Conference Proceedings of the Society for Experimental Mechanics Series

Robert Barthorpe · Roland Platz · Israel Lopez
Babak Moaveni · Costas Papadimitriou *Editors*

Model Validation and Uncertainty Quantification, Volume 3

Proceedings of the 35th IMAC, A Conference and Exposition on Structural Dynamics 2017

Springer

Conference Proceedings of the Society for Experimental Mechanics Series

Series Editor

Kristin B. Zimmerman, Ph.D.
Society for Experimental Mechanics, Inc.,
Bethel, CT, USA

More information about this series at http://www.springer.com/series/8922

Robert Barthorpe • Roland Platz • Israel Lopez • Babak Moaveni
Costas Papadimitriou
Editors

Model Validation and Uncertainty Quantification, Volume 3

Proceedings of the 35th IMAC, A Conference and Exposition on Structural Dynamics 2017

Springer

Editors
Robert Barthorpe
University of Sheffield
Sheffield, UK

Israel Lopez
Lawrence Livermore National Laboratory
Livermore, CA, USA

Costas Papadimitriou
University of Thessaly
Thessaly, Greece

Roland Platz
Fraunhofer Institute Darmstadt
Darmstadt, Germany

Babak Moaveni
Tufts University
Medford, MA, USA

ISSN 2191-5644 ISSN 2191-5652 (electronic)
Conference Proceedings of the Society for Experimental Mechanics Series
ISBN 978-3-319-54857-9 ISBN 978-3-319-54858-6 (eBook)
DOI 10.1007/978-3-319-54858-6

Library of Congress Control Number: 2017943322

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Model Validation and Uncertainty Quantification represents one of ten volumes of technical papers presented at the 35th IMAC, A Conference and Exposition on Structural Dynamics, organized by the Society for Experimental Mechanics and held in Garden Grove, California, on January 30–February 2, 2017. The full proceedings also include the following volumes: Nonlinear Dynamics; Dynamics of Civil Structures; Dynamics of Coupled Structures; Sensors and Instrumentation; Special Topics in Structural Dynamics; Structural Health Monitoring & Damage Detection; Rotating Machinery, Hybrid Test Methods, Vibro-Acoustics and Laser Vibrometry; Shock & Vibration, Aircraft/Aerospace, and Energy Harvesting; and Topics in Modal Analysis & Testing.

Each collection presents early findings from experimental and computational investigations on an important area within structural dynamics. Model Validation and Uncertainty Quantification (MVUQ) is one of these areas.

Modeling and simulation are routinely implemented to predict the behavior of complex dynamical systems. These tools powerfully unite theoretical foundations, numerical models, and experimental data which include associated uncertainties and errors. The field of MVUQ research entails the development of methods and metrics to test model prediction accuracy and robustness while considering all relevant sources of uncertainties and errors through systematic comparisons against experimental observations.

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

Sheffield, UK — Robert Barthorpe
Darmstadt, Germany — Roland Platz
Livermore, CA, USA — Israel Lopez
Medford, MA, USA — Babak Moaveni
Thessaly, Greece — Costas Papadimitriou

Contents

Chapter 1
Lateral Vibration Attenuation of a Beam with Piezo-Elastic Supports Subject to Varying Axial Tensile and Compressive Loads

Benedict Götz, Roland Platz, and Tobias Melz

Abstract In this paper, vibration attenuation of a beam with circular cross-section by resonantly shunted piezo-elastic supports is experimentally investigated for varying axial tensile and compressive beam loads. Varying axial beam loads manipulate the effective lateral bending stiffness and, thus, lead to a detuning of the beams resonance frequencies. Furthermore, varying axial loads affect the general electromechanical coupling coefficient of transducer and beam, an important modal quantity for shunt-damping. The beam's first mode resonance frequency and coupling coefficient are analyzed for varying axial loads. The values of the resonance frequency and the coupling coefficient are obtained from a transducer impedance measurement. Finally, frequency transfer functions of the beam with one piezo-elastic support either shunted to a RL-shunt or to a RL-shunt with negative capacitance, the RLC-shunt, are compared for varying axial loads. It is shown that the beam vibration attenuation with the RLC-shunt is less influenced by varying axial beam loads.

Keywords Piezo-elastic support • Resonant shunt • Vibration attenuation • Beam • Axial load

1.1 Introduction

Structural vibration may occur in mechanical systems leading to fatigue, reduced durability or undesirable noise. In this context, resonant shunting of piezoelectric transducers can be an appropriate measure for attenuating vibrations. Shunt-damping in general has been subject to research for several decades [1] and resulted in many diverse shunt concepts such as mono- or multi-modal resonant shunts [2], shunts with negative capacitances [3] or switched shunts [4]. Shunting a piezoelectric transducer with resistor and inductance, the RL-shunt, a tuned electrical oscillation circuit with the inherent capacitance of the transducer is created. This electromechanical system acts similar to a mechanical vibration absorber. RL-shunts are easy to implement and no stability limits or switching laws have to be taken into account. However, the achieved vibration attenuation significantly depends on the tuning of the shunt parameters and the amount of the general electromechanical coupling coefficient of transducer and structure [5]. By adding a negative capacitance, the RLC-shunt achieves higher vibration attenuation but stability issues have to be considered.

In mechanical and civil engineering, truss structures bear and withstand constant and variable loads that may lead to vibrations. Truss structures comprise truss members such as beams that are connected to each other via the relatively stiff truss supports. On the one hand, truss structures show global vibration modes with lateral moving or rotating truss supports. On the other hand, local modes exist that are dominated by the lateral vibration behavior of each beam. Additional varying quasi static loading may result in axial tensile and compressive loads of the beams. This affects the resulting lateral bending stiffness of the beams leading to a permanent change in the resonance frequencies. For vibration attenuation with resonant shunt-damping, the detuning of the resonance frequencies affects the achievable vibration attenuation capability. In truss structures, piezoelectric shunt-damping has been investigated in [6–8]. Axial piezoelectric stack transducers are integrated within one strut of the truss, resulting in compression and elongation of the transducer in normal axial direction of the strut [6, 8]. In [7], a beam support with integrated piezoelectric washers alongside the beam that are strained in shear under dynamic loading has been investigated and vibration attenuation of bending modes in a truss substructure was achieved. Due to a planar washer design, only one bending direction of the beams could be influenced. Nevertheless, research

B. Götz (✉) • T. Melz
System Reliability, Adaptive Structures, and Machine Acoustics SAM, Technische Universität Darmstadt, Magdalenenstraße 4, 64289 Darmstadt, Germany
e-mail: goetz@sam.tu-darmstadt.de

R. Platz
Fraunhofer Institute for Structural Durability and System Reliability LBF, Bartningstraße 47, 64289 Darmstadt, Germany

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_1

investigating the effect of varying quasi static axial beam loads on the vibration attenuation of beams with resonant shunt-damping are not known to the authors.

In this paper, a new concept of a piezo-elastic support [9] for vibration attenuation of truss structures comprising beams with circular cross-section by shunted transducers is investigated for varying axial tensile and compressive beam loads. Within the piezo-elastic support, deflections in both lateral directions of a beam are transferred into an axial deformation of the transducers that are arranged perpendicular to one free end of the beam. By that design, local and global truss modes can attenuated without manipulating the beam's surface. In the following only one beam with two piezo-elastic supports is investigated for varying axial tensile and compressive loads. One piezo-elastic support mechanically excites the beam laterally using white noise excitation while the second support is either shunted to a RL- or to a RLC-shunt. First, the experimental beam's first mode resonance frequency and its general electromechanical coupling coefficient are calculated from impedance measurements for different axial tensile and compressive loads. Changes in both, the resonance frequency and the coupling coefficient may influence the vibration attenuation. Considering the electrical transducer impedance in the frequency domain is a known method for identifying the transducer capacitance, structural resonance frequencies or the coupling coefficient and has several advantages. E.g., all important quantities are obtained from one measurement at one time, no structural transfer function has to be measured and the estimation of the resonance frequency as well as the coupling coefficient is less influenced by the used frequency resolution [10]. Second, the experimental frequency transfer functions in lateral direction of the excited beam in case one transducer is shunted to a RL-shunt and in case one transducer is shunted to a RLC-shunt are compared for uncertain axial tensile and compressive beam loads.

1.2 System Description

The investigated system is a beam under axial load made of aluminum alloy EN AW-7075 with length $l_\mathrm{b} = 400\,\mathrm{mm}$ and circular solid cross-section of radius $r_\mathrm{b} = 5\,\mathrm{mm}$, Fig. 1.1. The circumferential lateral stiffness is homogeneous and has no preferred direction of lateral deflection, so the beam may vibrate in any plane lateral to the longitudinal x-axis. The beam is supported by two piezo-elastic supports A and B at location $x = 0$ and location $x = l_\mathrm{b}$. Elastic membrane-like spring elements made of spring steel 1.1248 in both supports A and B at location $x = 0$ and $x = l_\mathrm{b}$ bear axial and lateral forces at the beam's ends in x-, y- and z-direction, and allow rotation φ_y and φ_z in any plane perpendicular to the x-axis, see Fig. 1.2. In Fig. 1.1, the membrane-like spring elements for both supports A and B are represented by axial stiffness $k_{x,\mathrm{A}} = k_{x,\mathrm{B}} = 1.53 \times 10^4\,\mathrm{N/mm}$, not shown in the figure, lateral stiffness $k_{y,\mathrm{A}} = k_{z,\mathrm{A}} = k_{y,\mathrm{B}} = k_{z,\mathrm{B}} = k_\mathrm{l} = 9.98 \times 10^4\,\mathrm{N/mm}$ in y- and z-direction and rotational stiffness $k_{\varphi_y,\mathrm{A}} = k_{\varphi_z,\mathrm{A}} = k_{\varphi_y,\mathrm{B}} = k_{\varphi_z,\mathrm{B}} = k_\mathrm{r} = 2.69 \times 10^5\,\mathrm{N\,mm/rad}$ around the y- and z-axes. All spring element stiffness values are obtained from a finite element simulation and they are not experimentally verified yet. In each piezo-elastic support A and B at $x = -l_\mathrm{ext}$ and $x = l_\mathrm{b} + l_\mathrm{ext}$, two piezoelectric stack transducers P_1 and P_2 as well as P_3 and P_4 are arranged in the support housing at an angle of 90° to each other orthogonal to the beam's x-axis, Fig. 1.1b. All transducers are mechanically prestressed by a stack of disc springs with stiffness $k_\mathrm{pre} = 2.6 \times 10^3\,\mathrm{N/mm}$. The transducers are connected to the beam via a relatively stiff axial extension made of hardened steel 1.2312 with length $l_\mathrm{ext} = 6.75\,\mathrm{mm}$ and edge length $t_\mathrm{ext} = 12\,\mathrm{mm}$. With that, lateral beam deflections in y- and z-direction due to vibration excitation are transformed into the stack transducer's axial deformation. Each piezoelectric transducer $\mathrm{P_1}$ to $\mathrm{P_4}$ is a PI P-885.51 stack transducers with the capacitance $C_\mathrm{p} = 1.65\,\mu\mathrm{F}$ at constant mechanical stress, internal series resistance $R_\mathrm{p} = 7\,\Omega$ and the mechanical stiffness $k_\mathrm{p} = 50 \times 10^3\,\mathrm{N/mm}$ with short circuited electrodes, defined as the ratio of the transducer's block force and the maximum free stroke. The input current to the transducer is $I(t)$ and the potential difference at the transducer electrodes is the voltage $U(t)$.

For vibration attenuation, a RL- and a RL-shunt with negative capacitance C, the RLC-shunt, are taken into account, Fig. 1.1c. In Fig. 1.1c, the RL-shunt is obtained by neglecting the negative capacitance. By adjusting the inductance L and the damping resistance R, the transducer shunted to a RL-shunt attenuates vibrations similar to a mechanical vibration absorber [1]. For vibration attenuation with shunt-damping, the generalized electromechanical coupling coefficient K_{33} is an important modal quantity indicating the vibration attenuation capability. The higher the value of K_{33} is, the higher the achievable vibration attenuation with a RL-shunt becomes. By adding a negative capacitance $C \approx -1.02 \cdot C_\mathrm{p}$ in series to R and L, the coupling coefficient K_{33} increases effectively and, hence, the vibration attenuation capability is significantly higher compared to the simple RL-shunt [5]. Apart from that, RLC-shunts may destabilize the beam vibration, therefore, stability limits for the value of C have to be considered. Basically, the achievable vibration attenuation capability with a RLC-shunt is almost independent of the coupling coefficient K_{33}. However, the smaller the coupling coefficient is, the close the value of the negative capacitance has to be chosen to the stability limit [5]. In the experiment, the inductance L is implemented by the use of a gyrator circuit and the negative capacitance by the use of negative admittance converter [2]. Both circuit designs are not discussed in detail in this paper.

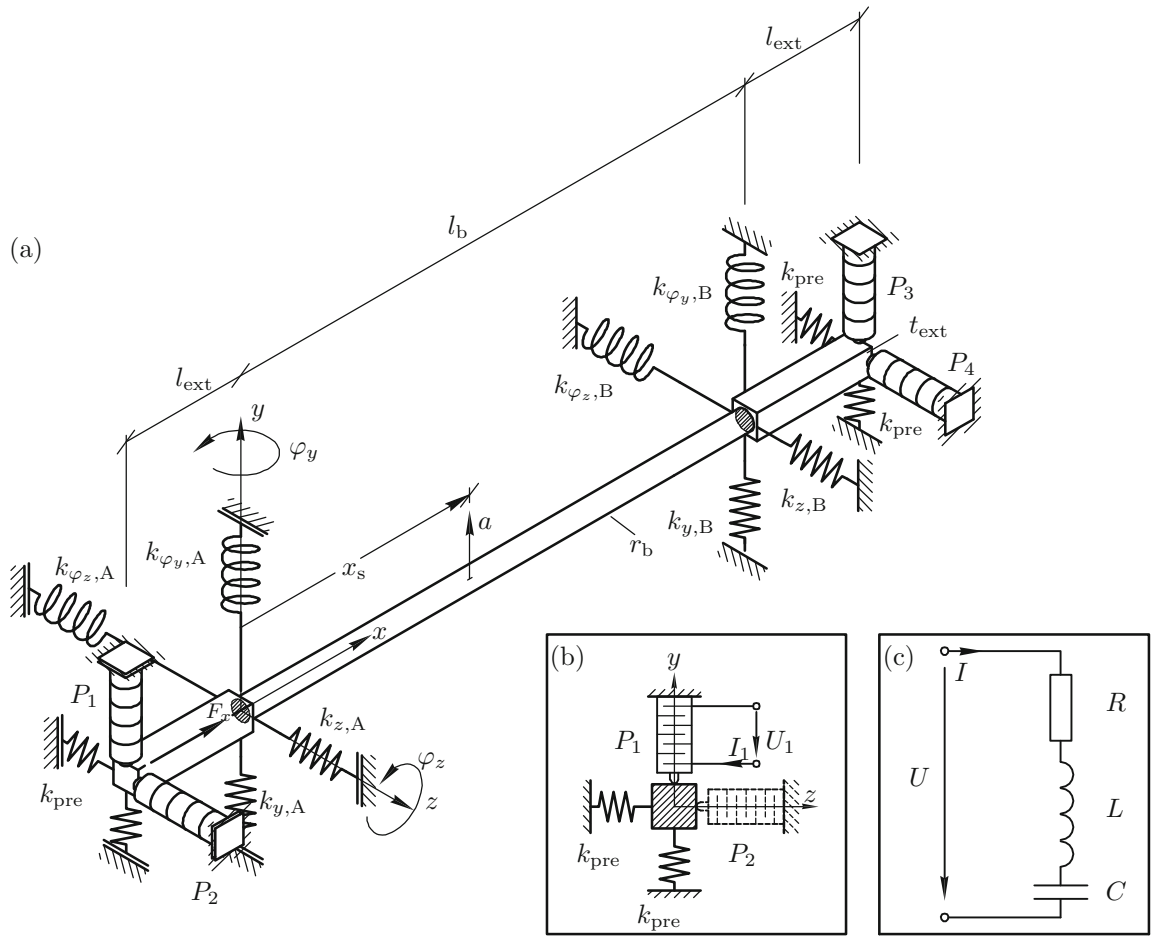


Fig. 1.1 Beam system, (**a**) beam with piezo-elastic supports A and B, (**b**) arrangement of piezoelectric transducers, (**c**) shunt circuit

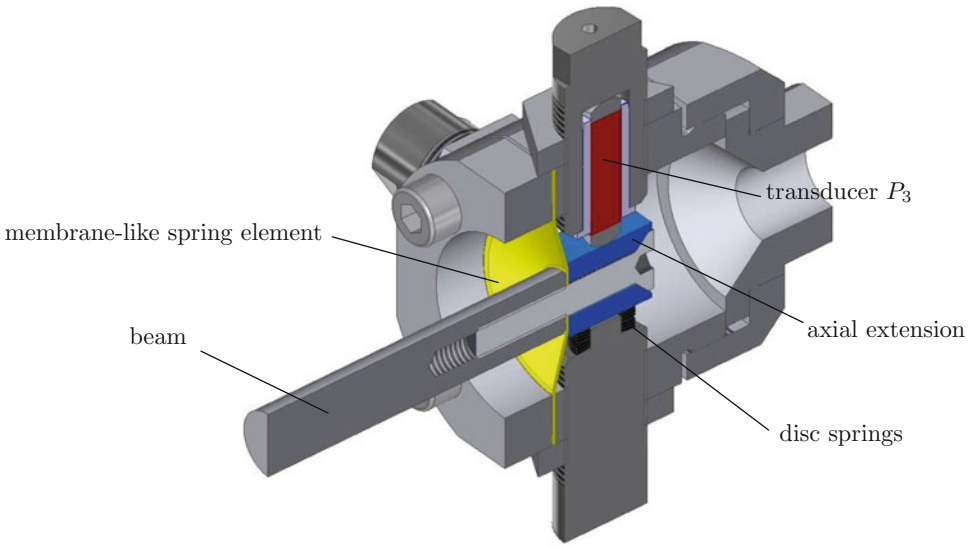


Fig. 1.2 Sectional view of piezo-elastic support [11]

For all experimental vibration attenuation investigations, the piezoelectric transducer P_1 excites the beam in y-direction via a controlled voltage signal $U_1(t)$. The transducer P_3, either shunted to the RL-shunt or the RLC-shunt, attenuates the vibration acceleration $a(t)$ at the sensor location x_s while the transducers P_2 and P_4 are operated with short circuited (sc) electrodes. Varying axial tensile and compressive loads $-1000\,\mathrm{N} \leq F_x \geq 1500\,\mathrm{N}$, with compressive loads in positive x-direction, are applied to the beam at $x = 0$ via a spindle-type lifting gear and measured by a force sensor.

1.3 The Beam's First Mode Eigenfrequency and Coupling Coefficient for Varying Axial Loads

The vibration attenuation capability bears on the piezoelectric transducers P_3 that is either shunted to a tuned RL- or a tuned RLC-shunt. Varying axial beam loads may change the beam's first mode resonance angular frequency ω_{sc} with short circuited (sc) transducer electrodes and the beam's first mode general coupling coefficient K_{33} of transducer P_3 in y-direction. Changes in both, ω_{sc} and K_{33} may influence the vibration attenuation with shunted transducers. To investigate the influence of varying loads F_x on ω_{sc} and K_{33}, the mathematical receptance model of the transducer P_3 is derived in frequency domain. The values of ω_{sc} and K_{33} for axial loads $F_x = [-1000, -500, 0, 500, 1000, 1500]\,\mathrm{N}$ are extracted from a least squares fit of the receptance model to the experimental data in frequency domain. As already shown by Kozlowski et al. [10], obtaining ω_{sc} and K_{33} from a curve fitting of the transducer receptance model results in a smaller error since the calculation of both parameters is less influenced by the used frequency resolution in the measurement.

1.3.1 Transducer Receptance Model

Figure 1.3 shows the electrical network representation of the piezoelectric transducer P_3 connected to the beam. The transducer P_3 is described by a gyrator-like two-port transducer network with its electrical capacitance C_p, a internal series resistance R_p and its transducer constant Y [12]. The vibration behavior of the beam's first mode with short circuited transducer electrodes is modelled by the modal mass m, the modal stiffness k and the assumed hysteretic damping with loss factor η resulting in the complex stiffness $k' = k\,(1 - i\,\eta)$.

The complex network receptance seen from the terminals 1 and 2 in Fig. 1.3 is obtained by

$$\alpha(\omega) = \frac{1}{i\,\omega\,Z(\omega)} \tag{1.1}$$

with excitation frequency ω. The impedance

$$Z(\omega) = \frac{U_3(\omega)}{I_3(\omega)} = \frac{Z_p(\omega)\,Z_1(\omega)}{Z_p(\omega) + Z_1(\omega)} \tag{1.2}$$

results from the parallel connection of the structural impedance $Z_1(\omega)$ of the first mode seen from the terminals 3 and 4 the transducer impedance

$$Z_p(\omega) = R_p - \frac{i}{\omega\,C_p}. \tag{1.3}$$

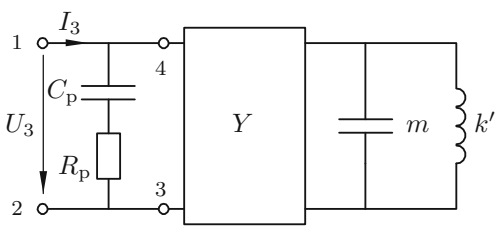


Fig. 1.3 Electrical network model of the piezoelectric transducers and the beam's first mode

The structural impedance of the first mode in y-direction

$$Z_1(\omega) = Y^2 \left(\frac{1}{Z_{\mathrm{m}}(\omega)} + \frac{1}{Z_{\mathrm{k}'}(\omega)} \right) = Y^2 \left(i\,\omega\,m + \frac{k\,(1+i\,\eta)}{i\,\omega} \right) \tag{1.4}$$

seen from the terminals 3 and 4 of the gyrator is the result of the parallel impedance of the modal mass $Z_{\mathrm{m}}(\omega) = \dfrac{1}{i\,\omega\,m}$ and the complex stiffness $Z_{\mathrm{k}'}(\omega) = \dfrac{i\,\omega}{k'}$. From (1.1) and (1.2), the receptance becomes

$$\alpha(\omega) = \frac{C_p}{\left(1 + \omega\,C_p\,R_{\mathrm{p}}\,i\right)} + \frac{1}{Y^2\,m\,\left(\omega_{\mathrm{sc}}^2 - \omega^2 + i\,\eta\,\omega_{\mathrm{sc}}^2\right)} \tag{1.5}$$

with the angular eigenfrequency $\omega_{\mathrm{sc}} = \sqrt{\dfrac{k}{m}}$. Furthermore, the term $\dfrac{1}{Y^2\,m}$ in (1.5) is replaced by $K_{33}\,C_{\mathrm{p}}\,\omega_{\mathrm{sc}}$, as suggested in [10], leading to the final expression of the transducer receptance

$$\alpha(\omega) = \frac{C_p}{\left(1 + \omega\,C_p\,R_{\mathrm{p}}\,i\right)} + K_{33}\,C_{\mathrm{p}}\,\frac{\omega_{\mathrm{sc}}}{\left(\omega_{\mathrm{sc}}^2 - \omega^2 + i\,\eta\,\omega_{\mathrm{sc}}^2\right)}. \tag{1.6}$$

1.3.2 Transducer Receptance Model Fit

In the model fit process, the parameters C_{p}, R_{p}, K_{33}, ω_{sc} and η in (1.6) are varied to solve the least squares curve fitting problem

$$\min_{C_{\mathrm{p}},\,R_{\mathrm{p}},\,K_{33},\,\omega_{\mathrm{sc}},\,\eta} ||\alpha(C_{\mathrm{p}},\,R_{\mathrm{p}},\,K_{33},\,\omega_{\mathrm{sc}},\,\eta,\,\omega) - \alpha_{\mathrm{exp}}(\omega)||_2^2 \tag{1.7}$$

where $\alpha_{\mathrm{exp}}(\omega)$ is the experimental data of the transducer receptance. Therefore, the *lsqnonlin* algorithm in MATLAB is used. Figure 1.4 shows the amplitude and phase response $|\alpha(\omega)|$ and $\arg\alpha(\omega)$ of the experimental data and the calculated receptance after the curve fitting for the axial load $F_x = 0\,\mathrm{N}$. Both, the model and the experimental data show a very good agreement.

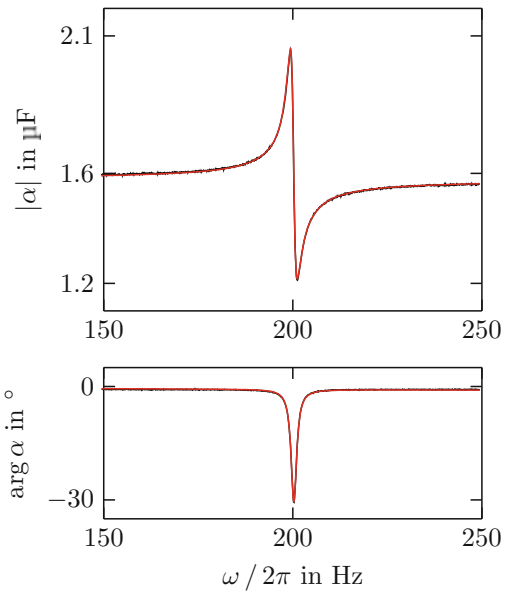


Fig. 1.4 Calculated transducer receptance (*red solid line*) with fitted parameters and experimental data (*black solid line*) for $F_x = 0\,\mathrm{N}$

1.3.3 Experimental Results of the First Eigenfrequency and Coupling Coefficient for Varying Axial Loads

As introduced before, a change of the resonance frequency ω_{sc} due to axial loads F_x will lead to a detuned vibration behavior of the RL- and RLC-shunt and, hence, the vibration attenuation capability will decrease. Apart from that, tensile and compressive axial loads may affect the electromechanical coupling coefficient K_{33} of transducer P_3. An increased coupling coefficient K_{33} will also increase the vibration attenuation capability with a RL-shunt, while a decreased K_{33} will also decrease the vibration attenuation potential. Furthermore, vibration stability issues with a RLC-shunt due to a changing coupling coefficient K_{33} are not investigated in this paper.

Figure 1.5a,b show the beam's first mode resonance frequency ω_{sc} and the beam's first mode coupling coefficient K_{33} of transducer P_3 extracted from the fitted receptance model (1.6) for axial tensile and compressive loads $F_x = [-1000, -500, 0, 500, 1000, 1500]$ N. In Fig. 1.5a, the resonance frequency ω_{sc} increases for an axial tensile load and decreases for an axial compressive load significantly. Compared to the axially unloaded beam with $F_x = 0$ N, the resonance frequency ω_{sc} increases by 6% and decreases by 10% for the extremes of the applied axial loads at $F_x = -1$ kN and $F_x = 1.5$ kN.

In contrast to the behavior of the frequency ω_{sc}, in Fig. 1.5b, the coupling coefficient K_{33} decreases for an axial tensile load and increases for an axial compressive load, as also shown by Lesieutre and Davis [13]. Compared to the axially unloaded beam with $F_x = 0$ N, the coupling coefficient K_{33} decreases by 6% and increases by 7% for the extremes of the applied axial loads at $F_x = -1$ kN and $F_x = 1.5$ kN.

Considering the absolute changes of the resonance frequency ω_{sc} and the coupling coefficient K_{33} in Fig. 1.5a,b axial tensile and compressive loads will decrease the vibration attenuation capability with the RL- and the RLC-shunt due to a detuning. For the RL-shunt with axial compressive loads, the increase in coupling coefficient and the associated theoretically increase of vibration attenuation capability will not be able to compensate the effects of a detuned ω_{sc} on the vibration attenuation capability, as it will be shown in the next section.

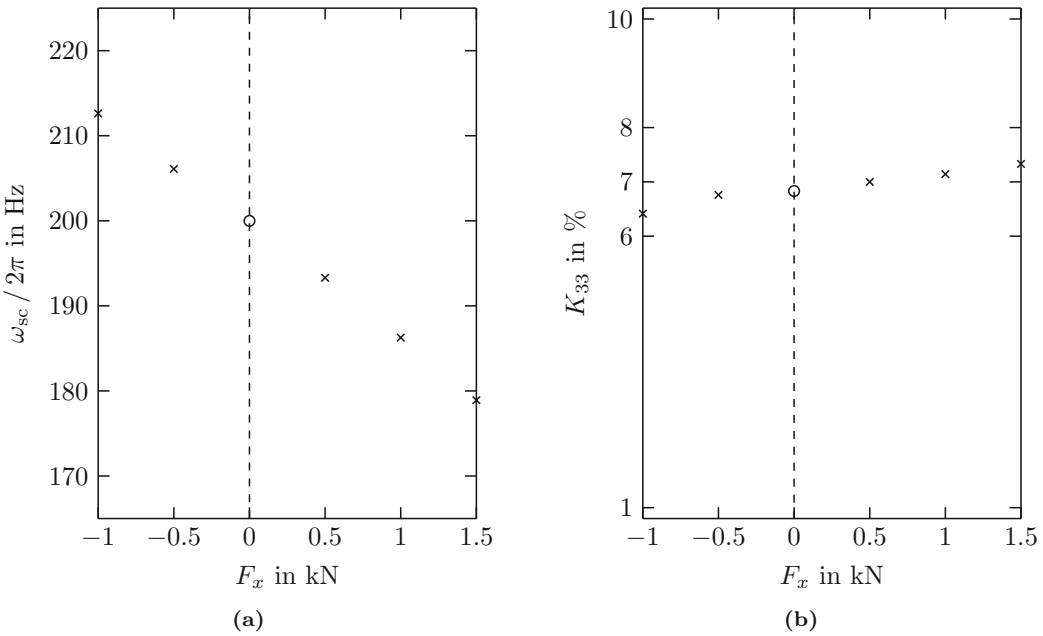


Fig. 1.5 Beam's first mode (**a**) resonance frequency ω_{sc} and (**b**) electromechanical coupling coefficient K_{33} of transducer P_3, for varying tensile and compressive axial loads F_x (*cross symbol*) and the axially unloaded beam $F_x = 0$ N (*open circle*)

1.4 Experimental Vibration Attenuation with RL- and RLC-Shunt for Varying Axial Loads

The vibration attenuation of the beam in case of the transducer P_3 is shunted to a RL-shunt and in case of P_3 is shunted to a RLC-shunt is compared for axial tensile and compressive loads $F_x = [-1000, -500, 0, 500, 1000]\,\mathrm{N}$, Fig. 1.6. As a measure for the vibration attenuation of the RL- and RLC-shunt, the frequency transfer function

$$H(\omega) = \frac{a(\omega)}{U_1(\omega)} \tag{1.8}$$

of the voltage excitation $U_1(\omega)$ of transducer P_1 to the beam vibration acceleration $a(\omega)$ is considered, Fig. 1.1. Additionally for the axially unloaded beam with $F_x = 0\,\mathrm{N}$, the vibration attenuation potential of the RL- and RLC-shunt is obtained from the comparison of $H(\omega)$ when the P_3 electrodes are short circuited and when shunted.

Figure 1.6a shows the amplitude and phase response $|H(\omega)|$ and $\arg H(\omega)$ when transducer P_3 is shunted to a RL-shunt with $R = 48.4\,\Omega$ and $L = 402\,\mathrm{mH}$. The values for R and L were tuned experimentally to give the highest reduction in amplitude in the considered frequency range. For the axially unloaded beam at $F_x = 0\,\mathrm{N}$, the vibration attenuation with RL-shunt is 15 dB. For varying axial tensile and compressive loads F_x, a significant decline in the vibration attenuation potential can be observed in the amplitude response. The decline in vibration attenuation is slightly higher for tensile loads. Furthermore, the detuning of ω_{sc} due to axial tensile and compressive loads is observed in the phase response by a shift of the 90° crossing frequency.

Figure 1.6a shows the amplitude and phase response $|H(\omega)|$ and $\arg H(\omega)$ when transducer P_3 is shunted to a RLC-shunt with $R = 4.8\,\Omega$, $L = 29.6\,\mathrm{mH}$ and $C = -1.69\,\mu\mathrm{F}$. The values for R, L and C were tuned experimentally to give the highest and stable reduction in amplitude in the considered frequency range. For the unloaded beam, the vibration attenuation with RCL-shunt is 26 dB and, as expected, is significantly higher than with the RL-shunt. For varying axial tensile and compressive loads F_x, the observed decline in the vibration attenuation capability is smaller compared to the RL-shunt. The decline in vibration attenuation, again, is slightly higher for tensile loads.

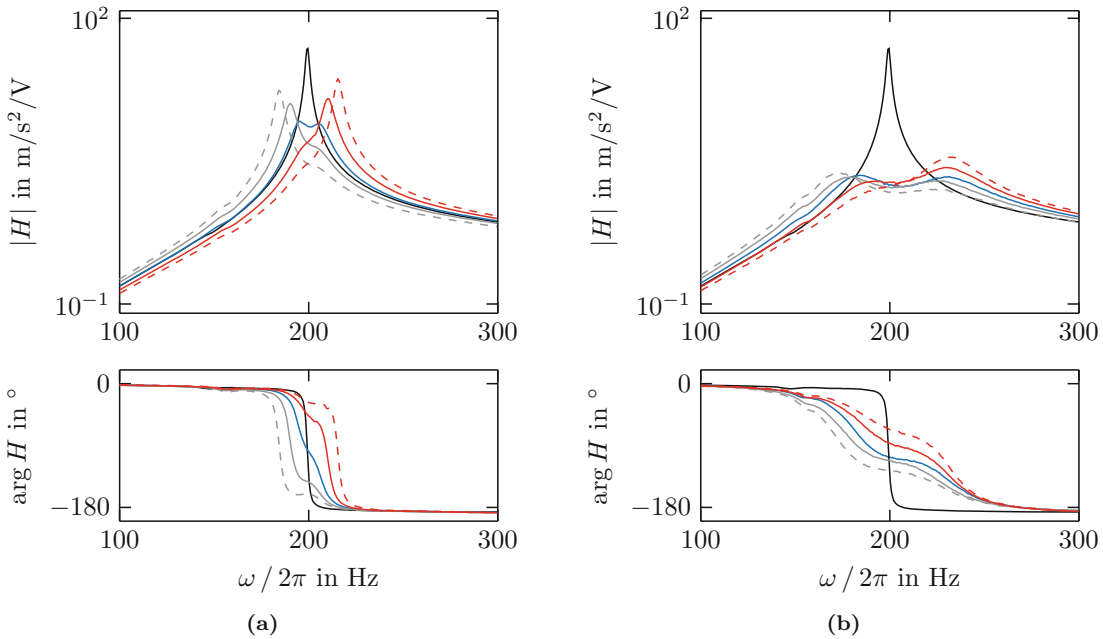


Fig. 1.6 Amplitude and phase response of $H(\omega)$, short circuited (*black solid line*), with (**a**) RL-shunt and (**b**) RLC-shunt, for the unloaded beam (*blue solid line*) with $F_x = 0\,\mathrm{N}$ and varying axial loads F_x : $+500\,\mathrm{N}$ (*gray solid line*), $+1000\,\mathrm{N}$ (*gray dashed line*), $-500\,\mathrm{N}$ (*red solid line*), $-1000\,\mathrm{N}$ (*red dashed line*)

To conclude, vibration attenuation with a RLC-shunt is less sensitive to varying axial tensile and compressive loads. However, using a negative capacitance may lead to stability issues, but they were not observed in the performed experiments and are not part of this paper.

1.5 Conclusion

Vibration attenuation of a beam with a circular cross-section by piezo-elastic supports with one transducer shunted to RL- or RLC-shunt subject to varying axial tensile and compressive loads is experimentally investigated. When no shunt is connected to the transducer, the first beam's mode resonance frequency significantly increases and decrease in consequence of axial tensile and compressive loads. Compared to the unloaded beam, the relative changes of the electromechanical coupling coefficient are in the same order of magnitude as for the resonance frequency, but, the absolute change of the coupling coefficient has no significant effect an the vibration attenuation with RL- and RLC-shunt. The resonance frequency detuning due to axial tensile and compressive loads results in a declined vibration attenuation when shunting the transducer to a RL- or a RLC-shunt. As observed in the measured frequency transfer functions, the RLC-shunt is less sensitive to uncertain axial loads then the RL-shunt.

Acknowledgements The authors like to thank the German Research Foundation DFG for funding this research within the SFB 805.

References

1. Hagood, N.W., von Flotow, A.H.: Damping of structural vibrations with piezoelectric materials and passive electrical networks. J. Sound Vib. **146**(2), 243–268 (1991)
2. Moheimani, S.O.R., Fleming, A.J.: Piezoelectric Transducers for Vibration Control and Damping. Springer, London (2006)
3. Beck, B.S.: Negative capacitance shunting of piezoelectric patches for vibration control of continuous systems. PhD thesis, Georgia Institute of Technology (2012)
4. Niederberger, D.: Smart damping materials using shunt control. PhD thesis, Swiss Federal Institute of Technology Zürich (2005)
5. Neubauer, M., Oleskiewicz, R., Popp, K., Krzyzynski, T.: Optimization of damping and absorbing performance of shunted piezo elements utilizing negative capacitance. J. Sound Vib. **298**, 84–107 (2006)
6. Hagood, N.W., Crawley, E.F.: Experimental investigation of passive enhancement of damping for space structures. J. Guid. Control. Dyn. **14**(6), 1100–1109 (1991)
7. Hagood, N.W., Aldrich, J.B., von Flotow, A.H.: Design of passive piezoelectric damping for space structures. NASA Contractor Report 4625 (1994)
8. Preumont, A., de Marneffe, B., Deraemaeker, A., Bossens, F.: The damping of a truss structure with a piezoelectric transducer. Comput. Struct. **86**, 227–239 (2008)
9. Götz, B., Schaeffner, M., Platz, R., Melz, T.: Lateral vibration attenuation of a beam with circular cross-section by a support with integrated piezoelectric transducers shunted to negative capacitances. Smart Mater. Struct. **25**(095045), 10 (2016)
10. Kozlowski, M.A., Cole, D.G., Clark, R.L.: A comprehensive study of the RL series resonant shunted piezoelectric: a feedback controls perspective. J. Vib. Acoust. **133**, 011012-1–011012-10 (2011)
11. Enss, G.C., Gehb, C.M., Götz, B., Melz, T., Ondoua, S., Platz, R., Schäffner, M.: Device for bearing design elements in lightweight structures (Festkörperlager) (2016)
12. Lenk, A., Ballas, R.G., Werthschützky, R., Pfeifer, G.: Electromechanical Systems in Microtechnology and Mechatronics. Number 978-3-540-89320-2. Springer, Berlin (2011)
13. Lesieutre, G.A., Davis, C.L.: Can a coupling coefficient of a piezoelectric device be higher than those of its active material? J. Intell. Mater. Syst. Struct. **8**, 859–867 (1997)

Chapter 2
Correlation of Non-contact Full-Field Dynamic Strain Measurements with Finite Element Predictions

Ibrahim A. Sever, Martyn Maguire, and Jose V. Garcia

Abstract It is highly desirable to have the capability to measure strain maps on components directly and in a full-field fashion that addresses shortcomings of conventional approaches. In this paper, use of a 3D laser measurement system is explored for direct and full-field dynamic strain measurements on compressor and turbine rotor blades. More importantly, the results obtained are numerically correlated to corresponding FE predictions in a systematic manner. The ability to measure strain maps on real engine hardware is demonstrated not only for low frequency fundamental modes, but also for challenging high frequency modes. Correlation results show a high degree of agreement between measured and predicted strains, demonstrating the maturity of the technology and the validity of the method of integration used here. The measurements are repeated for a number of different loading amplitudes to assess the variations in strain fields. Although the application of 3D laser systems to measurements of full-field strain were explored in previous studies, to the best knowledge of authors, full-field numerical correlation of full-field strain on a wide range of real, complex components to this extent is presented here for the first time.

Keywords Model validation • Full-field strain • 3D SLDV • Correlation • Non-contact

2.1 Introduction

The ability to measure dynamic strain on components subjected to high vibratory stresses is very important as these measurements then directly feed into all important endurance/life calculations. Historically this requirement has been fulfilled in two main ways. The first and most widely used approach is the application of strain gauges. Although a direct measurement and still a very popular practice; there are a number of shortcomings. Firstly, they are intrusive as they have to be bonded to the component. Typically only a few of these can be used which do not provide a representative spatial coverage nor are they enough to evaluate changing strain patterns due to load variations. Their nontrivial footprint means that they can only provide average strain under the area they cover. The second approach, albeit less common, is to validate a finite element model of a given component through direct measurements of displacement, velocity or acceleration and then to use that improved model for predictions of strain and stresses. Although effectively used, particularly in case of full-field measurement systems such as Scanning Laser Doppler Vibrometers (SLDV), good level of correlation with these measurements does not always translate to a good correlation in strain. Moreover, when it comes to components showing complex phenomena or those made from novel materials, the confidence in original FE models is often low or such FE models may not even exist; rendering strain predictions obtained this way even less reliable.

Non-contact and full-field measurements of 3D vibration responses have been explored via a number of different technologies over the years. Earlier systems using double-pulse Electronic Speckle Pattern Interferometry (ESPI) exploited different combinations of viewing and observation directions [1, 2] vibration measurements. Systems with 3-observation and 1-illumination directions as well as 3-illumination and 1-observation directions were explored however recovery of full-field dynamic strain for industrial applications and high frequency complex mode shapes were not reported. A creative way in

I.A. Sever (✉)
Rolls-Royce Plc, SinA-33, PO Box 31, Derby, DE24 8BJ, UK
e-mail: ibrahim.sever@rolls-royce.com

M. Maguire
Rolls-Royce MTOC, Kiefernstraβe 1, D-15827 Blankenfelde-Mahlow, Germany

J.V. Garcia
GP4-4, PO Box 3, Bristol, BS34 7QE

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_2

which a 1-D SLDV system is used in combination with a short-focus lens to recover 3D vibration information was given in [3] however obtaining full-field coverage this way is simply not practical. A detailed review of Digital Image Correlation (DIC) techniques applied to vibration measurements explored suitability of DIC technology compared with more conventional methods and SLDV based systems [4], however, similar to double pulse ESPI, the use of DIC based systems outside quasi static regime and for high frequency complex modes has been limited. Recently 3D SLDV based measurement systems have gained popularity due to their practicality for complete dynamic deformation field (i.e. 3D) measurements. A through study of strain measurements with a 3D SLDV system is given in [5]. Although comparison with FE and conventional strain gauge results are presented, these are done at the locations of strain gauges only, rather than in full-field sense, such as in the form of strain MAC and strain CoMAC.

The pursuit of more direct, high density and high accuracy measurements in this study is motivated in particular by their potential to provide better model validation opportunities. Valid models (i.e. models that are demonstrated to be adequate representation of real life behaviour) provide unique opportunities as they can enable simulation of behaviour for a wide range of parameter ranges and constraints that may not be practicable or cost effective to do through testing. Given the criticality of the use of these models, such as in estimating the stress and strain fields and ultimately the structural integrity of aero-engine components, ensuring that they are valid to an acceptable degree is essential. This is something that has been mainly done via measurements of displacement mode shapes as these are the easiest to measure. The inferred conclusion from such measurements is that when the displacement shapes are shown to match with a sufficient degree of correlation, the resulting stress and strain distributions will follow the same trend. However the more direct and the more detailed the measurements of parameters of interest are, such as strain and stress distributions, the higher will be the confidence one can have in simulation models these data are used to asses and, if necessary, to correct.

2.2 Measurement Campaign

Measurements of full-field strain on aero-engine blades using a 3D SLDV system were reported in an earlier publication [6]. Although in this paper the focus is on the correlation of results, particularly those of full-field strain, with the FE model predictions, it is worthwhile reviewing the basic principles of measurement system as well as hardware tested and the setup used.

2.2.1 Measurement System

A 3D Scanning LDV system is used in acquisition of displacement and strain measurements. A picture of the measurement system in use is given in Fig. 2.1a. The principle behind the operation of an LDV transducer can be explained simply as follows [7]: light produced by a laser source is split into two beams of the same amplitude by a beam splitter, one directed to a fixed reference and the other to the vibrating target. Following the same path back, the beams are combined by the same splitter and send to a photodetector. Since the light from the target is optically mixed with an equally coherent reference beam and heterodyned on the photodetector surface, the resolution of the sign of the vibration velocity is achieved by pre-shifting the reference beam's frequency by a known amount. The signal received by the photodetector is then frequency demodulated by a suitable Doppler processor and the vibration velocity of the target is worked out. The 3D SLDV system used during this measurement campaign was a PSV 500 3D. All simple out-of-plane 1D measurements were made using a Polytec PSV 400 HS. The 3D system consists of three independent SLDV heads as shown in Fig. 2.1a. Fundamental mode of operation for 1D and the 3D systems are identical in that each laser transducer captures the vibration response on the structure along its own line of sight. The fact that there are three such observation directions in the 3D case is being exploited to recover the complete vibration response in three orthogonal directions. This requires that all three SLDVs are coordinated and that the measurement surface is precisely aligned to a degree where laser beams from all SLDVs are coincident to within an acceptable tolerance. Alignment requirements are stricter for strain estimations than they are for the displacement measurements.

Fig. 2.1 (**a**) 3D SLDV measurement system in use, and (**b**) Intermediate pressure compressor blade (*left*) and intermediate pressure turbine blade (*right*)

2.2.2 Test Hardware

The measurement campaign is carried out on a number of aero-engine components including intermediate pressure compressor and turbine blades and a full-size fan blade (not shown here). Some of these components are shown in Fig. 2.1b. In case of compressor and turbine blades, tests are repeated for a number of different excitation levels. These components feature a number of different challenges in terms of clamping conditions, frequency range they cover and the complexity of modes of vibration they poses. As such they should provide appropriate coverage for demonstration of the capability being presented. Strain as well as displacement measurements are carried out for all components however in this paper the correlation of strain measurements is carried out for the compressor blade alone.

2.2.3 Test Environment and Setup

In order to eliminate the adverse effects that the environment might have on the measurements, the testing was carried out in a state-of-the-art vibration test facility. Vibration isolation is achieved through the use of large air-sprung bed plates, and the temperature is maintained at a suitable level. Thick, well insulated test cell walls ensure that there is no interference from external sources. The measurement process was largely automated which meant that once alignment was achieved no user intervention was required. As alignment was based on the engine coordinate system it was repeatable.

Components tested here were fixed at their roots with appropriate clamping mechanisms, mimicking similar boundary conditions to those present in engine. Various excitation techniques were used depending on the size of the component. Turbine blades were excited via an acoustic horn pressure unit and a bespoke piezoelectric resonator. Most tests on these components were performed using the pressure unit as it proved more effective. For larger components (i.e. compressor and fan blades) acoustic speakers were found to be more appropriate where suitable speakers were selected proportional to size of the components being excited.

Measurement grids on blade surfaces to be scanned were carefully optimised in a separate test planning process to maximise observability of the modes on interest (e.g. maximise ability to distinguish them without any ambiguity) using nominal FE models present. This ensured that the measurement grid was defined in the engine coordinate system. This is a major advantage as this grid is then transferred to the measurement system and measurement volume is calibrated in a way that corresponds to the FE environment, making the alignment and correlation of FE and test points much easier. Much denser grids were used in strain measurements, compared with the ones used in displacement mode shapes. Also a much more accurate laser alignment process had to be used in strain measurement case which in return made the strain measurements a longer campaign.

2.3 FE Model and Test Planning

Test planning for the measurement campaign was carried out using the nominal FE models for the blades tested. Provided that there are no fundamentally significant deviations, this is acceptable as the character of modes and the frequency ranges derived from the nominal models provide appropriate guidelines for defining the overall boundaries of the test campaign. However, it is well known that due to manufacturing tolerances, the physical parts show variations from their design intent. As these tolerances are often defined by manufacturability and performance constraints, their impact on structural dynamics may be non-trivial. The impact of these variations on the overall correlation study will be explored in future publications. For the sake of introducing the correlation methodology, all FE models used in this study will be those derived from the nominal geometry but with appropriate boundary conditions to reflect the test configuration.

Planning of the test campaign consists of defining the measurement grid and assessing the suitability of this grid in capturing vibration modes of interest. FE model of the compressor blade, measurement grids for displacement and strain mode shapes, and, auto correlation matrix of the displacement modes captured by identified measurement grid are all given in Fig. 2.2a, b and c, respectively.

In the case of displacement mode shapes, the effective independence method [8] is used for down selection of measurement points. Given that an optical measurement system is used, the candidate nodes to choose from are the ones that lie on the surface that can be measured (see Fig. 2.2a), rather the whole FE model. Suitability of this grid is confirmed by Auto Modal Assurance Criterion (autoMAC) plot given in Fig. 2.2c. Here the predicted modes of the nominal FE model

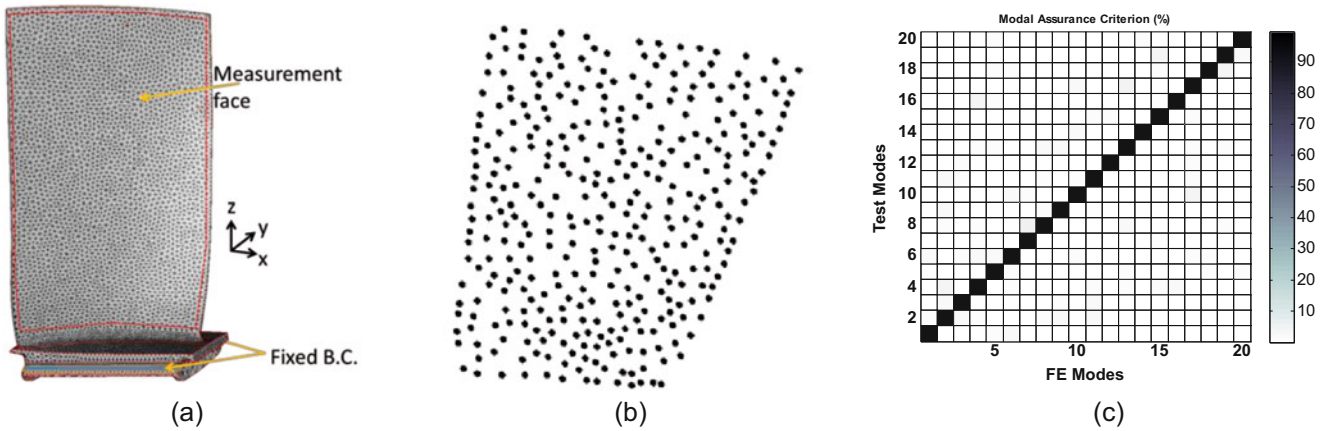


Fig. 2.2 (**a**) Compressor blade FE model, (**b**) Measurement grid used displacement mode shapes, (**c**) resulting auto-MAC plot

sampled at measurement nodes are compared to themselves. Cross-mode correlation amplitudes as evident from trivial off-diagonal terms (all below 10%) suggest that all measured modes should be uniquely identifiable.

Measurement grid for the strain mode shapes is significantly denser than the one for displacement mode shapes. This is required to capture the local strain variations faithfully. Note that by this stage the modes are already identified through correlation of displacement mode shapes. As such autoMAC check performed for displacements is not necessary to repeat for strain.

2.4 Correlation of Mode Shapes

Measurement grid given above was identified for the displacement mode shapes using the full deformation field. In other words all X, Y and Z displacement DOFs were used at each measurement grid as the same DOFs would be captured from the tests. Measurement of all DOFs at each point is a requirement for strain measurements but it is not essential for displacement mode shapes. In fact 1-D SLDVs, measuring a projection of total deformation field in the line of sight, have been used for decades. Nevertheless, availability of all DOFs brings significant advantages in the form of increased independent information which even in the case of displacement mode shapes can make a big difference.

Figure 2.3 shows a particular mode measured on the intermediate pressure turbine blade by 1-D and 3-D SLDV systems, together with the predicted FE mode shape where FE and the 3-D measured mode shape are almost identical. Although the 1-D measurement appears to be very different, a direct comparison is inappropriate. 1-D SLDV measures a projection of overall response in the viewing direction whereas distributions shown for the FE and the 3-D SLDV are for the resultant displacements from all DOFs computed and measured. A correct correlation in the case of 1-D SLDV measurements would be with FE predictions projected in a similar way to reflect the operation of 1-D SLDV system.

Having said that, the fact remains that the 3-D SLDV provides a lot more information (three times as much) about the deformation field, which in return allows better identification of measured mode shapes. This is demonstrated in Fig. 2.4. Here there are two correlation scenarios shown where mode shapes captured by 1-D and 3-D SLDV systems are correlated in the form of Modal Assurance Criterion (MAC) with their corresponding FE predictions in Fig. 2.4a and b, respectively. Significant off-diagonal values in 1-D SLDV case which lead to difficulties in identifying mode shapes unambiguously are greatly reduced in the 3-D case where the identification of the modes is now straight forward.

Displacement mode shape measurement campaign performed on the compressor blade is summarised in Fig. 2.5. As evident from the sub-set of measured and predicted mode shapes given in Fig. 2.5a, not only the global behaviour but also the local variations are extremely closely matched. MAC matrix given in Fig. 2.5b shows a remarkable degree of correlation between measurements and the predictions with MAC values at 95% and above, and, with all off-diagonal values below 10%. It is worth noting the extraordinary similarity between autoMAC plot generated in Fig. 2.2c and the MAC plot given in Fig. 2.5b. As such the FE model is demonstrated to be a good representation of the measurement hardware from mode shapes point of view.

This level of accuracy in the resultant correlation also demonstrates that tests are carried out as planned and that the alignment of FE model and the test model is performed adequately. The latter is a critical factor, particularly for mode shapes

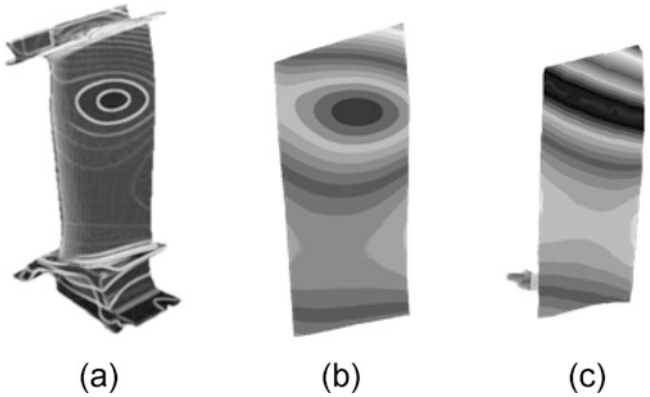


Fig. 2.3 (**a**) Predicted FE mode shape, (**b**) 3-D SLDV measurement, and (**c**) 1-D SLDV measurement

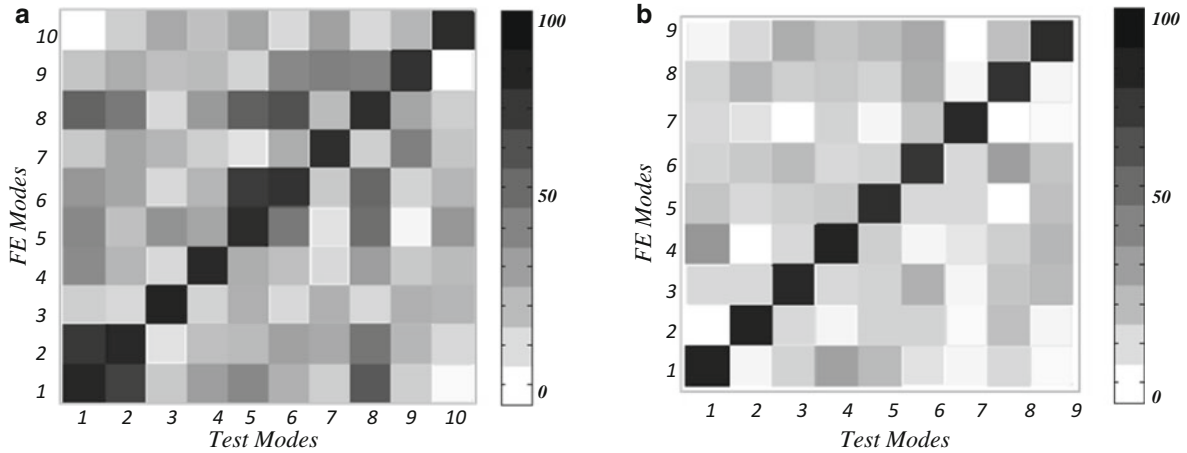


Fig. 2.4 Mode shape correlation (%) for the turbine blade using (**a**) 1-D SLDV, and (**b**) 3-D SLDV measurements

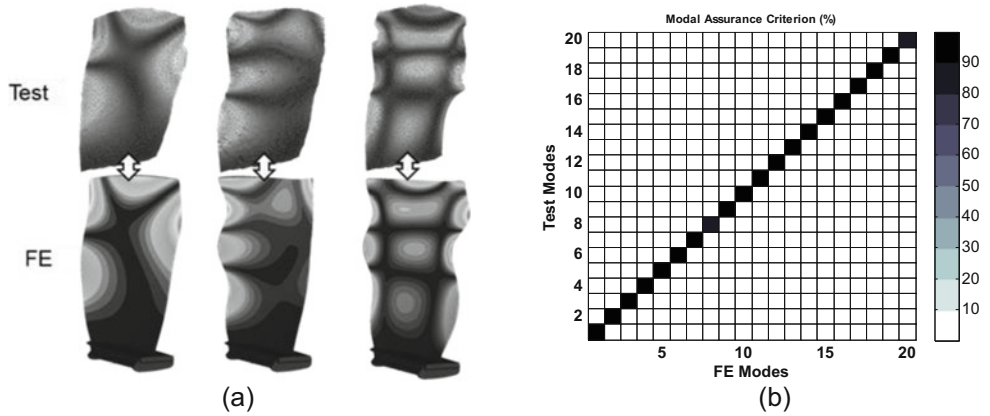


Fig. 2.5 (**a**) Sub-set of measured and predicted compressor blade mode shapes, (**b**) MAC matrix of displacement mode shapes

with complicated, localised variations and therefore needs appropriate care. One important point to note in the case full-field measurements of this kind is that, unlike conventional accelerometer or strain gauge etc. measurements where test grid is much less refined, the measurement grid can be denser than the corresponding distribution of the FE nodes. This is often not an issue for displacement mode shape measurements where the measurement grid is optimised using the FE model and therefore guaranteeing the alignment measurement points with the FE nodes. However the same is not necessarily the case for strain measurements where the density of test points can be much higher or when the model used in test planning and correlation are different (e.g. reversed engineered FE models).

2.5 Correlation of Full-Field Strain Measurements

Following the measurements of mode shapes on the turbine and compressor blades, a number of modes were identified for the measurement of full-field strains. Up to 12 modes covering the low, middle and high frequency range of modes identified in mode shape campaign were targeted. This enabled suitable coverage of complexity in terms of shapes targeted as well as allowing manageable measurement times. One important difference between displacement mode shape and strain mode shape measurements was the measurement grid used. Figure 2.6a shows the grid used for displacement measurements whilst Fig. 2.6b shows that used for strain. There are six times as many points in the strain grid as there are in the displacement grid. Since this density is much finer than the density of the FE nodes on the target surface, the new grid had to be created in the measurement system. The care was taken to ensure an overall regular distribution of points; however, the density was markedly increased around the edges and at areas where high gradient variations were expected.

Strictly speaking, mostly strain operational deflection shapes (ODSs) rather than true strain mode shapes are measured. Due to base and acoustic excitation techniques used in the tests, measuring the actual force driving the components was not possible. However, in case of strain measurements to be used in quantified correlation with FE predictions, the input voltage into generator was used as a reference signal.

The theory behind estimation of surface strain distributions from displacement measurements can be found in various publications [5, 9]. A triangular mesh is generated for the strain grid and virtual strain gauges are placed at each edge of each triangle as shown in Fig. 2.6c. For the grid shown in (b) this corresponds to having a total of 1800 virtual strain gauges. As described in [10], the displacements are transformed to the plane of surface triangles in the form of 2 in-plane and one out-of-plane components. Using these components strains are calculated for all surface triangles before being transformed back into global coordinate system.

Another critical difference between displacement and strain measurements is that the alignment precision required for strain measurements is much higher. The first step in measurements with the 3-D SLDV, regardless of whether displacement or strain is measured, is to perform the 3-D alignment. When the main vibration response is out-of-plane, such as is the case for turbine and compressor blades used in this work, potential crosstalk between out-of plane and in-plane components results in large errors, particularly for weak in-plane components. This is especially a challenge for strain calculations which relies on accurate resolution of weak in-plane, as well as strong out-of-plane displacements. To avoid these errors, alignment of all three laser beams with respect to measurement point should be ensured, i.e. all three beams should intersect on the measurement point. This, in Polytec 3-D SLDV system used in this study, is achieved via a process called video triangulation. More information on video triangulation can be found in [11] where the basic idea is the identification of the positions of the three beams in the video image by image processing. Once this is done, corrections to the mirror angles are applied until the beams meet precisely at the measurement point. To guarantee the best results, the video triangulation is performed at each such point. Since the alignment is done prior to measurements at each point, beam drift problems are eliminated, resulting in reliable and consistent alignment throughout the measurement session. Although performing video triangulation at each point results in considerably long measurement times, for a given setup, this needs to be done only once. Subsequent measurements can be performed much faster as the alignment information for each point is reused.

A sweep sine test was performed before strain measurements and this was repeated each time a configuration changed or ambient conditions were different to make sure that the exact natural frequencies at the time could be identified. Then the

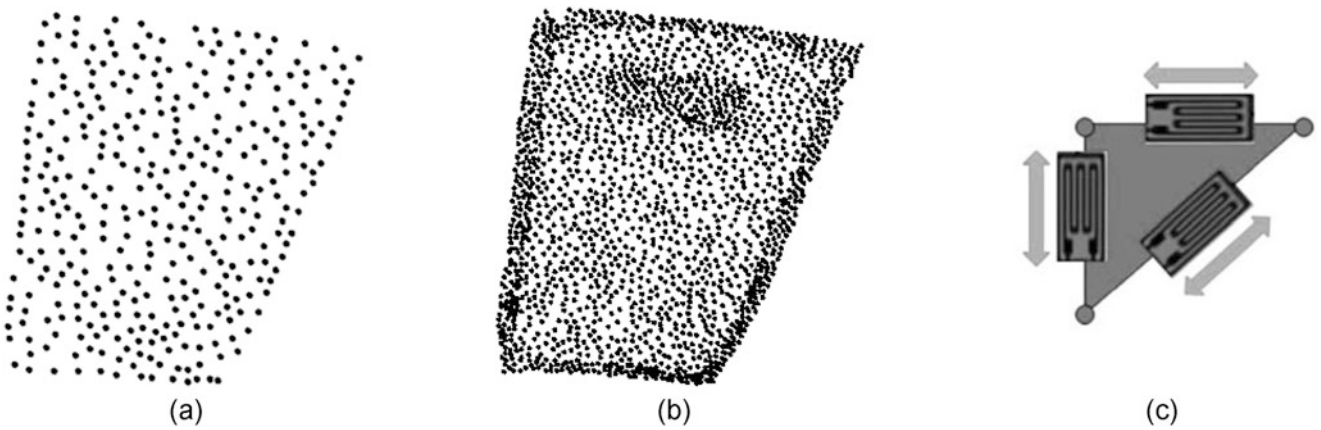


Fig. 2.6 (**a**) Displacement measurement grid, (**b**) strain measurement grid, and (**c**) placement of virtual strain gauges (taken from [10])

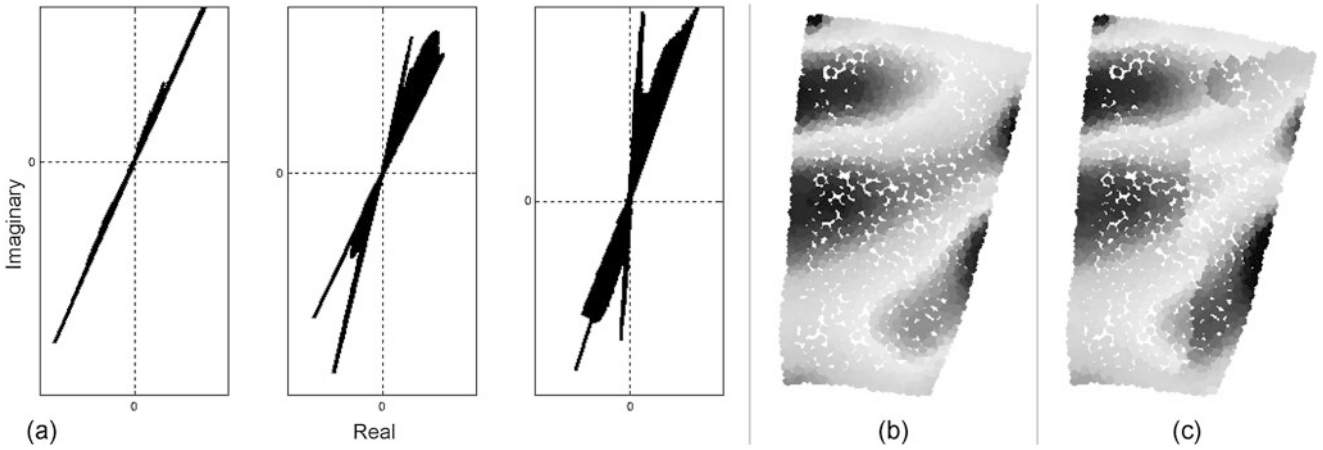


Fig. 2.7 (**a**) Strains on complex plane. Correct (**b**), and, incorrect (**c**) realisation of surface strain

strain measurements were performed in two ways. First, a number of selected modes were excited all at once by generating a multi-sine signal composed of natural frequencies of these modes. Since in this case the input energy was shared across all available components, each mode could only be excited to low amplitudes. In the second way, tests were repeated by concentrating on each mode separately. Since only one frequency was output at any given time, higher response amplitudes could be achieved. Performing measurements at different amplitudes this way was targeted as it would allow assessment of any potential dependency of strain distributions to response amplitudes. In this context multi-sine testing provided a time-efficient way of acquiring low-amplitude strain distributions for modes of interest.

Strain mode shapes estimated from the 3-D SLDV system are complex in mathematical sense and are often output in terms of real and imaginary parts. To be able to compare and correlate with FE predictions which are real-valued (e.g. signed amplitudes), a further processing of measured strains is needed to "realise" these otherwise complex quantities. Figure 2.7a shows Argand diagram representation of a number of strain mode shapes measured on the compressor blade together with results of some realisation attempts in Fig. 2.7b and c. It is evident from the co-linearity of the complex vectors that the underlying strain mode shapes are mainly real, albeit rotated in the complex plane. These vectors need to be appropriately "realised" by rotating them to align with x-axis and by removing remaining small complexities.

Care should be taken whilst performing the realisation to ensure unrealistic phase boundaries are not created, as illustrated in Fig. 2.7c, which in turn impact quantified correlation between measurements and the predictions. It is possible to do the realisation manually however it prolongs the process and does not allow batch processing in an automation environment. To this end, an algorithm is developed via which it is possible to identify the best-fit rotation of the vectors automatically for realisation of estimated surface strains.

Figure 2.8 shows strain measurement results on the compressor and the turbine blades for a number of modes as well as the predicted strain distributions from their nominal FE models. All six components of surface strain are estimated however only strains in z-direction are shown in the plots and used in the subsequent correlation processes.

Since the contouring used in the FE package in construction of predicted strain mode shapes is independent of the one used in the measurements, both sets of shapes may appear different at first sight. However, at a closer inspection the predicted and measured strain mode shapes given in Fig. 2.8 show a remarkable degree of similarity. This is true for high frequency complex modes (right hand side) as well as for low frequency fundamental modes (left hand side) where the evident similarity may more readily be expected.

Although visual examinations are often the first port of call, a more fundamental insight can only be obtained by performing a proper quantified correlation. To this end well documented displacement mode shape correlation metrics were used. Overall similarity of measured strain mode shapes with corresponding predictions was performed via MAC, and, the role of individual DOFs in this similarity, or lack of, was quantified via CoMAC. Utilising the same definition given for displacement mode shapes in [12]; MAC between measured, S_A, and a predicted, S_X, strain mode shapes is calculated as $MAC(A,X) = (|\{S_X\}^T\{S_A\}|^2)/(\{S_X\}^T\{S_X\})(\{S_A\}^T\{S_A\})$. This is a value between 0 and 1, with 1 representing identical mode shapes. However, in the plots and tables given here it is converted to percentages. Plot given in Fig. 2.9b shows the familiar MAC matrix, this time computed for the strain mode shapes. As mentioned before this is not strictly necessary as by this stage the mode pairing should already have been established through displacement mode shape MAC. Nevertheless it is given here to show that the mode pairing could equally be done using the strain MAC.

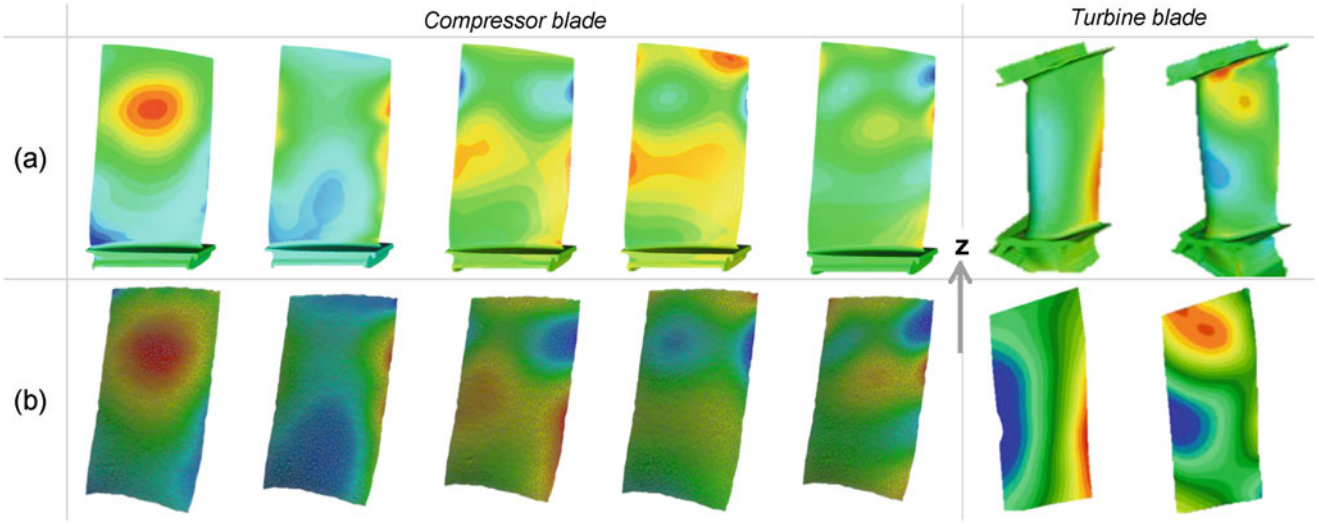
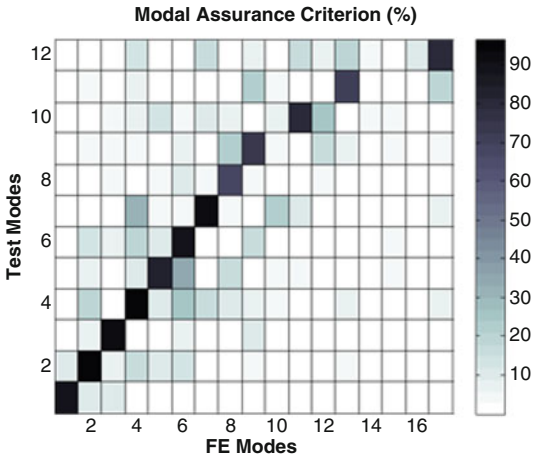


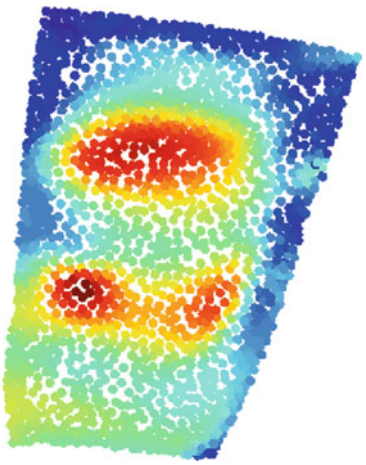
Fig. 2.8 (**a**) Predicted strain, (**b**) measured strain. *[both in z-direction]*

Pair #	ω_n diff (%)	MAC (%)
1	-1.3	89
2	-2.2	96
3	2.3	91
4	-0.1	97
5	3.2	84
6	-1.7	89
7	2.8	91
8	1.4	68
9	-1.2	73
10	-1.4	79
11	-1.5	71
12	-2.2	80

(a)



(b)



(c)

Fig. 2.9 (**a**) Quantified correlation of paired strain mode shapes and natural frequencies, (**b**) strain MAC matrix, and, (**c**) strain CoMAC (z-direction, *blue*➔low CoMAC, *red*➔high CoMAC)

Although the MAC is a very useful parameter, it is also a global similarity metric and does not necessarily give any insight into distribution of correlations across DOFs used. Coordinate MAC or CoMAC has been developed for this purpose [12], and for an individual DOF, i, it can be calculated across all correlated mode pairs as $CoMAC(i) = \left(\sum_{n=1}^{N} |(S_X)_{in}(S_A)_{in}|^2\right) \Big/ \left(\sum_{n=1}^{N} (S_X)_{in}^2 \cdot \sum_{n=1}^{N} (S_A)_{in}^2\right)$. This is a value between 0 and 1. Here N represents total number of paired modes. The list of paired modes used in strain CoMAC calculations for the compressor blade is given in Fig. 2.9a. Low values of CoMAC indicate DOFs that have relatively low contributions to overall correlation. Although locations of these DOFs are where the differences between the test and the mathematical model manifest themselves, they are not necessarily the locations responsible for the said differences.

Having warned against temptation to brand low CoMAC areas as the sources of discrepancies, the fact remains that the plot given in Fig. 2.9c presents a very regular distribution. Even if this cannot readily be identified form the plot, it is perfectly reasonable to presume a systematic discrepancy being present and this can be a very useful tool in identification of potential sources of such discrepancies between the FE model and the test. It is not hard to see that once such connections between CoMAC variations and underlying causes are established, effectiveness of model updating efforts can be significantly improved.

A CoMAC plot of this kind, to the best of authors' knowledge, is presented here for the first time. Therefore there is very little experience with what it might potentially reveal. However given that it is the result of direct strain measurements, it may point to potential discrepancies between simulation and test more directly. Nevertheless this is at best a speculation at this stage and requires a systematic study to establish with clear evidence.

Overall, the results given in Fig. 2.9 constitute a remarkable body of evidence for success of full-field non-contact strain measurements and more importantly for quantified correlation with their predicted counterparts. Up to 12 mode pairs, spanning the full operational range of the selected engine component, are identified with average MAC value of 84% with most of the modes near 90% or higher.

Nevertheless, some of the mode pairs have markedly low MAC values. There are a number of potential reasons for this. First of all, the measured results are compared with the nominal FE model constructed from the design intent. The test hardware has inevitable deviations from this design intent due to manufacturing tolerances. Furthermore, the blade model is simulated with fixed boundary conditions applied to its root whereas the actual fixture has a finite flexibility with potential nonlinear interactions due to contact and friction. Both of these causes need further exploration and they will be addresses in future studies.

One additional source for discrepancies has to do with the alignment errors during the process of matching measurement points with the FE nodes or in short; node-point pairing. In cases where the test planning for identification of measurement points is performed on the FE model that is used in eventual correlation, this errors should be very low. However as explained above, even when this is done model and real hardware geometries will not exactly match and there will be deviations. In the present case the test planning was carried out for the displacement mode shape measurements. As this grid was not dense enough for strain measurements, it was made denser in the measurement environment by manually adding new data points. As these new points could not be guaranteed to coincide with FE nodes, they were paired with the nearest ones which resulted in inevitable discrepancies. Node-point pairing done for strain correlation of the compressor blade is given in Fig. 2.10. Magnified portion given therein demonstrates the point being made here where measurement points are often outside circles representing FE node targets.

Although alignment errors can be reduced by generating the strain grids from the FE model, eliminating them completely is not viable. Inevitably, differences between simulated and real geometries will exist. Moreover, sometimes FE nodes present on the surface to be measured may not be dense enough for strain measurements whereby new points will have to be manually added. A better solution to this issue would be to map the measurement grid to the simulation surface as accurately as possible and then interpolate FE predictions at the actual measurement points. This will greatly reduce the alignment errors. Given that the density of the measurement points is often sufficiently high compared with the surface strain variations, the interpolations could be done linearly without having to resort to more sophisticated methods.

It is often assumed that the strain values scale up with vibration amplitudes linearly, preserving the shape of distribution. Although linearity checks are factored into any serious test campaign wherever this is assumed, it is not always straightforward to capture it with conventional strain measurement means such as with strain gauges. Amplitude dependent variations could be quite complex and distributed which render discrete measurements unsuited to the task. Full-field strain measurements provide a unique insight into capturing such variations, if present.

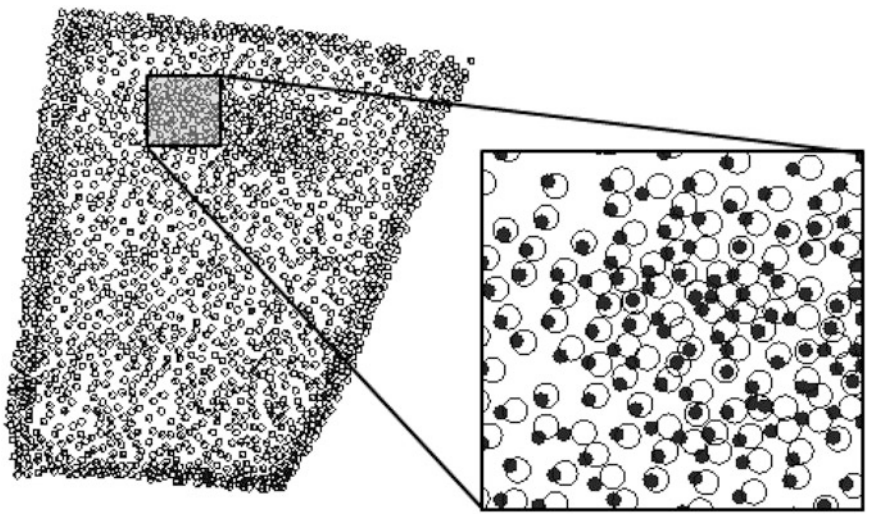


Fig. 2.10 Node-point alignment errors

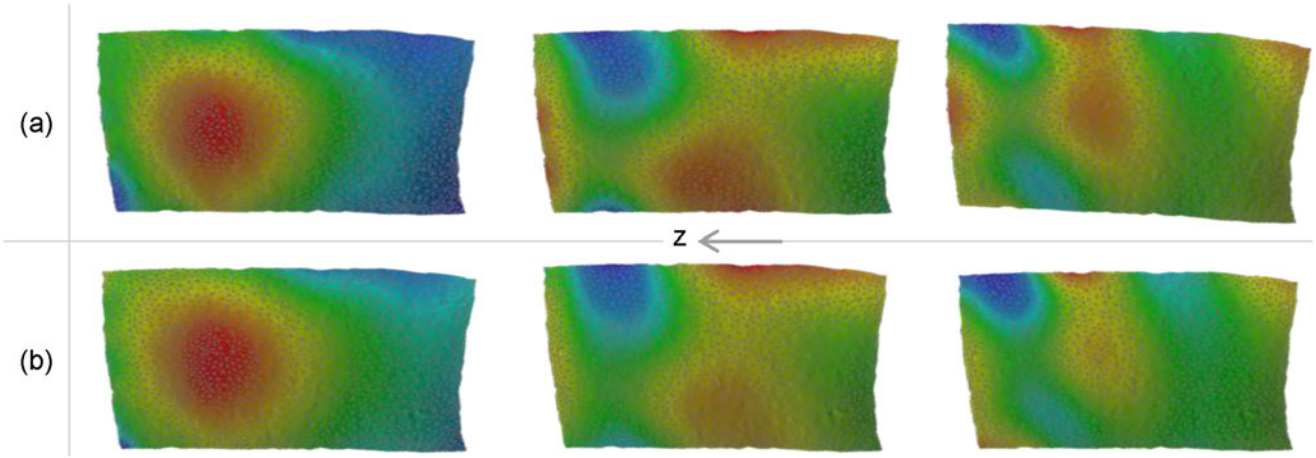


Fig. 2.11 (**a**) High-excitation, (**b**) low-excitation. [both in z-direction]

As explained earlier, the strain measurements were repeated at low and relatively high vibration amplitudes for this purpose. The resultant strain shapes for a number of selected compressor blade modes are given in Fig. 2.11. Strain plot contours in (a) for the high-excitation, and, in (b) for the low-excitation are scaled to their maximum amplitudes. This is useful for assessing in-plot patterns across low and high excitation amplitudes. Strain values in the high-excitation case were on average an order of magnitude higher but even then they were much smaller than operational levels. Despite that and despite overall similarity between the two cases, there are nontrivial features emerging in the high-excitation case that are not present in the low-excitation case. In particular, left-hand edges of middle and right modes in row (a) now feature strain hotspots that are absent from corresponding shapes in (b). It would be interesting to see how these variations evolve as the strain amplitudes approach those of the operating environment. This will be explored in future studies where the variations will also be numerically quantified.

2.6 Concluding remarks

A detailed study is presented where full-field non-contact surface strains are measured and integrated into simulation environment using advanced measurement and analysis techniques. This was demonstrated on a number of real engine hardware as well as for a wide dynamic range spanning from low frequency fundamental modes to high frequency complex modes.

Through careful post-processing and integration with simulation results, it has been shown that assessment of design intent can be effectively done using directly acquired full-field non-contact strain measurements. Traditionally displacements are measured and correlated in this manner as they are the most readily accessible quantities. The main objective inevitably is to get a handle on stresses and strains as they are the ultimate parameters used in assessment of the structural integrity. The ability to measure and then perform quantified correlation of strains directly is a major achievement and a step change from existing established practices.

Utilising existing modal analysis correlation metrics, not only the overall degree of similarity but also the coordinate specific contributions were quantified. A remarkable degree of match was obtained between measured and predicted strain mode shapes both visually and numerically. By repeating the measurements at low and high excitation levels, vibration amplitude dependence of strain mode shapes was evaluated. It is important to note that some of the changes were distributed and as such would not have been possible to capture with use of discrete strain measurement means (e.g. SGs).

Although on average the degree of correlation obtained was high, there were some modes where this was markedly low. There are a number of reasons. Some of them have to do with the discrepancies in representation of physics between the test and the simulation environments. Boundary conditions and interactions at these boundaries together with variations in actual and simulated geometries and alignment errors between test and FE coordinates were contributing factors.

This study will be expanded to account for these shortcomings. First of all more blades will be tested to form a statistical understanding. Reverse engineered FE models for each blade will be created by geometry scanning their actual hardware. This will significantly reduce uncertainties introduced due to mass and stiffness distribution discrepancies. Care will be

taken to get simulation models to replicate test boundary conditions more faithfully rather than using idealised constraints. In addition, alignment process performed to match test points with FE nodes will be improved considerably to eliminate artificial errors introduced with the existing method.

Acknowledgements The authors are grateful to Rolls-Royce plc for allowing the publication of this work.

References

1. Pedrini, G., Tiziani, H.J.: Double pulse electronic speckle interferometry for vibration analysis. Appl. Opt. **33**, 7857–7863 (1994)
2. Pedrini, G., Froning, P., Fessler, H., Tiziani, H.J.: Transient vibration measurements by using multi-pulse digital holography. Opt. Laser Technol. **29**(8), 505–511
3. Stanbridge, A.B., Martarelli, M., Ewins, D.J.: Scanning laser Doppler vibrometer applied to impact modal testing. Shock Vib. Digest. **32**, 35 (2000)
4. Beberniss, T.J., Ehrhardt, D.A.: High-speed 3D digital image correlation vibration measurement: Recent advancements and noted limitations. Mech. Syst. Signal Process. (2016). doi:10.1016/j.ymssp.2016.04.014
5. Vuye, C., Vanlanduit, S., Preseznia, F., Steenackers, G., Guillaume, P.: Optical measurement of the dynamic strain field of a fan blade using a 3D scanning vibrometer. Opt. Lasers Eng. **49**, 988–997 (2011)
6. Maguire M., Sever I. A.: Full-field strain measurements on turbomachinery components using 3D SLDV Technology. AIVELA conference, 28 June–1 July 2016, Ancona, Italy
7. Sever, I. A: Experimental validation of turbomachinery blade vibration predictions. Ph.D. thesis, Imperial College London (2004)
8. Kammer, D.C.: Sensor placement for on-orbit modal identification and correlation of large space structures. J. Guid. Control. Dyn. **14**(2), 251–259 (1991)
9. Cazzolato, B., Wildy, S., Codrington, J., Kotousov, A., Schuessler, M.: Scanning laser vibrometer for non-contact three-dimensional displacement and strain measurement. In: Proceedings of Acoustics 2008, 24–26 November 2008, Geelong, Australia (2008)
10. Polytec, "Polytec data sheet – Optical Measurement of Dynamic Stress and Strai", Waldbronn: Polytec (2011)
11. Polytec, "Importance of Video Triangulation for the PSV-3D Scanning Vibrometer", Application note VIB-G-19, (2010)
12. Ewins, D. J.: Modal Testing: Theory, Practice and Application, 2nd edn. Research Studies Press, Baldock (2000)
13. Polytec user manual – Vibrometer controller OVF 5000

Chapter 3
Nonlinear Prediction Surfaces for Estimating the Structural Response of Naval Vessels

Alysson Mondoro, Mohamed Soliman, and Dan M. Frangopol

Abstract Structural health monitoring (SHM) of naval vessels is essential for assessing the performance of the structure and the fatigue damage accrued over the service life. The direct integration of available SHM data may be useful in reducing the epistemic uncertainties arising from inaccuracies in the modeling and the variations in the as-built structural configuration from the initial design. Based on SHM data, fatigue damage indices can be predicted by implementing cell based approaches, such as the lifetime weighted sea method, that discretizes the operational conditions of the vessel into cells with specific wave height, heading angle, and speed. The integration of SHM data into the fatigue assessment using lifetime weighted sea method requires a complete set of data that covers the whole operational spectrum. However, technical malfunctions or discrete monitoring practices generate incomplete data sets. This paper proposes nonlinear prediction surfaces to estimate the ship structural response in unobserved cells based on available cell data. Expected theoretical variations of the structural response to changes in wave height, heading angle, and vessel speed are integrated in the development of the prediction surface. The proposed methodology is illustrated on the SHM data from a high speed aluminum catamaran.

Keywords Fatigue • Aluminum vessels • Structural health monitoring • Missing data • Nonlinear prediction

3.1 Introduction

SHM data can aid in the life-cycle management of structures by helping to identify the discrepancies between predicted and observed performance. The recorded SHM data provides an indication of the as-built condition of the ship and any variations in observed response from anticipated design conditions [1]. In both civil and marine structures, SHM data can be used to update design estimates for expected loads, structural responses, and fatigue life evaluation [2–5]. Fatigue cracking is a major concern in the life-cycle management of naval vessels. The constant fluctuations in loading, induced by the natural variability in sea surface, contributes to damage accumulation in fatigue sensitive details. If the observed conditions deviate substantially from their predicted values, fatigue damage may be either (a) significantly less than expected and lead to unnecessary and costly inspections, or (b) significantly higher than expected and may induce catastrophic failure with high consequences.

The lifetime weighted sea method, used to assess the fatigue life, is developed around the assumption that the operational condition can be discretized into cells where the response in each cell is stationary [6]. A cell is defined by a set of operational conditions which include wave height, ship speed, and heading angle. The total response is the summation of the response in each cell weighted by the probability of occurrence of the cell [7]. The prediction of fatigue life thus requires information to be available for all cells.

Missing data is a problem inherent in the use of SHM for fatigue analysis. First, discrete monitoring practices, while useful in limiting financial costs [8], can lead to some operational states (i.e., cells) not being recorded. Second, technical malfunctions can result in missing or unreliable data [9]. Lastly, even if data is recorded for all cells in the current operational theatre, the future operational conditions may not be similar to past ones; thus, there may be cells in the future profile that can be considered as missing data. Zhu [10] and Mondoro et al. [11] have begun to address the problem of missing data with respect to the structural response characterization of naval vessels. Linear surfaces were used to relate the operational

A. Mondoro • D.M. Frangopol (✉)
Department of Civil and Environmental Engineering, ATLSS Engineering Research Center,
Lehigh University, 117 ATLSS Dr., Bethlehem, PA 18015-4729, USA
e-mail: Dan.frangopol@lehigh.edu

M. Soliman
School of Civil and Environmental Engineering, College of Engineering, Architecture and Technology,
Oklahoma State University, 207 Engineering South, Stillwater, OK 74078-5033, USA

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_3

conditions (i.e. wave height, ship speed, and heading angle) to a structural response characteristic. The linear surface is useful for the ease of implementation. Additionally, it requires only a minimal amount of prediction model parameters to be estimated which limits the variations that arise based on availability of data. However, the linear surfaces lack a theoretical foundation.

This paper proposes a nonlinear surface for use in predicting unobserved data. The prediction is based on the theoretical relationship between operating conditions and the structural response given as a function of the vertical bending moment. Available data is discretized into cells, the low and high frequency content are separated and fit with response parameters as detailed in [11], and the goodness-of-fit of the theoretically-based nonlinear prediction surface is evaluated and compared with that of the linear surface. Furthermore, the performance of the theoretically-based nonlinear prediction surface is evaluated for several cases with different percentage of missing data. The methodology is applied to the SHM data obtained during the seakeeping trials of the HSV-2 Swift.

3.2 Available Data and Analysis

The full stress time-history of a structural detail is a nonstationary random process due to the exposure to various loading conditions associated with sea states, routes, and speeds. However, the full stress time-history response can be discretized into cells based on wave height H_s, ship speed V, and heading angle β [6]. This discretization leads to stationary processes for the stress time-history in each cell, for which the response spectrum can be estimated. Mondoro et al. [11] proposed that the SHM response spectrum can be fit with functional forms developed from wave spectra. The low and high frequency components were fit separately in order to account for the following actions: wave loads for low frequency, and slamming for high frequency. The Pierson-Moskowitz wave spectrum and the Joint North Sea Wave Observation Project (JONSWAP) were included as two representative wave spectra. The generalized variations of the Pierson-Moskowitz spectrum and the JONSWAP spectrum take the form [11]

$$S^{+}_{PM_{GEN}}(\omega) = \frac{A_{LF}}{\omega^5} e^{-B_{LF}\omega^{-4}} + \frac{A_{HF}}{\omega^5} e^{-B_{HF}\omega^{-4}} \tag{3.1}$$

$$S^{+}_{JONSWAP_{GEN}}(\omega) = \frac{C_{LF}}{\omega^5} \exp\left(-\frac{5}{4} {D_{LF}}^4 \omega^{-4}\right) {E_{LF}}^{\exp\left(-\frac{(\omega - D_{LF})^2}{2{D_{LF}}^2\sigma^2}\right)} + \frac{C_{HF}}{\omega^5} \exp\left(-\frac{5}{4} {D_{HF}}^4 \omega^{-4}\right) {E_{HF}}^{\exp\left(-\frac{(\omega - D_{HF})^2}{2{D_{HF}}^2\sigma^2}\right)} \tag{3.2}$$

where A_{LF} and B_{LF} are fitting parameters for the low frequency content and A_{HF} and B_{HF} are fitting parameters for the high frequency content of the complete generalized Pierson-Moskowitz function, $S^{+}_{PM_{GEN}}$; and C_{LF}, D_{LF}, and E_{LF} are the fitting parameters for the low frequency content and C_{HF}, D_{HF}, and E_{HF} are fitting parameters for the high frequency content of the complete generalized JONSWAP function,$S^{+}_{JONSWAP_{GEN}}$.

3.3 Development of Theoretical Prediction Surfaces

Discrete monitoring practices and technical malfunctions contribute to the missing data problem inherent in SHM. This presents a significant challenge in the fatigue analysis of naval vessels, which is dependent on complete data sets. Linear prediction surfaces have been employed in [10, 11] to relate structural response to wave height, ship speed, and heading angle. The linear prediction surface, Ψ^{lin}, is defined as

$$\Psi^{lin} = p_1 H_s + p_2 V + p_3 \cos(\beta) \tag{3.3}$$

where H_s is the wave height, V is the ship speed, and β is the heading angle. The linear surface is included in this paper for comparison purposes. The following subsections present the development of the theoretically-based nonlinear prediction surfaces. The discussion is framed around the HSV-2 Swift but is readily applicable to other naval vessels.

3.3.1 Operational Conditions and Theoretical Response

The theoretical relationship between the operational condition (i.e., wave height, ship speed, and heading angle) and ship response (i.e., vertical bending moment) is developed by investigating the response spectrum. For naval vessels, response spectrum can be decomposed into the wave spectrum and a transfer function which quantifies the structural response to a unit sinusoid at each frequency. The response spectrum for vertical bending moment S_{VBM} is defined as

$$S_{VBM}(\omega) = [\Phi_m]^2 S_\xi(\omega) \tag{3.4}$$

where Φ_m is the transfer function for vertical bending moment and S_ξ is the wave spectrum. There are many available forms for the wave spectrum. This paper uses the Pierson-Moskowitz wave spectrum [12]

$$S_{PM}(\omega) = \frac{\alpha g^2}{\omega^5} \exp\left(-\frac{5}{4}\left(\frac{\omega}{\Omega}\right)^{-4}\right) \tag{3.5}$$

where α is 8.10×10^{-3}, g is the gravitational constant (9.81 m/s^2), and Ω is the wave frequency. It is assumed that the general relationship between the wave frequency and wave height is $\Omega = 0.14\, g/U_{19.5}$ [12] where $U_{19.5} = (H_s/0.021)^{0.5}$ [13].

The transfer function for the vertical bending moment is derived using linear strip theory for a box shaped vessel [14]

$$\Phi_m = \kappa \frac{1 - kT}{(k_e L)^2}\left[1 - \cos\left(\frac{k_e L}{2}\right) - \frac{k_e L}{4}\sin\left(\frac{k_e L}{2}\right)\right] F_v F_c \sqrt[3]{|k\cos\beta|} \tag{3.6}$$

$$F_v = 1 + 3{F_n}^2 \tag{3.7}$$

$$F_c = (1 - \vartheta)^2 + 0.6\left(1 - F_n\sqrt{kL}\cos\beta\right)(2 - \vartheta) \tag{3.8}$$

$$F_n = \frac{V}{\sqrt{g_o L}} \tag{3.9}$$

$$k_e = |k\cos\beta| \tag{3.10}$$

$$\kappa = \exp(-k_e T) \tag{3.11}$$

where k is the wave number, which for deep water waves, $k \approx \omega^2/g$, T is draught of the ship, L is length of the ship, g_o is a general characteristic of the external field, and ϑ is a function of the block coefficient C_b [15].

The response spectrum for vertical bending moment is thus related to the significant wave height H_s through the wave spectrum, and, to β and V through the theoretical transfer function for vertical bending moment. The most probable extreme value for the response can be estimated as [6]

$$Q_p = \sqrt{2m_0 \ln\left(\frac{1800T}{\pi}\sqrt{\frac{m_2}{m_0}}\right)} \tag{3.12}$$

where m_0 and m_2 are the 0th and 2nd spectral moments, where the n^{th} spectral moment, m_n, of a spectrum, S_R, is defined as

$$m_n = \int_0^\infty \omega^n S_R(\omega)\, d\omega \tag{3.13}$$

The most probable vertical bending moment, M_p, can be related to H_s, V, and β. Figure 3.1 illustrates the procedure for developing the relationship between ship speed and the most probable vertical bending moment for the HSV-2 Swift using information from Brady et al. [16]. The ship speed is varied and ${\Phi_m}^2$ is calculated based on Eq. (3.6) (Fig. 3.1a). It is important to note that H_s, V, and β are assumed to be uncoupled in regards to M_p. Therefore, only a single parameter of $\{H_s, V, \text{and } \beta\}$ is varied at a time. Figure 3.1b shows the wave spectrum. Since S_{PM} is not a function of V, all lines depicting

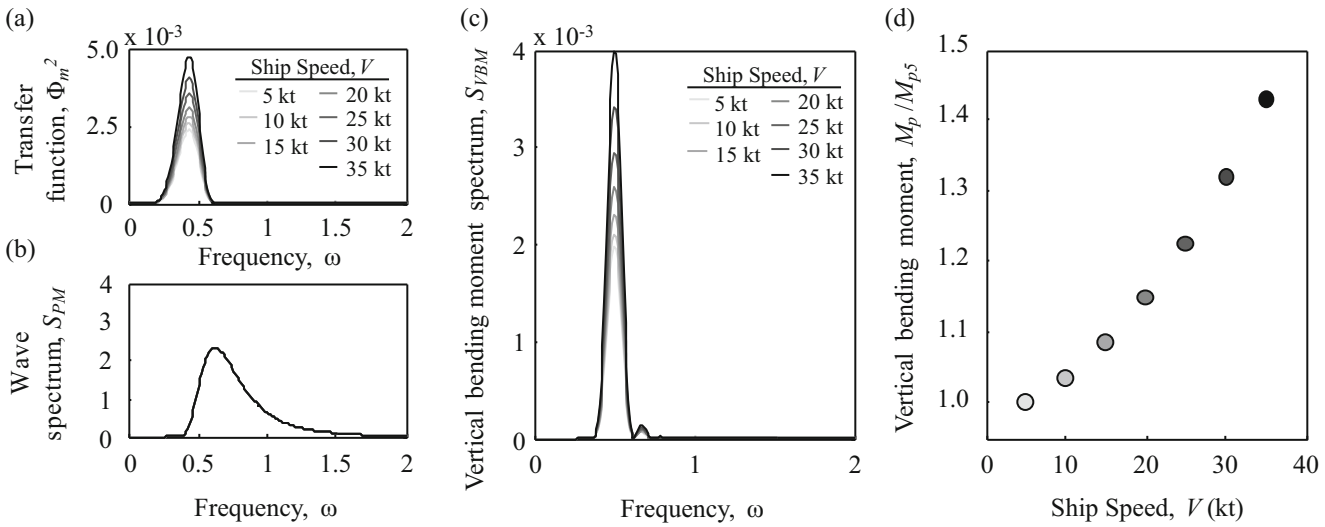


Fig. 3.1 The variations in the (**a**) transfer function $\Phi_m{}^2$, (**b**) wave spectrum S_{PM}, (**c**) response spectrum S_{VBM}, and (**d**) moment M_p/M_{p5} to changes in the speed of the ship

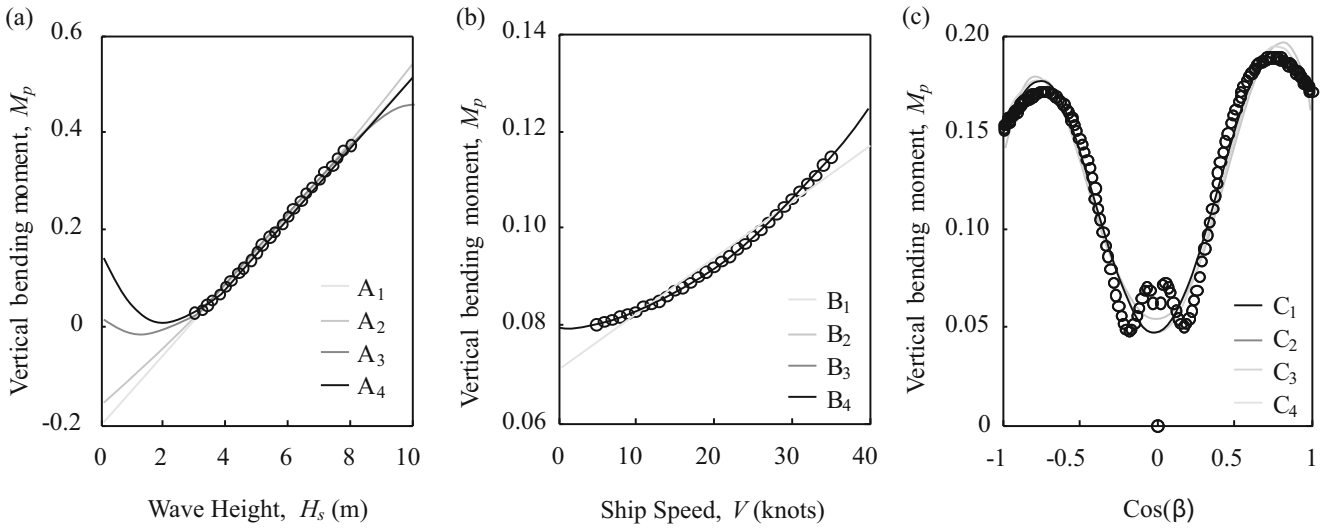


Fig. 3.2 Theoretical variations of bending moment with (**a**) wave height, (**b**) ship speed, and (**c**) heading angle (the theoretical values are shown as *black circles* and proposed functional forms are fit to each and shown as *solid lines*)

the variation of S_{PM} with V lie on top of each other. The response spectrum S_{VBM} is generated for each value of V using Eq. (3.4) and is shown in Fig. 3.1c. S_{VBM} is then used in Eqs. (3.12) and (3.13) to find the most probable vertical bending moment. The variation of M_p and V is depicted in Fig. 3.1d; the vertical bending moment at each ship speed is compared to the vertical bending moment at ship speed of 5 knots, M_{p5}. The same procedure is applied for H_s and β but figures are omitted for brevity.

3.3.2 Development of Functional Forms

A closed-form function to describe theoretical relationship between the operational condition (i.e., H_s, V, and β) and the response (i.e., M_p) is not readily available. As noted in Sect. 3.1, H_s, V, and β are assumed to be uncoupled in regards to M_p; therefore, M_p (H_s, V, β) can be decoupled into M_p (H_s), M_p (V), and M_p (β). The theoretical variations of bending moment with wave height, ship speed, and cosine of the heading angle are shown in Fig. 3.2a–c respectively. Four functional forms were developed for M_p (H_s), M_p (V), and M_p (β) and are listed in Table 3.1. Polynomial functions were used to describe M_p

Table 3.1 Proposed forms for $M_p\ (H_s)$, $M_p\ (V)$, and $M_p\ (\beta)$ and the Root Mean Square Error

Name	Form for $M_p\ (H_s)$	Root mean square error
A_1	$p_1 H_s + p_0$	0.0047
A_2	$p_2 {H_s}^2 + p_1 H_s + p_0$	0.0034
A_3	$p_3 {H_s}^3 + p_2 {H_s}^2 + p_1 H_s + p_0$	0.0007
A_4	$p_4 {H_s}^4 + p_3 {H_s}^3 + p_2 {H_s}^2 + p_1 H_s + p_0$	0.0001
Name	Form for $M_p\ (V)$	Root mean square error
B_1	$p_1 V + p_0$	0.0018
B_2	$p_2 V^2 + p_1 V + p_0$	3.78E−05
B_3	$p_3 V^3 + p_2 V^2 + p_1 V + p_0$	3.77E−09
B_4	$p_4 V^4 + p_3 V^3 + p_2 V^2 + p_1 V + p_0$	8.92E−11
Name	Form for $M_p\ (\beta)$	Root mean square error
C_1	$p_1 \sin(p_2 \cos(\beta) + p_3) + p_4 \sin(p_5 \cos(\beta) + p_6)$	0.0083
C_2	$p_1 \cos(\cos(\beta)) - p_2 \cos(p_3 \cos(\beta)) + p_4 \cos(\beta)$	0.0086
C_3	$p_4 \cos(\beta)^4 + p_3 \cos(\beta)^3 + p_2 \cos(\beta)^2 + p_1 \cos(\beta) + p_0$	0.0107
C_4	$p_6 \cos(\beta)^6 + p_5 \cos(\beta)^5 + p_4 \cos(\beta)^4 + p_3 \cos(\beta)^3 + p_2 \cos(\beta)^2 + p_1 \cos(\beta) + p_0$	0.0078

(H_s) and $M_p\ (V)$, while a two-term sinusoid, a summation of cosines and linear term, and a 4th and 6th order polynomials (similar to the 4th and 6th order Taylor series expansion for cosine) were used for $M_p\ (\beta)$. The goodness of fit is evaluated in terms of the root mean square error (RMSE) for each of the function for $M_p\ (H_s)$, $M_p\ (V)$, and $M_p\ (\beta)$ and listed in Table 3.1. In Fig. 3.2 the functions B_2, B_3, and B_4 lie on top of each other and fit the data points with a RMSE of less than 1.0E−4.

Based on the performance of the proposed functions for $M_P(H_s)$, $M_p(V)$, $M_P(\beta)$ as presented in Table 3.1, two nonlinear prediction surfaces are proposed. The first includes 2nd order polynomial functions for $M_P(H_s)$ and $M_p(V)$ (i.e. A_2 and B_2) and the summation of cosines and linear term for $M_P(\beta)$ (i.e. C_2). These functions were chosen in order to minimize the number of coefficients, p_i, while also having a low RMSE. The first nonlinear prediction surface takes the form

$$\Psi^{nonlin} = p_1 {H_s}^2 + p_2 H_s + p_3 + p_4 V^2 + p_5 V + p_6 \cos(\cos(\beta)) - p_7 \cos(p_8 \cos(\beta)) + p_9 \cos(\beta) \tag{3.14}$$

The second proposed nonlinear prediction surface restricts itself to the use of polynomial functions for H_s, V, and $\cos(\beta)$. The 2nd order polynomial functions A_2 and B_2 are used as the contributions for $M_P(H_s)$ and $M_p(V)$, respectively, and the 4th order polynomial (i.e. C_3) is used to account for $M_P(\beta)$. The polynomial based nonlinear prediction surface takes the form

$$\Psi^{nonlin-poly} = p_1 {H_s}^2 + p_2 H_s + p_3 + p_4 V^2 + p_5 V + p_6 \cos(\beta)^4 + p_7 \cos(\beta)^3 + p_8 \cos(\beta)^2 + p_9 \cos(\beta) \tag{3.15}$$

3.4 Application and Results

The HSV-2 Swift is an aluminum naval vessel which was instrumented with strain gauges and various other sensors for seakeeping trials. The ship was operated by systematically varying ship speeds and heading angles in different sea states [16]. The T2-4 sensor is located at a fatigue critical location on the HSV-2 Swift to specifically capture stresses induced by vertical bending. The SHM data for the T2-4 sensor were processed and fit with Eqs. (3.1) and (3.2). As a result, for each cell, the parameter set $\{A_{LF}, B_{LF}, A_{HF}, B_{HF}\}$ fully defines the structural response if Eq. (3.1) is used, and, $\{C_{LF}, D_{LF}, E_{LF}, C_{HF}, D_{HF}$, and $E_{HF}\}$ if Eq. (3.2) is used. In order to predict the response in missing cells, all parameters within the set must be extrapolated.

The performance of the nonlinear surfaces is first investigated with respect to their ability to capture the variation in the observed data. That is, all parameters in sets $\{A_{LF}, B_{LF}, A_{HF}, B_{HF}\}$ and $\{C_{LF}, D_{LF}, E_{LF}, C_{HF}, D_{HF}$, and $E_{HF}\}$, for Ψ^{lin}, Ψ^{nonlin}, and $\Psi^{nonlin\text{-}poly}$ are fit to the available data points and the mean square error (MSE) is calculated. This process is shown in Fig. 3.3 for C_{LF}. The data points for C_{LF} are plotted as a function of wave height and heading angle for a constant ship speed in Fig. 3.3a and as a function of ship speed and wave height for a constant heading angle in Fig. 3.3b. Figure 3.3a and b also show the fitted surfaces for Ψ^{lin}, Ψ^{nonlin}, and $\Psi^{nonlin\text{-}poly}$. The surfaces for Ψ^{nonlin} and $\Psi^{nonlin\text{-}poly}$ lie on top of each other in Fig. 3.3a and b and have comparable MSE, as listed in Table 3.2. The results for the MSE for Ψ^{lin}, Ψ^{nonlin}, and $\Psi^{nonlin\text{-}poly}$ for all parameters are listed in Table 3.2; similar figures can be generated for all other parameters but are omitted for brevity.

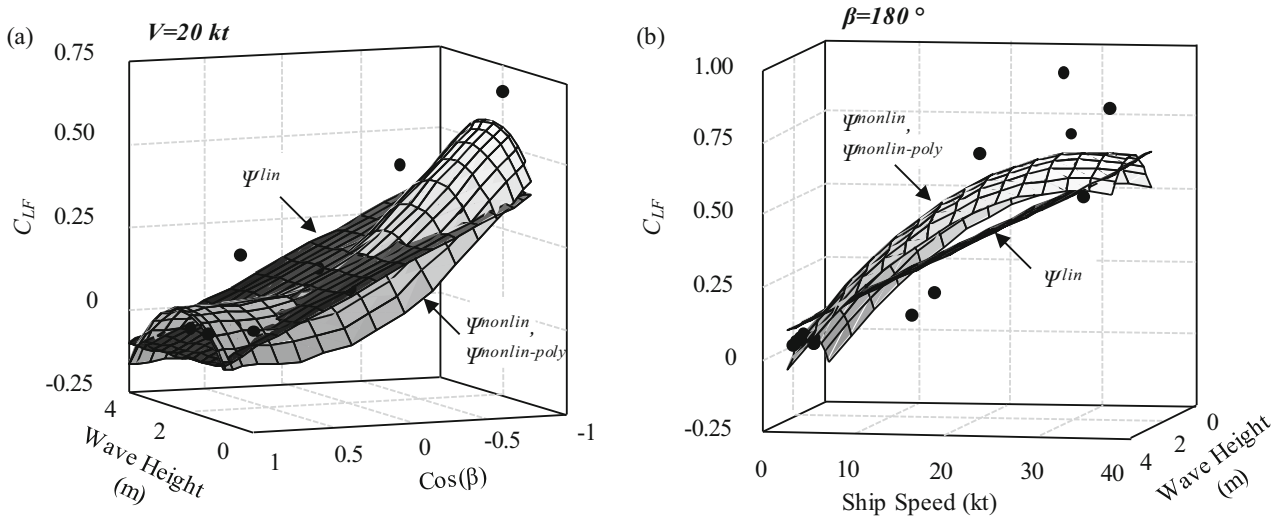


Fig. 3.3 Available data points for the response parameter C_{LF} (shown as *black* points) and the fitted surfaces for Ψ^{lin}, Ψ^{nonlin}, and $\Psi^{nonlin\text{-}poly}$ where (**a**) shows the variation of C_{LF} with wave height and heading angle for a ship speed of 20 kts and (**b**) shows the variation of C_{LF} with ship speed and wave height for a heading angle of 180°

Table 3.2 Evaluation of proposed surfaces with respect to observed data

Parameter	Mean square error Ψ^{lin}	Mean square error Ψ^{nonlin}	Mean square error $\Psi^{nonlin\text{-}poly}$	Parameter	Mean square error Ψ^{lin}	Mean square error Ψ^{nonlin}	Mean square error $\Psi^{nonlin\text{-}poyl}$
A_{LF}	0.022	0.022	0.020	C_{LF}	0.034	0.028	0.028
B_{LF}	0.034	0.022	0.017	D_{LF}	0.022	0.012	0.012
A_{HF}	0.103	0.099	0.099	E_{LF}	0.008	0.008	0.008
B_{HF}	0.083	0.087	0.086	C_{HF}	0.175	0.163	0.161
				D_{HF}	0.048	0.046	0.046
				E_{HF}	0.005	0.005	0.004

The theoretically-based nonlinear prediction surfaces typically outperform the linear surface for all low frequency parameters. The high frequency parameters show minimal or no improvement when compared to the linear surface. This can be attributed to the fact that the surfaces were developed based on wave bending moment, which is governed by the low frequencies. Slam impacts typically govern the high frequency response and would thus be better predicted by alternative nonlinear relationships.

The performance of the nonlinear surfaces is also investigated with respect to their ability to predict unobserved responses. In order to evaluate the performance of the prediction surfaces, data was removed from the initial data set (which includes all available sea trials). The data that was removed is referred to as the missing data set, while the remaining data is referred to as the available data set. Cases considering 70% available data and 50% available data were tested for predicting parameters in the sets $\{A_{LF}, B_{LF}, A_{HF}, B_{HF}\}$ and $\{C_{LF}, D_{LF}, E_{LF}, C_{HF}, D_{HF}$, and $E_{HF}\}$. C_{LF} is presented for further discussion; others are available but are omitted for brevity. Results of the prediction are shown in Table 3.2.

Ten sets of available data are included as representative data sets. Available data set 1 refers to the case of no missing data and it was found that Ψ^{nonlin} and $\Psi^{nonlin\text{-}poly}$ outperform Ψ^{lin}. The 70% available data and 50% available data cases are chosen randomly so as not to impart bias onto the available data set. Available data sets 2–5 use 70% of the available data and sets 6–10 use 50% available data. It can be seen that Ψ^{nonlin} and $\Psi^{nonlin\text{-}poly}$ outperforms Ψ^{lin} for available data sets 2–5 (i.e. 30% missing data) as shown in Table 3.3. The same can be said for available data sets 6–10 (i.e. 50% missing data).

Ψ^{nonlin} and $\Psi^{nonlin\text{-}poly}$ typically perform comparably for all available data sets. However, in some 50% available data cases, local fluctuations in the prediction surface can occur with Ψ^{nonlin} as shown in Fig. 3.4. While the MSE for Ψ^{nonlin} and $\Psi^{nonlin\text{-}poly}$ for available data set 6 are similar, $\Psi^{nonlin\text{-}poly}$ fits the overall trend while avoiding the local fluctuations shown in Fig. 3.4a. While such fluctuations in the surface Ψ^{nonlin} do not have a substantial impact on this specific case, they may have large impact on other case studies.

Table 3.3 MSE for predicting Parameter C_{LF} as a function of missing data

Available data set	Percent available data	Mean square error Ψ^{lin}	Mean square error Ψ^{nonlin}	Mean square error $\Psi^{nonlin\text{-}poly}$
1	100	0.034	0.028	0.028
2	70	0.034	0.029	0.029
3	70	0.035	0.031	0.031
4	70	0.034	0.030	0.031
5	70	0.035	0.029	0.029
6	50	0.035	0.030	0.034
7	50	0.036	0.032	0.032
8	50	0.034	0.029	0.029
9	50	0.034	0.030	0.031
10	50	0.039	0.033	0.039

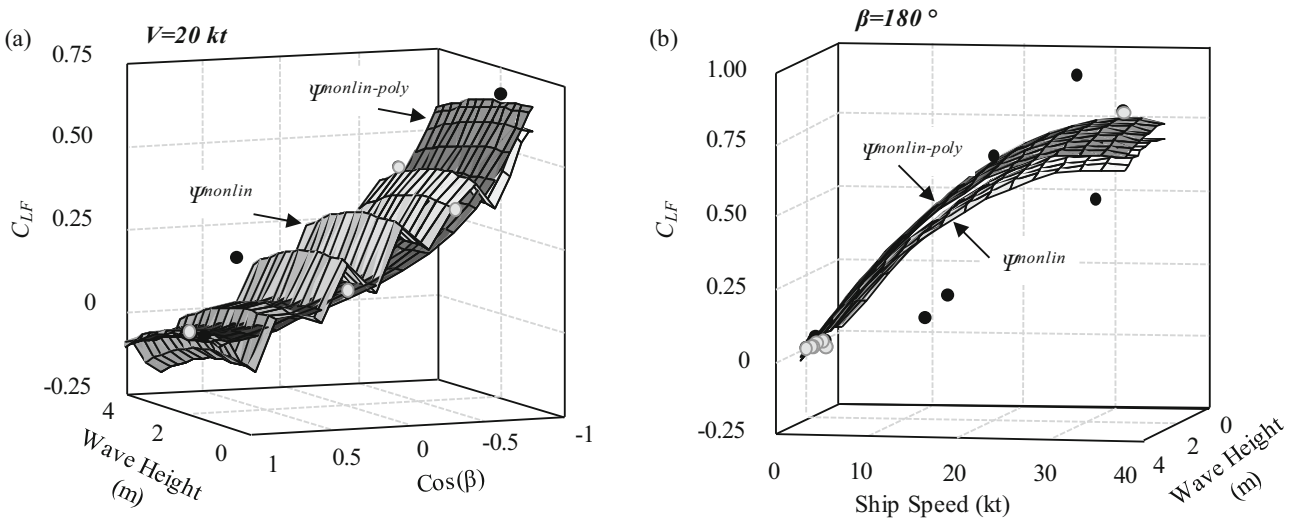


Fig. 3.4 Available data points for the response parameter C_{LF} (shown as *black* points), missing data points (shown in *light grey*), and the fitted surfaces for Ψ^{lin}, Ψ^{nonlin}, and $\Psi^{nonlin\text{-}poly}$ for available data set 6; (**a**) shows the variation of C_{LF} with wave height and heading angle for a ship speed of 20 kts and (**b**) shows the variation of C_{LF} with ship speed and wave height for a heading angle of 180°

3.5 Conclusions

Theoretically-based nonlinear prediction surfaces for the response of naval vessels under varying operational conditions are proposed and evaluated with respect to the structural response of the HSV-2 Swift. The nonlinear prediction surfaces are based on the theoretical relationships between the operational condition given in terms of the wave height, ship speed, and heading angle, and the ship response in terms of the vertical bending moment as derived from linear strip theory. Two theoretically-based nonlinear prediction surfaces are proposed and compared with linear prediction surfaces; both outperform linear surfaces in estimating low frequency response characteristics. For high frequency characteristics, the nonlinear and linear surfaces are comparable in their performance with linear surfaces occasionally outperforming the nonlinear ones. This is due to the fact that the nonlinear surfaces were developed based on the theoretical relationships of wave bending moments, which are governed by the low frequencies. However, slam impacts typically govern the high frequency response. The development of a separate surface for high frequency parameters which would focus on impact and slam loads may provide further enhancement to the prediction of the overall response. Additionally, the accuracy of the recorded SHM data is not included in the proposed methodology, which may contribute to the error in the developed surfaces.

Acknowledgments The support by grants from (a) the National Science Foundation (NSF) Award CMMI-1537926, (b) the U.S. Office of Naval Research (ONR) Awards N00014-08-1-0188, N00014-12-1-0023, and N00014-16-1-2299, and (c) the National Aeronautics and Space Administration (NASA) Award NNX10AJ20G is gratefully acknowledged. The opinions presented in this paper are those of the authors and do not necessarily reflect the views of the sponsoring organizations.

References

1. Lynch, J.P., Loh, K.J.A.: Summary review of wireless sensors and sensor networks for structural health monitoring. Shock Vib. Digest. **38**(2), 91–130 (2006)
2. Zhu, J., Collette, M.A.: Bayesian approach for shipboard lifetime wave load spectrum updating. Struct. Infrastruct. Eng. **13**(2), 298–312 (2016)
3. Kurata, M., Kim, J.H., Lynch, J.P., Law, K.H. and Salvino, L.W.: 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, pp. 741–750. American Society of Mechanical Engineers, New York (2010)
4. Zhu, B., Frangopol, D.M.: Reliability assessment of ship structures using Bayesian updating. Eng. Struct. **56**, 1836–1847 (2013)
5. Frangopol, D.M., Strauss, A., Kim, S.: Bridge reliability assessment based on monitoring. J. Bridg. Eng. **13**(3), 258–270 (2008)
6. Hughes, O.F.: Ship Structural Design: A Rationally-Based, Computer-Aided, Optimization Approach. Wiley-Interscience, Jersy City (1983)
7. Sikora, J.P., Dinsenbacher, A., Beach, J.E.: A Method for estimating lifetime loads and fatigue lives for swath and conventional monohull ships. Nav. Eng. J. **95**(3), 63–85 (1983)
8. Kim, S., Frangopol, D.M.: Cost-effective lifetime structural health monitoring based on availability. J. Struct. Eng. **137**(1), 22–33 (2010)
9. Iphar, C., Napoli, A. and Cyril, R.: Data quality assessment for maritime situation awareness. In: ISSDQ 2015-The 9th International Symposium on Spatial Data Quality, vol. 2, pp. 291–296. (2015).
10. Zhu, J.: Life cycle fatigue management for high-speed vessel using Bayesian updating approaches. Ph.D. thesis, Dept. of Naval Architecture and Marine Engineering, University of Michigan, Michigan, USA (2014)
11. Mondoro, A., Soliman, M., Frangopol, D.M.: Prediction of structural response of naval vessels based on available structural health monitoring data. Ocean Eng. **125**, 295–307 (2016)
12. Pierson Jr., W.J., Moskowitz, L.: A proposed spectral form for fully developed wind seas based on the similarity theory of SA Kitaigorodskii. J. Geophys. Res. **69**(24), 5181–5190 (1963)
13. Komen, G.J., Hasselmann, K., Hasselmann, K.: On the existence of a fully developed wind-sea spectrum. J. Phys. Oceanogr. **14**(8), 1271–1285 (1984)
14. Jensen, J.J., Mansour, A.E.: Estimation of ship long-term wave-induced bending moment using closed-form expressions. R. Inst. Nav. Archit. Trans. Part A. Int. J. Marit. Eng. 41–55 (2002)
15. Jensen, J.J., Mansour, A.E., Olsen, A.S.: Estimation of ship motions using closed-form expressions. Ocean Eng. **31**(1), 61–85 (2004)
16. Brady, T., Bachman, R., Donnelly, M., Griggs, D.: HSV-2 Swift Instrumentation and Technical Trials Plan. Naval Surface Warfare Center, Carderock Division (NSWCCD), West Bethesda (2004)

Chapter 4
A Case Study in Predictive Modeling Beyond the Calibration Domain

Philip Graybill, Eyob Tarekegn, Ian Tomkinson, Kendra Van Buren, François Hemez, and Scott Cogan

Abstract While numerical modeling is an important tool in many areas of engineering, caution must be exercised when developing and applying these models. This is especially true when models are developed under calibration conditions, which is referred to herein as the *calibration domain*, and applied to predict (or forecast) outcomes under a different set of conditions, which is referred to as the *forecasting domain*. This work discusses a case study of predictive capability of a simple model away from its calibration domain. The application is to predict the payload that a quadcopter is able to lift. Model development is supported by two calibration experiments. The first experiment measures displacements obtained by attaching masses to various springs; it is used to develop a model that predicts displacement as a function of weight. The second experiment measures displacements resulting from spinning propeller blades of various dimensions; it is used to develop a model that predicts displacement as a function of blade diameter and revolutions-per-minute. Both models are combined to predict the payload that a quadcopter can lift, which represents an extrapolated forecast because conditions of the quadcopter differ from those under which the models are calibrated. Finally the quadcopter is tested experimentally to assess the predictive accuracy of the model. This application illustrates a preliminary thought process to ultimately determine how models developed in calibration domains perform in forecasting domains. *(Approved for unlimited, public release, LA-UR-16-24484.)*

Keywords Calibration • Extrapolation • Uncertainity quantification

4.1 Introduction

Computational modeling is an essential technique in many areas of physics and engineering [1]. From the simulation of automobiles, to structures, to electronic circuits, models provide a means to estimate reality, shorten the design-prototype-test cycle and increase competitive advantage for product development. Models are especially useful to predict performance when testing an engineered system in its final configuration is not possible or practical. For example, the Neutral Buoyancy Laboratory at the National Aeronautics and Space Administration's Johnson Space Center simulates the weightlessness of

P. Graybill
Department of Electrical Engineering, Penn State University, 111H Electrical Engineering West, University Park, PA 16802, USA
e-mail: ppg5033@psu.edu

E. Tarekegn
Department of Physics and Astronomy, Clemson University, 820 College Avenue, Clemson, SC 29631, USA
e-mail: etareke@clemson.edu

I. Tomkinson
Department of Mechanical Engineering, Rice University, 23 Sunset Boulevard, Houston, TX 77005, USA
e-mail: ian.k.t1@gmail.com

K. Van Buren
"X" Computational Physics Division, Los Alamos National Laboratory, XCP-8, Mail Stop F644, Los Alamos, NM 87545, USA
e-mail: klvan@lanl.gov

F. Hemez (✉)
"X" Theoretical Design Division, Los Alamos National Laboratory, XTD-IDA, Mail Stop T087, Los Alamos, NM 87545, USA
e-mail: hemez@lanl.gov

S. Cogan
FEMSTO-ST, Département de Mécanique Appliquée, Université de Franche-Comté, 24 Chemin de l'Epitaphe, Besançon 25000, France
e-mail: scott.cogan@univ-fcomte.fr

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_4

space by testing equipment under water [2]. In this example, and many others like it, the predicted performance of a system in its configuration of interest is evaluated by combining testing and modeling conducted under conditions different than those of the final application. Conditions for which a model is developed and calibrated are herein referred to as those of the *calibration* domain to distinguish them from conditions of the final application, which are referred to as the *forecasting* domain [3–5].

Ideally a model is exercised under conditions that closely mimic—or cautiously interpolate or extrapolate—the conditions under which it was developed. Model development activities often include the calibration to experimental data. Occasionally the model must also be able to forecast outcomes of settings that are far from those used for calibration. While this work addresses the question of whether this strategy for model development is sound through a simple application, our ultimate goal is to propose a methodology to quantify the accuracy and robustness of simulation models in the broad context of computational physics and engineering [6–8].

Our application is to predict the payload capacity of a quadcopter. The calibration domain is comprised of two small-scale experiments performed to collect data that are the basis for model development. The first experiment measures deformations resulting from attaching masses to various springs to develop a model of displacement as a function of weight. The second experiment measures displacements resulting from spinning various propeller blades to develop a model of displacement as a function of rotation speed. The functions developed from these two small-scale experiments are combined to develop a model that predicts the payload lifting capacity of the quadcopter as a function of propeller size and speed. Predictions are obtained under conditions similar to, as well as different from, those of the calibration domain. The final step is to operate the quadcopter to measure how much payload it can lift and assess the accuracy of our predictions. Even though the application is simple, the observations made are sobering.

This application, while relatively simple, stresses the dangers of "blindly" trusting calibration as a way to establish the predictive accuracy of a model. Our representation of quadcopter lifting capability could be said to be naïve since it is mathematical in nature and little is done to account for the underlying physics as one steps outside the calibration domain. If mechanics or physics in the forecasting domain are poorly understood or significantly different from those of the calibration domain, then collecting more data for calibration is ineffective. Likewise calibrating a model might be a doomed strategy if the ultimate goal is to establish the accuracy of its predictions in a forecasting regime. Our quadcopter application illustrates these potential pitfalls in a simple, yet rigorous, manner.

4.2 Testing of a Mass-Spring System to Develop a Model that Predicts the Lifted Weight

The first experiment collects data to determine a relationship between force and displacement. We also generate a model that predicts the stiffness of a spring as a function of its physical parameters. The performance of this model is then assessed for springs "inside" and "outside" of the calibration domain. This application illustrates how rapidly the accuracy of a simple model can deteriorate as one generates a prediction away from conditions under which the model has been calibrated.

Figure 4.1 shows that the experimental setup is an Aluminum frame, supported by steel L-brackets, to which various combinations of masses and springs are attached. The ruler to the left of the mass is used to measure the displacement. Measurements of four different springs with varying physical parameters, namely, wire diameter (d), coil diameter (D) and number of coils (N), are collected. In order to exercise each spring over a large range of deflections, different sets of masses are attached. The smallest mass for each spring is selected such that the ruler can be used to measure a significant enough deflection. Likewise the largest mass for each spring is selected such that the resulting displacement does not exceed the clearance provided by the support frame. Each spring is tested three times by three experimenters, resulting in nine trials per spring with each trial consisting of five mass-deflection measurements. This procedure provides replication, which is essential to estimate the overall experimental variability.

For each spring considered, mass is converted to force ($F = m \cdot g$) using the gravitational constant $g = 9.81$ m/s^2. The force (F) versus displacement (x) data display strong linear trends, which makes it possible to best-fit simple relationships of the form:

$$F = F_0 + k \cdot x \tag{4.1}$$

for each spring, where (F_0; k) are the regression coefficients of the formula that can be calibrated using the deflections, x, observed in the experiment. Figure 4.2 is an example of force-deflection data collected where the measurements are indicated by cross symbols and the solid line suggests the regression model. The figure indicates, first, the excellent level of experimental repeatability and, second, the adequate goodness-of-fit of Eq. (4.1).

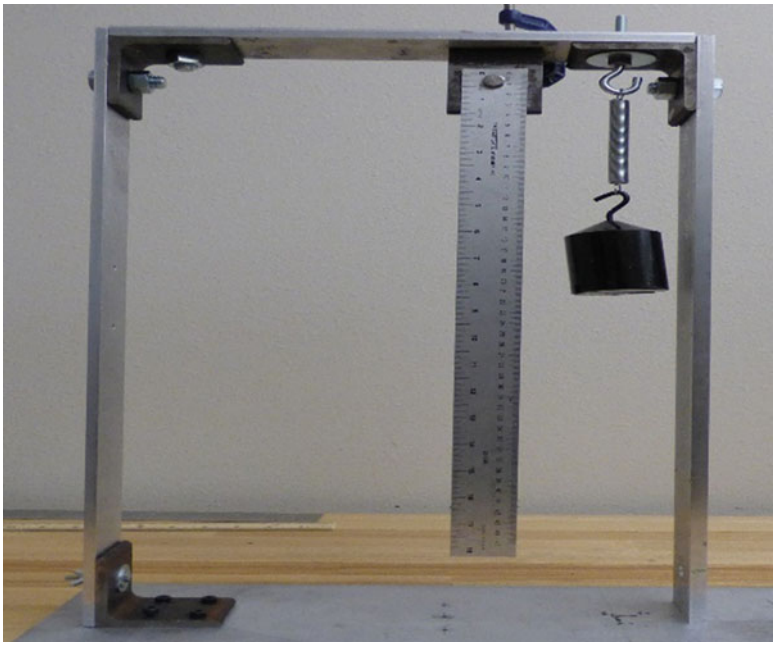


Fig. 4.1 Mass-spring test setup of the first experiment

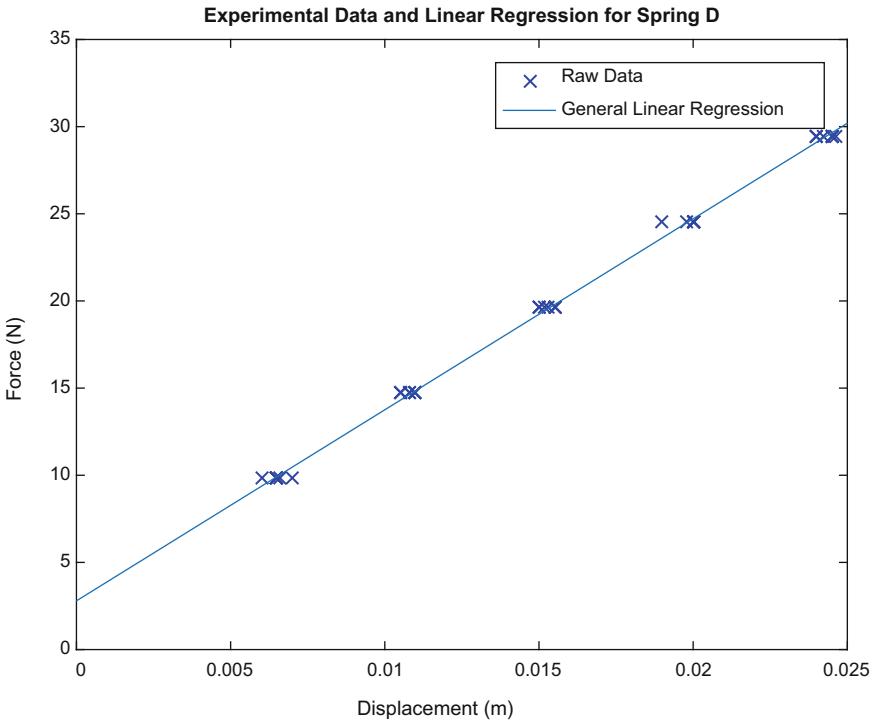


Fig. 4.2 Measurements of forces versus displacements for spring D

In addition to developing force-displacement models, the measurements are used to calibrate coefficients of a commonly used formula [9] that relates physical properties of a spring to its stiffness:

$$k = \alpha_K \cdot \frac{G \cdot d^p}{D^q \cdot N^r} \tag{4.2}$$

Table 4.1 Physical properties of calibration springs (A, B, C, D) and test springs (E, F)

Property	Range for springs (A, B, C, D)	Spring E	Spring F
Modulus of rigidity, G (GPa)	79.5	79.5	79.5
Wire diameter, d (mm)	1.12–1.58	1.30	0.45
Coil diameter, D (mm)	8.37–20.18	11.06	5.90
Number of coils, N (no unit)	36–108	98	51

Table 4.2 Predictions and measurements of stiffness coefficients for springs (E, F)

Property	Spring E	Spring F
Predicted stiffness using Eq. (4.2), k (N/m)	209	65.0
Measured stiffness using Eq. (4.1), k (N/m)	229	45.8
Percent error	−8.7%	41.9%

where G is the modulus of rigidity for the spring material (Pa), d is the wire diameter (m), D is the coil diameter (m), N is the number of coils (no unit) and α_K is a scaling factor (no unit if $p = q + 1$, otherwise units of m^{q+1-p}). The standard value of modulus of rigidity, $G = 79.5$ GPa, is used for all springs tested. Calibration yields the triplet of exponents $(p; q; r) = (4.5; 3.9; 0.8)$. They compare well to nominal values $(p_0; q_0; r_0) = (4.0; 3.0; 1.0)$ suggested in the literature. The scaling factor is $\alpha_K = 0.050\ \mathrm{m}^{0.4}$, which is a value that deviates significantly from the nominal value of 0.125. Such a deviation is not surprising since the two quantities exhibit different physical units.

As a case study for comparing model performance "inside" and "outside" of the calibration domain, two springs (labeled E and F) are selected with the same modulus of rigidity as the calibration springs but different combinations of $(d; D; N)$ properties. Table 4.1 defines these properties for the calibration springs (A, B, C, D) and the two springs (E, F) not used to calibrate Eq. (4.2). It can be observed from the table that spring E has $(d; D; N)$ properties that are well within the calibration domain because they fall within the range of values specified for springs A, B, C, and D in Table 4.1. Spring F has a value for N (number of coils) within the calibration domain but d and D (wire and coil diameters) fall outside of the calibration domain. Predicting the stiffness coefficient of spring F with Eq. (4.2), therefore, represents an extrapolated forecast of the model.

Table 4.2 compares the predictions of stiffness coefficients for springs (E, F), obtained using Eq. (4.2), to the experimentally determined values, obtained by applying the same measurement protocol as previously described for calibration springs (A, B, C, D). While a level of 8.7% error for spring E is deemed appropriate, predicting the stiffness coefficient of spring F with 41.9% error surely is not. This example, while simplistic, illustrates the caution that should be exercised when obtaining predictions in an extrapolative regime, that is, outside the domain where the model has been developed and calibrated.

4.3 Testing of a Single Propeller to Develop a Model that Predicts the Lifting Force

The second experiment investigates the lifting force of a single propeller. Using the relationship developed in the first experiment ($F = F_0 + k \cdot x$) and making some assumptions about the experimental setup, a model is developed for the lifting force generated by the propeller as a function of blade diameter and applied rotation-per-minute (RPM) speed.

The test apparatus illustrated in Fig. 4.3 is used to collect experimental data for different combinations of propeller blades and springs. It consists of a propeller motor on a horizontally sliding carriage, which is connected to a stationary point by a spring. The propeller speed is controlled with a throttle and measured by a tachometer. As the propeller speed increases, the carriage moves forward due to the lifting force generated by the spinning propeller, hence, stretching the spring. A ruler attached to the carriage provides a means of measuring carriage displacement (or change in spring elongation).

Four propellers with different blade diameters (25.4, 27.9, 30.5 and 33.0 cm) are each tested on all four calibration springs (A, B, C, D). A single trial to measure the displacement induced in the spring consists of gradually increasing and decreasing the RPM speed over a period of 40–60 s. Data are obtained by photographing the ruler displacement and tachometer in the same frame at multiple times throughout the trial. The procedure provides the pairs of (spring elongation; RPM speed) points needed to develop the model. Three trials are performed for each propeller-spring combination to estimate the experimental uncertainty, resulting in a total of 48 tests.

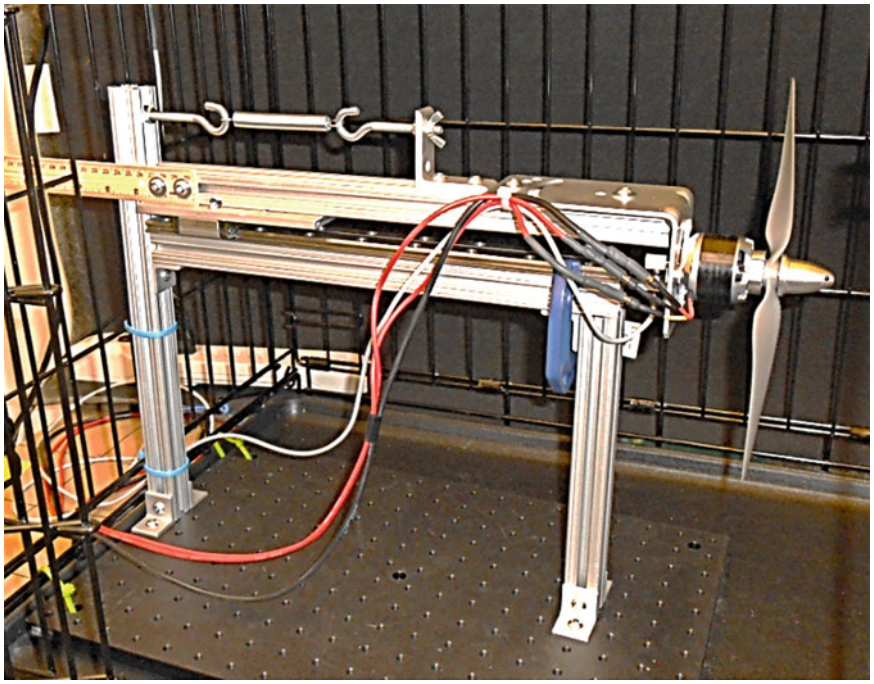


Fig. 4.3 Propeller test setup of the second experiment

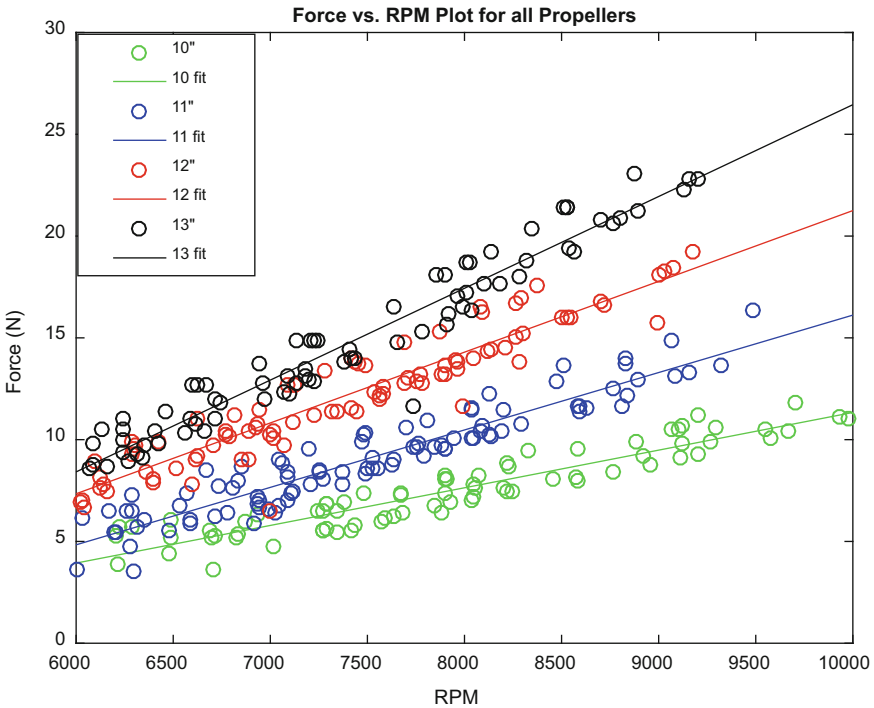


Fig. 4.4 Measurements of lifting forces versus RPM speeds for four propellers

Figure 4.4 depicts a sample of the data collected. The figure shows raw measurements (circle symbols) and predictions (solid lines) of the model shown in Eq. (4.3). Displacements are converted to forces using Eq. (4.1) for each calibration spring. Mapping all measurements from displacement to force allows data collected with different springs to be combined into a single data set for each propeller. The analysis is restricted to the range of propeller speeds from 6000 to 10,000 RPM because the data exhibit, for the most part, a bilinear trend. Figure 4.4 also suggests an acceptable level of experimental variability.

Using the data shown in Fig. 4.4, a predictive model for the lifting force of a single propeller is developed as a function of propeller diameter and speed. After evaluating the performance of several model forms (polynomials, exponentials and other mathematical functions), a polynomial that includes first-degree terms in RPM speed and second-degree contributions in propeller diameter is selected:

$$F_{Lift}\left(V_{RPM}; D_P\right) = \beta_0 + \beta_1 \cdot V_{RPM} + \beta_2 \cdot D_P + \beta_3 \cdot V_{RPM} \cdot D_P + \beta_4 \cdot D_P^2 \tag{4.3}$$

where the symbols V_{RPM} and D_P denote the RPM speed (m/s) and propeller diameter (m). The resulting lifting force is defined in Newtons. Equation (4.3) can be exercised to predict the lifting capacity, which is the force F_{Lift} in the left-hand side, as a function of propeller diameter and RPM speed.

Cross-validation is implemented to estimate the uncertainty of regression coefficients (β_0; β_1; β_2; β_3; β_4). Our experimental data are split into two sets where 70% of measurements define the training set used to calibrate the β_k-coefficients. The other 30% of data are kept separate for cross-validation, which assesses the accuracy of model predictions for combinations of control parameters (V_{RPM}; D_P) that have not been used to regress the β_k-coefficients. This entire procedure is repeated 100,000 times by randomly selecting different sets of training and cross-validation data.

This analysis yields, not a single set of regression coefficients (β_0; β_1; β_2; β_3; β_4), but an entire population that can be statistically sampled to account for experimental variability and, to some extent, model-form uncertainty when predicting the lifting force (F_{Lift}) with Eq. (4.3). A distribution of prediction errors is obtained from the 100,000 sets of cross-validation data. The absence of a systematic bias is inferred from the fact that the distribution is more-or-less centered at zero. In addition most of the probability mass is within $\pm 15\%$ error, which is deemed reasonable to accept the model as "valid."

4.4 Assessment of the Quadcopter Lifting Capacity Model in the Forecasting Regime

The final model, which predicts the payload-lifting capability of the quadcopter, is defined as:

$$M_{Total} = M_Q + M_P = \frac{4\ (Propellers)}{9.81\ \ (m \cdot s^{-2})} \cdot F_{Lift}\left(V_{RPM}; D_P\right) \tag{4.4}$$

where the total mass (M_{Total}) combines the quadcopter mass ($M_Q = 3.66$ kg) and its payload capacity (M_P), which we wish to estimate for a given combination of control parameters (V_{RPM}; D_P). The lifting force is obtained from Eq. (4.3). The factor of four in Eq. (4.4) accounts for the fact that the quadcopter is equipped with four nominally identical motors, and the constant of 9.81 m/s^2 converts force to mass.

This model makes several crucial assumptions such as, for example, postulating that each one of the four propellers generates an identical lifting force and any difference of the motor used in the second experiment and those used for the quadcopter can be neglected. Another assumption is that aerodynamic flow conditions generated when operating the test apparatus of Fig. 4.3 to measure a single propeller are similar to those encountered when the quadcopter is hovered.

Equations (4.3) and (4.4) are exercised to generate a distribution of predictions, as opposed to a single-point prediction, for any arbitrary combination of propeller speed (V_{RPM}) and diameter (D_P). Figure 4.5 shows one such distribution obtained by propagating the uncertainty of regression coefficients (β_0; β_1; β_2; β_3; β_4). For simplicity the remainder is illustrated with mean-value predictions only, such as $M_P = 4.72$ kg in Fig. 4.5, and prediction uncertainty is not considered in the discussion. The final step is to evaluate the predictive accuracy of the model by comparing its predictions to measurements obtained by flying the quadcopter.

The quadcopter shown in Fig. 4.6 is operated to measure its payload capacity as a function of propeller speed and blade diameter. The experimental procedure consists of attaching a known mass and gradually increasing the propeller speed until the quadcopter hovers steadily just below a given height threshold. The corresponding RPM level is recorded once the quadcopter has stabilized in this position. The procedure is replicated ten times using different experimenters to control the flight and record the RPM level. A range of RPM speeds that lift a given payload is generated from these experiments. At the same time the model, as defined in Eqs. (4.3) and (4.4), is used to predict the RPM speed needed to lift a given payload for comparison with the measurements. It is emphasized that these are true, "blind" forecasts obtained before the experiments are carried out.

Figure 4.7 summarizes the results of experimental testing and model forecasting of the quadcopter lifting capacity. The response surface shows model predictions over the calibration domain with training data indicated by turquoise crosses. Blue circles in the horizontal (M_P; D_P) plane are projections of the training data, which helps to visualize their coverage of the

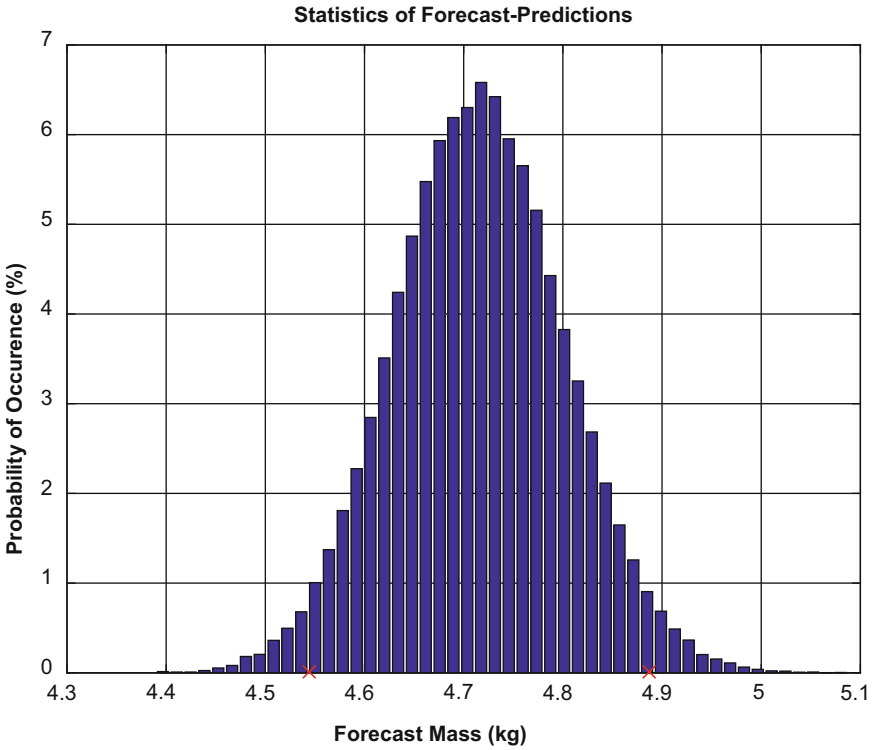


Fig. 4.5 Distribution of model predictions of quadcopter payload capacity (M_P)

Fig. 4.6 Quadcopter tested to assess the predictive accuracy of Eqs. (4.3) and (4.4)

calibration domain. Forecasts of the model are depicted with red crosses, and the corresponding measurements are shown with green crosses. Projections in the (M_P; D_P) plane show that these predictions and measurements extrapolate further and further away from the calibration domain. The experimental variability and modeling uncertainty are represented with multiple "x" symbols for the measurements, training points and forecasts in Fig. 4.7. The ranges that these replicates produce

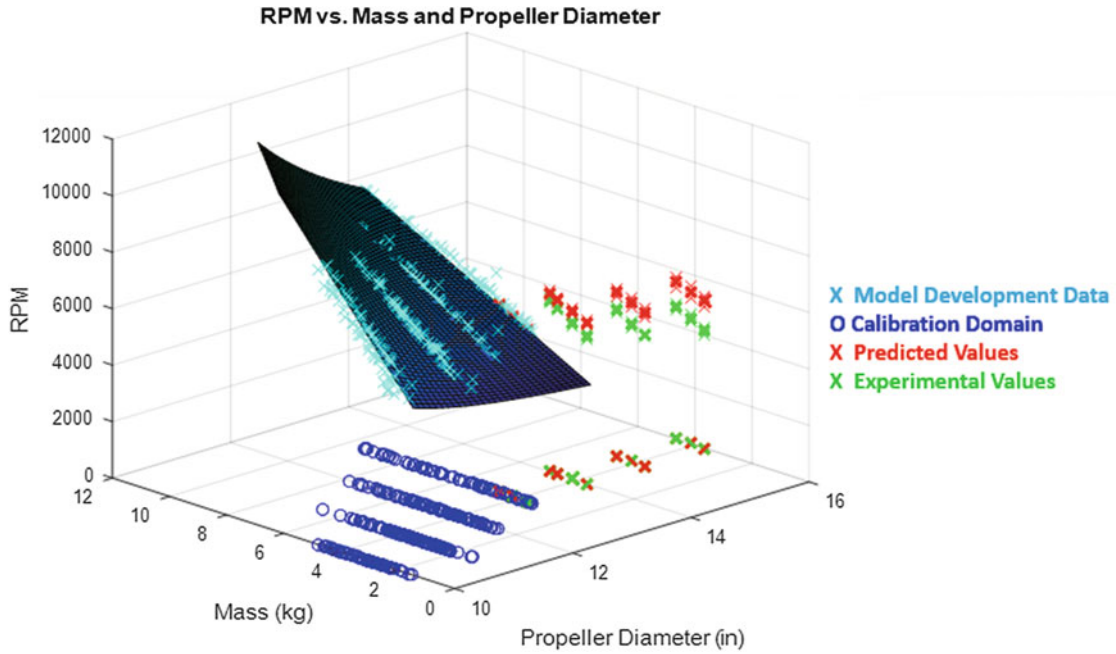


Fig. 4.7 Predictions and measurements of quadcopter payload lifting capacity (M_P)

Table 4.3 Ranges of predictions (column 3) and forecasting errors (column 4) of the payload capacity

Propeller diameter	Prediction regime	Range of predictions	Range of prediction errors
33.0 cm (13 in.)	Calibration	1.6%	½%-to-4%
35.6 cm (14 in.)	Extrapolation	3.0%	4%-to-9%
38.1 cm (15 in.)	Extrapolation	5.1%	12%-to-17%
40.6 cm (16 in.)	Extrapolation	7.5%	18%-to-22%

are small, which indicates that the trends of prediction accuracy discussed next are statistically significant in comparison to the experimental variability and modeling uncertainty.

It can be observed from Fig. 4.7 that measurements (green crosses) and predictions (red crosses) differ. Said differently, prediction errors increase as forecasts are performed further away from the calibration domain. For simplicity, Table 4.3 summarizes this trend as a function of propeller diameter (D_P) only. The third column lists the prediction uncertainty. It automatically grows as distance from the calibration domain increases, which is a benefit of our Monte Carlo procedure. The fourth column shows the ranges of prediction errors. While the formulae (Eqs. (4.3) and (4.4)) perform well at the edge of the calibration domain, which is 25.4 cm $\leq D_P \leq$ 33.0 cm, the model systematically over-predicts the measured RPM speeds. Results indicate that for this application, first, the accuracy of model predictions in the calibration domain is not representative of what might be achieved in the forecasting domain and, second, confidence in exercising the model should diminish as one extrapolates further from settings used to develop and calibrate it.

4.5 Conclusion

In this work we develop a model to predict the payload lifting capacity of a quadcopter. Two simple formulae are integrated to derive the model. The first one is a mass versus deflection function calibrated with experiments that measure deflections obtained by attaching masses to different springs. The second formula is a deflection versus rotational speed function calibrated by measuring the translational motion resulting from the lifting force generated when a single propeller is operated over a range of speeds. The model predicts the payload that the quadcopter should be able to lift when equipped with propeller blades of a given diameter and operated at a given speed. Our analysis quantifies the experimental variability and translates it to modeling uncertainty that can be propagated using Monte Carlo sampling.

While the model performs well in the calibration domain, it is also observed that its forecasts rapidly lose accuracy as they are extrapolated further away from settings used to calibrate it. Accuracy established in a calibration regime might not be representative of what can be achieved for forecasting. Collecting more data in the (same) calibration domain would not have improved our forecasting accuracy. Calibration is an ineffective strategy for model development if, firstly, the mechanics or physics in the forecasting domain are significantly different from those of the calibration domain and, secondly, the ultimate goal is to establish the accuracy of predictions in a forecasting regime.

Clearly not every computational model will exhibit the same degradation of accuracy away from the calibration domain. Whether or not this behavior is observed is application-specific. We nevertheless stress that the techniques employed in this case study can be broadly applied to predictive modeling. They include estimating experimental variability by performing as many replicate experiments as practically feasible, translating this variability to modeling uncertainty, and implementing Monte Carlo sampling or optimization to obtain a range of predictions as opposed to a single prediction. This methodology makes it possible to estimate prediction uncertainty both inside and outside of the calibration domain.

The single most detrimental assumption in our analysis might be that the physics remain unchanged when moving from the calibration domain to the region of forecasting. It does not seem to be the case here. For example, the height of the quadcopter (or distance between its propeller blades and ground) is shorter than the length of the room where the single-propeller testing apparatus is operated (or distance between the single propeller and wall). This offers the potential to produce different aerodynamic flow conditions, which invalidates the assumption that single-blade measurements are representative of flows encountered when operating the quadcopter. Flying the quadcopter in open air, as opposed to the indoor conditions of a laboratory, would be another level of extrapolation that could invalidate our modeling effort. Lateral wind gusts, precipitation and varying air densities, for example, are additional confounding factors that would need to be accounted for. Future work will investigate the possibility to bound these unknown phenomena and provide forecasts that are robust, as much as possible, to conditions that differ in the forecasting domain from what they might be in the calibration domain.

Acknowledgments The authors acknowledge the support of the 2016 Dynamics Summer School at Los Alamos National Laboratory (LANL), and the Advanced Scientific Computing program and Advanced Certification Campaign. Los Alamos National Security, L.L.C., operates LANL under contract DE-AC52-06NA25396 on behalf of the National Nuclear Security Administration of the U.S. Department of Energy.

References

1. Oden, J.T., Belytschko, T., Fish, J., Hughes, T.J.R., Johnson, C., Keyes, D., Laub, A., Petzold, L., Srolovitz, D., Yip, S.: Revolutionizing engineering science through simulation, National Science Foundation Blue Ribbon Panel on Simulation-Based Engineering (2006)
2. Strauss, S., Krog, R.L., Feiveson, A.H.: Extravehicular mobility unit training and astronaut injuries. Aviat. Space Environ. Med. **76**(5), 469–474 (2005)
3. Hemez, F.M.: Uncertainty quantification and the verification and validation of computational models. In: Damage Prognosis for Aerospace, Civil and Mechanical Systems, pp. 201–219. Wiley, London, Chichester (2005)
4. Williams, B.J., Stull, C.J.: Towards validating computational models in extrapolation regimes, Technical report LA-UR-11-3415, Los Alamos National Laboratory, Los Alamos, New Mexico (2011)
5. Van Buren, K.L., Cogan, S., Hemez, F.M.: A robust approach to quantifying forecasting uncertainty using proxy simulations. In: 34th International Modal Analysis Conference, Orlando (2016)
6. Ben-Haim, Y., Hemez, F.M.: Robustness, fidelity and prediction-looseness of models. Phys. Trans. R. Soc. Proc. A. (2011). doi:10.1098/rspa.2011.0050
7. Mollineaux, M.G., Van Buren, K.L., Hemez, F.M., Atamturktur, S.: Simulating the dynamics of wind turbine blades: part I, model development and verification. Wind Energy. **16**, 694–710 (2013)
8. Van Buren, K.L., Mollineaux, M.G., Hemez, F.M., Atamturktur, S.: Simulating the dynamics of wind turbine blades: part II, model validation and uncertainty quantification. Wind Energy. **16**, 741–758 (2013)
9. Marrs, J.: Machine Designer's Reference, p. 444. Industrial Press, New York (2012)

Chapter 5
A Brief Overview of Code and Solution Verification in Numerical Simulation

François Hemez

Abstract This manuscript is a brief overview of code and calculation verification in computational physics and engineering. Verification is an essential technology to assess the quality of discrete solutions obtained by running simulation codes that solve systems of ordinary or partial differential equations, such as the finite element and finite volume methods. Code verification assesses the extent to which a numerical method is implemented correctly, that is, without any programming mistake (or "bug") that would adversely affect the quality of computations. The centerpiece of code verification is the formulation of verification test problems that admit exact or manufactured solutions and which are used for comparison with approximate solutions obtained from the simulation software. Solution verification assesses the extent to which the discretization (in time, space, energy, modal basis, etc.) implemented to solve governing equations provides a sufficiently small level of truncation error. The keystone of solution verification is the practice of mesh refinement from which estimates of the (spatial) order of accuracy of the numerical method can be estimated. It is also possible to derive bounds of truncation error produced in the calculation. The discussion is presented in the context of Peter Lax's 1954 groundbreaking work on the convergence of discrete solutions. It is illustrated with a simple example of one-dimensional advection solver. *(Publication approved for unlimited, public release, LA-UR-16-24553, Unclassified.)*

Keywords Code verification • Solution verification • Solution convergence • Truncation error

5.1 Introduction

In computational engineering and physics, ordinary or partial differential equations that govern the evolution of state variables, such as pressure, temperature, velocity, or displacement, are discretized for implementation and resolution on finite-digitized arithmetic computer. The challenge of verification is to assess the extent to which approximate solutions of the discretized equations converge to the exact solution of the continuous equations. In addition to assessing how "closely" discrete solutions estimate the continuous solution, it is often important to verify that the observed rate-of-convergence matches the theoretical order of accuracy of the numerical method or solver. Simply speaking, verification is the first "V" of Verification and Validation (V&V) for predictions of numerical simulations [1].

Code and calculation (or solution) verification is defined as a *scientifically rigorous and quantitative process for assessing the mathematical consistency between continuum and discrete variants of partial differential equations* [2]. Verification involves comparing numerical solutions obtained on successively refined meshes (or grids) to an exact solution or "reference." The main difficulty is that the exact solution of continuous equations is not always known and available to define this reference. These exact solutions can be derived analytically only in a few cases that feature "smooth" dynamics, linearized operators, simple geometries, or combinations of the above. Well-known examples include the single degree-of-freedom harmonic oscillator, linear vibration of simple structural components, Poiseuille flow of 2D incompressible fluids, and the 1D Riemann problem for ideal gases. These are problems that have been extensively studied and for which exact or highly-accurate solutions of the continuous equations can be obtained. Exact solutions can also be produced with the technique of manufactured solutions [3]. The overall number of problems that offer closed-form solutions is, unfortunately, limited.

In general one talks of *code verification* when the solution of the continuous equations can be derived in closed form, which provides an exact reference to which the discrete solutions are compared. By "closed form," we mean either analytical formulae or simplified representations that can be accurately evaluated, e.g., reduction from a partial differential equation to an ordinary differential equation. If the system of equations and its initial and boundary conditions are complicated, then

F. Hemez (✉)
"X" Theoretical Design Division, Los Alamos National Laboratory, XTD-IDA, Mail Stop T087, Los Alamos, NM 87545, USA
e-mail: hemez@lanl.gov

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_5

a continuous solution cannot be derived analytically and one then talks of *solution verification*. In this latter case no exact reference is available to calculate the solution error due to discretization, which makes verification quite challenging.

The discussion presented herein offers a brief overview of technology applicable to code and solution verification alike. The only difference between the two cases is the availability (for code verification) or not (for solution verification) of a reference solution. The dominant paradigm for code and solution verification is that the numerical method provides discrete solutions that converge to the (possibly unknown) continuous solution with a specific rate-of-convergence (or order of accuracy). For example, a finite element analysis featuring quadratic shape functions should yield discrete solutions that converge with a rate of $p = 2$ when the mesh is refined. Common practice is to generate several discrete solutions from successively refined meshes, estimate an extrapolation of the continuous solution, verify the rate-of-convergence, and estimate bounds of numerical uncertainty. Discussion is restricted to well-established techniques and does not address recent developments that might challenge boundaries of the current state-of-the-practice. For a partial account of such developments see, for example, Refs. [4–10].

Successes of verification include the development of a formalism to study the convergence of discrete solutions for a wide range of applications, from linear, non-dissipative, elliptic equations to non-linear, shocked, hyperbolic equations. It means that the convergence, for example, of solutions for the resonant frequencies of a modal analysis can be studied with the same tools as the convergence of solutions for a blast wave propagating in ideal gases. The failures of verification are embodied by the many restrictions imposed on the way an analysis is typically carried out. These include restricting the studies to scalar quantities, further restricting them to spatial-only convergence, and not always accounting for how discrete solutions are constructed by the numerical methods and the specific properties that result.

5.2 The Consistency and Convergence of Modified Equations

Numerical methods, such as finite element and finite volume methods, solve conservation laws that balance the rate-of-change in time of state variables, gradient of their fluxes, and source terms. Without loss of generality these contributions can be written in a one-dimensional, Cartesian coordinate system as:

$$\frac{\partial y^{Exact}(x;t)}{\partial t} + \frac{\partial \mathrm{F}\left(y^{Exact}(x;t)\right)}{\partial x} = S(x;t), \tag{5.1}$$

where $y^{Exact}(x;\ t)$ is the exact solution, F(•) denotes the flux operator, and $S(x;\ t)$ is a source or forcing function that drives the dynamics of the system. These generic functions depend on space and time, labeled x and t, respectively. When solving systems of Eq. (5.1) with a numerical method, one seeks the best-possible approximation of the continuous, exact, and often unknown solution, y^{Exact}.

The numerical method discretizes the continuous Eq. (5.1) on a computational mesh to yield a discrete solution $y_k{}^n = y(k{\cdot}\Delta x;\ n{\cdot}\Delta t)$ where Δx and Δt are spatial and temporal resolutions. The approximation $y_k{}^n$ is obtained by solving a discretized system of equations that resembles, for example, something like:

$$\frac{y_k^{n+1} - y_k^n}{\Delta t} + \frac{F_{k+1/2}^n - F_{k-1/2}^n}{\Delta x} = S_k^n. \tag{5.2}$$

It is emphasized that Eq. (5.2) suggests a hypothetical discretization scheme (not necessarily the one used) where the temporal differentiation operator (∂•/∂t) is approximated by a forward Euler difference and the spatial differentiation operator (∂•/∂x) is approximated by central differences. The discrete values $y_k{}^n$, $F_k{}^n$ and $S_k{}^n$ can be obtained from finite difference, finite volume, or finite element approximations.

One observes that the discretized Eq. (5.2) seems "similar" to the original, continuous Eq. (5.1). The similarity, however, is misleading. Contrary to common belief the discrete solution $y_k{}^n$ does *not* approximate the continuous solution y^{Exact} of Eq. (5.1). Using the technique known as Modified Equation Analysis (MEA), see References [11, 12], it can be shown that the approximation $y_k{}^n$ converges to the solution of a *modified equation* that takes a form such as:

$$\frac{\partial y^{MEA}(x;t)}{\partial t} + \frac{\partial \mathrm{F}\left(y^{MEA}(x;t)\right)}{\partial x} = S(x;t) + \left(\frac{\partial^2 y^{MEA}(x;t)}{\partial t^2}\cdot\Delta t + \frac{\partial^3 y^{MEA}(x;t)}{\partial t^3}\cdot\Delta x^2 + \ldots\right). \tag{5.3}$$

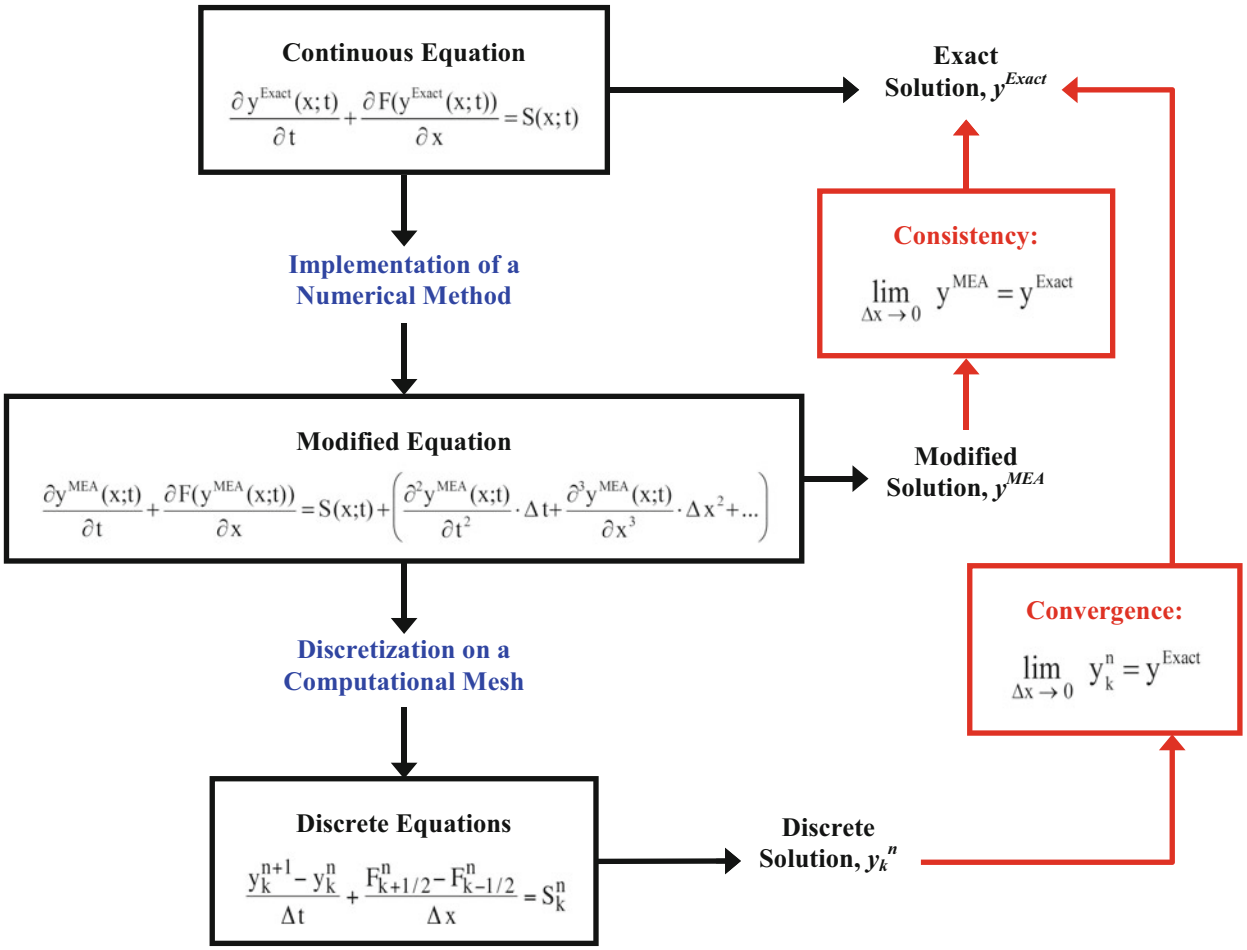


Fig. 5.1 The concepts of consistency and convergence of numerical solutions

Note that this example is conceptual and the correct form of the modified equation depends on properties of the original Eq. (5.1) and numerical method (5.2) implemented for its resolution. What is important to the discussion is that the continuous solution y^{MEA} of the modified Eq. (5.3) is different from y^{Exact} as long as the spatial and temporal resolutions remain finite (as long as $\Delta x \neq 0$, $\Delta t \neq 0$). The Lax equivalence theorem of Ref. [13] details conditions under which the discrete approximation y_k^n is *consistent* with, and *converges* to, the continuous solution y^{Exact}. Figure 5.1 illustrates these essential concepts.

Terms shown between parentheses in the right-hand side of the modified Eq. (5.3) represent an infinite series expansion that characterizes the *truncation error* of the numerical method or solver. Truncation is what explains the difference between the continuous solutions y^{MEA} and y^{Exact}. This is why it is essential to establish that the solution y^{MEA} of the modified Eq. (5.3) is consistent with the exact solution y^{Exact} of the original Eq. (5.1), as indicated in Fig. 5.1. Likewise one needs to understand the behavior of truncation error, that is, the extent to which the discrete solution y_k^n converges to the exact solution y^{Exact}.

In addition to defining the concept of modified equation, Eq. (5.3) also illustrates the order of accuracy of a numerical method. The leading-order term of spatial truncation, for example, is proportional to Δx^2, which means that the solver is 2nd-order accurate. Performing a mesh refinement should yield a truncation error between discrete solutions y_k^n and the exact-but-unknown solution y^{Exact} that converges at a quadratic rate. Likewise Eq. (5.3) suggests that the solver is 1st-order accurate in time since the leading-order term of temporal truncation is proportional to Δt. The next section discusses this asymptotic convergence.

5.3 The Regime of Asymptotic Convergence of Discrete Solutions

Because truncation error is often the main mechanism by which discrete solutions $y_k{}^n$ differ from the exact solution y^{Exact}, understanding its behavior is key to assess the numerical performance of simulation software. Verifying the quality of discrete solutions hinges on the *regime of asymptotic convergence*.

Different choices of discretization variables, such as Δx or Δt, induce different types and magnitudes of error. Plotting the solution error $||y^{Exact} - y_k{}^n||$ as a function of mesh size Δx gives a conceptual illustration of the main three regimes of the discretization. By definition the asymptotic regime is the region where truncation dominates the overall production of numerical error. Figure 5.2 provides a simplified illustration of these regimes. They are color-coded such that *red* is the regime where discretization is inappropriate, *green* denotes the regime of asymptotic convergence, and *grey* is where round-off error accumulates.

Going from right (larger values of Δx) to left (smaller values of Δx) in Fig. 5.2, the first domain shown in *red* is where the choice of element or cell size is not even appropriate to solve the discrete equations. This would be, for example, the case when elements are too coarse to resolve important geometrical features of a contact condition between components, or a numerical stability criterion is violated. Although it could be argued that discrete solutions should not be computed in this regime, the fact is undeniable that meshes analyzed in practical situations are often significantly under-resolved.

The second region shown in *green* is where truncation dominates the overall error: it is the regime of asymptotic convergence. Because truncation error dominates, the accuracy of a discrete solution y_k can be improved simply by performing the calculation with a smaller element size. We have just described the basic principle of conducting a mesh or grid refinement study.

Our conceptual illustration assumes that truncation error within the regime of asymptotic convergence is dominated by a single effect proportional to Δx^p, which appears as a straight line of slope p on the log-log representation of Fig. 5.2. The functional form of a typical modified equation, suggested in Eq. (5.3), motivates this assumption. It is also the reason why the behavior of truncation error is usually studied by formulating a simple model such as:

$$\varepsilon(\Delta x) = \left\| y^{Exact} - y_k^n(\Delta x) \right\| \approx \beta \cdot \Delta x^p + \text{H.O.T.}, \tag{5.4}$$

where $\varepsilon(\Delta x)$ denotes the difference, estimated in the sense of a user-defined norm $||\bullet||$, between the exact solution y^{Exact} of the continuous Eq. (5.1) and the discrete solution $y_k{}^n(\Delta x)$ obtained by executing the simulation code with mesh or grid size Δx. The pre-factor β represents a (constant) regression coefficient. The exponent p characterizes the rate at which the solution error decreases as the level of resolution is increased, that is, as $\Delta x \rightarrow 0$. It should match the theoretical order of

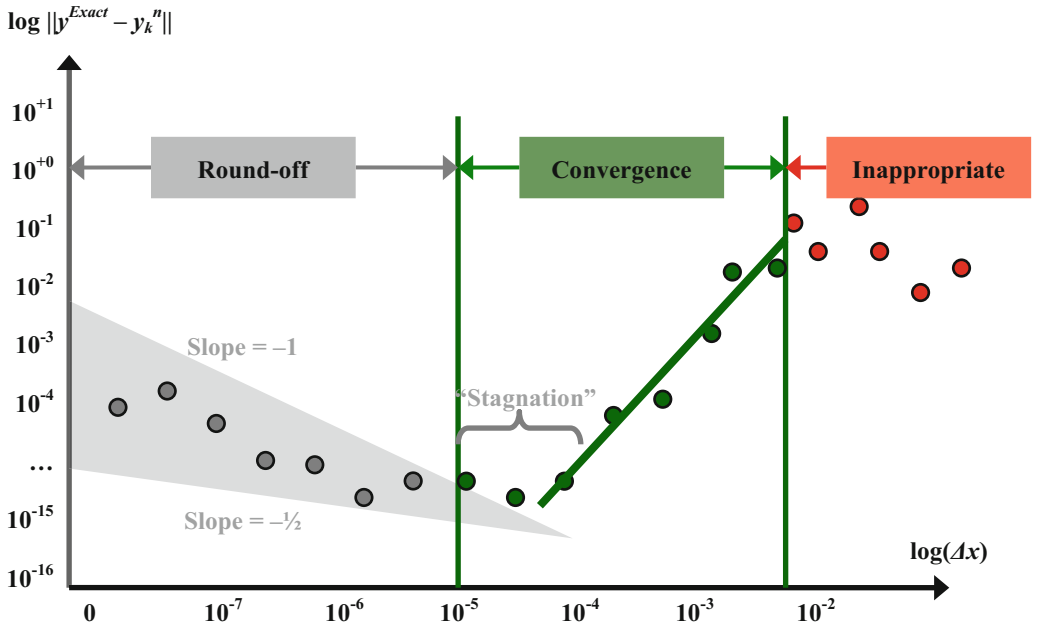


Fig. 5.2 The three regimes of solution error as a function of spatial discretization

accuracy of the numerical method. For example, a finite element calculation that uses linear shell elements should exhibit the rate of $p = 1$. Likewise the second-order accurate Gudonov scheme of a fluid dynamics solver should feature $p = 2$ for a laminar-like flow that does not develop a discontinuity (or shock).

The last region of Fig. 5.2 shown in *grey* is a limiting case for asymptotic convergence due to the fact that, with finite arithmetic implemented in a computer code, round-off errors eventually start to grow as $\Delta x \rightarrow 0$. Round-off could then accumulate to the point where it supplants truncation as the dominant mechanism that produces numerical error.

Understanding asymptotic convergence is important for two reasons. First, the formulae discussed in this manuscript, which provide the foundation of verification, are valid only within the regime of asymptotic convergence. Second, verifying that a discretization used in a calculation leads to a discrete solution in the asymptotic regime provides a strategy to reduce truncation error. If the error is too large for the application and needs to be reduced, performing the simulation with a smaller element or cell size automatically reduces it. This is, however, true only within the regime of asymptotic convergence.

5.4 State-of-the-Practice of Code and Solution Verification

This section summarizes the formalism and main equations used to study the convergence of discrete solutions. The discussion focuses on the state-of-the-practice in computational engineering and physics without paying tribute to some of the more advanced techniques. References [5, 9, 10] to list only a few, discuss issues that extend beyond the limited scope of this section.

The distinction needs to be emphasized between *code verification*, where the exact solution y^{Exact} of the continuous Eq. (5.1) is known analytically, or provided by an approximate yet highly accurate solution procedure, and *solution verification* where y^{Exact} is unknown. In the case of code verification, the error $\varepsilon(\Delta x)$ in the left-hand side of Eq. (5.4) can be computed given knowledge of the exact solution y^{Exact} and a discrete solution ${y_k}^n(\Delta x)$ obtained with the discretization size Δx. These exact and discrete solutions can be any scalar quantities, curves, or multi-dimensional fields. The unknowns of Eq. (5.4) are the parameters $(\beta; p)$. Two discrete solutions, one obtained with a "coarse" mesh of size Δx_C and another one obtained with a "fine" mesh of size Δx_F, where $\Delta x_C = R{\cdot}\Delta x_F$ (and $R > 1$), suffice to estimate β and p.

We now proceed to discuss the general case of solution verification where the difficulty is that the exact solution y^{Exact} of the continuous Eq. (5.1) is unknown. The challenge is that the procedure described in the previous paragraph cannot be implemented since the error $\varepsilon(\Delta x)$ cannot be explicitly computed. The starting point is to postulate an equation, referred to as an Ansatz model of numerical error, which describes how the error behaves in the regime of asymptotic convergence. The state-of-the-practice is to specialize Eq. (5.4) to scalar quantities such that is can be written as:

$$\varepsilon\left(\Delta x\right) = y^{Reference} - y_k^n\left(\Delta x\right) \approx \beta \cdot \Delta x^p, \tag{5.5}$$

where the unknown is the triplet of quantities $(y^{Reference}; \beta; p)$. The main difference with code verification defined in Eq. (5.4) is that the exact-but-unknown solution y^{Exact} is replaced by a reference $y^{Reference}$ that becomes a third unknown of the Ansatz model. Equation (5.5) also assumes that, first, the convergence of discrete solutions ${y_k}^n(\Delta x)$ towards the reference solution is monotonic and, second, the higher-order terms of the modified equation can be neglected. References [9, 10] address the case of non-monotonic convergence.

Because the formulation (5.5) features three unknowns, a minimum of three equations is required to solve it. These are provided by discrete solutions obtained from a coarse-mesh calculation (with element or cell size Δx_C), a medium-mesh calculation (Δx_M), and a fine-mesh calculation (Δx_F). These are written as:

$$\begin{aligned} y^{Reference} - y_k^n\left(\Delta x_C\right) &= \beta \cdot \Delta x_C^p \\ y^{Reference} - y_k^n\left(\Delta x_M\right) &= \beta \cdot \Delta x_M^p \\ y^{Reference} - y_k^n\left(\Delta x_F\right) &= \beta \cdot \Delta x_F^p \end{aligned}. \tag{5.6}$$

The system of Eq. (5.6) can be solved for the triplet of unknowns $(y^{Reference}; \beta; p)$ by combining the equations to eliminate two of the three unknowns. The value of the rate-of-convergence p, for example, is obtained by solving the following nonlinear equation:

$$p \cdot \log\left(R_{MF}\right) + \log\left(1 - R_{CM}^p\right) - \log\left(1 - R_{MF}^p\right) = \log\left(\frac{y_k^n\left(\Delta x_M\right) - y_k^n\left(\Delta x_C\right)}{y_k^n\left(\Delta x_F\right) - y_k^n\left(\Delta x_M\right)}\right), \tag{5.7}$$

where R_{CM} and R_{MF} are the successive mesh refinement ratios defined as:

$$R_{CM} = \frac{\Delta x_C}{\Delta x_M} \quad \text{and} \quad R_{MF} = \frac{\Delta x_M}{\Delta x_F}. \tag{5.8}$$

There is no closed-form solution to the nonlinear Eq. (5.7) when the refinement ratios R_{CM} and R_{MF} are different. The value of exponent p must be obtained through numerical optimization. With a constant mesh refinement ratio, which means that $R_{CM} = R_{MF} = R$, however, Eq. (5.7) is solved as:

$$p = \frac{\log\left(\dfrac{y_k^n(\Delta x_M) - y_k^n(\Delta x_C)}{y_k^n(\Delta x_F) - y_k^n(\Delta x_M)}\right)}{\log(R)}. \tag{5.9}$$

The next step is to extrapolate the discrete solutions $y_k{}^n(\Delta x_C)$, $y_k{}^n(\Delta x_M)$ and $y_k{}^n(\Delta x_F)$ to a reference $y^{Reference}$ that approximates the exact-but-unknown solution y^{Exact}. This is referred to as Richardson's extrapolation. Any combination of two meshes (coarse-medium or medium-fine) provides the same value for $y^{Reference}$:

$$y^{Reference} = y_k^n(\Delta x_F) + \frac{y_k^n(\Delta x_F) - y_k^n(\Delta x_M)}{R^p - 1} = y_k^n(\Delta x_M) + \frac{y_k^n(\Delta x_M) - y_k^n(\Delta x_C)}{R^p - 1}, \tag{5.10}$$

where the convergence rate p is the value obtained from Eq. (5.9). The regression coefficient β can be back-calculated from any one of Eq. (5.6):

$$\beta = \frac{y_k^n(\Delta x_C) - y_k^n(\Delta x_M)}{\Delta x_M^p \cdot (R^p - 1)} = \frac{y_k^n(\Delta x_M) - y_k^n(\Delta x_F)}{\Delta x_F^p \cdot (R^p - 1)} = \frac{y_k^n(\Delta x_C) - y_k^n(\Delta x_F)}{\Delta x_F^p \cdot (R^{2p} - 1)}. \tag{5.11}$$

In summary, Eqs. (5.7)–(5.11) provide expressions for the triplet of unknowns ($y^{Reference}$; β; p) in the case of either a variable or constant mesh refinement ratio. This formalism is applicable to code and solution verification alike, with the difference that the former case assumes knowledge of the exact solution y^{Exact}, which reduces the unknowns of the Ansatz model to the pair (β; p). Clearly this analysis technique is appropriate only to the extent that assumptions upon which it relies are satisfied. There are four essential assumptions summarized to conclude the discussion. (1) The mesh sizes used (Δx_C, Δx_M and Δx_F) provide discrete solutions located in the regime of asymptotic convergence. (2) The analysis is restricted to scalar quantities. (3) Because it is unknown, the exact solution y^{Exact} of the continuous equations is replaced by a to-be-estimated reference $y^{Reference}$. (4) Convergence of the discrete solutions $y_k{}^n(\Delta x)$ is monotonic.

5.5 The Bounds of Numerical Uncertainty

When the exact solution y^{Exact} of a system of partial or ordinary differential equations is unknown, the error $|y^{Exact} - y_k{}^n(\Delta x)|$ or $||y^{Exact} - y_k{}^n(\Delta x)||$ with which a discrete solution $y_k{}^n(\Delta x)$ is calculated becomes an *uncertainty*. This uncertainty cannot be explicitly computed since the exact solution is unknown and the approximation provided by the reference solution $y^{Reference}$ is itself uncertain. The best that one can achieve, therefore, is an upper bound $U(\Delta x)$ of truncation error such that:

$$\left|\frac{y^{Exact} - y_k^n(\Delta x)}{y_k^n(\Delta x)}\right| \leq U(\Delta x). \tag{5.12}$$

References [14, 15] derive such an upper bound and illustrate how it can be used to select an appropriate mesh size for the prediction of resonant dynamics of a wind turbine blade. The result is given as:

$$U(\Delta x) = \frac{1}{R^p - 1} \cdot \left|\frac{y_k^n(\Delta x) - y_k^n(R \cdot \Delta x)}{y_k^n(\Delta x)}\right|, \tag{5.13}$$

where R denotes the refinement ratio used between two successive calculations. By analogy with previous notation, the mesh size $R{\cdot}\Delta x$ represents a coarse resolution (previously denoted as Δx_C) and Eq. (5.13) defines the upper bound of truncation error for a fine-resolution (previously denoted as Δx_M) prediction.

The significance of the upper bound (5.13) is to indicate where the exact-but-unknown solution y^{Exact} might be located relative to the prediction $y_k{}^n(\Delta x)$ obtained at mesh resolution Δx. It is analogous to the Grid Convergence Index (GCI) proposed by Patrick Roache in 1994 to report the results of grid convergence studies in computational fluid dynamics [16]. The definition (5.13) does not, however, suffer from the disadvantage of having to select a somewhat arbitrary "factor-of-safety" F_S that appears in the GCI:

$$GCI\left(\Delta x\right) = \frac{F_S}{R^p - 1} \cdot \left| \frac{y_k^n\left(\Delta x\right) - y_k^n\left(R \cdot \Delta x\right)}{y_k^n\left(\Delta x\right)} \right|, \tag{5.14}$$

where the value of F_S depends on the application, characteristics of the code used and smoothness of the solution. Values $F_S = 1.25$ and 3.0 have been proposed for various applications in computational fluid dynamics and solid mechanics [2]. Some practitioners advocate to make the factor-of-safety variable to account for better or worse-than-expected convergence. Reference [17], for example, offers a definition that increases the upper bound if the observed rate-of-convergence is worse than anticipated:

$$F_S = \frac{R^{p^{Theory}} - 1}{R^p - 1}, \tag{5.15}$$

where p^{Theory} and p denote the theoretical and observed (in Eq. (5.7) or (5.9)) rates-of-convergence.

A small value of the upper bound $U(\Delta x)$, or the GCI, indicates that, first, the discrete solution obtained at mesh resolution Δx is converging to the exact-but-unknown solution y^{Exact} and, second, Richardson's extrapolation $y^{Reference}$ is a reasonably accurate estimation of y^{Exact}. What one considers "small enough" is subjective, usually, $U(\Delta x) = 10\%$ or less.

5.6 An Application of Solution Verification to a One-dimensional Advection Solver

An example is discussed to illustrate how solution verification is conducted in practice. Simplicity of our example should not obscure that the procedure would be similar for a real-world application in computational physics or engineering. Reference [14], for example, shows an application to wind turbine blade vibration using a finite element model.

The example is an advection equation solved in one-dimensional, Cartesian coordinates with a diffusion term added in the right-hand side:

$$\frac{\partial \psi\left(x;t\right)}{\partial t} + \alpha \cdot \frac{\partial \psi\left(x;t\right)}{\partial x} = \beta \cdot \frac{\partial^2 \psi\left(x;t\right)}{\partial x^2} + S\left(x;t\right), \tag{5.16}$$

where the computational domain is defined as $-1 \le x \le +1$. The initial (at $t = 0$) and boundary (at $x = \pm 1$) conditions are defined below. With the absence of a source function, that is, $S(x;t) = 0$ in Eq. (5.16), the waveform $\Psi(x;t)$ is simply advected (or translated) at constant velocity α across the computational domain. The coefficient β controls the amount of diffusion added to the solution. The solution $\Psi(x;t)$ is continuous, meaning that no discontinuity (or shock) or instability develops as time progresses.

An upwind scheme, which is 1st-order accurate, is selected to discretize the advection operator while central differences, which are 2nd-order accurate, are used for the diffusion term. These choices are written as:

$$\begin{gathered} \frac{\partial \psi\left(x;t\right)}{\partial t} \approx \frac{1}{\Delta t} \cdot \left(\psi_k^{n+1} - \psi_k^n\right), \quad \frac{\partial \psi\left(x;t\right)}{\partial x} \approx \frac{1}{\Delta x} \cdot \left(\psi_{k+1}^n - \psi_k^n\right). \\ \frac{\partial^2 \psi\left(x;t\right)}{\partial x^2} \approx \frac{1}{\Delta x^2} \cdot \left(\psi_{k+1}^n - 2\psi_k^n + \psi_{k-1}^n\right) \end{gathered} \tag{5.17}$$

The resulting finite volume scheme is fully explicit and implemented in a Matlab™ simulation code as:

$$\psi_k^{n+1} = \psi_k^n + \frac{\alpha \cdot \Delta t}{\Delta x} \cdot \left(\psi_{k+1}^n - \psi_k^n\right) + \frac{\beta \cdot \Delta t}{\Delta x^2} \cdot \left(\psi_{k+1}^n - 2\psi_k^n + \psi_{k-1}^n\right) + S_k^n, \tag{5.18}$$

where $\Psi_k{}^n$ denotes the discretized zone-averaged solution at location $x = k{\cdot}\Delta x$ and time $t = n{\cdot}\Delta t$:

$$\psi_k^n = \int_{x=\left(k-\frac{1}{2}\right)\Delta x}^{x=\left(k+\frac{1}{2}\right)\Delta x} \psi\left(x; n\cdot\Delta t\right)\cdot dx. \tag{5.19}$$

Modified equation analysis provides the continuous equation that the discrete solution $\Psi_k{}^n$ attempts to approximate. The derivations (not shown) indicate that it consists of the original Eq. (5.16) augmented with an infinite number of terms that represent the truncation error introduced by the particular choice of discretization scheme (5.17). The leading-order terms of the modified equation are:

$$\begin{aligned}\frac{\partial\psi}{\partial t} + \alpha\cdot\frac{\partial\psi}{\partial x} = \beta\cdot\frac{\partial^2\psi}{\partial x^2} + S + \frac{\alpha\Delta x}{2}\cdot(1-CFL)\cdot\frac{\partial^2\psi}{\partial x^2} - \frac{\alpha\Delta x^2}{6}\cdot\left(1-CFL^2\right)\cdot\frac{\partial^3\psi}{\partial x^3},\\ + \frac{\beta\Delta t}{2}\cdot\left(\alpha\frac{\partial^3\psi}{\partial x^3} - \beta\frac{\partial^4\psi}{\partial x^4}\right) - \frac{\beta\Delta t^2}{3}\cdot\left(\alpha^2\frac{\partial^4\psi}{\partial x^4} - \alpha\frac{\partial^5\psi}{\partial x^5} + \frac{\beta^2}{2}\frac{\partial^6\psi}{\partial x^6}\right) + \ldots\end{aligned} \tag{5.20}$$

with the Courant number defined as $CFL = \alpha{\cdot}\Delta t/\Delta x$. The fully explicit numerical scheme of Eq. (5.18) is stable only if the time-space discretization (Δt; Δx) satisfies $CFL < 1$. The modified equation clearly indicates that the proposed numerical method is 1st-order accurate in both time and space since the leading terms of truncation error are proportional to Δt (for the temporal error) and Δx (for the spatial error).

We now assume that modified equation analysis cannot be carried out. This is usually the case for real-world applications that feature multiple dimensions, large numbers of state variables, coupled mechanics or physics, and nonlinear governing equations, which makes it intractable to derive the modified equation. When confronted with this situation, the behavior of truncation error is studied through mesh refinement.

The exact solution of the homogeneous variant of Eq. (5.16), for which $S(x;t) = 0$, is:

$$\psi^{Exact}\left(x;t\right) = \psi_0\cdot e^{+\lambda t}\cdot\sin\left(kx+\omega t\right), \tag{5.21}$$

where the constants are defined as $\omega = -\alpha{\cdot}k$ and $\lambda = -\beta{\cdot}k^2$. In the numerical application discussed next, the coefficients are defined as $\alpha = 0.2$, $\beta = 0.002$, $\Psi_o = 1$ and $k = 15$. In addition, the initial and boundary conditions are defined to guarantee that $S(x;\, t)$ remains equal to zero at all locations and instants with:

$$\begin{aligned}\psi^{Exact}\left(x;t=0\right) &= \psi_0\cdot\sin(kx)\\ \psi^{Exact}\left(x=-1;t\right) &= \psi_0\cdot e^{-\beta k^2 t}\cdot\sin\left(-k\left(\alpha t+1\right)\right),\\ \psi^{Exact}\left(x=+1;t\right) &= \psi_0\cdot e^{-\beta k^2 t}\cdot\sin\left(-k\left(\alpha t-1\right)\right).\end{aligned} \tag{5.22}$$

Figure 5.3 illustrates exact solutions $\Psi^{Exact}(x;t)$ at times $t = 0$ (initial condition) and $t = 1$. The two physical effects can be observed with, first, advection that translates the solution across the domain (from the red curve to the blue curve) and, second, diffusion that removes energy from the solution, which tends to "spread" the waveform spatially and reduces its amplitude.

Table 5.1 defines six discrete solutions used to execute a mesh refinement study. The refinement ratio is uniform and equal to $R = 2$. Settings selected for these runs satisfy the stability criterion ($CFL < 1$), even for the most refined discretization with 3040 zones. Figure 5.4 illustrates the discrete solutions (color lines) and compares them to the exact solution (black dashed line) at $t = 1$. The figure clearly indicates that the discrete solutions converge to the exact solution (5.21) without, however, revealing at which rate.

Formal verification is carried out next. Code verification could be performed using the entire solution field by studying the behavior of truncation error $||\Psi^{Exact} - \Psi_k{}^n||$ since the exact solution is analytically known as shown in Eqs. (5.21) and (5.22). It is not, however, what is illustrated here. The exact solution is *not* used and solution verification is performed instead. To simplify the discussion, it is applied to peak values extracted from the discrete solutions at the final simulation time of $t = 1$, that is:

$$_{Max}\psi_k^n = \max_{-1\le x\le+1}\psi_k^n\left(x;1\right). \tag{5.23}$$

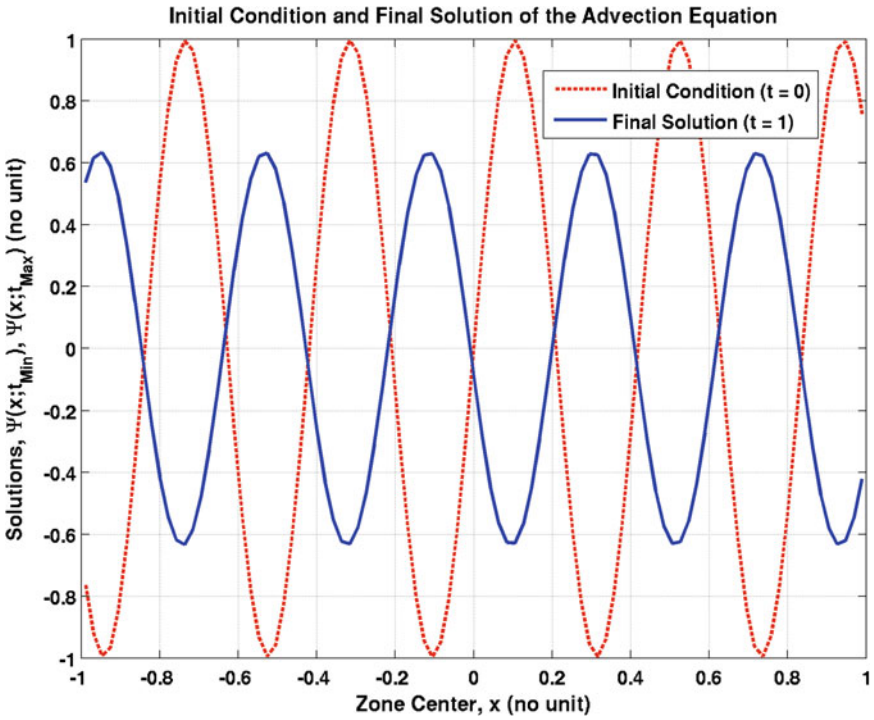


Fig. 5.3 Exact solutions Ψ^{Exact} at times $t = 0$ (*red* curve) and $t = 1$ (*blue* curve)

Table 5.1 Definition of settings used to generate six discrete solutions ${\Psi_k}^n$

Solution	N_{Zone}	Δx	Δt	CFL
1	95	21.052×10^{-3}	10^{-4}	9.5×10^{-4}
2	190	10.526×10^{-3}	10^{-4}	19.0×10^{-4}
3	380	5.263×10^{-3}	10^{-4}	38.0×10^{-4}
4	760	2.632×10^{-3}	10^{-4}	76.0×10^{-4}
5	1520	1.316×10^{-3}	10^{-4}	152.0×10^{-4}
6	3040	0.658×10^{-3}	10^{-4}	304.0×10^{-4}

The objectives of the analysis are to, first, verify if the peak values pictured in Fig. 5.4 converge with a 1st-order rate when the computational mesh is refined, as they should according to the modified Eq. (5.20); and, second, quantify the solution uncertainty at any level of mesh resolution.

Even though it would be possible to use triplets of discrete solutions to estimate the rate-of-convergence using Eq. (5.9) with $R = 2$, the behavior of truncation error is studied by simultaneously processing the six discrete solutions and solving Eq. (5.7) using a nonlinear optimization solver. It gives:

$$\left|{}_{Max}\psi^{Reference} - {}_{Max}\psi_k^n\,(\Delta x)\right| = 3.454 \cdot \Delta x^{0.998}, \tag{5.24}$$

with ${}_{Max}\Psi^{Reference} = 0.634$. The observed convergence rate of $p = 0.998$ clearly indicates 1st-order spatial convergence, which matches expectation given the theoretical properties of the numerical solver (5.18).

Table 5.2 lists the bounds of solution uncertainty $U(\Delta x)$ for the four finest levels of resolution. The fourth column gives the upper bounds as percentages of discrete solutions, as defined in Eq. (5.13), and the fifth column gives them in absolute magnitudes, that is, ${}_{Max}{\Psi_k}^n(\Delta x) \cdot U(\Delta x)$. These discrete solutions, all obtained with a minimum of 380 computational zones, feature less than 10% numerical uncertainty due to truncation error. Another factor, that Table 5.2 does not show, is time-to-solution. If an overall level of 10% error is admissible to predict the peak value, one might decide to pursue simulations using the 380-zone discretization because its solution can be computed faster than using a finer level of mesh resolution.

Figure 5.5 summarizes the results by illustrating the discrete solutions (blue squares), which are the peak values ${}_{Max}{\Psi_k}^n(\Delta x)$ obtained at different resolutions Δx, and their bounds of uncertainty (vertical green lines), which are the upper bounds $U(\Delta x)$ that result from mesh-induced truncation error. Note that the figure uses a log-log scale on its horizontal (mesh sizes) and vertical (solutions) axes. The red dashed line represents the 1st-order trend with which discrete solutions converge.

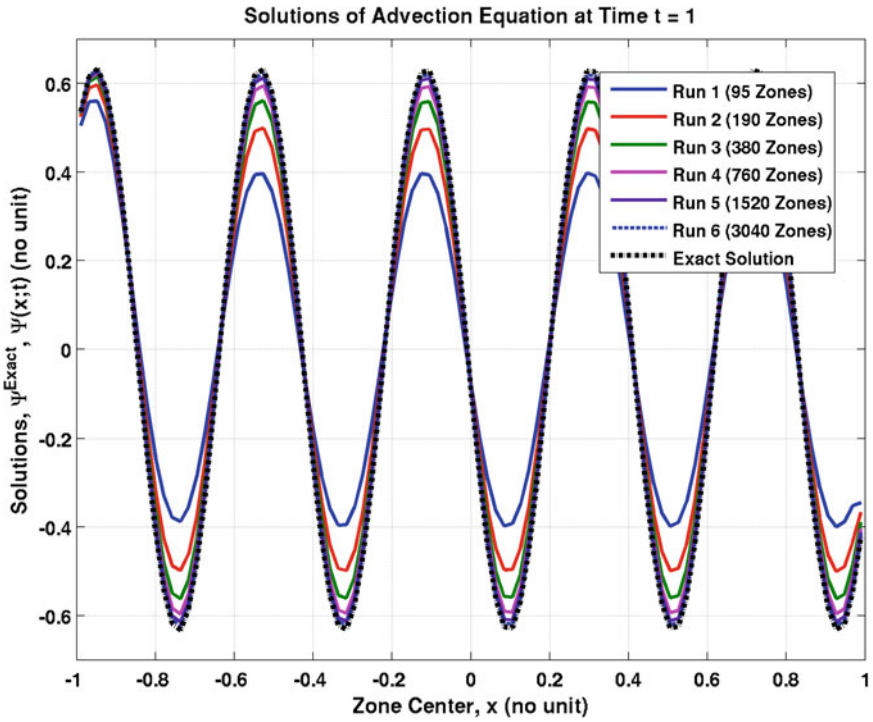


Fig. 5.4 Comparison of discrete $\Psi_k{}^n$ (*color lines*) and exact Ψ^{Exact} (*black dashed line*) solutions

Table 5.2 Bounds of solution uncertainty $U(\Delta x)$ at different mesh resolutions

Solution	N_{Zone}	Δx	$U(\Delta x)$	${}_{Max}\Psi_k{}^n(\Delta x)\bullet U(\Delta x)$
3	380	5.263×10^{-3}	9.12%	0.0561
4	760	2.632×10^{-3}	4.24%	0.0265
5	1520	1.316×10^{-3}	2.09%	0.0132
6	3040	0.658×10^{-3}	1.05%	0.0066

The horizontal black dashed line indicates the extrapolated solution ${}_{Max}\Psi^{\text{Reference}}$, which is the best estimate of the value that the numerical solver would converge to if a calculation could be performed at infinite resolution, that is, as $\Delta x \rightarrow 0$.

5.7 Conclusion

Verifying that a simulation code is free of programming mistakes, and the numerical methods that it implements deliver convergence properties which match theory, are essential activities to establish the quality of numerical solutions. Simply speaking, it is the first "V" of the Verification and Validation of numerical simulations. The discipline of code and solution verification in computational engineering and physics develops tools to assess the sources and severity of discretization and numerical errors. The discussion offered in this manuscript outlines simple procedures to evaluate the regime of asymptotic convergence, estimate a rate-of-convergence based on mesh (or grid) refinement, extrapolate discrete solutions, and develop bounds of numerical uncertainty due to discretization-induced truncation error.

Much of verification, however, remains incomplete. A non-exhaustive list of topics that warrant further research includes: extending the state-of-the-practice of verification to non-scalar quantities (curves, images, multiple-dimensional fields), developing technology to study the coupling between temporal and spatial discretization, defining a "reference" mesh for the estimation of solution error, and developing methods to verify the numerical quality of calculations which feature adaptive mesh refinement. It is our hope that computational engineers and physicists, while increasingly relying on modeling and simulation, will recognize the importance that numerical issues play in model validation and become advocates for best practices in the discipline of code and solution verification.

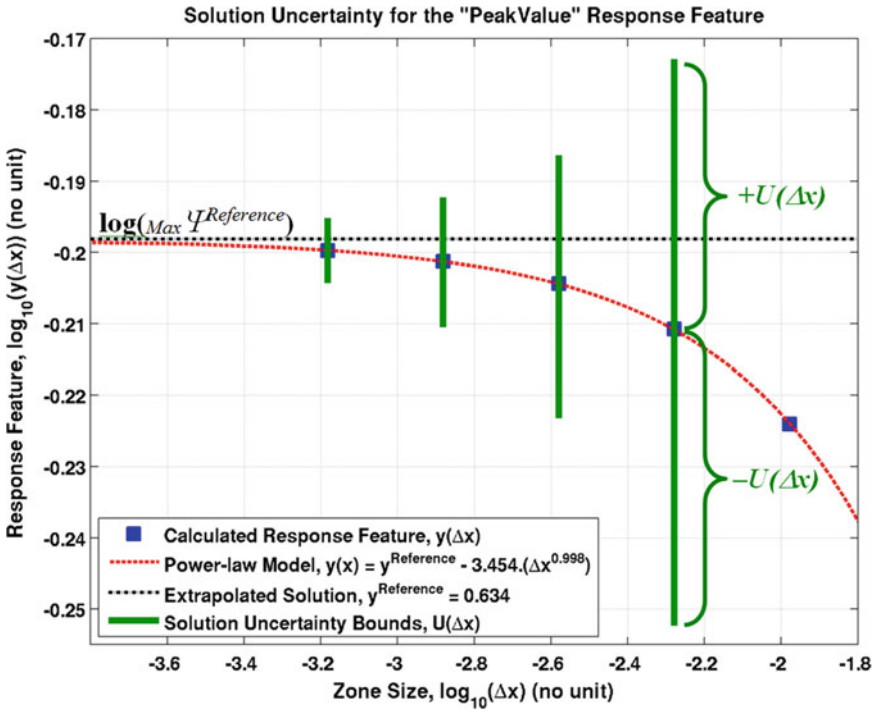


Fig. 5.5 Convergence of discrete solutions ${}_{Max}\Psi_k{}^n$ and their bounds of truncation error $U(\Delta x)$

Acknowledgments The author gratefully acknowledges the support of the Advanced Certification Campaign at the Los Alamos National Laboratory (LANL). LANL is operated by Los Alamos National Security, L.L.C., under contract DE-AC52-06NA25396 on behalf of the National Nuclear Security Administration of the U.S. Department of Energy.

References

1. Hemez, F.M., Doebling, S.W., Anderson, M.C.: A brief tutorial on verification and validation. In: 22nd SEM International Modal Analysis Conference, Dearborn, Michigan (2004)
2. Roache, P.J.: Verification in Computational Science and Engineering. Hermosa Publishers, Albuquerque, NM (1998)
3. Salari, K., Knupp, P.: Code verification by the method of manufactured solutions, Technical Report SAND-2000-1444, Sandia National Laboratories, Albuquerque, NM (2000)
4. Kamm, J.R., Rider, W.J., Brock, J.S.: Consistent metrics for code verification, Technical Report LA-UR-02-3794, Los Alamos National Laboratory, Los Alamos, NM (2002)
5. Kamm, J.R., Rider, W.J., Brock, J.S.: Combined space and time convergence analyses of a compressible flow algorithm. In: 16th AIAA Computational Fluid Dynamics Conference, Orlando, FL (2003)
6. Knoll, D.A., Chacon, L., Margolin, L.G., Mousseau, V.A.: On balanced approximations for time integration of multiple time scale systems. J. Comput. Phys. **185**(2), 583–611 (2003)
7. Buechler, M., McCarty, A., Reding, D., Maupin, R.D.: Explicit finite element code verification problems. In: 22nd SEM International Modal Analysis Conference, Dearborn, MI (2004)
8. Li, S., Rider, W.J., Shashkov, M.J.: Two-dimensional convergence study for problems with exact solution: uniform and adaptive grids, Technical Report LA-UR-05-7985, Los Alamos National Laboratory, Los Alamos, NM (2005)
9. Smitherman, D.P., Kamm, J.R., Brock, J.S.: Calculation verification: point-wise estimation of solutions and their method-associated numerical error. J. Aerosp. Comput. Inf. Commun. **4**, 676–692 (2007)
10. Hemez, F.M., Brock, J.S., Kamm, J.R.: Nonlinear error ansatz models in space and time for solution verification. In: 1st Non-deterministic Approaches Conference and 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Newport, RI (2006)
11. Warming, R., Hyett, B.: The modified equation approach to the stability and accuracy analysis of finite difference methods. J. Comput. Phys. **14**, 159–179 (1974)
12. LeVeque, R.J.: Numerical Methods for Conservation Laws. Birkhauser-Verlag Publishers, Basel (1990)
13. Lax, P.D., Richtmyer, R.D.: Survey of the stability of linear finite difference equations. Commun. Pure Appl. Math. **9**, 267–293 (1956)
14. Mollineaux, M.G., Van Buren, K.L., Hemez, F.M., Atamturktur, S.: Simulating the dynamics of wind turbine blades: part I, model development and verification. Wind Energy. **16**, 694–710 (2013)

15. Van Buren, K.L., Mollineaux, M.G., Hemez, F.M., Atamturktur, S.: Simulating the dynamics of wind turbine blades: part II, model validation and uncertainty quantification. Wind Energy. **16**, 741–758 (2013)
16. Roache, P.J.: Perspective: a method for uniform reporting of grid refinement studies. ASME J. Fluids Eng. **116**, 405–413 (1994)
17. Stern, F., Wilson, R., Shao, J.: Quantitative V&V of computational fluid dynamics (CFD) simulations and certification of CFD codes with examples. In: 2004 ICHMT International Symposium on Advances in Computational Heat Transfer, Norway (2004)

Chapter 6
Robust Optimization of Shunted Piezoelectric Transducers for Vibration Attenuation Considering Different Values of Electromechanical Coupling

Anja Kuttich, Benedict Götz, and Stefan Ulbrich

Abstract Structural vibration may occur in mechanical systems leading to fatigue, reduced durability or undesirable noise. In this context, shunting piezoelectric transducers to resonant shunts can be an appropriate measure for attenuating vibrations. The achieved vibration attenuation significantly depends on the tuning of the shunt parameters. Uncertainty in design and application of resonant shunted piezoelectric transducers may result in a detuned attenuation system and loss of attenuation performance. Therefore, we propose an approach based on robust optimization using the Bounded Real Lemma to contain the loss of vibration attenuation due to uncertainty. It is shown for resonant shunts, that for increasing electromechanical coupling coefficients the worst-case maximal vibration amplitudes for non-robust and robust optimization of shunt parameters converge. Furthermore by adding a negative capacitance to the resonant shunt, the worst-case maximal amplitude remains almost constant for all considered coupling coefficients for non-robust and robust optimization of the shunt parameters.

Keywords Robust optimization • Shunted piezoelectric transducers • Vibration attenuation

6.1 Introduction

Structural vibrations are (in many cases) undesirable and vibration reduction is an important goal in a variety of engineering application. An appropriate measure for attenuating vibrations are shunted piezoelectric transducers. This approach has been a research subject for several decades [1–4]. Generally, a piezoelectric transducer converts mechanical kinetic energy of a vibrating host structure into electrical energy. By connecting the electrodes of the transducer to an electrical circuit, the electromechanical impedance of the shunted transducer may achieve vibration attenuation. In [3], an overview of shunt-damping technologies such as mono- or multi-modal resonant shunts or more complex shunts with negative capacitance [1] and switched shunts [2] is given. With increasing shunt complexity, enhancements in vibration attenuation and system robustness are pursued. This paper focuses on mono-modal resonant shunts. Shunting the piezoelectric transducer with resistor and inductance, the resonant RL-shunt, an electrical oscillation circuit with the inherent capacitance of the transducer is created. This electromechanical system acts comparable to a mechanical vibration absorber. Resonant RL-shunts are easy to implement and no stability limits or switching laws have to be taken into account. On the other side, the achieved vibration attenuation significantly depends on the correct tuning of the shunt parameters and the amount of electromechanical coupling of transducer an structure [5]. By adding a negative capacitance, the resonant RLC-shunt achieves higher vibration attenuation. Furthermore, uncertainty in design and application of resonant shunted piezoelectric transducers may occur due to insufficient modeling, geometric and material deviations or deviations in the electrical quantities. During operation, this uncertainty may result in a detuned attenuation system and a significant loss of attenuation performance. Reducing the effects of uncertainty on the attenuation performance with resonant shunt damping, approaches such as robust optimization [6, 7] and shunts with adaptive inductors [8] have been investigated. So far, the investigations of vibration attenuation with shunted piezoelectric transducers under uncertainty were mainly limited to systems with a fixed electromechanical

A. Kuttich (✉) • S. Ulbrich
Technische Universität Darmstadt, Dolivostraße 15, 64293 Darmstadt, Germany
e-mail: kuttich@mathematik.tu-darmstadt.de

B. Götz
System Reliability, Adaptive Structures, and Machine Acoustics SAM, Technische Universität Darmstadt, Magdalenenstraße 4, 64289 Darmstadt, Germany

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_6

coupling of transducer and structure. However, different values of electromechanical coupling will result in different effects of uncertainty on the attenuation performance, as already investigated in own earlier studies [9]. In this paper, the vibration attenuation performance of resonant shunts, with and without negative capacitance, with robust optimization of shunt parameters is compared to the vibration attenuation performance with non-robust optimization of shunt parameters for different values of electromechanical coupling.

The investigated system is a directly excited single mass oscillator. For vibration attenuation, resonant RL- and RLC-shunts are connected to the transducer electrodes. Optimized shunt parameters are obtained by direct minimization of the $\mathcal{H}_\infty$-norm of the frequency transfer function of the structure with shunted transducer. Uncertainty is considered by detuning of the structural resonance frequency resulting, e.g., from uncertain structural stiffness. Then, we apply the concept of robust optimization to obtain a robust vibration attenuation for different values of electromechanical coupling. Using the Bounded Real Lemma, the optimization problem can be reformulated as nonlinear semidefinite programming problem and solved via an sequential semidefinite programming approach. It is shown by numerical simulation that the nonlinear sequential programming approach can be used for the robust optimization of shunt circuits. Furthermore, recommendations for the application of robust optimization of the parameters of systems with RL and RLC-shunts are presented taking into account the trade-off between robust and non-robust optimization influenced by the values of electromechanical coupling.

6.2 Frequency Transfer Function of the Single Mass Oscillator

The vibration attenuation with resonant RL-shunts and with resonant RLC-shunts including a negative capacitance is investigated for the non-robust and robust optimization of shunt parameters under uncertainty in the frequency domain. In this section, the frequency transfer function of the directly excited single mass system in Fig. 6.1a will be derived and used in the next section to calculate non-robust and robust optimized shunt parameters. In this section, we will only consider the case of RLC-shunts. It is easily seen that by neglecting C we obtain the corresponding model for RL-shunts.

The investigated system is a single mass system with mass m supported by a spring with stiffness k and a piezoelectric stack transducer, directly excited by the harmonic force $F(t)$, Fig. 6.1a. For the purpose of vibration attenuation, the stack transducer will be shunted to a resonant shunt comprising a resistance R, an inductance L and a negative capacitance C. The transducer's properties are the number of stack layers n_p, the piezoelectric material constant d_p, the stiffness k_p, all three lumped in $\Theta = n_\mathrm{p}\, d_\mathrm{p}\, k_p$, and the electrical capacitance C_p.

The transducer is represented by an uniaxial transducer model with electrical charges

$$q(t) = \Theta\, z(t) + C_\mathrm{p}\, u(t) \tag{6.1}$$

produced due to the deformation $z(t)$ and the applied electrical potential difference $u(t)$ between the electrodes. Furthermore, the transducer creates the force

$$F_\mathrm{p}(t) = -k_\mathrm{p}\, z(t) + \Theta\, u(t). \tag{6.2}$$

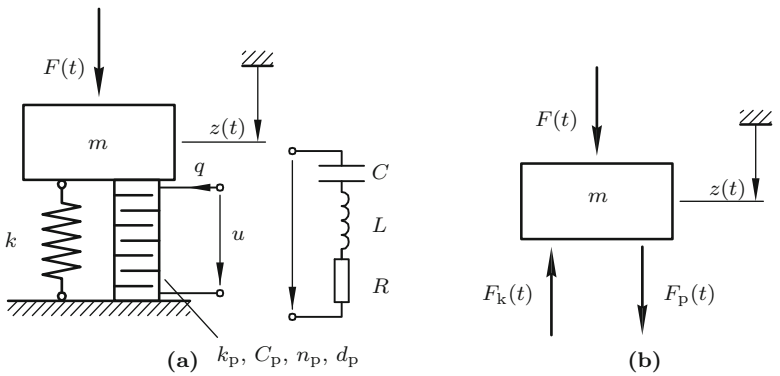


Fig. 6.1 (**a**) Single mass system excited via direct force and resonant RLC-shunt. (**b**) Free body diagram

Applying the principal of linear momentum, Fig. 6.1b, with (6.1) and (6.2) the equation of motion of the single mass system without shunt but with charge driven electrodes is

$$m\,\ddot{z}(t) + \left(k_{\mathrm{sc}} + \frac{\Theta^2}{C_{\mathrm{p}}}\right) z(t) - \frac{\Theta}{C_{\mathrm{p}}} q(t) = F(t) \tag{6.3}$$

where $k_{\mathrm{sc}} = k + k_{\mathrm{p}}$ is the transducer stiffness with short circuited (sc) electrodes. Furthermore, in (6.3), the term $k_{\mathrm{sc}} + \Theta^2\, C_{\mathrm{p}}^{-1} = k_{\mathrm{oc}}$ represents the transducer stiffness with open circuited (oc) electrodes [10]. For vibration reduction with shunted transducers, the generalized electromechanical coupling coefficient K_{33} is an important measure to evaluate the vibration attenuation capability. It characterizes the energy exchanged between the mechanical and the electrical domains of the piezoelectric transducer coupled to the structure. The coupling coefficient

$$K_{33} = \sqrt{\frac{\omega_{\mathrm{oc}}^2 - \omega_{\mathrm{sc}}^2}{\omega_{\mathrm{sc}}^2}} \tag{6.4}$$

may be calculated with the short and open circuited angular eigenfrequencies $\omega_{\mathrm{sc}} = \sqrt{\frac{k_{\mathrm{sc}}}{m}}$ and $\omega_{\mathrm{oc}} = \sqrt{\frac{k_{\mathrm{oc}}}{m}}$.

To describe the effect of resonantly shunted piezoelectric transducers on the mass vibration, a linear shunt is taken into account. Using the second KIRCHHOFF's law, the voltages across the shunt electrodes is

$$u(t) = -L\,\ddot{q}(t) - R\,\dot{q}(t) - \frac{1}{C}\,q(t). \tag{6.5}$$

With (6.1) and (6.5), the electromechanical shunt circuit equation of motion becomes

$$L\,C_{\mathrm{p}}\,\ddot{q}(t) + R\,C_{\mathrm{p}}\,\dot{q}(t) + \left(1 + \frac{C_{\mathrm{p}}}{C}\right) q(t) - \Theta\, z(t) = 0. \tag{6.6}$$

For a better condition of the system matrices, the non-dimensional time $\tau = \omega_{\mathrm{sc}}\, t$ and the normalized parameters

$$r(\tau) = \frac{q(\tau)}{\Theta} = z(\tau) + \frac{C_{\mathrm{p}}}{\Theta}\,u(\tau) \quad \text{and} \quad \gamma = K_{33}^2 = \frac{\Theta^2}{C_{\mathrm{p}}\,k_{\mathrm{sc}}}, \tag{6.7}$$

are introduced [5]. With the introduction of τ, the time derivatives in (6.3) become $\ddot{z}(t) = \omega_{\mathrm{sc}}^2\, z''(\tau)$ and $\dot{z}(t) = \omega_{\mathrm{sc}}\, z'(\tau)$. Analogue operations are performed for $r(\tau)$. Now, (6.3) and (6.5) with normalized coordinates and parameters read

$$\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & L\,C_{\mathrm{p}}\,\omega_{\mathrm{sc}}^2 \end{bmatrix}}_{\boldsymbol{M}} \begin{bmatrix} z''(\tau) \\ r''(\tau) \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & R\,C_{\mathrm{p}}\omega_{\mathrm{sc}} \end{bmatrix}}_{\boldsymbol{D}} \begin{bmatrix} z'(\tau) \\ r'(\tau) \end{bmatrix} + \underbrace{\begin{bmatrix} 1+\gamma & -\gamma \\ -1 & 1 + \dfrac{C_{\mathrm{p}}}{C} \end{bmatrix}}_{\boldsymbol{K}} \begin{bmatrix} z(\tau) \\ r(\tau) \end{bmatrix} = \begin{bmatrix} \dfrac{F(\tau)}{k_{\mathrm{sc}}} \\ 0 \end{bmatrix} \tag{6.8}$$

with mass matrix $\boldsymbol{M}$, damping matrix $\boldsymbol{D}$ and stiffness matrix $\boldsymbol{K}$.

The vibration attenuation capability of the shunt circuit is investigated by studying the frequency transfer function. Therefore, (6.8) is transferred into a system of first order differential equations and, then, the transfer function of the force $F(\tau)$ to the displacement $z(\tau)$ is derived in frequency domain. The system of first order differential equation of (6.8) is

$$\begin{aligned} \boldsymbol{x}'(\tau) &= \underbrace{\begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\boldsymbol{M}^{-1}\boldsymbol{K} & -\boldsymbol{M}^{-1}\boldsymbol{D} \end{bmatrix}}_{\boldsymbol{A}} \boldsymbol{x}(\tau) + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}}_{\boldsymbol{B}} \frac{F(\tau)}{k_{\mathrm{sc}}} \\ z(\tau) &= \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}}_{\boldsymbol{C}} \boldsymbol{x}(\tau) \end{aligned} \tag{6.9}$$

with system matrix $\boldsymbol{A}$, input matrix $\boldsymbol{B}$, output matrix $\boldsymbol{C}$ and the state space vector $\boldsymbol{x}(\tau) = [z(\tau), r(\tau), z'(\tau), r'(\tau)]^T$. From (6.9), the frequency transfer function

$$\boldsymbol{G}(j\eta) = \boldsymbol{C}\,(j\eta\boldsymbol{I} - \boldsymbol{A})^{-1}\,\boldsymbol{B} \tag{6.10}$$

is obtained relating the force input $F(\eta)$ to the displacement output $z(\eta)$ with the normalized excitation frequency $\eta = \dfrac{\Omega}{\omega_{\mathrm{sc}}}$.

6.3 Robust Optimization Approach

We are interested in optimal shunt parameters R and L for RL-shunts and R, L and C for RLC-shunts which minimize the $\mathcal{H}_\infty$-norm of the transfer function (6.10). In this section, we will only consider the case of RLC-shunts. It is easily seen that by neglecting C we obtain the corresponding model for RL-shunts. To get mechanical useful results we have to take into account some inequality constraints $\boldsymbol{g}(R, L, C)$ such as R and L have to be positive and C has to be negative. Furthermore we have to take into account the stability condition $\dfrac{C_{\mathrm{p}}}{C} > -\dfrac{1}{1+\gamma}$. The resulting optimization problem for optimal shunt parameters is the following

$$\begin{aligned} \min_{R,L,C} \quad & ||\boldsymbol{G}(R, L, C)||_{\mathcal{H}_\infty} \\ \text{s.t.} \quad & \boldsymbol{g}(R, L, C) \leq 0, \end{aligned} \tag{6.11}$$

which is equivalent to

$$\begin{aligned} \min_{R,L,C,\phi} \quad & \phi \\ \text{s.t.} \quad & ||\boldsymbol{G}(R, L, C)||_{\mathcal{H}_\infty} \leq \phi \\ & \boldsymbol{g}(R, L, C) \leq 0. \end{aligned} \tag{6.12}$$

To be able to consider uncertain input data in our optimization problem and to be able to efficiently solve it we have to reformulate (6.12). With the aid of the so-called Bounded Real Lemma the optimization problem (6.12) can be reformulated as nonlinear semidefinite programming problem.

Theorem 1 (Bounded Real Lemma). *Suppose $\boldsymbol{A}$ is stable and $(\boldsymbol{A} - s\boldsymbol{IB})$ is controllable. Then the following statements are equivalent*

(i) $||\boldsymbol{G}||_{\mathcal{H}_\infty} \leq \phi$

(ii) There exist $\boldsymbol{Y} \succ 0$ *with* $\begin{pmatrix} \boldsymbol{A}^\top\boldsymbol{Y} + \boldsymbol{YA} & \boldsymbol{YB} & \boldsymbol{C}^\top \\ \boldsymbol{B}^\top\boldsymbol{Y} & -\phi & 0 \\ \boldsymbol{C} & 0 & -\phi \end{pmatrix} \preceq 0.$

Proof. See [11, p. 37] and its references.

Since our system matrix $\boldsymbol{A}$ is stable (every eigenvalue of $\boldsymbol{A}$ has strictly negative real part) and $(\boldsymbol{A} - s\boldsymbol{IB})$ is controllable for all possible values of R, L and C the Bounded Real Lemma is applicable. Therefore, (6.12) is equivalent to

$$\begin{aligned} \min_{R,L,C,Y,\phi} \quad & \phi \\ \text{s.t.} \quad & \begin{pmatrix} \boldsymbol{A}(R, L, C)^\top\boldsymbol{Y} + \boldsymbol{YA}(R, L, C) & \boldsymbol{YB} & \boldsymbol{C}^\top \\ \boldsymbol{B}^\top\boldsymbol{Y} & -\phi & 0 \\ \boldsymbol{C} & 0 & -\phi \end{pmatrix} \preceq 0 \\ & -\boldsymbol{Y} \prec 0 \\ & \boldsymbol{g}(R, L, C) \leq 0. \end{aligned} \tag{6.13}$$

We use the concept of robust optimization to deal with uncertainty in the eigenfrequency ω_{sc}. The idea of robust optimization is to assume an uncertainty set $\mathcal{U}$ in which the uncertain parameter varies [12]. The optimization then aims to find the solution with the optimal worst-case value for all elements in the uncertainty set. In contrast to stochastic optimization the robust optimization approach does not consider certain probabilities for the parameters. The resulting optimization problem for varying $\omega_{\mathrm{sc}} \in \mathcal{U}$ is given by

$$\begin{aligned} \min_{R,L,C,\boldsymbol{Y}} \quad & \max_{\omega_{\mathrm{sc}} \in \mathcal{U}} \phi \\ \text{s.t.} \quad & \begin{pmatrix} \boldsymbol{A}(R,L,C,\omega_{\mathrm{sc}})^{\top}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{A}(R,L,C,\omega_{\mathrm{sc}}) & \boldsymbol{Y}\boldsymbol{B} & \boldsymbol{C}^{\top} \\ \boldsymbol{B}^{\top}\boldsymbol{Y} & -\phi & 0 \\ \boldsymbol{C} & 0 & -\phi \end{pmatrix} \preceq 0 \\ & -\boldsymbol{Y} \prec 0 \\ & g(R,L,C) \leq 0. \end{aligned} \tag{6.14}$$

For the choice of the underlying uncertainty set $\mathcal{U}$ we have to take into account the computationally tractability of the resulting robust nonlinear semidefinite optimization problem (6.14). Therefore we will use a finite set as uncertainty set based on different scenarios for ω_{sc}, namely $\mathcal{U} = \{\omega_{\mathrm{sc}}^1, \ldots, \omega_{\mathrm{sc}}^S\}$ for a finite number of scenarios S. Therefore, (6.14) simplifies to

$$\begin{aligned} \min_{R,L,C,\boldsymbol{Y}} \quad & \phi \\ \text{s.t.} \quad & \begin{pmatrix} \boldsymbol{A}(R,L,C,\omega_{\mathrm{sc}}^s)^{\top}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{A}(R,L,C,\omega_{\mathrm{sc}}^s) & \boldsymbol{Y}\boldsymbol{B} & \boldsymbol{C}^{\top} \\ \boldsymbol{B}^{\top}\boldsymbol{Y} & -\phi_s & 0 \\ \boldsymbol{C} & 0 & -\phi_s \end{pmatrix} \preceq 0 \qquad \forall s = 1,\ldots,S \\ & -\boldsymbol{Y} \prec 0 \\ & g(R,L,C) \leq 0 \\ & \phi_s - \phi \leq 0 \qquad \forall s = 1,\ldots,S. \end{aligned} \tag{6.15}$$

For sake of completeness, we state the complete optimization model with specified inequality constraints.

$$\begin{aligned} \min_{R,L,C,\boldsymbol{Y}} \quad & \phi \\ \text{s.t.} \quad & \begin{pmatrix} \boldsymbol{A}(R,L,C,\omega_{\mathrm{sc}}^s)^{\top}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{A}(R,L,C,\omega_{\mathrm{sc}}^s) & \boldsymbol{Y}\boldsymbol{B} & \boldsymbol{C}^{\top} \\ \boldsymbol{B}^{\top}\boldsymbol{Y} & -\phi_s & 0 \\ \boldsymbol{C} & 0 & -\phi_s \end{pmatrix} \preceq 0 \qquad \forall s = 1,\ldots,S \\ & -\boldsymbol{Y} \prec 0 \quad -\frac{1}{1+\gamma} - \frac{C_{\mathrm{p}}}{C} < 0 \\ & -R \leq 0 \quad -L \leq 0 \quad C \leq 0 \\ & \phi_s - \phi \leq 0 \qquad \forall s = 1,\ldots,S. \end{aligned} \tag{6.16}$$

The resulting optimization problem (6.16) is a nonlinear semidefinite optimization problem. In general, nonlinear semidefinite optimization problems are not easy to solve and are subject of current research in mathematical optimization. As a suitable solution approach we use sequential semidefinite programming based on [13] which we extend by the consideration of nonlinear inequality constraints. In each iteration of the sequential semidefinite programming algorithm a linearized subproblem of the nonlinear semidefinite program (6.15) is solved. To solve the linearized subproblems we use the solver SeDuMi which solves conic problems with a primal-dual interior-point algorithm [14]. For the sequential semidefinite programming algorithm a suitable initial point is needed. Here, we use the optimal shunt parameters according to [5]. For the RL-Shunt we use

$$R_{\mathrm{RL}} = \frac{2D_{\mathrm{RL}}}{C_{\mathrm{p}}\omega_{\mathrm{sc}}}, \qquad D_{\mathrm{RL}} = \sqrt{\frac{1}{2/\gamma + 4 + 2\gamma}}, \quad L_{\mathrm{RL}} = \frac{\lambda_{\mathrm{RL}}}{C_{\mathrm{p}}\omega_{\mathrm{sc}}^2}, \qquad \lambda_{\mathrm{RL}} = \frac{1}{1+\gamma} \tag{6.17}$$

as initial point. For the RLC-shunt the initial value is given by

$$
\begin{aligned}
R_{\mathrm{RLC}} &= \frac{2D_{\mathrm{RLC}}}{C_{\mathrm{p}}\omega_{\mathrm{sc}}}, & \delta_{\mathrm{RLC}} &= \frac{1.113\,\gamma - 1}{\gamma + 1}, & L_{\mathrm{RLC}} &= \frac{\lambda_{\mathrm{RLC}}}{C_{\mathrm{p}}\omega_{\mathrm{sc}}^2}, \\
D_{\mathrm{RLC}} &= (1 + \delta_{\mathrm{RLC}})\, D_{\mathrm{RL}}, & C_{\mathrm{RLC}} &= \frac{C_{\mathrm{p}}}{\delta_{\mathrm{RLC}}}, & \lambda_{\mathrm{RLC}} &= (1 + \delta_{\mathrm{RLC}})\, \lambda_{\mathrm{RL}}.
\end{aligned}
\tag{6.18}
$$

6.4 Numerical Results

In this section we present numerical results illustrating the benefit of robust optimization of shunt parameters for the single mass system with RL and RLC-shunts. It turns out that uncertain eigenfrequencies have much more impact for RL-shunts than for RLC-shunts. We compare the results for the non-robust optimization (6.13) and the robust optimization (6.16) for RL and RLC-shunts for the vibration amplitudes $|\boldsymbol{G}|$ of the frequency transfer function. Therefore, we first compare the nominal (non-robust) solution to the robust solution for the nominal case of the eigenfrequency for RL and RLC-shunts for a fixed electromechanical coupling coefficient K_{33}. Secondly, we compare the nominal solution for the worst-case eigenfrequency to the robust solution for the worst-case eigenfrequency, also for RL and RLC-shunts with a fixed K_{33}. We will see that for both shunts the robust solution achieves better vibration attenutation compared to the nominal solution for the worst-case eigenfrequency. To evaluate the influence of the electromechanical coupling coefficient on the robust optimization we then compare the worst-case objective value ϕ of the nominal and robust solution for RL and RLC-shunts for varying K_{33}. Our numerical results will show that electromechanical coupling coefficient has a positive influence on the benefit of robust optimization in case of RL-shunts an small values of K_{33}, but has nearly no impact on the robust solution for RLC-shunts. We use a Matlab implementation of the sequential semidefinite programming algorithm and the solver SeDuMi to solve the resulting linear semidefinite programs. As the uncertainty set, we choose a finite set with five scenarios and a maximum variance of $\pm 10\%$ around the nominal value of $\omega_{\mathrm{sc}} = 469.5542\,\frac{1}{\mathrm{s}}$, that is $\mathcal{U} = [0.9,\ 0.95,\ 1,\ 1.05,\ 1.1]\,\omega_{\mathrm{sc}}$. For the electromechanical coupling coefficient K_{33} we use discrete values between 0.05 and 0.3 with a step size of 0.01.

Figure 6.2a, b show the vibration amplitudes $|\boldsymbol{G}|$ with RL-Shunt for the nominal and robust solution for $K_{33} = 0.1$ for nominal and worst-case eigenfrequencies on a logarithmic scale. Figure 6.2a shows that the use of RL-shunts significantly reduces the vibration amplitudes compared to systems without shunt and short circuited electrodes, as expected. Considering uncertainty in the optimization, the robust solution takes into account all scenarios of the uncertainty set $\mathcal{U}$, hence, the robust

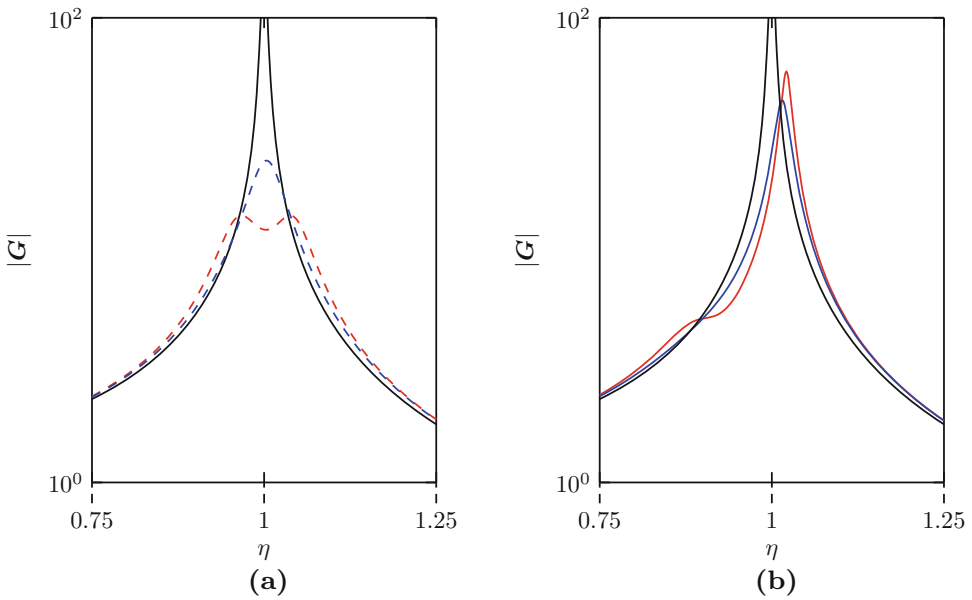


Fig. 6.2 Vibration amplitudes $|\boldsymbol{G}|$ with RL-shunt and $K_{33} = 0.1$ (**a**) nominal solution for nominal eigenfrequenxy (red dashed line), robust solution for nominal eigenfrequency (blue dashed line), (**b**) nominal solution for worst-case eigenfrequency (*red solid line*), robust solution for worst-case eigenfrequency (*blue solid line*); both: without shunt (*black solid line*)

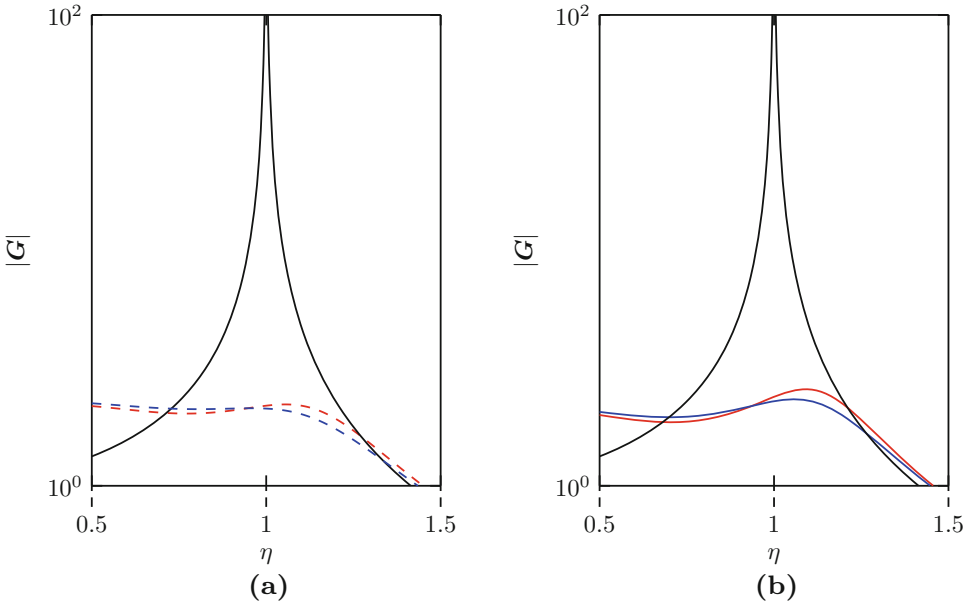


Fig. 6.3 Vibration amplitudes $|\boldsymbol{G}|$ with RLC-shunt with $K_{33} = 0.1$ (**a**) nominal solution for nominal eigenfrequenxy (*red dashed line*), robust solution for nominal eigenfrequency (blue dashed line), (**b**) nominal solution for worst-case eigenfrequency (*red solid line*), robust solution for worst-case eigenfrequency (*blue solid line*); both: without shunt (*black solid line*)

solution (blue dotted line) performs worse for the nominal eigenfrequency than the nominal solution (red dotted line), again as expected. However, in Fig. 6.2b, comparing the worst-case scenario for the nominal (blue line) and the robust (red line) solution one observes that the robust solution achieves a better worst-case vibration attenuation.

For the vibration amplitudes with RLC-shunt similar results are obtained, see Fig. 6.3a, b, but the changes in the vibration amplitudes for the nominal and the robust solution robust solution for the nominal and the worst-case eigenfrequency on a logarithmic scale are not as significant as for the RL-shunt.

In order to rate the benefit of the robust optimization approach with respect to the size of the electromechanical coupling coefficient K_{33}, we analyze the worst-case objective value ϕ (which is equivalent to the maximum vibration amplitude in the investigated frequency range) for different K_{33} for the single mass system with RL- and RLC-shunt. See Fig. 6.4a, b for results. Firstly, we want to point out that the optimization approach using (6.13) represented by the black lines always yields better vibration attenuation than the initial values for RL- and RLC-shunt (6.17) and (6.18) represented by the dashed black lines. This is mainly because the initial values according to [5] are derived by assuming that there exist fix-points for which the amplitudes of the frequency transfer function are equal. Using our optimization approach (6.13) we do not need such assumptions. Regarding RL-shunts in Fig. 6.4a, the robust optimization approach performs significantly better (up to 53%) for the worst-case scenario of the eigenfrequencies for small electromechanical coupling coefficient K_{33}, but thee advantage decreases with increasing K_{33}. Although the robust solution gives smaller maximal vibration amplitudes than the nominal solution for the worst-case eigenfrequency for $K_{33} \leq 0.15$, the advantage vanishes for $K_{33} > 0.15$. For RLC-shunts the difference between the maximal vibration amplitudes for the nominal and robust solution for the worst-case eigenfrequency is nearly constant for varying K_{33}, as depicted in Fig. 6.4b, (note the different scale of Fig. 6.4a, b). The maximal vibration amplitudes (worst-case objective values ϕ) for RLC-shunts are only 4% smaller for the robust solution for the worst-case eigenfrequency compared to the nominal solution also for the worst-case eigenfrequency and different values of K_{33} almost give the same maximal vibration amplitude, hence RLC-shunts are already much more robust with respect to uncertain eigenfrequencies than RL-shunts.

Finally, we want to point out that the worst-case objective value for the RL- and the RLC-shunt is not always obtained at the upper or lower bound of the uncertainty set. Consequently, it is not sufficient to only consider the lowest and highest eigenfrequency in the optimization approach to obtain robust solutions.

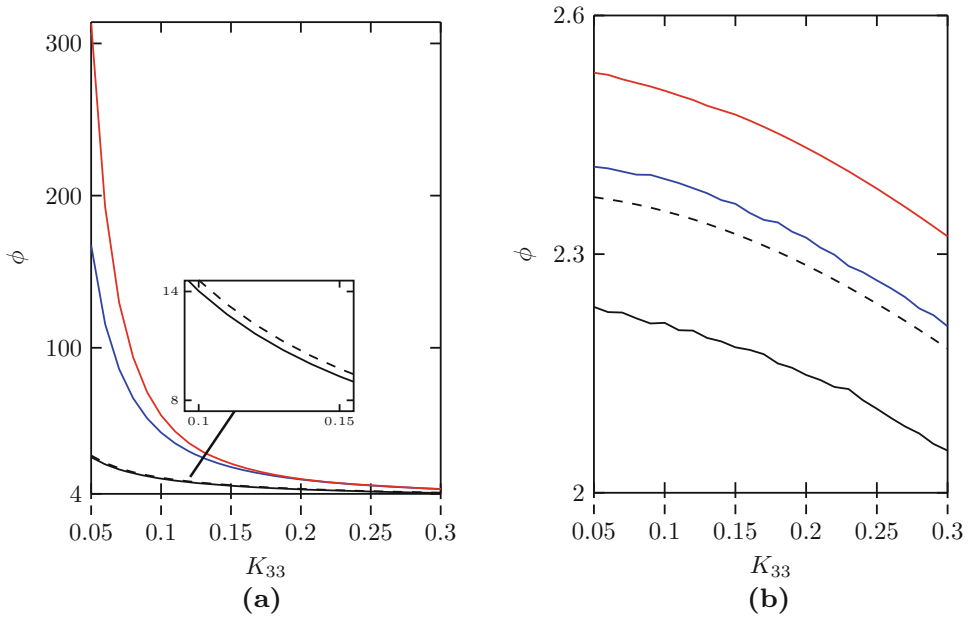


Fig. 6.4 Minimal objective function value ϕ for different values of K_{33} nominal solution for nominal eigenfrequency (black solid line), nominal solution for worst-case eigenfrequency (red solid line), robust solution for worst-case eigenfrequency (blue solid line)

6.5 Conclusion

We presented an optimization approach for the direct minimization of the $\mathcal{H}_\infty$-norm of the frequency transfer function of a single mass system with RL- and RLC-shunts. Furthermore, we introduced a robust optimization approach to deal with uncertainty in the structural eigenfrequency ω_{sc}. Our numerical results yield important information about the benefit of the robust optimization approach depending on the electromechanical coupling coefficient K_{33}. Whereas K_{33} influences the benefit of the robust optimization approach for systems with RL-shunts, K_{33} has nearly no impact on the robust solution for systems with RLC-shunts. To sum up, the robust optimization is useful for systems with RL-shunts considering an K_{33} smaller than 0.15. For systems with RLC-shunts the effort of robust optimization seems to be not necessary and, hence, uncertainty in the structural eigenfrequency may be neglected.

Acknowledgements The authors like to thank the German Research Foundation DFG for funding this research within the SFB 805.

References

1. Beck, B.S.: Negative capacitance shunting of piezoelectric patches for vibration control of continuous systems. PhD thesis, Georgia Institute of Technology (2012)
2. Niederberger, D.: Smart damping materials using shunt control. PhD thesis, Swiss Federal Institute of Technology Zürich (2005)
3. Moheimani, S.O.R., Fleming, A.J.: Piezoelectric Transducers for Vibration Control and Damping. Springer, London (2006)
4. Hagood, N.W., von Flotow, A.H.: Damping of structural vibrations with piezoelectric materials and passive electrical networks. J. Sound Vib. **146**(2), 243–268 (1991)
5. Neubauer, M., Oleskiewicz, R., Popp, K., Krzyzynski, T.: Optimization of damping and absorbing performance of shunted piezo elements utilizing negative capacitance. J. Sound Vib. **298**, 84–107 (2006)
6. Soltani, P., Kerschen, G., Tondreau, G., Deraemaeker, A.: Piezoelectric vibration damping using resonant shunt circuits: an exact solution. Smart Mater. Struct. **23**, 1–11 (2014)
7. Zambolini-Vicente, B.G.G.L., Silva, V.A.C., de Lima, A.M.G.: Robust design of shunt circuits for passive control of vibrations of composite structures. In: Proceedings of the 2nd International Symposium on Uncertainty Quantification and Stochastic Modeling (2014)
8. Mokrani, B., Burday, I., Tian, Z., Preumont, A.: Adaptive inductor for vibration damping in presence of uncertainty. In: Proceedings of SMART2015 7th ECCOMAS Thematic Conference on Smart Structures and Materials, Ponta Delgada, Azores, June 3–6 2015

9. Götz, B., Platz, R., Melz, T.: Consistent approach to describe and evaluate uncertainty in vibration attenuation using resonant piezoelectric shunting and tuned mass dampers. In: Proceedings of the 2nd International Symposium on Uncertainty Quantification and Stochastic Modeling Uncertainties 2014, Rouen/France, June 23–27 2014, pp. 51–64. INSA-Rouen
10. de Marneffe, B.: Active and passive vibration isolation and damping via shunted transducers. PhD thesis, Université Libre de Bruxelles (2007)
11. Scherer, C.: The Riccati inequality and state-space H_∞-optimal control. PhD thesis, Bayerische Julius Maximilians-Universität Würzburg (1990)
12. Nemirovksi, A., Ben-Tal, A., El Ghaoui, L.: Robust Optimization. Princeton University Press, Princeton (2009)
13. Correa, R., Ramirez, H.C.: A global algorithm for nonlinear semidefinite programming. SIAM J. Optim. **15**(1), 303–318 (2005)
14. Sturm, J.F.: Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones. Optim. Methods Softw. **11–12**, 625–653 (1999)

Chapter 7
Parameter Estimation and Uncertainty Quantification of a Subframe with Mass Loaded Bushings

Mladen Gibanica and Thomas J.S. Abrahamsson

Abstract In the automotive industry components are often connected through rubber bushings. The bushings' material properties are usually not well known. In computational models these properties are parametrised and their spread can be considerable. A good estimate of these parameters is important in various applications, including substructuring, and for uncertainty quantification of systems with connected components. This paper deals with the calibration of an industrial size finite element model of a car subframe with parametrised bushing models. Mass loading is used on the bushings to bring local bushing modes to a lower frequency region and impose a more realistic boundary condition in component testing. The model parameters can be calibrated in different ways. In this paper two approaches are considered. They are based on two test configurations, one with and one without mass loaded boundaries. In the first case only the bushing parameters are considered for the mass loaded boundary configuration. In the second case, consisting of two steps, the configuration without mass loaded boundaries is considered in which the bushing parameters are first fixed and other model parameters considered, and in the last step a subset of all parameters is considered. The calibration, validation and uncertainty quantification, using bootstrapping, have been performed using the open-source MATLAB tool FEMcali.

Keywords Uncertainty quantification • Model updating • Industrial application • Parametrised bushings • FEMcali

7.1 Introduction

In the automotive industry finalised products are commonly made up of thousands of components, which can in turn be made out of smaller components. These structures are increasingly studied with computer aided engineering (CAE) software, such as finite element (FE) software, due to the flexibility computer models bring in terms of modifying and analysing the structures. There exist many ways of analysing the structures. An approach that is often taken is to divide the structure into substructures and analyse every substructure separately, known as substructuring [1, 2]. The substructures results are assembled and the complete structure's responses can be solved for. For an accurate modelling it is important to understand the physics of every substructure, and especially the physical behaviour at the interfaces through which the substructures are assembled. Many different connection types exist, for different purposes, where bushing are commonly used when it is of interest to control, or suppress the vibration levels from one component to another [3]. Hence, they are commonly used in the automotive industry to connect, for example, subframes to the body in white (BIW). The bushings are partly constructed from rubber materials, and it is well known that the Young's modulus of rubber materials is temperature and frequency dependent and can exhibit large variations between samples [3]. Therefore, synthesised structures through rubber bushings can show a large variation, stemming from the uncertain rubber material properties in the bushings, but also poor prediction capabilities due to unknown, and not modelled, physics. Furthermore, for connected structures, through rubberised bushings, a good damping estimation can be important in the assembly's predictive capabilities. The damping is usually not modelled based on first principles due to its complexity, but rather by calibration of models to experimentally acquired data. Therefore, there is an interest in updating FE models towards experimentally obtained data from vibration tests [4], so that the FE model responses corresponds better to the specific component of interest, but also to obtain the damping characteristics of the structure.

M. Gibanica (✉)
Applied Mechanics, Chalmers University of Technology, Hörsalsvägen 7A, SE-41296 Göteborg, Sweden
Volvo Car Corporation, SE-40531 Göteborg, Sweden
e-mail: mladen.gibanica@chalmers.se

T.J.S. Abrahamsson
Applied Mechanics, Chalmers University of Technology, Hörsalsvägen 7A, SE-41296 Göteborg, Sweden

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_7

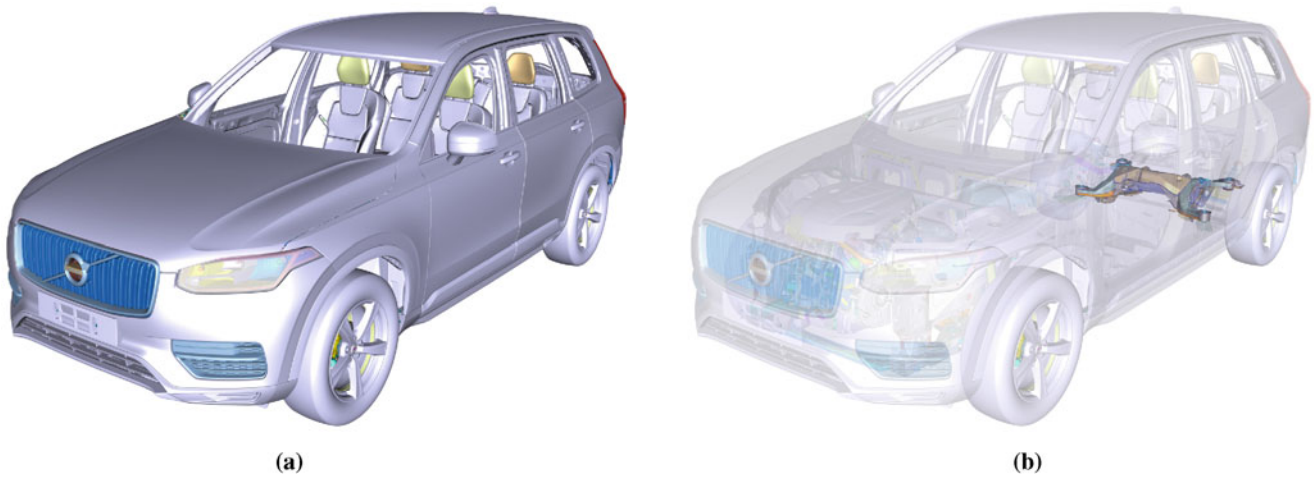


Fig. 7.1 The rear subframe connects the suspension link arms and the car body, and is thus an important part from a vibration perspective. (**a**) Parts of a Volvo XC90. (**b**) Rear subframe location in the car

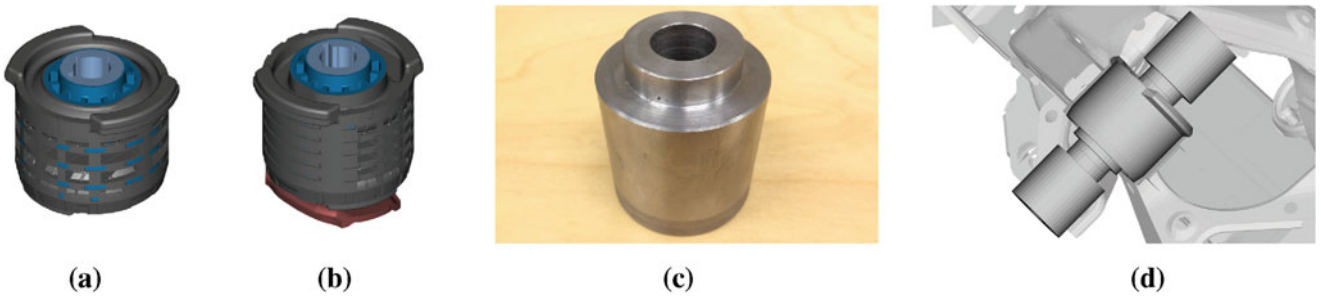


Fig. 7.2 In (**a**) one bushing is depicted and in (**b**) the other bushing. In (**c**) the additional mass is shown and in (**d**) the assembly of the additional mass and the rear subframe, for one bushing

In previous studies at Volvo Car Corporation three Volvo V40 subframes [5, 6] and three Volvo XC90 (2015) front subframes [7] have been calibrated towards experimental data. This paper, forming a second study on the Volvo XC90, is concerned primarily with the estimation of the bushing rubber material parameters in a mass loaded boundary configuration of a Volvo XC90 rear subframe, shown in Fig. 7.1, but also to update the rear subframe model in general. The rear subframe, like the front subframe is important from a vibration perspective as it connects the suspension link arms to the car body. In this paper model updating [8] is proposed as a tool for estimating the bushing model parameters by mass loading the bushings in an experimental modal analysis (EMA). By mass loading the bushings they are activated considerably and bring local bushing modes down in frequency where it is possible to study them with modal analysis. This also provides a good estimate of the additional damping provided by the bushings. Furthermore, mass loading provides a more realistic boundary condition on the bushings, i.e. closer to the assembled behaviour which is sought. In Fig. 7.2 the two different bushing types are shown along with the mass loading component in Fig. 7.2c and a small region of an FE representation of the mass loaded bushing boundary configuration in Fig. 7.2d.

In this paper the calibration is performed with the open-source application FEMcali, described in [9–11]. Two finite element models are used, one with and one without mass loaded boundary configuration. The FE models of both configurations consist of over half a million degrees of freedom each. In the calibration procedure a parametrised surrogate model of the FE model is first created. The calibration method then seeks to obtain a calibrated parameter setting such that the deviation between the FE model's frequency response functions (FRFs) and the experimentally identified model's FRFs is minimised. Normalised damping is imposed on both models to circumvent the mode pairing problem. A relatively high damping also yields a smooth deviation metric [9]. The calibrated parameter setting is then used as a starting point in an optimisation procedure using bootstrapping to quantify the parameter uncertainties. Previous studies using the same calibration procedure on various structures have produced mostly good results, see [5–7, 12, 13]. Pretest planning is performed on a coarse FE model using the effective independence (EfI) method [14, 15] with an added gramian rejection step, described further in [16]. In this paper a single input multiple output (SIMO) stepped sine testing procedure is used, with

multiple simultaneous sinusoidal inputs for faster test procedure utilising an improved frequency strategy proposed in [17] for subspace state-space system identification [18]. First the FE model with mass loading is calibrated towards experimental data for a good parameter estimation of the rubber's Young's modulus, i.e. all other model parameters are fixed at their nominal values. Then the configuration without mass loading is calibrated, with fixed but updated rubber parameters, for an overall good parameter estimation of the rear subframe FE model. In a final step a subset of all parameters are updated towards a higher frequency region for the configuration without mass loading to gain further physical understanding of the component.

Even though it is known that mass produced components exhibit varying dynamic and static properties, and it has been shown that a spread in the behaviour between front subframe individuals of a Volvo XC90 [7] is noticeable, this paper is concerned with only one rear subframe of said car model due to less observed variability between the rear subframes. This was experimentally verified from experiments on three rear subframes.

Rubber can be characterised as a viscoelastic material [3]. It is known that its Young's modulus is frequency and temperature dependent [3], which is also shown in [19] for some specific rubber materials. In this paper the rubber material is modelled with solid elements and an isotropic linear material model, which according to [3] is sufficient as the bushing parameters are estimated from the mass loaded configuration and the frequency region of interest is small enough for the Young's modulus frequency dependency to be disregarded. Small strains are assumed and non-linear effects are not regarded. Temperature effects are also disregarded.

The theory behind the model updating procedure used in this paper is presented briefly in Sect. 7.2. In Sect. 7.3 the vibration testing procedure, FE models and parameter selection are described. The results are presented in Sect. 7.4 and in Sect. 7.5 the work is concluded.

7.2 Theory

The theory behind the model updating procedure is briefly introduced here. For a full explanation of the theory behind the deterministic model calibration, see [9], and for the model parameter uncertainty quantification theory see [10, 11].

The goal of the calibration is to estimate a calibration parameter setting $\boldsymbol{p}_b$ that minimises the deviation between the experimentally obtained FRFs $\boldsymbol{H}_X(\omega_i) \in \mathbb{C}^{n_y \times n_u}$ and the FE model FRFs $\boldsymbol{H}_{FE}(\omega_i) \in \mathbb{C}^{n_y \times n_u}$ at n_f frequency steps ω_i where n_y and n_u denote the number of outputs and inputs, respectively. It is possible to express $\boldsymbol{H}_X(\omega_i)$, see [10, 11, 20, 21], as

$$\begin{aligned} \boldsymbol{H}_X(\omega_i) &= \boldsymbol{H}_R(\omega_i) + \boldsymbol{N}_o(\omega_i) = (\boldsymbol{H}_{FE}(\omega_i) + \boldsymbol{N}_m(\omega_i)) + \boldsymbol{N}_o(\omega_i) \\ &= \boldsymbol{H}_{FE}(\omega_i) + \boldsymbol{N}_G(\omega_i), \quad i = 1, \ldots, n_f \end{aligned} \tag{7.1}$$

where $\boldsymbol{H}_R(\omega_i) \in \mathbb{C}^{n_y \times n_u}$ denote the true frequency response of the real structure, $\boldsymbol{N}_G(\omega_i) \in \mathbb{C}^{n_y \times n_u}$ is the observed prediction error which can be represented as $\boldsymbol{N}_G(\omega_i) = \boldsymbol{N}_m(\omega_i) + \boldsymbol{N}_o(\omega_i)$ where $\boldsymbol{N}_m(\omega_i)$ denote the model prediction error and $\boldsymbol{N}_o(\omega_i)$ denote the measurement noise. Assuming that the bias introduced by the FE model is small, $\boldsymbol{N}_{G_{r,s}}$, for $r = 1, \ldots, n_y$ and $s = 1, \ldots, n_u$, can be modelled as an independent, zero mean, circularly complex normally distributed random variable with a known variance σ^2.

The calibration procedure consist of three steps, as explained in [10, 11]. First an EMA is performed in which the used discrete frequencies are selected according to [17] and a system $\boldsymbol{\Sigma}$ is identified from the experimental data using N4SID, such that its transfer function representation $\boldsymbol{H}_\Sigma$ asymptotically fulfils the criterion $\lim_{n_f \to \infty} ||\boldsymbol{H}_R - \boldsymbol{H}_\Sigma||_\infty = 0$ as shown in [18]. N4SID is a linear subspace state-space system identification method [18] implemented in MATLAB's System Identification toolbox [22] which enables the evaluation of the experimental system at arbitrary frequencies. The experimental data set $\boldsymbol{H}_X$ is split into two data sets after this stage, i.e. poorly identified channels are used as validation data in a validation set denoted $\boldsymbol{H}_V$ and the remaining channels in a calibration set denoted $\boldsymbol{H}_X$. The channels used in the validation data set are thus also removed from the identified system $\boldsymbol{\Sigma}$. Further details of the system identification is given in Sect. 7.3.1.

The second step consists of a deterministic calibration procedure in which a calibration parameter setting $\boldsymbol{p}^*$ is estimated from the nominal FE model parameter setting $\boldsymbol{p}_0$ towards the identified system $\boldsymbol{H}_\Sigma$, with equalised damping in both models. A deviation metric, that is smooth and weights high and low structural responses equally is used in the Levenberg–Marquardt optimisation algorithm [23–25], and can be formed as

$$\boldsymbol{p}^* = \arg\min_{\boldsymbol{p}} \frac{\boldsymbol{\epsilon}^H(\boldsymbol{p})\boldsymbol{\epsilon}(\boldsymbol{p})}{N} \qquad \text{with} \qquad \boldsymbol{\epsilon}(\boldsymbol{p}) = \log_{10}\text{vect}(\boldsymbol{H}_{FE}(\boldsymbol{p})) - \log_{10}\text{vect}(\boldsymbol{H}_\Sigma) \tag{7.2}$$

where $\boldsymbol{H}_{FE}(\boldsymbol{p})$ represents the FE model, at parameter setting $\boldsymbol{p}$, and the superscript H denote the conjugate transpose. Here vect$(\cdot)$ stands for the vectorisation operation in which a matrix is transformed into a vector by stacking the columns of the matrix such that the transfer function matrices in Eq. (7.2) are transformed into an $n_y n_u n_f \times 1$ column vector. The objective function is non-linear and thus many start parameter settings are used, generated by the Latin hypercube sampling technique [26], from which $\boldsymbol{p}^*$ is selected from the best calibration outcome. After the calibration a mode pairing algorithm [27] is used to map the experimental damping to the correct FE model modes. The theoretical framework for this step is explained in Sect. 7.2.1.

The last step considers the parameter uncertainty, in which the calibrated parameter configuration $\boldsymbol{p}^*$ is used as an initial start location in an undamped Gauss-Newton optimisation procedure [25] to find the parameter configuration $\boldsymbol{p}_b$ that minimises the deviation metric, at some parameter setting $\boldsymbol{p}$,

$$\boldsymbol{p}_b = \arg\min_{\boldsymbol{p}} \frac{\boldsymbol{\gamma}^H(\boldsymbol{p})\boldsymbol{\gamma}(\boldsymbol{p})}{N} \qquad \text{with} \qquad \boldsymbol{\gamma}(\boldsymbol{p}) = \log_{10}\text{vect}(\boldsymbol{H}_{FE}(\boldsymbol{p})) - \log_{10}\text{vect}(\boldsymbol{H}_X^b) \tag{7.3}$$

where $\boldsymbol{H}_X^b$ represent bootstrap samples [28] drawn from $\boldsymbol{H}_X$. This procedure is repeated n_b times on different data sets and hence the measurement noise influence on the model parameters can be obtained, which is further explained in Sect. 7.2.2.

7.2.1 Deterministic Model Updating Procedure

The calibration procedure utilises the mass and stiffness matrices formed from an FE representation of the structure of interest. Such mechanically vibrating systems can be written as [1]

$$\boldsymbol{M}\ddot{\boldsymbol{q}} + \boldsymbol{V}\dot{\boldsymbol{q}} + \boldsymbol{K}\boldsymbol{q} = \boldsymbol{f}(t) \tag{7.4}$$

where the dot notation is used for time differentiation and $\boldsymbol{M}$, $\boldsymbol{V}$ and $\boldsymbol{K} \in \mathbb{R}^{m\times m}$ represent the mass, damping and stiffness matrices, respectively. The general displacement vector is denoted by $\boldsymbol{q}$ and the external force vector by $\boldsymbol{f}(t)$.

Systems on second order form, such as in Eq. (7.4), can be cast into first order form by forming a state vector $\boldsymbol{x} = [\boldsymbol{q}^T, \dot{\boldsymbol{q}}^T]^T$ which gives

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u} \\ \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{u} \end{cases} \tag{7.5}$$

with $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ representing the system matrix, $\boldsymbol{B} \in \mathbb{R}^{n\times n_u}$ representing the input matrix, $\boldsymbol{C} \in \mathbb{R}^{n_y\times n}$ representing the output matrix, $\boldsymbol{D} \in \mathbb{R}^{n_y\times n_u}$ representing the force throughput matrix and $\boldsymbol{y}$ representing the system outputs. The load vector $\boldsymbol{f}$ in Eq. (7.4) can be formed from the input vector $\boldsymbol{u} \in \mathbb{R}^{n_u}$ with the Boolean transformation matrix $\boldsymbol{U}$ as $\boldsymbol{f} = \boldsymbol{U}\boldsymbol{u}$. The relationship between the state dimension n and degrees of freedom m is $n = 2m$. The system matrices can then be formed as

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\boldsymbol{M}^{-1}\boldsymbol{K} & -\boldsymbol{M}^{-1}\boldsymbol{V} \end{bmatrix} \qquad \boldsymbol{B} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{M}^{-1}\boldsymbol{U} \end{bmatrix} \tag{7.6}$$

where $\boldsymbol{C}$ and $\boldsymbol{D}$ are formed appropriately so that linear combinations of the system states $\boldsymbol{x}$ and inputs $\boldsymbol{u}$ form the system outputs $\boldsymbol{y}$. The transfer function matrix $\boldsymbol{H}$ can be formed from the system in Eq. (7.5), at frequency ω_i, as

$$\boldsymbol{H}(\omega_i) = \boldsymbol{C}(j\omega_i\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D} \tag{7.7}$$

where j given by $j^2 = -1$ is the imaginary number and $\boldsymbol{I}$ the identity matrix of appropriate dimension.

In the calibration procedure equalised damping is imposed on the experimentally identified system $\boldsymbol{\Sigma}_X$, defined as the quadruple $\{\boldsymbol{A}_X, \boldsymbol{B}_X, \boldsymbol{C}_X, \boldsymbol{D}_X\}$, and the FE system in order to avoid mode pairing, but also to obtain a smooth deviation metric [9]. To this end, the experimentally identified system $\boldsymbol{\Sigma}_X$ is transformed to diagonal form by a similarity transformation [29], e.g. the eigenvector matrix $\boldsymbol{X}$ of the eigenvalue problem $\boldsymbol{A}_X\boldsymbol{X} = \boldsymbol{X}\boldsymbol{\Lambda}$ where $\boldsymbol{\Lambda}$ is the diagonal eigenvalue matrix. Then $\boldsymbol{X}^{-1}\boldsymbol{A}_X\boldsymbol{X} = \bar{\boldsymbol{A}}_X = \text{diag}(\lambda_i)$ where λ_i is the ith complex valued system pole for $i = 1, \ldots, n$. For small damping the relative damping is $\xi_i = -\mathbb{R}(\lambda_i)/\,|\mathbb{C}(\lambda_i)|$ with $|\cdot|$ representing the absolute value and $\mathbb{R}(\cdot)$ and $\mathbb{C}(\cdot)$ representing the real

and complex part of an imaginary number, respectively. In the damping equalisation step, the modal damping is set to a fixed value for all modes, i.e $\xi_i = \xi_0 \forall i$. A new state space system is then obtained, replacing $\boldsymbol{\Sigma}_X$, with the quadruple set $\boldsymbol{\Sigma} = \{\tilde{\boldsymbol{A}}_X, \boldsymbol{X}^{-1}\boldsymbol{B}_X, \boldsymbol{C}_X\boldsymbol{X}, \boldsymbol{D}_X\}$ with $\tilde{\boldsymbol{A}}_X = \text{diag}(\tilde{\lambda}_i)$ and

$$\tilde{\lambda}_i = \mathbb{C}(\lambda_i)(-\xi_0 + i) \quad \forall \mathbb{C}(\lambda_i) > 0, \quad \tilde{\lambda}_i = \mathbb{C}(\lambda_i)(\xi_0 + i) \quad \forall \mathbb{C}(\lambda_i) < 0. \tag{7.8}$$

For the FE model with given mass $\boldsymbol{M}$ and stiffness $\boldsymbol{K}$ matrices it is possible to impose the same level of relative damping by forming the modal viscous damping as described in [1]

$$\boldsymbol{V} = \boldsymbol{M}\boldsymbol{X}\text{diag}(m_i^{-1})\text{diag}(2\xi_0 m_i \Omega_i)\text{diag}(m_i^{-1})\boldsymbol{X}^T\boldsymbol{M}, \quad i = 1, \ldots, m \tag{7.9}$$

where Ω_i are the eigenfrequencies, m_i the modal masses and $\boldsymbol{X}$ the eigenvector matrix from the undamped eigenvalue problem of the system in Eq. (7.4), $\boldsymbol{K}\boldsymbol{X} = \boldsymbol{M}\boldsymbol{X}\text{diag}(\Omega_i^2)$. It should be noted that in this paper equalised damping is only applied to the modes in the frequency range of interest $\omega_i \in [\omega_{low}, \omega_{high}]$ for the experimental model.

Reduced order models must be formed from the full FE system as a practical consequence of the large size of industrial FE models. The calibration procedure works with physical parameters and thus parametrised mass and stiffness matrices are formed, which can be expressed as $\boldsymbol{M} = \boldsymbol{M}(\boldsymbol{P})$ and $\boldsymbol{K} = \boldsymbol{K}(\boldsymbol{P})$, where the parameter vector $\boldsymbol{P}$ is related to a normalised parameter vector $\boldsymbol{p}$ and some fixed non-zero nominal parameter $\boldsymbol{P}_0$ so that $\boldsymbol{P} = \boldsymbol{P}_0(1 + \boldsymbol{p})$ holds [9]. The eigenvector matrix $\boldsymbol{T}_0$ at the nominal parameter setting is used as a reduction basis and is kept constant during the calibration procedure, which is formed from the undamped eigenvalue problem of the system in Eq. (7.4) such that residual modes, influencing the system in the frequency region of interest, are included [1]. The reduced mass and stiffness matrices at any parameter setting $\boldsymbol{p}$ can then be formed as

$$\bar{\boldsymbol{M}}(\boldsymbol{p}) = \boldsymbol{T}_0^T\boldsymbol{M}(\boldsymbol{p})\boldsymbol{T}_0, \quad \bar{\boldsymbol{K}}(\boldsymbol{p}) = \boldsymbol{T}_0^T\boldsymbol{K}(\boldsymbol{p})\boldsymbol{T}_0 \tag{7.10}$$

with the reduced mass and stiffness matrices at the nominal parameter setting $\boldsymbol{p}_0$ being represented by $\boldsymbol{M}_0$ and $\boldsymbol{K}_0$. Gradients of the reduced mass and stiffness matrices can be formed as

$$\bar{\boldsymbol{M}}_{,j} = \boldsymbol{T}_0^T \left(\frac{d\boldsymbol{M}}{dp_j}\Big|_{p=p_0} \right) \boldsymbol{T}_0, \quad \bar{\boldsymbol{K}}_{,j} = \boldsymbol{T}_0^T \left(\frac{d\boldsymbol{K}}{dp_j}\Big|_{p=p_0} \right) \boldsymbol{T}_0 \tag{7.11}$$

for the jth calibration parameter. From the first order expansions of the Taylor series of the reduced mass $\bar{\boldsymbol{M}}$ and stiffness $\bar{\boldsymbol{K}}$ matrices about $\boldsymbol{p}_o$ it is possible to form a surrogate model that is linear in the parameters.

$$\tilde{\boldsymbol{M}}(\boldsymbol{p}) = \bar{\boldsymbol{M}}_0 + \sum_{j=1}^{n_p} (p_j - p_{j,0})\,\bar{\boldsymbol{M}}_{,j}, \quad \tilde{\boldsymbol{K}}(\boldsymbol{p}) = \bar{\boldsymbol{K}}_0 + \sum_{j=1}^{n_p} (p_j - p_{j,0})\,\bar{\boldsymbol{K}}_{,j} \tag{7.12}$$

Here $p_{j,0}$ is the jth parameter at the nominal setting. The FE transfer function representation $\boldsymbol{H}_{FE}$, which is used in the calibration procedure, can now be formed with use of Eq. (7.7) from the state-space matrix quadruple $\boldsymbol{\Sigma}_{FE} = \{\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}, \tilde{\boldsymbol{D}}\}$, which is formed with a state transformation $\boldsymbol{x} = \boldsymbol{T}\boldsymbol{\xi}$ and with a viscous damping matrix $\tilde{\boldsymbol{V}}$ formed as in Eq. (7.9), such that

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\tilde{\boldsymbol{M}}^{-1}\tilde{\boldsymbol{K}} & -\tilde{\boldsymbol{M}}^{-1}\tilde{\boldsymbol{V}} \end{bmatrix} \qquad \tilde{\boldsymbol{B}} = \begin{bmatrix} \boldsymbol{0} \\ \tilde{\boldsymbol{M}}^{-1}\boldsymbol{T}^T\boldsymbol{U} \end{bmatrix} \qquad \tilde{\boldsymbol{C}} = \begin{bmatrix} \boldsymbol{T} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T} \end{bmatrix} \qquad \tilde{\boldsymbol{D}} = \boldsymbol{D}. \tag{7.13}$$

7.2.2 Model Parameter Uncertainties

From the deterministic calibration procedure a best parameter configuration $\boldsymbol{p}^*$ is obtained, with respect to the identified model $\boldsymbol{\Sigma}$. From a decision-making perspective it is important to asses the uncertainties in the obtained parameters such that confidence in the predictions can be obtained. In this paper bootstrapping [28] is used for stochastic model updating, proposed in [10, 11]. The procedure works by repeatedly drawing random datasets $\boldsymbol{H}_X^b$ with replacement from the original

data set $\boldsymbol{H}_X$, i.e. n_f random numbers $\boldsymbol{d} = \{d_1, \ldots, d_i, \ldots, d_{n_f}\}$ are drawn from the uniform distribution on $\mathbb{I} = \{1, 2, \ldots, n_f\}$ and with the frequency lines $\mathbb{F} = \{\omega_1, \ldots, \omega_i, \ldots, \omega_{n_f}\}$ in the experimental set $\boldsymbol{H}_X$, $\boldsymbol{H}_X^b$ can be formed such that

$$\boldsymbol{H}_X^b = \boldsymbol{H}_X(\omega_{d_i}) \quad \forall \mathbb{I}. \tag{7.14}$$

This procedure is repeated n_b times such that the deviation in Eq. (7.3) is minimised n_b times for the different data sets. Thus n_b vectors of calibrated parameters $\boldsymbol{p}_b$ will be obtained, from which the variance and expected value can be computed [10, 11].

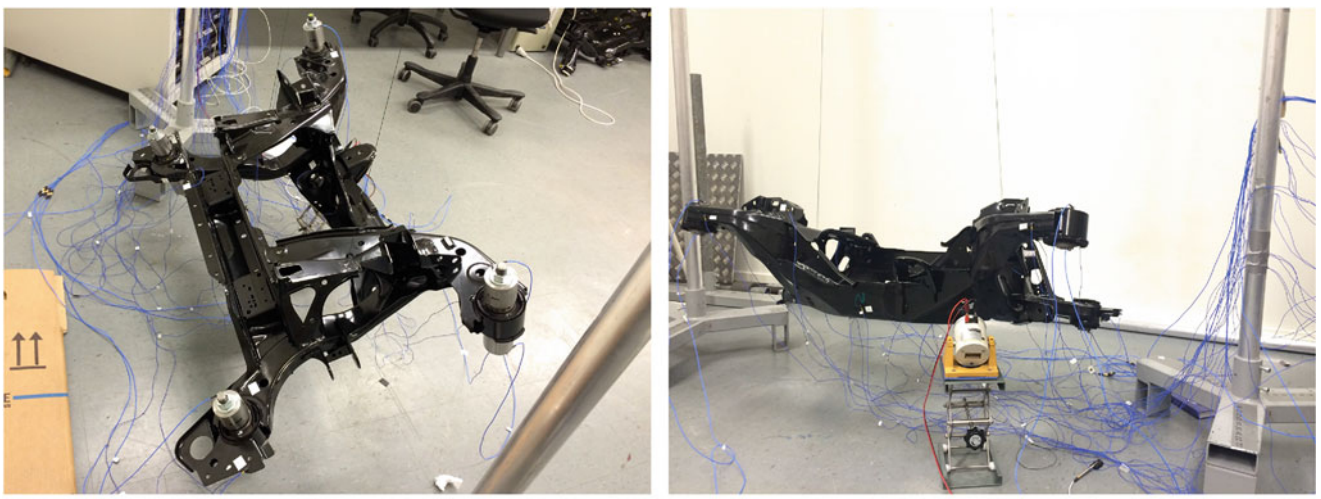
7.3 Model Preparation

In this section the experimental modal analysis is explained along with the two FE models and the calibration parameter selection.

7.3.1 Experimental Modal Analysis

Vibration tests have been performed at Chalmers University of Technology Vibration Lab on three rear subframes of the Volvo XC90 with and without mass loading. Experiments, and results, on one subframe is reported here, but the procedures for the two other components were identical. The rear subframe with mass loading is visible in Fig. 7.3a and without mass loading in Fig. 7.3b. The components were hung in long thin high-strength lines attached to steel springs on a supporting steel structure. Most of the quasi rigid body modes were kept under 2 Hz, while a support bouncing mode was kept under 5 Hz, i.e. the quasi rigid body modes were well under the first flexible mode of the component in both configurations.

An early version of the rear subframe FE model without modelled bushings was used in the pretest planing where 170 candidate locations were selected. Then 20 uniaxial sensor locations were found, out of the 170 candidate locations, with the set expansion EfI [14, 15] method, for the 20 first flexible modes, with an added gramian rejection step for rejection of locations with similar information [16]. Six more uniaxial accelerometer positions were added for visualisation purposes. An additional 10 triaxial accelerometers were placed on the bushings, where 8 triaxial accelerometer locations were different for the two configurations. All sensor placements can be seen in Fig. 7.4. In Fig. 7.4a, b the configuration without mass loading is shown while the mass loaded boundary configuration is shown in Fig. 7.4c, d. Note that circular markings indicate accelerometers while a rectangle marking (only location 24) indicates input, with an additional direct



(a) (b)

Fig. 7.3 Rear subframe shown in (**a**) configuration with mass loading, and in (**b**) configuration without mass loading, both with input at location 24

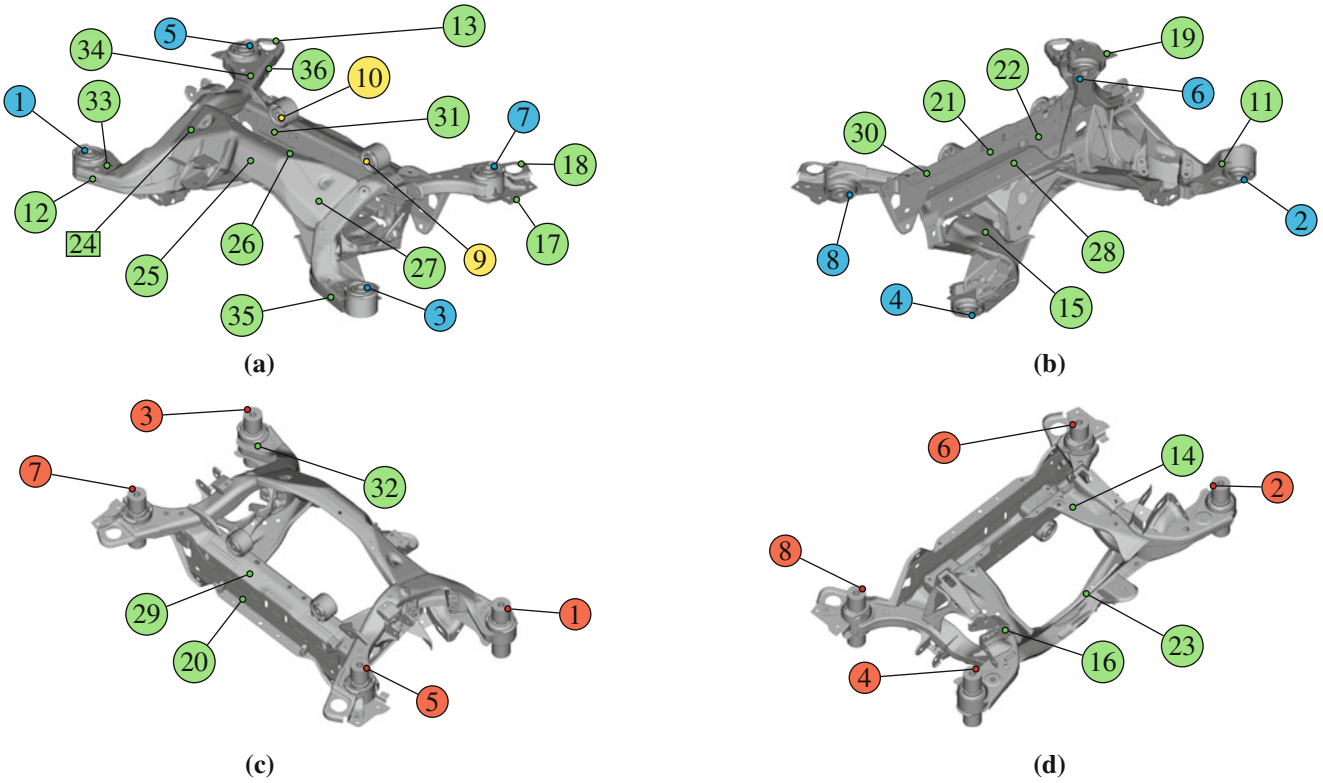


Fig. 7.4 Top view of configuration without mass loading in (**a**) and bottom view in (**b**). Top view of mass loaded configuration in (**c**) and bottom view in (**d**). *Circle markings* indicate accelerometer locations and a *rectangular marking* indicate the input position. Input was normal to the surface, with a direct accelerance force and accelerometer sensor configuration. *Green and yellow circles* represent uniaxial and triaxial accelerometer locations, respectively, equal for both configuration cases. *Cyan and red markings* indicate triaxial sensors for the configuration without and with mass loading, respectively

accelerometer placed on the opposite side of the sheet metal. This was considered best practice as the metal plate was flat and thin (approximately 2 mm). The green colour coding indicate uniaxial accelerometers (accelerometers 11 to 36), while yellow, cyan and red indicate triaxial accelerometers (accelerometers 1 to 10). Accelerometers 9 to 36 were placed identically on both configurations while accelerometers 1 to 8 were placed directly on the bushings in the configuration without mass loading and thus had to be placed on the additional component attached to the bushings in the mass loaded configuration. In Fig. 7.4 cyan markings indicate positions for the configuration without mass loading while red indicate positions for the mass loaded configuration.

Two type of accelerometers were used, 10 triaxial PCB Piezotronic type 356A03 weighting 1 gramme each and 26 uniaxial PCB Piezotronic type PCB 352C22/NC weighting 0.5 grammes each. The uniaxial accelerometers were attached with synthetic wax while the triaxial accelerometers were glued. The orientation of the triaxial accelerometers around its own z-axis was approximated by sight, therefore data gathered from the x and y components of these accelerometers must be considered less accurate than the z component, which was normal to the structure. The accelerometer masses were included in the FE models. The shaker used in the vibration test was of Ling Dynamic Systems make and of type V201, with a metallic stinger approximately 5 mm in length, when attached, as recommended in [4]. The excitation force was measured with a Brüel&Kjær force sensor of type 8203 with an IEPE converter 2647B, attached to the component through a stinger attachment plate which was glued and had a mass of around 0.2 g. The force sensor and stinger attachment plate masses were not included in the FE models.

Two excitation methods were used for both configurations. Periodic chirp tests at various amplitudes were performed to assess the linearity of the systems to obtain proper excitation loads. In both configurations a frequency range from 20 to 800 Hz was used. The calibration data was collected with a stepped sine tests with multiple sinusoidal components excited simultaneously for reduced noise influence on the estimated parameters in the model updating procedure. For the mass loaded configuration a frequency region from 40 to 500 Hz was selected with 2000 frequency lines, due to the relatively high damping in the system. In the configuration without mass loading a frequency region from 60 to 500 Hz was selected with 3000 frequency lines, due to lower damping. In [9] the authors advise in selecting the frequencies based on the half-band

width, $\Delta\omega_i$, of the eigenmodes. The half-band width for a damped structural resonance at a certain frequency ω_i can be expressed as $\Delta\omega_i = 2\xi_i\omega_i$ where ξ_i is the ith relative modal damping. This is used in the calibration procedure, and initially in the EMA to find the number of frequency lines necessary. The EMA frequency lines were selected based on the improved frequency selection [17] method for system identification with N4SID [18].

The calibration procedure initially obtains a parameter configuration that minimises the deviation between the FE model and the experimentally identified system. It is therefore important to obtain a good mathematical model from the measured data. For the configuration without mass loading all accelerometers were used in the identification, from 60 to 350 Hz with 20 states. Low and high frequency residual modes were added to the obtained system and the system's $\boldsymbol{B}$ and $\boldsymbol{C}$ matrices re-estimated with additional data, from chirp measurements for the low frequency region from 20 to 60 Hz and stepped sine data for the high frequency region from 350 to 470 Hz. This produced a system that accurately predicted both resonance and antiresonance behaviour. For the mass loaded configuration only the bushing accelerometers, accelerometer 1 to 8 in Fig. 7.4c, d, were used initially for the system identification due to the many local bushing modes identified from the measurements, not visible in other channels. Thus an identified system with 40 states from 40 to 225 Hz was identified for 24 channels (8 first triaxial accelerometers). A low and high frequency residual system pole was added to the system and additional data from a chirp test from 20 to 40 Hz and stepped sine test from 225 to 250 Hz was used to re-estimate the system's $\boldsymbol{B}$ and $\boldsymbol{C}$ matrices. In addition all 36 accelerometers, or 56 channels, throughout the frequency region from 20 to 250 Hz were used in the re-estimation for a system with more system outputs.

The channels for which the system identification model gave the poorest fit were used as validation data as raw data is used in the validation procedure. In Fig. 7.4, with x, y, and z components for the triaxial sensors, the calibration channels selected from the 56 available channels for the mass loaded configuration were $\{1_x, 1_z, 2_x, 2_z, 3_x, 3_z, 4_x, 4_z, 5_x, 5_z, 6_x, 6_y, 6_z, 7_z, 8_x, 8_z, 9_x, 11, 12, 16, 17, 18, 20, 21, 22, 25, 26, 28, 29, 33, 34, 35, 36\}$. For the configuration without mass loading the following set was selected $\{1_x, 1_y, 1_z, 2_x, 2_y, 2_z, 3_x, 3_y, 3_z, 4_x, 4_y, 4_z, 5_y, 5_z, 6_y, 6_z, 8_z, 10_x, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36\}$. Approximately 60% and 75% of the channels were used for calibration in the configuration with and without mass loaded boundaries, respectively.

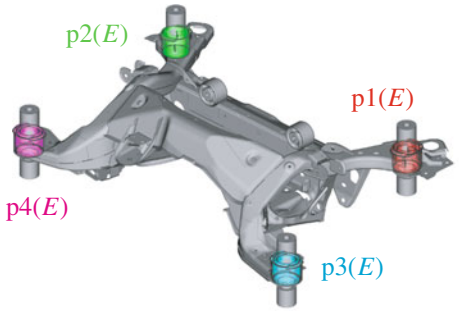
7.3.2 Finite Element Models

The rear subframe FE model, shown in Figs. 7.4 and 7.5, with and without mass loaded boundaries, consisted mainly of shell elements with the bushings modelled with an isotropic linear material model with solid elements, and consisted of over 600,000 degrees of freedom. The model was assumingly verified. The same FE model was used in the mass loaded boundary configuration where an additional heavy component was attached to the bushings. The additional component, shown in reality in Fig. 7.2c and as an FE model in the assembly in Fig. 7.2d, was designed to be simple to model with high accuracy. It is a simple cylinder with two different outer radii dimensions with a hole through its centre. It was made out of steel and each piece had a mass of 1 kg. Two such components were attached with bolts to each bushing of interest of the subframe, shown in Fig. 7.2d for one bushing, which had 6 bushings in total of which four were of interest in this study. MSC Nastran was used to establish the mass and stiffness matrices from the FE models.

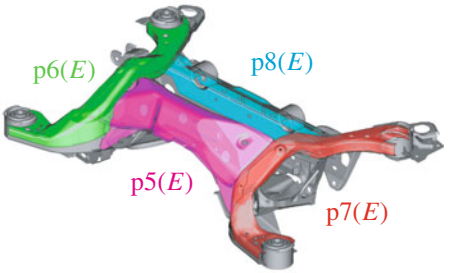
7.3.3 Parameter Selection

In the mass loaded boundary configuration only the rubber stiffness parameters were used in the calibration and all other model parameters were fixed to their nominal values. The parametrisation for the four bushings can be seen in Fig. 7.5a. For the configuration without mass loading 10 additional parameters were selected, in addition to the bushing parameters, which are shown in Fig. 7.5b–d. All other parameters, not highlighted, were fixed to their nominal values. The inverse Fisher information matrix (FIM) [30] was used in an identifiability analysis to fix parameters that would render the calibration problem unidentifiable. From this study it was found that parameter $p12$ should be fixed. Two parameter types were used, material stiffness and shell element thickness. In the deterministic calibration most parameters were constrained to approximately 25%, upper and lower bound, of their nominal setting, with some deviations.

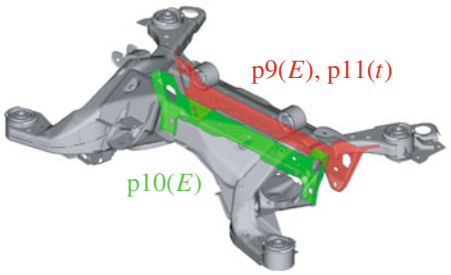
It is known that rubber is a nearly incompressible material and therefore with a Poisson's ratio approaching $\nu = 0.5$. With a value of ν very close to 0.5, linear material FE models become overly stiff [31]. It was found that ν was highly identifiable but did not give a smooth calibration metric and could thus not be used as a free parameter. A value of $\nu = 0.49$, as a best guess, was therefore fixed in all the model updating procedures.



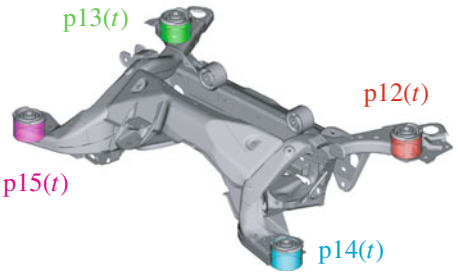
(a) Bushing rubber stiffness parameters p1(E), p2(E), p3(E) and p4(E)



(b) Steel stiffness parameters p5(E), p6(E), p7(E) and p8(E)



(c) Steel stiffness p9(E), p10(E) and thickness p11(t) parameters



(d) Steel thickness parameters p12(t), p13(t), p14(t) and p15(t)

Fig. 7.5 Parametrisation of the FE models with E representing stiffness and t thickness, in (**a**) for the mass loaded configuration and in (**b**), (**c**) and (**d**) for the configuration without mass loading. Parameters $p1$ to $p4$ were also parametrised in the configuration without mass loading

7.4 Calibration and Validation Results

The nominal and updated parameters from the deterministic calibration, and the mean value and coefficient of variation (COV, %) from the bootstrapping calibration are shown in Table 7.1 with the table order illustrating the calibration order, i.e. mass loaded configuration first followed by the configuration without mass loading up to 265 and 350 Hz in a second and final step, respectively. Thus, three calibrations were performed in total. Empty fields (–) denote fixed parameters in that particular calibration. In an initial step the mass loaded configuration was calibrated from 40 to 225 Hz so that the bushings' rubber stiffness could be estimated, i.e. parameters $p1$ to $p4$. All other parameters were fixed to their nominal values, where parameters $p12$ to $p15$ were set to 1.20 mm. These four parameters were manually updated to a nominal setting of 1.40 in the calibration of the configuration without mass loading. The configuration without mass loading was first calibrated from 60 to 265 Hz, with updated bushing parameters, by updating parameters $p5$ to $p11$. Updated parameters $p5$ to $p11$ were then used as nominal parameters in the final calibration from 60 to 350 Hz where parameters $p1$, $p2$, $p10$, $p13$, $p14$ and $p15$ were updated. Parameter $p12$ was kept fixed in all three calibrations. Throughout the results section the deterministic calibration parameters $\boldsymbol{p}^*$ are used in the results, denoted after their respective model as in Table 7.2.

In Table 7.1 it can be seen that the mean value of the bootstrapping results, which are based on $n_b = 20$ repeated calibrations, differ somewhat from the deterministic calibrated estimates. This is expected as the raw data is used in the bootstrapping calibration, which have significant noise levels around the low frequency anti-resonances for some channels, see Fig. 7.9b. Furthermore, the identified model used in the deterministic calibration does not perfectly fit test data in all channels, especially around the anti-resonances, and a difference is therefore to be expected. The parameters are stable, as indicated by the COV in Table 7.1, with parameters $p9$, $p10$ and $p11$ for calibration FE_{265} slightly higher than the others, approximately from 3% to 6%. Running at least $n_b = 100$ bootstrap calibrations is suggested in [28], which was not feasible due to the large datasets. It is likely that the variation would have decreased if n_b had been increased.

It can be seen in Table 7.1 that the two calibrated parameters $p1$ and $p2$ in the mass loaded configuration, visualised in Fig. 7.5a, are considerably stiffer than $p3$ and $p4$. This is a consequence of making the FE model fit experimental data so that the modes shown in Fig. 7.7b–j would correspond to what was found experimentally. This is possibly a model inadequacy given that the bushings are different for these two set of parameters, i.e. $p1$ and $p2$ correspond to bushing 2 shown in Fig. 7.2b while $p3$ and $p4$ correspond to bushing 1 shown in Fig. 7.2a. Bushing 2 has an additional mass attached to it, and is less stiff

Table 7.1 Nominal parameters, denoted *Nom.*, along with three calibration results

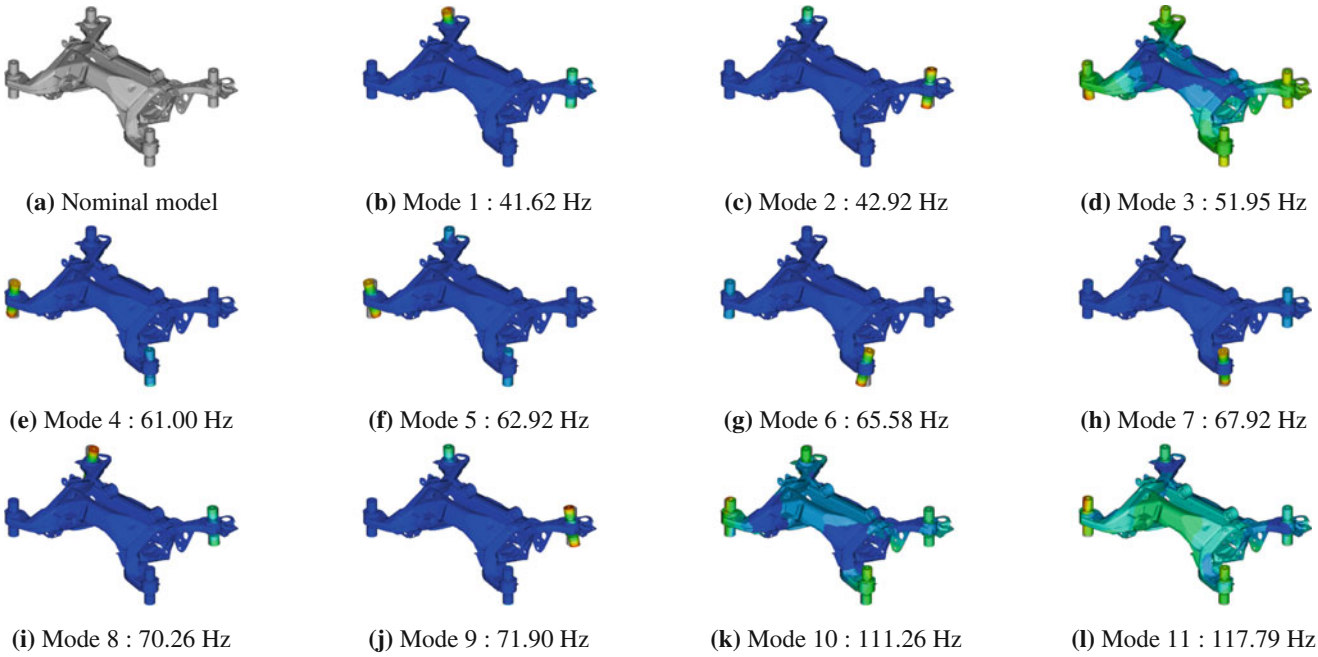
#	Parameter	*Nom.*	FE_{mass}	FE_{265}	FE_{350}	FE_{mass}^{μ}	FE_{mass}^{COV}	FE_{265}^{μ}	FE_{265}^{COV}	FE_{350}^{μ}	FE_{350}^{COV}
p1	E [MPa]	5.00	4.01	4.01	4.62	3.67	0.72	–	–	4.98	0.39
p2	E	5.00	3.82	3.82	7.01	3.48	0.85	–	–	5.25	0.11
p3	E	5.00	2.83	2.83	5.83	2.69	0.94	–	–	–	–
p4	E	5.00	2.41	2.41	5.41	2.94	1.19	–	–	–	–
p5	E [GPa]	210.00	210.00	192.32	192.32	–	–	196.44	0.99	–	–
p6	E	210.00	210.00	208.01	208.01	–	–	198.91	0.13	–	–
p7	E	210.00	210.00	240.73	240.73	–	–	228.60	0.09	–	–
p8	E	210.00	210.00	222.41	222.41	–	–	249.11	1.73	–	–
p9	E	210.00	210.00	250.82	250.82	–	–	264.38	3.86	–	–
p10	E	210.00	210.00	166.57	204.78	–	–	162.60	3.10	189.15	0.10
p11	t [mm]	1.80	1.80	1.46	1.46	–	–	1.35	5.81	–	–
p12	t	1.40	1.20	1.40	1.40	–	–	–	–	–	–
p13	t	1.40	1.20	1.40	3.67	–	–	–	–	2.56	0.71
p14	t	1.40	1.20	1.40	1.81	–	–	–	–	1.75	1.12
p15	t	1.40	1.20	1.40	1.61	–	–	–	–	1.80	0.74

Calibrated parameters for the mass loaded configuration denoted FE_{mass}, configuration without mass loading with calibration up to 265 Hz denoted FE_{265} and up to 350 Hz denoted FE_{350}. The mean value and coefficient of variation (%) of the bootstrapping procedure is denoted with superscript μ and COV, respectively, i.e FE_{mass}^{μ} and FE_{mass}^{COV}. Empty fields (–) represent fixed parameters in that particular calibration. In this table E represent material stiffness and t denote thickness

Table 7.2 First 11 flexible eigenfrequencies (Hz) for the two different configurations

Mode	Exp_{mass}	FE^{nom}_{mass}	FE_{mass}	Δ_{mass} [%]	Exp	FE^{nom}	FE_{265}	Δ_{265} [%]	FE_{350}	Δ_{350} [%]
1	40.95	46.45	41.62	1.64	76.82	78.20	76.86	0.05	76.90	0.10
2	–	48.16	42.92	–	160.54	158.18	160.53	0.01	160.63	0.06
3	52.07	53.92	51.95	0.23	193.08	194.63	193.13	0.03	192.62	0.24
4	59.12	78.60	61.00	3.18	195.25	196.97	195.40	0.08	196.24	0.51
5	63.21	80.03	62.92	0.46	204.94	207.37	206.29	0.66	205.59	0.32
6	66.07	82.50	65.58	0.74	210.84	208.84	210.84	0.00	210.71	0.06
7	–	84.96	67.92	–	240.93	241.10	239.62	0.54	240.13	0.33
8	71.84	86.42	70.26	2.20	254.36	254.27	253.34	0.40	255.99	0.64
9	74.64	88.03	71.90	3.67	–	293.76	268.52	–	290.45	–
10	110.59	116.27	111.26	0.61	306.89	304.23	278.47	9.26	310.18	1.07
11	118.19	126.17	117.79	0.34	320.09	316.98	304.01	5.02	315.46	1.45
Mass	37.18	37.01	37.01		26.80	26.73	26.51		26.87	

Experimental eigenfrequencies for the configuration with and without mass loading denoted Exp_{mass} and Exp, respectively. Nominal FE models for the two configurations denoted with superscript *nom*. FE_{mass} denote mass loaded calibration. FE_{265} denote calibration up to 265 Hz and FE_{350} calibration up to 350 Hz for the configuration without mass loading. Δ_{mass} (%) represent the relative difference between FE_{mass} and Exp_{mass}, and Δ_{265} (%) and Δ_{350} (%) the relative difference between Exp and FE_{265} and FE_{350}, respectively. Masses for the experimental set-ups and the different models is also shown



(a) Nominal model **(b)** Mode 1 : 41.62 Hz **(c)** Mode 2 : 42.92 Hz **(d)** Mode 3 : 51.95 Hz

(e) Mode 4 : 61.00 Hz **(f)** Mode 5 : 62.92 Hz **(g)** Mode 6 : 65.58 Hz **(h)** Mode 7 : 67.92 Hz

(i) Mode 8 : 70.26 Hz **(j)** Mode 9 : 71.90 Hz **(k)** Mode 10 : 111.26 Hz **(l)** Mode 11 : 117.79 Hz

Fig. 7.6 MAC correlation between the experimentally identified system and the nominal FE model FE^{nom} in (**a**) and calibrated FE model FE_{350} in (**b**), respectively

in one direction. It is also possible that the bushing parameters to some extent acted as surrogate parameters due to other model parameters being fixed.

In Table 7.2 the eigenfrequencies of the first eleven flexible modes are shown for the configuration with and without mass loading, along with their masses. It can be seen that for the mass loaded configuration the nominal results are very different from what is experimentally identified. It should be noted that mode 2 is very close to mode 1, in the experimental data, and could not be identified, i.e. that particular mode has mode multiplicity 2, show in Fig. 7.7b, c. The same holds for mode 6 and 7, shown in Fig. 7.7g, h. FE results indicate to that mode 4 and 5, and 8 and 9 are also paired, as seen from their modeshapes shown in Fig. 7.7. The calibrated FE_{mass} has a much better correspondence to experimental data compared to the nominal model FE^{nom}_{mass} which is also indicated by the deviation metric. The deviation metric between the experimental data Exp_{mass} and the nominal FE model FE^{nom}_{mass} is 0.90 and 1.02 for the calibration and validation, respectively. The updated model has a deviation metric of 0.44 and 0.50 for the calibration and validation, respectively. Thus an improvement of 51% is seen in both the calibration and validation metrics. Note that a deviation of zero indicates perfect fit, but there is no upper bound. Therefore, even if there exist a model inadequacy the updated model makes good predictions and the bootstrapping results show a low variation in the parameters, as seen in Table 7.1, and hence the updated parameters are considered reliable.

From Table 7.2 it looks as if the nominal configuration without mass loading, FE^{nom}, correlates well with the experimental model, but from the modal assurance criterion (MAC) [32] analysis in Fig. 7.6a it can be seen that there is a mode switch between mode 3 and 4, and mode 5 and 6. Also, some modes show a very low correlation and the last two modes have a high cross-correlation. The updated model FE_{350}, for which the MAC analysis is shown in Fig. 7.6b, shows a much better MAC correlation, even for the modes above 300 Hz. It should be noted that there is a mode in the FE results, not identified in the experiments around 290 Hz. It is shown in Fig. 7.8j and it can be seen that it is a very local mode in one bushing. It is therefore not controllable from the input position, and due to its high damping it would be very hard to identify from the raw data. Further in Fig. 7.8 it can be seen that in the last two modes, over 300 Hz, the bushings are considerably active. In order to capture the behaviour of the two modes above 300 Hz the stiffness parameter in the rubber bushings had to be increased. Thus parameters $p3$ and $p4$ were manually increased by 2 MPa and $p1$ and $p2$ used as free parameters in the last calibration step, FE_{350}. An explanation for this behaviour is that rubber is frequency dependent, i.e. its stiffness increases with frequency [3]. In the mass loaded configuration the bushing parameters are estimated towards a lower frequency compared to the calibration without mass loading, and therefore to capture the behaviour above 300 Hz for the configuration without mass loading the stiffness must be increased. The intermediate calibration step, FE_{265} does not show a good correlation for modes above the 8th flexible modes, as seen in Table 7.2, due to them not being included in the model updating. Because rubber elasticity is frequency dependent [3], not updating the rubber parameters but including the modes which activate the bushings in a

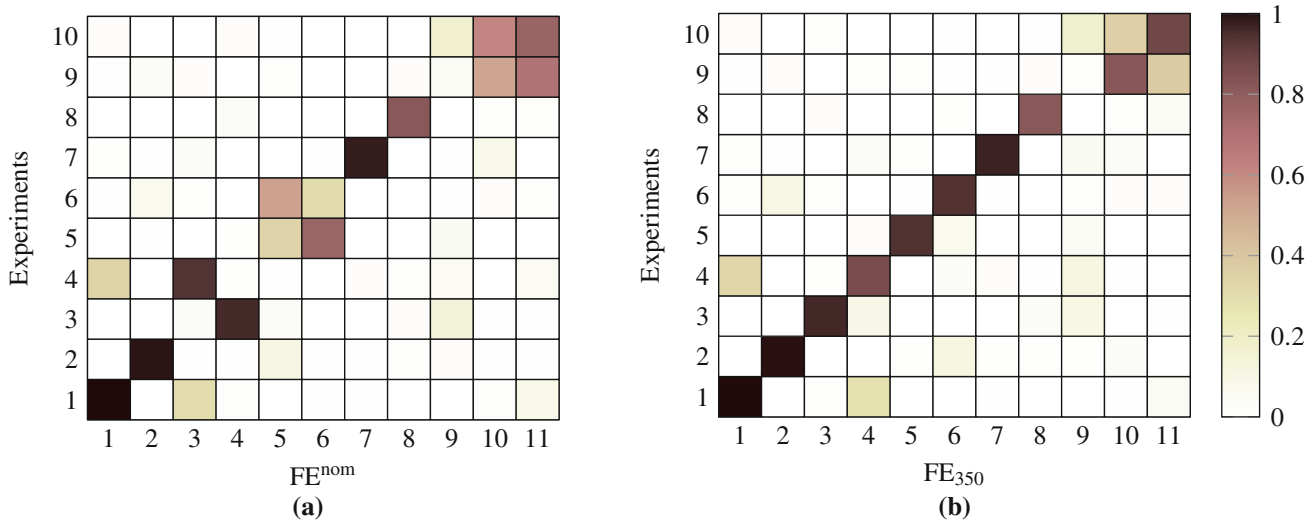
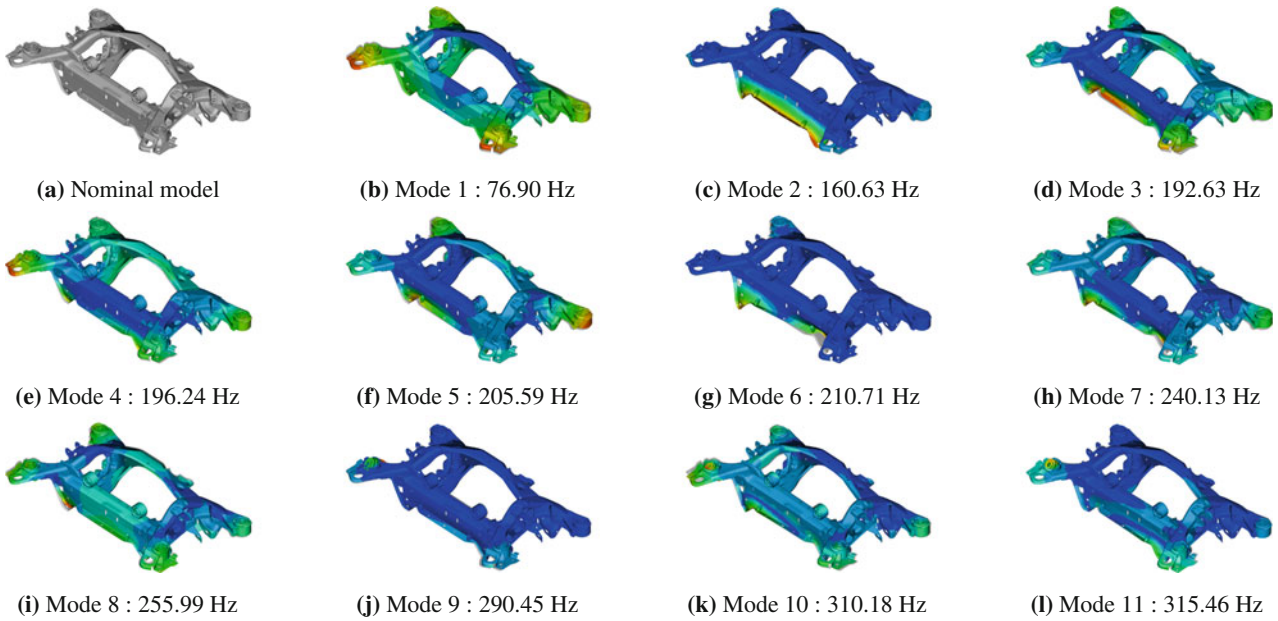


Fig. 7.7 The mass loaded configuration (**a**), along with the 11 first calibrated flexible modes, from mode 1 (**b**) to mode 11 (**l**). *Blue* indicate small modal motion and *red* large motion



(a) Nominal model **(b)** Mode 1 : 76.90 Hz **(c)** Mode 2 : 160.63 Hz **(d)** Mode 3 : 192.63 Hz

(e) Mode 4 : 196.24 Hz **(f)** Mode 5 : 205.59 Hz **(g)** Mode 6 : 210.71 Hz **(h)** Mode 7 : 240.13 Hz

(i) Mode 8 : 255.99 Hz **(j)** Mode 9 : 290.45 Hz **(k)** Mode 10 : 310.18 Hz **(l)** Mode 11 : 315.46 Hz

Fig. 7.8 The configuration without mass loading (**a**), along with the 11 first calibrated flexible modes for model FE_{350}, from mode 1 (**b**) to mode 11 (**l**). *Blue* indicate small modal motion and *red* large motion

frequency region not of interest would have caused the other parameters to act as surrogate parameters, compensating for the deviation above 300 Hz. The deviation metric for the intermediate model FE_{265} was improved with 60% for the calibration from 0.91 to 0.38. The validation metric was improved with 23% from 1.25 to 0.97. It should be noted that the nominal model from which FE_{265} was updated included the updated bushing parameters, and that this is reflected in the deviation metric. The deviation metric for the last model calibration FE_{350} was improved with 71% for the calibration from 0.88 to 0.26, with nominal model FE_{265}, but manually updated parameter $p3$ and $p4$. The validation metric was improved with 33% from 1.20 to 0.81. The improvement in the validation metric is notably lower compared to the calibration metric. One possible reason is that poorly estimated channels were placed in the validation data set.

In Fig. 7.9a one of the better channels is shown for raw data, identified model, nominal FE model, calibrated FE model up to 265 and 350 Hz. Figure 7.9b depicts one of the poorer channel estimates.

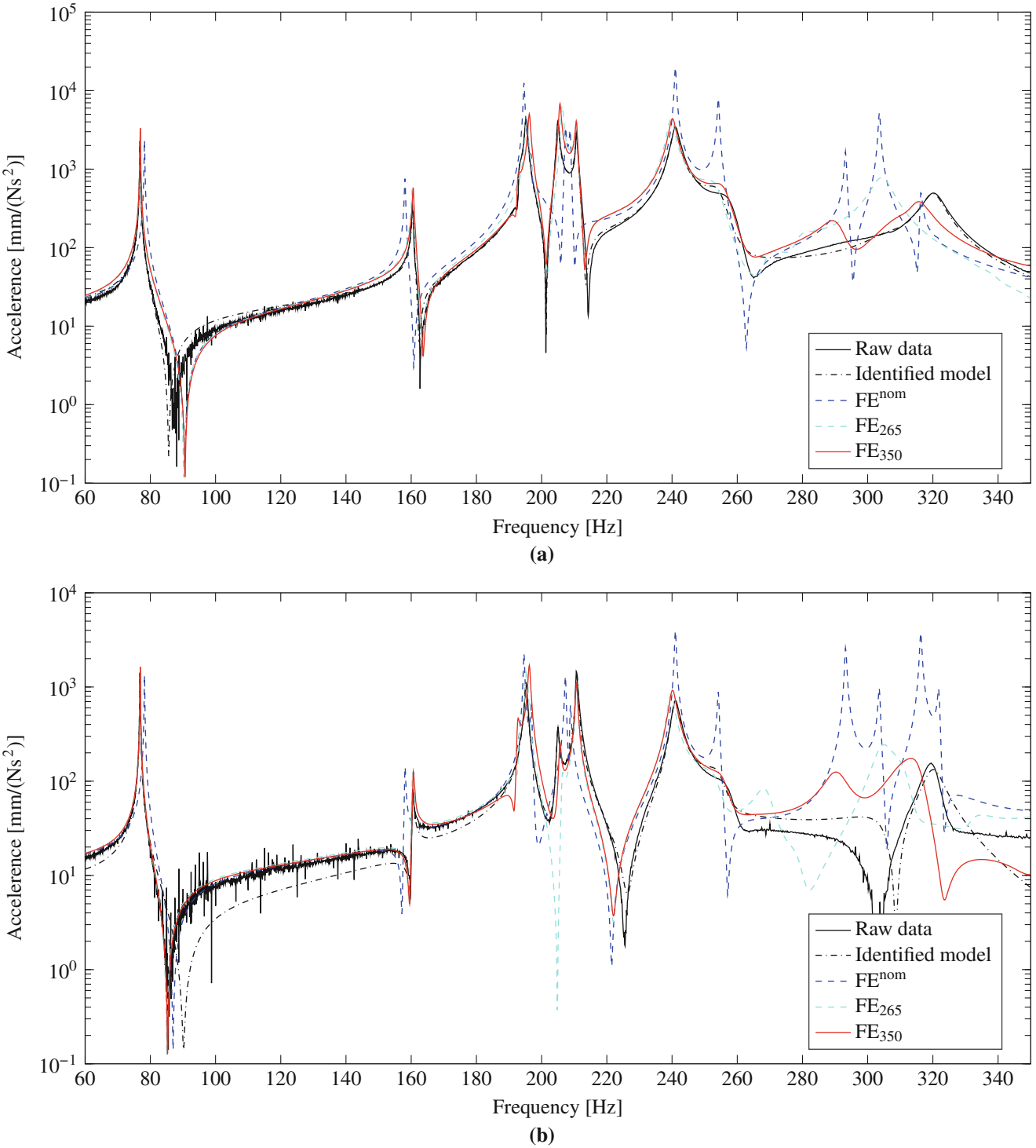


Fig. 7.9 Frequency response functions of configuration without mass loading with raw data (*black*), identified model (*dot dashed black*), FE^{nom} (*dashed blue*), FE_{265} (*dashed cyan*) and FE_{350} (*red*). In (**a**) the input is at position 24 with output at position 4_x and in (**b**) the input is at position 24 with output at position 30

7.5 Conclusions

This paper considers model updating and parameter uncertainty quantification of a rear subframe of a Volvo XC90 (2015), with parametrised bushings to gain physical insight in the structure at hand. A mass loaded boundary configuration was utilised to make more accurate estimates of the bushing's rubber stiffness such as they would be in a coupled configuration. A good estimate of the bushing parameters have been achieved from the mass loaded configuration. This parameter setting is thought to provide a more realistic representation of how the rear subframe will behave when coupled to the BIW, and should therefore be kept in the configuration without mass loading even though it has been shown that such an FE model produces poor results for frequencies above 300 Hz, where local bushing modes become active. The rubber stiffness must be increased if a good correlation for modes above 300 Hz for the configuration without mass loading is sought. Future work includes coupling the rear subframe to a BIW and assess the predictive capabilities of such an assembly from CAE. The frequency region of interest in such a study will be lower than 300 Hz, and therefore the mass loaded configuration should provide a better rubber parameter estimation.

It is likely that the bushings could have been modelled with bushing elements instead of solid elements, but because physical insight into the structure was sought a solid model was used, and the added degrees of freedom were negligible from a computational perspective. Further it would have been of interest to use Poison's ratio as a free parameter in the model updating procedure. The model was found to be very sensitive to it, but as it gave a non-smooth calibration metric it had to be fixed.

Model updating towards the two other measured components will be performed in a future study. It will be of interest to verify how well the parametrisation can capture the component variability, and asses the parameter uncertainty towards other experimental data. It is also of interest to find the correct bushing parameters for the other components, and asses how much they will differ from the component presented in this study.

The calibration, validation and bootstrapping has been performed using the open source MATLAB program FEMcali, downloadable from Mathwork's webpage at www.mathworks.com.

Acknowledgements Volvo Car Corporation is gratefully acknowledged for providing the funding for this paper.

References

1. Craig, R.R.J., Kurdila, A.J.: Fundamentals of Structural Dynamics, 2nd edn., p. 744. Wiley, Hoboken (2006)
2. Geradin, M., Rixen, D.J.: Mechanical Vibrations: Theory and Application to Structural Dynamics, 3rd edn., p. 616. Wiley, New York (2015)
3. Jones, D.I.G.: Handbook of Viscoelastic Vibration Damping, 1st edn., p. 410. Wiley, Chichester (2001)
4. Ewins, D.J.: Modal Testing: Theory Practice and Application, 2nd edn., p. 576. Wiley-Blackwell, Philadelphia (2000)
5. Abrahamsson, T.J.S., et al.: Calibration and validation of a car subframe finite element model using frequency responses. In: Mains, M. (ed.) Topics in Modal Analysis, vol. 10. Conference Proceedings of the Society for Experimental Mechanics Series, pp. 9–22. Springer, New York (2015)
6. Abrahamsson, T., et al.: Calibration and cross-validation of a car component using frequency response functions and a damping equalization technique. In: 26th International Conference on Noise and Vibration Engineering, ISMA 2014, Including the 5th International Conference on Uncertainty in Structural Dynamics, USD 2014, pp. 2625–2640 (2014)
7. Gibanica, M., Abrahamsson, T.J.S., Olsson, M.: Calibration, validation and uncertainty quantification of nominally identical car subframes. In: Atamturktur S., et al. (eds.) Model Validation and Uncertainty Quantification, vol. 3. Conference Proceedings of the Society for Experimental Mechanics Series, pp. 315–326. Springer, New York (2016)
8. Friswell, M.I., Mottershead, J.E.: Finite Element Model Updating in Structural Dynamics. Solid Mechanics and Its Applications, vol. 38. Springer, Dordrecht (1995)
9. Abrahamsson, T.J.S., Kammer D.C.: Finite element model calibration using frequency responses with damping equalization. Mech. Syst. Signal Process. **6263**, 218–234 (2015)
10. Khorsand Vakilzadeh, M.: Stochastic model updating and model selection with application to structural dynamics. Doctoral thesis, Chalmers University of Technology (2016)
11. Yaghoubi, V.: System identification of large-scale linear and nonlinear structural dynamic models. Doctoral thesis, Chalmers University of Technology (2016)
12. Larsson, K.-J., et al.: Calibration and cross-validation of a car component model using repeated testing. In: Atamturktur, H.S., et al. (eds.) Model Validation and Uncertainty Quantification, vol. 3. Conference Proceedings of the Society for Experimental Mechanics Series, pp. 339–350. Springer, New York (2015)
13. Echaniz Granado, I.: Model calibration of a vehicle tailgate using frequency response functions. http://publications.lib.chalmers.se/records/fulltext/218958/218958.pdf (2015) (visited on 10/22/2015)
14. Kammer, D.C.: Sensor placement for on-orbit modal identification and correlation of large space structures. J. Guid. Control. Dyn. **14**(2), 251–259 (1991)

15. Kammer, D.C.: Sensor set expansion for modal vibration testing. Mech. Syst. Signal Process. **19**(4), 700–713 (2005)
16. Gibanica, M., Abrahamsson, T.J.S., Kammer, D.C.: Redundant information rejection in sensor localisation using system gramians. In: Mains, M., (ed.) Topics in Modal Analysis & Testing, vol. 10. Conference Proceedings of the Society for Experimental Mechanics Series, pp. 325–333. Springer, New York (2016)
17. Vakilzadeh, M.K. et al.: Experiment design for improved frequency domain subspace system identification of continuous-time systems. In: IFAC-PapersOnLine 17th IFAC Symposium on System Identification SYSID 2015, Beijing, 19–21 October 2015, vol. 48, No. 28, pp. 886–891 (2015)
18. McKelvey, T., Akcay, H., Ljung, L.: Subspace-based multivariable system identification from frequency response data. IEEE Trans. Autom. Control **41**(7), 960–979 (1996)
19. Cunningham, J.R., Ivey, D.G.: Dynamic properties of various rubbers at high frequencies. J. Appl. Phys. **27**(9), 967–974 (1956)
20. Kennedy, M.C., O'Hagan, A.: Bayesian calibration of computer models. J. R. Stat. Soc. Ser. B (Stat. Methodol.) **63**(3), 425–464 (2001)
21. Simoen, E., De Roeck, G., Lombaert, G.: Dealing with uncertainty in model updating for damage assessment: a review. Mech. Syst. Signal Process. **5657**, 123–149 (2015)
22. Ljung, L.: System Identification: Theory for the User. 2nd edn., p. 672. Prentice Hall, Upper Saddle River (1999)
23. Levenberg, K.: A method for the solution of certain non-linear problems in least squares. Q. Appl. Math. **2**(2), 164–168 (1944)
24. Marquardt, D.W.: An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. **11**(2), 431–441 (1963)
25. Andreasson, N., Evgrafov A., Patriksson, M.: An Introduction to Continuous Optimization. 2nd edn., p. 498. Studentlitteratur AB, Lund (2013)
26. Mckay, M.D., Beckman, R.J., Conover, W.J.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics **42**(1), 55–61 (2000)
27. Yaghoubi, V., Abrahamsson, T.: The modal observability correlation as a modal correlation metric. In: Allemang, R., et al. (eds.) Topics in Modal Analysis, vol. 7. Conference Proceedings of the Society for Experimental Mechanics Series, pp. 487–494. Springer, New York (2014)
28. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning Data Mining, Inference, and Prediction. Springer Series in Statistics. Springer, New York (2009)
29. Golub, G.H., Loan, C.F.V.: Matrix Computations, p. 738. The Johns Hopkins University Press, Baltimore (2012)
30. Kay, S.M.: Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory, 1st edn., p. 625. Prentice Hall, Englewood Cliffs (1993)
31. Bathe, K.J.: Finite Element Procedures in Engineering Analysis, p. 736. Prentice Hall, Englewood Cliffs (1982)
32. Allemang, R.J., Brown, D.L.: A correlation coefficient for modal vector analysis. In: Proceedings of the 1st IMAC. International Modal Analysis Conference (1982)

Chapter 8
Vibroacoutsic Modelling of Piano Soundboards through Analytical Approaches in Frequency and Time Domains

B. Trévisan, K. Ege, and B. Laulagnet

Abstract The vibratory behavior and radiation of complex structures are a real challenge for many industrial domains. The increasing requirements of users and manufacturers justify the interest of the scientific community about this subject, particularly about ribbed structures. Initially, the design of such a structure is led by structural reasons and offers in the same time the possibility to reduce the weight and to reinforce the conception. Thus, they are common in many industrial domains, but also building and crafting sector. Among them, mention automotive, shipbuilding and aerospace industries or musical instruments too (see Fig. 8.1).

Keywords Orthotropic materials • Vibroacoustic of non-rectangular ribbed structures • Frequency and time domains • Coupled systems • Piano soundboard

However, their conception makes difficult the understanding of their vibroacoustical behavior contrary to simplest structures as beams or plates that are now well described. Since several decades, some studies tried to find elements of response to industrial needs and now allow understanding the behavior of more complex structures but remain far enough from real industrial case: finite rectangular ribbed plate, infinite ribbed cylindrical shells for examples. Moreover, the democratization of composite materials offers new possibilities of improvement, which is in good agreement with industrial considerations (like about reducing the weight of vehicles for example). So, taking into account the behavior of these new materials in addition to structures designs becomes a real challenge. For example, interactions between layers in the case of laminated composites or also honeycomb panels due to their reduced weight.

Actually, computing resources allow to describe various designs and to take into account material specificities with accurate results. However, these purely numerical methods need important hardware and imply long time computations for the most complex structures. Thus, they do not allow a step of comprehension of the phenomena and are rather adapted to evaluate performances after a first phase of design. Hence, it misses rules that allow to indicate the good practices to adopt as well as to anticipate their performances during the first outlines of design and at the time of optimization. Their modeling through lower cost methods remains a major challenge in order to do parametrical studies and so, highlight main tendencies that can be used in various domains. In that way, analytical approaches are particularly well adapted but imply simplified modeling as presented in [1]. However, simplifications applied to the method do not impact the main tendencies of the results. We are interested in orthotropic non-rectangular plates. To have a general method, the principal axis of orthotropy can make an angle with the edges of the structure. In the other case, we talk about a special orthotropy. Moreover, we consider some stiffeners/superstructures tied to the plate that can be oriented in the directions of plate's Young moduli, i.e. in two perpendicular directions.

The piano soundboard is a typical example of such a structure. Musical instruments making is a particularly interesting domain considering constrains due to conception and their consequences on the perceived sound. In the case of piano, many parameters must be considered from the wood used to the own design of the instrument. Among the mechanical aspects, we can quote: the design of the soundboard itself that have been studied at several times, the strings, the hammer and also interactions between them.

Thus, this is a complete field of investigations that involves subjective and perceptive aspects, which still are uncommon in industrial domains. The piano soundboard is the centerpiece of the instrument because it allows an effective acoustic radiation. Indeed, the strings are too small to radiate by themselves and they are tied to the soundboard through the bridges, exchanging energy. So, the soundboard has an essential function, justifying the several studies about it that have highlighted,

B. Trévisan (✉) • K. Ege • B. Laulagnet
Univ Lyon, INSA-Lyon, LVA EA677, 69621, Villeurbanne, France
e-mail: benjamin.trevisan@insa-lyon.fr

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_8

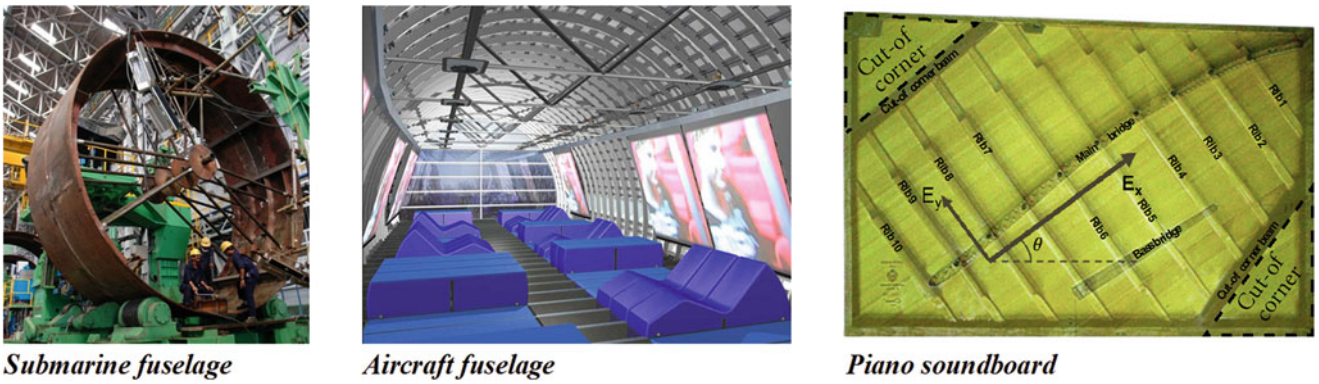


Fig. 8.1 Examples of ribbed structures in various domains

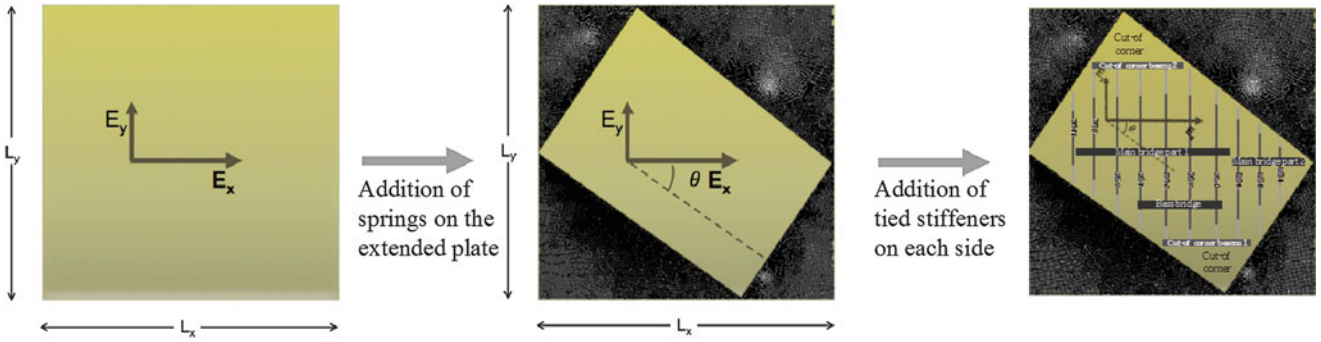


Fig. 8.2 Schematics of method outlines: adding of springs to describe non-rectangular edges and then, stiffeners in both directions

all the complexity of its vibroacoustic behavior: its non-rectangular edges, its crown, the use of wood, stiffeners, tapering or the downbearing induced by the strings are particular characteristics of its design. Indeed, its empirical conception is led by structural and dynamical constrains and the rules used by piano makers do not allow to anticipate the consequences of any changes on the quality of the perceived sound. Moreover, with limited financial resources, attempts of innovation are risky.

Despite that, piano makers often report a non-satisfaction about the timber and a good compromise sustain/radiated power is difficult to obtain in the treble zone. The major scientific studies focus on frequency response of the soundboard (modal shapes, mobility at the bridge for the most common) and cannot represent the perceptive aspect of the problem. To answer to purely musical questions, it is relevant to introduce the concept of coupled systems. So, by coupling the soundboard to a string and solving it in time domain, it is possible to listen to the sound produced and appreciate the difference between two soundboards. This way of solving offers new perspective of analysis and new indicators, which can be employed to complete frequency analysis.

In view to solve these issues, a modeling is developed using a variational approach that take its inspiration in [2, 3] and allows to describe orthotropic non-rectangular ribbed plates. This is done using an extended simply supported plate with special orthotropy and a modal decomposition on its modes. By an addition of several springs in order to block the transversal displacement, it is possible to describe any edges the user want (not limited to soundboard and musical domain) and angle of otrhotropy (see Fig. 8.2). Then, the stiffeners are added as straight stiffeners oriented in the directions of the Young moduli of the extended plate (see Fig. 8.2). The method takes into account the offset from the middle plane of the plate. By this way, we get a matrix formulation of the problem, which can be solved in frequency and time domains.

In the frequency domain, the method we propose for the acoustic radiation is an alternative to purely numerical method as for example finite element boundary method, Rayleigh integral or Perfectly Matched Layers. In that way, we calculate the acoustic radiation of the structure using the radiation impedances of the extended rectangular un-ribbed baffled plate, which have been studied several times [4, 5], but remains limited to rectangular ribbed plate [3]. It is now extended to the case of non-rectangular structure and finally, the method allows for example: solving the eigen-problem and so classified modal shapes into four particular families of modes (see Fig. 8.3); solving forced response of a ribbed structure with or without fluid/structure interactions and evaluate the acoustical radiation easily while being representative of real phenomenon on a

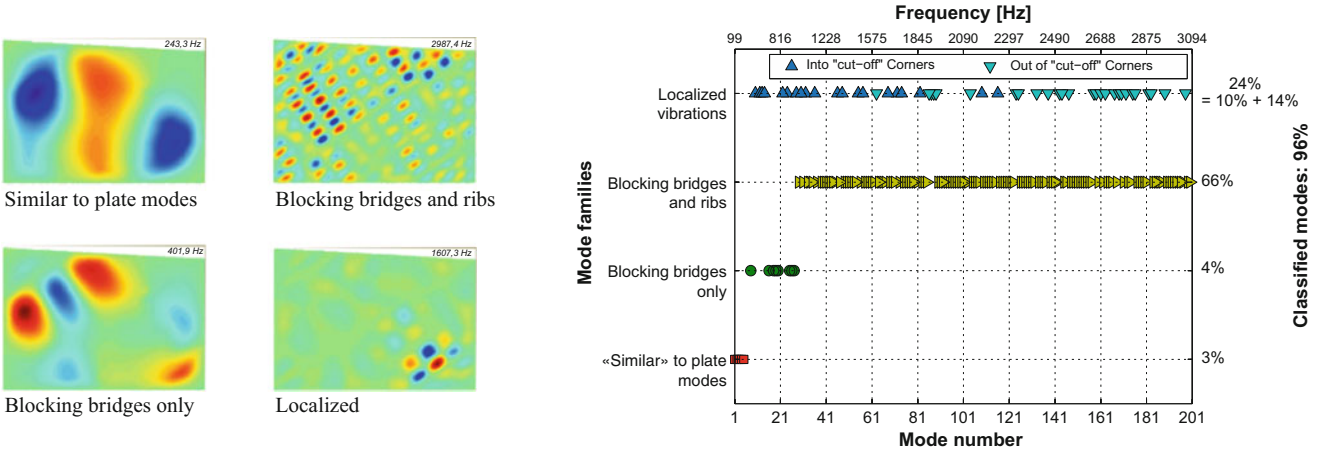


Fig. 8.3 Classification of modal shapes into different families

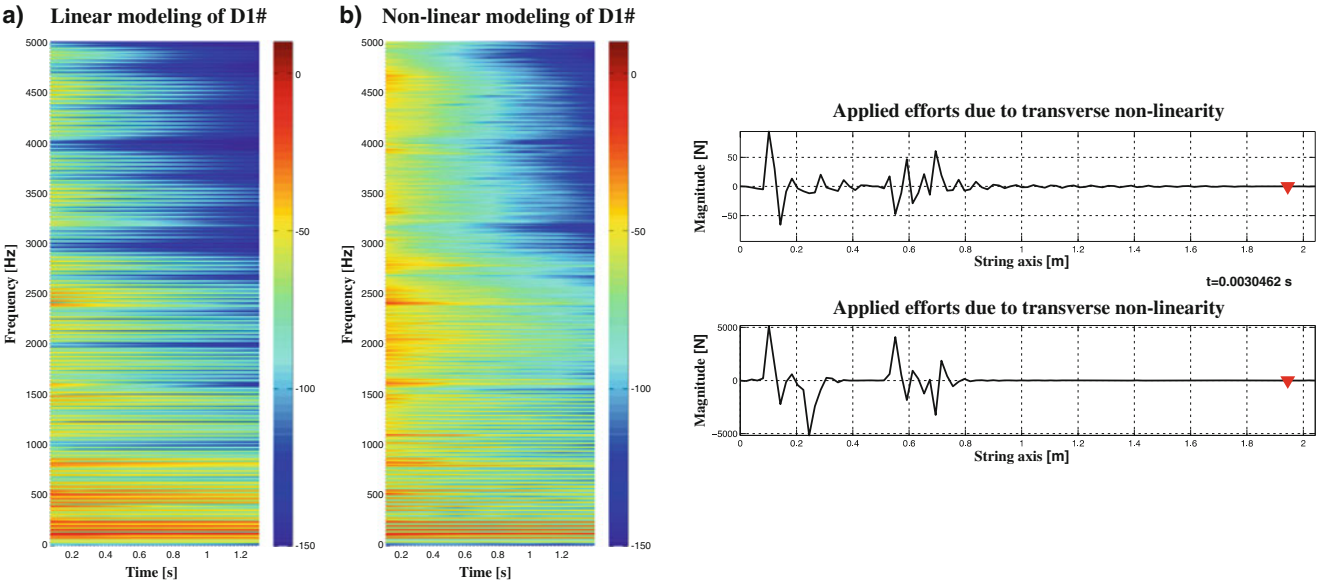


Fig. 8.4 Influence of the non-linearities into the string on the D1# radiated pressure: non-linearities induce efforts in the string that contribute to energy spectrum, especially in high frequencies.

large frequency band. Thus, it is simple to make parametrical studies and so to determine the main tendencies in order to get the rules useful to improve the vibro-acoustic behaviour of piano soundboards and more generally to any non-rectangular orthotropic ribbed plates.

As mentioned previously, perceptive aspects lead musical questions and we focus on strings/soundboard interactions. To answer these questions, the soundboard model is coupled to strings at the bridges and solved in time domain. The strings are modeled with transversal and longitudinal waves and it is also possible to take into account geometrical non-linearity of the strings. In such a calculi, the interaction efforts that ensure the continuity of displacements at the interface are unknown factors, as impact or friction problems [6]. That allows evaluating the influence of any structural changes on the coupling between the sub-systems. Then, it is possible to know the influence of the damping of the wood, the dead length of strings on the perceived sound and to highlight many coupling phenomenon and energy exchanges that were not easy to anticipate. We also show the influence of non-linearity (see Fig. 8.4), which are considered as small and so allow keeping notions of modes, on the timber of the instrument injecting several short stresses in the strings.

References

1. Trévisan, B., Ege, K., Laulagnet, B.: Vibroacoustics of orthotropic plates ribbed in both directions: application to stiffened rectangular wood panels. J. Acoust. Soc. Am. **139**(1), 227–246 (2016)
2. Guyader, J.-L., Laulagnet, B.: Sound radiation by finite cylindrical ring stiffened shells. J. Sound Vib. **138**(2), 173–191 (1990)
3. Guyader, J.-L., Laulagnet, B.: Structural acoustic radiation prediction: expending the vibratory response on a functional basis. Appl. Acoust. **43**(3), 247–269 (1994)
4. Wallace, C.E.: Radiation resistance of a rectangular panel. J. Acoust. Soc. Am. **51**(3), 946–952 (1972)
5. Stepanishen, P.R.: Modal coupling in the vibration of fluid-loaded cylindrical shells. J. Acoust. Soc. Am. **71**(4), 813–823 (1982)
6. Laulagnet, B.: Rattle noise generated by a beam tip impacting a baffled plate. Acta Acoust. **95**(1), 1–9 (2009)

Chapter 9
Combined Experimental and Numerical Investigation of Vibro-Mechanical Properties of Varnished Wood for Stringed Instruments

Sarah Louise Lämmlein, David Mannes, Francis Willis Mathew Schwarze, Ingo Burgert, and Marjan Sedighi Gilani

Abstract The obvious purpose of varnishing a stringed instrument is the protection against wear and relative humidity changes as well as enhancing its aesthetic appearance. Besides, it is known that varnishing changes the acoustic properties of the wood [1]. Unlike studies on the chemical compositions of old varnishes, in searching for the "secret" of well-known historical instruments [2], the vibro-mechanical properties of varnished wood have been less studied. Nevertheless, the vibrational and mechanical properties of wood are influenced after varnishing [3]. Vibrational experiments show that varnishing changes the specific stiffness and internal damping of wood. Therefore, varnishing also influences acoustic properties that are commonly used to evaluate the sound quality of an instrument such as speed of sound, characteristic impedance, sound radiation coefficient and emission ratio.

For the investigation of these temperature and humidity dependent properties, we propose a combined experimental and numerical analysis approach. X-ray microtomography images reveal density differences at the interface of wood and varnish, depending on its composition and application procedure. In combination with neutron imaging, this approach is used to evaluate the extent of varnish penetration and that of other particles (e.g. pumice powder) into the porous wood structure. Further, we study the vibrational and mechanical properties at the material to instrument scale. Additionally, numerical analysis improves the material property determination and helps evaluating the combination of different wood species and varnishes. Our aim is that these results will contribute to the scientific understanding of the variations in mechanics, vibrations and sorptivity of varnished wood and facilitate the design of higher quality varnish compositions and treatment processes.

Keywords Wood • Varnish • X-ray microtomography • Neutron based image analysis • Modal analysis • Numerical optimization

9.1 Introduction

Past investigations on the influence of varnish on the vibro-mechanical properties of wood, studied the changes in damping [4] as well as modifications in mass and stiffness [5]. Both came to the conclusion that the influence of varnish is relatively small compared to e.g. the wood thickness. Nevertheless, Schelleng (1968) mentioned that many luthiers experienced a better sound for their violins in the 'white', i.e. unvarnished state, than for the varnished instruments. Therefore, the influence of varnish on a previously perfectly shaped violin is large enough, to change these tuned properties; most commonly in an acoustically undesirable way [6]. These intrinsic material properties are normally measured either by a free-free flexural vibration method [3, 7] or with a tapping sound vibration apparatus [1]. The responses of small wood stripes to certain

S.L. Lämmlein (✉) • M. Sedighi Gilani
Applied Wood Material Laboratory, (EMPA) Swiss Federal Laboratories for Materials Science and Technology, Ueberlandstrasse 129, 8600 Dübendorf, Switzerland
e-mail: Sarah.Laemmlein@empa.ch

D. Mannes
Neutron Imaging and Activation Group, (PSI) Paul Scherrer Institute, 5232 Villigen, Switzerland

F.W.M. Schwarze
Applied Wood Material Laboratory, (EMPA) Swiss Federal Laboratories for Materials Science and Technology, Lerchenfeldstrasse 5, 9014 St. Gallen, Switzerland

I. Burgert
Wood Materials Science Institute for Building Materials, ETH Zürich, Stefano-Franscini-Platz 3, 8093 Zürich, Switzerland

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_9

vibration excitements are measured before and after varnishing. The measured properties correspond to averaged material properties. This implies that the varnish is assumed to be evenly distributed or spread over the wood cross-section and consequently is not considered as a separate material layer on top of the wood surface.

For changes in relative humidity respectively moisture content of the wood, the varnish layer acts as a barrier. It is presumed that the water adsorption is retarded [8], i.e. only takes place at the inner, unvarnished surface of the instrument [9].

This paper proposes an approach for further investigations of the varnish's influence on the physical and vibro-mechanical properties of wood. The focus lies on the interactions between the structure at the interface of wood and varnish, a more general mechanical characterization of common varnish materials, studies on the sorption behavior of varnished wood, the application of a different method for the material parameter determination and ultimately, the implementation of a numerical FEM model for the simulation of the vibro-mechanical properties of varnished wood.

9.2 Approach

The interactions between wood and varnish depend on different factors. On the one hand, these are the material properties of wood and applied varnishes. On the other hand, wood surface finishing methods and the varnishing procedure affect the final structure of the interface. Typically, the complete varnish application can be divided into three main steps:

- Preparing/finishing the wood surface for an optimal varnish application
- Grounding/filling/sealing the wood surface to avoid a deep penetration of the following varnish layers
- Applying the "main" varnish layers

In this context, the varnish procedure refers to all three steps. Nevertheless, it should be pointed out that each step influences the properties in a different way. The first step is not incorporating any new material, it "just" changes the geometry of the wood surface [10]. In the second step, a certain varnish or a particular ground is integrated into the upper wood cells. The aim is to close or fill the open wood cells to avoid deep penetration of the following varnish layers. In the third step, varnish is applied in several layers on the pretreated wood surface. These layers are comparable to thin films that finally lead to a closed coating. In contrast to the first two steps, the varnish layers are not affecting the wood properties directly. X-ray microtomography analysis can reveal the penetration of varnish into the wood cells. The method can be applied for different materials and varnishing procedures in order to gain some general insight and to evaluate the influence of each separate step. As it is indicated in Fig. 9.1, it is also possible to identify small particles within the varnish and wood structure.

Changes in the sorption behavior can be investigated with neutron imaging. As neutrons are very sensitive to hydrogen and hence water, the moisture content, respectively its gradient within the wood, can be determined [11]. Furthermore, the dynamics of varnish penetration into wood can be visualized. Fig. 9.2 shows the different penetration characteristics for an untreated Norway spruce wood sample (Fig. 9.2b1–c1) compared to a sample that has been pretreated with pumice powder (Fig. 9.2b2–c2). By comparing the varnish distribution after five minutes (Fig. 9.2c1–c2), the influence of the applied pretreatment becomes visible as the varnish penetrates much weaker into the wood and spreads across a larger surface area.

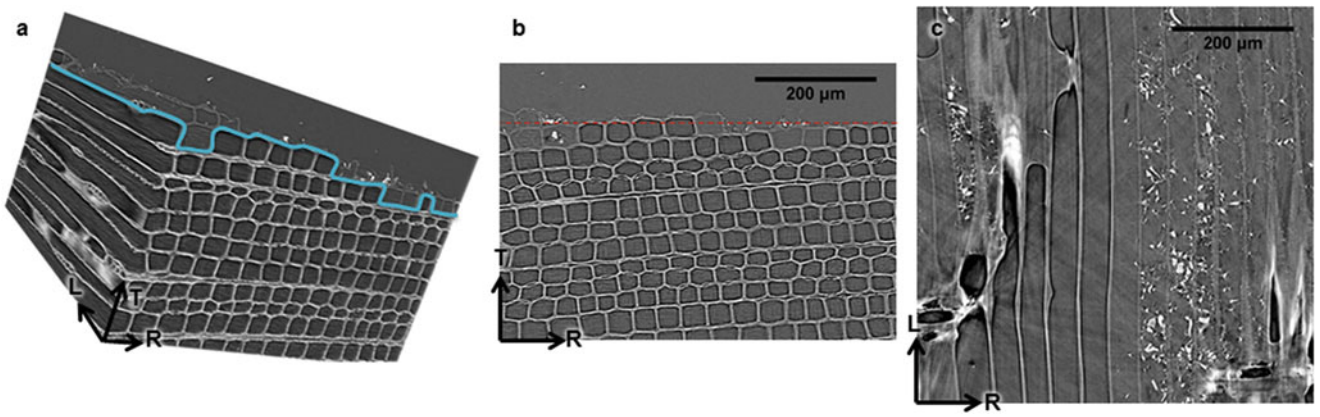


Fig. 9.1 X-ray tomography images showing the spatial penetration of varnish at the interface in (**a**) a 3D reconstruction of the structure (the blue line indicates the depth of varnish penetration), (**b**) a 2D cross-section of the TR-plane (the red dashed line represents the slice plane for **c**) and (**c**) a 2D cross-section of the LR-plane (SLS TOMCAT beamline)

Fig. 9.2 Neutron based images showing (**a**) the initial varnish droplet on the wood surface, (**b**) the TR-wood-cross-section before the application of a varnish droplet on a blank (**b1**) and a pretreated surface (**b2**), the difference in moisture content five minutes after the application of a varnish droplet on a blank (**c1**) and a pretreated surface (**c2**)

For the determination of wood material properties, a combined experimental and numerical method is proposed. The Eigen-frequencies and the corresponding internal damping values will be measured in a large frequency range by conducting a modal analysis with a laser vibrometer. The stiffness in longitudinal direction E_L and radial direction E_R are calculated from a numerical optimization by comparing the numerical Eigen-frequencies from a FEM model to the experimentally measured ones. The numerical values of E_L and E_R are adapted until the frequency difference is minimized. The same method can be applied to all material parameters, having an influence on the numerical results. A finite difference sensitivity analysis can reveal these improvable mechanical material properties. Compared to common measurements on 1D bar stripes, this technique increases the amount of measurable properties. All parameters can be determined in one measurement and on the same sample. Furthermore, as the measurement is conducted on a wide range of Eigen-frequencies, the technique can be used to evaluate frequency dependencies.

9.3 Conclusions

The described experimental and numerical methods enable further research on the influence of varnish on the properties of wood for musical string instruments. Particular attention is given to the consequences of different materials and different varnishing procedures on the wood/varnish interface, the sorption behavior and the vibro-mechanical properties. Further, factors such as relative humidity respectively moisture content and frequency dependencies can be taken into account.

References

1. Ono, T.: Effects of varnishing on acoustical characteristics of wood used for musical instrument soundboards. J. Acoust. Soc. Jpn. (E). **14**(6), 397–407 (1993)
2. Echard, J.P., et al.: The nature of the extraordinary finish of Stradivari's instruments. Angew. Chem. Int. Ed. **49**(1), 197–201 (2010)
3. Sedighi Gilani, M., et al.: Relationship of vibro-mechanical properties and microstructure of wood and varnish interface in string instruments. Appl. Phys. A. **122**(4), 1–11 (2016)
4. Meinel, H.: Regarding the sound quality of violins and a scientific basis for violin construction. J. Acoust. Soc. Am. **29**(7), 817–822 (1957)
5. Schelleng, J.C.: Acoustical effects of violin varnish. J. Acoust. Soc. Am. **44**(5), 1175–1183 (1968)
6. Haines, D.: On musical instrument wood Part II. Surface finishes, plywood, light and water exposure. CAS Newletter. **33**, 19–23 (1980)
7. Hearmon, R.F.S.: The influence of shear and rotatory inertia on the free flexural vibration of wooden beams. Br. J. Appl. Phys. **9**(10), 381 (1958)
8. Martínez, M.P., Poletti, P., Espert, L.G.: Vibration testing for the evaluation of the effects of moisture content on the in-plane elastic constants of wood used in musical instruments. Vib. Struct. Acoust. Anal. 21–57 (2011) http://link.springer.com/chapter/10.1007/978-94-007-1703-9_2
9. Thompson, R.: The effect of variations in relative humidity on the frequency response of free violin plates. J. Catgut Acoust. Soc. **32**, 25–27 (1979)
10. Barlow, C.Y., Woodhouse, J.: Microscopy of wood finishes. J. Catgut Acoust. Soc., Ser. II. **1**, 9–15 (1988)
11. Sedighi Gilani, M., Schwarze, F.W.M.R.: Hygric properties of Norway spruce and sycamore after incubation with two white rot fungi. Holzforschung. **69**(1), 77–86 (2015)

Chapter 10
Towards Robust Sustainable System Design: An Engineering Inspired Approach

Mario Holl and Peter F. Pelz

Abstract An engineering inspired method called multi-pole system analysis (MPSA) is presented and applied to an innovative wind-energy converter. The method offers a consecutive and structured guideline to determine optimal system designs in the tense interrelations of sustainability requirements, e.g. energetic efficiency, economic profitability and environmental quality. The method consists of the four steps of (1) system synthesis, (2) system analysis under uncertainty, (3) stochastic system optimization and (4) sensitivity analysis and addresses the involved uncertainty due to lack of information in the early stage of system design. As the results indicate, only a simultaneous consideration of the involved domains can truly lead to an optimal system design. By incorporating uncertainty aspects within the second step of the method and performing stochastic optimization, the disadvantage of missing robustness of previous deterministic optimal systems is overcome.

Keywords Multi-pole system analysis • Decision-making under uncertainty • Sensitivity analysis • Techno-economic analysis • Energy ship

Nomenclature

The nomenclature is shown in the dimension of length (L), mass (M), time (T) and currency (C).

Abbreviations

APF	Annuity-present value factor
CDF	Cumulative distribution function
CRF	Capital recovery factor
DBSA	Density-based sensitivity analysis
KS	Kolmogorov-Smirnov statistics
LCOH	Levelized costs of hydrogen
LPM	Lower-partial-moment
MPSA	Multi-pole system analysis
NPV	Net present value
PDF	Probability density function
RMSE	Root mean squared error
VBSA	Variance-based sensitivity analysis

M. Holl (✉) • P.F. Pelz
Chair of Fluid Systems, Technische Universität Darmstadt, Otto-Berndt-Straße 2, 64287 Darmstadt, Germany
e-mail: mario.holl@fst.tu-darmstadt.de

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_10

Accent

$(\dot{})$	Periodic variable
$(\hat{})$	Estimator
$(\tilde{})$	Random variable
$(\rightarrow)$	Vector defined in the x-, y-, z-space

Roman Symbols

$\mathbf{A}$	Matrix
A	Area in L^2
a	Dimensionless area
B	Operation time in T
b	Span in L
C	Costs in C
C_P	Coefficient of performance
c	Wind speed in L/T
c_D	Drag coefficient
c_L	Lift coefficient
d	Chord length in L
E	Internal Energy in $\mathrm{ML}^2\mathrm{T}^{-2}$
F	Distribution function
f	Specific price
G	Profit in C
H_{0,H_2}	Caloric value in $\mathrm{L}^2\mathrm{T}^{-2}$
I	Investment costs in C
K	Kinetic energy in $\mathrm{ML}^2\mathrm{T}^{-2}$
L	Vessel length in L
M	Mass conversion rate
$\dot{m}$	Mass flow rate in MT^{-1}
n	Operational lifetime in T
P	Power in $\mathrm{ML}^2\mathrm{T}^{-3}$
p	Pressure in $\mathrm{ML}^{-1}\mathrm{T}^{-2}$
$\dot{Q}$	Heat flow rate in $\mathrm{ML}^2\mathrm{T}^{-3}$
R	Revenue in C
T	Dimensionless sensitivity measure
t	t-statistics
U	Ocean current in L/T
u	Dimensionless ocean current
V	Vessel speed in L/T
v	Dimensionless vessel speed
W	Force in MLT^{-2}
w	Mass specific work in $\mathrm{L}^2\mathrm{T}^{-2}$
$\mathbf{X}$	Random input variable vector
X	Random input variable
$\mathbf{x}$	Deterministic input variable vector
$\underline{x}$	Deterministic input variable
$\overline{x}$	Multi-pole input vector
Y	Random output variable
$\underline{y}$	Deterministic output variable
$\overline{y}$	Multi-pole output vector
Z	Cash flow in C
z	Rate of interest

Greek Symbols

δ	Capacity factor
ε	Mass conversion rate
ζ	Axial induction factor
η	Efficiency factor
Λ	Aspect ratio
λ	Share of O&M costs
μ	Expected value of normal distribution
Ξ	Utility function
$\boldsymbol{\xi}$	Vector of random variables
Π	System efficiency
ϱ	Density in ML^{-3}
σ	Standard deviation of normal distribution
τ	Reference value of the LPM
ϕ	Degree of freedom t-distribution
$\boldsymbol{\Omega}$	System matrix
ω	Weighting factor of risk-aversion

Greek Symbols

Comp	Compressor
Des	Desalinator
Elec	Electrolyseur
g	Gaseous
l	Liquid
T	Turbine
V	Vessel

10.1 Introduction

Both, the growing sustainability requirements for technical products and the growing environmental awareness of the population lead inevitably to the incorporation of sustainability aspects in the phase of technical product design and development. The sustainability requirement is interdisciplinary and the involved disciplines are in most cases contradictionary, so that the question of the optimal technical system as a trade-off of the involved criteria arises.

In this paper a methodological approach is presented that leads in a structured manner to this optimal system. The method has been recently developed by the authors and is called Multi-pole system analysis (MPSA). By the adaption of well-known practices from engineering sciences and the consecutive projection into the involved scientific fields, one yields the method consisting of the four steps of (1) system synthesis, (2) system analysis under uncertainty, (3) stochastic system optimization and (4) sensitivity analysis. In the first step of the method the system as well as the system topology is defined. In this framework a system boundary is defined including the involved aspects, e.g. energetic efficiency, economic profitability and environmental quality, which are presented through adequate fluxes through the system. The system topology is defined by the number of components and their interconnections. By using multi-pole formalism a superior model of the system is gained showing all linkages of the considered criteria. Coupling effects of the cross-domain fluxes are already revealed. In the second step the system is analyzed in detail using the practices of the respective field. For instance, a generic physical model is derived in the technical domain which is in the next step mathematically modeled using first principles. In economics, methods of the investment analysis are used to describe the system economically. Due to the modelling process and the general lack of system information the description has to be performed taking uncertain aspects into account. In this paper, the most common approach of probabilistic uncertainty analysis is used by means of modeling uncertain input factors by adequate distribution functions. Following this train of thought, the system optimization of step (3) turns over into a stochastic

optimization problem. The optimization problem is solved using Monte-Carlo-Simulations. The model quality, the cause and effect relationship of the model as well as dominant input factors are determined by performing a sensitivity analysis as the last step of the method. The last step basically provides the information of how the uncertainty in the model output can be assigned to the respective uncertainty of the model input.

The method is presented in this paper on the example of an innovative wind-energy converter called the energy ship concept. The concept involves a wind-powered vessel with a hydrokinetic turbine. The vessel converts the kinetic energy of the wind into kinetic energy of the vessel. The wind is substituted with water for power generation, so that the resulting hydrokinetic turbine can be designed very small in comparison to a regular wind turbine. The electric energy is stored on board of the vessel by means of the electrolytic splitting of sea water into hydrogen and oxygen. The concept was first proposed by Salomon [1] in 1982, followed by Meller [2], Holder [3] and Gizara [4]. More quantitative analysis were performed by Platzer [5–9] and Kim [10–12]. Most recently the first and second author and Platzer [13–15] presented a physically based upper limit for the conversion of wind energy in mechanical energy of the presented concept. Based on this approach, the first and second author introduced a general method for holistic system analysis, the MPSA method [16].

The recent publications have the drawback of treating the energy converter deterministically. This is an assumption which truly does not represent reality sufficient and, thus, in this paper the next evolutionary step of the method is presented by incorporating system analysis under uncertainty to gain a certain degree of robustness in the phase of optimal system design. Consequently, the optimization problem turns over into a stochastic optimization problem. Following this train of thought and the given introduction, the paper is structured in the Sects. 10.2 and 10.3. In Sect. 10.2 the four steps of the method are gradually explained in general and then applied to the energy converter. Highlights and differences in contrast to the previous analysis are outlined as well as the advantages which go hand in hand with the incorporation of the uncertainty consideration. Section 10.3 closes the paper by summarizing the content and emphasizing the gained knowledge.

10.2 Multi-Pole System Analysis

10.2.1 System Synthesis

The first step of the MPSA focuses on the question of the shape and topology of the considered system. The definition of a system, according to Buchholz [17] requires a system boundary to delimit the "inner world" to the "outer world" and thus, defines the considered framework. The topology of the system is defined by the number of components within the system boundary and the interconnection through signals. We define these signals as various cross-domain fluxes, e.g. mass flows, energy flows and cash flows that are used for communication in between the components and across the system boundary. Focusing on a single component, the input fluxes are listed in the input vector $\underline{x}$ and the output fluxes in the output vector $\underline{y}$, respectively. The component itself is mathematically described by the matrix $\mathbf{A}$ and represents the effect of the component to each flux and also linkages of fluxes. Thus, the composite of input- and output fluxes and the component can be described as $\underline{x} = \mathbf{A}\underline{y}$. If all system components are series-connected, one can easily combine the component descriptions and yields the system description $\underline{x} = \prod_{i=1}^{N} \mathbf{A}_i \underline{y}$, with the system matrix $\mathbf{\Omega} = \prod_{i=1}^{N} \mathbf{A}_i$. If all components are parallel connected one yields for the system matrix $\mathbf{\Omega} = \sum_{i=1}^{N} \mathbf{A}_i$.

As can be seen in Fig. 10.1 the components of the energy converter are, relating to considered energy-, mass- and cash flows, series-connected. Thus, one can achieve the system matrix by multiplying all component matrices. An obvious interaction of the considered fluxes can be seen on the components desalinator, electrolyser and compressor, which require a certain amount of power for functional performance. Also, the periodic cash flow needs at least to be balanced through the periodic revenue, which is gained through the sale of the hydrogen.

The matrix notation of the multi-pole model is shown in detail in a recent publication [16] and is here recalled as

$$\begin{pmatrix} P_{\text{avail}} \\ \dot{m}_{\text{H}_2\text{O,s}} \\ \dot{C} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{C_{\text{P}}\eta_{\text{Gen}}}\left(\frac{H_{\text{o,H}_2}}{\eta_{\text{Elec}}} + \frac{w_{\text{Des}}}{\varepsilon_{\text{Elec}}} + w_{\text{Comp}}\right) & 0 \\ 0 & \frac{1}{\varepsilon_{\text{Des}}\varepsilon_{\text{Elec}}} & 0 \\ 0 & \delta T f_{\text{H}_2} & -1 \end{pmatrix} \begin{pmatrix} 0 \\ \dot{m}_{\text{H}_2} \\ \dot{G} \end{pmatrix}, \tag{10.1}$$

with the mass flows denoted as $\dot{m}$, efficiency factors denoted as η, mass conversion rates as ε and mass specific work as w. If one calculates in each specific domain the ratio of output and input quantity, one yields the system efficiency factor Π, the mass conversion rate M and the well-known economic indicator return on investment ROI, defined as the ratio of the periodic profit $\dot{G}$ and the periodic costs $\dot{C}$

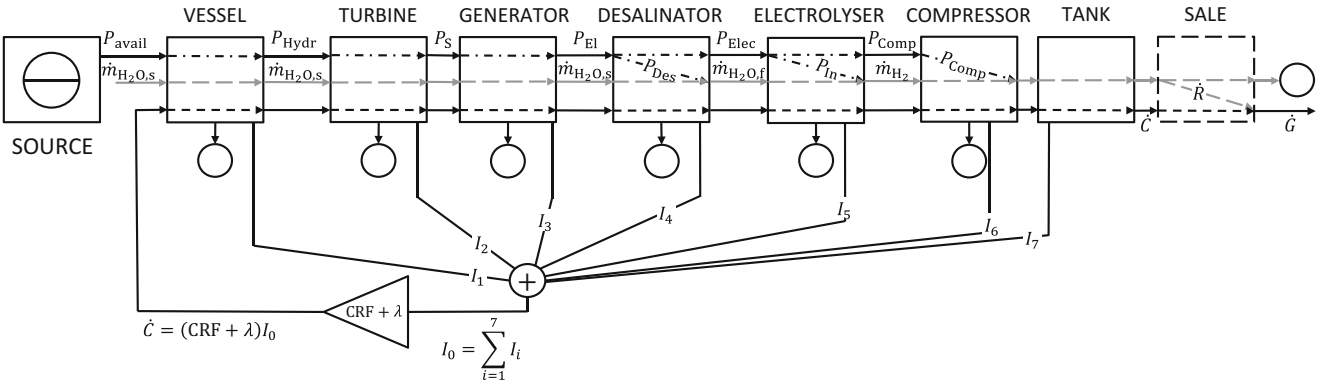


Fig. 10.1 Multi-pole model of the energy converter with energy-, mass- and cost flows [16]

$$\begin{aligned} \Pi &:= \frac{\dot{m}_{H_2} H_{o,H_2}}{P_{avail}} = \frac{C_P \eta_{Gen} H_{o,H_2}}{\frac{H_{o,H_2}}{\eta_{Elec}} + \frac{w_{Des}}{\varepsilon_{Elec}} + w_{Comp}}, \\ \mathrm{M} &:= \frac{\dot{m}_{H_2}}{\dot{m}_{H_2O,s}} = \varepsilon_{Des} \varepsilon_{Elec}, \\ \mathrm{ROI} &:= \frac{\dot{G}}{\dot{C}} = \frac{\delta T \dot{m}_{H_2} f_{H_2}}{\dot{C}} - 1. \end{aligned} \tag{10.2}$$

It can be seen, that the economic profitability is influenced by the energetic efficiency of the energy converter. Nevertheless, the gained multi-pole model represents only a superior model which will be further investigated in the detailed system analysis, the second step of the MPSA method.

10.2.2 Analysis Under Uncertainty

The second step of the MPSA deals with the detailed analysis of the considered system. The first step only provides a superior model, which highlights interactions of the fluxes. In engineering terms, a domain-specific white-box model of the system is derived. In the technical domain this is typically done by deriving a physical model of the real system. Since the real system is too complex to describe this model goes hand in hand with assumptions, simplifications and neglect of physical irrelevant phenomena. In the next step first principles, i.e. the conservation of mass, energy and momentum, etc., are applied to the physical model to determine the physics of the system. By further considering energy conversion systems as ideal one finds a physically based upper limit for power generation as presented by Carnot for thermodynamic power cycles [18], Betz for wind turbines [19], the second author for low-head hydro power [20] and the first and second author and Platzer for the here presented concept [15]. In contrast, one has to use economic models for an economic system evaluation, e.g. the method of net present value NPV, the annuity method or static methods like the amortization method. As one could already see in Fig. 10.1, all flows converge in the economic model and thus, one finds a combined holistic model of the system with the deterministic energetic and economic input factors $\mathbf{x}^T = [x_1, \ldots, x_M]$ and the deterministic model output y, $y = f(\mathbf{x})$. In a world with no uncertainty a specific system, characterized by $\mathbf{x}_0$, would lead to the deterministic model output y_0. In reality, systems cannot be characterized exactly since system information are never perfect, the system has been simplified through the modeling process and knowledge gaps exist. Especially in economics one has to deal with random price and cost alterations which can be faced with an uncertainty assignment for a more reliable system analysis. Following this train of thought, the deterministic model turns over into a stochastic model $Y = f(\mathbf{X})$, with the random input variable vector $\mathbf{X}$ and the random output variable Y. Due to the above mentioned reasons, the authors emphasize here that models generally should be considered uncertain. Various ways for uncertainty representation can be found in the literature [21–26]. Knight [27] was the first to differ between risk and uncertainty in 1921. The probabilistic uncertainty analysis is the most common approach. Each input factor is mathematically approximated by an adequate probability distribution function PDF. In general, an analytic calculation of the model output distribution is not possible and thus, the calculation is performed numerically

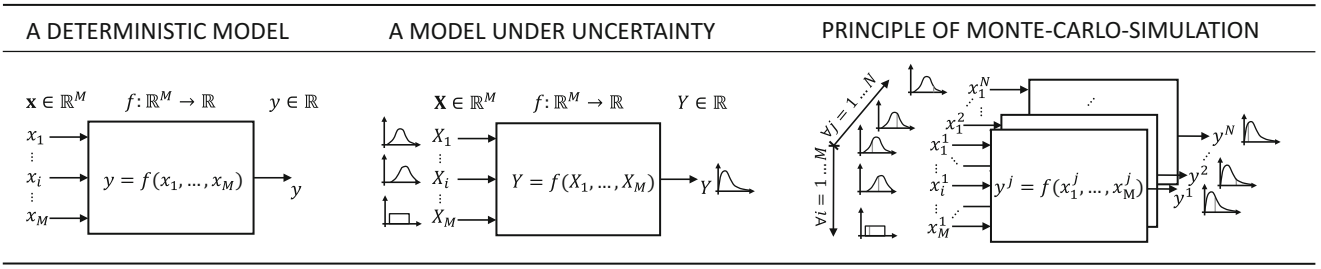


Fig. 10.2 Evolution of the considered models and calculation through Monte-Carlo-Simulation

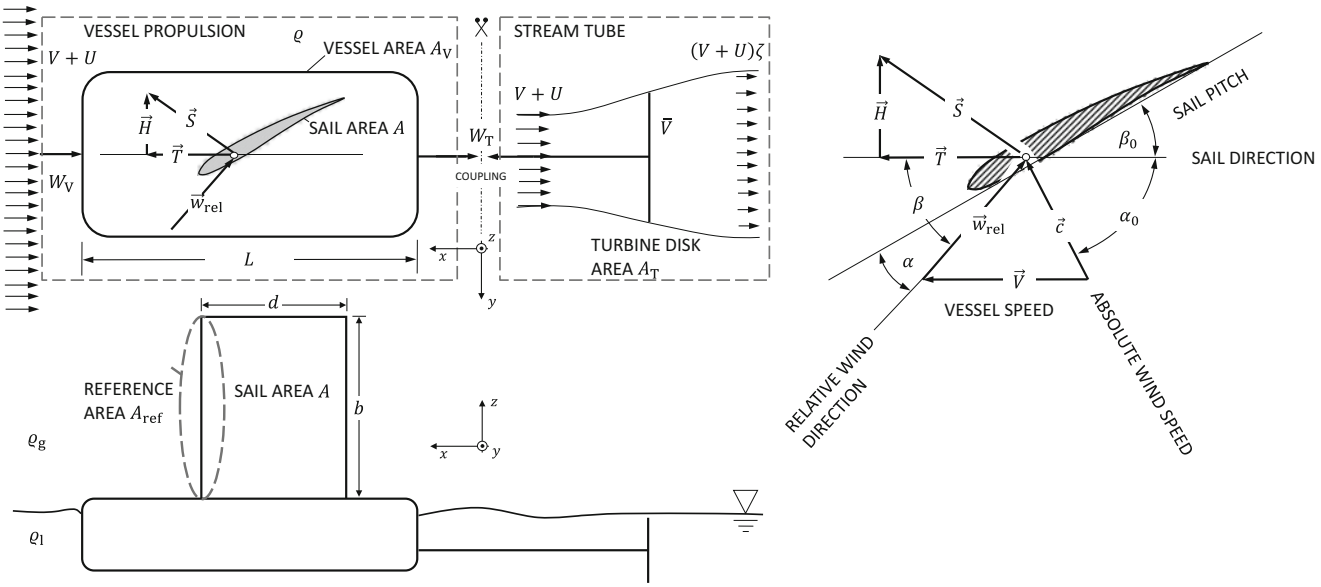


Fig. 10.3 Physical model of the energy converter

by discretizing each input factor distribution n times and calculating n model outputs. This principle is called Monte-Carlo-Simulation. Figure 10.2 shows the differences of a deterministic model, the mathematic model under uncertainty and the principle of calculating the latter model using Monte-Carlo-Simulation.

The energy conversion concept is in this paper modeled energetically and economically. Thus, the analysis is structured into these two parts. In the following system analysis random variables are mathematically denoted as (~).

10.2.2.1 Energetic System Analysis

According to the above mentioned process, the physical model of the energy converter can be seen in Fig. 10.3. The sail area is denoted as A, the wetted sail area as A_V and the turbine area as A_T. The vessel speed is denoted as V and the ocean current as U. The turbine is modelled as an ideal disc actuator.

To determine the physics of energy conversion, the first law of thermodynamics is applied. The first law of thermodynamics for a stream tube was scientifically first formulated by von Helmholtz [28] and reads

$$\frac{\mathrm{D}K}{\mathrm{D}t} + \frac{\mathrm{D}E}{\mathrm{D}t} = P + \dot{Q}, \tag{10.3}$$

with the kinetic energy K, the internal energy E, the power of body and surface forces P and the heat flow $\dot{Q}$. By applying the first law to the physical model one yields for the mechanical turbine power $\widetilde{P}_S = \widetilde{W}_T \left(\widetilde{V + U}\right) \widetilde{\eta}_T$, with the turbine resistance force $\widetilde{W}_T$, the turbine efficiency factor $\widetilde{\eta}_T$ and the average of far upstream and downstream velocity

$\overline{\widetilde{V}+\widetilde{U}} = \left(\widetilde{V}+\widetilde{U}\right)(1+\zeta)/2$, with ζ as the ratio of the far downstream and upstream velocity. The turbine resistance force results from the momentum equation, which reads

$$\frac{\mathrm{D}\overrightarrow{I}}{\mathrm{D}t} = \overrightarrow{F}, \tag{10.4}$$

with the impulse $\overrightarrow{I}$ and the forces $\overrightarrow{F}$. Thus, the turbine resistance force can be calculated as $\widetilde{W}_\mathrm{T} = \Delta\widetilde{p}A_\mathrm{T}$, with the pressure difference calculated by Bernoulli's equation $\Delta\widetilde{p} = \varrho_\mathrm{l}\left(\widetilde{V}+\widetilde{U}\right)^2\left(1-\zeta^2\right)/2$. It can be seen that the uncertainty of the vessel speed and ocean current leads to uncertainty in the turbine pressure drop, resistance force and the turbine power output. The total available power is defined as the power of the influenced air mass by the wing. Prandtl [29] showed that this air mass is equal the span b multiplied by twice the induced downwash velocity. Thus, the sail influences the air mass defined by the circular area A_ref and the available power is $\widetilde{P}_\mathrm{avail} := \varrho_\mathrm{g}\widetilde{c}^3 A_\mathrm{ref}/2$. In order to determine an upper limit for energy conversion of the proposed concept, the coefficient of performance $\widetilde{C}_\mathrm{P}$ is defined as the ratio of the mechanical turbine power and the available power

$$\widetilde{C}_\mathrm{P} := \frac{\widetilde{P}_\mathrm{S}}{\widetilde{P}_\mathrm{avail}} = \frac{4}{\varrho\pi\Lambda}\left(\widetilde{v}+\widetilde{u}\right)^3\widetilde{a}_\mathrm{T}\widetilde{\eta}_\mathrm{T}\frac{1+\zeta}{2}\left(1-\zeta^2\right), \tag{10.5}$$

with the dimensionless density ratio $\varrho := \varrho_\mathrm{g}/\varrho_\mathrm{l}$, the aspect ratio $\Lambda := b/d$, the dimensionless vessel speed $\widetilde{v} := \widetilde{V}/\widetilde{c}$, the dimensionless ocean current $\widetilde{u} := \widetilde{U}/\widetilde{c}$ and the dimensionless turbine area $\widetilde{a}_\mathrm{T} := A_\mathrm{T}/\widetilde{A}(L)$. The sail area as well as the wetted area of the vessel scale with the length of the vessel L. In a previous publication [16] the results of a market survey and the corresponding scaling laws were presented. To generally assess the assigned uncertainty of the estimated scaling law $\widehat{y}$, we calculate the corresponding 95%-prediction interval according to Sachs [30] and gain an upper and lower bound for all predictions

$$\widehat{y}_0 \pm \epsilon_i = \widehat{y}_0 \pm t_{1-\alpha/2,\phi-2}\cdot\mathrm{RMSE}\cdot\sqrt{1+\frac{1}{\phi}+\frac{(x_0-\underline{x})^2}{\sum_{i=1}^{\phi}(x_i-\underline{x})^2}}, \tag{10.6}$$

with the root mean square $\mathrm{RMSE} = \sqrt{\sum_{i=1}^{\phi}\frac{(y_i-\widehat{y}_i)^2}{\phi-2}}$, the average of the input factor $\underline{x}$ and the t-statistics, which is a function of the number of samples ϕ. The prediction interval is small if the RMSE is small, the input factor x_0 is close to the average $\underline{x}$ and the number of samples is high. Figure 10.4 shows the scaling laws of sail and wetted vessel area with their respective 95%-prediction interval.

Using these scaling laws a multitude of possible systems can be described with respect to their energetic efficiency. An energetic white-box model of the system is gained.

10.2.2.2 Economic System Analysis

To evaluate the economic quality the method of net present value is used in this paper. The method discounts all cash flows to a specific point of time to consider the time value of money. If the NPV is positive, one will pursue an investment and if the NPV is negative one will not. Investors in general are considered indifferent with regard to an investment, if the NPV = 0. According to Brockhoff [31] and Schneider [32] the first discounting formula has been used by Stevin [33] in 1582 but the first juristic and economic justification of the net present value method has been presented by the famous mathematician Leibniz [34] in 1682. The method of net present value reads

$$\widetilde{\mathrm{NPV}} = \sum_{t=0}^{n}\frac{\widetilde{\dot{Z}}_t}{\left(1+\widetilde{z}\right)^t} = -\widetilde{\dot{I}}_0 + \sum_{t=1}^{n}\frac{\widetilde{\dot{R}}_t - \widetilde{\dot{C}}_t}{\left(1+\widetilde{z}\right)^t}, \tag{10.7}$$

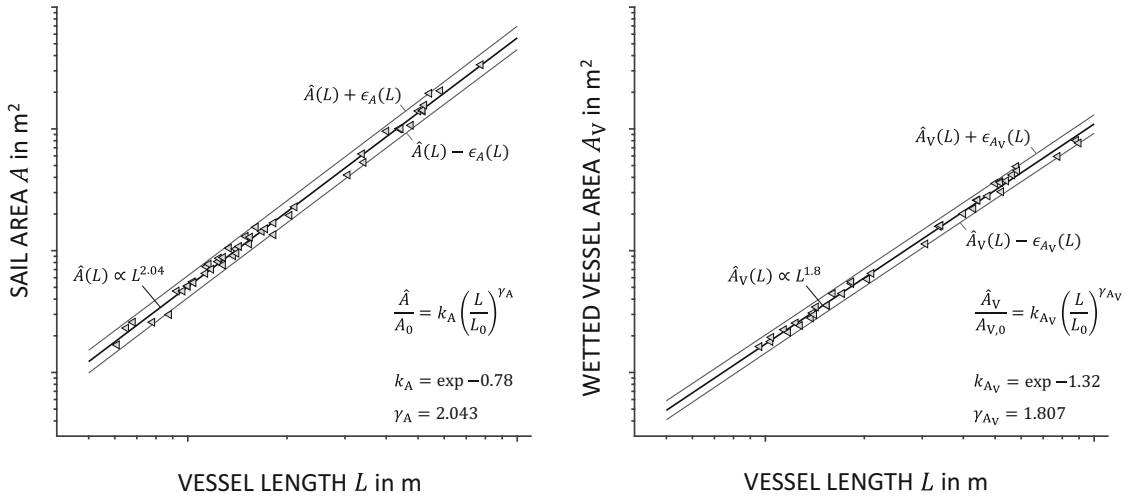


Fig. 10.4 Geometric scaling laws for the sail area A and the wetted vessel area A_v in logarithmic scale

with the periodic cash flows $\widetilde{\dot{Z}}_t$, which can further be divided into the periodic revenue $\widetilde{\dot{R}}_t$ and the periodic costs $\widetilde{\dot{C}}_t$, the operational time n and the rate of interest $\widetilde{z}$. If one further assumes the periodic revenues and costs as time-independent, one can calculate the NPV with the annuity present-value factor $\widetilde{\text{APF}}$, which is the reciprocal of the capital recovery factor $\widetilde{\text{CRF}} = \widetilde{z}\left(1+\widetilde{z}\right)^n / \left(\left(1+\widetilde{z}\right)^n - 1\right)$. While the initial costs $\widetilde{I}_0$ cover all the investment costs, the periodic costs cover the operation and maintenance (O&M) costs, which can be calculated as the percentage $\widetilde{\lambda}$ of the investment costs $\widetilde{I}_0$, $\widetilde{\dot{C}} = \widetilde{\lambda}\widetilde{I}_0$

$$\widetilde{\text{NPV}} = -\widetilde{I}_0 + \frac{1}{\widetilde{\text{CRF}}}\left(\widetilde{\dot{R}} - \widetilde{\lambda}\widetilde{I}_0\right). \tag{10.8}$$

With these assumptions the NPV method turns over into the annuity method so that one is able to calculate the periodic profit $\widetilde{\dot{G}}$ as the difference of the periodic revenue and the periodic costs

$$\underbrace{\widetilde{\text{CRF}} \cdot \widetilde{\text{NPV}}}_{\widetilde{\dot{G}}} = \widetilde{\dot{R}} - \underbrace{\left(\widetilde{\text{CRF}} + \widetilde{\lambda}\right)\widetilde{I}_0}_{\widetilde{\dot{C}}}. \tag{10.9}$$

The periodic revenue is generated through the sale of the periodically converted hydrogen $\widetilde{\dot{R}} = \widetilde{\delta} T \widetilde{\dot{m}}_{\text{H}_2} \widetilde{f}_{\text{H}_2}$, with the capacity factor $\widetilde{\delta} := \widetilde{B}/T$ as the ratio of the periodic system operation time $\widetilde{B}$ and the periodic total time T. The periodic revenue depends on the market price for hydrogen $\widetilde{f}_{\text{H}_2}$. To assess the techno-economic quality of an energy conversion technology, one calculates not the periodic revenue nor the net present value but the levelized production costs, which is the critical price of hydrogen for cost neutrality. Low levelized costs indicate a high techno-economic system quality. It is here referred to as "techno-economic" because the levelized costs of hydrogen LCOH considers on the one hand the investment and O&M-costs of the technology and on the other hand the periodically converted hydrogen, which depends on the energetic efficiency. Consequently, a hydrogen providing technology has small LCOH for a low investment and O&M costs, high system utilization δ and high system efficiency. Mathematically, the LCOH result of the condition $\widetilde{\text{NPV}} \overset{!}{=} 0$, so that Eq. (2.9) reads $0 = \widetilde{\delta} T \widetilde{\dot{m}}_{\text{H}_2} \widetilde{\text{LCOH}} - \widetilde{\dot{C}}$. Hence, the levelized costs can be calculated as

$$\widetilde{\text{LCOH}} = \frac{\left(\widetilde{\text{CRF}} + \widetilde{\lambda}\right)\widetilde{I}_0}{\widetilde{\delta} T \widetilde{\dot{m}}_{H_2}}, \tag{10.10}$$

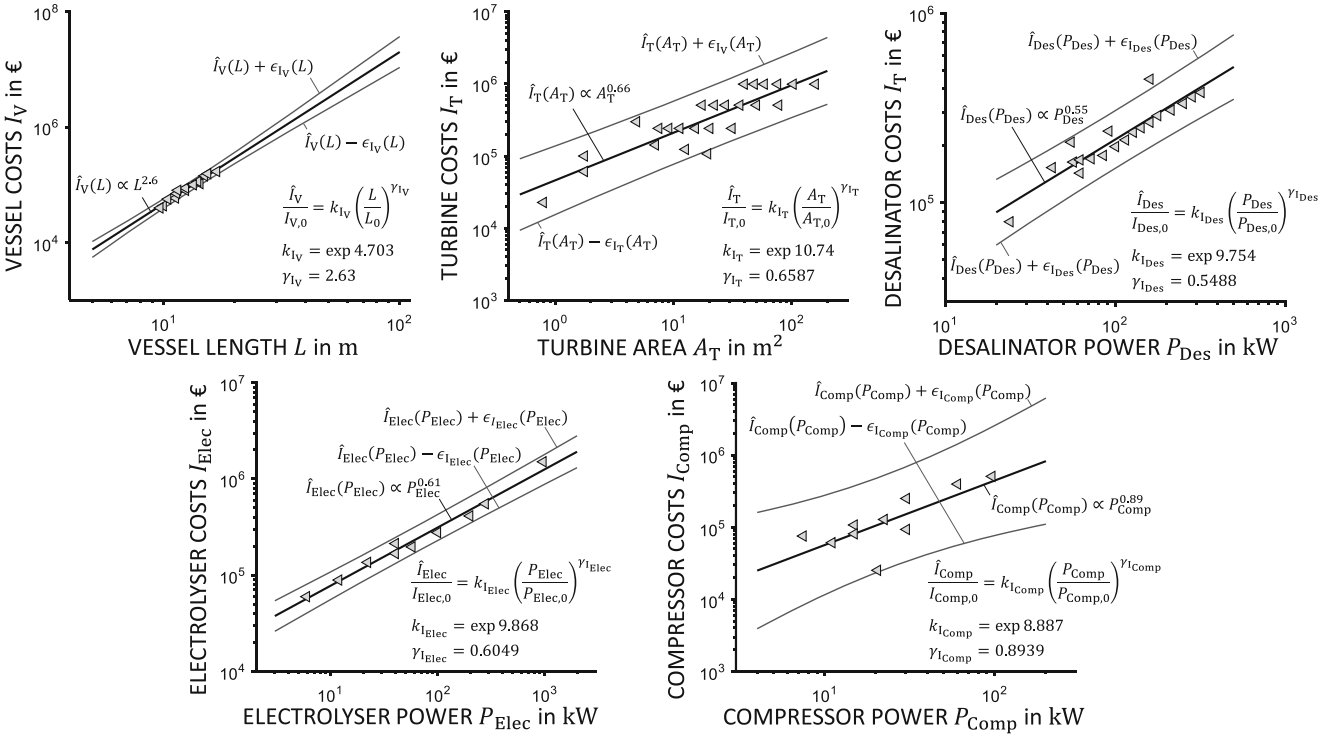


Fig. 10.5 Investment scaling functions for vessel, turbine, desalinator, electrolyser and compressor with respective 95%-prediction interval in logarithmic scale

and used here as the objective of the techno-economic optimization. By comparing the LCOH with the market price of hydrogen f_{H_2}, one can easily assess the profitability of the technology

$$\text{LCOH} \begin{cases} > f_{H_2}, \dot{G} < 0 \\ = f_{H_2}, \dot{G} = 0 \\ < f_{H_2}, \dot{G} > 0 \end{cases}. \tag{10.11}$$

In order to describe not only one but a multitude of possible systems, another market survey has been performed to derive scaling laws for the investment costs of the involved system components. Figure 10.5 shows the scaling laws for vessel, turbine, desalinator, electrolyser and compressor with their respective 95%-prediction interval.

As can be seen in Fig. 10.5 the vessel costs scale with the vessel length, the turbine costs with the turbine area and the desalinator, electrolyser and compressor with their respective consumed power. An economic white-box model results. Looking back on the system analysis, one can see several commonalities of the performed energetic and economic analysis, i.e. the usage of first principles in engineering and basic models in economics. In both cases, empirical scaling laws are used to describe a multitude of possible systems.

10.2.2.3 Uncertainty Assignment

The system analysis is performed using random variables. As the last step of the system analysis, these random variables are now assigned to adequate PDFs in order to mathematically model the corresponding input uncertainty. In this paper the three following PDF's are used:

(1) **Uniform distribution**

A random variable $X_i \sim \mathcal{U}(a, b)$ is uniformly distributed in the interval $[a, b]$ with $a < b$ and $a, b \in \mathbb{R}$ with the probability density function PDF

$$f_{\mathcal{U}}(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & else \end{cases} . \tag{10.12}$$

A uniform distribution is used if the input quantity can be assumed to appear in a specific interval with the same probability.

(2) **Normal distribution**

A random variable $X_i \sim \mathcal{N}(\mu, \sigma)$ is normally distributed with the expected value μ and the standard deviation σ and $\mu, \sigma \in \mathbb{R}$ with the symmetric PDF

$$f_{\mathcal{N}}(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma}\right). \tag{10.13}$$

A normal distribution is used to describe random processes which scatter around the expected value.

(3) ***t*-distribution**

A random variable $X_i \sim \mathcal{T}(\phi)$ is t-distributed with the degree of freedom ϕ. The PDF is given by

$$f_{\mathcal{T}}(x) = \frac{\Gamma\left(\frac{\phi+1}{2}\right)}{\Gamma\sqrt{\phi\pi}\left(\frac{\phi}{2}\right)}\left(1 + \frac{x^2}{\phi}\right)^{-\frac{\phi+1}{2}}, \tag{10.14}$$

with the gamma function Γ. The t-distribution plays an important role for the construction of confidence and prediction intervals and converges to the normal distribution for $\phi \to \infty$. Thus, if the sample size is small and the standard deviation is unknown, a greater uncertainty is assumed in comparison to an assumed normal distribution.

Figure 10.6 shows the three considered PDFs. The right side of the figure shows several t-distributions for varying degrees of freedom ϕ. It can be seen that the t-distribution converges to the normal distribution for greater sample sizes.

Table 10.1 shows the assigned PDF to the uncertain model parameters. The characteristic values of each PDF are based on literature reviews, market surveys and expert opinions.

Following the initial outlined motivation the concept has been energetically described by deriving a physical model and applying first principles, which leads to a physically based upper limit for energy conversion of the fluid system. The method of net present value has been used to describe the system economically. The LCOH has been identified as the objective for techno-economic optimal system design. The analysis has been performed assuming uncertain input factors instead of

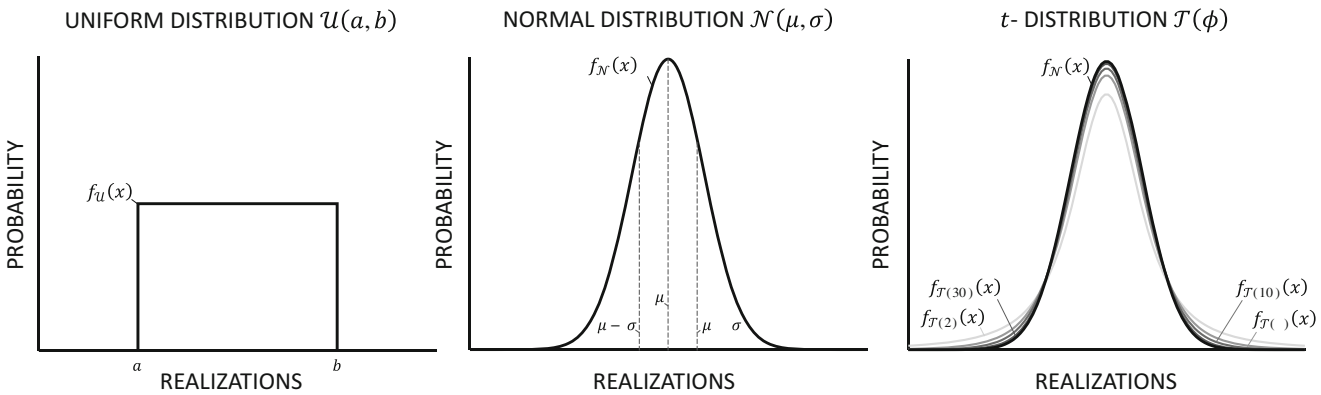


Fig. 10.6 Uniform- (*left*), normal- (*middle*) and t-distribution (*right*) for mathematical uncertainty modeling. The right PDF shows the t-distributions for the sample sizes $\phi^{\mathrm{T}} = [2, 5, 10, 30]$

Table 10.1 Uncertain input factors and assigned PDF

	Input factor	Symbol	Probability density function
Energetic input factors	Absolute wind direction	α_0	$\tilde{\alpha}_0 \sim \mathcal{U}(100,120)$
	Absolute wind speed	c	$\tilde{c} \sim \mathcal{N}(8, 1/\sqrt{20})$
	Dimensionless ocean current	u	$\tilde{u} \sim \mathcal{U}(0,0.2)$
	Sail lift coefficient	c_L	$\tilde{c}_\mathrm{L} \sim \mathcal{U}(1,1.8)$
	Vessel drag coefficient	c_D	$\tilde{c}_\mathrm{D} \sim \mathcal{U}(3.74 \cdot 10^{-3}, 7.7 \cdot 10^{-3})$
	Efficiency factor turbine	η_T	$\tilde{\eta}_\mathrm{T} \sim \mathcal{U}(0.8, 0.95)$
	Efficiency factor generator	η_Gen	$\tilde{\eta}_\mathrm{Gen} \sim \mathcal{U}(0.8, 0.95)$
	Efficiency factor electrolyser	η_Elec	$\tilde{\eta}_\mathrm{Elec} \sim \mathcal{U}(0.54, 0.74)$
	Efficiency factor compressor	η_Comp	$\tilde{\eta}_\mathrm{Comp} \sim \mathcal{U}(0.5, 0.8)$
	Inlet pressure compressor	p_1	$\tilde{p}_1 \sim \mathcal{U}(1, 80)$
	Outlet pressure compressor	p_2	$\tilde{p}_2 \sim \mathcal{U}(50, 400)$
	Volume specific work desalinator	w_Des	$\tilde{w}_\mathrm{Des} \sim \mathcal{U}(3.75, 5.95)$
Economic input factors	Lifespan	n	$\tilde{n} \sim \mathcal{U}(15, 25)$
	Rate of interest	z	$\tilde{z} \sim \mathcal{U}(0.03, 0.09)$
	Capacity factor	δ	$\tilde{\delta} \sim \mathcal{U}(0.6, 0.9)$
	O&M vessel	λ_V	$\tilde{\lambda}_\mathrm{V} \sim \mathcal{U}(0.04, 0.1)$
	O&M turbine	λ_T	$\tilde{\lambda}_\mathrm{T} \sim \mathcal{U}(0.043, 0.128)$
	O&M electrolyser	λ_Elec	$\tilde{\lambda}_\mathrm{Elec} \sim \mathcal{U}(0.01, 0.051)$
	O&M desalinator	λ_Des	$\tilde{\lambda}_\mathrm{Des} \sim \mathcal{U}(0.1, 0.2)$
	O&M compressor	λ_Comp	$\tilde{\lambda}_\mathrm{Comp} \sim \mathcal{U}(0.04,0.1)$
	O&M tank	λ_Ta	$\tilde{\lambda}_\mathrm{Ta} \sim \mathcal{U}(0.01, 0.5)$
	Specific costs hydrogen tank	f_Ta	$\tilde{f}_\mathrm{Ta} \sim \mathcal{U}(1, 4)$
Scaling functions	Aspect ratio	L_WL/L	$\tilde{L}_\mathrm{WL}/L \sim \mathcal{U}(0.7, 0.95)$
	Coefficient Taylor approximation	C_Taylor	$\tilde{C}_\mathrm{Taylor} \sim \mathcal{U}(2.5, 3)$
	Scaling function sail area	A	$\widehat{A} - \epsilon_\mathrm{A} \leq \tilde{A} \leq \widehat{A} + \epsilon_\mathrm{A}$
	Scaling function vessel displacement	D	$\widehat{D} - \epsilon_\mathrm{D} \leq \tilde{D} \leq \widehat{D} + \epsilon_\mathrm{D}$
	Scaling function vessel costs	I_V	$\widehat{I}_\mathrm{V} - \epsilon_{\mathrm{I_V}} \leq \tilde{I}_\mathrm{V} \leq \widehat{I}_\mathrm{V} + \epsilon_{\mathrm{I_V}}$
	Scaling function turbine costs	I_T	$\widehat{I}_\mathrm{T} - \epsilon_{\mathrm{I_T}} \leq \tilde{I}_\mathrm{T} \leq \widehat{I}_\mathrm{T} + \epsilon_{\mathrm{I_T}}$
	Scaling function electroyser costs	I_Elec	$\widehat{I}_\mathrm{Elec} - \epsilon_{\mathrm{I_{Elec}}} \leq \tilde{I}_\mathrm{Elec} \leq \widehat{I}_\mathrm{Elec} + \epsilon_{\mathrm{I_{Elec}}}$
	Scaling function desalinator costs	I_Des	$\widehat{I}_\mathrm{Des} - \epsilon_{\mathrm{I_{Des}}} \leq \tilde{I}_\mathrm{Des} \leq \widehat{I}_\mathrm{Des} + \epsilon_{\mathrm{I_{Des}}}$
	Scaling function compressor costs	I_Comp	$\widehat{I}_\mathrm{Comp} - \epsilon_{\mathrm{I_{Comp}}} \leq \tilde{I}_\mathrm{Comp} \leq \widehat{I}_\mathrm{Comp} + \epsilon_{\mathrm{I_{Comp}}}$

deterministic ones. Finally, empirical scaling laws have been presented in the energetic and economic system analysis as the result of market surveys and literature reviews to describe not only one, but a multitude of possible systems. Due to the imperfection of this data the respective 95%-prediction interval has been calculated to address the data uncertainty. In the last step the uncertain inputs are mathematically modelled by adequate PDFs. Thus, one yields a broad system description concerning the combined energetic and economic system quality under uncertainty. In the next step the optimal of these possible systems is identified.

10.2.3 Stochastic Optimization

Since uncertain model parameters have been considered in the second step, the optimization turns over into a stochastic optimization problem. The beginning of the mathematic field of stochastic programming goes back to Dantzig [35], who published a paper dealing with an linear optimization problem under uncertainty in 1955. Nowadays, the work of Kall [36] and Birge [37] are well known as standard literature in the field of stochastic optimization. Wets [38] gives as general form of a stochastic optimization problem

$$\begin{aligned} \min\ & \mathbb{E}\left\{ g_0\left(x, \widetilde{\boldsymbol{\xi}}\right)\right\} \\ \text{s.t. } & \mathbb{E}\left\{ g_i\left(x, \widetilde{\boldsymbol{\xi}}\right)\right\} \leq 0,\ \ \forall i = 1, \ldots, N \\ & \mathbb{E}\left\{ \mathrm{g_i}\left(\mathrm{x}, \widetilde{\boldsymbol{\xi}}\right)\right\} = 0,\ \ \forall \mathrm{i} = \mathrm{N}+1, \ldots, \mathrm{M} \\ & x \in X \subset \mathbb{R}^n, \end{aligned} \tag{10.15}$$

with the expected value $\mathbb{E}$, the objective function f_0 subjected to the constraints consisting of the equalities and inequalities g_i. The random variable vector is denoted as $\widetilde{\boldsymbol{\xi}}$. Generally, the constraints specify the considered system. Numerous methods for decision making under uncertainty can be found in several reviews [39–43], e.g. the Bernoulli-principle, the concept of stochastic dominance, downside-risk and mean-variance analysis. Nevertheless, the application specifies the chosen concept of decision making under uncertainty. A common ground of all methods is the statement of a utility function Ξ that is minimized or maximized. In this paper two different utility functions are derived and thus gain two different techno-economic optimization functions. If one recalls the objective of gaining minimal $\widetilde{\mathrm{LCOH}}$, this thought leads inevitably to the framework of expected value. This idea goes back to Bernoulli's work in 1738 for the assessment of gambling [44], which is nowadays referred to as the scientific fundamentals of normative decision-making. Later, Bernoulli's idea was revisited by von Neumann and Morgenstern [45], who were able to derive the axiomatic foundations of his approach. Thus, the framework of the expected value is often referred to as the Bernoulli-Principle or the Neumann-Morgenstern-utility. Consequently, the first utility function is defined as

$$\Xi_1 := \mathbb{E}\left\{\widetilde{\mathrm{LCOH}}\right\}, \tag{10.16}$$

and the corresponding objective function is min Ξ_1. The second utility function addresses the involved risk that goes hand in hand with a possible investment. In financial analysis, the cumulative distribution function CDF of the objective function is referred to as the risk profile [46]. If an agent is risk averse, he might pursue a system design with greater $\widetilde{\mathrm{LCOH}}$ but a lower assigned involved risk of exceeding the expected value. Because the involved risk only refers in this case to one side of the risk profile, often used methods of risk approximation using the variance of the risk profile are not suitable. Thus, the method of lower-partial-moments is used here. The method was scientifically first introduced by Bawa [47] and Jean [48] and later revisited by Fishburn [49], who presented the method in its most general way. The lower-partial-moment LPM(n, τ) of the degree n and the reference value τ is defined as

$$\mathrm{LPM}\,(n, \tau) := \int_{-\infty}^{\tau} (\tau - x)^n f(x) \mathrm{d}x. \tag{10.17}$$

The LPM of degree $n = 0$ is equal to the cumulative probability of the reference value τ, $\mathrm{LPM}\,(0, \tau) = \int_{-\infty}^{\tau} (\tau - x)^0 f(x) \mathrm{d}x = F\,(\tau)$. In order to address the involved risk, the second utility function is defined as a combination of the expected value and the zero-order LPM with the reference value $\mathbb{E}\left\{\widetilde{\mathrm{LCOH}}\right\}$

$$\Xi_2 := \mathbb{E}\left\{\widetilde{\mathrm{LCOH}}\right\} - \omega \cdot \mathrm{LPM}\left(0, \mathbb{E}\left\{\widetilde{\mathrm{LCOH}}\right\}\right). \tag{10.18}$$

It can be seen that the objective function is "rewarded" for greater cumulative probabilities of the expected values. The corresponding optimization objective is min Ξ_2. The weighting factor ω specifies the degree of risk aversion of an agent. The greater one chooses ω, the more important becomes the cumulative probability of the expected value in contrast to the expected value itself. Obviously, for $\omega = 0$ the utility function Ξ_2 becomes the first utility function Ξ_1. The constraints of the respective stochastic optimization problem consists of all the equations found in the second step of the MPSA, the system analysis. Because the system was analyzed energetically and economically, the constraints consist of these equations and forms therewith the stochastic techno-economic optimization problem. The two optimization problems are solved using Monte-Carlo-Simulations. The decision variables are the two design parameter vessel length L and the turbine area A_T. The results of the optimization are shown in Fig. 10.7 for the two optimization functions and varying values of ω.

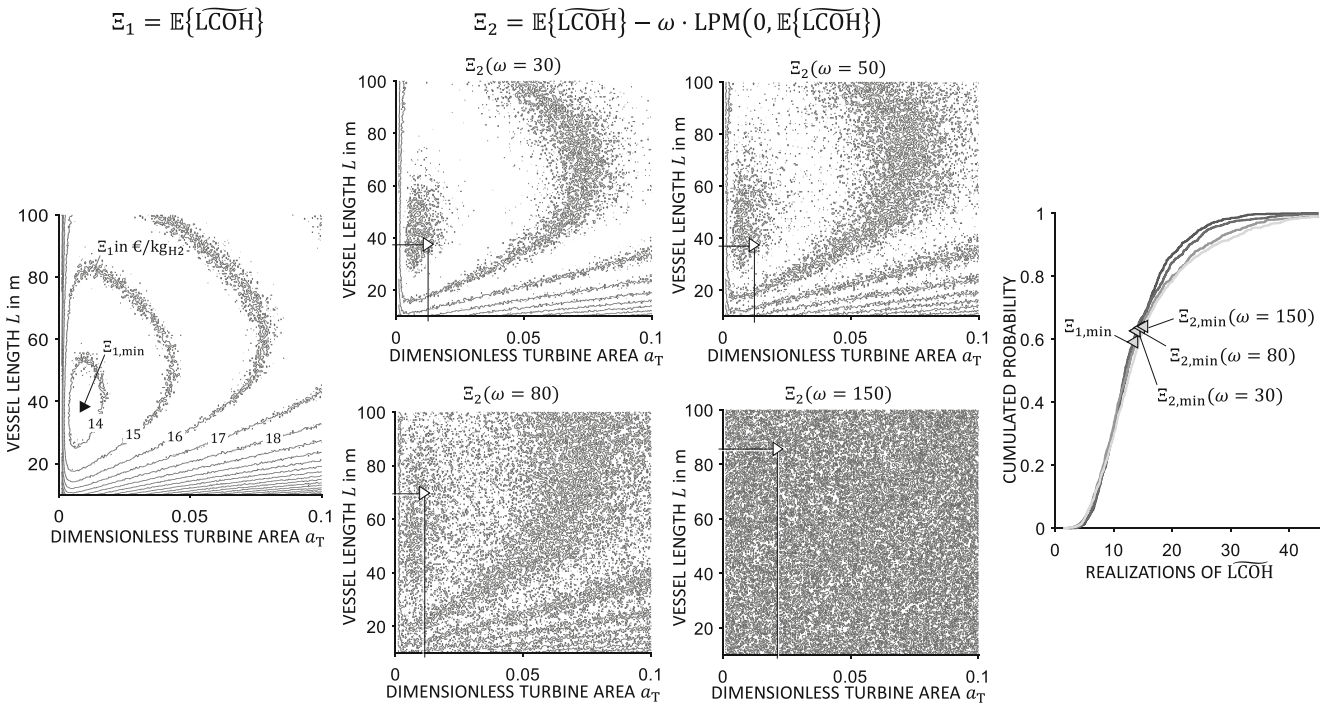


Fig. 10.7 Results of the two objective functions with varying weighting factors $\boldsymbol{\omega}^T = [30, 50, 80, 150]$ and their respective expected value (*right*)

The results of the objective function Ξ_1 (Bernoulli principle) can be seen in Fig. 10.7 (left) with the specific optimal design specified by $\Xi_{1,\min}$. The results have the shape of the deterministic optimal system, which has been presented in a previous paper [16] with the obvious disadvantage of missing robustness. The dispersion of the data is due to the statistics of the Monte-Carlo-Simulations. The optimal system designs for the objective function Ξ_2 (Combination of expected value and Lower-Partial-Moment) differ greatly according to their respective weighting factors ω. It can be seen, that the spread of the results increase for increasing weighting factors ω so that an agent faces greater indifference in choosing the optimal system. It shall be mentioned here that the absolute numbers of the objective Ξ_2 have no techno-economic meaning, whereas the objective Ξ_1 does. The defined objective Ξ_2 is only interpretable to determine the optimal system design under consideration of the involved risk. For instance, it can be seen that no difference of the optimal systems $\Xi_{2,\min}(\omega = 30)$ and $\Xi_{2,\min}(\omega = 50)$ exists. On the right side of Fig. 10.7, one can see the respective CDFs of the optimal solutions found and the corresponding expected value. It can be seen that the expected value of $\Xi_{1,\min}$ is the smallest, but the cumulative probability of the expected value is smaller compared to all other optimal systems. The figure also shows the increase of the cumulative probability of the expected value for the optimal solutions Ξ_2. But on the other hand this increase goes hand in hand with an increasing expected value. Thus, one is able to model the investment behavior of an agent by finding an appropriate weighting factor ω, e.g. a stringent risk-averse investment behavior can be modeled by choosing a great weighting factor ω. With this method, a specific techno-economic optimal system design can be assigned to every agent.

For the further analysis the system consideration is set to the optimal solution $\Xi_{2,\min}(\omega = 30)$, which is further evaluated in the last step of the method, the sensitivity analysis.

10.2.4 Sensitivity Analysis

In the last step of the MPSA method dominant input parameters are identified, the model quality is assessed and the cause and effect relationship of the model is revealed. The optimal system has been identified in the second step and the accompanying risk profile is known. Methods of sensitivity analysis provide information of how the uncertainty in the model output (risk profile) can be assigned to the different sources of uncertainty in the model input. Therefore, different methods exist and can be found in several reviews [50–53]. The most common methods are the method of variance-based sensitivity analysis (VBSA) and density-based sensitivity analysis DBSA. VBSA approximates uncertainty reduction as the reduction of variance for conditional and unconditional distribution functions and refers therewith to a symmetric statistical moment.

This can lead to misleading results, if out- or input distributions are not symmetric but skewed. Several examples can be found in the literature [50, 54, 55] as well as a recently published paper by the first and second author and Platzer [56], which presents this phenomenon on the here shown energy converter, treated as a deterministic optimization problem. DBSA on the other hand considers the entire output distribution and thus, does not refer to a statistical moment. Based on the advantage of moment-independence, only DBSA methods are considered here. The general form of a density-based sensitivity-index T_i, measuring the influence of the i-th parameter for uncertainty reduction, is given by Pianosi et al. [53]

$$T_i = \underset{x_i}{\text{stat}}\ \text{divergence}\left(f_Y | f_{Y|X_i=x_i}\right), \tag{10.19}$$

with the unconditional PDF of the model output f_Y and the conditional PDF of the model output $f_{Y\left|X_i=x_i\right.}$. The conditional PDF describes the PDF of the model output if the uncertainty of the input factor X_i disappears. The term 'stat' refers to some statistics, e.g. max, mean or median and 'divergence' denotes the divergence of the unconditional and conditional PDF. For instance, the δ-sensitivity method defines uncertainty as the area enclosed by unconditional and conditional PDFs [57]. In this paper the recently published PAWN method (named by the developers) is used, which uses the so called Kolmogorov-Smirnov statistics KS as uncertainty approximation [58–60]. The KS is attributed to Kolmogorov [61] and is a scale-invariant geometrically interpretable quantity that considers the maximum distance between conditional and unconditional CDFs

$$\text{KS}\,(x_i) = \max_y \left| F_Y(y) - F_{Y|X_i=x_i}(y) \right|. \tag{10.20}$$

The sensitivity index T_i is defined as the median of all calculated KS statistics. Thus, high scatterings of the conditional CDFs around the unconditional CDF implies a high sensitivity of the considered input variable

$$T_i := \text{median}\,(\text{KS}\,(x_i))\,. \tag{10.21}$$

The sensitivity indices can be seen sorted according to their respective influence and are shown in Fig. 10.8 (left) with their corresponding 95%-confidence intervals. This process is called **factor prioritization**. It can be seen that a reduction of the lift coefficient uncertainty $\widetilde{c}_\text{L}$ and the vessel drag coefficient uncertainty $\widetilde{c}_\text{D}$ will have the most significant impact on the uncertainty of the levelized costs $\widetilde{\text{LCOH}}$.

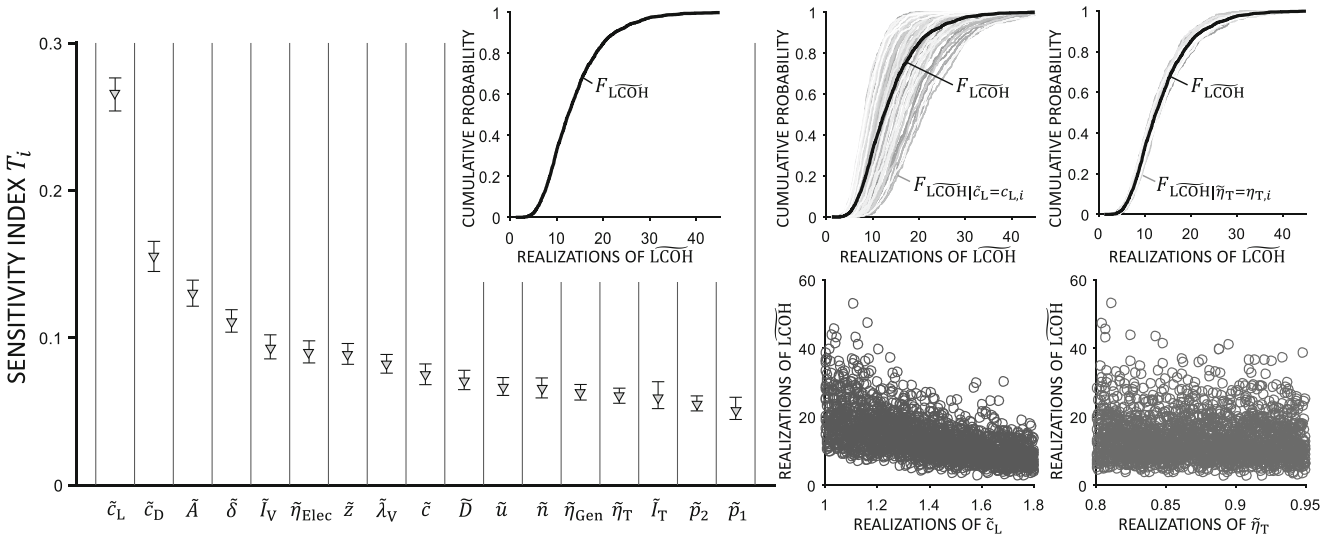


Fig. 10.8 Factor prioritization and exemplary illustration of the unconditional and conditional CDFs as well as scatter plots for a significant and a less significant input factor

Figure 10.8 (right) shows as an illustrative example for two input factors, the relevant input factor $\widetilde{c}_{\mathrm{L}}$ and the less relevant input factor $\widetilde{\eta}_{\mathrm{T}}$ the unconditional CDF $F_{\widetilde{\mathrm{LCOH}}}$ and the respective conditional CDFs. It can be seen on the example of the significant input factor $\widetilde{c}_{\mathrm{L}}$, that the unconditional CDFs scatter greatly around the conditional one, implying that the model output reacts very sensitive on input factor variations. This can also be seen on the corresponding scatter plot, also shown in Fig. 10.8. The model output decreases significantly for increasing lift coefficients. In contrast, the conditional CDFs of the turbine efficiency factor do not vary that much. It can also be seen on the scatter plot that the model output does not react very sensitive to variations of the turbine efficiency factor.

10.3 Conclusion

In this paper the method of multi-pole system analysis (MPSA) is presented in its most general form and also applied to an innovative wind-energy converter. The method consists of the four steps of (1) system synthesis, (2) system analysis under uncertainty, (3) stochastic optimization and (4) sensitivity analysis. The first step provides answers on the question of the considered framework of the system. The topology of the system is defined and a superior system model is achieved showing all the coupling effects of the modeled criteria. These coupling effects already impressively verify that an isolated system consideration is not suitable. In the second step, domain specific white-box models are derived by applying e.g., first principles in the technical domain or fundamental models in the economic domain. Empirical scaling laws, based on market surveys, literature review or expert opinions are used to describe not only one but a multitude of possible systems. In this paper we present for the first time a system analysis under uncertainty, using the probabilistic approach. Thus, each uncertain input factor is mathematically modeled using an adequate probability density function. In the third step, the optimal system of all described ones is identified by deriving a stochastic optimization problem, which consists of an objective function according to an agents preferences subjected to all the constraints, which have been found in the second step. By doing so, one is guided in a structured manner to a complete optimization problem which is solved here using Monte-Carlo-Simulations. The cause- and effect relationship, dominant input factors and the model quality is assessed by using methods of sensitivity analysis in the last step of the method. The results reveal how the uncertainty of the model output can be assigned to the uncertainty of the model inputs.

The method is presented in this paper on the example of an innovative wind-energy converter, which makes vast ocean wind energy accessible. It can be seen, that only a simultaneous consideration of energetic and economic aspects leads to a truly optimal system design. Due to the incorporation of uncertainty in the stage of early system design, different optimal systems are presented according to an agents preferences. It is shown that the optimal system design, resulting of the minimization of the expected utility differs greatly from the optimal system that provides minimal investment risk. Also, due to the incorporation of uncertainties, the disadvantage of missing robustness of the previous deterministic optimal system design is overcome.

Acknowledgement The authors would like to thank the Deutsche Forschungsgemeinschaft (DFG) for the financial support in the framework of the Excellence Initiative, Darmstadt Graduate School of Excellence Energy Science and Engineering (GSC 1070).

References

1. Salomon, R.E.: Process of converting wind energy to elemental hydrogen and apparatus therefor. US Patent 4,335,093, Temple University; 1982
2. Meller, M.: Wind-powered linear motion hydrogen production systems. US Patent 7,146,918 B2, 2006
3. Holder, K.L.: Einrichtung zum Umwandeln von Windenergie über dem offenen Wasser, insbesondere Ozean, in elektrische Energie. DE 10 2007 057 267 A1, 2008
4. Gizara, A.R.: Turbine-integrated hydrofoil. US Patent 7,298,056 B2, Integrated Power Technology Corporation, 2007
5. Platzer, M.F., Lennie, M., Vogt, D.M.: Analysis of the Conversion of Ocean Wind Power into Hydrogen. Proceedings of the World Renewable Energy Congress, Murdoch (2013)
6. Platzer, M.F., Sanz, W., Jericha, H.: Renewable power via energy ship and graz cycle. In: ISROMAC-15. Honolulu, USA, 2014
7. Platzer, M.F., Sarigul-Klijn, N.: A novel approach to extract power from free-flowing water and high altitude jet streams. In: ASME 2009 3rd International Conference on Energy Sustainability, San Francisco, California, USA, 2009, pp. 493–499

8. Platzer, M.F., Sarigul-Klijn, N.: A new oscillating-foil power generator for sailingship-based renewable energy generation. In: ASME 2010 4th International Conference on Energy Sustainability, American Society of Mechanical Engineers, 2010, pp. 909–916
9. Platzer, M.F., Sarigul-Klijn, N., Young, J., Ashraf, M.A., Lai, J.C.S.: Renewable hydrogen production using sailing ships. J. Energy Resour. Technol. **136**, 021203 (2014)
10. Kim, J.: Electric power generation system using hydro turbine tracted by paraglider. US Patent 2010/0001534 A1, 2010
11. Kim, J., Park, C.: Economy of hydrogen production by Parafoil-pulled ships. J. Energy Power Sour. **1**, 9–16 (2014)
12. Kim, J., Park, C.: Wind power generation with a parawing on ships, a proposal. Energy. **35**, 1425–1432 (2010)
13. Holl, M., Platzer, M., Pelz, P.F.: Techno-economical system optimisation and its application to an energy system. In: Energy-Science and Technologie EST, p. 406. KIT, Karlsruhe (2015)
14. Holl, M., Platzer, M., Pelz, P.F.: Optimal energy systems design applied to an innovative ocean-wind energy converter. In: 7th International Conference on Sustainable Development and Planning, pp. 547–557. WIT Press, Istanbul (2015)
15. Pelz, P.F., Holl, M., Platzer, M.: Analytical method towards an optimal energetic and economical wind-energy converter. Energy. **94**, 344–351 (2016)
16. Holl, M., Pelz, P.F.: Multi-pole system analysis (MPSA)—A systematic method towards techno-economic optimal system design. Appl. Energy. **169**, 937–949 (2016)
17. Buchholz, P.: Modellgestützte Analyse und Optimierung. Technische Universität Dortmund, Dortmund (2015)
18. Carnot, S.: Betrachtungen über die bewegende Kraft des Feuers. Wilhelm Engelmann, Leipzig (1892)
19. Betz, A.: Das maximum der theoretisch möglichen Ausnützung des Windes durch Windmotoren. Zeitschrift für das gesamte Turbinenwesen. **26**, 307–309 (1920)
20. Pelz, P.F.: Upper limit for hydropower in an open-channel flow. J. Hydraul. Eng. **137**, 1536–1542 (2011)
21. Aven, T.: Some reflections on uncertainty analysis and management. Reliab. Eng. Syst. Saf. **95**, 195–201 (2010)
22. Aven, T., Zio, E.: Some considerations on the treatment of uncertainties in risk assessment for practical decision making. Reliab. Eng. Syst. Saf. **96**, 64–74 (2011)
23. Devooght, J.: Model uncertainty and model inaccuracy. Reliab. Eng. Syst. Saf. **59**, 171–185 (1998)
24. Hoffman, F.O., Hammonds, J.S.: Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Anal. **14**, 707–712 (1994)
25. Nilsen, T., Aven, T.: Models and model uncertainty in the context of risk analysis. Reliab. Eng. Syst. Saf. **79**, 309–317 (2003)
26. Zio, E., Pedroni, N.: Literature review of methods for representing uncertainty. In: FonCSI, 2013
27. Knight, F.H.: Risk, Uncertainty and Profit. Hart, Schaffner and Marx, New York (1921)
28. von Helmholtz, H.L.F.: Über die Erhaltung der Kraft, eine physikalische Abhandlung. Reimer, Berlin (1847)
29. Tietjens, O.K.G., Prandtl, L.: Applied Hydro-and Aeromechanics: Based on Lectures of L. Prandtl. Courier Corporation, New York (1957)
30. Sachs, L., Hedderich, J.: Angewandte Statistik: Methodensammlung mit R. Springer, Berlin (2009)
31. Brockhoff, K.: Geschichte der Betriebswirtschaftslehre-Kommentierte Meilensteine und Originaltexte. Wiesbaden, Gabler Verlag (2000)
32. Schneider, D.: Betriebswirtschaftslehre, Band 4: Geschichte und Methoden der Wirtschaftswissenschaft. München, Wien, Oldenburg Verlag (2001)
33. Stevin, S.: Tafelen Van Interest, Midtgaders de Constructie der Selver. Christoffel Plantijn, Antwerp (1582)
34. Leibniz, G.W.: Meditatio juridico-mathematica de interusurio simplice. In: Knobloch, E., Graf von der Schulenberg, J.-M. (eds.) Hauptschriften zur Versicherungs-und Finanzmathematik. Oldenburg Akademieverlag, Berlin (2000)
35. Dantzig, G.B.: Linear programming under uncertainty. Manag. Sci. **1**, 197–206 (1955)
36. Kall, P., Wallace, S.W.: Stochastic Programming. Springer, New York (1994)
37. Birge, J.R., Louveaux, F.: Introduction to stochastic programming. Springer Science & Business Media, New York (2011)
38. Wets, R.J.-B.: Chapter VIII: Stochastic programming. Handbooks Oper. Res. Manag. Sci. **1**, 573–629 (1989)
39. Graf, M.: Financial Risk Management: State-of-the-Art [Bachelorthesis], 2005
40. Hildebrandt, P., Knoke, T.: Investment decisions under uncertainty—a methodological review on forest science studies. Forest Policy Econ. **13**, 1–15 (2011)
41. Kruschwitz, L., Husmann, S.: Finanzierung und Investition. Walter de Gruyter, Berlin (2012)
42. Poggensee, K.: Investitionsrechnung. Wiesbaden, Gabler (2008)
43. Trautmann, S.: Investitionen: Bewertung, Auswahl und Risikomanagement. Springer-Verlag, Berlin (2007)
44. Bernoulli D. Specimen Theoriae novae de Mensure Sortis. Cementer Academia Scientiarum Imperalis Petropolitanae. 1738:175–92.
45. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton, NJ (1944)
46. Baucells, M., Borgonovo, E.: Invariant probabilistic sensitivity analysis. Manag. Sci. **59**, 2536–2549 (2013)
47. Bawa, V.S.: Optimal rules for ordering uncertain prospects. J. Financ. Econ. **2**, 95–121 (1975)
48. Jean, W.H.: Comparison of moment and stochastic dominance ranking methods. J. Financ. Quant. Anal. **10**, 151–161 (1975)
49. Fishburn, P.C.: Mean-risk analysis with risk associated with below-target returns. Am. Econ. Rev. **67**, 116–126 (1977)
50. Borgonovo, E., Plischke, E.: Sensitivity analysis: a review of recent advances. Eur. J. Oper. Res. **248**, 869–887 (2016)
51. Campolongo, F., Saltelli, A.: Sensitivity analysis of an environmental model: an application of different analysis methods. Reliab. Eng. Syst. Saf. **57**, 49–69 (1997)
52. Kurowicka, D., Cooke, R.: Uncertainty Analysis with High Dimensional Dependence Modelling. Wiley, New York (2006)
53. Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., et al.: Sensitivity analysis of environmental models: a systematic review with practical workflow. Environ. Model. Softw. **79**, 214–232 (2016)
54. Borgonovo, E., Castaings, W., Tarantola, S.: Moment independent importance measures: new results and analytical test cases. Risk Anal. **31**, 404–428 (2011)
55. Liu, H., Chen, W., Sudjianto, A.: Relative entropy based method for probabilistic sensitivity analysis in engineering design. J. Mech. Des. **128**, 326–336 (2006)

56. Holl, M., Janke, T., Pelz, P.F., Platzer, M.: Sensitivity analysis of a techno-economic optimal wind-energy converter. In: 2nd International Conference on Next Generation Wind Energy, Lund, Sweden, 2016
57. Borgonovo, E.: A new uncertainty importance measure. Reliab. Eng. Syst. Saf. **92**, 771–784 (2007)
58. Pianosi, F., Sarrazin, F., Wagener, T.: A Matlab toolbox for global sensitivity analysis. Environ. Model. Softw. **70**, 80–85 (2015)
59. Pianosi, F., Sarrazin, F., Wagener, T.: SAFE Toolbox, 2016
60. Pianosi, F., Wagener, T.: A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. Environ. Model. Softw. **67**, 1–11 (2015)
61. Kolmogorov, A.N.: Sulla determinazione empirica delle leggi di probabilita. Giorn Ist. Ital. Attuari. **4**, 83–91 (1933)

Chapter 11
Linear Parameter-Varying (LPV) Buckling Control of an Imperfect Beam-Column Subject to Time-Varying Axial Loads

Maximilian Schaeffner and Roland Platz

Abstract In this paper, active buckling control of an imperfect slender beam-column with circular cross-section by piezo-elastic supports and Linear Parameter-Varying (LPV) control is investigated experimentally. The beam-column is loaded by a time-varying axial compressive load resulting in a lateral deflection of the beam-column due to imperfections. A finite element model of the beam-column under axial load is designed as an LPV system. A reduced and augmented modal model is used to design a quadratically stable gain scheduled LPV control. The control is implemented in an experimental test setup and the maximum bearable loads of the beam-column are obtained. Two cases are tested: with and without LPV control or, respectively, active and passive configuration. With the proposed active LPV buckling control it is possible to compensate the influence of beam-column imperfections and to compensate uncertainty in mounting and loading that in passive configuration without LPV control may lead to early buckling. Eventually, the maximum bearable axial compressive load is increased above the theoretical critical buckling load.

Keywords LPV control • Active buckling control • Uncertainty • Imperfect beam-column • Piezo-elastic supports

11.1 Introduction

For passive compressively loaded beam-columns, the theory of buckling has been thoroughly investigated for both ideal and imperfect structures [1]. Imperfect beam-columns exhibit large lateral deformations for axial loads considerably below the theoretical critical buckling load and, therefore, have lower maximum bearable axial loads. A general approach to passively increase the maximum bearable axial load is to change the geometry, e.g. length and cross-section area, or the material so that the beam-column withstands higher loads. This, however, is sometimes not desirable because of given design constraints. In these cases, active buckling control without significant change in the beam-column's geometry and material provides a suitable approach to increase the maximum bearable axial load of a given structure. Active buckling control of slender beam-columns with different boundary conditions has been investigated numerically and experimentally several times, [2–8]. The investigated structures all had rectangular cross-sections, relatively high slenderness ratios s, with $300 \leq s \leq 1760$ and low absolute values of the theoretical critical buckling load. Furthermore, the structures were subject to (quasi-) static axial loads and controllers were also static or switched manually. Often, surface bonded piezoelectric patches were applied to beam-columns with rectangular cross-section to induce active bending moments that counteract the deformation [2–6]. The active stabilization concept investigated by earlier own studies [7, 8] used piezoelectric stack actuators to apply active lateral forces near the base of a fixed-pinned beam-column with rectangular cross-section.

To the authors' knowledge, active buckling control of beam-columns with circular cross-section has not yet been investigated except in own works [9, 10]. In both investigations, active buckling control of a circular beam-column with relatively low slenderness ratio $s = 108$ with piezo-elastic supports at both ends and stabilized by integral linear quadratic control (LQR) was investigated numerically and experimentally for (quasi-)static axial loads. The integral LQR was able to compensate initial deformation that could lead to buckling and increase the maximum bearable load of the beam-column in numerical simulation and experiment. Control matrices were switched manually due to increasing axial loads and the maximum bearable axial load was limited by the theoretical critical buckling load of the beam-column with piezo-

M. Schaeffner (✉)
System Reliability, Adaptronics and Machine Acoustics SAM, Technische Universität Darmstadt, Magdalenenstraße 4, 64289 Darmstadt, Germany
e-mail: schaeffner@szm.tu-darmstadt.de

R. Platz
Fraunhofer Institute for Structural Durability and System Reliability LBF, Bartningstraße 47, 64289 Darmstadt, Germany

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_11

elastic supports at both ends. Now, in this paper, active buckling control of the same slender beam-column with improved piezo-elastic supports and LPV control is investigated experimentally. The beam-column is loaded by a time-varying axial compressive load resulting in a lateral deflection of the beam-column due to imperfections. The quadratically stable gain scheduled LPV control is designed and implemented in an experimental test setup to account for the axial load dependency of the beam-columns lateral stiffness.

11.2 System Description and Mathematical Model of Beam-Column System

In the following, first the investigated beam-column with piezo-elastic supports and the experimental test setup for active LPV buckling control are presented. Second, a mathematical finite element (FE) model of the axially loaded beam-column is derived and validated with experimental results. The investigated system is a slender beam-column made of aluminum alloy EN AW-7075 with length $l_\mathrm{b} = 400\,\mathrm{mm}$ and circular solid cross-section with radius $r_\mathrm{b} = 4\,\mathrm{mm}$, Fig. 11.1. It has a Young's modulus $E_\mathrm{b} = 75.8 \times 10^3\,\mathrm{N/mm^2}$ and density $\varrho_\mathrm{b} = 2.79 \times 10^{-3}\,\mathrm{g/mm^3}$, all assumed to be constant across the entire beam-column length.

The beam-column is embedded in two piezo-elastic supports A at $x = 0$ and B at $x = l_\mathrm{b}$. At support B, a time-varying axial load $F_x(t)$ is applied. Strain gauges at sensor position $x_\mathrm{s} = l_\mathrm{b}/2$ are used to measure the surface strain of the beam-column in y- and z-direction. Figure 11.2 shows sectional views of piezo-elastic support A [11]. The central element of the support are two concave-shaped elastic membrane springs made of spring steel 1.1248 that bear the axial compressive and lateral loads and allow rotations in any plane perpendicular to the x-axis, Fig. 11.2a, b. They are represented by rotational stiffness $k_{\varphi_y,\mathrm{A}} = k_{\varphi_z,\mathrm{A}} = k_{\varphi_y,\mathrm{B}} = k_{\varphi_z,\mathrm{B}} = k_\mathrm{r}$ and lateral stiffness $k_{y,\mathrm{A}} = k_{z,\mathrm{A}} = k_{y,\mathrm{B}} = k_{z,\mathrm{B}} = k_\mathrm{l}$ that are the same for both supports A and B and in y- and z-direction, Fig. 11.1b.

In each piezo-elastic support A and B at $x = -l_\mathrm{ext}$ and $x = l_\mathrm{b} + l_\mathrm{ext}$, two piezoelectric stack actuators *Piezomechanik PST 150* are arranged orthogonal to each other and orthogonal to the beam-column's x-axis, acting in y- and z-direction, Fig. 11.2c.

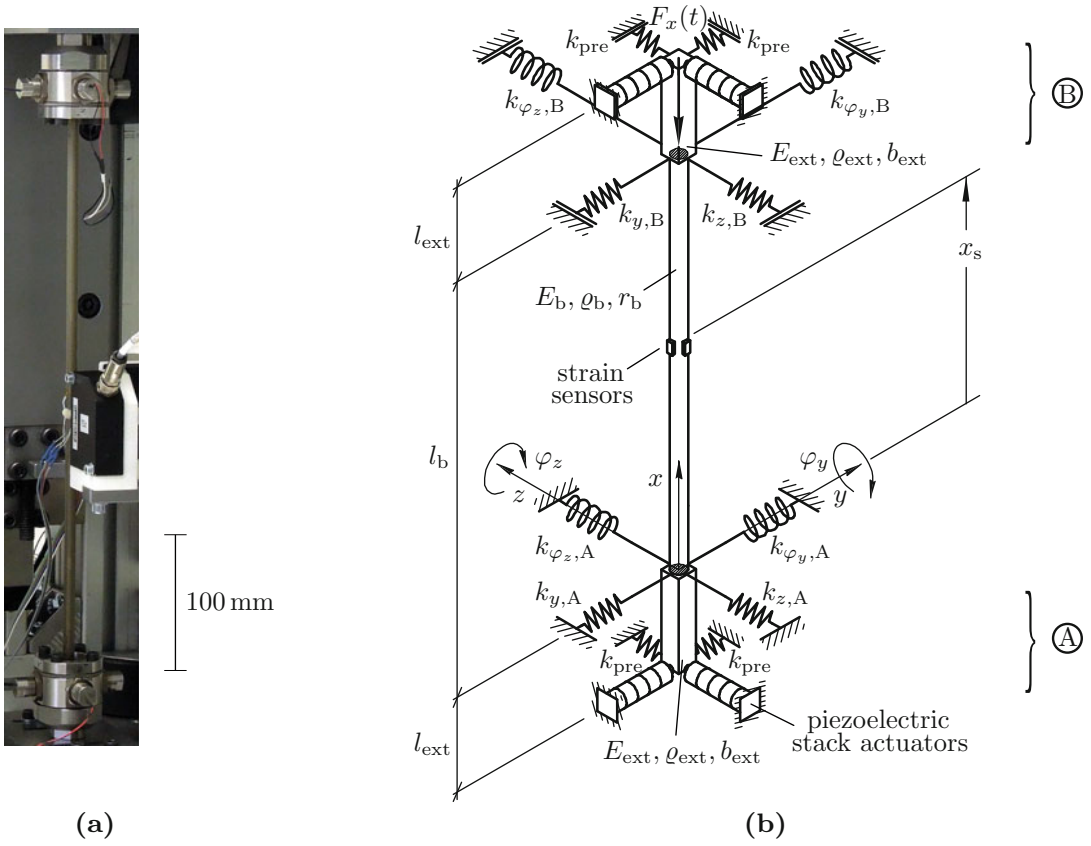


Fig. 11.1 Beam-column system, (**a**) beam-column with piezo-elastic supports for experimental test, (**b**) sketch of beam-column

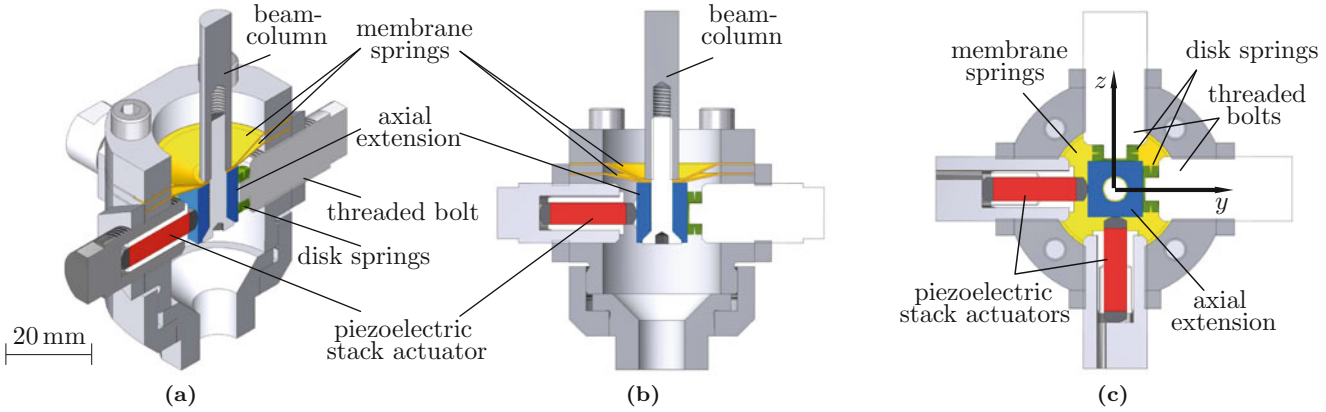


Fig. 11.2 Sectional views of piezo-elastic support A, (**a**) 3D section through x–y-plane, (**b**) 2D section through x–y-plane, (**c**) 2D section bottom view through actuator-plane

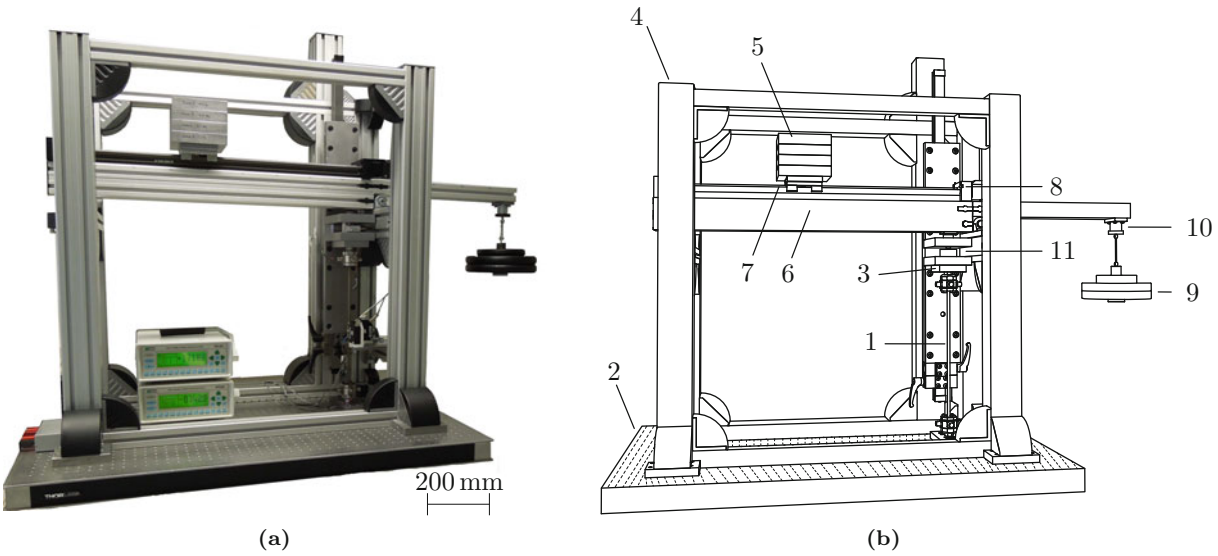


Fig. 11.3 Experimental test setup for active buckling control, (**a**) photo of test setup, (**b**) CAD sketch of test setup (without strain gauge amplifiers)

Each actuator is mechanically prestressed by an allocated stack of disc springs with stiffness $k_{\text{pre}} = 2.6 \times 10^3\,\text{N/mm}$. The initial deflection of the disk springs is adjusted by threaded bolts. The piezoelectric stack actuators exert lateral forces to the beam-column axial extension with quadratic cross-section and relatively high bending stiffness made of hardened steel 1.4301 with length $l_{\text{ext}} = 7.9\,\text{mm}$, edge length $b_{\text{ext}} = 6\,\text{mm}$, Young's modulus $E_{\text{ext}} = 210.0 \times 10^3\,\text{N/mm}^2$ and density $\varrho_{\text{ext}} = 7.81 \times 10^{-3}\,\text{g/mm}^3$. The extension works as cantilever beam end beyond the elastic membrane springs for each support A and B. This way, controlled active lateral forces by the two piezoelectric stack actuators act in arbitrary directions orthogonal to the beam-column's longitudinal x-axis. They result in bending moments acting in arbitrary directions at the beam-column ends in both piezo-elastic supports A and B.

For experimental investigations, the beam-column system is integrated in an experimental test setup. Figure 11.3 shows a photo and a CAD sketch of the experimental test setup for the active buckling control of the slender beam-column with circular cross-section and piezo-elastic supports, 1, Fig. 11.1a. The lower piezo-elastic support A is fixed on a baseplate 2, the upper piezo-elastic support B is fixed to a parallel guidance 3 that is connected to a stiff frame structure 4. The beam-column is loaded via a high lever ratio by a mass 5 of up to 30 kg that is placed on a hinged beam 6. The lever ratio and, therefore, the axial load are increased by shifting the mass in nearly horizontal plane via a linear axle 7, operated by a stepper motor 8. An additional optional weight 9 of up to 15 kg reduces the axial load on the beam-column. It can be released by an electromagnet 10, thus generating a step axial compressive load. The axial load is measured by a load cell 11. With the presented test setup, ramp-shaped axial loads between the static load $F_{\text{min}} = 330\,\text{N}$ of the parallel guidance 3, and $F_{\text{max}} \approx 5000\,\text{N}$ as well as step-shaped axial loads of up to $F_{\text{step}} = 800\,\text{N}$ can be generated. This is sufficient to generate

static and time-dependent axial loads $F_x(t)$ considerably above the theoretical EULER buckling load of the fixed-fixed beam-column with $F_{x,\mathrm{E}} = 4\,\pi^2 E_\mathrm{b} I_\mathrm{b}/l_\mathrm{b}^2 = 3760\,\mathrm{N}$ [1].

11.2.1 Finite Element State Space Model of Beam-Column System

To describe the lateral vibration of the axially loaded beam-column in state space, a finite element (FE) model is the basis. The beam-column and the stiff axial extensions are modeled by the FE equation of motion

$$\boldsymbol{M}\,\ddot{\boldsymbol{r}} + \boldsymbol{D}\,\dot{\boldsymbol{r}} + \left(\boldsymbol{K}_\mathrm{e} - F_x(t)\,\boldsymbol{K}_\mathrm{g}\right)\boldsymbol{r} = \boldsymbol{B}_0\,\boldsymbol{u} \tag{11.1}$$

and output equation

$$\boldsymbol{y} = \begin{bmatrix} \varepsilon_{\mathrm{s},y} \\ \varepsilon_{\mathrm{s},z} \end{bmatrix} = \boldsymbol{C}_0\,\boldsymbol{r}. \tag{11.2}$$

In (11.1), $\boldsymbol{r}$ is the $[4N \times 1]$ FE displacement vector for N nodes and $4N$ degrees of freedom with two translational and two rotational displacements in and around y- and z-direction. $\boldsymbol{M}$ is the mass matrix, $\boldsymbol{D}$ is the damping matrix, $\boldsymbol{K}_\mathrm{e}$ is the elastic stiffness matrix and $\boldsymbol{K}_\mathrm{g}$ is the geometric stiffness matrix that describes the influence of axial load $F_x(t)$ on the beam-column's lateral stiffness, all $[4N \times 4N]$. The stiffness of the elastic membrane springs, disk springs and piezoelectric stack actuators are included in the stiffness matrices.

The right side of (11.1) represents the active control forces acting simultaneously at both supports A and B that are summarized in control input vector

$$\boldsymbol{u} = \begin{bmatrix} F_{\mathrm{a},y} \\ F_{\mathrm{a},z} \end{bmatrix}, \tag{11.3}$$

and that are derived in Sect. 11.3.2. The $[4N \times 2]$ control input matrix

$$\boldsymbol{B}_0 = \begin{bmatrix} 1\,0\,0\,0 \cdots 1\,0\,0\,0 \\ 0\,1\,0\,0 \cdots 0\,1\,0\,0 \end{bmatrix}^T \tag{11.4}$$

allocates the active forces of the piezoelectric stack actuators to the lateral degrees of freedom of the first and last nodes of the FE model. In (11.2), the $[2 \times 4N]$ output matrix $\boldsymbol{C}_0$ allocates the surface strains $\boldsymbol{y}$ to the FE displacement vector $\boldsymbol{r}$ [9]. For convenience, (11.1) and (11.2) are written in state space representation

$$\begin{aligned} \dot{\boldsymbol{x}}_\mathrm{FE} &= \underbrace{\left(\begin{bmatrix} \boldsymbol{0} & \mathbf{I} \\ -\boldsymbol{M}^{-1}\boldsymbol{K}_\mathrm{e} & -\boldsymbol{D} \end{bmatrix} + F_x(t) \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{M}^{-1}\boldsymbol{K}_\mathrm{g} & \boldsymbol{0} \end{bmatrix} \right)}_{[8N \times 8N]} \boldsymbol{x}_\mathrm{FE} + \underbrace{\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{M}^{-1}\boldsymbol{B}_0 \end{bmatrix}}_{[8N \times 2]} \boldsymbol{u} \\ \boldsymbol{y} &= \underbrace{\begin{bmatrix} \boldsymbol{C}_0 & \boldsymbol{0} \end{bmatrix}}_{[2 \times 8N]} \boldsymbol{x}_\mathrm{FE}, \end{aligned} \tag{11.5}$$

with $[8N \times 1]$ FE state vector $\boldsymbol{x}_\mathrm{FE} = [\boldsymbol{r}, \dot{\boldsymbol{r}}]^T$ and zero and identity matrices $\boldsymbol{0}$ and $\mathbf{I}$ of appropriate dimensions. The full FE state space model of the beam-column as short form of (11.5) can be written as

$$\begin{aligned} \dot{\boldsymbol{x}}_\mathrm{FE} &= \boldsymbol{A}_\mathrm{FE}(F_x(t))\,\boldsymbol{x}_\mathrm{FE} + \boldsymbol{B}_\mathrm{FE}\,\boldsymbol{u} \\ \boldsymbol{y} &= \boldsymbol{C}_\mathrm{FE}\,\boldsymbol{x}_\mathrm{FE}, \end{aligned} \tag{11.6}$$

[9]. The influence of axial load $F_x(t)$ on the system behavior is described by system matrix $\boldsymbol{A}_\mathrm{FE}(F_x(t))$. In the following, the time-dependent axial force will be written as $F_x = F_x(t)$.

11.2.2 System Identification and Model Validation of Beam-Column System

To validate the FE beam-column model, the transfer behavior of model (11.6) and experiment are compared. The two frequency response functions

$$H_y(F_x, \Omega) = \frac{\varepsilon_{\mathrm{s},y}(\Omega)}{F_{\mathrm{a},y}(\Omega)} \qquad \text{and} \qquad H_z(F_x, \Omega) = \frac{\varepsilon_{\mathrm{s},z}(\Omega)}{F_{\mathrm{a},z}(\Omega)} \tag{11.7}$$

represent the surface strain responses $\varepsilon_{\mathrm{s},y/z}$ at $x = x_\mathrm{s}$ in y- and z-direction, Fig. 11.1b, to an excitation by the active forces $F_{\mathrm{a},y/z}$ at $x = -l_\mathrm{ext}$ and $x = l_\mathrm{b} + l_\mathrm{ext}$ in y- and z-direction with excitation angular frequency Ω. Figure 11.4a shows the amplitude $|H_{y/z}|$ and phase $\arg H_{y/z}$ of the frequency response functions H_y and H_z for the beam-column loaded with static axial load $F_x = 2000\,\mathrm{N}$ for model and experiment. Amplitude and phase of the frequency response functions from model (11.7) and both y- and z-direction in experiment coincide well for the considered frequency range $0\,\mathrm{Hz} < \Omega/2\pi < 250\,\mathrm{Hz}$. The first resonance frequency $\omega_1/2\pi = 142\,\mathrm{Hz}$ at the maximum amplitude $\max(|H_{y/z}|) = |H_{y/z}(F_x, \Omega = \omega_1)|$ also matches for numerical simulation and experiment.

With increasing static axial load F_x, the first resonance frequency of the beam-column decreases. Figure 11.4b shows the change of the first resonance frequency $\omega_1(F_x)/2\pi$ due to increasing static axial load of the experimental beam-column at five designated axial loads $F_x = 330, 1000, 1500, 2000, 2500\,\mathrm{N}$ and for the numerical simulation from $0\,\mathrm{N} < F_x < 4000\,\mathrm{N}$ for the beam-column with piezo-elastic supports and for ideal pinned-pinned and fixed-fixed supports. In the experiment, the system identification can only be performed up to axial loads of $F_x = 2500\,\mathrm{N}$ due to the high static deformations of the uncontrolled beam-column. In the numerical simulation, the first resonance frequency $\omega_1(F_x)/2\pi$ is calculated by solution of the well-known eigenvalue problem

$$\det\left[\left(\boldsymbol{K}_\mathrm{e} - F_x \boldsymbol{K}_\mathrm{g}\right) - \omega_1^2 \boldsymbol{M}\right] = 0. \tag{11.8}$$

As seen in Fig. 11.4b, the first resonance frequency decreases with increasing axial load F_x and the FE model (11.6) and the experiment agree well for the axial load range of $330\,\mathrm{N} < F_x < 2500\,\mathrm{N}$. In the numerical simulation, the first resonance frequency reaches zero for the critical buckling load that can, therefore, be calculated from the elastic and geometric stiffness matrices by solution of the modified eigenvalue problem (11.8) for $\omega_1 = 0$ by [12]

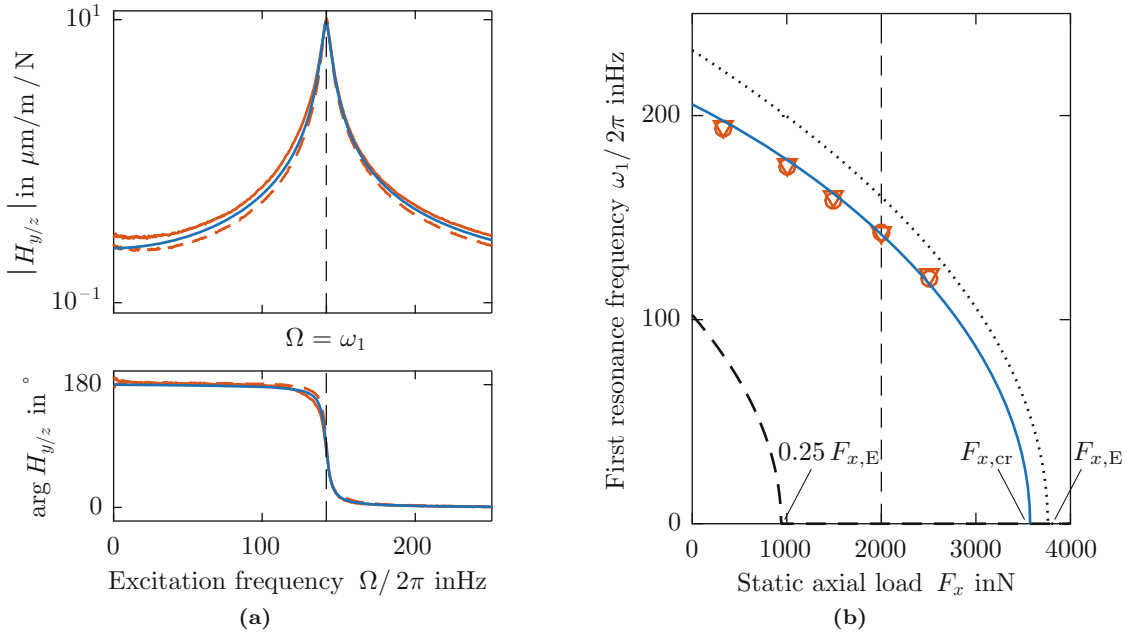


Fig. 11.4 System identification of beam-column system, (**a**) amplitude $|H_{y/z}|$ and phase $\arg H_{y/z}$ of frequency response functions $H_{y/z}$ for static axial load $F_x = 2000\,\mathrm{N}$ from numerical simulation (*blue solid line*) and experiment in y-direction (*red solid line*) and z-direction (*red dashed line*), (**b**) first resonance frequency $\omega_1/2\pi$ for varying axial load $0\,\mathrm{N} < F_x < 4000\,\mathrm{N}$ from simulation (*blue solid line*) and experiments in y-direction (*red inverted open triangle*) and z-direction (*red open circle*) as well as ideal pinned-pinned (*black dashed line*) and fixed-fixed (*black dotted line*) beam-column

$$\det\left[\boldsymbol{K}_\mathrm{e} - F_x\,\boldsymbol{K}_\mathrm{g}\right] = 0. \tag{11.9}$$

The decreasing course of the first resonance frequency $\omega_1/2\pi(F_x)$ for the beam-column with piezo-elastic supports in Fig. 11.4b is near the decreasing course of the first resonance frequency of the beam-column ideally fixed at supports A and B, whereas the distance between both curves decreases with increasing axial load. The slower decrease of the first resonance frequency with increasing axial load F_x results from a stiffening effect of the elastic membrane spring's lateral and rotational stiffness. Consequently, the lateral and rotational stiffness are modeled as

$$k_\mathrm{l} = k_\mathrm{l,e} + F_x\,k_\mathrm{l,g} \qquad \text{and} \qquad k_\mathrm{r} = k_\mathrm{r,e} + F_x\,k_\mathrm{r,g} \tag{11.10}$$

with constant stiffness for the unloaded beam-column $k_\mathrm{l,e} = 70.5 \times 10^3\,\mathrm{N/mm}$ and $k_\mathrm{r,e} = 388.8 \times 10^3\,\mathrm{N\,mm/rad}$ as well as axial force proportional component $k_\mathrm{l,g} = 9.5 \times 10^6\,\mathrm{N/mm}$ and $k_\mathrm{r,g} = 480.0\,\mathrm{N\,mm/rad}$. The lateral and rotational stiffness are added to the entries of the lateral and rotational degrees of freedom of nodes $n = 2$ and $n = N-1$ of the elastic and geometric stiffness matrices $\boldsymbol{K}_\mathrm{e}$ and $\boldsymbol{K}_\mathrm{g}$. For the given boundary conditions, the critical buckling load is determined to $F_{x,\mathrm{cr}} = 3573\,\mathrm{N}$ which is close to the theoretical critical buckling load of the fixed-fixed EULER beam-column $F_{x,\mathrm{E}} = 3760\,\mathrm{N}$, Fig. 11.4b.

11.3 Reduced State Space Control Model and LPV Control

In the following, first, the FE model of the axially loaded beam-column (11.6) will be reduced to a modal control model in order to reduce complexity and design the LPV control. The modal beam-column model is augmented to include integrated states to compensate imperfections of the beam-column. Second, the LPV control for active buckling control is derived.

11.3.1 Modal State Space Control Model

For the full state FE model (11.1), a number of $N = 35$ nodes resulting in $4N = 140$ degrees of freedom and modes is chosen to properly describe the maximum surface strains at sensor position x_s according to (11.2) and for adequate model validation in Fig. 11.4. For controller design, however, the FE model is reduced by modal truncation to only include the first lateral vibration mode for both y- and z-direction. Thus, the controller complexity is reduced considerably and the resulting modal beam-column model is observable and controllable with the given sensors and actuators, Sect. 11.2. Consequently, the FE displacement vector

$$\boldsymbol{r} \approx \boldsymbol{\Phi}\,\boldsymbol{q} = [\widehat{\boldsymbol{r}}_{1,y}, \widehat{\boldsymbol{r}}_{1,z}] \begin{bmatrix} q_{1,y} \\ q_{1,z} \end{bmatrix}. \tag{11.11}$$

is approximated by modal displacements $\boldsymbol{q}$ via the transformation with the $[4N \times 2]$ modal matrix $\boldsymbol{\Phi}$ that includes the first $[4N \times 1]$ eigenvectors $\widehat{\boldsymbol{r}}_{1,y}$ and $\widehat{\boldsymbol{r}}_{1,z}$ in y- and z-direction of the FE Model [13]. The modal matrix $\boldsymbol{\Phi}$ is normalized with respect to mass matrix $\boldsymbol{M}$ leading to the modal mass matrix $\boldsymbol{M}_\mathrm{m}$, modal elastic stiffness matrix $\boldsymbol{K}_\mathrm{m,e}$, modal geometric stiffness matrix $\boldsymbol{K}_\mathrm{m,g}$ and modal damping matrix $\boldsymbol{D}_\mathrm{m}$

$$\boldsymbol{M}_\mathrm{m} = \boldsymbol{\Phi}^T\boldsymbol{M}\,\boldsymbol{\Phi} = \mathbf{I}, \boldsymbol{K}_\mathrm{m,e} = \boldsymbol{\Phi}^T\boldsymbol{K}_\mathrm{e}\,\boldsymbol{\Phi}, \boldsymbol{K}_\mathrm{m,g} = \boldsymbol{\Phi}^T\boldsymbol{K}_\mathrm{g}\,\boldsymbol{\Phi} \text{ and } \boldsymbol{D}_\mathrm{m} = \boldsymbol{\Phi}^T\boldsymbol{D}\,\boldsymbol{\Phi} \tag{11.12}$$

with identity matrix $\mathbf{I}$, all $[2\times2]$. Using the $[4\times1]$ modal state vector with the modal displacements and velocities $\boldsymbol{x}_\mathrm{m} = [\boldsymbol{q}, \dot{\boldsymbol{q}}]^T$, the modal state space equations of first order are

$$\begin{aligned} \dot{\boldsymbol{x}}_\mathrm{m} &= \underbrace{\left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\boldsymbol{K}_\mathrm{m,e} & -\boldsymbol{D}_\mathrm{m} \end{bmatrix} + F_x \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \boldsymbol{K}_\mathrm{m,g} & \mathbf{0} \end{bmatrix} \right)}_{[4\times4]} \boldsymbol{x}_\mathrm{m} + \underbrace{\begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Phi}^T\boldsymbol{B}_0 \end{bmatrix}}_{[4\times2]} \boldsymbol{u} \\ \boldsymbol{y} &= \underbrace{\begin{bmatrix} \boldsymbol{C}_0\,\boldsymbol{\Phi} & \mathbf{0} \end{bmatrix}}_{[2\times4]} \boldsymbol{x}_\mathrm{m}, \end{aligned} \tag{11.13}$$

[13]. Equation (11.13) can also be written in short form

$$\begin{aligned} \dot{\boldsymbol{x}}_{\mathrm{m}} &= \boldsymbol{A}_{\mathrm{m}}(F_x)\,\boldsymbol{x}_{\mathrm{m}} + \boldsymbol{B}_{\mathrm{m}}\,\boldsymbol{u} \\ \boldsymbol{y} &= \boldsymbol{C}_{\mathrm{m}}\,\boldsymbol{x}_{\mathrm{m}}. \end{aligned} \tag{11.14}$$

Due to imperfections such as predeformation, eccentric loading or clamping moments that are present in a real beam-column system, the controller needs to have an additional integral term to avoid a static controller error. Therefore, the modal state vector $\boldsymbol{x}_{\mathrm{m}}$ is augmented by the integral of the modal displacements

$$\boldsymbol{x}_{\mathrm{int}} = \int_{t_{\mathrm{LPV}}}^{\infty} \boldsymbol{q}\,\mathrm{d}t \tag{11.15}$$

starting from t_{LPV} to get the new $[6 \times 1]$ state vector $\boldsymbol{x} = [\boldsymbol{x}_{\mathrm{m}}, \boldsymbol{x}_{\mathrm{int}}]^T$. With the first derivative of the new state vector

$$\dot{\boldsymbol{x}}_{\mathrm{int}} = \boldsymbol{q} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \boldsymbol{x}_{\mathrm{m}}, \tag{11.16}$$

the augmented state space system including the integral term can be written as

$$\begin{aligned} \dot{\boldsymbol{x}} &= \underbrace{\begin{bmatrix} \boldsymbol{A}_{\mathrm{m}}(F_x) & \mathbf{0} \\ \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} & \mathbf{0} \end{bmatrix}}_{[6\times6]} \boldsymbol{x} + \underbrace{\begin{bmatrix} \boldsymbol{B}_{\mathrm{m}} \\ \mathbf{0} \end{bmatrix}}_{[6\times2]} \boldsymbol{u} \\ \boldsymbol{y} &= \underbrace{\begin{bmatrix} \boldsymbol{C}_{\mathrm{m}} & \mathbf{0} \end{bmatrix}}_{[2\times6]} \boldsymbol{x}, \end{aligned} \tag{11.17}$$

[14]. In short form, the final state space system (11.17) of the beam-column system in Fig. 11.1 is written as

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{A}(F_x)\,\boldsymbol{x} + \boldsymbol{B}\,\boldsymbol{u} \\ \boldsymbol{y} &= \boldsymbol{C}\,\boldsymbol{x}. \end{aligned} \tag{11.18}$$

11.3.2 Quadratically Stable Gain-Scheduled LPV Control

The FE state space model (11.6) as well as the final controller state space model (11.18) are LPV systems in which the system matrices $\boldsymbol{A}_{\mathrm{FE}}(F_x)$ and $\boldsymbol{A}(F_x)$ depend on axial load F_x. There are different control approaches to deal with the parameter-dependency of LPV systems. One approach is to use robust control in which a single controller is used for all occurring axial load variations with respect to amplitude and time-dependency. Due to the large variation of axial loads and the transition from sub- to supercritical axial loads, robust control is not favorable and has not been used so far. The approach pursued in earlier own studies [7–10] calculated static control matrices for a number of different axial loads that were manually switched, resulting in discontinuities in control input (11.3).

Now, in this investigation, active buckling control of the circular beam-column is achieved by a quadratically stable gain-scheduled LPV control. In this approach, a continuous control input according to (11.3) with

$$\boldsymbol{u} = \boldsymbol{K}_{\mathrm{LPV}}(F_x)\,\boldsymbol{x}, \tag{11.19}$$

is achieved by $[2 \times 6]$ control matrix $\boldsymbol{K}_{\mathrm{LPV}}(F_x)$ as a linear function of the axial load F_x with

$$\boldsymbol{K}_{\mathrm{LPV}}(F_x) = \frac{F_{x,2} - F_x}{F_{x,2} - F_{x,1}}\,\boldsymbol{K}_1 + \frac{F_x - F_{x,1}}{F_{x,2} - F_{x,1}}\,\boldsymbol{K}_2. \tag{11.20}$$

In (11.20) $F_{x,1} = 300\,\mathrm{N}$ and $F_{x,2} = 4000\,\mathrm{N}$ are the minimum and maximum considered axial loads for the controller that define the vertices of the LPV system $\boldsymbol{A}_1 = \boldsymbol{A}(F_{x,1})$ and $\boldsymbol{A}_2 = \boldsymbol{A}(F_{x,2})$ and $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ are the control matrices calculated for the respective systems. The resulting controller stabilizes the beam-column for the entire range of considered axial loads [15].

In order to weight the different states and control inputs, an additional $[8 \times 1]$ performance variable

$$\boldsymbol{z} = \boldsymbol{C}_z(F_x)\,\boldsymbol{x} + \boldsymbol{D}_z\,\boldsymbol{u}, \tag{11.21}$$

is introduced. The $[8 \times 6]$ matrix $\boldsymbol{C}_z(F_x)$ includes weights on system state vector $\boldsymbol{x}$ and the $[8 \times 2]$ matrix $\boldsymbol{D}_z$ includes weights on control input $\boldsymbol{u}$. For the experimental active buckling control, the performance weight matrices in (11.21) are chosen as

$$\boldsymbol{C}_{z,1} = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 & 0 & 10 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{C}_{z,2} = \begin{bmatrix} 50 & 0 & 0 & 0 & 0 & 0 \\ 0 & 50 & 0 & 0 & 0 & 0 \\ 0 & 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \boldsymbol{D}_z = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 30 & 0 \\ 0 & 30 \end{bmatrix}. \tag{11.22}$$

Equations (11.18) and (11.21) define the LPV system for the controller. Consequently, the axial load dependent system matrices and state performance matrices are $\boldsymbol{A}_i = \boldsymbol{A}(F_{x,i})$ and $\boldsymbol{C}_{z,i} = \boldsymbol{C}_z(F_{x,i})$ for $i = 1, 2$.

The control law leads to the control input $\boldsymbol{u}$ (11.3) and (11.19) so that the expected value of the quadratic performance index

$$J = \lim_{t \to \infty} \mathbb{E}\left[z^T z\right] < \mu \tag{11.23}$$

is minimized [15]. The control for matrices $\boldsymbol{K}_i = \boldsymbol{K}(F_{x,i})$ for $i = 1, 2$ in (11.20) are calculated by

$$\boldsymbol{K}_i = \boldsymbol{L}_i\,\boldsymbol{X}^{-1} \tag{11.24}$$

In (11.24), $\boldsymbol{L}_i$ and $\boldsymbol{X}$ are the solution of the linear matrix inequalities (LMI)

$$\begin{aligned} \boldsymbol{A}_i\,\boldsymbol{X} + \boldsymbol{X}\,\boldsymbol{A}_i^T + \boldsymbol{B}\,\boldsymbol{L}_i + \boldsymbol{L}_i^T\,\boldsymbol{B}^T &\prec 0 \\ \begin{bmatrix} \boldsymbol{Z}_i & \boldsymbol{C}_{z,i}\,\boldsymbol{X} + \boldsymbol{D}_z\,\boldsymbol{L}_i \\ \boldsymbol{X}\,\boldsymbol{C}_{z,i}^T + \boldsymbol{L}_i^T\,\boldsymbol{D}_z^T & \boldsymbol{X} \end{bmatrix} &\succ 0 \end{aligned} \tag{11.25}$$

in which the left hand sides of 11.25 have to be negative and positive definite, respectively. The LMI's are solved by *YALMIP* [16], for all $i = 1, 2$ to minimize

$$\text{trace}\,(\boldsymbol{Z}_i) < \mu. \tag{11.26}$$

The resulting control input (11.19) provides the minimum performance variable (11.21) for the chosen weights (11.22).

11.4 Experimental Results for Active Buckling Control

Experimental studies to validate the LPV control (11.19) are conducted on the experimental test setup for active buckling control, Fig. 11.3. The axial load F_x is measured by load cell 11 in Fig. 11.3b as input for the control matrix $\boldsymbol{K}_{\text{LPV}}(F_x)$ in (11.20). For that, the surface strains according to (11.2) are measured with strain gauge sensors on the beam-column, Fig. 11.1, that are transformed into modal displacements of the beam-column in y- and z-direction. The amplified signal which is filtered by analogue elliptic lowpass anti-aliasing filters is fed into a real-time *dSPACE 1103* system. The deformation of the beam-column in y- and z-direction is approximated by the modal displacements $\boldsymbol{q}$ that, due to the invertibility of the $[2 \times 2]$ matrix $(\boldsymbol{C}_0\,\boldsymbol{\Phi})$, result in

$$\boldsymbol{q} = (\boldsymbol{C}_0\,\boldsymbol{\Phi})^{-1}\,\boldsymbol{y}. \tag{11.27}$$

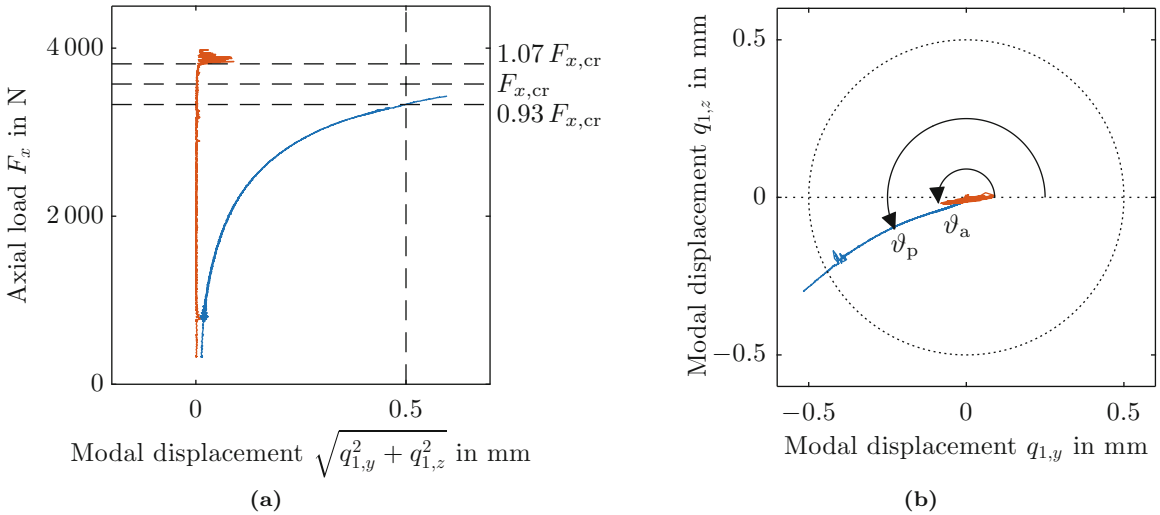


Fig. 11.5 Modal displacements $q_{1,y}$ and $q_{1,z}$ of the first modes in y- and z-direction for the beam-column with (*red solid line*) and without (*blue solid line*) active buckling control, (**a**) absolute modal displacement versus axial load F_x, (**b**) modal displacements in y–z-plane

A first order digital BUTTERWORTH highpass filter is used to approximate the modal velocities $\dot{\boldsymbol{q}}$ and a discrete integrator sums up the modal displacements $\boldsymbol{q}$ to get the integral state vector (11.15). The integration (11.15) starts with the beginning of active buckling control at t_{LPV} and continues indefinitely. The state vector $\boldsymbol{x}$ is used to calculate the control forces $\boldsymbol{u}$ via (11.19). The resulting *dSPACE* output voltages are filtered by analogue elliptic lowpass reconstruction filters before being amplified and applied to the piezoelectric stack actuators in piezo-elastic supports A and B.

The results of the beam-column system for ramp-shaped axial loads F_x with and without active buckling control are shown in Fig. 11.5. The absolute modal displacements $q_{1,y}$ and $q_{1,z}$ of the first modes in y- and z-direction are plotted with respect to the axial load F_x in Fig. 11.5a. Figure 11.5b shows the modal displacements $q_{1,y}$ and $q_{1,z}$ in the y–z-plane.

Without active buckling control, the load-displacement curve shows the typical continuous deformation with increasing axial load which is known from real, imperfect beam-columns [1]. Sudden buckling does not occur. In theory, the ideal beam-column with circular cross-section has infinite directions of buckling. The real beam-column, however, continuously deforms in an angle of $\vartheta_\mathrm{p} = 202°$ measured from the positive y-direction. The direction of buckling of the real uncontrolled system is mainly dependent on the predeflection of the beam-column and the assembly conditions of the test setup and varies accordingly. It is not possible to detect one single critical buckling load of the uncontrolled system, only a maximum admissible deformation can be set depending on the intended application. Here, a maximum admissible deformation of 0.5 mm is defined to avoid plastic deformation of the beam-column. It is reached for an axial load of $F_{x,0.5} = 3329\,\mathrm{N} = 0.93\,F_{x,\mathrm{cr}}$ which is less than the theoretical critical buckling load $F_{x,\mathrm{cr}} = 3573\,\mathrm{N}$ determined from the FE model via (11.9).

With active buckling control, the piezoelectric stack actuators in the piezo-elastic supports are able to initially force the beam-column into a straight position and then to reduce the beam-column deformation for the ramp-shaped axial load F_x. The LPV control is able to stabilize the beam-column up to an axial load of $F_{x,\max} = 3816\,\mathrm{N} = 1.07\,F_{x,\mathrm{cr}} = 1.14\,F_{x,0.5}$. Up to $F_{x,\max}$, the deformation in both y- and z-direction remains zero, Fig. 11.5a. At $F_{x,\max}$, the beam-column starts to oscillate and finally buckles at an angle of $\vartheta_\mathrm{a} = 194°$ in approximately the same direction as the passive beam-column, Fig. 11.5b.

11.5 Conclusion

Active Linear Parametric Varying (LPV) buckling control of an axially loaded beam-column with circular cross-section embedded in piezo-elastic supports is investigated experimentally. With the piezo-elastic supports, lateral forces of piezoelectric stack actuators are transformed into bending moments acting in arbitrary directions at the beam-column's ends. A modal model of the beam-column system based on a finite element (FE) model is validated by an experimental system identification. The modal model is augmented by integrated states and an LPV control is implemented. The experiments show that the axially loaded beam-column with circular cross-section without active buckling control exhibits high lateral deformations for axial loads considerably below the theoretical critical buckling load. With active buckling control using the presented piezo-elastic supports, stabilization of the beam-column in arbitrary direction is possible. The LPV control

is able to stabilize the beam-column for axial loads 14% higher than the passive beam-column and above the theoretical critical buckling load. In future investigations, active buckling control will be applied for step-shaped axial compressive loads where active buckling control is assumed to be very beneficial. Furthermore, system identification for higher axial loads may improve the controller model to avoid instability at higher axial loads.

Acknowledgements The authors like to thank the German Research Foundation (DFG) for funding this project within the Collaborative Research Center (SFB) 805.

References

1. Timoshenko, S.P., Gere, J.M.: Theory of Elastic Stability. McGraw-Hill, New York (1961)
2. Meressi, T., Paden, B.: Buckling control of a flexible beam using piezoelectric actuators. J. Guid. Control Dyn. **16**(5), 977–980 (1993)
3. Wang, Q.S.: Active buckling control of beams using piezoelectric actuators and strain gauge sensors. Smart Mater. Struct. **19**, 1–8 (2010)
4. Thompson, S.P., Loughlan, J.: The active buckling control of some composite column strips using piezoceramic actuators. Compos. Struct. **32**, 59–67 (1995)
5. Berlin, A.A., Chase, J.G., Yim, M., Maclean, J.B., Olivier, M., Jacobsen, S.C.: Mems-based control of structural dynamic instability. J. Intell. Mater. Syst. Struct. **9**, 574–586 (1998)
6. Zenz, G., Humer, A.: Stability enhancement of beam-type structures by piezoelectric transducers: theoretical, numerical and experimental investigations. Acta Mech. **226**(12), 3961–3976 (2015)
7. Enss, G.C., Platz, R., Hanselka, H.: Uncertainty in loading and control of an active column critical to buckling. Shock Vib. **19**, 929–937 (2012)
8. Enss, G.C., Platz, R.: Evaluation of uncertainty in experimental active buckling control of a slender beam-column using Weibull analysis. Mech. Syst. Signal Process. **79**, 123–131 (2016)
9. Schaeffner, M., Götz, B., Platz, R.: Active buckling control of a beam-column with circular cross-section using piezoelastic supports and integral LQR control. Smart Mater. Struct. **25**(6), 065008 (2016)
10. Schaeffner, M., Platz, R.: Active buckling control of an imperfect beam-column with circular cross-section using piezoelastic supports and integral LQR control. J. Phys. Conf. Ser. **744**, 012165 (2016)
11. Enss, G.C., Gehb, C.M., Goetz, B., Melz, T., Ondoua, S., Platz, R., Schaeffner, M.: Device for bearing design elements in lightweight structures (Festkörperlager) Patent. DE 10 2015 101 084 A1, July 2016
12. Przemieniecki, J.S.: Theory of Matrix Structural Analysis. McGraw-Hill, New York (1968)
13. Gawronski, W.: Advanced Structural Dynamic and Active Control of Structures. Springer, New York (2010)
14. Hendricks, E., Jannerup, O., Sørensen, P.H.: Linear Systems Control. Springer, Heidelberg (2008)
15. Pandey, A., Sehr, M., de Oliveira, M.: Pre-filtering in gain-scheduled and robust control. In: Proceedings of 2016 American Control Conference (ACC), Boston, MA, 6–8 July 2016
16. Löfberg, J.: Yalmip: a toolbox for modeling and optimization in MATLAB. In: Proceedings of the CACSD Conference, Taipei (2004)

Chapter 12
Quantification and Evaluation of Uncertainty in the Mathematical Modelling of a Suspension Strut Using Bayesian Model Validation Approach

Shashidhar Mallapur and Roland Platz

Abstract Mathematical models of a suspension strut such as an aircraft landing gear are utilized by engineers in order to predict its dynamic response under different boundary conditions. The prediction of the dynamic response, for example the external loads, the stress and the strength as well as the maximum compression in the spring-damper component aids engineers in early decision making to ensure its structural reliability under various operational conditions. However, the prediction of the dynamic response is influenced by model uncertainty. As far as the model uncertainty is concerned, the prediction of the dynamic behavior via different mathematical models depends upon various factors such as the model's complexity in terms of the degrees of freedom, material and geometrical assumptions, their boundary conditions and the governing functional relations between the model input and output parameters. The latter can be linear or nonlinear, axiomatic or empiric, time variant or time-invariant. Hence, the uncertainty that arises in the prediction of the dynamic response of the resulting different mathematical models needs to be quantified with suitable validation metrics, especially when the system is under structural risk and failure assessment. In this contribution, the authors utilize the BAYES factor as a validation metric to quantify the model uncertainty of a suspension strut system with similar specifications as actual suspension struts in automotive or aerospace applications for decision making in early design stage.

Keywords Bayes factor • Design concepts • Mathematical model • Model validation • Suspension strut • Uncertainty

12.1 Introduction

Since the mid-60s, mathematical simulation models have been used as a tool in the field of scientific research and design of engineering systems [1, 2]. In the field of structural dynamics in engineering, mathematical models are used to predict the dynamic response of the structural system. The predicted responses are used by engineers to evaluate the system's stability, strength and vibrational behavior under various operational conditions. However, different mathematical models based on different underlying assumptions may predict different dynamic responses. Hence, there arises a need to quantify the uncertainty in the predictive capability of the different forms of mathematical models. Under the framework of verification and validation, various validation metrics have been proposed for model validation under aleatory uncertainty [3, 4]. A validation metric gives a quantitative measure of the extent of agreement between the predicted model output and the experimentally observed output [5]. The metrics are used for model selection when alternative mathematical models are being considered. Under probabilistic methods, validation metrics such as BAYES factor which is based on BAYESIAN probabilistic approach can be utilized to judge the capability of prediction of a given mathematical model with the experimental data [6]. Moreover, a validation metric can also be developed by the normalized difference between the model prediction and the experimental observations [7]. Data uncertainty occurs when the input parameters, such as state variables or the system parameters of the mathematical models are uncertain. This can be represented using probabilistic and non-probabilistic methods [8]. Conversely, model uncertainty occurs due to uncertainty about which governing relations of data can predict the dynamic behavior of a structural system adequately within acceptable confidence limits.

S. Mallapur (✉)
System Reliability, Adaptive Structures and Machine Acoustics SAM, Technische Universität Darmstadt, Magdalenenstraße 4, 64289 Darmstadt, Germany
e-mail: mallapur@sam.tu-darmstadt.de

R. Platz
Fraunhofer Institute for Structural Durability and System Reliability LBF, Bartningstraße 47, 64289 Darmstadt, Germany

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_12

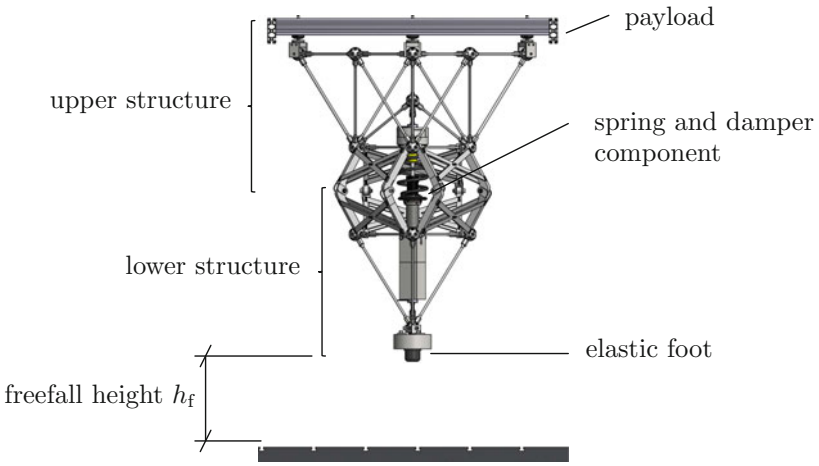


Fig. 12.1 CAD illustration of a suspension strut (MAFDS)

In this contribution, the authors consider the widely recognized BAYES factor as a model validation metric to quantify the uncertainty in the mathematical modelling of a suspension strut. However, and as a new approach, the authors present a way to adequately compare different models based on linear, nonlinear, axiomatic or empiric assumptions of functional relations. This work considers a suspension strut which is referred as the modular active spring-damper system (German acronym MAFDS). MAFDS is developed under the framework of German Collaborative Research Center SFB 805 at the Technische Universität Darmstadt. The concept of MAFDS is registered as a Patent DE 10 2014 106 858.A1 "Load transferring device" [9].

12.2 Suspension Strut MAFDS

As illustrated in Fig. 12.1, the MAFDS serves as an academic demonstrator to investigate the different methods and technologies to control uncertainty in all development stages such as the conceptual phase, design and optimization phase, production and assembly phase as well as the final operational phase. Uncertainty is investigated especially for MAFDS's main operating function and purpose: sustain stability, well balanced load distribution and ability to attenuate vibrations. Passive, semi-active and active technology approaches for stability, load distribution and vibration control to compensate uncertainty in the main functions mentioned above are integrated in a modular way within SFB 805 [10–13]. MAFDS has similar dynamic requirements of a typical suspension strut such as an aircraft landing gear. However it is not designed to be used as a substitute to an existing aircraft landing gear. Rather, the MAFDS is an academic example to study the uncertainty in the dynamic responses of suspension strut systems in a general way as explained above. One of the key dynamic responses of the MAFDS under uncertainty is the maximum relative compression of the spring damper component under different loading conditions. For this, drop tests with varying payload are conducted numerically via different mathematical models subject to experimental validation via test rig. For details about the test rig, refer Sect. 12.4. The maximum relative compression that the spring-damper component can undergo under different boundary and load conditions is limited to 0.1 m due to the design constraints of the spring-damper component.

12.3 Simplified 2DOF Mathematical Models of MAFDS

12.3.1 Motivation

As illustrated in Fig. 12.2, a simple two degree of freedom (2DOF) mathematical model representing MAFDS is utilized to predict the maximum relative compression $z_{r,max}$. The upper and lower structures of the MAFDS are simplified as two lumped masses m_u and m_l respectively. The spring-damper component may be represented by a stiffness parameter k and a suitable

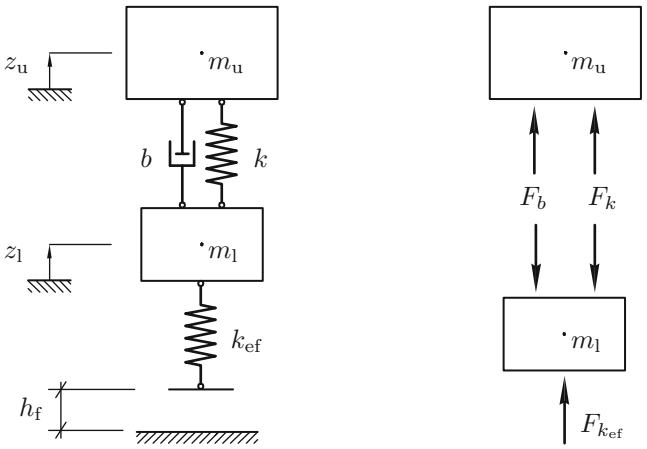


Fig. 12.2 Illustration of simplified 2DOF mathematical model of MAFDS

damping coefficient b. Similarly, the elastic foot may be represented by a stiffness parameter k_{ef}. The 2DOF mathematical model has two degrees of freedom that are defined by the local translational displacements z_u and z_l. The maximum relative compression

$$z_{r,\,max} = \max\,|z_u - z_l| \tag{12.1}$$

is defined as the maximum value of the difference between the local displacement z_u of the lumped mass m_u of the upper structure and the local displacement z_l of the lumped mass m_l of the lower structure. The mathematical model is excited by the initial conditions at time $t = 0$ that are valid for m_u and m_l. Thus the initial displacements are

$$z_u(0) = z_l(0) = 0, \tag{12.2}$$

and the initial velocities are

$$\dot{z}_u(0) = \dot{z}_l(0) = \sqrt{2gh_f}, \tag{12.3}$$

where h_f is the given freefall height. For the sake of convenience, the time variable (t) is neglected from the variables z and $\dot{z}$ in further explanations.

12.3.2 Five Different Approaches to Model the Stiffness and the Damping

The 2DOF mathematical model can be described with different underlying assumptions, for example linear or nonlinear, axiomatic or empiric assumptions. The authors present five different approaches to model the stiffness and the damping of the 2DOF mathematical model for the prediction of the maximum compression $z_{r,\,max}$ of the MAFDS. It is to be noted that the degrees of freedom for all the five approaches remain the same, however, the stiffness and the damping are modelled according to different underlying assumptions such as linear or nonlinear, axiomatic or empiric assumptions.

(a) Model with linear stiffness k and axiomatic damping b

From Fig. 12.2 the free body diagram forces are described as

$$F_b = b\,[\dot{z}_u - \dot{z}_l],\quad F_k = k[z_u - z_l],\quad F_{kef} = k_{ef}\,z_l \tag{12.4}$$

where F_b is the linear axiomatic viscous damping force, F_k is the linear axiomatic spring force and F_{kef} is the force exerted via the elastic foot. The governing differential equations of the 2DOF mathematical model for m_u is

$$m_u\,\ddot{z}_u \;+\; b\,[\dot{z}_u - \dot{z}_l] \;+\; k[z_u - z_l] = 0, \tag{12.5}$$

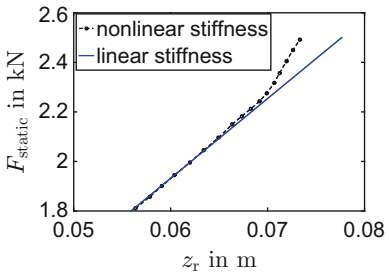


Fig. 12.3 Stiffness curve

and for m_{l} is

$$m_{\text{l}}\,\ddot{z}_{\text{l}} \;-\; b\,[\dot{z}_{\text{u}} - \dot{z}_{\text{l}}] \;-\; k[z_{\text{u}} - z_{\text{l}}] + k_{\text{ef}}\,z_{\text{l}} = 0. \tag{12.6}$$

with stiffness k and damping b. The matrix form representation of the governing differential equation is expressed as

$$\underbrace{\begin{bmatrix} m_{\text{u}} & 0 \\ 0 & m_{\text{l}} \end{bmatrix}}_{\boldsymbol{M}} \begin{pmatrix} \ddot{z}_{\text{u}} \\ \ddot{z}_{\text{l}} \end{pmatrix} + \underbrace{\begin{bmatrix} b & -b \\ -b & b \end{bmatrix}}_{\boldsymbol{B}} \begin{pmatrix} \dot{z}_{\text{u}} \\ \dot{z}_{\text{l}} \end{pmatrix} + \underbrace{\begin{bmatrix} k & -k \\ -k & k + k_{\text{ef}} \end{bmatrix}}_{\boldsymbol{K}} \begin{pmatrix} z_{\text{u}} \\ z_{\text{l}} \end{pmatrix} = 0, \tag{12.7}$$

where M, B and K are the global mass, damping and stiffness matrices.

(b) Model with nonlinear stiffness $k(z_{\text{r}})$ and axiomatic damping b

In this mathematical model, the stiffness of the spring $k(z_{\text{r}})$ is non-constant and is modelled as a function that is dependent upon the value of the relative displacement $z_{\text{r}} = (z_{\text{u}} - z_{\text{l}})$. As illustrated from Fig. 12.3, the relative displacement z_{r} increases nonlinearly for $z_{\text{r}} > 0.065$ m. The spring force is modelled as

$$F_{k(z_{\text{r}})} = k(z_{\text{r}})[z_{\text{u}} - z_{\text{l}}]. \tag{12.8}$$

The nonlinear stiffness curve in Fig. 12.3 is obtained from the experimental static relative displacement under varying static loads that is conducted via the MAFDS test setup.

The matrix form representation of the governing differential equation is expressed as

$$\underbrace{\begin{bmatrix} m_{\text{u}} & 0 \\ 0 & m_{\text{l}} \end{bmatrix}}_{\boldsymbol{M}} \begin{pmatrix} \ddot{z}_{\text{u}} \\ \ddot{z}_{\text{l}} \end{pmatrix} + \underbrace{\begin{bmatrix} b & -b \\ -b & b \end{bmatrix}}_{\boldsymbol{B}} \begin{pmatrix} \dot{z}_{\text{u}} \\ \dot{z}_{\text{l}} \end{pmatrix} + \underbrace{\begin{bmatrix} k(z_{\text{r}}) & -k(z_{\text{r}}) \\ -k(z_{\text{r}}) & k(z_{\text{r}}) + k_{\text{ef}} \end{bmatrix}}_{\boldsymbol{K(z_{\text{r}})}} \begin{pmatrix} z_{\text{u}} \\ z_{\text{l}} \end{pmatrix} = 0. \tag{12.9}$$

(c) Model with linear stiffness k and damping proportional to mass and stiffness ($\alpha\boldsymbol{M} + \beta\boldsymbol{K}$)

In cases where the knowledge about the axiomatic damping coefficient b is limited or unknown, RAYLEIGH damping is utilized to estimate the equivalent damping under the framework of linear modal analysis [14]. Under this framework, the damping Matrix B is proportionally equivalent to the mass matrix and the stiffness matrix with the help of RAYLEIGH damping coefficients α and β.

$$\boldsymbol{B} = \alpha\boldsymbol{M} \;+\; \beta\boldsymbol{K}. \tag{12.10}$$

Upon substituting (12.10) in (12.7), the matrix form representation of the governing differential equation is expressed as

$$\underbrace{\begin{bmatrix} m_{\text{u}} & 0 \\ 0 & m_{\text{l}} \end{bmatrix}}_{\boldsymbol{M}} \begin{pmatrix} \ddot{z}_{\text{u}} \\ \ddot{z}_{\text{l}} \end{pmatrix} + \underbrace{\begin{bmatrix} \alpha\, m_{\text{u}} + \beta\, k & -\beta\, k \\ -\beta\, k & \alpha\, m_{\text{l}} + \beta\,(k + k_{\text{ef}}) \end{bmatrix}}_{\boldsymbol{B = f(M,K)}} \begin{pmatrix} \dot{z}_{\text{u}} \\ \dot{z}_{\text{l}} \end{pmatrix} + \underbrace{\begin{bmatrix} k & -k \\ -k & k + k_{\text{ef}} \end{bmatrix}}_{\boldsymbol{K}} \begin{pmatrix} z_{\text{u}} \\ z_{\text{l}} \end{pmatrix} = 0. \tag{12.11}$$

The RAYLEIGH parameters α and β are estimated from the least squares optimization algorithm to get the best fit between the mathematical model's output and the experimental results of the relative compression velocity [15]. The model parameters α and β are estimated such that the squares of the differences between the compression velocity $\dot{z}_{\mathrm{r},n}$ calculated from the mathematical model (c) and the experimental data $\dot{z}_{\mathrm{r,\,E},n}$ is minimum for N number of time steps.

$$(\alpha, \beta) = \arg\min_{\alpha,\beta} \sum_{n=1}^{N} \left(\dot{z}_{\mathrm{r,\,E},n} - \dot{z}_{\mathrm{r},n}\right)^2 \tag{12.12}$$

The RAYLEIGH parameters $\alpha = 9.75$ and $\beta = 0.0068$ are estimated from the initial conditions in (12.2) and (12.3) for $h_{\mathrm{f}} = 0.02\,\mathrm{m}$.

(d) Model with nonlinear stiffness $k(z_{\mathrm{r}})$ and damping proportional to mass and stiffness ($\alpha\boldsymbol{M} + \beta\boldsymbol{K}$)

The governing relations are a combination of (12.9) and (12.13) with

$$\underbrace{\begin{bmatrix} m_{\mathrm{u}} & 0 \\ 0 & m_{\mathrm{l}} \end{bmatrix}}_{\boldsymbol{M}} \begin{pmatrix} \ddot{z}_{\mathrm{u}} \\ \ddot{z}_{\mathrm{l}} \end{pmatrix} + \underbrace{\begin{bmatrix} \alpha\, m_{\mathrm{u}} + \beta\, k(z_{\mathrm{r}}) & -\beta\, k(z_{\mathrm{r}}) \\ -\beta\, k(z_{\mathrm{r}}) & \alpha\, m_{\mathrm{l}} + \beta\,(k(z_{\mathrm{r}}) + k_{\mathrm{ef}}) \end{bmatrix}}_{\boldsymbol{B} = \boldsymbol{f}(\boldsymbol{M},\boldsymbol{K})} \begin{pmatrix} \dot{z}_{\mathrm{u}} \\ \dot{z}_{\mathrm{l}} \end{pmatrix} + \underbrace{\begin{bmatrix} k(z_{\mathrm{r}}) & -k(z_{\mathrm{r}}) \\ -k(z_{\mathrm{r}}) & k(z_{\mathrm{r}}) + k_{\mathrm{ef}} \end{bmatrix}}_{\boldsymbol{K}(\boldsymbol{z}_{\mathbf{r}})} \begin{pmatrix} z_{\mathrm{u}} \\ z_{\mathrm{l}} \end{pmatrix} = 0. \tag{12.13}$$

(e) Model with nonlinear stiffness $k(z_{\mathrm{r}})$ and velocity dependent damping $b \propto F_b(\dot{z}_{\mathrm{r}}(t))$

The damping behavior in the mathematical models (a) to (d) is based on linear assumptions. However, the damping force $F_b(\dot{z}_{\mathrm{r}})$ that is obtained from the experimental test of MAFDS depends nonlinearly on the relative velocity $\dot{z}_{\mathrm{r}}$.

$$\underbrace{\begin{bmatrix} m_{\mathrm{u}} & 0 \\ 0 & m_{\mathrm{l}} \end{bmatrix}}_{\boldsymbol{M}} \begin{pmatrix} \ddot{z}_{\mathrm{u}} \\ \ddot{z}_{\mathrm{l}} \end{pmatrix} + \underbrace{\begin{bmatrix} F_b(\dot{z}_{\mathrm{r}}) \\ -F_b(\dot{z}_{\mathrm{r}}) \end{bmatrix}}_{\boldsymbol{F_b}(\dot{\boldsymbol{z}}_{\mathbf{r}})} + \underbrace{\begin{bmatrix} k(z_{\mathrm{r}}) & -k(z_{\mathrm{r}}) \\ -k(z_{\mathrm{r}}) & k(z_{\mathrm{r}}) + k_{\mathrm{ef}} \end{bmatrix}}_{\boldsymbol{K}(\boldsymbol{z}_{\mathbf{r}})} \begin{pmatrix} z_{\mathrm{u}} \\ z_{\mathrm{l}} \end{pmatrix} = 0 \tag{12.14}$$

The relative velocity $\dot{z}_{\mathrm{r}}$ is shown in Fig. 12.4b. It is calculated from the measured relative displacement $z_{\mathrm{r,E}}$ via numerical differentiation over time, refer Fig. 12.4a. The damping force $F_b(\dot{z}_{\mathrm{r}})$ in (12.14) is obtained via experimental drop tests of MAFDS, refer Fig. 12.4c.

12.3.3 Solution of the Relative Compression Response of the 2DOF Mathematical Models (a) to (e)

The solution in time domain for (12.7), (12.9), (12.11), (12.13) and (12.14) are given in such a way that they represent the relative compression $z_{\mathrm{r,max}}$ in a similar manner for linear, nonlinear, axiomatic and empiric assumptions for stiffness and damping properties. The solutions are obtained by numerical integration via RUNGE KUTTA algorithm. From (12.5)

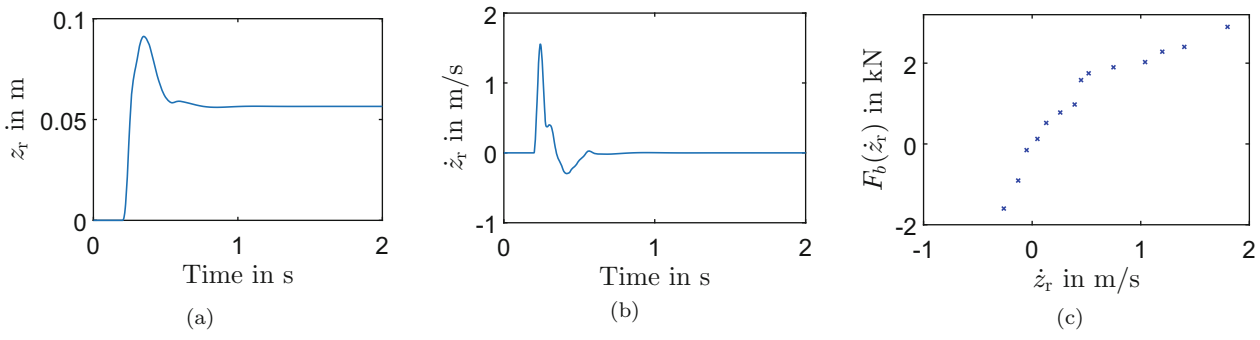


Fig. 12.4 (**a**) Relative displacement, (**b**) relative velocity and (**c**) velocity dependent damping force

and (12.6), the accelerations $\ddot{z}_\mathrm{u}$ and $\ddot{z}_\mathrm{l}$ are respectively described as

$$\ddot{z}_\mathrm{u} = \frac{1}{m_\mathrm{u}}\Big[-b\,[\dot{z}_\mathrm{u} - \dot{z}_\mathrm{l}] \;-\; k[z_\mathrm{u} - z_\mathrm{l}]\Big], \tag{12.15}$$

and

$$\ddot{z}_\mathrm{l} = \frac{1}{m_\mathrm{l}}\Big[b\,[\dot{z}_\mathrm{u} - \dot{z}_\mathrm{l}] \;+\; k[z_\mathrm{u} - z_\mathrm{l}] - k_\mathrm{ef}\,z_\mathrm{l}\Big], \tag{12.16}$$

which leads to the state-space form of (12.15) and (12.16), represented as

$$\underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\boldsymbol{x}} = \begin{bmatrix} z_\mathrm{u} \\ \dot{z}_\mathrm{u} \\ z_\mathrm{l} \\ \dot{z}_\mathrm{l} \end{bmatrix} \quad ; \quad \underbrace{\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix}}_{\dot{\boldsymbol{x}}} = \begin{bmatrix} \dot{z}_\mathrm{u} \\ \ddot{z}_\mathrm{u} \\ \dot{z}_\mathrm{l} \\ \ddot{z}_\mathrm{l} \end{bmatrix} = \begin{bmatrix} x_2 \\ \dfrac{1}{m_\mathrm{u}}\Big[-b\,[x_2 - x_4] \;-\; k[x_1 - x_3]\Big] \\ x_4 \\ \dfrac{1}{m_\mathrm{l}}\Big[b\,[x_2 - x_4] \;+\; k[x_1 - x_3] - k_\mathrm{ef}\,x_3\Big] \end{bmatrix}. \tag{12.17}$$

Similarly, the general time domain solution to the governing relations for linear, nonlinear, axiomatic and empiric assumptions in (12.7), (12.9), (12.11), (12.13) and (12.14) can be expressed in a generalized nonlinear state space form by

$$\dot{\boldsymbol{x}} = \begin{bmatrix} x_2 \\ \dfrac{1}{m_\mathrm{u}} f_\mathrm{u}(x_1, x_2, x_3, x_4) \\ x_4 \\ \dfrac{1}{m_\mathrm{l}} f_\mathrm{l}(x_1, x_2, x_3, x_4) \end{bmatrix}. \tag{12.18}$$

The summation of internal forces acting on m_u and m_l in (12.18) are described as internal force functions $f_\mathrm{u}(x_1, x_2, x_3, x_4)$ and $f_\mathrm{l}(x_1, x_2, x_3, x_4)$ respectively. The functions $f_\mathrm{u}(x_1, x_2, x_3, x_4)$ and $f_\mathrm{l}(x_1, x_2, x_3, x_4)$ for the mathematical models (a) to (e) are listed in Table 12.1. The state vector $\dot{\boldsymbol{x}}$ is numerically integrated using a fourth order RUNGE KUTTA algorithm to obtain the desired state variables x_1 and x_3, [16]. The numerical computation is started with the initial conditions in (12.2) and (12.3) along with the numerical time integration step $t_\mathrm{s} = 0.001$ s.

$$\begin{bmatrix} z_\mathrm{u} \\ z_\mathrm{l} \end{bmatrix} = \begin{bmatrix} 1\,0\,0\,0 \\ 0\,0\,1\,0 \end{bmatrix} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\boldsymbol{x}} \tag{12.19}$$

Hence, from (12.19), the desired relative maximum compression $z_\mathrm{r,max}$ is calculated according to (12.1).

Table 12.2 summarizes the model parameters for the mathematical models (a) to (e).

12.4 Experimental Setup of MAFDS

This section gives a brief description about the experimental setup of the MAFDS. It's purpose is to measure the maximum relative compression $z_{\mathrm{r,\,max,\,E}}$ via drop test from a specified height h_f. Moreover, other response variables such as loads, stresses, accelerations can also be measured in the experimental setup. However, these are not considered within the scope of this contribution.

Figure 12.5 shows the MAFDS, fixed at three points to the load frame 1 by means of three bearing supports 2. The hoist frame 3 and the load frame are connected by an electro-pneumatic coupling 4. The hoist and the load frames are fixed to the guide blocks 5 that allow smooth parallel movement of the MAFDS along the guide rails 6. The hoist frame is driven by two

Table 12.1 Functions $f_u(x_1, x_2, x_3, x_4)$ and $f_l(x_1, x_2, x_3, x_4)$ for different mathematical models

Model type	Functions $f_u(x_1, x_2, x_3, x_4)$ and $f_l(x_1, x_2, x_3, x_4)$
Model (a)	$f_u(x_1, x_2, x_3, x_4) = -\boldsymbol{b}\,[x_2 - x_4] - \boldsymbol{k}\,[x_1 - x_3]$
	$f_l(x_1, x_2, x_3, x_4) = \boldsymbol{b}\,[x_2 - x_4] + \boldsymbol{k}\,[x_1 - x_3] - k_{ef} x_3$
Model (b)	$f_u(x_1, x_2, x_3, x_4) = -\boldsymbol{b}\,[x_2 - x_4] - \boldsymbol{k}(\boldsymbol{z_r})\,[x_1 - x_3]$
	$f_l(x_1, x_2, x_3, x_4) = \boldsymbol{b}\,[x_2 - x_4] + \boldsymbol{k}(\boldsymbol{z_r})\,[x_1 - x_3] - k_{ef} x_3$
Model (c)	$f_u(x_1, x_2, x_3, x_4) = -[\boldsymbol{\alpha}\, m_u + \boldsymbol{\beta}\, \boldsymbol{k}] x_2 + \boldsymbol{\beta}\, \boldsymbol{k}\, x_4 - \boldsymbol{k}\,[x_1 - x_3]$
	$f_l(x_1, x_2, x_3, x_4) = \boldsymbol{\beta}\, \boldsymbol{k}\, x_2 - [\boldsymbol{\alpha}\, m_l + \boldsymbol{\beta}\, (\boldsymbol{k} + k_{ef})] x_4 + \boldsymbol{k}\,[x_1 - x_3] - k_{ef} x_3$
Model (d)	$f_u(x_1, x_2, x_3, x_4) = -[\boldsymbol{\alpha}\, m_u + \boldsymbol{\beta}\, \boldsymbol{k}(\boldsymbol{z_r})] x_2 + \boldsymbol{\beta}\, \boldsymbol{k}(\boldsymbol{z_r})\, x_4 - \boldsymbol{k}(\boldsymbol{z_r})\,[x_1 - x_3]$
	$f_l(x_1, x_2, x_3, x_4) = \boldsymbol{\beta}\, \boldsymbol{k}(\boldsymbol{z_r})\, x_2 - [\boldsymbol{\alpha}\, m_l + \boldsymbol{\beta}\, (\boldsymbol{k}(\boldsymbol{z_r}) + k_{ef})] x_4 + \boldsymbol{k}(\boldsymbol{z_r})\,[x_1 - x_3] - k_{ef} x_3$
Model (e)	$f_u(x_1, x_2, x_3, x_4) = -\boldsymbol{F_b}(\dot{\boldsymbol{z}}_\mathbf{r}) - \boldsymbol{k}(\boldsymbol{z_r})\,[x_1 - x_3]$
	$f_l(x_1, x_2, x_3, x_4) = \boldsymbol{F_b}(\dot{\boldsymbol{z}}_\mathbf{r}) + \boldsymbol{k}(\boldsymbol{z_r})\,[x_1 - x_3] - k_{ef} x_3$

Table 12.2 Model parameters used for different mathematical models

Parameter	Symbol	Value	SI unit	(a)	(b)	(c)	(d)	(e)
Mass of upper structure	m_u	184	kg	✓	✓	✓	✓	✓
Mass of lower structure	m_l	41	kg	✓	✓	✓	✓	✓
Linear stiffness	k	33×10^3	N/m	✓	×	✓	×	×
Nonlinear stiffness	$k(z_r)$	Fig. 12.3	N/m	×	✓	×	✓	✓
Elastic foot stiffness	k_{ef}	224×10^3	N/m	✓	✓	✓	✓	✓
Viscous damping constant	b	1800	N s/m	✓	✓	×	×	×
Rayleigh alpha constant	α	9.75	–	×	×	✓	✓	×
Rayleigh beta constant	β	0.0068	–	×	×	✓	✓	×
Velocity dependent damping force	$F_b(\dot{z}_r)$	Fig. 12.4c	N	×	×	×	×	✓
Freefall height	h_f	[0.01,..., 0.07]	m	✓	✓	✓	✓	✓

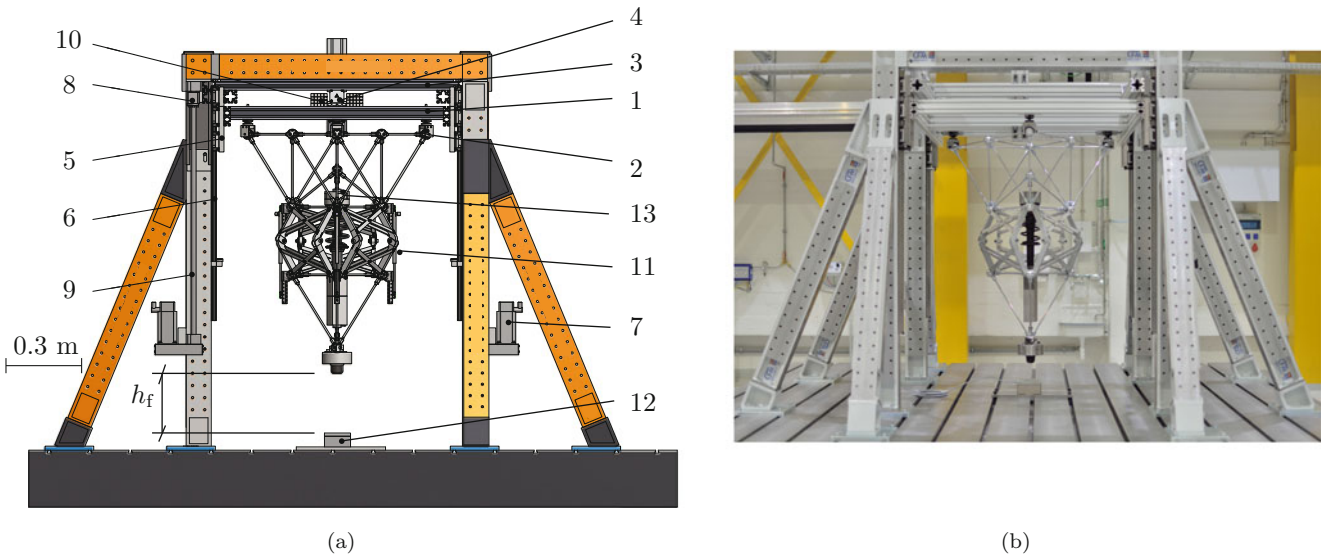


Fig. 12.5 Experimental test setup of MAFDS, (**a**) CAD illustration of test setup, (**b**) Photo of test setup

servomotors 7 that run synchronously. The ball screw 8 converts the rotational motion of the threaded spindle 9, which is powered by the servomotors into the translational parallel movement of the hoist frame. Once lifted to the desired height h_f, the MAFDS can be released by the pneumatic coupling for a freefall drop test. The maximum freefall height of MAFDS is 0.22 m. Additional payloads 10 can be fixed to the load frame, thereby adding static loads to the MAFDS.

The experimental relative compression $z_{r,\max,E}$ of the MAFDS is recorded by the displacement sensor 11 that is based on the linear displacement variable transformer (LVDT) principle. The total load exerted by the MAFDS on the ground is

recorded by the three-axial strain gauge force sensor 12. Similarly, the force exerted by the spring damper component of the MAFDS is recorded by the 1-axial strain gauge force sensor 13.

12.5 Approach to Evaluate Model Uncertainty

The five different mathematical models (a) to (e) introduced in Sect. 12.3 predict the maximum relative compression $z_{\mathrm{r,max}}$ differently according to the governing functions in Table 12.1. This leads to uncertainty regarding which mathematical model is suitable to predict $z_{\mathrm{r,max}}$ of MAFDS most adequately. For model uncertainty quantification, the authors validate the models with the experimental data via posterior probability by conducting the BAYESIAN probabilistic approach using BAYES factors [6, 17, 18]. The posterior probability is the probability that the predicted maximum relative compression event $H_{z_{\mathrm{r,max}}}$ is true under the condition of observing the experimental data, which is described as the measured maximum relative compression event $A_{z_{\mathrm{r,max}}}$.

The BAYES factor B is used as a validation metric to validate the prediction of the maximum relative compression event $H_{z_{\mathrm{r,max}}}$ by the different mathematical models (a) to (e) with the experimentally observed event $A_{z_{\mathrm{r,max}}}$. In general, $H_{z_{\mathrm{r,max}}}$ is the hypothesis event of maximum relative compression $z_{\mathrm{r,max}}$ that is computed using the mathematical models. Similarly, $A_{z_{\mathrm{r,max}}}$ is the aftereffect or the observation of maximum relative compression $z_{\mathrm{r,max}}$ that results from the experimental tests. The estimation of the BAYES factor B involves the calculation of the posterior and the prior probabilities of the hypothesis event $H_{z_{\mathrm{r,max}}}$. In the following, the prior, total and posterior probabilities as well as the likelihood used for the BAYESIAN approach to quantify uncertainty in the models (a) to (e) are briefly described.

12.5.1 Estimation of Posterior Probability of $z_{\mathrm{r,max}}$

The posterior probability

$$p(H_{z_{\mathrm{r,max}},n}|A_{z_{\mathrm{r,max}}}) = \frac{p(A_{z_{\mathrm{r,max}}}|H_{z_{\mathrm{r,max}},n})\, p(H_{z_{\mathrm{r,max}},n})}{p(A_{z_{\mathrm{r,max}}})}, \tag{12.20}$$

represents the probability of numerically computed hypothesis events $H_{z_{\mathrm{r,max}},n}$ under the condition that the observation event $A_{z_{\mathrm{r,max}}}$ is given, with the likelihood $p(A_{z_{\mathrm{r,max}}}|H_{z_{\mathrm{r,max}},n})$, the prior probability $p(H_{z_{\mathrm{r,max}},n})$ and the experimental data probability $p(A_{z_{\mathrm{r,max}}})$. The N hypothesis events $H_{z_{\mathrm{r,max}},n}$ are computed via mathematical models using (12.19) and (12.1) for $n = 1, 2, 3, \ldots, N$, where N is the total number of random numerical samples.

Prior Probability $p(H_{z_{\mathrm{r,max}},n})$

The prior probability

$$p(H_{z_{\mathrm{r,max}},n}) = \frac{1}{\sigma_H \sqrt{2\,\pi}}\, e^{-\frac{1}{2}\left(\frac{H_{z_{\mathrm{r,max}},n} - \mu_H}{\sigma_H}\right)^2}, \tag{12.21}$$

takes into account the random hypothesis events $H_{z_{\mathrm{r,max}},n}$ with normal density, with a mean μ_H and standard deviation σ_H. As mentioned earlier, N hypothesis events $H_{z_{\mathrm{r,max}},n}$ are computed via MONTE CARLO simulation of mathematical models using (12.19) and (12.1) based on random model parameter assumptions [19]. The random variation for N events of $H_{z_{\mathrm{r,max}},n}$ occurs due to the assumed random model parameters. As an assumed random varying model parameter, the mass of the upper structure m_{u} with an assumed mean $\mu_{m_{\mathrm{u}}} = 189\,\mathrm{kg}$ and an assumed standard deviation of $\sigma_{m_{\mathrm{u}}} = 1\%$ from the mean value is taken into account as a prior. The normal density masses $m_{\mathrm{u},n}$ are assumed and generated by a numerical random generator via MATLAB. The randomness in $m_{\mathrm{u},n}$ with a specified $\sigma_{m_{\mathrm{u}}}$ propagates through the mathematical models (a) to (e), resulting in random model hypothesis events $H_{z_{\mathrm{r,max}},n}$.

Likelihood $p(A_{z_{\mathrm{r,max}}}|H_{z_{\mathrm{r,max}},n})$

The likelihood

$$p(A_{z_{\mathrm{r,max}}}|H_{z_{\mathrm{r,max}},n}) = \prod_{k=1}^{K} \frac{1}{\sigma_H\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{A_{z_{\mathrm{r,max},k}} - H_{z_{\mathrm{r,max}},n}}{\sigma_H}\right)^2} \tag{12.22}$$

describes the probability of the observation events $A_{z_{\mathrm{r,max}}}$ that are obtained from experiments, when the hypothesis events $H_{z_{\mathrm{r,max}},n}$ predicted by the mathematical models are given. On the one hand, a set of $N = 1000$ hypothesis events $H_{z_{\mathrm{r,max}},n}$ with a mean μ_H and standard deviation σ_H are computed via MONTE CARLO simulation of the mathematical models. Conversely, a set of $K = 11$ experimental tests are conducted to obtain the observation events $A_{z_{\mathrm{r,max},k}}$ for $k = 1, 2, 3, \ldots, K$. The experiments are conducted with varying mass $m_{\mathrm{u},k}$ values that ranges between $[185, 186, \ldots, 194]$ kg. The mass of the upper structure m_{u} of MAFDS is varied using additional payloads, refer Sect. 12.4. The experimental tests for $h_{\mathrm{f}} = 0.03$ m yields K observation events $A_{z_{\mathrm{r,max},k}}$ with a mean value of $\mu_A = 0.0758$ m and an interval $[0.074, \ldots, 0.0778]$ m.

Total Probability $p(A_{z_{\mathrm{r,\,max}}})$

Figure 12.6 illustrates the total probability

$$p(A_{z_{\mathrm{r,max}}}) = \sum_{n=1}^{N} p(A_{z_{\mathrm{r,max}}}|H_{z_{\mathrm{r,max}},n})\, p(H_{z_{\mathrm{r,max}},n}) \tag{12.23}$$

used as the denominator in BAYES posterior formula described in (12.20). The total probability also takes into account the total number N of hypothetical events $\{H_{z_{\mathrm{r,max}},1}, H_{z_{\mathrm{r,max}},2}, \ldots, H_{z_{\mathrm{r,max}},N}\}$.

Posterior Probability $p(H_{z_{\mathrm{r,max}},n}|A_{z_{\mathrm{r,max}}})$

With (12.23), the posterior probability in (12.20) becomes

$$p(H_{z_{\mathrm{r,max}},n}|A_{z_{\mathrm{r,max}}}) = \frac{p(A_{z_{\mathrm{r,max}}}|H_{z_{\mathrm{r,max}},n})\, p(H_{z_{\mathrm{r,max}},n})}{\sum_{n=1}^{N} p(A_{z_{\mathrm{r,max}}}|H_{z_{\mathrm{r,max}},n})\, p(H_{z_{\mathrm{r,max}},n})}. \tag{12.24}$$

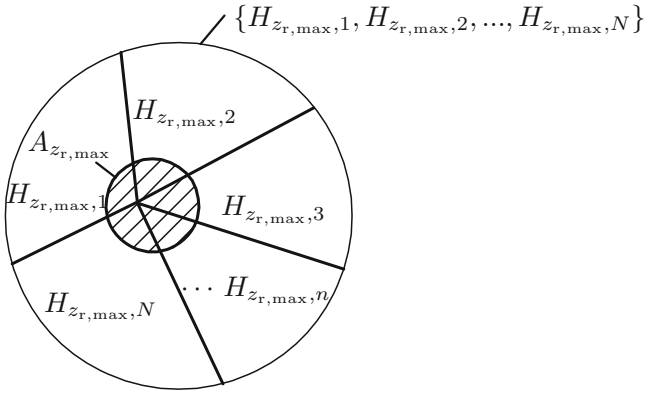


Fig. 12.6 Illustration of the law of total probability

It describes the probability of the hypothesis events $H_{z_{\mathrm{r,max}},n}$ predicted by the mathematical model, when the observation events $A_{z_{\mathrm{r,max}}}$ that are recorded by experiments are given. The denominator is a normalizing constant that ensures that the integral of the posterior probabilities equals unity.

12.5.2 Bayes Factor

The BAYES factor

$$B = \frac{p(\mu_H|A_{z_{\mathrm{r,max}}})}{p(\mu_H)} \tag{12.25}$$

is computed from the prior and posterior probabilities (12.24) and (12.21). The BAYES factor is defined as the ratio of the posterior probability $p(\mu_H|A_{z_{\mathrm{r,max}}})$ of the mean value μ_H of the hypothesis event to the prior probability $p(\mu_H)$ of the mean value μ_H of the hypothesis event [6].

12.5.3 Quantification of Uncertainty of Mathematical Models (a) to (e)

Deterministic Comparison of $z_{\mathrm{r,max}}$

A comparison of the prediction of experimental maximum relative compression $z_{\mathrm{r,\,max}}$ from the five mathematical models without taking uncertainty into account and under varying initial conditions $h_{\mathrm{f}} = 0.01, 0.02, \ldots, 0.07$ m is illustrated in Fig. 12.7. The experimental relative maximum compression $z_{\mathrm{r,\,max}} \approx 0.09$ m for $h_{\mathrm{f}} = 0.07$ m, which is close to the maximum relative compression limit of the spring-damper component of MAFDS. It is evident that only model (e) shows good agreement with the experimental data. Models (c) and (d) with RAYLEIGH damping show good agreement for a certain region of initial conditions $h_{\mathrm{f}} = 0.01, \ldots, 0.03$. The RAYLEIGH constants α and β are estimated for $h_{\mathrm{f}} = 0.02$. Hence, it can be inferred that the model with RAYLEIGH damping captures the dynamics of a system for only a certain region since it is based on linear damping assumptions. The models (a) and (b) show no agreement for all values of h_{f} and hence the axiomatic damping value b needs to be calibrated.

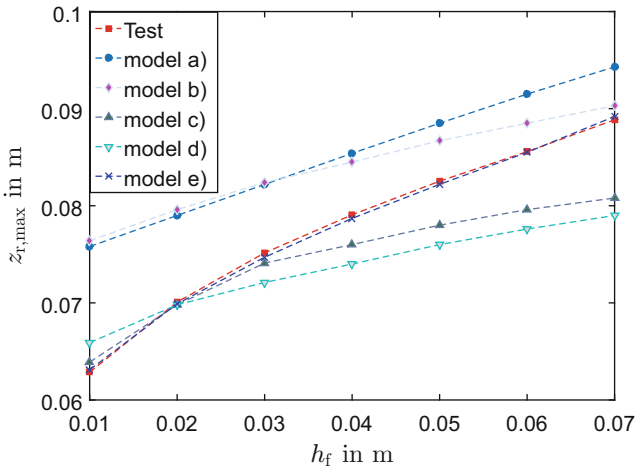


Fig. 12.7 Maximum compression $z_{\mathrm{r,\,max}}$ for all models under different initial conditions h_{f}

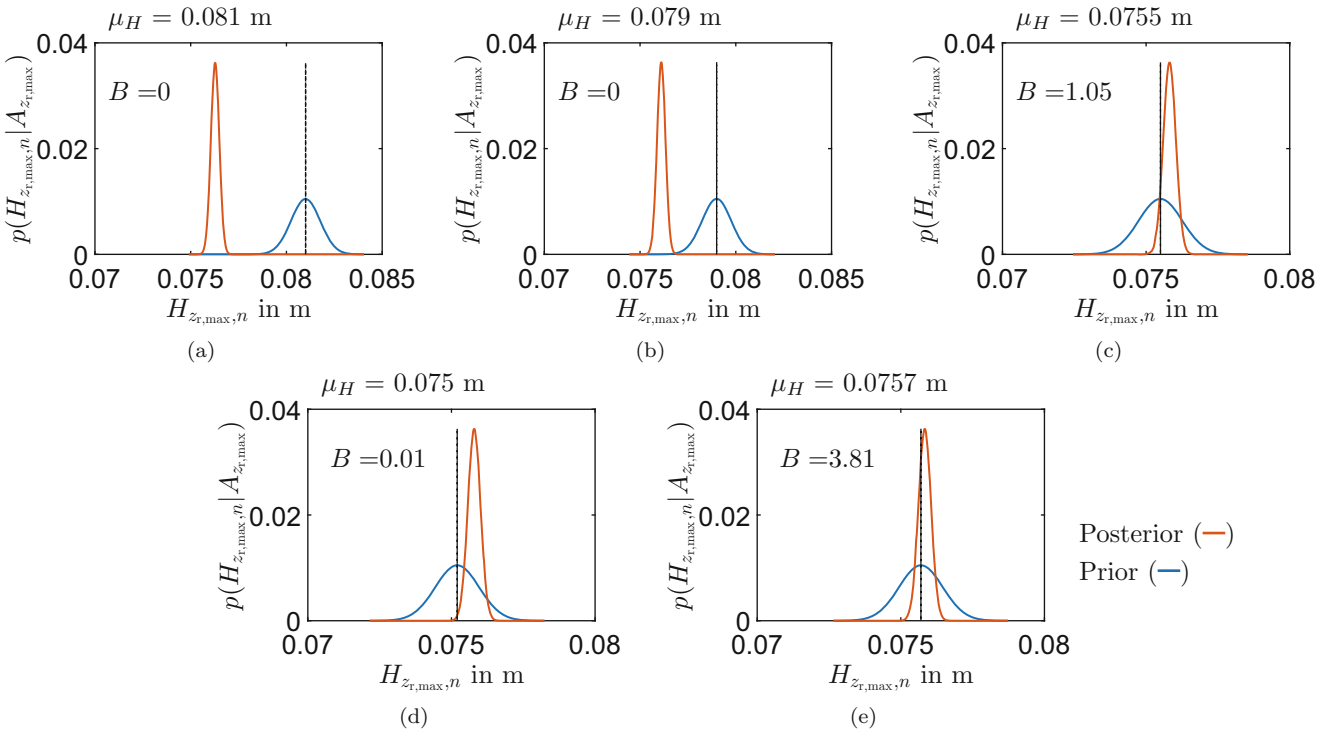


Fig. 12.8 BAYES factor B estimation for models (a) to (e) with posterior density (*red solid line*) and prior density (*blue solid line*) for 2DOF model with the assumptions: (**a**) linear stiffness and axiomatic damping, (**b**) nonlinear stiffness and axiomatic damping, (**c**) linear stiffness with RAYLEIGH damping, (**d**) nonlinear stiffness with RAYLEIGH damping, and (**e**) nonlinear stiffness with velocity dependent damping force for $h_f = 0.03$

Comparison of *B* for Mathematical Models (a) to (e)

Figure 12.8 illustrates the comparison of the posterior probability $p(H_{z_{r,max},n}|A_{z_{r,max}})$ and the prior probability $p(H_{z_{r,max},n})$. The calculations of posterior and prior probabilities are conducted according to the initial conditions (12.2) and (12.3) for given freefall height $h_f = 0.03$ m. As discussed in Sect. 12.5.2, the BAYES factor B is estimated for all the models (a) to (e). According to [6], if $B > 1$, then the mathematical model is adequate to represent the experimental data. Conversely, if $B < 1$, the mathematical model is considered as inappropriate. The BAYES factor for Model (c) and (e) is greater than 1, hence these two models are adequate to predict the compression of MAFDS under a given initial condition of $h_f = 0.03$. The BAYES factor for models (a) and (b) is 0 since more than 98% of $H_{z_{r,max}}$ events do not fall into the region where the experimental events $A_{z_{r,max}}$ lie. This is observed from the difference between the posterior and prior probabilities of model (a) and model (b), refer Fig. 12.8a, b.

12.6 Conclusion and Outlook

In this contribution, model uncertainty for a 2DOF mathematical model represented in five different approaches of using functional relations is quantified via BAYESIAN approach. The five different models are based on linear, nonlinear, axiomatic and empiric assumptions. The evaluation criteria is the differences in the prediction of the dynamic response, in this case, the maximum relative compression $z_{r,max}$. BAYES factor B is used to evaluate if the prediction of the mathematical model adequately represent the experimental observations. It is found that the mathematical model with velocity dependent damping force shows consistent agreement with the experimental data for different initial conditions. The damping in the other models (a) to (d) is based on linear assumptions. Hence, when these models are calibrated for certain initial conditions, they may

not show adequate prediction results for different initial conditions. In future work, the authors will explore the validation metrics that are based on both probabilistic and non-probabilistic approaches in order to validate the time history results of all the dynamic response variables, such as loads, displacements and stresses of the MAFDS.

Acknowledgements The authors thank the German Research Foundation DFG for funding this research within the SFB 805.

References

1. Oberkampf, W., Timothy, T.: Verification and validation in computational fluid dynamics. Prog. Aerosp. Sci. **38**(3), 209–272 (2002)
2. Oberkampf, W., Roy, C.: Verification and Validation in Scientific Computing. Cambridge University Press, Cambridge (2010)
3. Roy, C.: A Complete Framework for Verification, Validation, and Uncertainty Quantification in Scientific Computing. American Institute of Aeronautics and Astronautics, Inc., Reston (2010)
4. ASME: Guide for verification and validation in computational solid mechanics. American Society of Mechanical Engineers (2006)
5. Liu, Y., Wei, C., Arendt, P., Hong-Zhong, H.: Toward a better understanding of model validation metrics. J. Mech. Des. ASME **133**(7), 071005 (2011)
6. Sankararaman, S., Mahadevan, S.: Model validation under epistemic uncertainty. Reliab. Eng. Syst. Saf. **96**(9), 1232–1241 (2011)
7. Hills, R., Trucano, T.: Statistical validation of engineering and scientific models: background. Technical Report, SAND99-1256 (1999)
8. Melzer, C., Platz, R., Melz, T.: Comparison of methodical approaches to describe and evaluate uncertainty in the load-bearing capacity of a truss structure. In: Fourth International Conference on Soft Computing Technology in Civil, Structural and Environmental Engineering, paper 26. Civil-Comp Press, Stirlingshire (2015)
9. Enss, G., Gehb, C., Götz, B., Melz, T., Ondoua, S., Platz, R., Schäffner, M.: Load transferring device, Patent DE102014106858.A1 (2015)
10. Schaeffner, M., Goetz, B., Platz, R.: Active buckling control of a beam-column with circular cross-section using piezoelastic supports and integral LQR control. J. Smart Mater. Struct. **25**(6), 065008 (2016)
11. Gehb, C.,Platz, R., Melz, T.: Active load path adaption in a simple kinematic load-bearing structure due to stiffness change in the structure's supports. J. Phys. Conf. Ser. **744**, 012168 (2016)
12. Goetz, B., Schaeffner, M., Platz, R., Melz, T.: Lateral vibration attenuation of a beam with circular cross-section by a support with integrated piezoelectric transducers shunted to negative capacitances. J. Smart Mater. Struct. **25**(9), 095045 (2016)
13. Enss, G., Platz, R.: Evaluation of uncertainty in experimental active buckling control of a slender beam-column using Weibull analysis. J. Mech. Syst. Signal Process. **79**, 123–131 (2016)
14. Chowdhury, I., Dasgupta, S.: Dynamics of Structure and Foundation - A Unified Approach. CRC Press, Boca Raton (2008)
15. Marquardt, D.: An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. **11**(2), 431–441 (1963)
16. Gross, D., Werner, H., Schroeder, J., Wolfgang, A., Govindjee, S.: Engineering Mechanics, vol. 3. Springer, Berlin (2011)
17. Winkler, R.: Introduction to Bayesian Inference and Decision. Holt, Rinehart and Winston, New York (1972)
18. Rebba, R., Huang, S., Liu, Y., Mahadevan, S.: Statistical validation of simulation models. Int. J. Mater. Prod. Technol. **25**(1/2/3), 164–181 (2006)
19. Schuealler, G.: On the treatment of uncertainties in structural mechanics and analysis. Comput. Struct. **85**(5–6), 235–243 (2007)

Chapter 13
Unsupervised Novelty Detection Techniques for Structural Damage Localization: A Comparative Study

Zilong Wang and Young-Jin Cha

Abstract In recent decades, novelty detection has attracted considerable attention in the structural health monitoring field. Numerous machine-learning algorithms have been proposed to carry out novelty detection for structural damage detection and localization, owing to their abilities to identify abnormal data in large numbers of datasets. This paper introduces a number of unsupervised novelty detection methods, based on machine learning, and applies them to localize different types of damage in a laboratory scale (lab-scale) structure. The key concept behind unsupervised novelty detection is that the novelty detection model must be trained using only normal data. In this study, the model used to identify abnormal data in the testing datasets has been well trained using normal data. The unsupervised novelty detection methods in this paper include the Gaussian mixture method, one-class support vector machines, and the density peak-based fast clustering method, which was developed recently. To enable these methods to carry out novelty detection and to increase their localization accuracy, a number of improvements have been made to the original algorithms. In this comparative study of structural damage localization, two damage-sensitive features are extracted from the acceleration signals measured by the sensors installed on a complex lab-scale structure. The advantages and disadvantages of these methods are analyzed based on experimental comparative case studies of damage localization.

Keywords Unsupervised novelty detection • Machine learning methods • Structural damage localization • Clustering methods • Damage-sensitive features

13.1 Introduction

Many civil structures in service accumulate damage to varying degrees due to aging, overload, natural hazards, and other harmful effects. Thus, structural damage detection in an early stage of these structures is necessary to prevent their sudden collapse and prolong their service life [1, 2]. Vibration-based damage detection techniques are gaining popularity in the area of structural damage detection and localization. These techniques can be broadly classified into two branches: physical models-based methods and data-based methods [3]. Structural damage detection in data-based methods may be conducted in either supervised or unsupervised mode. Supervised learning methods require data from the structure in both its intact and damaged states for training. It is difficult to obtain sufficient abnormal data from the damaged large-scale civil structures due to scarce damaged infrastructures in service [4]. However, unsupervised learning methods only require the data from the intact structure in the training process. As a result, the unsupervised learning method is recommended to solve the problem of insufficient abnormal data. In this paper, many unsupervised novelty detection methods are applied to detect and locate damage in a lab-scale structure. Before applying these methods to damage detection, acceleration histories of the structural joints are measured, and two damage-sensitive features are extracted from them. Following the data acquisition and feature extraction, the normal features from each joint in the intact scenario are used to establish a statistical model operated by means of these methods. Then, the established normal models are used to detect novelty features in the testing feature set from the damaged scenario. In damage localization, the structural joint achieving the highest detection rate can be identified as the joint closest to the damage.

Z. Wang • Y.-J. Cha (✉)
Department of Civil Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada
e-mail: Young.cha@umanitoba.ca

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_13

13.2 Literature Review

A brief review of the unsupervised novelty detection methods proposed and applied to structural damage detection and localization in recent years is outlined below. According to the recent works, widely applied unsupervised novelty detection can be summarized according to three categories based on their detection techniques. The first category is the nearest neighbors technique. In the unsupervised mode of this technique, normal training points are seen as a cluster, and abnormal testing data points should lie far away from this normal cluster [4]. The K nearest neighbors (KNN) method is a common example in this technique and is widely used in damage detection. Walsh et al. [5] applied the KNN method to novelty detection for a real damaged bridge. The experimental results showed the applicability of KNN in damaged objective detection.

The second category is the clustering-based technique. In the unsupervised mode of this technique, the normal training points are first classified into a number of clusters. Next, decision boundary is created to determine whether the testing points belong to these normal clusters [4]. K-means, fuzzy C-means (FCM), and Gaussian mixture (GM) methods are well known in the clustering-based technique. For example, Manson et al. [6] applied the K-means method to cluster features extracted from the box girder of a bridge to detect its damage. The outcome of the experiment indicated that the K-means method was able to detect the cracks and trace their growth. Noh et al. [7] utilized a GM method to establish a statistical model and recognize the feature vectors extracted from the acceleration signals measured in a three-story and single-bay steel frame. The results of the experiment showed that the GM-based algorithm is effective in damage detection and localization under high-intensity and bidirectional loading cases.

The third category is the one-class support vector machine (OC-SVM) technique. An OC-SVM is a common branch of support vector machines (SVMs) for novelty detection [4]. The principle of novelty detection behind an OC-SVM is to seek an optimized hyperplane that encloses the majority of normal training points and allows few points outside the enclosed boundary. In the novelty detection process, the testing points sitting outside the boundary can be considered novelty points. Long and Buyukozturk [8] used an OC-SVM to detect various damage positions in a lab-scale steel structure. The experimental results indicated that the OC-SVM has high reliability in terms of damage detection but is weak at localizing damage positions. Khoa et al. [9] applied an OC-SVM to detect and localize damage in a three-story lab-scale building structure and in the Sydney Harbor Bridge in Australia. The experimental results showed good performance in both damage detection and localization.

13.3 Methodologies

In this section, two novelty detection methods in the previous literature review have been carefully selected and are improved in the unsupervised mode. In addition, a novel clustering method proposed in 2014 has been developed for the extensive comparative study of structural damage localization in Sect. 13.6.

13.3.1 GM Method

A Gaussian mixture model (GMM) is a weighted sum of several component Gaussian densities [10]. When a set of multidimensional vectors is scattered in a space, each component Gaussian density can be expressed as follows:

$$\mathrm{d}\left(\mathrm{x};\mu,\Sigma\right)=\frac{1}{2\pi^{\mathrm{D}/2}|\Sigma|^{1/2}}\exp\left|-\frac{1}{2}\left(\mathrm{x}\text{-}\mu\right)'\Sigma^{-1}\left(\mathrm{x}\text{-}\mu\right)\right| \tag{13.1}$$

where x is the multidimensional vector; D is the dimension of the multidimensional vector; μ and Σ are the mean vector and covariance matrix of the multidimensional vector dataset, respectively. Thus, the expression of a GMM can be formulated as:

$$P\left(x;w_i,\mu_i,\Sigma_i\right)=\sum_{i=1}^{c}w_i d\left(x;\mu_i,\Sigma_i\right) \tag{13.2}$$

where c is the number of clusters in the clustering cases; w is the mixture weight and satisfies the constraint: $\sum_{i=1}^{c}w_i=1$.

The data point clusters classified by the GM method have their own w_i, μ_i, and Σ_i. To achieve good clustering performance, these three parameters need to be optimized. The expectation maximization (EM) algorithm is widely used for parameter estimation [10]. The principle of the EM algorithm is to estimate new values of the three parameters based on their initial assumed values to make the new probability P in Eq. (13.2) greater than its previous value. This optimization process is terminated until the convergence threshold is reached.

After a set of normal data points is successfully clustered using the GM method, Masud et al.'s [11] method is used to give the trained normal model the ability to detect the novelty points. In this method, a circle is defined for each final formed normal cluster and to make it enclose the majority of points. The circle centers are assigned at the cluster centroids. The testing points outside of these circles can be considered novelty points. In Sect. 13.6, the radius of the circle for a cluster is the averaged distance from the normal points in the cluster to the cluster centroid.

13.3.2 OC-SVM

An OC-SVM is an extension of the original two-class (positive and negative) SVM for novelty detection. In the training process of novelty detection, an OC-SVM aims to build a decision boundary that encloses the majority of normal points. Few outliers outside the enclosed boundary can be considered noise points [8]. A well-trained OC-SVM model can solve the quadratic programming problem, which can be formulated as follows:

$$\begin{aligned} &\min_{w} && \frac{1}{2}w^2 + \frac{1}{vN}\sum_{i=1}^{N}\xi_i - b \\ &\text{subject to} && (w\cdot\tau(x_i)) \geq b - \xi_i, \quad \xi_i \geq 0 \end{aligned} \tag{13.3}$$

where v is a specified fraction of training objectives; N is the number of training objectives; ξ_i is the slack variable for ith objective; x_i is the ith training objective; $w\cdot\tau(x_i) - b$ is the function for optimization; and w and b are normal vector and bias terms, respectively. Without showing any more details of Lagrange multipliers and the kernel trick, the corresponding dual problem can be formulated as:

$$\begin{aligned} &\min_{\alpha} && \frac{1}{2}\sum_{i=1,j=1}^{N}\alpha_i\alpha_j k(x_j, x_i) \\ &\text{subjec to} && 0 \leq \alpha_i \leq \frac{1}{vN}, \quad \sum_{i=1}^{N}\alpha_i = 1 \end{aligned} \tag{13.4}$$

where $k(x_j, x_i)$ is the kernel function, typically the Gaussian kernel: $k(x, y) = \exp(-\gamma(x-y)^2)$; x_i is the ith training objective; and α_i is Lagrange coefficient of x_i.

In novelty detection, the testing objective x_t can be evaluated by the following sign functions:

$$f(x) = \text{sign}\left(\sum_{i}^{\#sv}\alpha_i k(x_i, x_t) - b\right) \tag{13.5}$$

$$b = \sum_{j=1}^{N}\alpha_j k(x_i, x_j) \tag{13.6}$$

where #SV is the index i of support vectors x_i, and the x_i is called a support vector if x_i is associated with non-zero α_i. When the sign function is negative, it means that the testing objective sits outside the enclosed boundary, which is defined by $\sum_{i}^{\#sv}\alpha_i k(x_i, x) - b = 0$. In addition, this testing objective can be classified as a novelty point.

13.3.3 Density Peaks-Based fast Clustering

The density peaks-based fast clustering (DPFC) method is a novel clustering method proposed by Rodriguez and Laio in 2014 [12]. The algorithm rule underlying this new method is composed of the following steps:

1. Calculate the local density ρ_i for point i. For the case studies in this paper, the Gaussian kernel function of radius is introduced to calculate the local density [13]. Its expression is as follows:

$$\rho_i = \sum_{j=1}^{N} \exp\left(-\left(d_{ij}-d_c\right)^2\right) \tag{13.7}$$

where N is the number of points for clustering in a space, d_c is a predefined cut-off distance, and d_{ij} is the distance between point i and j.

2. Calculate δ_i for point i. In this algorithm, δ_i is defined as the minimum distance between the point i and any other points with higher local density:

$$\delta_i = \min_{j:\rho_j>\rho_i}\left(d_{ij}\right) \tag{13.8}$$

For the point with the highest local density, take $\delta_i = \max_j\left(d_{ij}\right)$.

3. Select the points with a relatively high ρ_i and anomalously large δ_i as cluster centers. This choice makes the selected cluster centers surrounded by sufficient points. At the same time, these cluster centers are relatively far apart from each other.
4. Assign each remaining point to the same cluster as its nearest point of higher local density.
5. Identify halo points in the formed clusters. Halo points are defined as the set of points in one cluster having shorter distances than d_c to any points in the other clusters. Then, the halo point with the highest density in this cluster is found, and set its local density as threshold of this cluster for novelty detection, denoted by ρ_b. Compute the average value of ρ_b of all clusters, and set this value as a threshold of local density ρ_c.
6. Increase the value of d_c if the halo points cannot be detected in step 5. Introduce a weight for d_c to increase the updated cutoff distance wd_c until the halo points can be detected.
7. Add the testing points separately to the well-trained normal clusters formed in step 5. Calculate their local densities ρ_t and compare them with ρ_c. The testing points with lower local densities than ρ_c can be identified as novelty points.

13.4 Damage-Sensitive Features

In structural damage detection, the ideal features should be those that are sensitive to the presence of damage, but which are insensitive to operational and environmental variability in the normal range [1]. In Sect. 13.6, the acceleration histories $a(t_n)$ of structural joints are measured for feature extraction.

13.4.1 Crest Factor

The crest factor of the $a(t_n)$ is one feature that is verified as a sensitive quantity to the structural response before and after damage [1], and it can be calculated by:

$$\mathrm{CF} = \frac{|a|_{\max}}{a_{rms}} \tag{13.9}$$

where $|a|_{\max} = \max|a(t_n)|$; $a_{rms} = \sqrt{\frac{1}{n}\sum_n (a(t_n))^2}$.

13.4.2 Transmissibility

The second extracted feature sensitive to structural damage is the integral of the transmissibility curve in a frequency interval on the spectrum [8]. First, the raw acceleration response signals $a(t_n)$ need to be transformed into a frequency spectrum using fast Fourier transformation (FFT). The transmissibility curve is obtained by dividing the roving curve on the spectrum by the reference curve. The reference curve is a spectrum curve taken at the same degree of freedom (DoF) for all measurements of transmissibility. The transmissibility curve can be obtained by the following expression:

$$T_{ij}(\omega) = \frac{F_i(\omega)}{F_j(\omega)} \tag{13.10}$$

where ω is the frequency shown in the spectrum; $F_i(\omega)$ is the roving curve of spectrum at the ith DoF; and $F_j(\omega)$ is the reference curve of spectrum at the jth DoF.

It is essential to realize that the transmissibility curve $T(\omega)$ may be more sensitive at resonant frequencies of the structure. By checking spectra at a number of representative DoFs, some identified resonant frequencies are then used to define an appropriate frequency interval, which contains these identified resonant frequencies. Finally, the integral of the transmissibility curve is calculated in this interval, and it can be expressed by:

$$Int_i = \int_{\omega-\Delta\omega}^{\omega+\Delta\omega} T_{ij}(\omega) \tag{13.11}$$

where $\Delta\omega$ is the half length of the selected frequency interval; $T(\omega)$ is the transmissibility curve; and Int_i is the integral of the transmissibility curve at the ith DOF. By examining a number of spectrums at some representative joints, the selected frequency interval ranges from 40 to 50 Hz for the following case studies.

13.5 Experimental Setup

A three-story, two-bay, lab-scale steel structure is shown in Fig. 13.1 as the experimental structure [14]. The dimensions of the structural components, beams and columns, are 60cm × 5.08cm × 0.64cm. The components are bolted with four bolts at the structural joints, and the bottom of structure is bolted to a concrete foundation. Eighteen triaxial accelerometers are installed on the structure adjacent to the joints. The acceleration histories in X direction are used for case studies, as shown in Fig. 13.1. A small shaker attached at the top corner of structure, which is close to Joint 18, is used for excitation.

To carry out structural damage localization for the comparative studies in Sect. 13.6, two types of damage scenario (DS) are created. In the intact scenario, the bolts at all the structural joints are tightened. The bolts at Joint 9 are loosened in the DS 1. In the DS 2, a damaged column is used to replace the column between Joints 2 and 3. The experimental structure is tested 60 times for the intact scenario and 10 times for each damage scenario.

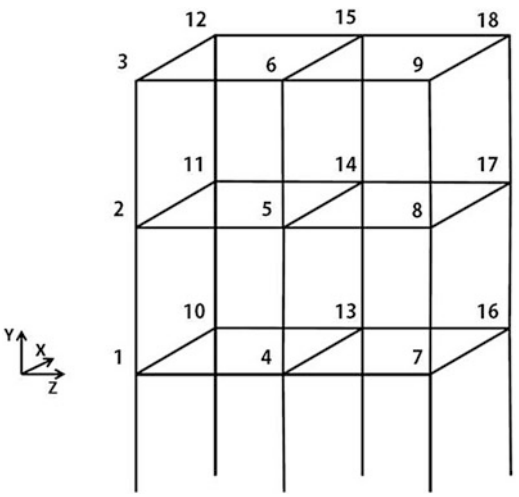


Fig. 13.1 Three-story two-bay structural model and joint positions

13.6 Comparative Case Studies

To evaluate damage detection and localization performance in the three representative methods introduced in Sect. 13.3, the two sets of damage-sensitive feature in Sect. 13.4 are processed into a two-dimensional feature vector set. Each vector corresponds to a sensor (joint) position in a single experimental test. The damage detection is carried out on a joint by joint basis. After mapping feature vectors into a space, we can obtain 60 intact points from the 60 tests in the intact scenario and 10 damage points from each damage scenario at each joint.

In damage localization, the joint that achieves the highest rate is closest to the damage. DS 1 can be localized to Joints 5 and 9 with the GM method, and damaged Joint 9 achieves one of the highest rates, as shown in Fig. 13.2. One weakness of the GM method is that in conditions of unknown distribution of training points, it is difficult to predefine an appropriate number of clusters. In addition, its algorithm lacks stability and reliability in novelty detection. Due to the randomly predefined initial cluster parameters, there is no way to avoid variability in the clustering results.

All 60 intact points are fitted into an OC-SVM to train an enclosed boundary. By tuning with the leave-one-out cross validation, the optimal value of parameter γ of the Gaussian kernel function in Eq. (13.4) can be obtained, which controls the shape of the boundary. The damage detection results are shown in Fig. 13.3. DS 1 is well localized to Joint 9, which is the exact damage position. DS 2 is also localized to Joint 2, which is one end of the damaged column. Although the OC-SVM performs better than the GM method in damage localization, its damage detection rates are obvious lower than the previous ones. It is anticipated that increasing the dimensions for feature vectors might lead to higher performance in damage detection.

Based on the damage detection rates shown in Fig. 13.4, DS 1 and DS 2 are all successfully localized to the damage positions using the DPFC method. The DPFC method shows several advantages in terms of damage localization compared to the other methods. The first advantage is its judicious selection of the number of clusters and cluster centers. The second is that the DPFC method is able to define a threshold local density for each formed normal cluster, which makes the established statistical models capable of identifying novelty points. However, one of the drawbacks of this method is that many parameters need to be artificially predefined and trained during the training process. For example, the minimum values of ρ in Eq. (13.7) and δ in Eq. (13.8) need to be carefully defined to form appropriate intact clusters. In addition, weight w in wd_c need to be adjusted to detect the halo points in the formed clusters, which is illustrated in step 6 of the DPFC method in Sect. 13.2.

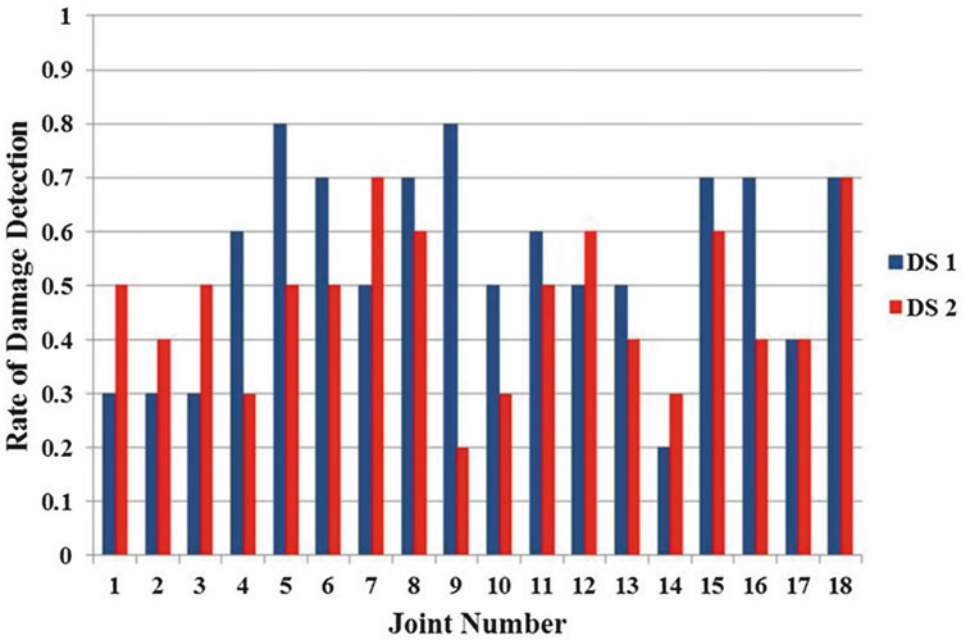


Fig. 13.2 Damage detection results at structural joints using the GM method

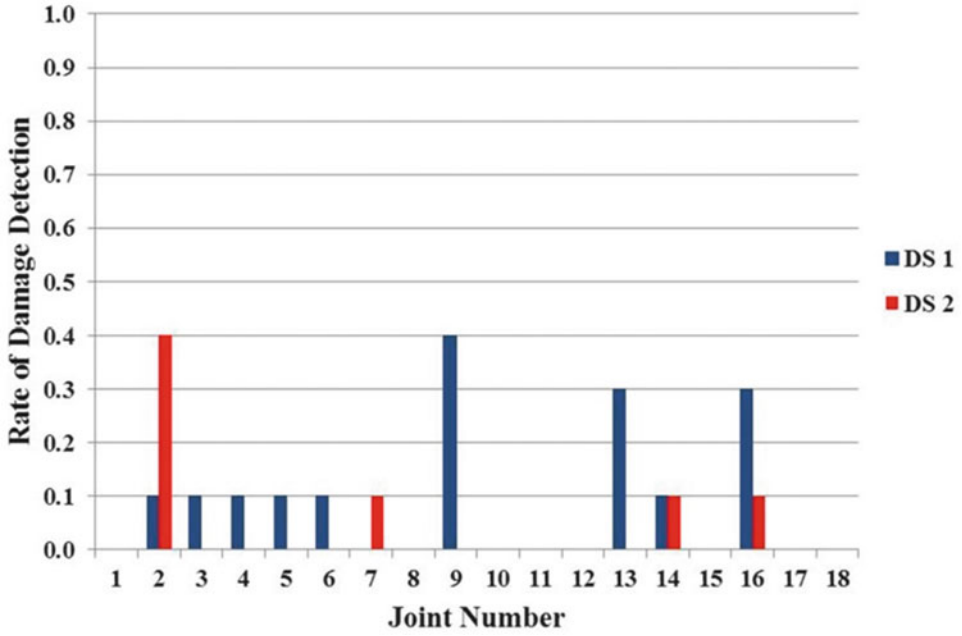


Fig. 13.3 Damage detection results at structural joints using an OC-SVM

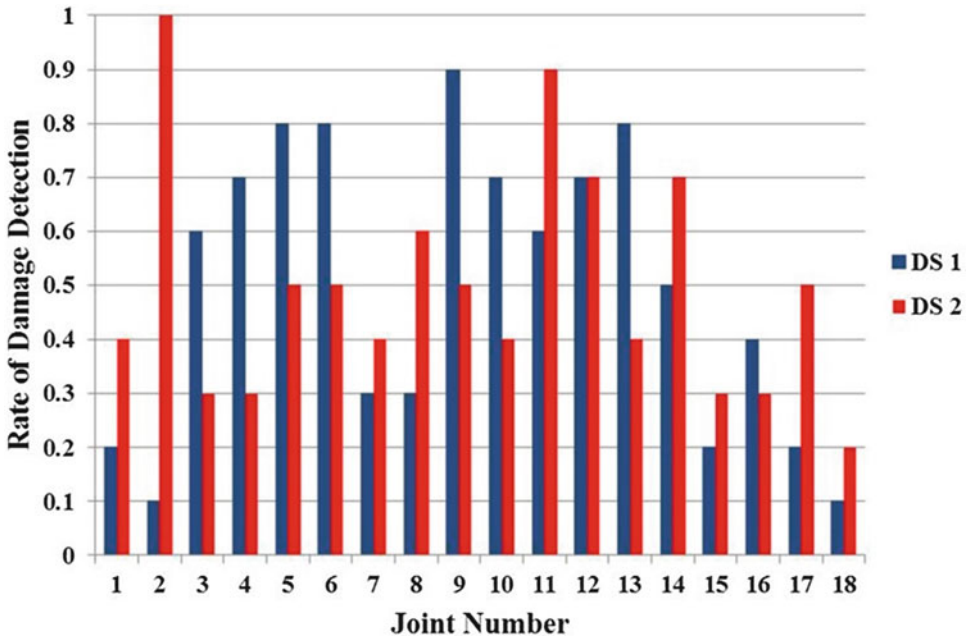


Fig. 13.4 Damage detection results at structural joints using the DPFC method

13.7 Conclusion

The main objective of this paper is to present and develop a number of unsupervised novelty detection methods and to apply them to damage localization in a lab-scale structure. The DPFC method achieved the best performance in terms of damage localization [15]. The OC-SVM technique can provided reliable information in damage localization, but improvements are needed to obtain higher rates of damage detection. Compared to the other methods, the damage localization results obtained using the GM method was worse than expected. In addition, two damage-sensitive features were extracted for the comparative studies, and they performed well in the proposed methods in detecting and locating damage in a complexity multi-bay and multi-story structural model. Thus, further research on the unsupervised novelty detection method should first focus on

extracting more appropriate damage-sensitive features from the acceleration histories. In addition, more research is needed to improve the efficiency of the DPFC method during the training process. The objective of such improvement is to reduce the workload of artificial parameter setting in the DPFC method and to apply this method to actual constructed buildings and bridges in real-life damage scenarios.

References

1. Farrar, C.R., Worden, K.: Structural Health Monitoring: A Machine Learning Perspective. Wiley, Hoboken, NJ (2012)
2. Cha, Y.J., Trocha, P., Buyukozturk, O.: Field measurement based system identification and dynamic response prediction of a unique MIT building. Sensors. **16**(7), 1016 (2016)
3. Barthorpe, R.J.: On model-and data-based approaches to structural health monitoring. Ph.D. thesis, University of Sheffield, Sheffield, UK (2010)
4. Ding, X., Li, Y., Belatreche, A., Maguire, L.P.: An experimental evaluation of novelty detection methods. Neurocomputing. **135**, 313–327 (2014)
5. Walsh, S.B., Borello, D.J., Guldur, B., Hajjar, J.F.: Data processing of point clouds for object detection for structural engineering applications. Comput. Aided Civ. Inf. Eng. **28**(7), 495–508 (2013)
6. Manson, G., Worden, K., Holford, K., Pullin, R.: Visualisation and dimension reduction of acoustic emission data for damage detection. J. Intell. Mater. Syst. Struct. **12**(8), 529–536 (2001)
7. Noh, H.Y., Nair, K.K., Kiremidjian, A.S., Loh, C.H.: Application of time series based damage detection algorithms to the benchmark experiment at the National Center for Research on Earthquake Engineering (NCREE) in Taipei, Taiwan. Smart Struct. Syst. **5**(1), 95–117 (2009)
8. Long, J., Buyukozturk, O.: Automated structural damage detection using one-class machine learning. Dyn. Civil Struct. **4**, 117–128 (2014)
9. Khoa, N.L.D., Zhang, B., Wang, Y., Chen, F., Mustapha, S.: Robust dimensionality reduction and damage detection approaches in structural health monitoring. Struct. Health Monit. **13**(4), 406–417 (2014)
10. Reynolds, D.: Gaussian mixture models. In: Encyclopedia of Biometrics, pp. 827–832. Springer, New York (2015)
11. Masud, M.M., Gao, J., Khan, L., Han, J., Thuraisingham, B.: Classification and novel class detection in concept-drifting data streams under time constraints. IEEE Trans. Knowl. Data Eng. **23**(6), 859–874 (2011)
12. Rodriguez, A., Laio, A.: Clustering by fast search and find of density peaks. Science. **344**(6191), 1492–1496 (2014)
13. Benoudjit, N., Archambeau, C., Lendasse, A., Lee, J.A., Verleysen, M.: Width optimization of the Gaussian kernels in radial basis function networks. In: Proceedings of the 10th European Symposium on Artificial Neural Networks, vol. 2, pp. 425–432 (2002)
14. Cha, Y.J., Buyukoztur, O.: Structural damage detection using modal strain energy and hybrid multiobjective optimization. Comput. Aided Civ. Inf. Eng. **30**(5), 347–358 (2015)
15. Cha, Y.J., Wang, Z.: Unsupervised novelty detection-based structural damage localization using a density peaks-based fast clustering algorithm. Struct. Health Monit. (2017). doi:10.1177/1475921717691260

Chapter 14
Global Load Path Adaption in a Simple Kinematic Load-Bearing Structure to Compensate Uncertainty of Misalignment Due to Changing Stiffness Conditions of the Structure's Supports

Christopher M. Gehb, Roland Platz, and Tobias Melz

Abstract Load-bearing structures with kinematic functions enable and disable degrees of freedom and are part of many mechanical engineering applications. The relative movements between a wheel and the body of a car or a landing gear and an aircraft fuselage are examples for load-bearing systems with defined kinematics. In most cases, the load is transmitted through a predetermined load path of a host structure to the structural support interfaces. However, uncertainty due to unexpected load peaks or varying health condition, e.g. changes in stiffness or damping parameters over time of the structure's components may require an adjustment of the load path for safety reasons. Load paths transmitted through damaged or weakened components can be the reason for reduced comfort or even failure. For example, reduced support stiffness can lead to uncertain and undesirable misalignment in the structure. In this paper, a two mass oscillator, a translatoric moving mass connected to a rigid beam by a spring-damper system, is used to numerically investigate the capability of load path adaption due to controlled semi-active guidance elements with friction brakes. The mathematical friction model will be derived by the LuGre approach. The rigid beam is embedded on two supports and is initially aligned with evenly distributed loads in beam and supports by the same stiffness condition. However, if uneven support stiffness occurs, e.g. by damage or fatigue, the beam becomes misaligned. One sided lowering of the beam may follow. Two auxiliary kinematic guidance elements are used to redirect the load path depending on the beam's alignment condition. With the semi-active auxiliary kinematic guidance elements it is possible to provide additional forces to relieve one of the beam's support if it changes its stiffness. The beams's misalignment is calculated numerically for varying stiffness parameters of the supports and is compared with and without semi-active auxiliary kinematic guidance elements. The structure is loaded with a force according to a step-function and a simple signal-based feedback PID-controller is designed to induce additional forces in the auxiliary guidance elements to bypass portions of loading away from supports with decreasing stiffness. Thus, uncertainty due to unacceptable misalignment caused by varying stiffness conditions of the structure's supports can be reduced by shifting load between the supports during operation.

Keywords Adaptive system • Kinematic elements • Load redistribution • Structural Health Control SHC • LuGre friction • PID-control

14.1 Introduction

Defined kinematics are often an important part of the functional performance in load-bearing systems with a specified movement of structural components. In most cases, the load is transmitted from one or more transmission points through a predetermined load path to the structural support interfaces. Example is compression stroke of landing gear or suspension strut in airplanes or vehicles. On the one hand, a spring-damper system often determines the main kinetic properties. On the other hand, the desired compression stroke is supported by kinematic guidance elements like torque-links or other suspension links as an auxiliary structure that link two or more parts of a primer load-bearing structure for stability reasons. The load is distributed to the structural support interfaces as predetermined and, mostly, not subject to any change during the structure's

C.M. Gehb (✉) • T. Melz
System Reliability, Adaptive Structures, and Machine Acoustics SAM, Technische Universität Darmstadt, Magdalenenstraße 4, 64289 Darmstadt, Germany
e-mail: gehb@sam.tu-darmstadt.de

R. Platz
Fraunhofer Institute for Structural Durability and System Reliability LBF, Bartningstraße 47, 64289 Darmstadt, Germany

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_14

lifetime. However, if system properties, e.g. damping and stiffness or strength of the supports are uncertain or vary over time, load path adaption to bypass portions of loading away from weakened structural components could be an option to prevent the structure from failure or reduced comfort. Accordingly, it might be useful to change the load path and shift the load to the stronger support interfaces.

An optimal design for the load path in a structural dynamic system like a truss with well known loading conditions and without moving structural components has been investigated thoroughly in literature, e.g. in [1, 2]. In general, typical optimization parameters are specific material and geometry values like Young's modulus, cross section areas of truss members and the shape of trusses, [3]. Another approach for optimal truss design is possible via damage tolerance, [4]. In this particular approach, a simple 18-bar truss structure is designed to not collapse despite initial damage. These approaches improve the dynamic structural behavior and are based mainly on the design of passive structures without any external energy that is fed into the structure for adapting purposes. With additional energy fed into a structure, for example with actuators creating counterforces to stabilize equilibrium conditions or attenuate vibrations, a structure becomes active. Active approaches found in literature mostly aim for an improvement in vibration or damping control to enhance particularly the dynamic behavior of truss members. For example, in [5, 6], the authors used semi-active joints with piezoelectric washers and stack actuators to improve the damping of two connected beams and a cantilever truss structure. The change of the load capacity, though, was not quantified. However, some research was conducted for enhanced load path distribution or change with active approaches in trusses. Studies to enhance the load capacity that could lead to load path adaption were made in [7] for a simple 9-bar truss with hydraulic jacks to apply internal forces and neuronal network as control strategy to react on unexpected high static load. In other studies, several beams of truss structures were substituted with idealized actuators in an academic way to enhance the load capacity by modifying the load path and to react to unknown loads, [8, 9]. Both approaches used actuators which were able to change their axial length and, hence, the bending stiffness of the examined truss to create a fully stressed state of all the beams in the truss. In systems with free to move but guided structural components within a defined trajectory like landing gears or car-suspension, (semi-)active systems are used for vibration control, mostly controlled by the sky-hook control strategy, [10, 11]. Load path adaption is, again, not addressed. Local load path adaption was conducted in [12] to reduce crack propagation with inducing active compression forces near the crack tip. By this approach, the stress intensity at the crack tip could be reduced significantly to achieve a 20% increase of durability. However, this local approach did not take into account global adaption of load paths through a structure.

In this work, a 2D two mass oscillator, a translatoric moving mass connected to a rigid beam by a spring-damper system, is used to show numerically the possibility of global load path adaption with friction brakes in semi-active kinematic guidance elements. Load path adaption might be necessary when parts of the structure become misaligned, for example the rigid beam by stiffness change in the supports due to damage leading to lowering the beam in a undesirable non symmetric way. The basic idea is to shift loads that normally go through the spring-damper but now go through the kinematic guidance elements, [13]. Load path adaption will be possible if semi-active friction forces are applied in the joints or hinges of the kinematic guidance elements that now become kinetic. The 2D two mass oscillator is derived from a complex load-bearing system that is currently developed in the German Collaborative Research Center (SFB) 805 "Control of Uncertainties in Load-Carrying Structures in Mechanical Engineering" at the Technische Universität Darmstadt. This load-bearing system is a Modular Active Spring Damper System (German acronym MAFDS) and serves as an example for a modular structure with passive, semi-active and active modules to study uncertainty in different approaches to control stability and vibration as well as adapt load paths.

14.2 Truss Structure Example MAFDS

Figure 14.1a shows the 3D design model of the MAFDS. The main parts are the upper and lower truss structure, a suspension strut from a mid-range car with a coil spring and a viscous damper combined to a spring-damper system, and kinematic guidance elements to realize a defined compression stroke trajectory between the upper and lower truss. The ratio of the masses of the upper and lower truss is similar to a typical quarter-car-model. The architecture of the MAFDS is modular, for example it is possible to replace the passive suspension strut by an active spring-damper system [14] or use piezoelastic supports for truss members for stability and vibration control [15, 16].

For the following approach to compensate misalignment of the upper structure, the 3D-structure of MAFDS in Fig. 14.1a is approximated by a simplified and downscaled 2D two mass oscillator structure, Fig. 14.1b. The two guidance elements in the 2D structure replace the three guidance elements in the 3D structure. Assuming the MAFDS to be upside-down, the lumped mass m_A approximates the lower structure, the beam m_B approximates the upper structure of the MAFDS. The lumped mass m_A is connected to the rigid beam m_B by a spring-damper-system and the two guidance elements.

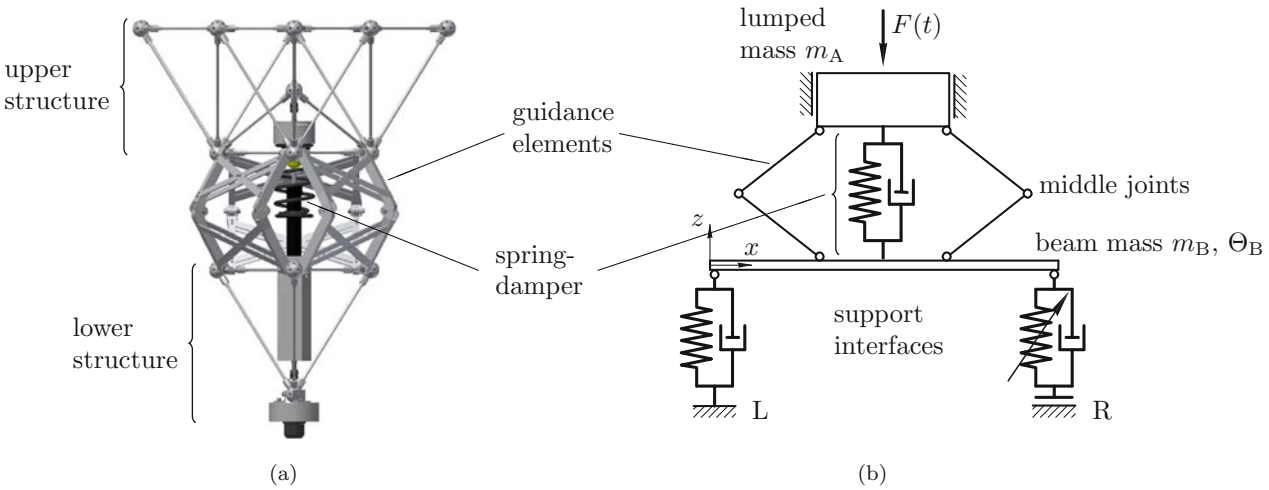


Fig. 14.1 (**a**) 3D design model of the MAFDS and (**b**) simplified 2D two mass oscillator

The rigid beam m_B is supported at its ends by support interfaces L and R. The stiffness characteristic of support R may vary. For example, if support R is damaged, its stiffness may decrease. In this case, the kinematic guidance elements will be augmented by semi-active kinetic functions with controlled friction forces in the middle joints to compensate beam misalignment.

In case the guidance elements fulfill only kinematic tasks, their main function is to enable the defined up-and-down compression stroke trajectory in z-direction, Fig. 14.1b. They do not carry forces between the two masses m_A and m_B in z-direction. The semi-active guidance elements become kinetic when they generate friction forces and, respectively, torques in the middle joints of the guidance elements. This way, an additional load path can be generated, for example with an electrodynamic actuator operating a disc brake.

In the following Sect. 14.3, the mathematical model of the two mass oscillator is derived and the LuGre-friction model according [17] and [18] is introduced to generate kinetic functionality of the guidance elements. In the numerical simulation in Sect. 14.4, the stiffness in support R is assumed to be less than the stiffness in support L, resulting into misalignment of the beam. If both support's stiffnesses are the same, the rigid beam will not misalign. Shifting the load away from support R with less stiffness towards support L with higher stiffness will restore alignment of the beam.

14.3 Mathematical Model of the 2D Two Mass Oscillator

The model parameters of the two mass oscillator with three degrees of freedom are shown in Fig. 14.2a. The oscillator has a mass m_A and a rigid beam with the mass m_B and the moment of inertia Θ_B in x–z-plane and the associated vertical displacements in z_A- and z_B-direction and rotation in φ-direction. The masses m_A and m_B are symmetrically connected with each other by angular guidance elements ① to ④ on each side of the beam. Elements ② and ④ are connected to the beam at the contact points $x = a$ and $x = l - a$. A spring-damper system using a coil spring and a viscous damper with stiffness and damping coefficients k_S and b_S is connected to the beam at $x = l/2$. The relative compression travel is $z_B - z_A$. The beam is connected to the ground at $x = 0$ and $x = l$ via two elastic supports L and R represented by two springs and two viscous dampers with stiffnesses k_L and k_R and damping coefficients b_L and b_R. The spring-damper system and the angular guidance elements are assumed to be free of mass. The linear time dependent displacements

$$z_L = -\varphi \frac{l}{2} + z_B \qquad \text{and} \qquad z_R = \varphi \frac{l}{2} + z_B \tag{14.1}$$

represent the compression travel of the supports for small angles φ with $\sin\varphi \approx \varphi$.

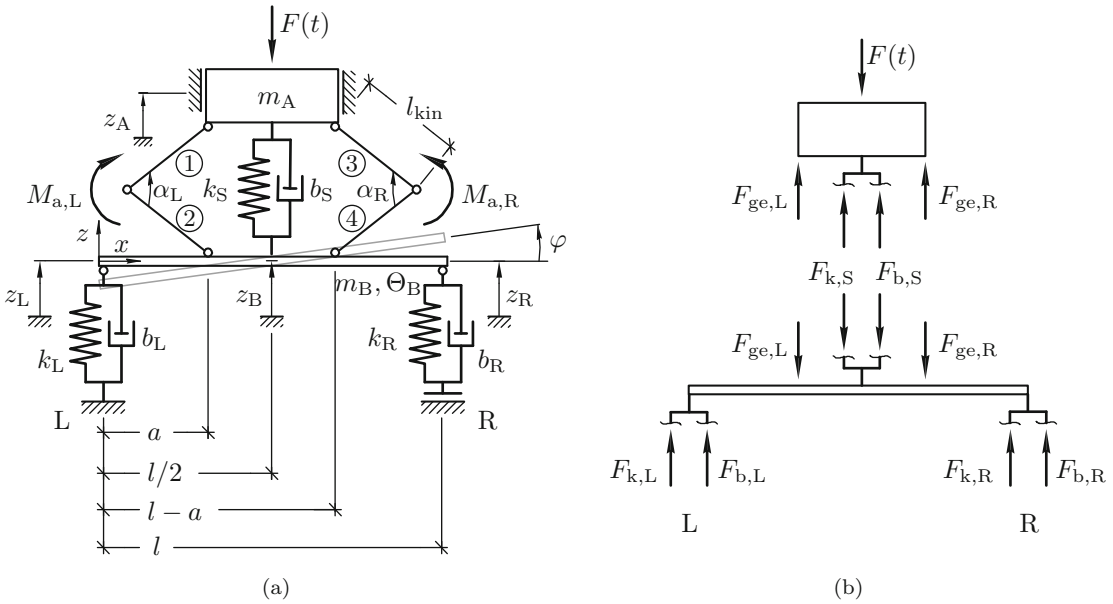


Fig. 14.2 (**a**) 2D two mass oscillator with semi-active moments $M_{a,L}$ and $M_{a,R}$ in the kinematic guidance elements and (**b**) internal and external forces

14.3.1 Internal and External Forces

According to Fig. 14.2b, the internal spring and damping forces of the spring-damper system are

$$F_{k,S} = k_S\,[z_B - z_A], \tag{14.2a}$$

$$F_{b,S} = b_S\,[\dot{z}_B - \dot{z}_A] \tag{14.2b}$$

and the internal spring and damping forces of the supports L and R are

$$F_{k,L} = -k_L\,z_L \quad \text{and} \quad F_{b,L} = -b_L\,\dot{z}_L \tag{14.3a}$$

$$F_{k,R} = -k_R\,z_R \quad \text{and} \quad F_{b,R} = -b_R\,\dot{z}_R. \tag{14.3b}$$

The combined support reaction and spring-damper system forces are

$$F_L = F_{k,L} + F_{b,L}, \tag{14.4a}$$

$$F_R = F_{k,R} + F_{b,R}, \tag{14.4b}$$

$$F_S = F_{k,S} + F_{b,S}. \tag{14.4c}$$

The external excitation force

$$F(t) = \begin{cases} 0 & \text{for } t < 0.1\,\text{s}, \\ \widehat{F} & \text{for } t \geq 0.1\,\text{s} \end{cases} \tag{14.5}$$

is assumed as a step-function, Fig. 14.3. When assuming the two mass oscillator to be passive, the guidance elements have only a kinematic function. When assuming the two mass oscillator to be semi-active with load path adaption, the kinematic guidance elements provide additional moments $M_{a,L}$ and $M_{a,R}$, Fig. 14.2a, resulting in additional kinetic functions. $M_{a,L}$ and $M_{a,R}$ result in vertical guidance element forces $F_{ge,L}$ and $F_{ge,R}$, Fig. 14.2b. They are coupled by the kinematic transmission

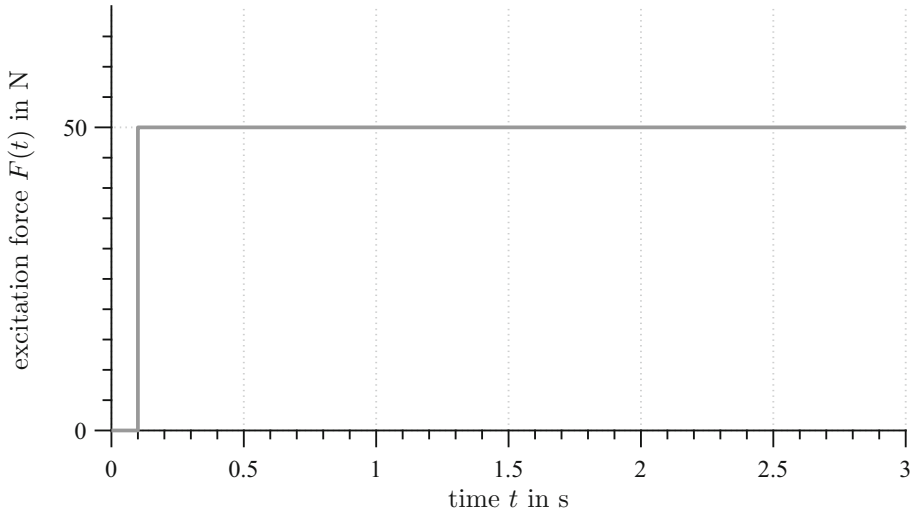


Fig. 14.3 Assumed step function $F(t)$

$$F_{\mathrm{ge,L}} = \frac{1}{l_{\mathrm{kin}} \cos\alpha_{\mathrm{L}}} M_{\mathrm{a,L}}, \tag{14.6a}$$

$$F_{\mathrm{ge,R}} = \frac{1}{l_{\mathrm{kin}} \cos\alpha_{\mathrm{R}}} M_{\mathrm{a,R}} \tag{14.6b}$$

with the length l_{kin} of each guidance component ① to ④ and the relative angle α_{L} between ① and ② as well as the relative angle α_{R} between ③ and ④, Fig. 14.2a.

14.3.2 Equation of Motion System

According to the direction of internal forces (14.2) and (14.3) in Fig. 14.2b, the linear equation of motion for the upper mass m_{A} becomes

$$m_{\mathrm{A}}\ddot{z}_{\mathrm{A}} - b_{\mathrm{S}}\left[\dot{z}_{\mathrm{B}} - \dot{z}_{\mathrm{A}}\right] - k_{\mathrm{S}}\left[z_{\mathrm{B}} - z_{\mathrm{A}}\right] = -F + F_{\mathrm{ge,L}} + F_{\mathrm{ge,R}} \tag{14.7}$$

and for the beam m_{B} and Θ_{B}

$$\begin{aligned} m_{\mathrm{B}}\ddot{z}_{\mathrm{B}} + b_{\mathrm{S}}\left[\dot{z}_{\mathrm{B}} - \dot{z}_{\mathrm{A}}\right] + k_{\mathrm{S}}\left[z_{\mathrm{B}} - z_{\mathrm{A}}\right] + k_{\mathrm{L}}\left[-\varphi\,\frac{l}{2} + z_{\mathrm{B}}\right] \ldots \\ + b_{\mathrm{L}}\left[-\dot{\varphi}\,\frac{l}{2} + \dot{z}_{\mathrm{B}}\right] + k_{\mathrm{R}}\left[\varphi\,\frac{l}{2} + z_{\mathrm{B}}\right] + b_{\mathrm{R}}\left[\dot{\varphi}\,\frac{l}{2} + \dot{z}_{\mathrm{B}}\right] = -F_{\mathrm{ge,L}} - F_{\mathrm{ge,R}} \end{aligned} \tag{14.8a}$$

$$\begin{aligned} \Theta_{\mathrm{B}}\ddot{\varphi} - k_{\mathrm{L}}\,\frac{l}{2}\left[-\varphi\,\frac{l}{2} + z_{\mathrm{B}}\right] - b_{\mathrm{L}}\,\frac{l}{2}\left[-\dot{\varphi}\,\frac{l}{2} + \dot{z}_{\mathrm{B}}\right] + k_{\mathrm{R}}\,\frac{l}{2}\left[\varphi\,\frac{l}{2} + z_{\mathrm{B}}\right] \ldots \\ + b_{\mathrm{R}}\,\frac{l}{2}\left[\dot{\varphi}\,\frac{l}{2} + \dot{z}_{\mathrm{B}}\right] = \left(\frac{l}{2} - a\right) F_{\mathrm{ge,L}} - \left(\frac{l}{2} - a\right) F_{\mathrm{ge,R}} \end{aligned} \tag{14.8b}$$

for translational z_{A}- and z_{B}-directions as well as for rotational φ-direction. Equations (14.7) and (14.8) are combined to a linear equation of motion system to become

$$\underbrace{\begin{bmatrix} m_\mathrm{A} & 0 & 0 \\ 0 & m_\mathrm{B} & 0 \\ 0 & 0 & \Theta_\mathrm{B} \end{bmatrix}}_{\boldsymbol{M}} \underbrace{\begin{bmatrix} \ddot{z}_\mathrm{A} \\ \ddot{z}_\mathrm{B} \\ \ddot{\varphi} \end{bmatrix}}_{\ddot{\boldsymbol{r}}} + \underbrace{\begin{bmatrix} b_\mathrm{S} & -b_\mathrm{S} & 0 \\ -b_\mathrm{S} & b_\mathrm{S}+b_\mathrm{L}+b_\mathrm{R} & -\frac{l}{2}b_\mathrm{L}+\frac{l}{2}b_\mathrm{R} \\ 0 & -\frac{l}{2}b_\mathrm{L}+\frac{l}{2}b_\mathrm{R} & \frac{l^2}{4}b_\mathrm{L}+\frac{l^2}{4}b_\mathrm{R} \end{bmatrix}}_{\boldsymbol{D}} \underbrace{\begin{bmatrix} \dot{z}_\mathrm{A} \\ \dot{z}_\mathrm{B} \\ \dot{\varphi} \end{bmatrix}}_{\dot{\boldsymbol{r}}} + \dots$$

$$\underbrace{\begin{bmatrix} k_\mathrm{S} & -k_\mathrm{S} & 0 \\ -k_\mathrm{S} & k_\mathrm{S}+k_\mathrm{L}+k_\mathrm{R} & -\frac{l}{2}k_\mathrm{L}+\frac{l}{2}k_\mathrm{R} \\ 0 & -\frac{l}{2}k_\mathrm{L}+\frac{l}{2}k_\mathrm{R} & \frac{l^2}{4}k_\mathrm{L}+\frac{l^2}{4}k_\mathrm{R} \end{bmatrix}}_{\boldsymbol{K}} \underbrace{\begin{bmatrix} z_\mathrm{A} \\ z_\mathrm{B} \\ \varphi \end{bmatrix}}_{\boldsymbol{r}} = \underbrace{\begin{bmatrix} -F + F_{\mathrm{ge,L}} + F_{\mathrm{ge,R}} \\ -F_{\mathrm{ge,L}} - F_{\mathrm{ge,R}} \\ (\frac{l}{2}-a)\,F_{\mathrm{ge,L}} - (\frac{l}{2}-a)\,F_{\mathrm{ge,R}} \end{bmatrix}}_{\boldsymbol{F}}. \tag{14.9}$$

In (14.9), $\boldsymbol{M}$, $\boldsymbol{D}$ and $\boldsymbol{K}$ are the global mass, damping and stiffness matrices, $\ddot{\boldsymbol{r}}$, $\dot{\boldsymbol{r}}$, $\boldsymbol{r}$ and $\boldsymbol{F}$ are the translational and angular acceleration, velocity, displacement vectors and the force vector. The force vector $\boldsymbol{F}$ in (14.9) contains the external excitation force $F = F(t)$ according to (14.5) without extra declaration of time t, and the forces $F_{\mathrm{ge,L}}$ and $F_{\mathrm{ge,R}}$ provided by the semi-active guidance elements in Fig. 14.2b.

14.3.3 LuGre *Friction Model*

Experimentally, the semi-active moments

$$M_{\mathrm{a,L}} = \mu_\mathrm{L}(q, \dot{q}, \Delta\dot{x}_{\mathrm{ge,L}})\, F_{\mathrm{a,L}}\, r_{\mathrm{brake}}, \tag{14.10a}$$

$$M_{\mathrm{a,R}} = \mu_\mathrm{R}(q, \dot{q}, \Delta\dot{x}_{\mathrm{ge,R}})\, F_{\mathrm{a,R}}\, r_{\mathrm{brake}}. \tag{14.10b}$$

may be generated by electrodynamic friction brakes in the guidance elements middle joints with controlled brake forces $F_{\mathrm{a,L,R}}$, see Sect. 14.3.4, and an average brake disk radius r_{brake}. The state dependent friction coefficients $\mu_{\mathrm{L,R}}$ are modeled by the dynamic LuGre model [17, 18]. This model is capable of predicting relevant friction effects and thereby predicting the non-linear force generating behavior of the friction brake. In the LuGre model, the contact dynamics of the interacting surfaces are derived by assuming contact by bristles. The average lateral deflection of the bristles is represented by a new time dependent state variable q which is governed by a first-order differential equation

$$\dot{q} = \Delta\dot{x}_{\mathrm{ge,L,R}} - \frac{|\Delta\dot{x}_{\mathrm{ge,L,R}}|}{g(\Delta\dot{x}_{\mathrm{ge,L,R}})} q \tag{14.11}$$

with the relative velocity $\Delta\dot{x}_{\mathrm{ge,L,R}}$ at the friction interface. The function $g(\Delta\dot{x}_{\mathrm{ge,L,R}})$ considers different coefficients for static and dynamic friction μ_s and μ_d, it contains information about the velocity dependence of the resulting friction force. A reasonable choice of $g(\Delta\dot{x}_{\mathrm{ge,L,R}})$ proposed in [17] leads to

$$g(\Delta\dot{x}_{\mathrm{ge,L,R}}) = \frac{1}{\sigma_0}\left[\mu_\mathrm{d} + (\mu_\mathrm{s} - \mu_\mathrm{d}) e^{-(\Delta\dot{x}_{\mathrm{ge,L,R}}/v_\mathrm{S})^2}\right] \tag{14.12}$$

with the Stribeck velocity v_S and the characteristic bristle stiffness σ_0. Finally, the state dependent friction coefficients $\mu_{\mathrm{L,R}}$ according to [17] can be stated as

$$\mu_{\mathrm{L,R}}(q, \dot{q}, \Delta\dot{x}_{\mathrm{ge,L,R}}) = \sigma_0 q + \sigma_1 \dot{q} + \sigma_2 \Delta\dot{x}_{\mathrm{ge,L,R}} \tag{14.13}$$

with the bristle damping coefficient σ_1 and viscous friction σ_2. The brake forces are the contact normal forces leading to the friction forces.

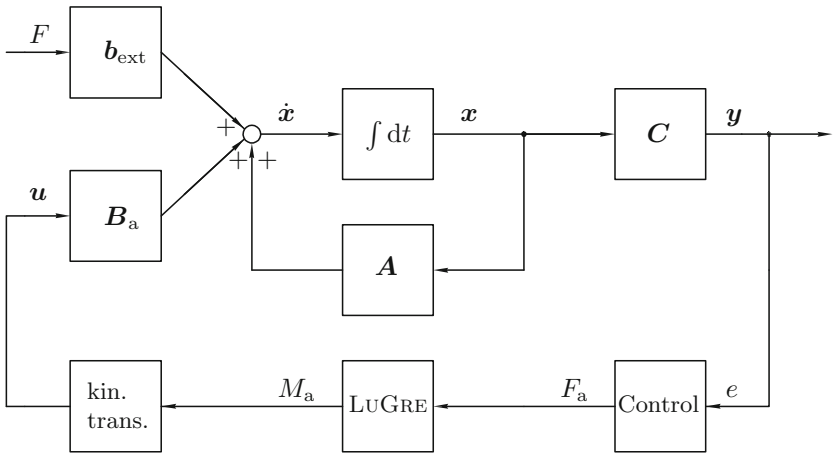


Fig. 14.4 Block diagram for state space model with control

14.3.4 Controller for Semi-active Friction Force

A signal-based feedback PID-controller is implemented to control the semi-active brake forces $F_{a,L,R}$ or, respectively, moments $M_{a,L,R}$ according to (14.10). The time dependent control input variable is the actual control deviation

$$e(t) = z_L(t) - z_R(t) \tag{14.14}$$

representing the beam's misalignment.

As mentioned before, semi-active control will be realized by electrodynamic actuators in the guidance elements. However, the actuator's dynamic behavior is neglected and it is assumed that the idealized actuators convert the controller output directly into the forces $F_{a,L}$ and $F_{a,R}$. Hence, the controller's output forces with $e_L(t)$ for the left and $e_R(t) = -e_L(t)$ for the right guidance elements are

$$F_{a,L}(t) = \begin{cases} \left| K_P\, e_L(t) + K_I \int e_L(t)\,dt + K_D \dfrac{d\, e_L(t)}{dt} \right| & \text{for} \quad \operatorname{sgn}(\mu_L) = \operatorname{sgn}(F_{a,L}) \\ 0 & \text{for} \quad \operatorname{sgn}(\mu_L) \neq \operatorname{sgn}(F_{a,L}), \end{cases} \tag{14.15a}$$

$$F_{a,R}(t) = \begin{cases} \left| K_P\, e_R(t) + K_I \int e_R(t)\,dt + K_D \dfrac{d\, e_R(t)}{dt} \right| & \text{for} \quad \operatorname{sgn}(\mu_R) = \operatorname{sgn}(F_{a,R}) \\ 0 & \text{for} \quad \operatorname{sgn}(\mu_R) \neq \operatorname{sgn}(F_{a,R}). \end{cases} \tag{14.15b}$$

In (14.15), K_P is the gain of the proportional component, K_I is the gain of the integral component and K_D is the gain of the derivation component. K_P, K_I and K_D are the tunable parameters of the PID-controller. They are determined by the Ziegler-Nichols tuning method and from that base fine-tuned empirically, [19] and Table 14.1.

The case analysis in the control law (14.15) is required to consider the effect of the relative translational movement between the mass m_A and the contact points of the guidance elements. The semi-active forces $F_{a,L,R}$ are controlled friction forces and are always positive. They act in z-direction according to (14.6) and (14.10). Their effect depends on the relative translational movement between the mass m_A and the contact points at $x = a$ and $x = l - a$ in Fig. 14.2a of the guidance elements. It is considered in $\mu_{L,R}$ according to (14.10).

14.3.5 State Space Model with Control

Figure 14.4 shows the block diagram for the proposed state space model with control to reduce the beam's misalignment by semi-active measures. All state variables and vectors are time dependent, even though time t is not pointed out, e.g. control deviation $e = e(t)$.

The support's reaction forces F_L and F_R and displacements z_L and z_R of the supports L and R at $x = 0$ and $x = l$ need to be calculated to investigate the ability to load path adaption, or respectively, redirection to reduce possible misalignment of the rigid beam, Fig. 14.3. For that, it is useful to solve (14.9) in time domain numerically. In state space representation, (14.9) becomes

$$\dot{\boldsymbol{x}} = \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\boldsymbol{M}^{-1}\boldsymbol{K} & -\boldsymbol{M}^{-1}\boldsymbol{D} \end{bmatrix}}_{[6\times6]} \boldsymbol{x} + \underbrace{\begin{bmatrix} \mathbf{0} \\ \boldsymbol{M}^{-1}\boldsymbol{B}_\mathrm{a} \end{bmatrix}}_{[6\times2]} \boldsymbol{u} + \underbrace{\begin{bmatrix} \mathbf{0} \\ \boldsymbol{M}^{-1}\boldsymbol{b}_\mathrm{ext} \end{bmatrix}}_{[6\times1]} F \tag{14.16}$$

with matrix dimensions given in brackets.

The semi-active forces $F_\mathrm{ge,L}$ and $F_\mathrm{ge,R}$ provided by the guidance elements as a result of (14.15), (14.10) and (14.6) are summarized in the control input vector

$$\boldsymbol{u} = \begin{bmatrix} F_\mathrm{ge,L} \\ F_\mathrm{ge,R} \end{bmatrix}. \tag{14.17}$$

Eventually, and by the control, the output vector

$$\boldsymbol{y} = \begin{bmatrix} 0 & -1 & \frac{l}{2} & 0 & 0 & 0 \\ 0 & 1 & \frac{l}{2} & 0 & 0 & 0 \\ 0 & -k_\mathrm{L} & \frac{l}{2}k_\mathrm{L} & 0 & -d_\mathrm{L} & \frac{l}{2}d_\mathrm{L} \\ 0 & k_\mathrm{L} & \frac{l}{2}k_\mathrm{L} & 0 & d_\mathrm{L} & \frac{l}{2}d_\mathrm{L} \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} z_\mathrm{L} \\ z_\mathrm{R} \\ F_\mathrm{L} \\ F_\mathrm{R} \end{bmatrix} \tag{14.18}$$

contains the displacements z_L and z_R needed to minimize the control deviation $e(t)$ in (14.14) and, therefore, needed for the load path adaption as well as the reaction forces F_L and F_R. In (14.16) and (14.18), $\boldsymbol{x} = [\boldsymbol{r}, \dot{\boldsymbol{r}}]^T$ is the $[6 \times 1]$ state vector with translational and angular displacement and velocity vectors in (14.9), the zero and identity matrices $\mathbf{0}$ and $\mathbf{I}$ are of appropriate dimensions. The forces $F_\mathrm{ge,L}$ and $F_\mathrm{ge,R}$ and the external excitation force F are allocated to the system by the $[3 \times 2]$ control input matrix and the $[3 \times 1]$ excitation input vector

$$\boldsymbol{B}_\mathrm{a} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \\ (\frac{l}{2} - a) & -(\frac{l}{2} - a) \end{bmatrix} \quad \text{and} \quad \boldsymbol{b}_\mathrm{ext} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \tag{14.19}$$

In short form, (14.16) and (14.18) become

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{A}\,\boldsymbol{x} + \boldsymbol{B}_\mathrm{a}\,\boldsymbol{u} + \boldsymbol{b}_\mathrm{ext}\,F \\ \boldsymbol{y} &= \boldsymbol{C}\,\boldsymbol{x} \end{aligned} \tag{14.20}$$

representing the state space model with control of the two mass oscillator with three degrees of freedom and semi-active forces as well as an excitation force.

14.4 Numerical Simulation of Load Path Adaption

In this section, numerical simulations of the two mass oscillator in Fig. 14.2 and according to (14.20) are presented to prove its load path adaption capability. The misalignment $e = z_\mathrm{L-}z_\mathrm{R}$, the support reaction forces F_L and F_R and the corresponding guidance forces $F_\mathrm{ge,L}$ and $F_\mathrm{ge,R}$ during semi-active load shift are shown in Figs. 14.5, 14.6, and 14.7. Three cases are investigated:

(i) both supports L and R have equal stiffnesses $k_L = k_R$, the system is passive and assumed to be without kinetic function of the guidance elements,
(ii) support R is assumed to be damaged with reduced stiffness and $k_R = 0.5\,k_L$, the system is passive and assumed to be without kinetic function of the guidance elements,
(iii) support R is assumed to be damaged with reduced stiffness and $k_R = 0.5\,k_L$, the system is semi-active and assumed to be with auxiliary kinetic function of the guidance elements.

The aim is to minimize the misalignment $e \to \min$ according to (14.14), after exciting mass m_A with the excitation force $F(t)$ according to (14.5) with $\widehat{F} = 50\,\mathrm{N}$ applied to the system at $t = 0.1\,\mathrm{s}$. Support R is assumed to have reduced stiffness $k_R = 0.5\,k_L$ for case (ii) and (iii). The load path adaption due to the semi-active kinematic guidance elements is used to shift load away from support R with less stiffness towards support L with the original higher stiffness to counteract the resulting misalignment in the passive system. Table 14.1 summarizes all properties of the two mass oscillator and the LuGre-Model.

Figure 14.5 shows the misalignment according to external excitation mentioned above for simulation time $t = 0$–$3\,\mathrm{s}$. For the three cases

(i) the misalignment is $e = z_{L-}z_R = 0$ for the total simulation time $t = 0$–$3\,\mathrm{s}$ even with the beginning of excitation (14.5). No overshoot and no misalignment oscillations can be observed,
(ii) the misalignment is $e = z_{L-}z_R \neq 0$ from time $0.1 \leq t \leq 3\,\mathrm{s}$ with the beginning of excitation (14.5). A constant steady state misalignment for $t > 2\,\mathrm{s}$ after high oscillations as a result of the step excitation is observed. An overshoot and misalignment oscillations after step excitation can be observed,
(iii) the misalignment is again $e = z_{L-}z_R \neq 0$ from time $0.1 \leq t \leq 3\,\mathrm{s}$ with the beginning of excitation (14.5). With the controlled semi-active forces according to (14.6a) and (14.6b) it is possible to reduce the steady state misalignment $e = z_{L-}z_R$ by 45% at $t > 2\,\mathrm{s}$. The overshoot and the oscillations are more damped when compared to case (ii).

The semi-active application in case (iii) leads to a misalignment reduction in comparison to the passive system in case (ii) but still steady state misalignment of the beam.

Figure 14.6 shows the support's reaction forces F_L and F_R. For case

(i) the reaction forces are $F_L = F_R$ from time $t = 0$–$3\,\mathrm{s}$ and similar to case (ii). Therefore, they are not plotted in Fig. 14.6,
(ii) the reaction forces are $F_L \approx F_R$ from time $t = 0$–$3\,\mathrm{s}$ and, therefore, are plotted in one graph in Fig. 14.6. The excitation load is evenly split into $F_L = F_R = 25\,\mathrm{N}$ in the steady state $t > 2\,\mathrm{s}$ even though, the misalignment for system (ii) is $z_{L-}z_R \neq 0$, cf. Fig. 14.5,

Table 14.1 Parameters for 2D two mass oscillator

Property	Symbol	Value	Unit
Damping coefficient	b_S	5	N s/m
Stiffness coefficient	k_S	1000	N/m
Stiffness support left	k_L	5000	N/m
Stiffness support right	k_R	5000	N/m
Stiffness support right, damaged	k_R	2500	N/m
Mass A	m_A	1	kg
Mass B	m_B	0.5	kg
Moment of inertia	Θ	0.0417	kg m^2
Beam length	l	1	m
Guidance element length	l_{kin}	0.2	m
Contact point	a	0.25	m
Excitation step amplitude	$\widehat{F}$	50	N
LuGre bristle stiffness	σ_0	10^5	m^{-1}
LuGre damping coefficient	σ_1	25	s/m
LuGre viscous damping	σ_2	0	s/m
LuGre static friction coefficient	μ_s	1	–
LuGre dynamic friction coefficient	μ_d	0.8	–
LuGre stribeck velocity	v_S	10^{-3}	m/s
Proportional control gain	K_P	$1.5 \cdot 10^4$	N/m
Integral control gain	K_I	10^3	N/s m
Derivation control gain	K_D	10	N s/m

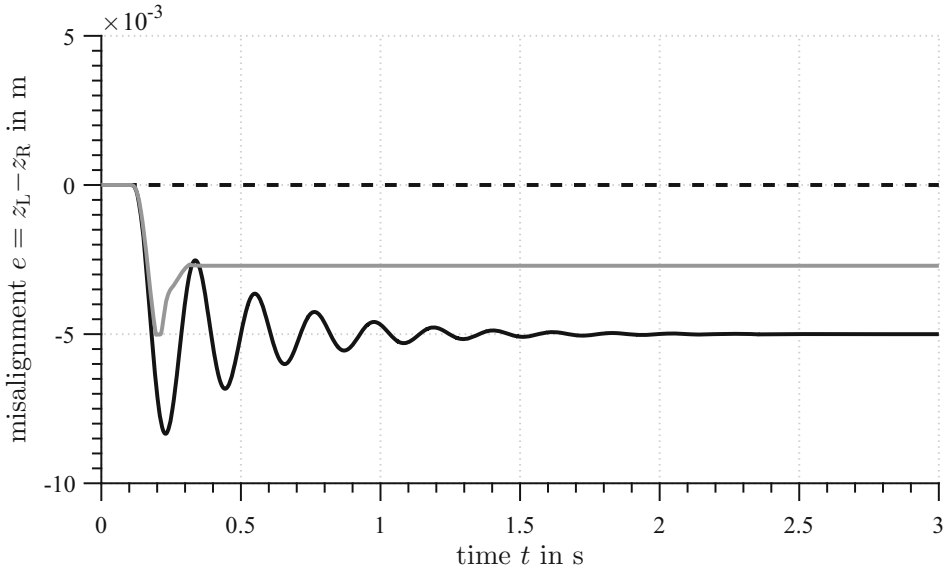


Fig. 14.5 Numerical simulation of misalignment $e = z_{\mathrm{L}-z_{\mathrm{R}}}$ for cases (i) (*black dashed line*), (ii) (*black solid line*) and (iii) (*gray solid line*)

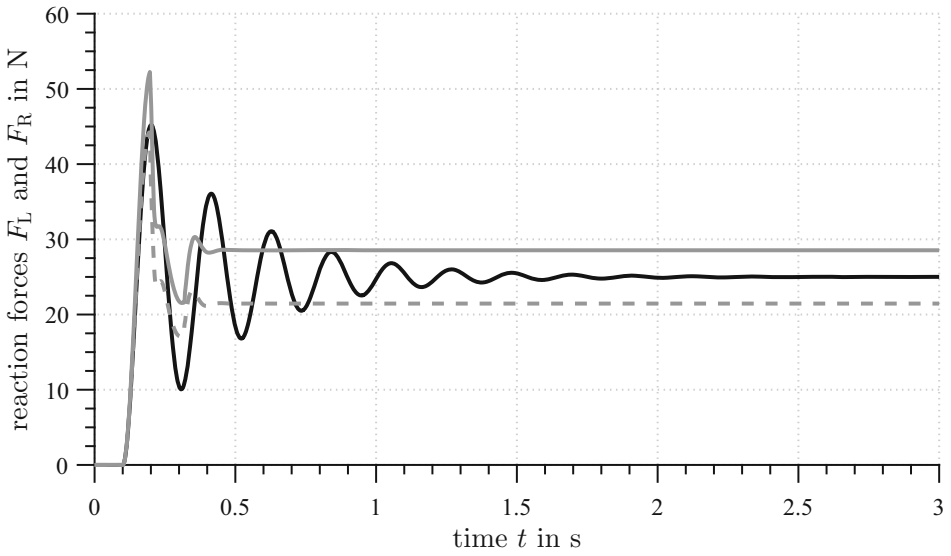


Fig. 14.6 Numerical simulation of the support reaction forces for case (i) $F_{\mathrm{L}} = F_{\mathrm{R}}$ (not plotted), case (ii) $F_{\mathrm{L}} \approx F_{\mathrm{R}}$ (*black solid line*) and case (iii) F_{L} (*gray dashed line*) and F_{R} (*gray solid line*)

(iii) the reaction forces are $F_{\mathrm{L}} \geq F_{\mathrm{R}}$ for $0 \leq t \leq 3$ s. The excitation load is not evenly split to both supports L and R anymore due to semi-active change of the load path between the transmission point of the excitation force and the supports.

Beside the ability of load path adaption in case (iii), oscillations are damped out faster compared to case (ii). According to the prescribed aim of the semi-active guidance elements, the load is shifted from support R to support L.

Figure 14.7 shows the semi-active induced guidance element forces $F_{\mathrm{ge,L}}$ and $F_{\mathrm{ge,R}}$. For the case

(i) the guidance element forces are $F_{\mathrm{ge,L}} = F_{\mathrm{ge,R}} = 0$ from time $t = 0$–3 s and, therefore, are plotted in Fig. 14.7,
(ii) the guidance element forces are again $F_{\mathrm{ge,L}} = F_{\mathrm{ge,R}} = 0$ from time $t = 0$–3 s and, therefore, are plotted in Fig. 14.7,
(iii) the guidance element forces are $F_{\mathrm{ge,L}} \geq 0$ and $F_{\mathrm{ge,R}} \leq 0$ from time $t = 0$–3 s.

The guidance element forces $F_{\mathrm{ge,L}}$ and $F_{\mathrm{ge,R}}$ are equal to 0 if it is not possible to shift load to reduce control deviation e or if control deviation e, respectively, misalignment is equal to 0, cf. Fig. 14.5.

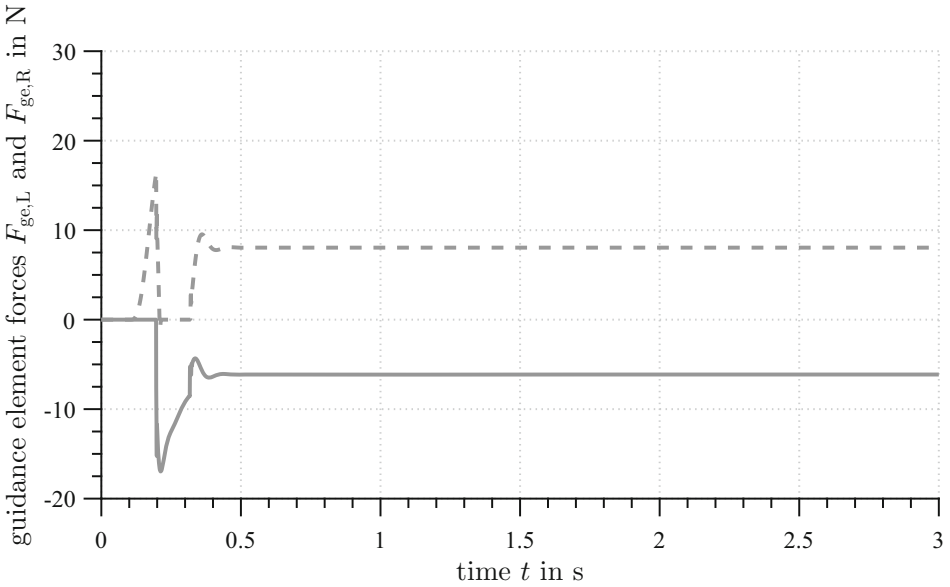


Fig. 14.7 Numerical simulation of the guidance element forces $F_{ge,L}$ (*gray dashed line*) and $F_{ge,R}$ (*gray solid line*) for case (iii)

For a semi-active approach the controller output for each guidance element depends, on the one hand, on the control deviation according to (14.14), and, on the other hand, on the relative movement between the mass m_A and the connection points of the guidance elements at the beam. When the mass m_A and the connection points are moving towards each other, e.g. for $0.1\,s < t < 0.2\,s$, the left guidance element is activated with $F_{ge,L} > 0$ to shift load towards support L. The right guidance element can not shift load towards support L during this movement and $F_{ge,R} = 0$. When the upper mass m_A and the connection points are moving away from each other, e.g. for $0.2\,s < t < 0.3\,s$, the right guidance element is activated with $F_{ge,R} < 0$ to shift load towards support L. The left guidance element can not shift load towards support L during this movement and $F_{ge,L} = 0$. If there is no movement between the mass m_A and the connection points for $t > 0.4\,s$, both guidance elements have potentially the ability to shift load towards support L depending on the previous movement.

14.5 Conclusion

This paper shows numerically the potential of semi-active auxiliary kinetic guidance elements in a 2D two mass oscillator to adapt the load path according to uncertain stiffnesses of the supports. The stiffness change may represent changing health condition of the elastic supports. It can be stated that semi-active kinetic guidance elements have the ability to provide an alternative load path in a load-bearing system and can be used to shift load from one structural element to another. The guidance elements become semi-active by using a friction brakes with controlled normal forces in the middle joints of the guidance elements. Friction is modeled by the LuGre friction model and a PID-controller is used as control strategy. If one support is assumed to have lower stiffness than the other, e.g. due to damage, load path adaption is used to reduce misalignment of the beam by shifting the load away from the support with less stiffness towards the support with remaining higher stiffness. The stiffness difference leads to misalignment of the beam if the load path is not controlled. With the semi-active guidance elements the misalignment of the beam can be reduced and thereby uncertainty is controlled. It is not possible to totally cancel out misalignment because of the semi-active approach without feeding large amount of energy into the system. In future work, the mathematical model will be more detailed e.g. the actuator behavior and a non-dimensional model will be derived to speed up simulation time. Moreover, different control strategies will be tested to improve the performance of shifting load. Finally, the results of this paper will be verified with experimental results.

Acknowledgements The authors like to thank the German Research Foundation (DFG) for funding this project within the Collaborative Research Center (SFB) 805 “Control of Uncertainties in Load-Carrying Structures in Mechanical Engineering”.

References

1. Bendsøe, M.P., Ben-Tal, A., Zowe, J.: Optimization methods for truss geometry and topology design. Struct. Optim. **7**, 141–159 (1994)
2. Rajan, S.: Sizing, shape, and topology design optimization of trusses using genetic algorithm. J. Struct. Eng. **121**, 1480–1487 (1995)
3. Farajpour, I.: Constrained optimization of structures with displacement constraints under various loading conditions. Adv. Eng. Softw. **41**, 580–589 (2010)
4. Marhadi, K.S., Venkataraman, S., Wong, S.A.: Load redistribution mechanism in damage tolerant and redundant truss structure. Struct. Multidiscip. Optim. **44**, 213–233 (2011)
5. Gaul, L., Nitsche, R.: Vibration control by interface dissipation in semi-active joints. Z. Angew. Math. Mech. **80**, 45–48 (2000)
6. Gaul, L., Hurlebaus, S., Wirnitzer, J., Albrecht, H.: Enhanced damping of lightweight structures by semi-active joints. Acta Mech. **195**, 249–261 (2008)
7. Joghataie, A.: Active control of trusses under heavy static loads. Neural Netw. World **11**(3), 285–292 (2001)
8. Lemaitre, C.: Topologieoptimierung von adaptiven Stabwerken (engl. Topology Optimization of adaptive Trusses). Ph.D. thesis, University Stuttgart, Dept. Civil and Environmental Engineering (2008)
9. Teuffel, P.: Entwerfen adaptiver Strukturen - Lastpfadmanagement zur Optimierung tragender Leichtbaustrukturen (engl. Design of adaptive structures - Load path management to optimize load-bearing lightweight structures). Ph.D. thesis, University Stuttgart, Dept. Civil and Environmental Engineering (2004)
10. Unger, A.F.: Serientaugliche quadratisch optimale Regelung für semiaktive Pkw-Fahrwerke (engl. Series-ready quadratic optimal control for semi-active passenger car chassis). Ph.D. thesis, Technische Universität München (2012)
11. Do, A.L., Spelta, C., Savaresi, S., Sename, O.: An lpv control approach for comfort and suspension travel improvements of semi-active suspension systems. IEEE Conf. Decis. Control **49**, 5560–5565 (2010)
12. Platz, R., Stapp, C., Hanselka, H.: Statistical approach to evaluating active reduction of crack propagation in aluminum panels with piezoelectric actuator patches. Smart Mater. Struct. **20**, 085009 (2011)
13. Gehb, C.M., Platz, R., Melz, T.: Approach to prevent locking in a spring-damper system by adaptive load redistribution in auxiliary kinematic guidance elements. In: Proc. SPIE 9433, Industrial and Commercial Applications of Smart Structures Technologies, San Diego , 8–12 March 2015, vol. 9433, pp. 94330G–94330G-9
14. Bedarff, T., Hedrich, P., Pelz, P.: Design of an active air spring damper. In: Proceedings 9. IFK, vol. 3 (2014)
15. Schaeffner, M., Götz, B., Platz, R.: Active buckling control of a beam-column with circular cross-section using piezoelastic supports and integral LQR control. Smart Mater. Struct. **25**(6), 065008 (2016)
16. Götz, B., Schaeffner, M., Platz, R., Melz, T.: Lateral vibration attenuation of a beam with circular cross-section by a support with integrated piezoelectric transducers shunted to negative capacitances. Smart Mater. Struct. **25**, 095045, 10 (2016)
17. Canudas de Wit, C., Olsson, H., Åström, K.J., Lischinsky, P.: A new model for control of systems with friction. IEEE Trans. Autom. Control **40**(3), 419–425 (1995)
18. Berger, E.: Friction modeling for dynamic system simulation. Appl. Mech. Rev. **55**(6), 535–577 (2002)
19. Visioli, A.: Practical PID Control. Springer, London (2006)

Chapter 15
Assessment of Uncertainty Quantification of Bolted Joint Performance

Nedzad Imamovic and Mohammed Hanafi

Abstract In this paper an investigation of uncertainty quantification of bolted-joint beams is presented. Series of very high frequency measurements were carried out on three simple beams, one continuous beam, one fully spliced beam with structural joint consisting of two connecting plates and bolts, and one continuous (unspliced) beam with the same connecting plates and bolts. Frequency Response Functions were measured in high frequency range for all beams under varying conditions of torque applied to bolts in the structural joints. Model validation of continuous beam is carried out by comparing predicted and measured FRFs, and limit of accuracy of measured FRFs is derived by considering uncertainty quantification of various sources of uncertainty from testing. Sensitivity of measured FRFs to varying torque of structural bolts is investigated and directly compared to variability of measured FRFs related to measurement accuracy. Roughness of surfaces that are in contact at structural joints is measured and considered as source of uncertainty with discussion of characterization of bolted joint as deterministic or stochastic model, with an attempt to identify sources of uncertainty quantification in both testing and modelling.

Keywords Uncertainty quantification • FRF variability • Bolted joints performance • Model updating • Modal testing

15.1 Introduction

Uncertainty Quantification (UQ) is a process of characterisation and reduction of uncertainties in theoretical predictions and experimental measurements according to Ralph Smith [1]. Uncertainties are equally present in theoretical predictions and experimental measurements. Theoretical predictions uncertainties originate from two sources, and these are: (1) discrepancies and errors in models used to predict behaviour of systems, and (2) input uncertainties which come from uncertain input parameters, boundary and initial conditions, and excitation functions. Experimental uncertainties can be described as two types, and these are: (1) incomplete or limited data, and (2) limited accuracy or resolution of measured data.

15.1.1 Uncertainties in Theoretical Predictions

Predicting dynamic behaviour of continuous system is primarily implemented through application of finite element methodology to model real life structures. Continuum systems are discretised into finite elements with assumed known displacement field function (shape function), which is applied to derive stiffness element between connected degrees-of-freedom (topology and types of elements dependent) in the discretised model resulting in the governing equation of motion for the discrete system of the form

$$[M]\{\ddot{x}(t)\} + [C]\left\{\dot{x(t)}\right\} + [K]\{x(t)\} = \{f(t)\} \tag{15.1}$$

where $[M], [C], [K] \in R_{n,n}$ are mass, damping and stiffness matrices respectively.

N. Imamovic (✉)
Independent Consultants Group, 30 Nursery Road, London, N14 5QB, UK
e-mail: nimamovic@hotmail.com

M. Hanafi
Department of Civil, Environmental and Geomatic Engineering, University College London, London, WC1E 6BT, UK

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_15

Theoretical models can be used to find solutions for dynamic response depending on the excitation function $\{f(t)\}$. If the excitation function is a pure harmonic excitation function $\{f(t)\} = \{F\}e^{i\omega t}$ at a single degree-of-freedom then response of a linear structure at an arbitrary degree-of-freedom will be a pure harmonic function at exactly the same frequency $\{x(t)\} = \{X\}e^{i\omega t}$. The ratio of amplitudes and phases of response and excitation for all known frequencies is known as Frequency Response Function (FRF) between response (j) and excitation (k) degree-of-freedoms $\alpha_{j,k}(\omega) = \left(\frac{X_j}{F_k}\right), F_m{=}0, m{=}1, 2 \ldots n, m \neq k.$

In order to calculate FRFs values it is necessary to transfer the Eq. (15.1) above into frequency domain and then find the inverse of a system matrix in the form

$$[\alpha(\omega)] = \left([K] + i\omega[C] - \omega^2[M]\right)^{-1} \tag{15.2}$$

The inversion of full system matrix as given in Eq. (15.2) at successive frequencies can be prohibitively expensive for systems that have large degrees-of-freedoms. If, however, normal modes solution of Eq. (15.1) is found then it is possible to calculate FRF between two degrees-of-freedom using the following expression

$$\alpha_{j,k}(\omega) = \sum_{r=1}^{n} \frac{(\phi_{j,r})(\phi_{k,r})}{\omega_r^2 - \omega^2 + i\eta\omega_r^2} \tag{15.3}$$

The above summation will be equivalent to Eq. (15.2) if all normal modes are calculated (which is never the case in reality for large systems), but in case where only a limited number of modes are calculated then Eq. (15.3) would only represent an approximation of Eq. (15.2), with larger discrepancies away from resonances in FRF.

Uncertainties in theoretical model can be applied as uncertainties on mass, damping and stiffness matrix elements as well as uncertainties in terms of discretisation and topology of the continuum model, resulting in a different topology of the system matrices depending on the boundary conditions and structural joints.

15.1.2 Uncertainties in Experiments

These characterisations of uncertainties in experimental measurements can be recognised as limited sampling frequency of measured time domain signals and resolution of analogue-to-digital (ADC) conversion of measured data by data acquisition system. At the present time, the digital resolution of analogue-to-digital conversion of general oscilloscope is 8 bits, although there are some oscilloscopes which use 12 bit ADC but these are very expensive and suffer from low sampling rate due to slow processing speed. 8 bit ADC digitiser has only $2^8 = 256$ values to represent complete measurement range and it might seem that this number of values is not sufficient to accurately reproduce measured signal. However, the main purpose of oscilloscope is to visually represent measured signal and resolution of digitised signal is deliberately compromised against slow processing speed which would be completely unacceptable. The higher the resolution of ADC converter the longer it takes to process and digitise a measured signal, which means that sampling frequency is inversely proportional to accuracy resolution. Data Acquisition Systems (DAQ) at present time generally use 24 bits as resolution accuracy at around 200 kHz sampling frequency. The processing time of this system is measured in seconds and this kind of processing speed would not be acceptable for interactive devices such as oscilloscopes as it would take too long to process and visually display the signal. 24 bit ADC converter has $2^{24} = 16{,}777{,}216$ number of values to represent the measured range.

Measurement accuracy is also limited by accuracy resolution and limitations of the sensor in terms of measurement range.

Damage Detection, Structural Health Monitoring and Dynamic Response Variability are in process of transition from theoretical methods into practical applications with increased number of practical applications either in Non-Destructive Testing (NDT) or Vibrations (Modal Analysis). NDT Testing requires high sampling frequencies but it generally works with small excitation energies in the continuum domain, whereas vibration testing generally requires lower sampling frequencies but at higher energies of the whole structural dynamic behaviour.

Use of variability of measured dynamic data from structures due to different loading and environmental conditions is limited by measurement accuracy of measured dynamic data. The two types of measurement uncertainties mentioned above, limited or incomplete data and limited accuracy of measured data, which translate into two main bottlenecks identified as sampling frequency and resolution of analogue-to-digital converters, are the principal factors which will affect the use of variability of measured dynamic data for the purpose of damage detection and structural health monitoring.

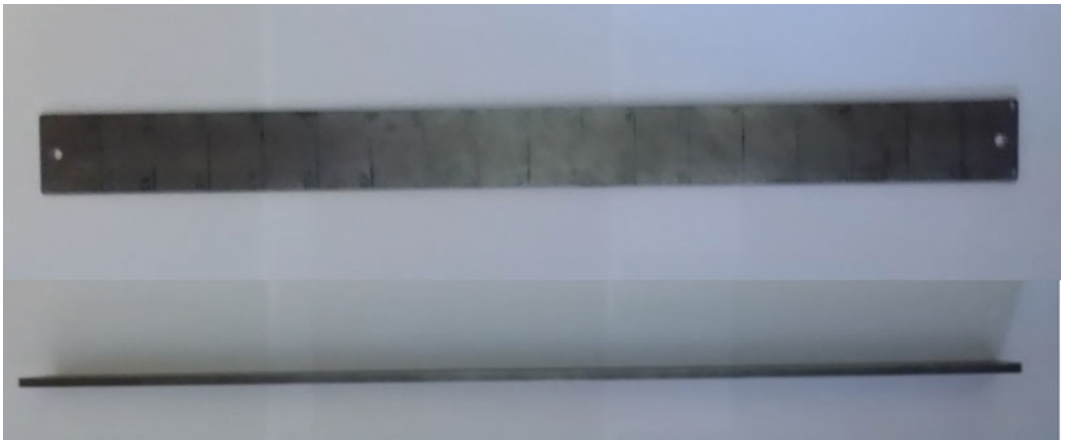


Fig. 15.1 Simple beam

Practical engineering structures will exhibit changes in measured dynamic data due to changes in loading, changes in environmental conditions and changes in material properties and structural joints. The real difficulty in application of variability of measured dynamic data is to distinguish and separate the effects of those different causes of variability of measured dynamic data.

15.1.3 Uncertainties Comparison Between Theoretical Predictions Using Experimental Data

The overwhelming conclusion from the above discussion is that neither theoretical predictions nor experimental measurements can produce completely accurate and reliable data. Both theoretical predictions and experimental measurements can only be verified between them in order to assess the degree of correlation between the two data sets, a process known as model validation using test data. In order to practically demonstrate validation of FRF for theoretical predictions and experimental measurements, a single continuum structure, consisting of a simple beam as shown in Fig. 15.1 was modelled and experimentally tested in order to generate predicted and measured FRF, respectively, which are compared in Fig. 15.2.

Close inspection of correlation between theoretically predicted and experimentally measured FRFs shown in Fig. 15.2 initially leads to conclusion that there is overall trend of both curves following each other, more closely at lower frequencies with discrepancies between the curves increasing at higher frequencies. The actual frequency range is 12 kHz, which is considered almost beyond frequency range of piezo accelerometers. It can be seen that both resonances and antiresonance follow each other on both curves for frequencies up to 5 kHz, whereas for higher frequencies resonances between two FRFs seem to be better correlated than antiresonances. Poor correlation at antiresonances at higher frequencies is likely to be due to limited number of mode shapes taken for calculation of theoretically predicted FRFs (Eq. 15.3).

Measured FRF has clear visible signs of noise, particularly at antiresonances and at higher frequencies. The reason for this is that signal-to-noise ratio at antiresonant frequencies is significantly higher than at other regions where there is larger amplitude of vibrations. This problem progressively increases at higher frequencies because excitation of higher frequencies is significantly lower than excitation at lower frequencies.

The measurements are always contaminated with noise even if other uncertainties are somehow either eliminated or reduced. Impact hammer testing is sensitive to variability of exact impact location during testing, which can never be guaranteed to be the same. Each impact force is significantly different for different impact. The hammer tip determines the excitation frequency range and the amount of energy transferred to structure, with soft tips (such as rubber and plastic) of the hammer generally exciting low frequencies with high energy and hard tip of the impact hammer (metal) generally exciting significantly higher frequency range but with lower energy. Use of hard tips of impact hammer also increase the likelihood of having double or multiple impacts as the structure responds with speed of elastic wave propagation (which for steel is about $c_{wave} = \sqrt{\frac{E}{\rho}} \approx 6000\ \frac{m}{s}$, where E and ρ are Young's Modulus and Density, respectively). An example of double hit effect is shown in Fig. 15.3. The original signal was processed using rectangular window to remove noise and majority of double hits.

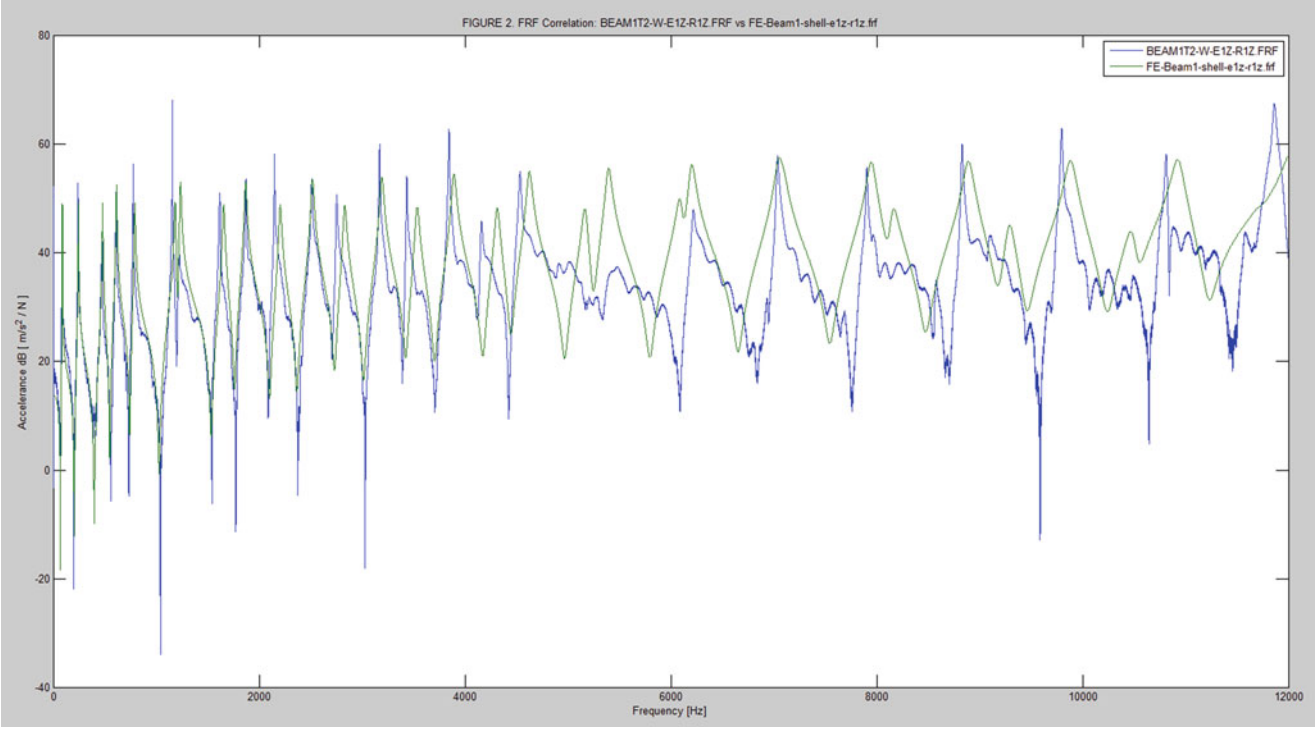


Fig. 15.2 Comparison between measured and predicted FRFs for a single continuous beam

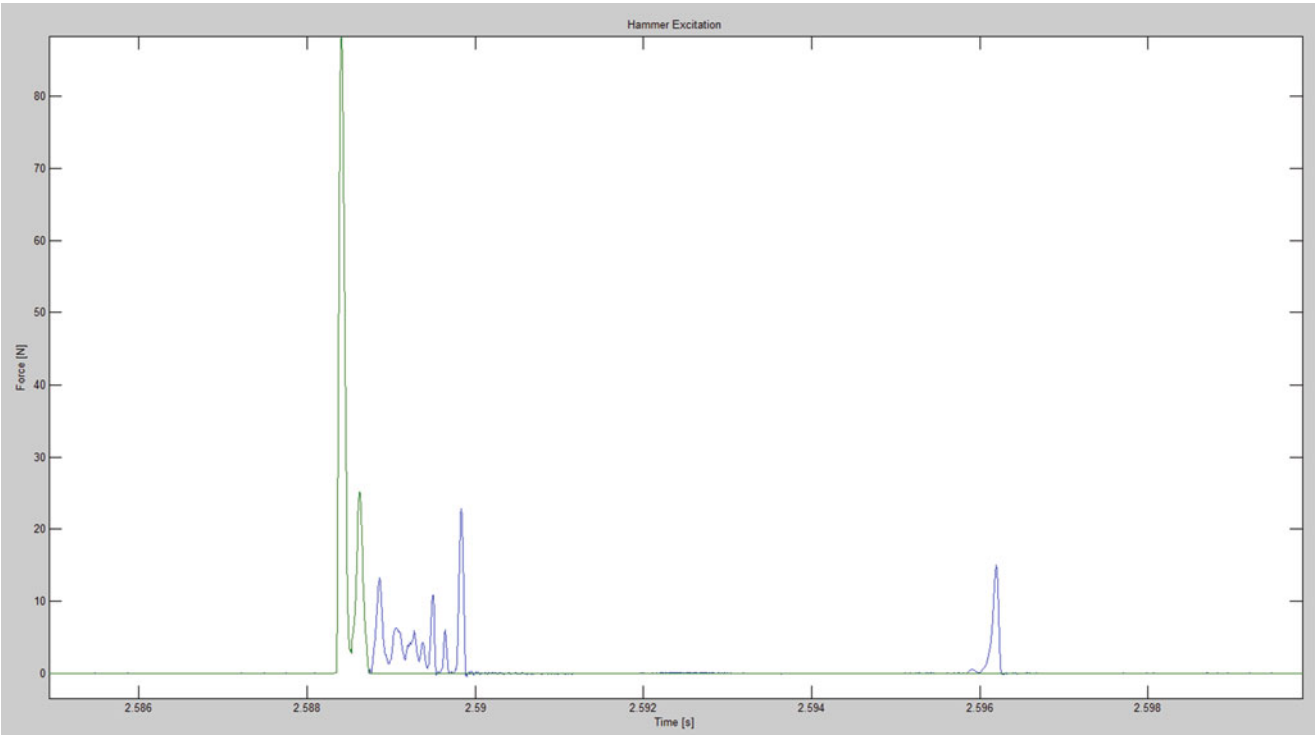


Fig. 15.3 Double hit effect using impact hammer (*blue* is raw signal and *green* is windowed signal)

Figure 15.4 shows the time domain signals from Fig. 15.3 in frequency domain. The sampling frequency of raw transient signal was carried out at 216 kHz resulting with Nyquist frequency of 108 kHz for signal in frequency domain shown in Fig. 15.4. It can be seen from Fig. 15.4 that lower frequencies were excited with higher energies that higher frequencies.

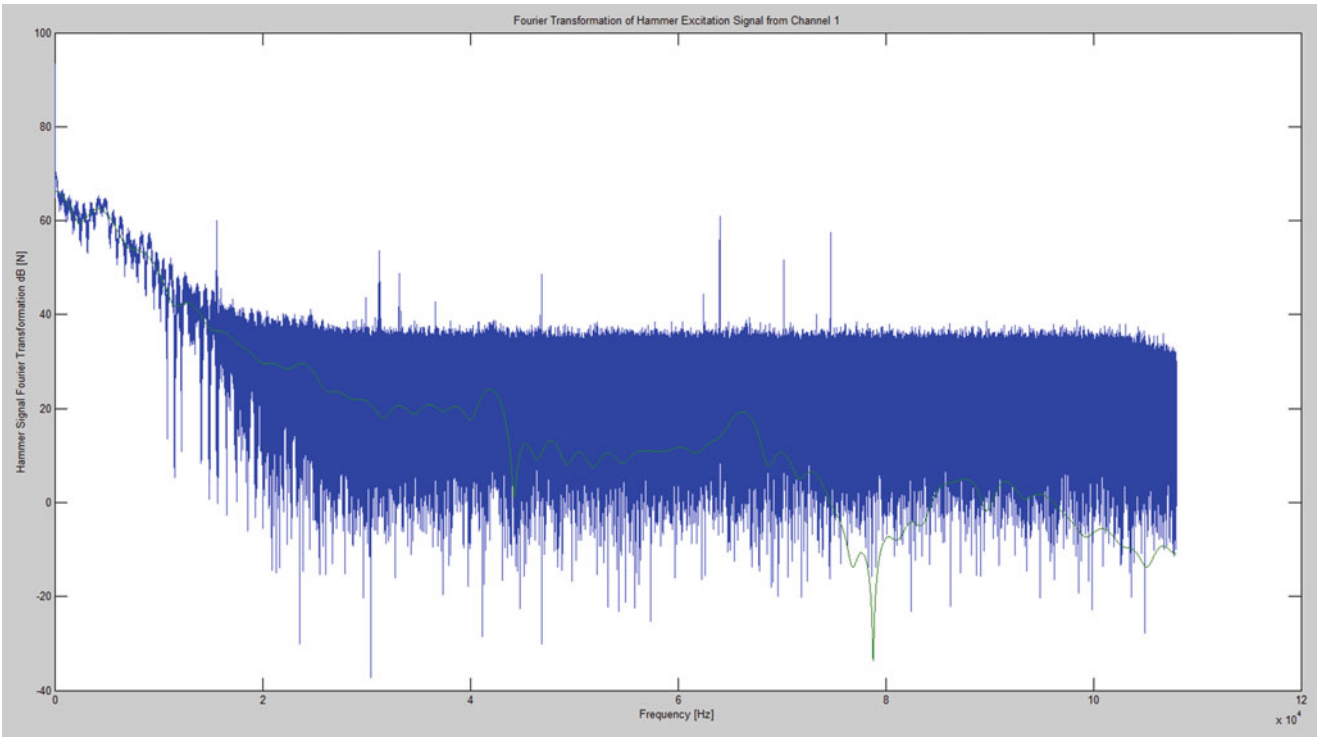


Fig. 15.4 Combined effect of double hit and noise in frequency domain (blow raw signal)

This explains partly poor correlation between measured and predicted FRFs at higher frequencies. The raw signal in frequency domain in Fig. 15.4 shows two trends of variability, one almost pure harmonic trend (in frequency domain!) which can be attributed to presence of double hits, which after being removed resulted in green curve which did not show the harmonic-like variability (in frequency domain!). The second variability trend is clear link between noise is time domain (which is invisible and unnoticeable in time domain in Fig. 15.3) with noise on the transferred signal in frequency domain (which is highly visible particularly at higher frequencies).

The comparison between estimated FRFs using raw data and windowed data is shown in Fig. 15.5 below. The applied windowing consists of removal of noise at frequencies where signal-to-noise ratio was low, and removal of majority of double hit effects as shown in Fig. 15.3.

The combined effect of double hit effect and random noise presence in frequency domain can be seen in Fig. 15.5. It can be concluded that both presence of random double hits during testing, and/or contamination of signal in time domain with random noise results in considerable variability of measured FRFs.

15.1.4 Introduction of Variation of Uncertainties in Testing

In order to demonstrate variability of FRFs due to input parameters which are partly controllable, the beam from Fig. 15.1 was spliced and connected using connecting plates as shown in Fig. 15.6.

Validation of predicted and measured FRFs for spliced beam is presented in Fig. 15.7, from which it can be seen that there was significant deterioration of correlation for the case of spliced beam when compared to validation of a single continuous beam. The reduction in correlation can be attributed to introduction of uncertainties due to presence of splice. Spliced beam is not a single continuum and this increases the complexities and uncertainties for both theoretical predictions and measurements resulting in higher variability of measured and predicted FRFs.

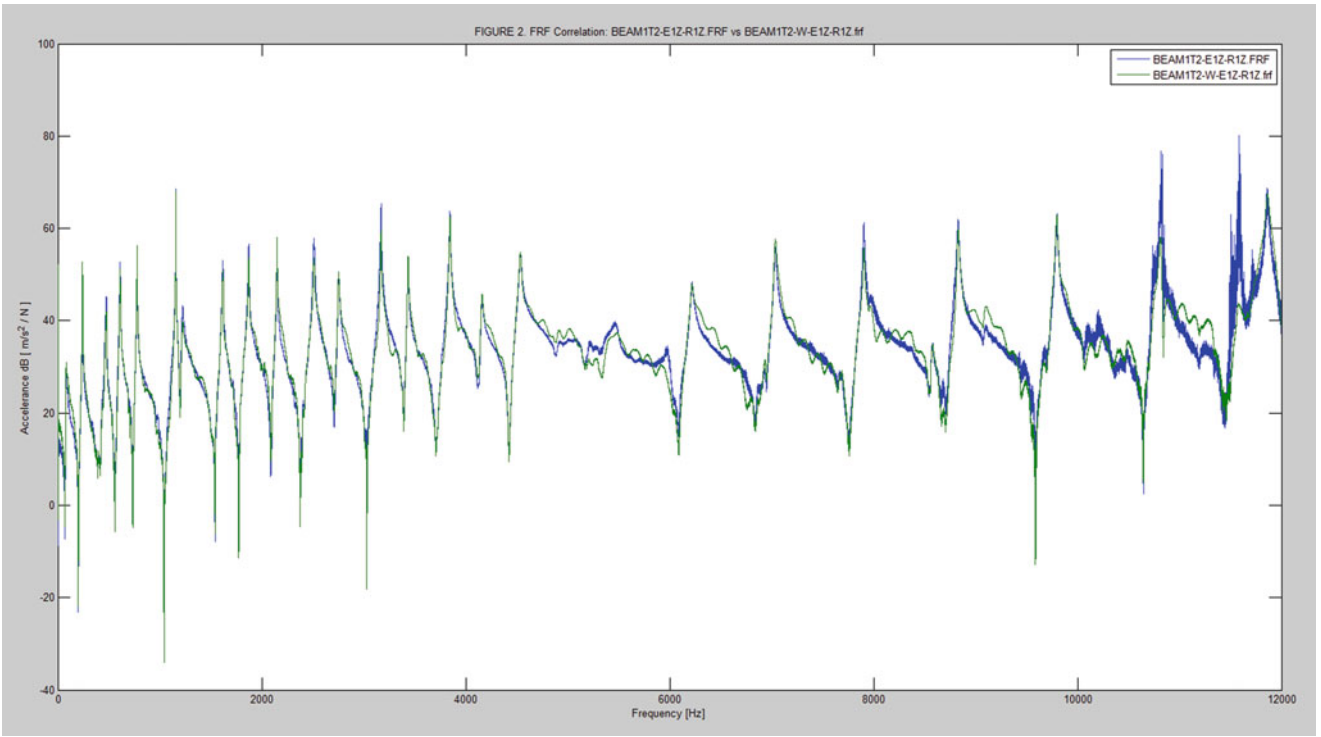


Fig. 15.5 Comparison of two estimated FRFs using the same raw data (*blue* is raw data, *green* is windowed data)

Fig. 15.6 Spliced beam

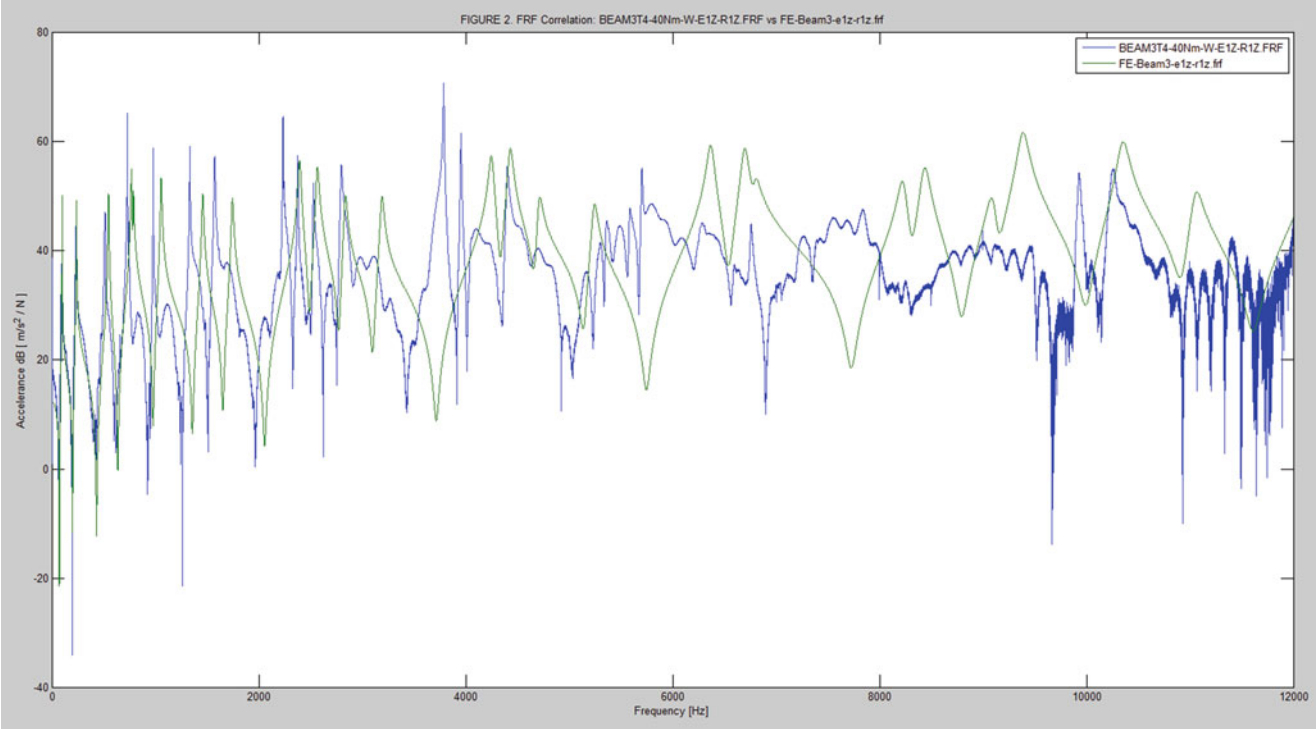


Fig. 15.7 Validation of predicted and measured FRFs for spliced beam

15.1.5 Control of Uncertainty by Varying Torque Applied to Structural Bolts

Situation can be further complicated by changing one of the uncertain parameters of spliced beam, such as torque setting of structural bolts, from 10 Nm to 40 Nm to check what kind of FRF variability this torque variation is causing. Controlling the torque that was applied to structural bolts of the spliced beam will change contact pressure between connecting plates and beams. Higher torque will translate into higher tension of the bolts which will cause high contact pressure across the contact surface. High torque will also result in larger residual compressive stress of the connecting plates and beams in the vicinity of the structural bolts. The results of this investigation are shown in Figs. 15.8, 15.9, 15.10, 15.11, and 15.12, starting from 12 kHz frequency range and reducing down to a single resonance zoomed frequency range. It can be seen from Figs. 15.8, 15.9, 15.10, and 15.11 that there is significantly more noise on FRFs of spliced beams across all frequencies and particularly at higher frequencies. The reason for this is that the introduction of the splice itself significantly increased the friction in the structure compared to single continuum beam, and the amount of energy from impact hammer was not sufficient to excite these higher frequencies, thus resulting in subdued transient response of the structure with high signal-to-noise ratio in transient signal which then propagated into noise in measured FRFs.

Further close observations of individual resonant frequencies reveals that there is almost consistent trend of increase of resonant frequency with increase of torque at structural bolt. This higher torque applied to structural bolts results in stiffer joint which then results in higher resonant frequency. This increase is not fully consistent for all mode shapes as different more shapes would have different deformation level and strain energy in the splice. There is also a visible trend of change of response amplitude for different torque levels, which can be explained by change in frictional energy dissipation at contact surfaces for different torques.

15.1.6 Limits of Control of Uncertainties

There are practical limits in terms of how much control of the uncertainties is possible in real applications. Some of the mentioned testing uncertainties can be controlled and reduced by using shaker which is attached to a fixed point and can

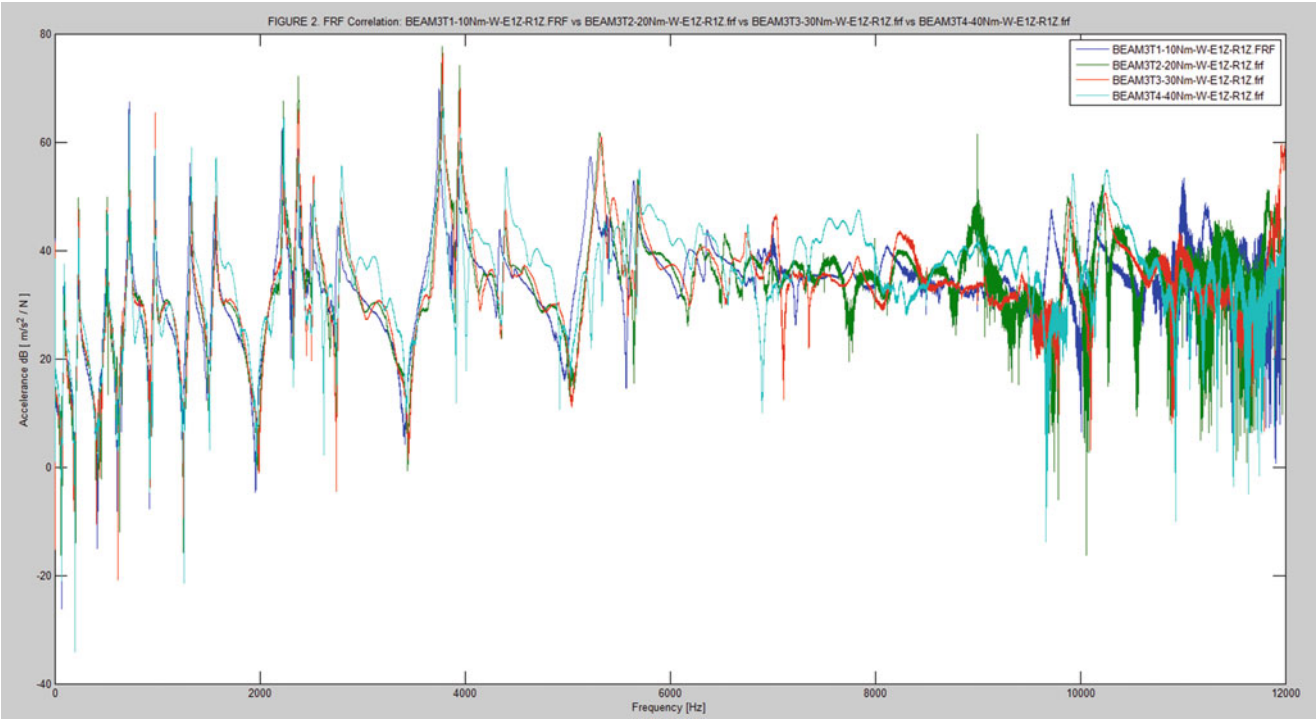


Fig. 15.8 FRF Variability due to change in torque applied to structural bolts (12 kHz frequency range)

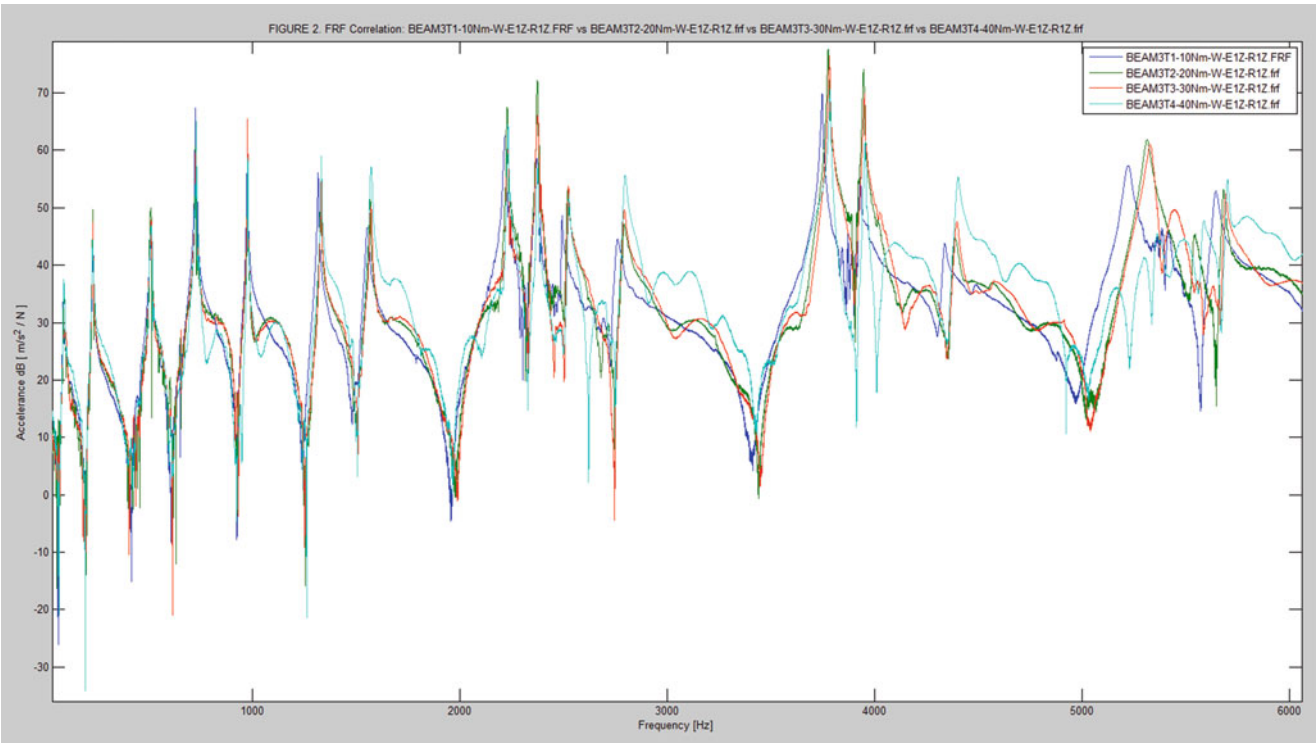


Fig. 15.9 Variability due to change in torque applied to structural bolts (6 kHz frequency range)

reproduce the same excitation signal. In this case, if other conditions such as boundary and environmental conditions during testing are kept closely constant, one would assume that repeated measurements would result in almost identical FRFs, but this is not the case in practice. Measured FRFs are contaminated by noise which was added to measured signal in time

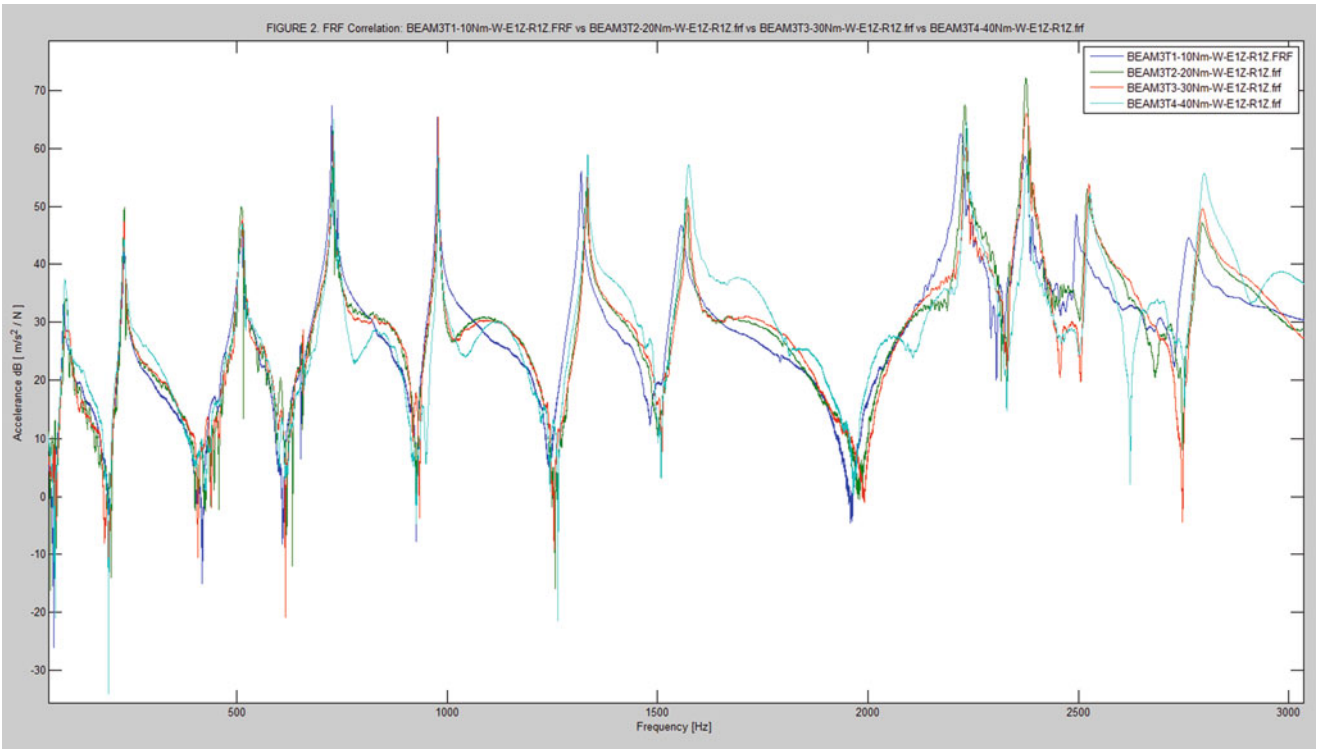


Fig. 15.10 Variability due to change in torque applied to structural bolts (3 kHz frequency range)

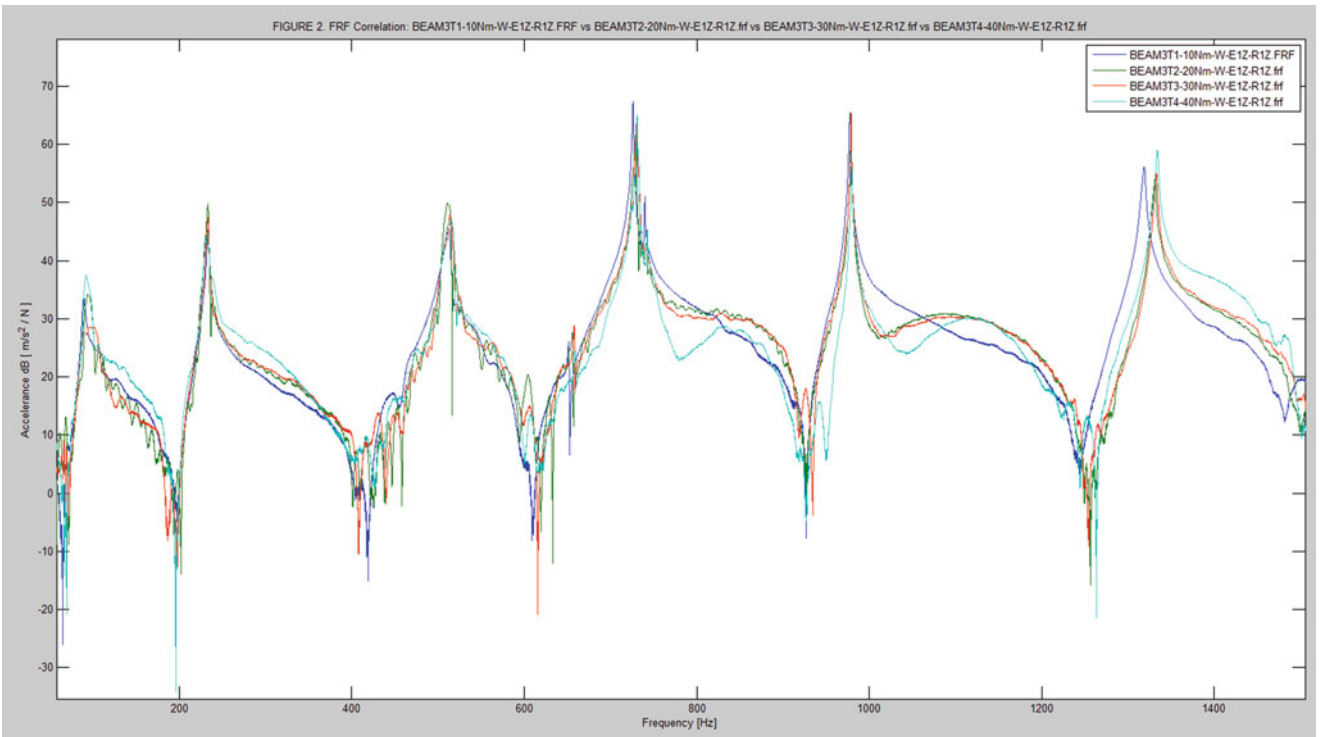


Fig. 15.11 Variability due to change in torque applied to structural bolts (1.5 kHz frequency range)

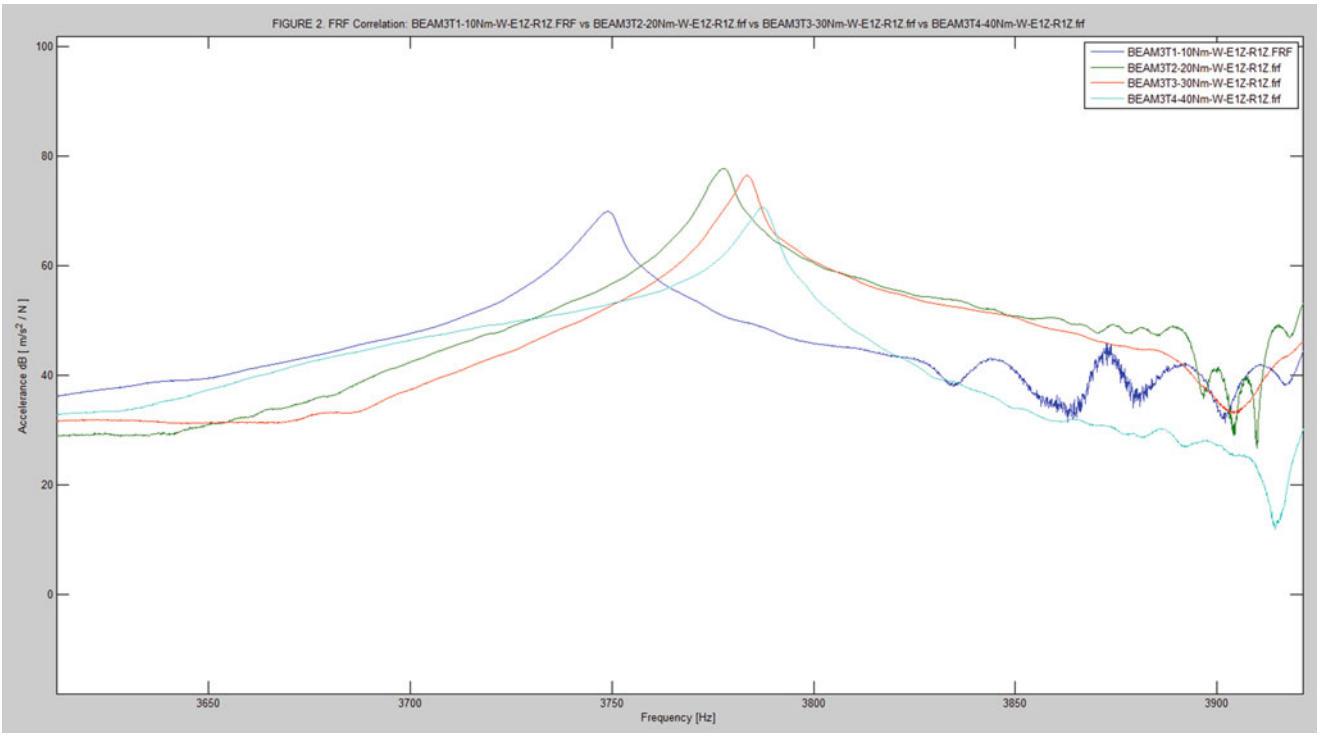


Fig. 15.12 Variability due to change in torque applied to structural bolts (zoomed frequency range)

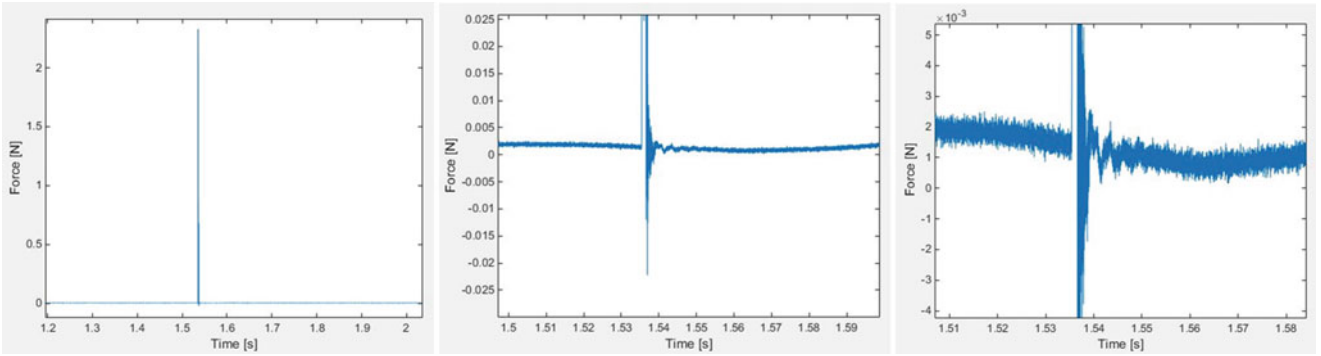


Fig. 15.13 Actual noise in the measurements (zoomed views)

domain and this noise propagates into FRFs estimate. For this reason it is impossible to reproduce measured FRFs with exact accuracy, and any repeated measurements of FRFs will result in variability of measured FRFs, depending on the actual number and level of uncertainties present during a testing.

Actual noise recorded in the transient signal cannot be visually seen unless zoomed view of transient signal is produced, as given in Fig. 15.13, which shows actual measurements and noise in the transient signal.

Statistical analysis of the noise isolated from impact signal reveals that mean value of the actual noise is 0.0018 of unscaled units, with standard deviation of 0.00022232 value. The noise distribution seems to follow 'Normal or Gaussian Distribution' which is expected considering that this noise is generated through some sampling of error propagations. Although the actual noise parameters as recorded are relatively insignificant to maximum value of the signal recorded, which is about 2.5 N, (the standard deviation of the actual noise compared to maximum transient signal recorded is $\frac{0.00022232}{2.5} x100\% = 0.009\%$), the uncertainty propagation due to noise is significant in FRFs estimation mainly because Discrete Fourier Transformation (DFT) is highly sensitive to discontinuities of transferred functions.

Theoretically predicted FRFs also suffer from non-repeatability. For instance, two different finite element models, one using shell and other using solid elements, will result in different FRFs. Even if the type of elements used is kept constant, two different meshes will result in variability of predicted FRFs. The thickness of the beam at Fig. 15.1 will vary at different locations and its material properties might not be homogenious and isotropic throughout the domain of the structure. All these uncertainties will also propagate and result in variability of theoretically predicted FRFs.

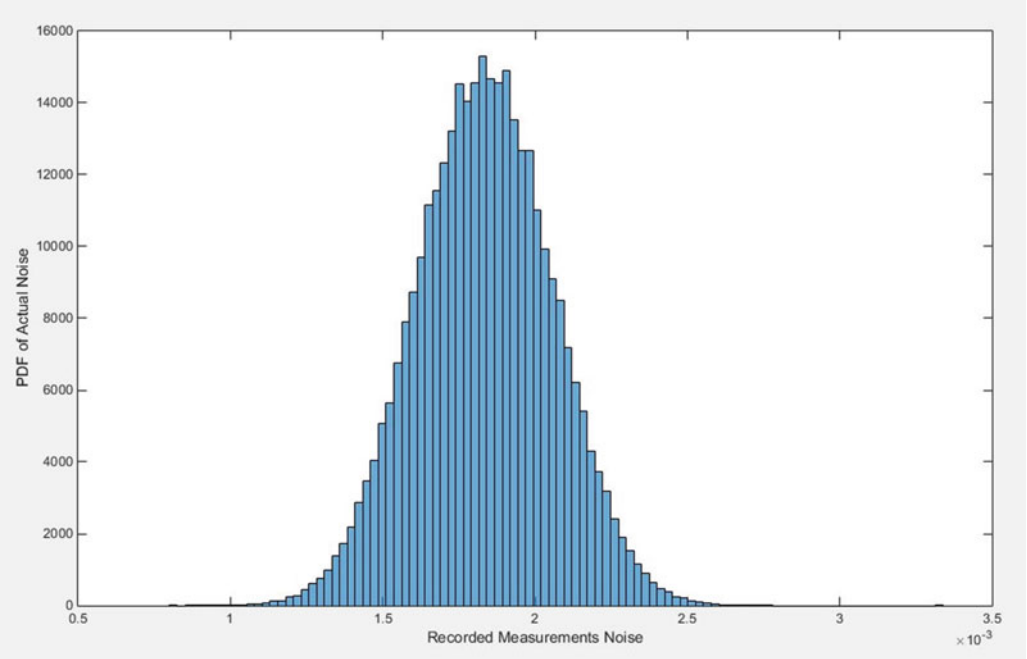


Fig. 15.14 Histogram of the actual noise in the measurements

Natural question that arise from this discussion about validation between predicted and measured FRFs is how variability of either measured or predicted FRFs can be assessed and assigned to particular source of uncertainty. The overwhelming conclusion of the uncertainty propagation in both experimental and theoretical parts is that underlying level of uncertainty can only be reduced but never completely removed in practice, which then leads to conclusion that both measured and predicted FRFs will have certain degree of uncertainty that need to be understood and as much as possible controlled.

Estimation of FRFs can be done using so-called H1 and H2 estimators, which allow elimination of noise from one single channel, through use of convolution of excitation and response signals. There is also a procedure which allows elimination of noise on both signals but only using random excitation vibration testing and with minimum of two simultaneous channel measurements, and this estimator of FRF is known as H3 estimator (see Ewins (2000)). However, for a single impact transient measurement in single channel response, noise will always be present at least on one channel.

15.2 Literature Review

FRF variability is an important integral part of structural health monitoring process as summarised by Claus-Peter Fritzen [2] in his review of vibration-based structural health monitoring concepts and applications. One of the most comprehensive literature reviews of Structural Health Monitoring was conducted by Sohn et al. [3] of Los Alamos National Laboratory in which 'Data Normalisation' was identified (page 7) as a procedure "to normilize data sets so that signal changes caused by operational and environmental variations of the systemcan be separated from structural changes of interest, such as structural deterioration and degradation". Specific review of literature survey of inverse methods applied to damage identification by Friswell [4] also identifies (page 404) "environmental and other non-stationary effects" as a difficult aspect. Since then various papers had been written on variability of FRFs due to changes of system masses Adhikari et al. [5] where variability of FRFs was measured for changing masses. In this study several small masses were attached at random location to a clamped beam and the structure was excited by a shaker in order to measure FRFs for variable random masses distributions. It is assumed in this study that all other parameters were kept constant and that uncertainty of FRFs was only due to randomness of masses, but there was no a single plot of repeatability of FRFs for the same mass distribution (see conclusions at page 490). Uncertainty quantification of FRFs was also studies by Mao and Todd [6] by analysing error propagation from the uncertain measurements into estimated FRFs. FRF uncertainty was also studied in terms of polynomial chaos expansion by Manan and Cooper [7] and they verified polynomial chaos expansion with Monte Carlo simulation.

15.3 Conclusions

This study was originally inspired as investigation of results published by Adhikari et al. [5] in which measured FRF variability study was conducted on the basis that "*Special measures have been taken so that the uncertainty in the response* ***only arise*** *from the randomness in the mass and oscillator locations . . .* " as claimed by the authors. Adhikari et al. [5] claimed in the publications that they were able to reproduce measurement FRFs with complete accuracy so that variability of FRFs was only due to uncertainty in the mass and oscillator locations. However, Adhikari et al. [5] never published repeatability of measurement FRFs, which would had put their claim on test. This paper investigates whether at all and to what extent it is possible to estimate FRFs accurately from measured dynamic data, (i.e. to achieve repeatability of measured FRFs). The main objective of this paper is to investigate the limits of uncertainty of measurement FRFs in order to test the claim of repeatability of measured FRFs by Adhikari et al. [5]. The conclusion is that Adhikari et al. [5] only measured and observed combined effect of uncertainty in mass and actual noise present during the measurement.

One possible advancement taken from this study is to develop a procedure that will allow to estimate the isolated and combined effect of known actual measured noise during the experiment. Noise can be easily measured during the experiment and statistically assessed but unfortunately it cannot be removed from the actual measurements. This study will attempt to assist in understanding the propagation of noise and their contribution in FRF variability.

Studying and understanding of uncertainty propagation in FRFs is more complex than it might seem. Uncertainties originate from several sources and it is almost impossible to isolate uncertainties propagation from various sources.

Experimental methods of measuring FRFs have several sources of uncertainties, but generally FRFs are estimated from measured transient signals which always have noise on them that propagates as uncertainty into FRFs. For this reason all measured FRFs have certain amount of noise on them which cannot be removed.

Some attempts had been made to subdue propagation of uncertainties from some sources but these attempts are not convincing. Apart from noise, measuring transient signals is susceptible to uncertainties due to impact position and direction of the excitation force, variation in boundary conditions, variation of initial conditions during testing, nonlinearities of structures, variation of performance of structural joints in the structures, variability of unknown processes of energy dissipation and so on.

The damping or energy dissipation plays a crucial role in studying uncertainty propagation. Transient response of a structure is highly sensitive to damping and energy dissipation in the structure. Damping and energy dissipation generally comes from material damping, viscous damping caused by fluid trapped in voids of rough surfaces in structural joints, viscous damping of surrounding fluid where structure is vibrating, friction due to macro and micro slip in structural joints, plastic deformation, sound and heat generation and other effects such as fretting and similar.

The general thinking methodology in uncertainty propagation is to assume validity of the theoretical model and statistically study uncertainty propagation by varying mass, stiffness and damping properties of the theoretical model. Another approach is to conduct experiments in which certain isolated parameters are varied and measure variability of FRFs. Both approaches have flaws in that theoretical models, particularly damping models, tend to approximate damping and energy dissipation through a simple crude coefficient often neglecting other sources of energy dissipation and damping. Experimental approach tend to rely on assumption that somehow effects of other sources of uncertainties are removed or subdued and that measured variability in FRFs is assumed to be due to varying parameters in experiments. Variability in structural joints is generally completely neglected whereas it is known that structural joints are single most significant sources of uncertainty.

It has been demonstrated in this paper that structural joint are significant source of uncertainty propagation in measured and predicted FRFs. Measured FRFs were shown to be sensitive to torque of structural bolts in structural joints even though theoretical model assumes continuum at the structural joint. The real question that this paper is proposing is whether structural joints can be represented as deterministic models, i.e. whether the performance of structural joints is generally predictable. Structural joints are in reality contacts of two rough surfaces with general roughness or surface topology being unknown with also unknown relative position in contact. This in itself represents sufficient uncertainties to cause uncertainty propagation into measured FRFs which theoretically is currently not modelled and little understood.

The recommended way forward in terms of research is to research sources of energy dissipation in the real model, classify energy dissipation into material or structural damping, viscous damping, friction due to macro and micro slips, energy dissipation due to sound and heat conversion and verify or validate models with real test data.

The primary conclusion is that various scientific data that had been published with claims that the sources of uncertainty were isolated should be carefully considered and concluded that recorded variability is resulting effect of several sources of uncertainties rather than single isolated uncertainties.

References

1. Smith, R.C.: Uncertainty Quantification: Theory, Implementation and Applications. SIAM, Philadelphia, PA (2014)
2. Fritzen, C.-P.: Vibration-based structural health monitoring–concepts and applications. Eng. Mater. **293–294**, 3–20 (2005)
3. Sohn, H. et al.: A review of structural health monitoring literature: 1996–2001, Los Alamos National Laboratory, LA-13976-MS (2004)
4. Friswell, M.: Damage identification using inverse methods. Phil. Trans. R. Soc. A. **365**, 393–410 (2007)
5. Adhikari, S., Friswell, M.I., Lonkar, K., Sarkar, A.: Experimental case studies for uncertain quantification in structural dynamics. Probab. Eng. Mech. **24**, 473–492 (2009)
6. Mao, Z., Todd, M.: Statistical modelling of frequency response function estimation for uncertainty quantification. MSSP. **38**(2013), 333–345 (2013)
7. Manan, A., Cooper, J.E.: Prediction of uncertain frequency response function bounds using polynomial chaos expansion. J. Sound Vib. **329**(2010), 3348–3358 (2010)
8. Davies, I.C., Wicks, A.L.: Sources of FRF variation in damage detection. M.Sc. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA (2002)
9. Friswell, M., Penny, J.E.T.: Crack modelling for structural health monitoring. SAGE. **12**, 0139–0148 (2002)
10. Friswell, M., Penny, J.E.T.: Is damage location using vibration measurements practical. In: EUROMECH 365 International Workshop, DAMAS 97, Sheffield, UK (1997)

Chapter 16
Sensitivity Analysis and Bayesian Calibration for 2014 Sandia Verification and Validation Challenge Problem

Ming Zhan, Qin-tao Guo, Lin Yue, and Bao-qiang Zhang

Abstract As the wide use of modeling and simulation in many fields, the credibility of model is of great importance to predict the model response. In this paper, taking the 2014 Sandia V&V Challenge problem for example, some detail works were done. Firstly, a parametric model was established and influential parameters were selected based on global sensitivity analysis. Bayesian updating was conducted and the uncertain parameters were identified. Uncertainty propagation was investigated and all kinds of uncertainties were quantified based on the corresponding experimental results, and the credibility of model was assessed. Then, a relationship between displacement and strain was constructed and responses were predicted based on the validated model.

Keywords Model validation • Parameter identification • V&V challenge problem

Modeling and simulation technology has been the third pillar of science and has been widely used in aerospace, mechanical, and civil engineering, etc. The 2014 Sandia Verification and Validation (V&V) Challenge Problem is a typical example that actually exist in industrial practice. The challenge problem focused on the overall strategies to address the detail aspects of challenge problem and also fostered discussion about how these strategies used in a V&V activity can establish or assess the credibility of a specific prediction. The Challenge Problem contains three main aspects which are uncertainty parameters identification, uncertainty quantification and validation assessment. This paper mainly focuses on the part of model validation and presents some approaches that can be used to address these challenges.

Firstly, a synthesis model of tanks under different load conditions is established to represent the relationship between input parameters and model output results. Model parameters in the challenge problem are dimensional sizes, material properties, working loads, and mesh size and model output response are displacement and stress at different locations. In general, the established model differs from the real world more or less. In order to catch the parameters that have significant influences on the model output, a global sensitivity analysis is conducted in the design space of all possible parameters. Then, for the sake of narrowing the gap between simulation and the corresponding experimental measurements, the model updating approach based on Bayesian inference theory is adopted to calibrate the model parameters used in the initial model.

Secondly, the selected parameters that have obvious impact on the model output are regarded as uncertain. All of these uncertain model parameters are divided into random uncertainty, epistemic uncertainty or mixed uncertainty categories. According to the characteristics of each kind of uncertain parameters, sampling points are generated utilizing the corresponding sampling methodologies. Uncertainties in the input model parameters are propagated through the updated synthesis model and the stochastic characteristics are obtained by data processing. The kernel concept of model validation is the process that evaluating the ability of model, to which extent, can represent the real world. Model validation of tank displacement is conducted by comparing the model output and the released experimental data, and then credibility of tank model is estimated based on validation metrics.

However, in the 2014 V&V Challenge problem, the quantity of interests are stress at some special locations while the quantities that can be tested are displacements at the corresponding locations. In order to evaluate model credibility of the predicted stress using the model which is validated using displacement, a bridge that describes the relationship between displacement and stress is needed. Surrogate model is a favorable tool which combines experimental design and approximation methodology, and is commonly used in model updating and model validation. The Latin hypercube

M. Zhan (✉) • Q.-t. Guo • L. Yue
College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
e-mail: zhanming@nuaa.edu.cn

B.-q. Zhang
College of Aeronautics and Astronautics, Xiamen University, Xiamen 361005, China

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_16

sampling and radial basis function approximation approaches are investigated and surrogate models which represent the relationship between displacement and stress are constructed. The premise to use a surrogate in an engineering practice is that the model has sufficient accuracy. Determine coefficient and relative root mean square errors are adopted to evaluate the precision of surrogate in this paper. Prediction of stress can be achieved utilizing the surrogate models which satisfy precision requirements. Finally, the ability of model to predict stress is validated by the measurements offered by the workshop.

Acknowledgement This research has been supported by the National Natural Science Foundation of China (Grant No. 51275240).

Chapter 17
Non-probabilistic Uncertainty Evaluation in the Concept Phase for Airplane Landing Gear Design

Roland Platz and Benedict Götz

Abstract Predicting the kinematic and dynamic behavior of complex load bearing structures with high safety requirements such as landing gears is time consuming. For that, mathematical analytic, finite element or multi body surrogate models are needed for numeric simulation purposes. Today, these models take into account both deterministic and non-deterministic approaches. However, before adequate and verified simulation begins, the modeling of the mathematical surrogates requires most of the time for adequate prediction, including model verification, before even more costly experimental testing phase begins. This contribution investigates an approach based on INFO-GAP analysis to predict critical performance requirements of major landing gear design alternatives in an early design stage. This analysis uses only simple analytical but comparable and sufficient adequate models for four major design concept alternatives according to basic design rules found in relevant literature. The concepts comprise one telescopic and three different trailing link designs. It is the aim to make decisions in selecting the most suitable design as early as possible in the design stage with taking into account uncertainty—before time consuming efforts in modeling finite element and multi body models for detailed prediction are conducted. Particularly, the authors evaluate the robustness to uncertainty or how much of an uncertainty horizon by means of uncertain compression stroke ability due to varied stiffness properties can be tolerated with the four different concepts, until the absolute maximum allowable compression stroke limit is reached. This contribution continues the authors' prior work presented at IMAC 2016. In there, the authors evaluated and compared the performance requirements like compression stroke ability and ride quality, elastic force retention, structure strength, and weight of mechanisms for main and nose landing gears resulting from the four significant structural design concepts in mathematical physical models in an analytic deterministic way.

Keywords Landing gear • Concept evaluation • Uncertainty • INFO-GAP analysis • Decision making

17.1 Introduction

Today, numerical simulation strategies to predict the dynamic behavior comprise deterministic and statistical methods. According to [11], *all* required methods are available even in commercial codes and for sophisticated structures, geometries and materials in aerospace application. The actual challenge is caused by the high efforts of modeling set-up, management of the large amount of required input data, the management of different simulation approaches and material models over the required frequency range as far an appropriate experimental approach for model verification is concerned [11]. In the prior work [12], the authors took a first step to clarify pros and cons of different design concepts of landing gears with respect to major and typical performances requirements. It was the aim to evaluate and compare the performance requirements compression stroke ability and ride quality, elastic force retention, structure strength, and weight of mechanisms for main and nose landing gears resulting from four significant structural design concepts in mathematical physical models according to basic design rules found in [3] and [10] in an analytic deterministic way. The authors only took a first glimpse on a simple uncertainty evaluation for each concept, relating defined stiffness worst case variations in the elastic strut to the loading and compression stroke ability. The focus of the paper mainly was, though, to introduce an adequate mathematical-physical model of each concept that allows, at all, the adequate and consistent performance comparison—but without uncertainty evaluation for decision making.

R. Platz (✉)
Fraunhofer Institute for Structural Durability and System Reliability LBF, Bartningstraße 47, 64289 Darmstadt, Germany
e-mail: roland.platz@lbf.fraunhofer.de

B. Götz
System Reliability, Adaptive Structures, and Machine Acoustics SAM, Technische Universität Darmstadt, Magdalenenstraße 4, 64289 Darmstadt, Germany

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_17

Energy absorption and compliance properties are major features when evaluating the performance of landing gears in aerospace application. From the beginning of the twentieth century, airplane landing gears experience major enhancements with the invention of the telescopic oleo-pneumatic shock absorber for non retractable landing gears [1]. Different designs for non-retractable and retractable landing gears have been improved for optimal shock absorbing potential as well as enhancing strength and light-weight design for high quality and the fulfillment of allowances in aerospace industry [8]. Today, simple non retractable leaf-spring steel or composite designs for small airplanes and retractable oleo struts with and without trailing links for main and nose landing gear for small and large aircraft lead to a great variety of design concepts. Relevant specifications for landing gear design are sufficient absorption of kinetic energy and overall compliance to prevent damaging impact as well as adequate steering, stability and ride quality on the ground, high reliability, low cost, overall aircraft integration, airfield compatibility, compact design for storage, low weight etc., [4–6].

Until today, optimal damping properties of the oleo strut and optimal structural reliability and fatigue life as well as new ways of conceptional design are aspired and documented in several contributions from an academic point of view, e.g. [13–16]. However, only a few contributions discuss the pros and cons of principal landing gear concepts and give useful recommendations for the engineer who he has to decide between the concepts and oppose their requirements and benefits. For that, the engineer has to make early choices in geometry, kinematics and dynamics, e.g. whether the landing gear should have a trailing link or not. If a trailing link is in favor, further decisions have to be made concerning the position and stroke capability of the oleo strut. Yet, the position and compression stroke ability of the oleo strut affects the strength of the surrounding and supporting struts. For example, ride quality on uneven ground may be one of many criteria. The trailing link landing gear offers good ride quality since the kinematics allow larger deflection of the overall mechanism [7]. However, the cost is higher weight.

In [12], the authors proposed an approach to discuss the pros and cons for the four major landing gear design concepts

(a) telescopic design with fixed elastic strut and hinged upper/lower struts as torque links
(b) trailing link with fixed elastic strut and hinged upper/lower struts
(c) trailing link with hinged elastic strut, fixed upper strut, hinged lower strut, and high leverage
(d) trailing link with hinged elastic strut, fixed upper strut, hinged lower strut, and small leverage

in early design stage and, eventually, evaluated and compared the concepts to each other with respect to stroke ability and ride quality, elastic force retention, structure strength and total weight.

For the sake of decision making in early design stage, the authors assumed that properties from the detailed design do not alter the structural conceptional general performance [12]. In order to compare the four major landing gear design concepts (a) to (d), simplified mathematical models were introduced in [12] to ensure comparability as far as possible—knowing that usually detailed design approaches according to the guidelines in [3, 10] differ a lot and are unique for every airplane. Yet, it was the goal of that paper to prepare applicable and adequate deterministic models for early evaluation of uncertainty that comes along with different design concepts, *before* the designer gets lost in time consuming design details and *before* it might be too late to conduct major changes in the design. The mathematical models stayed simple, they all include a wheel, supporting struts and an elastic strut for elastic force retention, but arranged differently in common approaches that are included in the four major design concepts (a) to (d). However, wheel stiffness is neglected in this study.

Now, since simple but adequate deterministic mathematical modeling of the four concepts was completed in [12], in this present contribution, the four design concepts' uncertainty is evaluated using INFO-GAP analysis as a non-deterministic and non-probabilistic tool [2]. Particularly, the authors evaluate how much of an uncertainty horizon by means of uncertain absolute compression stroke ability due to varying stiffness properties can be tolerated with the four different concepts, until the absolute maximum allowable compression stroke limit is reached? How robust are the concepts against uncertainty? The tolerable uncertainty horizon is related to the total mass needed for each concept and the expected comfort or, respectively, ride quality for the different design approaches. This way, specific tolerated uncertainty will be evaluated to help decision making in selecting the most adequate design approach for landing gears in an early design stage.

17.2 INFO-GAP Theory

BEN-HAIM proposed the INFO-GAP theory that helps to quantify robustness of mathematical models based decisions despite lack of knowledge in the models [2]. Estimating the uncertainty horizon by means of uncertain absolute compression stroke ability due to varying stiffness properties for landing gear design in this present contribution takes four basic steps to apply INFO-GAP theory as shown in several other works, for example in [9] with respect to [2]. The four steps are extended by the authors for the present contribution by a specific robustness to uncertainty:

- derivation of an mathematical model $\mathcal{M}(\beta)$ with the vector β that comprises (i) model parameters such as material and geometry assumptions as well as properties like damping and stiffnesses, and (ii) state variables such as displacements, velocities, accelerations, forces and moments assumptions
- derivation of an uncertainty model

$$\mathcal{U}(\beta_0, h) \;=\; \left\{ \beta \;:\; \frac{|\beta - \beta_0|}{\beta_0} \;\leq\; h \right\}, \quad h \geq 0 \tag{17.1}$$

  taking into account assumed uncertainty in the vectors β around the vector's nominal entities β_0. The variation of β is limited by the uncertainty horizon h. A simple uncertainty model could be an interval $|\beta - \beta_0|$ related the nominal value β_0
- performance requirement

$$\mathcal{P}\Big(\mathcal{M}(\beta)\Big) \;\leq\; \mathcal{P}_{\text{crit}} \tag{17.2}$$

  to specify a critical level such as limit loads etc.
- robustness to uncertainty

$$\widehat{h} \;=\; \max\{h \;:\; \max \mathcal{P}\Big(\mathcal{M}(\beta)\Big) \;\leq\; \mathcal{P}_{\text{crit}}\} \tag{17.3}$$

  with the highest tolerable uncertainty horizon $\widehat{h}$.

17.3 Uncertainty Quantification in Landing Gear Design Concepts via INFO-GAP Approach

17.3.1 Guideline

For each landing gear design concept (a) to (d), the authors derive a mathematical model

$$\mathcal{M}(\beta) = z_{\text{a}}(p_x, v_x) \tag{17.4}$$

to calculate the absolute compression stroke progress $z_{\text{a}}(p_x, v_x)$ as a function of model parameters p_x such as density of material, geometric dimensions and other properties like stiffness k as well as state variables v_x such as static absolute loading F_{a} assumed to act on all models of the four landing gear design concepts (a) to (d). All model parameters and state variables are listed in [12]. Table 17.1 lists only a selection of model parameters and state variables that are relevant for the uncertainty evaluation in this present contribution.

Next, the uncertainty model

$$\mathcal{U}(k_0, h) \;=\; \left\{ k \;:\; \frac{|k - k_0|}{k_0} \;\leq\; h \right\}, \quad h \geq 0 \tag{17.5}$$

comprises only one varied model parameter in β: the stiffness k of the suspension rod, Sect. 17.3.2.

The performance requirement for the landing gear design,

$$\mathcal{P}\Big(z_{\text{a}}\big(p_x(k), v_x\big)\Big) \;\leq\; \mathcal{P}_{\text{crit}} = z_{\text{a, max}} \tag{17.6}$$

specifies the critical or maximum allowable absolute compression stroke $z_{\text{a, max}}$ as a result of increasing static loading F_{a}. For $z_{\text{a}}\big(p_x(k), v_x\big)$ in (17.6), only the stiffness k out of all model parameters and state variables listed in [12] is assumed to be uncertain and is varied according to (17.5). Of course, all other model parameters and state variables may vary too. To the authors' opinion, however, relative to varied stiffness of the suspension strut, the variation of material and geometric parameters can be neglected due to today's high precision and accuracy in manufacturing and assembling. Stiffness variations are considered to be more likely to occur due to influence of temperature, fatigue etc.

Eventually, the robustness to uncertainty

$$\widehat{h} \;=\; \max\left\{ h \;:\; \max\left(\left| z_{\mathrm{a}}\left(p_x(k), v_x\right) - z_{\mathrm{a,\,max}} \right| \right) \;\leq\; 0 \right\} \tag{17.7}$$

discloses the maximum uncertainty horizon $\widehat{h}$ for reaching the maximum allowable absolute static compression stroke $z_{\mathrm{a,\,max}}$ at static loading $z_{\mathrm{a}}\left(p_x(k), v_x\right)$ that will be tolerated for each design concept (a) to (d) when the elastic strut compliance or, respectively, stiffness k is uncertain according to (17.5).

17.3.2 Mathematical Modeling and Achieving Comparability Between the Concepts

Figure 17.1 shows real examples of the general four concepts (a) to (d) for main and nose landing gear as introduced in [12]. Their differences are based on different mounting conditions of the landing gear elements such as elastic strut and the supporting upper and lower struts or, respectively, torque links. These conditions lead to different landing gear mechanisms with their characteristic kinematics and resulting dynamics.

The concepts were simplified in [12] as schematic diagrams in Fig. 17.2 with the basic landing gear elements: elastic strut cylinder 1, elastic strut piston rod 2, upper supporting strut 3, lower supporting strut 4 or, respectively, upper torque link strut 3 and lower torque link strut 4 in (a), and wheel 5. The elastic strut's piston rod is directly connected with the wheel 5 in concept (a) or to the lower supporting strut 4 in (b) to (d). For comparative reasons, the four concepts have the same absolute installation length l_{a}, same vertical nominal static and maximum stroke ability $z_{\mathrm{a,\,stat}}$ and $z_{\mathrm{a,\,max}}$ of the wheel, and same nominal vertical static absolute load $F_{\mathrm{a,\,stat}}$ that represents the airplane's weight. In addition, the distance $x_{\mathrm{c,\,us}}$ between the elastic and upper struts support, the height h_{hinge} of the upper strut's hinge of upper strut for (b) to (d), and the distance d_{hinge} between the piston rod low end and the hinge of upper/lower strut are given for this study.

In addition to [12], the authors now investigate the uncertainty for the relation between the absolute static load F_{a} that leads to the resulting compression stroke $z_{\mathrm{a,\,stat}} \leq z_{\mathrm{a}} \leq z_{\mathrm{a,\,max}}$ due to uncertain stiffness k. As a performance design specification and constraint, the maximum compression stroke $z_{\mathrm{a,\,max}}$ can not be exceeded. The central question is: How much uncertainty of compliance variation or stiffness decrease for each concepts can be tolerated in order to fulfill the major requirement of not exceeding the maximum compression stroke due to static loads that are close to the static limit loads?

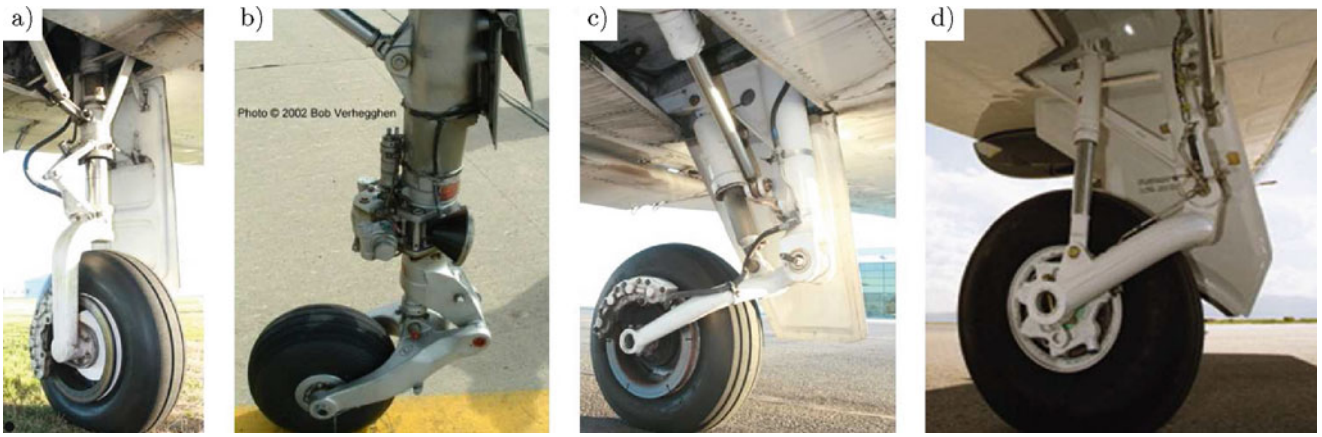


Fig. 17.1 Examples of general concepts for main and nose landing gear, (**a**) telescopic with fixed elastic strut and hinged upper/lower struts as torque links (from WIKIPEDIA), (**b**) trailing link with fixed elastic strut and hinged upper/lower struts, (from BOB VERHEGGHEN), (**c**) trailing link with hinged elastic strut, fixed upper strut, hinged lower strut, and high leverage (from WIKIPEDIA), (**d**) trailing link with hinged elastic strut, fixed upper strut, hinged lower strut, and small leverage (from PILATUS BUSINESS AIRCRAFT) [12]

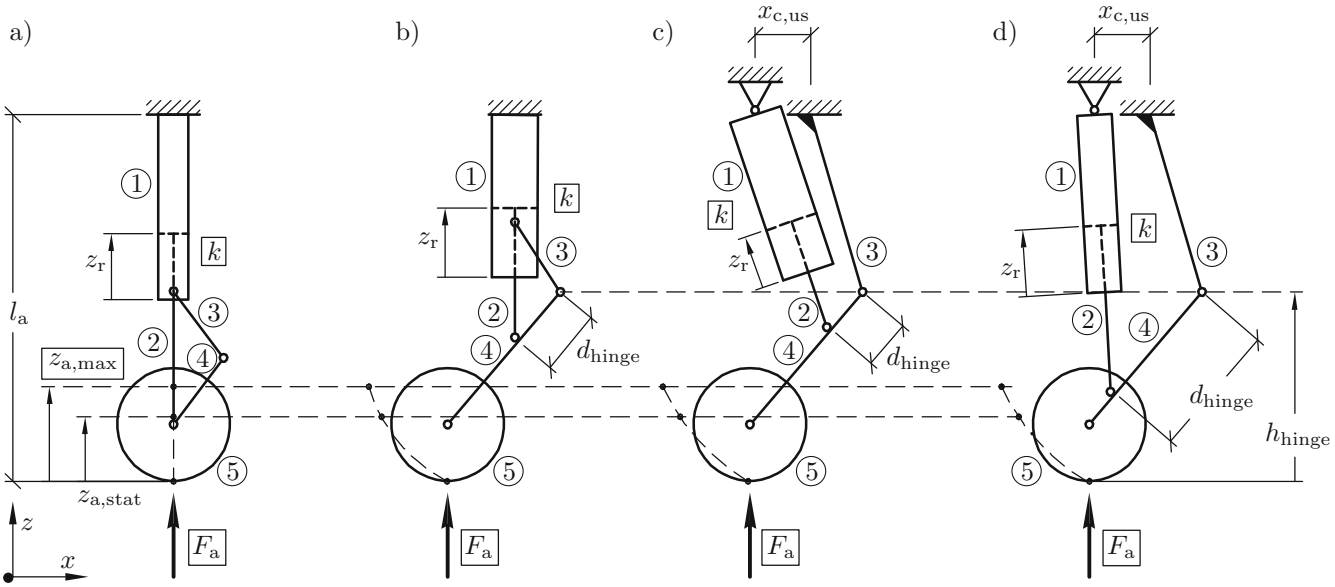


Fig. 17.2 Simplified general concepts according to [12] for main and nose landing gear with: (**a**) telescopic with fixed elastic strut (cylinder 1 and piston rod 2), hinged upper/lower struts 3/4 as torque links and wheel 5, (**b**) trailing link with fixed elastic strut (cylinder 1 and piston rod 2), hinged upper/lower struts 3/4 and wheel 5, (**c**) trailing link with hinged elastic strut (cylinder 1 and piston rod 2), fixed upper strut 3, hinged lower strut 4 with high leverage, and wheel 5, (**d**) trailing link with hinged elastic strut (cylinder 1 and piston rod 2), fixed upper strut 3, hinged lower strut 4 with small leverage, and wheel 5, [12]

17.3.2.1 Achieving Comparability Between the Concepts

Achieving comparability between different design concepts with simple models is only possible if simplifications do not oversimplify and, therefore, falsify the landing gear conceptual characteristics. The prior assumptions to achieve comparability between the concepts (a) to (d) with respect to materials, dynamics and geometry in are listed completely in [12]. Out of them, most relevant for this contribution's uncertainty evaluation are, according to Fig. 17.2:

- equal nominal static absolute axial or, respectively, nominal vertical load $F_{a,\,stat}$, referring to basic design guidelines in [3]
- equal total absolute installation length l_a as the distance between bottom wheel and elastic strut mount
- equal maximum allowable absolute vertical compression stroke ability $z_{a,\,max}$ of the mechanisms
- equal static absolute axial or, respectively, vertical nominal compression stroke $z_{a,\,stat}$ due to static airplane weight of the mechanisms

17.3.2.2 Selected Properties for Comparing the Concepts' Compression Stroke Capability Under Uncertainty

In [12], a complete list of material, dynamic and geometric properties for modeling the simplified concepts (a) to (d) was given. Table 17.1 shows a selection of the properties for comparing the concepts' compression stroke capability under uncertainty using the INFO-GAP approach. Relevant prior assumptions are listed in the first half, calculated relevant properties are listed in the second half in Table 17.1 due to the prior assumptions, including the resulting weight of the overall mechanisms. Of course, for implementation in a real airplane, further detailed design work will and must follow according to the guidelines such as in [3] and [10] to meet the manufacturer's and the regulating authorities' requirements.

The assumed linear elastic strut's stiffness nominal value becomes

$$k_0 = \frac{F_r}{z_r} \tag{17.8}$$

for each concept model (a) to (d) due to relative elastic strut loads F_r and strokes z_r, with the identity to absolute loads and strokes $F_r = F_a$ and $z_r = z_a$ in (a). The relation in (17.8) has been derived by adapting the relative load and stroke relation to the overall requirement of equal nominal static absolute vertical load $F_{a,\,stat}$ and equal nominal static absolute vertical stroke $z_{a,\,stat}$ that are assumed for all four concepts and listed in Table 17.1 and according to [12]. The rather complex modeling of the kinematic relations has been done numerically with finite element calculation and is not shown here.

Table 17.1 Dynamic, geometric and material properties as well as resulting weight for the simplified concepts (a) to (d) according to Figs. 17.1 and 17.2 to ensure comparability, density steel: $\rho = 7.850\,\mathrm{kg/m^3}$, material for all elements: D6AC 4335V, mass of wheel: neglected, n.s. = not specified [12]

Property	Variable	Unit	(a)	(b)	(c)	(d)
Prior assumptions to achieve comparability						
Load, nominal static absolute	$F_{\mathrm{a,\,stat}}$	kN	30.0	30.0	30.0	30.0
Height hinge of upper strut (b) to (d)	h_{hinge}	m	n.s.	0.450	0.450	0.450
Distance between mounting of cylinder and upper strut	$x_{\mathrm{c,\,us}}$	m	–	–	0.125	0.125
Distance piston rod low end to hinge upper/lower strut	d_{hinge}	m	n.s.	0.158	0.118	0.315
Compression stroke, maximum vertical static absolute wheel	$z_{\mathrm{a,\,max}}$	m	0.250	0.250	0.250	0.250
Compression stroke, nominal vertical static absolute wheel	$z_{\mathrm{a,\,stat}}$	m	0.170	0.170	0.170	0.170
Calculated properties due to prior assumptions						
Stiffness, elastic strut	k_0	kN/m	120.0	660.0	1260.0	241.2
Load, maximum absolute	$F_{\mathrm{a,\,max}}$	kN	44.118	40.410	37.328	42.000
Load, maximum relative elastic strut	$F_{\mathrm{r,\,max}}$	kN	44.118	85.085	119.289	52.095
Load, static relative elastic strut	$F_{\mathrm{r,\,stat}}$	kN	30.000	59.097	85.791	36.297
Stroke, maximum relative elastic strut	$z_{\mathrm{r,\,max}}$	m	0.250	0.129	0.095	0.216
Stroke, static relative elastic strut	$z_{\mathrm{r,\,stat}}$	m	0.170	0.090	0.068	0.150
Mass, total system	m	kg	19.82	32.96	53.00	29.49

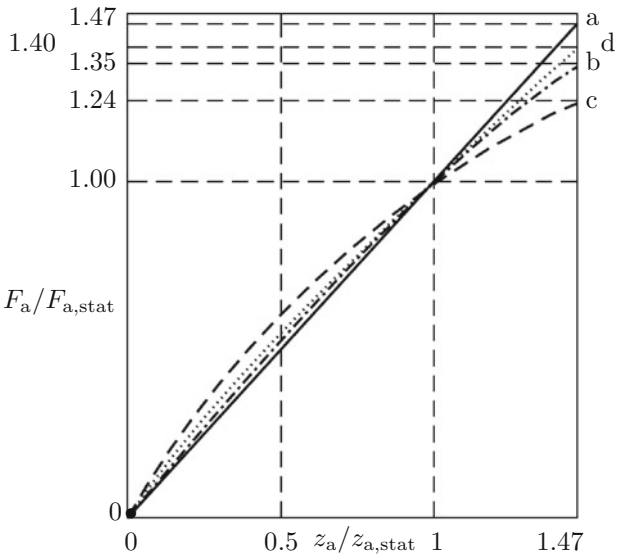


Fig. 17.3 Static load F_{a} vs. static compression stroke z_{a}, normalized to the nominal static absolute load $F_{\mathrm{a,\,stat}}$ and to the nominal vertical static absolute stroke $z_{\mathrm{a,\,stat}}$ for concept (a) *black solid line*, (b) *black dot-dashed line*, (c) *black dashed line*, and (d) *black dotted line*

17.3.2.3 Deterministic Comparison of Static Compression Stroke Behavior

In this section and at first, the authors repeat to evaluate and compare the distinctive compression stroke ability $z_{\mathrm{a}}(p_x, v_x)$ without varied stiffness k due to static loading F_{a} resulting from significant structural design concepts (a) to (d) from [12].

Figure 17.3 shows the progress of load vs. stroke relations F_{a} vs. z_{a} for the absolute vertical wheel stroke z_{a} for concepts (a) to (d), Fig. 17.2 [12]. The relations are normalized to the nominal static absolute load $F_{\mathrm{a,\,stat}}$ and nominal vertical static absolute stroke $z_{\mathrm{a,\,stat}}$. Uncertainty is not taken into account, the relations are derived from deterministic calculation. In case of the absolute load-stroke relation F_{a} vs. z_{a} of the overall mechanism, the relation in concept (c) is most nonlinear due to the high leverage effect. Concept (c) ensures smoothest ride quality of all, concept (a) leads to rather bumpy rides. In case the absolute stroke exceeds the absolute static stroke by the factor of 1.47 due to higher loading than the static load, the maximum allowable stroke is reached. Concept (a) allows highest absolute loading than concept (c) with lowest absolute loading capability.

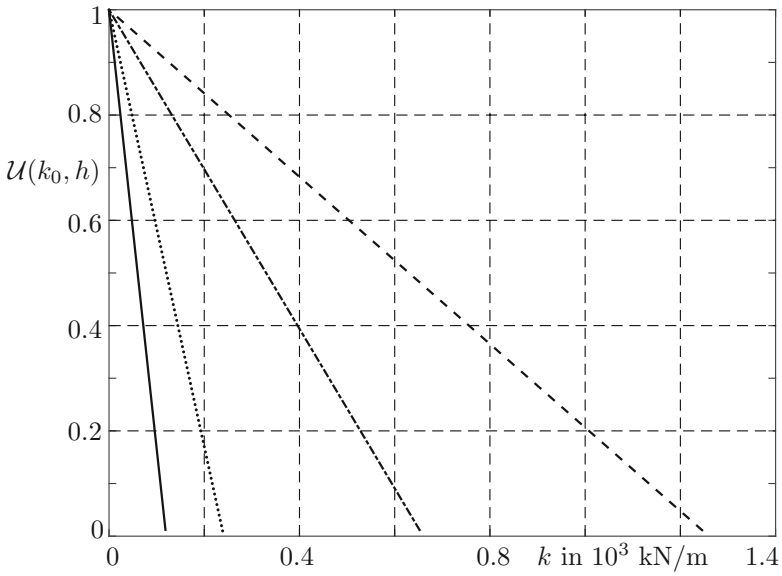


Fig. 17.4 Uncertainty for varied decreasing stiffnesses k when deviating from the nominal values k_0 according to Table 17.1 for each concept (a) *black solid line*, (b) *black dot-dashed line*, (c) *black dashed line*, and (d) *black dotted line*

17.3.3 Uncertainty Model

Each concept (a) to (d) leads to an uncertainty model (17.5) with different varied stiffnesses k according to different nominal stiffness values k_0 listed in Table 17.1. Figure 17.4 shows the progress of decreasing stiffness k for the concepts. For $\mathcal{U}(k_0, h) = 0$, the stiffnesses $k = k_0$ for concepts (a) to (d). For $\mathcal{U}(k_0, h) = 1$, all stiffnesses would be zero, that would be a total failure of the suspension rod and unlikely to happen.

From an engineering point of view, only the uncertainty in decreasing stiffness is of interest, since decreasing stiffness means a higher relative compression stroke z_r for a given constant relative loading F_r, see (17.8) and Fig. 17.2. Higher relative static compression stroke, in turn, means lowering the safety margins in reaching the maximum absolute static compression $z_{\mathrm{a,\,max}}$

17.3.4 Performance Requirement

The effectual performance requirement (17.6) has been thoroughly explained in the guideline Sect. 17.3.1. The critical or maximum allowable absolute compression stroke $z_{\mathrm{a,\,max}}$ should not be reached.

17.3.5 Robustness to Uncertainty

Figure 17.5 shows the robustness or maximum uncertainty horizon $\widehat{h}$ as expelled maximum values of h according to (17.7) for varied stiffnesses k when the maximum allowable static absolute compression stroke $z_{\mathrm{a,\,max}} = 1.47 \cdot z_{\mathrm{a,\,stat}}$, Table 17.1, is reached for all landing gear design concepts (a) to (d) and for different absolute static loads F_a in four cases (A) to (D).

For case (A), a constant static absolute load $F_\mathrm{a} = 1 \cdot F_{\mathrm{a,\,stat}}$ is assumed. For all design concepts (a) to (d), robustness $\widehat{h}$ is greater than 0.19. That means that for the nominal static load $F_{\mathrm{a,\,stat}}$, a safety margin for at least 19% robustness is given for varied stiffnesses for all concepts. The highest robustness 32% is given for concept (a), the lowest with 19% for concept (c).

For case (B), a constant static absolute load $F_\mathrm{a} = 1.10 \cdot F_{\mathrm{a,\,stat}}$ is assumed. For all design concepts (a) to (d), robustness $\widehat{h}$ is greater than 0.11. That means that for the static load $1.10 \cdot F_{\mathrm{a,\,stat}}$, a safety margin for at least 11% robustness is given for varied stiffnesses for all concepts. The highest robustness 25% is given for concept (a), the lowest with 11% for concept (c).

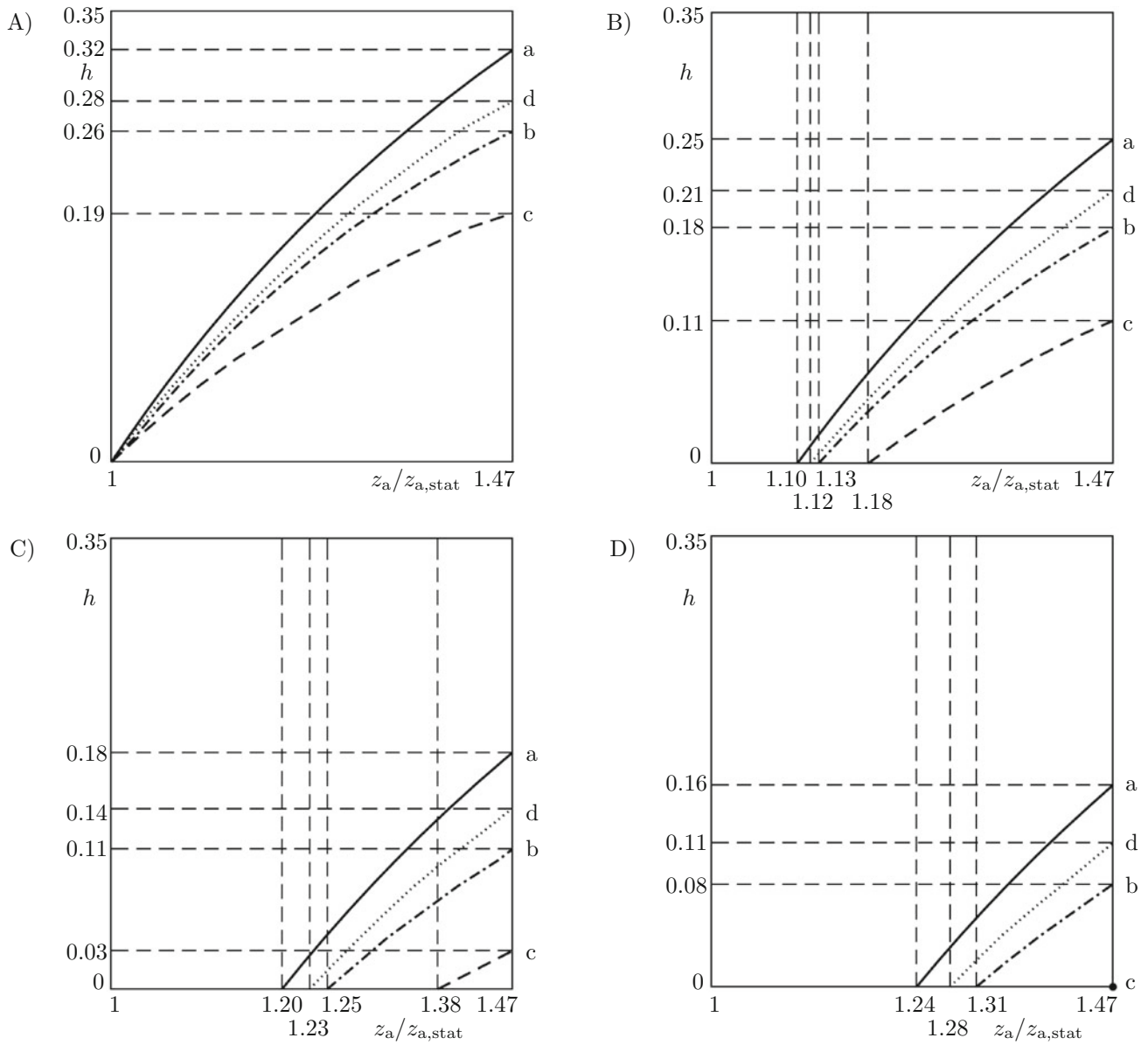


Fig. 17.5 Uncertainty horizon h as a result of varied decreasing stiffness k according to (17.5) vs. the compression stroke z_a, normalized to the nominal static absolute stroke $z_{a,\,stat}$ for (**A**) static load $F_a = 1 \cdot F_{a,\,stat}$, (**B**) $F_a = 1.10 \cdot F_{a,\,stat}$, (**C**) $F_a = 1.20 \cdot F_{a,\,stat}$, and (**D**) $F_a = 1.24 \cdot F_{a,\,stat}$ with respect to the landing gear concepts (a) *black solid line*, (b) *black dot-dashed line*, (c) *black dashed line*, and (d) *black dotted line*

For case (C), a constant static absolute load $F_a = 1.20 \cdot F_{a,\,stat}$ is assumed. For all design concepts (a) to (d), robustness $\widehat{h}$ is greater than 0.03. That means that for the static load $1.20 \cdot F_{a,\,stat}$, a safety margin for at least 3% robustness is given for varied stiffnesses for all concepts. The highest robustness 18% is given for concept (a), the lowest with 3% for concept (c).

For case (D), a constant static absolute load $F_a = 1.24 \cdot F_{a,\,stat}$ is assumed. Only for the design concepts (a), (b) and (d), robustness $\widehat{h}$ is greater than 0.08. That means that for the static load $1.24 \cdot F_{a,\,stat}$, a safety margin for at least 8% robustness is given for varied stiffnesses for these concepts. The highest robustness 16% is given for concept (a), no robustness with 0% for concept (c).

17.4 Conclusion

This contribution verifies the usefulness of the INFO-GAP approach to quantify uncertainty for major landing gear design concepts in a non-probabilistic way in early design stage. The concepts comprise one telescopic and three different trailing link designs. They are modeled mathematically in a simple but adequate analytic way according to guidelines found in distinctive literature. It is assumed that the stiffness of the integrated elastic strut is varied. The uncertainty in stiffness has major impact on the static compression stroke ability of the landing gears. The robustness to uncertainty of reaching the maximum allowable static compression stroke as a result to varied stiffness, and not as a result of exceeding maximum loading, could be quantified. The uncertainty quantification shows that the concept with the telescopic design with fixed elastic strut and hinged upper/lower struts as torque links is most robust against varied stiffness, whereas the concept with trailing link, hinged elastic strut, fixed upper strut, hinged lower strut, and high leverage is least robust. As an outlook, the authors will have a closer look to additional varied parameters to get an even more distinctive view on the uncertainty in the design concepts at early design stage.

Acknowledgements The authors like to thank the German Research Foundation DFG for funding this research within the SFB 805.

References

1. Barfield, N.A.: Fifty Years of Landing Gear Development: a survey of the pioneering contribution to aircraft undercarriage development made by Vickers and B.A.C. at Waybridge. Aircr. Eng. Aerosp. Technol. **40**(1), 17–22 (1968)
2. Ben-Haim, Y.: Info-Gap Decision Theory, Decisions Under Severe Uncertainty, 2nd edn. Elsevier, New York (2006)
3. Bruhn, E.F.: Analysis and Design of Flight Vehicle Structures. Jacobs Publishing, Inc., Indianapolis (1973)
4. Chai, S.T., Mason, W.H.: Landing gear integration in aircraft conceptual design, Report MAD 96-09-01, Multidisciplinary Analysis and Design Center for Advanced Vehicles, Virginia Polytechnic Institute and State University Blacksburg, Virginia, 24061-0203 (1997)
5. Conway, H.G.: Landing Gear Design. Chapman and Hall, London (1958)
6. Currey, N.S.: Aircraft Landing Gear Design: Principles and Practices. AIAA Education Series. AIAA, Washington (1988)
7. Gudmundsson, S.: General Aviation Aircraft Design. Elsevier, New York (2014)
8. Hoare, R.G.: Aircraft landing gear: an account of recent developments in undercarriage design an the materials used. Aircr. Eng. Aerosp. Technol. **40**(1), 6–8 (1968)
9. Hot, A., Cogan, S., Foltête, E., Kerschen, G. Buffe, F., Buffe, J., Behar, S.: Design of uncertain prestressed space structures: an info-gap approach. In: 30th IMAC Annual Conference on Experimental and Applied Mechanics, Topics of Model Validation and Uncertainty Quantification, Volume 4, Conference Proceedings of the Society for Experimental Mechanics Series, pp. 13–20 (2012)
10. Niu, M.C.Y.: Airframe Structural Design. Conmilit Press Ltd., Hong Kong (1988)
11. Peiffer, A.: Full frequency vibro-acoustic simulation in the aeronautics industry. In: Proceedings (Keynote) of ISMA 2016 including USD 2016 International Conference on Uncertainty in Structural Dynamics, Leuven, pp. 1–15 (2016)
12. Platz, R., Götz, B.: Approach to evaluate and to compare basic structural design concepts of landing gears in early stage of development under uncertainty. In: 34th IMAC Annual Conference on Experimental and Applied Mechanics, Topics of Model Validation and Uncertainty Quantification, Volume 3, Conference Proceedings of the Society for Experimental Mechanics Series, pp. 167–175. Springer, New York (2016)
13. Pragadheswaran, S.: Conceptual design and linear static analysis of nose landing gear. In: International Conference on Robotics, Automation, Control and Embedded Systems – RACE 2015, 18–20 February 2015, Hindustan University, Chennai (2015)
14. Veaux, J.: New design procedures applied to landing gear development. J. Aircr. **25**(10), 904–910 (2014)
15. Wang, H., Xue, C.J.; Jiang, W.T.: Fuzzy fatigue reliability analysis for a landing gear structure. In: IEEE Conference Publications, Computational Science and Optimization (CSO), 2010 Third International Joint Conference on Computational Science and Optimization, vol. 1, pp. 112–115 (2010). doi:10.1109/CSO.2010.130
16. Xue, C.J., Dai, J.H., Wei, T., Lui, B.: Structural optimization of a nose landing gear considering its fatigue life. J. Aircr. **49**(1), 225–236 (2012)

Chapter 18
Modular Analysis of Complex Systems with Numerically Described Multidimensional Probability Distributions

J. Stefan Bald

Abstract To analyze complex systems, many teams have to collaborate, often from different disciplines. The presented approach allows a systematic modular description of such systems (esp. systems, that are influenced by uncertain values, e.g. from human behavior or from mechanical/material uncertainty). The model of the system is made up from a network of modules, which are connected by interfaces. Each interface is a set of values, each described by Numerically Described Multidimensional Probability Distributions (NDMPD). This allows a description even of scattered values. The behavior of each module can be described by

- formulas (the related computer tool will evaluate it for parameters given as NDMPD and provide the result as NDMPD)
- NDMPD, gathered by
  - surveys,
  - evaluations of the module as a subsystem,
  - many other methods, e.g. Monte Carlo,
  - appropriate assumptions.

If a module is considered as a subsystem, it can be hierarchically modeled by sub-modules. The behavior of the subsystem can be described with NDMPD, which are derived from the behavior of its (sub-)modules. As the system behavior is described numerically, computers can evaluate it. This allows automatic reevaluation, e.g. after a change of a single module (after a refinement, for new findings, for other hypothesis etc.).

Keywords System analysis • Cooperation • Interdisciplinary work • Dealing with uncertainty • Probability distributions

18.1 Introduction

Actual work on technical system is linked with increasing complexity. In addition, extensive "optimization to the limit" of systems and material requires very differentiated treatment.

The complexity of the systems requires cooperation of numerous teams, often from different disciplines. They have to deal with their own subject in a very detailed manner, while the other teams need a more comprehensive representation of their findings. Thus it is important, that there are scalability of the work results and efficient means to put the individual results of the teams together to an overall solution.

The "optimization to the limit" requires to deal with the whole range of values. The relevant combination of different values is not necessarily one of extreme values, as the probability for such combination is often very, very low. Experience shows, that most failures occur, if many parameters are on a bad medium range, while only one or two get extremely bad. Therefore the combination of values and their probability has to be considered as well as the eventual interdependency between the values. The increasing use of methods to deal with uncertainty (fuzzy logic, Monte Carlo methods, neural networks) shows the need for such analysis.

In this paper a method shall be presented which enables organizations and teams to deal with these challenges.

J.S. Bald (✉)
Road and Pavement Engineering, Technische Universität Darmstadt, Otto-Bernd-Str. 2, 64287 Darmstadt, Germany
e-mail: jsbald@sw.tu-darmstadt.de

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_18

18.2 Development of Research

Searching for better analysis methods to deal with road safety, a literature study [1] has been conducted, especially evaluating the German Risk Analysis of Nuclear Power Plants [2] and a collection on analysis of important safety hazards of industrial sites [3]. The result of the study was, that the methods, that were applied there, were well suited to be applied to road safety problems, especially the methods, that described the systems systematically with processes and situations (that were later renamed to "passive" and "active" elements); the only problem to deal with was human behavior, which was tried to eliminate in the industrial processes, but is a major issue in road traffic, where it cannot be eliminated and therefore has to be integrated into the analysis process.

The author developed such a method in his doctoral dissertation [4]. Road traffic is governed by automatic human behavior. Such behavior may be described by probability distributions, if the influence of boundary conditions can be taken into account. A method was to be found to systematically describe arbitrary probability distributions related to other values (which itself could be distributions), e.g. boundary conditions. In his doctoral dissertation, the author proposed to use collections of Bayes' distributions (later [5] named "Numerically Described Multidimensional Probability distribution (NDMPD)") to describe system variables. He developed a language to describe systematic collections of such distributions, described an algorithm to combine and evaluate such NDMPD and provided a software to easily work with them. With some examples he proved the applicability.

In later research, the software was transferred to Linux/UNIX-Systems [6, 7], further developed to use human behavior [8] and applied to material projects [9].

Actual research is using the method to analyze road state development functions (for Pavement Management Systems PMS) and to conduct structural analysis. In addition, the program code is overhauled to allow more cooperation (web orientation, database integration) and to obtain higher speed (multiprocessing).

18.3 General Considerations on Systems and Modules

This paper concentrates on the ability of the method to enhance the analysis of complex systems by many different teams.

A system is a virtually or really separated part of reality, which is perceived as an entity. A system is separated from its environment by a system boundary, through which there is limited interaction through well defined interfaces only. The concept of systems is well suited to analyze and improve somehow limited structures as machines, ecosystems or mental models. As a rule, systems are complex. While analyzing them systematically, it is normally easier to consider internal dependencies than connections to the system environment. Therefore one tends to enlarge systems to consider more influences of the reality. Such large systems require many teams to be analyzed, often from different disciplines.

One approach to deal with such complex systems is to split them up into parts (modules), that can be analyzed easier. These modules may be split up further to sub-modules and so on, until the size of the sub-modules is small enough to be analyzed by a team of specialists. The problem, which arises from this splitting process, is, that with every split one creates new interfaces between (sub-)modules with exactly that limitations, that lead to the enlargement of the system before.

The Darmstadt Risk Analysis Method DRAM [6, 7] has been developed to analyze such complex systems by subdividing them into sub-systems, to analyze these sub-systems systematically, and to compile the partial results to an overall result systematically and automatically. Originally it was developed to address risk analysis of road system (therefore the name), but in the last years it proved to be applicable for other system related problems with scattering variable values, too. Due to its numerical structure, its precision and complexity is limited only by the capacity of the data crunching machine.

The objectives of the design are given in the left column of Fig. 18.1.

As components of the method were identified (Fig. 18.1):

- assessment by using probability values
- systematically describing the system by using active and passive elements
- numerically described multidimensional probability distributions (NDMPD)
- cooperation and management of knowledge

To facilitate the handling, two tools have been designed (and exemplarily realized):

- a computer program Darmstadt Risk Analysis Tool DRAT to deal with NDMPD
- a database of knowledge DoKn to systematically file the knowledge collected with the method

The components and the tolls shall be described in the following as detailed as necessary to deal with the topic of this paper.

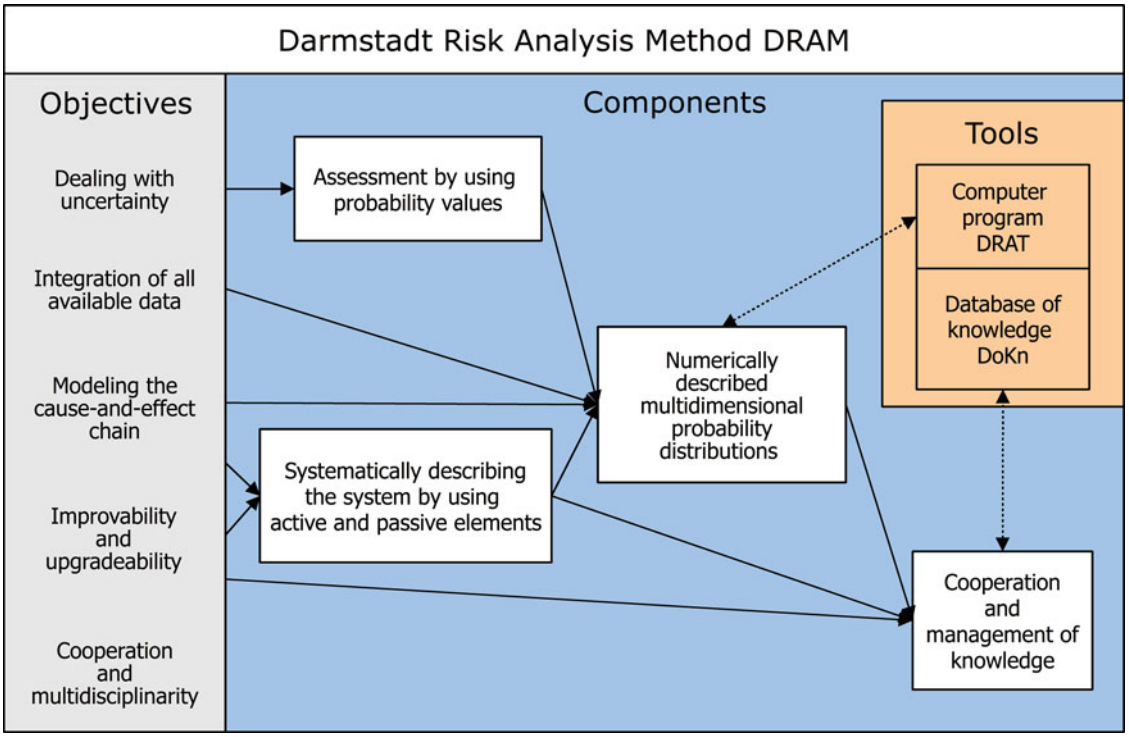


Fig. 18.1 Objectives, components and tools of Darmstadt Risk Analysis Method DRAM [6]

18.4 Using Probability Values

The advantages of using probability distributions to describe the values of parameters has already been addressed in the introduction. In the context of this paper it seems not necessary to deepen this argumentation. The concept is broadened by considering certain dependencies, what eventually leads to multidimensional probability distributions.

18.5 Numerical Described Multidimensional Probability Distributions (NDMPD)

Numerical Described Multidimensional Probability Distributions (NDMPD) describe such multidimensional probability distributions by numerical values. In principle they are a systematic collection of probability values.

An one-dimensional distribution can be described by a scale (names for nominal and ordinal scales; numbers as class limits for cardinal scales), which divides the value range into classes, and probability values for each class. If such a distribution describes a variable, the probability values indicate for each class the probability, that the value of the variable will lay in this class. Depending on the scale type, additional classes are added for very high, very low and undefined values.

Multidimensional distributions have an additional dimension for each dependency, that shall be considered. A two-dimensional NDMPD is effectively a set of one-dimensional distributions, one for each class of the dependency scale. Then, the probability values are not to be interpreted absolutely, but Bayes' probabilities in relation to the dependency (p | dependency). This is continued for each dependency (Fig. 18.2). Thus, a five-dimensional NDMPD then consists of 5 scale descriptions and–with about 10 classes per scale/dimension–of $10^5 = 100{,}000$ (related) probability values. That sounds much, but is no problem for actual computers. To the contrary: the systematic structure of NDMPD allows the treatment of complex NDMPD by machine, even if many NDMPD are connected by dependencies.

The method is designed in that way, that the algorithm recognizes, if two NDMPD are referenced to the same scale and considers the entity with this scale as mutual dependency of the two variables described by the NDMPD, and considers it in the calculation process.

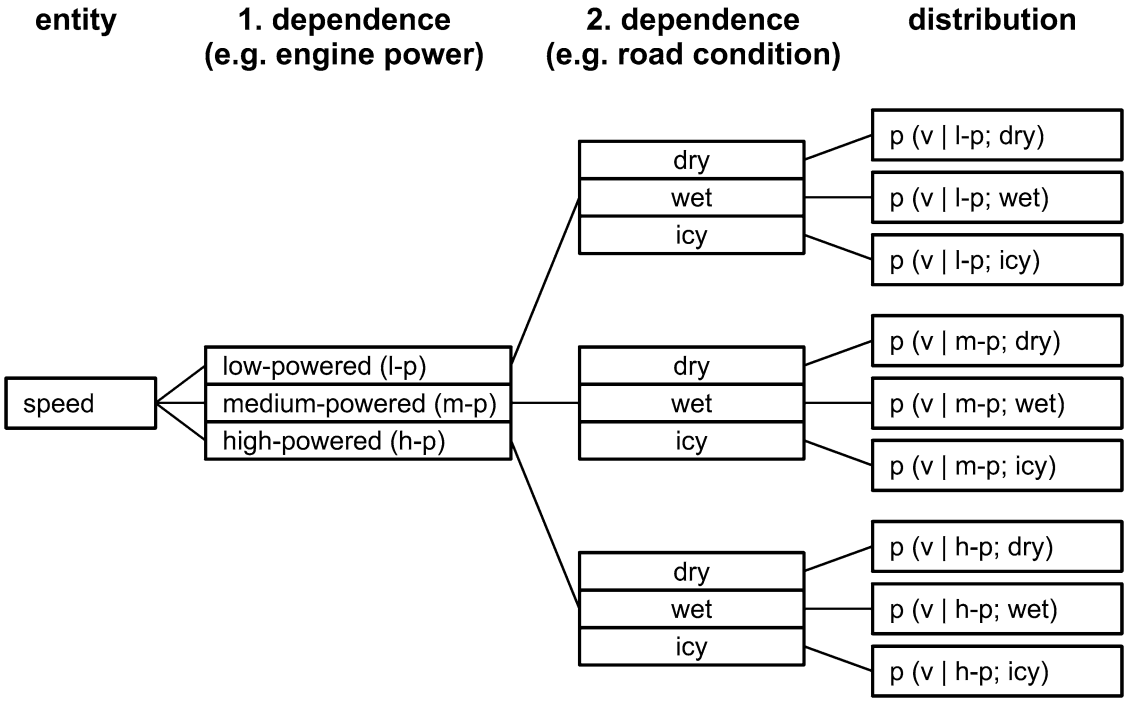


Fig. 18.2 Hierarchical Structure of NDMPD [4]

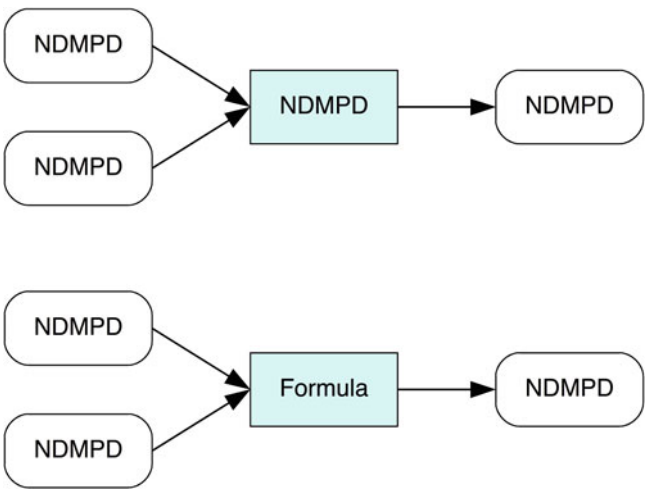


Fig. 18.3 Interfaces are described by NDMPD, module behavior by NDMPD or by formulas

18.6 Active and Passive Elements, Modules and Interfaces

The generic approach of DRAM starts from the idea, that each system can be described by a branched network of active elements (modules), that are connected by passive elements (interfaces), that describe more and less the cause–and–effect-structure of the system.

The system behavior is described by the module behavior, by means of formulas or by NDMPD (as result dependent of the input parameters and/or boundary conditions, see below; Fig. 18.3). The interfaces act as clue between the modules and allow the exchange of module descriptions without affecting bigger parts of the system description. They also should be described by NDMPD allowing to consider complex interdependencies from intermediate results and boundary conditions.

As the caretakers of the modules tend to need a much more detailed view on his/her module than the caretakers of the surrounding modules, they may (and in general will) split the modules in sub-modules, thus generating new interfaces.

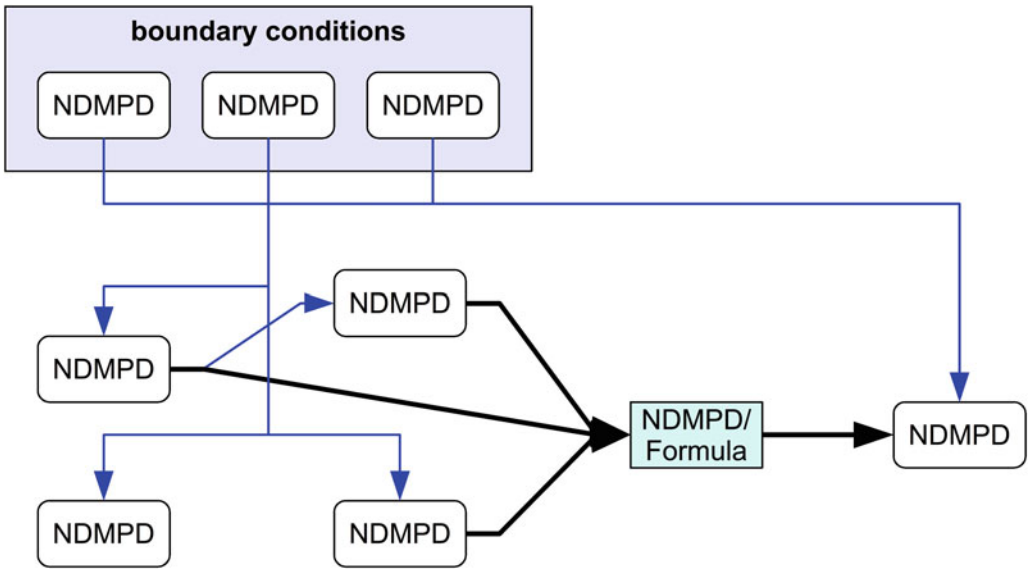


Fig. 18.4 Dependencies may create a complicated network

DRAM is designed in a way to allow to eliminate these additional interfaces at the end of the analysis and provide an integrated NDMPD for the whole module, that is systematically derived from the network of sub-modules.

The information of the NDMPD, that describe the sub-modules, may be created by theoretical considerations, by surveys, by evaluations of the module as a subsystem or by many other methods, e.g. by Monte Carlo methods or by applying assumptions.

The approach of DRAM allows the caretaking group of experts to analyze the module with a method, it considers as appropriate, but to provide their results in a flexible, but standardized format (as NDMPD), that can be easily integrated into a general analysis.

NDMPD are well suited to describe the values at the interfaces, as dependencies to other results or to boundary conditions may be considered. They may be used to describe the outcome of (sub-)module behavior, too, especially, if this behavior is a summarizing result of a detailed analysis. The value of the outcome of the module may be obtained by combining (convoluting) the module behavior NDMPD with the NDMPD of the inputs and of the boundary conditions (Fig. 18.4). In the case, that the module behavior is described by a formula whose variables are described by NDMPD, the formula is evaluated considering the statistical nature of the variables. In all cases it may be appropriate to retain the dependency to certain intermediate results or to boundary conditions, e.g. to take it into account at subsequent combinations. For the calculations, the author uses an approach, that is based on interval arithmetic, but any other algorithm may be used (e.g. Monte Carlo simulation). Which algorithm is used, is mostly a question of precision and efficacy, which still has to be investigated systematically. However NDMPD are a suitable way to exchange data between divers modules or to describe the behavior of particular modules (and document and archive it), especially if these data depend on joint dependencies.

18.6.1 Working with NDMPD as Description of System Behavior

NDMPD lead to a large amount of numerical data, that requires automatic processing by computers. But the systematic and hierarchic structure of the data allows such processing. In this way, the use of NDMPD is somehow comparable to Finite Element Methods, where also a large amount of systematic structured data enables to analyze systems, whose analysis seemed impossible thirty years ago. In both cases the human describes the system and the machine cares about the calculation.

If NDMPD are used to exchange data between modules, this data can also be stored as intermediate results, allowing to continue the calculation with different assumptions or scenarios (after a refinement, for new findings, for other hypothesis etc.) for the following modules, thus enabling improvability and upgradability, as demanded at the beginning.

18.6.2 Database of Knowledge

NDMPD are also suited to document findings of research. Concepts to systematically set up such databases of Knowledge are under investigation.

18.6.3 Summary and Conclusions

Modeling complex systems systematically with active and passive description elements (modules and interfaces) is an appropriate way to deal with the requirements of the analysis of such systems, especially if the interfaces between the modules and eventually the module behavior itself are described with "Numerically Described Multidimensional Probability Distributions" (NDMPD).

References

1. Durth, W., Bald, J.S.: Risikoanalysen im Straßenwesen (Risk analysis in road engineering). In: Schriftenreihe Forschung Straßenbau und Straßenverkehrstechnik Heft, vol. 531, pp. 1–68, BMV (German DOT) (1988/1997)
2. Div.: Deutsche Risikostudie Kernkraftwerke (German risk analysis of nuclear power plants). Verlag TÜV Rheinland 2. Ed. (1980)
3. Div.: Große technische Gefahrenpotentiale (Large technical risks). Springer (1983)
4. Bald, J.S.: Grundlagen für die Anwendung von Risikoanalysen im Straßenwesen (Basics for the application of risk analysis methods in road engineering). Doctoral Dissertation, Technische Hochschule Darmstadt (1991). In: Schriftenreihe Forschung Straßenbau und Straßenverkehrstechnik Heft, vol. 531, pp. 69–196, BMV (German DOT) (1988/1997)
5. Bald, J.S., Stumpf, K., Wallrabenstein, T.: Modelling human behaviour by numerically described multidimensional probability distributions. In: Advances in Transportation Studies, An International Journal, vol. 16 (2008), University of Roma Tre http://host.uniroma3.it/riviste/ats/
6. Bald, J.S., Stumpf, K., Wallrabenstein, T., Huyen, L.T.: Road risk analysis tools. Deliverable 3.2 of the 6th Framework EU Project In-Safety (2008) (http://www.insafety-eu.org)
7. Bald, J.S., Stumpf, K., Wallrabenstein, T., Huyen, L.T.: Managing the risks. Road risk analysis tools. In: Bekiaris, E., Wiethoff, M., Gaitanidou, E. (eds.) Infrastructure and Safety in a Collaborative World, pp. 143–152. Springer, Berlin (2011)
8. Bald, J.S., Heimbecher, F.: Darmstadt risk analysis method (DRAM): a generic method for modular, systematic, quantitative and interdisciplinary risk assessments considering human behavior. In: Cacciabue, P.C., Hjälmdahl, M., Luedtke, A., Riccioli, C. (eds.) Human Modelling in Assisted Transportation, pp. 199–206. Springer, Berlin (2011)
9. Rosauer, V.: Abschätzung der herstellungsbedingten Qualität und Lebensdauer von Asphaltdeckschichten mit Hilfe der Risikoanalyse (Assessment of quality and longevity of asphalt wearing courses with risk analysis). Doctoral dissertation, TU Darmstadt (2010)

Chapter 19
Methods for Component Mode Synthesis Model Generation for Uncertainty Quantification

A.R. Brink, D.G. Tipton, J.E. Freymiller, and B.L. Stevens

Abstract Component mode synthesis (CMS) is a widely employed model reduction technique used to reduce the computational cost associated with the dynamic analysis of complex engineering structures. To generate CMS models, specifically the formulation of Craig and Bampton, both normal fixed-interface modes and constraint modes of the component's structure are calculated. These modes are used in conjunction with the component level mass and stiffness matrices to generate reduced mass and stiffness matrices used in the final analyses. For some component models, the most computationally expensive part of this procedure is calculating the component normal modes information. Several different approaches are utilized to investigate the sensitivity of system level responses to variations in several aspects of the CMS models. One approach evaluates changes due to modifications of the reduced mass and stiffness matrices assuming that the mode shapes do not change. The second approach assumes that the mode shapes change but the reduced mass and stiffness matrices do not change. An example is presented to show the influence of these two approaches.

Keywords Component mode synthesis • CMS • Craig-Bampton • Uncertainty quantification • UQ

19.1 Introduction

As engineering structures increase in complexity, cost and importance, the level of rigor in the computational analyses of these structures must increase proportionally. These computational analyses must include quantification of the uncertainties (UQ) associated with the structure and capture the response with a high level of accuracy. One common method for determining the effect that model uncertainty has on the predicted structural response is to define ranges for the uncertain parameters, then use Monte Carlo, Latin Hypercube, or other similar sampling methods to explore the parameter space [1]. All of these sampling techniques require the numerical model to be rebuilt with each new parameter set assignment and the full analyses suite re-calculated. This method is often extremely computationally expensive for large scale numerical models. To alleviate some of the computational cost, model reduction techniques, such as component mode synthesis, are employed. Component mode synthesis (CMS) has been a heavily utilized model reduction tool since its inception in the early 1960s. It is useful not only for its ability to reduce computational cost during the analysis of complex structures but also in facilitating the sharing of major component assemblies among multiple design agencies. Currently, the most heavily used derivation of CMS is that proposed by Craig and Bampton, which relies on fixed interface modes. A review of the Craig-Bampton method is provided in Sect. 19.2. Since the use of CMS models is required to reduce the computational burden while maintaining high response accuracy, they must also be subjected to a UQ with similar rigor as that applied to the full structure. Recreating the Craig-Bampton reduction for each new parameter set introduces a new computational cost associated with generating the fixed-interface modes of the structure. Many researchers are working to investigate how to handle this UQ while not introducing a large computational expense using parameterized reduced order models (PROMs) [2–7]. According to [2], PROMs include the derivative of parametric terms while building the CMS model. By including these derivatives, uncertain parameters are directly included in the CMS formulation, thus it is not required to rebuild the model with each new parameter set. While these PROM techniques are mathematically elegant and more exact than what is presented here, a method is sought that is easily implemented into existing FE codes. This paper explores two techniques to recreate the Craig-Bampton reduction *without* recomputing the mode shapes each time. One method, developed by Tipton

A.R. Brink (✉) • D.G. Tipton • B.L. Stevens
Sandia National Laboratories, Albuquerque, NM, USA
e-mail: arbrink@sandia.gov

J.E. Freymiller
ATA Engineering, San Diego, CA, USA

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_19

and Brink, calculates the mode shapes for the nominal condition, then uses the same mode shapes for all realizations of the structure. The mass and stiffness matrices are updated to reflect the new parameter set. The second method, developed by ATA Engineering, instead retains the mass and stiffness matrices of the nominal structure and perturbs the mode shape matrix with a cross-orthogonality matrix, corresponding to the expected change in mode shapes resulting from the new parameter set. Both of these methods are developed in Sects. 19.3.1 and 19.3.2, respectively. They are then applied to a prototypical structure, and their results compared for validity to full Craig-Bampton reductions, which include recalculation of the modes for each parameter set.

19.2 A Brief Review of Craig-Bampton Reduced Order Models

Craig and Bampton's landmark work [8], first published in 1968, is a simplification of work done by Hurty [9] earlier in the same decade. This simplification involves the way in which interface degrees of freedom are handled. The procedure for generating a Craig-Bampton reduced order model is as follows [10]:

1. Generate a detailed numerical model (finite element or otherwise) of the system or component that is to be reduced.
2. Identify all interface and internal degrees of freedom.
3. Calculate the *fixed-interface modes* modal matrix for the system with all interface degrees of freedom fixed. Engineering judgment and convergence studies dictate how many of these modes to retain.
4. Calculate the *constraint modes* of the interface degrees of freedom.
5. Assemble the *Craig-Bampton transformation matrix*.
6. Calculate the reduced mass and stiffness matrices.
7. Assemble the reduced order model into the system numerical model.

19.2.1 Model Generation and DOF Identification

Generating the numerical model to be reduced is often time consuming. The analyst must build a full model capable of capturing all of the relevant dynamics. For baseline Craig-Bampton reduced order models, the model is completely linear, but can still be quite complex. To perform UQ analysis involving geometric parameters, this model needs to be rebuilt for each new parameter set. However, if changes are small, such as those attributed to geometric tolerances, then the original model can be tweaked without major effort. Identifying the interface and internal degrees of freedom is straight forward. Any degree of freedom that couples to the main structure of interest is considered an interface degree of freedom. All remaining degrees of freedom are considered internal.

19.2.2 Fixed-Interface Modes

In the Craig-Bampton formulation, fixed-interface modes are defined as the normal modes of the internal degrees of freedom with all interface degrees of freedom fixed. This requires a straight forward Eigen analysis. The modal matrix is stored as $\mathbf{\Phi}$. As stated earlier, the analyst must decide the appropriate number of modes to retain to achieve an acceptable level of accuracy.

19.2.3 Constraint Modes

Constraint modes are calculated by statically applying a unit displacement to one interface degree of freedom at a time, while fixing all other interface degrees of freedom. This not only helps to define the proper interface stiffness, but also helps to enforce inter-component compatibility during assembly into the complete structure. The constraint modes of the substructure, are calculated with

$$[\boldsymbol{\Psi}] = -[\mathbf{K_{nn}}]^{-1}[\mathbf{K_{na}}], \tag{19.1}$$

where a represents interface degrees of freedom and n the interior degrees of freedom.

19.2.4 Craig-Bampton Transformation Matrix

The Craig-Bampton transformation matrix, $\mathbf{W}$ is assembled using the component and constraint modes. This matrix transforms the generalized degrees of freedom used to synthesize the reduced order model back into the original system's degrees of freedom. This takes the form

$$\begin{Bmatrix} \mathbf{u_n} \\ \mathbf{u_a} \end{Bmatrix} = \underbrace{\begin{bmatrix} \boldsymbol{\Phi} & \boldsymbol{\Psi} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}}_{\mathbf{W}} \begin{Bmatrix} \mathbf{p_k} \\ \mathbf{u_a} \end{Bmatrix}, \tag{19.2}$$

where $\mathbf{p_k}$ is the vector of modal coordinates, and $\mathbf{u_n}$ and $\mathbf{u_a}$ are displacements in generalized and physical coordinates, respectively.

19.2.5 Reduced Stiffness and Mass Matrices

Once the Craig-Bampton transformation matrix is assembled, it is used to calculate the reduced mass and stiffness matrices for the component. These matrices are then used to assemble the reduced order model back into the main system and include both interface as well as modal degrees of freedom. They are calculated as

$$\left[\hat{\mathbf{M}}_{\mathbf{CB}}\right] = [\mathbf{W}]^T[\mathbf{M}][\mathbf{W}] \tag{19.3}$$

$$\left[\hat{\mathbf{K}}_{\mathbf{CB}}\right] = [\mathbf{W}]^T[\mathbf{K}][\mathbf{W}]. \tag{19.4}$$

19.3 Craig-Bampton Generation for UQ Studies

As was mentioned in Sect. 19.1, there is significant computational cost associated with calculating the fixed-interface modes, which is outlined in Sect. 19.2.2. It is also noted that for extremely large component models, there can be significant cost associated with inverting the stiffness matrix in Eq. (19.1). However, for this study it is assumed that the component model is a small enough size such that the fixed-interface mode calculation dominates. For a UQ study involving hundreds or thousands of parameter sets, re-calculating the fixed-interface modes for each set would put too much computational burden on the analysis. Presented in the next two subsections are methods which perturb the Craig-Bampton model to account for uncertain parameter sets without recalculating the fixed-interface modes.

19.3.1 REMAP Technique

The first technique presented is further separated into two sub techniques, REMAP1 and REMAP2. Both of these technique use the nominal fixed-interface modes when calculating the Craig Bampton transformation matrix. The assumption is that the perturbed state, due to the uncertain parameter set, is small enough that major changes to the fixed-interface modal matrix do not occur. For REMAP1, both the constraint modes, $\boldsymbol{\Psi}$, as well as the component mass and stiffness matrices are updated to reflect the UQ parameters. The Craig Bampton transformation matrix and reduced mass and stiffness matrices are shown in Eqs. (19.5), (19.7) and (19.8), respectively. REMAP2 uses all nominal entries in the Craig Bampton transformation matrix,

and only updates the component mass and stiffness matrices, as shown in Eqs. (19.7) and (19.8). Since the old modes are used to map the updated mass and stiffness matrices to the old basis, the technique is call REMAP. This is also short for *RE*duced *MA*trix *P*erturbation.

$$\mathbf{W}_{updated} = \begin{bmatrix} \mathbf{\Phi}_{nominal} & \mathbf{\Psi}_{updated} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{19.5}$$

where

$$[\mathbf{\Psi}]_{updated} = -[\mathbf{K_{nn}}]^{-1}_{updated}[\mathbf{K_{na}}]_{updated}. \tag{19.6}$$

The stiffness terms in Eq. (19.6) are based on the perturbed parameter sets that are generated during the UQ, as are the mass and stiffness matrices used in Eqs. (19.7) and (19.8), respectively.

$$\left[\hat{\mathbf{M}}_{\mathbf{CB}}\right]_{updated} = [\mathbf{W}]^{T}_{updated}[\mathbf{M}]_{updated}[\mathbf{W}]_{updated} \tag{19.7}$$

$$\left[\hat{\mathbf{K}}_{\mathbf{CB}}\right]_{updated} = [\mathbf{W}]^{T}_{updated}[\mathbf{K}]_{updated}[\mathbf{W}]_{updated}. \tag{19.8}$$

Since this method requires knowledge of the unreduced structure to generate the mass and stiffness matrices, it is ideal for UQ analysis within the same design agency. If the Craig-Bampton reduced order model is transferred to another design agency for implementation into a larger system model, then this technique is not applicable or requires multiple realizations to be transferred. An analysis on a prototypical structure using this technique is presented in Sect. 19.4.

19.3.2 COMP Technique

Where the REMAP technique perturbs the mass and stiffness matrices of the reduced order model, the COMP technique instead perturbs the modal matrix, and leaves the mass and stiffness matrices in their nominal condition. To perturb the modal matrix, it is pre-multiplied by a cross-orthogonal modal matrix. Hence the acronym *C*ross-*O*rthogonal *M*odal *P*erturbation. Consider if the component modal matrices were known for both the nominal state and the state that corresponds to an updated parameter set from UQ. Then, the cross-orthogonality matrix of the two modal matrices is

$$[c_{\Phi}] = [\Phi]^{T}_{nominal}\,[M]_{nominal}\,[\Phi]_{updated}. \tag{19.9}$$

If both modal matrices are identical, then $[c_{\Phi}]$ is the identity matrix. The Craig-Bampton transformation matrix used in the COMP technique is then

$$\mathbf{W}_{updated} = \begin{bmatrix} \mathbf{c}_{\Phi}\,\mathbf{\Phi}_{nominal} & \mathbf{\Psi}_{nominal} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{19.10}$$

For the example problem shown in Sect. 19.4, the cross-orthogonality matrix is generated with a priori knowledge of how the UQ parameter set changes the modal matrix. If the matrix is to be perturbed without this knowledge, care must be taken by the analyst to ensure realistic mode shapes are generated by the perturbation. In addition to perturbing the modal matrix, it is also possible using this technique to perturb the Eigen vector, although this is not considered in this paper.

19.4 Application of REMAP and COMP Techniques

In this section the REMAP1, REMAP2 and COMP techniques are applied to the structure shown in Fig. 19.1. The nominal parameters of the structure are given in Table 19.1. In addition to the linear beams that make up the structure, a point mass/inertia is added to the center of each 'X' structure. The nominal mass is 2.863 kg and the nominal inertia is

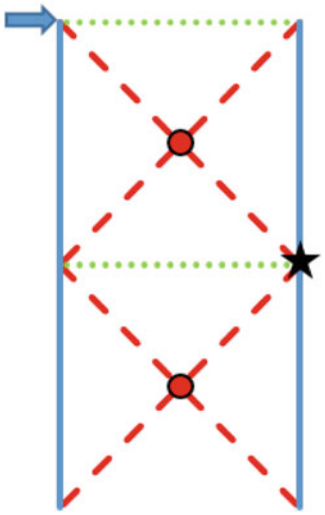


Fig. 19.1 Structure under consideration. The two 'X' shaped structures, assembled with the *red dashed lines*, are the reduced substructures. The *solid blue/green frame* is not reduced

Table 19.1 Nominal beam properties for the example structure shown in Fig. 19.1

Beam type	D (mm)	L (mm)	E (Pa)	ρ (kg/m^3)
Frame (blue/solid)	12.50	300.00	69×10^9	2700.0
Frame (green/dotted)	12.50	300.00	69×10^4	2700.0
'X' (red/dashed)	6.25	212.13	69×10^9	2700.0

0.8052 kg·m^2. Each node of the beams has two translational and one rotational degrees of freedom. The two red/dashed 'Xs' are reduced using a full Craig-Bampton reduction (i.e. full recalculation of the Craig-Bampton transformation matrix), as well as the REMAP1, REMAP2 and COMP techniques. To simulate a UQ study, each parameter listed in Table 19.1, as well as the point mass/inertia and the angle of the beams, are perturbed to form a Gaussian distribution, as shown in Fig. 19.2. This parameter perturbation only occurs on the 'X' structure being reduced. The remainder of the frame is left in its nominal state.

To conduct the simulation, a random parameter set is generated. This same parameter set is then propagated through the full, REMAP1, REMAP2 and COMP Craig-Bampton reduction techniques. Each realization is then assembled into the frame structure, from which modal information as well as frequency response functions (FRFs) are extracted. For this study, 300 unique parameter sets are analyzed. All fixed-interface modes are retained for this study and a structural damping factor of 3% is used for all post-processing. Looking first at the FRFs generated, Fig. 19.3 shows the upper and lower bounds of the FRF computed using each reduction technique. Figure 19.4 magnifies two of the FRF peaks for clarity. These figures show that the REMAP1 technique produces nearly identical bounds on the problem as the full Craig-Bampton reduction does. REMAP2 tracks the true envelope closely for lower frequencies, then tends to collapse on the COMP technique plot. This indicates that the constraint modes are important in this frequency range. The FRF bounds computed by the COMP method are collapsed nearly identically to the nominal parameter set. This indicates that the FRF is insensitive to the fixed-interface modes and instead relies mainly on the constraint modes to produce the response. The drive point FRF is shown in Fig. 19.5 and a magnification of the two regions of interest in Fig. 19.6. This figure shows that up to 450 Hz, the constraint modes dominate, between 450 Hz and 950 Hz, there is a transition to reliance on the fixed-interface modes. This is evident in the collapse of the full Craig-Bampton reduction away from the REMAP1 and REMAP2 solutions and onto the COMP solution. For this drive point FRF, the REMAP1 and REMAP2 techniques produce nearly identical bounds on the problem.

To compare the modal data, the cross-orthogonality of the modes calculated by the REMAP and COMP techniques are compared to the full Craig-Bampton reduction. Figure 19.7 shows the statistics for the diagonal terms in the cross-orthogonality matrix. Figure 19.8 shows the maximum off diagonal term in the matrix. If the modes are exactly orthogonal, then the diagonal term is unity and the off diagonal terms are zero. The red horizontal line indicates the mean value, the blue bounding box shows the 25th and 75th percentile bounds, and the end of the whiskers show the maximum and minimum values, excluding statistical outliers. These plots show that the each technique calculates certain modes properly and others not as well. Which modes it calculates correctly is problem dependent. Note that REMAP1 and REMAP2 produce nearly identical graphs, for this reason, only the REMAP1 plot it shown.

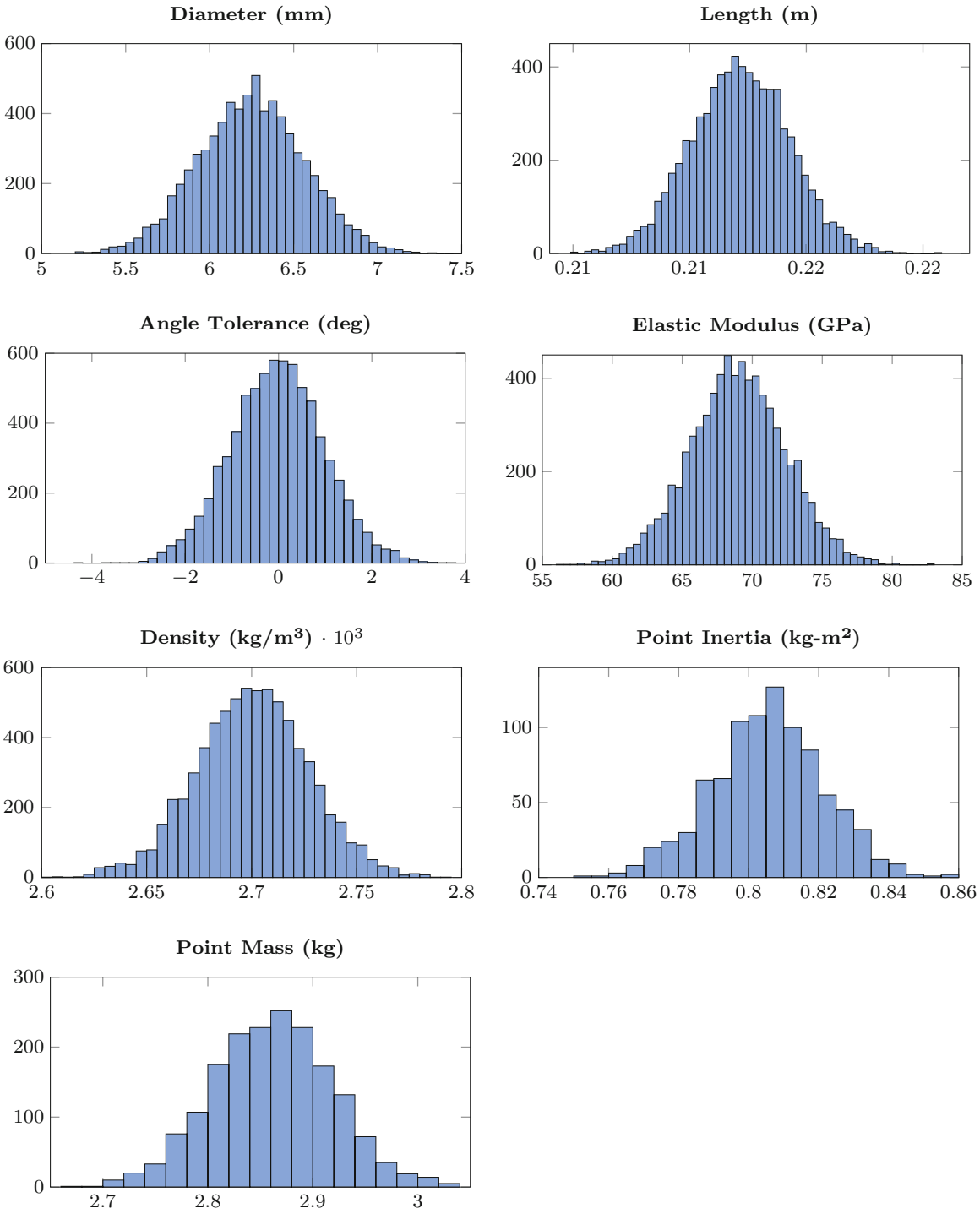


Fig. 19.2 Distribution of parameters

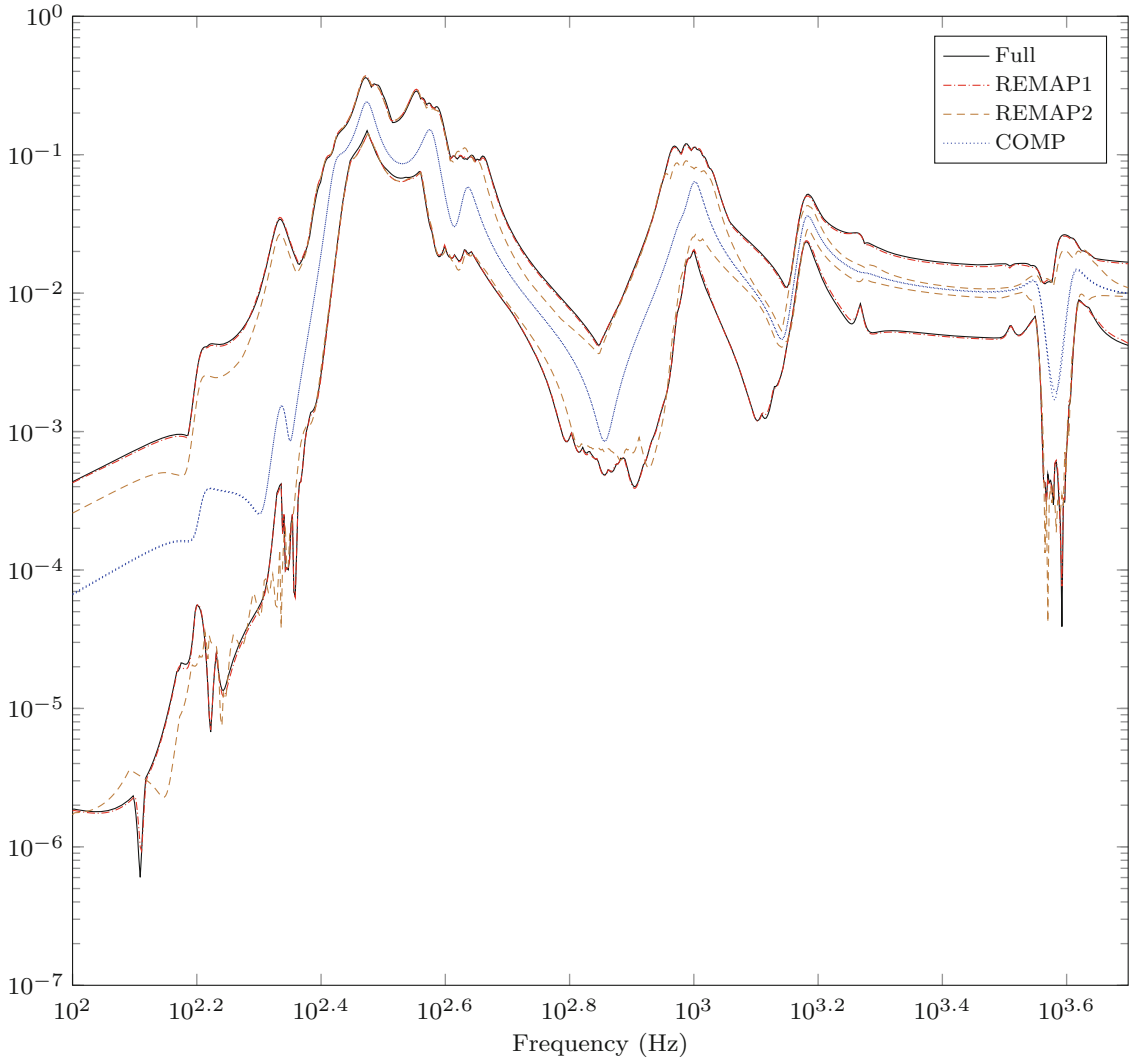


Fig. 19.3 Referencing Fig. 19.1, FRF envelopes for force in at the arrow and acceleration out at the star

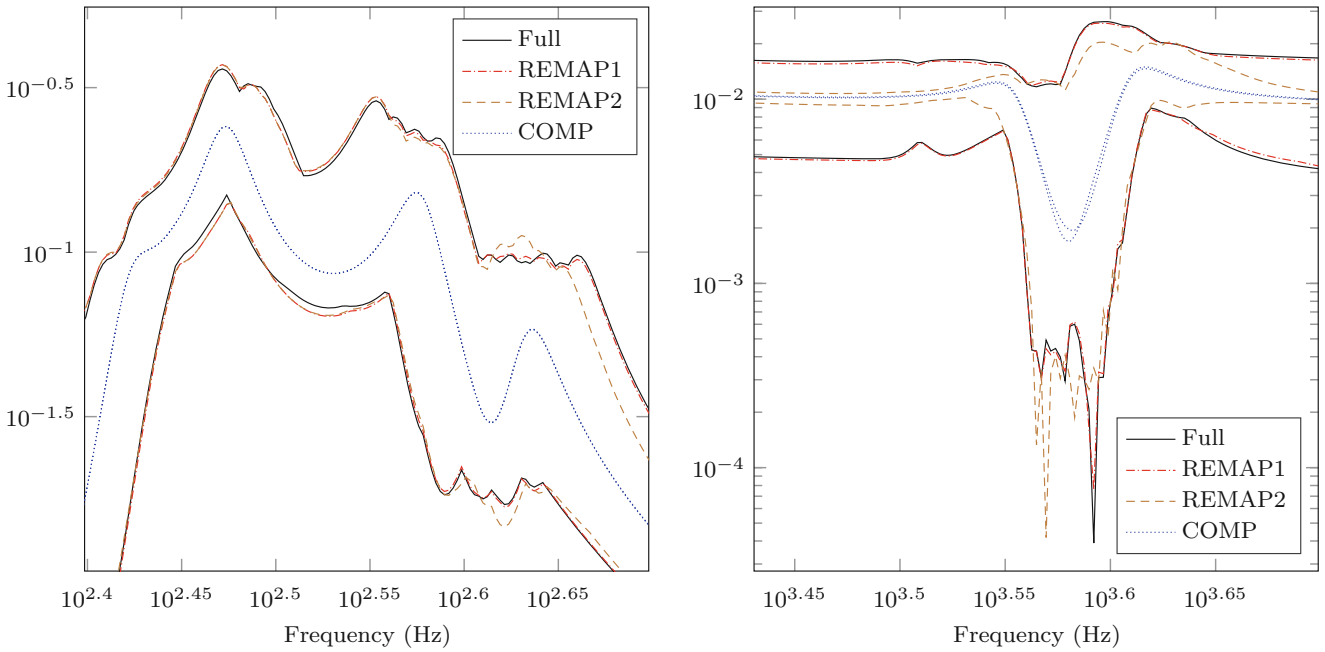


Fig. 19.4 Magnification of two regions of interest from the FRF in Fig. 19.3

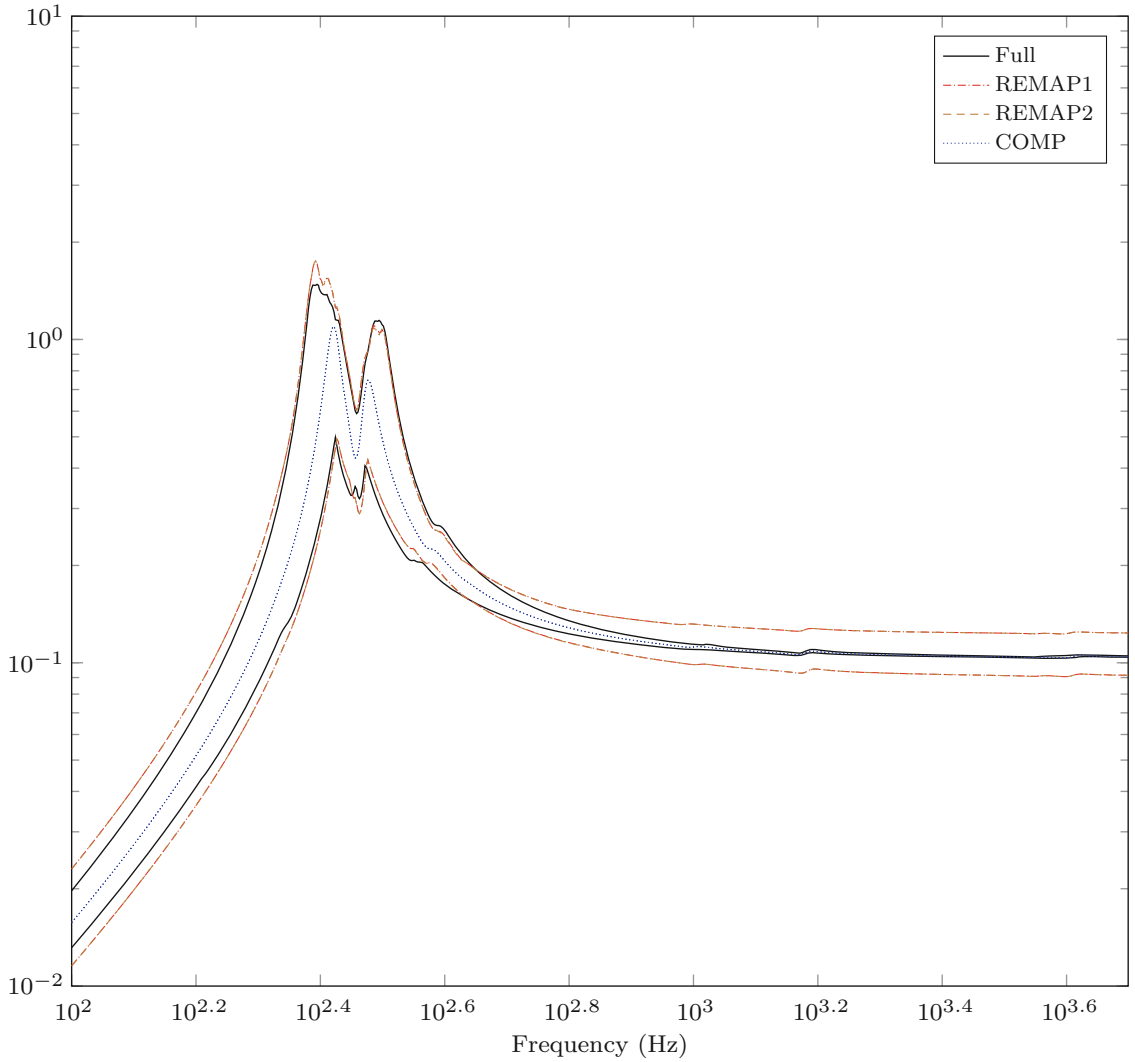


Fig. 19.5 Drive point FRF

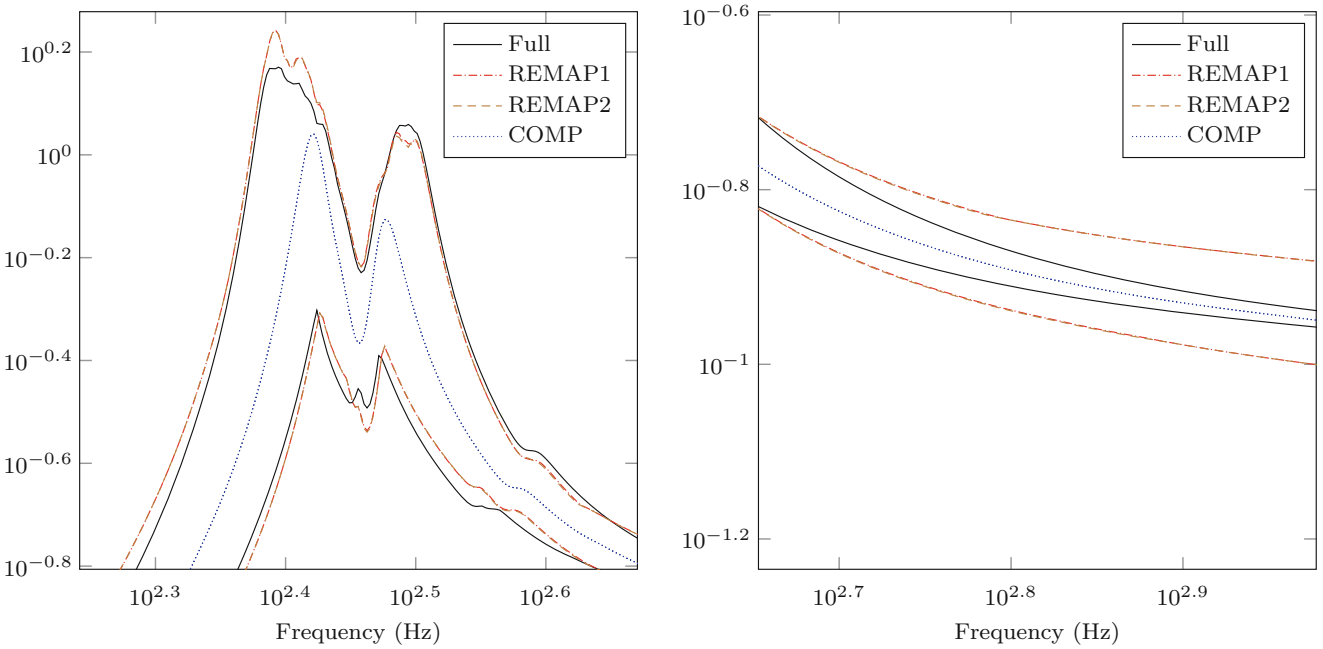


Fig. 19.6 Magnification of drive point FRF in frequency ranges of interest

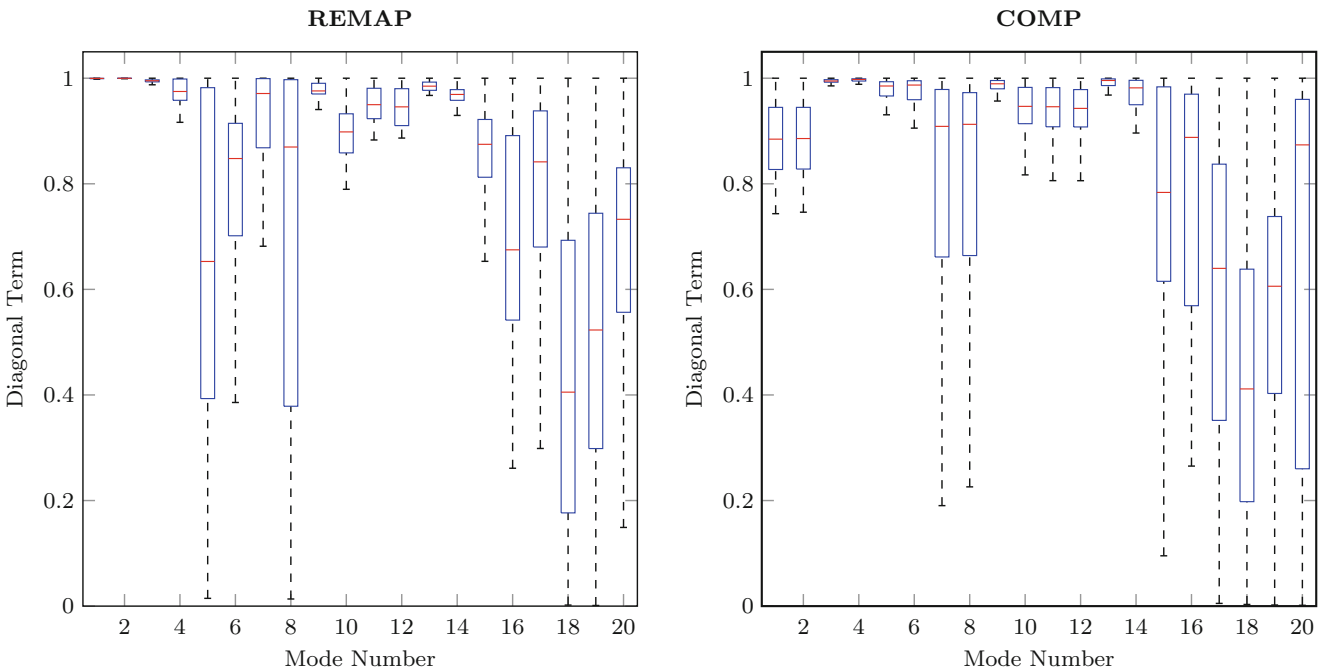


Fig. 19.7 Statistical analysis of the diagonal terms from the cross-orthogonality check between the REMAMP and COMP techniques compared to the full Craig-Bampton reduction

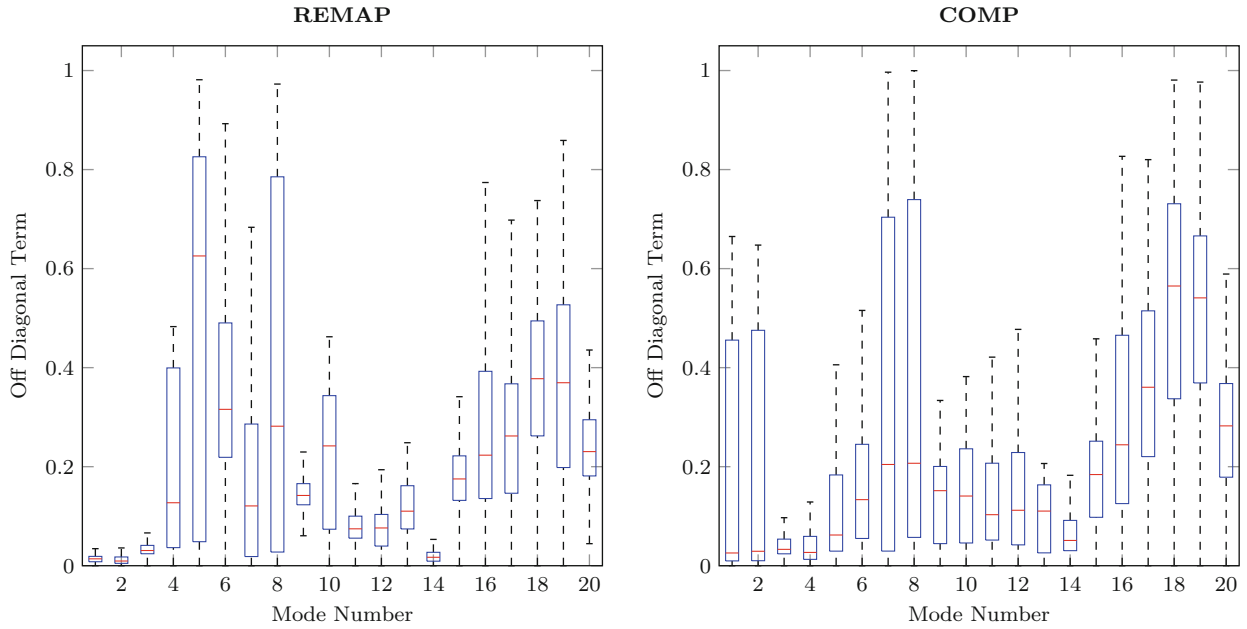


Fig. 19.8 Statistical analysis of the maximum off diagonal terms from the cross-orthogonality check between the REMAMP and COMP techniques compared to the full Craig-Bampton reduction

19.5 Conclusions

This paper presents three techniques to increase the computational efficiency when creating many Craig-Bampton reductions of the same component for UQ studies. The REMAP1 technique, which retains nominal fixed-interface modes and perturbs the constraint modes, the REMAP2 technique, which uses a fully nominal Craig-Bampton transformation matrix and only updates the component mass and stiffness matrices, and the COMP technique, which retains the nominal constraint modes and perturbs the fixed-interface modes. An example problem, which compared FRFs generated for a sample frame structure, shows that the REMAP1 technique accurately bounds the UQ problem, where as the REMAP2 and COMP techniques do not. A major concern for the COMP technique is how to properly perturb the component modal matrix such that permissible mode shapes are calculated and that also bound the UQ space. Both of these techniques require access to portions of the Craig-Bampton transformation matrix, which excludes them from being used if only the reduced mass and stiffness matrices are known. The choice of perturbing either the component or constraint modes is likely a problem specific choice that must be explored by the analyst prior to committing to a particular technique. It is suggested that a sensitivity study be carried out to determine the best technique for the system/problem of interest. Future work will explore these techniques on additional and more complex systems.

Acknowledgements Sandia National Laboratories is a multi-mission laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

1. Haldar, A., Mahadevan, S.: Probability, Reliability and Statistical Methods in Engineering Design. Wiley, Hoboken, NJ (2000)
2. Brake, M., Fike, J., Topping, S.: Parameterized reduced order models from a single mesh using hyper-dual numbers. J. Sound Vib. **371**, 370–392 (2016)
3. Hong, S., Epureanu, B., Castanier, M.: Joining of components of complex structures for improved dynamic response. J. Sound Vib. **331**, 4285–4298 (2012)
4. Hong, S., Epureanu, B., Castanier, M.: Next-generation parametric reduced-order models. Mech. Syst. Signal Process. **37**, 403–421 (2013)
5. Kammer, D., Krattiger, D.: Propagation of uncertainty in substructured spacecraft using frequency response. AIAA J. **52**, 353–361 (2013)

6. Kammer, D., Nimityongskul, S.: Propogation of uncertainty in test-analysis correlation of substructured spacecraft. J. Sound Vib. **330**, 1211–1224 (2011)
7. Momin, F., Millwater, H., Osborn, R., Enright, M.: A non-intrusive method to add finite element-based random variables to a probabilistic design code. Finite Elem. Anal. Des. **46**, 280–287 (2010)
8. Bampton, M.C., Craig Jr., R.R.: Coupling of substructures for dynamic analyses. AIAA J. **6**(7), 1313–1319 (1968)
9. Hurty, W.C.: Dynamic analysis of structural systems using component modes. AIAA J. **3**(4), 678–685 (1965)
10. Craig Jr., R.R., Kurdila, A.J.: Fundamentals of Structural Dynamics, 2nd edn. Wiley, Hoboken, NJ (2006)

Chapter 20
Parameterization of Large Variability Using the Hyper-Dual Meta-model

Matthew S. Bonney and Daniel C. Kammer

Abstract One major problem in the design of aerospace components is the nonlinear changes in the response due to a change in the geometry and material properties. Many of these components have small nominal values and any change can lead to a large variability. In order to characterize this large variability, traditional methods require either many simulation runs or the calculations of many higher order derivatives. Each of these paths requires a large amount of computational power to evaluate the response curve. In order to perform uncertainty quantification analysis, even more simulation runs are required. The hyper-dual meta-model is used to characterize the response curve with the use of basis functions. The information of the curve is generated with the utilization of the hyper-dual step to determine the sensitivities at a few number of simulation runs. This paper shows the accuracy of this method for two different systems with parameterization at different stages in the design.

Keywords Surrogate model • Uncertainty quantification • Hyper-duel • Hyper-duel meta-model • Distribution propagation

20.1 Motivation

During the design phase of a product, there are many engineering decisions that must be made. One of the most important decision is the physical characteristics of the product, such as physical dimensions and material choices. Some of these decisions are based on previous designs, available materials and equipment, and if the component will be combined with other products. Since these criteria can be a rough estimate on the design, the actual model can have a large amount of variability. Another important aspect of this variability is the nominal values of the design and the other modeling choices. For products used for the aerospace field, a nominal value, such as skin thickness, can be relatively small that implies that any small change can have a large variability in the response.

In order to characterize this variability, typically a Taylor series expansion of the response is used with sensitivities determined using the finite difference approach. When the variability becomes larger, more derivatives are required to get an accurate representation. This requires more simulation evaluations and can have high computational costs.

Once the response is clearly mapped, an uncertainty quantification (UQ) analysis is typically performed to determine the margin of safety or other reliability information due to the tolerances. This is normally performed with a Monte-Carlo (MC) propagation method. A MC analysis is performed by taking a random sample from the input distribution, such as physical distance or material property, and perform the simulation at that random input. This produces a distribution on the output that can be analyzed based on statistics. A typically MC analysis uses on the order of 10,000 random samples to produce an accurate distribution [1, 2].

As can be expected, this is a large computational cost. There are many methods available to reduce the computational burden of this design process. One commonly use reduction technique is to characterize the response due to an uncertain input by using a surrogate model or reduced order model (ROM). Typically, a surrogate model is a mathematic expression of the response and a ROM uses some physics of the system to reduce the size of the model. A example of a surrogate model is to use a Taylor series to characterize the response then to use the Taylor series in the MC analysis. A commonly used ROM

M.S. Bonney (✉)
University of Wisconsin-Madison, Madison, WI, USA
e-mail: msbonney@wisc.edu

D.C. Kammer
Engineering Physics, University of Wisconsin-Madison, Madison, WI, USA
e-mail: kammer@engr.wisc.edu

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_20

uses the modal information and the superposition of modes to reduce the system. The most commonly used ROM is the Craig-Bampton ROM [3–5]. Another commonly used reduction technique is to use a reduced sampling method such as Latin hypercube sampling [6–8] that separates the distribution into equally probable sections and takes a single point from each section. This can be used independently or combined with the ROM/surrogate model to reduce to total computational time.

This paper presents a new method to characterize the response of the system due to a large variability in the input parameters, the Hyper-Dual Meta-Model (HDM). The main advantage that this method presents, compared to more traditional approaches, is the usability for large variations and using the extra information of the derivatives with the use of the hyper-dual (HD) step. This paper is setup as follows: Sect. 20.2 gives the definition of the HDM in a general sense that uses a few simulation runs with the sensitivities at each simulation run given. In Sect. 20.3, several methods for determining the sensitivity of the output to a change in the input. Section 20.4 describes a couple methods of how to characterize the response by using different basis functions. Numerical examples are presented in Sect. 20.5. These examples include two different systems that are parameterized at different levels with varying complexity. Finally, some concluding remarks are presented in Sect. 20.6.

20.2 Hyper-Dual Meta-model Formulation

The main motivation for this technique is to create an accurate and inexpensive representation of the response of a system that undergoes a large change in the input parameters to be further used in sampling techniques such as MC or Latin hypercube. One of the main variability that causes a very nonlinear change in the response is a geometric variation, in particular for lengths that are small, such as in aerospace components. Typically, each change in geometric variation requires a re-computation of the finite element mesh. This requires a large amount of man-hours and computational power for each simulation run. Due to this, a numerical sampling method becomes computationally prohibitive. This lead to the development of the HDM.

The HDM uses the finite element code at a few selected values of the input variable. Along with the desired output, the information about the sensitivity of the output to the varied input is also needed. This sensitivity can be generated via many different methods. An in-depth discussion of the numerical derivatives is presented in Sect. 20.3. The main development is based on methods that determine the output and the sensitivities with a single code evaluation, such as the HD step. For the analytical example that is presented in this section, the analytical derivatives are used via beam theory.

Once the output and the sensitivities are known, the HDM characterizes this data with the use of basis functions. These functions can be any function that matches the output and derivatives at all of the code evaluations. Some of the possible basis functions are discussed in Sect. 20.4. This selection can be based on a known relationship or as a simple mathematic function such as a polynomial. For the analytical example, a simple matrix polynomial is used.

Another important engineering decision is where to parameterize the model within the simulation. This is an important decision and is based on how the simulations are performed. If the geometry is complex and the analysis is simple, this parameterization is more effective at the system level, such as the mass and stiffness matrices. One important aspect if this parameterization is used is that the matrices must be the same size and must have the same degrees of freedom (DOFs) that correspond to the same relative physical location on the structure. This eliminates the meshing portion of the analysis. The analytical example is parameterized at the stiffness matrix level. Another possible level of parameterization is at the output level. One major advantage of this parameterization is that the mesh for each simulation run can contain different number of DOFs. This is important for large geometric variation that change the system. One example is if a hole within the material that expands larger than the mesh size such that entire elements are removed. A disadvantage of this is that the solver must be able to calculate the derivative. All solvers are able to calculate derivatives with a finite difference approach, but if a more accurate result is required, the solver must be able to perform in either multi-complex math or HD math. This parameterization is expected to be more accurate, but typically require special solvers.

20.2.1 Analytical Example

In order to show the core concept of this method, a simple analytical example of a cantilever beam is presented. This system can be seen in Fig. 20.1. The variability expressed in this system is the total length of the beam and the desired result is the tip deflection due to a constant load P applied at the tip.

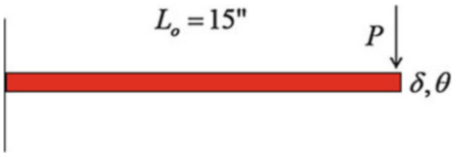


Fig. 20.1 Example cantilever system

Due to the simplicity of this system, an analytical expression for the tip displacement for a linear isotropic material is used. The tip deflection can be determined by

$$EI_x \begin{bmatrix} 12/L^3 & 6/L^2 \\ 6/L^2 & 4/L \end{bmatrix} \begin{pmatrix} \delta \\ \theta \end{pmatrix} = \begin{pmatrix} P \\ 0 \end{pmatrix}, \tag{20.1}$$

where E is the Young's modulus, I_x is the area moment of inertia, L is the length of the beam, δ is the tip deflection, and θ is the slope at the tip. This is thought of as a classic linear system of $KX = F$. Since the tip deflection is known analytically, the derivative with respect to the changing length is also known analytically.

This system contains an uncertain length of the beam. For this example, the beam length is described by the interval of [10, 20] in. Since the response is very nonlinear with respect to the length of the beam, along with the large variability, the traditional Taylor series require many higher order derivatives to be accurate. This leads to the implementation of the HDM. For this example, the stiffness matrix is evaluated at three different lengths, 10, 15, and 20 in., and the first derivatives are also calculated at each of those lengths. This information is used to determine the HDM using a polynomial fit that matches the stiffness matrix and the derivative at each length. The stiffness matrix is expressed as

$$K = K_0 + K_1\gamma + K_2\gamma^2 + K_3\gamma^3 + K_4\gamma^4 + K_5\gamma^5, \tag{20.2}$$

with K_i are matrix coefficients and γ is a non-dimensional length. This is defined as $\gamma = \frac{L-L_0}{5}$ for this system, such that the maximum and minimum have values of unity and negative unity respectively. The reason for this is for numerical conditioning for determining the matrix coefficients. These are determined by solving the linear equation of

$$\begin{bmatrix} I & -I & I & -I & I & -I \\ I & 0 & 0 & 0 & 0 & 0 \\ I & I & I & I & I & I \\ 0 & I & -2I & 3I & -4I & 5I \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & I & 2I & 3I & 4I & 5I \end{bmatrix} \begin{bmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \\ K_4 \\ K_5 \end{bmatrix} = \begin{bmatrix} K(L_{-1}) \\ K(L_0) \\ K(L_1) \\ K'(L_{-1}) \\ K'(L_0) \\ K'(L_1) \end{bmatrix}, \tag{20.3}$$

with I being the identity matrix, K' being the derivative of the stiffness matrix with respect to the non-dimensional length γ, and L_γ is the length associated with the non-dimensional length γ.

The linear system in Eq. (20.3) can then be solved for the unknown matrix coefficients. Those coefficients are used in Eq. (20.2) for the stiffness matrix then used to determine the tip deflection at any given length. This method is compared to a traditional Taylor series with the first two derivatives used at the nominal length. The results are shown in Fig. 20.2. This shows that the HDM very closely matches the truth data while the Taylor series is accurate close to the nominal length but becomes very inaccurate for large changes in length.

20.3 Determining Parameter Sensitivity

One of the main differences between the HDM and other response surface methods is with the additional information of the parameter sensitivity at each data point. The calculation of this parameter sensitivity can be calculated via multiple methods. Each method has its own advantages and disadvantages. As can be informed by the naming of this method, HDM,

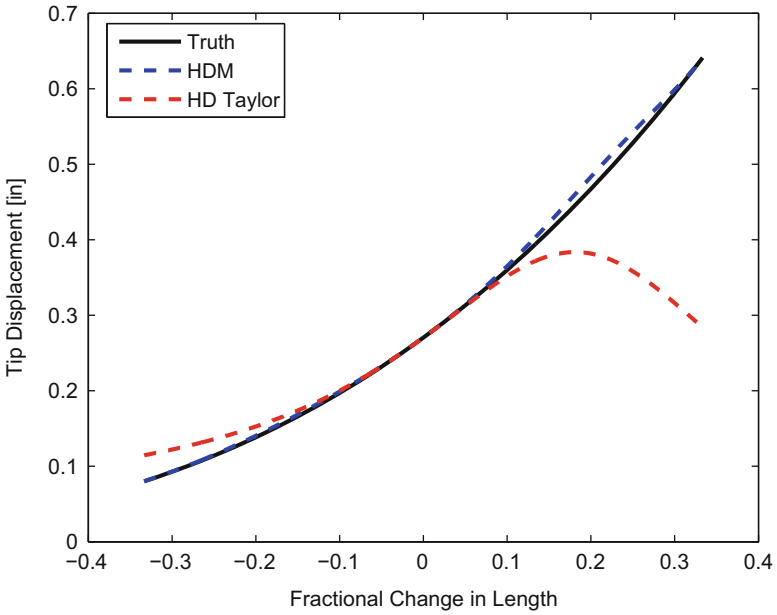


Fig. 20.2 Tip deflection as a function of length

the preferred method is the HD step. This allows for exact calculations where the other methods are approximations. The methods explained in this section are not an exhaustive study of the techniques, but does give a general overview of the basic and most commonly used methods.

20.3.1 Finite Difference

The simplest and most commonly used method for determining gradients is the finite difference approach. This is based on the Taylor series expansion about a design point. The Taylor series can be written as

$$f(x+h) = f(x) + hf'(x) + \frac{h^2 f''(x)}{2!} + \cdots + \frac{h^n f^{(n)}(x)}{n!} + \ldots \tag{20.4}$$

where x is the design point, h is a perturbation from the design point, $f(.)$ is the functional such as a polynomial or finite element code, and $f^{(n)}(.)$ is the nth derivative of the function. This infinite series can be truncated to give an approximation, but this introduces truncation error. The magnitude of this error varies based on the size of the perturbation and the order of the truncation.

The finite difference scheme uses this expansion at multiple perturbation points to estimate values of the derivatives. This assumes that the perturbation is small such that the series can be truncated. These multiple evaluations are then combined in a linear combination to eliminate the undesirable components of Eq. (20.4). The linear combination of the evaluations introduces subtractive errors when the perturbation is small. Subtractive error is the error associated with the bit resolution of the computer. Since the computer can only store sixteen digits of a number, for any two numbers that are very close in value, the computer interprets the numbers as the same and results in a zero difference.

There are three main categories that finite difference methods can be categorized as: forward, backwards, and central. The forward methods use perturbations where h is positive, while the backwards methods use perturbations where h is negative. Central difference approaches uses both positive and negative perturbation. The selection of which category to use involves engineering judgment of the function, number of function evaluations, and desired error. One important consideration is the continuity of the design space. If the design point is close to the discontinuity, then a directional finite difference method should be used. This should be noted such that any perturbation taken is still within the desired sample space. One possible example is a specified length that cannot physically have a negative distance, so if the distance is small, a directional derivative can be more efficient and accurate.

For a specific amount of function evaluations, the central difference methods tend to have the best accuracy. The first derivative using two function evaluations can be determined by

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + \mathcal{O}(h^2). \tag{20.5}$$

In Eq. (20.5), the Taylor series is truncated to the second derivative. The notation of $\mathcal{O}(h^2)$ represents that the infinite series is truncated for terms that contain higher orders of the step size. Any term that contains h^3 or higher in Eq. (20.4) is assumed to be small and is neglected. This determination only requires two function evaluations at $(x + h)$ and $(x - h)$. If more function evaluations are used, the truncation error can decrease.

Since the derivative requires taking the difference of two different function evaluations, subtractive error is introduced when the step size becomes small. There is a trade-off between decreasing the truncation error and the subtractive error. This trend can be seen later in the comparison section.

Higher order derivatives can also be found with similar methods. By using higher order derivatives, larger steps in an optimization routine can be used and thus account for larger variability on the design parameters. A second order central difference for the second derivative can be calculated as

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} + \mathcal{O}(h^2). \tag{20.6}$$

Using the HDM typically only uses the first two derivatives at each design point. Determining the coefficients for any arbitrary derivative and arbitrary order of truncation error can be calculated with the algorithm developed in [9]. This gives both the algorithm and the solution of the algorithm for the first four derivatives at several orders of truncation error for both central difference and directional difference.

This technique is straight forward and requires no special solver to determine the gradient of a functional. Current finite element codes, such as Ansys and Patran, use this formation to perform sensitivity analysis. While this method does not give the most accurate solution, it is a robust method that is simple to implement in current commercial finite element code.

20.3.2 Complex and Multi-complex Step

One way to eliminate the subtractive error that the finite difference approach introduces is to determine the gradient by using a single function evaluation. This can be accomplished by using a perturbation step in the imaginary direction [10, 11]. By using an imaginary step, Eq. (20.4) can be rewritten as

$$f(x+ih) = f(x) + ihf'(x) - \frac{h^2 f''(x)}{2!} + \cdots + \frac{(ih)^n f^n(x)}{n!} + \ldots \tag{20.7}$$

where i is the imaginary number defined as $i^2 = -1$. The determination of the first derivative can be calculated with only a single function evaluation thus eliminating the subtractive error. There is still an induced error due to the truncation of the Taylor series. The first derivative can be calculated as

$$f'(x) = \frac{Imag[f(x+ih)]}{h} + \mathcal{O}(h^2). \tag{20.8}$$

Compared to the finite difference approach, the same truncation error is introduced but the complex step does not contain the subtractive errors and thus produces a more accurate calculation of the first derivative.

The second derivative using the complex step introduces a subtractive error. This is due to the fact that the real component of the expansion contains both the nominal function value and the second derivative without anyway to differentiate between the values. In order to determine the second derivative without subtractive error, a more complicated expression is needed. With two function evaluations, the second derivative can be calculated as

$$f''(x) = \frac{2(f(x) - Real[f(x+ih)])}{h^2} + \mathcal{O}(h^2). \tag{20.9}$$

While this derivative requires two function evaluations, using the finite difference approach for the second derivative requires three function evaluations. The fewer function evaluations implies less subtractive errors since there is less algebra in the expression.

The complex step shown contains a step purely in the imaginary direction. However, there are methods that use a combination of the finite difference and the complex step. This is presented in [12] for cases of the step being a complex number. When the step is purely in the imaginary direction, this method reduces to the complex step method. The same can be said when the step is in the real direction; the method reduces to the finite difference method. This combination allows for a reduction in the induced errors relative to either method individually but requires more code evaluations.

One current method of determining the higher order derivatives with complex numbers is with the use of multi-complex numbers. These are slightly different than quaternions and are able to produce higher derivatives will less error. The methodology was first introduced in [13] explaining the mathematics. Millwater has introduced this mathematics into a complex step analysis in [14] and into a time integrator in [15]. The multi-complex numbers are similar to HD in several aspects. These introduce independent imaginary numbers. While this is the same as quaternions, these are communicative such that $i_1 i_2 = i_2 i_1$. This simple difference allows for calculation of the second derivative to be calculated with a single code evaluation, which is impossible with quaternions. Using multi-complex numbers eliminates the subtractive error in the determination of the higher-order derivatives, but it still is subject to truncation error. Typically a step of 10^{-16} is used, but is still problem dependent and thus needs to be verified that truncation error is small.

The use of the complex step allows for evaluation of the first derivative with a single function evaluation. This method, however, requires multiple function evaluations for the second derivative. With the single evaluation, the model solver must be able to handle complex number and the arithmetic associated with this. This may be difficult with some commercial software but is fairly simple with programs such as Matlab. The complex step still introduces truncation error and sometimes subtractive error. There are some techniques that combine the complex step and finite different methods, but not explicitly discussed. In order to calculate higher order derivatives with a single code evaluations, multi-complex numbers can be used. Multi-complex numbers are very similar to HD in the application, but are less accurate since there is still truncation error.

20.3.3 Hyper-Dual

The study of HD numbers originated in 2011 in [16]. It is used for optimization techniques in [17, 18] and further developed in [19]. The use of HD numbers within finite elements along with substructuring is analyzed in [20]. HD numbers are not specified to a certain number of non-real parts, but it is discussed similar to the quaternions with two independent directions and one dependent direction. There is a clear possibility for expanding the method to more independent direction, but the programming is not as straight forward.

With the use of HD numbers, the Taylor series becomes

$$f(x + h_1\epsilon_1 + h_2\epsilon_2 + 0\epsilon_{12}) = f(x) + h_1 f'(x)\epsilon_1 + h_2 f'(x)\epsilon_2 + h_1 h_2 f''(x)\epsilon_{12}, \tag{20.10}$$

where ϵ_k is an independent dual number and ϵ_{kl} is the dependent dual number. One main advantage of using HD numbers is the commutative property of these numbers. This property can be expressed as $\epsilon_1\epsilon_2 = \epsilon_2\epsilon_1 = \epsilon_{12}$. Using HD numbers for a single uncertain parameter allows for the determination of higher order derivatives. The second derivative can be determined as

$$f''(x) = \frac{\epsilon_{12} Part[f(x + h_1\epsilon_1 + h_2\epsilon_2 + 0\epsilon_{12})]}{h_1 h_2}. \tag{20.11}$$

This derivative is by definition exact. There is no truncation error of the Taylor series and is not subject to subtractive error. One interesting aspect is that the accuracy of the derivative is not dependent on the step size due to the lack of truncation and subtractive errors.

Extending this concept to more dimensions is not theoretically complex. This would only involve including more non-real dimensions, such as ϵ_3. Including this also involves programming the cross directions, ϵ_{13}, ϵ_{23}, and ϵ_{123}. These extra directions increase the computational time and storage requirements to perform the calculations. The current programming is classified as either two or three independent dual numbers. Current research is being conducted to program an arbitrary number of independent dual numbers.

For the most part, the algebra using HD numbers is similar to multi-complex numbers. This allows for simple programming in Matlab. Some examples of this algebra and functions are given in [16]. All of the analysis done in this proposal uses Matlab with some validation and output from other commercial codes. More complicated functions require some extra considerations. Some of the functions that are used include but not limited to: square root, inverse, sort, logical arguments, and eigen analysis. A couple of the required functions that are used were derived theoretically while others are reduced to the basic algorithm and let the elementary functions, such as add and multiply, perform the calculations. The extension of this method to more directions is simple for functions that use the original algorithm but is more complex for functions that are derived theoretically.

In Matlab, both the second and third order HD numbers are programmed. Along with the basic algebra, several functions are also programmed. The speed of running these functions is based heavily on the skill of the programmer and the platform. Most of the functions used by the HD numbers are programmed by engineers and not professional programmers. It is believed that some improvements to the programming can be made. Due to this fact, the main criterion that is used is the accuracy of the analysis. Another important criterion is time to perform the analysis, but is subject to programming efficiency and thus is not always reported. For some instances, both are reported but the main criterion is the accuracy of the technique.

20.3.4 Comparison of Methods

Three different methods are explained in the section: finite difference, complex step, and HD step. Each of these methods has its own advantages and disadvantages. To compare these methods, the accuracy of the first derivative is analyzed. One of the main considerations of using any of these methods is the step size. For two of the methods, the Taylor series is truncated based on step size. This implies that the accuracy of the truncation is highly dependent on the magnitude of the step size.

To examine the error on the derivatives, a polynomial function is used. This choice allows for quick calculations and a known true value of the derivative. The polynomial is not a commonly used output function, but does show the trend and is able to show the trade-off between the truncation and subtractive errors along with the relative errors between the three methods. A normalized error is used and the error the first derivative can be seen in Fig. 20.3.

The first trend is for step sizes of 10^{-4} or larger, the complex step and the central finite difference have the same errors. One important comparison is the accuracy of the HD step. The error is considered a numerical zero. Since the HD step is not subject to truncation error or subtractive error, the total error does not depend on step size and is numerically zeros for every

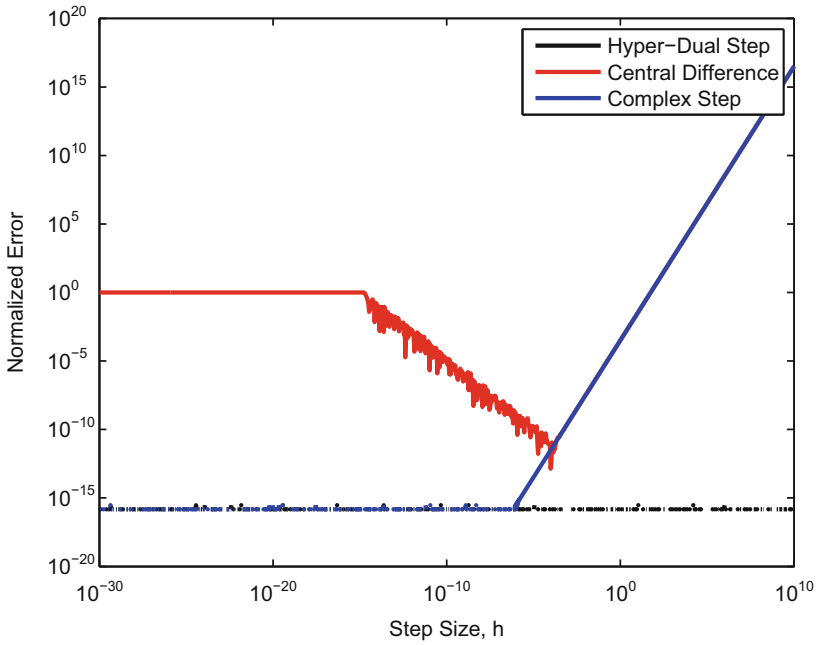


Fig. 20.3 First derivative error as a function of step size

step size tested. Another important trend that is noticed is on the central finite difference. For large step sizes, the main error is the truncation error, but as the step size decreases, the main error is the subtractive error. If the step size is smaller than around 10^{-15}, the difference between the function evaluations are zero due to the bit resolution.

20.4 Selection of Basis Function

One of major decision in the use of the HDM is what basis function(s) to use in the surrogate. While there are many candidate functions, some are discussed here and one is implemented. The first basis function is the use of a polynomial function. This is a classic function that is simple to generate and evaluate. The current use of this polynomial is to match the data from the full simulation runs exactly. This forces the surrogate model to match both the nominal value and derivatives at each data point. Using this type of formulation increases the accuracy of the surrogate, but can require many unknown matrix coefficients. One simple example of how to use this polynomial is shown in Sect. 20.2 in the analytical example. This uses a polynomial that matches the values and derivatives. Another interesting aspect of this example is the use a dimensionless parameter γ. The main reason for this is the numerical conditioning of the equations used to determine the matrix coefficients. If the uncertain parameter has a nominal value that is small, some numerical issues arise during the inversion.

One other possibility that can be used is the match the data in a least-squares sense. This is not recommended since it doesn't use the information gathered in the high-fidelity model. One possible use of a least-squared analysis is with the higher order derivative information. It is possible to match the nominal value and the first derivative exactly, then match the second derivatives in a least-squares sense. This has not been implemented, but is conceptually possible.

While this formulation of a polynomial is simple, there are other possible basis functions to use. Another possible basis function is the use of orthogonal polynomials. There are many possible orthogonal polynomials, such as Legrange, Hermite, or Jacobi. This is similar to the formulation of polynomial chaos [21]. These would utilized a matrix coefficient for each polynomial to generate the HDM. Computationally, this would require more time for a single evaluation compared to previous use of polynomials, but can be more accurate. This formulation has not previously been implemented.

The functions discussed in this section are only a selected few candidate functions that are simple to use. These are by no means an extensive survey of functions, but show a couple examples of polynomial type of basis functions. Other possible classes of functions can be used. One example of this would be similar to a Fourier transform, where the response is described with the use of sine functions. Another class of functions would be that of splines or exponentials.

The previous basis functions are candidate functions that do not consider any information about the physics of the system. These can be used for any variable on any system. If some information about how the response changes with the uncertain input, this can be used to increase the accuracy of the system. As seen in the analytical example, the stiffness has a $1/L^3$ dependency. Selecting a basis function of negative polynomials will tend to be more accurate since it is matching the physics of the system.

20.5 Numerical Examples

In order to evaluate this new surrogate model, two different systems are evaluated. Each system contains a different type of uncertain parameter. The study utilizes the results from a high-fidelity code named SIERRA SD [22, 23]. This is a code that is focused on parallel computing and has been implemented in select versions to be able to evaluate structural dynamics problems using HD numbers. The implementation of HD numbers into SIERRA is done at an object overload level. This replaces the simple arithmetic within the code to accommodate HD numbers. An alternative method to implement HD numbers in commercial code is to use a matrix notation for the HD number. This notation has been partially implemented into Ansys by Millwater's research group along with multi-complex numbers [14].

The first system that this method is tested on is the Brake-Reuß beam. This system is a simple beam that contains a lap joint and can be seen in Fig. 20.4. The use of this system is as a test bed for analyzing a simple joint. For this analysis, the uncertain parameter is the Young's modulus, E, of the aluminum material. While the response follows an expected trend, a simple polynomial is used for the basis function. In this system, the natural frequencies are used for the interpolation. This study was originally used in [24] as a comparison of different parameterized ROMs.

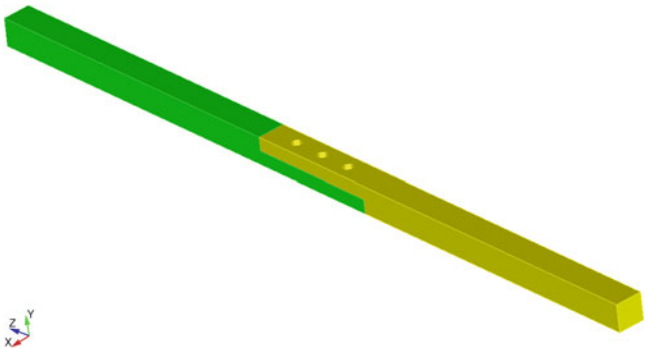


Fig. 20.4 Brake-Reuß beam schematic

20.5.1 Brake-Reuß Beam

For this system, two different ranges of variability was used: one for small variability $E \in [60, 80]$ GPa, and one for large variability $E \in [50, 300]$ GPa. For both of these ranges, three code evaluations are used with up to the first three derivatives. Since this is the first practical application of the HDM, many engineering choices are explored. The two main choices that are explored are: location of evaluation points within the sample space, and the number of derivatives to use.

In order to evaluate the accuracy and cost savings of each range, a parameter sweep is performed. During this sweep, two main criteria are used: root-mean-squared (RMS) accuracy, and the computational time to perform the evaluation sweep. This type of validation check is important for new models, but it is not the only use of surrogate models. Another important use of this model is for distribution propagation. With a specified distribution on the input parameters, either from experimental data or engineering knowledge, the output distribution is desired. This is particularly useful when attempting to determine a failure criteria, such as a natural frequency approaching the driving frequency, or the tensile stress reaching the yield or ultimate strength. In order to test the HDM against other surrogate models, this distribution propagation is performed on the large variability scenario.

20.5.1.1 Small Parameter Sweep

The first result to show is the small variability parameter sweep. This sweep is performed on the range of $E \in [60, 80]$ GPa. The determination of this range was due to a believed Young's modulus of 70 GPa. This was proven incorrect as investigation of the Brake-Reuß Beam continued, but does show interesting characteristics for poor choice of evaluation points and number of derivatives. One important thing to note about these points is that they do not span the sample space well. They are clustered about the expected value very closely.

In order to compare the HDM to typical surrogate models, another surrogate model and a truth curve are used. The truth curve is generated with the high-fidelity finite element code at increments in the modulus of 0.1 GPa; this process was automated for convenience. A surrogate model used as a comparison is based off of the Taylor series expansion. This expansion uses the information about the derivatives at the nominal point to setup the basis for the expansion. The model uses the HD step at the nominal value. These derivatives are exact and therefor more accurate than a typical finite difference approach. These curves are presented in Fig. 20.5 for the first natural frequency using only the first derivative information and are shown as a difference between the surrogate models and the truth data. The vertical dotted lines represent where the HDM is performed at.

One of most interesting results for this small sweep comes when higher order models are used. The same first natural frequency can be seen in Fig. 20.6. For the range where the modulus is outside the evaluation points used to determine the HDM, large extrapolation errors occur. This is primarily thought to be due to the order on the polynomial used for the basis function. As more data is used in the HDM generation, more coefficients are used and requires higher order polynomials. This can be classified as over-fitting of the data.

While the qualitative results comparing the models and the truth data is useful, quantitative results are very useful to demonstrates which model is better to the casual observer. This data is available in Table 20.1. The RMS error measurement is a RMS of the difference between each model and the truth data for each value of the Young's modulus and for each of the first 14 natural frequencies. Time is measured as the time to generate the parameter sweep curve after the model is generated.

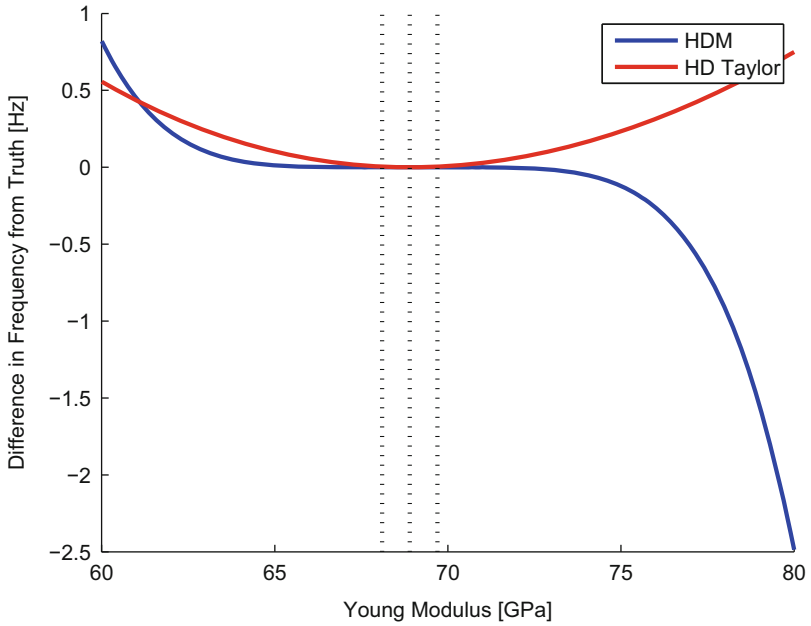


Fig. 20.5 First natural frequency for small sweep with 1st order models

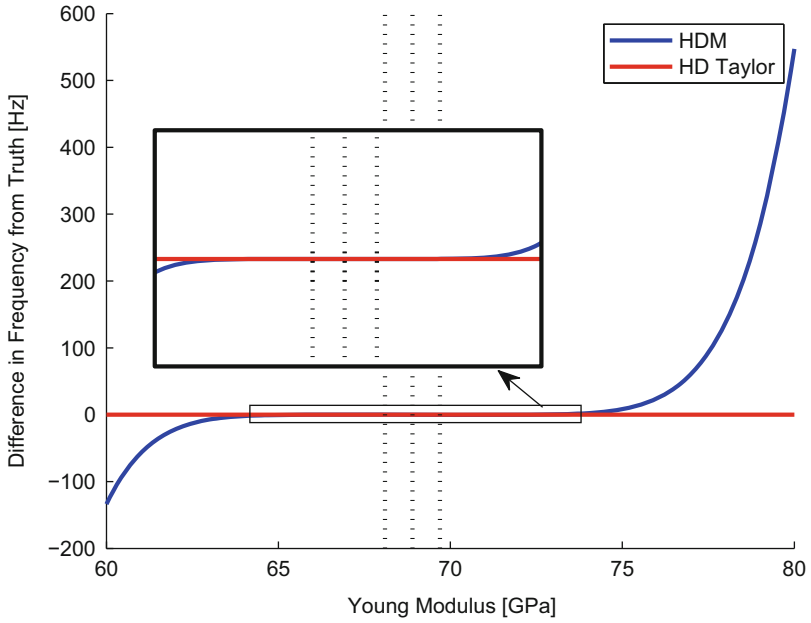


Fig. 20.6 First natural frequency for small sweep with 2nd order models

Table 20.1 Numerical results of accuracy and speed

Order	RMS error (HDM) [%]	RMS error (HD) [%]	Time (HDM) [s]	Time (HD) [s]
1	0.73	0.12	0.90	14.97
2	3727.8	7.53E−2	0.65	11.56
3	Large	6.44E−4	0.64	11.77

The time to generate the models are not measure, but the HDM takes more time to generate due to the matrix inversion and multiple code evaluations. This is thought of as a one time cost since the model can be used for both a parameter sweep to check the accuracy and to perform a MC analysis.

This example is a good case where a method is used inappropriately. The HDM is a proposed interpolation type surrogate model. Therefor the evaluation points for the model should span a majority of the sample space. The Taylor series surrogate model is an extrapolation type model since it is based off of a single point. So to correctly compare these two methods, a new parameter sweep is performed.

20.5.1.2 Large Parameter Sweep

This new parameter sweep is performed on the range of $E \in [50, 300]$ GPa. One main difference between this sweep and the previous one is the location of the code evaluation. These new locations are close to edges of the sweep. This allows for an accurate comparison between the two surrogate models. After more investigation into the material of the beam, the nominal value of the Young's modulus is changed to 150 GPa. This reflects the material used of aluminum. The values of simulation runs for the HDM is selected to by 70, 150, and 230 GPa. This is selected more on the lower end of the range but is centered about the nominal value and the HD Taylor series is selected at the nominal value.

The first comparison for this sweep is the first natural frequency. This can be seen in Fig. 20.7 along with the truth data. The HDM almost matches the truth data exactly except for at the large end of the range.

The same trend can be seen for higher order models and other natural frequencies and therefor not shown. Figure 20.7 also shows the location of the function evaluations. Since these are farther out to sample the entire space, the accuracy of the HDM has greatly increased compared to the small sweep. A quantitative result of the accuracy is given in Table 20.2. This shows that the second order model is the most accurate for the HDM. The over-fitting of the data for the higher order model creates a reduction in the model accuracy since there is some extrapolation of the model compared to the simulation runs.

These results, compared to the small sweep, show that the HDM is more accurate and requires less time to compute. This also shows that a smart selection of the evaluation points is required for the HDM to be the most effective.

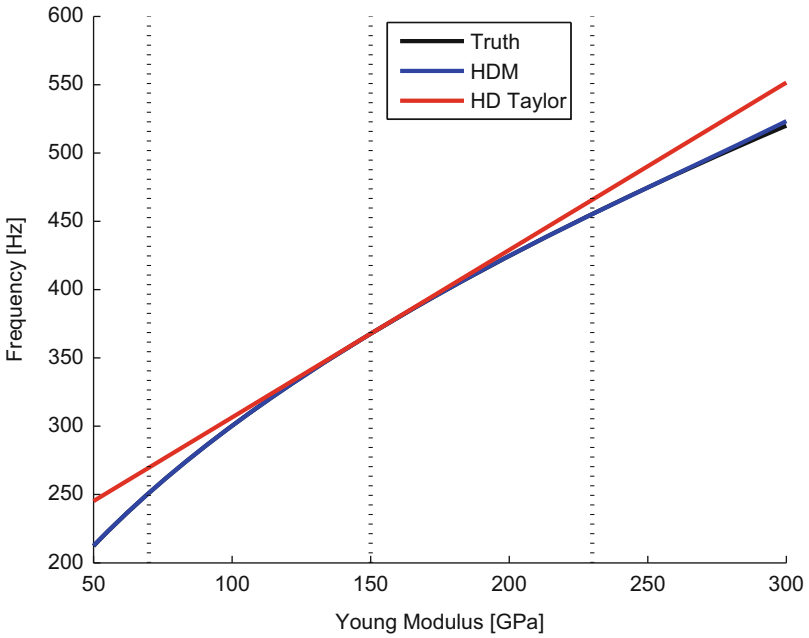


Fig. 20.7 First natural frequency for small sweep with 1st order models

Table 20.2 Numerical results of accuracy and speed for parameter sweep

Order	RMS error (HDM) [%]	RMS error (HD) [%]	Time (HDM) [s]	Time (HD) [s]
1	0.1274	4.3038	0.26	32.07
2	0.0524	1.4877	0.20	33.94
3	0.2681	0.6712	0.21	36.31

20.5.1.3 Distribution Propagation

A parameter sweep is typically a preliminary analysis for validation and model calibration. Once the model is calibrated, the sensitivity to the calibration is desired, typically done by assigning a distribution to the input parameters and propagating those distributions to the output. This tells the design engineer an idea on the tolerances that need to be specified when production begins. Along with tolerances, this information can also be useful for failure analysis and also non-technical information such as insurance or possible recalls.

For this analysis, a simple MC sampling analysis is used. With the selection of this sampling method, the number of samples must be selected. For convergence and accuracy, 100,000 samples are selected. Along with the number of samples, the distributions must also be chosen. This information can be chosen either from experimental data or from engineering judgment. For this system, the Young's modulus is chosen to be a log-normal distribution. This distribution has a support of $(0, \infty)$ and can never reach zero. The main reason for this is that the log-normal distribution spans the same range as the Young's modulus. For the log-normal distribution, two parameters are required, the mean and the coefficient of variation. The mean is chosen to be 150 GPa and the coefficient of variation is chosen to be 14%. This variability is chosen to fully test the range of the HDM that was shown to be accurate from the range of [50, 300] GPa.

One of the main choices in this type of analysis is what the output quantity is along with how to quantify the distribution of the output. For this section, the fundamental frequency is selected. Similar results can be seen for any of the first fourteen frequencies since the HDM and HD Taylor models are generated for each frequency independently, the performance of the model does not depend heavily on which natural frequency is selected.

In order to quantify the accuracy, two methods are used: qualitative distribution, and quantitative information on the statistical moments. The qualitative distribution can be seen in Fig. 20.8.

By looking at the distribution, there is only a slight difference between the distributions. The truth distribution is generated from a linear interpolation from the parameter sweep. Since the difference is difficult to see, a point-by-point difference between the methods and the truth distribution can be seen in Fig. 20.9.

The main take away from Fig. 20.9 is that the HDM is almost perfect while the HD Taylor model varies from the truth data more in the lower frequency range. While the qualitative results are good for a visual interpretation, numerical results are very helpful. For this numerical result, the first four statistical moments are determined from the distribution and the RMS of those first four moments and the first fourteen natural frequencies is computed. These results can be seen in Table 20.3. This shows that the HDM provides nearly perfect results at about 1% of the time compared to the HD Taylor model.

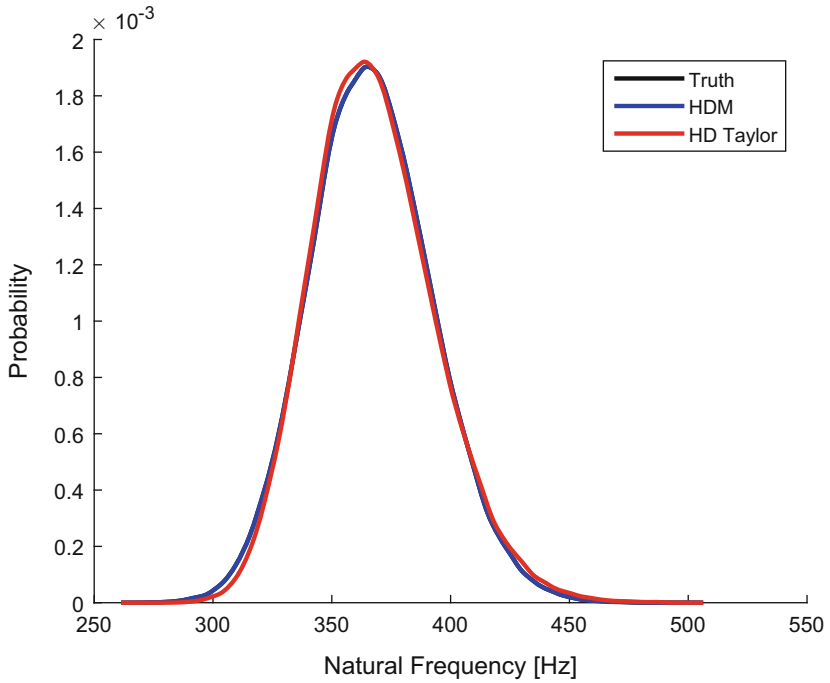


Fig. 20.8 Distribution of the fundamental frequency

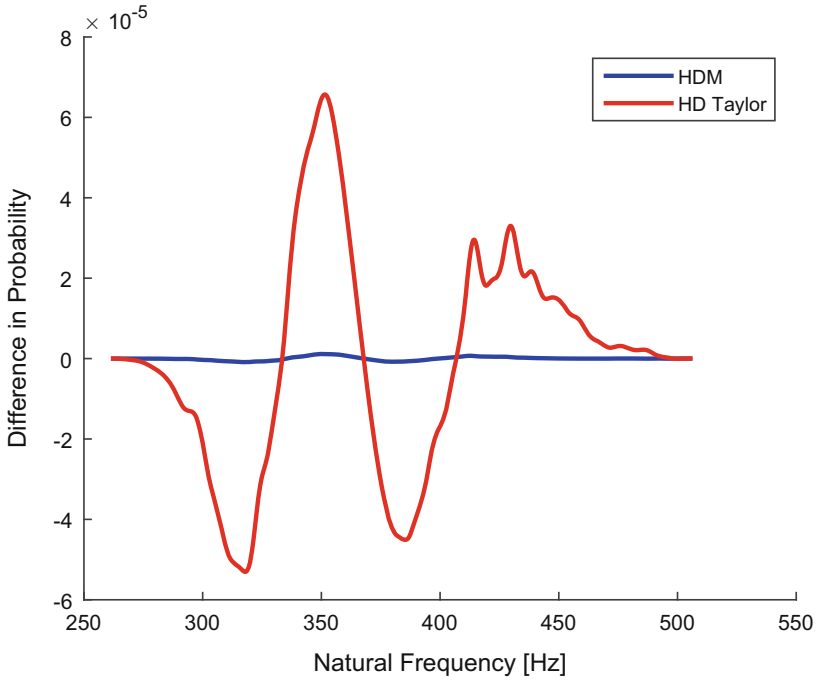


Fig. 20.9 Point-by-point difference of distribution of the fundamental frequency

Table 20.3 Numerical results of accuracy and speed for distribution moments

Order	RMS error (HDM) [%]	RMS error (HD) [%]	Time (HDM) [s]	Time (HD) [s]
1	0.0089	0.6952	30.4	3697
2	0.0003	0.0318	30.9	3721
3	0.0003	0.0149	30.7	3765

20.5.2 Geometric Change

A variation on the Young's modulus of 100% is not typical, but is necessary to show the usefulness of the HDM for that parameter. This is due to how the natural frequencies vary with the Young's modulus, a square root proportionality. In a more realistic system, geometric parameters are uncertain. This uncertainty can come from an unknown design or physical issues such as manufacturing tolerances. The relationship between the output and the geometric change is much more difficult to determine. In order to explore this relationship, a simplified system with a geometric change is presented. The system that is varied is presented in Fig. 20.10. The geometric variation of this system is the length of the appendage on the left side of the system. This nominal length is 12 in. and is varied by a substantial amount.

This system has been studied previously in [20, 25, 26]. In [20], this system explores the use of traditional HD numbers compared to the finite difference method to determine sensitivities. This research is an expansion of that research by applying the HDM to the system. In [25, 26], this system is treated as two substructures and the model form error is calculated using a non-parametric, maximum entropy approach. The system in Fig. 20.10 is generated using a custom-made Matlab finite element code. This allows for the generation of the system matrices in HD and provide up to the first three derivatives of the mass and stiffness. The code used is available upon request to the primary author.

The HDM is applied to this system at two different levels of the analysis and at two different variabilities representative of different uses. Firstly, the HDM can be applied at any level in the analysis. In this analysis, the HDM is applied at the matrix level and the output level. This is done to show the possibilities of the HDM and how it can be used based on the software available. Applying the HDM to the output of the system is the simplest to implement if the analysis is able to handle HD numbers. For the system used, the Eigen analysis is programmed in HD numbers, so this output can be done fully in HD and is typically much smaller in size than the full system. The other analysis is applied at the system matrix level. This applies the basis functions to the mass and stiffness matrix. As one can expect, this is a much larger than the output. This shows that if the solver is not programmed in HD numbers, the HDM can still be applied and can give reasonable results.

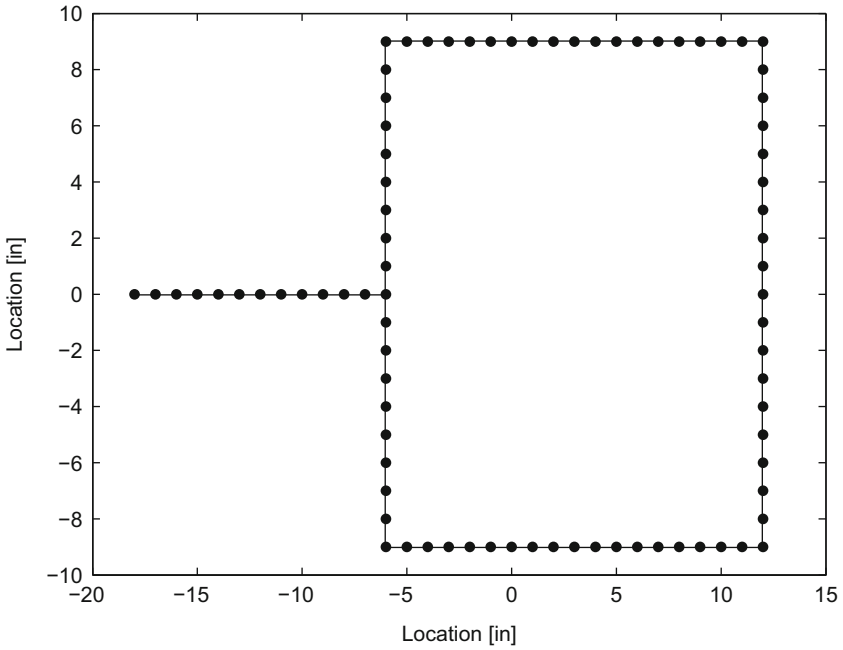


Fig. 20.10 Nominal geometric system

For instance, if the time history is desired, current programming does not allow for time integration. But if the basis function is applied before the integration, the system can be reduced to the real domain then sent through whatever solver is required. This provides the possibility of using advanced codes and solvers without a complete reprogramming of the algorithms. As a side note, if the algorithm is expressed simply, a redefinition of the elementary functions, such as plus and multiply, can be implemented to allow for easy conversion of the complicated solver.

This analysis is also applied for two different levels of variation, small and large. The small variation is a representative analysis for a production variation analysis. This would be representative of manufacturing tolerances and thus be able to determine reliability information such as failure probability. The other analysis is representative of a design stage. If the design engineer desires to vary geometric components to best match requirements, such as nuclear qualification or compartment size requirements, then the HDM can be used as a surrogate model to decrease the design time and thus reduce design costs.

20.5.2.1 Reliability Analysis

This type of analysis is based on a completed design process to quantify the effect of tolerances and defects on the final output. How this is implemented is with smaller variations. This variations can be interpreted as tolerance build-up or a defect if the end of the appendage gets struck and disconnects. Mathematically, this is defined as a length variation of 10%. The HDM uses three data points, one at nominal length of 12 in., and one at ± 1.2 in. At each of the data-points, the first two derivatives are computed. This allows for up to the second order HDM generation. The main output investigated is the fundamental frequency along with the second elastic natural frequency to investigate the scalability of the method. The results show two versions of the HDM: one where the parameterization is performed at the system level, and one parameterized at the eigenvalue level. In this simulation, the computational time for each version of the HDM is comparable. The majority of the computational time is within the finite element code.

The first result is the fundamental frequency of the system as a function of the length of the appendage using the first derivative information and can be seen in Fig. 20.11. First thing to note is the limits on the length of the appendage. The length is varied from 8 to 15 in. that indicates a 33.3% range.

In Fig. 20.11, the range within the evaluation points is zoomed in the center. The difference between the parameterizations and the truth data is small. This shows that either type of parameterization is accurate for this type of range. Once the model is extrapolated, the differences between the parameterizations becomes more apparent. This is better displayed as the a point-by-point difference in the fundamental frequency as shown in Fig. 20.12.

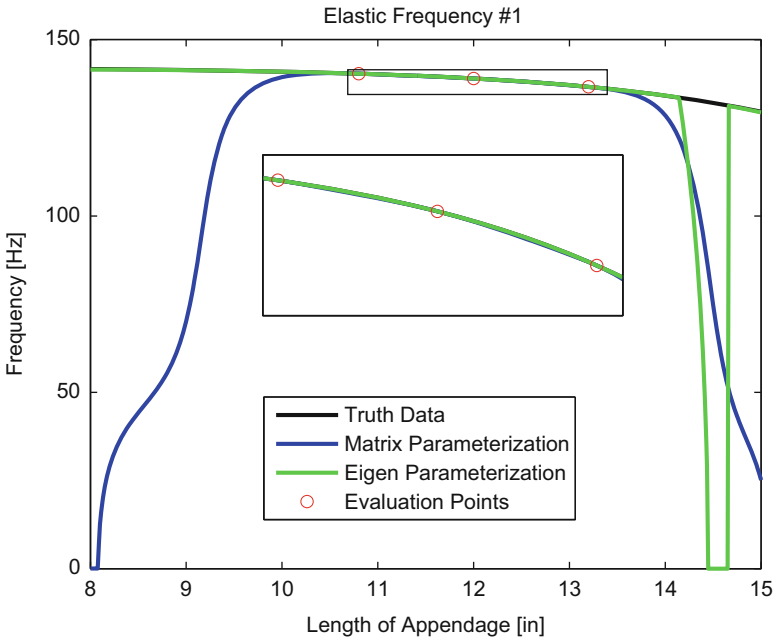


Fig. 20.11 Fundamental frequency of 10% variation

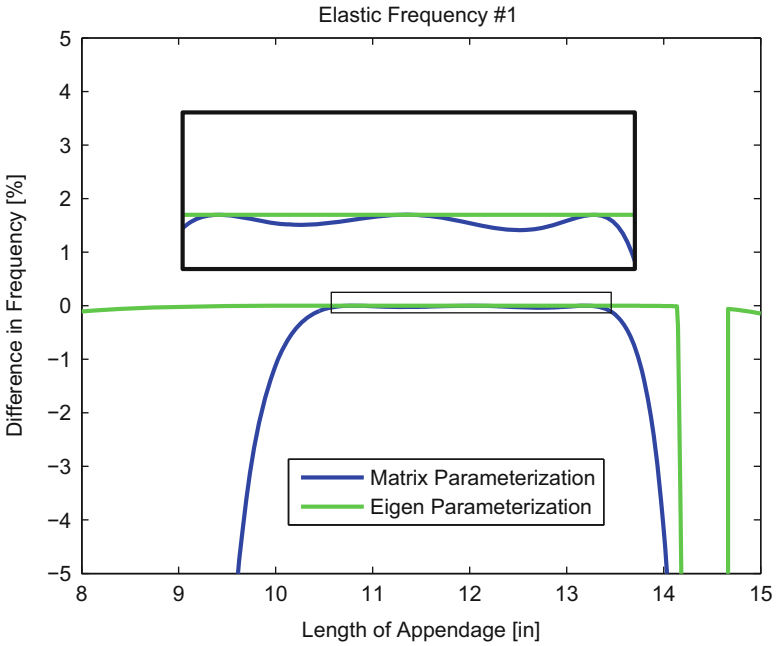


Fig. 20.12 Difference in fundamental frequency of 10% variation

This curve also shows a zoomed region to show the small differences within the model evaluations. The difference is very small, but does show that the Eigen parameterization is more accurate. As also can be seen is that the Eigen parameterization is valid for a much larger range than the system matrix parameterization. For a 95% accurate model, the matrix parameterization is valid from $[9.6, 14]$ in. and the Eigen parameterization is valid from $(< 8, 14.1]$ in. Around 14.5 in., there is some numerical issues that occurred within the model. The reason for this is being investigated, but is primarily believed to be due to ill-conditioning and mode switching.

Along with the fundamental frequency, the second elastic natural frequency is also evaluated. This was done to evaluate higher natural frequencies and to see if the more physics based matrix parameterization if more effective in the higher natural frequencies. The second natural frequency is found in Fig. 20.13.

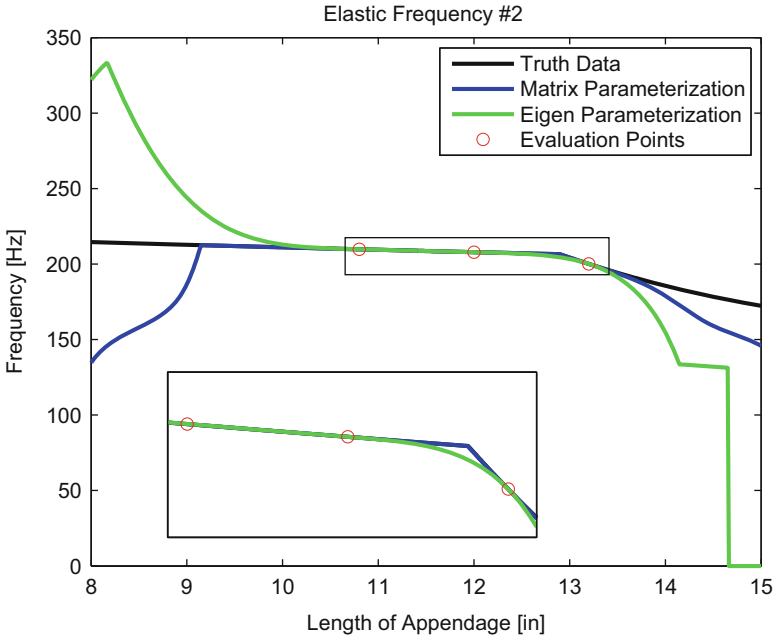


Fig. 20.13 Second elastic natural frequency with 10% variation

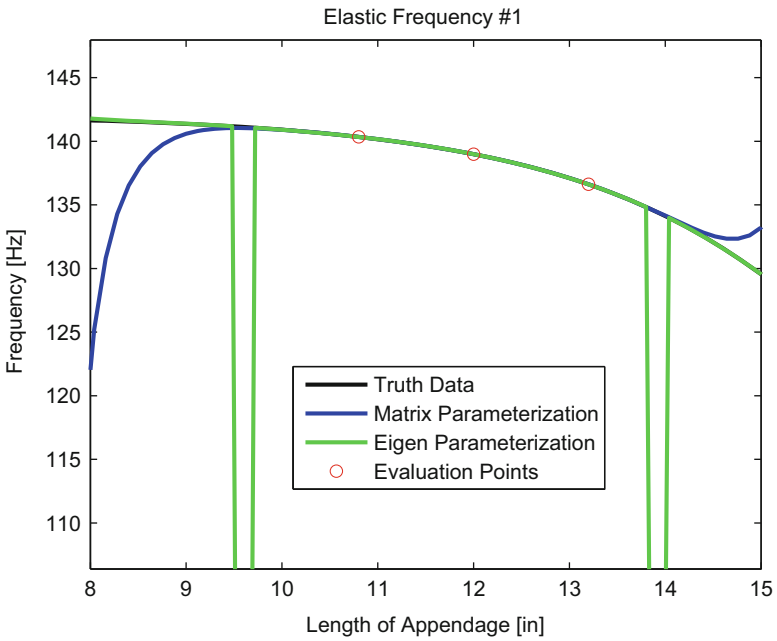


Fig. 20.14 Fundamental frequency using 2nd order HDM with 10% variation

One of the most interesting aspect of Fig. 20.13 is that the matrix parameterization produces accurate results for a larger range than the Eigen parameterization. For the matrix parameterization, the 95% accuracy range is [9, 14.1] in. and the range for the Eigen parameterization is [9.5, 13.7] in. Another interesting aspect is that the matrix parameterization is conservative over the accuracy range. This is not true for the Eigen parameterization, particularly in the shorter length appendages.

The results shown are for only taking the nominal values and the first derivative at each data-point, called the first order HDM. This is the most expected version of the HDM since this can be done for a single uncertain variable with the use of a simple complex step. Higher order models and multi-dimensional systems require the use of HD or multi-complex numbers to produce accurate derivatives. Using this extra information produces a more accurate model within the evaluation points. This accuracy can be seen in Fig. 20.14. One interesting aspect of this higher order is that there is more numerical errors

with the Eigen parameterization. This does not affect the matrix parameterization. One issue with the use of the higher order HDM is the possibility of over-fitting the data. This can be seen near the tails for the matrix parameterization. Over-fitting is thought of as applying higher order polynomial than is more than the physics entail. This is primarily seen when extrapolating outside of the evaluation points. The values within the evaluation points tend to get more accurate with more information since the information is implicitly contained via the derivatives.

Overall, for this type of analysis, the HDM performs excellent over the desired range. One interesting aspect found is that the extrapolation past the evaluation points is still reasonably accurate. For this amount of variation, 10% of the nominal length, both type of parameterization produce accurate results for length values change of at least 15%. This range is not uncommon for manufacturing tolerances, particularly for small nominal values. The accuracy of the model does depends on what parameterization is used, but for this range of values, there is only a small difference between the parameterizations. This selection depends heavily on the computational capabilities are available, if the solver is capable of evaluating the solution in HD or multi-complex math.

20.5.2.2 Design Analysis

For this design analysis, an assumption is made that the length of the appendage is unknown and therefor a large variation must be tested. The main difference between the two sections is the location of the evaluation points on the sample space. For this analysis, the variation is performed at $\pm 20\%$. This value was selected to explore a much larger sample space since the final design is not determined. The first result for this analysis is the fundamental frequency as a function of the appendage length. This can be seen in Fig. 20.15.

One of the main difference between the design analysis and the reliability analysis is the comparison between the two parameterizations. In the reliability analysis, both parameterizations yielded very accurate results within the evaluation points, while the difference between the two parameterizations appear in the design analysis. A point-by-point difference can be seen in Fig. 20.16.

The first difference between Figs. 20.12 and 20.16 is that the Eigen parameterization does not have a numerical issue around 14.5 in. for the length. This is believed to because there is an evaluation point around that range. One noticeable aspect in Fig. 20.16 is that the matrix and Eigen parameterization show different trends. The matrix parameterization still has a 95% accuracy range within the evaluation points. This range is the reason that 20% variation was selected for this design analysis. When the range is increased to 25%, the accuracy drops below 95% within the evaluation points.

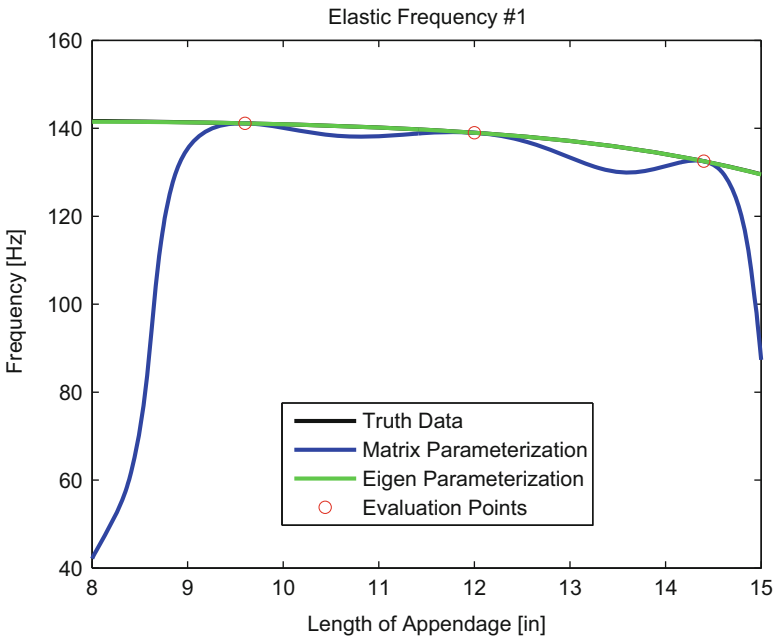


Fig. 20.15 Fundamental frequency with 20% variation

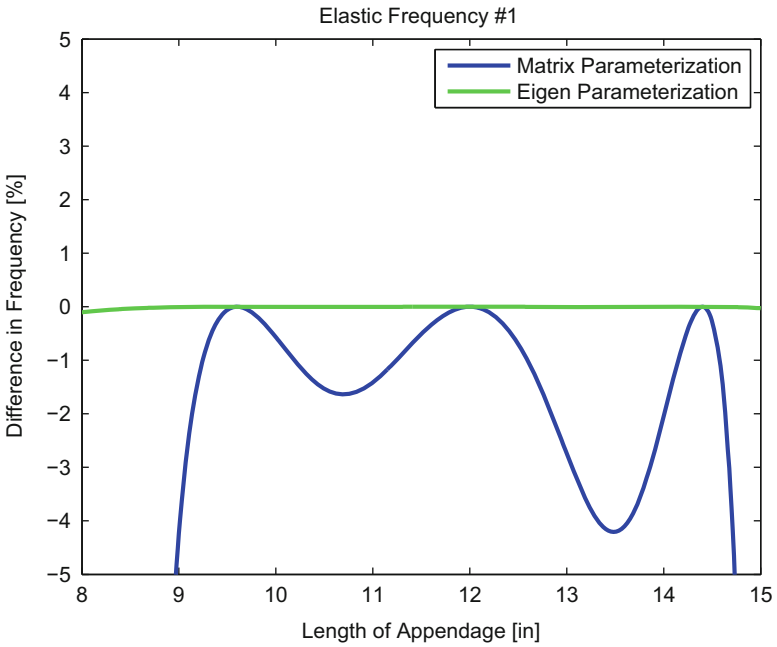


Fig. 20.16 Difference in fundamental frequency with 20% variation

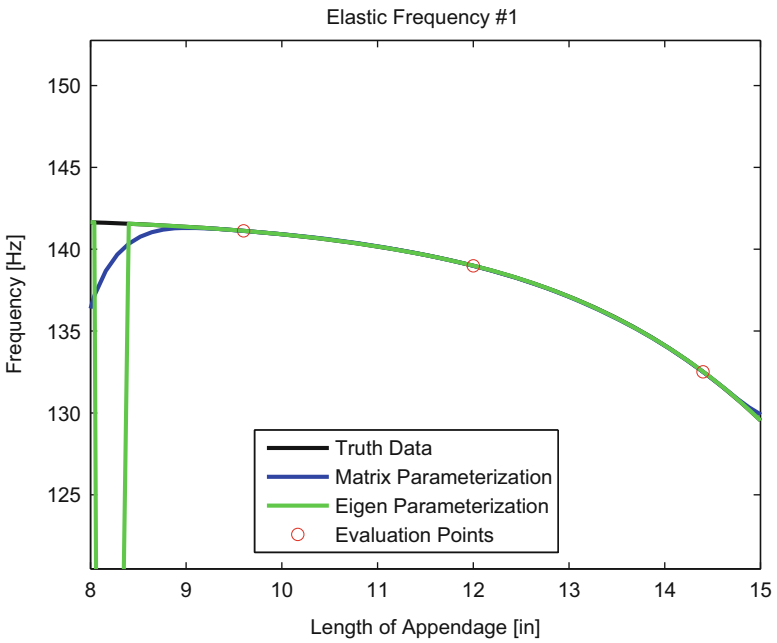


Fig. 20.17 Fundamental frequency using 2nd order HDM with 20% variation

The second order model improves the model compared to the first order model. This can be seen in Fig. 20.17. Using the second order model has more of an effect on the matrix parameterization compared to the Eigen parameterization. This reduces the error within the evaluation points greatly and also increases the accuracy of the extrapolation. The Eigen parameterization suffers from some numerical issues around lengths of 8.25 in.

For this design analysis, a large variation of the sample space is explored. Among the two parameterization, the Eigen level parameterization is the most accurate as long as the solution is available with HD or multi-complex numbers, but is not always available depending on the analysis. Using a higher order HDM improves the matrix parameterization greatly and allows a large possible range of variability for a given accuracy.

20.6 Conclusions

For systems that can contain a large variation, either in design phase or small nominal values, traditional Taylor series response surface methods are inaccurate or computationally expensive. The Hyper-Dual Meta-Model uses a select number of data points along with the information about the sensitivities to generate an accurate and inexpensive surrogate model. This surrogate model is most effective is the data points span the majority of the desired sample space, no matter the size of the sample space. The model can be applied at any point in the analysis and able to produce accurate results. This is explored using some numerical examples with both a geometric variation and a material variation. For the geometric variation, the parameterization is applied at both the system and the output level. These results show that both parameterization are accurate for a large region, although the Eigen parameterization is more accurate as long as the solver is able to handle hyper-dual numbers. The model presented in this paper is shown to be both accurate and computationally efficient as a surrogate model by using the information of the gradient to better represent the response surface with only a few selected data points.

Acknowledgements Sandia is a multi-mission laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

1. Metropolis, N., Ulam, S.: The monte carlo method. J. Am. Stat. Assoc. **44**(247), 335–341 (1949)
2. Shapiro, A., Homem-de Mello, T.: On the rate of convergence of optimal solutions of monte carlo approximations of stochastic programs. SIAM J. Optim. **11**(1), 70–86 (2000)
3. Craig, R.R., Bampton, M.C.C.: Coupling of substructures for dynamic analysis. AIAA J. **6**, 1313–1319 (1968)
4. Craig, R.R.: Coupling of substructures for dynamic analyses - an overview. In: 41st Structures, Structural Dynamics, and Materials Conference and Exhibit, Structures, Structural Dynamics, and Materials and Co-located Conferences (2000)
5. Kammer, D.C., Triller, M.J.: Selection of component modes for Craig-Bampton substructure representations. ASME J. Vib. Acoust. **188**(2), 264–270 (1996)
6. Helton, J.C., Davis, F.J.: Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab. Eng. Syst. Saf. **81**(1), 23–69 (2003)
7. McKay, M.D., Beckman, R.J., Conover, W.J.: Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics **21**(2), 239–245 (1979)
8. Michael, S.: Large sample properties of simulations using Latin hypercube sampling. Technometrics **29**(2), 143–151 (1987)
9. Fornberg, B.: Generation of finite difference formulas on arbitrarily spaced grids. Math. Comput. **51**(184), 699–706 (1988)
10. Martins, J.R.R.A., Sturdza, P., Alonso, J.J.: The connection between the complex-step derivative approximation and algorithmic differentiation. AIAA Paper **921**, 2001 (2001)
11. Martins, J.R.R.A., Sturdza, P., Alonso, J.J.: The complex-step derivative approximation. ACM Trans. Math. Soft. **29**(3), 245–262 (2003)
12. Lai, K.L., Crassidis, J.L.: Extensions of the first and second complex-step derivative approximations. J. Comput. Appl. Math. **219**(1), 276–293 (2008)
13. Lantoine, G., Russell, R.P., Dargent, T.: Using multicomplex variables for automatic computation of high-order derivatives. ACM Trans. Math. Softw. **38**(3):16:1–16:21 (2012)
14. Garza, J., Millwater, H.: Sensitivity analysis in structural dynamics using the ZFEM complex variable finite element method. In: 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, p. 1580 (2013)
15. Garza, J., Millwater, H.: Multicomplex newmark-beta time integration method for sensitivity analysis in structural dynamics. AIAA J. **53**(5), 1188–1198 (2015)
16. Fike, J.A., Alonso, J.J.: The development of hyper-dual numbers for exact second-derivative calculations. AIAA Paper **886**, 124 (2011)
17. Fike, J.A.: Multi-objective optimization using hyper-dual numbers. PhD thesis, Stanford university (2013)
18. Fike, J.A., Jongsma, S., Alonso, J.J., Van Der Weide, E.: Optimization with gradient and hessian information calculated using hyper-dual numbers. AIAA Paper **3807**, 2011 (2011)
19. Fike, J.A., Alonso, J.J.: Automatic differentiation through the use of hyper-dual numbers for second derivatives. In: Recent Advances in Algorithmic Differentiation, pp. 163–173. Springer, Berlin (2012)
20. Bonney, M.S., Kammer, D.C., Brake, M.R.W.: Fully parameterized reduced order models using hyper-dual numbers and component mode synthesis. In: Proceedings of the ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, p. 46029 (2015)

21. Wiener, N.: The homogeneous chaos. Am. J. Math. **60**(4), 897–936 (1938)
22. Edwards, H.C.: Sierra framework version 3: core services theory and design. SAND Report No. SAND2002-3616 (2002)
23. Reese, G.M., et al.: Sierra structural dynamics user's notes. Technical Report, Sandia National Laboratories (SNL-NM), Albuquerque (2015)
24. Bonney, M.S., Brake, M.R.W.: Determining reduced order models for optimal stochastic reduced order models. Technical Report SAND2015-6896, Sandia National Laboratories, Albuquerque, NM (2015)
25. Bonney, M.S., Kammer, D.C., Brake, M.R.W.: Determining model form uncertainty of reduced order models. In: Model Validation and Uncertainty Quantification, vol. 3, pp. 51–57. Springer, New York (2016)
26. Bonney, M.S., Kammer, D.C., Brake, M.R.W.: Numerical investigation of probability measures utilized in a maximum entropy approach. In: Uncertainty in Structural Dynamics, pp. 4307–4321 (2016)

Chapter 21
Similitude Analysis of the Frequency Response Function for Scaled Structures

Mohamad Eydani Asl, Christopher Niezrecki, James Sherwood, and Peter Avitabile

Abstract Full-scale testing of the large structures is time consuming and expensive. Scaled-down structures expedite and facilitate the testing process of the large structures. To design a scaled-down structure, a deliberate methodology should be used to ensure the similarity between the full-scale structure and its scaled-down model. Comparison of the frequency response functions of the full-scale structure and its scaled-down model can verify the similarity between two scales. In this study, three composite I-beams with different scales (i.e. small, medium and large) are designed and manufactured. The designed composite I-beams replicate the spar caps and the shear web of a utility-scale wind turbine blade. Frequency response functions of the three beams are measured using a hammer impact test. The scaling laws of the frequency response function are derived using similitude analysis. The derived scaling laws are used to map the frequency response functions of the large and medium beams to that of the small beam. The comparison between the FRF curve of the small beam and the mapped FRF curves of the large and medium beams indicates the similarity of the laminated beams with different scales and demonstrates the effect of the scaling on shear and flexural stiffnesses.

Keywords Wind turbine blade • Similitude • Frequency response function • Composite I-beam • Scaling

21.1 Introduction

Certification of the utility-scale wind turbine blades starts with coupon testing of the materials and culminates with a final full-scale testing of the blade. Coupon testing is not considered a full representative of the structural durability of the blade and full-scale testing although necessary, is time-consuming and very expensive. Subcomponent testing can bridge the gap between the coupon and the full-scale testing by expediting and facilitating the testing of the new materials and designs for the wind turbine blades. The design of subcomponent tests of wind turbine blades has been the scope of several studies. Mandell et al. [1] fabricated composite I-beams representative of the cross section geometry of a wind turbine blade and loaded the beams in a four-point bending test. The resulting measurements were in good correlation with the finite element model and the simple beam theory. However, the similarity between the designed specimens and the parent I-beam geometry inside the utility–scale wind turbine blade was not studied. Cairns et al. [2] designed a subcomponent test for the root stud section of a blade. The root studs sections were manufactured and loaded in a pull-out test replicating those of a wind turbine blade. Although the primary focus of the study was manufacturing, a large amount of fatigue and static test data were generated. Mandell et al. [3] pursued a study to determine the optimal transition angle between the sandwich panels and the skin-stiffener of a wind turbine blade. Subcomponent specimens were designed to measure the fracture loads in the tension tests. Sayer et al. [4] proposed a subcomponent design to study the performance of the bonding paste used in the bonded joints of the blade under complex loading configuration. Later, the effect of the bond-line thickness on the flexural strength of the proposed Henkel beam design was investigated [5]. Zarouchas et al. [6] designed a subcomponent test for the spar caps and the shear web of the blade using two symmetric and asymmetric I-beams to investigate the dominant failure modes in a four-point bending test. However, the asymmetric design had unpredicted failure modes on the location of the loading supports. Jensen [7] performed a flap-wise bending test on a 35 meters box girder of a wind turbine blade to investigate the dominant failure modes. Although the size of the test was larger than what is considered a laboratory-scale test, the dominant failure modes were characterized using different measurement techniques. White et al. [8] used resonance excitation with forced hydraulic loading to develop a dual-axis fatigue test on a truncated wind turbine blade using the.

M.E. Asl (✉) • C. Niezrecki • J. Sherwood • P. Avitabile
Department of Mechanical Engineering, University of Massachusetts Lowell, Lowell, MA, USA
e-mail: Mohamad_eydaniasl@student.uml.edu

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_21

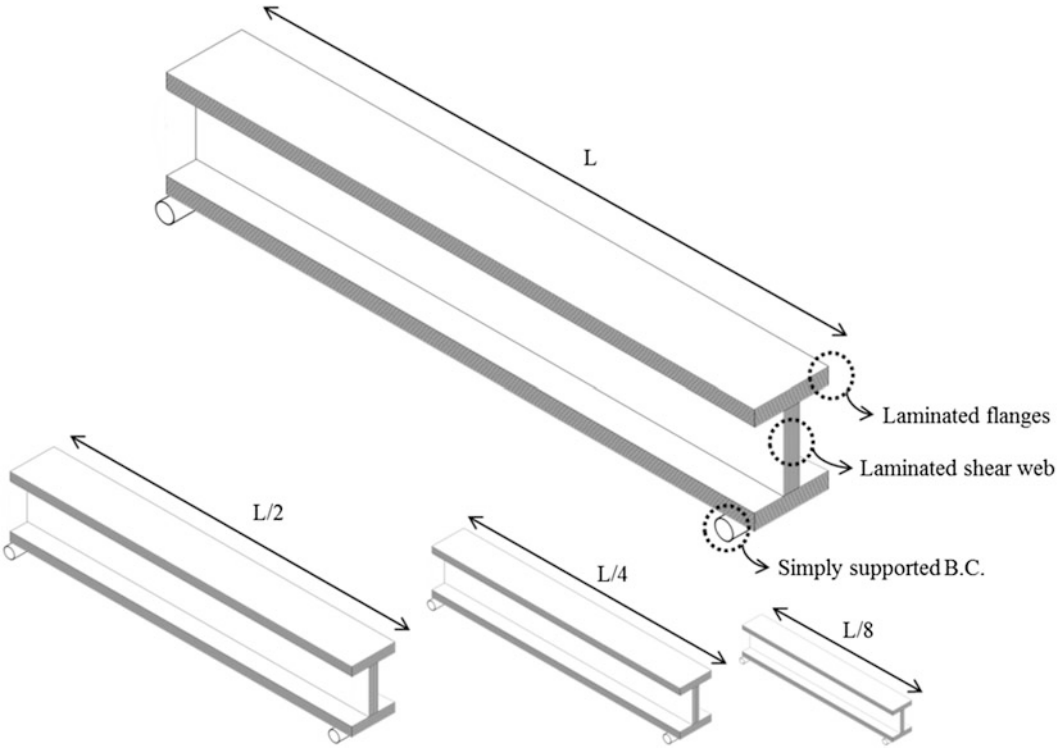


Fig. 21.1 Prototype (*top*) and Models with different scales and layups (*bottom*)

Although the previous efforts on subcomponent testing of wind turbine blades include a wide range of designs in size and detail, the proposed designs lack a systematic methodology to guarantee the correlation and similarity between the parent components and the designed subcomponents.

The designed subcomponent (referred to as the "model") regardless of the size and complexity needs to be correlated with the full-scale component (referred to as "prototype", see Fig. 21.1). The correlation between the scaled model and the prototype must be based on the structural parameters that describe the response of the system under study [9]. Similitude theory can extract scaling laws from the governing equations of the system to predict the response of the prototype from those of the scaled-down model. In order to take advantage of the scaling laws, a properly scaled model should be designed to work well with the derived scaling law. The designed scaled model needs to predict the response of the prototype accurately using the derived scaling laws. Otherwise, the experimental data of the scaled model cannot be correlated with the prototype and therefore the scaled-down model will not be representative of its corresponding prototype. As the model is not always an exact scaled-down replica of the prototype, a deliberate methodology should be implemented to design scaled models. The design of a scaled-down laminated composite structure is a challenge because the thickness of the plies in the laminates is fixed cannot be scaled to any desired ratio. Therefore, a deliberate methodology needs to be applied to design a scaled composite structure which is representative of its full-scale parent. To study the degree of similarity between the scaled-down model and the full-scale prototype, the structural response of the scaled model can be mapped to those of the prototype using the scaling laws. Comparison of the response of the prototype with the mapped response of the model indicates the degree of the similarity between the scaled model and the full-scale prototype. A scaled model which is structurally similar to its full-scale prototype can be used for mechanical testing, verification and validation and system identification of the large structures [10, 11].

Similitude analysis can provide the necessary and sufficient conditions for the similarity of the two structures with different scales. The similarity conditions can be obtained using either dimensional analysis or direct use of the governing equations. To study the buckling of the laminated plates using the scaled models, Simitses and Rezaeepazhand [12] derived the similarity conditions for the buckling of the orthotropic and symmetric cross-ply laminated plates using the governing equations to demonstrate the credibility of similitude analysis. They applied the same approach to extract the scaling laws for the vibration of scaled laminated rectangular plates [13] and to investigate the effect of axial and shear loads on the stability of the scaled laminated rectangular plates [14, 15]. Later, the same method was extensively used in their works regarding the prediction of vibration response of laminated shells [16, 17]. According to their results, ply-level scaling of a laminated plate

yields scaled models with perfect accuracy in predicting the response of the prototype. However, a systematic approach to design the partially similar scaled models was lacking. Ilbeigi et al. [18] developed a method to design reduced scale models for an euler-bernoulli beam and later they extended their approach to study reduced scale models for systems with disparate spatial and temporal scales [19]. In two previous works by the authors [20, 21], a clear methodology was proposed to design the partially similar models for a simply supported laminated plate. The derived scaling laws were used as a metric to measure the correlation between the scaled models and the prototype. The accuracy of the designed scaled models in predicting the response of the prototype was shown to be proportional to the discrepancy of the models in satisfying the requirements of the similarity. The same approach was improved in the later studies [22] to design scaled composite I-beams representative of the spar caps and the shear web of a utility-scale wind turbine blade using partial similarity. Experimental validation of the similarity of the strain fields in the designed I-beams using the distorted layup scaling technique was demonstrated in a three-point quasi-static bending test [23].

The best candidates for the design of the subcomponent test for wind turbine blade are the portions of the blade that are structurally critical. The spar caps and the shear web of the wind turbine blade carry the majority of the aerodynamic loads and make an interesting case study for subcomponent testing of the blade. In this study, the I-beam geometry of a utility-scale wind turbine blade [24] near its max chord is considered as the prototype and scaled-down beams are designed from the prototype in three different scales (i.e. small, medium and large) [25, 26]. This area was chosen for analysis because many failures in blades occur near maximum the chord, and the area of interest is responsible for carrying most of the aerodynamic loads. The designed I-beams are manufactured and the frequency response functions of the beams are measured using a hammer impact test. The frequency response functions of the medium and large beams are mapped to that of the small beam using the scaling laws to investigate the degree of the similarity among the scaled beams and to study the effect of scaling on the shear and flexural stiffnesses. According to the experimental results, frequency response functions of all the beams show a very good correlation for the first five modes which indicates the similarity of the dynamic behavior of the three scales. However, the degree of correlation decreases for the higher modes which suggest that the scaling has a higher impact on the shear stiffness of the beams compared to the flexural stiffness.

21.2 Governing Equations

In this section, the governing equations for vibration of a laminated composite I-beam are presented. The closed form solution for the natural frequencies of the beam in free-free boundary conditions is derived. The similarity transformation is introduced and the scaling laws for natural vibration frequencies of the composite beam are derived. The flexural vibration of a shear deformable thin-walled composite I-beam shown in Fig. 21.2 in the absence of thermal effects for a symmetric I-beam is considered.

Neglecting all coupling effects due to the symmetric geometry, the governing equations for free flexural vibration in y-direction are given by [27] which are well known Timoshenko beam equations:

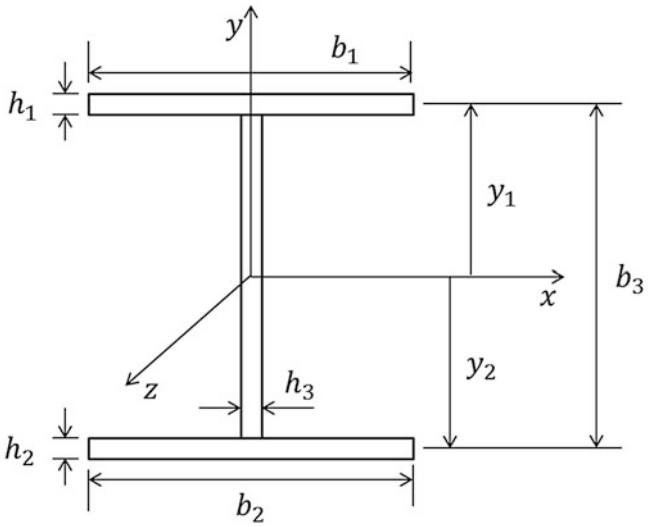


Fig. 21.2 Geometry and coordinate of the I-beam [22]

$$(GA_x)_{com}\left(V'' + \Psi_x'\right) = \rho A\ddot{V} \tag{21.1a}$$

$$(EI_x)_{com}\Psi''_x - (GA_x)_{com}\left(V' + \Psi_x\right) = \rho I_x\ddot{\Psi}_x \tag{21.1b}$$

where Ψ_x denotes the rotation of the cross section with respect to x axis shown in Fig. 21.2, V the displacement in y direction, q the distributed load, the prime ($'$) is used to indicate differentiation with respect to z and ($\cdot$) is time differentiation. Density and area of cross section are expressed by ρ and A respectively and I_x is moment of inertia with respect to x-axes. $(GA_x)_{com}$ and $(EI_x)_{com}$ are shear and flexural rigidity of thin walled composite with respect to x, respectively which could be expressed as:

$$(EI_x)_{com} = \left[A_{11}^{\alpha}\, y_{\alpha}^{(2)} - 2B_{11}^{\alpha}\, y_{\alpha} + D_{11}^{\alpha}\right] b_{\alpha} + \frac{b_3^{(3)}}{12}A_{11}^{3} \tag{21.1c}$$

$$(GA_x)_{com} = A_{55}^{\alpha} b_{\alpha} + A_{66}^{3} b_3 \tag{21.1d}$$

where A_{11} , A_{66} , A_{55} , B_{11} and D_{11} are elements of extensional, coupling and bending stiffness matrices for a composite layup [27]. The superscript in the parenthesis ( ) denotes the power of the exponent, and the repeated index denotes summation. Index α varies from 1 to 3 where the indices 1 and 2 represent the top and bottom flanges, and 3 is for the web, respectively as shown in Fig. 21.2 and b_{α} denotes width of the flanges and web. The closed-form solution for flexural vibration frequencies in the y-direction may be directly calculated for the free-free boundary condition as [28]:

$$\omega_{y_n} = \sqrt{\left[\frac{\rho A}{(EI_x)_{com}}\frac{L^4}{(n+0.5)^4\pi^4} + \frac{\rho A}{(GA_x)_{com}}\frac{L^2}{(n+1)^2\pi^2}\right]^{-1}} \tag{21.2}$$

Natural frequencies for vibration of a shear deformable composite I-beam in free-free boundary conditions are described by Eq. (21.2). To derive the scaling laws, it is assumed that all the variables of the governing equations for the prototype (x_p) can be connected to their corresponding variables in a scaled model (x_m) by a one to one mapping. Then, the scale factor for each variable can be defined as $\lambda_x = x_p/x_m$ which is ratio of each variable of the prototype to that of the scaled model. Rewriting Eq. (21.2) for the model and prototype and applying similarity transformation, the scaling laws can be extracted as the following based on the standard similitude procedure [12]:

$$\lambda_l{}^2 = \frac{\lambda_{EI}}{\lambda_{GA}} \tag{21.3}$$

$$\lambda_{\omega} = \sqrt{\frac{\lambda_n{}^4\lambda_{EI}}{\lambda_{\rho}\lambda_A\lambda_l{}^4}} \tag{21.4a}$$

$$\lambda_{\omega} = \sqrt{\frac{\lambda_n{}^2\lambda_{GA}}{\lambda_{\rho}\lambda_A\lambda_l{}^2}} \tag{21.4b}$$

Eq. (21.3) which is referred to as the design scaling law is a prerequisite for deriving constitutive response scaling laws Eqs. (21.4a) and (21.4b). Design scaling law Eq. (21.3) denotes that for having complete similarity between two shear deformable beams, ratio of the flexural to shear stiffness must be equal to the square of the length for the two scales. This condition is shown to be satisfied if the prototype and model have a unidirectional layup in their flanges and also the dimensions of the model are proportional to those of the prototype [25]. Having Eq. (21.3) satisfied, the ratio of natural frequencies between two scales can be obtained using constitutive scaling laws Eqs. (21.4a) and (21.4b).

21.3 Experimental Results

Based on the similitude analysis in the previous section, three I-beams having three different scales (i.e. small, medium and large) were manufactured from the prototype. The prototype geometry and lay-up scheme considered in this study emulate the spar-cap flanges of the Sandia 100 m wind turbine blade [24] near its max chord. The prototype is assumed

Table 21.1 Geometry and layup of the small, medium and large beams.

Beam scale	Dimensions ($l \times b_3 \times b_1$) m^3	Layup of the flanges	Layup of the shear web
Small	$0.58 \times 0.05 \times 0.03$	$[0]_2$	$[\pm 45]$
Medium	$1.17 \times 0.10 \times 0.06$	$[0]_4$	$[\pm 45]_2$
Large	$3.51 \times 0.32 \times 0.18$	$[0]_{12}$	$[\pm 45]_6$

Fig. 21.3 Hammer impact test on the small (*left*), medium (*middle*) and large (*right*) beams

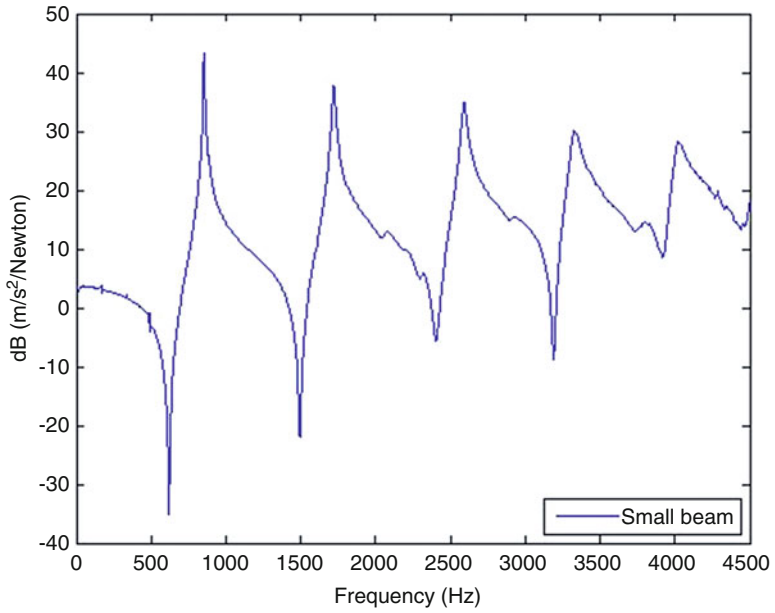


Fig. 21.4 Frequency response function of the small beam for the drive point

to be an I-beam consisting of two identical laminated plates and a sandwich shear web having an overall dimension of 29.3 m × 2.446 m × 1.5 m with layup $[0\ ]_{100}$ and ply thickness of t = 1.36 mm. The shear web is a sandwich panel with the foam thickness of 160 mm and two face sheets with layup $[\pm 45]_{50}$ with an overall thickness of 17.5 mm. Because the blade is tapered throughout its length and the geometry of the structure changes significantly, the geometry of the full-scale component (i.e. the first 1/3 of the Sandia blade from its root) is assumed to have a uniform cross-section whose height along its length was averaged to simplify the analysis. All computations are implemented for the glass/epoxy materials with following material characteristics [24]: $E_1 = 41.8$ GPa, $E_2 = E_3 = 14$ GPa, $G_{12} = G_{13} = 2.63$ GPa, $G_{23} = 1.83$ GPa, $\upsilon_{12} = \upsilon_{13} = 0.28$, $\upsilon_{23} = 0.47$. The sizes of the designed beams were chosen based on the manufacturing capabilities and limitation of the test equipment. The scale of the small, medium and large beams were selected to be 1/50, 1/25 and 1/8.33 of the prototype respectively. The flanges of the small beam consisted of two unidirectional plies and the shear web foam was wrapped in a $[\pm 45]$ glass fabric. Medium and large beams had 4 and 12 unidirectional plies in their flanges and their shear webs have the layup $[\pm 45]_2$ and $[\pm 45]_6$ respectively as summarized in Table 21.1.

The beams were manufactured and their frequency response functions were measured using a hammer impact test in free-free boundary conditions as shown in Fig. 21.3. The FRF curves of the drive points of the all three beams were measured as shown in Figs. 21.4, 21.5, and 21.6. The comparison of the FRF curves for all three beams within their frequency range are shown in Fig. 21.7. The first five flexural bending modes were measured for all three beams. Only modes in the y-direction are considered in this work; however, the actual measurements shows these five major modes in the y-direction but there are

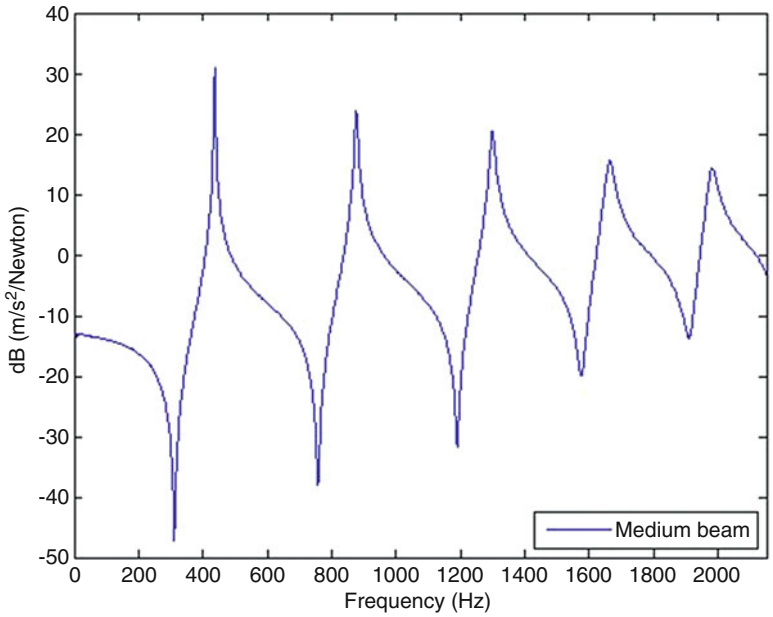


Fig. 21.5 Frequency response function of the medium beam for the drive point

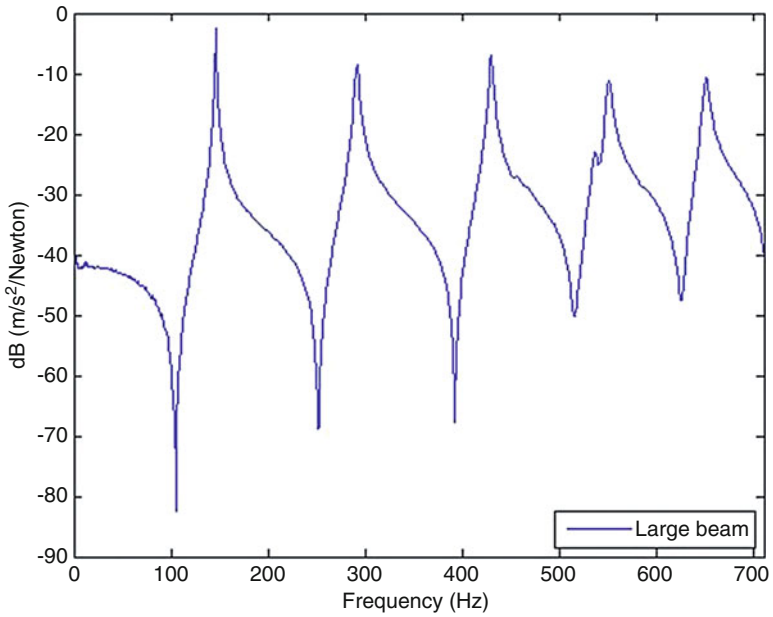


Fig. 21.6 Frequency response function of the large beam for the drive point

also some off axis motion in the x-direction (perpendicular to the y-direction that are not of interest in this study). Because the designed I-beams satisfy the conditions for the complete similarity described by Eq. (21.3), both of the response scaling laws Eq. (21.4a) and Eq. (21.4b) can be used to scale the FRF curves. The FRF of the medium and large beams are scaled to that of the small beam using the derived response scaling law Eq. (21.4a) for comparison of the FRF curves of the designed beams. The stiffness, density and geometry ratios (i.e. λ_{EI}, λ_{ρ}, λ_{A} and λ_{l}) were calculated based on Eqs. (21.1c) and (21.1d) and the dimensions of the beams assuming that the modes of the large and medium beams are mapped to their corresponding modes of the small beams (i.e. $\lambda_{n} = 1$). The frequency scale factor λ_{ω} for the medium and large beams were calculated based on Eq. (21.4b) and were applied to FRF curves of the medium and large beams.

The FRF of the small beam and the scaled FRFs of the medium and large beams are shown in Fig. 21.8. According to Fig. 21.8, there is a very good correlation among the FRFs of the three beams for the first five flexural bending modes which

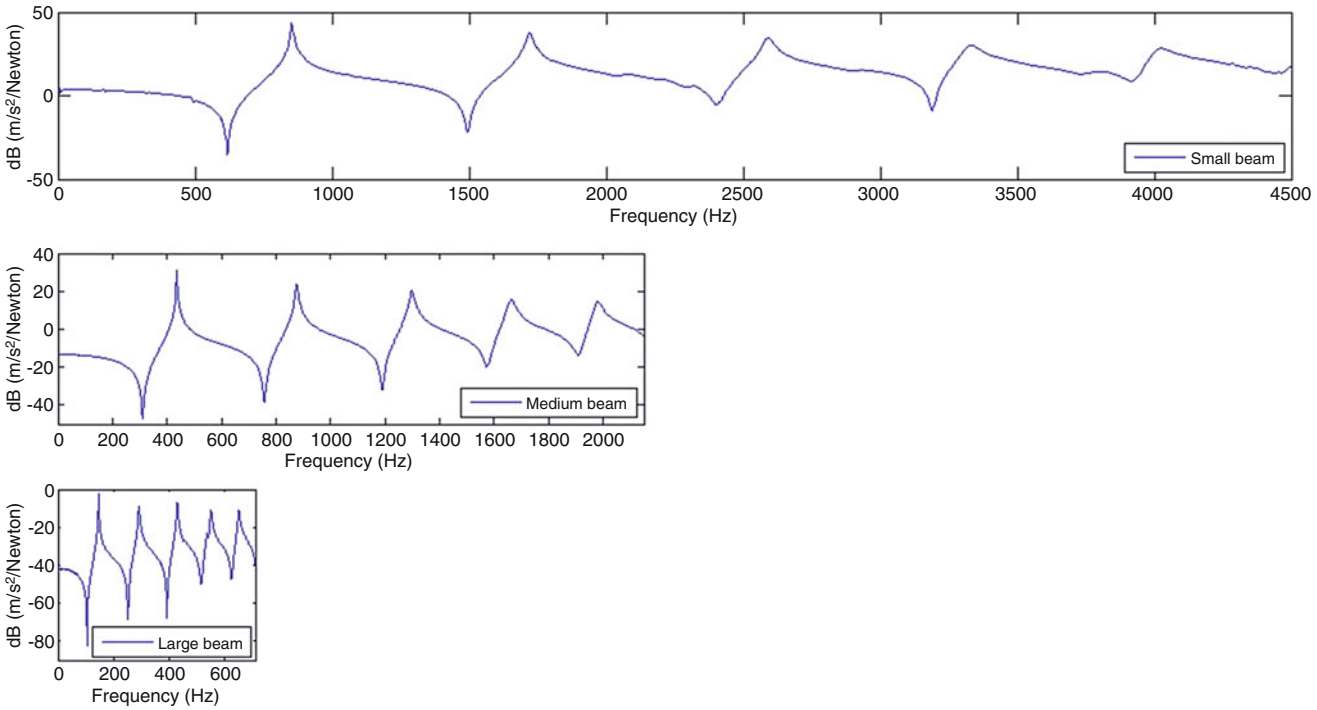


Fig. 21.7 Comparison of the unscaled frequency response functions of the small (*top*), medium (*middle*) and large (*bottom*) beams

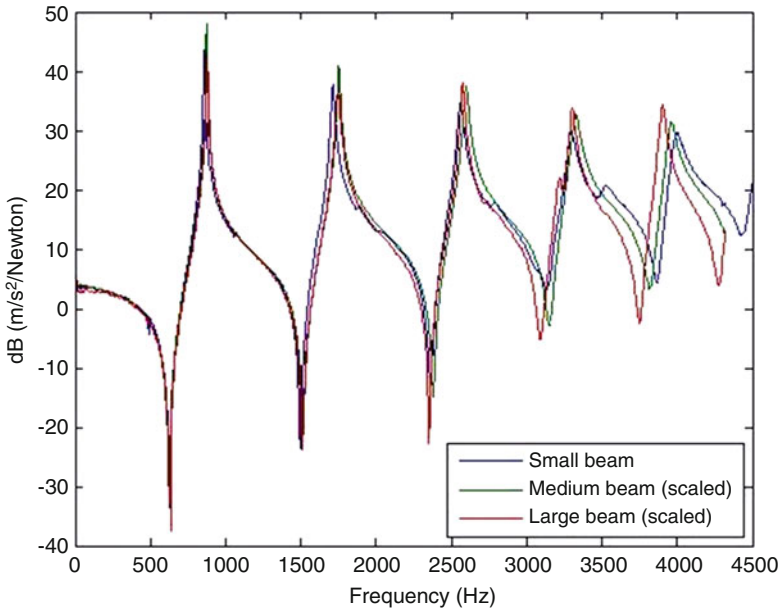


Fig. 21.8 Comparison of the FRF of the small beam with the scaled FRFs of the medium and large beams

suggests the similarity of the dynamic response for the different scales of the I-beams. However, the degree of correlation decreases for the higher modes. According to Eq. (21.2) and the length to height ratios of the beams, both flexural and shear stiffnesses are important in describing the natural frequencies of the beams. The effect of the flexural stiffness is dominant for the first two modes and the shear stiffness becomes dominant for the higher modes. The excellent correlation of the FRFs for the first two modes suggests that scaling has a negligible impact on the flexural stiffness of the beams. However, the decreased level of correlation for the higher modes indicates that scaling has a more significant impact on the shear stiffness of the beams. For fifth flexural bending mode where the shear stiffness is dominant, the resonance peaks shifts toward the higher frequencies for the smaller scales. This observation indicates that the shear stiffness of the designed beams decrease

for the larger scales. The rotary inertia effect is also a factor and becomes important for the higher modes. Although the rotary inertia term was neglected in the governing equations to derive the closed form scaling laws, the frequency values of the higher modes are decreased due to the rotary inertia effects. This effect has a higher impact on the frequency values of the beams with the larger scale which is in agreement with the shifting of the fifth mode as shown in Fig. 21.8.

21.4 Conclusions

Scaled-down models of the large laminated structures can expedite and facilitate the certification procedure, especially for utility-scale wind turbine blades. In this study, three laminated composite I-beams were designed in three different scales replicating the spar caps and the shear web geometry of a utility-scale wind turbine blade. The designed beams were manufactured and tests using modal impact hammer measurements were performed to investigate the similarity of the designed beams having different scales and the scaling effects on the dynamic behavior of the beams. The FRFs of the small, medium and large beams were measured for the first five flexural bending modes. Then, the FRF curves of the medium and the large beams were scaled using the derived scaling laws and were mapped to that of the small beam. A very good correlation was observed among the FRF curves of the small, medium and large beams. There was a small decrease in the correlation of the beams at higher frequencies that is likely due to either stiffness or rotary inertia effects that are not entirely captured by the scaling laws and will be studied in future work. Overall, there was a very good correlation seen in the FRFs of the small, medium and large beams and shows promise for using this approach to predict frequency response measurements from smaller scaled systems.

Acknowledgments This material is based upon work supported by the National Science Foundation under Grant Number 1230884 (Achieving a Sustainable Energy Pathway for Wind Turbine Blade Manufacturing). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Special thanks go to Stephen Nolet from TPI Composites Incorporation to help define the composition of the specimens and the tests conducted. We also thank Andy Schoenberg from the Maine Composites Alliance – Composites Engineering Research Laboratory for fabricating the I-beam specimens.

References

1. Mandell, J.F., Combs, D.E., Samborsky, D.D.: Fatigue of fiberglass beam substructures. Wind Energy. **16**, 99 (1995)
2. Cairns, D.S., Skramstad, J.D., Mandell, J.F.. Evaluation of hand lay-up and resin transfer molding in composite wind turbine blade structures. In: International SAMPE Symposium and Exhibition: SAMPE, pp. 967–980 (1999).
3. Mandell, J.F., Creed Jr., R.J., Pan, Q., Combs, D.W., Shrinivas, M.: Fatigue of fiberglass generic materials and substructures. Wind Energy. **15**, 207 (1994)
4. Sayer, F., Post, N., Van Wingerde, A., Busmann, H.G., Kleiner, F., Fleischmann, W., et al.. Testing of adhesive joints in the wind industry. In: European Wind Energy Conference and Exhibition 2009. pp 288–315 (2009).
5. Sayer, F., Antoniou, A., van Wingerde, A.: Investigation of structural bond lines in wind turbine blades by sub-component tests. Int. J. Adhes. Adhes. **37**, 129–135 (2012)
6. Zarouchas, D.S., Makris, A.A., Sayer, F., Van Hemelrijck, D., Van Wingerde, A.M.: Investigations on the mechanical behavior of a wind rotor blade subcomponent. Compos. Part B. **43**, 647–654 (2012)
7. Jensen, F.M., Falzon, B.G., Ankersen, J., Stang, H.: Structural testing and numerical simulation of a 34 m composite wind turbine blade. In: Fifteenth International Conference on Composite Materials ICCM-15, vol. 76, pp. 52–61 (2006)
8. White, D., Musial, W., Engberg, S.. Evaluation of the B-REX fatigue testing system for multi-megawatt wind turbine blades. In: Collection of the 2005 ASME Wind Energy Symposium Technical Papers at the 43rd AIAA Aerospace Sciences Meeting and Exhibit. pp 52–65.
9. Ilbeigi, S., Chelidze, D.: Reduced order models for systems with disparate spatial and temporal scales. In: Rotating Machinery, Hybrid Test Methods, Vibro-Acoustics & Laser Vibrometry, vol. 8, pp. 447–455. Springer International Publishing, Orlando (2016)
10. Behmanesh, I., Yousefianmoghadam, S., Nozari, A., Moaveni, B., Stavridis, A.: Effects of prediction error bias on model calibration and response prediction of a 10-story building. In: Model Validation and Uncertainty Quantification, vol. 3, pp. 279–291. Springer International Publishing (2016)
11. Yousefianmoghadam, S., Song, M., Stavridis, A., Moaveni, B.: System identification of a two-story infilled rc building in different damage states. Improving the Seismic Performance of Existing Buildings and Other Structures 2015, pp. 607–618. American Society of Civil Engineers (ASCE), San Francisco (2015)
12. Simitses, G.J., Rezaeepazhand, J.: Structural similitude for laminated structures. Compos. Eng. **3**, 751–765 (1993)
13. Rezaeepazhand, J., Simitses, G.J.: Use of scaled-down models for predicting vibration response of laminated plates. Compos. Struct. **30**, 419–426 (1995)
14. Rezaeepazhand, J., Simitses, G.J., Starnes Jr., J.H.: Design of scaled down models for stability of laminated plates. AIAA J. **33**, 515–519 (1995)

15. Simitses, G.J., Rezaeepazhand, J.: Structural similitude and scaling laws for buckling of cross-ply laminated plates. J. Thermoplast. Compos. Mater. **8**, 240–251 (1995)
16. Rezaeepazhand, J., Simitses, G.J., Starnes Jr., J.H.: Design of scaled down models for predicting shell vibration response. J. Sound Vib. **195**, 301–311 (1996)
17. Rezaeepazhand, J., Simitses, G.J.: Structural similitude for vibration response of laminated cylindrical shells with double curvature. Compos. Part B. **28**, 195–200 (1997)
18. Ilbeigi, S., Chelidze, D.: Model order reduction of nonlinear Euler-Bernoulli beam. In: Nonlinear Dynamics, vol. 1, pp. 377–385. Springer International Publishing (2016)
19. Ilbeigi, S., Chelidze, D.: Reduced order models for systems with disparate spatial and temporal scales. Rotating Machinery, Hybrid Test Methods, Vibro-Acoustics & Laser Vibrometry, vol. 8, pp. 447–455. Springer International Publishing (2016)
20. Asl, M.E., Niezrecki, C., Sherwood, J., Avitabile, P.: Predicting the vibration response in subcomponent testing of wind turbine blades. In: Special Topics in Structural Dynamics, vol. 6, pp. 115–123. Springer (2015)
21. Asl, M.E., Niezrecki, C., Sherwood, J., Avitabile, P.: Application of structural similitude theory in subcomponent testing of wind turbine blades. Proceedings of the American Society for Composites, pp. 8–10. San Diego (2014)
22. Asl, M.E., Niezrecki, C., Sherwood, J., Avitabile, P.: Similitude analysis of composite I-beams with application to subcomponent testing of wind turbine blades. In: Experimental and Applied Mechanics, vol. 4, pp. 115–126. Springer (2016)
23. Asl, M., Niezrecki, C., Sherwood, J., Avitabile, P.. Similitude analysis of the strain field for loaded composite I-beams emulating wind turbine blades. In: Proceedings of the American Society for Composites: Thirty-First Technical Conference (2016).
24. Griffith, D.T., Ashwill, T.D.. The Sandia 100-meter all-glass baseline wind turbine blade: SNL100–00. Sandia National Laboratories, Albuquerque, Report NoSAND2011–3779 (2011)
25. Asl, M.E., Niezrecki, C., Sherwood, J., Avitabile, P.: Design of scaled-down composite I-beams for dynamic characterization in subcomponent testing of a wind turbine blade. In: Shock & Vibration, Aircraft/Aerospace, Energy Harvesting, Acoustics & Optics, vol. 9, pp. 197–209. Springer International Publishing (2016)
26. Eydani, A.M., Niezrecki, C., Sherwood, J., Avitabile, P.: Vibration prediction of thin-walled composite I-beams using scaled models. Thin-Walled Struct. **113**, 151–161 (2017)
27. Reddy, J.N.: Mechanics of Laminated Composite Plates And Shells: Theory and Analysis. Boca Raton, FL. CRC, London (2003)
28. Vo, T.P., Lee, J.: Flexural–torsional coupled vibration and buckling of thin-walled open section composite beams using shear-deformable beam theory. Pergamon, p. 631

Chapter 22
MPUQ-b: Bootstrapping Based Modal Parameter Uncertainty Quantification—Fundamental Principles

S. Chauhan and S.I. Ahmed

Abstract It is well known that modal parameters play a key role towards understanding the dynamics of a structure. Their estimation, by means of experiments, forms the crux of modal analysis. Modal parameters not only help in characterizing the dynamics of the structure but are also used for several other purposes including, finite element model updating, design optimization, sensitivity analysis, etc. It is therefore important to estimate them accurately and several modal parameter estimation techniques have been developed over the years for this purpose. Despite advance methods, estimation of modal parameters is always accompanied with certain uncertainty, which can be attributed to several factors including, noisy measurements, complexities inherent in the structure, modeling errors etc. Remarkably, the usual practice is to provide the estimated modal parameters as they are, without providing any means to validate the accuracy of these estimates. In other words, the estimation procedure often does not include, or overlooks, the measures for quantifying the uncertainty associated with estimated modal parameters.

In this work, a methodology to quantify uncertainty associated with the modal parameters estimated using Experimental Modal Analysis techniques is developed. This methodology is termed as Bootstrapping based Modal Parameter Uncertainty Quantification (MPUQ-b). The proposed methodology utilizes the technique of Bootstrapping, which is a computer intensive method for statistical inference. In this first paper, the fundamentals of this methodology are laid out. The paper focuses on illustrating the characteristics of Bootstrapping and its effectiveness for the intended use of modal parameter uncertainty quantification. By means of studies conducted on a simple single degree of freedom system, it is shown how Bootstrapping can be employed for quantifying the uncertainty associated with estimated modal parameters by providing such measures of accuracy as variance, confidence intervals, bias etc. Incorporation of bootstrapping in modal parameter estimation procedure forms the essence of MPUQ-b and is detailed in a subsequent paper.

Keywords Bootstrapping • Uncertainty quantification • Modal parameter estimation • Statistical inference • Standard error

Abbreviations

CI	Confidence intervals
EMA	Experimental modal analysis
FRF	Frequency response function
IRF	Impulse response function
MPUQ-b	Bootstrapping based Modal Parameter Uncertainty Quantification
OMA	Operational modal analysis
SDOF	Single degree-of-freedom
SNR	Signal-to-noise ratio
SSI	Stochastic subspace identification

S. Chauhan (✉)
Bruel & Kjær Sound and Vibration Measurement A/S, Skodsborgvej 307, DK 2850, Naerum, Denmark
e-mail: schauhan@bksv.com

S.I. Ahmed
Mechanical Engineering Department, Danmarks Tekniske Universitet, Copenhagen, Denmark
e-mail: shahrukh3307@gmail.com

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_22

Nomenclature

x	Original data sample
x^i_{BS}	i[th] bootstrap sample drawn from **x** with replacement
m , c *and* k	Physical mass, damping and stiffness parameters of the SDOF system
ω_{theo} , ζ_{theo}	Theoretical modal frequency and damping
NB	Block size
N_{avg}	Number of averages
H_{avg}	Averaged FRF estimated using H_1 technique on original data
$\overline{MP}_{est}$	Modal parameters estimated on the basis of H_{avg}
ω_{est} , ζ_{est}	Modal frequency and damping estimated on the basis of averaged H_{avg}
$\mathbf{G}^i_{XF}$, $\mathbf{G}^i_{FF}$	Input-output cross and input-input auto power spectra for individual blocks used for estimating H_{avg}
$\mathbf{H}^i_{BS}$	i[th] averaged bootstrap FRF
MP^i_{BS}	Modal parameters estimated on the basis of $\mathbf{H}^i_{BS}$
N_{boot}	Number of bootstrap repetitions. Same for number of times Monte Carlo simulations are run.
$\overline{\omega}_{BS}$, $\overline{\zeta}_{BS}$	Mean of N_{boot} bootstrapped estimates of modal frequency and damping
$\overline{\omega}_{MC}$, $\overline{\zeta}_{MC}$	Mean of N_{boot} Monte Carlo estimates of modal frequency and damping

22.1 Introduction

Modal parameter estimation procedure is typically a two stage procedure. In the first stage, acquired raw data is processed into characteristic functions (such as Frequency response functions (FRF) or Impulse response functions (IRF)), using signal processing techniques. The subsequent stage involves application of modal parameter estimation algorithms to the characteristic function estimated in the first stage. The above mentioned two-stage procedure can be also viewed within a statistical framework, as steps involved in either stages (for e.g. use of averaging techniques to reduce random errors while calculating characteristic functions, utilization of least squares methodology for estimation of modal parameter) are based on solid statistical foundation.

In spite of this, general practice is to provide estimated modal parameters as they are, without much statistical insight. The experimental nature of modal analysis however, necessitates the need to provide statistical evaluation of the modal parameter estimation procedure and underline its importance. It is well understood that entire modal parameter estimation procedure is subject to issues related to accuracy and precision. This is due to several factors that cannot be entirely accounted for in the estimation procedure. Starting with the raw data, which might have errors associated with data acquisition, to complexities of the structure being analyzed that cannot be modeled sufficiently, from inherent limitations of the models used to characterize the structure, to issues associated with signal processing techniques involved [1], there are several reasons that affect the accuracy of estimated modal parameters. Thus, it is not incorrect to say that estimated modal parameters always have some degree of uncertainty associated with them. It is therefore vital to include some measure for quantifying this uncertainty to provide a true reflection of how accurate and reliable these estimates are.

While there has been some work with regards to estimation of variance associated with estimated characteristic functions like FRFs ([2] is a classic work listing formulas for calculating variance in FRF estimates in terms of ordinary coherence functions), quantifying uncertainty in estimated modal parameters is a relatively newer subject of research. In [3], covariance matrices of estimates obtained from Maximum Likelihood and prediction error based parameter estimation methods are used to obtain the uncertainty bounds. Similar approach is extended to Operational Modal Analysis (OMA) in [1]. The suggested method is developed around Stochastic Subspace Identification algorithm (SSI) [4]. Dohler et al. [5] builds on this and provides for a fast, efficient multi-order method for quantifying uncertainty associated with SSI estimates. This approach is also extended to multi-setup scenario in [6].

A common aspect of these papers is that they are based on the principle of error propagation [7]. The method involves estimation of standard deviation in characteristic functions (such as FRF in case of EMA or output response covariance in case of OMA), as a measure of error associated with the characteristic function, and then propagates it to the modal parameters estimated from these characteristic functions. This error propagation is done by exploiting the mathematical relationship between the characteristic function and modal parameters. In simple words, estimated standard deviation in characteristic functions is used to estimate standard deviation in modal parameters by exploiting the functional mathematical relationship between the two. Equation (22.1) forms the basis of this method [7] and can be described as follows. If θ and β are two variables such

that θ is a function of β and variance of beta ($\Sigma_{\widehat{\beta}}$) is known, then variance of θ ($\Sigma_{\widehat{\theta}}$) can be obtained using Eq. (22.1), by mathematically computing the gradient of θ with respect to β $\left(\left(\frac{\partial\theta}{\partial\beta}\right)_{\widehat{\beta}}\right)$ and evaluating it at point estimate of β, i.e. $\widehat{\beta}$.

$$\Sigma_{\widehat{\theta}} = \left(\frac{\partial\theta}{\partial\beta}\right)_{\widehat{\beta}}^{\mathrm{T}} \Sigma_{\widehat{\beta}} \left(\frac{\partial\theta}{\partial\beta}\right)_{\widehat{\beta}} \tag{22.1}$$

This approach is however, not simple and straight forward to apply for several reasons. The process of obtaining modal parameters from characteristic functions is algorithm dependent and involves several intermediary steps and hence the error propagation is mathematically overwhelming and cumbersome. Modal parameters are often indirect result of parameter estimation procedure. Most algorithms involve estimation of polynomial coefficients, or state transition matrix, as intermediate steps and they are then utilized to estimate modal parameters. Further, modal parameters are nonlinear functions of polynomial coefficient matrices (or state transition matrices), which means that Eq. (22.1) is approximate and only holds up to higher terms [3].

In comparison to the error propagation based approach, which has been central to the research efforts in this field, this paper highlights a resampling based statistical method called *Bootstrapping* for the purpose of assigning such measures of accuracy as variance, bias, confidence intervals etc. to estimated modal parameters. This method thus, provides an alternate technique to quantify uncertainty associated with them. Despite some initial efforts along these lines during the 1990s (when bootstrapping was gaining foothold within statistical community), application of bootstrapping within the field of modal analysis has been quite limited. This was perhaps due to the associated computational demands, which are easily manageable in current scenario. [8] investigated bootstrapping in the context of providing statistical insights in scenarios where input excitation/responses are not distributed normally. In [9], Farrar et al. carried out similar investigations in this direction and extended it to modal parameter estimates, comparing the results of bootstrapping with those from Monte Carlo simulations.

The overall goal of this work, which is presented in two parts, is to showcase a Bootstrapping based methodology for uncertainty quantification of modal parameters estimated from experimentally acquired data. This methodology is termed as Bootstrapping based Modal Parameter Uncertainty Quantification or MPUQ-b. This first paper highlights the fundamentals of the suggested methodology, starting with an introduction to the concept of bootstrapping, underlining its capabilities in terms of providing statistical inference. The paper then focuses on illustrating various characteristics of bootstrapping that enable it to be used effectively for the purpose of providing statistical inference with respect to estimated modal parameters. Finally, an exhaustive single degree-of-freedom (SDOF) system based numerical study demonstrates the usefulness of bootstrapping in context of modal parameter estimation procedure. The study tests bootstrapping under variety of conditions to verify its suitability and robustness. The study also brings to fore several qualitative and quantitative tools that bootstrapping enables.

22.2 Bootstrapping

22.2.1 Basic Principles and Procedure

Bootstrapping is a computer based resampling method of statistical inference using random sampling with replacement. Starting with Bradley Efron's seminal work [10], Bootstrapping principles were developed throughout the 1980s and are well documented in the form of a textbook [11] written by Bradley Efron. To understand how bootstrapping works, it is better to start with a small detour on statistical inference. Unlike descriptive statistics that concerns only with collected (or observed) data, inferential statistics aims at inferring properties of the population from which the data is collected. The collected data is, in general, smaller in comparison to the population (in other words is a sample obtained from the population). Thus one is not just interested in an estimate summarizing the collected data, but also in measures (such as variance, bias, confidence intervals, etc.) quantifying the accuracy of the estimate, its distribution (referred as sampling distribution), etc. Thus, the main goal of statistical inference is to say what one has learned about the population based on the collected sample(s).

An oft quoted example with regards to explaining statistical inference in context of bootstrapping is that of the sample mean. Suppose the task is to calculate the mean (average) age of the entire population of a certain region. Since the age related data for entire population cannot be collected, the obvious step is to collect a sample from the population and calculate the statistic of interest, which happens to be mean age in this case. For inferential purposes, this process of taking random sample from the population is repeated number of times and mean age is calculated for each of these samples. After sufficient repetitions, one can now find the distribution of mean age and represent it graphically by means of a histogram. This distribution is called *Sampling Distribution* [12]. Other measures of accuracy, such as standard errors, variance, etc. can also be calculated.

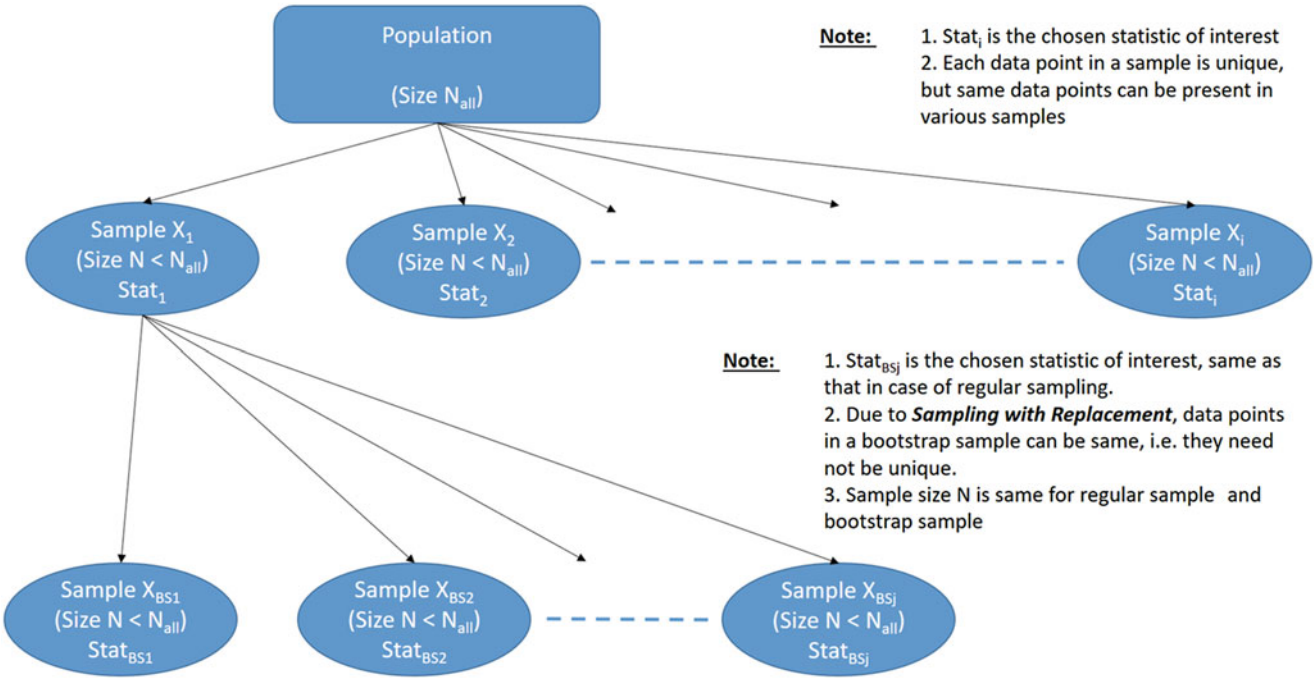


Fig. 22.1 Bootstrapping

This is the usual process of statistical inference, but there are several limitations that this approach poses, making it difficult to implement in practice.

1. The process of taking multiple samples is difficult, costly and time consuming. This is akin to performing multiple experiments and it is well-known that carrying out experiments is often an expensive and time-intensive proposition.
2. There are several situations where one does not have the luxury of repeating experiments; take for example a ground vibration test, where one has limited time during which the access to the test object (in case of Ground Vibration Test, an aircraft) is available. So the experiment can only be performed once. In other words, one can only collect one sample.
3. There are situations where the population may no longer be existing once the sample has been drawn.

It is in circumstances such as these that bootstrapping comes handy as it paves the way to statistical inference even when only a single sample is available (in other words, experiment is performed only once). The main idea behind bootstrapping is that inference about the population can be made on the basis of inference on the sample. This is referred as the *Plug-In principle* [11]*: To estimate a parameter describing the population, use the corresponding statistic for the sample*. This can be explained by continuing the mean age example (Fig. 22.1).

Once a sample (say x) has been drawn from the population, the sample is now treated as population and several samples, called *bootstrap samples*, are drawn from it by employing '*sampling with replacement*' [11]. Sampling with replacement means that a data point selected to form a part of the bootstrap sample (say x_{BS}^1) is placed back and can be therefore selected multiple times in a bootstrap sample (see Fig. 22.2). This is in contrast to the '*sampling without replacement*' scheme adopted while drawing the original sample x from the population, where a data point was only selected once. Thus using sampling with replacement, one can create several bootstrap samples (x_{BS}^i) and the statistic of interest (in this case, average age) can be calculated for each of these bootstrap samples. This allows for statistical inference on similar lines as mentioned before. To differentiate this approach, based on bootstrap samples, from the normal approach based on collecting several actual samples from the population, the distribution estimated in this manner is referred as *Bootstrap Distribution* [11]. This basic bootstrapping procedure is illustrated in Fig. 22.1.

It should be noted that bootstrap distribution is not same as sampling distribution. However, it has been shown that bootstrap distribution has approximately the same shape and spread as the sampling distribution. The main reasoning behind this is that if a sample is a good approximation of the population, bootstrapping will result in producing bootstrap distribution that is a good approximation of sampling distribution. For a more mathematically rigorous explanation, readers are referred to texts dedicated to thorough mathematical analysis of bootstrapping [13]. To summarize, here are the steps involved in bootstrapping (See Fig. 22.1):

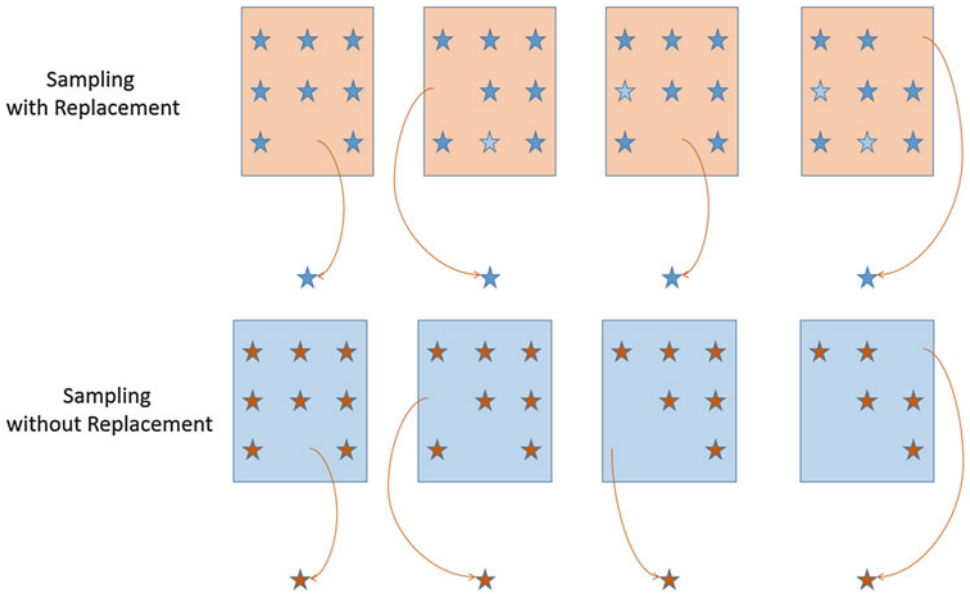


Fig. 22.2 Sampling with replacement

1. *Perform experiment*: Conduct an experiment to obtain a sample x of data.
2. *Estimate statistic of interest*: Calculate the statistic of interest (like mean age in above example) for the collected sample.
3. *Bootstrapping*
   a. *Resample*: Draw a random *sample with replacement* (bootstrap sample x^i_{BS}) from sample x. Typically size of bootstrap sample x^i_{BS} is same as that of originally collected sample x.
   b. *Bootstrap estimate*: Calculate the statistic of interest, using same settings as used in step 2, for the bootstrap sample x^i_{BS}.
   c. *Iterate*: Repeat steps 3.a and 3.b a large number of time to obtain a number of bootstrap estimates.
4. *Statistical inference*: Use bootstrap estimates for the purpose of statistical inference. Depending on the requirements, this can include approximation of sampling distribution by bootstrap distribution, calculation of standard error on statistic of interest, variance and bias calculation, estimation of confidence intervals etc.

22.2.2 Advantages and Limitations

This section discusses features of bootstrapping and highlights some of its advantages and limitations.

1. *Simplicity*: The simplicity of bootstrapping is its greatest strength and paves the way for its growing popularity. The technique is easy to adopt and understand and, in general, free of assumptions.
2. *Ease of use in complicated situations*: In complicated situations, where no straightforward mathematical formulas exist without making significant assumptions, bootstrap provides means to perform statistical inference and calculate measures of accuracy in a straightforward and 'automatic' way without employing theoretical calculations or any complicated mathematical formulas. This makes it a tool of choice no matter how complicated the statistic of interest is.

   To add to this point further, traditional methods for statistical inference rely on mathematical formulas of the measure of accuracy (for e.g. variance) or its approximation. Due to approximation, one typically requires a large sample size to correctly estimate this measure of accuracy. Further, these mathematical formulas are based on an underlying model, which may or may not be entirely correct. Yet another issue is that the mathematical formula is problem dependent, thus one has to derive a new formula for every problem. In the world of modal parameter estimation, this means deriving a different formula for different parameter estimation algorithm. This can become very tedious and, it should be further underlined, requires good knowledge of mathematics and theoretical statistics. In fact, as mentioned earlier, in certain cases, derivation of theoretical formulas, or its approximation, can be difficult or even impossible.

3. *Ease of statistical inference*: One of the significant advantages of bootstrapping, especially in context of modal analysis, is that it allows for great deal of statistical inference even in situations where experiment is conducted only once; in other words, even when only one sample is available. Not only that, bootstrapping also allows for answering a wider range of statistical inference questions in comparison to the error propagation approach. For e.g. it is possible to not only compute variance but also bias. One can plot a histogram to visually see the underlying distribution instead of making assumptions about the same. Similarly, one can make a q-q plot and check how normal the underlying distribution is. Bootstrapping also provides a means to calculate confidence intervals even in situations where underlying distribution is not known or complex enough such that it cannot be approximated by means of well-known standard distributions such as Gaussian distribution.

Despite above-mentioned advantages it is not as if bootstrapping is a cure-for-all panacea of all statistical inference problems. To begin with, a substitute or alternate more experiments or acquiring a large dataset (i.e. having a large sample). It should be looked upon as a tool that will bring value when such options of conducting multiple experiments do not exist. There are other situations where bootstrapping might not yield satisfactory results [14].

1. *Statistics based on limited observations*: Typically bootstrapping does not work well for statistics that are based on one or two observations, for e.g. median or sample maximum.
2. *Sample issues*: It does not overcome the issues related to small sample size. In such circumstances, one is not sure if the sample is a good enough reflection of the population or not and hence statistical inference can be prone to error. It should be noted however, that these issues, related to faulty sample, are not addressed by traditional methods either. Perhaps this fits with the general paradigm of '*garbage-in, garbage-out*'.
3. *Duplicate data*: If the statistic of interest is sensitive to duplicate data then bootstrapping might not be a solution. In some sense, this issue goes back to the first point. This is however not an issue with averaging, which forms the cornerstone of modal parameter estimation practice.

With this background, it is worth adopting bootstrapping for uncertainty quantification purposes in modal parameter estimation procedure. Next section is devoted to the preliminary work done in this regard by means of studies conducted on a simple single degree-of-freedom (SDOF) system.

22.3 Studies on a SDOF System

In typical EMA scenario, the time to conduct experiment is often limited on account of factors ranging from complexities involved in carrying out experiment to financial constraints. Bootstrapping can be a good option under these circumstances for providing statistical insights into obtained results and for quantifying the associated uncertainty. In this study, a SDOF system is chosen to understand and validate a bootstrapping based approach. The devised bootstrapping approach is validated by means of Monte Carlo simulations. This section starts by describing the SDOF system used before describing the design and procedure of the study including the details of bootstrapping and Monte Carlo approaches. The results and accompanying analysis are provided in Sect. 22.4.

22.3.1 SDOF System

A simple SDOF system is used in this study. This system is characterized by mass $m = 10$ Kg, damping $c = 2$ Ns/m and stiffness $k = 1600$ N/m. Based on these physical parameters, the corresponding theoretical modal parameters of the system are

Damped natural frequency (Hz): 2.0131 Hz
Modal damping (%): 0.79%

22.3.2 Numerical Study: Design and Procedure

The SDOF system based numerical study is designed to replicate general EMA procedure, which is a three step procedure:

1) Data Acquisition—Acquire input force and output response raw time histories,
2) FRF Estimation—Estimate FRFs (or other such characteristic functions) using signal processing techniques such as averaging, windowing etc., and

3) Modal Parameter Estimation—Estimate modal parameters from FRFs estimated in previous step using a chosen modal parameter estimation algorithm.

The study is aimed at replicating a typical shaker test scenario and uses H_1 algorithm [15] for FRF estimation. Additionally, the study is designed to look into effect of three different factors as they are varied;

1) Number of averages
2) Frequency resolution
3) Signal-to-Noise ratio (SNR)

The following is the procedure for bootstrapping.

1) Time history generation
   a. Generate a long time history by exciting the SDOF system with a random force signal (having unit magnitude and random phase).
   b. Depending on the chosen SNR scenario, add noise to the generated output time history.
2) FRF estimation
   a. Choose a specific block size (NB), dictated by chosen frequency resolution, and process the signals by choosing parameters such as number of averages (N_{avg}), windowing etc. to obtain averaged FRF (H_{avg}) using H_1 algorithm.
3) Data preparation
   a. In the process of calculating averaged FRFs using H_1 technique, one calculates input-output cross power spectra ($\mathbf{G}^i_{XF}$) and input auto power spectra ($\mathbf{G}^i_{FF}$) for individual blocks, which are averaged to obtain $\overline{\mathbf{G}}_{XF}$ and $\overline{\mathbf{G}}_{FF}$ before using Eq. (22.2) to obtain H_{avg}.

$$\mathbf{H}_{\text{avg}} = \overline{\mathbf{G}}_{XF}\overline{\mathbf{G}}_{FF}^{-1} \tag{22.2}$$

   General practice is to only save the averaged FRFs, H_{avg}, without saving intermediate functions such as $\mathbf{G}^i_{XF}$ and $\mathbf{G}^i_{FF}$. However, for bootstrapping, power spectra estimates corresponding to individual blocks are also saved.
4) Bootstrapping
   a. Generate a random integer sequence with repetition between 1 and N_{avg}. This is the essence of random sampling with replacement.
   b. Use individual blocks of $\mathbf{G}^i_{XF}$ and $\mathbf{G}^i_{FF}$ corresponding to random sequence generated in previous step and average them together to calculate averaged bootstrapped FRF ($\mathbf{H}^i_{BS}$). Note that, due to sampling with replacement strategy, a block might be used multiple times while calculating the averaged bootstrap FRF.
   c. Repeat this procedure N_{boot} number of times, where N_{boot} is the desired number of bootstrap samples. The result of this step will be N_{boot} number of averaged bootstrapped FRFs ($\mathbf{H}^i_{BS}$).
5) Modal parameter estimation
   a. Estimate modal parameters $\overline{MP}_{est}$ (modal frequency, ω_{est} and damping, ζ_{est}) corresponding to averaged FRF H_{avg} calculated in step 2.a. This is used to calculate the bias.
   b. Estimate modal parameters MP^i_{BS} (modal frequency, $\overline{\omega}_{BS}$ and damping, $\overline{\zeta}_{BS}$) corresponding to each of N_{boot} averaged bootstrapped FRFs $\mathbf{H}^i_{BS}$ (Step 4.c). This is used to calculate the bias. It is important to ensure that modal parameter estimation procedure chosen is set up consistently throughout Step 5. This means that same algorithm is used and various parameters of the algorithm are set the same way while estimating modal parameters from H_{avg} and $\mathbf{H}^i_{BS}$.
6) Statistical inference
   a. Based on estimated modal parameters in step 5, provide statistical inference.

In this particular case, statistical inference is focused on quantifying the uncertainty associated with these parameters. Hence, measures and visualization tools such as standard error, histogram, bias etc. are used in this study. It should be noted that this is not an exhaustive list and based on the requirements other measures and visualizations can be used.

The bootstrapping approach suggested above is validated by means of Monte Carlo simulations. In case of this numerical study, Monte Carlo simulations provide a good way to simulate the scenario where one has the possibility of performing the same experiment multiple times, which is analogous to drawing necessary statistical inference in true way by means of sampling multiple times from the population. The procedure for performing Monte Carlo simulations is as follows:

1) Time history generation
   a. Generate a long time history by exciting the SDOF system with a random force signal (having unit magnitude and random phase).
   b. Depending on the chosen SNR scenario, add noise to the generated output time history.
2) FRF estimation
   a. Choose a specific block size (NB), dictated by chosen frequency resolution, and process the signals by choosing parameters such as number of averages (N_{avg}), windowing etc. to obtain averaged FRF (H_{avg}) using H_1 algorithm.
3) Modal parameter estimation
   a. Estimate modal parameters $\overline{MP}_{est}$ (modal frequency, $\overline{\omega}_{MC}$ and damping, $\overline{\zeta}_{MC}$) corresponding to averaged FRF H_{avg} calculated in step 2.a. Make sure that modal parameter estimation is done in the same way as in case of bootstrapping.
4) Iterate
   a. Repeat above steps same number of times as in case of bootstrapping (i.e. N_{boot}).
5) Statistical inference
   a. Based on estimated modal parameters in step 4, provide statistical inference for comparison with suggested bootstrap methodology.

The above described bootstrapping and Monte Carlo procedure is shown schematically in Fig. 22.3.

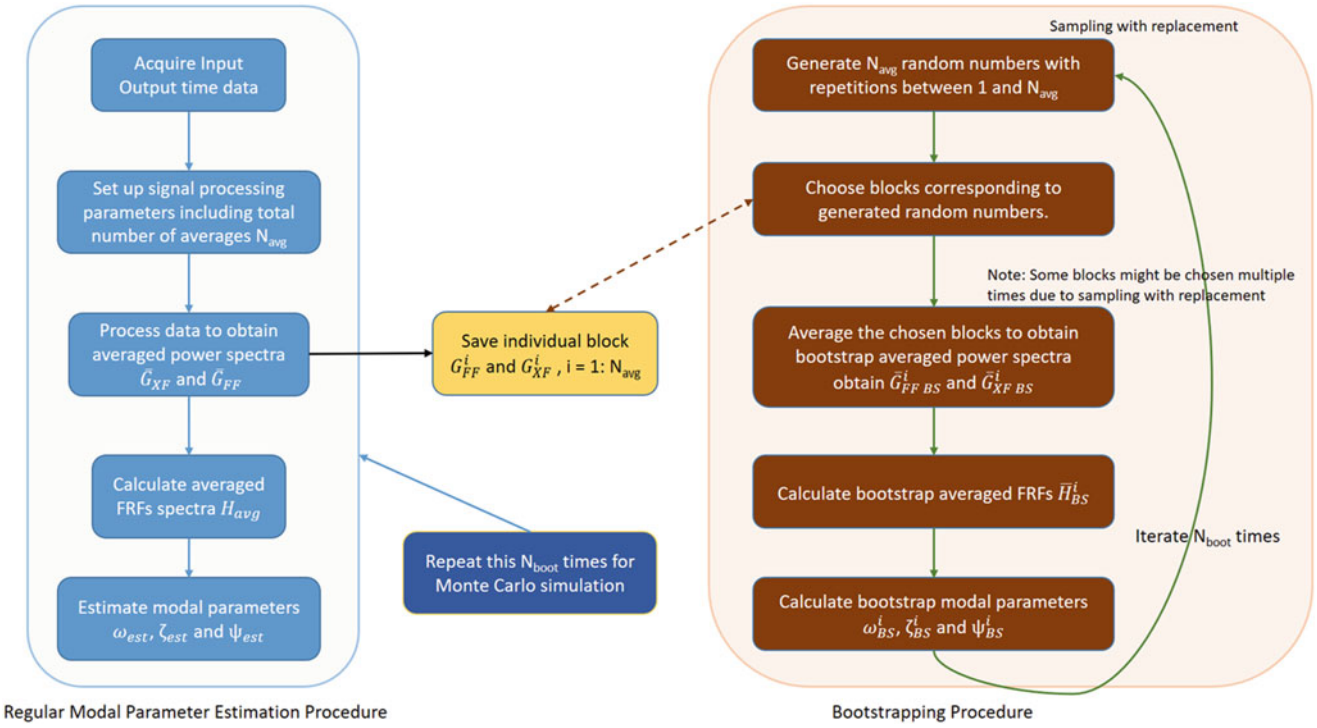


Fig. 22.3 Numerical study procedure

22.4 Results and Analysis

As mentioned in previous section, the study is aimed to validate the suggested approach by means of comparing bootstrapping with Monte Carlo simulations on the basis of three different factors that are varied; number of averages (N_{avg}), frequency resolution (NB) and SNR. The following parameters are kept constant in all the simulations.

Number of bootstrapping iterations (N_{boot})	100
Window	Rectangular
Overlapping	None
Sampling frequency (Hz)	20 Hz
FRF estimation algorithm	H_1
Modal parameter estimation algorithm	Complex Exponential Algorithm [16]
Model order	2
Number of averages	Varying as per case
Block size (Frequency Resolution)	Varying as per case

The modal parameter estimation algorithm used in this case is the Complex Exponential algorithm [16]. The algorithm is set up the same way for all calculations.

As explained in Sect. 22.2, bootstrapping results in generating several bootstrapped average FRFs (in case of this study, 100). Figure 22.4 shows these bootstrapped averaged FRFs ($\mathbf{H}^i_{BS}$), along with the averaged FRF (H_{avg}) corresponding to the sample on which bootstrapping has been performed. Due to availability of $\mathbf{H}^i_{BS}$ one can also quickly calculate *95% bootstrap percentile confidence intervals*, which is the interval between the 2.5th and 97.5th percentiles of the bootstrap distribution of a statistic. These confidence intervals (CI) for the FRF magnitude are also calculated and shown in Fig. 22.5 for illustration purposes. It should be noted that this is not the typical method for calculating confidence intervals in bootstrapping as it is very simple and applicable only in simple scenarios. In general, more advanced methods, like bootstrap bias-corrected accelerated (BCa) [11] and bootstrap tilting [11], are used for the purpose of calculating more accurate confidence intervals. However, the simpler percentile confidence intervals are calculated in this study for demonstrating that bootstrapping can also provide estimates of CIs if needed.

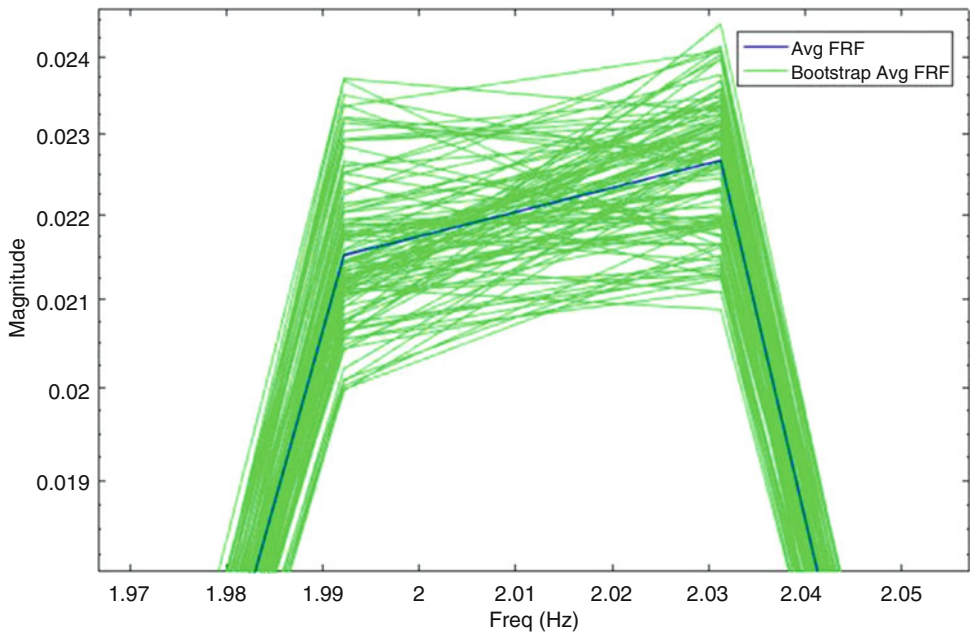


Fig. 22.4 Bootstrapped FRFs $\mathbf{H}^i_{BS}$ along with H_{avg}

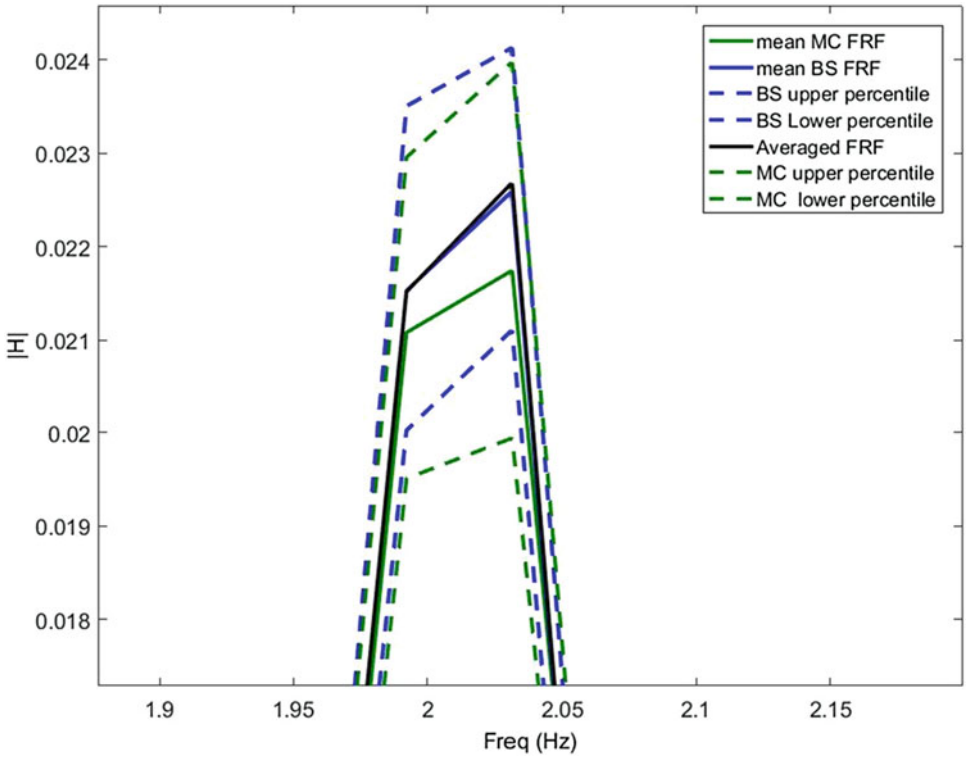


Fig. 22.5 Bootstrap and Monte Carlo means FRFs and Confidence Intervals along with H_{avg}

Figure 22.5, also shows bootstrap CIs and mean value compared with those obtained from Monte Carlo simulations along with averaged FRF (H_{avg}). This plot shows that bootstrap mean value is expectedly closer to the H_{avg}. which is based on the same sample that is used for bootstrapping. The CIs from two methods don't really overlap, but their spread is similar. It should be noted that these plots are developed for block size of 1024 and 100 averages.

22.4.1 Effect of Number of Averages

Few terminologies and symbols are introduced before the start of this section to facilitate easier understanding.

ω_{theo} , ζ_{theo}– Theoretical modal frequency and damping,

$\overline{\omega}_{BS}$, $\overline{\zeta}_{BS}$– Mean of N_{boot} bootstrapped estimates of modal frequency and damping,

$\overline{\omega}_{MC}$, $\overline{\zeta}_{MC}$– Mean of N_{boot} Monte-Carlo estimates of modal frequency and damping, and

ω_{est} , ζ_{est}—Modal frequency and damping estimated on the basis of averaged H_{avg}, following the general modal parameter estimation procedure.

A general understanding related to averaging is that it reduces random error, hence more the number of averages, closer one gets to the correct estimates. In statistical terms, number of averages is equivalent to size of the sample used to characterize the population. The bigger the sample size, the closer one is to overall population statistics. If bias errors are not present, this also means that statistics calculated on the basis of bigger samples are bound to be more accurate and precise in comparison to comparatively smaller sample. This section is devoted to understanding these aspects in relation to suggested bootstrapping procedure. Note that a block size of 1024 is used for this study.

Figures 22.6 and 22.7 show a bar plot comparing bootstrap standard errors with those from Monte Carlo simulation with increasing number of averages, for modal frequency and damping estimates. The *Bootstrap Standard Error* of a statistic is the standard deviation of the bootstrap distribution of that statistic [11]. Standard error represents how far the sample mean is likely to be from population mean. For uncertainty quantification purposes, it provides an idea about how certain one can be about the estimated statistic. A low standard error signifies that the estimate does not have much dispersion and is close to the true population statistic. As expected, these plots show that standard error reduces as the number of averages increases. This can be attributed to the growing sample size. As the sample size increases, it is expected to be more and more close

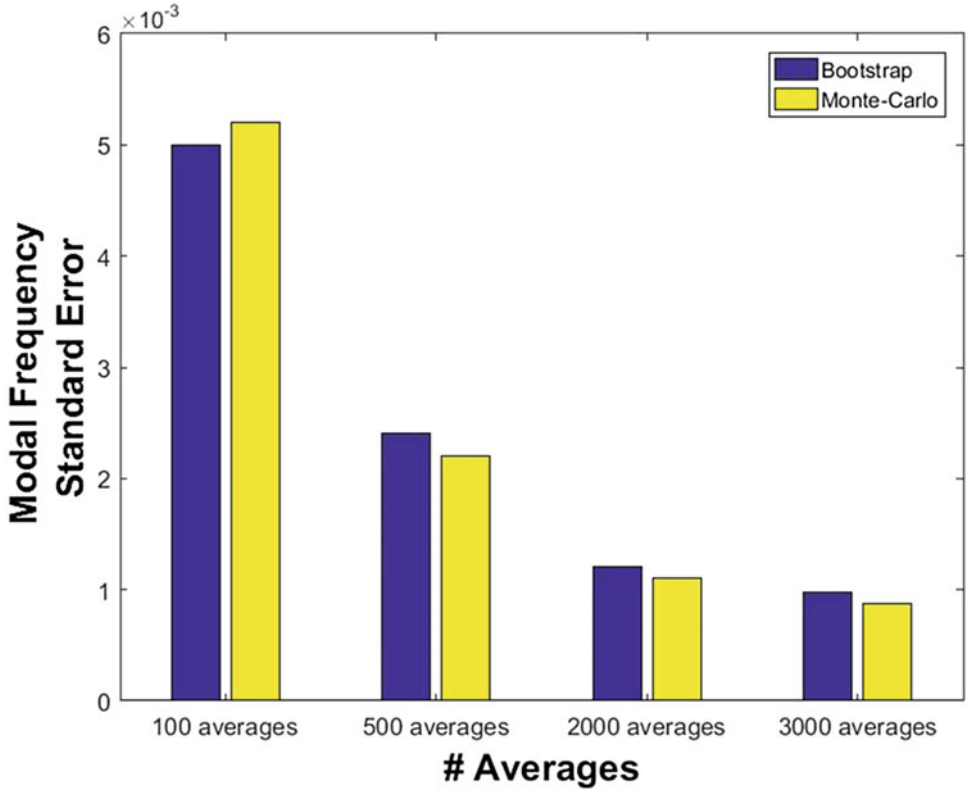


Fig. 22.6 Comparison of modal frequency standard error due to varying number of averages

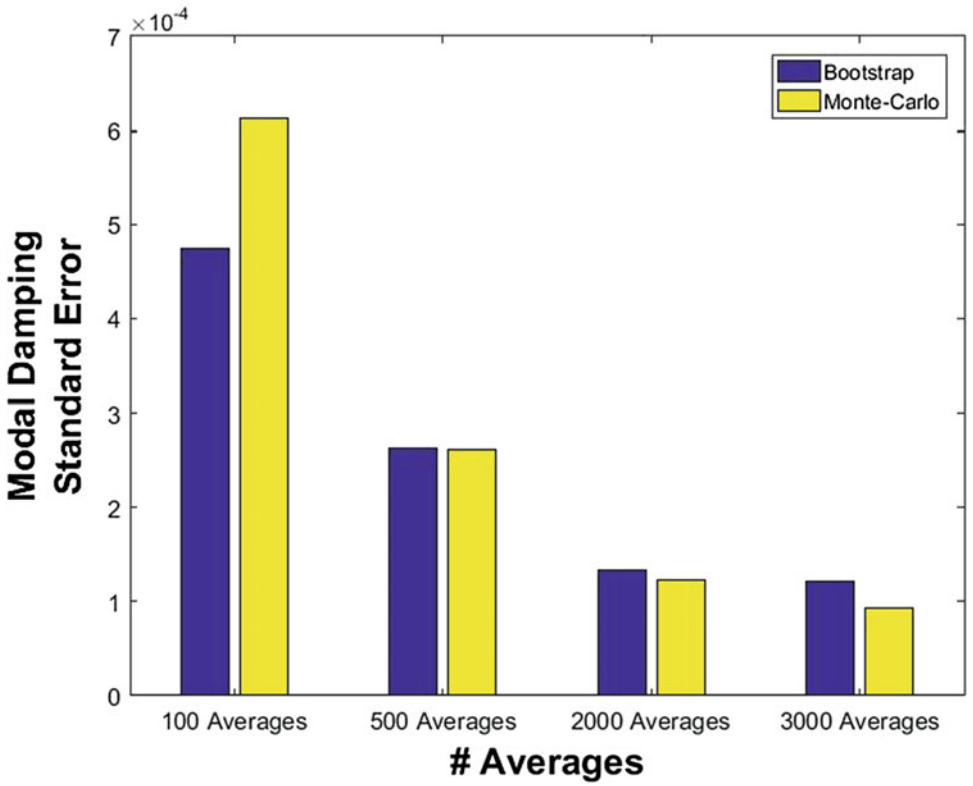


Fig. 22.7 Comparison of modal damping standard error due to varying number of averages

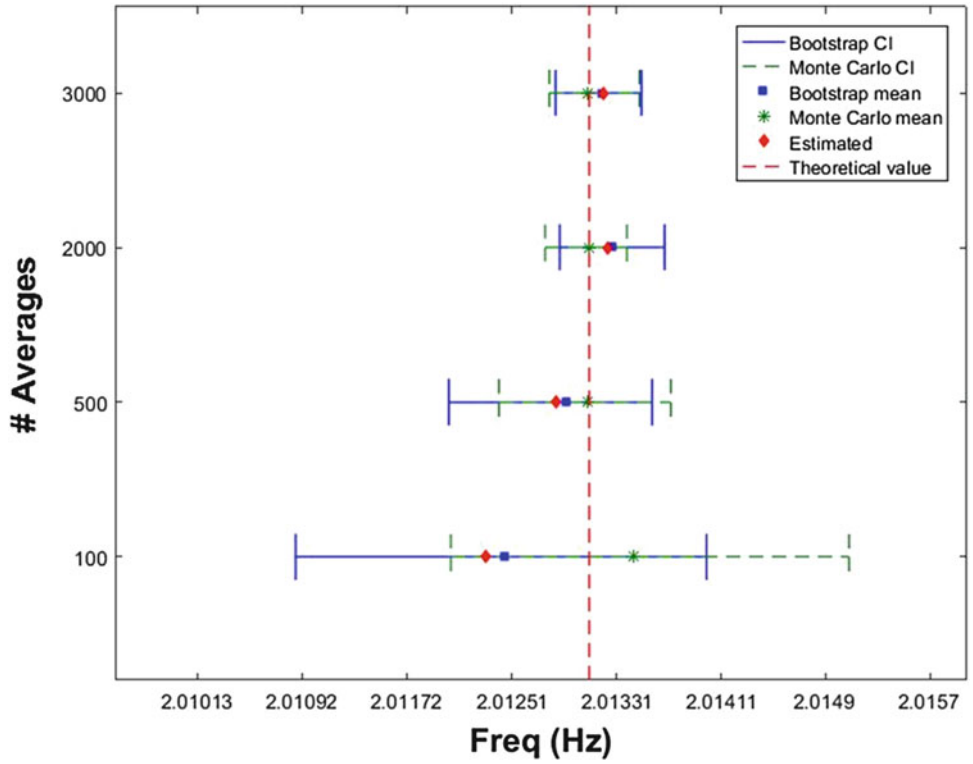


Fig. 22.8 Comparison of modal frequency 95% percentile confidence intervals due to varying number of averages

to actual population and hence the dispersion in estimated statistic will become smaller. There is also very good agreement between bootstrap and Monte Carlo results, instilling more confidence in the suggested procedure.

This effect is also seen in 95% percentile confidence intervals (Fig. 22.8). This is expected as confidence intervals (CIs) are functions of standard deviation. Unlike standard error though, bootstrap CIs do not coincide with Monte-Carlo CIs, although the range is similar. This can be attributed to the fact that whereas Monte Carlo distribution is bound to be closer to the actual sampling distribution (especially as the sample size becomes bigger), and is centered around the population mean (ω_{theo} in this case, although this need not necessarily be the case if estimation procedure introduces some bias), bootstrap distribution is centered around ω_{est}. This behavior underlines the important bootstrap principle that bootstrapping distribution mimics shape, spread and bias of the sampling distribution but not its center.

Figure 22.9 shows similar plot for damping estimates. Note that damping estimates are given as damping ratio (and not in terms of %). What stands out in this plot is that, unlike frequency estimates, damping estimates, from either bootstrapping or Monte Carlo simulations, are significantly different from the theoretical value. This brings to fore the discussion and understanding about accuracy and how it is affected by errors associated with modeling and signal processing. Both bootstrapping and error-propagation are methods for quantifying uncertainty in estimated modal parameters. These methods provide measures of accuracy of these estimates. It should however be noted that these methods view modal parameter estimation algorithm related settings and associated signal processing as intrinsic to the process. Hence, issues related to setting up algorithm correctly and using correct signal processing steps cannot be addressed by these methods. Uncertainty quantification methods only provide statistics, measures of accuracy, etc. for the specified procedure (characterized by algorithm and signal processing parameters), if this procedure is repeated multiple times. This aspect will be revisited in next section where it is shown how frequency resolution is a governing factor in getting damping estimates closer to the theoretical value.

One of the advantages of bootstrapping is that it provides various means to visually investigate shape and spread of the distribution, which is not possible with mathematical error propagation approach. One such tool is histogram as shown in Fig. 22.10 (for 3000 averages case). The 95% bootstrap percentile confidence intervals are indicated on histogram along with mean value of bootstrap estimates ($\overline{\omega}_{BS}$). In case of bootstrap, ω_{est} is also indicated. The difference in these two quantities ($\omega_{est} - \overline{\omega}_{BS}$) is an indication of bias, which is negligible in this case. To restate, *the bootstrap estimate of the bias of a statistic is difference between mean of bootstrap estimates for this statistic (in other words, mean of bootstrap distribution) and the statistic estimated from the original sample* [11].

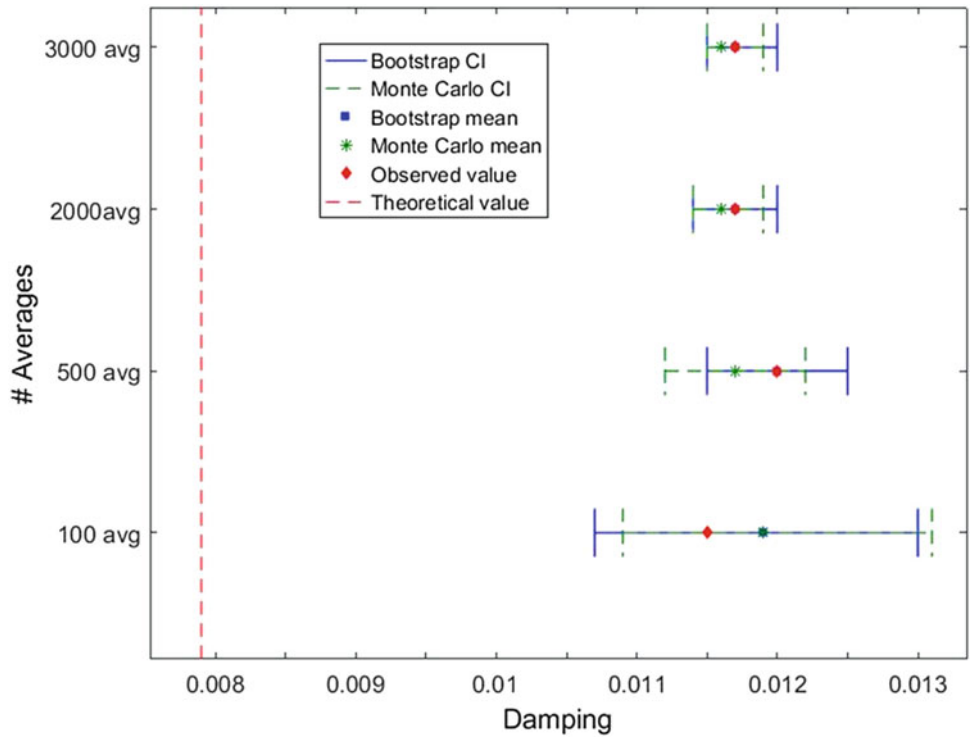


Fig. 22.9 Comparison of modal damping 95% percentile confidence intervals due to varying number of averages

Since the sample data that goes into calculating the statistic of interest is available, one can use other visualization tools like scatter plots as shown in Fig. 22.11 for 100 estimates. A quick multivariate normality test, like Henze-Zirkler's test [17], on the estimates shows that the frequency and damping estimates are multivariate normal for both bootstrapping and Monte Carlo cases. This is yet another advantage of the bootstrap approach over error propagation approach as it allows statistical tests to be performed and confirm whether the assumptions regarding data normality can be made or not.

Since the estimates are distributed normally, one can plot error ellipses (as shown in Fig. 22.11) with given confidence levels. It should be noted that the error ellipses in this case use a confidence level of 95%, which governs the scale of the ellipse. Error ellipses show the confidence curves for bivariate normal distribution and also indicate the correlation between the two variables. The error ellipse defines the region that contains 95% of all data points that can be drawn from the underlying Gaussian distribution. Figure 22.11 shows that the shape of the two ellipses is similar, although they are not centered around the same point (which is expected as bootstrapping only mimics shape and spread and not the center). It can also be said that bootstrapping frequency and damping estimates seem to be more correlated in comparison to Monte Carlo estimates (as evident from the slight tilt of the ellipse), although this difference is not significant. Note that an error ellipse with axis parallel to x or y axis signifies that the bivariate data is uncorrelated. Since a confidence level of 95% is chosen to draw error ellipses, it is expected that 95% data points (in this case frequency and damping estimates of a mode) should be inside the region defined by the ellipse. This can be easily verified by noting the number of bootstrap estimates outside the ellipse. Since $N_{boot} = 100$, it is expected that only 5 estimates should be outside the ellipse, which is nearly the case (only 4 are outside the ellipse). This is also the case with Monte Carlo simulations, as can be observed from Fig. 22.11. This is yet another aspect of bootstrapping approach over error propagation approach that due to availability of actual data, it is possible to verify the assumptions being made regarding the distribution of the estimated modal parameters.

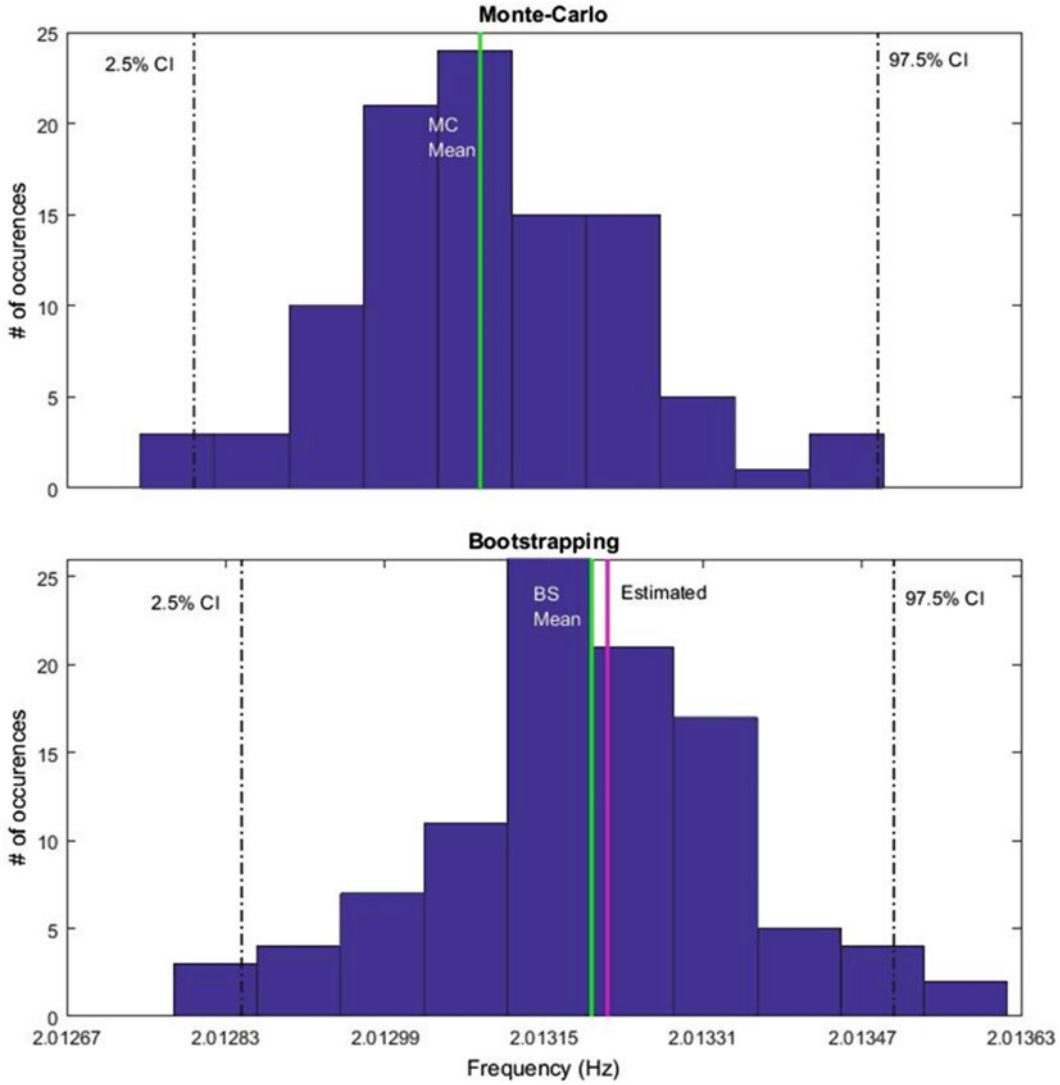


Fig. 22.10 Comparison of bootstrapping and Monte Carlo histograms for frequency estimates

22.4.2 Effect of Frequency Resolution

Previous study, aimed at understanding the effect of number of averages, is a good example highlighting that bigger sample size is beneficial as it is expected to be a good representation of the population and hence the sample statistic is expected to be close to the population statistic (or true value). This of course depends on adoption of correct procedures (such as using correct and optimal settings for signal processing, modal parameter estimation algorithms in case of EMA) for calculating the statistic of interest. If this is not the case, bias errors, for e.g. errors due to modeling, sub-optimal signal processing etc. will be introduced and resulting statistics will not be close to theoretical values.

In case of modal parameter estimation, one of the variables, and there can be many, that affects the estimated parameters is frequency resolution. The purpose of this part of the study is to show that bootstrapping is not a panacea for issues related to how parameter estimation algorithm is set up and how data is analyzed. These factors play a big role not just in variability associated with the parameters (which can be captured through bootstrapping) but also in how close they are to the true or theoretical values (which cannot be assessed as true values are not known a priori). This distinction is important to understand, as from statistical perspective, the population which one is trying to describe is already characterized by chosen signal processing parameters and parameter estimation set up, which might have introduced some bias. Hence the statistic associated with such a population need not be the same as theoretical value.

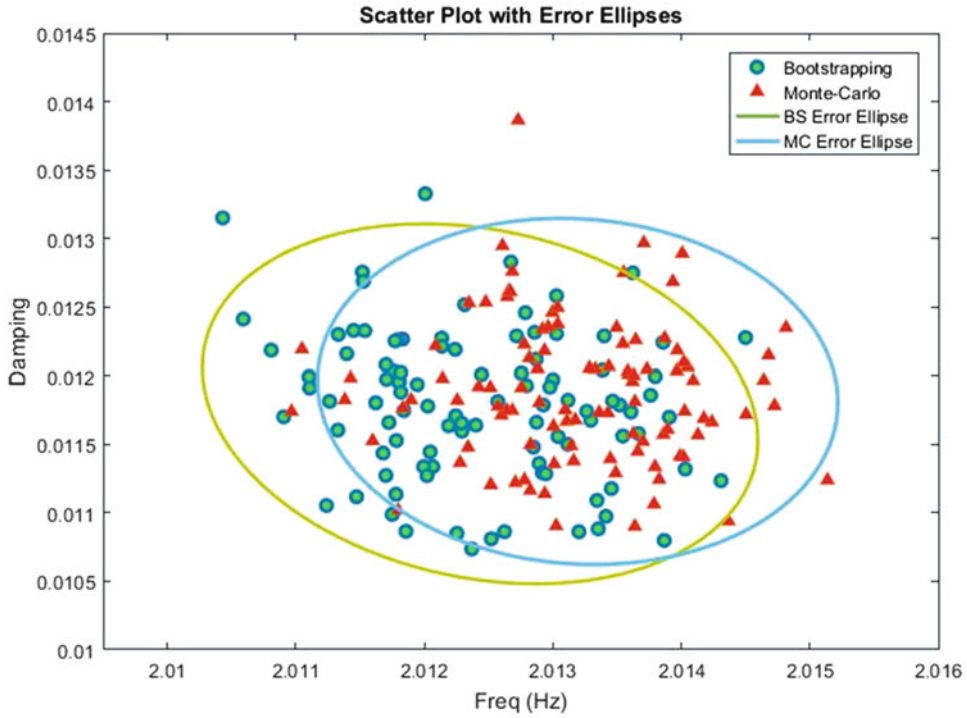


Fig. 22.11 Bootstrapping and Monte Carlo scatter plots along with error ellipses

For this study, the block size is varied (1024, 5120, 20480, 30720) resulting in subsequently finer frequency resolution (0.0391, 0.0078, 0.0020, 0.0013 Hz) and the number of averages is kept constant at 100. Figure 22.12, shows the effect of increasing frequency resolution not just on 95% percentile confidence intervals, but also how finer resolution results in damping estimates that are closer to the theoretical values. From modal analysis perspective, this highlights that signal processing and parameter estimation algorithm settings play a major role in terms of ultimate accuracy of the estimated parameters and their importance cannot be underestimated, circumvented or compensated by means of error quantification techniques such as bootstrapping or error propagation approach. Figure 22.12 also shows a good agreement between results of bootstrapping and Monte Carlo simulations. Similar agreement in results is obtained for modal frequency estimates as well (though they are not illustrated here). The narrowing of CIs with better frequency resolution is again on expected terms. As the quality of data improves, the estimates become more and more precise. Except for the lowest resolution (1024 block size), the bootstrap mean ($\overline{\zeta}_{BS}$) and estimated values (ζ_{est}) more or less coincide with each other, indicating that there's no bias in the damping estimates.

Figures 22.13 and 22.14 show the comparison on the basis of estimated standard errors on modal frequency and damping estimates. As these figures show, there is a good general agreement between bootstrapping and Monte Carlo simulations, validating the suggested approach.

22.4.3 Effect of Noise

Final part of this study is devoted to verifying whether suggested approach works when dealing with noisy data. Since H_1 estimator is being used in this study, white Gaussian noise of given SNR is added to the output time response signal. In addition to the studies conducted in previous sections that had no additional noise, three more 'noisy' scenarios are considered, with SNR of 50 dB, 30 dB and 10 dB.

Figures 22.15 and 22.16 show estimated bootstrap and Monte Carlo standard errors in modal frequency and damping estimates for varying SNR and number of averages. It can be observed that for most cases there is a good agreement between standard error estimated using bootstrapping and Monte Carlo simulations, except when noise is very significant (SNR 10

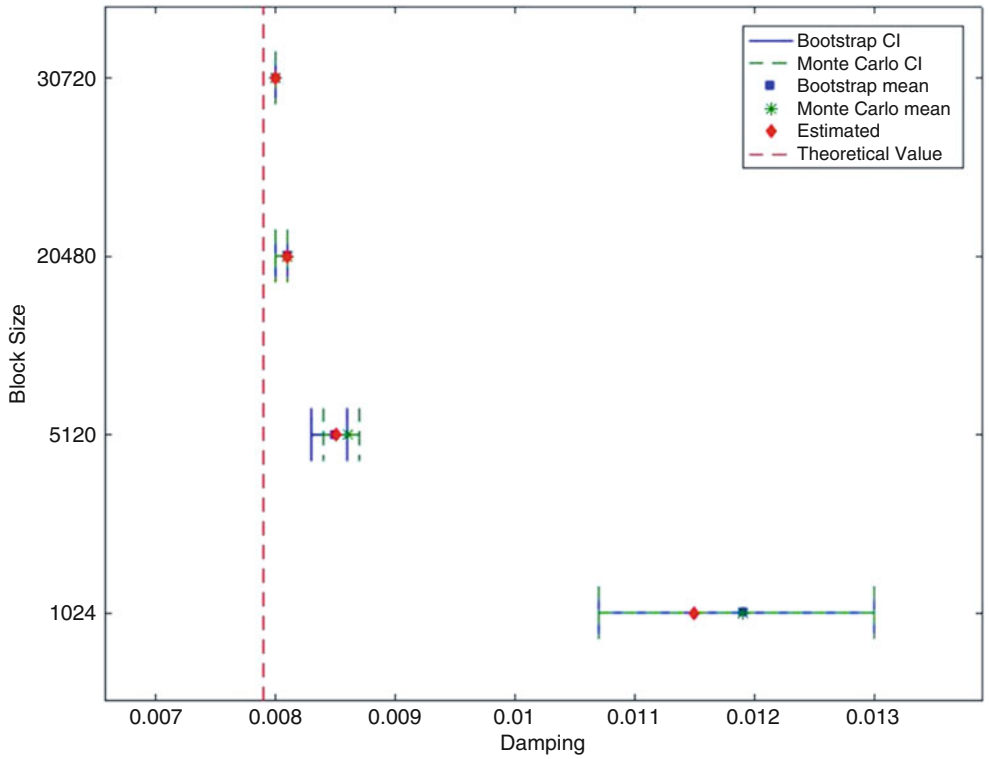


Fig. 22.12 Comparison of modal damping 95 % percentile confidence intervals due to varying frequency resolution

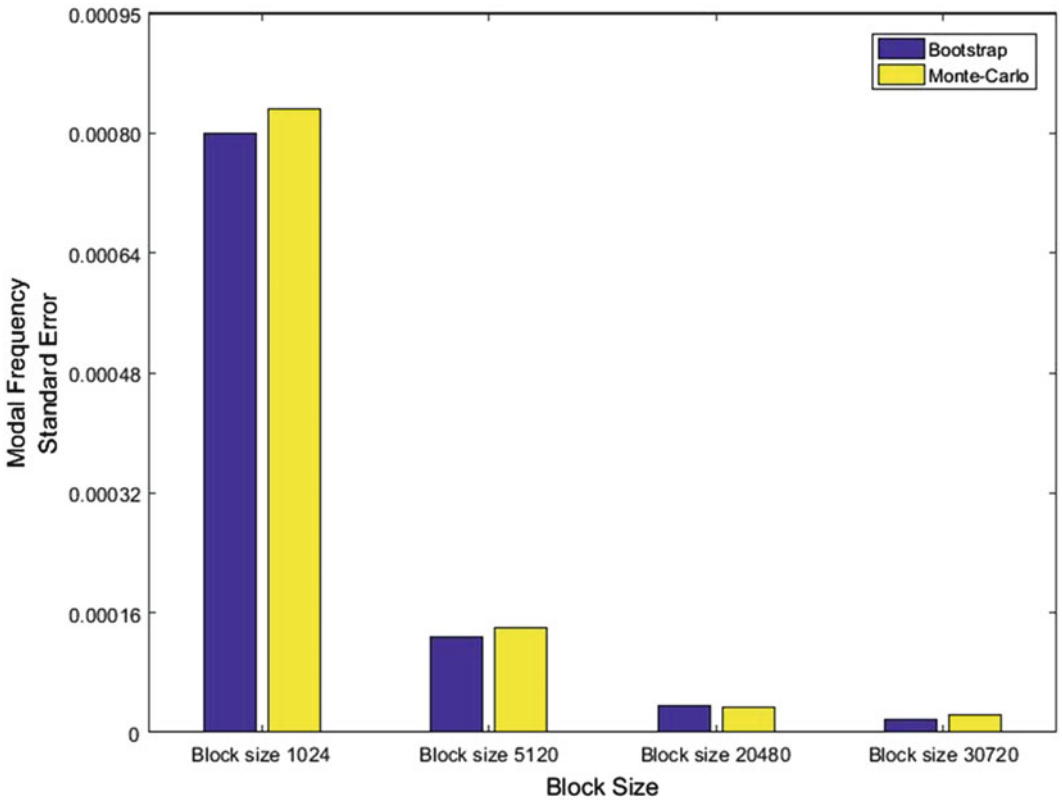


Fig. 22.13 Comparison of modal frequency standard error due to varying frequency resolution

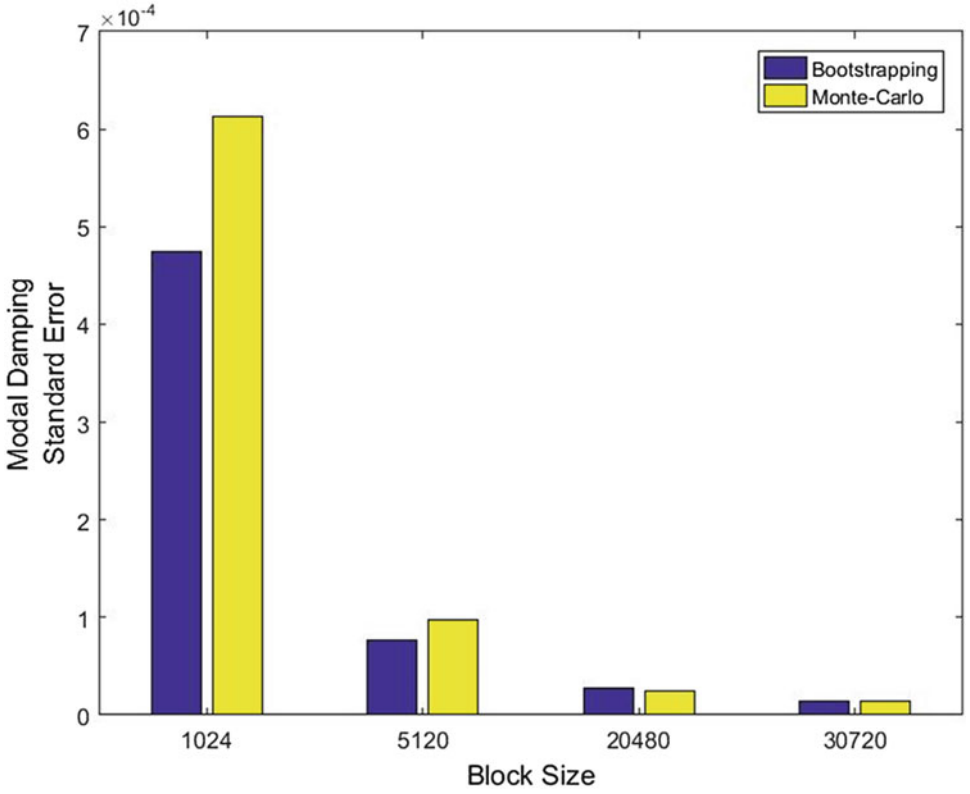


Fig. 22.14 Comparison of modal damping standard error due to varying frequency resolution

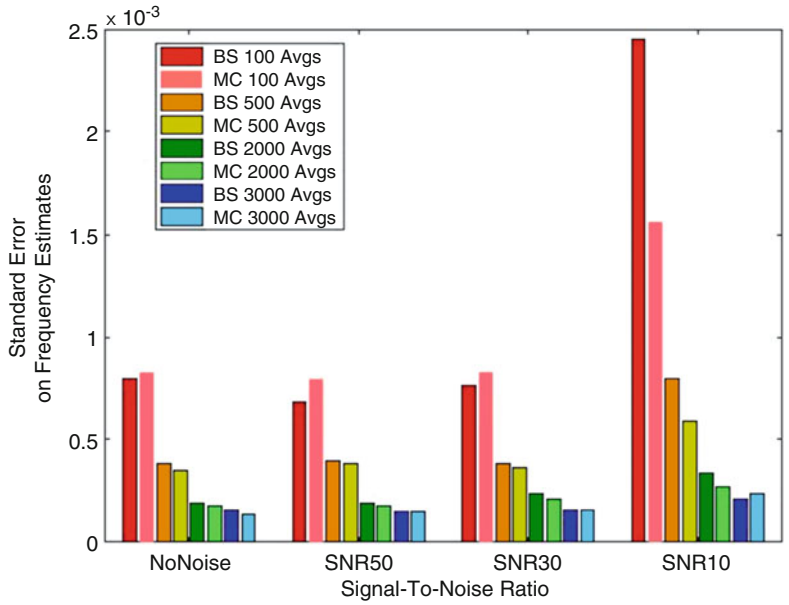


Fig. 22.15 Comparison of modal frequency standard error due to varying SNR and number of averages

case), in which case the two estimates differ from each other. This difference however, diminishes as the number of averages is increased. This is another good indication of how a bigger sample is more likely to capture the essence of the population.

This noise related study was also repeated for varying frequency resolution case and the observed results were on similar lines as this case. For the sake of conciseness, these results are not reported in this paper.

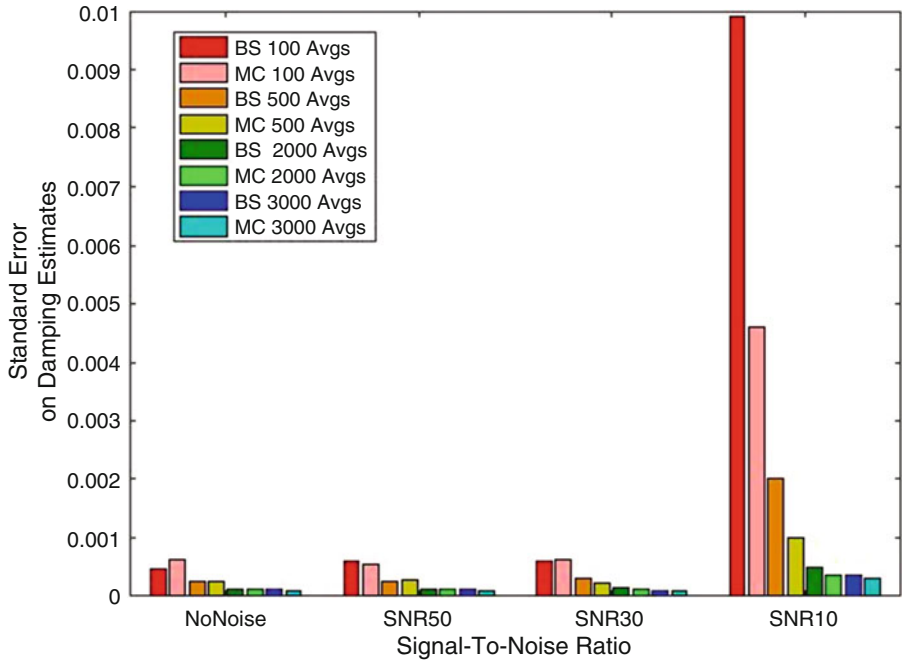


Fig. 22.16 Comparison of modal damping standard error due to varying SNR and number of averages

22.5 Conclusions

Quantifying uncertainty in estimated modal parameters is an important task that is gaining increasing attention in structural dynamics research community. In this regard, a bootstrapping based methodology is introduced in this paper. The paper highlights the fundamental principles of bootstrapping, which is a statistical method based on resampling the original sample. The methodology is presented by explaining bootstrapping in the context of modal parameter estimation, thereby also providing a statistical perspective to parameter estimation procedure.

It is highlighted in the paper how this approach differs from error propagation approach of uncertainty quantification and it is shown that suggested bootstrapping methodology provides a good estimate of shape, spread and bias of the underlying distribution. Significantly, unlike the error-propagation approach, it does not make any assumptions or simplifications while calculating the statistics of interest. Another advantage of this method is that it provides an estimate of bias, which is not possible with error-propagation approach. The suggested methodology also makes it possible to utilize several tools of statistical inference, including visualization tools, such as histograms, scatter plots, box plots, etc. Suggested approach has been validated by means of Monte Carlo simulations and a good agreement in results augurs well for the proposed technique. Overall, studies conducted in this work show that bootstrapping is a simple yet powerful technique that can be adopted in modal analysis domain for the purpose of uncertainty quantification.

References

1. Reynders, E., Pintelon, R., De Roeck, G.: Uncertainty bounds on modal parameters obtained from stochastic subspace identification. Mech. Syst. Signal Process. **22**, 948–969 (2008)
2. Bendat, J.S., Piersol, A.G.: Engineering Applications of Correlation and Spectral Analysis, Second edn. John Wiley, NY, USA (1993)
3. Pintelon, R., Guillaume, P., Schoukens, J.: Uncertainty calculation in (operational) modal analysis. Mech. Syst. Signal Process. **21**, 2359–2373 (2007)
4. Brincker, R., Andersen, P.: Understanding stochastic subspace identification. In: Proceedings of 24th International Modal Analysis Conference (IMAC), St. Louis (MO), USA (2006)
5. Dohler, M., Lam, X.B., Mevel, L.: Efficient multi-order uncertainty computation for stochastic subspace identification. Mech. Syst. Signal Process. **38**, 346–366 (2013)
6. Dohler, M., Lam, X.B., Mevel, L.: Uncertainty quantification for modal parameters from stochastic subspace identification on multi-setup measurements. Mech. Syst. Signal Process. **36**, 562–581 (2013)
7. Taylor, J.R.: An Introduction to Error Analysis, 2nd edn. University Science Books, Sausalito, CA, (1997)

8. Hunter, N.F., Paez, T.L.: Application of the bootstrap to the analysis of vibration test data. In: 66th Shock and Vibration Symposium, Biloxi, MS (1995)
9. Farrar, C.R., Doebling, S.W., Cornwell, P.J.: A comparison study of modal parameter confidence intervals computed using the Monte Carlo and bootstrap techniques. In: Proceedings of the 16th IMAC, Santa Barbara, CA, USA (1998)
10. Efron, B.: Bootstrap methods: another look at the jackknife. Ann. Stat. **7**, 1–26 (1979)
11. Efron, B., Tibshirani, R.J.: An Introduction to the Bootstrap. Chapman and Hall, New York (1993)
12. Lefebvre, M.: Applied Probability and Statistics. Springer, New York (2006)
13. Horowitz, J.L.: The bootstrap. Handb. Econ. **5**, 3159–3228 (2001)
14. Haukoos, J.S., Lewis, R.J.: Advanced statistics: bootstrapping confidence intervals for statistics with "difficult" distributions. Acad. Emerg. Med. **12**(4), 360–365 (2005)
15. Heylen, W., Lammens, S., Sas, P.: Modal analysis theory and testing. PMA Katholieke Universteit, Leuven (1995)
16. Brown, D.L., Allemang, R.J., Zimmerman, R., Mergeay, M.: Parameter estimation techniques for modal analysis. In: SAE Paper No. 790221, SAE Transactions, vol. 88, pp. 828–846 (1979)
17. Henze, H., Zirkler, B.: A class of invariant consistent tests for multivariate normality. Commun. Stat. Theory Methods. **19**(10), 3595–3617 (1990)

Chapter 23
MPUQ-b: Bootstrapping Based Modal Parameter Uncertainty Quantification—Methodology and Application

S. Chauhan

Abstract Building on a previous paper that illustrated the fundamental principles of bootstrapping, and how it can be employed in context of quantifying uncertainty in modal parameter estimation, a new methodology for uncertainty quantification in estimated modal parameters is proposed. This methodology, termed as Bootstrapping based Modal Parameter Uncertainty Quantification (MPUQ-b), utilizes bootstrapping for the purpose of statistical inference regarding estimated modal parameters. This second paper, elaborates and demonstrates MPUQ-b methodology and shows how bootstrapping can be incorporated in modal parameter estimation process. Suggested method is validated by means of comparison with Monte Carlo simulation studies on a numerical five degrees-of-freedom system. It is highlighted in the paper, how MPUQ-b differs from other methods available in the literature and its advantages and limitations are discussed.

Keywords Bootstrapping • Uncertainty quantification • Modal parameter estimation • Statistical inference • Standard error

Abbreviations

DOF	Degrees of Freedom
EMA	Experimental Modal Analysis
ERA	Eigensystem Realization Algorithm
FFT	Fast Fourier Transformation
FRF	Frequency Response Function
IQR	Interquartile Range
MIMO	Multiple Input, Multiple Output
MPUQ-b	Bootstrapping based Modal Parameter Uncertainty Quantification
OMA	Operational Modal Analysis
PTD	Polyreference Time Domain algorithm
RFP-z	Rational Fraction Polynomial Z domain
SDOF	Single Degree-of-Freedom
SNR	Signal-to-Noise Ratio
SSI-Cov	Covariance driven Stochastic Subspace Identification algorithm

Nomenclature

$\mathbf{M}$, $\mathbf{C}$ *and* $\mathbf{K}$	Mass, damping and stiffness matrices
ω_{theo}, ζ_{theo}	Theoretical modal frequency and damping
$\mathbf{N_{avg}}$	Number of averages
$\mathbf{H_{avg}}$	Averaged FRF estimated using H_1 technique on original data
ω_{est}, ζ_{est} and ψ_{est}	Modal frequency, damping and mode shape estimated on the basis of averaged FRF, $\mathbf{H_{avg}}$
$\mathbf{G}^i_{XF}$, $\mathbf{G}^i_{FF}$	Input-output cross and input-input auto power spectra for individual blocks used for estimating $\mathbf{H_{avg}}$

S. Chauhan (✉)
Bruel & Kjær Sound and Vibration Measurement A/S, Skodsborgvej 307, DK 2850, Naerum, Denmark
e-mail: schauhan@bksv.com

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_23

$\mathbf{H}_{BS}^{i}$	i^{th} averaged bootstrap FRF
ω_{BS}^{i}, ζ_{BS}^{i} and ψ_{BS}^{i}	Modal frequency, damping and mode shape estimated on the basis of i^{th} averaged bootstrap, $\mathbf{H}_{BS}^{i}$
N_{boot}	Number of bootstrap repetitions. Same for number of times Monte Carlo simulations are run.
σ	Standard deviation of physical system parameters perturbation
$\overline{\omega}_{BS}$, $\overline{\zeta}_{BS}$	Mean of $\mathbf{N_{boot}}$ bootstrapped estimates of modal frequency and damping
$\sigma_{BS\omega}$, $\sigma_{BS\zeta}$	Standard deviation of $\mathbf{N_{boot}}$ bootstrapped estimates of modal frequency and damping
$\overline{\omega}_{MC}$, $\overline{\zeta}_{MC}$	Mean of $\mathbf{N_{boot}}$ Monte Carlo estimates of modal frequency and damping
$\sigma_{MC\omega}$, $\sigma_{MC\zeta}$	Standard deviation of $\mathbf{N_{boot}}$ Monte Carlo estimates of modal frequency and damping

23.1 Introduction

When a structure is being designed from dynamics point of view, it is important to characterize its dynamics, or in other words ascertain its dynamic properties. A dynamic system can be characterized by several means, including modal parameters. As Eqs. (23.1.a), (23.1.b), and (23.1.c) show, it is possible to characterize a MIMO dynamic system in terms of its physical parameters, **M**, **C** and **K** matrices (Physical model representation), or in terms of frequency response functions, **H** (Response model representation) or associated modal parameters (Modal model representation). The dynamics design process often comprises of two stages. In the first stage, a numerical model of the system is created. This model makes assumptions with regards to the physical properties of the system and uses the physical model to design the system and obtain corresponding modal parameters by solving the characteristic equation as shown in Eq. (23.1.a).

$$-\omega^2\,\mathbf{M} + j\omega\,\mathbf{C} + \mathbf{K} = 0 \tag{23.1.a}$$

$$\mathbf{H}(\omega) = \frac{\mathbf{X}(\omega)}{\mathbf{F}(\omega)} = \frac{1}{(-\omega^2\,\mathbf{M} + j\omega\,\mathbf{C} + \mathbf{K})} \tag{23.1.b}$$

$$\mathbf{H}(\omega) = \sum_{i=1}^{N} \frac{\mathbf{A}_i}{j\omega - \lambda_i} + \frac{\mathbf{A}_i^*}{j\omega - \lambda_i^*} \tag{23.1.c}$$

This process would have sufficed but for the fact that typically these numerical models are too simplistic to represent the system in real world scenario. For e.g., it is difficult to accurately model the damping in the system or the boundary conditions and joints. Thus, it is necessary to validate these models before they can be utilized. This leads to the second stage and brings into effect the experimental side of modal analysis. Equation (23.1.b) forms the basis of this stage. Unlike the first stage, that utilized the physical model, here one uses the response model representation. Thus an experiment is conducted by exciting the structure by a known force and the response of the structure to this excitation is measured. The two measured quantities are used to obtain frequency response functions by using Eq. (23.1.b). Subsequently, a modal parameter estimation algorithm is used that exploits the relationship in Eq. (23.1.c), or its equivalent, to obtain the modal parameters from experimental data. These experimentally obtained modal parameters are now used for comparing and validating the results provided by the numerical model. Thus, modal parameters play a vital role in this process by providing the bridge between physical and response models. The use of modal parameters is not just limited to this, they are utilized for several other purposes including finite element model updating, design optimization, troubleshooting, damage detection etc. It is therefore important that they are estimated as accurately as possible.

Over the years, several algorithms and signal processing methods have been developed that have enabled modal analysis to progress beyond research table and being adopted as a standard tool in industry. These advance algorithms aim at estimating modal parameters reliably and accurately. Yet, the very nature of experimental modal analysis indicates that no matter what, there will be some variability associated with the results obtained. The measured data always has some intrinsic uncertainty associated with it on account of several factors including associated measurement noise, variable material properties, complicated boundary conditions and several others. Thus, modal parameters obtained from this data will also have some uncertainty associated with it. Signal processing methods and further assumptions made during the modal parameter estimation stage (processing measured data to obtain modal parameters), add further uncertainty to the estimated parameters. Despite this understanding, the common practice is to provide modal parameters as they are, without any means to quantify, or qualify, the associated uncertainty.

In recent years, however, there have been efforts to address this aspect. The general methodology of most of the works looking at uncertainty quantification in estimated modal parameters is that of error propagation by means of sensitivity

analysis. In case of experimental modal analysis (EMA), where both input excitation and output response is measured, the starting point of the error propagation approach is typically the frequency response functions (FRF) estimated from the measured data. Based on the work presented in [1], the standard deviation of these FRFs can be estimated using the coherence functions [2]. Standard deviation is a measure of variability (or uncertainty) present in the estimated FRFs. The task is then to propagate this variability in FRFs to modal parameters estimated from these FRFs. The general framework is that of error propagation by means of exploiting the mathematical functional relationship between two variables [3]. Mathematically this is expressed by Eq. (23.2), which shows that if θ and β are two variables such that θ is a function of β and variance of beta ($\Sigma_{\widehat{\beta}}$) is known, then variance of θ ($\Sigma_{\widehat{\theta}}$) can be obtained using this relation by mathematically computing the gradient of θ with respect to β $\left(\left(\frac{\partial\theta}{\partial\beta}\right)_{\widehat{\beta}}\right)$ and evaluating it at point estimate of β, i.e. $\widehat{\beta}$.

$$\Sigma_{\widehat{\theta}} = \left(\frac{\partial\theta}{\partial\beta}\right)^{\mathrm{T}}_{\widehat{\beta}} \Sigma_{\widehat{\beta}} \left(\frac{\partial\theta}{\partial\beta}\right)_{\widehat{\beta}} \tag{23.2}$$

In case of modal analysis, β is frequency response functions, $\mathbf{H}$ and θ are modal parameters, ω_{est}, ζ_{est} and ψ_{est} (Modal frequency, damping and mode shapes). As can be seen from Eq. (23.2), the main task involved in this approach is to obtain the expression for the gradient ($\left(\frac{\partial\theta}{\partial\beta}\right)$) and its evaluation. The approach is also referred to as sensitivity analysis, a technique for calculating the sensitivity of a quantity with respect to change or variability in another quantity (which essentially means evaluating the gradient term in Eq. (23.2)).

Though simple in terms of understanding, the error propagation approach can be quite complicated and challenging. Most modal parameter estimation algorithms do not estimate modal parameters directly from FRFs. Instead, they first estimate an intermediary quantity from which modal parameters are estimated (Fig. 23.1). For e.g., in case of Polyreference Time Domain [4, 5] (PTD), the FRFs are processed to obtain polynomial coefficients and modal parameters are estimated from these coefficients. Thus, Eq. (23.2) is effectively applied twice and the error is propagated from FRFs to modal parameters in various stages. More importantly, the functional relationship that relates intermediary functions with the modal parameters is not linear. As a consequence, Eq. (23.2) is approximate, only holds up to higher terms and might fail in case of low SNR [6]. The assumptions that one has to make in order to apply this methodology are thus more stringent.

The other issue with this approach is that it is algorithmic specific. The derivation of the gradient term differs from algorithm to algorithm; a process which is mathematically overwhelming, cumbersome and not very intuitive. It should

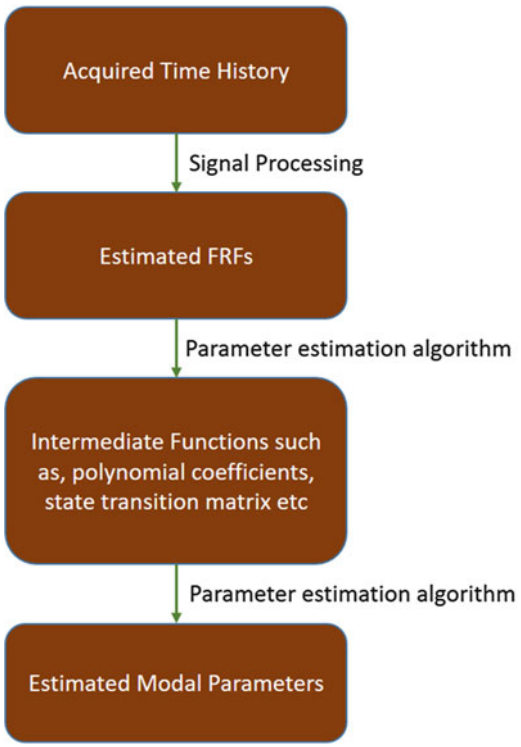


Fig. 23.1 Modal parameter estimation procedure

however be mentioned that once the expression for the gradient has been derived, the error propagation is fast and computationally efficient. Finally, this method assumes that errors in input forces and responses are distributed normally and so is the case with estimated modal parameters, an assumption that is key to error propagation and may or may not be true in real world scenarios.

This approach was used in [7] to develop expressions for variance and bias in modal parameters estimated using Eigensystem Realization algorithm [8]. Similar approach is adopted in [9], where covariance matrices of estimates obtained from Maximum Likelihood and prediction error based parameter estimation methods are used to obtain the uncertainty bounds. In [6], this approach is utilized within Operational modal analysis framework for Covariance driven Stochastic Subspace Identification (SSI-Cov) [10] algorithm. The paper provides expressions of covariance of system matrices (state transition matrix etc.) as well as those of modal parameters based on the error propagation approach. The standard deviation of output correlation functions is propagated first to system matrices (characterizing the state space model), and then subsequently propagated to the modal parameters identified using these system matrices. This work provides the basis for [11, 12], which suggest a fast efficient approach for efficiently computing uncertainty related information in case of SSI and how it can be utilized in a multi-setup scenario. Above mentioned works adopt the error-propagation approach to derive expressions specific to various modal parameter estimation algorithms in both experimental and operational modal analysis framework, highlighting the algorithm specific nature of this approach.

The work presented in this paper departs from the mathematically intensive error propagation approach and instead uses a statistical approach to quantify the uncertainty associated with estimated modal parameters. The method used in this paper is based on a resampling technique called *Bootstrapping*. [13] explained fundamentals of bootstrapping and highlighted its various aspects in relation to modal analysis by means of studies conducted on a single degree-of-freedom (SDOF) system. The paper illustrates how despite being a simple technique, bootstrapping provides ways to calculate several measures of accuracy that bring statistical perspective to the modal parameter estimation procedure and shed valuable insights regarding their accuracy, enabling quantification of associated uncertainty. For more general information about bootstrapping, readers should refer [14], which is an established resource on bootstrapping.

This paper is an extension of [13] as it delineates a bootstrapping based uncertainty quantification approach for modal parameters. This approach is termed as *Bootstrapping based Modal Parameter Uncertainty Quantification* (MPUQ-b). The methodology builds on the concepts developed in [13] and presents a step-by-step strategy to attune bootstrapping within modal analysis parameter estimation framework. Section 23.2 of the paper is devoted to MPUQ-b, where this methodology is described along with its features vis-à-vis error propagation approach. The methodology is then validated by means of numerical studies conducted on a 5 degrees-of-freedom (DOF) system. It is shown that MPUQ-b provides similar results as those obtained using Monte-Carlo simulations. The key highlight of the suggested methodology is that it enables variety of statistical data analysis and visualization tools that help to not just quantify the uncertainty but also gain valuable insights about estimated modal parameters. The paper ends with concluding remarks regarding MPUQ-b.

23.2 MPUQ-b: Bootstrapping Based Modal Parameter Uncertainty Quantification

23.2.1 Procedure

Bootstrapping is a computer based resampling method of statistical inference using random sampling with replacement. It facilitates calculation of measures of accuracy, such as variance, bias, confidence intervals, sampling distribution etc. It also makes it possible to use several data visualization tools, some of which, like histograms, scatter plots etc., have been introduced in [13]. MPUQ-b adapts bootstrapping within the modal parameter estimation framework and is described in terms of following steps. It should be noted that the procedure described here is designed for Shaker based testing.

1. *Data acquisition*: Conduct a modal test and acquire the input excitation and output response data.
2. *FRF estimation and bootstrapping data preparation*: Estimate FRFs from data acquired in step 1. Typically, FRFs are estimated using a chosen estimator (H_1, H_2 or H_v) [2] and by employing signal processing techniques like averaging and windowing etc. The result of this step is estimated averaged FRFs ($\mathbf{H_{avg}}$). In general, it is only these averaged FRFs that are saved for further use and individual blocks of auto and cross power spectra that go into estimating these averaged FRFs, are not saved. However, these individual blocks are saved in case of MPUQ-b as they form the basis of performing bootstrapping.
3. *Modal parameter estimation*: Estimate modal parameters i.e. modal frequency, damping and mode shapes (ω_{est}, ζ_{est} and ψ_{est}) from the averaged FRFs calculated in step 2 using a parameter estimation algorithm of choice. Step 2 and 3 are

very critical from the point of view of accurate estimation of modal parameters as these steps dictate the modeling error associated with estimated modal parameters. Hence, it is important that signal processing and parameter estimation algorithm settings are chosen carefully, so that the modeling error is minimized. This is also important from the perspective of bootstrapping because bootstrapping employs same signal processing and parameter estimation algorithm settings, as chosen in this step. Thus, one needs to follow general guidelines and adopt well established best practices (such as use of correct frequency resolution, choosing correct model order, etc.) for accurately estimating modal parameters.

4. *Bootstrapping*: This step is as explained in section 3.2 of [13]. A random integer sequence with repetition is generated between 1 and N_{avg}, where N_{avg} is number of averages used in step 2. Individual auto and cross power spectra blocks ($\mathbf{G}^i_{XF}$, $\mathbf{G}^i_{FF}$) corresponding to generated random integer sequence are averaged together to calculate averaged bootstrapped FRFs ($\mathbf{H}^i_{BS}$) and this process is repeated N_{boot} number of times, N_{boot} being the desired number of bootstrap samples. This results in N_{boot} number of averaged bootstrapped FRFs ($\mathbf{H}^i_{BS}$, i = 1: $\mathbf{N_{boot}}$).
5. *Bootstrap modal parameters* ($\omega^i_{BS}, \zeta^i_{BS}$ and ψ^i_{BS}): Estimate modal parameters corresponding to all N_{boot} averaged bootstrapped FRFs $\mathbf{H}^i_{BS}$ from Step 4. It is important to ensure that modal parameter estimation algorithm is set up in similar manner as step 3.

It is important to understand that in real world situations parameter estimation procedure results not only in estimation of system modes but also of computational modes. This might be due to several factors, including presence of noise, modeling errors, etc. One of the tools that is often utilized to help make a proper decision in this regard is *Stabilization Diagram* [2]. MPUQ-b utilizes an approach similar to stabilization diagram for the purpose of filtering computational modes. Unlike modal parameter estimation in regular scenario (as used in step 3), where stabilization diagram is constructed on the basis of varying model order, in case of MPUQ-b, stabilization diagram is constructed on the basis of varying bootstrap datasets. Thus at each iteration, a different set of averaged bootstrap FRFs ($\mathbf{H}^i_{BS}$) are used. This approach helps in performing a cluster based analysis of the estimated modal parameters, paving way for straight forward statistical inference.

6. *Statistical inference*: Bootstrapped modal parameters estimated in step 5 are used for gaining statistical insights into the modal parameter estimation procedure used for estimating these parameters. Various qualitative and quantitative methods of statistical inference can be now utilized depending on the requirements.

This procedure is described graphically in Fig. 23.2.

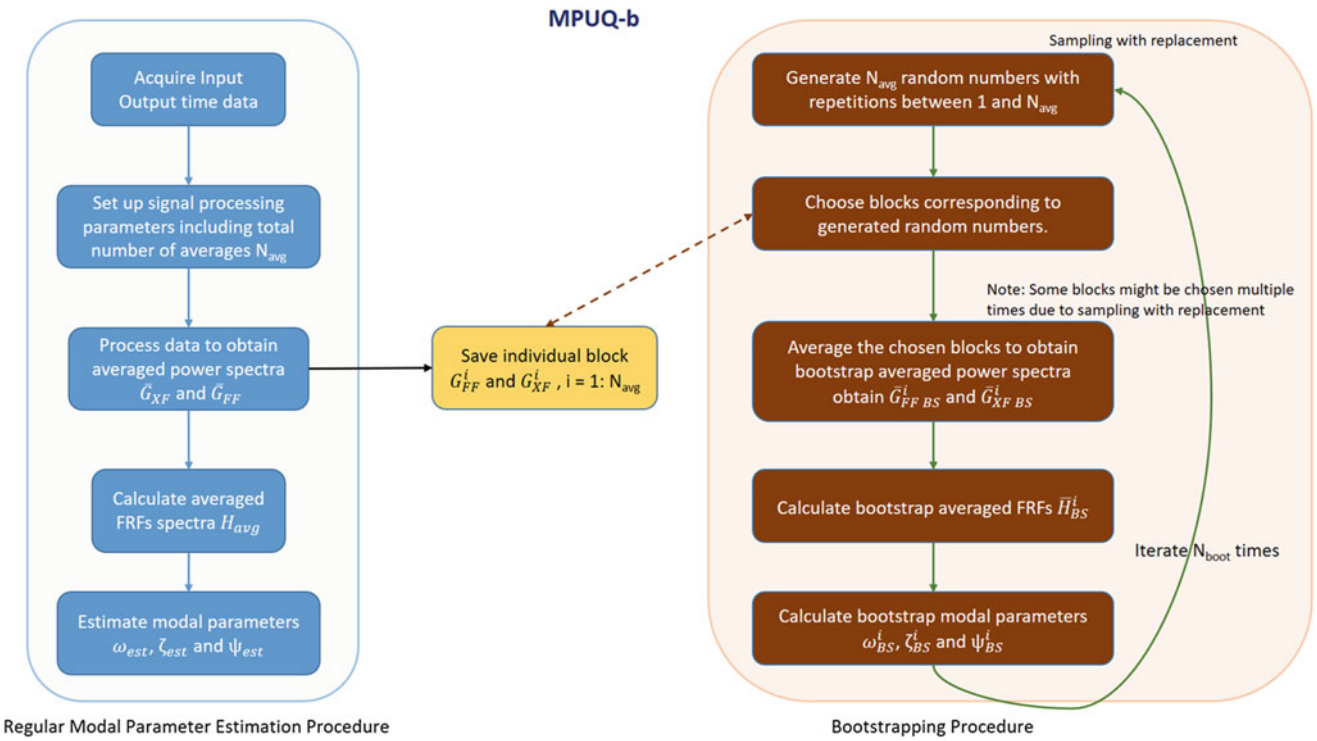


Fig. 23.2 Bootstrap based modal parameter uncertainty quantification (MPUQ-b) procedure

23.2.2 Features

MPUQ-b has several features that make it an attractive option for uncertainty quantification purposes. This section highlights these aspects of proposed methodology as well as its limitations. In the process, vital differences of MPUQ-b with respect to error propagation approach are also pointed out.

1. *Simple*: The most striking aspect of MPUQ-b is its simplicity, a property inherent to bootstrapping. Unlike error propagation approach, MPUQ-b relies more on computation than mathematical equations to answer statistical questions, such as those pertaining to uncertainty. In essence MPUQ-b simulates, by means of computer, what one would like to achieve in reality, that is conduct multiple experiments. That it achieves this without extensive modeling and mathematical analysis, makes it an appealing option that is easy to follow and can be applied in automatic manner.
2. *Algorithm independent*: Unlike error-propagation approach, MPUQ-b is not specific to any modal parameter estimation algorithm. One does not need to derive algorithm specific mathematical formulas for calculating the gradient (Eq. (23.2)) for MPUQ-b. The procedure remains the same as that described in Sect. 23.2.1 irrespective of the algorithm.
3. *Bias calculation*: Yet another strength of MPUQ-b is that it allows for calculation of bias, which is not possible with error-propagation approach. The bias of a statistic (an estimate of a parameter) is calculated as difference between the mean of the sampling distribution and the true parameter. However, due to practical constraints mentioned earlier, it is often neither possible to estimate the sampling distribution or the true value of the parameter. In these circumstances, bootstrap paves the way to estimate bias in the same way as it can be utilized to get an impression of the sampling distribution. Using bootstrapping, it is possible to check if the bootstrap distribution (analogous to sampling distribution) of the estimated modal parameter is centered at the modal parameter estimated from the original complete sample on which the bootstrapping has been performed. In mathematical terms, the bootstrap estimate of bias in estimated modal parameters is expressed as shown in Eqs. (23.3.a) and (23.3.b). It is easy to see that due to absence of data to create multiple samples in statistical sense, it is not possible to calculate bias through error propagation approach.

$$Bias_{\omega} = \omega_{est} - \overline{\omega}_{BS} \tag{23.3.a}$$

$$Bias_{\zeta} = \zeta_{est} - \overline{\zeta}_{BS} \tag{23.3.b}$$

4. *Enables the use of several tools of statistical inference*: The data resampling oriented inherent nature of bootstrapping that allows to create virtual experiments, enables MPUQ-b to use variety of statistical tools for inferential purposes and allows MPUQ-b to assess the uncertainty associated with estimated parameters in several ways. One can, for e.g. use a wide range of visualization tools like histograms, scatter plots, box plots, etc. in conjunction with various measures of accuracy such as variance, bias, sampling distribution, confidence intervals, normality checks etc. This is yet another advantage over error propagation approach, which is limited in comparison in this regard.
5. *Less stringent assumptions*: Unlike error propagation approach, which requires making significant assumptions and simplifications that may or may not be valid in all scenarios (for e.g. Eq. (23.2) is a result of linearization and t-ratios used for calculating confidence intervals are assumed to follow certain distribution), MPUQ-b does not involve any such assumptions. As noted earlier, performing bootstrapping is analogous to performing multiple experiments, except that the procedure is still based on data acquired in a single experiment. Thus, the most significant assumption in case of MPUQ-b, is that the acquired data is a good representation of the population, otherwise inference drawn on the basis of bootstrapping will suffer. However, this assumption applies to error propagation approach as well because it also operates on data collected in a single experiment assuming that the experiment is conducted in the best possible manner. In fact, it goes back to classic axiom in modal world, "Garbage In, Garbage Out" and MPUQ-b is not alienated from this.
6. *Computational aspects*: One advantage of the error propagation approach is that gradients can be calculated as a part of parameter estimation procedure and hence one can simultaneously estimate modal parameters and associated uncertainty. Computation of gradients does add to the overall parameter estimation procedure but this computational effort is not significant. MPUQ-b differs in this aspect as it performs signal processing and modal parameter estimation on each set of averaged bootstrapped FRFs ($\mathbf{H}^{i}_{BS}$). This process does take more time in comparison to error propagation but it can be optimized. Moreover, the numerous advantages that MPUQ-b provides justifies the additional computational effort.
7. *Data storage*: The resampling aspect of bootstrapping requires that in case of MPUQ-b individual blocks of power spectra have to be stored. This is a major difference from the usual practice of only saving the averaged FRF. This is however, not

an issue in current 'Big Data' age where processing power and storage capabilities of computers are more than equipped to handle such requirements related to computational effort and data storage.

23.3 Validation Studies on a Numerical System

23.3.1 Description of Numerical Experiment

To validate the proposed MPUQ-b methodology, a numerical experiment is designed to compare bootstrapping based statistical inference to that using Monte Carlo simulations. This section describes the numerical experiment in detail. The validation exercise is carried out by means of studies performed on a 5 DOF system shown in Fig. 23.3. Primary emphasis is on validating whether standard error (chosen measure of uncertainty) of frequency and damping estimates from two approaches compare well or not. It should be noted that MPUQ-b can also be used for mode shapes as well, however this paper focuses on modal frequency and damping for illustrating the proposed approach.

The physical parameters of the system are shown in Table 23.1.

This system can be solved numerically to obtain theoretical modal parameters and theoretical Frequency Response Functions (FRFs) can also be generated for a chosen frequency resolution. For the purpose of these studies, the theoretical FRFs are chosen to have a frequency resolution of 0.25 Hz, with a sampling rate of 256 Hz. Theoretical modal frequencies and damping values for this system are given in Table 23.2.

In order to validate the suggested bootstrapping methodology, a perturbation approach is adopted. The details of this approach are discussed next. A typical modal experiment involves acquiring several blocks of input force and output response data. Each block is processed using signal processing techniques like windowing, Fast Fourier transformation (FFT) etc. and in conjunction with averaging, one finally obtains averaged FRFs that are ultimately used for calculating modal parameters.

To simulate above scenario a perturbation approach is employed in which mass, damping and stiffness values (as listed above) are perturbed randomly. These perturbations are characterized by zero mean and a known standard deviation (σ). For each set of perturbed parameters, corresponding FRFs are generated. These FRFs are generated so that they have same block size, sampling frequency and hence frequency resolution as the theoretical FRFs. These perturbations are carried out as many number of times as the number of averages (N_{avg}) required. From bootstrapping perspective, this is all the data that is needed. Rest of the procedure remains same as that described in Sect. 23.2.1 (Steps 2-6). It should be noted that normally FRFs are estimated from collected input-output data by processing them using signal processing techniques such as

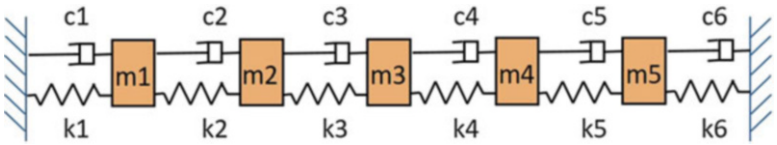


Fig. 23.3 5 DOF system

Table 23.1 Physical parameters of 5 DOF system

Mass (Kg)		Damping (Ns/m)		Stiffness (N/m) ($\times 10^6$)	
m1	250	c1	3000	k1	4
m2	350	c2	250	k2	5
m3	30	c3	200	k3	6
m4	450	c4	120	k4	6.5
m5	50	c5	70	k5	8
		c6	200	k6	7

Table 23.2 Modal frequencies and damping of the 5 DOF analytical system

Frequency (Hz)	Damping (% Critical)
12.5263	1.1486
22.083	1.0589
34.8635	2.172
88.5238	0.4872
104.7787	0.8473

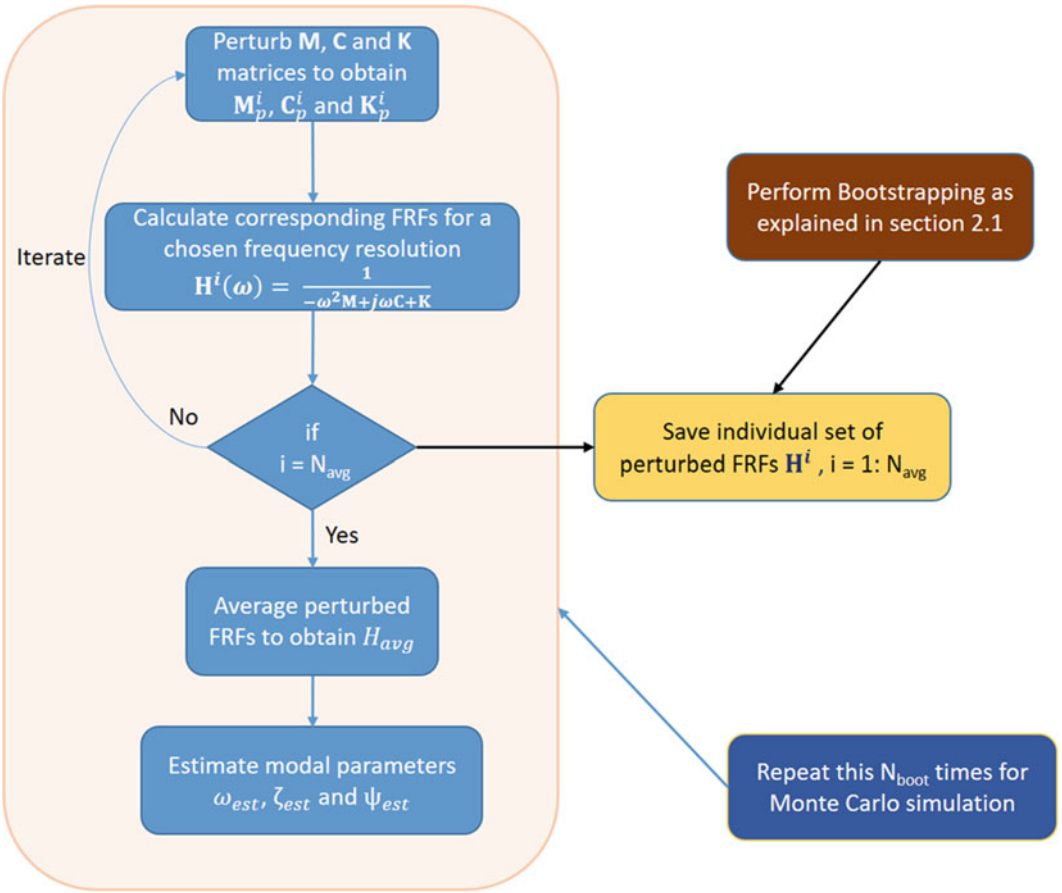


Fig. 23.4 Perturbation approach used in the study

windowing, averaging etc. In fact, it is typically power spectra that is averaged and FRFs are estimated as a ratio of averaged input-output cross power spectra and input auto power spectra (for H_1 estimator). Thus perturbation method adopted in this study differs from the usual practice but it is done to avoid effects arising from signal processing. More importantly, this has no consequence with regards to applicability of MPUQ-b or its validation.

Unlike bootstrapping, where only one sample (comprising of individual FRFs) is used to create N_{boot} bootstrap samples, for Monte Carlo simulations, the numerical experiment is repeated in its entirety N_{boot} times. More specifically, in each Monte Carlo simulation, the physical parameters of the system are perturbed and perturbation and other (sampling frequency, frequency resolution etc.) parameters are same as to those used in Bootstrapping procedure. Each simulation consists of N_{avg} number of perturbations, resulting in N_{avg} number of 'perturbed' FRFs. These N_{avg} FRFs are then averaged to obtain averaged FRFs. This simulation process is then repeated multiple times, each time with a different realization of perturbation (keeping perturbation and other parameters constant). For comparison purposes, the number of simulations is equal to number of bootstraps (N_{boot}). In this study the statistical inference is concentrated on modal frequency and damping estimates. Thus primary focus is on comparing mean, standard error and distribution of these quantities. Note that standard error is essentially standard deviation but calculated in context of the statistic of interest. Unlike standard deviation, which signifies the dispersion in the data, standard error signifies the dispersion in the sampling distribution of the statistic of interest. The perturbation approach explained above is shown schematically in Fig. 23.4.

23.3.2 Results and Discussions

The numerical experiments carried out in this study use following parameters:

Number of bootstrapping iterations (N_{boot})	500
Sampling Frequency (Hz)	256 Hz
Modal parameter estimation algorithm	Polyreference Time Domain (PTD) [4, 5], Rational Fraction Polynomial Z domain (RFP-z) [15]
Model order	2, 6 (Two different cases)
Number of averages (N_{avg})	100
Block size (Frequency Resolution)	1024
Perturbation Standard Deviation (σ)	0.01, 0.5 (Two different cases)

Table 23.3 Case 1, $\sigma = 0.01$, Model order = 2, Algorithm = PTD

Bootstrapping				Monte Carlo			
Modal frequency (Hz)		Modal damping (%)		Modal frequency (Hz)		Modal damping (%)	
Mean $\overline{\omega}_{BS}$	Std error $\sigma_{BS\omega}$	Mean $\overline{\zeta}_{BS}$	Std error $\sigma_{BS\zeta}$	Mean $\overline{\omega}_{MC}$	Std error $\sigma_{MC\omega}$	Mean $\overline{\zeta}_{MC}$	Std error $\sigma_{MC\zeta}$
12.5264	1.29E-05	1.148742	2.48E-06	12.52641	1.17E-05	1.148743	2.38E-06
22.08305	1.61E-05	1.059009	4.05E-06	22.08304	1.82E-05	1.059016	4.17E-06
34.86343	5.00E-05	2.17162	1.05E-05	34.86342	4.74E-05	2.171605	1.04E-05
88.52447	0.000848	0.487468	3.94E-05	88.52379	8.67E-04	0.487482	4.26E-05
104.7823	0.001838	0.844675	0.00015	104.7829	0.001576	0.844571	0.000112

Table 23.4 Case 2, $\sigma = 0.01$, Model order = 6, Algorithm = PTD

Bootstrapping				Monte Carlo			
Modal frequency (Hz)		Modal damping (%)		Modal frequency (Hz)		Modal damping (%)	
Mean $\overline{\omega}_{BS}$	Std error $\sigma_{BS\omega}$	Mean $\overline{\zeta}_{BS}$	Std error $\sigma_{BS\zeta}$	Mean $\overline{\omega}_{MC}$	Std error $\sigma_{MC\omega}$	Mean $\overline{\zeta}_{MC}$	Std error $\sigma_{MC\zeta}$
12.52635	6.35E-04	1.149143	0.0132949	12.52633	0.000657	1.149077	0.014491
22.08305	6.55E-05	1.058923	0.0003995	22.08304	8.09E-05	1.058926	0.000371
34.86354	5.01E-05	2.171991	1.104E-05	34.86353	4.74E-05	2.171976	1.13E-05
88.52465	0.002831	0.487501	0.0006183	88.5259	0.002932	0.487548	0.001767
104.778	0.001829	0.843186	0.0002183	104.7786	0.001571	0.843113	0.000175

Table 23.5 Case 3, $\sigma = 0.5$, Model order = 2, Algorithm = PTD

Bootstrapping				Monte Carlo			
Modal frequency (Hz)		Modal damping (%)		Modal frequency (Hz)		Modal damping (%)	
Mean $\overline{\omega}_{BS}$	Std error $\sigma_{BS\omega}$	Mean $\overline{\zeta}_{BS}$	Std error $\sigma_{BS\zeta}$	Mean $\overline{\omega}_{MC}$	Std error $\sigma_{MC\omega}$	Mean $\overline{\zeta}_{MC}$	Std error $\sigma_{MC\zeta}$
12.52727	0.000574	1.150573	0.000281	12.52634	0.000626	1.150672	0.000301
22.08283	0.000833	1.060528	0.000318	22.08295	0.000965	1.060752	0.000327
34.86487	0.002222	2.1762	7.55E-04	34.86302	0.0025	2.176861	9.31E-04
88.4925	0.100675	1.039849	0.079193	88.49422	0.105022	1.067984	0.091387
–	–	–	–	–	–	–	–

This section discusses and presents results of the numerical studies from four different perspectives, illustrating the possibilities of statistical inference in context of modal analysis using MPUQ-b.

23.3.2.1 Quantitative Analysis

It is stressed again that MPUQ-b and Monte Carlo simulations work with identical settings of parameter estimation algorithm. The mean and standard error for the 500 bootstrap and Monte Carlo estimates is shown in Tables 23.3, 23.4, 23.5, and 23.6 and it is observed that there is very good agreement between the results. Please note that four different cases are considered here by utilizing various combination of model order, perturbation standard deviation and chosen estimation algorithm. Since this is a 5 DOF system and the complete FRF matrix (5×5) is utilized, a model order of 2 should suffice for identification purposes. However, as the perturbation is increased the required model order to correctly model the system needs to be increased and hence a model order of 6 is also used.

In general, the results show very good agreement between the two approaches. It can be observed from these results that standard error in modal frequencies and damping ratios increases as perturbation (σ) increases. The criticality of choice of

Table 23.6 Case 4, $\sigma = 0.5$, Model order = 6, Algorithm = RFPz

Bootstrapping				Monte Carlo			
Modal frequency (Hz)		Modal damping (%)		Modal frequency (Hz)		Modal damping (%)	
Mean $\overline{\omega}_{BS}$	Std error $\sigma_{BS\omega}$	Mean $\overline{\zeta}_{BS}$	Std error $\sigma_{BS\zeta}$	Mean $\overline{\omega}_{MC}$	Std error $\sigma_{MC\omega}$	Mean $\overline{\zeta}_{MC}$	Std error $\sigma_{MC\zeta}$
12.53199	0.000608	1.090088	0.0019844	12.53108	0.000645	1.088029	0.002423
22.08049	0.000832	1.05613	0.0003073	22.08063	0.000966	1.056195	0.000286
34.86505	0.002229	2.1798	0.000518	34.86323	0.002498	2.179993	0.000576
88.51925	0.045466	0.672944	0.0173422	88.55324	0.042238	0.667636	0.017393
104.8044	0.075018	1.194212	0.0300125	104.6061	0.08807	1.230987	0.035009

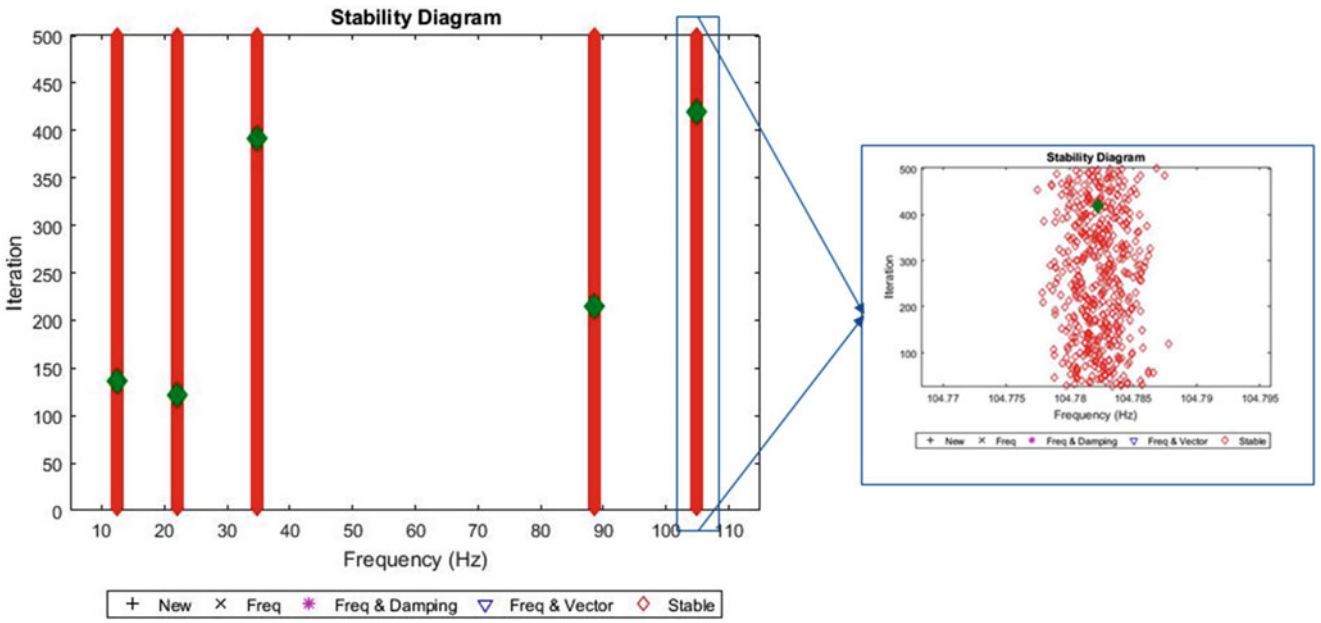


Fig. 23.5 Stabilization diagram for Case 1

correct model order is also highlighted in these results. For the σ = 0.01 case, when perturbation is low and hence model order of 2 is appropriate (as it is the correct order), the standard error in identified parameters is lower than the ones estimated using a higher model order of 6. This is however not true for the case when perturbation is more significant ($\sigma = 0.5$ case). In this case, model order of 2 is not even sufficient for identification of all 5 modes (Last mode is not identified). It is at model order of 6 that all modes are estimated and it is also observed that standard error of these modes is similar to those estimated using lower model order (Note that this comparison is only for the first four modes that are estimated in both case).

23.3.2.2 Qualitative Analysis

This section is devoted to showcasing the various visualization tools that MPUQ-b enables to gain more insights into estimated modal parameters. The stabilization diagram for case 1 ($\sigma = 0.01$, order = 2), shown in Fig. 23.5 is pretty clear, which is expected. As mentioned in Sect. 23.2, the stabilization diagram drawn here iterates over various set of averaged bootstrapped FRFs, unlike the conventional practice where it is the model order that is iterated. The figure also shows, by means of a zoomed version, the spread of one of the modal frequencies (around 104.78 Hz) from one set of bootstrapped averaged FRFs to another. In comparison, stabilization diagram for case 3 ($\sigma = 0.5$, order = 2), indicates poor estimation of 5th mode (Fig. 23.6).

Yet another way to look at these results is by means of a cluster diagram, which is essentially a frequency vs damping plot. This is akin to scatter plot, which can give a quick assessment about the correlation between the represented quantities. Figure 23.7 shows the cluster diagram for case 3 ($\sigma = 0.5$, model order 2). It clearly shows that 4th and 5th mode have more variability in comparison to other three modes. Zooming around 5th mode shows that damping variation is very high, indicating unreliable estimation of this mode.

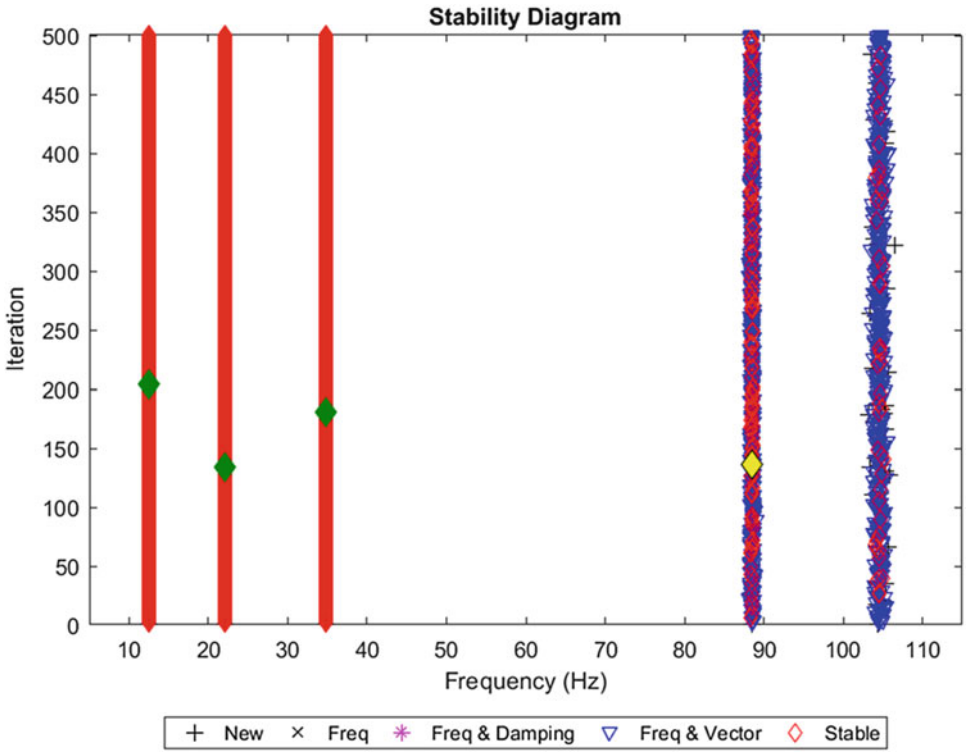


Fig. 23.6 Stabilization diagram for Case 3

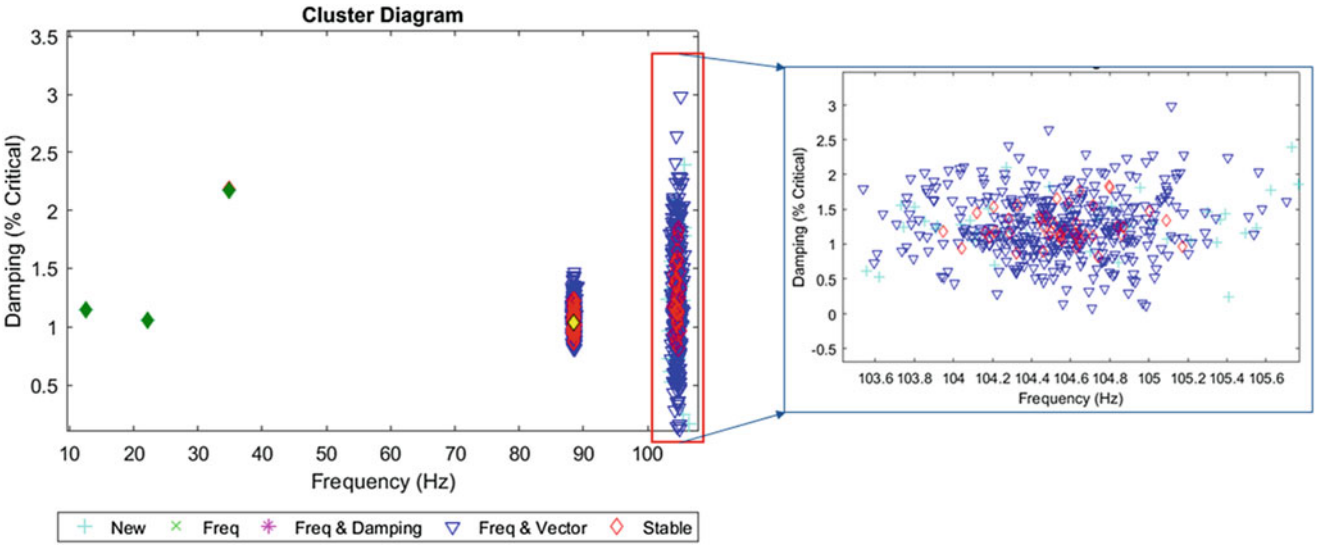


Fig. 23.7 Cluster diagram for Case 3

Such qualitative appreciation can also be gained by means of other plots like histogram (Fig. 23.8) and box plots (Fig. 23.9). The histogram plot (pertaining to case 4) in Fig. 23.8 also provides an idea about the distribution of the estimated quantities. Along with the distribution, mean of bootstrapped estimates ($\omega_{BS}^{i}, \zeta_{BS}^{i}$), parameters ($\omega_{est}, \zeta_{est}$) estimated from averaged FRFs (H_{avg}) and theoretical modal frequency and damping values ($\omega_{theo}, \zeta_{theo}$) are also indicated in the histogram. There is no significant difference between ($\overline{\omega}_{BS}, \overline{\zeta}_{BS}$) and ($\omega_{est}, \zeta_{est}$), indicating that the estimates are free of bias. However, there is considerable difference from theoretical parameters, indicating some modeling or signal processing related issues. In this case, since there is no signal processing involved, this might be an issue due to chosen frequency resolution, as it is damping estimates in most cases that differ from the theoretical values.

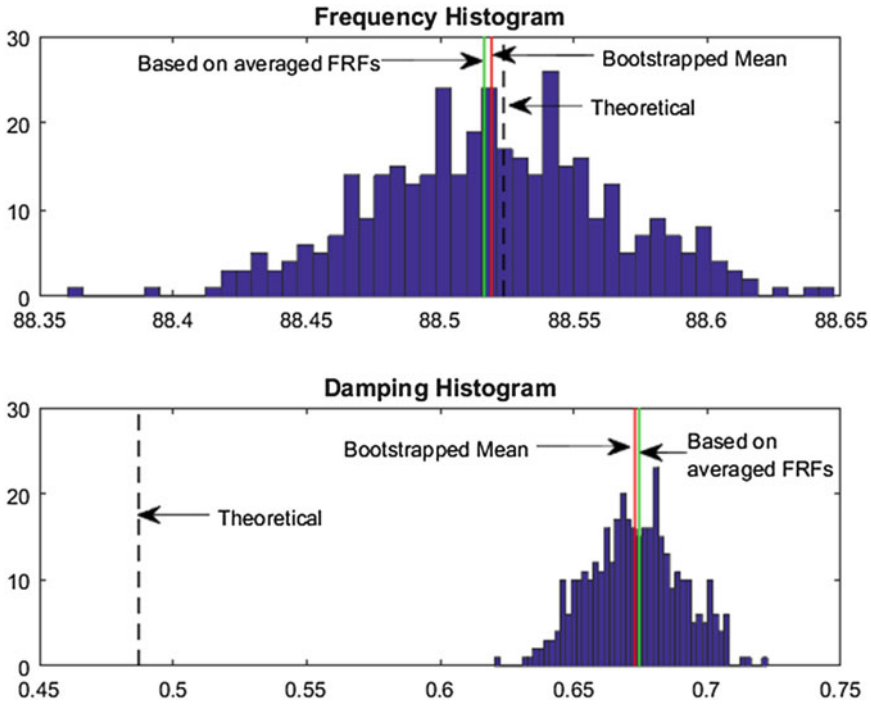


Fig. 23.8 Modal frequency and damping histograms for Case 4

Figure 23.9a shows the estimates of modal frequency and damping for case 1 in box plot representation [16]. For sake of better readability, box plot of damping estimate of 5th mode is shown separately in Fig. 23.9b. A box plot shows the data in terms of its quantiles and is a tool for quick visual summarization of data summaries like minimum, first quantile Q_1 (25%), median, third quantile Q_3 (75%), Interquartile range and maximum. The rectangle or the box shows the interquartile range (IQR), spanning first and third quartiles (IQR = Q_3–Q_1), along with a horizontal segment showing the median. In this case instead of showing the minimum and maximum of the data, the 1.5 IQR whiskers are shown and the points outside the whiskers provide an indication of outliers.

23.3.2.3 Normality Checks

Yet another aspect of MPUQ-b is its ability to verify the distribution of estimated parameters, specifically whether it is distributed normally or not. Unlike error propagation approach that is based on this assumption, and yet has no computational tool to verify the same, MPUQ-b provides several means to achieve this, both qualitatively and quantitatively. Qualitatively, one can look at histograms, such as the one shown in Fig. 23.8 to get some inclination about the distribution of the estimates. One can also use q–q plots [16] (Fig. 23.10) to get a visual impression of normality (or lack of it). A q–q plot is a tool to compare two distributions by plotting the quantiles and is often used to compare observed distribution with theoretical normal distribution. If the two distributions are similar than the points lie on a straight line. Figure 23.10, is the q–q plot for modal frequency around 12.52 Hz (Case 1) compared with theoretical normal quantiles. The linearity of this plot shows that the distribution is close to normal.

Instead of considering modal frequency as a univariate variable, one can look at frequency and damping estimates in unison by considering them as one bivariate variable (Fig. 23.11). In this case as well, the distribution is close to normal.

While histograms and q–q plots are good visualization tools for normality assessment, MPUQ-b also allows for more concrete measures in the form of normality tests that can be performed to gain a quantitative evaluation regarding normality of the estimates. For e.g. one can perform Henze-Zirkler's Multivariate Normality Test [17] on the estimated bootstrap modal frequency and damping estimates, considering them a bivariate variable. If multivariate data is distributed normally, the test statistic i.e. HZ value is approximately log-normally distributed. Such a test performed for first mode estimated in Case 1 results in following test statistic value (HZ value) at significance level 0.05 and the *p*-value, which is the significance value of this test statistic or in other words, significance of multivariate normality. Since *p*-value is much greater than 0.05, it can be concluded that the estimates of modal frequency and damping, treated as a bivariate variable, are distributed normally.

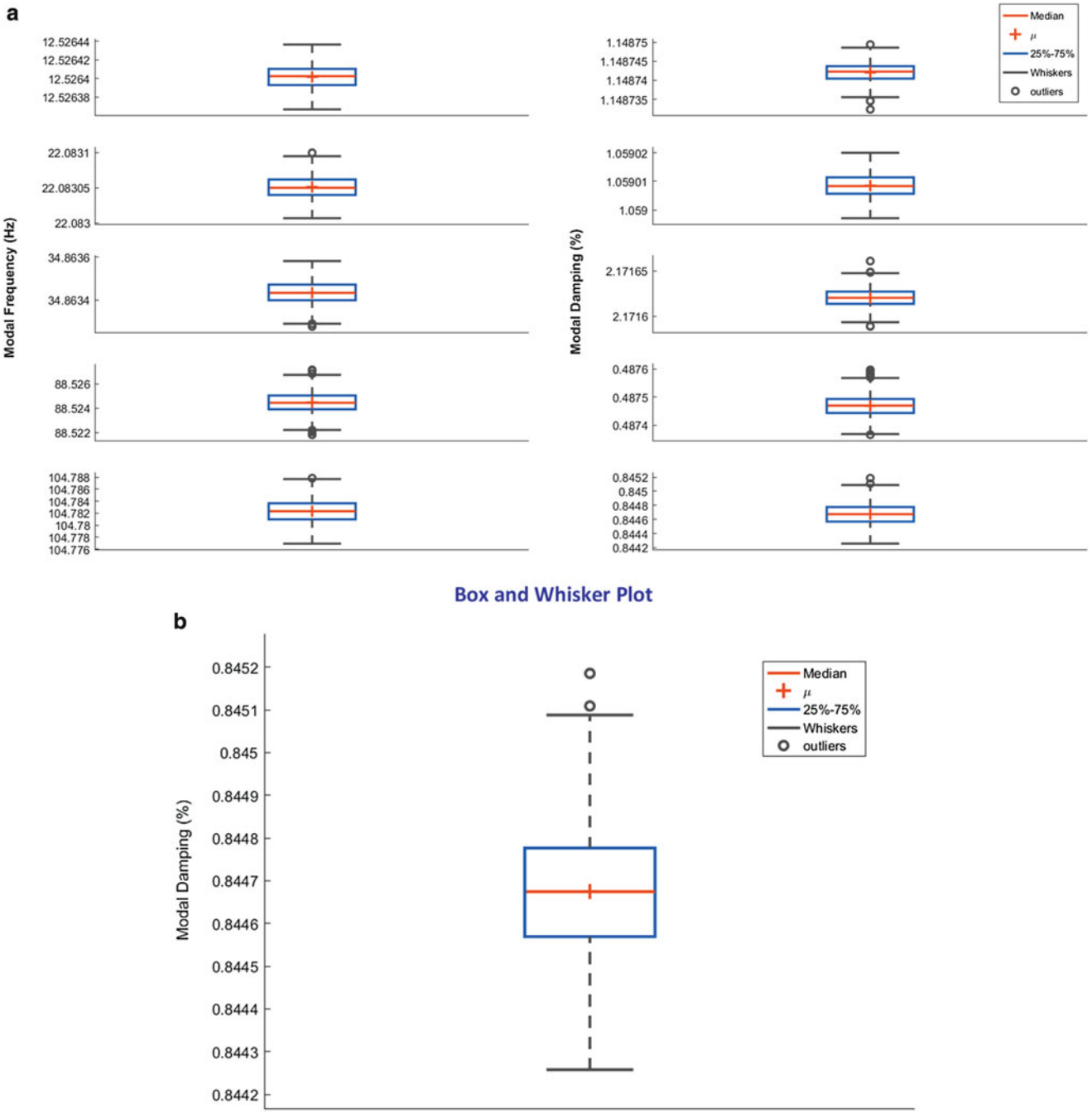


Fig. 23.9 (**a**) Box plot for estimated modal parameters (Case 1), (**b**) Zoom in for modal damping *of* 5th mode

HZ value: 0.6207425
p-value: 0.6380939

23.3.2.4 Comparison with Monte Carlo Simulations

Tables 23.3, 23.4, 23.5, and 23.6 provide estimate of standard error estimated using MPUQ-b and Monte Carlo simulations and show that there is good agreement between the two. It is well understood that Bootstrapping mimics the shape and spread of original sampling distribution but it is centered around the single sample statistic and hence its center might not coincide

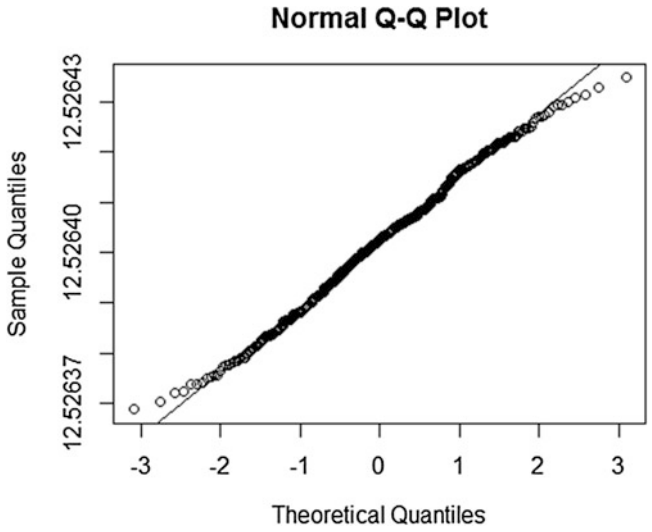


Fig. 23.10 Q–Q plot for one of the estimated modal frequencies (Case 1)

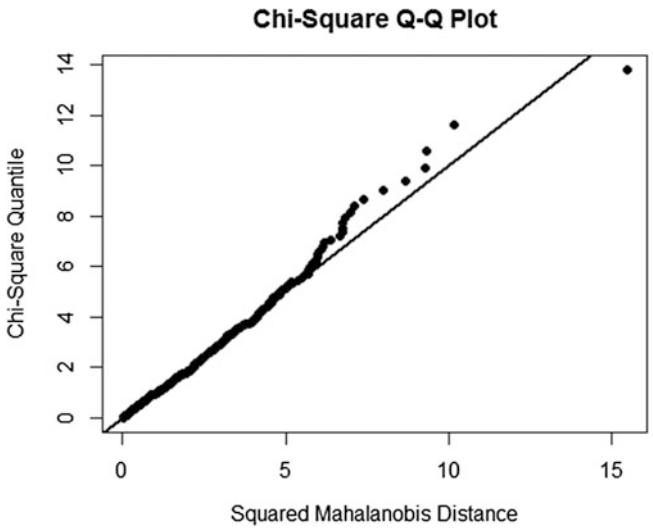


Fig. 23.11 Q–Q plot considering modal frequency and damping as bivariate variable (Case 1)

with the sampling distribution, which in absence of bias is most likely to be centered around population mean. This can be observed from Fig. 23.12 where histograms of frequency and damping estimates of 3nd mode (Case 1) from MPUQ-b and Monte Carlo are plotted on top of each other for illustration. The histograms have been normalized such that they plot an estimate of the probability distribution function of the estimates.

Expectedly, the histograms look very similar in shape and spread. However, they don't overlap each other completely, which is again on expected lines. This is more evident in damping histograms in comparison to frequency histograms. These histograms also look normal in terms of distribution and a normality check by means of Shapiro-Wilk's normality test [18] reveals that the estimates have univariate normal distributions at significance level 0.05. Note that the figures also show the normal distribution characterized by mean and standard deviation of estimated bootstrap parameters i.e. $\overline{\omega}_{BS}$, $\sigma_{BS\omega}$ and $\overline{\zeta}_{BS}$, $\sigma_{BS\zeta}$ on top of the histograms. This again provides a quick visual impression about how closely the histograms follow normal distribution.

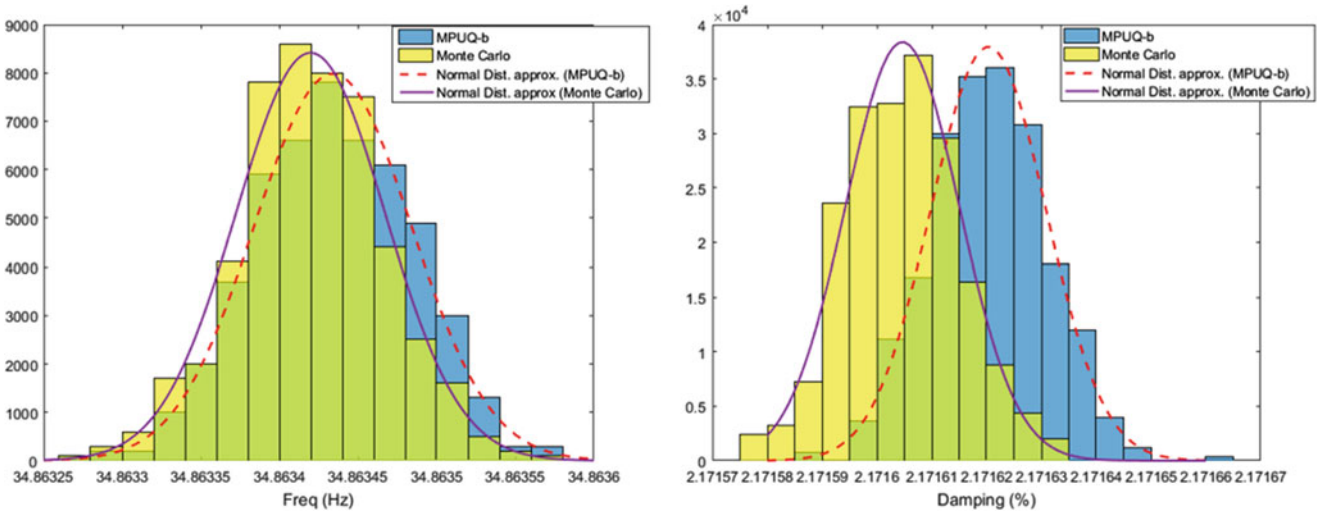


Fig. 23.12 Comparison of MPUQ-b and Monte Carlo histograms of 3rd mode (Case 1)

23.4 Conclusions

This paper showcased Bootstrapping based modal parameter uncertainty quantification (MPUQ-b) methodology. The methodology is described in step by step manner and its characteristics, advantages and limitations are discussed. Comparisons are also presented with respect to error propagation approach. It is shown how MPUQ-b combines the strength of bootstrapping technique with standard modal parameter estimation tools like stabilization and cluster diagrams to devise a simple, yet very effective and powerful technique for uncertainty quantification purposes. Biggest strength of this approach is that it does not require any rigorous mathematical treatment, making it readily applicable, irrespective of the modal parameter estimation algorithm or signal processing procedure. Yet another favorable property of MPUQ-b is that it does not involve any demanding assumptions. It does depend on the quality of the data collected but this is true for any method of uncertainty quantification.

Validation exercise, described in the paper, shows that MPUQ-b results are comparable to those from Monte Carlo simulations, which mimic the multiple experiment scenario, aimed at getting the sampling distribution. It is also highlighted that MPUQ-b, or for that matter any other uncertainty quantification technique, does not account for errors due to modelling and signal processing. Established good practices in this regard should be adhered to avoid or minimize them.

As illustrated, resampling characteristic of bootstrapping enables MPUQ-b to generate bootstrap samples to work with and also paves the way for the use of several data visualization tools, providing useful insight about the estimated modal parameters. With wide range of both qualitative and quantitative tools of statistical inference, MPUQ-b is a promising approach and encouraging results provide assurances towards its application to real-life scenarios.

References

1. Bendat, J.S., Piersol, A.G.: Engineering Applications of Correlation and Spectral Analysis, Second edn. John Wiley, New York (1993)
2. Heylen, W., Lammens, S., Sas, P.: Modal analysis theory and testing. PMA Katholieke Universteit, Leuven (1995)
3. Taylor, J.R.: An Introduction to Error Analysis, 2nd edn. University Science Books, Sausalito, CA (1997)
4. Vold, H., Rocklin, T.. The numerical implementation of a multi-input modal estimation algorithm for mini-computers. In: Proceedings of the 1st IMAC, Orlando, FL, November (1982).
5. Vold, H., Kundrat, J., Rocklin, T., Russell, R.: A multi-input modal estimation algorithm for mini-computers. SAE Trans. **91**(1), 815–821 (1982)
6. Reynders, E., Pintelon, R., De Roeck, G.: Uncertainty bounds on modal parameters obtained from stochastic subspace identification. Mech. Syst. Signal Process. **22**, 948–969 (2008)
7. Longman, R.W., and Juang, J., A variance based confidence criterion for ERA identified modal parameters. In: AAS PAPER 87-454, AAS/AIAA Astrodynamics Conference, MT, United States (1988).
8. Juang, J.N., Pappa, R.S.: An Eigensystem realization algorithm for modal parameter identification and model reduction. AIAA J. Guid. Control Dyn. **8**(4), 620–627 (1985)

9. Pintelon, R., Guillaume, P., Schoukens, J.: Uncertainty calculation in (operational) modal analysis. Mech. Syst. Signal Process. **21**, 2359–2373 (2007)
10. Brincker, R., Andersen, P.: Understanding Stochastic Subspace Identification, In: Proceedings of 24th International Modal Analysis Conference (IMAC), St. Louis (MO), USA, (2006).
11. Dohler, M., Lam, X.B., Mevel, L.: Efficient multi-order uncertainty computation for stochastic subspace identification. Mech. Syst. Signal Process. **38**, 346–366 (2013)
12. Dohler, M., Lam, X.B., Mevel, L.: Uncertainty quantification for modal parameters from stochastic subspace identification on multi-setup measurements. Mech. Syst. Signal Process. **36**, 562–581 (2013)
13. Chauhan, S., Ahmed, S.I., MPUQ-b: bootstrapping based modal parameter uncertainty quantification – fundamental principles, In: Proceedings of 35th International Modal Analysis Conference (IMAC), CA, USA, (2017).
14. Efron, B., Tibshirani, R.J.: An Introduction to the Bootstrap. Chapman and Hall, New York (1993)
15. Guillaume, P., Verboven, P., Vanlanduit, S., H. Van Der Auweraer, Peeters, B; A poly-reference implementation of the least-squares complex frequency-domain estimator, In: Proceedings of the 21st IMAC, Kissimmee (FL), USA, (2003).
16. Navidi, W.: Statistics for Engineers and Scientists, 3rd edn. McGraw Hill, New York (2010)
17. Henze, H., Zirkler, B.: A class of invariant consistent tests for multivariate normality. Commun. Stat. Theory Methods. **19**(10), 3595–3617 (1990)
18. Shapiro, S.S., Wilk, M.B.: An analysis of variance test for normality (complete samples). Biometrika. **52**(3–4), 591–611 (1965)

Chapter 24
Evaluation of Truck-Induced Vibrations for a Multi-Beam Highway Bridge

Kirk A. Grimmelsman and John B. Prader

Abstract Vibrations induced by trucks crossing bridges can add uncertainty to the identification of the structure's dynamic characteristics, and in some cases, can negatively impact the serviceability performance of the structure. Furthermore, the nature of the vibrations induced by trucks and other vehicles crossing highway bridges is not well-understood in practice and has primarily been evaluated using analytical models and limited experimental studies. This paper presents the results of an experimental study on the characteristics of the vibrations induced by trucks crossing a multi-beam highway bridge. The objective of the bridge monitoring program was to assess the prevalence of overweight truck crossings due to rapid growth of fatigue cracks that had been observed on the steel beams. The bridge monitoring system was triggered to record strain measurements along with camera images of the traffic on the bridge during each truck crossing event. The strain response of the bridge beams were recorded for several thousand truck crossings along with snapshot images of the trucks over a six week period. Several thousand truck crossing events were recorded over the entire monitoring program. The strain measurements and images recorded for over 400 different truck crossing events recorded over a 24 hour period were evaluated in the context of the resulting bridge vibrations and truck characteristics including gross vehicle weight and the spatial lane position. This analysis and results are presented and discussed in this paper.

Keywords Bridges • Trucks • Strain • Vibrations • Fatigue

24.1 Introduction

Vehicle-induced dynamic effects are considered in current design practice for highway bridges through the use of a dynamic load allowance (DLA), which is a factor that increases the static load effects on the bridge components due to the design truck load/design tandem, by different percentages for strength, service and fatigue limit states. For the design of most of the bridge components, the DLA ranges from 15% for the fatigue and fracture limit state to 33% in all other limit states [1]. From a design perspective, the DLA is a simple and efficient way to bound and account for complex dynamic load effects of vehicles that may result from the discontinuities in the pavement surface on or approaching the bridge and due to dynamic interactions between the vehicle and the overall structure. The characterization and evaluation of the actual bridge vibration responses due to the random vehicular live loads that cross these structures is significantly more challenging, and is an important consideration for the dynamic characterization of these structures by vibration monitoring for evaluating their health and condition, and for evaluating many of the performance problems that arise for bridge components during their actual service lives. The measured vibration response of most common highway bridges due to truck load crossing consists of both forced and free vibration stages which are affected by the mass, stiffness, velocity and other geometrical characteristics of the truck load which define its dynamic characteristics, its time-varying spatial position along the span as it crosses the bridge and the position of the lane it occupies relative to the underlying superstructure, the condition of the roadway surface, and the mass, stiffness and damping characteristics of the overall structure. The influence of these various parameters on the vibration responses of highway bridges has been studied to various degrees by many researchers, either analytically [2, 3, 4] or experimentally [5, 6]. Analytical studies generally must incorporate idealizations related to the loads and responses. The findings from such studies are more challenging to consider in the health and performance evaluation of in-service bridges given the random nature of truck loads and considering the myriad of other load- and structural-related uncertainties that inevitably cloud the evaluation process. The direct measurement and characterization of bridge vibration

K.A. Grimmelsman (✉) • J.B. Prader
Intelligent Infrastructure Systems, 3001 Market Street, Suite 200, Philadelphia, PA 19104, USA
e-mail: kgrimmelsman@iisengineering.com

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_24

responses due to actual truck loads can be more directly beneficial for informing the evaluation of specific bridge health or performance problems, but such studies can be complex to execute and interpret and thus, are somewhat less common in the literature.

The original motivation for the experimental study presented in this paper was quite different from the final direction in which measurement results ultimately led the analysis. A bridge owner was alerted to fatigue cracks in several steel multi-beam bridges located on the same highway route through a recent visual inspection of the structures. The fatigue cracks were observed to have developed and propagated rapidly considering that most of the cracks had not been identified in the previous biennial visual inspections of these structures. The engineers who identified the fatigue cracks suspected that large numbers of overweight trucks were crossing these bridges and were responsible for responsible for the rapid fatigue the crack propagation and growth. An experimental program was designed and executed by the writers at the request of the bridge owner to evaluate the prevalence of overweight trucks on these structures. A representative bridge from the group of structures was selected for instrumentation with strain gages and monitoring to assess the prevalence of overweight trucks crossing the structures. Analysis of the resulting strain measurements indicated that quantity of overweight trucks was not particularly excessive and that vibration related stress cycles associated with truck crossing events were probably a more significant influence on the fatigue crack growth. The nature of the vibration-related stress response was observed to be variable for different truck crossing events, and the measurement data were further analyzed to study and evaluate how the vibration-related stresses were influenced by different truck characteristics. Specifically, the writers wanted to evaluate how the weight of the truck and the lane used in crossing the bridge influenced the vibration-related stress cycles measured for the steel beams.

24.2 Test Structure and Experimental Program

The bridge selected for monitoring was generally representative of the group of bridges located on a particular stretch of highway that had similar ages and design attributes, and that were all observed to have developed fatigue cracks at a detail that was common to all of these bridges. The test bridge was a steel multi-beam bridge with a reinforced concrete bridge deck on rolled steel beams with cover plates welded to the bottom flange. The bridge crossing was located on a state highway and carried two lanes of single-direction traffic that included a relatively large number of trucks. The overall bridge crossing consisted of two simple girder spans with span lengths of 27.4 m measured from center to center of the bearings. The transverse width of reinforced concrete deck was 12.2 m and was striped for two 3.7 m wide traffic lanes, a 3.7 m wide shoulder adjacent to the slow speed lane and a 1.2 m wide shoulder adjacent to the high speed lane. The reinforced concrete deck was supported on six W36x302 rolled steel beams with 38 mm thick cover plates welded to the underside of the bottom flange. The arrangement of the steel framing plan was such that girder line 2 was located nearly under the center of the high speed (left) traffic lane, and girder line 4 was located nearly under the center of the slow speed (right) traffic lane, and girder line 3 was located approximately under the boundary between the high speed and slow speed traffic lanes. The remaining three girder lines were located under the shoulders on the right and left sides of the two traffic lanes.

A single span from the two-span bridge was instrumented with strain gages and monitored for live load strains. A total of 17 strain gages were installed on the top of the bottom flanges of the 6 girders at 7 different cross section locations, denoted as Sections A through E in Fig. 24.1. Sections C and F are each located approximately 1 m away from the midspan location, and Sections B and F are each located approximately 2 m away from the midspan location. The strain gages used for the testing were weldable vibrating wire strain transducers from Geokon and these were measured with dynamic vibrating wire data acquisition hardware from Campbell Scientific. A low-cost IP camera was also installed on the bridge railing near the middle of the adjacent bridge span to capture snapshot images of the truck traffic crossing the instrumented bridge span during the monitoring program. The strain gages were sampled continuously at 20 Hz and measurements were recorded based on a trigger threshold established by controlled truck load testing of the instrumented span to record the strains for any truck crossing event, whether the truck was loaded or empty. The pre- and post-trigger durations for the triggered strain measurements were arbitrarily set at 30 seconds each.

A controlled truck load test was performed on the instrumented bridge span using a tandem axle dump truck with known axle weights and distances. The gross vehicle weight (GVW) of the test truck was 302 kN, and the truck was positioned at known locations in both the slow speed and high speed lanes to calibrate the measured girders strains to the known GVW and enable the screening of GVW of truck traffic crossing the bridge during the continuous monitoring. The trigger threshold value used for identifying and recording strains associated with truck crossing events and the snapshot images of the traffic were also established based on the strain analysis results from the truck load test.

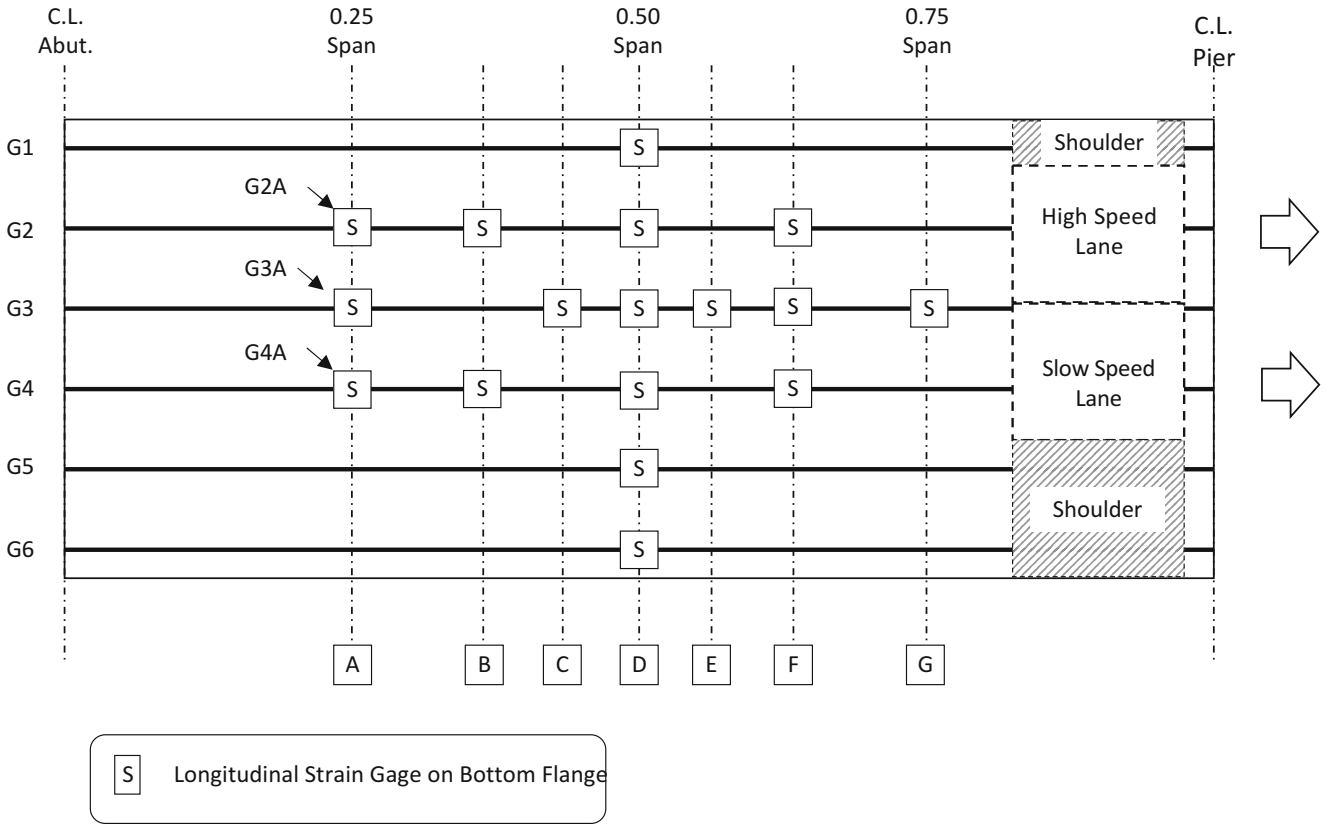


Fig. 24.1 Strain gage locations on test bridge span

24.3 Data Analysis and Results

The primary objective of the bridge monitoring program was to evaluate the prevalence of overweight truck crossing events. The recorded strains were analyzed in conjunction with the recorded traffic snapshot images and truck load test results to accomplish this objective. A review of the recorded traffic images indicated that the triggered recording scheme employed for the monitoring program was very effective as all of the images consisted of trucks of various types and configurations and did not include any images of regular passenger vehicles. Each truck image was reviewed and aligned with the strain recordings associated with the event for analysis and validation of the measurements. The low-cost IP camera used did not include an infrared illuminator, so the truck images recorded between sunset and sunrise were not as useful for analysis. Fortunately, this timeframe also generally coincided with lower instances of truck crossing events. The strain gage located on Girder 4 at midspan (Section D) malfunctioned after its installation and was excluded from all analysis.

The girder strain measurements recorded for each truck crossing event were analyzed to estimate the GVW of the truck using the measured strains for the known truck load and positions recorded during the load test. Lane presence was evaluated by comparing strains recorded from girder lines 2 and 4 and validated with the traffic images. The peak strains recorded for each truck event were divided by the peak strains recorded for truck in each lane during the load test to estimate the GVW. The estimated GVW values were determined to be reasonable approximations to the actual GVW and were thus suitable for truck load screening purposes, but not for weight enforcement. The recorded strains indicated that the typical truck crossing events included both a forced and free vibration components and that multiple, and not insignificant, stress cycles were occurring during the free vibration stage of the truck crossing events. This observation indicated that these stress cycles could be a more significant contributor to the fatigue crack growth than the number of overweight truck crossings. The characteristics of the vibration response observed from the strain recordings also motivated the writers to take a deeper look at the measurement data to evaluate if certain truck characteristics would result lead to larger and longer free vibration responses and corresponding stress cycles.

The strain data analyzed in this paper was recorded over a single 24 h period during the overall monitoring program. The strains recorded for each truck event were assessed in conjunction with the snapshot images to extract specific truck events

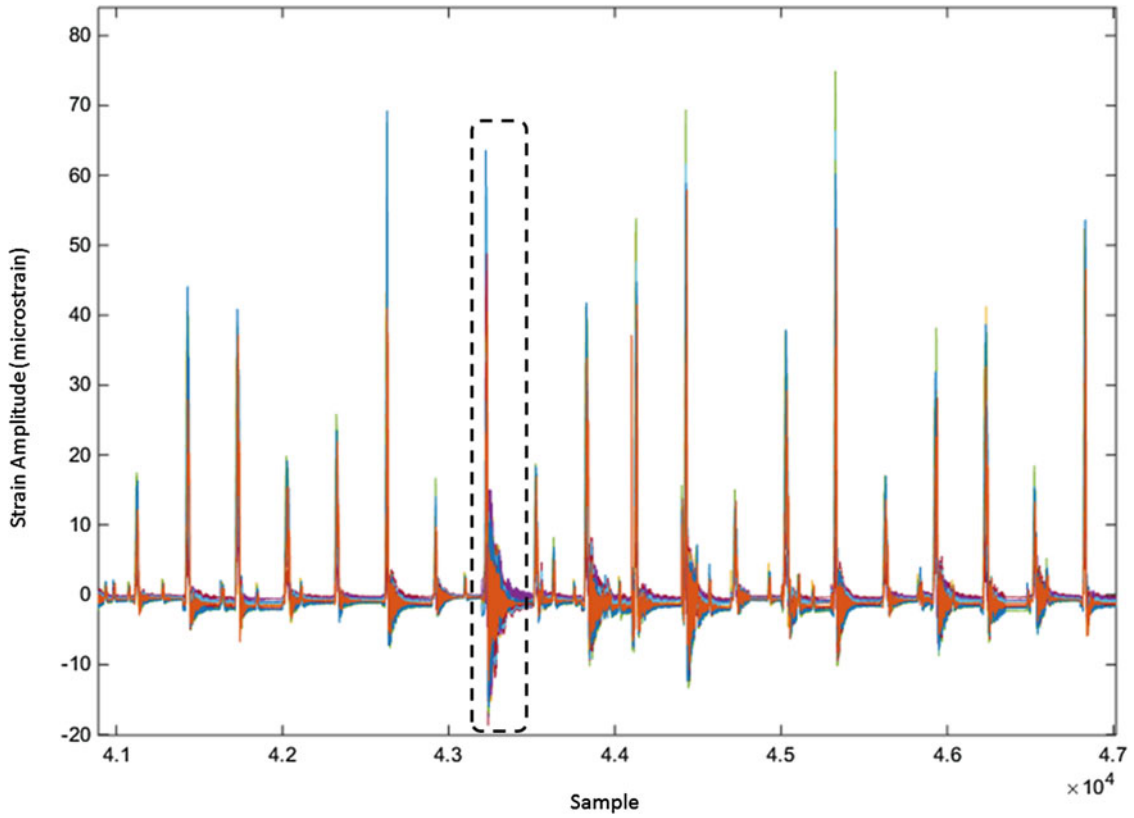


Fig. 24.2 Subset of strain records for truck crossing events

most suitable for evaluating the bridge's vibration response. A specific goal was to identify isolated truck crossing events in which a single truck of any weight crossed the bridge with adequate time for characterizing the free vibration responses before another truck crossed. The nature of the truck traffic crossing the bridge was quite random, and truck crossing events that occurred within 14 seconds or less of the most previous truck crossing event were excluded from this analysis. This analysis resulted in the identification of 419 truck events that could be characterized as being isolated from each other. Each truck event strain record consisted of 300 samples for each of the 16 bottom flange strain gages located on the bridge. It should be noted that the triggered recording scheme employed could not exclude passenger vehicle crossing events that may have occurred within the same time window of the extracted truck crossing events, but the effects of such lighter weight vehicles on the global vibration response of the bridge induced during a truck crossing event was evaluated to represent a relatively minor source of uncertainty.

Figure 24.2 shows the time domain strain responses for a subset of about 20 of the total 419 truck crossing events extracted from the measurement data for analysis. A magnified view of the strains for the single truck event circled in this figure is shown in Fig. 24.3 and the forced and free vibration stages of the strain response are clearly visible. The strain measurements associated with the 419 events were further analyzed as described in the following.

24.3.1 RMS Analysis

The GVW of the truck associated with each of the 419 events was estimated using the procedure described above. The lane occupation for each truck event was first established by comparing the peak strain from the same cross section location for Girder 2 and Girder 4. These results were validated from the image snapshots of each event and proved to be very reliable. The peak strains recorded from the controlled truck load test for lane associated with the truck event were then used with the known GVW of the test truck to estimate the GVW for the truck. The GVW values calculated using the event peak strains for gages G2F, G3F and G4F were then averaged together to obtain a final estimate of the GVW for the truck.

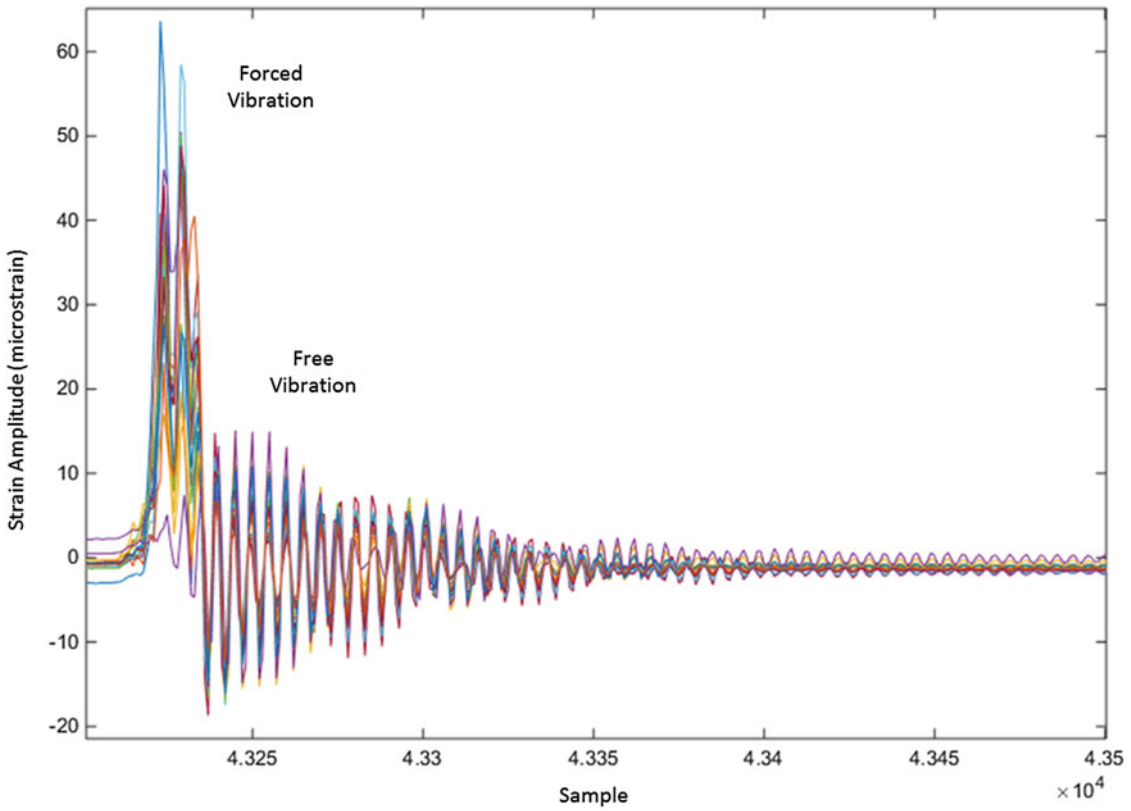


Fig. 24.3 Forced and free vibration related strains for a single truck crossing event

An attempt was made to estimate the truck velocity using the peak strains from gages at different cross section locations on the same girder line. The time elapsed between the peaks in two strain gages on the same girder was used with the known distance between the cross section locations to compute the truck velocity. This was attempted for multiple strain gage combinations, but the results proved inconsistent. These results indicated that the 20 Hz sampling rate was not fast enough to obtain a reliable estimate of the truck velocity for each event.

The influence of truck weight on the vibration response of the bridge was evaluated through analysis of the root mean squared (RMS) values computed from the strain records from the 16 gages for each event. The RMS strain amplitude was determined to be a straightforward means for comparing the nature of the bridge vibrations induced by each truck crossing event. The RMS strain amplitudes computed for each strain gage for a given event were then summed together to obtain a single index, total RMS, that would reflect the global vibration of the bridge. Larger and/or longer vibration responses would be associated with the larger total RMS strain amplitudes computed for each truck event. The RMS analysis was performed for each truck event for both the complete measurement record (forced and free vibration stages) and for the portion of the record associated with free vibration response. The total RMS strain amplitudes for each event are compared against the GVW in Figs. 24.4 and 24.5.

The total RMS strain amplitudes for the forced vibration portions of each truck event are nearly linear with increasing GVW. This indicates that the strain response is dominated by the forced vibration component and is as would be expected. There is significantly less scatter in the relationship between total RMS amplitude and truck GVW. When only the free vibration portion of the strain records are considered, the relationship between the liveliness of the bridge vibration response and the truck GVW becomes less steep and the data have a much larger scatter for larger GVW values than it does for smaller GVW values. There is a marked increase in the scatter of this relationship once the GVW exceeds 155 kN (35 kips). The free vibration response appears to be less correlated to GVW as the truck weight approaches or exceeds the legal load of 356 kN (80 kips). Although truck velocities could not be computed from the strain data, it is reasonable that truck velocities decrease with increasing GVW and the vibrations induced by these trucks tend to reach a more constant value. It can be observed from these results that the velocity of the truck will influence the resulting vibration response of the bridge to a larger extent than will the GVW.

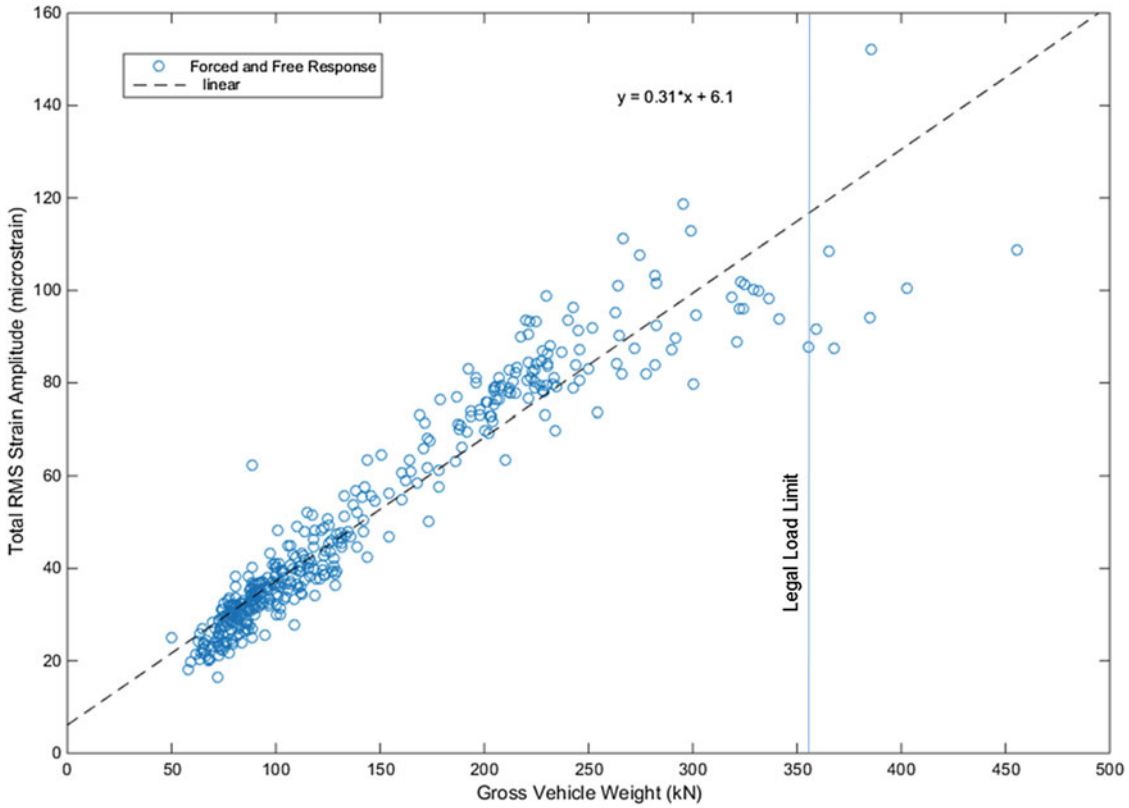


Fig. 24.4 Total RMS strain amplitude versus GVW from total strain record (forced and free vibration)

24.4 Conclusions and Future Work

A bridge monitoring program was designed and executed for a typical steel multi-beam highway bridge that had been identified to have fatigue cracks that were rapidly developing and growing. The original goal of the program was to assess the prevalence of overweight trucks crossing the bridge. The bridge was instrumented with strain gages and the strains and snapshot images of trucks crossing events were recorded on a triggered basis. The strain results were used in conjunction with those recorded from a controlled truck test to estimate GVW values for trucks crossing the bridge. The analysis did not indicate that there were large numbers of overweight truck crossings, but the strain records did indicated multiple and often large stress cycles were occurring with the truck crossings as a result of the free vibrations of the bridge. These stress cycles are very likely contributing to the fatigue crack growth, particularly since only a single stress cycle would be expected to be associated with each truck crossing event for the simple span structure.

The writers were thus motivated to evaluate the influence of truck weight on the overall vibration response of the bridge. RMS analysis of the strains records associated with each truck crossing event indicated that the total strain response is dominated by the quasi-static behavior while the truck physically on the structure. When only the free vibration response of the bridge is evaluated from the RMS strain values, the relationship between total vibration response of the bridge and the weight of the truck is better correlated for smaller GVW values. As the GVW of the crossing trucks increases, the relationship between the GVW and the global free vibrations is less consistent and appears to follow a different relationship. It is generally understood that with increasing truck GVW, truck velocities will decrease (due to lower posted speed limits for trucks or due to performance limitations of the truck itself when it is more heavily loaded). The 20 Hz sampling rate employed in this monitoring program was adequate for estimating the GVW for the random trucks crossing the bridge, but it was too slow to obtain reliable estimates of the truck speeds.

Although 419 truck crossing events were analyzed in this paper, this represents a very small sample of the total number of truck crossing events recorded during the monitoring program. Analysis of the recorded strain measurements is ongoing and spectral analysis of the strains will be performed to further evaluate the influence of truck characteristics on the vibration response of the bridge. Stress cycle analysis, such as rainflow analysis, will also be performed to better characterize the contributions of the girder strains associated with the free vibration response to the fatigue fracture behavior of the bridge.

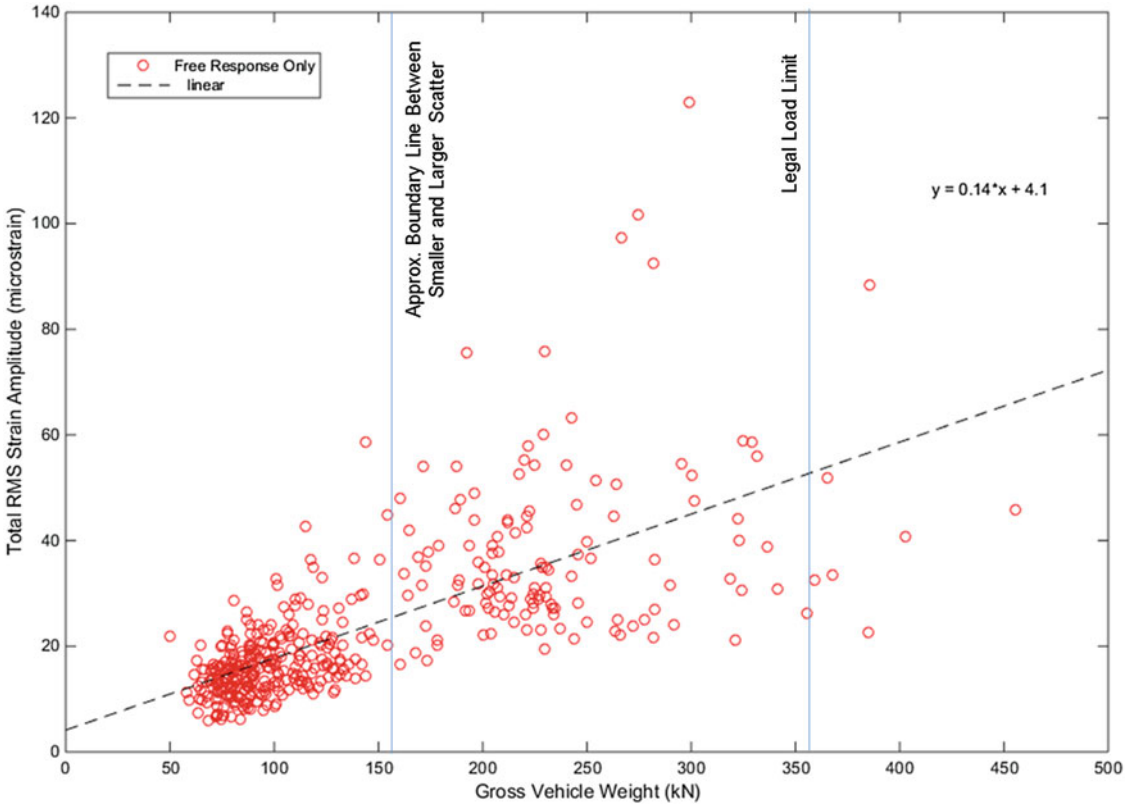


Fig. 24.5 Total RMS strain amplitude versus GVW from only the free vibration response

References

1. AASHTO LRFD: Bridge Design Specifications, 6th edn. American Association of State Highway and Transportation Officials, Washington, D.C (2012)
2. Green, M.F., Cebon, D.: Dynamic interaction between heavy vehicles and highway bridges. Comput. Struct. **62**(2), 253–264 (1997)
3. O'Brien, E.J., Rattigan, P., Gonazalez, A., Dowling, J., Zindaric, A.: Characteristic dynamic traffic load effects in bridges. Eng. Struct. **31**, 1607–1612 (2009)
4. Brady, S.P., O'Brien, E.J., Znidaric, A.: Effect of vehicle velocity on dynamic amplification of a vehicle crossing a simply supported bridge. J. Bridg. Eng. **11**(2), 241–249 (2006)
5. Cantineni, R.: Dynamic behaviour of highway bridges under the passage of heavy vehicles. In: Swiss Federal Laboratories for Materials Testing and Research, EMPA, Report No. 220, (1992).
6. Kim, J., Lynch, J.P.: Experimental analysis of vehicle-bridge interaction using a wireless monitoring system and a two-stage system identification technique. Mech. Syst. Signal Process. **28**, 3–19 (2012)

Chapter 25
Innovations and Info-Gaps: An Overview

Yakov Ben-Haim and Scott Cogan

Abstract Info-gap theory is a method for modeling and managing severe uncertainty and for prioritizing the options facing a decision maker. We consider uncertainty both in parameters and in the shape of functions. The analysis reveals intuitively meaningful concepts: the trade off between robustness and requirements, cost of robustness, zero robustness of predicted outcomes, innovation dilemmas, preference reversals resulting from uncertainty, and opportuneness from uncertainty. These insights underlie decision making under severe uncertainty. We consider an example of modeling a mechanical system.

Keywords Decision making • Uncertainty • Info-gap theory • Robustness • Opportuneness • Innovation dilemma

25.1 Info-Gap Theory: A First Look

Info-gap theory is a method for prioritizing options and making choices and decisions under severe uncertainty [1, 2]. The options might be operational alternatives (design a system, choose a budget, decide to launch or not, etc.) or more abstract decisions (choose a model structure, make a forecast, design a project flow chart, etc.). Decisions are based on data, scientific theories, empirical relations, knowledge and contextual understanding, all of which are one's *models*, and these models often recognize and quantify uncertainty.

Info-gap theory has been applied to decision problems in many fields, including various areas of engineering [3–5] biological conservation [6] economics [2, 7] medicine [8] homeland security [9] public policy [10] and more (info-gap.com). Info-gap robust-satisficing has been discussed non-technically elsewhere [11–13].

Uncertainty is often modeled with probability distributions. If the probability distributions are correct and comprehensive then one can exploit the models exhaustively to reliably achieve stochastically optimal outcomes, and one doesn't need info-gap theory. However, if one's models will be much better next year when new knowledge has become available (but you must decide now), or if processes are changing in poorly known ways, or if important factors will be determined beyond your knowledge or control, then one faces severe uncertainties and info-gap theory might help. This section presents a concise intuitive discussion of two basic ideas of info-gap theory: satisficing and robustness. Simple examples are presented in the next section.

Frank Knight [14] distinguished between what he called "risk" (for which probability distributions are known) and "true uncertainty" (for which probability distributions are not known). Knightian ("true") uncertainty reflects ignorance of many things, including underlying processes, functional relationships, strategies or intentions of relevant actors, future events, inventions, discoveries, surprises and so on. Info-gap models of uncertainty provide a non-probabilistic quantification of Knightian uncertainty. An info-gap is the disparity between what you *do know* and what you *need to know* in order to make a reliable or responsible decision. An info-gap is not ignorance per se, but rather those aspects of one's Knightian uncertainty that bear on a pending decision and the quality of its outcome.

An info-gap model of uncertainty is particularly suitable for representing uncertainty in the shape of a function. For instance, one might have an estimate of the stress-strain curve, or of a probability density function, but the shape of the function (e.g. the shape of the elastic-plastic transition or the tails of the pdf) may be highly uncertain. Info-gap models are also widely used to represent uncertainty in parameters or vectors or sometimes uncertainty in sets of such entities.

Y. Ben-Haim
Yitzhak Moda'i Chair in Technology and Economics, Technion—Israel Institute of Technology, Haifa, Israel

S. Cogan (✉)
FEMSTO-ST, Département de Mécanique Appliquée, Université de Franche-Comté, 24 Chemin de l'Epitaphe, Besançon 25000, France
e-mail: scott.cogan@univ-fcomte.fr

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_25

Decision makers often try to optimize the outcome of their decisions. That is usually approached by using one's best models to predict the outcomes of the various options, and then choosing the option whose predicted outcome is best. The aspiration for excellence is usually to be commended, while recognizing that outcome-optimization may be costly, or one may not really need the best possible outcome. Schwartz [15] discusses the irrelevance of optimal outcomes in many situations.

Outcome-optimization—the process of using one's models to choose the decision whose predicted outcome is best—works fine when the models are pretty good, because exhaustively exploiting good models will usually lead to good outcomes.

However, when one faces major info-gaps one's models contain major errors, and exhaustively exploiting the models can be unrealistic, unreliable, and can lead to undesired outcomes [11]. Under severe uncertainty it is better to ask: what outcomes are critical and must be achieved? This is the idea of *satisficing* introduced by Herbert Simon [16]: achieving a satisfactory or acceptable, but not necessarily optimal, outcome.

Planners, designers and decision makers in all fields have used the language of optimization (the lightest, the strongest, the fastest, ...) for ages. In practice, however, satisficing is very wide spread though not always recognized as such. Engineers satisfy design specifications (light enough, strong enough, fast enough, ...). Stock brokers, business people, and investors or all sorts don't really need to maximize profits; they only need to beat the competition, or improve on last year, or meet the customer's demands. Beating the competition means satisficing a goal.

Once the decision maker identifies the critical goals or outcomes that must be achieved, the next step is to make a decision or choose an action that will achieve those goals despite current ignorance or unknown future surprises. A decision has high *robustness* if it satisfices the performance requirements over a wide range of unanticipated contingencies. Conversely, a decision has low robustness if even small errors in our knowledge can prevent achievement of the critical goals. The *robust-satisficing* decision maker prioritizes the alternatives in terms of their robustness against uncertainty for achieving the critical goals.

The decision methodology of robust-satisficing is motivated by the pernicious potential of the unknown. However, uncertainty can be propitious, and info-gap theory offers a method for prioritizing one's options with respect to the potential for favorable surprises. The idea of "windfalling" replaces the concept of satisficing. The opportune windfalling decision maker prioritizes the alternatives in terms of their potential for exploiting favorable contingencies. We will illustrate this later. We explain these ideas with an example.

25.2 Gap-Closing Electrostatic Actuators

The non-linear force-displacement relation for the gap-closing electrostatic actuator in Fig. 25.1 is fairly well represented by:

$$F = kx - \frac{\varepsilon A V^2}{2(g-x)^2} \tag{25.1}$$

where F is the applied force, x is the displacement, ε is the dielectric constant, A is the area of the plates, V is the electric potential on the device, k is the spring stiffness and g is the initial gap size.

Clever mechanical design can circumvent the complex non-linearity of Eq. (25.1). Figure 25.2 shows a mechanically linearized modification of the device in Fig. 25.1 for which the force-displacement relation is, nominally, linear:

$$F = Kx \tag{25.2}$$

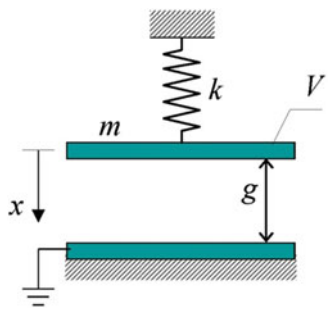


Fig. 25.1 Gap-closing electrostatic actuator. The figure is reproduced here with the permission of Prof. David Elata, head, Mechanical Engineering Micro Systems (MEMS) lab, Technion—Israel Institute of Technology

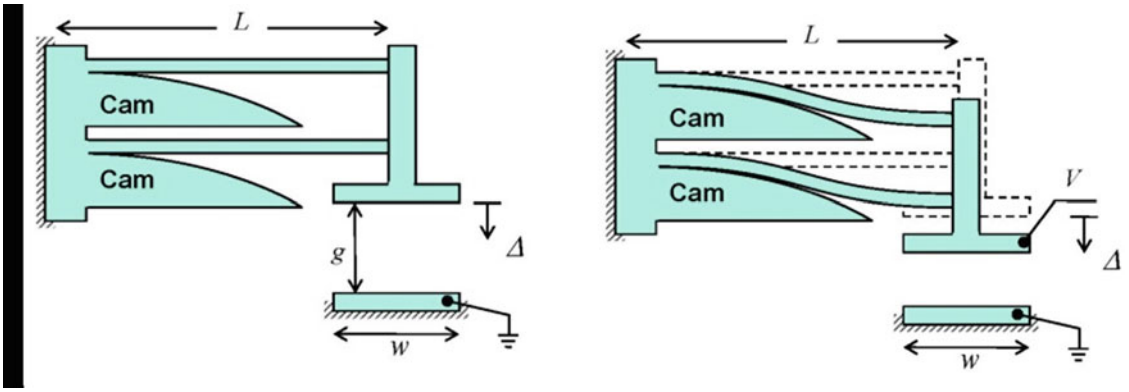


Fig. 25.2 Mechanically linearized gap-closing electrostatic actuator. The figure is reproduced here with the permission of Prof. David Elata

where K is a constant stiffness coefficient. The degree of linearity depends on the shapes of the cams and on the degree of mechanical and structural uniformity of the pair of beams. We will explore the robustness to uncertainty in the stiffness coefficient of the linearized beam. We will also explore robustness to uncertainty in a probabilistic model. Finally we will consider opportuneness. We consider F and x to be positive.

In our first approach to this problem we suppose that our knowledge of the stiffness coefficient, K, is quite limited. We know an estimated value, $\tilde{K}$, and we have an estimate of the error, s, but the most we can confidently assert is that the true stiffness, K, deviates from the estimate by $\pm s$ or more. We do not know a worst case or maximum error, and we have no probabilistic information about K.

There are many types of info-gap models of uncertainty [1]. A fractional-error info-gap model is suitable to this state of knowledge:

$$U(h) = \left\{ K : \quad K > 0, \quad \left| \frac{K - \tilde{K}}{s} \right| \le h \right\}, \quad h \ge 0 \tag{25.3}$$

The info-gap model of uncertainty in Eq. (25.3) is an unbounded family of sets of possible values of the uncertain entity, which is the stiffness coefficient K in the present case. For any non-negative value of h, the set $U(h)$ is an interval of K values. Like all info-gap models, this one has two properties: nesting and contraction. 'Nesting' means that the set, $U(h)$, becomes more inclusive as h increases. 'Contraction' means that $U(h)$ is a singleton set containing only of the known nominal value, $\tilde{K}$, when $h = 0$. These properties endow h with its meaning as an 'horizon of uncertainty'.

Info-gap robustness is based on three components: a system model, an info-gap uncertainty model, and one or more performance requirements. In this present case, Eq. (25.2) is the system model and Eq. (25.3) is the uncertainty model. Our performance requirement is that the displacement, x, be no less than the critical value x_c.

The info-gap robustness is the greatest horizon of uncertainty, h, up to which the system model obeys the performance requirement:

$$\widehat{h}(x_c) = \max \left\{ h : \quad \left(\min_{K \in U(h)} x \right) \ge x_c \right\} \tag{25.4}$$

Reading this equation from left to right we see that the robustness $\widehat{h}$ is the maximum horizon of uncertainty, h, up to which all realizations of the uncertain stiffness K in the uncertainty set $U(h)$ result in displacement x no less than the critical value x_c.

Robustness is a useful decision support tool because more robustness against uncertainty is better than less. Given two options that are more or less equivalent in other respects but one is more robust than the other, the robust-satisficing decision maker will prefer the more robust option. In short, "bigger is better" when prioritizing decision options in terms of robustness.

Derivation of the robustness function is particularly simple in this case. From the system model we know that $x = \frac{F}{K}$. Let $m(h)$ denote the inner minimum in Eq. (25.4), and note that this minimum occurs, at horizon of uncertainty h, when $K = \tilde{K} + sh$. The robustness is the greatest value of h up to which $m(h)$ is no less than x_c:

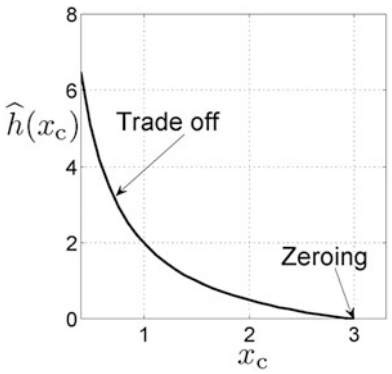


Fig. 25.3 Robustness curve of Eq. (25.5). $F/s = 3$. $\tilde{K}/s = 1$

$$m(h) = \frac{F}{\tilde{K} + sh} \geq x_c \qquad \Rightarrow \qquad \widehat{h}(x_c) = \frac{1}{s}\left(\frac{F}{x_c} - \tilde{K}\right) \tag{25.5}$$

or zero if this is negative which occurs if the performance requirement, x_c, is too large to be obtained with the nominal system.

The robustness function in Eq. (25.5) demonstrates two fundamental properties that hold for all info-gap robustness functions: trade off and zeroing, illustrated in Fig. 25.3.

The performance requirement is that the displacement, x, be no less than the critical value x_c. This requirement becomes more demanding as x_c increases. We see from Eq. (25.5) and Fig. 25.3 that the robustness decreases as the requirement becomes more demanding. That is, robustness trades off against performance: the robustness can be increased only by relaxing the performance requirement. The negative slope in Fig. 25.3 represents the *trade off between robustness and performance:* strict performance requirements, demanding very good outcome, are less robust against uncertainty than lax requirements. This trade off quantifies the intuition of any healthy pessimist: more demanding requirements are more vulnerable to surprise and uncertainty than more modest requirements.

The second property illustrated in Fig. 25.3 is *zeroing*. Our best estimate of the stiffness is $\tilde{K}$, so the predicted displacement is $x = F/\tilde{K}$. Eq. (25.5) shows that the robustness becomes zero precisely at the value of the critical displacement, x_c, that is predicted by the nominal model: $x_c = F/\tilde{K}$, which equals 3 for the parameter values in Fig. 25.3. Stated differently, the zeroing property asserts that best-model predictions have no robustness against error in the model. Like trade off, this is true for all info-gap robustness functions.

Models reflect our best understanding of the system and its environment. Nonetheless, the zeroing property means that model predictions are not a good basis for design or planning decisions, because those predictions have no robustness against errors in the models. Recall that we're discussing situations with large info-gaps. If your models are correct (no info-gaps), then you don't need robustness against uncertainty.

The zeroing property asserts that the predicted outcome is not a reliable characterization of the design. The trade off property quantifies how much the performance requirement must be reduced in order to gain robustness against uncertainty. The slope of the robustness curve reflects the cost of robustness: what decrement in performance "buys" a specified increment in robustness. Outcome-quality can be exchanged for robustness, and the slope quantifies the cost of this exchange. In Fig. 25.3 we see that the cost of robustness is very large at large values of x_c, and decreases as the performance requirement is relaxed.

The robustness function is useful as a response to the pernicious side of uncertainty. In contrast, the opportuneness function is useful in exploiting the potential for propitious surprise. We now discuss the info-gap opportuneness function.

Info-gap opportuneness is based on three components: a system model, an info-gap model of uncertainty, and a performance aspiration. The performance aspiration expresses a desire for better-than-anticipated outcome resulting from propitious surprise. This differs from the performance requirement for robustness that expresses an essential or critical outcome without which the result would be considered a failure.

We will illustrate the opportuneness function by continuing the same example, with positive F and x. The robust-satisficing decision maker requires that the displacement be no less that x_c. The opportune windfalling decision maker recognizes that a larger displacement would be better, especially if it exceeds the anticipated displacement, $F/\tilde{K}$. For the opportune windfaller, the displacement would be wonderful if it is as large as x_w, which exceeds the anticipated displacement. The windfaller's aspiration is not a performance requirement, but it would be great if it occurred.

Achieving a windfall requires a favorable surprise, so the windfaller asks: what is the lowest horizon of uncertainty at which windfall is possible (though not necessarily guaranteed)? The answer is the opportuneness function, defined as:

$$\widehat{\beta}\,(x_w) = \min\left\{h: \quad \left(\max_{K\in U(h)} x\right) \geq x_w\right\} \tag{25.6}$$

Reading this equation from left to right we see that the opportuneness, $\widehat{\beta}$, is the minimum horizon of uncertainty, h, up to which at least one realization of the uncertain stiffness K in the uncertainty set $U(h)$ results in displacement x at least as large as the wonderful windfall value x_w. The opportuneness function, $\widehat{\beta}\,(x_w)$, is the complement of the robustness function $\widehat{h}\,(x_c)$ in Eq. (25.4). We see, for example, that the min and max operators in these two equations are reversed. This is the mathematical manifestation of the inverted meaning of these two functions. Robustness is the *greatest* uncertainty that *guarantees* the *required* outcome, while opportuneness is the *lowest* uncertainty that *enables* the *aspired* outcome.

Opportuneness is useful for decision support because a more opportune option is better able to exploit propitious uncertainty than a less opportune option. An option whose $\widehat{\beta}$ value is small is opportune because windfall can occur even at low horizon of uncertainty. The opportune windfaller prioritizes options according to the smallness of their opportuneness function values: an option with small $\widehat{\beta}$ is preferred over an option with large $\widehat{\beta}$. That is, "smaller is better" for opportuneness, unlike robustness for which "bigger is better". Once again we note the logical inversion between robustness and opportuneness.

Whether a decision maker prioritizes the options using robustness or opportuneness is a methodological decision that may depend on the degree of risk aversion of the decision maker. Furthermore, these methodologies may or may not prioritize the options in the same order.

The opportuneness function is derived in a manner analogous to the derivation of Eq. (25.5), yielding:

$$\widehat{\beta}\,(x_{\mathrm{w}}) = \frac{1}{s}\left(\tilde{K} - \frac{F}{x_{\mathrm{w}}}\right) \tag{25.7}$$

or zero if this is negative which occurs when x_{w} is so small, modest, and unambitious that it is possible even with the nominal design and does not depend on the potential for propitious surprise.

The robustness and opportuneness functions, Eqs. (25.5) and (25.7), are plotted in Fig. 25.4. The opportuneness function displays zeroing and trade off properties whose meanings are the reverse of those for robustness. The opportuneness function equals zero at the nominal outcome, $x = F/\tilde{K}$, like the robustness function. However, for the opportuneness function this means that favorable windfall surprise is not needed in order to enable the predicted outcome. The positive slope of the opportuneness function means that greater windfall (larger x_{w}) is possible only at larger horizon of uncertainty.

The robustness and opportuneness functions may respond differently to proposed changes in the design, as we now illustrate with Eqs. (25.5) and (25.7). From the first of these equations we note that $\widehat{h}$ decreases as the nominal stiffness, $\tilde{K}$, increases. From the second equation we see that $\widehat{\beta}$ increases as $\tilde{K}$ increases:

$$\frac{\partial \widehat{h}}{\partial \tilde{K}} < 0, \qquad \frac{\partial \widehat{\beta}}{\partial \tilde{K}} > 0 \tag{25.8}$$

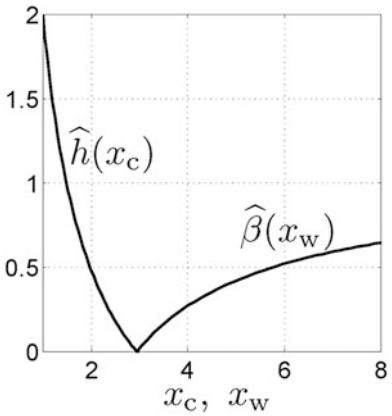


Fig. 25.4 Robustness and opportuneness curves of Eqs. (25.5) and (25.7). $F/s = 3$. $\tilde{K}/s = 1$

Recall that "bigger is better" for robustness while "smaller is better" for opportuneness. We see that any increase in $\tilde{K}$ will make both robustness and opportuneness worse, and any decrease in $\tilde{K}$ will improve them both. We summarize this by saying the robustness and opportuneness are *sympathetic* with respect to change in stiffness.

Now consider the estimated error, s in the info-gap model of Eq. (25.3). A smaller value of s implies greater confidence in the estimate, $\tilde{K}$, while a larger s implies a greater propensity for error in the estimate. From Eq. (25.5) we see that the robustness improves ($\widehat{h}$ increases) as s decreases: better estimate of $\tilde{K}$ implies greater robustness against uncertainty in K. In contrast, from Eq. (25.7) we see that the opportuneness gets worse ($\widehat{\beta}$ increases) as s decreases: lower opportunity for windfall as uncertainty of the estimate declines. In short:

$$\frac{\partial \widehat{h}}{\partial s} < 0, \qquad \frac{\partial \widehat{\beta}}{\partial s} < 0 \tag{25.9}$$

A change in the estimated error acts differently on robustness and opportuneness: by reducing the error of the estimated stiffness one increases the robustness but diminishes the opportuneness; increasing the error acts in the reverse. In short, robustness and opportuneness are *antagonistic* with respect to error in the estimated stiffness.

An innovation dilemma occurs when the decision maker must choose between two options, where one is putatively better but more uncertain than the other. Technological innovations provide the paradigm for this dilemma. An innovation is supposedly better than the current state of the art, but the innovation is new so there is less experience with it and its behavior in practice may turn out worse than the current state of the art. We will illustrate an innovation dilemma with the previous example, demonstrating its resolution using the robustness functions of the two options.

Consider two alternative designs of the linear elastic system, Eq. (25.2), one of which has lower estimated stiffness than the other:

$$\tilde{K}_1 < \tilde{K}_2 \tag{25.10}$$

Both designs will operate under the same positive force, F, so the predicted displacement, $x = F/\tilde{K}$, is greater with option 1. Thus option 1 is preferred based on the estimated stiffnesses and the requirement for large displacement.

However, the putatively better option 1 is based on innovations for which the actual stiffness, in operation, is more uncertain than for option 2 which is the state of the art. Referring to the uncertainty estimate, s, in the info-gap model of Eq. (25.3), we express this as:

$$s_1 > s_2 \tag{25.11}$$

The dilemma is that option 1 is putatively better, Eq. (25.10), but more uncertain, Eq. (25.11). This dilemma is manifested in the robustness functions for the two options, which also leads to a resolution, as we now explain.

To illustrate the analysis we evaluate the robustness function for each option, Eq. (25.5), with the following parameter values: $F = 1$, $\tilde{K}_1 = 1/6$, $s_1 = 1$, $\tilde{K}_2 = 1/3 = s_2$. The robustness curves are shown in Fig. 25.5.

The innovative option 1 in Fig. 25.5 (dashed curve) is putatively better than the state of the art option 2 (solid curve) because the predicted displacement of option 1 is $F/\tilde{K}_1 = 6$ while the predicted displacement for option 2 is only 3. However, the greater uncertainty of option 1 causes a stronger trade off between robustness and performance than for option 2. The cost of robustness is greater for option 1 for x_c values exceeding about 2, causing the robustness curves cross one another at about $x_c = 2.4$.

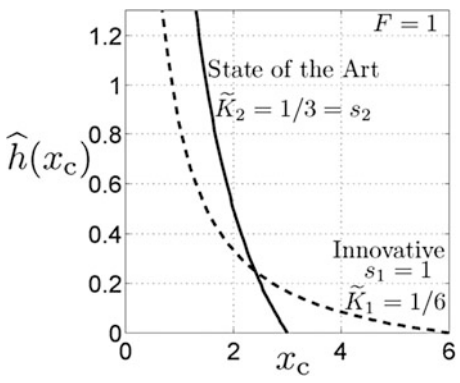


Fig. 25.5 Robustness curves for innovative and state of the art options

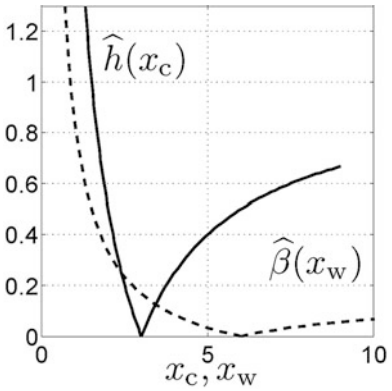


Fig. 25.6 Robustness and opportuneness curves for innovative (*dashed*) and state of the art (*solid*) options

The innovation dilemma is manifested graphically by the intersection of the robustness curves in Fig. 25.5, and this intersection is the basis for the resolution. Innovative option 1 is more robust than option 2 for highly demanding requirements, $x_c > 2.4$, and hence option 1 is preferred for this range of performance requirements. Likewise, state-of-the-art option 2 is more robust for more modest requirements, $x_c < 2.4$, and hence option 2 is preferred for this lower range of performance requirements.

The robust-satisficing designer will be indifferent between the two options for performance requirements at or close to the intersection value of $x_c = 2.4$. Considerations other than robustness can then lead to a decision. Figure 25.6 shows the robustness curves from Fig. 25.5 together with the opportuneness curves, Eq. (25.7), for the same parameter values. We note that the innovative (dashed) option is more opportune (smaller $\widehat{\beta}$) than the state of the art (solid) for all values of x_w exceeding the nominal innovative value. Designers tend to be risk averse and to prefer robust-satisficing over opportune windfalling. Nonetheless, opportuneness can "break the tie" when robustness does not differentiate between the options at the specified performance requirement.

Functional uncertainty. We have discussed the info-gap robustness function and its properties of trade off, zeroing, and cost of robustness. We have illustrated how these concepts support the decision making process, especially when facing an innovation dilemma. Finally, we have described the info-gap opportuneness function and its complementarity to the robustness function. These ideas have all been illustrated in the context of a one-dimensional linear system with uncertainty in a single parameter. In most applications with severe uncertainty the info-gaps include multiple parameters as well as uncertainty in the shapes of functional relationships. We now extend the previous example to illustrate the modeling and management of functional uncertainty. This will also illustrate how uncertain probabilistic models can be incorporated into an info-gap robust-satisficing analysis.

Let the stiffness coefficient, K in Eq. (25.2), be a random variable whose estimated probability density function (pdf), $\tilde{p}(K)$, is normal with mean μ and variance σ^2. We are confident that this estimate is accurate for K within an interval around μ of known size $\pm\delta_s$. However, outside of this interval of K values the fractional error of the pdf is unknown. In other words, we are highly uncertain about the shape of the pdf outside of the specified interval. This uncertainty derives from lack of data with extreme K values and absence of fundamental understanding that would dictate the shape of the pdf. We face "functional uncertainty" that can be represented by info-gap models of many sorts, depending on the type of information that is available. Given the knowledge available in this case we use the following fractional-error info-gap model:

$$U(h) = \Big\{ p(K) : \quad \int_{-\infty}^{\infty} p(K)\ dK = 1, \quad p(K) \geq 0 \ \text{ for all } \ K, \\ p(K) = \tilde{p}(K) \quad \text{for} \quad |K - \mu| \leq \delta_s, \\ \left| \frac{p(K) - \tilde{p}(K)}{\tilde{p}(K)} \right| \leq h \quad \text{for} \, |K - \mu| > \delta_s \Big\}, \quad h \geq 0 \tag{25.12}$$

The first row of this info-gap model states that the set $U(h)$ contains functions $p(K)$ that are normalized and non-negative, namely, mathematically legitimate pdf's. The second line states the these functions equal the estimated pdf in the specified interval around the mean, μ. The third line states that, outside of this interval, the functions in $U(h)$ deviate fractionally from the estimated pdf by no more than h. We add the technical simplification that the pdf's in $U(h)$ are non-atomic: they contain no delta functions. In short, this info-gap model is the unbounded family of nested sets, $U(h)$, of pdf's that are known within the interval $\mu \pm \delta_s$ but whose shapes are highly uncertain beyond it. This is one example of an info-gap model for uncertainty in the shape of a function.

The system fails if $x < x_c$ where $x = F/K$ and F is a known positive constant. x is now a random variable (because K is random) so the performance requirement is that the probability of failure not exceed a critical value P_c. We will explore the robustness function. We consider the special case that $F/x_c \geq \mu + \delta_s$, meaning that the failure threshold for K lies outside the interval in which the pdf of K is known.

The probability of failure is:

$$P_f(p) = \text{Prob}\,(x < x_c) = \text{Prob}\,(K > F/x_c) = \int_{F/x_c}^{\infty} p(K)\ \mathrm{d}K \tag{25.13}$$

For the estimated pdf, $\tilde{p}(K)$, one finds the following expression for the estimated probability of failure:

$$P_f\,(\tilde{p}) = 1 - \Phi\left(\frac{(F/x_c) - \mu}{\sigma}\right) \tag{25.14}$$

where $\Phi(\cdot)$ is the cumulative probability distribution function of the standard normal variate.

The robustness function, $\widehat{h}\,(P_c)$, is the greatest horizon of uncertainty h up to which all pdf's $p(K)$ in the uncertainty set $U(h)$ do not have failure probability $P_f(p)$ in excess of the critical value P_c:

$$\widehat{h}\,(P_c) = \max\left\{h:\ \left(\max_{p \in U(h)} P_f(p)\right) \leq P_c\right\} \tag{25.15}$$

After some algebra one finds the following expression for the robustness:

$$\widehat{h}\,(P_c) = \begin{cases} 0 & \text{if} \quad 0 \leq P_c < P_f\,(\tilde{p}) \\ \frac{P_c}{P_f(\tilde{p})} - 1 & \text{if} \quad P_f\,(\tilde{p}) \leq P_c \leq 2P_f\,(\tilde{p}) \\ \infty & \text{otherwise} \end{cases} \tag{25.16}$$

The robustness curve in Eq. (25.16) is illustrated in Fig. 25.7 for $[(F/x_c) - \mu]/\sigma = 3$, meaning that the failure threshold is three standard deviations above the mean. Hence the estimated probability of failure is $P_f\,(\tilde{p}) = 0.00135$. The trade off property is evident in this figure: lower (better) required probability of failure, P_c, entails lower (worse) robustness $\widehat{h}\,(P_c)$. We see the discontinuous jump of robustness to infinity at $P_c = 2P_f\,(\tilde{p})$. This is because the actual probability of failure, $P_f(p)$, cannot exceed more than twice the estimated value, $P_f\,(\tilde{p})$. This results from the constraints on the pdf's in the info-gap model of Eq. (25.12). The zeroing property is expressed by robustness equaling zero when the performance requirement, P_c, equals the estimated value $P_f\,(\tilde{p})$.

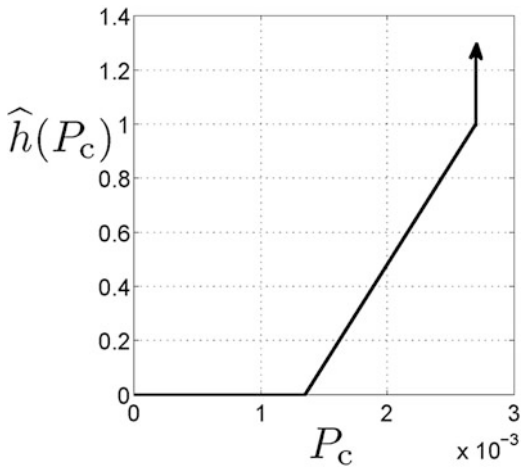


Fig. 25.7 Robustness curve for Eq. (25.16)

25.3 Conclusion

We have discussed quantitative tools—based on mathematical analysis—for assessing severe uncertainty and for prioritizing the options facing a decision maker. We have considered uncertainty both in parameters and in the shape of functions. The theorems underlying these tools reveal intuitively meaningful concepts: the trade off between robustness and requirements, cost of robustness, zero robustness of predicted outcomes, innovation dilemmas and preference reversals, and opportuneness. These insights underlie decision making under severe uncertainty.

References

1. Ben-Haim, Y.: Info-Gap Decision Theory: Decisions Under Severe Uncertainty, 2nd edn. Academic Press, London (2006)
2. Ben-Haim, Y.: Info-Gap Economics: An Operational Introduction. Palgrave-Macmillan, London (2010)
3. Chinnappen-Rimer, S., Hancke, G.P.: Actor coordination using info-gap decision theory in wireless sensor and actor networks. Int. J. Sens. Netw. **10**(4), 177–191 (2011)
4. Harp, D.R., Vesselinov, V.V.: Contaminant remediation decision analysis using information gap theory. Stoch. Env. Res. Risk A. **27**(1), 159–168 (2013)
5. Kanno, Y., Takewaki, I.: Robustness analysis of trusses with separable load and structural uncertainties. Int. J. Solids Struct. **43**(9), 2646–2669 (2006)
6. Burgman, M.: Risks and Decisions for Conservation and Environmental Management. Cambridge University Press, Cambridge (2005)
7. Knoke, T.: Mixed forests and finance—methodological approaches. Ecol. Econ. **65**(3), 590–601 (2008)
8. Ben-Haim, Y., Zetola, N.M., Dacso, C.: Info-gap management of public health policy for TB with HIV-prevalence. BMC Public Health. **12**, 1091 (2012). doi:10.1186/1471-2458-12-1091
9. Moffitt, L.J., Stranlund, J.K., Field, B.C.: Inspections to avert terrorism: robustness under severe uncertainty. J. Homel. Secur. Emerg. Manag.. **2**(3), (2005). http://www.bepress.com/jhsem/vol2/iss3/3.
10. Hall, J.W., Lempert, R.J., Keller, K., Hackbarth, A., Mijere, C., McInerney, D.J.: Robust climate policies under uncertainty: A comparison of robust decision making and info-gap methods. Risk Anal. **32**(10), 1657–1672 (2012)
11. Ben-Haim, Y.: Doing our best: optimization and the management of risk. Risk Anal. **32**(8), 1326–1332 (2012a)
12. Ben-Haim, Y.: Why risk analysis is difficult, and some thoughts on how to proceed. Risk Anal. **32**(10), 1638–1646 (2012b)
13. Schwartz, B., Ben-Haim, Y., Dacso, C.: What makes a good decision? Robust satisficing as a normative standard of rational behaviour. J. Theory Soc. Behav. **41**(2), 209–227 (2011)
14. Frank Knight, H.: Risk, Uncertainty and Profit, Houghton Mifflin Co., Re-issued by University of Chicago Press, 1971. (1921)
15. Schwartz, B.: Paradox of Choice: Why More Is Less. Harper Perennial, New York (2004)
16. Simon, H.: Rational choice and the structure of the environment. Psychol. Rev. **63**(2), 129–138 (1956)

Chapter 26
Bayesian Optimal Experimental Design Using Asymptotic Approximations

Costas Argyris and Costas Papadimitriou

Abstract Bayesian optimal experimental design (OED) tools for model parameter estimation and response predictions in structural dynamics include sampling (Huan and Marzouk, J. Comput. Phys., 232:288–317, 2013) and asymptotic techniques (Papadimitriou et al., J. Vib. Control., 6:781–800, 2000). This work compares the two techniques and discusses the theoretical and computational advantages of asymptotic techniques. It is shown that the OED based on maximizing the expected Kullback-Leibler divergence between the prior and posterior distribution of the model parameters is equivalent, asymptotically for large number of data and small model prediction error, to minimizing asymptotic estimates of the robust information entropy measure introduced in the past (Papadimitriou et al., J. Vib. Control., 6:781–800, 2000; Papadimitriou, J. Sound Vib., 278:923–947, 2004; Papadimitriou and Lombaert, Mech. Syst. Signal Process., 28:105–127, 2012) for structural dynamics applications. Based on the asymptotic approximations, techniques are proposed to overcome the sensor clustering. In addition, an insightful analysis is presented that clarifies the effect of the variances of Bayesian priors on the optimal design. Finally the importance of uncertainties in nuisance model parameters is pointed out and the expected utility functions are extended to take into account such uncertainties. A heuristic forward sequential sensor placement algorithm (Papadimitriou, J. Sound Vib., 278:923–947, 2004) is effective in solving the optimization problem in the continuous physical domain of variation of the sensor locations, bypassing the problem of multiple local/global optima manifested in optimal experimental designs and providing near optima solutions in a fraction of the computational effort required in expensive stochastic optimization algorithms. The theoretical and computational developments are demonstrated for optimal sensor placement designs for applications taken from structural mechanics and dynamics areas. Examples covering the optimal sensor placement design for parameter estimation and response predictions are covered.

Keywords Bayesian inference • Kullback-Leibler divergence • Information entropy • Parameter estimation • Response prediction

26.1 Optimal Experimental Design

Consider a structural model and an excitation model for a structure and let $\underline{\theta} \in R^{N_\theta}$ be the vector of model parameters to be estimated using a set of measured data $\underline{y} \equiv \underline{y}(\underline{\delta}) \in R^N$, $\underline{\varphi}$ be the set of nuisance parameters, and $\underline{g}\left(\underline{\theta}, \underline{\varphi}; \underline{\delta}\right) \in R^N$ be the predictions from the model for given $\underline{\theta}$ and $\underline{\varphi}$. The measured data depend on design variables $\underline{\delta}$ related to the location of sensors placed in a structure or to the characteristics of the excitations. Prior probability distributions $\pi\left(\underline{\theta}, \underline{\varphi}\right)$ are assigned to model the uncertainty in parameters $\underline{\theta}$ and $\underline{\varphi}$, and the Bayesian inference is used to update uncertainties in the parameter $\underline{\theta}$ given the data by assuming a zero-mean Gaussian additive prediction error $\underline{e} = \underline{y} - \underline{g}(\underline{\theta}; \underline{\delta}) \sim N(\underline{0}, \Sigma)$, where $\Sigma = \Sigma(\underline{\sigma}) \in R^{N \times N}$ is the prediction error covariance matrix with parameters included in the nuisance set $\underline{\varphi}$ to yield the posterior distribution $p\left(\underline{\theta} | \underline{y}, \underline{\varphi}, \underline{\delta}\right)$ of the model parameters.

The objective of an experiment is to maximize the information content in the data for estimating the parameters of the model or the response quantities of interest (QoI). Extending Lindley's [1] work to include the uncertainty in the nuisance parameters, one maximizes an expected utility function

C. Argyris • C. Papadimitriou (✉)
Department of Mechanical Engineering, University of Thessaly, Volos, Greece
e-mail: costasp@uth.gr

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_26

$$U\left(\underline{\delta}\right) = \int_{\Phi}\int_{\Upsilon}\int_{\Theta} u\left(\underline{\delta};\underline{\theta},\underline{\varphi},\underline{y}\right)\ p\left(\underline{\theta},\underline{y}|\underline{\varphi},\underline{\delta}\right)\ d\underline{\theta}\ d\underline{y}\ d\underline{\varphi} \tag{26.1}$$

that quantifies the usefulness of the experiment in estimating the model parameters or predicting an output QoI, where $u\left(\underline{\delta};\underline{\theta},\underline{\varphi},\underline{y}\right)$ is the utility function given a particular value of the model and nuisance parameter sets $\underline{\theta}$ and $\underline{\varphi}$ and the outcome $\underline{y}$ from the experiment, $p\left(\underline{\theta},\underline{y}|\underline{\varphi},\underline{\delta}\right) = p\left(\underline{\theta}|\underline{y},\underline{\varphi},\underline{\delta}\right)\ p\left(\underline{y}|\underline{\varphi},\underline{\delta}\right)$, $p\left(\underline{\theta}|\underline{y},\underline{\varphi},\underline{\delta}\right)$ is the posterior uncertainty in the model parameters given the outcome $\underline{y}$ and $p\left(\underline{y}|\underline{\varphi},\underline{\delta}\right)$ is the uncertainty in the data. A rational choice of the utility function is the information gained by the experiment quantified by the Kullback-Leibler divergence between the prior and posterior probability distribution given an outcome $\underline{y}$ obtained from an experimental design $\underline{\delta}$. The expected utility function is an average of the utility function over all possible values of the model and nuisance parameters as they are inferred from the data, and all the possible outcomes of the experiment.

Sampling techniques [2] used to evaluate the multi-dimensional integrals involve a nested sampling procedure [2] resulting in excessive computational effort since a very large number of model runs is required. Instead, asymptotic estimates can be used to significantly reduce the computational burden and gain insight into the factors that affect optimal design [3, 4]. It can be shown that asymptotically for large enough number of data and small enough prediction error the multidimensional integral takes the form

$$U\left(\underline{\delta}\right) = c + \frac{1}{2}\int_{\Phi}\int_{\Theta} \ln\det\left\{\left[\underline{\nabla}_{\theta}\underline{g}\left(\underline{\theta},\underline{\varphi},\underline{\delta}\right)\right]^{T}\ \Sigma^{-1}\left(\underline{\varphi}\right)\ \underline{\nabla}_{\theta}\underline{g}\left(\underline{\theta},\underline{\varphi},\underline{\delta}\right) + Q_{\pi}\left(\underline{\theta}\right)\right\}\ \pi\left(\underline{\theta},\underline{\varphi}\right)\ d\underline{\theta}\ d\underline{\varphi} \tag{26.2}$$

where $Q_{\pi}\left(\underline{\theta}\right)$ is the inverse of the covariance of the prior distribution of the model parameters evaluated at $\underline{\theta}$. The integral (26.2) represents the robust information entropy over all possible values of the model and nuisance parameters quantified by the prior PDF. The multidimensional integral can be evaluated using Monte Carlo techniques or sparse grid methods [5]. In this case it can be seen that the experimental design is based on the average sensitivity of the output quantities of interest that are to be measured from the experimental setup with respect to the model parameters over the prior uncertainty domain in the parameter space $\underline{\theta}$ and $\underline{\varphi}$. This result is same as the result in [3] obtained using the concept of robust information entropy, generalized herein to handle the uncertain nuisance parameter case. Simplifications for the special case of linear models with respect to the model parameters and its significance in structural dynamics is discussed in Argyris and Papadimitriou [6]. The asymptotic approximations developed are useful for analytically investigating the problem of sensor clustering phenomena manifested in structural dynamics for dense finite element meshes [7]. Techniques are proposed to overcome the sensor clustering. In addition, an insightful analysis is presented that clarifies the effect of the variances of Bayesian priors on the optimal design.

In optimal experimental design one maximizes the expected utility function $U\left(\underline{\delta}\right)$ with respect to the he design variables $\underline{\delta}$

$$\underline{\delta}_{opt} = \arg\max_{\underline{\delta}} U\left(\underline{\delta}\right) \tag{26.3}$$

Optimal experimental design problems involving the design the position of sensors often results in multiple local and local solutions. Also the gradient of the objective function with respect to the design variables in most cases of practical interest cannot be evaluated analytically. To avoid premature convergence to a local optimum and the evaluation of sensitivities of the utility function with respect to the design variables, stochastic optimization algorithms such as CMA-ES [8] can be used to find the optimum. Heuristic forward sequential sensor placement algorithms [4] are effective in solving the optimization problem in the continuous physical domain of variation of the sensor locations. The heuristic algorithm bypasses the problem of multiple local/global optima manifested in optimal experimental designs, providing near optima solutions in a fraction of the computational effort required in expensive stochastic optimization algorithms.

26.2 Applications

The applications are selected to give insight into the optimal experimental design methods and compare the effectiveness of the asymptotic approximations in relation to sampling techniques. For this, analytical and insightful results are obtained for structural mechanics and dynamics problems. In structural mechanics, we consider the problem of parameter estimation

and response prediction. The objective is to design the location of displacement and strain measurements for estimating the location and magnitude of loads. The effect of uncertainty in nuisance parameters, related to stiffness boundary condition parameters that could be considered uncertain and modeled based on engineering experience, is investigated. A formulation is presented to treat any type of structure. Results are presented for simplified structural elements in order to gain insight into the OED as well as compare asymptotic and sampling algorithms. Such simplified elements include bar and beam models. Truss and frame structures are used to test the effectiveness of the sequential sensor placement strategies by comparing it to exact results obtained from stochastic optimization method such as the CMA-ES.

In structural dynamics, a number of important problems for optimal sensor location and optimal design of excitation characteristics have been considered in past studies [3, 4, 7]. Selected cases are revisited in this work to demonstrate the re-formulation of the optimization problem in continuous space, the efficiency of sequential sensor placement algorithms applied over the continuous space of design variables, the effect of uncertainties in nuisance parameters and the use of prior variances to control the optimal design based on the importance of the uncertain parameters or response predictions. In particular formulating the problem of optimally locating the sensors in continuous space allows one to simultaneously optimize with respect to sensor location and excitation characteristics for both parameter estimation and response predictions. Furthermore, the effect of the uncertainty in nuisance parameters in optimal experimental design is explored by optimally designing the location of sensors for displacement, acceleration and/or strain sensor locations for response reconstruction. Kalman filter techniques are used for response reconstruction and the optimal sensor placement designs proposed are robust to uncertainties in the nuisance parameters such as stiffness and other structural parameters considered to be uncertain during model specification.

26.3 Conclusions

Asymptotic methods, valid for small prediction errors and relatively large number of data, are shown to be convenient, computationally tractable and sufficiently accurate as compared to sampling algorithms that require the specification of the magnitude of prediction error of the used model, a quantity not known during the experimental design phase. Asymptotic methods also provide valuable insight into the dependence of the optimal design on the sensitivity of the response quantities on the model parameters, the effect of priors in the design, as well as the effect of uncertainties in the nuisance parameters. Developments are demonstrated for optimal sensor placement designs of selected structural mechanics and dynamics problems.

References

1. Lindley, D.V.: On a measure of the information provided by an experiment. Ann. Math. Stat. **27**, 986–1005 (1956)
2. Huan, X., Marzouk, Y.M.: Simulation-based optimal Bayesian experimental design for nonlinear systems. J. Comput. Phys. **232**(1), 288–317 (2013)
3. Papadimitriou, C., Beck, J.L., Au, S.K.: Entropy-based optimal sensor location for structural model updating. J. Vib. Control. **6**(5), 781–800 (2000)
4. Papadimitriou, C.: Optimal sensor placement methodology for parametric identification of structural systems. J. Sound Vib. **278**(4), 923–947 (2004)
5. Gerstner, T., Griebel, M.: Numerical integration using sparse grids. Numer. Algorithms. **18**, 209–232 (1998)
6. Argyris, C., Papadimitriou, C.: A Bayesian framework for optimal experimental design in structural dynamics. In: 34th International Modal Analysis Conference (IMAC), January 24–28, 2016, Orlando, Florida
7. Papadimitriou, C., Lombaert, G.: The effect of prediction error correlation on optimal sensor placement in structural dynamics. Mech. Syst. Signal Process. **28**, 105–127 (2012)
8. Hansen, N., Muller, S.D., Koumoutsakos, P.: Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). Evol. Comput. **11**(1), 1–18 (2003)

Chapter 27
Surrogate-Based Approach to Calculate the Bayes Factor

Ramin Madarshahian and Juan M. Caicedo

Abstract Calculation of the Bayes factor for comparison of computationally expensive models is not a straight forward process. To obtain the model's evidence, numerical integration methods are often implemented. To apply these methods, many samples should be obtained from the model, requiring the model to be run thousands of times. In this paper, we propose to use a probability distribution function as a surrogate model for the evidence. Considering the second axiom of probability, a modified scale factor used for fitting the surrogate model to the integrand would be an estimate of the model's evidence, and no further integration is needed after the surrogate model is fitted to the integrand. Two structural models are compared to explore the advantages and disadvantages of the proposed method. The Bayes factor is estimated for both models using Monte Carlo integration and the proposed method. Results show that fewer numbers of samples of the structural models are needed when the proposed method is applied. Therefore, this method reduces the computational cost required to compare models in a probabilistic sense.

Keywords Model evidence • Metamodeling • Bayesian inference • MCMC • Computational cost

27.1 Introduction

Application of Bayesian modeling in engineering problems has recently become popular [1–5]. The computational cost of numerical models is an important issue for today's scientific world [6, 7]. Moreover, when faced with a complex problem, several computational models could be proposed, and one of them should be selected as the final model [8]. The Bayes factor is a metric used to select a model considering both the data and the experience of the analyst [9]. Calculating the Bayes factor for computational expensive models, however, is compositionally expansive. To calculate the Bayes factor, a complicated integral, which also contains the model in its equation, should be solved.

Surrogate models are used to approximate a complex model and reduce its computational cost. So far, many different metamodeling techniques have been introduced in the science and engineering fields [10]. The idea is to get samples from the expensive model and fit the appropriate metamodel based on these samples. The expensive model would then, be replaced by the metamodel in other needed analysis. In Bayesian model selection, the expensive model can be replaced by a metamodel [11, 12]. In this paper, we propose the use of a surrogate model after forming the Bayesian inference relationship. The evidence is obtained by calculating the integral of the product of the likelihood of a Bayesian model and its parameters' priors. This integrand, called Evidence Density Function (EDF) in this paper, has some of the characteristics of probability distribution functions (PDFs) [13, 14]. For example, EDFs are always positive. Therefore, we propose to use a PDF to model the EDF. The main advantage of our proposed technique is the simplicity of the evidence calculation after metamodeling. To fit the metamodel to the EDF a scale factor is needed. Then the evidence would be a scale factor multiplied by a modification factor, which was obtained from the CDF of the metamodel. This is explained in detail in the following sections.

R. Madarshahian (✉) • J.M. Caicedo
University of South Carolina, 300 Main St., Columbia, SC 29208, USA
e-mail: mdrshhn@email.sc.edu; caicedo@cec.sc.edu

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_27

27.2 Methodology

Considering the model, M, and the data, D_r, used to infer the model parameters, θ, the Bayesian equation is formed as:

$$P(\theta|D_r, M) = \frac{P(D_r|\theta)P(\theta)}{\int_{\Omega_\theta} P(D_r|\theta)P(\theta)d\theta} \tag{27.1}$$

Where Ω_θ is the domain of parameters. The denominator in Eq. (27.1) is the Bayesian evidence for model M. The term in the integral is modeled by a suitable PDF, $f(\theta, \beta)$:

$$P(D_r|\theta)P(\theta) = \rho f(\theta, \beta) \tag{27.2}$$

Where β are the PDF's parameters. Then:

$$\int_{\Omega_\theta} \rho f(\theta, \beta)d\theta = \rho F|\Omega_\theta \tag{27.3}$$

Where F is the CDF of the $f(\theta, \beta)$. To obtain β and ρ, another Bayesian equation could be formed. The samples from the EDF, D, are used for the new inference. In the example discussed in the next section, the standard deviation of the posterior predictive samples are used to form a sampling strategy for D.

27.3 Example

27.3.1 Problem Definition

Figure 27.1 shows the model representing the "true" structure. Data is obtained by measuring the end displacement of the beam when a point load P, or moment M, is exerted to the end of the beam. The parameter θ is the ratio of the cantilever part of the beam to the whole length of the beam. The "true" model has a value of $\theta = 0.5$. A normally distributed number with a mean zero and a standard deviation of 0.005 m is added to the end displacement as the measurement's noise. The parameter of the interest is θ, and two models are proposed. The first model is the same as the "true" structure. The second model considers constraining the rotational DOF of the internal support. Table 27.1 shows different loadings and their correspond end displacement of the beam.

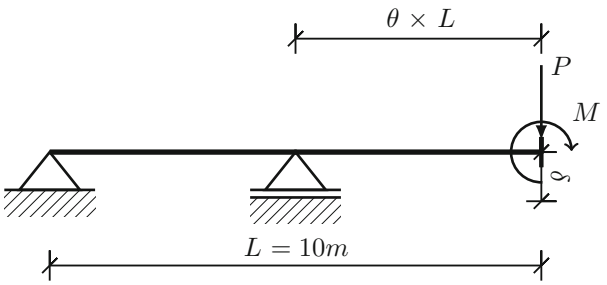


Fig. 27.1 The first model with a simple support within the length of the beam

Table 27.1 Data used in the example

P (N)	M (N m)	Displacement (m)
20,000	0	0.0855
20,000	0	0.0817
0	50,000	0.0642
0	50,000	0.0508

Table 27.2 Monte Carlo integration

Iteration	Obtained value	Error	*P*-value	Total number of samples
Model 1				
1	89,798	10,479	1.00	1000
2	94,862	2506	0.25	2000
3	95,166	1023	0.13	3000
Model 2				
1	40,007	9028	1.00	1000
2	52,346	2280	0.16	2000
3	52,777	587	0.36	3000

27.3.2 Monte Carlo Estimate

Bayes factor is the ratio on the Bayesian evidence of the two models. Here, normal distribution is assumed for the likelihood, so standard deviation of the likelihood, σ, is considered as a free parameter.

To calculate the evidence for each model Monte Carlo estimate of each integral using the Vegas algorithm is obtained [15]. The algorithm is stopped after three iterations, when the error is approximately 10% of the final result. A total of 3000 model evaluations are used in this process. The Bayes factor is calculated as the ratio of the evidences, $B12 = \frac{95166}{52777} = 1.803$, which shows a preference for model 1 as expected. Table 27.2 summarizes the obtained results using the Vegas algorithm.

27.3.3 Proposed Method

The evidence in our example contains two parameters (θ, σ). Therefore, a bivariate distribution is used to model the EDF. Based on our experience, we know that the distribution for σ is usually skewed. So we selected a bivariate distribution with a normal distribution for θ and an inverse Gaussian distribution for σ. Seven parameters are needed to be updated including: Mean and standard deviation for normal distribution and mean and scale and shape parameters for the inverse Gaussian distribution, the scale factor to fit the metamodel, and standard deviation of the likelihood of a Bayesian model used to fit the metamodel. Samples of the EDF, requiring evaluations of the structural model, are drawn using the following strategy. First, 30 samples are randomly selected from priors of θ and σ. Then, a mesh of points are constructed in the θ and σ domain, and posterior predictive samples are obtained for each of these points. The next sample of the EDF is obtained from the cell with the highest standard deviation of the predictive samples. The process is stopped when the estimated evidence does not change significantly between steps. 50 and 45 samples are obtained for model 1 and 2 respectively. Figure 27.2 shows the samples drawn for each model. Figure 27.3 shows the fitted metamodels on EDF of model 1 and 2. Finally, MCMC chains are obtained for all parameters of the metamodel, and the Bayes factor is calculated from these samples as indicated in a histogram is shown in Fig. 27.4.

27.3.4 Concluding Remarks

The Bayes factor comparing the two structural models is calculated using two different methods. (1) Using a traditional MCMC algorithm (Vegas) using 3000 evaluation of the structural model, and (2) Using a proposed technique that uses a metamodel of the EDF and only required 45, and 50 evaluations of the structural model 1 and 2 respectively. Results of the proposed technique show that there is a 95% probability that the Bayes factor is between 1.792 and 1.816. These results includes the value of 1.803 obtained with the Vegas algorithm. The proposed methodology has the potential to reduce the computational cost of model selection of computationally expensive models as it only uses a fraction of the samples than traditional methods.

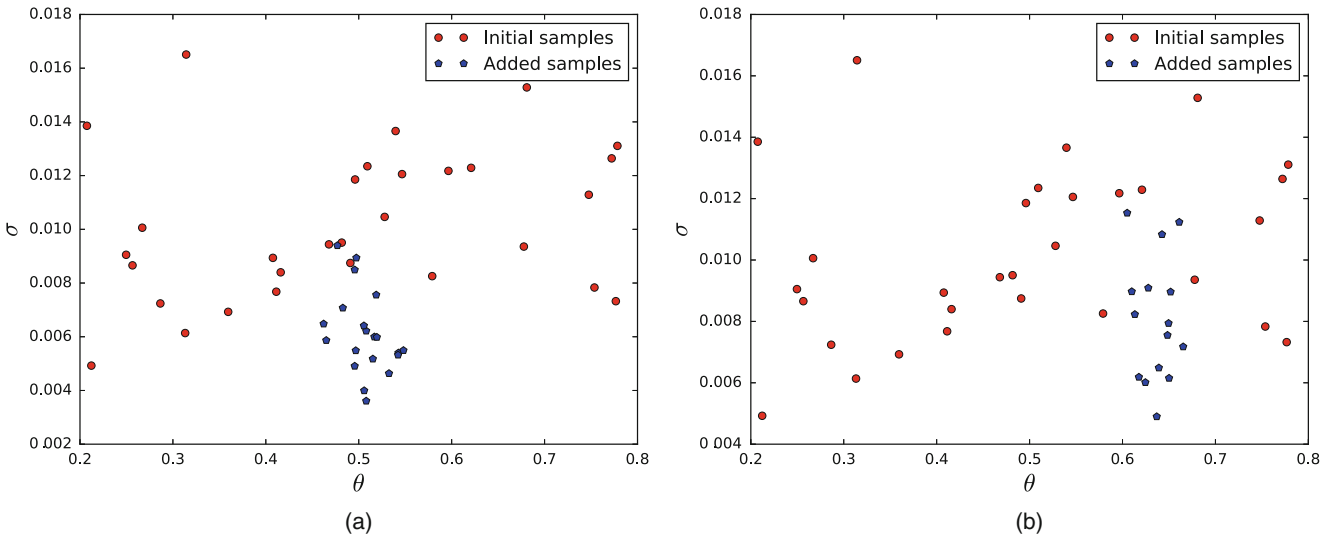


Fig. 27.2 Sampling for both models. (**a**) Model 1. (**b**) Model 2

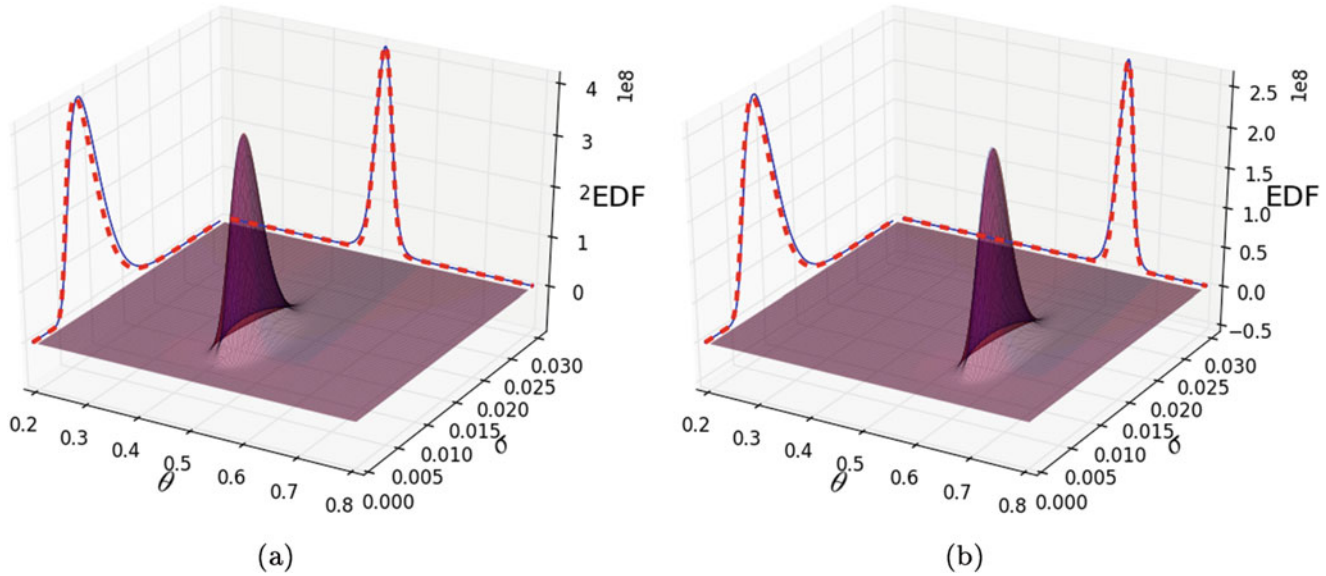


Fig. 27.3 Fitted metamodel for model 1 EDF (*red surface* and *dashed red lines* are related to the metamodel). (**a**) Model 1. (**b**) Model 2

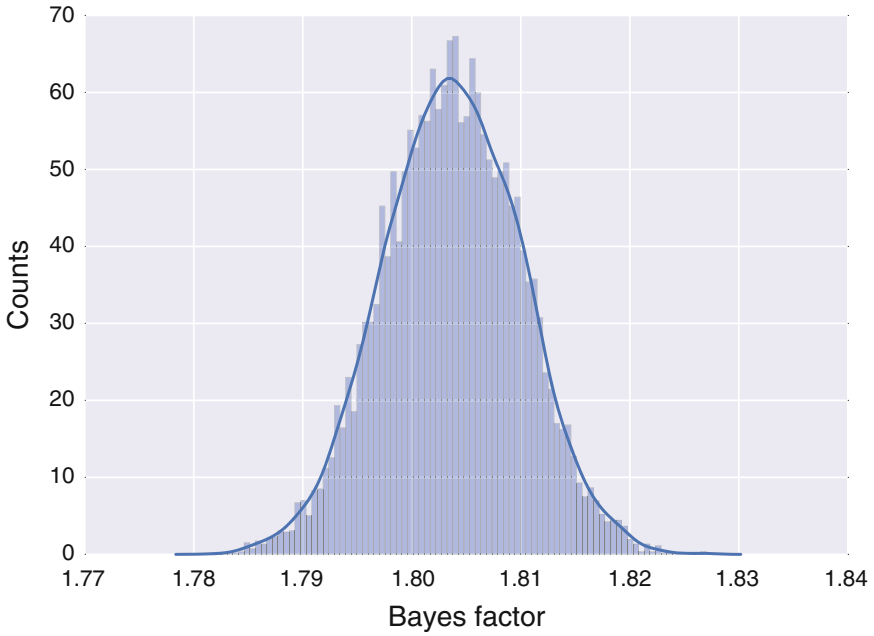


Fig. 27.4 Histogram for Bayes factor

References

1. Vakilzadeh, M.K., Huang, Y., Beck, J.L., Abrahamsson, T.: Approximate Bayesian computation by subset simulation using hierarchical state-space models. Mech. Syst. Signal Process. **84**, 2–20 (2017)
2. Ortiz, A.R., Madarshahian, R., Caicedo, J.M., Rizos, D.: Transfer length probabilistic model updating in high performance concrete. In: Dynamics of Civil Structures, vol. 2, pp. 325–330. Springer, Cham (2016)
3. Papadimitriou, C.: Bayesian uncertainty quantification and propagation (UQ+ P): state-of-the-art tools for linear and nonlinear structural dynamics models. In: Identification Methods for Structural Health Monitoring, pp. 137–170. Springer, Cham (2016)
4. Behmanesh, I., Moaveni, B.: Accounting for environmental variability, modeling errors, and parameter estimation uncertainties in structural identification. J. Sound Vib. **374**, 92–110 (2016)
5. Miles, P., Hays, M., Smith, R., Oates, W.: Bayesian uncertainty analysis of finite deformation viscoelasticity. Mech. Mater. **91**, 35–49 (2015)
6. Boateng, L.K., Madarshahian, R., Yoon, Y., Caicedo, J.M., Flora, J.R.: A probabilistic approach for estimating water permeability in pressure-driven membranes. J. Mol. Model. **22**(8), 185 (2016)
7. Kennedy, G.J., Martins, J.R.: A parallel finite-element framework for large-scale gradient-based design optimization of high-performance structures. Finite Elem. Anal. Des. **87**, 56–73 (2014)
8. Oliver, J., Huespe, A., Sanchez, P.: A comparative study on finite elements for capturing strong discontinuities: E-FEM vs X-FEM. Comput. Methods Appl. Mech. Eng. **195**(37), 4732–4752 (2006)
9. Wasserman, L.: Bayesian model selection and model averaging. J. Math. Psychol. **44**(1), 92–107 (2000)
10. Wang, G.G., Shan, S.: Review of metamodeling techniques in support of engineering design optimization. J. Mech. Des. **129**(4), 370–380 (2007)
11. Matott, L.S., Rabideau, A.J.: Calibration of complex subsurface reaction models using a surrogate-model approach. Adv. Water Resour. **31**(12), 1697–1707 (2008)
12. Govers, Y., Khodaparast, H.H., Link, M., Mottershead, J.: A comparison of two stochastic model updating methods using the {DLR} {AIRMOD} test structure. Mech. Syst. Signal Process. **52–53**, 105–114 (2015)
13. Madarshahian, R., Caicedo, J.M.: Metamodeling of Model Evidence, pp. 307–313. Springer, Cham (2016)
14. Madarshahian, R., Caicedo, J.M.: Reducing MCMC computational cost with a two layered Bayesian approach. In: Model Validation and Uncertainty Quantification, vol. 3, pp. 291–297. Springer, Cham (2015)
15. Lepage, G.P.: A new algorithm for adaptive multidimensional integration. J. Comput. Phys. **27**(2), 192–203 (1978)

Chapter 28
Vibrational Model Updating of Electric Motor Stator for Vibration and Noise Prediction

M. Aguirre, I. Urresti, F. Martinez, G. Fernandez, and S. Cogan

Abstract In order to improve the comfort of passengers in electrical vehicles, it is increasingly important to consider the vibroacoustic behavior of electrical machines during the design phase. In this work, a weakly coupled multiphysical model for electrical machine vibration and noise prediction is presented and applied to a 75 kW railway traction motor. The main objectives of the model are to obtain firstly the vibrational level and secondly the acoustic pressure level predictions. The multiphysical model includes an electromagnetic 2D model, a 3D structural vibrational model and an acoustic model, all of them based on the finite element method. The work is focused on the validation of the modal analysis and vibrational models, using a bottom-up approach. Experimental modal analyses at different assembly stages are performed in order to update uncertain input parameters of the structural model at those levels. An anisotropic damping model is developed and updated in order to obtain adequate FRF amplitudes and the mean squared error (MSE) metric is employed to quantify the correlation between the experimental and numerical results. Finally, vibrational spectra under nominal operational conditions of the motor are used to demonstrate the adequacy of the vibrational model.

Keywords Electrical machine • Multiphysical model • Vibrational model • Anisotropic damping • Mean squared error metric

28.1 Introduction

For some years now, noise has become a major factor in the ambient quality. Several laws and standards have been written to limit the total amount of noise to a safe and comfortable level [1, 2]. This is why engineers are constantly challenged to reduce the levels of vibration and noise, increase the life expectancy of components and improve the efficiency of machines. Since electrical machines are complex systems where a great amount of physical phenomena are produced simultaneously, a detailed multidisciplinary approach taking into account the coupling between different physical fields is needed in order to implement a fast, accurate, reliable and optimized design methodology. This paper presents a sequentially coupled multiphysical model where electromagnetic, structural vibrational and acoustic aspects are taken into account for a Permanent Magnet Synchronous Machine (PMSM). This type of calculation has been presented and implemented by several authors. For example, Rainer et al. [3] computed the dynamic response of a skewed induction machine and studied the accuracy in the frequency domain. Pellerey et al. [4] also applied this methodology to a wound rotor synchronous motor. Dupont et al. [5] applied to electric motors from automotive industry trying to validate the dynamic response of the stator, and Abrahamsson et al. [6] calibrated a finite element model of a car front subframe against test data.

In the field of structural dynamics, reliable finite element (FE) response predictions are becoming increasingly important to industry and there is a real interest to improve these in the light of measured frequency response functions (FRFs). Model updating using FRFs is a modern design technology which improves the predictive capabilities of computer-based models of structural dynamics problems. Grafe [7] emphasized the importance of using FRF based correlation and model updating formulations instead of modal-based formulations for reliable finite element (FE) response predictions. In this paper three different FRF correlation metrics are presented and an amplitude correlation based metric is employed in the analysis. This correlation measure may be used across the full measured frequency range and any complex response can be

M. Aguirre (✉) • I. Urresti • F. Martinez • G. Fernandez
IK4-Ikerlan Technology Research Centre, Mechanics Area, P J.M. Arizmendiarrieta, 2, 20500, Mondragón, Spain
e-mail: maguirre@ikerlan.es; iurresti@ikerlan.es; Felix.Martinez@ikerlan.es; gfernandez@ikerlan.es

S. Cogan
FEMSTO-ST, Département de Mécanique Appliquée, Université de Franche-Comté, 24 Chemin de l'Epitaphe, Besançon 25000, France
e-mail: scott.cogan@univ-fcomte.fr

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_28

summarized with an error value. Moreover, the importance of giving suitable damping properties is highlighted in this paper. An anisotropic damping model is developed and updated in order to obtain adequate FRF amplitudes. This innovative form of applying different damping values at different directions allows to better correlate the experimental data.

Therefore, the main objectives of this paper are: (a) implement a weakly coupled multiphysical model where electromagnetic, structural vibrational and acoustic aspects are taken into account, (b) perform an experimental analysis in the so called IkerMAQ electrical machine both in free and operational conditions, (c) update mass and stiffness of the model to correlate better the peak frequency values, (d) develop the anisotropic damping model to give directional dependent damping values, (e) perform a sensitivity analysis to analyze the influence of the different parameters and (f) adjust parameter values to improve experimental correlation.

28.2 Multiphysical Model

To model the PMSM from a global point of view, a multiphysical design methodology is developed where elements of electromagnetic finite element calculations, dynamic finite element calculations and acoustic finite element calculations are combined. To implement this calculation methodology, the ANSYS WORKBENCH software [8] for finite element modeling is used. Figure 28.1 shows the three main groups composing the model. In the first group a 2D electromagnetic Maxwell model is defined where electromagnetic forces are calculated. These results are then inserted into the dynamic model where modal and harmonic simulations are performed. The last step consists in the acoustic simulation. This paper focuses on the second group where the vibrational behavior of the electrical machine is simulated and updated.

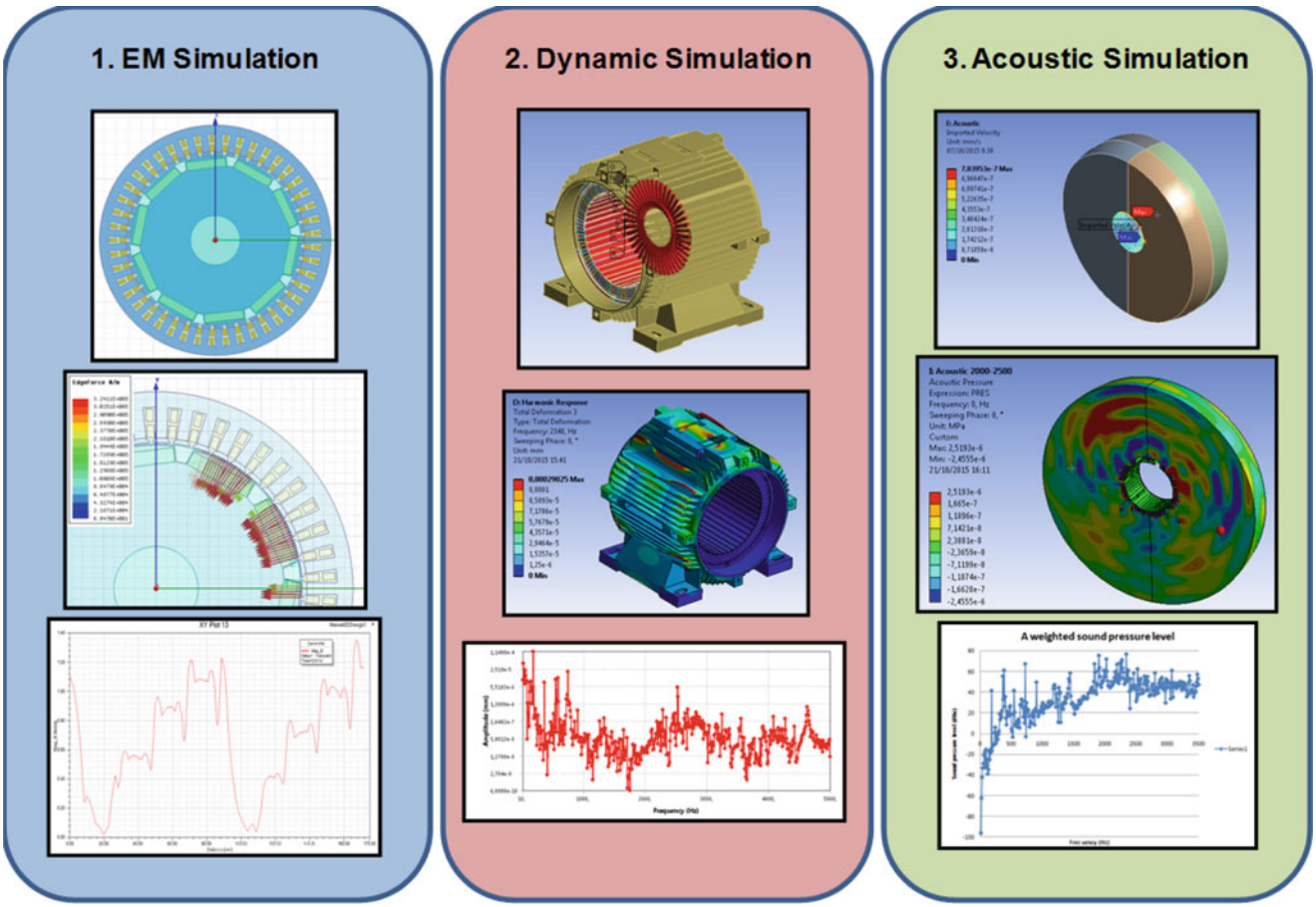


Fig. 28.1 Multiphysical model for PMSM complete modeling

28.2.1 Vibrational Model

Among all components of the electrical machine, the stator is the major responsible of the acoustic noise; this is why the vibrational model analysis is focused on this component. The model includes the housing and the stator core with coils. The material of each element is properly selected; therefore realistic orthotropic behavior of the structure is taken into account. This fact affects the dynamic behavior of the structure, so the damping properties are appropriately defined as presented in Sect. 28.5. Solid elements (solid 186) have been used to build up this model with a total number of 193,225 nodes and 68,437 elements. Harmonic analysis is performed taking into account the electromagnetic forces calculated in the Maxwell model.

28.3 Experimental Campaign

For the experimental campaign, a PMSM (Permanent Magnet Synchronous Machine) was fully designed in IK4-Ikerlan. After an extensive study, the final prototype, called IkerMAQ [9] is a machine with 45 slots and 5 pole pairs, this is, a Q45p5 machine. The nominal electrical output power is 75 kW with a nominal speed of 1080 rpm; this establishes a nominal torque of 700 N.m. The chosen magnets are N40H.

In the work described in this paper, a bottom-up approach is employed. First, single components are measured, correlated and adjusted. After that, the same procedure is applied to subassemblies with increasing number of components. Finally, the whole system is analyzed. For the study of the vibratory behavior of the structure, two different modal analyses are performed: Experimental Modal Analysis (EMA) and Operational Modal Analysis (OMA). EMA is carried out using a shock hammer to apply the external force, while OMA is done during IkerMAQ's normal operation. During the tests the Brüel & Kjær platform PULSE is used and piezoelectric triaxial accelerometers are placed at different locations.

The EMA's main objective is the identification of the modal shapes of the structure under test including the damping factors and resonance frequencies of each mode. Two stator EMAs were carried out during IkerMAQ manufacture process.

Once the test campaign is done, it is possible to identify the different experimental modal shapes with its resonant frequencies and the damping factors and correlate them with those simulated in FEM. All the studied modes are vibrating only in the radial direction and they have both, tangential and axial dependency. For this particular situation, the only modes of interest are the radial ones as they have a major contribution in noise emission.

As vibrational behavior and, hence noise, are going to be maximum at nominal values, a test campaign at this working conditions is performed with the PMSM in generator mode. Figure 28.2 shows the test configuration.

28.4 Baseline Model Definition

For a reliable model of the electrical machine, the first step consists on building a suitable FE baseline model. If the frequency values of the radial forces are close to any eigenfrequency of the stator, the stator could vibrate in resonance propitiating deformations, vibrations and acoustic noise. Therefore, the first step for the complete modelization of the electrical machine

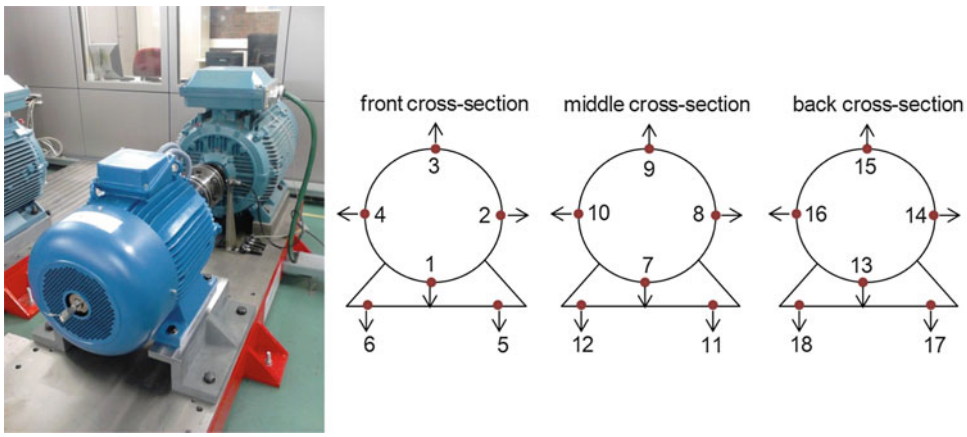


Fig. 28.2 Operational modal test in the IkerMAQ machine with accelerometers position and direction schedule

is the prediction of the natural frequencies and mode shapes by a modal analysis. For this reason, a model for vibroacoustic simulation in the frequency domain is developed. The updating of the mass, stiffness and damping values are performed taking into account EMA results.

The winding and core are the two elements which have more contribution on the weight of the stator. Also, considering that the stator has an axis-symmetrical geometry, it is considered that the mass has a slight and homogeneous influence on the modal frequencies and shapes. Knowing the total mass of the stator (available at IK4 Ikerlan) and taking into account the different elements it is composed with, the densities of the different parts of the model are defined.

The second step is the adjustment of stiffness in the model. For that purpose, first a sensitivity analysis is performed to adjust the materials elastic properties. Sensitivity analysis is the systematic investigation of the model responses to quantitative (input parameters) or qualitative (structure, etc) factors of disturbance. Once the most influential parameters are selected, an optimization process is performed via FEMTools [10] and an adjusted model is obtained.

The natural frequencies of the model depend on the mass and stiffness of the materials. However, the amplitude of vibration, and thus the noise emitted at each frequency, depends largely on the damping values introduced to the model. Therefore, there exists a need to design a model adjusted both in frequency and in amplitude to be able to predict the vibrational behavior and thus, the noise.

The updating procedure of the damping is not as straightforward as with mass and stiffness values. Figure 28.3 shows an example of the FRF curves obtained both in the experimental campaign (red line) and in simulation (blue line) at a specific point. The simulated FRF has no damping; therefore, there are sharp frequency peaks. However, analyzing the experimental curve, it is possible to see that there exist some peaks with non-negligible damped values. These peaks are related through the longitudinal direction of the stator; the resin that there is placed between the different sections makes to increase the damping in this direction. For that reason, damping is directional dependent and this fact is mandatory to take it into account.

Ansys has an option to insert a global damping to the model called constant damping ratio [11]. This is the simplest way of specifying damping in a structure. It represents the ratio of actual damping to critical damping. However, this parameter applies the damping globally in the model and thus, its suitability is limited. Therefore it seems evident the necessity of defining an anisotropic damping that can give different damping values in the different directions.

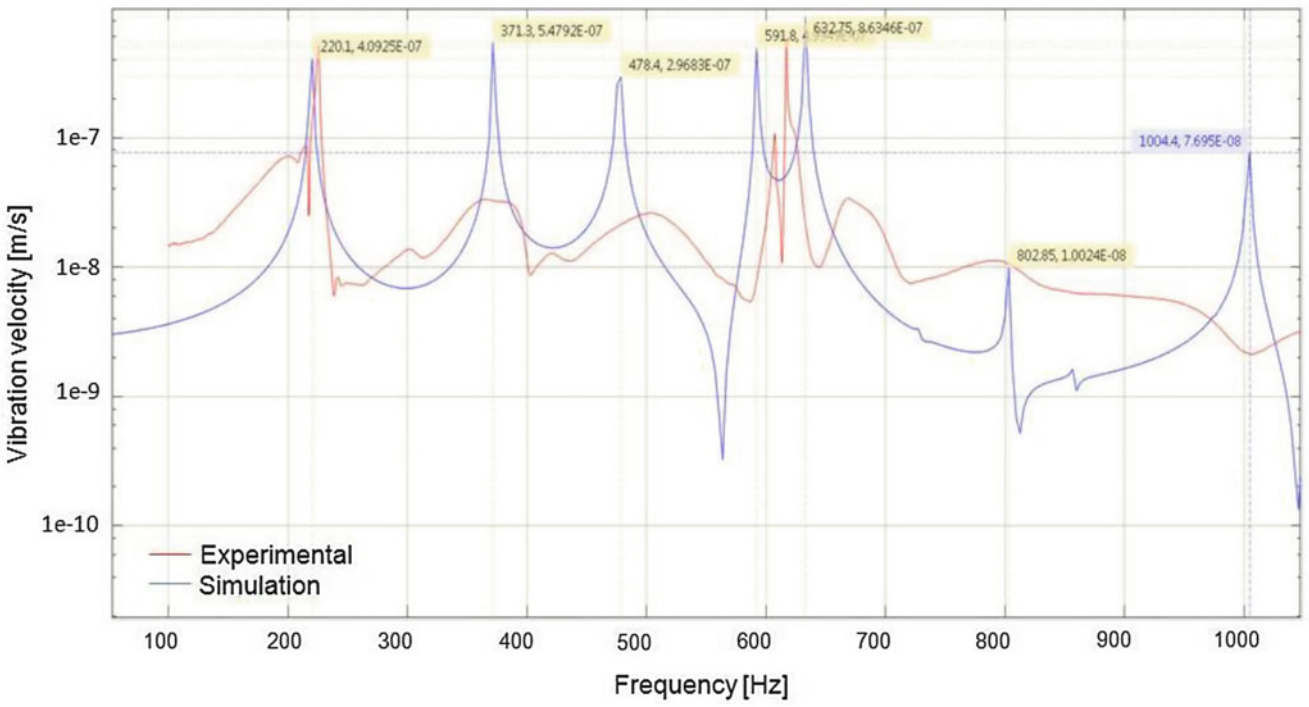


Fig. 28.3 Experimental and numerical (without damping) FRF curves from 0 to 1000 Hz

28.5 Anisotropic Damping

The most appropriate solution to have different damping values at different directions is to develop a FE model with an anisotropic damping. For that purpose, in parallel to the original model, a virtual material is created [12]. These two parallel materials are connected at every node.

$$[K]\{x\} + [C]\{\dot{x}\} + [M]\{\ddot{x}\} = \{0\} \tag{28.1}$$

$$[K]'\{x\} + [C]'\{\dot{x}\} + [M]'\{\ddot{x}\} = \{0\} \tag{28.2}$$

where $[K]$ is the damping matrix, $[C]$ the stiffness matrix and $[M]$ the mass matrix of the original model, $[K]$', [C]', $[M]$' are the damping, stiffness and mass matrices respectively of the virtual material and $\{x\}, \{\dot{x}\}, \{\ddot{x}\}$ are the displacement, velocity and acceleration vectors respectively.

The mass and stiffness matrices of the first model are already updated as described in the previous section. Therefore, the mass and stiffness values of the virtual material need to be negligible, so: $[K]' << [K]$ and $[M]' << [M]$.

On the other hand, the first material has a negligible damping, while the damping of the virtual material is defined by the stiffness and mass damping matrices:

$$[C] = [0] \tag{28.3}$$

$$[C]' = \beta'[K]' + \alpha'[M]' \tag{28.4}$$

As directional dependence is needed, stiffness damping is only used, $\alpha' = 0$, thus:

$$[K]\{x\} + \beta'[K]'\{\dot{x}\} + [M]\{\ddot{x}\} = \{0\} \tag{28.5}$$

β' is defined as a constant number while in $[K]$' there are 4 independent parameters that need to be adjusted: a1 defines the Young's Modulus in x direction ($Ex = Ey$), a2 is the Young's Modulus in z direction (Ez), a3 refers to the shear modulus in xy direction (Gxy) and a4 is the shear modulus in xz direction ($Gyz = Gxz$). Taking into account the mode shapes from 0 to 1000 Hz from EMA test campaign, an updating procedure of the parameters β' and $[K]$' is performed. This time, the Frequency Response Function obtained in the simulation is closer to the experimental data than before (see Fig. 28.4).

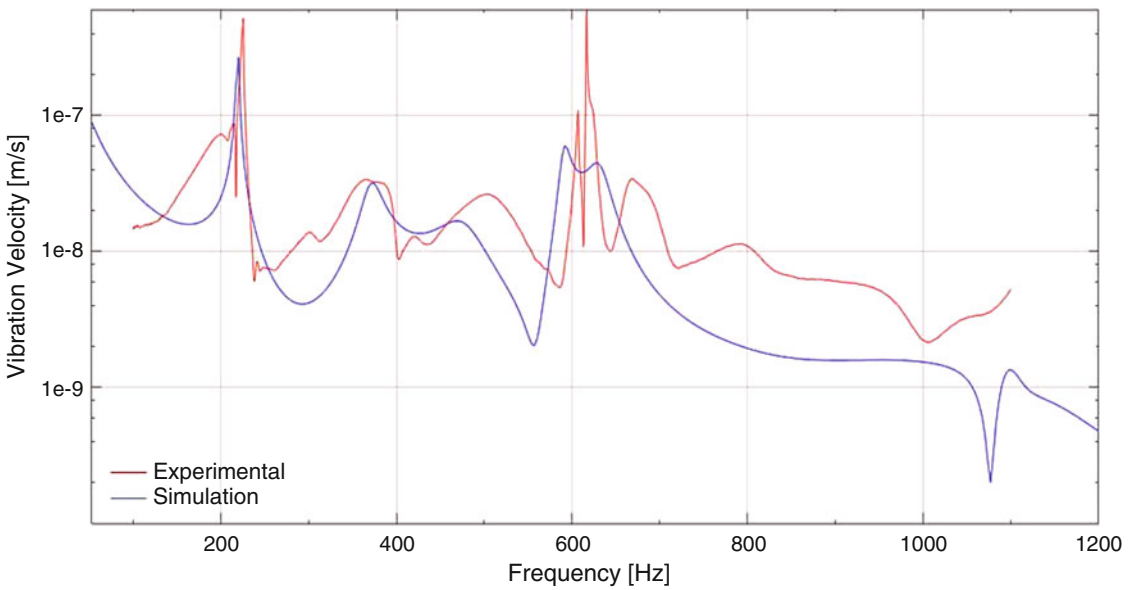


Fig. 28.4 Experimental and damping updated numerical FRF curves from 0 to 1000 Hz

It can be concluded that when the anisotropic damping is defined, the correlation between the simulated and experimental data is better. Next step consists on the correlation and posterior updating analysis between the OMA results and the simulated model.

28.6 Operational Correlation and Updating Analysis

To obtain a reliable FE model of the electrical machine compared to OMA results, different steps need to be followed. First, it is mandatory to define a valuable metric. This metric will give a quantitatively measure of the difference between the simulated and the experimental data. Three different metrics are described in this paper and the most suitable one is chosen for this specific case. After that a sensitivity analysis is carried out to identify the most influential parameters to finally improve the correlation between the simulated and the experimental data. As experimental data, the response in the radial direction of 18 accelerometers located at different places of the machine is chosen (see Fig. 28.2).

28.6.1 Frequency Response Calibration Metrics

This section describes three different metrics for frequency responses. They are described and the most suitable one is chosen for this specific case.

28.6.1.1 Frequency Response Assurance Criterion (FRAC)

Two frequency response functions, simulation (A) and experimental (X), representing the same input-output relationship, can be compared using a technique known as the Frequency Response Assurance Criterion (FRAC) which was proposed by Heylen and Lammens [13]. The basic assumption is that the measured frequency response function and the synthesized frequency response function should be linearly related (unity scaling coefficient) at all frequencies. The FRFs can be compared over the full or partial frequency range of the FRFs as long as the same discrete frequencies are used in the comparison. This procedure is particularly effective when the shape comparison of the FRFs is the major issue and the amplitude correlation is left aside.

Instead of handling the 18 FRF signals individually, which results interpretation can be of high complexity, the 75 percentile is calculated for all signals frequency by frequency and a unique FRF signal, ${H_{75}}^X(\omega)$, is then obtained and evaluated. The same procedure is applied to the simulation results obtaining one FRF signal, ${H_{75}}^A(\omega)$. It should be noted that the frequency spectrum ω goes from 0 to 5000 Hz with a resolution of 10 Hz.

$$FRAC = \frac{\sum_\omega \left| {H_{75}}^X(\omega) \cdot \overline{{H_{75}}^A(\omega)} \right|^2}{\left(\sum_\omega {H_{75}}^X(\omega) \cdot \overline{{H_{75}}^X(\omega)} \right) \cdot \left(\sum_\omega {H_{75}}^A(\omega) \cdot \overline{{H_{75}}^A(\omega)} \right)} \tag{28.6}$$

For identical FRFs, the FRAC value is unity and zero if the responses are uncorrelated. The authors, however, point out that a global shift in frequency between the experimental (X) and analytical (A) FRFs leads to a biased correlation value even if the FRFs are otherwise identical.

28.6.1.2 Square Deviation (SD)

This metric instead of using the 75 percentile of the signals, all signals are arranged in one vector one after another, such as:

$$H^X(\omega) = \begin{bmatrix} {H_1}^X(\omega) \\ {H_2}^X(\omega) \\ \vdots \\ {H_{18}}^X(\omega) \end{bmatrix} \tag{28.7}$$

The same procedure is applied to the signals from FE calculations. This metric [6] does not discriminate against deviations at frequencies where the structural response is small is the quadratic functional:

$$\delta = \frac{\sum_{\omega} \overline{\varepsilon(\omega)} \cdot \varepsilon(\omega)}{N} \tag{28.8}$$

where N is the number of elements of that vector. The deviation vector ε is computed for each frequency ω as:

$$\varepsilon(\omega) = \log_{10}\left(H^A(\omega)/H^X(\omega)\right) \tag{28.9}$$

28.6.1.3 Mean Squared Error (MSE)

The third metric used is the Mean Squared Error (MSE). This time again, percentile 75 of the signals are calculated; ${H_{75}}^X(\omega)$ for experimental and ${H_{75}}^A(\omega)$ for simulation.

$$MSE = \frac{\sum_{\omega} \left|{H_{75}}^A(\omega) - {H_{75}}^X(\omega)\right|^2}{\sum_{\omega} \left|{H_{75}}^X(\omega)\right|^2} \tag{28.10}$$

28.6.1.4 Correlation Metric Selection

The FRAC metric (Sect. 28.6.1.1) is mainly concerned to the shape comparison of the FRFs, while the SD metric (Sect. 28.6.1.2) gives more relevance to the relation between amplitudes of the FRFs. However, there is still a drawback in this metric. The correlation between the resonant peaks needs to be more important than the correlation between the anti-resonant peaks as the priority of this study is the correlation of the vibration that most influence in acoustic. This is why the Mean Squared Error (MSE) metric is the best option in this case. As the differences between the experimental and the simulated values are squared, an error in a peak will be more important than an error in the valleys.

28.6.2 Construction Parameters and Boundary Conditions

Before proceeding with the sensitivity analysis itself, it is mandatory to fix some parameters first. These are the eccentricity of the rotor, the boundary conditions and the introduction of the covers to the stator. The influence of these three design parameters has been studied first. Regarding eccentricity of the rotor, it is possible to introduce this effect in the electromagnetic calculation part of the analysis (see Fig. 28.1). Different eccentricity values have been applied and the value that gives the minor error MSE value has been chosen as reference for further research. Eccentricity values of 0.03 mm, 0.05 mm, 0.07 mm, 0.1 mm, 0.2 mm, 0.4 mm and 0.8 mm corresponding to 1.3%, 2.2%, 3.0%, 4.3%, 8.7%, 17.4% and 34.8% of the air gap have been calculated and 0.05 mm eccentricity has been chosen. So as to boundary conditions, in the experimental set-up the machine is bonded to the ground. However, this joint is not very stiff. To evaluate this joint, the same calculation is performed twice, one using bonded conditions and another one in free-free conditions. MSE values are calculated and in overall, the errors are lower in the free-free configuration. Therefore, the following analysis is performed in free-free conditions. Last point to take into account is that in the experimental set-up the IkerMAQ machine has covers at both sides (see Fig. 28.2). In order to see if this variable affect to the responses, the same calculation is done, one with covers and another one without them. Analyzing MSE values of the different points, it can be concluded that there is no much difference in using them or not. Thus, in order to simplify the FE model, the following calculations are performed without covers.

28.6.3 Sensitivity Analysis of Damping Parameters

Once the FE baseline model is completely defined, it is possible to perform a sensitivity analysis taking into account the different damping parameters as input parameters. In this paper, a One-Factor-at-a-Time (OFAT) approach has been employed

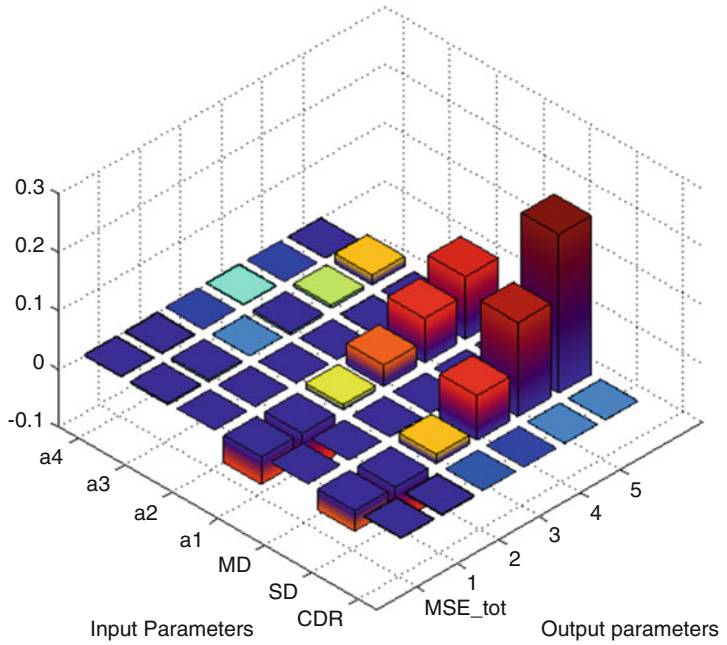


Fig. 28.5 Input parameter influence in the response

to analyze the effect of one input parameter at a time. In total, apart from the baseline model calculation, seven different calculations are performed as there are seven different damping input parameters: the Constant Damping Ratio (CDR), the Stiffness Damping (SD), the Mass Damping (MD) and four parameters that defined the Anisotropic Damping (a1, a2, a3, a4). The reference values for the baseline model are obtained from the conclusions in Sect. 28.4.

Figure 28.5 shows the sensitivity analysis results. The results are given in normalized values to avoid unit problems. A normalized sensitivity shows the percentage change of the response value for one percent change of the parameter value. Normalized sensitivities are dimensionless and thus ideal to be used when comparing different combinations of responses and parameters. The *x* axis represents the different input parameters (CDR, SD, MD, a1, a2, a3, a4). The *y* axis represents the Mean Squared Error (MSE) split into several frequency ranges; the first one is the total error obtained from full range (0–5000 Hz), second one is the MSE value from 0–1000 Hz, the third one from 1000–2000 Hz and so on. Finally, the *z* axis represents the values of the errors. Analyzing the figure it is possible to conclude that the most influential parameters are the stiffness damping (SD) and the anisotropic damping a1 and a3. Therefore, these are the parameters that are going to use in the model calibration.

The sensitivity result shows that parameters SD, a1 and a3 have higher influence in output parameters than the rest. Thus, these three input parameters will be used in the next sections to improve FE model's correlation with the experimental measurements.

28.6.4 Surrogate Models

To obtain the most adequate values for the three input parameters defined in the previous section, this is, to perform an optimization procedure, it is necessary to build a metamodel or surrogate model. A surrogate model can be defined as an approximation model of the original model (in this case, a FE model), which is built from sampled data obtained by randomly probing the design space (called sampling via Design of Experiment (DoE)). Once the surrogate model is built, the optimization procedure is easier to execute as the computational cost associated with the search based on the surrogate model is generally negligible. A suitable approach to choose sample points in the design space is the Design of Experiments (DoE). In this case the Latin Hypercube Sampling procedure from Matlab [14] has been employed and a total number of 30 simulations have been carried out. There are several approaches to build a surrogate model. Maybe the most popular ones are polynomial response surfaces (RSM), kriging, support vector machines, space mapping and artificial neural networks. In this paper two approaches are used and compared; RSM and kriging.

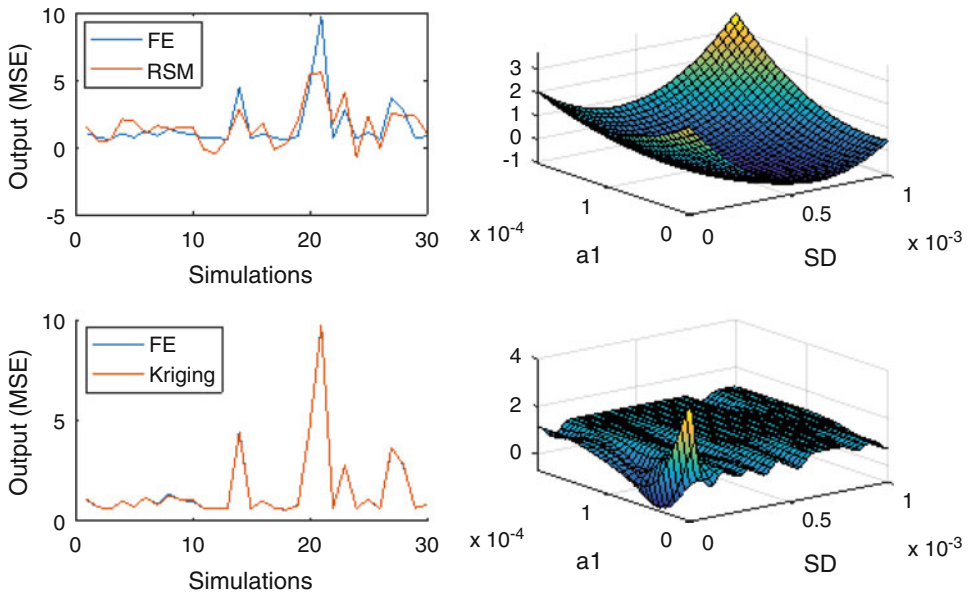


Fig. 28.6 Surrogate model precision with (**a**) RSM and (**c**) kriging, and the obtained response surface keeping the third parameter constant in its reference value for (**b**) RSM and (**d**) kriging

RSM [15] denotes a polynomial approximation model in which the sampled data is fitted by a least-square regression technique. In RSM-based optimization applications, the quadratic polynomial model usually provides the best compromise between the modeling accuracy and computational expense, when compared with the linear or higher order polynomial models. An advantage of RSM is that it can smooth out the various scales of numerical noise in the data while captures the global trend of the variation, which makes it very robust and thus well suited for optimization problems in engineering design. In this case, Fig. 28.6a shows that the generated surrogate model cannot approximate exactly the FE responses. Nevertheless, as a practical example, plotting two input parameters (SD and a1) and keeping the third one (a3) in its reference value, it is possible to see that smooth surface is obtained applying this methodology (see Fig. 28.6b).

Kriging [16] is an interpolating method which features the observed data at all sample points. This method provides a statistical prediction of an unknown function by minimizing its Mean Squared Error (MSE). It can be equivalent to any order of polynomials and is thus well suited for a highly-nonlinear functions with multiple extremums. This time, the surrogate model can resemble the FE responses perfectly (see Fig. 28.6b). However, looking at the response surface generated from two input parameters, keeping the third one constant, it is possible to see that this methodology sacrifices smooth responses in order to resemble the FE responses as much as possible (see Fig. 28.6d). This type of surrogate model generally gives local minimums and they are more complicated to use.

28.6.5 FRF Updating

Once surrogate models are built, they are employed in the optimization app from Matlab [14]. The objective of the optimization procedure is to find the minimum of the constrained nonlinear multivariable function, in order to minimize the MSE. Minimum and maximum bounds are defined to avoid having results outside the design space. Using the RSM surrogate model, regardless the starting point, a minimum is reached. These parameters values are then inserted into FE model in order to assess that the obtained results are in concordance with the FE results. For the case of the kriging surrogate model, as there are many local minimums, the starting point influences the obtained result. Therefore three different starting points are chosen and the optimization process gives three different results. These parameter values are then inserted into the FE model and simulation is run. Although the four simulation options give better results that the first baseline model, thus, the MSE value is reduced, there is one DOE simulation that has a lower error value. Therefore it seems that these 2 surrogate models are not accurate enough. In this paper, the DOE simulation that gives minimum error is selected. Figure 28.7 shows the obtained 75 percentile FRF signals for the baseline model and for the minimum MSE simulation case. These two curves

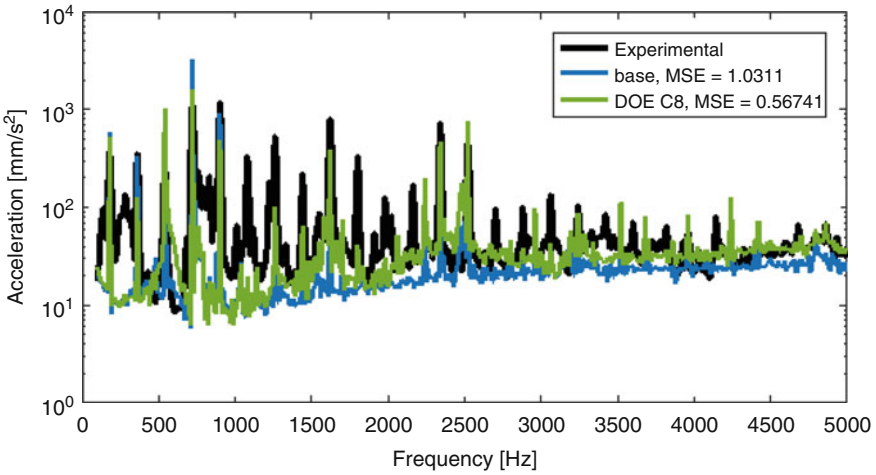


Fig. 28.7 FRF comparison between baseline model and minimum MSE simulation against experimental data

are compared against test data. It is possible to see that in the first half of the graph, this is, from 0–2500 Hz, the peak amplitudes are closer to the experimental ones for the DOE case simulation and this is which most influences on the obtained error value.

28.7 Conclusions

A weakly coupled multiphysical methodology for the calculation of the dynamic response of an electric motor stator has been presented in this paper. This method can also lead to the calculation of the acoustic power radiated by the motor. However, the first step, and the main objective of this paper, is to correlate the vibrational response of the machine. For that purpose, an electrical machine called IkerMAQ is built and a bottom-up approach is employed. First, single components are measured, correlated and adjusted. Finally, the whole system is analyzed testing it at nominal speed values.

As the electric machine is built with different materials with different properties, the damping values are directionally dependent. Therefore, the main contribution of this paper is to develop an anisotropic damping model that can apply different damping values at different directions. This model provides a better approximation to the available test data.

A complete sensitivity analysis is performed and the influence of the diverse simulation parameters is analyzed in this paper. Although the generated surrogate models are not capable of correctly representing the FE model, the objective of decreasing the MSE value is reached. Therefore, it can be concluded that the vibrational response of the electrical machine is correlated relatively well. Nevertheless, for future work the development of a suitable surrogate model is recommended. Once the vibrational model is calibrated, the next step consists in correlating the acoustic model.

To study the global vibrational behavior of the experimental and simulation cases, the Root Mean Square (RMS) values are calculated. In the case of the experimental 75 percentile signal, a RMS value of 130.0 mm/sˆ2 is obtained. On the other hand, the baseline 75 percentile signal gives a RMS of 164.1 mm/s^2 while the optimized one (minimum MSE signal) gives a RMS of 114.3 mm/s^2.

In summary, the overall suitability of the study carried out in this research, the errors between the experimental RMS and simulated RMS are calculated. While the baseline error is 26.2%, the optimized signal error is 12.1%, thus reducing the error considerably.

References

1. International Organization for Standardization (ISO), Acoustics-normal equal loudness-level contours, ISO 226:2003 -BS 3383. 2003
2. International Organization for Standardization (ISO), Mechanical vibration and shock – evaluation of human exposure to whole-body vibration. Part 1: general requirements, ISO 2631-1: 1997.

3. Rainer, S., Biro, O., Weilharter, B., Stermecki, A.: Weak coupling between electromagnetic and structural models for electrical machines. IEEE Trans. Magn. **46**, 2807–2810 (2010)
4. Pellerey, P., Lanfranchi, V., Friedrich, G.: Vibratory simulation tool for an electromagnetically excited non skewed electrical motor, case of the Wound Rotor Synchronous Machine. In: ELECTRIMACS 2011, Cergy-Pontoise, France (2011)
5. Dupont, J., Bouvet, P.: Noise radiated by an electrical powertrain: multiphysical simulation. In: 21eme Congres Francais de Mecanique, Bordeaux (2013)
6. Abrahamsson, T., Bartholdsson, F., Hallqvist, M., Olsson, K., Olsson, M., Sällström, Å.: Calibration and validation of a car subframe finite element model using frequency responses. In: Mains, M. (ed.) Topics in Modal Analysis, vol. 10, pp. 9–22. Springer (2015)
7. Grafe, H.: Model updating of large structural dynamics models using measured response functions. University of London, London (1999)
8. ANSYS® Academic Research, Release 16.2.
9. Rodríguez, A.L.: Development of a Multidisciplinary and Optimized Design Methodology for Surface Permanent Magnets Synchronous Machines. University of Santiago de Compostela (USC) (2015). http://hdl.handle.net/10347/14788
10. FEMtools Theoretical Manual, Version 3.8.2, Dynamic Design Solutions, Leuven, Belgium, (2015).
11. ANSYS® Academic Research, Release 16.2, Help System, Mechanical APDL, ANSYS, Inc.
12. Rydberg S., Prediction of vibrational amplitude in composite sandwich structures, Master's thesis in Applied Mechanics, Chalmers University of Technology, 2013. Available at: http://publications.lib.chalmers.se/records/fulltext/187196/187196.pdf Accessed 10 Sept 2016.
13. Heylen, W., Lammens, S.: FRAC: a consistent way of comparing frequency response functions. Proceedings of the Conference on Identification in Engineering Systems. 48–57 (1996)
14. MATLAB and Statistics Toolbox Release 2016a, The MathWorks, Inc., Natick, Massachusetts, United States.
15. Response surface methodology. Available at: http://www.brad.ac.uk/staff/vtoropov/burgeon/thesis_luis/chapter3.pdf Accessed 10 Sept 2016.
16. Gano S.E, Kim H, Brown D.E. II, Comparison of three surrogate modeling techniques: datascaper, kriging, and second order regression. In: 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 6–8 September 2006.

Chapter 29
A Comparison of Computer-Vision-Based Structural Dynamics Characterizations

Aral Sarrafi, Peyman Poozesh, and Zhu Mao

Abstract As a specific modern non-contact sensing technology, optical/video information is getting more and more attention employed to interpret structural responses and system status awareness. By means of processing the acquired video, a full-field system information is available which may be applied later to Experimental Modal Analysis (EMA), Structural Health Monitoring (SHM), System Identification (SI), etc., while at the same time, there is no influence to the structural testing such as mass loading and stiffness change. There are numerous technologies to extract the dynamic response of structures from acquired videos. In this paper, several point tracking algorithms are particularly compared, including Lucas-Kanade tracker, Hungarian registration algorithm and particle filter. These computer vision algorithms are implemented to extract the natural frequencies of a lab-scale structure, and the efficiency of each method is investigated regarding the consistency in estimating the natural frequencies and computational time. The recorded video contains external noise caused by lighting change during the experiment, as well as the intrinsic uncertainty on the photosensitive devices. Therefore, the natural frequencies estimated via different algorithms will have different values. An overall comparison between several computer vision algorithms are made in this paper in terms of precision, and computational load.

29.1 Introduction

Structural dynamics characterization has significant importance in several fields of research and engineering applications including aerospace, civil and mechanical structures [1]. Experimental modal analysis (EMA) and operational modal analysis (OMA) are the most well developed tools being utilized for structural dynamics identification purposes [2]. The experimentally identified dynamics could be used later for model validation, model updating [3] or structural health monitoring applications. The conventional EMA and OMA requires contacts sensor systems such as accelerometers to be attached to the structures to record the response of the structure due to the external loads [4]. Although accelerometers are precise and reliable sensor systems with broad frequency range, instrumentation of large scale structures with accelerometers is a tedious and labor intensive task [5]. Moreover, mass loading effects induced by mounted accelerometers on the light weight structures can change the true dynamics of the subjected structure dramatically [6]. Therefore, utilizing accelerometers and contact sensors are not the best choice for all modal testing case studies, and other alternative sensor systems should be considered as options. Laser vibrometer can be an excellent measurement equipment since the measurements are non-contact and mass loading effects are avoided, and the instrumentation of the structure and wiring the accelerometers are also by passed. However, the laser vibrometer is only capable of making measurements sequentially which often takes a long time, especially when high spatial resolution is needed, and the number of measurement points are relatively large.

Recent advances in production and design of high speed and high resolution digital cameras has introduced a new measurement technique referred to as optical measurement systems. Optical measurement system is a new sensor system that combines the sequence of images captured from a mechanical system with advanced computer vision algorithms to extract worthwhile information [7]. Optical measurement systems as a non-contact sensing method are able to make the measurements without inducing any changes to the structures. Furthermore, unlike laser vibrometer all the full field measurements are made simultaneously which decrease the testing time dramatically [8]. 3D Digital image correlation (3D DIC) and 3D point tracking (3DPT) are considered as the most reliable optical measurement techniques which has been commercialized, and are being used widely in industrial and research applications [9–13]. Although recording the

A. Sarrafi • P. Poozesh • Z. Mao (✉)
Structural Dynamics and Acoustic Systems Laboratory, Department of Mechanical Engineering, University of Massachusetts Lowell, One University Avenue, Lowell, MA 01854, USA
e-mail: zhu_mao@uml.edu

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_29

images from a scene of interest is a fairly straightforward task, extracting quantities such as displacement that could be used by engineers is computationally intensive, and it often takes relatively long time to process the sequence of images [5]. The efficiency of any signal processing technique has significant importance for engineering applications to save time and accelerate the decision making procedure [14]. The main objective of this paper is to investigate the accuracy and efficiency of other computer vision algorithms which could be beneficial to apply for modal analysis and structural health monitoring applications. In this perspective, several point tracking algorithms that have been developed in computer vision are utilized to extract the natural frequencies of a cantilever beam. Iterative Lucas-Kanade, Hungarian registration algorithm and particle filter point tracking methods are studied within this context.

In the first section of the paper the experimental test-bed is explained which consists of a cantilever beam and a high speed camera. The cantilever beam has been selected as the case study because of its simplicity and enables the discussion to stay focused on computer vision algorithms rather than structural modeling and testing complications. Followed by a brief literature review of motion estimation in computer vision and the theoretical background on Lucas-Kanade optical flow estimation and point tracking in Sect. 29.3, natural frequencies of the subjected structure are extracted using the iterative Lucas-Kanade point tracker. Overview of Hungarian registration algorithm for point tracking are discussed in Sect. 29.4. Section 29.5 is focused on the theoretical background and the results for particle filter point tracking method. The results provided in this paper reveal the capabilities of other computer vision point tracking methods in structural dynamics identification. It is also worth mentioning that these methods should be considered as research stage methods and could not be compared to commercialized software packages available in the market for 3D DIC or 3D point tracking in terms of accuracy and reliability. In general, for structural health monitoring applications the accuracy and precision provided by 3D DIC and 3D point tracking is not always necessary, and any reasonable measurement that can be post-processed to narrow down the decision making regarding the status of the structure would be enough. Therefore, other motion estimation techniques with lower computational costs presented in this paper could be an excellent candidate.

29.2 Experimental Test-Setup

In order to apply the computer vision algorithms for structural dynamics identification, a lab scale test setup has been prepared. A contrived cantilever beam is selected as the case study, since the dynamics of the cantilever beam is well defined, and there is no need for further modeling and validation. The sequence of images from the vibrating cantilever beam is captured by a CMOS high speed camera shown in Fig. 29.1. The sequence of images was recorded with a sampling rate of 2500 frames per second, and the Nyquist criteria indicates that the natural frequencies and operating mode shapes can be extracted up to 1250 Hz without facing any aliasing. Since there is no anti-aliasing filtering included in optical measurements the aliased higher order frequencies can be detected often, and the post-processing of the data should be handled more carefully.

29.3 Iterative Lucas-Kanade Optical Flow Estimation and Point Tracking

Optical flow methods could be considered as one of the earliest approaches in computer vision for motion estimation [15]. A large group of optical flow estimation methods including Lucas-Kanade are based on brightness consistency assumption and the fact that the motion in two consecutive frames are relatively small and Taylor series expansion could be used to obtain a linear sets of equations which are straightforward to solve with digital computers [15]. Since Lucas-Kanade trackers are based on Lucas-Kanade optical flow estimation, a brief discussion on the theoretical background of Lucas-Kanade optical flow estimation is presented in this section. Assuming that the image intensity for the frame at time t is $I(x, y, t)$ and $I(x + \Delta x, y + \Delta y, t + \Delta t)$ at time $t + \Delta t$, in order to satisfy the assumption of the brightness consistency the Eq. (29.1) should be true.

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t). \tag{29.1}$$

Assuming that the motion between two consecutive frames are small the Taylor series expansion could be utilized to expand the right side of the Eq. (29.1) as follows.

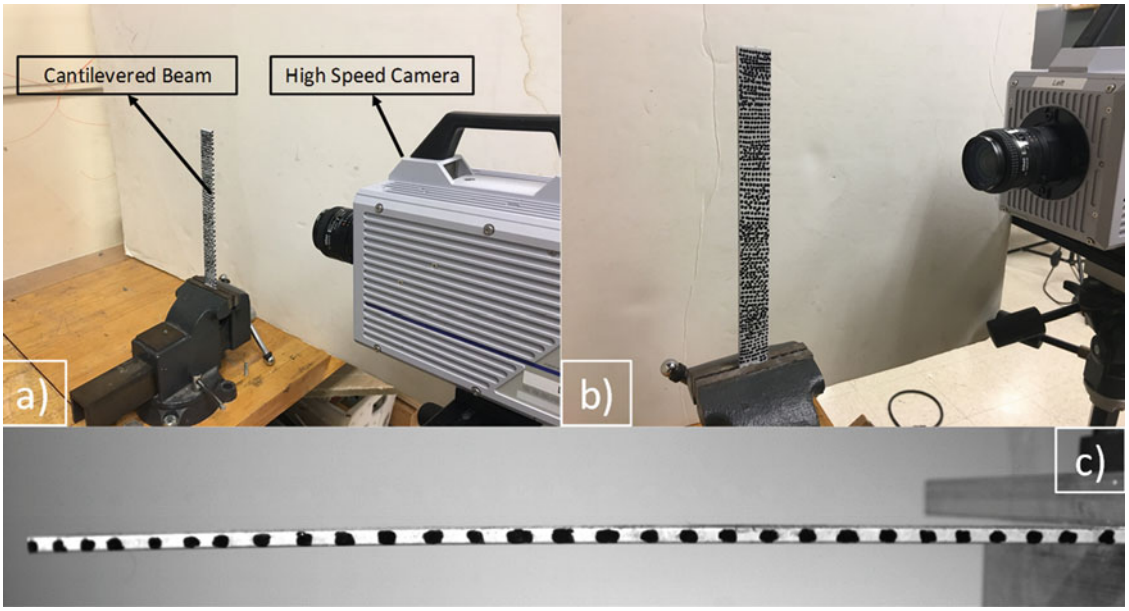


Fig. 29.1 (**a**, **b**) The CMOS high speed camera and the cantilever beam, (**c**) field of view for the camera capturing the sequence of images of the vibrating beam

$$I\left(x+\Delta x, y+\Delta y, t+\Delta t\right)=I\left(x, y, t\right)+\frac{\partial I}{\partial x}\Delta x+\frac{\partial I}{\partial y}\Delta y+\frac{\partial I}{\partial t}\Delta t+H.O.T. \tag{29.2}$$

Substituting (29.2) into (29.1) and dividing by Δt will result the flowing equations for each pixel which is known as the optical flow general equation or referred to as the aperture problem in computer vision community.

$$\frac{\partial I}{\partial x}\frac{\Delta x}{\Delta t}+\frac{\partial I}{\partial y}\frac{\Delta y}{\Delta t}+\frac{\partial I}{\partial t}\frac{\Delta t}{\Delta t}=0. \tag{29.3}$$

It is more conventional to write Eq. (29.3) in the following format in which V_x and V_y are the components of optical flow vectors at each pixel.

$$\frac{\partial I}{\partial x}V_x+\frac{\partial I}{\partial y}V_y+I_t=0. \tag{29.4}$$

As it is clear there is only one governing equation at each pixel and two unknowns V_x and V_y, as a result there is no unique solution for the aperture problem to estimate the optical flow components. Several approximate methods have been introduced in order to obtain the optical flow vector field including Lucas-Kanade [16] and Horn-Schunck [17]. In this paper iterative Lucas-Kanade method has been used to solve the aperture problem to extract the motion field between two consecutive frames of a sequence of images. The underlying assumption in Lucas-Kanade method is that the nearby pixels have equal optical flow vector components. In this case additional constraints are introduced to the problem and made the aperture problem an over constrained set of equations. This over constraint set of equations are then solved using the least squares algorithm. After estimation of the optical flow vectors at each pixel, interest points are selected on the first frame of the captured video from the structure using Harris corner detection algorithm. The point tracking is performed by moving the interest points with respect to the estimated motion filed provided by Lucas-Kanade algorithm frame by frame. As a result, the interest points will follow the optical flow vectors from one frame to another and point tracking is performed frame by frame. In order to make the results of the Lucas-Kanade algorithm more reliable and robust to noise, disturbances and also to be able to handle large motions in the scene, the algorithm is modified to work in an iterative fashion. The theoretical background of performing iterative Lucas-Kanade is beyond the scope of this paper and the details for the method could be found in references.

Figure 29.2 shows three snap shots of the Lucas-Kanade point tracking performance. The points are following the motion pattern as expected. Although the tracking is not performed perfectly, the natural frequencies of the cantilever beam can

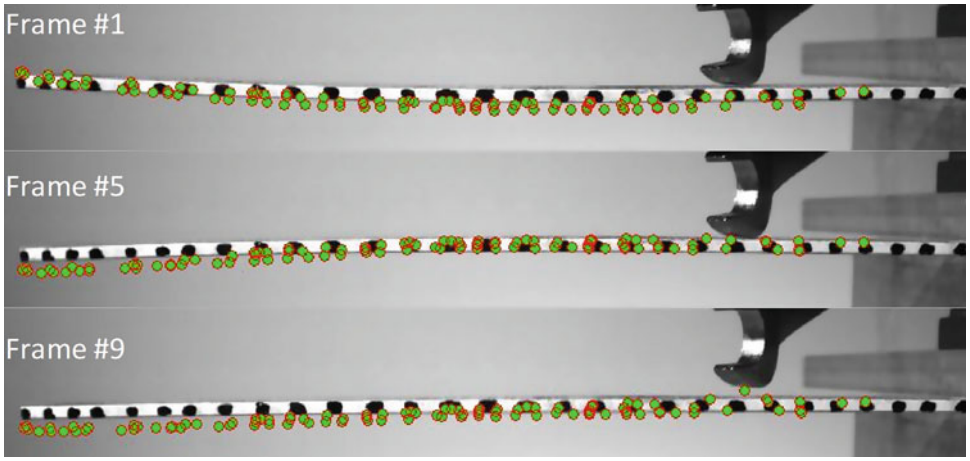


Fig. 29.2 Three snap shots of the Lucas-Kanade point tracking on the cantilever beam

Table 29.1 First four natural frequencies of the beam estimated by different computer vision algorithms

Mode number	Lucas-Kanade (Hz)	Hungarian registration (Hz)	Particle filter (Hz)
1	36.62	35.32	35.55
2	222.36	221.44	219.78
3	645.99	635.99	637.56
4	720	690.25	694.39

be extracted since the tracking pattern matches with true displacements on the beam. In summary Lucas-Kanade algorithm is computationally intensive since the motion filed needs to be calculated at each pixel. It is also sensitive to noise and external disturbances such as lighting variations, since the coefficients of aperture problem $(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, I_t)$ are extracted by means of applying 2D discrete wavelets on image intensity $I(x, y, t)$ which causes the noise to propagates through the analysis [18, 19]. Therefore, the Lukas-Kanade is not the first choice for structural dynamics identification in comparison with other point tracking and computer vision techniques, but it is able to extract the natural frequencies with in an acceptable range of uncertainty Table 29.1.

29.4 Hungarian Registration Algorithm

In previous section the point tracking was performed by estimating the optical flow in sequence of images, and the interest points were being moved with respect to the motion vectors. Unlike Lucas-Kanade point tracker Hungarian registration algorithm does not require the computation of the motion field which is a computationally intensive task. In order to apply Hungarian registration algorithm points with specific criteria are found by means of fundamental image processing algorithms, this approach is quite similar to the procedure which is followed by conventional point tracking algorithms. In this case study the black dots on the beam are considered as the specific points and by tuning the parameters computers are able to detect them at each frame. Detecting the points at each frame does not mean that the point tracking is performed since the point are not labeled yet, and further analysis should be considered to find out that which point at each frame corresponds to which point at another frame. Once the correspondence between the points are found the point tracking problem is solved completely. In order to find the correspondence between the points from one frame to another Hungarian registration algorithm [20] is employed. The underlying assumption in this analysis is that since the videos are captured with high-speed cameras at 2500 frames per second (fps) and the subjected scene is a vibrating beam the displacements of the points in between two consecutive frames are relatively small. Therefore, by minimizing the Euclidian distance between the sets of point in two consecutive frames the correspondence between the points could be found easily. In other words, two points in two consecutive frames with minimum Euclidian distance corresponds to each other and should have the same label. This procedure should be followed from the first frame of the sequence to the last frame till all the correspondence information is recovered.

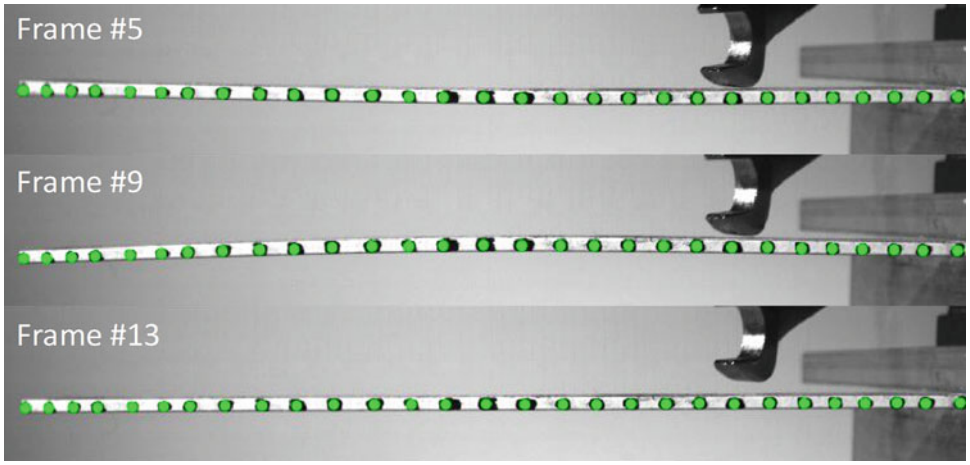


Fig. 29.3 Three snap shots of the Hungarian registration point tracking algorithm on the cantilever beam

Three snap shots of the performance of the point detection combined with Hungarian registration algorithm is illustrated in Fig. 29.3. Since this method is based on finding specific features in the sequence of images in this case the high contrast black dots, the tracking is performed with high precision. This method is similar to traditional point tracking algorithms, and the Hungarian registration is combined with it in order to find the correspondence between the detected points in two consecutive frames. Natural frequencies of the vibrating beam estimated by this method are in Table 29.1.

29.5 Particle Filters for Point Tracking

The particle filters are one of the most general and practical tools in estimation theory, and have numerous applications in all aspect of engineering especially in estimating the states of the dynamic systems such as navigation systems, robotic systems and smart structures [21]. In computer vision particle filters could be used for point tracking applications [22]. Unlike Kalman filters, particle filters could be used to estimate the states of a non-linear dynamic system with non-Gaussian noise while conventional Kalman filters are only able to handle the state estimation for linear systems with Gaussian distributed process and measurement noise. Extensions of Kalman filters such as Unscented Kalman Filter (UKF) and Extended Kalman Filter (EKF) are able to estimate the states of a non-linear system, but they are still limited by Gaussian noise. Since the scope of this paper is focused on the applications of the computer vision algorithms in structural dynamics presenting the details regarding the performance and formulation of the particle filters is avoided, and only key points of this method is highlighted. Particle filters estimate the states of a system; in this case the states are defined as the location of the interest points. Particle filters use a prediction dynamic model to predict the states in the next time step and then updates the estimation using the measurement information via a likelihood distribution. In this paper constant speed motion model is assumed between two consecutive frames representing the dynamic model which is used by particle filter to predict the states in the next frame, as shown in the following equation.

$$\begin{Bmatrix} x_n \\ y_n \\ \dot{x}_n \\ \dot{y}_n \end{Bmatrix} = \begin{pmatrix} 1\,0\,1\,0 \\ 0\,1\,0\,1 \\ 0\,0\,1\,0 \\ 0\,0\,0\,1 \end{pmatrix} \begin{Bmatrix} x_{n-1} \\ y_{n-1} \\ \dot{x}_{n-1} \\ \dot{y}_{n-1} \end{Bmatrix} + \begin{Bmatrix} n_x \\ n_y \\ \dot{n}_x \\ \dot{n}_y \end{Bmatrix}. \tag{29.5}$$

x_n, y_n are the location of the targets and $\dot{x}_n, \dot{y}_n$ are the speed of the target at frame n. The likelihood distribution function is selected based on the Euclidian distance between the RGB values of the target and the rest of the pixels in the scene.

$$d = \sqrt{(r - r_t)^2 + (g - g_t)^2 + (b - b_t)^2}. \tag{29.6}$$

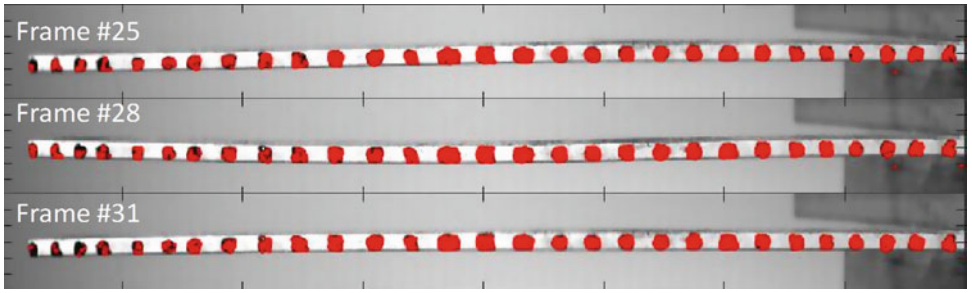


Fig. 29.4 Three snap shots of the particle filter point tracking algorithm on the cantilever beam

Table 29.2 Computational time for different point tracking algorithms

Algorithm	Computation time (min)
Lucas-Kanade	37
Hungarian registration	14
Particle filter	9

d is the Euclidian distance between the pixel RGB values and target RGB in which r, g, b are the RGB values of the pixels in the scene and r_t, g_t, b_t are the RGB values of the specific targets that are desired to be tracked in the scene.

$$P\left(Y_n|X^k{}_{n|n-1}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{d^2}{2\sigma^2}\right). \tag{29.7}$$

The likelihood distribution is selected as in Eq. (29.7) to resample the predicted states and update the estimation. As RGB values of the pixels get closer to the RGB values of the target the likelihood of the particles become larger, and the algorithm updates the particles respectively which result in tracking of the targets.

Figure 29.4 shows different frames of the vibrating beam with particle filter point tracking. The red dots are clusters of numerous particles which are tracking the targets based on the RGB values that controls the likelihood distribution and the constant speed model which is used for prediction stage. Particle filters are efficient in terms of computational load if the number of particles are selected properly, and the natural frequencies can be estimated with an acceptable accuracy presented in Table 29.1.

The computational time for different algorithms are presented in Table 29.2 as mentioned before since in iterative Lucas-Kanade algorithm the motion field needs to be estimated at each pixel for all the frames the computation time for this method is higher than other algorithms. Hungarian registration algorithm which has a similar approach as the conventional 3D point tracking is ranked second in terms of computational load. The particle filter point tracker has the least computational time in this analysis, but the computational load may increase if more particles were selected for the analysis. It should be also considered that the MATLAB scripts used for the analysis has not been optimized in terms of efficient computation, and the computation time for the algorithms can be reduced by means of avoiding unnecessary variables and computational effort.

29.6 Conclusion

In this paper several computer vision algorithms are utilized for structural dynamics identification. The major goal of this context is to evaluate the capabilities of other point tracking methods developed in computer vision in estimating the natural frequencies of structures which may be later used for structural health monitoring applications. The performance of these algorithms were compared regarding the accuracy in estimating the natural frequencies, and computational effort. In real-time structural health monitoring applications minimizing the computational time is essential, and accuracy could be sacrificed to reach to a computationally effective algorithm. Conventional 3D DIC and 3D point tracking methods can provide very precise results, but the computational time will not allow them to be applied in real-time applications. Particle filter point tracking presented in this analysis has way less computational time, and it has a great potential to be applied for structural dynamics identification, and real time structural health monitoring. Moreover, using optimum number of particles and also optimizing the operating MATLAB script can decrease the computational effort. Iterative Lucas-Kanade point tracking is very sensitive to noise and light changing, and it also has higher computational load since estimating the optical flow vector

at each pixel is necessary for this method. Therefore, optical flow based approaches could not be the first choice. Hungarian registration algorithm is also a feature based point tracking method similar to traditional point tracking methods, so the time consuming motion field estimation is bypassed. This method has a precise performance, but in comparison with particle filters the result are showing that particle filters are more efficient regarding the computational effort. The prelaminar result presented in this paper are applied on to a simple cantilever beam, and in order to make solid conclusion regarding the choice of point tracking method the algorithms needs to be utilized on more complex structures in future.

References

1. Doebling, S.W., Farrar, C.R., Prime, M.B.: A summary review of vibration-based damage identification methods. Shock Vib. Digest. **30**, 91–105 (1998)
2. Ewins, D. J.: Modal Testing: Theory and Practice, vol. 15. Research Studies Press, Letchworth (1984)
3. Madarshahian, R., Caicedo, J. M., Sun, Z.: Direct inverse finite element model updating. In: 2012 Joint Conference of the Engineering Mechanics Institute and the 11th ASCE Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability, Notre Dame (2012)
4. Reynders, E.: System identification methods for (operational) modal analysis: review and comparison. Arch. Comput. Meth. Eng. **19**, 51–124 (2012)
5. Yang, Y., Dorn, C., Mancini, T., Talken, Z., Kenyon, G., Farrar, C., et al.: Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification. Mech. Syst. Signal Process. **85**, 567–590 (2017)
6. Dorn, C. J., Mancini, T. D., Talken, Z. R., Yang, Y., Kenyon, G., Farrar, C., et al.: Automated extraction of mode shapes using motion magnified video and blind source separation. In: Michael, M. (ed.) Topics in Modal Analysis & Testing, vol. 10, pp. 355–360. Springer (2016)
7. Chen, J.G., Wadhwa, N., Cha, Y.-J., Durand, F., Freeman, W.T., Buyukozturk, O.: Modal identification of simple structures with high-speed video using motion magnification. J. Sound Vib. **345**, 58–71 (2015)
8. Poozesh, P., Baqersad, J., Niezrecki, C., Avitabile, P., Harvey, E., Yarala, R.: Large-area photogrammetry based testing of wind turbine blades. Mech. Syst. Signal Process. **86**, 98–115 (2016)
9. Baqersad, J., Poozesh, P., Niezrecki, C., Avitabile, P.: Photogrammetry and optical methods in structural dynamics–a review. Mech. Syst. Signal Process. **86**, 17–34 (2016)
10. Poozesh, P., Baqersad, J., Niezrecki, C., Avitabile, P.: A multi-camera stereo DIC system for extracting operating mode shapes of large scale structures. In: Jin, H., Sciammarella, C., Yoshida, S., Lamberti, L. (eds.) Advancement of Optical Methods in Experimental Mechanics, vol. 3, pp. 225–238. Springer (2016)
11. Baqersad, J.: A non-contacting approach for full field dynamic strain monitoring of rotating structures using the photogrammetry, finite element, and modal expansion techniques (2015)
12. Baqersad, J., Poozesh, P., Niezrecki, C., Avitabile, P.: Full-field strain monitoring of a wind turbine using very limited set of displacements measured with three-dimensional point tracking. IN: ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, pp. V008T13A100–V008T13A100 (2015)
13. Baqersad, J., Poozesh, P., Niezrecki, C., Avitabile, P.: Extracting full-field dynamic strain response of a rotating wind turbine using photogrammetry. In: SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring, pp. 94371O–94371O-10 (2015)
14. Madarshahian, R., Estekanchi, H., Mahvashmohammadi, A.: Estimating seismic demand parameters using the endurance time method. J. Zhejiang Univ., Sci., A. **12**, 616–626 (2011)
15. Fleet, D., Weiss, Y.: Optical flow estimation. In: Paragios, N., Chen, Y., Faugeras, O.D. (eds.) Handbook of Mathematical Models in Computer Vision, pp. 237–257. Springer (2006)
16. Baker, S., Matthews, I.: Lucas-kanade 20 years on: a unifying framework. Int. J. Comput. Vis. **56**, 221–255 (2004)
17. Horn, B.K., Schunck, B.G.: Determining optical flow. Artif. Intell. **17**, 185–203 (1981)
18. Sarrafi, A., Mao, Z.: Probabilistic uncertainty quantification of wavelet-transform-based structural health monitoring features. In: SPIE Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring, pp. 98051N–98051N-10 (2016)
19. Sarrafi, A., Mao, Z.: Statistical modeling of wavelet-transform-based features in structural health monitoring. In: Atamturktur, H.S., Moaveni, B., Papadimitriou, C.,Schoenherr, T. (eds.) Model Validation and Uncertainty Quantification, vol. 3, pp. 253–262. Springer (2016)
20. Burkhard, R., Dell'Amico, M., Martello, S.: Assignment Problems (Revised reprint), ed: SIAM (2012)
21. Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T.: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. Signal Process. **50**, 174–188 (2002)
22. Schulz, D., Burgard, W., Fox, D., Cremers, A.B.: Tracking multiple moving targets with a mobile robot using particle filters and statistical data association. In: Robotics and Automation, 2001, Proceedings 2001 ICRA. IEEE International Conference on, 2001, pp. 1665–1670 (2001)

Chapter 30
Sequential Gauss-Newton MCMC Algorithm for High-Dimensional Bayesian Model Updating

Majid K. Vakilzadeh, Anders Sjögren, Anders T. Johansson, and Thomas J.S. Abrahamsson

Abstract Bayesian model updating provides a rigorous framework to account for uncertainty induced by lack of knowledge about engineering systems in their respective mathematical models through updates of the joint probability density function (PDF), the so-called posterior PDF, of the unknown model parameters. The Markov chain Monte Carlo (MCMC) methods are currently the most popular approaches for generating samples from the posterior PDF. However, these methods often found wanting when sampling from difficult distributions (e.g., high-dimensional PDFs, PDFs with flat manifolds, multimodal PDFs, and very peaked PDFs). This paper introduces a new multi-level sampling approach for Bayesian model updating, called Sequential Gauss-Newton algorithm, which is inspired by the Transitional Markov chain Monte Carlo (TMCMC) algorithm. The Sequential Gauss-Newton algorithm improves two aspects of TMCMC to make an efficient and effective MCMC algorithm for drawing samples from difficult posterior PDFs. First, the statistical efficiency of the algorithm is enhanced by use of the systematic resampling scheme. Second, a new MCMC algorithm, called Gauss-Newton MCMC algorithm, is proposed which is essentially an M-H algorithm with a Gaussian proposal PDF tailored to the posterior PDF using the gradient and Hessian information of the negative log posterior. The effectiveness of the proposed algorithm for solving the Bayesian model updating problem is illustrated using three examples with irregularly shaped posterior PDFs.

Keywords Bayesian model updating • Markov chain Monte Carlo • Optimal scale factor tuning • Hessian • Gauss-Newton • Uncertainty quantification

30.1 Introduction

To solve the problem of structural model updating through the Bayesian approach, a key idea is to describe the uncertain behavior of a system by constructing a *stochastic model class* which consists of two fundamental probability distributions [1]. First, a set of parameterized probability models $p(\boldsymbol{D}|\boldsymbol{\theta})$ is established to partially quantify the relative plausibility of the possible values of the system output $\boldsymbol{D} \in \mathbb{R}^{n_o}$ given the vector of uncertain parameters $\boldsymbol{\theta} \in \mathbb{R}^{n_p}$. This can be constructed by *stochastic embedding* of any deterministic model, e.g., a state-space model or a finite element model, of the system that gives the relationship between the parameter vector $\boldsymbol{\theta}$ and the model output $\boldsymbol{S} \in \mathbb{R}^{n_o}$. To this end, an uncertain prediction error can be introduced:

$$\boldsymbol{D} = \boldsymbol{S}(\boldsymbol{\theta}) + \boldsymbol{e} \tag{30.1}$$

to connect the output of the deterministic model to the observable system output. The probability model of the prediction error $\boldsymbol{e}$ is chosen to be a Gaussian white noise $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Gamma}_e)$ based on the Principle of Maximum (Information) Entropy [2] under the first and second moment constraints. Then, the predictive PDF for the observable system output is given by:

M.K. Vakilzadeh (✉) • A.T. Johansson
Applied Mechanics, Chalmers University of Technology, S-41296 Göteborg, Sweden
e-mail: khorsand@chalmers.se

T.J.S. Abrahamsson
Applied Mechanics, Chalmers University of Technology, Hörsalsvägen 7A, SE-41296 Göteborg, Sweden

A. Sjögren
Volvo Car Corporation, S-40531 Göteborg, Sweden

Mathematical Sciences, Chalmers University of Technology, S-41258 Göteborg, Sweden

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_30

$$p(\boldsymbol{D}|\boldsymbol{\theta}) = \frac{1}{(2\pi)^{n_o/2}|\boldsymbol{\Gamma}_e|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{D} - \boldsymbol{S}(\boldsymbol{\theta}))^T \boldsymbol{\Gamma}_e^{-1} (\boldsymbol{D} - \boldsymbol{S}(\boldsymbol{\theta})) \right] \tag{30.2}$$

where $|.|$ denotes the determinant of a matrix. For the second part of the stochastic model class, a *prior* distribution $p(\boldsymbol{\theta})$ over the parameter space $\boldsymbol{\Theta} \in \mathbb{R}^{n_p}$ is selected that encodes the initial relative degree of plausibility of each probability model $p(\boldsymbol{D}|\boldsymbol{\theta})$. Here, we assume that the prior distribution has a PDF and also that the negative log-prior has a positive semidefinite Hessian over $\boldsymbol{\Theta}$.

When the measurement data $\boldsymbol{D}^x$ is available from the dynamic system, the predictive PDF in (30.2) gives the likelihood function $p(\boldsymbol{D}^x|\boldsymbol{\theta})$ which can be used to update the prior probability for the uncertain parameters $\boldsymbol{\theta}$ through Bayes' Theorem to obtain the *posterior* PDF as:

$$p(\boldsymbol{\theta}|\boldsymbol{D}^x) = \frac{p(\boldsymbol{D}^x|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{D}^x)} \tag{30.3}$$

where $p(\boldsymbol{D}^x) = \int_{\boldsymbol{\Theta}} p(\boldsymbol{D}^x|\boldsymbol{\theta})p(\boldsymbol{\theta})\, d\boldsymbol{\theta}$ denotes the *evidence*, or marginal likelihood. The Markov chain Monte Carlo (MCMC) methods are currently the most popular approaches for generating samples from the posterior PDF. However, these methods face two central difficulties when solving large-scale Bayesian model updating problems; first, the large size of forward models makes the evaluation of the posterior PDF at any point in the parameter space computationally involved. Different means are used in the literature to reduce the computational cost of the forward simulations [3, 4] or in general the MCMC sampling [5]. Second, the posterior PDF is often defined over a high-dimensional parameter space or its support has a complex geometry which renders its sampling very wanting. In the field of structural dynamics, the proposed MCMC methods that can deal with such posterior PDFs essentially fall into three distinct categories. The first category includes methods that interpret a Bayesian model updating problem as an equivalent rare-event simulation problem to adopt the Subset Simulation technique for drawing samples from the posterior PDF [6–9]. The second category consists of the "Data Annealing" method proposed by Green [10] and the third category involves methods that are based on the ideas from simulated annealing, e.g., Adaptive Metropolis-Hastings (AMH) [11] and Transitional Markov Chain Monte Carlo (TMCMC) algorithm [12]. In this paper, the potential problems of the TMCMC algorithm to sample from difficult posterior distributions are discussed and improvements are proposed to make it more effective and efficient in solving high-dimensional model updating problems.

30.1.1 Sequential MCMC Algorithm

In this section, we present an overview of a novel sequential MCMC algorithm for solving the Bayesian model updating problem named Sequential Gauss-Newton. This algorithm, which is in spirit analogous to the Resample-move algorithm [13] and the TMCMC algorithm [12], is devised to facilitate sampling from difficult posterior distributions by sequentially constructing a series of intermediate distributions:

$$\pi_j(\boldsymbol{\theta}) \propto \pi(\boldsymbol{\theta})\pi(\boldsymbol{D}^x|\boldsymbol{\theta})^{\beta_j}, \quad \text{for } 0 = \beta_0 \leq \beta_1 \leq \cdots \leq \beta_m = 1 \tag{30.4}$$

that interpolates between the prior and posterior PDFs. Here j denotes the simulation level number. The idea here is that most often the geometry of the posterior PDF is dramatically different from the prior PDF which renders generating samples from posterior PDF cumbersome. However, a series of intermediate distributions starting from the prior PDF and converging to the posterior PDF can be designed such that $\pi_j(\boldsymbol{\theta})$ can be readily sampled based on samples from $\pi_{j-1}(\boldsymbol{\theta})$.

The algorithm starts by generating independent and identically distributed (*i.i.d*) samples $\{\boldsymbol{\theta}_0^k : k = 1, \ldots, N_t\}$ from the prior PDF $\pi(\boldsymbol{\theta})$. Then, at each simulation level $j \geq 1$, based on samples $\{\boldsymbol{\theta}_{j-1}^k : k = 1, \ldots, N_t\}$ from the previous level $j-1$, the same number of samples $\{\boldsymbol{\theta}_j^k : k = 1, \ldots, N_t\}$ is generated such that the samples are approximately distributed as $\pi_j(\boldsymbol{\theta})$ (N_t is the total number of samples in each level). This can be done using the following steps:

1. *Importance resampling step*: the *importance weights* can be computed for samples $\{\boldsymbol{\theta}_{j-1}^k : k = 1, \ldots, N_t\}$ generated at simulation level $j-1$ with respect to the intermediate distribution $\pi_j(\boldsymbol{\theta})$ at level j as follows:

$$w_{j-1}^k = \frac{\pi_j(\boldsymbol{\theta}_{j-1}^k)}{\pi_{j-1}(\boldsymbol{\theta}_{j-1}^k)} = \pi(\boldsymbol{D}^x|\boldsymbol{\theta}_{j-1}^k)^{\beta_j - \beta_{j-1}} \tag{30.5}$$

where $w_{j-1}^k \equiv w(\boldsymbol{\theta}_{j-1}^k)$. The samples $\{\boldsymbol{\theta}_{j-1}^k : k = 1, \ldots, N_t\}$ are then resampled according to their normalized importance weights:

$$\bar{w}_{j-1}^k = \frac{w_{j-1}^k}{\sum_{l=1}^{N_t} w_{j-1}^l} \tag{30.6}$$

to generate the set of N_t samples $\{\tilde{\boldsymbol{\theta}}_j^k : k = 1, \ldots, N_t\}$ that are approximately distributed according to $\pi_j(\boldsymbol{\theta})$ [14].

2. *Move step*: starting from the samples $\{\tilde{\boldsymbol{\theta}}_j^k : k = 1, \ldots, N_t\}$ run MCMC chains of length 1 or more to populate the support of $\pi_j(\boldsymbol{\theta})$ with N_t samples $\{\boldsymbol{\theta}_j^k : k = 1, \ldots, N_t\}$.

This procedure is proceeded until N_t samples from the posterior PDF ($\beta_m = 1$) are drawn.

The success of this sequential algorithm to a large extent depends on the schedule of the exponent parameters β_j used in the simulation. A common approach is to use adaptive schedules [12, 15, 16], in which the exponents β_j are defined based on a characteristic measure of the samples at the previous level. One such measure is the coefficient of variation (c.o.v) of the normalized importance weights $\bar{w}_{j-1}^k$, which indicates how similar adjacent PDFs $\pi_{j-1}(\boldsymbol{\theta})$ and $\pi_j(\boldsymbol{\theta})$ are. Thus, the exponent β_j can be adaptively determined so that the coefficient of variation of the normalized importance weights $\bar{w}_{j-1}^k$ is equal to a predefined threshold value δ_w.

In the next sections, we first propose to use the systematic resampling scheme within the aforementioned sequential algorithm to enhance the statistical efficiency of the algorithm. We further propose a novel MCMC sampling technique which exploits the local derivative information of the posterior PDF to accelerate the sampling algorithm, the Gauss-Newton MCMC algorithm. Finally, we summarize the proposed methods as a complete sampling technique which we call Sequential Gauss-Newton algorithm.

30.1.2 Importance Resampling

In the described sequential algorithm, the samples $\{\boldsymbol{\theta}_{j-1}^k, k = 1, \ldots, N_t\}$ generated in the simulation level $j-1$ along with their associated normalized importance weights $\bar{w}_{j-1}^k$ provide a weighted approximation density for the intermediate distribution $\pi_j(\boldsymbol{\theta})$ in form of [17]:

$$\hat{\pi}_j(\boldsymbol{\theta}) = \sum_{k=1}^{N_t} \bar{w}_{j-1}^k \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_{j-1}^k) \tag{30.7}$$

The aim of the importance resampling step is to produce samples $\{\tilde{\boldsymbol{\theta}}_j^k, k = 1, \ldots, N_t\}$ from $\hat{\pi}_j(\boldsymbol{\theta})$ that are approximately distributed as $\pi_j(\boldsymbol{\theta})$ by statistically repeating the most relevant samples to the extent of their relevance and eliminating the less relevant samples with lower $\bar{w}_{j-1}^k$. Therefore, at each simulation level j, $\hat{\pi}_j(\boldsymbol{\theta})$ is replaced by an empirical distribution:

$$\tilde{\pi}_j(\boldsymbol{\theta}) = \sum_{k=1}^{N_t} \frac{1}{N_t} \delta(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}_j^k) = \sum_{k=1}^{N_t} \frac{\tilde{n}_j^k}{N_t} \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_{j-1}^k) \tag{30.8}$$

under the following constraints:

$$\begin{aligned} &\sum_{k=1}^{N_t} \tilde{n}_j^k = N_t \\ &\mathbb{E}\left[\tilde{n}_j^k \middle| \bar{\boldsymbol{w}}_{j-1}\right] = N_t \bar{w}_{j-1}^k \end{aligned} \tag{30.9}$$

Here, $\tilde{n}_j^k$ is the duplication counts of the sample $\boldsymbol{\theta}_{j-1}^k$ in the resampled set of samples $\{\tilde{\boldsymbol{\theta}}_j^i, i = 1, \ldots, N_t\}$, $\delta(.)$ denotes the Dirac delta function, and $\bar{\boldsymbol{w}}_{j-1} = [\bar{w}_{j-1}^1, \ldots, \bar{w}_{j-1}^{N_t}]^T$.

Numerous resampling schemes [17, 18] are proposed in the literature which satisfy the constraints in (30.9). Two of such resampling schemes are the multinomial resampling [19], and systematic resampling [20] schemes. These two schemes start

Algorithm 1 Pseudo code for the systematic resampling scheme at simulation level j

Inputs
$\{\boldsymbol{\theta}_{j-1}^k, k = 1, \ldots, N_t\} \sim \pi_{j-1}(\boldsymbol{\theta})$
$\{\bar{w}_{j-1}^k, k = 1, \ldots, N_t\}$ (the importance weights corresponding to $\boldsymbol{\theta}_{j-1}^k$'s),
Algorithm
Permute randomly the samples $\{\boldsymbol{\theta}_{j-1}^k : k = 1, \ldots, N_t\}$,
Sample $u_1 \sim \mathscr{U}[0, 1/N_t]$,
Compute $u_k = \frac{k-1}{N_t} + u_1$ for $k = 2, \ldots, N_t$,
for $k = 1, \ldots, N_t$ **do**
  Set $\tilde{\boldsymbol{\theta}}_j^k = \boldsymbol{\theta}_{j-1}^l$ such that index l satisfies (30.10),
end for
Outputs
$\{\tilde{\boldsymbol{\theta}}_j^k, k = 1, \ldots, N_t\} \sim \pi_j(\boldsymbol{\theta})$

by drawing a random number u_i from a uniform distribution $\mathscr{U}[0, 1]$. The sample $\tilde{\boldsymbol{\theta}}_j^i = \boldsymbol{\theta}_{j-1}^l$ is then selected such that l satisfies:

$$\sum_{k=1}^{l-1} \bar{w}_{j-1}^k \leq u_i \leq \sum_{k=1}^{l} \bar{w}_{j-1}^k, \tag{30.10}$$

where l is a dummy index. Repeating this procedure for $i = 1, \ldots, N_t$ generates a new set of samples $\{\tilde{\boldsymbol{\theta}}_j^i, i = 1, \ldots, N_t\}$ that are approximately distributed according to $\hat{\pi}_j(\boldsymbol{\theta})$. The difference between these resampling schemes arises from the way they generate the sequence of random numbers $\{u_i, i = 1, \ldots, N_t\}$. The multinomial resampling scheme, as the most straightforward scheme, draws N_t independent random numbers $u_i \sim \mathscr{U}[0, 1]$, for $i = 1, \ldots, N_t$ [19]. On the other hand, the systematic resampling partitions the interval $(0, 1]$ into N_t disjoint sets, $(0, 1] = (0, 1/N_t] \cup \cdots \cup ((N_t - 1)/N_t, 1]$, and draws the first random number $u_1 \sim \mathscr{U}[0, 1/N_t]$. Then, the other random numbers are deterministically linked to u_1 as:

$$u_k = \frac{(k-1)}{N_t} + u_1 \quad \text{for } k = 2, \ldots, N_t \tag{30.11}$$

Therefore, the outcome of the systematic resampling scheme is deterministically dependent on the order in which samples $\{\boldsymbol{\theta}_{j-1}^k, k = 1, \ldots, N_t\}$ are sorted. To remove such deterministic link, one can randomly permute the samples before resampling [18]. See Algorithm 1 for a detailed implementation of the systematic resampling scheme, which is included to complete the series of pseudo codes for the proposed sequential MCMC algorithm.

The resampling schemes introduce a variability to the duplication count $\tilde{n}_j^k$ of samples [21]. To reduce this, in this study we propose to employ the systematic resampling scheme within the described sequential algorithm.

30.1.3 MCMC Sampling

The importance resampling step generates a set of samples $\{\tilde{\boldsymbol{\theta}}_j^k, k = 1, \ldots, N_t\}$ approximately distributed according to $\pi_j(\boldsymbol{\theta})$ [14]. Assume that this set includes $N_j^{(0)}$ distinct samples $\{\boldsymbol{\theta}_j^{(0),k}, k = 1, \ldots, N_j^{(0)}\}$, each duplicated $n_j^{(0),k}$ times. In the move step, an MCMC chain of length $n_j^{(0),k}$ is initiated from each $\boldsymbol{\theta}_j^{(0),k}$ to draw new samples from the target distribution $\pi_j(\boldsymbol{\theta})$. Because of this we hereafter call the $\boldsymbol{\theta}_j^{(0),k}$s seeds for MCMC chains. This leads to an enriched set of N_t samples $\{\boldsymbol{\theta}_j^k, k = 1, \ldots, N_t\}$ from the support of the intermediate distribution $\pi_j(\boldsymbol{\theta})$. The role of MCMC sampling in the proposed sequential algorithm is thus to hinder the accumulation of the importance weights on just a few samples as the simulation level j increases [13] and, as a result, helps to avoid the sample impoverishment.

The challenge for MCMC algorithms is to devise a proposal PDF that is a good approximation of the posterior distribution and at the same time is inexpensive to manipulate. In this section, we first present a new MCMC algorithm that exploits the local gradient and Hessian information of the negative log posterior to accelerate the MCMC sampling. Afterwards, by exploiting the special structure of the likelihood function (30.2), we propose an augmented Gauss-Newton approximation of the Hessian.

Algorithm 2 Pseudo code for the Metropolis-Hastings algorithm to sample from target PDF $\pi_j(\boldsymbol{\theta})$

Inputs
$\boldsymbol{\theta}_j^{(0),k}$ (the kth seed in the simulation level j),
$n_j^{(0),k}$ (number of duplications associated with the kth seed in the simulation level j),
Algorithm
Compute $\pi_j(\boldsymbol{\theta}_j^{(0),k})$,
for $i = 1, \ldots, n_j^{(0),k}$ **do**
Draw a candidate sample from proposal density $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_j^{(i-1),k})$
Compute $\pi_j(\boldsymbol{\theta}^*)$, $q(\boldsymbol{\theta}_j^{(i-1),k}|\boldsymbol{\theta}^*)$ and $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_j^{(i-1),k})$
Compute $\alpha(\boldsymbol{\theta}_j^{(i-1),k}, \boldsymbol{\theta}^*) = \min\{1, \frac{\pi_j(\boldsymbol{\theta}^*)q(\boldsymbol{\theta}_j^{(i-1),k}|\boldsymbol{\theta}^*)}{\pi_j(\boldsymbol{\theta}_j^{(i-1),k})q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_j^{(i-1),k})}\}$
Set

$$\boldsymbol{\theta}_j^{(i),k} = \begin{cases} \boldsymbol{\theta}^* & \text{with probability } \alpha(\boldsymbol{\theta}_j^{(i-1),k}, \boldsymbol{\theta}^*) \\ \boldsymbol{\theta}_j^{(i-1),k} & \text{with probability } 1 - \alpha(\boldsymbol{\theta}_j^{(i-1),k}, \boldsymbol{\theta}^*) \end{cases}$$

end for
Outputs
Samples $\{\boldsymbol{\theta}_j^{(i),k} : i = 1, \ldots, n_j^{(0),k}\}$ from the intermediate distribution $\pi_j(\boldsymbol{\theta})$.

30.1.3.1 Hessian Informed Metropolis-Hastings Algorithm

The Metropolis-Hastings algorithm, first proposed by Metropolis [22] and then improved by Hastings [23], is probably the most popular MCMC algorithm to draw samples from a target distribution. Algorithm 2 presents a pseudo code for the M-H algorithm. The M-H algorithm attains its popularity due to the minimal requirements it imposes on the proposal density $q(\boldsymbol{\theta}^*|\boldsymbol{\theta})$ and the target distribution [24]. One of the simplest choices for the proposal density is the isotropic Gaussian PDF:

$$q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_j^{(i),k}) = \frac{1}{(2\pi\sigma^2)^{n_p/2}} \exp\left[-\frac{(\boldsymbol{\theta}^* - \boldsymbol{\theta}_j^{(i),k})^T(\boldsymbol{\theta}^* - \boldsymbol{\theta}_j^{(i),k})}{2\sigma^2}\right] \tag{30.12}$$

where $\boldsymbol{\theta}_j^{(i),k} \in \mathbb{R}^{n_p}$ denotes the ith state of a Markov chain initiated from the kth seed $\boldsymbol{\theta}_j^{(0),k}$ at simulation level j and σ^2 denotes the proposal variance. While it is easy to draw samples from this proposal density, this choice results in poor performance of the M-H algorithm due to the discrepancy between the proposal and the target PDFs. At the other end of the spectrum, an ideal choice for the proposal density is the target PDF itself for which the acceptance probability is 1. For most model updating problems the target PDF is not known a priori, so a direct sampling from it is not feasible. To achieve an efficient sampling scheme, thus, the challenge is to find a proposal distribution that is both easy to sample from and a good local approximation of the target PDF.

Such a proposal density can be constructed by exploiting the local structure of the target distribution. To this end, assume that the target PDF $\pi_j(\boldsymbol{\theta})$ has continuous and bounded derivatives up to order two. Its negative logarithm can be expanded in a second-order Taylor series about $\boldsymbol{\theta}_j^{(i),k}$ as follows:

$$-\log \pi_j(\boldsymbol{\theta}) \approx -\log \pi_j(\boldsymbol{\theta}_j^{(i),k}) + \boldsymbol{g}_j(\boldsymbol{\theta}_j^{(i),k})(\boldsymbol{\theta} - \boldsymbol{\theta}_j^{(i),k}) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_j^{(i),k})^T \boldsymbol{H}_j(\boldsymbol{\theta}_j^{(i),k})(\boldsymbol{\theta} - \boldsymbol{\theta}_j^{(i),k}) \tag{30.13}$$

Here, $\boldsymbol{H}_j(\boldsymbol{\theta}) = -\nabla^2 \log \pi_j(\boldsymbol{\theta})$ and $\boldsymbol{g}_j(\boldsymbol{\theta}) = -\nabla \log \pi_j(\boldsymbol{\theta})$ denote the Hessian and the gradient of $-\log \pi_j(\boldsymbol{\theta})$, respectively. We can rewrite this expression in the form:

$$-\log \pi_j(\boldsymbol{\theta}) \approx \left(\boldsymbol{\theta} - \boldsymbol{\theta}_j^{(i),k} + \boldsymbol{H}_j(\boldsymbol{\theta}_j^{(i),k})^{-1}\boldsymbol{g}_j(\boldsymbol{\theta}_j^{(i),k})\right)^T \boldsymbol{H}_j(\boldsymbol{\theta}_j^{(i),k})\left(\boldsymbol{\theta} - \boldsymbol{\theta}_j^{k} + \boldsymbol{H}_j(\boldsymbol{\theta}_j^{(i),k})^{-1}\boldsymbol{g}_j(\boldsymbol{\theta}_j^{(i),k})\right) \tag{30.14}$$

which gives a local Gaussian approximation, i.e.,

$$\bar{\pi}_j(\boldsymbol{\theta}) = \frac{|\boldsymbol{H}_j(\boldsymbol{\theta}_j^{(i),k})|^{1/2}}{(2\pi)^{n_p/2}} \exp\left[-\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{m}_i)^T \boldsymbol{H}_j(\boldsymbol{\theta}_j^{(i),k})(\boldsymbol{\theta} - \boldsymbol{m}_i) \right] \tag{30.15}$$

for the target PDF centered on the point $\boldsymbol{m}_i = \boldsymbol{\theta}_j^{(i),k} - \boldsymbol{H}_j(\boldsymbol{\theta}_j^{(i),k})^{-1} \boldsymbol{g}_j(\boldsymbol{\theta}_j^{(i),k})$ with the covariance $\boldsymbol{H}_j(\boldsymbol{\theta}_j^{(i),k})^{-1}$. Taking $\bar{\pi}_j(\boldsymbol{\theta})$ as the proposal density for the M-H algorithm, i.e., $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_j^{(i),k}) = \bar{\pi}_j(\boldsymbol{\theta}^*)$, generates an MCMC method with a proposal density that not only can be readily sampled, but also is aligned according to the local curvature of the target distribution.

A careful look at the mean value $\boldsymbol{m}_i$ reveals that it has an additional deterministic term compared to the formulation of the isotropic Gaussian proposal. This term, also called drift term [25], can be interpreted as a so-called Newton step from an optimization perspective. Provided that the Hessian $\boldsymbol{H}_j(\boldsymbol{\theta})$ is positive definite, the drift term tends to drive the current state of the Markov chain to a point with a higher value of $\pi_j(\boldsymbol{\theta})$. This results in a faster exploration of the highly probable region of $\pi_j(\boldsymbol{\theta})$ and, as a consequence, speeds up the convergence of the Markov sampling process.

In fact, there exist two challenges when using the Hessian for speeding up MCMC sampling within the proposed sequential algorithm: (1) the exact Hessian matrix is an expensive piece of information to obtain, and (2) convergence is guaranteed only if the Hessian matrix is positive definite, but that is not generally guaranteed at an arbitrary point in the parameter space. In the next section, we propose an approximation for the Hessian which alleviates the aforementioned challenges.

30.1.3.2 Gauss-Newton Approximation of Hessian

In the setting of the proposed sequential algorithm, the Hessian of the negative logarithm of an intermediate PDF $\pi_j(\boldsymbol{\theta})$ can be written as:

$$\boldsymbol{H}_j(\boldsymbol{\theta}) = \beta_j \boldsymbol{H}_\ell(\boldsymbol{\theta}) + \boldsymbol{H}_{pr}(\boldsymbol{\theta}) \tag{30.16}$$

where $\boldsymbol{H}_\ell(\boldsymbol{\theta})$ denotes the Hessian of the negative log-likelihood function and $\boldsymbol{H}_{pr}(\boldsymbol{\theta})$ denotes the Hessian of the negative log-prior. Since $\boldsymbol{H}_{pr}(\boldsymbol{\theta})$ is assumed to be known, the computationally expensive part of (30.16) is the evaluation of $\boldsymbol{H}_\ell(\boldsymbol{\theta})$. Therefore, we here exploit tools established in the context of optimization to provide a robust and computationally efficient approximation for the Hessian of the negative log-likelihood. In addition, we modify the resulting approximation of the Hessian $\boldsymbol{H}_j(\boldsymbol{\theta})$ so that it is positive definite over the support of $\pi_j(\boldsymbol{\theta})$.

In the Bayesian model updating problem as described in Sect. 30.1, the negative log-likelihood has the following special form:

$$-\log \pi(\boldsymbol{D}^x|\boldsymbol{\theta}) = \frac{1}{2}(\boldsymbol{D}^x - \boldsymbol{S}(\boldsymbol{\theta}))^T \boldsymbol{\Gamma}_e^{-1} (\boldsymbol{D}^x - \boldsymbol{S}(\boldsymbol{\theta})) + \text{constant} \tag{30.17}$$

By decomposing the noise covariance matrix as $\boldsymbol{\Gamma}_e = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^T$, using either the symmetric square root or Cholesky factorization, this expression can be rewritten as:

$$-\log \pi(\boldsymbol{D}^x|\boldsymbol{\theta}) = \frac{1}{2}\left[\boldsymbol{\Gamma}^{-1}(\boldsymbol{D}^x - \boldsymbol{S}(\boldsymbol{\theta}))\right]^T \left[\boldsymbol{\Gamma}^{-1}(\boldsymbol{D}^x - \boldsymbol{S}(\boldsymbol{\theta}))\right] + \text{constant} = \frac{1}{2}\boldsymbol{r}(\boldsymbol{\theta})^T \boldsymbol{r}(\boldsymbol{\theta}) + \text{constant} \tag{30.18}$$

where $\boldsymbol{r}(\boldsymbol{\theta}) = [\boldsymbol{\Gamma}^{-1}(\boldsymbol{D}^x - \boldsymbol{S}(\boldsymbol{\theta}))] \in \mathbb{R}^{n_o}$. Assuming that the log-likelihood function is twice continuously differentiable, its gradient and Hessian can then be expressed as [26]:

$$\boldsymbol{g}_\ell(\boldsymbol{\theta}) = -\nabla \log \pi(\boldsymbol{D}^x|\boldsymbol{\theta}) = \boldsymbol{J}(\boldsymbol{\theta})^T \boldsymbol{r}(\boldsymbol{\theta}) \tag{30.19}$$

$$\boldsymbol{H}_\ell(\boldsymbol{\theta}) = -\nabla^2 \log \pi(\boldsymbol{D}^x|\boldsymbol{\theta}) = \boldsymbol{J}(\boldsymbol{\theta})^T \boldsymbol{J}(\boldsymbol{\theta}) + \sum_{j=1}^{n_o} \boldsymbol{r}_j(\boldsymbol{\theta}) \nabla^2 \boldsymbol{r}_j(\boldsymbol{\theta}) \tag{30.20}$$

where $\boldsymbol{J}(\boldsymbol{\theta}) \in \mathbb{R}^{n_o \times n_p}$ denotes the *Jacobian* matrix defined by:

$$\boldsymbol{J}(\boldsymbol{\theta}) = \begin{bmatrix} \nabla \boldsymbol{r}_1(\boldsymbol{\theta})^T \\ \nabla \boldsymbol{r}_2(\boldsymbol{\theta})^T \\ \vdots \\ \nabla \boldsymbol{r}_{n_o}(\boldsymbol{\theta})^T \end{bmatrix}^T \tag{30.21}$$

The special structure of the log-likelihood function (30.16) and its second derivative (30.20), suggests to approximate the Hessian of the negative log-likelihood function by its Gauss-Newton portion, i.e., $\boldsymbol{H}_\ell(\boldsymbol{\theta}) \approx \tilde{\boldsymbol{H}}_\ell(\boldsymbol{\theta}) = \boldsymbol{J}(\boldsymbol{\theta})^T \boldsymbol{J}(\boldsymbol{\theta})$. It has been previously used by Vakilzadeh et al. [27].

Substituting $\tilde{\boldsymbol{H}}_\ell(\boldsymbol{\theta})$ in (30.16) gives the following approximation for the Hessian of the negative logarithm of $\pi_j(\boldsymbol{\theta})$:

$$\boldsymbol{H}_j(\boldsymbol{\theta}) \approx \tilde{\boldsymbol{H}}_j(\boldsymbol{\theta}) = \beta_j \boldsymbol{J}(\boldsymbol{\theta})^T \boldsymbol{J}(\boldsymbol{\theta}) + \boldsymbol{H}_{pr}(\boldsymbol{\theta}) \tag{30.22}$$

We refer to $\tilde{\boldsymbol{H}}_j(\boldsymbol{\theta})$ as the Gauss-Newton Hessian. This simple modification of the Hessian matrix gives two immediate advantages over $\boldsymbol{H}_\ell(\boldsymbol{\theta})$. First, it eliminates the need for evaluation of the second derivatives of $\boldsymbol{r}(\boldsymbol{\theta})$ in (30.20) and thus constructing $\tilde{\boldsymbol{H}}_\ell(\boldsymbol{\theta})$ only requires n_p evaluations of the forward problem. Second, whenever $\boldsymbol{J}(\boldsymbol{\theta})$ has full column rank and the gradient $\boldsymbol{g}_\ell(\boldsymbol{\theta})$ is nonzero, $\tilde{\boldsymbol{H}}_\ell(\boldsymbol{\theta})$ is a positive definite matrix. This guarantees that the drift term in the proposal PDF (30.15) drives the current state towards a point with a higher value of the posterior PDF.

The Gauss-Newton Hessian $\tilde{\boldsymbol{H}}_j(\boldsymbol{\theta})$ gives a close approximation to $\boldsymbol{H}_j(\boldsymbol{\theta})$ when $\boldsymbol{J}(\boldsymbol{\theta})^T \boldsymbol{J}(\boldsymbol{\theta})$ dominates in (30.20). This is the case when the forward problem $\boldsymbol{S}(\boldsymbol{\theta})$ is an almost linear function of the uncertain parameters, the noise is small or the number of observations n_o is large [26, 28]. When the Gauss-Newton Hessian $\tilde{\boldsymbol{H}}_j(\boldsymbol{\theta})$ is not an adequate approximation of $\boldsymbol{H}_j(\boldsymbol{\theta})$, the Markov chain remains still in local regions for long periods of time and, hence, the average acceptance rate is low. To avoid this problem, we propose to replace $\tilde{\boldsymbol{H}}_j(\boldsymbol{\theta})$ by its augmented form:

$$\tilde{\boldsymbol{H}}_{Gt,j}(\boldsymbol{\theta}) = \beta_j \boldsymbol{J}(\boldsymbol{\theta})^T \boldsymbol{J}(\boldsymbol{\theta}) + \boldsymbol{H}_{pr}(\boldsymbol{\theta}) + \lambda_j \boldsymbol{I}_{n_p} \tag{30.23}$$

where $\boldsymbol{I}_{n_p}$ is the $n_p \times n_p$ identity matrix. This is known as a *trust-region* technique in the context of optimization [26]. The rationale behind it is that if $\tilde{\boldsymbol{H}}_j(\boldsymbol{\theta})$ can only be trusted as a good approximation for $\boldsymbol{H}_j(\boldsymbol{\theta})$ in a small region around $\boldsymbol{\theta}$, the *trust-region parameter* λ_j will be selected appropriately large such that $\lambda_j \boldsymbol{I}_{n_p}$ dominates the other terms in (30.23) and, hence, the drift term in (30.15) coincides with the steepest descent direction from an optimization perspective. In addition, when neither the likelihood function nor the prior PDF contain sufficient information (meaning that $\boldsymbol{H}$ is close to singular), a suitable λ_j can be imposed to make $\tilde{\boldsymbol{H}}_{Gt,j}(\boldsymbol{\theta})$ positive definite.

30.1.4 Summary of the Sequential Gauss-Newton Algorithm

The proposed sequential sampling algorithm for solving the Bayesian model updating problem is summarized in Algorithm 3.

Algorithm 3 Summary of the Sequential Gauss-Newton algorithm

Inputs

N_t (the total number of samples to be generated at each simulation level),
δ_w (prescribed threshold value for the c.o.v of the importance weights),

Algorithm

Draw independent samples $\{\boldsymbol{\theta}_0^k \in \mathbb{R}^{n_p},\ k = 1, \ldots, N_t\}$ from the prior PDF $\pi(\boldsymbol{\theta})$,
Set $j = 1$, $\beta_0 = 0$,
while $\beta_{j-1} \leq 1$ **do**

- Find β_j such that the c.o.v of $\bar{w}_{j-1}^k(\beta_j)$, $k = 1, \ldots, N_t$, is equal to the threshold value δ_w,
- Calculate the normalized importance weights $\bar{w}_{j-1}^k$ using (30.6) for the samples $\{\boldsymbol{\theta}_{j-1}^1, \ldots, \boldsymbol{\theta}_{j-1}^{N_t}\}$ at level $j-1$,
- Using the systematic resampling scheme, resample $\{\boldsymbol{\theta}_{j-1}^1, \ldots, \boldsymbol{\theta}_{j-1}^{N_t}\}$ according to their associated weights $\bar{w}_{j-1}^k$ using Algorithm 1 to obtain $N_j^{(0)}$ seeds $\{\boldsymbol{\theta}_j^{(0),1}, \ldots, \boldsymbol{\theta}_j^{(0),N_j^{(0)}}\}$ each duplicated $n_j^{(0),k}$ times,
- Starting from $\{\boldsymbol{\theta}_j^{(0),k},\ k = 1, \ldots, N_j^{(0)}\}$, run Algorithm 2 to simulate N_t samples $\{\boldsymbol{\theta}_j^k,\ k = 1, \ldots, N_t\}$ from $\pi_j(\boldsymbol{\theta})$ using the Gauss-Newton MCMC algorithm,
- $j = j + 1$.

end while

Outputs

Samples $\{\boldsymbol{\theta}_m^k,\ k = 1, \ldots, N_t\}$ from the posterior PDF $\pi(\boldsymbol{\theta}|\boldsymbol{D}^x)$.

30.2 Illustrative Examples

In this section, three examples are presented to illustrate the capability of the Sequential Gauss-Newton algorithm to draw samples from parameter posterior PDFs with challenging topologies. The first example is the classic Rosenbrock function from deterministic optimization, which has the following effective target distribution [29]:

$$p(\boldsymbol{D}^x|\boldsymbol{\theta}) \propto \exp\left[\frac{-(1-\theta_1)^2 - R^2(\theta_2 - \theta_1^2)^2}{2}\right] \tag{30.24}$$

Here, $R > 0$ is the "Rosenbrock" constant. The reason behind studying this example is three-fold. First, the parameters have a nonlinear correlation. Second, the Hessian matrix can become ill-conditioned. Third, the target distribution cannot be well approximated by the Gaussian PDF (30.15). We study the behavior of the sequential Gauss-Newton algorithm for three different values of the Rosenbrock constant, $R = 1$, $R = 5$, and $R = 10$. The prior PDF for the parameters θ_1 and θ_2 is defined as the union of independent uniform distributions $\mathscr{U}[-20, 20]$.

The second example is selected form Haario et al. [30]. In this example, the proposed algorithm is applied to draw samples from a "banana-shaped" target distribution as follows:

$$p(\boldsymbol{D}^x|\boldsymbol{\theta}) \propto \exp\left[-\frac{1}{200}\theta_1^2 - \frac{1}{2}(\theta_2 + B\theta_1^2 - 100B)^2\right] \tag{30.25}$$

where $B > 0$ denotes the "bananicity" constant. This example is selected here since the existing strong nonlinear correlation between the parameters renders the exploration of the target distribution very challenging for MCMC algorithms. Here, we consider three different values of the bananicity constant, i.e., $B = 0$, $B = 0.05$, and $B = 0.1$. For $B = 0$, the banana-shaped distribution gives a Gaussian distribution. By increasing B the nonlinear dependency between the components of the unknown parameter vector increases. The prior PDF over the parameter vector θ is selected as the product of uniform distributions for each component $\theta_i \in [-40, 40]$, for $i = 1, 2$.

The third example is in spirit similar to the example studied by Beck and Au [11]. In this example, a moment-resisting frame structure is modeled using a two-degree-of-freedom shear building in order to identify the interstory stiffnesses based on the measured eigenfrequencies. The shear building has known story masses $m_1 = 16.5 \times 10^3$ kg and $m_2 = 16.1 \times 10^3$ kg. The unknown interstory stiffnesses are parameterized as $k_1 = \bar{k}_1\theta_1$ and $k_2 = \bar{k}_2\theta_2$ where $\bar{k}_1 = \bar{k}_2 = 29.7 \times 10^6$ N/m are the actual interstory stiffnesses for the first and second stories. Damping is not considered in this example. The prior PDFs for θ_1 and θ_2 are selected as independent uniform distributions $\mathscr{U}[0, 3]$. Following Beck and Au [11], the likelihood is defined as:

$$p(\boldsymbol{D}^x|\boldsymbol{\theta}) \propto \exp\left[-\frac{\sum_{i=1}^{n_o}(f_i^2/\tilde{f}_i^2 - 1)^2}{2\sigma_\epsilon^2}\right] \tag{30.26}$$

where f_i denotes the ith eigenfrequency predicted by the model, $\tilde{f}_1 = 3.13$ Hz and $\tilde{f}_1 = 9.83$ Hz denote the measured eigenfrequencies and $\sigma_\epsilon^2 = 1/16$ denotes the variance of the prediction error. Here, two levels of identifiability are studied: (1) Case I (unidentifiable): there are infinitely many maximum likelihood estimates since data includes only the fundamental eigenfrequency, i.e., $\boldsymbol{D}^x = \tilde{f}_1$, (2) Case II (locally identifiable): there is more than one maximum likelihood estimate since data contains both eigenfrequencies, i.e., $\boldsymbol{D}^x = [\tilde{f}_1, \tilde{f}_2]^T$.

30.3 Numerical Results and Discussion

In this section, we first present numerical results that show the influence of the choice of the importance resampling scheme on the statistical efficiency of the proposed sequential algorithm. Afterwards, the proposed sequential algorithm is used to draw samples from the target distribution of the examples presented in the previous section.

– *Importance resampling scheme*: In order to compare the performance of various importance resampling schemes within the Sequential Gauss-Newton algorithm and to demonstrate the merits of the systematic resampling scheme, the Sequential Gauss-Newton algorithm is implemented for all examples mentioned in the previous section. For each case, we perform 50 runs each of which generates $N_t = 2000$ samples at each simulation level. Here, $\delta_w = 0.1$ is selected to force the proposed algorithm to take a large number of simulation levels in order to converge to the target PDF. This makes the effect of the

Table 30.1 Normalized effective sample size for various resampling schemes with automatic tuning of the trust-region parameter λ where the threshold value for the c.o.v of the importance weights set to $\delta_w = 0.1$

	Banana-shape			Rosenbrock			2DOF	
	$B = 0$	$B = 0.05$	$B = 0.1$	$R = 1$	$R = 5$	$R = 10$	Case I	Case II
Multinomial	0.507	0.105	0.050	0.285	0.088	0.052	0.043	0.027
Residual	0.571	0.185	0.130	0.417	0.146	0.111	0.109	0.068
Stratified	0.582	0.201	0.171	0.482	0.160	0.087	0.144	0.067
Systematic	0.806	0.458	0.409	0.881	0.417	0.347	0.502	0.174

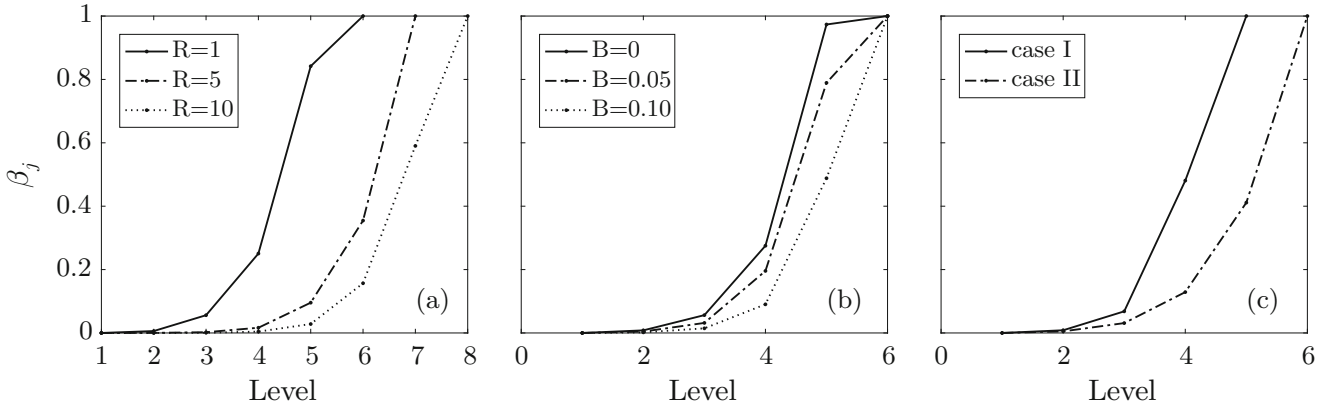


Fig. 30.1 The value of the exponent parameter β_j for different levels of the Sequential Gauss-Newton algorithm when applied to (**a**) the Rosenbrock distribution, (**b**) the banana-shaped distribution, and (**c**) the 2DOF linear structure

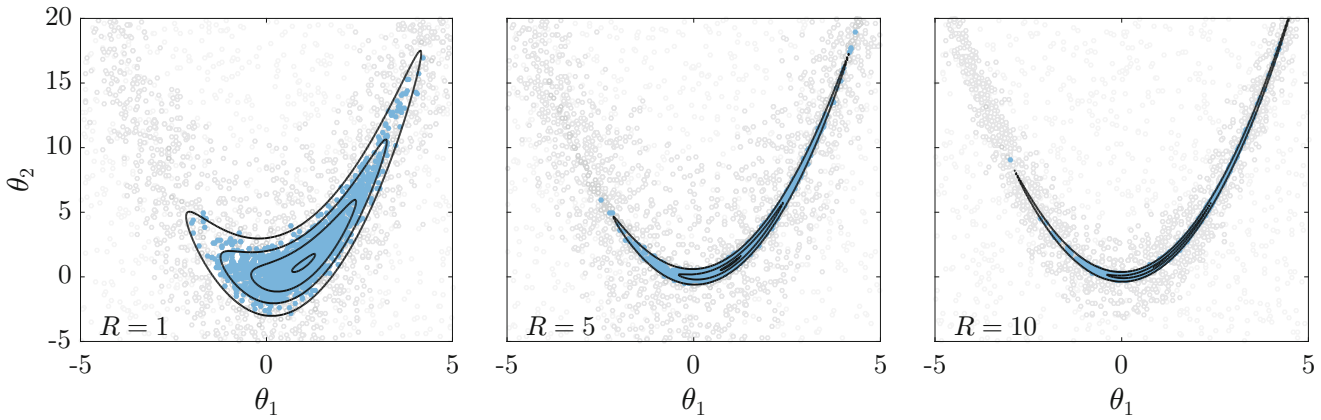


Fig. 30.2 Scatter plot of 2000 samples from the Rosenbrock distribution for the intermediate levels (in *gray*) and the final level (in *blue*) for different values of the Rosenbrock parameter R. The *black lines* are the exact contour plots of the Rosenbrock distribution

resampling scheme on the performance of the proposed algorithm more pronounced. Table 30.1 presents the normalized ESS. It can be observed that the systematic resampling scheme gives the largest ESS for all cases. This implies that the systematic resampling outperforms other resampling schemes in terms of the efficiency of the estimator $\tilde{h}$.

– *Performance of the Sequential Gauss-Newton algorithm*: The Sequential Gauss-Newton algorithm with the systematic resampling and $\delta_w = 1$, where the δ_w value is taken from Ching and Cheng [12], is implemented to draw 2000 samples from the target distribution for all examples. Figure 30.1 demonstrates the evolution of the exponent parameter β_j as the simulation level j increases. Figures 30.2, 30.3, and 30.4 plot 2000 samples of the parameters from the target distribution for all cases. Figure 30.5 presents the empirical marginal cumulative distribution function for the first component of the parameter vector of the linear 2DOF structure over 50 simulation runs.

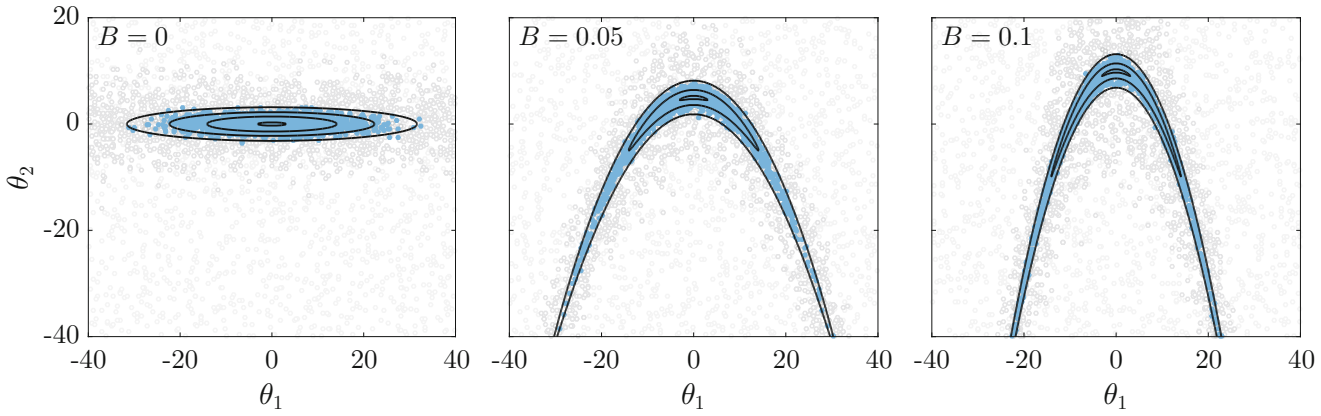


Fig. 30.3 Scatter plot of 2000 samples from the banana-shaped distribution for the intermediate levels (in *gray*) and the final level (in *blue*) for different values of the bananicity parameter B. The *black lines* are the exact contour plots of the banana-shaped distribution

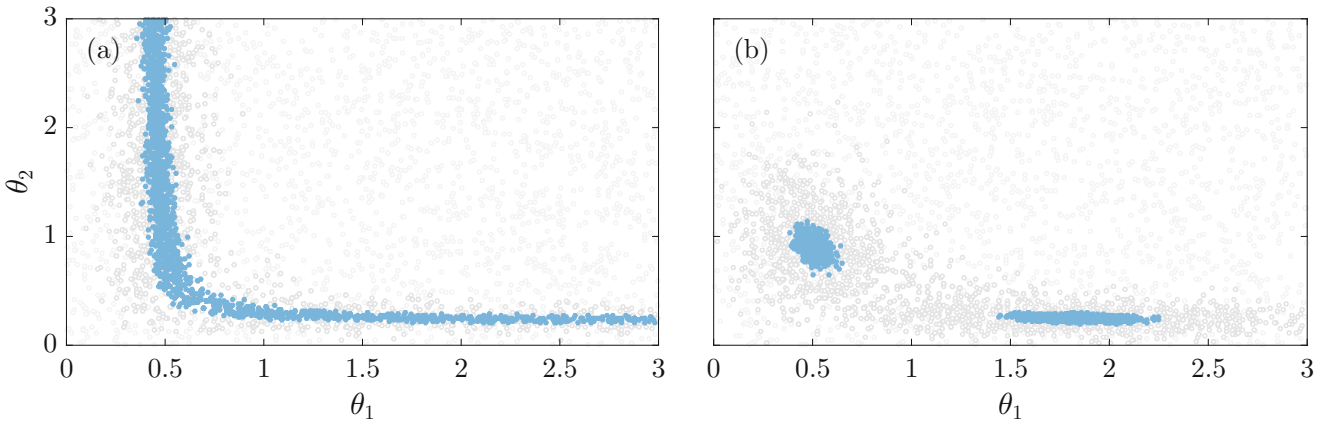


Fig. 30.4 Scatter plot of 2000 samples from the uncertain parameters of the linear two DOF structure for the intermediate levels (in *gray*) and the final level (in *blue*) for different levels of identifiability: (**a**) Case I: unidentifiable, and (**b**) Case II: locally identifiable

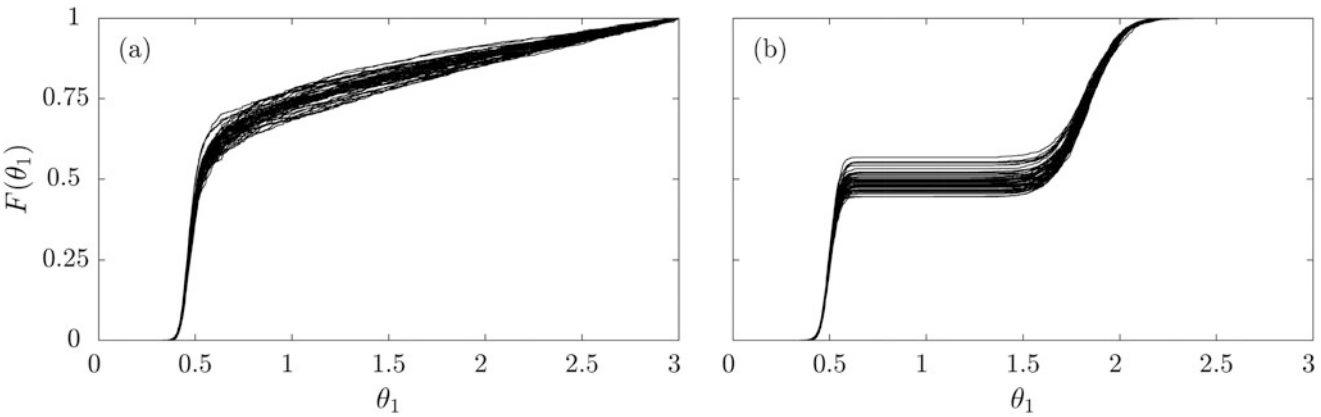


Fig. 30.5 The posterior cumulative distribution function for the first parameter θ_1 of the linear 2DOF structure over 50 runs of the Sequential Gauss-Newton algorithm for (**a**) Case I: unidentifiable, and (**b**) Case II: locally identifiable

30.4 Conclusion

In this paper, a new multi-level algorithm, called the Sequential Gauss-Newton algorithm, for sampling from difficult posterior distributions appearing in the Bayesian updating of structural models is introduced. The algorithm replaces sampling from difficult posterior PDFs by sampling from a sequence of intermediate distributions converging to the posterior PDF. The Sequential Gauss-Newton algorithm adopts the systematic resampling algorithm from the particle filter community to avoid the problem of sample impoverishment and also to enhance the statistical efficiency of estimates for the posterior expectation of uncertain quantities of interest.

At the heart of this paper is a new MCMC algorithm called the Gauss-Newton MCMC algorithm which is essentially an M-H algorithm with a Gaussian proposal PDF that locally approximates the posterior PDF. The Gauss-Newton approximation of the Hessian is proposed as it is well known within classical optimization for being numerically robust, computationally efficient and semi-positive definite over the parameter space given certain conditions. Furthermore, a diagonal matrix is added to the Hessian matrix to control the step length of the MCMC algorithm, thereby constraining the step length when the target distribution cannot be well approximated by a Gaussian distribution.

The performance of the Sequential Gauss-Newton MCMC algorithm is studied using three illustrative numerical examples with challenging irregularly shaped target PDFs. The next step is to apply the Sequential Gauss-Newton algorithm to Bayesian model updating of more complex structural systems, in particular systems with a high-dimensional vector of uncertain parameters.

Acknowledgements The first author would like to thank Prof. James L. Beck from California Institute of Technology for his comments and suggestions on the early stages of this work.

References

1. Beck, J.L.: Bayesian system identification based on probability logic. Struct. Control Health Monit. **17**(7), 825–847 (2010)
2. Jaynes, E.T.: Information theory and statistical mechanics. Phys. Rev. **106**(4), 620 (1957)
3. Vakilzadeh, M.K., Rahrovani, S., Abrahamsson, T.: Modal reduction based on accurate input-output relation preservation. In: Topics in Modal Analysis, Springer, New York, **7**, pp. 333–342 (2014)
4. Yaghoubi, V., Vakilzadeh, M.K., Abrahamsson, T.: A parallel solution method for structural dynamic response analysis. In: Dynamics of Coupled Structures, Springer International Publishing, New York, **4**, pp. 149–161 (2015)
5. Madarshahian, R., Caicedo, J.M.: Reducing MCMC computational cost with a two layered Bayesian approach. In: Model Validation and Uncertainty Quantification, Springer International Publishing, New York, **3**, pp. 291–297 (2015)
6. Chiachio, M., Beck, J.L., Chiachio, J., Rus, G.: Approximate Bayesian computation by subset simulation. SIAM J. Sci. Comput. **36**(3), A1339–A1358 (2014)
7. Vakilzadeh, M.K., Huang, Y., Beck, J.L., Abrahamsson, T.: Approximate Bayesian computation by subset simulation using hierarchical state-space models. Mech. Syst. Signal Process. **84**(part B), 2–20 (2017)
8. Straub, D., Papaioannou, I.: Bayesian updating with structural reliability methods. J. Eng. Mech. **141**(3), 04014134 (2014)
9. Au, S.-K., DiazDelaO, F.A., Yoshida, I.: Bayesian updating and model class selection with Subset Simulation. arXiv preprint arXiv:1510.06989 (2015)
10. Green, P.: Bayesian system identification of a nonlinear dynamical system using a novel variant of simulated annealing. Mech. Syst. Signal Process. **52**, 133–146 (2015)
11. Beck, J.L., Au, S.-K.: Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. J. Eng. Mech. **128**(4), 380–391 (2002)
12. Ching, J., Chen, Y.-C.: Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging. J. Eng. Mech. **133**(7), 816–832 (2007)
13. Gilks, W.R., Berzuini, C.: Following a moving target-Monte Carlo inference for dynamic Bayesian models. J. R. Stat. Soc. Ser. B Stat. Methodol. **63**(1), 127–146 (2001)
14. Smith, A.F, Gelfand, A.E.: Bayesian statistics without tears: a sampling–resampling perspective. Am. Stat. **46**(2), 84–88 (1992)
15. Beck, J.L., Zuev K.M.: Asymptotically independent Markov sampling: a new Markov chain Monte Carlo scheme for Bayesian inference. Int. J. Uncertain. Quantif. **3**(5), 445–474 (2013)
16. Neal, R.M.: Probabilistic inference using Markov chain Monte Carlo methods Technical report CRG-TR-93-1, Department of Computer Science, University of Toronto, Toronto, CA (1993)
17. Hol, J.D., Schön, T.B., Gustafsson, F.: On resampling algorithms for particle filters. In: 2006 IEEE Nonlinear Statistical Signal Processing Workshop, pp. 79–82. IEEE, Piscataway, NJ (2006)
18. Douc, R., Cappé, O.: Comparison of resampling schemes for particle filtering. In: Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, 2005, pp. 64–69. IEEE, Piscataway, NJ (2005)
19. Gordon, N.J., Salmond, D.J., Smith, A.F.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In: IEEE Proceedings F Radar and Signal Processing, vol. 140(2), pp. 107–113 (1993). IET

20. Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T.: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. Signal Process. **50**(2), 174–188 (2002)
21. Doucet, A., Johansen, A.M.: A tutorial on particle filtering and smoothing: fifteen years later. In: Handbook of Nonlinear Filtering, vol. 12, pp. 656–704. Oxford University Press, Oxford (2009)
22. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller A.H., Teller, E.: Equation of state calculations by fast computing machines. J. Chem. Phys. **21**(6), 1087–1092 (1953)
23. Hastings, W.K.: Monte Carlo sampling methods using Markov chains and their applications. Biometrika **57**(1), 97–109 (1970)
24. Robert, C., Casella, G.: Monte Carlo Statistical Methods. Springer Science and Business Media, New York (2013)
25. Girolami, M., Calderhead, B.: Riemann manifold Langevin and Hamiltonian Monte Carlo methods. J. R. Stat. Soc. Ser. B Stat. Methodol. **73**(2), 123–214 (2011)
26. Nocedal, J., Wright, S.: Numerical Optimization. Springer Science and Business Media, New York (2006)
27. Vakilzadeh, M.K., Yaghoubi, V., Johansson, A.T., Abrahamsson, T.: Manifold Metropolis adjusted Langevin algorithm for high-dimensional Bayesian FE. In: 9th International Conference on Structural Dynamics (EURODYN), Porto, 30 Jun–02 Jul 2014, pp. 3029–3036 (2014)
28. Bui-Thanh, T., Ghattas, O., Martin, J., Stadler, G.: A computational framework for infinite-dimensional Bayesian inverse problems Part I: the linearized case, with application to global seismic inversion. SIAM J. Sci. Comput. **35**(6), A2494–A2523 (2013)
29. Martin, J., Wilcox, L.C., Burstedde, C., Ghattas, O.: A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion. SIAM J. Sci. Comput. **34**(3), A1460–A1487 (2012)
30. Haario, H., Saksman, E., Tamminen, J.: Adaptive proposal distribution for random walk Metropolis algorithm. Comput. Stat. **14**(3), 375–396 (1999)

Chapter 31
Model Calibration with Big Data

Guowei Cai and Sankaran Mahadevan

Abstract The aim of this paper is to address computational efficiency in model calibration in the presence of big data. Model calibration refers to the adjustment of model parameters so that the model output matches well with the observation data. Improvement of sensing technology makes large volume data available, which challenges the capability of current Bayesian calibration techniques, especially when model runs are expensive. Often researchers have used reduced-order models or surrogate models in calibration to save the computational cost; however, this leads to loss of accuracy. This paper applies the MapReduce technique to parallelize the model calibration, in order to make the computation more efficient without lowering the accuracy.

This paper applies MapReduce in two steps of the Bayesian calibration process: (1) surrogate model construction, and (2) computation of the posterior distribution of the calibration parameters. In order to build a high quality surrogate model, many training runs should be performed. Therefore the collection of training data is first parallelized in this paper using MapReduce. After the surrogate model is obtained, one commonly used Bayesian updating methods is considered for computing the posterior distribution of the calibration parameters, namely, particle filtering (PF). This paper uses MapReduce to parallelize both steps, and implements them in the Spark platform. The methodology is illustrated using one numerical example. The numerical example is the calibration of the thermal conductivity of concrete, with field temperature observed from infrared thermography (IR) camera. In this example, the large volume of data occurs in the spatial domain.

Keywords Big data analytics • MapReduce • Bayesian calibration • Finite element analysis • Surrogate model

31.1 Introduction

Most civil engineering infrastructure is built using concrete, for performance, safety and economic reasons. However, concrete structures suffer different types of degradations during their life: physical damage such as freeze-thaw cycling or irradiation; chemical degradation such as chloride penetration or alkali-silica reaction (ASR); and mechanical damage such as impact or fatigue [1]. Damage growth can be influenced by the structure's usage and environment, as well as the properties and current state of the structure. Prognosis and risk analysis of the structure is the primary goal of structural health monitoring. An accurate model will be needed for this purpose. Model calibration is required since sometimes the parameters are not exactly known, and needed to be calibrated based on the comparison between model output and observation.

Model calibration refers to the adjustment of model parameters so that the model output matches well with the field data. When full field observations are available (spatially or temporarily), different options are available for calibration. The most natural thought is to calibrate parameters using only some locations. For example, McFarland & Mahadevan [2] used only temperature observation at the final time step to calibrate thermal parameters. Madsen [3] estimated parameters of hydrological catchment model using observations from multiple locations in a 440 km^2 area. Lefevre et al. [4] calibrated thermal conductivity for hot wire based dc scanning thermal microscopy by measurements from different tip temperatures. Some researchers turned to perform model calibration use dimension reduction methods. For example, Higdon et al. [5, 6] used basis representations (e.g., principal components) to reduce the dimensionality of the problem and speed up the computations required for exploring the posterior distribution. Higdon et al. [5, 6] also used singular value decomposition (SVD) to reduce the dimension. On the other hand, some researchers applied full field measurements to update the model

G. Cai
Department of Civil Engineering, Vanderbilt University, Nashville, TN, 37235, USA

S. Mahadevan (✉)
Department of Civil and Environmental Engineering, Vanderbilt University, 400 24th Avenue South, Nashville, TN 37212, USA
e-mail: sankaran.mahadevan@vanderbilt.edu

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_31

parameters. For example, Roux & Bouchard [7] calibrated a ductile damage model using measurements from the full displacement field.

It can be observed that the main reason that researchers choose not to use full field observations to do parameter calibration is due to computational cost. However, the price people must pay is the loss of information and accuracy, since such a strategy implies that the model parameters do not vary over space and time. For the general, heterogeneous case where model parameters vary over space and time (e.g., material properties), full-field calibration would be high dimensional. Since calibration using full field observations is time consuming, parallel and distributed computing can help to reduce the time cost of data analytics, without causing any accuracy loss. A parallel computing approach in the context of big data is MapReduce. Utilizing a cluster of nodes, MapReduce performs two essential functions—it parcels out work to various nodes within the cluster or map, and it organizes and reduces the results from each node into a cohesive answer to a query [8].

To the authors' knowledge, only a few studies were found on the MapReduce application on model calibrations can be found. Humphrey et al. [9] parallelized calibration of parameters in watershed models, which was realized on Windows Azure cloud computing platform. Zhang et al. [10] realized cloud-based calibration of hydrologic model on Hadoop platform. The related researches only parallelized calibration process to certain application (hydrological model), also they did not handle large volume of observations. In this paper, a novel application of MapReduce on model calibration is presented. Here we focus on handling big data issue in model calibration. It is known that numerical models are sometimes too expensive to be repeatedly run in calibration process, which calls for the construction and use of surrogate models. The training point collection and the training of the surrogate model are also parallelized in this paper. The proposed methodology is general, and applies to variations over both space and time.

The rest of this paper is organized as follows. Section 31.2 provides a background review of the basic concepts related to model calibration and big data analytics. Section 31.3 develops the big data analytics approach for the calibration process. Section 31.4 implements the proposed approach using an illustrative example of thermal conductivity calibration of a concrete slab with holes using observations from infrared thermography, and discusses the performance of the MapReduce methodology. Section 31.5 provides concluding remarks.

31.2 Background

This section provides a review of basic concepts of model calibration, surrogate model, and big data analytics.

31.2.1 Bayesian Calibration

As we explained before, model calibration refers to the adjustment of model parameters so that the model output matches well with the field data. In practice, Bayesian calibration is popular due to its robustness. Here we will use Bayesian inference as an example technique of calibration. Bayes' rule describes the parameter update process:

$$f''(\theta) = \frac{L(\theta) f'(\theta)}{\int L(\theta) f'(\theta)\, d\theta} \tag{31.1}$$

Here θ represents the uncertain parameters of interest, Y represents the observations. Like the observations are continuous as well as the design parameters. Also $f'(\theta)$ is the prior density, $L(\theta)$ is the likelihood function, $f''(\theta)$ is the posterior density, and $\int L(\theta) f'(\theta) d\theta$ is the evidence.

Several numerical techniques are available to construct the posterior distribution, such as Markov Chain Monte Carlo simulation, and particle filter. The latter is used in this paper. In the particle filter approach, the weights of the particles (samples) define the probability distribution; these weights are updated with observations using the Bayes' rule above, to get the samples that correspond to the posterior distribution.

31.2.2 Gaussian Process (GP) Surrogate Model

Since Bayesian updating requires repeated runs of computer model, surrogate model is preferred to reduce the computational cost. Gaussian process surrogate model is chosen for this purpose [11]. Many types of surrogate modeling techniques are available in the literature, such as polynomial regression radial basis functions, Kriging (Gaussian process) and neural networks. The Gaussian process surrogate model is used in this paper for the sake of illustration.

A Gaussian process is completely specified by its mean function and covariance function and is a generalization of the multivariate normal distribution. We define the mean function $m(x)$ and the covariance function $K(x, x')$ of a random process $f(x)$ as $m(x) = E[f(x)]$ and $K[x, x'] = E[(f(x) - m(x))(f(x') - m(x'))]$ respectively. The process $f(x)$ can then be denoted as $f(x) \sim GP(m(x), K(x, x'))$. This definition implies a requirement that $(y_T, y_P) \sim N(m, K)$. In prediction, the joint distribution of the training outputs y_T and the prediction y_P is:

$$\begin{bmatrix} y_T \\ y_P \end{bmatrix} \sim N\left(\begin{bmatrix} m_T \\ m_P \end{bmatrix}, \begin{bmatrix} K_{TT} & K_{TP} \\ K_{PT} & K_{PP} \end{bmatrix}\right) \tag{31.2}$$

where T indicates training and P indicates prediction. The prediction (conditioned on the training points) $(y_P|y_T)$ follows a Gaussian distribution:

$$y_P \Big| y_T \sim N(m, S) \tag{31.3}$$

in which,

$$m = K_{PT} K_{TT}^{-1} y_T \tag{31.4}$$

$$S = K_{PP} - K_{TT}^{-1} K_{PT}^T \tag{31.5}$$

A number of common functions can be used as kernels to construct the covariance matrices. As an example, the commonly used squared exponential function is used here:

$$K(x_i, x_j) = \sigma^2 e^{-\frac{1}{2}\left(\frac{x_i - x_j}{l}\right)^2} \tag{31.6}$$

in which l is the length scale (which controls the correlation decay with distance) and σ^2 is the magnitude of variance. Based on the training data, these parameters can be estimated by the maximum-likelihood estimation (MLE) method.

31.2.3 MapReduce Framework

MapReduce is a framework designed for processing large datasets, by utilizing multiple nodes (machines) for the computations. It takes key/value pairs as inputs and generates the other key/value pairs as outputs. As mentioned earlier, the MapReduce framework can be split into two steps: map and reduce, both of which are created by the user. Before applying the MapReduce model, the user will need to write the input as the key/value pair. The key/value pair (k1, v1) will then be input to the map function, which will generate the intermediate key/value pairs (k2, v2). Then the intermediate key/value pairs are passed to the reduce function, which merges together these values to form a smaller set of values. The process is shown below, which allows us to handle lists that are too large to fit into the memory:

$$\text{map} \ \ (\text{k1}, \text{v1}) \geq \text{list}\,(\text{k2}, \text{v2})$$

$$\text{reduce} \ \ (\text{k2}, \text{list}\,(\text{v2})) \geq \text{list}\,(\text{v3})$$

A cluster of computers (nodes) are used to implement this framework (Fig. 31.1). One of them is the master node and the others are slave nodes. As shown in Fig. 31.1, the master node talks to the user program, and assign the tasks to the slave nodes (workers). First, the input files are parsed and split into smaller pieces (size 16–64 MB). The master will pick the idle workers and assign each with a map task or reduce task. Then each worker will do its own task and when all tasks completed, the output files will be obtained and synthesized.

While there are different implementations of MapReduce, Apache Spark [12] is the one chosen in this study. Spark is an open source cluster computing framework, and API (Application Program Interface) for Java, Scala and Python are available, which is convenient for non-computer science programmers. Beside the basic capability of using the MapReduce methodology, Spark employs Resilient Distributed Datasets (RDD) that enable efficient data reuse in a broad range of applications.

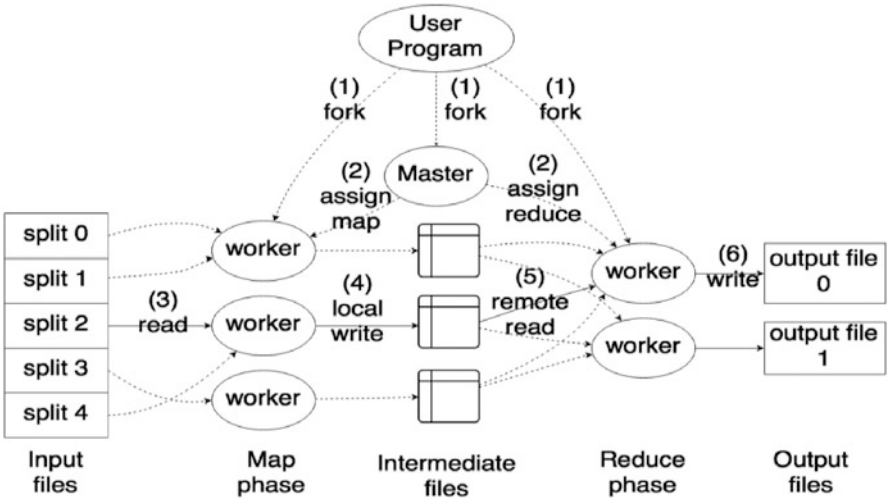


Fig. 31.1 MapReduce execution overview

31.3 Proposed Methodology

This section will describe the application of MapReduce to do model calibration in the context of big data. As we explain before, a physical model will be used to do prognosis, whose parameters are needed to be calibrated sometimes. To save computational cost, surrogate model will be applied in the process of calibration. However, the preparation of training points for surrogate model is still expensive. So MapReduce will be used to parallelized the running of physical model to generate of training data. After the training data are prepared, the surrogate model of training process can also be parallelized. The third level parallelization would be handling big data issue in calibration, which is the likelihood functions calculation. Last, PF can also be parallelized as the actual training process. So there are totally four parallelization happening in the whole calibration process.

The basic idea of the application of processing in MapReduce is to divide the files into different partitions (each partition contains multiple files), and then perform the mapping and reducing operations separately. To fully use the resources, the number of partitions is always greater than the number of instances (i.e. cores, of which each node might contain a multiple). For example, if the number of files to be analyzed is 100, and the number of cores available is 20, the number of partitions should be at least 20. Otherwise some of the cores will be idle. For each task, a separate processed result is obtained, without combination. Pseudocode 1 shows the implementation of MapReduce.

Pseudocode:

```
function mapper(x):
     function InputData = ReadData(x);
     function OutputData = Processing(InputData);
     function WriteData(OutputData);
     return (x, 0)

SparkContext(appName="myApp").parallelize(range(N), N).map(mapper).count()
```

The above pseudocode has two steps. First, a Map function is defined (‘mapper’), within which all the actual data processing functions are defined (reading, filtering, writing). The argument ‘x’ is the task (can be FEA runs, surrogate model construction, likelihood calculation and calibration) to be analyzed, which is assigned by the task manager. As discussed previously, since there is only the Map function, the input file can be mapped with any value (here we mapped x to 0). The second step, SparkContext, represents the connection to the cluster, which is the main class in Spark; “parallelize” is the method to split the input files into N partitions; and “map” is the method to call the Map function defined in the first step and to pass the input file to it. The “count” method is used to count the number of outputs. The number of outputs is not of interest, since the result has already been obtained in the mapper function. All four levels of parallelization can be realized using the pseudocode proposed above.

In summary, this paper proposes a framework of performing model calibration with large volume field data. Instead of simply parallelize only the parameter updating process, this paper considers all tasks related to model calibration. MapReduce

in four steps: [9] multiple runs of the original physics model to generate training points for an inexpensive surrogate model, [1] training of the surrogate model to be used in calibration, [8] construction of likelihood function for large volume observations, and [5] Bayesian calibration using the surrogate model and likelihood function.

31.4 Numerical Example

This example illustrates the proposed methodology for big data analytics in model calibration. The purpose of this example is to calibrate the thermal conductivity of concrete structure. To be more general, concrete structure with damage is considered, where the damage is simulated by drilled holes (Fig. 31.2). Holesof 1/2, 3/8, and 5/16 in. diameter (all of them 4.45 in. deep) were drilled into the side of a 15.5 × 15.5 × 2 inch concrete slab. Here we heated the slab from the bottom, and observe the top surface temperature by infrared thermography camera.

Once the structure becomes damaged, the parameter of interest (thermal conductivity) becomes heterogeneous, thus we need to use different conductivity coefficient values at different locations in order to use in future prognosis of the structure. In a realistic structure, the damaged area could be quite irregular; thus an averaged value or a structured representation of property variation may not be accurate. As a result, we may need to discretize the entire domain into many sub-domains (consistent with the FEA model) and calibrate the property for each sub-domain. In that case, calibration becomes a high dimensional problem if many sub-domains need to be considered.

31.4.1 Experimental Setup

The mechanics of damage detection using infrared thermography is based on the differences in heat transfer properties of different materials. The air in the drilled holes in the structure has much lower thermal conductivity coefficient than concrete, which will lead to a 'lagging' phenomenon, i.e., the heating and cooling time of the hole are slower than the surrounding solid region. Each thermal cycle has a total duration of 70 min. The heating profile is shown in Fig.31.2.

31.4.2 Data Processing

The cropped image is shown in Fig. 31.3. The image shows boundary effects, where additional heat may be introduced from the area around the slab, since the reflective material may not block all of the heat from the thermal blanket. It is seen that there is a heterogeneity of the concrete slab.

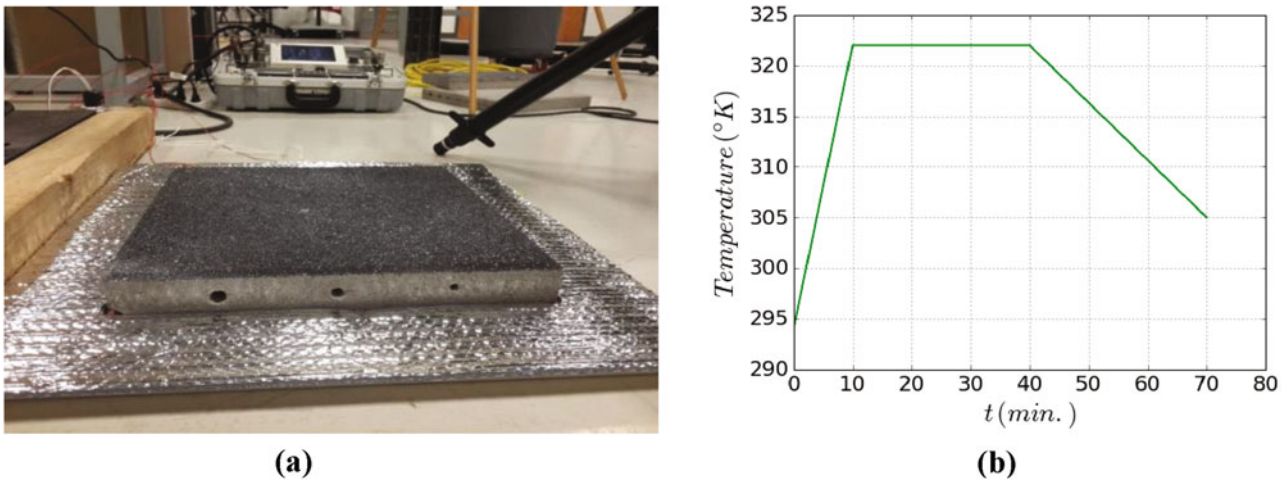


Fig. 31.2 The specimen to be monitored and thermal loading time history (scaled values)

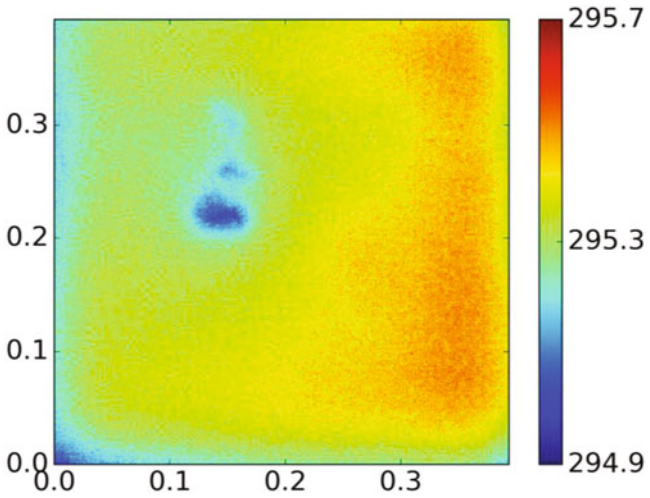


Fig. 31.3 Cropped image ($t = 60\ s$)

31.4.3 Finite Element Model

The physics modeling of the concrete structure is done using an FEA model, and implemented in commercial software Abaqus. Instead of performing the calibration by observation on each top surface node, for the purpose of illustration, a coarse grid (21×21) for calibration is used. For each calibration block, the thermal conductivity is considered to be constant. We assume there are 400 ($=20 \times 20$) observation points in the top surface, and temperature value on each observation point is effected by the four neighboring blocks.

Since we know the FEA model is too expensive for Bayesian calibration, we use a surrogate model to replace it. Training points of the surrogate model can be obtained via design of experiment (DOE) methods. Here Latin-hypercube design is chosen, and 5 DOE points are obtained for conductivity of each block. Note that for each observation point, the number of DOE points are 625 ($= 5 \times 5 \times 5 \times 5$).

Since in each FEA run, the temperature at all locations can be obtained at the same time, the total number of FEA runs will be 625. These 625 runs can be parallelized via MapReduce process described Sect. 31.3.

31.4.4 Surrogate Model Training

Based on the inputs (DOE points) and outputs (nodal temperature value in each run), the GP surrogate model can be obtained. For each FEA output, we will have a series of output for 70 time steps (70 min). For each spatial location i, if we create a surrogate model for each time step, we will lose the correlation between each time step. In order to capture the correlation, and also reduce the dimension, singular value decomposition (SVD) is applied. After SVD, outputs (625×70) at each spatial location i will be decomposed to a multiplication between coefficients and bases. Two bases are used here as an example, and the corresponding coefficient for each DOE output will have a dimension of 1×2.

Since for each spatial location, the bases are the same, one surrogate model will be created for each of the 2 coefficients, and the inputs are the 4 neighboring ks. Thus the total number of surrogate models will be 800 ($=4 \times 400$). The training of those surrogate model can be parallelized follow the procedure in Sect. 31.3.

31.4.5 Calibration

After all the surrogate models are trained for each spatial location, the parameter calibration can be performed by using Bayesian network. Figure 31.4 shows part of the network, in which ellipses are random variables, while squares are observations. Red ellipses are the random variables represent conductivity coefficients $\{k_j, j = 0, 1, \ldots, 440\}$ to be calibrated, while the yellow ones are random variables represent SVD coefficients for each spatial location $\{b_{0,i}, b_{1,i}, b_{2,i}, b_{3,i}, i = 0, 1, 399\}$,

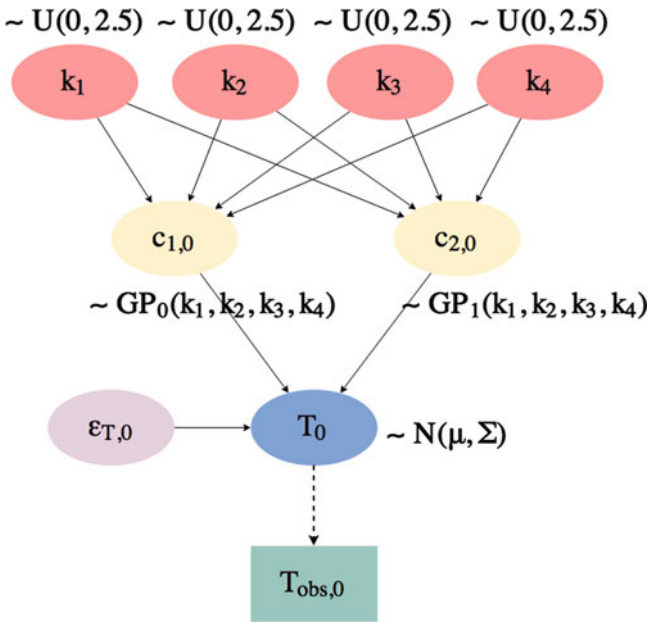


Fig. 31.4 Bayesian network for calibration (only shows the T_0 part)

which can be obtained from corresponding surrogate model. Each blue ellipse represents a temperature random variable T_i for a spatial location i, where $i = 0, 1, \ldots, 399$. Note here each $\{T_i\}$ follows a multivariate normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = [E[T_{i,l}]], l = 0, 1, \ldots, 69$ and $\boldsymbol{\Sigma} = [Cov[T_{i,l}, T_{i,m}]], l = 0, 1, \ldots, 69; m = 0, 1, \ldots, 69$. Here E means the expectation function and Cov means the covariance function. The likelihood function value given corresponding observation can be calculated in parallel following the procedure in Sect. 31.3. Finally, thermal conductivity coefficients at each spatial location can be updated via parallelization of PF.

31.5 Conclusion

This paper investigates the MapReduce technique to parallelize the model calibration process in a high-dimensional parameter space and in the presence of big data, in order to make the computation more efficient without lowering the accuracy. This paper investigates MapReduce in four steps of the model calibration process for mechanics problems: [9] multiple runs of the original physics model to generate training points to build an inexpensive surrogate model, [1] training of the surrogate model to be used in calibration, [8] construction of likelihood functions for large volume observations, and [5] Bayesian posterior construction (via the PF algorithm) using the surrogate model and likelihood function. The methodology is illustrated for the estimation of thermal conductivity at different locations in a damaged concrete structure, using data from infrared thermography (IR).

Acknowledgement This study was partly supported by funding from the U.S. Department of Energy (DOE) through the Light Water Reactor Sustainability (LWRS) Program (Monitors: Vivek Agarwal and Bruce Hallbert, Idaho National Laboratory). The support is gratefully acknowledged.

References

1. Stewart, M.G., Wang, X., Nguyen, M.N.: Climate change impact and risks of concrete infrastructure deterioration. Eng. Struct. **33**(4), 1326–1337 (2011)
2. McFarland, J., Mahadevan, S.: Multivariate significance testing and model calibration under uncertainty. Comput. Methods Appl. Mech. Eng. **197**(29), 2467–2479 (2008)
3. Madsen, H.: Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives. Adv. Water Resour. **26**(2), 205–216 (2003)

4. Lefevre, S., Volz, S., Saulnier, J.B., Fuentes, C., Trannoy, N.: Thermal conductivity calibration for hot wire based dc scanning thermal microscopy. Rev. Sci. Instrum. **74**(4), 2418–2423 (2003)
5. Higdon, D., Gattiker, J., Williams, B., Rightley, M.: Computer model calibration using high-dimensional output. J. Am. Stat. Assoc. **103**, 570–583 (2008a)
6. Higdon, D., Nakhleh, C., Gattiker, J., Williams, B.: A Bayesian calibration approach to the thermal problem. Comput. Methods Appl. Mech. Eng. **197**(29), 2431–2441 (2008b)
7. Roux, E., Bouchard, P.O.: On the interest of using full field measurements in ductile damage model calibration. Int. J. Solids Struct. **72**, 50–62 (2015)
8. Dean, J., Ghemawat, S.: MapReduce: simplified data processing on large clusters. Commun. ACM. **51**(1), 107–113 (2008)
9. Humphrey, M., Beekwilder, N., Goodall, J. L., Ercan, M. B.: Calibration of watershed models using cloud computing. In 2012 IEEE 8th International Conference on E-Science (e-Science), pp. 1–8. IEEE (2012)
10. Zhang, J., Qiu, H., Shamsabadi, S. S., Birken, R., & Schirner, G.: SIROM3—A Scalable Intelligent Roaming Multi-modal Multi-sensor Framework. In Computer Software and Applications Conference (COMPSAC), 2014 IEEE 38th Annual, pp. 446–455. IEEE (2014)
11. Rasmussen, C.E.: Gaussian Processes for Machine Learning. MIT Press, Cambridge (2006)
12. Zaharia, M., Chowdhury, M., Franklin, M.J., Shenker, S., Stoica, I.: Spark: cluster computing with working sets. HotCloud. **10**, 10–10 (2010)

Chapter 32
Towards Reducing Prediction Uncertainties in Human Spine Finite Element Response: In-Vivo Characterization of Growth and Spine Morphology

E.S. Doughty and N. Sarigul-Klijn

Abstract Computational modeling and simulation studies involving human spine using finite element methods typically encompass fixed geometry and material properties. The accuracy and reliability of the response obtained from these simulations are largely influenced by the input data from human spine. Input data uncertainties associated with natural variations in geometry and material characteristics of the human spine cause response to be imprecise. In addition if human is a pediatric human, growth rate uncertainties influence the predictions and their validations against experimental findings. The end result is the uncertainties in computationally predicted response. Some prediction uncertainties can be reduced by additional data collection and characterization. This paper presents a unique in-vivo noninvasive method and characterizes geometry and growth rate of human spine. The results will serve useful in reducing prediction uncertainties in finite element simulations of human spine. This study is the first extensive study of the pediatric spinal structure dimensions to include such a large variety of bony structures for this particular age group.

Keywords Data uncertainties • Pediatric growth rate • Spinal canal • Vertebral height

32.1 Introduction

Input data uncertainties associated with natural variations in geometry and material characteristics of the human spine cause computational simulation response to be imprecise. In addition if human is a pediatric human, growth rate uncertainties influence the predictions and their validations against experimental findings. Some prediction uncertainties can be reduced by additional data collection and characterization. More information can be found in References [1, 2] on the mathematical and statistical methods in analyzing uncertainty. As it is emphasized in these references that uncertainties can cause infeasibilities, false diagnosis and very imprecise prognosis if not correctly taken into account.

This paper presents a unique in-vivo noninvasive method and characterizes geometry and growth rate of human spine. The results will serve useful in reducing prediction uncertainties in finite element simulations of human spine. This study is the first extensive study of the pediatric spinal structure dimensions to include such a large variety of bony structures for this particular age group. Additionally, CT images provide insight into the transverse plane of the spine that was not available in the single-film radiograph studies. The current study provides growth and dimension information applicable in the 21st century where we have seen increases in growth due to enhanced nutrition information. The knowledge and data gathered in this study is useful for evaluating pediatric patient growth, as well as, developing finite element models of the pediatric spine at different ages.

The many structures in the human spine develop at different rates during growth. Few studies have been published describing the variance in structure size for the adult age group. Even fewer in-depth longitudinal studies describe or model the growth for these different structures for the pediatric age group [3–10]. Knowledge of the growth rates of the variable structures of the spine can help with determining normal growth for pediatric patients. Design of implants for the pediatric population would benefit from normal growth data of the various spinal structures. This information could also be used to create a finite element model that simulates growth from infancy to adulthood including the closing of growth plates, development of bony end plates, and changing of facet angle orientation.

E. Doughty • N. Sarigul-Klijn (✉)
Mechanical and Aerospace Engineering Department, Biomedical Engineering Department, University of California Davis, Davis, CA 95616-5294, USA
e-mail: nsarigulklijn@ucdavis.edu

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_32

The current gold standard for dimensioning anatomical structures is mechanical measurement of cadaveric specimens. The difficulty in acquiring pediatric post-mortem specimens has led to an absence of spine structure measurements for the skeletally immature age group. In lieu of cadaveric measurements, investigators have used medical imaging techniques to catalogue the dimensions of some structures [4, 5, 8, 10]. However, the existing literature is lacking in comprehensive data for the many structures of the pediatric spine.

Anthropometrics provide noninvasive and inexpensive measurement techniques of the spine but are limited to acquiring sitting height and approximate spinal curvature at various ages. A retrospective longitudinal study using radiographs to measure sagittal plane dimensions exists; however, the scans were acquired from the early 1900s and do not provide insight into current growth characteristic [9]. Additionally, plain radiographs are limited with single plane acquisition and depth of field. To remove this source of error and get information from the transverse plane, magnetic resonance imaging (MRI) studies have been performed to acquire information on spinal structures [5, 7]. While MRI studies avoid radiation exposure, computed tomography (CT) scans can create a three-dimensional illustration of the spine and highlight bony structures.

Due to the risks associated with radiation in the skeletally immature and the high cost of medical imaging, a longitudinal growth study of the human pediatric spine via CT scans has not been published. An even harder avenue for acquiring growth data is through pediatric cadavers. A high number of subjects would need to be collected at each age to produce statistically significant data. Therefore, the objective of this study was to document the growth of various spinal structures with non-invasive methods to create growth models describing the phases of human development.

32.2 Methods

Retrospective acquisition of CT scans resulted in 68 normal scans for cataloguing growth under the following criteria: male, 2- to 17-years of age, cervical-thoracic-lumbar segments of the spine, scans collected from 01/01/2005 to 01/17/2011, and exclusion of patients with Myelomeningocele and extreme abnormalities such as vertebral fusion, severe scoliosis, and fracture. Myelomeningocele is a common form of spina bifida found in 1% of U.S. births [11]. Owing to the growth differences between males and females, the scope of the study was limited to males only due to the larger availability of male scans. Due to the low subject number and variance between available scans for the different regions and ages, statistical analysis was not performed on the data. Instead the data is described with trends. When necessary to complete the data for a particular age, data was averaged from the surrounding ages.

For the entire spine, the curvature for each regional segment was recorded. Segment curvature was measured by finding the angle between the superior surface of the most superior segment vertebra and the inferior surface of the most inferior segment vertebra. For each vertebra the following features were measured: vertebral body height, spinal canal width and depth, and intervertebral disc height. Digital measurements were taken using medical imaging software (Mimics v11.11, Materialise, Leuven, Belgium).

32.3 Results

Age-dependent segment curvature equations for the cervical, thoracic, and lumbar regions of the spine are presented in Table 32.1. While the measurements varied, the general trend was an increase in cervical and lumbar lordosis with an increase in thoracic kyphosis. The cervical curvature showed a decrease from 2- to 8-years of age with an increase until maturity. These increases in curvature enhance the "S" shape of the spine with age as it succumbs to the force of gravity over time. The coefficient of determination (r^2) is used to describe the accuracy of a trend line to represent the given data. The r^2 value for the polynomial trends of the curvature data with the raw data shows wide spread, variable data in the cervical region ($r^2 = 0.28$) with an increasingly tighter fit of the data with the lumbar ($r^2 = 0.44$) and then the thoracic curvature

Table 32.1 Regression equations describing spinal curvature with age

Region	Curvature (degrees) with age in years	R^2
Cervical	$0.33*(age)^2–5.16*(age) + 26.78$	0.28
Thoracic	$2.45*(age) + 13.05$	0.74
Lumbar	$1.68*(age) + 12.29$	0.44

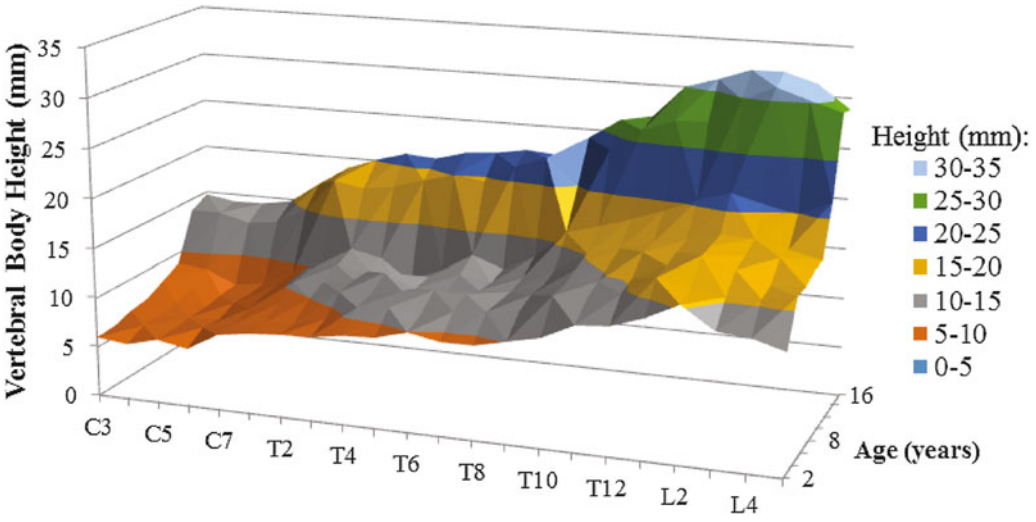


Fig. 32.1 Vertebral body height by age and vertebra

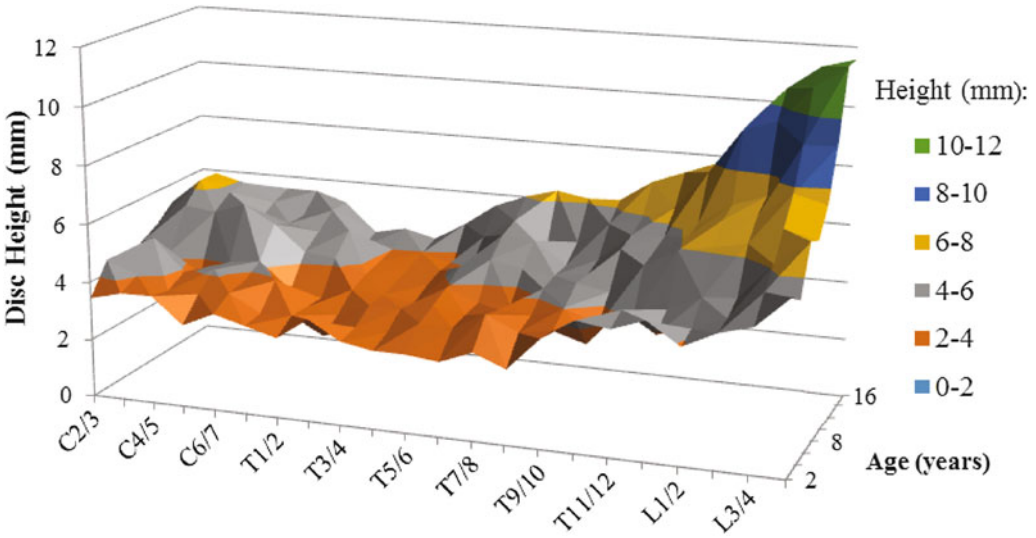


Fig. 32.2 Intervertebral disc height by age and spinal location

($r^2 = 0.74$). According to the linear trends, the thoracic and lumbar curvatures increase yearly by 2.4° and 1.7°, respectively. Linearization of the cervical lordosis trendline indicates a yearly increase in curvature of 2.7° from 8° to maturity.

Vertebral body height showed trends based off of location within the spine and age. A surface plot demonstrating these changes in height by vertebra and age is shown in Fig. 32.1. Non-linear growth occurs in the vertebral body during development. A large increase in vertebral body height occurs from 2- to 4-years of age followed by a slow increase in height from 4- to 12-years of age. A spurt in growth occurs from 12- to 16-years of age indicative of overall height changes seen during puberty.

Vertebral body height also varied by location. The cervical vertebrae had similar heights. An increase in height occurred from C6 to T3. From T3 to T8 there is minimal but steady increase in height. A steep increase in vertebral body height was seen from T8 to L2. The maximum body height occurred at the L2 and L3 vertebrae. Vertebral body height decreased from L3 to L5. In summary, the lumbar region had the tallest vertebral bodies followed by the thoracic region. The cervical vertebrae were the shortest.

Disc height showed trends with changes in age and spinal segment (Fig. 32.2). The discs in the lumbar segment have the greatest height while the shortest discs reside in the thoracic region. The trend line indicates a parabolic pattern in disc height by segment with a decrease in disc height from C3 to T4 followed by an increase in disc height from T4 to L5. For the 2-year

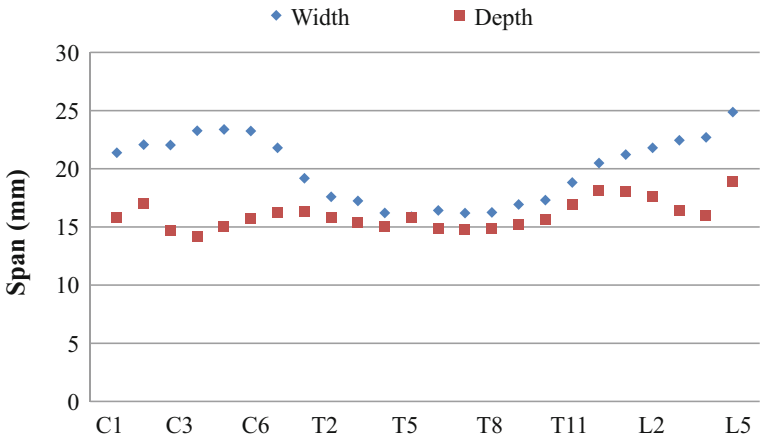


Fig. 32.3 Spinal canal width versus depth provides insight into the canal shape

old age group, the minimal disc height was 2.6 mm at T6/7 and the maximum disc height at L4/5 was more than doubled with 5.84 mm. For the 16-year old age group, the smallest disc was 4.2 mm thick at T3/4 and the L4/5 disc had a height of 11.6 mm.

An increase in disc height was seen for all segments as age progressed; however, the segments showed varied growth characteristics. The cervical discs do not show a distinct change in height from 2- to 8-years of age, but then jump up in size around 10-years old. The thoracic discs had a steady near-linear change in height during development. On the other hand, the lumbar discs grew at a more rapid rate, indicated by the higher slope in the linear region, and experienced a tapering off of disc height growth around 14-years of age. From 2- to 16-years of age, the lumbar disc doubled in height while the smallest disc, located in the upper thoracic region, only increased by 1.5 times.

The spinal canal did not show significant changes in size with age after 2 years. However, the spinal canal did show segmental trends in shape and area. The thoracic segment had the smallest and most uniform canal areas. The cervical and lumbar canals showed more variance in size with the largest canals occurring at the most proximal and distal regions of the spine on C1 and L5. The thoracic segment had no change in canal size with age. There was a slight increase in canal size in the cervical and lumbar segments with an increase in age.

In order get a better idea of the canal shape variance by segment, the width and depth for one was plotted as a function of vertebra in Fig. 32.3. In the cervical and lumbar segments, the width is greater than the depth indicating an elliptical canal that is wider in the medial-lateral plane than the anterior-posterior plane. The width and depth are very similar in the thoracic segment implying that the canal shape is circular. The thoracic spinal canal is narrower than the lumbar and cervical canals in the medial-lateral plane.

The segmental height was calculated by combining the vertebral height and intervertebral disc height for each spinal segment: cervical, thoracic, and lumbar. Due to the characteristic nature of the cervical spine and the unusual C1/C2 joint, the height for the cervical segment was calculated from the tip of the odontoid to the midpoint of the C7/T1 disc. The overall height is the summation of the cervical, thoracic, and lumbar segments. The segmental height is plotted against age in Fig. 32.4. The three primary growth phases are evident in the curves: fast growth from 2- to 4-years old, slow growth from 4- to 12-years old, and fast growth from 12- to 16-years old.

Growth models were created to describe the segmental vertical growth during the three primary growth phases (Table 32.2). The slopes of the linear models, 2- to 4-years old and 4- to 12-years old, show that the thoracic and lumbar segments are growing faster in the younger age group. The cervical segment does not show much change for the two age groups. When creating a linear relationship between cervical height and a 2- to 12-year old age group, the slope was 3.2 mm/year with r^2 of 0.97. Quadratic functions were used to model the growth from 12- to 16-years of age. The growth in this age group is very fast from 12 to 14 with a reduction in rate from 14 to 16. The rates of change of segmental height at 13, found by taking the first derivative of the growth model, were 18.3 mm/year, 47.1 mm/year, and 26.8 mm/year for the cervical, thoracic and lumbar segments, respectively. Compared to the first fast growth phase, the third growth phase has a much higher rate of change therefore growth is occurring fastest in the third growth phase from 12- to 16-years of age.

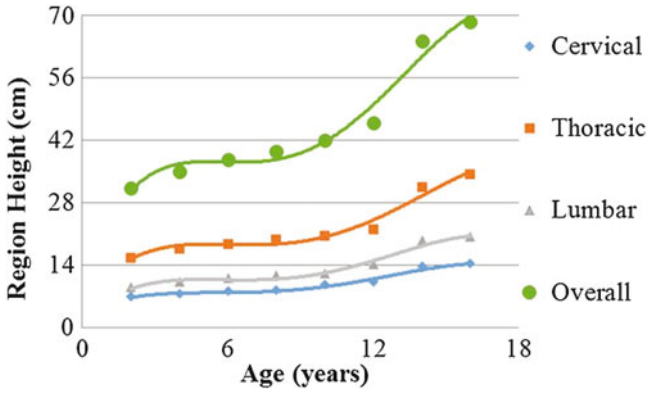


Fig. 32.4 Vertical growth by region

Table 32.2 Segmental growth models. 'y' is segmental height in mm. 'x' is age in years

Age (years)	Cervical height (mm)	Thoracic height (mm)	Lumbar height (mm)
2–4	$y = 2.9x + 62.24$	$y = 10.111x + 135.89$	$y = 6.3088x + 75.57$
4–12	$y = 3.3166x + 59.665$	$y = 5.1936x + 155.62$	$y = 4.4403x + 81.336$
12–16	$y = -3.9988x^2 + 122.22x - 790.95$	$y = -7.9774x^2 + 254.56x$ - 1686.1	$y = -5.6062x^2 + 172.56x - 1123.9$

32.4 Discussion

Linear regressions of all three segments showed an increase in curvature with age. The spine is S-shaped when viewed in the sagittal plane. The angles of lordosis in the cervical and lumbar segments and kyphosis in the thoracic segment change throughout development. The degree of lordosis and kyphosis increase as the spine matures progressing into a deeper S-shape [10–14].

Various measurement techniques have been used to collect segmental curvature data such as mechanical measurement with a kyphometer and dimensioning of medical images taken in the sagittal plane. The kyphometer requires manual manipulation for locating bony landmarks to identify the top and bottom of a spinal segment. Human error is a possibility with this measurement technique. Medical images are often collected while the patient is in a prone position which can reduce the "S" effect of curvature created under gravity. Contrary to the current results shown in Fig. 4.4, literature shows a decrease in thoracic curvature from 5- to 8-years of age [12]. The higher curvature at younger ages in literature could be due to the proportionality of the head size for young subjects when measurements are taken in an erect position. The subjects in the current study were in a prone position for scan acquisition resulting in some innate curvature changes compared to an erect position.

The few data points in the current study resulted in large percent error between the current data collected and the literature data for yearly curvature increase [12–16]. The yearly curvature increase predicted from the current study is higher than expected. The low sample size reduced the accuracy of the curvature data to be used as a predictive tool for the pediatric population. Many of the scans only contained a portion of a spinal segment, in order to make an accurate curvature measurement inclusion of the entire spinal segment in one scan was necessary.

The vertebral body showed continued growth following the three primary growth phases. The lumbar vertebrae obtained the largest vertebral height at maturity and the cervical vertebrae obtained the least. Maximum vertebral height occurred at approximately the L2/3 vertebrae. The lumbar vertebrae had the fastest growth rates during puberty compared to the other segments. On average, the vertebral body height at 16 was 2.3 times the height at 2-years of age.

In relation to the neighboring vertebral body height, the disc is larger in infancy compared to other points of development. This ratio decreases until 2-years of age [6]. Due to the small disc size in the thoracic segment, the thoracic ratios are higher showing that the disc height was small compared to the vertebral body height. The upper cervical segment has an almost 1:1 ratio with the disc size comparable to the vertebral body size. Similarly, the low ratios in the lumbar segment point to a larger disc since the lumbar vertebrae have the largest vertebral bodies. Changes in the ratio are indicative of differences in growth characteristics between the disc and the vertebral body. An increase in the ratio with age indicates that the vertebral body had a greater increase in height compared to the disc. On the other hand, a reduction in the ratio indicates that the disc had the greater increase in height. Overall, the vertebral body height to disc ratio decreased until 10- to 12-years of age then

sharply increased during puberty. This indicates that the discs were growing at a faster rate than the vertebral body until 12 and then at a slower rate than the vertebral body growth during puberty. Sze *et al.* had a lumbar vertebral body to disc ratio of 3.70 for the 1.5- to 2-year age group which is greater than the 3.42 ratio acquired for the L2 vertebral body height to L2/3 disc height in the current study [7].

The cervical spinal canal, measured in the transverse plane, had the largest area for the three segments of the spine, followed by the lumbar spine and, finally, the thoracic vertebrae had the smallest spinal canal area. The cervical and lumbar spinal canals were elliptical in shape and the thoracic spine had a circular canal. Cadaveric specimen observation supports the canal shapes determined in the current study [17]. The larger elliptical shape of the cervical spine allows room for the cervical enlargement of the spinal cord. The thoracic spine has minimal movement due to rib stabilization reducing the need for a larger canal. The thoracic spinal cord is also narrower than the cervical cord [17]. Due to the smaller canal area, the thoracic spine could be at a greater risk of spinal canal stenosis. There was not a significant change in spinal canal area with age which is supported by pediatric cadaveric observation which states that 90% of spinal canal growth is completed by late infancy, the starting age for the current study [18]. Due to the young age at which spinal canal growth is completed, pre- and post-natal health are considered to be important factors in determining final canal size development. Maternal smoking *in utero* has been linked to decreases in spinal canal size and associated back pain [18].

Two existing literature studies have published cervical spinal canal dimensions for the pediatric population [9, 10]. The results for the current study have larger measurement values. The Wang *et al*. study was a retrospective study performed on radiographs collected from the 1920s to 40s which could result in difficulty in acquiring measurement accuracy [10]. Vara and Thompson used sensitive and modern measurement techniques on cadaveric specimens to determine spinal canal dimensions [9]. However, they used specimens collected from the 1890s to 1938. The current study investigated pediatric patients in the 2000s. It is possible that the current pediatric population is larger in size than the pediatric population 100 years ago resulting in larger measurements for the current study when compared to these existing studies.

According to literature, there are three primary phases of growth. Phase 1 is a rapid growth phase from birth to 5 years of age. During this phase, the cervical, thoracic and lumbar segments grow by 3.3, 7 and 3 cm. Phase 2, from 5–10 years of age, is a slower growth phase with a 2 cm/year increase in sitting height [19]. The cervical, thoracic and lumbar segments see 1.5, 4 and 2.5 cm of growth between age 5 and 10. Phase 3 is a rapid growth phase that occurs from age 10 until skeletal maturity with the cervical spine gaining 3.5 cm in length, the thoracic gaining 7 cm and the lumbar gaining 3 cm. Fifty percent of skeletal growth is completed by age 2 and 66% by age 5. Eighty percent of the child's growth is completed by age 10 [19–21]. Comparatively, the current study measured the spinal height at 2-years old as 45% of adult spinal height, 55% at age 5, and 66% at age 10. These differences could be attributed to the greater average overall adult height acquired in the current study. Total spinal height for the adult age group averaged more than 100 mm taller in the current study than in literature. When the 2-, 5-, and 10-year old data from the current study are compared to the adult height from literature, they measure 53.9, 62.3, and 73.5% of adult values, respectively. These values are much closer to the literature values of 50, 66, and 80% [19, 20].

The rate of growth of the vertebral body increases as a function of thoracic spine level [6]. The differences within a single segment indicate that a high-fidelity growth model of the pediatric spine should incorporate individual growth models for each vertebra and structure. Literature states that the lumbar vertebrae grow twice as fast as the thoracic vertebrae from 2- to 16-years of age [17]. If the annual growth rate for the 4- to 12-year group for the current growth models is divided by the number of vertebra in a segment, each thoracic vertebra grew at a rate of 0.43 mm/year and each lumbar vertebra at 0.88 mm/year. The lumbar growth rate per vertebra in the current study was twice that of the thoracic vertebra.

32.5 Concluding Remarks

This paper presented a unique in-vivo noninvasive method and characterized geometry and growth rate of human spine. The results will serve useful in reducing prediction uncertainties in finite element simulations of human spine via reducing uncertainties in input data. This study is the first extensive study of the pediatric spinal structure dimensions to include such a large variety of bony structures for this particular age group. The non-invasive methods used to gather this data facilitated characterizing the many spinal structures of pediatrics eliminating the need to collect pediatric cadaveric spines. Additionally, CT images provide insight into the transverse plane of the spine that was not available in the single-film radiograph studies. Existing cadaveric measurements have been acquired from specimens collected at the turn of the 20th century. The current study provides growth and dimension information applicable in the 21st century where we have seen increases in growth due to enhanced nutrition information.

The knowledge and data gathered in this study is useful for evaluating pediatric patient growth, as well as, developing finite element models of the pediatric spine at different ages. The results will serve useful in reducing prediction uncertainties in finite element simulations of human spine.

References

1. Lopez, I., Sarigul-Klijn, N.: Effects of dimensional reduction techniques on structural damage assessment under uncertainty. ASME. J. Vib. Acoust. **133**(6), 061008–061008-12 (2011). doi:10.1115/1.4003592
2. Lopez, I., Sarigul-Klijn, N.: A review of uncertainty in flight vehicle structural damage monitoring, diagnosis and control: Challenges and opportunities. Prog. Aerosp. Sci. **46**(7), 247–273 (2010)
3. Anderson, M., Hwang, S., Green, W.T.: Growth of the normal trunk in boys and girls during the 2nd decade of life: related to age, maturity, and ossification of the iliac epiphyses. J. Bone Joint Surg. Am. **47**, 1554–1564 (1965)
4. Hellsing, E., Reigo, T., McWilliam, J., Spangfort, E.: Cervical and lumbar lordosis and thoracic kyphosis in 8, 11, and 15-year-old children. Euro. J. Orthod. **9**, 129–138 (1987)
5. Knirsch, W., Keutz, C., Haffner, N., Langer, M., Kececioglu, D.: Normal values of the sagittal diameter of the lumbar spine (vertebral body and dural sac) in children measured by MRI. Pediatr. Radiol. **35**, 419–424 (2005)
6. Lord, M.J., Ogden, J.A., Ganey, T.M.: Postnatal development of the thoracic spine. Spine. **20**, 1692–1698 (1995)
7. Sze, G., Baierl, P., Bravo, S.: Evolution of the infant spinal column: evaluation with MR Imaging. Radiology. **181**, 819–827 (1991)
8. Taylor, J.R.: Growth of human intervertebral discs and vertebral bodies. J. Anat. **120**, 49–68 (1970)
9. Vara, C.S., Thompson, G.H.: A cadaveric examination of pediatric cervical pedicle morphology. Spine. **31**, 1107–1112 (2006)
10. Wang, J.C., Nuccion, S.L., Feighan, J.E., Cohen, B., Dorey, F.J., Scoles, P.V.: Growth and development of the pediatric cervical spine documented radiographically. J. Bone Joint Surg. **83-A**, 1212–1218 (2001)
11. Centers for Disease Control and Prevention (CDC): Spina bifida and anencephaly before and after folic acid mandate—United States, 1995–1996 and 1999–2000. MMWR Morb. Mortal. Wkly Rep. **53**, 362–365 (2004)
12. Kasai, T., Ikata, T., Katoh, S., Miyake, R., Tsubo, M.: Growth of the cervical spine with special reference to its lordosis and mobility. Spine. **21**, 2067–2073 (1996)
13. Kuo, Y.L., Tully, E.A., Galea, M.P.: Video analysis of sagittal spinal posture in healthy young and older adults. J. Manip. Physiol. Ther. **32**, 210–215 (2009)
14. Papadakis, M., Papadokostakis, G., Stergiopoulos, K., Kampanis, N., Katonis, P.: Lumbar lordosis in osteoporosis and in osteoarthritis. Eur. Spine J. **18**, 608–613 (2009)
15. Widhe, T.: Spine: posture, mobility and pain. A longitudinal study from childhood to adolescence. Eur. Spine J. **10**, 118–123 (2001)
16. Willner, S., Johnson, B.: Thoracic kyphosis and lumbar lordosis during the growth period in children. Acta Paediatr. Scand. **72**, 873–878 (1983)
17. Cramer, G.D.: Clinical anatomy of the spine, spinal cord, and ANS. Elsevier, St. Louis (2014)
18. Clark, G.A., Panjabi, M.M., Wetzel, F.T.: Can infant malnutrition cause adult vertebral stenosis? Spine. **10**, 165–170 (1985)
19. Dimeglio, A.: Growth in pediatric orthopaedics. J. Pediatr. Orthop. **21**, 549–555 (2001)
20. Lovell, W., Winter, R.: Lovell and Winter's Pediatric Orthopedics, 6th edn. Lippincott, Williams & Wilkins, Philadelphia (2005)
21. Smith, J.T.: The use of growth-sparing instrumentation in pediatric spinal deformity. Orthop. Clin. North Am. **38**, 547–552 (2007)

Chapter 33
Structural Damage Detection Using Convolutional Neural Networks

Nur Sila Gulgec, Martin Takáč, and Shamim N. Pakzad

Abstract Detection of the deficiencies affecting the performance of the structures has been studied over the past few decades. However, with the long-term data collection from dense sensor arrays, accurate damage diagnosis has become computationally challenging task. To address such problem, this paper introduces convolutional neural network (CNN), which has led to breakthrough results in computer vision, to the damage detection challenge. CNN technique has the ability to discover abstract features which are able to discriminate various aspect of interest. In our case, these features are used to classify "damaged" and "healthy" samples modeled through the finite element simulations. CNN is performed by using a Python library called Theano with the graphics processing unit (GPU) to achieve higher performance of these data-intensive calculations. The accuracy and sensitivity of the proposed technique are assessed with a cracked steel gusset connection model with multiplicative noise. During the training procedure, strain distributions generated from different crack and loading scenarios are adopted. Completely unseen damage setups are introduced to the simulations while testing. Based on the findings of the proposed study, high accuracy, robustness and computational efficiency are succeeded for the damage diagnosis.

Keywords Structural health monitoring • Damage detection • Sensitivity analysis • Convolutional neural networks • Deep neural network

33.1 Introduction

The main difficulty of the structural damage detection is defining the unknown relation between the measurements and damage patterns. For this reason, damage diagnosis has been a challenging inverse problem in structural health monitoring for last few decades. In the 1990s, neural networks were proposed as a remedy for such poorly defined problems [1], and since then, they have been practiced to diagnose damages from the measurement data or its features without estimating the structure characteristics [2]. Literature has some studies which feed the neural network by using the inputs such as modal analysis of vibration response [3, 4], statistical parameters of vibration data [5], frequency response functions (FRFs) [6], wavelet transform coefficients of the acceleration data [2], and static displacements [7, 8]. In some studies, uncertainties in measurements are employed to improve the network architecture [9, 10]. Nevertheless, as mentioned in [11], in order to use the full potential of neural networks, researchers need to have more complex architectures. This complexity brings computational difficulty with it, when long-term data collection from dense sensor arrays are considered [12–14].

Proposed study addresses all these issues by developing a convolutional neural network (CNN) for damage identification. First of all, instead of creating hand-designed features like in traditional methods, our method allows the network to learn its damage features from the raw input. In other words, our study is capable of learning sophisticated damage features and create complex classifier boundaries. Secondly, by training a variety of loading cases, damage scenarios, and measurement noise levels, the proposed architecture accomplishes detecting damages even from the unseen damages with high accuracy. Lastly, the computational needs of the methodology are decreased by exploiting CNN's shared parameterization and GPU's massively parallel architecture.

N.S. Gulgec (✉) • S.N. Pakzad
Department of Civil and Environmental Engineering, Lehigh University, Imbt Labs, 117 ATLSS Drive, Bethlehem, PA 18015, USA
e-mail: nsg214@lehigh.edu

M. Takáč
Department of Industrial and Systems Engineering, Lehigh University, Harold S. Mohler Laboratory, 200 West Packer Avenue, Bethlehem, PA 18015, USA

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_33

The rest of the paper is organized as follows. First, a brief explanation of CNNs is provided in Sect. 33.2; then, the data preparation is described in Sect. 33.3. In Sect. 33.4, description of proposed CNN architecture and the main findings are presented. Conclusions and future directions are given in Sect. 33.5.

33.2 Deep Neural Networks

Machine learning (ML) is a scientific field that is gradually developed from pattern recognition [15] (recognition of patterns and regularities in data) and learning theory in artificial intelligence. Deep learning (deep neural network or DNN) is a branch of machine learning which attempts to build a model using a deep graph organized in multiple linear layers and non-linear transformations [16]. In Fig. 33.1, a small DNN which consists of an input layer, two hidden layers, and an output layer is shown. The input layer feeds the input instance $x = (x_1, \ldots, x_p)^T$ to the network, and the output is the result of the network. In Fig. 33.1, each circle represents one neuron. An arrow represents a connection from the output of one neuron to the input of another. Each arrow is associated with a weight. The output of the neuron in a hidden layer is composed of a weighted sum of inputs composed with a non-linear mapping, e.g., sigmoid, tanh, or others.

For given input $x \in \mathbb{R}^p$, ML algorithms try to build a prediction function $\theta(x; w)$ parametrized by w (frequently called weights). In the simplest case this function is considered to be linear, i.e., $\theta(x; w) = x^T w$. After the family of the prediction function is fixed, a loss function, which measures the error between a prediction and the true value, is chosen. The most elementary loss function is $\ell(\theta(x; w), y) = \|\theta(x; w) - y\|^2$, where $y \in \mathbb{R}^c$ is the true observed value (sometimes called the label) of the input query x. Other commonly used loss functions are the so-called soft-max and cross entropy functions. It is assumed that there is a distribution of inputs and corresponding labels (X, Y). The learning problem then boils down to finding the best possible instance of the prediction function from the selected family, which amounts to finding the best possible values of the weights w. Mathematically speaking, one solves the following optimization problem [17–20]:

$$\min_w \mathbb{E}_{(X,Y)}[\ell(\theta(x; w), y)], \tag{33.1}$$

where the expectation is taken over the true distribution of inputs and labels (X, Y). Unfortunately, in practice, it is impossible to have exact knowledge about the true distribution. The common practice is to sample n data points $\{(x_i, y_i)\}_{i=1}^n$ (frequently called training data) from the unknown distribution, and minimize the empirical loss instead:

$$\min_w \frac{1}{n} \sum_{i=1}^n \ell(\theta(x_i; w), y_i). \tag{33.2}$$

Convolutional Neural Networks Convolutional neural networks (CNN) are one of the most widely used types of deep neural networks and has been a breakthrough in visual and speech recognition for the last few years. The framework of CNN was first proposed by LeCun et al. [21] to classify handwritten digits. It did not work well with the large-scale image and video classification due to the lack of computing power and large training data at that time [22]. With the introduction of a highly parallel programmable unit called graphics processing unit (GPU), large scale visual recognition has become more pronounced. With the development in computing power and large-scale hierarchical image database [23], CNN architectures

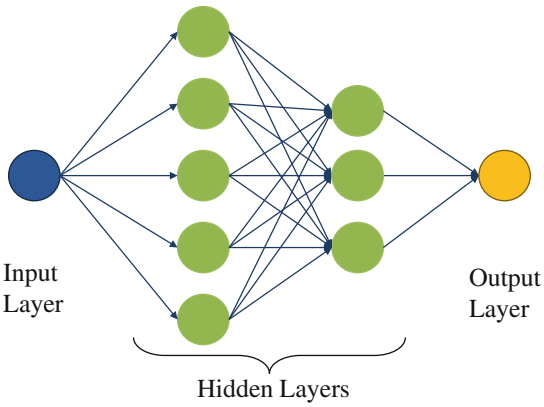


Fig. 33.1 A DNN with two hidden layers

kept evolving [24–26]. The performance improved drastically as the networks became more complex and deeper [27, 28]. The reason behind such success was the ability to keep temporal features of the input and using fewer parameters to reduce memory requirements [29].

Convolutional neural networks have three architectural frameworks such as local receptive fields, shared weights, and spatial sub-sampling [21]. Local receptive fields allow extracting multiple feature maps by sliding the same set of units all over the input. This property makes CNN robust to the translation and distortion of the input (i.e. the feature map will be shifted by the same amount of input shifting). Moreover, these feature maps use shared weights and biases, which reduce the learned parameters as well as the memory needs. Lastly, spatial sub-sampling reduces the resolution of the feature maps to avoid sensitivity of the outputs under shifts and rotations.

CNN architecture consists of three main types of layers:

1. **Convolutional layer** parameters are learnable weights and biases which are shared in the depth of the input. Feature maps are formed as these weights (or namely filters) are slid through the entire input, and the dot product between the filter and input are computed. This operation is called "convolution". Then, nonlinear activation functions activate the feature maps.
2. **Pooling layer** performs a downsampling operation (using e.g. maximum, average, sum operations) in the feature maps which reduce the dimensionality.
3. **Fully-connected layers** get the stacked convolutional layer outputs and compute the weighted sum of inputs composed with a non-linear mapping as described in Sect. 33.2.

33.3 Data Preparation

In order to have accurate damage diagnosis, the networks should be trained with well-known damage states [30]. Therefore, an analytic model of a steel gusset plate connection is simulated in ABAQUS by using shell elements. As shown in Fig. 33.2, the structural connection consists of two C8×11.5 channels and a steel plate ($28 \times 14 \times 1/4$ in.) welded together. Each connecting member is 20 in. long and has 8-in. overlap with the main plate. The model has the mesh size of 0.5 in. The behavior of the material is defined as elastic-perfectly plastic which has a yield strength of 36 ksi. Uniformly distributed load changing from 120 kips tensile to 100 kips compression is applied to the end of the channel members. Strain field in the direction of loading of the gusset plate is used to analyze the proposed methodology.

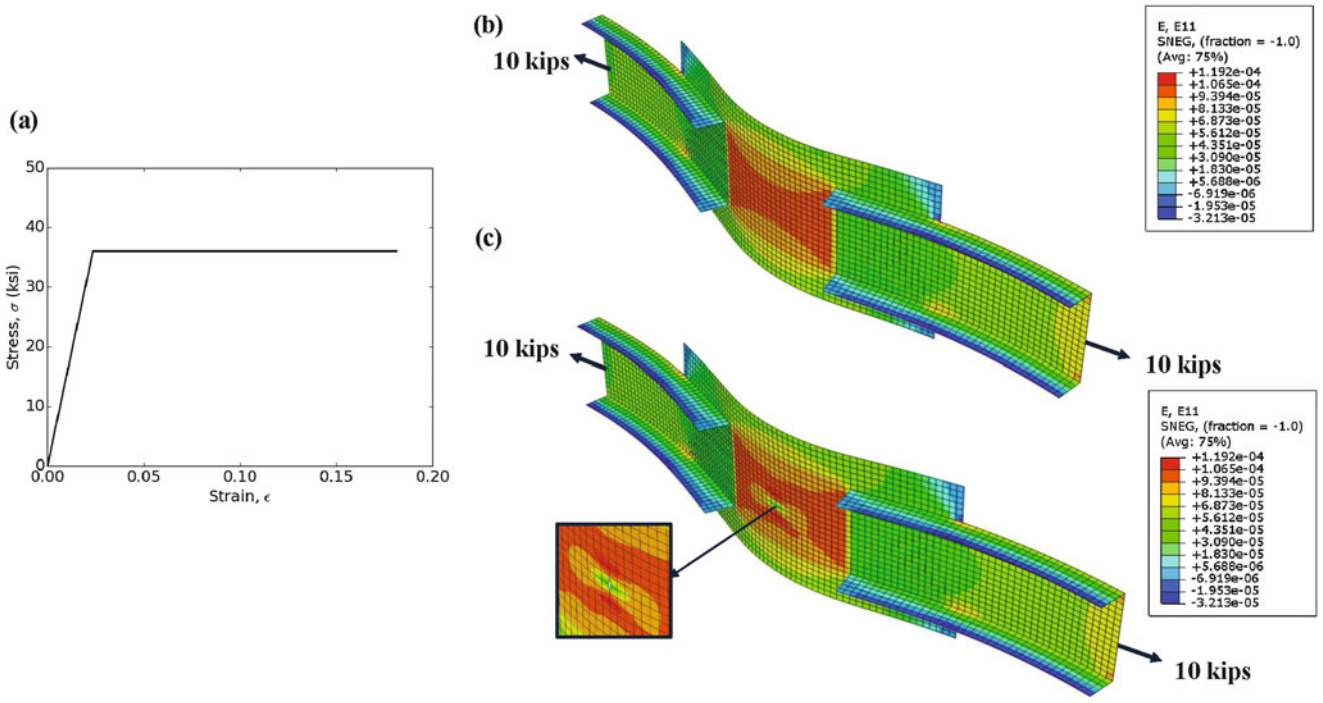


Fig. 33.2 (**a**) Material behavior; the setup of the (**b**) healthy and (**c**) single damaged gusset-plate

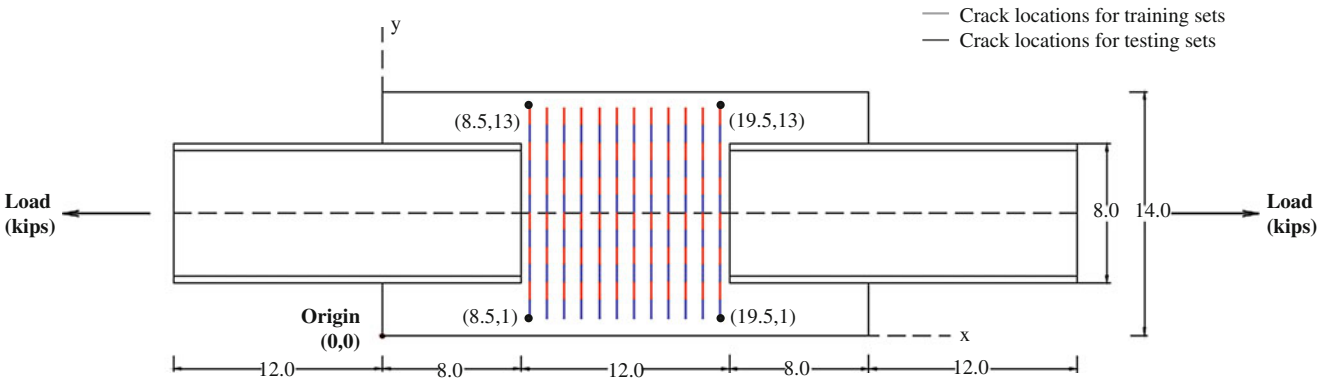


Fig. 33.3 The damage locations for training and test data sets

Variability in training, validation, and test data sets are formed by exploiting different loading cases, damage scenarios and noise levels. Damages are simulated as 1-in. long cracks and the crack locations are chosen randomly at the beginning of the analysis with a specified load in the range of $U[-100, 200]$ kips. The coordinates of the cracks changing between $(8.5, 1)$ and $(19.5, 13)$ are shown in Fig. 33.3. None of the coordinates of training samples are used in the testing samples to be able to evaluate the methodology with completely unseen damaged samples. Furthermore, inevitable uncertainty in the measurement and modeling processes is simulated as multiplicative noise in the samples. Total 24,000 "healthy" and 24,000 "single-damaged" unique samples are generated and distributed to the training, validation, and test data sets. The strain field of the gusset plate is represented as $28 \times 56 \times 1$ tensors as inputs to the network. These inputs are normalized by subtracting the data mean and dividing the data maximum.

33.4 Proposed Architecture for Damage Detection

33.4.1 CNN Architecture

The convolutional neural network can be built in various ways by using the sequence of convolutional layers (CONV), pooling layers (POOL) and fully connected layers (FC). Therefore, hyperparameter search mechanism should be performed to find the structure of optimal CNN architecture. In this study, a total of 50 networks are constructed with different hyperparameters such as learning rate, the number of convolutional layers, the number of FC layers, the size of hidden layers, filter and kernel sizes. The first run is performed for few epochs, then, 10% of the networks which have the worst test score are removed. The remaining networks are run for another set of epochs. The runs are computed until the best 10 networks remain. The variability in the hyperparameters are listed as: filter size: [(3x3), (5x5)], max-pooling size: (2x2), learning rate: $[2-2^{-8}]$, kernel size: $[2, 4, 8, 16, 32]$, the number of convolutional layer: $[1-3]$, the number of fully connected layer: $[1-3]$ and randomly selected hidden layer sizes.

The best network obtained as a result of the search mechanism is shown in Fig. 33.4. The network classifies 28×56 inputs of normalized strain distribution as "undamaged" and "damaged". The proposed architecture has two convolutional networks with kernels of $[16, 16]$ and filter sizes of [(5x5), (5x5)]. The feature maps of the last pooling layer, which are stacked together in an array, feed the fully connected layers with two hidden layers $[1388, 1212]$. Nonlinear function tanh() is used for the activation of the layers.

For a given input, prediction function can be parametrized by w (namely weights). In this study, the label of the final output is found by softmax classifier. The class i of the input x is predicted by choosing the maximum probability of softmax function defined as follows:

$$[\text{softmax}(\theta(x; w))]_i = \frac{e^{[\theta(x;w)]_i}}{\sum_j e^{[\theta(x;w)]_j}}, \qquad y_{prediction} = \underset{i}{\text{argmax}}([\text{softmax}(\theta(x; w))]_i). \tag{33.3}$$

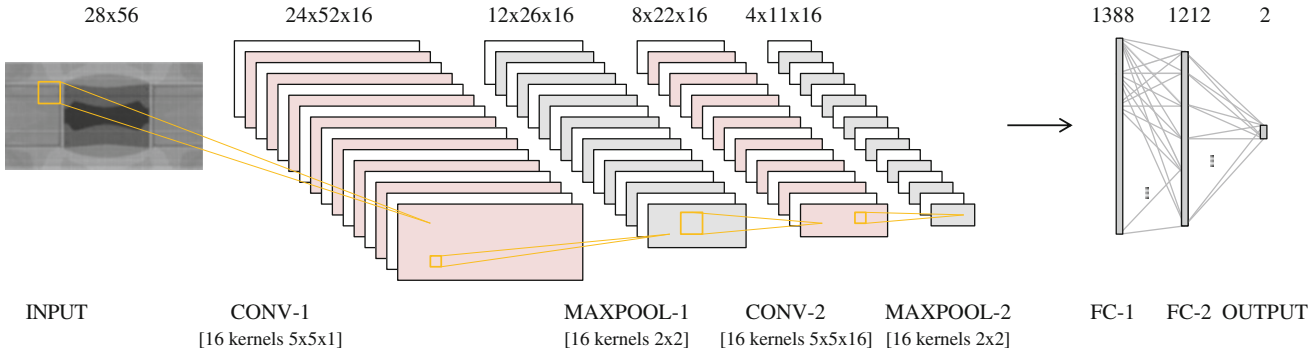


Fig. 33.4 Proposed CNN architecture

Table 33.1 Sensitivity analysis of the CNN network

	Noiseless	1% noise	2% noise
Testing error %	2.06	2.56	3.51
False negative ratio %	2.06	2.56	3.34
False positive ratio %	0.00	0.00	0.17

33.4.2 Training

Proposed study is trained by using a Python library called Theano to optimize the mathematical expressions including multi-dimensional arrays [31]. Higher performance of these data-intensive calculations is accomplished by NVIDIA Tesla K80 GPUs which enable parallelism for data processing. Furthermore, in order to reduce variations between the layers, weights are initialized by Xavier initialization for tanh function [32]. Weight initialization of ith layer is set to be in the uniform distribution in the interval $\left[-\sqrt{\frac{6}{n_{i-1}+n_i}}, \sqrt{\frac{6}{n_{i-1}+n_i}}\right]$ where n_{i-1} and n_i are the are the number of units in the $(i-1)$th and ith layer.

The practical advantage of choosing CNN as a model is that it can be trained efficiently using a stochastic gradient descent algorithm [33], where the gradients are computed by back-propagation [34]. The proposed model adopts the negative log-likelihood as the loss function of back-propagation process, where optimal model parameters θ^* are learned by maximizing the likelihood of the data set. In this study, stochastic gradient descent method is implemented to a batch size of 512 samples. Once the gradients are computed with stochastic gradient descent, the model is updated with the learning rate, $\eta = 0.033$. In order to prevent overfitting, validation set performance is monitored in every epoch. When model's performance is improved sufficiently on the validation set, no further optimization is implemented.

33.4.3 Results

Trained network is tested on previously unseen damage states which are not used for training procedure. The findings of the proposed CNN architecture is presented in this section. As shown in Table 33.1, this study identifies the unseen damages with 2.06% error. This error rate represents that the CNNs are capable of learning the damage features if enough training cases are provided. Furthermore, the introduction of different noise levels during the training helps network to learn damage features under uncertainty. As presented in Table 33.1, test error doesn't change drastically under the 1% and 2% multiplicative noise introduced to the samples. Therefore, our methodology is robust under the slight noise. The validation and test error sensitivity of the samples that are without noise and with 1% and 2% noise are also shown in Fig. 33.5.

As presented in Table 33.1, the majority of the misclassified observations belong to the damage status that is detected as "healthy". In SHM, incorrect classification of damaged samples (namely false negative) is not desirable; thus, the damage locations of misclassifications are further investigated. The gusset plate area shown in Fig. 33.3 is converted into a matrix having the number of false negatives in each damage location for test data set (Fig. 33.6). For example, the element in the first column of the first row represents the misclassified crack from (8.5, 11) to (8.5, 12); the element in the first column of the second row represents the misclassified crack from (8.5, 10) to (8.5, 11), and so on. The number of false negatives is also color coded as (green: < 5, yellow: 5–10, red: >10) to provide better visualization. As can be seen from Fig. 33.6, most

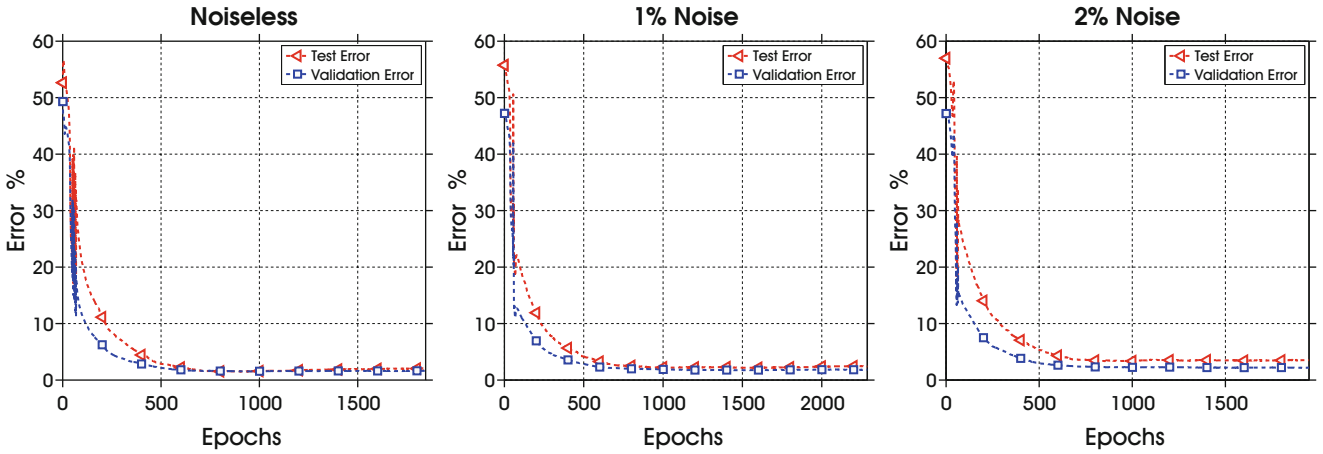


Fig. 33.5 Error sensitivity of the test samples with different noise levels

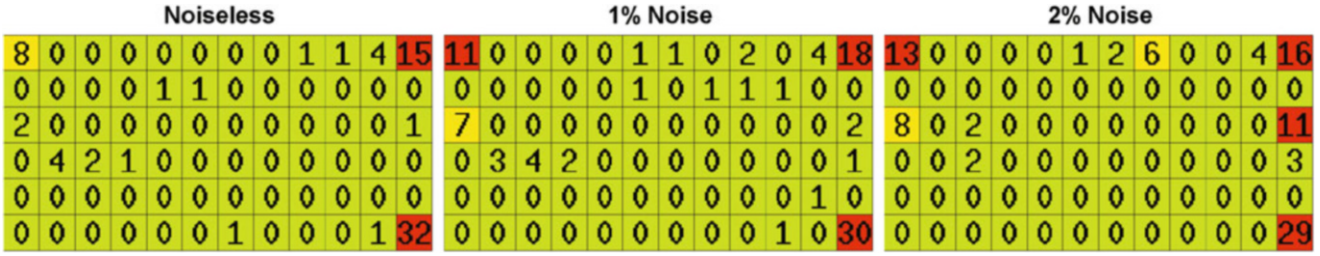


Fig. 33.6 The false negatives per each simulated crack location for test data set with 0% noise (*left*), 1% noise (*middle*) and 2% noise (*right*)

of the misclassifications occur in the regions where the stress concentration is lower compared to the center of the plate. Therefore, false negative samples do not fall in high-risk region where damages can propagate failure of the structure.

33.5 Conclusion

The major challenge of the structural damage detection is finding the unknown relation between the measurements and damage patterns. To address such problem, this paper introduces convolutional neural network (CNN) which has the ability to discover abstract features that are able to discriminate various aspect of interest. In our study, these feature maps are used to classify "damaged" and "healthy" cases modeled through the analytic simulations. The computational needs of the methodology are decreased by exploiting CNN's shared parameterization and GPU's massively parallel architecture.

Based on the results of the proposed study, high accuracy, robustness, and computational efficiency are succeeded for the damage diagnosis challenge. However, further research is necessary to be able to benefit more from convolutional neural networks. Future work aims to (1) determine the location and severity of the damages, in addition to the detection of damage existence, (2) reduce the number of false negatives by penalizing it, (3) test the network on the real experimental setup.

Acknowledgements Research funding is partially provided by the National Science Foundation through Grant No. CMMI-1351537 by Hazard Mitigation and Structural Engineering program, and by a grant from the Commonwealth of Pennsylvania, Department of Community and Economic Development, through the Pennsylvania Infrastructure Technology Alliance (PITA). Martin Takáč was supported by National Science Foundation grant CCF-1618717.

References

1. Flood, I., Kartam, N.: Neural networks in civil engineering. II: systems and application. J. Comput. Civ. Eng. **8**(2), 149–162 (1994)
2. Shi, A., Yu, X.-H.: Structural damage detection using artificial neural networks and wavelet transform. In: 2012 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA) Proceedings, pp. 7–11. IEEE, New York (2012)
3. Hearn, G., Testa, R.B.: Modal analysis for damage detection in structures. J. Struct. Eng. **117**(10), 3042–3063 (1991)
4. Hadzima-Nyarko, M., Nyarko, E.K., Morić, D.: A neural network based modelling and sensitivity analysis of damage ratio coefficient. Expert Syst. Appl. **38**(10), 13405–13413 (2011)
5. Shu, J., Zhang, Z., Gonzalez, I., Karoumi, R.: The application of a damage detection method using artificial neural network and train-induced vibrations on a simplified railway bridge model. Eng. Struct. **52**, 408–421 (2013)
6. Fang, X., Luo, H., Tang, J.: Structural damage detection using neural network with learning rate improvement. Comput. Struct. **83**(25), 2150–2161 (2005)
7. Szewczyk, Z.P., Hajela, P.: Damage detection in structures based on feature-sensitive neural networks. J. Comput. Civ. Eng. **8**(2), 163–178 (1994)
8. Zhao, J., Ivan, J.N., DeWolf, J.T.: Structural damage detection using artificial neural networks. J. Infrastruct. Syst. **4**(3), 93–101 (1998)
9. Bakhary, N., Hao, H., Deeks, A.J.: Damage detection using artificial neural network with consideration of uncertainties. Eng. Struct. **29**(11), 2806–2815 (2007)
10. Nazarko, P., Ziemiański, L.: Application of artificial neural networks in the damage identification of structural elements. Comput. Assist. Mech. Eng. Sci. **18**(3), 175–189 (2011)
11. Flood, I.: Towards the next generation of artificial neural networks for civil engineering. Adv. Eng. Inform. **22**(1), 4–14 (2008)
12. Yao, R., Pakzad, S.N., Venkitasubramaniam, P.: Compressive sensing based structural damage detection and localization using theoretical and metaheuristic statistics. Struct. Control Health Monit. **24**, e1881 (2017). doi:10.1002/stc.1881
13. Yao, R., Pakzad, S.N., Venkitasubramaniam, P., Hudson, J.M.: Iterative spatial compressive sensing strategy for structural damage diagnosis as a big data problem. In: Dynamics of Civil Structures, vol. 2, pp. 185–190. Springer, New York (2015)
14. Shahidi, S.G., Gulgec, N.S., Pakzad, S.N.: Compressive sensing strategies for multiple damage detection and localization. In: Dynamics of Civil Structures, vol. 2, pp. 17–22. Springer, New York (2016)
15. Bishop, C.M.: Pattern recognition. Mach. Learn. **128**, 1–58 (2006)
16. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
17. Shalev-Shwartz, S., Ben-David, S. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, New York (2014)
18. Nilsson, N.J.: Introduction to machine learning. An Early Draft of a Proposed Textbook (1996)
19. Smola, A., Vishwanathan, S.V.N.: Introduction to Machine Learning, pp. 32–34. Cambridge University, New York (2008)
20. Alpaydin, E.: Introduction to Machine Learning. MIT, Cambridge (2014)
21. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
22. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G.: Recent advances in convolutional neural networks. arXiv preprint, arXiv:1512.07108 (2015)
23. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L.: Imagenet: a large-scale hierarchical image database. In: IEEE Conference on Computer Vision and Pattern Recognition, 2009 (CVPR 2009), pp. 248–255. IEEE, New York (2009)
24. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems. Curran Associates Inc. (2012)
25. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint, arXiv:1409.1556 (2014)
26. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: European Conference on Computer Vision, pp. 818–833. Springer, Berlin (2014)
27. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
28. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. arXiv preprint, arXiv:1512.03385 (2015)
29. LeCun, Y., Bengio, Y.: Convolutional Networks for Images, Speech, and Time-Series. MIT Press, Cambridge (1995)
30. Elkordy, M.F., Chang, K.C., Lee, G.C.: Neural networks trained by analytically simulated damage states. J. Comput. Civ. Eng. **7**(2), 130–145 (1993)
31. Theano Development Team. Theano: a Python framework for fast computation of mathematical expressions. arXiv e-prints, abs/1605.02688, May (2016)
32. Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. In: Aistats, vol. 9, pp. 249–256 (2010)
33. Robbins, H., Monro, S.: A stochastic approximation method. Ann. Math. Stat. **22**(3), 400–407 (1951)
34. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Cogn. Model. **5**(3), 1 (1988)

Chapter 34
Experimental Model Validation of an Aero-Engine Casing Assembly

D. Di Maio, G. Ramakrishnan, and Y. Rajasagaran

Abstract This work presents an experimental model validation of an aero-engine casing assembly. The objective is to show how to perform an experimental model validation of a casing assembly with bolted flanges. One of the major challenges was given by the unknown geometrical dimensions and material properties of the structure at the outset. This goal could be reached by using modal updating so as to identify the best geometrical and material parameters for the model. Experimental testing was a key contributor in this process, since the test data were used to benchmark the updating process. The updating process highlighted that the model could not be correctly updated until the joints were included. Therefore, the FE model required to be upgraded, by inclusions of joint flexibility, before it could be updated. This paper will present both the upgrading-updating process and the experimental testing carried out on the casing assembly.

Keywords Upgrading • Updating • Experimental model validation

34.1 Introduction

This research work was developed using an aero-engine casing assembly the finite model of which was not available at the beginning. The hardware was used in past researches [1] where the model was developed and supported by the project sponsor. Unfortunately, this information was not available in this present research. So, the objective and challenge was to exploit the potential of the model updating to experimentally validate a FE model. Ewins et al. [2] suggested that any FE model is good as much as it includes the correct physics which enables the right calculation of the dynamic response. Ewins also suggests that a model must be upgraded if physical elements are not yet taken into account for the calculation of the dynamic response. So, any FE model can be updated as much as possible but it will inevitably fail to simulate the correct dynamics until all important flexible elements are included, for instance joints. Therefore, under these circumstances, a model should be upgraded and then updated. For example, irregular changes of thickness in the physical structure can be included into the model updating process where the mass and stiffness are corrected for better values. However, if a joint flexibility is not included the model must be upgraded before any updating process takes place. To prove this assumption, the cases presented in here will discuss about a FE model created with and without the inclusion of the joint flexibility. Considering that both geometrical and material properties were unknown the paper will show that the model can be updated up to find the best values for predicting some of the modes measured during the experimental campaign. But, only the inclusion of the joint flexibility into the model will deliver the correct response modes. The process might look trivial but the model updating software, used in this project, requires the complete modelling of the bolts in order to perform the correct updating. And, this type of modelling is not a trivial task when approx. 150 bolts must be included, which creates redundancy and increases the computational cost. Despite the entire work is based on linear dynamic response, so nonlinearity is not even attempted, the process to upgrade and update a linear model showed not to be such a straightforward process.

D. Di Maio (✉) • G. Ramakrishnan • Y. Rajasagaran
Department of Mechanical Engineering, University of Bristol, University Walk, BS8 1TR, Bristol, UK
e-mail: dario.dimaio@bristol.ac.uk

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_34

34.2 Experimental Setup

The test structure was suspended horizontally by using turn buckles and elastic chords as shown in Fig. 34.1. An electromagnetic shaker was also suspended and attached on the top part of the casing on the middle section. The dynamic range selected was between 150 and 350 Hz. The measurement points for the test were the same already selected for the validation carried out in [1]. The modal testing and analysis were carried out by LMS TestLAB (SIEMENS). The mode shapes are presented in the APEENDIX. The modal data were correlated with the ones obtained during a past research project in which the aero-engine casing was supported vertically rather than horizontally as in the presented case. Figure 34.2 shows the correlation map between the two sets of modal data. The map shows that the correlation does not achieve 100% for all the modes selected. The reason can be due to several factors but the most important seems to be the effect of the mass loading (due to the gravity) on the joints. In fact, the vertical configuration would equally load all the joints, which is not the case for the horizontal setup.

Fig. 34.1 Test setup

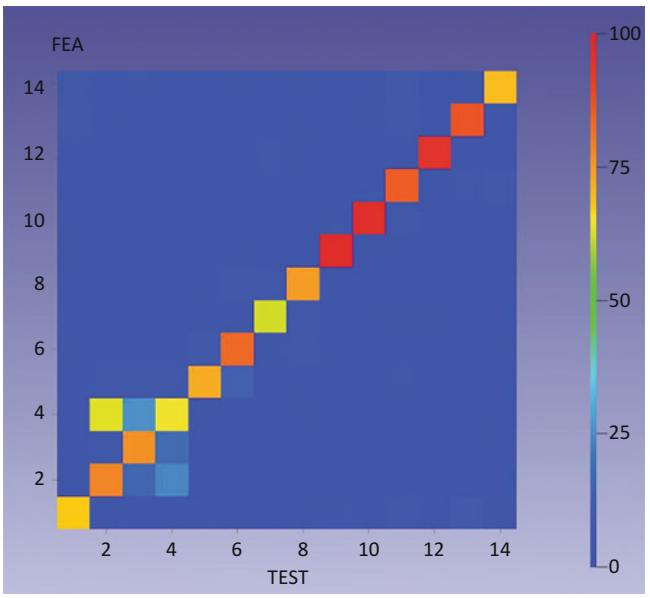


Fig. 34.2 Correlation between the modal data obtained from vertical (Y-axis) and horizontal (X-axis) setup

34.3 Finite Element Modelling of the Aero-Engine Casin Assembly

The aim of the model development was to create an FE model that was simple yet able to simulate the underlying linear behaviour of the test structure. To achieve this, the configuration and parameter errors in the model were needed to be improved. This was done by progressively increasing the complexity of the model. At each stage, mode shapes from the experimental data were matched to mode shapes from the FE model based on MAC and natural frequency difference (NFD) values. Lower NFD values are preferred as they indicate the model is able to accurately predict a mode shape at the right frequency. The model was then subjected to an updating procedure using FEMTools (DDS). Here, the software attempts to match the model results to the experimental by manipulating the parameters. The results from the updating procedures were the updated final parameters and mode shape pairs, with NFD and MAC values. Mode shape pairs in this section are defined as a mode shape pair correlated between the experiment data and the numerical model. These results were studied and a decision was made on the changes to be made to improve the model. This changed model was subjected to the same process until the ideal number of mode shape pairs with ideal NFD and MAC values were achieved: at least 10 mode shape pairs with the experimental data with MAC and NFD values above 50% and below 15% respectively. The frequency of interest was 150–350 Hz. Usually, MAC values of 80% or higher are preferred. However due to the size and complexity of the structure, even the slightest mis-match between an experimental and model node behaviour resulted in a MAC below 80%. In improving the model at each stage, the parameters and/or configuration were changed. The main variables involved were the material properties (Young's Modulus, (E), and density (ρ)), the thickness h of the model, and the physical make-up (configuration) of the model. These were dependent on one another. For example, the thickness and Young's Modulus affected the overall stiffness of the casing, and the density and thickness affected the overall mass. The configuration of the model affects the mass and stiffness as well. As such they cannot be changed independently. The initial objective was to identify the material properties. The configuration of the model was initially kept simple. It was also allowed to be flexible to fit the requirements of other parameters.

The aero-engine assembly was modelled by ABAQUS (Dassault Systems). The casings were weighted and the overall assembly was 256 Kg. The major issue was to identify the material properties of the assembly. Two materials were assumed by visually inspecting the casings, these were titanium and stainless steel. Clearly, these properties could not be retrieved by the standard libraries nevertheless the nominal values were used as guess for starting the model updating. Rough measurements of the dimensions of the casings' thicknesses were also made, as shown in Fig. 34.3. However, several features had to be neglected and assume the updating process to be able to cope with these, like the one showed in Fig. 34.4.

It is not scientifically interesting to explain the trial and error procedure to validate the model since it is rather a procedural thing. It is, instead, useful to explain that the several attempts achieved some level of agreement between the experimental and theoretical model. The error between the two was caused by lack of information about aforesaid properties, which the model updating managed to identify at its best as presented in Table 34.1.

At this stage all three casings were rigidly bonded together. This was one of the reasons why the updating could not improve the correlation beyond an acceptable level. For example as presented in Fig. 34.5. A simulation with the model updating software was carried out to verify where the most sensitive areas in the structure were. Figure 34.6 shows that the jointed flanges were the most sensitive locations were the model required refinements, which indicated the largest errors to be around the connections between the parts of the casing.

Fig. 34.3 Measurement of the casing thickness

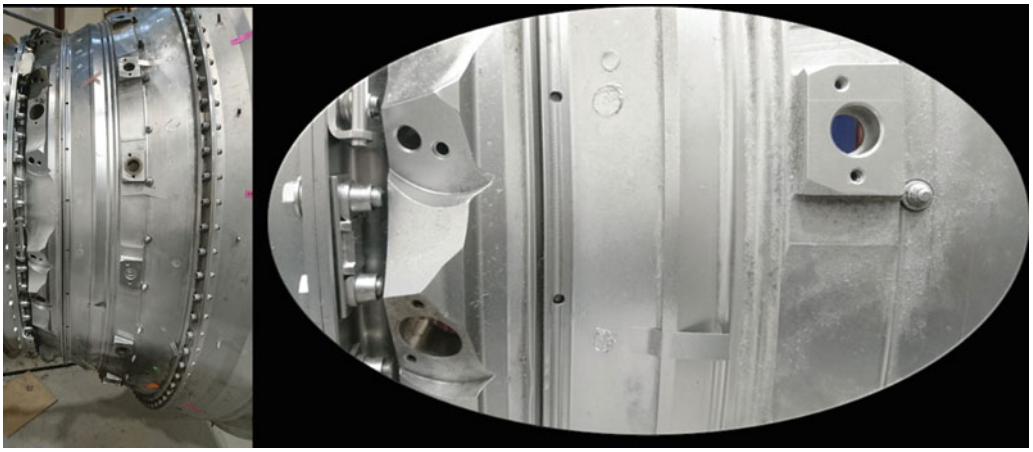


Fig. 34.4 Geometrical features neglected in the model

Table 34.1 Geometrical and material parameters

Parameter	Section	Initial	Updated	Difference (%)
E	Cylinder	1.14E+05	1.14E+05	−1.02E-03
E	Cone 1	1.14E+05	1.14E+05	−8.75E-02
E	Cone 2	2.03E+05	1.52E+05	−2.53E+01
ρ	Cylinder	4.40E-09	4.40E-09	2.80E-04
ρ	Cone 1	4.40E-09	4.40E-09	7.40E-03
ρ	Cone 2	8.00E-09	1.01E-08	2.68E+01
Thickness	Cylinder	1.00E+01	1.00E+01	−1.97E-03
Thickness	Cone 1	2.00E+01	2.00E+01	−1.32E-01
Thickness	Cone 2	1.00E+01	9.26E+00	−7.43E+00

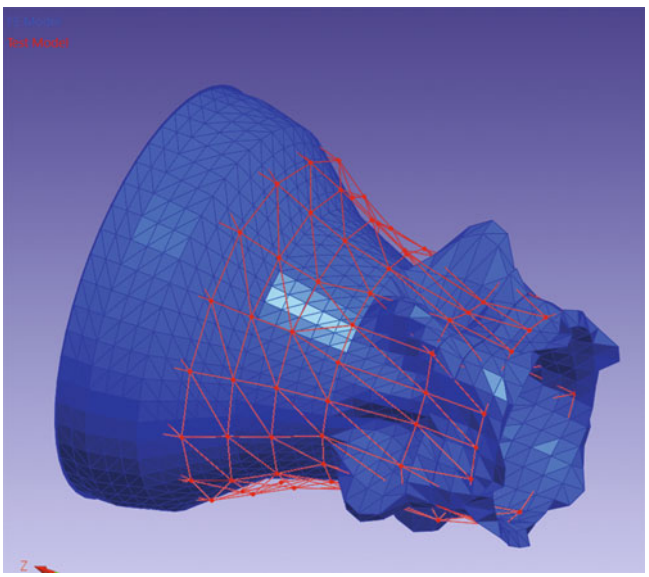


Fig. 34.5 Correlation between measured and theoretical mode

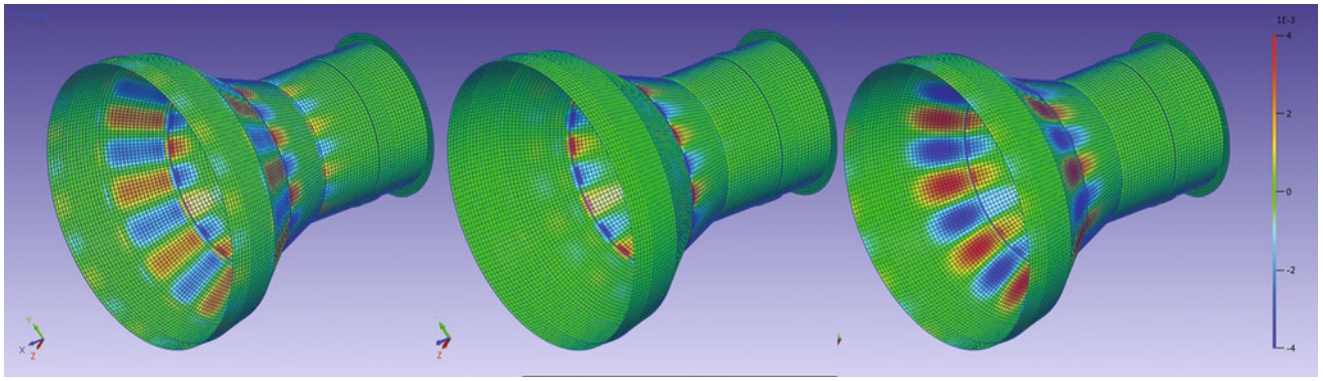


Fig. 34.6 Sensitive areas identified by the model updating software

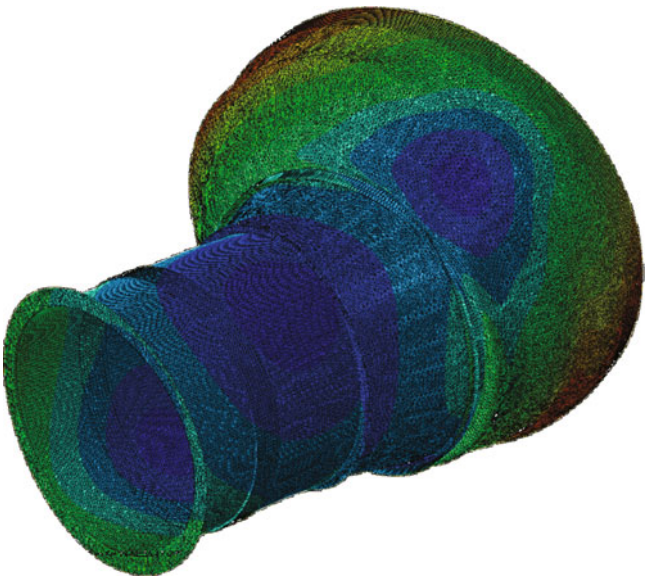


Fig. 34.7 Theoretical mode shape with inclusion of joints

Another undesired effect of these types of constrains was the absence of a theoretical mode which, instead was measured during the experiments. That mode shape was located at much higher frequency because the unrealistic stiffness created at the bolted flanges. So, it was necessary to model the joints and this was carried out by using fasteners, which are linear springs. By creating a bushing 'Connector Section' in ABAQUS, a fastener could be given properties that allowed six DOFs. The fasteners can also be assigned multiple stiffness in all six DOFs with an initial value of 50,000 Nmm^{-1}. The benefits of fasteners over modelling the bolt itself are that fasteners are mesh-independent, easily manipulated and less computationally demanding. Partitions were made on the flanges of the model with 80 connection points. These were created on the partitions which were assigned the bushing connector section with a radius of influence of 8 mm. This method was used as the number of points could be changed easily. Due to limited access to model updating, the model was manually updated. A range of frequencies of interest were selected and the parameters were changed individually and the resultant change and/or shift in mode shapes were studied. The mode shapes were monitored visually with the ABAQUS environment to ensure there was no large shift in frequency in between iterations. Figure 34.7 shows the theoretical mode shape which did not appear in previous calculations because beyond the dynamic range selected. The increased flange flexibility has lowered the natural frequency of the mode shape (Fig. 34.7).

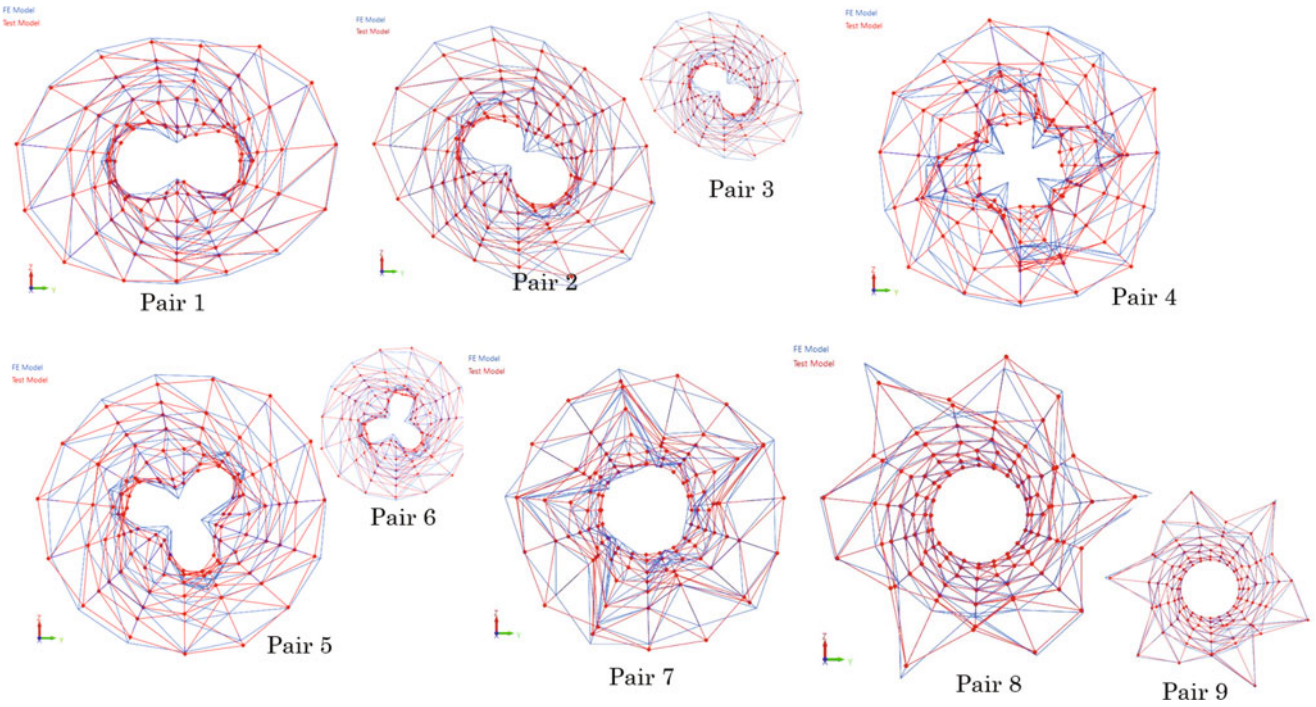


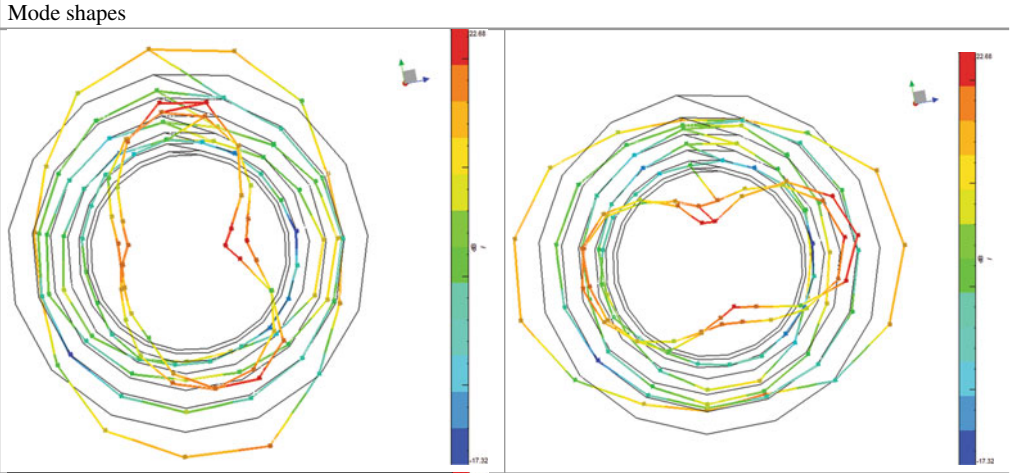
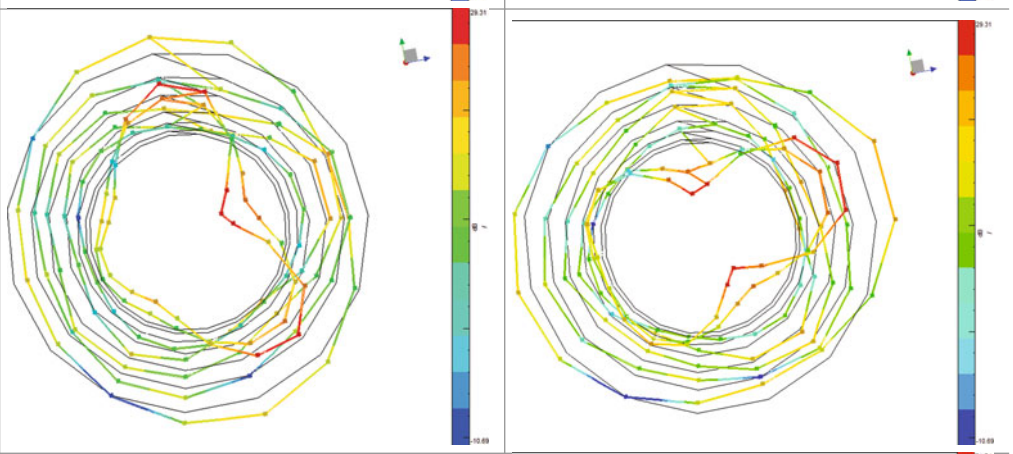
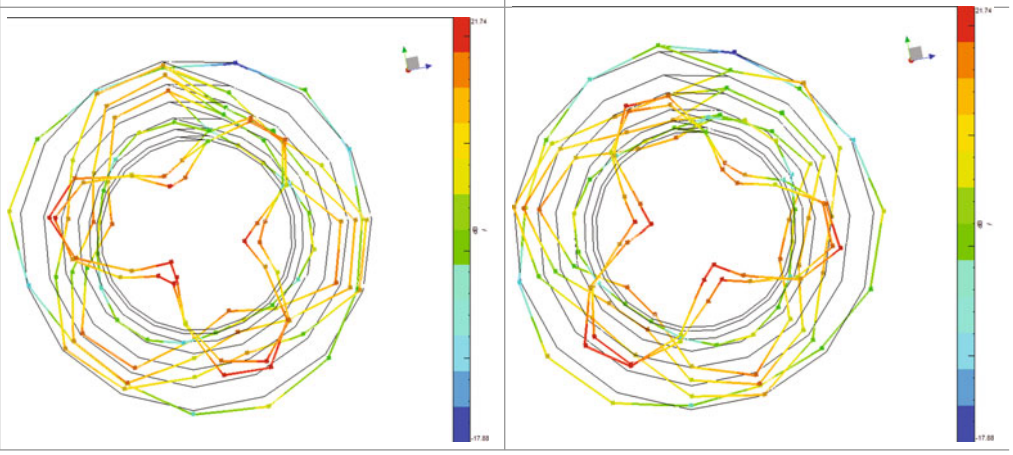
Fig. 34.8 Final updated mode shapes

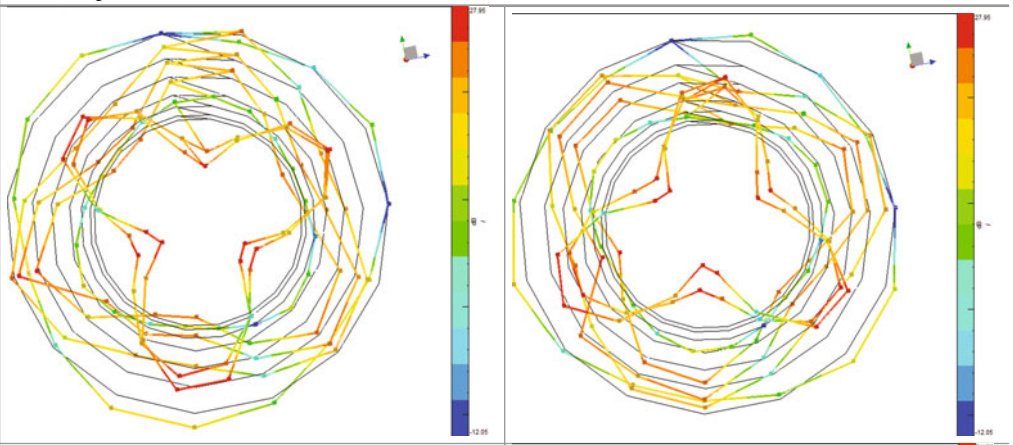
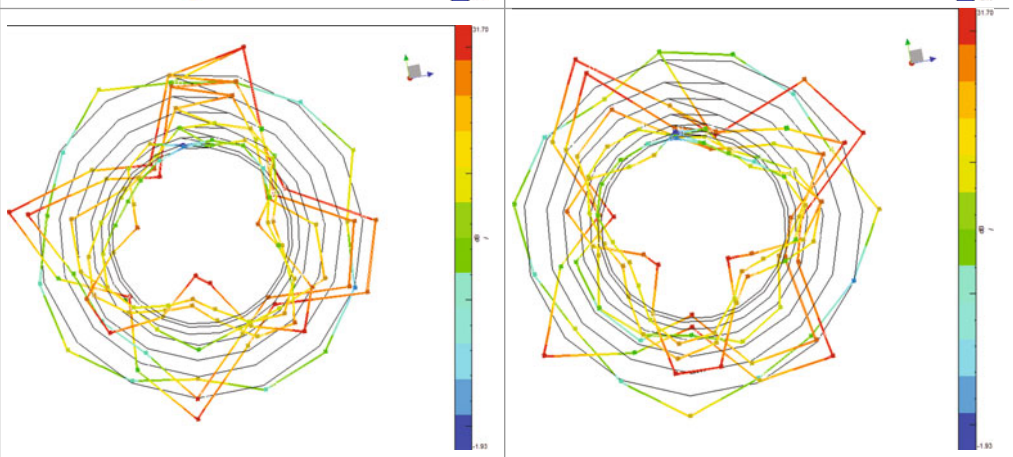
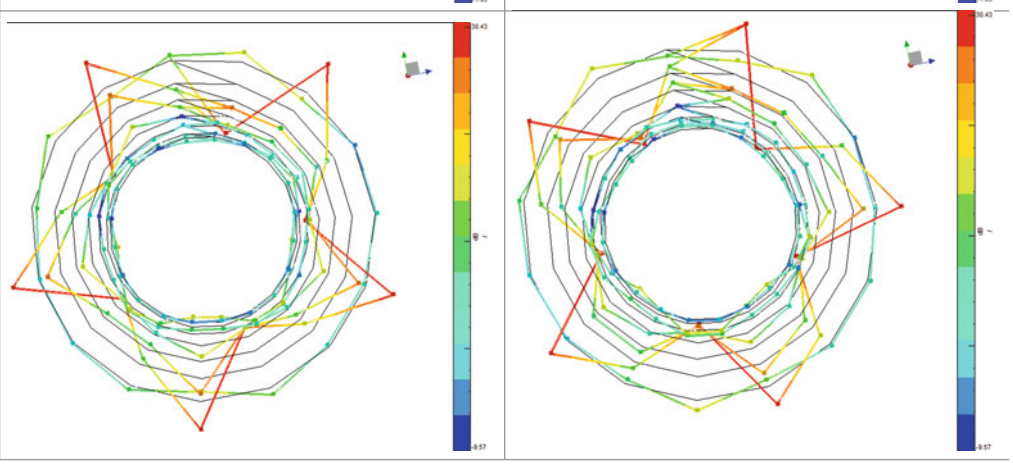
Having identified the most suitable stiffness to give to the joints a new model updating was eventually run with the upgraded model. Figure 34.8 shows the updated mode shapes after the upgrading FE model with the joints.

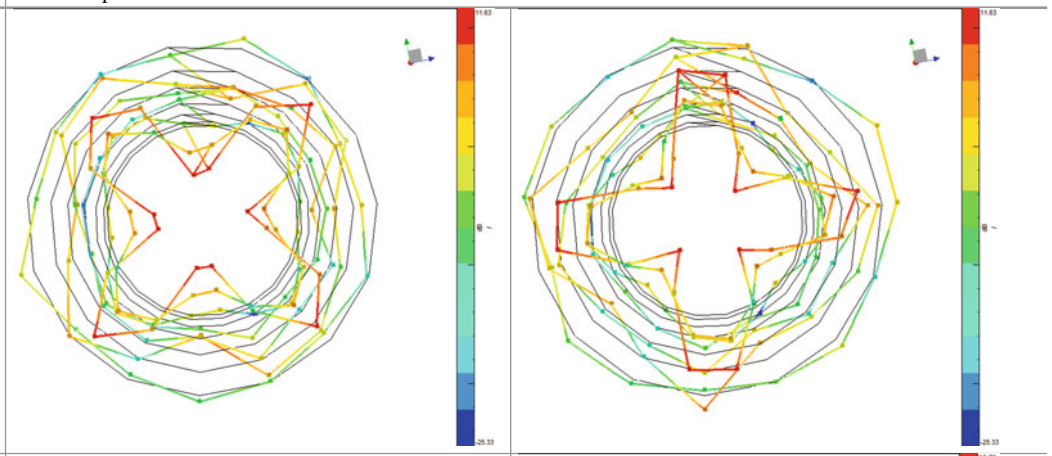
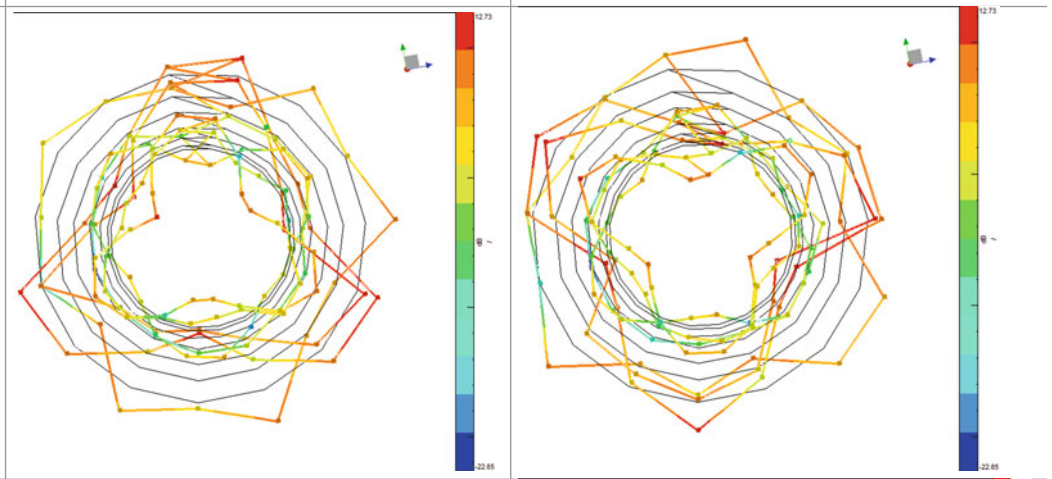
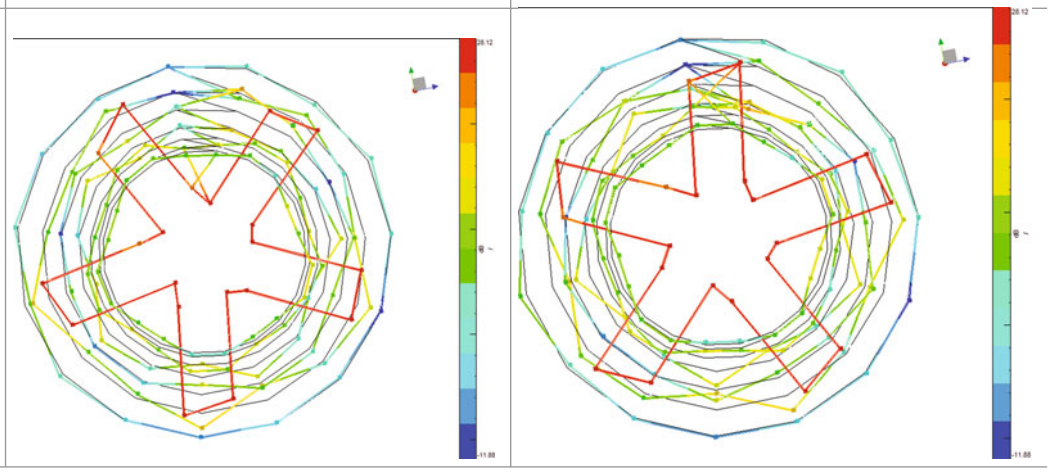
34.4 Conclusions

This piece of research work has demonstrated how important is joints modelling. The finite element modelling of the casing assembly was done and model updating of the mass and stiffness matrix was also carried out by using reference test data. The major challenge was to create a valid FE model by starting with unknown material and geometrical properties. The experimental identification helped the modal updating software to converge to values which produced some acceptable MAC results. However, it was shown that if the bolts are not included into the model then it is impossible to obtain an adequate model validation. It was clear that meshing several bolts was not an option, which left the model updating software unable to perform the updating. Eventually, bolts were included by using fasteners (springs). Manual updating of the spring parameters was carried out and new model updating was gain attempted thus making the FE model of the casing validated. In ultimate instance, it worth mentioning the model updating has provided an excellent support. However, the updating process was limited until the model was upgraded with the inclusions of the stiffness of the joints.

A.1 Appendix

Frequency (Hz)	Mode shapes
188.06	
194.99	
200.88	

Frequency (Hz)	Mode shapes
203.95	
232.95	
282.72	

Frequency (Hz)	Mode shapes
289.46	
318.83	
343.79	

References

1. Di Maio, D., Bennett, P., Schwingshackl, C., Ewins, D.: Experimental non-linear modal testing of an aircraft engine casing assembly. In: Kerschen, G., Adams, D., Carrella, A. (eds.) Topics in Nonlinear Dynamics, Volume 1 SE—2, vol. 35, pp. 15–36. Springer, New York (2013)
2. Ewins, D.J., Weekes, B., delli Carri, A.: Modal testing for model validation of structures with discrete nonlinearities. Phil. Trans. R. Soc. A. **373**(2051), 20140410 (2015)

Chapter 35
Damage Detection in Railway Bridges Under Moving Train Load

Riya C. George, Johanna Posey, Aakash Gupta, Suparno Mukhopadhyay, and Sudib K. Mishra

Abstract In vibration based structural health monitoring, measured response data is used to detect structural damage. This study considers monitoring of railway bridges using response data under moving train loads. The effect of train-bridge interaction, especially with heavy trains, makes the train-bridge system time-varying, and modal identification challenging. The problem becomes even more complex when only the bridge response data is available, and characteristics of the train load (mass, speed etc.) are unknown. To avoid this complexity we engage into a strictly data based technique. Signal energies of the measured responses from healthy and damaged systems are compared statistically to detect the damage in the system. This comparison accounts for: (a) operational variability from different train masses and speeds, and (b) uncertainty from limited instrumentation and unknown input. The technique is validated though numerical simulations and the results promise faster detection of damage.

Keywords Damage detection • Railway bridge • Operational variability • Signal energy • Mahalanobis distance

35.1 Introduction

Structural Health Monitoring (SHM) is the process of implementing a damage detection strategy for engineering infrastructure [1]. With the development of new technologies, it has become plausible to monitor the state of structural components for defects and impairments that may result in a failure. The present state of the structure is compared to a baseline or healthy state of the structure. This allows for the opportunity to repair or replace the damaged part before the failure occurs, which is economically beneficial and significantly improves safety. Early methods for SHM involved visual inspection of the structure, followed by the use of non-destructive testing methods like x-ray and infrared. These methods could be used only when the probable locations of damage are known in advance.

Vibration-based SHM uses sensors, e.g. accelerometers, to measure the vibrations of the structure [2]. The vibration data can be used along with a modal identification algorithm to identify physical properties (e.g. mass and stiffness matrices) of the structure using an analytical model of the real system and the identified modal parameters [3–5]. However, such analytical model based identification techniques would face significant challenges in case of time-varying systems whose modal properties rapidly change with time. Moreover, most model based methods face challenges posed by incomplete information and environmental/operational variability [3]. In recent years, Feature Based Damage Detection (FBDD) methods have emerged as an alternative, which are ideally suited to account for the uncertainties induced by environmental/operational variability. Such techniques do not rely on any analytical physics based model of the structure. Instead, damage sensitive features are extracted from the measured vibration responses of the healthy structure; and the structure is continuously monitored for any changes in these extracted features overtime to indicate damage in the structure [6, 7]. However, in these methods, since no physical model of the structure is identified, the prediction of location, type, and extent of damage become increasingly difficult. While initially developments in vibration-based damage detection relied on structural testing with known applied excitations, recent studies have explored the use of ambient or operational forces as the excitation source [4]; such operational testing has the potential to continuously monitor the structure without having to halt normal operations.

Since railway bridges form a very crucial component of our infrastructure, their structural monitoring continue to garner interest [8, 9]. In this paper we propose a method for structural health monitoring of railway bridges using the acceleration

R.C. George • A. Gupta • S. Mukhopadhyay (✉) • S.K. Mishra
Department of Civil Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India
e-mail: suparno@iitk.ac.in

J. Posey
Department of Civil Engineering, Texas A&M University, College Station, TX 77843, USA

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_35

responses under normal train traffic loads. The modal identification based monitoring of such systems becomes challenging, as the heavy mass of the moving train makes the train-bridge system time varying. Moreover, varying masses and speeds of trains introduce significant operational variability. Hence, we explore here a solely data based method using signal energies of the measured accelerations as damage sensitive features. The signal energies calculated from the responses of the healthy/baseline structure under various mass and speed conditions are compared statistically to the energies of the responses obtained from an unknown state of the structure, to decide whether the unknown state is still healthy or has suffered damage.

35.2 Numerical Simulation

A simplified numerical model of the railway bridge and train system is considered for the simulation of responses under train traffic (Fig. 35.1). The bridge is modeled as a four span continuous beam of 200 m length. The train is modeled as a moving uniformly distributed load of length 100 m. Mass per unit length of the bridge is 12,750 kg/m and flexural rigidity is 10^{15} kgm^3/s^2. Multiple masses and speeds of the train are used as listed in Table 35.1. Mass ratio in Table 35.1 denotes the ratio of mass per unit length of train to that of bridge.

35.2.1 Finite Element Model

The finite element model used for the simulation of the response of the bridge to moving train is obtained by modeling the continuous beam using Euler Bernoulli beam elements. Four elements are chosen per span of the bridge. The Galerkin's method with Hermite polynomial shape functions is used for the finite element formulation. The train load is treated as a distributed dynamic moving mass (see Fig. 35.1), and modeled using Heaviside step functions [10]. To account for damping, a Rayleigh damping model is used with 2% modal damping in the first and second modes. Time integration is performed using the average Newmark-Beta integration scheme, and the mass matrix of the train-bridge system is updated at each time step to account for the moving train mass.

35.2.2 Effects of Operational Variability on Modal Properties of Train-Bridge System

When the train moves over the bridge, it influences the dynamic properties of the train-bridge system, the effects of which are prominent if the mass of the train is comparable to that of the bridge. As the train moves, the location of the train mass on the bridge changes, and this leads to changes of the natural frequencies of the train-bridge system with time. The amount of change in the natural frequencies is governed by the magnitude of the mass of the train. The ratio of the natural frequency of train-bridge system to the natural frequency of bridge is denoted as normalized natural frequency. Figure 35.2 shows the changes in the normalized fundamental natural frequency of the system with time, for different masses of train with a speed

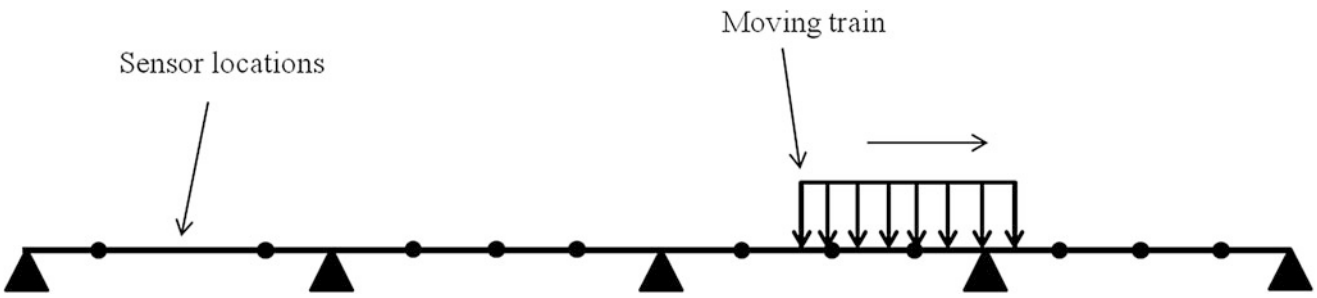


Fig. 35.1 Bridge and train model

Table 35.1 Variability in train mass and speed

Mass ratio	0.25	0.50	0.75	1.00	1.25
Speed (m/s)	20		30		40

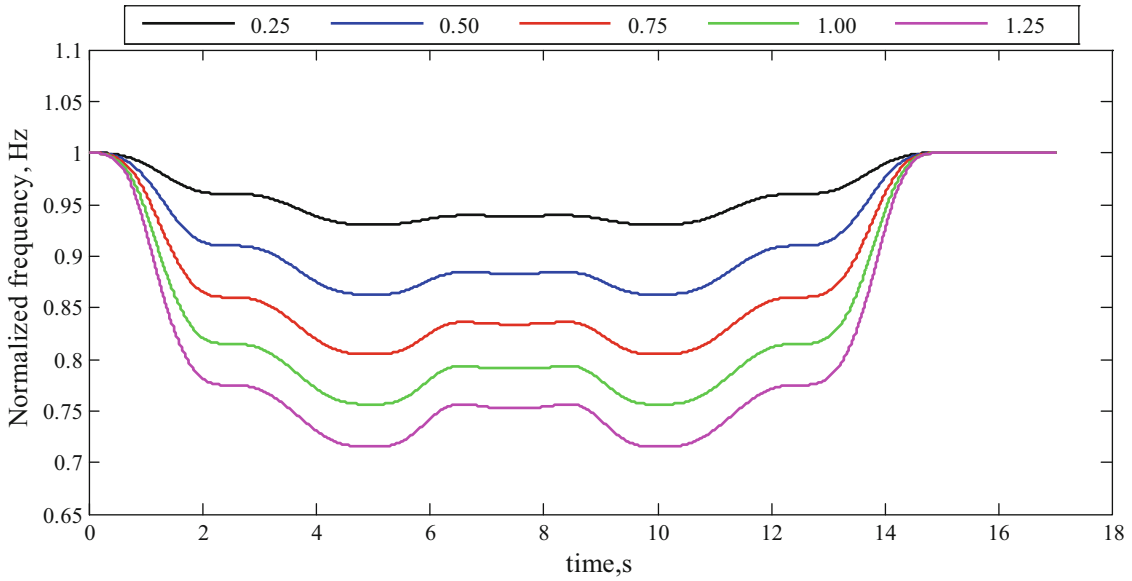


Fig. 35.2 Variation of normalized frequencies of the train-bridge system in time at different masses of train with speed of 20 m/s

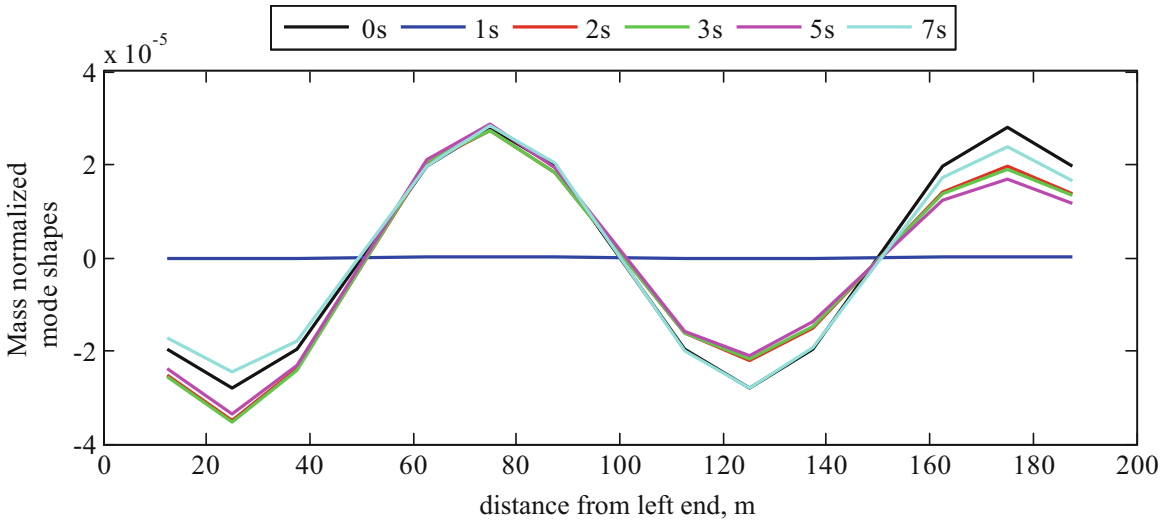


Fig. 35.3 Mass normalized mode shapes at different instances of time for the train load of mass ratio 0.25 and speed 20 m/s

of 20 cm/s. The constant values at the beginning and the end of the duration in Fig. 35.2 represent the natural frequencies of the bridge alone. As expected, as the mass of the train increases, the frequencies of the train-bridge system deviates more from the frequencies of the bridge. When the speed changes, the duration in which the frequencies vary, change, but the amplitude of the frequency variations remains the same. Figure 35.3 shows the variation in the mass normalized mode shapes of the bridge for the translational degree of freedoms of the FE model. The changes of the modal parameters of the train-bridge system at every instant of time as the train passes over the bridge, as observed in Figs. 35.2 and 35.3, illustrate the complexity one would face in engaging in system identification based SHM technique for such systems. The measured accelerations from the bridge will include the effect of these rapidly changing modal parameters, thereby making modal identification of the bridge a challenging problem. To circumvent this issue, we engage into a data based damage detection framework, as discussed in the following section.

35.3 Signal Energy Based Damage Detection

The energy of a discrete time signal can be expressed as:

$$E_S = \sum_{i=1}^{nt} {x_i}^2 \tag{35.1}$$

where, x_i is the signal value at the ith time instant, and nt is the total number of data points in the signal. The signal energies of the recorded acceleration responses at the sensor locations are utilized here as damage sensitive features to detect the presence of damage in the bridge. We assume that acceleration sensors are located at all the unrestrained vertical degrees of freedom of the bridge, but the rotational degrees of freedom are unmeasured as would be expected in real situations. Hence, with d vertical degrees of freedom, we will have a d dimensional feature vector.

35.3.1 Normalization of Signal Energy

Multiple sets of acceleration responses of the healthy bridge, under moving train loads of various mass and speed, are simulated as explained in Sect. 35.2. Ten percent random noise is added to each simulated acceleration time history as measurement noise. To address the uncertainty induced by operational variability, the signal energies are normalized by comparing the energies of each set of responses with the corresponding energies of every other set of responses, as done in [3] with identified stiffness parameters. These normalized signal energies form the training data set. When a new set of acceleration response is obtained from an unknown (still healthy or damaged) state of the structure, the signal energies of these new set are compared with the corresponding energies of every set of responses of the healthy bridge, to obtain the normalized signal energies of the unknown state. This normalization of signal energies reduces the effects of operational variability induced by different masses and speeds of trains running on the healthy bridge and bridge in unknown state.

35.3.2 Identification of Damage

The mean of the vectors of normalized signal energies in the unknown state are compared with the normalized energies of the healthy bridge responses, i.e. the training data points, using outlier analysis with the Mahalanobis squared distance metric [6]. If the Mahalanobis squared distance between the vector of mean normalized signal energies in the unknown state and the training data exceeds a threshold, the unknown state of the bridge is declared damaged. The threshold value is given by,

$$\Gamma = \frac{\left(n_{tr}^2 - 1\right) d}{n_{tr}\left(n_{tr} - d\right)} \left[(1-\alpha) - \textit{quantile of } F_{d,n_{tr}-d}\right] \tag{35.2}$$

where, n_{tr} is number of training data points, d is size of the feature vector (number of sensors in this case), F represents the F distribution, and α is chosen as 1% as in [6].

35.4 Results

Five different masses of the train and three different speeds, as listed in Table 35.1, are considered in the simulation of the bridge responses. Five trials are carried out for each mass-speed combination for the healthy bridge structure. Thus, 75 sets of training signal energies are available. For the bridge considered, we have a total of 12 measured vertical degrees of freedom; hence each set of signal energies is a 12 dimensional vector. Each of these signal energy vectors is compared to the other 74 vectors. This leads to a total of 5550 training data points (12 dimensional vectors) representing the healthy state of the structure. The signal energy vector from a single set of response from the unknown state of the system is compared with the 75 training signal energy vectors, leading to 75 data points representing the unknown state. The mean vector of these latter 75 data points is now compared with the former 5550 training vectors using Mahalanobis squared distance.

Table 35.2 Damage detection using proposed method

Scenario	MSD_{TH}	MSD	Damage induced	Damage detected by the proposed method
Case 1	18.604	12.206	No damage	No damage
Case 2	18.604	701.999	20% in element 6	Presence of damage

Two scenarios are considered to evaluate the efficiency of this method. In Case 1, response from a healthy structure is used as the unknown state's response. This is because, it is essential that damage detection methods should not give false alarms and identify a healthy system as damaged [11], so as to avoid inconveniences to the end user and lack of trust in the methodology. In the simulation of the Case 1 unknown state data, the healthy bridge is excited with a train load of mass ratio 1.25 and train speed 40 m/s (see Table 35.1); this mass-speed combination leads to the maximum response amplitude among all the combinations considered, and is thus most susceptible to be diagnosed wrongly as a case of damage. In Case 2, damage is induced in the 6th element from left of the bridge by reduction the stiffness of the element by 20%. Response under mass ratio of 0.25 and 20 m/s speed is used for damage detection. The performance of the proposed method in both the scenarios is tabulated in Table 35.2. As is evident, the method performs well in both detecting the unknown healthy state as healthy, and the unknown damaged state as damaged.

35.5 Conclusions

An energy based method is proposed to detect the damage in a rail bridge under train traffic load. Mahalanobis squared distance between the normalized signal energies of acceleration responses from healthy bridge structure, and any unknown (healthy/damaged) state of the structure is found to be successful in detecting the presence of damage in the bridge. The proposed methodology is data based and output-only. Hence, it can be used in the structural health monitoring of railway bridges without any interruption to the normal train traffic. This method works well even with incomplete instrumentation of the structure, viz. without requiring measurements from the rotational degrees of freedom. Only the vertical accelerations at three locations in each span are used in this study. Since the method is not based on modal parameters, it circumvents the problem induced by the time-varying nature of the train-bridge system. Hence, this method can also be applied even when the mass of train is comparable to the mass of the bridge, and the influence of this is prominent in the rapidly changing modal properties of the train bridge system. The method also accounts for the operational variability induced by trains of different masses and speeds. Usage of responses of the healthy structure under all expected variations in mass and speed of trains ensure that the possibility of false damage alarms is minimized, when the responses change due to changes in the operational conditions of the traffic. The biggest advantage of the method lies in its computational efficiency. Calculation of signal energies from acceleration time histories and calculation of Mahalanobis squared distance are both computationally very cheap. However, the method has been applied only to detect the presence of damage in the bridge. The possibility of extending the approach to detect damage location still need to be explored.

References

1. Sohn, H., Farrar, C.R., Hemez, F.M., Shunk, D.D., Stinemates, D.W., Nadler, B.R., Czarnecki, J.J.: A Review of Structural Health Monitoring Literature: 1996–2001. Los Alamos National Laboratory, Los Alamos, NM (2003)
2. Doebling, S.W., Farrar, C.R., Prime, M.B., Shevitz, D.W.: Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review, Technical Report LA-13070-MS, Los Alamos National Laboratory, Los Alamos, New Mexico (1996)
3. Mukhopadhyay, S., Luş, H., Betti, R.: Probabilistic structural health assessment with identified physical parameters from incomplete measurements. ASCE-ASME J. Risk Uncertain. Eng. Syst., Part A: Civ. Eng. **2**(3), B4015003 (2015)
4. Duan, Z., Yan, G., Ou, J., Spencer, B.: Damage detection in ambient vibration using proportional flexibility matrix with incomplete measured DOFs. Struct. Control Health Monit. **14**(2), 186–196 (2007)
5. Pandey, A.K., Biswas, M.: Damage detection in structures using changes in flexibility. J. Sound Vib. **169**(1), 3–17 (1994)
6. Balsamo, L., Betti, R., Beigi, H.: A structural health monitoring strategy using cepstral features. J. Sound Vib. **333**(19), 4526–4542 (2014)
7. Gul, M., Catbas, F.N.: Statistical pattern recognition for structural health monitoring using time series modeling: theory and experimental verifications. Mech. Syst. Signal Process. **23**(7), 2192–2204 (2009)
8. Kim, R.E., Moreu, F., Spencer, B.F.: Hybrid model for railroad bridge dynamics. ASCE J. Struct. Eng. **142**(10), 04016066 (2016)

9. Feng, D., Feng, M.: Model updating of railway bridge using in situ dynamic displacement measurement under trainloads. ASCE J. Bridg. Eng. **20**(12), 04015019 (2015)
10. Brewick, P.T., Smyth, A.W.: An investigation of the effects of traffic induced local dynamics on global damping estimates using operational modal analysis. Mech. Syst. Signal Process. **41**(1–2), 433–453 (2013)
11. Farrar, C.R., Worden, K.: Structural Health Monitoring: A Machine Learning Perspective. Wiley, Chichester (2013)

Chapter 36
Multi-Fidelity Calibration of Input-Dependent Model Parameters

G.N. Absi and S. Mahadevan

Abstract The aim of this research is to investigate the use of non-linear structural dynamics computational models with multiple levels of fidelity for the calibration of input dependent system parameters. Non-linear materials often lead to system parameters that are input dependent (function of time, temperature, loading, etc.). Different types of models may also be available for the estimation of unmeasured system properties, with different levels of physics fidelity, mesh resolution and boundary condition assumptions. In order to infer these system properties, Bayesian calibration uses information from multiple sources (including experimental data and prior knowledge), and comprehensively quantifies the uncertainty in the calibration parameters. Estimating the posteriors is done using Markov Chain Monte Carlo sampling, which requires a large number of computations, thus making the use of a high-fidelity model for calibration prohibitively expensive. On the other hand, use of a low-fidelity model could lead to significant error in calibration and prediction. Therefore, this paper develops an approach for input-dependent model parameter calibration with a low-fidelity model corrected using higher fidelity simulation data.

The methodology is illustrated for a curved panel located near a hypersonic aircraft engine, subjected to acoustic loading, where the damping properties of the panel are dependent on the acoustic loading magnitude. Two models (a frequency response analysis and a full time history analysis) are combined to estimate the damping characteristics of the panel. The aim of this study is to develop a novel approach of fusing information from models of different levels of fidelity in the Bayesian calibration of input dependent model parameters.

Keywords Input dependent parameters • Bayesian calibration • Multi-fidelity • Information fusion • Hypersonic vehicles

36.1 Background

36.1.1 Non-Linearity in Structural Dynamics

Although non-linearity is a lot more common in nature than linear behavior, non-linearity is not as well understood and established because of the complex behavior it exhibits, even in weakly non-linear systems (e.g. modal interactions where linear modal superposition becomes inadequate). Typical sources of non-linearity in structural dynamics are geometric non-linearity (large deformations in a system), material non-linearity (non-linear stress/strain constitutive law), damping dissipation (dry friction—contact and sliding between bodies, and hysteretic damping [1] effects), boundary conditions (e.g., surface/fluid interactions), external non-linear body forces (e.g., hydrodynamic forces), etc. [2]. In presence of non-linear behavior, the use of linear superposition to compute the response is erroneous, i.e. the response to a combination of loads in a system does not equal the sum of the responses from individual loads. Although attempts at developing non-linear superposition methods have been reported [3], such methods are limited to simple systems. Reduced order models (ROMs) like convolution integrals, frequency response functions, etc. used in linear analysis cannot be directly applied to non-linear analysis, and the high fidelity non-linear analyses are often prohibitively expensive to run. An alternative approach to deal with non-linear behavior is to develop accurate inexpensive surrogate models (such as neural networks and kriging) trained by low-fidelity model runs but corrected by higher fidelity model simulations.

G.N. Absi • S. Mahadevan (✉)
Department of Civil and Environmental Engineering, Vanderbilt University, 400 24th Avenue South, Nashville, TN 37212, USA
e-mail: sankaran.mahadevan@vanderbilt.edu

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_36

36.1.2 Damping Calibration

Damping is defined as the dissipation of energy from a vibrating structure, and is dependent on the active physical mechanisms in the structure. The types of damping present depend on which mechanism dominates in a given structural configuration [4]. For example, dry friction in structure joints is represented by Coulomb damping (also known as frictional damping). Viscous damping is assumed and derived from the shape of the frequency response curve in several ways [5]. One way is to use the half-power bandwidth method in which the viscous damping ratio is determined from the frequencies for which the power input is half the input at resonance. Test results [6] have shown a broadening of the resonant peak of the frequency response curve under increased acoustic loading, leading to a hypothesis of damping being an input-dependent parameter. Mei and Prasad [7] had shown analytically that for a clamped beam, resonant peak broadening was a direct result of non-linear damping behavior.

36.1.3 Bayesian Calibration

Bayesian calibration of model parameters may be expressed as

$$f_{\boldsymbol{\Theta}}\left(\boldsymbol{\Theta}|\boldsymbol{Y}_{obs}\right) \propto f_{\boldsymbol{\Theta}}\left(\boldsymbol{\Theta}\right).L\left(\boldsymbol{\Theta}\right) \tag{36.1}$$

where $f_{\boldsymbol{\Theta}}(\boldsymbol{\Theta}\,|\,\boldsymbol{Y}_{obs})$ is the posterior distribution of the parameter $\boldsymbol{\Theta}$ after calibration using the data $\boldsymbol{Y}_{obs}$, $f_{\boldsymbol{\Theta}}(\boldsymbol{\Theta})$ is the prior distribution of $\boldsymbol{\Theta}$ (assumed by the analyst), and $L(\boldsymbol{\Theta})$ is the likelihood function (i.e., probability of observing the data $\boldsymbol{Y}_{obs}$, given a value of the calibration parameter). Samples of the posterior can be constructed using a Markov Chain Monte Carlo (MCMC) algorithm. Slice sampling [8] is used in the numerical example in this paper to evaluate Eq. (36.1).

36.1.4 Surrogate Models

In this paper, a Polynomial Chaos Expansion (PCE) model [9] is used for the sake of illustration to replace the original FEA model for inexpensive sampling during the calibration process. PCE is a regression-based surrogate model that represents the output of a model with a series expansion in terms of standard random variables (SRVs).

In all calibration calculations in this paper, the surrogate model error is calculated [10] and added to the output by randomly sampling a value from its distribution.

36.1.5 Model Calibration Under Uncertainty

The experimental observation $\boldsymbol{Y}_{obs}$ is expressed in terms of the input $\boldsymbol{X}$, parameters $\boldsymbol{\theta}(\boldsymbol{X})$, the errors and the model output as follows:

$$\boldsymbol{Y}_{obs} + \boldsymbol{\varepsilon}_{obs} = G\left(\boldsymbol{X} + \boldsymbol{\varepsilon}_{in}, \boldsymbol{\theta}\left(\boldsymbol{X}\right)\right) + \boldsymbol{\varepsilon}_{surr} + \boldsymbol{\varepsilon}_{d}(\boldsymbol{X}) \tag{36.2}$$

where $\boldsymbol{\varepsilon}_{in}$: vector of input measurement errors, one for each input measured; $\boldsymbol{\varepsilon}_{obs}$: vector of measurement errors for the multiple outputs with zero mean and standard deviation $\boldsymbol{\sigma}_{obs}$; $\boldsymbol{\varepsilon}_{d}(\boldsymbol{X})$: vector of model discrepancy terms (discrepancy between prediction and observation, function of input $\boldsymbol{X}$), one for each output; and $\boldsymbol{\varepsilon}_{surr}$: vector of surrogate model errors, one for each output.

The system parameters $\boldsymbol{\theta}(\boldsymbol{X})$ (designated as $\boldsymbol{\theta}$ in the following equations, for simplicity) are calibrated using Bayes' theorem as:

$$\pi\left(\boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_{d}|\boldsymbol{y}_{D}\right) = \frac{A}{B} \tag{36.3}$$

where $A = L(\boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_{d})\ \pi(\boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_{d})$,

$$B = \int L(\boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_d)\ \boldsymbol{\pi}(\boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_d)\ d\boldsymbol{\theta}\ d\boldsymbol{\sigma}_{obs}\ d\boldsymbol{\varepsilon}_d, \text{and}$$

$$L(\boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_d) \propto \prod_{i=1}^{m} \pi\ (\boldsymbol{Y}_{obsi} = \boldsymbol{y}_{Di} | \boldsymbol{x}_{Di}, \boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_d)$$

$\pi(*)$ denotes the joint probability density function (PDF) of the variables, $\pi(\boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_d)$ is the prior joint PDF of $\boldsymbol{\theta}$, $\boldsymbol{\sigma}_{obs}$ and $\boldsymbol{\varepsilon}_d$; $\pi(\boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_d | \boldsymbol{y}_D)$ is the joint posterior PDF of $\boldsymbol{\theta}$, $\boldsymbol{\sigma}_{obs}$ and $\boldsymbol{\varepsilon}_d$ given $\boldsymbol{y}_D$; and $L(\boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_d)$ is the joint likelihood function of $\boldsymbol{\theta}, \boldsymbol{\sigma}_{obs}, \boldsymbol{\varepsilon}_d$. $\boldsymbol{y}_{Di}$ is the vector of the n outputs $\boldsymbol{Y}_{obsi}$ for each individual input setting $\boldsymbol{x}_{Di}$, and $i = 1$ to m, where m is the number of settings for which the outputs are observed. Note that the likelihood function is based on the joint PDF of the observations $\boldsymbol{y}_{Di}$, conditioned on $\boldsymbol{\theta}$, $\boldsymbol{\sigma}_{obs}$ and $\boldsymbol{\varepsilon}_d$.

36.2 Multi-Fidelity Calibration Method for Input-Dependent System Parameters

In this section, the concept of calibration is extended from a simple calibration using experimental data with a single model, to a two-step approach that combines models of different fidelities.

The proposed method is an extension of our previous method [11] that did not consider input-dependent parameters. The approach avoids building a surrogate model for the high-fidelity (HF) simulations, and the high surrogate model error that comes with it. First, a surrogate model is built using low-fidelity (LF) model outputs $S_1(\boldsymbol{X}, \boldsymbol{\theta}(\boldsymbol{X}))$. HF model simulations are then used to update the distributions of the parameters in the LF surrogate model. The updated distributions are subsequently used as *stronger physics-informed priors* for calibration with actual experimental data. The assumption is that the HF model captures the physics that might not be present in the LF model, and might influence the values of the parameters as well as the discrepancy, thus providing stronger physics-informed priors for calibration.

The proposed multi-fidelity calibration algorithm is as follows:

1. *Run the low ($G_1(\boldsymbol{X}, \theta_i(\boldsymbol{X}))$) and high ($G_2(\boldsymbol{X}, \theta_i(\boldsymbol{X}))$) fidelity models to obtain N_1 and N_2 sets of outputs, respectively.*
2. *Build $S_1(\boldsymbol{X}, \theta_i(\boldsymbol{X}))$, the surrogate model replacing $G_1(\boldsymbol{X}, \theta_i(\boldsymbol{X}))$. In this step, the variance of $\boldsymbol{\varepsilon}_{surr}$ of $S_1(\boldsymbol{X}, \theta_i(\boldsymbol{X}))$ is also calculated to account for the surrogate model prediction uncertainty.*
3. *Define the priors of the calibration parameters $\theta_i(\boldsymbol{X})$, and the discrepancy between the models $\boldsymbol{D}_{2,1}(\boldsymbol{X})$.*
4. *Update the parameters of the low-fidelity model as well as the discrepancy term with the high-fidelity simulation results, i.e., use the relationship*

$$\boldsymbol{Y}_{HF} = S_1(\boldsymbol{X}, \theta_i(\boldsymbol{X})) + \boldsymbol{\varepsilon}_{surr} + \boldsymbol{D}_{2,1}(\boldsymbol{X}) \tag{36.4}$$

   to compute the posterior distributions of $\theta_i(\boldsymbol{X})$ and $\boldsymbol{D}_{2,1}(\boldsymbol{X})$, denoted as $\theta_i'(\boldsymbol{X})$ and $\boldsymbol{D}'_{2,1}(\boldsymbol{X})$ respectively.

5. *Define the corrected low-fidelity surrogate model with the updated parameters $\theta_i'(\boldsymbol{X})$ and the updated discrepancy $\boldsymbol{D}'_{2,1}(\boldsymbol{X})$ as*

$$\boldsymbol{LF}_{corr} = S_1\left(\boldsymbol{X}, \theta_i'(\boldsymbol{X})\right) + \boldsymbol{D}'_{2,1}(\boldsymbol{X}) \tag{36.5}$$

6. *Assume a prior distribution for the model discrepancy term $\boldsymbol{\varepsilon}_d(\boldsymbol{X})$ (i.e., the difference between model prediction and experimental observation).*
7. *Re-calibrate the dynamics model parameters along with $\boldsymbol{\varepsilon}_d(\boldsymbol{X})$ with the available experimental data, using the "corrected" low-fidelity model and the relashionship below:*

$$\boldsymbol{Y}_{obs} = S_1\left(\boldsymbol{X}, \theta'_i(\boldsymbol{X})\right) + \boldsymbol{\varepsilon}_{surr} + \boldsymbol{D}'_{2,1}(\boldsymbol{X}) + \boldsymbol{\varepsilon}_d(\boldsymbol{X}) \tag{36.6}$$

   Note that the posteriors of the dynamics model parameters $\theta'_i(\boldsymbol{X})$ from step iv. are used as priors here, in order to compute updated parameters $\theta_i''(\boldsymbol{X})$ and the updated model discrepancy $\boldsymbol{\varepsilon}'_d(\boldsymbol{X})$. (However, $\boldsymbol{\varepsilon}_{surr}$ and $\boldsymbol{D}'_{2,1}(\boldsymbol{X})$ are fixed in this step based on the results of step ii and step v respectively).

The multi-fidelity calibration approach is particularly useful when models of varied fidelities capture different physical behaviors of the system. When calibrating the low-fidelity model using high-fidelity simulations, the parameters to which the high-fidelity model are most sensitive see a substantial shift in their distributions.

If the data on inputs and corresponding model parameters is directly available, then we can simply construct the relationship between inputs and model parameters. However, in general, model parameters are not directly measured but are inferred (calibrated) based on measurements of model outputs (and inputs). In this case, we need to assume a mathematical form of the relationship between inputs and model parameters, and calibrate the coefficients of this assumed relationship along with other calibration quantities. In the numerical example below, we assume an algebraic relationship between the input and the parameters for illustration. The discrepancy also takes into account the physics of the model (e.g. using a functional dependence, $\boldsymbol{\epsilon}_d(\mathbf{X}) = h(\mathbf{X})$).

36.3 Numerical Example

36.3.1 Problem Description

The example problem is of a hypersonic airplane fuselage panel located next to the engine, subjected to dynamic acoustic loading (P). The panel is curved, as shown in Fig. 36.1, and is modeled using the FEA software ANSYS.

The objective is to calibrate the damping properties of the panel (represented by two coefficients, as explained below) as a function of the acoustic loading P, using experimental strain data observed under varying acoustic loading levels, under room temperature. The strain is recorded at seven different locations of the panel. Three strain gages are placed on the top (SG1, SG2 and SG3), two on the bottom (SG4 and SG5) as shown in Fig. 36.1, with the center gage (SG3) recording strain in three different directions. SG1, SG2 and SG4 are used in the calibration. No repeated measurements under the same input are available. This hinders the ability to calculate any spatial correlation between the strain gages. However, repeated runs of the high-fidelity simulations allow the calculation of the spatial correlation between the outputs at those locations. The correlation coefficient is found to be equal to 0.85 among the model outputs at the three strain gage locations, and we used the same correlation coefficient among the observations, observation errors, and model discrepancies at the three locations.

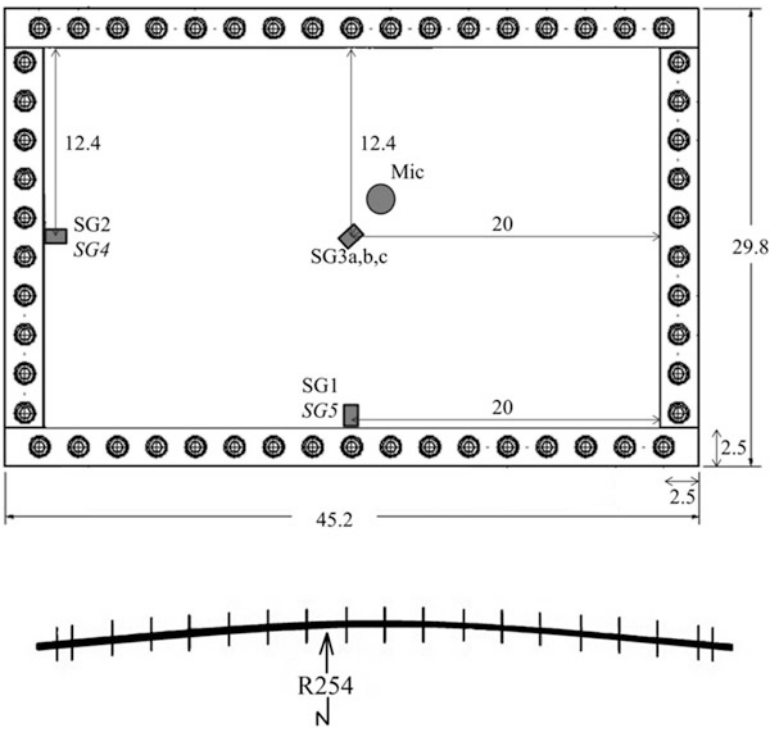


Fig. 36.1 Curved panel dimensions and strain gage locations (units: cm)

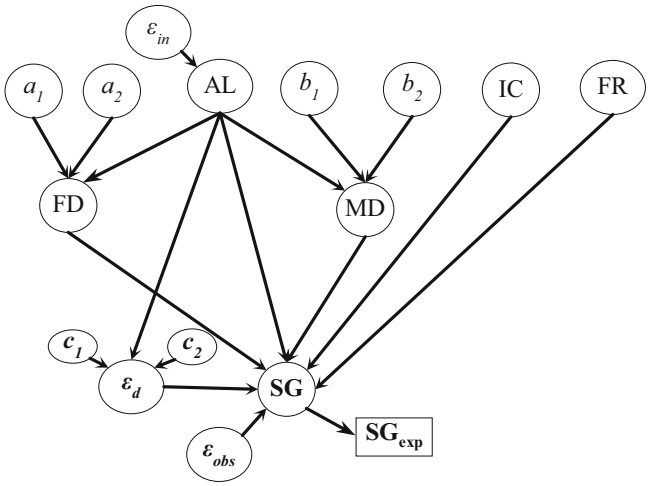


Fig. 36.2 Bayesian network used in calibration

The experimental strain measurement was saved for a duration of 60 s, at a frequency of 50,000 recordings per second, for a total of three million data points. In order to compare the experimental signal to the simulation outputs, the strain signal in the time domain is transformed into a power spectral density (PSD) in the frequency domain, and the energy under the PSD curve, i.e., the root mean square (RMS) value of the signal (the area under the curve) is calculated.

A microphone, located near the center of the panel, captures the acoustic load seen by the panel as it vibrates. This measurement is used as the input load on the simulated structure.

Two types of damping are considered here. First, energy is dissipated via sliding friction at the bolts. This form of energy loss is defined as Coulomb damping [12], and we denote it as frictional damping (FD) for a width of $1''$ around the perimeter of the panel (the edge of the plate is sandwiched between two $1''$ wide metal strips, and bolted to the test frame, as shown in Fig. 36.1). It is modeled as a linear function of the acoustic loading P: $\text{FD} = a_1 * \text{P} + a_2 + \varepsilon_{FD}$, where $\varepsilon_{FD} \sim \text{N}(0, \sigma_{FD})$, the residual in our frictional damping model, represents our uncertainty with respect to the linearity of the assumed function. The second type of damping considered is material damping (MD), i.e., viscous damping throughout the panel, which is also modeled as a linear function of the acoustic loading: $\text{MD} = b_1 * \text{P} + b_2 + \varepsilon_{MD}$, $\varepsilon_{MD} \sim \text{N}(0, \sigma_{MD})$. In this application, we simplify the calibration by assuming exact linear functions for the FD and MD (i.e. $\varepsilon_{FD} = 0$ and $\varepsilon_{MD} = 0$). The boundary fixity is also a calibration variable, accounting for the uncertainty in building the test setup (e.g. loose bolts), and is described by a fixity ratio $FR =$ length of fixed plate boundary/total boundary length. The discrepancy is modeled as $\varepsilon_d = \boldsymbol{c_1} * \text{P} + \boldsymbol{c_2}$ where we assume a linear trend function with respect to the input ($\boldsymbol{c_1}$ and $\boldsymbol{c_2}$ are considered (3×1) vectors for the three strain gage locations used in the calibration). Finally, the strain recordings did not start at the beginning of the experiment, thus initial conditions (IC) in the form of a constant uniform pressure are added to the model and calibrated. The Bayesian network associated with this calibration is shown in Fig. 36.2.

Two models of different fidelities were considered: Model 1 consists of a power spectral density analysis, which is a linear combination of mode shape effects (referred to as low-fidelity model—LF), and model 2 is a full transient analysis where the acoustic loading is applied as a dynamic time history input (referred to as high-fidelity model—HF). A simple sensitivity analysis showed that the strain output from the HF model is highly sensitive to the variations of the frictional damping, whereas the strain output from the LF fidelity model is more sensitive to the material damping. This observation motivates the need to use information from both models for parameter calibration.

36.3.2 Results

The priors for the calibration parameters are listed in Table 36.1. The same priors are assumed for the parameters at all strain gage locations. The surrogate model used in the calibration is a second-order polynomial chaos expansion model, built with 100 training points sampled using an Optimum Symmetric Latin Hypercube sampling technique [13].

The FD (Fig. 36.3) and MD (Fig. 36.4) posteriors are shown for the calibration using only a LF model and then using a corrected LFcorr model. The FR and IC posteriors are shown in Fig. 36.5.

Table 36.1 Calibration parameters priors

Variable	a_1	a_2	b_1	b_2	IC	FR	c_1	c_2	ϵ_{obs}	σ_{obs}	ϵ_{in}
Distribution type	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Normal	Uniform	Normal
Distribution parameters	$[10^{-8}, 6.10^{-5}]$	$[10^{-6}, 10^{-4}]$	$[5.10^{-10}, 6.10^{-7}]$	$[5.10^{-8}, 5.10^{-7}]$	$[10^{-6}, 10^{-4}]$	[0.5, 1]	$[-7.10^{-4}, 7.10^{-4}]$	$[-10^{-4}, 10^{-4}]$	$(0, \sigma_{obs})$	$[0, 5.10^{-3}]$	(0, 1.5)

Note: Distribution parameters refers to the lower and upper bounds for the uniform distribution, and mean and standard deviation for the normal distribution

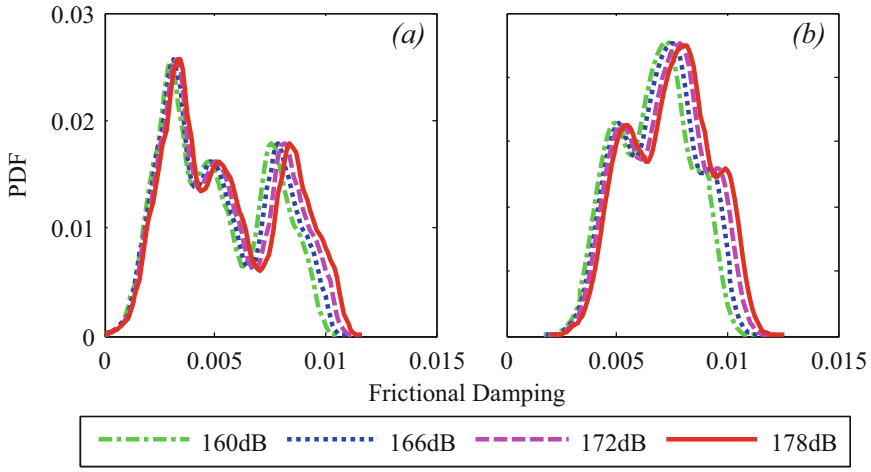


Fig. 36.3 Posteriors of frictional damping using (**a**): LF and (**b**): LFcorr

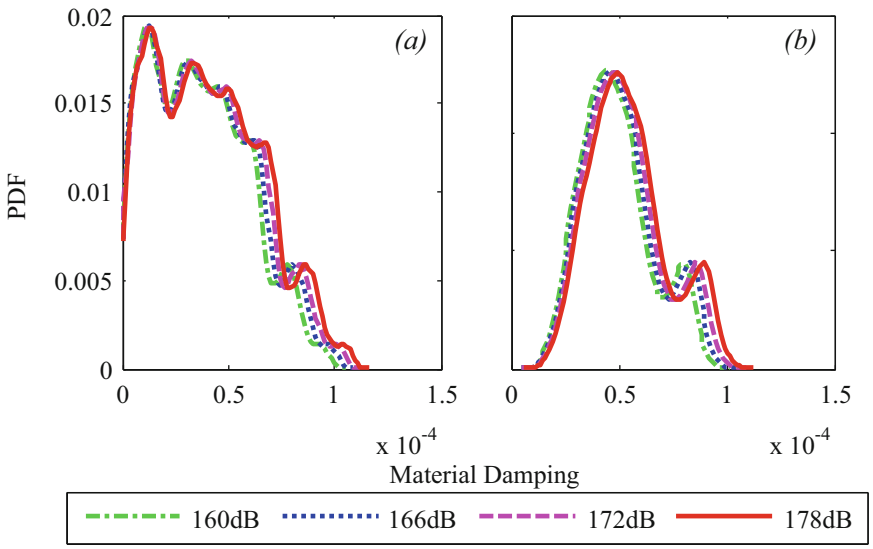


Fig. 36.4 Posteriors of material damping using (**a**): LF and (**b**): LFcorr

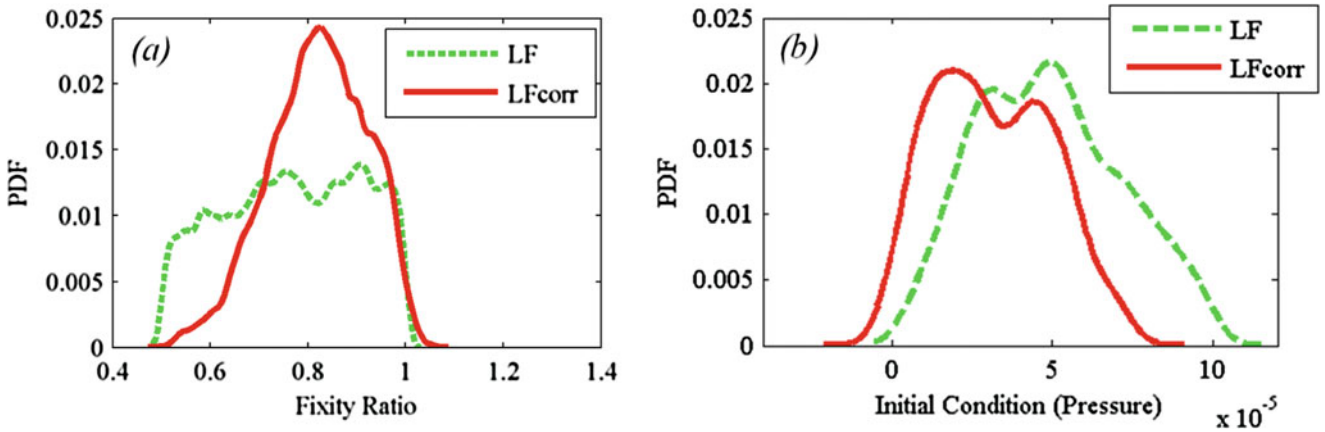


Fig. 36.5 Posteriors of (**a**): fixity ratio (FR) and (**b**): initial condition (IC) using LF (*green*) and LFcorr (*red*)

36.3.3 Discussion

From the results, it is seen that the multi-fidelity method results in significant reduction in uncertainty in the posteriors compared to calibration using only the low-fidelity model. The "correction" of the LF surrogate with HF simulations as a first step results in a reduced space for sampling priors in the second step, helping the convergence of the posterior for the frictional damping towards one dominant mode. A significant reduction in uncertainty for the material damping posterior is also seen. Most importantly, this Bayesian calibration application allowed the expression of the frictional and material damping as functions of the acoustic loading input.

The discrepancy term in the LFcorr model in the application problem is specific to the locations of the strain gages considered. With additional experimental data, the discrepancy term could be improved to include a random field component that takes into account spatial variability across the panel.

36.4 Conclusion

This paper investigated a multi-fidelity approach for the Bayesian calibration of input-dependent model parameters. The correction of a low-fidelity surrogate model with a few high-fidelity simulations provided an efficient way to calibrate non-linear system parameters. Ongoing research includes the additional effect of temperature in the calibration methodology, since temperature affects both the model parameters and the strain gage reading.

Acknowledgments The research was partly supported by funds from the Air Force Office of Scientific Research (Project Manager: Dr. Fariba Fahroo) through subcontract to Vextec Corporation (Investigators: Dr. Robert Tryon, Dr. Animesh Dey). Valuable discussions with Dr. Joseph Hoelkampf at the Air Force Research Laboratory (Wright Patterson Air Force Base) are also acknowledged.

References

1. Caughey, T.K., Vijayaraghavan, A.: Free and forced oscillations of a dynamic system with "linear hysteretic damping" (non-linear theory). Int. J. Non Linear Mech. **5**, 533–555 (1970)
2. Kerschen, G., Worden, K., Vakakis, A.F., Golinval, J.-C.: Past, present and future of nonlinear system identification in structural dynamics. Mech. Syst. Signal Process. **20**, 505–592 (2006)
3. Anderson, R.L., Harnad, J., Winternitz, P.: Systems of ordinary differential equations with nonlinear superposition principles. Physica D. **4**, 164–182 (1982)
4. Adhikari, S.: Damping Models for Structural Vibration. Trinity College, University of Cambridge, Cambridge (2000)
5. Clough, R.W., Penzien, J.: Dynamics of Structures. McGraw Hill, New York City, New York (1975)
6. Gordon, R.W., Hollkamp, J.J.: Reduced-order models for acoustic response prediction, DTIC Document (2011)
7. Mei, C., Prasad, C.B.: Effects of non-linear damping on random response of beams to acoustic loading. J. Sound Vib. **117**, 173–186 (1987)
8. Neal, R.M.: Slice sampling. Ann. Stat. **31**(3), 705–767 (2003)
9. Ghanem, R.G., Spanos, P.: Stochastic Finite Elements: A Spectral Approach. Springer, Berlin (1991)
10. Seber, G.A.F., Wild, C.J.: Nonlinear Regression. John Wiley & Sons, Inc., New York City, New York (1989)
11. Absi, G.N., Mahadevan, S.: Multi-fidelity approach to dynamics model calibration. Mech. Syst. Signal Process. **68–69**, 189–206 (2016)
12. Walshaw, A.G.: Mechanical Vibration with Applications. Ellis Horwood Ltd, New York (1984)
13. Park, J.-S.: Optimal Latin-hypercube designs for computer experiments. J. Stat. Plann. Inference. **39**, 95–111 (1994)

Chapter 37
Empirically Improving Model Adequacy in Scientific Computing

Sez Atamturktur, Garrison N. Stevens, and D. Andrew Brown

Abstract In developing mechanistic models, we establish assumptions regarding aspects of the system behavior that are not fully understood. Such assumptions in turn may lead to a simplified representation or omission of some underlying phenomena. Although necessary for feasibility, such simplifications introduce systematic bias in the model predictions. Often times model bias is non-uniform across the operational domain of the system of interest. This operational domain is defined by the *control parameters*, i.e., those that can be controlled by experimentalists during observations of the system behavior. The conventional approach for addressing model bias involves empirically inferring a functional representation of the discrepancy with respect to control parameters and accordingly bias-correcting model predictions. This conventional process can be considered as experimental data fitting informed by theoretical knowledge, only providing a one-way interaction between simulation and observation. The model calibration approach presented herein recognizes that assumptions established during model development may require omission or simplification of interactions among model input parameters. When prediction accuracy relies on the inclusion of these interactions, it becomes necessary to infer the functional relationships between the input parameters from experiments. As such, this study demonstrates a two-way interaction in which theoretical knowledge is in turn informed by experimental data fitting. We propose to empirically learn previously unknown parameter interactions through the training of functions emulating these relationships. Such interactions can be posed in the form of reliance of model input parameter values on control parameter settings or on other input parameters. If the nature of the interactions is known, appropriate parametric functions may be implemented. Otherwise, nonparametric emulator functions can be leveraged. In our study, we use nonparametric Gaussian Process models in the Bayesian paradigm to infer the interactions among input parameters from the experimental data. The proposed approach will equip model developers with a tool capable of identifying the underlying and mechanistically-relevant physical processes absent from engineering models. This approach has the potential to not only significantly reduce the systematic bias between model predictions and experimental observations, but also further engineers' knowledge of the physics principles governing complex systems.

Keywords Bayesian inference • Discrepancy bias • Gaussian process • Markov Chain Monte Carlo • Parameter calibration • Uncertainty quantification

37.1 Introduction

Scientific computing is being used now more than ever in the development of breakthroughs in all areas of science and engineering design and analysis, as well as to drive missions of national security and energy independence. Success with these programs relies upon our ability to develop fully predictive, mechanistic models of complex systems. These models simulate physics principles which link operational and compositional state variables to the for the purpose of evaluating the behaviors under vastly different operational conditions, often times in settings where we lack a fully understanding of all influential mechanism. The complex nature of these scientific processes often times leads to changing physical properties and characteristics during operation, hence requiring a flexibility for suitable values for input parameters to change throughout

S. Atamturktur (✉)
Glenn Department of Civil Engineering, Clemson University, Clemson, SC, 29631, USA
e-mail: sez@clemson.edu

G.N. Stevens
MET-1 Actinide Engineering and Science, Los Alamos National Laboratory, Los Alamos, NM, 87545, USA

D.A. Brown
Department of Mathematical Sciences, Clemson University, Clemson, SC, 29631, USA

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_37

simulations. Critical aspects of the governing physics principles can be lost in these models if the state-dependent input parameters are simplified using fixed, averaged values throughout the lifespan of the fuel. The omission of such state-dependencies leads to systematic biases between model predictions and observations, which become more pronounced as the complexity of the system increases and knowledge regarding the underlying physics becomes increasingly incomplete.

For example, the next-generation computational capabilities developed as part of Department of Energy's Nuclear Energy Advanced Modeling and Simulation toolkit model the behavior of nuclear fuels by linking the state variables (temperature, fuel burn up, etc.) to relevant output responses (fuel performance, fuel lifespan, etc.) so as to predict fuel performance under various reactor operational conditions [1]. In simulating the underlying fuel behavior, physical parameters that represent the physical properties of the fuel (i.e. density, thermal conductivity) are introduced to the model. With current model calibration methodologies, when these parameters are uncertain they are treated as purely adjustable "knobs" empirically derived from data [2].

The goal of this research is to generate an *empirical approximation* of unmodeled parameter interactions to improve accuracy and understanding of numerical models in scientific computing. We begin with a discussion of current methods in model parameter calibration and inference of discrepancy bias. Section 37.3 presents a new methodology for extending current calibration methods to account for functional relationships amongst parameters, posed in the context of Bayesian inference. Section 37.4 provides a demonstration of the functional calibration capability. Finally, concluding remarks and a path for future work are included in Sect. 37.5.

37.2 Current State of the Art in Calibration of Models Against Experiments

In numerical models, *state variables,* $\boldsymbol{x}$, define the operational domain of the system [3]. The settings of these variables are typically controlled during the experiments and model developers often have knowledge regarding their appropriate ranges. Beyond state variables, numerical models are also dependent upon *physical parameters* that define properties of the system. As systems become more complex, physical parameters, $\boldsymbol{t}$, may not be possible to directly measure or control during experimental tests; yet they are necessary, and often influential, in execution of the simulation.

The computer output can then be denoted as a function of these two separate inputs, $\eta(\boldsymbol{x},\boldsymbol{t})$. When such influential parameters are uncertain the computer model η may be calibrated against experiments, which entail finding the values of $\boldsymbol{t}$ that yield model predictions that best correspond to measurements for all values of $\boldsymbol{x}$ [4–6]. If we let $\boldsymbol{\theta}$ denote the "best" parameter values (i.e., those that result in predictions best corresponding to the experiments), our best computer model yields output $\eta(\boldsymbol{x},\boldsymbol{\theta})$. Unless the calibrated computer model is an exact representation of the true process, however, there will be a systematic bias, δ, between the computer model and the true physical process even at these best-fit values [6–9]. We can then represent the true physical process, $\zeta(\boldsymbol{x})$, as:

$$\zeta(\boldsymbol{x}) = \eta(\boldsymbol{x},\boldsymbol{\theta}) + \delta(\boldsymbol{x}) \tag{37.1}$$

where $\delta(\boldsymbol{x})$ is an empirical function representing the systematic disagreement between the computer model and reality as a function of state variables, as demonstrated in Fig. 37.1 [6, 7, 9, 10].

Although the treatment of discrepancy bias as shown in Eq. (37.1) has proven useful in previous applications, there are also many situations where this formulation is insufficient. This inadequacy stems from the fact that implementing the formulation given in Eq. (37.1) for model calibration *requires us to assume the best-fit values for parameters representing the physical properties,* $\boldsymbol{\theta}$*, to be fixed during the operation.* As a result, the calibration process would search for a single best "averaged" definition of $\boldsymbol{\theta}$ for the entire domain of $\boldsymbol{x}$. However, in many realistic scientific applications—, such as nuclear energy models where materials in fuel rods are subjected to severe radiation thus altering their thermo-mechanical properties [1]— this assumption may not always hold. In the case that influential parameter interactions are omitted during model development, traditional calibration will therefore never completely capture reality and the remaining inadequacies will be lumped into an indecipherable, black-box model bias term.

The limitations of Eq. (37.1) in the presence of state-dependent model input parameters can be analytically demonstrated through a hypothetical computer model and its corresponding experiments (Fig. 37.1). Consider the case where the computer model includes a state-dependent calibration parameter that varies as a function of a control variable, $\boldsymbol{\theta(x)}$ (for instance, the way thermal conductivity varies as a function of burn up in a nuclear fuel rod [2]). If the modeler was unaware of this dependency (or chose to simplify it), the model would be built with fixed values of $\boldsymbol{\theta}$ for all values of $\boldsymbol{x}$. Figure 37.1 (left) illustrates an ensemble of model predictions obtained by sampling constant values of $\boldsymbol{\theta}$. Clearly, while some of these parameter values provide a better match to experiments than others none of the ensemble members are able to completely

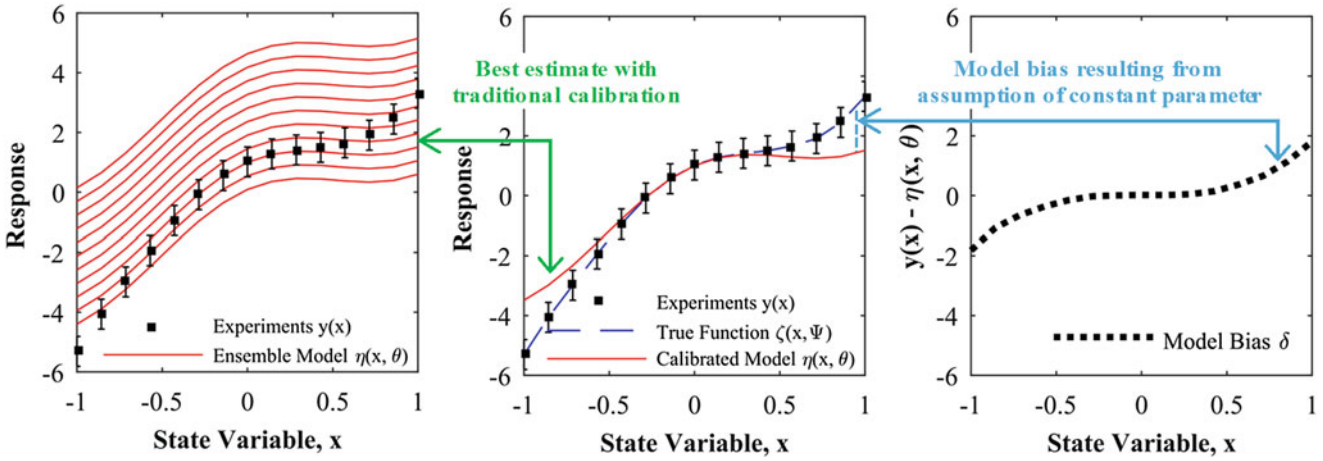


Fig. 37.1 (*Left*) Ensemble of model predictions obtained by sampling with a set of constant values. (*Middle*) Comparison of the true response to the best estimate model when the traditional approach of a constant-valued uncertain parameter is used. (*Right*) Model bias resulting from modeling assumption that θ is independent of the state variable x

recover the effects of the missing parameter interaction, as shown by the best estimate model in Fig. 37.1 (center). Thus, the model would exhibit a systematic discrepancy with experiments as shown in Fig. 37.1 (right). It is also important to note that in the formulation given in Eq. (37.1), any unmodeled state-dependency of $\boldsymbol{\theta}$ is absorbed into $\delta(\boldsymbol{x})$. Hence, this general formulation treats the systematic discrepancy bias as a black-box without physical interpretability. In other words, *this approach identifies that the model has a deficiency but does not explain why or what it the cause*, making the underlying parametric relationships indecipherable.

37.3 Research and Methods

37.3.1 Methodology: Statistically Rigorous Framework for Model Calibration

We propose to extend the framework detailed in Eq. (37.1) by allowing the best estimates for calibrated parameters, $\boldsymbol{\theta}$, to be functions of other model input parameters. The following methodology is presented with calibration parameters functionally dependent upon state variables so that $\boldsymbol{\theta} \equiv \boldsymbol{\theta}(\boldsymbol{x})$, however, the same concept can be extended to infer relationships between parameters that are not state parameters. *This novel approach addresses the fact that the 'correct' value of* $\boldsymbol{\theta}$ *may evolve over the operational domain of a system, in accordance with the omitted (yet important) relationships between model parameters.*

With this new methodology, missing relationships that were previously grouped together into the black-box bias term will identified, increasing our understanding of underlying physics as well as reducing the model bias. If a sufficient portion of the missing interactions (as well as the associated $\boldsymbol{\theta}$ and $\boldsymbol{x}$ variables) are correctly accounted for, the need for a separate black-box term for discrepancy bias could be eliminated altogether. In that case, we obtain a model that is uniformly valid throughout the domain with discrepancy $\delta^*(\cdot) \approx 0$, leading to:

$$\zeta(\boldsymbol{x}) = \eta(\boldsymbol{x}, \boldsymbol{\theta}(\boldsymbol{x})) \tag{37.2}$$

However, if some of the missing physics in the computer model manifests as missing model parameters (i.e., incomplete parameterization), a level of discrepancy bias, $\delta^*(\cdot)$, may still remain. In this case, the true system response could be represented by:

$$\zeta(\boldsymbol{x}) = \eta(\boldsymbol{x}, \boldsymbol{\theta}(\boldsymbol{x})) + \delta^*(\boldsymbol{x}) \tag{37.3}$$

It is important to note that in Eq. (37.3), for any discrepancy function, $\delta^*(\cdot)$, a best-fitting $\boldsymbol{\theta}(\cdot)$ could be found to adjust the model to better fit experiments, possibly resulting in non-physical solutions. Thus, the model discrepancy and the calibration parameters may be *confounded*, posing a challenge in distinguishing between model error (due to model

inadequacy) and parameter error (due to inaccurate model calibration). A large body of work has demonstrated the ability of Bayesian calibration to successfully estimate parameters, in spite of these confounding effects, through the definition of prior distributions (e.g., [11]). Furthermore, imposing an appropriate level of smoothness for $\boldsymbol{\theta}(\boldsymbol{x})$ makes it possible to acquire information about $\delta^*(\cdot)$ and $\boldsymbol{\theta}(\cdot)$ through a reasonable number of experiments [12].

Using Bayesian inference for determining the function $\boldsymbol{\theta}(\boldsymbol{x})$ will allow the incorporation of all available prior knowledge concerning the functional form of $\boldsymbol{\theta}(\boldsymbol{x})$, its reasonable values and the uncertainty associated with these values. Such knowledge may be obtained through expert judgment, empirical estimation, or predictions obtained from lower-scale code. However, if such prior knowledge regarding the nature of $\boldsymbol{\theta}(\boldsymbol{x})$ is not available, a nonparametric model (such as a Gaussian process model, see the next section) capable of adapting its complexity to best represent the data and avoid imposing restrictions on the functional form of $\boldsymbol{\theta}(\boldsymbol{x})$ can be used.

37.3.2 Gaussian Process Models for Emulating $\delta(\cdot)$

There is likely to be limited information, if any, about the functional form of $\delta(\cdot)$. To alleviate this, Kennedy and O'Hagan [7] suggested modeling the discrepancy nonparametrically with a stationary Gaussian process (GP) and eliminating the need for assumptions regarding a distributional family from which $\delta(\cdot)$ is drawn. We can represent the discrepancy, for instance, through a GP with a mean of zero[1] over the domain of applicability. We would thus have $\delta(\cdot) \sim GP\left(0, \lambda_\delta^{-1} R_\delta(\cdot,\cdot)\right)$, where λ_δ is the precision and $R^\delta(\cdot,\cdot)$ is the correlation function.

The use of GPs to model unknown functional forms is motivated by the fact that the GP is nonparametric and hence capable of adapting to a very wide range of smooth functions while maintaining computational tractability [13–16]. While parametric models are preferable when appropriate, their utility is contingent upon the chosen parametric form being approximately correct. This is a strong, often unverifiable assumption when little is known about the system behavior being modeled. Thus, while certain parametric models yield desirable properties such as knowledge of higher-order moments, their properties can be quite misleading when the model is wrong, giving a researcher a false sense of security in the model's predictions. Nonparametric models such as the GP make no assumptions about particular functional forms and are thus more resistant incorrect distributional assumptions by giving the data more freedom in determining the shape of the appropriate function. In fact, certain parametric models, including linear models, can actually be represented as special cases of GPs [17], meaning that the GP can be thought of as a generalization of a broad class of parametric models. In general, it is more difficult to quantify the predictive uncertainty associated with nonparametric models than their parametric counterparts. However, an appealing property of the GP is the availability of an expression for the variance of predicted values with which the prediction precision can be estimated. Further, the Bayesian approach using GP priors enables a modeler to provide tolerance bounds around model predictions, within which true observations would fall with a specified probability. These desiderata persist even when a computer model is applied to predict at settings for which no experimental observations are available (provided there exists some concomitant information from similar scenarios).

37.3.3 Gaussian Process Models for Emulating $\eta(\cdot,\cdot)$

In the case of a fast-running computer model, it may not be necessary to emulate the $\eta(\cdot,\cdot)$ since simulations can be quickly performed to determine the actual model prediction. When the code is computationally expensive, however, it might be necessary to create an emulator (also known as a surrogate model, response surface model, or a meta-model) of the computer output to estimate what the output would be at certain input settings where actual output from the computationally expensive model is not available. Once again Kennedy and O'Hagan [7] suggested the use of an emulator that does not restrict $\eta(\boldsymbol{x},\boldsymbol{t})$ to a specific functional form. Hence, if computational demands associated with the engineering system model become prohibitive, we can use another GP in lieu of $\eta(\cdot,\cdot)$. For instance, we can specify $\eta(\cdot,\cdot) \sim GP\left(\boldsymbol{h}^T(\cdot)\boldsymbol{\beta}_\eta, \lambda_\eta^{-1} R_\eta(\cdot,\cdot)\right)$, where $\boldsymbol{h}^T(\boldsymbol{x})\boldsymbol{\beta}_\eta$ is a mean function of $\boldsymbol{x}$ that is linear in the unknown regression parameters $\boldsymbol{\beta}_\eta$. When the model is to be used only for

[1] A nonzero constant mean could be used as well. However, experience and published literature show that the bias is typically encouraged to center around 0, allowing a correct computer model to emerge with little bias when supported by the data.

predicting output within the range of state variables for which experiments are available, a constant mean is a suitable choice ($\boldsymbol{h}^T(\cdot) \equiv 1\,, \boldsymbol{\beta}_\eta \equiv \mu \in R^1$) [6]. A more complex function for the mean may be necessary if the GP is desired to extrapolate the model outside this range [9].

37.4 Conceptual Demonstration

Here, we demonstrate an extension of the new model calibration approach as described by Eq. (37.2). Let us illustrate this concept using simulated field data $\boldsymbol{y}$, generated using known "true" values of all model input parameters. Recognizing that numerical models will have parameters beyond $\boldsymbol{x}$ and $\boldsymbol{\theta}(\boldsymbol{x})$, we introduce a separate control parameter, $\boldsymbol{z}$ (which experiments and the model depend upon, but no input parameters have an interaction with) as well as an uncertain parameter, $\boldsymbol{\theta}$, that is known to be constant and does not have a functional interaction with other parameters. Data are generated supposing that:

$$y = \boldsymbol{\theta}_1(\boldsymbol{x}) \sin(\boldsymbol{z}) + \boldsymbol{\theta}_2 \boldsymbol{x} + \cos(3\boldsymbol{x}) + \epsilon \tag{37.4}$$

where $\boldsymbol{\theta}_1(\boldsymbol{x})$ is an uncertain, functional input parameter and $\boldsymbol{\theta}_2$ is an uncertain, constant input parameter. True values for these parameters, used to generate the data for calibration, are $\boldsymbol{\theta}_1(\boldsymbol{x}) = \boldsymbol{x}^3 - 2\boldsymbol{x} + 5$ and $\boldsymbol{\theta}_2 = 2.5$. Variability in the data due to unavoidable experimental error is represented by $\epsilon \sim N(0, 0.05^2)$. Figure 37.2 clearly reveals the inaccuracies in the model when $\boldsymbol{\theta}_1$ is assumed to be a constant value throughout the $\boldsymbol{x}$ domain.

In this problem we consider limited prior information regarding the true parameter values to be known. Uniform distributions were assumed *a priori* for both calibration parameters, with upper and lower bounds of [0 10] and [0 4] for $\boldsymbol{\theta}_1(\boldsymbol{x})$ and $\boldsymbol{\theta}_2$, respectively. Following the formulation of Eq. (37.2) along with a GP representation $\boldsymbol{\theta}_1(\boldsymbol{x})$ as outlined in Sect. 37.3.2, the model is calibrated considering both forms of parametric uncertainty as well as experimental uncertainty simultaneously.

Posterior density of the uncertain parameter that is known to remain constant, $\boldsymbol{\theta}_2$ is shown in Fig. 37.3. The posterior distribution of the $\boldsymbol{\theta}_1(\boldsymbol{x})$ function is shown in Fig. 37.4, where grey lines illustrate 3000 realizations and the dashed red lines indicates the "true" parameter interaction that was used to generate the synthetic data. To ensure convergence of the sampling, three sample chains of 3000 samples, retaining every 3rd draw (resulting in 1000 samples per chain), were carried out from different starting values. Notice that not all of the state-variable settings with available experimental data were used in the parameter calibration (Fig. 37.4 left). The portion of the data kept as a holdout is used to evaluate the accuracy of the empirically inferred functional form at untested settings (Fig. 37.4 right). Improvements to the predictive capability of the model through the new model calibration approach are demonstrated in Fig. 37.5, in which model predictions are clearly better fitted to experimental data points throughout the entire domain than previously when $\boldsymbol{\theta}_1$ was assumed to be constant.

These preliminary results demonstrate the promise for better model calibration techniques that lies within the new methodology for model calibration. As seen, the proposed approach is capable of functionally identifying unforeseen or

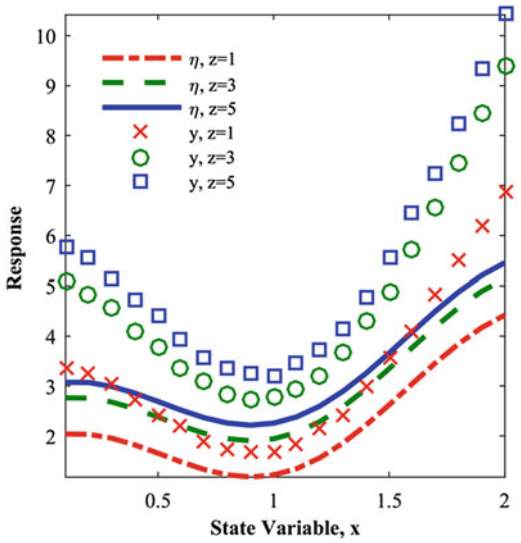


Fig. 37.2 Mismatch of model predictions to experimental data throughout the operational domain due to neglecting the parameter interaction θ(x)

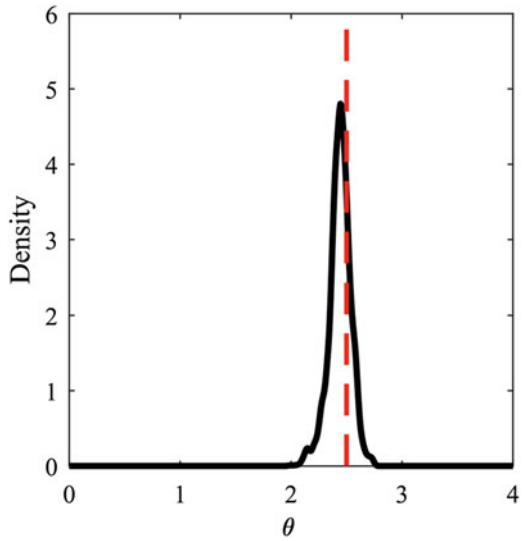


Fig. 37.3 Posterior distribution of the constant-valued calibration parameter, centered around the true value

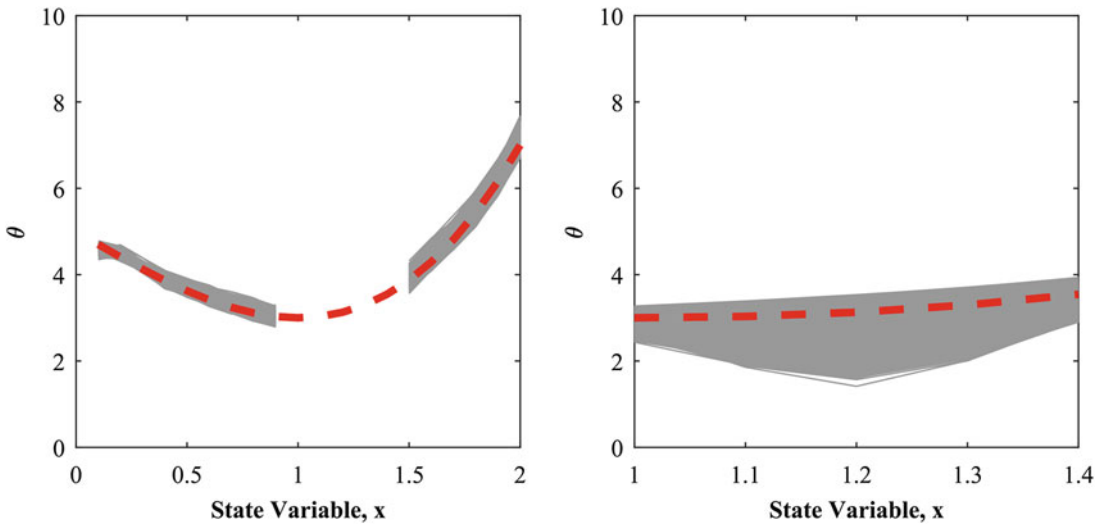


Fig. 37.4 Posterior distribution of the functional calibration parameter at calibrated (*left*) and holdout (*right)* settings, compared to the true function

insufficiently understood relationships between input parameters and state variables. Such lack of knowledge is typically exacerbated as the system complexity increases and as the available knowledge of the associated mechanistic behavior becomes increasingly insufficient, which are characteristics of many engineering applications of complex systems driving our world today.

37.5 Discussion and Conclusion

Current statistical analysis techniques can identify systematic bias throughout the operational domain of a system, however, there is no existing capability to explain the root causes of such bias. The calibration method presented herein provides an entirely new statistical analysis procedure capable of revealing the inherent model deficiencies responsible for the discrepancy bias in model predictions. Functional model calibration achieves this capability by inferring, from readily available experimental measurements, the functional relationships that reflect the state dependencies of individual model input parameters. This new approach introduces multiple benefits including (1) increasing the amount of information gained from experiments when calibrating our models, (2) yielding a mathematics-based description of the omitted, but mechanistically relevant, parameter interactions and (3) producing models with significantly reduced discrepancy bias.

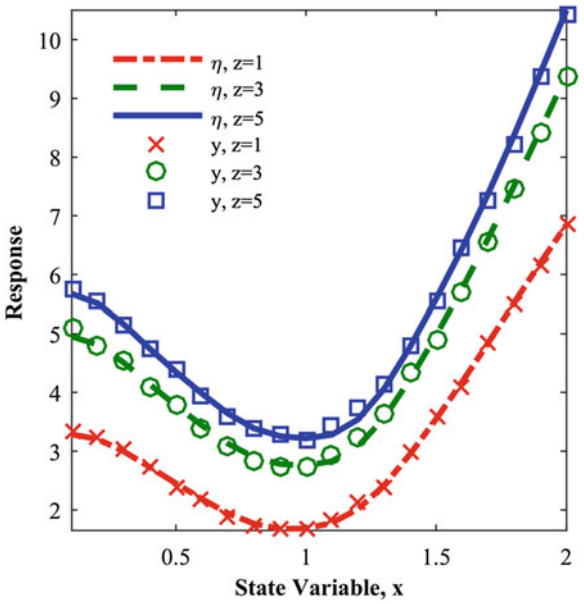


Fig. 37.5 Improved predictive capability of the functionally calibrated model, demonstrated by better match to experiments across the operational domain from Fig. 37.2

Impacts of this work will continue to grow as efforts continue to expand capabilities of the algorithm. One anticipated challenge is that the computational demands of our proposed calibration method to infer state dependencies can grow very quickly with the dimension of the parameter space, therefore exploration of efficient sampling methods for high dimensional problems would be high impact. Further considering the computational demands of functional parameter calibration, a means of identifying the influential parameters in a model that may remain constant and that should be state-dependent is necessary. Developments, for example a sensitivity analysis to make such assertions, would streamline the calibration to focus on those parameter interactions that are most influential. With the potential demonstrated in this preliminary work, we contend the proposed model calibration approach will greatly enhance the management of complex science and engineering problems.

References

1. Williamson, R.L., et al.: Multidimensional multiphysics simulation of nuclear fuel behavior. J. Nucl. Mater. **423**(1–3), 149–163 (2012)
2. Unal, C., Williams, B.J., Yacout, A., Higdon, D.M.: Application of advanced validation concepts to oxide fuel performance codes: LIFE-4 fast-reactor and FRAPCON thermal-reactor fuel performance codes. Nucl. Eng. Des. **263**, 102–128 (2013)
3. Hemez, F., Atamturktur, H.S., Unal, C.: Defining predictive maturity for validated numerical simulations. Comput. Struct. **88**(7–8), 497–505 (2010)
4. Reese, C.S., Wilson, A.G., Hamada, M., Martz, H.F., Ryan, K.J.: Integrated analysis of computer and physical experiments. Technometrics. **46**(2), 153–164 (2004)
5. Kennedy, M.C., Anderson, C.W., Conti, S., O'Hagan, A.: Case studies in Gaussian process modelling of computer codes. Reliab. Eng. Syst. Saf. **91**(10–11), 1301–1309 (2006)
6. Higdon, D., Nakhleh, C., Gattiker, J., Williams, B.: A Bayesian calibration approach to the thermal problem. Comput. Methods Appl. Mech. Eng. **197**(29–32), 2431–2441 (2008)
7. Kennedy, M.C., O'Hagan, A.: Bayesian calibration of computer models. J. R. Stat. Soc. Ser. B Stat. Methodol. **63**(3), 425–464 (2001)
8. Williams, B., Higdon, D., Gattiker, J., McKay, M., Keller-McNulty, S., Moore, L.: Combining experimental data and computer simulations, with an application to flyer plate experiments. Bayesian Anal. **1**(4), 765–792 (2006)
9. Bayarri, M.J., et al.: A framework for validation of computer models. Technometrics. **49**(2), 138–154 (2007)
10. Atamturktur, S., Hegenderfer, J., Williams, B., Egeberg, M., Lebensohn, R.A., Unal, C.: A resource allocation framework for experiment-based validation of numerical models. Mech. Adv. Mater. Struct. **22**(8), 641–654 (2015)
11. Gustafson, P.: On model expansion, model contraction, identifiability and prior information: two illustrative scenarios involving mismeasured variables. Stat. Sci. **20**(2), 111–140 (2005)
12. Fricker, T.E., Oakley, J.E., Urban, N.M.: Multivariate Gaussian process emulators with nonseparable covariance structures. Technometrics. **55**(1), 47–56 (2013)
13. O'Hagan, A., Kingman, J.F.C.: Curve fitting and optimal design for prediction. J. R. Stat. Soc. Ser. B Methodol. 1–42 (1978)
14. Neal, R.M.: Regression and classification using Gaussian process priors. Bayesian Stat. **6**, 1–16 (1998)
15. Santner, T.J., Williams, B.J., Notz, W.I.: The Design and Analysis of Computer Experiments. Springer Science & Business Media, New York (2003)
16. Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. MIT Press, Cambridge, MA (2006)
17. Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B.: Bayesian Data Analysis, vol. 2. Chapman & Hall, London (2014)

Chapter 38
Mixed Geometrical-Material Sensitivity Analysis for the Study of Complex Phenomena in Musical Acoustics

R. Viala, V. Placet, and S. Cogan

Abstract The behavior of stringed musical instruments have historically been investigated primarily through experimental studies and simplified analytical models. In this study, a physics-based model of a violin is used to determine the most influential design and environmental factors impacting its vibrational behavior. A global sensitivity analysis is performed based on the frequency responses of the violin and includes geometric and material properties as well as climatic conditions. In particular, this study attempts to provide a better physical understanding of the bridge hill effect.

Keywords Lutherie • Bridge hill • Tonewood • Dynamical behavior • Sensitivity analysis

38.1 Introduction

Historically, the physics underlying the behavior of musical instruments have been studied primarily using analytical or experimental methods. While these approaches have proven useful in understanding global tendencies, they are not well adapted to the detailed study of the impact of design modifications on the dynamic behavior of the instrument. Virtual prototyping based on physics-based numerical models has proven to be an effective tool in industry for supporting design decisions through effects screening analysis, design optimization, and uncertainty quantification for complex dynamical systems. For fields typically relying on more traditional and experience-based approaches, virtual prototyping may provide a useful tool for attaining a better understanding of the phenomena in play. This study explores the potential of virtual prototyping in its application to stringed instrument making and conservation. Traditional approaches are generally limited by the irreversibility of structural modifications, the high cost of raw materials and construction time, as well as the ethical objections to modifying old and valuable instruments. Virtual prototyping can provide a solution to many of these issues by proposing a decision support tool for both instrument makers and museums curators allowing them to explore new designs, new presets, potential wood substitutions, as well as providing an essential ingredient for model-based or hybrid approaches to sound synthesis. Moreover, virtual prototyping can be particularly useful to investigate a wide range of new designs and the impact of the natural variability in the component material properties. The vibrational behavior of violins has been widely studied and exhibits large variabilities in behavior [1]. Numerical models have been developed over the last decades, from [2] and [3] who used the nascent finite element method to more recent work with refined numerical models that exhibit behavior close to real instruments [4–8]. The objective of the present study is to investigate the effects of violin design and environmental conditions on its dynamical behavior and in particular on the bridge hill phenomenon [9]. The Morris sensitivity method [10] is used to rank the influence of the design parameters including material properties, geometry, and relative humidity.

R. Viala (✉) • V. Placet
FEMTO-ST Institute, University of Bourgogne-Franche-Comté, 24 rue de l'Épitaphe, 25000 Besançon, France
e-mail: romain.viala@univ-fcomte.fr

S. Cogan
FEMSTO-ST, Département de Mécanique Appliquée, Université de Franche-Comté, 24 Chemin de l'Epitaphe, Besançon 25000, France
e-mail: scott.cogan@univ-fcomte.fr

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_38

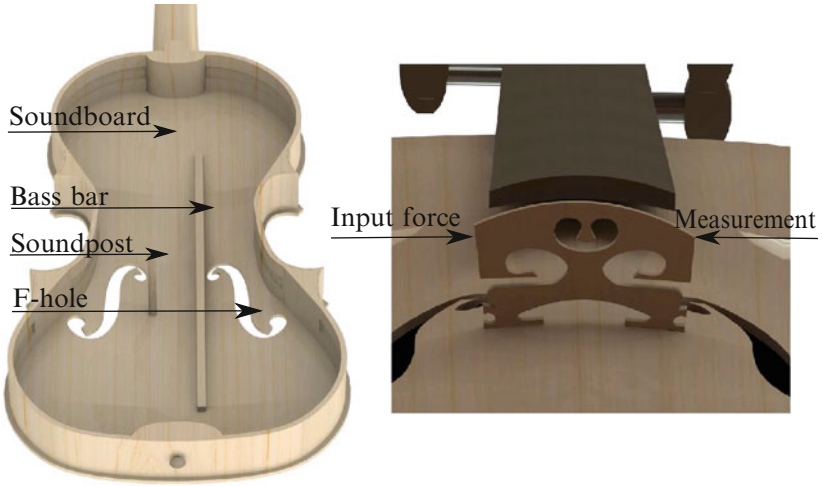


Fig. 38.1 CAD of violin and bridge admittance scheme

38.2 Numerical Modeling of Violin

A parametric computer aided design (CAD) of a violin, shown in Fig. 38.1, is prepared to carry out the geometrical modifications of the violin. Two different kinds of geometrical parameters are taken into account in this study. First, there are the parameters that can not be changed once the violin is built (e.g., the thickness of the soundboard, the position of the bass bar, and the shape of the f-holes). Secondly, there are the parameters that are considered as presets and adjustable by the instrument maker, such as the position of the soundpost and the shape of the bridge. This set of geometrical modifications is not exhaustive but constitutes a starting point, as the effects of these parameters have long been a source of intrigue for instrument makers. Once the CAD is prepared, a detailed numerical model is constructed using the finite element method based on the commercially available software MSC-NASTRAN. The mesh is created using quadratic tetrahedral elements of 4 mm length, and orthotropic material parameters are defined based on the results of earlier studies. All of the violin parts are assumed here to be made of wood, mainly spruce and maple for the body and the neck, and ebony for the fingerboard. It should be noted that wood naturally exhibits both viscoelastic and hygroscopic behaviors that should ultimately be taken into account in the simulation. In the case of spruce, which is used for the soundboard, a finely selected wood, known as tonewood, is used for its desirable characteristics (acoustic radiation, high specific modulus, straighteness of the grain, lack of defects, etc.) as described in [11], which lists the material properties values for woods used in musical instruments. The large number of material parameters associated with the different parts of the violin are implemented in the model and an effects screening has been performed in an earlier study [12] to determine the most influential material parameters. Out of a total of more than 100 material parameters, a set of 15 parameters has been retained for the sensitivity analysis. These parameters include the longitudinal and radial (L and R) Young's moduli, the LR shear modulus and density for the soundboard, the back, the neck and the fingerboard. The moisture content, which is sensitive to the relative humidity, is related to the water absorbed by the wood. It is one of the most influential material parameters as it affects stiffness, density, geometric dimensions and damping. The remaining parameters displaying a relatively weak impact have been fixed and taken from [13]. The subset of influential parameters have been identified from 40 wood samples using a numerical-experimental inverse method discussed in [14]. The response features of interest for the sensitivity analysis are based on a modal analysis of the model in the frequency range from 1 to 5000 Hz, which results in approximately 250 modes. They include the matched eigenfrequency and eigenvector errors as well as the bridge admittance errors between the baseline and modified models. The latter has been experimentally measured in numerous studies [15–18] and consists in the ratio between the in plane displacement at the treble side of the bridge and the input force applied at its bass side, as described in Fig. 38.1. The frequency response function is computed using the modal superposition method with an assumed modal damping of 0.8%, which corresponds to the average value modal damping identified experimentally on four different violins, either new or old.

Table 38.1 Geometrical cases considered and evaluation of their changes with respect to the baseline model

Study case	Description	Averaged MEE (%)	Averaged MAC
1	Baseline geometrical parameters	–	–
2	Change in soundpost position, 2 mm in each direction	0.1	90
3	Carving of the bridge	0.2	78
4	Change in bass bar position, 5 mm closer to the f-hole	0.4	66
5	Both cases 3 and 4	0.3	63
6	Case 5 and change in soundboard thickness (2.3 mm)	1.9	60
7	Case 6 and f-hole displacement (5 mm closer to each edges)	2.9	63
8	Case 7 and bassbar length (30 mm) and height (5 mm) extension	2.2	64

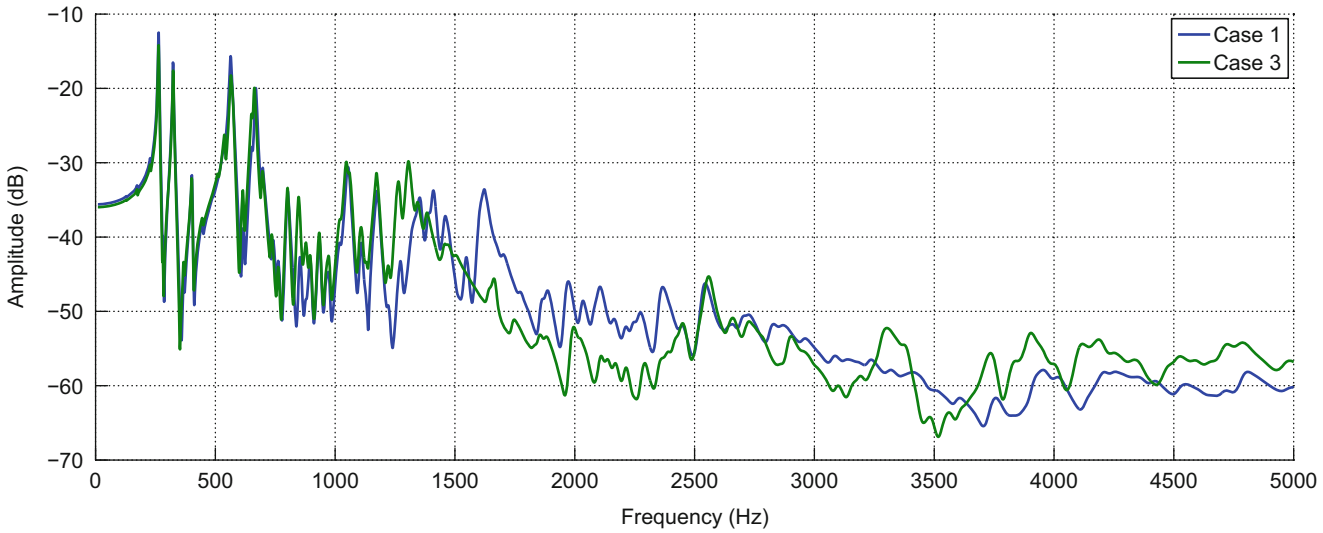


Fig. 38.2 Bridge admittances for cases 1 and 2

38.3 Results and Conclusion

In preparation for the full global sensitivity analysis, a preliminary investigation has been performed based on different geometric configurations shown in Table 38.1. This table also shows the change in the averaged matched eigenfrequencies error (MEE) and eigenvectors between the different modified models and the baseline.

These results provide a preliminary indicator that allow the different geometrical designs to be ranked based on their effects on the modal behaviour of the structure. However, it only covers a small range of the design space and does not take into account the coupling effects that may occur. Figure 38.2 shows the bridge admittances for two cases, baseline geometrical parameters and carved bridge. This highlights certain phenomena not showed in Table 38.1, such as the bridge hill effect described in [9] and [19]. This effect consists in the enhancement of the response of the bridge between 2000 and 3500 Hz and is a phenomenon that is not observed on all violins. According to our results, the bass bar position and bridge carving seem to have an effect on the bridge hill, as this phenomenon appears when these geometrical parameters are changed for cases 3–5, and its amplitude is changed for cases 6–8. Usually it is considered that the soundpost position will adjust the global level of the admittance. Another observation is that the same global trend is seen for all the cases up to 1200 Hz. Above this value the bridge admittance changes significantly for each case. For a more precise ranking, the Morris sensitivity analysis will be used and this will require an automation of the design sensitivity process. Future work will focus on the sensitivity analysis for both geometrical and material parameters changes in order to rank these parameters for the three types of response features. The objective is to highlight whether the variability of the wood or the work of the instrument maker will mainly drive the vibratory behavior of the musical instruments, taking into account the climatic conditions that the violin will encounter.

References

1. Woodhouse, J.: The acoustics of the violin: a review. Rep. Prog. Phys. **77**(11), 115901 (2014). http://doi.org/10.1088/0034-4885/77/11/115901
2. Schwab, H.L., Chen, K.C.: Finite element analysis of a guitar soundboard. Catgut Acoust. Soc. **24**, 13 (1974)
3. Knott, G.A.: A modal analysis of the violin using MSC/NASTRAN and PATRAN. M.S Thesis, Naval Postgraduate School (1987)
4. Gough, C.: Acoustic characterization of violin family signature modes by internal cavity measurements. In: Proceedings of the Stockholm Music Acoustics Conference, Stockholm (2013)
5. Gough, C.: Vibrational modes of the violin family. In: Proceedings of the Stockholm Music Acoustics Conference, Stockholm (2013)
6. Pyrkosz, M.A.: Reverse engineering the structural and acoustic behavior of a Stradivari violin. Dissertation, Michigan Technological University (2013)
7. Buksnowitz, C., Evans, R., Müller, U., Teischinger, A.: Indented rings (hazel growth) of Norway spruce reduce anisotropy of mechanical properties. Wood Sci. Technol. **46**(6), 1239–1246 (2012). http://doi.org/10.1007/s00226-012-0480-0
8. Konopka, D., Gebhardt, C., Kaliske, M.: Numerical modelling of wooden structures. J. Cult. Herit. (2015). http://doi.org/10.1016/j.culher.2015.09.008
9. Woodhouse, J.: On the bridge-hill of the violin. Acta Acust. United Acust. **91**(1), 155–165 (2005)
10. Morris, M.D.: Factorial sampling plans for preliminary computational experiments. Technometrics **33**(2), 161–174 (1991). http://doi.org/10.2307/1269043
11. Bremaud, I., Gril, J., Thibaut, B.: Anisotropy of wood vibrational properties: dependence on grain angle and review of literature data. Wood Sci. Technol. **45**(4), 735–754 (2011). http://doi.org/10.1007/s00226-010-0393-8
12. Viala, R., Placet, V., Cogan, S., Foltete, E.: Model-based effects screening of stringed instruments. In: Conference Proceedings of the Society for Experimental Mechanics Series, vol. 3, pp. 151–156 (2015). http://doi.org/10.1007/978-3-319-15224-0
13. Guitard, D.: Mecanique du materiau bois et composites. Cellection Nabla, Cepadus-Editions, Toulouse (1987)
14. Tam, J.H., Ong, Z.C., Ismail, Z., Ang, B.C., Khoo, S.Y.: Identification of material properties of composite materials using non-destructive vibrational evaluation approaches: a review. Mech. Adv. Mater. Struct. **6494** (2016). http://doi.org/10.1080/15376494.2016.1196798
15. Elie, B., François, G., Bertrand, D.: Analysis of bridge mobility of violins (2013). Retrieved from http://hal.inria.fr/hal-01060528/
16. Maestre, E., Scavone, G.P., Smith, J.O.: Digital modeling of bridge driving-point admittances from measurements on violin-family instruments. In: Bresin, R., Askenfelt, A. (eds.) Proceedings of the Stockholm Music Acoustics Conference, pp. 101–108. Logos Verlag, Berlin (2013)
17. Rodgers, O.E., Masino, T.R.: The effect of wood removal on bridge frequencies. Catgut. Acoust. Soc. **1**, 6–10 (1990). Retrieved from http://www.oosterhofonline.net/pdf/Bridge_wood_removal_Rodgers.pdf
18. Zhang, C., Zhang, G., Ye, B., Liang, L.: Violin bridge mobility analysis under in-plane excitation, 15290–15306 (2013). http://doi.org/10.3390/s131115290
19. Jansson, E.V.: Violin frequency response - bridge mobility and bridge feet distance. Appl. Acoust. **65**(12 Spec. Iss.), 1197–1205 (2004). http://doi.org/10.1016/j.apacoust.2004.04.007

Chapter 39
Experimental Examples for Identification of Structural Systems Using Degree of Freedom-Based Reduction Method

Heejun Sung, Seongmin Chang, and Maenghyo Cho

Abstract Identification method on various structural system has been introduced in many numerical ways to validate complex structures by FEM using experimentally measured data. The objection of this study is to figure out how to identify a perturbed structure model by comparing with measured data based on original FEM data. Identified structure will improve the accuracy and reality to the numerical model by minimizing those differences between two models. Base-line model (original model) is constructed by FEM and will be compared with perturbed model (real model) by IPM (Inverse Perturbation Method). Measured dynamic responses, which is eigenvalues and eigenvectors, will be applied to satisfy the equilibrium and minimize the difference of dynamic responses between base-line model and perturbed model. In experimental examples, due to lack of number of sensor locations which will be located on the model, condensation method is used to restore full model. The equilibrium equation is expressed in terms of measured (primary) and unmeasured (secondary) degree of freedom. In the present study, influence of selection of sensor location and the convergence are considered and selection of sensor algorithm is applied to identification method. Experimental examples demonstrate that the proposed method improves accuracy of identifying perturbed structure model.

Keywords System identification • Sensor location selection method • Inverse perturbation method • Degrees-of-freedom reduction method

39.1 Introduction

As the possibility of numerical calculation increase over and over, demands of calculation skill have been increased evidently. Concept of calculation skill begins with the object of how to minimize the calculation time. Ever-increasing capacity of calculations of computers enabled to calculate the models with complicated designs. Therefore, requirement of models became more complicated and difficult and at the same time it became to be a challenge for engineers. Finite element method (FEM) was one of the most progressive process that have been grown with calculation skills. On the other hand, however, FEM was dismissed due to differences between its real models. Minimizing the difference between real model and analysis model became critical object for engineers. Structural system identification is a method to describe real model on computer analysis. It can be used for damage detection or model update method or any kind of process that needs to compare analysis model with real model.

In this study, structural system identification is used to update numerical model with data from real model. Vibration test data were used to compare its differences. Previous works about structural system identification were compared with only eigenvalues, therefore, only simple models were applied for identification process, such as truss model or beam model [1–3]. Truss or beam models have reasonable number of parameters to solve the identification problems [4]. However, in this study, continuum model was applied to extend its coordinate scale. To apply continuum model, handling improvement of parameters is a crucial point to consider. Also, due to improvement of parameters, eigenvalues from each mode shape seemed to be insufficient. Eigenvectors from each mode shape has been measured from each sensor points and compared with corresponding eigenvectors from numerical model. The concept of degree of freedom reduction method were used to restore full eigenvector data from measured data [5–10]. Few methodologies for structural system identification have been introduced and utilized [10–13]. Inverse perturbation method (IPM) is used to compare numerical model and real model

H. Sung • M. Cho (✉)
Department of Mechanical Engineering, Seoul National University, Seoul, 08826, Republic of Korea
e-mail: heejunsung@snu.ac.kr; mhcho@snu.ac.kr

S. Chang
Korea Atomic Energy Research Institute, Daejeon, 34057, Republic of Korea

R. Barthorpe et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3*,
Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-54858-6_39

[14–16]. Selection of sensor locations [16, 17] were one of the most considerable process of this study and proposed method is provided by comparing node energy from considering mode shapes. By this process, structural differences of real model can be identified and analysis model can be updated. Validations of experimental examples compared to numerical model will be followed.

39.2 Experiment

To compare numerical model from the real model, each model should be designed either in computer or in real first. Then both models should be possible to gather necessary data throughout it. Numerical model can be designed easily with CAE programs and allowed to gather any kind of answers from analysis. However on the other hand, real model has more difficulties in modeling and gathering data from itself. The main reason of considering only eigenvalues in previous works is that the model is simple and it is possible to compare the differences by comparing changes of eigenvalues only.

Eigenvalue and eigenvector can be easily measured from vibration test equipment. Single axis accelerometer, impact hammer and multi-purpose 6-channel input module are used for the vibration test. Frequency response function (FRF) can be measured from each sensor locations and shows different eigenvector from each sensor locations. However eigenvalues from different locations doesn't change by its location. It gives same value throughout one model. Coherence is a value that gives reliability to the experiment. Frequency response function (FRF) can be calculated by dividing the output of frequency domain data and the input frequency domain data from impact hammer which erases effectiveness of hammer. Mode shape of the model which are eigenvector of the model can be constructed by the data from vibration test restoring by degree of freedom-reduction method.

The necessity of sensor location selection algorithm should be considered and confirmed. Due to the large amount of number of parameters, reduction method is applied to restore the full model. When restoring the full model, influence of error from unpredicted points has possibility to ruin the data. By considering sensor location problem, however, it can advance the algorithm of identification of structural system.

39.3 Analysis

As structural system identification is a process to match numerical model with real model, the meaning of this process will decrease without an experimental data. However, after gathering experimental data, the methodology how to compare experimental data to numerical data is another significant process for identification. Eigenvalues and eigenvectors are used as an indicator for this study and the geometry (thickness) has been selected as a parameter.

Inverse perturbation method (IPM) has been used as a methodology to compare two different models. IPM repeats comparing numerical model with the data from experiment. In every step, numerical model's thickness changes and gives different eigenvalue and eigenvector. Then they are to be compared with the experimental data and calculates residual error which shows the differences between two models. While this iteration repeats, the method will find the minimum residual error and will provide the response as thickness value. When these two models have a zero residual error value, it means that these two model has same vibration behavior, which means that they are same model.

To improve the convergence of the problem, some eigenvector were used as compared sensor location. Eigenvector from the comparing sensor location will be compared to the eigenvector from same location which are restored from other sensor location's eigenvector. When this comparison increases, it will increase the residual error and make the calculation to iterate more. This comparison will improve the convergence and help to find more exact geometrical answer to the identification problem.

Simple aluminum plate were suggested as an identification example. One was designed in CAE program with FEM, and the other was designed same as CAE model and modeled with real aluminum with different parameters on some elements. Figure 39.1a shows experimental model and numerical model that were used for structural system identification. Process of identification was verified by the results that are shown in Fig. 39.1b. Parameters were grouped to 21 different groups with same thicknesses. As it is shown in Fig. 39.1b, difference of geometry were founded. Both the location and the amount of thickness are identified accurately.

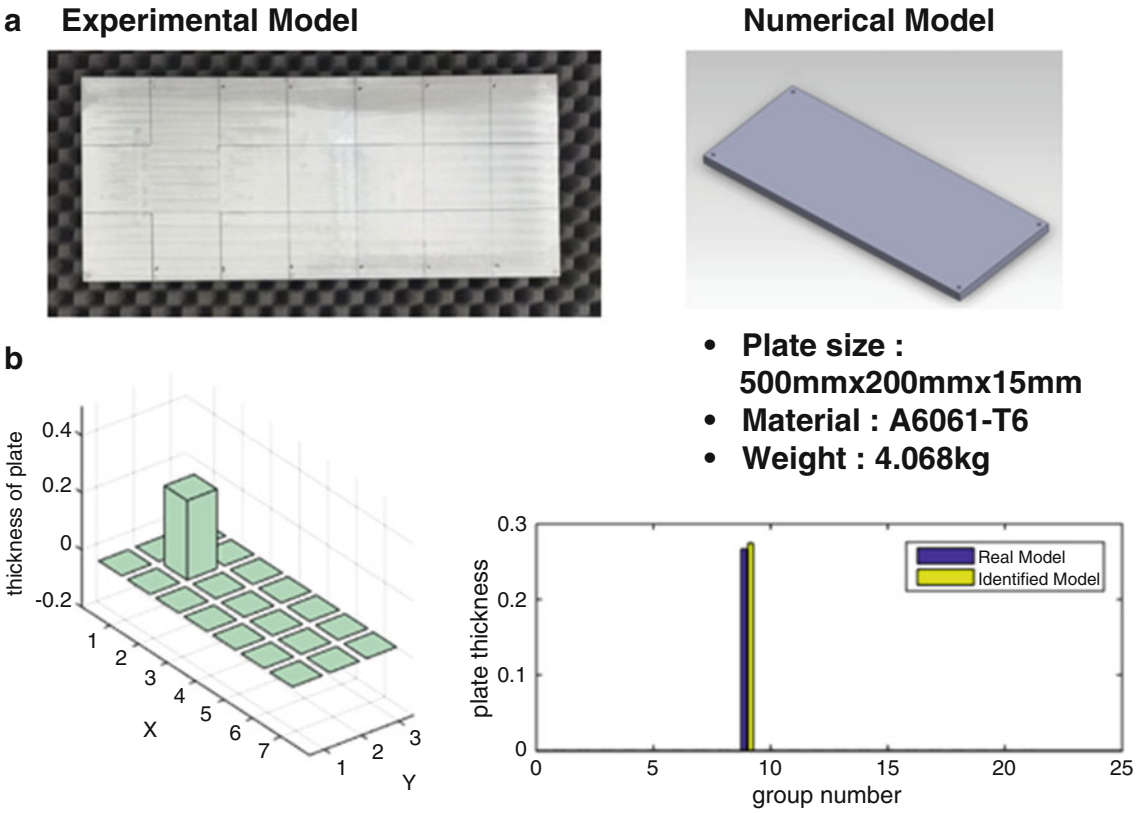


Fig. 39.1 (**a**) Structural system identification model and (**b**) results of identification

39.4 Summary

Identification method for structural system based on finite element method (FEM) has been constructed by this study, especially for continuum model. Difficulties of identification on continuum model was that they have too much parameters to handle. Lots of numbers of parameters were handled by using degree of freedom reduction method. Degree of freedom reduction method is a method that expresses full model accurately by only using some degrees of freedom. As continuum model has huge number of degrees of freedom, it can't be measured from all of them. Therefore, this reduction method is used to make the measurement realistic. Comparison of the eigenvectors are also significant part of the study. IPM was applied for comparison and data from other sensors were used to improve the convergence of IPM. Numerical example was suggested and verified.

Acknowledgement This work was supported by a grant from the National Research Foundation of Korea (NRF) funded by the Korea government (MSIP) (No. 2012R1A3A2048841).

References

1. Liu, P.: Identification and damage detection of trusses using modal data. J. Struct. Eng. **121**, 599–608 (1995)
2. Hajela, P., Soeiro, F.J.: Structural damage detection based on static and modal analysis. AIAA J. **28**(6), 1110–1115 (1989)
3. Alvin, K.F., Park, K.C.: Second-order structural identification procedure via state-space-based system identification. AIAA J. **2010**, 398621 (2010)
4. Chen, H., Kurt, M., Lee, Y.S., McFarland, D.M., Bergman, L.A., Vakakis, A.F.: Experimental system identification of the dynamics of a vibro-impact beam with a view towards structural health monitoring and damage detection. Mech. Syst. Signal Process. **46**, 91–113 (2014)
5. Guyan, R.: Reduction of stiffness and mass matrices. AIAA J. **3**(2), 380 (1965)
6. Friswell, M., Garvey, S., Penny, J.: Model reduction using dynamic and iterated IRS techniques. J. Sound Vib. **186**(2), 311–323 (1995)
7. Kim, H., Cho, M.: Two-level scheme for selection of primary degree of freedom and semi-analytic sensitivity based on the reduced system. Comput. Methods Appl. Mech. Eng. **195**, 4244–4268 (2006)

8. Kim, K., Choi, Y.: Energy method for selection of degree of freedom in condensation. AIAA J. **38**(7), 1253–1259 (2000)
9. O'Callahan, J., Avitabile, P., Riemer, R.: System equivalent reduction expansion process. In: Seventh International Modal Analysis Conference, Las Vegas (1989)
10. Chang, S., Baek, S., Kim, K., Cho, M.: Structural system identification using degree of freedom-based reduction and hierarchical clustering algorithm. J. Sound Vib. **346**, 139–152 (2015)
11. Robertson, A.N., Park, K.C., Alvin, K.F.: Extraction of impulse response data via wavelet transform for structural system identification. J. Vib. Acoust. **120**(1), 252–260 (1998)
12. Weng, S., Xia, Y., Zhou, X.Q., Xu, Y.L., Zhu, H.P.: Inverse substructure method for model updating structures. J. Sound Vib. **331**, 5449–5468 (2012)
13. Alvin, K.F., Park, K.C.: Extraction of substructural flexibility from global frequencies and mode shapes. AIAA J. **37**, 1444–1451 (1999)
14. Kim, K.O., Anderson, W.J., Sandstorm, R.E.: Nonlinear inverse perturbation method in dynamic analysis. AIAA J. **21**, 1310–1316 (1982)
15. Hoff, C.J., Bernitsas, M.M., Sandatorm, R.E., Anderson, W.J.: Inverse perturbation method for structural redesign with frequency and mode shape constraints. AIAA J. **22**, 1304–1309 (1983)
16. Park, Y.C., Choi, Y.J., Cho, J.Y., Kim, K.O.: Inverse perturbation method and sensor location for structural damage detection. Int. Counc. Aeronaut. Sci. **4**, 31–38 (2003)
17. Kammer, D.C., Tinker, M.L.: Optimal placement of triaxial accelerometers for modal vibration tests. Mech. Syst. Signal Process. **18**, 29–41 (2004)