ADVANCED MACHINE LEARNING APPROACHES FOR FORCE RECONSTRUCTION AND STATE ESTIMATION IN STRUCTURAL DYNAMICS

SHUO HAO

PhD

The Hong Kong Polytechnic University

2025

The Hong Kong Polytechnic University

Department of Civil and Environmental Engineering

Advanced Machine Learning Approaches for Force Reconstruction and State Estimation in Structural Dynamics

Shuo Hao

A Thesis Submitted in Partial Fulfillment of Requirements for

the Degree of Doctor of Philosophy

December 2024

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

____________________(Signed)

Shuo Hao (Name of student)

Dedicated to my family

ABSTRACT

Forward problems in the structural dynamics typically involve simulating structural responses under specified conditions through constructed structural models, which are fundamental problems from the design of new structures and assessment of existing structures. Conversely, inverse problems in structural engineering often serve as dual counterparts to many forward problems, with their models typically evolving from various forward problem models. In existing structures, modern sensing technologies can effectively acquire structural information, enabling the inference of valuable structural insights based on measured data and inverse problem models. Generally, valuable structural information refers to structural parameters, external forces acting on the structures, and responses at unmeasured locations of structures.

The first part of the thesis introduces a nonparametric Bayesian multi-task learning framework for time-domain force reconstruction. Traditional methods often struggle with ill-posed deconvolution problems and uncertainties. By assigning the Gaussian process (GP) priors to force functions within a Bayesian context, the proposed method effectively mitigates these issues. This approach leverages the relationship between loads and responses through the convolution operator, resulting in responses that also follow a GP. A joint Gaussian distribution across multiple tasks enables closed-form posterior distributions of the forces based on measured responses. The framework is validated through simulations on a truss bridge and experiments on a frame structure subjected to impact loads, demonstrating high accuracy and efficient uncertainty quantification.

Building on this framework, the thesis extends its application to reconstruct transient aerodynamic loads on high-speed maglev vehicles using onboard acceleration measurements. Traditional estimation methods are often time-consuming and costly. The novel framework

utilizes an inverse mathematical model derived from a calibrated maglev vehicle model and integrates the multi-task GP algorithm to reconstruct aerodynamic loads efficiently. The method treats the reconstructed loads as multiple GP, allowing for closed-form posterior calculations. Validation with data from maglev trains passing through double-track tunnels confirms the framework's effectiveness.

To enhance state estimation capabilities, the thesis presents two complementary approaches. The first is a Bayesian-based multi-fidelity GP method that employs a time-delayed GP model to capture the statistical correlations between sensor observations from preceding time steps and the current structural states. Training data for this model is generated using a simulated finite element model subjected to synthetic excitations. To further enhance the accuracy, the proposed multi-fidelity GP model integrates data from both simulations and actual measurements. This integration effectively combines the low-fidelity information from simulations with high-fidelity direct measurements to provide posterior estimations with quantified uncertainties. Validation examples confirm the effectiveness of multi-fidelity GP in providing precise state estimations, thereby expectedly boosting the prediction reliability across various applications in structural dynamics.

The second approach explores a recurrent neural network (RNN) method enhanced with transfer learning. While GP offer probabilistic estimations, the RNN provides high-accuracy point estimations by learning correlations in multi-output problems. The RNN is trained using extensive response data from a calibrated finite element model under synthetic excitations. Transfer learning adapts the RNN to real-structure predictions using actual measurement data. A fine-tuning strategy adjusts parameters in the RNN initial layers while keeping output layers fixed, ensuring effective convergence. Numerical and experimental validations show that the RNN models significantly outperform traditional methods, especially in complex, multi-output state predictions.

LIST OF PUBLICATIONS

Journal articles:

S. Hao, S.M. Wang, Y.Q. Ni, Z.W. Chen, M.O. Adeagbo, Multi-task Gaussian Processes based transient aerodynamic load reconstruction for maglev vehicle using acceleration response. *Journal of Sound and Vibration*, 596 (2025) 118754. https://doi.org/10.1016/j.jsv.2024.118754.

S. Hao, H.W. Li, Y.Q. Ni, W. Zhang, L. Yuan, State estimation in structural dynamics through RNN transfer learning. *Mechanical Systems and Signal Processing*, 233 (2025) 112767. https://doi.org/10.1016/j.ymssp.2025.112767.

W. Zhang, S.M. Wang, Y.Q. Ni, X. Yuan, Y. Feng, L. Yuan, **S. Hao**, Physics-enhanced multi-fidelity neural ordinary differential equation for forecasting long-term creep behavior of steel cables. *Thin-Walled Structures*, 208 (2025) 112846. https://doi.org/10.1016/j.tws.2024.112846.

H.W. Li, **S. Hao**, Y.Q. Ni, Y.W. Wang, Z.D. Xu, Hybrid structural analysis integrating physical model and continuous-time state-space neural network model. *Computer-Aided Civil and Infrastructure Engineering*, 40 (2024) 166-180. https://doi.org/10.1111/mice.13282.

H.W. Li, J. Zhou, **S. Hao**, Y.Q. Ni, Z.D. Xu, Dynamic modeling and substructuring analysis leveraging long short-term memory neural network. *Structures*, 70 (2024) 107602. https://doi.org/10.1016/j.istruc.2024.107602.

E.Z. Rui, G.Z. Zeng, Y.Q. Ni, Z.W. Chen, **S. Hao**, Time-averaged flow field reconstruction based on a multifidelity model using physics-informed neural network (PINN) and nonlinear information fusion. *International Journal of Numerical Methods for Heat & Fluid Flow*, 34 (2024) 131-149. https://doi.org/10.1108/hff-05-2023-0239.

S. Hao, Y.Q. Ni, S.M. Wang, Probabilistic identification of multi-DOF structures subjected to ground motion using manifold-constrained Gaussian processes. *Frontiers in Built Environment*, 8 (2022) 932765. https://doi.org/10.3389/fbuil.2022.932765.

L. Yuan, Y.Q. Ni, X. Deng, **S. Hao**, A-PINN: Auxiliary physics informed neural networks for forward and inverse problems of nonlinear integro-differential equations. *Journal of Computational Physics*, 462 (2022) 111260. https://doi.org/10.1016/j.jcp.2022.111260.

S.M. Wang, G.F. Jiang, Y.Q. Ni, Y. Lu, G.B. Lin, H.L. Pan, J.Q. Xu, **S. Hao**, Multiple damage detection of maglev rail joints using time-frequency spectrogram and convolutional neural network. *Smart Structures and Systems*, 29 (2022) 625-640. https://doi.org/10.12989/sss.2022.29.4.625.

Conference papers:

S. Hao, S.M. Wang, Z.W. Chen, W. Zhang, Y.Q. Ni, Inverse reconstruction of unsteady aerodynamic loads acting on railway vehicles. In *14th International Workshop on Structural Health Monitoring*, San Francisco, 2023.

L. Yuan, Y.Q. Ni, **S. Hao**, W. Zhang, From initial to final state: Single step prediction of structural dynamic response. In *14th International Workshop on Structural Health Monitoring*, San Francisco, 2023.

W. Zhang, Y.Q. Ni, L. Yuan, **S. Hao**, S.M. Wang, Structural parameter identification with a physics-informed neural networks-based framework. In *14th International Workshop on Structural Health Monitoring*, San Francisco, 2023.

S. Hao, S.M. Wang, Y.Q. Ni, Kernel ridge regression-based force identification in the time domain. In *Engineering Mechanics Institute Conference*, Atlanta, 2023.

S. Hao, S.M. Wang, Y.Q. Ni, Gaussian Process-based non-uniform Fourier transform. In *8th World Conference on Structural Control and Monitoring*, Orlando, 2022.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my heartfelt gratitude to my advisor, Prof. Yi-Qing Ni, who has served as a guiding light throughout my more than four years of PhD studies. When I first joined the research group, he taught me to focus on fundamental key technologies in scientific research and to grasp their underlying philosophies. He also consistently encouraged me to stay informed about cutting-edge science and technology, and to consider the potential benefits of applying them to our field. Through over four years of his patient and insightful mentorship, I have gained immensely from Prof. Ni's guidance. As I approach graduation, I sincerely thank Prof. Yi-Qing Ni; his spirit and visions as a researcher would consistently motivate me on my future way of life.

I am also grateful to Dr. Su-Mei Wang for taking over the role as my co-supervisor. I deeply appreciate having her supervision that not only cares for my well-being but also fosters my self-discipline, enabling me to achieve this final accomplishment.

I want to thank the financial supports from The Hong Kong Polytechnic University. I would like to express my special thanks to Dr. Xiangyun Deng. Without his initial encouragement and guidance, I would not have had the opportunity to pursue my fulfilling studies at The Hong Kong Polytechnic University. To me, Dr. Deng is like a respected elder brother, and I deeply admire his noble demeanor and approach to life and work.

I would like to extend my sincere appreciation to my colleagues at the National Rail Transit and Electrification and Automation Engineering Technology Center (Hong Kong Branch). I am particularly thankful to Mr. Wing-Hong Kwan, Mr. Tai-Tung Wai, Mr. Guang Zhou, Mr. Yang Lu, Mr. Xiang-Xiong Li, and Mr. Qi-Fan Zhou for their unreserved support in conducting numerous experiments. My sincere thanks also go to Dr. Hong-Wei Li, Dr. Zheng-Wei Chen, and Dr. Mujib Adeagbo for their invaluable guidance and insightful suggestions for my PhD

studies. Dr. Hong-Wei Li instructed me with vibrational time series modeling using deep learning, Dr. Zheng-Wei Chen provided precious suggestions on aerodynamic effects of vehicle systems, Dr. Mujib Adeagbo helped on modal analysis and Bayesian model updating.

I am grateful to my dear friends, Dr. Lu Zhou, Dr. You-Wu Wang, Dr. Siqi Ding, Dr. Si-Xin Chen, Dr. Bei-Yang Zhang, Dr. Yunke Luo, Dr. Si-Yi Chen, Dr. Xin Ye, Dr. Yue Dong, Dr. Duo Zhang, Mr. Zhen Lin, Mr. Zhengyang Li, Mr. Xiang-Tao Sun, Mr. Zijian Hu, Ms. Mingxi Li, Ms. Qi Zhu, Mr. Yan-Ke Tan, Mr. Da-Zhi Dang, Mr. Yu-Ling Wang, Mr. You-Liang Zheng, Mr. Weijia Zhang, Mr. Lei Yuan, Mr. En-Ze Rui, Mr. Gao-Feng Jiang, Mr. Shengyuan Liu, Mr. Yingnan Hu, Ms. Yuanheng Zhang, Mr. Binyu Yang, Ms. Yinghong Lin, whose friendship and encouragement have been a source of strength and joy throughout my doctoral journey. Their understanding, laughter, and support provided me with the motivation and resilience needed to overcome challenges and stay focused on my goals. Thank you all for being an integral part of this journey.

Lastly, I would like to thank my wife, Mrs. You Li, for her unwavering love and support. I am also deeply grateful to my mother, Mrs. Yan-Hua Shi; my father, Mr. Zhi-Huan Hao; and my brother, Mr. Ze Hao, for their constant encouragement and belief in me. The steadfast support of my family has been instrumental in the completion of this thesis.

TABLE OF CONTENTS

Chapter 6. Transfer learning of recurrent neural networks for enhanced state estimation in structural dynamics 153

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

The abbreviations used in this thesis apply to both the singular and plural forms of the terms listed.

AKF	Augmented Kalman filter
BiRNN	Bidirectional recurrent neural network
CNN	Convolutional neural network
CSD	Cross spectral density
DKF	Dual Kalman filter
DOF	Degree of freedom
FBG	Fiber Bragg grating
FDD	Frequency domain decomposition
FE	Finite element
FRF	Frequency response function
GP	Gaussian process
GPR	Gaussian process regression
HFGP	High-fidelity Gaussian process
LFGP	Low-fidelity Gaussian process
LSTM	Long short-term memory
MFGP	Multi-fidelity Gaussian processes
MLP	Multi-layer perceptron
MTGP	Multi-task Gaussian processes
NRMSE	Normalized root mean square error
PLL	Predictive log likelihood
PSD	Power spectral density
RNN	Recurrent neural network
RKHS	Reproducing kernel Hilbert space
SHM	Structural health monitoring
SNR	Signal-to-noise ratio
SVD	Singular value decomposition
TL-(Bi)RNN	Transfer learning of (bidirectional) recurrent neural network

Chapter 1.

Introduction

1.1. Research background

Inverse problems in structural dynamics, particularly force reconstruction and state estimation, are crucial for understanding and predicting the behavior of complex engineering structural systems. These problems primarily deal with deducing unseen causes, such as external forces or unmeasured internal states, based on observable measurements (e.g., accelerations, strains, and displacements). Successfully solving these inverse problems is essential, as they underpin various practical applications including design validation, structural control, structural health monitoring (SHM), and the deployment of proactive maintenance strategies (Gallet et al., 2022).

Force reconstruction allows engineers to determine the dynamic loads a structure experiences during various events, such as earthquakes (Taher et al., 2021), wind gusts (Pham et al., 2024), and operational vibrations (Firus et al., 2022). The reconstructed force information is critical for validating computational models and optimizing structural designs. Conversely, state estimation involves determining internal states including displacements and velocities based on measurable responses. Effective state estimation could allow for continuous structural control and health monitoring (Ou et al., 2015; Panda et al., 2021; Galaz-Palma et al., 2022; Kamariotis et al., 2023), facilitating the mitigation of undesirable vibrations and the early detection of damage, degradation, or potential failures. The proactive approach not only enhances the safety and resilience of infrastructure but also extends the service life of structures by enabling timely interventions.

Despite their importance, inverse problems in structural dynamics are inherently challenging due to several factors (Stuart, 2010; Turco, 2015), including challenges in adapting inverse models to specific applications, the ill-posed nature of solutions, system uncertainties, and high computational requirements. Traditional methods for solving inverse problems, such as various regularization techniques (Engl et al., 1996) and Kalman filtering (Iglesias et al., 2013), have been extensively employed with varying degrees of success. However, these approaches often struggle to improve solution accuracies, efficiently handle uncertainties, and integrate information from multiple sources, thereby limiting their scalability and applicability to many real-world scenarios.

In recent years, machine learning has emerged as a transformative tool in engineering (Bishop, 2006; Murphy, 2012, Goodfellow, 2016), offering innovative solutions to the persistent challenges of inverse problems in structural dynamics (Thai 2022; Cunha et al., 2023). Advanced machine learning approaches, including probabilistic machine learning and various deep learning methods, have demonstrated significant potential in enhancing both force reconstruction and state estimation processes.

Probabilistic machine learning methods, such as nonparametric Bayesian approaches, provide an efficient framework for uncertainty quantification (Rasmussen and Williams, 2006; Ghahramani, 2015). These methods enable more reliable and interpretable inferences from noisy and incomplete data by explicitly modeling the uncertainties inherent in both the measurements and the inverse problems. This probabilistic perspective facilitates robust decision-making and risk assessment in structural engineering applications, ensuring that the reconstructed forces and estimated states are not only accurate but also trustworthy under varying conditions.

Deep learning techniques excel at capturing high-dimensional, complex, nonlinear relationships inherent in excitations and structural responses (Ye et al., 2019; Cha et al., 2024). Deep learning architectures can learn intricate patterns from large datasets, thereby improving the accuracy and robustness of state estimations and force reconstructions. Additionally, advancements in deep learning have enabled the development of models that can adapt to diverse structural dynamic conditions, enhancing their applicability across various engineering contexts.

Multi-fidelity modeling and transfer learning further extend the capabilities of machine learning approaches for inverse problems in structural dynamics (Weiss et al., 2016; Brevault et al., 2020; Cunha et al., 2023). These methods leverage data from various sources with differing levels of fidelity and computational and measurement costs, effectively balancing the trade-off between data quality and resource utilization. By integrating information from multiple fidelities, these techniques enhance the prediction accuracy of machine learning models for force reconstruction and state estimation, making them more versatile and efficient.

The integration of advanced machine learning approaches into force reconstruction and state estimation offers several advantages over traditional methods, including enhanced accuracy and robustness, efficient uncertainty quantification, and adaptability to multi-source and multi-fidelity data. This thesis aims to advance the resolution of these critical inverse problems in structural dynamics through the application and development of state-of-the-art machine learning methodologies. By leveraging probabilistic machine learning, deep learning, multi-fidelity modeling, and transfer learning, the research seeks to contribute more accurate, reliable, and efficient solutions for a wide array of practical applications in the field of inverse problems in structural dynamics.

1.2. Research objectives

The thesis aims to advance the field of typical inverse problems in structural dynamics through the development and application of sophisticated machine learning methodologies. The following objectives delineate the scope and direction of this study.

First, to construct a framework for efficient uncertainty quantification of forces given the noisy measurements of structures. Existing research in force reconstruction has utilized Bayes' theorem to derive posterior distributions of forces based on structural measurements. However, these approaches often encounter significant computational challenges due to the high-dimensional nature of force representations, making the posterior calculations exceedingly resource-intensive. To address this limitation, there is a need to develop an efficient method for uncertainty quantification of forces that circumvents the drawbacks of high-dimensional vector representations. Instead, forces should be modeled as random processes, which inherently capture their stochastic nature while reducing dimensional complexity. By adopting this probabilistic framework, the posterior calculation process becomes more tractable and computationally efficient, enabling accurate and reliable uncertainty quantification even in the presence of noisy measurements. This advancement will enhance the feasibility of implementing Bayesian-based force reconstruction methods in practical structural dynamics applications.

Second, enhancing the accuracy of force reconstruction through advanced prior modeling of forces. The precision of force reconstruction is heavily dependent on the effectiveness of prior models used in Bayesian inference. Traditional prior models often fail to capture the intricate distributions and functional behaviors of real-world forces, leading to less accurate posterior estimates. This objective focuses on enhancing the accuracy of force reconstruction by developing advanced prior models that treat reconstructed forces as continuous functions rather than discrete vectors. Utilizing nonparametric Bayesian techniques, such as Gaussian process (GP), allows for the incorporation of prior knowledge about the smoothness and continuity of forces. These advanced priors can better represent the temporal correlations inherent in structural forces, resulting in more accurate and reliable force estimations.

Third, constructing effective inverse models of state estimation using machine learning-based techniques. Although rigorous in mathematical formulations, traditional methods for state estimation can be limited by their reliance on some specific assumptions and predefined system models, which may not adequately capture the complexities of real-world structures. This research objective aims to construct robust inverse models for state estimation using advanced machine learning-based techniques, including Gaussian process regression (GPR) and recurrent neural networks (RNN). By training these models on extensive datasets derived from both simulated and experimental structural responses, the inverse models can learn to accurately infer internal states under various loading and boundary conditions. Machine learning approaches offer the flexibility to handle nonlinearities and adapt to diverse structural configurations, thereby enhancing the precision and applicability of state estimation processes.

Fourth, improving the accuracy of state estimation via multi-fidelity modeling and transfer learning. Multi-fidelity data are frequently encountered in various structural engineering practices, originating either from model simulations or multi-source measurements of structures. These multi-fidelity data differ in both size and quality, with each fidelity level offering distinct advantages and drawbacks. This objective aims to enhance the accuracy of state estimation by effectively utilizing the different levels of data through multi-fidelity modeling and transfer learning techniques. Specifically, the proposed multi-fidelity Gaussian processes (MFGP) approach will advance GPR models by integrating data of varying fidelities, thereby leveraging the strengths of each fidelity level to improve estimation precision. Additionally, RNN models will be adapted to handle multi-fidelity data through transfer learning, allowing the models to transfer knowledge from low-fidelity datasets to enhance performance on high-fidelity ones. By utilizing multi-fidelity modeling and transfer learning, this research seeks to significantly boost the robustness and scalability of state estimation methods. The enhanced methods are expected to be more versatile and effective across a wide range of applications in structural dynamics, enabling more reliable and efficient monitoring of complex structural systems.

1.3. Main contributions

To achieve the research objectives described above, several advanced machine learning approaches adaptable tailored to the goals have been developed. The primary contributions of the thesis are listed below.

An innovative force reconstruction method based on Bayesian nonparametric multi-task learning has been developed. This approach assigns GP priors to the force functions to address inverse problems and uncertainty quantification. By exploiting the convolution operator between the force and the structural response, the method demonstrates that the responses also follow a GP, with covariance functions derived from the forces' covariances. This correlation enables the formation of a joint Gaussian distribution over multiple variables (from multiple GP), incorporating both measurable structural responses and forces. Consequently, the closed-form posterior distribution of the forces given the measurements is obtained. Compared to existing Bayesian approaches for force reconstruction, this method offers a more efficient quantification of uncertainty propagation from measurements to forces.

A structured framework for aerodynamic load reconstruction in vehicular systems has been constructed. Specifically, the time-domain force reconstruction is applied to transient aerodynamic load reconstruction for maglev vehicles using onboard acceleration data. An inverse mathematical model, linking measured accelerations to external aerodynamic forces, is derived from a precisely calibrated maglev vehicle model. To tackle the inherent challenges in solving this inverse problem, the proposed nonparametric Bayesian multi-task learning approach is adopted. Validation using transient vibration data from a maglev train operating through a double-track tunnel indicates that the framework is both cost-effective and efficient for deriving aerodynamic loads, thereby proving its utility in high-speed maglev technology field tests.

The applicability of a time-delayed probabilistic modeling approach using GPR has been investigated for precise state estimation. This method leverages measured data from preceding time steps as inputs to estimate the current state, framing the development as a machine learning problem. Both training and test datasets are generated using a FE model subjected to synthetic excitations. By learning the mapping from time-delayed inputs to system states, the GPR model, with only a few parameters to optimize, effectively addresses the state estimation challenge while also quantifying the associated uncertainties in its predictions.

A novel GP-based approach that integrates multi-fidelity data has been developed to enhance the accuracy of state estimations. This method combines high-fidelity data (from precise physics-based models or actual measurements) with abundant low-fidelity data from simplified models. The resulting multi-fidelity GP (MFGP) models consistently outperform single-fidelity models, particularly in settings where both the availability and fidelity of data vary spatially and quantitatively.

A transfer learning framework for RNN has been proposed to achieve effective physics-data fusion for multi-output state estimation. Using training and test datasets generated from FE models with synthetic excitations, the RNN are initially trained and validated for state estimation accuracy. To improve predictions for real-world structures, the framework integrates data from both physics-based FE models and actual structural measurements. The dual-purpose strategy, where a portion of the collected sensor data is used as RNN input while the entire dataset supports transfer learning, enables the RNN models to adapt efficiently for real-structure state prediction tasks.

1.4. Thesis outline

Following the Introduction, this thesis is organized into six chapters; each part is briefly outlined below:

Chapter 2 presents a comprehensive review of the two investigated inverse problems, namely force reconstruction and state estimation, along with the fundamental machine learning approaches utilized in this study. This chapter examines the development of methods addressing these inverse problems and summarizes the challenges associated with their practical applications. To mitigate these challenges, the thesis adopts probabilistic machine learning techniques and deep learning. Consequently, the subsequent sections of Chapter 2 delve into the fundamental rationale behind these two methodological frameworks, summarizing their advantages and disadvantages and providing a foundation for the selection of corresponding methods in the subsequent chapters.

Chapter 3 explores a nonparametric Bayesian multi-task learning approach designed to address time-domain force reconstruction problems, which are characterized as ill-posed deconvolution challenges. This method incorporates GP priors for the reconstructed forces, effectively framing the deconvolution problem within a Bayesian framework. Within this framework, the posterior distributions of the forces, given the response measurements, become analytically tractable. The integration of GP into force reconstruction successfully manages the propagation of uncertainties from measurements to the reconstructed forces. To demonstrate the effectiveness of the proposed method, the chapter presents a numerical example involving a truss structure and an experimental example using a frame structure. The results highlight the high accuracy and efficient uncertainty quantification achieved by the nonparametric Bayesian multi-task learning approach.

Chapter 4 builds upon the methodology introduced in Chapter 3 by presenting a systematic framework for transient aerodynamic load reconstruction on maglev vehicles using onboard acceleration measurements. This framework encompasses multiple procedures, including maglev vehicle system modeling and calibration, construction of an inverse force reconstruction model, application of nonparametric Bayesian multi-task learning, and evaluation of force reconstruction results. The proposed framework is applied to a scenario where a maglev train passes through a double-track tunnel, utilizing field-collected data. The results validate the framework's effectiveness and underscore its practical applicability for future aerodynamic field testing of maglev vehicles.

Chapter 5 introduces the time-delayed MFGP approach for state estimation in structural dynamic systems. The time-delayed model for state estimation is conceptualized as a machine learning problem, where GPR is employed to learn the mapping. GPR facilitates probabilistic state estimation based on data generated from a physics-based FE model. To enhance state estimation by integrating additional measurements, the MFGP is proposed, allowing the GP model to leverage both low-fidelity and high-fidelity data for more accurate estimations of the structural system state. Two numerical examples are provided: a generic mass-spring-damper system and a 45-floor building subjected to wind excitations. Comparisons between the estimated states and reference data demonstrate that MFGP effectively maps high-dimensional, multi-fidelity, time-delayed data to the system state.

Chapter 6 complements the Bayesian-based time-delayed MFGP method within a deep learning framework. It introduces an RNN model for state estimation that excels in learning and exploiting correlations inherent in multi-output problems. To address the challenge of multi-fidelity data in practical state estimation, transfer learning is applied to a pre-trained RNN model based on multi-sensor measurements in SHM systems. In this study, the full dataset serves dual purposes: a portion of the data acts as input for the RNN model, while the complete dataset facilitates the transfer learning process. A novel method is proposed to ensure the convergence of the transfer learning process by fine-tuning parameters within the RNN cells at the network's front end while keeping those near the output layers frozen. This approach deviates from conventional transfer learning methods typically used for other neural network architectures and proves particularly beneficial for RNN models tailored for state estimation. Numerical and experimental studies validate that the proposed transfer learning of RNN (TL-RNN) approach seamlessly integrates both model-generated and actual measurement data. Under identical data acquisition conditions, TL-RNN models achieve significantly higher accuracy compared to state estimation models that rely solely on FE models.

Chapter 7 provides a summary of the current research work and offers recommendations for future investigations.

Chapter 2.

Literature review

This chapter presents a comprehensive review of the force reconstruction and state estimation problems, alongside the advanced machine learning techniques foundational to the research. It begins by exploring inverse problems within the domains of structural dynamics, with a particular emphasis on the inverse models employed for force reconstruction and state estimation. Building upon this foundational understanding, the chapter delves into probabilistic machine learning methods, highlighting their applications in efficient probabilistic modeling and multi-output estimation. Subsequently, the discussion advances to deep learning-based approaches, detailing sophisticated neural network architectures and the advantages of transfer learning in addressing practical engineering problems.

2.1. Inverse problems in structural dynamics

2.1.1. Definition and scope

Throughout the past several centuries of scientific and technological advancement, the formulation of first principles has enabled researchers and engineers to develop effective mathematical equations tailored to specific problems within their respective fields. These equations delineate the relationships between known and unknown variables. When combined with modern computational methods, they facilitate the determination of physical quantities of interest in complex systems, thereby enabling the prediction, estimation, or inference of phenomena. In structural engineering, a set of motion equations that characterize the behavior of structural systems could be formulated using the first principles. Forward problems are typically defined as the determination of structural responses given known geometric and material parameters, boundary conditions, and external excitations by solving these structural motion equations. Conversely, inverse problems are formulated in relation to forward problems, where the objective is to determine other quantities within the structural motion equations based on partially known structural responses (Gallet et al., 2022; Xu et al., 2023). The unknowns may include physical quantities preset in the forward problem or unmeasured responses (Lam et al., 2015; Mayes et al., 2020; Hou and Xia, 2021).

Inverse problems are inherently more complex than forward problems due to several factors (Lesnic, 2022). Firstly, the input for inverse problems consists of real-world measurements, which are often sparse in both space and time. Secondly, the acquired data are typically contaminated with noise, introducing uncertainty into the problem. Thirdly, the motion equations used to model the structure are approximations of the true behavior, thereby enhancing the uncertainty associated with the solutions. The sparsity of measurements results in low-rank issues when directly applying the structural motion equations to inverse problems. Additionally, noise in the data can render inverse problems ill-posed, where even minimal noise can significantly distort the results. Lastly, the inherent approximations in the motion equations contribute to further uncertainty, thereby impacting the accuracy and reliability of the inverse problem solutions.

In this thesis, addressing the key technical requirements within the fields of structural dynamics, two primary types of inverse problems are of concern: 1) force reconstruction and 2) state estimation of structural dynamic systems. In some existing research, force reconstruction is commonly referred to as load/force identification (Wu and Law, 2010; Yang et al., 2023), input estimation (Nayek and Narasimhan, 2020) as well, while state estimation is also termed as dynamic displacement/velocity estimation (Bhowmick et al., 2020; Gulgec et al., 2020). For the sake of clarity and consistency, the thesis will uniformly refer to them as force reconstruction and state estimation throughout this thesis.

Both force reconstruction and state estimation are pivotal inverse problems within the domains of structural dynamics. The foremost consideration in addressing these inverse problems is the determination of the appropriate input information to be utilized. As previously discussed, for inverse problems commonly encountered in the field of structural engineering, the inputs typically refer to structural data that can be directly measured using sensors or instrumentation devices. This characteristic inherently diversifies inverse problems relative to the forward problems, as the modeling of inverse problems varies depending on the type of measured data used to estimate the physical quantities of interest. Consequently, the methodologies employed to solve these models also differ accordingly. It is essential to delineate that the scope of force reconstruction and state estimation investigated in this thesis is situated within a vibration-based SHM framework (Avci et al., 2021; Kamariotis et al., 2022). Within this context, the primary inputs for these inverse problems consist predominantly of acceleration data. Although, in numerous instances, other vibration-related information, such as dynamic displacements and velocities of the structure (Narazaki et al., 2021;

Shao et al., 2021), strains of structural components (Torres et al., 2011; Karatas et al., 2020), or relative displacements between nodes (Li and Hao, 2016), etc., can also be measured through strategies proposed in various studies, the current state of technological advancement renders acceleration measurement techniques the most mature and reliable (Abdulkarem et al., 2020; Hassani and Dackermann, 2023). Moreover, many existing SHM systems have deployed a substantial number of accelerometers throughout the structure (Ni et al., 2009; Tcherniak and Mølgaard, 2017). The algorithms proposed in this thesis for force reconstruction and state estimation are designed to be directly implemented within these established systems.

Additionally, the emphasis on acceleration data as the majority input stems from the fact that some of the algorithms developed in this thesis also incorporate multi-sensor data fusion techniques to enhance the accuracy of inverse problem solutions (Wu and Jahanshahi 2020; Hassani et al., 2024). Therefore, in certain algorithms, a limited amount of additional measurement data beyond acceleration is also considered in some chapters. This approach leverages the strengths of multiple data sources to mitigate the inherent challenges associated with inverse problem-solving, such as measurement noise and deviation between structural FE model and real-world structure, thereby improving the precision and reliability of the estimations.

2.1.2. Advancements of force reconstruction problems and challenges

Structures are subjected to a variety of forces throughout their service life. The force information during specific conditions is essential for dynamic analysis and SHM (Uhl, 2007; Zhang et al., 2022). However, directly measuring these critical dynamic forces often faces numerous limitations (Ronasi and Nielsen, 2013; Niu et al., 2015). Consequently, load reconstruction is formulated as an inverse problem, which leverages other structural measurements, e.g., accelerations, to compute the force history that generates these measured responses, thereby estimating the critical force histories.

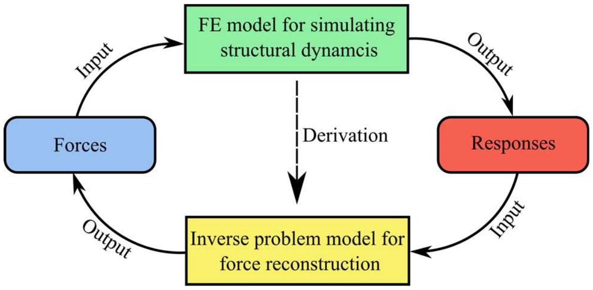


Figure 2.1. Demonstration of forward problem and inverse force reconstruction.

In the forward counterpart of this inverse problem, the FE model plays a pivotal role. As illustrated in Figure 2.1, the development of a physics-based FE model enables the computation of structural responses given specific forces. To perform force reconstruction, it is necessary to integrate the relationship between forces and partially measurable responses as represented in the FE model. This integration leads to the construction of an inverse problem model derived from the FE model, facilitating the backward calculation required for accurate force estimation.

Many studies have focused on the development of inverse problem models derived from FE models. These inverse problem models can be primarily categorized into frequency domain-based and time domain-based load reconstruction models, depending on the specific representations of the external forces required.

Within the frequency domain-based force reconstruction framework, the relationship between the frequency representations of structural responses and forces is typically established using frequency response functions (FRF). These FRF are derived from the FE model and serve as operators for computations at specified frequencies (Liu and Shepard, 2005; Law et al., 2011). Many frequency domain-based force reconstruction models consider the sparsity character of forces represented in frequency domain. To name a few, Rezayat et al. (2016) introduced a group-sparsity approach that effectively localizes and reconstructs dynamic forces in the presence of measurement noise by employing a structured penalty function and iterative algorithms. Qiao et al. (2016) proposed a general sparse regularization model using dictionary for force reconstructions, where the sparse reconstruction by separable approximation with different dictionaries is developed to solve the sparse regularization problem.

The time-domain force reconstruction models are normally constructed with two distinct approaches. The first approach is based on recursive computing, in which forces and measurements are treated as states at each time step. By recursively inputting the measurements into the force reconstruction models, the corresponding forces are calculated at each time step (Lourens et al., 2012; Naets et al., 2015). The second approach leverages the convolutional relationship between system responses and forces, deriving a convolutional operator based on the impulse responses of the FE model and ultimately abstracting force reconstruction to a deconvolution process (Jacquelin et al., 2003; Li and Lu, 2018).

The selection of inverse problem models for force reconstruction may largely depend on the force characteristics involved in the problem. Frequency domain-based force reconstruction models are particularly advantageous for stationary random forces, where the primary concern is not the exact force amplitude but rather the statistical properties of energy distribution with respect to frequencies (Xie et al., 2013). The recursive computing-based time-domain method is suitable for both stationary and non-stationary cases. Some recursive models even allow for online reconstruction, making them highly effective for real-time applications (Wang et al., 2021). The convolution-based inverse problem model for force reconstruction offers the best accuracy for non-stationary and transient forces, as reported by Ronasi et al. (2011), leading to the widespread development of convolution models in various transient load identification systems. Exhaustively evaluating all potential models incurs substantial computational costs, necessitating the judicious selection of inverse problem models based on practical requirements. Given that many critical applications demand the identification of non-constant transient dynamic loads in the field of structural dynamics (Choi et al., 2007; Ronasi et al., 2013; De Simone et al., 2019; Qiao and Rahmatalla, 2020), this study concentrates on convolution-based inverse problem models to better support the practical implementation of load reconstruction methodologies.

The deconvolution problems associated with force reconstruction are notorious for their ill-posed property, which is characterized by the lack of uniqueness or stability in the solutions (Kabanikhin, 2008). Directly inverting the transfer matrix for discretized deconvolution to calculate forces can result in numerically intractable issues or solutions that are highly sensitive to the system noise. To mitigate these challenges, one common strategy is the use of regularization methods. These methods impose additional constraints on the objective function to ensure that the solution resides within a specified vector space, thereby enhancing both stability and uniqueness. For example, Tikhonov regularization has been applied in several studies (Choi et al., 2007; Li et al., 2022), in which the l_2-norm is incorporated as a penalty term in the objective function, making the reconstruction less sensitive to measurement noise. Other fundamental penalty terms are also widely used according to the characteristics of the forces. For instance, the l_1-norm is more effective for sparse representations of forces (Zhang et al., 2022), and the fractional l_q-norm can incorporate adaptivity to the force profile, striking a balance between the l_2-norm and l_1-norm (Li and Lu, 2018). More complex regularization techniques, such as multiplicative regularizations (Aucejo and De Smet, 2017) and higher-order regularizations (Qiao and Rahmatalla, 2020), have proven to be more robust for force reconstruction, though they involve more parameters, complicating the analysis. Other methods such as truncated singular value decompositions (SVD) (Liu et al., 2014), iterative optimizations (Aucejo and De Smet, 2019), and recent advances in deep learning-based approaches (Wang et al., 2020a) also serve as effective forms of regularization for force reconstruction. These methods introduce constraints or additional information that stabilize the calculated result, offering promising solutions for addressing the ill-posedness of force reconstruction problems.

Recent research on force reconstruction has emphasized understanding the impact of multi-source uncertainties, which are broadly categorized into epistemic (resulting from incomplete knowledge, such as modeling errors) and aleatory (arising from inherent variability, such as measurement noise). These uncertainties critically influence the reliability of reconstructed forces, necessitating methods to rigorously quantify their propagation. Such quantification ensures the accuracy of downstream analyses and decisions dependent on these forces.

Many approaches have been developed, with Bayesian frameworks gaining prominence due to their ability to address both uncertainty types (Zhang et al., 2012). First, Bayesian methods allow epistemic uncertainties (e.g., imperfect model assumptions) to be encoded via prior probability distributions, offering intrinsic regularization. Second, aleatory uncertainties (e.g., sensor noise) are naturally embedded within the likelihood function, enabling a unified probabilistic treatment of multi-source errors. To enhance adaptability, recent contributions have tailored force priors to specific applications (Li and Lu, 2018; Aucejo and De Smet, 2019; Feng et al., 2021). However, inferring high-dimensional posterior force distributions often demands computationally intensive sampling techniques, such as Hamiltonian Monte Carlo, due to complex statistical dependencies.

Beyond probabilistic frameworks, set-theoretic methods provide an alternative paradigm by modeling uncertainties—particularly epistemic ones, as unknown-but-bounded parameters (Yang, 2024; Liu et al., 2023). These methods propagate uncertainty intervals through system models, avoiding distributional assumptions. The computational simplicity and interpretability of interval-based analysis make it appealing for scenarios where probabilistic data is scarce or assumptions are untenable (Wang et al., 2023).

To effectively address the requirements of practical engineering applications, force reconstruction methods must undergo further enhancements in two critical areas, building upon the advancements previously discussed. First, there is a need to achieve higher reconstruction accuracy. This involves refining existing algorithms and developing more sophisticated techniques to ensure that the forces being measured and reconstructed closely aligned with the actual forces present in the system. Accurate force reconstruction ensures that engineers can make informed decisions, predict system behaviors reliably, and maintain the safety and integrity of structures and machinery.

Second, it is essential to reduce the computational cost associated with uncertainty quantification. Uncertainty quantification is a vital process that assesses the reliability and confidence in the reconstructed force measurements by accounting for various sources of error and variability inherent in engineering systems. However, current methods for uncertainty quantification can be computationally intensive, requiring significant processing power and time, which may not be feasible for real-time or large-scale applications. By developing more efficient algorithms and leveraging advancements in computational technologies, it is possible to lower the computational burden. This reduction enables quicker assessments of uncertainty, facilitating faster decision-making and allowing force reconstruction methods to be more widely applied in dynamic and resource-constrained engineering environments.

2.1.3. Advancements of state estimation problems and challenges

The accurate state information of structural systems is crucial for various structural control and health monitoring applications, such as digital twin construction (Torzoni et al., 2024), virtual sensing technology (Song et al., 2022; Teymouri et al., 2022), active vibration control of structures (Moradi et al., 2021), and structural damage detection (Papadimitriou et al., 2011). Similar to the force reconstruction problems, direct measuring comprehensive measurements of state for real-world structures often presents significant challenges. Deploying a large number of transducers across numerous degrees of freedom (DOF) can be impractical due to high costs, extensive space requirements, and increased risk of damage. As a result, the problem of state estimation, inferring a structure's complete state from limited measurements based on the inverse problem model, has attracted considerable interest in the field of structural dynamics.

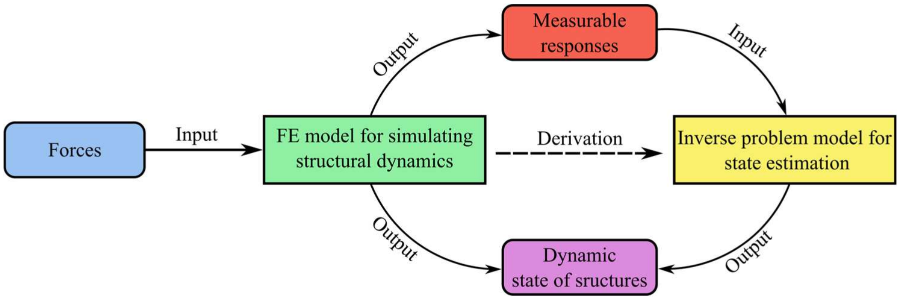


Figure 2.2. Demonstration of forward problem and inverse state estimation.

A simple and indirect approach for state estimation is based on load reconstruction models. As shown in Figure 2.1, the reconstructed forces from measured responses can serve as inputs to the FE model that simulates the dynamics of structures. This method is employed in various applications. For example, Liu and Li (2018) proposed a reconstruction strategy that utilizes a subspace identification algorithm and a fast iterative shrinkage-thresholding algorithm to reconstruct impact forces. The reconstructed forces are then input into a linear operator to compute the responses at unmeasured locations. This theory was applied to a simplified artillery testbed to verify the reconstruction strategy under three scenarios: single-source, two-source, and consecutive impact cases. Similarly, Li et al. (2021a) combined SVD with a pseudo-inverse operator to estimate the dynamic loads experienced by a continuous beam structure. These dynamic forces were subsequently input into an FE model to compute the global response of the continuous beam under these loads. The two-stage approach to state estimation allows for the direct and efficient utilization of existing FE models and force reconstruction models. However, error analysis becomes challenging due to the accumulation of measurement errors when calculating the complete state through multiple models.

In contrast, many studies focus on directly constructing inverse problem models for state estimation, with the model construction methods and application patterns illustrated in Figure 2.2. Due to the relationship between external loads and the state, many methods support the simultaneous calculation of forces and state based on measurements. In this thesis, to simplify the problem, we separate these two functions and focus on studying them individually. Regarding the state estimation problem, most prior studies on state estimation have primarily utilized various models derived from the principle of Kalman filtering. By incorporating the equations of motion of the FE model and embedding priors on external unknown forces and measurements, the constructed Kalman filter model can recursively output the complete system state using noisy measurements of system response. Early work focused on simplified linear structural systems. Papadimitriou et al. (2011) employed the standard Kalman filter, which uses accelerations from sparsely distributed accelerometers to estimate the full state of the system, enabling the estimation of strain at locations where no measurements were taken. Lourens et al. (2012) developed the augmented Kalman filter (AKF) for joint input-state estimation, where unknown forces are incorporated into the state vector. This approach allows for more effective prior assumptions regarding stochastic loads, measurement, and modeling errors compared to the standard Kalman filter solution. Consequently, the AKF can estimate the augmented vector, including both forces and states. Furthermore, Eftekhar Azam et al. (2015) introduced a dual implementation of the Kalman filter algorithm, known as the dual Kalman filter (DKF), for full-state estimation. The DKF employs a sequential structure of two Kalman filters, which circumvents the numerical issues related to unobservability and rank deficiency encountered in the AKF. As a result, the DKF yields more precise state estimation results. Recent efforts have focused on developing state estimation algorithms for nonlinear structures, such as the extended Kalman filter, unscented Kalman filter, and ensemble Kalman filter approaches, as well as various variants of these filters dedicated to the corresponding nonlinear structures (Erazo and Hernandez, 2015; Roohi et al., 2021; Paul et al., 2022; Impraimakis and Smyth, 2022).

In recent years, machine learning models have been increasingly explored and applied to state estimation problems, complementing the mainstream use of Kalman filters. Both machine learning models and Kalman filter-based models share the similarity of predicting the state at a specific time point based on the information collected over a surrounding time window (Cunha et al., 2023). For some online state estimation models, the information used only includes data collected at the current time point as well as data from a preceding time window for prediction. The primary difference lies in their development approaches: while Kalman filter-based models derive their governing equations from FE models, machine learning models are developed through the training and optimization of parameters. The training and test datasets for machine learning models are generated from the responses of an FE model subjected to numerous, and if necessary, synthetic excitations.

Advancements in machine learning technology have resulted in a greater variety of available models compared to Kalman filter models, offering more flexibility in their construction. This flexibility is particularly evident when the FE model cannot be explicitly formulated (Lv et al., 2007) or when the FE model includes hybrid components functioning as black boxes (Li et al., 2024a; Li et al., 2024b). In such cases, Kalman filter methods become inapplicable unless modifications are made to the FE model, whereas machine learning models remain valid. Furthermore, the success of constructing machine learning models depends on the loss values of the training and test datasets. When the training process converges to a relatively low test loss, the predictive performance of the machine learning-based model is considered sufficient for subsequent state estimation tasks. Consequently, with properly prepared training and test datasets, machine learning-based state estimation frameworks can handle both linear and nonlinear structural systems with comparable complexity. In contrast, state estimation for nonlinear structures using Kalman filter methods requires more complex designs, and many existing Kalman filter-based models cannot be directly applied to nonlinear structural systems.

Predicting the state of structures using data collected over a surrounding time window is effectively abstracted as a high-dimensional mapping problem. Numerous machine learning models have been employed to learn this high-dimensional mapping. To date, the vast majority of these models adopt neural network architectures within deep learning frameworks, taking into account the potential curse of dimensionality inherent in high-dimensional mappings. With advancements in technology, many studies have continuously innovated across various neural network architectures. For example, Yang and Lee (1997) demonstrated the effectiveness of a multi-layer perceptron (MLP), where a four-step delay MLP with a single hidden layer yielded accurate estimates of composite beam tip displacement and velocity within a 5% error margin. Wu and Jahanshahi (2018) presented a deep convolutional neural network (CNN) approach for state estimation of a linear single-DOF system, a nonlinear single-DOF system, and a full-scale three-story multi-DOF steel frame. The results, compared with those of a MLP, indicated that the proposed CNN approach is more robust against noise-contaminated data. Kumar et al. (2022) utilized an autoencoder coupled with long short-term memory (LSTM) neural networks to construct an inverse problem model for state estimation. The introduction of the autoencoder effectively learned the nonlinear manifold, while the LSTM neural networks leveraged the temporal dependencies inherent in the state data. Zhou et al. (2024) developed an LSTM model to estimate the dynamic displacement of a 420-meter-high building during super typhoon Mangkhut. The estimated displacements exhibited a high level of agreement with those measured by GPS, highlighting the reliability and accuracy of the developed LSTM model.

It is crucial to recognize that, despite the continuous advancements in machine learning technologies and the introduction of numerous sophisticated neural network architectures, the improvement in accuracy for state identification problems achieved by merely adopting more advanced neural networks is exceedingly limited. In many instances, the training of complex machine learning models becomes increasingly intricate, and the computational demands of forward propagation escalate, rendering many of these advanced algorithms impractical for real-world applications. To better address the needs of practical engineering, this thesis primarily focuses on the following two issues as entry points. First, the reliability of computational results derived from machine learning-based state estimation models needs to be quantified. Machine learning models are trained on datasets, often high-dimensional, and data collected from real structures are processed as inputs to these models. If the input data distribution deviates from that of the original training set, the model's predictions should inherently reflect higher uncertainty. This quantification of uncertainty provides a foundation for subsequent evaluation and decision-making processes, ensuring that the state estimation remains robust even when faced with unforeseen data variations.

Second, and most importantly, existing research typically operates under the assumption that the FE model established based on physical principles and its surrounding environment differs minimally from actual conditions. As illustrated in Figure 2.2, inverse problem models are derived from FE models. Whether these models are Kalman filter-based or machine learning-based, inaccuracies inherent in the FE model itself will lead to imprecise state estimations. This discrepancy frequently arises in real-world applications, where actual structural systems are often highly complex. Environmental influences and interactions can result in significant differences between the idealized FE model and the real structure. Even with advancements in simulation techniques and model updating approaches, acquiring an FE model that closely matches the real structure incurs substantial costs. Therefore, there is a pressing need for a methodology that ensures accurate state estimation even when the FE model is not sufficiently precise.

2.2. Probabilistic machine learning

2.2.1. Nonparametric Bayesian machine learning

Nonparametric Bayesian machine learning is a widely recognized approach in probabilistic machine learning and is frequently employed to address various uncertainty quantification challenges within the field of structural engineering (Wan et al., 2014; Teimouri et al., 2017; Tognan et al., 2022). To optimize the uncertainty quantification problems associated with force reconstruction and state estimation discussed earlier, several methods grounded in existing nonparametric Bayesian frameworks are proposed in this thesis. Consequently, this section primarily introduces the fundamental theories and applications of nonparametric Bayesian methods and evaluates the advantages and limitations of this approach in solving inverse problems in structural engineering.

Nonparametric Bayesian machine learning is a counterpart to parametric Bayesian machine learning. As widely recognized, Bayes' theorem provides a mathematical framework for inverting conditional probabilities, thereby enabling the determination of the probability of a cause given its observed effect. In machine learning models, predictions are governed by parameters; different parameter sets yield varying model predictions. Consequently, the model's prediction can be viewed as the effect generated by the underlying parameters within the machine learning model. In parametric Bayesian machine learning models, the distributions of these parameters are quantified given the observations, allowing for probabilistic interpretations and uncertainty quantification of the model parameters.

In contrast, nonparametric Bayesian machine learning does not assume a fixed number of parameters. Instead, it allows for an infinite-dimensional parameter space, providing greater flexibility in modeling complex data structures. This flexibility is particularly advantageous in inverse problems, where the relationship between observations and underlying states may be highly intricate and not well-captured by a finite set of parameters.

In this thesis, the GPR is used as a foundational tool in nonparametric Bayesian machine learning (Rasmussen and Williams, 2006). In probability theory, a GP is a stochastic process (a collection of random variables indexed by time or space), for which any finite subsets follow a multivariate Gaussian distribution. In essence, GP is a generalization of multivariate Gaussian distributions to infinite-dimensional spaces.

The distribution of a GP is the joint distribution of all the random variables, and thus the GP is a distribution over functions with a continuous domain.

A GP can be fully characterized by two functions: the mean function and the covariance function, analogous to a multivariate Gaussian distribution that is specified by a mean vector and a covariance matrix. The mean function provides the expected value of the GP at any given spatial or temporal indices. The covariance function, which defines the correlation between any two variables drawn from the GP, is typically represented by a positive definite kernel function whose inputs are two indices within the domain. The reason for adopting kernel functions is that they offer a flexible and mathematically robust framework for specifying the covariance structure of the GP. Kernel functions ensure that the resulting covariance matrix is positive definite, which is essential for the validity and stability of the GP. Moreover, kernel functions allow for the incorporation of prior knowledge about the properties of the underlying function, such as smoothness, periodicity, or linearity. This flexibility enables the modeling of complex relationships and dependencies within the data, facilitating more accurate and meaningful predictions.

A GP f over a domain χ is denoted as $f(x) \sim \mathcal{GP}(m(x), k(x, x'))$, where $x \in \chi$, $m: \chi \to \mathbb{R}$ is the mean function, and $k: \chi \times \chi \to \mathbb{R}$ is the covariance function. By modifying these two functions, the properties of the GP change accordingly. For example, let $m(x)$ be defined either as the zero function or as a polynomial function $m(x)=4x^2-2x+1$, and let the kernel function be the commonly used squared exponential kernel:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right), \tag{2.1}$$

where σ_f^2 and ℓ are hyperparameters of the covariance function, known as the output variance and length scale, respectively. As shown in Figure 2.3, realizations of GP with varying mean functions and hyperparameters for the squared exponential kernel exhibit distinct patterns. In this context, realizations refer to samples drawn from the specified GP, with each red solid line in the figure representing an individual sample. The two upper subfigures demonstrate that the length scale parameter ℓ influences the distance over which the sampled function values are strongly correlated. The two left subfigures illustrate that the output covariance σ_f^2 determines the range of variation in the samples. Additionally, the two diagonal subfigures show that the mean function $m(x)$ affects the overall trend of the sampled functions.

The demonstration illustrates that by selecting appropriate mean and covariance functions, a GP can effectively characterize functions in a probabilistic manner. Consequently, GP are employed as Bayesian priors over functions. According to Bayes' theorem, the posterior distribution of function values can be computed given observations of the functions, thereby enabling the learning of functions from data—this framework is known as GPR. Throughout this process, there is no need to define a parametrized model; instead, only the mean and covariance functions for the target are specified. This characteristic is why GPR is referred to as a nonparametric Bayesian method.

In many practices of GPR, determining the mean function $m(x)$ is of less concern compared to the covariance function $k(x, x')$. Specifically, a zero-mean function is often employed for convenience without loss of generality. This is because any non-zero mean can be incorporated into the GP framework by appropriately adjusting the covariance function or by pre-processing the data to center it around zero.

Consequently, the GP can focus on modeling the deviations from the mean, simplifying the analysis without compromising the flexibility or expressiveness of the model.

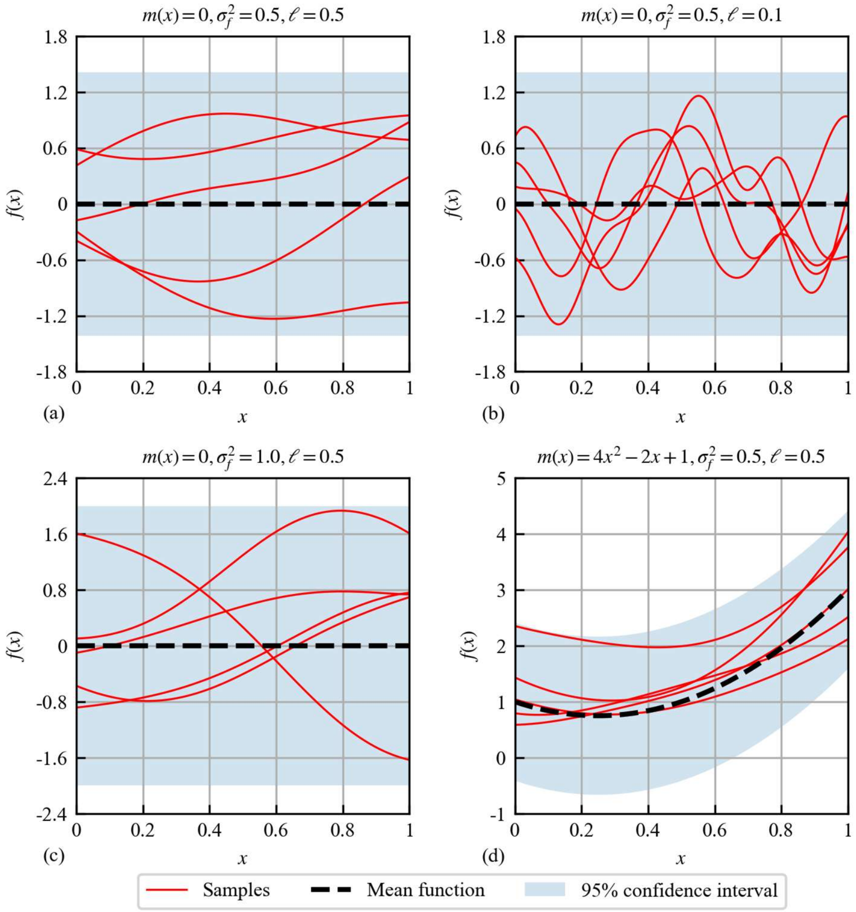


Figure 2.3. Realizations of GP with specified mean functions and hyperparameters for the squared exponential kernel.

To illustrate how GPR operates, consider a training dataset $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{n}$ where $y_i = f(x_i) + \epsilon_i$. Here, $f(x_i)$ denotes the noise-free observation over the function f at the input point x_i, and ϵ_i represents independent Gaussian noise with variance σ_1^2. The bold character $\boldsymbol{y}$ is used to denote the vector containing all y_i. By assuming the function $f(x)$ is a GP, GPR enables the computation of the conditional probability of any function value $f_* = f(x_*)$ at any new input point $x_* \in \chi$, leveraging the property that any finite set of function values follows a multivariate Gaussian distribution. The mean vector and covariance matrix of this distribution are derived from the GP's mean and covariance functions. Consequently, the posterior probability $p(f_*|\boldsymbol{y})$ is obtained via Bayes' theorem and is expressed as:

$$p(f_*|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|f_*)p(f_*)}{p(\boldsymbol{y})}, \tag{2.2}$$

which ultimately allows GPR to update beliefs about the function f in light of observed data, providing a principled approach to regression that quantifies uncertainty in the predictions.

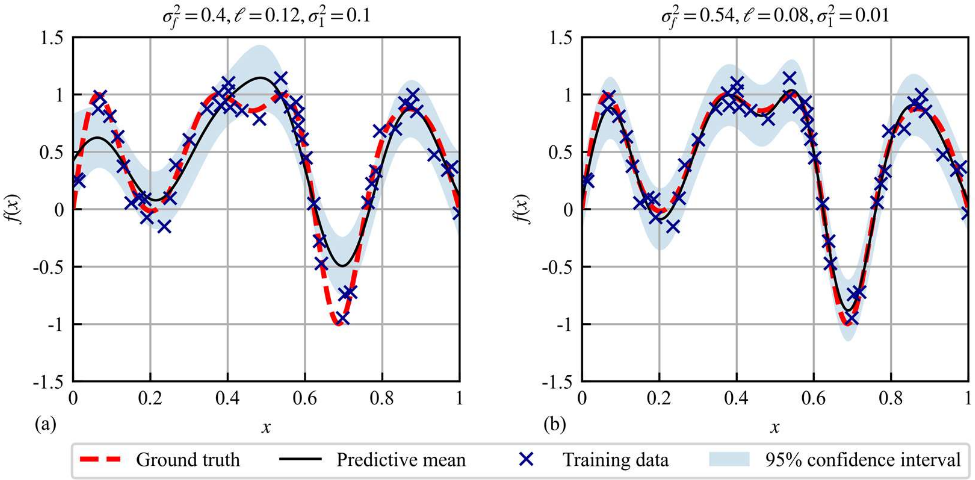


Figure 2.4. GPR predicted posterior with randomly initialized hyperparameters and optimized hyperparameters.

Figure 2.4 illustrates two scenarios in which the aforementioned procedures are employed to predict the function $f(x) = \sin(2\pi x + 2\sin(3\pi x))$ given 50 label pairs contaminated with noise. Under identical dataset conditions, the posterior distributions of the function predictions in these two cases are indeed different. This discrepancy arises from the differing priors selected for Bayesian inference as outlined in Eq. (2.2). In both scenarios, the GP utilizes a zero-mean function; however, the hyperparameters of the covariance function and those estimating the noise measurement vary between the two cases. Consequently, the prediction in Figure 2.4(b) outperforms that in Figure 2.4(a) significantly. This outcome underscores the critical importance of selecting an appropriate GP prior for the successful application of GPR.

Several methods are available for estimating optimal GP priors for effective GPR (Karvonen and Oates, 2023), including maximum likelihood estimation, maximum a posteriori, cross-validation, Bayesian inference, and kernel flows. Among these, maximum likelihood estimation is the most widely adopted, mostly because of the straightforward analytical formulation of the optimization problem, $\max p(\boldsymbol{y})$. Currently, most GPR software, including Rasmussen and Nickisch (2010), GPy (since 2012), Matthews et al. (2017), and Gardner et al. (2018), default to maximum likelihood estimation for optimizing GP priors. The methods proposed in this thesis would also primarily utilize maximum likelihood estimation for optimizing GP priors.

2.2.2. Nonparametric Bayesian machine learning applications in structural engineering

GPR for probabilistic modeling

The GPR grounded in Bayes' theorem enables systematic probabilistic modeling based on labeled data pairs. The posterior distributions generated by these probabilistic models provide high-quality information assurance for subsequent analysis and decision-making processes. In structural engineering, many applications demand models that deliver probabilistic analysis results, making nonparametric Bayesian machine learning particularly prevalent. For instance, Fricker et al. (2011) developed a

GP-based emulator for computationally intensive FE models. This emulator not only facilitates rapid predictions of FE model outputs but also accounts for the additional uncertainty arising from limited model evaluations. Their study demonstrated that a multivariate GP emulator outperforms independent univariate GP emulators in quantifying uncertainty across multiple outputs by leveraging dependencies between outputs, thereby significantly reducing uncertainty propagated through subsequent analyses.

Wan and Ren (2015) introduced a residual-based GP framework for updating FE models. This framework characterizes the relationship between residuals, the discrepancies between model predictions and experimental measurements, and selected model parameters. By implementing variance-based global sensitivity analysis within this framework, they quantitatively assessed and selected influential parameters, enhancing the efficiency of the parameter selection process. Additionally, the analytical derivation of gradients and Hessians from the residual-based GP accelerated the optimization process, enabling rapid convergence to optimal parameter values.

Aye and Heyns (2017) applied GPR to predict the remaining useful life of slow-speed bearings using a novel degradation assessment index derived from acoustic emission signals. They meticulously designed the mean and covariance functions by integrating multiple individual mean and covariance functions, resulting in an integrated GPR model. Their approach allowed for the simultaneous estimation of a composite mean function and a composite covariance function, leading to more flexible and accurate predictions of bearing lifespan.

Chakraborty and Adhikari (2021) explored the probabilistic modeling of the time evolution of structural system parameters using GPR and an advanced “mixture experts of GP” approach. This method effectively handled temporal multi-scale data, and their results demonstrated the ability to accurately predict the probability distributions of mass and stiffness parameters in future time steps.

Wang et al. (2022) employed heteroscedastic GP for probabilistic modeling of strains in a large-scale suspension bridge during typhoons. Their study successfully estimated the noise level changes induced by typhoons, providing a foundational basis for assessing the structural health of large-scale bridges under extreme weather conditions.

Most recently, Xian and Wang (2024) utilized GPR to enhance surrogate modeling for high-dimensional rare event simulations. They adopted a heteroscedastic GP to model the error correction between low-fidelity physics-based surrogates and original high-fidelity models, effectively capturing input-dependent uncertainties. Furthermore, they integrated uncertainty estimates within an active learning framework to iteratively refine the training dataset, ensuring that the surrogate model maintained high correlation and low bias in critical regions relevant to rare events.

Multi-output GP-based modeling

In addition to directly performing probabilistic regression using the GPR framework, numerous studies leverage the inherent properties of GP and Bayes’ theorem to explore the interrelationships among multiple GP, thereby enabling more sophisticated functionalities. These models are collectively referred to as multi-output GP.

Multi-output GP have been widely adopted in multi-fidelity modeling, where the framework enables probabilistic estimation by leveraging datasets of varying fidelities. In many studies, target functions at different fidelity levels are modeled as GP, and by assuming functional relationships between these GP, they can be effectively correlated

(Le Gratiet and Garnier, 2014; Perdikaris et al., 2017; Cutajar et al., 2019). Based on the multi-output GP framework, information from low-fidelity data can inform and enhance the predictions of high-fidelity models, thereby improving overall accuracy and reducing computational costs. In multiple applications of structural engineering, variations in fidelity arise from different measurement techniques applied to real structures. Consequently, GP-based multi-fidelity modeling has seen extensive application in these areas. For instance, Jin et al. (2021) addressed the trade-off between accuracy and spatial resolution in strain measurements by combining point strain sensors, which measure accurate strains (high accuracy) at discrete positions (low spatial resolution), with distributed strain sensors that provide quasi-continuous distributed measurements (high spatial resolution) but with less accurate strains (low accuracy). An autoregressive dual-fidelity GP was employed in the study to achieve accurate strain distribution, effectively combining high accuracy with high spatial resolution. The experiments demonstrated the feasibility of multi-fidelity data fusion, and the approach has the potential to monitor strain across entire structural systems within limited budgets.

Guo et al. (2021) proposed an MFGP approach to identify flame frequency responses. Currently, two methods are commonly used in the community: harmonic excitation, which provides accurate FFR estimates even in the presence of significant noise but only at discrete frequencies, and broadband excitation, which offers complete flame frequency responses over the frequency range of interest but may introduce higher levels of uncertainty in the results. The proposed multi-fidelity method leverages the strengths of both approaches by merging flame frequency responses identification results from short-time broadband excitation (low-fidelity) with harmonic excitations at selected frequencies (high-fidelity). This integration avoids the weaknesses of each individual method, enhancing the overall accuracy and reliability of flame frequency responses identification.

Additionally, due to differences in precision between FE models and real structures, as well as discrepancies in the volume of data generated from FE models compared to experimental measurements, many studies have introduced multi-fidelity modeling to address these challenges. Li and Jia (2020a) proposed an MFGP approach to integrate high-fidelity experimental data with low-fidelity analytical and numerical modeling data. This approach was successfully applied to predict the deformation capacity of reinforced concrete, demonstrating the effectiveness of MFGP models in enhancing prediction accuracy while managing computational resources efficiently.

GP have been also employed to solve various linear and nonlinear ordinary and partial differential equations (Särkkä, 2011; Raissi et al., 2017; Raissi et al., 2018; Chen et al., 2021a). Essentially, differential equations define relationships that implicitly embed correlations between multiple functions. By defining the involved functions as GP a priori, the correlations among these GP can be derived from the differential equations and subsequently reflected in the mean and covariance functions of all GP. When solving a specific differential equation, some GP are fully observed through specified initial and boundary conditions, while others have no direct data. The solution process is framed within a Bayesian formulation analogous to Eq. (2.2), where the estimation of GP without measurements is informed by the GP with complete observations. The posterior distributions obtained through this process effectively represent the solution to the differential equation. This approach allows for a coherent integration of prior knowledge and observed data, ensuring that the correlations and uncertainties inherent in the system are accurately captured and propagated through the solution. Tondo et al. (2023) applied a similar approach to beam deflection prediction problems by leveraging the Timoshenko beam theory. The theory facilitates the formulation of differential equations governing deflections, rotations, strains, bending moments, shear forces, and applied loads. Using GP framework, unobserved responses to applied loads are successfully predicted.

In addition to the aforementioned representative applications of multi-output GP, any scenario that allows for an analytically expressible functional relationship can leverage GP and Bayes' theorem for probabilistic analysis. In many studies, the physical quantities being analyzed do not exhibit clear physical correlations but do show statistical dependencies. By constructing parameterized functional relationships, the reliability of predictions can be enhanced, as the parameters within these functions can be directly integrated into maximum likelihood estimation to find optimal values. For example, Wan and Ni (2018) employed this approach to perform multi-output modeling of temperature and acceleration data in super-tall building structures. Their constructed model utilized partial sensor information to calculate measurements at other locations, thereby enabling data reconstruction for damaged sensors. This method effectively enhances the robustness and reliability of SHM by compensating for sensor failures. Similar methodologies have also been applied in dam SHM systems, where statistical correlations between different measured quantities are exploited to maintain accurate and reliable monitoring despite potential sensor damages or data gaps (Li et al., 2021b).

2.2.3. Advantages and limitations of nonparametric Bayesian machine learning

Nonparametric Bayesian machine learning offers significant advantages for addressing inverse problems in structural engineering, yet it also presents certain limitations that must be carefully considered. A foremost benefit of GPR is its intrinsic ability to quantify uncertainty, delivering not only point estimates but also confidence intervals for predictions. This probabilistic nature is essential for decision-making processes related to structural safety and maintenance, as it allows engineers to comprehensively assess the reliability of reconstructed forces or estimated states. Additionally, GPR exhibits remarkable flexibility in modeling complex, nonlinear relationships without presupposing a fixed functional form. This adaptability is especially valuable in structural dynamics, where the interactions between structural responses and underlying forces or states are often intricate and nonlinear. Furthermore, nonparametric Bayesian methods adeptly incorporate prior knowledge through the specification of mean and covariance functions. By integrating established physical principles or empirical observations, GPR enhances the accuracy and reliability of inverse problem solutions. The smoothness and continuity imposed by commonly used covariance functions, such as the squared exponential kernel, ensure that the estimated quantities exhibit physically plausible behavior, which is crucial for accurately capturing the dynamic responses of structures under varying loads and conditions. Moreover, multi-output GP approaches facilitate the integration of data from diverse sources with varying accuracies and resolutions, thereby improving overall model performance while efficiently managing computational resources.

However, a primary limitation of GPR is its computational complexity, particularly as the number of data points increases. The inversion of large covariance matrices scales cubically with dataset size, rendering it computationally prohibitive for large-scale structural monitoring systems equipped with extensive sensor networks. Although sparse approximations and inducing point methods can alleviate some computational burdens, they often introduce additional approximations that may compromise model accuracy. Additionally, nonparametric Bayesian methods can struggle with high-dimensional input and output spaces, a common scenario in complex structural systems with numerous degrees of freedom and multiple sensor inputs. While multi-output GP can handle correlated outputs, the increasing dimensionality exacerbates the "curse of dimensionality", hindering the effective modeling of dependencies and correlations and potentially degrading performance in state estimation and force reconstruction.

2.3. Deep learning

2.3.1. Deep learning

Building upon the challenges of the inverse problem discussed in Section 2.1, some investigations in the thesis employ deep learning techniques to address the state estimation problem. Therefore, this subsection provides a brief review of deep learning. Over the past decade, deep learning research has experienced explosive growth, underpinned by profound theoretical advancements and a wide range of applications. This subsection adopts a relatively broad perspective, presenting an overview of some theories and applications of deep learning in the fields of structural engineering from a practitioner’s standpoint. The aim is to demonstrate the necessity of deep learning for the problems investigated in this thesis through this review.

Deep learning is a specialized branch of machine learning that focuses on algorithms inspired by the architecture and function of the human brain, commonly implemented through deep neural networks (Goodfellow et al., 2016). Mathematically, a fundamental deep neural network consists of multiple layers of interconnected nodes or neurons, where each connection is associated with a weight parameter. These networks are typically organized into an input layer, several hidden layers, and an output layer. Each neuron in a layer performs a linear transformation of its inputs, followed by a nonlinear activation function, which introduces the necessary nonlinearity to model complex relationships. Formally, for ith given layer, the transformation can be expressed as:

$$\boldsymbol{x}^{(i)} = \sigma\left(\mathbf{W}^{(i)}\boldsymbol{x}^{(i-1)} + \boldsymbol{b}^{(i)}\right), \tag{2.3}$$

where $\boldsymbol{x}^{(i-1)}$ is the activation from the previous layer; $\mathbf{W}^{(i)}$ and $\boldsymbol{b}^{(i)}$ are weight matrix and bias vectors for the ith layer, respectively, containing the parameters to be optimized; $\sigma(\cdot)$ is the activation function, e.g., ReLU, tanh, or sigmoid. With the forward propagation from layers to layers, the mappings from input domain to the output domain are implemented consequently. With multiple label pairs to compose the training, validation and test dataset, the training of deep neural networks could be conducted, which involves optimizing the weights and biases to minimize a loss function, typically using the gradient-based optimizers, e.g., stochastic gradient descent, Adam, etc.

Figure 2.5(a) illustrates the fundamental deep neural network, commonly known as the MLP. In the early stages of deep learning research, this basic MLP was frequently employed as a universal approximator to model a wide variety of functions, given sufficient depth and complexity (Hornik et al., 1989). In recent years, numerous studies on physics-informed neural networks have leveraged this characteristic of MLP to solve partial differential equations (Raissi et al., 2019). By incorporating multiple loss functions derived from governing equations, boundary conditions, and measurement data, the MLP is continuously optimized to accurately estimate the solutions of these partial differential equations.

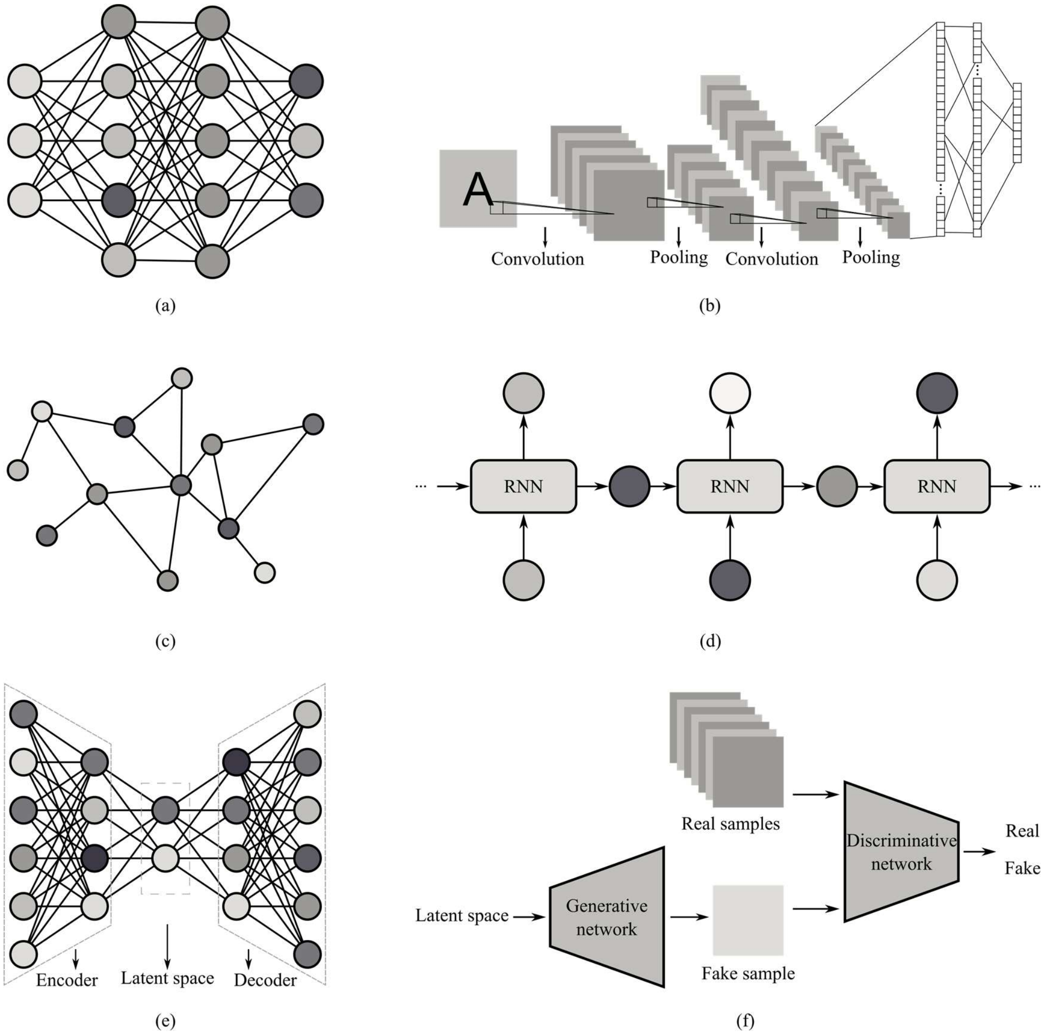


Figure 2.5. Neural network architectures for different applications.

Furthermore, the widespread application of deep learning is largely attributable to various specialized neural network architectures that have evolved from the MLP, as depicted in Figures 2.5(b)-(f). These include CNN, graph neural networks, RNN, autoencoders, and generative adversarial networks. Each of these architectures is tailored to address specific problems that traditional machine learning methods have been unable to solve effectively. For example: CNN are extensively used in applications requiring image and signal recognition, such as human face recognition (Mehdipour Ghazi and Kemal Ekenel, 2016) and autonomous driving (Li et al., 2019). Graph neural network are suitable for tasks involving spatial or temporal topological relationships, such as predicting properties based on protein structures (Jiang et al., 2021). RNN excel in handling sequence-based problems (Van Houdt et al., 2020), including language translation and time-series forecasting. Autoencoders are effective for non-linear compression of high-dimensional data into lower-dimensional representations (Tschannen et al., 2018). Generative adversarial networks are employed in tasks related to image and signal generation, enabling the creation of realistic synthetic data (Creswell et al., 2018). In many contemporary artificial intelligence applications, the strengths of these advanced neural network architectures are often integrated to collaboratively achieve sophisticated intelligent objectives. This integration allows for more robust and versatile solutions, harnessing the unique capabilities of each network type to address complex, multifaceted problems across various domains.

It is important to note that, compared to traditional machine learning models, the advanced neural network architectures offer significant advantages in addressing high-dimensional, multi-input, and multi-output problems. Traditional approaches often encounter challenges such as the curse of dimensionality, which can hinder their performance and scalability when dealing with complex datasets. In contrast, deep learning architectures, with their multiple layers and vast number of parameters, are inherently designed to capture and model intricate patterns and relationships within high-dimensional data. Moreover, deep learning models excel in handling scenarios with multiple inputs and outputs, enabling simultaneous processing and prediction of diverse variables. This capability is particularly beneficial for the problems investigated in this thesis, where the complex structure systems typically involve numerous interacting components and a multitude of measurements. By leveraging the strengths of deep learning, the inverse problems could be abstracted to a multidimensional problem that is beyond the reach of traditional machine learning techniques.

Another pivotal framework within deep learning is transfer learning, which enables the leveraging of pre-trained models to enhance performance on new, related tasks. Transfer learning capitalizes on the knowledge acquired from large-scale datasets and complex models trained in extensive computational resources, allowing this expertise to be applied to different but related problems with limited data (Weiss et al., 2016; Zhuang et al., 2020). This approach is particularly advantageous in scenarios where acquiring labeled data is challenging or expensive, as it reduces the need for extensive training from scratch. By fine-tuning pre-trained networks, transfer learning not only accelerates the training process but also often results in superior performance compared to models trained solely on the target dataset. In the context of structural engineering, transfer learning can facilitate the adaptation of models trained on generic structural data to specific applications. This adaptability underscores the versatility and efficiency of deep learning frameworks, making transfer learning an invaluable tool for advancing research and practical implementations in complex, high-dimensional environments.

2.3.2. Deep learning applications in structural engineering

The integration of deep learning techniques into structural engineering has revolutionized the way engineers approach complex problems related to structural analysis, health monitoring, and predictive maintenance. By leveraging the powerful pattern recognition and data-driven capabilities of deep neural networks, structural engineers can achieve higher accuracy, efficiency, and reliability in various applications. In the initial stages, applications primarily relied on MLP.

Berke et al. (1992) applied MLP to capture structural design expertise, demonstrating the potential of neural networks as expert designers in aerospace structures. In this investigation, an advanced structural optimization code alongside the neural network code NETS are utilized to generate optimal designs for various aerospace components, including a trussed ring and two types of aircraft wings, under both static and dynamic constraints. By processing the optimum design data into input-output pairs, they trained the MLP to learn the complex relationships between design parameters and optimal configurations. The trained network successfully predicted optimal designs for new scenarios with minimal computational effort, highlighting the efficiency of neural networks compared to traditional optimization methods.

Ni et al. (2006) presented an experimental investigation of seismic damage detection of a 38-story tall building using MLP. In this study, the principal component analysis-compressed FRF data were used as inputs to MLP for damage identification, including the damage evaluation and damage location identification. It was shown that the identification results by means of the FRF projections on a few principal components are much better than those directly using the measured FRF data for inputs of MLP, and the identification results agreed well with the visual inspection in the experiment.

Zhou et al. (2010) developed an MLP-based modal information prediction of large infrastructures using temperature sensing data. The temperature data at multiple locations were processed through principal component analysis, and the results at each sampling step will be used for input of the MLP. 770-hour modal frequency and temperature data were utilized to train, validate, and test several pre-designed MLP models with different hidden nodes. The results revealed strong alignment of predictions with respect to the natural frequencies from measurement.

Continuing the exploration of deep learning applications in structural engineering, subsequent advancements moved beyond the foundational MLP to incorporate more specialized neural network architectures. Atha and Jahanshahi (2018) adopted a CNN architecture for corrosion detection of structural metallic surfaces. The effect of different color spaces, sliding window sizes, and CNN architectures were discussed. To this end, the proposed CNN architecture was evaluated, and it was shown that CNN outperform state-of-the-art vision-based corrosion detection approaches that were developed based on texture and color analysis using a simple MLP.

Yu et al. (2019) proposed a bidirectional recurrent neural network (BiRNN) based autoencoder scheme to estimate the remaining useful life. A BiRNN based autoencoder was trained in the investigation to convert the multi-sensor readings collected from historical run-to-failure instances to low-dimensional embeddings, which were adopted as matrices to reflect various health degradation patterns of the instances. The BiRNN was demonstrated to successfully enrich the input information to enhance the effectiveness of predictions.

Zhou et al. (2019) developed a deep RNN specifically designed for impact load identification in nonlinear structural systems. The architecture of the deep RNN comprised two LSTM layers, one bidirectional LSTM layer, and an MLP. Utilizing a comprehensive dataset of dynamic response and impact load pairs, the deep RNN was trained to learn the intricate inverse mapping between the structural responses and the corresponding impact loads. To validate the effectiveness of their model, three distinct nonlinear structural systems were presented in the investigation: a damped Duffing oscillator, a nonlinear three-degree-of-freedom system, and a nonlinear composite plate. The results demonstrated that the deep RNN exhibited a strong capacity for accurately identifying impact loads, even in scenarios where the exact impact location remained unknown.

Qin et al. (2021) presented a hybrid deep neural network aimed at accurately predicting cutterhead torque in shield tunneling machines by utilizing operational and status parameters of the equipment. This hybrid model seamlessly integrates a 1D CNN, LSTM networks, and a MLP to extract both implicit and sequential features from the input data. By combining these architectures, the model effectively harnesses deep information, enabling precise torque prediction. Furthermore, to enhance prediction performance and address the vanishing gradient problem typically encountered in deep-layer network training, the authors incorporated a residual network module into the proposed neural network architecture. The effectiveness and superiority of the proposed method were validated using fifteen distinct datasets derived from actual project data, demonstrating its robust performance and practical applicability in real-world scenarios.

From the aforementioned application examples, it is evident that a diverse array of neural networks and their combinations can be employed to address practical problems in structural engineering. While these combinations often result in highly complex architectures that are challenging to theoretically justify, the paramount concern for practitioners is possessing a model that effectively resolves real-world problems. Although it may be difficult to conduct a detailed theoretical evaluation of these models, their performance on test datasets provides clear evidence of their applicability and effectiveness in practical scenarios. In simple terms, this philosophy can be encapsulated as: deep learning models are valuable because they work.

Additionally, transfer learning has become widely adopted in structural engineering. The primary motivation for utilizing transfer learning is the insufficient volume of data required to train complex neural networks for specific tasks. As a result, many studies endeavor to adapt pretrained neural networks to address these specialized applications. From a neuroscience perspective, this approach leverages existing extensive knowledge to learn a particular and specific task. The broad knowledge base can be considered to be low fidelity, characterized by abundant data, whereas the data for specific tasks are high-fidelity but often limited in quantity. Many applications in structural engineering encounter this data imbalance issue. According to research findings, numerous studies have effectively implemented transfer learning to successfully accomplish specific tasks within the field.

Chen et al. (2021b) introduced a deep transfer learning framework designed to progressively assess the structural integrity of rail systems by leveraging acoustic emission monitoring data combined with knowledge transferred from an acoustic database. Specifically, the approach utilizes the lower-level layers of a CNN that has been pre-trained on extensive audio datasets for effective feature extraction within the predictive model for rail systems. In contrast to traditional transfer learning methods, which typically incorporate knowledge from models pre-trained on standard image data, this novel approach is better suited for handling acoustic emission spectrograms.

Bao and Mahadevan (2022) adopted a CNN and transfer learning approach for effective structural internal damage detection, e.g., voids and cracks. Addressing the persistent challenge of insufficient training data, particularly prevalent in many SHM applications, the study utilized FE computer simulations to generate a substantial volume of training data encompassing various damage shapes and locations. These simulated datasets were integrated with pre-trained convolutional cores from an advanced computer vision-based deep CNN, thereby facilitating effective transfer learning. This hybrid approach enabled the CNN to automatically extract and learn relevant features for damage diagnosis, eliminating the need for manual feature engineering typically required in traditional image processing techniques.

In this study, transfer learning plays a crucial role due to the data-intensive nature of certain problems. A key detail to note is that different neural networks require distinct transfer learning strategies tailored to their specific tasks, and the chosen strategy is pivotal for the successful transfer of knowledge. Within the domain of structural engineering, the majority of existing research has focused on CNN-based transfer learning. These CNN-based approaches have been validated through numerous studies, demonstrating their efficacy as a versatile method. Conversely, transfer learning strategies for other types of neural networks cannot simply replicate the CNN methodology. Instead, they must be specifically designed to align with the unique architectures of the respective neural networks, ensuring effective knowledge transfer in structural engineering applications.

2.3.3. Advantages and limitations of deep learning

Deep learning has emerged as a transformative technology in structural engineering, offering numerous advantages that significantly enhance the ability to analyze and manage complex engineering problems. One of the most prominent strengths of deep learning lies in its capacity to handle and learn from multi-dimensional data. Structural engineering often involves intricate interactions between various parameters such as stress, strain, vibration, and material properties, all of which are inherently multi-dimensional. Deep neural networks, with their multiple layers and non-linear activation functions, can effectively model these complex relationships, capturing subtle patterns and dependencies that traditional analytical methods might overlook.

Additionally, deep learning frameworks offer a wide array of neural network architectures, including CNN, RNN, autoencoders, and transformer models. This diversity allows engineers to select and tailor specific architectures that best suit the nature of the problem at hand. For instance, CNN are highly effective for image-based tasks such as crack detection on concrete surfaces, while RNN and their variants like LSTM networks are suitable for analyzing time-series data from vibration sensors. Moreover, the ability to combine different architectures in hybrid models enhances the flexibility and adaptability of deep learning solutions, enabling the development of customized models that address unique structural engineering challenges. The transferability of deep learning models further amplifies their utility in this field. Existing research has demonstrated effective transfer learning strategies, where knowledge from models trained on related tasks or larger datasets is leveraged to improve performance on specialized structural engineering applications.

The advancements in deep learning infrastructure and computational power have also played a crucial role in its adoption within structural engineering. open-source libraries like PyTorch, TensorFlow, and JAX are continuously being developed, providing robust platforms that ease the development, training, and application of neural networks (Dixit et al., 2018). These tools abstract complex algorithms into user-friendly interfaces, making the deep learning implementation more accessible to researchers and engineers in structural engineering. Concurrently, hardware improvements, particularly GPU and TPU technology, are enhancing computational power for training deep neural networks (Wang et al., 2019). The progress not only accelerates the training process but also democratizes access to these advanced facilities. As a result, building and training neural networks, as well as fine-tuning them using the transfer learning, are becoming more convenient and effective.

Despite its powerful capabilities, deep learning also presents certain limitations that must be carefully considered. One major drawback is the substantial requirement for large volumes of data, making it most effective in scenarios where extensive datasets are available. This high data demand can be a barrier in applications where data collection is costly or impractical. Additionally, deep learning models typically involve a vast number of parameters, resulting in prolonged training times compared to traditional machine learning methods. Unlike the traditional machine learning approaches which have fewer parameters and can be trained relatively quickly, deep neural networks require significant computational resources and extended training periods to converge to optimal solutions. This complexity can be a drawback in scenarios where computational power is limited or rapid model development is required. Moreover, extensive training time may impede the iterative process of model refinement and deployment in dynamic engineering environments.

Another significant limitation is the necessity for experienced practitioners to effectively train deep learning models. Achieving high performance with deep learning requires a deep understanding of various neural network architectures, optimization algorithms, and regularization techniques (Glorot and Bengio, 2010; Srivastava et al., 2014; Abdulkadirov et al., 2023). Practitioners must adeptly tune hyperparameters and navigate common pitfalls such as overfitting and underfitting. The process of selecting appropriate architectures, adjusting learning rates, and configuring network layers is often empirical and relies heavily on the practitioner's intuition and experience. This requirement can be a substantial barrier for organizations lacking skilled personnel, limiting the widespread adoption of deep learning technologies in structural engineering fields. Additionally, the opaqueness of deep learning models, often referred to as the "black box" problem (Castelvecchi, 2016), can hinder their interpretability and trustworthiness in critical engineering applications where understanding the decision-making process is essential.

2.4. Summary

This chapter provides a comprehensive review of the fundamental concepts and methodologies that underpin the contributions of this thesis. It begins with an overview of general inverse problems, establishing the foundational framework for a detailed exploration of two specific inverse problems: force reconstruction and state estimation in structural dynamical systems. The discussion of these problems focuses on defining each one, summarizing the approaches and solutions documented in existing literature, and identifying the current challenges that remain unresolved. Subsequently, the chapter delves into two key methodologies utilized in this research: probabilistic machine learning and deep learning. For each methodology, the basic theoretical principles, historical development, and diverse applications are thoroughly examined. Additionally, the advantages and disadvantages of probabilistic machine learning and deep learning are critically analyzed to provide a balanced understanding of their respective strengths and limitations.

Building upon the challenges identified in the literature review, this thesis will address key issues in force reconstruction and state estimation using probabilistic machine learning and deep learning techniques. Specifically, an accurate and efficient probabilistic force reconstruction framework is developed by leveraging multi-task nonparametric Bayesian learning. This framework enables the estimation of forces from structural measurements while providing confidence intervals for force values at any time step. The theoretical foundation of this proposed method is detailed in Chapter 3, with its practical application to maglev dynamic system illustrated in Chapter 4. For the state estimation problem, this thesis contributes two distinct yet complementary approaches to overcome existing limitations: a Bayesian-based MFGP method and an RNN approach enhanced by transfer learning. The theoretical underpinnings of these methods are presented in Chapter 5, followed by their practical applications in Chapter 6. Together, these contributions demonstrate the effective integration of probabilistic machine learning and deep learning techniques to advance force reconstruction and state estimation in structural dynamical systems.

Chapter 3.

Nonparametric Bayesian multi-task learning for time-domain force reconstruction

3.1. Introduction

This chapter introduces the theory of time-domain force reconstruction for structural dynamic systems based on the proposed nonparametric Bayesian multi-task learning method. Generally, the reconstruction of dynamic forces is modelled via convolution between force and response (Sanchez and Benaroya, 2014). However, the deconvolution to compute the force often encounters ill-posed problems and multi-source uncertainties. To tackle these challenges, the proposed method assigns GP priors to the force functions using a nonparametric Bayesian approach. By exploiting the convolution operator between force and response, it is shown that the responses are also GP, of which covariance functions can be derived from those of the forces. To handle the correlated GP, the concept of multi-task learning is applied. This enables the construction of a joint Gaussian distribution among variables from multiple GP. Consequently, the measured responses and forces are jointly characterized by a multivariate Gaussian distribution, and the posterior distribution of the force given the measurement is derived in closed form. The performance of the proposed method is demonstrated through simulations on a truss bridge and experiments on a frame structure subjected to impact loads. The results demonstrate the high accuracy and efficient uncertainty quantification achieved by the nonparametric Bayesian multi-task learning method.

3.2. Problem description of time-domain force reconstruction

The equation of motion of an n DOF structural system subject to external excitations can be expressed as:

$$\mathbf{M}\ddot{\boldsymbol{x}}(t) + \mathbf{C}\dot{\boldsymbol{x}}(t) + \mathbf{K}\boldsymbol{x}(t) = \mathbf{L}\boldsymbol{f}(t), \tag{3.1}$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K} \in \mathbb{R}^{n\times n}$ represent the mass, damping, and stiffness matrices, respectively; $\ddot{\boldsymbol{x}}(t)$, $\dot{\boldsymbol{x}}(t)$ and $\boldsymbol{x}(t) \in \mathbb{R}^{n}$ are the acceleration, velocity and displacement vectors at timestamp t. $\mathbf{L} \in \mathbb{R}^{n\times n_f}$ is the excitation matrix that maps the input force vector $\boldsymbol{f}(t) \in \mathbb{R}^{n_f}$ at time t to the n dimensional force vector. n_f is the number of forces acting on the structure.

By defining a state vector $\widetilde{\boldsymbol{x}}(t) = [\boldsymbol{x}(t); \dot{\boldsymbol{x}}(t)]$, Eq. (3.1) can be converted to continuous-time state space representation:

$$\dot{\widetilde{\boldsymbol{x}}}(t) = \mathbf{A}_{\mathrm{c}}\widetilde{\boldsymbol{x}}(t) + \mathbf{B}_{\mathrm{c}}\boldsymbol{f}(t), \tag{3.2}$$

with

$$\mathbf{A}_{\mathrm{c}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix}, \text{ and } \mathbf{B}_{\mathbf{c}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}\mathbf{L} \end{bmatrix}, \tag{3.3}$$

where $\mathbf{I} \in \mathbb{R}^{n\times n}$ denotes the identity matrix. Given the assumption that accelerations are measured at n_s number of DOF of the structure, the n_s dimensional observation vector $\boldsymbol{y}(t)$ at time t is defined as:

$$\boldsymbol{y}(t) = \mathbf{J}\ddot{\boldsymbol{x}}(t), \tag{3.4}$$

where $\mathbf{J} \in \mathbb{R}^{n_s\times n}$ is a binary matrix associated with the sensor locations. Each row of $\mathbf{J}$ exactly contains one non-zero element (set to 1), whose index corresponds to the DOF where an accelerometer is installed, with all other elements being 0. With the substitution of Eq. (3.4) to the continuous-time state space representation, the observation vector can be obtained:

$$\boldsymbol{y}(t) = \mathbf{C}_{\mathrm{c}}\tilde{\boldsymbol{x}}(t) + \mathbf{D}_{\mathrm{c}}\boldsymbol{f}(t), \tag{3.5}$$

where

$$\mathbf{C}_{\mathrm{c}} = [-\mathbf{J}\mathbf{M}^{-1}\mathbf{K} \quad -\mathbf{J}\mathbf{M}^{-1}\mathbf{C}], \text{ and } \mathbf{D}_{\mathrm{c}} = \mathbf{J}\mathbf{M}^{-1}\mathbf{L}. \tag{3.6}$$

The observation equation articulates the functional relationship among the observations, system states, and forces. This relationship hinges on precise knowledge of the system matrices of the structural system, as well as the specific locations of applied forces and accelerometers. However, it is not applicable to directly using information from $\boldsymbol{y}(t)$ to compute representations of $\boldsymbol{f}(t)$, because the system state functions $\tilde{\boldsymbol{x}}(t)$ cannot be isolated based on the above-mentioned equations. To eliminate the system state in the formulation, assume the input signal remains constant in each sampling period of length Δt (zero-order hold assumption), and n_t time steps are involved. This allows expressing the system states, forces, and observations as:

$$\begin{gathered} \tilde{\boldsymbol{x}}(t) = \tilde{\boldsymbol{x}}(k\Delta t) = \tilde{\boldsymbol{x}}_k, \boldsymbol{f}(t) = \boldsymbol{f}(k\Delta t) = \boldsymbol{f}_k, \boldsymbol{y}(t) = \boldsymbol{y}(k\Delta t) = \boldsymbol{y}_k; \\ \forall t \in [k\Delta t, (k+1)\Delta t), k = 0, 1, 2, \dots n_t - 1, \end{gathered} \tag{3.7}$$

Then, the state space representation and observation equation can be transformed from continuous to discretized form:

$$\tilde{\boldsymbol{x}}_{k+1} = \mathbf{A}_d\tilde{\boldsymbol{x}}_k + \mathbf{B}_d\boldsymbol{f}_k, \tag{3.8}$$

$$\boldsymbol{y}_k = \mathbf{C}_d\boldsymbol{x}_k + \mathbf{D}_d\boldsymbol{f}_k, \tag{3.9}$$

where the matrices $\mathbf{A}_d, \mathbf{B}_d, \mathbf{C}_d$ and $\mathbf{D}_d$ are calculated by:

$$\mathbf{A}_d = \exp(\mathbf{A}_{\mathrm{c}}\Delta t), \mathbf{B}_d = (\mathbf{A} - \mathbf{I})\mathbf{A}_{\mathrm{c}}^{-1}\mathbf{B}_{\mathrm{c}}, \mathbf{C}_d = \mathbf{C}_{\mathrm{c}}, \text{ and } \mathbf{D}_d = \mathbf{D}_{\mathrm{c}}. \tag{3.10}$$

For zero initial conditions in a structural dynamic system, the observation at step k can be obtained via iterative substitutions of Eq. (3.8) to Eq. (3.9):

$$\boldsymbol{y}_k = \sum_{i=0}^{k-1} \mathbf{C}_d \mathbf{A}_d^{i-1} \mathbf{B}_d \boldsymbol{f}_{k-i} + \mathbf{D}_d \boldsymbol{f}_k. \tag{3.11}$$

Consequently, the relationship between full measured acceleration data and external forces could be formulated based on a discrete convolutional operator:

$$\begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{n_t-1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{H}_1 & \mathbf{H}_0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{H}_{n_t-1} & \mathbf{H}_{n_t-2} & \cdots & \mathbf{H}_0 \end{bmatrix} \begin{bmatrix} \boldsymbol{f}_0 \\ \boldsymbol{f}_1 \\ \vdots \\ \boldsymbol{f}_{n_t-1} \end{bmatrix}, \tag{3.12}$$

where on the left-hand side is the series of observed acceleration data, represented as $\boldsymbol{y} \in (\mathbb{R}^{n_s})^{n_t}$, and on right-hand side is the convolutional operator $\mathbf{H} \in \mathbb{R}^{n_s n_t \times n_f n_t}$ in Toeplitz matrix form, and the force vector $\boldsymbol{f} \in (\mathbb{R}^{n_f})^{n_t}$ that contains values of forces from time step 0 to $n_t - 1$. Here, $\mathbf{H}_i \in \mathbb{R}^{n_s \times n_f}$ represents the ith block of the convolutional operator, it is given by:

$$\mathbf{H}_0 = \mathbf{D}_d, \text{ and } \mathbf{H}_i = \mathbf{C}_d \mathbf{A}_d^{i-1} \mathbf{B}_d;\ i = 1, 2, 3, \dots n_t - 1. \tag{3.13}$$

Note that all $\mathbf{H}_i$ are intrinsically derived from system matrices, therefore, with the specification of the structural dynamical system, and indices of DOF correspond to input and output, the convolutional operator $\mathbf{H}$ is determined.

The time-domain force reconstruction is a typical deconvolution problem that necessitates the use of discretized measurements $\boldsymbol{y}$ to inversely compute results of $\boldsymbol{f}$. The convolutional operator $\mathbf{H}$ is notoriously ill-conditioned, making direct inversion impossible to be directly inverted (Kabanikhin, 2008). Traditionally, this inverse problem has been addressed using various regularization techniques, such as SVD and Tikhonov regularization (Liu et al., 2022). Nevertheless, the reconstruction of accuracy and stability remains a concern, and further improvements are needed. Furthermore, real-world measurements are inevitably uncertain, and it is essential to effectively propagate these uncertainties to the estimated forces (Zhang et al., 2012). Therefore, a method that not only solves the deconvolution problem but also provides uncertainty quantification is of great interest in this chapter.

3.3. Multi-task learning with nonparametric Bayesian approach for force reconstruction

Nonparametric Bayesian multi-task learning leverages GP with shared information to model the covariance structure across different but related tasks. This shared covariance structure allows the model to utilize strength from all available data, thereby enhancing prediction performance for each task through pooled information (Wan and Ni, 2018; Wan and Ni, 2019). In the specific application of force reconstruction, where direct force data may not be measurable, nonparametric Bayesian multi-task learning adapts the principles of GPR to address complex problems involving multiple related tasks simultaneously.

For a structural dynamic system subjected to external excitations, the applied force and response histories in the time domain are interdependent tasks. The relationships among tasks can be expressed based on the equations of motion of the FE model and eventually encapsulated in a convolutional mapping, as shown in Eq. (3.12). Figure 3.1 demonstrates the mechanism of nonparametric Bayesian multi-task learning, referred to here as multi-task Gaussian processes (MTGP), and compares it with traditional GPR. In a typical GPR setting, each task is modeled independently with a single GP, yielding a task-specific predictive posterior distribution. In contrast, the MTGP framework can handle tasks lacking direct measurements. By leveraging the convolutional relationships derived from the FE model, the posterior distributions of such tasks can be inferred by conditioning on data from related tasks.

To enable the multi-task learning of forces, the n_f forces are initially modeled as independent GP. Each force $f_i(t)$ is represented as $f_i(t) \sim \mathcal{GP}(0, k_i(t, t'))$ for $i = 1,2, \dots, n_f$. In this framework, any force value at a specific time is a Gaussian variable, and any finite set of force values follows a multivariate Gaussian distribution. According to Eq. (3.12), derived from the governing equations of the FE model, the relationship $\boldsymbol{y} = \mathbf{H}\boldsymbol{f}$ holds, where $\boldsymbol{y}$ represents the acceleration response data as linear combinations of these Gaussian variables. Consequently, any finite set of data drawn from the acceleration responses will also adhere to a multivariate Gaussian distribution. This indicates that the acceleration responses can also be modeled as GP, with covariance functions dependent on those of the forces.

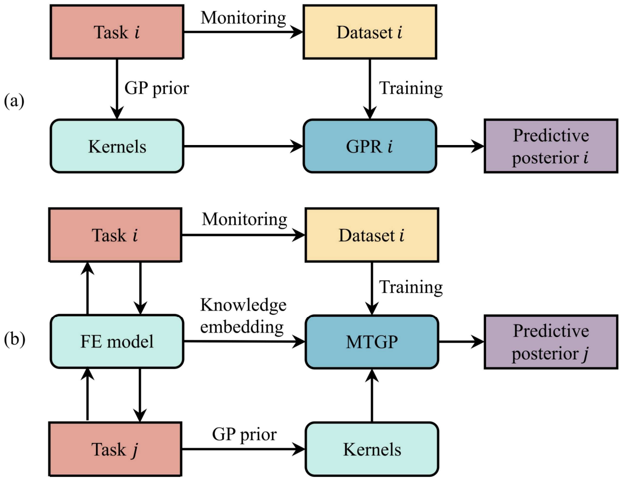


Figure 3.1. The working mechanism of MTGP for force reconstruction and comparison with GPR.

The primary goal of multi-task learning for force reconstruction is to leverage information from tasks with measurements (accelerations) to infer tasks lacking direct data (forces), that is, computing the posterior distribution $p(\boldsymbol{f}|\widehat{\boldsymbol{y}})$, where $\widehat{\boldsymbol{y}}$ represents the measured accelerations. When all tasks are GP with interdependent kernels, a multivariate Gaussian distribution involving any function values from force and acceleration histories can be constructed, as follows:

$$\begin{bmatrix} \boldsymbol{f} \\ \widehat{\boldsymbol{y}} \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{K}_{ff} & \mathbf{K}_{fy} \\ \mathbf{K}_{yf} & \mathbf{K}_{yy} \end{bmatrix}\right), \tag{3.14}$$

The mean vector of the multivariate distribution is $\mathbf{0}$ because the mean functions for the forces are zero functions, and it is assumed that measured acceleration data in $\widehat{\boldsymbol{y}}$ differs from the function values $\boldsymbol{y}$ by additive Gaussian noise $\boldsymbol{w}$, thus the mean vector of $\widehat{\boldsymbol{y}}$ is:

$$\mathbb{E}\{\widehat{\boldsymbol{y}}\} = \mathbb{E}\{\mathbf{H}\boldsymbol{f} + \boldsymbol{w}\} = \mathbf{H}\mathbb{E}\{\boldsymbol{f}\} + \mathbb{E}\{\boldsymbol{w}\} = \mathbf{0}. \tag{3.15}$$

As a critical feature of GP regression, the posterior distribution $p(\boldsymbol{f}|\widehat{\boldsymbol{y}})$ is also Gaussian and could be analytically formulated from the covariance matrices and $\widehat{\boldsymbol{y}}$. Since the forces are modeled as GP, $\mathbf{K}_{ff}$ could be directly obtained from the covariance kernel functions. The other covariance matrices could be derived from $\mathbf{K}_{ff}$ using the FE model that correlates different tasks:

$$\mathbf{K}_{fy} = \mathbf{K}_{yf}^{\mathrm{T}} = \mathbb{E}\{\boldsymbol{f}\widehat{\boldsymbol{y}}^{\mathrm{T}}\} = \mathbb{E}\{\boldsymbol{f}\boldsymbol{f}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \boldsymbol{f}\boldsymbol{w}\} = \mathbf{K}_{ff}\mathbf{H}^{\mathrm{T}}, \tag{3.16}$$

$$\mathbf{K}_{yy} = \mathbb{E}\{\mathbf{H}\boldsymbol{f}\boldsymbol{f}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \boldsymbol{w}\boldsymbol{w}^{\mathrm{T}} + \mathbf{H}\boldsymbol{f}\boldsymbol{w}^{\mathrm{T}} + \boldsymbol{w}\boldsymbol{f}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\} = \mathbf{H}\mathbf{K}_{ff}\mathbf{H}^{\mathrm{T}} + \boldsymbol{\Psi}, \tag{3.17}$$

where $\boldsymbol{\Psi}$ is an identical matrix representing the variance of noise, i.e., $\mathbb{E}\{\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}\}$. Assuming that the noise contaminating the acceleration data from each sensor has the same variance, the noise covariance matrix $\boldsymbol{\Psi}$ can be written as $\mathrm{diag}([\sigma_1^2\mathbf{I}, \sigma_2^2\mathbf{I}, \dots, \sigma_{n_s}^2\mathbf{I}])$, where $\sigma_1^2, \sigma_2^2, \dots, \sigma_{n_s}^2$ are the noise variances for each sensor, and $\mathbf{I}$ is the identity matrix. Consequently, the mean and covariance of the posterior $p(\boldsymbol{f}|\widehat{\boldsymbol{y}})$ can be analytically derived as:

$$\bar{\boldsymbol{f}} = \mathbf{K}_{fy}{\mathbf{K}_{yy}}^{-1}\widehat{\boldsymbol{y}} = \mathbf{K}_{ff}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{K}_{ff}\mathbf{H}^{\mathrm{T}} + \boldsymbol{\Psi}\right)^{-1}\widehat{\boldsymbol{y}}, \tag{3.18}$$

$$\begin{aligned} \mathrm{cov}(\boldsymbol{f}) &= \mathbf{K}_{ff} - \mathbf{K}_{fy}{\mathbf{K}_{yy}}^{-1}\mathbf{K}_{yf} \\ &= \mathbf{K}_{ff} - \mathbf{K}_{ff}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{K}_{ff}\mathbf{H}^{\mathrm{T}} + \boldsymbol{\Psi}\right)^{-1}\mathbf{H}\mathbf{K}_{ff} \end{aligned} \tag{3.19}$$

Note that the posterior distribution is governed by the $\mathbf{K}_{ff}$ and $\boldsymbol{\Psi}$, where $\mathbf{K}_{ff}$ is determined by the hyperparameters in the covariance kernel functions characterizing the GP for forces, $\boldsymbol{\Psi}$ is controlled by noise-related hyperparameters. The optimal hyperparameter could be determined by minimizing the negative log marginal likelihood, with the objective function and its partial derivatives formulated as:

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{2}\widehat{\boldsymbol{y}}^{\mathrm{T}}{\mathbf{K}_{yy}}^{-1}\widehat{\boldsymbol{y}} + \frac{1}{2}\log\left|\mathbf{K}_{yy}\right| + \frac{n_t\left(n_s + n_f\right)}{2}\log 2\pi, \tag{3.20}$$

$$\frac{\partial\mathcal{L}(\boldsymbol{\theta})}{\partial\theta_i} = \frac{1}{2}\mathrm{tr}\left(\left({\mathbf{K}_{yy}}^{-1} - \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\right)\frac{\partial\mathbf{K}_{yy}}{\partial\theta_i}\right), \tag{3.21}$$

with

$$\boldsymbol{\alpha} = {\mathbf{K}_{yy}}^{-1}\widehat{\boldsymbol{y}}. \tag{3.22}$$

3.4. Numerical example: a truss bridge

3.4.1. Model description

Figure 3.2 depicts a 3-dimensional truss bridge, a typical example of the Pratt style bridge widely used in railway infrastructure. The structure consists of multiple components, including lower and upper chord, end raker, vertical and diagonal, portal girder, stringer, cross girder, and chord bracing, all labeled in Figure 3.2. Each component is prefixed with a serial number indicating its cross-sectional index. This truss bridge, which is subject to external impacts, is analyzed in this section using the proposed nonparametric Bayesian multi-task learning approach for force reconstruction.

The FE model of the truss bridge is constructed based on the Euler-Bernoulli beam theory. It treats all components as 3-dimensional, 2-node beam elements, where each node possesses 6 DOF: displacements and rotations with respect to three axes. In accordance with practical engineering constraints for truss bridges, nodes at positions 0, 8, 27, and 35 are constrained. The vertical DOF and rotational DOF along the x and z axes are all constrained for these nodes. Furthermore, the displacement DOF along the x and y axes are specifically constrained at node 0, along the y axis at node 8, and along the x axis at node 27. The bridge encompasses 50 nodes and 120 elements, resulting in a total of 284 DOF upon applying boundary constraints. The parameters for the FE models are detailed in Table 3.1, employing letters “A” to “F” to represent six different cross-section configurations for the structural components depicted in Figure 3.2. Note that the densities of the stringers and cross girders are modeled three times greater than those of other components to account for possible additional non-structural mass typically present in truss bridges. Based on the configuration of the FE model and the specified parameter values, the global mass and stiffness matrices are constructed. Subsequently, a primary modal analysis is conducted to validate the effectiveness of the FE model. The analysis confirms that the natural frequencies and mode shapes closely align with the existing literature (Dai et al., 2016). The first five natural frequencies for the constructed model are 3.13, 6.07, 7.86, 10.57 and 11.60 Hz, respectively. To model the energy dissipation mechanism, Rayleigh damping is utilized, which is expressed as a linear combination of the mass and stiffness matrices: $\mathrm{C} = \alpha\mathrm{M} + \beta\mathrm{K}$. The coefficients α and β can be derived using the formula:

$$\zeta_i = \frac{\alpha}{2\omega_i} + \frac{\beta\omega_i}{2}, \tag{3.23}$$

where ζ_i and ω_i are the modal damping ratio and radius natural frequency of the ith mode, respectively. For this study, a damping ratio of 2% is assigned to both the first and second modes ($\zeta_1 = \zeta_2 = 2\%$). Consequently, the values for α and β are determined to be $5.19\times10^{-1}$and 6.92×10^{-4}, respectively.

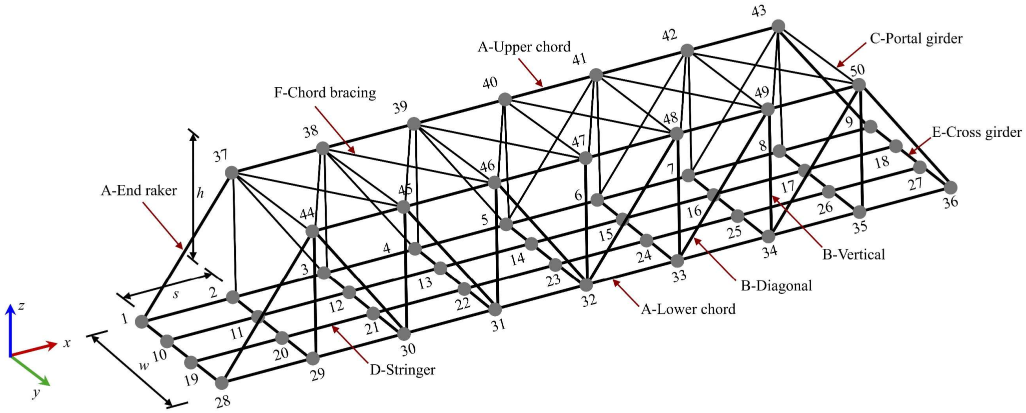


Figure 3.2. General view of Pratt truss bridge.

After defining the motion equations for the Pratt truss bridge, structural responses are analyzed subject to the external impact loads. This simulation not only serves to validate the proposed algorithm but also closely resembles real-world scenarios where impacts are caused by external vehicles or ships. If the impact load can be successfully calculated using the proposed algorithm after a collision incident, it would greatly assist in the subsequent decision-making process. In this study, the impact load is simulated as a point load applied to node 31 of the truss bridge, with its time-varying profile defined by the following equation:

$$f(t) = F_0(t - t_0)^2 e^{-\lambda_0(t-t_0)}. \tag{3.24}$$

In this equation, F_0, t_0, and λ_0 represent three control parameters associated with the impact force $f(t)$. F_0 governs the peak value of the impact force, λ_0 regulates the frequency band of the excitation, and t_0 determines the time at which the impact force reaches its maximum amplitude. As shown in Figure 3.3(a), the impact force history is obtained by configuring F_0, t_0, and λ_0 as 1×10^8, 0.3, and 12, respectively. Subsequently, a transient dynamic analysis is conducted with the 4th order Runge Kutta method, yielding the complete responses of the structure when subjected to the impact force.

Table 3.1. (a) Geometrical and material parameters of the truss bridge FE model.

Parameter	Symbol	Value	Unit	Parameter	Symbol	Value	Unit
Width	w	7.00	m	Elastic modulus	E	2.10×10^{11}	N/m^2
Height	h	5.00	m	Shear modulus	G	8.10×10^{10}	N/m^2
Spacing	s	5.00	m	Density (excluding stringer and cross girder)	ρ_1	7.85×10^{3}	kg/m^3
Density (stringer and cross girder)	ρ_2	3.14×10^{4}	kg/m^3				

Table 3.1. (b) Geometrical and material parameters of the truss bridge FE model.

Parameter	Symbol	Unit	Value					
			A	B	C	D	E	F
Cross section areas	$A_{A\sim F}$	m^2	4.56×10^{-2}	2.01×10^{-2}	1.04×10^{-2}	3.94×10^{-2}	2.68×10^{-2}	2.94×10^{-3}
Polar moment of inertias	$J_{A\sim F}$	m^4	2.35×10^{-3}	5.52×10^{-4}	9.36×10^{-5}	8.88×10^{-6}	6.75×10^{-6}	6.52×10^{-6}
Moment of inertias (horizontal local axis)	$I_{yA\sim F}$	m^4	1.10×10^{-3}	4.60×10^{-4}	6.04×10^{-5}	3.20×10^{-4}	2.14×10^{-4}	4.18×10^{-6}
Moment of inertias (vertical local axis)	$I_{zA\sim F}$	m^4	2.48×10^{-3}	2.93×10^{-4}	6.04×10^{-5}	2.22×10^{-3}	9.57×10^{-4}	4.18×10^{-6}

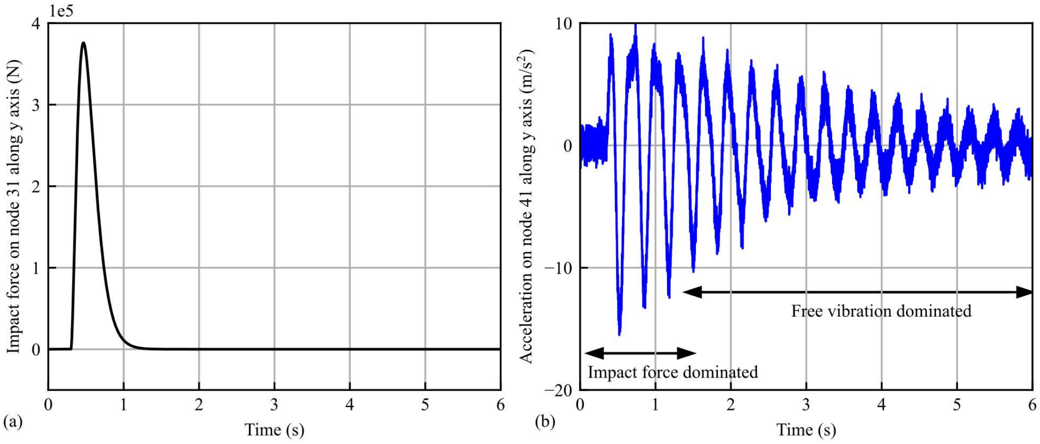


Figure 3.3. (a) The applied force over node 31 along y axis; (b) The acceleration response on node 41 along y axis with added noise.

Since the practical measured acceleration data are prone to noise contamination, different levels of white Gaussian noise are added to the simulated acceleration data $\tilde{\boldsymbol{y}}$ from each individual channel to obtain the simulated measurements $\hat{\boldsymbol{y}}$. The mean of the added noise is zero, while the variance σ_n^2 is determined based on the following equation:

$$\sigma_n^2 = \frac{\mathbb{E}(\tilde{\boldsymbol{y}} \odot \tilde{\boldsymbol{y}})}{\mathrm{SNR}}, \tag{3.25}$$

where $\mathbb{E}(\tilde{\boldsymbol{y}} \odot \tilde{\boldsymbol{y}})$ represents the expected value of the element-wise product of $\tilde{\boldsymbol{y}}$ itself, and SNR denotes the signal-to-noise ratio. Eventually, the simulated measurements $\hat{\boldsymbol{y}}$ could be obtained by adding the randomly sampled noise with the simulated acceleration data $\tilde{\boldsymbol{y}}$. Figure 3.3(b) demonstrates the simulated acceleration measurement of an SNR of 30. It is important to note that the vibration data is segmented into domains dominated by impact force and free vibration. To minimize computational burden and avoid the use of superfluous data, only the data from the time intervals dominated by impact force are utilized for subsequent force reconstruction (De Simone et al., 2019).

3.4.2. Force reconstruction for truss structure

Based on the constructed FE model of the truss bridge, the **H** matrix, as described in Eq. (3.12), can be formulated. This matrix is subsequently utilized as prior knowledge embedded to the MTGP. The force history in the MTGP model is modelled as a GP with a squared exponential kernel function, defined as:

$$k(x, x'; \boldsymbol{\theta}) = \sigma_f^2 \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right). \tag{3.26}$$

The kernel function includes two hyperparameters: the lengthscale ℓ and the output covariance σ_f^2. Additionally, it is assumed that the measured data consists of the ground truth combined with Gaussian noise, where the noise variance is denoted as σ_1^2. These three hyperparameters, ℓ, σ_f^2, and σ_1^2, are used to parameterize the negative log marginal likelihood shown in Eq. (3.32). Substituting the simulated measurements $\hat{\boldsymbol{y}}$ with a sampling frequency of 2,000 Hz over 1.5 s into this formula, optimization is performed using the gradient-based L-BFGS-B algorithm. The results of the optimization are displayed in Figure 3.4, which illustrates that the negative log marginal likelihood values exhibit local convexity over the examined range. A star marks the location of the optimal hyperparameters for the task of force reconstruction.

To verify the effectiveness of the nonparametric Bayesian multi-task learning under varying noise conditions, simulated acceleration measurements are generated at three different SNRs: 20, 30, and 40. It is anticipated that the posterior distributions of the forces derived from these measurements will vary; specifically, the confidence interval for the force calculated at an SNR of 40 is expected to be narrower than that from an SNR of 20. This difference arises because measurements with a lower SNR of 20 contain more uncertainties and errors compared to those with higher SNR.

Consequently, using data with higher uncertainty for inference leads to larger confidence intervals in the posterior distributions. By substituting the optimal hyperparameters obtained from the minimizing the negative log marginal likelihood (as detailed in Eqs. (3.18) and (3.19)), the posterior distributions of forces under three noise conditions are derived. Figure 3.5 illustrates the calculated posteriors, showing that the confidence intervals of the posteriors widen as the SNR decreases, indicating an increase in noise dominance. The result demonstrates the capability of the proposed nonparametric Bayesian multi-task learning approach to effectively capture and quantify the uncertainties present in the original data. The probabilistic nature of this approach in force identification offers valuable insights into the reliability of the measured acceleration data for conducting time-domain force identification.

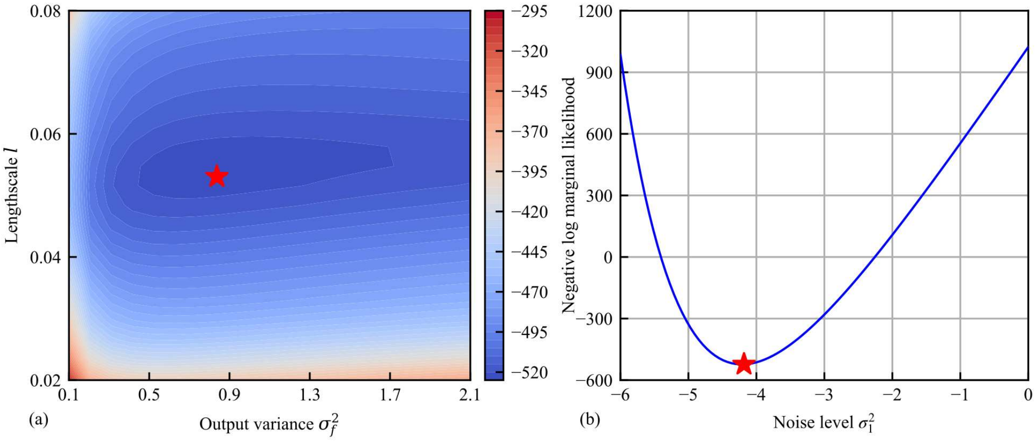


Figure 3.4. Negative log marginal likelihood function with respect to parameters of the squared exponential kernel, where the star indicates the optimal solution.

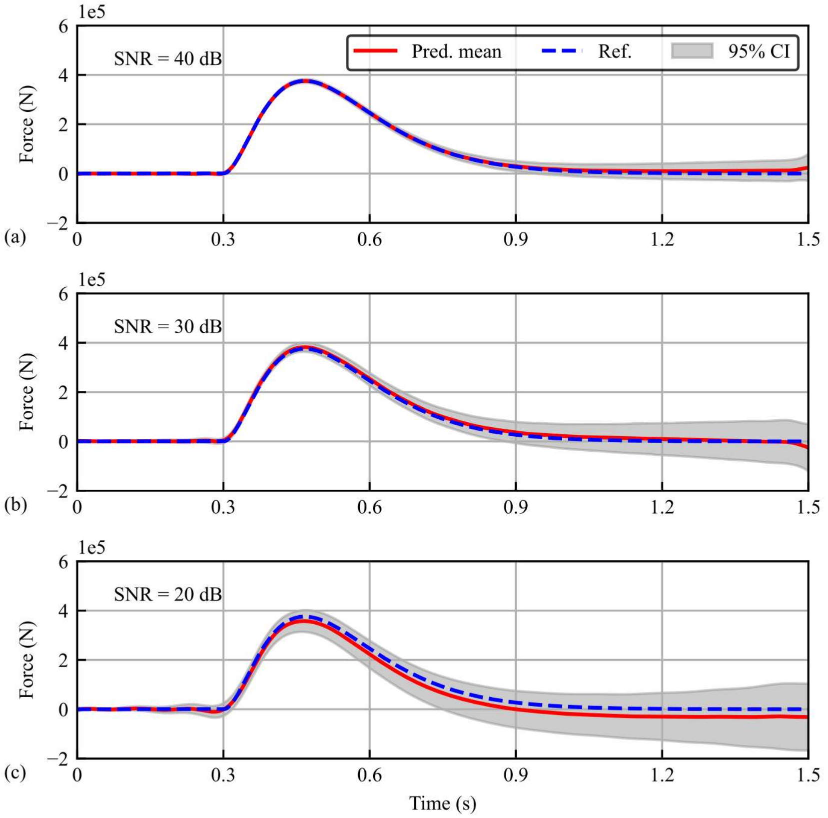


Figure 3.5. Posteriors of forces obtained using the proposed method from noisy acceleration response at node 41 along the *y* axis.

The force posterior derived from limited accelerometer data can be instrumental in reconstructing the full system response. In this case, the mean function of the identified force, calculated at an SNR of 30, serves as the input for the FE model. The estimated impact force is assumed to activate within a 1.5 s interval, with force values set to zero beyond this period. As a result, full system responses are generated, and acceleration data from four samples are compared with reference data from the initial impact simulation. Figure 3.6 shows this comparison, demonstrating that the identified force provides a valuable basis for computations of the full system response. Notably, through the constructed FE model and the proposed nonparametric Bayesian multi-task learning method, the original acceleration data from node 41 along the y axis can be mapped to responses of the whole structure.

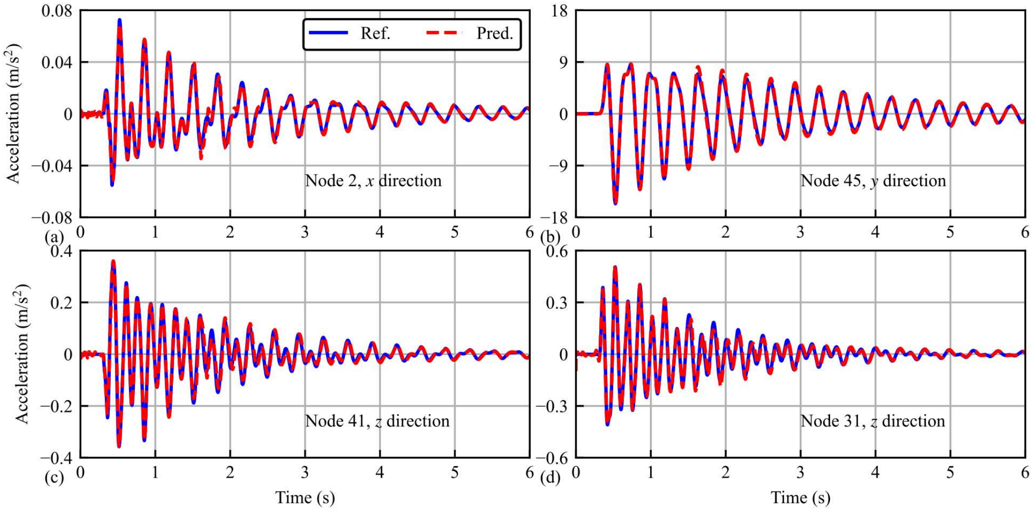


Figure 3.6. Samples of the calculated structure responses using the reconstructed force with the ground truth.

The accuracy of the proposed approach is compared with the conventional Tikhonov regularization method using three matrices, normalized root mean squared error (NRMSE), mean squared error (MSE), and mean absolute error (MAE). These matrices comprehensively quantify the deviation between the predicted results (from the MTGP) and the Tikhonov-regularized solution w.r.t. the reference force vector $\boldsymbol{f}_{\text{ref}}$. Their definitions are as follows:

$$\text{NRMSE}\left(\boldsymbol{f}_{\text{ref}}, \bar{\boldsymbol{f}}_{\text{pred}}\right) = \frac{\left\|\boldsymbol{f}_{\text{ref}} - \bar{\boldsymbol{f}}_{\text{pred}}\right\|_2}{\left\|\boldsymbol{f}_{\text{ref}} - \text{mean}(\boldsymbol{f}_{\text{ref}})\right\|_2}, \tag{3.27}$$

$$\text{MSE}\left(\boldsymbol{f}_{\text{ref}}, \bar{\boldsymbol{f}}_{\text{pred}}\right) = \frac{\left\|\boldsymbol{f}_{\text{ref}} - \bar{\boldsymbol{f}}_{\text{pred}}\right\|_2^2}{N} \tag{3.28}$$

$$\text{MAE}\left(\boldsymbol{f}_{\text{ref}}, \bar{\boldsymbol{f}}_{\text{pred}}\right) = \frac{\left\|\boldsymbol{f}_{\text{ref}} - \bar{\boldsymbol{f}}_{\text{pred}}\right\|_1}{N} \tag{3.29}$$

where $\boldsymbol{f}_{\text{ref}}$ is the reference force vector, $\bar{\boldsymbol{f}}_{\text{pred}}$ denotes the predictive results, $\|\cdot\|_1$ and $\|\cdot\|_2$ respectively represents the l_1-norm and l_2-norm, and N is the dimension of $\boldsymbol{f}_{\text{ref}}$. Lower values of NRMSE, MSE, and MAE indicate higher prediction accuracy, as they reflect smaller deviations from the reference. The performance of the two methods is evaluated across six SNR ranging from 15 to 45, corresponding to decreasing noise levels in the measured data. Measurement and force configurations are consistent with those illustrated in Figure 3.3, with 1,000 acceleration data points collected over a 1.5 s period. For each force reconstruction test, Tikhonov regularization parameters are optimized using the L-curve method. Figure 3.7 presents the results, demonstrating that under each SNR setting, the proposed nonparametric Bayesian multi-task learning method, which incorporates a GP prior for the force function, consistently outperforms Tikhonov regularization. The advantage of the proposed method becomes more pronounced at higher SNR values.

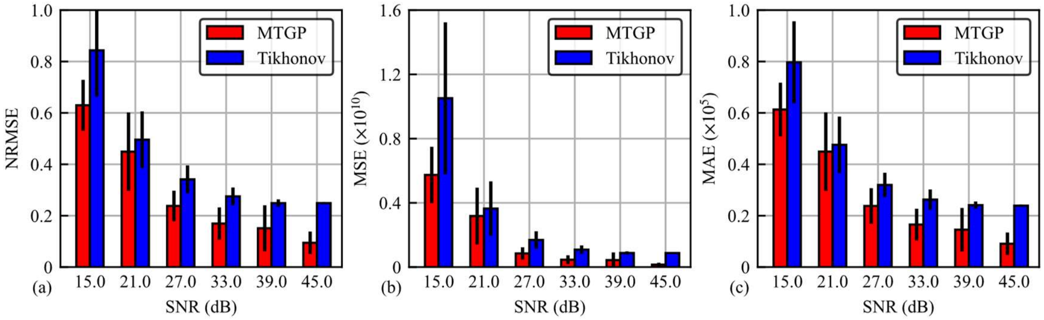


Figure 3.7. Force reconstruction accuracy comparison between the proposed method and Tikhonov regularization under various SNR.

In addition to noise, the number of data points (sampled within a 1.5-second window) used for reconstruction significantly influences the accuracy of results. To evaluate this effect, multiple experiments were conducted using y-axis acceleration data from Node 7 (SNR = 40) to identify the force on Node 31. The number of data points varied from

200 to 700 across trials, with computational time recorded for each case. For every data point quantity, 10 trials were performed, and the posterior mean forces were calculated alongside their corresponding NRMSE, MSE, and MAE. Figure 3.8 illustrates the mean and one standard deviation of the computational time, NRMSE, MSE, and MAE across the ten trials. When fewer than 550 data points are used, the three metrices exhibits a decreasing trend, suggesting improved predictive performance of the MTGP algorithm. Beyond 500 data points, the NRMSE stabilizes at consistently low values, indicating convergence. Conversely, computational time increases substantially as the number of data points grows from 200 to 700. Specifically, while the number of data points increases by a factor of 3.5, the average computational time rises by a factor of 14.3, highlighting a nonlinear scaling relationship.

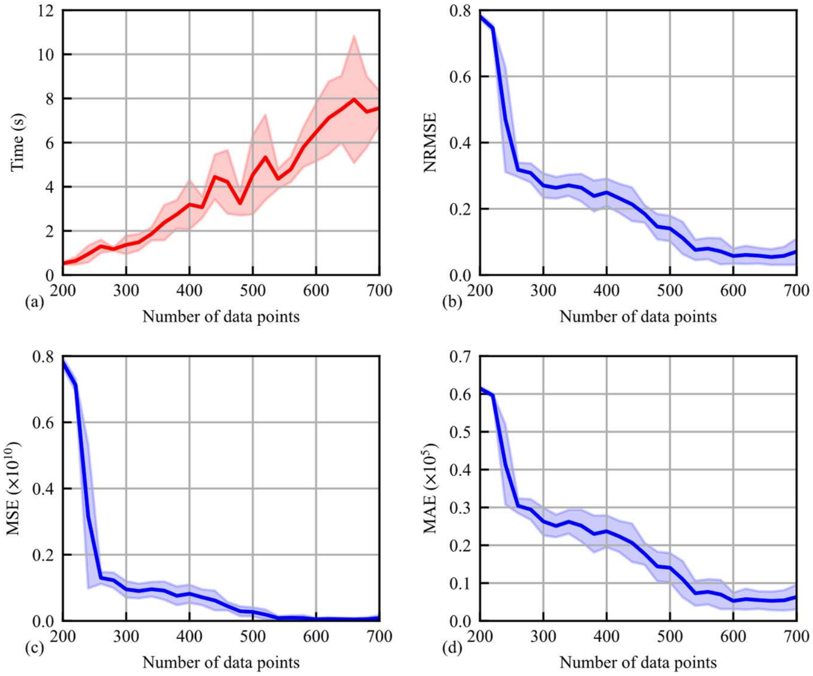
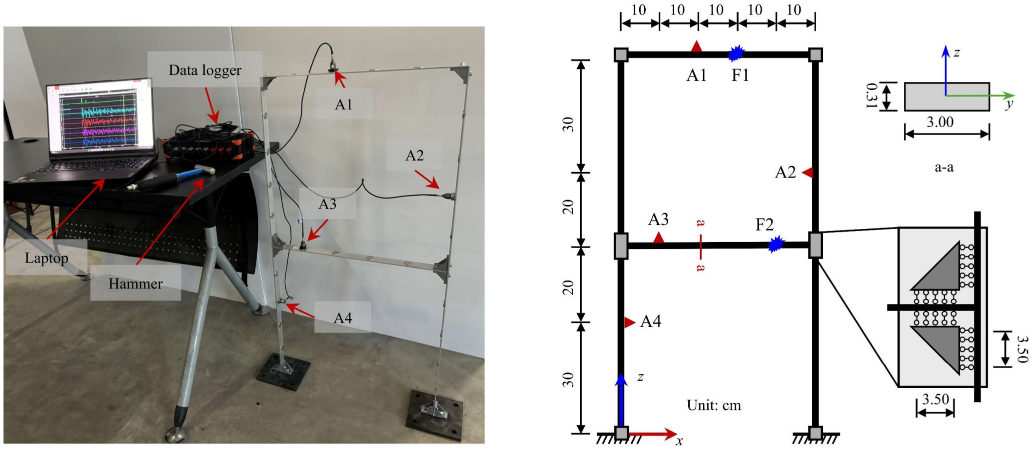


Figure 3.8. Force reconstruction accuracy and time consuming versus the volume of acceleration data used as input.

3.5. Experimental example: a frame structure

3.5.1. Measurement setup

A force reconstruction experiment is performed based on a frame structure. As depicted in Figure 3.9(a), the experimental setup consists of six aluminum beams connected by angle brackets and bolts to form the complete frame. A hammer is used to strike predetermined points on the structure, and the resulting accelerations are measured by four instrumented accelerometers, labeled A1 through A4. Simultaneously, the impact load is captured by a load cell integrated into the hammer. The data from these five channels are subsequently stored on a laptop for analysis at a sampling frequency of 20,000 Hz. A detailed illustration of the structural configuration, including accelerometer and impact locations and geometry of the frame, is provided in Figure 3.9(b). In this experiment, the recorded acceleration data serves as inputs to a nonparametric Bayesian multi-task learning model, and the computed posteriors of force histories will be compared with the measured impact load to validate the effectiveness of the proposed approach.



(a)

(b)

Figure 3.9. (a) Experimental setup for force reconstruction on a frame structure; (b) Idealized model of the frame structure with detailed configurations of accelerometers and forces.

3.6.2. FE model construction and calibration

An FE model is constructed for the frame structure using 2-dimensional, 2-node Euler-Bernoulli beam elements. Each node has 3 DOF: two translational DOF along the x and z axes, and one rotational DOF about the y axis. The beam elements have a length of 2.5 cm, leading to a total of 120 nodes and 120 elements in the FE model. Two nodes connected to the ground are fully constrained in all DOF. Additionally, for nodes located at each angle bracket (as shown in Figure 3.9(b)), all their rotational DOF are coupled. This results in unsymmetrical system mass and stiffness matrices, $\mathbf{M}$ and $\mathbf{K}$. After implementing all constraints and couplings, the FE model has 342 DOF. To achieve a refined damping modelling for the FE model, the modal damping for each mode is assumed to constitute the system damping matrix $\mathbf{C}$:

$$\mathbf{C} = \mathbf{\Phi}_l^{\mathrm{T}} \mathbf{Z} \mathbf{\Phi}_l^{\mathrm{T}} \mathbf{M}, \tag{3.30}$$

where $\mathbf{Z}$ is the modal damping matrix, a diagonal matrix with each term calculated by the damping ratio of the ith mode multiplied by its corresponding radial natural frequency; $\mathbf{\Phi}_l$ is the left eigenvector matrix obtained through eigen analysis of $\mathbf{M}^{\mathrm{T}^{-1}} \mathbf{K}^{\mathrm{T}}$.

An experimental modal analysis is conducted to measure modal information of the frame structure. Four accelerometers are evenly instrumented on each beam, batch by batch, and the hammer is utilized to generate a single peak excitation for each test case over a specific location, 10 cm vertically below the right upper corner of the frame. A total of six batches of measurements are obtained, each containing one channel for the measured force history and four channels for accelerations. For each batch of measurement, four FRF are calculated, resulting in a total of 24 FRF for the frame structure, denoted as $\hat{h}_i(f)$, where $i = 1,2,\ldots,24$. Using these FRF, the first 10 damping ratios are evaluated based on the 3dB method. The damping ratios for higher order modes are assumed to be the same as the 10th damping ratio. Consequently, the modal damping matrix $\mathbf{Z}$ is determined based on these measurements.

Beyond the directly measurable parameters for the FE model, the system matrices are governed by three parameters: the elastic modulus E, and densities ρ of the idealized beam, and the equivalent added mass m for each accelerometer. Incorporating the equivalent added mass for the accelerometers in the FE model is crucial because initial calculations indicate that the total mass of the frame structure is approximately 800 g, while the total mass of the accelerometers (including magnetic bases) weighting on a scale exceeds 100 g. Neglecting this added mass could significantly impact the fidelity of the FE model. Note that for each batch of measurements, the locations of the accelerometers differ, necessitating adjustments to the added mass to compose the system mass matrix $\mathbf{M}$. To calibrate the FE model, a model updating is conducted based on the calculated FRF. The evaluated frequency band is constrained from 10 to 250 Hz, with a grid length of 1 Hz. The optimization problem is defined as:

$$\min \sum_{i=1}^{24} \sum_{j=10}^{250} \left[\log_{10} \frac{|\hat{h}_i(j)|}{|h_i(j;E,\rho,m)|} \right], \tag{3.31}$$

where $|\cdot|$ denotes the magnitude operator; $h_i(j;E,\rho,m)$ is the ith FRF at frequency j predicted from the FE model given the parameters E, ρ, and m, respectively. The L-BFGS-B algorithm is adopted for optimization, with the initial guessed values for elastic modulus, densities and equivalent added mass being 6.5×10^{10} N/m^2, 3×10^3 kg/m^3, and 3.5×10^{-2} kg, respectively. Upon convergence of the optimization, the scaling factors for the three parameters are 0.894, 0.831, and 0.518, respectively. Figure 3.10 shows the comparison between four FRF from the measurement and the FE model before and after updating. Figure 3.11 shows the measured natural frequencies versus the FE model computed natural frequencies before and after model updating. Both results indicate that the calibration process has improved the accuracy of FE model to the reality. The predictive performance of the FE model is better in lower frequency than in higher frequency, as the high-frequency vibration of the frame structure is more sensitive to the fidelity of the FE model. With the optimal parameters determined for the FE model, force reconstruction can be implemented in the following subsection.

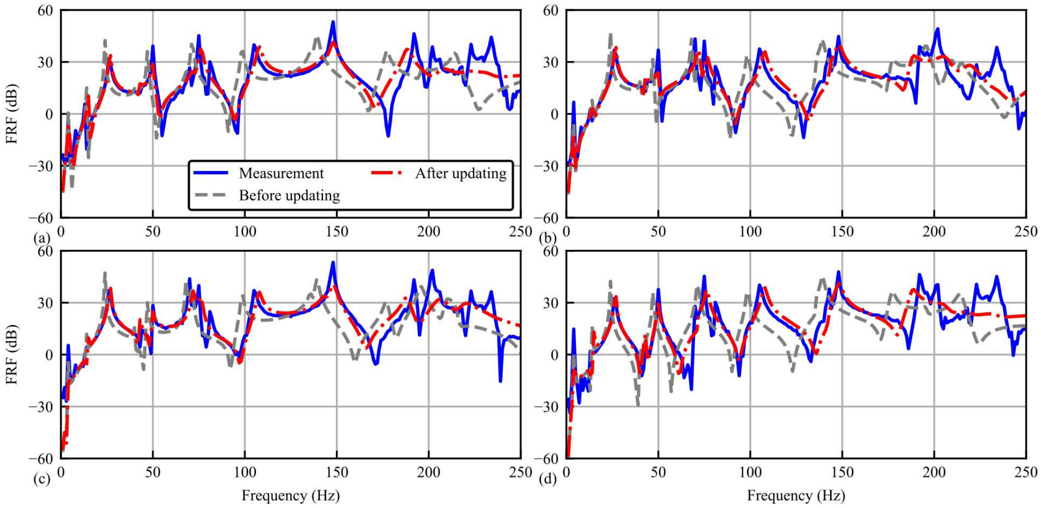


Figure 3.10. Measured and FE model (before and after updating) output FRF.

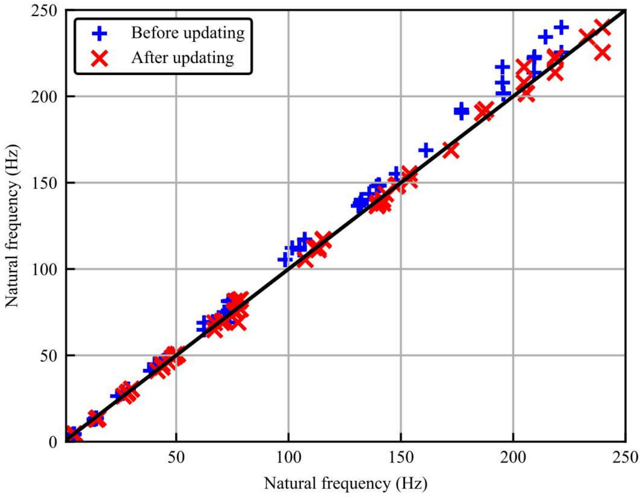


Figure 3.11. Measured and calculated natural frequencies before and after model updating.

3.5.3. Force reconstruction in the presence of epistemic noise

In contrast to the simulation case, force reconstruction for the real-world structure must always consider the deviation between the calibrated FE model and the actual structure. Figures 3.10 and 3.11 highlight this discrepancy, which can be challenging and costly to mitigate. However, this difference can also be viewed as the epistemic noise (error, uncertainty) inherent in the structural dynamical system. The propagation of this epistemic uncertainty can be quantified using the proposed nonparametric Bayesian multi-task learning model. The posteriors over the forces can reveal the extent of uncertainty at a specific time step, taking into account both the FE model's imperfect simulation of the real-world structure and the measurement noise.

To enhance the clarity of epistemic uncertainty, consider the data from impact tests as outlined in the experimental setup depicted in Figure 3.12 illustrates a pair of time histories for force (applied at F2) and acceleration (measured at A3). One could first assume that the force measurement is perfectly accurate and serves as a benchmark. When this measured force is input into the FE model to calculate responses at locations

A1 to A4, deviations are observed between the calculated responses and the actual measurements. As depicted in Figure 3.13, the measured acceleration time histories are represented by blue solid lines, while the red dotted lines represent the calculated structural responses. Notably, the measured accelerations tend to exhibit greater energy at higher frequencies compared to the computed responses. These discrepancies can be attributed to epistemic errors between the real-world tested frame structure and the calibrated FE model. Under such conditions, measurement noise, also known as aleatory noise, appears to play a secondary role in influencing the accuracy of force reconstruction.

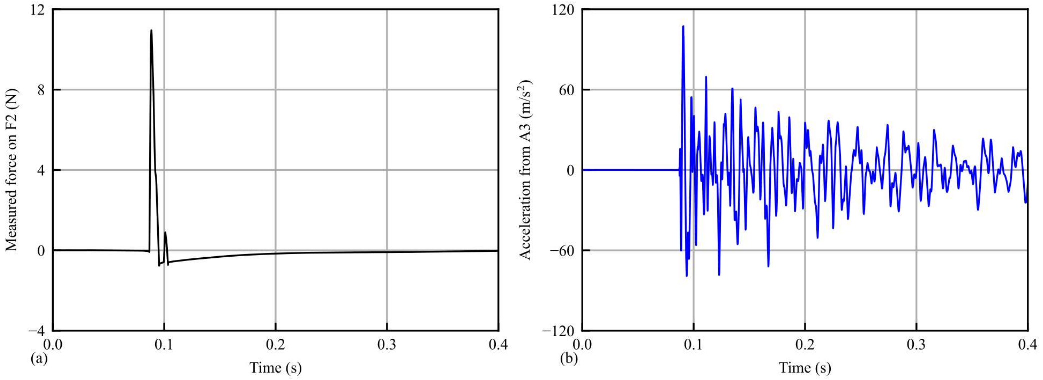


Figure 3.12. (a) Applied Impact force on F2; (b) Collected acceleration data from A3.

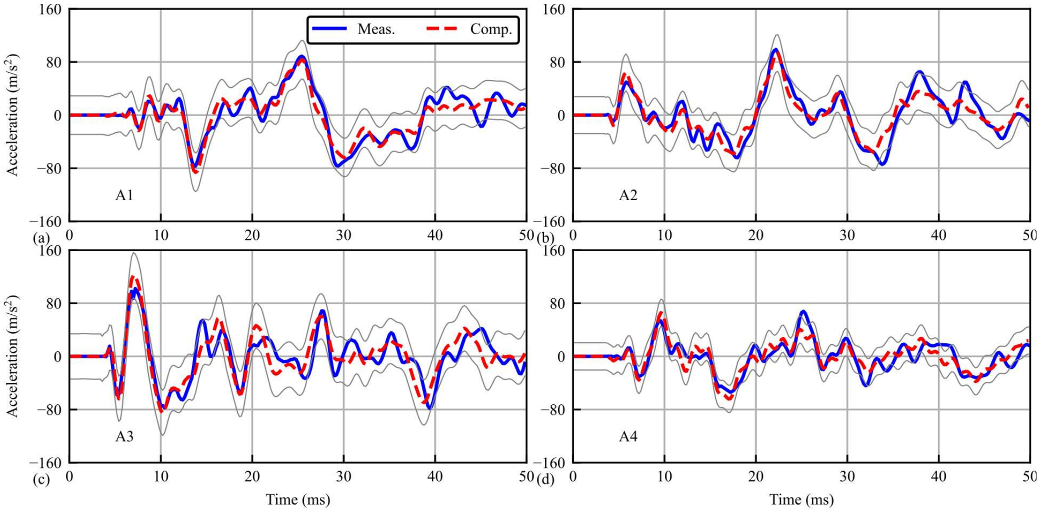


Figure 3.13. Comparison between the measured accelerations with the calculated acceleration responses based on the calibrated FE model.

If the analysis only considers measurement noise and neglects epistemic uncertainty, the results from our proposed nonparametric Bayesian multi-task learning approach can lead to two scenarios in multiple tests: one where the optimization process fails due to an ill-conditioned problem setting, and another where there is a high level of confidence in the posterior estimates of forces that may not be accurate. To address this issue, it is crucial to incorporate epistemic noise into the calculation of force posteriors. To effectively manage this, a prior assumption is applied on the epistemic noise, quantified using an SNR index. Specifically, the SNR index is used to gauge the deviations between measured and computed values. For example, if setting the SNR at 10 for the time histories depicted in Figure 3.13, the variance can be calculated using Eq. (3.19). The gray solid line in Figure 3.13 illustrates the bounds of two standard deviations, indicating that by adding Gaussian noise to the measurements, epistemic noise can be accurately included. Consequently, the propagation of this uncertainty to the calculated forces can be quantified using the approach proposed in this chapter.

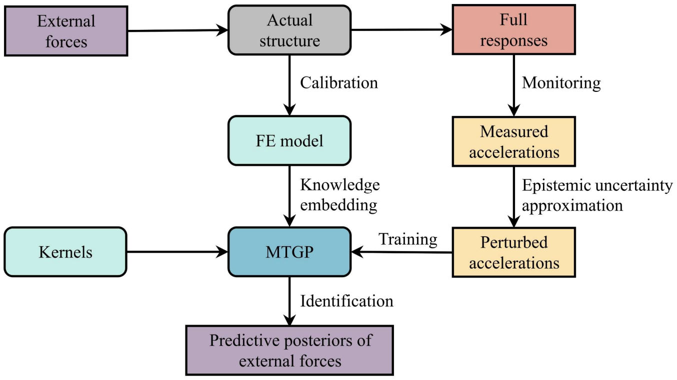


Figure 3.14. Workflow for force reconstruction in the presence of epistemic noise.

Figure 3.14 outlines the workflow for addressing epistemic uncertainty in real-world scenarios. This workflow enables effective quantification of the posteriors of external forces. To validate this approach, four dedicated impact loading tests are conducted.

The first two tests apply impact forces on F1, while the last two target F2. Each test involved impacts that exceeded a single peak force. By embedding knowledge from the FE model and incorporating perturbed accelerations from all four channels into the MTGP, the posteriors of impact histories are computed for each test, as depicted in Figure 3.15. The results demonstrate that the proposed approach can effectively capture the characteristic peaks of the impacts. Additionally, the 95% confidence interval provides insight into the force uncertainties, given the measurements and the calibrated FE model.

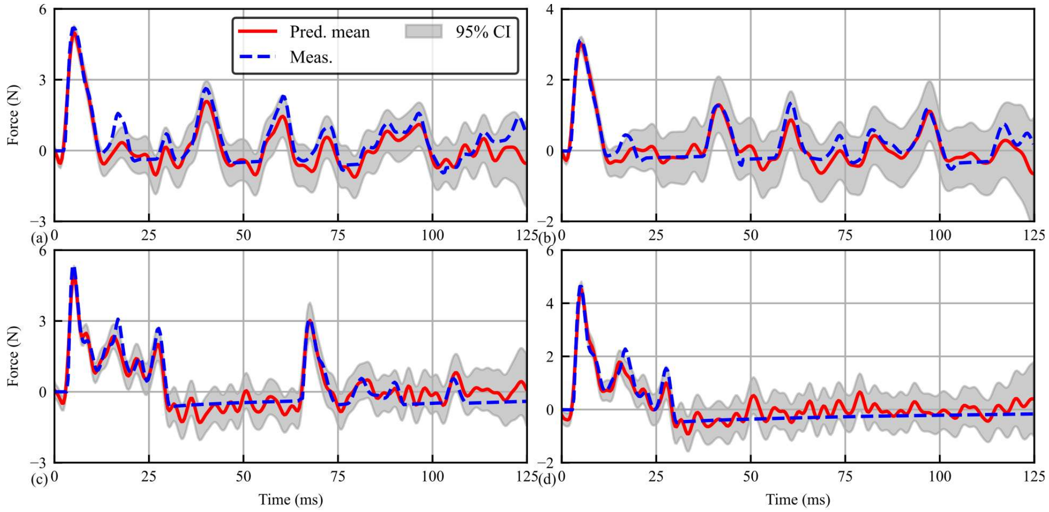


Figure 3.15. Measured forces and reconstructed forces based on the proposed nonparametric Bayesian multi-task learning approach.

This study evaluates the influence of accelerometer configurations on time-domain force reconstruction, with a focus on identifiability—the ability to uniquely determine forces from sensor data—when reducing sensor counts. Identifiability is critical in underdetermined systems, where insufficient sensors may fail to resolve unique force solutions. Fifteen accelerometer configurations (Table 3.2), varying in sensor numbers (1–4) and positions (A1–A4), were tested using five impact trials at force locations F1 and F2. Force histories were reconstructed from measured accelerations, with accuracy quantified via NRMSE, MSE, and MAE relative to ground-truth hammer measurements.

Key findings demonstrate that sensor placement and redundancy directly govern identifiability. Configurations with fewer than three accelerometers (e.g., A1 alone, Table 3.2, Index 1) exhibit high NRMSE (0.320–0.716 for F1, 0.340–0.716 for F2), indicating poor identifiability due to insufficient spatial resolution. Conversely, configurations including A1, positioned to capture dominant structural modes, consistently yield lower errors (e.g., A1+A4, Index 7: NRMSE = 0.265 for F1, 0.369 for F2). This underscores A1's critical role in resolving identifiability by providing essential modal information. While increasing sensors improves accuracy (e.g., four-sensor configurations reduce F2 MAE to 0.454, Index 15), diminishing returns emerge beyond four sensors, as redundancy compensates for identifiability limitations rather than introducing new information.

Kernel selection in the GP-based Bayesian framework further interacts with identifiability constraints. The white kernel (Table 3.3a–b), which assumes force independence, performs poorly (e.g., A2;F1: NRMSE = 0.793, MSE = 1.317, MAE = 0.930), as it fails to model temporal correlations, exacerbating identifiability issues in sparse configurations. In contrast, the squared exponential kernel enforces smoothness, implicitly regularizing underdetermined systems (e.g., A1,A4;F1: NRMSE = 0.261, MSE = 0.154). Matérn kernels, with tunable flexibility, improve identifiability as sensor counts increase (e.g., Matérn 5/2 for A1,A3,A4;F2: NRMSE = 0.351, MAE = 0.538), leveraging spatial correlations to resolve forces. These results emphasize that kernel design must align with identifiability constraints imposed by sensor configurations.

Table 3.2. Force reconstruction accuracy comparison via different number and position of accelerometers.

Index	Acc. config.	NRMSE (F1)	MSE (F1)	MAE (F1)	NRMSE (F2)	MSE (F2)	MAE (F2)
1	A1	0.320	0.211	0.357	0.408	0.770	0.629
2	A2	0.487	0.495	0.538	0.560	1.416	0.868
3	A3	0.716	1.193	0.771	0.340	0.602	0.572
4	A4	0.696	1.121	0.768	0.522	1.665	0.860
5	A1, A2	0.245	0.126	0.277	0.412	0.788	0.646
6	A1, A3	0.333	0.219	0.360	**0.262**	**0.315**	**0.410**
7	A1, A4	0.265	**0.147**	0.296	0.369	0.646	0.558
8	A2, A3	0.424	0.356	0.453	0.386	0.663	0.604
9	A2, A4	0.475	0.466	0.515	0.450	0.920	0.678
10	A3, A4	0.558	0.674	0.610	0.418	0.929	0.683
11	A1, A2, A3	0.266	**0.147**	0.291	0.286	0.393	0.450
12	A1, A2, A4	**0.244**	**0.147**	**0.278**	0.336	0.576	0.528
13	A1, A3, A4	0.268	0.154	0.297	0.315	0.448	0.470
14	A2, A3, A4	0.470	0.473	0.520	0.341	0.535	0.530
15	A1, A2, A3, A4	0.251	**0.147**	0.280	0.307	0.439	0.454

Table 3.3 (a). Performance comparison of nonparametric Bayesian multi-task learning using different kernels.

Kernel	Kernel function expressions	NRMSE			
		A2; F1	A1; F2	A1, A4; F1	A1, A3, A4; F2
Squared exponential	$k(x,x') = \sigma_f^2 \exp\left(-\frac{r^2}{2\sigma_l^2}\right)$	0.488	**0.411**	0.261	**0.341**
White	$k(x,x') = \sigma_f^2 \delta(x,x')$	0.793	0.601	0.587	0.493
Exponential	$k(x,x') = \sigma_f^2 \exp\left(-\frac{r}{\sigma_l}\right)$	0.627	0.521	0.350	0.387
Rational quadratic	$k(x,x') = \sigma_f^2 \left(1 + \frac{r^2}{2\alpha\sigma_l^2}\right)^{-\alpha}$	**0.432**	0.416	0.258	0.464
Matérn 3/2	$k(x,x') = \sigma_f^2 \left(1 + \frac{\sqrt{3}r}{\sigma_l}\right) \exp\left(-\frac{\sqrt{3}r}{\sigma_l}\right)$	0.600	0.475	0.294	0.389
Matérn 5/2	$k(x,x') = \sigma_f^2 \left(1 + \frac{\sqrt{5}r}{\sigma_l} + \frac{5r^2}{3\sigma_l^2}\right) \exp\left(-\frac{\sqrt{5}r}{\sigma_l}\right)$	0.464	0.456	**0.246**	0.351

Table 3.3 (b). Performance comparison of nonparametric Bayesian multi-task learning using different kernels.

Kernel	MSE				MAE			
	A2; F1	A1; F2	A1, A4; F1	A1, A3, A4; F2	A2; F1	A1; F2	A1, A4; F1	A1, A3, A4; F2
Squared exponential	0.498	**0.782**	0.154	0.604	0.538	**0.630**	0.297	**0.537**
White	1.317	1.618	0.740	1.141	0.930	0.911	0.708	0.776
Exponential	0.881	1.435	0.257	0.723	0.709	0.836	0.379	0.620
Rational quadratic	**0.386**	0.860	0.132	0.957	**0.478**	0.666	0.280	0.704
Matérn 3/2	0.801	0.863	0.182	0.739	0.660	0.704	0.329	0.598
Matérn 5/2	0.464	1.003	**0.124**	**0.595**	0.512	0.725	**0.270**	0.538

3.6. Summary

In this chapter, a nonparametric Bayesian-based framework for time-domain force reconstruction is proposed, leveraging the concept of multi-task learning. By incorporating GP priors for forces and embedding knowledge from FE models, the method enables the analytical calculation of force posterior distributions given measured acceleration data. Compared to previously developed Bayesian-based force reconstruction methods, the proposed nonparametric Bayesian multi-task learning approach is highly efficient in quantifying the propagation of uncertainties from measurements and the constructed FE model to time-domain forces. Additionally, compared to conventional regularization methods, the proposed method reconstructs forces more accurately by embedding the priors that correlate different force function values. The effectiveness of the proposed approach in uncertainty quantification and prediction accuracy improvement is demonstrated through two examples: a spatial truss numerical example and a frame structure experimental example. The results from both examples show that the force reconstruction problem can be successfully addressed using the proposed method. Furthermore, the ill-posed nature of the inverse problem is effectively managed.

Chapter 4.

Transient aerodynamic load reconstruction for maglev vehicles using onboard acceleration measurements

4.1. Introduction

This chapter presents the application of time-domain force reconstruction to the maglev vehicles subject to transient aerodynamic loads. Accurate estimation of aerodynamic loads is crucial for developing high-speed maglev trains. Traditionally, this estimation has been achieved through computational aerodynamic simulation or direct pressure measurement, both of which can be time-consuming and expensive. A novel framework is developed for reconstructing transient aerodynamic loads on maglev vehicles using on-board acceleration measurements. In this framework, an inverse mathematical model correlating the measured acceleration with external aerodynamic loads is derived from a well-calibrated maglev vehicle model. To address the ill-posedness inherent in solving the inverse mathematical model, the MTGP algorithm proposed in Chapter 3 is adopted. This approach treats all reconstructed transient aerodynamic loads as GP, enabling the calculation of closed-form posterior distributions of these aerodynamic loads. To validate the proposed framework, transient vibration data collected from an operational maglev train passing through a double-track tunnel are utilized for load reconstruction. The results demonstrate that the framework offers a cost-effective and efficient means to obtain aerodynamic loads, highlighting its practical relevance for aerodynamic field testing in the context of evolving high-speed maglev technologies.

4.2. Problem description of aerodynamic load reconstruction for maglev vehicles

4.2.1. Transient vibration of maglev vehicles

Transient vibrations of maglev vehicles refer to temporary, non-steady-state vibrations occurring in the vehicle system due to sudden changes or disturbances (Jönsson and Johansson, 2005). These vibrations are typically caused by two vehicles passing each other, a single vehicle moving through a double-track tunnel, or a single vehicle encountering crosswinds (Rocchi et al., 2018; Li et al., 2020b; Chen et al., 2019). The occurrence of transient vibrations may significantly influence the performance, comfort, and safety of the maglev system under extreme circumstances. In-situ testing reveals sudden changes in vehicle acceleration when the vehicle enters and exits a double-track tunnel. This is attributed to an abrupt change in the aerodynamic field around the vehicle as it transitions from a confined environment to an open environment. Figure 4.1 displays GPS velocity data alongside three sets of acceleration data collected from the front right of the maglev vehicle during three distinct intervals. As depicted in Figure 4.1(a), the intervals correspond to the following phases: the train moving in a straight line, entering a double-track tunnel, and exiting the tunnel. When the maglev vehicle passes through a double-track tunnel, the GPS signal may weaken or disappear completely. Almost simultaneously, significant unsteady vibrations are observed, as shown in Figure 4.1(c) and (d). Compared to the vibrations under normal operational conditions illustrated in Figure 4.1(b), the vibrations in Figure 4.1(c) and (d) predominantly occur at low frequencies. These are primarily caused by the rigid motion of the maglev vehicle assembly, rather than by the elasticity of the vehicle structures. This phenomenon suggests that transient variations in the pressure field surrounding the vehicle, characterized by aerodynamic loads, could be responsible for the abrupt changes in acceleration observed. Based on this observation, the objective is to use onboard measured acceleration data to inversely reconstruct the aerodynamic loads affecting the vehicle.

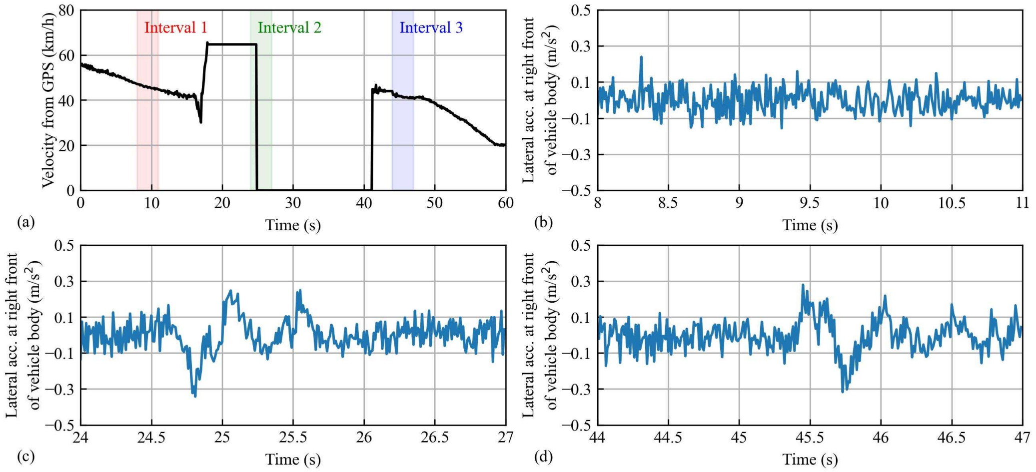


Figure 4.1. (a) Maglev operational velocity recorded via GPS; (b) Lateral accelerations in an open environment; (c) Lateral acceleration upon entering the tunnel; (d) Lateral acceleration upon exiting the tunnel.

4.2.2. Aerodynamic load reconstruction: an inverse problem

Generally, forward computation refers to calculating the full system response (output) under the action of aerodynamic loads (input) using an accurately calibrated multi-DOF model. The model can be calibrated using acceleration data collected from the operational maglev vehicle. However, reconstructing aerodynamic loads constitutes an inverse problem, i.e., determining the input functions from part of the system's output. Despite having an accurate model elucidating the relationship between input and output, reconstructing the aerodynamic load requires formulating an explicit mapping between the load functions and recorded accelerations with a limited number of DOF measured. Moreover, managing the ill-posed nature of the inverse problem is a significant challenge to ensure the algorithm yields stable solutions for aerodynamic loads by using noisy acceleration data. Without appropriate strategies to address the ill-posedness, minor data variations could lead to significant discrepancies in the reconstructed loads. Additionally, the noise present in the acceleration data can potentially affect the results. Therefore, understanding the mechanism of uncertainty propagation from noise to the reconstructed loads is a crucial area of study.

Figure 4.2 displays the technical flowchart for the proposed framework to reconstruct the transient aerodynamic load when a maglev vehicle enters or exits a double-track tunnel. Firstly, a model describing the motion of a maglev train is established. Despite numerous simplifications in the modeling process, the objective is to achieve a balance between model complexity and effective implementation of aerodynamic load reconstruction. Then, certain parameters within this model are calibrated using stationary random vibration data collected from a monitoring system installed on the vehicle. To calibrate the model, operational modal analysis is conducted to obtain the mode shapes and natural frequencies of the vehicle. The specific patterns of the mode shapes serve as a benchmark to assess the model's results, and associated natural frequencies are used for model validation. Based on the calibrated model, inverse analysis is conducted to establish a time-domain mathematical model correlating aerodynamic loads and measured acceleration data. To carefully handle issues of ill-posedness and error propagation from measurements in the established model, the proposed MTGP algorithm is used to derive the closed-form posterior distribution of the reconstructed aerodynamic loads using the noisy transient vibration data collected when the vehicle enters or exits the double-track tunnel.

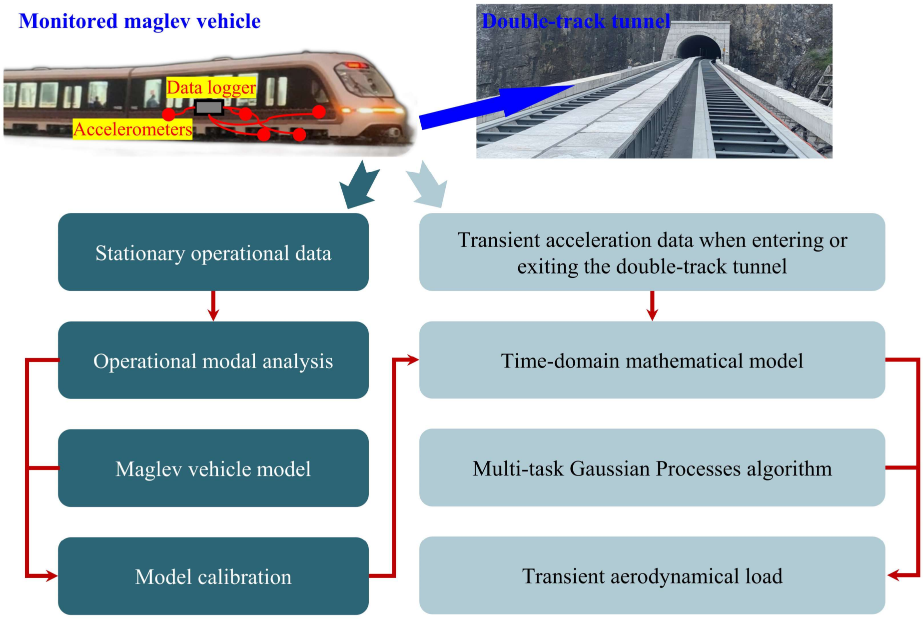


Figure 4.2. Technical flowchart for transient aerodynamic load reconstruction using acceleration data.

4.3. Maglev vehicle system modelling

The maglev vehicle considered in this chapter comprises a car body and five levitation suspension bogies. The car body is considered rigid and has five DOF: movements along the y and z axes, as well as roll, pitch, and yaw motions. The car body is elastically attached to the levitation bodies with a pair of secondary suspensions, providing lateral, longitudinal, and vertical stiffness and damping. Each pair of levitation bogies consists of two levitation modules connected on two sides via anti-rolling bars, decoupling the dynamics of each side bogie and limiting the rolling rotation of the bogies. Hence, each levitation module is assumed to be a 5-DOF rigid body. Moreover, four electromagnets are rigidly connected to individual levitation modules and interact with the guideway to realize levitation functions. Thus, the maglev vehicle-guideway system model involves:

i. One vehicle body, modeled as a rigid body with 5 DOF including vertical and lateral displacements (i.e., y_c and z_c, respectively), and pitching, yawing and rolling rotations (i.e., ϕ_c, θ_c, and ψ_c respectively);
ii. Ten levitation modules, each modeled as a rigid body with 5 DOF, i.e., y_{bi}, z_{bi}, ϕ_{bi}, θ_{bi}, and ψ_{bi}, respectively for $i = 1, \dots, 10$;
iii. Twenty secondary suspensions, each modeled as spring-dashpot elements with stiffness k_{sx}, k_{sy} and k_{sz}; and damping c_{sy}, c_{sy} and c_{sz};
iv. Forty primary electromagnetic forces modeled as linear spring-dashpot elements with stiffness k_{pz} and k_{py}; and damping c_{pz} and c_{py};
v. Ten anti-rolling bars modelled as linear spring pair elements with stiffness k_r each.

Hence, the total number of DOF involved in the multi-rigid-body dynamics sums up to 55.

As depicted in Figure 4.3, the maglev vehicle model simplifies essential dynamic calculations for understanding key aspects of vibrations when subjected to aerodynamic loads. This is accomplished by focusing on eliminating complex deformation interactions, which are considered less critical according to previous research (Dou et al., 2017; Liu et al., 2017). Importantly, transient aerodynamic loads, treated as external excitations, predominantly manifest energy content below 5 Hz (Liu et al., 2017; Zeng et al., 2021). This frequency range significantly overlaps with the frequency band of rigid body motions of the maglev vehicles being studied but does not overlap with the frequency bands associated with elastic vibrations caused by material elasticity. In other words, the aerodynamic loads primarily induce rigid body motion rather than causing structural deformations. Furthermore, the natural frequencies of the vehicle's elastic vibrations generally exceed those of rigid body motion. This characteristic allows for the practical separation of these two types of motion in the collected signals using filtering techniques. By employing filters, rigid body motions can be isolated from higher frequency elastic vibrations, enabling more targeted analyses. Consequently, analysis is focused within this specific frequency threshold, offering a focused and practical approach to examining the dynamic responses of the maglev vehicle under operational conditions, with an emphasis on the effects on rigid body dynamics while effectively filtering out elastic vibrations.

With these assumptions, the equation of motion for the maglev vehicle model is derived based on the Euler-Lagrange equation. The Lagrangian identity is defined as:

$$L = T - V, \tag{4.1}$$

here T represents kinetic energy and V stands for potential energy. Both are functionals with respect to velocity or displacement. The total kinetic energy of the system is the summation of vehicle body's kinetic energy T_c and each bogie's kinetic energy T_{bi}, which is given by:

$$T = T_c + \sum_{i=1}^{10} T_{bi}\,, \tag{4.2}$$

with:

$$T_c = 0.5\left(m_c \dot{y}_c^2 + m_c \dot{z}_c^2 + J_{cx}\dot{\phi}_c^2 + J_{cy}\dot{\theta}_c^2 + J_{cz}\dot{\psi}_c^2\right), \tag{4.3}$$

$$T_{bi} = 0.5\left(m_b \dot{y}_{bi}^2 + m_b \dot{z}_{bi}^2 + J_{bx}\dot{\phi}_{bi}^2 + J_{by}\dot{\theta}_{bi}^2 + J_{bz}\dot{\psi}_{bi}^2\right). \tag{4.4}$$

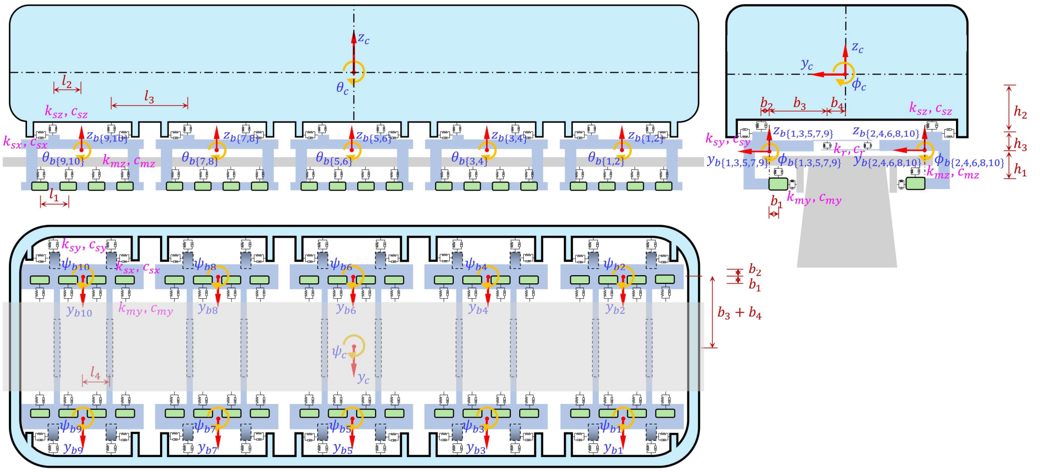


Figure 4.3. Constructed maglev vehicle model for transient aerodynamic load reconstruction.

The system's potential energy is composed of six components, corresponding to the variations in secondary suspensions in three directions (V_{sx}, V_{sy}, and V_{sz}), the primary electromagnetic forces in two directions (V_{py} and V_{pz}), and the anti-rolling bars (V_r). Therefore, the potential energy is given by:

$$V = V_{sx} + V_{sy} + V_{sz} + V_{my} + V_{mz} + V_r, \tag{4.5}$$

with each component of the system potential energy expressed as:

$$V_{sx} = 0.5k_{sx}\sum_{i=1}^{10}\sum_{j=1}^{2}[(\theta_c h_2 + \theta_{bi}h_3 + (-1)^{i-1}\psi_c(b_2 + b_3 + b_4) + (-1)^i\psi_{bi}b_2)^2], \tag{4.6}$$

$$V_{sy} = 0.5k_{sy}\sum_{i=1}^{10}\sum_{j=1}^{2}\left[\left(y_c + \phi_c h_2 + \psi_c p_{(2i+j-2)} - y_{bi} + \phi_{bi}h_3 + (-1)^{j-1}\psi_{bi}l_2\right)^2\right], \tag{4.7}$$

$$V_{sz} = 0.5k_{sz}\sum_{i=1}^{10}\sum_{j=1}^{2}\left[\left(z_c + (-1)^{i-1}\phi_c(b_2 + b_3 + b_4) - \theta_c p_{(2i+j-2)} - z_{bi} + (-1)^j\theta_{bi}l_2 + (-1)^i\phi_{bi}b_2\right)^2\right], \tag{4.8}$$

$$V_{my} = 0.5k_{my}\sum_{i=1}^{10}\sum_{j=1}^{4}[(y_{bi} + \phi_{bi}h_1 + (j-2.5)\psi_{bi}l_1)^2], \tag{4.9}$$

$$V_{mz} = 0.5k_{mz}\sum_{i=1}^{10}\sum_{j=1}^{4}[(z_{bi} + (-1)^i\phi_{bi}b_1 + (j-2.5)\theta_{bi}l_1)^2], \tag{4.10}$$

$$\begin{aligned} V_r = 0.5k_r\sum_{i=1}^{5}\sum_{j=1}^{2}\Big[&\big(z_{b(2i-1)} - z_{b(2i)} + \phi_{b(2i-1)}b_3 - \phi_{b(2i)}(b_3 + 2b_4) \\ &+ (-1)^{j-1}\theta_{b(2i-1)}l_2 + (-1)^j\theta_{b(2i)}l_2\big)^2 + \big(z_{b(2i-1)} - z_{b(2i)} \\ &+ \phi_{b(2i-1)}(b_3 + 2b_4) - \phi_{b(2i)}b_3 + (-1)^{j-1}\theta_{b(2i-1)}l_2 \\ &+ (-1)^j\theta_{b(2i)}l_2\big)^2\Big], \end{aligned} \tag{4.11}$$

where p_k in Eqs. (4.7) and (4.8) is of the form:

$$p_k = v_{1k}l_2 + v_{2k}l_3, k = 1, 2, \dots, 20\,, \tag{4.12}$$

with:

$$v_{1k} = \{5,3,5,3,3,1,3,1,1,-1,1,-1,-1,-3,-1,-3,-3,-5,-3,-5\}, \tag{4.13}$$

$$v_{2k} = \{2,2,2,2,1,1,1,1,0,0,0,0,-1,-1,-1,-1,-2,-2,-2,-2\}. \tag{4.14}$$

The Python symbolic mathematics toolbox *SymPy* is utilized to construct the expressions for T and V. The Lagrangian identity and its variations with respect to displacement and velocity follow the equation:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0, i = 1,2,\dots,55\,, \tag{4.15}$$

where q_i and $\dot{q}_i$ represent the displacement and velocity of the ith DOF, respectively. The stiffness and mass matrices $\mathbf{K}$ and $\mathbf{M}$ are subsequently obtained. For the damping matrix $\boldsymbol{\Upsilon}$, deriving reliable estimates for each component's damping coefficient is challenging. Therefore, it is constructed from the stiffness and mass matrices by incorporating modal damping factors identified from operational data, specifically using Rayleigh damping. This approach simplifies the damping model and balances the realism of the maglev vehicle model with the practical feasibility of model calibration. The original spring-damper model offers greater descriptive power, enabling the capture of non-proportional damping effects within the structure. However, the complexity of calibrating such models renders them less practical for operational use. Analysis of mode shapes from actual acceleration measurements suggests that the structure's damping behavior is predominantly proportional. Consequently, Rayleigh damping is adopted in the model, aligning well with these findings and simplifying the analysis without sacrificing significant accuracy.

Table 4.1 presents the nominal values of the main parameters, with those having corresponding scaling factors identified as more critical parameters requiring further validation. By specifying the parameters of the established maglev vehicle model, the vehicle's modal properties are determined. The natural frequencies and mode shapes can be determined by solving the eigenvalue and eigenvector problem, given by:

$$(\mathbf{M}^{-1}\mathbf{K})\boldsymbol{\phi}_i = \omega_i^2\boldsymbol{\phi}_i, \tag{4.16}$$

In this equation, ω_i is the ith angular natural frequency and $\boldsymbol{\phi}_i \in \mathbb{R}^n$ represents the ith mode shape. However, accelerometers are not exactly installed at the mass center of the rigid bodies, resulting in identified mode shapes from field-measured data deviating from the output $\boldsymbol{\phi}_i$. To formulate the deviation, a linear transformation matrix $\boldsymbol{\Psi}$ is constructed based on the relative positions between the accelerometers and the center of the vehicle body. This matrix allows mapping the theoretical mode shapes $\boldsymbol{\phi}_i$ to designated mode shapes $\widehat{\boldsymbol{\phi}}_i \in \mathbb{R}^{n_s}$, with n_s represents the number of accelerometers. This is expressed as:

$$\widehat{\boldsymbol{\phi}}_i = \boldsymbol{\Psi}\boldsymbol{\phi}_i. \tag{4.17}$$

The matrix $\boldsymbol{\Psi}$ is a sparse matrix comprising various geometric parameters related to the positions of accelerometers. For instance, mapping the theoretical mode shapes $\boldsymbol{\phi}_i$ to the vertical component of a sensor installed on the front of the car body in designated mode shapes $\widehat{\boldsymbol{\phi}}_i$ can be expressed by:

$$\hat{z}_c = z_c - \theta_c l_5 + \phi_c l_6. \tag{4.18}$$

Here, l_5 and l_6 are the distances of accelerometer positions to the mass center of car body. The other mappings can be formulated similarly, allowing $\boldsymbol{\Psi}$ to be fully determined.

Table 4.1. Nominal values of main parameters of the maglev vehicle model.

Parameter	Notation	Scaling factor	Nominal value	Unit
Vehicle body mass	m_c	-	2.83×10^4	kg
Bogie mass	m_b	-	9.60×10^2	kg
Magnetic support stiffness along *y*-axis	k_{m_y}	$\theta_{k_{my}}$	3.00×10^6	N/m
Magnetic support stiffness along *z*-axis	k_{m_z}	$\theta_{k_{mz}}$	2.00×10^5	N/m
Rotation stiffness	k_r	θ_{k_r}	2.00×10^6	N/m
Suspension support stiffness along *x*-axis	k_{s_x}	$\theta_{k_{sx}}$	3.00×10^6	N/m
Suspension support stiffness along *y*-axis	k_{s_y}	$\theta_{k_{sy}}$	7.00×10^4	N/m
Suspension support stiffness along *z*-axis	k_{s_z}	$\theta_{k_{sz}}$	3.00×10^4	N/m
Vehicle body's moment of inertia about *x*-axis	J_{cx}	-	3.26×10^4	kg·m
Vehicle body's moment of inertia about *y*-axis	J_{cy}	-	6.22×10^5	kg·m
Vehicle body's moment of inertia about *z*-axis	J_{cz}	-	6.32×10^5	kg·m
Bogie's moment of inertia about *x*-axis	J_{bx}	-	4.00×10^1	kg·m
Bogie's moment of inertia about *y*-axis	J_{by}	-	5.55×10^2	kg·m
Bogie's moment of inertia about *z*-axis	J_{bz}	-	5.57×10^2	kg·m

Beyond its modal properties, the maglev vehicle model can also be directly employed to predict responses given external excitations, such as aerodynamic loads (Wang et al., 2020b). The equation of motion is adopted for the inverse analysis of the relationship between external loads and responses at specified locations. This analysis is undertaken in the following subsections.

4.4. Model calibration based on operational data

To calibrate the established maglev model, operational modal analysis is first performed using acceleration data collected during normal train operation. Specifically, the frequency domain decomposition (FDD) method is adopted due to its straightforward implementation and computational efficiency in handling large datasets (Brincker et al., 2001). This method extracts information concerning the vehicle's natural frequencies and mode shapes.

Based on the FDD method, the acceleration functions $\boldsymbol{y}(t) \in \mathbb{R}^{n_s}$ are expressed in the modal expansion:

$$\boldsymbol{y}(t) = \widehat{\boldsymbol{\Phi}}\boldsymbol{p}(t), \tag{4.19}$$

where $\widehat{\boldsymbol{\Phi}}$ is the modal matrix derived from the designated mode shapes, and $\boldsymbol{p}(t)$ is the modal coordinate vector. The correlation of the response $\boldsymbol{y}(t)$ is given by:

$$\mathbf{R}_{yy}(\tau) = \mathbb{E}[\boldsymbol{y}(t)\boldsymbol{y}^{\mathrm{T}}(t+\tau)] = \widehat{\boldsymbol{\Phi}}\mathbf{R}_{pp}(\tau)\widehat{\boldsymbol{\Phi}}^{\mathrm{T}}, \tag{4.20}$$

where $\mathbf{R}_{pp}$ denotes the correlation matrix of the modal coordinate vector. Then, cross spectral density (CSD) matrix $\mathbf{G}_{YY}$ can be obtained by applying the Fourier transform to both sides of the above equation, that is:

$$\mathbf{G}_{YY}(\omega) = \widehat{\boldsymbol{\Phi}}\left[\int_{-\infty}^{+\infty} \mathbf{R}_{pp}(\tau)\,\mathrm{e}^{-\mathrm{i}\omega\tau}\mathrm{d}\tau\right]\widehat{\boldsymbol{\Phi}}^{\mathrm{T}} = \widehat{\boldsymbol{\Phi}}\mathbf{G}_{PP}(\omega)\widehat{\boldsymbol{\Phi}}^{\mathrm{T}}. \tag{4.21}$$

Since the modal coordinates are commonly considered uncorrelated, the off-diagonal elements $\mathbf{G}_{PP}(\omega)$ are set to zeros, making $\mathbf{G}_{PP}(\omega)$ a diagonal positive definite matrix. Additionally, as mode shapes can contain complex numbers, the transpose notation $^{\mathrm{T}}$

in the above equation is replaced by the Hermitian transpose notation †. Finally, the equation is rewritten as

$$\mathbf{G}_{YY}(\omega) = \widehat{\mathbf{\Phi}}[g_n^2(\omega)]\widehat{\mathbf{\Phi}}^\dagger, \tag{4.22}$$

where $g_n^2(\omega)$ is the nth diagonal element of $\mathbf{G}_{PP}(\omega)$ that represents the PSD of modal coordinates at frequency ω. The CSD matrix $\mathbf{G}_{YY}(\omega)$ can be constructed by computing the CSD function between any two random vibration signals using Welch's method. SVD is then applied to the calculated $\mathbf{G}_{YY}(\omega)$, resulting in:

$$\mathbf{G}_{YY}(\omega) = \widehat{\mathbf{\Phi}}'[s_n^2(\omega)]\widehat{\mathbf{\Phi}}'^{\mathrm{T}}, \tag{4.23}$$

where the singular values s_n^2 in the diagonal positions are the power spectral densities (PSD) of modal coordinates, and the singular vectors in $\widehat{\mathbf{\Phi}}'$ are interpreted as the mode shapes. Figure 4.4 illustrates the flowchart for computing natural frequencies and mode shapes using the FDD method. It is important to determine the rough frequency interval based on the peaks in the spectrums, ensuring that each interval includes only one natural frequency. Thus, the $\omega^\star$ producing the largest singular value within each interval is identified as the natural frequency, and the corresponding mode shape is obtained accordingly.

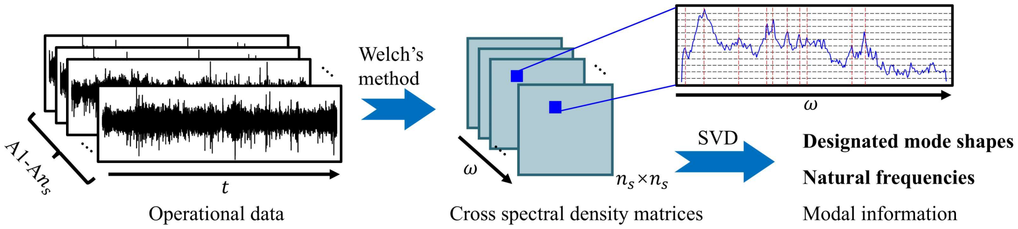


Figure 4.4. Flowchart of computing natural frequencies and designated mode shapes using FDD.

Multiple redundant peaks in the spectrums, primarily caused by elastic vibrations of different parts of the maglev vehicles not included in the constructed maglev vehicle model, may mislead mode identification. To address this, a meticulous analysis of the mode shapes is necessary to identify each detectable mode. Subsequent identification of the mode shapes enables discernment of the associated modal characteristics. Only modes with well-defined characteristics are selected to calibrate the maglev vehicle model. Using the well-identified natural frequencies, the parameters of the maglev vehicle model that are critical are adjusted to minimize the difference between the measured natural frequencies and the model-predicted natural frequencies. This optimization is expressed as:

$$\min\sum_{i=1}^{n}\frac{1}{\alpha_i}\left(\bar{\lambda}_i-\lambda_i(\boldsymbol{\theta})\right)^2, \tag{4.24}$$

where $\bar{\lambda}_i$ denotes the measured natural frequency of mode i, $\lambda_i(\boldsymbol{\theta})$ is the model-predicted natural frequency of mode i, α_i is a constant for normalization purposes, and $\boldsymbol{\theta}$ represents the scaling factor vector containing the updating parameters $\theta_{k_{my}}\theta_{k_{mz}}, \theta_{k_r}, \theta_{k_{sx}}, \theta_{k_{sy}}, \theta_{k_{sz}}$. These parameters are updated until the objective function is minimized. The small error confirmed through this calibration process attests to the model's capability to accurately replicate the dynamic behavior of the vehicle, thereby providing a solid foundation for subsequent investigations on accurately reconstructing transient aerodynamic loads.

4.5. Transient aerodynamic load reconstruction using MTGP

A well-calibrated maglev vehicle model enables the forward computation of vehicle responses to any given load history. Transient aerodynamic load reconstruction constitutes an inverse time-domain force reconstruction problem, where the aerodynamic loads on the vehicle are derived from direct pressure measurements and translated into forces acting on multiple rigid bodies. It is assumed that no aerodynamic loads are present on the bogies. Consequently, the problem is defined as follows: given specific acceleration measurements at certain locations of the maglev system during the vehicle's entry and exit from a double-track tunnel, the objective is to derive five aerodynamic load histories acting on the vehicle body. These five load histories correspond to the five DOF of the vehicle body.

Let the measured acceleration data at the *i*th time step represented by $\boldsymbol{y}_i \in \mathbb{R}^{n_s}$, for $i = 0,1, \dots, (n_t - 1)$. The concatenation of all $\boldsymbol{y}_i$ is denoted as $\boldsymbol{y} \in (\mathbb{R}^{n_s})^{n_t}$. Based on the time-domain force reconstruction theory illustrated in Chapter 3, the convolutional operator $\mathbf{H}$ is derived from calibrated system matrices. Consequently, the force histories represented by a concatenated vector $\boldsymbol{f} \in (\mathbb{R}^5)^{n_t}$, can be related to $\boldsymbol{y}$ through the convolutional operator $\mathbf{H}$, such that:

$$\boldsymbol{y} = \mathbf{H}\boldsymbol{f}. \tag{4.25}$$

To quantify the force histories based on the measured acceleration data, the MTGP algorithm is employed. This approach effectively calculates the posterior distributions of the five forces, providing insights into the uncertainties associated with the estimated forces.

4.6. Results

Field-measured acceleration data from a maglev line are used to validate the proposed framework. As shown in Figure 4.5, an on-board monitoring system with 25 uniaxial accelerometers is installed on the maglev vehicle to continually assess its dynamic performance. The measured locations include vertical and lateral vibrations of the vehicle body, the vertical vibrations of bogies 1–4, 6, 8, and 10, as well as lateral vibrations of bogies 3 and 4. For the field operational data, piezoelectric accelerometers with sensitivities of approximately 500 mV/g and measurement ranges of ±1g (accelerometers in the vehicle body) and ±10g (accelerometers in bogies) are used over a period of 180 seconds at a sampling frequency of 100 Hz. A sample of the stationary time-domain data for vehicle body and bogies' vertical vibrations is presented in Figure 4.6.

The operational data are analyzed using the FDD method. Figure 4.7 presents the computed PSD with frequencies ranging from 1/180 Hz to 50 Hz with an increment of 1/180 Hz. Numerous frequency intervals are determined for FDD based on the detected peaks in the PSD. Subsequently, the mode shapes and natural frequencies of different modal orders are calculated. The mode shapes are analyzed to exclude modes caused by the elasticity of the maglev vehicle. Table 4.2 lists nine successfully identified modes with well-defined characteristics. The natural frequencies of modes related to the motions of vehicle bodies range from 0.68 to 1.42 Hz, and those related to bogies range from 4.35 to 34.55 Hz. The damping ratio for each mode is estimated using the -3 dB points method, with the first two modes being approximately 3%. The damping matrix $\boldsymbol{\Upsilon}$ could be calculated from the linear combination of calibrated stiffness and mass matrices.

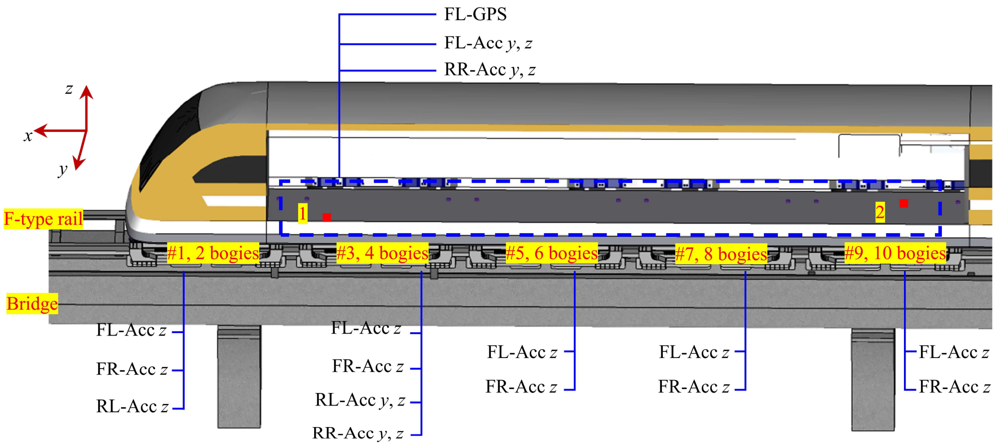


Figure 4.5. On-board monitoring system layout. F: Front, R: Rear, L: Left, R: Right, Acc: Accelerometer, Disp: Laser Displacement Sensor.

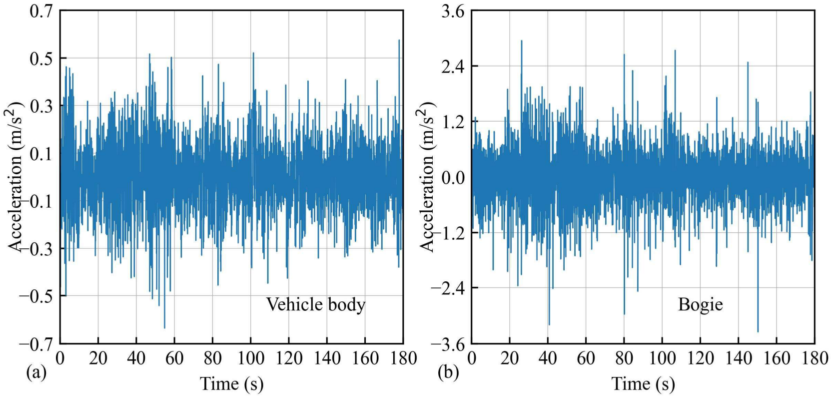


Figure 4.6. Measured acceleration data for operational modal analysis.

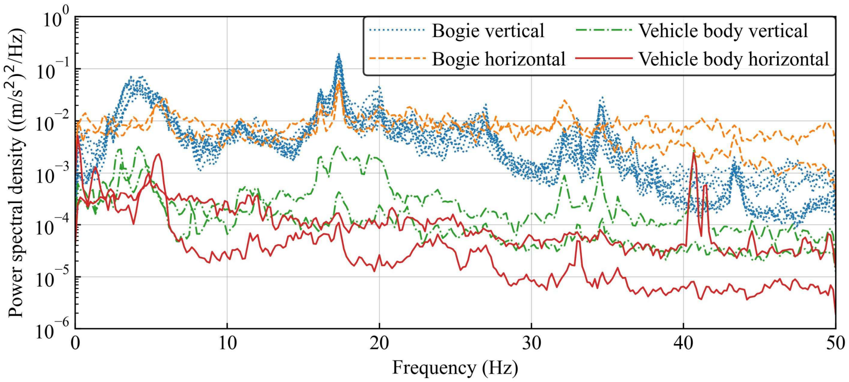


Figure 4.7. PSD from the measured operational data.

Table 4.2. Natural frequencies from operational modal analysis and maglev vehicle model.

Mode number	Natural frequency	Mode key feature description
1	0.68	Lateral movement of vehicle body
2	0.75	Vertical motion of vehicle body
3	1.39	Roll of vehicle body
4	1.42	Pitch of vehicle body
5	4.36	Vertical movement of bogie
6	5.41	Lateral movement of bogie
7	16.05	Roll of vehicle body
8	17.35	Pitch of bogie
9	34.55	Yaw of bogie

By applying the L-BFGS-B optimization algorithm to solve Eq. (4.24), the discrepancy between the model-predicted natural frequencies and the measured natural frequencies is minimized, as shown in Table 4.2. Figure 4.8(a) illustrates the evaluated parameters at each iteration. It is observed that 12 iteration steps are carried out, with convergence occurring at step 6. Five parameters, including $\theta_{k_{sz}}$, $\theta_{k_{sx}}$, $\theta_{k_{mz}}$, $\theta_{k_{my}}$, and $\theta_{k_{sr}}$ are found to be underestimated in the initial setup, while the parameter $\theta_{k_{sy}}$ is overestimated. The trajectory of the objective function's value is presented in Figure 4.8(b), which gradually converges to 0.06 within 6 steps, indicating that the discrepancy between the model-predicted natural frequencies and measured natural frequencies begins to level off. To further verify this, the natural frequencies obtained from the model and those derived via operational modal analysis are compared in Figure 4.9. The natural frequencies obtained from the measured acceleration data closely match those obtained by the calibrated model, confirming the accuracy of the calibration process.

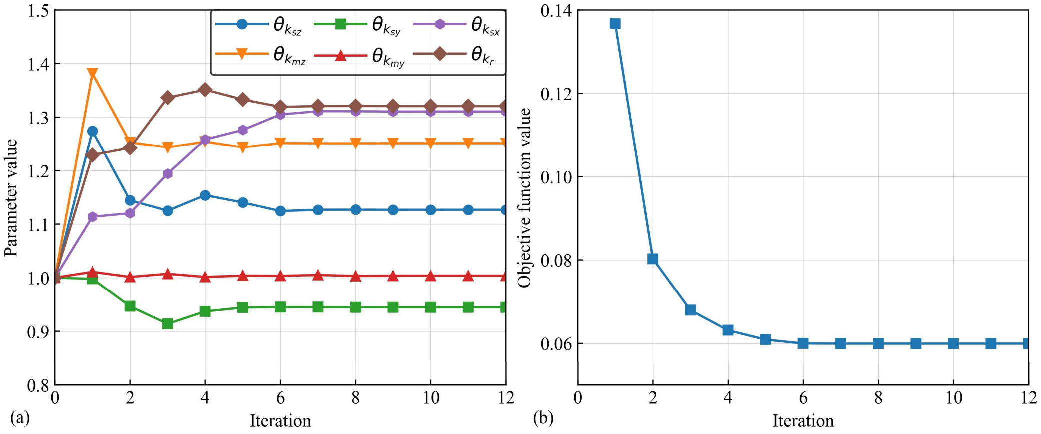


Figure 4.8. (a) Evolution of parameters at each iteration. (b) Natural frequency discrepancy obtained at each iteration.

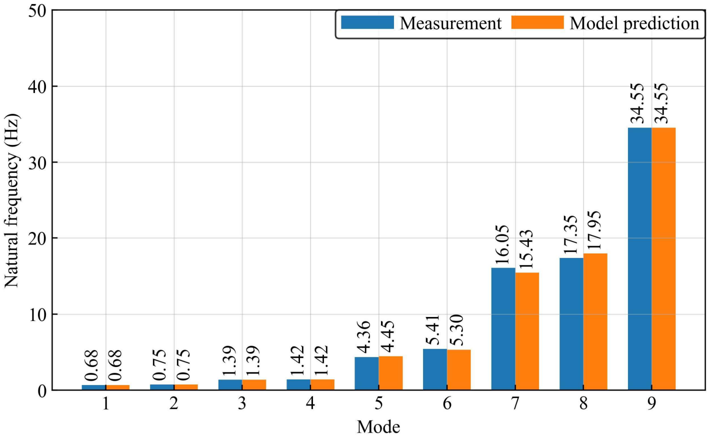


Figure 4.9. Bar chart for comparison of the updated model's natural frequencies and that from the operational data.

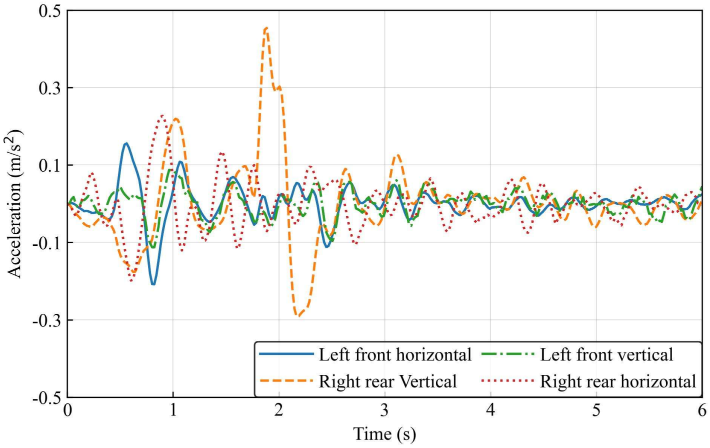


Figure 4.10. Utilized acceleration data for aerodynamic load reconstruction.

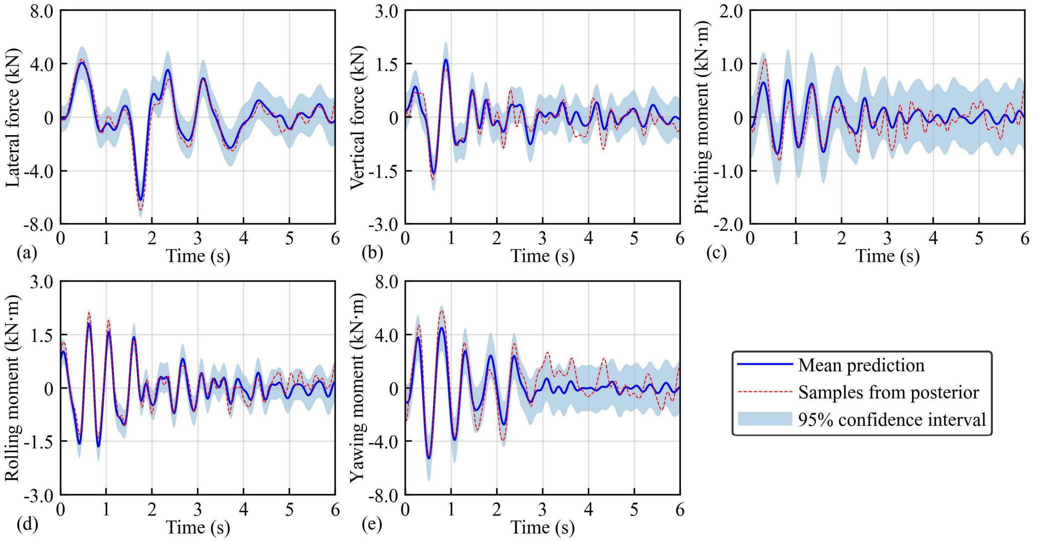


Figure 4.11. Predicted posteriors concerning lateral force, vertical force, pitching moment, rolling moment, and yawing moment by the proposed MTGP.

To corroborate the accuracy of aerodynamic load reconstructions, the technique flowchart shown in Figure 4.12 is proposed. Since the time-domain mathematical model, together with the MTGP algorithm, outputs the posterior of transient aerodynamic loads as multivariate Gaussians, samples of aerodynamic loads are drawn from the posterior distribution. The red lines in Figure 4.11 depict samples from the predicted posterior. These samples are then input into the calibrated maglev model to calculate the full system responses, from which acceleration responses at various accelerometer-installed positions are derived. Four measured accelerations are utilized to test the accuracy of aerodynamic load reconstruction. The concordance between model predictions and empirical measurements, as evidenced in Figure 4.13, provides strong validation for the reconstructed loads, affirming their reliability for practical applications.

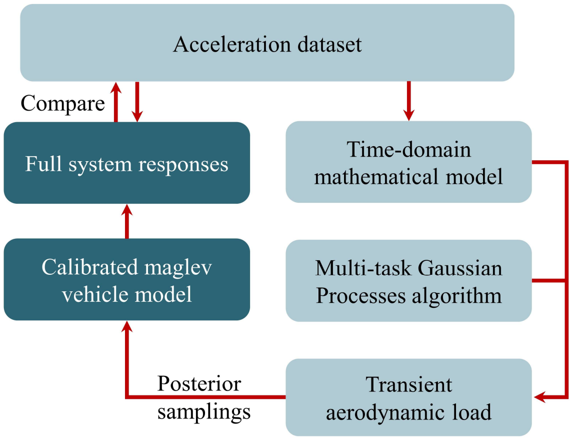


Figure 4.12. Corroboration of aerodynamic load reconstruction.

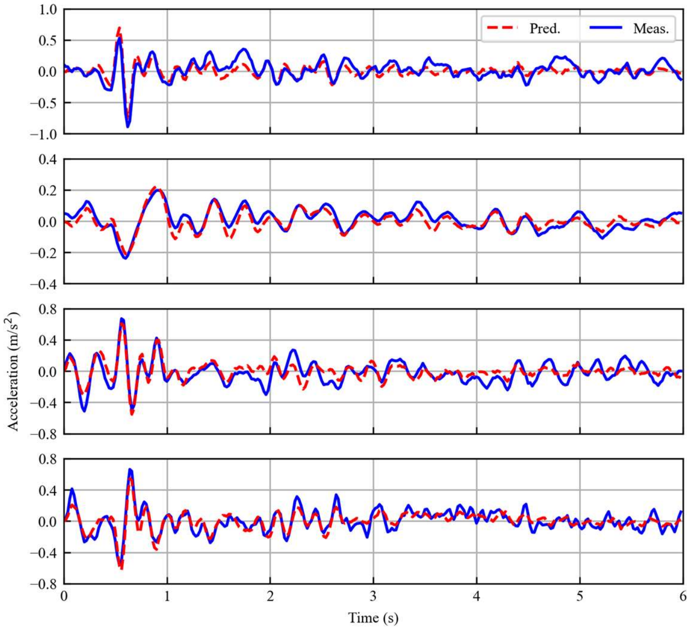


Figure 4.13. Comparison of predicted and measured accelerations.

4.7. Summary

This chapter introduces a novel framework designed to reconstruct transient aerodynamic loads on maglev vehicles using observed acceleration responses. The investigation leverages data from a maglev test line. Specifically, stationary random vibration data captured during maglev operation are utilized to calibrate the model. Subsequently, acceleration data recorded as the maglev enters and exits a double-track tunnel are analyzed. The proposed algorithms calculate the posterior distributions of these transient aerodynamic loads, providing robust validation. Remarkably, the results indicate that transient aerodynamic loads can be quantified through the vehicle's response to external excitation. This method represents a significant shift from traditional approaches, which rely heavily on direct measurements from pressure sensors or on simulation techniques. These findings suggest promising applications in field experiments, particularly in advancing high-speed train systems. Moreover, the MTGP method introduced in Chapter 3 not only estimates the posterior distribution of the reconstructed loads but also imparts confidence in the solutions derived from acceleration data. This method surpasses previous methodologies by providing analytical uncertainty propagation formulas, thereby eliminating the need for sampling-based techniques traditionally employed to determine the posterior and streamlining the load reconstruction process. However, for real-world applications involving force reconstruction, the elimination of sampling-based techniques introduces inherent costs and limitations that warrant careful consideration. The analytical formulation of the posterior distribution of forces relies on assumptions that may oversimplify scenarios where the underlying system dynamics or force characteristics deviate from idealized conditions. For instance, when random forces to be reconstructed are not suitable to be described by a GP, analytical framework may fail to capture the true uncertainty structure. This can lead to biased reconstructions or artificially narrow confidence intervals.

Chapter 5.

Time-delayed multi-fidelity Gaussian processes for state estimation of structural dynamic systems

5.1. Introduction

Predicting the structural state, including displacement and velocity at any given location, using limited sensor measurements is a critical inverse problem in many structural control and health monitoring systems (Papadimitriou et al., 2011; Song et al., 2022; Teymouri et al., 2022). This chapter introduces the theory of state estimation for structural dynamic systems employing a time-delayed GP approach. Additionally, it explores methods to enhance estimation accuracy by leveraging multi-fidelity data, which could be derived from both physics-based models and direct measurements. First, an effective GPR model is developed to incorporate sensor data from preceding time steps to estimate the current state. The training and test datasets are all generated from the FE model. Then, the adaptation of MFGP for mapping multi-fidelity time-delayed time series data to the structural system state is explored. To validate the proposed methodology, two examples of dynamic systems of structures are presented, including a generic mass-spring-damper dynamic system subject to stochastic excitations, and a 45-story tall building subject to wind excitations. The results indicate that the time-delayed MFGP method successfully manages the high-dimensional mapping of multi-fidelity time-delayed data to accurately estimate the system state.

5.2. Problem description of state estimation

Consider an n-DOF structural dynamic system formulated through spatial and temporal discretization of the governing differential equations:

$$g_1(\boldsymbol{x}_k, \boldsymbol{f}_k, \boldsymbol{x}_{k+1}; \boldsymbol{\theta}) = \mathbf{0}, \tag{5.1}$$

in which $\boldsymbol{x}_k \in \mathbb{R}^{2n}$ is the state vector at time instant k, containing the displacement and velocity of each DOF of the structure; $\boldsymbol{f}_k \in \mathbb{R}^{n_f}$ represents the external excitation vector; $\boldsymbol{\theta} \in \mathbb{R}^{n_p}$ denotes the parameters that completely define the system; n_f and n_p are the numbers of forces and parameters, respectively; $g_1: \mathbb{R}^{2n} \times \mathbb{R}^{n_f} \times \mathbb{R}^{2n} \to \mathbf{0} \in \mathbb{R}^{2n}$ is the function that governs the dynamical evolution of the state vector. The observation process of this n-DOF structure subjected to external excitation is of the form:

$$g_2(\boldsymbol{x}_k, \boldsymbol{f}_k, \boldsymbol{y}_k) = \mathbf{0}, \tag{5.2}$$

where $\boldsymbol{y}_k \in \mathbb{R}^{n_s}$ is the observation vector involves measurements of limited DOF, obtained from the sensor array installed on the structure; n_s denotes the number of sensors; $g_2: \mathbb{R}^{2n} \times \mathbb{R}^{n_f} \times \mathbb{R}^{n_s} \to \mathbf{0} \in \mathbb{R}^{n_s}$ represents the observation function.

In general, state estimation in structural dynamic systems typically involves developing a model that provides accurate estimates of state vectors $\boldsymbol{x}_k$, denoted as $\widetilde{\boldsymbol{x}}_k$, throughout a specified time interval $[0, T]$. This estimation is based on a dataset $\mathcal{D} = \{(t_k, \widehat{\boldsymbol{y}}_k)\}_{k=1}^{N}$, where $t_k \in [0, T]$ is the timestamp corresponds to the measurement $\widehat{\boldsymbol{y}}_k \in \mathbb{R}^{n_s}$ over structure, and T is the total duration of the estimation period. Generally, the measurement $\widehat{\boldsymbol{y}}_k$ corresponds to the observation vector $\boldsymbol{y}_k$ described in Eq. (5.2) and may include various physical quantities measurable in actual structures, such as accelerations, displacements, strains at specified DOF. In this chapter, the measurements are assumed to be accelerations, as their measurement techniques are the most mature and convenient to deploy.

5.3. Time-delayed GP for real-time state estimation

The time-delayed GP developed in this chapter utilize the measured data from preceding time steps as input to estimate the state of the current step. The methodology is formalized within the framework of reproducing kernel Hilbert spaces (RKHS), providing a rigorous mathematical foundation. Essentially, the time-delayed GP model is formulated within an RKHS $\mathcal{H}$, which is associated with a positive definite kernel function $k\colon \mathbb{R}^{(n_y+1)n_s} \times \mathbb{R}^{(n_y+1)n_s} \to \mathbb{R}$. Here, n_y denotes the number of preceding steps considered. The RKHS framework facilitates the modeling of complex temporal dependencies and correlations in the measurement data. For each ith component state vector $\boldsymbol{x}_k$. The time-delayed GP model is defined as:

$$\tilde{x}_k^i(\boldsymbol{y}) \sim \mathcal{GP}(0, k(\boldsymbol{y}, \boldsymbol{y}')) \,, \tag{5.3}$$

where $\tilde{x}_k^i(\boldsymbol{y})$ is the estimated ith component of the state vector at time step k; $\boldsymbol{y} = \left(\widehat{\boldsymbol{y}}_k, \widehat{\boldsymbol{y}}_{k-1}, \dots, \widehat{\boldsymbol{y}}_{k-n_y}\right) \in \mathbb{R}^{(n_y+1)n_s}$ is the concatenated measurement vector from the preceding n_y time steps and the current time step; $k(\boldsymbol{y}, \boldsymbol{y}')$ defines the inner product in the RKHS $\mathcal{H}$, capturing the similarity between different measurement vectors. The RKHS $\mathcal{H}$ leverages the reproducing property, which ensures that any function $f \in \mathcal{H}$ satifies:

$$f(\boldsymbol{y}) = \langle f, k(\boldsymbol{y}, \cdot) \rangle_{\mathcal{H}} \, , \tag{5.4}$$

where $\langle \cdot \rangle_{\mathcal{H}}$ denotes the inner product in the RHKS. This property is instrumental in efficiently mapping the high-dimensional measurement data into a feature space where linear relationships can be exploited for state estimation.

The training of time-delayed GP involves the optimization of covariance structure $k(\boldsymbol{y}, \boldsymbol{y}')$ from a dataset $\mathcal{B}$, which is defined as:

$$\mathcal{B} = \left\{ \left(t_k, \widehat{\boldsymbol{y}}_k, \widehat{\boldsymbol{y}}_{k-1}, \dots, \widehat{\boldsymbol{y}}_{k-n_y}, \boldsymbol{x}_k^i \right) \right\}_{k=n_y+1}^{N} . \tag{5.5}$$

During the training, the hyperparameters $\boldsymbol{\theta}$ of the kernel function k are optimized by maximizing the marginal likelihood of the observed data, i.e.,

$$\boldsymbol{\theta}^{\star} = \arg\max_{\boldsymbol{\theta}} \log p(\mathcal{B} | \boldsymbol{\theta}) \, . \tag{5.6}$$

This optimization ensures that the GP model accurately captures the relationship between historical measurements and the corresponding system states. In this chapter, the gradient based optimizer L-BFGS optimizer is adopted. Once trained, the time-delayed GP model could perform state estimation by conditioning on the input measurements from the preceding n_y time steps and dataset $\mathcal{B}$ to predict the current state $\widetilde{\boldsymbol{x}}^i$. The prediction not only provides an estimate of the ith state vector component, but also quantifies the uncertainty associated with each estimate. This probabilistic output, inherent to GP, is crucial for assessing the reliability and confidence of the state estimates, thereby enhancing the robustness and dependability of SHM and control systems.

5.4. Enhancing state estimation through MFGP

Figure 5.1 illustrates the comparison of the observation process g_2 of the structural dynamic system and time-delayed GP state estimation. In this context, state estimation can be regarded as an inverse of the observation process. However, two primary differences between the forward (observation) and inverse (state estimation) processes present significant challenges: First, there is no available external excitation data for state estimation. In many cases, the model often assumes that excitations are stationary Gaussian white noise, a simplification that may not accurately reflect real-world excitation patterns. This lack of direct excitation data complicates the accurate reconstruction of the system state from measurements alone. Second, the given structural dynamic system is mostly rooted in first principle, while the state estimation model originates from data. This necessitates the creation of comprehensive training and test datasets to optimize the GP model's parameters effectively. Moreover, the inverse problem must reconcile the deterministic nature of the forward model with the probabilistic framework of the GP model. Any modeling errors in the forward process can propagate and adversely affect the accuracy of the state estimation.

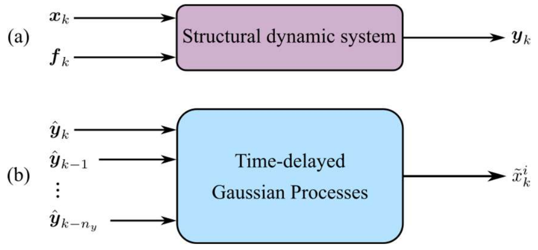


Figure 5.1. Comparison of forward process (observation) based on structural dynamic system and inverse process (state estimation) using time-delayed GP.

To address these challenges, the concept of multi-fidelity modeling is integrated into the time-delayed GP framework. Multi-fidelity modeling combines data from multiple sources of varying fidelity to enhance state estimation accuracy (Parussini et al., 2017). Figure 5.2 depicts the schematic of using MFGP to improve state estimation. This section considers two levels of fidelity: (a). High-fidelity structural dynamic systems: These include well-refined models or actual structures that closely represent the physical system. High-fidelity data can effectively overcome the aforementioned limitations by providing accurate training and test datasets. However, high-fidelity data is often limited due to the high costs associated with modeling and measurement. (b). Low-fidelity structural dynamic systems: These models are constructed based on certain simplifications, which may make the data less consistent with real-world structures but are more readily available and less expensive to obtain. Low-fidelity data can supplement high-fidelity data by providing a larger dataset for training the time-delayed GP model, thereby enhancing its generalization capabilities. For the sake of brevity, the time-delayed GP trained solely on low-fidelity data and high-fidelity data are respectively denoted as low-fidelity Gaussian process (LFGP) model and high-fidelity Gaussian process (HFGP) model, respectively, and the models trained from more than two levels of fidelities are called MFGP model.

The MFGP approach extends the single-fidelity GP framework by incorporating multiple data sources with varying degrees of fidelity (Gratiet and Garnier, 2014; Perdikaris et al., 2017; Cutajar et al., 2019). This integration allows the model to leverage both high-fidelity and low-fidelity data to enhance state estimation accuracy. Formally, let $\mathcal{B}_H$ and $\mathcal{B}_L$ represent the high-fidelity and low-fidelity datasets, respectively. The MFGP aims to learn the mapping from measurement inputs to state outputs by leveraging both datasets $\mathcal{B}_H$ and $\mathcal{B}_L$. To achieve this, a two-step generative model for the multi-fidelity outputs is assumed: First, sample from a GP:

$$u(\boldsymbol{y}) \sim \mathcal{GP}\big(0, k_u(\boldsymbol{y}, \boldsymbol{y}')\big), \tag{5.7}$$

where $u(\boldsymbol{y})$ is an underlying latent function capturing the shared structure between fidelities, and $k_u(\boldsymbol{y}, \boldsymbol{y}')$ is the covariance function. Second, obtain the high-fidelity and low-fidelity outputs through linear transformation of sampled $u_1(\boldsymbol{y})$:

$$\tilde{x}_{kH}^{i}(\boldsymbol{y}) = a_{H1}^{i} u_1(\boldsymbol{y}), \tag{5.8}$$

$$\tilde{x}_{kL}^{i}(\boldsymbol{y}) = a_{L1}^{i} u_1(\boldsymbol{y}). \tag{5.9}$$

where $\tilde{x}_{kH}^{i}(\boldsymbol{y})$ and $\tilde{x}_{kL}^{i}(\boldsymbol{y})$ represent the sampled high-fidelity state and low-fidelity state functions respectively; a_{H1}^{i} and a_{L1}^{i} are scalar coefficients that scale the latent function $u_1(\boldsymbol{y})$ for the high-fidelity and low-fidelity outputs. Therefore, as long as $u(\boldsymbol{y})$ is a GP, both the high-fidelity state estimation function $\tilde{x}_{kH}^{i}(\boldsymbol{y})$ and low-fidelity state estimation function $\tilde{x}_{kL}^{i}(\boldsymbol{y})$ would be GP as well. A multivariate Gaussian distribution among any function values of $\tilde{x}_{kH}^{i}(\boldsymbol{y})$ and $\tilde{x}_{kL}^{i}(\boldsymbol{y})$ can be constructed, and the mean vector and covariance matrix could be derived based on the relationship defined in Eqs. (5.7)-(5.9).

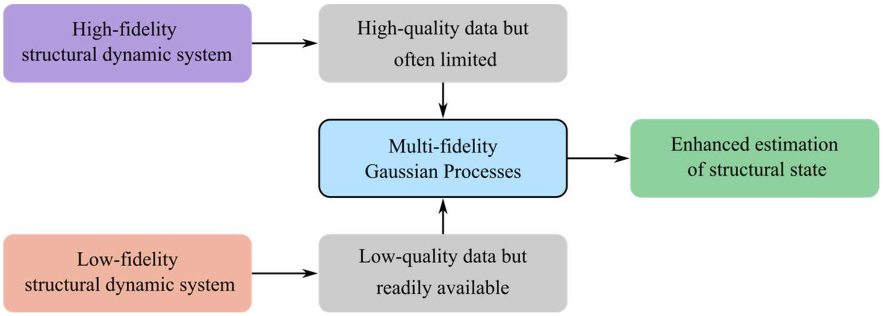


Figure 5.2. Overview of enhancing state estimation accuracy through MFGP.

To formalize the relationship between high-fidelity and low-fidelity data, the intrinsic coregionalization model with the GP framework is introduced in this chapter. Coregionalization allows the modeling of multiple outputs by capturing the interdependencies between different fidelity levels through a shared covariance structure. The combined MFGP kernel k_{MFGP} is constructed as the product of shared input kernel $k_u(\boldsymbol{y}, \boldsymbol{y}')$ and a coregionalization kernel $k_c(f, f')$:

$$k_{\text{MFGP}}\big((\boldsymbol{y}, f), (\boldsymbol{y}', f')\big) = k_u(\boldsymbol{y}, \boldsymbol{y}')k_c(f, f'), \tag{5.10}$$

where $\boldsymbol{y}, \boldsymbol{y}' \in \mathbb{R}^{(n_y+1)n_s}$ are the measurement input vectors; $f, f' \in \{H, L\}$ denote the fidelity levels; $k_c(f, f')$ captures the covariance between different fidelity levels. In this chapter, for a two-fidelity scenario, the coregionalization kernel $k_c(f, f')$ can be derived based on the definition of covariance between any two function values from $\tilde{x}_{kH}^i(\boldsymbol{y})$ and $\tilde{x}_{kL}^i(\boldsymbol{y})$, which is given by:

$$k_c(f, f') = \begin{cases} {a_{H1}^i}^2 & \text{if } f = f' = H \\ a_{H1}^i a_{L1}^i & \text{if } f \neq f' \\ {a_{L1}^i}^2 & \text{if } f = f' = L \end{cases}. \tag{5.11}$$

To further enhance the performance of the MFGP state estimation model, the sampling step can be iterated multiple times. In this chapter, while only two fidelity levels are considered, the GP $u(\boldsymbol{y})$ are sampled twice to obtain independent latent functions $u_1(\boldsymbol{y})$ and $u_2(\boldsymbol{y})$. The high-fidelity and low-fidelity outputs are then generated through linear transformations of these latent functions as follows:

$$\tilde{x}_{kH}^i(\boldsymbol{y}) = a_{H1}^i u_1(\boldsymbol{y}) + a_{H2}^i u_2(\boldsymbol{y}), \tag{5.12}$$

$$\tilde{x}_{kL}^i(\boldsymbol{y}) = a_{L1}^i u_1(\boldsymbol{y}) + a_{L2}^i u_2(\boldsymbol{y}). \tag{5.13}$$

where a_{H1}^i, a_{H2}^i, a_{L1}^i ,and a_{L2}^i are the scalar coefficients for latent functions. By introducing multiple latent functions $u_1(\boldsymbol{y})$ and $u_2(\boldsymbol{y})$, the MFGP model can capture more complex and nuanced relationships between the high-fidelity and low-fidelity data sources. This multi-sample approach allows the model to account for non-linear dependencies and variability that a single latent function might miss, thereby improving the fidelity and accuracy of state estimations.

Under these conditions, the MFGP kernel k_{MFGP} would follow the form as illustrated in Eq. (5.10), while the coregionalization kernel $k_c(f,f')$ is reformulated to accommodate the more intricate relationships between state estimation functions and latent functions. Specifically, for a two-fidelity scenario, the coregionalization kernel $k_c(f,f')$ is defined as:

$$k_c(f,f') = \begin{cases} {a_{H1}^{i}}^2 + {a_{H2}^{i}}^2 & \text{if } f = f' = H \\ a_{H1}^{i} a_{L1}^{i} + a_{H2}^{i} a_{L2}^{i} & \text{if } f \neq f' \\ {a_{L1}^{i}}^2 + {a_{L2}^{i}}^2 & \text{if } f = f' = L \end{cases} . \tag{5.14}$$

Based on the number of samplings the generative model, the initial model defined by Eqs. (5.8) and (5.9) is referred to as the level-one model, while the extended model defined by Eqs. (5.12) and (5.13) is termed the level-two model. For both models, the corresponding multi-fidelity kernel structure k_{MFGP} is composed of the same shared kernel $k_u(\boldsymbol{y},\boldsymbol{y}')$ and the coregionalization kernel $k_c(f,f')$. In the following examples, all shared kernel $k_u(\boldsymbol{y},\boldsymbol{y}')$ in the following examples are determined to be the squared exponential kernel for the sake of broad applicability and simplicity. While this choice ensures generality, it may not always be optimal for every application. More sophisticated methods, such as kernel flows, could be employed to tailor the kernel for specific cases, but such optimization falls outside the scope of this work. . Further discussion on kernel selection is therefore omitted. Regarding the coregionalization kernel $k_c(f,f')$, there are two hyperparameters a_{H1}^{i} and a_{L1}^{i} for level-one model, and four hyperparameters a_{H1}^{i}, a_{H2}^{i}, a_{L1}^{i},and a_{L2}^{i} for level-two model. Although utilizing multiple latent functions with coregionalization allows the MFGP model to scale to more complex systems without a proportional increase in computational complexity, the level-two model does not necessarily perform better than the level-one model in practice, because the practical performance of the two different MFGP depends on how closely their idealized assumptions match the actual system. It remains necessary to determine which model to use based on cross validation and Bayesian model evidence comparison.

The training of the MFGP model involves optimizing hyperparameters of both shared kernel and coregionalization kernel. This optimization is still based on maximizing the joint marginal likelihood of the combined high-fidelity and low-fidelity datasets $\mathcal{B}_H$ and $\mathcal{B}_L$:

$$\boldsymbol{\theta}^{\star} = \arg\max_{\boldsymbol{\theta}} \log p(\mathcal{B}_H, \mathcal{B}_L | \boldsymbol{\theta}) \, . \tag{5.15}$$

where $\boldsymbol{\theta}$ encompasses all hyperparameters in k_{MFGP}. Ultimately, a well-trained MFGP model can leverage the advantages of both high-fidelity and low-fidelity data to provide enhanced state estimations based on the input measurements.

5.5. Numerical example 1: a mass-spring-damper dynamic system

5.5.1. State estimation for mass-spring-damper dynamic system

In this section, a linear 5-DOF mass-spring-damper dynamic system is firstly used to demonstrate the effectiveness of the proposed time-delayed GP on state estimation. As depicted in Figure 5.3, the system comprises five masses interconnected by linear springs and viscous dampers. The specific parameters of this system, including mass values, spring constants, and damping coefficients, are presented in Table 5.1. The system exhibits five natural frequencies at 0.64, 1.88, 2.86, 4.14, and 4.81 Hz, with corresponding damping ratios of 0.54%, 1.63%, 2.68%, 2.86%, and 3.61%, respectively.

A random excitation $f(t)$ is applied on m_1, and its time-domain representation along with PSD is shown in Figure 5.4. The PSD reveals that the excitation energy is predominantly concentrated between 0.1 and 10 Hz, encompassing the five natural frequencies of the mass-spring-damper system. Consequently, the complete dynamic responses, including displacements, velocities, and accelerations, of each DOF are obtained by numerically solving the equations of motion using an implicit Runge-Kutta method.

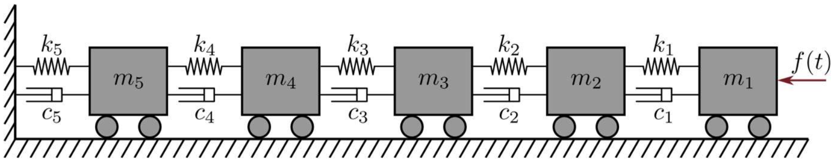


Figure 5.3. A linear 5-DOF mass-spring-damper dynamic system.

Table 5.1. Parameters for the mass-spring-damper dynamic system.

Symbol	Value	Unit	Symbol	Value	Unit	Symbol	Value	Unit
m_1	12	kg	c_1	6	Ns/m	k_1	2×10^3	N/m
m_2	14	kg	c_2	6	Ns/m	k_2	3×10^3	N/m
m_3	8	kg	c_3	6	Ns/m	k_3	2×10^3	N/m
m_4	12	kg	c_4	6	Ns/m	k_4	3×10^3	N/m
m_5	10	kg	c_5	6	Ns/m	k_5	2×10^3	N/m

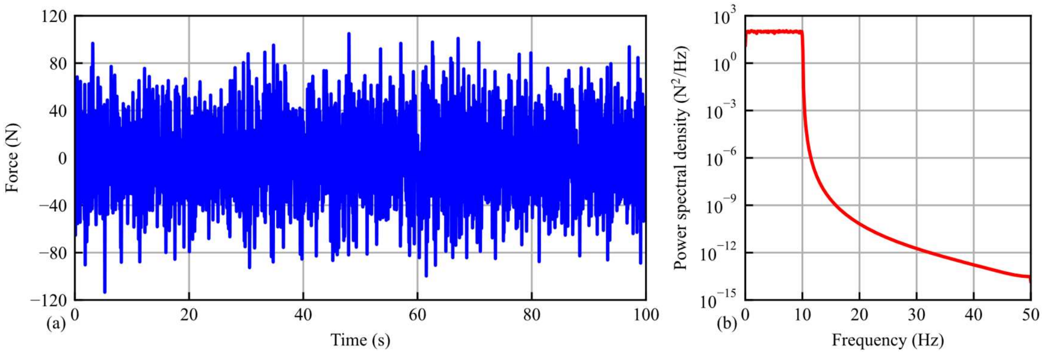


Figure 5.4. (a) Time-domain representation of the random excitation $f(t)$; (b) PSD of the random excitation.

In this example, it is assumed that only the acceleration of mass m_1 is measured. This acceleration data serves as the primary input for training and validating the proposed time-delayed GP model for state estimation. By utilizing the measured acceleration from the preceding n_y time steps, the time-delayed GP model estimates the current state vector, which includes the displacement and velocity of each mass in the system.

Figure 5.5 shows the partial responses of the mass-spring-damper dynamic system subject to the random excitation $f(t)$. For demonstration purposes, the first 24 s of the colored data are used to train the time-delayed GP model, while the remaining 76 s are reserved for testing. Before delving into the specifics of the time-delayed model, Figure 5.6, based on the data from Figure 5.5, depicts the relationships among the measured acceleration, displacement, and velocity. The plot reveals that there is almost a negligible correlation between the current acceleration measurement and the system state (displacement and velocity) at the same time step. This finding emphasizes the importance of incorporating information from preceding measurements when predicting the state at the current step.

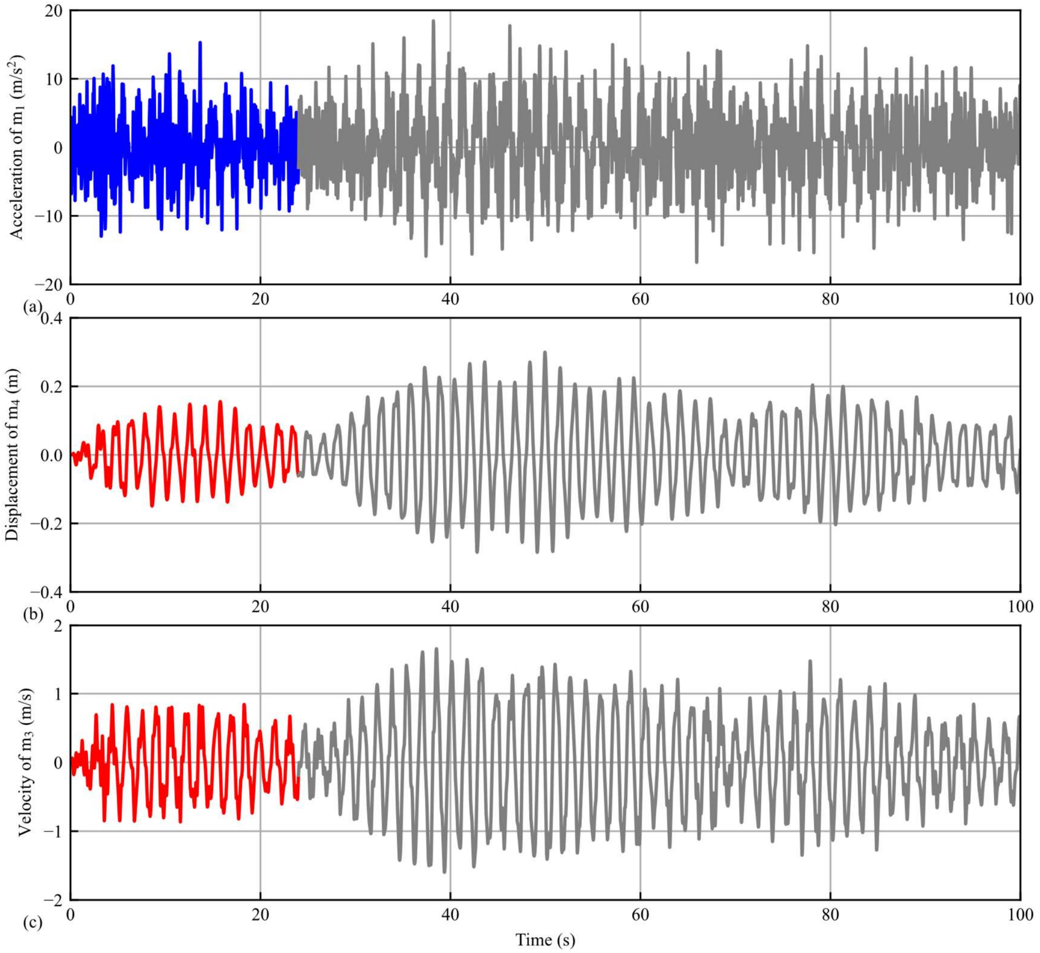


Figure 5.5. Mass-spring-damper dynamic system response subject to the random excitation: (a). acceleration of m_1, (b) displacement of m_4, and (c) velocity of m_3.

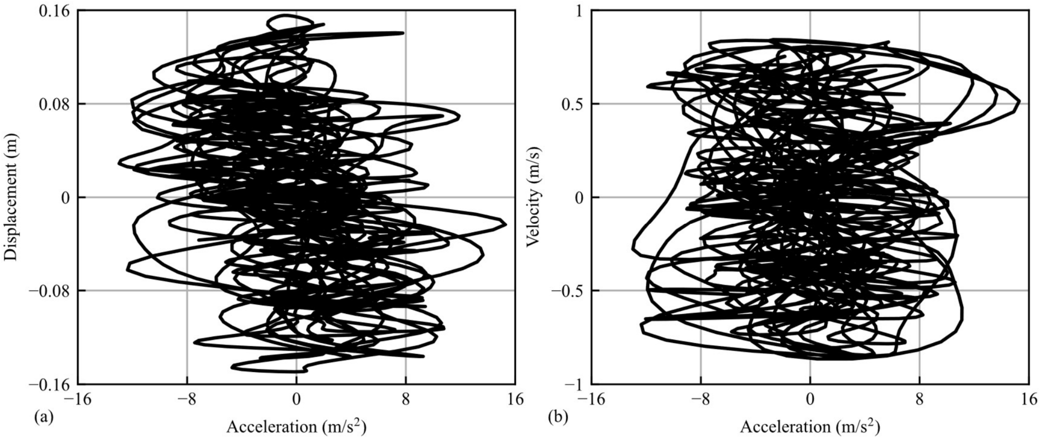


Figure 5.6. (a) Displacement-acceleration variation during the random excitation process. (b) Velocity-acceleration variation during the random excitation process.

Using the data presented in Figure 5.5, time-delayed GP models for the displacement and velocity prediction are developed. The delay step is initially set to 30, and the training and test datasets are correspondingly generated according to Eq. (5.5). The GP models are trained by maximizing the marginal likelihood as described in Eq. (5.6), with the hyperparameters of the squared exponential kernel optimized to best fit the training dataset. The training process takes 4.36 s and 5.59 s for the displacement and velocity models, respectively.

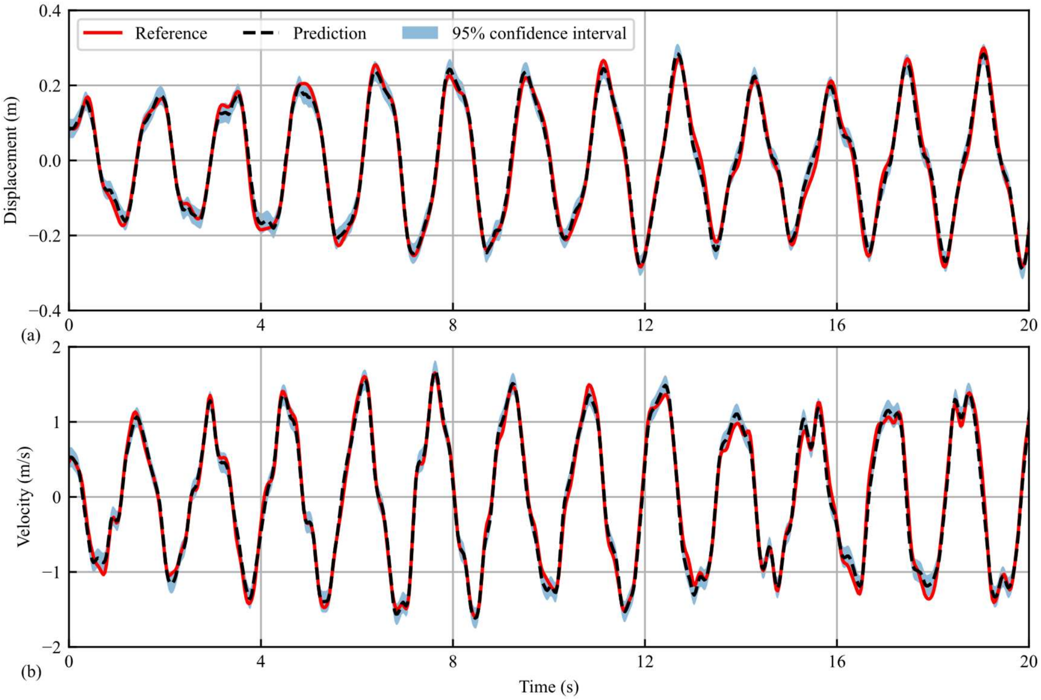


Figure 5.7. Comparison of time-delayed GP model predictions with reference: (a) displacement of m_4, (b) velocity of m_3.

The well-trained GP models are validated using the test dataset. Figure 5.7 compares the predictions from the time-delayed GP model with reference data for displacement and velocity over the time interval from 20 s to 44 s, as shown in Figure 5.5. The results demonstrate that GP can efficiently and effectively learn the high-dimensional mapping from the time-delayed data series to the system state. Furthermore, the time-delayed model successfully performs accurate state estimation.

To investigate the performance of the proposed time-delayed GP approach across various delay step settings, two matrices are employed: NRMSE and predictive log likelihood (PLL). These matrices are defined as follows:

$$\mathrm{NRMSE} = \frac{\sqrt{\sum_{k=n_y+1}^{N}\left(x_k^i - \mu_k\right)^2}}{\sqrt{\sum_{k=n_y+1}^{N}\left(x_k^i - \bar{x}^i\right)^2}}, \tag{5.16}$$

$$\mathrm{PLL} = \sum_{k=n_y+1}^{N}\left(-\frac{1}{2}\log(2\pi\sigma_k^2) - \frac{\left(x_k^i - \mu_k\right)^2}{2\sigma_k^2}\right). \tag{5.17}$$

in which x_k^i is the reference value of the ith component of the state vector at time step k; μ_k is predicted mean of state at time step k from the GP model; $\bar{x}^i$ is the mean of ith component of the state vector from time steps $n_y + 1$ to N; σ_k is the predicted standard deviation of state at time step k from the GP model. Overall, the NRMSE metric assesses the accuracy of the predicted mean values. It is calculated by taking the square root of the sum of squared differences between the reference values and the predicted means, normalized by the square root of the sum of squared differences between the reference values and their mean. Mathematically, a lower NRMSE indicates higher prediction accuracy. The PLL metric quantifies how well the GP model's posterior distributions explain the reference data. It is defined as the sum of the log probabilities of each individual reference value under their corresponding predictive posterior distributions. A higher PLL value signifies a better prediction of the posterior distribution by the time-delayed GP model.

By setting the delay step from 1 to 40, the performance of the proposed time-delayed GP is systematically evaluated using the NRMSE and PLL. For each delay step d within the range, the models incorporate d previous time steps of acceleration to predict the current displacement and velocity. The training the testing processes are repeated for each d, enabling a comparative analysis of how different delay steps influence prediction accuracy and probabilistic forecasting.

The results, depicted in Figure 5.8, illustrate the variation of NRMSE and PLL across different delay steps when predicting the displacement of m_4 and velocity of m_3. It is observed that the delay step increases from 1 to an optimal value (25 for displacement and 20 for velocity), the NRMSE decreases, indicating improved prediction accuracy. This improvement can be attributed to the model's enhanced ability to capture temporal dependencies and incorporate relevant historical information, thereby enriching the feature set used for making predictions. Concurrently, the PLL increases, signifying that the model's predictive distributions are better aligned with the observed data, offering more reliable uncertainty quantification.

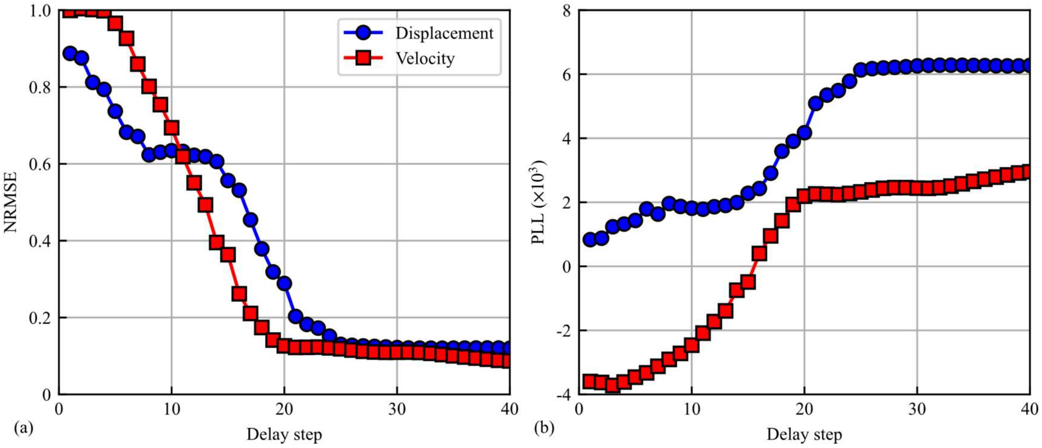


Figure 5.8. NRMSE and PLL for prediction: (a) displacement of m_4, (b) velocity of m_3.

However, beyond the optimal delay overlap, further increases in d lead to a slight gradual decrease in NRMSE and slight increase in PLL, indicating a convergence trend. This suggests that measurements sampled significantly earlier than the optimal delay contribute marginally to the prediction accuracy and uncertainty quantification. Conversely, very small delay steps fail to provide sufficient historical context, limiting the model's ability to capture underlying patterns and dependencies. This insufficiency leads to higher NRMSE values and lower PLL scores, underscoring the importance of selecting an appropriate delay step to balance the capture of essential temporal dynamics while avoiding the incorporation of superfluous data.

5.5.2. Enhanced state estimation for nonlinear mass-spring-damper dynamic system

To validate the effectiveness of time-delayed MFGP model in enhancing state estimation accuracy, a nonlinear 5-DOF mass-spring-damper system, depicted in Figure 5.9, is considered. This system represents the high-fidelity model, while the simplified model presented in Figure 5.1 serves as its low-fidelity approximation.

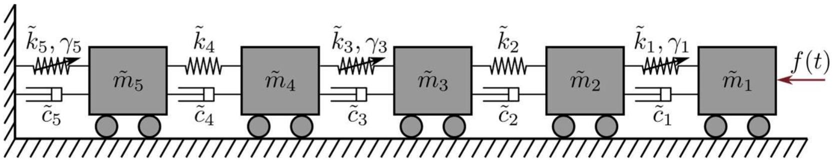


Figure 5.9. A nonlinear 5-DOF mass-spring-damper dynamic system.

Table 5.2. Parameters for the nonlinear mass-spring-damper dynamic system.

Symbol	Value	Unit	Symbol	Value	Unit	Symbol	Value	Unit
$\tilde{m}_1$	13.19	kg	$\tilde{c}_1$	5.44	Ns/m	$\tilde{k}_1$	1.91×10^3	N/m
$\tilde{m}_2$	13.60	kg	$\tilde{c}_2$	5.44	Ns/m	$\tilde{k}_2$	3.95×10^3	N/m
$\tilde{m}_3$	9.04	kg	$\tilde{c}_3$	6.29	Ns/m	$\tilde{k}_3$	2.31×10^3	N/m
$\tilde{m}_4$	15.66	kg	$\tilde{c}_4$	3.70	Ns/m	$\tilde{k}_4$	2.72×10^3	N/m
$\tilde{m}_5$	9.53	kg	$\tilde{c}_5$	3.93	Ns/m	$\tilde{k}_5$	2.22×10^3	N/m
γ_1	2×10^5	N/m^2	γ_3	2×10^5	N/m^2	γ_5	2×10^5	N/m^2

The nonlinear 5-DOF mass-spring-damper system comprises five masses interconnected by springs and dampers, with nonlinear characteristics introduced through spring nonlinearities. Specifically, the elastic force for each spring is modeled using a second-order polynomial function characterized by parameters $\tilde{k}_i$ and γ_i for linear and quadratic term respectively, with $i = 1,3,5$. Additionally, the system parameters for the nonlinear high-fidelity model are detailed in Table 5.2. Beyond the distinctions in spring properties, differences in damping coefficients and mass values also exist between the high-fidelity and low-fidelity models.

A random excitation with similar statistical properties in the frequency domain, as described in the previous subsection, is applied to $\widetilde{m}_1$. The measured acceleration data from $\widetilde{m}_1$ is utilized for the state estimation of the nonlinear system. Initially, the time-delayed GP models trained using the low-fidelity dataset $\mathcal{B}_L$ are employed to estimate the state of the high-fidelity system. Using the measured acceleration data from $\widetilde{m}_1$ as input, the displacement of $\widetilde{m}_4$ and velocity of $\widetilde{m}_3$ are predicted. The prediction results over a 10 s interval are presented in Figure 5.10(a) and (b). These results demonstrate that the performance of the LFGP models is significantly inferior when applied to the nonlinear high-fidelity dynamic system, highlighting the limitations of relying solely on low-fidelity data for accurate state estimation in complex dynamics.

Subsequently, a high-fidelity dataset $\mathcal{B}_H$, comprising 20% of the data volume of the low-fidelity dataset $\mathcal{B}_L$, is generated. Time-delayed GP models are trained exclusively on $\mathcal{B}_H$, and the corresponding prediction outcomes over the test dataset are illustrated in Figure 5.10(c) and (d). It is observed that HFGP models exhibit improved prediction performance compared to LFGP models, despite the smaller training data volume.

Specifically, for displacement estimation, the HFGP models achieve NRMSE of 0.46 and a PLL of 4619, compared to the LFGP models' NRMSE of 0.66 and PLL of -1123. Similarly, for velocity estimation, the HFGP models attain an NRMSE of 0.42 and a PLL of 256, outperforming the LFGP models' NRMSE of 0.62 and PLL of -8388. Additionally, uncertainty quantification is notably enhanced in the HFGP models, providing more reliable uncertainty estimations.

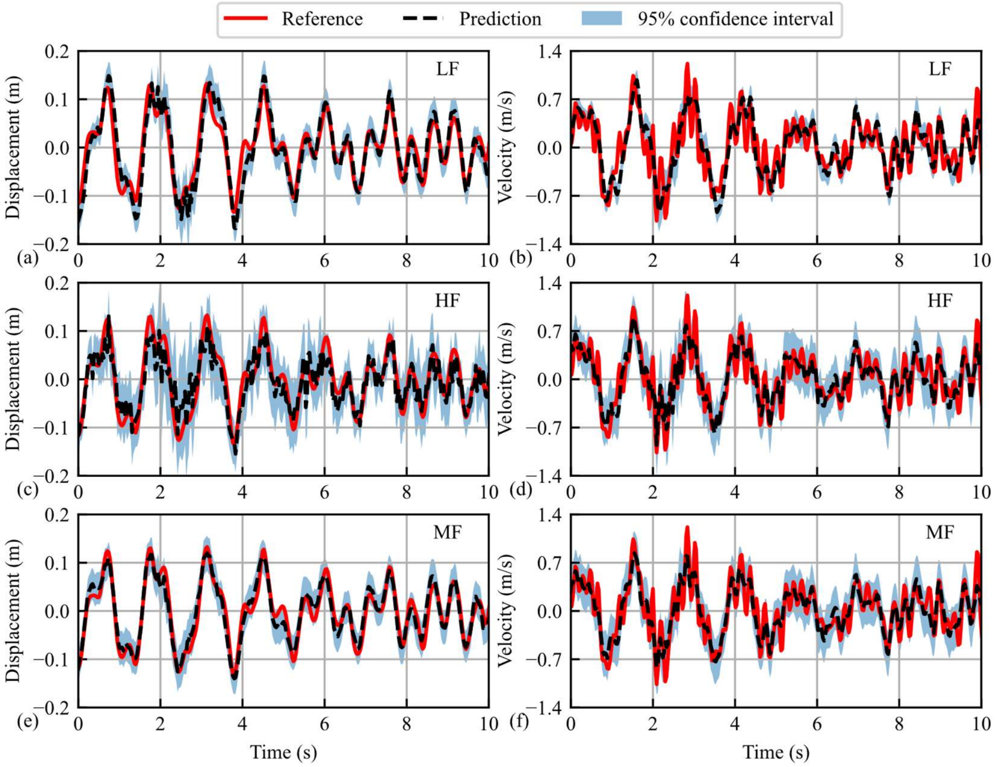


Figure 5.10. Prediction of displacement of $\tilde{m}_4$ and velocity of $\tilde{m}_3$ using low-fidelity, high-fidelity and multi-fidelity data trained time-delayed GP models.

To capitalize on the strengths of both low-fidelity and high-fidelity datasets, a time-delayed MFGP model is employed, integrating the extensive Low-fidelity data with the limited but more precise high-fidelity data. In this study, all multi-fidelity results are derived from a level 2 multi-fidelity model. The prediction results using the MFGP model are presented in Figure 5.10(e) and (f). These results exhibit superior prediction accuracy compared to both the low-fidelity-only and high-fidelity-only models. Specifically, the MFGP model achieves an NRMSE of 0.24 for displacement and 0.39 for velocity, outperforming the high-fidelity models' NRMSE of 0.46 and 0.42, respectively. In terms of PLL, the MFGP model attains values of 5839 for displacement and 343 for velocity, surpassing the high-fidelity models' PLL of 4619 and 256, and significantly outperforming the low-fidelity models' PLL of -1123 and -8388. The posterior distributions generated by the MFGP model more accurately align with the reference data compared to the low-fidelity-only model, and the confidence intervals are narrower compared to the high-fidelity-only model, rendering the predictions more informative and reliable.

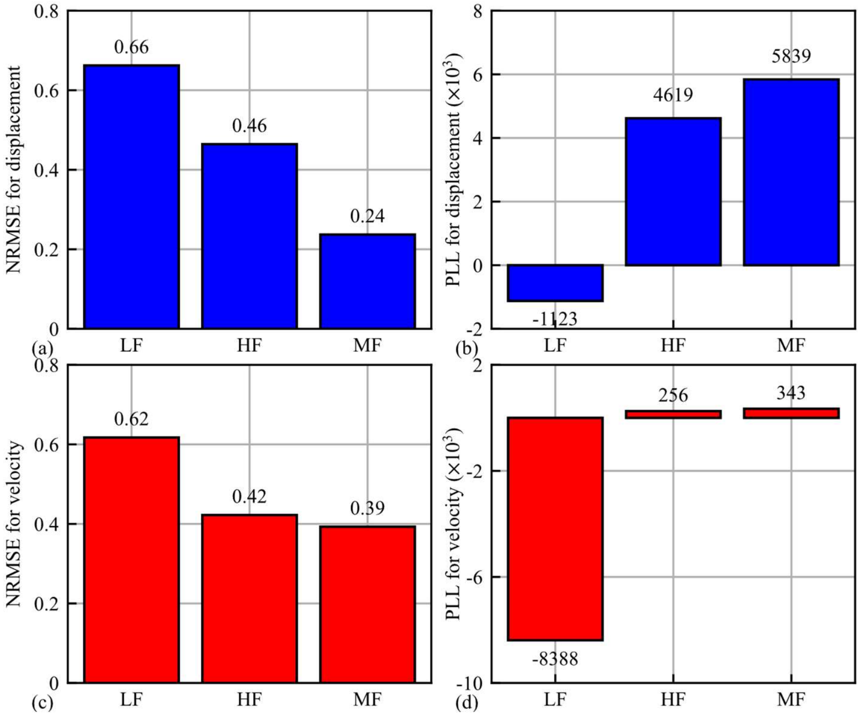


Figure 5.11. Quantitative comparison of state estimations by LFGP, HFGP, and MFGP models.

Quantitative evaluation results of the three models' performance are concluded and presented in Figure 5.11. In this example, the MFGP model not only enhances prediction accuracy but also provides robust uncertainty quantification. The MFGP model effectively leverages the comprehensive coverage of the low-fidelity dataset and the precision of the high-fidelity dataset, resulting in state estimations that are both accurate and trustworthy. This comprehensive performance improvement underscores the efficacy of the multi-fidelity approach in overcoming the limitations inherent in single-fidelity models, particularly in complex nonlinear dynamic systems.

5.6. Numerical example 2: a 45-story tall building

5.6.1. Building structure and health monitoring system description

The building structure selected for this example is a steel-framed tower comprising 45 stories, representing a typical high-rise construction in urban environments. The detailed vertical sections of the building's framing are depicted in Figure 5.12. The tower's structural steel framing system consists of moment-resisting frames in the east-west direction and a combination of eccentrically braced frames and moment resisting frames in the north south direction. The eccentrically braced frames are traditional moment-resisting frames with diagonal bracing in the inner two bays, and therefore could enhance the stiffness of the building in the north-south direction. There are four plane views of the tower as shown in Figure 5.13, each representing a different level from the base to the top. Figure 5.13(a) displays the ground floor layout, illustrating the foundational steel grid that supports the entire structure. As moving upward through Figure 5.13(b), (c), and (d), each subsequent plan view details the progressive arrangement and reinforcement of steel beams and columns, reflecting the building's scalability and adaptability to various load conditions.

A detailed FE model is constructed for the 45-story steel-framed tower using ANSYS APDL, comprising a total of 2,360 nodes and 4,552 elements. This comprehensive FE model enables accurate simulation of the building's structural behavior under various loading conditions, serving as the foundation for subsequent dynamic analysis and state estimation studies. In this numerical example, the FE model effectively represents the "real-world" structure, and the measurements derived from this model are utilized to validate the proposed approaches discussed in this chapter.

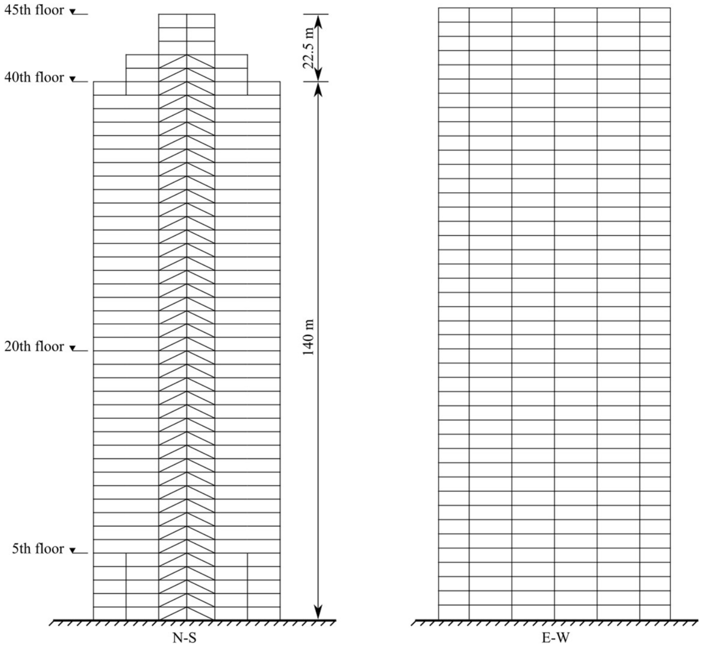


Figure 5.12. Vertical sections of the 45-story building framing.

To obtain measurements across the building structure, an embedded SHM system is required. Figure 5.14 illustrates the details of the SHM system, which includes the installation of eight accelerometers and radar displacement measurements conducted at four different levels to capture dynamic displacements. Measured accelerations are used as inputs for the time-delayed GP model, while the measured displacements serve as references to validate the model's predictions. Additionally, the coordinates shown in Figure 5.14 are used for all illustrations of this example.

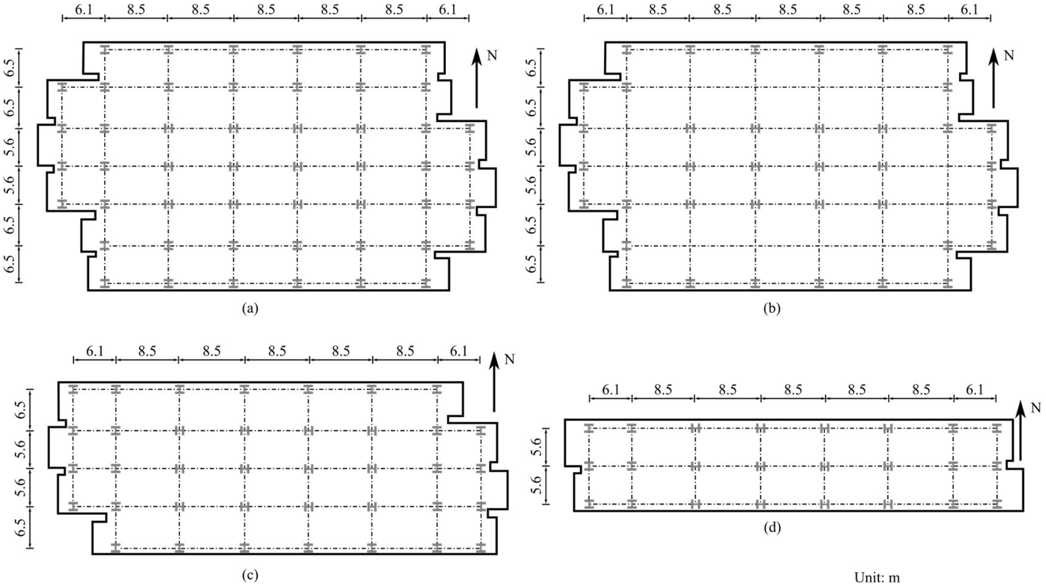


Figure 5.13. Plan views of the 45-story building framing.

The subsequent subsections address a practical challenge: estimating the building's dynamic displacement using measured acceleration data. As detailed in Section 5.3, both measured acceleration and displacement data series typically contribute to the displacement estimation process. However, external excitations are often limited to relatively low levels, which can constrain the accuracy of state estimations under more severe loading conditions. In this study, radar measurements are conducted exclusively during slight wind disturbances. The corresponding measured acceleration and displacement data obtained under these minimal excitations are then utilized to predict structural behavior under stronger wind disturbances. Notably, direct data from the "real-world" structure are considered as high-fidelity data, providing reliable and accurate inputs for the estimation process.

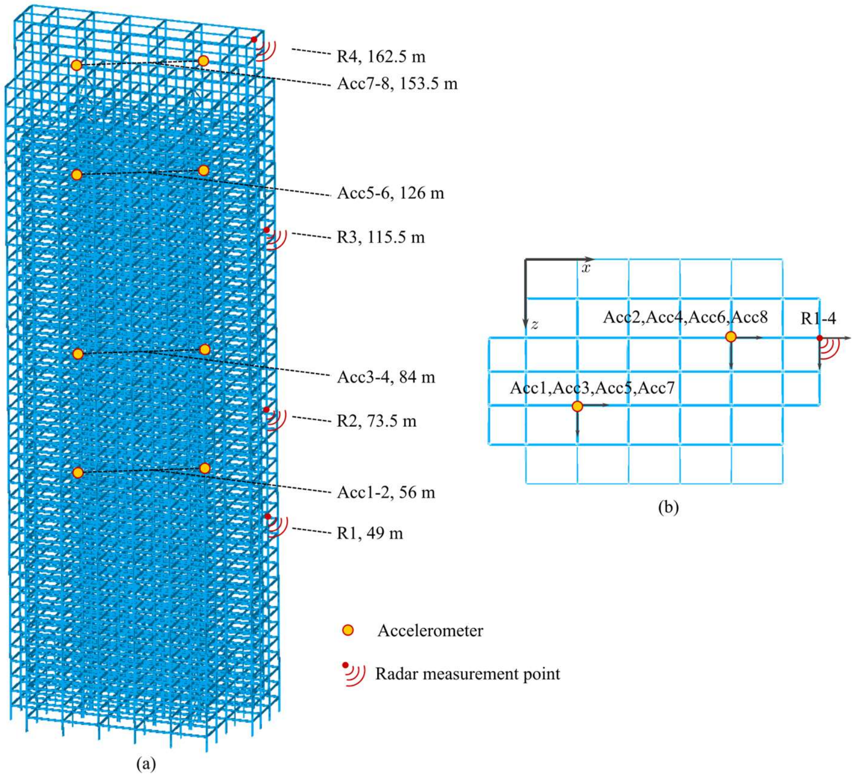


Figure 5.14. SHM system for the 45-story building structure.

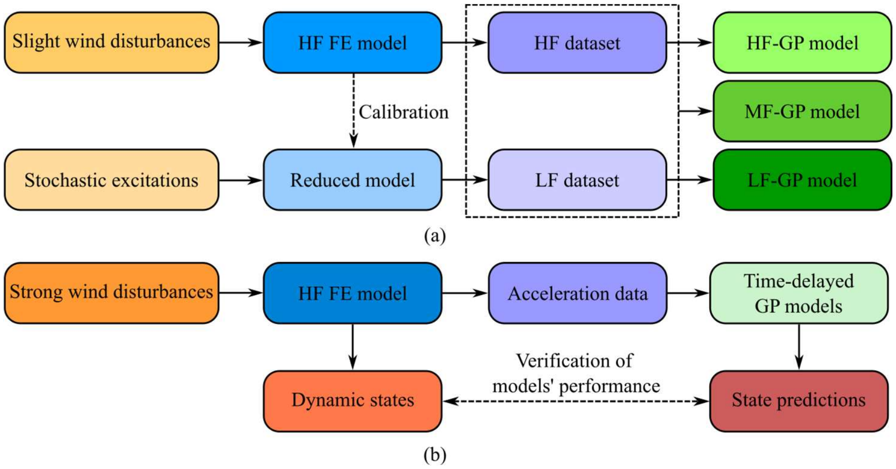


Figure 5.15. (a) Building time-delayed GP via different sources of data; (b) Validation of different time-delayed GP models.

The central question explored in the ensuing subsections is whether these high-fidelity measurements can accurately predict structural responses under significantly different excitation conditions. Addressing this question, the proposed approach demonstrates an effective method for performing dynamic displacement estimation, as illustrated in Figure 5.15. This investigation is crucial for validating the efficacy of the proposed methodologies within this chapter, ensuring that accurate state estimations can be achieved even when the measurement conditions differ from those under which the data are originally collected.

5.6.2. Building structure subjected to wind excitations

The von Kármán spectrum is adopted to generate wind excitations for the building structure (Solari and Piccardo, 2001). Specifically, the wind speeds in both the along-wind and cross-wind directions are characterized by PSD that follow the von Kármán turbulence model. This model accurately represents the statistical distribution of wind speed fluctuations across various frequencies and spatial scales, making it particularly suitable for dynamic structural analyses of high-rise buildings. The PSD are mathematically defined as follows:

$$\frac{nS_u(z,n)}{u_*^2} = \frac{4\beta\left(\frac{nL_u(z)}{U(z)}\right)}{\left[1+70.8\left(\frac{nL_u(z)}{U(z)}\right)^2\right]^{\frac{5}{6}}}, \tag{5.18}$$

$$\frac{nS_v(z,n)}{u_*^2} = \frac{4\beta\left(\frac{nL_v(z)}{U(z)}\right)\left(1+755.2\left(\frac{nL_v(z)}{U(z)}\right)^2\right)}{\left[1+283.2\left(\frac{nL_v(z)}{U(z)}\right)^2\right]^{\frac{11}{6}}}, \tag{5.19}$$

where S_u and S_v denotes the PSD in the along-wind and cross-wind directions, respectively; n represents the frequency in Hz; z is the spatial height in meters; u_* is the friction velocity, set at 1.26 m/s; β is a factor related to roughness length z_0; $L_e(z)$

is the integral length scale of the turbulence component in the e direction, where $e \in \{u, v\}$; $U(z)$ is the mean wind profile, defined by an exponential form:

$$U(z) = U_g \left(\frac{z}{H_g} \right)^{\alpha}. \tag{5.20}$$

Here, U_g, H_g, and α are control parameters set to 30 m/s, 100 m, and 0.22, respectively. Here, the integral scale length scale $L_u(z)$ is represented as:

$$L_u(z) = Cz^m, \tag{5.21}$$

where C and m can be determined based on the roughness length z_0, which is dependent on the surface type. In this example, z_0 is selected as 0.03 m, corresponding to areas with low vegetation and isolated obstacles. Consequently, the C and m are set to be 100 and 0.2, respectively. According to Simiu and Yeo, (2021), for the cross-wind length scale $L_v(z)$, it is set to $0.33L_u(z)$; the roughness length also influences the factor β, which is determined to be 6.0.

Using these parameters, the von Kármán spectrum effectively models the wind speed at each height across the structure. To capture the spatial correlations between wind speeds at different heights, the cross-spectral density $S_{e_1 e_2}(z_1, z_2, n),\ e \in \{u, v\}$ is formulated with:

$$S_{e_1 e_2}(z_1, z_2, n) = \sqrt{S_e(z_1, n) S_e(z_2, n)} \exp(-\hat{f}), \tag{5.22}$$

where

$$\hat{f} = \frac{2n\left[\left(C_z^2 + C_y^2\right)\left(z_1^2 - z_2^2\right)\right]^2}{\left[U(z_1) + U(z_2)\right]}. \tag{5.23}$$

In these equations, $S_e(z_1, n)$, $S_e(z_1, n)$ represent the PSD in the along-wind or cross-wind direction at heights z_1 and z_2, respectively; $\hat{f}$ quantifies the decay of correlation between wind speeds as the height difference increases, incorporating the integral length scales and mean wind profiles at each height. In this example, C_z and C_y are control parameters for $\hat{f}$, set to 1×10^{-2} and 2×10^{-3}, respectively.

Given the set involves heights of 45 floors of the investigated, at each frequency, a CSD matrix is calculated by randomly determining the phase of the wind histories. The inverse Fourier transformation is then applied to the series of CSD matrices. Consequently, the wind speed histories in both along-wind and cross-wind direction at the specified heights are generated.

Once the wind speed histories have been generated, they are transformed into wind load histories based on the principle that wind pressure is proportional to the square of wind speed. Specifically, let $s(z_i, t), i \in \{1,2,\dots,45\}, t \in \mathbb{R}$ denotes the calculated wind speed time history, the wind load at each floor is calculated using the relationship:

$$f(z_i, t) = \eta k(z_i) \mathrm{sgn}(s(z_i, t)) s(z_i, t)^2. \tag{5.24}$$

where η is a variational coefficient to adjust the amplitude of wind excitations, $k(z_i)$ is a coefficient dependent on the exposed area and specific characteristics of ith floor facing the wind load direction. $\mathrm{sgn}(\cdot)$ is the sign function that incorporates the directionality of wind excitation, which ensures that the wind loads are correctly oriented based on the prevailing wind speed relative to the building's orientation.

Figure 5.16 depicts the distributions of the normalized wind loads applied to the building structure. In this example, the parameter η is adjusted to scale the wind load , simulating conditions of both slight and strong wind disturbances. Consequently, the aerodynamic loads applied to each floor are proportional to the wind speed histories, accurately reflecting the dynamic forces acting on the building. The resulting wind load histories are integrated into the FE model to simulate the building’s dynamic response under realistic wind conditions. These simulated responses are then collected and utilized for further state estimation tasks using various time-delayed GP models.

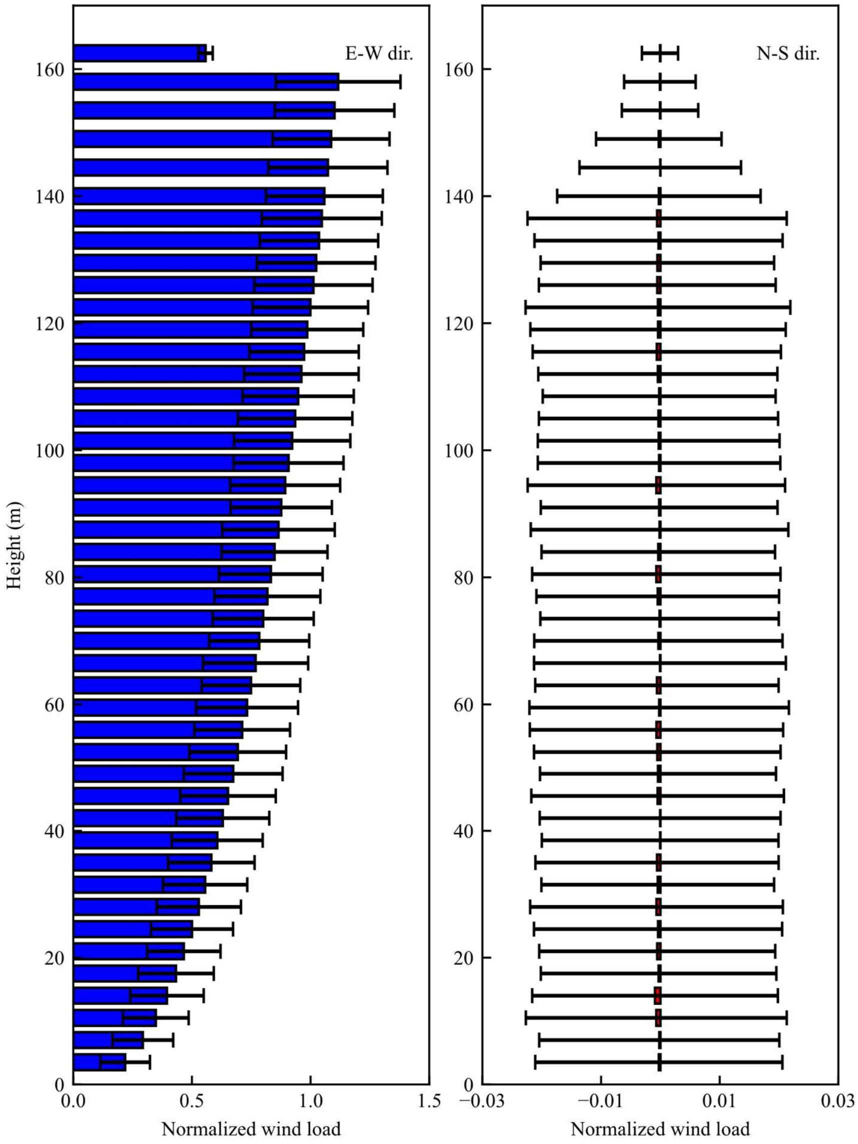


Figure 5.16. Distributions of the normalized wind loads applied to the building structure.

5.6.3. State estimation of the building structure solely based on high-fidelity data

Following the layout of the SHM system, the recorded accelerations and displacements at the corresponding DOF are meticulously extracted from the comprehensive response data obtained from the FE model subjected to slight wind disturbances. The sampling frequency for all sensors is set to 4 Hz. To enhance signal quality and reduce noise, acceleration data collected from sensors located on the same floor and oriented in the same directional axis are averaged. This averaging process ensures that displacement measurements are consistent and reflective of the overall structural behavior, mitigating the influence of localized anomalies or sensor-specific noise. Furthermore, the calculated displacement data are filtered to eliminate quasi-static displacements caused by wind, thereby isolating the dynamic component of the displacement. Figure 5.17 presents a subset of processed data, spanning from 0 to 500 s, which is utilized to demonstrate the state estimation process. As observed in the figure, the acceleration and displacement in the *z*-direction are significantly greater than those in the *x*-direction. Additionally, the vibrations along the *x*-axis appear more stationary compared to the dynamic behavior observed in the *z*-axis vibrations.

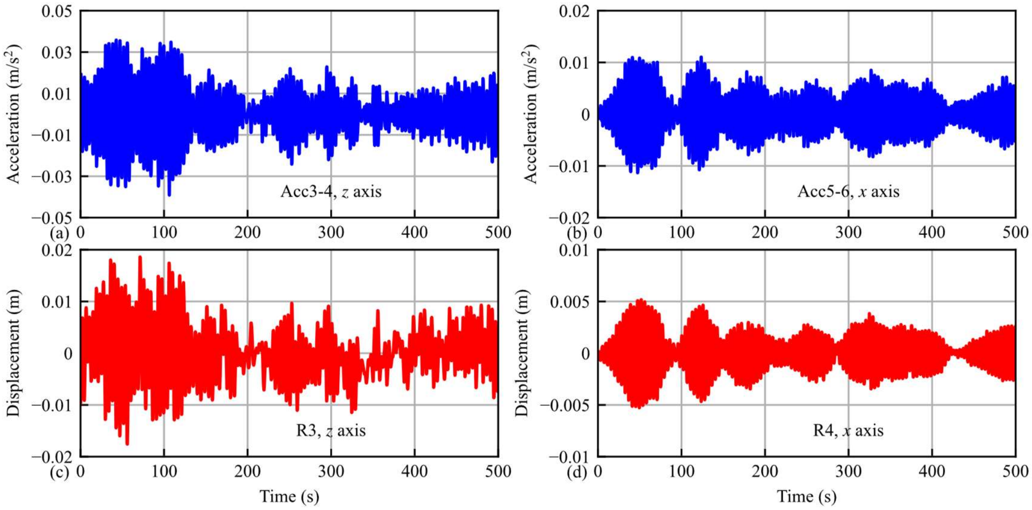


Figure 5.17. Processed high-fidelity acceleration and displacement data for demonstration of state estimation.

The time-delayed GP models are trained using these high-fidelity data. Specifically, acceleration data from accelerometers Acc3-4 are used to predict displacements from reference point R3, while acceleration data from accelerometers Acc5-6 are utilized to predict displacements from reference point R4. For the 500 s dataset, the first half is allocated for training the models, and the remaining half is reserved for testing the validity of the time-delayed GP models.

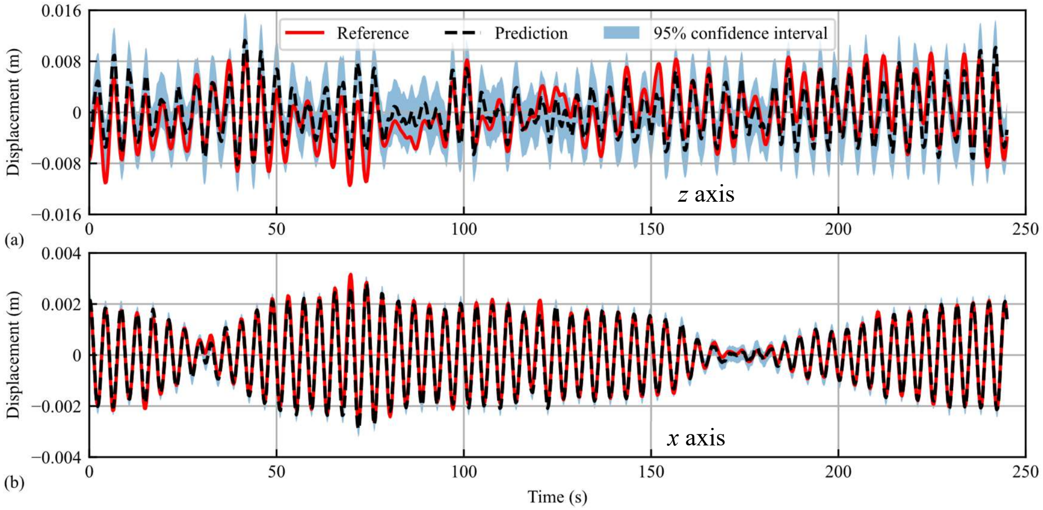


Figure 5.18. Predictive results of time-delayed GP models on test dataset.

After optimizing the hyperparameters of the GP models, the well-trained time-delayed GP models are evaluated using the test dataset. Figure 5.18 illustrates the prediction results alongside the reference displacement data. The results indicate that the prediction quality for *x*-axis displacements is superior to that of *z*-axis displacements. Specifically, the posterior distributions for *z*-axis displacements exhibit wider confidence intervals, and the predictive mean shows less accuracy compared to the predictions for *x*-axis displacements. Nonetheless, the posterior predictions from both time-delayed GP models effectively align with the reference data, demonstrating the models' capability to capture the underlying displacement dynamics.

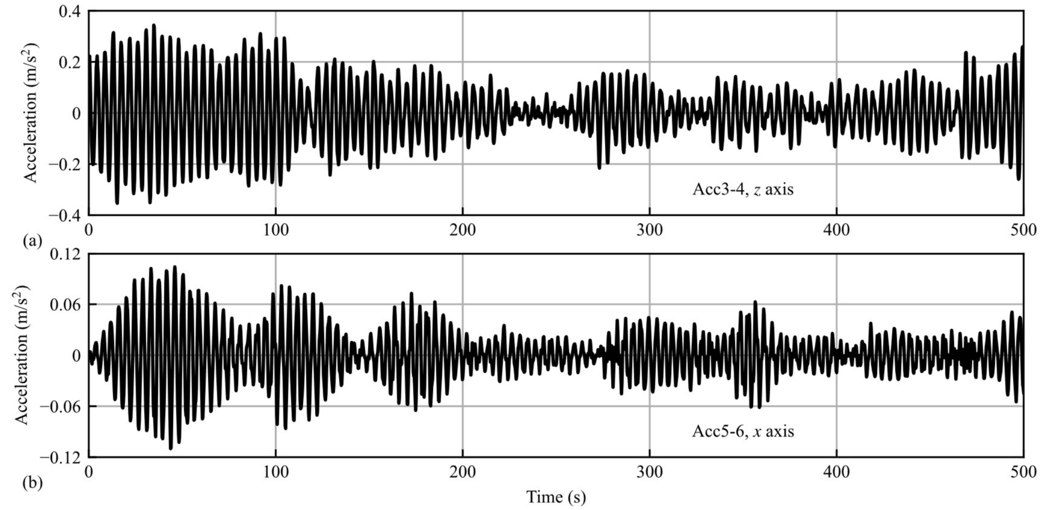


Figure 5.19. Processed acceleration data when building undergoes a strong wind disturbance.

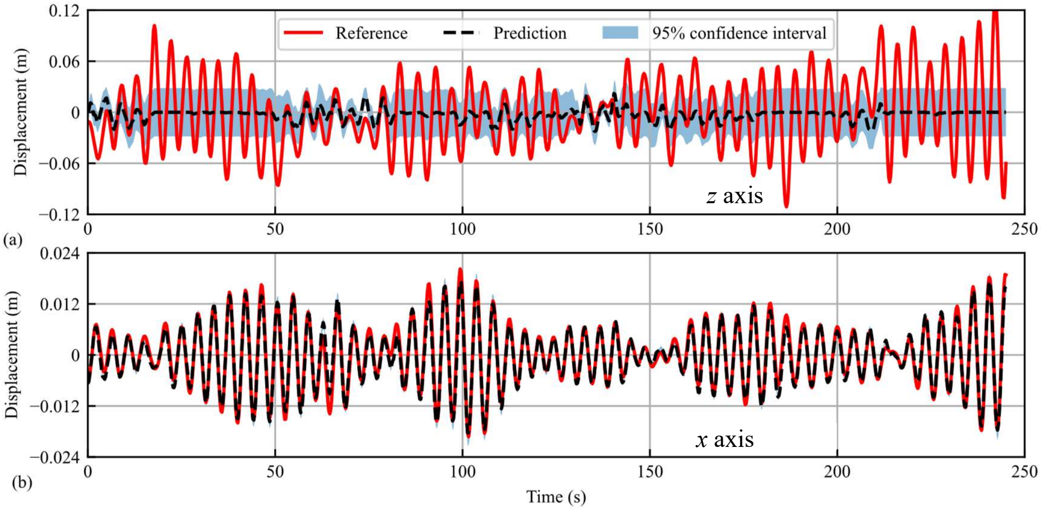


Figure 5.20. Displacement prediction using high-fidelity data trained time-delayed GP model when building undergoes a strong wind disturbance.

Figure 5.19 displays acceleration data recorded during a strong wind disturbance, where the amplitude of acceleration significantly exceeds the data used for training the time-delayed GP models. By inputting these higher-valued accelerations into the well-trained GP models, the resulting displacement predictions are shown in Figure 5.20. The results reveal that for *z*-axis displacement predictions, the model's performance is suboptimal, indicating that the time-delayed GP model is effective only under conditions of relatively small displacements. In contrast, for x-axis displacement predictions, although there is a decrease in prediction accuracy compared to the test dataset, the performance remains within acceptable ranges for practical applications.

5.6.4. Enhanced state estimation of the building structure using low-fidelity and multi-fidelity data

A reduced model of the building structure is developed to generate low-fidelity data for state estimation. As illustrated in Figure 5.21, the entire tall building is divided into 45 segments, each modeled as a 3D beam-like element. Consequently, the reduced model is represented as a cantilever beam with 46 nodes and 45 elements. Each node possesses five DOF: two horizontal translational DOF and three rotational DOF. Therefore, each element has 10 DOF, and the entire model comprises 225 DOF in total, with all DOF of node 1 constrained.

The element mass and stiffness matrices are initialized based on the high-fidelity FE model. Half of the mass of each segment between two adjacent nodes is assigned to the upper node and the remaining half to the lower node. Similarly, the equivalent rotational inertia of each segment relative to the node is calculated, with half allocated to the upper node and half to the lower node (Xia et al., 2008).

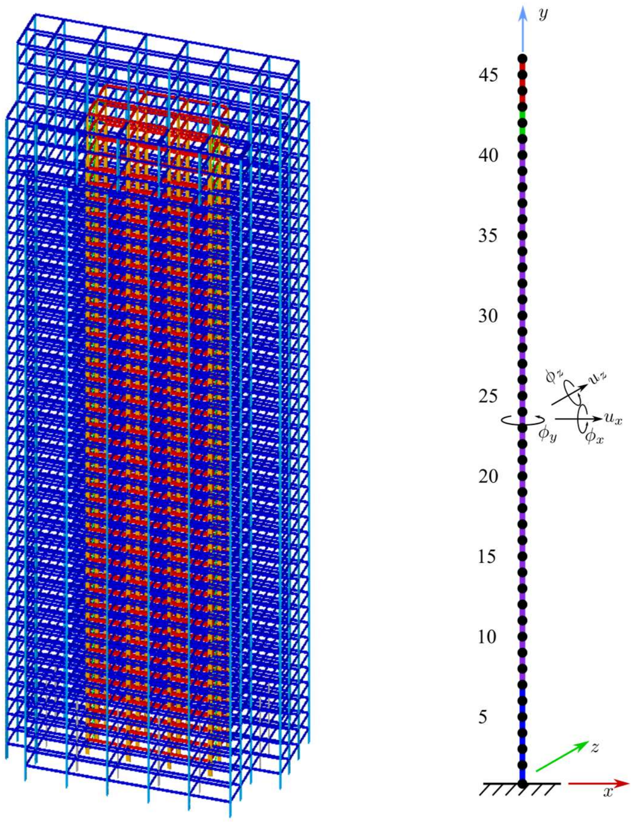
Element stiffness matrices are derived using the displacement method through the following steps:

i. Each beam element is aligned with a corresponding segment from the high-fidelity FE model. Each segment consists of two horizontal sections (planes), with multiple nodes located within each plane. In each plane, a central node is generated. All nodes within a plane are rigidly connected to their respective central node using rigid beam connections. These rigid beams are assigned materials with significantly higher elastic moduli and sections with much larger cross-sections to ensure that the connections behave as rigid links.

ii. To compute the stiffness coefficients of the *j*th DOF, a displacement of 1×10^{-2} m or rad is applied to the *j*th DOF of one central node, while restricting all other potential displacements of this central node and another central node. This displacement is deliberately kept small to avoid nonlinear geometric effects. The index *j* ranges from 1 to 10, where *j*=1 to 5 correspond to lateral translational displacements in the *x* and *z* directions and rotations about the *x*, *z*, and *y* axes at the lower node, respectively. Similarly, *j*=6 to 10 represent the corresponding displacements at the upper node. Figure 5.22(a) to (e) depicts the cases for *j*=1 to 5, with Figure 5.22(e) specifically showing the rigid beam connection to the defined node.

iii. The resulting generalized force at the *i*th DOF of the central node due to the applied displacement at the *j*th DOF is calculated. The forces are multiplied by 1×10^{2} and recorded as the *i*th row and *j*th column of the element stiffness matrix. For each specified displacement, ten resisting forces are obtained, requiring ten iterations per element to fully determine the stiffness matrix.

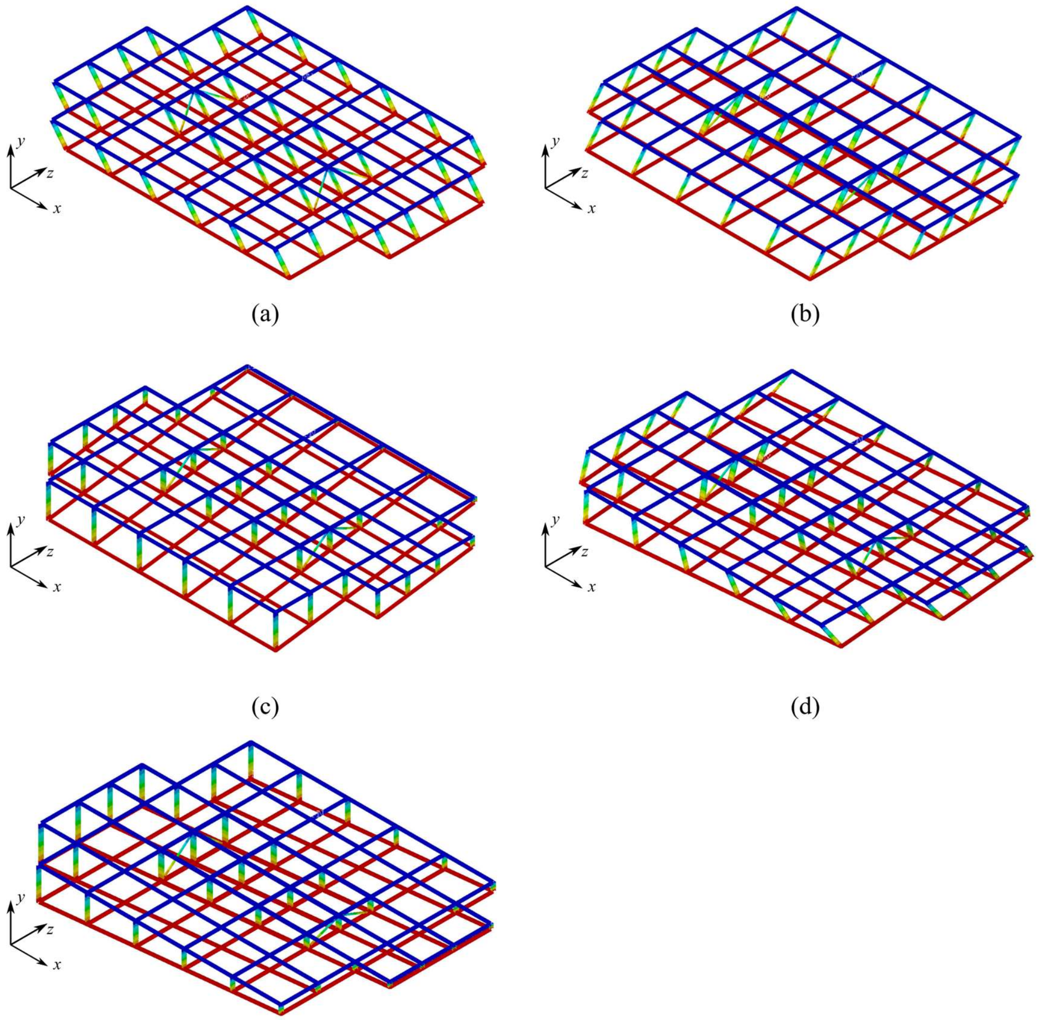
iv. The procedures from i to iii are repeated for each element, resulting in all element stiffness matrices. These matrices are then assembled into the global stiffness matrix. By applying the constraints on all DOF of node 1, the stiffness matrix of the reduced model is fully constructed.



(a) (b)

Figure 5.21. The high-fidelity FE model and reduced model of the investigated building structure.

The initialization of the reduced model lays the foundation for the subsequent model updating process, which aims to enhance the accuracy of state estimation. It is crucial to emphasize that the calibration of this reduced model adheres to actual engineering procedures. Although the reduced model operates at a low fidelity, it is essential to align it with the characteristics of the actual structure to ensure practical applicability.



(a)

(b)

(c)

(d)

(e)

Figure 5.22. Positive displacement in different directions: (a) *x* direction; (b) *z* direction; (c) rotation along *x* axis; (d) rotation along *z* axis; (e) rotation along *y* axis.

Initial calculations reveal that the dynamic properties of the reduced model differ from those of the high-fidelity FE model due to the large simplifications. To reduce the error, the residual vectors for the first six modes, denoted as $\boldsymbol{r}_{nf} \in \mathbb{R}^6$ and $\boldsymbol{r}_{ms} \in \mathbb{R}^6$ are defined. Here, $\boldsymbol{r}_{nf}$ represents discrepancies in natural frequencies, and $\boldsymbol{r}_{ms}$ quantifies the mode shape similarities. The ith components of these two vectors are calculated as follows:

$$r_{nf}^{i} = \left(\frac{\hat{f}_i - f_i(\boldsymbol{\kappa})}{\hat{f}_i}\right)^2, \tag{5.25}$$

$$r_{ms}^{i} = 1 - \frac{\widehat{\boldsymbol{\phi}}_i^{\mathrm{T}} \boldsymbol{\phi}_i(\boldsymbol{\kappa})}{\left(\widehat{\boldsymbol{\phi}}_i^{\mathrm{T}} \widehat{\boldsymbol{\phi}}_i\right)\left(\boldsymbol{\phi}_i^{\mathrm{T}}(\boldsymbol{\kappa}) \boldsymbol{\phi}_i(\boldsymbol{\kappa})\right)}. \tag{5.26}$$

where $\hat{f}_i$ and $\widehat{\boldsymbol{\phi}}_i$ denotes the *i*th natural frequency and mode shape of the high-fidelity FE model, respectively. Conversely, $f_i(\boldsymbol{\kappa})$ and $\boldsymbol{\phi}_i(\boldsymbol{\kappa})$ is the *i*th natural frequency and mode shape of the reduced model, given an updating parameter set $\boldsymbol{\kappa}$. The parameter set $\boldsymbol{\kappa}$ comprises two parameters, κ_E and κ_G, for each element stiffness matrix. Specifically, κ_E represents the modulus variation coefficient affecting all components of the element stiffness matrix, and κ_G is associated only with bending and rotational DOF.

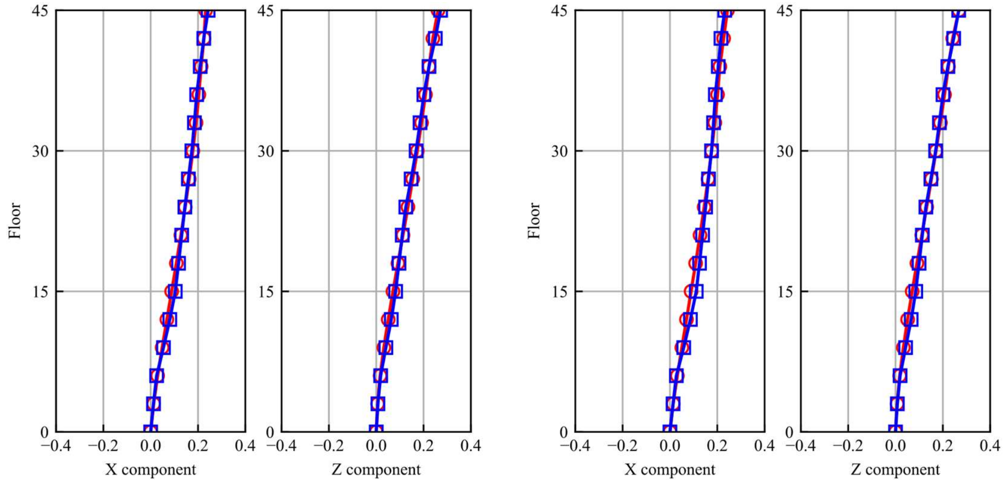
Therefore, for the element stiffness matrix $\mathbf{K}^i$ of *i*th element, the updated matrix can be expressed as:

$$\mathbf{K}_u^i = \kappa_E \mathbf{K}^i \odot \begin{bmatrix} \boldsymbol{\chi}_G & \boldsymbol{\chi}_G \\ \boldsymbol{\chi}_G & \boldsymbol{\chi}_G \end{bmatrix}, \tag{5.27}$$

with

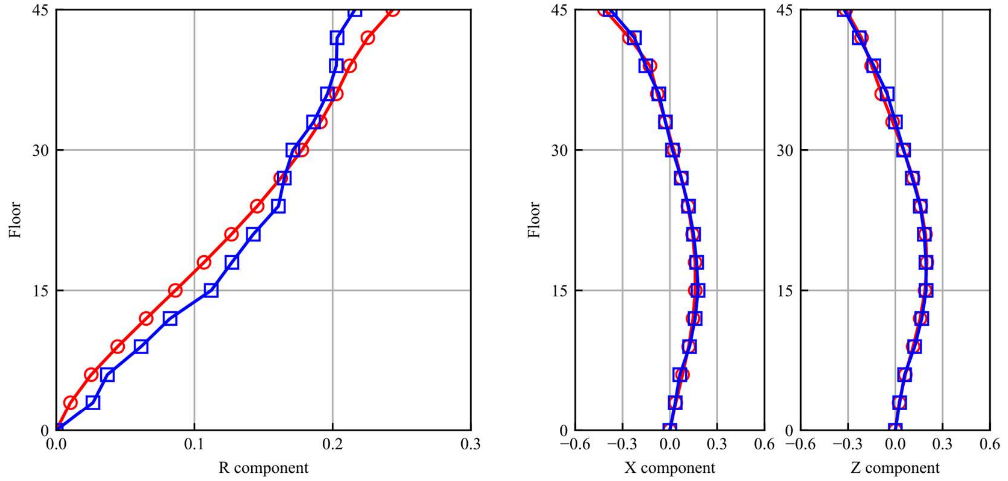
$$\boldsymbol{\chi}_G = \begin{bmatrix} 1 & 1 & \kappa_G & \kappa_G & \kappa_G \\ 1 & 1 & \kappa_G & \kappa_G & \kappa_G \\ \kappa_G & \kappa_G & \kappa_G & \kappa_G & \kappa_G \\ \kappa_G & \kappa_G & \kappa_G & \kappa_G & \kappa_G \\ \kappa_G & \kappa_G & \kappa_G & \kappa_G & \kappa_G \end{bmatrix}, \tag{5.28}$$

where "⊙" denotes the element-wise product.



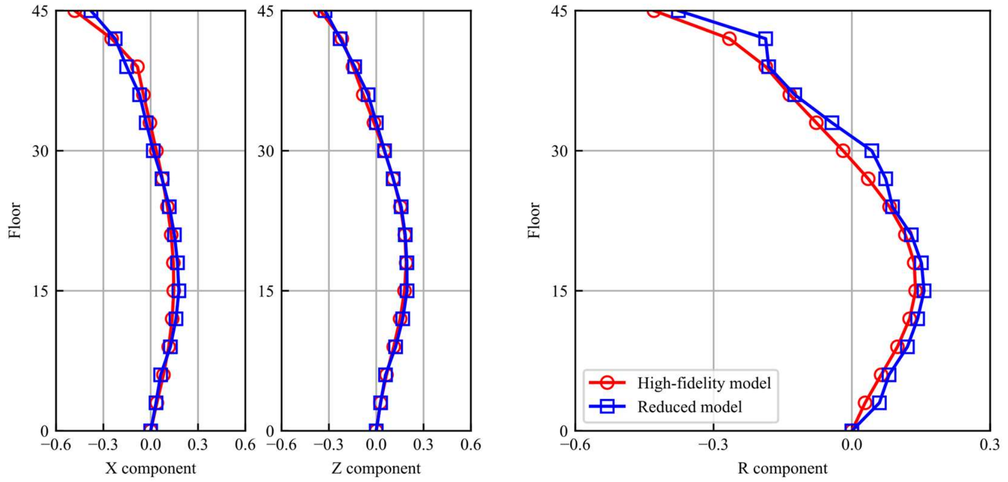
Mode 1: 1st short-axis bending

Mode 2: 1st long-axis bending



Mode 3: 1st torsion

Mode 4: 2nd short-axis bending



Mode 5: 2nd long-axis bending

Mode 6: 2nd torsion

Figure 5.23. Comparison of the mode shapes between high-fidelity FE model and reduced model.

Table 5.3. Natural frequency comparison between high-fidelity model and reduced model.

Mode	Natural frequency (Hz)		
	High-fidelity model	Reduced model	Difference (%)
1	0.231	0.230	0.4
2	0.245	0.245	0.0
3	0.306	0.305	0.3
4	0.711	0.700	1.5
5	0.765	0.777	1.6
6	0.888	0.888	0.0

The model updating is conducted by minimizing loss function $J(\boldsymbol{\kappa})$ defined in Eq. (5.29), utilizing the L-BFGS-B optimizer for the optimization process. By effectively adjusting stiffness matrices $\mathbf{K}_u^i$ based on the parameter set $\boldsymbol{\kappa}$, the reduced model could more accurately represent the dynamic behavior of the actual high-fidelity FE model. Consequently, the calibrated reduced model serves as a robust foundation for efficient and precise state estimation in subsequent analyses.

$$J(\boldsymbol{\kappa}) = \sum_{i=1}^{6} r_{nf}^i + r_{ms}^i. \tag{5.29}$$

To validate the model updating process, a comparison between the high-fidelity FE model and the reduced model was conducted. As presented in Table 5.3, the natural frequencies of the two models are in close agreement, with the maximum discrepancy being 1.6% among the first six modes. Figure 5.23 illustrates the mode shapes, encompassing various orders of bending and torsion. The results confirm that the reduced model effectively represents the high-fidelity FE model, although some distinctions between the models persist.

The reduced model is developed to estimate the state of tall buildings subject to strong wind disturbances. As discussed in Subsection 5.6.3, high-fidelity data is unavailable for constructing a time-delayed GP model to predict the state under such extreme wind conditions. However, the calibrated reduced model can generate a substantial dataset that encompasses these scenarios. Additionally, the measured acceleration data presented in Figure 5.19 are utilized in this analysis as input for time-delayed GP models. In this section, the contribution of combining low-fidelity data generated by the reduced model with the available high-fidelity data is demonstrated to achieve accurate state prediction.

To generate low-fidelity data, it is essential to design dynamic loads for the reduced model. Wind loads on tall buildings predominantly contain energy at lower frequencies. Therefore, a parametrized roll-off PSD function is employed, defined by:

$$S(\omega) = \frac{\alpha_v \omega_c^2}{(\omega^2 + \omega_r^2)(\omega^2 + \omega_c^2)}, \tag{5.30}$$

where α_v, ω_r, and ω_c are parameters, ω is the angular frequency in rad/s. By tuning these parameters and applying inverse Fourier transformations, corresponding load histories can be generated.

Assuming that all nodes in the same directional axis are subjected to identical dynamic loads. That is, two PSD functions, one for *x*-axis loads and another for *z*-axis loads, must be determined. The measured acceleration data presented in Figure 5.19 are utilized to roughly calibrate the PSD parameters. Specifically, these parameters are adjusted to generate dynamic loads that produce acceleration outputs from the reduced model with comparable dispersions to the measured data. Ultimately, the three parameters are selected as 2×10^9, 2, and 4 for *x*-axis loads; and 4×10^{10}, 3, and 12 for *z*-axis loads. These two PSD are illustrated in Figure 5.24.

Given the input to the reduced model, the complete responses are computed. The low-fidelity data for training of time-delayed GP model is generated based on the SHM system configuration shown in Figure 5.14. For demonstration purposes, Figure 5.25 illustrates the processed low-fidelity acceleration and displacement data, where the accelerations obtained from accelerometers of the same floor and direction are averaged. As can be seen, the amplitudes of the acceleration and displacement in the low-fidelity data are much higher than those observed in high-fidelity data depicted in Figure 5.17. Additionally, the range of acceleration data in low-fidelity data closely matches that of the data depicted in Figure 5.19. This alignment enables the time-delayed GP model trained with low-fidelity data to perform competitively in state estimation under strong wind disturbances when compared to the GP model trained with high-fidelity data.

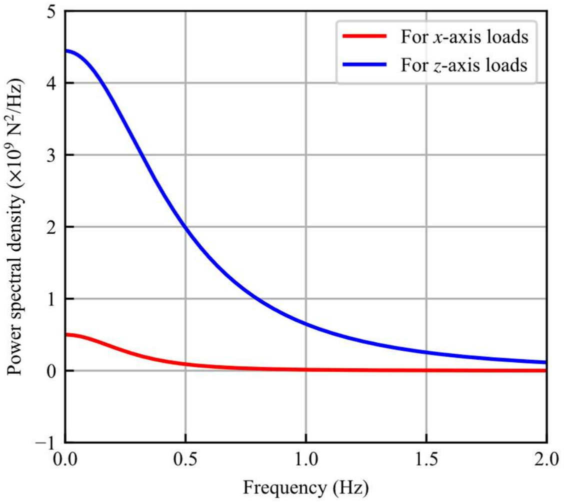


Figure 5.24. PSD for x- and z-axis loads.

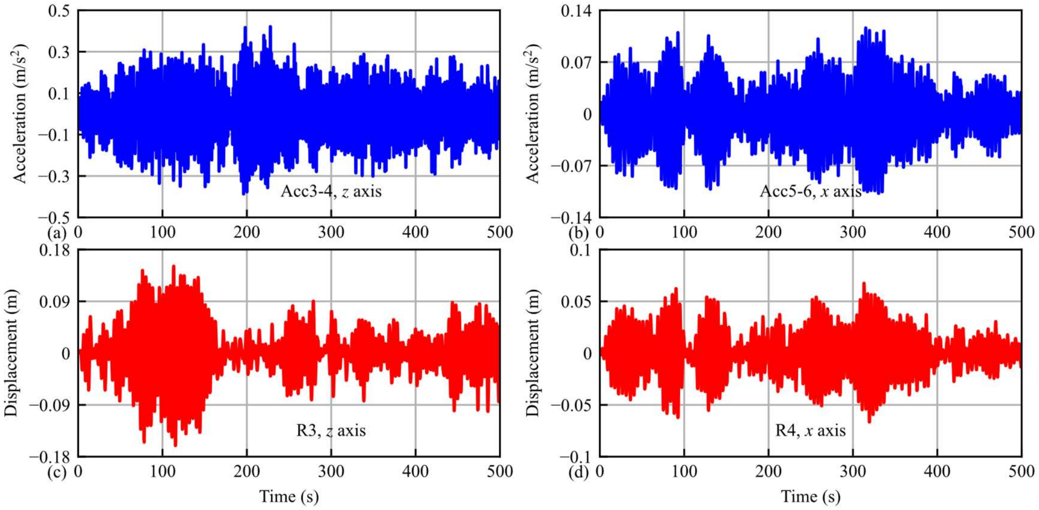


Figure 5.25. Processed low-fidelity acceleration and displacement data for demonstration of state estimation.

Subsequently, the time-delayed GP models are developed using the generated low-fidelity data. In particular, acceleration measurements from accelerometers Acc3-4 are employed to forecast displacements at reference point R3, whereas data from accelerometers Acc5-6 are utilized to predict displacements at reference point R4. For the 500 s dataset, the initial 250 s are designated for training the GP models, while the final 250 s are used to evaluate their performance and validate their predictive capabilities.

After optimizing the hyperparameters, the well-trained time-delayed LFGP models were evaluated using the test dataset. Figure 5.26 presents the predicted posterior distributions of displacements at reference points R3 (*z*-axis) and R4 (*x*-axis). The alignment between the predicted displacements and the reference data confirms the successful training and accurate predictive performance of the time-delayed LFGP models.

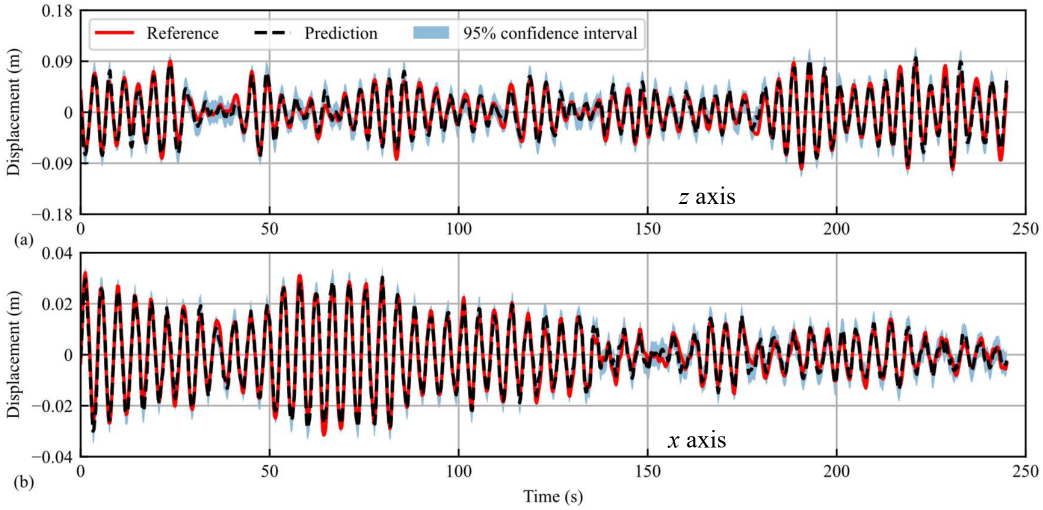


Figure 5.26. Predicted results of low-fidelity time-delayed GP models on test dataset.

As previously discussed, the objective of constructing LFGP models is to enable accurate state estimation of the high-fidelity FE model. By inputting the measured acceleration data from the high-fidelity FE model, as depicted in Figure 5.19, into the well-trained time-delayed LFGP models, the dynamic displacements at reference points R3 and R4 are calculated. The results, shown in Figure 5.27, demonstrate that the predicted posterior distributions from the LFGP models outperform those from the HFGP models. Specifically, for the *x*-axis displacement prediction illustrated in Figure 5.27(b), the LFGP model effectively captures the reference displacement, despite the confidence interval exhibiting greater uncertainty compared to predictions based solely on high-fidelity data. In contrast, for the *z*-axis displacement prediction, the LFGP model provides effective results, although the accuracy is somewhat reduced compared to the *x*-axis prediction. This decreased accuracy may be influenced by the nonstationary excitation in the *z*-axis direction. Furthermore, the HFGP model fails to accurately predict displacements in the *z*-axis, highlighting its limitations under certain conditions. Therefore, the LFGP model proves to be a superior choice for state estimation of the high-fidelity FE model, particularly in scenarios involving strong wind disturbances.

GP models trained based on low-fidelity data can be further enhanced by incorporating high-fidelity data. Although, as previously mentioned, the performance of high-fidelity data alone is quite poor, it is directly obtained from the high-fidelity FE model. Therefore, the MFGP model is constructed using the proposed multi-fidelity modeling approach. Based on testing results, the displacement prediction accuracy in the *x*-axis direction is already excellent when using either high-fidelity or low-fidelity data. Consequently, the focus is solely on predicting displacements in the z-axis direction at four reference points, R1 to R4. Using acceleration data collected from adjacent floors as inputs, four MFGP models are developed. Additionally, both level 1 and level 2 multi-fidelity models were tested based on high-fidelity data, with level 2 demonstrating better adaptability for this application.

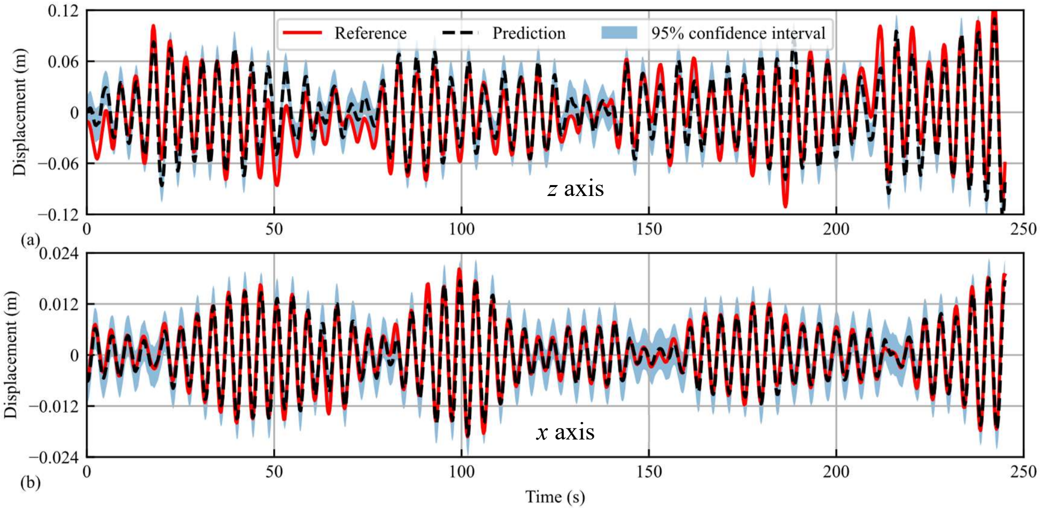


Figure 5.27. Displacement prediction using low-fidelity data trained time-delayed GP model when building undergoes a strong wind disturbance.

After optimizing the hyperparameters of the MFGP models, the processed acceleration data are input into the models, and the predicted displacements at points R1 to R4 are presented in Figure 5.28(a) to (d). To quantify the prediction accuracy, the previously defined NRMSE metric is employed to compare the predicted means with reference data. As shown in Figure 5.29, the prediction accuracy of the multi-fidelity models has improved compared to the low-fidelity model after integrating high-fidelity data. Although the accuracy improvement depicted in the figure is limited, with the most notable being a 3.2% reduction in NRMSE for the *z*-axis displacement prediction on the 45th floor, this case study aims to illustrate the benefits of multi-fidelity modeling for prediction. In practical applications, state estimation typically utilizes LFGP models. However, to further enhance the accuracy of state estimation, introducing high-fidelity data and employing a multi-fidelity modeling approach proves to be a favorable option.

Additionally, it is noteworthy that in the displacement prediction results shown in Figure 5.28, the posterior confidence intervals are noticeably smaller and do not align as well with reference data compared to the LFGP model. This may be due to the incorporation of high-fidelity data leading to an overly idealized prediction model. Therefore, regarding prediction uncertainty in this case, the LFGP model is actually more precise. This conclusion appears to differ from that observed in the first case study, and this discrepancy will be further explored in the next subsection.

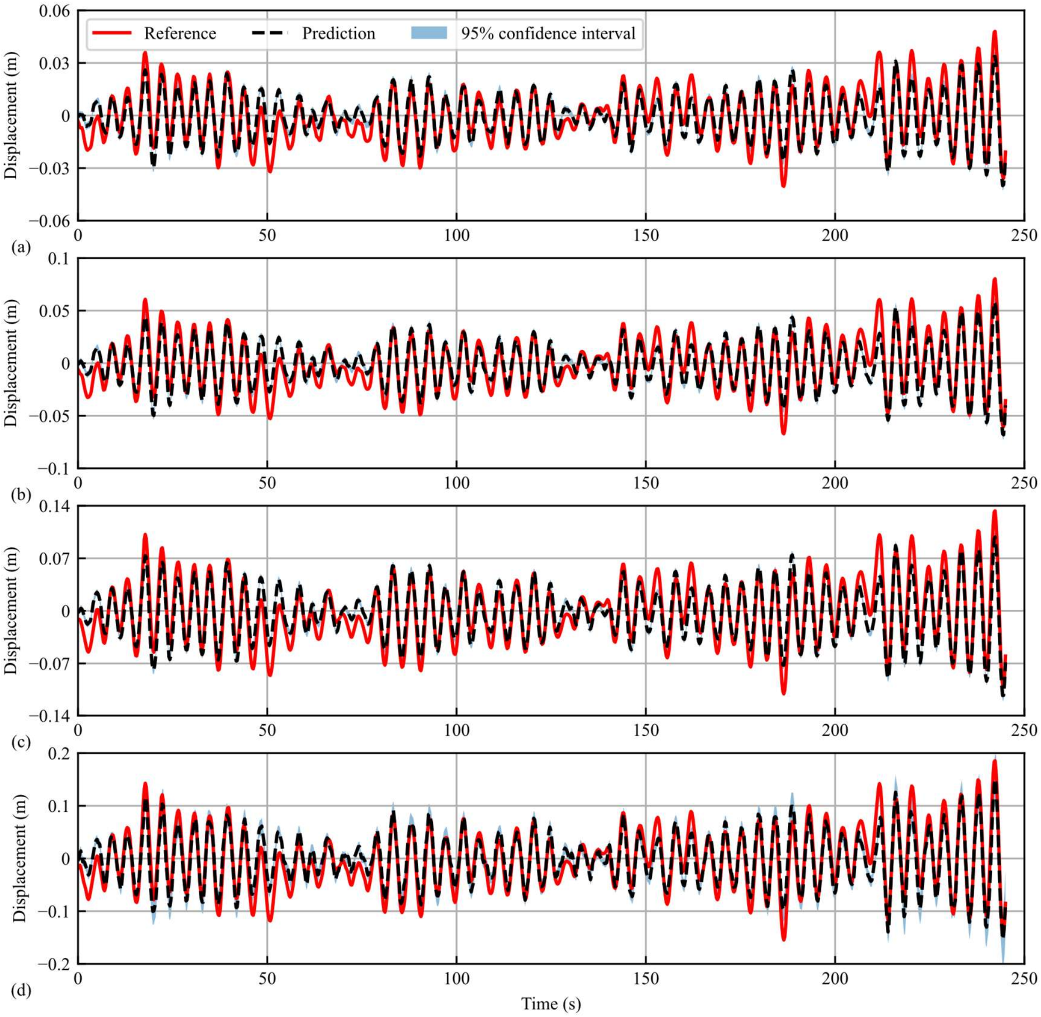


Figure 5.28. *z*-axis displacement prediction using MFGP models.

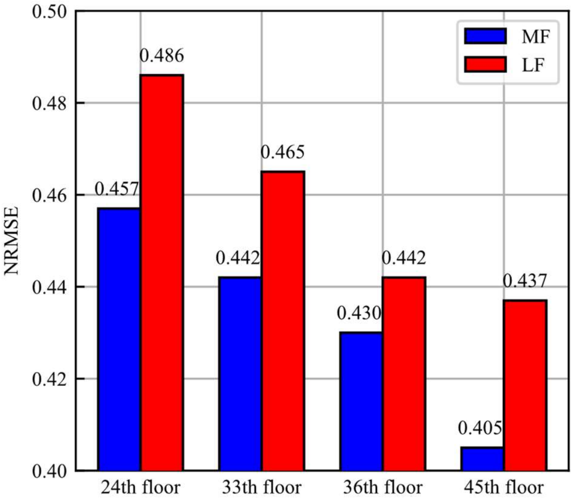


Figure 5.29. Quantitative comparison of state estimations by LFGP and MFGP models using NRMSE.

5.7. Discussion

In this chapter, two numerical case studies are presented to validate that time-delayed GP models can perform state estimation efficiently and accurately. Additionally, the adoption of a multi-fidelity modeling approach further enhances the precision of state estimation. Based on the results from these case studies, we focus on discussing the applicability and feasibility of MFGP models in practical applications.

Figure 5.30 visually represents the distribution of high-fidelity and low-fidelity data within a two-dimensional space. The area enclosed by the black boundary delineates the regions where different types of data are available. Blue square markers indicate the presence of low-fidelity data points, while red circles denote high-fidelity data points. The datasets X_1 and X_2 correspond to high-fidelity and low-fidelity datasets, respectively. In scenario Figure 5.30(a), high-fidelity data do not exhibit a disadvantage in terms of quantity compared to low-fidelity data, and their spatial coverage is also comparable. Under these conditions, it is evident that relying solely on high-fidelity data for constructing time-delayed state estimation models is the most effective approach. In this context, multi-fidelity modeling does not offer additional advantages.

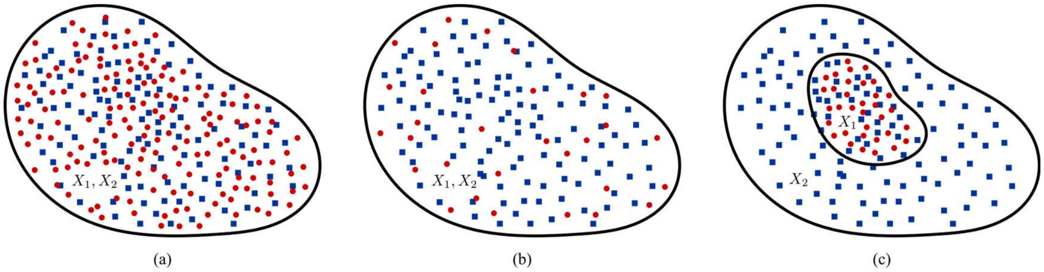


Figure 5.30. Three different distributions of high-fidelity and low-fidelity data.

Figure 5.30(b) corresponds to the first case study discussed in this chapter, where the volume of high-fidelity data is limited, yet its spatial coverage matches that of low-fidelity data. The results from this case study, as illustrated in Figures 5.10 and 5.11, demonstrate that integrating both high-fidelity and low-fidelity data through the multi-fidelity approach significantly enhances prediction accuracy and the posterior distributions of the GP model compared to models that utilize only low-fidelity or high-fidelity data individually. Therefore, in this scenario, multi-fidelity modeling is essential for achieving improved state estimation accuracy.

Conversely, Figure 5.30(c) illustrates the situation addressed in the second case study, which is also commonly encountered in practical applications. Here, the quantity of high-fidelity data is not substantially lower than that of low-fidelity data; however, the spatial coverage of high-fidelity data is considerably smaller. The findings from this case study reveal that using solely high-fidelity data for state estimation in regions with limited high-fidelity data leads to model inadequacies. While models relying exclusively on low-fidelity data outperform those using only high-fidelity data in this scenario, the fusion of both data types within an MFGP model results in enhanced prediction accuracy compared to an LFGP model alone. Nevertheless, regarding the effectiveness of posterior predictions, the results indicate that the LFGP model exhibits higher posterior accuracy than the MFGP model. This discrepancy likely stems from the significant variation in spatial coverage between datasets X_1 and X_2. Addressing this issue may require more advanced multi-fidelity modeling techniques to improve the posterior prediction capabilities of MFGP models.

5.8. Summary

In this chapter, a comprehensive framework for real-time state estimation of structural dynamic systems using time-delayed GP has been developed and validated.

The proposed methodology leverages both high-fidelity and Low-fidelity data through a multi-fidelity modeling approach, significantly enhancing the accuracy and reliability of state estimations. The chapter first introduces the theoretical underpinnings of time-delayed GP models within the RKHS framework, emphasizing their capability to capture complex temporal dependencies and correlations in structural measurements. The integration of multi-fidelity data is meticulously detailed, showcasing how combining high-fidelity data, derived from precise physics-based models, with abundant low-fidelity data, obtained from simplified models, can overcome the inherent limitations of single-fidelity approaches. Two numerical case studies are presented to validate the effectiveness of the proposed MFGP models, including a mass-spring-damper dynamic system, and a 45-story tall building subject to wind excitations. Throughout these case studies, key performance metrics such as NRMSE and PLL were employed to quantitatively assess the models' accuracy and uncertainty estimation capabilities. The results consistently highlight that MFGP models offer notable improvements over single-fidelity models, particularly in environments where data availability and fidelity vary spatially and quantitatively. To sum up, the adoption of time-delayed MFGP presents a robust and effective approach for real-time state estimation in structural dynamic systems. By intelligently combining diverse data sources, MFGP models enhance both the precision and reliability of state predictions, making them highly suitable for complex engineering applications.

Chapter 6.

Transfer learning of recurrent neural networks for enhanced state estimation in structural dynamics

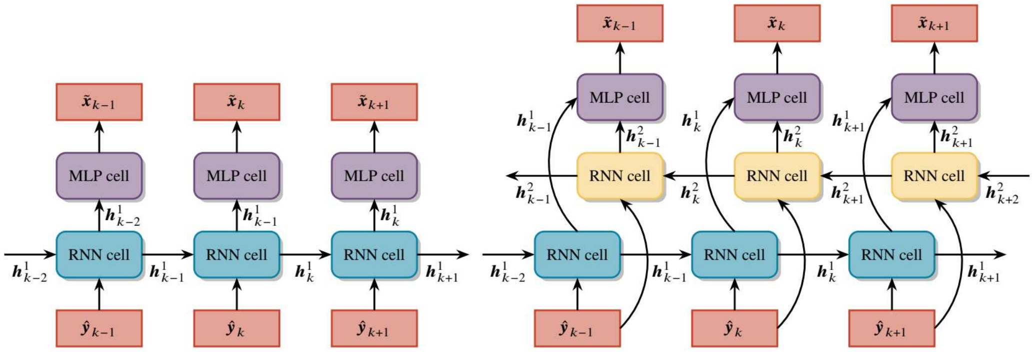
6.1. Introduction

This chapter presents a method for precise state identification by leveraging transfer learning within a deep learning framework, specifically utilizing RNN. While the Bayesian-based time-delayed MFGP model introduced in Chapter 5 demonstrated exceptional prediction accuracy and robust uncertainty quantification capabilities in state prediction, it typically models each DOF state individually. This approach necessitates that the output of the corresponding high-dimensional mapping process remains singular. Consequently, for complex scenarios requiring joint predictions across multiple DOF, the time-delayed MFGP demands multiple separate modeling, which can adversely affect task completion efficiency. To address these limitations, this chapter introduces a complementary approach based on deep learning. An efficient point estimation framework is adopted, which excels in learning and exploiting correlations inherent in multi-output problems. Given the simplicity and effectiveness of RNN in time series modeling, this neural network architecture is employed for the state estimation tasks discussed herein. Additionally, the chapter tackles the challenge of multi-fidelity data often encountered in practical state estimation scenarios by incorporating the concept of transfer learning to enhance the predictive accuracy of RNN models. Initially, a calibrated FE model is utilized to generate extensive response data under synthetic excitations. This data is subsequently processed and integrated to train an RNN model specifically designed for state estimation. Recognizing the presence of multiple sensors in real-world structural monitoring, this study innovatively employs the collected data for dual purposes. A portion of the data serves as input for the RNN model, while the complete dataset facilitates the transfer learning process for the RNN model. This dual-purpose strategy enables the RNN model to adapt effectively to real-structure state prediction tasks. To ensure effective convergence during transfer learning, a novel method is proposed in which parameters within the RNN cells at the network's front end are fine-tuned, whereas those near the output layers are frozen. This approach deviates from conventional transfer learning methods typically used for other neural network architectures and proves particularly beneficial for RNN models tailored for state estimation. Numerical and experimental studies validate that the proposed TL-RNN approach can seamlessly integrate both model-generated and actual measurement data. Under identical data acquisition conditions, the TL-RNN models achieve significantly higher accuracy compared to state estimation models that rely solely on FE models.

6.2. RNN for structural state estimation

RNN constitute a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence. This architecture allows the RNN to exhibit temporal dynamic behavior and enables the network to act as a form of memory. As a result, RNN could be exceptionally well-suited for applications in state estimation, where the states of a structural dynamic system are in a state of flux, continuously influenced by their historical states and external excitation. By assimilating the data from practical measurements over a structure, a well-trained RNN is expected to output the states of the system accurately.

In the present study, two general architectures of RNN for state estimation are investigated, i.e., the vanilla RNN and BiRNN. Although there have been advancements in RNN design, such as the LSTM and the gated recurrent unit, the simulation results indicate that the simplest form of RNN achieves outstanding performance in structural state estimation. Therefore, these more complex RNN structures are not adopted in this paper. Figure 6.1 displays the general architectures of RNN and BiRNN. It is pertinent to note that the RNN facilitates real-time state computation using measurement data. In contrast, computations of BiRNN are offline but often yield higher prediction accuracy than RNN by leveraging additional information from future measurements in the data sequence.



(a)

(b)

Figure 6.1. The general architecture of RNN and BiRNN for state estimation.

In Figure 6.1, the RNN cell in the *i*th layer processes the measurements from the structure, denoted as $\hat{\boldsymbol{y}}_k$, and the previous hidden state, $\boldsymbol{h}^i_{k-1}$, to produce the current hidden state, $\boldsymbol{h}^i_k$. The computation at each time step follows the equation:

$$\boldsymbol{h}_k^i = \tanh\left(\mathbf{W}_1^i \hat{\boldsymbol{y}}_k + \mathbf{W}_2^i \boldsymbol{h}_{k-1}^i + \boldsymbol{b}^i\right). \tag{6.1}$$

Here, k ranges from 1 to N, where N is the total number of time steps in the state estimation process. The matrices $\mathbf{W}_1^i$ and $\mathbf{W}_2^i$ represent the weights respectively for measurements and hidden states, while the vector $\boldsymbol{b}^i$ includes the corresponding biases. In an RNN, the hidden state output from the RNN cell at each step is the direct input to a MLP cell. In contrast, for a BiRNN, the hidden states from forward and backward RNN cells are concatenated before being input into the MLP cell. The MLP cell, composed of densely connected neurons with nonlinear activation functions, ultimately provides the estimated state at each timestamp.

It is suggested to set all bias vectors in both RNN and MLP cells to zero when performing state estimation tasks. This ensures that the RNN model outputs are primarily influenced by the inputs and hidden states, aligning with the expectation of zero dynamic states when measurements are zero.

As a data-driven approach, the training and test datasets for the RNN model are constructed from the vibration responses of an idealized FE model. Both the input features and target outputs for the RNN model correspond to the response variables obtained from the simulations. To create label pairs for the RNN model, it is necessary to apply synthesized external excitations to the idealized FE model, thereby generating the requisite input-output data pairs alien to the RNN architecture. The optimization of parameters is carried out by minimizing the loss function across all time steps, as demonstrated in Eq. (6.2), where $\mathcal{L}_1$ denotes the loss function for the first training of RNN. The test loss generated from the test dataset reflects the expected performance of the trained RNN model on the state estimation task.

$$\min \sum_{k=1}^{n} \mathcal{L}_1(\widetilde{\boldsymbol{x}}_k, \boldsymbol{x}_k). \tag{6.2}$$

It is also important to note that the artificial specification of excitation for the FE model implies a prior embedding in the context of state reconstruction. Traditional Kalman filter-based approaches for state estimation are typically formulated under the prior assumption that external loads are stationary white noise. However, RNN-based methodologies exhibit increased adaptability in their prior assumptions. The datasets for training and testing can be derived from not only stationary white noise but also non-stationary excitation endowed with particular statistical properties. Such versatility allows the RNN model to provide improved predictive performance in state estimation when the prior assumptions are more closely aligned with reality.

6.3. TL-RNN for enhanced state estimation

6.3.1. Apply transfer learning to enhance pre-trained RNN

All models in structural dynamics are simplifications of actual structures they attempt to represent. These approximations, by their nature, cannot capture every nuance of their real-world counterparts. The inaccuracies can be particularly stark when structures are subjected to extreme operational conditions. The divergence between the model predictions and actual behavior is often due to a range of factors, including but not limited to, oversimplified assumptions, parameter uncertainties, ignored physical mechanisms, and the constraints of computational techniques (Liu and Quek, 2013). Even though the calibration of the FE model through sensitivity analysis, uncertainty quantification, and model updating techniques improve its reliability (Moaveni et. al., 2009), discrepancies with the actual structure could not be thoroughly eliminated.

Moreover, the accuracy of state estimation in real-world structures relies heavily on the fidelity of the FE models. Since the mathematical model for state estimation is typically derived from the equations of motion or trained using datasets created from the FE model, any discrepancies in the simplified model can significantly impact the precision of state estimations. In situations where the structure experiences extreme conditions which are not accounted for in the idealized FE model, the errors in state estimation are likely to be amplified. Consequently, the reliability of the state estimation is inherently linked to the representativeness of the idealized FE model: an imprecise idealized FE model can lead to poor state estimation outcomes.

Acknowledging the impracticality of creating an exhaustive model for complex structures due to computational and modeling constraints, this chapter advances the field by proposing an approach referred to as TL-RNN that harnesses transfer learning in conjunction with RNN models for state estimation. Transfer learning leverages the power of pre-trained models by applying knowledge acquired from previous tasks to new, similar tasks (Pan and Yang 2009). This approach significantly reduces the need for extensive data typically required to achieve high performance in new learning tasks, addressing the data-hungry problem ubiquitous in deep learning.

In the context of this study, the pre-trained RNN is initially trained using data from an FE model. The subsequent transfer learning task involves improving this model using data from the actual structure. In practice, generating full input-output label pairs for the RNN based on the actual structure is challenging, as it is usually infeasible to observe all states of the structure. Moreover, the volume of observation data available from the actual structure is substantially less than what can be simulated using the idealized FE model, thereby providing an insufficient basis for training an accurate

RNN model for state estimation. A practical strategy is that beyond observing the first portion of data (e.g., accelerations) that would be used as input to RNN, measurements from other transducers on the structure together with the acceleration data would be useful to fine-tune the pre-trained RNN model. For example, in multi-story buildings, various techniques can be implemented to effectively measure inter-story displacements and velocities (Chang and Huang, 2020). Similarly, for certain structural elements, strain measurements are readily available. These measurements are compiled into a vector, denoted as $\hat{\boldsymbol{z}}_k \in \mathbb{R}^{n_e}$ at step k, where n_e denotes the number of extra transducers rather than those used for $\hat{\boldsymbol{y}}_k$. The measurement vector $\hat{\boldsymbol{z}}_k$ is intrinsically linked to the structural states predicted by the RNN. It can be derived from the state vector through a specific linear transformation, represented here by the matrix $\mathbf{B} \in \mathbb{R}^{n_e \times 2n}$. As a result, a novel loss function can be introduced, minimizing the discrepancies between $\mathbf{B}\tilde{\boldsymbol{x}}_k$ and $\hat{\boldsymbol{z}}_k$, as depicted Eq. (6.3):

$$\min \sum_{k=1}^{n} \mathcal{L}_2(\mathbf{B}\tilde{\boldsymbol{x}}_k, \hat{\boldsymbol{z}}_k), \tag{6.3}$$

where $\mathcal{L}_2(\cdot)$ represents the specified loss specified loss function for transfer learning. The minimization process aligns the pre-trained RNN model, derived from the idealized model, with the high-fidelity data of the actual structure, thus effectively fine-tuning the model to improve state estimation accuracy.

To preserve the knowledge learned from the idealized model, as shown in Figure 6.2, the parameters in the MLP layers are kept constant, i.e., the MLP layers are “frozen”. The re-training process is focused exclusively on the parameters in the RNN cells, utilizing the new loss function as described in Eq. (6.3). This selective training approach contrasts with that commonly employed in transfer learning for CNN, where it is standard practice to re-train the final few layers while the initial layers, which capture more generic features, remain unchanged. By maintaining the frozen MLP layers, the model retains its foundational understanding of system constraints and universal structural behavior derived from the idealized FE model. Meanwhile, retraining the RNN cells enables targeted refinement of temporal dynamics specific to the actual structure, ensuring the model adapts to field-measured data without overwriting pre-learned patterns. This dual mechanism mitigates catastrophic interference, a risk in conventional retraining, by incrementally adjusting only the time-sensitive components of the network. This distinction highlights a tailored transfer learning strategy for RNN in the context of state estimation in structure dynamics. By re-training only the RNN cells, the model is fine-tuned to capture the temporal dynamics characteristic of the actual structure, while the knowledge embedded in the MLP layers, likely representing more universal aspects of structural behavior, is retained without modification. This allows the TL-RNN to adapt to the specificities of the field-measured data from the actual structure without needing to re-learn fundamental patterns and relationships already established during the initial training phase on the idealized FE model. The frozen MLP layers act as a stabilizing anchor, preserving domain-general knowledge, while the retuned RNN cells focus on domain-specific temporal dependencies, ensuring both stability and adaptability. Moreover, from the practice of optimization based on Eq. (6.3), the execution of freezing MLP layers and re-training RNN cells leads to the convergence of the algorithm. This proposed strategy strikes a balance between retaining prior knowledge and accommodating new structural conditions, which is critical for reliable state estimation in real-world applications.

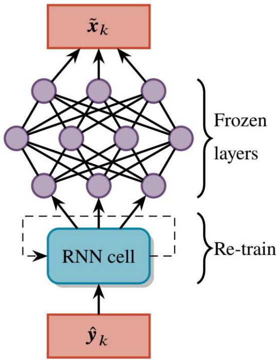


Figure 6.2. Apply transfer learning for RNN.

6.3.2. State estimation procedures using TL-RNN

The workflow of the proposed TL-RNN is shown in Figure 6.3. For the sake of convenience, the full monitoring data are divided into accelerations and additional measurements, with accelerations serving as the primary information for state estimation. The major steps are summarized as follows:

i. Idealized model construction. Construct an idealized, simplified parametric FE model to approximate the dynamic of the actual structure. Incorporate existing information to preliminarily determine the model parameters as accurately as possible.

ii. System identification and model updating. Calculate the vibrational characteristics of the actual structure through field-monitoring data. Refine the parameters of the idealized FE model to align the model results with the measurement, thereby further reducing the discrepancies between the model and the actual structure.

iii. Data preparation. Delineate a multitude of external excitation time histories as input for the calibrated FE model. Compute the resultant responses to these excitations. For each timestamp, the acceleration responses of designated DOF are extracted from the global responses. These selected acceleration responses are utilized as input to the RNN, while the system's state vector is designated as the output. These corresponding pairs of input and output are methodically compiled to constitute the training and test datasets.

iv. RNN construction, training, and validation. Build the RNN or BiRNN model based on the general architecture shown in Figure 6.1. Begin by selecting suitable hyperparameters, which should include, but are not limited to, the learning rate, number of training epochs, type of loss function and optimizer, regularization techniques and associated parameters, dimension of hidden state in the RNN cell, and the configuration of the MLP cell. Then, train the RNN with various hyperparameter combinations to minimize the training loss. Concurrently, monitor the RNN performance on the test dataset to guard against overfitting. The aim is to identify the set of hyperparameters that results in the minimum loss on the test dataset after training has converged. This optimal hyperparameter set should be considered as the one that enhances the model's generalization ability on unseen state estimation tasks.

v. RNN Transfer learning. Divide the multi-sensor measurements into two sets at each time step. For example, if there are four strain gauges on the structure, one set (the training dataset) includes the measurements from three of these gauges, while the second set (the test dataset) contains data from the remaining gauge. The corresponding acceleration data, which acts as the input to the RNN, is prepared in alignment with the strain measurements of both datasets. Then, a loss function is formulated for both datasets according to Eq. (6.3). The pre-trained RNN is fine-tuned on the training dataset following the regulation as shown in Figure 6.2, during which the hyper-parameters are adjusted. The RNN model that achieves the best performance on the test dataset is chosen as the fine-tuned model. This model is expected to fit well to measured data while retaining the essential knowledge from the pre-training phase and is then deployed for state estimation on the actual structure. The fine-tuned model should now be capable of making accurate predictions of the structural states based on the input acceleration data, informed by the multi-sensor measurements.

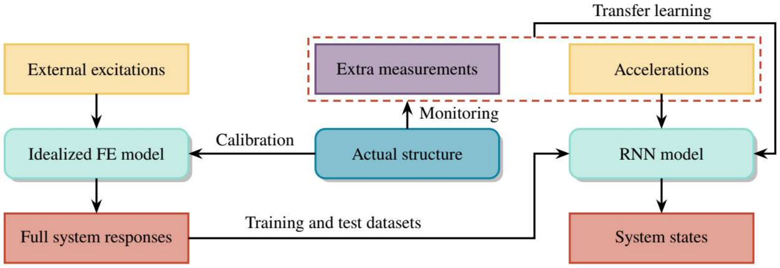


Figure 6.3. Workflow of the TL-RNN for state estimation.

6.4. Numerical example: a base-isolated building structure

6.4.1. Base-isolated building structure and simplified model

A 13-story base-isolated shear-type building shown in Figure 6.4(a) is utilized to represent the high-fidelity real-world structure, where the superstructure is simulated as a linear shear-type structure, and New Zealand (N-Z) bearings are adopted between the foundation and superstructure (Matsagar and Jangid, 2003). As shown in Figure 6.4(b), the dominant feature of the N-Z bearing is the parallel action of a linear spring, damper, and hysterics. Therefore, the Bouc-Wen model is well-suited to characterize the force-deformation relationship (Matsagar and Jangid, 2004). The restoring force F_b provided by the N-Z bearing is given by:

$$F_b = \alpha k_1 u_1 + c_1 \dot{u}_1 + (1-\alpha) F_y \nu, \tag{6.4}$$

where F_y is the yield strength of the bearing; u_1 and $\dot{u}_1$ are displacement and velocity of the base floor, respectively; α denotes the ratio of post- to pre-yielding stiffness; k_1 and c_1 represent the stiffness and viscous damping factor of the base story, respectively; ν is the dimensionless hysteric displacement component that satisfies the following nonlinear first order differential equation:

$$q\nu = A\dot{u}_1 + \beta|\dot{u}_1||\nu|^{n-1}\nu - \gamma\dot{u}_1|\nu|^n, \tag{6.5}$$

where q denotes the yield displacement; A, β, γ, and n are dimensionless quantities, among which n must be a positive integer to control the smoothness of transition from elastic to plastic response.

The equations of motion for the 13-story base-isolated structure subjected to external excitation are given by:

$$\mathbf{M}_s \ddot{\boldsymbol{u}}_s + \mathbf{C}_s \dot{\boldsymbol{u}}_s + \mathbf{K}_s \boldsymbol{u}_s = -\mathbf{M}_s \boldsymbol{r} \ddot{u}_1 + \boldsymbol{f}_s, \tag{6.6}$$

$$m_1 \ddot{u}_1 + F_b - c_2 \dot{u}_2 - k_2 u_2 = f_b, \tag{6.7}$$

where $\mathbf{M}_s$, $\mathbf{C}_s$, and $\mathbf{K}_s \in \mathbb{R}^{12\times12}$ are mass, damping, and stiffness matrices of the superstructure, respectively; $\boldsymbol{u}_s$, $\dot{\boldsymbol{u}}_s$, and $\ddot{\boldsymbol{u}}_s \in \mathbb{R}^{12}$ are relative floor displacement, velocity, and acceleration vectors with respect to the base story, respectively; $\boldsymbol{r} \in \mathbb{R}^{12}$ is a vector with all elements equal to 1; $\boldsymbol{f}_s \in \mathbb{R}^{12}$ is the external excitation vector, and $f_b \in \mathbb{R}$ denotes the external load applied to the base story. The complete parameters that define the base-isolated structural dynamic system are listed in Table 6.1 by referring to Matsagar and Jangid, (2003) and Yi and Song, (2021). The governing equations in Eqs. (6.6) and (6.7) could be explicitly formulated with the determination of these parameters. Upon defining the time-varying external forces acting on the base-isolated building, the responses of the whole system are numerically calculated step by step based on the Newmark-β method ($\beta_{\mathrm{Newmark}} = 1/6, \gamma_{\mathrm{Newmark}} = 1/12$).

The linear shear-type structure depicted in Figure 6.4(c) serves as a simplified model to approximate the vibration of the base-isolated building shown in Figure 6.5(a). Each floor of the simplified model is assigned a mass equivalent to its base-isolated building counterpart. However, replication of the base-isolated building's stiffness and damping characteristics necessitates a process of system identification and subsequent model refinement. Despite these efforts to calibrate the parameters of the simplified model, discrepancies between the model and the base-isolated building are inevitable. Consequently, reliance on the simplified model for state estimation is likely to introduce a measure of inaccuracy.

Table 6.1. Parameters for defining the base-isolated building structure.

Parameter	Symbol	Value	Unit
Mass of 1st ~13th story	$m_1 \sim m_{13}$	1.25×10^5	kg
Stiffness of 1st story	k_1	1.56×10^5	kN/m
Stiffness of 2nd~4th story	$k_2 \sim k_4$	1.44×10^5	kN/m
Stiffness of 5th~9th story	$k_5 \sim k_9$	9.60×10^4	kN/m
Stiffness of 10th~13th story	$k_{10} \sim k_{13}$	6.00×10^4	kN/m
Damping factor of 1st ~4th story	$c_1 \sim c_4$	3.60×10^2	kNs/m
Damping factor of 5th ~9th story	$c_5 \sim c_9$	2.88×10^2	kNs/m
Damping factor of 10th ~13th story	$c_{10} \sim c_{13}$	1.80×10^2	kNs/m
Yield displacement	q	0.01	m
Yield strength	F_y	1.20×10^3	kN
Ratio of post- to pre-yield stiffness	α	0.70	-
Bearing dimensionless quantity	β	0.50	-
Bearing dimensionless quantity	γ	0.50	-
Bearing dimensionless quantity	n	2	-
Bearing dimensionless quantity	A	1.00	-

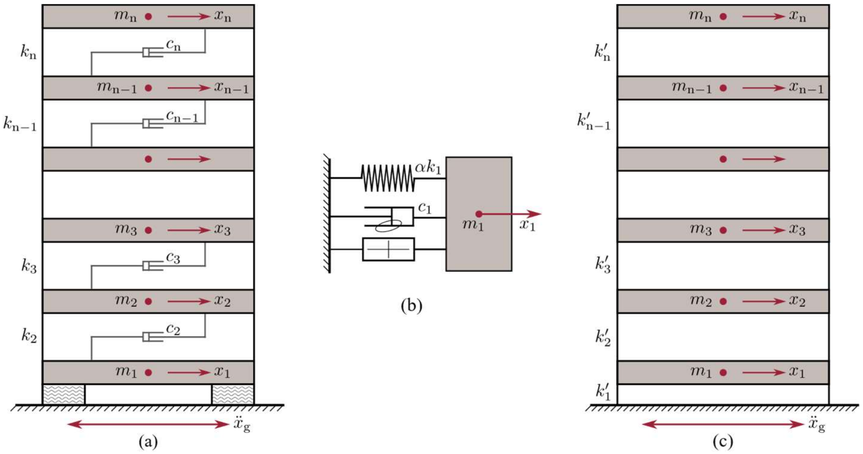


Figure 6.4. A base-isolated building structure and simplified shear-type structure.

6.4.2. System identification and model updating from ambient vibration

Stationary white Gaussian noise processes as external excitation are applied to each story of the base-isolated building to simulate the ambient vibration. These processes are synthesized by first establishing one-sided, flat PSD with a constant value of 1.44×10^4 N^2/Hz ranging from 0 to 10 Hz, then the time histories are generated by applying the inverse Fourier transform on the spectral components derived from the flat PSD function. The sampling frequency and total duration of the white Gaussian noise process are 20 Hz and 1,000 s, respectively. The system responses of the base-isolated building are computed accordingly with the excitation input to the governing equation in Eqs. (6.6) and (6.7). It is worth noting that although simulated ambient excitation could only result in a relatively low amplitude of structure vibration, e.g., the peak acceleration at the top story during the whole loading process is merely 0.065 m/s^2, the hysteresis from the bearing still occupies a certain role in contributing the base shear force. Figure 6.5 presents a comparison between the elastic force and hysteretic force in the first 100 s, indicating that the hysteretic force is not negligible even under the ambient vibration condition. Thus, there could be certain deviations between the calibrated shear-type structure and the base-isolated building.

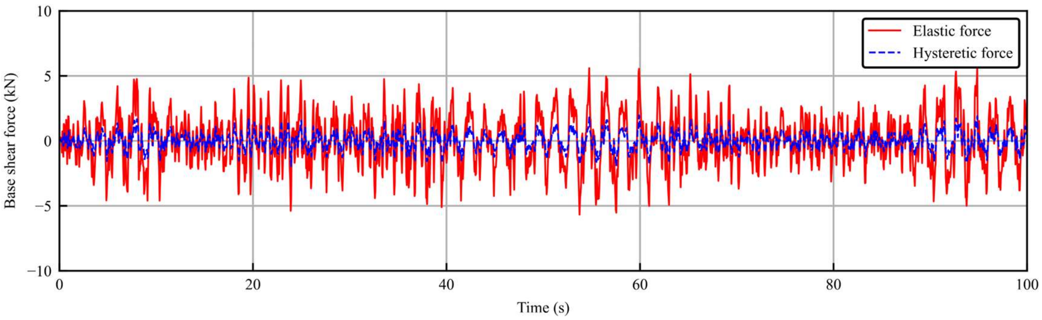


Figure 6.5. Comparison of elastic force and hysteretic force from the bearing under ambient vibration.

The output-only system identification is virtually performed on the base-isolated building. This process involves equipping each story with an accelerometer to capture the lateral acceleration. As a result, ambient vibration data for 1,000 s from every floor is collected. The collected data undergoes frequency domain decomposition (Pasca et al., 2022) to extract the building's modal properties including modal frequency, mode shape, and damping ratio for each vibration mode. Figure 6.6 presents the various orders of singular values, where the black line denotes the predominant singular value at every frequency. Modal frequencies and damping ratios are estimated by fitting the single DOF spectral bell over the black line. Mode shapes are obtained from singular vectors corresponding to the peaks in the bells. The identified first five modal frequencies and mode shapes are presented in Figure 6.7(b) and (c), and the corresponding damping ratios are respectively 2.14 %, 2.25 %, 3.05 %, 4.51 %, and 4.41 %.

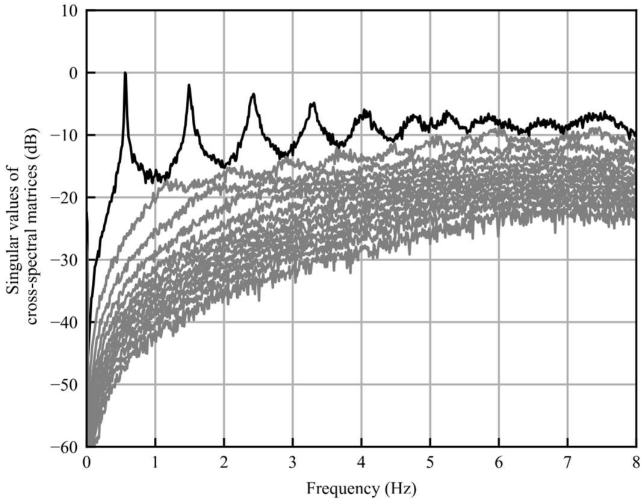


Figure 6.6. Singular values of cross-spectral matrices.

The linear shear-type structure's stiffness parameters are refined by aligning them with the modal frequencies and mode shapes obtained from measurements. To achieve the closest possible correlation between the simplified shear-type structure and the base-isolated building, a goodness-of-fit function $J(\boldsymbol{\tau})$ is minimized (Lam et al., 2015), which is given by:

$$J(\boldsymbol{\tau}) = \sum_{i=1}^{r}\left[\left(\frac{\hat{f}_i - f_i(\boldsymbol{\tau})}{\hat{f}_i}\right)^2 + \left(1 - \frac{\left|\widehat{\boldsymbol{\phi}}_i^{\mathrm{T}}\boldsymbol{\phi}_i(\boldsymbol{\tau})\right|}{\left(\widehat{\boldsymbol{\phi}}_i^{\mathrm{T}}\widehat{\boldsymbol{\phi}}_i\right)\left(\boldsymbol{\phi}_i^{\mathrm{T}}(\boldsymbol{\tau})\boldsymbol{\phi}_i(\boldsymbol{\tau})\right)}\right)\right], \tag{6.8}$$

where $r = 5$ is the total order of modal information used for model updating; $\hat{f}_i$ and $\widehat{\boldsymbol{\phi}}_i$ are the measured ith modal frequency and mode shape, respectively; $f_i(\boldsymbol{\tau})$ and $\boldsymbol{\phi}_i(\boldsymbol{\tau})$ are the corresponding model prediction results under the stiffness parameter set $\boldsymbol{\tau}$. The optimization process employs the standardized L-BFGS-B algorithm. Initial values of stiffness parameters for the optimization are based on the parameters k_1 to k_{13} listed in Table 6.1.

The optimal stiffness parameters are displayed in Figure 6.7(a). Although at the model level, the shear-type structure and the base-isolated structure differ only in the bearings at the base, this distinction is manifested in the global stiffness parameters identified during the parameter identification process. The updated parameters are observed to be slightly lower than the true parameters. This discrepancy can be attributed to the presence of hysteretic behavior in the bearings of the base-isolated structure, which absorbs a portion of the total energy. Consequently, the amount of elastic energy is reduced, leading to a decrease in the stiffness of the base-isolated structure when subjected to dynamic loads. Figure 6.7(b) and (c) illustrate the comparisons of modal frequencies and mode shapes. The red patterns represent those calculated from the updated linear shear-type structure model, while the blue patterns are obtained from the measured acceleration data on the base-isolated building. The differences between the predicted and measured modal frequencies are all smaller than 3 %, and the correlation between the predicted and measured mode shapes are all higher than 99.98 %. Therefore, the updated simplified model demonstrates a high degree of accuracy in replicating the dynamic behavior of the base-isolated structure.

For the damping characteristics of the simplified model, the Rayleigh damping approach is used. In this method, the damping matrix for the shear-type structural system is derived from a linear combination of the mass matrix and the updated stiffness matrix. The Rayleigh damping coefficients, $\alpha_{\mathbf{M}}$ for mass and $\beta_{\mathbf{K}}$ for stiffness, are determined using the equations below:

$$\alpha_{\mathbf{M}} = 4\pi f_1 f_2 \frac{\zeta_1 f_2 - \zeta_2 f_1}{f_2^2 - f_1^2}, \tag{6.9}$$

$$\beta_{\mathbf{K}} = \frac{\zeta_2 f_2 - \zeta_1 f_1}{\pi(f_2^2 - f_1^2)}, \tag{6.10}$$

where ζ_1 and ζ_2 are the damping ratios for the first two modes of vibration, while f_1 and f_2 are the corresponding measured modal frequencies in Hz.

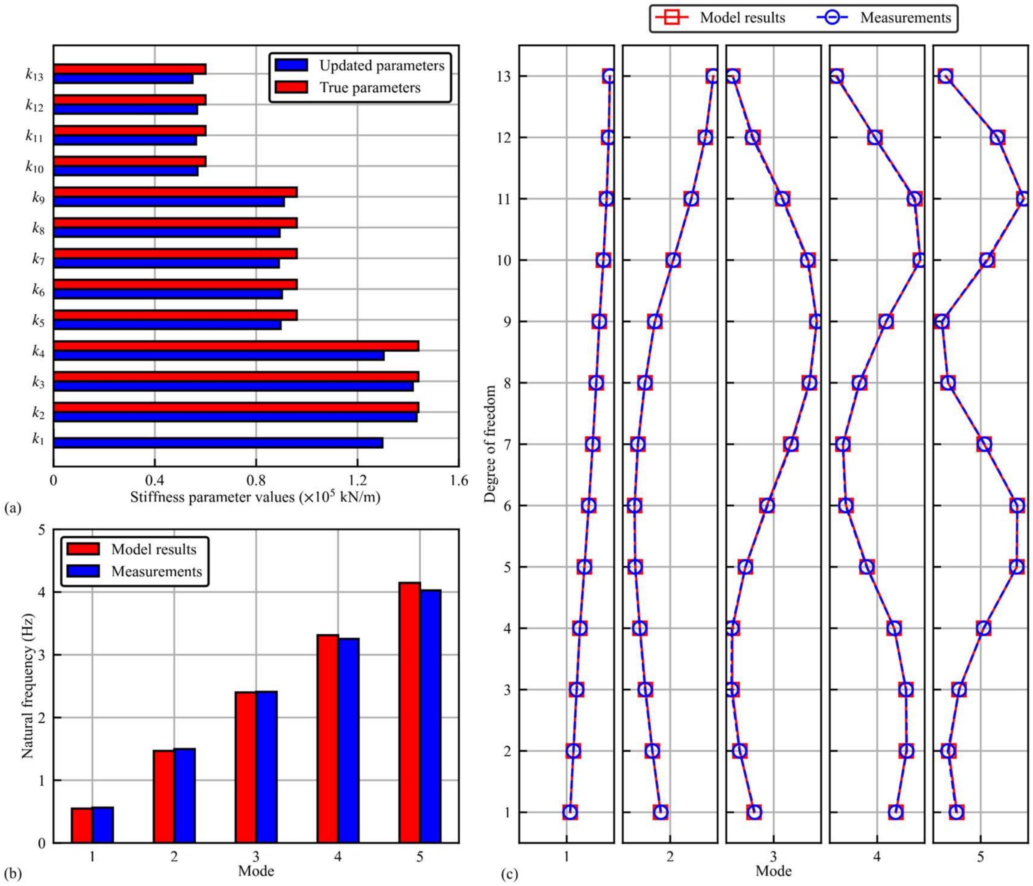


Figure 6.7. Optimal stiffness parameter values and comparison of model results with the measurements.

6.4.3. Pre-training of RNN based on the calibrated shear-type structure model

Training and test datasets for the RNN and BiRNN are created based on the calibrated shear-type structure model. To begin with, 100 robust ground acceleration time histories are synthesized to act as external excitations for the shear-type structure. Earthquake-induced ground acceleration is inherently a non-stationary process. This research synthesizes the time histories by combining stationary white Gaussian noise processes with a time-varying envelope function. These stationary processes stem from a one-sided flat PSD from 0 to 10 Hz with random phases, set to a constant value of 0.003 g^2/Hz where g is equivalent to 9.8 m/s^2. This setting ensures that the generated random excitations have a standard deviation of approximately 0.17 g, aligning with the procedure in Kalman filter-based methods where a prior assumption about the external excitations must be determined. Each stationary process is sampled at a frequency of 20 Hz and spans a total duration of 40 s. The envelope function $I(t)$ that imparts non-stationarity is given by the following equation:

$$I(t) = \begin{cases} t/(b_1 t_{\max}) & t/t_{\max} \le b_1 \\ 1 & b_1 < t/t_{\max} \le b_2, \\ \exp[-\kappa(t/t_{\max} - b_2)] & t/t_{\max} > b_2 \end{cases} \tag{6.11}$$

Where b_1, b_2, and κ are parameters drawn from uniform distributions with $b_1 \in [0.1, 0.2]$, $b_2 \in [0.4, 0.6]$, and$\kappa \in [3, 5]$. The variable $t_{\max}$represents the total duration of each time history, i.e., 40 s. Figure 6.8(a) illustrates a representative synthesized ground acceleration alongside the corresponding window function. Figure 6.8(b) displays the continuous wavelet transform results of a sample, confirming the non-stationary nature of the synthesized time history.

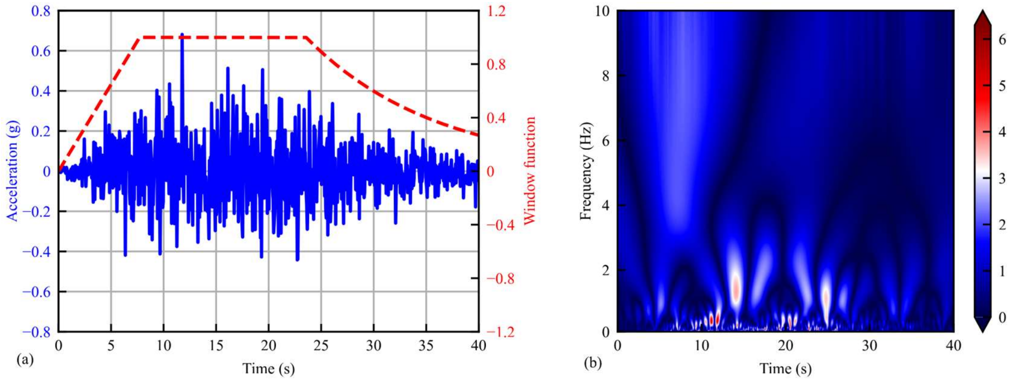


Figure 6.8. A sample of synthesized ground acceleration, window function, and continuous wavelet transformation of the ground acceleration.

The acceleration, velocity, and displacement responses of the idealized model subjected to 100 synthesized ground motions are calculated and retained. Assuming that accelerometers are mounted on stories 1-5, acceleration responses from these stories are extracted from the overall response. These selected acceleration responses serve as input data for RNN and BiRNN, with the state vector defined as the output target. These input-output pairs are compiled to create the training and test datasets. Specifically, the training dataset is composed of 90 batches of sequences, while the test dataset includes the remaining 10 batches.

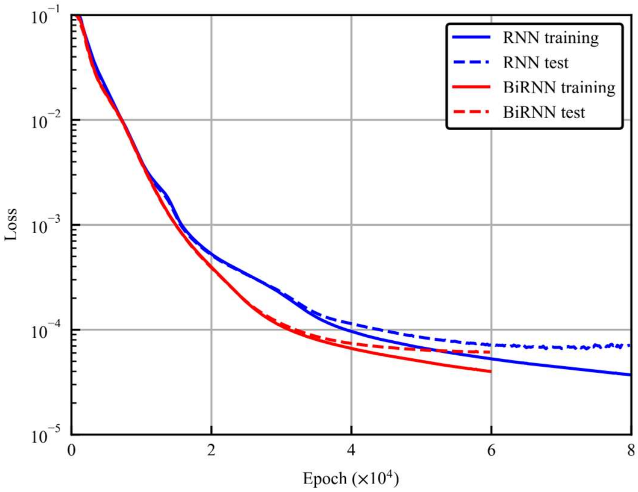


Figure 6.9. Loss curves from the pre-training of RNN and BiRNN for state estimation of the shear-type structure.

RNN and BiRNN models are established according to the architecture shown in Figure 6.1. Multiple training iterations are performed for each model to determine the optimal hyper-parameters that yield the best test dataset performance. The mean square error function is employed as the loss function for both models, with the Adam optimizer used for optimization. The hidden state dimension for both models is set to 30. Within the MLP cell, there are 4 layers, each hidden layer being equipped with 26 neurons, and the Tanh activation function is utilized. The RNN and BiRNN models contain 3,858 and 5,688 trainable parameters, respectively. For optimal performance, the training process involves 80,000 epochs for the RNN and 60,000 epochs for the

BiRNN. Additionally, their training configurations differ in the learning rates: the RNN is set at 1×10^{-5} and the BiRNN at 8×10^{-6}. The training of RNN and BiRNN is based on a personal computing system with an AMD Ryzen 3700X 8-Core Processor, a NVIDIA GeForce GTX 1660 Super graphical card, and 32 GB of RAM. On the software front, the models are implemented using Python 3.10.14 with Pytorch 2.3.0 serving as the deep learning framework. The training durations are recorded to be 364.0 s for the RNN and 334.1 s for the BiRNN. Figure 6.9 shows the loss curve from the training of both the RNN and the BiRNN, illustrating that by the end of the training, the BiRNN's performance is better than that of the RNN.

The performance of the trained RNN and BiRNN models can be assessed by visualizing their state estimation outputs on the test dataset and comparing these outputs with references. For comparison purposes, two established methods, AKF (Lourens et al., 2012) and DKF (Eftekhar Azam et al., 2015), are employed to address the same state estimation problem. The hyperparameters in the DKF and AKF algorithms, which include parameters for process noise, measuring noise, and forces modeled as zero-mean GP, are tuned using the training dataset to ensure optimal performance. When fed with identical acceleration data from the test dataset, both DKF and AKF are tasked with performing state estimations comparable to those of the RNN models. The outcomes of this comparative analysis are depicted in Figure 6.10, highlighting the estimated displacement at the 9th story and velocity at the 13th story. The preliminary visualization of the state estimation data indicates that the four approaches, RNN, BiRNN, DKF, and AKF, deliver acceptable results when applied using the calibrated idealized model.

The NRMSE is utilized to quantitatively compare the prediction accuracy to the actual reference values. Eq. (6.12) illustrates the NRMSE formula, where $\tilde{\boldsymbol{x}}_i$ represents the predicted state sequence (displacement or velocity) at the ith story and $\boldsymbol{x}_i$ is the reference sequence. The notation $\|\cdot\|_2$ signifies the 2-norm, or Euclidean norm. A lower NRMSE value indicates a higher prediction accuracy.

$$\text{NRMSE}(\boldsymbol{x}_i, \tilde{\boldsymbol{x}}_i) = \frac{\|\boldsymbol{x}_i - \tilde{\boldsymbol{x}}_i\|_2}{\|\boldsymbol{x}_i - \text{mean}(\boldsymbol{x}_i)\|_2}. \tag{6.12}$$

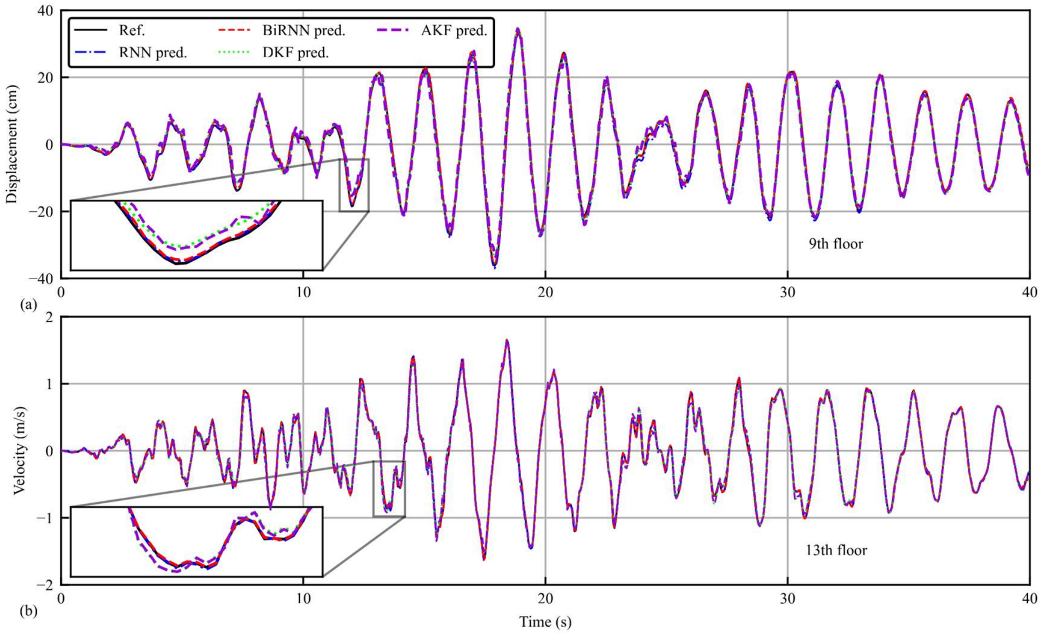


Figure 6.10. Comparison of state prediction samples from RNN, BiRNN, DKF, and AKF models; these predictions are based on acceleration inputs from a shear-type structural model and are benchmarked against the reference output of the same model.

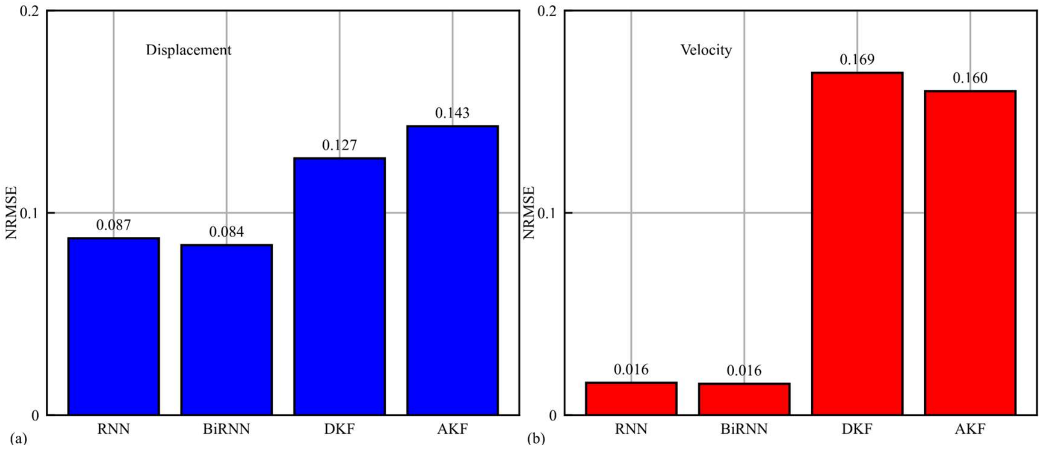


Figure 6.11. Averaged NRMSE of RNN, BiRNN, DKF, and AKF models against the shear-type structural model for displacement and velocity predictions.

Within the test dataset, there are 10 sequence batches, with each containing 13 displacement and 13 velocity sequences. An NRMSE score is computed for each sequence. The average NRMSE scores for displacement and velocity predictions are calculated separately. The results, presented in Figure 6.11, indicate that the average NRMSEs for both the RNN and BiRNN models are lower than those for two Kalman filter-based methods. Furthermore, the BiRNN model exhibits a marginally better performance compared to the RNN model. Notably, for both RNN-based models, the predictions for velocity are more accurate than those for displacement, which contrasts with the results from the Kalman filter-based methods.

6.4.4. TL-RNN for state estimation of base-isolated building under seismic excitation

The calibrated shear-type structure-based RNN and BiRNN models are employed for state estimation of the base-isolated building under seismic excitation. To begin, ground motion histories from four distinct regions are collected and normalized using specific scaling factors, as illustrated in Figure 6.12. These ground motions are much stronger than the synthesized ambient excitations, and in such cases, they can excite a higher portion of nonlinear dynamic responses of the base-isolated building which might not be reflected in the linear shear-type structure. Moreover, the records display variations from synthesized ground acceleration in peak values, duration, and PSD. Nevertheless, the bulk of the energy is concentrated below 10 Hz, which is similar to that of the synthesized acceleration for training of RNN. Subsequently, the responses of the base-isolated building to these real-world strong ground motions are computed. Accelerations from the 1st to the 5th floors are extracted from the responses to serve as input to the pre-trained RNN and BiRNN. The states from these calculated responses provide a benchmark for assessing the performance of the neural networks. Concurrently, the DKF and AKF given in the previous subsection are applied as well, utilizing the same acceleration data to conduct state estimation for the building.

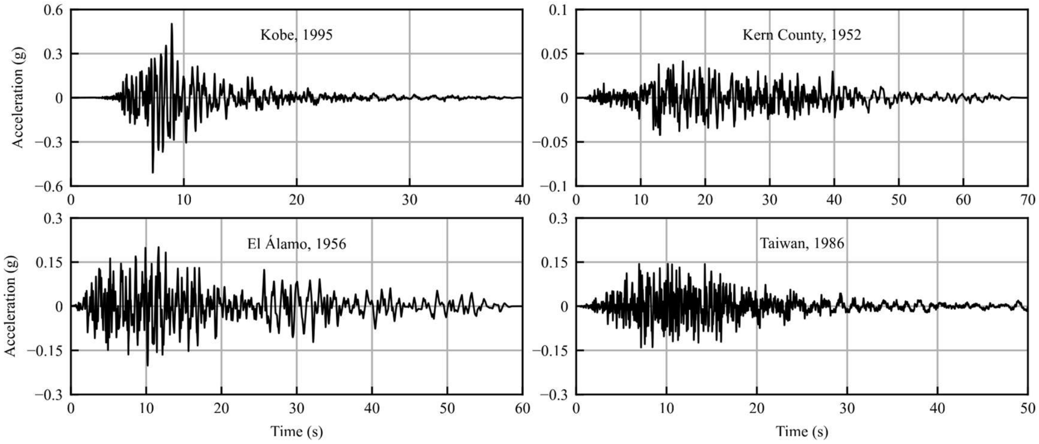


Figure 6.12. Ground motion records from Kobe, Kern County, El Álamo, and Taiwan earthquakes.

A sample of the velocity estimation results is depicted in Figure 6.13, illustrating that all four methods yield satisfactory performance. For displacement estimation, a sample of results from the four methods is presented in Figure 6.15, revealing that each method exhibits varying degrees of deviation in displacement prediction. The BiRNN demonstrates the least deviation, while the RNN shows the most. This discrepancy in prediction is attributed to the inherent differences between the simplified shear-type structure and the base-isolated building structure, even though the simplified structure has been calibrated with the monitoring data. The RNN and BiRNN are trained and tested using data derived from the simplified structure, while the DKF and AKF are directly deduced from the motion equations of the simplified structure. Therefore, the introduction of the TL-RNN framework is essential to enhance the precision of state estimation.

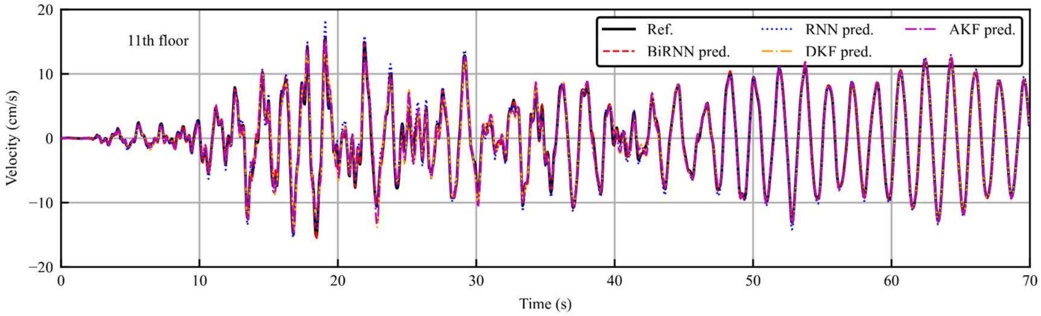


Figure 6.13. Velocity estimation results of the 11th floor of the base-isolated structure subject to the Kern County earthquake.

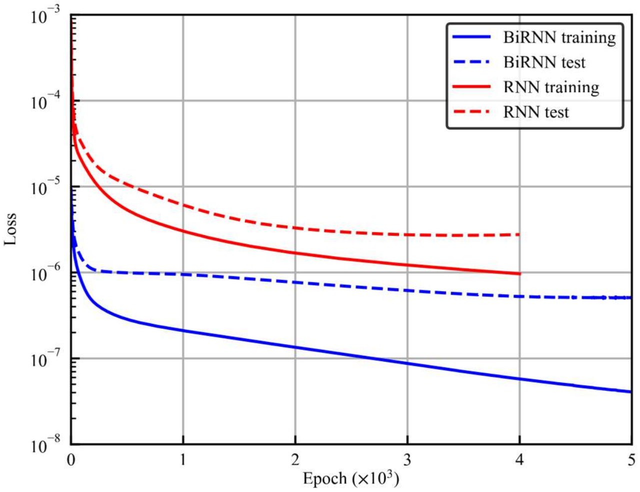


Figure 6.14. Loss curves from the RNN and BiRNN transfer learning under the Kobe earthquake.

In this study, inter-story displacements (Chang and Huang, 2020) between adjacent floors from the 1st to 5th floors/DOF are assumed to be monitored during the earthquake. Therefore, there are extra measurements regarding 4 inter-story displacements beyond the acceleration measurements. The extra measurements are divided into two sets at each time step, one set (the training dataset) contains information on 1-2, 2-3, and 3-4 inter-story displacements, while the second set (the test dataset) contains the 4-5 inter-story displacement. The RNN and BiRNN are fine-tuned based on the training and test datasets. The parameters in the MLP cells are kept constant, and only those on the RNN cells are optimized based on Eq. (6.3). In total, there are 1050 parameters for RNN and 2100 parameters for BiRNN to be re-trained. The learning rate is set as 1×10^{-5}. The training and test loss curves corresponding to the ground motion in the case of Kobe earthquake are shown in Figure 6.14, where the test loss from BiRNN is lower than that of RNN by the end of transfer earning. In comparison with the training of RNN and BiRNN, the fine-tuning process takes much fewer epochs to converge. Under the same computing environment, for the RNN model, the time cost for transfer learning is 43.9 s, 105.0 s, 68.1 s, and 70.1 s for the four ground motion cases, respectively; for the BiRNN model, the time cost is 61.6 s, 160.2 s, 88.67 s, and 56.47 s, respectively.

For comparison purposes, the extra measurements are combined with acceleration measurements to serve as input to the DKF and AKF for state estimation, resulting in the methods termed Integrated DKF and AKF. The procedures for implementing integrated DKF and AKF are similar to those of the traditional DKF and AKF. It is important to note that TR-RNN and integrated prediction methods fundamentally differ: the TL-RNN modifies the pre-trained state estimation RNN model by using the extra measurements, whereas the integrated prediction methods provide the pre-trained state estimation model with additional reference data for forecasting.

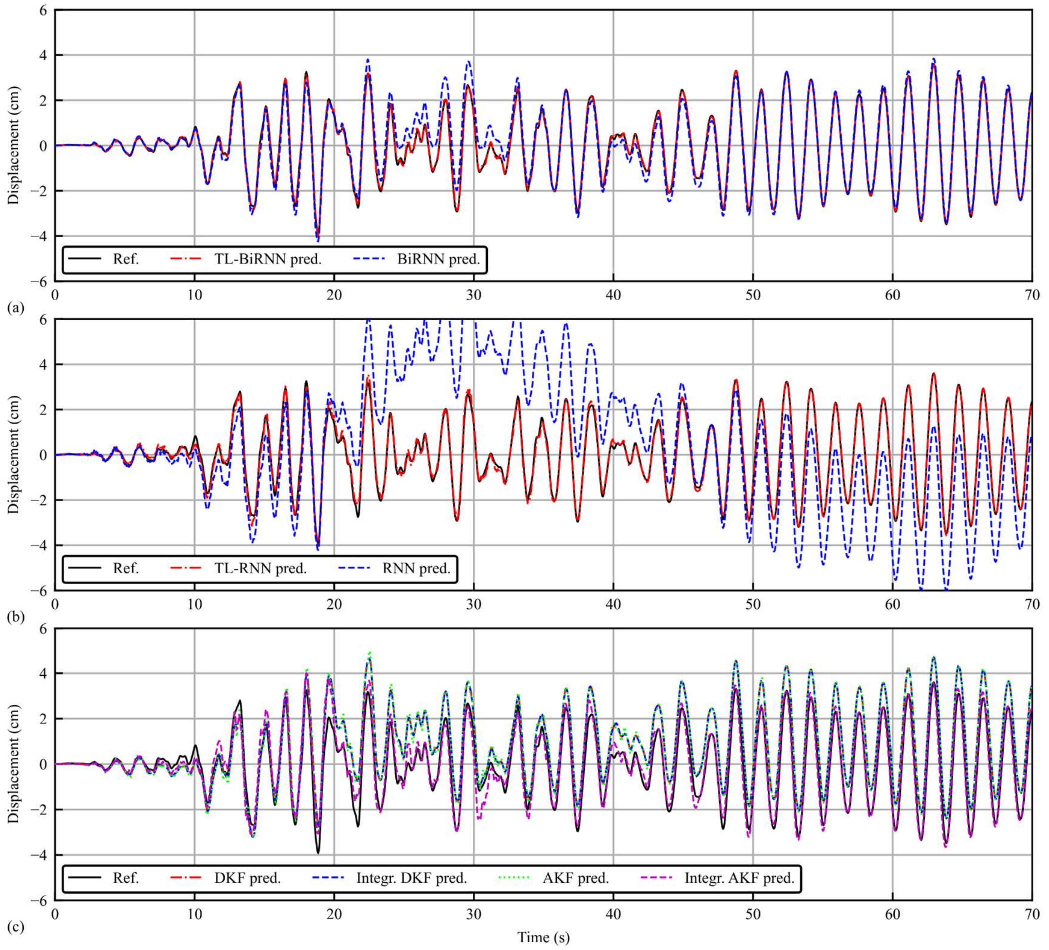


Figure 6.15. Displacement estimation results of the 11th floor of the base-isolated structure subject to the Kern County earthquake.

The displacement prediction samples for the base-isolated building using the extra measurements are shown Figure 6.15. Due to the inherent limitation of the DKF algorithm, the inclusion of extra information does not alter the DKF predictions, whereas the other three methods experience varying degrees of precision enhancement. The transfer learning of bidirectional recurrent neural network (TL-BiRNN) outperforms the other methods, followed by the TL-RNN, and then the Integrated AKF. Even though the extra measurements have enhanced the predictive capability of the AKF, this enhancement is often evident in corrections to the overall trend. During strong seismic events, its state predictions still exhibit significant discrepancies when compared with reference data. In contrast, the TL-RNN and TL-BiRNN maintain a certain level of predictive accuracy throughout the entire seismic event. This may be due to the pronounced nonlinearity exhibited by the base-isolated building under earthquake conditions, particularly in the interval between 10 s and 30 s, as observed in the Kern County earthquake. Such nonlinearity cannot be represented by the calibrated shear-type structure. Therefore, the refinement of the state prediction model itself (i.e., the transfer learning) based on additional measurement data is necessary to achieve better predictive accuracy. In summary, when the structural model does not align with the actual conditions, the accuracy of state estimation can be significantly improved by fine-tuning the RNN model using real-world measurements. This method of refinement yields better results than simply incorporating extra measurements into the existing model as additional inputs for prediction.

Quantitative evaluations of model performance with the inclusion of extra measurements are carried out using the NRMSE metric. In this context, the assessment of the integrated DKF predictive results is omitted, as additional measurements do not impact its predictions. For each of the four strong ground motion cases, NRMSE for all predicted displacement sequences against their corresponding references are calculated. These NRMSE values are then averaged for each case, and the results are presented in Figure 6.16. In cases where the simplified model diverges from the actual structure, the proposed TL-RNN framework has proven to be effective in ensuring the accuracy of state estimation. Following the application of transfer learning, the performance of the BiRNN is outstanding across all four cases, while the TL-RNN also demonstrates a commendable level of accuracy. On the other hand, simply supplying the AKF with additional input information falls short of enhancing state estimation accuracy, yielding much greater NRMSE values than those from the TL-RNN and TL-BiRNN models.

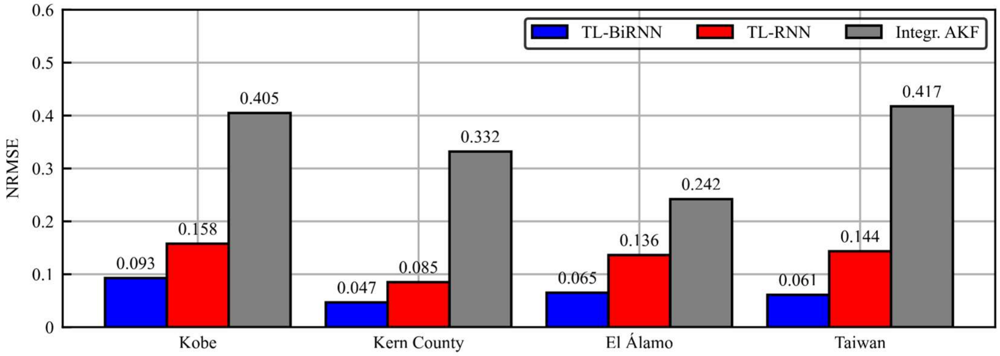


Figure 6.16. Averaged NRMSE of TL-BiRNN, TL-RNN, and Integrated AKF for displacement prediction under the four earthquakes.

6.5. Experimental example: A two-span continuous beam

6.5.1. Measurement setup

In the experiment, the investigation focuses on a 1.26 m two-span aluminum beam that is continuously supported. Figure 6.17 presents an overview of the experimental setup. The primary components of this setup are numbered from ① to ⑥, including fiber Bragg gratings (FBG), bearings with bolted connections, a load cell, accelerometers, a shaker, and a laser distance sensor. From left to right, the three FBGs and three accelerometers are abbreviated as FBG1-FBG3 and A1-A3. Specifically, FBG1-FBG3 have wavelengths of 1563 nm, 1555 nm, and 1527 nm, respectively, with a grating length of 5 mm, a bandwidth of less than 0.4 nm, a side-mode suppression ratio greater than 12 dB, and a reflectivity of over 75 %. The accelerometers (A1-A3), load cell, and laser distance sensor are all piezoelectric types that convert collected voltage signals into data. The shaker, connected to a power amplifier and a waveform generator, is controlled by digitally-generated voltage signals from a computer to deliver continuous stochastic excitations at a specific location on the beam. Concurrently, the sensor array is designed to continuously record original data at their respective locations with a 5000 Hz sampling frequency. As a result, at each time step, the system captures measurements from the three FBG and three accelerometers, along with load measurement from the load cell and deflection measurement from the laser distance sensor. This array of sensors ensures comprehensive monitoring and accurate state estimation of the beam.

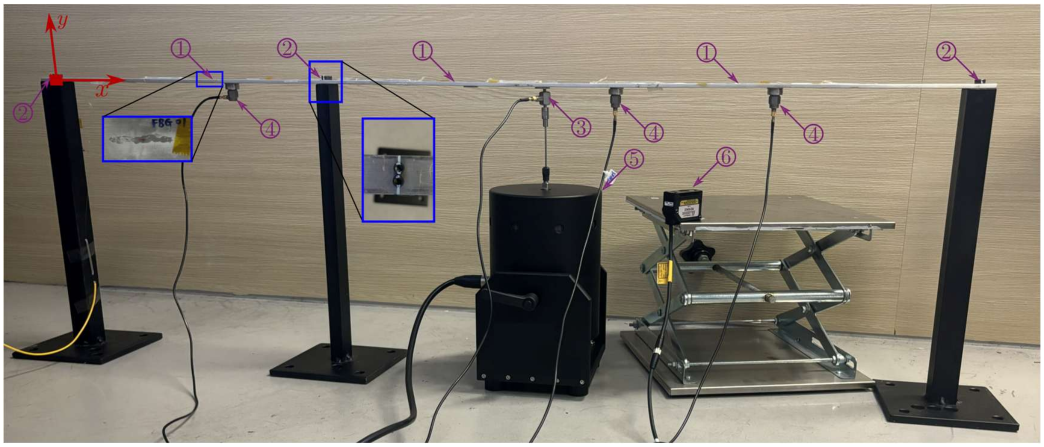


Figure 6.17. Overview of the experimental setup where key components are highlighted and numbered: ① FBG, ② bearings with bolted connections, ③ load cell, ④ accelerometers, ⑤ shaker, and ⑥ laser distance sensor.

6.5.2. Model calibration

In this experiment, the bolts at both ends of the beam are securely tightened, while those in the middle are deliberately left somewhat looser. This arrangement introduces controlled imperfections in the boundary conditions of the beam. Figure 6.18 illustrates the idealized model of the continuous beam, detailing the structure and instrument configurations. As shown in Figure 6.18, the two ends of the beam are transversely fixed, while the rotations of the three supports are constrained by rotational springs with stiffness values$k_{\theta 1}$, $k_{\theta 2}$, and$k_{\theta 3}$. Additionally, a linear transversal spring is added to the middle support, denoted by its stiffness k_w. In this setup, the Euler-Bernoulli beam element is utilized for the FE model, incorporating two DOF at each node, i.e., the transversal deflection and rotation. Each element is 2 cm long, resulting in a total of 63 elements, 64 nodes, and 126 DOF for the FE model. To fully define the FE model, except for the geometric parameters, the elastic modulus E, density ρ, rotational spring parameters$k_{\theta 1}$to $k_{\theta 3}$, and transversal spring parameter k_w must be determined based on the monitoring data from the beam. Moreover, the damping model employed is Rayleigh damping, with the assumption that the first and second modes share the same damping ratio. Hence, the single parameter damping ratio ζ for these two modes needs to be calibrated as well.

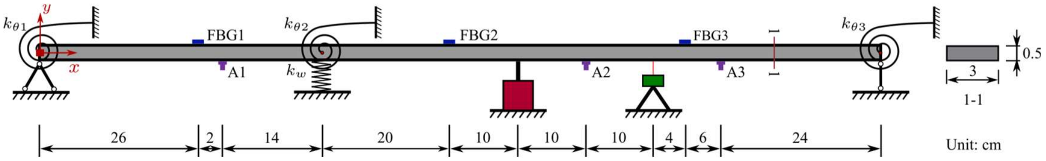


Figure 6.18. Idealized model for the beam with instruments positioning.

An experimental modal test is conducted on a beam. A digital voltage waveform, set to a 2 V peak-to-peak amplitude, is generated by a computer and input into a waveform generator. The resultant output is a band-limited white noise signal, confined to a frequency range of 10 Hz to 100 Hz. The duration of this signal is fixed at 25 seconds. The power amplifier settings are adjusted so that the standard deviation of the output force is approximated to around 1 N. Concurrently, accelerations at points A1 to A3 on the beam are recorded. Figure 6.19 displays the measured acceleration and force data over a 1 s interval.

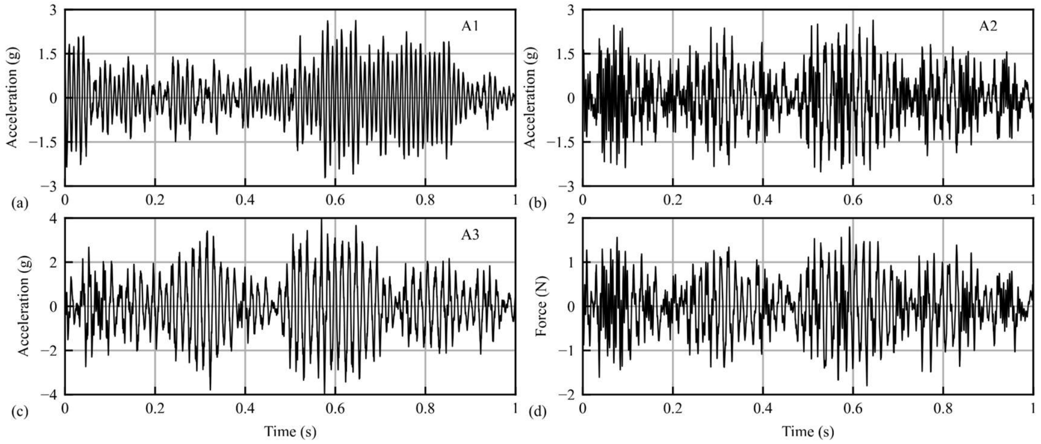


Figure 6.19. Measured acceleration and force data in experimental modal testing.

Table 6.2. Nominal values and scaling factors for parameters in the beam model.

Parameter	Symbol	Nominal value	Scaling factor	Unit
Elastic modulus	E	6.85×10^{10}	0.957	N/m^2
Density	ρ	2.70×10^{3}	1.278	Kg/m^3
Rotational stiffness of left support	$k_{\theta 1}$	1.00×10^{4}	1.487	N/rad
Rotational stiffness of middle support	$k_{\theta 2}$	1.00×10^{1}	0.313	N/rad
Rotational stiffness of right support	$k_{\theta 3}$	1.00×10^{4}	0.902	N/rad
Transversal stiffness of middle support	k_w	1.00×10^{5}	2.951	N/m
Damping ratio of 1st and 2nd mode	ζ	8.00×10^{-3}	1.415	-

With all the collected data over the 25 s, averaged Fourier transforms of excitation and acceleration signals are computed using a sliding window of 5 s per block. Subsequently, the FRF, denoted as $\widehat{H}_i(f)$ where $i = 1,2,3$ and $f = 0.2k$ (with $k \in \mathbb{N}_0$ and $k < 1.25 \times 10^4$), are evaluated via the ratio of the Fourier transform results of acceleration to excitation. Only the FRF data ranging from 10 Hz to 100 Hz is utilized for updating the parameters of the FE model (Imregun et al., 1995). The optimization problem is formulated as follows:

$$\min \sum_{k=50}^{500} \sum_{i=1}^{3} \left[\left(\log \frac{|\hat{H}_i(0.2k)|}{|H_i(0.2k;\boldsymbol{\eta})|} \right)^2 + \left(\angle \hat{H}_i(0.2k) - \angle H_i(0.2k;\boldsymbol{\eta}) \right)^2 \right]. \tag{6.13}$$

Here, $\boldsymbol{\eta} = \{E, \rho, k_{\theta1}, k_{\theta2}, k_{\theta3}, k_w, \zeta\}$ represents the set of parameters to be updated; $|\cdot|$ and $\angle$ denote the magnitude and phase operators, respectively; $H_i(0.2k;\boldsymbol{\eta})$ is the ith FRF predicted from the FE model given $\boldsymbol{\eta}$. The L-BFGS-B algorithm is employed for optimization, and it converges with the function values in Eq. (6.13) being 0.148. The nominal values and optimal scaling factors for each element in $\boldsymbol{\eta}$ are presented in Table 6.2. Visualizations comparing the single-sided FRF from the updated model and measurements are shown in Figure 6.20. The results indicate that the principal trends between the model and measurements are almost identical, although some discrepancies exist, e.g., the first modal frequency being higher than measured, and the second being lower. Nevertheless, the calibrated model closely approximates the real structure, and it will be used for further state estimation tasks in the following subsections.

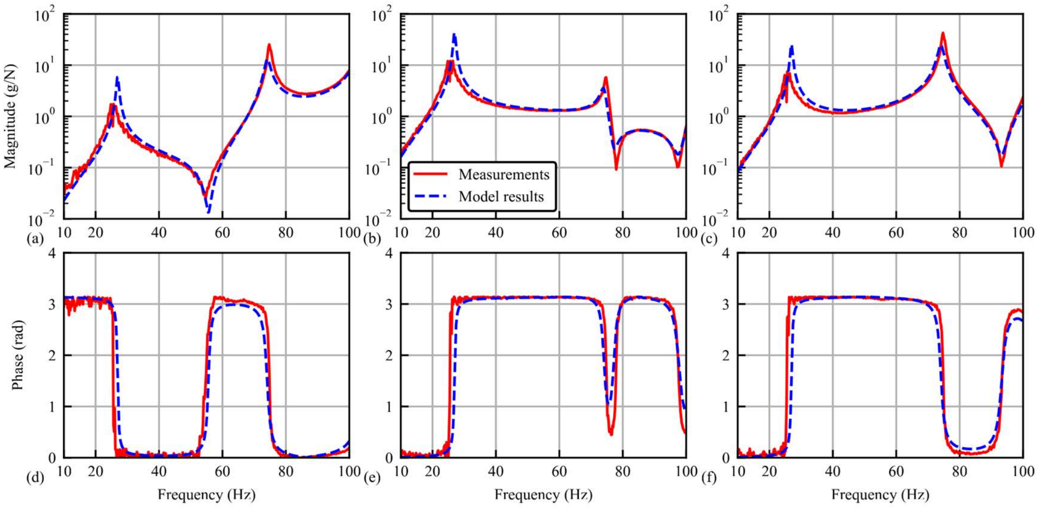


Figure 6.20. Measured and calibrated model output single-sided FRF.

6.5.3. Pre-training of RNN based on the calibrated beam model

The training of the RNN and BiRNN follows the procedure outlined in Figure 6.21. Initially, a stochastic force generator is programmed to produce white noise force histories within the frequency range of 10 Hz to 410 Hz, based on a flat PSD. In this process, only two force histories are used: one for the training dataset and one for the test dataset, both with a standard deviation of 0.5 N. These force histories are then fed into the calibrated FE model, which computes the system response, including acceleration, velocity, and displacement at each DOF. The duration for each batch of response is 4 s, with a sampling frequency of 5000 Hz. That is, there are 20,000×3 acceleration data fed into the RNN model, and the model will output 20,000×252 data of structural state. To enhance the convergence of RNN training, all response data are normalized. The normalized displacement and velocity data are treated as labeled states at each node. Noise is subsequently added to the acceleration data at locations A1 to A3, which is similar to priors embedded in noise as in Kalman filter-based methods. In this case study, the noise-to-signal ratio is presumed to be 10 %. The noisy, normalized acceleration data are used as input to the RNN models, enabling the RNN to predict states at each node of the FE model. Accordingly, the mean square error loss function is formulated based on the predicted state and the labeled state, and the Adam optimizer is utilized to efficiently tune parameters in both RNN and BiRNN models.

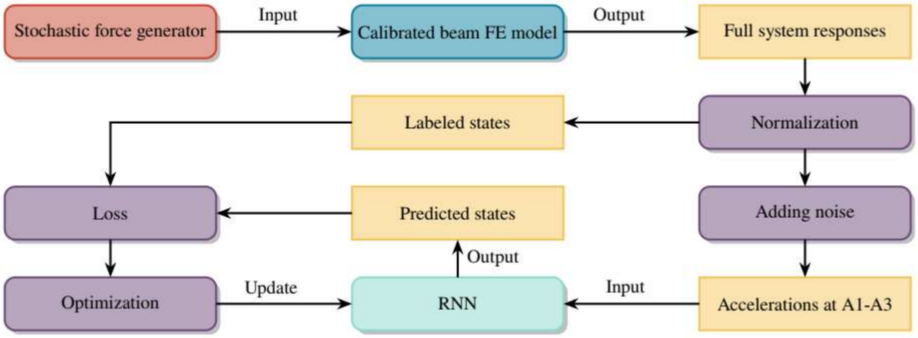


Figure 6.21. Training procedures of RNN and BiRNN for the beam.

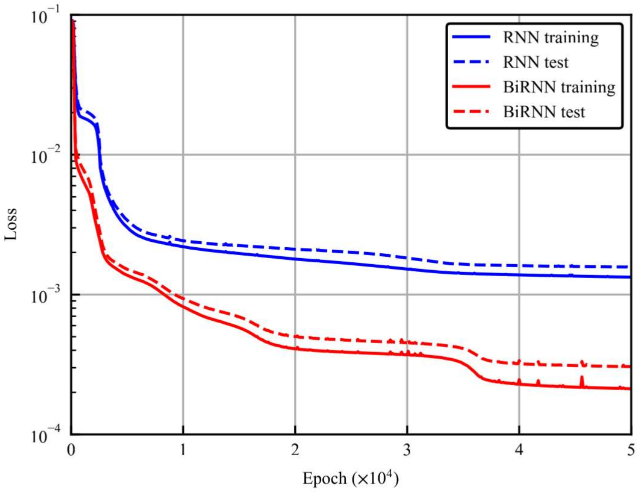


Figure 6.22. Loss curve from the training of RNN and BiRNN for the beam.

The hidden state dimension is uniformly set at 36 for both the RNN and BiRNN models. Each model includes an MLP component with three layers. The RNN configuration features 36 neurons in each of the first two layers and 126 neurons in the third layer. In contrast, the BiRNN is equipped with 72 neurons in each of the first two layers and 126 neurons in the third layer. Biases in the propagation paths of both the RNN and BiRNN have been removed. In total, there are 8,532 parameters in RNN and 22,248 parameters in BiRNN to be trained. Both RNN and BiRNN models are trained over 50,000 epochs using a consistent learning rate of 3×10^{-5}. The training and test loss curves, presented in Figure 6.22, demonstrate that the RNN model's test loss converges at 1.33×10^{-3}, while the BiRNN model's test loss converges at 2.12×10^{-4}. These results indicate that, with the simple test dataset generated from the calibrated model, the BiRNN significantly outperforms the RNN in terms of state estimation accuracy.

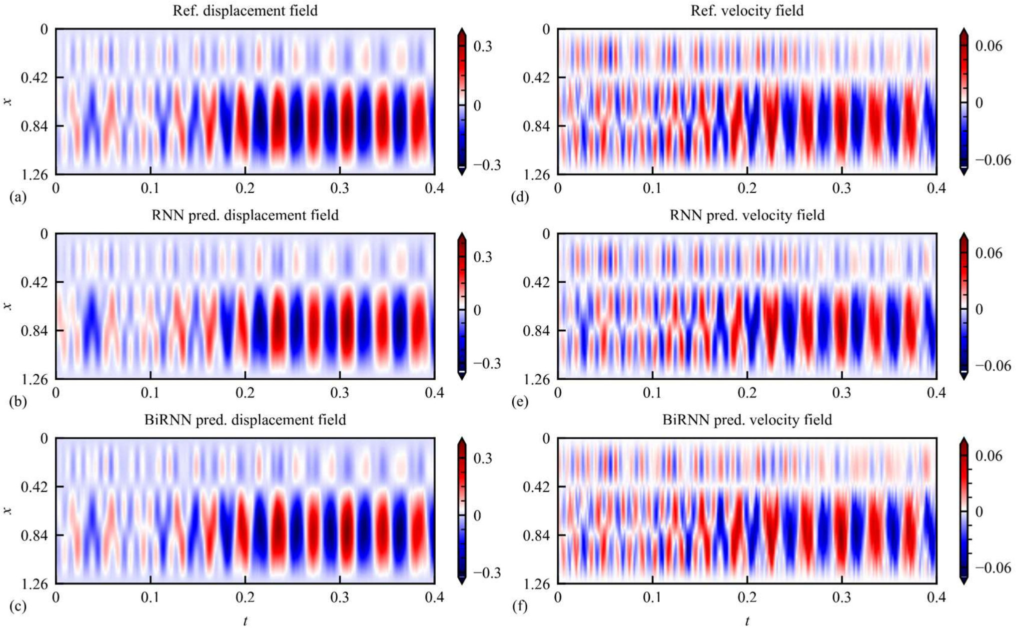


Figure 6.23. Predicted displacement and velocity field over the beam in 0.4 s.

To visualize the models' performance, the predicted displacement and velocity fields over 0.4 s are displayed for the test dataset in Figure 6.23. The horizontal axis represents time, while the vertical axis corresponds to the spatial coordinates of the beam, illustrating the full range of states. Both models yield satisfactory state estimations. Figure 6.24 further illustrates the predicted deflection at 0.76 m and rotational speed at 0.52 m during the initial 2 s of the test period. While the RNN model shows slight deviations in amplitude from the reference data, the BiRNN model displays superior accuracy and closer adherence to the expected outcomes.

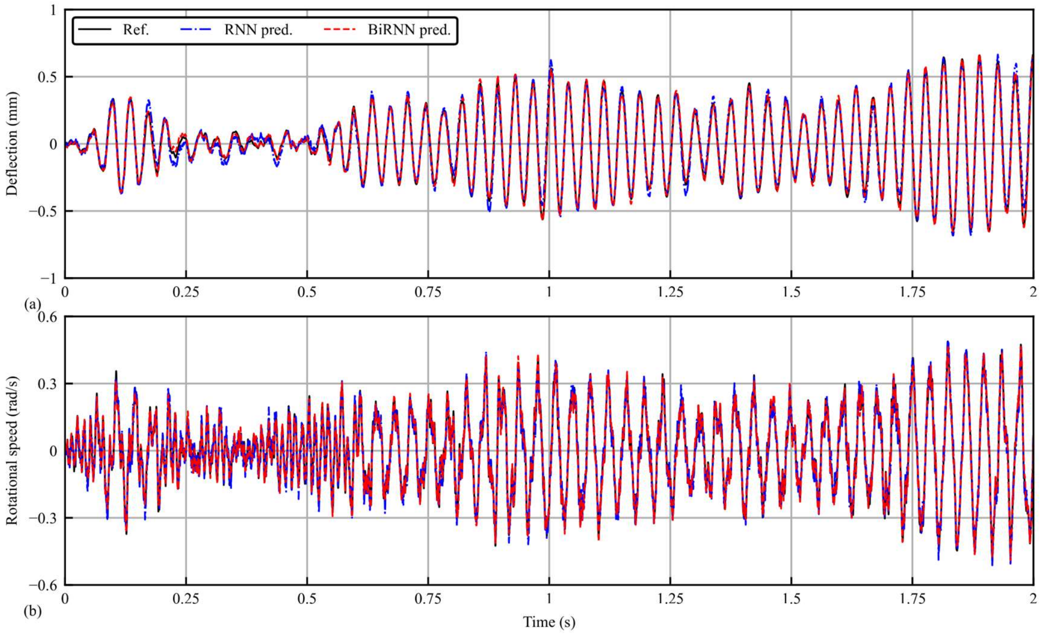


Figure 6.24. Prediction results of deflection at 0.76 m and rotational speed 0.52 m of the beam.

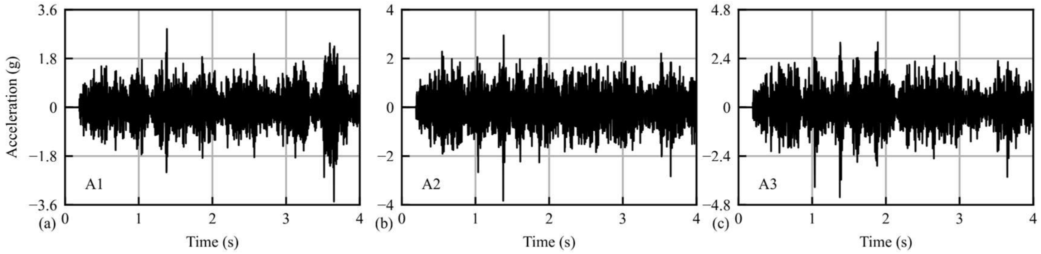


Figure 6.25. Measured acceleration data from A1 to A3.

The acceleration data measured from sensors A1 to A3, as displayed in Figure 6.25, are fed into the well-trained RNN and BiRNN models. It is important to note that the acceleration data presented are not as ideal as those in the training and test datasets; the noise, in this case, can be categorized into two distinct types: noise originating from the voltage signal and noise resulting from the discrepancy between the model and the actual structure. Furthermore, Figure 6.25 demonstrates that there could be some outliers distorting the signal (Gul and Catbas, 2009). These imperfections present a significant challenge to accurate prediction of the beam's system state by the RNN and

BiRNN models. To quantitatively assess the performance, the deflection measured by a laser distance sensor at a distance of 92 cm is used for comparison with the results from the RNN and BiRNN models. A preliminary laboratory test revealed that this laser distance sensor could only measure vibrations up to approximately 35-40 Hz.

Consequently, a second-order low-pass Butterworth filter with a cutoff frequency of 38 Hz has been implemented to filter the outputs from the RNN and BiRNN. Performance is evaluated using NRMSE, detailed in Eq. (6.12). The NRMSE for the RNN is 0.545, while for the BiRNN, it is 0.221, showing that the BiRNN model achieves superior prediction accuracy over the RNN model. Such differences have been reflected during the training of RNN and BiRNN models using the training set. When the input acceleration data is subject to noise interference, the bidirectional propagation of information and computation of the state can effectively reduce such interference. In contrast, the unidirectional approach appears to be less effective. However, the unidirectional propagation allows the RNN model to perform online state estimation, which is not achievable with the BiRNN.

6.5.4. Using transfer learning to enhance model performance

The sensor array, as illustrated in Figure 6.17, captures comprehensive data during the vibration process. This includes not only acceleration and deflection measurements but also strain data from three specific locations on the beam. The strain information, in conjunction with the measured acceleration data, is valuable for the transfer learning process discussed in this subsection.

Let $\hat{\boldsymbol{\epsilon}}_k \in \mathbb{R}^2$ denote the measured strain from FBG1-FBG2, and $\tilde{\boldsymbol{x}}_k \in \mathbb{R}^{252}$ represent predicted state by the pre-trained RNN, where $k = 1,2, \dots, N$, and N is the total number of timestamps. The predicted state vector $\tilde{\boldsymbol{x}}_k$ can be linearly transformed into the predicted strain vector $\tilde{\boldsymbol{\epsilon}}_k$ using the linear operator $\mathbf{B}_\epsilon \in \mathbb{R}^{2\times 252}$. The operator, which is derived from the second derivatives of the four shape functions associated with the Euler-Bernoulli beam element, relates solely to the geometry of the beam and the location of FBGs. The loss function for transfer learning can be formulated as:

$$\min \sum_{k=1}^{N} \mathcal{L}(\mathbf{B}_\epsilon \tilde{\boldsymbol{x}}_k, \hat{\boldsymbol{\epsilon}}_k). \tag{6.14}$$

During the training process, the parameters of the pre-trained RNN are fine-tuned. The training dataset incorporates measured strain data from FBG1 and FBG2, while strain data from FBG3 is reserved for the test dataset to prevent overfitting in the transfer learning process. Figure 6.26 illustrates the procedures for implementing RNN transfer learning to enhance the performance of RNN prediction.

As detailed in subsection 6.3.2, during the transfer learning process, all parameters except those within the RNN cells are kept constant. Consequently, there are 1,404 parameters in the RNN and 2,808 in the BiRNN that require retraining using the measured strain data. This represents 16.46 % of the parameters in the RNN and 12.62 % in the BiRNN will undergo transfer learning. The learning rate is set at a very low 1×10^{-6} because the success of transfer learning in this scenario is highly sensitive to the learning rate. In this problem, it is found that stability in the optimization process can only be achieved when the learning rate is significantly lower than that used in the initial training phase of the RNN and BiRNN. The transfer learning for RNN and BiRNN converge after 1,420 epochs and 2,420 epochs of training, respectively. NRMSE from the TL-RNN and TL-BiRNN predicted deflection at 92 cm of the beam are calculated as well, and comparisons are made with those from the RNN and BiRNN in the previous subsection, as presented in Figure 6.27. The results indicate that through the use of two FBGs' information, the transfer learning-based RNN and BiRNN models significantly outperform the vanilla RNN and BiRNN. For the RNN, the NRMSE decreased from 0.545 to 0.273, a reduction of 49.9 %. In contrast, the BiRNN started with higher initial accuracy, with an NRMSE of 0.221. After the transfer learning, the NRMSE is reduced to 0.168, providing a decrease of 24.0 %. Although the accuracy of the TL-RNN does not surpass that of the FE model-based training of BiRNN, it likely constitutes the achievable upper boundary for models designed to conduct online state estimation using only three accelerometer inputs. Finally, having demonstrated that the TL-RNN method proposed in this chapter indeed enhances the performance of RNN models, the deflection results at 92 cm for both TL-RNN and TL-BiRNN are presented in Figure 6.28. Comparison with the measured values shows superior accuracy of TL-BiRNN in state estimation problems. At the same time, the results confirm that the performance of TL-RNN is acceptable and can be applied to online monitoring and digital twin systems, despite minor discrepancies between predicted and actual values at some time steps.

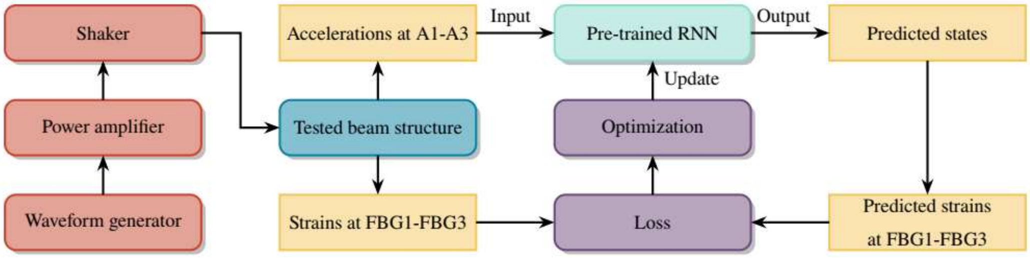


Figure 6.26. Detailed procedures for RNN transfer learning to enhance the performance of beam state estimation.

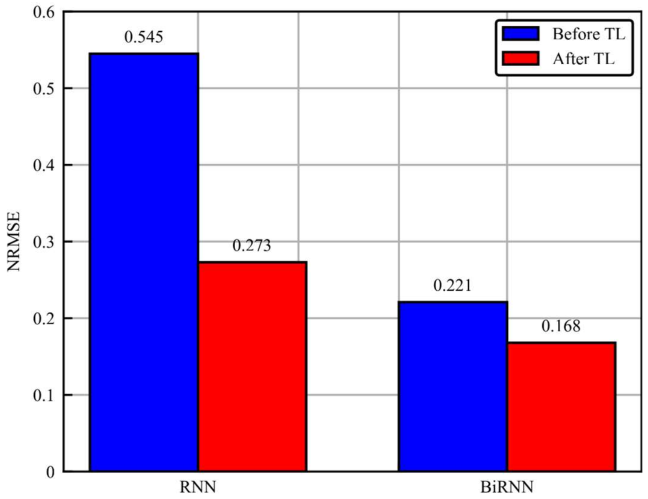


Figure 6.27. NRMSE values of RNN, BiRNN, TL-RNN, and TL-BiRNN models for the beam deflection prediction.

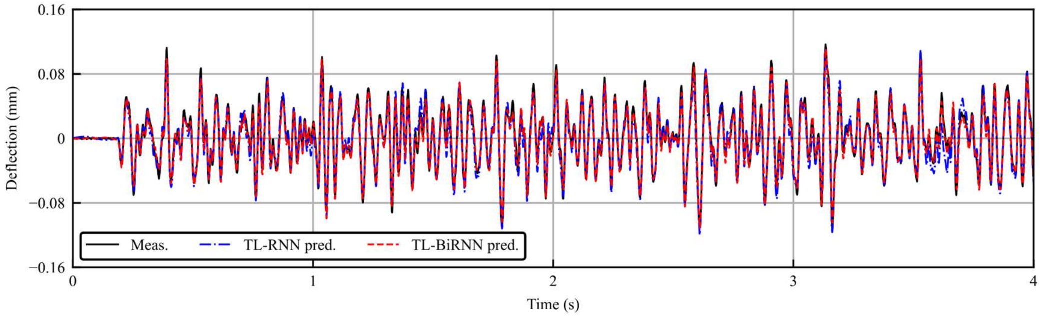


Figure 6.28. Comparison of measured and predicted beam deflections using TL-RNN and TL-BiRNN models.

6.6. Discussion

The research presented in this chapter demonstrates the efficacy of RNN transfer learning for state estimation in structural dynamics, providing a commendable alternative to conventional methods. The improved state estimation performance of TL-RNN can be intuitively understood by analyzing how it leverages information compared to traditional methods. Figure 6.29 illustrates three approaches for state estimation by using different types of information and their combinations, where Type *A* represents commonly used acceleration data, Type *B* encompasses additional measurement data from the structure, and Type C pertains to the estimated state information. Models $\mathcal{M}_1$ to $\mathcal{M}_3$, despite their differences, output identical results—Type C. Traditionally, the approach depicted in Figure 6.29(a) is most prevalent, with extensive research focusing on effectively constructing the $\mathcal{M}_1$ model to ensure accurate state prediction. Some methods originate directly from the equations of motion formulated by the FE model, while others derive from the data generated by the FE model.

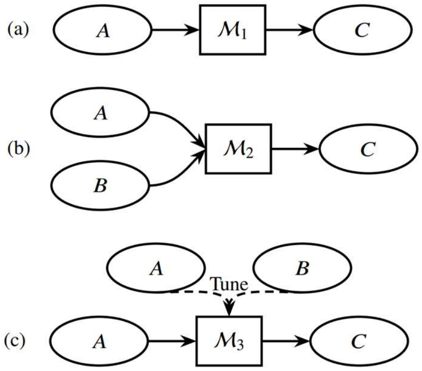


Figure 6.29. Three approaches on using information for state estimation in structural dynamics.

The interest of this study lies in optimizing the use of increased monitoring data available through sensor arrays in monitoring systems, a scenario that aligns with real-world applications where multiple sensors often collaborate in monitoring tasks. In the numerical example illustrated above, both approaches depicted in Figure 6.29(b) and (c) are investigated. Either a larger model is constructed with diverse inputs for predicting the structural state or the model is refined using both Types A and B data. The present work indicates that employing all available information as per the approach in Figure 6.29(c) substantially outperforms that in Figure 6.29 (b). This is primarily because models $\mathcal{M}_2$ and $\mathcal{M}_3$ initially require calibration using an FE model, which, even when calibrated, still differs from the actual structure. Such discrepancies can be exacerbated under specific conditions, leading to inaccuracies in the initial outputs of $\mathcal{M}_2$ or $\mathcal{M}_3$. This is evidenced in the numerical example where a multi-story linear model is calibrated using ambient vibration data from the base-isolated building. During an earthquake, the sliding mechanism at the base of the building causes its response to become predominantly nonlinear, rendering the initial prediction of $\mathcal{M}_2$ and $\mathcal{M}_3$ inaccurate. Despite additional data from Type B, the results displayed in Figure 6.15 still show low prediction accuracy. Contrarily, using the mode depicted in Figure 6.29(c), where the input remains simply Type A but is fine-tuned using both Types A and B data, results in accurate prediction of Type C. This suggests that incorporating additional information into model adjustments, rather than directly into model predictions, might yield more precise results. In the experimental validation, the approach shown in Figure 6.29(c) is replicated. The FE model of the continuous beam is calibrated using FRF data obtained from experimental modal analysis. The training and test datasets are generated based on the random vibration of the calibrated FE model; and both RNN and BiRNN models for state estimation are trained with those generated structural random vibration data, establishing the information flow from Types A to C as shown in both Figure 6.29(a) and (c). Subsequently, the strain at three locations is calculated based on the measured wavelength changes in FBG1 to FBG3, constituting Type B data, and RNN and BiRNN models are fine-tuned accordingly. This results in model $\mathcal{M}_3$, which, despite having the same input information as $\mathcal{M}_1$, demonstrates superior performance, thereby confirming the practical efficacy of $\mathcal{M}_3$ over $\mathcal{M}_1$ in future state prediction tasks.

While the advantages of the TL-RNN model are evident, it remains essential to evaluate its appropriateness for specific problems in comparison with traditional methods. Without introducing additional measurements, a comparison between traditional Kalman filter-based approaches and RNNs highlights the greater flexibility of RNNs. This flexibility offers both benefits and drawbacks. In particular, training RNN models involves many more parameters and requires careful tuning of numerous hyperparameters compared to the Kalman filter-based models commonly used for state estimation. Consequently, developing an effective state estimation model becomes considerably more time-consuming and labor-intensive. However, for practitioners with extensive experience in deep learning, a well-constructed RNN model can surpass the benchmarks set by the Kalman filter-based models. This superiority arises from the empirical nature of deep learning: its value lies in its efficacy. Moreover, a state estimation model based on the proposed TL-RNN framework may be better suited to complex structures where FE models, based on existing physical laws, fail to capture the dynamics accurately. In scenarios where model assumptions align closely with actual conditions, TL-RNN may not necessarily outperform existing methods. Conversely, in cases where there is a significant mismatch between model assumptions and real-world conditions, a common occurrence given the complexity of real-world structures, TL-RNN can be a highly effective choice, potentially leading to substantial improvements in prediction accuracy.

Looking ahead, open-source libraries like PyTorch, TensorFlow, and JAX are continuously evolving, providing robust platforms that ease the development, training, and application of neural networks. These tools abstract complex algorithms into user-friendly interfaces, making RNN implementation more accessible to researchers and engineers in structural engineering. Concurrently, hardware improvements, particularly GPU technology, are enhancing computational power for training deep neural networks. This progression not only accelerates the training process but also democratizes access to these advanced facilities. As a result, building and training RNN models, as well as fine-tuning them using the transfer learning approach proposed in this article, are becoming more convenient and effective for state estimation in structural dynamics. This integration of advanced computational tools and methodologies significantly contributes to the precision and efficiency of dynamic state estimation, aligning with the latest trends in data-centric engineering research.

6.7. Summary

This chapter proposes an RNN transfer learning approach for accurate estimation of structural state. There are two original innovations introduced. First, when multiple types of monitoring information are available, a new paradigm for information integration and utilization in constructing state estimation model is illustrated, which truly achieves the physics-data fusion. Second, contrary to the traditional approach of freezing the latter several layers of the model, the fine-tuning of the RNN model in state estimation requires locking the parameters in earlier RNN cells. The innovations presented, along with acknowledgment that discrepancies between FE models and actual structure are inevitable, suggest that integrating the constructed physics-based model (low-fidelity) with measured data (high-fidelity) is the optimal approach for solution. The concept is exemplified by the TL-RNN and TL-BiRNN models developed herein. The effectiveness of the innovations in improving prediction accuracy is demonstrated through two state estimation examples: a numerical example using a linearized shear-type structure as an approximation of a base-isolated building, and an experimental example employing a simplified beam model to represent the laboratory structure. Despite the inherent discrepancies between the FE model and the actual structure, the state estimation results from both numerical and experimental case studies, based on the TL-RNN framework, prove the efficacy of the proposed innovations in enhancing prediction accuracy. Additionally, as a complement to the time-delayed MFGP approach discussed in Chapter 5, the TL-RNN within the deep learning framework effectively captures correlations in multi-output data. This capability underscores the flexibility and robustness of the transfer learning RNN method for enhanced state estimation.

Chapter 7.

Conclusions and future work

7.1. General summary

This thesis introduces a series of advanced machine learning methods aimed at addressing two critical inverse problems in the field of structural dynamics: 1) force reconstruction and 2) state estimation. By reviewing existing approaches to these challenges and considering the specific requirements of load reconstruction and state identification in the industrial sector, this study delves into several key issues within these inverse problems. These include the reliability assessment of conclusions derived from measurement data and the inaccuracies in predictions caused by discrepancies between FE models and actual structures, etc. With the continuous advancement of machine learning technologies, this work leverages nonparametric Bayesian methods and deep learning techniques, integrating physics-driven FE models with multi-source real-world measurement data to propose comprehensive solutions. The proposed methods not only enhance the precision of force reconstruction and state estimation but also improve the model's generalization capabilities and the reliability of its practical applications. Through extensive experiments and case studies, the proposed methods demonstrate superior performance across various application scenarios, providing robust support for the advancement of SHM technologies. The primary contributions of the thesis are summarized as follows:

7.2. Suggestions for future research

7.2.1 Bayesian-based force reconstruction with uncertain system parameters

This study advances time-domain force reconstruction by integrating GP priors for forces, thereby establishing a probabilistic inverse model that links force and structural response. Structural systems often exhibit inherent uncertainties, characterized by multiple uncertain parameters. To achieve more realistic force characterizations, this approach comprehensively accounts for multi-source uncertainties originating from both the structural dynamic system and the measurement processes. By doing so, the reconstructed forces better reflect real-world conditions. Furthermore, the uncertainty associated with unknown system parameters can be incorporated into the Bayesian framework. However, deriving analytical expressions for the posterior distributions of forces in the presence of these uncertainties was not attainable within the scope of this thesis. Consequently, the posterior force distributions remain intractable analytically. Future research will focus on developing efficient algorithms to accurately derive force posteriors while accounting for uncertain system parameters. The advancements will enhance the robustness and applicability of Bayesian-based force reconstruction methods in complex structural systems.

7.2.2 Mathematical rationale behind the success of time-delayed GP for state estimation

The time-delayed modeling approach for state estimation is fundamentally straightforward, involving the preparation of labeled pairs for training and test datasets, which enables the model to function effectively as a state estimator. However, as the number of delay steps increases, the problem transitions into a high-dimensional learning challenge. Traditionally, GPR is more commonly applied in lower-dimensional spaces due to the curse of dimensionality, which generally hampers its performance in high-dimensional settings. Despite this conventional limitation, this study demonstrates that GPR can successfully capture the necessary mappings based on the data, even in high-dimensional scenarios. This success is particularly noteworthy given that the kernels employed in the proposed examples are simple squared exponential kernels supplemented by a noise hyperparameter, resulting in a GPR model with only three parameters. Remarkably, this streamlined GPR model is capable of effectively learning complex high-dimensional mappings without succumbing to the typical performance degradation associated with increased dimensionality. This phenomenon suggests that, particularly for structural dynamic systems, especially those exhibiting linear behavior, there may be underlying mathematical principles that facilitate the success of GPR within the context of time-delayed state estimation. The linearity of the system could create inherent structures or symmetries in the data that GPR can exploit, allowing it to maintain high performance despite the high dimensionality introduced by multiple delay steps. Future research should explore the geometry of the time-delayed series to further investigate and elucidate the factors contributing to the effectiveness of GPR in this setting. Understanding these geometric properties could provide deeper insights into why GPR performs well despite the high-dimensional nature of the problem and could lead to the development of more refined models or kernel functions tailored to exploit these underlying mathematical rationales.

7.2.3 Transfer learning of more sophisticated RNN models for state estimation

This thesis employed a fundamental RNN model to demonstrate its capability in performing state estimation and its transferability to measured data. With the rapid advancements in deep learning techniques, a diverse array of RNN models has been developed, offering enhanced sophistication in time series modeling. Advanced RNN architectures, such as multiple LSTM networks and gated recurrent units, are particularly adept at capturing long-term dependencies and handling complex temporal dynamics. In scenarios where structural dynamics are highly complex, these advanced RNN models can provide more accurate and reliable state estimations compared to their fundamental counterparts. Their ability to model intricate patterns and dependencies within the data enables them to better understand and predict the state of complex structural systems. However, the transfer learning of these advanced RNN models presents challenges not encountered with fundamental RNNs. Specifically, determining which parameters in the RNN model should be frozen during the transfer process requires empirical validation and practical experimentation. To address these challenges, future work will focus on developing and validating transfer learning strategies tailored for sophisticated RNN models.

References

R. Abdulkadirov, P. Lyakhov, N. Nagornov, Survey of optimization algorithms in modern neural networks. *Mathematics*, 11 (2023) 2466.

M. Abdulkarem, K. Samsudin, F.Z. Rokhani, M.F. A Rasid, Wireless sensor network for structural health monitoring: A contemporary review of technologies, challenges, and future direction. *Structural Health Monitoring*, 19 (2019) 693–735.

D.J. Atha, M.R. Jahanshahi, Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. *Structural Health Monitoring*, 17 (2018) 1110–1128.

M. Aucejo, O. De Smet, A multiplicative regularization for force reconstruction. *Mechanical Systems and Signal Processing*, 85 (2017) 730–745.

M. Aucejo, O. De Smet, On a general iteratively reweighted algorithm for solving force reconstruction problems. *Journal of Sound and Vibration*, 458 (2019) 376–388.

O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, D.J. Inman, A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. *Mechanical Systems and Signal Processing*, 147 (2021) 107077.

S.A. Aye, P.S. Heyns, An integrated Gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission. *Mechanical Systems and Signal Processing*, 84 (2017) 485–498.

Y. Bao, S. Mahadevan, Diagnosis of interior damage with a convolutional neural network using simulation and measurement data. *Structural Health Monitoring*, 21 (2022) 2312–2328.

L. Berke, S.N. Patnaik, P.L.N. Murthy, Optimum design of aerospace structural components using neural networks. *Computers & Structures*, 48 (1993) 1001–1010.

S. Bhowmick, S. Nagarajaiah, Z. Lai, Measurement of full-field displacement time history of a vibrating continuous edge from video. *Mechanical Systems and Signal Processing*, 144 (2020) 106847.

C.M. Bishop, Pattern recognition and machine learning, Springer, New York, 2006.

L. Brevault, M. Balesdent, A. Hebbal, Overview of Gaussian process based multi-fidelity techniques with variable relationship between fidelities, application to aerospace systems. *Aerospace Science and Technology*, 107 (2020) 106339.

R. Brincker, L. Zhang, P. Andersen, Modal identification of output-only systems using frequency domain decomposition. *Smart Materials and Structures*, 10 (2001) 441.

D. Castelvecchi, Can we open the black box of AI? *Nature News*, (2016).

Y.J. Cha, R. Ali, J. Lewis, O. Büyüköztürk, Deep learning-based structural health monitoring. *Automation in Construction*, 161 (2024) 105328.

S. Chakraborty, S. Adhikari, Machine learning based digital twin for dynamical systems with multiple time-scales. *Computers & Structures*, 243 (2021) 106410.

C.Y. Chang, C.W. Huang, Non-contact measurement of inter-story drift in three-layer RC structure under seismic vibration using digital image correlation. *Mechanical Systems and Signal Processing*, 136 (2020) 106500.

Z. Chen, T. Liu, M. Li, M. Yu, Z. Lu, D. Liu, Dynamic response of railway vehicles under unsteady aerodynamic forces caused by local landforms. *Wind and Structures*, 29 (2019) 149–161.

Y. Chen, B. Hosseini, H. Owhadi, A.M. Stuart, Solving and learning nonlinear PDEs with Gaussian processes. *Journal of Computational Physics*, 447 (2021a) 110668.

S.X. Chen, L. Zhou, Y.Q. Ni, X.Z. Liu, An acoustic-homologous transfer learning approach for acoustic emission–based rail condition evaluation. *Structural Health Monitoring*, 20 (2021b) 2161–2181.

H.G. Choi, A.N. Thite, D.J. Thompson, Comparison of methods for parameter selection in Tikhonov regularization with application to inverse force determination. *Journal of Sound and Vibration*, 304 (2007) 894–917.

A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A.A. Bharath, Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35 (2018) 53–65.

B.Z. Cunha, C. Droz, A.M. Zine, S. Foulard, M. Ichchou, A review of machine learning methods applied to structural dynamics and vibroacoustic. *Mechanical Systems and Signal Processing*, 200 (2023) 110535.

K. Cutajar, M. Pullin, A. Damianou, N. Lawrence, J. González, Deep Gaussian processes for multi-fidelity modeling. *arXiv preprint arXiv:1903.07320*, (2019).

K. Dai, Y. Wang, A. Hedric, Z. Huang, Dynamic behaviors of historical wrought iron truss bridges: a field testing case study. In *SPIE Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring*, Las Vegas, 2016.

M.E. De Simone, F. Ciampa, M. Meo, A hierarchical method for the impact force reconstruction in composite structures. *Smart Materials and Structures*, 28 (2019) 085022.

M. Dixit, A. Tiwari, H. Pathak, R. Astya, An overview of deep learning architectures, libraries and its applications areas. In *International Conference on Advances in Computing, Communication Control and Networking*, Greater Noida, 2018.

W. Dou, L. Zhang, S. Stichel, Analysis of aerodynamic load on high speed train carbody utilizing fluid-structure interaction method. In *1st International Railway Symposium*, Aachen, 2017.

S. Eftekhar Azam, E. Chatzi, C. Papadimitriou, A dual Kalman filter approach for state estimation via output-only acceleration measurements. *Mechanical Systems and Signal Processing*, 60–61 (2015) 866–886.

H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Springer Science & Business Media, Berlin, 1996.

K. Erazo, E.M. Hernandez, Uncertainty quantification of state estimation in nonlinear structural systems with application to seismic response in buildings. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 2 (2015) B5015001.

W. Feng, Q. Li, Q. Lu, C. Li, B. Wang, Element-wise Bayesian regularization for fast and adaptive force reconstruction. *Journal of Sound and Vibration*, 490 (2021) 115713.

A. Firus, R. Kemmler, H. Berthold, S. Lorenzen, J. Schneider, A time domain method for reconstruction of pedestrian induced loads on vibrating structures. *Mechanical Systems and Signal Processing*, 171 (2022) 108887.

T.E. Fricker, J.E. Oakley, N.D. Sims, K. Worden, Probabilistic uncertainty analysis of an FRF of a structure using a Gaussian process emulator. *Mechanical Systems and Signal Processing*, 25 (2011) 2962–2975.

R. Galaz-Palma, B. Targui, O. Hernández-González, G. Valencia-Palomo, M.E. Guerrero-Sánchez, Robust observer for input and state estimation in building structure systems. *Journal of Vibration and Control*, 29 (2022) 4422–4438.

A. Gallet, S. Rigby, T.N. Tallman, X. Kong, I. Hajirasouliha, A. Liew, D. Liu, L. Chen, A. Hauptmann, D. Smyl. Structural engineering from an inverse problems perspective. *Proceedings of the Royal Society A*, 478 (2022) 20210526.

J. Gardner, G. Pleiss, K.Q. Weinberger, D. Bindel, A.G. Wilson, GPyTorch: Blackbox matrix-matrix Gaussian process inference with GPU acceleration. In *Advances in Neural Information Processing Systems*, Montréal, 2018.

Z. Ghahramani, Probabilistic machine learning and artificial intelligence. *Nature*, 521 (2015) 452-459.

X Glorot, Y Bengio, Understanding the difficulty of training deep feedforward neural networks. In *13th International Conference on Artificial Intelligence and Statistics*, Sardinia, 2010.

I. Goodfellow, Y. Bengio, A. Courville, Deep learning, MIT Press, Cambridge, 2016.

GPy, GPy: A Gaussian process framework in python, http://github.com/SheffieldML/GPy. Since 2012.

M. Gul, F.N. Catbas, Statistical pattern recognition for structural health monitoring using time series modeling: Theory and experimental verifications. *Mechanical Systems and Signal Processing*, 23 (2009) 2192–2204.

N.S. Gulgec, M. Takáč, S.N. Pakzad, Structural sensing with deep learning: Strain estimation from acceleration data for fatigue assessment. *Computer-Aided Civil and Infrastructure Engineering*, 35 (2020) 1349–1364.

S. Guo, C.F. Silva, W. Polifke, Robust identification of flame frequency response via multi-fidelity Gaussian process approach. *Journal of Sound and Vibration*, 502 (2021) 116083.

S. Hassani, U. Dackermann, A systematic review of advanced sensor technologies for non-destructive testing and structural health monitoring. *Sensors*, 23 (2023) 2204.

S. Hassani, U. Dackermann, M. Mousavi, J. Li, A systematic review of data fusion techniques for optimized structural health monitoring. *Information Fusion*, 103 (2024) 102136.

K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators. *Neural Networks*, 2 (1989) 359–366.

R. Hou, Y. Xia, Review on the new development of vibration-based damage identification for civil engineering structures: 2010–2019. *Journal of Sound and Vibration*, 491 (2021) 115741.

M.A. Iglesias, K.J.H. Law, A.M. Stuart, Ensemble Kalman methods for inverse problems. *Inverse Problems*, 29 (2013) 045001.

M. Impraimakis, A. W. Smyth, An unscented Kalman filter method for real time input-parameter-state estimation. *Mechanical Systems and Signal Processing*, 162 (2022) 108026.

M. Imregun, W. Visser, D. Ewins, Finite element model updating using frequency response function data: I. Theory and initial investigation. *Mechanical Systems and Signal Processing*, 9 (1995) 187–202.

C. Introini, S. Cavalleri, S. Lorenzi, S. Riva, A. Cammi, Stabilization of Generalized Empirical Interpolation Method (GEIM) in presence of noise: A novel approach based on Tikhonov regularization. *Computer Methods in Applied Mechanics and Engineering*, 404 (2023) 115773.

E. Jacquelin, A. Bennani, P. Hamelin, Force reconstruction: Analysis and regularization of a deconvolution problem. *Journal of Sound and Vibration*, 265 (2003) 81–107.

D. Jiang, C.Y. Hsieh, Z. Wu, Y. Kang, J. Wang, E. Wang, B. Liao, C. Shen, L. Xu, J. Wu, D. Cao, T. Hou, InteractionGraphNet: A novel and efficient deep graph representation learning framework for accurate protein–ligand interaction predictions. *Journal of Medicinal Chemistry*, 64 (2021) 18209−18232.

S.S. Jin, S.T. Kim, Y.H. Park, Combining point and distributed strain sensor for complementary data-fusion: A multi-fidelity approach. *Mechanical Systems and Signal Processing*, 157 (2021) 107725.

P. Jönsson, Ö. Johansson, Prediction of vehicle discomfort from transient vibrations. *Journal of Sound and Vibration*, 282 (2005) 1043–1064.

S.I. Kabanikhin, Definitions and examples of inverse and ill-posed problems. *Journal of Inverse and Ill-Posed Problems*, 16 (2008) 317–357.

A. Kamariotis, E. Chatzi, D. Straub, Value of information from vibration-based structural health monitoring extracted via Bayesian model updating. *Mechanical Systems and Signal Processing*, 166 (2022) 108465.

A. Kamariotis, E. Chatzi, D. Straub, A framework for quantifying the value of vibration-based structural health monitoring. *Mechanical Systems and Signal Processing*, 184 (2023) 109708.

C. Karatas, B. Degerliyurt, Y. Yaman, M. Sahin, Fibre Bragg grating sensor applications for structural health monitoring. *Aircraft Engineering and Aerospace Technology*, 92 (2018) 355–367.

T. Karvonen, C.J. Oates, Maximum likelihood estimation in Gaussian process regression is ill-posed. *Journal of Machine Learning Research*, 24 (2023) 1–47.

A.M. Kosikova, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Bayesian structural identification using Gaussian Process discrepancy models. *Computer Methods in Applied Mechanics and Engineering*, 417 (2023) 116357.

Y. Kumar, P. Bahl, S. Chakraborty, State estimation with limited sensors—a deep learning based approach. *Journal of Computational Physics*, 457 (2022) 111081.

H.F. Lam, J. Yang, S.K. Au, Bayesian model updating of a coupled-slab system using field test data utilizing an enhanced Markov chain Monte Carlo Simulation algorithm. *Engineering Structures*, 102 (2015) 144–155.

S.S. Law, J. Li, Y. Ding, Structural response reconstruction with transmissibility concept in frequency domain. *Mechanical Systems and Signal Processing*, 25 (2011) 952–968.

L. Le Gratiet, J. Garnier, Recursive co-kriging model for design of computer experiments with multiple levels of fidelity. *International Journal for Uncertainty Quantification*, 4 (2014) 365–386.

D. Lesnic, Inverse problems with applications in science and engineering, Chapman & Hall/CRC, Boca Raton, 2022.

J. Li, H. Hao, Health monitoring of joint conditions in steel truss bridges with relative displacement sensors. *Measurement*, 88 (2016) 360–371.

Q. Li, Q. Lu, Time Domain Force identification based on adaptive ℓq regularization. *Journal of Vibration and Control*, 24 (2018) 5610–5626.

P. Li, X. Chen, S. Shen, Stereo R-CNN based 3D object detection for autonomous driving. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 2019.

M. Li, G. Jia, Multifidelity Gaussian process model integrating low-and high-fidelity data considering censoring. *Journal of Structural Engineering*, 146 (2020a) 04019215.

W. Li, T. Liu, Z. Chen, Z. Guo, X. Huo, Comparative study on the unsteady slipstream induced by a single train and two trains passing each other in a tunnel. *Journal of Wind Engineering and Industrial Aerodynamics*, 198 (2020) 104095.

Y. Li, H. Huang, W. Zhang, L. Sun, Structural full-field responses reconstruction by the SVD and pseudo-inverse operator-estimated force with two-degree multi-scale models. *Engineering Structures*, 249 (2021a) 112986.

Y Li, T Bao, Z Chen, Z Gao, X Shu, K Zhang, A missing sensor measurement data reconstruction framework powered by multi-task Gaussian process regression for dam structural health monitoring systems. *Measurement*, 186 (2021b) 110085.

Y. Li, L. Sun, Y. Xia, L. Luo, A. Wang, X. Jian, General Tikhonov regularization-based load estimation of bridges considering the computer vision-extracted prior information. *Structural Control and Health Monitoring*, 29 (2022) e3135.

H.W. Li, Y.Q. Ni, Y.W. Wang, Z.W. Chen, E.Z. Rui, Z.D. Xu, Modeling of forced-vibration systems using continuous-time state-space neural network. *Engineering Structures*, 302 (2024a) 117329.

H.W. Li, S. Hao, Y.Q. Ni, Y.W. Wang, Z.D. Xu, Hybrid structural analysis integrating physical model and continuous-time state-space neural network model. *Computer-Aided Civil and Infrastructure Engineering*, published online, early view, (2024b).

Y. Liu, W.S. Shepard, Dynamic force identification based on enhanced least squares and total least-squares schemes in the frequency domain. *Journal of Sound and Vibration*, 282 (2005) 37–60.

G.R. Liu, S.S. Quek, The finite element method: A practical course, Butterworth-Heinemann, Oxford, 2013

J. Liu, X. Sun, X. Han, C. Jiang, D. Yu, A novel computational inverse technique for load identification using the shape function method of moving least square fitting. *Computers & Structures*, 144 (2014) 127–137.

J. Liu, B. Li, A novel strategy for response and force reconstruction under impact excitation. *Journal of Mechanical Science and Technology*, 32 (2018) 3581–3596.

R. Liu, E. Dobriban, Z. Hou, K. Qian, Dynamic load identification for mechanical systems: A review. *Archives of Computational Methods in Engineering*, 29 (2022) 831-863.

Y. Liu, L. Wang, B.F. Ng, Load-independent multi-objective sensor placement method for localization and reconstruction of external excitations under interval uncertainties. *Computer Methods in Applied Mechanics and Engineering*, 416 (2023) 116344.

E. Lourens, E. Reynders, G. De Roeck, G. Degrande, G. Lombaert, An augmented Kalman filter for force identification in structural dynamics. *Mechanical Systems and Signal Processing*, 27 (2012) 446–460.

X. Lv, X. Y. Huang, X. G. H., X. H. Su, A matrix-free implicit unstructured multigrid finite volume method for simulating structural dynamics and fluid-structure interaction. *Journal of Computational Physics*, 225 (2007) 120–144.

V.A. Matsagar, R. Jangid, Seismic response of base-isolated structures during impact with adjacent structures. *Engineering Structures*, 25 (2003) 1311–1323.

V.A. Matsagar, R. Jangid, Influence of isolator characteristics on the response of base-isolated structures. *Engineering Structures*, 26 (2004) 1735–1749.

A.G.G. Matthews, M. van der Wilk, T. Nickson, K. Fujii, A. Boukouvalas, P. León-Villagrá, Z. Ghahramani, J. Hensman, GPflow: A Gaussian process library using TensorFlow. *Journal of Machine Learning Research*, 18 (2017) 1–6.

R. Mayes, L. Ankers, P. Daborn, Predicting system response at unmeasured locations. *Experimental Techniques*, 44 (2020) 457–474.

M. Mehdipour Ghazi, H. Kemal Ekenel, A comprehensive analysis of deep learning based representation for face recognition. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Caesars Palace, 2016.

B. Moaveni, J.P. Conte, F.M. Hemez, Uncertainty and sensitivity analysis of damage identification results obtained using finite element model updating. *Computer-Aided Civil and Infrastructure Engineering*, 24 (2009) 320–334.

S. Moradi, S. Eftekhar Azam, M. Mofid, On Bayesian active vibration control of structures subjected to moving inertial loads. *Engineering Structures*, 239 (2021) 112313.

K.P. Murphy, Machine learning: a probabilistic perspective, MIT Press, Cambridge, 2012.

F. Naets, J. Croes, W. Desmet, An online coupled state/input/parameter estimation approach for Structural Dynamics. *Computer Methods in Applied Mechanics and Engineering*, 283 (2015) 1167–1188.

Y. Narazaki, F. Gomez, V. Hoskere, M.D. Smith, B.F. Spencer, Efficient development of vision-based dense three-dimensional displacement measurement algorithms using physics-based graphics models. *Structural Health Monitoring*, 20 (2020) 1841–1863.

R. Nayek, S. Narasimhan, Extraction of contact-point response in indirect bridge health monitoring using an input estimation approach. *Journal of Civil Structural Health Monitoring*, 10 (2020) 815–831.

Y.Q. Ni, X.T. Zhou, J.M. Ko, Experimental investigation of seismic damage identification using PCA-compressed frequency response functions and neural networks. *Journal of Sound and Vibration*, 290 (2006) 242–263.

Y.Q. Ni, Y. Xia, W.Y. Liao, J.M. Ko, Technology innovation in developing the structural health monitoring system for Guangzhou New TV Tower. *Structural Control and Health Monitoring*, 16 (2009) 73–98.

Y. Niu, C. Fritzen, H. Jung, I. Buethe, Y. Ni, Y. Wang, Online simultaneous reconstruction of wind load and structural responses—theory and application to Canton Tower. *Computer-Aided Civil and Infrastructure Engineering*, 30 (2015) 666–681.

A.C. Ogren, B.T. Feng, K.L. Bouman, C. Daraio, Gaussian process regression as a surrogate model for the computation of dispersion relations. *Computer Methods in Applied Mechanics and Engineering*, 420 (2024) 116661.

G. Ou, A.I. Ozdagli, S.J. Dyke, B. Wu, Robust integrated actuator control: experimental verification and real - time hybrid - simulation implementation. *Earthquake Engineering & Structural Dynamics*, 44 (2015) 441-460.

L. Parussini, D. Venturi, P. Perdikaris, G.E. Karniadakis, Multi-fidelity Gaussian process regression for prediction of random fields. *Journal of Computational Physics*, 336 (2017) 36–50.

P.K. Paul, A. Dutta, S.K. Deb, Comparison of the performance of nonlinear Kalman filter based algorithms for state-parameter identification of base isolated structures. *Structural Control and Health Monitoring*, 29 (2022) e3029.

S.J. Pan, Q. Yang, A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22 (2009) 1345–1359.

J. Panda, S. Chakraborty, S. Ray - Chaudhuri, A novel servomechanism based proportional–integral controller with Kalman filter estimator for seismic response control of structures using magneto-rheological dampers. *Structural Control and Health Monitoring*, 28 (2021) e2807.

C. Papadimitriou, C.-P. Fritzen, P. Kraemer, E. Ntotsios, Fatigue predictions in entire body of metallic structures from a limited number of vibration sensors using Kalman filtering. *Structural Control and Health Monitoring*, 18 (2011) 554–573.

D.P. Pasca, A. Aloisio, M.M. Rosso, S. Sotiropoulos, PyOMA and PyOMA GUI: A Python module and software for operational modal analysis, *SoftwareX*. 20 (2022) 101216.

D. Paul, K. Roy, Application of bridge weigh-in-motion system in bridge health monitoring: A state-of-the-art review. *Structural Health Monitoring*, 22 (2023) 4194–4232.

P. Perdikaris, M. Raissi, A. Damianou, N.D. Lawrence, G.E. Karniadakis, Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling. *Proceedings of the Royal Society A*, 473 (2017) 20160751.

J.V. Pham, O. Ghattas, N.T. Clemens, K.E. Willcox, Real-time aerodynamic load estimation for hypersonics via strain-based inverse maps, In *AIAA SCITECH Forum*, Orlando, 2024.

B. Qiao, X. Zhang, C. Wang, H. Zhang, X. Chen, Sparse regularization for force identification using dictionaries. *Journal of Sound and Vibration*, 368 (2016) 71–86.

G. Qiao, S. Rahmatalla, Moving load identification on Euler-Bernoulli beams with viscoelastic boundary conditions by Tikhonov regularization. *Inverse Problems in Science and Engineering*, 29 (2020) 1070–1107.

C. Qin, G. Shi, J. Tao, H. Yu, Y. Jin, J. Lei, C. Liu, Precise cutterhead torque prediction for shield tunneling machines using a novel hybrid deep neural network. *Mechanical Systems and Signal Processing*, 151 (2021) 107386.

C.E. Rasmussen, C.K.I. Williams, Gaussian processes for machine learning, MIT Press, Cambridge, 2006.

C.E. Rasmussen, H. Nickisch, Gaussian processes for machine learning (GPML) toolbox. *Journal of Machine Learning Research*, 11 (2010) 3011–3015.

M. Raissi, P. Perdikaris, G.E. Karniadakis, Machine learning of linear differential equations using Gaussian processes. *Journal of Computational Physics*, 348 (2017) 683–693.

M. Raissi, P. Perdikaris, G.E. Karniadakis, Numerical Gaussian processes for time-dependent and nonlinear partial differential equations. *SIAM Journal on Scientific Computing*, 40 (2018) A172–A198.

M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378 (2019) 686-707.

A. Rezayat, V. Nassiri, B. De Pauw, J. Ertveldt, S. Vanlanduit, P. Guillaume, Identification of dynamic forces using group-sparsity in frequency domain. *Mechanical Systems and Signal Processing*, 70–71 (2016) 756–768.

D. Rocchi, G. Tomasini, P. Schito, C. Somaschini, Wind effects induced by high speed train pass-by in open air. *Journal of Wind Engineering and Industrial Aerodynamics*, 173 (2018) 279–288.

H. Ronasi, H. Johansson, F. Larsson, A numerical framework for load identification and regularization with application to rolling disc problem. *Computers & Structures*, 89 (2011) 38–47.

H. Ronasi, J.C.O. Nielsen, Inverse identification of wheel–rail contact forces based on observation of wheel disc strains: An evaluation of three numerical algorithms. *Vehicle System Dynamics*, 51 (2013) 74–90.

M. Roohi, K. Erazo, D. Rosowsky, E.M. Hernandez, An extended model-based observer for state estimation in nonlinear hysteretic structural systems. *Mechanical Systems and Signal Processing*, 146 (2021) 107015.

J. Sanchez, H. Benaroya, Review of force reconstruction techniques. *Journal of Sound and Vibration*, 333 (2014) 2999–3018.

G. Solari, G. Piccardo, Probabilistic 3-D turbulence modeling for gust buffeting of structures. *Probabilistic Engineering Mechanics*, 16 (2001) 73-86.

Y. Shao, L. Li, J. Li, S. An, H. Hao, Computer vision based target-free 3D vibration displacement measurement of structures. *Engineering Structures*, 246 (2021) 113040.

E. Simiu, D.H. Yeo, Wind Effects on Structures: Modern Structural Design for Wind. Wiley-Blackwell, Hoboken, 2021.

M. Song, S. Christensen, B. Moaveni, A. Brandt, E. Hines, Joint parameter-input estimation for virtual sensing on an offshore platform using output-only measurements. *Mechanical Systems and Signal Processing*, 170 (2022) 108814.

N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15 (2014) 1929−1958.

A.M. Stuart, Inverse problems: a Bayesian perspective. *Acta Numerica*, 19 (2010) 451–559.

S.A. Taher, J. Li, H. Fang, Earthquake input and state estimation for buildings using absolute floor accelerations. *Earthquake Engineering and Structural Dynamics*, 50 (2021) 1020–1042.

D. Tcherniak, L.L. Mølgaard, Active vibration-based structural health monitoring system for wind turbine blade: Demonstration on an operating Vestas V27 wind turbine. *Structural Health Monitoring*, 16 (2017) 536–550.

H. Teimouri, A.S. Milani, J. Loeppky, R. Seethaler, A Gaussian process-based approach to cope with uncertainty in structural health monitoring. *Structural Health Monitoring*, 16 (2016) 174–184.

D. Teymouri, O. Sedehi, L. S. Katafygiotis, C. Papadimitriou, A Bayesian Expectation-Maximization (BEM) methodology for joint input-state estimation and virtual sensing of structures. *Mechanical Systems and Signal Processing*, 169 (2022) 108602.

H.T. Thai, Machine learning for structural engineering: A state-of-the-art review. *Structures*, 38 (2022) 448–491.

A. Tognan, L. Laurenti, E. Salvati, Contour method with uncertainty quantification: A robust and optimised framework via Gaussian process regression. *Experimental Mechanics*, 62 (2022) 1305–1317.

G.R. Tondo, S. Rau, I. Kavrakov, G. Morgenthal, Stochastic stiffness identification and response estimation of Timoshenko beams via physics-informed Gaussian processes. *Probabilistic Engineering Mechanics*, 74 (2023) 103534.

B. Torres, I. Payá-Zaforteza, P.A. Calderón, J.M. Adam, Analysis of the strain transfer in a new FBG sensor for Structural Health Monitoring. *Engineering Structures*, 33 (2011) 539–548.

M. Torzoni, M. Tezzele, S. Mariani, A. Manzoni, K.E. Willcox, A digital twin framework for civil engineering structures. *Computer Methods in Applied Mechanics and Engineering*, 418 (2024) 116584.

M. Tschannen, O. Bachem, M. Lucic, Recent advances in autoencoder-based representation learning. *arXiv preprint arXiv:1812.05069*, (2018).

E. Turco, Tools for the numerical solution of inverse problems in structural mechanics: review and research perspectives. *European Journal of Environmental and Civil Engineering*, 21 (2016) 509–554.

S. Särkkä, Linear operators and stochastic partial differential equations in Gaussian process regression. In *Artificial Neural Networks and Machine Learning*, Espoo, 2011.

M. Song, S. Christensen, B. Moaveni, A. Brandt, E. Hines, Joint parameter-input estimation for virtual sensing on an offshore platform using output-only measurements. *Mechanical Systems and Signal Processing*, 170 (2022) 108814.

T. Uhl, The inverse identification problem and its technical application. *Archive of Applied Mechanics*, 77 (2006) 325–337.

G. Van Houdt, C. Mosquera, G. Nápoles, A review on the long short-term memory model. *Artificial Intelligence Review*, 53 (2020) 5929–5955.

H.P. Wan, Z. Mao, M.D. Todd, W.X. Ren, Analytical uncertainty quantification for modal frequencies with structural parameter uncertainty using a Gaussian process metamodel. *Engineering Structures*, 75 (2014) 577–589.

H.P. Wan, W.X. Ren, A residual-based Gaussian process model framework for finite element model updating. *Computers & Structures*, 156 (2015) 149–159.

H.P. Wan, Y.Q. Ni, Bayesian modeling approach for forecast of structural stress response using structural health monitoring data. *Journal of Structural Engineering*, 144 (2018) 04018130.

H.P. Wan, Y.Q. Ni, Bayesian multi-task learning methodology for reconstruction of structural health monitoring data. *Structural Health Monitoring*, 18 (2019) 1282-1309.

Y.E. Wang, G.Y. Wei, D. Brooks, Benchmarking TPU, GPU, and CPU platforms for deep learning. *arXiv preprint arXiv:1907.10701*, (2019).

L. Wang, Y. Liu, K. Gu, T. Wu, A radial basis function artificial neural network (RBF ANN) based method for uncertain distributed force reconstruction considering signal noises and material dispersion. *Computer Methods in Applied Mechanics and Engineering*, 364 (2020a) 112954.

S.M. Wang, J.D. Yau, Y.F. Duan, Y.Q. Ni, H.P. Wan, S.K. Wu, E.C. Ting, Prediction of crosswind-induced derailment of train–rail–bridge system by vector mechanics. *Journal of Engineering Mechanics*, 146 (2020b) 04020132.

L. Wang, J. Guo, I. Takewaki, Real-time hysteresis identification in structures based on restoring force reconstruction and Kalman filter. *Mechanical Systems and Signal Processing*, 150 (2021) 107297.

Q.A. Wang, C. Zhang, Z.G. Ma, Y.Q. Ni, Modelling and forecasting of SHM strain measurement for a large-scale suspension bridge during typhoon events using variational heteroscedastic Gaussian process. *Engineering Structures*, 251 (2022) 113554.

L. Wang, H. Xu, Y. Liu, A novel dynamic load identification approach for multi-source uncertain structures based on the set-theoretical wavelet transform and layered noise reduction. *Structures*, 51 (2023) 91–104.

K. Weiss, T.M. Khoshgoftaar, D.D. Wang, A survey of transfer learning. *Journal of Big data*, 3 (2016) 9.

S.Q. Wu, S.S. Law, Moving force identification based on stochastic finite element model. *Engineering Structures*, 32 (2010) 1016–1027.

R.T. Wu, M.R. Jahanshahi, Data fusion approaches for structural health monitoring and system identification: Past, present, and future. *Structural Health Monitoring*, 19 (2018) 552–586.

R.T. Wu, M.R. Jahanshahi, Deep convolutional neural network for structural dynamic response estimation and system identification. *Journal of Engineering Mechanics*, 145 (2019) 04018125.

Y. Xia, Y.Q. Ni, J.M. Ko, H.B. Chen, Development of a structural health monitoring benchmark problem for high-rise slender structures. *Advances in Science and Technology*, 56 (2008) 489–494.

J. Xian, Z. Wang, A physics and data co-driven surrogate modeling method for high-dimensional rare event simulation. *Journal of Computational Physics*, 510 (2024) 113069.

S.L. Xie, Y.H. Zhang, Q. Xie, C.H. Chen, X.N. Zhang, Identification of high frequency loads using statistical energy analysis method. *Mechanical Systems and Signal Processing*, 35 (2013) 291–306.

C. Xu, B.T. Cao, Y. Yuan, G. Meschke, Transfer learning based physics-informed neural networks for solving inverse problems in engineering structures under different loading scenarios. *Computer Methods in Applied Mechanics and Engineering*, 405 (2023) 115852.

S. M. Yang, G. S. Lee, Vibration control of smart structures by using neural networks. *Journal of Dynamic Systems, Measurement, and Control*, 119 (1997) 34–39.

H. Yang, J. Jiang, G. Chen, J. Zhao, Dynamic load identification based on deep convolution neural network. *Mechanical Systems and Signal Processing*, 185 (2023) 109757.

C. Yang, Interval strategy-based regularization approach for force reconstruction with multi-source uncertainties. C*omputer Methods in Applied Mechanics and Engineering*, 419 (2024) 116679.

X.W. Ye, T. Jin, C.B. Yun, A review on deep learning-based structural health monitoring of civil infrastructures. *Smart Structures and Systems*, 24 (2019) 567–586.

S.R. Yi, J. Song, First-passage probability estimation by poisson branching process model. *Structural Safety*, 90 (2021) 102027.

W. Yu, I.I.Y. Kim, C. Mechefske, Remaining useful life estimation using a bidirectional recurrent neural network based autoencoder scheme. *Mechanical Systems and Signal Processing*, 129 (2019) 764–780.

E. Zhang, J. Antoni, P. Feissel, Bayesian force reconstruction with an uncertain model. *Journal of Sound and Vibration*, 331 (2012) 798–814.

Z.H. Zhang, W.Y. He, W.X. Ren, Moving force identification based on learning dictionary with double sparsity. *Mechanical Systems and Signal Processing*, 170 (2022) 108811.

X.H. Zeng, H.M. Shi, H. Wu, Nonlinear dynamic responses of high-speed railway vehicles under combined self-excitation and forced excitation considering the influence of unsteady aerodynamic loads. *Nonlinear Dynamics*, 105 (2021) 3025-3060.

H.F. Zhou, Y.Q. Ni, J.M. Ko, Constructing input to neural networks for modeling temperature-caused modal variability: Mean temperatures, effective temperatures, and principal components of temperatures. *Engineering Structures*, 32 (2010) 1747–1759.

J.M. Zhou, L. Dong, W. Guan, J. Yan, Impact load identification of nonlinear structures using deep Recurrent Neural Network. *Mechanical Systems and Signal Processing*, 133 (2019) 106292.

Q. Zhou, Q.S. Li, B. Lu, Displacement estimation for a high-rise building during Super Typhoon Mangkhut based on field measurements and machine learning. *Engineering Structures*, 307 (2024) 117947.

X.Q. Zhu, S.S. Law, Recent developments in inverse problems of vehicle–bridge interaction dynamics. *Journal of Civil Structural Health Monitoring*, 6 (2016) 107–128.

F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109 (2021) 43–76.