x)$. The MH method works even if $p(x)$ is too complicated to sample from directly. First of all, one assumes that one can at least *evaluatep*(x), given a candidate value for x; in fact, one only needs to evaluate a density $p\ ^*\ (x)$, where $p\ ^*\ (x) = Z_p p(x)$ i.e. the density of interest up to a multiplicative constant. The method hinges on the use of a *proposal distributionq*(x), which is simpler than $p(x)$ and can also be evaluated up a multiplicative constant. The use of the proposal distribution is common to many methods like importance and rejection sampling [13]. The unique feature of MCMC methods is that a sequence of samples or states x^i are generated, with the proposal distribution a function of the state at any given time; one estimates the probability for the next state in the sequence (a Markov chain—hence the name) by conditioning on the current state, the proposal distribution is $q(x' \quad x^i)$. Another way of regarding q is as a transition probability density for a jump in the state/sample space. Often q is taken as a simple distribution, i.e. a Gaussian. Note that if the proposal distribution has a small variance, this will lead to small jumps in the sample space and will produce (at least locally) correlated states/samples.

The MH algorithm proceeds as follows. Assume one is at iteration i in the process:

1. Sample from the proposal density $q(x' \quad x^i)$ to generate a candidate state x'.
2. Evaluate the quantity,

$$a = \frac{p_*(x')q(x^i|x')}{p_*(x^i)q(x'|x^i)}$$

3. If $a \geq 1$, the new state is accepted, *otherwise* the new state is accepted with probability a.
4. If the new state is accepted, set $x^{i+1} = x'$, else set $x^{i+1} = x^i$.

A full description of why this procedure leads to sampling from $p(x)$ is beyond the scope of this paper, the curious reader should consult [13].

It is possible to show that for any q such that $q(x' \quad x) > 0$ for all x, x', the probability density of the x^i will tend to $p(x)$ as $i \rightarrow \infty$ [13]. Two issues arise with the iteration above. The first concerns initial conditions. As in any iterative process, initial conditions matter, depending on these there will be a transient period before the sequence becomes stationary and is consistently generating samples from $p(x)$. In order to allow for this, the algorithm is usually run for a *burn-in* period before samples are drawn. The second issue is concerned with independence, as observed above, the proposal density may only allow small jumps in the state/sample space, so states close together in the sequence will be correlated. To obtain independent samples, one only saves every n_tth state; this process is called *thinning*. Finally, another issue becomes important when one uses the algorithm in higher dimensions, one should choose the proposal distribution in such a way that the algorithm explores the posterior space quickly without exhibiting too much *random walk* behaviour. This is the problem of *mixing* There are no hard-and-fast rules for choosing a suitable proposal distribution, though [12] offers some helpful guidance.

6.3.3 MCMC for System Identification

In order to implement MCMC, it is necessary to evaluate a density $p^*(\underline{w}|D)$ which is the required $p(\underline{w}|D)$ up to a constant multiple (forgetting the noise variance for the moment). Using Bayes theorem, one has,

$$p(\underline{w}|D) = \frac{p(D|\underline{w})p(\underline{w})}{p(D)} \tag{6.16}$$

and as $p(D)$ is fixed by the data (and cannot be evaluated anyway), one has,

$$p^*(\underline{w}|D) = p(D|\underline{w})p(\underline{w}) \tag{6.17}$$

as the required density (p, up to a constant). This is the product of a likelihood for the data (which can be computed under the assumption that the noise is Gaussian—the log likelihood is then related to the least-squared error between predictions using the candidate parameters and the measured data) and a prior for the parameters. The only thing needed then for the MH algorithm is the proposal distribution $q(\underline{x}'|\underline{x}^i)$ and a simple prescription for this can be a spherical Gaussian distribution,

$$q(\underline{x}^{'}|\underline{x}^{i}) \sim N(\underline{x}^{i}, \sigma_p I) \tag{6.18}$$

controlled by a single scale parameter σ_p (I is the identity matrix).

One feature of the MCMC algorithm is that it simplifies the computation of the DIC discussed earlier; one simply calculates the average of $D(\underline{w})$ over the sequence of parameter samples and evaluates D at the mean of those samples $\overline{\underline{w}}$.

6.4 Results

The training data for the current study were generated by simulation. The coupled equations (6.1) and (6.3) were integrated forward in time using Matlab [14] using a fixed-step fourth-order Runge–Kutta scheme for initial value problems [15]. The parameters for the baseline system adopted here were: $m = 1$, $c = 20$, $\alpha = 1.\ 5$, $\beta = -1.5$, $A = 6680.\ 0$ and $n = 2$. The excitation was a Gaussian random sequence with mean zero and standard deviation 10.0. The step-size (or sampling interval) was taken as 0.004 s, corresponding to a sampling frequency of 250 Hz. Noise of RMS 1% of the response was added to the velocity response data which were used, with the corresponding samples of excitation force, as ‘training data’ for the algorithm. The training set, or identification set, used here was composed of 500 points corresponding to a record duration of 2 s. The training set, force and displacement is illustrated in Fig. 6.1.

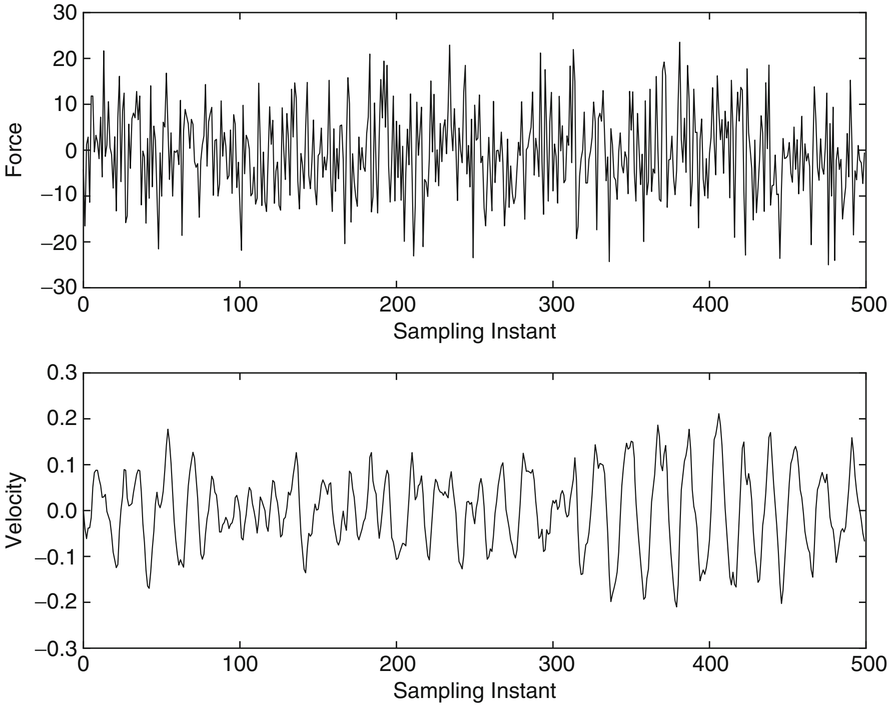


Fig. 6.1 Training data for the continuous-time Bouc-Wen system

Once the data had been generated, the MCMC algorithm was applied to the identification problem. The parameter n was not included for direct estimation; the reason being that different values of n here will be treated as generating different models from a given class and model selection will be carried out by using the DIC as discussed in the last section.

The MCMC identification routine was coded in the Python language using the package PyMC [16]. A number of identification runs were made using values of the Bouc-Wen n parameter from 1 to 4 in order to later investigate the possibility of model selection using the DIC as discussed earlier. As well as the seven parameters: a_1, a_2, b_1, b_2, α, β and γ_2, as discussed earlier, the MCMC routine also attempted to learn the output noise variance; however, for convenience this was encoded as a precision parameter $B = 1/\sigma^2$. Because of the differences in scale between the parameters, log values were used in the algorithm to improve mixing as discussed earlier. All of the model parameters were given uniform priors with the bounds set one order of magnitude below and above the 'true' parameters. As the data are from a continuous-time Bouc-Wen system yet the model fitted is discrete-time, the true parameters are not known, so in this case the parameters were intialised with estimates obtained from the data using the Self-Adaptive Differential Evolution (SADE) algorithm as described in [1]. The initial parameter values were thus set as $a_1 = 1.819, a_2 = -0.92, b_1 = 0.00344, b_2 = -0.00361, \alpha = -0.0000104$, $\beta = -0.0000222, \gamma_2 = 13.6$. The proposal distribution q was essentially set as a spherical Gaussian with variance of 0.1 in each direction. The MCMC algorithm was set to run for 25,000 samples, thinning by taking every tenth sample and a burn-in of 5,000 samples was used. The results from the MCMC run are shown in Fig. 6.2. Each of the subfigures shows the sequence of MC samples for the parameter of interest, followed by a frequency histogram giving a coarse view of the parameter density.

The first observation one can make from Fig. 6.2 is that the chains for the parameters are all largely stationary immediately following the burn-in period, although some are still showing a little drift (and these drifts are clearly correlated between the parameters). When the means and standard deviations for the parameters were estimated over the samples in each chain, the results were as summarised in Table 6.1. As mentioned above the only comparison one can make is with the results obtained when the data were identified using the SADE algorithm.

The results are excellent. The 3σ bounds (99.7% confidence intervals if the parameter densities are Gaussian) in almost every case, encompass the true values of the parameters estimated using the SADE algorithm. The one exception is the γ_2 parameter which is still not subject to significant bias. Also, in contrast to an identification scheme where the parameter distributions are *assumed* Gaussian, the parameter samples are available here so that a test can be carried out to check. Using the last 50 samples from the chains, the predicted responses were computed and are shown superimposed on the true response in Fig. 6.3.

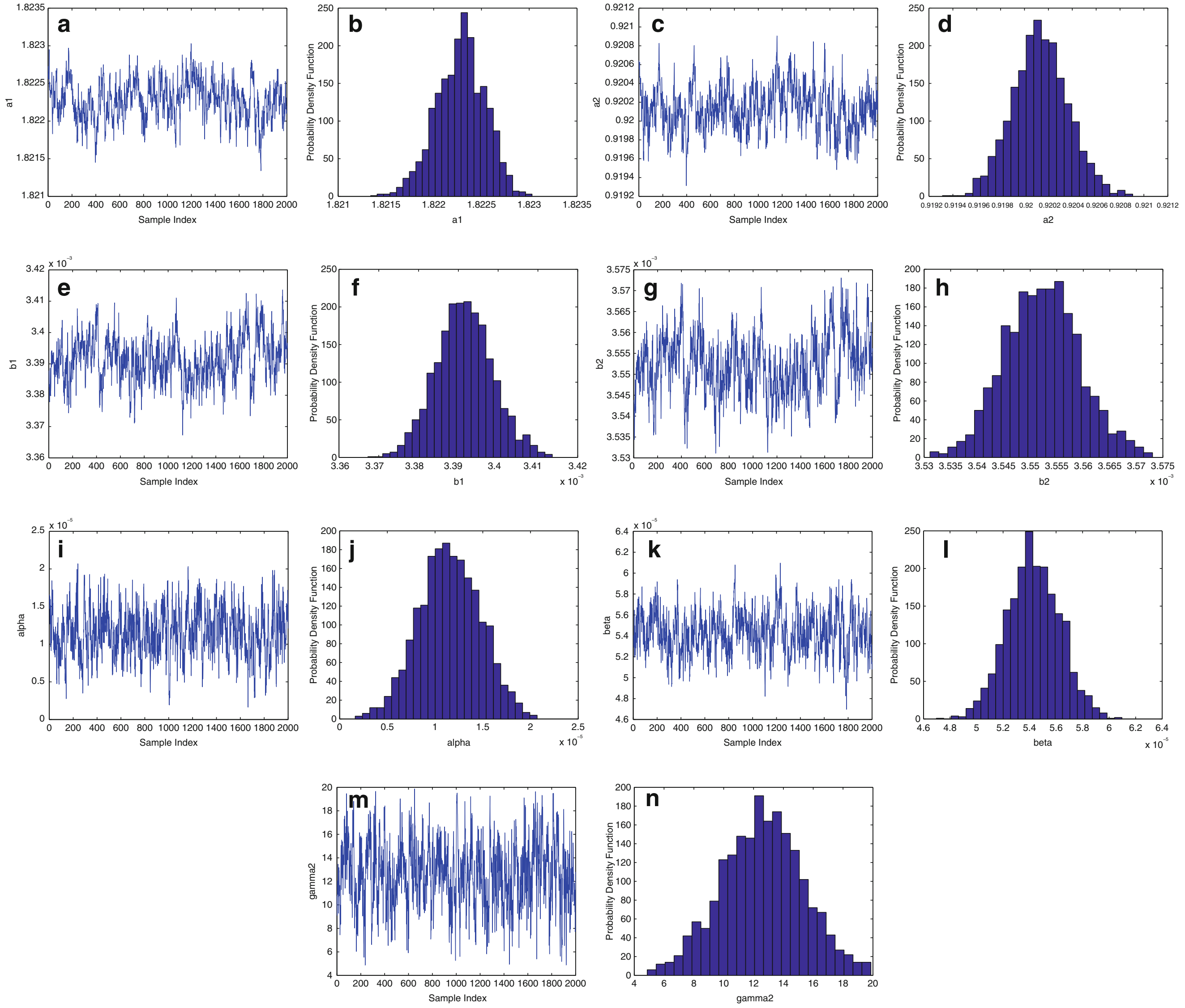


Fig. 6.2 Results from MCMC identification of NARX Bouc-Wen system with $n = 2$. (**a**) a_1: Samples, (**b**) a_1: PDF, (**c**) a_2: Samples, (**d**) a_2: PDF, (**e**) b_1: Samples, (**f**) b_1: PDF, (**g**) b_2: Samples, (**h**) b_2: PDF, (**i**) α: Samples, (**j**) α: PDF, (**k**) β: Samples, (**l**) β: PDF, (**m**) γ_2: Samples, and (**n**) γ_2: PDF

Table 6.1 Parameter estimates for NARX Bouc-Wen model with $n = 2$ (correct n)

Parameter	SADE value	MC mean	MC Std
a_1	1.83	1.82228	0.00025
a_2	-0.92	-0.92012	0.00024
b_1	0.00344	0.0033915	0.0000071
c	-0.00361	-0.0035522	0.0000071
α	-1.04×10^{-5}	-1.152×10^{-5}	0.321×10^{-5}
β	-2.22×10^{-5}	-5.430×10^{-5}	0.199×10^{-5}
γ_2	13.6	12.594	2.740

It is clear that the predictions have very little variation between them, consistent with the low standard deviations of the parameters. This is, to an extent, a consequence of the way the predictions have been made and is on the optimistic side. By integrating the differential equations across the entire time with a single realisation of the parameters, the predictions ignore the component of uncertainty associated with *state estimation*. This effect is discussed in some detail in [6].

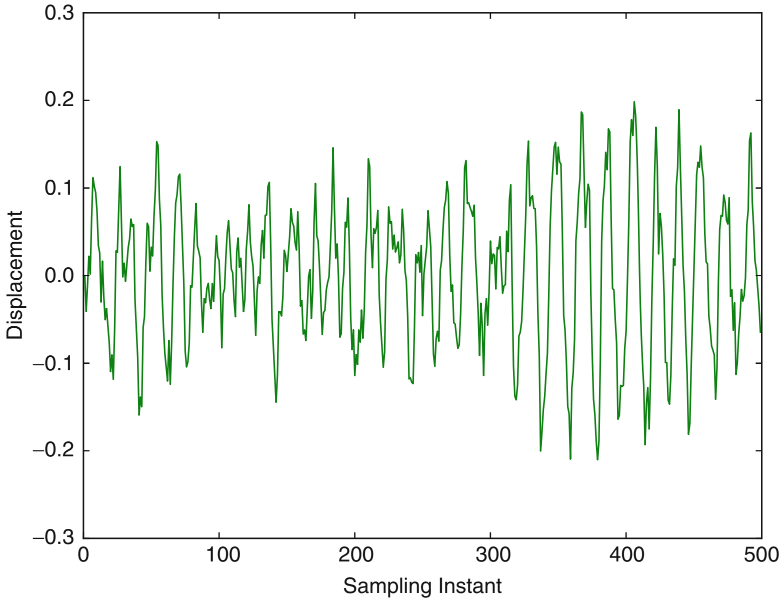


Fig. 6.3 Predictions using parameters from 50 MCMC parameter samples, superimposed on observed data

Table 6.2 Parameter estimates for NARX Bouc-Wen model with different n values

n	a_1	a_2	b_1	b_2	α	β	γ_2	DIC	DIC *
Linear	1.817	− 0.921	0.00341	− 0.00358				21301.6	0
1	1.819	− 0.921	0.00342	− 0.00358	-9.595×10^{-5}	-1.978×10^{-4}	45.42	20,505	− 796
2	1.822	− 0.920	0.00339	− 0.00355	-1.152×10^{-5}	-5.430×10^{-5}	12.59	20,216	− 1085
3	1.821	− 0.921	0.00339	− 0.00355	-1.012×10^{-6}	-2.486×10^{-6}	8.908	20,415	− 887
4	1.701	− 0.765	0.00346	− 0.00325	-1.000×10^{-6}	-2.000×10^{-6}	1.351	24,7757	22,6455

As discussed earlier, one can also use the Bayesian approach to guide model selection. As the DIC is readily available from MCMC calculation and is actually computable within the PyMC software, it will be considered here. NARX Bouc-Wen models with n equal to 1–4 were fitted, and also a linear ARX model of the form,

$$v_i = \sum\nolimits_{j=1}^{2} a_j v_{i-j} + \sum\nolimits_{j=1}^{2} b_j x_{i-j} \tag{6.19}$$

The resulting (mean) parameters and corresponding DICs are given in Table 6.2. In all cases, the analysis produced small confidence intervals for the parameters and the predictions were equally good apart from the $n = 4$ model which is also the only one showing significant deviations in the linear parameters from those extracted by the SADE algorithm. If one defines a normalised mean-square-error by,

$$MSE(y^*) = \frac{100}{N\sigma_y^2} \sum\nolimits_{i=1}^{N} (y_i - y_i^*)^2 \tag{6.20}$$

where the y_i are the observed outputs and the y_i * are the predictions, the first three Bouc-Wen models gave MSE values of about 5.0, while the $n = 4$ model gave 50.6 and the linear model gave a value of 5.2. This is understandable, the difference between the Bouc-Wen models and the linear model shows that the hysteretic effects are small but observable. Given the extra parameters in the Bouc-Wen models, they are all able to fit the data better (with the exception of the $n = 4$ model, somewhat surprisingly). However, the DIC shows that the $n = 2$ model is clearly preferred. The results from the DIC can be shown in a slightly clearer manner by setting a reference value for the linear model at 0 and stating the differences as DIC * .

6.5 Conclusions

It has been shown here that a Bayesian MCMC system identification scheme provides an effective means of estimating parameters for NARX Bouc-Wen hysteretic systems. The scheme provides natural confidence intervals for the parameter estimates. A further advantage of the Bayesian approach is that it provides access to various means of model selection. The DIC approach shown here is not one of the most discriminating; however, it was shown to select the correct model structure here from a set of competing Bouc-Wen models.

References

1. Worden K, Barthorpe RJ (2012) Identification of hysteretic systems using NARX models, Part I: evolutionary identification. In: This proceedings
2. Bouc R (1967) Forced vibration of mechanical system with hysteresis. In: Proceedings of 4th conference on nonlinear oscillation, Prague
3. Wen Y (1976) Method for random vibration of hysteretic systems. ASCE J Eng Mech Div 102:249–263
4. Leontaritis IJ, Billings SA (1985) Input–output parametric models for nonlinear systems, part I: deterministic nonlinear systems. Int J Control 41:303–328
5. Leontaritis IJ, Billings SA (1985) Input–output parametric models for nonlinear systems, part II: stochastic nonlinear systems. Int J Control 41:329–344
6. Worden K, Hensman JJ (2011) Submitted to Mechanical System and Signal Processing. Bayesian inference and nonlinear system identification
7. Ljung L (1987) System identification: theory for the user. Prentice Hall, Englewood Cliffs
8. Söderstrom T, Stoica P (1988) System identification. Prentice-Hall, London
9. Worden K, Tomlinson GR (2001) Nonlinearity in structural dynamics: detection, modelling and identification. Institute of Physics, Bristol/Philadelphia
10. Girolami M (2008) Bayesian inference for differential equations. Theor Comput Sci 408:4–16
11. Calderhead B, Girolami M, Higham DJ (2010) Is it safe to go out yet? Statistical inference in a zombie outbreak model. Preprint, University of Strathclyde, Department of Mathematics and Statistics
12. Gelman A, Carlin JB, Stern HS, Rubin DB (2004) Bayesian data analysis, 2nd edn. Chapman and Hall/CRC, Boca Raton
13. Mackay MJC (2003) Information theory, inference and learning algorithms. Cambridge University Press, Cambridge/New York
14. The Mathworks Inc., (2004) MATLAB version 7 Natick, Massachusetts
15. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical recipes: the art of scientific computing, 3rd edn. Cambridge University Press, Cambridge/New York
16. Patil A, Huard D, Fonnesbeck CJ (2010) PyMC: Bayesian stochastic modelling in Python. J Stat Softw 35:81

Chapter 7
Bayesian Methods for Uncertainty Quantification in Multi-level Systems

Shankar Sankararaman, Kyle McLemore, and Sankaran Mahadevan

Abstract This paper develops a Bayesian methodology for uncertainty quantification and test resource allocation in multi-level systems. The various component, subsystem, and system-level models, the corresponding parameters, and the model errors are connected efficiently using a Bayes network. This provides a unified framework for uncertainty analysis where test data can be integrated along with computational models and simulations. The Bayes network is useful in two ways: (1) in a forward problem where the various sources of uncertainty are propagated through multiple levels of modeling to predict the overall uncertainty in the system response; and (2) in an inverse problem where the model parameters of multiple subsystems are calibrated simultaneously using test data. The calibration procedure leads to a decrease in the variance of the model parameters, and hence, in the overall system performance prediction. Then the Bayes network is used for test resource allocation where an optimization-based procedure is used to identify tests that can effectively reduce the uncertainty in the system model prediction are identified. The proposed methods are illustrated using two numerical examples: a multi-level structural dynamics problem and a multi-disciplinary thermally induced vibration problem.

Keywords Bayes network • Model calibration • Resource allocation • Sensitivity analysis • Testing

7.1 Introduction

Engineering systems are often composed of multiple levels of models. These levels may correspond to subsystems and/or components of the overall system. The subsystem/component may represent isolated features, or isolated physics, or simplified geometry of the original system. Typically, along the hierarchy of a multi-level system, the complexity of the governing physics increases, and hence, the complexity of the model increases, the cost of testing increases, and hence, the amount of available experimental data decreases. At the system-level, full-scale testing may not even be possible to predict the system performance under actual operating conditions. It is essential to quantify the overall uncertainty in the system-level prediction using the models and data available at all levels. The field of "quantification of margins and uncertainties" (QMU) has the goal of enabling this overall capability [1].

The first challenge in this regard is to efficiently connect the model inputs, parameters, outputs, model errors, and available data across all the hierarchy. This is not straightforward because there are several sources of uncertainty – physical variability, data uncertainty, and model uncertainty – at each level of the multi-level system. The Bayes network serves as an efficient and powerful tool to integrate multiple levels of models, associated sources of uncertainty, and experimental data at multiple levels. It can used for uncertainty propagation (forward problem), where the system-level prediction uncertainty is quantified by propagating all the sources of uncertainty at lower levels through the Bayes network. The Bayes network is also useful for model parameter calibration (inverse problem), where the data available at all levels can be simultaneously used to calibrate the underlying parameters at different levels of models. In addition, it can also be used to quantitatively judge the predictive capability of a model, which is the focus of model validation; the issue of model validation is not addressed in this paper.

S. Sankararaman • K. McLemore • S. Mahadevan (✉)
Vanderbilt University, Nashville, TN 37235, USA
e-mail: sankaran.mahadevan@vanderbilt.edu

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_7,

The above capabilities of the Bayes network are then exploited to guide in test resource allocation. Test data available at multiple levels of system hierarchy are used for model parameter calibration, which in turn leads to a reduction of uncertainty in the model parameters; this reduced uncertainty is represented through the posterior distributions of the model parameters. When these posterior distributions are propagated through the Bayes network, the uncertainty in the system-level response also decreases. Thus, testing can be used to evolve the system performance prediction and a common concern is to select that test design which leads to maximum reduction in the uncertainty (usually expressed through variance) of the system performance prediction. The tests need to be selected and designed with adequate precision (measurement error and resolution), and the simulations need to be developed with adequate resolution (model fidelity) to achieve the project requirements. This can be performed by embedding the Bayes network within an optimization algorithm where the decision variables correspond to the test data. This formulation is very interesting because, in the past [2], model parameters have been calibrated with available test data; the difference now is that it is required to perform Bayesian calibration and assess the reduction in uncertainty in the system-level response even before actual testing is performed.

Two types of questions are answered in this paper: (1) what tests to do? and (2) how many tests to do? Tests at different levels of the system hierarchy have different costs and variance reduction effects. Hence, the test selection is not trivial and it is necessary to identify an analytical procedure that helps in the optimum test resource allocation. However, current practice for this is, at best, based on simplified analysis and relevant experience, and at worst based on ad hoc rules, any of which may or may not result in truly conservative estimates of the margin and uncertainty. For multi-level systems, a rational test selection procedure should also incorporate the information from component-level and subsystem-level tests towards overall system level performance prediction.

Section 7.2 describes the Bayes network-based methodology for integrating multiple models and the different sources of uncertainty. The Bayes network can be used in two ways: (1) in a forward problem, where the different sources of uncertainty are propagated through multiple levels using multiple models to compute the uncertainty in the overall system performance prediction, and (2) in an inverse problem, where test data at a higher level can be used to calibrate the model parameters at lower levels. Section 7.3 describes the proposed approach for test resource allocation using optimization. Sections 7.4 and 7.5 implement the proposed methodology in two illustrative examples – a multi-level structural dynamics problem (Sect. 7.4) and a multi-physics thermal vibration problem (Sect. 7.5).

7.2 Bayes Network for Uncertainty Quantification

A Bayes network [3] is a graphical representation of relationship between uncertain quantities. Each uncertain quantity is represented as a node and successive links are connected to each other using unidirectional arrows that express dependence based on conditional probabilities. In a Bayes network, an ellipse corresponds to a source of uncertainty and a rectangle represents observed data. A solid line arrow represents a conditional probability link, and a dashed line arrow represents the link of a variable to its observed data if available.

Figure 7.1 shows a conceptual Bayes network that aids in uncertainty quantification across multiple levels of models. In this Bayes network, $\boldsymbol{\Phi}$ and $\boldsymbol{\Omega}$ are model parameters, and z and y are model outputs. Experimental data $\boldsymbol{Z}$ and $\boldsymbol{Y}$ are available for comparison with the respective model predictions z and y. Let X denote a higher level system response quantity that depends on the lower level model predictions z and y.

The Bayes network shown in Fig. 7.1 can be used in two different problems – a forward problem [4], and an inverse problem [5]. In the forward problem, the probability distributions of the parameters $\boldsymbol{\Phi}$ and $\boldsymbol{\Omega}$ are known and these can be used to calculate the probability distribution of z, y, and hence X, using methods of uncertainty propagation such as Monte Carlo

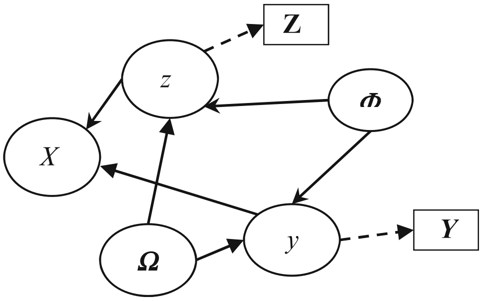


Fig. 7.1 Bayes network

sampling (MCS), first-order reliability method (FORM), second-order reliability method (SORM), etc. The experimental data available, i.e. Z and Y are not used. The distribution of the higher level output quantity X can be calculated as:

$$f_X(x) = f_X(x|y,z)f(y,z|\Phi,\Omega)f(\Phi)f(\Omega)d\Phi d\Omega. \tag{7.1}$$

In the inverse problem, the probability densities of the parameters (***Ω*** and ***Φ***) can be updated using the data ***Y*** and ***Z*** as:

$$f(,\Phi|Y,Z) \propto L(,\Phi)f(,\Phi). \tag{7.2}$$

In (7.2), the first term on the right hand side represents the likelihood function and the second term refers to the prior distribution of the parameters (***Ω*** and ***Φ***).

The updated statistics of these parameters (***Ω*** and ***Φ***) can then be used to estimate the updated statistics of the system response quantity X. The Bayes network thus links the various models and corresponding experimental observations in multiple levels to accomplish uncertainty quantification and propagation through multiple levels of modeling, thereby aiding in the uncertainty quantification of the overall system model prediction.

7.3 Test Resource Allocation

Using the Bayes network, the model parameters are updated using test data and this updating procedure is likely to lead to a decrease in the variance of the calibration quantity. When the updated distribution of the calibration quantity is propagated to calculate the overall system performance prediction, the variance of the overall system performance prediction is also likely to decrease with more data.

The fact that data can be collected at multiple levels translates to the fact that data can be collected across several nodes in a Bayes network. As each source of uncertainty corresponds to a node in the Bayes network, the ranking of importance of different source of uncertainty corresponds to the ranking of nodes in the Bayes network. Hence, it would be ideal to collect experimental data related to those nodes that contribute significantly to the overall uncertainty in the system prediction.

Suppose that there are multiple nodes related to which data may be collected. Let the number of data points to be collected in the *ith* node be equal to N_i and the cost of each corresponding experiment be C_i. Let D denote this entire set of data. The test resource allocation problem is to find optimal values of N_i, given the cost values C_i. The objective is be to minimize the variance of the system-level prediction as a result of the above collected data.

If the data D were available, then it is possible to include this available data in the Bayes network, and recalculate the probability distribution of the system response quantity. Let $\mathit{\Delta}(D)$ denote the reduction in variance as a result of the updating using this data D. Note that this reduction in variance is a function of the data set D. However, this quantity $\mathit{\Delta}$ cannot be calculated because $\mathit{\Delta}$ depends on the test data set D which is not available before testing.

This difficulty is overcome by calculating an estimate of $\mathit{\Delta}$ by generating "imaginary" bootstrap samples for the test data set (D) and calculating the probability distribution of $\mathit{\Delta}$ as a function of N_i, i.e. $f_{\mathit{\Delta}}(\mathit{\Delta}/N_i)$. Note that the variability in $\mathit{\Delta}$ is due to generating multiple samples of D. The optimal choice of testing can be selected by solving the optimization problem explained in Fig. 7.2.

Note that this optimization is a complicated integer optimization problem, where the Bayesian updating procedure has to be repeated multiple times for every evaluation of the objective function. Two different approaches have been used in this paper to solve this optimization. First, a grid search algorithm which computes the expected value of $\mathit{\Delta}$ for every feasible combination of N_i's is used. Then, an adaptive test selection algorithm which is based on a what-to-test-next approach is used. In this adaptive method, the test that immediately leads to the maximum reduction in variance in chosen at every step and this procedure is repeated until the budget is exhausted or the target variance reduction in reached. The proposed methods are illustrated through two numerical examples in the following two sections.

Minimize Expected Value of $\mathit{\Delta}$

$\Sigma\ (C_i\ N_i) <$ Total Budget

Given C_i where i denotes multiple possible tests
(component level, subsystem level, system level, etc.)

Fig. 7.2 Formulating the test resource allocation problem

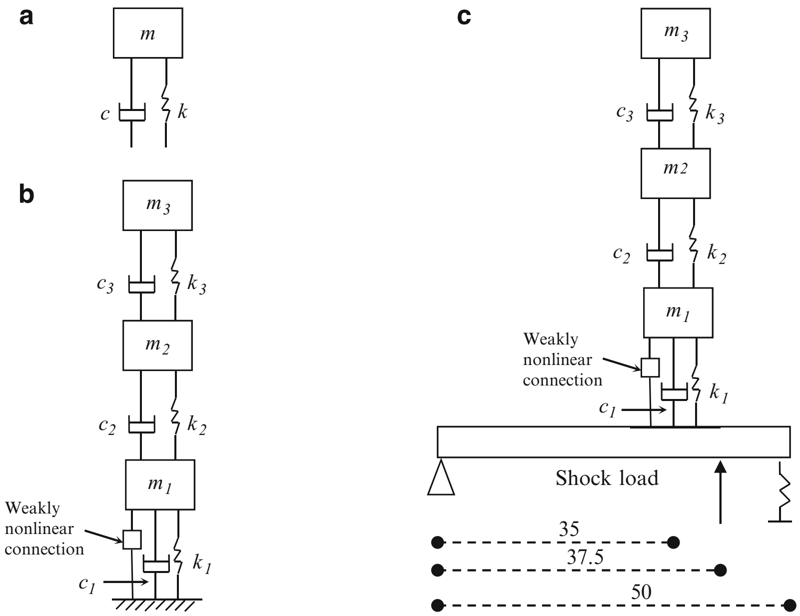


Fig. 7.3 Structural dynamics problem (**a**) Level 1: spring mass damper (**b**) Level 2: three spring mass dampers (**c**) Level 3: three spring mass dampers on a beam

7.4 Numerical Example I – Structural Dynamics

This section illustrates the proposed methods through the structural dynamics validation challenge problem proposed at the Sandia National Laboratories [7]. Note that the focus of this paper is not on validation and only the definition of the structural dynamics system is used in this paper. This system consists of three levels. The first level consists of a simple spring mass damper. The second level integrates three spring mass dampers from the previous level. The third level integrates the previous level with a beam. These three levels are shown in the Fig. 7.3.

The solutions to the subsystems in the first two levels are straightforward [9] and the solution to the third level can be obtained using the Belytschko method. Red-Horse and Paez [7] describe in detail the modeling and simulation of this system.

In this paper, the system response quantity (R) whose variance needs to be minimized is the acceleration in the first mass of the spring mass damper system. Let x_{11}, x_{12}, x_{13} denote the response quantities for the three different masses in the 1st level; x_2 and x_3 denote the response in the 2nd and 3rd level respectively; let ε be the model uncertainty in each level. Then the Bayes network for this problem is shown in Fig. 7.4.

In this numerical example, the stiffness values of the three masses, i.e. k_1, k_2, and k_3 are identified to be the calibration parameters. Prior distributions are assumed for these quantities and later updated with testing data. The numerical values of all the quantities are summarized in Table 7.1.

The mass of the beam is taken to be 0.1295. The excitation applied on all the levels is a sine function force with an amplitude of 10,000 with a frequency of $\omega = 100$There are five possible test cases. The costs of testing, for the sake of illustration, are assumed to be $100, $500 and $1,000 in the 1st, 2nd, and 3rd level respectively. For the purpose of test resource allocation, future test data is generated by drawing samples out of a uniform distribution, based on expert opinion.

The grid search algorithm yielded results of 20 1st level tests (1 for the 1st mass, 14 for the 2nd mass, and 5 for the 3rd mass). While the computational difficultly involved in executing the grid search algorithm increases with the number of possible test cases (5, in this numerical example), alternative solution approaches like the adaptive search method must also be investigated. While the grid search algorithm leads to a global minimum, the adaptive method might not. In this case, the adaptive search method led to an optimum of 15 tests at the 1st level (4 for m_1, 7 for m_2, and 4 for m_3) and one test at the 2nd level, thus exhausting the the $2,000 budget. However, the resultant variance using the adaptive test selection procedure was higher than the global optimum variance reduction attained through the grid search algorithm, suggesting that the adaptive

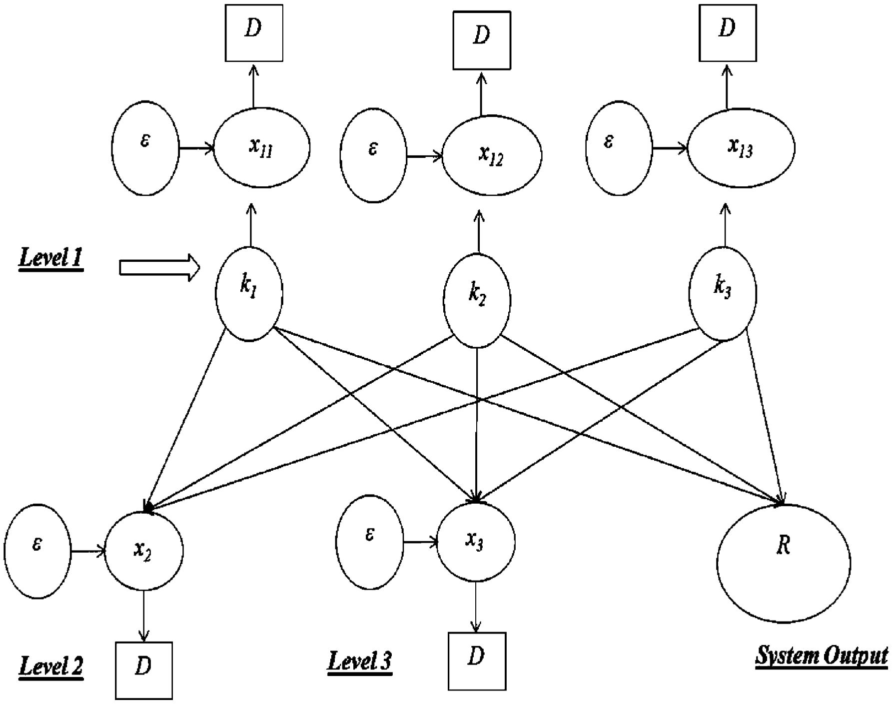


Fig. 7.4 Bayes network: structural dynamics problem

Table 7.1 Numerical values

			Priors (stiffness)	
No.	Mass (m)	Damping (c)	μ_k	CoV of k
1	0.012529	0.023466	5,600	0.260187
2	0.019304	0.021204	11,000	0.280557
3	0.035176	0.031216	93,000	0.279296

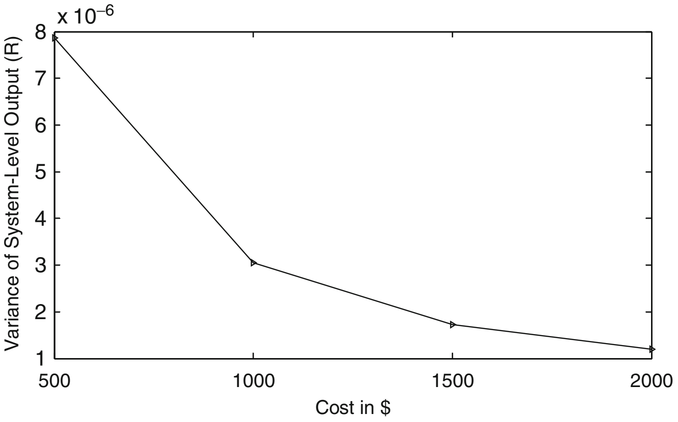


Fig. 7.5 Decrease of variance with increasing cost

test selection procedure might lead to a local minimum. The reduction of variance with increasing cost (for the adaptive test selection algorithm) is shown in Fig. 7.5.

It is seen that the reduction in variance using the last $500 (i.e. from $1,500 to $2,000) was insignificant compared to the reduction in variance using the initial $1,500. Hence, from an economics-perspective, it is advantageous to use only $1,500 for testing as the extra $500 does not lead to further significant reduction in variance.

7.5 Numerical Example II – Thermally Induced Vibration

This coupled-physics example illustrates a laboratory experiment to simulate the behavior of spacecraft booms. In this experiment, a thin walled circular tube is rigidly attached at its top and supports a concentrated mass at its bottom. The tube and the mass are initially at rest and a constant heat flux is applied along the length of the tube. The application of the heat flux causes an increase in the temperature on the incident surface while the unheated side remains at the initial temperature. The temperature gradient causes the beam to bend rapidly away from the lamp and triggers continuous vibration. The set up of this experiment is shown in Fig. 7.6.

The temperature at the tip mass T_m is given by the following differential equation:

$$\frac{\partial T_m}{\partial t} + \frac{T_m}{\tau} = \frac{T^*}{\tau}\left(1 - \frac{v(x,t)}{\beta^*}\right). \tag{7.3}$$

In (7.3), *v(x, t)* represents the displacement of the beam as a function of length and time. The other parameters T*, τ, β* are constants and Thornton [8] explains how to calculate them.

The displacement *v(x, t)* can be related to the displacement of the tip mass *V(t)* as:

$$v(x,t) = \left(\frac{3x^2}{2l^2} - \frac{x^3}{2l^3}\right)V(t). \tag{7.4}$$

The tip mass displacement, in turn, depends on the forcing function as follows:

$$\ddot{V} + (2\xi\omega_0)\dot{V} + \left(\omega_0^2 + \frac{6g}{5l}\right)V = \frac{F(t)}{m}. \tag{7.5}$$

In (7.5), ξ is the damping ratio, and ω_0 is the frequency. The forcing function *F(t)* depends on the thermal moment which in turn depends on the temperature, thereby casing coupling between the thermal equation and the structural equation. These relations are shown in (7.6) and (7.7) respectively.

$$F(t) = -\frac{3}{l^3}\,{}^{lx}_{00} M(u,t)dudx. \tag{7.6}$$

$$M(x,t) = \int E\alpha T_m(x,t)\cos(\Phi)ydA \tag{7.7}$$

Equation 7.7, *E* is the elastic modulus, α is the coefficient of thermal expansion, *Φ* is the angle of incident flux on the cross section, *y* is the distance from the center of the cross section and the integral is over the area of the cross section *A*. Refer Thornton [8] for a detailed description of this problem. The Bayes network corresponding to this problem is shown in Fig. 7.7.

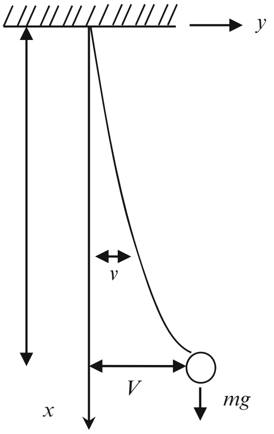


Fig. 7.6 Thermally induced vibration [8]

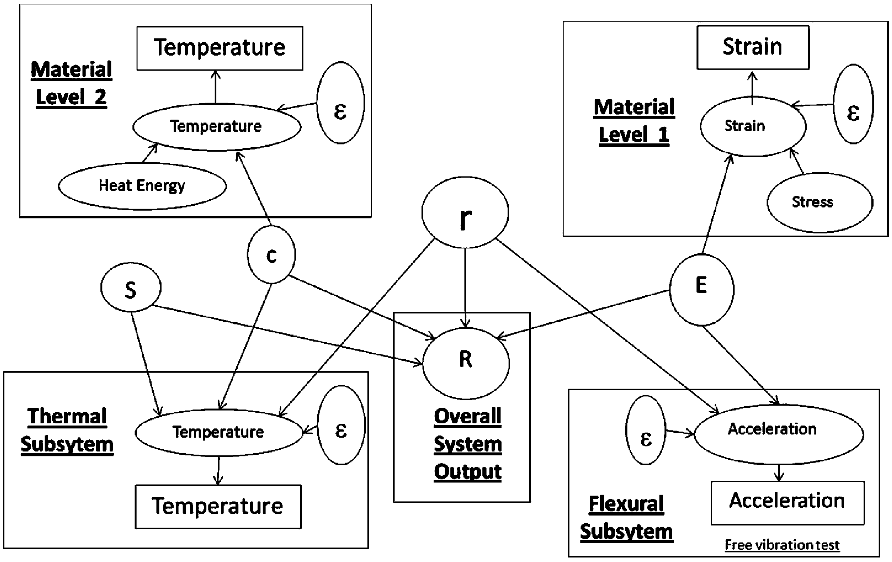


Fig. 7.7 Bayes network: thermally induced vibration

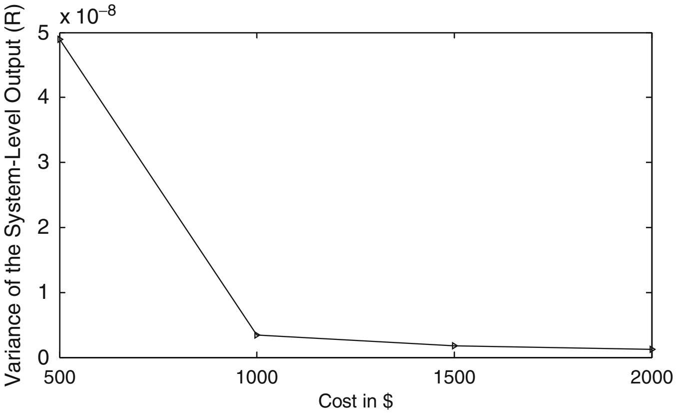


Fig. 7.8 Decrease of variance with increasing cost

In this numerical example, three quantities, the elastic modulus (*E*), the damping constant (c), and the radius of the mass (r) are chosen as calibration quantities. The other quantities are assumed to be known. Their numerical values are identical to those given by Thornton [8]. The priors for the calibration quantities are chosen to be normal, with means based on Thornton [8], and 10% coefficient of variation. The overall system output (*R*) is chosen to be ratio of the amplitudes at two different time instants for the coupled system which is subjected to thermally induced vibrations. If $R < 1$, then it implies that the system is stable with oscillations diminishing as a function of time. If $R > 1$, then it implies that the system is unstable, commonly referred to as flutter, an undesirable scenario.

There are four possible test cases – two material level tests (stress–strain test, and temperature test), a thermal subsystem test (without structural effects), and a flexural subsystem test (without thermal interaction). It is assumed that there is no testing possible for the coupled thermally induced vibration system. For the purpose of illustration, each material level test is assumed to cost $100 while each subsystem level is assumed to cost $500.

The test resource allocation problem was solved using the grid search algorithm and the adaptive test selection procedure. For this problem, both the algorithms yielded identical solutions: 15 temperature-based material level tests, and one flexural subsystem test, thereby exhausting the $2,000 budget. The sensitivity of the modulus (*E*) was low compared to the sensitivity of the other two calibration quantities (*r* and *c*), and hence no stress–strain material level tests were required. The decrease of variance with cost is shown in Fig. 7.8.

Similar to the structural dynamics example in Sect. 7.4, it is seen that the reduction in variance using the last $1,000 (i.e. from $1,000 to $2,000) was insignificant compared to the reduction in variance using the initial $1,000. Hence, from an economics-perspective, it is advantageous to use only $1,000 for testing as the exta $1,000 does not lead to further significant reduction in variance.

7.6 Conclusion

This paper proposed a Bayesian methodology for uncertainty quantification in multi-level systems. The various component, subsystem, and system models, and their inputs, parameters, and outputs, and experimental data were efficiently connected through a Bayes network. Further, the various sources of uncertainty – physical variability, data uncertainty, and model uncertainty were also included in the Bayes network. The Bayes network was then used for three different tasks: (1) calibrate the parameters of models at multiple levels using all available test data from multiple levels; (2) propagate the various sources of uncertainty (including the previously estimated model parameters) through the Bayes network to predict the overall uncertainty in the system-level prediction; and (3) aid in resource allocation for test selection and identify the most effective tests to reduce the overall uncertainty in the system-level prediction. The procedure for test resource allocation required Bayesian calibration and assessment of system-level prediction uncertainty even before actual testing was performed. This was achieved by generating multiple samples of test data and estimating the expected reduction in variance of the system-level prediction.

The algorithm for test resource allocation leads to several insights. A lower level test can easily isolate individual components and hence, the model parameters can be effectively updated, leading to a significant reduction in the variance of the system-level prediction. However, such a test would not account for the interaction between the higher level models and the corresponding parameters. In contrast, a higher level test would readily include the effects of interaction between multiple subsystem-level and component-level models. However, the calibration of parameters across multiple models may be difficult and may not lead to a significant reduction in the variance of the system-level prediction. The proposed test resource allocation procedure trades off between lower level tests and higher level tests by accounting not only for the resultant reduction in variance of the system-level prediction but also for the costs involved in testing.

Acknowledgment The research described in this paper was carried out at Vanderbilt University and the Jet Propulsion Laboratory, California Institute of Technology, under a contract (No. RSA 1400821, P. I. Dr. Lee Peterson) with the National Aeronautics and Space Administration.

References

1. Helton J (2009) Conceptual and computational basis for the quantification of margins and uncertainty. Sandia Report. SAND2009-3055
2. Urbina A, Mahadevan S, Paez T (2010) Resource allocation using quantification of margins and uncertainty. In: Proceedings of the 51st AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and materials conference, Orlando, 12–15 Apr 2010
3. Jensen FV, Jensen FB (2001) Bayesian networks and decision graphs. Springer, New York
4. Sankararaman S, Ling Y, Shantz C, Mahadevan S (2009) Uncertainty quantification in fatigue damage prognosis. In: Proceedings of the 1st annual conference of the prognostics and health management society, San Diego, 27 Sept–1 Oct 2009
5. Sankararaman S, Ling Y, Shantz C, Mahadevan S (2010) Parameter estimation and inversion of fracture-mechanics-based crack growth models under uncertainty. In: Proceedings of the inverse problems, design, and optimization symposium, Joao Pessoa, 25–27 Aug 2010
6. Haldar A, Mahadevan S (2000) Probability, reliability and statistical methods in engineering design. Wiley, New York
7. Red-Horse JR, Paez TL (2008) Sandia national laboratories validation workshop: structural dynamics application. Comput Meth Appl Mech Eng 197:2578–2584
8. Thornton EA (1996) Thermal structures for aerospace applications. In: Przemieniecki JS (ed) AIAA education series, AIAA, Reston, VA
9. Chopra AK (1995) Dynamics of structures: Theory and applications to earthquake engineering. Vol. 2. Prentice Hall, New Jersey

Chapter 8
Sampling Techniques in Bayesian Finite Element Model Updating

I. Boulkaibet, T. Marwala, L. Mthembu, M.I. Friswell, and S. Adhikari

Abstract Recent papers in the field of Finite Element Model (FEM) updating have highlighted the benefits of Bayesian techniques. The Bayesian approaches are designed to deal with the uncertainties associated with complex systems, which is the main problem in the development and updating of FEMs. This paper highlights the complexities and challenges of implementing any Bayesian method when the analysis involves a complicated structural dynamic model. In such systems an analytical Bayesian formulation might not be available in an analytic form; therefore this leads to the use of numerical methods, i.e. sampling methods. The main challenge then is to determine an efficient sampling of the model parameter space. In this paper, three sampling techniques, the Metropolis-Hastings (MH) algorithm, Slice Sampling and the Hybrid Monte Carlo (HMC) technique, are tested by updating a structural beam model. The efficiency and limitations of each technique is investigated when the FEM updating problem is implemented using the Bayesian Approach. Both MH and HMC techniques are found to perform better than the Slice sampling when Young's modulus is chosen as the updating parameter. The HMC method gives better results than MH and Slice sampling techniques, when the area moment of inertias and section areas are updated.

Keywords Bayesian • Finite element model updating • Hybrid monte carlo method • Markov chain monte carlo • Metropolis-hastings method • Sampling • Slice sampling method

8.1 Introduction

Finite element models (FEM) use a numerical method to provide approximate solutions for complex engineering problems [1, 2]. The FEM method is recognised as a powerful method for computing displacements, stresses and strains in structures under a set of loads. However, accurate solutions obtained by an FEM are only possible for few simple model cases. In complex systems FEM results are different from those obtained from experiments [3, 4]. These differences can be a result of the modelling errors and/or the uncertainties associated with the process of constructing the FEM of the structure. Therefore, the models obtained from the finite element method need to be updated to match the measurements. In the recent years, Bayesian model updating techniques have shown promising results in systems identification research [4–6].

The Bayesian approach provides accurate solutions for complex systems, due to its ability to characterize the uncertainties of these complex systems. In this paper, the Bayesian approach is used to update the uncertain parameter of the FEM where these can be represented as random vectors with a joint probability density function (pdf). This posterior (pdf) characterises

---

I. Boulkaibet (✉) • T. Marwala • L. Mthembu
The Centre For Intelligent System Modelling (CIMS), Electrical and Electronics Engineering Department, University of Johannesburg, 524, Auckland Park 2006, South Africa
e-mail: iboulkaibet@postgrad.uj.ac.za

M.I. Friswell
Aerospace Structures, College of Engineering, Swansea University, Singleton Park, Swansea SA2 8PP, UK

S. Adhikari
Aerospace Engineering, College of Engineering, Swansea University, Singleton Park, Swansea SA2 8PP, UK

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_8, © The Society for Experimental Mechanics, Inc. 2012

the uncertainty in the model parameters. The use of Bayesian techniques becomes useful when an analytical solution to this function is not available. This is often the case because of the high dimensionality of the parameter search space. In these situations some Markov chain Monte Carlo (MCMC) simulation methods provide the only practical solution [4, 6].

In this paper three sampling techniques, the Metropolis-Hastings (MH), the Slice Sampling (SS) and the Hybrid Monte Carlo (HMC) are investigated on their ability to sample the posterior pdf of FEM updating parameters. These three techniques are tested by updating a structural beam model. The efficiency, reliability and limitations of each technique are investigated when a Bayesian approach is implemented on a FEM updating problem.

In the next section, the finite element model background is presented. In Sect. 8.3, an introduction to the Bayesian framework is introduced where the posterior distribution of the uncertain parameters of FEM is also presented. Section 8.4 introduces the sampling techniques used to predict the posterior distribution. Section 8.5 presents an implementation of the Bayesian FEM updating on a simple cantilever steel beam, where two cases of study will be provided. Section 8.6 concludes the paper.

8.2 Finite Element Model Background

In finite element modelling, an N degree of freedom dynamic structure may be described by the matrix equation of motion [7]

$$\mathbf{M}\ddot{x}(t) + \mathbf{C}\ddot{x}(t) + \mathbf{K}x(t) = \mathbf{f}(t) \tag{8.1}$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are the mass, damping and stiffness matrices of size $N \times N$. $\mathbf{x}(t)$ is the vector of N degrees of freedom and $\mathbf{f}(t)$ is the vector of forces applied to the structure. In case that no external forces are applied to the structure and if the damping terms are neglected ($\mathbf{C} \sim 0$), the dynamic equation may be written in the modal domain (natural frequencies and mode shapes) where the error vector for the ith mode is obtained from

$$\left[-\left(\omega_i^m\right)^2\mathbf{M} + \mathbf{K}\right]\Phi_{\mathbf{i}}^{\mathbf{m}} = \varepsilon_i \tag{8.2}$$

ω_i^m is the *ith* measured natural frequency, $\phi_{\mathbf{i}}^{\mathbf{m}}$is the *ith* measured mode shape vector and $\varepsilon_{\mathbf{i}}$ is the *ith* error vector. In (8.2), the error vector $\varepsilon_i$is equal to 0if the system matrices $\mathbf{M}$ and $\mathbf{K}$ correspond to the modal properties (ω_i^m and ϕ_i^m). However, ε_i is a non-zero vector if the system matrices obtained analytically from the finite element model do not match the measured modal properties ω_i^m and $\phi_{\mathbf{i}}^{\mathbf{m}}$. In the case that (8.2) contains both measured vectors, ω_i^m and ϕ_i^m, and the matrices obtained from the analytic methods $\mathbf{M}$ and $\mathbf{K}$, the error obtained in (8.2) represents the difference between the measured and the analytic modal properties. Another problem might arise in many practical situations, where the dimension of the mode shape vector does not match the dimension of the system matrices. This is because the measured modal coordinates are fewer than the finite element modal coordinates. To ensure the compatibility between the system matrices and the mode shape vectors, the dimension of the system matrices is reduced using Guyan reduction [8].

8.3 Bayesian Inferences

In order to update the mathematical models, the uncertain parameters have to be identified. The Bayes theorem offers this possibility, where the uncertain parameters can be determined from their measurements [4, 9]. In this work, the Bayesian method is used to solve the finite element updating problem based on modal properties. Bayesian approaches are governed by Bayes rule [9, 10]:

$$P(\mathbf{E}|D) = \frac{P(D|\mathbf{E})P(E)}{P(D)} \tag{8.3}$$

where $\mathbf{E}$ represent the vector of updating parameters and the mass $\mathbf{M}$ and stiffness $\mathbf{K}$ matrices are functions of the updating parameters $\mathbf{E}$. The quantity $P(\mathbf{E})$, known as the prior probability distribution, is a function of the updating parameters in the absence of the data. D is the measured modal properties; the natural frequencies ω_i and mode shapes $\phi_{\mathbf{i}}$. The quantity $P(\mathbf{E}|D)$ is the posterior probability distribution function of the parameters in the presence of the data D. $P(D|\mathbf{E})$ is the likelihood probability distribution function and $P(D)$ is a normalization factor [4, 9, 10].

8.3.1 The Likelihood Distribution Function

The likelihood distribution can be seen as the probability of the modal measurements in the presence of uncertain parameters. The likelihood function $P(D|\mathbf{E})$ can be defined as the normalized exponent of the error function that represents the differences between the measured and the analytic frequencies.

This function can be written as follows

$$P(D|\mathbf{E}) = \frac{1}{Z_D(\beta)} exp\left(-\beta \sum_j^F \sum_i^{N_m} \left(\varepsilon_{ij}^2\right)\right) \tag{8.4}$$

where, β is the coefficient of the measured modal data contribution to the error. ε_{ij} represents the error between the measured and analytical frequencies where i indicates the ith modal properties and subscript j represents the jth measurement position. The superscript F is the number of measured mode shape coordinates; N_m is the number of measured modes and $Z_D(\beta)$ is a normalising constant given by

$$Z_D(\beta) = exp\left(-\beta \sum_j^F \sum_i^{N_m} \left(\varepsilon_{ij}^2\right)\right) d[D] = \left(\frac{\pi}{\beta}\right)^{FN_m/2} \tag{8.5}$$

8.3.2 Prior Distribution Function

The prior pdf represents the prior knowledge of the updating parameters **E** . It quantifies the uncertainty of the parameters **E**. In this paper, some parameters are updated more intensely than others. For example, parameters next to joints should be updated more intensely than for those corresponding to smooth surface areas far from joints. The prior probability distribution function for parameters **E** is assumed to be Gaussian and is given by [10, 11]

$$P(\mathbf{E}) = \frac{1}{Z_E(\alpha)} exp\left(-\sum_i^Q \frac{\alpha_i}{2} ||\mathbf{E}||^2\right) \tag{8.6}$$

where Q is the number of groups of parameters to be updated, and α_i is the coefficient of the prior pdf for the ith group of updating parameters. The $||*||$ is the Euclidean norm of $*$. In (8.6), if α_i is constant for all of the updating parameters, then the updated parameters will be of the same order of magnitude. Equation 8.6 may be viewed as a regularization parameter [12]. In (8.6), Gaussian priors were conveniently chosen because many natural processes tend to have a Gaussian distribution. The function $Z_{\mathbf{E}}(\alpha)$is a normalization factor given by [4, 10]

$$Z_{\mathbf{E}}(\alpha) = exp\left(-\sum_i^Q \frac{\alpha_i}{2} ||\mathbf{E}||^2\right) d[\mathbf{E}] = (2\pi)^{Q/2} \prod_{i=1}^Q \frac{1}{\sqrt{\alpha_i}} \tag{8.7}$$

8.3.3 Posterior Probability Distribution Function

The posterior distribution function of the parameters **E**given the observed data D is denoted as $P(\mathbf{E}|D)$and obtained by applying the Bayes' theorem as represented in (8.3). The distribution $P(\mathbf{E}|D)$is calculated by substituting (8.4) and (8.6) into (8.3) to give [4]

$$P(\mathbf{E}|D) = \frac{1}{Z_s(\alpha, \beta)} exp\left(-\beta \sum_j^F \sum_i^{N_m} \left(\varepsilon_{ij}^2\right) - \sum_i^Q \frac{\alpha_i}{2} ||\mathbf{E}||^2\right) \tag{8.8}$$

where

$$Z_s(\alpha, \beta) = Z_D(\beta).Z_{\mathbf{E}}(\alpha) \tag{8.9}$$

8.4 Sampling Techniques

Sampling techniques are very useful numerical methods which can be employed to simplify the Bayesian inference by providing a set of random samples from posterior distribution [6, 9, 10, 13]. Suppose that Y is the observation of certain parameters at different discrete time instants; the probabilistic information for the prediction of the future responses Y at different time instants is contained in the robust predictive PDF which is given by the Theorem of Total Probability as

$$P(Y|D) = P(Y|\mathbf{E})P(\mathbf{E}|D)d[\mathbf{E}] \tag{8.10}$$

Equation 8.10 depends on the posterior distribution function which is very difficult to solve analytically due to the dimension of the updating parameters. Therefore, sampling techniques, such as Markov chain Monte Carlo (MCMC) methods, are employed to predict the updating parameter distribution and subsequently to predict the modal properties. Given a set of N_s random parameter vector drawn from a pdf $P(\mathbf{E}|D)$, the expectation value of any observed function Y can be easily estimated.

The integral in (8.10) is solved using three different sampling techniques: the MH algorithm, the slice sampling algorithm and the HMC algorithm [6, 9, 14, 15]. These algorithms are used to generate a sequence of vectors $\{\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_{N_s}\}$ where N_s is the number of samples and these vectors can be used to form a Markov chain. This generated vector is then used to predict the form of the stationary distribution function$P(\mathbf{E}|D)$. The integral in (8.10) can be approximated as

$$\widetilde{Y} \cong \frac{1}{N_s}\sum\nolimits_{i=1}^{N_s} G(\mathbf{E}_i) \tag{8.11}$$

where G is a function that depends on the updated parameters $\mathbf{E}_i$. As an example, if $\mathbf{G} = \mathbf{E}$ then $\widetilde{Y}$ becomes the expected value of $\mathbf{E}$. Generally, $\widetilde{Y}$is the vector that contains the modal properties and N_s is the number of retained states.

8.4.1 The Metropolis-Hastings Algorithm

The MH algorithm is one of the simplest sampling methods, and is related to rejection and importance sampling [9]. To sample from the posterior distribution function $P(\mathbf{E}|D)$, where $\mathbf{E} = \{E_1, E_2, \ldots, E_{N_m}\}$ is a N_m-dimensional parameter vector, a proposal density distribution $q(\mathbf{E}|\mathbf{E}_{t-1})$ is introduced in order to generate a random vector $\mathbf{E}$ given the value at the previous iteration of the algorithm. The MH algorithm consists of two basic stages: the draw from the proposed density stage and the retained/rejected stage. The MH algorithm can be summarized as:

1. An initial value $\mathbf{E}_0$ is used to initiate the algorithm.
2. At iteration t, $\mathbf{E}^*$ is drawn from the proposed density $q(\mathbf{E}|\mathbf{E}_{t-1})$, where $\mathbf{E}_{t-1}$is the parameter value at the previous step.
3. Update the FEM to obtain the new analytic frequencies, then compute the acceptance probability, given by

$$a(\mathbf{E}^*, \mathbf{E}_{t-1}) = \min\left\{1, \frac{P(\mathbf{E}^*|D)q(\mathbf{E}_{t-1}|\mathbf{E}^*)}{P(\mathbf{E}_{t-1}|D)q(\mathbf{E}^*|\mathbf{E}_{t-1})}\right\}$$

4. Draw u from the uniform distribution $U(0, 1)$.
5. If $u \le a(\mathbf{E}^*, \mathbf{E}_{t-1})$ accept $\mathbf{E}^*$. Otherwise, reject $\mathbf{E}^*$
6. Return to step 2.

8.4.2 The Slice Sampling Algorithm

The slice sampling method is a simple technique that can provide an adaptive step size, which is automatically adjusted to match the characteristics of the posterior distribution function [9, 14]. In this method, the goal is to sample uniformly from the area under the posterior distribution $P(\mathbf{E}|D)$ where $\mathbf{E} = \{E_1, E_2, \ldots, E_{N_m}\}$. The algorithm of this technique can be described as [14]:

```
1. Draw $Y$ from the uniform distribution $U(0, P(\mathbf{E}_0|D))$.
2. Initiate the interval around $\mathbf{E}_0$ as follows:
        For $i = 1$ to $N$
            $U_i \sim Uniform(0,1)$
            $L_i \leftarrow E_{0,i} - w_i U_i$
            $R_i \leftarrow L_i + w_i$
        End
3. Sample from the interval $I = (R, L)$ and do the following:
   Repeat:
        For $i = 1 to N$
            $U_i \sim Uniform(0,1)$
            $E_{1,i} \leftarrow L_i + U_i(R_i - L_i)$
        End
        IF $Y \leq P(\mathbf{E}_0|\mathbf{E})$ Then exit loop
        For $i = 1 to N$
            IF $E_{1,i} < E_{o,i}$ Then $L_i \leftarrow E_{1,i}$ Else $R_i \leftarrow E_{1,i}$
4. Repeat step 3 to get $N_s$ samples.
```

8.4.3 Hybrid Monte Carlo

The Hybrid Monte Carlo method, known as the Hamiltonian Markov Chain method, is a good method for solving higher-dimensional complex problems [6, 9, 15]. In HMC, a new dynamical system is considered in which auxiliary variables, called momentum, $\mathbf{p} \in R^N$, are introduced and the uncertain parameters, $\mathbf{E} \in R^N$, in the target posterior distribution function are treated as displacements. The total energy, Hamiltonian function, of the new dynamical system is defined by $H(\mathbf{E}, \mathbf{p}) = V(\mathbf{E}) + W(\mathbf{p})$, where the potential energy is defined by $V(\mathbf{E}) = -\ln(P(\mathbf{E}|D))$ and the kinetic energy $W(\mathbf{p}) = \mathbf{p}^{\mathbf{T}}\mathbf{M}^{-1}\mathbf{p}/2$ depends only on $\mathbf{p}$ and some chosen positive definite matrix $\mathbf{M} \in R^{N \times N}$. Using Hamilton's equations, the evolution of (E, p)through time t and time step δt is given by

$$\mathbf{p}\left(t + \frac{\delta t}{2}\right) = \mathbf{p}(t) - \frac{\delta t}{2} \nabla V[\mathbf{E}(t)] \tag{8.12}$$

$$\mathbf{E}(t + \delta t) = \mathbf{E}(t) + \delta t \mathbf{M}^{-1}\mathbf{p}\left(t + \frac{\delta t}{2}\right) \tag{8.13}$$

$$\mathbf{p}(t + \delta t) = \mathbf{p}\left(t + \frac{\delta t}{2}\right) - \frac{\delta t}{2} \nabla V[\mathbf{E}(t + \delta t)] \tag{8.14}$$

where ∇V is obtained numerically by finite difference as

$$\frac{\partial V}{\partial E_i} = \frac{V(\mathbf{E} + \Delta h) - V(\mathbf{E} - \Delta h)}{2h\Delta_i} \tag{8.15}$$

$\Delta = [\Delta_1, \Delta_2, \ldots, \Delta_N]$is the perturbation vector, the distribution of which is user-specified and h is a scalar which dictates the size of the perturbation of **E**. After each iteration of Equations (8.12)–(8.14), the resulting candidate state is accepted or rejected according to the Metropolis criterion based on the value of the Hamiltonian $H(\mathbf{E}, \mathbf{p})$. Thus, if $(\mathbf{E}, \mathbf{p})$ is the initial state and $(\mathbf{E}^*, \mathbf{p}^*)$ is the state after the equations above have been updated, then this candidate state is accepted with probability $\min(1, exp\{H(\mathbf{E}, \mathbf{p}) - H(\mathbf{E}^*, \mathbf{p}^*)\})$**.** The new vector **E** will be used for the next iteration and the algorithm will be stopped when a N_s samples of **E**are provided.

8.5 Beam Example

An experimental cantilever steel beam is updated based on the measurements of Kraaij [16]. The beam has the following dimensions: length 500 mm, width 60 mm and thickness 10 mm. $E = 2.1 \times 10^{11}\mathrm{N/m^2}$, $v = 0.3$ and $\rho = 7850\mathrm{kg/m^3}$. Three accelerometers were used in the experiment, which were all located 490 mm from the clamped end. This location is chosen because the response on this point is large [16]. Each accelerometer has a mass of 40 g; the middle accelerometer is of type 303*A*3, the outer accelerometers are of type 303*A*2 (see [16] for more details of experimental set-up).

To test the updating methods, the beam was modeled using Version 6.3 of the Structural Dynamics Toolbox SDT® for MATLAB. The beam was divided into 50 Euler–Bernoulli beam elements and excited at various positions. The measured natural frequencies of interest of this structure are: 31.9 Hz, 197.9 Hz, 553 Hz, 1082.2 Hz and 1781.5 Hz, which correspond to modes 1, 3, 5, 7, and 9, respectively. For the first set of experiments, the Young's modulus of the beam elements was used as an updating parameter where for every 10 elements a different Young's modulus is allocated. Thus, the parameters to be updated can be represented by a vector of 5 variables $\mathbf{E} = \{E_1, E_2, E_3, E_4, E_5\}$. In the second set of experiments, the moments of inertia and the section areas are updated. This is done by associating an area moment of inertia, I_x, and an area, A_x, to every 25 elements of the beam (this will reduce the number of the parameters to be updated to four parameters). The updated parameters vector is thus $\mathbf{E} = \{I_{x1}, I_{x2}, A_{x1}, A_{x2}\}$.

8.5.1 *Updating Young's modulus*

In this section, a vector of 5 parameters $\mathbf{E} = \{E_1, E_2, E_3, E_4, E_5\}$ is updated using the Bayesian approach. The reason for using a large number of updating parameters is to determine the performance and the convergence speed of each sampling technique when a large number of variables are introduced in the updating process. N_s samples of the vector **E** were generated from the posterior distribution function, $P(\mathbf{E}|D)$ mentioned in (8.8). The constant β in (8.8) was set equal to 1, and the coefficients α_i were set equal to $\frac{1}{\sigma_i^2}$, where σ_i^2 is the variance of the parameterE_i. Since the updating parameter vector contains only the Young's modulus, all σ_i were set equal to 2×10^{11} (a large value of $\sigma_i$to weight the value of E). The updating parameters E_i were bounded with a maximum equal to 2.5×10^{11} and a minimum equal to 1.7×10^{11}. The number of samples N_s was 1,000, for all techniques, and the initial vector of **E** is $\{2.4 \times 10^{11}, 2.4 \times 10^{11}, 2.4 \times 10^{11}, 2.4 \times 10^{11}, 2.4 \times 10^{11}\}$. Instead of using the mean steel value of **E** as an initial value, a large value of the initial parameter vector is chosen to highlight the updating process. The results for the updating vector and the frequencies are given in Tables 8.1 and 8.2.

The MH technique updates all vector parameters simultaneously and gives results that are close to the mean value for steel of $2.1 \times 10^{11}\mathrm{N/m^2}$. The same comment can be made for the HMC technique where three of these parameters are close to the mean value. The Slice sampling method gives updating parameters that are far from the mean, because of the way that the Slice technique generates samples. The Slice sampling technique sequentially updates individual vector entries as oppose to updating all entries simultaneously, see Sect. 8.4. In this case, a small adjustment of the first parameter can cause a significant updating of the rest of the updating vector.

The results in Table 8.2 show that the three sampling techniques give results that on average are better than the initial FEM. The three algorithms weight the measured frequencies in different ways because each algorithm has a different way of generating samples. In the Slice sampling technique, each variable is changed one at a time which is not the case with the M-H algorithm where all the parameters are varied at once. The HMC technique uses additional parameters to evaluate the sampling.

The MH and HMC algorithms give better results than those obtained by using the Slice Sampling method. The error between the third measured natural frequency and that of the initial model was 7.55%. When the MH method was used, this error was reduced to 2.15% and in using the Slice sampling method it was only reduced to 3.21%. The same comment can be

Table 8.1 The updated vector of Young's modulus using MH, Slice Sampling and HMC techniques

	Young's modulus (N/m^2)			
	Initial	M-H method	Slice sampling method	HMC method
E_1	2.4×10^{11}	2.177×10^{11}	2.387×10^{11}	1.727×10^{11}
E_2	2.4×10^{11}	2.151×10^{11}	1.851×10^{11}	2.344×10^{11}
E_3	2.4×10^{11}	1.944×10^{11}	1.849×10^{11}	1.838×10^{11}
E_4	2.4×10^{11}	1.785×10^{11}	1.853×10^{11}	2.052×10^{11}
E_5	2.4×10^{11}	2.036×10^{11}	1.864×10^{11}	2.004×10^{11}

Table 8.2 Frequencies and errors when MH, slice sampling and HMC techniques used to update Young's modulus

Modes	Measured frequency (Hz)	Initial frequency (Hz)	Error (%)	Frequencies M-H method (Hz)	Error (%)	Frequencies slice sampling method (Hz)	Error (%)	Frequencies HMC method (Hz)	Error (%)
1	31.9	32.7	2.51	30.8	3.38	30.8	3.44	29	9
2	197.9	209.4	5.83	190.3	3.85	190.9	3.51	186.2	5.92
3	553	594.8	7.55	541.1	2.15	535.2	3.21	548	0.9
4	1082.2	1237.2	14.32	1123.6	3.83	1127.6	4.2	1099.99	1.64
5	1781.5	1961.7	10.12	1798.2	0.93	1772.4	0.51	1773.6	0.45

made for the first and fourth natural frequencies. The errors for the second and fifth modes are targeted by the Slice sampling method, where the error is a little smaller than that obtained from the MH method. However,, the overall result shows that the MH method gave better results than the Slice sampling method in terms of errors. Also the convergence of the MH method was faster than the Slice sampling method. The results show that Slice sampling is inefficient in sampling Young's modulus when 1,000 samples are generated. The HMC method converges faster than the other methods (MH and Slice sampling techniques). In addition, the HMC method gives a bad error for the two first modes where the error between the first measured natural frequency and that of the initial model was 2.51%. When the HMC method was used, this error was increased to 9% (similarly for the second mode). However, for the other three modes the HMC method gave better results than both the MH and Slice sampling methods. For example, the error between the third measured natural frequency and that of the initial model was 7.55%; the HMC method reduced this error to 0.9% whereas the Slice sampling method only reduced it to 3.21%. The same comment can be made in using MH algorithm where the error for the third mode only reduced to 2.15%. The three methods did not improve the first natural frequency because the same coefficient, β, was set for all of the natural frequencies. Choosing β as a vector, to weight the natural frequencies, would improve the first natural frequency results. In general, the error between the measured frequency and these obtained by the three algorithms is a bit high when the Young's modulus is updated. In this case, conclusions about the most efficient method are difficult. For this reason, another updating exercise is performed, where the area moments of inertia and cross-section area are updated and the results obtained are discussed in Sect. 8.5.2.

8.5.2 *Updating Area Moments of Inertia and Cross-section Area*

In this section, four parameters $\mathbf{E} = \{I_{x1}, I_{x2}, A_{x1}, A_{x2}\}$ are updated using the Bayesian approach. N_s samples of the vector $\mathbf{E}$ were generated from the posterior distribution function $P(\mathbf{E}|D)$. The constant β is equal 1, and all coefficients α_i are set equal to $\frac{1}{\sigma_i^2}$, where σ_i^2 is the variance of the parameterE_i. Since the updating parameter vector contains area moments of inertia and section areas, the vector of σ_i is defined as $\sigma = \left[5 \times 10^{-9}, 5 \times 10^{-9}, 5 \times 10^{-4}, 5 \times 10^{-4}\right]$. The Young's modulus is set to 2.1×10^{11}N/m^2 as opposed to 2.4×10^{11} N/m^2 in the previous section. The updating parameters E_i are bounded by maximum values equal to $\left[7.5 \times 10^{-9}, 7.5 \times 10^{-9}, 9 \times 10^{-4}, 9 \times 10^{-4}\right]$ and minimum values equal to $\left[3.5 \times 10^{-9}, 3.5 \times 10^{-9}, 4.5 \times 10^{-4}, 4.5 \times 10^{-4}\right]$. The number of samples N_s is set to 1,000. The results are given in Tables 8.3 and 8.4.

The MH and HMC techniques update all parameters simultaneously and again give results that are closer to the the initial parameter vector than those obtained by using the Slice sampling technique (see Table 8.3). The Slice sampling method gives updating parameters far from the mean value of $\mathbf{E}$ and this because of the way that the Slice sampling technique updates (the Slice sampling technique updates each parameter in turn).

Table 8.3 The updated vector using MH, Slice Sampling and HMC techniques

	Initial **E**	**E** vector, M-H method	**E** vector, slice sampling method	**E** vector, HMC method
I_{x1}	5×10^{-9}	7.05×10^{-9}	7.88×10^{-9}	6.96×10^{-9}
I_{x2}	5×10^{-9}	6.36×10^{-9}	3.67×10^{-9}	6.41×10^{-9}
A_{x1}	8×10^{-4}	8.28×10^{-4}	5.62×10^{-4}	8.92×10^{-4}
A_{x2}	8×10^{-4}	8.46×10^{-4}	8.09×10^{-4}	8.13×10^{-4}

Table 8.4 Frequencies and errors when MH, slice sampling and HMC techniques used to update momentums and area sections

Modes	Measured frequency (Hz)	Initial frequency (Hz)	Error (%)	Frequencies M-H method (Hz)	Error (%)	Frequencies slice sampling method (Hz)	Error (%)	Frequencies HMC method (Hz)	Error (%)
1	31.9	32	0.32	31.3	1.97	32.9	3.14	31.6	1.08
2	197.9	203.5	2.58	195.2	1.34	183.3	7.1	193.8	2.06
3	553	575.7	4.1	552	0.19	556.4	0.62	545.3	1.4
4	1082.2	1136.8	5.04	1087.6	0.5	1,084	0.16	1,077	0.48
5	1781.5	1889.6	6.06	1811.2	1.67	1799.4	1.00	1790.7	0.51

In this section, the results obtained give lower errors than those in Table 8.2. The results in Table 8.4 show that the HMC and HM methods give better updating than the Slice sampling algorithm. This can be seen by comparing the errors for modes 1 and 2. The error between the second measured natural frequency and that of the initial model was 2.58%. When the MH and HMC methods were used, the error was reduced to 1.34% and 2.06% respectively, and in using the Slice sampling method it was increased to 7.1%. Moreover, the results in Table 8.2 give a priority to HMC over MH and Slice sampling algorithms in term of convergence time since the HMC approach converges faster than the other methods. In addition, the HMC technique gives a small error compared to that obtained by using MH algorithm, where in modes 1, 4 and 5 the HMC method gives small frequency errors compared to the MH method. Despite the errors that may occur when the gradients are evaluated numerically, the HMC algorithm shows good results.

8.6 Conclusion

In this paper the finite element model updating problem is posed as a Bayesian problem. This means the uncertainty associated with the model parameters and the distribution of the data with these parameters is concisely formulated in Bayes theorem. To evaluate the resultant high dimensional posterior distribution three MCMC sampling techniques are implemented; the Metropolis-Hastings algorithm, Slice Sampling and the Hybrid Monte Carlo technique.

These sampling techniques are tested on a simple beam structure. In the first simulation the Hybrid Monte Carlo and Metropolis-Hastings technique gave more accurate results than the Slice sampling method. In addition the HMC method converges faster than both the MH and Slice sampling algorithms. In the second simulation the HMC method gave better results than both the MH and slice sampling methods and has more attractive convergence properties than the latter. Further work includes testing these algorithms on more complicated dynamic structures.

References

1. Onãte E (2009) Structural analysis with the finite element method. Linear statics, vol 1, Basis and solids. Springer, Dordrecht/London
2. Rao SS (2004) The finite element method in engineering, 4th edn. Elsevier Butterworth Heinemann, Burlington
3. Friswell MI, Mottershead JE (1995) Finite element model updating in structural dynamics. Kluwer, Dordrecht/Boston
4. Marwala T (2010) Finite element model updating using computational intelligence techniques. Springer, London
5. Yuen KV (2010) Bayesian methods for structural dynamics and civil engineering. Wiley, New York
6. Cheung SH, Beck JL (2009) Bayesian model updating using hybrid Monte Carlo simulation with application to structural dynamic models with many uncertain parameters. J Eng Mech 135(4):243–255
7. Ewins DJ (1984) Modal testing: theory and practice. Research Studies, Letchworth

8. Guyan RJ (1965) Reduction of stiffness and mass matrices. Am Inst Aeronaut Astronaut 11(5):380–388
9. Bishop CM (2006) Pattern recognition and machine learning. Springer, New York
10. Marwala T, Sibisi S (2005) Finite element model updating using Bayesian approach. In: Proceedings of the international modal analysis conference, Orlando, ISBN: 0-912053-89-5
11. Bishop CM (1995) Neural networks for pattern recognition. Oxford University Press, Oxford, UK
12. Vapnik VN (1995) The nature of statistical learning theory. Springer, New York
13. Ching J, Leu SS (2009) Bayesian updating of reliability of civil infrastructure facilities based on condition-state data and fault-tree model. Reliab Eng Syst Saf 94(12):1962–1974
14. Neal RM (2000) Slice sampling. Technical Report, No. 2005, Department of Statistics, University of Toronto
15. Hanson KM (2001) Markov Chain Monte Carlo posterior sampling with the Hamiltonian method. Proc SPIE 4322:456–467
16. Kraaij CS (2007) Model updating of a ‘clamped’-free beam system using FEMTOOLS. Technische Universiteit Eindhoven

Chapter 9
Bayesian Model Updating Approach for Systematic Damage Detection of Plate-Type Structures

Masahiro Kurata, Jerome P. Lynch, Kincho H. Law, and Liming W. Salvino

Abstract This paper presents a model-based monitoring framework for the detection of fatigue-related crack damages in plate-type structures commonly seen in aluminum ship hulls. The monitoring framework involves vibration-based damage detection methodologies and finite element modeling of continuum plate structures. A Bayesian-based damage detection approach is adopted for locating probable damage areas. Identifying potential damage locations by evaluating all possible combinations of finite elements in the model is computationally infeasible. To reduce the search space and computational efforts, initial knowledge of the probable damage zones and a heuristic-based branch-and-bound scheme are systematically included in the Bayesian damage detection framework. In addition to an overview of the model-based monitoring framework, preliminary results from numerical simulations and experimental tests for a plate specimen with a welded stiffener are presented to illustrate the Bayesian damage detection approach and to demonstrate the potential application of the approach to detect fatigue cracks in metallic plates.

Keywords Aluminum plate • Bayesian model updating • Branch and bound technique • Damage detection

9.1 Introduction

The adoption of aluminum alloy materials for ship structures presents many challenges to the naval engineering community. In addition to being light-weight, aluminum alloys have high corrosion resistance but often exhibit fatigue-related micro-cracks. High-speed aluminum hulls can remain in operation even after the initiation of micro-cracks because of the ductile mechanical characteristic of the material. Therefore, to enhance the maintenance of ships and to prevent catastrophic failure of ship's hull, one key component in monitoring aluminum ship structures is the early detection of fatigue-induced cracks.

This paper presents a preliminary investigation of a computationally efficient and practical framework for the monitoring and performance assessment of aluminum plate structures. Specifically, aluminum plate specimens have been designed and built with welded assemblies to facilitate the investigation of system identification and damage detection methods. The area of damage in the structure is estimated by a model-based damage detection methodology that compares the structural characteristics of the "true" structure (damaged or undamaged) to finite element "trial" models. The probability associated with a hypothesized damage state (e.g., location and size) is evaluated through the calculation of an error between the "true" and "trial" models. Incorporating with prior knowledge and computational heuristics, the probable damage area (i.e., fatigue crack path) is identified by repeatedly applying a Bayesian inference algorithm [1]. The Bayesian probabilistic model-updating framework is applied to validate its potential applicability for damage identification around critical weld zones in the plate specimens. Numerical simulations and experimental tests are conducted to examine the model-based approach to the testing of aluminum plates with crack damages.

M. Kurata • J.P. Lynch (✉)
Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109-2125, USA
e-mail: jerlynch@umich.edu

K.H. Law
Civil and Environmental Engineering, Stanford University, Stanford, CA 94305-4020, USA

L.W. Salvino
Naval Surface Warfare Center, Carderock Division, West Bethesda, MD 20817-5700, USA

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_9,

9.2 Methodology

9.2.1 *Model Updating for Damage Detection*

A general class of damage detection algorithms is based on model updating. Model parameters are varied until the model approximates the behavior of the true observed system. To identify an optimal model, the outputs from the observed system are used to evaluate a pre-defined objective (or error) function. An appropriate objective function is one that takes into account the fundamental behavior of the system in both its damaged and undamaged states, yet is compatible with experimental data available. Changes in the model parameters are then correlated to the condition (damage vs. undamage) of the structure.

Consider a simple plate structure. The governing equation describing the dynamic behavior of a vibrating plate can be written in the following form [2]:

$$D\nabla^4 w + \rho h \frac{\partial^2 w}{\partial t^2} = q(x, y, t) \tag{9.1}$$

where D is the flexural rigidity and is defined as:

$$D = \frac{Eh^3}{12(1 - v^2)} \tag{9.2}$$

Here, w denotes the vertical displacement of the plate, $q(x,y,t)$ is the normal load distribution function on the top of the plate, ρ is the plate density (mass per unit area), h is the plate thickness, E is the Young's (elastic) modulus and v is the Poisson ratio.

For a finite element (FE) model with N elements, the elastic moduli for the elements can be denoted as:

$$E = \{E_1, E_2, \cdots, E_i, E_{i+1}, \cdots, E_{i+n}, \cdots, E_{N-1}, E_N\} \tag{9.3}$$

For modeling purposes, fatigue crack damage is represented by a change in the flexural rigidity or stiffness of the plate, which, in turn, is reflected by a reduction in the elastic modulus of the damaged elements. If the model contains n damaged elements, their elastic moduli $\{E_{i+1}, \cdots, E_{i+j}, \cdots, E_{i+n}\}$ are replaced by

$$E_d = \{k_{i+1}E_{i+1}, \cdots, k_{i+j}E_{i+j}, \cdots, k_{i+n}E_{i+n}\} \tag{9.4}$$

where k_{i+j} (<1), $j = 1,\ldots,n$, denotes the reduction factor for the $(i + j)$th damaged element.

The model updating problem is posed as a combinatorial optimization problem that seeks to find the optimal set of elastic moduli that minimizes a defined objective function by comparing the FE model output and the output derived from actual measurements taken from the structure. In other words, the inverse problem can be posed as a combinatorial optimization problem for finding the optimal set of n elements, whose effective elastic moduli represent the damage state of the real structure as denoted in (9.4).

9.2.2 *A Flexibility-Based Objective Function*

Conceptually, damage detection methods seek to identify changes (damages) in a structure by using certain structural characteristics as the basis for evaluation. Vibration data are among the most common sensor information easily obtained from dynamic excitations. Vibration-based damage identification methods using mode-shape information have been proposed [3]. One approach is to detect damages directly based on the changes in structural characteristics between damaged and undamaged structures. Examples of structural characteristics include curvature mode shape [4], flexibility parameters [5, 6], strains and modal strain energy [7]. Although current research in damage detection has made substantial progress, the methods developed so far are primarily restricted to simple beam or frame structures with a limited number of degrees-of-freedom (DOFs). In contrast, continuum systems, such as plate-like structures, pose many challenges for current methods because of the large number of DOFs that are required to properly model the structure. Additionally, it is not realistic to place sensors to measure the system responses at all of the DOFs. An alternative approach is to construct a system model (such as a FE model) of a target structure using measurement data. This "model-based" approach updates the system model by modifying the structural properties of the elements until the model resembles the dynamic characteristics estimated from sensor information. The objective is to define an objective (error) function and to compute the difference (error) between the trial FE model and the target structure. Damages, if any, are identified and revealed through the changes in the model parameters during the model-updating process.

For the plate problem, our study indicates that the use of flexibility matrices constructed based on modal properties (modal frequencies and mode shapes), as opposed to the direct use of modal properties, is an appropriate choice for the objective function [8]. The inverse relationship between the flexibility matrix and the square of modal frequency renders the flexibility matrix as less sensitive to high frequency modes which are difficult to identify in vibration tests. This unique characteristic allows for the inclusion of lower order modes in a truncated flexibility matrix. This feature has attracted many researchers to explore flexibility as a core element in developing structural damage detection algorithms [5–7, 9]. In the "Flexibility-Based Approach" (FBA) described below, the objective (error) function is expressed in terms of the difference between the flexibility matrices that correspond to the true (measured) and trial (FE) models.

When the mode shapes are mass-normalized (i.e., $\bar{\phi}^T M \bar{\phi} = I$), the flexibility matrix F can be expressed in terms of the modal properties as follows:

$$F = \bar{\phi}\Omega^{-1}\bar{\phi}^T = \sum\nolimits_{i=1}^{N} \frac{1}{\omega_i^2}\bar{\phi}_i\bar{\phi}_i^T \tag{9.5}$$

where M is the mass matrix, $\bar{\phi} = [\bar{\phi}_1\bar{\phi}_2 \ldots \bar{\phi}_N]$ is the mode shape matrix, $\Omega = \mathrm{diag}(\omega_i^2)$ is the spectral matrix consisting the square of the modal frequencies ω_i, and N is the number of DOFs in the system. The mass-normalized modal vector $\bar{\phi}_i$ is related to the arbitrarily scaled mode shape ϕ_i as:

$$\bar{\phi}_i = \phi_i d_i \tag{9.6}$$

where $d_i = \frac{1}{\sqrt{\phi_i^T M \phi_i}}$ is a mass normalization constant for the ith mode.

Suppose only a few (lower) modes are available (e.g., from experimental tests), a truncated flexibility matrix is obtained as:

$$F_{\mathrm{trun}} = \sum\nolimits_{i=1}^{n} \left(\frac{d_i}{\omega_i^2}\right)^2 \bar{\phi}_i\bar{\phi}_i^T \tag{9.7}$$

where n denotes the number of modes available.

Let's define the difference (ΔF_{trun}) between the flexibility matrices of the true (damaged) structure and the trial FE model as:

$$\Delta F_{trun} = F_{trun}^{true} - F_{trun}^{trial} \tag{9.8}$$

When a trial FE model reasonably resembles the damaged structure with true damage, the difference in flexibility matrices is close to zero (exactly zero if there is no measurement noise nor modeling error). The scalar magnitude on the difference in the flexibility matrices can be measured by calculating the Frobenius norm of ΔF_{trun}:

$$\|\Delta F_{trun}\|_F = \sqrt{\sum\nolimits_i \sum\nolimits_j x_{ij}^2} \tag{9.9}$$

which vanishes when all matrix elements x_{ij} in ΔF_{trun} are zero (i.e., the FE model perfectly matches the observed structure).

The difference in the flexibility matrices can also be further decomposed into singular values by singular value decomposition (SVD):

$$\Delta F_{trun} = USV^T \tag{9.10}$$

where V^T and U are matrices of singular vectors, S is the diagonal matrix whose elements are the singular values, s_i. Since the Frobenius norm is invariant under unitary transformation, the SVD of the ΔF_{trun} yields

$$\|\Delta F_{trun}\|_F = \left\|USV^T\right\|_F = \|S\|_F = \sqrt{s_1^2 + s_2^2 + \cdots + s_R^2} \tag{9.11}$$

where R is the rank of ΔF_{trun}. Thus, the objective is to search for a trial model that minimizes the Frobenius norm shown in (9.11).

Since the mode shapes experimentally obtained are arbitrary scaled, the mass normalization constants (d_i) are required to properly compute the flexibility matrix of the real, true structure. One approach to extracting the mass normalization constants is based on testing the structure with a perturbed mass matrix (by adding a known mass at a certain location) and examining the sensitivity of the eigenvalues [10, 11]. In this study, mass normalization constants are estimated using the FE model and applied to the experimentally obtained (i.e., not mass normalized) mode shapes.

9.2.3 Bayesian Probabilistic Approach

The model updating procedure adopted in this study is based on a Bayesian probabilistic approach which utilizes the parameters measured or estimated from a series of collected vibration signals or data. Unlike a deterministic optimization formulation, the state space search must reflect the relative degree of belief on the estimates of the optimal subset (i.e., E_d, in this case). Let M denote the hypothesized damage states of the model. The calculation of the error that exists between the "true" structure and the FE "trial" model is based on the measured or estimated structural parameters. The updated estimates on the damage of the structure are expressed as the posterior distribution based on Baye's rule as follows

$$p(M|s) = \frac{p(s|M)p(M)}{p(s)} \propto p(s|M)p(M) \tag{9.12}$$

where $p(M|s)$ is the posterior distribution function for a hypothesis M given the measured or estimated parameters, s. $p(s|M)$ is termed the likelihood function, $p(M)$ is the prior probability of the hypothesis and $p(s)$ is treated here as some normalization constant. By collecting the likelihood function, the posterior distribution $p(M|s)$ would progressively become a better estimate than the prior probability $p(M)$ as the process goes on. For instance, if the initial estimate of the probable area of damage is assumed to be uniformly distributed over the structure before the model updating procedure, the most likely damaged areas are revealed with relatively higher posterior estimates by the repeated applications of the Bayesian inference process.

The selection of the most probable events from all conceivable possibilities using the Bayesian probabilistic approach can be systematized using various "optimal" search methods; otherwise a random search in optimal subspace becomes a computationally intractable task. To sample the posterior parameter distribution, the Bayesian approaches implemented with the Markov Chain Monte Carlo (MCMC) method and genetic algorithms (GA) have been reported with successful performance in detecting structural damage [12–14]. However, these approaches are computationally expensive and they do not guarantee convergence to an optimal solution. To reduce the computational effort, the Bayesian damage detection algorithm proposed herein is enhanced with a branch-and-bound (BB) search technique where the search space is systematically narrowed through enumeration and pruning of candidate model solutions [1].

The BB technique is a general search method originally developed for discrete optimization problems and is a powerful technique for controlling the size of a search space used in model updating [15]. As illustrated in Fig. 9.1, the BB algorithm initially starts its search from some random subspaces (i.e., leaf nodes associated with Branch 1 in Fig. 9.1a). At each step, the algorithm takes an additional sample (i.e., an additional element in the hypothesized subset) at each leaf node and aims to improve its estimates (this process is called branching). As the search proceeds, the branches associated with large "errors" are pruned and the search is bounded by evaluating the remaining branches (this process is called pruning). Figure 9.1b schematically shows the application of BB technique in FE updating of a plate-type structure. Here, the crack is assumed to damage the entire section of the plate in the thickness direction and represented in the model by a set of finite elements with substantially reduced elastic modulus for simplicity. The collection of survived leaf nodes (or "trial" model with small error) is mapped to reveal a potential damage area at every branching process. Therefore, the probable damage area in a system can

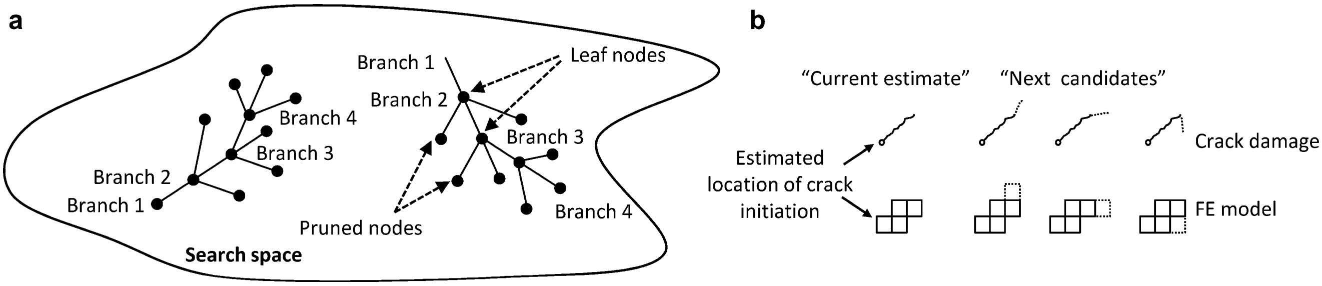


Fig. 9.1 Application of the branch and bound (BB) technique for model-updating: (**a**) concept of the branch-and-bound search scheme, (**b**) candidates for probable damage state

be systematically updated by implementing the BB technique in the Bayesian formulation. The computational effort and accuracy of the estimates for the optimal subset (i.e., E_d) can be controlled by pruning the less likely candidates during the branching process.

9.2.4 Summary of the Model Based Bayesian Damage Detection Framework

In summary, the model-based structural monitoring framework consists of three phases. The first phase is the construction of the FE model with structural properties and boundary conditions close to that of a target structure. In the second phase, the monitoring system acquires information about the status of the target structure through a network of sensors deployed to the structure. Once the variation in the modal properties exceeds a pre-defined threshold value, the third phase involving the Bayesian-based damage detection algorithm proceeds to compute the error associated with an initial set of hypotheses using the flexibility-based objective function. Engineering judgment based on knowledge about common damage patterns can be very useful to account for the selection of the initial set of hypotheses; for instance, heat affected zones around welds or notches along edges are good initial candidates for hypothetical probable damaged areas. The branching process adds one more damaged element to each hypothesis following the branching rule shown in Fig. 9.1b. Hypotheses with relatively large error are pruned before the next branching process. The branching, error computing and pruning processes make a finite loop until the damage identification algorithm terminates. At the end of each loop, the elements survived in the current and previous pruning processes are classified and stored in bins based on their error. The histogram-like plot of survived elements serves as a damage map showing the probable area(s) of damage. The search process is terminated when the deviation between the current and last damage areas converges within a predefined tolerance.

9.3 Numerical Simulation and Experimental Validations

The Bayesian model-updating procedure is applied to the problem of damage detection on a stiffened aluminum plate as shown in Fig. 9.2a. The design of the structural plate is intended to include the geometric complexity that is commonly found in aluminum ship hull structures. The aluminum plate includes an area with high stress concentration due to the presence of a welded stiffener. Such areas are the likely location for fatigue-related damage (i.e., fatigue cracking). Knowledge of this fact allows one to customize the model updating algorithm to prioritize the search of this area. The aluminum plate is 24 in. × 48 in. and is ¼ in. thick. In addition, the plate has a 2 in. × 18 in. × ¼ in. stiffener plate; the base plate and the stiffener plate are rigidly welded together. Crack paths in the main plate initiating from the heat affected zone (HAZ) around the welded stiffener plate are considered as example damage scenario cases.

The plate structure is modeled using ABAQUS, a general-purpose FE analysis program. The base plate and stiffener plate are modeled with 4-node reduced integration, doubly curved shell elements with hourglass control (*S4RSW*). These plates are assumed to be rigidly connected. The mesh size of the plate elements are 1 in. × 1 in. The elastic modulus E of the undamaged elements is 10,300 ksi while for the cracked elements, the elastic modulus is reduced 10^{-6} times the original

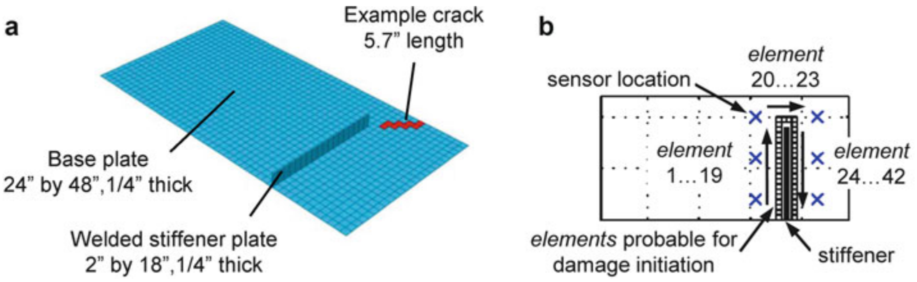


Fig. 9.2 A testbed plate structure: (**a**) schematics of the test-bed plate structure, (**b**) sensor locations and probable damage zone

modulus ($E_d = E \times 10^{-6}$). The mass density of the aluminum is assumed to be 2.489×10^{-7} slug/in.^{3} The *Lanczos* frequency analysis method in ABAQUS is employed to compute the modal properties (i.e., modal frequency and mode shape) of the base structure as well as the structures with different damage scenarios. Only the first five modes are considered in both the numerical and experimental simulations. Additionally, sensors are placed (or assumed to be placed for numerical simulations) at strategic locations as shown in Fig. 9.2b.

9.3.1 Numerical Simulations

For the numerical simulation study, the test-bed plate structure as shown in Fig. 9.2a is employed for evaluating the model-based damage detection algorithm. Two separate cases are considered. First, damage detection of the plate with a single stiffener having individual cracks around the stiffener is investigated. Second, the plate is assumed to have an additional welded stiffener and the initial probable damage zone is expanded to include areas around both stiffeners and along the edges of the plate.

Damage detection of individual crack path with a single stiffener: Numerical simulations of the plate with a single stiffener are studied first. The three crack paths (i.e. the targeted area) considered in the numerical studies are shown in Fig. 9.3. The modal properties of the undamaged plate structure and of the hypothetical damaged plate are obtained first. Table 9.1 summarizes the modal frequencies obtained for the undamaged and damaged plates. Using the computed modal frequencies and mode shapes, the flexibility matrices of the damaged models are constructed using (9.7).

The Bayesian-based damage detection procedure, as described in phase 3 of the model-based monitoring framework, is employed to identify the individual cracks. Figure 9.3a illustrates the basic process of locating crack path 1 (the “Target”) as shown. Starting from the probable damaged elements in the vicinity of the weld as shown in Fig. 9.2b, at each branching step, the probabilistic branch-and-bound scheme proceeds to search for the most likely set of damaged elements based on the objective function shown in (9.11), which evaluates the difference between the trial FE model and the true damaged FE model. The results shown in the figure is based on a pruning rate of 50% (i.e., only 50% of the hypotheses which have smaller error are retained in the next branching process). In the damage maps, the elements marked with darker color have the higher probabilities of being damaged. As shown in Fig. 9.3a, the procedure progressively converges to the crack region and is terminated when the variation between 8th and 9th branching results becomes negligible.

The probabilistic model-based updating procedure is also applied for the detections of crack paths 2 and 3 as shown in Fig. 9.3b, c respectively. These two cracks are slightly shorter than crack 1 and they are located internally inside the plate boundaries. The detection of these two cracks is more challenging because of the relatively small changes in the modal properties between the undamaged and damaged plates as shown in Table 9.1. As shown in Fig. 9.3b, the model-updating algorithm is able to identify accurately the location of crack path 2. It can be seen from the damage map that the search for probable damaged elements converges quickly with only a modest number of branching steps. Crack path 3 is a particularly challenging case because the crack is oriented in the longitudinal direction (i.e., along the longer side of the plate). It may also be interesting to point out that the finite elements are uniformly and regularly distributed in the mesh. As shown in Table 9.1, there are very little changes in the modal properties between the damaged and the undamaged plates. Nonetheless, as shown in Fig. 9.3b, the model-updating procedure is able to locate the probable damage area that includes the crack but

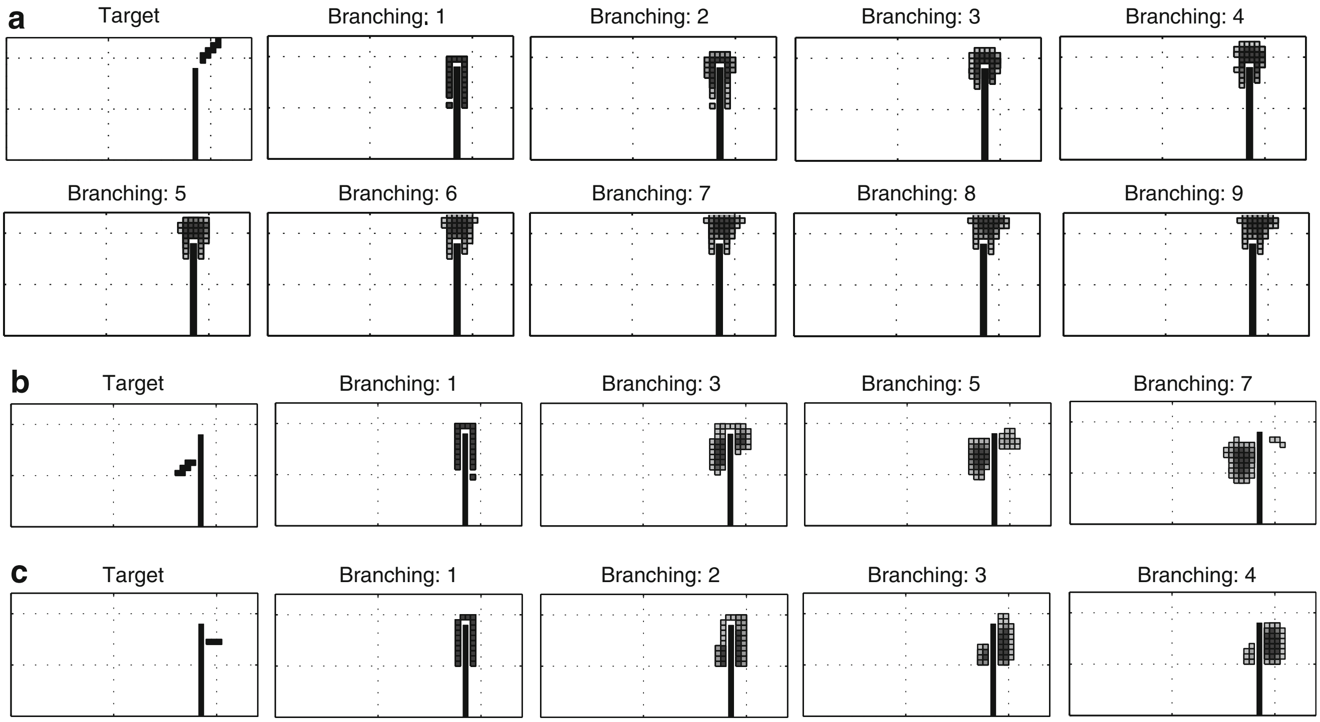


Fig. 9.3 Histogram-like damage map showing the basic process for locating probable damage region: (**a**) damage detection for crack path 1, (**b**) damage detection for crack path 2, (**c**) damage detection for crack path 3

Table 9.1 Modal frequencies of the damaged and undamaged plates

Mode	No. crack (Hz)	Crack 1 (Hz)	Crack 1 Diff (%)	Crack 2 (Hz)	Crack 2 Diff (%)	Crack 3 (Hz)	Crack 3 Diff (%)
1	26.88	26.52	−1.34	26.82	−0.22	26.86	−0.05
2	41.76	39.89	−4.48	41.60	−0.39	41.69	−0.18
3	72.87	70.89	−2.72	72.02	−0.18	72.82	−0.08
4	94.93	89.49	−5.73	94.88	−0.05	94.85	−0.09
5	128.95	126.93	−1.57	128.64	−0.24	128.93	−0.02

the area is distributed in the transverse direction (i.e. along the shorter side of the plate). Again, the probable damage area converges very quickly with very few branching steps.

Damage detection of individual cracks with multiple stiffeners: The model-updating procedure is applied to detect damages in a plate with two welded stiffeners as shown in Fig. 9.4a. Similar to the previous cases, the three crack paths as shown in Fig. 9.3 are considered individually. However, the initial probable damage zone is expanded to include the weld toes around the two stiffeners and along the longitudinal edges of the plate. Figure 9.4b shows the probable damage areas identified. For the damage detection for crack path 1, areas around the longer stiffener and along the top edge are suspected for possible damage very early in the process. The hypothetical branches grow both from the weld toe and edges and later formed the most probable damage area around the vicinity of crack 1. For the detection of crack 2, the model-updating algorithm is not able to pin-point the damage area with high probability. Instead, the detection result suggests two damage prone areas as shown in Fig. 9.4b. Nevertheless, the area around crack 2 is included in one of the probable damage areas. For detecting crack path 3, the model-updating procedure identifies the damage area successfully with high probability, as noted with darker color elements shown in Fig. 9.4b. All in all, the damage detection results show the ability of the probabilistic model-updating algorithm in detecting cracks on a plate with multiple stiffeners and the rapid convergence of the method on identifying the damage areas even when the initial probable damage areas includes relatively widespread regions around the stiffeners and along the edges of the plate.

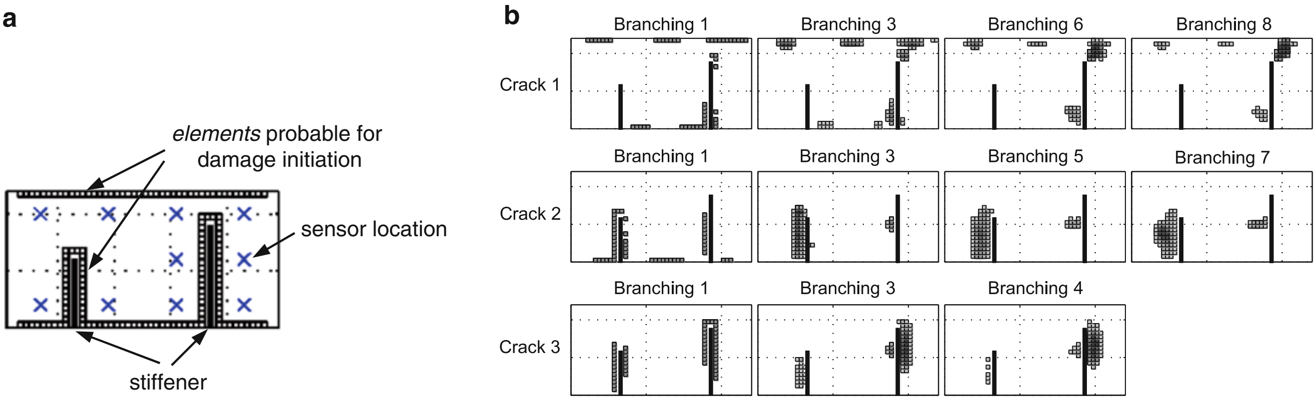


Fig. 9.4 Damage detection of cracks on a plate with multiple welded stiffeners: (**a**) a baseline structure with two stiffeners, (**b**) damage detection results with individual cracks and two stiffeners

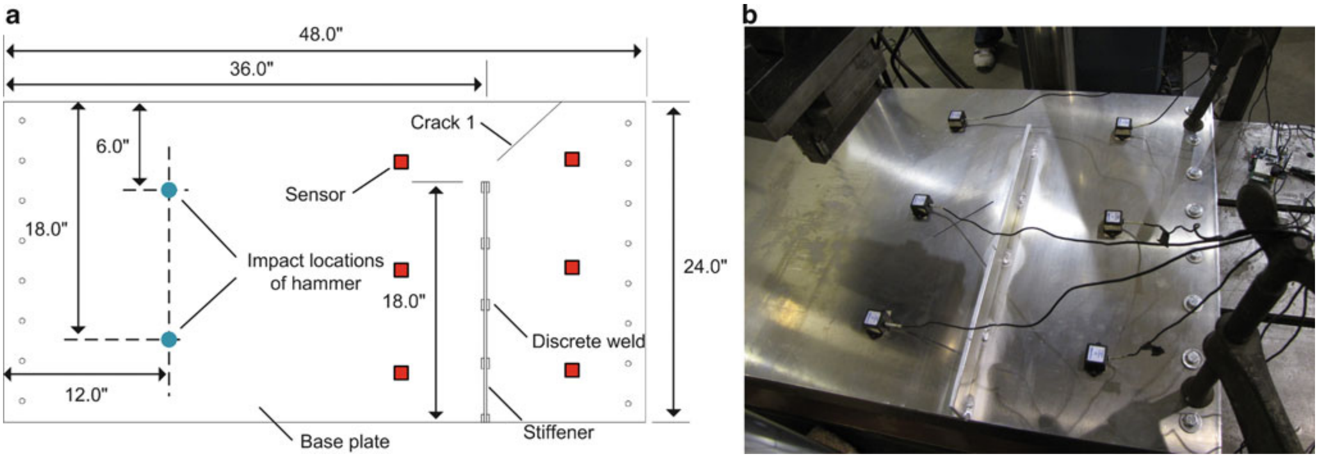


Fig. 9.5 Setup for experimental tests on plate specimens with stiffener and crack: (**a**) schematics of the experimental plate specimens, (**b**) specimen with crack and accelerometers

9.3.2 *Experimental Simulation*

A number of plate specimens have been fabricated in full-scale (according to the schematic of the testbed structure shown in Fig. 9.2a) to test the applicability of the probabilistic model-updating procedure for crack damage detection. The experimental setup is shown in Fig. 9.5. The specimens are made of marine grade aluminum alloy (alloy 5086). As shown in Fig. 9.5a, an 18 in. long stiffener plate has been welded to each of the tested plates with 0.625 in. long discrete tungsten inert gas (TIG) welds at five locations (with a spacing of 4.5 in.) to avoid excessive distortions that might result from weld heat. The base plates have been pre-heated using a gas torch to expedite the welding process. As shown in Fig. 9.5b, the base plates are fixed along the shorter edges using 8–3/8 in. stainless screws with round aluminum washers; furthermore, the assemblage is rigidly connected to a large steel beam post-tensioned to the concrete strong floor. Figure 9.5b shows a diagonally cut (resembling crack path 2 as shown in Fig. 9.3b) adjacent to the stiffener. As shown in the figure, MEMS-based accelerometers (Crossbow CXL02LF1Z) are used to measure the vibration of the specimen during the tests. Furthermore, the Narada wireless system developed at the University of Michigan, which has been successfully implemented for the monitoring of civil infrastructures, is deployed as the primary data acquisition system for collecting the acceleration data [16].

Modal properties extraction and calibration: Both impact hammer tests and hand tapping tests were applied on each plate specimen. For each plate specimen, albeit the undamaged plate or the plates with crack, the hammer test is applied six times with impact locations away from the weld and crack damage areas as shown in Fig. 9.5a. Additionally, each plate specimen

Table 9.2 Summary of modal properties extracted for specimen with crack path 1

	Impact hammer test			Hand-tapping test		
	Modal frequency		Mode shape	Modal frequency		Mode shape
Mode	Test (Hz)	Updated FEM (Hz)	MAC	Test (Hz)	Updated FEM (Hz)	MAC
1	24.22	25.17	0.998	24.22	25.17	0.999
2	35.89	37.89	0.983	35.94	37.89	0.996
3	65.43	67.36	0.992	65.43	67.36	0.985
4	85.57	85.14	0.982	85.64	85.14	0.976
5	122.07	120.43	0.932	121.97	120.43	0.935

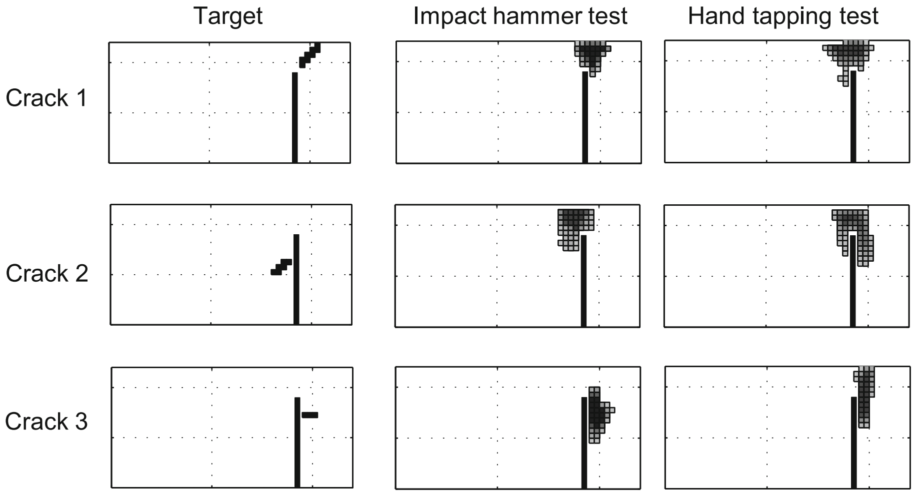


Fig. 9.6 Damage identification results from experimental test specimens

is hand tapped five times at random locations to simulate broad-band ambient excitation. The frequency domain decomposition (FDD) technique is employed to extract the modal properties of the plate specimens using the measured acceleration data. The FDD technique is widely used for modal parameter estimation and is based on the classical complex mode identification function [17].

For the calibration of the finite element model, the modal properties of the undamaged plate specimen (without cracks) are extracted from the vibration tests. Slight differences in modal frequency between the experimental results and the original FE model (used in the numerical simulation) has been observed. Such a difference is expected because of the variation in the idealized finite element model, the material properties in the plate specimens, and boundary conditions in the experimental setup. To better calibrate between the numerical model and the experimental specimens, the elastic (Young's) modulus used in the finite element model is reduced to 90% of the nominal values (10,300 ksi) so that the first mode frequencies of the undamaged plate from both the finite element model and the experimental tests are closely matched. Table 9.2 summarizes the modal frequencies of the plate structure with crack path 1 obtained from the experimental hammer and hand tapping tests as well as from the updated finite element model. It can be seen that the results match very well with the values of modal assurance criteria (MAC) well above 90%.

Damage identification results: The model-updating procedure is applied to identify the damage area (cracks) based on the modal properties extracted from the plate specimens. Figure 9.6 shows the damage identification results for the three crack paths from both impact hammer and hand tapping tests. It can be seen that the model-updating procedure successfully detects crack paths 1 and 3. However, for crack path 2, although the most probable damage area is identified near the damage area, the procedure fails to pin-point the actual crack location. It can also be observed that, in general, the damage identification results based on impact hammer tests are more precise, as illustrated from the darker colors shown in the damage maps. It should be cautioned that, as in any experimental tests, the fabrication qualities of the plate specimens and the welds do vary. The modal frequencies could vary up to 5–7%. Further study on the impact of modeling uncertainties is needed. Preliminary investigation into the factors of noise to the probabilistic model-updating procedure is currently underway.

9.4 Summary and Discussion

In this paper, a model-based monitoring framework for the detection of fatigue-related crack damage in plate-type structures commonly seen in aluminum ship hulls is presented. The monitoring framework involves vibration-based damage detection methodologies and finite element modeling of continuum plate structures. Modal properties derived from sensor information are employed to update the finite element models. A Bayesian-based approach is adopted for model updating. Inherent to the finite element modeling of a continuum structure is the large number of elements and degrees-of-freedom that are required to properly model the structure. To examine all possible combinations of elements to identify potential damage locations is computationally infeasible. To reduce the computational efforts involved in searching for possible damage locations, initial knowledge of probable damage areas and a heuristic-based branch-and-bound scheme are included in the Bayesian-based model-updating framework. Numerical simulations and experimental tests have been conducted to illustrate the model-based monitoring framework and to validate the potential application of the probabilistic-based damage detection approach. Preliminary numerical and experimental results for the plate specimens studied indicate that the proposed procedure is able to successfully identify crack damage areas, especially when the size of the crack is relatively large. However, further study is needed to enhance the damage detection algorithm for accurately locating the exact location of small cracks. Furthermore, modeling uncertainties, sensor noise, variations in loading excitations, and their effects on the proposed damage detection method are interesting subjects of future research.

Acknowledgments The authors would like to gratefully acknowledge the support offered by the Office of Naval Research under Contract Numbers N00014-09-1-0567 and N00014-10-1-0613 awarded to University of Michigan and N00014-10-1-0384 awarded to Stanford University. The advice and suggestions offered by Dr. Paul Hess are also gratefully acknowledged.

References

1. Sohn H, Law KH (2002) A Bayesian probabilistic approach for structure damage detection. Earthquake Eng Struct Dyn 26:1259–1281
2. Chakraverty S (2008) Vibration of plates. CRC Press/Taylor & Francis, Boca Raton
3. Doebling SW, Farrar CR, Prime MB (1998) A summary review of vibration-based damage identification methods. Shock Vib Dig 30:91–105
4. Pandey AK, Biswas M, Samman MM (1991) Damage detection from changes in curvature mode shapes. J Sound Vib 145:321–333
5. Pandey AK, Biswas M (1995) Experimental verification of flexibility difference method for locating damage in structures. J Sound Vib 184(2):311–328
6. Bernal D (2006) Flexibility-based damage localization from stochastic realization results. J Eng Mech 132(6):651–658
7. Zonta D, Bernal D (2006) Strain-based approaches to damage localization in civil structures. In: Proceedings of XXIV international modal analysis conference, Saint Louis
8. Kurata M, Kim J-H, Lynch JP, Law KH, Salvino LW (2010) A probabilistic model updating algorithm for fatigue damage detection in aluminum hull structures. In: ASME 2010 conference on smart materials, Adaptive structures and intelligent systems, Philadelphia
9. Gao Y, Spencer BF Jr (2002) Damage localization under ambient vibration using changes in flexibility. Earthquake Eng Eng Vib 1(1):136–144
10. Brinker R, Andersen P (2002) A way of getting scaled mode shapes in output-only modal testing. In: Proceedings of 21st modal analysis conference (IMAC XXI), Orlando
11. Parloo E, Verboven P, Cuillame P, Overmeire MV (2002) Sensitivity-based operational mode shape normalization. Mech Struct Signal Process 16(5):757–767
12. Cheung SH, Beck JL (2009) Bayesian model updating using hybrid Mote Carlo simulation with application to structural dynamic models with many uncertain parameters. J Eng Mech-ASCE 135(4):243–255
13. Stull CJ, Earls CJ, Koutsourelakis P-S (2009) Model-based structure health monitoring using parallel stochastic search methods to enable inverse solutions. In: Proceedings of the 7th international workshop on structural health monitoring, Stanford University, Stanford, pp 1959–1969
14. Nichols JM, Link WA, Murphy KD, Olson CC, Bucholtz F, Mchalowicz JV (2009) A Bayesian approach to identifying and tracking damage in structures. In: Proceedings of the 7th international workshop on structural health monitoring, Stanford University, Stanford, pp 1951–1958
15. Norkin VI, Pflug GCh, Ruszczynski A (1998) A branch and bound method for stochastic global optimization. Math Program Ser A and B 83(3), 425–450
16. Swartz RA, Lynch JP (2009) Strategic network utilization in a wireless structural control system for seismically excited structures. J Struc Eng 135(5):597–608
17. Brinker R, Zhang L, Andersen P (2001) Modal identification of output-only systems using frequency domain decomposition. Smart Mater Struct 10(3):441–445

Chapter 10
On the Legitimacy of Model Calibration in Structural Dynamics

François M. Hemez and Christopher J. Stull

Abstract In structural dynamics, a finite element model is often calibrated to better reproduce experimental measurements collected from the structure. When the agreement between measurements and predictions is unsatisfactory, the model is parameterized and calibrated to improve its overall accuracy. We argue that model calibration may not always be legitimate to improve test-analysis correlation, because a calibration study attempts to compensate for parametric errors when, in fact, the disagreement between measurements and predictions may originate from other sources (e.g. discretization error). In this work, a scaled model of a three-story frame structure that responds mostly in bending is tested experimentally and modeled with finite elements. The agreement between measurements and predictions is assessed relative to the overall level of experimental variability. Truncation error is quantified by performing mesh refinement studies. Guidance on the legitimacy of model calibration is formulated by comparing the overall levels of truncation error, parametric uncertainty, and experimental variability. It is concluded that, while useful, model calibration is a technique that should be deployed only after other sources of modeling error have been rigorously quantified and adjusted for. *(Publication approved for unlimited, public release on September 29, 2011, LA-UR-11-5608, Unclassified.)*

Keywords Model calibration • Mesh discretization • Sensitivity analysis • Truncation error

10.1 Introduction

In structural dynamics, a Finite Element (FE) model is often calibrated to better reproduce experimental measurements collected from the structure. Features commonly used for test-analysis correlation are the resonant frequencies, mode shapes, and modal damping ratios of the linear vibration response. When the agreement between experimental measurements and numerical predictions is unsatisfactory, the model can be parameterized and calibrated to improve its accuracy. Model-updating techniques have been developed for and successfully applied to linear [1, 2] and nonlinear dynamics problems [3].

FE model-updating techniques and more broadly, calibration methods offer the advantage that the optimization of parameters of a numerical simulation leads to predictions that better reproduce the experimental measurements. This is appealing to analysts who must demonstrate that their numerical simulations are appropriate representations of a reality of interest. Matching experimental observations also appeals to decision-makers who need to rely on numerical simulations to guide their judgments. Model calibration, however, may lead to a false sense of confidence, because there are

---

F.M. Hemez (✉)
Los Alamos National Laboratory, XTD-3, Mail Stop T087, Los Alamos 87545, NM, USA
e-mail: hemez@lanl.gov

C.J. Stull
The Engineering Institute, Los Alamos, USA

Los Alamos National Laboratory, INST-OFF, Mail Stop T001, Los Alamos 87545, NM, USA
e-mail: stull@lanl.gov

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_10,

fundamental trade-offs between: (1) the fidelity-to-data, or ability of predictions to match measurements; (2) the robustness to lack-of-knowledge, or influence on predictions of assumptions made during the model development; and (3) the predictive power, or consistency of predictions provided by a family of equally-robust models [4]. The conjecture that we wish to explore in this work is that calibration may lead to accurate predictions for the wrong reasons. It is emphasized that the previous statement is *not* a criticism of calibration; the authors rely on model calibration in their everyday work and understand its strengths and weaknesses. Our intent is instead to draw attention to potential abuses and suggest methods that can be implemented to guarantee that the recourse to calibration is legitimate.

There are many reasons why calibrating a model may not be appropriate to improve fidelity-to-data, or the ability of predictions to match experimental observations. The main four candidates are: (1) the lack of resolution in the simulation; (2) an inappropriate mathematical idealization; (3) an ineffective parameterization; and (4) the lack of prediction sensitivity. This work focuses exclusively on demonstrating the effect that a potential lack-of-resolution may exercise, first, on the global sensitivity of a prediction to parameter changes and, second, on the calibration. The lack of resolution aspect refers to the fact that the discretization of a FE model may not provide enough resolution to yield "*asymptotically converging*" discrete solutions. The immediate consequence of inappropriate convergence is that truncation effects yield large numerical uncertainty in the predictions. The danger that we wish to expose is that model calibration may attempt to reduce the discrepancy between numerical predictions and experimental measurements by adjusting parameters when this discrepancy may instead be due (in part) to truncation effects. The second aspect (idealization) refers to what is commonly known as "*modeling error*" or the selection of an inappropriate mathematical model to represent the reality of interest. An example would be the simulation of deformation of a wall structure, due to Earthquake vibration, using membrane elements that carry in-plane stresses only when the actual response involves shear stresses.

A parameterization may be ineffective (third aspect) because it fails to adequately compensate for the discrepancy between predictions and measurements. If this happens, then adjusting the parameters may not translate into a reduction of discrepancy. An example would be attempting to correct a mistake in the moments of inertia of a non-symmetric, slender structure by adjusting its modulus of elasticity. An ineffective choice of parameters may originate from a mathematical idealization that does not define the "*correct*" parameters, therefore, leaving the analyst with no choice but to calibrate parameters that cannot compensate for the discrepancy. It may also be that the parameterization appropriately represents the reality of interest but the selection made for calibration is incorrect. The effectiveness of the parameterization may, or not, be coupled to a lack of sensitivity of predictions to changes in parameter values (fourth aspect). A lack-of-sensitivity translates into the inability to influence the predictions by varying model parameters. If this happens, then numerical ill-conditioning is likely to be generated which, if not addressed using some sort of pre-conditioning or regularization, further accentuates the problem.

The manuscript is organized as follows. The frame structure analyzed to illustrate our conjecture is described, together with the results of vibration testing, in Sect. 10.2. The FE model developed for this study is explained in Sect. 10.3. Section 10.4 briefly summarizes the quantification of solution uncertainty caused by truncation effects, and the results of our application to the three-story frame structure are presented in Sect. 10.5.

10.2 The Three-Story Frame Structure and its Vibration Testing

The application presented in this study is a three-story frame designed to mimic conditions encountered when column-to-floor connections of a building are subjected to vibrations, such as those generated during an earthquake. Section 10.2.1 briefly describes the structure and the experimental campaign of vibration testing is summarized in Sect. 10.2.2.

10.2.1 The Three-Story Frame Structure

The three-story frame designed and tested at Los Alamos National Laboratory (LANL) is illustrated in Fig. 10.1. The structure is designed to represent a scaled model of a shear building that can be used in structural health monitoring research [5], as well as to generate datasets for model validation activities [6]. The structure consists of aluminum columns and plates assembled using bolted joints. Each floor measures 30.5 × 30.5 × 2.5 cm, and is connected to the floors above and/or below by four aluminum columns measuring 17.7 × 2.5 × 0.6 cm. The frame slides on rails that allow movement in one direction which coincides with the weak bending direction of columns. A center column measuring 15.0 × 2.5 × 2.5 cm is suspended from the third floor, as shown in the left panel of Fig. 10.1, and a contact mechanism is attached to the second floor (see Fig. 10.1, right).

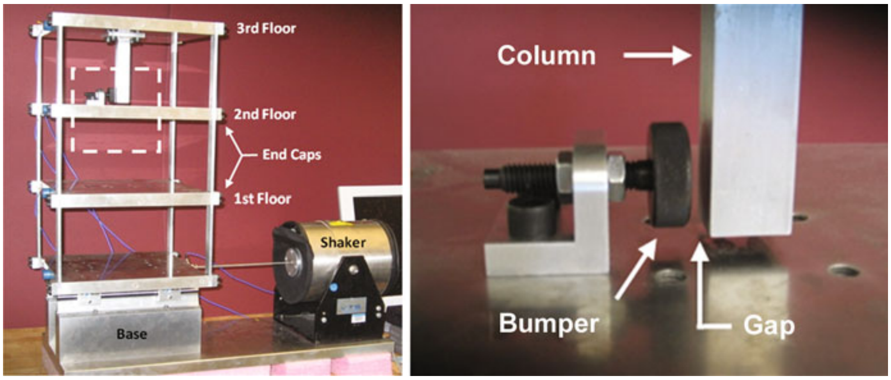


Fig. 10.1 The LANL three-story frame structure (*left*) and its contact mechanism (*right*)

Nonlinearity may be introduced into the system by adjusting the amplitude of the oscillations and/or the gap between the center column and back-stop, also referred to as the "bumper," such that contact occurs between these two components. The position of the bumper can be adjusted to vary the severity of impacts that occur at a given level of excitation. As explained in Reference [5], this particular design is motivated by several real-world examples that have been published, which report opening and closing of bending and shear cracks under dynamic loading conditions such as those encountered during an earthquake event.

10.2.2 *Vibration Testing and Identification of the Linear Response*

An experimental campaign of vibration testing is performed to acquire datasets for different configurations of the three-story frame. Four factors are varied: (1) the type of input excitation, (2) magnitude of input force excitation, (3) activation of the nonlinear contact mechanism (on/off), and (4) level of clearance between the center column and the back-stop. The combinations of these four factors result in a total of 42 datasets, without accounting for the replicates performed for each configuration tested. Because we are interested in correlating the modal response of a linear FE model, only the non-contacting response is investigated in this study.

The input force is applied by an electromagnetic shaker connected to the base floor. The structure and shaker are mounted together on the base floor and the entire system rests on rigid foam, where the foam is intended to minimize extraneous sources of unmeasured excitation from being introduced through the base. A force transducer with nominal sensitivity of 2.2 mV/N is attached to the end of the stinger to measure the input force from the shaker.

Time-history responses are measured for each type of excitation signal. These signals include random, chirp, harmonic, and transient excitations. A total of four accelerometers with nominal sensitivities of 1,000 mV/g are attached at the centerline of each horizontal floor on the opposite side from the excitation side. A Dactron Spectrabook fast Fourier transform analyzer is used for data collection and analysis. The output channel that provides the excitation signal is connected to a Techron 5530 power amplifier that drives the shaker. For the case of random excitation analyzed here, the analog signals are discretized with 4,096 time-history points at the sampling rate of 640 Hz. A band-limited random excitation in the range of 20–200 Hz is used to define the input signals, so as to avoid exciting the rigid body modes present below 20 Hz. The Root Mean Square (RMS) excitation level is varied from 1.5 V RMS to 2.5 V RMS to perform linearity checks that turn out conclusive.

For experimental modal analysis, the Frequency Response Function (FRF) curves are defined with 1,024 samples, 50% overlap, and a Hanning window. Ten replicates of time-history data are collected to perform the averaging. FRF curve-fitting is carried out in the frequency domain within the 20–80 Hz bandwidth. Figure 10.2 compares the mode shapes of the first three bending modes identified using different configurations where the type and magnitude of input excitation signals are varied. These comparisons, combined with the linearity check, indicate that the extraction of modal properties is high-quality.

The left side of Table 10.1 lists the identified resonant frequencies and modal damping ratios that result from the experimental modal analysis, where only the first three bending modes are considered in this study. The overall level of experimental variability is then quantified in the right side of Table 10.1. Variability is assessed by replicating the modal tests and varying the configuration tested with different types of excitation signals and different levels of excitation magnitudes. These values serve as a "*reference*" for the test-analysis correlation of numerical models discussed in Sect. 10.3.

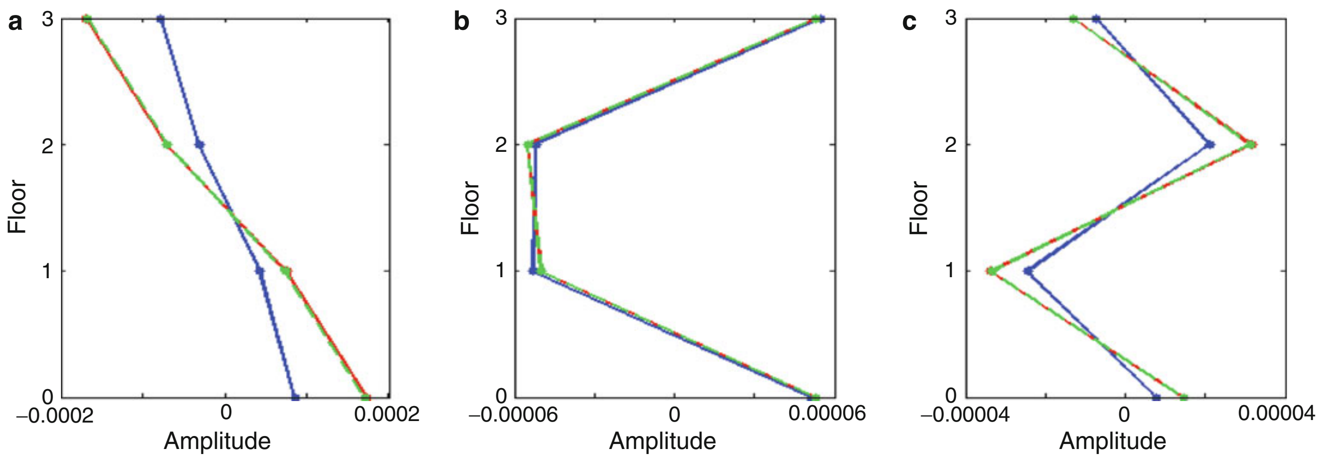


Fig. 10.2 Mode shape deflections identified from different testing configurations. (**a**) First bending. (**b**) Second bending. (**c**) Third bending

Table 10.1 Identified bending frequencies and damping ratios of the three-story frame

Mode	Frequency (Hz)	Damping ratio (%)	Frequency standard deviation[a] (σ, Hz)	Frequency coefficient of variance[a] (σ/μ, %)
2	29.40	3.26	0.30	1.00
3	54.25	1.07	0.45	0.82
4	71.34	0.73	0.20	0.28

[a] Statistics based on testing different configurations, with ten replicates each time

10.3 Computational Models of the Three-Story Frame Structure

This section describes the model used to simulate vibrations of the frame structure. The rationale of model development is introduced in Sect. 10.3.1 by discussing a one-dimensional, shear-building model. In Sect. 10.3.2, the two-dimensional model, that is the focus of this investigation, is discussed.

10.3.1 One-Dimensional Shear Model for Linear or Nonlinear Vibrations

Irrespective of the type of modeling implemented, the equation of motion is described in a matrix notation using the master FE mass (M), stiffness (K), and damping (C) matrices as:

$$\mathrm{M}\,\frac{\mathrm{d}^2\mathrm{x(t)}}{\mathrm{dt}^2}+\mathrm{C}\frac{\mathrm{dx(t)}}{\mathrm{dt}}+\mathrm{K\ x(t)}\ +\mathrm{F}_{\mathrm{NL}}(\mathrm{t})=\mathrm{F}_{\mathrm{EXT}}(\mathrm{t}), \tag{10.1}$$

where the vectors x(t), $\mathrm{F}_{\mathrm{EXT}}(\mathrm{t})$, and $\mathrm{F}_{\mathrm{NL}}(\mathrm{t})$ denote the displacement solution, external (or applied) forces, and nonlinear forces, respectively. To simulate time-domain datasets, the displacement, velocity, and acceleration solutions of Eq. 10.1 are obtained with numerical integration using an integration scheme such as the Newmark or Runge–Kutta method. Proportional damping is assumed, which implies that the master matrix can be defined as $\mathrm{C} = \alpha\mathrm{M} + \beta\mathrm{K}$ where the two coefficients (α; β) are related to the *k*th-mode resonant frequency ω_k and damping ratio ξ_k as:

$$\zeta_\mathrm{k}=\frac{1}{2}\left(\frac{\alpha}{\omega_\mathrm{k}}+\beta\ \omega_\mathrm{k}\right). \tag{10.2}$$

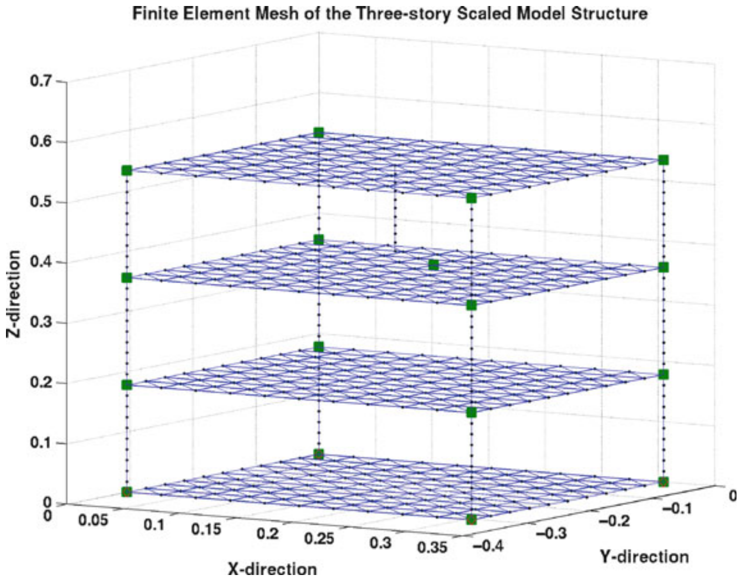


Fig. 10.3 Two-dimensional model discretized with $N_{FE} = 10$ elements per component

Equation 10.2 shows that coefficients (α; β) can be best-fitted given a minimum of two identified modes. Finally, it is noted that $F_{NL} = 0$ in Eq. 10.1 whenever the linear system is simulated.

The one-dimensional model represents the three-story frame as a shear building discretized with four degrees-of-freedom (DOF). Masses are lumped (or concentrated) while the stiffness coefficients are obtained by adding the bending stiffness values of the four columns treated as one-dimensional beams. The base-to-first-floor stiffness coefficient is assigned a low-magnitude value because friction between the support rails and frame structure is negligible. The structure also includes a center column suspended from the 3rd floor and a contact mechanism attached to the 2nd floor (see Fig. 10.1, left). Masses of the center beam and back-stop are added to the appropriate DOF of the one-dimensional model. Even though this model is rather simplistic, it provides accurate predictions of the first three bending modes of the three-story frame structure (see Reference [6]). It is used as starting point to develop the two-dimensional model.

10.3.2 Two-Dimensional Finite Element Model for Modal Analysis

The numerical model employed in this study is composed of linear, Euler-Bernouilli beams to describe the bending of vertical columns and triangular shell elements to describe the membrane behavior of the floors. Although the geometry of the three-story frame is represented with fidelity, this model is referred to as "*two-dimensional*" because of the nature of elements used in the mathematical idealization.

Figure 10.3 illustrates a computational mesh composed of ten beam elements per column and, likewise, ten triangular shell elements per side of a floor. Green square symbols denote the added point-masses at column-to-floor connections and the back-stop mass on the second floor. For modal analysis, eigenmodes are extracted from the master mass and stiffness matrices. Boundary condition springs are included in the model to account for the potential flexibility of the support rails. This capability, however, is not exercised here because the six DOFs of boundary nodes are kept clamped.

Figure 10.4 illustrates the mode shapes (w/ modal frequencies) associated with the first three bending modes as predicted by the "*nominal*," two-dimensional FE model. This model implements the mass and stiffness coefficients proven acceptable by the one-dimensional shear-building model. It also discretizes the geometry with a seed of $N_{FE} = 20$ elements. Each column is discretized by 20 beam elements and, likewise, each side of a horizontal floor is discretized by 20 triangular membrane elements, for a total of 12,048 elements.

Direct comparison between Table 10.1 (identified frequencies) and the predicted frequencies given in Fig. 10.4 indicates that the initial test-analysis correlation of the nominal model is acceptable. Frequency errors for the first three bending modes are in the order of 10%. These large errors, relative to the overall level of experimental variability (see Table 10.1), would normally justify the recourse to parameter calibration to improve the test-analysis correlation. Before calibrating the model, the numerical uncertainty is first quantified in Sect. 10.4.

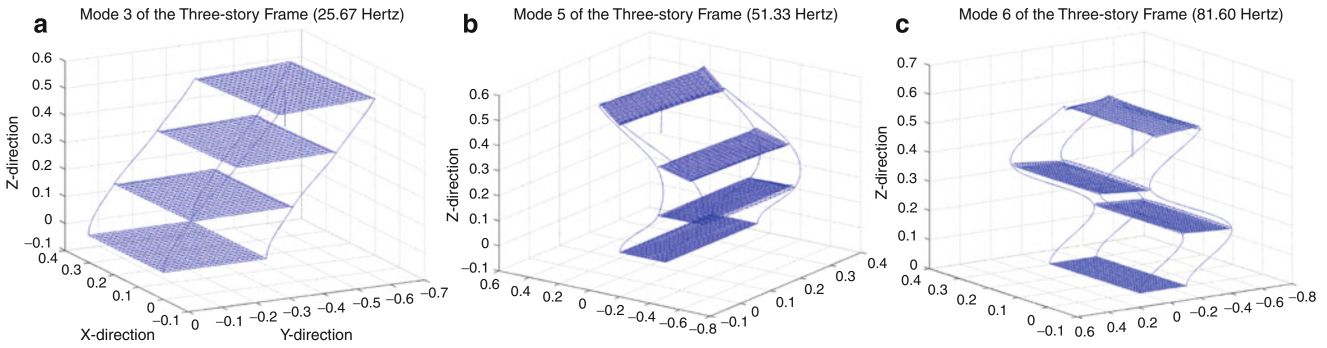


Fig. 10.4 First six mode shape deflections of the nominal, two-dimensional FE model. (**a**) 1st bending (25.67 Hz) (**b**) 2nd bending (51.33 Hz) (**c**) 3rd bending (81.60 Hz)

10.4 Asymptotic Convergence of Bending Frequency Predictions

This section discusses numerical uncertainty. The theory, briefly summarized in Sect. 10.4.1, is illustrated with two mesh refinement studies: one where the predictions are converging asymptotically and one where this is not the case. In Sect. 10.4.2, mesh refinement is applied to only the vertical columns of the model while the discretization of horizontal floors remains unchanged. The second study, where the vertical beams *and* horizontal columns are refined simultaneously, is discussed in Sect. 10.4.3.

10.4.1 Brief Overview of Numerical Uncertainty

It is well known that the discrete solutions obtained by discretizing the equations of motion, or conservation laws, of a physical system do not converge to the exact solution of the Partial Differential Equations (PDE). For example, y^{Exact} is the exact solution of the following equation in the case of a simple, one-dimensional conservation law:

$$\frac{\partial y^{\text{Exact}}}{\partial t} + \frac{\partial F(y^{\text{Exact}})}{\partial x} = S(x;t). \tag{10.3}$$

Equation 10.3 expresses that the rate-of-change and the flux of the solution are in equilibrium with the source on the right-hand side. When geometry, boundary conditions, or nonlinearity prevent the resolution in closed-form, a numerical solution can be obtained by discretizing equation (10.3) using, for example, finite elements. As the level of resolution in the calculation is increased, denoted as "$\Delta x \rightarrow 0$" for short, the discrete solution $y(\Delta x)$ converges to the exact solution $\hat{y}$ of the *modified equation* [7]. The modified equation is an infinite Taylor series-like expansion that explicitly identifies the truncation error introduced by the particular choice of numerical method implemented. For example:

$$\underbrace{\frac{\partial \hat{y}}{\partial t} + \frac{\partial F(\hat{y})}{\partial x} = S(x;t)}_{\text{OriginalEquation}} + \underbrace{\frac{\partial^2 \hat{y}}{\partial t^2}\Delta t + \frac{1}{2}\frac{\partial \hat{y}}{\partial x}\Delta x^2 + \frac{1}{6}\frac{\partial^2 \hat{y}}{\partial x^2}\frac{\partial \hat{y}}{\partial t}\Delta t \Delta x^3 + \ldots}_{\text{TruncationError}}. \tag{10.4}$$

Deriving a modified equation analytically, such as suggested by Eq. 10.4, is possible only in a few cases that combine simple PDE, geometry, boundary conditions, and numerical method. Modified equations are generally inaccessible in solid mechanics or structural dynamics where the FE method is applied. When the exact functional form of truncation is unknown, the best that one can hope for is to obtain an upper bound of solution error, simply written as:

$$\left| \frac{y^{\text{Exact}} - y(\Delta x)}{y(\Delta x)} \right| \leq U(\Delta x), \tag{10.5}$$

Table 10.2 Definition of the column-only refinement of the two-dimensional FE model

	FE discretization		Statistics of meshes		Predicted bending frequencies		
Mesh number	Column discretization	Number of nodes	Number of elements	Number of DOFs	Mode-3 frequency (Hz)	Mode-5 frequency (Hz)	Mode-6 frequency (Hz)
(I)	5	1,813	3,265	10,878	25.67	51.33	81.60
(II)	10	1,878	3,330	11,268	25.68	51.35	81.62
(III)	20	2,008	3,460	12,048	25.68	51.35	81.62
(IV)	40	2,268	3,720	13,608	25.68	51.35	81.62
(V)	80	2,788	4,240	16,728	25.68	51.35	81.62
(VI)	160	3,828	5,280	22,968	25.68	51.35	81.62

where $y(\Delta x)$ denotes a prediction, such as a resonant frequency or peak stress, obtained at the level of resolution Δx and $U(\Delta x)$ is the corresponding upper bound of solution error.

When the exact solution y^{Exact} is unknown, one can no longer talk of "*error*." The difference in the left-hand side of Eq. 10.5, $|y^{Exact} - y(\Delta x)|$, becomes an uncertainty caused by truncation effects. This uncertainty can be estimated with a formalism that derives from modified equation analysis [7], postulating that the difference between exact and discrete solutions is defined as:

$$y^{Exact} \approx y(\Delta x) + \beta \cdot \Delta x^{p}, \tag{10.6}$$

where β is a pre-factor coefficient and the exponent p denotes the rate-of-convergence of the numerical method. Equation 10.6 is a simplistic approximation of the modified equation where the higher-order terms are ignored, the pre-factor β is assumed to remain constant, convergence is monotonic as $\Delta x \to 0$, and the analysis is restricted to scalar-valued quantities. Using mesh refinement where the same problem is solved several times with different levels of resolution, References [8, 9] independently arrive at a similar upper bound:

$$U(\Delta x) = F_S \cdot \frac{|y(R \cdot \Delta x) - y(\Delta x)|}{(R^{p} - 1)}, \tag{10.7}$$

where Δx is a characteristic mesh size, R is the refinement ratio, and F_S denotes a user-defined safety factor. By definition, the refinement ratio relates the coarse and fine levels of resolution, that is, $\Delta x_C = R \cdot \Delta x$ and $\Delta x_F = \Delta x$, respectively. The safety factor of the Grid Convergence Index (GCI) proposed in Reference [8] is a user-defined value within $1 \le F_S \le 3$. Reference [9] further shows that this scaling is not necessary to arrive at an upper bound, which means that $F_S = 1$ can be used in Eq. 10.7 without loss of generality.

10.4.2 Mesh Refinement Applied to the Vertical Columns Only

The first mesh convergence study refines the vertical columns (beam elements), while the discretization of horizontal floors (shell elements) is kept constant. Table 10.2 defines properties of the six meshes analyzed. Since the refinement is limited to beam elements, the number of DOFs does not increase significantly with higher resolution. Resonant frequencies of the first three bending modes are listed in Table 10.2 for each of the meshes. It can be observed that the predictions "*stabilize*" as the level of resolution is increased. However, without a *quantitative* analysis, this observation remains pointless.

Because the refinement ratio is not kept constant, numerical optimization is used to solve for the unknown pre-factor and rate-of-convergence (β; p) of Eq. 10.6. These coefficients are listed in Table 10.3 for the three frequencies analyzed. Also tabulated are the extrapolated solutions, which are the best-possible estimates of solutions $\hat{y}$ of the modified equation (10.4) that the code is attempting to converge to as $\Delta x \to 0$.

As expected, the super-convergent, second-order accuracy is recovered in Table 10.3. This is due to the fact that, even though the Euler-Bernouilli element implements linear shape functions, this low-order interpolation suffices to capture well the bending of vertical columns. Attempting to predict, on the other hand, a torsional mode would reduce the accuracy back to first order. The rates-of-convergence are then used to calculate the upper bounds $U(\Delta x)$ listed in Table 10.4.

Table 10.4 indicates that the solution uncertainty is very low when mesh refinement is applied only to the beam elements of the vertical columns. The numerical uncertainty is smaller than 0.02 Hz, even for the coarsest mesh labeled (II), while the

Table 10.3 Asymptotic behavior observed with the column-only refinement

Mode	Pre-factor coefficient, β (Hz/mm^p)	Rate of convergence, p	Extrapolation, $\hat{y}$ (Hz)	Residual MSE (%)
3	5,397.9	2.027	25.68	0.003
5	15,000.8	2.009	51.35	0.000
6	17,621.1	2.002	81.62	0.000

Table 10.4 Upper bounds of uncertainty estimated from the column-only refinement

	Absolute uncertainty, U(Δx)			Percent uncertainty, U(Δx)/y(Δx)		
Mesh	Mode 3 × 10^{-4} (Hz)	Mode 5 × 10^{-4} (Hz)	Mode 6 × 10^{-4} (Hz)	Mode 3 (%)	Mode 5 (%)	Mode 6 (%)
(II)	36.39	116.87	144.14	0.014	0.023	0.018
(III)	9.09	29.01	35.97	0.004	0.006	0.004
(IV)	2.27	7.24	8.99	0.001	0.001	0.001
(V)	0.58	1.81	2.25	0.000	0.000	0.000

Table 10.5 Definition of the entire-mesh refinement of the two-dimensional FE model

	FE discretization		Statistics of meshes	
Mesh number	Floor discretization	Column discretization	Number of elements	Number of DOFs
(I)	5	5	265	1,158
(II)	10	10	930	3,588
(III)	20	20	3,460	12,048
(IV)	40	40	13,320	43,368
(V)	60	60	29,580	93,888
(VI)	80	80	52,240	163,608

overall experimental variability is one order of magnitude greater (see Table 10.1). These results indicate that the predictions of bending frequency are not sensitive to the level of resolution used to discretize the vertical columns, at least when using a minimum of ten elements per column.

10.4.3 Mesh Refinement Applied to the Entire Mesh

The insensitivity of predictions to mesh resolution, observed in Sect. 10.4.2, justifies the recourse to model calibration to adjust the material properties and moments of inertia of beam elements. One should, however, stop short of drawing a similar conclusion when it comes to the shell elements of the horizontal floors since their discretization is unchanged in the first study.

We proceed with another study where the entire mesh is refined uniformly. The difficulty comes from the fact that beam and shell elements exhibit different asymptotic behaviors. While the beam is second-order accurate, as observed previously, the shell is expected to be only first-order accurate. The two behaviors are in competition in the (unknown) modified equation of the problem and it not possible to assess a priori what will be the dominant form of truncation error.

Table 10.5 defines properties of the six meshes analyzed. It is similar to Table 10.2 with the exception that resolution is, here, increased for the entire mesh. It can be seen that refining the triangular shells rapidly becomes expensive because the computational cost of a modal extraction grows approximately as the square of the number of DOFs.

Table 10.6 summarizes the asymptotic behavior of truncation error observed when the *entire* mesh is refined. It is similar to Table 10.3 in the case of beam-only refinement. The triplet (β; p; $\hat{y}$) of unknown coefficients is, as before, estimated through numerical optimization of Eq. 10.6. Although the extrapolated solutions differ from those of Table 10.3 by only a few percentage points, the convergence rates clearly suggest first-order accuracy, indicating that the overall asymptotic behavior of the calculation is dominated by the truncation error of shell elements.

For the analysis of meshes (II) to (V) in Table 10.5, the first-order accuracy of the two-dimensional model leads to 25–36% prediction uncertainty for mode 3, 5–9% uncertainty for mode 5, and 4–5% uncertainty for mode 6. These percentages of solution uncertainty yield, on average, 7 Hz for mode 3, 3 Hz for mode 5, and 4 Hz for mode 6, which is *larger* than the experimental variability by at least one order of magnitude (see Table 10.1).

Table 10.6 Asymptotic behavior observed with the entire-mesh refinement

Mode	Pre-factor coefficient, β (Hz/mm^p)	Rate of convergence, p	Extrapolation, $\hat{y}$ (Hz)	Residual MSE (%)
3	288.87	1.112	22.34	2.014
5	88.95	1.035	49.93	2.125
6	168.17	1.133	79.86	1.768

Instead of tabulating the upper bounds U(Δx) of solution uncertainty, they are represented as a function of mesh size Δx in Fig. 10.5. The figure also compares the two refinement studies. To emphasize the different asymptotic behavior, the left side of Fig. 10.5 (column-only refinement) represents numerical solutions y(Δx) while the right side (refinement of the entire mesh) depicts solution errors $|\hat{y} - y(\Delta x)|$ relative to the best-guess, extrapolated solutions $\hat{y}$. In both cases, the vertical blue lines indicate the ± U(Δx) bounds of solution uncertainty. Clearly, the refinement of the entire mesh generates more prediction uncertainty than restricting it to the vertical columns.

10.5 Influence of Mesh Resolution on Sensitivities of Model Predictions

Section 10.4 demonstrates that the refinement of shell elements leads to a significant level of prediction uncertainty that is not reduced as resolution increases. This arises from the sub-optimal formulation of the particular shell implemented for this study, where it is noted that implementing an improved formulation could reduce the prediction uncertainty. These results nevertheless beg the question of whether calibration is a legitimate recourse to improve the correlation between experimental measurements and numerical predictions. To answer this question, the sensitivities of bending frequencies to mesh resolution are studied. First, the connection between model calibration and global sensitivity analysis is stressed. Then, the results observed are discussed to shed light on the legitimacy of model calibration.

10.5.1 Calibration of Model Parameters and Sensitivity of Predictions

It is well known that calibration depends on the sensitivity of the response being correlated to the parameters being adjusted. This is because a calibration algorithm generally solves a first-order approximation such as:

$$\omega_k^{\mathrm{Test}} \approx \omega_k(p) + \begin{bmatrix} \frac{\partial \omega_k(p)}{\partial p_1} & \frac{\partial \omega_k(p)}{\partial p_2} & \cdots & \frac{\partial \omega_k(p)}{\partial p_N} \end{bmatrix} \cdot \begin{bmatrix} \Delta p_1 \\ \Delta p_2 \\ \vdots \\ \Delta p_N \end{bmatrix}, \tag{10.8}$$

where corrections $\Delta p = [\Delta p_1; \Delta p_2; \ldots \Delta p_n]$ are brought to the parameters $p = [p_1; p_2; \ldots p_n]$ such that the FE model predicts resonant frequencies that better match the measurement $\omega_k{}^{\mathrm{Test}}$ on the right-hand side. Writing Eq. 10.8 for several modes, or augmenting it with similar equations written for other responses of the structure, leads to a linear system of equations from which the adjustment Δp can be calculated.

Equation 10.8 indicates that an adjustment Δp_n depends on the sensitivities of the prediction to all parameters considered for calibration. Therefore, the calibration results are likely to change if these sensitivities vary relative to one another. The level of mesh resolution is one of the factors that could change these sensitivities, which is the contention that we wish to demonstrate next. In the application to the two-dimensional FE model presented in Sect. 10.3.2, eight parameters are considered for calibration. Table 10.7 defines the nominal values of these parameters, with their ranges of variability. Prior to starting the calibration, we seek to learn the sensitivities of bending frequencies such that only the most influential parameters are adjusted.

Table 10.8 shows the total variability of the first three resonant frequencies predicted by the model when the eight parameters of Table 10.7 vary simultaneously within their respective ranges. The table indicates 24–27% prediction variability, which is more important than the experimental variability of 1%, at most, in Table 10.1. These variance statistics are estimated using a two-level, full-factorial design that requires $2^8 = 256$ runs of the FE model at the nominal mesh resolution.

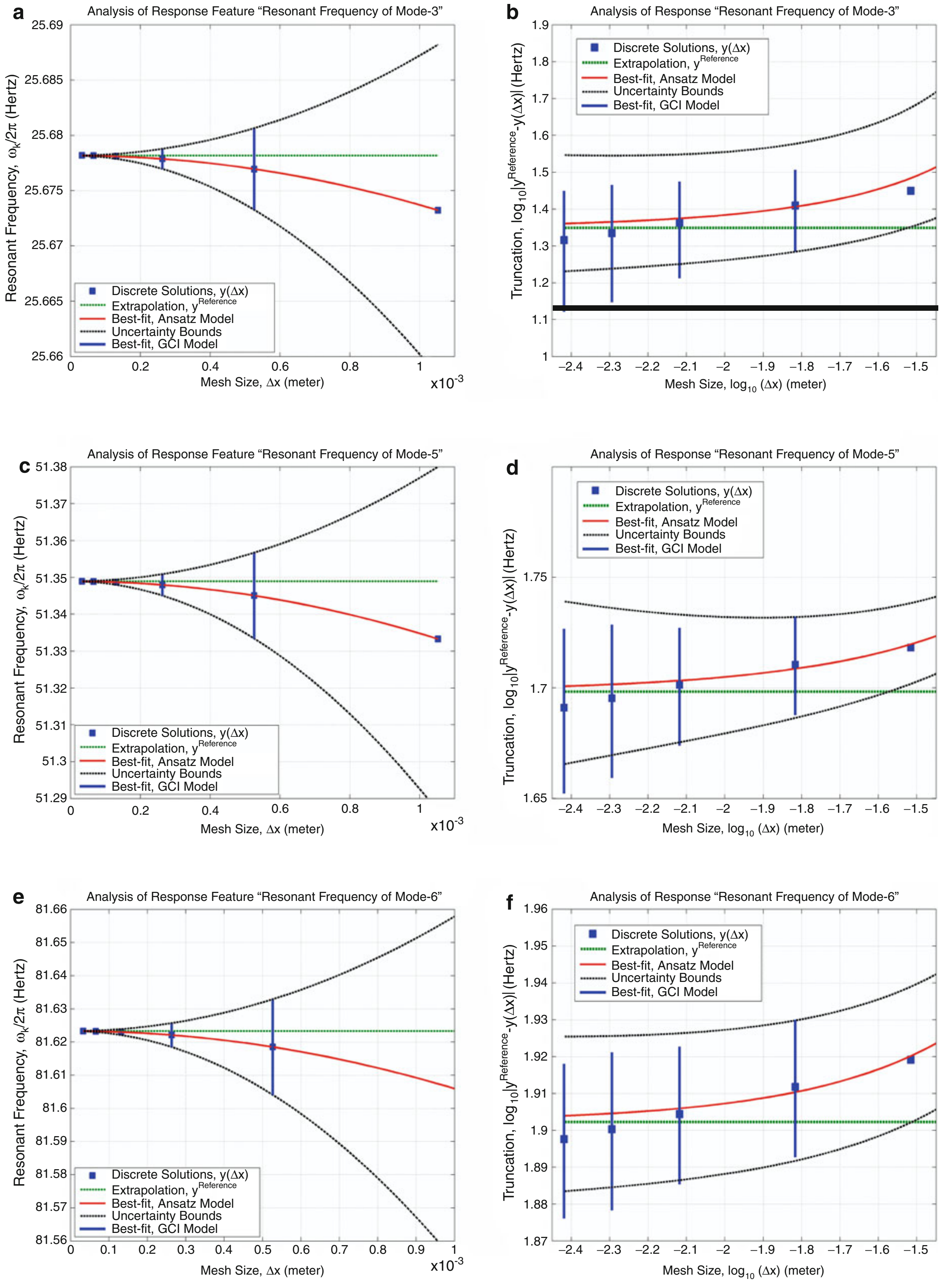


Fig. 10.5 Asymptotic convergence of the first three bending frequencies. (**a**) First bending (column-only refinement). (**b**) First bending (entire-mesh refinement). (**c**) Second bending (column-only refinement). (**d**) Second bending (entire-mesh refinement). (**e**) Third bending (column-only refinement). (**f**) Third bending (entire-mesh refinement)

Table 10.7 Parameterization of the two-dimensional model for sensitivity analysis

Symbol	Label of Input parameter	Nominal value	Lower bound (%)	Upper bound (%)
A	"ElasticityModulusAllFloors"	69.00 GPa	−50	+50
B	"MaterialDensityAllFloors"	2,700.00 kg/m^3	−20	+20
C	"ElasticityModulusAllColumns"	69.00 GPa	−50	+50
D	"MaterialDensityAllColumns"	2,700.00 kg/m^3	−20	+20
E	"CornerMassAllFloors"	21.94 g	−50	+50
F	"NonLinearBeamElasticityModulus"	69.00 GPa	−50	+50
H	"NonLinearBeamMaterialDensity"	2,700.00 kg/m^3	−50	+50
K	"NonLinearBeamBackStopMass"	140.70 g	−50	+50

Table 10.8 Total variability of FE model predictions of first three bending frequencies

Mode	Mean frequency, μ (Hz)	Standard deviation, σ (Hz)	Coefficient $C_{VAR} = \sigma/\mu$ (%)
3	24.67	5.88	23.82
5	50.03	13.52	27.02
6	79.58	20.51	25.77

Coefficients of global sensitivity are estimated from an Analysis-of-Variance (ANOVA), where the ANOVA decomposes the total prediction variance into contributions due to each parameter, their linear interactions, and higher-order interactions. This decomposition can be written symbolically as:

$$\sigma^2(\omega_k) = \sum_{n=1\ldots N} S_n^2 + \sum_{p=1\ldots N q\geqslant p} S_{p,q}^2 + \sum_{p=1\ldots N q\geqslant p, r\geqslant q} S_{p,q,r}^2 + \ldots + S_{1,2,\ldots,N}^2, \tag{10.9}$$

where $S_n{}^2$ is the contribution to total variance of the *n*th variable only, $S_{p,q}{}^2$ is the contribution due to interaction between the *p*th and *q*th parameters, etc. The coefficients of partial variance reflect the statistical influence of specific effects [10]. A simplification of Eq. 10.9 is:

$$\sigma^2(\omega_k) \approx \sum_{n=1\ldots N} \left(\left. \frac{\partial \omega_k(p)}{\partial p_n} \right|_{p=p_{Mean}} \right)^2 \cdot \sigma^2(p_n), \tag{10.10}$$

where the total variance $\sigma^2(\omega_k)$ of the resonant frequency is expressed as a linear combination of the variability $\sigma^2(p_n)$ of the individual parameters. In this simplification, the coefficient of influence is simply the square of the local derivative estimated at the mean parameter value. Equation 10.10, however, is correct only for a linear model, $\omega_k = \alpha_0 + \alpha_1 p_1 + \alpha_2 p_2 + \ldots + \alpha_N p_N$, with uncorrelated parameters. The ANOVA decomposition performed next does not suffer from these restrictions.

10.5.2 Global Sensitivity Analysis of the Bending Frequencies

The coefficients of global sensitivity S_k of Eq. 10.9 are estimated using a two-level, full factorial designs-of-computer-experiments for mesh resolutions similar to those of Table 10.5. It means that, at each resolution, 256 calculations are performed considering all combinations of lower and upper bounds defined in Table 10.7. Results of a main-effect ANOVA are summarized in Table 10.9 for the first three bending frequencies.

Columns of Table 10.9 identify the mesh resolution at which an ANOVA is carried out, where the symbol N_{FE} denotes the number of elements that discretize a vertical column or horizontal floor. The rows refer to a specific parameter of the FE model. For simplicity, only the three parameters labeled A, B, and C in Table 10.7 are listed while the other factors (D to K) are grouped together. A large R^2 statistic indicates a variable that, when it is varied by itself while the other variables are kept unchanged, tends to influence significantly the prediction of resonant frequency.

It can be observed from Table 10.9 that, as expected, the modulus of elasticity of the vertical columns is the most statistically significant parameter of the model. This does not come as a surprise since the analysis is limited to the Y-direction bending dynamics, which the columns influence greatly. The second most influential variable is the density

Table 10.9 Main-effect ANOVA for predictions of the first three bending frequencies

Main-effect, R^2 statistics of the first bending frequency (mode 3)

Input parameter label	$N_{FE} = 5$ (%)	$N_{FE} = 10$ (%)	$N_{FE} = 20$ (%)	$N_{FE} = 30$ (%)	$N_{FE} = 40$ (%)	$N_{FE} = 60$ (%)
"ElasticityModulusAllFloors"	10.11	10.37	8.25	6.38	4.95	3.02
"MaterialDensityAllFloors"	16.35	16.52	16.12	15.65	15.19	14.39
"ElasticityModulusAllColumns"	73.47	73.05	75.58	77.92	79.81	82.54
All other variables combined	0.07	0.06	0.05	0.05	0.05	0.05

Main-effect, R^2 statistics of the second bending frequency (mode 5)

Input parameter label	$N_{FE} = 5$ (%)	$N_{FE} = 10$ (%)	$N_{FE} = 20$ (%)	$N_{FE} = 30$ (%)	$N_{FE} = 40$ (%)	$N_{FE} = 60$ (%)
"ElasticityModulusAllFloors"	0.64	0.63	0.51	0.39	0.29	0.17
"MaterialDensityAllFloors"	12.81	12.85	12.73	12.56	12.40	12.14
"ElasticityModulusAllColumns"	86.51	86.49	86.72	87.01	87.27	87.66
All other variables combined	0.04	0.03	0.04	0.04	0.04	0.03

Main-effect, R^2 statistics of the third bending frequency (mode 6)

Input parameter label	$N_{FE} = 5$ (%)	$N_{FE} = 10$ (%)	$N_{FE} = 20$ (%)	$N_{FE} = 30$ (%)	$N_{FE} = 40$ (%)	$N_{FE} = 60$ (%)
"ElasticityModulusAllFloors"	2.63	2.18	1.69	1.42	1.24	0.99
"MaterialDensityAllFloors"	14.24	13.94	13.61	13.40	13.24	13.00
"ElasticityModulusAllColumns"	83.07	83.83	84.64	85.13	85.47	85.96
All other variables combined	0.06	0.05	0.06	0.05	0.05	0.05

of the horizontal floors. The contribution of the floors to the bending deformation comes essentially in the form of rigid masses, which explains why the density variable is influential. The modulus of elasticity of horizontal floors is a distant third effect because the floor rigidity plays little role in the bending behavior. Finally, the group of five remaining variables is almost insignificant to explain how the resonant frequencies vary.

Table 10.9 also tabulates the influence of mesh size on the global sensitivity to model parameters. This is particularly striking for the first bending frequency where varying the level of resolution reduces the influence of the floor modulus of elasticity by a factor of three while, at the same time, it increases the sensitivity to the column modulus of elasticity by 9% points. For each main effect investigated, these variations can be quantified as a change of the R^2 statistic relative to the number of elements used in the calculations:

$$\Delta_k = \left| \frac{S_k^2(\Delta x^F) - S_k^2(\Delta x^C)}{N_{FE}^F - N_{FE}^C} \right|, \tag{10.11}$$

where Δx and N_{FE} represent the resolution and total number of elements used in a calculation, while superscripts $(\bullet)^C$ and $(\bullet)^F$ denote the coarse-mesh and fine-mesh resolutions, respectively. Figure 10.6 illustrates these incremental changes as a function of the total number of elements used in the calculations. Because the scale is logarithmic, the slope of curves Δ_k-versus-N_{FE} should be consistent with the rate-of-convergence of the truncation error (see Sect. 10.4.1). A theoretical, first-order accuracy is inserted in the triangular shapes for reference.

Figure 10.6 provides the evidence needed to confirm our hypothesis that the level of resolution in the calculation exercises a significant influence on the sensitivities of resonant frequencies. This is especially true for runs performed with fewer than $N_{FE} = 10$ elements to discretize the columns and floors. The figure also indicates that, as the level of resolution increases, these incremental changes in sensitivity start to behave according to the theoretical rate of first-order accuracy. When Δ_k reaches 10^{-4}, or less, any variation due to mesh resolution becomes so small that the global sensitivities "*stabilize,*" eventually rendering them meaningful.

10.6 Conclusion

The practice of calibrating a FE model when the agreement between physical measurements and numerical predictions is unsatisfactory is common in structural dynamics. Our contention is that calibration may not always be a legitimate recourse to improve test-analysis correlation because it attempts to compensate for parametric errors when, in fact, the source of the disagreement between measurements and predictions may be elsewhere.

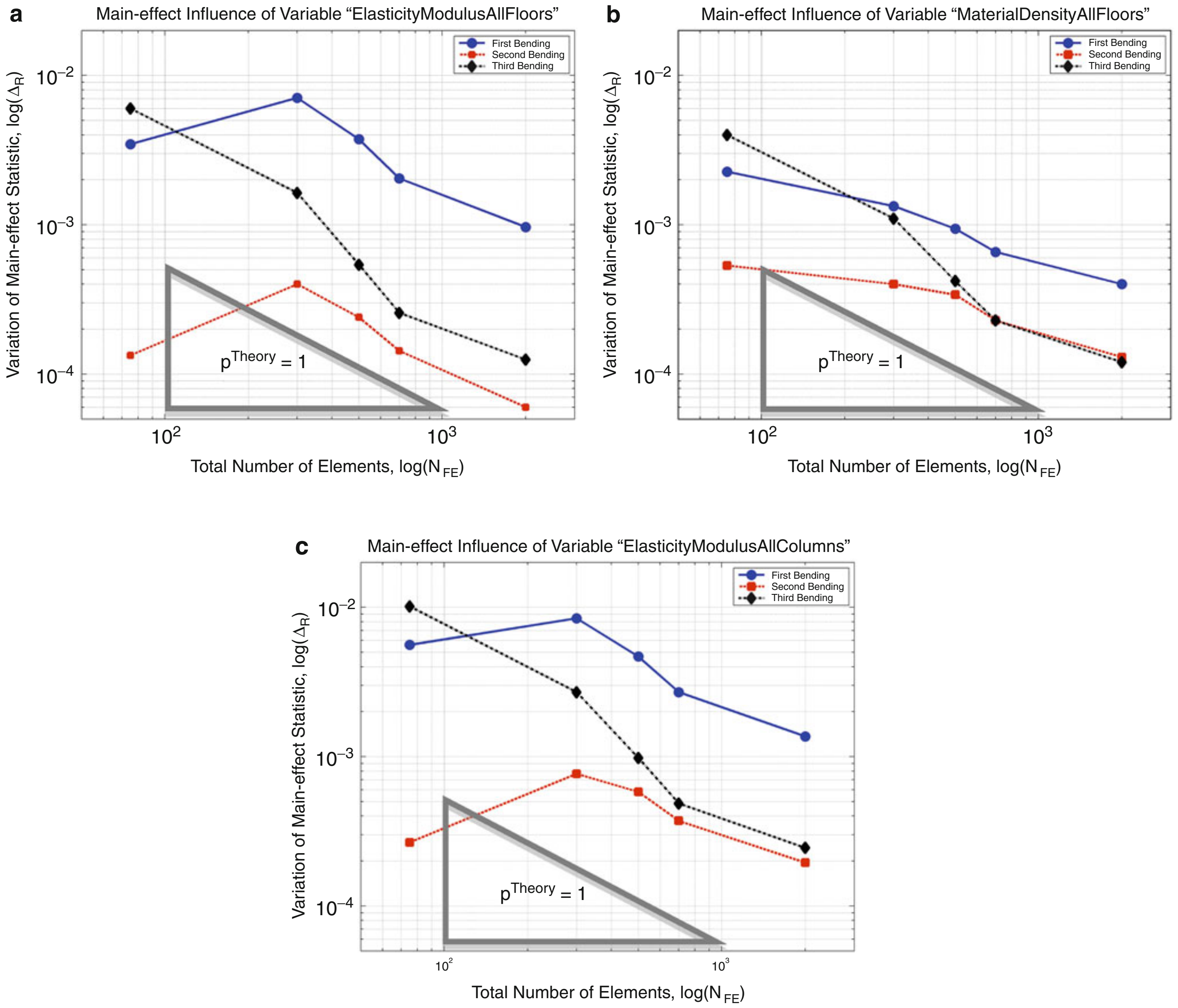


Fig. 10.6 Incremental changes of sensitivity, Δ_k, as a function of mesh resolution. (**a**) Sensitivity to “ElasticityModulusAllFloors”. (**b**) Sensitivity to “MaterialDensityAllFloors”. (**c**) Sensitivity to “ElasticityModulusAllColumns”

This research focuses attention on the truncation error caused by a lack-of-resolution of the spatial discretization of the equations of motion. When it is significant enough, truncation error introduces a systematic bias between numerical predictions and “*true values*” of the responses being simulated. Attempting to resorb this gap through model calibration is the wrong remedy. The mechanism of this commonly encountered mistake is explored by showing that truncation error adversely affects the global sensitivities of numerical predictions. Incorrect sensitivities are, in turn, fed to the calibration, resulting in an erroneous inference of model parameters.

This mechanism is illustrated by analyzing the bending deformations of a scaled model of three-story frame structure. Experimental modal analysis and finite element modeling are carried out to measure and predict, respectively, the vibration response. Mesh refinement is performed to quantify the numerical uncertainty with which resonant frequencies of the structure are predicted by the finite element model. The second step is an analysis-of-variance to estimate the global influence, or sensitivities, that parameters of the model exercise on frequency predictions. The analysis indicates that a relatively fine level of discretization is necessary before the sensitivities start to “*stabilize*.” It suggests that any attempt to calibrate calculations performed with a coarse discretization yields incorrect parameter values because the model sensitivities are erroneous.

The conclusion is that, while useful, the calibration of a numerical model should be attempted only after other sources of modeling error have been rigorously quantified and adjusted for. In particular, the overall level of truncation error should be quantified through mesh refinement. The guidance suggested is twofold. First, one should proceed with calibration only if it can be demonstrated that the global sensitivities of the model are "*stable*" under the conditions imposed by mesh refinement. Second, the numerical uncertainty of predictions should be "*small*" relative to the experimental variability. The risk of not meeting these conditions is to compensate, via calibration, for the systematic bias introduced by truncation error instead of learning parameter values from the inference process. Using these incorrect parameters in conjunction with another discretization, or another structure, would then yield erroneous predictions. The continuation of this work assesses the potentially adverse consequences for the calibration of two- and three-dimensional models of the three-story frame structure.

Acknowledgements This research is performed under the auspices of the Advanced Scientific Computing program at the Los Alamos National Laboratory (LANL). The authors express their gratitude to Dr. Mayuko Nishio, University of Tokyo, Japan, for completing the experimental modal analysis of the three-story frame structure. LANL is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396.

References

1. Cogan S (2006) Model updating for validation: a tutorial. In: 24th SEM international modal analysis conference, St. Louis, 29 Jan–3 Feb 2006
2. Mottershead JE, Friswell MI (1993) Model updating in structural dynamics: a survey. J Sound Vibrat 162(2):347–375
3. Hemez FM, Doebling SW (2001) Review and assessment of model updating for nonlinear, transient dynamics, (Los Alamos technical report LA-UR-00-0091). Mechanical systems and signal processing, vol 15, No. 1, Jan 2001, pp 45–74
4. Ben-Haim Y, Hemez FM (2011) Robustness, fidelity and prediction-looseness of models, (Los Alamos technical report LA-UR-11-0497). Royal society proceedings A, Sep 2011
5. Figueiredo E, Park G, Figueiras J, Farrar C, Worden K (2009) Structural health monitoring algorithm comparisons using standard data sets, (Technical report LA-14393). Los Alamos National Laboratory, Los Alamos
6. Nishio M, Hemez FM, Worden K, Park G, Takeda N, Farrar CR (2011) Feature extraction for structural dynamics model validation. In: 29th SEM international modal analysis conference, Jacksonville, 31 Jan–3 Feb 2011
7. Warming RF, Hyett BJ (1974) The modified equation approach to the stability and accuracy analysis of finite difference methods. J Comput Phys 14:159–179
8. Roache PJ (1994) Perspective: a method for uniform reporting of grid refinement studies. ASME J Fluid Eng 116:405–413
9. Mollineaux MG, Van Buren KL, Hemez FM, Atamturktur S (2011) Simulating the dynamics of wind turbine blades: part I, Model development and verification, (Los Alamos technical report LA-UR-11-4996). Wind energy, Sep 2011 Submitted
10. Saltelli A, Chan K, Scott M (2000) Sensitivity analysis. John Wiley & Sons, NY, New York

Chapter 11
Possibilistic Interpretation of Mistuning in Bladed Disks by Fuzzy Algebra

H. Çağlar Karataş, Ender Ciğeroğlu, and H. Nevzat Özgüven

Abstract In the study of effect of mistuning on the dynamic response analysis of bladed disk systems, in literature, probabilistic methods are used. Conversely, in this paper, mistuning will be investigated by possibilistic analysis where mistuning parameters are modeled as fuzzy variables possessing possibility distributions. Fuzzy forced response characteristics of mistuned bladed disk systems are determined by mathematical basis of fuzzy sets. In order to do so, extension principle solution of fuzzy functions is used which overcomes the dependency issue problem observed on interval arithmetic solutions; hence, enhancing solution accuracy. Membership function distributions are digitized using alpha-cut methodology, slicing distributions to levels of confidence. Bounds of fuzzy variables in each and every level of confidence are determined using genetic optimization. Using this method, fuzzy forced response characteristics of a cyclically symmetric lumped parameter bladed disk model is determined. The possibilistic interpretation of mistuning is exemplified by determining the bounds of the possible maximum blade forced responses of the system for different orders of engine order excitation and by determining worst-possible case.

Keywords Mistuning • Possibilistic analysis • Fuzzy modeling • Maximum possible forced response • Engine order excitation • Energy localization • Extension principle solution • Genetic algorithm

11.1 Introduction

Mistuning – deviation from cyclical symmetry in bladed disks - is a phenomenon which results in failure of blades much earlier than their predicted life due to localization of vibration energy on a particular blade. In order to avoid possible failures, the system needs to be analyzed by considering uncertainties in system parameters that are arisen from reasons like manufacturing tolerances, wear in service and material non-uniformities. Depending on the type of analysis, these uncertainties can either be treated as random [1, 2] or imprecise. Mathematics and methodologies considering uncertain variables as random variables have already been studied broadly in literature. To the best of authors' knowledge, there is no study on mistuning where the uncertain parameters are treated as imprecise parameters and modeled as fuzzy variables. Therefore, the aim of this paper is to define the possibilistic interpretation of mistuning by explaining the methodology for solving dynamic equations of fuzzy forced response problem which enables the determination of the maximum possible forced response among blades within each and every uncertainty range.

In order to explain the solution procedure, first of all, details of the process of modeling uncertain variables as fuzzy variables are briefly summarized. Up on the investigation of solution methods to fuzzy equations, proper method is chosen and adapted to fuzzy forced response problems. Finally, the methodology is exemplified by determining the worst-case scenario among different engine order excitations.

H.Ç. Karataş • E. Ciğeroğlu (✉) • H.N. Özgüven
Mechanical Engineering Department, Middle East Technical University, Ankara 06800, Turkey
e-mail: ender@metu.edu.tr

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_11,

11.2 Methodology

11.2.1 Fuzzy Modeling and Fuzzy Algebra

For the analysis of engineering systems with vague and imprecise uncertainties, fuzzy set theory and possibility theory are utilized together. For modeling uncertain parameters as fuzzy variables possessing possibility distributions, possibility theory [3] is used and required mathematical operations regarding fuzzy variables are governed by the fuzzy arithmetic, based on the rules of fuzzy set theory [4], all developed by A. Lutfi Zadeh.

For instance, the proposition p: 'deviation in stiffness parameters of structure Y is ±5% at most' carries fuzziness on the values that are going to be assigned to the imprecise variable, deviation in stiffness parameters of structure Y, D_k. Due to the fuzziness in meaning of the proposition p, the imprecise parameter D_k is modeled as fuzzy variable possessing possibility distribution Π_k.

Let us suppose that the universe of discourse M with a generic element m, carries all possible values of any stiffness value. Then, the proposition '±5% at most' defines a unique fuzzy subset F which is characterized by the membership function μ_F, governing the compatibility of m with the proposition. Namely, regarding to the definition of fuzzy sets [4], every value m takes a membership function value $\mu_F(takes)$ from 0 to 1 denoting its degree of membership to the fuzzy set F with zero membership function value meaning m does not belong to fuzzy set F and the closer the membership function value to unity the higher the grade of membership of m in F. Consequently, the proposition p can be transformed into the proposition "D_k is F".

According to the definition of the concept of possibility distribution [3], the proposition p induces a possibility distribution Π_{k_Y} which equates the possibility of D_k taking value m to $\mu_F(m)$. In this way D_k becomes a fuzzy variable which is associated with the possibility distribution Π_k.

For instance, let us suppose that the stiffness parameters are designed to be of value 4 kN/m. Regarding the proposition p, stiffness parameters can take values between 3.8 and 4.2 kN/m. Then in order to represent the possibility distribution Π_k of fuzzy variable D_k a triangular fuzzy number $\bar{A}$ (3.8/4/4.2) can be used as shown in Fig. 11.1; where for a triangular fuzzy number $\bar{A}(x/y/z)$, the three numbers $x<y<z$ identifies the base of the triangular possibility distribution by $[x, z]$ and its vertex by y [5]. Figure 11.1 states that the stiffness parameters are designed to be of value 4 kN/m and they are not expected to be 4.25 or 3.75 kN/m. Furthermore, it is observed that regarding proposition p, assessor is more confident to assign 3.9 kN/m to the stiffness value than 4.15 kN/m since 3.9 kN/m has larger possibility value.

Similarly, in this paper, uncertain system parameters of the bladed disk structure are modeled as triangular fuzzy variables possessing ±5% deviation at most. These fuzzy variables are used in fuzzy forced response analysis in order to obtain worst-possible blade response in bladed disk.

Mathematically, fuzzy forced response analysis is composed of solutions to fuzzy equations and basic algebraic operations of fuzzy variables. There are two methods for basic algebraic operations of fuzzy variables and solution to fuzzy equations; extension principle solution and alpha-cuts and interval arithmetic solution.

For two fuzzy numbers $\bar{A}$ and $\bar{B}$, if $\bar{A} \bullet \bar{B} = \bar{C}$, $\bullet$ denoting addition, multiplication, subtraction or division, then extension principle defines membership function for $\bar{C}$ as;

$$\bar{C}(z) = \max_{x,y} \{\min(\bar{A}(x), \bar{B}(y)) | x \bullet y = z\} \tag{11.1}$$

Extension principle solution, in definition, searches for all $[x, y]_i$ pairs that constitute z_i and records minimum membership function value within the pairs $\min(\bar{A}(x_i), \bar{B}(y_i))$ as $\bar{C}(z_i)$.The membership function value for $\bar{C}(z_i)$ is finally assigned as the maximum among the records, $\max(\bar{C}(z_i))$.

In contrast, interval arithmetic method uses alpha-cuts in definition to calculate basic algebraic operations. Alpha-cuts are slices of possibility distributions producing confidence levels as non-fuzzy, crisp intervals, as shown in Fig. 11.1 by squared marks. For the closed-bounded intervals of alpha-cuts of two fuzzy numbers $\bar{A}$ and $\bar{B}$, $\bar{A}[\alpha] = [a_1(\alpha), a_2(\alpha)]$ and $\bar{B}[\alpha] = [b_1(\alpha), b_2(\alpha)]$ (for instance, the least confidence level of fuzzy variable in Fig. 11.1 can be written as $\bar{A}[0] = [3.8, 4.2]$), interval arithmetic solution calculates addition, subtraction, multiplication and division using interval arithmetic rules [6]:

$$\bar{C}[\alpha] = \bar{A}[\alpha] \bullet \bar{B}[\alpha] \tag{11.2}$$

Due to the rules of interval arithmetic, compared with respect to extension principle solution, larger solution intervals are expected. These rules conflict with the physical structure integrity and creates dependency problem [6]. In contrast,

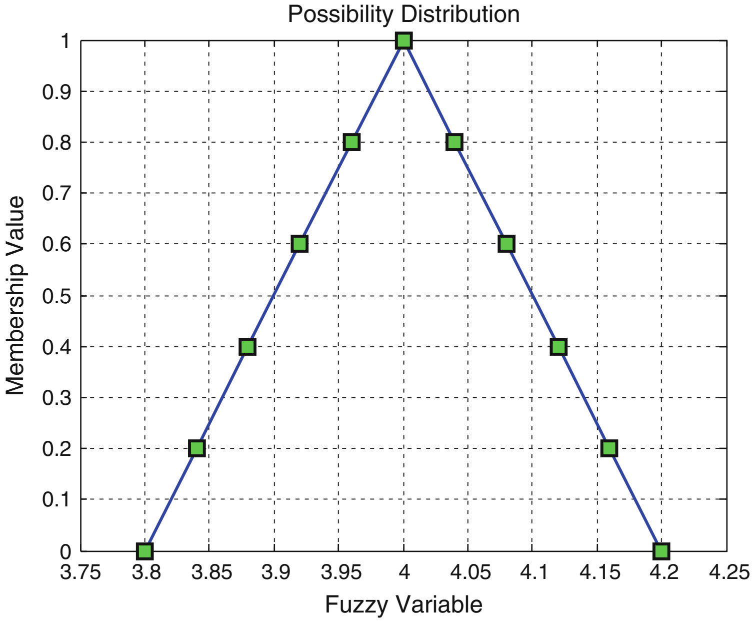


Fig. 11.1 A triangular membership function distribution

extension principle solution can be adapted to alpha-cut methodology [5]; hence, approximates the true solution practically and without creating dependency issue. Therefore, for the definition of fuzzy functions and for the solutions of fuzzy equations extension principle solution is used in this study.

11.2.2 Fuzzy Functions and Solutions of Fuzzy Equations

Mathematically speaking, fuzzy functions are mapping from fuzzy numbers to fuzzy numbers. They are usually extended from real valued functions. According to extension principle, a real-valued function $h : [x, y] \rightarrow \mathbf{R}$ can be extended to the corresponding fuzzy function $H(\bar{X}, \bar{Y}) = \bar{Z}$ as follows:

$$\bar{Z}(z) = \max_{x,y}\{\min(\bar{X}(x), \bar{Y}(y)) | h(x, y) = z, a_1 \leq x \leq b_1, a_2 \leq y \leq b_2\} \tag{11.3}$$

According to Eq. 11.3, in order to be able to assess the possibility of the dependent variable $\bar{Z}(z)$, all pairs of the independent variables $[x, y]$ that are mapped into z should be determined regardless of their confidence levels. This makes use of extension principle impossible on computer since it is impossible to determine all pairs for all types of equations. However, Buckley et.al. [5] proposes that, if the real-valued function h is continuous, then upper and lower bounds of the alpha-cuts of fuzzy function $\bar{Z}(z)$ can be determined as follows:

$$\begin{aligned} z_1(\alpha) &= \min\{h(x, y) | x \in \bar{X}[\alpha], y \in \bar{Y}[\alpha]\} \\ z_2(\alpha) &= \max\{h(x, y) | x \in \bar{X}[\alpha], y \in \bar{Y}[\alpha]\}; 0 \leq \alpha \leq 1 \end{aligned} \tag{11.4}$$

Equation 11.4 uses alpha-cut representation of possibility distributions. For instance, the fuzzy variable in Fig. 11.1 can be expressed in terms of its alpha-cuts as $\bar{A}[\alpha] = [3.8 + 0.2\alpha, 4.2 - 0.2\alpha], 0 \leq \alpha \leq 1$. So, instead of declaring the shape of the possibility distribution as $\bar{A}(x/y/z)$, the values that share the same confidence level is declared as intervals to represent the fuzzy variable.

By using alpha-cut representation of fuzzy variables on solutions to fuzzy equations, Eq. 11.4 approximates the true solution by utilizing independent variables from the same confidence level. Namely, instead of searching all possible independent variable pairs from any confidence level, the ones that share the same possibility distribution value $\bar{X}[\alpha]$ and $\bar{Y}[\alpha]$ are used to determine the interval of dependent variables with the same possibility distribution value, $\bar{Z}[\alpha]$. This approximation makes the use of extension principle possible on computer and hence, establishes the basis of solution to fuzzy functions and equations.

In order to determine upper and lower bounds of alpha-cuts of the dependent variable, $\bar{Z}[\alpha] = [z_1(\alpha), z_2(\alpha)]$, Eq. 11.4 requires a minimization and/or maximization operation which is an optimization procedure. Determination of global optimum requires a "good" initial guess if gradient based optimization procedures are used which is not practical; hence, the formulation given above is suitable to be used with genetic algorithm which can be utilized to determine the bounds of the alpha-cuts. In conclusion, in this paper for the solution of fuzzy forced response equations, extension principle is used by utilizing alpha-cut representation of fuzzy variables and by using genetic algorithm for determining bounds of alpha-cuts.

11.2.3 *Fuzzy Forced Response Analysis*

Mathematically, determining the worst-possible blade forced response requires solution to dynamic equations of fuzzy forced response problem which can be done by modal analysis. For a crisp (non-fuzzy) lumped-parameter system with structural damping, the crisp equation of motion is given as follows:

$$[M]\{\ddot{q}\} + [K^*]\{q\} = \{F\} \tag{11.5}$$

where $[K^*]$ is the complex stiffness matrix including structural damping, $[M]$ is the mass matrix, $\{F\}$ is the forcing vector, $\{q\}$ and $\{\ddot{q}\}$ represents displacement and acceleration vectors, respectively. Since mass and stiffness matrices reflect physical systems, the solution to equation of motion needs to be continuous. Therefore, the equation that is used for determining alpha-cuts of the solution, Eq. 11.4, is valid.

Consequently, let Φ be an algorithm that computes the forced vibration responses of the system by modal analysis where it accepts mass and stiffness matrices and forcing vector $[M]$, $[K^*]$, $\{F\}$as inputs and returns displacement of the nodes of the lumped parameter model $\{q\}$, as the output.If the system elements are considered as imprecise so as fuzzy variables, the equation of motion transforms into the fuzzy equation of motion:

$$[\bar{M}]\{\ddot{\bar{q}}\} + [\bar{K}^*]\{\bar{q}\} = \{\bar{F}\} \tag{11.6}$$

where over-bar expresses fuzzy matrices and vectors. Then, correspondingly the algorithm Φ would accept alpha-cuts of elements fuzzy correspondence of the inputs $[\bar{M}]$, $[\bar{K}^*]$, $\{\bar{F}\}$; $\bar{m}_{ij}[\alpha]$, $\bar{k}_{ij}[\alpha]$, $\bar{f}_i[\alpha]$ and would return the alpha-cuts of elements of fuzzy displacement vector $\{\bar{q}\}$, $\bar{q}_i[\alpha]$ as the output. Then, according to Eq. 11.4, assuming the excitation force vector is exactly known, i.e. crisp, alpha-cuts of the blade displacement forced response values with respect to deviations in system parameters are calculated as follows:

$$\begin{aligned} \bar{q}[\alpha] &= [q_1(\alpha), q_2(\alpha)] \\ q_1(\alpha) &= \min\{\Phi(k_{11}, \ldots, k_{nn}, m_{11}, \ldots, m_{nn}) | k_{ij} \in \bar{k}_{ij}[\alpha], m_{ij} \in \bar{m}_{ij}[\alpha]\} \\ q_2(\alpha) &= \max\{\Phi(k_{11}, \ldots, k_{nn}, m_{11}, \ldots, m_{nn}) | k_{ij} \in \bar{k}_{ij}[\alpha], m_{ij} \in \bar{m}_{ij}[\alpha]\} \end{aligned} \tag{11.7}$$

So, for a vibrating system with imprecise parameters regarding the proposition 'deviation in stiffness parameters of structure is ±5% at most' can be solved for alpha-cuts of the blade fuzzy forced response elements practically and without creating dependency issue using Eq. 11.7.

This concludes explanations to methodology of fuzzy forced response modal analysis. In the following section, given methodology is exemplified by determining the worst-case blade displacement of a cyclically symmetric lumped parameter mistuned bladed disk model under the excitation of an engine order forcing.

11.3 Numerical Example

Numerical example considered in this section is designed to exemplify the effects of mistuning on a cyclically symmetric integrated bladed disk system (bladed disk) by determining the worst-case scenario among different possible engine order excitations. The worst-case scenario for mistuning considerations is the case where the energy is localized on a particular blade. Since energy localization results in much higher displacement values on that particular blade compared to the rest of

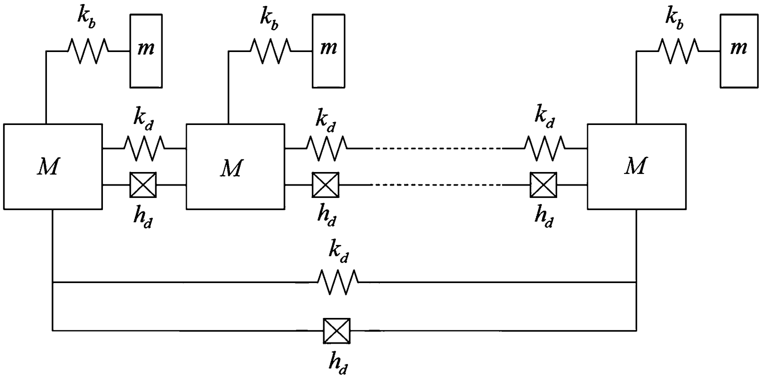


Fig. 11.2 Cyclically symmetric lumped-parameter jet-engine bladed disk model

Table 11.1 Physical values of the lumped parameter model

Parameter	Value
Disk mass, M	4 kg
Disk stiffness, k_d	60,000 N/m
Blade mass, m	2 kg
Blade stiffness, k_b	7,000 N/m
Structural damping coefficient, γ	0.01

the blades, the maximum possible blade displacements for each engine order excitation is sought and compared with the rest of the blade displacement values in order to determine the worst-case situation where the energy localization occurs.

Energy localization results from destruction of cyclical symmetry due to uncertainties like manufacturing tolerances, wear in service and material non-uniformities. In this study, these uncertainties are treated as imprecise and assumed to be effecting stiffness parameters only. The stiffness parameters are expected to deviate at most ±5% from the tuned case stiffness parameters of a cyclically symmetric lumped parameter bladed disk model shown in Fig. 11.2.

Here, M, m, k_d, k_b, h_d are the disk mass, blade mass, disk stiffness, blade stiffness and structural damping, respectively and ±5% deviation in imprecise parameter is named as ±5% mistuning level in this study. The assumption regarding imprecision in system parameters requires fuzzy modeling of stiffness parameters and the use of possibilistic analysis in order to determine maximum possible blade displacement value. Therefore, possibility distributions of fuzzy variables is determined and used in fuzzy forced response analysis using extension principle solution that utilizes alpha-cut methodology and genetic algorithm as stated in Eq. 11.7 with fuzzy stiffness parameters $\bar{k}_d[\alpha]$,$\bar{k}_b[\alpha]$ only.

Let us suppose that elements of blisk model are realized with non-fuzzy, crisp values given in Table 11.1. All system parameters except stiffness parameters are used in possibilistic analysis with the tuned case values given in Table 11.1 and stiffness parameters are modeled as triangular fuzzy variables. A triangular possibility distribution has three components, i.e. two base values and a vertex value. For fuzzy stiffness values, tuned case value given is assigned to the vertex of the corresponding fuzzy variable and base values are chosen to be ±5% deviated from the tuned case value compatible with the assumption, as shown in Fig. 11.3.

Let us further suppose that the blisk spins in a static pressure field hence experiences engine order excitation which can be modeled as [7]:

$$f_j = F_m \cos(m\Omega t + \frac{2\pi mj}{n}), \tag{11.8}$$

where j, m, n, Ω and F_m are the blade number, engine order, number of blades, rotational speed and excitation amplitude, respectively. The effects of engine orders 1, 2, 3 and 4 of 1 N forcing amplitude are sought for the worst-case analysis. In order to determine the mode that each type of excitation excites, the frequency term $m\Omega$ is swept within the frequency range of interest. For 0.9 confidence level (corresponding to ±0.5% mistuning level), at each forcing frequency, the displacement of the blade corresponding to the maximum response is given for different engine order excitations in Fig. 11.4.

From Fig. 11.4 it can clearly be seen that different engine order excitations excite different modes of vibration at different resonant frequencies. Also, the figure highlights the maximum possible blade displacements recorded within the blisk for a forcing of 1 N amplitude and ±0.5% mistuning level within the forcing frequency range of interest.

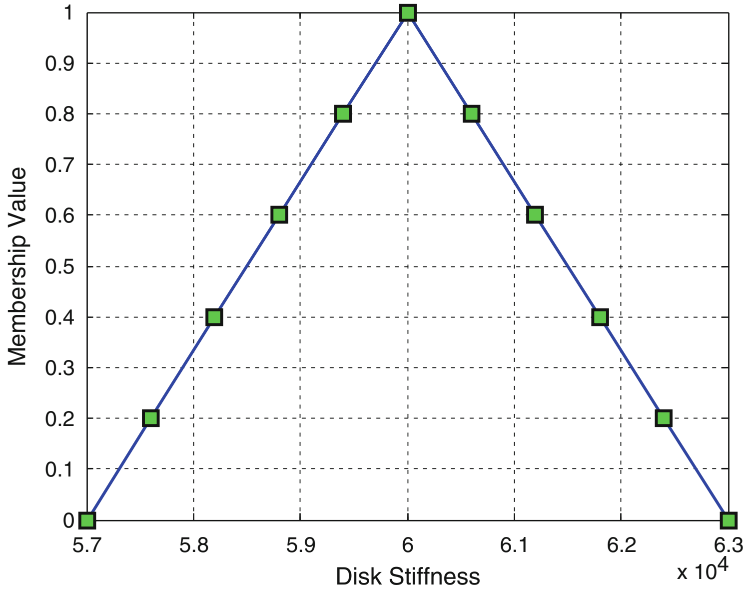


Fig. 11.3 Possibility distribution for disk stiffnesses

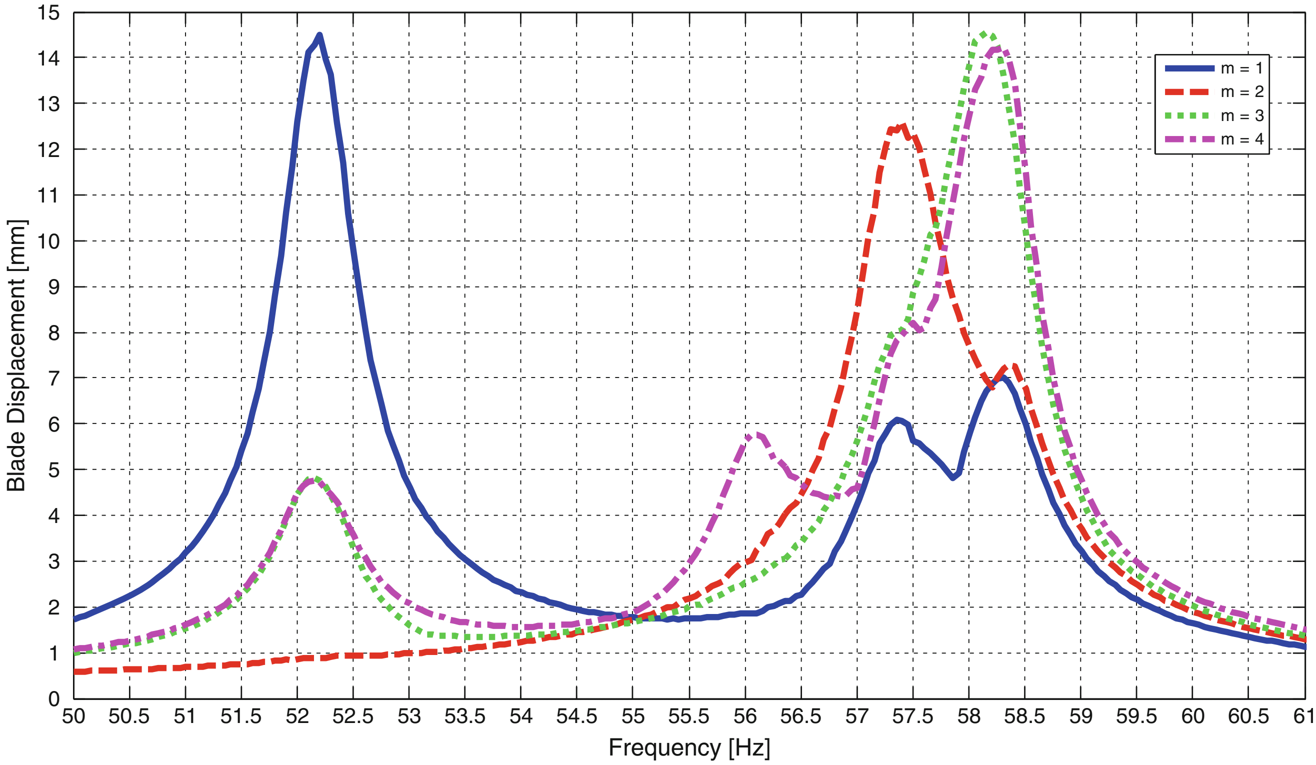


Fig. 11.4 Displacement of the blade corresponding to the maximum response vs. frequency for engine orders 1, 2, 3 and 4 (±0.5% mistuning level)

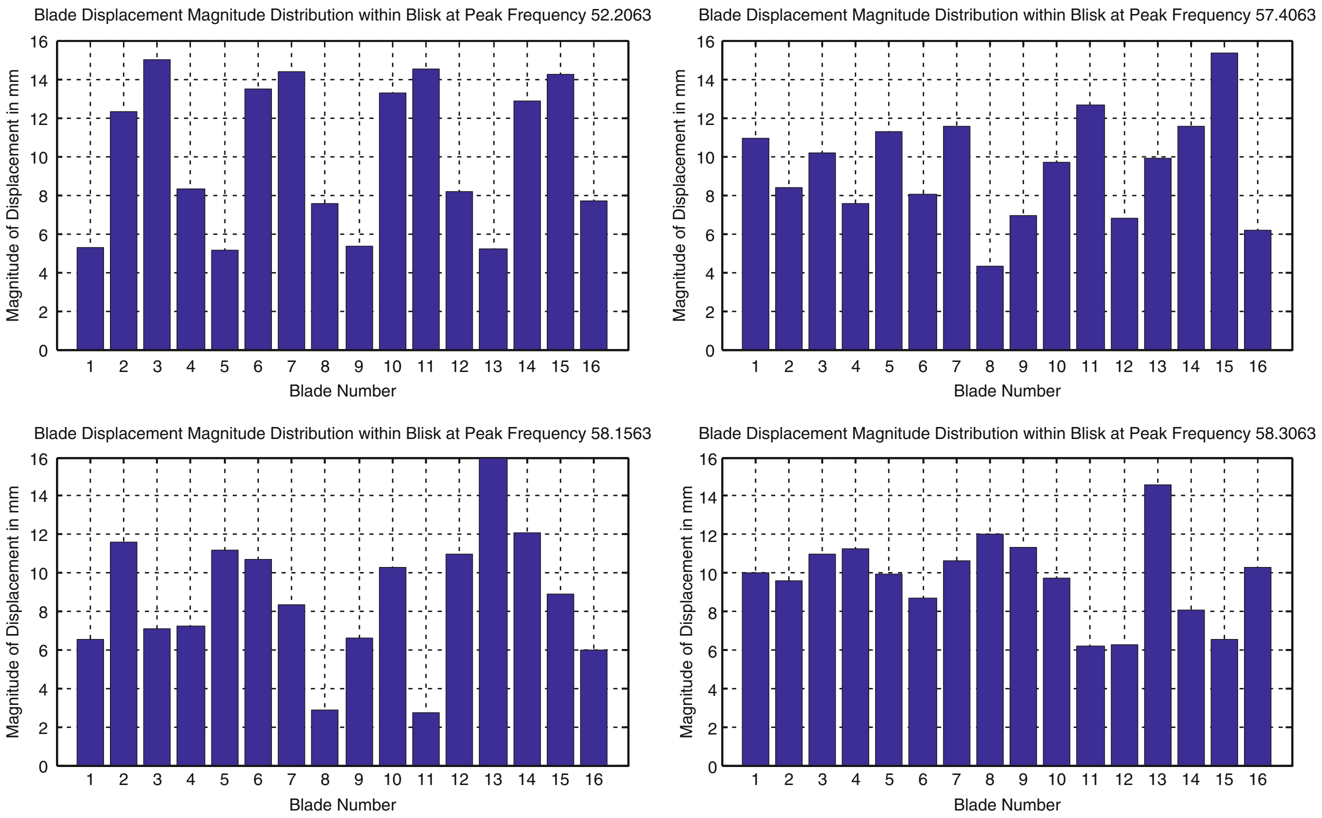


Fig. 11.5 Maximum response of blades for engine orders 1, 2, 3 and 4 (±5% mistuning level)

In order to observe the effect of mistuning, together with the blade that has maximum displacement, the vibration amplitudes of rest of the blades are given in Fig. 11.5 as bar plots for four different engine orders at ±5% mistuning level. From Fig. 11.5, among engine orders 1, 2, 3 and 4, the energy localization for engine orders 3 and 4 can clearly be seen in the figure where a single blade has a much larger response compared to all other blades. Moreover, engine order 3 shows the highest maximum blade forced response compared to other cases. Therefore, the worst-possible case is determined to be the case where the blisk is excited by engine order 3 for ±5% mistuning level.

This concludes the determination of worst-case scenario for the blisk model considered for different engine order excitations. First, the mode that each type of excitation excites is selected by pursuing a frequency sweep within the range of interest and resonant frequencies are determined for ±0.5% mistuning level since different engine order excitations excite different modes of vibration. Then the behavior of the blisk excited by four different engine orders at the corresponding resonant frequencies is investigated and the case where the blisk is excited by engine order 3 is determined to be the worst-possible case for the system considered. Furthermore, In Fig. 11.6, behavior of the blisk excited by engine order 3 for different mistuning levels is detailed at different excitation frequencies. The plot contains representative mistuning levels to show the change in the maximum blade displacement with respect to increasing percent deviations in stiffness parameters. It is observed that the possibility of obtaining a higher maximum forced response increases as the mistuning level increases. Therefore, it can be concluded that around the crisp resonant frequency, the possibility of resonance increases as the mistuning level increases.

11.4 Conclusion

In conclusion, this study explains and exemplifies the methodology to treat uncertainties in mistuning problems as imprecise and the way to pursue fuzzy forced response analysis. The proper fuzzy arithmetic method is suggested in order to overcome the dependency issue problem. Adapting fuzzy solution tools to modal analysis, solution to fuzzy forced response problem is

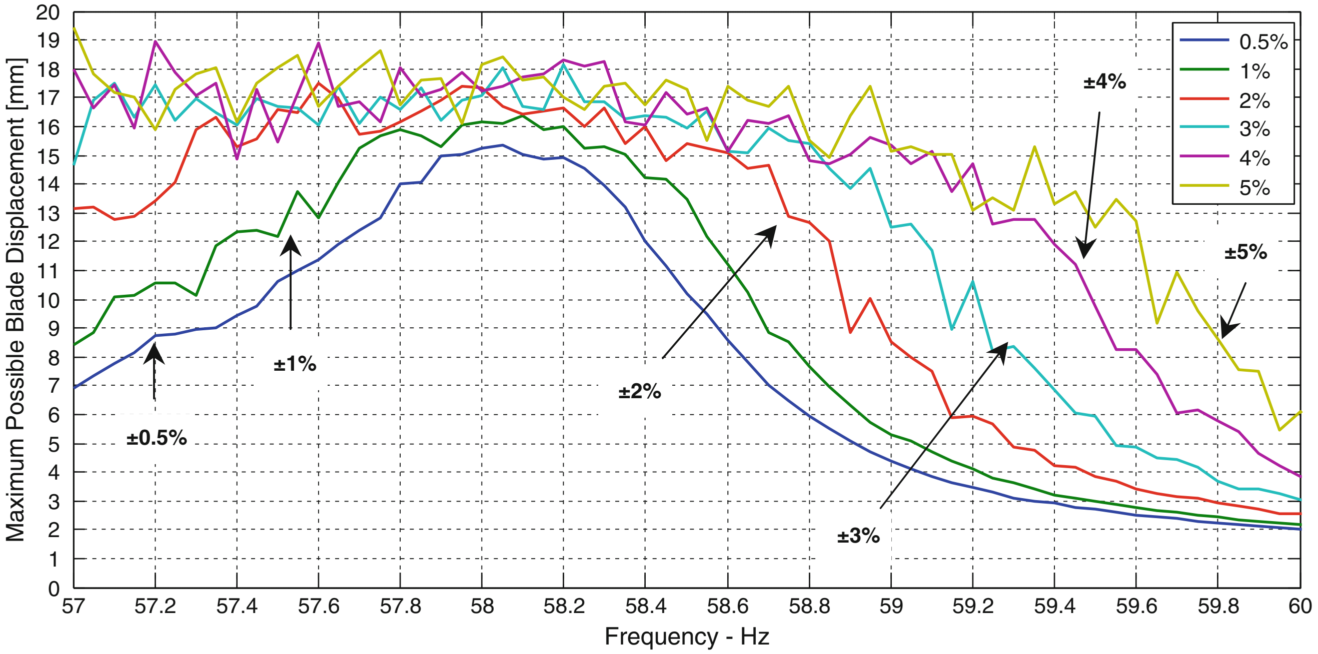


Fig. 11.6 Maximum blade displacement as a function of frequency and mistuning level (Engine order 3)

proposed. Using fuzzy forced response analysis that uses extension principle solution which utilizes alpha-cut methodology and genetic algorithm, the worst-possible effect of mistuning on a cyclically symmetric lumped parameter model that is excited by different engine order excitations is determined. It is observed that not all engine order excitations result in energy localization. Therefore, in order to determine the critical one, several engine order excitations are required to be checked separately. Finally, the trend of change of maximum blade displacement with respect to changes in frequency and mistuning level is obtained. It is concluded that possibility of resonance increases as the mistuning level increases around the tuned case resonant frequency.

References

1. Griffin JH, Hoosac TM (1984) Model development and statistical investigation of turbine blade mistuning. J Vib Acoust Stress Reliab Des 106 (2):204–210
2. Sinha A (1986) Calculating the statistics of forced response of a mistuned bladed disk assembly. AIAA J 24(11):1797–1801
3. Zadeh LA (1999) Fuzzy sets as a basis of a theory of possibility. Fuzzy Set Syst 100(1):9–34, North-Holland
4. Zadeh LA (1965) Fuzzy Sets. Inform Control 8:338–353
5. Buckley JJ, Eslami E, Feuring T (2002) Fuzzy mathematics in economics and engineering. Physica, Heidelberg/New York
6. Moore RE, Kearfott RB, Cloud MJ (2009) Introduction to interval analysis. Society for Industrial and Applied Mathematics, Philedelphia
7. Ciğeroğlu E, Özgüven HN (2006) Nonlinear vibration analysis of bladed disks by friction dampers. J Sound Vib 295:1028–1043, Elsevier

Chapter 12
FEM Sensitivity Technique for Dynamic Response Uncertainty Analyses

Robert N. Coppolino

Abstract Parametric variation of large order finite element models is required for bracketing dynamic response uncertainties of complex systems. It is generally accepted that structural joint interface stiffness parameters are uncertain over parametric ranges that may span orders of magnitude. Therefore incremental sensitivity analysis techniques are not appropriate. Recent development of a vector-based sensitivity analysis technique that utilizes residual shape functions provides an efficient, accurate tool for vibration mode uncertainty studies. This paper introduces an extension of the sensitivity technique for dynamic response uncertainty studies. The key innovation lies in an efficient quasi-static stiffness sensitivity formulation that is closely related to mathematical principles employed in the vibration mode sensitivity technique. An end-to-end sensitivity formulation, appropriate for response to transient, steady-state and random environments, is defined for displacement, member load and stress recovery based on the mode displacement method with quasi-static residual vector augmentation. A simple illustrative example is provided to demonstrate accuracy and efficiency of the technique.

Keywords Dynamic response sensitivity method • Finite element method • Modal sensitivity method

12.1 Introduction

An efficient method for computation of structural dynamic modal frequency and mode shape sensitivities due to variation of physical stiffness and mass parameters was published at IMAC XXIX [1]. At the heart of the method is the formation of residual vectors that describe the distributed effect of both stiffness and mass deviations with respect to the baseline system's modes. The combined set of baseline modes and residual vectors define a transformation matrix which is used to form a greatly reduced order sensitivity model that accurately tracks changes in system modes and natural frequencies driven by large stiffness and mass variations. This technique was successfully applied to effect reconciliation of an International Space Station Component's finite element model with modal test data as early as 2001 [2]. A second, important application of sensitivity analysis is estimation of system dynamic response variations (accelerations, displacements, strains and stresses) due to stiffness and mass uncertainties. Additional residual vectors, however, must be defined in order to accurately and efficiently track changes in system dynamic response.

Two equally valid strategies for accurate computation of system dynamic response have been widely employed for many years, namely (a) the mode acceleration method [3], and (b) the residual mode vector method [4]. Both approaches account for "quasi-static" response contributions of high frequency modes that are not part of the selected (truncated) set of low frequency modes. Details associated with implementation of the above techniques in finite element analysis are provided in Ref. [5]. In order to efficiently and accurately track changes in dynamic response loads due to stiffness and mass variations, sensitivity of "quasi-static" contributions of high frequency modes must be taken into account. Formation of vectors describing sensitivity of the "quasi-static" residual vectors is the primary innovation of this paper.

R.N. Coppolino (✉)
Measurement Analysis Corporation, 23850 Madison Street, Torrance 90505, CA, USA
e-mail: rcoppolino@macorp.net

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_12,

12.2 Nomenclature

Variable	Definition	Equation	Variable	Definition	Equation
[A]	Generalized mass	13.31	[ΔB]	Damping alteration	13.1
[B]	Damping	13.1	[ΔK]	Stiffness alteration	13.1
{F(t)}	Physical applied forces	13.1	[ΔM]	Mass alteration	13.1
[I]	Identity matrix	13.22	[Δk]	Generalized stiffness alteration	13.6
[K]	Stiffness	13.1	[Δm]	Generalized mass alteration	13.6
[M]	Mass	13.1	[Φ]	Eigenvectors, modes	13.2
[Q]	Generalized coordinate	13.24	[Γ]	Load allocation	13.1
[U]	Static displacement	13.11	[Ψ]	Trial vectors	13.4
[k]	Generalized stiffness	13.5	[φ]	Generalized eigenvector	13.7
[m]	Generalized mass	13.5	[λ]	Eigenvalues	13.2
p	Sensitivity parameter	13.1	{σ(t)}	Stresses	13.21
{q}	Modal displacement	13.19	[ω]	Frequency (rad/s)	13.10
{u}	Physical displacements	13.1	ζ	Modal critical damping ratio	13.19

12.3 Low Frequency Mode Sensitivity

Low frequency mode sensitivity was the primary focus of Ref. [1]. The matrix equation set describing dynamics of a baseline or altered system, excited by applied dynamic loads is

$$\left[M_0 + \sum_{i=1}^{N} p_i \cdot \Delta M_i\right]\{\ddot{u}\} + \left[B_0 + \sum_{i=1}^{N} p_i \cdot \Delta B_i\right]\{\dot{u}\} + \left[K_0 + \sum_{i=1}^{N} p_i \cdot \Delta K_i\right]\{u\} = [\Gamma]\{F(t)\} \tag{12.1}$$

When all "p$_i$" are null, the system is "baseline". Attention shall be limited to sensitivities due to stiffness and mass changes.

The low frequency undamped modes of the baseline system are solutions of the eigenvalue problem

$$[K_O][\Phi_{0L}] - [M_O][\Phi_{0L}][\lambda_{0L}] = [0] \tag{12.2}$$

Definition of residual vectors describing parametric variations in (12.1) is accomplished utilizing the lowest frequency mode shapes of the baseline structure as well as the lowest mode shapes associated with each independent alteration of the structure

$$[K_O + \overline{p}_i \Delta K_i][\Phi_{iL}] - [M_O + \overline{p}_i \Delta M_i][\Phi_{iL}][\lambda_{iL}] = [0] \quad (\text{for i} = 1, \ldots, \text{N}), \tag{12.3}$$

where $\bar{p}_i$ is a finite (rather than infinitesimal parametric perturbation). An initial set of trial vectors that redundantly encompass all low frequency altered system mode shapes is

$$[\Psi] = [\, \Phi_{1L} \quad \Phi_{2L} \quad \ldots \quad \Phi_{NL} \,] \tag{12.4}$$

The redundant set of trial vectors is reduced to a linearly independent "modal" set, $[\bar{\Phi}_{OL}]$, by following the methodology described in Appendix A. The resulting approximate generalized sensitivity model is

$$\left[k_O + \sum_{i=1}^{N} p_i [\Delta k_i]\right][\phi] - \left[m_O + \sum_{i=1}^{N} p_i [\Delta m_i]\right][\phi][\lambda] = [0], \tag{12.5}$$

where the reduced stiffness and mass matrix components are

$$[k_O] = \left[\bar{\Phi}_{OL}^T K_O \bar{\Phi}_{OL}\right], \quad [m_O] = \left[\bar{\Phi}_{OL}^T M_O \bar{\Phi}_{OL}\right], \quad [\Delta k_i] = \left[\bar{\Phi}_{OL}^T \Delta K_i \bar{\Phi}_{OL}\right], \quad [\Delta m_i] = \left[\bar{\Phi}_{OL}^T \Delta M_i \bar{\Phi}_{OL}\right], \tag{12.6}$$

The low frequency physical modes for the altered dynamic system are recovered using the relationship

$$[\Phi_L] = [\bar{\Phi}_{OL}][\phi] \tag{12.7}$$

12.4 System Dynamic Response and “Quasi-Static” Residual Vectors

The matrix equations describing exact forced response of a linear structural model are

$$[M]\{\ddot{u}\} + [B]\{\dot{u}\} + [K]\{u\} = [\Gamma]\{F(t)\} \tag{12.8}$$

Truncation inaccuracies associated with forced response analysis (when the lowest frequency modes are included) are remedied by incorporation of “quasi-static” residual response of the higher frequency modes [4, 5]. MacNeal’s approximation incorporating “quasi-static” residual response is

$$\{u(t)\} = [\Phi_L]\{q_L\} + \left[\Psi_\rho\right]\{F(t)\}, \tag{12.9}$$

where the “quasi-static” residual vectors are defined as

$$\left[\Psi_\rho\right] = [\Phi_H]\left[\omega_H^2\right]^{-1}\left[\Phi_H^T\Gamma\right] = [K]^{-1}[\Gamma] - [\Phi_L]\left[\omega_L^2\right]^{-1}\left[\Phi_L^T\Gamma\right] \tag{12.10}$$

A more convenient alternative to the above expression utilizes the residual vectors, $[\Psi_\rho]$, as pseudo-modes that are used to complete the augmented trial vector set.

In order to assure accurate calculation of system response, in the presence of stiffness and mass changes, sensitivity of the “quasi-static” residual trial vectors, defined by (12.10), must be formulated. This is accomplished by collection of static response solutions for baseline and altered systems which are excited by the load distributions, $[\Gamma]$. The static solutions are

$$[U_O] = [K_O]^{-1}[\Gamma], \quad [U_i] = [K_O + \bar{p}_i \cdot \Delta K_i]^{-1}[\Gamma], \quad (\text{for i} = 1, \ldots, \text{N}) \tag{12.11}$$

An initial set of trial static displacement vectors that adequately (and perhaps redundantly) spans all potential altered system static deflection shape variations is

$$[\Psi_S] = [\,U_0 \quad U_1 \quad \ldots \quad U_N\,] \tag{12.12}$$

A minimum rank augmented of the trial vector set spanning (a) low frequency mode variations (defined in Sect. 12.3), and (b) static deflection shape variations is defined using the general procedure (described in Appendix A) as follows:

$$\left[\Psi'_S\right] = [\Psi_S] - [\bar{\Phi}_{OL}]\left[\bar{\Phi}_{OL}^T K_O \bar{\Phi}_{OL}\right]^{-1}\left[\bar{\Phi}_{OL}^T K_O \Psi_S\right] \quad (\text{residual ``quasi - static'' trial vectors}) \tag{12.13}$$

$$\left[\bar{\Psi}_{OS}\right] = \left[\,\bar{\Phi}_{OL} \quad \Psi'_S\,\right] \ (\text{augmented modal sensitivity trial vectors}) \tag{12.14}$$

$$\left[\bar{\Psi}_{OS}^T K_O \bar{\Psi}_{OS}\right][\varPsi] - \left[\bar{\Psi}_{OS}^T M_O \bar{\Psi}_{OS}\right][\varPsi][\lambda_\varPsi] = [0], \quad [\bar{\Phi}_{OS}] = \left[\bar{\Psi}_{OS}\right][\varPsi] \quad (\text{orthogonalized trial vectors}) \tag{12.15}$$

The response sensitivity analysis relationships are therefore

$$\left[k_O + \sum_{i=1}^{N} p_i[\Delta k_i]\right][\phi] - \left[m_O + \sum_{i=1}^{N} p_i[\Delta m_i]\right][\phi][\lambda] = [0], \tag{12.16}$$

Where

$$[k_O] = \left[\bar{\Phi}_{OS}^T K_O \bar{\Phi}_{OS}\right], \quad [m_O] = \left[\bar{\Phi}_{OS}^T M_O \bar{\Phi}_{OS}\right], \quad [\Delta k_i] = \left[\bar{\Phi}_{OS}^T \Delta K_i \bar{\Phi}_{OS}\right], \quad [\Delta m_i] = \left[\bar{\Phi}_{OS}^T \Delta M_i \bar{\Phi}_{OS}\right] \tag{12.17}$$

The physical modes for the altered dynamic system (with "quasi-static" residual enhancement) are

$$[\Phi] = [\bar{\Phi}_{OS}][\phi] \tag{12.18}$$

Forced response of the sensitized dynamic system is calculated by solution of the uncoupled modal equations

$$\ddot{q}_i(t) + 2\zeta_i\omega_i\dot{q}_i(t) + \omega_i^2 q_i(t) = \left[\Phi_i^T\Gamma\right]\{F(t)\} \tag{12.19}$$

The system dynamic displacement, velocity, and acceleration responses are

$$\{u(t)\} = [\Phi]\{q(t)\},\ \{\dot{u}(t)\} = [\Phi]\{\dot{q}(t)\},\ \{\ddot{u}(t)\} = [\Phi]\{\ddot{q}(t)\} \tag{12.20}$$

Provided a stress-displacement matrix, $[K_\sigma]$, is available (also a function of the sensitivity parameters), the system stresses are

$$\{\sigma(t)\} = [K_\sigma]\{u(t)\} \tag{12.21}$$

12.5 Illustrative Example: Rectangular Plate

The illustrative example structure, selected to demonstrate the qualities of the dynamic response sensitivity technique, consists of the 8.5″ width, 6.3″ height, 0.06″ thick aluminum ($E = 10^7$ psi, $v = 0.3$, $\rho g = 0.3$ lb/in.^{3}) rectangular plate, illustrated in Fig. 12.1. The plate is supported on its lower face and two side faces; the upper face is free.

The plate, modeled with 108 grid points, is shown in its third mode with its lower and side faces simply supported. A unit integrated applied load distribution (each point load is 1/12 units) to be employed in forced response analysis is also illustrated. The boundary stiffness variation distributions are parameterized as follows:

- Lower boundary: $p_1 \times \Delta K(i,i) = p_1 \times [20{,}000]$ on all lower boundary grid points on DOF 5
- Side boundaries: $p_2 \times \Delta K(i,i) = p_2 \times [20{,}000]$ on all side boundary grid points on DOF 6

Accuracy of the modal sensitivity technique (described in Sect. 12.3) is illustrated by comparison of natural frequency sensitivities for three methods of analysis, namely, (a) exact modal analysis, (b) approximate analysis based on ten baseline system modes plus 16 residuals (based on (12.4)–(12.7)), and (c) approximate analysis based on 26 baseline system modes

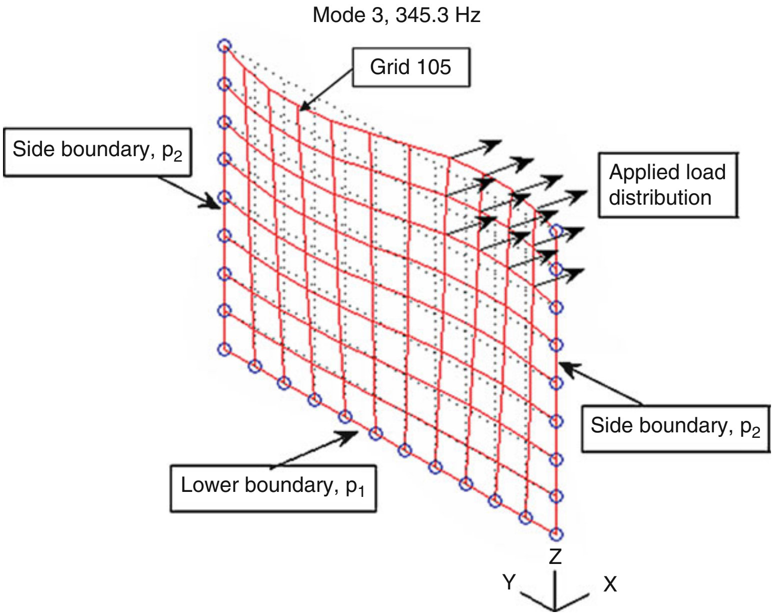


Fig. 12.1 Illustrative example rectangular plate (mode 3, simply supported on three faces)

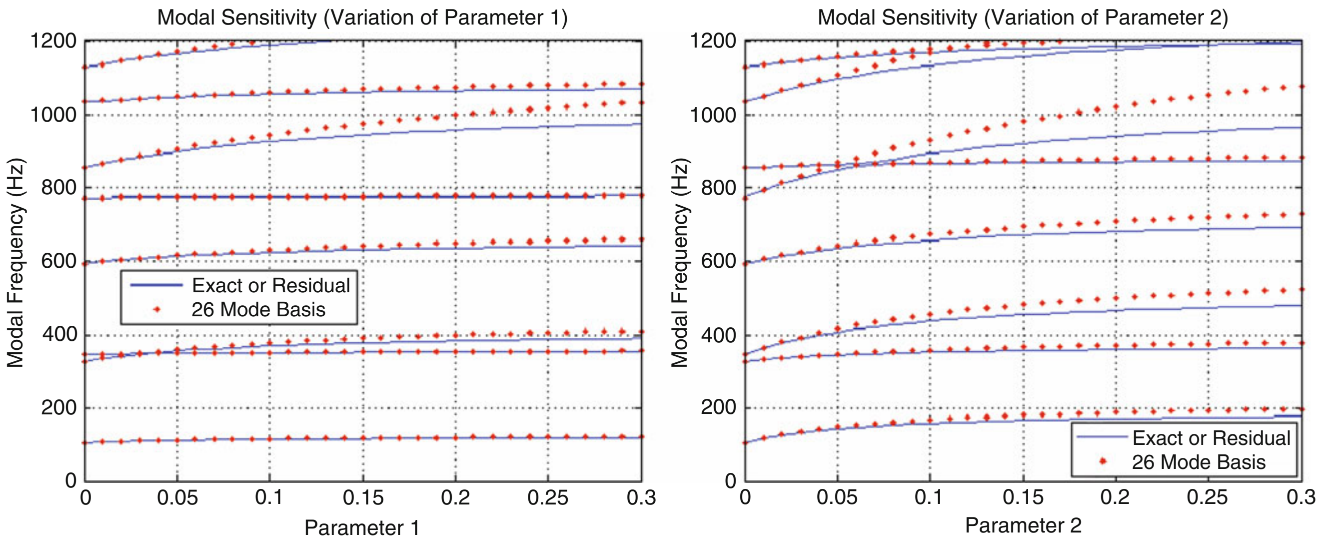


Fig. 12.2 Rectangular plate modal frequency sensitivities

Table 12.1 Rectangular plate modal sensitivities ($p_1 = p_2 = 0.1$)

	Exact solution											
			1	2	3	4	5	6	7	8	9	10
	Mode	Freq (Hz)	174.14	403.75	458.24	703.30	921.79	956.98	1,182.60	1,250.70	1,611.80	1,740.90
10 Modes + Residuals	1	174.15	100.00									
	2	403.80		100.00				0.01	0.01			0.01
	3	458.28			100.00	0.01						
	4	703.51			0.01	100.00				0.05		
	5	921.87					100.00	0.01	0.01			0.01
	6	975.19		0.01				100.00	0.07			0.05
	7	1,182.90		0.01			0.01	0.07	100.00			0.08
	8	1,251.10				0.04				100.00	0.03	
	9	1,611.90								100.00	0.03	
	10	1,741.80		0.01			0.01	0.05	0.08			99.99

	Exact solution											
			1	2	3	4	5	6	7	8	9	10
	Mode	Freq (Hz)	174.14	403.75	458.24	703.30	921.79	956.98	1,182.60	1,250.70	1,611.80	1,740.90
26 Modes	1	189.04	99.95	0.95			2.61	1.04	0.51			0.13
	2	424.19	0.96	99.91			0.70	2.76	2.29			0.48
	3	487.19			99.93	0.36				0.98	3.28	
	4	734.14			0.34	99.92				1.87	0.55	
	5	978.82	1.85	1.72			90.37	42.35	2.37			0.52
	6	995.98	2.01	2.21			42.22	90.39	2.55			2.39
	7	1,243.60	0.56	2.21			3.08	1.34	99.74	0.01		0.01
	8	1,290.90			0.89	1.87			0.01	99.85	1.89	
	9	1,688.50			3.20	0.33				1.78	99.68	
	10	1,803.60	0.16	0.38			1.32	1.93	0.18			99.77

(no residuals). Results of sensitivity analyses for which p1 and p2 are separately varied, summarized below in Fig. 12.2, indicate a close match of exact analysis and modal sensitivity technique, confirming validity of the modal sensitivity technique [1].

Forced response sensitivity analysis (described in Sect. 12.4) is demonstrated by study of modal parameter and dynamic response changes associated with $p_1 = p_2 = 0.1$. Table 12.1 provides a comparison of exact and approximate modal frequencies and mode shapes (via cross-orthogonality [5]).

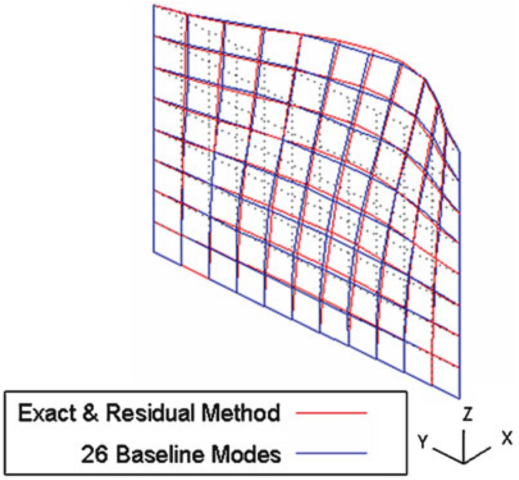
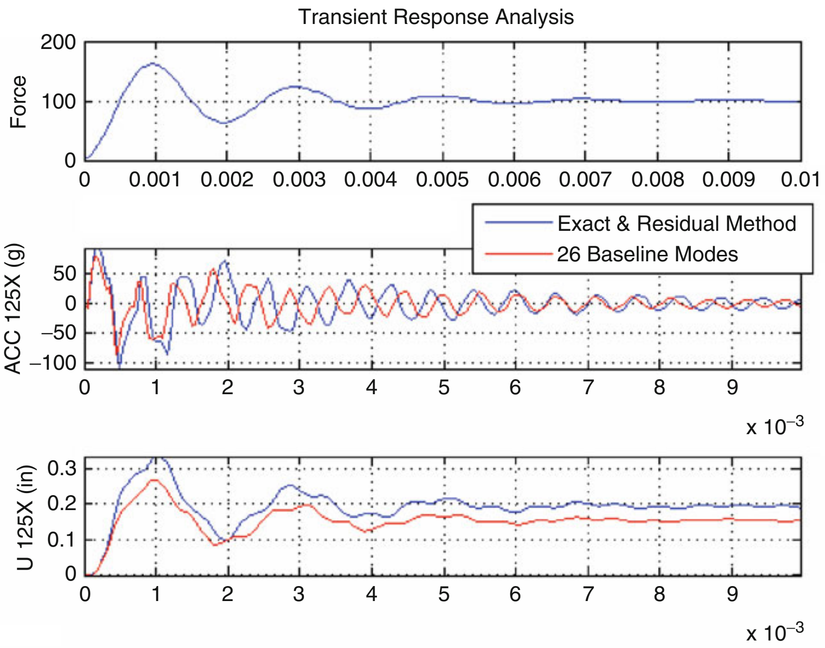


Fig. 12.3 Quasi-static deflection accuracies (26 baseline mode estimate has ~5% error)



Case	max|A105X| (g)	max(U125X) (in)	U125X (in) @ t=0.5 sec
Exact Analysis	110.13	0.3333	0.1934
Residual Method	109.19	0.3332	0.1933
26 Baseline Modes	85.79	0.2681	0.1540

Fig. 12.4 Rectangular plate response to a step-transient applied force

The numerical results provide further confirmation of the accuracy of the modal sensitivity technique. Note its superiority over the 26 baseline mode set sensitivity approximation. Additional "quasi-static" residual vector sensitivity data is required (see Sect. 12.4) for accurate transient response estimation. It was found that three additional trial vectors, appended to the original 26 trial vectors, permit accurate tracking of "quasi-static" response of the plate. Ability of the resulting trial vector set to track plate static deflection, subjected to the unit amplitude load patch in Fig. 12.1, is illustrated below in Fig. 12.3.

Performance of the forced response sensitivity technique is evaluated for the situation in which a distributed transient load is applied to the patch shown in Fig. 12.1. Acceleration and displacement responses at grid point 105 (see Fig. 12.1) to a transient, step-up load are provided below in Fig. 12.4

The above results demonstrate that modal sensitivity analysis using the residual sensitivity technique yields accurate acceleration and displacement estimates. Note that the 26 baseline modes (no residuals) approximation yields erroneous peak and steady-state displacements (which are required for recovery of structural member loads and stresses).

12.6 Concluding Remarks

This paper provides a natural extension of the modal sensitivity method introduced in Ref. [1] for estimation of accurate dynamic response sensitivities. The key innovation of this work is systematic definition of trial vectors that permit efficient tracking of (1) natural frequency and mode shape changes, and (2) static response sensitivities. The trial vectors are used to define a compact, reduced order set of dynamic equations that, when solved, closely approximate exact system dynamic response sensitivities.

Demonstrated accuracy of the technique indicates that studies of dynamic response and load sensitivity, for large parametric alterations, may be reliably conducted using extremely small order approximate dynamic models. Such capability offers a timely, cost-effective means for estimation of peak dynamic loads in the presence of large uncertainties in multiple system parameters, typical of distributed structural joints.

Implementation of the dynamic response sensitivity method requires explicit knowledge of mass and stiffness matrices. This presents a challenge when FEM codes that do not make stiffness and mass matrices accessible are used.

Appendix A: General Procedure for Generating Residual Vector Sets

Consider the general task of relating an arbitrary set of trial vectors, $[\Psi]$, to a set of the lowest frequency system modes, $[\Phi_{\mathrm{OL}}]$, of an undamped dynamic system

$$[K_O][\Phi_{0L}] - [M_O][\Phi_{0L}][\lambda_{0L}] = [0] \quad \text{where} \quad \left[\Phi_{OL}^T M_O \Phi_{OL}\right] = [\mathrm{I}], \left[\Phi_{OL}^T K_O \Phi_{OL}\right] = [\lambda_{OL}] \tag{12.22}$$

In addition, the stiffness matrix is related to the low and high frequency modes according to

$$[K_O] = [M_O \Phi_{OL}][\lambda_{OL}]\left[\Phi_{OL}^T M_O\right] + [M_O \Phi_{OH}][\lambda_{OH}]\left[\Phi_{OH}^T M_O\right] \tag{12.23}$$

The set of trial vectors is the sum of (a) a linear combination of system modes, $[\Phi_{\mathrm{OL}}]$ and (b) residual vectors, $[\Psi_{\mathrm{R}}]$

$$[\Psi] = [\Phi_{OL}][Q] + [\Psi_R] \tag{12.24}$$

Employing weighted least squares, and enforcing orthogonality of the residuals with respect to the system modes,

$$\left[\Phi_{OL}^T K_O \Psi\right] = \left[\Phi_{OL}^T K_O \Phi_{OL}\right][Q] + \left[\Phi_{OL}^T K_O \Psi_R\right], \tag{12.25}$$

which results in,

$$[Q] = \left[\Phi_{OL}^T K_O \Phi_{OL}\right]^{-1}\left[\Phi_{OL}^T K_O \Psi\right] \tag{12.26}$$

Thus the residual vectors are,

$$[\Psi_R] = [\Psi] - [\Phi_{OL}]\left[\Phi_{OL}^T K_O \Phi_{OL}\right]^{-1}\left[\Phi_{OL}^T K_O \Psi\right] = [\Psi] - [\Phi_{OL}][\lambda_{OL}]^{-1}\left[\Phi_{OL}^T K_O \Psi\right] \tag{12.27}$$

Due to the assumption in (12.24), the residual vectors are orthogonal with respect to the stiffness matrix as well as the mass matrix, as shown below:

$$\left[\Phi_{OL}^T M_O \Psi_R\right] = \left[\Phi_{OL}^T M_O \Psi\right] - \left[\Phi_{OL}^T M_O \Phi_{OL}\right][\lambda_{OL}]^{-1}\left[\Phi_{OL}^T K_O \Psi\right] = \left[\Phi_{OL}^T M_O \Psi\right] - [\lambda_{OL}]^{-1}\left[\Phi_{OL}^T K_O \Psi\right] \tag{12.28}$$

Substituting (12.23) into the above results in

$$\left[\Phi_{OL}^T M_O \Psi_R\right] = \left[\Phi_{OL}^T M_O \Psi\right] - [\lambda_{OL}]^{-1}\left[\left[\Phi_{OL}^T M_O \Phi_{OL}\right][\lambda_{OL}]\left[\Phi_{OL}^T M_O\right] + \left[\Phi_{OL}^T M_O \Phi_{OH}\right][\lambda_{OH}]\left[\Phi_{OH}^T M_O\right]\Psi\right] \tag{12.29}$$

$$\left[\Phi_{OL}^T M_O \Psi_R\right] = \left[\Phi_{OL}^T M_O \Psi\right] - [\lambda_{OL}]^{-1}\left[[\lambda_{OL}]\left[\Phi_{OL}^T M_O\right]\Psi\right] \equiv [0]$$

It can also be proven that mathematically equivalent residual vectors may be derived using the mass matrix as a weighting matrix,

$$[\Psi_R] = [\Psi] - [\Phi_{OL}]\left[\Phi_{OL}^T M_O \Phi_{OL}\right]^{-1}\left[\Phi_{OL}^T M_O \Psi\right] = [\Psi] - \left[\Phi_{OL}\Phi_{OL}^T M_O \Psi\right] \tag{12.30}$$

While the residual vectors are orthogonal to the low frequency modes, they are not necessarily linearly independent of one another. A reduced order linearly independent residual vector set, however, may be estimated using singular value decomposition. This is accomplished by solving the following algebraic eigenvalue problem,

$$[A] = \left[\Psi_R^T M_O \Psi_R\right], \quad [A]\left[\phi_\rho\right] = \left[\phi_\rho\right][\lambda_\rho], \quad \lambda_{\rho 1} \geq \lambda_{\rho 2} \geq \lambda_{\rho 3} \geq \ldots.. \tag{12.31}$$

A suitable cut-off criterion, noted below, that has been employed over the past 10 years with good success in defining the suitable reduced trial vector set, is

$$\frac{\lambda_{\rho N}}{\lambda_{\rho 1}} \leq 10^{-5}. \tag{12.32}$$

The set of linearly independent residual vectors and an augmented trial vector set, respectively, are defined as

$$\left[\Psi_R'\right] = [\Psi_R]\left[\phi_\rho\right] \text{ and } \left[\bar{\Psi}_{OL}\right] = \left[\Phi_{OL} \quad \Psi_R'\right] \tag{12.33}$$

Finally, a mutually orthogonal trial vector set may be defined based on the complete (not truncated) solution of the following reduced eigenvalue problem:

$$\left[\bar{\Psi}_{OL}^T K_O \bar{\Psi}_{OL}\right][\mathit{\Psi}] - \left[\bar{\Psi}_{OL}^T M_O \bar{\Psi}_{OL}\right][\mathit{\Psi}][\lambda_{\mathit{\Psi}}] = [0], \tag{12.34}$$

where

$$[\bar{\Phi}_{OL}] = \left[\bar{\Psi}_{OL}\right][\mathit{\Psi}] \tag{12.35}$$

References

1. Coppolino R (2011) Sensitivity method for uncertainty and reconciliation analysis. In: Proceedings of the 29th international modal analysis (IMAC) conference, Feb 2011, Jacksonville, FL
2. Coppolino R (2002) International space station P5 modal survey: test planning through FEM reconciliation, In: Proceedings of the 20th international modal analysis (IMAC) conference, Feb 2002, Los Angeles, CA
3. Williams D (1945) Dynamic loads on aeroplanes under given impulsive loads with reference to landing and gust loads on a large flying boat, Great Britain Royal Aircraft Establishment Reports, SME 3309 and 3316
4. MacNeal R (1971) A hybrid method for component mode synthesis. Comput Struct 1:581–601
5. Piersol A, Paez T (2010) Harris' shock and vibration handbook (Ch. 23, R. Coppolino), 6th edn. McGraw-Hill, New York, 2010

Chapter 13
Uncertainty Quantification of Weighted Residual Method in Loads Estimation

Colin M. Haynes, Michael D. Todd, and Kevin L. Napolitano

Abstract This paper presents an uncertainty analysis for a new methodology to predict reaction forces in weapon store connections of naval aircraft. The proposed methodology utilizes a strain-gage-based measurement technique in which a series of sensors are affixed to the connector and calibrated with a set of known loading configurations. The calibration matrix relating the measured strains to the loads can then be used to estimate unknown loads from measured strains. In this way, the system can be used to monitor the force through the connector when the aircraft is in service.

The primary objective in this work is to quantify the inherent uncertainties due to noise, miscalibration, or the estimation process itself in this loads estimation approach. First, the relationship between variance on the sensor measurements and the variance on the predicted force is characterized. A technique of sensor fusion called the weighted residual method is used to minimize the variance on force components that are the most important to the user of the system. The effect of thermal loading on the system is also explored via finite element simulation.

Keywords Uncertainty quantification • Loads estimation • Optimal sensor placement

13.1 Introduction

Quantifying the load transfer between two structural components is a challenge that has widespread interest in the structural health monitoring (SHM) community. These connection regions tend to be critical to the safe operation of the structure, and being able to measure experimentally the loads going through them facilitates the application of a damage prognosis strategy. Such load monitoring techniques are in fact a key part of any predictive failure modeling effort, since they can be used to define the operational load profiles to which the connection is subjected. However, the estimation of these load profiles is subject to a substantial degree of uncertainty that must be properly understood and quantified in order to produce meaningful prediction.

A new approach to predicting reaction forces in the weapon store connections of naval aircraft is presented in this paper, along with a methodology for quantifying and prioritizing the various sources of uncertainty in the prediction. The force prediction is accomplished by a strain-based trained network approach that uses known training loads and corresponding measured strains to determine the calibration matrix. This process is modeled using finite element analysis, and an optimal sensor placement scheme is implemented. Finally, several sources of uncertainty are considered – noise on strain gage measurements, noise on force measurements, and uncertainty from uncompensated thermal change – and the contributions of each to the total uncertainty in the force estimate are analyzed.

C.M. Haynes (✉) • M.D. Todd
Department of Structural Engineering, University of California San Diego, San Diego, CA, USA
e-mail: cmhaynes@ucsd.edu

K.L. Napolitano
ATA Engineering, San Diego, CA, USA

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_13,

13.2 Force Prediction Procedure

The proposed methodology for predicting reaction forces is a trained network approach using strain measurements. Using a given sensor arrangement, a set of known training loads (F_{cal}) is applied, resulting in a measured matrix of strains (E_{cal}). For a linear network, this yields a calibration matrix (C) which can be used to predict future unknown forces from measured strains. The procedure is summarized in (13.1):

$$\begin{aligned} F &= C^{*}E \\ F &= F_{cal}{}^{*}pinv(E_{cal})^{*}E \end{aligned} \tag{13.1}$$

where F is the estimated vector of forces, and E is the vector of measured test strains.

The general trained network procedure may also be extended to the nonlinear case by using a tool like an artificial neural network. However, in the present application example, it is not yet assumed that the structure will exhibit sufficiently nonlinear behavior so as to warrant using such an approach, so this paper will consider only the linear case of (13.1).

13.3 Optimal Sensor Placement Procedure

13.3.1 Modeling

In order to develop a scheme for optimally placing a limited number of sensors, it is necessary to develop a model of the structure. A finite element model (FEM) was previously developed for the forward bracket alone [1]. Analysis subsequent to that initial model development revealed that a more global approach taking into account all three connecting brackets must be pursued. As a result, a new FEM was developed as shown in Fig. 13.1. However, the focus at this phase is still to instrument and predict the interface forces for the forward connector.

In order to get an idea of the local strain fields surrounding the forward connector, six unit loads – bounce, lateral, and fore/aft forces, in addition to roll, pitch, and yaw moments (defined in Fig. 13.2) – were separately applied to the connector. The other two connectors were held fixed along the portion that they are clamped to the missile launcher. The surface strains from the areas on which gages may safely be applied were then extracted from the FEM model. The strains are then averaged to over areas equivalent to the strain gages' active gage length (~1 cm) and passed to the optimization routine. Furthermore, since the strain/force relationship is assumed to be linear, any arbitrary force state may be reconstructed by a linear combination of the strain fields obtained in this way.

13.3.2 Optimization Algorithm

The first step in determining an optimal sensor arrangement is deciding upon a metric that quantifies the performance of the given arrangement so that it may be compared to other arrangements in a meaningful way. In this case, the so-called

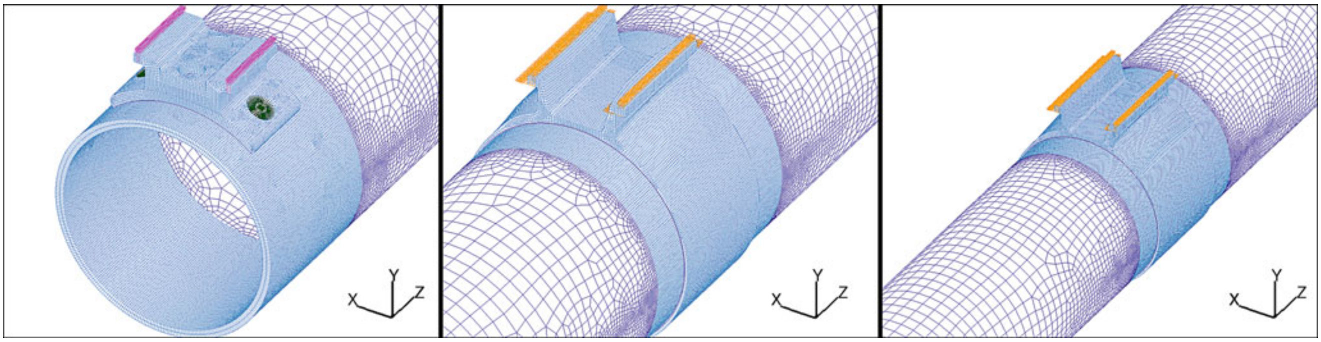


Fig. 13.1 Views of the new FEM of the missile – forward bracket (*left*), middle bracket (*middle*), and aft brackets (*right*). Forces are shown in *pink*, restraints in *orange*

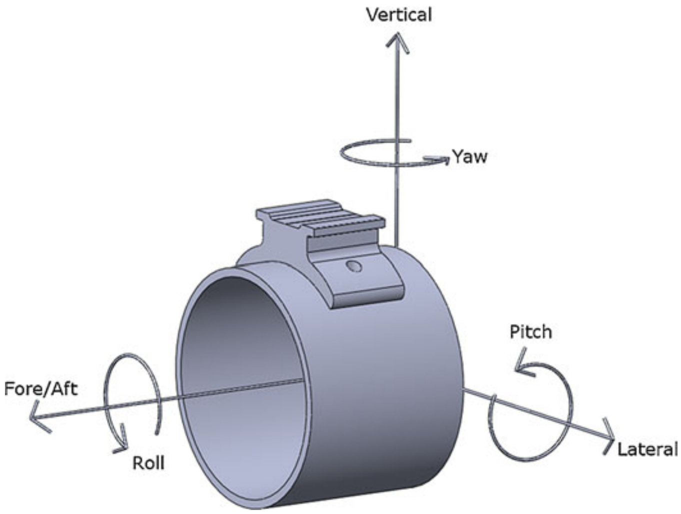


Fig. 13.2 Definition of forces and moments applied to bracket

weighted residual method is used, as presented previously in Ref. [1]. This method is an extension of Kammer's method [2], seeking to minimize the weighted trace of the covariance matrix, or equivalently to maximize the weighted trace of the Fisher information. Extending Kammer's method to the current case, the following model of the force prediction is used:

$$\beta = pinv(C) \tag{13.2}$$

Therefore:

$$\varepsilon = \beta F + N \tag{13.3}$$

$$\hat{F} = \left(\beta^T \beta\right)^{-1} \beta^T \varepsilon \tag{13.4}$$

where $N \sim \mathrm{N}(0, \sigma^2)$ and $\hat{F}$ is the estimated force vector.

For an efficient unbiased estimator, Kammer has already shown that the covariance matrix of $\hat{F}$ is given by

$$P = \left[\beta^T \left(\sigma^2\right)^{-1} \beta\right]^{-1} = Q^{-1} \tag{13.5}$$

where Q is the Fisher information matrix. With the covariance matrix derived, Kammer suggests that the trace is the appropriate matrix norm to minimize. However, rather than using the trace, the diagonal components of the covariance matrix are multiplied by arbitrary coefficients representing their relative importance and subsequently summed. This allows the user of the system to design a sensor arrangement that emphasizes accuracy on the most important load components.

Finally, with the new performance metric defined, a new algorithm for finding the optimal placement of sensors was developed. There are approximately 3,500 possible sensor locations, based on the discretization of the FEM mesh. This FEM size is prohibitively large to implement a simple rote search. Therefore, a hybrid between a genetic algorithm and a forward greedy algorithm was developed. The genetic algorithm is based on a slight modification of the built-in MATLAB toolbox. A greedy algorithm is one in which a sensor is added at each step of the process according to which one improves the performance most at that stage (the locally optimal choice), and iterates until all sensors have been selected [3]. In the hybrid scheme, the genetic algorithm is run to select the first six sensors because that is the minimum number to predict six independent load cases, and thus the lowest number of sensors for which the performance metric is valid. From that point, a forward greedy algorithm is used to select the remaining sensors because of its computational efficiency – running a genetic algorithm on the full number of sensors would be up to a factor of 65 slower and does not necessarily improve performance. Furthermore, in the implementation of this hybrid algorithm, the genetic algorithm mutates such that most of the time, each sensor shifts slightly in space or in orientation, and only a few times is allowed to mutate to a completely new and random location. The starting locations are randomized only within a limited subset of user-selected locations. Both of these modifications cause the algorithm to produce arrangements that are more in line with the new, stringent constraints on sensor location.

At this point, equal weights were applied to all of the load cases, and the algorithm was used to compute a sensor arrangement. The fitness value of the genetic algorithm is plotted versus each generation is shown in Fig. 13.3, and the

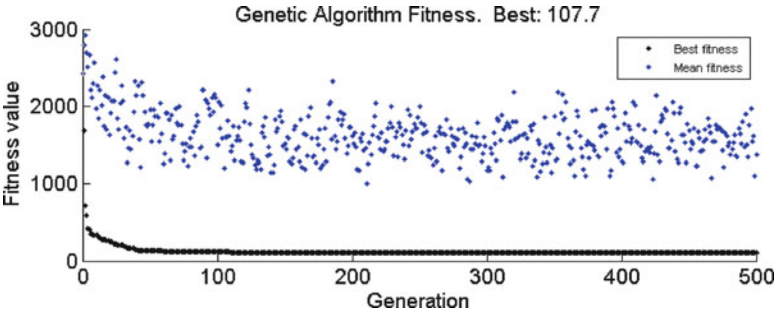


Fig. 13.3 Results of genetic algorithm

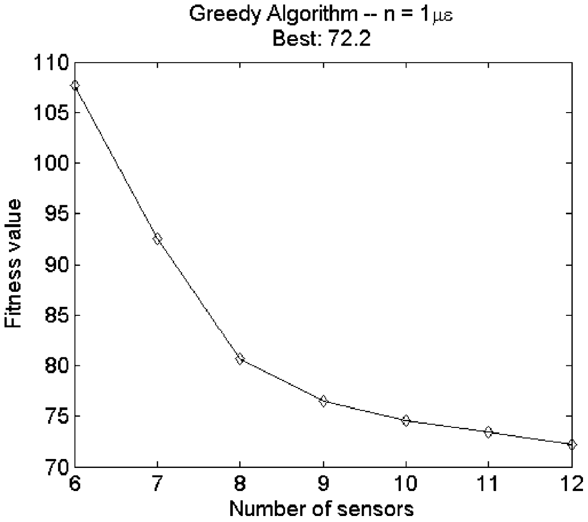


Fig. 13.4 Results of greedy algorithm

subsequent greedy algorithm results (fitness value vs. number of sensors) is given in Fig. 13.4. The final arrangement of the sensors is shown in Fig. 13.5, which contains four different views of the forward bracket. Note that the sensors are numbered by the order they were selected – i.e. the lower the number, the more critical the sensor is to the final force prediction.

13.4 Uncertainty Quantification

Whenever any sort of prediction is being made, it is important to understand the uncertainty inherent in that prediction. For this system, the uncertainty in the location, orientation, and magnitude of the training loads are all of significant concern, since these sources of error will dramatically affect the realized subsequent predictions. The uncertainty in the location of the strain sensors, while not affecting the force prediction directly, makes it difficult to place sensors at the optimal location to maximize the independence of the strain measurements. Finally, the random noise on the sensors, both the strain gages and the force transducer, will contribute significantly to the prediction uncertainty. The model used to represent these sources is given in the following section.

13.4.1 Noise Model

It is necessary to develop a model of the uncertainty of each input in order to begin to estimate the output of the prediction. A very simple model of the noise affecting the measurements is given in (13.6):

$$F_{est} = (F_{cal} + u)^{*} pinv(E_{cal} + v)^{*}(E + w) \tag{13.6}$$

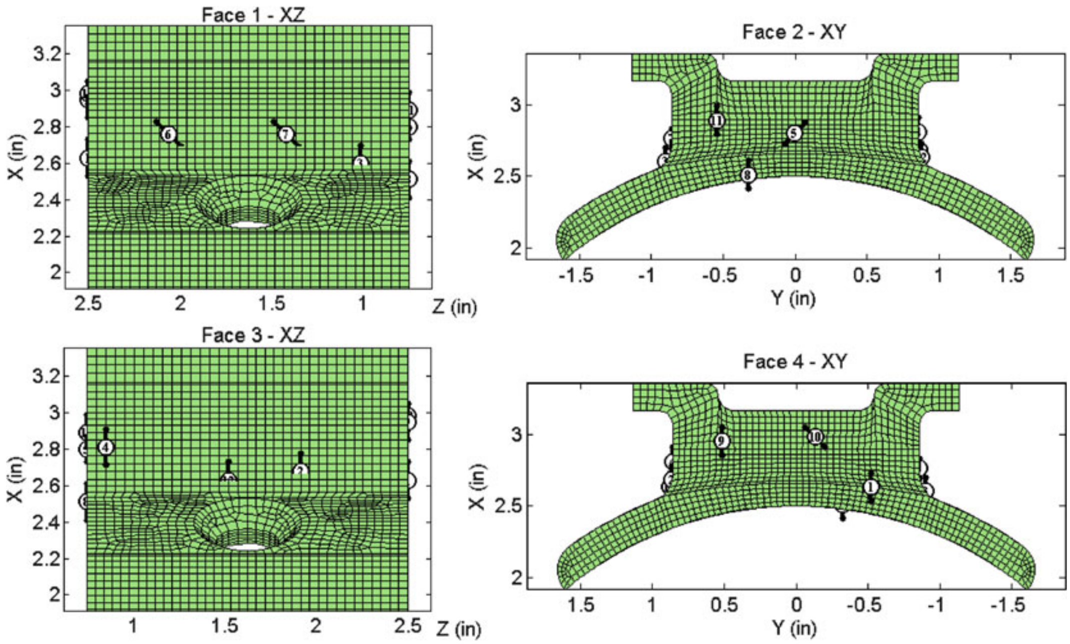


Fig. 13.5 Sensor placement selected by the hybrid algorithm for equal weights on all load components. Sensors are numbered in the order they were selected

Note that (13.6) is a matrix equation. F_{cal} is a vector of calibration forces, E_{cal} and E are the matrices of calibration and test strains, with a size of number of forces x number of strain sensors. F_{est} is the estimated force vector, u is the uncertainty in the measurement of the training forces, and v and w are two realizations of the noise on the strain gage measurement. It is assumed that u, v, and w are normally-distributed, zero-mean, statistically independent additive noise components. Since v and w are associated with the same sensors, it is generally assumed that they have the same variance, which is distinct from the variance of u. It is further assumed that there is no correlation in error between individual sensors, so the covariance matrices of u, v, and w are all diagonal.

The first step is to simply take the expectation of each side of (13.6) and see what can be learned. However, due to the fact that three matrices fully populated with independent random variables are multiplied together, the algebra quickly becomes difficult. It becomes intractable when realizing that the middle matrix is actually the pseudoinverse of a normal random matrix. The distribution of such a matrix is addressed in Ref. [4], but to continue propagating the error through the rest of the matrices proved too difficult a process to be attempted here. Instead, a numerical approach based on Monte Carlo simulations was undertaken to quantify the impact that each source of noise has on the total variance of the force estimate.

13.4.2 Monte Carlo

Numerical simulation can be used to generate good estimates of the uncertainty parameters given a sufficient number of iterations. In order to understand the relative importance of the various sources of error in this system, a Monte Carlo simulation was run using different noise levels for the strain and force sensors. The results, broken down by each load component, are presented in Fig. 13.6. It is clear that the variance in the overall estimate is much more sensitive to noise on the strain sensors than noise on the force sensors. The value of that sensitivity varies from component to component. A separate Monte Carlo simulation was also run using only one noise variable (i.e. u, v, or w) at a time. Thus, the percentage of the total estimation variance represented by each noise variable can be determined, and these percentages are plotted in Fig. 13.7. It is clear that w, the noise on the test strain, is the dominant source of variance, accounting for about 80% of the variance on average. Note that the totals are not quite 100% because the Monte Carlo simulation is approximate.

The results of performing the Monte Carlo analysis on the entire (13.6) are shown in Fig. 13.8. Despite the fact that three normal random matrices are passed through (13.6), which is a nonlinear operation on the noise, the results seem well approximated by the normal distributions overlaid in red. It is also interesting to note that the overall estimated force variance changes with both the mean and variance of the input parameters.

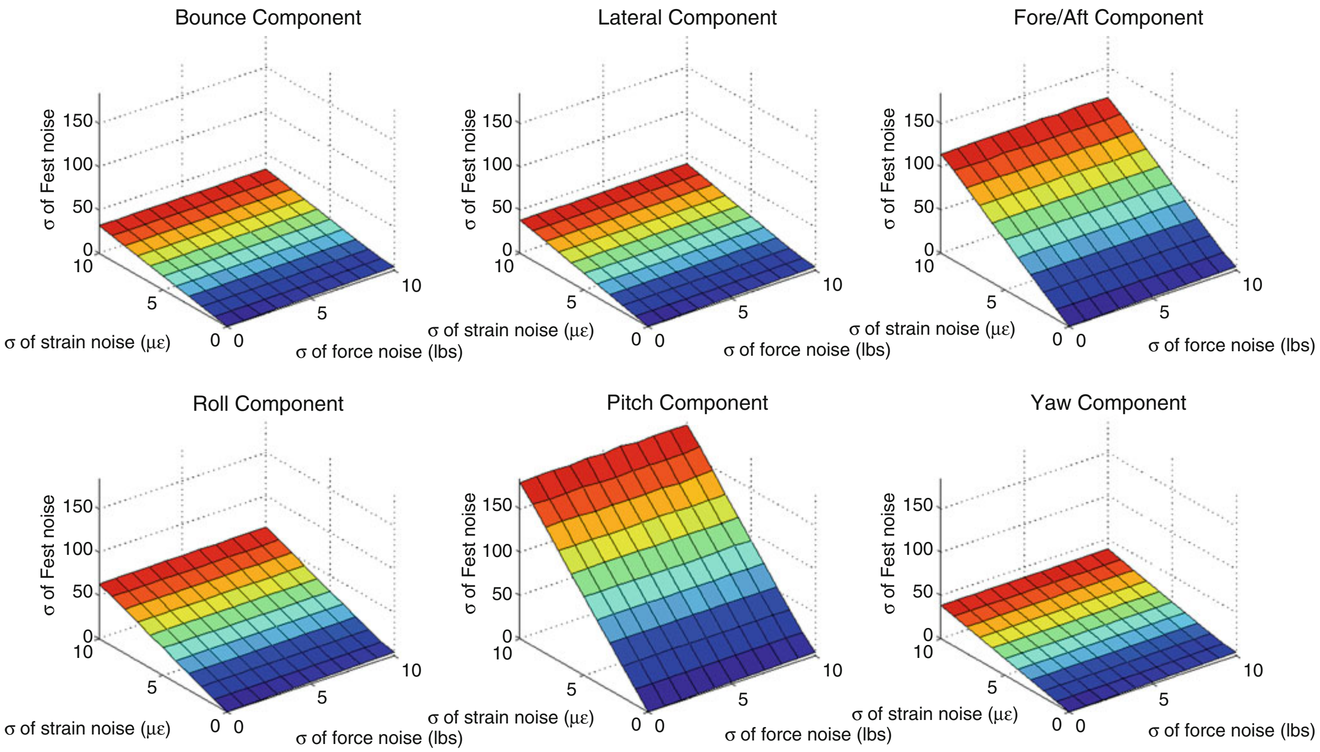


Fig. 13.6 Surface plots of standard deviation of estimation versus standard deviation of strain and force measurements, by load component

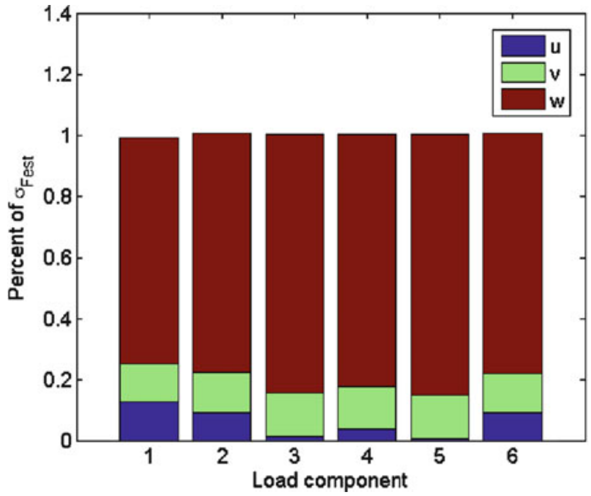


Fig. 13.7 Graphical representation of the contribution of each noise term to the total variance of the estimate

13.4.3 Uncertainty from Temperature Change

Given that this system is designed with in-flight sensing in mind, it is important to quantify the effects that temperature changes will have on the force prediction, if unaccounted for. Therefore, the FE model was used to run a thermal simulation, and the strains for a change of 1°F were output. Using the sensor arrangement obtained from the optimal sensor placement procedure, this strain matrix was fed into the trained network to see what magnitude of forces would be produced. The result is presented in Table 13.1. The magnitude of the errors are quite significant, highlighting the need for effective thermal compensation to be integrated into the procedure. The error tolerance for this load estimation system has been defined as 5%,

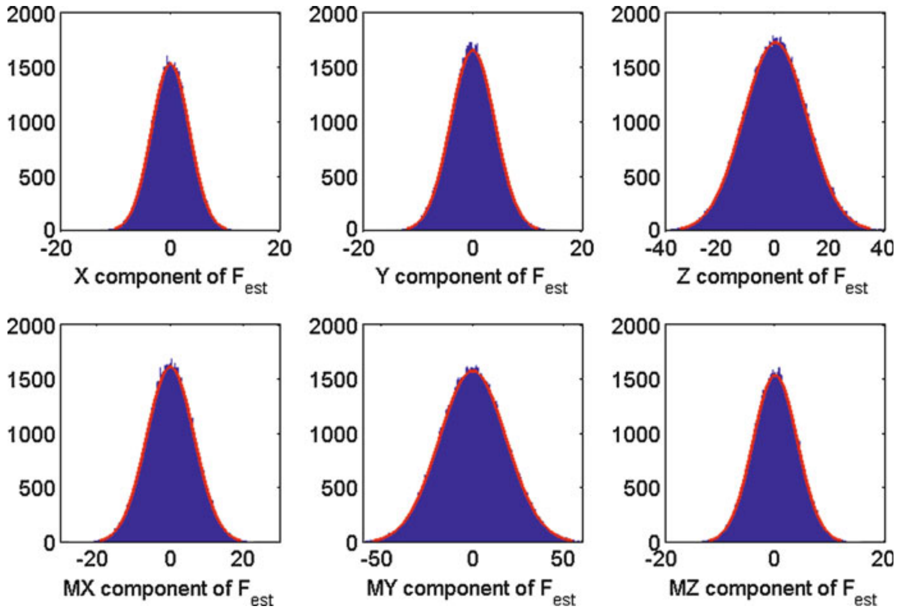


Fig. 13.8 Histograms of the error on each component of the force prediction, from a Monte Carlo simulation of the process

Table 13.1 Force errors resulting from uncompensated thermal change

	X (Lateral)	Y (Bounce)	Z (Fore/Aft)	MX (Pitch)	MY (Yaw)	MZ (Roll)
Force (lb/°F)	−22.82	11.45	−23.20	14.18	−15.93	−8.03

with loads anticipated to go as high as 2,000 lb. That means that if all of the permissible error were to come from uncompensated temperature change, the maximum temperature change would be approximately 4°F. Temperature variation in the flight environment is expected to be much greater, perhaps as much as 130°F – therefore, a thermal compensation system accurate to 4°F or better must be implemented.

13.5 Conclusion

A new loads estimation procedure for structural connectors has been introduced. The principle of using a strain-based trained network was applied to a new FEM of the entire missile and all three connecting brackets. An updated scheme for optimal sensor placement is presented and used to find a sample arrangement.

Finally, the uncertainty of the final force estimation is considered. A model of the noise process was developed, and the uncertainty was propagated through it via Monte Carlo methods. This analysis reveals that the variance of the estimate is much more sensitive to noise on the strain sensors than the force sensors, with approximately 80% of the total variance coming from the noise on the test strain, w. The entire process even appears to fit a Gaussian distribution very well, despite the fact that the noise process involves the nonlinear pseudoinverse function. Results from thermal loading were considered, and it seems clear that thermal compensation will be a necessary component of the loads estimation system.

Acknowledgments This project is funded under a Small Business Technology Transfer program awarded by NAVAIR and is being conducted in collaboration with ATA Engineering in San Diego, California. Special thanks are extended to our partners Ron Merritt and Shawn Hertz at Naval Air Warfare Center Weapons Division (NAWCWD) for supporting this effort.

NAVAIR Public Release 11-1305.

Approved for public release; distribution is unlimited.

References

1. Haynes C, Todd M, Napolitano K, Matt H (2011) Optimal sensor placement and weighted residual method in loads estimation supporting structural health monitoring. In: International workshop on structural health monitoring, Standford
2. Kammer DC (1991) Sensor placement for on-orbit modal identification and correlation of large space structures. J Guid Control Dyn 14 (2):251–259
3. Black PE (2005) Greedy algorithm [Internet]. Dictionary of algorithms and data structures, [cited Aug 22, 2011]. http://www.nist.gov/dads/HTML/greedyalgo.html
4. Zhang Z (2007) Pseudo-inverse multivariate/matrix-variate distributions. J Multivariate Anal 98(8):1684–1692

Chapter 14
Rapid Structural Condition Assessment Using Transmissibility with Quantified Confidence for Decision Making

Zhu Mao and Michael Todd

Abstract The motivation of this paper is to achieve a rapid condition assessment and identification for aero-space structures. For such a circumstance, transmissibility is a good vibration-domain feature for damage detection and localization, highlighting the sensitivity of the relative dynamics between any two interest points on the structure to many forms of damage (such as connection losses). With a quantified uncertainty model of transmissibility available, test results may be compared to baseline (undamaged) statistics, and statistical significance under certain confidence level may be found. Receiver operating characteristic curves suggest the optimized threshold of detection that balances detections and false alarms. In this paper, three such models are presented, and the aforementioned statistical identifications are validated through a plate structure with a single-edge clamped boundary condition. The benchmark test results show very good consistency with the model prediction and the robustness to different noise contamination situations.

Keywords Uncertainty quantification • Statistical modeling • Damage detection • Receiver operating characteristics

14.1 Introduction and Background

In the past couple of decades, transmissibility has been proven to be a good feature for assessing structural condition, because of its potential sensitivity to small and local parameter changes [1, 2]. Among all kinds of damage detection methodologies, changes in transmissibility are often considered as the indicator of damage occurrence. However, in real implementations with different types of variability or uncertainty, the measurements are necessarily noisy, and the transmissibility estimations are non-deterministic. These sources of uncertainty and noise may cause misinterpretation of test results by similarly affecting transmissibility, namely false alarms (Type-I error).

Most of the works involved with uncertainty quantification are focused on the modeling of the power density and input–output transfer function uncertainties [3]. In this paper, three statistical models are proposed to quantify the uncertainty of transmissibility magnitude estimations from fundamental Gaussian random processes. Quantified uncertainty models help to improve the confidence in decision making, where decision thresholds for optimal sensitivity and specificity can be determined for damage detection via receiver operating characteristics.

Z. Mao (✉) • M. Todd
Department of Structural Engineering, University of California San Diego, La Jolla, CA 92093-0085, USA
e-mail: zmao@ucsd.edu; mdtodd@ucsd.edu

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_14,

14.2 Uncertainty Quantification Models

14.2.1 Transmissibility and Estimation

We simplify the data flow into a single-input multiple-output (SIMO) chart shown in Fig. 14.1, in which transmissibility can be mathematically defined as the ratio between two outputs at different locations, in the spectral domain:

$$T_{ij}(\omega) = \frac{V_i(\omega)}{V_j(\omega)}, \tag{14.1}$$

where $x(t)$ is the input, $v_i(t)$ is the response at arbitrary position i, and $V_i(\omega)$ is the Fourier transform of $v_i(t)$. For a more realistic condition, extraneous noise *n(t)* is also considered.

Under random broad-band excitation, Welch's method will usually be adopted to average multiple segments of measurements in order to obtain a better quality of estimation [3, 4]. Equation 14.2 shows the estimation of cross-power density function as an example with n_d times of averages via Welch's algorithm:

$$\hat{G}_{xy}(\omega) = \hat{X}^*(\omega) \cdot \hat{Y}(\omega) = \frac{1}{n_d T} \sum_{k=1}^{n_d} \tilde{X}_k^*(\omega, T) \cdot \tilde{Y}_k(\omega, T). \tag{14.2}$$

From the power spectral estimations, we will have different forms of transmissibility expressions, or estimators, which are more robust and practical compared to the original mathematical definition in (14.1). Three such estimators for transmissibility magnitude are shown in (14.3):

$$\left|\hat{T}_{ij}\right| = \frac{\left|\hat{G}_{xy_i}\right|}{\left|\hat{G}_{xy_j}\right|} = \frac{\left|\hat{H}_i\right|}{\left|\hat{H}_j\right|} = \sqrt{\frac{\hat{G}_{y_i y_i}}{\hat{G}_{y_j y_j}}}. \tag{14.3}$$

14.2.2 Statistical Model of Transmissibility Magnitude

Previous works [5, 6] have presented more detailed discussions of the three estimators in (14.3) and have compared their accuracy under various modeling assumptions. In the first approach, a perturbation approach is used, and (14.4) gives the approximate variance of transmissibility magnitude estimation in terms of the coherence function $\gamma^2_{y_i y_j}$ between two output responses

$$\mathrm{Var}\left[\left|\hat{T}_{ij}\right|\right] \approx \frac{\left(1 - \sqrt{\gamma^2_{y_i y_j}}\right)}{\sqrt{\gamma^2_{y_i y_j}} \cdot n_d} \left|T_{ij}\right|^2, \tag{14.4}$$

where it was assumed that the coherences were balanced, i.e. $\gamma^2_{v1y2} = \gamma^2_{y1v2} = \sqrt{\gamma^2_{y1y2}}$. This model gives the order statistics of transmissibility estimation and approximately quantifies the uncertainty bounds without sophisticated probability density derivation.

For the second estimator in (14.3), $\left|\hat{T}_{ij}\right| = \left|\hat{H}_i\right| / \left|\hat{H}_j\right|$, exact probability density function was derived, (14.5), where the frequency response functions are regarded as Gaussian random variables themselves:

$$\begin{aligned} f(z) = {} & \frac{1}{\pi} \cdot \frac{\sigma_i \sigma_j \sqrt{1 - r^2}}{\sigma_j^2 - 2rz\sigma_i\sigma_j + z^2\sigma_i^2} \cdot e^{-\frac{1}{2}\cdot\frac{1}{1-r^2}\left(\frac{\mu_i^2}{\sigma_i^2} - 2r\frac{\mu_i\mu_j}{\sigma_i\sigma_j} + \frac{\mu_j^2}{\sigma_j^2}\right)} \\ & + \frac{\sigma_j\left(r\mu_j\sigma_i - \mu_i\sigma_j\right) + z\sigma_i\left(r\mu_i\sigma_j - \mu_j\sigma_i\right)}{\pi\sqrt{\left(\sigma_j^2 - 2rz\sigma_i\sigma_j + z^2\sigma_i^2\right)^3}} \cdot e^{-\frac{1}{2}\cdot\frac{\left(\mu_j - z\mu_i\right)^2}{\sigma_j^2 - 2rz\sigma_i\sigma_j + z^2\sigma_i^2}} \cdot \int_0^{\frac{\sigma_j\left(r\mu_j\sigma_i - \mu_i\sigma_j\right) + z\sigma_1\left(r\mu_i\sigma_j - \mu_j\sigma_i\right)}{\sigma_i\sigma_j\sqrt{\left(1-r^2\right)\left(\sigma_j^2 - 2rz\sigma_i\sigma_j + z^2\sigma_i^2\right)}}} e^{-\frac{1}{2}u^2}\, du \end{aligned} \tag{14.5}$$

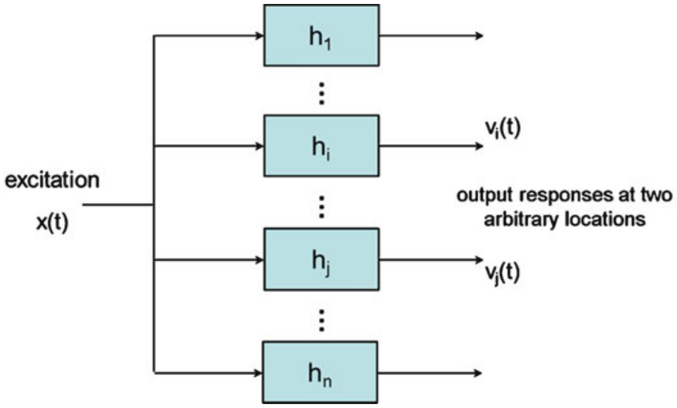


Fig. 14.1 SIMO signal flow

where

$$\mu_i = \mathrm{E}\left[\left|\hat{H}_i\right|\right] = \frac{G_{xv_i}}{G_{xx}} \quad \text{and} \quad \mu_j = \mathrm{E}\left[\left|\hat{H}_j\right|\right] = \frac{G_{xv_j}}{G_{xx}};$$

$$\sigma_i = \varepsilon\left[\left|\hat{H}_i\right|\right] \cdot \mu_i \quad \text{and} \quad \sigma_j = \varepsilon\left[\left|\hat{H}_j\right|\right] \cdot \mu_j;$$

$$\varepsilon\left[\left|\hat{H}_i\right|\right] \approx \sqrt{\frac{1-\gamma_{xy_i}^2}{2 \cdot n_d \cdot \gamma_{xy_i}^2}} \quad \text{and} \quad \varepsilon\left[\left|\hat{H}_j\right|\right] \approx \sqrt{\frac{1-\gamma_{xy_j}^2}{2 \cdot n_d \cdot \gamma_{xy_j}^2}}.$$

This approach models the transmissibility magnitude in a Gaussian bivariate way, and ends up with the expression in error function. Instead of giving only order statistics as in the perturbation approach, this approach gives expression of entire probability density, and the distribution of estimation is fully described.

A third approach uses the magnitude estimator $\left|\hat{T}_{ij}\right| = \sqrt{\hat{G}_{y_i y_i}/\hat{G}_{y_j y_j}}$, and it can be shown that the auto-power density estimations is theoretically Chi-square distributed if the measurement series is Gaussian. Therefore, the exact probability density function for this estimation can be derived through a Chi-square bivariate approach as shown in (14.6):

$$p_t(\hat{T}) = \frac{2^{2n_d}\hat{T}^{(2n_d-1)}\left((1-\rho^2)\sigma_{Y_i}^2\sigma_{Y_j}^2\right)^{n_d}\Gamma\left(n_d+\frac{1}{2}\right)}{\sqrt{\pi}\left(\sigma_{Y_i}^2+\hat{T}^2\sigma_{Y_j}^2\right)^{2n_d}\left(1-\frac{4\hat{T}^2\rho^2\sigma_{Y_i}^2\sigma_{Y_j}^2}{\left(\sigma_{Y_i}^2+\hat{T}^2\sigma_{Y_j}^2\right)^2}\right)^{n_d+\frac{1}{2}}\Gamma(n_d)}, \tag{14.6}$$

where the symbols $\sigma_{Y_i}^2$, $\sigma_{Y_j}^2$ represent the variance of complex Gaussian random variables Y_i and Y_j, ρ is the correlation coefficient, and $\Gamma(\cdot)$ is Gamma function. This uncertainty model describes the transmissibility magnitude uncertainty in the most accurate way, but as addressed in Ref. [5], for specific noise-contaminated conditions, the estimator itself fails for low signal-to-noise levels.

14.3 Plate Structure Experimental Validation

The uncertainty quantification models in (14.4)–(14.6) are compared and validated by experimental structural tests. Because this validation is deployed in a laboratory, most of the uncertainties come from the sensing and calculation, without significant influence from external influences of environmental and operational variability. For a better simulation of in-situ test condition, we intentionally add extraneous noise to the measured time series in order to validate the statistical models under a more strict condition.

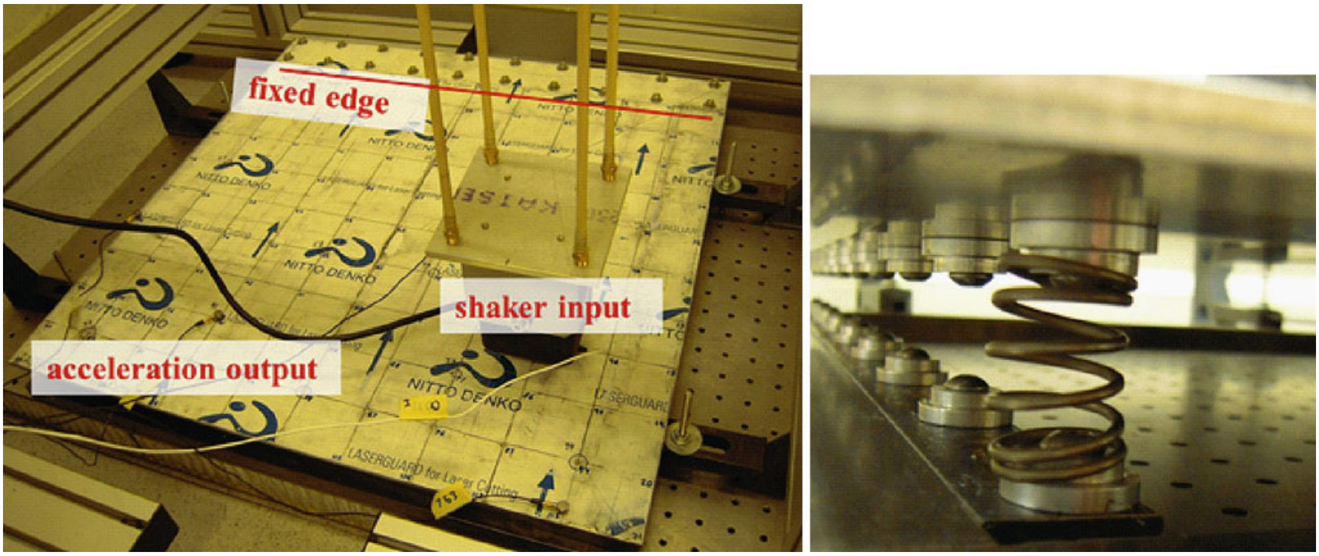


Fig. 14.2 Test structure and spring damage

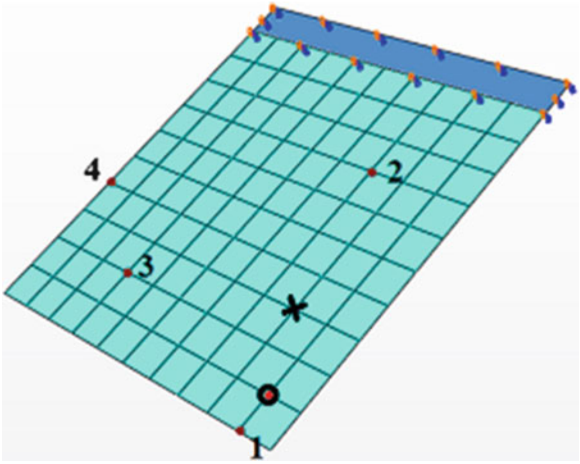


Fig. 14.3 Illustration of plate test bed, *cross*: position of excitation; *circle*: damage location; *dots*: sensing positions

14.3.1 Test Structure

Validation is implemented on a clamped aluminum plate shown in Fig. 14.2, and by placing an additional external spring under the plate, the structure can be regarded as “damaged” under stiffness change. Gaussian noise is applied on to the plate, and four channels of acceleration responses are recorded. Figure 14.3 shows a mesh model of the structure with one fixed end boundary condition and all the positions of the shaker, sensors, and spring addition.

14.3.2 Test Results

Several transmissibility estimations between position 2 and 4 are plotted in Fig. 14.4, with corresponding confidence bounds calculated from the aforementioned three uncertainty models. The left column of Fig. 14.4 represent the extraneous noise

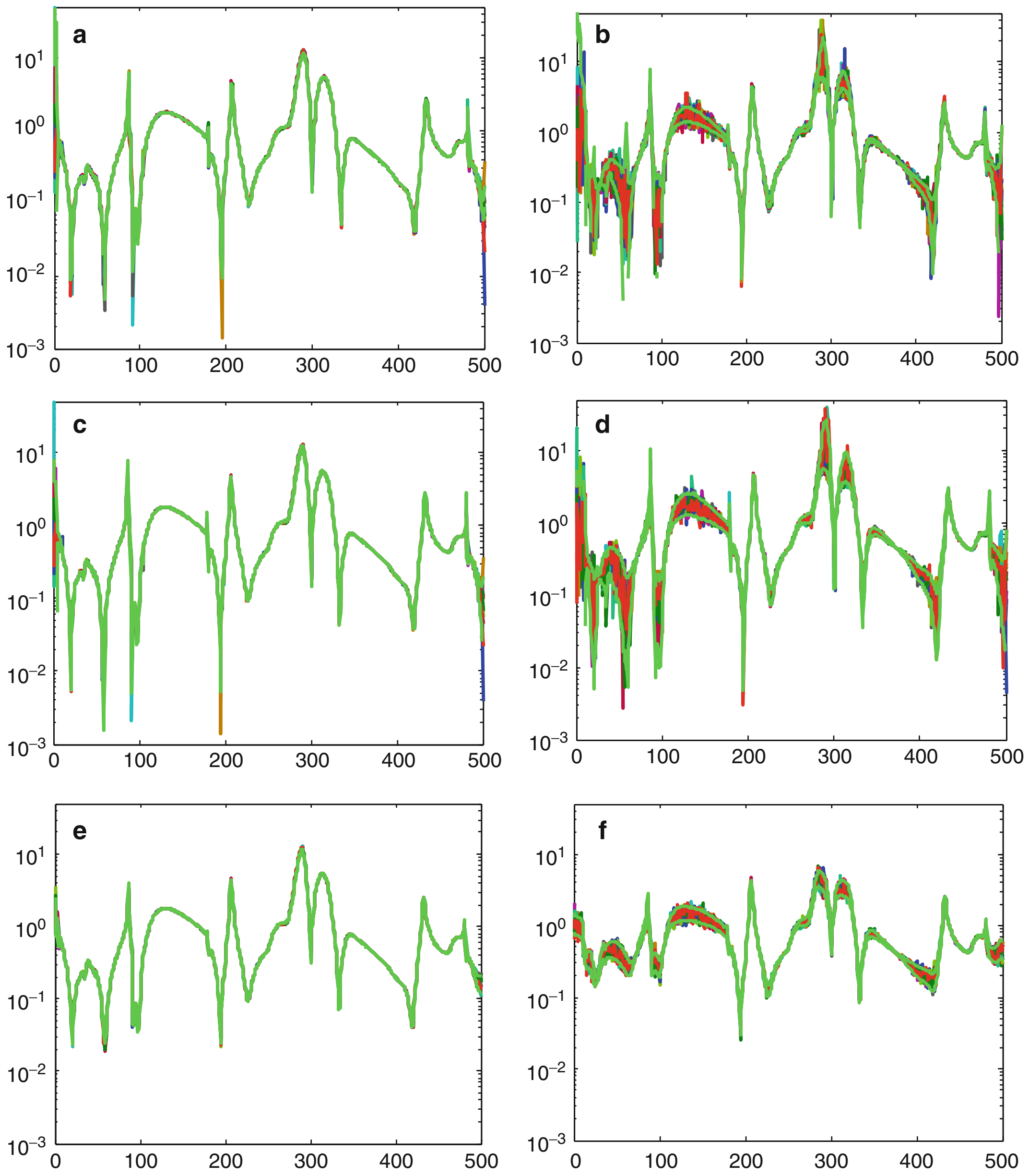


Fig. 14.4 Transmissibility estimation between position 2 and 4, with confidence bounds plotted. (**a**) ±1σ bounds, perturbation approach, noise free. (**b**) ±1σ bounds, perturbation approach, noise contaminated. (**c**) 90% confidence bounds, Gaussian bivariate, noise free. (**d**) 90% confidence bounds, Gaussian bivariate, noise contaminated. (**e**) 90% confidence bounds, Chi-square bivariate, noise free. (**f**) 90% confidence bounds, Chi-square bivariate, noise contaminated

free condition, which means the uncertainty is primarily from measurement noise and the estimation algorithm itself; as a result, the estimations are very accurate, and all the confidence bounds are very close to each other. The right column of plots in Fig. 14.5 represent the condition with surrogate Gaussian white noise added in order to simulate a more realistic environment with other sources of ambient noise and variability. For the noise- contaminated conditions, transmissibility magnitude estimations are corrupted with obviously larger uncertainty, but the quantification models also capture the change and reflect wider uncertainty intervals.

One observation from Fig. 14.4f is that the estimator involved with output auto-power density function does not require information from input. Consequently, when the extraneous noise is added onto outputs, the estimation gets much poorer without the reference from input series, and peaks sharp peaks and notches are not observed as in Fig. 14.4b, d.

Specific performance comparison among the methods may be achieved by comparing experimental outliers to predicted outliers (based on percentage). A 500-round Monte Carlo test was implemented, and at every frequency line the uncertainty

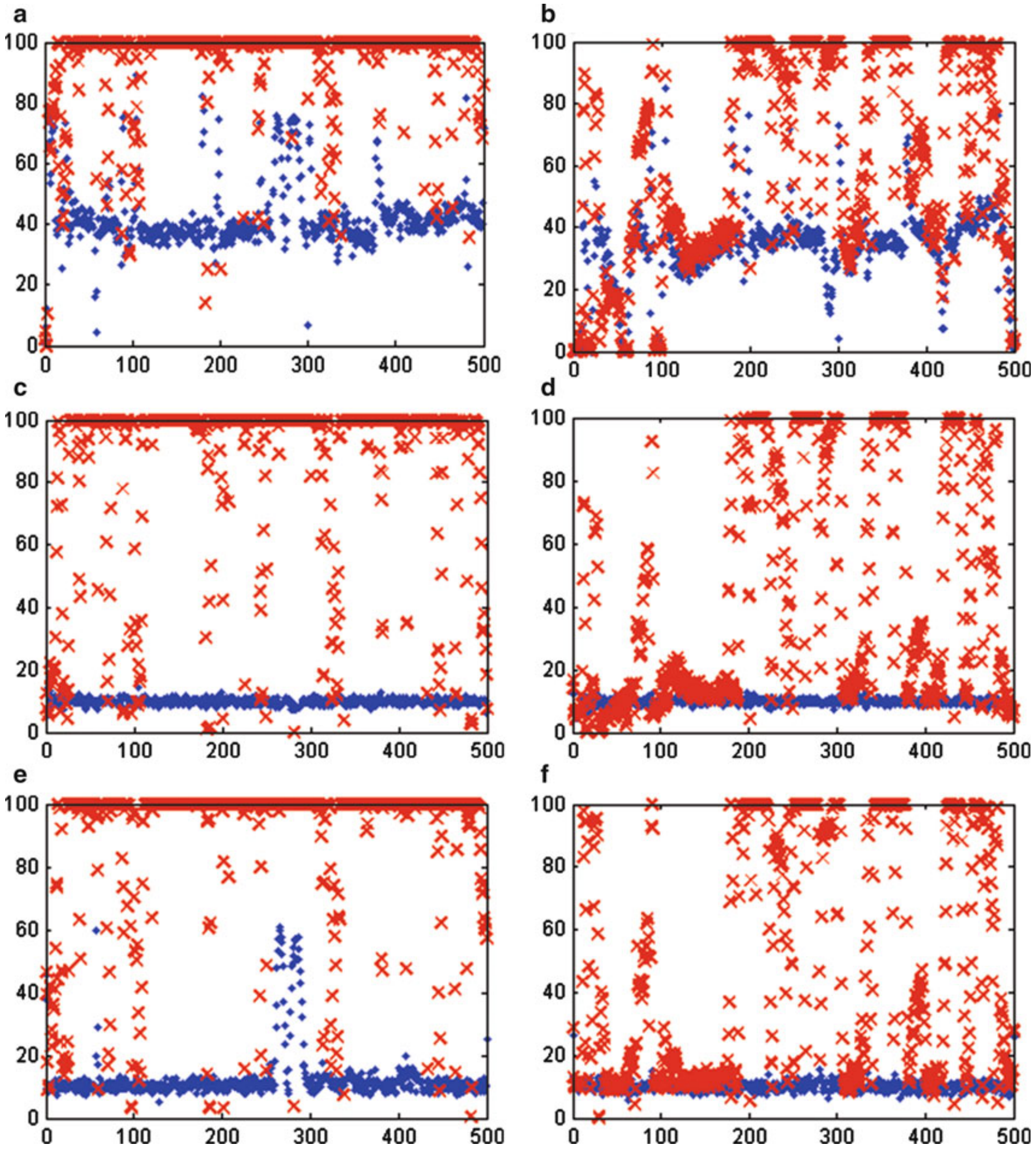


Fig. 14.5 Transmissibility estimation outlier percentage at each frequency line, *dot*: undamaged condition; *cross*: damaged condition. (**a**) Approach, noise free. (**b**) Perturbation approach, noise contaminated. (**c**) Gaussian bivariate, noise free. (**d**) Gaussian bivariate, noise contaminated. (**e**) Chi-square bivariate, noise free. (**f**) Chi-square bivariate, noise contaminated

models were evaluated. The percentage of outliers was calculated and plotted in Fig. 14.5. For statistical models established via Gaussian and Chi-square bivariate approaches, probability density functions are available, and we set the confidence level to be 90%, anticipating 10% outliers. For the statistical model given by perturbation approach, we have only standard deviation of estimations available, without full knowledge of the actual underlying distribution; therefore we plot in Figs. 14.4a, b and 14.5a, b the estimations and outlier percentages bounded by one standard deviation limits for fair comparison.

From the outlier percentage plots in Fig. 14.5, we can conclude that all the uncertainty models are stable in different frequency lines and different noise contamination conditions, and all are sensitive to capturing the structural parameter changes. For extraneous noise-free conditions, a dramatic increase of outlier observations is present, no matter which quantification model is applied. When the measurements are suffered to extraneous noise, the outlier detection can still capture most of the statistical significances.

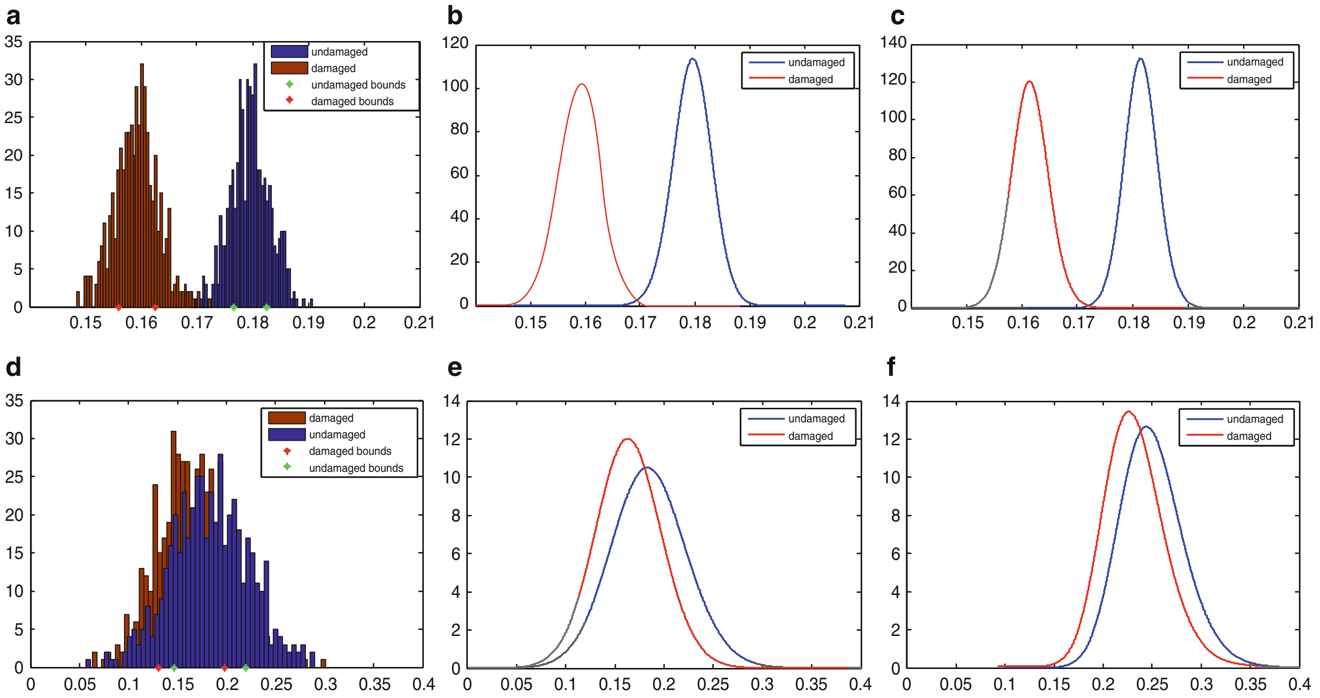


Fig. 14.6 Statistical model visualization for transmissibility estimation at an arbitrary frequency. (**a**) Histogram observation and 1-σ bounds of perturbation model, noise free. (**b**) Probability density function of Gaussian bivariate model, noise free. (**c**) Probability density function of Chi-square bivariate model, noise free. (**d**) Histogram observation and 1-σ bounds of perturbation model, noise contaminated. (**e**) Probability density function of Gaussian bivariate model, noise contaminated. (**f**) Probability density function of Chi-square bivariate model, noise contaminated

There is a very narrow frequency region below 300 Hz for the noise free condition in Chi-square outlier percentage plot (Fig. 14.5e) where baseline values are not as consistent with the pre-set confidence threshold. This is because of a violation of the stationary Gaussian time series requirement, and therefore weak satisfaction of Chi-square characteristics in auto-power density magnitude. By increasing the length of measurement or adding extraneous Gaussian noise, the stationarity will become much better, and this can be seen from Fig. 14.5f.

From the entire frequency domain point of view, we can draw the conclusion that Gaussian bivariate approach is the most stable/reliable method in quantifying the confidence. However, once extraneous noise is present in measurements or the record is long enough to show stationarity, the Chi-square model also has very good performance.

14.3.3 Receiver Operating Characteristics

Instead of looking at the entire frequency domain for a global damage detection purpose, we sometimes focus on limited frequency bands in order to distinguish the statistical significance between two distributions. Figure 14.6 shows the visualization of each model for damaged/undamaged structure at an arbitrary frequency line, with both noise-free and noise-contaminated data. For each model, there is overlap between the two distributions, although it may appear that the chi-square model performs the best purely based on visual observation.

In order to quantify the detection performance, we will calculate and compare the true positive rate (true detection) versus false positive rate (false alarm), which is called receiver operating characteristics (ROC).

Receiver operating characteristics for the two approaches using analytical probability density functions are plotted in Fig. 14.7. The two curves with noise contamination are much closer to the random guess line (45° line) than the two noise-free ROC curves, as would be expected. Comparing the results between the models, we can see that Chi-square results are furthest from the random guess line, which means for a certain number of false alarms there will be more true detections. Therefore, different from the conclusion we made in Sect. 14.3.2 for entire frequency domain, we see a more sensitive and specific detection at single frequency line through Chi-square bivariate distribution modeling.

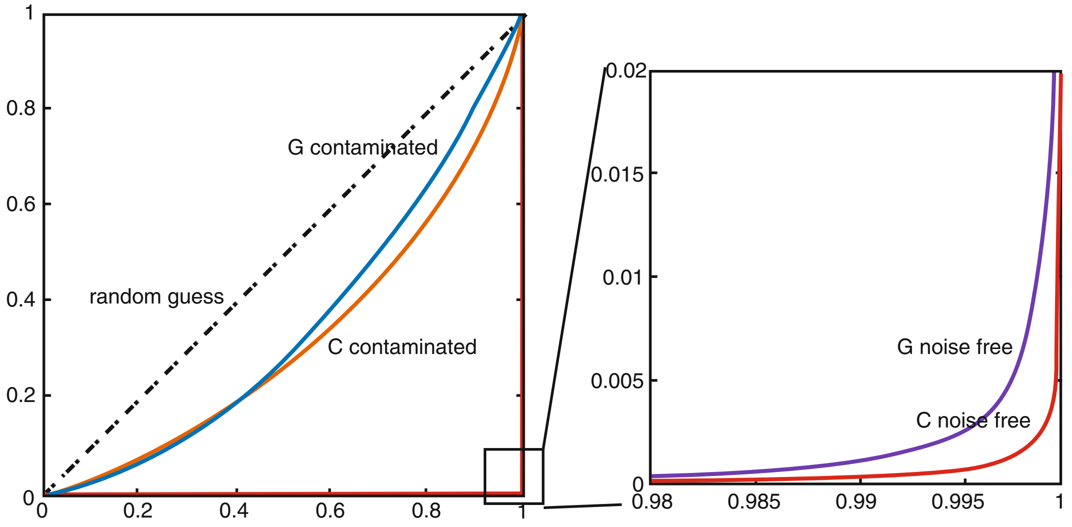


Fig. 14.7 Receiver operating characteristics (ROC) for Gaussian bivariate (G) detection and Chi-square bivariate (C) detection

14.4 Conclusion

In this paper, three models of transmissibility magnitude uncertainty quantification are investigated and compared experimentally using shaker test on a clamped plate. To be more realistic, the robustness of the quantification models is also tested with extraneous noise added. One model gives variance of the estimations, and the other two describe the full probability density characteristics. All the models are validated and compared quantitatively via outlier percentage observations, and test results show a very good consistency between the observed percentage and corresponding pre-set confidence level. These models, with different computational complexity and accuracy, all successfully quantify the transmissibility uncertainty, and detect the change of transmissibility in a more quantified statistical way so that help to improve the quality of decision making in the applications of structural health monitoring. Receiver operating characteristics suggest a quantitative sensitivity and specificity option that certain detect threshold can be given by which the detection quality is optimized to trade-off between true detections and false alarms.

Acknowledgments This material is based on research sponsored by the Army Research Laboratory and managed by the Air Force Research Laboratory, Space Vehicles Directorate, under agreement number FA9453-09-1-0315 P00001. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Army Research Laboratory, Air Force Research Laboratory, or the U.S. Government.

References

1. Worden K (1997) Structural fault detection using a novelty measure. J Sound Vib 201(1):85–101
2. Adams DE, Johnson TJ (2002) Transmissibility as a differential indicator of structural damage. Trans ASME J Vib Acoust 124(4):634
3. Bendat JS, Piersol AG (1986) Random data: analysis and measurement procedures, 2nd edn. Wiley, New York
4. Welch PD (1967) The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoust 15:70–73
5. Mao Z, Todd MD (2010) A structural transmissibility measurements-based approach for system damage detection. Proc SPIE 7650:76500G–76500G-11
6. Mao Z, Todd MD (2011) A model of uncertainty quantification in the estimation of noise-contaminated transmissibility measurements for system identification. In: Proceedings of 29th society of experimental mechanics (SEM) international modal analysis conference (IMAC – XXIX), Jacksonville

Chapter 15
Simulating Unbalance for Future IVHM Applications

Ryan Walker, Sureshkumar Perinpanayagam, and Ian Jennions

Abstract Unbalance is among the most common mechanical faults in rotating machinery, and is of particular interest to the aviation industry. A state of the art machine fault simulator has been used in order to recreate a range of unbalance faults which have been studied in detail from the perspective of fault localisation. High fidelity finite element models have been created in NASTRAN NX and experimentally validated against results for the machine fault simulator. The applicability of such simulations is discussed from an IVHM perspective, along with the potential for such research to influence the development of future engine health management with respect to improved safety and maintenance.

Keywords Rotordynamics • Localisation • Simulation • IVHM

15.1 Introduction

As the aviation industry continues to work towards implementing complete IVHM (Integrated Vehicle Health Management) systems, the need for accurate fault diagnosis and localisation techniques for industrial applications is becoming increasingly urgent. One of the key areas of interest is rotating machinery, specifically gas turbines. Rotordynamics is a subject which has been studied in detail for many years; however techniques for the diagnosis and, particularly, localisation of faults have yet to be widely implemented in the aviation industry.

Gas turbines can suffer from a wide range of rotordynamic faults, from bearing failures to fluid induced instability. One of the most prevalent (and potentially serious) faults is unbalance [1]. Whilst unbalance faults in rotating machinery have been extensively studied from the perspective of diagnosis ([2, 3] being two of many recent examples), techniques for the localisation of this fault are little studied [4].

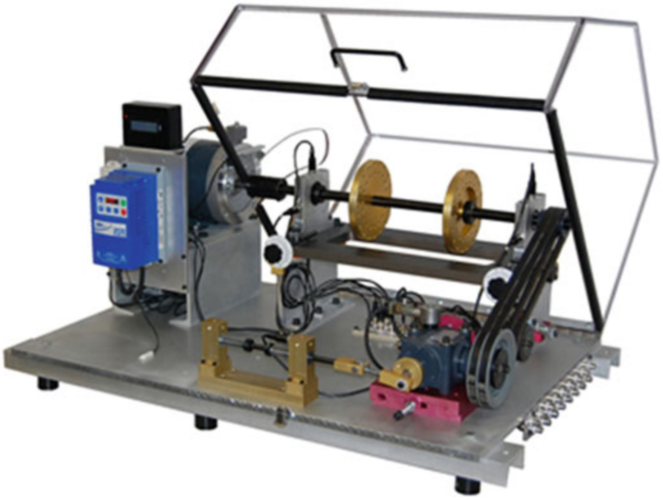
The object of this paper is to describe studies into localisation of a range of unbalance faults from a simulation perspective. As 'seeding' unbalance faults into gas turbines is difficult and costly a state of the art machine fault simulator (MFS) has been used as the basis of the studies and for model validation. The machine fault simulator (see Fig. 15.1) enables unbalance weights to be added and removed quickly and easily, and a number of other faults can also be simulated for study. Simulation of the MFS has been performed using NASTRAN NX finite element software. This paper takes reference from and builds upon the work by Machorro-López et al. [5], who used a similar setup to study damaged shafts from a simulation perspective.

NASTRAN finite element models have been constructed and validated against the MFS, this takes the form of extensive hammer testing to identify modes of both the rotating parts and the support structure. In addition to this the complex eigenvalue solution for comparison against the rotating system with and without unbalance faults at various speeds has been considered. Based upon these experimental results, the FEA models were updated, validated and then expanded in order to study fault localisation methods not possible/practical to do experimentally – even on the MFS, thus taking full advantage of the benefits of simulation.

R. Walker (✉) • S. Perinpanayagam • I. Jennions
IVHM Centre, Cranfield University, Conway House, Medway Court, University Way, Cranfield Technology Park, Cranfield MK43 0FQ, UK
e-mail: r.b.walker@cranfield.ac.uk

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_15,

Fig. 15.1 Spectraquest machine fault simulator



The initial studies of this ongoing project which are outlined in this paper are concentrated on mode shapes and modal frequencies, with the aim of fully understanding the system and the effects of rotor unbalance for the purposes of localisation. Based upon the results collated here, a discussion is included that is aimed at outlining future work into fault localisation, including design optimisation for future IVHM systems and considerations required for implementing fault diagnosis and localisation techniques into industrial IVHM applications.

15.2 High-Fidelity Model Validation

In order to correctly validate finite element models for further simulation, a high degree of accuracy is required (at least initially), involving a combination of accurate material properties, precise dimensions and appropriate constraints. Of particular relevance to rotating machinery are the bearing points, and providing stiffness and damping factors which sufficiently represent the system in question. In order to achieve the desired aims, the modelling procedures outlined in Ref. [5] provided a starting point, from which increasingly accurate models have been developed by incorporating accurate geometry and further developing bearing stiffness and damping factors for validation by hammer testing. Figure 15.2 displays the finite element model in meshed and constrained format. Initial models began with a simple constraint system on the bearing points (five fixed degrees of freedom, one free rotational). This was then developed by the addition of one-dimensional stiffness/damping elements linking the shaft with the fixed constraints.

Hammer testing was performed in free-free and fixed-fixed configurations, with the first ten modes identified from the finite element model initially studied (resulting in a frequency range of 0–1,000 Hz). Figure 15.3 demonstrates an example correlation between the experimental hammer test and the FEA (finite element analysis) results. Table 15.1 displays the first ten modes from fixed-fixed simulation configuration against the hammer test results. Not all predicted modes from the FEA analysis are visible in the hammer test, however those that could be correlated were catalogued and used to tune the finite element model in order to produce the results displayed in Table 15.1. A similar study was produced using a free-free setup for the purposes of validating the initial rotor model before applying the bearing constraints. By altering accelerometer positions around the system it was determined that the detected mode at 110 Hz was the first bending mode (shown in NASTRAN simulation in Fig. 15.3), the third bending mode was detected to be at 353 Hz which displayed the clearest indication of any mode detected by hammer test.

After the rotor-shaft system was validated for normal operating conditions, a series of faults were induced into the model and a complex eigenvalue analysis performed in order to study the effects on the rotating system. Experimental results are detailed in a spectrogram (Fig. 15.4) for normal (left) and faulty (5.8 g static unbalance, right) operating conditions. These experimental results display frequency shifts which can be correlated to those predicted from the complex eigenvalue simulation analysis, in this case the rotational speed was 25 Hz. Table 15.2 displays the results of the complex eigenvalue analysis frequency shifts from normal to faults conditions for a rotational speed of 25 Hz.

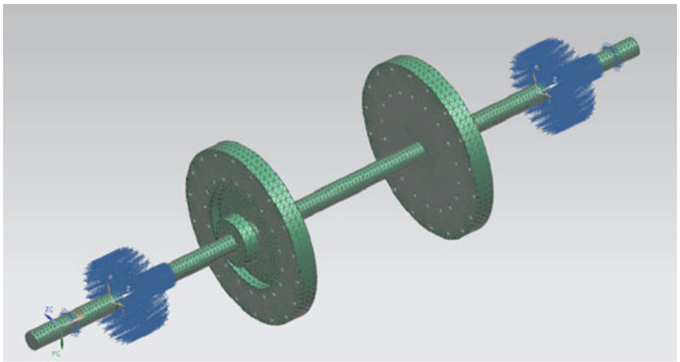


Fig. 15.2 High-fidelity FEA model

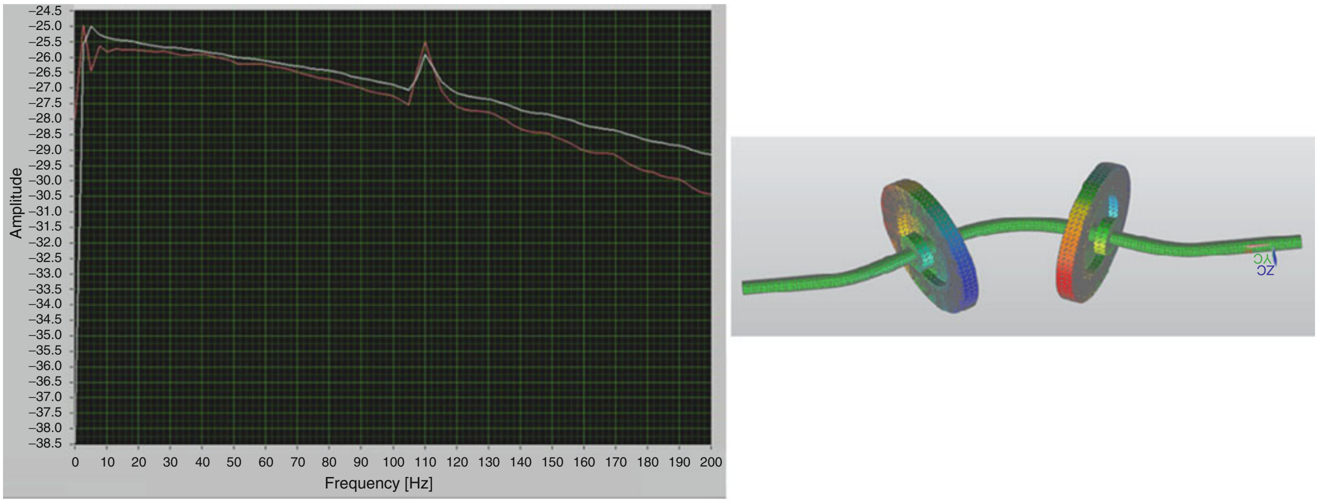


Fig. 15.3 Hammer test experimental and simulation correlation

Table 15.1 Hammer test experimental and simulation correlation

Mode number	NASTRAN NX (Hz)	Fixed-fixed hammer test (Hz)
1	112.4	110
2	157.9	Not detected
3	158.0	Not detected
4	196.2	192
5	360.1	353
6	360.1	353
7	581.8	580
8	581.9	580
9	635.3	640
10	635.4	640

Repetitions of these experiments at different rotational speeds were used to validate the simulations against results from the MFS for the frequency range of 0–1,000 Hz, covering the first ten modes. Building upon these results, a finite element model of the MFS structure was created in order to identify structural modes and provide a rounded view of the experimental information collected from the experimental results. The structural model can be viewed in Fig. 15.5.

Due to the large number parts connected to the support structure (gearbox, electric motor, reciprocating 'piston' etc.) identifying all of the structural modes would be a complex and unnecessary task. Despite this, identifying the key structural modes present in experimental studies was important so that they could be separated from the behaviour of the rotating

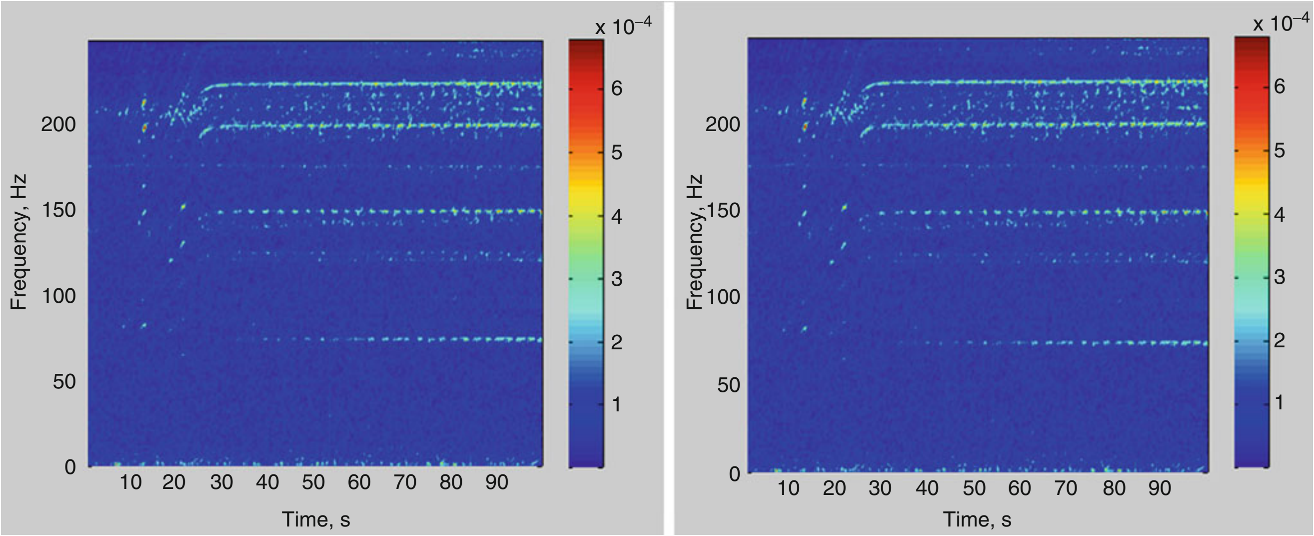


Fig. 15.4 MFS 25 Hz spectrograms without fault (*left*) and with fault (5.8 g unbalance, *right*)

Table 15.2 Modal frequency shifts from normal operating conditions for 5.8 g

Mode	1	2	3	4	5	6	7	8	9	10
Frequency shift (Hz)	1.16	1.50	1.30	2.7	8.9	4.8	2.5	1.4	4.8	4.0

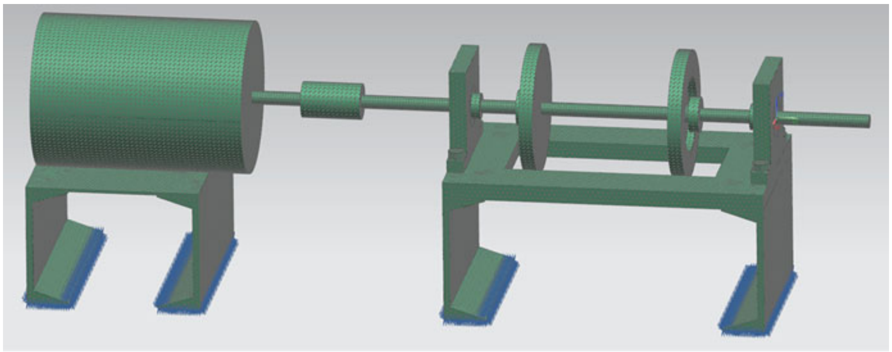


Fig. 15.5 Full MFS finite element model

system. With this in mind a number of finite element studies were performed using the aforementioned model, intended to interpret structural modes identified from hammer testing. In Fig. 15.6, two accelerometers were placed on the structure, one on the motor-side bearing housing (white) and one on the centre of the support plate (red). In this case the complex nature of the response can be seen, however from matching with the simulation results, structural bending modes were identified at 140 and 560 Hz. The frequency range of 0–1,000 Hz was again studied for the purposes of this test, with repeat experiments carried out using accelerometers at different positions around the support structure.

Using the data collected from experimental studies, the final calibration of the finite element models could be performed. This in turn led to the simulations assisting in identifying the phenomena which were observed from the experimental studies. The result indicated that the final set of finite element models could be considered to be validated to a suitable degree of accuracy, taking into account the desire to equate numbers between simulation and experiment along with the need to use simulations to produce future trending data.

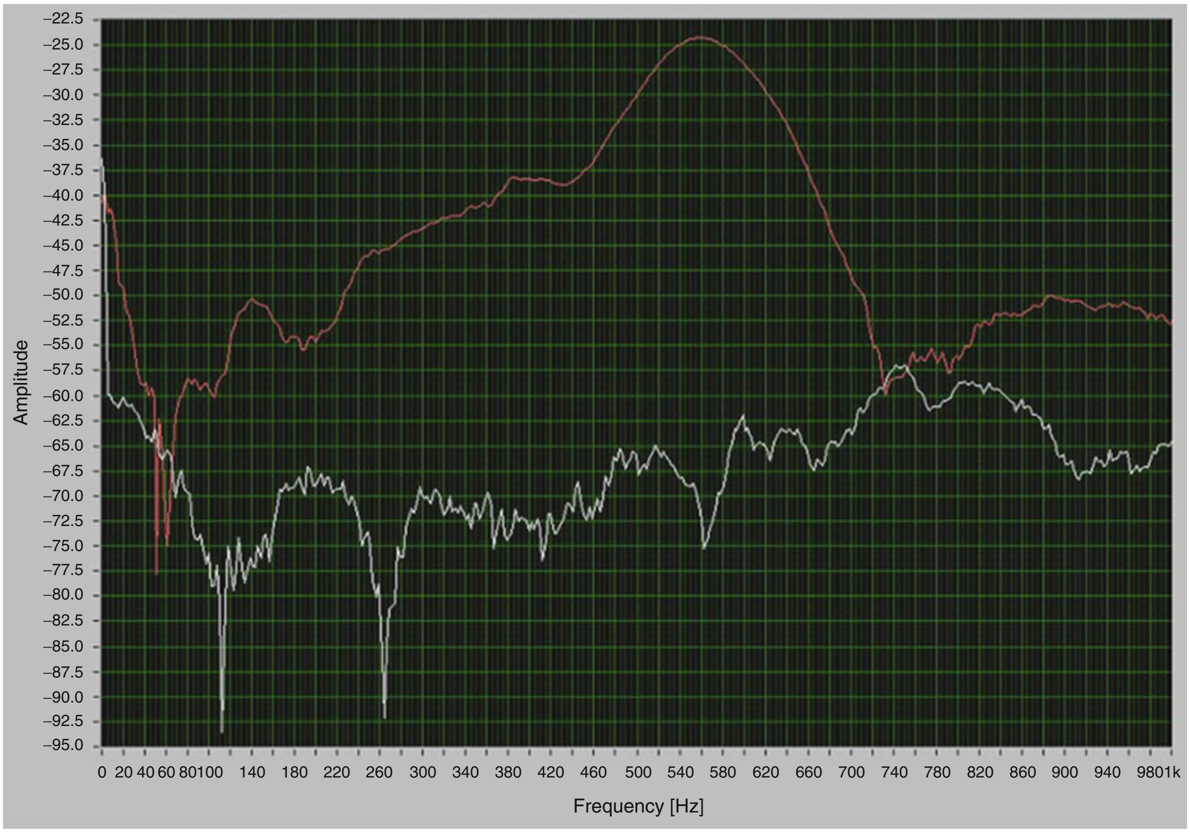


Fig. 15.6 MFS support structure hammer test, 0–1,000 Hz

15.3 Fault Localisation

Using the set of validated finite element models, the next stage of simulation was to study mode shapes and modal frequencies for potential use in localising unbalance faults. In this regard, simulation presented several advantages over an experimental MFS. From an experimental perspective, the MFS contains two rotors – opening the possibility to localise a fault to one (or both) of these rotors. Using simulation, it is possible to extend this to a situation where 12 rotors are fitted. This is of interest as many rotating machines, gas turbines providing one example, contain many rotors in several ‘stages’. The ability to localise a fault to one section of a machine could produce several enhancements with maintenance and safety – including the ability to speed up fault finding and part replacement.

In order to study this aspect, a model was created which maintained the shaft and bearing dimensions of the original model (and thus actual MFS), however applied 12 equally-spaced rotors along the length of the shaft. A simple 13.3 g static unbalance was applied to each rotor in turn, monitoring the shifts in modal frequency depending on the position of the unbalance within the machine. The 12 rotor model can be seen displayed in Fig. 15.7, with the results displayed in Fig. 15.8.

In Fig. 15.8, rotor 1 refers to the rotor positioned closest to the motor and rotor 12 refers to the rotor positioned farthest from the motor. A detailed, finely-meshed, simulation was used in order to reduce model error – however from repeat experiments it is estimated that a model error of ±0.1 Hz is incorporated in these results, and therefore only shifts in frequency greater than 0.1 Hz can be considered valid. It is also worth noting that, although the rotors are equally spaced, due to the bearing points and motor position the model is not symmetrical. The results produced in Fig. 16.10 display clear defining features when the unbalance is applied to certain rotors, with less well defined features on others. It would not be possible to accurately localise an unbalance based solely on these results, however in combination with further studies these results could potentially prove to be very useful.

Unbalance faults can occur on one rotor in a variety of forms; these include static, dynamic and couple unbalance [6]. With this in mind, another form of fault localisation could be considered to be localising an unbalance within one rotor, or,

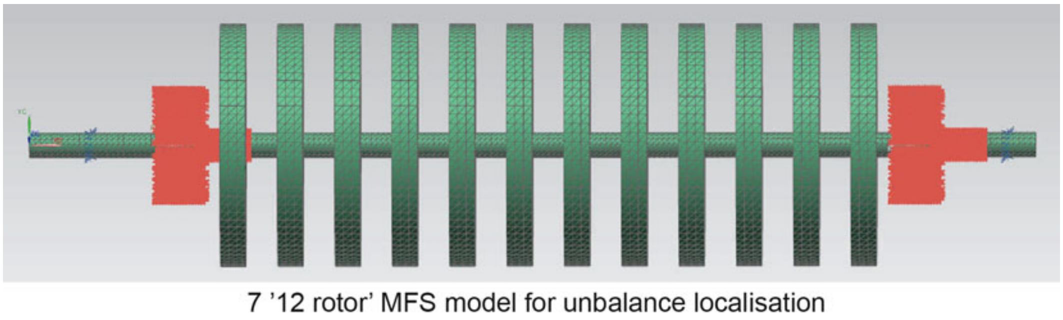


Fig. 15.7 '12 rotor' MFS model for unbalance localisation

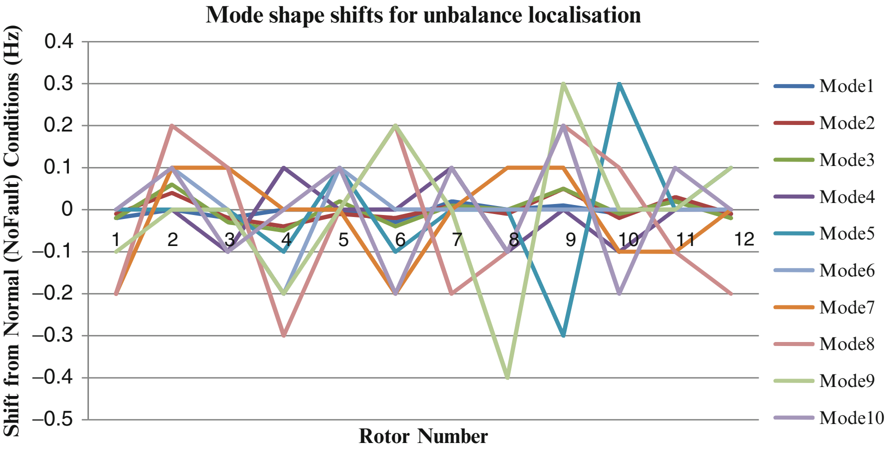


Fig. 15.8 Modal frequency shifts for '12 rotor' model. Rotor '1' closest to motor end, rotor '2' farthest from motor end

more specifically, identifying the type of unbalance which has occurred. In order to study this further simulation studies were performed. For this case, the original finite element model of the MFS was used, with a variety of unbalance faults applied to the motor-side rotor. Starting at 'top dead centre' a static (single) unbalance was applied. This unbalance was then incrementally separated in 20° segments until the unbalance was separated by 180° (coupled unbalance). The resulting model displayed the characteristics shown in Fig. 15.9. Repeat studies estimated that this model contained a variation (error) of approximately ±0.2 Hz, and therefore only results outside of 0.2 Hz variation can be considered a good indication of variance from the static unbalance case. It is also of note that due to a shaft-rotor coupling, each rotor, whilst balanced, is not symmetrical. Once again the results do not provide a conclusive method of localising unbalance; however in combination with additional studies this can potentially be achieved.

15.4 Other Unbalance Studies and Observations

In addition to the results detailed above, some additional studies performed provide information valuable to future works. Unbalance as a fault can occur as an individual item or as a function of another underlying problem – such as a shaft misalignment. It is therefore pertinent to investigate if unbalance faults can be identified as an individual fault or a function of another malfunction. Table 15.3 displays the differences from normal operating conditions between mode shifts of the first ten modes for unbalance and unbalance + misalignment cases. The unbalance applied was 5.8 g applied to both rotors; the misalignment was a 1 mm misalignment at the rotor-end bearing point. The results are based on the same conditions as detailed for Table 15.2.

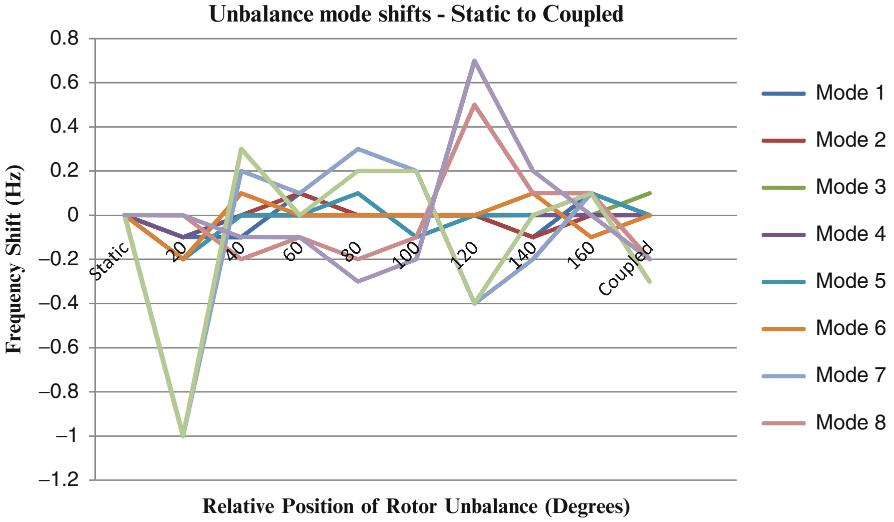


Fig. 15.9 Modal frequency shifts from static unbalance to couple unbalance in 20° increments

Table 15.3 Modal frequency shifts from normal operating conditions for 5.8 g unbalance and 5.8 g unbalance plus 1 mm bearing point misalignment cases

Mode	1	2	3	4	5	6	7	8	9	10
Unbalance	1.16	1.50	1.30	2.7	8.9	4.8	2.5	1.4	4.8	4.0
Unbalance + misalignment	−0.5	1.9	1.5	3.0	5.6	3.4	4.4	9.0	3.9	1.6

The results in Table 15.3 indicate a relatively large difference between the plain unbalance and the misalignment cases. However, as only one simple combination of misalignment and unbalance has been studied, this alone is not enough to state that unbalance can be identified as an individual fault or as part of a misalignment. Further investigation would be required in this area to correctly understand fault differentiation.

Also of note is the effect of the size of the unbalance on detection ability. All unbalances simulated for this paper are based upon those available for application on a MFS. It should be noted that in rotating machinery terms these are relatively large unbalances. In gas turbines, a much smaller unbalance than has been studied here can be deemed unacceptable. As size of the unbalance (as a function of rotor mass) decreases the variation in modal frequencies also decreases, thus complicating the ability to localise such faults.

15.5 Discussion

The experimental and simulation work outlined in this paper indicates the first steps towards the production of a strategy for fault localisation. The work concerns mostly model validation and lower-order modal investigations. Accurate, high fidelity models have been used as a basis for validation of simulation for the purposes of future study. A level of validation has been achieved which has been deemed acceptable in these circumstances, following general and popular model validation techniques (namely hammer testing). The next stage in the investigation of fault localisation involved experimental and simulation studies to determine the effect of a variety of unbalance faults on modal frequencies, with the intention to use this information for fault localisation. The studies concentrated on lower-order modes, in the case of the MFS the first ten, or roughly 0–1,000 Hz.

In this frequency range, it is possible to relate experimental results to mode shapes and defined modal frequencies. The results obtained from simulation indicate that the possibility exists in order to develop such localisation methods (or strategies) using techniques such as these. However, due to a number of factors including the relatively small frequency shifts involved the results to not represent a complete method for localising unbalance.

In order for this aim to be achieved, it is necessary to look at a broader frequency spectrum and alternate methods of assessing frequency-spectrum changes. By incorporating medium and high-frequency information with that which has been obtained during this study, the possibility to localise unbalance faults is greatly increased. In addition to this, some other studies performed in this paper indicate at directions for future research. In order to provide suitable unbalance localisation techniques which are valid in industrial situations a number of factors need to be considered. These include considering a wide range of possible faults so that overlap does not occur, with the potential for false localisation readings to be produced.

Other considerations for future studies include the applicability of FEA models to determine the information necessary for fault localisation. High-fidelity models generate large amounts of data and take a long time to solve. Repeating the procedures here, along with any additional required studies in order to develop localisation techniques for large structures such as a gas turbine could quickly get unwieldy. With this in mind, techniques such as model order reduction could prove valuable for effective simulation and data gathering of larger structures. Another important factor discussed was the ability to differentiate unbalance from other common faults such as misalignment; as such cases complicate fault localisation and render it less effective.

Finally, for future unbalance fault localisation techniques to prove usable, data acquisition methods and sensor positioning must be considered. In a simulation/lab based environment this does not prove as large an issue as with large, complex systems. Therefore any resulting localisation techniques need to be compatible in systems where sensors cannot necessarily sit on or close to the desired section of machine, thus resulting in a key consideration for future.

15.6 Conclusion

This study in the simulation of unbalance in rotating machinery for fault localisation has produced a number of results relevant for the development of future IVHM systems. High-fidelity model validation provided the basis for accurate simulation results, with simulations and experimental work centred on a state of the art machine fault simulator. Lower order modes and modal frequencies were used as the basis for unbalance localisation studies. Summarising the results in can be said that with further research into medium and high-order frequency spectra, combined with such considerations as sensor placement could result in a comprehensive strategy for fault localisation for future IVHM systems.

References

1. Domes B (2008) Vibration phenomena in aero-engines. In: 9th international conference on vibrations in rotating machinery, Exeter, UK. IMechE 1, 15–32
2. Ganeriwala SN, Schwarz B, Richardson MH (2009) Operating deflection shapes detect unbalance in rotating equipment. J Sound Vib 43:11
3. Yang T, Hsu HW (2009) An efficient diagnosis technique for variations of shaft-bow and unbalance. In: Computers and information in engineering conference, San Diego, ASME, 1, 57–66
4. Walker RB, Perinpanayagam S, Jennions I (2011) Physics-based simulation for health management of rotating machinery. In: Proceedings of the 8th international conference on condition monitoring and machinery failure prevention technologies – CM/MFPT 2011, vol 1. BINDT, Cardiff, 20–22 Jun 2011
5. Machorro-López JM, Adams DE et al (2009) Identification of damaged shafts using active sensing-simulation and experimentation. J Sound Vib 327:368–390
6. Bently DE (2002) Fundamentals of rotating machinery diagnostics, 1st edn. ASME, Canada

Chapter 16
Inverse Eigensensitivity Approach in Model Updating of Avionic Components

Elvio Bonisoli, Carlo Rosso, Cristiana Delprete, and Fabio Stratta

Abstract This paper presents an application of a well-known model updating approach based on the inverse eigensensitivity method.

The methodology is applied to an avionic equipment that supports different electronic instruments. In order to predict the operative dynamic behavior, a modal analysis is necessary, but high modal density and effects of nonlinearities of links, joints or rivets in the structure cannot allow an adequate correlation from test and numerical analysis. The structure is composed of two plates of aluminum connected by frame links with different sections and thickness to improve the stiffness on the borders and in the centre where the instruments are positioned. The different parts are connected with rivets. The target is the gap minimization between measured and computed modal parameters through the adjustment of a small number of physical parameters in the FE model.

The original contribution of the paper consists of substructuring and numerical strategies of coupling and select eigenfrequencies and eigenvectors in the iterative parameter updating. The approach demonstrates that few iterations are required to get a good correlation between EMA and FEA.

Keywords Avionic equipment • Inverse eigensensitivity method • Model updating

16.1 Introduction

The inverse eigensensitivity method has become relatively more popular than other methods and it has been successfully applied to improve analytical model, according to experimental data. These methods have been developed to reduce the inaccuracies of mathematical models. The sources of errors are different and appear most often combined one with another, principally when there are complex structures where joints, boundary conditions and damping have to be modeled.

Finite element model updating has emerged in the 1990s as shown in the surveys by Imregun and Visser [1]. The problem of updating a numerical model by using data acquired from a physical vibration test is richly handled by Mottershead and Friswell [2]. They showed how many issues are to be addressed to produce the desired improvement. The totally of method can be classified especially into direct methods and iterative methods. The latter are based on minimizing an objective function that is generally a non-linear function. Often eigenvalues and eigenvectors are used to construct this function. The use of eigendata sensitivity for analytical model updating in an iterative framework was first proposed by Collins et al. [3], instead Chen and Garba [4] used matrix perturbation technique for recomputation of eigensolution and evaluation of eigendata sensitivities. The effect of the improvement including second order sensitivities was studied by Kim et al. [5]. Another type of method was proposed by Lin et al. [6] to employ both the analytical and the experimental modal data for evaluating sensitivity coefficients with the objective of improving convergence to cases where there is a higher error magnitude, thing that happens in complex analytical model. It is also possible to directly use the measured frequency response function (FRF) data for identifying the system matrices as done in Fritzen [7], or still using this experimental

---

E. Bonisoli (✉)
Department of Production Systems, Politecnico di Torino, Torino, Italy
e-mail: elvio.bonisoli@polito.it

C. Rosso • C. Delprete • F. Stratta
Department of Mechanics, Politecnico di Torino, Torino, Italy

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_16, © The Society for Experimental Mechanics, Inc. 2012

information to apply a method named response function method (RFM) that was developed by Lin and Ewins [8]. Imregun et al. [9–11] conducted several studies using simulated and experimental data to evaluate the effectiveness of this approach as a matter of common.

Although several different approaches have been proposed and successfully applied on different structures, the authors of this paper would propose their experience on an avionic equipment. The aim is to define an appropriate and updated FE model for modal analysis (FEA) and to match the numerical results with an experimental modal analysis (EMA) campaign. The inverse eigensensitivity approach [2, 12] is proposed as an iterative model updating technique with respect to an experimental modal test campaign. Its validity and quick convergence has been demonstrated in Literature, but modal truncation and experimental inaccuracy effects may represent an interesting task for model updating of this kind of structures. The principal task of this article is the comparison of two different versions of the method used to reach a good matching between the modal characteristics of an avionic structure modeled with a linear FE approach and substructured in components. The mathematical bases are the same for both methods but the difference is defined in the objective functions to minimize, and this difference goes to affect the dimensions of the sensitivity matrix that are smaller for the new method than for the previous one. The classical objective function is the combination of errors between experimental and numerical eigenvalues and eigenvectors as is customary [12, 13]:

$$\Delta \omega_r^2 = \frac{\omega_{\exp,r}^2 - \omega_r^2}{\omega_r^2}, \quad \left\{\Delta \Phi^{(r)}\right\} = \left\{\Phi_{\exp}^{(r)}\right\} - \left\{\Phi^{(r)}\right\}. \tag{16.1}$$

After this the updating of the parameters $\{\Delta Par\} = \left\{\Delta a_1; \ldots; \Delta a_{pm}; \Delta b_1; \ldots; \Delta b_{pk}\right\}^T$ of the analytical model are done by solving a linear system where the sensitivity $[S]$ is calculated analytically from the FE model:

$$\{\Delta Eig\} = \left\{ \begin{array}{c} \dfrac{\omega_{\exp,r}^2 - \omega_r^2}{\omega_r^2} \\ \left\{\Phi_{\exp}^{(r)}\right\} - \left\{\Phi^{(r)}\right\} \end{array} \right\} = [S]\ \{\Delta Par\}. \tag{16.2}$$

The idea of trying a method modification arises because, neglecting the meaning of each row of (16.2), errors of eigenvector degree of freedom can affect the parameter corrections in the same way of an eigenvalue error. The fact of modifying (16.2) is to weight the eigenvector error with a scalar estimate and minimize the sensitivity matrix $[S]$. In particular, in the identified mode-shapes is common having some qualitative misalignments between experimental nodes and, although these small discrepancies do not affect the goodness of experimental modes, they can produce undesired effects on the updating parameters computations. The proposed modification tries to minimize this problem without producing worse updating results.

To evaluate the efficiency of both methods, before applying them to a real avionic structure, some numerical tests have been performed, taking into account that the equivalent experimental data were actually a reduced part of analytical results obtained from some modifications of substructures parameters in the FE model.

16.2 Modified Inverse Eigensensitivity Model Updating

In order to obtain an updated FE model, the inverse eigensensitivity methodology using arbitrarily chosen macro-elements is implemented. It relies on the minimization of the discrepancies between measured and computed modal parameters through the adjustment of a number of uncertain parameters in the finite element model. In particular, the inverse eigensensitivity methodology takes into account modal parameters obtained from a modal test, which are, in most cases, not compatible with those from the analytical model, because of the limited number of experimental modes and measured coordinates. Therefore, a strong modal truncation is applied on the error estimation of eigenvectors.

The success of the iterative procedure based on the inverse eigensensitivity methodology depends heavily on the choice of the submatrices, i.e. macro-elements, which the structure is divided in. The fastness of convergence and the stability of the iterative procedure depend on this choice. It was observed that global parameters are highly more efficient, instead of local parameters, to get a better updated model in few iterations and a following refinement may be applied, by using more autonomous substructures [13].

The mass and stiffness matrices of the FE model can be expressed through the sum of macro-substructures, expanded to global dimensions and weighted by independent parameters a_i, b_i

$$[M] = \sum_{i=1}^{pm} a_i [M_i], \quad [K] = \sum_{i=1}^{pk} b_i [K_i] \tag{16.3}$$

where an overall of *pm* inertial plus *pk* elastic updating parameters for the FE model are considered.

Regarding (16.2), both non-dimensional eigenvalues and eigenvectors errors $\{\Delta\, EigenV\}$ between experimental and analytical model are considered, in order to update the FE model respect to the vector of correction coefficients $\{\Delta\, Par\} = \left\{\Delta\, a_1; ...; \Delta\, a_{pm}; \Delta\, b_1; ...; \Delta\, b_{pk}\right\}^T$ in the neighborhood of the unitary vector, corresponding to the analytical configuration. The proposed modification of the approach acts on the vector $\{\Delta\, Eig\}$. It reduces the number of independent error estimations related to the sensitivity matrix. Instead of having one independent row for each point of measure, the error on each involved eigenvector is computed as a scalar value, pre-multiplying corresponding part of the error vector $\{\Delta\, Eig\}$ and the sensitivity matrix $[S]$ by the experimental-numerical eigenvector error $\left\{\Delta\, \Phi^{(r)}\right\}^T$. Although this modification of the original approach combines experimental and numerical estimations in the sensitivity matrix $[S]$, a numerical improvement is expected in numerical stability and insensitive properties on erroneous identification of some components of experimental eigenvectors.

The *r*th eigenvalue and eigenvector of the FE model can be expanded in a Taylor series; by neglecting the second and higher-order terms, both eigenvalues and eigenvectors residuals can be approximated as

$$\begin{Bmatrix} \frac{\omega^2_{\exp,1} - \omega^2_1}{\omega^2_1} \\ \frac{\left\{\Delta\Phi^{(1)}\right\}^T \left\{\Delta\Phi^{(1)}\right\}}{\left\|\Phi^{(1)}\right\|^2} \\ \vdots \\ \frac{\omega^2_{\exp,r} - \omega^2_r}{\omega^2_r} \\ \frac{\left\{\Delta\Phi^{(r)}\right\}^T \left\{\Delta\Phi^{(r)}\right\}}{\left\|\Phi^{(r)}\right\|^2} \end{Bmatrix} = \begin{bmatrix} \frac{\partial\, \omega^2_1}{\partial\, a_1} \Big/ \omega^2_1 & \cdots & \frac{\partial\, \omega^2_1}{\partial\, a_{pm}} \Big/ \omega^2_1 & \frac{\partial\, \omega^2_1}{\partial\, b_1} \Big/ \omega^2_1 & \cdots & \frac{\partial\, \omega^2_1}{\partial\, b_{pk}} \Big/ \omega^2_1 \\ \frac{\left\{\Delta\Phi^{(1)}\right\}^T}{\left\|\Phi^{(1)}\right\|^2} \frac{\partial \left\{\Phi^{(1)}\right\}}{\partial\, a_1} & \cdots & \frac{\left\{\Delta\Phi^{(1)}\right\}^T}{\left\|\Phi^{(1)}\right\|^2} \frac{\partial \left\{\Phi^{(1)}\right\}}{\partial\, a_{pm}} & \frac{\left\{\Delta\Phi^{(1)}\right\}^T}{\left\|\Phi^{(1)}\right\|^2} \frac{\partial \left\{\Phi^{(1)}\right\}}{\partial\, b_1} & \cdots & \frac{\left\{\Delta\Phi^{(1)}\right\}^T}{\left\|\Phi^{(1)}\right\|^2} \frac{\partial \left\{\Phi^{(1)}\right\}}{\partial\, b_{pk}} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial\, \omega^2_r}{\partial\, a_1} \Big/ \omega^2_r & \cdots & \frac{\partial\, \omega^2_r}{\partial\, a_{pm}} \Big/ \omega^2_r & \frac{\partial\, \omega^2_r}{\partial\, b_1} \Big/ \omega^2_r & \cdots & \frac{\partial\, \omega^2_r}{\partial\, b_{pk}} \Big/ \omega^2_r \\ \frac{\left\{\Delta\Phi^{(r)}\right\}^T}{\left\|\Phi^{(r)}\right\|^2} \frac{\partial \left\{\Phi^{(r)}\right\}}{\partial\, a_1} & \cdots & \frac{\left\{\Delta\Phi^{(r)}\right\}^T}{\left\|\Phi^{(r)}\right\|^2} \frac{\partial \left\{\Phi^{(r)}\right\}}{\partial\, a_{pm}} & \frac{\left\{\Delta\Phi^{(r)}\right\}^T}{\left\|\Phi^{(r)}\right\|^2} \frac{\partial \left\{\Phi^{(r)}\right\}}{\partial\, b_1} & \cdots & \frac{\left\{\Delta\Phi^{(r)}\right\}^T}{\left\|\Phi^{(r)}\right\|^2} \frac{\partial \left\{\Phi^{(r)}\right\}}{\partial\, b_{pk}} \end{bmatrix} \begin{Bmatrix} \Delta\, a_1 \\ \vdots \\ \Delta\, a_{pm} \\ \Delta\, b_1 \\ \vdots \\ \Delta\, b_{pk} \end{Bmatrix} \tag{16.4}$$

where $\left\|\Phi^{(r)}\right\|^2 = \left\{\Phi^{(r)}\right\}^T \left\{\Phi^{(r)}\right\}$ and the matrix is the balanced sensitivity matrix $[S]$. In order to estimate the coefficients of the sensitivity matrix $[S]$, by taking derivatives of the *M* and *K*-orthogonality properties of eigenvectors, it results [14, 15]:

$$\frac{\partial\ \omega^2_r}{\partial\ a_i} = -\omega^2_r \left\{\Phi^{(r)}\right\}^T [M_i] \left\{\Phi^{(r)}\right\}, \quad \frac{\partial\ \omega^2_r}{\partial\ b_i} = \left\{\Phi^{(r)}\right\}^T [K_i] \left\{\Phi^{(r)}\right\}, \tag{16.5}$$

$$\frac{\partial\ \left\{\Phi^{(r)}\right\}}{\partial\ a_i} = \sum_{s=1}^{N} \alpha_i^{r,s} \left\{\Phi^{(s)}\right\}, \quad \frac{\partial\ \left\{\Phi^{(r)}\right\}}{\partial\ b_i} = \sum_{s=1}^{N} \beta_i^{r,s} \left\{\Phi^{(s)}\right\}, \tag{16.6}$$

where

$$\begin{cases} \alpha_i^{r,s} = \dfrac{-\omega^2_r\ \left\{\Phi^{(s)}\right\}^T [M_i]\ \left\{\Phi^{(r)}\right\}}{\omega^2_r - \omega^2_s}, & r \neq s \\ \alpha_i^{r,s} = -\dfrac{1}{2}\ \left\{\Phi^{(s)}\right\}^T [M_i]\ \left\{\Phi^{(r)}\right\}, & r = s \end{cases}, \quad \begin{cases} \beta_i^{r,s} = \dfrac{\left\{\Phi^{(s)}\right\}^T [K_i]\ \left\{\Phi^{(r)}\right\}}{\omega^2_r - \omega^2_s}, & r \neq s \\ \beta_i^{r,s} = 0, & r = s \end{cases}, \tag{16.7}$$

hence the eigenvectors derivative are expressed by using the expansion theorem and considering eigenvectors as a base to describe those terms.

The updating procedure has been achieved by solving the following linear system:

$$\{\Delta\, Eig\} = [S]\ \{\Delta\, Par\} = [U]\ [\Sigma]\ [V]^T\ \{\Delta\, Par\}, \tag{16.8}$$

where the number of updating parameters are usually much smaller than the number of modal residual equations in the original inverse eigensensitivity method, while $[S]$ is more similar to a square matrix in the proposed modification.

Therefore the updating parameters are obtained in a least square sense, by a Singular Value Decomposition technique

$$\{\Delta Par\} = [V]\ [\Sigma]^{+}[U]^{T}\ \{\Delta EigenV\}, \tag{16.9}$$

where $[\Sigma]^{+}$ is the Moore-Penrose generalized inverse matrix of $[\Sigma]$.

The mass and stiffness matrices of the updated FE model result thus:

$$[M] = \sum_{i=1}^{pm} (a_i + \Delta a_i)\ [M_i], \quad [K] = \sum_{i=1}^{pk} (b_i + \Delta b_i)\ [K_i]. \tag{16.10}$$

Finally, before or after the updating procedure, independently to the eigenvectors, it is possible to minimize the non-dimensional eigenvalues errors $\Delta\omega$ by using a global rescale of the FE model, in order also to control the mass of the system; in a least square sense, the inertial and elastic updating parameters may be evaluated through

$$\Delta a_i = 0, \quad \Delta b_i = b_i\,(\varepsilon^2 - 1), \tag{16.11}$$

where $\Delta\omega = \min_{\varepsilon} \left(\sum_{r=1}^{n} \left| \frac{\omega_{\exp,r} - \varepsilon\ \omega_r}{\varepsilon\ \omega_r} \right| \right)$ gives $\varepsilon = \sum_{r=1}^{n} \left[\frac{\omega_{\exp,r}}{\omega_r} \right]^2 \Big/ \sum_{r=1}^{n} \frac{\omega_{\exp,r}}{\omega_r}$.

The improvement of the inverse eigensensitivity model updating of (16.4) is mainly to reduce the size of the sensitivity matrix, so even the mathematical errors to solve the linear system in a mean least square sense are minor. But on the other hand experimental information are lost because there is only one value that groups the errors of each identified eigenvector. The difference from the two different methods can be quantified with controlled test that are developed in the rest of the paper.

In order to test effectively the inverse eigensensitivity approach, the FE model must be prepared and be modified. In the next paragraph the avionic structures under test is presented.

16.3 Avionic Structure Description

The object under study is an avionic equipment inside aircraft, whose rule is to operate as support for electronic devices. The structure is a typical rig composed of aluminum plates with several stiffening components. Numerous rivets are used to connect the components. Numerous and close modes of the plates are expected, depending heavily on connections properties. The 3D CAD model is shown in Fig. 16.1.

The prototype of the object is simplified as shown in Fig. 16.1, trying to keep intact as much as possible its inertia and stiffness properties, and preserving geometrical dimensions. In Fig. 16.2 the corresponding model of the realized prototype is shown. Specifically, the assembly is therefore consisting of the following elements:

- Two frame links with Z section,
- Two frame links with T section,
- Two frame links with C section,
- One upper plate,
- One bottom plate.

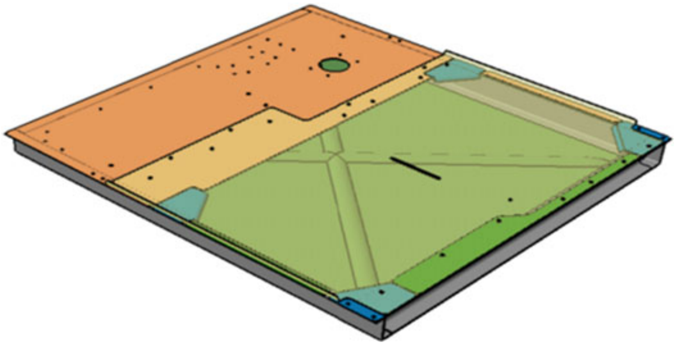


Fig. 16.1 3D drawing of avionic equipment

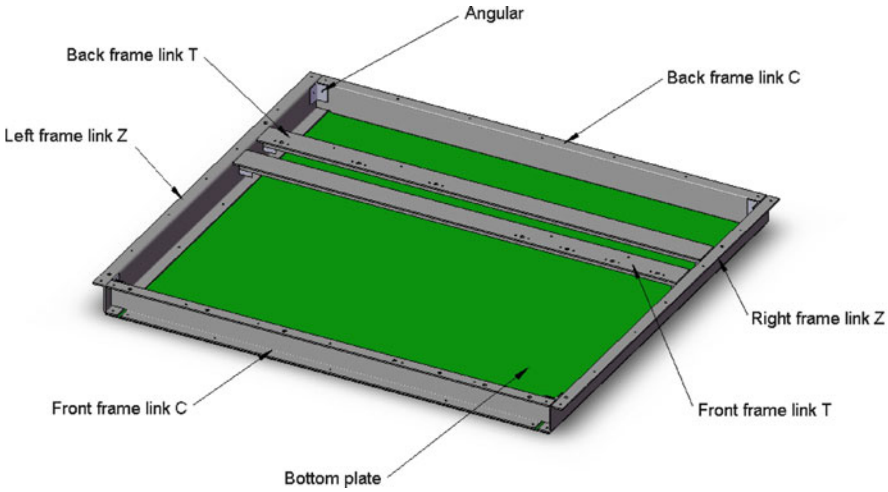


Fig. 16.2 Isometric view with components name of the avionic equipment prototype

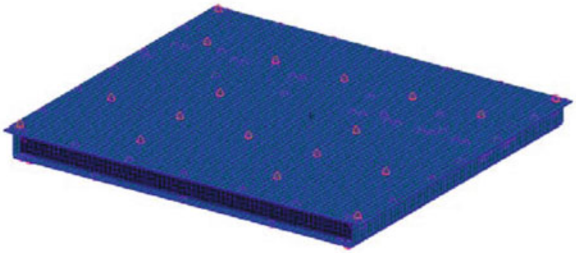


Fig. 16.3 Corresponding FE model of the prototype

Table 16.1 FEM characteristics

Nodes	1,925
Elements	1,573
CONM2	31
RBE2	87
RBE3	31

The simplified geometry is imported, using the IGES interface, in the preprocessing for the FE analysis PATRAN 2005. This structure is discretized using two-dimensional shell elements quad4 and tria3 with equivalent thickness properties of the prototype. The material is a hard aluminum class T2024. The connections are composed of RBE2 elements and accelerometers are been taken into account with CONM2 elements for completeness. The FE model and its characteristics are reported in Fig. 16.3 and resumed in Table 16.1.

The experimental tests are defined by positioning 31 mono-axial accelerometers on both plates; the boundary conditions are assumed as free-free by means of a suspension structure of soft springs, as shown in Fig. 16.4. Typical hammer tests [16] are performed to identify modal data through SIMO (single input, multi output) frequency domain techniques implemented in LMS software.

Having the modal data from the experimental campaign and the predictions arising from the numerical model, it was possible to perform a numerical–experimental comparison. To evaluate the mode-shape correlation, the index MAC (Modal Assurance Criterion) is adopted [16]:

$$MAC(j,k) = \frac{\left|\left\{\Phi^{(j)}\right\}^T\left\{\Phi_{\exp}^{(k)}\right\}\right|^2}{\left(\left\{\Phi^{(j)}\right\}^T\left\{\Phi^{(j)}\right\}\right)\left(\left\{\Phi_{\exp}^{(k)}\right\}^T\left\{\Phi_{\exp}^{(k)}\right\}\right)} \tag{16.12}$$

Fig. 16.4 Experimental test rig with accelerometer setup on both plates

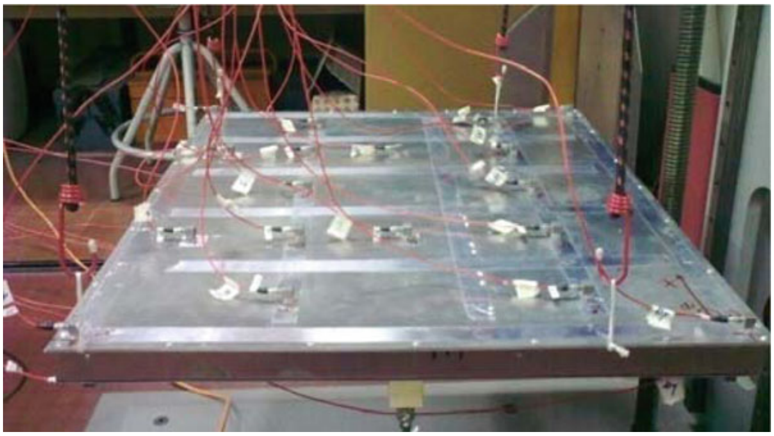
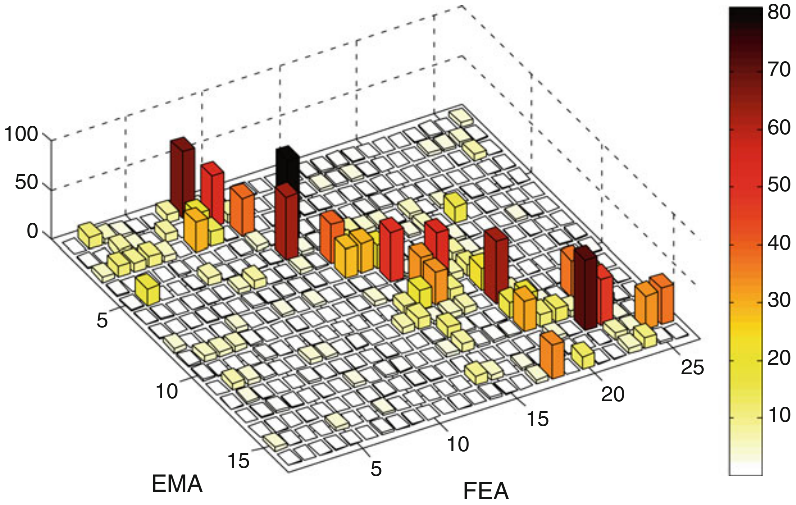


Fig. 16.5 EMA–FEA comparison: calculation of MAC index



From this first comparison, reported in Fig. 16.5, two main inconsistencies are noted:

- The absence of the first mode of bottom plate from experimental tests,
- A not-satisfactory comparison for the first six modes.

These two problems are discussed and tackled separately. The first problem can be solved adding rivet damping to the model. In particular, this mode is over-damped, due to the strong dissipative effects of the numerous rivets on the boundary of the structure. This problem is not handled in this paper, while for the first mode-shapes the model updating technique, based on the modification of the inverse eigensensitivity approach, is the aim.

16.4 Methodology Application

By applying the method to the structure, the avionic equipment is divided into five sub-structures (Fig. 16.6). Material Young's modulus and density for each component are the updating parameters. In this case it is appropriate to put the two plates separated into two independent subgroups (named part 3 for the top plate and part 4 for the bottom one). Since other parts of the model come in pairs (there are two frame links of section C, two Z and two T) three different subgroups could be defined; the better choice is to bring together the C and Z sections because they represent the structural board of the structure (named part 1) and, if some modifications are expected, these should affect the entire subgroup. Whereas the frame links T (named part 2) are independent because they deeply interact on the vibration modes of the upper plate. The steffening corners, fifth subgroup (named part 5), contain the eight angulars that form the connection between the various parts. It is defined for the structure completeness, but its effect could be omitted and its parameters could not be updated. Figure 16.6 shows the exploded view of the assembly to highlight the five subgroups.

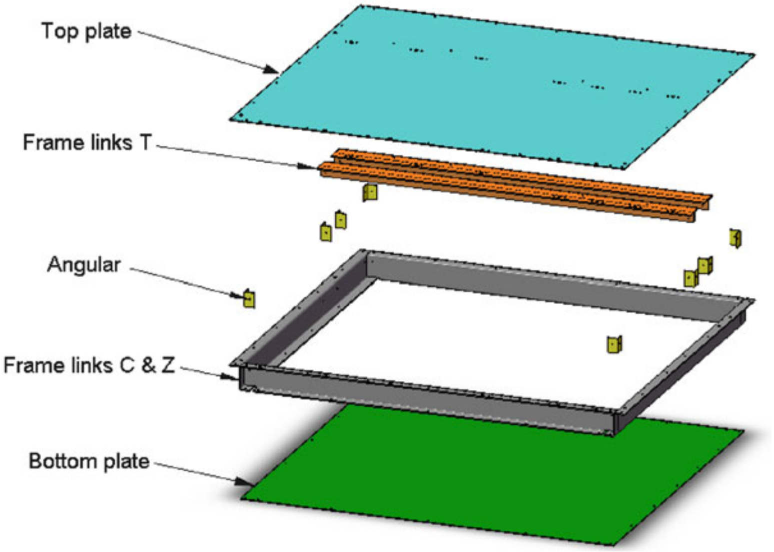


Fig. 16.6 Division of the model in five subgroups

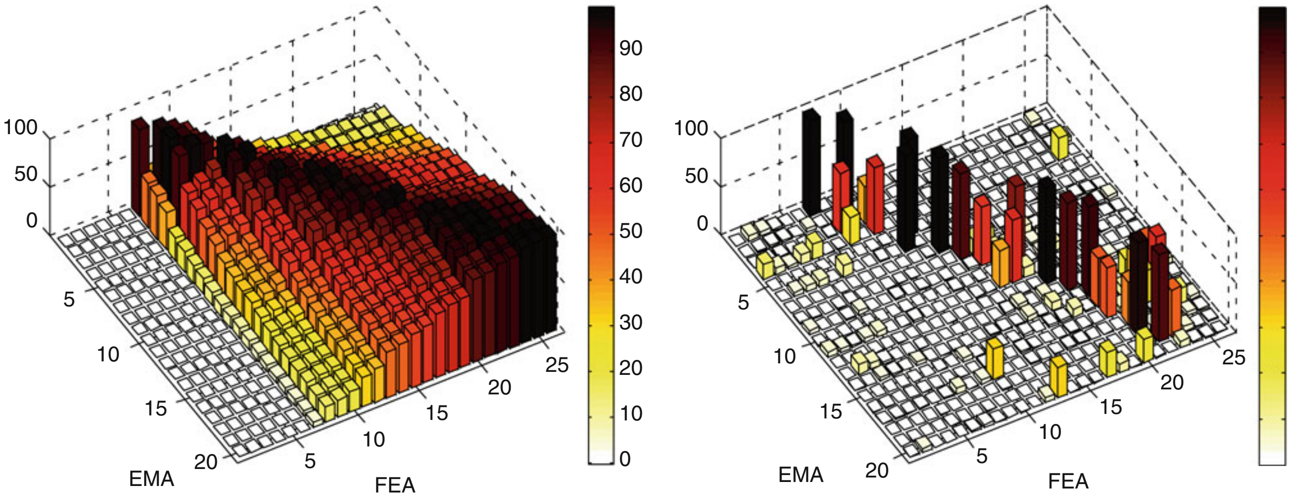


Fig. 16.7 Test 1: eigenvalues errors (*left*) and MAC matrix (*right*) between EMA and FEA

16.5 Numerical and Experimental Tests

16.5.1 Test 1: EMA Obtained from Numerical Modifications

To be confident of the convergence, physical validity and robustness of both methods, controlled tests are analyzed. Some parameters values are numerically modified in the FE model. This modified model is adopted, through modal truncation, as the equivalent experimental modal data to update as a reference instead of the actual data derived from experimental modal analysis. Therefore, a physical solution of the updating technique is definitely consistent.

The starting point is the errors on eigenvalues and the evaluation of the MAC index for eigenvectors discrepancies. Both are represented like matrices in Fig. 16.7.

The progressive iterations of density and Young's modulus for each part are shown in Figs. 16.8, 16.9, and 16.10, using the original version of the inverse eigensensitivity approach, while Figs. 16.11, 16.12, 16.13, and 16.14 show the results obtained through the modification of (16.4). The first six flexible modes are controlled at each iteration. The known existing updated model is represented as the target of this test.

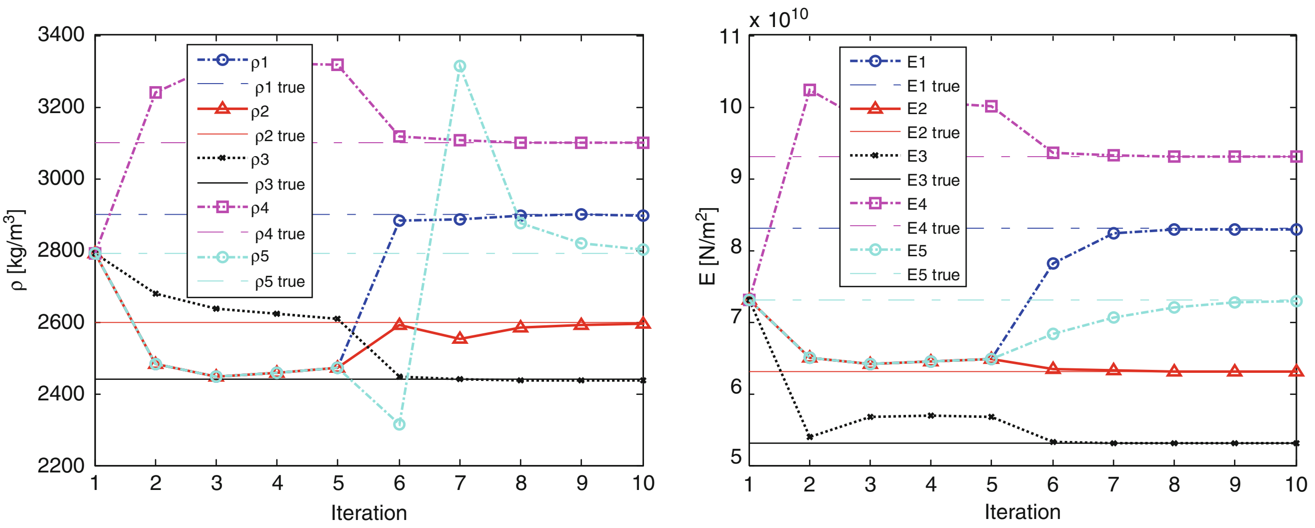


Fig. 16.8 Test 1: density (*left*) and Young's modulus (*right*) during the progressive iterations of the five substructures with the inverse eigensensitivity approach in the original version

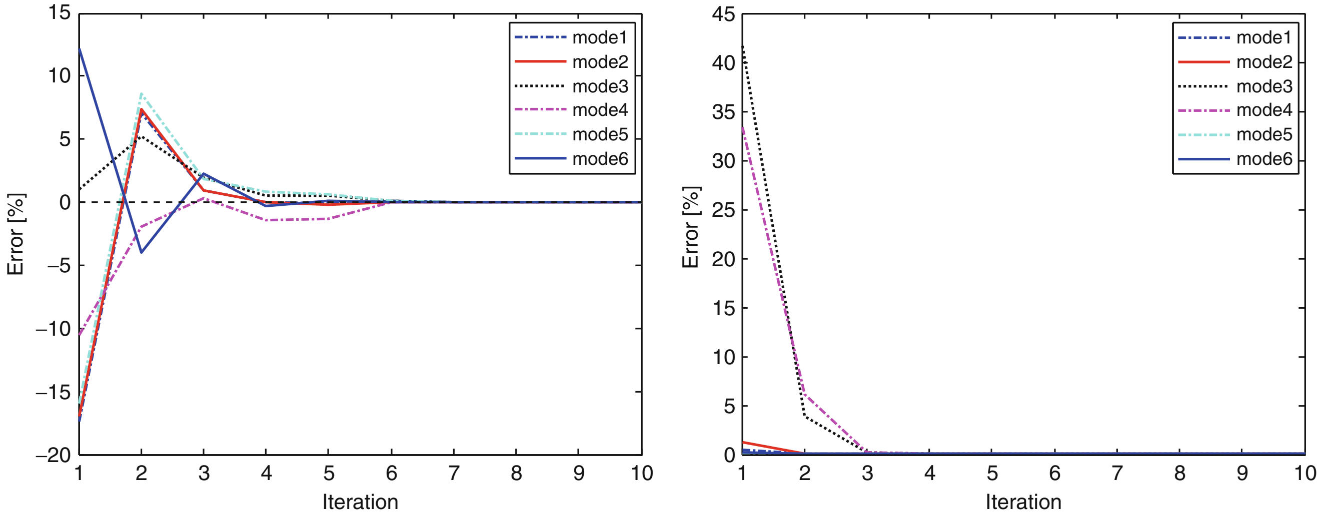


Fig. 16.9 Test 1: eigenvalues (*left*) and eigenvectors errors (*right*) of the first six modes during the progressive iterations with the inverse eigensensitivity approach in the original version

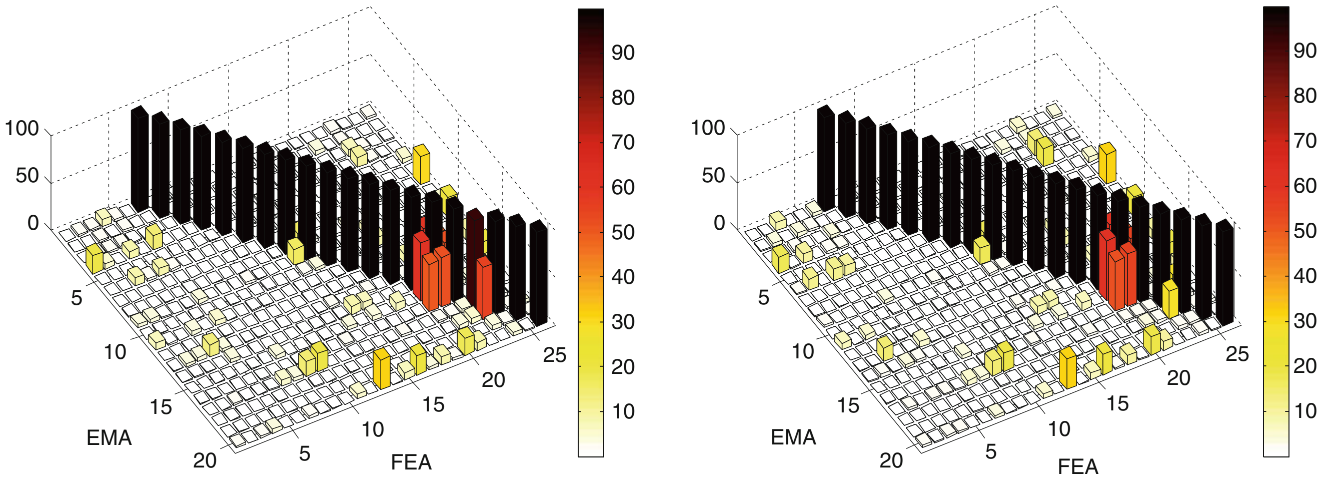


Fig. 16.10 Test 1: MAC matrices between EMA and FEA after 5 iterations (*left*) and 10 iterations (*right*) with the inverse eigensensitivity approach in the original version

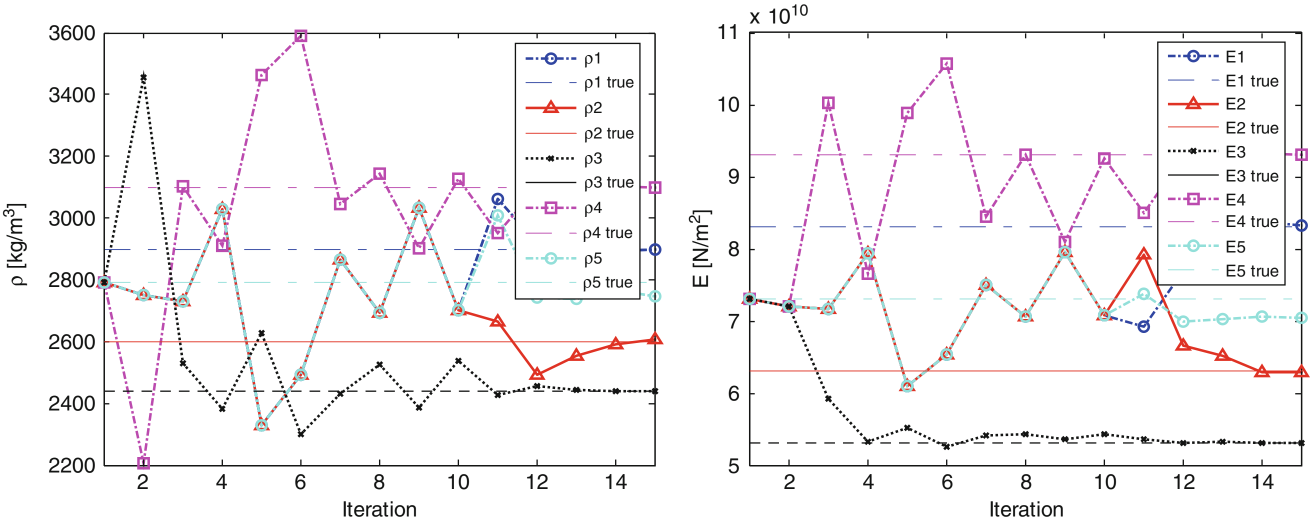


Fig. 16.11 Test 1: density (*left*) and Young's modulus (*right*) during the progressive iterations of the five substructures with the inverse eigensensitivity approach in the modified version

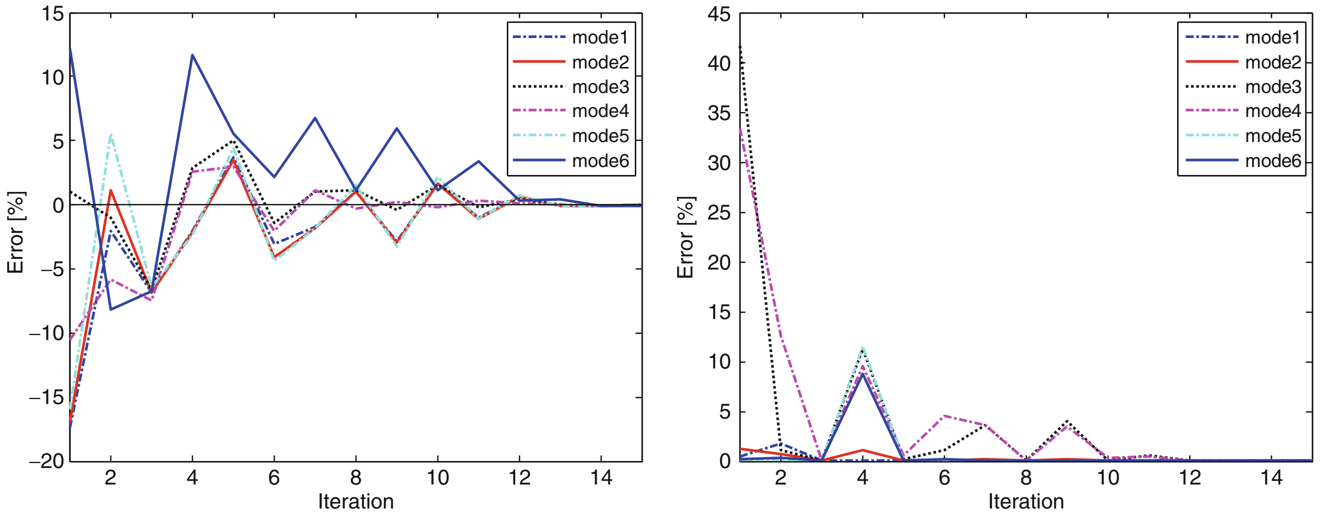


Fig. 16.12 Test 1: eigenvalues (*left*) and eigenvectors errors (*right*) of the first six modes during the progressive iterations with the inverse eigensensitivity approach in the modified version

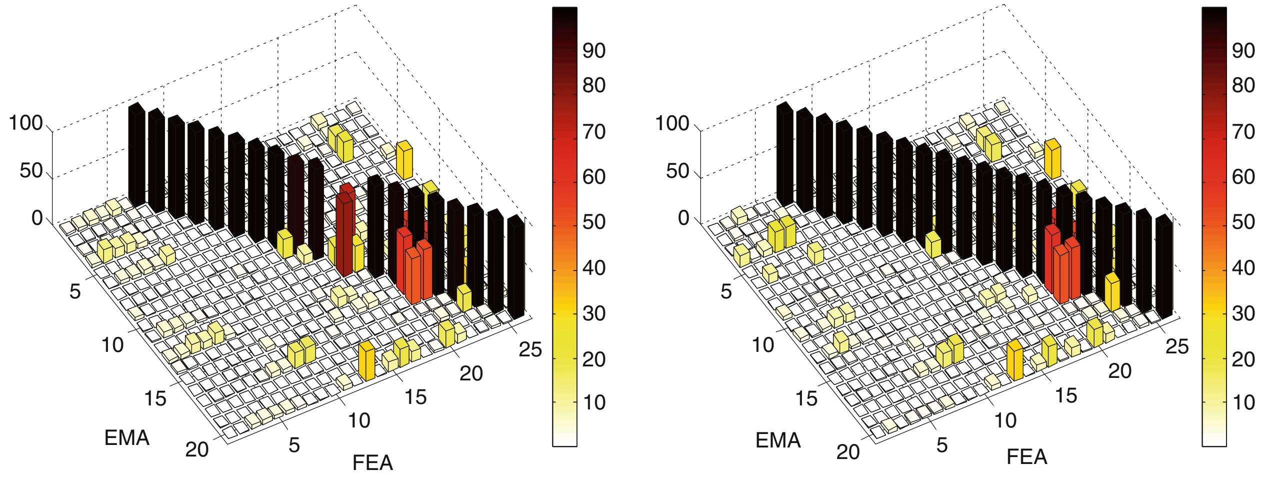


Fig. 16.13 Test 1: MAC matrices between EMA and FEA after 10 iterations (*left*) and 15 iterations (*right*) with the inverse eigensensitivity approach in the modified version

16.5.1.1 Updating with the Original Method

The procedure is necessarily iterative, but engineering experience is suitable for avoiding not physical solutions. Therefore the iterative algorithm is not automatic. Furthermore two degrees of freedom for piloting the iterations are available at each iteration: the number of modes to control and how many parameters can be accepted or neglected in the updating parameters. In fact in this first test the first six modes are always taken into account for all the iterations, but for the first five iterations only the more important four parameters are updated to tune the modal characteristics (density ρ and Young's modulus E of the upper and bottom plates). After the first five iterations, all the ten parameters (density ρ and Young's modulus E for each substructure) are then considered. It is worth noting that already at the fifth iteration the method leads to an equivalent system that has the same modal characteristics of the true model used to withdraw the reference data. So after this point the method allows to update all the ten parameters that in few steps reach the known values.

16.5.1.2 Updating with the Modified Method

The procedure and results are similar to the original method, but more iterations are needed. All the ten parameters are active in the updating technique after the tenth iteration. A negative note is that, although even this method converges at true values, it is more unstable and requires more attention to decide which parameters update at each iteration. With this first test the efficiency of both methods is demonstrated.

16.5.2 Test 2: EMA Obtained from Numerical Modifications with Noise

The next step is to introduce a Gaussian distribution of noise in the data taken as a reference. This test allows to have a measure of the robustness of the proposed modification.

The initial parameters are identical to the first test. Here, however, experimental values are taken as the eigenvalues and eigenvectors that come from this model and are slightly dirtied through the introduction of a percentage of noise: a random Gaussian noise of 10% with respect to the eigenvector RMS is assumed. The added noise must necessarily be small because otherwise both methods diverge; it means that the model cannot replicate the information that the modal values are taken as experimental.

The starting point is represented in the matrices shown in Fig. 16.14.

The effect of noise can be underlined with the use of an auto-MAC between eigenvectors before and after the introduction of noise. A high correspondence of the two sets of eigenvectors is anyway evinced; in fact, the index MAC has a mean value of about 98% for all modes on the diagonal and some correlations between similar mode-shapes are emphasized. Thus, mainly for the first flexible modes, random noise has a relative weight on eigenvector discrepancies. Higher negative effects can be produced on eigenvalues, where random noise could produce unphysical systems with respect to the updating parameters.

Also in this test, the first six flexible modes are controlled at each iteration.

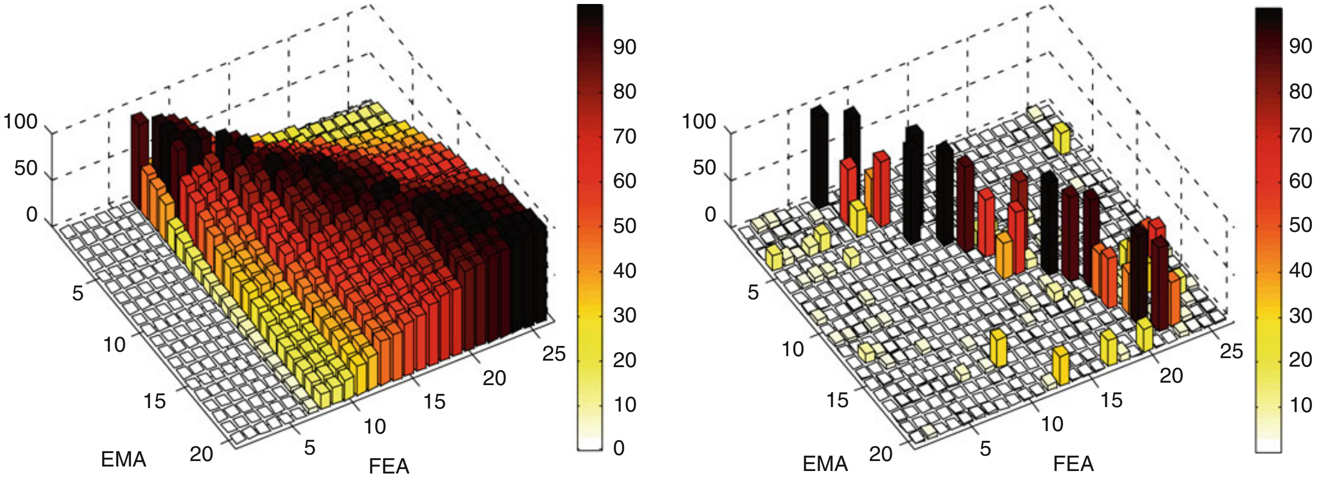


Fig. 16.14 Test 2: eigenvalues errors (*left*) and MAC matrix (*right*) between EMA and FEA

16.5.2.1 Updating with the Original Method

With the introduction of the random noise in the reference data, the method fails to converge to the true values of density and Young's modulus (Fig. 16.15), although both approaches still leads to a physical system with the same modal information except for modes 9 and 10 that are not well described (Figs. 16.16, 16.17).

It is worth noting that inertial and elastic properties of both plates are detected after six iterations, where all the ten updating parameters are active and independent.

16.5.2.2 Updating with the Modified Method

The modification is more unstable as discussed in the previous test (Figs. 16.18, 16.19). It is further from the true values even if it also leads to an equivalent dynamic system. The MAC index shown in Fig. 16.20 has equivalent correspondence to the original method results. For all the iterations the parts 1, 4 and 5 are not updated. Similar conclusions of the previous test can be resumed.

At the end of the considered iterations, both methods have eigenvalues and eigenvectors errors that are not negligible; the modified method demonstrated minor suitability, but its trend is similar to the original approach.

16.5.3 Test 3: EMA from Experimental Campaign

After checking the reliability of both methods, a real experimental case is chosen to update the FE model of the avionic equipment described previously.

The main characteristics of the test are:

- There is no guarantee that a physical system exists and possesses the identified experimental modal data;
- Moreover, if the boundary conditions between the substructures are not well modeled or the updating parameters are not suitable to describe the modal modification according to the modal errors, the iterative updating procedure could be not effective;
- Finally, due to the relative weight of updating an FE model for each iteration, the computing cost is quite relevant, the iterative procedure depends from the starting point and it is not possible to know if an optimal iterative solution is the best physical solution. In practice, the approach does not guarantee to compute a global optimization of the modal errors.

The starting point is represented in the matrix shown in Figs. 16.21.

16.5.3.1 Updating with the Original Method

In both cases only the density and Young's modulus of both plates are updated, because these four parameters affect more the modal characteristics of the structure. The low frequencies modes depends only on the first plate modes. In order to reduce the difference between EMA and FEA, after numerous attempts, it has been useless go to activate other parameters.

16.5.3.2 Updating with the Modified Method

Similar strategy is adopted for the modified method. The final MAC is practically identical for both methods. This confirms that the modified method comes to similar results of the original method. There are differences on the values of density and Young's modulus of the subgroups but, of course, it is not possible saying if one approach is better than the other one.

Analogously to the MAC comparison of Fig. 16.24 for the original method, Fig. 16.27 shows the EMA and FEA results for the modified one: the inversion of frequency order between the first two modes is removed, but the mean MAC value is similar to the previous results.

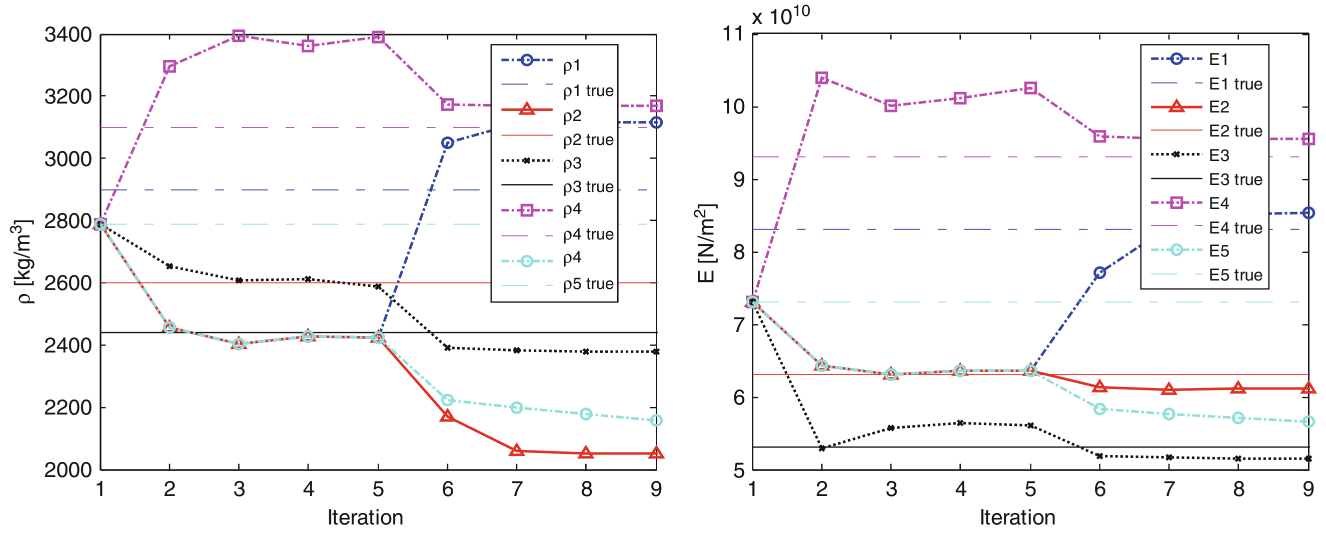


Fig. 16.15 Test 2: density (*left*) and Young's modulus (*right*) during the progressive iterations of the five substructures with the inverse eigensensitivity approach in the original version

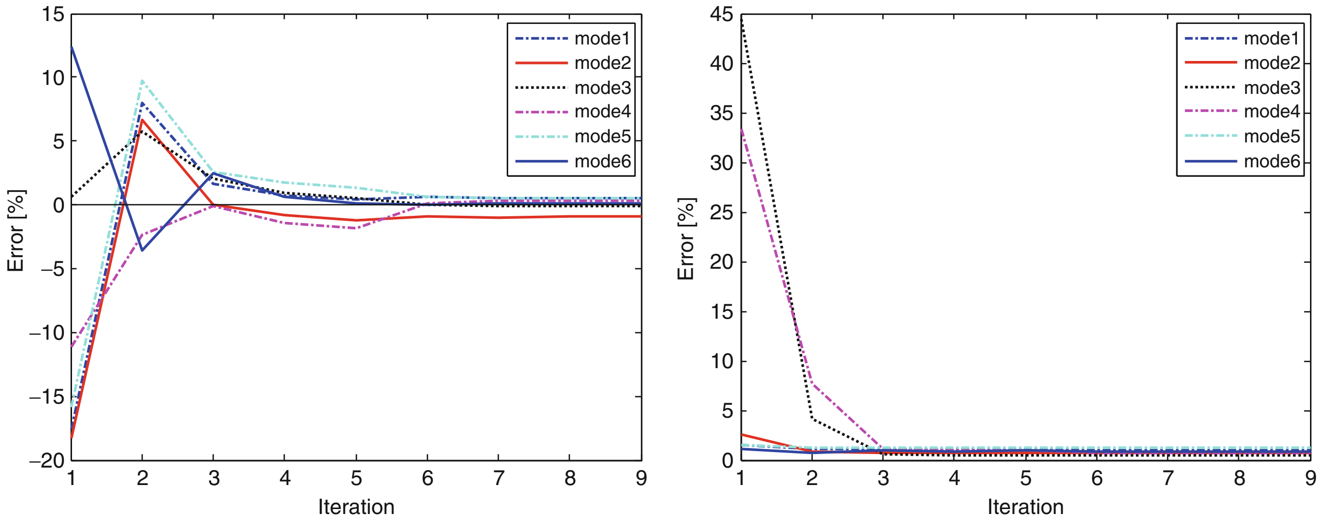


Fig. 16.16 Test 2: eigenvalues (*left*) and eigenvectors errors (*right*) of the first six modes during the progressive iterations with the inverse eigensensitivity approach in the original version

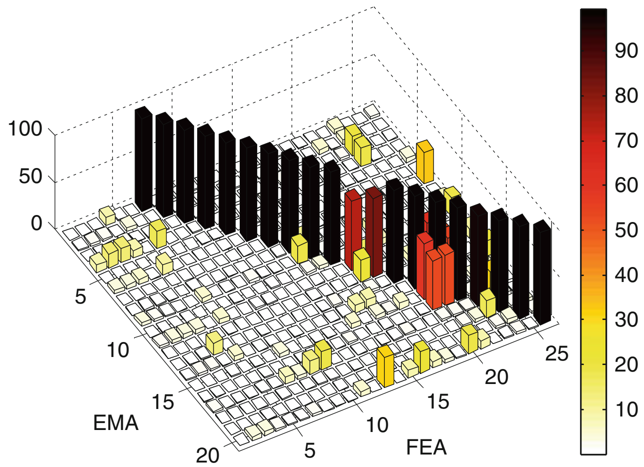


Fig. 16.17 Test 2: MAC matrix between EMA and FEA after nine iterations with the inverse eigensensitivity approach in the original version

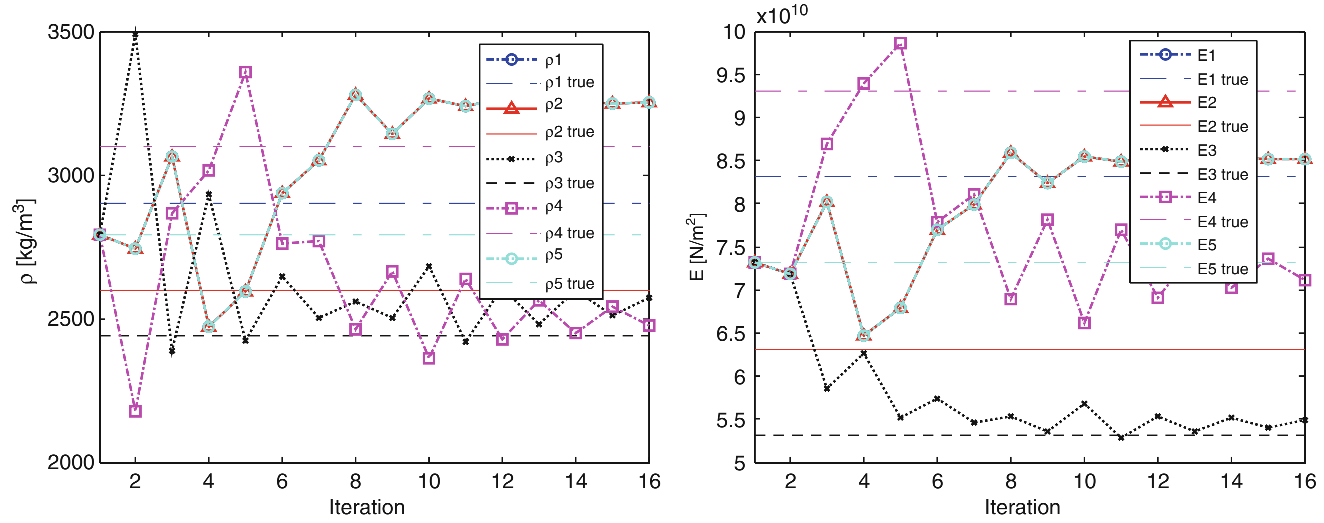


Fig. 16.18 Test 2: density (*left*) and Young's modulus (*right*) during the progressive iterations of the five substructures with the inverse eigensensitivity approach in the modified version

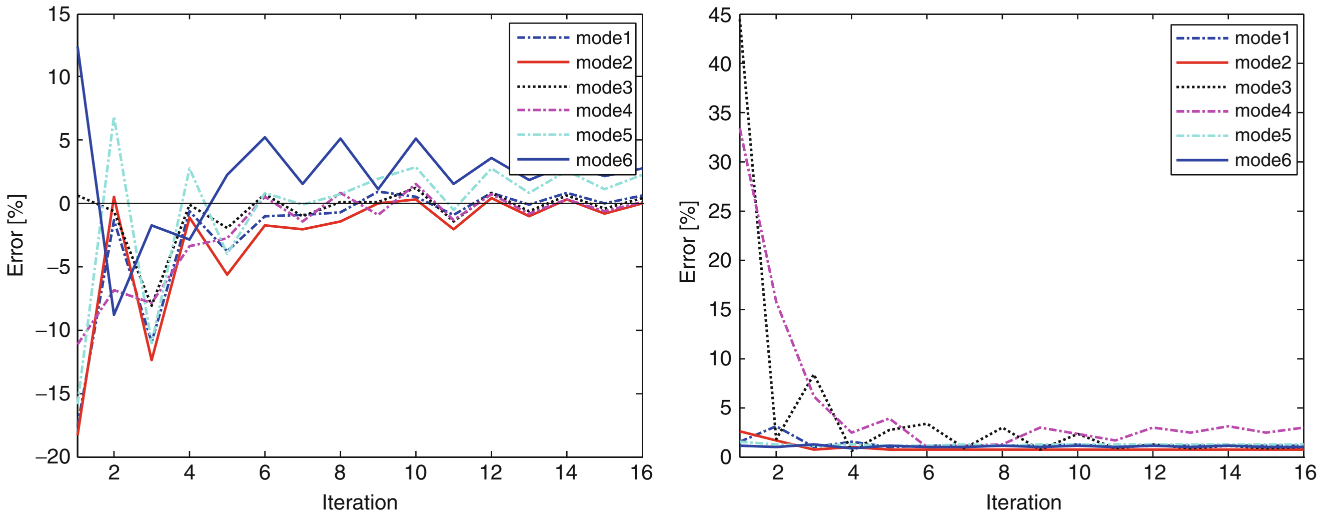


Fig. 16.19 Test 2: eigenvalues (*left*) and eigenvectors errors (*right*) of the first six modes during the progressive iterations with the inverse eigensensitivity approach in the modified version

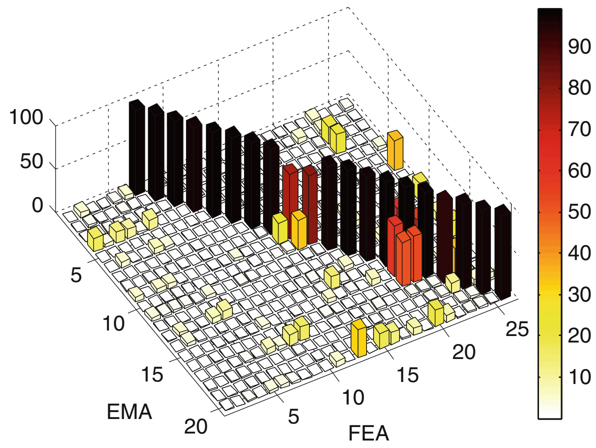


Fig. 16.20 Test 2: MAC matrix between EMA and FEA after 10 iterations with the inverse eigensensitivity approach in the modified version

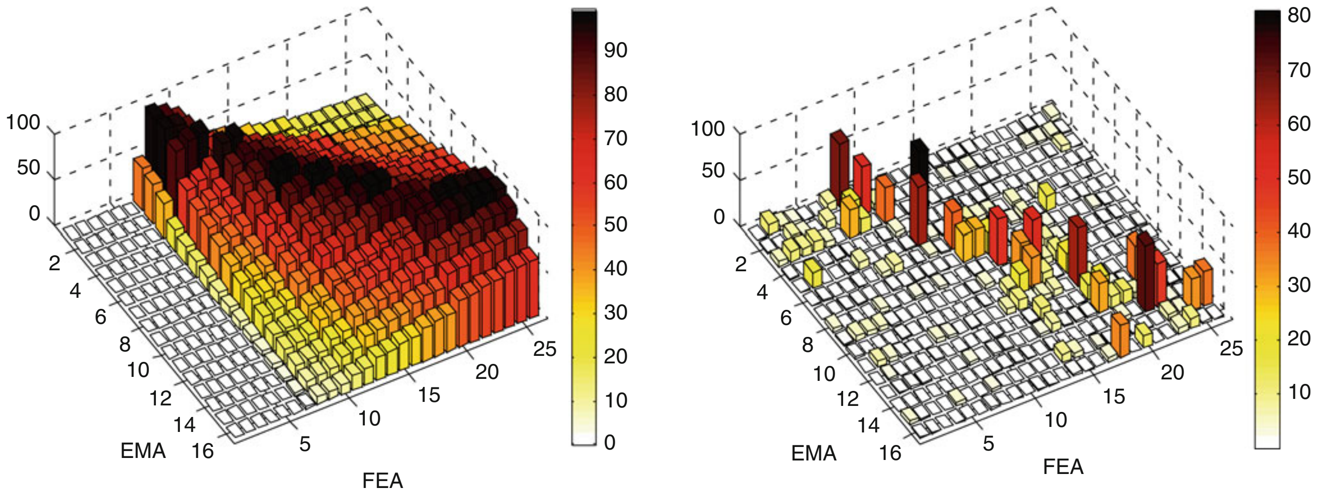


Fig. 16.21 Test 3: eigenvalues errors (*left*) and MAC matrix (*right*) between EMA and FEA

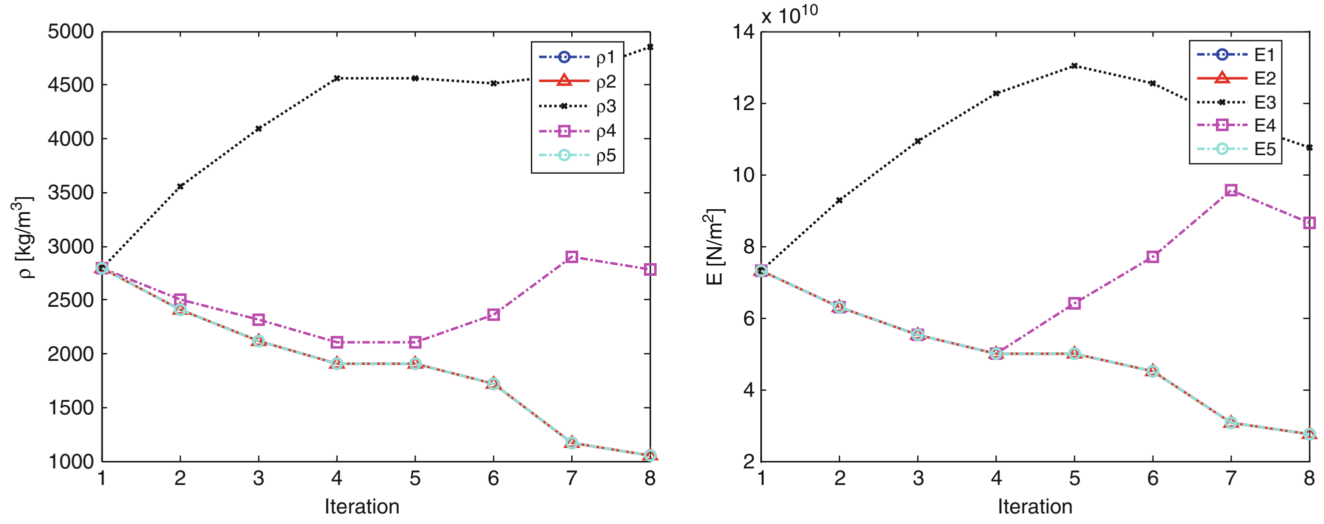


Fig. 16.22 Test 3: density (*left*) and Young's modulus (*right*) during the progressive iterations of the five substructures with the inverse eigensensitivity approach in the original version

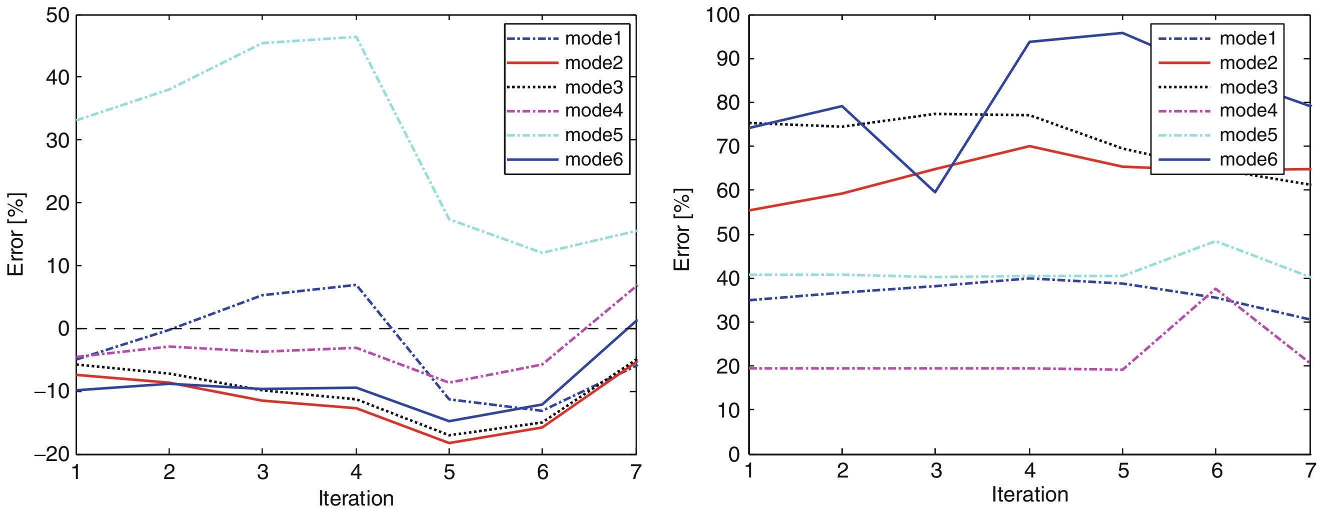


Fig. 16.23 Test 3: eigenvalues (*left*) and eigenvectors errors (*right*) of the first six modes during the progressive iterations with the inverse eigensensitivity approach in the original version

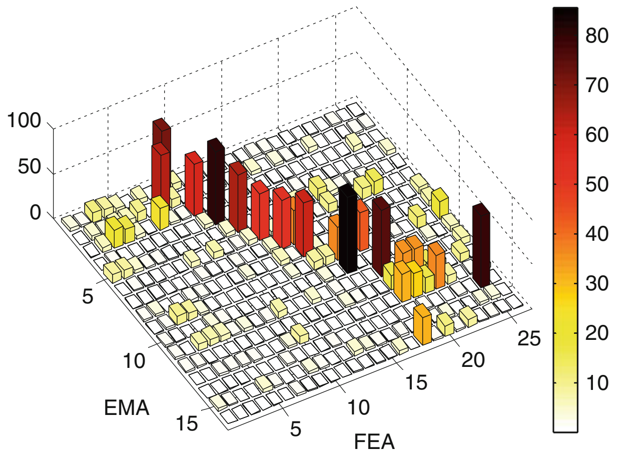


Fig. 16.24 Test 3: MAC matrix between EMA and FEA after seven iterations with the inverse eigensensitivity approach in the original version

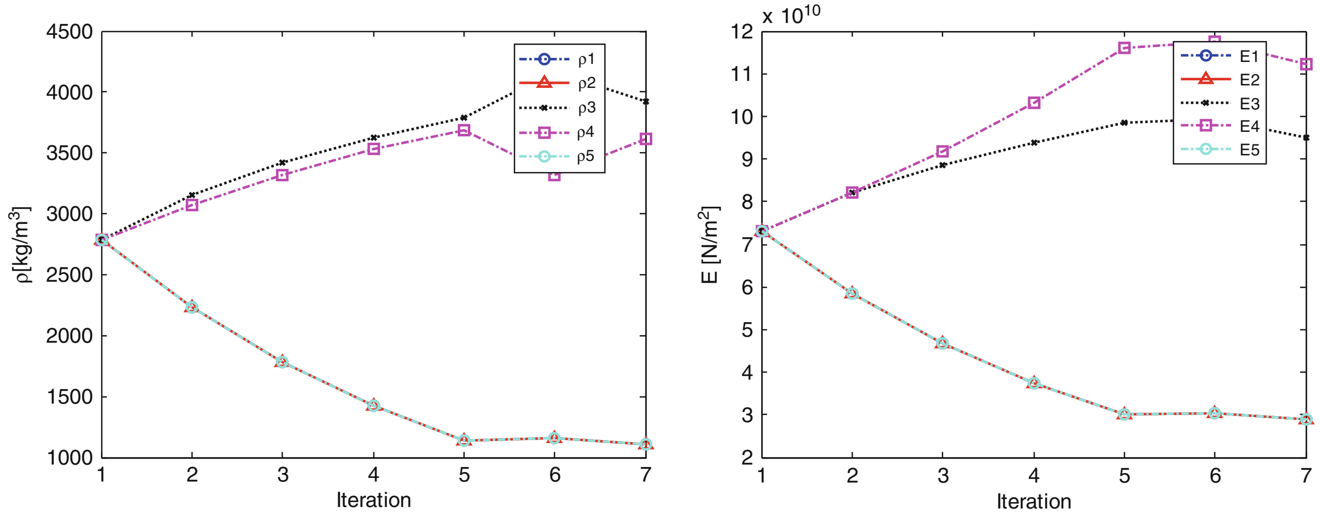


Fig. 16.25 Test 3: density (*left*) and Young's modulus (*right*) during the progressive iterations of the five substructures with the inverse eigensensitivity approach in the modified version

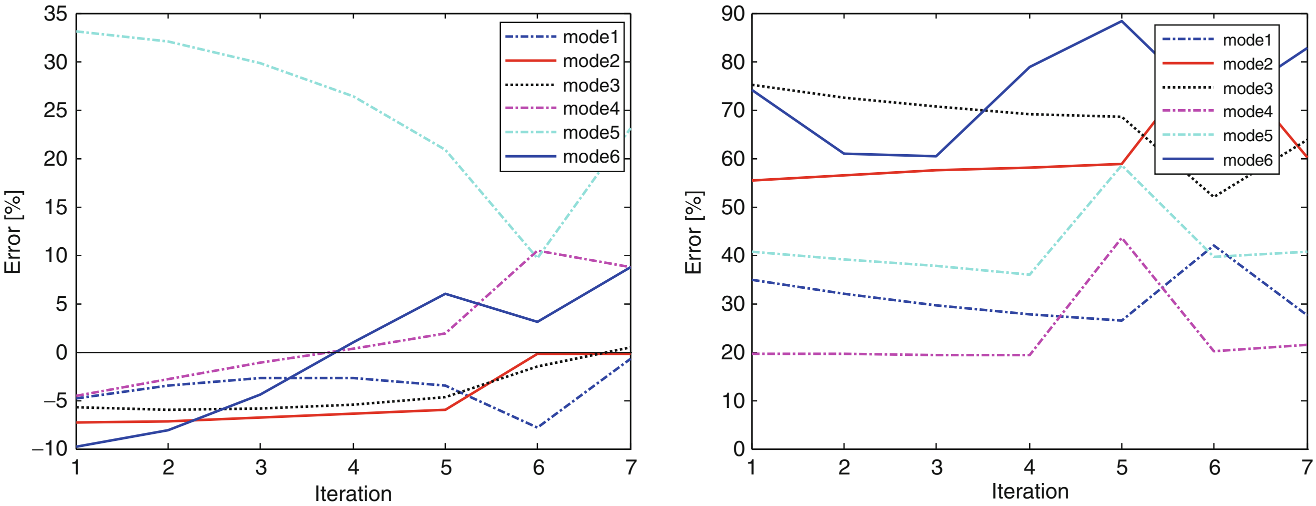


Fig. 16.26 Test 3: eigenvalues (*left*) and eigenvectors errors (*right*) of the first six modes during the progressive iterations with the inverse eigensensitivity approach in the modified version

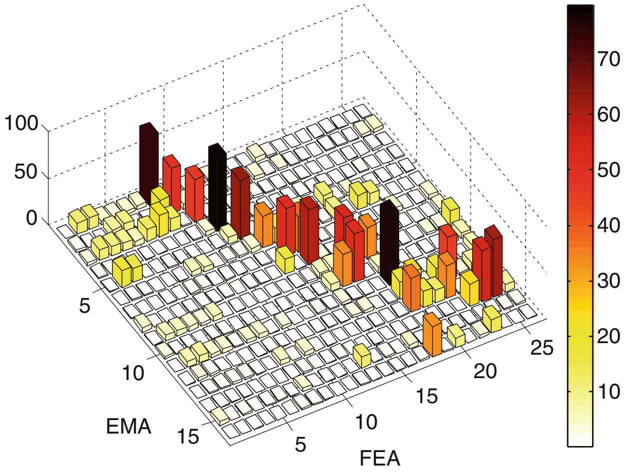


Fig. 16.27 Test 3: MAC matrix between EMA and FEA after seven iterations with the inverse eigensensitivity approach in the modified version

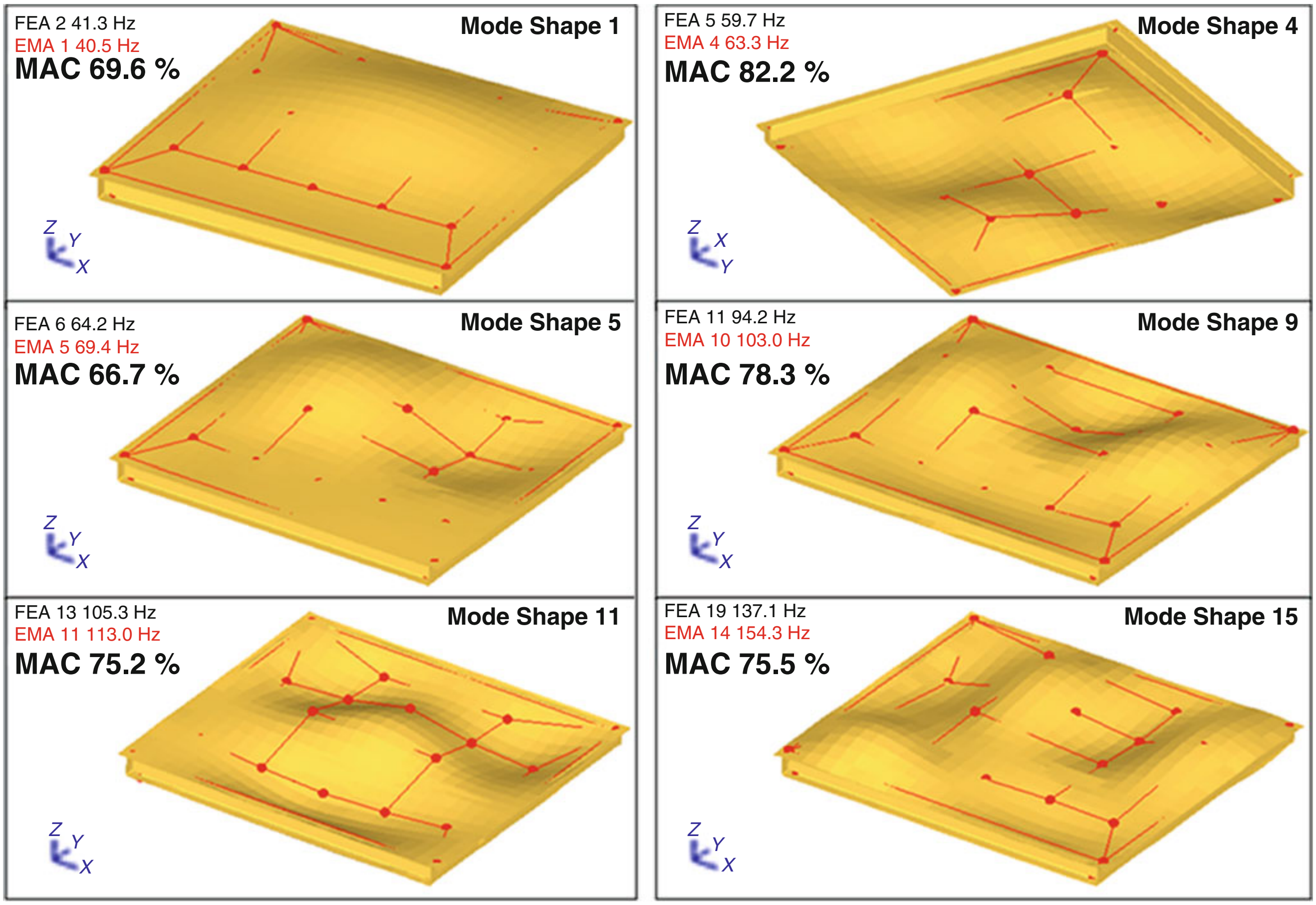


Fig. 16.28 Comparison of the first six EMA and FEA mode-shapes of the top and bottom plates

As a result of the conducted tests, it can be said that the modified method converges like the previous one. The new method is certainly the most unsuitable and leads to less good results with regard to the characteristics of the materials, but the modal characteristics of the equivalent model that is to be found are very similar for the two methods.

In Fig. 16.28, the comparison of the first six EMA and FEA mode-shapes of the top and bottom plates is shown. It confirms the right matching of the modes taken in consideration and the complexity of the model to reproduce the experimental modal characteristics.

16.6 Discussions and Conclusions

On the basis of the above results the following conclusions are drawn:

- In the case of the first test, both methods have eliminated the errors completely;
- In the presence of noise, the convergence for both methods is adversely affected making the iterations relatively unstable especially for the modified method; adding even more noise of about 20%, the results (not shown) are worst; it is reasonable, due to modal data that is not compatible with the considered physical system;
- In the case of the real application both methods lead to a systems not so different, and the results are satisfactory because 12 modes have a good matching after the updating.

In all tests the iterations are stopped when the ratio between time and quality of results is optimized. In conclusion the two methods produce similar results, although the second one is slightly worse because it requires a greater numbers of iterations to converge, but it has also an important advantage of giving the same weight for errors on eigenvalues and eigenvectors. Thus the numerical approximation involved in the sensitivity matrix inversion is reduced.

References

1. Imregun M, Visser WJ (1991) A review of model updating techniques. Shock Vib Dig 23(1):9–20
2. Mottershead JE, Friswell MI (1993) Model updating in structural dynamics: a survey. J Sound Vib 167(2):347–375
3. Collins JD, Hart GC, Hasselman TK, Kennedy B (1974) Statistical identification of structures. AIAA J 12(2):185–190
4. Chen JC, Garba J (1980) Analytical model improvement using modal test results. AIAA J 18(6):684–690
5. Kim KO, Anderson WJ, Sandstorm RE (1983) Non-linear inverse perturbation method in dynamic analysis. AIAA J 21(9):1310–1316
6. Lin RM, Lim MK, Du H (1995) Improved inverse eigensensitivity method for structural analytical model updating. Trans ASME J Vib Acoust 117:192–198
7. Fritzen CP (1986) Identification of mass, damping and stiffness matrices of mechanical systems. ASME Trans J Vib Acoust Stress Reliab 108 (1):9–16
8. Lin RM, Ewins DJ (1990) Model updating using FRF data. In: Proceedings of 15th international modal analysis seminar, KU Leuven, Belgium, pp 141–163
9. Imregun M, Visser WJ, Ewins DJ (1995) Finite element model updating using frequency response function data – I: theory and initial investigation. Mech Syst Signal Process 9(2):187–202
10. Imregun M, Sanliturk KY, Ewins DJ (1995) Finite element model updating using frequency response function data – II: case study on a medium size finite element model. Mech Syst Signal Process 9(2):203–213
11. Imregun M, Ewins DJ, Hagiwara I, Ichikawa TA (1994) Comparison of sensitivity and response function based updating techniques. In: Proceedings of 12th international modal analysis conference, Hawaii, pp 1390–1400
12. Bonisoli E, Fasana A, Garibaldi L, Marchesiello S (2002) Sensitivity analysis and damage location over the Z24 bridge. ONERA, Structural health monitoring, Paris, pp 1007–1015
13. Bonisoli E, Fasana A, Garibaldi L, Marchesiello S, Sabia D (2003) Advanced in identification and fault detection on bridge structures. Key Eng Mater 245–246:339–348
14. Fox RL, Kapoor MP (1968) Rate of change of eigenvalues and eigenvectors. AIAA J 12(6):2426–2429
15. Adhiakri S (1999) Rates of change of eigenvalues and eigenvectors in damped dynamic system. AIAA J 39(11):1452–1457
16. Maia NM, Silva JMM (1997) Theoretical and experimental modal analysis. Research Studies Press, Letchworth

Chapter 17
Shape Optimization of Plates for Desired Natural Frequencies from Coarse Grid Results

Eduardo B.M.R. Germano and Rodrigo Nicoletti

Abstract The design of structures and machines must consider the restrictions imposed by the boundary conditions. Such conditions can be of dynamic nature, thus limiting the frequency ranges that the structure can operate. Among the different design tools available for dealing with dynamic restrictions, shape optimization is a way of deviating natural frequencies from problematic ranges during the design process. In this work, a cantilever plate is optmized aiming at desired natural frequencies. A coarse grid finite element model is correlated to results from experimental modal analysis, and the optimization is done with help of Nastran software. The main contribution is showing that an interpolation of the coarse grid optimization results leads to the solution of the fine grid optimization problem. Hence, it is possible to obtain the smooth geometry that optimizes natural frequencies in plates from coarse grid models, thus not requiring high computational costs.

17.1 Introduction

Metallic panels are widely used in industry as a part of the machine structure. In many cases, panel structures are subjected to weight and dynamic restrictions, and these restrictions must be taken into account in the design process. This is the case of aeronautic and automotive systems. To cope with the dynamic restrictions in panels (frequency ranges and frequency response limits), several dampening mechanisms have been successfully developed, for example bitumen layers, deadeners, vibration absorbers, and active piezo patches. However, the disadvantage of such mechanisms is the addition of mass into the system. For this reason, the study of panel shapes and thickness variation of plates has been motivated, aiming at designing structures with desired dynamic characteristics (natural frequencies, mode shapes, frequency response functions).

The early works on the subject focused on the sole calculation of natural frequencies and mode shapes of plates with given thickness variation. Initially, stepped thickness variation was considered, where areas of the plate had discrete thickness values with sudden local variations [1]. Proper shape optimization of plates aiming at desired eigenfrequencies and mode shapes began to be investigated in the middle of the 1990s. A methodology for the optimization of plate thickness in structures of revolution is presented in [2]. In this work, the authors adopt cubic splines to represent the thickness variation along the plate, and discuss different parameters that can be used as design variables: key points, position vectors, or splines coefficients. Examples of numerical application are presented for conical shells, circular plates, branched shells, and bells. Similar analyzes are presented in [3] and [4] for the optimization problem of prismatic structures formed by folded plates.

The optimization of plates with non uniform variation of thickness (non continuous functions) started with the idea of stepped variation. In this case, the thickness itself is used as the design variable at different regions of the plate (regions to be optimized), with sudden change in thickness values between regions. Analytical procedures and numerical results were obtained for concentric rectangular plates [5] and linear segmented circular plates [6]. Finite element algorithms can also be used and are naturally suitable for the stepped variation approach because finite element models divide the plate area in discrete regions of constant thickness.

E.B.M.R. Germano • R. Nicoletti (✉)
Department of Mechanical Engineering, University of São Paulo, School of Engineering of São Carlos,
Trabalhador São Carlense 400, 13566-590 São Carlos, Brazil
e-mail: bandeiraedu@yahoo.com.br; rnicolet@sc.usp.br

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_17,

The most recent investigations on the shape optimization of plates for desired eigenfrequencies are related to laminated (composite) plates. First attempts on this direction used the orientation of fibers as design variables [7, 8]. Later, other parameters began to be investigated, like the number of plies, number of layers, and the type of composites in the layers [9–11].

Considering the information available in literature, much has been done for eigenfrequency optimization of plates in the numerical field, but few works deal with experimental validation. Actually, the only comparison to experimental results is found in [12]. In the present work, the shape optimization of a cantilever plate aiming at desired natural frequencies is validated against experimental data. In the optimization procedure, one adopts the optimization toolbox of Nastran software and plate local thickness is used as design variables (stepped thickness variation). Results show that the errors in natural frequency values remained below 3%. In addition, and most importantly, it is shown that the interpolation of coarse grid results can be used as an approximate solution to the problem of fine grid stepped thickness variation. Consequently, computational cost is reduced in the optimization procedure, and convergence is more easily assured by the adoption of coarser grids.

17.2 Cantilever Plate Modal Analysis and Model Correlation

The aluminum cantilever plate used in the optimization has dimensions of 150×185 mm, nominal thickness of 2 mm, and material density of 2674 kg/m^3. The identification of the dynamic characteristics of the plate is performed through classical modal analysis procedure [13]. The plate is clamped as shown in Fig. 17.1. The excitation force from the electrodynamic shaker (1) is measured by a load cell (2) connected to the corner of the plate. The plate response is measured by a 2.4 g piezo resistive accelerometer (3) in a grid of 25 measurement points (5×5 grid). The adopted excitation signal is a chirp signal ranging from 10 to 350 Hz in 2 s, in a total of 20 repetitions and 40 s period of acquisition. The adopted acquisition rate is 2 kHz. The frequency response functions are calculated by the H_1 estimator [13] and several tests are performed until clamping condition does not affect significantly the results.

The obtained results in terms of the first three natural frequencies are listed in Table 17.1. The variation observed in the results of the 25 measuring points is mainly caused by the additional mass of the accelerometer over the plate. Despite the uncertainties introduced by the mass of the accelerometer, the variation of the values remained below 2% for the 96% confidence interval (2σ confidence interval). The mode shapes associated to the first three natural frequencies of the plate are obtained by the Ewins-Gleeson Method [13], and the results are shown in Fig. 17.2.

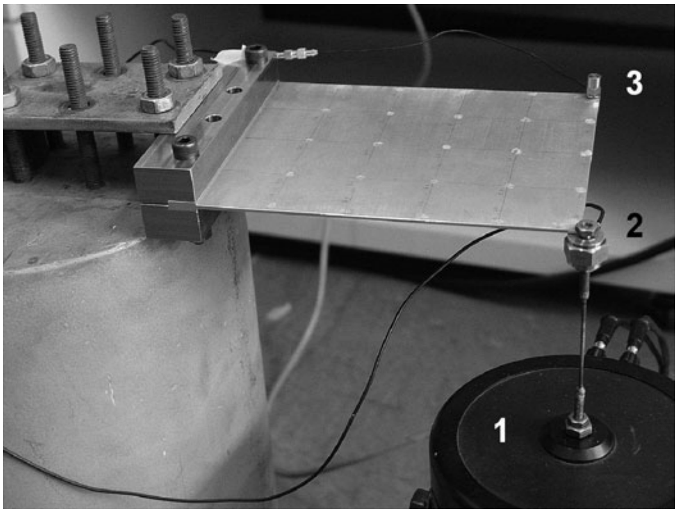


Fig. 17.1 Configuration for the experimental identification of the plate: (*1*) shaker; (*2*) load cell; (*3*) accelerometer

Table 17.1 Natural frequencies identified experimentally for the cantilever plate

Natural frequency	1	2	3
Mean value (Hz)	50.76	135.52	277.48
Standard deviation	0.436	1.558	2.387
96% confidence interval (Hz)	50.8 ± 0.9	136 ± 3	278 ± 5

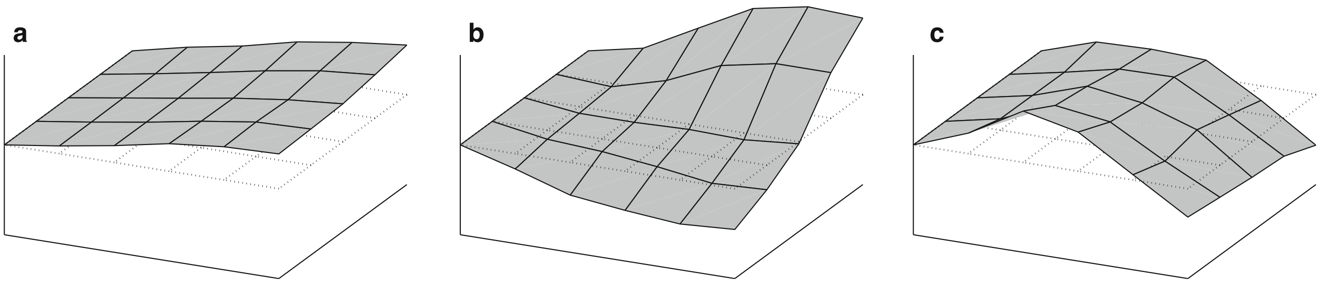


Fig. 17.2 Experimental mode shapes of the cantilever plate. (a) First mode (50.8 Hz). (b) Second mode (136 Hz). (c) Third mode (278 Hz)

Table 17.2 Natural frequencies of the 4×5 and 48×60 mesh models

	4×5			48×60			48×60		
Mesh model	No add mass			No add mass			5 g mass		
Mode	1	2	3	1	2	3	1	2	3
Natural frequency (Hz)	49.3	134.7	291.9	50.2	144.2	310.7	47.4	127.7	298.5
Error (%)	2.9	0.6	5.0	1.2	6.0	11.7	6.7	6.1	7.5

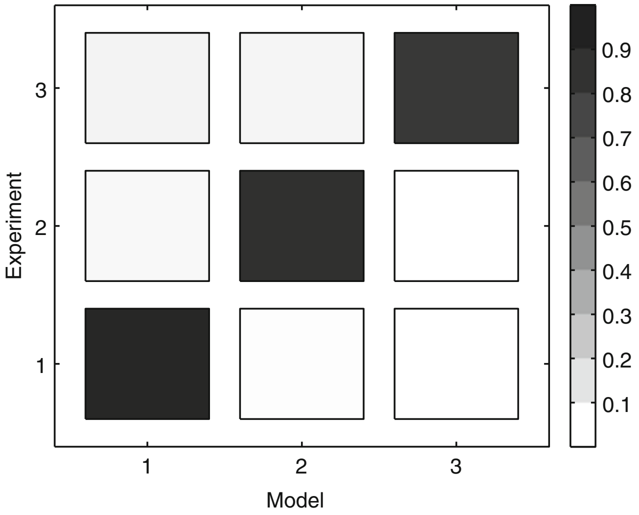
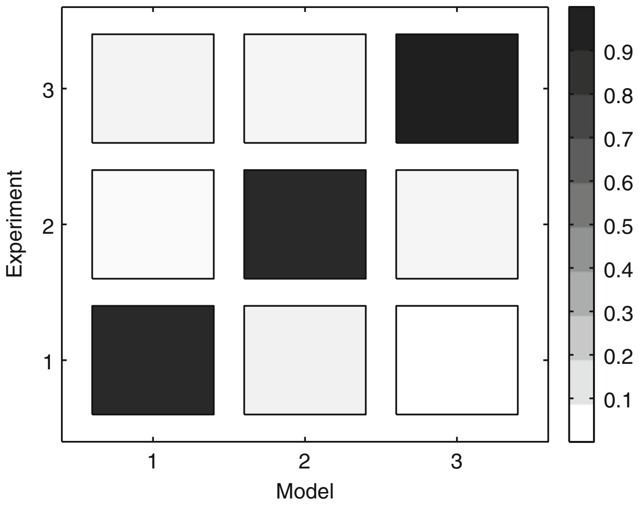


Fig. 17.3 MAC matrix of the 4×5 mesh model

In order to optimize the plate thickness towards desired natural frequencies, a finite element model is developed using MSC Nastran v.2005 finite element algorithm. In this case, a 4×5 element mesh is used. Elements are CQUAD4 shell elements with 2 mm thickness and material is considered isotropic (MAT1), with Young modulus of 7.1×10^{10} N/m^2 and Poisson ratio of 0.3. All nodes in the clamped edge of the plate are fully constrained (zero displacements and rotations in all directions). The natural frequencies and associated mode shapes are obtained with solution 103 of Nastran software.

By adopting the 4×5 mesh model and calculating the respective natural frequencies and mode shapes, one obtained the results presented in Table 17.2. The error between numerical and experimental results remained below 5% for the first three natural frequencies. Mode shapes are also coherent to those experimentally measured as evidenced by the MAC matrix (Modal Assurance Criterion) shown in Fig. 17.3, whose strong diagonal preponderance gives an indication of the similarity between the numerical and the experimental modes [14]. Considering that natural frequency errors are within acceptable margins, the model is thought to be a fairly good representation of the cantilever plate in study.

Fig. 17.4 MAC matrix of the 48×60 mesh model with 5 g additional mass



Similar results are obtained with a fine mesh model of 48×60 elements. In this case, better results are obtained for a model with additional 5 g mass in the excitation position (Table 17.2, Fig. 17.4). For this reason, the 48×60 element model with 5 g additional mass at the excitation position is also considered to be a good representation of the system in study (cantilever plate).

17.3 Shape Optimization

The optimization of the plate thickness aiming at desired natural frequencies is performed with solution 200 of Nastran software. The adopted objective function is:

$$f = \left(1 - \frac{\omega_{n1}}{\Omega_1}\right)^2 + \left(1 - \frac{\omega_{n2}}{\Omega_2}\right)^2 + \left(1 - \frac{\omega_{n3}}{\Omega_3}\right)^2 \tag{17.1}$$

where ω_{ni} is the i-th natural frequency of the plate (numerical) and Ω_i is the i-th desired natural frequency for the plate. The thickness of every element in the model (t_i) is a design variable of the optimization problem, and they are subjected to the side constraints:

$$1.5 \leq t_i \leq 5.0\ mm \tag{17.2}$$

The initial values of t_i are 2 mm.

The natural frequencies of the plate are 50.8, 136, and 278 Hz, as identified previously. Let's change the second natural frequency to 100 Hz, keeping the other frequencies at the same values. Thus, the desired natural frequencies are: 50.8, 100, and 278 Hz. By adopting the finite element model described earlier (4×5 model), and solving the optimization problem ((17.1) and (17.2)) with help of Nastran software, one obtains the results shown in Table 17.3 and Fig. 17.5. The results shown in Fig. 17.5 refer to a thickness equally distributed in relation to the plate middle surface (t_i / 2 in the bottom surface and t_i / 2 in the upper surface of the plate).

By checking the optimized solution (Fig. 17.5), one can see that mass is concentrated in the free tips of the plate (upper left and right corners) and in the center region near the clamped edge (bottom). This makes sense, considering that the objective was to lower the torsion natural frequency without affecting bending natural frequencies. The concentrated masses in the free corners of the plate will reduce torsion natural frequency. In order to keep the bending frequencies unchanged, more mass is located near the clamping to balance the additional mass in the free edge of the plate and to increase bending stiffness.

Table 17.3 Natural frequencies of the optimized model

	4×5 model		
	1	2	3
Natural frequency (Hz)	48.7	101.6	274.3
Error (%)	4.1	1.6	1.3

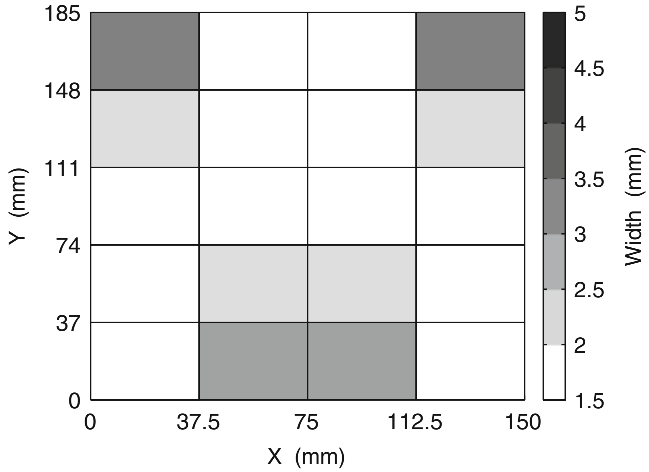


Fig. 17.5 Optimized thickness distribution over the cantilever plate for desired frequencies of 50.8, 100, and 278 Hz (clamped edge in the *bottom*) – 4×5 model

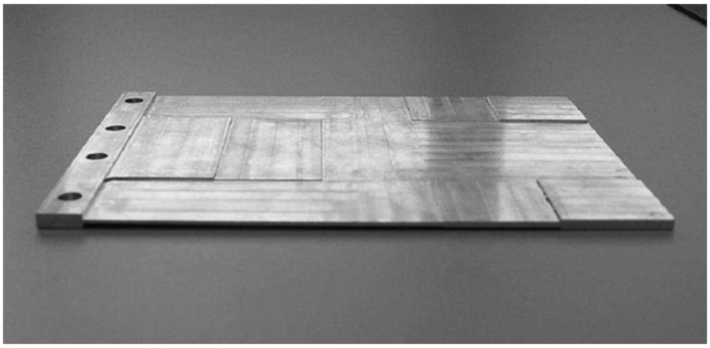


Fig. 17.6 Aluminum cantilever plate with optimized thickness (4×5 element configuration)

The time required to run the optimization algorithm remained below 5 min in a desktop microcomputer Pentium 4 with 3.2 GHz processor and 2.5 GB RAM memory, taking five design cycles to converge to a solution. The optimization problem using the fine mesh model with 48×60 elements did not converge to a solution. Hence, it was possible to find optimized shapes for the plate in a stepped thickness configuration (coarse grid results), but an optimized smooth shape of the plate was not obtained, as it would be expected from the fine mesh model with 48×60 elements (fine grid results). In the later case, the number of design variables (2,880 variables) was too large for the algorithm to cope with (insufficient computer memory).

17.4 Optimized Plate Modal Analysis

In order to experimentally check the optimization results, the solution obtained for the 4×5 element model (Fig. 17.5) was machined with the same material sample used to machine the original cantilever plate (Fig. 17.6). The identification of the dynamic characteristics of the plate is performed according to the same experimental procedure used in the identification of the original plate. Again, several tests are performed until clamping condition does not affect significantly the results.

The obtained results in terms of the first three natural frequencies are listed in Table 17.4. The variation observed in the results of the 25 measuring points is mainly caused, again, by the additional mass of the accelerometer over the plate. The variation of the values remained below 3% for the 96% confidence interval. Comparing the experimental results to those predicted by the numerical simulation (Table 17.3), a remarkably good agreement is found, with bigger error in the

Table 17.4 Natural frequencies identified experimentally for the optimized cantilever plate

Natural frequency	1	2	3
Mean value (Hz)	48.75	102.40	275.96
Standard deviation	0.577	1.581	2.300
96% confidence interval (Hz)	49.8 ± 1.2	102 ± 3	276 ± 5
Predicted in Table 17.3 (Hz)	48.7	101.6	274.3
Error (%)	2.2	0.4	0.6
Desired (Hz)	50.8	100	278
Error (%)	1.9	2.0	0.7

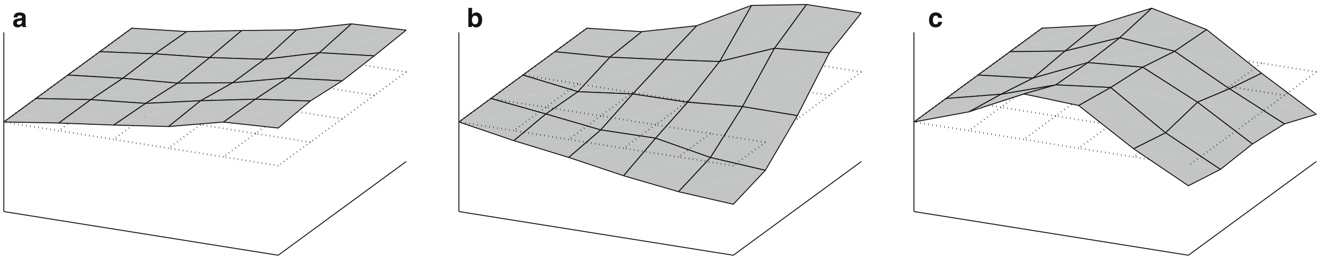


Fig. 17.7 Experimental mode shapes of the optimized cantilever plate. (**a**) First mode (49.8 Hz). (**b**) Second mode (102 Hz). (**c**) Third mode (276 Hz)

prediction of the first natural frequency (2.2%). Finally, comparing the obtained natural frequencies to the desired ones, again good agreement is achieved, remaining the error below 2%.

The mode shapes associated to the first three natural frequencies of the optimized plate are also obtained by the Ewins-Gleeson Method. The results in Fig. 17.7 show that the mode shapes did not change in comparison to those of the original plate (Fig. 17.2).

Hence, the placement of natural frequencies of a cantilever plate, by changing its shape, was successfully achieved using a 4×5 finite element model (stepped thickness distribution). A faster manufacturing solution is stamping a smooth surface on the plate. For that, one must know the surface to be stamped, which could be obtained from the results of the optimization problem using a finer mesh model. The problem is that finer mesh models have a large number of design variables, and this can jeopardize algorithm convergence as observed in the case of the 48×60 element model. An alternative solution is interpolating the results obtained for the coarse grid, which are computationally cheap and achieve good experimental results, into a fine grid of points to form the smooth surface of the plate.

17.5 Interpolation of Coarse Grid Results

The interpolation of the optimized results is done with help of command interp2 of MatLab software. This interpolation command is suitable for functions in two dimensions (surfaces) and bases on the Delaunay triangulation of data using the Quickhull algorithm [15]. One can adopt linear and cubic interpolating functions in the procedure.

By interpolating the results obtained for the coarse grid model (4×5 model), one obtained the results shown in Fig. 17.8. In both cases (linear and cubic interpolation), one considered the thickness in the center position of each element on the plate. The results were interpolated into a fine 48×60 element grid. As one can see in Fig. 17.8, there is a clear difference between the linear and the cubic interpolation due to the coarse nature of the original thickness distribution. The smoother thickness distribution obtained by interpolating the coarse grid optimization results is used to calculate the natural frequencies of the plate (now in a 48×60 element model). The results are presented in Table 17.5.

Good results are obtained from the interpolation of the coarse grid optimized results. Errors below 10% are achieved for all natural frequencies. The linear interpolation presents better results for the second and third natural frequencies, whereas the cubic interpolation gives better results for the first natural frequency.

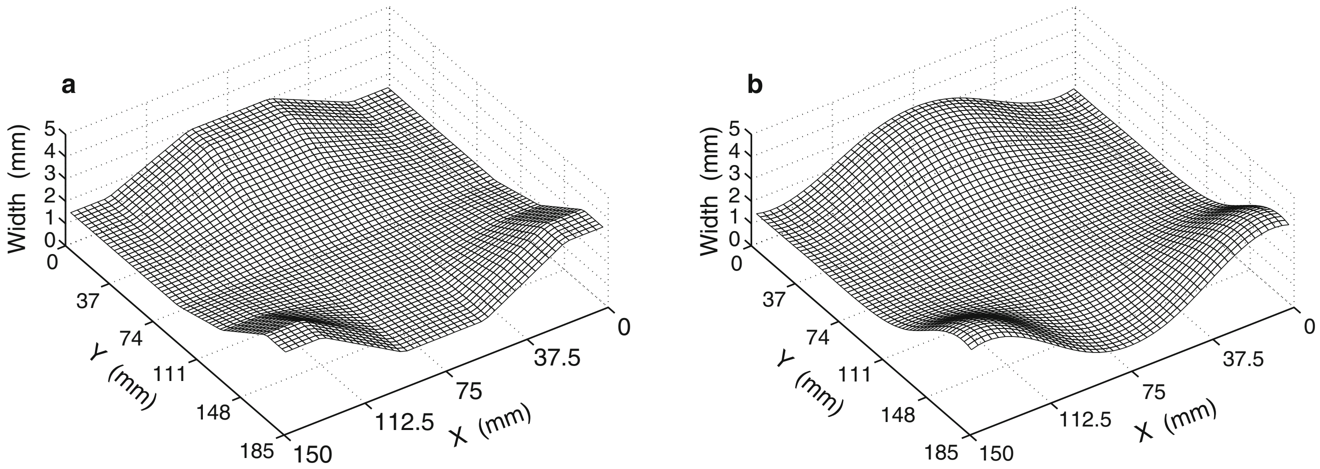


Fig. 17.8 Interpolating surfaces of the optimized results obtained from the coarse grid model into a 48×60 element grid. (**a**) Linear interpolation. (**b**) Cubic interpolation

Table 17.5 Natural frequencies calculated with interpolated thickness distribution with a 48×60 element model

Natural frequency	1	2	3
Desired (Hz)	50.8	100	278
4×5 linear (Hz)	46.6	100.9	289.0
Error (%)	8.3	0.9	3.9
4×5 cubic (Hz)	46.8	96.5	290.3
Error (%)	7.9	3.5	4.4

17.6 Conclusion

The application of the optimization procedure for tuning the natural frequencies of a metal plate with varying thickness resulted in the following conclusions:

- One can achieve successful results from the optimization of plate thickness using a commercial finite element software. The error between the desired and the experimental values of natural frequencies remained below 3%.
- The interpolation of coarse grid results from the finite element analysis leads to the approximated solution of the optimization problem for the fine grid model.
- The linear interpolation leads to slightly better results than those obtained with the cubic interpolation.

It is important to note that good model correlation is crucial to obtain successful results, especially in the shape optimization procedure. Such correlation is strongly affected by the boundary conditions of the plate (supporting assumption), which in the studied case (clamped-free-free-free) was simple enough to be well represented and updated in the model.

References

1. Irie T, Yamada G, Ikari H (1980) Natural frequencies of stepped thickness rectangular plates. Int J Mech Sci 22:767–777
2. Özakça M, Hinton E (1994) Free vibration analysis and optimization of axisymmetric plates and shells – II: shape optimization. Comput Struct 52:1199–1211
3. Afonso SMB, Hinton E (1995) Free vibration analysis and shape optimization of variable thickness plates and shells – II: sensitivity analysis and shape optimization. Comput Syst Eng 6:47–66
4. Hinton E, Özakça M, Rao NVR (1995) Free vibration analysis and shape optimization of variable thickness plates, prismatic folded plates and curved shells, part 2: shape optimization. J Sound Vib 181:567–581
5. Gutierrez RH, Rossi RE, Laura PAA (1995) Determination of the fundamental frequency of transverse vibration of rectangular plates when the thickness varies in a discontinuous fashion. Ocean Eng 22:663–668

6. Chou FS, Wang CM (1996) Optimization of linear segmented circular Mindlin plates for maximum fundamental frequency. Struct Optim 11:128–133
7. Dianconu CG, Sato M, Sekine H (2002) Layup optimization of symmetrically laminated thick plates for fundamental frequencies using lamination parameters. Struct Multidiscip Optim 24:302–311
8. Topal U (2009) Frequency optimization of laminated general quadrilateral and trapezoidal thin plates. Mater Des 30:3643–3652
9. Wang X, Oguamanam DCD, Hansen JS (2005) Shape optimization of stiffers in stiffened composite plates with thermal residual stresses. Strut Multidiscip Optim 30:38–42
10. Bargh HG, Sadr MH (2011) Optimal design by elitist-genetic algorithm for maximum fundamental frequency of fiber metal laminated plates. Key Eng Mater 471–472:331–336
11. Karakaya S, Soykasap Ö (2011) Natural frequency and buckling optimization of laminated hybrid composite plates using genetic algorithm and simulated annealing. Struct Multidiscip Optim 43:61–72
12. Ercoli L, Laura PAA, Gil R, Carnicer R, Sanzi HC (1990) Fundamental frequency of vibration of rectangular plates of discontinuously varying thickness with a free edge: analytical and experimental results. J Sound Vib 141:221–229
13. Maia NMM, Silva JMM (1997) Theoretical and experimental modal analysis. Research Studies Press, Taunton
14. Allemang RJ (2003) The modal assurance criterion – twenty years of use and abuse. Sound Vib 8:14–21
15. Barber CB, Dobkin DP, Huhdanpaa HT (1996) The Quickhull algorithm for convex hulls. ACM T Math Softw 22:469–483

Chapter 18
Model Updating of Complex Assembly Structures Based on Substructures-Joint Parameters

Morteza H. Sadeghi, Parivash Soleimanian, and Hamed Samandari

Abstract To study the dynamic behavior of complex assembled structures consisting of several substructures and real joints connecting them, an updated finite element model of the associated structure is required. This paper presents a new technique to create an accurate updated finite element model of such structures. Given the fact that modal testing of real joints (such as bolt with some washers) are almost impossible; in this research the updated model of the assembled structures is constructed by utilizing parametric finite element model of the joint in conjunction with modal testing of the assembly structure and its substructures. In this paper, eigen-sensitivity method (used for characterizing cost function) and genetic algorithm (used for minimization scheme) are employed to update the assembled structure as well as substructures. A laboratory-scale unsymmetrical cross-beam is employed as the case study. The actual dynamic properties of the joint (including stiffness, mass and damping matrix) of this structure were estimated. The accuracy of the estimated parameters of the model was examined by comparison of the FRFs of the real assembled structure with the ones of the updated model. By achieving full compliance between these FRFs, the accuracy and efficiency of the proposed method, in a wide frequency range, is demonstrated.

Keywords Model updating • Genetic algorithm • Eigen-sensitivity • FRF • Substructure • Finite element

18.1 Introduction

Most of the assembly structures consisting of some substructures connected by joints play a significant role in dynamic properties of these systems. FE model updating has become a viable approach to increase the correlation between the dynamic response of a structure and the predictions from a model. It is well known that a mathematical FE model will be incorrect due to unavoidable difficulties in modeling of joints, damping and boundary conditions. It is clear that the experimental data are generally considered to be more accurate. Extracting a model in experimental approach also confronts problems due to limited number of measured degrees of freedom (DOFs), limited frequency range and difficulty in measurement of rotational-torsional DOFs [1, 2]. This has led to the progress of model updating techniques which aims at reducing the inaccuracies present in the analytical model in the light of measured dynamic test data.

M.H. Sadeghi • P. Soleimanian (✉)
Tabriz University, Tabriz, Iran
e-mail: morteza@tabrizu.ac.ir; soleimanianp@yahoo.com

H. Samandari
Middle East Technical University, Ankara, Turkey
e-mail: Hamed.samandari@metu.edu.tr

T. Simmermacher et al. (eds.), *Topics in Model Validation and Uncertainty Quantification, Volume 4*,
Conference Proceedings of the Society for Experimental Mechanics Series 29,
DOI 10.1007/978-1-4614-2431-4_18,

The substructuring method is preferable to cope with complex structures such as Cross-Beam. In general, the individual substructures are analyzed separately to obtain designated solutions, which are subsequently assembled to recover the properties of the global structure by constraining the interfaces of the adjacent substructures [3]. The substructuring method is advantageous mainly in three aspects. First, as the global structure is replaced by smaller and more manageable substructures, it is much easier and quicker to analyze the small system matrices. Second, the separated substructures are analyzed independently. When one or more substructures are modified, only the modified substructures need to be re-analyzed while the others remain unchanged [4]. This property can be promising when applied to model updating. When the uncertain parameters are localized within parts of a structure, only one or more substructures containing those parts are re-analyzed during model updating, and the other substructures are untouched [5, 6]. The substructuring method will be more efficient when some identical substructures exist or when the substructuring method is incorporated with parallel computation. Finally, the substructuring method is helpful to be combined with the model reduction technique in calculating the eigensolutions and eigensensitivities [7, 8].

Model updating is a procedure in which the errors in the computational model are corrected by minimizing the discrepancy between the predicted and the actually measured response data [9]. The most commonly used experimental data for model updating are the dynamic modal data. Because the dynamic modal data are inherently dependent upon the structural properties, minimizing the difference between the predicted and measured dynamic properties will lead to the identification of the correct physical model parameters.

The traditional techniques for performing dynamic model updating are mostly iterative-based [10, 11], and they usually employ the sensitivity analysis to directly update the structural physical parameters to achieve a match between the analytical and experimental data. Some example applications can be found in Refs. [12, 13]. However, iterative-based updating methods bear the limitation that the searching procedure may be trapped into a local optimum instead of the actual global solution; and moreover the modal data sensitivities must be computed in each iteration, which is very time-consuming and troublesome due to the high non-linearity of mode shape sensitivity with respect to the physical parameters [14, 15]. The recent development in artificial intelligence algorithms provides desirable alternatives to sensitivity-based optimization approaches in searching for the true optimum. Amongst them, genetic algorithms (GAs) appear to be particularly attractive because they operate to globally optimize a given objective function and hence are capable of finding a global solution. The GA searches for the candidate solution from a large population of points rather than a single point as in traditional optimization methods. It does not require the continuity or differentiability of the objective function. Thus, constructing the objective (error) function becomes very flexible. Recently, a few researchers have applied GAs in model updating and damage detection problems [16–19]. For example, Levin and Lieven [18] applied GA and the blended simulated annealing (BSA) technique for FE model updating, using directly the frequency response function (FRF) data. They also compared the binary GA (BGA) with BSA and discussed the appropriate choice of updating parameters after some trial analyses. David and Keng [19] also applied the binary coding GA approach to solve a model refinement problem. They proposed a new method for the mode shape correlation in the formulation of the GA fitness function to tackle the problems associated with closely spaced modes of vibration, and also touched on the solution quality in connection with the relationship between the chosen response and the physical parameters. The binary coding scheme adopted in the above studies, however, is computationally intensive and hence could become intractable if the number of parameters to be updated is large such as in real life structures. However, the real number coding GA, which encodes each parameter using only one real number in the chromosome, can substantially reduce the computational effort. In addition, the real number representation also prevents the loss of precision which could be encountered in the binary-coding GA [20]. Most of the updating methods neglect the damping. In complex structures such as Cross-Beam, these methods cannot be used for predicting complex FRFs and damping especially in a wide frequency range. All structures have some form of damping, but despite a large study on the damping, it still remains one of the apprehensions in general vibration analysis. Adhikari and Woodhouse [21] identified the damping of the system as viscous damping. However, viscous damping is not only damping of a model.

In this paper, GA, Eigen-Sensitivity and FRF-Based approach, is presented for the dynamic model updating of complex assembly structures consisting of substructures and real joints based on the measured dynamic properties. At the outset of this paper, the Eigen-sensitivity based approach using GA algorithm is incorporated to perform the updated model of substructures. In the following by using the obtained updated results of substructures and considering properties of joint as updated variables, the updated model of assembly structure is created by using GA algorithm in a wide frequency range.

Damping of joint is updated by using FRF-based model updating approach and minimizing related objective function. In the final step, the updated FRFs of assembly structure is achieved by using updated damping of joint and damping of substructures, and the assembly structure extracted from the measured data.

18.2 Eigensensitivity-Based Model Updating Method

Model updating by nature is to make adjustment on the variable structural parameters so that the predicted responce by the model best matches the measured counterparts. In sensitivity-based model updating procedure, different objective functions which minimized by a sensitivity-based iterative procedures are used. In this paper the following objective function is used as a cost function that optimized by genetic algorithm (GA).

$$J(p) = \sum_{i=1}^{N_f} \left\| \frac{(\omega_x)_i^2 - (\omega_A(p))_i^2}{(\omega_x)_i^2} \right\| \tag{18.1}$$

Where N_f is number of measured frequency points, p is the vector of design variables and $(\omega_A)_i$ and $(\omega_x)_i$ are the ith corresponding analytical and experimental eigenvalues, while $J(p)$ is a function of updating parameters.

In this model updating method the eigensolutions are calculated from the system matrices of the global structure, based on the eigenequation:

$$\mathbf{K}\{\varphi_i\} = \lambda_i \mathbf{M}\{\varphi_i\} \tag{18.2}$$

Where **K** and **M** are the stiffness and mass matrices, and λ_i, $\{\varphi_i\}$ are the ith eigenvalue and eigenvector, respectively.

Updated mass and stiffness matrices are the results of this section. The objective function is minimized by continuously adjusting the parameters {r} of the initial FE model through optimization searching techniques. The optimization genetic algorithm (GA) is used to minimize the objective function in a trust region.

18.3 Overview of the Genetic Algorithm

In GAs, the searching procedure derives from the process of natural selection and evolution. The distinguished feature of genetic algorithms lies in their inherent advantage that following the rule of 'survival of the fittest' they can arrive at the global optima. GAs differ from calculus-based search procedures in several ways [22], namely, (a) they search from a population of points, not a single point; (b) they make use of information about the objective function itself, not derivatives or other auxiliary knowledge; and (c) they adopt probabilistic rather than deterministic transition rules. The basic elements in GAs are 'strings', or in genetic terminology 'chromosomes', which code the information about the phenotype of variables. An individual chromosome is composed of a number of 'genes'. Different numbers of genes may be used to represent different variables. The choice of a proper form of genes for a particular type of problem can be vital to the success of the optimization. For example, when dealing with discrete problems, binary genes may be preferred, while for continuous problems (e.g. the FE model updating problems concerned in this paper) the use of real number genes is favoured for efficiency consideration [22]. For each chromosome, a quantity called 'fitness' is defined to measure the quality of the solution represented by this chromosome. The 'fitness' of chromosomes refers to the corresponding values of the defined objective function. The genetic optimization procedure proceeds step-by-step. After the initialization stage the genetic operations, i.e. selection, crossover and mutation, are activated repeatedly.

18.3.1 Initialization

First, the GA is initialized by establishing the starting population, which can be generated either randomly or heuristically. The choice of the population size, i.e. the number of strings, is a trade-off decision between computing time and the requirement on the precision of the optimal solution. In general, a larger population size can result in a better solution. As an illustration, assuming the underlying problem has four variables, the initial population with real number genes can take the following form:

Chromosome 1: 0.43 0.61 0.85 0.92
Chromosome 2: 0.52 0.67 0.28 0.52
Chromosome 3: 0.36 0.49 0.65 0.47
Chromosome 4: 0.97 0.57 0.63 0.59

In practice, the population size could be far larger than the above. After the initial population is established, a new population is to be produced through the genetic operations, as discussed in what follows.

18.3.2 Selection

In this process, a decision is made on the number of copies for each chromosome into the next generation. Typical selection strategies include roulette wheel selection and ranking selection. The roulette wheel selection, owing to its full mapping scheme, limits the GA to only perform a maximization problem. In contrast, ranking selection methods only require the objective function to map the solutions to a partially ordered set. They work by assigning selection probability pi for each individual based on the rank of its fitness, and thus minimization problems and negative objective functions can be tackled. In the present study, the ranking selection is adopted. Via the selection operation, the number of chromosomes with fitness above the average will increase exponentially in future generations.

18.3.3 Crossover

This operation reproduces the offspring population by recombining the information from two parent chromosomes. The simplest type of crossover is single point crossover, in which two parent chromosomes are selected randomly and cross over their genes at a randomly chosen position to produce two offspring. For example, suppose the two parents consisting of m genes $x = [x_1 \quad x_2 \quad .... \; x_m]$ and $y = [y_1 \quad y_2 \quad .... \; y_m]$. If a crossover point is chosen at the kth position counting from the left side, then the resulting two offspring are:

$$x' = [x_1 \quad x_2 \quad .... \; x_k \quad y_{k+1} \quad .... \quad y_m] \quad y' = [y_1 \quad y_2 \quad .... \; y_k \quad x_{k+1} \quad .... \quad x_m]$$

For the simple case mentioned earlier, if Chromosome 1 and Chromosome 3 are chosen as parents and the crossover point is randomly selected at the second position, the two descendants from crossover would be:

Offspring 1: 0.43 0.61 0.65 0.47
Offspring 2: 0.36 0.49 0.85 0.92

Depending on the coding schemes, there are many other ways to perform this operation, such as the two-point and the uniform crossover in binary-coding GAs and the arithmetic crossover in real-coding GAs. This paper, however, adopts the so-called ‘heuristic’ crossover as it can generally result in offspring with improved fitness in real-coding GA. Assuming that parent X is better than Y in terms of fitness, the heuristic crossover [23] operates such that

$$\begin{aligned} X' &= X + r(X \quad Y) \\ Y' &= X \end{aligned}$$

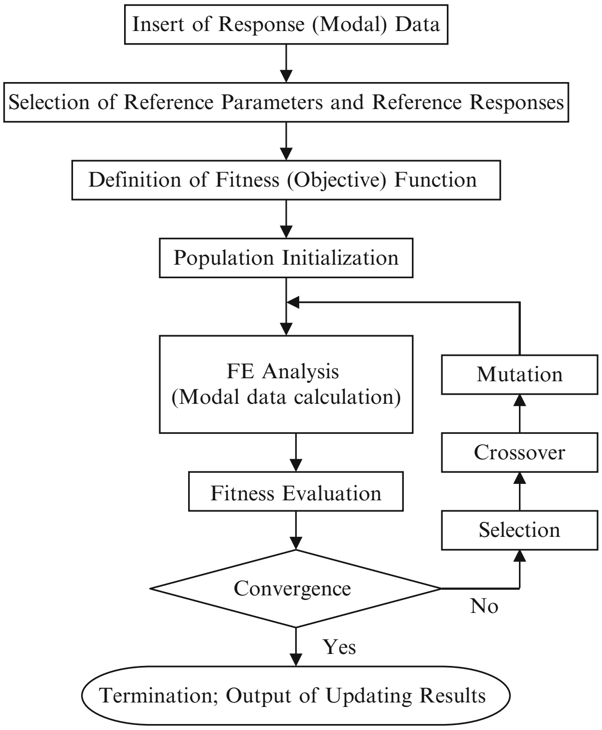
where X' and Y' are the two newborn descendants and r takes a random number between 0 and 1. In case any variable coded in X' violates the corresponding boundary condition, the above process has to be repeated until a reasonable X' emerges.

18.3.4 Mutation

This final operator is motivated to help further increase the genetic diversity and improve the global nature of the search. It simply alters one gene (representing a variable) in a chromosome based on a certain rule. For example, boundary mutation in a real-coding GA only replaces the value of a randomly chosen gene by one of its two bounds. There exist many mutation methods for different application and encoding schemes. In this paper, uniform mutation is used such that the selected gene is superseded with a real number randomly drawn within its prescribed boundary.

Following the selection, crossover and mutation operations, a new generation is produced. Typically, the process will iterate for many generations until the convergence criterion is satisfied, which may be a maximum number of generations allowed, or when there is no increase in the fitness over a few generations. Besides, a replacement strategy called elitism strategy [22] is incorporated into GA operations in this study to ensure the best members are retained from one generation to the next.

Fig. 18.1 Flowchart of GA-based dynamic model updating procedure



18.3.5 Implementation of the GA to Minimize the Defined Objective Function

Model updating by nature is to make adjustment on the variable structural parameters so that the predicted response by the model best matches the measured counterparts. Therefore, the updating procedure can be ascribed to performing optimization on an error function. In the present study, the available measurements are assumed to be the modal data, i.e. the resonance frequencies. Subsequently, the objective or error function is defined as consisting of the error in the natural frequency, $J(p)$. It should be mentioned that the representation of the structural parameters identified through the updating procedure to the real physical parameters depends upon the overall adequacy of the FE model for the structure. Where a FE model is a much simplified form of the actual structure, the parameters identified should be interpreted as an equivalent measure of the actual physical properties.

It has to be pointed out that the allocation of weights has considerable influence on the updating process and the results by GA. There is no definite method for the selection of weighting factor $W_{\omega i}$. A general rule is that the weighting factors may be chosen as inversely proportional to the variance of the corresponding measurements.

Having specified the objective function, a searching space for the individual variables needs to be defined before the GA can start. In a model updating problem, the variables are the structural parameters to be updated and the searching space for the GA should be defined according to the possible variation range of each parameter. The evaluation of the fitness involves a structural analysis procedure to calculate the eigendata. Figure18.1 shows the flowchart of a GA-based model updating procedure.

In the actual implementation, it is necessary to choose a proper configuration for the GA. Studies by DeJong [24], Grefenstette [25] and Schaffer et al. [26] indicate that for binary-encoding GAs, a population size of 20–30 chromosomes, a crossover rate of 0.6–0.9 and a mutation rate of 0.01–0.02 perform reasonably well in most cases. However, for a real-coding GA, theoretically a lower crossover rate and a higher mutation probability should be considered. In the present study, the configuration of the GA is chosen following a preliminary trial procedure, in which the mutation rate was selected between 0.0 and 1.0, and 20 runs with different seeds were performed. The running average error upon the final GA solution and the variance value are used to gauge the quality of the updating. The chosen configuration details are shown in Table 18.1.

Table 18.1 GA configuration

Encoding scheme	Population size	Termination criterion	Selection method
Real number coding	100	Maximum generations 250	Linear ranking
Crossover	Mutation	Replacement strategy	Fitness scaling
Heuristic crossover	Uniform mutation	Elitist strategy, replacement rate	No
Probability $p_c = 0.80$	Probability $p_m = 0.024$	$\lambda = 0.10$	

18.4 FRF-Based Model Updating for Identifying Damping of Joint

Joints are the major source of damping in complex assembly structures (such as Cross-Beam). In this paper a new iterative method introduced which uses measured FRF data directly to identify damping of joint. The mass and stiffness matrices of joint are $\mathbf{M}_j$ and $\mathbf{K}_j$ that updated by presented GA algorithm. The following proportional damping considered for joint:

$$\mathbf{C}_j = \alpha_j \mathbf{M}_j + \beta_j \mathbf{K}_j \tag{18.3}$$

The damping coefficients of joint α_j and β_j are identified by minimizing the following objective function:

$$\varepsilon(p) = \sum_{i=1}^{N_f} \mathrm{w_{ii}} \log \left\| \mathrm{H_A}(\omega_i) - \mathrm{H_x}(\omega_i) \right\| \tag{18.4}$$

Where N_f is number of measured frequency points, $\mathrm{w_{ii}}$ weighting coefficient, $\mathrm{H}(\omega_i) = \mathrm{H}(\omega_i, p)$, $\mathrm{H_A}(\omega_i)$ and $\mathrm{H_x}(\omega_i)$ are the ith corresponding analytical and experimental FRFs respectively.

The minimization algorithm of objective function is described in Table 18.2. The process described in this table, starts with creating a sample of $n + 1$ random points in the search space in line 2. Here, the create Pop operation must ensure that these samples form a no degenerated simplex. Notice that apart from the creation of the initial simplex, all further steps are deterministic and do not involve random numbers. In each search step, the points in the simplex S are arranged in ascending order according to their corresponding objective values (line 4). Hence, the best solution candidate is $S[0]$ and the worst is $S[n]$. We then compute the center **m** of the **n** best points in line 5 and then reflect the worst solution candidate $S[n]$ through this point in line 6, obtaining the new point **r**. The reflection parameter α is usually set to 1.

In the case that **r** is somewhere in between of the points in the current simplex, i.e., neither better than $S[0]$ nor as worse as $S[n]$, we directly replace $S[n]$ with it. This simple move was already present in the first simplex algorithm defined by Spendley et al. [27]. The contribution of Nelder and Mead [28] was to turn the simplex search into an optimization algorithm by adding new options. These special operators were designed for speeding up the optimization process by deforming the simplex in way that they suggested would better adapt to the objective functions [29].

If **r** is better than the best solution candidate $S[0]$, one of these operators is to expand the simplex further into this promising direction (line 11). we obtain the point **e** with the expansion parameter γ set to 1. We now choose the better one of these two points in order to replace $S[n]$. If **r** was no better than $S[n]$, we the simplex is contracted by creating a point **c** somewhere in between of **r** and **m** in line 17. Then the contraction parameter ρ was set to 1/2.

We substitute $S[n]$ with **c** only if **c** is better than **r**. When everything else fails, we shrink the whole simplex in line 23 by moving all points (except $S[0]$) into the direction of the current optimum $S[0]$. The shrinking parameter σ normally has the value 1/2.

18.5 Experimental Case Study

Cross-Beam laboratory-scale structure is conducted to demonstrate the effectiveness and efficiency of the above mentioned approach in structural model updating. The test assembly is a Cross-Beam laboratory-scale structure, consisting of two steel beams with identical lengths and cross sections ($800, 30 \times 10$ mm), which were connected (asymmetrically) together by an elastic joint, as shown in Fig.18.2. The joint consisted of a nut, bolt, rubber and steel washers, as shown in Fig.18.3. ($\mathrm{U1}_z - \mathrm{U9}_z$) and ($\mathrm{D1}_z - \mathrm{D9}_z$) are measured points of substructures 1 and 2 respectively.

Table 18.2 Minimization algorithm in FRF-sensitivity

```
Input: f: the objective function subject to minimization

Input: $n$: the dimension of the search space

Input: $\alpha$, $\rho$, $\gamma$, $\sigma$: the reflection, the expansion, the contraction, and the shrink coefficient

Data: $S$: the simplex

Data: $\mathbf{m}$: the centroid of the simplex

Data: $\mathbf{r}$: the reflection

Data: $\mathbf{e}$: the expansion

Data: $\mathbf{c}$: the contraction

Data: $i$: a counter variable

Output: $x^*$: the best solution candidate found
begin
    S ⟸ createPop(n + 1)
    while termination Criterion() do
        $S \Longleftarrow \text{sortList}_a(S, f)$
        $\mathbf{m} \leftarrow \frac{1}{n}\sum_{i=0}^{n-1} S[i]$        // Reflection: reflect the worst point over m
        $\mathbf{r} \leftarrow \mathbf{m} + \alpha(\mathbf{m} - S[n])$
        if $f(S[0]) < f(\mathbf{r}) < f(S[0])$ then
            $S[n] \leftarrow \mathbf{r}$
        else
            if $f(\mathbf{r}) < f(S[0])$ then  // Expansion: try to search farther in this direction
                $\mathbf{e} \leftarrow \mathbf{r} + \gamma(\mathbf{r} - \mathbf{m})$
                $f(\mathbf{e}) < f(\mathbf{r})$ then $S[n] \leftarrow \mathbf{e}$
                else $S[n] \leftarrow \mathbf{r}$
            else
                $b \leftarrow$ true
                if $f(\mathbf{r}) \geq f(S[n-1])$ then  // Contraction: a test point between r and m
                    $\mathbf{c} \leftarrow \rho\mathbf{r} + (1 - \rho)\mathbf{m}$
                    if $f(\mathbf{c}) \leq f(\mathbf{r})$ then
                        $S[n] \leftarrow \mathbf{c}$
                        $b \leftarrow$ false
                if $b$ then      // Shrink towards the best solution candidate $S[0]$
                    for $i \leftarrow \mathbf{n}$ down to 1 do
                        $S[i] \leftarrow S[0] + \sigma(S[i] - S[0])$
    return $S[0]$
end
```

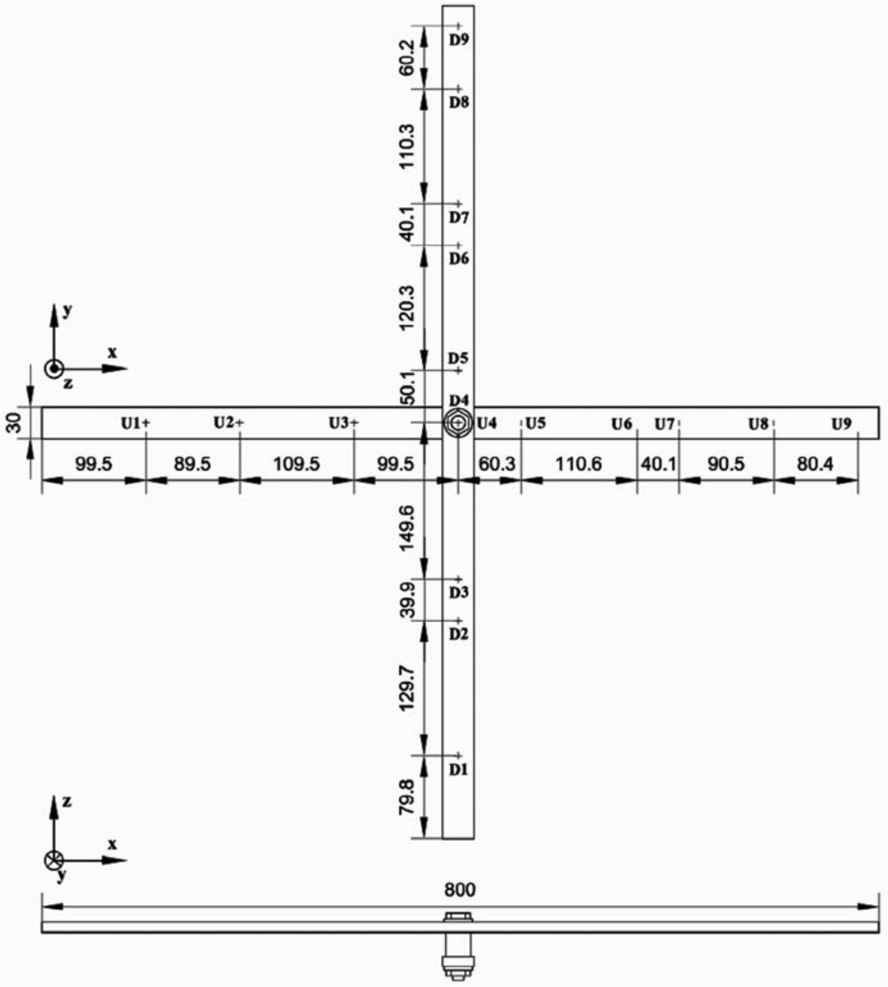


Fig. 18.2 The plot of real assembly

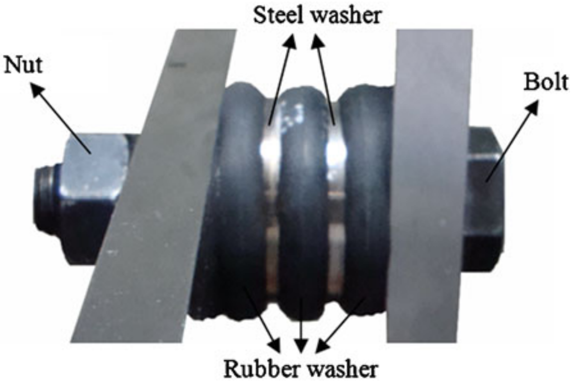


Fig. 18.3 The real joint composition

A *B&K 4809* shaker powered by a *B&K 2706* amplifier was used to random excitation of the system. The applied force was measured by a force transducer connected to a charge amplifier (*B&K 8200 + 2647a*) and the response was measured by a *DJB A/120/V* accelerometer as shown in Fig.18.4. The accelerometer and force signals were collected at nine locations on each substructure, with totally 18 points on the assembly, as shown in Fig.18.2. The structure was hanged from one point by

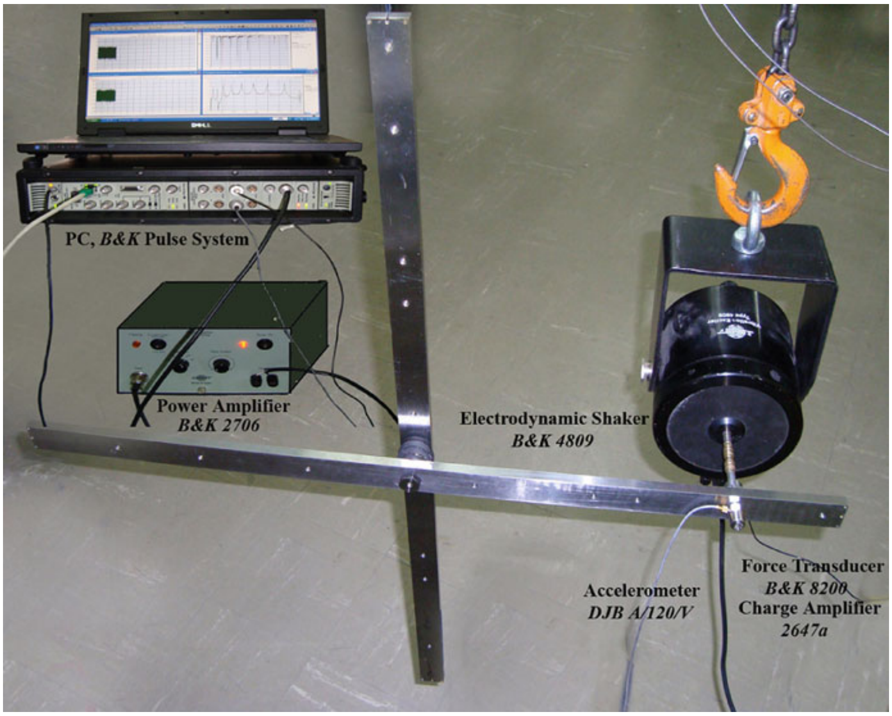


Fig. 18.4 Experimental test setup for FRF measurement

a thin string resembling the free-free condition as much as possible. In order to combine the obtained FRFs from the experiments and the ones from the updated FE model, the obtained accelerants from measurements converted to receptances by multiplying them to $(-1/\omega^2)$. Using the setup shown in Fig.18.4, the FRFs associated only with the TDOFs were obtained. An example of a measured FRF is shown in Fig.18.5.

18.6 Updating of Assembly Structure Based on Substructures-Joint Parameters

As mentioned before, model updating of complex assembly structures that consisting of several substructures and real joints, is considerably difficult. So to reach the exact and acceptable results, the updated model of substructures must be made by using measured modal data and presented GA algorithm. Then by considering the characteristics of joint as variables of updating and using updated results of substructures, the updated FE model of assembly structure can be achieved.

In this paper, three dimensional Euler-Bernoulli beam elements and three dimensional parametric Euler-Bernoulli beam elements having two-node/12-DOF in each element are used for FE model of the substructures and joint respectively. Finite element model updating is obtained by using GA approach in 3 steps: In first step the substructures (Upper and Down) are update using presented GA approach by considering module of elasticity (E_s) and density (ρ_s) as update variables. In the second step, the updated model of the assembly structure is constructed by utilizing parametric finite element model of joint and updated results of substructures. So by considering characteristics of joint (module of elasticity E_j, density ρ_j and moment of inertia I_j) as update variables and using presented GA approach, the updated model of assembly structure can be obtained. At the end of these two steps, undamped updated mass and stiffness matrix of substructures, joint and assembly structure are achieved. Figure18.6 shows Comparison of the measured FRF and analytical FRF curves before updating. Figure18.7 shows the performance curve of the GA. Damping matrix of the substructures and assembly structure were extracted from modal parameters, estimated from the measured data. The sample updated results for damped substructures and undamped assembly structure are presented in Figs.18.8 and 18.9. Table 18.3 shows the updated variables before and after updating for assembly structure.

Fig. 18.5 (**a**) Sample receptance and (**b**) coherence ($U9_z/U1_z$)

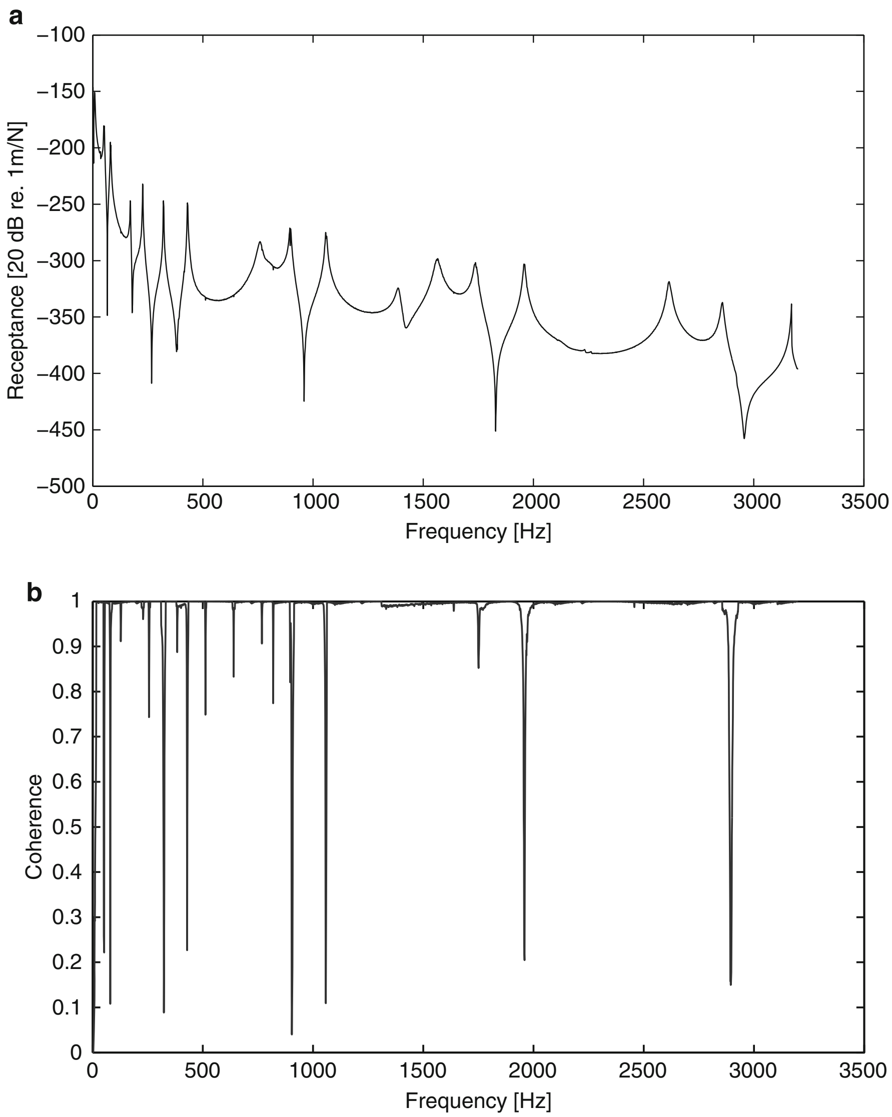
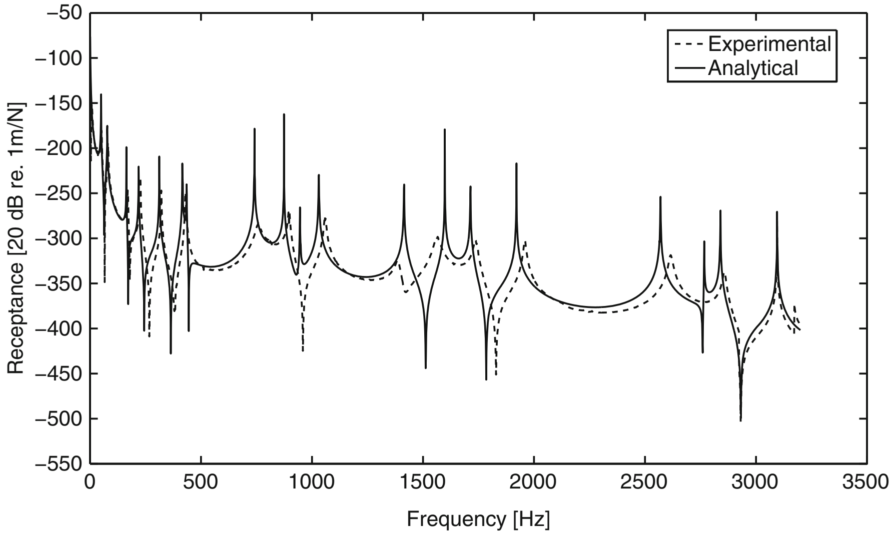


Fig. 18.6 FRFs for assembly structure before updating



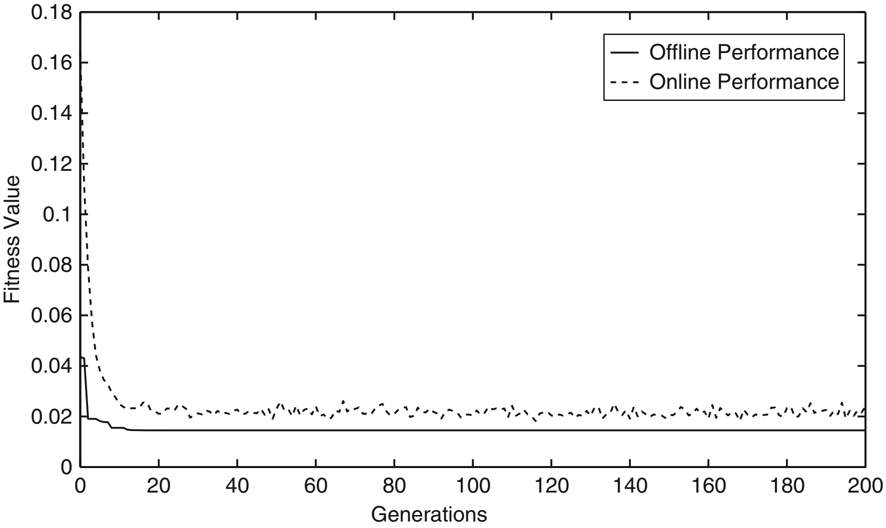


Fig. 18.7 Performance curve of the GA

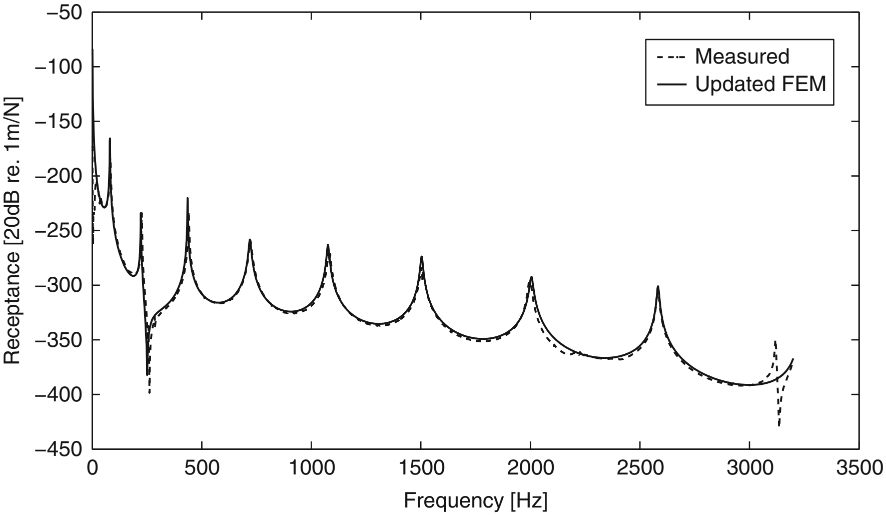


Fig. 18.8 FRFs for upper substructure ($U9_z/U1_z$)

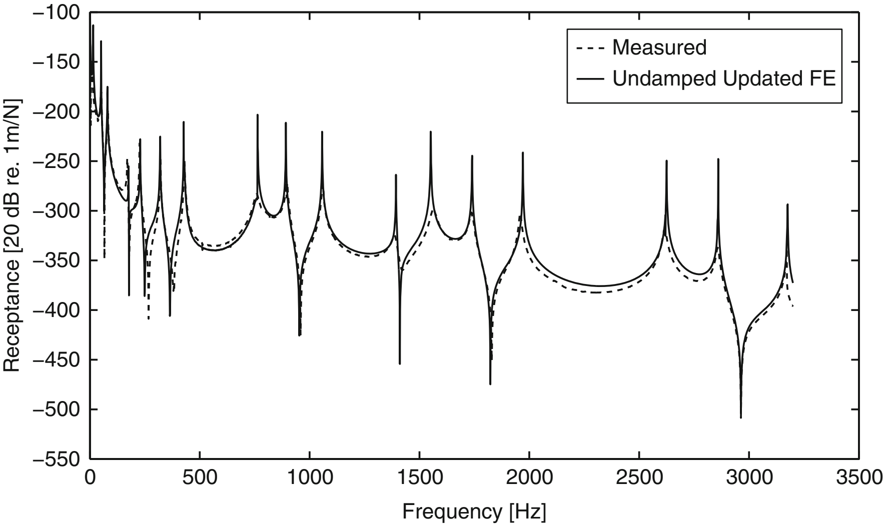


Fig. 18.9 FRFs for assembly structure ($U9_z/U1_z$)

Table 18.3 Variables before and after updating

	Initial value	After updating
E_j(Pa)	1e10	0.6748e10
ρ_j(kg/m^3)	1e3	3.7275e3
I_j(m^4)	1e-8	0.4392e-8

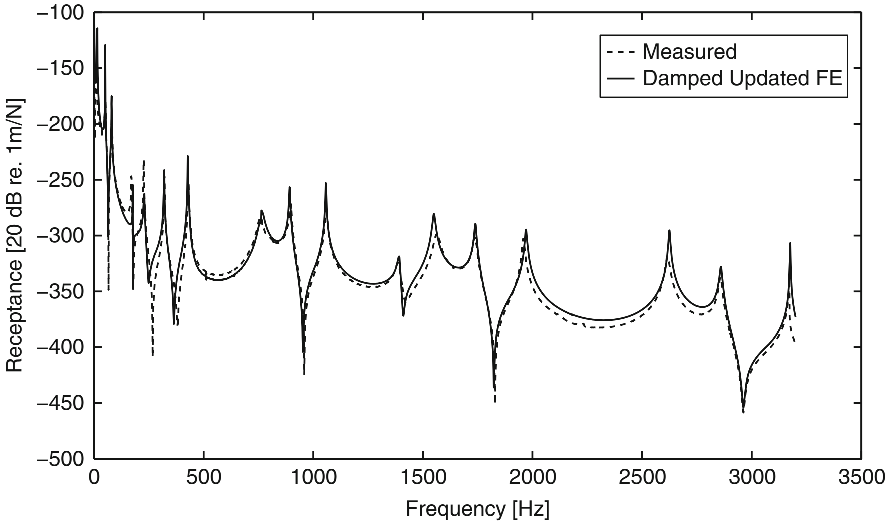


Fig. 18.10 Comparison of FRFs for assembly structure (*damped*)

Table 18.4 Effect of damping of joint on sample resonance

	Amplitude of resonance	
Natural frequency (Hz)	$\alpha_j = 10,\ \beta_j = 4e-7$	$\alpha_j = 5,\ \beta_j = 2e-7$
81	209.9	200.9
1,057	273.6	267.8
1,392	309.3	308.1
1,971	305.9	303.1
3,176	362.1	360.9

It is seen that the FRFs of undamped updated model has a very good match with the measured FRF except at the resonance and antiresonance frequencies. In the following the damped FRFs of assembly updated structure obtained and compared with experimental FRFs, as shown in Fig.18.10. This figure shows that the damped updated model FRF has a very good match with the measured FRF except specific resonance and antiresonance frequencies (80, 1,057, 1,392, 1,971 and 3176 Hz) that are depend on damping of joint.

18.7 Identification of Damping of Joint

Using objective function of Eq. 18.4 and minimization algorithm according to Table 18.2 and changing damping coefficients of joint (α_j and β_j) manually we see that only resonance and antiresonance at specific frequencies of assembly structure FRFs (80, 1,057, 1,392, 1,971 and 3176 Hz) are changed (Table 18.4).

Consequently damping coefficient of joint must be identified using presented FRF-based model updating. Figure18.11 compares experimental and analytical FRF curves near by 80 Hz natural frequency before updating of damping coefficients and Fig.18.12 shows the progress of updating process. Figures18.13 and 18.14 show the convergence of damping coefficients of joint from its initial value respectively. Comparison of experimental and analytical FRF curves near by 80 Hz natural frequency after updating of damping coefficients is shown in Fig.18.15. Table 18.5 shows the damping coefficients of joint before and after updating.

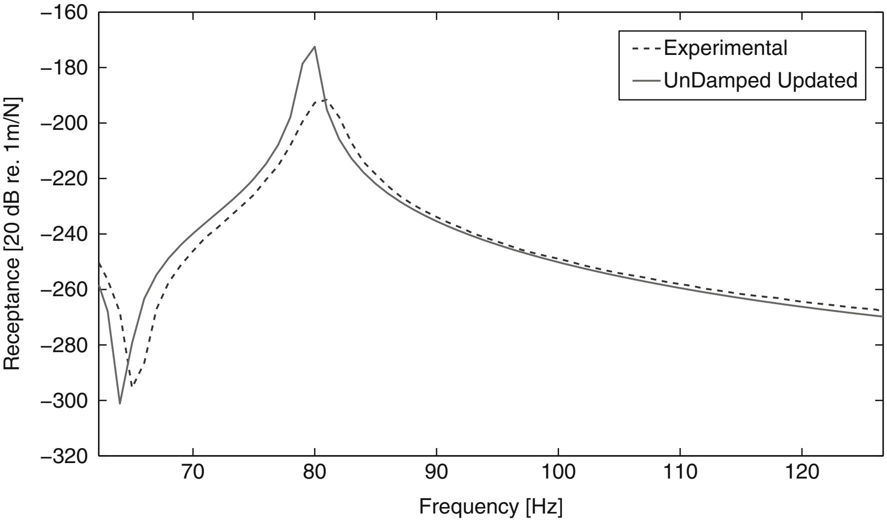


Fig. 18.11 Comparison of FRFs near by 80 Hz, before updating of joint damping

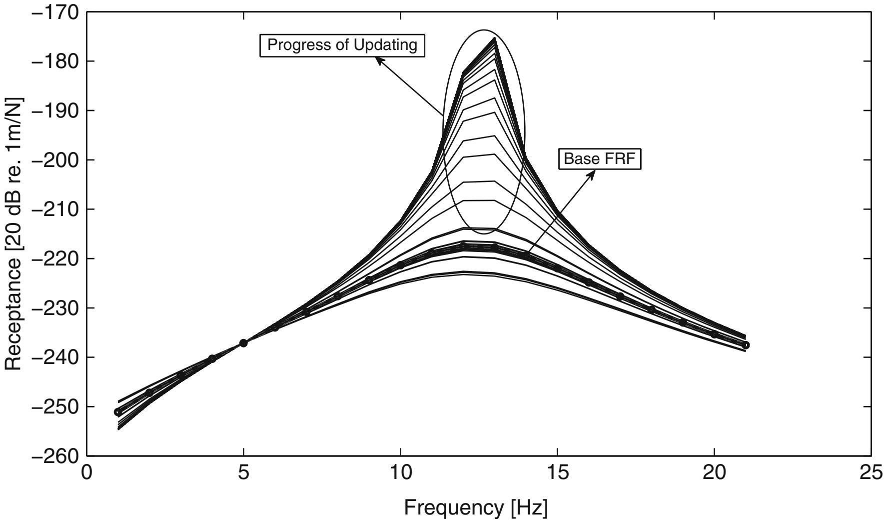


Fig. 18.12 Progress of updating near by 80 Hz

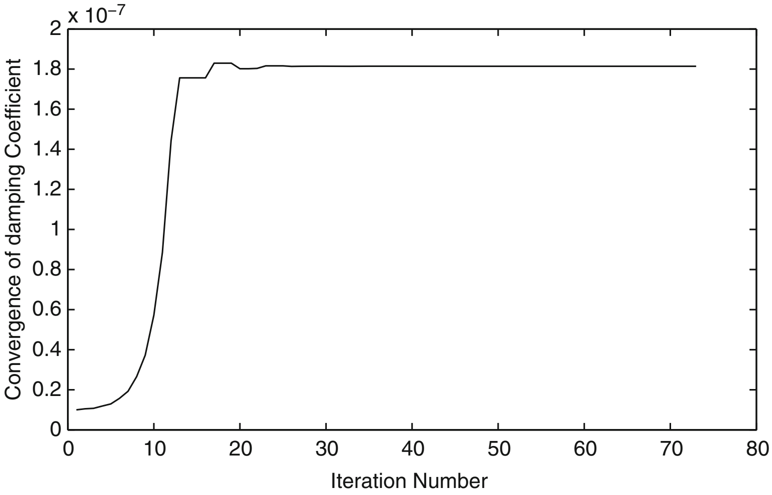


Fig. 18.13 Convergence of damping coefficient of joint (β_j)

Fig. 18.14 Convergence of damping coefficient of joint (α_j)

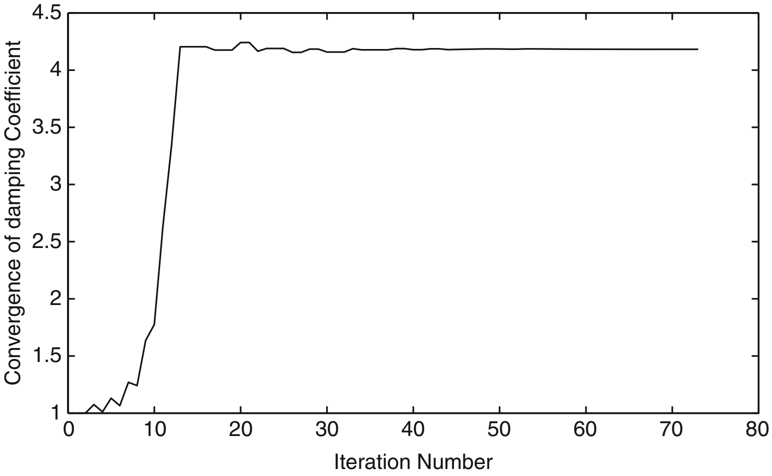


Fig. 18.15 Comparison of FRFs near by 80 Hz, after updating of joint damping

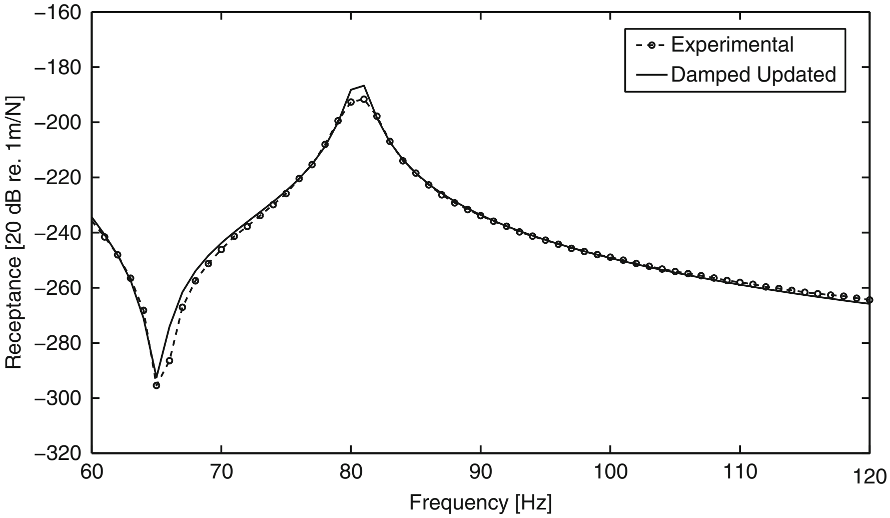


Table 18.5 Damping coefficient of joint

	Initial value	After updating
α_j	1	4.1821
β_j	1e-8	18.1435e-8

It is seen that the damped updated model FRF has a very good match with the measured FRF. It is observed from Fig.18.15 and Table 18.5 that after damping identification the predicted dynamic characteristics are closer to the measured characteristics of the modified structure even at resonance and antiresonance frequencies. In the third step using damping of assembly structure and updated damping of joint, FRFs of assembly updated structure obtained by using GA and compared with experimental FRFs (Fig.18.16).

18.8 Conclusions

In this paper, GA, Eigen-Sensitivity and FRF-Based approach, is presented for the dynamic model updating of complex assembly structures consisting of substructures and real joints based on the measured dynamic properties. The modal testing of real joints, such as real bolted joints, is almost impossible. So, model updating of the mentioned assembly structures in a

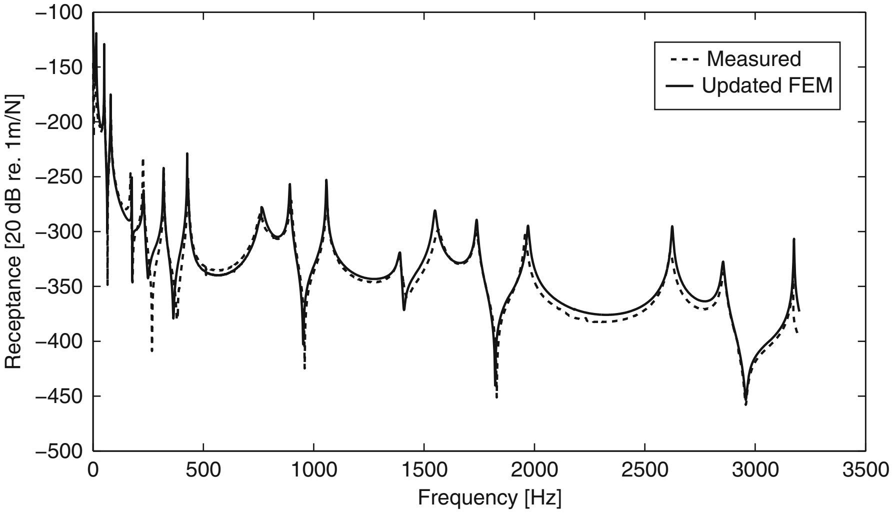


Fig. 18.16 Comparison of FRFs for damped assembly structure

wide frequency range, will be difficult. To tackle this problem, at the outset of this paper, the Eigen-sensitivity based approach using GA algorithm is incorporated to perform the updated model of substructures. In the following by using the obtained updated results of substructures and considering properties of joint as updated variables, the updated model of assembly structure is created by using GA algorithm in a wide frequency range. In the process of updating, the damping matrix of substructures and assembly structure were extracted from modal parameters, which were estimated by measured data. The obtained results show that there are good agreements between the FRFs obtained from the updated model (substructure and assembly structure) and the ones from the measured data, except for specific resonance and antiresonance frequencies that depend on the damping of joint. Damping of joint is updated by using FRF-based model updating approach and minimizing related objective function. In the final step, the updated FRFs of assembly structure is achieved by using updated damping of joint and damping of substructures, and the assembly structure extracted from the measured data. The updated FRFs show a full fitness with the measured ones, which demonstrates the superior performance of the proposed approach in comparison with other methods.

References

1. Ewins DJ (2000) Modal testing: theory, practice and application. Research Studies, Hertfordshire
2. Maia NMM, Silva JMM (1997) Theoretical and experimental modal analysis. Research Studies, Taunton
3. Craig RR (2000) Coupling of substructures for dynamic analyses: an overview. In: AIAA/ASME/ASCE/AHS/ASC 41st structures, structural dynamics, and materials conference, Atlanta, pp 3–17
4. Klerk D, Rixen DJ, Voormeeren SN (2008) General framework for dynamic substructuring: history, review, and classification of techniques. AIAA J 46:1169–81
5. Perera R, Ruiz A (2008) A multistage FE updating procedure for damage identification in large-scale structures based on multiobjective evolutionary optimization. Mech Syst Signal Process 22:970–991
6. Ko JM, Sun ZG, Ni YQ (2002) Multi-stage identification scheme for detecting damage in cable-stayed Kap Shui Mun Bridge. Eng Struct 24:857–868
7. Xia Y, Lin RM (2004) Improvement on the iterated IRS method for structural eigensolutions. J Sound Vib 270:713–727
8. Xia Y, Lin RM (2004) A new iterative order reduction (IOR) method for eigensolutions of large structures. Int J Numer Meth Eng 59:153–172
9. Mottershead JE, Friswell MI (1998) Editorial. Mech Syst Signal Process 12(1):1–6
10. Chen JC, Kou CP, Garba JA (1983) Direct structural parameter identification by modal test results. In: 24th AIAA/ASME/ASCE/AHS structural dynamics and materials conference, pp 44–49
11. Friswell MI, Mottershead JE (1995) Finite element model updating in structural dynamics. Kluwer, Dordrecht
12. Mottershead JE, Mares C (2000) Selection and updating of parameters for an aluminium space-frame model. Mech Syst Signal Process 14 (6):923–944
13. Jones KW (2000) Finite element model updating using antiresonance frequencies. MS thesis, AFIT/GA/ENY/00-M08. Air Force Institute of Technology (AU), Wright-Patterson AFB, Ohio
14. Mottershead JE, Friswell MI (1993) Model updating in structural dynamics: a survey. J Sound Vib 167(2):347–375

15. Doebling SW, Farrar CR, Prime MB, Shevitz DW (1996) Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review, LA-13070-MS, Los Alamos National Laboratory, Los Alamos
16. Friswell MI, Penny JET, Garvey SD (1998) A combined genetic and eigensensitivity algorithm for the location of damage in structures. Comput Struct 69:547–556
17. Chow J-H, Ghaboussi J (2001) Genetic algorithm in structural damage identification. Comput Struct 79:1335–1353
18. Levin RI, Lieven NAJ (1998) Dynamic finite element model updating using simulated annealing and genetic algorithms. Mech Syst Signal Process 12(1):91–120
19. David Z, Keng Y (1999) Evolutionary approach for model refinement. Mech Syst Signal Process 13(4):609–625
20. Wright AH (1991) Genetic algorithms for real parameter optimization. In: Rawlins GJE (ed) Foundation of genetic algorithms. Morgan Kaufmann, San Mateo
21. Adhikari S, Woodhouse J (2000) Identification of damping, part 1, viscous damping. J Sound Vib 243:43–61
22. Goldberg DE (1989) Genetic algorithms in search, optimization and machine learning. Addison-Wesley Longman, Reading
23. Christopher RH, Jefery AJ, Michael GK (1995) A genetic algorithm for function optimization: a matlab implementation. NCSU-IE Technical Report 95-09, North Carolina State University, Raleigh
24. DeJong KA (1975) Analysis of the behavior of a class of genetic adaptive systems. Ph.D. Dissertation, University of Michigan, Ann Arbor
25. Grefenstette JJ (1986) Optimization of control parameters for genetic algorithms. IEEE Trans Syst Man Cybern SMC-16(1):122–128
26. Schafer JD, Caruana RA, Eshelman LJ, Das R (1989) A study of control parameters affecting online performance of genetic algorithms. In: Proceedings of the 3rd international conference on genetic algorithms, San Francisco, pp 51–60
27. Spendley W, Hext GR, Himsworth FR (1962) Sequential application of simplex designs in optimization and evolutionary operation. Technometrics 4(4):441–461
28. Nelder JA, Mead RA (1965) A simplex method for function minimization. Comput J 7:308–313
29. Lewis RM, Torczon VJ, Trosset MW (2000) Direct search methods: then and now. Technical report NASA/CR-2000-210125 and ICASE Report No. 2000-26, Institute for computer applications in science and engineering, NASA Langley Research Center, Hampton (Operated by Universities Space Research Association), National Aeronautics and Space Administration Langley Research Center, Hampton, Virginia 23681-2199