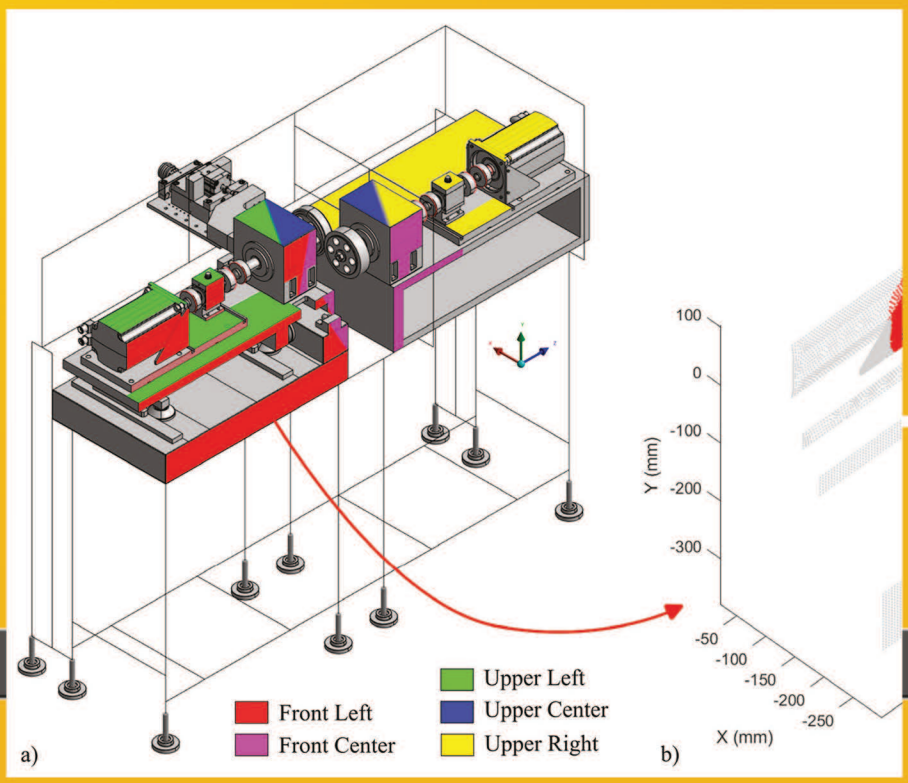
River Rapids

Conference Proceedings of the Society for Experimental Mechanics Series

Model Validation and Uncertainty Quantification, Volume 3

Garrison Flynn
Kyle Neal
Scott Ouellette



Proceedings of the 44th IMAC, A Conference and Exposition on Structural Dynamics 2026

River Publishers

Conference Proceedings of the Society for Experimental Mechanics Series

Series Editor

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

The Conference Proceedings of the Society for Experimental Mechanics Series presents early findings and case studies from a wide range of fundamental and applied work across the broad range of fields that comprise Experimental Mechanics. Series volumes follow the principle tracks or focus topics featured in each of the Society's two annual conferences: IMAC, A Conference and Exposition on Structural Dynamics, and the Society's Annual Conference & Exposition and will address critical areas of interest to researchers and design engineers working in all areas of Structural Dynamics, Solid Mechanics and Materials Research.

Garrison Flynn · Kyle Neal · Scott Ouellette
Editors

Model Validation and Uncertainty Quantification, 2026, Vol. 3

Proceedings of the 44th IMAC, A Conference and Exposition on Structural Dynamics 2026



River Publishers

Published, sold and distributed by:

River Publishers
Broagervej 10
9260 Gistrup
Denmark

www.riverpublishers.com

ISBN 97887-438-1442-9 (Hardback)

ISBN 97887-438-1414-6 (eBook)

https://doi.org/10.13052/rp-9788743814146

Conference Proceedings of the Society for Experimental Mechanics

An imprint of River Publishers

© The Society for Experimental Mechanics, Inc. 2026

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, or reproduction in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use. The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Preface

Model Validation and Uncertainty Quantification represents one of ten volumes of technical papers presented at the 44th IMAC, A Conference and Exposition on Structural Dynamics, organized by the Society for Experimental Mechanics, and held January 19-22, 2026. The full proceedings also include volumes on Nonlinear Structures & Systems; Dynamics of Civil Structures; Dynamic Substructuring & Transfer Path Analysis; Special Topics & Experimental Techniques in Structural Dynamics; Computer Vision & SLDV for Structural Dynamics; Dynamic Environments & Durability Testing; Topics in Modal Analysis & Parameter Identification in Structural Dynamics; Data Science in Engineering; and Structural Health Monitoring, Metamaterials, Metastructures & Additive Manufacturing and Dynamics of Biomedical and Bioinspired Systems.

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 MVUQ Technical Division at the Society for Experimental Mechanics would like to thank the authors, presenters, session organizers, and session chairs for their participation and effort in this track.

Editors:

Garrison Flynn – Los Alamos National Laboratory, Los Alamos, NM, USA;
Kyle Neal – Sandia National Laboratories, Albuquerque, NM, USA;
Scott Ouellette – Los Alamos National Laboratory, Los Alamos, NM, USA

Contents

1 Intelligent Digital Twin Assembly Leveraging Bayesian Classification 1
Shishir Sharma, Nicholas Daly, Stephen Jay Wright, Laura Redmond, Christopher S. McMahan, John R Wagner, and Matthew P. Castanier

2 Bayesian Finite Element Model Updating in the Modal Domain via Integration of Normalizing Flows with Markov Chain Monte Carlo 17
Lin Sun, Zhen Hu, and Michael D. Todd

3 Correlation & Orthogonality – What Do the Numbers Mean? – Use, Misuse and Confusion 25
Peter Avitabile

4 Anti-Resonance-Based Surrogate Modelling for Stochastic FRF Prediction of Finite Element Models 49
Enora Denimal Goy

5 Reliability Analysis of Nonlinear Dynamic Systems via Importance Sampling and Surrogate Modeling 55
Zhi-wei Wang, Michael D. Todd, and Zhen Hu

6 Adaptive Reduced Order Model Selection in Engineering Decision Support 63
Konstantinos Vlachas, Antonios Kamariotis, and Eleni Chatzi

7 Overcoming Data Scarcity for Load Identification and Full-Field Response Reconstruction 73
Aohan Wen and Alana Lund

8 Digital Twin-Based Computer Vision Method for Robust Structural Vibration Estimation in the Underwater Environment 81
Enjian Cai, Zihan Wu, Michael D. Todd, and Zhen Hu

9 Evolution of Modal Characteristics and Model Comparison for a 10-Story Cold-Formed Steel Framed Building under Seismic Excitation 89
Jiachen (Charlie) Zhang, Shokrullah Sorosh, Amanpreet Singh, Benjamin W. Schafer, and Tara C. Hutchinson

10 Modal Identification of Rooftop Mechanical Equipment with Different Attachment Designs on a Full-Scale Ten-Story Shake Table-Tested Building 97
Xianzhao Zhang, Jiachen Zhang, Shokrullah Sorosh, and Tara C. Hutchinson

11 From Data to Decision: Evaluating the Dangers and Benefits of Reliance on a Flexible Discrepancy Term in Model Calibration 105
Alexander Knecht and Sez Russcher

12 System Identification of a 10-story Cold-Formed Steel Building: Systematic Study Considering Sensor Distribution and Changes in Ambient Conditions 109
Shokrullah Sorosh, Jiachen Zhang, M. Khalid Saifullah, Derek Skolnik, and Tara C. Hutchinson

13 Automated Correlation of Simulations and Experiments for Digital Twin Development 115
G. Brunbauer, A. M. Puhwein, and M. J. Hochrainer

Chapter 1
Intelligent Digital Twin Assembly Leveraging Bayesian Classification

Shishir Sharma, Nicholas Daly, Stephen Jay Wright, Laura Redmond, Christopher S. McMahan, John R Wagner, and Matthew P. Castanier

Abstract Digital twins can be used for system performance optimization, health monitoring, predictive maintenance, and investigating system anomalies. A ground vehicle digital twin may be assembled using representations of the subsystem models (tires, suspension system, powertrain, etc.) of different fidelity. The objective of this study is to demonstrate the application of Bayesian Classification within a novel tool for selecting subsystem models to construct suitable digital twins. This tool consists of two phases: first, each subsystem model is calibrated to a set of experimentally collected field or laboratory data (or against the highest fidelity model if test data is not available); second, Bayesian classification is used to select a suitable digital twin based on the likelihood of observing the experimental data from that model. If the probability of observing the data from a digital twin assembled from lower fidelity subsystem models is similar to the probability of observing the data from a digital twin composed of higher fidelity subsystem models, then using the lower fidelity option has several significant advantages, including computational time and engineering cost. The proposed tool can assess many possible mixtures of fidelities among subsystems and thus select only the high-fidelity representations where they are truly needed. A demonstration of the Bayesian Classification method is provided using a wheeled ground vehicle dynamics model

Shishir Sharma
Graduate Research Assistant, Department of Civil Engineering, Clemson University, S Palmetto Blvd, Clemson, SC 29634
e-mail: shishis@clemson.edu

Nicholas Daly
Graduate Research Assistant, Clemson University, Fluor Daniel Building, 216 South Palmetto Boulevard, Clemson, SC 29634-0921
e-mail: ndaly@g.clemson.edu

Stephen Jay Wright
Graduate Research Assistant, Clemson University, Lowry Hall: 306 S. Palmetto Blvd, Clemson, SC 29634
e-mail: swrigh7@g.clemson.edu

Laura Redmond
Associate Professor, School of Civil and Environmental Engineering and Earth Sciences, and School of Mechanical and Automotive Engineering, Clemson University, 116 Lowry Hall, Clemson, SC, 29634
e-mail: lmredmo@clemson.edu

Christopher S. McMahan
Professor, School of Mathematical and Statistical Sciences,
Clemson University, O-102 Martin Hall, 220 Parkway Drive, Clemson, SC 29634-0975
e-mail: mcmaha2@clemson.edu

John R Wagner
Professor of Mechanical Engineering, PLM Center Director, College of Engineering, Computing and Applied Sciences,
Clemson University, 212 Fluor Daniel Engineering Innovation Building, 216 South Palmetto Boulevard, Clemson, SC 29634-0921
e-mail: jwagner@clemson.edu

Matthew P. Castanier
Research Mechanical Engineer, US Army DEVCOM Ground Vehicle Systems Center,
6501 E. 11 Mile Rd., MS 157, Warren, MI 48397-5000
e-mail: matthew.p.castanier.civ@army.mil

DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited. OPSEC10040

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

created in Matlab/Simulink. Simulated field data was generated using the highest fidelity model with added white noise. The numerical results show that the innovative classification tool can identify the best combination of tire and suspension subsystem models within a vehicle dynamics model to represent the simulated field data. Future work will include expanding the demonstration to include more digital twin subsystem models and a demonstration with physical data.

Keywords Digital twin · Model selection · Bayesian classification · Mixture modeling · Uncertainty quantification

Introduction

A Digital Twin is a virtual representation of an actual physical asset in operation from the microatomic level to the macro-geometric level that can be used to evaluate the current condition of the asset and predict future behavior, refine control, or optimize operation [3],[13]. It can be developed and used throughout the product life cycle, from design to manufacturing, through deployment operations [19] and for system upgrade evaluations. Digital Twins can effectively reduce costs by enabling virtual testing, integrating physical test outcomes, and facilitating timely, data-driven decisions by providing comprehensive insights into individual components and the asset as a whole [20].

A complex system, such as a ground vehicle, usually has multiple Digital Twins compiled into a single universal Twin [25],[21]. A ground vehicle consists of multiple subsystems like the tires, suspension system, chassis, powertrain, etc. The digital twin to represent a ground vehicle may be assembled using different fidelity representations of such subsystem models. Often a developer's ambition to build a Digital Twin that is identical to the physical system, leads to the creation of an overly complex Digital Twin [26]. However, the fidelity of the subsystem models impacts the computational cost of the Digital Twin. An efficient Digital Twin must balance the fidelity of predictions with the computational demands (time, power and complexity) [2][21].

The selection of an appropriate Digital Twin model has often been conducted through decision-based algorithms selecting the best Digital Twin according to use cases. Such approaches can determine the overall class of model that match the intended service or performance requirements, while not directly considering the appropriateness of the underlying physics or system architecture to adequately capture the output of interest. For example, [14] formulates a decision model for quantification of the suitability of DT applications and targeted selection of data-driven DT applications based on value-effort analysis. Similarly, [1] focuses on prioritizing DT applications based on stakeholders' needs, technical feasibility, business needs, and risk factors.

Physics-based Digital Twin selection methodologies have also been put forth that aim at selecting Digital Twin models based on their ability to replicate the actual physical system behavior. Such methods are often data-driven, developed to match the Digital Twin output to some experimental data from the physical system. Efforts have been made toward the creation and selection of the optimal set of data-based reduced-order component models to comprise a digital twin using optimal decision trees, an interpretable machine learning method [8]. The key differences between this work and that proposed in [8] is that the decision tree does not provide statistical insight into its decision, which would allow a user to select the next best model if the marginal cost in computation time for the increased accuracy of the optimal model was deemed too high. Other notable contributions in this area include Digital Twin selection via dynamic Bayesian networks [7] and reinforcement learning [17]. Unlike, [7] which primarily focuses on updating unknown parameters throughout the lifetime of an asset to account for physical degradation, our approach seeks to identify and calibrate different digital twin assemblies and more aptly accounts for uncertainty propagation throughout the selection process by explicitly accounting for the uncertainty associated with the calibration of each assembly. Further, all evaluations of the digital twin in the proposed approach can be parallelized thereby reducing the total time required to make the selection. In contrast, the technique in [17] must assess a posited assembly at each iteration of their approach; these serial computations would render their proposal intractable for computationally intensive assemblies. Further, given the use of approximate Bayesian computation techniques, the compute time is highly dependent on the number of unknown parameters and their ranges.

We propose a novel mixture-based Bayesian machine learning framework for intelligent selection of Digital Twin assembly by identifying and selecting the appropriate combination of the subsystem models with varying fidelities. While Bayesian methods have been implemented in the Digital Twin domain for model updating [23] or calibrating model parameters [6], the proposed Bayesian Assembly Tool integrates Bayesian model calibration and Bayesian classification to systematically assess Digital Twin models comprised of different combinations of different fidelity subsystem models and select the DT with the most appropriate mixture of subsystem model fidelities while also providing statistical insight to the decision. By embedding engineering models of varying fidelities as components in a mixture model and deploying a reduction mechanism that promotes lower-fidelity model, this framework delivers data-driven, uncertainty-aware, and computationally efficient digital twin assemblies, thereby enhancing parsimony.

The proposed method has five major steps. The first step is to acquire laboratory test data for calibration. The second step involves creating Digital Twin assemblies with combinations of different fidelity subsystem models and calibrating to the lab- oratory test data. The calibration process essentially creates a set of representative output data from each digital twin assembly that is not affected by the improper selection of the uncertain input parameters. Griddy Gibbs calibration methodology [6] has been implemented in this demonstration, which produces posterior distributions of the uncertain model parameters. In the third step, a set of field test data is obtained for classification. This data can be the same as the laboratory test data used for cali- bration or a new one under different known parameter settings. Next, synthetic field data are generated from the Digital Twin Assemblies for the unknown parameters sampled from the posterior distributions obtained from the Griddy Gibbs calibration. The synthetic field data from each assembly is fitted to a multivariate normal distribution with a unique mean and covariance structure. The final step is Bayesian classification, where the probability of observing the field test data given the mean and co- variance of each assembly is evaluated. The model with the highest probability is the best-fitting model selected by the Bayesian Assembly Tool. In addition, the Assembly tool can be used to analyze if more than one combination of the subsystem models fits the observed data based on the probability of observing the data, allowing the users to select the Digital Twin assembly that is least computationally demanding.

Methodology

The workflow of the Bayesian Assembly Tool for Digital Twin selection involves five basic steps, and is schematically shown in Figure 1.

1. *Obtain Laboratory Test Data for Calibration*: Laboratory test data $\boldsymbol{Y} = (\boldsymbol{Y}_1, \boldsymbol{Y}_2, ..., \boldsymbol{Y}_n)$ are obtained from experiments in the physical vehicle or generated from a high-fidelity Digital Twin simulation with added measurement error under known parameters $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2, ..., \boldsymbol{X}_n)$. The laboratory test data $\boldsymbol{Y}$ under the known parameters $\boldsymbol{X}$ is denoted by $\boldsymbol{D}$.
2. *Griddy Gibbs Model Calibration*: A suite of Digital Twin assemblies $\boldsymbol{M}_1, \boldsymbol{M}_2, ...\boldsymbol{M}_m$ with different fidelity subsystem models is created, and each of them is calibrated to the laboratory test data $\boldsymbol{Y}$. The Griddy Gibbs calibration method

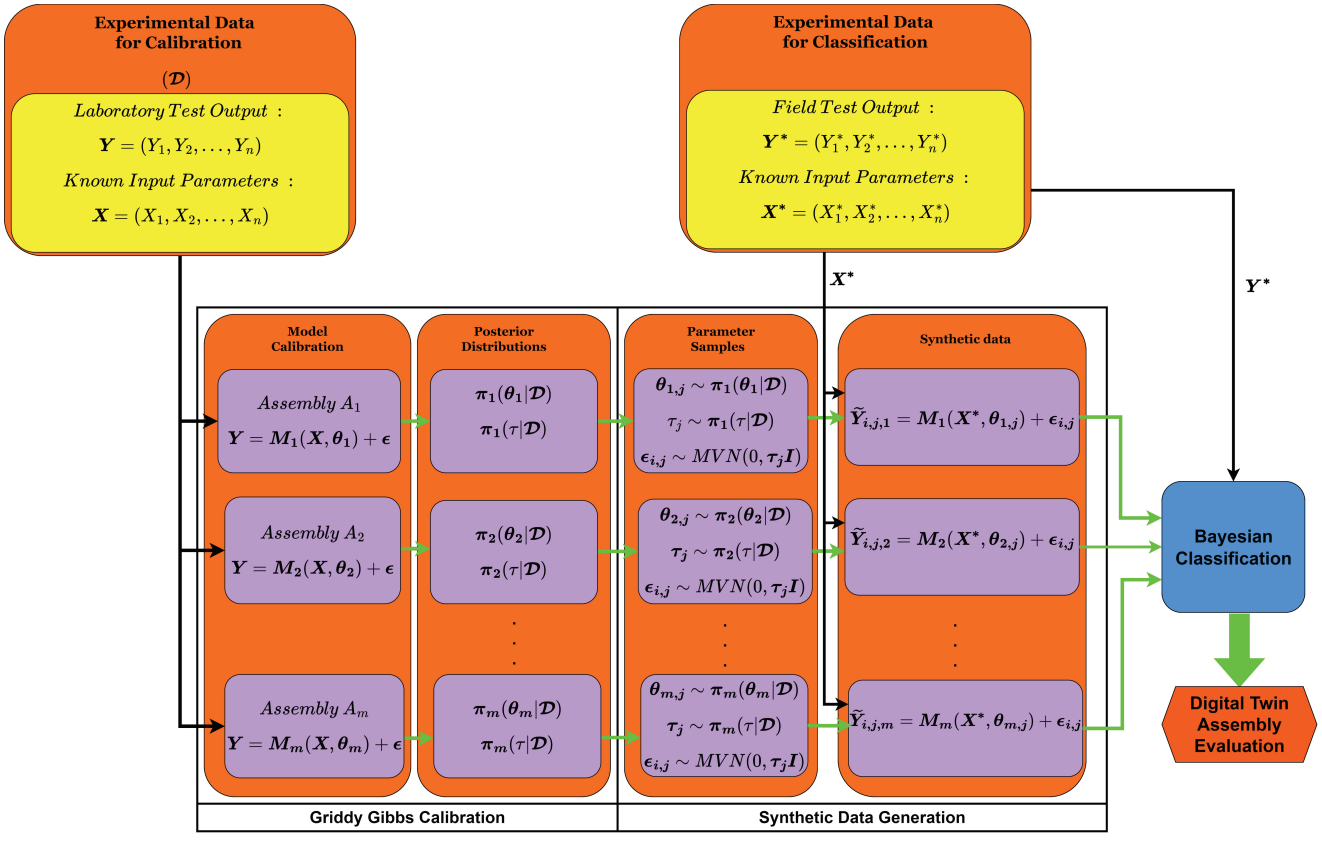


Fig. 1 Bayesian Assembly Tool working flowchart

is used, which produces the posterior distributions $\pi_m(\boldsymbol{\theta}_m|\boldsymbol{D})$ of the uncertain parameters $\boldsymbol{\theta}_m$ and the posterior distributions $\pi_m(\tau_m|\boldsymbol{D})$ of the precision parameter τ_m governing the error distribution from each assembly m.

3. *Obtain Field Test Data for Classification*: A set of field test data, $\boldsymbol{Y}^* = (\boldsymbol{Y}^*_1, \boldsymbol{Y}^*_2, ..., \boldsymbol{Y}^*_n)$, is obtained, which may be the same used for the calibration or a new set of data under different known parameter settings $\boldsymbol{X}^* = (\boldsymbol{X}^*_1, \boldsymbol{X}^*_2, ..., \boldsymbol{X}^*_n)$.
4. *Synthetic Field Data Generation from Digital Twin Assemblies*: The uncertain parameters $\boldsymbol{\theta}_{m,j}$ for each Digital Twin Assembly m are sampled from their posterior distributions $\pi_m(\boldsymbol{\theta}_m|\boldsymbol{D})$. Then, the Digital Twin Assemblies are evaluated at these parameters $\boldsymbol{\theta}_{m,j}$ under the known parameters $\boldsymbol{X}^*$ to obtain the model output $\boldsymbol{M}_m(\boldsymbol{X}^*, \boldsymbol{\theta}_{m,j})$. The precision parameter $\tau_{m,j}$ is sampled from the posterior distribution $\pi_m(\tau_m|\boldsymbol{D})$, and a Gaussian noise $\boldsymbol{\epsilon}_{m,i,j}$ is sampled from a multivariate normal distribution with $\mathbf{0}$ mean and covariance $\tau^{-1}_{m,j}\,\mathbf{I}$. This noise is added to $\boldsymbol{M}_m(\boldsymbol{X}^*, \boldsymbol{\theta}_{m,j})$ to obtain large sets of synthetic data $\tilde{\boldsymbol{Y}}_{m,i,j}$ from each assembly. Then, a multivariate normal distribution is fitted to the synthetic data, and the corresponding mean vector and covariance matrix are estimated for each of the Digital Twin assemblies.
5. *Model Selection by Bayesian Classification*: The final step is to evaluate the likelihood of obtaining the data $\boldsymbol{Y}^*$ given the mean and covariance structure of each assembly. Based on the relative probability of each assembly and the computational cost associated with the assembly (given different fidelity levels), the optimal Digital Twin assembly can be selected.

Each step of the methodology is described in detail in the subsequent sections.

Obtain Laboratory Test Data for Calibration

A set of experimental laboratory data $\boldsymbol{Y}$ is acquired from the physical vehicle system under known input parameters $\boldsymbol{X}$ (true parameters). This data is used to calibrate the Digital Twin assemblies with various possible combinations of subsystem models of different levels of fidelity. An example of such data could be the time history of normal force on the tire when the vehicle passes over a speed bump. The data can be acquired from laboratory tests of vehicle systems or a controlled field test. However, if such a test is not feasible, simulated laboratory test data can be generated from the highest-fidelity Digital Twin Assembly, and the output can be used to calibrate the other Digital Twin Assemblies under consideration.

Time domain data should be processed to obtain lower resolution data (such as local maxima/minima, data at significant time points, integrated parameters, or frequency-domain responses) to avoid artificially reducing the uncertainty in the posterior distributions given by the Griddy Gibbs calibration.[6]

Griddy Gibbs Model Calibration

The next step is calibrating the Digital Twin assemblies $\boldsymbol{M}_1, \boldsymbol{M}_2, ..., \boldsymbol{M}_m$ of varying fidelity, with the laboratory test data $\boldsymbol{Y}$. There are numerous deterministic and statistical model calibration methods that could be used for this purpose; e.g., see [24].[18],[9]. We have implemented the Bayesian Griddy Gibbs model calibration method developed in [6], except for a slight modification in the optimization process used to identify the feasible region. In particular, in the optimization process used herein, the sum of squared error between laboratory test data $\boldsymbol{Y}$ and the Digital Twin assembly output $\boldsymbol{M}_m(\boldsymbol{X}, \boldsymbol{\theta}_m)$ for unknown parameter settings $\boldsymbol{\theta}_m$ and known parameter settings $\boldsymbol{X}$ is taken as the objective function. By minimizing this objective function, via a non-linear optimization algorithm, we identify an optimal solution $\boldsymbol{\theta}_{mo}$. The normal likelihood of observing the laboratory test data given the model output at the optimal solution as the mean and the maximum likelihood estimate of the covariance as the covariance is evaluated. This likelihood corresponds to the maximum likelihood. Next, to evaluate the feasible region around $\boldsymbol{\theta}_{mo}$, each dimension of $\boldsymbol{\theta}_{mo}$ is iteratively varied by a certain fixed step, keeping the other dimensions fixed at the optimal solution, and the likelihood is calculated at each step. When the likelihood drops to 1% of the maximum likelihood, the algorithm is stopped, and the same process is repeated for the other dimensions of $\boldsymbol{\theta}_{mo}$. In this way, the upper and lower boundaries for $\boldsymbol{\theta}_m$ are computed.

Obtain Field Test Data for Classification

For the final classification of the Digital Twin Assemblies, another data set is required. This data set would preferably come from a field test, but if these data are not available, the classification could be conducted based on the laboratory test data, or simulated field test data can be generated from the highest-fidelity Digital Twin Assembly. Unlike the laboratory test data, the field test data used for classification do not need to be down-sampled or summarized.

Synthetic Field Data Generation from Digital Twin Assemblies

The next step after calibration is to generate synthetic field data from each of the calibrated assemblies. The uncertain parameters $\boldsymbol{\theta}_{m,j}$ are sampled from their respective posterior distributions $\pi_m(\boldsymbol{\theta}_m|\boldsymbol{D})$ obtained from calibration of assembly $\boldsymbol{M}_m$. The number of parameter samples should be sufficient to represent the posterior distributions, which may be less than the number of Griddy Gibbs samples forming the posterior distributions. Digital Twin simulations are run for those parameters under the same known parameter settings $\boldsymbol{X}^*$ as in the field test data to obtain the model output $\boldsymbol{M}_m(\boldsymbol{X}^*, \boldsymbol{\theta}_{m,j})$. The precision parameters $\tau_{m,j}$ are sampled from their posterior distributions $\pi_m(\tau_m|\boldsymbol{D})$, and the Gaussian noise $\boldsymbol{\epsilon}_{m,ij}$ is sampled from a multivariate normal distribution with $\mathbf{0}$ mean and covariance $\tau^{\ 1}_{m,i,j}\ \mathbf{I}$. This noise is added to the simulation output from the assemblies to obtain the simulated field data $\boldsymbol{Y}^*$ from each assembly, i.e., the synthetic field data is given by:

$$\widetilde{\mathbf{Y}}^*_{m,i,j} = \boldsymbol{M}_m\left(\boldsymbol{X}^*_i, \boldsymbol{\theta}_{m,j}\right) + \boldsymbol{\epsilon}_{m,i,j} \tag{1}$$

A multivariate normal distribution is then fitted to these data, and estimates of the mean vector and covariance matrix are obtained for each of the assemblies. These multivariate normal distributions are used during Bayesian classification for model selection.

Model Selection by Bayesian Classification

The final step is the selection of the appropriate Digital Twin assembly using the Bayesian classification method. This involves evaluating the probability of observing the field test data given the aforementioned multivariate normal distributions of each assembly. This measure provides the relative likelihood of observing the field test data ($\boldsymbol{Y}^*$) under each of the potential assem- blies; i.e., which assembly best aligns with the field test data. The assembly corresponding to the highest likelihood is identified as the best representation for the field data by the Bayesian Assembly Tool.

However, the Digital Twin assembly with the highest probability is not necessarily the best choice overall. In selecting the optimal Digital Twin, the fidelity of different subsystem models should also be considered, along with how frequently the assembly is chosen by the Bayesian Assembly Tool. If a relatively lower-fidelity assembly yields a comparable likelihood of being selected to that of a higher-fidelity one, the lower-fidelity version may be a more efficient option, given the lower computational costs and demands associated with the lower-fidelity model.

Demonstration Problem: Vehicle Dynamics Model in MATLAB Simulink

A Digital Twin model of a wheeled vehicle was created for this demonstration in MATLAB Simulink. Four Digital Twin assemblies were used, consisting of combinations of two tire models and two suspension models, while rest of the subsystems were identical.

Tire Models

The first tire model used was the Fiala tire. The model is available in the Simulink library as "Fiala Wheel 2DOF" block. The Fiala tire model considers the tire acts as a beam on an elastic foundation. It assumes that the contact patch is rectangular and pressure is uniform across it. It neglects the effects of the tire's camber angle. The model calculates the normal force using equation (2), where the notation mirrors the Matlab documentation; see [5] for further discussion.

$$F_{ztire} = \rho_z k - b\dot{z} \tag{2}$$

The second tire model used was the PAC2002 tire which is relatively more complex than the Fiala model. The "Combined Slip Wheel 2DOF" block available in the Simulink library is used to model the Pac2002 tire [4]. In the PAC2002 model, the vertical tire response is modeled as a spring and a damper in parallel. The contact patch is found assuming the tire acts as a rigid disc. The model calculates the normal force using equation(3).

$$F_z = \left\{1 + q_{V2}\left|\Omega\right|\frac{R_o}{V_o} - \left(q_{Fcx}\frac{F_x}{F_{zo}}\right)^2 - \left(q_{Fcy}\frac{F_y}{F_{zo}}\right)^2\right\}.\left\{\left(q_{Fz1} + q_{Fz3}\gamma^2\right)\frac{\rho_z}{R_o} + q_{Fz2}\frac{\rho_z^2}{R_o^2}\right\}.\left(1 + p_{pFz1}dp_i\right).F_{zo} \tag{3}$$

See Table 3 in Appendix for descriptions of the variables used in equation (3) and [16],[22] for a more discussion.

Suspension Models

The two suspension models were used in the Digital Twin assemblies: MacPherson suspension and solid axle suspension.

The MacPherson suspension model used in the Digital Twin uses a linear spring and damper to model the vertical dynamic ef- fects of the suspension system. Using the relative positions and velocities of the vehicle and wheel carrier, the block calculates the vertical suspension forces on the wheel and vehicle. The block uses a linear equation that relates vertical damping and com- pliance to the suspension height, suspension height rate of change, and absolute value of the steering angles. The MacPherson

suspension is available in the Simulink library as “Independent Suspension - MacPherson”. The block implements equation (4) to calculate the suspension force, where the notation mirrors the Matlab documentation; see [11] for a further discussion.

$$F_{wz_{a,t}} = F_{z0_a} + k_{z_a}\left(z_{v_{a,t}} - z_{w_{a,t}} + m_{h\mathrm{steer}_a}\left|\delta_{\mathrm{steer}_{a,t}}\right|\right) + c\left(\dot{z}_{v_{a,t}} - \dot{z}_{w_{a,t}}\right) + F_{zhstop_{a,t}} + F_{zaswy_{a,t}} \tag{4}$$

See Table 4 in Appendix for descriptions of the variables used in equation (4).

Solid axle suspension block used in the Digital Twin models the suspension compliance, damping, and geometric effects as functions of the wheel positions and velocities, with axle-specific compliance and damping parameters. Using wheel position and velocity, the block calculates the vertical wheel position and suspension forces on the vehicle and wheel. The "Solid axle Suspension - Coil Spring” block is used to model the solid axle suspension in Simulink. The block implements equation (5), where the notation mirrors Matlab documentation; see [12] for further discussion.

$$F_{wz_{a,t}} = -F_{wa_{z0}} + k_{wa_z}\left(z_{w_{a,t}} - z_{s_{a,t}}\right) - c_{w_{az}}\left(\dot{z}_{w_{a,t}} - \dot{z}_{s_{a,t}}\right) \tag{5}$$

See Table 5 in Appendix for descriptions of the variables used in equation (5).

Implementation of the Bayesian Assembly Tool

Simulated Laboratory Data Generation

To validate the Bayesian assembly tool, four different Digital Twin assemblies were created by combining the two tire and two suspension systems:

- M_1: Pac2002 tire and MacPherson suspension (high-fidelity assembly)
- M_2: Pac2002 tire and Solid axle suspension
- M_3: Fiala tire and MacPherson suspension
- M_4: Fiala tire and Solid axle suspension

The simulated laboratory data was generated from the Digital Twin assembly M_1. The input parameters (true parameters) are shown in Table 1

Table 1 True parameter setting for simulated experimental data generation

Parameter	True value
Tire Stiffness $(\theta(1))$	$180000\ \mathrm{N/m}$
Tire Damping $(\theta(2))$	$3000\ \mathrm{Ns/m}$
Suspension Stiffness $(\theta(3))$	$30000\ \mathrm{N/m}$
Suspension Damping $(\theta(4))$	$2500\ \mathrm{Ns/m}$

The simulated laboratory data was generated under the scenario where the Digital Twin accelerated along a straight path from 0 to 20 mph in the first five seconds and then continued forward with a constant velocity. The road profile used for the simulation was smooth from 0 to 6 seconds. At the sixth second, the front wheel of the vehicle encountered a speed bump 200 mm high and 100 mm long. The road profile was flat for the remainder of the simulation. The road profile used for the simulation is shown in Figure 2, and the reference velocity profile is shown in Figure 3. The model output considered for calibration was the normal force on the left front wheel hub. White noise representing the measurement error was added to the Digital Twin simulation output. The noise was estimated by assuming the measurement would be taken with a $20kN$ force transducer assuming a measurement error of 2%, we took the 3σ error to be $400N$.

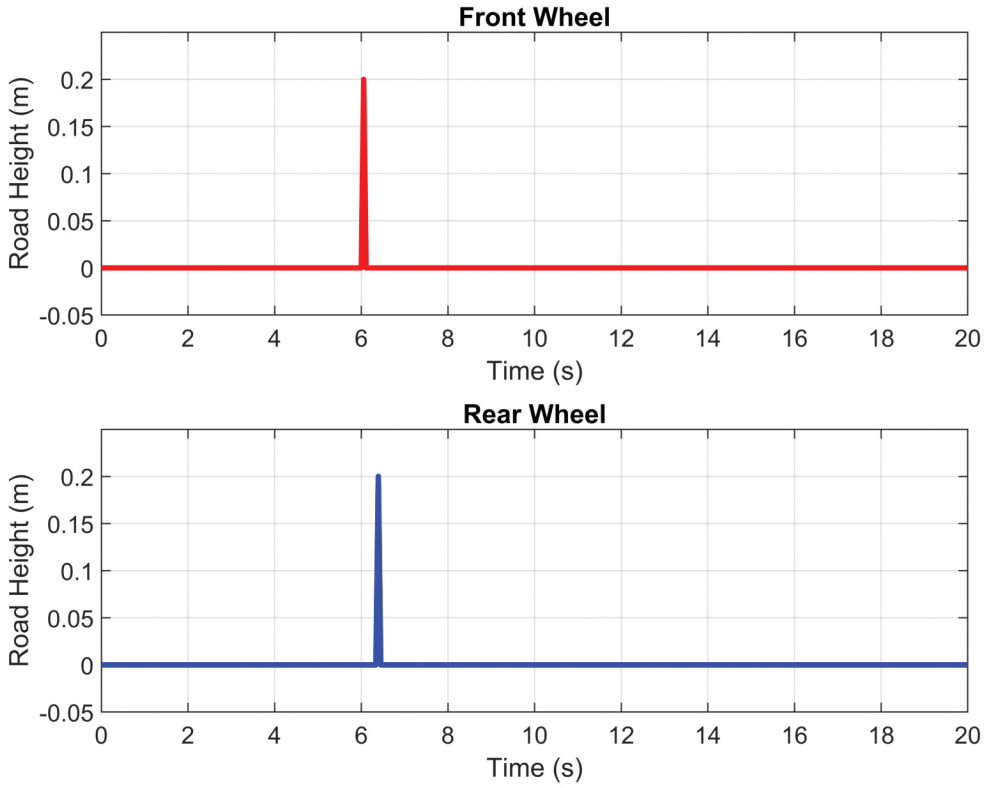


Fig. 2 Road profile for the front and rear wheel (y-axis is road elevation in meters) used in the simulation

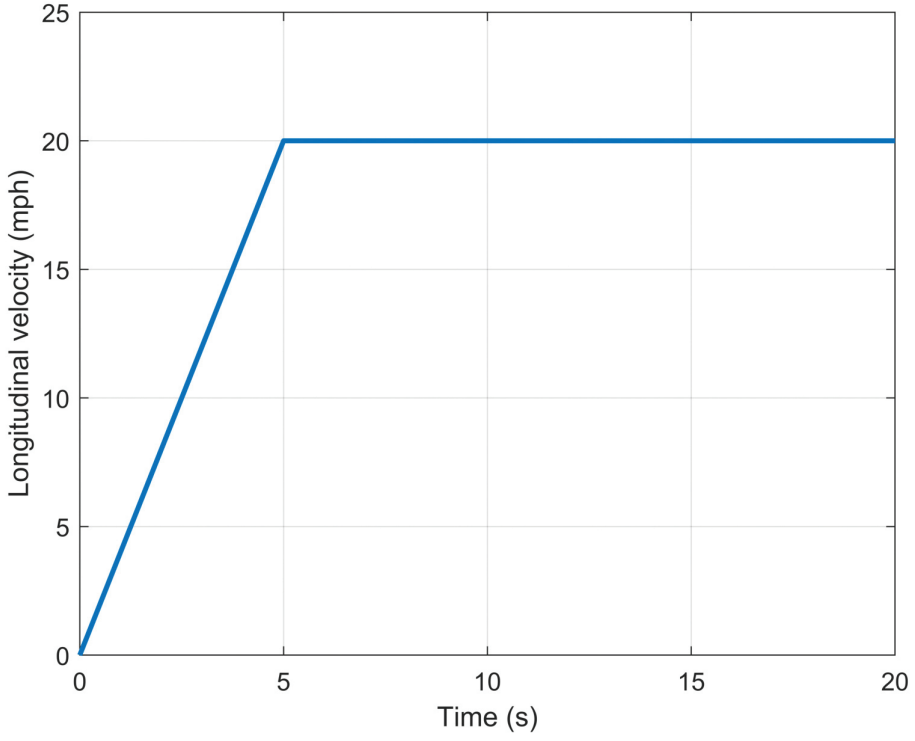


Fig. 3 Longitudinal velocity reference (y-axis is the velocity in mph)

This Gaussian noise was sampled ten times from a normal distribution with 0 mean and 18000N variance and added to the normal force output of the assembly M_1 under the aforementioned true parameters and simulation settings to obtain the ten sets of high-resolution simulated laboratory data, as shown in Figure 4.

The normal force at the wheel hub of the front left tire was originally sampled every 0.001 seconds. Calibrating the Digital Twin assemblies to such high-resolution data would produce a narrow posterior distribution due to an artificial reduction in the uncertainty in the uncertain parameter estimation. To avoid this, the full-time history data was condensed to include only the response at a few significant time points near where local maxima and minima typically occur in the models. Those timesteps were 6.422, 6.558, 6.689, 6.973, and 7.291 seconds. This was the actual simulated laboratory data used for calibration, which is shown in Figure 5.

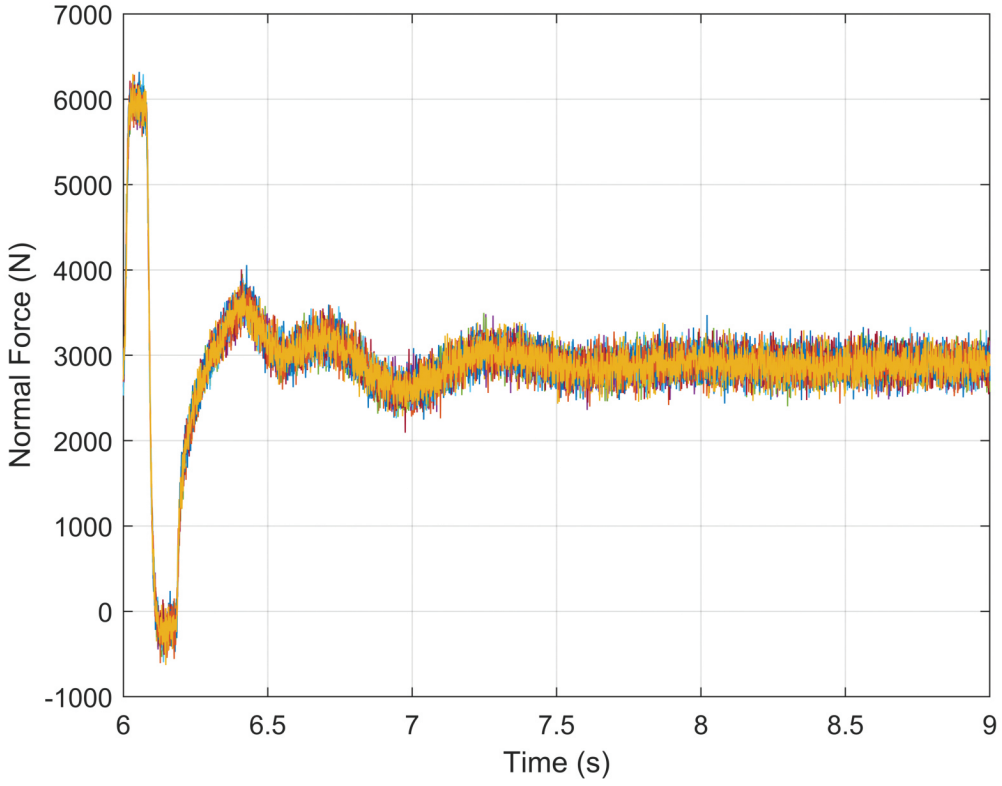


Fig. 4 Simulated experimental data (full time history)

Griddy Gibbs Model Calibration

The four Digital Twin assemblies were calibrated to the condensed simulated laboratory data. First, for each assembly, the full time history of the normal force from the left front wheel hub in the model was down-sampled to only the response at the five time points of the reduced simulated laboratory data. A non-linear optimization algorithm, 'fmincon' [10] in MATLAB, was used to determine the feasible region for the Griddy Gibbs model calibration. The sum of squared error between the simulated experimental data Y and the Digital Twin output $f(\theta, X)$ was the objective function minimized. The constraints for the optimization were the upper and lower boundaries of the four parameters given in Table 2.

After identifying an optimal solution at which to begin defining the feasible region, the upper and lower limits of the feasible region were determined for each parameter, one at a time, keeping the other three parameters at their optimal solution. The boundary of the feasible region was determined by finding the parameter value where the likelihood of observing the output drops to 1% of the maximum likelihood (at the optimal solution). Once the boundaries for all four parameters were set, a grid of 2000 set of parameters was created using Latin Hypercube Sampling and the Digital Twin assembly was evaluated for those parameters. Then, the Griddy Gibbs calibration was run, where 50000 Griddy Gibbs samples of the four uncertain parameters and the precision parameter were taken, 15000 samples were burnt, and the remaining samples formed

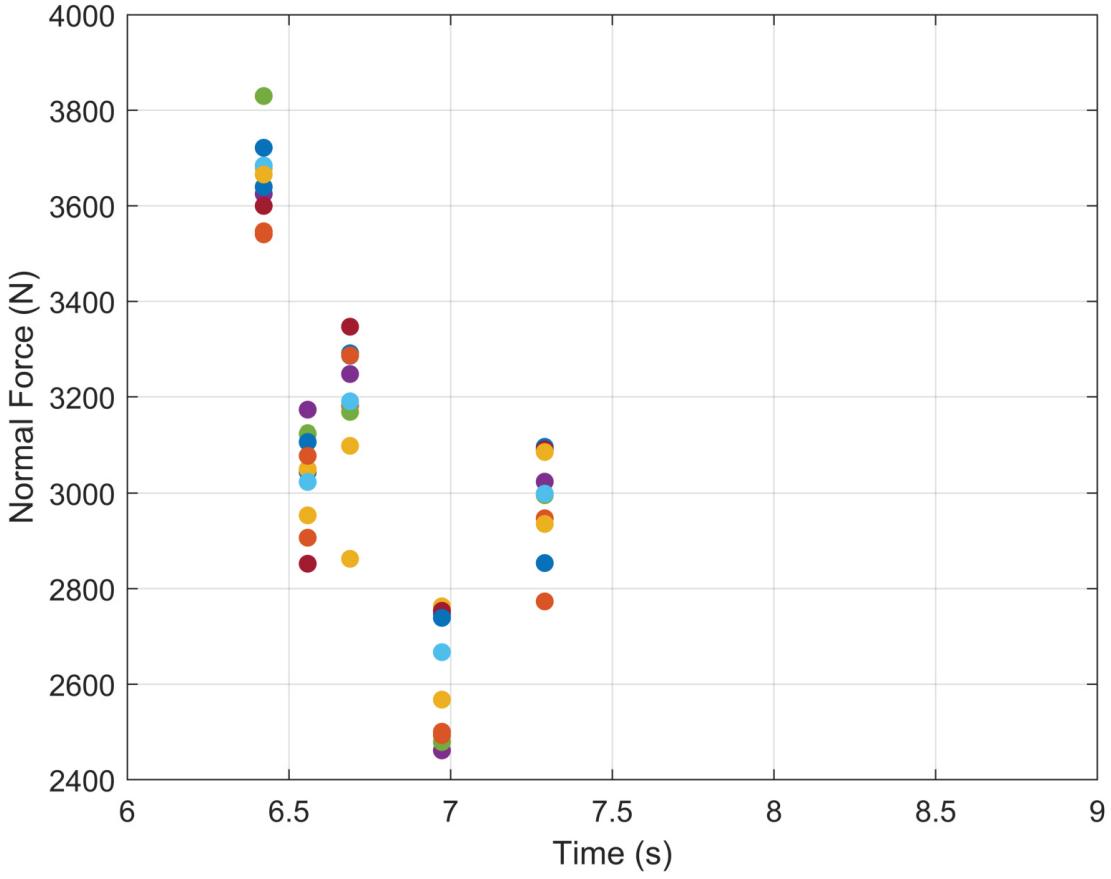
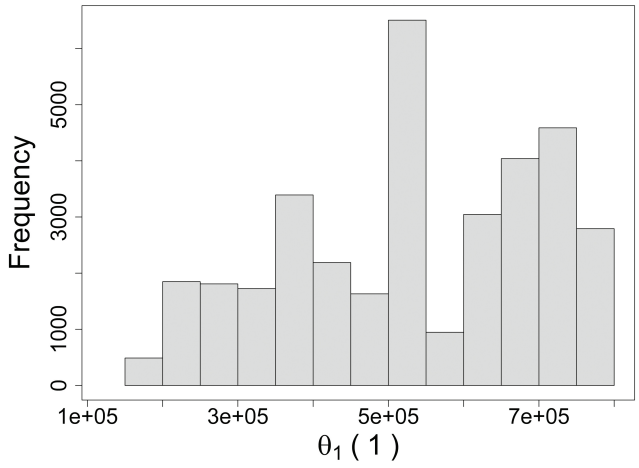


Fig. 5 Simulated laboratory data

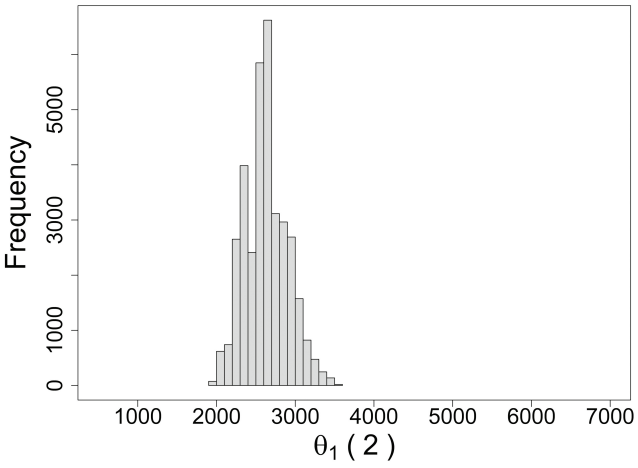
Table 2 Calibration Parameters and Ranges for Different Digital Twin Assemblies

Digital Twin Assembly	Calibrating Parameters	Range
Assembly 1 (MacPherson suspension and Pac2002 tire, high-fidelity)	Pac2002 Tire stiffness ($\theta_1(1)$)	120000 to 800000 N/m
	Pac2002 Tire damping ($\theta_1(2)$)	500 to 7000 N·s/m
	Suspension stiffness ($\theta_1(3)$)	10000 to 180000 N/m
	Suspension damping ($\theta_1(4)$)	500 to 30000 N·s/m
Assembly 2 (MacPherson suspension and Fiala tire)	Fiala Tire stiffness parameter ($\theta_2(1)$)	400 to 8000 N/m
	Fiala Tire damping parameter ($\theta_2(2)$)	3000 to 180000 N·s/m
	Suspension stiffness ($\theta_2(3)$)	10000 to 180000 N/m
	Suspension damping ($\theta_2(4)$)	500 to 30000 N·s/m
Assembly 3 (Solid axle suspension and Pac2002 tire)	Pac2002 Tire stiffness ($\theta_3(1)$)	120000 to 800000 N/m
	Pac2002 Tire damping ($\theta_3(2)$)	500 to 7000 N·s/m
	Suspension stiffness ($\theta_3(3)$)	10000 to 180000 N/m
	Suspension damping ($\theta_3(4)$)	500 to 30000 N·s/m
Assembly 4 (Solid axle suspension and Fiala tire)	Fiala Tire stiffness parameter ($\theta_4(1)$)	400 to 8000 N/m
	Fiala Tire damping parameter ($\theta_4(2)$)	3000 to 180000 N·s/m
	Suspension stiffness ($\theta_4(3)$)	10000 to 180000 N/m
	Suspension damping ($\theta_4(4)$)	500 to 30000 N·s/m

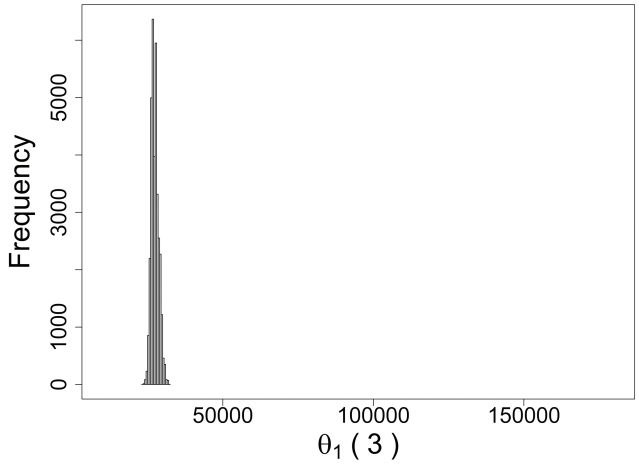
the parameters' posterior distributions. This whole process of optimization, boundary evaluation and the Griddy Gibbs calibration was repeated for all four assemblies. The posterior distributions of the four uncertain parameters and the precision parameter of assembly M_1 are shown in Figure 6.



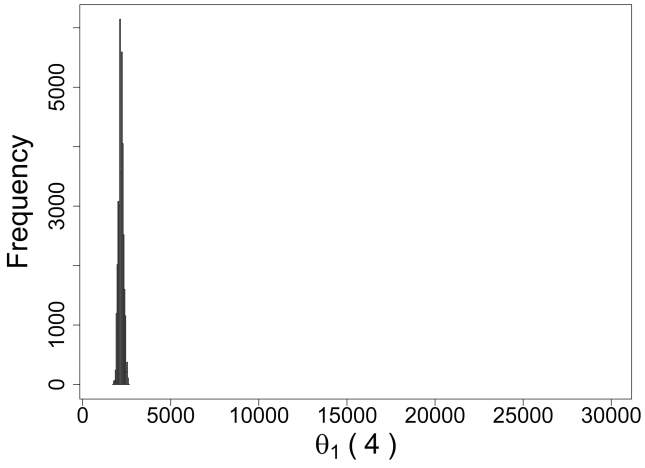
(a) Posterior distribution for $\theta_1(1)$



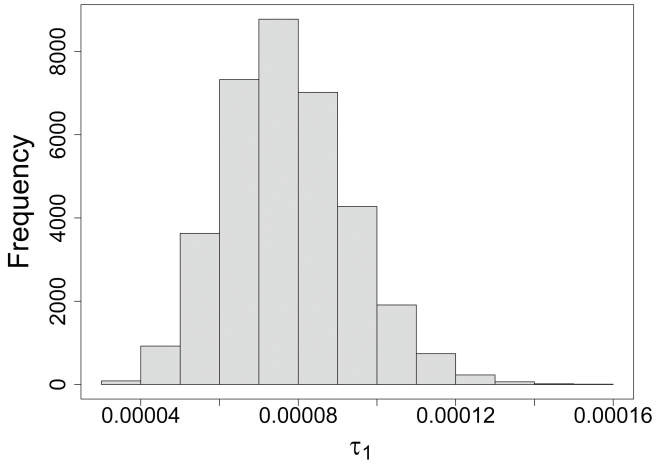
(b) Posterior distribution for $\theta_1(2)$



(c) Posterior distribution for $\theta_1(3)$



(d) Posterior distribution for $\theta_1(4)$



(e) Posterior distribution for τ_1

Fig. 6 Posterior distributions for Assembly 1

Simulated Field Test Data for Classification

For Bayesian classification, a set of simulated field data was generated from the high-fidelity Digital Twin simulation under the same true parameter settings used for the generation of simulated laboratory data. The high-fidelity Digital Twin was run using the same longitudinal velocity. However, instead of a speed bump road profile, stochastic road profiles with surface roughness $G_d = 1024 * 10^{-6} m^3/cycles$ conforming to class D road according to ISO 8608 [15] were used. An example of such road profile is shown in Figure 7.

500 such road profiles were generated, and high-fidelity Digital Twin simulations were run on each of those road profiles. The simulation time was 15 seconds, with the first 6 seconds under smooth road conditions and then rough road (Class D) from 6 to 15 seconds. The normal force time history between 10 and 15 seconds, sampled at 0.001 seconds, was taken. The Gaussian noise with the same mean and variance as for the laboratory data was sampled and added to the model outputs to obtain 500 sets of simulated field data.

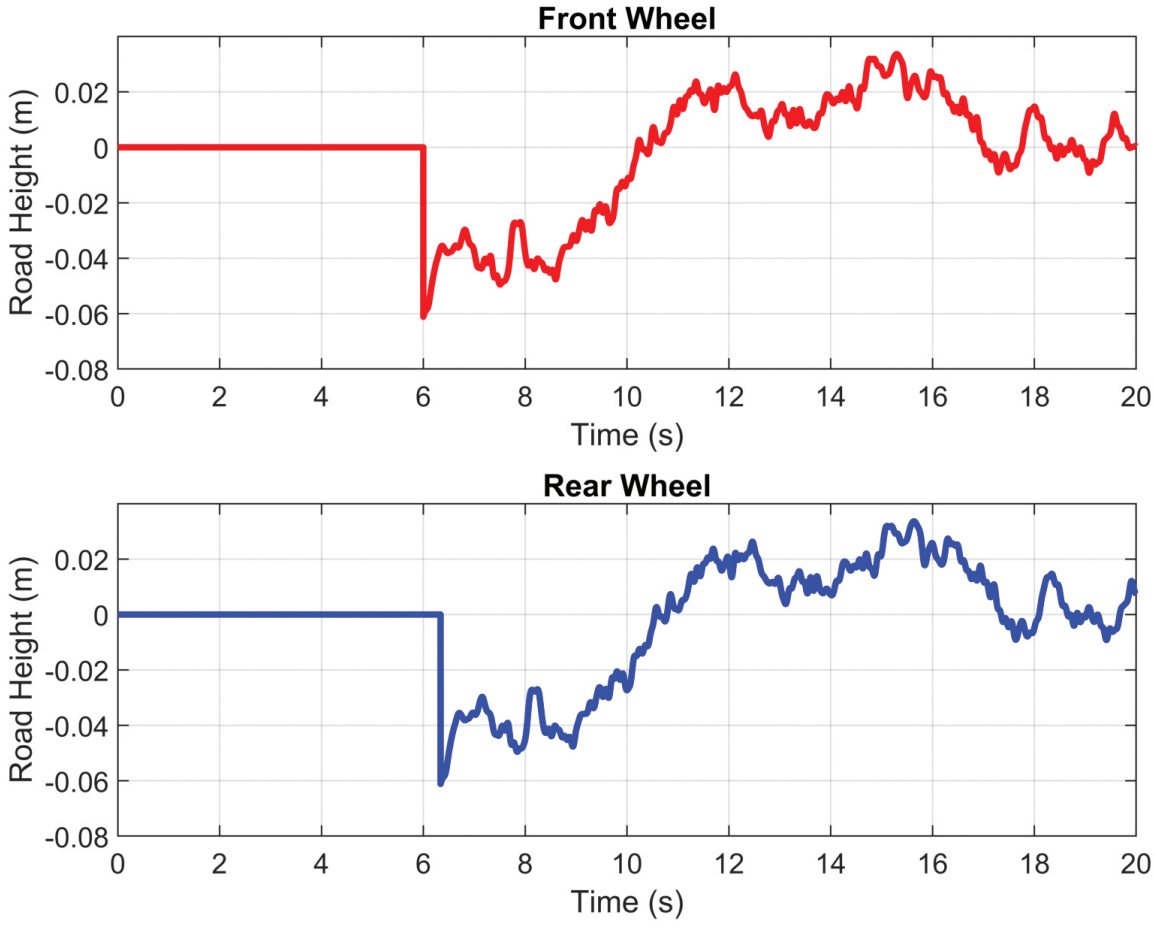


Fig. 7 Example of Stochastic Road Profile

Synthetic Field Data Generation

From the 35000 samples in the posterior distributions of each assembly, 10000 random samples were taken (10000 samples were found to be sufficient to represent the respective posterior distributions). A library of 50000 ISO 8608 class D road profiles was created. For each of the 10000 samples from the posterior distributions of each assembly, the Digital Twin simulation was run under 5 randomly selected stochastic road profiles, generating a total of 50000 model outputs from each assembly. The precision parameters for each assembly were sampled from their posterior distribution, and the error was sampled from a normal distribution with 0 mean and variance equal to the inverse of the precision sampled. The errors were

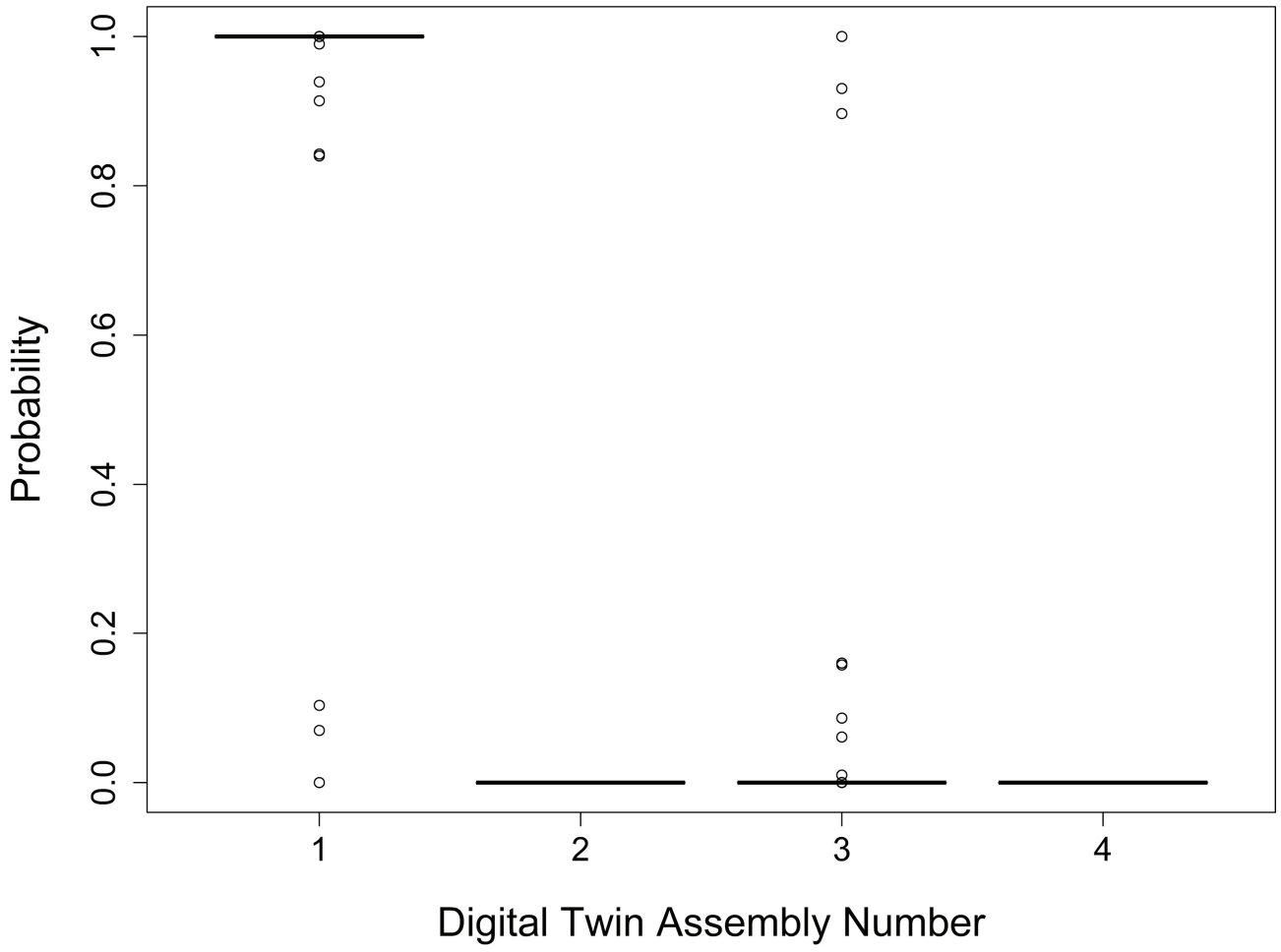


Fig. 8 Boxplot of the probability of selecting each assembly

added to each of those 50000 model outputs from the corresponding assembly to obtain the synthetic field data. The data were then fitted to a multivariate normal distribution. This yielded a unique mean and covariance corresponding to each of the assemblies.

Bayesian Classification

Finally, the probability of observing each of the 500 sets of simulated field data, given that the data came from each of the assemblies (in other words, the probability of observing the data given the mean and covariance structure of each assembly) were evaluated. It was observed that the probability of selecting the true assembly (assembly M_1) is highest for most of those 500 runs, as shown in the box and whisker plot 8 below. This shows that the Bayesian assembly tool can select the correct model with a very high probability.

Conclusion

A mixture-based Bayesian Assembly Tool is proposed that selects the set of Digital Twin subsystems with different levels of fidelities that best fit observed data. First, laboratory test data is acquired either from a laboratory test or from a simulation using a higher fidelity Digital Twin assembly. The Digital Twin assemblies with combinations of different fidelity level subsystem models are created and calibrated to the experimental data. The Griddy Gibbs calibration methodology based on the Bayesian framework is adopted. The result of calibration is posterior distributions of the uncertain parameters and the precision parameter associated with the model error. These posterior distributions are used to inform the Bayesian As-

sembly Tool on the probability distributions of the uncertain parameters of the Digital Twin assemblies. Finally, based on the likelihood of observing the field test data given those distributions, the best-fit model is selected. The application of the Bayesian Assembly Tool is validated using four sets of Digital Twin assemblies with a combination of two tire models, the Fiala model and the PAC2002 model, and two suspension models, MacPherson suspension and solid axle suspension. The Bayesian Assembly Tool was able to select the correct Digital Twin assembly with high accuracy.

Acknowledgments This work was supported by Clemson University's Virtual Prototyping of Autonomy Enabled Ground Systems (VIPR-GS) under Cooperative Agreement W56HZV-21-2-0001 with the US Army DEVCOM Ground Vehicle Systems Center (GVSC).

DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited. OPSEC10040

References

1. Sara Blasco Román and Till Böttjer. "A Methodological Approach to Prioritize Digital Twin Development in Manufac- turing". In: *Applied Stochastic Models in Business and Industry* 41.3 (2025), e2889. DOI: https://doi.org/10.1002/asmb.2889. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asmb.2889. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/asmb.2889.
2. National Academy of Engineering, Engineering National Academies of Sciences, and Medicine. *Foundational Research Gaps and Future Directions for Digital Twins*. Washington, DC: The National Academies Press, 2024. ISBN: 978-0-309-70042-9. DOI: 10.17226/26894. URL: https://nap.nationalacademies.org/catalog/26894/foundational-research-gaps-and-future-directions-for-digital-twins.
3. Michael Grieves and John Vickers. "Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems". In: *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*. Ed. by Franz-Josef Kahlen, Shannon Flumerfelt, and Anabela Alves. Cham: Springer International Publishing, 2017, pp. 85–113. ISBN: 978-3-319-38756-7. DOI: 10.1007/978-3-319-38756-7_4. URL: https://doi.org/10.1007/978-3-319-38756-7_4.
4. The MathWorks Inc. *Combined Slip Wheel 2DOF*. URL: https://www.mathworks.com/help/vdynblks/ref/combinedslipwheel2dof.html.
5. The MathWorks Inc. *Fiala Wheel 2DOF*. URL: https://www.mathworks.com/help/vdynblks/ref/fialawheel2dof.html.
6. Stephen Jay Wright et al. "An Efficient Statistical Inference Approach for Model Calibration Using Griddy Gibbs Sam- pling". In: *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg* 11.3 (2025).
7. M.G. Kapteyn et al. "Data-driven physics-based digital twins via a library of component-based reduced-order models". In: *International Journal for Numerical Methods in Engineering* 123.13 (2022), pp. 2986–3003. DOI: https://doi.org/10.1002/nme.6423. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.6423. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.6423.
8. Michael G. Kapteyn, David J. Knezevic, and Karen Willcox. "Toward predictive digital twins via component-based reduced-order models and interpretable machine learning". In: *AIAA Scitech 2020 Forum*. DOI: 10.2514/6.2020- 0418. eprint: https://arc.aiaa.org/doi/pdf/10.2514/6.2020-0418. URL: https://arc.aiaa.org/doi/abs/10.2514/6.2020-0418.
9. Guesuk Lee et al. "Review of statistical model calibration and validation—from the perspective of uncertainty structures". In: *Structural and Multidisciplinary Optimization* 60 (Oct. 2019). DOI: 10.1007/s00158-019-02270-2.
10. MathWorks. *fmincon*. Accessed: 2025-03-22. n.d. URL: https://www.mathworks.com/help/optim/ug/fmincon.html.
11. MathWorks. *Independent Suspension - MacPherson*. Accessed: 2025-03-20. n.d. URL: https://www.mathworks.com/help/vdynblks/ref/independentsuspensionmacpherson.html.
12. MathWorks. *Solid Axle Suspension - Coil Spring*. n.d.
13. MathWorks. *What is a Digital Twin?* Accessed: 2025-03-20. n.d. URL: https://www.mathworks.com/discovery/digital-twin.html.
14. Johannes Mayer et al. "A decision-making methodology for selecting digital twin applications in the product service phase considering value and effort". In: *Production Engineering* (May 2025). DOI: 10.1007/s11740-025-01350-1.
15. Internationale Organisation für Normung. *ISO 8608 : Mechanical Vibration, Road Surface Profiles, Reporting of Measured Data: Inte* International Organization for Standardization, ISO, 1995. URL: https://books.google.com/books?id=1kyioAEACAAJ.
16. Hans B. Pacejka. *Tire and Vehicle Dynamics*. 3rd. Oxford: Elsevier, 2012. ISBN: 978-0123820301.
17. T.G. Ritto, S. Beregi, and D.A.W. Barton. "Reinforcement learning and approximate Bayesian com- putation for model selection and parameter calibration applied to a nonlinear dynamical system". In: *Mechanical Systems and Signal Processing* 181 (2022), p. 109485. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2022.109485. URL: https://www.sciencedirect.com/science/article/pii/S0888327022005970.
18. Simon Rouchier, Mickaël Rabouille, and Pierre Oberlé. "Calibration of simplified building energy models for param- eter estimation and forecasting: Stochastic versus deterministic modelling". In: *Building and Environment* 134 (2018), pp. 181–190. ISSN: 0360-1323. DOI: https://doi.org/10.1016/j.buildenv.2018.02.043. URL: https://www.sciencedirect.com/science/article/pii/S036013231830115X.
19. Concetta Semeraro et al. "Digital twin paradigm: A systematic literature review". In: *Computers in Industry* 130 (2021), p. 103469. ISSN: 0166-3615. DOI: https://doi.org/10.1016/j.compind.2021.103469. URL: https://www.sciencedirect.com/science/article/pii/S0166361521000762.
20. Angira Sharma et al. "Digital Twins: State of the art theory and practice, challenges, and open research questions". In: *Journal of Industrial Information Integration* 30 (2022), p. 100383. ISSN: 2452-414X. DOI: https://doi.org/10.1016/j.jii.2022.100383. URL: https://www.sciencedirect.com/science/article/pii/S2452414X22000516.
21. Sumit Singh et al. "Challenges of Digital Twin in High Value Manufacturing". In: Oct. 2018. DOI: 10.4271/2018-01- 1928.
22. Hannah Stewart. "Creation and Validation of a Novel Bayesian Calibration Method with Griddy Gibbs Sampling". All Theses. Paper 4177. Master's thesis. Clemson University, 2023. URL: https://open.clemson.edu/all_theses/4177.

23. Thomas Titscher et al. “Bayesian model calibration and damage detection for a digital twin of a bridge demonstrator”. In: *Engineering Reports* 5.11 (2023), e12669. DOI: https://doi.org/10.1002/eng2.12669. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/eng2.12669. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/eng2.12669.
24. Rui Tuo and C. F. Jeff Wu. *A theoretical framework for calibration in computer models: parametrization, estimation and convergence* 2015. arXiv: 1508.07155 [stat.ME]. URL: https://arxiv.org/abs/1508.07155.
25. Seongjin Yun, Jun-Hong Park, and Won-Tae Kim. “Data-centric middleware based digital twin platform for dependable cyber-physical systems”. In: *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*. 2017, pp. 922–926. DOI: 10.1109/ICUFN.2017.7993933.
26. Lin Zhang, Longfei Zhou, and Berthold K.P. Horn. “Building a right digital twin with model engineering”. In: *Journal of Manufacturing Systems* 59 (2021), pp. 151–164. ISSN: 0278-6125. DOI: https://doi.org/10.1016/j.jmsy.2021.02.009. URL: https://www.sciencedirect.com/science/article/pii/S0278612521000455.

Appendix

Table 3 Description of Variables Used in the Equation (3)

Variable	Description
q_{v2}	Tire stiffness variation coefficient with speed
Ω	Wheel rotational velocity
R_0	Unloaded radius
Vo	Measurement speed at test bench
q_{FCX}	Load response to longitudinal force
F_x	Longitudinal force
F_{z0}	Normal wheel load
q_{Fcy}	Load response to lateral force
F_y	Lateral force
q_{Fz1}	Tire vertical stiffness coefficient (linear)
q_{Fz3}	Camber dependency of the tire vertical stiffness
v	Camber (inclination) angle
ρ_z	Tire vertical deflection
q_{Fz2}	Tire vertical stiffness coefficient (quadratic)
p_{pFz1}	Factor to include effect of tire pressure
d_{pi}	Normalized inflation pressure
λ_{cz}	Scale factor for vertical tire stiffness
K_z	Tire vertical damping
qreo	Correction factor for measured unloaded radius
q_{Fcx1}	Tire stiffness interaction with F_x
q_{FCY1}	Tire stiffness interaction with camber

Table 4 Description of Variables Used in the Equation (4)

Variable	Description
$F_{wz_{a,t}}$	Suspension force applied to the wheel on axle a, wheel t along wheel-fixed z-axis.
$F_{z\mathrm{O}_a}$	Vertical suspension spring preload force applied to the wheels on axle a.
k_{z_a}	Vertical spring constant applied to wheels on axle a.
m_{hsteer_a}	Steering angle to vertical force slope applied at wheel carrier for wheels on axle a.
$\delta_{\text{steer}_{a,t}}$	Steering angle input for axle a, wheel t.
C	Vertical damping constant applied to wheels on axle a.
$F_{\text{zhstop}_{a,t}}$	Vertical hardstop force at axle a, wheel t, along the inertial-fixed z-axis.
$F_{zaswya,t}$	Vertical anti-sway force at axle a, wheel t, along the inertial-fixed z-axis.
$z_{v_{a,t}}, z\dot{}v_{a,t}$	Vehicle displacement and velocity at axle a, wheel t, along the inertial-fixed z-axis.
$z_{w_{a,t}}, z\dot{}w_{a,t}$	Wheel displacement and velocity at axle a, wheel t, along the inertial-fixed z-axis.

Table 5 Description of Variables Used in the Equation (5)

Variable	Description
$F_{wz_{a,t}}$	Suspension force applied to the wheel on axle a, wheel t along the wheel-fixed z-axis
$F_{wa_{z0}}$	Wheel and axle interface compliance constant
k_{wa_z}	Wheel and axle interface compliance constant
$z_{w_{a,t}}$	Wheel displacement at axle a, wheel t along the inertial-fixed z-axis
$z_{s_{a,t}}$	Suspension displacement at axle a, wheel t along the inertial-fixed z-axis
$\boldsymbol{C_{wa_z}}$	Wheel and axle interface damping constant
$z'w_{a,t}$	Wheel velocity at axle a, wheel t along the inertial-fixed z-axis
${z^*}_{s_{a,t}}$	Suspension velocity at axle a, wheel t along the inertial-fixed z-axis

Table 6 Non-default parameter values of Fiala Tire model

Parameter	Label	Unit	Value
Brake type	-	-	Disc
Rolling Resistance	-	-	Magic Formula
Vertical Motion	-	-	Mapped stiffness and damping
Kinematic friction	muMin	-	.5
Static friction	muMax	-	.8
Wheel mass	MASS	kg	29.46
Maximum pressure	PRESMAX	Pa	1000000
Minimum pressure	PRESMIN	Pa	10000
Max allowable slip ratio (absolute	KPUMAX	-	1.5
Minimum allowable slip ratio (absolute)	KPUMIN	-	-1.5
Max allowable slip angle (absolute)	ALPMAX	rad	1.5
Minimum allowable slip angle (absolute)	ALPMIN	rad	-1.5
Maximum allowable camber angle	CAMMAX	rad	0.2
Minimum allowable camber angle	CAMMIN	rad	-0.2

Table 7 Non-default parameter values of Pac2002 Tire model

Parameter	Label	Unit	Value
Brake type	-	-	Disc
Maximum pressure	PRESMAX	Pa	1000000
Minimum pressure	PRESMIN	Pa	10000
Max allowable slip angle (absolute)	ALPMAX	rad	1.5
Minimum allowable slip angle (absolute)	ALPMIN	rad	-1.5
Maximum allowable camber angle	CAMMAX	rad	0.2
Minimum allowable camber angle	CAMMIN	rad	-0.2
Nominal pressure	NOMPRES	Pa	220000
Nominal normal force	FNOMIN	N	4000
Wheel mass	MASS	kg	29.46
Longitudinal stiffness	LONGITUDINAL_STIFFNESS	N/m	$1e7$
Lateral stiffness	LATERAL_STIFFNESS	N/m	$4.5e4$

Table 8 Non-default parameter values of MacPherson Suspension Model

Parameter	Label	Unit	Value
Toe angle at steering center	Toe	rad	0
Roll steer angle vs suspension height slope	RollStrgSlp	rad/m	0
Toe angle vs steering angle slope	ToeStrgSlp	-	0
Caster angle at steering center	Caster	rad	0
Caster angle vs suspension height slope	CasterHslp	rad/m	0
Caster angle vs steering angle slope	CasterStrgSlp	-	0
Camber angle at steering center	Camber	rad	0
Camber angle vs suspension height slope	CamberHslp	rad/m	0
Camber angle vs steering angle slope	CamberStrgSlp	-	0

Table 9 Non-default parameter values of Solid Axle Suspension model

Parameter	Label	Unit	Value
Axle and wheels lumped principal moment of inertia about longitudinal axis	AxlIxx	$kg * m^2$	62.37
Axle and wheels lumped mass	AxlM	kg	29.46
Wheel and axle interface compliance constant	KzWhlAxl	N/m	138031.61
Wheel and axle interface compliance preload	F0zWhlAxl	N	2907
Wheel and axle interface damping constant	CzWhlAxl	Ns/m	11490.97
Suspension spring preload	F0z	N	2907
Suspension maximum height	Hmax	m	0.25
Toe angle at steering center	Toe	rad	0
Roll steer angle vs suspension height slope	RollStrgSlp	rad/m	0
Toe angle vs steering angle slope	ToeStrgSlp	-	0
Caster angle at steering center	Caster	rad	0
Caster angle vs suspension height slope	CasterHslp	rad/m	0
Caster angle vs steering angle slope	CasterStrgSlp	-	0

Chapter 2
Bayesian Finite Element Model Updating in the Modal Domain via Integration of Normalizing Flows with Markov Chain Monte Carlo

Lin Sun, Zhen Hu, and Michael D. Todd

Abstract Bayesian finite element (FE) model updating is a powerful framework for calibrating FE models by quantifying uncertainties in model parameters. Markov Chain Monte Carlo (MCMC) methods are commonly used to draw samples from the posterior distribution of these parameters. MCMC operates by accepting or rejecting proposed samples generated from a transition kernel that satisfies the detailed balance condition, ensuring convergence to the target distribution. However, conventional MCMC methods face several limitations: (1) they often require a large number of iterations to reach high-probability regions, (2) they may struggle to explore all modes in a multi-modal posterior distribution, and (3) they incur significant computational cost when evaluating the likelihood function for large-scale and/or nonlinear FE models. To address these challenges, the MCMC framework can be enhanced by integrating a machine learning technique known as normalizing flows (NFs). The resulting algorithm runs multiple Markov chains in parallel, with proposals alternately drawn from a standard Gaussian distribution and a learned NF-based distribution. We demonstrate the proposed method on a multi-degree-of-freedom (DOF) dynamic system, using modal properties (natural frequencies and mode shapes). Results indicate that the NF-enhanced MCMC approach achieves fast convergence and improved sampling accuracy. This study introduces a novel and effective strategy for improving the efficiency and robustness of Bayesian FE model updating through machine learning.

Keywords Bayesian finite element model updating · Normalizing flow · Markov Chain Monte Carlo

List of Notations and Abbreviations

Finite element: FE
Bayesian finite element model updating: BFEMU
Markov Chain Monte Carlo: MCMC
Normalizing flow: NF
Degree of freedom: DOF
Metropolis-Hastings: MH
Metropolis-adjusted Langevin algorithm: MALA
Kullback–Leibler: DL
D: measured response data
θ: unknown parameter vector
λ: eigenvalue
$\boldsymbol{\Phi}$: eigenvector / mode shape

Lin Sun · Michael D. Todd
Department of Structural Engineering, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
e-mail: lsun@ucsd.edu; mdtodd@ucsd.edu

Zhen Hu
Department of Industrial and Manufacturing Systems Engineering,
University of Michigan-Dearborn, 4901 Evergreen Road, Dearborn, MI 48128, USA
e-mail: zhennhu@umich.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

σ: standard deviation
N_m: number of modes
N_D: number of datasets
N_{chain}: number of Markov chains
N_{iter}: number of iterations
f: natural frequency
∂: scaling factor
$p_{base}(\boldsymbol{\alpha})$: base distribution
$F(\cdot)$: one-to-one mapping
$p(\boldsymbol{\beta})$: target distribution
k_{local}: switching step
$s(\cdot)$: scale function
$t(\cdot)$: translation function
$\boldsymbol{\alpha}(\cdot)$: acceptance probability
$L(\cdot)$: loss function
$\mathbf{w}$: weights
$\mathbf{b}$: bias
ε: learning rate

Introduction

The finite element (FE) method is a cornerstone of structural engineering, widely used to simulate the behavior of complex systems such as buildings, bridges, and mechanical assemblies [1, 2]. However, due to modeling assumptions, parameter uncertainties, and simplifications, discrepancies often arise between FE predictions and experimental or field measurements. Bayesian finite element model updating (BFEMU) has emerged as a powerful framework to calibrate model parameters and quantify uncertainty, providing more reliable and interpretable models [3].

Traditional BFEMU algorithms, such as Markov Chain Monte Carlo (MCMC) [4, 5], are effective but face notable limitations. When applied to large-scale and/or nonlinear models, they incur high computational costs due to repeated FE model evaluations. Additionally, MCMC samplers may require excessive iterations to reach high-probability regions—particularly when the prior is distant from the posterior—and often struggle with multi-modal posteriors. To address these challenges, researchers have begun integrating machine learning techniques into inference frameworks. One promising approach is normalizing flows (NFs), a class of invertible neural networks capable of learning complex probability distributions [6, 7]. While NF has shown success in general-purpose Bayesian inference, its application to BFEMU remains limited.

This study applies an NF-enhanced MCMC sampler to BFEMU in the modal domain. The method is evaluated on a simulated four-degree-of-freedom shear building subjected to white noise excitation. Modal properties—such as natural frequencies and mode shapes—can be obtained either by perturbing the "true" model parameters or through linear system identification methods [8, 9]. This work contributes a novel integration of machine learning and Bayesian inference for structural model calibration, offering a scalable and accurate alternative to conventional methods.

Bayesian Finite Element Model Updating in Modal Domain

BFEMU is commonly performed using Bayes' theorem:

$$p(\boldsymbol{\theta}|\mathbf{D}) = \frac{p(\mathbf{D}|\boldsymbol{\theta}) \times p(\boldsymbol{\theta})}{p(\mathbf{D})} \sim p(\mathbf{D}|\boldsymbol{\theta}) \times p(\boldsymbol{\theta}) \tag{1}$$

where $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$ denotes the set of selected FE model parameters, and $\mathbf{D} \in \mathbb{R}^{n_D}$ represents the measured response data. In this expression, $p(\boldsymbol{\theta})$ is the prior distribution of the parameters, $p(\mathbf{D}|\boldsymbol{\theta})$ is the likelihood function, and $p(\mathbf{D})$ is the evidence term. The posterior distribution $p(\boldsymbol{\theta}|\mathbf{D})$ characterizes the updated belief about the model parameters after observing data $\mathbf{D}$. Since the evidence $p(\mathbf{D})$ is a normalization constant, the posterior is proportional to the product of the likelihood and the prior.

In this study, the data $\mathbf{D}$ used for BFEMU consists of modal properties, including natural frequencies and mode shapes. These modal parameters are typically obtained through system identification techniques. Assuming statistical independence

across N_D datasets and N_M modes, the likelihood function can be expressed as [10]:

$$p\left(\mathbf{D}_{1:N_D}|\boldsymbol{\theta}\right) = \prod_{i=1}^{N_D}\prod_{j=1}^{N_M}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{e_{i,j,\lambda}^2}{2\sigma_{i,j,\lambda}^2} - \frac{\mathbf{e}_{i,j,\boldsymbol{\Phi}}^T\mathbf{e}_{i,j,\boldsymbol{\Phi}}}{2\sigma_{i,j,\boldsymbol{\Phi}}^2}\right\}, \tag{2}$$

where λ represents the eigenvalue (related to the natural frequency$f$ via $\lambda = (2\pi f)^2$), and $\boldsymbol{\Phi}$ denotes the eigenvector/mode shape. The terms $\sigma_{i,j,\lambda}$ and $\sigma_{i,j,\boldsymbol{\Phi}}$ denote the standard deviations associated with the eigenvalue and eigenvector errors, respectively. The eigenvalue error is defined as:

$$e_{i,j,\lambda} = \lambda_{i,j}^{\exp} - \lambda_j\left(\boldsymbol{\theta}\right), \tag{3}$$

where $\lambda_{i,j}^{\exp}$ is the experimentally estimated eigenvalue for the jth mode from the ith dataset, and $\lambda_j\left(\boldsymbol{\theta}\right)$ is the corresponding FE-predicted eigenvalue. The eigenvector error term defined as:

$$\boldsymbol{e}_{i,j,\boldsymbol{\Phi}} = \frac{\boldsymbol{\Phi}_{i,j}^{\exp}}{\left\|\boldsymbol{\Phi}_{i,j}^{\exp}\right\|} - \partial_j\frac{\boldsymbol{\Phi}_j\left(\boldsymbol{\theta}\right)}{\left\|\boldsymbol{\Phi}_j\left(\boldsymbol{\theta}\right)\right\|}, \tag{4}$$

where $\|\cdot\|$ denotes the L_2 norm. Here, $\boldsymbol{\Phi}_{i,j}^{\exp}$ is the experimental mode shape for the ith dataset and j^{th} mode, and $\boldsymbol{\Phi}_j\left(\boldsymbol{\theta}\right)$ is the corresponding FE-predicted mode shape. To eliminate the effect of arbitrary scaling in the mode shapes, a scaling factor ∂_j is applied and defined as:

$$\partial_j = \frac{\boldsymbol{\Phi}_{i,j}^{\exp}\cdot\boldsymbol{\Phi}_j\left(\boldsymbol{\theta}\right)}{\left\|\boldsymbol{\Phi}_{i,j}^{\exp}\right\|\left\|\boldsymbol{\Phi}_j\left(\boldsymbol{\theta}\right)\right\|}. \tag{5}$$

This formulation ensures proper comparison between experimental and numerical mode shapes during the updating process.

Integration of Normalizing Flow with Markov Chain Monte Carlo

NFs are a class of generative models that transform a simple base distribution $p_{base}\left(\boldsymbol{\alpha}\right)$ into a complex target distribution $p\left(\boldsymbol{\beta}\right)$ using a sequence of invertible and differentiable mappings. This transformation is modeled as a one-to-one mapping function $\boldsymbol{\beta} = F\left(\boldsymbol{\alpha}\right)$. This formulation enables the use of the change-of-variables formula to compute the target distribution:

$$\hat{p}\left(\boldsymbol{\beta}\right) = p_{base}\left(\hat{F}^{-1}\left(\boldsymbol{\alpha}\right)\right)\left|\det\left(\nabla_{\boldsymbol{\alpha}}\hat{F}^{-1}\left(\boldsymbol{\alpha}\right)\right)\right|, \tag{6}$$

where $\hat{F}\left(\cdot\right)$ is an approximation of the optimal mapping $F^*\left(\cdot\right)$; $\hat{F}^{-1}\left(\cdot\right)$ is its inverse; $\nabla_{\boldsymbol{\alpha}}\hat{F}^{-1}\left(\boldsymbol{\alpha}\right)$ is denotes the Jacobian matrix of the inverse transformation, and $\det\left(\cdot\right)$ represents the determinant operator.

In practice, the design of the mapping $\hat{F}\left(\cdot\right)$ must enable efficient density evaluation, which requires computing Jacobian determinants and inverses in a tractable manner. To achieve this, coupling layers—known for their computational efficiency and structural simplicity—are often employed in NF architectures [6]. A key design objective in NF models is to ensure these computations remain easily tractable. Addressing these challenges has been a central focus in the literature [6].

Coupling layers typically adopt alternating patterns to construct the one-to-one transformation $\hat{F}\left(\cdot\right)$ [7]. Let $\boldsymbol{\alpha}^{(i)} \in \mathbb{R}^D$ and $\boldsymbol{\beta}^{(i)} \in \mathbb{R}^D$ denote the input and output vectors in the ith coupling layer, respectively. The transformation between $\boldsymbol{\alpha}^{(i)}$ and $\boldsymbol{\beta}^{(i)}$ is defined as:

$$\begin{cases} \boldsymbol{\beta}_{1:d}^{(i)} = \boldsymbol{\alpha}_{1:d}^{(i)} \\ \boldsymbol{\beta}_{d+1:D}^{(i)} = \boldsymbol{\alpha}_{d+1:D}^{(i)} \odot \exp\left(s\left(\boldsymbol{\alpha}_{1:d}^{(i)}\right)\right) + t\left(\boldsymbol{\alpha}_{1:d}^{(i)}\right) \end{cases}, \tag{7}$$

where d partitions the input vector into two subsets, $s\left(\cdot\right)$ and $t\left(\cdot\right)$ are scale and translation functions, respectively, and $\odot$ denotes element-wise multiplication. Since the Jacobian matrix of the transformation in Eq. (7) is triangular, its determinant simplifies to:

$$\det\left(\frac{\partial\boldsymbol{\beta}^{(i)}}{\partial\boldsymbol{\alpha}^{(i),T}}\right) = \operatorname{diag}\left(\exp\left(s\left(\boldsymbol{\beta}_{1:d}^{(i)}\right)\right)\right). \tag{8}$$

The inverse transformation is also tractable:

$$\begin{cases} \boldsymbol{\alpha}_{1:d}^{(i)} = \boldsymbol{\beta}_{1:d}^{(i)} \\ \boldsymbol{\alpha}_{d+1:D}^{(i)} = \left(\boldsymbol{\beta}_{d+1:D}^{(i)} - t\left(\boldsymbol{\beta}_{1:d}^{(i)}\right)\right) \odot \exp\left(-s\left(\boldsymbol{\beta}_{1:d}^{(i)}\right)\right) \end{cases} . \tag{9}$$

By stacking n such coupling layers in alternating order, the full transformation can be constructed as (see Figure 1):

$$F = F_1 \circ \cdots \circ F_n. \tag{10}$$

The scale and translation functions in each layer are typically parameterized by neural networks. The model is trained by minimizing the Kullback–Leibler (KL) divergence between the transformed distribution and the target distribution. The objective function is given by:

$$D_{KL}\left(\hat{p}\left(\boldsymbol{\beta}\right)||p_*\left(\boldsymbol{\beta}\right)\right) = \int_{\Omega_{\boldsymbol{\beta}}} \left[-\log\left(p_*\left(\boldsymbol{\beta}\right)\right) + \log\left(\hat{p}\left(\boldsymbol{\beta}\right)\right)\right] \hat{p}\left(\boldsymbol{\beta}\right) d\boldsymbol{\beta} \ . \tag{11}$$

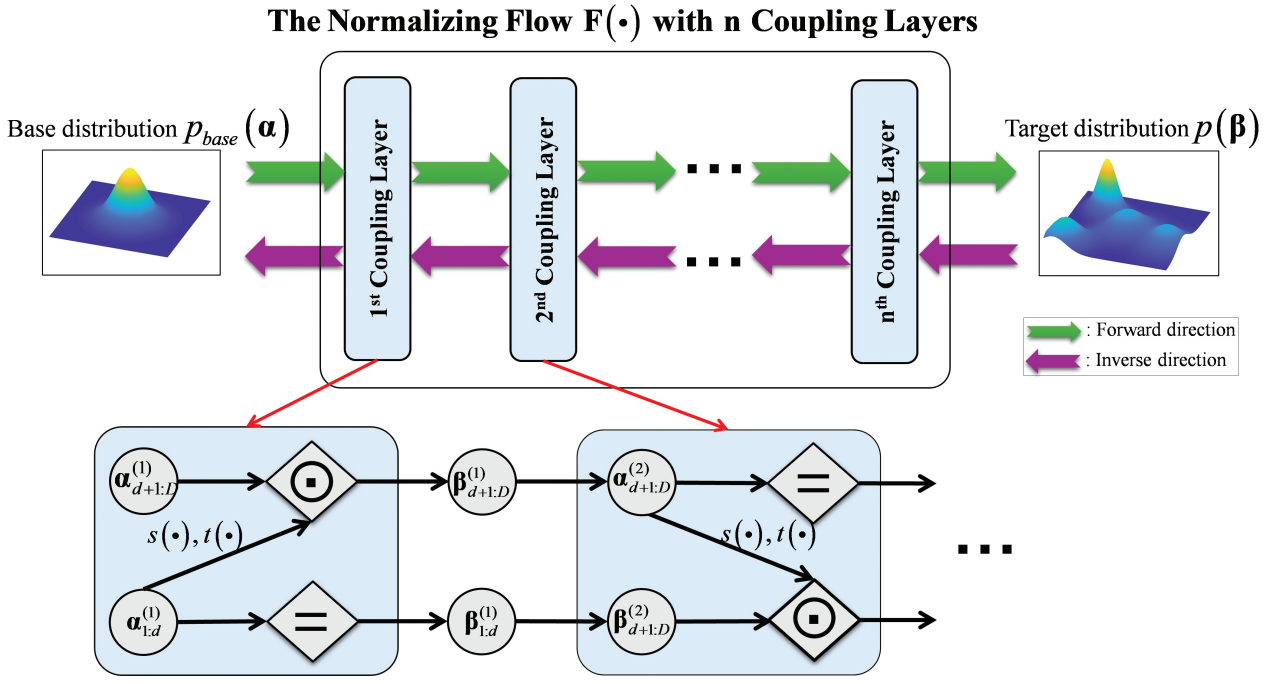


Fig. 1 Normalizing flow model with n alternating coupling layers

MCMC algorithms are widely used to generate samples from a target posterior distribution. In the context of BFEMU, the goal is to sample from $p\left(\boldsymbol{\theta}|\mathbf{D}\right)$. A central challenge in MCMC is the design of effective transition kernels $\pi\left(\boldsymbol{\theta},\ \boldsymbol{\theta}'\right)$ that enable efficient exploration of complex, high-dimensional distributions while maintaining detailed balance. Traditional MCMC methods such as Metropolis–Hastings (MH) [11] and Metropolis-adjusted Langevin algorithm (MALA) [12] are commonly employed. However, these methods can exhibit slow convergence when the posterior has multiple modes or nonlinear correlations. To address these limitations, recent approaches integrate NFs into MCMC frameworks, leveraging their capacity to learn expressive, invertible transformations that map simple base distributions to complex targets.

Gabrie et al. [13] proposed an alternating MCMC scheme that switches between a local Gaussian kernel and a global NF-based kernel. This hybrid design enables the sampler to maintain robustness during local exploration while benefiting from the expressivity of the NF in global moves. Crucially, proposals from the global kernel are independent of previous samples in the chain, leading to improved mixing and reduced autocorrelation.

Let N_{iter} denote the number of MCMC iterations and N_{chain} the number of parallel chains. The algorithm generates samples $\{\boldsymbol{\theta}_{i,j};\ i = 0,\ 1,\ ...,\ N_{chain};\ j = 0, 1, ..., N_{iter}\}$ from the posterior distribution $p\left(\boldsymbol{\theta}|\mathbf{D}\right)$. Each chain is initialized with $\{\boldsymbol{\theta}_{i,0};\ i = 0,\ 1,\ ...,\ N_{chain}\} \sim p\left(\boldsymbol{\theta}\right)$. At each iteration j, a decision is made to use either the global or local kernel based on a predefined schedule (e.g., alternating every k_{local} steps).

If the global kernel is selected, a base sample $\theta'_{base,i} \sim p_{base}(\theta)$ is drawn from a standard normal distribution and transformed via the NF mapping $\hat{F}(\cdot)$ to obtained the proposal θ'_i (see step 2 in Table 1). The proposed sample θ'_i is accepted with the acceptance probability:

$$\alpha_{global}\left(\theta_{i,j},\ \theta'_i\right) = \min\left(1,\ \frac{\hat{p}(\theta_{i,j})\, p\left(\theta'_i|\mathbf{D}\right)}{\hat{p}\left(\theta'_i\right) p(\theta_{i,j}|\mathbf{D})}\right), \tag{12}$$

where $\hat{p}(\theta)$ denotes the probability density estimated by the NF model. Notably, $\hat{p}(\theta)$ does not necessarily match the target distribution $p(\theta|\mathbf{D})$ as the NF model may not be optimal. If the local kernel is selected, the proposed sample θ'_i is drawn from the conditional Gaussian distribution (see step 5 in Table 1), and accepted with the standard MH probability $\alpha_{local}\left(\theta_{i,j},\ \theta'_i\right)$.

If the jth iteration is completed, the NF model is updated by minimizing the negative log-likelihood loss $L(\mathbf{w},\ \mathbf{b})$ over samples, where $(\mathbf{w},\ \mathbf{b})$ are the trainable parameters of the scale and translation networks. The model is updated via gradient descent with learning rate ε. The training process ensures that future proposals from the NF model better approximate the target posterior distribution.

Table 1 General procedure for integrating NF with MCMC

input: $\left\{\boldsymbol{\theta}_{i,0};\ i=0,1,...,N_{chain}\right\},\ \sigma,\ \varepsilon$

while $j < N_{iter}$:

  for $i=1,...,N_{chain}$

    if j mod $k_{local}=0$:

      $1:\boldsymbol{\theta}'_{base,i} \sim p_{base}\left(\boldsymbol{\theta}\right)$

      $2:\boldsymbol{\theta}'_i = \hat{F}\left(\boldsymbol{\theta}'_{base,i}\right)$

      3 : accept: $\boldsymbol{\theta}_{i,j+1}=\boldsymbol{\theta}'_i$ with probability $\alpha_{global}\left(\boldsymbol{\theta}_{i,j},\boldsymbol{\theta}'_i\right)$

      4: otherwise: $\boldsymbol{\theta}_{i,j+1}=\boldsymbol{\theta}_{i,j}$

    else:

      $5:\boldsymbol{\theta}'_i \sim N\left(\boldsymbol{\theta}_{i,j},\sigma^2\mathbf{I}\right)$

      6 : accept: $\boldsymbol{\theta}_{i,j+1}=\boldsymbol{\theta}'_i$ with probability $\alpha_{local}\left(\boldsymbol{\theta}_{i,j},\boldsymbol{\theta}'_i\right)$

      7: otherwise: $\boldsymbol{\theta}_{i,j+1}=\boldsymbol{\theta}_{i,j}$

  8 : evaluate loss function: $L\left(\mathbf{w},\mathbf{b}\right)=-\frac{1}{N_{chain}}\sum_{i=1}^{N_{chain}}\log\left(\hat{p}\left(\boldsymbol{\theta}_{i,j+1}\right)\right)$

  9 : update the mapping: $\left(\mathbf{w},\mathbf{b}\right)=\left(\mathbf{w},\mathbf{b}\right)-\varepsilon\nabla L\left(\mathbf{w},\mathbf{b}\right)$

$j=j+1$

return $F(\cdot)$, $\left\{\boldsymbol{\theta}_{i,j};\ i=0,1,...,N_{chain};\ j=0,1,...,N_{iter}\right\}$

Case Study: Numerically Simulated 3-DOF Dynamic System

Figure 2(a) illustrates a numerically simulated three-degree-of-freedom (3-DOF) dynamic system. The nominal story mass is 625 tons, and the nominal story stiffness is 1×10^9 N/m. In the context of structural health monitoring, story stiffness is of particular interest, as it reflects the condition and potential damage within a building. To account for possible structural changes, the story stiffness is treated as an unknown parameter vector θ:

$$\theta = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}^T. \tag{13}$$

Eigenvalues and mode shapes are extracted via eigen-analysis from the system's stiffness and mass matrices. To generate modal data for BMU, the "true" eigenvalues and mode shapes are computed using the nominal mass and stiffness matrices and a set of "true" parameter values θ^{true}. For numerical stability, each eigenvalue is normalized by the largest eigenvalue (corresponding to the third mode), and each mode shape is normalized by its maximum nodal displacement. To simulate measurement noise, 100 datasets are generated by perturbing the "true" eigenvalues and mode shapes with white noise. The eigenvalues are perturbed with Gaussian noise having a standard deviation equal to 5% of their corresponding "true" values. Similarly, each nodal value in the mode shapes is perturbed with Gaussian noise with a standard deviation of 5% of its respective normalized value (see Figure 2(b) and (c)). The prior distribution for the parameters is assumed to be Gaussian, with a mean equal to 1.5 times the corresponding "true" value, a coefficient of variation of 20%, and no correlation among parameters.

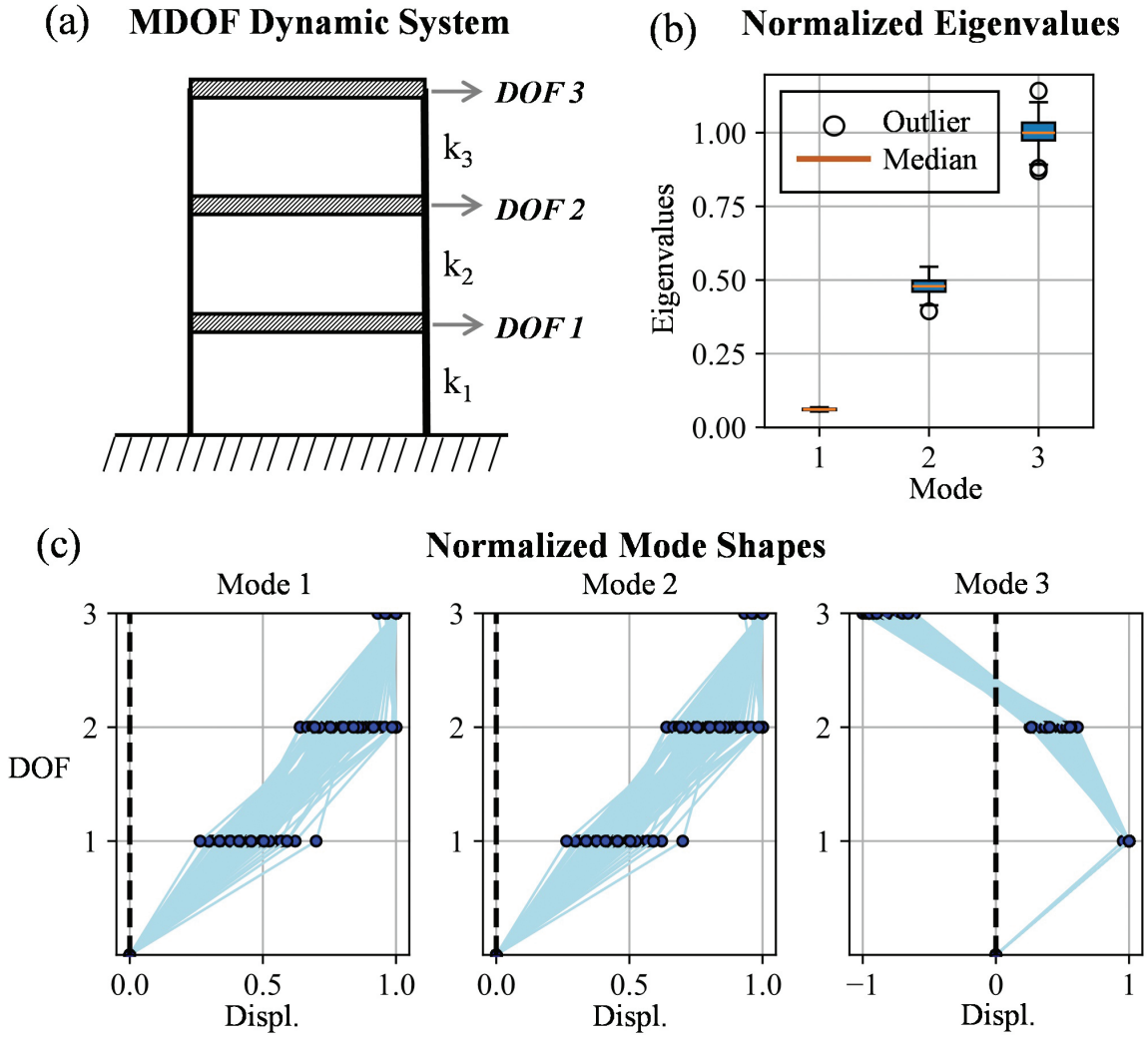


Fig. 2 (a) Schematic of the 3-DOF dynamic system; (b) normalized eigenvalues; (c) normalized mode shapes

We run three Markov chains using the MCMC algorithm integrated with the NF. The initial positions of each Markov chain are drawn from the prior distribution of the parameters. The switching step k_{switch} is set to 2, and the total number of iterations for each chain is set to 4000. This means the transition kernel switches between the local and global kernels at each iteration. Specifically, 2000 iterations use the local kernel, and the other 2000 iterations use the global kernel for each chain.

Figure 3(a) displays the 3D trace plot of the three Markov chains in the parameter space (k_1, k_2, k_3). The red stars represent the initial positions of the chains, while the red dot indicates the "true" parameter values. The plot shows that after a sufficient number of iterations, all three chains converge toward the true parameter values, i.e., the high-probability region.

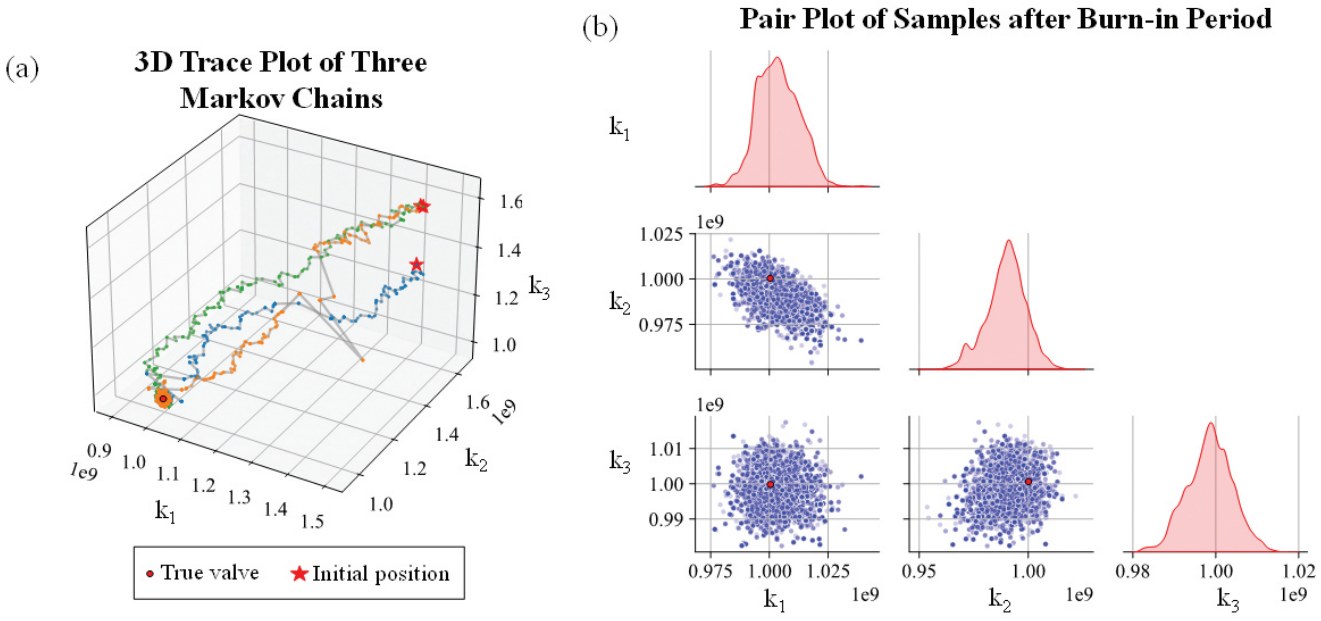


Fig. 3 3D trace plot of the three Markov chains in the parameter space (k_1, k_2, k_3); (b) pair plot of the samples after the burn-in period

After discarding the burn-in period for each chain, the collected samples are shown in the pair plot in Figure 3(b). It is observed that the sample pair (k_1, k_2) exhibits a weak negative correlation, while the pairs (k_1, k_3) and (k_2, k_3) show minimal correlation. Furthermore, the marginal distributions of k_1, k_2 and k_3 indicate that the samples are concentrated near their respective true values.

Conclusions

In this study, we applied a Markov Chain Monte Carlo (MCMC) algorithm integrated with a normalizing flow (NF) to perform Bayesian inference on a 3-degree-of-freedom (3-DOF) dynamic system in modal domain. The three story stiffnesses are elected as parameters to update. By running three independent Markov chains with a total of 4000 iterations each—switching between a local kernel and a global kernel—we observed robust convergence behavior in the parameter space.

The 3D trace plot demonstrated that all chains effectively converged to the high-probability region corresponding to the "true" parameter values. After discarding the burn-in period, the pair plot revealed a slight negative correlation between k_1 and k_2, while (k_1, k_3) and (k_2, k_3) showed minimal correlation. Moreover, the marginal distributions of each parameter were sharply peaked around the true values, confirming the accuracy and efficiency of the inference. These results highlight the effectiveness of the NF-enhanced MCMC framework in capturing complex posterior distributions and ensuring rapid convergence in multi-parameter settings.

Acknowledgments Funding for this work by the U.S. Army Corps of Engineers through the U.S. Army Engineer Research and Development Center Research Cooperative Agreement W9132T-22-20014 is gratefully acknowledged. The opinions and findings in this study are those of the authors and do not necessarily reflect the views of the sponsor.

References

1. K.J. Bathe, Finite Element Procedures, Prentice Hall, 2007.
2. R.D. Cook, D.S. Malkus, M.E. Plesha, R.J. Witt, Concepts and Applications of Finite Element Analysis, 4th ed., Wiley, 2001.
3. T. Marwala, L. Mdlazi, S. Sibisi, "Finite element model updating using Bayesian approach," 2007.
4. W.K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," Biometrika 57 (1970) 97–109. https://doi.org/10.1093/biomet/57.1.97.

5. J. Ching, Y.C. Chen, "Transitional Markov Chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging," J. Eng. Mech. 133 (2007) 816–832. https://doi.org/10.1061/(ASCE)0733-9399(2007)133:7(816).
6. G. Papamakarios, E. Nalisnick, D.J. Rezende, S. Mohamed, B. Lakshminarayanan, "Normalizing flows for probabilistic modeling and inference," 2019. arXiv:1912.02762.
7. L. Dinh, J. Sohl-Dickstein, S. Bengio, "Density estimation using Real NVP," 2016. arXiv:1605.08803.
8. L. Sun, Linear System Identification and Bayesian Finite Element Model Updating of Civil Structural and Substructural Systems, Ph.D. Dissertation, University of California, San Diego, 2024. https://escholarship.org/uc/item/2090j6fz
9. L. Sun, J.P. Conte, M.D. Todd, R. Astroza, Y. Bock, G. Offield, F. Vernon, "Linear system identification of the UC San Diego Geisel Library building under ambient vibration," J. Civ. Struct. Health Monit. (2024). https://doi.org/10.1007/s13349-024-00889-4.
10. M. Song, L. Renson, J. Noël, B. Moaveni, G. Kerschen, "Bayesian model updating of nonlinear systems using nonlinear normal modes," Struct. Control Health Monit. 25 (2018) e2258. https://doi.org/10.1002/stc.2258.
11. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, "Equation of state calculations by fast computing machines," J. Chem. Phys. 21 (1953) 1087–1092. https://doi.org/10.1063/1.1699114.
12. G. Roberts, R. Tweedie, "Exponential convergence of Langevin distributions and their discrete approximations," Bernoulli 2 (1996) 341–363.
13. M. Gabrié, G.M. Rotskoff, E. Vanden-Eijnden, "Adaptive Monte Carlo augmented with normalizing flows," Proc. Natl. Acad. Sci. U.S.A. 119 (2022) e2109420119. https://doi.org/10.1073/pnas.2109420119.

Chapter 3
Correlation & Orthogonality – What Do the Numbers Mean? – Use, Misuse and Confusion

Peter Avitabile

Abstract Experimental modal testing is performed for a variety of reasons – one of which is to validate or compare to a finite element model. Metrics such as the Modal Assurance Criteria and Mass Orthogonality are vector based techniques used to compare an analytically developed mode shape with an experimentally obtained mode shape. These vector based numerical values are then evaluated to some criteria that may be mandated by contract or imposed by some specification or maybe be a company internally agreed upon number. Beyond vector correlation metrics, the shapes can also be compared on a degree of freedom to degree of freedom (point to point) comparison to further interrogate the level of correlation – but often these further checks are not normally performed or mandated as part of the process. Unfortunately, these degree of freedom checks are very important to really understand the vector correlation results and without any assessment, the true correlation is really not clearly understood. This paper will discuss many of these aspects that are very important to the actual correlation achieved. Basic theoretical equations are presented and examples of many important issues are presented. In addition, general commentary is provided because there is a general misunderstanding and misconceptions as to what all the numbers really mean.

Keywords Orthogonality · Model validation · MAC · POC

Introduction

Experimental modal tests are performed for many purposes. Understanding the system frequencies and shapes is of paramount importance in designing better structures to allow for improved performance. But many times, a modal test may be conducted to help verify or update a finite element model. When this is the case, many correlation tools may be deployed to check both mode shapes as well as individual degree of freedom (DOF) correlation. At times, these correlation tools are deployed for the engineer to better understand their models and adjust or improve the model to better reflect the reality of the actual hardware. And at other times, these tools may be mandated by contract or agency responsible for the acceptance of particular hardware or components used as part of a larger assembly.

When correlation is mandated by contract, then an arbitrary value may be set. If the target is met then the structure is considered "validated" – if the target is not met, then the structure is considered to be "not validated". But the specific number is somewhat arbitrary at times and then also, somewhat questionable". The general theory is presented, and in some cases expanded, to clearly understand exactly how the numbers are computed and the actual contribution to the overall calculation to achieve a particular value; this helps in the ability to better understand the values and what they really mean in comparison to the stated "value to meet".

This paper only addresses real normal modes and the understanding of the vector correlation metrics for real normal modes. The inclusion of complex modes only further complicates these issues. Once some of the misunderstandings and general misconceptions are clarified, the next step is to address this with complex modes as part of the mix. But this paper

Peter Avitabile
Structural Dynamics and Acoustic Systems Laboratory, University of Massachusetts Lowell,
One University Avenue, Lowell, Massachusetts 01854
e-mail: peter_avitabile@uml.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

just takes the first step to clarify these points because there seems to be "old wives' tales" that have masked some of the reality of what really should be of concern when performing any type of correlation or orthogonality study.

Theory is summarized first followed by examples to illustrate many misconceptions. Some general observations are made followed by conclusions.

Theory Review

The very basic equations related to MAC, POC, CORTHOG and DCAF are summarized in this section; more detailed theory is contained in the references identified.

Modal Assurance Criteria (MAC)

The MAC [1] is really nothing more than a simple dot product that helps to identify how correlated two vectors are to each other (or not correlated to each other depending on how one looks at it). The MAC can be written as

$$\mathrm{MAC}_{ij} = \frac{\left[\{u_i\}^T \{e_j\}\right]^2}{\left[\{u_i\}^T \{u_i\}\right]\left[\{e_j\}^T \{e_j\}\right]}; \mathrm{MAC}_{ij} = \frac{\left[\{e_i\}^H \{e_j\}\right]^2}{\left[\{e_i\}^H \{e_i\}\right]\left[\{e_j\}^H \{e_j\}\right]} \tag{1}$$

One formulation is used for real normal vectors (when comparing to analytical models) and the other is used for data that is comprised of complex vectors with both magnitude and phase. As the MAC approaches 1, there is a high indication that the two vectors are very correlated. As the MAC approaches 0, there is a high indication that the two vectors are very uncorrelated.

There are many reasons why the MAC may fall less than 1; some typical reasons usually cited are:

- The modal vectors extracted are different from each other
- There may be noise that contaminates the vectors being compared
- The vectors may be affected by nonlinearities in the measured data
- There is an insufficient number of measured points to characterize the shape

For all the cases presented in this paper, the modes shapes are considered real normal modes or proportional modes. A sample MAC matrix is seen in Figure 1 to show a correlation as typically observed comparing an experimental set of modal vectors to an analytical (FEA) set of modal vectors.

Notice that the diagonal has very strong values for the first seven modes indicating a high correlation between each of the mode shape pairs whereas the off-diagonal terms are much smaller approaching zero. Note that the MAC result is often presented in a matrix type display but the calculation is really just individual vector dot products with the results of each dot product arranged in a matrix format.

The biggest strength of the MAC is that the mass matrix is not required. However, that is also the biggest drawback to the MAC because the mass matrix is not used. Mass scaling is critically important to clearly and deeply understand the correlation between an analytical and experimental model. MAC is widely used and an extremely important tool for correlation assessment but does lack the important mass scaling.

Coordinate Modal Assurance Criteria (CoMAC)

The Coordinate Modal Assurance Criteria (CoMAC) [2] is nothing more than the vector comparison but without mass scaling and formulated in a similar fashion as the MAC for correlated mode pairs. Without mass scaling, the CoMAC (like the MAC) is not as robust a correlation tool. Interpretation is similar to MAC. The CoMAC is given as

$$CoMAC(k) = \frac{\left[\sum_{c=1}^{m} \left| u_k^{(c)} \cdot e_k^{(c)} \right|\right]^2}{\sum_{c=1}^{m} \left(u_k^{(c)}\right)^2 \cdot \sum_{c=1}^{m} \left(e_k^{(c)}\right)^2} \tag{2}$$

The interpretation follows in a similar fashion to the MAC.

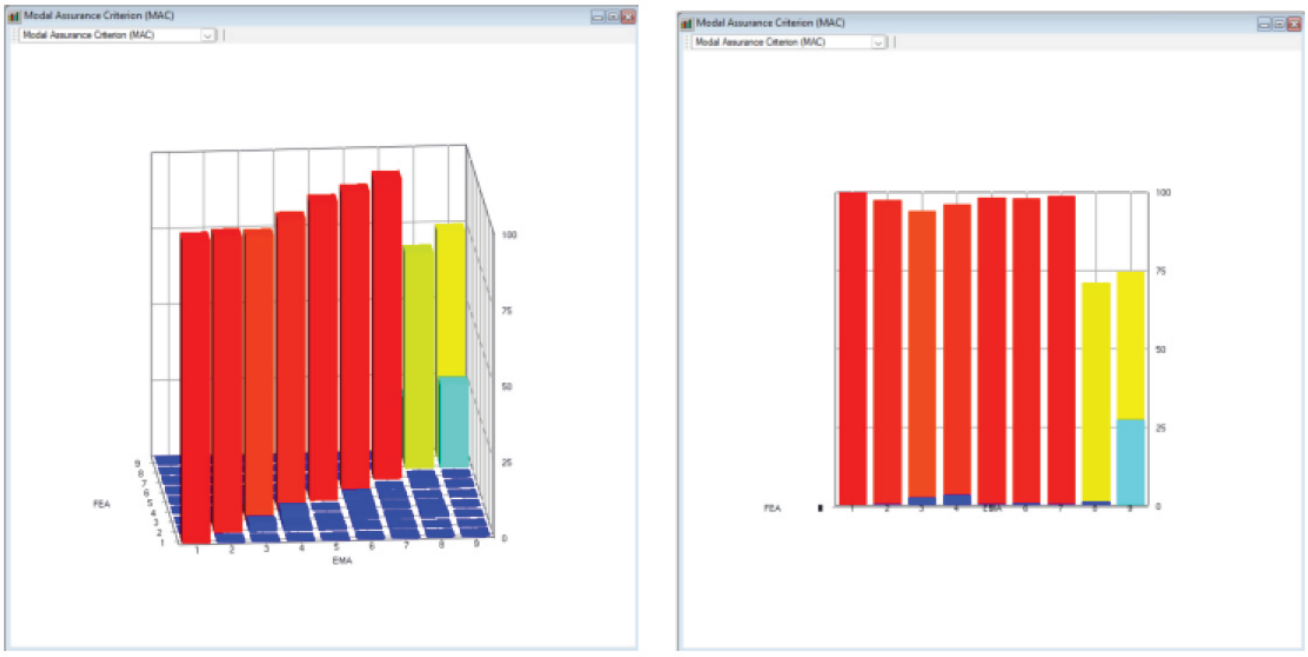


Fig. 1 An Example MAC Matrix – oblique view (left) and side view (right)

Test Response Assurance Criteria (TRAC)

At times, two different time responses may need to be compared. The MAC formulation can also be used for the comparison and is called the Test Response Assurance Criteria (TRAC) (3) and is given as

$$TRAC = \frac{\left[\{x_{n1}\}^T \{x_{n2}\}\right]^2}{\left[\{x_{n1}\}^T \{x_{n1}\}\right]\left[\{x_{n2}\}^T \{x_{n2}\}\right]} \tag{3}$$

The interpretation is the same but one additional item is that the plot can be formed like a regression analysis where a line fit and regression coefficient is shown. This provides significant additional graphical interpretation to the process as shown in Figure 2. Note that the regression coefficient is the calculation value from the TRAC (or MAC).

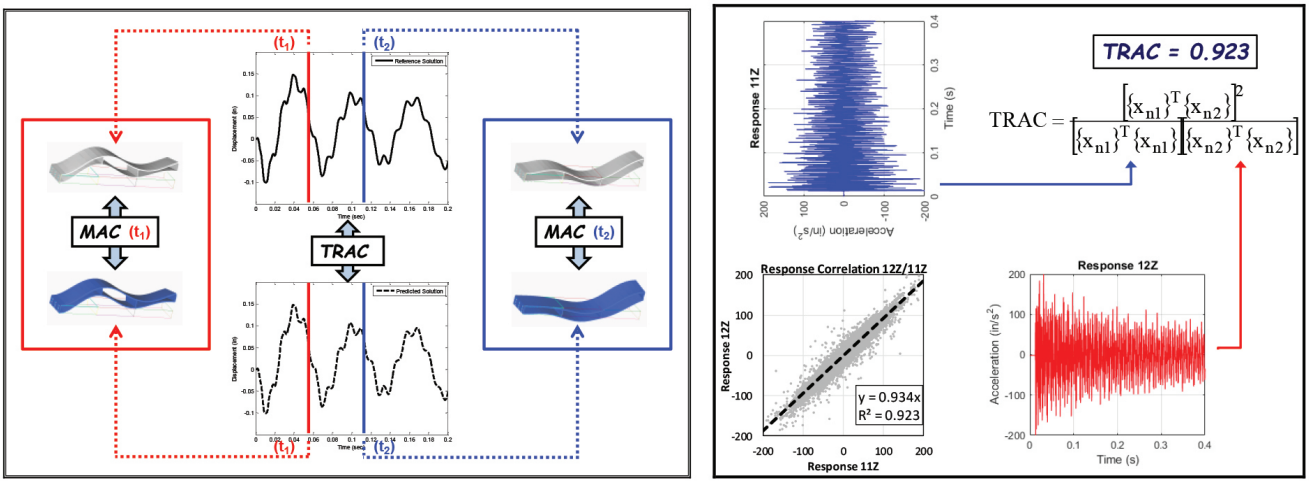


Fig. 2 Illustration of TRAC (with MAC) (left) and TRAC shown as a Regression Analysis (right)

Pseudo Orthogonality Criteria (POC)

The Mass Orthogonality can be stated as

$$\mathrm{ORT} = [\mathrm{U}]^{\mathrm{T}}[\mathrm{M}][\mathrm{U}] = [\mathrm{I}] \tag{4}$$

Now this leads to the Pseudo Orthogonality Criteria (POC) [4] stated as

$$\mathrm{POC} = [\mathrm{E}]^{\mathrm{T}}[\mathrm{M}][\mathrm{U}] \stackrel{?}{=} [\mathrm{I}] \tag{5}$$

This is stated to be "pseudo" or "false" because the definition of the orthogonality can only be with respect to the analytical mass with the analytical vectors.

A typical Pseudo Orthogonality is shown in Fig 3. Note that the Mass Orthogonality (Not shown) would clearly be the identity matrix as expected from the formulation in (4) whereas the POC is a similar result where the diagonals are close

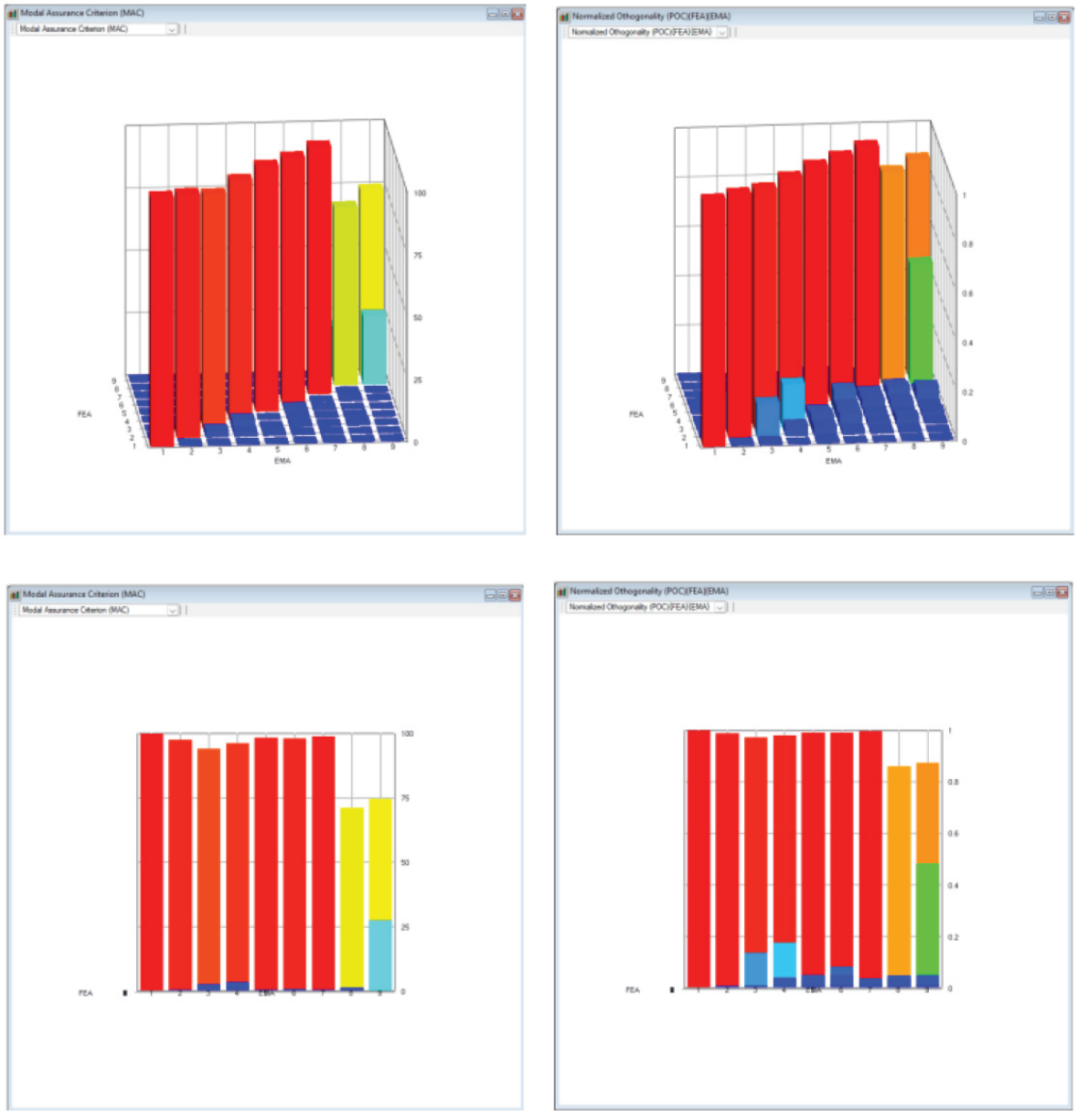


Fig. 3 Typical Comparison of MAC and POC for Two FEA Models (one with 4 times mesh density)
Top Left MAC oblique view – Lower Left MAC side view
Top Right POC oblique view – Lower Right POC side view

to 1.0 for a well correlated model with off-diagonal terms which are very low approaching zero. Figure 3 shows the POC with the MAC for the same models compared. Notice that the MAC values on the diagonal are slightly smaller than the POC values due to the inclusion of mass in the POC calculation; a side view is also shown to more clearly see these differences. And the off-diagonal terms for the POC value are larger than the corresponding MAC values, again due to the inclusion of mass in the POC calculation.

Now all individual terms of the resulting POC do come from a matrix calculation for each of the individual values. But there are many individual terms that are computed and added together to determine each term of the matrix.

Breakdown Orthogonality Calculation

So the Mass Orthogonality can be written out (for a lumped mass matrix) as

$$\mathrm{ORT_{ij}} = (\mathrm{u_{1i}\ m_{11}u_{1j}} + \mathrm{u_{2i}\ m_{22}u_{2j}} + \mathrm{u_{3i}\ m_{33}u_{3j}} + \ldots) = \begin{matrix} 1(\text{ for i} = \text{j}) \\ 0(\text{ for i} \neq \text{j}) \end{matrix} \tag{6}$$

And the POC can also be written in a similar fashion as

$$\mathrm{POC}_{ij} = (e_{1i}m_{11}u_{1j} + e_{2i}m_{22}u_{2j} + e_{3i}m_{33}u_{3j} + \ldots) = \begin{matrix} 1 & (\text{ for } i = j) \\ 0 & (\text{ for } i \neq j) \end{matrix} \tag{7}$$

When written this way, there is more clarity that there are many terms that form the entire resulting POC matrix. This leads the way to the Coordinate Orthogonality Check (CORTHOG) [5–7].

Reduction of Finite Element Mass/Stiffness Matrix and Expansion of Modal Vectors

There are many techniques available to do this. Guyan [8] and Improved Reduced System (IRS) [9] are possible choices. However, the System Equivalent Reduction and Expansion Process (SEREP) [10] has been shown to be the most accurate approach and is used here.

The reduced mass and stiffness transformation matrix (which is also used for expansion) can be formed from the full space eigenvectors obtained from an eigensolution using the finite element model. The transformation matrix [T] is used to project the full mass and stiffness matrices to a smaller size. The reduced matrices can be formulated as

$$[M_a] = [T]^T [M_n] [T] \tag{8}$$

$$[K_a] = [T]^T [K_n] [T] \tag{9}$$

The SEREP transformation is formed as

$$[T_U] = [U_n] [U_a]^g \tag{10}$$

Coordinate Orthogonality Check (CORTHOG)

The Coordinate Orthogonality Check (CORTHOG) [5–7] is simply the comparison of what should have been obtained analytically for each dof in an orthogonality check to what was actually obtained for each dof in a POC from test.

The most important thing to note is that each of the DOF are compared but the calculation includes the mass as a very important scaling component to the entire calculation. This mass scaling proves to be an important consideration that enables the proper scaling for each DOF.

Now if SEREP is used for the formulation of the reduced mass or the expanded vectors, then the orthogonality and pseudo-orthogonality become much simpler expressions. If the SEREP technique is utilized, then the standard mass orthogonality can be written as

$$\mathrm{ORT} = [\mathrm{U_n}]^{\mathrm{T}} [\mathrm{M_n}] [\mathrm{U_n}] = [\mathrm{U_a}]^{\mathrm{T}} [\mathrm{M_a}] [\mathrm{U_a}] = [\mathrm{U_a}]^{\mathrm{g}} [\mathrm{U_a}] \tag{11}$$

and the Pseudo Orthogonality Check can be written as

$$\mathrm{POC} = [\mathrm{U_n}]^{\mathrm{T}} [\mathrm{M_n}] [\mathrm{E_n}] = [\mathrm{U_a}]^{\mathrm{T}} [\mathrm{M_a}] [\mathrm{E_a}] = [\mathrm{U_a}]^{\mathrm{g}} [\mathrm{E_a}] \tag{12}$$

Now when considering one particular "ij" term, the orthogonality and pseudo-orthogonality can be easily written (using just the last term without the mass terms) as

$$\mathrm{ORT}_{ij} = u^{g}_{i1}u_{1j} + u^{g}_{i2}u_{2j} + u^{g}_{i3}u_{3j} + \cdots + u^{g}_{ia}u_{aj} = \begin{matrix} 1(\text{ for } i = j) \\ 0(\text{ for } i \neq j) \end{matrix} \tag{13}$$

$$\mathrm{POC}_{ij} = u^{g}_{i1}e_{1j} + u^{g}_{i2}e_{2j} + u^{g}_{i3}e_{3j} + \cdots + u^{g}_{ia}e_{aj} = \begin{matrix} 1(\text{ for } i = j) \\ 0(\text{ for } i \neq j) \end{matrix} \tag{14}$$

In this form, a simple comparison, term by term, shows the DOF correlation. The CORTHOG manipulation is the comparison of each of the terms from the relationships in Eq. 13 and Eq. 14 for corresponding degrees of freedom.

$$\begin{array}{lccccc} \mathrm{POC}_{ij} & = u^{g}_{i1}e_{1j} + u^{g}_{i2}e_{2j} + u^{g}_{i3}e_{3j} + \cdots + u^{g}_{ia}e_{aj} & & & & \\ & \Updownarrow & \Updownarrow & \Updownarrow & \Updownarrow & \mathrm{CORTHOG}_{ij} \\ \mathrm{ORT}_{ij} & = u^{g}_{i1}u_{1j} + u^{g}_{i2}u_{2j} + u^{g}_{i3}u_{3j} + \cdots + u^{g}_{ia}u_{aj} & & & & \\ & \vdots & \vdots & \vdots & & \\ & \mathrm{dof1} & \mathrm{dof2} & \mathrm{dof3} & & \end{array} \tag{15}$$

Equation (13) and (14) above show the individual contribution for each DOF. The comparison of these (as a simple difference or other scaled forms) clearly shows the contribution that each DOF makes to the individual terms of the ORT or POC matrix and the comparison clearly show the differences between the FEA model and the Test results.

A simple 3 dof example will identify the concept very easily. One misconception is that the closer the off-diagonal terms are to zero, the less correlated the vectors are to each other. The problem is from the equation above, the sum of a large number of terms are required to obtain any one off-diagonal term. The numbers themselves can be large or small, but the sum of all of them will equal zero if the modes are orthogonal to each other. While this is true for analytical statement of orthogonality, this may not necessarily be true when combining analytical and experimental data sets. Two vectors are compared – one is the ORT and the other is the POC. Now in Figure 4, the individual terms for each calculation sum up to be zero but notice that the comparison on a term by term basis shows there are differences. So while the off-diagonal term may have a summed value of zero, this does not necessarily imply that the two vectors are indeed orthogonal to each other and there may be discrepancies looking on a term by term basis.

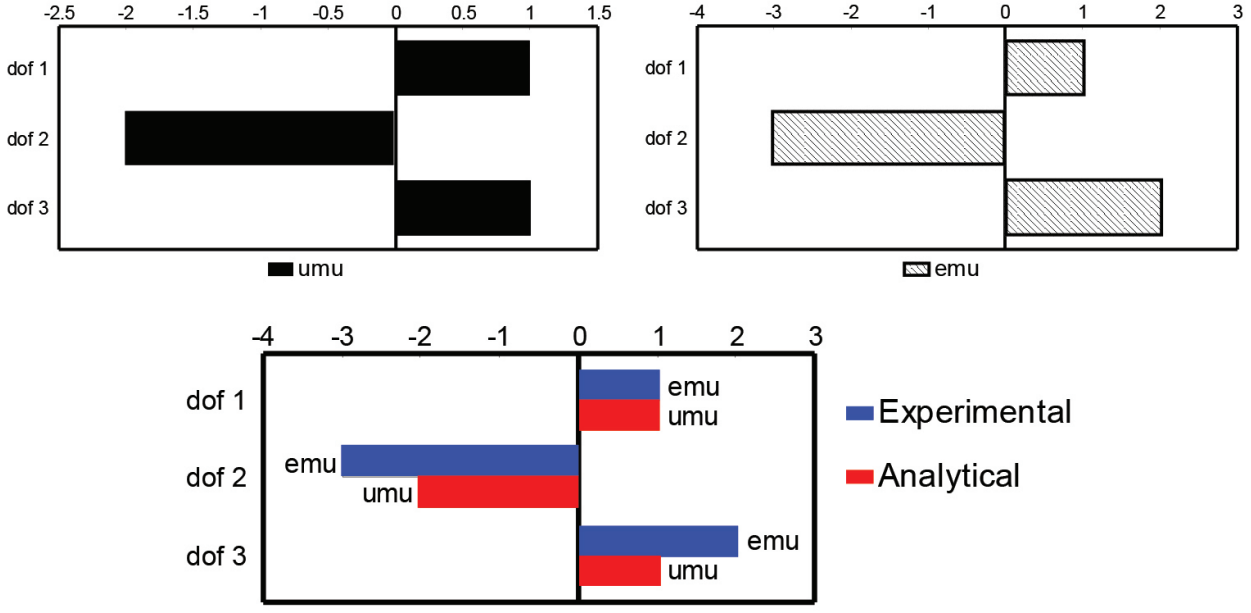


Fig. 4 CORTHOG for Simple 3 DOF System

Data Consistency Assessment Function (DCAF)

The Data Consistency Assessment Function (DCAF) [11, 12] is a tool that enables the identification of items in a data base that may not seem to be correlated with the rest of the data set. Basically, the DCAF systematically removes individual data points from a data set and uses synthesis to recreate the removed point which is then compared to the original data to see how well the removed data point compares. Obvious poor data points can be easily identified and bad data points can be removed from the data set. Figure 5 schematically shows the process for assessment of mode shapes for instance. (In essence, the DCAF is nothing more than removal of outliers as typically done in fitting of data.)

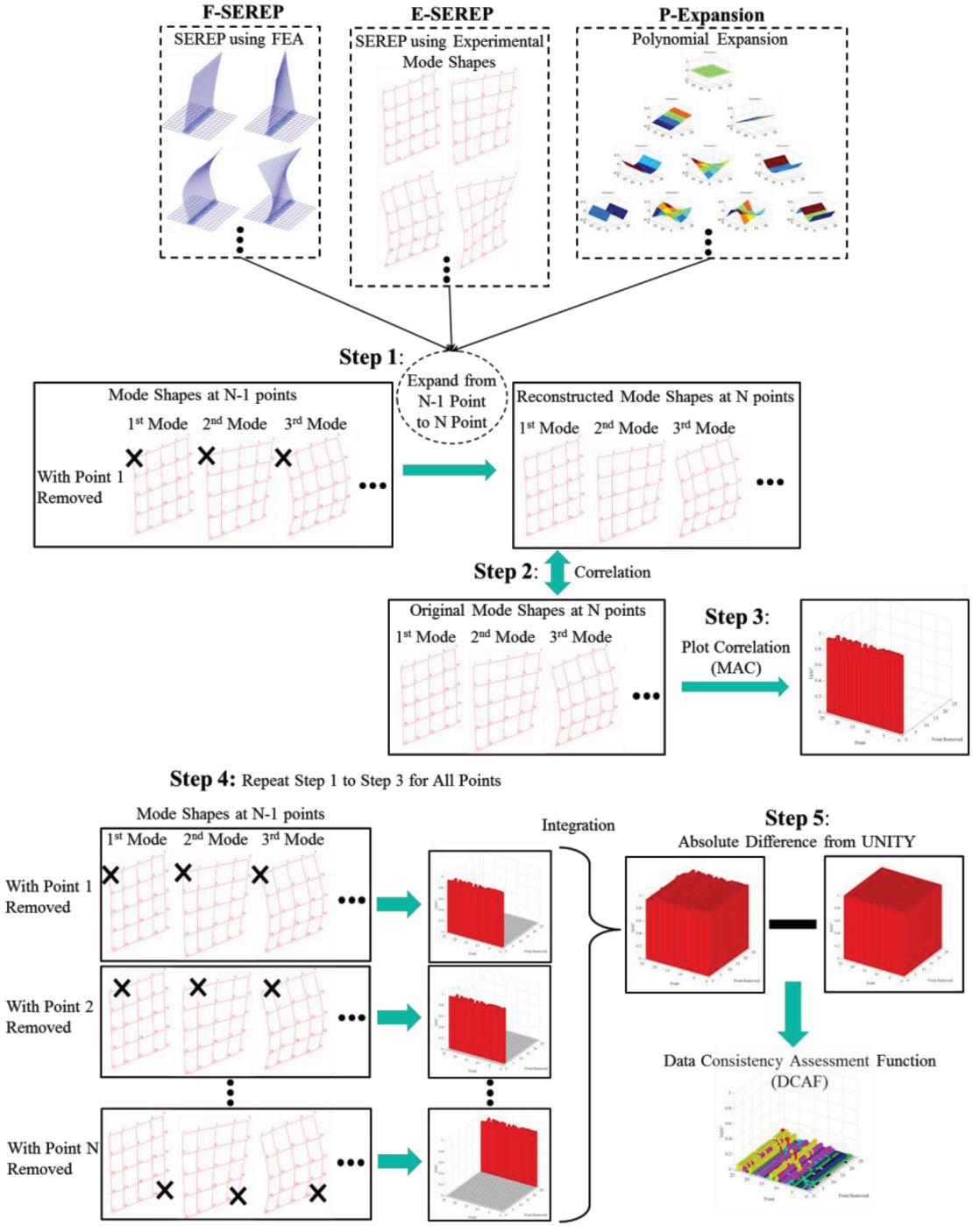


Fig. 5 Process of DCAF for mode shape assessment

Examples

In order to illustrate some of the differences and peculiarities of all the tools, several examples are included. In all cases, real normal modes are used to keep the numerics simple to show the strengths, benefits and quirks of the different tools.

Modal Assurance Criteria – Example of a Good MAC

The first example is one for a structure where the general MAC calculation produces results as would be expected. A MAC was performed for a finite element model for a simple Base Upright (BU) structure and the experimental modal test results obtained from a good modal test. Figure 6 shows the MAC matrix for an example where there are no major issues that arise and the MAC matrix provides good insight into the correlation between the test and FEA.

FEA Mode Number	Experimental Mode Number	FEA Frequencies (Hz)	Experimental Frequencies (Hz)	Frequency % Difference	MAC Value (%)
1	1	26.03	26.17	-0.54	99.8
2	2	70.70	67.34	4.75	98.5
3	3	77.69	75.89	2.32	99.3
4	4	108.80	104.48	3.97	97.1
5	5	158.01	161.55	-2.24	98.9
6	6	269.88	270.00	-0.04	98.9
7	7	304.40	299.93	1.47	96.5
8	8	350.45	354.60	-1.18	98.5
9	9	419.36	419.22	0.03	94.3
10	10	457.13	482.15	-5.47	92.7
11	12	541.29	547.84	-1.21	96.6
12	11	563.47	538.20	4.48	86.3
13	13	777.29	771.76	0.71	90.3
14	14	779.61	791.18	-1.48	97.7
15	15	862.76	871.90	-1.06	84.3

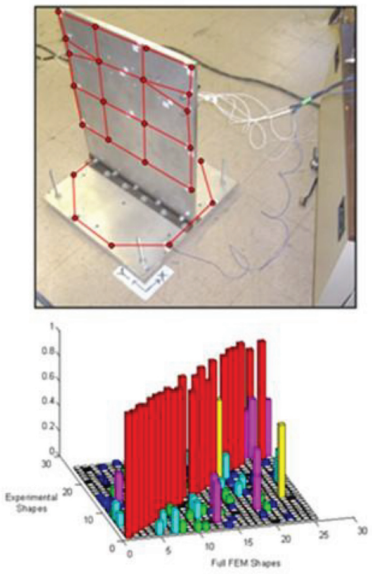


Fig. 6 Comparison of FEA and Modal Test for the Base-Upright Structure

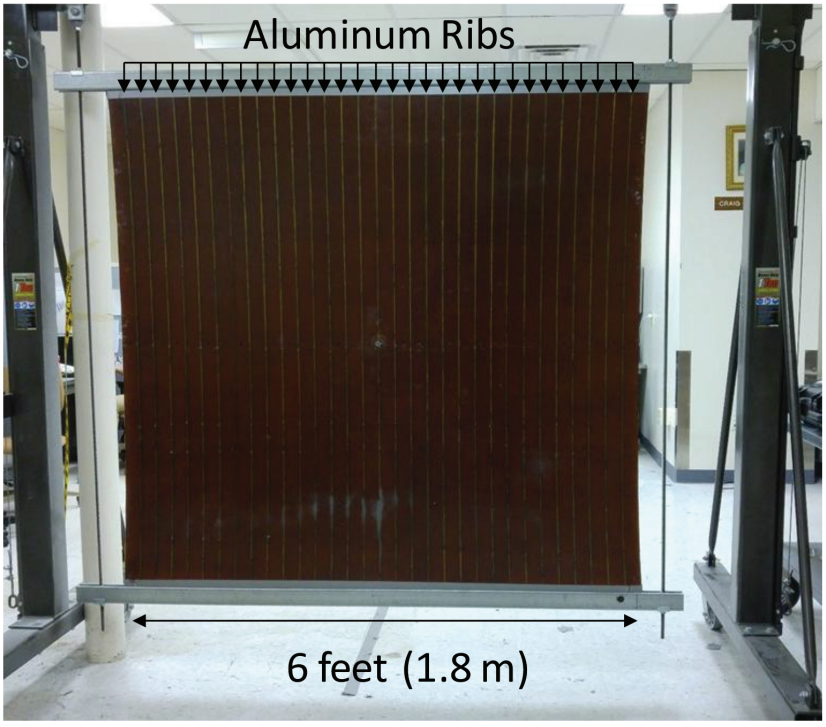
Modal Assurance Criteria – When Just a Portion of the Structure Lacks Correlation

So a question was raised as to how good a finite element model was that was used for predictions for Bragg scattering studies. A modal test was performed to compare to a very preliminary finite element model of a rib-stiffened urethane plate structure; the way the structure was designed and fabricated, the clamping arrangement (and modeling of that) was unclear and not straight-forward. The structure is shown in Figure 7 along with the mode shape of concern.

For one of the lower order modes, a MAC of 94% was identified from the correlation. Without really looking at the mode shapes in detail, a MAC value this high would generally indicate a good shape correlation. But looking at the mode shape more deeply and closely viewing the overlaid mode shapes, there appeared to be about 75% of the structure that was very correlated between the FEA and test results. But there was about 10% of the structure that really didn't show any correlation to the naked eye. And the portion that was most different was situated in the region where the clamping arrangement was of concern. So this is a good case where having a very high MAC value does not really guarantee that the entire mode shape is well correlated. (Of course, a mass orthogonality would be more telling but that was not done or required as part of the user initiated validation of the model.)

Modal Assurance Criteria – When Not All of the Structure is Measured

So a number of years ago, some work was done on jet engine turbine blades. As part of the certification of remanufactured jet engine turbine blades, 25 OEM blades and 25 re-engineered blades are subjected to correlation studies for the first 25 modes of the blades. This is typically up to 20 KHz to 30 KHz ranges. The criteria is that the MAC must be above 90% for all modes of all blades in order for the certification of the new blades to be accepted. Well, there was significant effort to accomplish



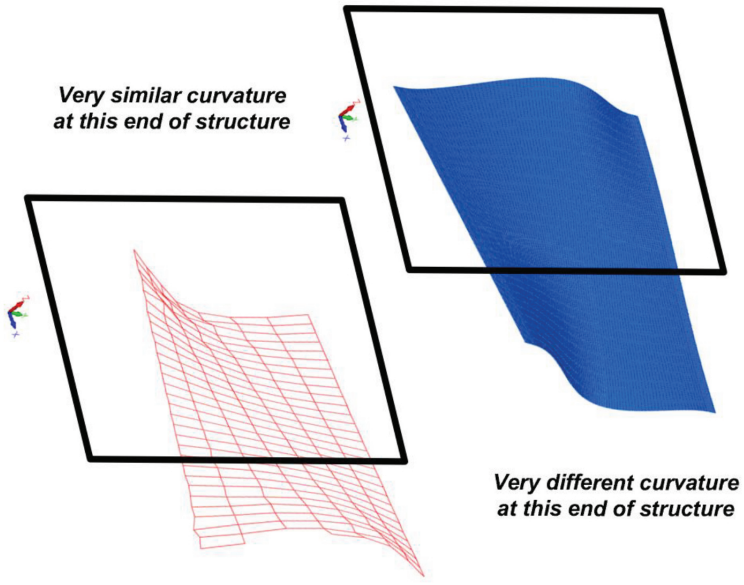
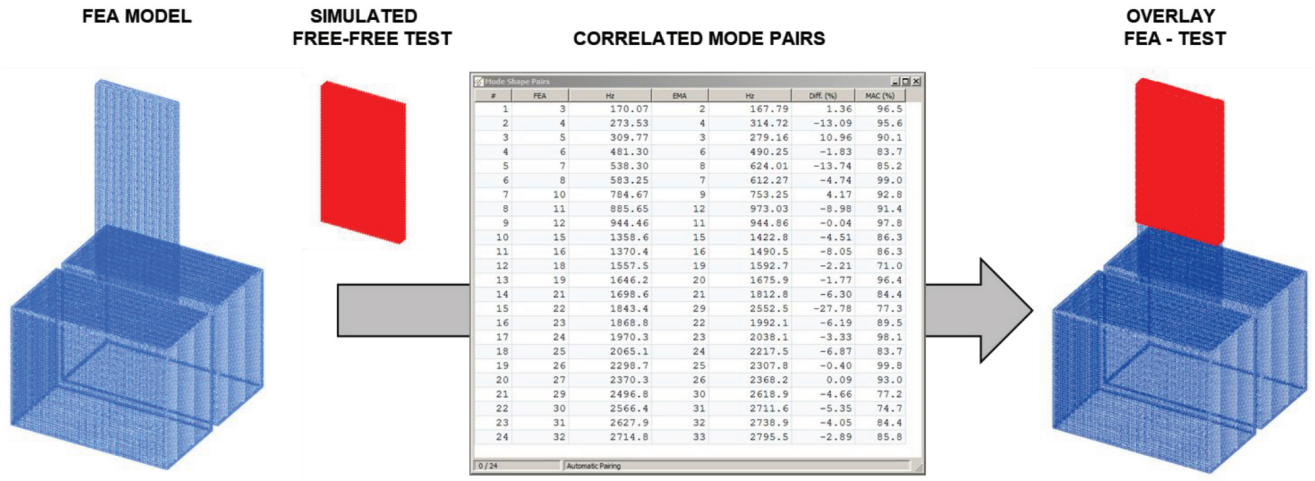


Fig. 7 Large Plate Structure with Mode Shape Comparison

this and along the way a number of questions arose as to how to handle variability in the blades and how address what the geometric correlation should be (considering that every blade is really slightly different) and how to actually perform the correlation. The FAA folks were adamant that a MAC of 90% was critical to assure that the blades were the "same" (whatever that means).

In an effort to help everyone understand MAC and level of correlation, a simple cantilevered plate FEA model was developed with a fixed BC. An equivalent FEA model was generated with a free-free BC to simulate test data. Because test typically cannot measure as many points, only the upper two-thirds of the model was used to simulate test data; this was mainly done because a laser holography system was being used to capture mode shapes and due to line of sight issues the entire blade could not be captured in the field of view of the optical system.

Now clearly comparing free-free modes to cantilever modes should result in an obvious conclusion that the modes are not very similar at all. However, the correlation showed something quite different. And this was intentionally done to highlight the fact that the MAC correlation might produce questionable results. Now the correlation results shown in Figure 8 shows that 20 of 24 modes correlated with a MAC of above 80% AND more than 10 modes correlated with MAC values above 90% with 7 modes above 95%. So using the MAC can provide questionable results when comparing modes from two very different configurations – one free-free and the other built in at one end.



#	FEA	Hz	EMA	Hz	Diff. (%)	MAC (%)
1	3	170.07	2	167.79	1.36	96.5
2	4	273.53	4	314.72	-13.09	95.6
3	5	309.77	3	279.16	10.96	90.1
4	6	481.30	6	490.25	-1.83	83.7
5	7	538.30	8	624.01	-13.74	85.2
6	8	583.25	7	612.27	-4.74	99.0
7	10	784.67	9	753.25	4.17	92.8
8	11	885.65	12	973.03	-8.98	91.4
9	12	944.46	11	944.86	-0.04	97.8
10	15	1358.6	15	1422.8	-4.51	86.3
11	16	1370.4	16	1490.5	-8.05	86.3
12	18	1557.5	19	1592.7	-2.21	71.0
13	19	1646.2	20	1675.9	-1.77	96.4
14	21	1698.6	21	1812.8	-6.30	84.4
15	22	1843.4	29	2552.5	-27.78	77.3
16	23	1868.8	22	1992.1	-6.19	89.5
17	24	1970.3	23	2038.1	-3.33	98.1
18	25	2065.1	24	2217.5	-6.87	83.7
19	26	2298.7	25	2307.8	-0.40	99.8
20	27	2370.3	26	2368.2	0.09	93.0
21	29	2496.8	30	2618.9	-4.66	77.2
22	30	2566.4	31	2711.6	-5.35	74.7
23	31	2627.9	32	2738.9	-4.05	84.4
24	32	2714.8	33	2795.5	-2.89	85.8

Fig. 8 Correlated Mode Pairs – Built-in Plate vs Free-Free Plate

Modal Assurance Criteria – Numbers can give a false impression (excerpt from Modal Space Articles)

But often times, people state that MAC values lower than 90% are acceptable with values as low as 80% being accepted; many state that real structures are much different than academic structures evaluated in an educational environment. Well, this researcher is very sensitive to the validity of this common statement heard mainly due to the fact that the mass matrix is not included in the evaluation.

Two examples are presented to show some MAC values that result from modes that are clearly not similar at all. The first case is to compare the rocking rigid body mode of a free free beam with the first cantilever mode of that beam. Figure 9 shows that there is about 60% similarity indicated by MAC.

These two vectors have nothing to do with each other but MAC shows 60% correlation. How could that possibly happen. Well looking at the lower right inset in Figure 9, notice that the values of the cantilever mode shape from the base to midspan of the beam are very small and their contribution in the MAC calculation are very small. And looking at the mode shape from the midspan to the tip of the beam, the values are much larger and with the naked eye you can see that the rocking rigid body mode looks quite similar. And that is what the MAC is indicating. But we know these two modes are not similar at all. Of course, there is no reason to compare these modes, but this was done to illustrate what a 60% MAC value indicates. So, the numbers can be very misleading.

Now let's proceed on with two cantilever beams – one with a uniform mass distribution and one with an additional lumped mass at the center of the beam (20% of the weight of beam). Now look at the MAC between Mode 1 and Mode 2 with the uniform mass distribution, the MAC will show little correlation as expected. But when the MAC is performed on the beam with the additional lumped mass, the MAC between Mode 1 and Mode 2 is almost 80%. *Now this is very disturbing!* But looking at the uniform beam modes in the upper right portion of Figure 2 and comparing them to the modes in the lower right portion in Figure 10, the same type of issue as discussed in the earlier case can be easily seen. The mode shape values close to the built in end have very little contribution to the MAC. And the values of the shape towards the end of the beam for Mode 1 and Mode 2 look very similar. And that is what MAC is indicating again. But these two modes are orthogonal to each other but the MAC completely breaks down here. However, the mass orthogonality clearly identifies the vector status because the proper mass distribution was included in the orthogonality check. The MAC has no way to account for the uneven mass distribution for this case.

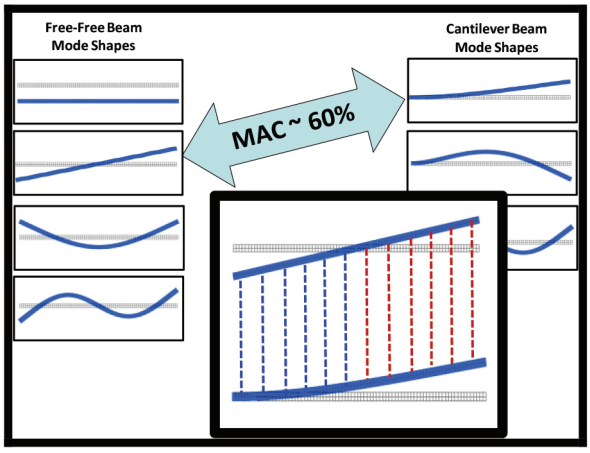


Fig. 9 Mode Shapes for Free-Free Beam (left) and Cantilever Beam (right)

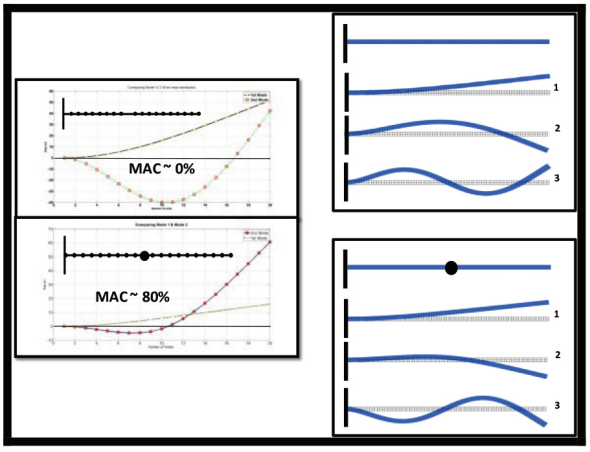


Fig. 10 Cantilever Beam Mode 1 and Mode 2 with Uniform Mass Distribution and Center Lumped Mass

Modal Assurance Criteria – When Geometry Doesn't Exactly Match

So back to the FAA turbine blade certification project. The earlier example showed a cantilever beam/plate structure compared to its free-free counterpart. Obviously, this was expected to show little correlation but the analysis showed something quite different. In this example [13], a simple cantilevered plate structure used for academic studies (related to fixture calibration work) is used to show correlation between an analytical model and a modal test data set. To illustrate the effects of geometry differences that were faced in regards to certifying 25 re-engineered blades, comparing the mode shapes amongst all the blades, the test model was dithered a small amount (one element width of the FEA) to illustrate how sensitive the MAC value can be. This is critically important when identifying node point pairs as part of the correlation process. Small geometry differences can affect the node correlation process when each of the test turbine blades are NOT the exact same dimensions. From an earlier study over 10 years ago, this simple graphic illustrated the problem faced when the geometry changes from blade to blade as shown in Figure 11. There is concern on exactly how to identify the node pair points for the correlation process. Four different cases were shown that, in general, had the same trends but the values for the correlated modes showed differences.

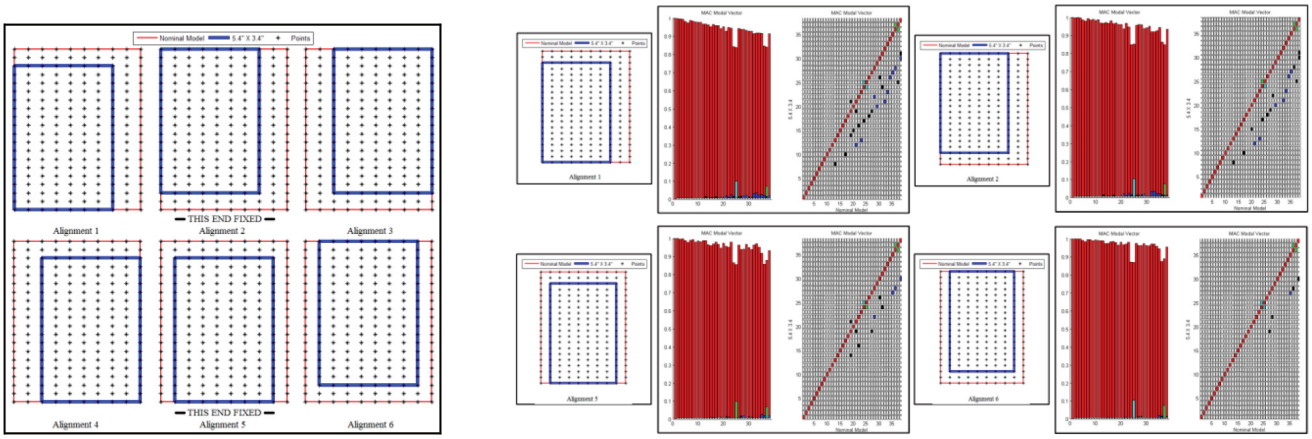
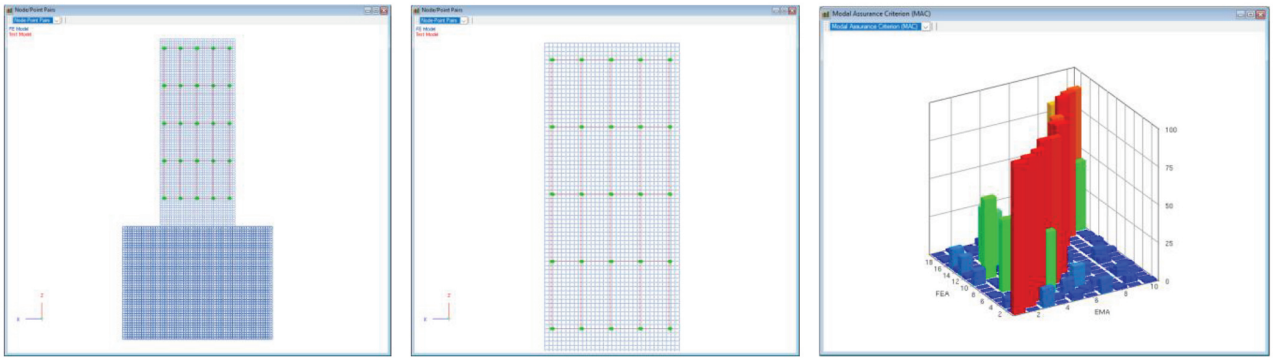


Fig. 11 Comparison of MAC Degradation Resulting from Node Pair Adjustments

In order to put numbers to the study a simple cantilevered plate, bolted to a larger more massive block, was used to simulate a cantilever boundary condition (which was similar to the manner in which the turbine blades are fixed and tested). The model was generated and an experimental modal test was performed. An initial correlation was performed with what seemed to be the proper node pair correlation based on the test performed. But a second correlation was performed where the test model was shifted one element up in the vertical direction. Node point pairs were identified and correlation was performed. The results are shown in Figure 12.



Orig Mode Shape Pairs

#	FEA	Hz	EMA	Hz	Diff.	MAC
1	1	364.65	1	370.75	-1.65	99.8
2	2	1586.43	2	1587.06	-0.04	98.9
3	3	1963.11	3	1966.74	-0.18	98.6
4	5	4994.55	4	4998.96	-0.09	98.8
5	6	5347.91	5	5351.97	-0.08	95.9
6	8	8714.53	6	9147.56	-4.73	98.5
7	12	10202.15	7	10322.81	-1.17	91.8
8	14	12668.24	8	13151.85	-3.68	98.5
9	16	14138.67	9	14274.81	-0.95	95
10	18	15048.8	10	15504.16	-2.94	90.7

Mode Shape Pairs - Element Shift

#	FEA	Hz	EMA	Hz	Diff.	MAC
1	1	364.65	1	370.75	-1.65	99.7
2	2	1586.43	2	1587.06	-0.04	98.9
3	3	1963.11	3	1966.74	-0.18	99.8
4	5	4994.55	4	4998.96	-0.09	99.5
5	6	5347.91	5	5351.97	-0.08	99.7
6	9	9103.49	6	9147.56	-0.48	99.5
7	12	10202.15	7	10322.81	-1.17	98.9
8	14	12668.24	8	13151.85	-3.68	98.8
9	16	14138.67	9	14274.81	-0.95	98.7
10	18	15048.8	10	15504.16	-2.94	92.7

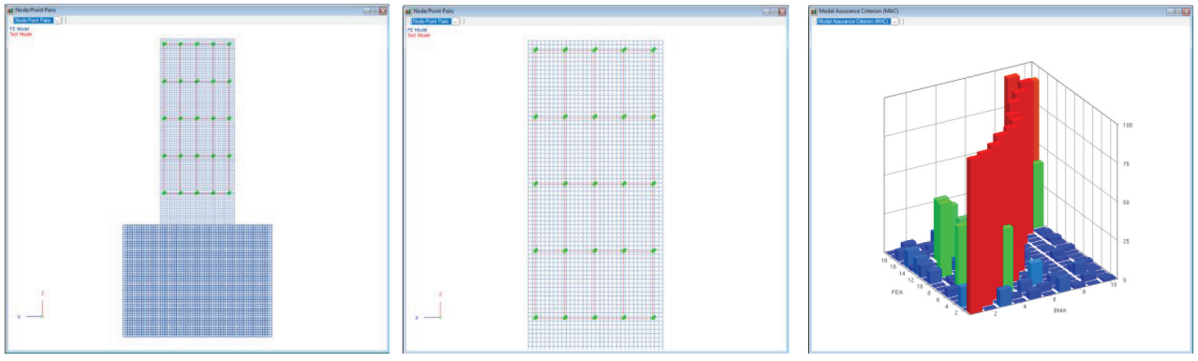


Fig. 12 Effect on MAC Values due to Shifting of Node Pairs By One Finite Element

So here out of the 10 correlated mode shape pairs, for 6 mode shape pairs, the MAC value changed by less than 2%. But 3 mode shape pairs changed by 2% to 4%. But most importantly 1 mode shape pair changed by over 7%. So, this seriously puts in question how good is a MAC threshold of 90% as a criteria for acceptance. Now this example simply illustrated with just a shifting of the model and test data by one element width in a very controlled environment how dramatic the results can be. But now consider the problem when many prototype re-manufactured blades are compared to each other. The manufacturing tolerances alone bring into view the fact that the node point paring can be arbitrarily "adjusted" to provide better correlation – or if care is not taken, blades that have so called good correlation may be questioned.

The MAC is a tool that is used often but these examples show just how much care and caution needs to be given in order to obtain useful results. For this blade correlation work, a different approach is needed which is not as sensitive to these geometric differences that can strongly affect the MAC results. Work on polynomial based correlation techniques provide an alternative for this type of correlation. Work done in this area has shown significant improvement in the correlation determination but "heritage" hinders the acceptance of newer approaches and concepts.

Orthogonality – Pseudo Orthogonality

Now orthogonality checks are a relatively simple calculation. However, the mismatch between the number of finite element DOF and number of measurement points is where the difficulty arises. There are essentially three options possible:

1. Reduce the finite element mass matrix down to the set of test DOF
2. Expand the mode shapes to the finite element set of DOF
3. A combination of reduction and expansion to a common set of DOF

A simple portal frame structure is used to present some test cases [14–16]. The model is shown in Figure 13 with some simulated measured data used for the evaluations presented here; one model uses a lumped mass formulation whereas the second model uses a consistent mass matrix. While the frequencies and mode shapes are very similar, this simple difference helps to see some very important aspects of correlation. Further discussion can be seen in the original reference papers covering these examples.

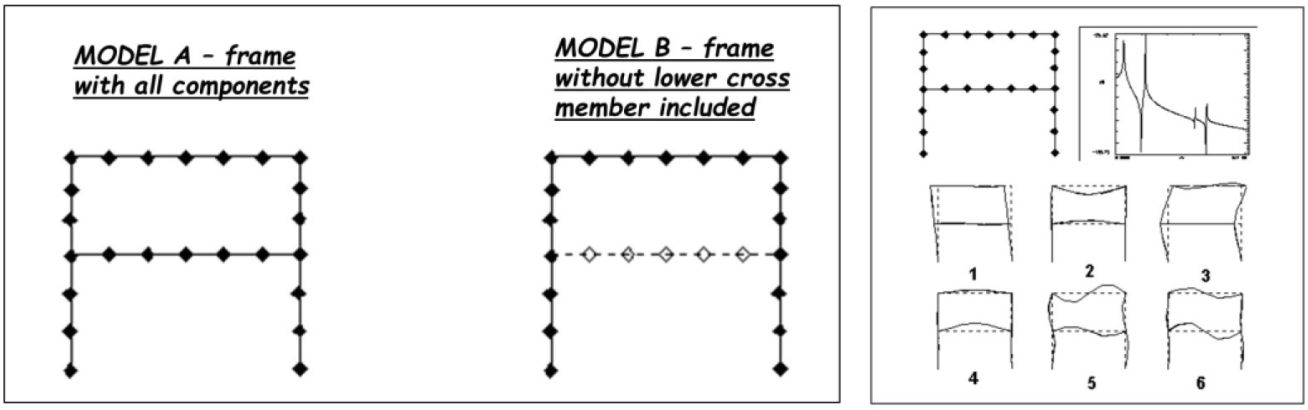


Fig. 13 Portal Frame with Measured FRFs

In the first case, all DOF available from test are used to correlate to the finite element model. The results are shown in Figure 14 on the left and without the interior cross piece on the right. Recall that these two models are essentially the same and only have differences due to the formulation of the mass matrix. The analytical vectors used here have purity in that there is no experimental variations seen.

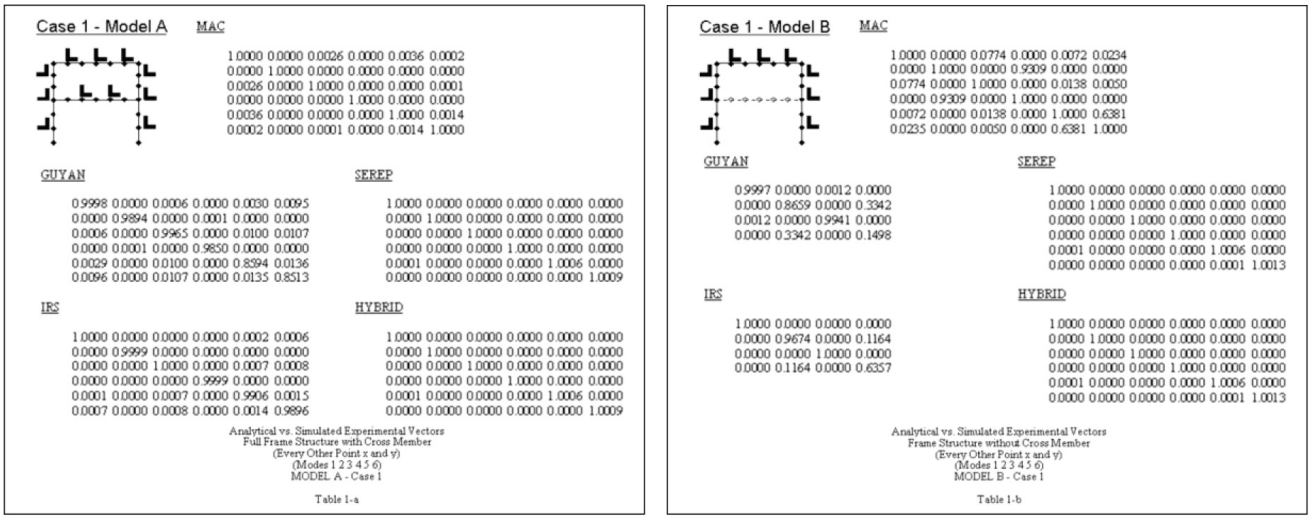


Fig. 14 Portal Frame with Two Orthogonality Checks Using Simulated Experimental Data

Several important items need to be noted here regarding the MAC and orthogonality checks shown [14–16].

- MAC provides useful results providing that sufficient dofs are used to spatially identify the modal vectors. Comparing the left plot the MAC shows very good correlation for the main modes of the system. However, on the right, the MAC shows significant correlation to the other modes. This is due to the fact that the cross piece, which has significant modal contribution, is not included in the measurement. The MAC is prone to this when the complete modal vector is not identified.

- As far as the different reduction techniques, there is a general improvement of the vector correlation identified using the IRS technique rather than Guyan. This is expected because the Guyan reduction is only based on the reduction of the stiffness matrix and the mass effects are not included. The IRS technique adjusts the reduced matrices to account for the embedded inertia related to the DOF that are not retained in the process.
- There is a general improvement of the vector correlation identified using the SEREP and Hybrid techniques rather than the IRS and Guyan techniques. There does not appear to be any benefit of the Hybrid technique rather than the SEREP technique; this discussion is beyond the scope of this paper (see references for additional info)
- One very important item to note is that the diagonal terms of this orthogonality check can actually be greater than 1.0. This is not true for an analytical mass orthogonality due to the inherent formulation. However, there is no guarantee as to how the experimental vectors are scaled; in addition, the analytical properties such as density and Young's modulus can introduce a scale that is clearly observed in this calculation. Some software "forces" the scale to be no greater than 1.0. This is not suggested because this can mask obvious material property differences that may exist in the model; rescaling the vectors tends to mask some of the critically important info.

In the second case, actual experimental data is used in Figure 15. Essentially the same remarks above are applicable here. The case above allows for very clear comments to be extracted from the case evaluated that is not affected by experimental variation that may occur in actual measurement scenarios. In this case experimental variations do exist but the same general trends are observed; but these are not a result of experimental variations and are due to expected differences.

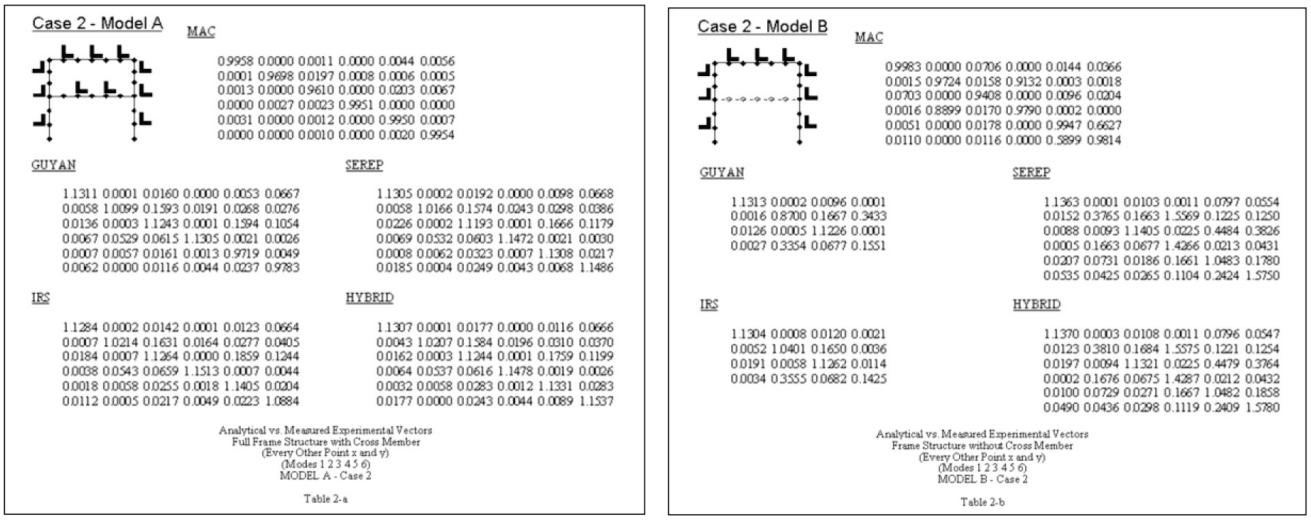


Fig. 15 Portal Frame with Two Orthogonality Checks Using Actual Experimental Data

This particular paper [14–16] goes on to make several different evaluations in regards to the selection of DOF to include in the orthogonality checks and several items were noted (see additional content in the original papers).

- Selection of DOFs used for correlation studies be well thought out; the DOFs must be significant in all of the modal vectors
- This point is particularly important if the experimental modal vectors are used in an expansion process such as SEREP; inclusion of bad data points will have a direct effect on the accuracy of the expanded modal vector.
- Use of all measured dof may actually degrade the orthogonality check, resulting in poor vector correlation
- DOF which are points of high response for some modes may be poor for other modes, especially for structures with directional modes
- Using a carefully selected set of dof for each mode may result in a more accurate indication of correlation
- Due to the directional nature of the modes, confidence in the accuracy of ALL of the measured 'a' dof for ALL of the six modes is not high
- MAC is good correlation indicator for large shape components
- Expect the largest shape dofs are the most accurate
- Use of all dof for all modes for orthogonality checks in reduced space may not necessarily provide a good indication of vector correlation

- Difficulty in selecting an appropriate set of 'a' dof that are good for all modes evaluated
- Using a more select set of dofs may produce a better pseudo orthogonality check
- Correlation studies are most likely to produce the best results if the modes are expanded in sets using an optimum set of 'a' dofs for each set of modes

These results then pointed to the fact that a more elaborate assessment of DOF correlation needs to be performed. However, tools such as CoMAC do not include mass scaling. The CORTHOG was specifically developed to assist in a more robust evaluation of DOF contribution to the correlation process. A few examples in the next section highlight this.

DOF Correlation – CoMAC and CORTHOG

There are several misconceptions regarding the correlation of DOF for modal vectors. The next several examples [17–19] aim to clarify some of these.

A simple 4 dof cantilever beam is used to quickly and easily gain an appreciation of contribution of DOF towards the Pseudo Orthogonality Check in Figure 16. The analytical vectors were perturbed with a slight amount of noise. The vectors themselves appear to look the same and the MAC values are almost identical. But the POC clearly shows one off-diagonal term (the 2-1 term with 0.056) that is very large. The CORTHOG reveals that DOF 3 has significant discrepancy. In fact, if the noise is removed from all DOF, except DOF 3, the off-diagonal term only changes from 0.056 to 0.054. This clearly shows how one DOF can have a dramatic effect on the off-diagonal terms. Note that this DOF is associated with the node of the second mode. This can be a more difficult DOF to measure and will not be identified with the MAC or CoMAC. Because the CORTHOG is mass weighted, the contribution of this very small mode shape value can be very important to the POC computed.

A second example involves a structure with very directional modes. The experimental modal test involved almost 100 measurement DOF (using 33 triaxial accelerometers. Looking at the mode shapes, there are clearly very directional modes

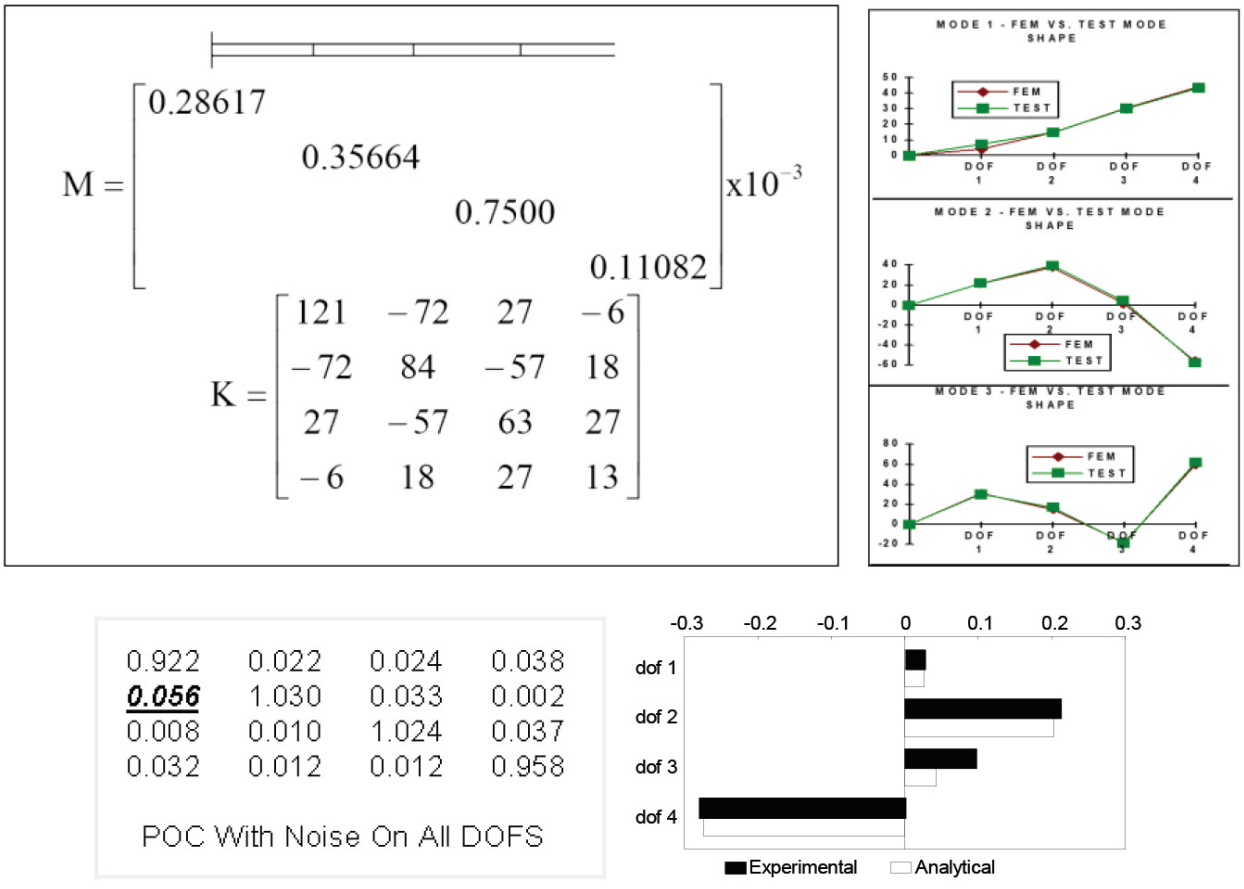
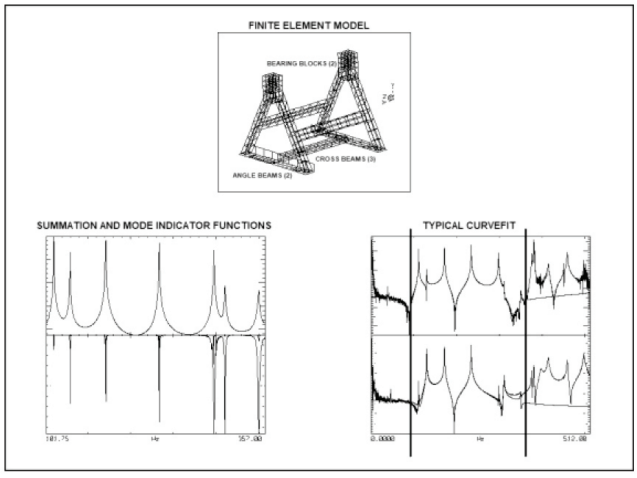


Fig. 16 Simple 4 DOF Cantilever Showing Mass Effect

related to the structure. Using triaxial accelerometers is common practice but obviously there are certain modes where one of the directions measured will not have significant response and will have small shape values more easily contaminated by noise. And each of the different directional modes will have different directions where this problem will occur. That is why an earlier the statement was made "due to the directional nature of the modes, confidence in the accuracy of ALL of the measured 'a' dof for ALL of the modes is not high". This is a common problem testing many engineering frame type structure – these directional modes will fall into this category all the time. The structure is shown in Figure 17 with sample FRFs and extracted mode shapes which are overlaid on the finite element model. Clearly the mode shapes appear to compare well.



MAC

0.85	0.00	0.06	0.00
0.00	0.94	0.00	0.00
0.00	0.00	0.87	0.00
0.00	0.00	0.00	0.94

Diagonal terms good with small off-diagonal terms

POC

1.138	0.048	0.009	0.026
0.008	1.169	0.007	0.004
0.054	0.025	1.257	0.042
0.205	0.045	0.039	1.169

Orthogonality diagonals greater than 1.0 with some off-diagonals values

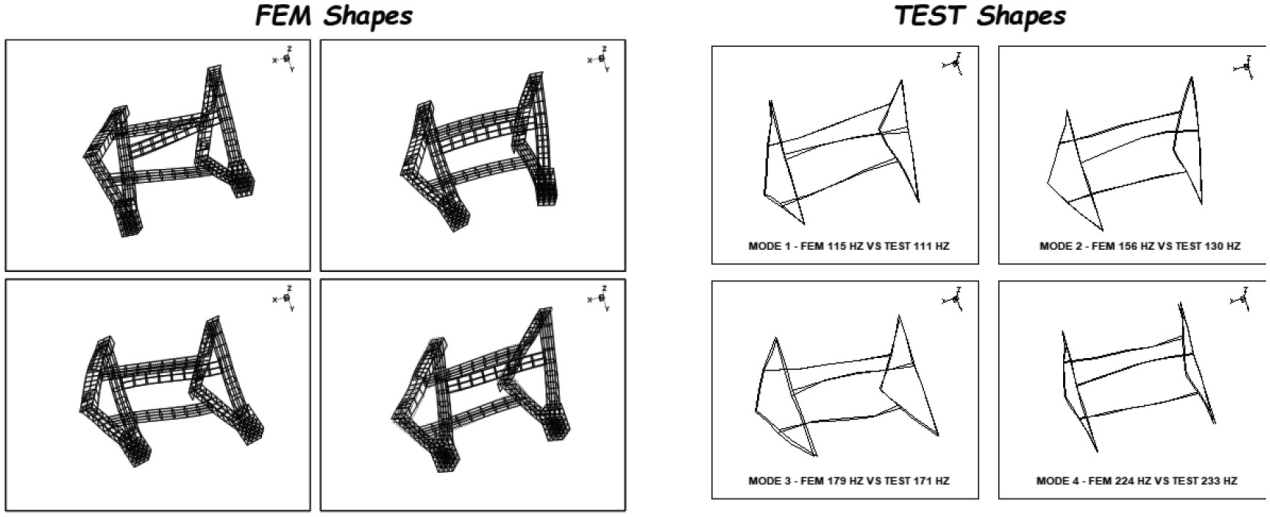


Fig. 17 GTE Frame with Directional Modes

Two modes are further studied using the CORTHOG. One is a comparison of two directional modes (4th experimental mode with 1st analytical mode). Here the off diagonal term showed a high value of 0.205. Certainly there are enough DOF to describe the mode shape. But the problem is that there are five locations on each leg that have essentially zero response but due to noise and mass scaling, their contribution to that off-diagonal term is significant. In fact, over 50% of the error in that off-diagonal term is related to these poorly measured DOF – that is less than 10% of the entire set of measurements which causes that poor off-diagonal term. And looking at the mode shapes, those particular DOF are expected to not be measured well at all. Including them in the POC calculation puts a heavy mark on that off-diagonal term. The CORTHOG clearly shows this as seen in Figure 18.

These results reveal that the DOF along both angle bracket legs have the most discrepancy and account for over 50% of the error in the total summation of absolute simple differences. These particular DOFs, which comprise over 50% of the total error, represent less than 10% of the total number of measured DOF. In reviewing mode 1 and mode 4, these two modes are somewhat directional in nature. Mode 1 is a twisting mode of the structure and mode 4 is a shearing mode. The main contribution to the off-diagonal term is due to these DOFs that are not measured well because of the directional nature of the modes. (The error is likely due to cross axis sensitivity of the transducers used.)

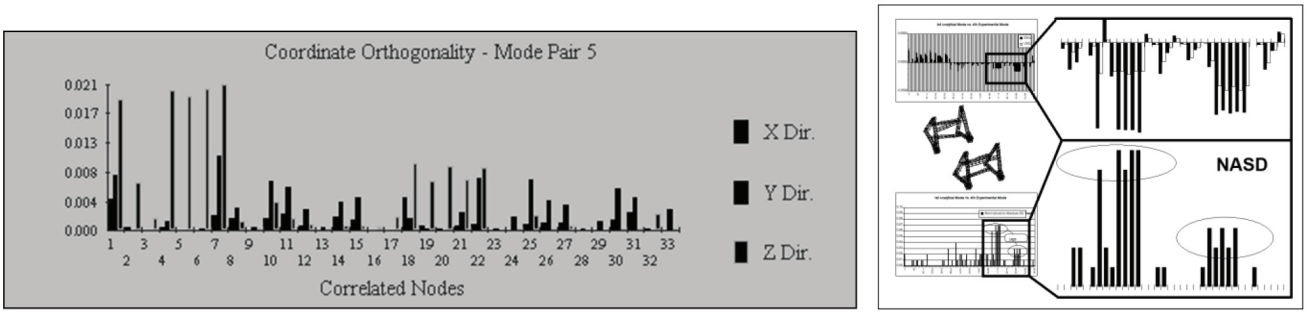


Fig. 18 GTE Frame with Directional Modes – CORTHOG Mode 1 to 4

The other comparison is in regards to the second bending mode of the system shown in Figure 19. The diagonal term shows difference. At first, consideration was given to the fact that the DOF geometric coordinates may not have been correct and the node correlation was to a different point in the model. But further investigation revealed that the mass in the pillar blocks were incorrectly input into the finite element model. The CORTHOG clearly showed this discrepancy; note that the CoMAC failed to identify this serious difference because mass scaling is not included in the CoMAC formulation.

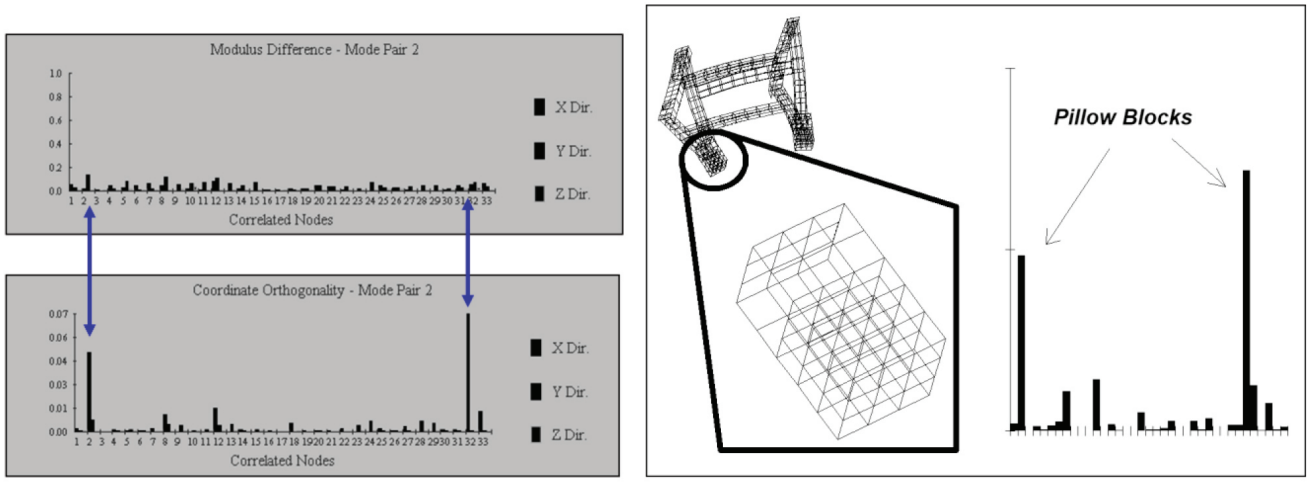


Fig. 19 GTE Frame with Directional Modes – CORTHOG Mode Pair 2

The CORTHOG plots clearly show the large discrepancy between the two degrees of freedom associated with the bearing blocks. This discrepancy may have occurred for several reasons such as the actual mass of the bearing blocks not being well represented, or the joint between the bearing block and the structure is not integral as modeled, or the geometric point for the analytical and experimental models are not truly at the same geometric point in space. While the difference between the mode shapes may not appear significantly different and cannot be identified using the CoMAC, ECoMAC or Modulus Difference, the CORTHOG shows that these two DOFs are significant contributors to the POC. This is due to the fact that these two DOFs have a significant amount of mass associated with them. While it is likely that the discrepancy is due to an analytical modeling error and not an experimental error, there is a definite indication that these two degrees of freedom are not correlated very well when compared to other degrees of freedom in the modal vector. Due to lack of mass scaling, both MAC and COMAC are unable to clearly identify this discrepancy.

Data Consistency Assessment Function (DCAF)

Tools are necessary to help interrogate the data collected. The measurement process is riddled with possible places where errors can crop into the measurement. Unfortunately, the measurement process is not perfect. So there needs to be a critical assessment of the experimental data acquired in order to assure its usefulness. Two specific tools used in the collection and reduction of data are discussed next but unfortunately they do not really interrogate the data in a significant meaningful fashion.

During the measurement process, the coherence is always evaluated. But that coherence may be deceptive because the coherence can be 1.0 and yet the measurement may be faulty. Issues such as phase distortion or phase lag between channels may cause a measurement to be faulty even though the coherence is 1.0; that is because all of the output was due to the measured input. A calibration error will never show up in the coherence function so the measurement may be wrong yet the coherence is 1.0. An accelerometer may be mounted in the wrong direction and the coherence could be 1.0 but the measurement is not associated with the proper point and dirrection. So each of these examples are a few of the common examples where the coherence may give a false sense of security as to the adequacy of the measurement.

Once data is collected and the data reduced with the modal parameter estimation tools, the next step is normally to synthesize the data from the model extracted and compared to the measured data. This provides feedback as to how well the data fits to the model. While this step is very important to the overall process, the problem is that each measurement is only compared individually; there is no association with all of the measurements included in the process. So while a measurement may have a very good synthesized comparison, there is no guarantee that the measurement is appropriate when compared to all the other measurement points describing the model. In a multi-reference test with many accelerometers, simple issues related to calibration, mounting direction, etc. fall into the same dilemma as the coherence.

Other tools are available to assist in further interrogation of the data such as Amplitude Weighted Assessment Function (AWAF) and Enhanced Amplitude Weighted Assessment Function (EAWAF) [20]. These tools help to more clearly show where errors may exist in the entire database describing the mode shapes for the structure. In a recent modal test on a fully dressed truck, there were two separate tests performed with different shaker configurations. Due to sensitivity of the data, a mock up structure [21] was developed to illustrate the situation at hand. One test using 4 vertical only shakers and the other test deployed skewed shakers which allowed for better excitation of all the modes of the structure. The EAWAF clearly showed where errors as part of the modal parameter estimation process were clearly evident. The modes extracted from the skewed shaker configuration were much better than the modes extracted with the vertical only shaker configuration. Two resulting EAWAF plots are shown to highlight the results obtained in Figure 20.

But one additional newer tool developed is the Data Consistency Assessment Function (DCAF) which interrogates all the modal data as a complete set to identify outliers in the data set. DCAF has been used on modal vectors, FRFs, operating data, optical data and general time data sets as well as strain data from optical systems. Basically, the DCAF systematically removes individual data points from a data set and uses synthesis to recreate the removed point which is then compared to the original data to see how well it compares. Obviously poor data points can be easily identified and bad data points can be removed from the data set. This is much different than other correlation tools which only assess individual measurements and never assess the relationship with the rest of the data set.

One recent application used an optical set of time response data which has thousands of data points. Subsets of data were evaluated and there were several “truth” measurement where the correlation was less than 90%. Using the DCAF, five points were clearly identified and removed from the data set thereby increasing the correlation significantly; additional passes on the data set removed additional outliers. Figure 21 shows the DCAF results from the first pass as an example.

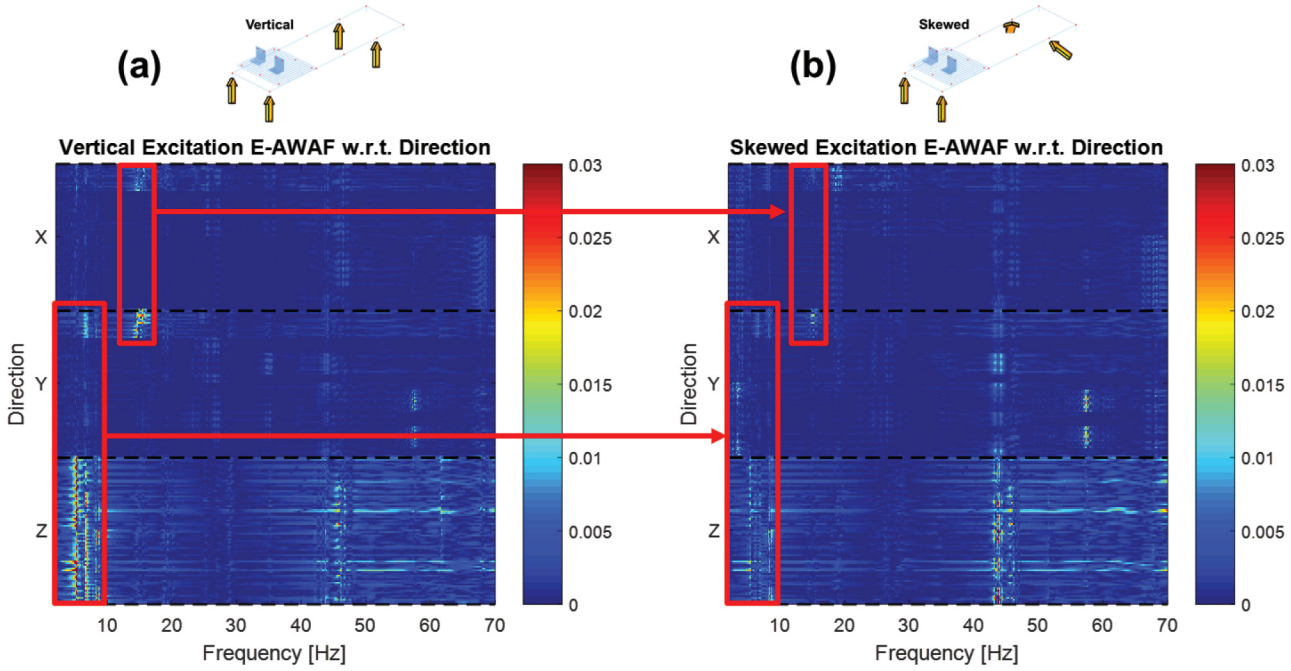


Fig. 20 Comparison of the E-AWAF computed from the (a) vertical configuration data and (b) skewed configuration data. Red boxes superimposed on the E-AWAF color maps to highlight areas of improvement using skewed excitation tests

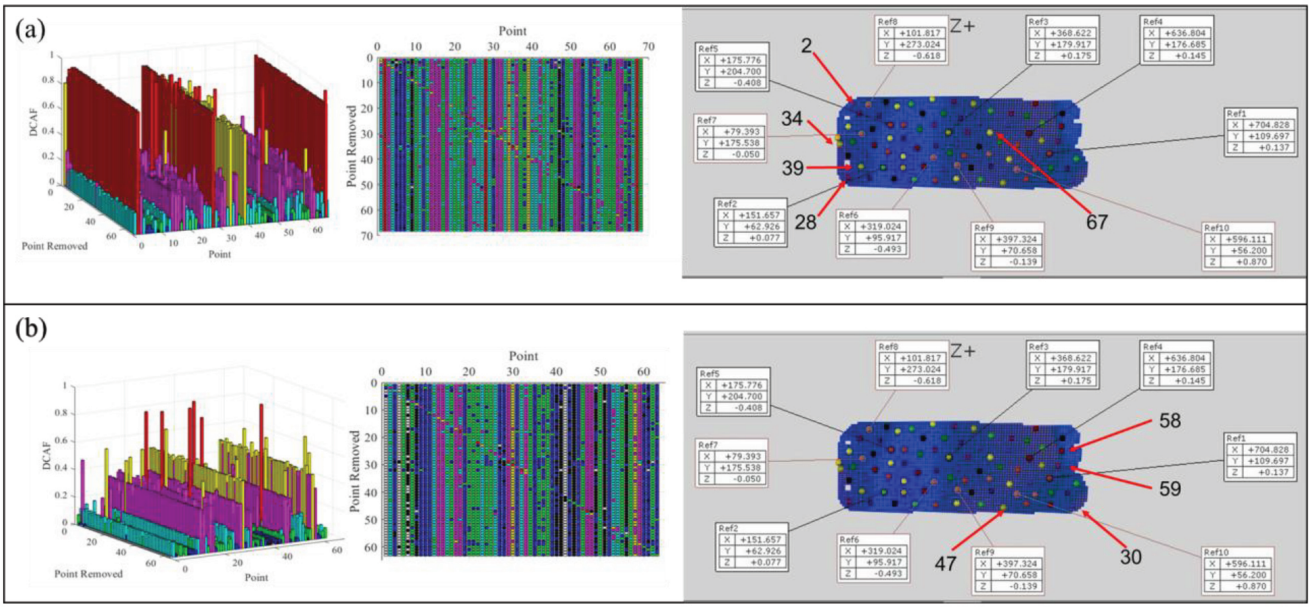


Fig. 21 (a) DCAF results for 68 active elements strain measurements with 5 inconsistent measurements identified (b) DCAF results for 63 active elements strain measurements with 4 inconsistent measurements identified

Observations and Misconceptions

So much has been presented in this paper to start to show some of the problems with the tools that are used for correlation. Using a simple number as a threshold may be necessary for contractual work but the actual correlation may be quite different. So the user must be very careful in the interpretation of the numbers. Suffice it to say that picking a threshold for accept or reject can be very difficult to say the least.

So this paper has presented several important concepts. But beyond that there are several areas where additional discussion is necessary to highlight additional problems extending from what has been presented

- Is all the measured data correct
- Test & Analytical Engineer awareness
- Complex modes – are they real? (no pun intended)

Is all the measured data correct?

Often times, data is collected with a contractual requirement or a mandate identified from above. Data, at times, may be collected without the real understanding of why the data is needed, how it is going to be used or what needs to be extracted from the study. When this is the case, the data may get acquired but the specific reason might cause the test engineer to acquire data in a different manner.

Most importantly, the consistency of the data to itself is rarely, if ever, checked. As mentioned in the paper, issues of coherence or synthesis of data as a validation tool is very important but this only checks the data with regards to itself. The coherence only shows that the particular measurement was collected in a way that the output is linearly related to the input signal. This just attests to the conditions of noise on the data or if a nonlinear measurement may be of concern. While this helps to validate the particular measurement, there is no consistency identified for that one data point in relation to all the other data points that are part of the entire data set. And while using synthesis following curvefitting is always a good step to take, this does not indicate if the data point is consistently related to all the other data points in the data set – this just validates this one individual point.

There are many issues related to measurements with a list too long to include here but a few glaring issues related to leakage, windows, time variance, DAQ dynamic effects, accelerometer cross axis sensitivity, inconsistent excitation (hammer hits for instance) are all very important issues that can have a significant effect on model correlation.

But at the end of the day, the analyst needs to use the data for the study necessary. Important details about the data may cause the test to be conducted differently or for the analysts to use the data in a more judicious manner.

And as stated in this paper before, “Are all the measured DOF for all modes measured with the same accuracy”. Hopefully, from all the items discussed thus far, clearly the measured DOF are not measured with the same accuracy for all modes. Yet when an analyst uses the raw data (in its entirety) for a correlation study, this is absolutely assuming that all the measured data is equally accurate when in fact this is never the case. But almost all correlation software packages treat all the data in unison and the correlation tools are deployed with no understanding that some data points are very good and some data points are absolutely terrible.

Without a very judicious assessment of all the DOF for each modal vector, then all the data is considered equally valid. When correlation studies are performed in this manner then the results will be considered very questionable.

Test & Analytical Engineer awareness?

Often times, there is a disconnect between the test folks and the analytical folks. In most organizations they are in separate groups that may not even be in the same physical location. Both are very unaware of the problems faced by the other. Neither has done any detailed work in the other’s area. There is often a complete lack of understanding of the problems faced by each other in the development of their model or test data. When this is the case, the tools are used in a very formal fashion but lacking the understanding of the limitations and without any respect for the detailed use of the tools. When this is the case, a correlation is preformed with numbers generated but without any real understanding of what the results mean.

Often times, the analytical engineer is passed data from the test engineer and then blindly starts to use that data without any real assessment of the validity of the data. On a recent project over a decade ago, test data was provided with an analytical model with the specific intent to develop a highly reduced order model to be updated with the measured experimental data and to predict accurate response with that measured data. The assumption was that the models and test data base were suitable to undertake the project. After many hours spent with various models, an acceptable model could not be developed. After extensive scrutiny of the test data, the measured data showed significant inconsistencies with itself. Of course, this was the root cause of all the difficulty with the project. Once the data was properly adjusted to maintain consistency, the analysis went forward without a hitch.

There was another case at the turn of the century when some test analysis correlation work was done on some helicopter missile firing systems. A consulting company was hired and spent several weeks evaluating the correlation between several sets of data and models developed. The level of correlation was poor and there was no insight as to what might be viable ways to achieve better correlation. All the typical software correlation tools were deployed but the software is just a tool and the insight must come from the engineer deploying the tools. A few months later different studies were performed but

with a more hands-on approach to understanding the models developed and the test data acquired. Some very basic fundamental steps were taken to achieve far superior results within several days. In the earlier study, data was passed from test engineer to analyst but the data was never assessed before using it with the model. Correlation proved to be a difficult process until some of the basic aspects of the underlying test data were clearly understood and incorporated into the analysis.

The most critical item here is that the test and analytical engineers need to work in unison and very clearly identify all the particulars that can have a very distinct effect on the results of the correlation process. While this is obvious, at times there is a slight disconnect between the parties involved which can be detrimental to the project results.

Complex Modes – are they real?

Now in all the material presented here, real normal modes were assumed. This was necessary so that simple examples could be used to portray some of the difficulties with the tools. Complex modes just make all the numbers more difficult and more complicated but the same exact issues exist with complex modes.

But on the point of complex modes, there are several issues that must be addressed. First and foremost, are the measured modes really complex? Or are there other issues that may lead the engineer to believe the modes are complex, when in fact they may not be complex. Unfortunately, just because the software shows a complex mode does not really confirm the modes are complex – the software just indicates that with the measurements provided, the indication is that there is some complexity to modes based on the measured FRFs used to extract data.

In one recent large modal test on a fully dressed truck structure, complex modes were revealed. And with such a large structure, just about everyone was convinced that complex modes would exist. But that was mainly because everyone convinced themselves that complex modes would exist – but with no strong evidence that complex modes really did exist. The original test utilized a four vertical shaker configuration but the overall measurements were not of spectacular quality with MCOH that was not as good as needed. But everyone was strongly stating that complex modes for this type of structure were clearly expected. After much debate and much discussion, the shakers were arranged in a skewed orientation in the hopes to better excite all the modes of the structure. Overall, the FRFs were greatly improved and the MCOH was also greatly improved. Upon extracting modal parameters, the majority of modes that were classified as complex became far less complex and started to look more like real normal modes. So the way in which the structure is excited may have a direct effect on the measurements and ultimately the modes that are extracted.

This author has seen many, many, many instances where people insisted that complex modes existed. But with further scrutiny of the measurements, measurement technique, instrumentation used and techniques deployed, many of these so called “complex modes” are really just an artifact of the lack of consistent and proper measurement techniques. In the majority of the situation observed, there has almost always been some detail of measurement acquisition, manner of collection of data, and numerous other causes that have manifested themselves as complex behavior that is not related to the structure but more related to the lack of pristine measurements acquired.

Now certainly complex modes can exist from a mathematical standpoint and allowing for complex modes in the development of a modal model is always good practice. But before those so called “complex modes” are used in any further analysis (such as correlation) the experimental data needs to be scrutinized.

Attached here is a short list of some of the common issues observed (in close to half a century of modal testing) that have a strong influence on the extraction of “so-called” complex modes

- Leakage associated with measurement acquisition
- Inappropriate use of windows to minimize leakage
- Time invariance has been seen especially when tests are conducted over several days
- Suspension system (air bags) with bleeding air system
- Bungie cords stretching over time with an inconsistent stiffness
- Cords and cables not secured and allowed to freely vibrate
- Using different accelerometers where the phase roll-off varies from accel to accel
- Channel drift and phase distortion
- Accelerometer cross axis sensitivity
- Inconsistent hammer hits in impact testing

These are just a list of some of the more common problems that come to mind.

Conclusions

This paper attempts to address the use of correlation tools for model validation. The differences between vector correlation tools and DOF correlation tools was presented. The difference between mass scaled and non-mass scaled tools was also presented to illustrate the importance of mass scaling.

Many times there is a lack of understanding of what the tools can and cannot provide. Some of this stems from "old wives' tales" that do not have any credibility but have been blindly accepted for decades as truth.

A series of examples were presented to highlight some of the common misconceptions related to the tools.

References

1. "A Correlation Coefficient for Modal Vector Analysis", Allemang,R.J., Brown,D.L., First International Modal Analysis Conference, Orlando, Florida, November 1982, pp. 110-116]
2. "Spatial Correlation of Mode Shapes, the Coordinate Modal Assurance Criteria (COMAC)", Lieven,N.A.J., Ewins,D.J., Sixth International Modal Analysis Conference, Orlando, Florida, February 1988, pp. 690-695
3. "Development of Efficient Reduced Models for Flexible Body Dynamic Simulation", P.Avitabile, T.VanZandt, N.Wirkkala, B.Birdsong, Twenty-Fifth International Modal Analysis Conference, Orlando, FL, Feb 2007
4. "Model Correlation and Orthogonality Criteria", Avitabile,P., O'Callahan,J.C., Milani,J., Sixth International Modal Analysis Conference, Orlando, Florida, February 1988, pp. 1039-1047
5. "The Coordinate Orthogonality Check (CORTHOG)", P.Avitabile, F.Pechinsky, Mechanical Systems and Signal Processing, Vol 12, No. 3, pp 395-414, May 1998
6. "Coordinate Orthogonality Check (CORTHOG)", Pechinsky,F., Avitabile,P., Twelfth International Modal Analysis Conference, Honolulu, Hawaii, February 1993, pp. 753-760
7. "Evaluation of DOF Correlation using the Coordinate Orthogonality Check", Pechinsky,F., Avitabile,P., Thirteenth International Modal Analysis Conference, Nashville, Tennessee, February 1995, pp. 1313-1320
8. Guyan, R.J., "Reduction of Stiffness and Mass Matrices", AIAA Journal, Vol. 3, No 2, 1965, p. 380
9. O'Callahan,J.C., "A Procedure for an Improved Reduced System (IRS)", Seventh International Modal Analysis Conference, Las Vegas, Nevada, February 1989, pp. 17-21
10. O'Callahan,J.C., Avitabile,P., Riemer,R., "System Equivalent Reduction Expansion Process", Seventh International Modal Analysis Conference, Las Vegas, Nevada, February 1989, pp. 29-37
11. "Data Consistency Assessment Function (DCAF)", Y.Chen, J.Dodson, P.Avitabile, *Proceedings of the Thirty-Eighth International Modal Analysis Conference*, Houston, Texas, Feb 2020
12. "Data Consistency Assessment Function (DCAF)", Y Chen, P Avitabile, J Dodson, Mechanical Systems and Signal Processing (2020), Mechanical Systems and Signal Processing 141 (2020) 106688
13. "Use of Zernike Polynomials for Modal Vector Correlation of Small Turbine Blades", J.Salerno, P.Avitabile, *Proceedings of the Thirty-Second International Modal Analysis Conference*, Orlando, FL, Feb 2014
14. "Study of Modal Vector Correlation using Various Techniques for Model Reduction", P.Avitabile, F.Pechinsky, *Proceedings of the Tenth International Modal Analysis Conference*, San Diego, California, Feb 1992
15. "Study of Modal Vector Correlation using Various Techniques for Model Reduction", P.Avitabile, F.Pechinsky, *Proceedings of the Tenth International Modal Analysis Conference*, San Diego, California, Feb 1992
16. "Comparison of Reduction vs. Expansion for Orthogonality Checks", P.Avitabile, T.Foster, J.O'Callahan, *Proceedings of the Twelfth International Modal Analysis Conference*, Honolulu, Hawaii, Feb 1994
17. "Evaluation of DOF Correlation using the Coordinate Orthogonality Check", F.Pechinsky, P.Avitabile, *Proceedings of the Thirteenth International Modal Analysis Conference*, Nashville, Tennessee, Feb 1995
18. "Coordinate Orthogonality Check (CORTHOG)", P.Avitabile, F.Pechinsky, *Proceedings of the Twelfth International Modal Analysis Conference*, Honolulu, Hawaii, Feb 1994
19. "Understanding the Effects of DOF Contribution to Orthogonality Checks", F.Pechinsky, P.Avitabile, T.Foster, *Proceedings of the Eleventh International Modal Analysis Conference*, Orlando, Florida, Feb 1993
20. "Enhanced Amplitude Weighted Assessment Function (EAWAF)", P.Avitabile, Proceedings of the Twenty-Third International Modal Analysis Conference, Orlando, Florida, Feb 2005
21. "Appropriate Excitation for Extraction of Modes", C.Beale, D.Joffre, P.Avitabile, Proceedings of the Thirty-Seventh International Modal Analysis Conference, Orlando, Florida, Jan 2019

References to papers where generic equations are available

- "A Correlation Coefficient for Modal Vector Analysis", Allemang,R.J., Brown,D.L., First International Modal Analysis Conference, Orlando, Florida, November 1982, pp. 110-116]
- "Spatial Correlation of Mode Shapes, the Coordinate Modal Assurance Criteria (COMAC)", Lieven,N.A.J., Ewins,D.J., Sixth International Modal Analysis Conference, Orlando, Florida, February 1988, pp. 690-695

- Finite Element Model Correlation, SEM Handbook of Experimental Structural Dynamics, DOI:10.1007/978-1-4614-4547-0_34, July 2022
- Guyan, R.J., "Reduction of Stiffness and Mass Matrices", AIAA Journal, Vol. 3, No 2, 1965, p. 380
- O'Callahan,J.C., "A Procedure for an Improved Reduced System (IRS)", Seventh International Modal Analysis Conference, Las Vegas, Nevada, February 1989, pp. 17-21
- O'Callahan,J.C., Avitabile,P., Riemer,R., "System Equivalent Reduction Expansion Process", Seventh International Modal Analysis Conference, Las Vegas, Nevada, February 1989, pp. 29-37
- "Comparison of Reduction vs. Expansion for Orthogonality Checks", P.Avitabile, T.Foster, J.O'Callahan, *Proceedings of the Twelfth International Modal Analysis Conference*, Honolulu, Hawaii, Feb 1994
- "Model Correlation and Orthogonality Criteria", Avitabile,P., O'Callahan,J.C., Milani,J., Sixth International Modal Analysis Conference, Orlando, Florida, February 1988, pp. 1039-1047
- "Mass and Stiffness Orthogonality Checks without a Mass or Stiffness Matrix", Avitabile,P., O'Callahan,J., Thirteenth International Modal Analysis Conference, Nashville, Tenn, February 1995, pp. 1515-1519
- "The Coordinate Orthogonality Check (CORTHOG)", P.Avitabile, F.Pechinsky, Mechanical Systems and Signal Processing, Vol 12, No. 3, pp 395-414, May 1998
- "Coordinate Orthogonality Check (CORTHOG)", Pechinsky,F., Avitabile,P., Twelfth International Modal Analysis Conference, Honolulu, Hawaii, February 1993, pp. 753-760
- "Evaluation of DOF Correlation using the Coordinate Orthogonality Check", Pechinsky,F., Avitabile,P., Thirteenth International Modal Analysis Conference, Nashville, Tennessee, February 1995, pp. 1313-1320
- "Data Consistency Assessment Function (DCAF)", Y.Chen, J.Dodson, P.Avitabile, *Proceedings of the Thirty-Eighth International Modal Analysis Conference*, Houston, Texas, Feb 2020
- "Data Consistency Assessment Function (DCAF)", Y Chen, P Avitabile, J Dodson, Mechanical Systems and Signal Processing (2020), Mechanical Systems and Signal Processing 141 (2020) 106688

References to papers using examples presented in this paper

- "Use of Zernike Polynomials for Modal Vector Correlation of Small Turbine Blades", J.Salerno, P.Avitabile, *Proceedings of the Thirty-Second International Modal Analysis Conference*, Orlando, FL, Feb 2014
- "Non-Model Based Expansion from Limited Points to an Augmented Set of Points Using Chebyshev Polynomials", Y.Chen, P Logan, P Avitabile, J Dodson, Experimental Techniques (2019): 1-23.
- "Appropriate Excitation for Extraction of Modes", C.Beale, D.Joffre, P.Avitabile, Proceedings of the Thirty-Seventh International Modal Analysis Conference, Orlando, Florida, Jan 2019
- "Enhanced Amplitude Weighted Assessment Function (EAWAF)", P.Avitabile, Proceedings of the Twenty-Third International Modal Analysis Conference, Orlando, Florida, Feb 2005
- "Strain Shape Expansion Applied to Optically Measured Data", T.Doveno, Z.Mao, D.Soine, P.Avitabile, Proceedings of the Thirty-Eighth International Modal Analysis Conference, Houston, Texas, Feb 2020
- "Study of Modal Vector Correlation using Various Techniques for Model Reduction", P.Avitabile, F.Pechinsky, *Proceedings of the Tenth International Modal Analysis Conference*, San Diego, California, Feb 1992
- "Study of Modal Vector Correlation using Various Techniques for Model Reduction", P.Avitabile, F.Pechinsky, *Proceedings of the Tenth International Modal Analysis Conference*, San Diego, California, Feb 1992
- "Comparison of Reduction vs. Expansion for Orthogonality Checks", P.Avitabile, T.Foster, J.O'Callahan, *Proceedings of the Twelfth International Modal Analysis Conference*, Honolulu, Hawaii, Feb 1994
- "Evaluation of DOF Correlation using the Coordinate Orthogonality Check", F.Pechinsky, P.Avitabile, *Proceedings of the Thirteenth International Modal Analysis Conference*, Nashville, Tennessee, Feb 1995

- "Coordinate Orthogonality Check (CORTHOG)", P.Avitabile, F.Pechinsky, *Proceedings of the Twelfth International Modal Analysis Conference*, Honolulu, Hawaii, Feb 1994
- "Understanding the Effects of DOF Contribution to Orthogonality Checks", F.Pechinsky, P.Avitabile, T.Foster, *Proceedings of the Eleventh International Modal Analysis Conference*, Orlando, Florida, Feb 1993

References to other online training materials

- Modal Space articles, P.Avitabile, https://www.uml.edu/research/sdasl/education/modal-space.aspx
- Structural Dynamics and Vibrations course (https://www.learnsv.com/)
- Experimental Modal Analysis course (https://www.learnsv.com/)

References related to software mention in paper

- FEMAP – Finite Element Modeling and Postprocessing, Version 10.0.2, UGS Corp.
- FEMtools 3.8 – Dynamic Design Solutions, Leuven, Belgium.
- MATLAB - Matrix Analysis Software, The MathWorks, Inc., Natick MA
- Siemens/LMS Test.LAB, Leuven, Belgium

Chapter 4
Anti-Resonance-Based Surrogate Modelling for Stochastic FRF Prediction of Finite Element Models

Enora Denimal Goy

Abstract Quantifying uncertainties in dynamic systems is crucial to ensure the safety and integrity of structures. When addressing real-world problems, numerous uncertain parameters exist, and the computational cost of determining the stochastic response increases significantly, potentially becoming unaffordable. This is particularly true when predicting the frequency response function (FRF). Surrogate models are a well-known strategy for reducing this numerical cost. However, direct application to the frequency response function results in convergence issues at the resonance peaks. This study proposes an advanced numerical framework for constructing of the surrogate models of FRF based on anti-resonance properties. Building on recent preliminary results obtained on a 2 dof system, this work presents an extension for finite element models, where the antiresonances may appear and disappear. The proposed approach is applied to a finite element model of a cantilever beam. It is demonstrated that the convergence issues at FRF peaks are eliminated.

Keywords Frequency response function · Uncertainty propagation · Antiresonance· Gaussian process · Finite element model

Introduction

The presence of uncertainties highly impacts dynamic response of mechanical structures, so they must be considered in the design stage. However, the computation of the stochastic response can quickly become expensive when the number of uncertain parameters increases or when one wants to predict the the structure full frequency response function (FRF). In this context, surrogate models are a powerful tool as they construct a mathematical approximation of the response, which can then be evaluated at almost no numerical cost. For FRF prediction, direct approaches based on constructing one surrogate for the full FRF directly proved to experience convergence issues at the resonance peaks [1]. In a recent work, a strategy where the FRF is cut in intervals based on numerical properties to construct one surrogate per interval [2]. However, those approaches fail to deal with complex mechanical structures with more complex dynamical behaviour. In [3], preliminary results prove that exploiting the resonance and antiresonance properties of the mechanical structures to cut the FRF in different intervals to construct surrogate models on each interval was a promising strategy. The results were limited to discrete systems, where the existence of antiresonance is always ensured and doesn't depend on coordinates. The current work proposes to extend this work to Finite Element Models (FEM) where antiresonance existence depends on the model parameters, but also on the location on the structure. The work proposes to answer this problematic with a two-steps surrogate modelling strategy. In a first time, based on classification surrogate modelling, the existence of the antiresonance is assessed. It makes possible to cut the FRF in different clusters. In a second time, a regressive surrogate model is constructed on each cluster to predict the FRF amplitude. The method is illustrated on the 3D FEM of a beam and gives promising results for more complex test cases.

Enora Denimal Goy
Inria, CMAP, Ecole Polytechnique, Palaiseau, France
e-mail: enora.denimal-goy@inria.fr

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

Proposed methodology

General objective

Let's consider a linear dynamical system described by the following equation of motion:

$$\mathbf{M}\ddot{\mathbf{X}} + \mathbf{C}\dot{\mathbf{X}} + \mathbf{K}\mathbf{X} = \mathbf{F} \tag{1}$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are the mass, damping and stiffness matrices. $\mathbf{X}$ is the displacement vector, the dot denotes the derivative over time and $\mathbf{F}$ the excitation vector of the form $\mathbf{F}_e \cos \omega t$. The damping is supposed to be low. For such system, it is classical to study the Frequency Response Function (FRF), i.e. the response of the system when ω varies on an interval for each location, i.e. each dof. This work proposes a new and efficient strategy to predict the amplitude response at each dof, for any excitation frequency ω and when the system is uncertain (or depends on a set of parameters). The general workflow is explained first, and the different steps are then detailed.

General workflow

In the following, the uncertain parameters are denoted ξ and x denotes the coordinates of the point at which the FRF is computed. $\omega_R^{(i)}$ and $\omega_{AR}^{(j)}$ denote the i-th and j-th resonance and antiresonance, respectively. In the following, the surrogate of a quantity of interest z is referred with a: $\hat{z}$.

- Step 1: the surrogate models of the resonances $\hat{\omega}_R^{(i)}(\xi)$ are constructed using a GPR and a training set $\left\{ \left(\xi_k, \omega_R^{(i,k)} \right)_{k \in [1,N]} \right\}$.
- Step 2: the surrogate models of the antiresonances $\omega_{AR}^{(j)}(x, \xi)$ are constructed using a GPC (to predict the existence) and a GPR (to predict the pulsation frequency), using a training set $\left\{ \left((x_k, \xi_k), \omega_{AR}^{(i,k)} \right)_{k \in [1,N]} \right\}$.
- Step 3: a training set of the FRF is generated. It is composed of N_{FRF} input points $((\omega_k, x_k, \xi_k))_{k \in [1, N_{FRF}]}$ and their evaluation $\mathbf{y}_k$.
- Step 4: for each of this point, the resonance and antiresonance frequencies at (x_k, ξ_k) are computed. Comparing ω_k to those frequencies allow to attribute a cluster to the point.
- Step 5: for each cluster, a surrogate model of the FRF is constructed.

Antiresonance and resonance frequencies computation

The antiresonance at dof i can be obtained by resolving a modified eigenvalue problem where the i-th column of the sitffness matrix is replaced by the force vector, and where the i-th column of the mass matrix is replaced by a vector of $\mathbf{0}$ [4]. This strategy is employed in Step 2 to generate the training set of the antiresonance surrogates. For the resonance, a classical eigenvalue analysis is performed to obtain the natural frequencies of the system.

Gaussian process regression and classification

GPs are used for the creation of the different surrogate models. For GPR, the data are assumed to be generated by a process $Y_i = f(X_i) + \epsilon_i$ with $i \sim \mathcal{N}(0, \sigma^2)$. From a training set of N points (X_i, Y_i), a GP is constructed with the covariance function κ (which depends on hyperparameters θ) and the mean function m. It defines the matrix $K_{ij} = \kappa(\theta; X_i, X_j)$ and the vector $\mu = (f(X_1), \ldots, f(X_N))$. The hyperparameters θ are optimised to minimise the likelihood. Then, the prediction at a new points X_* is $\mu_* = K_*^T K^{-1} f$ with K_* the covariance matrix between X_* and the training set.

GPC are used for binary classification, to predict the existence or not of the different antiresonances. It is constructed with a training set of inputs $\mathbf{Z} = [\mathbf{z_1}, \ldots, \mathbf{z_N}]$ and outputs $\mathbf{y} = [\mathbf{y_1}, \ldots, \mathbf{y_N}]$ with $y_k \in \{0, 1\}$, which defines the class of membership of $\mathbf{z_k}$. The prediction of the membership of a new point is done by considering a latent function f approximated by a GP and mapped on the unitary interval with a sigmoid function.

The GPs are constructed with the GPflow python package [5].

Application on beam Finite Element Model

Model presentation

The model studied is a circular beam of radius R and length L simply supported at each end. It is shown in Figure 1. The Young modulus is 2e1 Pa, the shear modulus is 7.1e10 Pa, the density is 7800 kg/m^3, the Poisson ratio is 0.3, the radius of the cross section is 0.05 m, the length is 1 m. A Rayleigh damping is used with parameters $\alpha = 0.66$ and β =1.2e-6. The beam is discretised in 30 finite elements with four dof at each node. The excitation is located at dof 5 and is applied in both vertical and horizontal directions. The excitation amplitude is $1/\sqrt{2}$, and the frequency range of interest is $[100, 2500]$ rad/s. In the following, the Young modulus is considered as uncertain and can vary between $[1.75, 2.25\text{e}11$ Pa.

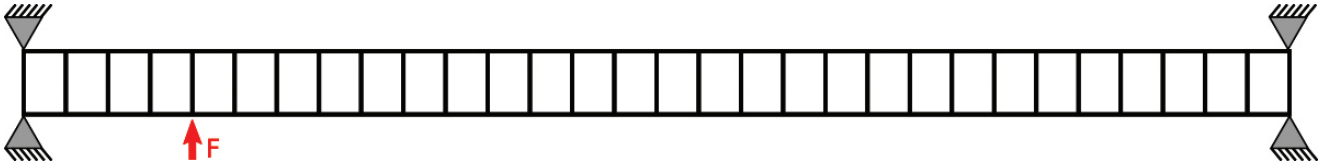
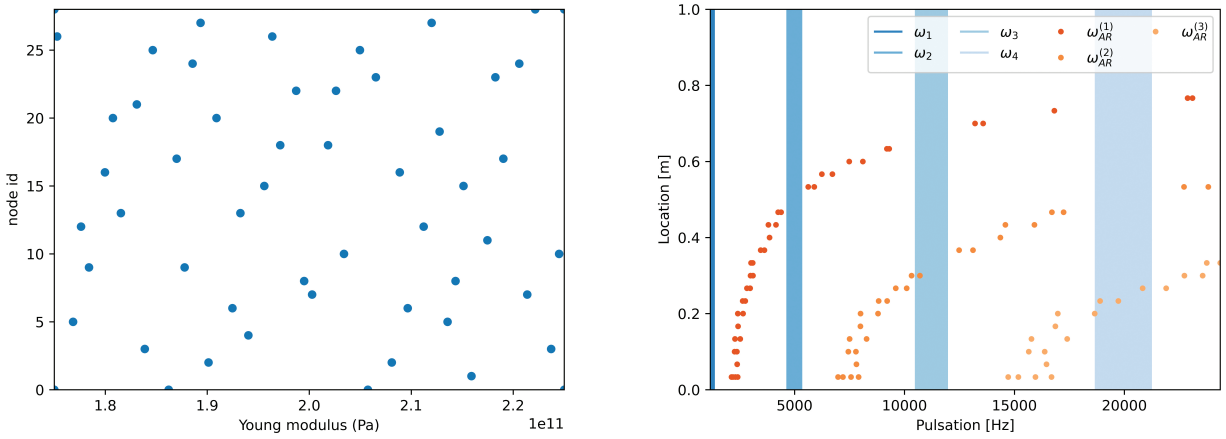


Fig. 1 Model under study

Results

In a first step, the training set for the resonance and antiresonance frequencies training are constructed. As an illustration, an input set of size 50 is shown in Figure 2(a). It is generated with a Halton sequence. For each of these points, the resonance and antiresonance frequencies are computed. They are displayedin Figure 2(b). As expected, the antiresonance frequencies depend on the location, and so are not always defined (e.g. near $L = 1$, there is no antiresonance).



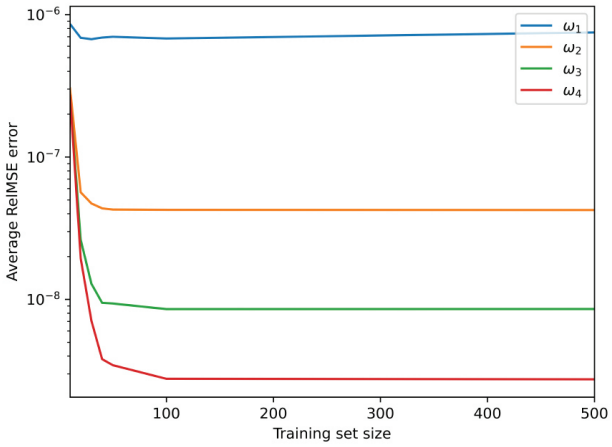
(a) Thrust dimensionless coefficient variation in inertial range

(b) Thrust dimensionless coefficient variation in stiffness range

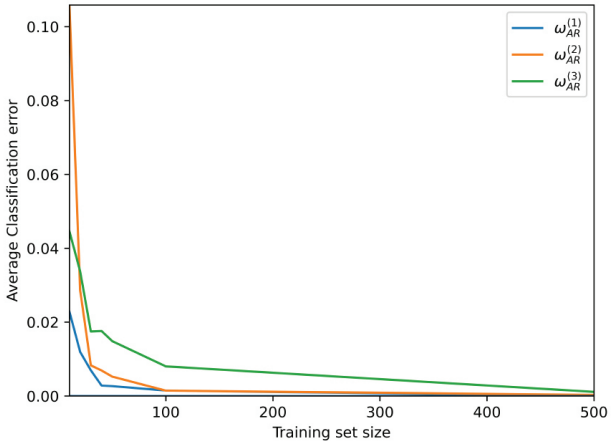
Fig. 2 Example of training set: input set of 50 points (left) and corresponding resonance and antiresonance frequencies (right)

GPR are constructed for each resonance frequencies, and for different size of training set. The latter can vary from 10 to 500 points, and for each size, 20 sets are generated. Each time, the relMSE error is computed using a validation set of 10000 points, and averaged over the 20 sets. The evolution of this error is given in Figure 3(a). One can see that the errors are always very low and that the convergence is quickly reached.

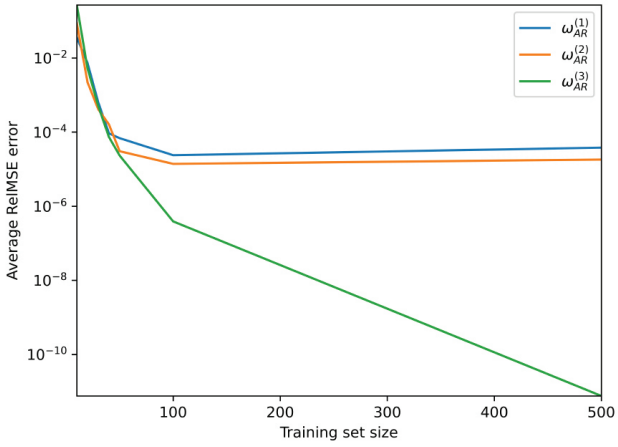
For the antiresonance frequencies, a first GPC is constructed to predict the existence or not the corresponding antiresonance. The error is computed as the average number of points misclassified on a validation set of 10000 points. Similarly to the resonance, training set size varies from 10 to 500 points, and 20 sets of each size are considered. The average error on



(a) Average error on resonance frequency prediction



(b) Average classification error on antiresonances

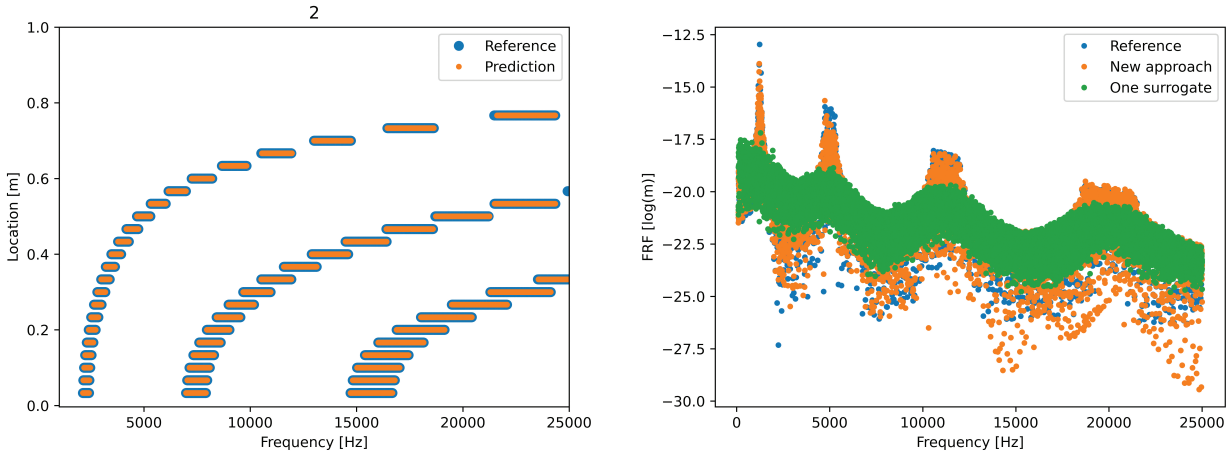


(c) Average regression error on antiresonances

Fig. 3 Evolution of the regression error on the resonance frequencies (a) and of the classification (b) and regression (c) error on the antiresonance frequencies vs the size of the training set

the 20 sets is computed and displayed in Figure 3(b). One can see that low classification error are reached with a low number of training points. Moreover, the convergence is quicker on the two first antiresonances. Finally, similarly to the resonances, the regression error is computed on the prediction of the antiresonance frequencies. The evolution is given in Figure 3(c). The error reached are low even with a low number of training points. As a validation, the predictions are compared to values from the validation set in Figure 4(a) for the three antiresonances. The surrogate used are constructed with 74 training points. Once can see that for each of them, prediction looks are very accurate.

Finally, the surrogate of the full FRF is constructed. The full FRF is divided in clusters for which boundaries are marked with the resonance and antiresonance frequencies (i.e. cut at the peaks). One GPR is constructed on each cluster. This strategy is compared to a strategy where all points from the training set are used to create a unique surrogate of the FRF. A validation set of 10000 points is validated and the average absolute error is computed, and with the proposed approach, the error of prediction is divided by 6 compared to the naïve approach. As a validation, the predicted FRF with both approaches are displayed in Figure 4(b), where they are compared to the reference values. Once can see that the proposed approach outperforms the one-surrogate approach as resonance peaks and anti-resonance peaks are well approximated. With the proposed approach and at high frequencies, the amplitude of the FRF near antiresonance is underestimated. This is due to some misclassification, which pushes the surrogate in extrapolation and bad prediction. This can be circumvented by adding more points to train the classification surrogate as small mistake there can have dramatic consequences.



(a) Antiresonance frequency predictions

(b) FRF prediction

Fig. 4 (a) Comparison of the reference and prediction antiresonance frequencies and (b) Comparison of the reference and predicted FRF with one surrogate and the new approach

Conclusion

The present study has for objective to propagate uncertainty for FRF prediction of linear mechanical system. A new strategy is proposed by cutting the training set in subsets based on physical properties, namely resonance and antiresonance, and creating a surrogate for each of this subset. The computation of the antiresonance is done with a modified eigenvalue analysis. As antiresonance frequencies are not always defined, their existence is learned with a classification surrogate and traditional GPR are used to predict their frequency. The strategy is compared to the case where a unique surrogate model is constructed to predict the full FRF. This is applied to a FEM of a beam, and the results proved the efficiency of the approach as lower error level and quicker convergence are observed.

Acknowledgments The author acknowledges the ANR – FRANCE (French National Research Agency) for its financial support of the ANR JCJC MeMoRa project n°ANR-23-CE51-0006.

References

1. Jacquelin, E., Adhikari, S., Sinou, J.J., and Friswell, M.I. “Polynomial Chaos Expansion and Steady-State Response of a Class of Random Dynamical Systems”. *Journal of Engineering Mechanics*, 141(4):04014145 (2015)ifnextchar.gobble.
2. Yaghoubi, V., Marelli, S., Sudret, B., and Abrahamsson, T. “Sparse polynomial chaos expansions of frequency response functions using stochastic frequency transformation”. *Probabilistic Engineering Mechanics*, 48:39–58 (2017)ifnextchar.gobble.
3. Goy, E.D. “A surrogate modelling approach based on anti-resonance properties for frf prediction of uncertain dynamical systems”. In *International Modal Analysis Conference - IMAC*, Orlando, FL (2025)ifnextchar.gobble. Society for Experimental Mechanics.
4. Jung, J., Goo, S., and Kook, J. “Predicting anti-resonance frequencies using a novel eigenvalue formulation”. *Finite Elements in Analysis and Design*, 191:103525 (2021)ifnextchar.gobble.
5. Matthews, A.G.d.G., Van Der Wilk, M., Nickson, T., Fujii, K., Boukouvalas, A., Le, P., Ghahramani, Z., Hensman, J., et al. “Gpflow: A gaussian process library using tensorflow”. *Journal of Machine Learning Research*, 18(40):1–6 (2017)ifnextchar.gobble.

Chapter 5
Reliability Analysis of Nonlinear Dynamic Systems via Importance Sampling and Surrogate Modeling

Zhi-wei Wang, Michael D. Todd, and Zhen Hu

Abstract Reliability analysis of nonlinear dynamic systems under uncertainty remains a critical yet challenging task. Traditional methods, such as the Monte Carlo simulation (MCS) and analytical approximations, often suffer from limited accuracy and high computational cost, particularly when dealing with high-dimensional nonlinear dynamic systems. This study develops a surrogate-assisted importance sampling (IS) framework for reliability analysis of nonlinear dynamical systems subjected to stochastic excitations. The proposed method employs an attention long short-term memory (aLSTM) feature extractor combined with a spectral-normalized neural Gaussian process (SNGP) to form an uncertainty-aware surrogate model (aLSTM-SNGP). This model directly predicts the structural response and its uncertainty, thereby avoiding costly nonlinear time-domain simulations. A particle filtering approach is introduced to approximate the optimal proposal distribution, and an active learning strategy with fine-tuning is applied to iteratively improve surrogate accuracy in critical failure regions. Numerical studies on nonlinear shear-building models under stochastic seismic excitations demonstrate that the framework provides accurate failure probability estimates with greatly reduced computational cost compared to the MCS method. The results highlight the potential of integrating advanced surrogate modeling and adaptive sampling strategies for reliability analysis of complex structural systems.

Keywords Reliability analysis · Nonlinear dynamic systems · Importance sampling · Surrogate modeling · Uncertainty quantification

Introduction

The reliability analysis of nonlinear dynamic structural systems under stochastic excitations is a fundamental problem in structural engineering. Conventional Monte Carlo simulation (MCS) provides unbiased estimates of failure probability but suffers from prohibitive computational costs when dealing with rare-event probabilities and nonlinear dynamic responses [1, 10]. To improve efficiency, importance sampling (IS) has been widely adopted as it reallocates sampling effort toward critical failure regions while retaining unbiasedness [2, 9]. However, the efficiency of IS critically depends on constructing an effective proposal distribution, which remains challenging in high-dimensional and strongly nonlinear settings [6, 11].

To mitigate these challenges, surrogate models have been increasingly integrated into IS frameworks. Polynomial chaos expansions [5], Gaussian process regression [8], and Kriging-based meta-modeling [4] have been employed to approximate

Zhi-wei Wang
Research Fellow, Department of Industrial and Manufacturing Systems Engineering, University of Michigan-Dearborn, 4901 Evergreen Rd, Dearborn, MI 48128, USA
e-mail: wzhiwei@umich.edu

Michael D. Todd
Professor, Department of Structural Engineering, University of California San Diego, 9500 Gilman Dr, La Jolla, CA 92093, USA
e-mail: mdtodd@ucsd.edu

Zhen Hu
Associate Professor, Department of Industrial and Manufacturing Systems Engineering, University of Michigan-Dearborn, 4901 Evergreen Rd, Dearborn, MI 48128, USA
e-mail: zhennhu@umich.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

computationally expensive dynamic simulations. More recently, machine-learning-based surrogates such as NARX models [12], hybrid Kriging–NARX frameworks [3], and physics-informed neural networks [14] have been proposed to capture nonlinear system behavior with improved generalization. Deep learning models, including recurrent and convolutional architectures, have also demonstrated potential in approximating nonlinear dynamic responses under seismic or wind loading [13, 15]. Despite these advances, existing surrogates often face difficulties in maintaining accuracy near failure boundaries, where reliability estimates are most sensitive. In recent years, efforts have focused on adaptive and uncertainty-aware surrogate models. For example, active learning strategies have been used to enrich surrogate training datasets near critical regions, while Gaussian process (GP) models with uncertainty quantification have provided confidence-aware reliability estimates [7, 16]. Nevertheless, balancing surrogate accuracy, uncertainty quantification, and computational efficiency remains a pressing issue.

This study contributes to this field by proposing a surrogate-assisted IS framework tailored for nonlinear dynamic reliability analysis. The framework integrates an uncertainty-aware surrogate model with an adaptive sampling strategy, aiming to enhance accuracy near the failure boundary while retaining computational efficiency. A nonlinear shear-building model under stochastic seismic excitations is employed to demonstrate the effectiveness of the proposed method.

Methodology

IS-based Reliability Analysis for Nonlinear Dynamic Systems

In the reliability analysis of nonlinear dynamical systems, the structural response is obtained from a simulation model $\mathcal{M}$, which maps the stochastic dynamic excitation $\mathbf{s}$ (e.g., ground motions, wind loads) and random structural parameters $\boldsymbol{\theta}$ (e.g., stiffness, damping, and material properties) to the dynamic response time history $\mathbf{y}(t)$. Specifically, $\mathbf{y}(t) = \mathcal{M}(\mathbf{s};\boldsymbol{\theta})(t)$. Since the extreme responses are of interest, the performance (limit state) function is defined based on the maximum absolute value of the response $y_{\max}(\mathbf{s},\boldsymbol{\theta})$ within the time horizon $[0,T]$, namely,

$$g(\mathbf{s},\boldsymbol{\theta}) = d - y_{\max}(\mathbf{s},\boldsymbol{\theta}) = d - \sup_{t\in[0,T]} \big|\,\mathbf{y}(t)\,\big|, \tag{1}$$

where d is a prescribed threshold.

A straightforward approach to estimate the failure probability is the crude MCS, which draws samples $(\mathbf{s}^{(i)},\boldsymbol{\theta}^{(i)})$ from their true distributions and evaluates the failure probability, namely,

$$\widehat{P}_f^{\mathrm{MCS}} = \frac{1}{N}\sum_{i=1}^{N} I\Big[g\Big(\mathbf{s}^{(i)},\boldsymbol{\theta}^{(i)}\Big)\le 0\Big], \quad \mathbf{s}^{(i)}\sim f_{\mathbf{S}}(\mathbf{s}),\ \boldsymbol{\theta}^{(i)}\sim f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}). \tag{2}$$

where $I[\cdot]$ is the indicator function; $f_{\mathbf{S}}$ and $f_{\boldsymbol{\Theta}}$ are the probability density functions of the stochastic dynamic excitation and random structural parameters, respectively.

However, when the event of interest is rare, MCS requires an excessively large number of samples to obtain accurate estimates, leading to prohibitive computational costs in dynamic analyses. To overcome this limitation, the IS method is introduced here. The key idea is to sample from an auxiliary distribution that places more weight in regions contributing most to failure. In our formulation, the dynamic excitation samples are still drawn from their true distribution $f_{\mathbf{S}}(\mathbf{s})$, while the structural parameters are sampled from a proposal distribution $h(\boldsymbol{\theta})$. The failure probability can then be reformulated as

$$P_f = \int_{\mathbb{R}^{d_s}}\int_{\mathbb{R}^{d_\theta}} I[g(\mathbf{s},\boldsymbol{\theta})\le 0]\ \frac{f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})}{h(\boldsymbol{\theta})}\, f_{\mathbf{S}}(\mathbf{s})\, h(\boldsymbol{\theta})\, d\mathbf{s}\, d\boldsymbol{\theta}. \tag{3}$$

Analogizing with the calculation formula of $\widehat{P}_f^{\mathrm{MCS}}$, the IS estimator of the failure probability is given by

$$\widehat{P}_f^{\mathrm{IS}} = \frac{1}{N}\sum_{i=1}^{N} I\Big[g(\mathbf{s}^{(i)},\boldsymbol{\theta}^{(i)})\le 0\Big]\ \frac{f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}^{(i)})}{h(\boldsymbol{\theta}^{(i)})}, \quad \mathbf{s}^{(i)}\sim f_{\mathbf{S}}(\mathbf{s}),\ \boldsymbol{\theta}^{(i)}\sim h(\boldsymbol{\theta}). \tag{4}$$

By choosing a suitable proposal density $h(\boldsymbol{\theta})$, importance sampling greatly improves the efficiency of failure probability estimation compared to crude MCS, especially for rare-event problems in dynamic reliability analysis. Despite its advantages, the practical application of IS in the reliability analysis of nonlinear dynamic systems still faces several key challenges: (i) strong nonlinearity of structural response; (ii) high-dimensional random inputs; (iii) heavy computational burden; and (iv) design of the proposal distribution.

Surrogate-assisted IS Reliability Analysis Framework

aLSTM-SNGP Surrogate Model

To address the above challenges, we developed a surrogate-assisted IS reliability analysis framework based on the proposed aLSTM-SNGP model. A surrogate modeling strategy is adopted to approximate the complex input–output mapping of the nonlinear dynamic systems. Specifically, we develop a hybrid model termed the aLSTM-SNGP model, as illustrated in Fig. 1, which serves as an efficient and uncertainty-aware surrogate $\mathfrak{S}(\mathbf{s},\boldsymbol{\theta})$ to the original simulation model $\mathcal{M}(\mathbf{s};\boldsymbol{\theta})$. The proposed surrogate model $\mathfrak{S}$ is directly trained to approximate the maximum absolute response $y_{\max}$ and to provide an associated uncertainty measure for its prediction, that is,

$$\sup_{t\in[0,T]}\left|\mathcal{M}(\mathbf{s};\boldsymbol{\theta})(t)\right| \;\approx\; \mathfrak{S}(\mathbf{s},\boldsymbol{\theta}) \;=\; \left[\,\hat{y}_{\max}(\mathbf{s},\boldsymbol{\theta}),\ \text{uncertainty}\,\right]. \tag{5}$$

The input of the aLSTM-SNGP model consists of the stochastic dynamic excitation $\mathbf{s}$ and random structural parameters $\boldsymbol{\theta}$. These inputs are first processed by an attention Long Short-Term Memory (aLSTM) network, which acts as a feature extractor. By compressing dynamic excitation histories and numerous structural parameters into compact latent representations, the aLSTM implicitly performs dimension reduction, which alleviates the curse of dimensionality that commonly arises in nonlinear dynamical reliability analysis. The Spectral-normalized Neural Gaussian Process (SNGP) module is employed as the probabilistic prediction component of the aLSTM-SNGP surrogate model. It receives compact feature representations from the aLSTM encoder and processes them through several residual layers with spectral normalization. These layers are followed by a random Fourier feature (RFF) mapping that approximates kernel functions. At the output stage, a neural Gaussian process (GP) layer is attached, which replaces the conventional dense output with a probabilistic predictor. Unlike a standard deterministic output, this layer produces not only the predictive mean $\hat{\mu}_{\max}$ but also the predictive uncertainty $\hat{\sigma}_{\max}$, effectively quantifying epistemic uncertainty in the model outputs. Formally, the surrogate model $\mathfrak{S}$ provides both the predictive mean estimate of the maximum absolute response $y_{\max}(\mathbf{s},\boldsymbol{\theta})$ and the associated predictive uncertainty:

$$\hat{y}_{\max}(\mathbf{s},\boldsymbol{\theta}) \;=\; \mathfrak{S}(\mathbf{s},\boldsymbol{\theta}) \;\sim\; \mathcal{N}\Big(\hat{\mu}_{\max}(\mathbf{s},\boldsymbol{\theta}),\ \hat{\sigma}^2_{\max}(\mathbf{s},\boldsymbol{\theta})\Big). \tag{6}$$

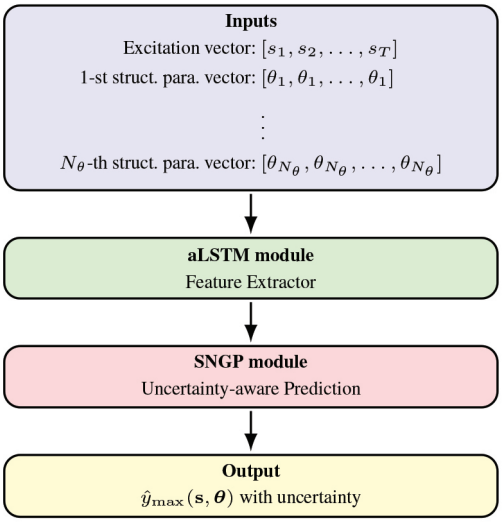


Fig. 1 Overall architecture of the proposed aLSTM-SNGP surrogate model

Particle Filtering for Approximating Proposal Distribution

A central difficulty in implementing IS for nonlinear dynamical reliability analysis lies in designing an effective proposal distribution $h(\boldsymbol{\theta})$. The distribution must concentrate sampling effort in regions that contribute significantly to the failure probability. We employ a surrogate-assisted particle filtering strategy, which adaptively updates the sampling distribution based on the surrogate model predictions. Specifically, we first draw a set of prior samples (particles) of structural parameters, $\mathcal{X}_{\text{prior}} = \{\boldsymbol{\theta}^{(i)}_{\text{prior}}\}_{i=1}^{N}$, $\boldsymbol{\theta}^{(i)}_{\text{prior}} \sim f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$, together with a pre-generated library of dynamic excitations $\mathcal{E} = \{\mathbf{s}^{(i)}\}_{i=1}^{N}$, $\mathbf{s}^{(i)} \sim$

$f_{\mathbf{S}}(\mathbf{s})$. For each particle pair $(\mathbf{s}^{(i)}, \boldsymbol{\theta}_{\text{prior}}^{(i)})$, the likelihood is evaluated through the surrogate model $\mathfrak{S}$. This likelihood is defined as:

$$\varphi(\mathbf{s}, \boldsymbol{\theta} \mid \mathfrak{S}) \;=\; \Phi\left(\frac{\hat{\mu}_{\max}(\mathbf{s}, \boldsymbol{\theta}) - d}{\hat{\sigma}_{\max}(\mathbf{s}, \boldsymbol{\theta})}\right), \tag{7}$$

where $\hat{\mu}_{\max}(\mathbf{s}, \boldsymbol{\theta})$ and $\hat{\sigma}_{\max}(\mathbf{s}, \boldsymbol{\theta})$ are the surrogate-predicted mean and standard deviation of the maximum absolute response, and d is the failure threshold. This quantity reflects the surrogate-informed conditional failure probability of a particle.

The normalized Bayesian weight of each particle is then defined as:

$$w(\mathbf{s}^{(i)}, \boldsymbol{\theta}_{\text{prior}}^{(i)} \mid \mathfrak{S}) \;=\; \frac{\varphi(\mathbf{s}^{(i)}, \boldsymbol{\theta}_{\text{prior}}^{(i)} \mid \mathfrak{S})}{\sum_{j=1}^{N} \varphi(\mathbf{s}^{(j)}, \boldsymbol{\theta}_{\text{prior}}^{(j)} \mid \mathfrak{S})}, \tag{8}$$

ensuring that particles closer to the failure region receive higher importance. By resampling according to these weights, the particle set gradually concentrates in the critical domain, resulting in the posterior set $\mathcal{X}_{\text{post}} = \{\boldsymbol{\theta}_{\text{post}}^{(i)}\}_{i=1}^{N}$ of structural parameters.

Next, an analytic approximation of the proposal distribution, denoted as $h^*(\boldsymbol{\theta} \mid \mathfrak{S})$, is constructed from $\mathcal{X}_{\text{post}}$, for example by fitting a Gaussian mixture model (GMM). This approximate distribution serves as a practical substitute for the unknown optimal $h(\boldsymbol{\theta})$. Drawing from this distribution yields a set of important samples of structural parameters, namely, $\mathcal{X}_{\text{IS}} = \{\boldsymbol{\theta}_{\text{IS}}^{(i)}\}_{i=1}^{N_{\text{IS}}}, \boldsymbol{\theta}_{\text{IS}}^{(i)} \sim h^*(\boldsymbol{\theta} \mid \mathfrak{S})$. Based on Eq. (4), the IS estimator is then evaluated as:

$$\hat{P}_f^{\text{IS}} \;=\; \frac{1}{N_{\text{IS}}} \sum_{i=1}^{N_{\text{IS}}} I\big[\hat{g}(\mathbf{s}^{(i)}, \boldsymbol{\theta}_{\text{IS}}^{(i)} \mid \mathfrak{S}) \le 0\big] \, \frac{f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}_{\text{IS}}^{(i)})}{h^*(\boldsymbol{\theta}_{\text{IS}}^{(i)} \mid \mathfrak{S})}, \qquad \boldsymbol{\theta}_{\text{IS}}^{(i)} \in \mathcal{X}_{\text{IS}}, \;\; \mathbf{s}^{(i)} \in \mathcal{E}, \tag{9}$$

where $\hat{g}(\mathbf{s}, \boldsymbol{\theta} \mid \mathfrak{S}) = d - \hat{y}_{\max}(\mathbf{s}, \boldsymbol{\theta}) = d - \hat{\mu}_{\max}(\mathbf{s}, \boldsymbol{\theta})$ is the surrogate-based limit state function.

Training Strategy for aLSTM-SNGP Model

The surrogate model $\mathfrak{S}$ trained using a limited initial dataset is often not sufficiently accurate in the critical failure regions (the initial dataset is generated by evaluating the simulation model $\mathcal{M}$ and denoted as $\mathcal{D}_0 = \{[(\mathbf{s}_i, \boldsymbol{\theta}_i),\ y_{\max,i}]\}_{i=1}^{N_0}, \mathbf{s}_i \in \mathcal{E}$, $\boldsymbol{\theta}_i \in \mathcal{X}_{\text{prior}}$, where the size N_0 is typically small, e.g., 120). Consequently, the approximated proposal distribution $h^*(\boldsymbol{\theta} \mid \mathfrak{S})$ and the corresponding IS estimates may lack precision. Hence, iterative training is necessary to progressively improve the accuracy of $\mathfrak{S}$ in the failure regions and enhance the precision of the IS estimates.

Active Learning-based Sample Augmentation. We adopt an active learning strategy to adaptively refine the surrogate by selectively adding new training samples that are most informative for failure probability estimation. Candidate samples are formed as pairs $(\mathbf{s}, \boldsymbol{\theta}) \in \mathcal{E} \times \mathcal{X}_{\text{IS}}$. The informativeness of each candidate is measured by the following learning function:

$$\mathcal{L}(\mathbf{s}, \boldsymbol{\theta} \mid \mathfrak{S}) \;=\; \frac{\left|\hat{\mu}_{\max}(\mathbf{s}, \boldsymbol{\theta}) - d\right|}{\hat{\sigma}_{\max}(\mathbf{s}, \boldsymbol{\theta})}. \tag{10}$$

Instead of selecting a single new sample, a batch of N_{new} candidates (e.g., $N_{\text{new}} = 5$) with the smallest values of $\mathcal{L}(\mathbf{s}, \boldsymbol{\theta} \mid \mathfrak{S})$ are selected at each iteration:

$$\{(\mathbf{s}_{\text{next}}^{(j)}, \boldsymbol{\theta}_{\text{next}}^{(j)})\}_{j=1}^{N_{\text{new}}} = \underset{(\mathbf{s},\boldsymbol{\theta}) \in \mathcal{E} \times \mathcal{X}_{\text{IS}}}{\overset{\text{select } N_{\text{new}}}{\arg\min}} \; \mathcal{L}(\mathbf{s}, \boldsymbol{\theta} \mid \mathfrak{S}) \tag{11}$$

so that the surrogate model is refined in regions where the prediction is close to the failure boundary and the uncertainty remains high. By repeatedly enriching the training dataset with samples near the failure boundary, the surrogate model $\mathfrak{S}$ gradually improves its predictive accuracy in critical failure regions, thereby improving the quality of the approximated proposal distribution $h^*(\boldsymbol{\theta} \mid \mathfrak{S})$ and enhancing the reliability of IS-based failure probability estimation.

Hybrid Training with Fine-Tuning. In addition to the active learning strategy, we adopt a hybrid training strategy to efficiently update the surrogate model $\mathfrak{S}$ during the iterative refinement process. Specifically, after the full training performed on the initial dataset, the training process alternates between the following two modes:

(i) periodic full retraining, where every fixed number of iterations (e.g., every 5 iterations) the surrogate $\mathfrak{S}$ is reinitialized and trained from scratch on the current entire dataset, thereby preventing error accumulation and model drift; and

(ii) lightweight fine-tuning, where in the remaining iterations the surrogate $\mathfrak{S}$ is refined using the newly augmented samples (or their cumulative set) after the last full training, starting from the most recent set of model weights obtained from the last full training, which allows efficient assimilation of new information while significantly reducing training budget.

This procedure is conducted until convergence of $\hat{P}_f^{\text{IS}}$ is achieved. The above hybrid training strategy achieves a balance between stability (through full retraining) and efficiency (through fine-tuning) in updating the surrogate model, allowing $\mathfrak{S}$ to remain accurate in the critical failure regions without incurring prohibitive computational cost.

Case Study: Nonlinear Shear-Building Model under Stochastic Seismic Excitations

Nonlinear Shear-Building Model Introduction

To validate the proposed surrogate-assisted IS reliability analysis framework, a six-story nonlinear shear-building model subjected to stochastic seismic excitations is investigated, as illustrated in Fig. 2. The governing equations of motion for an n-story shear building are given by:

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \mathbf{f}_s(\mathbf{u}(t), \mathbf{z}(t)) = -\mathbf{M}\mathbf{1}\ddot{u}_g(t), \tag{12}$$

where $\mathbf{u}(t) \in \mathbb{R}^n$ denotes the vector of floor displacements relative to the ground, $\mathbf{M} \in \mathbb{R}^{n\times n}$ is the diagonal mass matrix, $\mathbf{C} \in \mathbb{R}^{n\times n}$ is the damping matrix, $\mathbf{f}_s(\mathbf{u}, \mathbf{z}) \in \mathbb{R}^n$ is the vector of nonlinear restoring forces, $\mathbf{z}(t) \in \mathbb{R}^n$ is the vector of hysteretic internal variables, and $\mathbf{1} \in \mathbb{R}^n$ is the vector of ones ensuring that all stories are subjected to the same ground acceleration input $\ddot{u}_g(t)$.

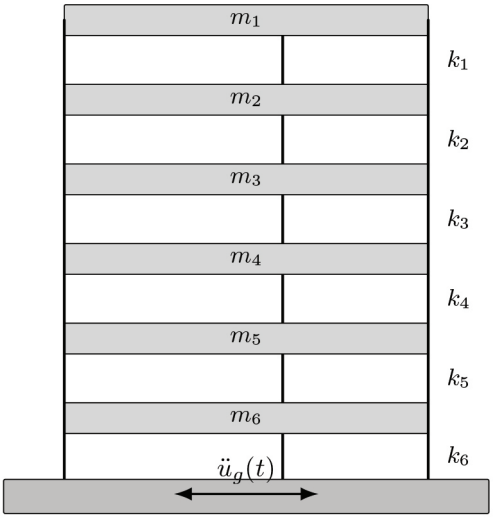


Fig. 2 Six-story nonlinear shear-building model under stochastic seismic excitation

The restoring force of the i-th story is modeled by the Bouc–Wen hysteretic model:

$$f_{s,i}(t) = k_i\big(\alpha u_i(t) + (1-\alpha)z_i(t)\big), \qquad i = 1, \ldots, n, \tag{13}$$

with the hysteretic evolution equation:

$$\dot{z}_i(t) = A\dot{u}_i(t) - \beta|\dot{u}_i(t)||z_i(t)|^{\eta-1}z_i(t) - \gamma\dot{u}_i(t)|z_i(t)|^{\eta}, \tag{14}$$

where k_i is the inter-story stiffness, α is the post-yield ratio, and (A, β, γ, η) are Bouc–Wen parameters governing hysteresis shape. In this study, they are set as $\alpha = 0.3$, $A = 1.0$, $\eta = 1.0$, and $\beta = \gamma = A/(2x_y^{\eta})$, and $x_y = 0.01\,\text{m}$ is the yield displacement.

In this case study, structural uncertainties are represented by the inter-story stiffness parameters $\{k_i\}_{i=1}^6$, which are modeled as independent Gaussian random variables with mean 4.5×10^7 N/m and standard deviation 9.0×10^6 N/m for each story. The lumped masses are deterministic and identical across all stories, each set to 3.4×10^4 kg. The performance measure for reliability analysis is defined as the maximum absolute top-floor displacement $y_{\max}$, and failure occurs if $y_{\max}$ exceeds a prescribed threshold d.

Results

In this example, the failure threshold $d = 0.060\,\mathrm{m}$ is considered. The aLSTM-SNGP surrogate model $\mathfrak{S}$ is trained initially on a dataset of $N_0 = 120$ samples. At each active learning iteration, a batch of $N_{\mathrm{new}} = 5$ new samples is selected. The training procedure alternates between full retraining every 5 iterations and fine-tuning in the intermediate iterations. For the surrogate-assisted IS reliability estimation at each iteration, 100,000 Monte Carlo samples are employed for posterior sampling and 100,000 importance samples are used in the IS estimator. The convergence of surrogate model training and failure probability estimate over successive iterations is shown in Fig. 3. The results demonstrate that the surrogate-assisted IS estimator converges to the MCS reference $\widehat{P}_f^{\mathrm{MCS}}$ within 50 iterations, using a total of 370 evaluations of the dynamic simulation model $\mathcal{M}$. The results confirm that the proposed method achieves accurate estimates with significantly reduced computational cost.

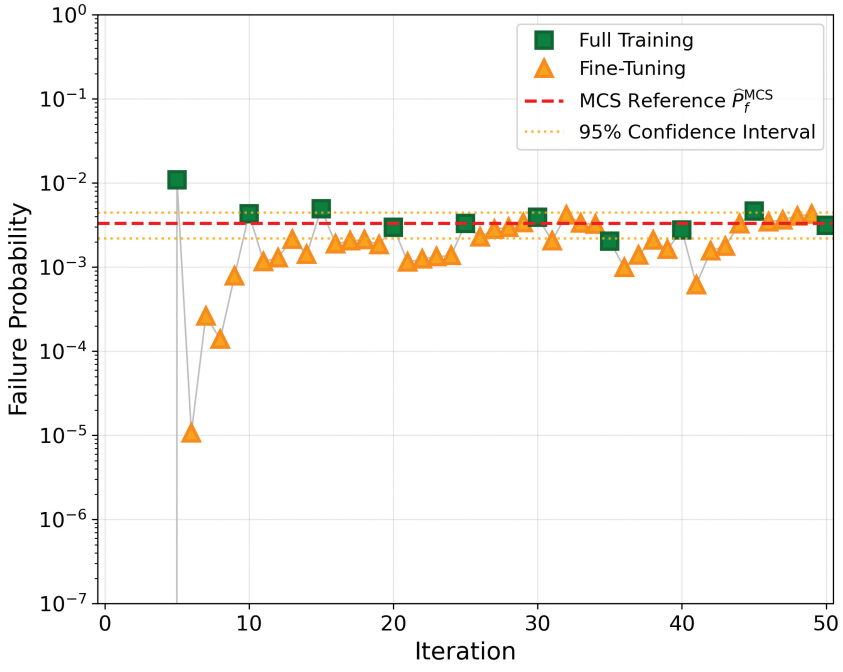


Fig. 3 Convergence of surrogate model training and failure probability estimate (Threshold $d = 0.060\,\mathrm{m}$)

To illustrate the approximation result of the proposal distribution $h(\boldsymbol{\theta})$, Fig. 4 shows the comparison of the prior distribution, the posterior distribution obtained from particle filtering, and the GMM-fitted proposal distribution for each stiffness parameter k_i. It can be clearly observed that the posterior distributions of the stiffness parameters k_i shift to the left compared to their priors, indicating smaller stiffness values. Physically, this corresponds to larger structural displacements under seismic excitations, which brings the system closer to the failure domain. This directional migration of the distributions confirms that the particle filtering process effectively concentrates the sampling effort in regions most relevant to failure events. Furthermore, the fitted proposal distribution provides a smooth and flexible representation that follows the posterior while reducing sampling variance, thereby ensuring both accuracy and efficiency in the IS estimator.

Figure 5 further illustrates the evolution of the surrogate model predictions on the important sample set drawn from the fitted proposal distribution across iterations. In the first iteration [Fig. 5(a)], the predictive variance $\hat{\sigma}_{\max}$ is relatively large, particularly near the failure threshold $d = 0.06\,\mathrm{m}$, indicating the limited accuracy of the initially trained surrogate in the critical region. After iterative active learning and fine-tuning, the variance is significantly reduced in the last iteration [Fig. 5(b)], especially in the vicinity of the failure boundary, which demonstrates the improved reliability of the surrogate for discriminating between safe and failure states. At the same time, the distribution of important samples is seen to migrate towards the failure boundary as iterations progress. This migration also reflects the fact that the surrogate-guided particle filtering adaptively allocates sampling effort to regions where the mean structural response $\hat{\mu}_{\max}$ approaches the failure threshold.

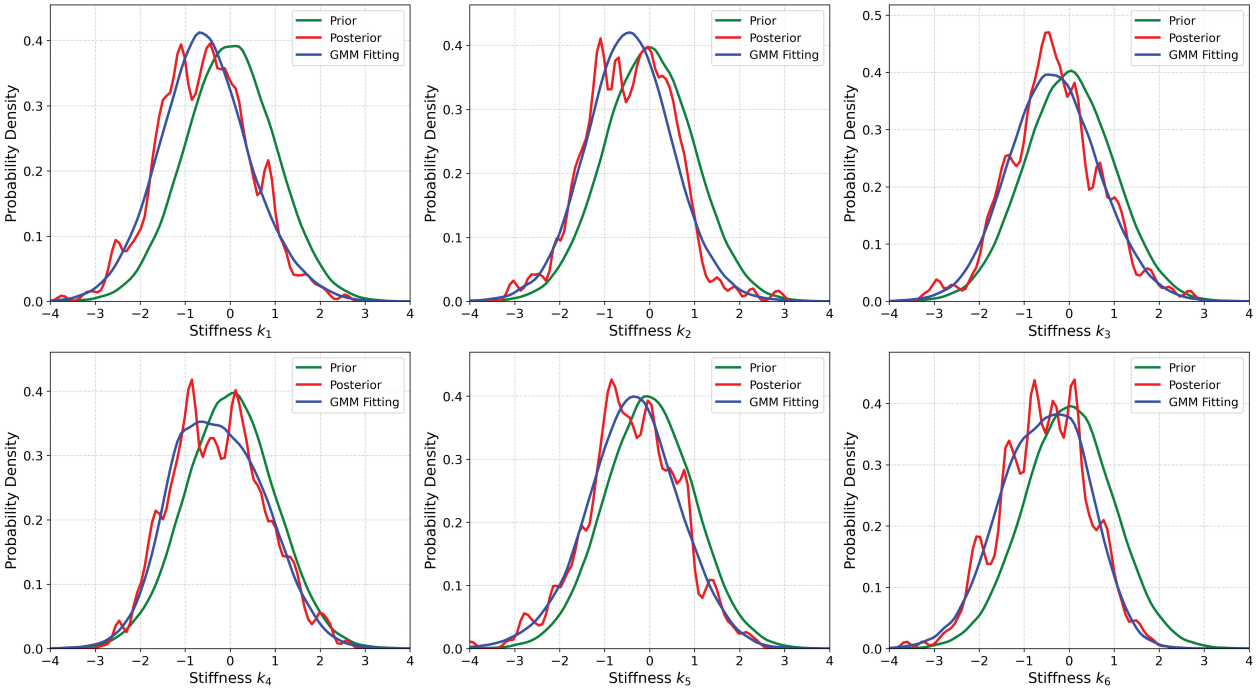


Fig. 4 Comparison of prior, posterior, and fitted proposal distributions of stiffness parameters at the last iteration

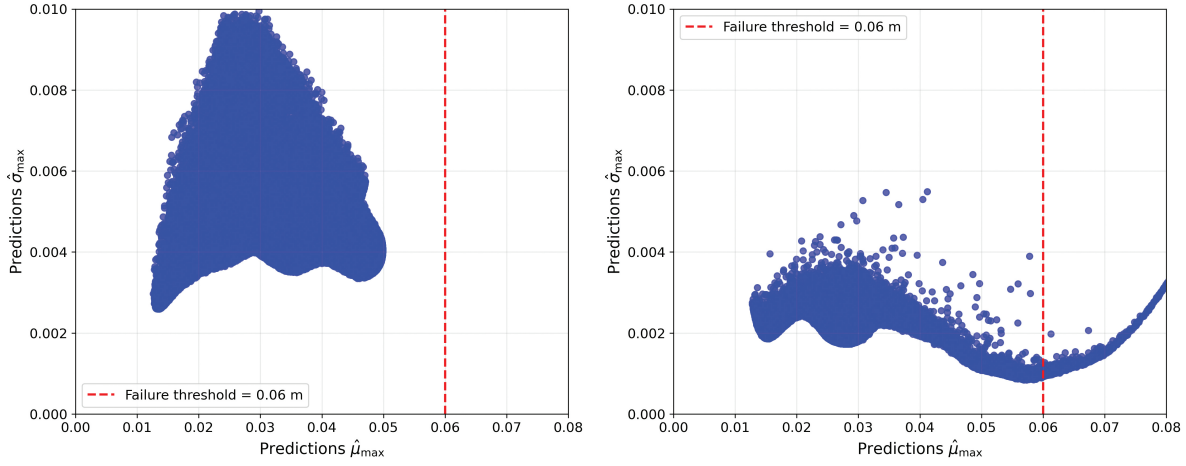


Fig. 5 The surrogate model predictions on the important sample set: (a) at the first iteration; and (b) at the last iteration

Conclusions

This work presents a surrogate-assisted importance sampling (IS) framework for efficient reliability analysis of nonlinear dynamic systems under stochastic excitations. The framework integrates three essential components: an aLSTM–SNGP surrogate model that predicts the maximum structural response together with predictive uncertainty, a particle filtering scheme that adaptively guides sampling toward regions relevant to failure, and an active learning strategy with hybrid training that incrementally improves surrogate accuracy. The numerical experiment on the six-story shear-building model confirms that the method can accurately estimate rare-event probabilities at a fraction of the cost of direct MCS. The study demonstrates that combining uncertainty-aware machine learning surrogates with adaptive sampling strategies offers a powerful tool for reliability analysis of dynamical systems, with potential applications to large-scale civil infrastructures under multi-hazard environments.

References

1. Au, S. K. and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277.
2. Au, S. K. and Beck, J. L. (2003). Subset simulation and its application to seismic risk based on dynamic analysis. *Journal of Engineering Mechanics*, 129(8):901–917.
3. Bhattacharyya, S., Ghosh, S., Roy, B., and Mahata, S. (2020). Hybrid kriging-narx model for nonlinear system reliability analysis. *Mechanical Systems and Signal Processing*, 135:106385.
4. Bichon, B. J., Eldred, M. S., Swiler, L. P., Mahadevan, S., and McFarland, J. M. (2008). Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA Journal*, 46(10):2459–2468.
5. Blatman, G. and Sudret, B. (2010). Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliability Engineering & System Safety*, 95(11):1216–1229.
6. Bucher, C. G. (1988). Adaptive sampling—an iterative fast monte carlo procedure. *Structural Safety*, 5(2):119–126.
7. Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, 10(3):273–304.
8. Echard, B., Gayton, N., and Lemaire, M. (2011). AK-MCS: An active learning reliability method combining kriging and monte carlo simulation. *Structural Safety*, 33(2):145–154.
9. Engelund, S. and Rackwitz, R. (1993). A benchmark study on importance sampling techniques in structural reliability. *Structural Safety*, 12(4):255–276.
10. Katafygiotis, L. S. and Cheung, S. H. (2005). Application of importance sampling in estimation of first-passage probabilities. *Probabilistic Engineering Mechanics*, 20(2):123–133.
11. Melchers, R. E. (1989). Importance sampling in structural systems. *Structural Safety*, 6(1):3–10.
12. Spiridonakos, M. D. and Chatzi, E. N. (2015). Metamodeling of dynamic nonlinear structural systems through polynomial chaos NARX models. *Computers & Structures*, 157:99–113.
13. Wang, H. and Wu, T. (2020). Knowledge-enhanced deep learning for wind-induced nonlinear structural dynamic analysis. *Journal of Structural Engineering*, 146(11):04020235.
14. Zhai, W., Tao, D., and Bao, Y. (2023). Parameter estimation and modeling of nonlinear dynamical systems based on Runge–Kutta physics-informed neural networks. *Nonlinear Dynamics*, 111:21117–21130.
15. Zhang, R., Liu, Y., and Sun, H. (2020). Physics-informed multi-LSTM networks for metamodeling of nonlinear structures. *Computer Methods in Applied Mechanics and Engineering*, 369:113226.
16. Zhao, Y., Jiang, C., Vega, M. A., Todd, M. D., and Hu, Z. (2023). Surrogate modeling of nonlinear dynamic systems: A comparative study. *Journal of Computing and Information Science in Engineering*, 23(1):011001.

Chapter 6
Adaptive Reduced Order Model Selection in Engineering Decision Support

Konstantinos Vlachas, Antonios Kamariotis, and Eleni Chatzi

Abstract Model-based approaches offer a structured, physics-informed framework to support engineering decisions related to the operation & maintenance of infrastructure assets. The core element of such frameworks are numerical models that distill complex systems into computationally tractable forms, while capturing essential dynamic behavior, variability and uncertainty. Within this context, Reduced Order Models (ROMs) facilitate real-time operations by reconciling accelerated model evaluations with accurate estimations. This balance can be achieved through techniques that embed physical insight into the ROM structure allowing for extrapolation and generalization capabilities. However, the utility of ROMs hinges on a delicate trade-off between computational savings and fidelity relative to Full Order Models (FOMs); a critical consideration when model outputs inform high-stakes decisions. In this work, we leverage real-time sensor data to propose an adaptive approach to dynamically select the appropriate model resolution, allowing the representation to evolve in step with the monitored system. First, a set of candidate ROMs is derived, each incorporating a generic parametrization to account for environmental and operational variability as well as for evolving phenomena, such as damage onset/progression or deterioration processes. Next, a sequential Bayesian inference framework is developed for quantification of the posterior uncertainty associated with each considered candidate ROM of a given resolution, thereby guiding the selection of the most suitable model representation. In turn, the generated ROMs are employed as predictive surrogates within the assembled inference framework, which assimilates monitoring information in a recursive manner. As a result, the selection among the generated models is posed as a problem of decision-making under uncertainty. The objective function is formulated to favor the resolution that maximizes the expected utility, defined as a trade-off between two competing criteria: i) precision in terms of recovering target quantities of interest and ii) computational efficiency measured by runtime. Subsequently, the selected resolution can be employed to support response prediction tasks and inform decision-making, while explicitly quantifying the associated uncertainty. The proposed integrated methodology enables adaptive model management by determining whether to accept the current model's prediction, switch to a different ROM instance or parameter realization, or trigger an on-the-fly retraining process based on real-time vibration measurements.

Keywords Reduced order models (ROMs) · Bayesian inference · Uncertainty quantification · Continuous data integration · Decision support

Introduction

A typical challenge in twining of structural infrastructure assets are their large dimensionality and the underlying complex dynamics that are characterized by stochasticity and uncertainty. Thus, for the derived virtual assets to deliver on their promise, they often rely on models that are both precise and computationally tractable [1]. Without loss of generality, in this work we focus on model-based techniques, and specifically Reduced Order Models (ROMs) [2]. These offer a structured and physics-based framework that provides better interpretability and more reliable extrapolation compared to purely data-driven

Konstantinos Vlachas · Eleni Chatzi
Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, Stefano-Franscini-Platz-5, 8093, Zurich, Switzerland
e-mail: vlachas@ibk.baug.ethz.ch; chatzi@ibk.baug.ethz.ch

Antonios Kamariotis
Hitachi Energy AG, Baden, Switzerland
e-mail: antoniskam@hotmail.com

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

surrogates when tasked to support engineering decisions related to the operation & maintenance (O&M) of infrastructure assets. This has been demonstrated via utilizing ROMs as forward simulators for an array of tasks that necessitate repeated model evaluations, such as fault detection [3], input estimation [4], or identification under uncertainty [5].

In this context, the serviceability of ROMs lies in navigating the balance between computational cost and fidelity relative to recovering target Quantities of Interest (QoI). In this work we tackle this challenge by proposing an adaptive framework that integrates monitoring data continuously to dynamically select the optimal model for the downstream task at hand. Existing hybrid-twinning methodologies that attempt to address such adaptations in changing environments typically rely on passive data assimilation [6, 7], or problem-specific approaches that use suitable indicators to trigger retraining or refinement steps [8, 9], thus limiting their adaptability to continuously evolving conditions. In short, any model adaptation or updating happens in a static and one-off manner and the updated model's fidelity is not weighed against operational constraints, leading to potentially sub-optimal representations. In contrast, our work introduces a value-of-information-driven model selection [10], which occurs in a continuous and online manner, enabling the recursive identification of the task-optimized model to be utilized. Thus, the proposed approach enables the recursive selection of the lowest-cost model that can deliver predictions of the required fidelity for the system of interest in the context of interest, maximizing the overall benefit at any given time [11].

The proposed implementation initiates by deriving ROMs of different fidelity levels. As mentioned, ROMs are employed given their ability to facilitate real-time operations by reconciling accelerated model evaluations with accurate estimations. The developed ROMs incorporate a generic parametrization scheme that accounts for environmental variability and evolving conditions like structural damage [12]. Next, the candidate ROMs are incorporated into a sequential Bayesian inference framework, which quantifies the posterior uncertainty associated with each considered model of a given resolution, thereby guiding the selection of the optimal representation [13]. The assimilation of sensing data happens recursively and in a continuous manner leading to an adaptive framework able to follow along the evolving asset. The corresponding selection process aims to maximize the expected utility by weighing two competing factors: how precisely we can recover our quantities of interest, and how efficiently we can compute them to maintain real-time feasibility. The developed integrated methodology is illustrated in Fig. 1 and is employed to decide whether to accept the hybrid model's prediction, switch to a different generated instance or parameter realization, or trigger a retraining event, thus enabling an adaptive treatment.

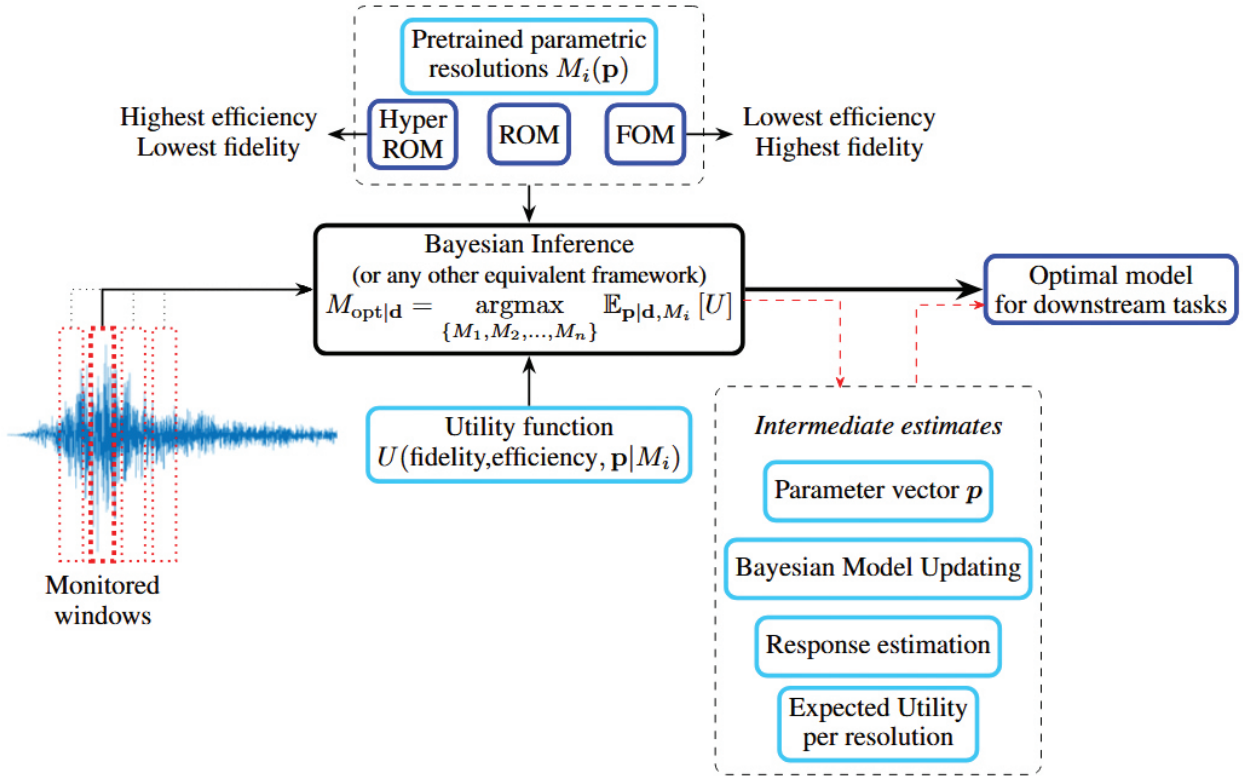


Fig. 1 Graphical abstract illustrating the proposed framework. The continuous monitoring and recursive model updating leads to an adaptive treatment

Problem Statement

Our work assumes availability of a Full-Order Model (FOM), which corresponds to a finite element (FE) model that is dependent on the input vector $\boldsymbol{p} = [p_1, ..., p_k]^{\mathrm{T}} \in \Omega \subset \boldsymbol{R}^k$, which captures all system- and excitation-relevant parameters. Thus, the response is described by the following set of governing equations:

$$\boldsymbol{M}\ddot{\boldsymbol{u}}(t) + \boldsymbol{g}\left(\boldsymbol{u}(t), \dot{\boldsymbol{u}}(t), \boldsymbol{p}\right) = \boldsymbol{F}(t, \boldsymbol{p}), \tag{1}$$

where $\boldsymbol{u}(t) \in \boldsymbol{R}^n$ denotes the response in terms of displacements, $\boldsymbol{M} \in \boldsymbol{R}^{n \times n}$ the mass matrix, and $\boldsymbol{F}(t, \boldsymbol{p}) \in \boldsymbol{R}^n$ the induced excitation. For the sake of simplicity, the parametric dependency in the mass matrix has been dropped. Nonlinear behavior is captured via the restoring force term $\boldsymbol{g}\left(\boldsymbol{u}(t), \dot{\boldsymbol{u}}(t), \boldsymbol{p}\right) \in \boldsymbol{R}^n$. The full-order dimension n represents the total number of degrees of freedom and indirectly expresses the computational efficiency of the model, as necessary operations scale with n.

VpROM: A conditional Variational Autoencoder (cVAE) boosted ROM

The reduction approach employed herein relies on a Galerkin projection scheme and the underlying assumption that the dynamic behavior, namely the solution of Eq. 1, lies in a low-order subspace of dimension r, with $r \ll n$. Therefore, the following holds:

$$\boldsymbol{u}\left(\boldsymbol{p}\right) \approx \boldsymbol{V}\left(\boldsymbol{p}\right)\boldsymbol{q} \tag{2}$$

where $\boldsymbol{V} \in \mathbb{R}^{N \times r}$ represents the projection basis and $\boldsymbol{q} \in \mathbb{R}^r$ is the low-order coordinate vector. Via substitution of $\boldsymbol{u}$ into Eq. 1 and after multiplying the governing set with $\boldsymbol{u}^T$, thus performing a Galerkin projection, the following can be derived:

$$\tilde{\boldsymbol{M}}\ddot{\boldsymbol{q}}\left(t\right) + \tilde{\boldsymbol{g}}\left(\ddot{\boldsymbol{q}}, \dot{\boldsymbol{q}}, \boldsymbol{p}\right) = \tilde{\boldsymbol{F}}\left(\boldsymbol{p}, t\right) \tag{3}$$

where $\tilde{\boldsymbol{M}} = \boldsymbol{V}^T\boldsymbol{M}\boldsymbol{V}$, $\tilde{\boldsymbol{g}} = \boldsymbol{V}^T\boldsymbol{g}$ and $\tilde{\boldsymbol{F}} = \boldsymbol{V}^T\boldsymbol{F}$. A number of techniques have been proposed for the assembly of $\boldsymbol{V}$ [2]. Herein we employ the Proper Orthogonal Decomposition [12], which relies on a collection of FOM evaluations for a training set of parameters and assembles $\boldsymbol{V}$ via truncating the corresponding decomposition modes using suitable error measures. The Energy Conserving Mesh Sampling and Weighting hyper-reduction technique is also employed herein, which refers to a second-tier approximation required to tackle the computational bottleneck of updating $\tilde{\boldsymbol{g}}$ in Eq. 3 efficiently [14].

To address variability, a typical strategy is to assemble a collection of local projection bases $\boldsymbol{V}_i$, each corresponding to FOM snapshots for a specific realization $\boldsymbol{p}_i$, and utilize suitable interpolation or clustering techniques [15]. In this context, the proposed framework utilizes a nonlinear generative model in the form of a cVAE to capture a generalizable nonlinear mapping in the parameter-basis relation. This is shortly presented here and the reader can refer to [12, 16] for a more detailed discussion. First, a collection of local bases is obtained as described and the following system is solved in the least-squares sense:

$$\boldsymbol{V}_i\left(\boldsymbol{p}_i\right) = \boldsymbol{V}_{global} * \boldsymbol{X}_i\left(\boldsymbol{p}_i\right) \tag{4}$$

where $\boldsymbol{V}_i \in \mathbb{R}^{n \times r}$ is an instance of the collection, $\boldsymbol{V}_{global} \in \mathbb{R}^{n \times \tilde{r}}$ captures the dynamics across the entire domain and $\boldsymbol{X}_i \in \mathbb{R}^{\tilde{r} \times r}$ is a coefficient matrix. $\tilde{r}$ signifies the total number of truncated modes retained on $\boldsymbol{V}_{global}$. In turn, the proposed approach uses a conditional VAE as a nonlinear approximator of coefficients $\boldsymbol{X}$, and thus local projection-subspaces $\boldsymbol{V}$ and local ROMs, conditioned on known system parameters. Regarding the implementation aspects, the parametric features $\boldsymbol{p}$ are concatenated with the inputs of the VAE $\boldsymbol{X}$, and its latent space $\boldsymbol{Z}$. This is represented by an (inferred) probabilistic distribution, which enables the extraction of uncertainty bounds on the response estimates by sampling it and propagating the uncertainty on the estimates to the output through the ROM. After training the cVAE, a generative model is derived that can be sampled through its probabilistic latent space to produce the reduced basis coefficients $\boldsymbol{X}_i$ in Eq. 4. Concretely, once the trained decoder is available, and given the inferred parameter vector realization $\hat{\boldsymbol{p}}$ via Bayesian inference, the variational distribution of the latent space can be sampled. In turn, the generated samples are passed through the decoder portion and predictions of the coefficients $\boldsymbol{X}$ can be made. The generation process is illustrated in Fig. 2

Model selection via Bayesian decision theory

The framework outlined in the previous section can be employed to generate different model resolutions to assist with decision-making and virtual sensing tasks. As highlighted in Fig. 1, three candidates are considered here, each one corresponding to a trade-off between computational efficiency and fidelity: a) The full-order FE model of the system, termed

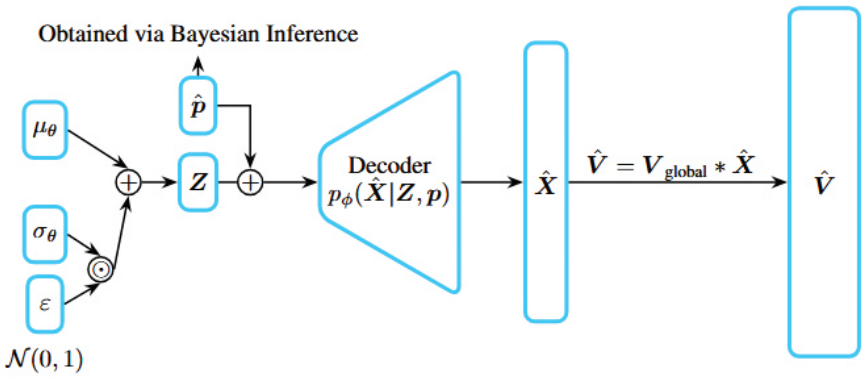


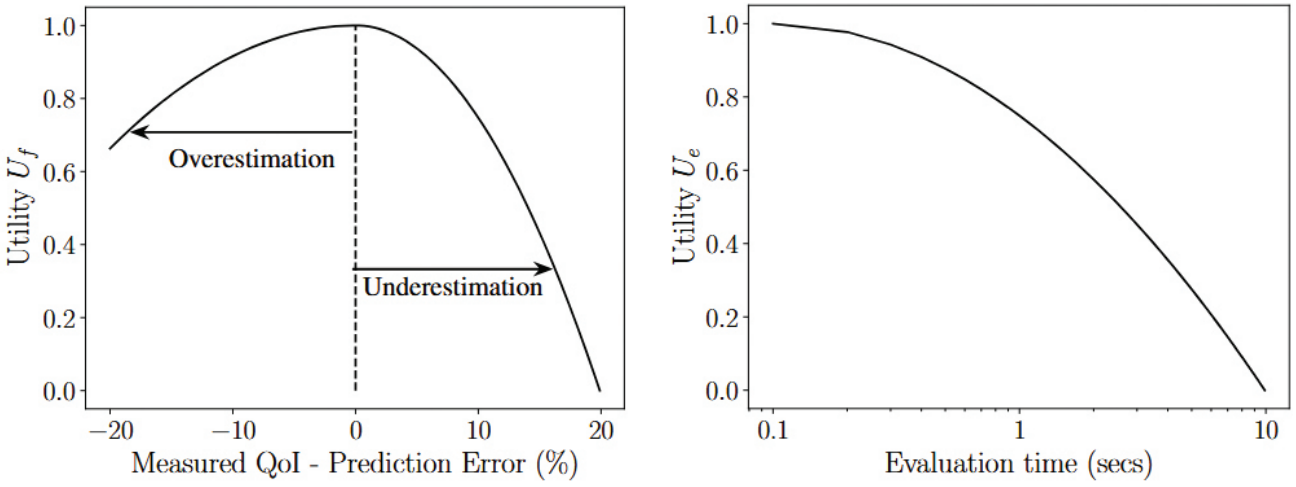
Fig. 2 The proposed framework in generation mode. Dimensionality serves only illustrative purposes

FOM, which offers the highest possible precision, b) the VpROM that strike a balance between precision and efficiency, and c) the H-VpROM, which corresponds to VpROM equipped with hyper-reduction, and potentially offers the highest efficiency. For selecting the best model available conditioned on incoming response measurements Bayesian decision theory is employed, whereby the optimal decision is to employ the representation that maximizes the expected utility [17]. The optimal model to be selected M_{opt} conditional on the monitoring data $\mathbf{d}$ can be obtained as:

$$M_{\text{opt}|\mathbf{d}} = \underset{\{M_1, M_2, \ldots, M_n\}}{\text{argmax}} \; \mathbb{E}_{\hat{\mathbf{p}}|\mathbf{d}, M_i}\left[U_i\right], \tag{5}$$
$$\text{where } U_i = U(\text{fidelity}|M_i, \text{efficiency}|M_i, \hat{\mathbf{p}}|M_i)$$

where U is the considered utility function and $\hat{\mathbf{p}}$ the inferred model parameters. M_i represents the decision to employ the i-th candidate model for recovering the dynamic behavior of the engineered system of interest.

Two competing attributes that contribute to the utility U are considered here, namely the model's precision and efficiency. Since the Bayesian inference framework is also employed for parameter estimation, the utility is evaluated using readings from redundant sensor measurements that are not employed on the inference of $\hat{\mathbf{p}}$. The two branches of the utility function are illustrated in Fig. 3. In 3a underestimation is penalized more as it might lead to unexpected failures and catastrophic



(a) Utility U_f as a function of the model fidelity

(b) Utility U_e as a function the model runtime

Fig. 3 Example utility functions for the considered competing attributes

events. The overall utility can be computed as a weighted combination of the competing attributes as follows

$$U_i(\text{fidelity}|M_i, \text{efficiency}|M_i, \hat{\mathbf{p}}|M_i) = \gamma * U_i(\text{fidelity}|M_i, \hat{\mathbf{p}}|M_i) + (1-\gamma) * U_i(\text{efficiency}|M_i, \hat{\mathbf{p}}|M_i) \tag{6}$$

$$U_i(\text{fidelity}|M_i, \hat{\mathbf{p}}|M_i) = U_f(\hat{\mathbf{p}}|M_i) = \begin{cases} 1 - w_1 * \left(\frac{y_f - \hat{y}_f}{y_f}\right)^2 * (w_1 * \epsilon_{max})^{-1}, & (y_f - \hat{y_f}) > 0 \\ 1 - w_2 * \left(\frac{y_f - \hat{y_f}}{y_f}\right)^2 * (w_2 * \epsilon_{max})^{-1}, & (y_f - \hat{y_f}) \leq 0 \end{cases} \tag{7}$$

$$U_i(\text{efficiency}|M_i, \hat{\mathbf{p}}|M_i) = U_e(\hat{\mathbf{p}}|M_i) = 1 - \frac{\log^2(t_{M_i})}{\log^2(t_{max})} \tag{8}$$

where y_f denotes the observed QoI and $\hat{y}$ the model-approximated ones using the inferred $\hat{\boldsymbol{p}}$ and the selected resolution M_i. The variable t denotes the evaluation time. The weights w_1, w_2 are chosen properly to account for over- and under-estimation effects ($w_1 > w_2$) and to normalize the corresponding utility. Regarding the model updating task, the following probabilistic model is assumed for the discrepancy between the observed acceleration signals $\mathbf{y}$ and the model-predicted response $\hat{\mathbf{y}}$:

$$nRMSE(\hat{\mathbf{p}}, M_i) = \sqrt{\frac{\sum_{j=1}^{n_f} \left(\mathbf{y}_j - \hat{\mathbf{y}}_j(\hat{\mathbf{p}}, M_i)\right)^2}{\sum_{j=1}^{n_f} \mathbf{y}_j^2}} \in \mathbb{R}^{n_f} \sim \mathcal{N}\big(\mathbf{0}, \boldsymbol{\Sigma} = \text{diag}(c^2 ||\mathbf{y}||^2)\big) \tag{9}$$

where $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ denotes the multivariate normal distribution with a zero-mean vector and covariance matrix $\boldsymbol{\Sigma}$. A diagonal covariance matrix is assumed, with the variance of each component assumed proportional to the L_2-norm of $\mathbf{y}$. The factor $c = 0.10$ can be regarded as a coefficient of variation, and its chosen value reflects the total prediction error. The likelihood function can be then written as:

$$L(\boldsymbol{p}; \mathbf{y}) \sim \mathcal{N}\big(\boldsymbol{\eta}; \mathbf{0}, \boldsymbol{\Sigma}\big), \tag{10}$$

where $\mathcal{N}(\cdot; \mathbf{0}, \boldsymbol{\Sigma})$ denotes the value of the MVN density function at a specified location. Thus, the expected fidelity utility U_f to employ model M_i can be approximated using n_s posterior samples as:

$$\mathbb{E}_{\hat{\mathbf{p}}|\boldsymbol{y}, M_i}[U_f(\hat{\mathbf{p}}|M_i)] \approx \frac{1}{n_s} \sum_{j=1}^{n_s} U_f\big(M_i, \hat{\mathbf{p}}_j\big), \tag{11}$$

The equivalent applies for U_e. Regarding the numerical implementation of the Bayesian inference components outlined in this section, the improved Transitional Markov Chain Monte Carlo (iTMCMC) method is employed [18].

Benchmark Shear Frame with Hysteretic Links

The numerical validation focuses on an open-source benchmark of a three-dimensional shear frame with nodal couplings that exhibit Bouc-Wen hysteresis, available in [19]. This example is chosen due to its relevance to real–life structures and its inherent ability to parametrically model multiple simultaneously activated instances of nonlinearity, representing damage or condition deterioration. The proposed framework is utilized to identify the optimal model that maximizes the cost-to-efficiency tradeoff, defined as the utility, to be used for response inference at any given time. The setup is summarized in Tab. 1 and a graphical illustration is offered in Fig. 4a, where the colored columns indicate the regions where evolving damage is introduced.

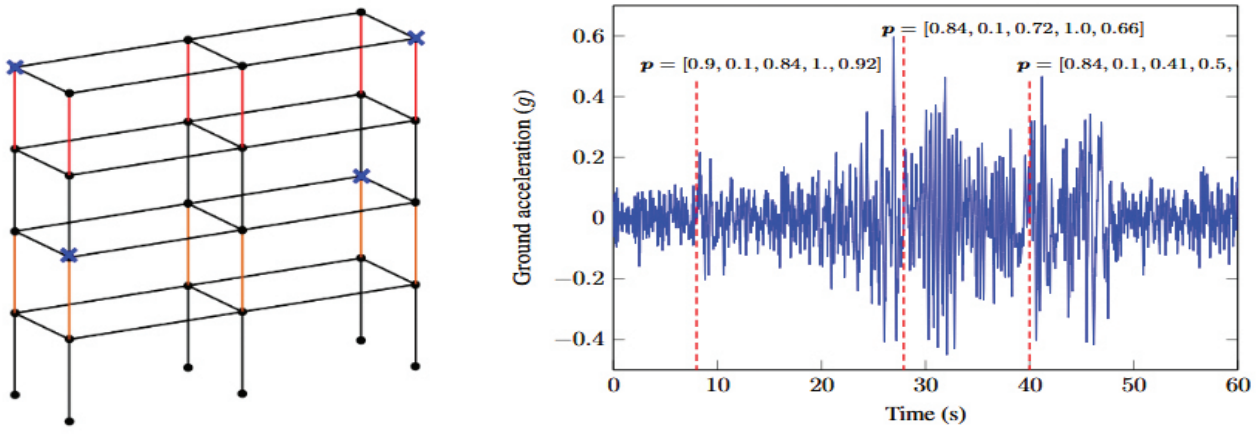
Due to the limited availability of experimental or actual monitoring measurements, simulated measurements are utilized instead. These are obtained via an independent numerical FE model, termed perturbed FOM, and by applying the following measures to avoid the so-called inverse crime:

- A finer mesh is used for the FE model representing the perturbed FOM
- A random perturbation of the Young's Modulus E is introduced.
- Measurement noise equal to 8% of the root-mean-square value of each acceleration signal is added.

Our setup assumes damage occurs in the basement or second floor columns, as illustrated in 4a, via modifying the stiffness in the corresponding Bouc-Wen springs. We also consider the Young modulus E as a varying parameter, to address

Table 1 Details of the case study and its numerical setup

Scenario	Structure behaves linearly at first. Structure and column joints get damaged (unevenly) due to ground motion.
Objective	Select model to be used for O&M downstream tasks
Monitoring Data	Nodal accelerations
	Four nodes monitored (blue crosses in Fig. 4a)
QoIs for Utility computation	Fidelity: Inter-story drift (Normalized RMSE)
	Efficiency: Avg. model evaluation time (s)
Utility functions	Equations 6-8, 3, $t_{max} = 10s, w_1 = 3, w_2 = 1, \epsilon_{max} = 20\%$
Monitoring window	1 second rolling windows



(a) Illustration of the shear frame

(b) Testing input ground motion accelerogram

Fig. 4 Shear frame diagram and the testing ground motion accelerogram with damage events indicated by red lines. Initially the frame is assumed undamaged, and thus $\boldsymbol{p} = [1.0, 1.0, 1.0, 1.0, 1.0]$

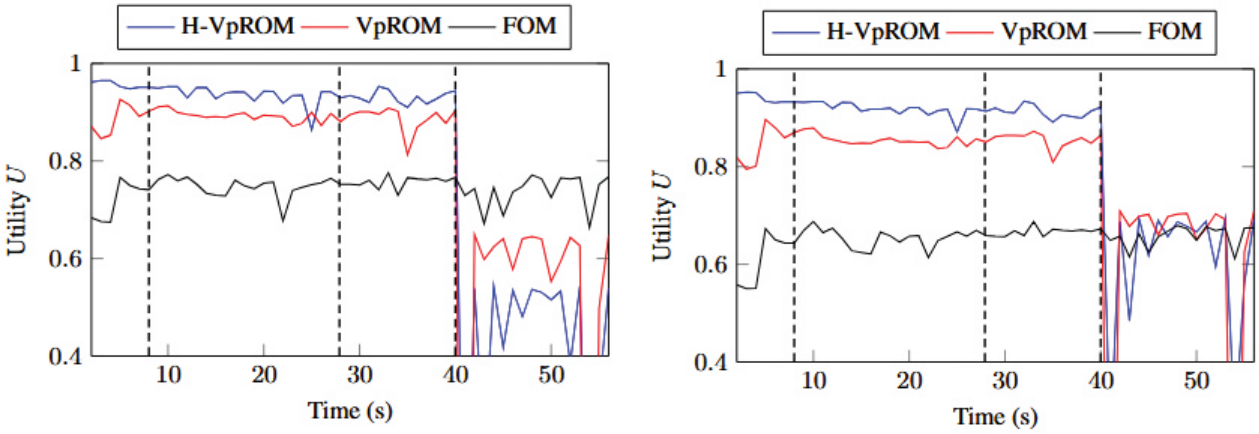
global damage. Thus, the parameter vector to be inferred in Eq. 5 is $\boldsymbol{p} = [E, \alpha, k_1, k_2, k_3]$, where k_i indicates the k parameter in the links of floor i. Parameter α models the initially linear frame ($\alpha = 1.0$) that experiences damage via the activation of the nonlinear springs ($\alpha = 0.1$). The range and distribution of the parameters is summarized in Tab. 2. Regarding material properties, steel HEA cross sections are used, and the frame has $l = 15$ m, $w = 5$ m, and all stories have $h = 3.2$ m.

Forcing is applied to the frame as a base excitation scenario representing an earthquake. Base accelerograms are applied at an angle of $\theta = \pi/4$ with respect to the longitudinal axis. During training, synthetic accelerograms obtained from white noise signals are used, along with 500 training samples of $\boldsymbol{p}$. The framework is tested using the example input sequence in Fig. 4b and three condition deterioration events. Specifically, the frame is initially assumed healthy, and thus linear. After the first damage event, the hysteretic links are activated ($\alpha = 0.10$) with the corresponding stiffness coefficients ($k_1 = 0.84, k_2 = 1.0, k_3 = 0.92$). In the second event, the stiffness coefficients of the hysteretic springs in the first and last floor reduce unevenly, reflecting localized damage while the E modulus also reduces. Lastly, the third event reduces all three stiffness coefficients k_1, k_2, k_3 to values outside the operational regime that was considered for deriving the ROMs. This extreme case is designed to challenge the accuracy of the available reduced representations and test whether the framework will signal a retraining step.

Table 2 System properties and range of parameter values for the training of the ROMs. Variable α changes during operation to signify the activation of the nonlinear links, implying initiation of damage

Parameter:	E ($\times$210 GPa)	$k_1, k_2 (\times 1e^8)$	α	β	γ	δ_η, δ_ν	A, w
Range:	$[0.70, 1.0]$	$[0.6, 1.0]$	$\in \{0.1, 1\}$	3.0	4.0	0.0	1.0

The total utility readings for the example scenario are reported in Fig. 5. The damage events are also illustrated using black dotted lines in the corresponding timestamps. As expected, the hyper-reduced model resolution, termed H-VpROM provides high accuracy response estimates under parametric variability and accelerated model evaluations, thus achieving the maximum utility before the $40s$ timestamp, despite the few low-amplitude oscillations observed. However, the FOM outperforms the ROM variants after the $40s$ damage event, when the integrated parametric traits assume values well-beyond the training domain. Despite the well-documented extrapolation capabilities of physics-based ROMs, this performance is expected, especially under extreme events. As a result, the proposed framework signals successfully that the FOM is to be preferred for tasks requiring high precision. In addition, a threshold can be defined in the utility curves that can serve as an indicator that signals the need for retraining a representation and triggers the corresponding calibration event. Naturally the optimal model selection that depends on the maximization of the utility function also reflects a decision on the importance of the competing attributes. Thus, in Fig. 5a, fidelity and precision are equally weighted, and the framework indicates that the H-VpROM and the FOM are the best models available before and after the $40s$ extreme event. For the sake of demonstration, in Fig. 5b computational runtime is favored over precision by setting $\gamma = 0.3$ in Eq. 6, and, thus, the VpROM seems the most suitable model after the $40s$ event.

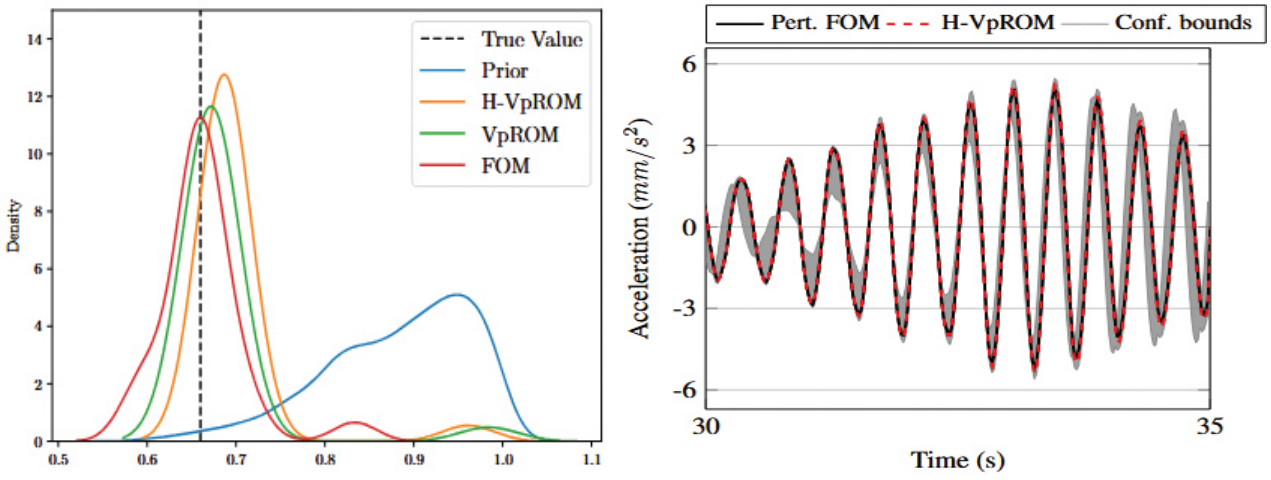


(a) Equally important utility attributes ($\gamma = 0.5$)

(b) Efficiency prioritized over precision ($\gamma = 0.3$)

Fig. 5 Total utility evaluations for different weighted combinations of the considered attributes in Eq. 6

To further scrutinize the performance of the proposed framework, the prior and posterior distributions of the k_1 uncertain model parameter are reported in Fig. 6a after the second damage event. As shown in 6a, the performance on the parameter



(a) Stiffness coefficient k_1 inference

(b) Response approximation using H-VpROM

Fig. 6 Output response approximation and error bounds at the 30s mark,including a 5s-ahead response prediction

inference task depends on the underlying model functioning as forward simulator. The FOM exhibits the highest precision, followed by the VpROM and lastly the hyper-reduced H-VpROM. However, the H-VpROM is the optimal candidate, as reported in Fig. 5a, which further highlights why the utility should be the critical quantity for identifying the optimal model, as it weights both model fidelity and efficiency. The corresponding time history response approximation is reported in Fig. 6b, along with the uncertainty bounds obtained. Interestingly, these bounds contain the model's response and can be treated as indicative of the error and confidence that may be attributed to the response estimates, thus offering a quantification of the uncertainty that propagates into related decision tasks.

Discussion and Limitations

Our work addresses a key challenge in engineering decision support, namely the adaptive selection of the optimal model for downstream O&M tasks in an online and continuous manner. The proposed framework relies on a generative framework that outputs rapid but reliable ROMs of different resolutions conditioned on parametric inputs. These models are integrated as forward simulators of the system's dynamics into a Bayesian framework tasked to infer the underlying parametric traits from monitoring data, and, quantify the resulting utility of the models for the decision-maker. This is achieved via maximizing a decision-theoretic objective function that considers two competing attributes: i) precision, which takes into account the over/under-estimation of target quantities, and (ii) computational runtime.

Furthermore, the developed integrated methodology can be employed to suggest whether to accept a candidate model's prediction, switch to a different generated instance or parameter realization, or trigger a retraining event, thus enabling an adaptive treatment. The framework will be validated on large-scale examples and case studies involving experimental data as next to pinpoint its potential and limitations for real-time O&M.

Acknowledgments The authors gratefully acknowledge the funding from the European Commission under the Horizon Europe funding guarantee, for the projects 'ReCharged - Climate-aware Resilience for Sustainable Critical and interdependent Infrastructure Systems enhanced by emerging Digital Technologies' (grant agreement No: 101086413) and 'TURING - Trustworthy Unified Robust Intelligent Generative Systems' (grant agreement No: 101215032).

References

1. Chinesta, F. and Cueto, E. . Empowering engineering with data, machine learning and artificial intelligence: a short introductive review. *Advanced Modeling and Simulation in Engineering Sciences*, 9(1):21, 2022.
2. Benner, P. , Ohlberger, M. , Cohen, A. , and Willcox, K. . *Model reduction and approximation: theory and algorithms*. SIAM, 2017.
3. Azzi, M. J. and Farhat, C. . Enhanced Multimodal Nonparametric Probabilistic Method for Model-Form Uncertainty Quantification and Digital Twinning. *AIAA Journal*, pages 1–15, 2024.
4. Tatsis, K. E. , Agathos, K. , Chatzi, E. , and Dertimanis, V. K. . A hierarchical output-only Bayesian approach for online vibration-based crack detection using parametric reduced-order models. *Mechanical Systems and Signal Processing*, 167:108558, 2022.
5. Kamariotis, A. , Vlachas, K. , Ntertimanis, V. , Koune, I. , Cicirello, A. , and Chatzi, E. . On the consistent classification and treatment of uncertainties in structural health monitoring applications. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 11(1):011108, 2025.
6. Moya, B. , Badías, A. , Alfaro, I. , Chinesta, F. , and Cueto, E. . Digital twins that learn and correct themselves. *International Journal for Numerical Methods in Engineering*, 123(13):3034–3044, 2022.
7. Rocha, I. B. , Meer, F. , van der, and Sluys, L. J. . An adaptive domain-based POD/ECM hyper-reduced modeling framework without offline training. *Computer Methods in Applied Mechanics and Engineering*, 358:112650, 2020.
8. Etter, P. A. and Carlberg, K. T. . Online adaptive basis refinement and compression for reduced-order models via vector-space sieving. *Computer Methods in Applied Mechanics and Engineering*, 364:112931, 2020.
9. Peherstorfer, B. and Willcox, K. . Online adaptive model reduction for nonlinear systems via low-rank updates. *SIAM Journal on Scientific Computing*, 37(4):A2123–A2150, 2015.
10. Kamariotis, A. , Chatzi, E. , and Straub, D. . Value of information from vibration-based structural health monitoring extracted via Bayesian model updating. *Mechanical Systems and Signal Processing*, 166:108465, 2022. Publisher: Elsevier.
11. Worden, K. , Cross, E. J. , Barthorpe, R. J. , Wagg, D. J. , and Gardner, P. . On digital twins, mirrors, and virtualizations: Frameworks for model verification and validation. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 6(3):030902, 2020.
12. Vlachas, K. , Tatsis, K. , Agathos, K. , Brink, A. R. , and Chatzi, E. . A local basis approximation approach for nonlinear parametric model order reduction. *Journal of Sound and Vibration*, 502:116055, 2021.
13. Kamariotis, A. and Chatzi, E. . Bayesian decision-theoretic model selection for monitored systems. *arXiv preprint arXiv:2310.10485*, 2023.

14. Farhat, C. , Avery, P. , Chapman, T. , and Cortial, J. . Dimensional reduction of nonlinear finite element dynamic models with finite rotations and energy-based mesh sampling and weighting for computational efficiency. *International Journal for Numerical Methods in Engineering*, 98(9):625–662, 2014.
15. Amsallem, D. and Haasdonk, B. . PEBL-ROM: Projection-error based local reduced-order models. *Advanced Modeling and Simulation in Engineering Sciences*, 3:1–25, 2016.
16. Vlachas, K. , Simpson, T. , Garland, A. , Quinn, D. D. , Farhat, C. , and Chatzi, E. . Reduced order modeling conditioned on monitored features for response and error bounds estimation in engineered systems. *Mechanical Systems and Signal Processing*, 226:112261, 2025.
17. Keeney, R. L. and Raiffa, H. . *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press, 1993.
18. Betz, W. , Papaioannou, I. , and Straub, D. . Transitional markov chain monte carlo: observations and improvements. *Journal of Engineering Mechanics*, 142(5):04016016, 2016.
19. Vlachas, K. , Agathos, K. , Tatsis, K. E. , Brink, A. R. , and Chatzi, E. . Two-story frame with Bouc-Wen hysteretic links as a multi-degree of freedom nonlinear response simulator. In *5th Workshop on Nonlinear System Identification Benchmarks*, 2021.

Chapter 7
Overcoming Data Scarcity for Load Identification and Full-Field Response Reconstruction

Aohan Wen and Alana Lund

Abstract Evidence of a substantial population of aging, deteriorated, or otherwise hazardous bridge infrastructure across the world has necessitated the development of new techniques for bridge health monitoring that can interpret behavior, localize damage, and prioritize repairs across bridge networks. Long-term continuous monitoring of bridges through a dense, stationary sensor network represents a common approach in the research community to rapidly interpret bridge condition, which has been shown to be successful in cases such as the Jindo Bridge (Korea) or the St. Anthony Falls Bridge (USA). However, achieving an estimate of full-field dynamic structural behavior using sparse and sub-optimal placed sensors remains a major challenge. This paper presents a novel approach that integrates a generalized structural model of a bridge based on the unit load method with non-negative Bayesian linear regression to enable real-time reconstruction of the full-field structural responses. We demonstrate this approach on a simulated 12-meter beam to estimate the loading scenario, bending moments and deflections using only sparsely deployed strain gauges in an underdetermined identification context. The proposed method is compared with traditional regularization methods and shows great accuracy in structural response reconstruction and load identification. This approach presents opportunities to reduce costs in long-term monitoring systems and allow for those systems to be more easily repurposed toward different monitoring goals.

Keywords Load identification · Sparse sensing · Non-negative regularization constraints · Response reconstruction

Introduction

Accurate load identification is important for structural health monitoring of infrastructure systems. Based on accurate load, the response of structure can be then reconstructed in real time to further monitor changes in structural parameters and prevent failure.

Indirect measurement methods based on strain response are commonly used in load identification. However, identifying or estimating structure input loads using system characteristics and responses is an inverse problem [1]. In practical deployments with sparse sensing, the number of parameters needed to characterize the load's spatial distribution and magnitudes often exceeds the number of available sensors, yielding an underdetermined, ill-posed inverse problem with non-unique and unstable solutions [2]. To address this problem, many regularization techniques have been utilized to achieve a unique and stable solution, like Tikhonov [3], Lasso [4] and Bayesian regularization methods [5].

Bayesian regularization methods are particularly useful because of their ability to estimate weights together with statistical information. A prominent instance is the Automatic Relevance Determination (ARD) method. Within a hierarchical Bayesian framework, ARD method employs evidence-driven hyperparameter learning to automatically select relevant variables, yielding sparse solutions that mitigate ill-posedness and enhance stability and identifiability in underdetermined settings [6]. However, to ensure conjugacy with the Gaussian noise in the likelihood, ARD method typically places a Gaussian prior on the weights, which allows the inferred weights to be negative. This characteristic hinders the applicability of ARD methods when non-negative parameters are required, such as in the case of identifying gravity induced loads, material densities, or strength parameters. Rather than imposing a non-negative prior distribution, we enforce non-negativity by modifying the feature-pruning scheme, resulting in a non-negative ARD (NN-ARD) approach.

Aohan Wen · Alana Lund
Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
e-mail: a25wen@uwaterloo.ca; alund@uwaterloo.ca

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

This study first formulates a least-squares problem for load identification with a non-negative support set via the unit-load method. These loads are then identified using NN-ARD approach, and model selection is performed by minimizing the negative log-likelihood (NLL). The structural displacement response is reconstructed using the response design matrix. Uncertainties in the load and measurement noise are estimated and propagated to the displacement response.

Unit load method

For load identification from structural responses, the possible loading scenario on the structure is first discretized into a series of finite-length uniformly distributed loads to approximate the true loading. As shown in Fig. 1, a beam under an arbitrary distributed load $W(x)$ with pinned support on the left and fixed support on the right is shown in the left figure and its discretized uniformly distributed loads is shown in the right figure.

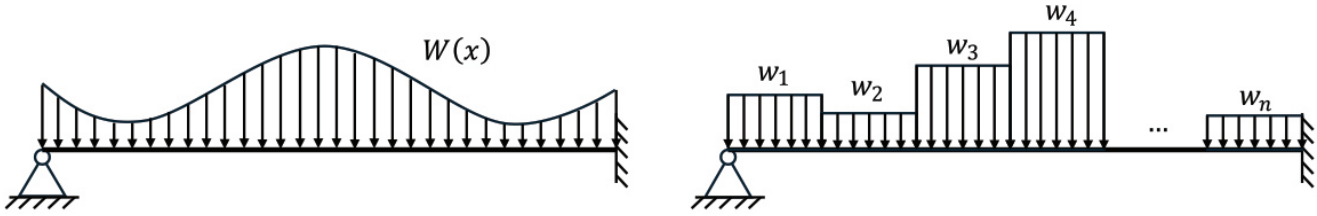


Fig. 1 Beam under an arbitrary distributed load $W(x)$ and beam under discretized uniformly distributed load segments $(w_1, w_2, ..., w_n)$

Under the assumption of linear elastic and small deformation (hence superposition holds), the beam's response under a series of loading segments $w_1,\ w_2,\ \cdots, w_n$ can be approximated as the sum of responses under each of discretized uniformly distributed load segments. Then, given linearity, the response of the beam to each discretized uniformly distributed load segment of intensity w_n is exactly the response of unit distributed load $\overline{w}_n = 1$ on that segment scaled by w_n. Therefore, the structural response can be represented as weighted summation of structural response under unit distributed loads, as shown in Eq. (1) below.

$$\boldsymbol{\mathcal{R}}(W) = \sum_{i=1}^{n} \boldsymbol{\mathcal{R}}(w_i) = \sum_{i=1}^{n} w_i \boldsymbol{\mathcal{R}}(\overline{w}_i) \# (1) \tag{1}$$

Treating the structural response under unit distributed loads as a column vector, the response under unit distributed loads applied at different locations can be assembled into a design matrix $\boldsymbol{\Phi}$.

$$\boldsymbol{\Phi} = \left[\begin{array}{cccc} \boldsymbol{\mathcal{R}}(\overline{w}_1) & \boldsymbol{\mathcal{R}}(\overline{w}_2) & \ldots & \boldsymbol{\mathcal{R}}(\overline{w}_n) \end{array} \right] \# (2) \quad \boldsymbol{w} = \left[\begin{array}{cccc} w_1 & w_2 & \ldots & w_n \end{array} \right]^{\top} \# (3) \tag{2}$$

where, $\boldsymbol{\mathcal{R}}(W) \epsilon \mathbb{R}^{m \times 1}$ is the structural response at m measurement stations under true distributed load $W(x)$, $\boldsymbol{\mathcal{R}}(\overline{w}_i) \epsilon \mathbb{R}^{m \times 1}$ is the structural response at m measurement stations under unit load $\overline{w}_i = 1$, $\boldsymbol{\Phi} \epsilon \mathbb{R}^{m \times n}$ is the design matrix with n segments along the beam, $\boldsymbol{w} \epsilon \mathbb{R}^{n \times 1}$ is the weight or scaling factor of n segments.

In our research problem, the objective is to identify the true load $W(x)$ acting on the structure. The structural responses $\boldsymbol{\mathcal{R}}(W)$ at the measurement stations are observable — for example, the bending moment of a beam at a given location can be inferred from strain. We approximate $W(x)$ by a linear combination of n uniformly distributed load segments with weights $\boldsymbol{w} \epsilon \mathbb{R}^{n \times 1}$. The design matrix $\boldsymbol{\Phi} \epsilon \mathbb{R}^{m \times n}$ is formed by taking the structural response $\boldsymbol{\mathcal{R}}(\overline{w}_i)$ to a normalized gravity-induced unit distributed load on segment i as its i-th column. In other words, each $\boldsymbol{\mathcal{R}}(\overline{w}_i)$ serves as a basis vector of the response space and directly constitutes one column of the design matrix. These responses can be obtained analytically, numerically or experimentally. This yields a general linear model, where $\boldsymbol{\varepsilon}$ represents the noise:

$$\boldsymbol{\mathcal{R}}(W) = \boldsymbol{\Phi w} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} \epsilon \mathbb{R}^{m \times 1} \# (4) \tag{3}$$

There are two important features for this specific general linear model. One is non-negativity on weights. The other one is underdetermined problem. For non-negativity, since the loads under consideration are gravity-induced and the columns of design matrix $\boldsymbol{\Phi}$ are constructed from responses to normalized gravity loads, the coefficients represent non-negative load intensities. Consequently, the admissible solution is supported on the non-negative domain, and we impose the constraint $\boldsymbol{w} \geq \boldsymbol{0}$. For underdetermined problems, with sparse sensing, the number of measurement stations m is limited. To better capture the spatial variation of $W(x)$, the load is discretized more finely so that $n > m$. The identification is therefore underdetermined and ill-posed. To stabilize the problem and obtain a unique, physically meaningful solution, different regularization methods are applied.

Regularization Method

The problem of load identification can be formulated as a least-squares (LS) optimization framework, wherein the objective function is defined by the discrepancy between the predicted structural responses, $\boldsymbol{\Phi w}$, and the experimentally measured responses, $\boldsymbol{\mathcal{R}}(W)$. However, the direct solution of such an LS formulation is often hindered by the intrinsic ill-posedness of inverse problems. To address this challenge, regularization methods are commonly employed, which enable the incorporation of prior knowledge to provide a stable approximation of ill-posed (pseudo-) inverses [7]. Different prior knowledge applied to solutions will lead to different regularization techniques:

Tikhonov regularization, also known as ℓ_2 regularization, is a widely used technique to stable inverses by penalizing the ℓ_2-norm on solution to the LS problem. The regularization parameter, denoted as λ, controls the penalty magnitude. In particular, ℓ_2 regularization penalizes large coefficient values, thereby preventing individual solution components from dominating the fit. This not only alleviates instability inherent to ill-posed problems but also promotes smoother and more robust solutions [8].

$$\min_{\boldsymbol{w}} \|\boldsymbol{\mathcal{R}}(W) - \boldsymbol{\Phi w}\|_2^2 + \lambda \|\boldsymbol{w}\|_2 \,\#\,(5) \tag{4}$$

Lasso regularization, often referred to as ℓ_1 regularization, imposes a penalty proportional to the absolute value of the coefficients. Unlike ℓ_2 regularization, which shrinks all coefficients smoothly towards zero, ℓ_1 regularization encourages sparsity by driving some coefficients exactly to zero. This property not only stabilizes the solution of ill-posed problems but also enables automatic feature selection, leading to more interpretable models [9].

$$\min_{\boldsymbol{w}} \|\boldsymbol{\mathcal{R}}(W) - \boldsymbol{\Phi w}\|_2^2 + \lambda \|\boldsymbol{w}\|_1 \,\#\,(6) \tag{5}$$

Bayesian regularization provides a probabilistic framework for stabilizing ill-posed problems by introducing prior distributions over model parameters. Instead of imposing penalties in an explicit optimization form (as in ℓ_1 or ℓ_2 regularization), Bayesian methods encode assumptions about parameter magnitudes and sparsity through priors and update these beliefs using observed data via Bayesian theorem. A notable example is the Automatic Relevance Determination (ARD) approach, which assigns each model weight an independent Gaussian prior with zero mean and its own precision (inverse variance) hyperparameter. By iteratively updating these hyperparameters, ARD suppresses irrelevant parameters while preserving influential ones, thereby achieving both regularization and feature selection within a Bayesian framework [6], [8]. An important advantage of this method is that it not only identifies the weights but also provides their associated statistical information, such as covariance of weights and noise. Furthermore, Bayesian regularization eliminates the need for complex tuning of regularization parameters, since the process naturally incorporates adaptive and automatic parameter selection.

Non-negative ARD Method

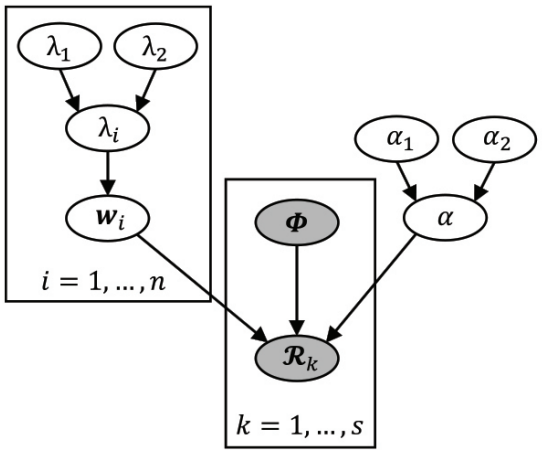
In the standard ARD framework, as shown in Fig. 2, each weight is assigned a zero-mean Gaussian prior with an individual precision hyperparameter. While this setting facilitates automatic feature selection, it also implies that the posterior estimates of the weights are distributed around zero. As a result, negative weight values may appear in the solution, which may be undesirable in applications where the underlying weights are inherently non-negative.

In the Bayesian framework, non-negativity can be introduced through the choice of priors, like the truncated Gaussian distribution [10] or the exponential distribution [11]. However, these priors break conjugacy, which means that the posterior distribution of the weights $\boldsymbol{w}$ no longer has a closed-form analytical expression. Consequently, approximate inference techniques such as variational inference (VI) are required to get posterior distribution of weights [10], [12]. This not only increases the methodological complexity but also entails substantial computational cost, making such non-negative Bayesian approaches difficult to apply in online identification tasks.

In this study, non-negativity is achieved by altering the feature-pruning rule rather than modifying the prior distribution of the weights. A zero-mean Gaussian prior is still adopted, which, combined with a Gaussian likelihood, guarantees the existence of a closed-form analytical posterior solution due to conjugacy. For the nth weight, let the posterior distribution in ARD method to be $w_n^+ \sim \mathcal{N}(\mu_n, \Sigma_{nn})$ with standard deviation $\sigma_n = \sqrt{\Sigma_{nn}}$. Then the standardized variable $Z = \frac{w_n^+ - \mu_n}{\sigma_n}$ follows the standard normal distribution, i.e. $Z \sim \mathcal{N}(0,1)$. Define the one-sided positivity probability as $P_n = P(w_n^+ > 0)$:

$$P_n = P(w_n^+ > 0) = P\left(\tfrac{w_n^+ - \mu_n}{\sigma_n} > \tfrac{0-\mu_n}{\sigma_n}\right) = P\left(Z > -\tfrac{\mu_n}{\sigma_n}\right) = 1 - \Phi\left(-\tfrac{\mu_n}{\sigma_n}\right) = \Phi\left(\tfrac{\mu_n}{\sigma_n}\right) \#\,(7) \tag{6}$$

Keeping the feature n if $P_n \geq \tau$; otherwise, the feature will be pruned. One choice, which will be used in the remainder of this study, is $\frac{\mu_n}{\sigma_n} = 0.5$ ($\tau \approx 0.69$).



Prior of weights:

$$\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Lambda}^{-1}), \qquad \boldsymbol{\Lambda} = diag(\boldsymbol{\lambda}), \qquad \boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)$$

Likelihood:

$$\boldsymbol{\mathcal{R}}_k \mid \boldsymbol{w} \sim \mathcal{N}(\boldsymbol{\Phi w}, \alpha^{-1}\boldsymbol{I}), \qquad k = 1, \dots, s$$

Hyperparameters:

$$p(\alpha) = \text{Gamma}(\alpha \mid \alpha_1, \alpha_2) \propto \alpha^{\alpha_1} \exp(-\alpha_2 \alpha)$$

$$p(\lambda_i) = \text{Gamma}(\lambda_i \mid \lambda_1, \lambda_2) \propto \lambda_i^{\lambda_1} \exp(-\lambda_2 \lambda_i)$$

Fig. 2 Probabilistic graphical model and underlying principles of ARD method

Numerical Case Study

A constant cross-section beam with the left end pinned and the right end fixed is adopted as a numerical case study for this method. This beam has a length of 12m. Homogeneous material with elastic modulus of 2E11Pa is applied and the cross-sectional moment of inertia is 6.572E-5 m^4. Bending moments along the beam are inferred from strain measurements by considering the flexural rigidity of cross sections and the relative vertical positions between neutral axis and strain gauges. Positions of measurement stations along the beam are illustrated in Fig. 3.

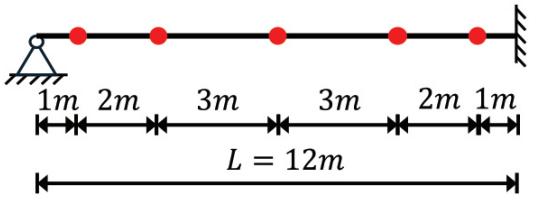


Fig. 3 Positions of measurement stations (strain gauges) along the beam

Different numbers of segments are examined, and the corresponding design matrices under unit uniform loads are obtained numerically. The structural responses considered include bending moment and deflection. Each line in Fig. 4 illustrates base vectors (BVs) of those responses produced by unit uniform loads applied at different positions for 1,2, and 3 segments.

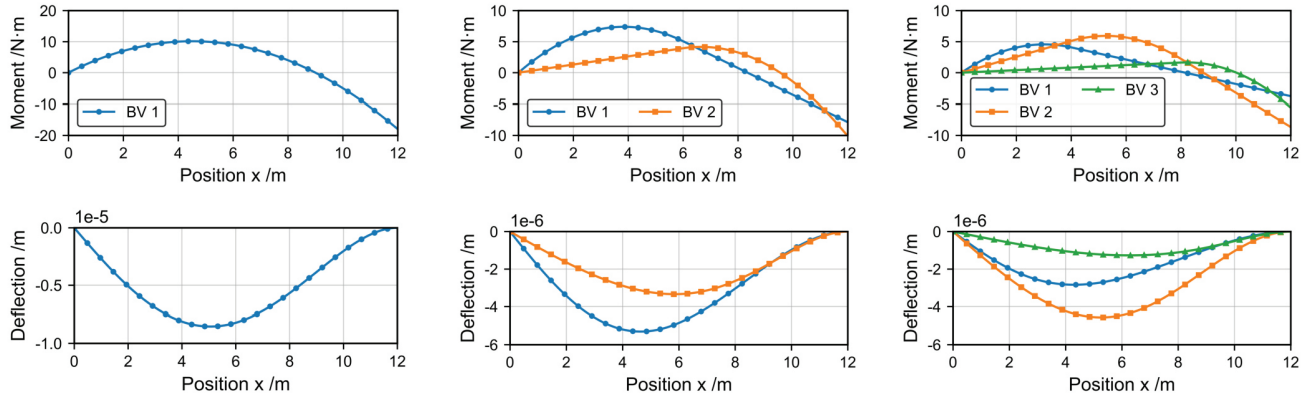
In the process of load identification, different numbers of segments can be applied to identify the loading scenario. To make the regularization method adaptive and select the optimal number of segments using only the identified result, an index called negative log-likelihood (NLL) is utilized [13]. NLL is the log form of likelihood of prediction model. The expression of NLL is presented in Eq. (7) under the assumption that the noise level is constant at different measurement stations and observations are independent and identically distributed:

$$NLL = \frac{1}{2} log(2\pi\widehat{\sigma}^2) + \frac{1}{2ms\widehat{\sigma}^2} \sum_{k=1}^{s} \sum_{j=1}^{m} (y_j^{(k)} - \hat{y}_j)^2 \# (8) \tag{7}$$

where $\widehat{\sigma}^2$ is the estimation of noise variance, $\hat{y}_j$ is the model prediction at measurement station j, $y_j^{(k)}$ is the kth observation at measurement station j, m is the number of measurement stations on the structure, and s is the number of observations in each measurement station. By taking the partial derivative of NLL to the estimation of noise variance $\widehat{\sigma}^2$ and setting it to zero, we obtain:

$$\widehat{\sigma}^2 = \frac{\sum_{k=1}^{s} \sum_{j=1}^{m} \left(y_j^{(k)} - \hat{y}_j\right)^2}{ms} \# (9) \tag{8}$$

Equation (8) demonstrates that the estimated variance of measurement noise in the NLL is interpreted as the mean discrepancy between the observations and the model predictions. A lower NLL then indicates that residuals are well explained by the noise model, implying better predictive fit. We fit models with varying numbers of segments and select the segment count that minimizes NLL. Computed via Eq. (7), NLL assesses consistency between the identified and measured responses under the estimated noise variance and does not require true response; by contrast, the root mean square error (RMSE),



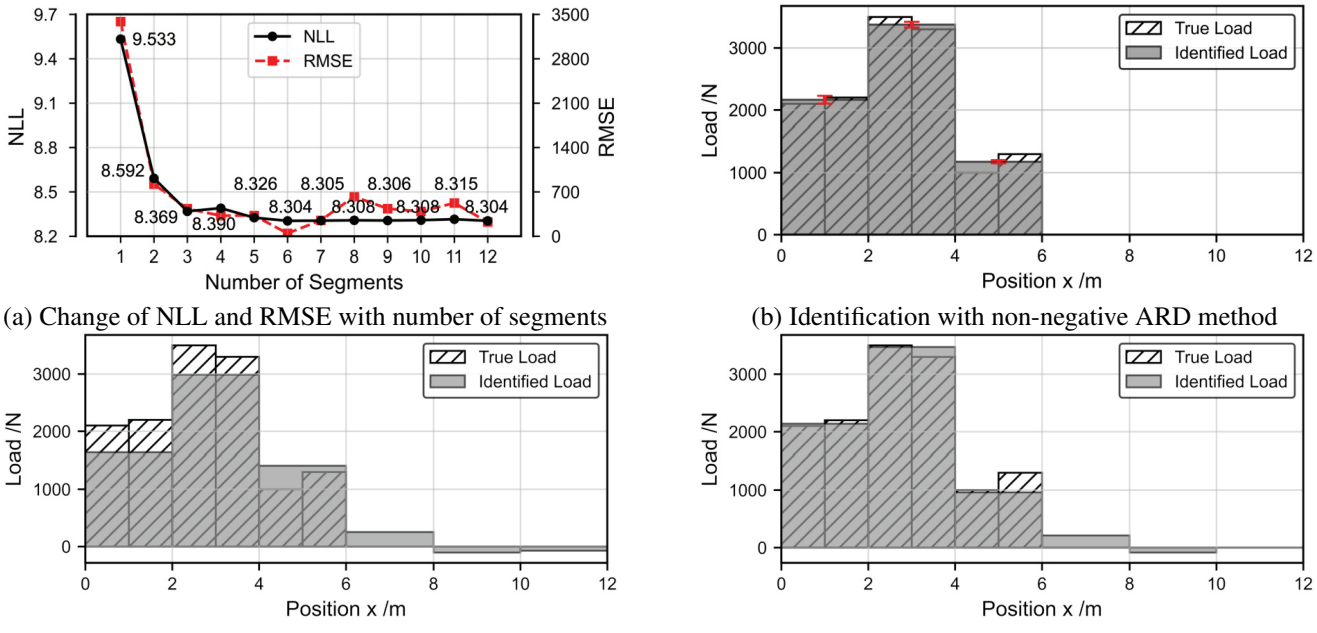
(a) BVs under 1 segment (b) BVs under 2 segments (c) BVs under 3 segments

Fig. 4 Base vectors of structural response under 1/2/3 segments

which is utilized for comparison, evaluates the difference between the true and identified responses and is often an infeasible metric in practice. However, NLL does not penalize model complexity and therefore cannot, by itself, prevent overfitting. In the case studies, we mitigate overfitting and improve generalization by promoting sparsity and preferring simpler models.

Case 1: Half span distributed load

Assume 500 sets of bending moments are inferred from strain measurements with an imposed additive noise sampled from $N(0, 1000\ (Nm)^2)$. The true load is discretized into 12 uniformly distributed loading sections across the beam. Models possessing between 1 and 12 segments are considered in the load identification process. The value of NLL and RMSE as functions of the number of segments are presented in Fig. 5(a), where values of NLL are recorded and presented at each number of segments. The true load and identification result by NN-ARD method are shown in Fig. 5(b). For comparison, the load identification results with ℓ_2 and ℓ_1 regularization are illustrated in Fig. 5(c) and (d) respectively. The regularization parameters λ are set to 1.0 in each case. Increasing λ strengthens the penalty, which driving ℓ_1 regularized solutions towards more sparsity and ℓ_2 regularized solutions towards more smoothness.



(a) Change of NLL and RMSE with number of segments

(b) Identification with non-negative ARD method

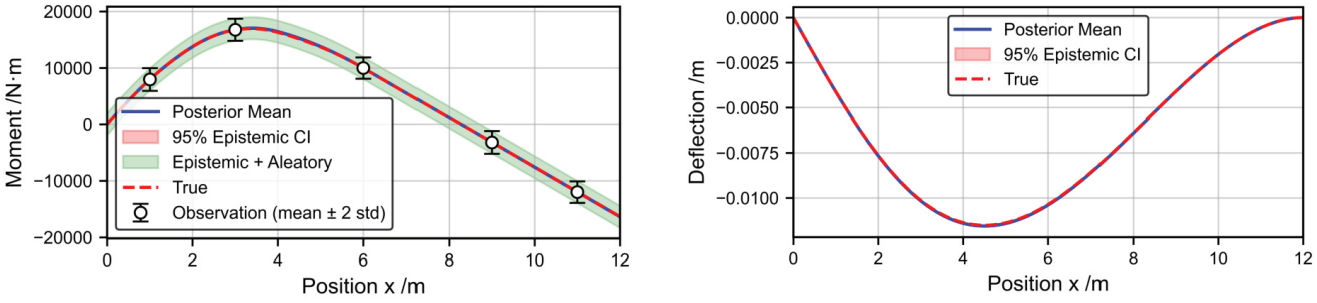
(c) Identification with ℓ_2 regularization

(d) Identification with ℓ_1 regularization

Fig. 5 Load identification results under half span distributed load

As shown in Fig. 5(a), both NLL and RMSE attain their minimum value when the number of segments is 6, indicating the NLL can correctly identify the number of segments that yields the best predictive performance. It's worth noting that, although the NLL is also comparatively low when the number of segments is 12, this comes with increased model complexity. In line with Occam's razor, we prefer a simpler model, i.e. a model with fewer segments, even if this requires accepting a marginal NLL increase to reduce the number of segments. The NN-ARD method yields sparse, non-negative load estimates with quantified uncertainty. By contrast, Tikhonov and Lasso produce physically inadmissible negative coefficients; with $\lambda = 1.0$, the $\ell 1$ penalty is too weak to induce meaningful sparsity, implying a larger λ is required.

The reconstructed response of moment and deflection and their propagated uncertainties from NN-ARD method are further shown in Fig. 6. Two kinds of uncertainties are considered here. Epistemic uncertainty corresponds to estimated variance from posterior distribution of weights which can be propagated into deflection by $\boldsymbol{\Phi}_D \boldsymbol{\Sigma} \boldsymbol{\Phi}_D^\top$, where $\boldsymbol{\Phi}_D$ is the design matrix of deflection and $\boldsymbol{\Sigma}$ is the posterior covariance of weights. Aleatory corresponds to identified variance of noise. The standard deviation of estimated noise is 978.12 Nm.



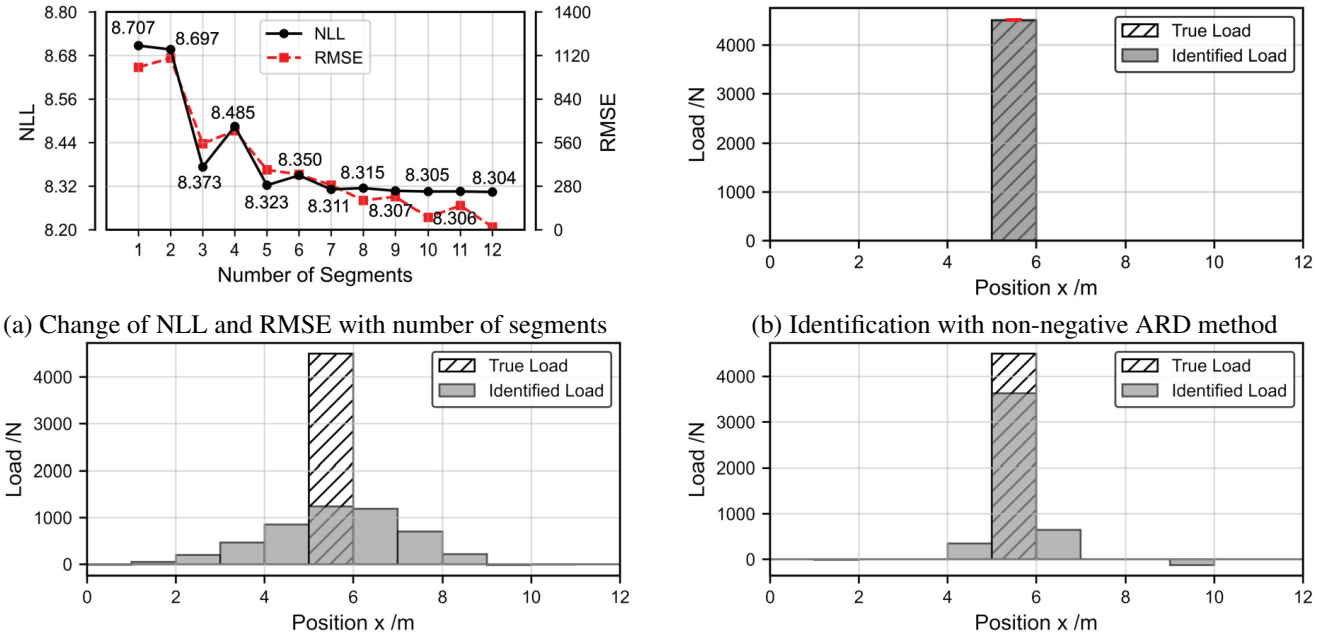
(a) Moment response reconstruction

(b) Deflection response reconstruction

Fig. 6 Response reconstruction and uncertainty propagation under half span distributed load

Case 2: Concentrated load

Instead of a point load, we apply a uniformly distributed load over a 1-m segment to approximate a concentrated load. This setup enables a more direct and interpretable comparison between the true load and the identified load. As shown in Fig. 7(a),



(a) Change of NLL and RMSE with number of segments

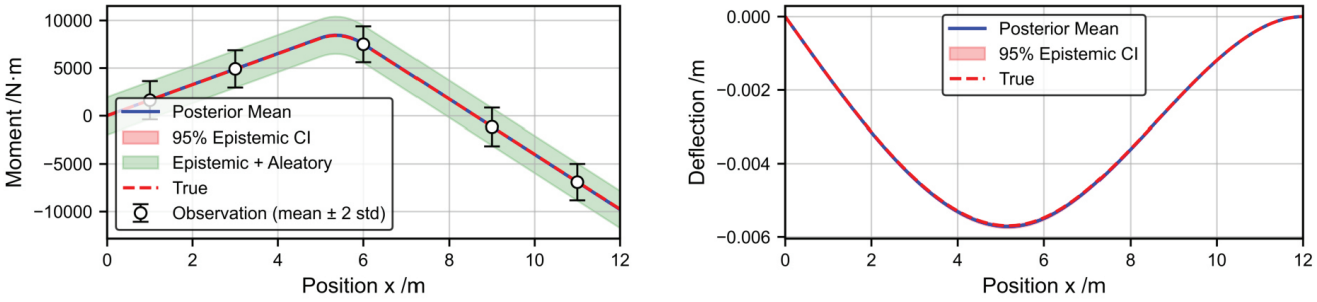
(b) Identification with non-negative ARD method

(c) Identification with ℓ_2 regularization

(d) Identification with ℓ_1 regularization

Fig. 7 Load identification results under concentrated load

both NLL and RMSE attain their minimum value when the number of segments is 12. In Fig. 7(b), a sparse solution with estimated uncertainty in weights are obtained by NN-ARD method. Identification with ℓ_2 regularization tends to present a smooth solution, while a sparse solution is given under ℓ_1 regularization. The reconstructed response of bending moment and deflection and their propagated uncertainties are shown in Fig. 8.



(a) Moment response reconstruction

(b) Deflection response reconstruction

Fig. 8 Response reconstruction and uncertainty propagation under concentrated load

Thanks to the accurate load estimation, the structural response can be reconstructed with high fidelity. As shown in Fig. 8(b), the posterior mean of the deflection almost coincides with the true response, resulting in a very narrow 95% credible interval for epistemic uncertainty. The standard deviation of estimated noise in this case is 977.88 Nm.

Conclusion

The proposed non-negative ARD (NN-ARD) method accurately identifies both the magnitude and location of loads while quantifying their uncertainties and eliminating the negative weights that often arise with conventional regularization methods. Given observed and estimated responses, the negative log-likelihood (NLL) criterion enables effective model selection by maximizing fitting likelihood. Moreover, full-field structural responses can be reconstructed with high fidelity from the response design matrix and the inferred weights, with proper propagation of uncertainty to the reconstructed responses. Future work will apply this method to real structures for in-situ load identification and full-field response reconstruction.

Acknowledgments We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number RGPIN-2023-03651]. 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 NSERC.

Cette recherche a été financée par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), [numéro de référence RGPIN-2023-03651]. Les opinions, conclusions et recommandations exprimées dans ce document sont celles de l'auteur (des auteurs) et ne reflètent pas nécessairement les opinions de CRSNG.

References

1. R. Liu, E. Dobriban, Z. Hou, and K. Qian, "Dynamic Load Identification for Mechanical Systems: A Review," Arch. Comput. Methods Eng., vol. 29, no. 2, pp. 831–863, Mar. 2022, doi: 10.1007/s11831-021-09594-7.
2. D. Xiao, Z. Sharif-Khodaei, and M. H. Aliabadi, "Impact force identification for composite structures using adaptive wavelet-regularised deconvolution," Mech. Syst. Signal Process., vol. 220, p. 111608, Nov. 2024, doi: 10.1016/j.ymssp.2024.111608.
3. M. Aucejo, "Structural source identification using a generalized Tikhonov regularization," J. Sound Vib., vol. 333, no. 22, pp. 5693–5707, Oct. 2014, doi: 10.1016/j.jsv.2014.06.027.
4. C.-D. Pan, L. Yu, H.-L. Liu, Z.-P. Chen, and W.-F. Luo, "Moving force identification based on redundant concatenated dictionary and weighted l1-norm regularization," Mech. Syst. Signal Process., vol. 98, pp. 32–49, Jan. 2018, doi: 10.1016/j.ymssp.2017.04.032.
5. M. Aucejo and O. De Smet, "An optimal Bayesian regularization for force reconstruction problems," Mech. Syst. Signal Process., vol. 126, pp. 98–115, Jul. 2019, doi: 10.1016/j.ymssp.2019.02.021.
6. D. J. MacKay and R. M. Neal, "Automatic relevance determination for neural networks," Proc. Tech. Rep. Prep., 1994.
7. M. Benning and M. Burger, "Modern regularization methods for inverse problems," Acta Numer., vol. 27, pp. 1–111, May 2018, doi: 10.1017/S0962492918000016.

8. C. M. Bishop, Pattern recognition and machine learning. in Information science and statistics. New York: Springer, 2006.
9. R. Tibshirani, “Regression Shrinkage and Selection Via the Lasso,” J. R. Stat. Soc. Ser. B Stat. Methodol., vol. 58, no. 1, pp. 267–288, Jan. 1996, doi: 10.1111/j.2517-6161.1996.tb02080.x.
10. Y. Liu, W. Dong, W. Song, and L. Zhang, “Bayesian Nonnegative Matrix Factorization with a Truncated Spike-and-Slab Prior,” in 2019 IEEE International Conference on Multimedia and Expo (ICME), Shanghai, China: IEEE, Jul. 2019, pp. 1450–1455. doi: 10.1109/ICME.2019.00251.
11. M. N. Schmidt, O. Winther, and L. K. Hansen, “Bayesian Non-negative Matrix Factorization,” in Independent Component Analysis and Signal Separation, T. Adali, C. Jutten, J. M. T. Romano, and A. K. Barros, Eds., Berlin, Heidelberg: Springer, 2009, pp. 540–547. doi: 10.1007/978-3-642-00599-2_68.
12. R. Yuan, C. Leng, S. Zhang, J. Peng, and A. Basu, “Bayesian non-negative matrix factorization with Student’s t-distribution for outlier removal and data clustering,” Eng. Appl. Artif. Intell., vol. 132, p. 107978, Jun. 2024, doi: 10.1016/j.engappai.2024.107978.
13. K. P. Murphy, Machine Learning - A Probabilistic Perspective. in Adaptive Computation and Machine Learning. Cambridge: MIT Press, 2014.

Chapter 8
Digital Twin-Based Computer Vision Method for Robust Structural Vibration Estimation in the Underwater Environment

Enjian Cai, Zihan Wu, Michael D. Todd, and Zhen Hu

Abstract Structural Health Monitoring (SHM) of large civil infrastructures in inland waterways presents significant challenges, primarily because the critical components for damage assessment are often submerged. Even though sensors like strain gauges can provide valuable insights into structural dynamic response, their effectiveness is affected by the limited number of sensors and how the sensors are placed. To address these limitations, Computer Vision (CV)-based vibration estimation methods have been developed, offering full-field measurements and overcoming the spatial constraints of sparse sensors. However, the efficacy of CV-based approaches declines drastically in underwater environments due to poor visibility conditions, which significantly affect image quality and measurement reliability. Motivated by overcoming the challenge of underwater SHM using CV-based methods, this paper proposes a Digital Twin (DT)-based approach for Underwater Image Enhancement (UIE) to enable accurate sub-pixel vibration estimation. First, a DT-based Physics-Based Graphics Model (PBGM) is employed to generate sufficient datasets for the various water scenarios, addressing the scarcity of underwater images and ground-truth labels. Then, an Edge-Preserving Dual Attention Transformer (EDUT) with an edge-preserving loss is designed to eliminate image degradation while retaining fine structural details. Finally, the sub-pixel template matching method is applied for vibration estimation. The performance of the proposed method is demonstrated through a case study using the Greenup miter gate in the United States, achieving an average PSNR of 27.81 dB for UIE and 0.0485 to 0.2319 sub-pixel accuracy in vibration estimation.

Keywords Digital twin · Computer vision · Underwater image enhancement · Vibration estimation

Introduction

Structures like dams and bridges in inland waterways are essential infrastructure components. However, aging, deterioration, and climate change-induced flooding threaten structural integrity, making cost-effective underwater inspection crucial [1]. Current underwater inspection primarily uses human divers, sonar scanning, and Computer Vision (CV) techniques for static surface defect detection [2]. While CV integrated with Unmanned Underwater Vehicles (UUVs) offers automated and convenient capabilities and eliminates safety risks [3], these surface-based methods cannot reveal internal structural damage, which requires vibration-based identification to capture dynamic characteristics and modal properties [4].

Existing underwater structural vibration detection research is extremely limited [5]. Terrestrial methods using contact sensors face complex underwater installation and corrosion issues, providing only localized measurements [6]. LiDAR systems suffer from water absorption and scattering effects, limiting their underwater effectiveness. CV methods offer potential

Enjian Cai · Zhen Hu(✉)
Department of Industrial and Manufacturing Systems Engineering, University of Michigan-Dearborn,
4901 Evergreen Road, Dearborn, MI 48128, USA
e-mail: enjcai@umich.edu; zhennhu@umich.edu

Zihan Wu · Michael D. Todd
Department of Structural Engineering, University of California San Diego,
9500 Gilman Drive Mail Code 0085, La Jolla, CA 92093, USA
e-mail: z5wu@ucsd.edu; mdtodd@ucsd.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings
of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

using UUVs with 2K+ cameras, but vibration identification requires sub-millimeter precision and sub-pixel measurement capabilities [7]. However, underwater images may be severely degraded by color casts, low contrast, and blurred details due to light effects. In this regard, existing Underwater Image Enhancement (UIE) methods face limitations: requiring extensive realistic datasets difficult to obtain in varying conditions [2], and relying on Cycle-Consistent Generative Adversarial Networks (CycleGAN) approaches with training instability, limited enhancement precision, and domain adaptation challenges [8].

To this end, this paper proposed a comprehensive Digital Twin (DT)-based underwater dataset generation method and an Edge-Preserving Dual Attention Transformer (EDUT) approach. Main contributions include: (1) enabling sub-pixel vibration measurement for underwater structural dynamic monitoring, (2) developing a Physics-Based Graphics Model (PBGM) creating the realistic three-dimensional (3D) underwater scenarios, and (3) introducing EDUT with Dual Attention Transformer Block (DATB) and edge-preserving loss function to address CycleGAN and Transformer limitations and preserve fine details [8, 9].

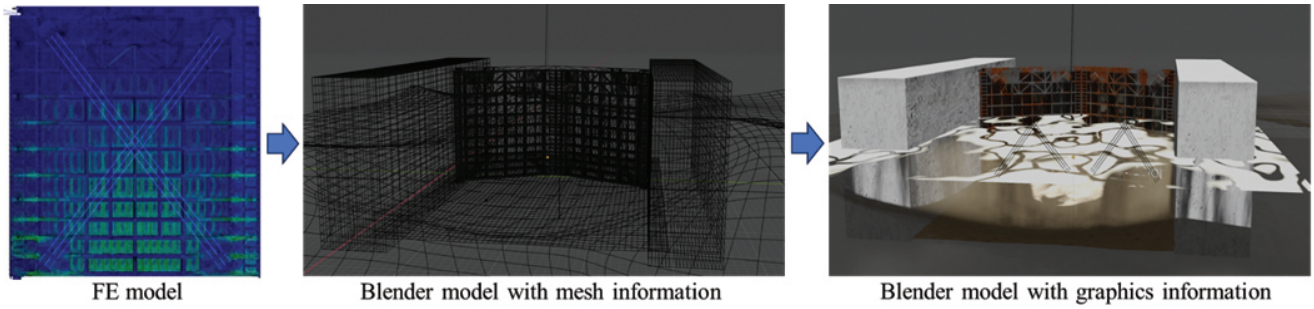


Fig. 1 Framework of DT-based underwater data generation

DT-Based Underwater Data Generation

The implementation of PBGM-based water-quality simulation follows a comprehensive methodology integrating structural modeling with aquatic environment reproduction. A Finite Element (FE) model of the underwater structure developed in Abaqus is first validated to capture structural behaviors, with geometric specifications input into Blender 4.0 using principled BSDF for material appearance. The framework, as shown in Figure 1, employs computational fluid-dynamics principles with physically accurate optical modeling to generate turbid water conditions based on the Greenup miter gate environment, representing sediment-laden waterways affected by erosion. The methodology involves four key stages. First, mesh validation uses Gerstner wave mathematics for surface topology. Second, the material setup applies warm base colors and volume scattering at 0.015 density. Third, lighting uses warm illumination at 4.0 energy units. Finally, turbid water simulation employs muddy substrates with sparse vegetation and 30 bubbles using increased drag coefficients (0.15) for sediment effects.

Edut Method for Effective UIE

The EDUT method architecture, as shown in Figure 2, employs a hierarchical encoder-decoder structure with three encoder layers and three decoder layers connected by skip connections for multi-scale feature processing. This design enables enhancement of objects at different underwater distances where scattering effects vary significantly. A bottleneck layer captures global pixel dependencies while maintaining computational efficiency for large-scale underwater structures.

EDUT addresses underwater degradation through three main components. Firstly, the Channel Self-Attention Transformer (CSAT) component handles wavelength-dependent color distortion by weighting channel importance, recognizing that red light penetrates only 3-4 m while blue-green light reaches deeper depths. This selective attention prevents single-channel dominance, crucial for detecting materials and corrosion in varying water conditions. Secondly, the Shifted Window Pixel Self-Attention Transformer (SW-PSAT) component addresses distance-dependent degradation through local window-based attention, enabling cross-window connections essential for continuous structural features like pipelines and hull sections that span multiple processing windows. The connection scheme and layers for CSAT and SW-PSAT are consistent with those in [10].

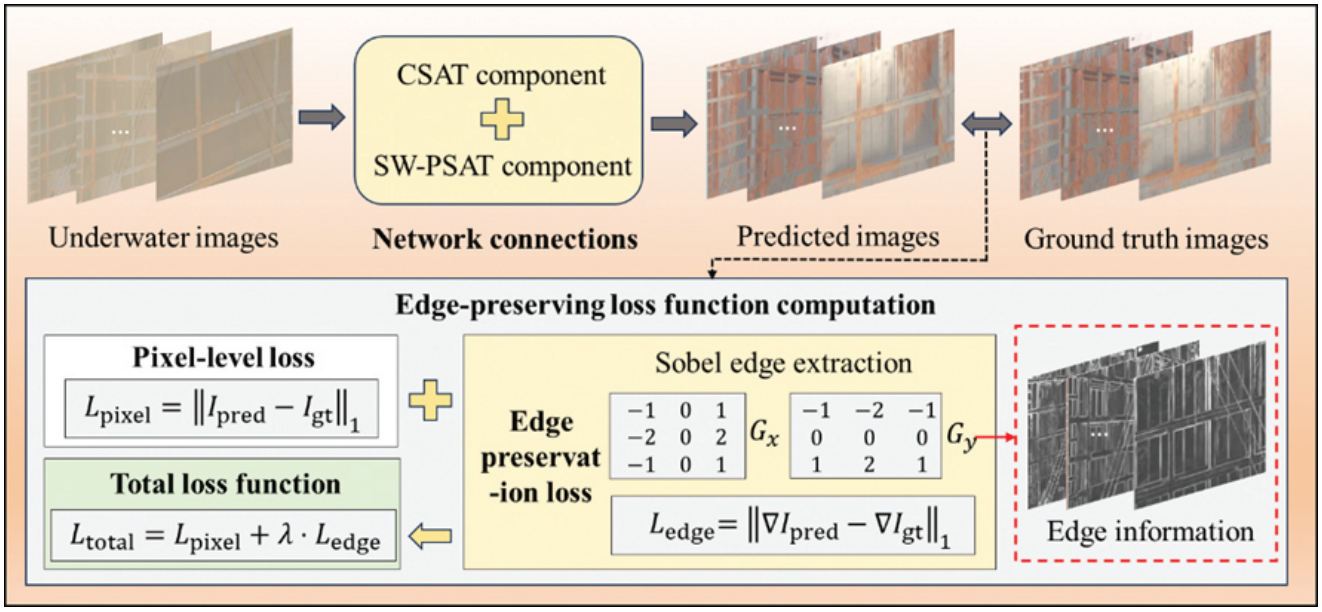


Fig. 2 Architecture of the EDUT method for effective UIE

Thirdly, the proposed edge-preserving loss function addresses the limitations of standard L_2 loss methods that treat all pixels equally, overlooking structural boundaries essential for underwater inspection. Underwater environments present unique degradation challenges: bubble occlusion creates false edges and obscures structural boundaries, scattering blur from suspended particles softens edges, reducing detection accuracy, and non-uniform degradation causes structural elements to appear with varying clarity depending on distance and water conditions.

To address these challenges, the proposed edge-preserving loss function incorporates both pixel-level reconstruction and structural boundary preservation. The pixel-level reconstruction loss $L_{\text{pixel}} = \|I_{\text{pred}} - I_{\text{gt}}\|_1$ uses L_1 norm for its robustness to outliers, while edge information is computed using the Sobel operator $\nabla I = \sqrt{(G_x * I)^2 + (G_y * I)^2}$ where G_x and G_y are horizontal and vertical Sobel kernels. The edge preservation loss $L_{\text{edge}} = \|\nabla I_{\text{pred}} - \nabla I_{\text{gt}}\|_1$ ensures the network maintains both edge strength and directional accuracy. The complete loss function $L_{\text{total}} = L_{\text{pixel}} + \lambda \cdot L_{\text{edge}}$ combines these components, where the weight parameter $\lambda = 0.2$ is carefully calibrated to balance edge preservation without introducing artificial sharpening that could be mistaken for structural defects. This formulation forces the EDUT network to maintain sharp, geometrically accurate structural boundaries while preserving fine details crucial for sub-pixel underwater vibration detection.

Sub-Pixel Underwater Structural Vibration Estimation

After EDUT enhancement, sub-pixel template matching is used to estimate structural vibrations from the enhanced video sequence, where the first frame serves as a reference, as shown in Figure 3. Mean Squares (MS) is employed as the similarity

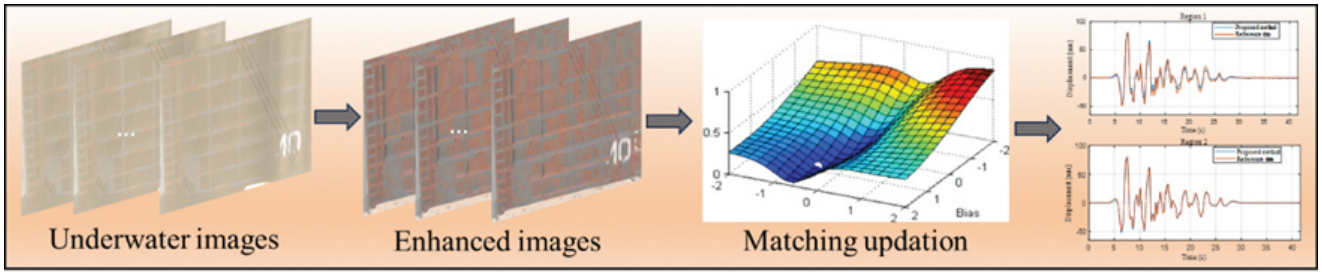


Fig. 3 Architecture of the sub-pixel template matching for underwater structural vibration estimation

metric for its superior noise resistance in intensity-based matching applications. The MS between reference frame R_0 and current frame R_i incorporates spatial displacement parameters to quantify structural movement. The similarity metric is computed as $\mathrm{MS}(u,v) = (1/N)\Sigma\Sigma\left[R_0(x,y) - R_{\mathrm{i}}(x+u,y+v)\right]^2$, where u and v represent horizontal and vertical structural vibrations in pixels, $R_0(x,y)$ and $R_i(x+u,y+v)$ are intensity values of corresponding structural regions, and N is the number of pixels in the overlapping area. This formulation captures template correlation across frames, accounting for translational displacement. Since structures primarily undergo rigid-body translation, optimization focuses only on translational parameters. The optimal structural vibrations with units of pixels are determined by minimizing the MS metric $(\hat{u},\hat{v}) = \arg\min \mathrm{MS}(u,v)$ to identify displacement values with sub-pixel accuracy. Physical structural vibrations are then obtained by multiplying pixel vibrations with calibration factors derived from known structural dimensions.

Case Study

The proposed method was validated using the Greenup miter gate on the Ohio River, Kentucky, United States. A high-fidelity graphical model was generated in Blender 4.0 from the validated FE model, with yellow turbid water conditions simulating the gate's actual environment, as already shown in Figure 1. Virtual camera trajectories were systematically planned to simulate UUV navigation, capturing both gate sides from multiple viewpoints with 3000 frames at 1080×1920 resolution and 24 Hz sampling frequency. Comprehensive comparisons were conducted with state-of-the-art approaches: enhanced Underwater (U)-GAN [11] and U-Shape Transformer (UST) [9] method. Additionally, the EDUT trained on the authoritative UIE Benchmark (UIEB) dataset [12] was compared to assess the advantages of the proposed dataset. All networks were trained for 120 epochs with identical settings, using 256×256 resolution during training. For vibration validation, structural excitations along the Z-axis were applied to both gates using the CNP196 component of the 1994 Northridge earthquake [13]. Testing videos comprised 1001 frames captured at the same resolution and sampling frequency, with the initial frames of the vibration videos and two measurement regions shown in Figure 4. Pixel-to-physical conversion factors were calculated as 9.0765 mm/pixel and 11.0869 mm/pixel for left and right gates, respectively. All experiments were conducted on a workstation with an Intel Xeon W-2125 CPU, 64 GB RAM, and dual NVIDIA GeForce RTX 2080 Ti GPUs.

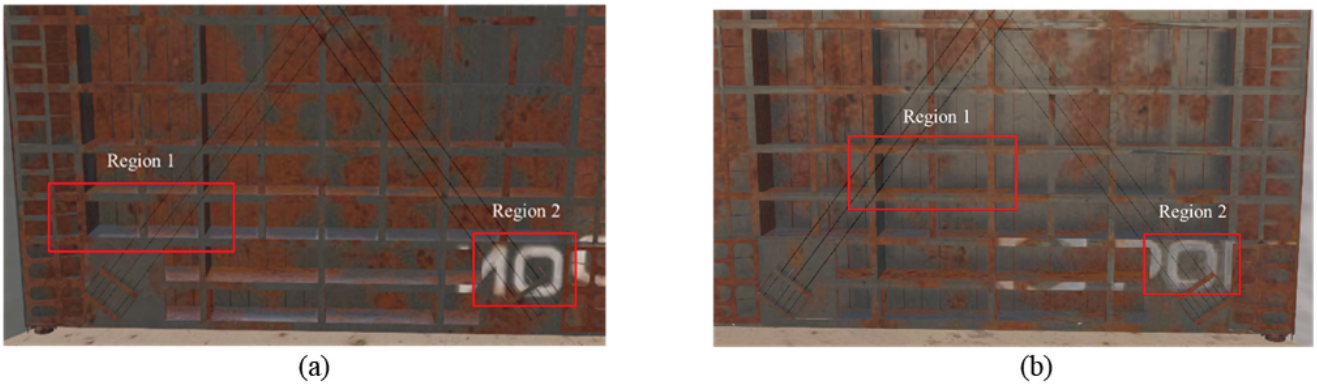


Fig. 4 Rendered frames recorded by two virtual cameras during the vibration at the 1st frame: (a) left gate and (b) right gate

After generating the yellow turbid 3D underwater environment, two virtual cameras were positioned along navigation routes to capture videos, producing 3000 frames per gate at a 24 Hz sampling frequency. The corresponding underwater videos and ground truth sequences are shown in Video 1. It could be observed that, with PBGM integration, the generated dataset realistically reflected variations in water color, bubble behavior, and lighting as the camera location and viewpoint changed. These images were used to train four models: DT-trained U-GAN, DT-trained UST, UIEB-trained EDUT, and DT-trained EDUT. The resulting models were applied to vibration videos of each gate in the yellow turbid environment, each containing 1001 frames. To save pages, only the left gate results are shown in Figure 5, with only the 500th frame of enhanced images shown, while full sequences enhanced by DT-trained EDUT for both left and right gates are available in Video 2 and Video 3. As observed, DT-trained U-GAN and UST enhanced images were close to the ground truth but failed to fully recover gate textures. U-GAN results appeared blurred due to high-frequency detail smoothing to satisfy the

discriminator, while UST outputs were overly bright because global brightness adjustment with patch-level attention reduced local control and amplified uneven illumination. UIEB-trained EDUT removed only a small portion of water effects due to dataset limitations. In contrast, DT-trained EDUT produced enhanced images closely matching ground truth with well-preserved gate edge textures, achieved through PBGM-based dataset generation and shifted-window pixel-level attention with edge-preserving loss. In addition, DT-trained EDUT reached the highest Peak Signal-to-Noise Ratio (PSNR) values, averaging 27.81 dB.

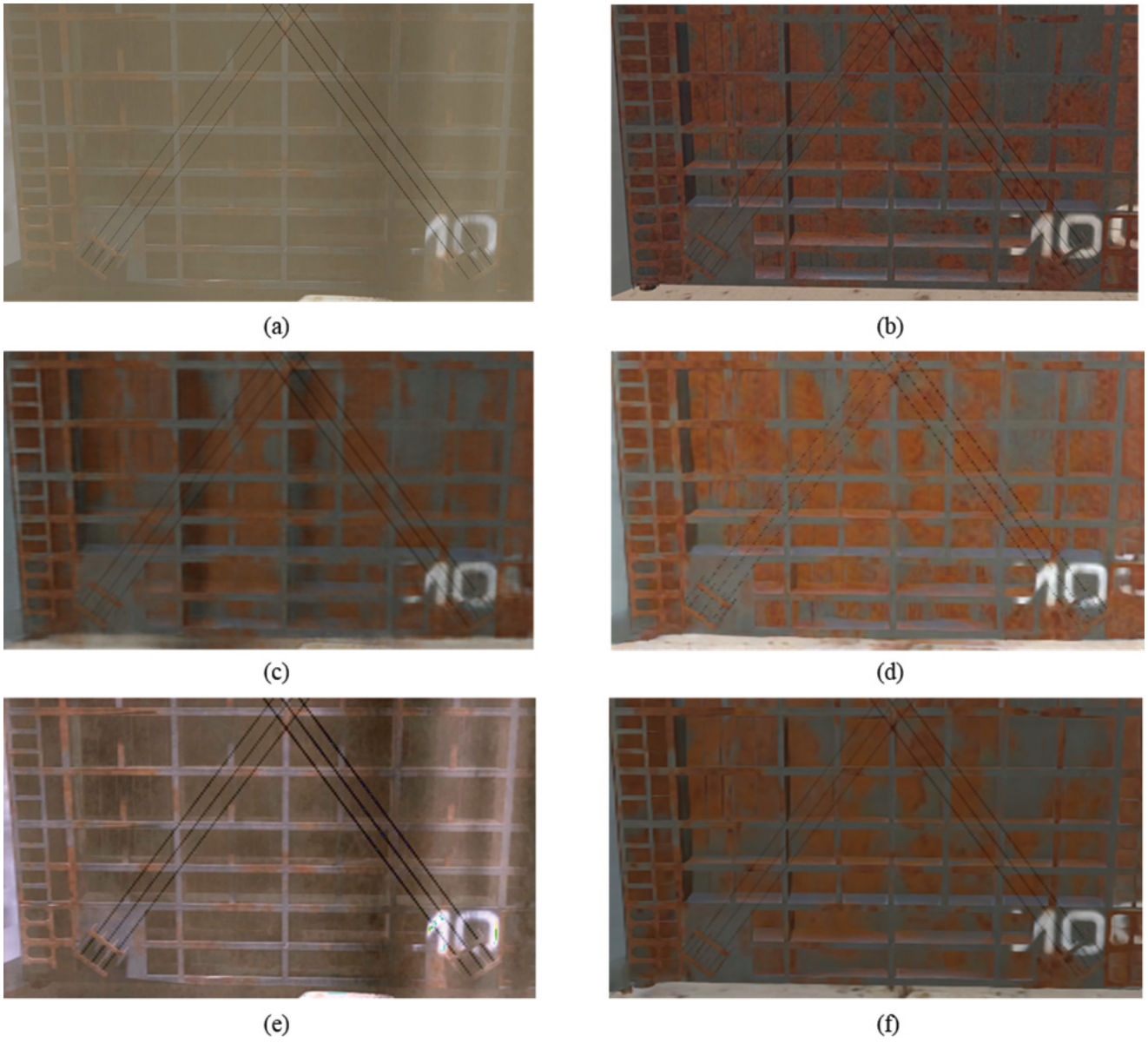


Fig. 5 The 500th recorded and enhanced frames from the 3D yellow turbid underwater environment when the left gate was vibrating: (a) raw underwater image, (b) ground truth image, and images enhanced by models of (c) DT-trained U-GAN, (d) DT-trained UST, (e) UIEB-trained EDUT, and (f) DT-trained EDUT

Then, sub-pixel template matching was applied to the enhanced videos to estimate structural vibrations, with resulting time-history curves shown in Figure 6, Video 4, and Video 5. Only videos enhanced by DT-trained EDUT were used for estimation due to superior performance. The estimated vibrations closely matched reference data, with Correlation Coefficients (CCs) above 0.9790 in all regions and reaching 0.9970 in Region 2. Mean Absolute Errors (MAEs) were 0.4400 mm and 0.8342 mm in Region 2, and 2.1326 mm and 2.5705 mm in Region 1, corresponding to sub-pixel accuracy of approximately 0.0485 to 0.2319 pixels based on conversion factors of 9.0765 mm/pixel and 11.0869 mm/pixel for left and right gates,

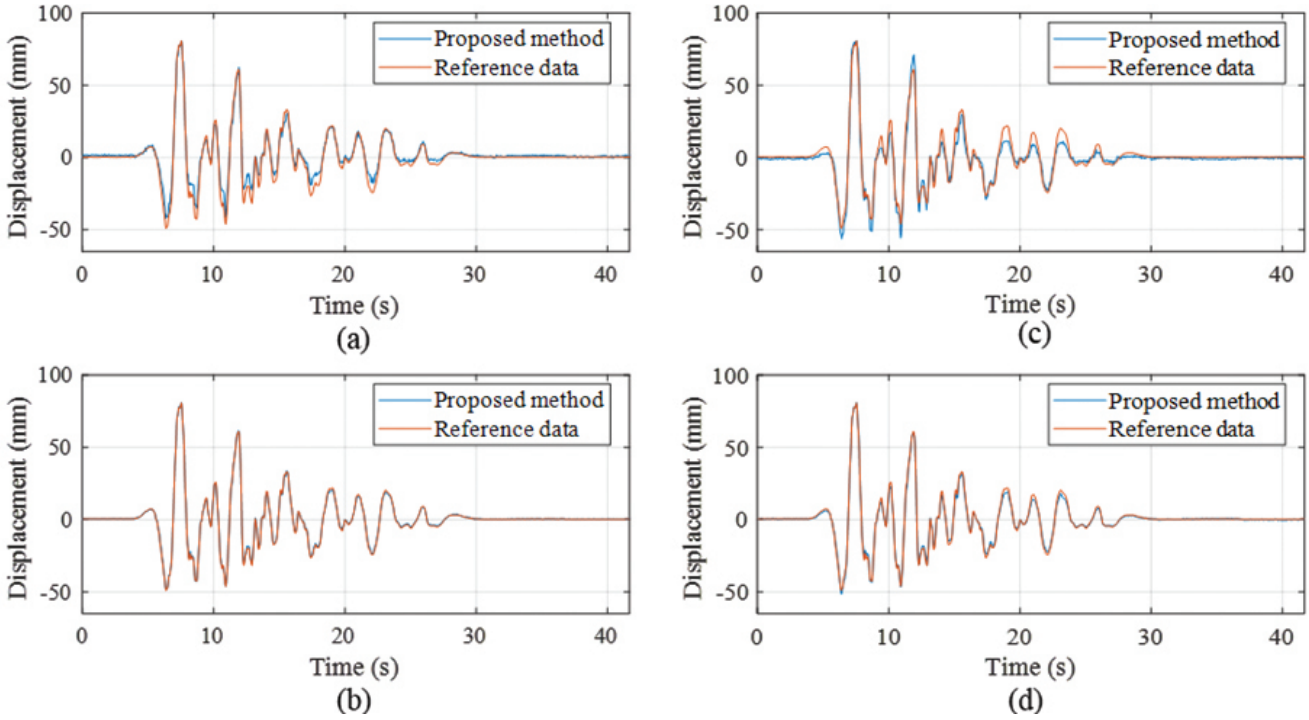


Fig. 6 Estimated structural vibrations from the 3D yellow turbid underwater environment in (a) Region 1, (b) Region 2 of the left gate, and (c) Region 1, (d) Region 2 of the right gate

respectively. The accuracy of estimated structural vibrations in Region 2 showed higher accuracy than that in Region 1, due to less complex textures that facilitated better UIE and reduced measurement artifacts.

Conclusions

In this paper, a DT-based method was proposed for effective underwater sub-pixel structural vibration estimation, by addressing severe image degradation in underwater scenes. A DT-based data generation method integrated with PBGM was proposed to generate sufficient underwater datasets for yellow turbid water conditions characteristic of sediment-laden rivers. Then, EDUT was proposed for effective UIE while preserving detailed edges, and sub-pixel template matching was applied to estimate structural vibrations with high precision. Through application on the Greenup miter gate under the yellow-turbid water condition, the proposed method could achieve an average PSNR of 27.81 dB for UIE with structural-vibration accuracy of approximately 0.0485 to 0.2319 sub-pixel. Future studies will extend the proposed method for out-of-plane vibration identification and elimination of camera motions from water surface disturbances.

Acknowledgments This work was supported by the US Army Corps of Engineers through the U.S. Army Engineer Research and Development Center Research Cooperative Agreement W9132T-22-20014. The support is gratefully acknowledged.

References

1. H. Tan, L. Zheng, C. Ma, Y. Xu, Y. Sun, Deep learning-assisted high-resolution sonar detection of local damage in underwater structures, Automation in Construction, 164 (2024) 105479.
2. S. Talamkhani, K. Liu, Underwater vision-enhanced image segmentation for supporting automated inspection of underwater bridge components, Automation in Construction, 175 (2025) 106230.
3. Z. Wu, Z. Hu, M.D. Todd, Uncertainty Quantification for Deep Learning–Based Automatic Crack Detection in the Underwater Environment, Springer Nature Switzerland, Cham, 2025, pp. 133-136.
4. E. Cai, Y. Zhang, X. Lu, P. Li, T. Zhao, G. Lin, W. Guo, A target-free video structural motion estimation method based on multi-path optimization, Mechanical Systems and Signal Processing, 198 (2023) 110452.

5. T. Chu, T. Nguyen, H. Yoo, J. Wang, A review of vibration analysis and its applications, Heliyon, 10 (2024) e26282.
6. J. Li, G. Chen, M.F. Antwi-Afari, Recognizing sitting activities of excavator operators using multi-sensor data fusion with machine learning and deep learning algorithms, Automation in Construction, 165 (2024) 105554.
7. E. Cai, Y. Zhang, S.T. Quek, Visualizing and quantifying small and nonstationary structural motions in video measurement, Computer-Aided Civil and Infrastructure Engineering, 38 (2023) 135-159.
8. Z. Lv, S. Dong, Z. Xia, J. He, J. Zhang, Enhanced real-time detection transformer (RT-DETR) for robotic inspection of underwater bridge pier cracks, Automation in Construction, 170 (2025) 105921.
9. L. Peng, C. Zhu, L. Bian, U-Shape Transformer for Underwater Image Enhancement, IEEE Transactions on Image Processing, 32 (2023) 3066-3079.
10. Z. Shen, H. Xu, T. Luo, Y. Song, Z. He, UDAformer: Underwater image enhancement based on dual attention transformer, Computers & Graphics, 111 (2023) 77-88.
11. C. Fabbri, M.J. Islam, J. Sattar, Enhancing Underwater Imagery Using Generative Adversarial Networks, 2018 IEEE International Conference on Robotics and Automation (ICRA), 2018, pp. 7159-7165.
12. C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, D. Tao, An Underwater Image Enhancement Benchmark Dataset and Beyond, IEEE Transactions on Image Processing, 29 (2020) 4376-4389.
13. A. Singh, T. Hutchinson, X. Wang, Z. Zhang, B. Schafer, F. Derveni, H. Castaneda, K. Peterman, Wall Line Tests: Phase 1 – Shake Table Tests, Designsafe-CI, 2023.

Chapter 9
Evolution of Modal Characteristics and Model Comparison for a 10-Story Cold-Formed Steel Framed Building under Seismic Excitation

Jiachen (Charlie) Zhang, Shokrullah Sorosh, Amanpreet Singh, Benjamin W. Schafer, and Tara C. Hutchinson

Abstract A 10-story cold-formed steel-framed building, coined CFS10, was constructed and tested at the 6-DOF Large High-Performance Outdoor Shake Table (LHPOST6) at UC San Diego in the summer of 2025. Seismic tests targeted Service-Level Earthquake (SLE), Design Earthquake (DE) and risk-targeted Maximum-Considered Earthquake (MCE_R) hazards and were followed by live fire compartment tests. This landmark full-scale building specimen was designed beyond current code limits with advances in steel sheet sheathed shear walls and heavy chord stud details proofed in complementary component test programs, while also incorporating a wide range of realistic nonstructural systems (stairs, fire sprinkler, gas piping, glazing, etc.). The shake table testing captures the accumulation of damage to the building and its nonstructural elements during realistic seismic scenarios. In preparation for predicting building performance and to aid in ground motion selection and test protocols, a practice-oriented, phenomenological finite element (FE) model was developed. This model aimed for efficient simulation run speeds, with only essential degrees-of-freedom defined, while still seeking to capture the salient features of the specimen's behavior to earthquake excitation. This paper discusses the modal characteristics of the CFS10 building and how those characteristics evolve as damage accumulates during the seismic testing phase. Importantly, a preliminary comparison is made between the experimentally extracted modal characteristics, which evolve during the test program, and the aforementioned FE model predictions.

Keywords Cold-formed steel · Shake table test · Modal characteristics · Damage accumulation · Numerical model predictions

Introduction

As a capstone to the National Science Foundation (NSF) primary sponsored project on cold-formed steel (CFS) framed buildings, a full-scale 10-story CFS building, coined CFS10, was tested in the summer of 2025 at the 6-degree-of-freedom Large High-Performance Outdoor Shake Table (LHPOST6) [1] at the University of California San Diego. The landmark CFS10 building specimen was designed to above-code height limits with 10-stories and a total height of 30.5 m and a floor plan measuring 6.86 × 11.58 m, see Fig 1. The primary goal of this test program was to evaluate the seismic performance of a tall CFS building with steel sheet sheathed shear walls exceeding the ASCE 7-22 [2] height limit of 19.8 m, or 6 stories. A detailed discussion on the seismic performance of the building is provided in Hutchinson et al. [3]. A summary of the design details of the structural system of this building is provided in Singh et al. [4]. In addition, this building was outfitted with various nonstructural components and systems, including a resilient stair system [5, 6], a fire riser and sprinkler system, suspended ceiling [7], gas piping system, windows, doors, and rooftop equipment [8].

Jiachen (Charlie) Zhang · Shokrullah Sorosh · Tara C. Hutchinson
Department of Structural Engineering University of California San Diego, La Jolla, CA 92093
e-mail: jiz149@ucsd.edu; ssorosh@ucsd.edu; tara@ucsd.edu

Amanpreet Singh · Benjamin W. Schafer
Department of Civil & Systems Engineering Johns Hopkins University, Baltimore, MD 21218
e-mail: amanprs@jhu.edu; schafer@jhu.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

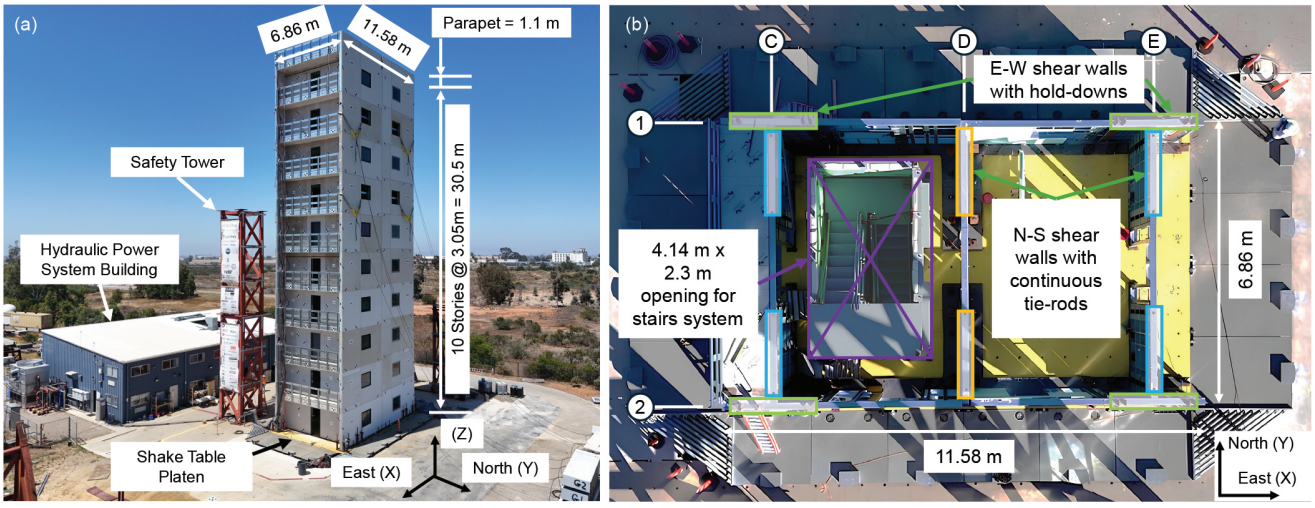


Fig. 1 CFS10 building (a) isometric view and (b) plan view

Seismic Test Protocol

Table 1 presents the seismic test protocol for the CFS10 program. A detailed discussion on the development of the seismic test protocol is provided in Zhang et al. [9]. Low-amplitude white noise (WN) tests were conducted between input ground motions to capture the evolution of the building's dynamic characteristics. Fig 2 presents the details of the test days, including WN and seismic excitations in chronological order. In total, 18 ground motions, including 15 targeting Service-Level Earthquake (SLE), one targeting Design Earthquake (DE), and two targeting Risk-Targeted Maximum-Considered Earthquake (MCE_R) hazards, as well as 37 WNs were applied to the test specimen over a duration of 7 test days.

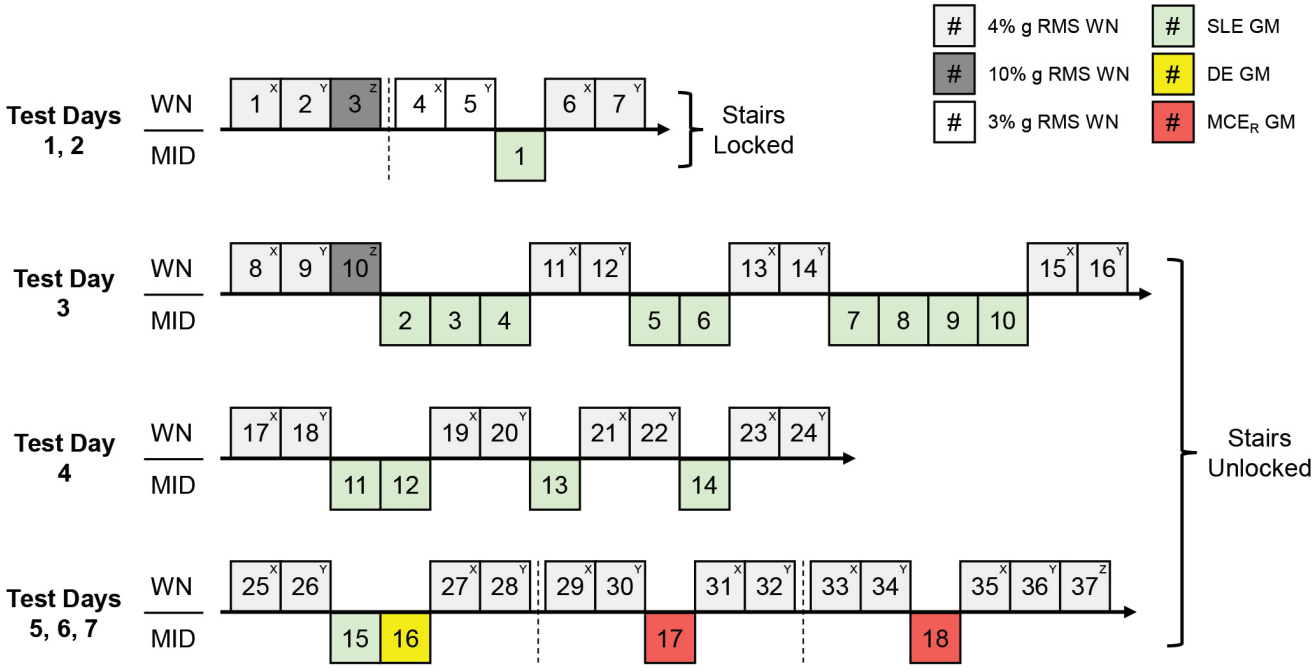


Fig. 2 CFS10 seismic test days in chronological order

Table 1 CFS10 seismic test protocol

Hazard level	MID	Earthquake event				Input DOF	Components		$D_{5\text{-}95}$ [sec]	PIA_x [g]	PIA_y [g]	PIA_z [g]
		Name, location	Year	M_w	Station		X	Y				
SLE	1*	Loma Prieta, California	1989	6.9	Gilroy Array #3	XYZ	90°	0°	11.4	0.126	0.122	0.082
	2					XYZ	90°	0°	11.4	0.118	0.114	0.096
	3					X	90°	-	11.4	0.133	0.041	0.025
	4					Y	-	0°	6.4	0.042	0.109	0.024
	5	Northridge, California	1994	6.7	LA City Terrace	XYZ	90°	180°	11.9	0.125	0.121	0.081
	6					XYZ	180°	90°	11.9	0.184	0.094	0.084
	7					X	180°	-	11.9	0.168	0.027	0.023
	8					Y	-	90°	11.6	0.026	0.093	0.021
	9	Cape Mendocino, California	1992	7.0	Rio Dell Overpass	XYZ	270°	360°	15.3	0.084	0.109	0.099
	10					XYZ	360°	270°	15.3	0.078	0.122	0.151
	11**	Kahramanmaras, Türkiye	2023	7.8	4611	XYZ	E	N	43.4	0.095	0.075	0.050
	12					XYZ	E	N	43.4	0.117	0.110	0.064
	13	Elbistan, Türkiye	2023	7.5	4611	XYZ	E	N	35.1	0.100	0.061	0.052
	14	Maule, Chile	2010	8.8	Curico Hospital	XY	EW	NS	45.2	0.097	0.091	0.033
	15	Kahramanmaras, Türkiye	2023	7.8	4611	XYZ	E	N	43.4	0.119	0.089	0.063
DE	16	Loma Prieta, California	1989	6.9	Gilroy Array #3	XY	90°	0°	11.4	0.454	0.388	0.014
MCE_R	17	Loma Prieta, California	1989	6.9	Gilroy Array #3	XY	90°	0°	11.4	0.723	0.619	0.016
	18	Cape Mendocino, California	1992	7.0	Rio Dell Overpass	XY	360°	270°	15.3	0.782	0.336	0.025

SLE = service level earthquake, DE = design earthquake, and MCE_R = risk-targeted maximum considered earthquake.

PIA = peak input acceleration (achieved) using shake table platen center of mass accelerometers

$D_{5\text{-}95}$ = significant duration of strong shaking computed as the time interval between 5% and 95% Arias Intensity per original record (max component)

*MID1 was conducted with the stairs in a locked configuration

**MID11 target to 75% SLE

Description of the Practice-oriented Finite Element Model

In preparation for predicting building performance and to aid in ground motion selection and test protocols, a pre-test practice-oriented, phenomenological finite element (FE) model, as shown in Fig 3, was developed. This model was developed with only essential degrees of freedom to facilitate efficient simulation run times, while also aiming to capture the salient features of the specimen's behavior to earthquake excitation. Shear walls and gravity walls were modeled to match full hysteretic behavior with both backbone degradation and pinching. A detailed description of this FE model, which was developed in OpenSeesPy [10], can be found in Zhang et al. [11]. In addition to the features presented in [11], zero-length shear springs that reflect the stiffness of the stairs were added to the center of the stairs to model the additional stiffness and strength contributions from the stairs. The details of modeling the stairs system can be found in [5].

FE Model Eigenvalue Analyses Considering Sequential Run-style Analysis

As-achieved acceleration histories of the input motions during the "stairs unlocked" configuration (i.e., MID2 through MID18) were used as ground motion inputs for the FE model and sequentially run to replicate the actual sequence of how these motions were applied to the building. That is, the model state was preserved following individual nonlinear time

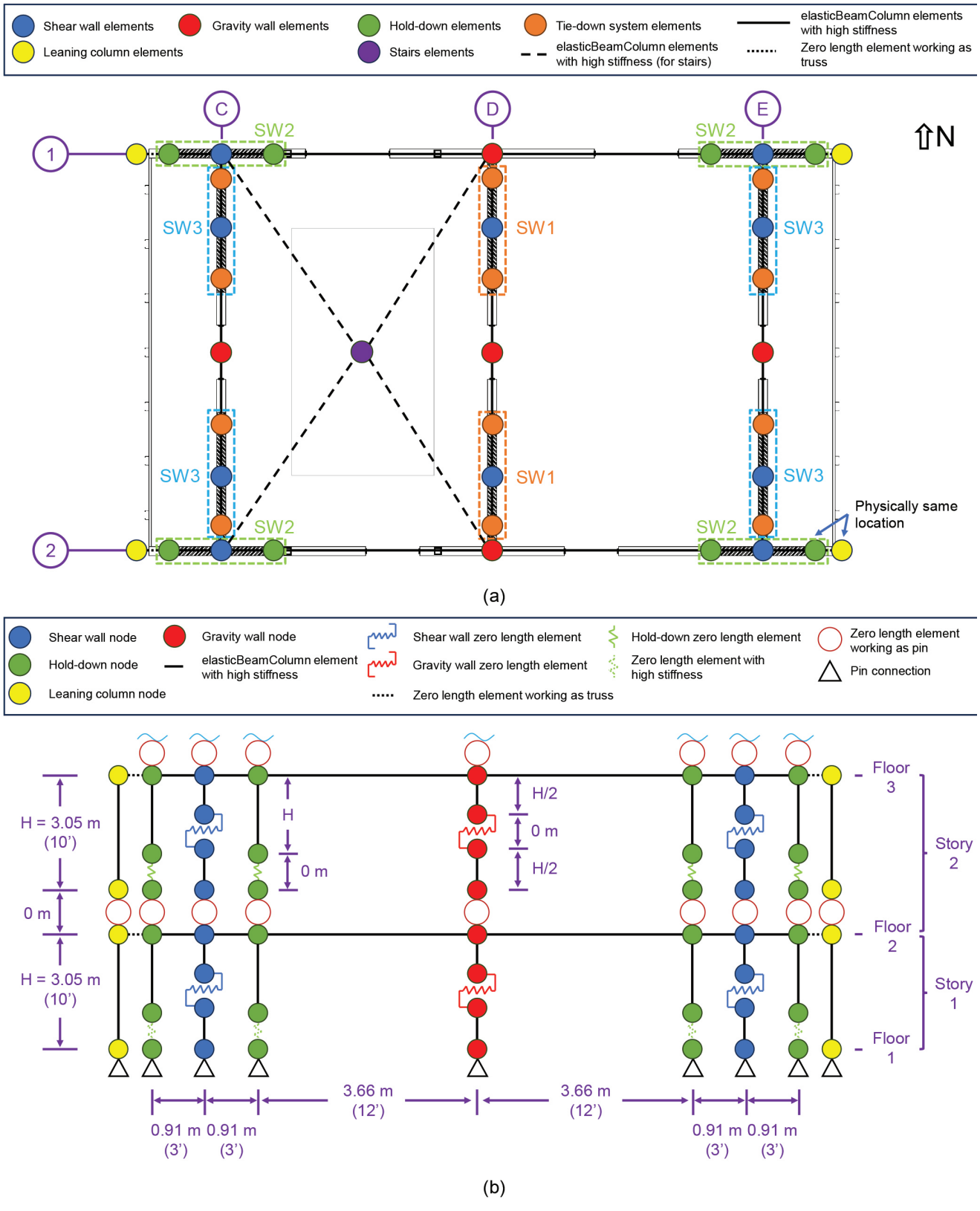


Fig. 3 Schematic (a) plan view and (b) representative elevation view of wall lines 1 and 2 of FE model

history analyses, allowing the accumulated damage to be captured through the modeling of the shear walls and gravity walls. Following each ground motion run, an additional 30 seconds were introduced in the numerical model to allow free vibration of the building model to dampen out prior to performing an eigenvalue analysis. The eigenvalue analyses were performed with the aim of replicating the system identification process, which during testing was achieved using low-amplitude unidirectional white noise tests, following seismic excitations, see Fig 2. Fig 4 summarizes the building modal period evolution for the first three fundamental modes, i.e., 1st longitudinal (1-X), 1st transverse (1-Y), and 1st torsional (1-θ). These analyses capture the general gradual increase in periods across the primary modes anticipated during sequentially increasing earthquake demand, with the largest increases, as anticipated, during the MCE_R simulated MID17 and MID18.

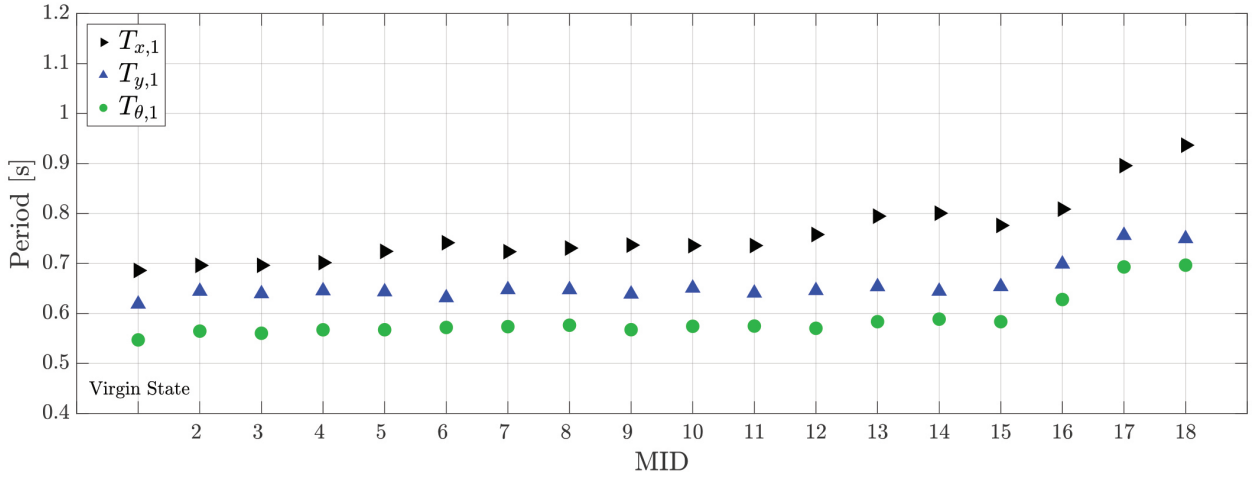


Fig. 4 Building modal period evolution based on FE model eigenvalue analyses results

Experimental System Identification (SI) Results

Using WN test data, the building period was estimated using the Stochastic Subspace Identification (SSID) technique [12]. In this analysis, 40 channels of acceleration time series data recorded at the estimated center of mass and the northwest corner of the building at each floor, and four channels of acceleration data on the shake table platen were used. In the identification of the period corresponding to the flexural-dominant (X-direction) and torsional modes, the data from WN tests running in the X-direction were used. Similarly, to determine the period of the flexural-dominant (Y-direction) mode, the data from the WN tests running in the Y-direction were used. Fig 5 summarizes the building's periods corresponding to the first three vibration

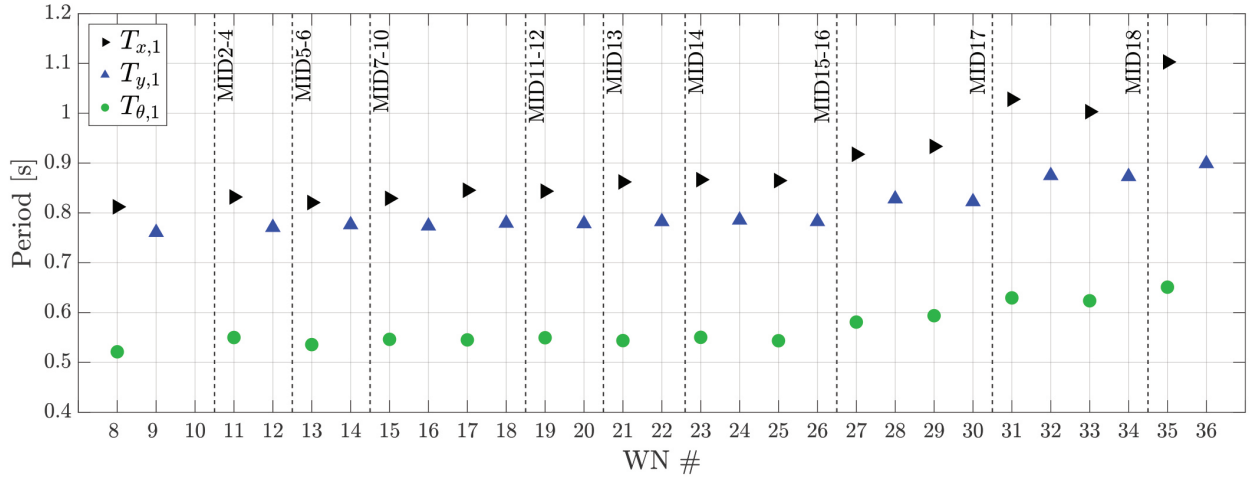


Fig. 5 Building modal period evolution based on experimental SI results

modes throughout the "stairs unlocked" configuration test phase. Interested readers may find useful the preliminary period evolution results throughout the entire test sequence documented in [3] and the detailed discussion on the impact of stair locking on the vibration period of the building in [6]. These analyses, similar to the FE model, capture the generally increasing amplitude trend in the primary modes, with larger amplitude increases following MCE_R excitations of MID17 and MID18.

FE Model Eigenvalue Analyses and Experimental SI Results Comparison

Figs. 6 and 7 compare the evolution of building periods as predicted by the sequential FE model eigenvalue analyses and experimentally determined SI results, shown as absolute values (Fig 6) and percentage change (Fig 7). It can be observed that the FE model consistently predicts smaller periods than the SI-based building periods for both the longitudinal (X) and transverse (Y) modes, reflecting stiffer lateral stiffness assumptions in the FE model, whereas the FE model consistently produces larger periods than the SI-based results for the torsional (θ) modes. The deviations between the FE-model-based results and SI-based results are much smaller for the torsional mode than the translational modes.

Expressing the period evolution as percentage change results in improved comparison between the FE-model-based and SI-based results. Generally, the FE-model-based results produce larger percentage change than the SI-based results, although the difference remains within 10%. In the longitudinal (X) direction, the percentage change after MID4 (using WN11 for

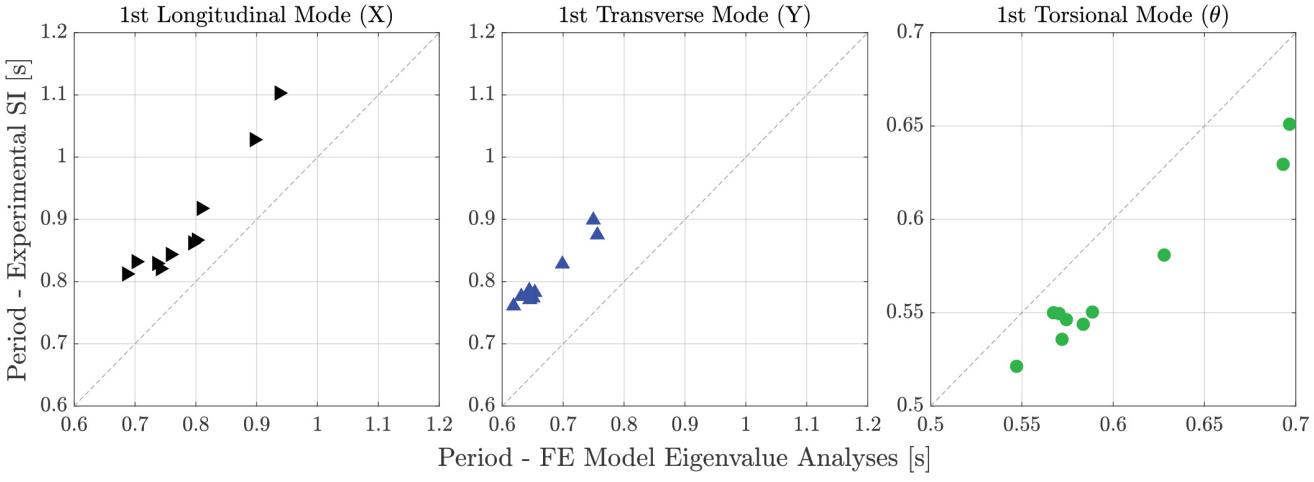


Fig. 6 Period evolution comparison between experimental SI-based and FE-model-based results

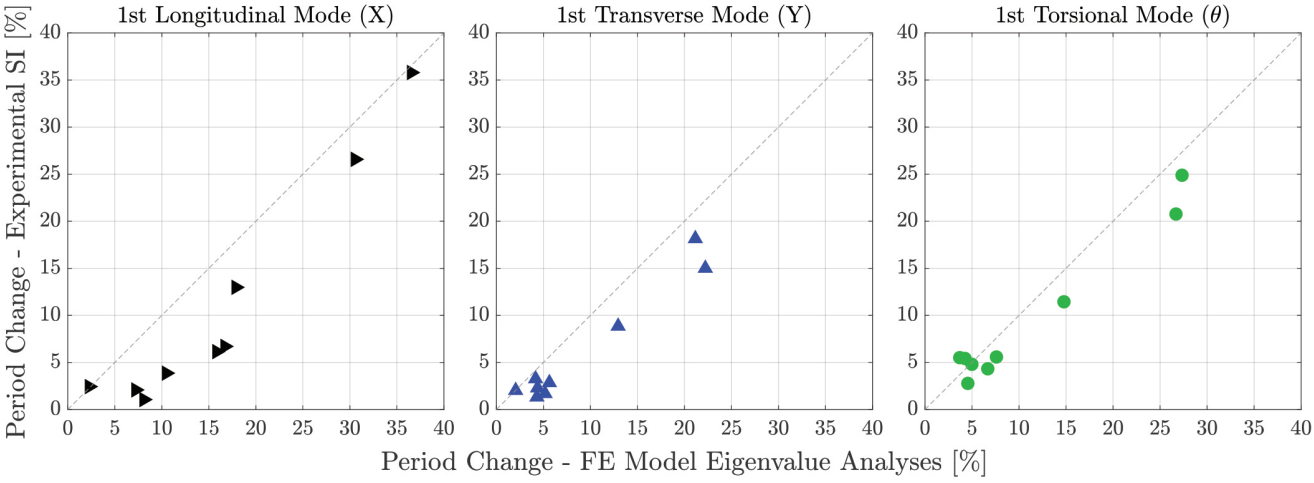


Fig. 7 Percent change in period, compared with virgin state, as estimated using experimental data (SI-based, WN) and numerical data (FE-model-based, eigenvalue analyses)

SI) and accumulated percentage change after MID18 (using WN35 for SI) were almost the same, indicating a very robust performance of the FE model in terms of capturing accumulated damage. In the transverse (Y) direction, the percentage change after MID6 (using WN14) was almost identical, while the difference between the FE-model-based results and SI-based results diverges as the motion intensities increase.

Concluding remarks

This paper examines the evolution of modal characteristics of the CFS10 building as damage accumulates under seismic excitations with increasing intensity. A preliminary comparison is presented between the experimentally extracted modal characteristics, obtained from the utilization of the extensive instrumentation of the test specimen, and predictions from a phenomenological FE model. Such comparison indicates that while the current FE model yields larger lateral stiffness assumptions, which signals the need for further study of uncertainties in material properties definitions in the initial development of the FE model as well as considering as-built conditions of the building, the FE model can reasonably capture the accumulated damage in terms of the relative increase in the fundamental building periods.

Acknowledgments This program is supported by the National Science Foundation (NSF) grants CMMI 1663569 and 1663348 and numerous other federal/state and industry partners, see: cfs10.ucsd.edu/sponsors-partners. The efforts of all undergraduate and graduate students and visiting scholars listed at (cfs10.ucsd.edu/people) as well as the LHPOST6 facility staff (esec.ucsd.edu/about/personnel) are greatly appreciated. Findings, opinions, and conclusions are those of the authors and do not necessarily reflect those of the sponsoring organizations.

References

1. McCartney, J. S., Van Den Einde, L., Morrison, M., Palermo, A., Lotfizadeh, K., Lai, C.-T., and Conte, J. P. (2025). Multi-degree of freedom shake table testing of large-scale structural and geotechnical systems with NHERI-UCSD LHPOST6. *Frontiers in the Built Environment*, 11:1573390. DOI: https://doi.org/10.3389/fbuil.2025.1573390.
2. ASCE/SEI (2022). *Minimum Design Loads and Associated Criteria for Buildings and Other Structures*, ASCE/SEI 7-22. American Society of Civil Engineers, Reston, VA.
3. Hutchinson, T. C., Eladly, M., Ji, R., Rivera, D., Rukundo, F. C., Singh, A., Sorosh, S., Zhang, J., Zhang, X., Padgett, L., Jones, H., Karns, J., and Schafer B.W. (2026). Advancing use of cold-formed steel framing systems for mid-rise residential construction in seismic zones via full-scale shake table testing. *Proceedings of the 8th Residential Building Design & Construction Conference*, State College, PA.
4. Singh, A., Zhang, J., Rivera, D., Eladly, M., Jones, H., Kovac, A., Padgett, L., Rivera, D., Smith, K., Torabian, S., Peterman, K., Schafer, B. W., and Hutchinson, T. C. (2024). CFS-NHERI 10-Story Building Shake Table Test Specimen: Design Updates. *Proceedings of the 18th World Conference on Earthquake Engineering (WCEE2024)*, Milan, Italy.
5. Sorosh, S., Zhang, J., Kovac, A., Smith, K., and Hutchinson, T. C. (2025). Design Strategies and Modeling Verification for Steel Stairs with Drift-Compatible Connections. *Proceedings of the 2025 Structural Engineers Association of California (SEAOC) Convention*, San Diego, CA.
6. Sorosh, S., Zhang, J., Kovac, A., Smith, K., Schafer, B. W., and Hutchinson, T. C. (2026). Influence of Stair System Flexibility on Building Vibration Period: Insights from Full-Scale Building Shake Table Tests at Low-Amplitude Excitation. *Proceedings of the 6th International Workshop on the Seismic Performance of Non-Structural Elements (SPONSE)*, Yokohama, Japan.
7. Singh, A., Rukundo, F. C., Ameen, J. R., Fathali, S., Jaramillo, J., Torbin, B., Huntress, J., Szakats, G., Cook, D. T., Speicher, M. S., Sattar, S., Hutchinson, T. C., and Schafer, B. W. (2026). Interaction of Suspended Ceilings and Pressurized Piping in a Full-Scale 10-Story Cold-Formed Steel Building Shake Table Test Program. *Proceedings of the 6th International Workshop on the Seismic Performance of Non-Structural Elements (SPONSE)*, Yokohama, Japan.
8. Zhang, X., Zhang, J., and Hutchinson, T. C. (2026). Shake Table Testing of Rooftop Mechanical Equipment with Different Attachment Strategies on A Full-Scale Ten-Story Building. *Proceedings of the 6th International Workshop on the Seismic Performance of Non-Structural Elements (SPONSE)*, Yokohama, Japan.
9. Zhang, J., Sorosh, S., Ji, R., Lotfizadeh, K. H., Schafer, B. W., and Hutchinson, T.C. (2026). Full-Scale 10-Story Cold-Formed Steel Building Shake Table Test: Protocol Development and Shake Table Performance. *Proceedings of the 13th National Conference on Earthquake Engineering*, Portland, OR.
10. Zhu, M., McKenna, F., and Scott, M. H. (2018). OpenSeesPy: Python library for the OpenSees finite element framework. *Software X*, 7: 6-11. DOI: https://doi.org/10.1016/j.softx.2017.10.009.
11. Zhang, J., Singh, A., Eladly, M., Schafer, B. W., and Hutchinson, T. C. (2024). Pre-Test Numerical Modeling of the CFS-NHERI 10-story Capstone Building. *Proceedings of the 18th World Conference on Earthquake Engineering (WCEE2024)*, Milan, Italy.
12. Van Overschee, P., and De Moor, B. (1996). *Subspace Identification for Linear Systems: Theory–Implementation–Applications*. Kluwer Academic Publishers, Norwell, MA.

Chapter 10
Modal Identification of Rooftop Mechanical Equipment with Different Attachment Designs on a Full-Scale Ten-Story Shake Table-Tested Building

Xianzhao Zhang, Jiachen Zhang, Shokrullah Sorosh, and Tara C. Hutchinson

Abstract Modal properties are critical for estimating the seismic acceleration and displacement demands of nonstructural components, whose satisfactory performance during earthquakes can help ensure the continued functionality of buildings. While vibration-based system identification techniques have advanced significantly for large structural systems, their application to real-world nonstructural components subjected to increasing earthquake-induced load scenarios remains limited due to a lack of recorded data. To address this gap, five lightweight mechanical equipment units, comprising three slender and two squat aspect ratios, were installed on the rooftop of a full-scale, ten-story cold-formed steel building, and subsequently tested on the UC San Diego six-degree-of-freedom outdoor shake table. The equipment's base attachments were strategically designed to evaluate the impact of different boundary conditions, such as direct (fixed) attachment or the use of vibration isolation interfaces. The variation in base attachment leads to different dynamic properties reflective of realistic installation variability. Fundamental periods and damping ratios of each unit were extracted from white noise test data using three output-only system identification algorithms, with the aim of correlating local damage to changes in global modal features.

Keywords Experimental modal analysis · Modal parameters · Nonstructural components · System identification · Shake table test

Introduction

The six-degree-of-freedom (6-DOF) Large High-Performance Outdoor Shake Table (LHPOST6) at the UC San Diego offers a unique opportunity to extract dynamic behavior-critical modal parameters of civil structures through experimental and operational modal analysis techniques. The first landmark system identification study of an as-built structure tested on the UC San Diego shake table was conducted by Moaveni et al. [1], who evaluated the modal parameters of a seven-story reinforced concrete structural wall building slice using input-output and output-only system identification methods, and correlated the identified natural frequencies with different damage states. Subsequently, Astroza et al. [2, 3] carried out modal identification of a five-story reinforced concrete building furnished with a variety of nonstructural components and systems, while it was both fixed and base-isolated at its attachment to the shake table. The high-damping rubber isolation system was found to elongate the predominant period and increase damping of the building-nonstructural system. The first comprehensive system identification study of a metal structure tested on the UC San Diego shake table was conducted by Wang and Hutchinson [4]. This study embarked to understand the evolution of modal characteristics of a six-story cold-formed steel (CFS) building under a multi-hazard loading scenario, namely involving both earthquake and live-fire tests. Mugabo et al. [5] explored changes in natural frequencies and damping ratios of a two-story mass timber building as shaking intensity increased and further examined the effects of repair and retrofit measures during the experimental program.

Significant progress has been made in system identification of structures tested on 1-DOF shake tables. However, these studies face limitations as the imposition of excitation along a single DOF limits the characterization. Uncertainty analyses

Xianzhao Zhang · Jiachen Zhang · Shokrullah Sorosh · Tara C. Hutchinson
Department of Structural Engineering, University of California San Diego, La Jolla, CA 92093, USA
e-mail: xiz217@ucsd.edu; j-charlie-zhang@ucsd.edu; ssorosh@ucsd.edu; tara@ucsd.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

of system identification results [6] revealed that excitation amplitude had the greatest influence on the variability of the identified modal parameters, particularly the first-mode natural frequency. Consequently, for the structural axis that did not experience sufficient excitation, the identified modal parameters may be less accurate.

Moreover, although various system identification methods have been applied to building structures with different construction materials, heights, hazard scenarios, and damage states, few studies have focused on estimating the modal parameters of nonstructural components. Amongst the most relevant to the present study, the work of Astroza et al. [7], using base excitation (pulse, white noise, and earthquake) and pull back tests investigated a rooftop-mounted cooling tower supported by four isolation mounts. Pull back tests were conducted while the tower was both empty and fully water-filled. The presence of the isolation mounts led to a vertical direction dominant mode with a fundamental period of 0.21 sec.

Description of the Rooftop Equipment

As shown in Figure 1(a), five lightweight air-conditioning condenser units, representing two different models, were installed on roof curbs atop a full-scale ten-story CFS building [8], tested at LHPOST6. Figures 1(b, c) illustrate the geometric features of the slender and squat units, respectively. The slender unit measured 1,000 mm × 368 mm in plan and 1,365 mm in height, with a resulting maximum aspect ratio (AR) of 3.7, while the squat unit measured 902 mm × 902 mm in plan and 972 mm in height, with a resulting AR of 1.1. The slender unit was supported by two manufacturer-prefabricated steel tracks with four feet attached to its bottom face, whereas the squat units had no such installation aids. The weights of the slender and squat units were estimated at 1,434 N and 1,339 N, respectively, assuming a full refrigerant charge.

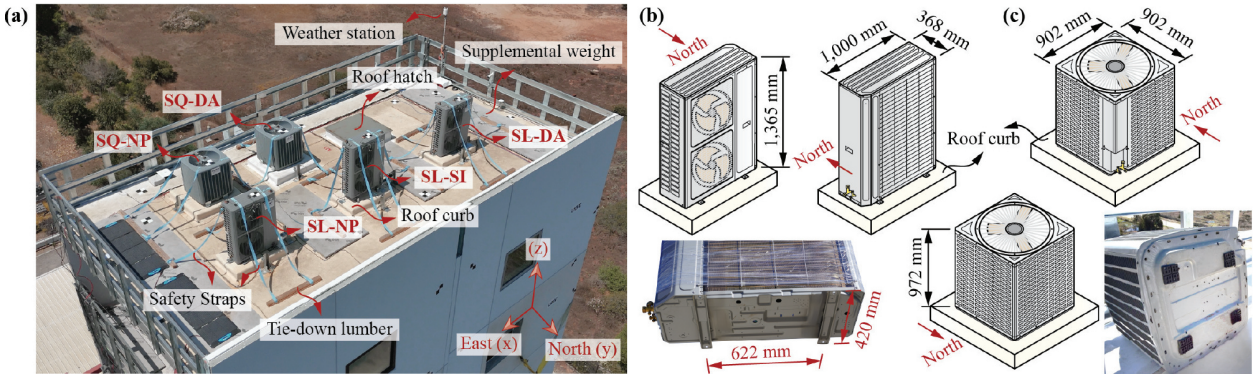


Fig. 1 Overview of rooftop mechanical equipment: (a) aerial view of the equipment layout, view from the northeast corner, (b) slender unit geometry, and (c) squat unit geometry

Two types of commonly used vibration and acoustic mitigation hardware, neoprene pads and spring isolators, were installed at the interface between the slender units and a roof curb. As a benchmark configuration, one slender unit was directly screw-mounted at its four manufacturer-provided feet to CFS studs of the roof curb. To modify its base boundary conditions, two alternative attachment strategies were implemented. In the first case, four neoprene pads, each placed beneath a shim plate, were installed under the equipment feet. In the second case, each equipment foot was supported by a spring isolator. For the squat units, both were installed using four steel angle support clips. One of the squat units was additionally supported by four neoprene pads placed beneath the equipment. Each clip was fastened to the equipment housing with three self-drilling screws, and to the CFS studs of the roof curb with two self-drilling screws. For brevity, SL-DA denotes the slender unit with direct attachment, SL-NP the slender unit with neoprene pads, SL-SI the slender unit with spring isolators, SQ-DA the squat unit with direct attachment, and SQ-NP the squat unit with neoprene pads. Further details about the rooftop mechanical equipment and their attachment designs are provided in [9].

Instrumentation Plan and Dynamic Tests

A total of 36 analog sensors were used to monitor the rooftop mechanical equipment, including 16 accelerometers and 20 linear potentiometers. The accelerometers were micro-electro-mechanical systems-based, with a full-scale range of ±10 g, a frequency bandwidth of 350 Hz, and a dynamic range of 76 dB. These sensors were connected to a distributed multi-node

data acquisition system, operating at a sampling rate of 256 Hz. The layout of accelerometers for the equipment is shown in Figure 2. For the slender unit, two uniaxial accelerometers oriented in the y-direction were mounted at the northwest and southeast corners of the top surface, approximately 76 mm from the edges. In addition, one uniaxial accelerometer in the x-direction and another in the z-direction were installed at the center of the top face. For the squat unit, one uniaxial accelerometer in each of the x- and y-directions was installed at the center of the top surface, while no z-direction accelerometers were deployed. The recorded acceleration time series was baseline-corrected and filtered using a fourth-order bandpass Butterworth filter with cutoff frequencies of 0.2 Hz and 35 Hz, to remove low- and high-frequency signal noise while retaining the frequency range containing the primary vibration modes that contributed significantly to the equipment response.

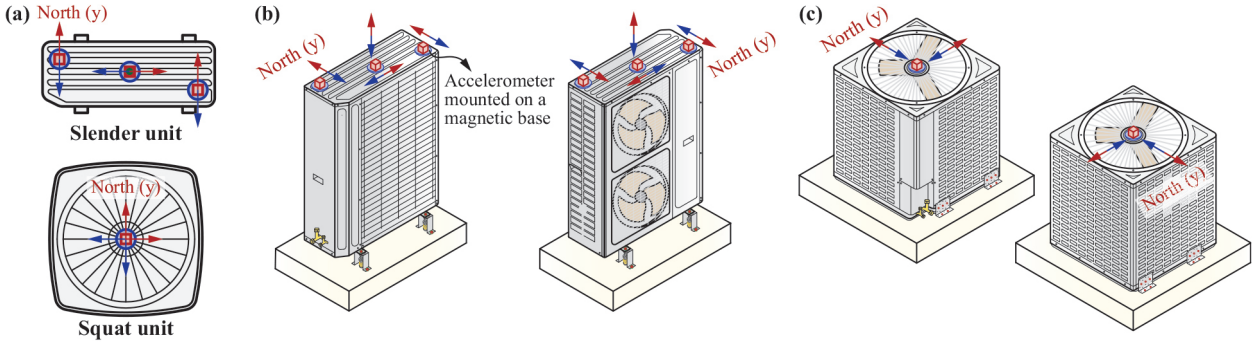


Fig. 2 Accelerometer instrumentation: (a) top view of slender and squat units, and representative layout of accelerometers for (b) SL-SI and (c) SQ-NP

A sequence of low-amplitude vibration tests, including white noise (WN) tests and ambient vibration (AV) tests, were performed before, during, and after multi-directional earthquake tests. The earthquake test suite included 15 service-level earthquakes (SLEs), one design-level earthquake (DE), and two risk-targeted maximum considered earthquakes (MCE_Rs) [10]. Translational (unidirectional) WN tests with a target 3%–4% g root-mean-square (RMS) acceleration, and vertical WN tests with a target 10% RMS acceleration, were imposed sequentially at the base of the building. Each WN test consisted of multiple 16-second windows within a frequency band of 0.125–25 Hz. Except for WN tests 1–3 and 8–10, which had durations of 160 seconds, the remaining WN tests had a duration of 64 seconds. In total, 37 WN tests were conducted, including 17 in the x-direction, 17 in the y-direction, and three in the z-direction, along with 13 AV tests. The timeline of the low-amplitude vibration test protocol is summarized in Figure 3. The focus of this paper is on system

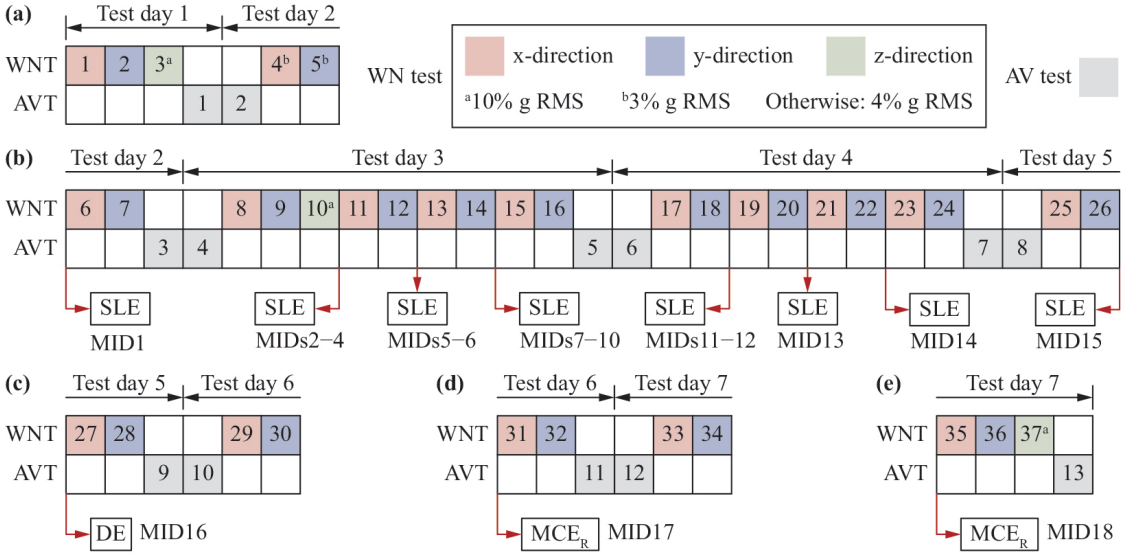


Fig. 3 Timeline of the low-amplitude vibration test protocol: (a) pre-earthquake testing, (b) during SLE phases, (c) post-DE and pre-first MCE_R, (d) post-first and pre-second MCE_R, and (e) post-second MCE_R. MID = motion identification (see [10])

identification using data obtained from WN tests because the equipment experienced sufficient and relatively consistent excitation.

Modal Parameters Identified from WN Tests

Three state-of-the-art output-only system identification methods were applied to the WN test data. These methods include the Enhanced Frequency Domain Decomposition (EFDD) [11], the Natural Excitation Technique combined with the Eigensystem Realization Algorithm (NExT-ERA) [12, 13], and the Covariance-Driven Stochastic Subspace Identification method (SSI-COV) [14]. All three methods assume a linear time-invariant model of the equipment, with energy dissipation represented by linear viscous damping. Accordingly, the identified natural periods and damping ratios correspond to the modal parameters of an equivalent linear viscoelastic system.

Figure 4 shows the identified fundamental periods, T_a from the WN test data for three slender units, SL-DA, SL-NP, and SL-SI, in both the x- and y-directions, along with normalized values expressed relative to the initial T_a. At the beginning of the test sequence, the fundamental periods T_a, in the x-direction were 0.16 sec, 0.18 sec, and 0.77 sec for SL-DA, SL-NP, and SL-SI, respectively, while in the y-direction they were 0.24 sec, 0.25 sec, and 0.73 sec. These values reflect the average across the three system identification methods. An elongation of T_a was observed as the excitation intensity increased from

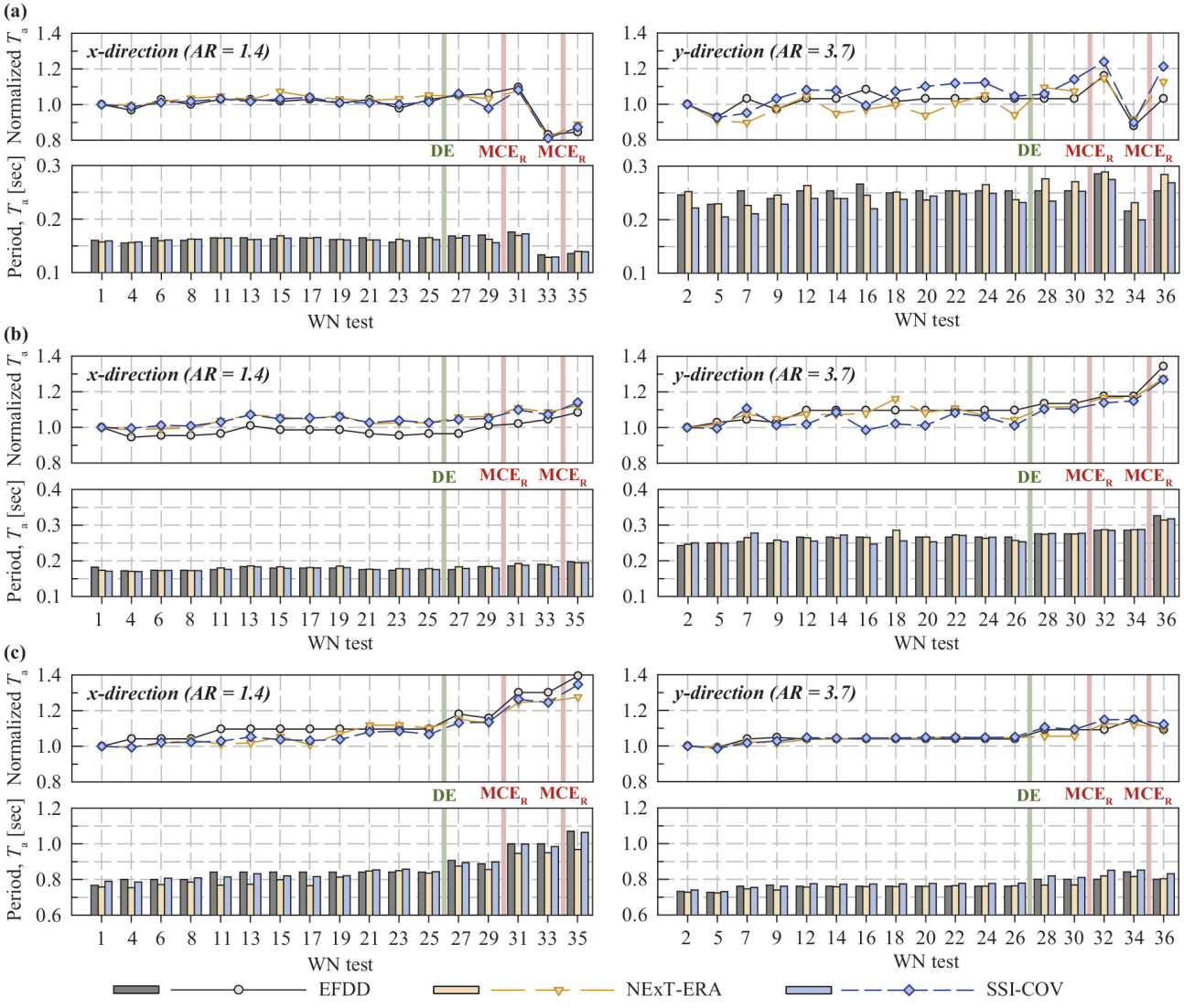


Fig. 4 Variation of natural periods of slender units from WN test data: (a) SL-DA, (b) SL-NP, and (c) SL-SI

the DE to the MCE_R test phase. For SL-DA, T_a increased by 8.6% in the x-direction and 18.2% in the y-direction after the first MCE_R test phase, relative to the initial T_a at the start of testing. The elongation of T_a is attributed to anchorage failure at one attachment point [15]. After retrofitting the failed anchorage, a noticeable reduction in T_a was observed prior to the second MCE_R test phase.

For SL-NP, no visible damage was identified during the post-test inspection. However, T_a increased by 11.6% and 29.7% in the x- and y-directions, respectively, by the end of the testing program. This elongation may be attributed to subtle damage mechanisms not easily detected visually, such as micro-slippage or wear of screw threads, localized bearing damage between screws and CFS studs, or minor crushing of the cementitious panel around the fasteners. For SL-SI, the average increases in T_a were 33.9% in the x-direction and 10.6% in the y-direction after the second MCE_R test phase. This elongation is associated with localized deformation of the spring isolator anchorages and permanent bending of the supporting track [15]. Similarly, the variation of T_a from the WN test data for the two squat units, SQ-DA and SQ-NP is shown in Figure 5. A subtle increase of 7.9% in the y-direction T_a was observed for SQ-NP. However, post-test inspection revealed no visible damage to either squat unit. As shown in Figure 6, T_a associated damping ratios, ζ_a exhibited much larger variations across the different system identification methods and test sequences. For example, for SL-SI during the SLE test phase, which can be considered a damage-free condition, the coefficient of variation of the damping ζ_a was 28.3% in the x-direction and 14.5% in the y-direction. This level of variability suggests that relying solely on identified damping ζ_a may not be sufficient to characterize physical damage.

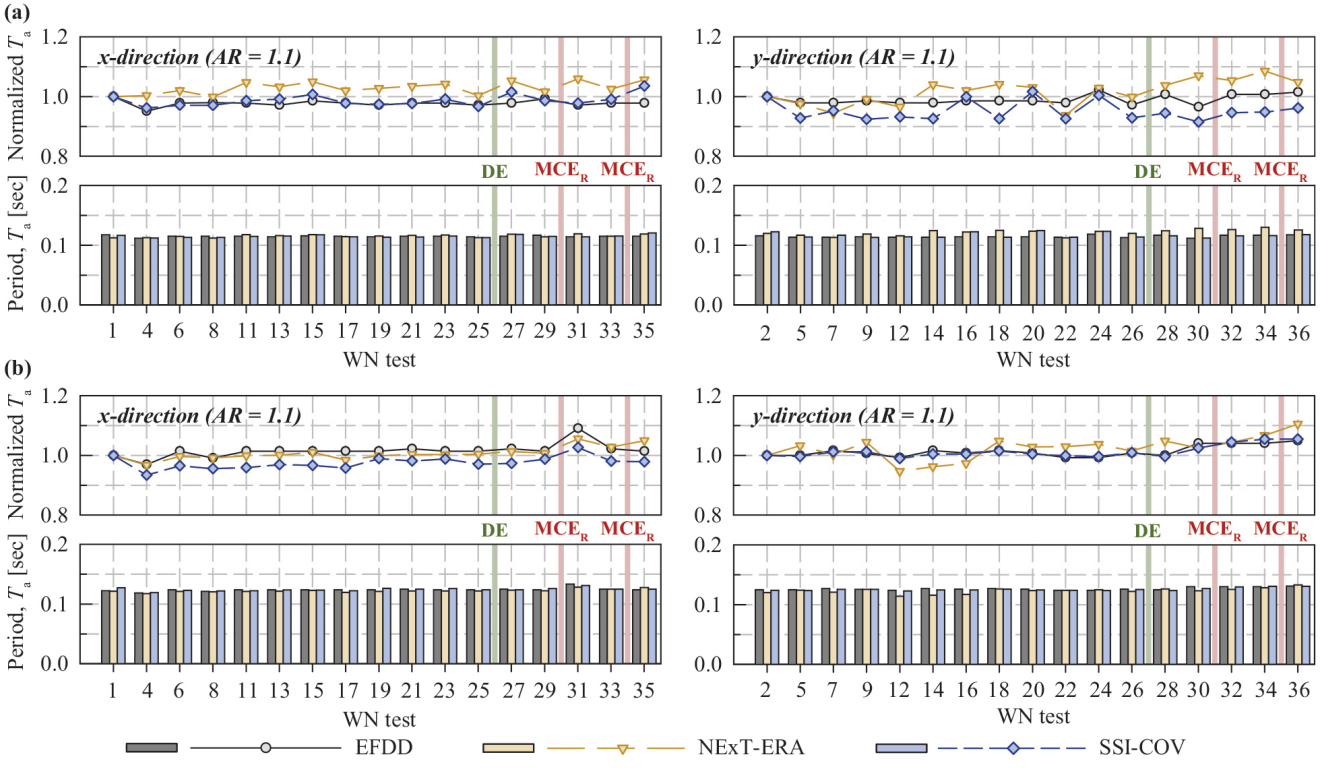


Fig. 5 Variation of natural periods of squat units from WN test data: (a) SQ-DA and (b) SQ-NP

Figures 4–6 aim to show the changes in identified modal parameters of the slender and squat units using different system identification methods at various test phases. When floor response spectra are employed to analyze uncoupled nonstructural components, knowledge of T_a and ζ_a is important and necessary. From the variations of T_a shown in Figures 4–5 it can be inferred that the modal parameters of the equipment in its intact condition can be established from system identification prior to the DE test phase. In other words, the initial undamaged state of the equipment can be characterized using the first 12 x-direction WN tests and the first 12 y-direction WN tests.

Figure 7 summarizes the mean values of T_a and ζ_a obtained from three system identification methods using WN test data prior to the design level test. For the slender unit, the use of neoprene pads for acoustic and vibration mitigation will increase T_a by 12.5% and 8.3% in the x- and y-directions, respectively, while ζ_a increased by 25.7% and 24.2%. When spring isolators were used, which can be considered a more effective vibration mitigation strategy, T_a increased dramatically by 406.3% and 216.7% in the x- and y-directions, respectively. Such a large elongation poses a potential risk of resonance with

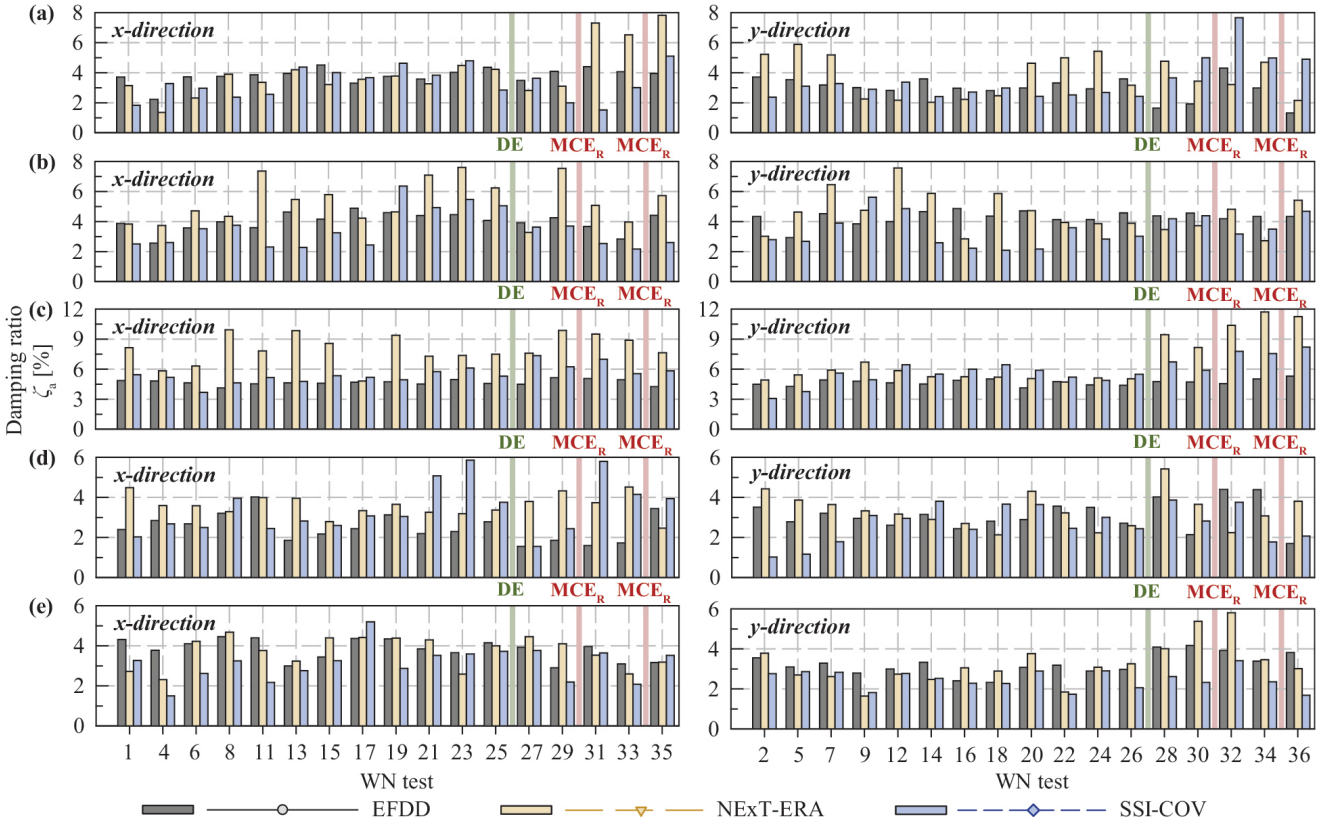


Fig. 6 Variation of damping ratios of slender and squat units from WN tests data: (a) SL-DA, (b) SL-NP, (c) SL-SI, (d) SQ-DA, and (e) SQ-NP

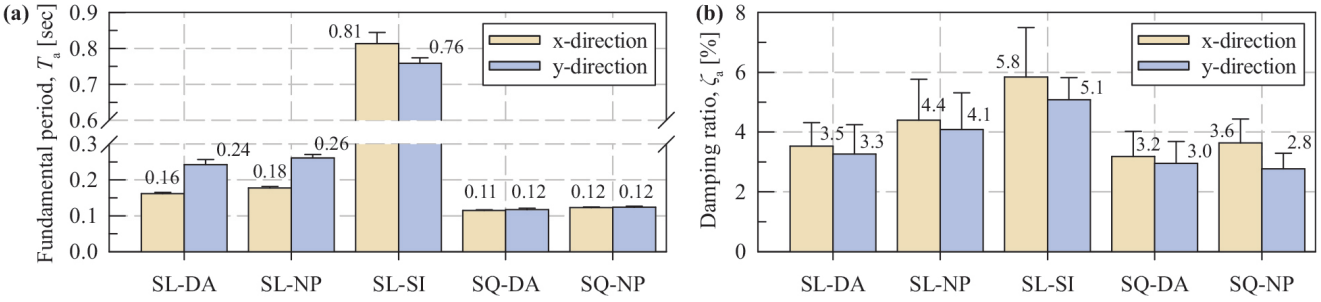


Fig. 7 (a) Natural periods and (b) damping ratios identified using WN test before design level test

the supporting structure, even though it also beneficially increased ζ_a by 65.7% and 54.5%. For the squat unit, T_a in both the x- and y-directions remained small (on the order of about 0.1 sec), and the use of neoprene pads appeared to have negligible influence.

Conclusions

This paper examines the dynamic characteristics of lightweight rooftop mechanical equipment with different attachment strategies, installed on a full-scale ten-story CFS building. The building was tested at the UC San Diego outdoor shake table under sequentially increasing multi-directional earthquake scenarios. Modal parameters were identified from WN tests using three output-only system identification methods. Results highlight the important role that the equipment unit's attachment design plays in dynamic characteristics, influencing both natural periods and damping ratios.

For the slender units, the directly attached configuration (SL-DA) experienced elongation of period, T_a, due to anchorage failure, which was recovered after retrofit. The neoprene-pad-supported unit (SL-NP) showed moderate period elongation without visible physical damage, suggesting the presence of possible subtle degradation mechanisms. The spring-isolator-supported unit (SL-SI) exhibited period elongation linked to anchorage and support rack deformation. Squat units (SQ-DA and SQ-NP) had the shortest periods ($\sim$0.12 sec) and only minor elongation by the end of testing. Damping ratios, ζ_a, displayed significant variability across methods and test phases in the WN tests, though all units attained damping values $> \sim$3%. While vibration and acoustic mitigation hardware increased damping, it also introduced dramatic elongation of periods, raising the potential risk of resonance with the supporting structure.

Acknowledgments This project was executed with support of the National Science Foundation under Grants CMMI-1663569 and CMMI-1663348, together with numerous other federal, state, and industry partners. Specific to the rooftop mechanical equipment discussed herein, additional support was provided by Mason West, Inc., Mason Industries, Inc., and Wells Plumbing and Heating Supplies, Inc. A complete list of sponsors and partners is available at: cfs10.ucsd.edu/sponsors-partners. The authors gratefully acknowledge all undergraduate and graduate students and visiting scholars (cfs10.ucsd.edu/people), as well as the LHPOST6 facility staff (esec.ucsd.edu/about/personnel), for their contributions to the project. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsoring organizations.

References

1. Moaveni, B., He, X., Conte, J.P., Restrepo, J.I., and Panagiotou, M. "System identification study of a 7-story full-scale building slice tested on the UCSD-NEES shake table". *Journal of Structural Engineering*, 137.6: 705−717 (2011) https://doi.org/10.1061/(ASCE)ST.1943-541X.0000300
2. Astroza, R., Ebrahimian, H., Conte, J.P., Restrepo, J.I., and Hutchinson, T.C. "System identification of a full-scale five-story reinforced concrete building tested on the NEES-UCSD shake table". *Structural Control and Health Monitoring*, 23.3: 535−559 (2016) https://doi.org/10.1002/stc.1778
3. Astroza, R., Conte, J.P., Restrepo, J.I., Ebrahimian, H. and Hutchinson, T. "Seismic response analysis and modal identification of a full-scale five-story base-isolated building tested on the NEES@ UCSD shake table". *Engineering Structures*, 238: 112087 (2021) https://doi.org/10.1016/j.engstruct.2021.112087
4. Wang, X. and Hutchinson, T.C. "Evolution of modal characteristics of a mid-rise cold-formed steel building during construction and earthquake testing". *Earthquake Engineering & Structural Dynamics*, 49.14: 1539−1558 (2020) https://doi.org/10.1002/eqe.3316
5. Mugabo, I., Barbosa, A.R., Sinha, A., Higgins, C., Riggio, M., Pei, S., van de Lindt, J.W. and Berman, J.W. "System identification of UCSD-NHERI shake-table test of two-story structure with cross-laminated timber rocking walls". *Journal of Structural Engineering*, 147.4: 04021018 (2021) https://doi.org/10.1061/(ASCE)ST.1943-541X.0002938
6. Moaveni, B., Barbosa, A.R., Conte, J.P. and Hemez, F.M. "Uncertainty analysis of system identification results obtained for a seven-story building slice tested on the UCSD-NEES shake table". *Structural Control and Health Monitoring*, 21.4: 466−483 (2014) https://doi.org/10.1002/stc.1577
7. Astroza, R., Pantoli, E., Selva, F., Restrepo, J.I., Hutchinson, T.C., and Conte, J.P. "Experimental evaluation of the seismic response of a rooftop-mounted cooling tower". *Earthquake Spectra*, 31.3: 1567−1589 (2015) https://doi.org/10.1193/071513EQS205M
8. Hutchinson, T.C., Eladly, M., Ji, R., Rivera, D., Rukundo, F.C., Singh, A., Sorosh, S., Zhang, J., Zhang, X., Padgett, L., Jones, H., Karns, J. and Schafer, B.W. "Advancing use of cold-formed steel framing systems for mid-rise residential construction in seismic zones via full-scale shake table testing". In *Proceedings of the 8th Residential Building Design & Construction Conference*, State College, Pennsylvania (2026)
9. Zhang, X., Zhang J., and Hutchinson, T.C. "Shake table testing of rooftop mechanical equipment with different attachment strategies on a full-scale ten-story building". In *Sixth International Workshop on the Seismic Performance of Non-Structural Elements*, Yokohama, Japan (2026)
10. Zhang J., Sorosh S., Ji R., Lotfizadeh K.H., Schafer B.W., Hutchinson T.C. "Full-scale 10-story cold-formed steel building shake table test: Protocol development and shake table performance". In *Proceedings of the 13th National Conference on Earthquake Engineering*, Earthquake Engineering Research Institute, Portland, Oregon (2026)
11. Brincker, R., Zhang, L., and Andersen, P. "Modal identification of output-only systems using frequency domain decomposition". *Smart Materials and Structures*, 10.3: 441−445 (2001) https://doi.org/10.1088/0964-1726/10/3/303
12. Juang, J.N. and Pappa, R.S. "An eigensystem realization algorithm for modal parameter identification and model reduction". *Journal of Guidance, Control, and Dynamics*, 8.5: 620−627 (1985) https://doi.org/10.2514/3.20031
13. Carne, T.G., and James III, G.H. "The inception of OMA in the development of modal testing technology for wind turbines". *Mechanical Systems and Signal Processing*, 24.5: 1213−1226 (2010) https://doi.org/10.1016/j.ymssp.2010.03.006
14. Magalhães, F., Cunha, A., and Caetano, E. "Online automatic identification of the modal parameters of a long span arch bridge". *Mechanical Systems and Signal Processing*, 23.2: 316−329 (2009) https://doi.org/10.1016/j.ymssp.2008.05.003
15. Zhang, X., Zhang J., and Hutchinson, T.C. "Physical damage to rooftop mechanical equipment considering variations in attachment designs". In *Proceedings of the 13th National Conference in Earthquake Engineering*, Earthquake Engineering Research Institute, Portland, Oregon (2026)

Chapter 11
From Data to Decision: Evaluating the Dangers and Benefits of Reliance on a Flexible Discrepancy Term in Model Calibration

Alexander Knecht and Sez Russcher

Keywords Model Calibration, Bayesian Calibration, Uncertainty Quantification, Model Discrepancy, Verification and Validation

Conventional approaches for calibrating constitutive models of engineering systems have tended to rely on deterministic optimization methods to find a single set of parameters that produce the best fit to data. However, such deterministic approaches are inherently flawed due to their inability to incorporate uncertainties inherent in model development, ranging from input parameters to model form errors to measurement noise [1]. A Bayesian approach for model calibration has gained popularity as it offers a framework that treats model parameters as random variables with appropriately defined prior probability distributions. Rather than chasing single value parameter estimates, the goal becomes to construct a complete posterior probability distribution for each of the parameters. This distribution represents an updated state of knowledge, formally synthesizing prior beliefs about the parameters with the evidence contained in the experimental data by using the log-likelihood function [2]. The resulting posterior distribution offers a more comprehensive characterization of parameters and their associated uncertainties, a significantly more useful result than deterministic methods can provide. [3]

A crucial consideration in modern model calibration, the explicit handling of model discrepancy, has been popularized by Kennedy and O'Hagan (2001), originally referred to as model inadequacy [4]. Their framework suggests that forcing a potentially flawed model to perfectly calibrate to a limited experimental data set may lead to highly biased parameter estimates, stripping them of any predictive significance, a phenomenon known as parameter confounding [5],[8]. To mitigate this, they proposed augmenting the statistical model with the addition of the discrepancy term, comprehensively representing all systemic inadequacies introduced through experimental error, computational model limitations, or prior biases. In many cases, this framework is implemented via Gaussian Processes (GP) models, allowing the surrogate to learn model error, and calibrate much like a parameter [5]. The introduction of a discrepancy term, however, may present additional challenges, primarily risking overfitting and non-identifiability [6],[8].

We have evaluated this on a simple polynomial and trigonometric function. A function is produced for the purpose of this evaluation, with 10 experimental points chosen along the curve, $f(x) = 2x - 2(x)^2 + .8sin(2 * pi * x) + .8cos(2 * pi * x)$. The simulation is then produced, with the removal of the *cosine* term to simulate model form error. Table 1 presents preliminary results used in the simulation calibration, with a baseline test (no model error, no discrepancy) depicted on the left, and increasing complexity moving from left to right (with model error but no discrepancy in the center, with model error and with a discrepancy term to the right). Evident from this table is the dramatically better curve fit produced by the model with discrepancy, despite similarity in the parameter predictions and higher standard deviations (or lower precision) among the data calibrated with discrepancy.

In Table 1, the risk of overfitting is evident; when calibrating to a limited set of experimental data, the introduction of a highly flexible discrepancy term may naturally dominate the effects of true parameters, overly forcing the curve to fit the presented data. While this produces a very accurate representation of the currently available observed points, it handicaps future predictive power by over-constraining the model to a limited set of data. The second issue is that of non-identifiability; while a flexible discrepancy term allows the model to absorb the influence of the physical parameters, it can become difficult to distinguish between the effects of the discrepancy term and the physical parameters, particularly when data is sparse [7]. This issue, as explored by Arendt, Apley, and Chen (2012), describes the effect where a flexible discrepancy term may replicate or even counteract the impacts of the physical model parameters, thereby inhibiting parameter prediction, as seen

Alexander Knecht · Sez Russcher(✉)
Clemson University
e-mail: knecht@clemson.edu; sez@clemson.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

Table 1 Post-Calibration Results of 3 distinct calibration scenarios, the true curve matched to the simulation curve (left), a curve with error but no discrepancy (center), a curve with error and with a discrepancy term (right)

	Perfect Match, no Discrepancy		With Model Error, without Discrepancy		With Model Error, with Discrepancy	
True Parameter Value:	Mean Prediction:	Standard Deviation:	Mean Prediction:	Standard Deviation:	Mean Prediction:	Standard Deviation:
2	1.9998	0.0399	1.996	0.2008	1.990	0.9728
-2	-2.0003	0.0200	-1.998	0.1968	-1.828	0.6624
.8	.80032	0.0183	0.798	0.0794	0.799	0.2678
RMSE:	0.0002		1.2866		0.2240	
|E|:	0.0002		1.8308		0.3275	

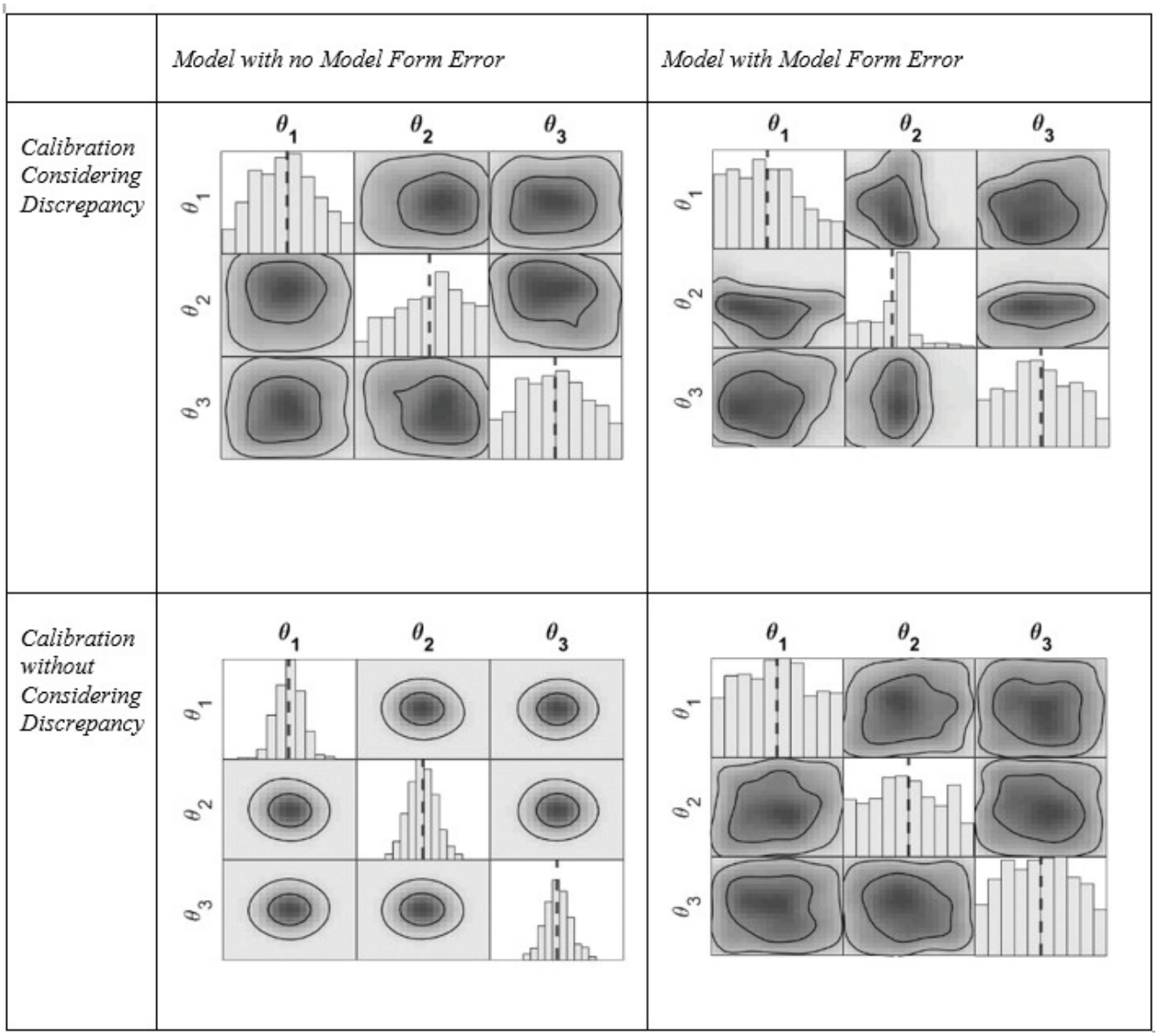


Fig. 1 Posterior prediction histograms produced in the calibration process, with and without model form error, with and without discrepancy terms

in Table 1. This can result in a posterior parameter distribution that is excessively wide or skewed, rendering it physically meaningless in calibration processes [7]. For this reason, we will present model calibration results executed both with and without the discrepancy consideration, to observe how the inclusion of a flexible, GP based discrepancy term impacts model calibration.

Using the same simulation function explored in Table 1, Figure 1 outlines the results of various calibration processes both with and without discrepancy terms and with and without model form error, to observe the effects of flexible discrepancy consideration on the distribution of posterior predictions.

Quadrants 1 and 3 represent the logical selections based on inference alone. For example, in a model with known limitations, the modeler will tend to prefer a method of accounting for those in the calibration process, therefore selecting a calibration with discrepancy when known model form error exists. Alternatively, if the model is known to be an exceptional representation of reality, the best choice will likely be a calibration without the flexible discrepancy term, choosing to accept a highly biased model for the sake of a tight posterior distribution.

References

1. Beck, J. L., & Au, S. K. (2002). Bayesian updating of structural models and their uncertainties. I: Theory. Journal of engineering mechanics, 128(4), 380-391.
2. ASME. (2006). Standard for Verification and Validation in Computational Solid Mechanics (ASME V&V 10-2006). American Society of Mechanical Engineers.
3. Sankararaman, S., & Mahadevan, S. (2015). Bayesian methodology for uncertainty quantification in material property estimation. International Journal of Materials Research, 106(1), 3-13.
4. Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(3), 425–464.
5. Brynjarsdóttir, J., & O'Hagan, A. (2014). Learning about physical parameters: The importance of model discrepancy. Inverse Problems, 30(11), 114007.
6. McFarland, J., & Mahadevan, S. (2008). Error and variability in robust design. AIAA Journal, 46(10), 2586-2597.
7. Arendt, P. D., Apley, D. W., & Chen, W. (2012). Quantification of model uncertainty: Calibration, model discrepancy, and identifiability. Journal of Mechanical Design, 134(10).
8. Maupin, K. A., & Swiler, L. P. (2020). Model discrepancy calibration across experimental settings. Reliability Engineering & System Safety, 200, Article 106818

Chapter 12
System Identification of a 10-story Cold-Formed Steel Building: Systematic Study Considering Sensor Distribution and Changes in Ambient Conditions

Shokrullah Sorosh, Jiachen Zhang, M. Khalid Saifullah, Derek Skolnik, and Tara C. Hutchinson

Abstract A 10-story cold-formed steel (CFS) building was constructed and tested on the six-degree-of-freedom Large High-Performance Outdoor Shake Table at UC San Diego (NHERI@UC San Diego). The primary objective of this testing program was to evaluate the seismic performance of a building with steel-sheathed shear walls beyond the current code height limit. The test specimen was also detailed with various nonstructural components and systems, including 9 stories of prefabricated steel stairs with drift-compatible connections, a fire sprinkler system with horizontal and vertical runs, rigid and flexible gas pipes, and rooftop equipment. These subsystems increased the fidelity of the specimen and its overall potential to capture true-to-field functional buildings. Prior to the seismic testing phase, the building's response under ambient conditions was recorded using various sensor systems and networks, including six rapid-deployment triaxial accelerographs. The unique combination of portability (wireless), high-fidelity (24-bit resolution), and precise time-synchronization (GPS timestamp) of these sensors allowed for convenient and speedy data collection prior to the seismic test phase. Using this pre-test data, this paper discusses the application of several output-only system identification techniques, including conventional Power Spectral Density (PSD) analysis, Frequency Domain Decomposition (FDD), and Data-driven Stochastic Subspace Identification (SSID), to determine the building's modal characteristics under ambient conditions. Observations regarding the SI results are presented, and the variation in the building's modal characteristics under changing ambient conditions is evaluated.

Keywords System identification · Dynamics · Ambient vibration · Shake table test · Operational modal analysis

Introduction

Recently, a 10-story cold-formed steel building, referred to as CFS10, was tested on top of the six-degree-of-freedom Large High-Performance Outdoor Shake Table (LHPOST6) at UC San Diego, see Figure 1a. This landmark building specimen measured 30.5 m in height and 6.86 m × 11.58 m in floor plan. The lateral force resisting system of this test specimen was designed with steel-sheet-sheathed shear walls. The primary intention of this test effort was to test a cold-formed steel-framed building with steel-sheet-sheathed shear wall beyond the current code (ASCE 7-22 [1]) height limit of 19.8 m. The building was subjected to a total of 18 earthquake motions, scaled to various risk-targeted hazard levels [2]. Hutchinson et al. [3] provide an overview of the building performance under the seismic test phase, while Singh et al. [4] summarize the structural design details of this building. In addition, this building featured various nonstructural components and systems, including a 9-story stair tower [5, 6], fire riser and sprinkler systems, windows, gas pipes [7], and roof-mounted mechanical equipment with various attachment details [8].

Shokrullah Sorosh · Jiachen Zhang · Tara C. Hutchinson
Department of Structural Engineering
University of California San Diego, La Jolla, CA
e-mail: ssorosh@ucsd.edu; j-charlie-zhang@ucsd.edu; tara@ucsd.edu

M. Khalid Saifullah · Derek Skolnik
Kinemetrics Inc., Pasadena, CA
e-mail: mks@kmi.com; das@kmi.com

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

Numerous studies have investigated the use of system identification techniques to determine the modal characteristics of test structures under white noise base excitation or ambient conditions [9–12]. The primary objective of the present study was to identify the fundamental modal characteristics of the test specimen using ambient vibration data collected prior to the seismic testing phase. This step is essential for validating numerical simulation tools before conducting seismic tests. It is worth noting that no discussion of the finite element model validation is included in this paper. Zhang et al. [2] provide a detailed discussion on the evolution of the test protocol throughout the seismic testing phase of the CFS10 program. The finite element model of the building plays a crucial role in the ground motion iteration and scaling process. Therefore, identifying the modal characteristics of the building using ambient vibration (AV) data prior to the seismic test phase is critical for ensuring confidence in the finite element model results. The use of portable wireless-based sensors is feasible and convenient to conduct such a study, as it renders an efficient strategy for quickly providing response data. Alternatively, wired sensors would be installed within the building specimen, a highly time-consuming effort. In this study, six portable tri-axial accelerograph units were deployed to collect sufficient data to characterize the first three modal responses of the 10-story building. The vibration data was recorded for a duration of 23 hours and 35 minutes with 200 samples per second (SPS). With 200 SPS, the Nyquist frequency is 100 SPS, which is far above the range of interest for building modal characteristics. The deployment strategy was varied, considering various horizontal and vertical placement options, though the present article focuses on a sample of these results with one strategic layout. Results herein focus on documenting the change in building modal characteristics under varying ambient conditions.

Instrumentation

The acceleration responses of CFS10 under ambient conditions were recorded using six tri-axial accelerographs. These accelerographs were of model ETNA 2 with 3 sensor channels with a high fidelity internal EpiSensor triaxial deck, 24-bit resolution, and 155 dB+ dynamic range. As shown in Figure 1b, the sensors in the present discussion were placed at the northeast and southwest corners of the building at Levels L3, L7, and L11/roof. These locations along the height of the building were particularly chosen to avoid the modal nodes and to capture flexural dominant and torsional modes. In addition, the ambient conditions, such as temperature, humidity, and wind speed, were recorded using a weather station installed on the control building at the south side of the CFS10. Furthermore, the weather data were retrieved from Elliot station of the National Oceanic and Atmospheric Administration (NOAA) [13], which is located approximately 1.6 km southeast of the CFS10.

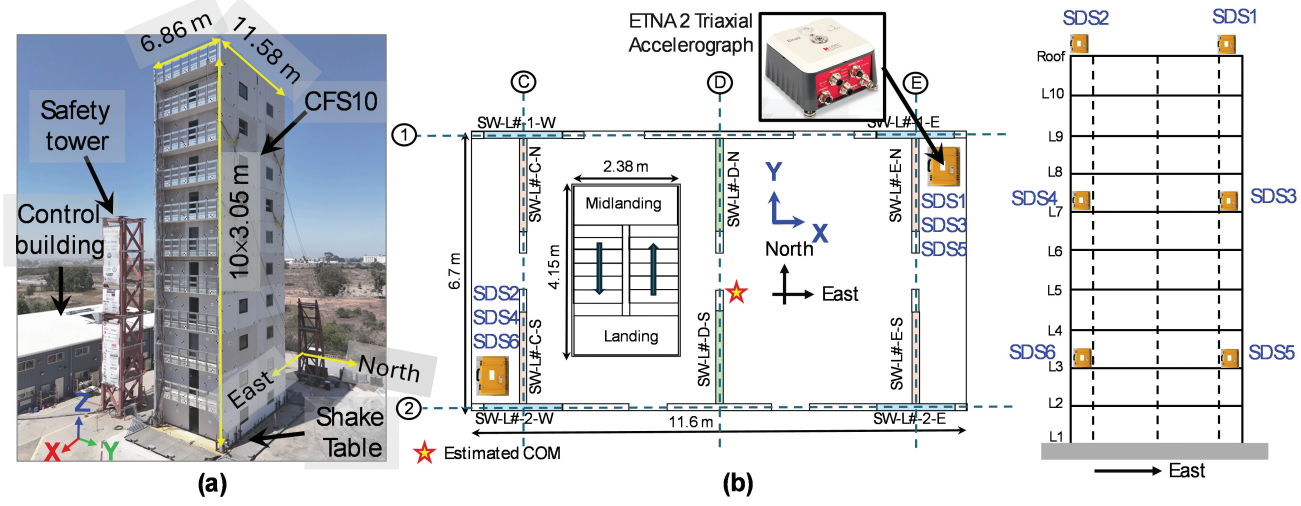


Fig. 1 (a) Isometric view of CFS10, and (b) instrumentation plan (SDS: Serial Data Stream)

Ambient Vibration Data

Data collection with the sensor deployment plan shown in Figure 1b began at 15:00 local time on April 28 and ended at 14:35 local time on April 29. Figure 2a shows the ambient conditions recorded during this period, based on measurements

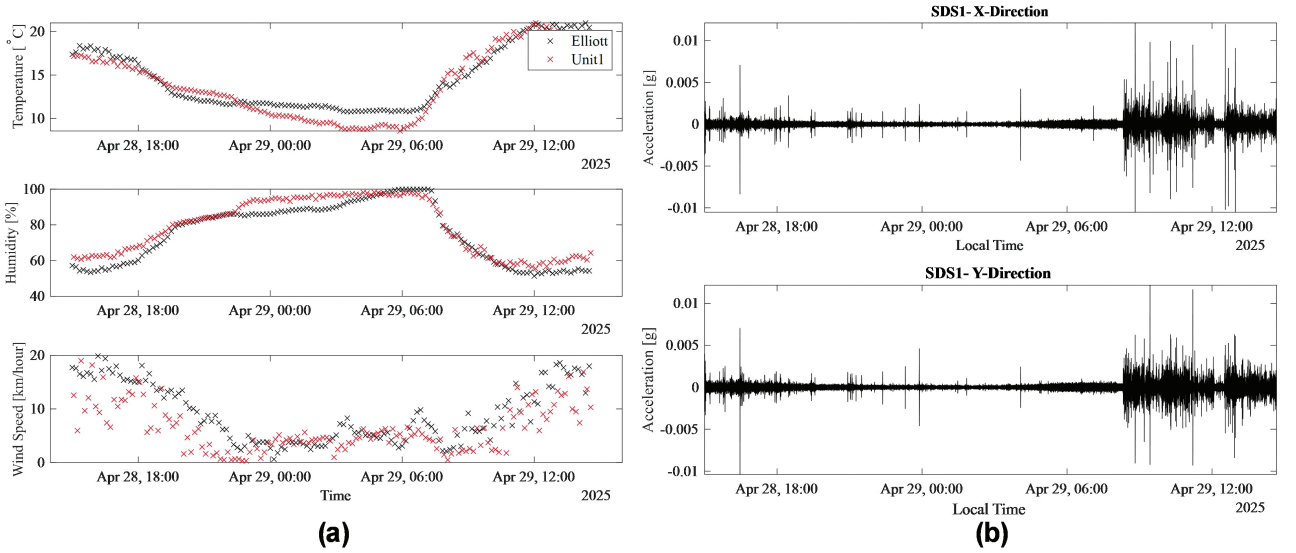


Fig. 2 (a) Ambient conditions and (b) acceleration time series during data collection with sensor deployment plan shown in Figure 1b. (Note: local time is based on Pacific Daylight Time (PDT))

from both the Elliot station (ground surface sensor) and the sensor attached to the control building (approximately 9m from ground level). The weather data from the two sources show good agreement, with some minor variations expected due to spatial differences between the stations. Figure 2b shows the acceleration time series in the X and Y directions recorded by sensor SDS1, located at the northeast corner of the building on the roof. In general, larger acceleration amplitudes are seen during daytime hours, primarily due to human activity in the building during ongoing cable and sensor installation. In addition, the wind speed during the daytime was approximately two to three times larger than at nighttime. The contribution of wind-induced loading on the lateral mode excitation of buildings with enclosed exteriors is significant.

Identified Modal Parameters

In this study, the vibration frequencies of the CFS10 building were identified using three output-only system identification techniques. The first method was based on the power spectral density (PSD) analysis, where the vibration frequencies were determined from the peaks of the PSD of the acceleration time series computed using Welch's method [14]. In addition, the Frequency Domain Decomposition (FDD) [15] and Stochastic Subspace Identification (SSID) [16] techniques were used to estimate the building's modal characteristics, including vibration frequencies and damping ratios. It is noted that, prior to running the system identification analyses, the acceleration time series were baseline corrected and filtered using a 4th-order Butterworth bandpass filter with cutoff frequencies of 0.075 Hz and 50 Hz. Figure 3a shows the acceleration time histories recorded at the northeast corner of the roof of the building by the SDS1 sensor during the first five minutes of data collection (i.e., April 28, 15:00-15:05 PDT). The corresponding PSD of this window is shown in Figure 3b, where each peak in the PSD identifies a vibration mode. The vibration frequencies identified from the SSID and FDD methods are shown by vertical dashed and solid lines, respectively. The frequencies estimated from the three methods show a good agreement.

The FDD and SSID techniques were used to identify the modal parameters throughout the duration of the deployment using continuous non-overlapping 5-minute data segments. Figure 4 shows the identified vibration periods and corresponding damping ratios for the first three vibration modes of the building during this deployment period. The first vibration mode is a flexural-dominant mode in the Y direction with a period ranging from 0.55 to 0.57 seconds. The second mode is a flexural-dominant mode in the X direction with a period between 0.53 and 0.55 seconds. The torsional mode is the third vibration mode with a period ranging from 0.37 to 0.38 seconds. As shown in Figure 4b, the damping ratios corresponding to these modes were generally estimated between 0.5% and 2%. A comparison between the identified vibration periods shown in Figure 4a and the temperature variation shown in Figure 2a demonstrates a gradual decrease in ambient temperature from approximately 18°C during the day to 9°C at midnight, accompanied by a slight reduction in the building's flexural mode period from 0.545 s during the day to 0.535 s at midnight in the X direction, and from 0.565 s during the day to 0.555 s at midnight in the Y direction. Although this 0.01-second change in the building period is relatively small, the observed

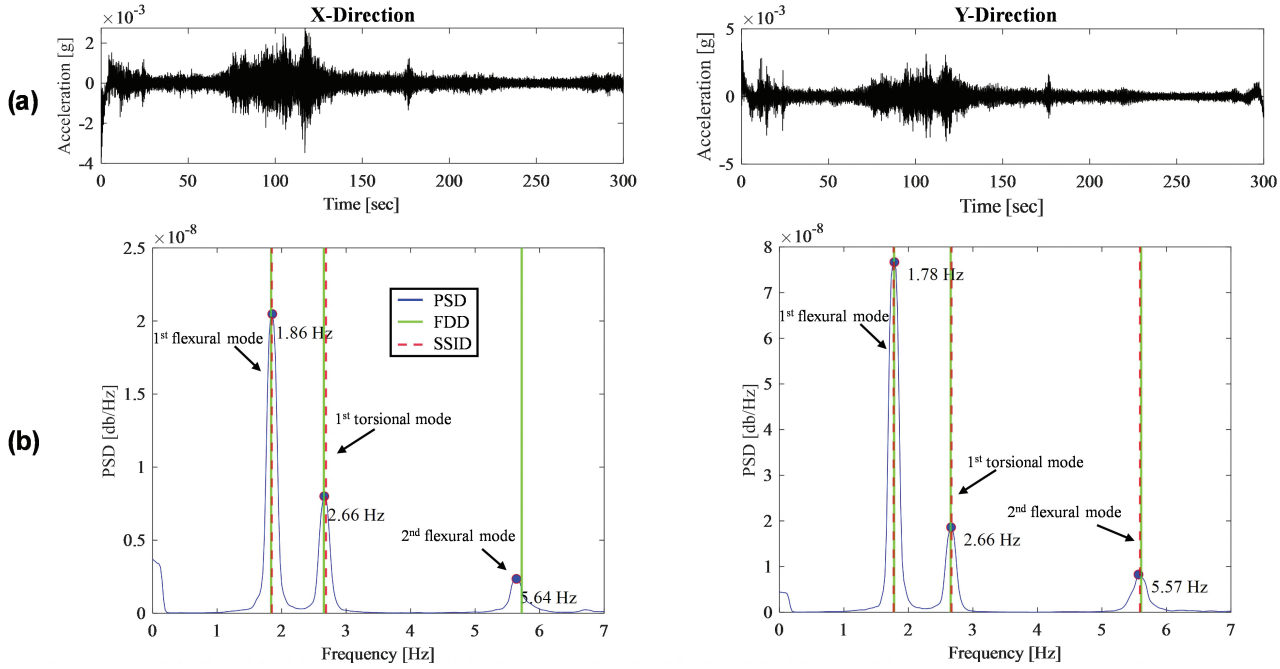


Fig. 3 (a) Acceleration time series from SDS1 for the first time window (April 28, 15:00-15:05 PDT), and (b) corresponding power spectral density with identified modal frequencies using each of the three studied methods highlighted

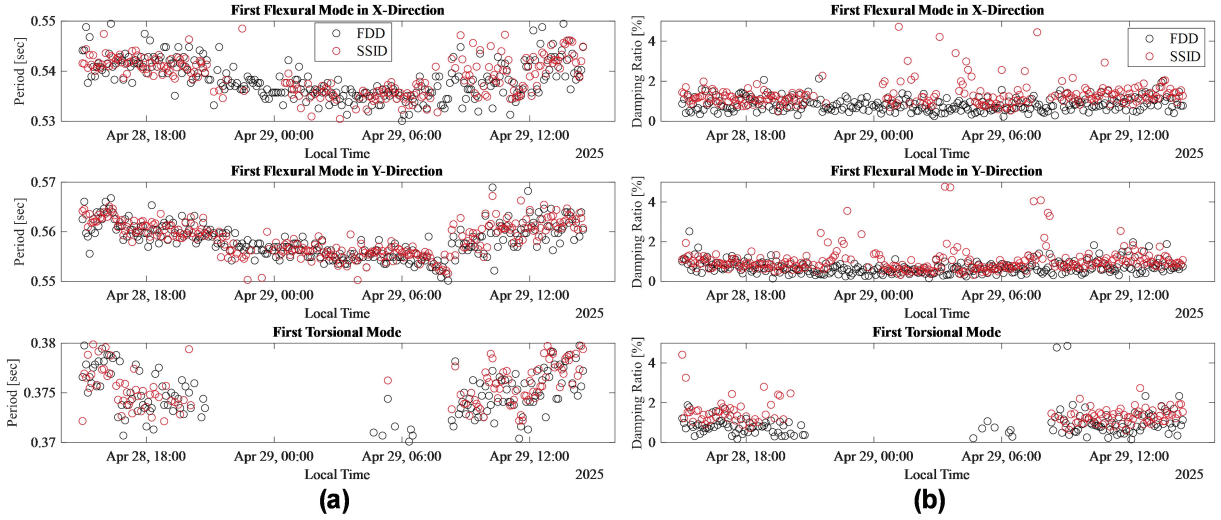


Fig. 4 (a) Identified vibration period (b) damping ratios of CFS10 corresponding to the first three fundamental modes

trend shows that the building becomes slightly stiffer under colder conditions. In addition, this reduction in building period is partially due to lower wind speed during the night as compared to wind speed during the daytime. An increase in humidity would also increase the mass of some buildings, such as wood-framed buildings, but cold-form-steel-framed buildings are expected to be less affected by the humidity level. Furthermore, observing the torsional mode shows that both FDD and SSID identification techniques failed to identify the modal parameters during the calmer hours of the day, which were approximately between 8:00 PM PDT April 28 and 8:00 AM PDT April 29. During this period, the wind speed was relatively low, ranging from about 2 to 4 kilometers per hour.

Figure 5 shows a comparison of identified modal parameters based on the FDD and SSID methods, along with a histogram of the distribution of identified modal parameters. It is noted that the histogram is developed based on a total of 283

samples. Observing the period's scatter plots, the data points are observed closely around the diagonal reference line. Both scatter plots and histograms show a strong agreement between FDD and SSID in estimating the vibration periods. However, the damping ratios show relatively larger scatter, particularly for the X and Y directions, suggesting larger discrepancies in damping ratios obtained using the two methods. Nonetheless, these results show the robustness of both methods in identifying the vibration period while also emphasizing the discrepancies in estimated damping ratios, which are typically very sensitive to noise and system identification modeling assumptions, thus an anticipated observation given the use of AV-data.

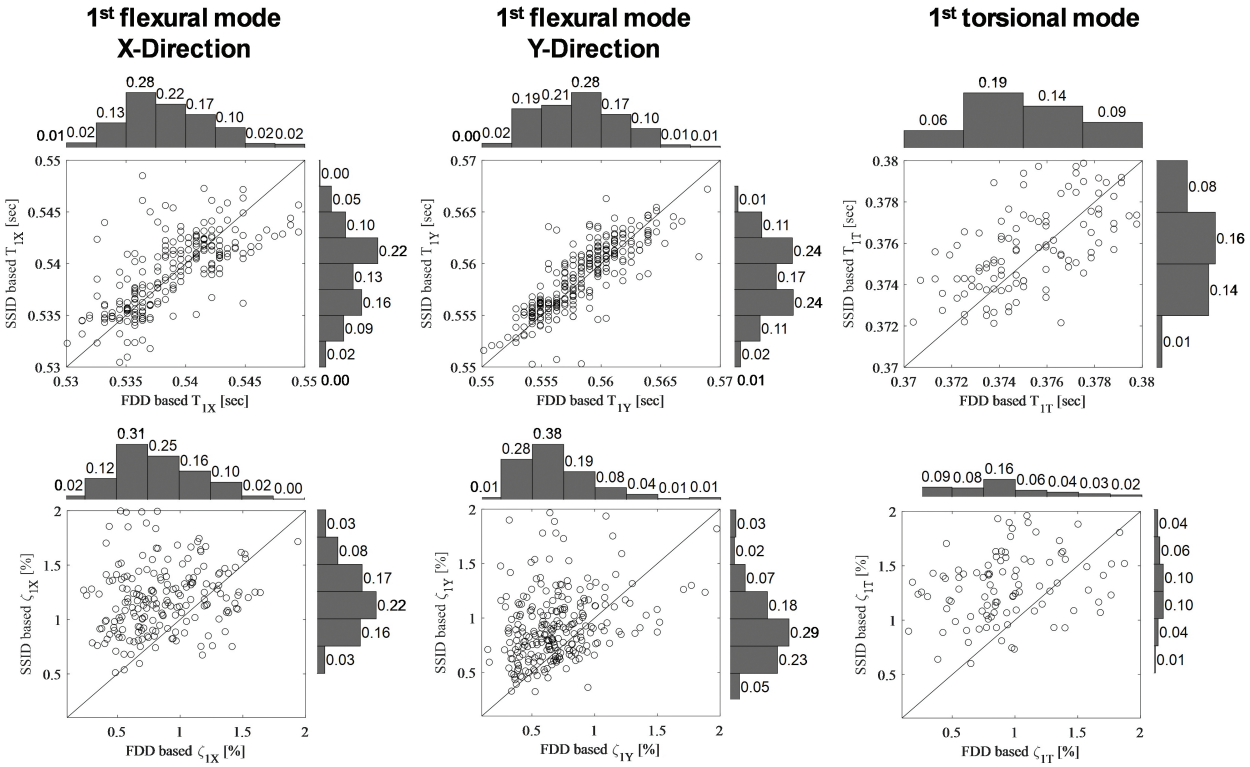


Fig. 5 Comparison of identified period and damping ratios estimated using SSID and FDD (Note: (1) for each case, a total of 283 windows of 5-minute length AV data is considered (total test duration was 23 hours and 35 minutes), and (2) histogram frequency bins are at 0.0025sec and 0.25% for period and damping, respectively

Conclusions

This paper discussed the use of portable wireless-based accelerograph units in vibration data collection of a full-scale shake table test building. The vibration data of the building under ambient conditions were recorded for 23 hours and 35 minutes and deployed in a variety of configurations. Utilizing one of the deployment configurations, these data were used to identify the modal parameters, including the vibration periods and the damping ratios of the building, using three output-only system identification techniques, namely, Power Spectral Density (PSD) analysis, Frequency Domain Decomposition (FDD), and Stochastic Subspace Identification (SSID). The analysis results showed that all three methods are robust in identifying the first three vibration periods of the building. However, comparing the identified damping ratios based on FDD and SSID techniques reveals that there are slight discrepancies between the estimated damping ratios. This is likely because damping ratios are very sensitive to noise and system identification model assumptions. In addition, monitoring the modal parameters for the 23 hours and 35 minutes data collection period showed that the building period decreases during the nighttime, potentially due to a minor increase in stiffness during colder hours, though this reduction in stiffness is not significant. Furthermore, it was concluded that the torsional mode of the building was only predictable when the wind speed was relatively larger, while during the calmer hours of the night, the torsional mode of the building was not identified. It is worth noting that conducting such system identification analysis is essential prior to the shake table test phase of such a large

specimen, as such results provide information regarding the state of the building considering its as-built configuration, and eventually, these results ensure confidence in the results of the developed finite element models of the building.

Acknowledgments The CFS10 project was funded through the National Science Foundation (NSF) grants CMMI 1663569 and CMMI 1663348, project entitled: Collaborative Research: Seismic Resiliency of Repetitively Framed Mid-Rise Cold-Formed Steel Buildings, and numerous other federal/state and industry partners. A complete list of sponsors/partners to the CFS10 program may be found at: https://cfs10.ucsd.edu/sponsors-partners. The authors appreciate the aforementioned support. The efforts of all undergraduate and graduate students and visiting scholars listed at https://cfs10.ucsd.edu/people/ are greatly appreciated. The opinions and findings of this study are those of the authors and do not necessarily reflect those of the sponsors.

References

1. ASCE. (2022). *Minimum design loads and associated criteria for buildings and other structures (ASCE/SEI 7-22)*. Structural Engineering Institute of the American Society of Civil Engineers. Reston, VA.
2. Zhang, J., Sorosh, S., Ji, R., Schafer, B. W., and Hutchinson, T. C. (2026). "Full-scale 10-story cold-formed steel building shake table test: Protocol development and shake table performance." In *Proceedings of the 13th National Conference on Earthquake Engineering*, Earthquake Engineering Research Institute, Portland, OR.
3. Hutchinson, T. C., Eladly, M., Ji, R., Rivera, D., Rukundo, F. C., Singh, A., Sorosh, S., Zhang, J., Zhang, X., Padgett, L., Jones, H., Karns, J., and Schafer, B. W. (2026). "Advancing use of cold-formed steel framing systems for mid-rise residential construction in seismic zones via full-scale shake table testing." In *Proceedings of the 8th Residential Building Design & Construction Conference*, State College, PA.
4. Singh, A., Zhang, J., Rivera, D., Eladly, M., Jones, H., Kovac, A., Padgett, L., Smith, K., Torabian, S., Peterman, K. D., Schafer, B. W., and Hutchinson, T. C. (2024). "CFS-NHERI 10-story building shake table test specimen: Design updates." In *Proceedings of the 18th World Conference on Earthquake Engineering (WCEE2024)*, Milan, Italy.
5. Sorosh, S., Zhang, J., Kovac, A., Smith, K., and Hutchinson, T. C. (2025). "Design strategies and modeling verification for steel stairs with drift-compatible connections." In *Proceedings of the 2025 Structural Engineers Association of California (SEAOC) Convention*, San Diego, CA, USA.
6. Sorosh, S., Zhang, J., Kovac, A., Smith, K., Schafer, B. W., and Hutchinson, T. C. (2026). "Influence of stair system flexibility on building vibration period: Insight from full-scale building shake table tests at low-amplitude excitation." In *Proceedings of the ATC/SPONSE Sixth International Workshop on the Seismic Performance of Non-Structural Elements (SPONSE)*, Yokohama, Japan.
7. Singh, A., Rukundo, F. C., Ameen, J. R., Fathali, S., Jaramillo, J., Torbin, B., Huntress, J., Szakats, G., Cook, D. T., Speicher, M. S., Sattar, S., Hutchinson, T. C., and Schafer, B. W. (2026). "Interaction of Suspended Ceilings and Pressurized Piping in a Full-Scale 10-Story Cold-Formed Steel Building Shake Table Test Program." In *Proceedings of the 6th International Workshop on the Seismic Performance of Non-Structural Elements (SPONSE)*, Yokohama, Japan.
8. Zhang, X., Zhang, J., and Hutchinson, T. C. (2026). "Shake table testing of rooftop mechanical equipment with different attachment strategies on a full-scale ten-story building." In *Proceedings of the 6th International Workshop on the Seismic Performance of Non-Structural Elements (SPONSE)*, Yokohama, Japan.
9. Moaveni, B., He, X., Conte, J. P., Restrepo, J. I., and Panagiotou, M. (2011). System identification study of a 7-story full-scale building slice tested on the UCSD-NEES shake table. *Journal of Structural Engineering, 137*(6), 705–717. https://doi.org/10.1061/(ASCE)ST.1943-541X.0000300
10. Astroza, R., Ebrahimian, H., Conte, J. P., Restrepo, J. I., and Hutchinson, T. C. (2016). System identification of a full-scale five-story reinforced concrete building tested on the NEES-UCSD shake table. *Structural Control and Health Monitoring, 23*(3), 535–559. https://doi.org/10.1002/stc.1778
11. Astroza, R., Conte, J. P., Restrepo, J. I., Ebrahimian, H., and Hutchinson, T. C. (2021). Seismic response analysis and modal identification of a full-scale five-story base-isolated building tested on the NEES@UCSD shake table. *Engineering Structures, 238*, 112087. https://doi.org/10.1016/j.engstruct.2021.112087
12. Wang, X., and Hutchinson, T. C. (2020). Evolution of modal characteristics of a mid-rise cold-formed steel building during construction and earthquake testing. *Earthquake Engineering & Structural Dynamics, 49*(14), 1539–1558. https://doi.org/10.1002/eqe.3316
13. National Oceanic and Atmospheric Administration. (2025). *Weather*. Retrieved from https://www.noaa.gov/weather
14. MathWorks. (2025). *pwelch, Welch's power spectral density estimate*. In *MATLAB Signal Processing Toolbox Documentation*. Retrieved from https://www.mathworks.com/help/signal/ref/pwelch.html
15. Brincker, R., Zhang, L., and Andersen, P. (2001). Modal identification of output-only systems using frequency domain decomposition. *Smart Materials and Structures*, *10* 441 445 https://doi.org/10.1088/0964-1726/10/3/303
16. Overschee, P.V., and De Moor, B.L. (1996). *Subspace Identification for Linear Systems: Theory–Implementation–Applications*, Kluwer Academic Publishers, Norwell, MA.

Chapter 13
Automated Correlation of Simulations and Experiments for Digital Twin Development

G. Brunbauer, A. M. Puhwein, and M. J. Hochrainer

Abstract Accurate experimental validation of finite element models with 3D scanning laser Doppler vibrometry measurements is crucial for reliable structural-dynamics simulations, especially for large or complex systems. When entire test rigs are measured, experimental modal analysis of individual components is infeasible and the finite element model must inevitably include non-trivial boundary conditions. Although commercial software tools offer automated correlation algorithms, these often fail when applied to large models with point masses, rigid bodies, contacts or mismatched geometries. Therefore a transparent workflow is presented to validate complex numerical models using 3D vibrometry data. It covers import and processing of simulated and measured modal data, alignment of measurement grids and FE mesh, robust eigenvector interpolation and correlation via modal assurance criterion. Unlike black-box tools, the proposed open implementation allows full control of each step within the validation process. The workflow proposed has proven to work reliably on a large FE model of a twin-disc tribometer and builds a suitable basis for realistic digital twin modelling.

Keywords Digital Twin · Correlation · FEM · Modal Analysis · SLDV

Introduction

Modern high fidelity finite element (FE) models are of high practical value for the design and optimisation of mechanical structures. However this is restricted to models which behave similar to the real structure, thus simulation parameters must be matched with measurements. Correlating FE analysis with 3D scanning laser Doppler vibrometry (3D-SLDV) is an state of the art approach to tune and validate numerical models in order to predict the system dynamics [1].

Applications reach from aviation [1, 2] and aerospace [3, 4] to automotive [5] and mechanical engineering [6]. Correlating simulations and measurements is important in tribology, since direct measurements in the contact zone are rarely feasible [7]. Dedicated friction test rigs (tribometers) enable controlled experiments [8], but the influence of machine dynamics on experimental results often remains unknown [9]. Therefore, numerical models of tribometers are used to study contact forces and structural vibrations to improve tribometer design and test quality [10].

Modern FE tools allow the modelling of complex structures by integrating e.g. bolt pretension, nonlinear contacts, nonlinear joints, or rigid bodies [11]. Reduced modal models can be embedded in multi body simulations with application-dependent contact models to form digital twins [10].

However, validation of FE models against measurements is challenged by (i) spatial limitations due to line-of-sight constraints, (ii) limited frequency resolution due to relatively short acquisition times in point-by-point scanning, (iii) mismatched geometries and (iv) multiple datasets with rather weak excitations [1, 2, 3]. Furthermore, correlation of large numerical models can cause stability or runtime problems.

To address these issues, a flexible workflow for validating FE models with multiple 3D-SLDV datasets is introduced. It provides, consistent import and processing of simulation and measurement results, spatial transformation of datasets into the

G. Brunbauer
AC2T research GmbH, Viktor Kaplan-Straße 2/C, 2700 Wiener Neustadt, Austria
e-mail: gottfried.brunbauer@ac2t.at

A.M. Puhwein · M.J. Hochrainer
University of Applied Sciences Wiener Neustadt Johannes Gutenberg-Strasse 3, 2700 Wiener Neustadt, Austria
e-mail: anton.puhwein@fhwn.ac.at; markus.hochrainer@fhwn.ac.at

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Model Validation and Uncertainty Quantification, 2026, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814146

FE reference frame, efficient interpolation of aligned mode shapes, and correlation using modal assurance criterion (MAC) [12]. The open, well structured implementation enables transparency and control throughout validation.

The approach is demonstrated for a twin-disc tribometer, showing reliable model validation as a base for model tuning and realistic digital-twin development.

Correlation Workflow

Figure 1 illustrates the proposed workflow: First, the simulation and identification datasets j from varying perspectives i are matched (step a). After reducing the data for efficient computation, measurement data sets are transformed by aligning scan grid and FE mesh. Then, simulated eigenvectors are interpolated at the scan points and correlation is quantified via MAC (step b). The various MAC matrices are finally assembled and reduced via mode matching (step c), see Section for a detailed description.

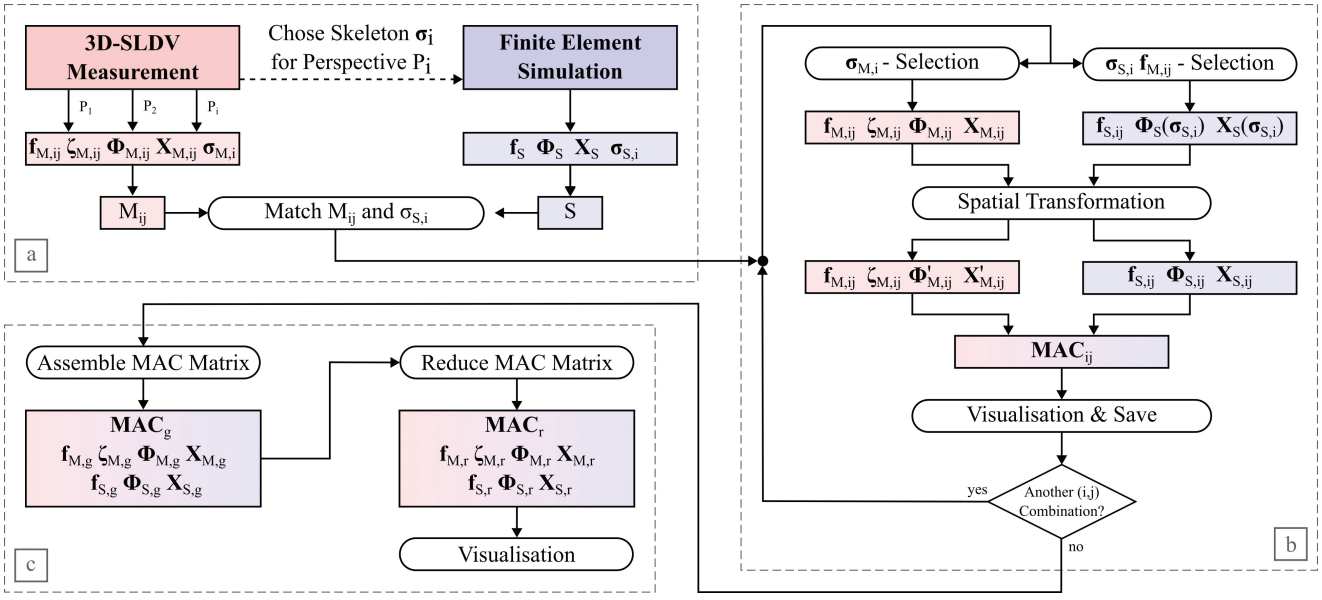


Fig. 1 Correlation workflow: a) load simulation S and measurement M_{ij} data, match FE node selection $\boldsymbol{\sigma}_{S,i}$ per perspective with M_{ij}, b) select frequency range ($\mathbf{f}_{M,ij}$ selection) and spacial distribution ($\boldsymbol{\sigma}_{M,i}$ and $\boldsymbol{\sigma}_{S,i}$ selection), transform M_{ij}, interpolate FE mode shapes to scan point coordinates, compute $\boldsymbol{MAC}_{ij}$ per combination and c) assemble/reduce results. $\mathbf{f}$ are natural frequencies, $\boldsymbol{\zeta}$ modal damping ratios, $\boldsymbol{\Phi}$ is the modal matrix, $\mathbf{X}$ the coordinate frame and $\boldsymbol{\sigma}$ a point-ID set.

Use Case: Twin-Disc Tribometer

The workflow is demonstrated for a large-scale FE model (∼900 k DoF) of a twin-disc tribometer, see Figure 2 a), including nonlinear contacts, pretensioned bolts, nonlinear joints, springs and rigid bodies. The linear modal analysis of the nonlinear FE model contains 400 modes, rendering frequencies up to 2.8 kHz. The tribometer consists of two discs, mounted on independently driven shafts, pressed against each other with a defined force, see [13, 14] for details.

To capture the dynamics of the tribometer a 3D-SLDV measurement setup with triaxial accelerometer references is used. Impulse hammer measurements are captured for multiple components from five perspectives, see Figure 2 a). Furthermore, operational measurements are carried out with and without slippage between the discs [14].

The identification of modal parameters is based on stochastic subspace identification for both operational and experimental modal analysis [15].

Correlation Workflow

Regardless of the identification type, the data set

$$M_{ij} = \{\mathbf{f}_{M,ij},\, \boldsymbol{\zeta}_{M,ij},\, \boldsymbol{\Phi}_{M,ij},\, \mathbf{X}_{M,ij}\},$$

is stored for each perspective–identification pair (i, j). An optional user defined mask $\boldsymbol{\sigma}_{M,i}$, consisting of scan point-IDs, can also be used to reduce the size of each M_{ij}. Vector $\mathbf{f}$ contains natural frequencies, $\boldsymbol{\zeta}$ modal damping ratios, $\boldsymbol{\Phi}$ is the modal matrix and $\mathbf{X}$ the coordinate frame.

To export the modal simulation data

$$S = \{\mathbf{f}_S, \boldsymbol{\Phi}_S, \mathbf{X}_S\}$$

from the FE result file one can either use FE-tool scripting or extract the required information from the result file directly using e.g. an application programming interface [16]. For each scanning perspective i, FE node-ID sets $\boldsymbol{\sigma}_{S,i}$ are defined to restrict subsequent processing to the relevant geometry regions. Geometric features (surfaces, edges) can be used to define the regions of interest.

Prior to the spatial transformation of the measurement results (alignment of point clouds), the datasets M_{ij} and S are matched. Each M_{ij} is mapped to the corresponding $\boldsymbol{\sigma}_{S,i}$ based on a user-defined map. S is then reduced for each M_{ij} to the $\boldsymbol{\sigma}_{S,i}$ node-IDs and to the frequency range of interest (pass band $\mathbf{f}_{M,ij}$), yielding the reduced simulation data $\mathbf{f}_{S,ij}$, $\boldsymbol{\Phi}_S(\boldsymbol{\sigma}_{S,i})$ and $\mathbf{X}_S(\boldsymbol{\sigma}_{S,i})$.

To determine the used point cloud alignment the least squares problem

$$L_i = \sum_k \left\| \mathbf{x}_{S,k} - \left(\mathbf{R}_i \, \mathbf{x}_{M,k} + \mathbf{t}_i \right) \right\|^2 \rightarrow \, min,$$

is solved for every perspective i for k manually selected reference points $\mathbf{x}_{S,k}$ and $\mathbf{x}_{M,k}$, rendering the rotation matrix $\mathbf{R}_i$ as well as the transformation vector $\mathbf{t}_i$. M_{ij} is aligned to S by the rigid transform

$$\boldsymbol{\Phi}'_{M,ij} = \mathbf{R}_i \, \boldsymbol{\Phi}_{M,ij}, \qquad \mathbf{X}'_{M,ij} = \mathbf{R}_i \, \mathbf{X}_{M,ij} + \mathbf{t}_i.$$

To minimise computational costs and inaccurate extrapolation, FE nodes without corresponding measurement points are removed yielding the reduced modal matrix $\boldsymbol{\Phi}_{S,ij}$ and coordinates $\mathbf{X}_{S,ij}$, see Figure 2 b).

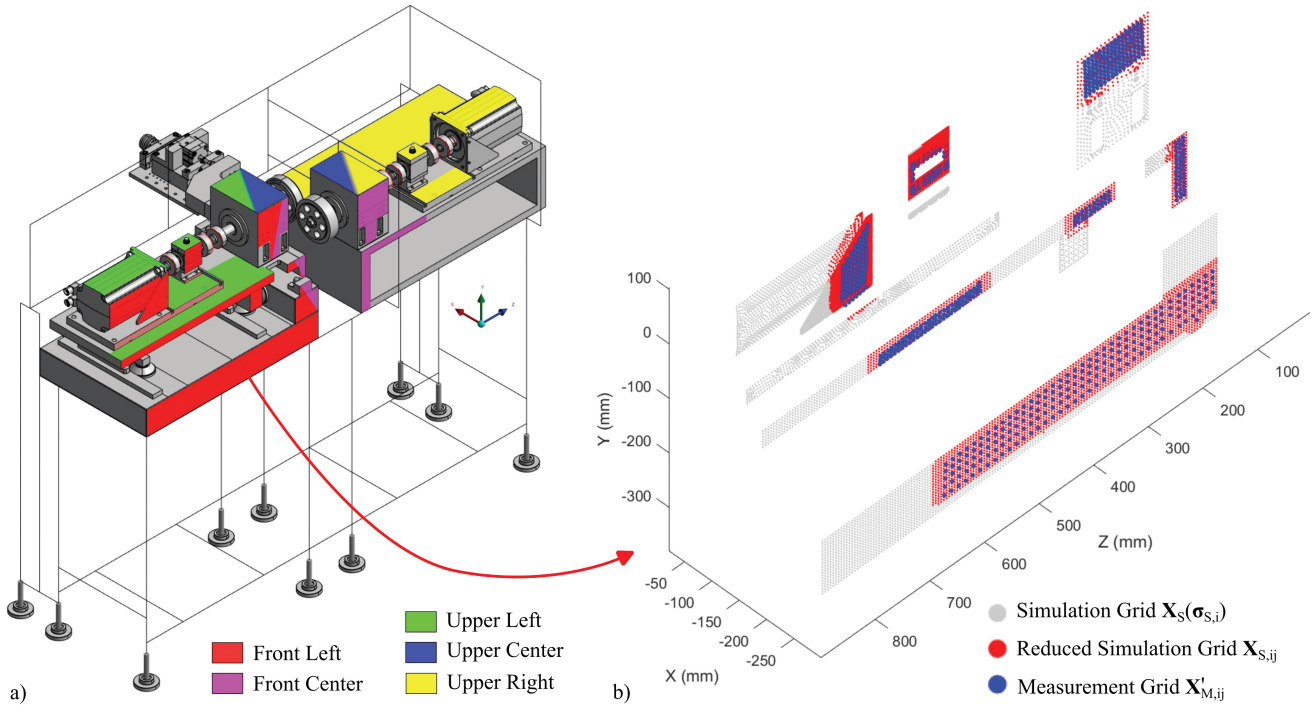


Fig. 2 a) FE model of the twin-disc tribometer with highlighted scanning perspectives, $\boldsymbol{\sigma}_{S,i}$ respectively. b) Exemplary transformation result for the perspective *Front Left*: The measurement grid $\mathbf{X}'_{M,ij}$ is already transformed to the simulation mesh $\mathbf{X}_S(\boldsymbol{\sigma}_{S,i})$ yielding the reduced coordinates $\mathbf{X}_{S,ij}$ for a distance threshold of 20 mm.

Because the FE nodes and measurement points do not coincide, the FE results $\boldsymbol{\Phi}_{S,ij}$ are interpolated at the transformed

measurement points $\mathbf{X}'_{M,ij}$ yielding the interpolated modal matrix $\tilde{\mathbf{\Phi}}_{S,ij}$. For each identification dataset the

$$\mathbf{MAC}_{ij} = \frac{\left| \left(\mathbf{\Phi}'_{M,ij}\right)^{\top} \tilde{\mathbf{\Phi}}_{S,ij} \right|^2}{\left(\left(\mathbf{\Phi}'_{M,ij}\right)^{\top} \mathbf{\Phi}'_{M,ij}\right) \left(\tilde{\mathbf{\Phi}}_{S,ij}^{\top} \tilde{\mathbf{\Phi}}_{S,ij}\right)},$$

is computed. Node ID selection (spatial reduction) and frequency-band selection (frequency reduction), decrease processing requirements substantially.

All $\boldsymbol{MAC}_{ij}$ are assembled to a global matrix $\mathbf{MAC}_g$, by clustering the subsets and thereby eliminating the different perspectives. Then, the reduced matrix $\mathbf{MAC}_r$ is obtained from $\mathbf{MAC}_g$ in a three step procedure. First, the size is reduced by neglecting mode pairs whose MAC is below a predefined threshold value τ. Next, combinations with a relative natural frequency deviation of

$$\frac{|f_M - f_S|}{f_S} \leq \delta,$$

are selected as valid matches where δ defines the frequency corridor. Finally, a square matrix is obtained by keeping unique mode pairs, based on the maximum MAC values, while ensuring each simulated and each measured mode is used only once.

Results and Discussion

Figure 3 a) shows the valid matches for the twin-disc tribometer within a frequency corridor of $\delta = 0.3$ and a value $MAC \geq 0.7$ up to 500 Hz. The figure clearly illustrates whether modes with high correlation cluster along the main diagonal or deviate significantly in frequency. In the study 38 unique mode pairs can be selected within the valid matches and are further investigated. Figure 3 b) shows the selected mode pairs of $\mathbf{MAC}_r$ in matrix form. Two exemplary images of the

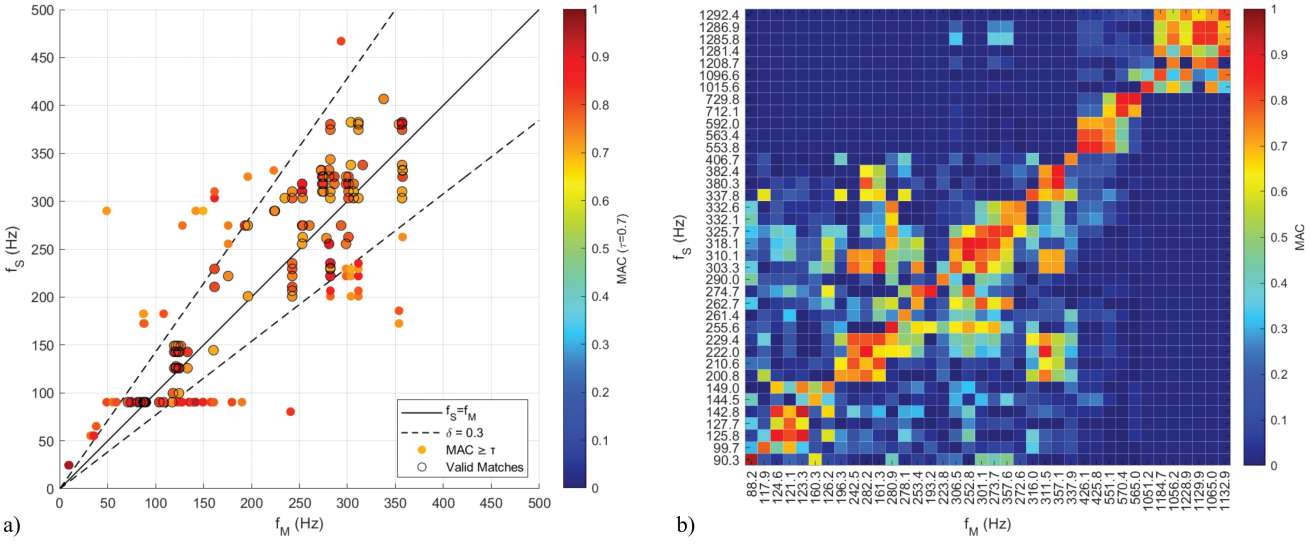


Fig. 3 Results for twin-disc tribometer for $\delta = 0.3$ and $\tau = 0.7$ a) Mode pairs with $MAC \geq 0.7$ and valid mode matches within δ corridor b) Reduced matrix $\mathbf{MAC}_r$ for unique mode pairs within valid matches.

mode animation with normalised and scaled amplitudes is given in Figure 4, where good qualitative agreements of the mode shapes can be observed.

The results show, that the proposed workflow robustly recovers mode pairs across five perspectives and different identification data sets. The entire computation, including data loading and user inputs, only takes a few minutes. Main practical limitations arise from incomplete scan coverage, high local mode density, noisy measurement data and varying reference sensors, preventing combined identification. These limitations can be further mitigated by using the FE model to focus on measurement coverage in frequency ranges and geometry areas with high mode density and by increasing measurement duration per scan point. The reduced set $\mathbf{MAC}_r$ provides compact input for visualisation and targeted model updating.

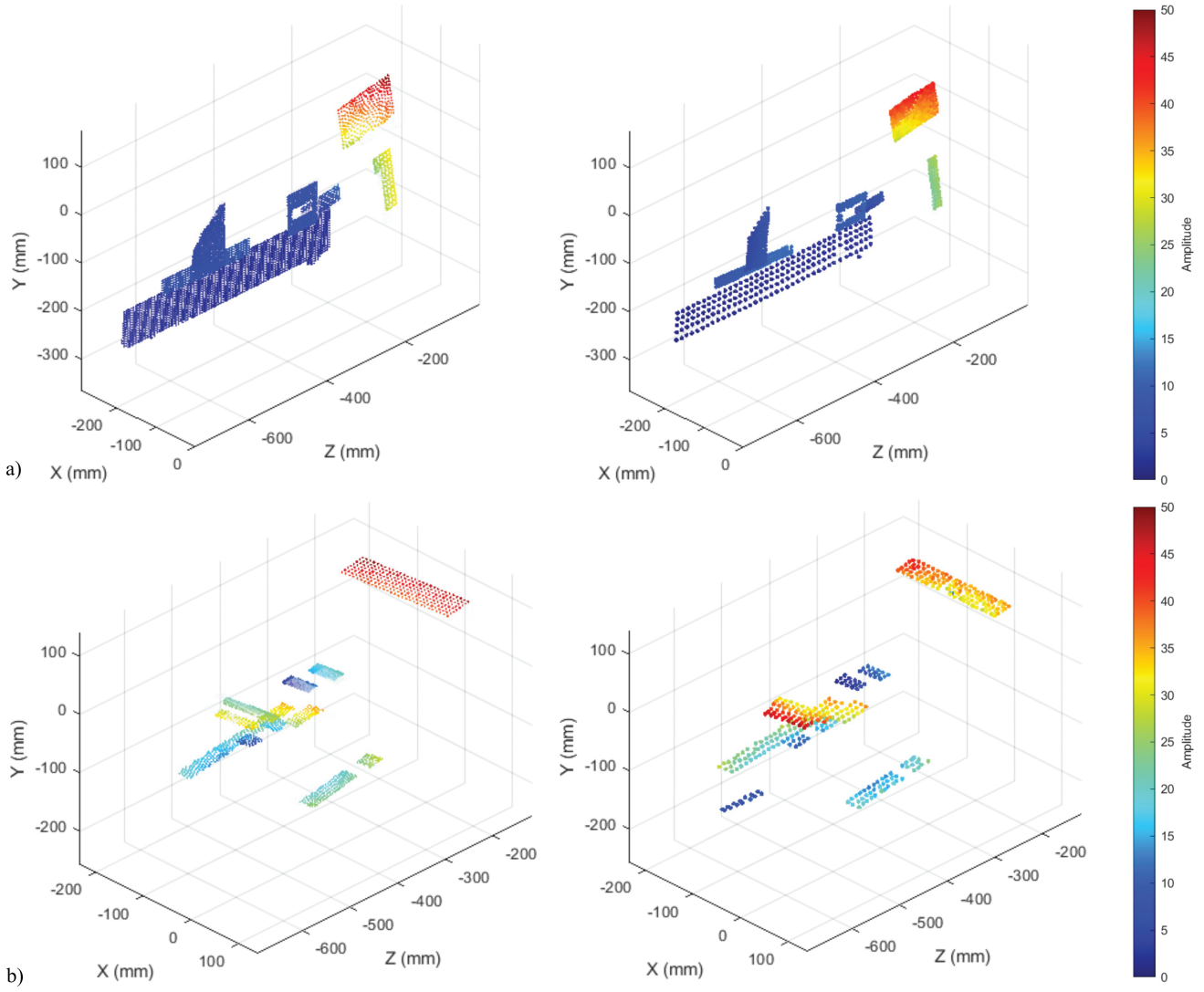


Fig. 4 Exemplary normalised visualisation of matched mode pairs between simulation (left) and measurement (right), scale factor 50
a) Perspective *Front Left*, $MAC = 0.96$, $f_S = 90.30\,Hz$, $f_M = 80.15\,Hz$, $\delta = -2.38\%$
b) Perspective *Upper Left*, $MAC = 0.89$, $f_S = 125.80\,Hz$, $f_M = 124.56\,Hz$, $\delta = -0.98\%$

Conclusion

A workflow for validating complex FE models with 3D-SLDV data is presented, which processes simulation and identification results, aligns mismatched coordinate frames, interpolates FE mode shapes to the scan grids and quantifies correlation via MAC. User input is kept to a minimum and is mainly required for coordinate transformation. On a large FE model of a twin-disc tribometer the workflow proved robust and practical, enabling reliable model validation and a compact one-to-one set of matched modes for visual inspection and further analysis. The correlation strongly depends on the measurement quality and the spatial scan coverage. For large structures with high modal density and localised modes, sparse scan coverage, or regions with line-of-sight restrictions, the correlation quality may be limited, resulting in non-diagonal matches or gaps in the assembled MAC matrix.

Acknowledgments The presented results were realized as part of the COMET Centre InTribology (FFG no. 906860), a project of the “Excellence Centre for Tribology” (AC2T research GmbH). InTribology is funded within the COMET – Competence Centres for Excellent Technologies Programme by the federal ministries BMIMI and BMWET as well as the federal states of Niederösterreich and Vorarlberg based on financial support from the project partners involved. COMET is managed by The Austrian Research Promotion Agency (FFG). The authors acknowledge the "Gesellschaft für Forschungsförderung Niederosterreich m.b.H." through its FTI PhD Funding Programme (FTI23-D-028) for the financial support.

References

1. Ben Weekes and David Ewins. Effective Use of Scanning Laser Doppler Vibrometers for Modal Testing. In Randall Allemang, editor, *Topics in Modal Analysis II, Volume 8*, pages 43–58. Springer International Publishing, Cham, 2014. Series Title: Conference Proceedings of the Society for Experimental Mechanics Series.
2. Jerome Eichenberger and Joerg Sauer. Validating Complex Models Accurately and Without Contact Using Scanning Laser Doppler Vibrometry (SLDV). In Dario Di Maio and Javad Baqersad, editors, *Rotating Machinery, Optical Methods & Scanning LDV Methods, Volume 6*, pages 113–124. Springer International Publishing, Cham, 2022. Series Title: Conference Proceedings of the Society for Experimental Mechanics Series.
3. Song Yan, Bin Li, Feng Li, and Binchao Li. Finite element model updating of liquid rocket engine nozzle based on modal test results obtained from 3-D SLDV technique. *Aerospace Science and Technology*, 69:412–418, October 2017.
4. Randolph H. Cabell. Vibration Response Models of a Stiffened Aluminum Plate Excited by a Shaker. In *Proceedings of Noise-Con 2008*, Dearborn, MI, USA, July 2008.
5. Luca Pagan and Kelvin Lake. Modal Analysis: A Comparison between Finite Element Analysis (FEA) and Practical Laser Doppler Vibrometer (LDV) Testing. In *2017 UKSim-AMSS 19th International Conference on Computer Modelling & Simulation (UKSim)*, pages 75–80, Cambridge, UK, April 2017. IEEE.
6. G. Martin, E. Balmes, G. Vermot des Roches, and T. Chancelier. Squeal measurement with 3D Scanning Laser Doppler Vibrometer: handling of the time varying system behavior and analysis improvement using FEM expansion. In *Proceedings of ISMA2018 and USD2018*, Leuven, Belgium, 2018. KU Leuven, Department of Mechanical Engineering.
7. Horst Czichos and Karl-Heinz Habig, editors. *Tribologie-Handbuch: Tribometrie, Tribomaterialien, Tribotechnik*. Springer Fachmedien Wiesbaden, Wiesbaden, 2015.
8. Anthony Ngo and Michael Handschuh. An Experimental Investigation on the Effect Operating Procedure Has on Tribometer Measurements. In *Volume 10: 2024 International Power Transmission and Gearing Conference (PTG)*, page V010T10A014, Washington, DC, USA, August 2024. American Society of Mechanical Engineers.
9. Josef Prost, Guido Boidi, Thomas Lebersorger, Markus Varga, and Georg Vorlaufer. Comprehensive review of tribometer dynamics-Cycle-based data analysis and visualization. *Friction*, 10(5):772–786, May 2022.
10. Zilin Li, Xiaocui Wang, Qi Zhang, Zhenqun Guan, Jiliang Mo, and Huajiang Ouyang. Model reduction for friction-induced vibration of multi-degree-of-freedom systems and experimental validation. *International Journal of Mechanical Sciences*, 145:106–119, September 2018.
11. Christof Gebhardt. *Praxisbuch FEM mit ANSYS Workbench: Einführung in die lineare und nichtlineare Mechanik*. Hanser eLibrary. Hanser, München, 3., aktualisierte auflage edition, 2018.
12. David J. Ewins. *Modal testing: theory, practice and application*. Number 10 in Mechanical engineering research studies engineering dynamics series. Research Studies Press, Baldock, 2. edition, 2000.
13. A. Mario Puhwein, Balazs Jakab, Christoph Haslehner, Widder Florian, Viktoria Thalhammer, Markus J. Hochrainer, Georg Vorlaufer, and Markus Varga. Investigation of the Dynamic Influences of a Two-Disc Tribometer on Wear via High-End Cameras and Vibrometers. In Javad Baqersad, Dario Di Maio, and Dan Rohe, editors, *Computer Vision & Laser Vibrometry, Vol. 6*, pages 69–83. Springer Nature Switzerland, Cham, 2025. Series Title: Conference Proceedings of the Society for Experimental Mechanics Series.
14. A. M. Puhwein, A. Biederbeck, B. Jakab, M. Varga, and M. J. Hochrainer. Experimental Modal Analysis of a Twin-Disc Tribometer Using a 3D-Scanning Laser Doppler Vibrometer. In *Proceedings of the 43rd IMAC, A Conference and Exposition on Structural Dynamics 2025*, Orlando, USA, February 2025. Manuscript submitted.
15. Rune Brincker and Carlos E. Ventura. *Introduction to Operational Modal Analysis*. Wiley, 1 edition, August 2015.
16. ANSYS, Inc. ansys.dpf.core (PyDPF-Core) Python Library, 2025.

River Rapids

Model Validation and Uncertainty Quantification, Volume 3

Garrison Flynn ■ Kyle Neal ■ Scott Ouellette

Model Validation and Uncertainty Quantification, Volume 3: Proceedings of the 44th IMAC, A Conference and Exposition on Structural Dynamics, 2026, the third volume of ten from the Conference brings together contributions to this important area of research and engineering. The collection presents early findings and case studies on fundamental and applied aspects of Model Validation and Uncertainty Quantification, including papers on:

- Model Validation for Digital Twinning
- Challenges on Model Form Uncertainty
- Surrogate Modeling
- Uncertainty Quantification and Propagation in Structural Dynamics
- Bayesian Methods
- MVUQ for Complex Systems

ISBN 978-87-438-1442-9



River Publishers