PROBABILISTIC MACHINE LEARNING APPROACHES TOWARDS DIAGNOSTIC AND PROGNOSTIC OF DEGRADING ENGINEERING SYSTEMS

A thesis submitted in the fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

SHIVAM OJHA

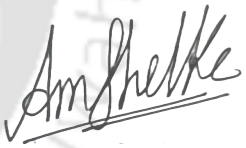
DEPARTMENT OF CIVIL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY, GUWAHATI
NORTH GUWAHATI - 781039, ASSAM , INDIA

April, 2025

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
DEPARTMENT OF CIVIL ENGINEERING

CERTIFICATE

It is certified that the work contained in the thesis entitled **"Probabilistic machine learning approaches towards diagnostic and prognostic of degrading engineering systems"** by **Shivam Ojha**, a research scholar, in the Department of Civil Engineering, Indian Institute of Technology, Guwahati, India for the award of the degree of the Doctor of Philosophy, has been carried out under my supervision and, that this work has not been submitted elsewhere for the degree.



Date: 16 April 2025

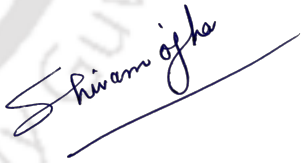
Dr. Amit Shelke
Associate Professor
Department of Civil Engineering
Indian Institute of Technology, Guwahati
Assam, India

Declaration

I, **Shivam Ojha**, declare that the work presented in the thesis entitled, **"Probabilistic machine learning approach towards diagnostic and prognostic of degrading engineering systems"** in partial fulfillment of the requirement for the award of the degree of Doctor of Philosophy is the outcome of the research work, performed by myself, else stated, under the guidance of Dr. Amit Shelke.

Any part of this work has not been submitted for the award of any degree, diploma, associate-fellowship, fellowship or its equivalent to any university or institution. Further, I confirm that:

- This work was conducted primarily during my candidature at this Institute.
- All sources used, including direct quotations, are appropriately cited.
- This thesis represents my original work except for clearly indicated contributions from others.
- All significant assistance received is acknowledged.
- For collaborative work, my specific contributions are specified.



Date: 16 April 2025

Shivam Ojha
196104026
Department of Civil Engineering
Indian Institute of Technology Guwahati
Assam, India

Acknowledgement

I am deeply grateful to Lord Shiva and Mata Parvati for their unwavering blessings, which guided me through challenging times and ultimately led to the successful completion of this dissertation.

This dissertation marks a significant milestone in my academic journey, transitioning from student to researcher. Its realization would not have been possible without the invaluable support of numerous individuals. I extend my sincere thanks to everyone who contributed directly or indirectly to this endeavor.

First of all, I want to thank my supervisor, Dr. Amit Shelke for giving endless important suggestions and critical reviews of my work through their guidance. His endless reassurance, recommendations and co-operation have been a great motivation for me while carrying out my research work. I will always be grateful to him for the knowledge he conveyed from his enormous experience. Besides my supervisor, I would like to acknowledge the doctoral committee members, Dr. Kaustubh Dasgupta , Dr. Budhaditya Hazra and Dr. Deepak Sharma for their valuable recommendations and insightful remarks throughout the course of research activities, which has constantly engaged me towards undertaking a meaningful thesis work.

Special thanks to Dr. Anowarul Habib from UiT, The Arctic University of Norway, for his valuable advice and contributions to my research. Additionally, I am grateful to Prof. Yongming Liu of Arizona State University and Dr. Tishun Peng of the Gas Technology Institute for conducting the experiment on aluminium lap joint and making the data available online (PHM 2019) for the structural health monitoring research community.

Further, special appreciation for the senior technician and support staff of the Department of Civil Engineering for their valuable help namely, Mr. Biswajit Debnath, Mr. Pranab Hazarika, Mr. Saurabh Mudoi, and Mr. Suresh Boro. I would also gratefully acknowledge the financial support provided by the Indian Space Research Organization (ISRO) (ISRO/RES/STC/IITG/2021-22), Government of India.

I am truly grateful to my colleagues, friends and seniors: Dr. Anupoju Rajeev, Dr. Lavish Gobind Pamwani, Dr. Pallab Jyoti Das , Dr. Pranjal Tamuly, Dr. Saroj Kumar Sahu,

Tapas Tripura, Ankush Gogoi, Satyam Panda, Mussa Kalimullah, Nilanjan Samanta, Ayush Thakur, Naveen Jangid, Abhishek Maharishi, Anushka Vashistha, Arpita Ghosh, Tori Basar, Shehnaz Deep, Dhananjoy Deo, Sahil Kumar, Shobhit Thakur for their enriching my PhD journey and for valuable discussions and contribution towards my dissertation.

Further, special thanks to Puru Mittal and Ajay Rajauriya for always being available to share my feelings and for providing the emotional support when I was feeling low.

Lastly, I extend my deepest gratitude to my sister, Prachi Sharma whose invaluable support was a cornerstone during the challenging periods of my journey. My parents, Mr. Mahesh Chand Ojha and Mrs. Reena Ojha, along with my sister, Shivani Ojha, deserve heartfelt thanks for their unwavering guidance, support and encouragement that propelled me forward every step of the way.

Shivam Ojha
Indian Institute of Technology Guwahati
Assam, India

Dedicated to my parents

Reena Ojha
Mahesh Chand Ojha

Abstract

Damage detection and prognosis are critical to managing degrading systems in the engineering infrastructure. The damage detection relying on sensor data and prognosis on historical failure data, usually both are treated separately. However, a holistic approach is needed to capture degradation at initial stages and prognosis. When implementing a machine learning framework for remaining useful life (RUL) estimation, two primary scenarios emerge: (1) utilizing a complete run-to-failure dataset for training, and (2) relying on a partial dataset. The latter presents significant challenges due to the necessity for extrapolation. Furthermore, limited research has explored extending damage detection results to RUL estimation by modeling the underlying degradation process using a surrogate measure. Moreover, few studies extend damage detection results to remaining useful life (RUL) estimation by modeling the underlying degradation process using a surrogate measure. An integrated probabilistic machine learning framework that systematically undertakes diagnosis, prognosis, and RUL estimation is lacking in the literature. Additionally, parameters for the surrogate degradation model used for prognosis are often estimated using conjugate gradient methods, which often fall into local minima. These surrogate models also ignore general population trends.

The main objectives of this thesis are threefold: First, a comprehensive prognosis framework is developed that integrates Gaussian Process Regression (GPR) with a robust solution to address extrapolation challenges, ensuring accurate and reliable predictions beyond the training data range. Second, the set of GPR hyperparameters for achieving the global optimum value of maximum likelihood is obtained through the proposed entropy-assisted genetic algorithm. This incorporates Renyi's entropy to diversify the initial population, achieve global convergence, and mitigate the risk of premature optimization. Finally, the proposed probabilistic framework is applied to real-world scenarios, specifically the prognosis of elastomeric rubber bearings and fatigue crack monitoring in lap joints. The aim is to demonstrate the framework's practical applicability, effectiveness in predicting RUL, and ability to account for uncertainties in different engineering domains.

Damage initiation in these applications is often followed by a sudden change in the degradation path. With this shock, the degradation path changes its trajectory, leading to

multi-stage degradation that requires shifting the adaptive mean function. Hence, this thesis attempts to locate change points through a Bayesian approach, presenting theoretical formulations for multi-functional forms in two phases (pre and post-change point) with damage detection in the context of structural health monitoring.

For case-specific surrogate measures, the development of a system identification tool using the two-stage constraint unscented Kalman filter (CUKF) is proposed for complex structural systems like elastomeric high damping rubber bearings. These are modeled through the Biaxial Bouc-Wen model to account for bidirectional effects observed in structures under seismic excitations. The effectiveness of the two-stage CUKF in parameter estimation is validated through numerical case studies and experimental validation, establishing a robust foundation for subsequent prognostics. Additionally, a case study from the aviation industry is investigated where a combination of automated wavelet feature extraction and a probabilistic Bayesian neural network directly correlates Lamb wave signatures with crack widths, addressing both aleatoric and epistemic uncertainties. This method provides probabilistic estimates of crack width and automates fatigue life prognosis, validated through extensive experimental data.

Results from numerical and experimental data show that the proposed surrogate degradation modeling framework is effective in performing prognosis. The analysis reveals that utilizing more degradation data post-change point leads to timely fault detection, improved model hyperparameter estimates, variance reduction, and reasonably accurate RUL predictions. Experimental validations underscore the robustness and reliability of the methods across diverse applications, paving the way for future innovations in the probabilistic machine learning diagnosis and prognosis of engineering systems.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Acronyms	Meaning
ARMA	Auto-regressive moving average
ANN	Artificial neural network
BW	Bouc-Wen model
BBW	Biaxial Bouc-Wen model
BGPR	Basis function Gaussian process regression
BNN	Bayesian neural network
CUKF	Constraint unscented Kalman filter
DL	Deep learning
DI	Damage index
EKF	Extended Kalman filter
EGA	Entropy assisted genetic algorithm
EMD	Empirical mode decomposition
GA	Genetic algorithm
GP	Gaussian process
GPR	Gaussian process regression
HDRB	High damping rubber bearing
HMC	Hamiltonian Monte Carlo
HMM	Hidden markov model
IMF	Intrinsic mode functions
Lin	Linear kernel
MCMC	Markov chain Monte Carlo
MODWT	Maximal overlap discrete wavelet transform
NDT	Non-destructive testing
NN	Neural network
PBNN	Probabilistic Bayesian neural network
PCA	Principal component analysis
PER	Periodic kernel
PHM	Prognosis and health management

Acronyms	Meaning
PML	Probabilistic machine learning
POD	Principal orthogonal decomposition
PZT	Peizoelectric
RMSE	Root mean square error
RUL	Remaining useful life
RQ	Rational quadratic
SE	Squared exponential
SHM	Structural health monitoring
SVM	Support vector machine
SSA	Singular spectral analysis
TVAR	Time varying auto-regression
UQ	Uncertainty quantification
UKF	Unscented Kalman filter
VI	Variational Inference
WT	Wavelet transform

Chapter 1

Introduction

1.1 Overview

The structural integrity of engineering systems is paramount in terms of the economic growth and development of the nation. However, these engineering systems are prone to degradation due to various causes like aging, fatigue, corrosion, successive damage due to operating loads and sudden damages. Ensuring their safety, reliability, and longevity is essential for protecting lives, safeguarding valuable assets, and minimizing economic disruptions. Structural Health Monitoring (SHM) plays a pivotal role in this endeavor, encompassing the continuous or periodic assessment of structural components and systems. It comprises two key components (a) diagnosis (Sohn et al., 2003), and (b) prognosis (Kim et al., 2017). This comprehensive approach aids in the early detection of degradation, allowing for trend analysis of damage progression that ultimately contributes to the anticipation of future component reliability. The complexities of mechanical and civil engineering structures present significant challenges for SHM approaches. The behavior of composites under complex loading scenarios leads to barely visible damage and can cause delamination, fiber breakage, and matrix cracking which can progress under operational conditions and unpredictable environmental factors (Staszewski et al., 2004; Coverley and Staszewski, 2003). Likewise, civil engineering infrastructure, owing to their inherent vulnerability, are at risk from aging, fatigue, and deterioration processes resulting from aggressive chemical attacks and other physical damage mechanisms (Ellingwood, 2005; Brownjohn, 2007). The detrimental effects of these phenomena can lead over time to unsatisfactory structural performance under service loading or accidental actions and extreme events like earthquake. Thus, the aging infrastructure and necessary monitoring for extending lifecycles put great emphasis on reliable and accurate SHM techniques (Farrar and Worden, 2012*a*). The SHM frameworks can be based on vibration methods, strain monitoring methods, model approaches, and guided wave methods, which observe changes in natural frequencies, stress-strain relationships, scattering and reflections for identifying the damage (Doebling et al., 1998; Friswell, 2007). However, these are sensitive to environmental and operational variability, face difficulty in addressing complex damage

modes, and cannot account for the uncertainties that pervade real-world engineering systems (Beck and Au, 2002; Sohn et al., 2003). The inevitable variations in material properties, the stochastic nature of loads, imperfections in sensor data and sparsity of the measurements contribute to aleatoric and epistemic uncertainties that can undermine the accuracy of diagnosis as well as prognosis (Sakin and Ay, 2008; Peng et al., 2015). Incorporating both aleatory and epistemic uncertainties into the SHM framework is essential for robust and reliable monitoring. Further, the most common causes of prolonged failure are fatigue and corrosion which are often encountered in SHM applications. Fatigue results from repetitive, cyclic loading and can lead to the initiation and growth of cracks or other damage modalities. The stochastic nature of fatigue processes, coupled with uncertainties in material properties and usage patterns, makes deterministic modeling of fatigue-driven damage progression inherently difficult (Suresh, 1998). The probabilistic methods provide a systematic approach to quantifying and managing these uncertainties, enabling more accurate assessment of damage severity, rate of growth, and remaining useful life. Recent advancements in machine learning (ML) and deep learning (DL) techniques, especially when integrated with probabilistic frameworks, have opened new avenues for enhancing Structural Health Monitoring (SHM). While traditional statistical approaches and deterministic ML methods have shown promise, their limitations in handling complex uncertainties and capturing nonlinear damage evolution have motivated a shift towards probabilistic ML models. These models not only excel at pattern recognition within vast datasets but also quantify prediction uncertainty, thus enabling risk-informed decision-making (Farrar and Worden, 2012*a*). This thesis contributes to this evolving landscape by advancing probabilistic SHM methodologies with a special focuses on the development of uncertainty-aware diagnosis and prognosis framework, leveraging Bayesian approaches like Gaussian Processes and Bayesian Neural Networks to address the challenges of computational complexity and data scarcity. An integrated probabilistic framework is proposed, incorporating Gaussian Process Regression (GPR) optimized with an Entropy-Assisted Genetic Algorithm (EGA) considering the adaptive mean functions and Bayesian change point detection. This framework is applied to the prognosis of multi-stage degradation processes, with a particular emphasis on damage evolution due to fatigue loading. The effectiveness of this approach is demonstrated through detailed case studies on elastomeric rubber bearings and fatigue crack monitoring in lap joints.

1.2 Background

Past decades are evidence of innumerable accidents and mishaps due to the sudden failure of structural operations and the performance of the civil systems. These failures underscore the subtle ways in which structures can weaken over time, leading to

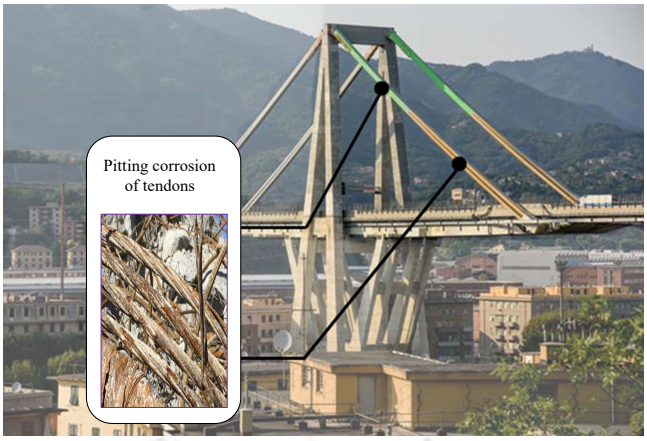


Figure 1.1: Illustration of primary reason of collapse in Morandi Bridge in Genoa, Italy in 2018 (Morgese et al., 2020)

unexpected and potentially devastating consequences. The iconic Tacoma Narrows Bridge in 1940 collapsed due to wind-induced oscillations. Further, Mianus River Bridge collapsed in 1983 in which corrosion and metal fatigue led to the failure of the pin and hanger assembly in the bridge's suspension system (Gorlov, 1984). This drives an increased focus on bridge inspection, but even now, hidden corrosion in post-tensioned concrete and other challenging areas underscores the need for vibration based SHM. The I-35W Bridge collapse in Minneapolis 2007 reminds us of the catastrophic failure, despite of periodic retrofitting, attributed to undersized gusset plates on the bridge's steel truss. Next, the collapse of a balcony constructed partly with composite beams in Dee Why Collaroy Surf Life Saving Club in Australia, in 2016 raised concerns about the long-term durability of composites in exposed marine environments. Recently, the collapse of the Morandi Bridge in Genoa, Italy in 2018 was attributed to a combination of factors, including design flaws, lack of maintenance, and environmental corrosion (Milillo et al., 2019; Calvi et al., 2019). Figure 1.1 shows the corrosion of pre-stressed tendons which fails due to prolonged fatigue, causing failure of the bridge. The collapse led to an increased focus on the maintenance and inspection of aging infrastructure, especially for bridges designed in the mid-20th century. Moreover, the tragic collapse of the Nicoll Highway in Singapore (2004) exposed shortcomings in traditional monitoring. This incident took place due to pre-collapse wall movements, emphasizing the need for automated SHM to potentially catch such warnings (Whittle and Davies, 2006). Further, the use of composites in wind turbine blades has dramatically improved their efficiency, but delamination, leading-edge erosion, and lightning damage are common occurrences. These failures affect power generation and, in extreme cases, pose safety hazards. These cases illustrate the critical need and role of SHM in ensuring the safety and integrity of infrastructure projects (Brownjohn, 2007).

Similar narratives exist within the aerospace industry. The tragic events of De Havilland

Comet crash in 1954 exposed the need to understand the impact of repeated pressurization cycles on aircraft structures, leading to advancements in fatigue analysis. The Aloha Airlines Flight 243 incident in 1988 is again a stark reminder of the consequences of undetected corrosion in aging aircraft, demanding changes in inspection protocols (Hendricks, 1991). In 2002, a Boeing 747 disintegrated mid-flight due to a previously undetected repair on the tail section, which was likely not in compliance with standards (Kumar and Padture, 2018). The crash highlighted the need for meticulous SHM throughout an aircraft's lifespan. Recently, the failure of a titanium component near a composite air intake contributed to the fatal crash of an F-22 in 2010. These accidents could have been easily averted by detailed inspection and routine maintenance. Apart from the case studies discussed so far, the structures generally undergo aging and degradation of strength and stiffness that requires long-term periodic collection and evaluation of structural responses. Also, under catastrophic events such as earthquake loading, blast, shock, and impact loading, the structures are often subjected to large magnitudes of loads (Krauthammer, 1999). These extreme loads cause severe damage to the structures and their components that affect their overall performance. The decision-making for reinstating the utility and safe operation of critical infrastructure is often questionable in the absence of the availability of structural health monitoring and damage assessment framework (Balageas et al., 2010). By employing sensor networks and advanced data analysis, SHM systems can monitor a structure's behavior in real time. Thus, we should take a vital step towards safeguarding the critical structures by embracing SHM .

1.3 Degradation of engineering systems

The failure in the engineering systems can be classified as (a) Hard failures, and (b) Soft failures. The hard failures are sudden in nature and generally triggered by external shocks or events, irrespective of the current system's health. This can occur due to natural disasters, or unanticipated overload conditions. Conversely, the failures resulting from gradual and internal degradation processes are termed soft failures. These failures occur when a system's deteriorating condition reaches a critical threshold, rendering it incapable of fulfilling its intended function (Nelson, 2009; Gorjian et al., 2010). Soft failures are a major concern in engineering structures due to their potential for gradual, often predictable deterioration. It can arise from factors such as fatigue, corrosion, and material degradation over extended periods of service. For instance, the components exposed to cyclic loading during flight operations experience fatigue which induces cracks in aluminum lap joints in the structural frame, gradually compromising structural integrity (Peng et al., 2015). Similarly, exposed body parts experience corrosion due to exposure to harsh environmental conditions, moisture, or chemical pollutants, which starts gradual deterioration and eventually leads to failure (Li and Chakik, 2021; Tavares and Castro, 2017). Infrastructures also face similar challenges related to corrosion and fatigue. The

steel structures as well as steel reinforcement within concrete rusts, compromise structural strength, particularly in coastal areas or environments with deicing salts. Further, concrete itself degrades through processes like alkali-silica reaction (ASR) and carbonation, while fatigue from traffic or wind can create cracks in bridges and other structures (Biondini and Frangopol, 2016*a*). While talking to bridge components, laminated rubber bearings are subjected to small cyclic deformation with an average shear strain of 0.5% by traffic loads (Fukada et al., 1998; Abe et al., 2004) leading to fatigue degradation. The details of the various cases of long term degradation in civil engineering from relevant literature is shown in Figure 1.2. In many of such scenarios, the underlying damage mechanisms themselves might not be directly observable, or obtaining accurate system parameters proves challenging. However, their effects can often be indirectly measured. For example, fatigue-induced stiffness loss in a metal connection could manifest in altered strain or acceleration readings, while compromised bonding in reinforced concrete could be revealed through changes in deflection. This underscores the importance of data-driven approaches, particularly surrogate modeling, to assess and predict the progression of degradation mechanisms like fatigue and corrosion, which are primary contributors to long-term structural deterioration

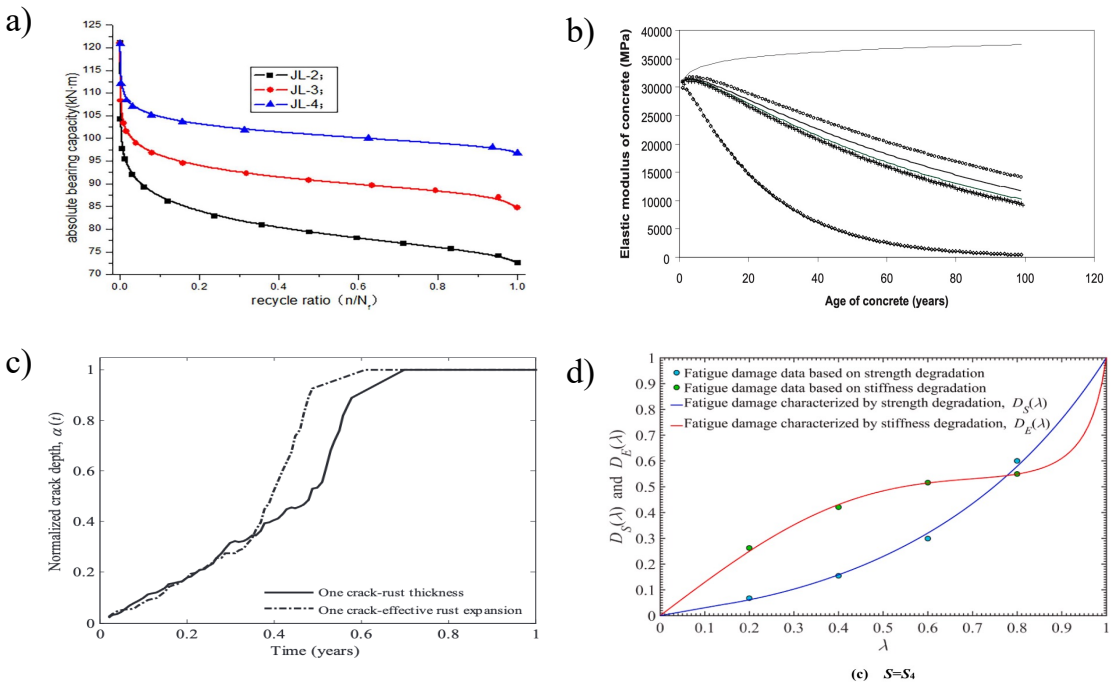


Figure 1.2: It illustrates the degradation in various engineering system in a long term (a) Reduction of stiffness as well as bearing capacity of RCC beam under long term fatigue (Liu and Yan, 2018), (b) Reduction in the elastic modulus of concrete due aging (important in case of earthen dams) (Burman et al., 2011), (c) Development of crack in the RCC structures due to prolonged corrosion (Zhong et al., 2010), and (d) Degradation in stiffness and strength in the composites due to fatigue (Gao et al., 2022)

1.4 Structural health monitoring

Structural Health Monitoring (SHM) is a systematic and multidisciplinary approach to evaluating the integrity of in-service structures and their components. It involves strategically deployed sensors, data collection, and sophisticated analysis to detect, locate, and quantify signs of structural damage or deterioration. The core goal of SHM is to enable proactive maintenance and intervention, minimizing the risk of failures and extending the lifespan of valuable assets like buildings, bridges, and aircraft (Sohn et al., 2003). SHM techniques often integrate elements of vibration analysis, statistical pattern recognition, and model-based simulations for comprehensive damage assessment. The various stages of structural health monitoring are demonstrated to form SHM cycle as shown in Figure 1.3, whose primary steps with increasing level of difficulty are as follows:

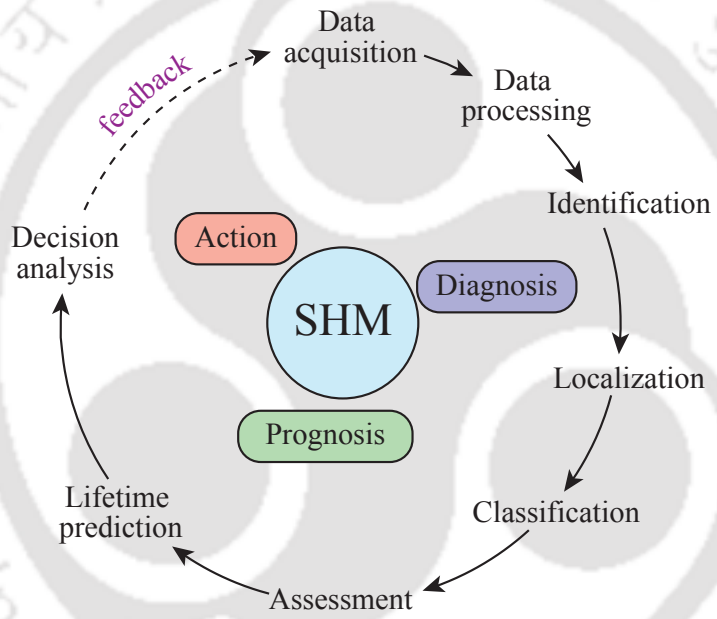


Figure 1.3: SHM Cycle illustrating the chain of steps involved in diagnosis and prognosis

- Level 1: Identification - It determines whether damage occurred on a global scale.
- Level 2: Localization - It identifies the location and coordinates of the damage.
- Level 3: Assessment - It determines the intensity of damage in various components.
- Level 4: Lifetime prediction - It forecasts the future state behaviour and estimates the structure's remaining life.

Within SHM, the concepts of diagnosis and prognosis are central. Diagnosis focuses on interpreting collected data to determine the existence, location, and severity of damage within a structure. Prognosis takes this further by predicting how existing damage might evolve and forecasting the structure's remaining useful life (RUL). The first two levels are defined as diagnosis steps, and the next two as prognosis levels. Importantly, levels 1 and 2 can be evaluated without using any model, but level 3 requires numerical modeling. The

last level requires processes such as fracture mechanics, fatigue-life analysis, reliability analysis, or structural design assessment. Additionally, level 4 stems from level 3, where the assessment of fracture parameters is utilized to achieve fatigue life analysis to specify the structure's remaining life. The diagnosis and prognosis together lead to the maintenance transformation and development of a complex regime of structural health monitoring called prognosis and health management (PHM) (Kim et al., 2017; Lee et al., 2014). The fundamentals of the diagnosis and prognosis are discussed in following sections.

1.4.1 Diagnosis of system

The diagnosis of the system identifies whether the damage has occurred or not in the global domain and then identifies the precise location of the damage. The first step of diagnosis is visual inspection which is carried out by trained professionals who carefully examine the structure for visible signs of damage, deterioration, or any unusual changes since previous inspections. This initial assessment helps pinpoint areas of concern that might need closer examination. To delve deeper, non-destructive testing (NDT) techniques are employed. These methods, such as ultrasonic testing, ground-penetrating radar, or infrared thermography, allow engineers to probe the structure's internal condition without causing harm. They provide valuable data on the structure's health and performance, enabling the detection of hidden defects or weaknesses. Moreover, structural health monitoring (SHM) systems will be installed for specific monitoring and condition assessment of the structure. These involve strategically placed sensors that continuously gather data on strain levels, vibrations, temperature, and other factors. The process begins with data acquisition using sensors to gather information about the structure's behavior. Subsequently, feature extraction from the sensor data is conducted to identify relevant patterns or anomalies that could indicate damage. After feature extraction, statistical modeling is employed to detect and analyze any detected damage, enabling a comprehensive assessment of the structure's health through the development of different physical or statistical health indicators (Doebling et al., 1996; Kim et al., 2006; Park et al., 2012). Moreover, system identification is carried out to assess the current state of the degradation (Yang et al., 2006). A significant limitation of traditional diagnostic approaches lies in their deterministic nature, often overlooking the inherent uncertainties present in real-world systems. This thesis addresses this limitation by introducing machine learning frameworks that adopt a probabilistic approach to diagnosis like Bayesian neural network for crack width estimation. These frameworks not only provide point estimates of damage occurrence, intensity, location, and quantification but also offer valuable confidence intervals and output distributions. This probabilistic information enables seamless integration with prognosis frameworks, in turn, enhances the reliability of

predictions regarding the remaining useful life of the structure.

1.4.2 Prognosis of degrading systems

Prognostics is the process of predicting the future reliability of a product by assessing the extent of derivation or degradation of the product from its expected norm operating conditions (Kim et al., 2017). It is the prediction of the future state of system's health based on current and historical conditions. The graphical illustration of prognosis is shown in Figure 1.4

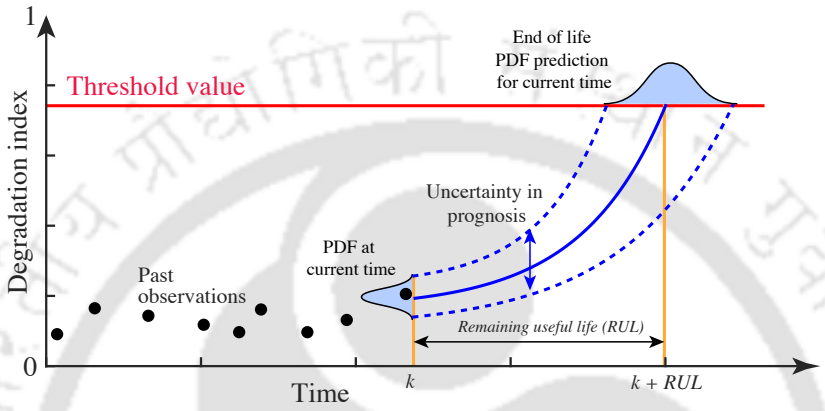


Figure 1.4: Prognosis fundamental

To monitor the evolution of degradation due to various mechanisms, a system may be subjected to intermittent online health monitoring. The acquired data, collectively known as degradation data, serve as a basis for degradation modeling and subsequent prognosis. While degradation modeling focuses on the selection, calibration, and validation of competing models against the acquired degradation data, prognosis is the systematic framework for assessing the future condition or performances of an infrastructure asset, given its present health condition. This is often done by projecting the degradation model to a predefined threshold. The process of predicting the future behaviour of degradation includes numerous sources of uncertainty as discussed in Section 3.2. Therefore, it is natural to consider the predicted degradation as a statistical distribution. That is, even if degradation is measured as a single point at the current cycle, the degradation is represented as a distribution due to uncertainty. Therefore, the time to reach the threshold is uncertain and can be represented as a probability distribution. Then, for a given time, the area under the distribution up to the time is considered as probability of failure. The failure probability increases as cycle increases. The end-of-life (EOL) is the corresponding cycle when the degradation reaches to the threshold and the maintenance should be ordered for the system. Since the degradation is a distribution, the cycle reaching to the threshold is not a deterministic value. The RUL, the remaining time to the maintenance from the current time, is also predicted as a distribution by subtracting the current cycle

from the EOL distribution, as

$$t_{RUL} = t_{EOL} - t_k \tag{1.1}$$

While numerous machine learning (ML) frameworks have been successfully applied to Structural Health Monitoring (SHM), their focus has primarily been on deterministic approaches (Salehi and Burgueño, 2018; Ochella et al., 2022; Xu and Saleh, 2021; Gordan et al., 2022). These methods often provide valuable insights into damage diagnosis and prognosis, but they typically neglect the inherent uncertainties that permeate real-world systems. This limitation motivates the development of advanced probabilistic ML frameworks capable of quantifying the uncertainty associated with their predictions. Such frameworks not only offer more comprehensive and reliable assessments of structural health but also enable risk-informed decision-making by explicitly considering the potential variability in damage evolution and remaining useful life estimates.

1.5 Motivation towards uncertainty quantified diagnostics and prognosis

Probabilistic machine learning (PML) is a branch of machine learning that focuses on explicitly representing and quantifying uncertainty. Unlike traditional deterministic approaches, PML methods explicitly account for uncertainty, providing a probabilistic framework that quantifies confidence in predictions. This richer information empowers better decision-making, especially in high-stakes scenarios like infrastructure maintenance where the consequences of error can be significant. Probabilistic machine learning has applications across diverse domains, including predictive modeling, classification, and regression. It is particularly valuable in safety-critical applications, where understanding and managing uncertainty are paramount for ensuring structural integrity and safety. The foundation of PML lies in the integration of probability theory with machine learning (Ghahramani, 2015*a*). The basic rules of probability theory are used to infer unobserved quantities given observed data. This learning process transforms prior probability distributions (defined before observing the data) into posterior distributions (after observing the data). This application of probability theory to learning is called Bayesian learning, which follows Bayes' theorem. Graphical models, in various forms like Bayesian networks, Markov networks, and mixed graphs, provide a powerful way to represent complex probabilistic relationships. Apart from Bayesian inferences and probabilistic graph models, Gaussian processes offer flexibility in modeling, providing both predictions and confidence intervals around them. Finally, deep probabilistic models combine the expressiveness of deep neural networks with probabilistic reasoning, enabling uncertainty modeling even in highly complex scenarios. The details about mathematical integration is discussed in Chapter 3. The primary advantage of uncertainty-aware diagnostics and

prognosis is the ability to provide more robust and reliable predictions. By accounting for the variability and noise in sensor data, these methods reduce the risk of false positives and false negatives. For example, in fatigue crack monitoring, an uncertainty-aware model can better distinguish between actual crack growth and noise-induced variations in sensor readings. This leads to more accurate estimations of the remaining useful life (RUL) of the structure, enabling timely maintenance actions and preventing unexpected failures. Another key benefit is the enhanced interpretability of the results. Probabilistic models provide a clear measure of confidence in their predictions, which can be invaluable for decision-makers. This transparency allows engineers to understand the risks associated with different maintenance strategies and make informed choices based on the likelihood of various outcomes. Moreover, uncertainty-aware diagnostics and prognosis facilitate the incorporation of expert knowledge and domain-specific insights into the modeling process. This integration is particularly useful in SHM, where historical data and expert opinions can significantly enhance the predictive performance of models. For instance, in the case of elastomeric rubber bearings, incorporating prior knowledge about material properties and loading conditions can improve the accuracy of the prognosis.

1.6 Scope and objectives of the thesis

This thesis explores and proposed probabilistic machine learning (PML) techniques to Structural Health Monitoring (SHM) with a focus on uncertainty-aware diagnostics and prognosis. The primary scope encompasses the development, implementation, and validation of advanced probabilistic frameworks aimed at enhancing the reliability and accuracy of diagnosing and predicting the health of engineering systems. The thesis integrates concepts from Bayesian inference, Gaussian process regression, and state-space modeling to address the inherent uncertainties in SHM data and provide robust, risk-informed decision-making tools. The global objectives of the thesis are illustrated as follows:

1. **Develop an uncertainty-aware degradation indicator** : This objective to develop and evaluate probabilistic surrogate models that explicitly quantify uncertainties in the prediction of structural health. By leveraging Bayesian approaches like Gaussian Processes and Bayesian Neural Networks, the goal is to enhance the reliability and robustness of prognosis for complex engineering systems, especially in scenarios with limited data or high computational costs.
2. **Develop an integrated probabilistic framework for prognosis**: Building on the foundation of uncertainty-aware diagnosis, the primary objective is to develop a comprehensive probabilistic framework for the prognosis of multi-stage degradation. This framework will integrate Gaussian Process Regression (GPR), optimized with an Entropy-Assisted Genetic Algorithm (EGA), along with adaptive mean functions and

Bayesian change point detection. The ultimate goal is to provide a holistic and accurate assessment of the remaining useful life of deteriorating structures, taking into account various sources of uncertainty.

3. **Validate and demonstrate practical application of the framework**: To validate the effectiveness and applicability of the proposed probabilistic framework through real-world case studies of elastomeric rubber bearings and fatigue crack monitoring in lap joints. These case studies will showcase the framework's ability to provide accurate and reliable predictions.

1.7 Overall methodology

The overall methodology for the diagnosis and prognosis adopted in the thesis is illustrated in Figure 1.5.

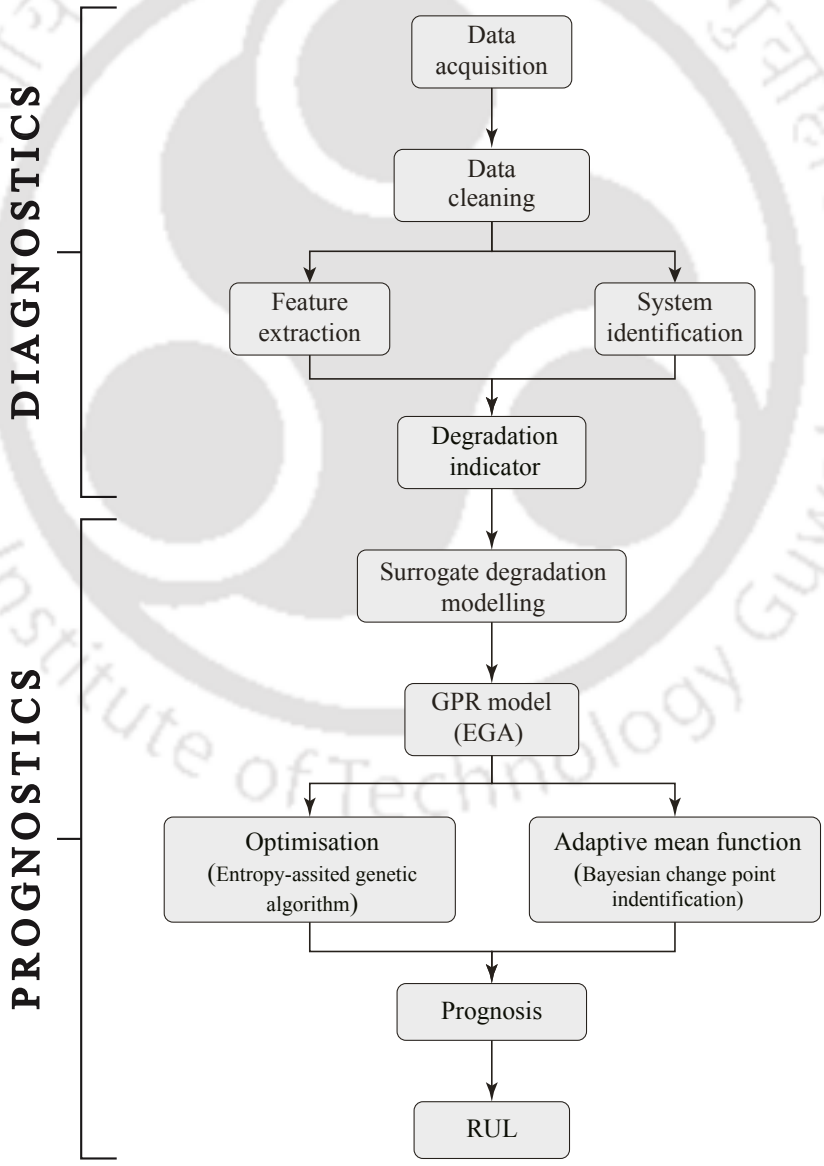


Figure 1.5: Illustration of overall methodology applied in the thesis for the diagnosis and prognosis of engineering systems

In the proposed methodology, first, data acquisition is followed by signal filtering to eliminate noise and irrelevant data. This filtered signal is then processed to obtain either system parameters or statistical damage indices, which are fundamental to the subsequent stages of SHM. For system identification and damage quantification, particularly in complex systems with intricate dynamics, a state-space model-based approach is proposed. Additionally, a statistical damage index is derived and proposed, utilizing vibrational data analysis for a comprehensive understanding of structural degradation. Moreover, for diagnosis, a Bayesian machine learning framework is introduced to leverage probability distributions and effectively account for both aleatoric and epistemic uncertainties, thus enhancing the accuracy and reliability of predictions. Next, the identified system parameters or statistical damage indices then serve as degradation indicators, forming the basis for prognosis through surrogate modeling. Gaussian process regression (GPR) is employed as the surrogate model, and its limitations in prognosis are addressed by incorporating entropy-assisted genetic optimization (EGA) and adaptive mean functions. Once the surrogate model is adequately formed and optimized, it enables accurate prognosis and the determination of the remaining useful life (RUL), culminating the SHM process with valuable insights into the structural health and longevity.

1.8 Organization of the thesis

A flow chart showing the thesis organization is given in Figure 1.6, which contains total seven chapters as follows:

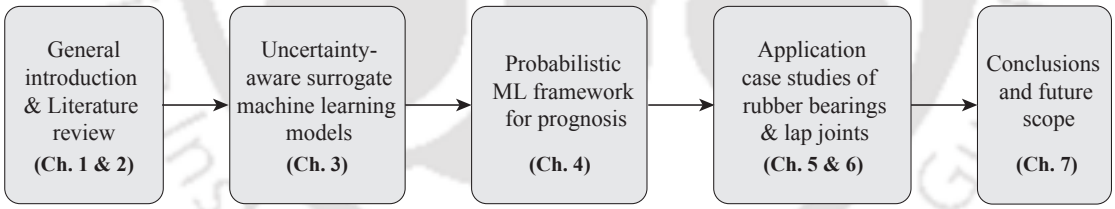


Figure 1.6: Organization of the thesis illustrating different part and chapters

- Chapter 1 introduces the need for the probabilistic health monitoring approaches and provides a brief introduction to the structural health monitoring explaining the diagnosis and prognosis for critical engineering system. Later, the motivation and a summary of research objectives, and overall methodology adopted in the thesis are presented. Finally, chapter concludes with the organization of the thesis.
- Chapter 2 provides an extensively detailed background and literature review on damage detection and prognosis methods for engineering systems, especially focussed on the civil infrastructure and industry applications, primarily related to the fatigue degradation. This chapter has two parts, first contain the literature review of the diagnosis methodologies and second part contains the survey of prognosis

methods. Both of these part provide extensive survey of probabilistic machine learning frameworks that are developed earlier. However, for the prognosis, they are provided in well organized tabular form. Lastly, the specific objectives of this thesis are described following the identified gaps in the literature.

- Chapters 3 introduces the mathematics of uncertainty informed machine learning models which originates from the fusion of probability theory and machine learning. Further, the importance of the uncertainty quantification and its mathematical expressions in the closed form is presented. The Bayesian inference, limitations of Gaussian process regression for the prognosis and extrapolation are discussed in details. Finally, this chapter forms the basis for the development of the prognosis framework and case studies which are presented in subsequent chapters.
- Chapter 4 introduces a novel data-driven framework designed to address the challenges of predicting the remaining useful life (RUL) of engineering systems undergoing multi-stage degradation. The methodology leverages Gaussian Process Regression (GPR), optimized with an Entropy-Assisted Genetic Algorithm (EGA) and adaptive mean functions, to model the complex, nonlinear degradation processes for prognosis, typical in structural health monitoring applications.
- Chapter 5 presents a case study of diagnosis and prognosis on elastomeric rubber bearing. It proposes a two-stage constraint unscented Kalman filter for the parameter estimation of degrading complex systems under biaxial excitations. This two stage framework helps in handling computational complexity and overconfidence issue which evaluating variance in large number of states. Further, identified hysteresis is used to develop degradation indicator and prognosis is achieved through proposed framework.
- Chapter 6 presents another case study of fatigue prognosis in aluminium lap joints. Here, probabilistic crack width is examined through Bayesian machine learning frameworks and wavelet transform. Further, the estimated crack width is used as degradation indicator and prognosis is achieved through proposed framework.
- Chapter 7, finally discussed the number of conclusions resulting from the presented work. Several recommendations for future study are also discussed.

Chapter 2

Literature review

2.1 Introduction

This chapter explores the field of probabilistic machine learning frameworks for the diagnosis and prognosis of engineering systems. The literature survey emphasizes the use of these frameworks to identify the damage, and to predict future system health and remaining useful life. Probabilistic machine learning frameworks are highlighted for their ability to address the uncertainties inherent in complex engineering systems. The chapter begins by providing a background on diagnosis and prognosis, offering a concise overview of relevant SHM methodologies and the specific challenges within this domain. Next, various probabilistic machine learning frameworks for prognosis are reviewed and presented in the tabular form. Finally, the chapter concludes by outlining promising future directions and highlighting existing gaps to further advance the field of probabilistic diagnosis and prognosis for engineering systems.

2.2 Diagnosis methodologies for engineering systems

A wide range of non-destructive techniques have been established for damage detection in civil infrastructure, including acoustic emission, eddy current, thermal field, radiography, strain measurements, magnetic field methods, and ambient vibration response monitoring (Mutlib et al., 2016; Siringoringo and Fujino, 2008; Sun et al., 2015). The underlying principle is that structural damage manifests through alterations in measurable quantities, such as stiffness, damping, or mass. These changes are reflected in vibration data and modal properties (frequencies, mode shapes, and modal damping) (Brownjohn et al., 2011; Doebling et al., 1998; Wei and Pizhong, 2011). By tracking the evolution of these parameters, engineers can infer the presence and potential location of damage. Three main categories of diagnosis methods exist in literature:

1. **Physics-based methods:** These methods rely on governing equations to create a mathematical model of the structure. This model is then compared with measured vibration data to quantitatively assess damage (Pandey et al., 1991; Shi et al., 2000).

While effective, these models often require complex finite element analysis, which can be computationally expensive for large structures or require analytical expression which need extensive research to develop.

2. **Data-driven methods:** These methods leverage statistical principles and measured data to construct models (input-output models like autoregressive (AR) family of time-series models) representing observations. These models are computationally less expensive and can be applied to various measurement types (Figueiredo et al., 2011; Manson et al., 2001; Worden and Lane, 2001). However, a key drawback is that the measured quantities might not directly correspond to the underlying damage mechanism, potentially hindering accurate damage detection. To avoid the potential drawbacks, several data driven mechanisms are developed to achieve problem specific solutions. Few examples of data-driven tools for damage detection include:

   (a) Auto-regressive moving-average (ARMA) time series models (Nair et al., 2006; Omenzetter and Brownjohn, 2006; Sohn and Farrar, 2001)

   (b) Support vector machines (SVM)(Bornn et al., 2009; Worden and Lane, 2001)

   (c) Neural networks (NNs) (Fang et al., 2005)

   (d) Self-organizing maps (SOMs) (Tibaduiza et al., 2013)

   (e) Principal component analysis (PCA) (Manson et al., 2001)

   (f) Statistical process control (SPC) (Sohn et al., 2000; Wang and Zhang, 2008)

3. **Hybrid methods:** These approaches to damage detection in engineering systems integrate multiple diagnostic techniques to overcome the limitations of single-method approaches. These approaches often blend low-frequency structural vibration data with advanced nonlinear dynamic analysis to enhance sensitivity to local defects and robustness against uncertainty.

The development of damage index rely on surrogate measures like Mahalanobis distance, residual error, or features extracted from autoregressive moving average coefficients to indirectly assess damage. The classification of the diagnosis methodologies is shown in Figure 2.1 and the details of various types of the data-driven diagnosis approaches are presented in next section.

2.2.1 Data-driven diagnosis methods

The data-driven methods forms the fundamentals of damage detection from ambient vibration data, frequency response function and the system modal parameter updates. Further, machine learning based frameworks are developed for damage detection. The details are presented below:

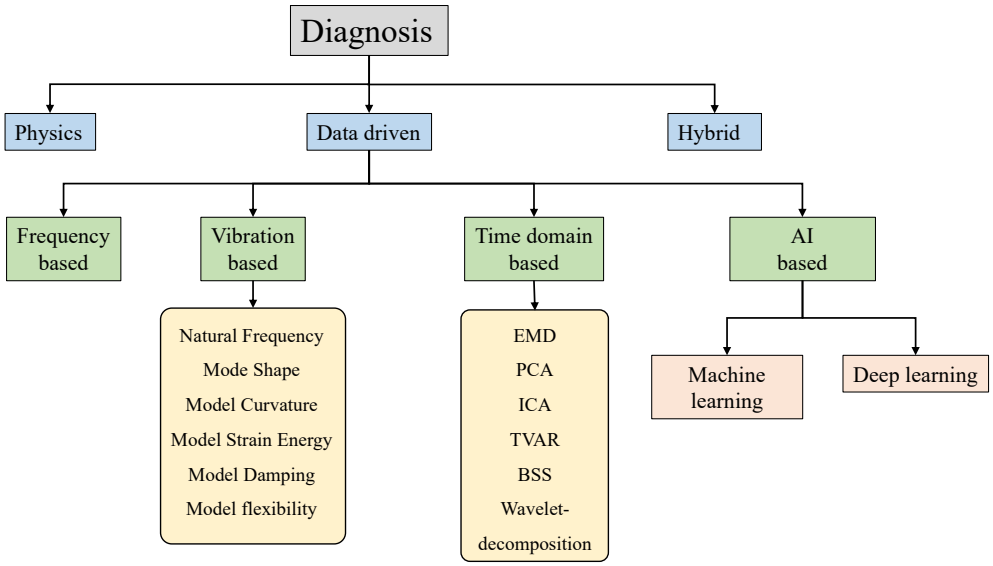


Figure 2.1: Illustration of classification of diagnosis methodologies in structural health monitoring

Vibration based methods

Ambient vibration data is integral to the structural health monitoring (SHM), providing crucial insights into the physical properties of structures and their functional integrity. These data are typically derived from modal parameters such as natural frequencies, mode shapes, and modal damping ratios, which help identify damages by highlighting changes in mass, stiffness, and damping properties of the structure. The sensitivity of these parameters to damage varies, with damping and mass being the most and least sensitive respectively. Despite the wide application of these methods, limitations persist including their sensitivity to specific damage types, dependency on baseline data from an undamaged structure, and the necessity for extensive data storage, particularly in complex structures like composites (He et al., 2018; Kamal and Taha, 2010; Luengo et al., 2016). This vibration data is used for extraction of system characteristic features which indirectly represents the damage in the system. These characteristic features area listed below:

1. **Natural Frequency:** Reduction in structural stiffness due to damage typically results in decreased natural frequencies. These frequencies are often easily and economically measurable, making them popular in damage detection studies (Doebling et al., 1996; Lifshitz and Rotem, 1969). Despite their susceptibility to environmental influences, such as temperature variations, their sensitivity to structural changes makes them valuable for assessing damage severity and location.

2. **Mode Shapes:** Mode shapes are less affected by environmental factors and provide spatial information crucial for localizing damage. Techniques like the Modal Assurance Criterion (MAC) and its derivatives utilize mode shapes to effectively detect and localize structural damages across various engineering disciplines

(Allemang, 1982; Salawu and Williams, 1995).

3. **Modal Curvature:** This method focuses on the second derivative of mode shapes to detect stiffness loss due to damage (Moughty and Casas, 2017; Pandey et al., 1991). Although highly sensitive, the requirement for multiple sensors and potential noise amplification limit its application (Capecchi et al., 2016).

4. **Modal Strain Energy:** Based on the energy stored in deformed structures, this method shows promise in localizing and quantifying damages, particularly in beam structures modelled using finite element methods. However, the need for extensive sensor networks and the influence of noise remain challenges (Cornwell et al., 1999; Duffey et al., 2001; Farrar and Worden, 2012*b*).

5. **Modal Damping:** Although less commonly utilized due to its complex nature and modelling challenges, studies have shown that damping variations can indicate damage locations effectively, even without baseline data (Ay et al., 2019; Cao et al., 2017; Chandra et al., 1999; Mustafa et al., 2018).

6. **Modal Flexibility:** Using the inverse of the stiffness matrix, modal flexibility methods offer high sensitivity to damage. These methods benefit from fewer mode requirements and have shown potential in damage localization and assessment through dynamic simulations (Lu et al., 2002; Shih et al., 2009).

Frequency based methods

Frequency response functions (FRFs) represent a broader spectrum of frequencies than modal data and are valuable in damage detection for their detailed response profiles. They are typically used in conjunction with optimization-based model updating problems to refine the structural models based on observed damages (Hassani and Shadan, 2022). However, the selection of appropriate excitation frequencies and the need for precise knowledge of input forces are significant challenges (Limongelli, 2010).

Time domain based methods

The vibration based methods often fall short in handling the non-stationary nature of damage-induced signals. Advanced time domain analysis techniques have therefore been developed to better represent energy variations within such signals, crucial for accurate damage detection. These techniques are:

1. **Principal component analysis (PCA):** It identifies changes in structures by decomposing vibration data into principal components that highlight anomalies. This statistical method efficiently isolates the significant features, facilitating early detection of structural impairments.

2. **Singular spectral analysis (SSA):** It is applied in damage detection to decompose time

series data into its singular components, enabling the identification of subtle structural changes. This technique isolates noise from the underlying signal patterns, enhancing the accuracy of detecting potential damages. SSA's ability to reconstruct the data with selected components makes it a powerful tool for early diagnosis and monitoring of structural health.

3. **Empirical Mode Decomposition (EMD):** It decomposes complex signals into intrinsic mode functions (IMFs), which represent natural oscillation modes. Despite its potential, EMD suffers from mode mixing issues, which have been addressed by variants like Ensemble EMD (EEMD) and Multivariate EMD (MEMD), improving noise robustness and applicability to multi-channel signals (Barbosh et al., 2020; Barile et al., 2020; Huang et al., 1998; Wu and Huang, 2009).

4. **Time varying autoregressive modelling (TVAR):** TVAR models are pivotal in damage detection, adapting to structural changes by capturing temporal variations in dynamic data. This method allows for the continuous assessment of structural integrity by analyzing how regression coefficients evolve over time. TVAR's sensitivity to subtle shifts in data makes it effective for early warning, real-time monitoring and preventative maintenance in engineering structures.

5. **Wavelet Transformation (WT):** WT has been widely adopted in structural health monitoring (SHM) due to its sensitivity to anomalies in vibration signals. Techniques such as Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT) have been used for spatial damage detection across various structures (Chang and Chen, 2005; Liew and Wang, 1998; Zhong and Oyadiji, 2007). WT is favored for its ability to localize damage through detailed analysis of wavelet coefficients, often combined with neural networks to enhance damage detection accuracy (Wu et al., 2021; Hein and Feklistova, 2011).

2.2.2 Artificial Intelligence Approaches

AI is revolutionizing diagnosis by mimicking human intelligence to surpass human efficiency in complex tasks (Goodfellow et al., 2016). While initially successful in rule-based problems, traditional AI struggled with tasks like facial recognition due to the difficulty of codifying human intuition (Ghahramani, 2015*b*; Müller and Guido, 2016). To address this, machine learning (ML) was developed, enabling systems to learn from data without explicit programming (Santos et al., 2016). ML extracts key data features to train models, and is categorized into two main types:

- **Supervised Learning:** This involves training algorithms on a labelled dataset, where the desired outcomes are known, allowing the model to learn the relationships between inputs and outputs. This is particularly effective in regression and

classification tasks, where the algorithm predicts a numeric value or classifies data into categories, respectively (Kubat, 2017; Figueiredo et al., 2011).

- **Unsupervised Learning:** Unlike supervised learning, unsupervised learning algorithms do not require labelled data. They work by uncovering hidden patterns or intrinsic structures in input data, which can be particularly useful for discovering unknown phenomena in the data or for organizing data in meaningful ways (Yegnanarayana, 2009).

Further, the deep learning (DL), a subset of ML, has revolutionized the field by reducing the dependence on manually extracted features. Deep Neural Networks (DNNs), a key component of DL, automatically extract and learn features directly from raw data (Kwon et al., 2019; Pouyanfar et al., 2019). This capability allows DL to handle complex, high-dimensional data in a way that earlier ML techniques could not. By learning hierarchical feature representations, starting from simple, low-level features and building up to high-level, abstract ones, DL models can perform sophisticated diagnostic tasks with high accuracy. DL is particularly renowned for its performance in areas where feature representation plays a critical role, such as image and speech recognition or complex decision-making tasks. Its ability to learn from large volumes of data and its capacity to improve continually as it is exposed to more data make it an invaluable tool in the diagnostic processes across various fields (Kiranyaz et al., 2015; Meruane et al., 2021; Zhao et al., 2018).

2.3 Probabilistic frameworks in diagnosis of engineering systems

Till now, we have explored the general categories of diagnosis in structural health monitoring, addressing various damage identification methods that yield a single quantitative outcome. However, in practice, results are not always definitive due to the numerous uncertainties inherent in structural health monitoring. In response to these challenges, researchers have sought to probabilistically diagnose systems by incorporating Bayesian statistics into the field. The next section will elaborate the probabilistic framework developed in the past related to the diagnosis of engineering systems.

2.3.1 Statistical Bayesian framework for damage detection

In early 90s, the probabilistic system identification is carried out by Beck where the system health is identified through changes in stiffness distribution and model updating using vibration measurement. This study uses Bayesian formulation and provide asymptotic expression for the posterior distribution of model parameters (Beck and Katafygiotis, 1992). Further, Sohn (1999) presented a thesis on Bayesian probabilistic approach to damage detection which discusses diagnosis with a few measurement points and modes,

systematically extracting information from continuously obtained test data. The methodology involves a Bayesian probabilistic framework for damage detection, formulation of relative posterior probability based on output error, a branch-and-bound search scheme, integration of load-dependent Ritz vectors, and a procedure for extracting Ritz vectors from vibration tests.

Vanik et al. (2000) proposes the approach to real time monitoring through time variation of a probabilistic damage measure. The paper emphasizes the importance of structural health monitoring, the incorporation of theoretical development in the method, and its ability to account for uncertainty and be suitable for on-line application. The method is applied to simulated data to test its effectiveness.

Sohn and Farrar (2001) proposes a probabilistic approach for damage detection in civil structures, specifically applied to predict the location of plastic hinge deformation in a reinforced-concrete bridge column using experimental data from vibration tests. The method included a simplified analytical model and the updating of damage probabilities with new test data.

Ching and Beck (2004) proposes a two-step approach for probabilistic structural health monitoring, including modal identification, damage assessment using a Bayesian model updating algorithm, addressing incomplete mode shape information, and utilizing the Expectation-Maximization algorithm. The optimization problem is simplified into two quadratic optimization problems The approach was illustrated using the Phase II Simulated Benchmark problems, demonstrating its effectiveness in structural health monitoring and showed promise in addressing structural health monitoring problems with incomplete mode shape information.

Saito et al. (2005) develops a framework for probabilistic estimation of structural damage, utilizing system identification of a building with observed vibration data, applying the method to vibration data from shaking table tests, accurately estimating story stiffness, and evaluating the reliability of estimated values.

Yang et al. (2006) introduced an adaptive extended Kalman filter for state and parameter estimation in structural systems. This method was validated through simulations that tracked changes in nonlinear elastic, hysteretic, and linear benchmark structures, demonstrating its effectiveness in damage detection using vibration data. Advanced Kalman filters like unscented Kalman filter (UKF), in particular, generates sigma points to capture the mean and covariance of the system's posterior distribution, which are then transformed back to update the states. To simplify calculations and align solutions with the physical properties of the system, a variation known as the constrained Kalman filter has been developed. Applications of constrained EKF and UKF for the system

identification of Bouc-Wen model are detailed in various studies (Li and Wang, 2021; Calabrese et al., 2018; Yang and Ma, 2003).

Ciampoli et al. (2008) suggests that structural health monitoring requires both damage detection and assessment of the effects of damages on reliability. The methodology involves a complete procedure for structural health monitoring, utilizing techniques such as damage detection with unknown input, Bayesian model updating for quantifying structural damage, and subset simulation for assessing the probability of exceeding response levels while considering uncertainties.

Cangru (2009) proposes the Bayesian wavelet probabilistic approach for nonparametric damage detection, data cleansing using a Bayesian discrete wavelet packet transform-based denoising approach, a nonparametric system identification method based on fuzzy wavelet neural networks for predicting dynamic responses. Bayesian hypothesis testing for assessing differences between sensed data and model prediction, and construction of a probability density function for the Bayesian assessment metric using Monte Carlo simulation to account for uncertainties.

Lombaert et al. (2009) carried out vibration-based finite element (FE) model updating, and Bayesian FE model updating are used for damage identification, where the stiffness of substructures is updated to identify and localize damage. Bayesian inference is employed to compute the posterior joint probability density of the stiffness in the substructures.

Mahadevan et al. (2008) presents probabilistic methodologies for designing a reliable structural health monitoring system, optimizing sensor layout for enhanced damage detection reliability, and introducing Bayesian methods to address uncertainties in sensor data and model predictions.

Sankararaman and Mahadevan (2011) provides the development and comparison of classical statistics-based methods and Bayesian statistics-based methods for quantifying uncertainty in all three steps of damage diagnosis, with a focus on combining uncertainties to quantify overall uncertainty in diagnosis. Further, a metric based on least squares is proposed to assess the uncertainty in damage localization when the damage signatures fail to localize the damage uniquely. The uncertainty in damage quantification is evaluated through statistical non-linear regression, resulting in confidence bounds for the damage parameter. The uncertainties in damage detection, isolation and quantification are combined to quantify the overall uncertainty in diagnosis.

Rabiei (2011) presented thesis on Bayesian framework for structural health management using acoustic emission monitoring and periodic inspections. A novel structural health management framework was proposed using Bayesian knowledge fusion, enabling the estimation of important quantities for structural health management. The methodology

involves conducting fatigue crack growth tests with acoustic emissions monitoring, developing a statistical model, employing Bayesian regression technique for parameter estimation, and using recursive Bayesian estimation for online integration of assessment information.

Zhang et al. (2011) suggests probabilistic simultaneous identification method which can decrease false positives of damages and achieve accurate identification for structural unknown input. The methodology involves a probabilistic approach combined with a deterministic simultaneous identification method to identify damages in structures with uncertainties under unknown input. It includes investigating the effect of uncertainties, calculating statistical parameters, deriving a damage index, and determining the probability of identified damage.

Haynes and Todd (2012) applied statistical damage detection techniques to a three-story structure with bolted joints, using a sparse network of PZT sensor-actuators bonded to the structure. Receiver operating characteristics are used to quantify detector performance, and the detection rate is compared across different types and levels of damage.

Simoen et al. (2013) applied Bayesian linear finite-element (FE) model updating for uncertainty quantification (UQ) in the vibration based damage assessment of a 7-story RC building slice. The Bayesian UQ approach and subsequent resolution analysis are shown to be effective in assessing uncertainty in FE model updating. Furthermore, it is demonstrated that the Bayesian FE model updating approach provides insight into the regularization of its often ill-posed deterministic counterpart.

M. Dirbaz (2013) presents a thesis on development of Bayesian methods (BBCA and BSCA) for assessing bridge conditions based on subjective ratings and sensor/modal data, with demonstrated applicability to highway bridges. Huang and Beck (2013) presented sparse Bayesian approach that eliminates the need for a threshold for issuing a damage alarm and reduces false-positive and false-negative damage detection in the presence of modeling error.

Tee et al. (2012) introduces a quantile regression approach using vibration data, based on quantile autoregressive time series models, to study differences between distributions of undamaged and damaged structures. Performance evaluated through simulation studies and experimental validation.

Rabiei and Modarres (2013) presents a Bayesian framework for integrating structural health assessment information, utilizes an AE-based monitoring approach for crack growth rate observations, and uses updated parameters for prognosis based on a future usage profile.

Zheng and Yu (2013) proposed probabilistic framework can effectively draw statistical samples that approximate the updated probabilistic distributions of uncertain model parameters and provide model evidence that can quantify probabilities of uncertain model classes. The framework integrates the transitional Markov chain Monte Carlo algorithm with the finite element model, considering uncertainties in truss joint configurations and model parameters, and examining the framework's effectiveness through an application.

Hussein and Haldar (2015) developed a new method, UKF-UI-WGI, which was found to be superior to the previously used GLIS-EKF-UI for the rapid assessment of infrastructure health. The new method is recommended for use at the beginning of an inspection when defects and their severity are unknown. The methodology involved system-identification techniques with ignored excitation information, finite element representation of structures, tracking stiffness parameter changes, verification through investigations, comparison of GLIS-EKF-UI and UKF-UI-WGI, and recommendation of UKF-UI-WGI for rapid health assessment.

Yang et al. (2016) implemented Bayesian inference method using random walks to estimate the reliability of structures, a significant reduction in uncertainties associated with the estimation, and validation through numerical and practical examples through estimating the reliability of structures.

Mukhopadhyay et al. (2016) discusses a probabilistic method for assessing structural damage, considering variability in stiffness estimates and using experimental data from a 3-story sliding base frame. The methodology used in the study is a probabilistic method to assess structural damage location and severity, considering variability in healthy and damaged states through empirical distributions of stiffness estimates.

Cai and Mahadevan (2018) explores big data analytics in uncertainty quantification for structural diagnosis and prognosis, utilizing MapReduce technique and Monte Carlo simulation for efficient handling of sensor data and predicting damage growth. The methodology involves using the MapReduce technique for parallelizing data analytics, Bayesian methods for diagnosis, and Monte Carlo simulation for predicting damage growth and remaining useful life, with a focus on parallelizing FEA runs.

Prakash and Narasimhan (2018) aims to develop a mathematical framework using two-phase gamma process for damage detection and prognosis using a surrogate measure of degradation, integrate historical failure-time data with monitored data for RUL estimation and maintenance planning, and utilize change point identification for two-phase degradation modeling and future degradation level prediction.

Kim et al. (2018) adopted ARX model to consider environmental and operational changes led to higher accuracy compared the linear regression without consideration of the effects.

The methodology involves identifying damage indicators, performing regression analysis using Bayesian regression to remove environmental and operational influences, conducting decision-making analysis through Bayesian hypothesis testing, utilizing a state space model for the Kalman filter, employing an ARX model to handle external disturbances, and utilizing Bayesian statistics to estimate the probability of damage.

Liu et al. (2018) introduces a novel approach using state-space probability models to create a damage metric robust to environmental and operational variations. The research uses state-space probability models for structural health monitoring, defining algorithms for training and prediction, and explaining the application of the algorithm in the presence or absence of measurements for external conditions.

Davíð (2019) implemented the machine learning approach and successfully detected realistic damage stages in a bridge structural health monitoring system, making it a viable option for operational bridge monitoring. The methodology involved using the Z-24 bridge benchmark dataset for validation, training a Bayesian autoencoder neural network to reconstruct sensor data sequences, comparing reconstruction errors with a healthy-state error distribution, and successfully detecting various damage stages.

Vega and Todd (2020) uses a Bayesian neural network with variational inference to learn a damage feature from a high-fidelity finite element model, demonstrating its application to miter gates. The architecture includes 2 layers with 50 neurons in the hidden layer, with regularization to penalize architectures with more parameters. Activation functions like sigmoid and soft plus are used in the hidden and output layers. The main findings include the use of a degradation model based on real inspection data, evaluation of the added value of SHM in miter gates in terms of maintenance cost savings, and a demonstrated maintenance cost reduction of 11.1% to 35.1% when using OCA ratings with a surrogate model based on a physical model.

Tehrani et al. (2021) proposes the Bayesian multiple modeling approach is used for improving data fitting precision and increasing stability of the identification algorithm. The proposed algorithm is successfully implemented on buildings with varying numbers of unknowns. The method can simultaneously select the optimal model class and identify its unknown states and parameters in an online manner. The methodology involves utilizing the Bayesian multiple modeling approach, a time-domain method with the central difference Kalman filter, testing on buildings with different numbers of unknowns, and conducting robust simulations for verification.

Sajedi and Liang (2021) proposed framework that can be applied to future deep vision SHM frameworks to incorporate model uncertainty in the inspection processes. The methodology involved using Bayesian inference and Monte Carlo dropout sampling for deep vision SHM

models, conducting three case studies, and proposing a framework for future applications.

Amer and Kopsaftopoulos (2022) involves using damage indices and Gaussian process regression models for structural damage quantification, comparing standard and variational GPRMs, training the models with DI values, applying a state prediction framework, and assessing performance in different test cases.

Hence, system identification in structural health monitoring (SHM) has been extensively developed using Bayesian approaches. The early studies focused on identifying structural health through changes in stiffness distribution and model updating using vibration measurements, which led to the development of Bayesian probabilistic frameworks for damage detection, systematically extracting information from limited measurement points and modes. The methodologies incorporate Bayesian model updating, relative posterior probability formulation, and search schemes for identifying damage.Real-time monitoring approaches emerged, leveraging probabilistic damage measures and incorporating load-dependent Ritz vectors. Advanced Bayesian methods were applied for predicting plastic hinge deformation in reinforced concrete structures, demonstrating effectiveness in damage probability updating. The development of Bayesian model updating techniques addressed incomplete mode shape information using optimization and expectation-maximization algorithms.Further advancements integrated Kalman filtering techniques for state and parameter estimation, particularly for nonlinear structural systems. Probabilistic damage estimation frameworks combined system identification and Bayesian inference to track stiffness changes and assess reliability. Hybrid methods incorporated nonparametric wavelet transforms, machine learning, and Monte Carlo simulations to enhance damage detection accuracy.These methods have improved SHM accuracy, uncertainty quantification, and maintenance decision-making.

2.3.2 Probabilistic machine learning diagnosis approaches

Machine learning (ML) diagnosis frameworks are increasingly essential alongside traditional statistical methods due to several advantages. ML excels in handling complex, high-dimensional data and adapts over time, enhancing diagnostic accuracy as new data emerges. Its predictive capabilities allow for proactive maintenance strategies, crucial for extending infrastructure lifespan. ML models are particularly adept at detecting anomalies and nonlinear relationships, crucial for identifying subtle damages. Additionally, ML can be automated and scaled easily, integrates well with modern digital technologies like IoT, and offers customization to fit specific needs. This makes ML a versatile and powerful tool in structural health monitoring, complementing and enhancing the capabilities of statistical approaches. A brief review of machine learning approaches for the damage detection is presented as follows.

Worden and Manson (2006) illustrates the utility of data-driven approaches for damage identification through case studies and explains the statistical model used in structural health monitoring. Brehm and Deraemaeker (2014) presents a methodology for structural damage detection using artificial immune systems, which successfully detects damages in the experimental setup. The methodology involves using artificial immune systems for structural damage detection, inspecting the structure with a sensor network at different actuation phases, defining baselines using data from a healthy structure, and detecting damage in an unknown structural state.

Mechbal et al. (2015) proposes a metric for condition-based maintenance decisions, applying a damage detection method to a composite plate with piezoelectric buzzers, and introducing the SVM-GPDT as an original approach to multi-class probabilistic damage classification.

Beck and Huang (2017) include the effective use of a Bayesian approach to quantify modeling uncertainty, the advantage of Sparse Bayesian learning in incorporating sparseness during model updating, and the ability of the Bayesian probabilistic framework to quantify the relative plausibility of different model parameters.

Yuhang et al. (2017) proposes a dynamic data-driven hierarchical Bayesian degradation model to estimate the damage growth trend efficiently, validated through systematic case analyses. The methodology involves a dynamic data-driven hierarchical Bayesian degradation model that integrates a physical finite element model and a data-driven Bayesian framework to predict structural damage growth.

Beck and Huang (2017) focuses on using a Bayesian learning method for structural damage assessment of a benchmark problem, aiming to provide effective regularization and probabilistic prediction for reliable decision-making. The methodology used in the study is the Bayesian learning method, specifically employing Bayesian regression with automatic relevance determination prior for structural damage assessment. The method aims to achieve effective regularization and probabilistic prediction, expanded to provide vector outputs for monitoring.

Fang et al. (2018) found that Bayesian inference is beneficial for damage identification, resulting in a stable Markov chain with qualified parameter samples, and the method was validated using numerical and experimental beams. The methodology involved using approximate Bayesian computation (ABC) with the Metropolis Hastings sampling (MHS) algorithm and stochastic response surface (SRS), obtaining a stable Markov chain, and implementing a nested Bayesian updating procedure for fast estimation convergence.

Bull et al. (2019) found that active learning improves online classification performance for damage detection and classification. Active learning through uncertainty sampling leads to a clear increase in diagnostic performance of the probabilistic classifier. Active

learning reduces the variability of classification performance. The methodology involves an overview of partially supervised learning in SHM, defining a probabilistic model and guided sampling framework, integrating the model into an online SHM strategy, and demonstrating improvements through active learning on three datasets.

Chen et al. (2020*a*) proposed hierarchical Bayesian learning methodology with sensitivity analysis is robust for structural damage identification even in challenging conditions such as high measurement noise and limited sensor recordings. The methodology involves classical Bayesian learning and sparse Bayesian learning. Incomplete modal quantities are used for classical Bayesian learning to update a baseline model, followed by sparse Bayesian learning for accurate damage sparsity identification. Sensitivity analysis of model parameters is included to compensate for errors and optimize results.

Thus, machine learning (ML) frameworks have gained significant traction in structural health monitoring (SHM) due to their ability to handle complex, high-dimensional data, detect nonlinear relationships, and adapt over time for improved diagnostic accuracy. Unlike traditional statistical approaches, ML enables proactive maintenance strategies, anomaly detection, and seamless integration with digital technologies like IoT, making it a robust tool for SHM. Several studies have explored ML techniques for damage detection. Data-driven approaches have demonstrated effectiveness in structural damage identification through statistical modeling. Artificial immune systems have been employed to inspect structures under different actuation phases, successfully identifying damages in experimental setups. Support vector machines (SVMs) and probabilistic classification techniques, such as the SVM-GPDT framework, have enhanced condition-based maintenance decision-making. Bayesian methods have played a crucial role in quantifying modeling uncertainties and regularizing damage assessment models.Hierarchical Bayesian frameworks have been used to track damage progression efficiently by integrating finite element models with data-driven learning techniques. Approximate Bayesian computation (ABC) combined with Metropolis-Hastings sampling has improved convergence in parameter estimation for damage identification. Additionally, hierarchical Bayesian learning approaches incorporating sensitivity analysis have proven robust even under high noise and sparse sensor data. Furthermore, deep learning models such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have exhibited remarkable potential in automated feature extraction and time-dependent anomaly detection. Hybrid approaches integrating physics-based models with ML algorithms are being explored to enhance interpretability and reliability in damage prognosis. These advancements highlight the growing synergy between ML and SHM, complementing traditional statistical methods and offering enhanced predictive capabilities for infrastructure health management.

2.4 Prognosis methodologies in engineering systems

While damage detection in infrastructure is well-studied, damage prognosis remains a less explored area. Common prognosis focuses on fatigue cracking in steel, rebar corrosion in reinforced concrete, and age-related stiffness loss. Prognosis predicts a structure's future performance based on its current condition, often by projecting damage trends to a failure threshold. Results are expressed as remaining useful life (RUL), health index, or probability of future failure. Notable research includes work on fatigue crack growth in riveted joints (Singh et al., 1994), system identification for overall structural health (Abdullah and Haldar, 2017), and expectation-maximization algorithms for predicting damage due to natural hazards (Wang et al., 2016). Extreme value theory (EVT) shows promise for applications like bridge girder cover plate degradation and steel rebar corrosion (Li and Zhang, 2014). Other notable frameworks include sensor-driven prognosis for maintenance scheduling (Li et al., 2013). Reliability-based frameworks have found success in diverse applications such as nuclear piping (Pandey et al., 2009; Yuan et al., 2008), oil and gas pipelines (Pandey, 1998), highway bridges (Biondini and Frangopol, 2016*b*; Frangopol et al., 2001), and corrosion-affected concrete beams (Enright and Frangopol, 1998). These studies underscore the importance of real-time monitoring to refine remaining useful life (RUL) predictions. RUL estimation involves modelling the degradation process and projecting it until a predefined failure criterion is reached. This RUL distribution then informs the maintenance model to optimize replacement intervals. The details of the degradation modelling and literature survey of prognosis are presented in next sections.

2.4.1 Degradation modelling

Most engineered system failures are attributed to a gradual and irreversible accumulation of damage during operation. This degradation can be tracked through both directly like measuring crack length and indirectly through observing vibration, acoustic, magnetic, and thermal signals. These collected signals, known as degradation signals, form the basis for mathematical modelling of degradation processes and the estimation of RUL based on established failure thresholds (Nelson, 2009). Direct measurement-based degradation modelling has been applied to various scenarios including crack growth in metals (Lu and Meeker, 1993), fatigue, nuclear piping (Pandey et al., 2009), oil and gas pipelines (Pandey, 1998), and flow-accelerated corrosion (Yuan et al., 2008). In cases where direct measurement isn't possible, surrogate measures from vibrations (Ali et al., 2015; Nikulin et al., 2010; Zhang et al., 2019), strains (Mata, 2011; Prakash and Narasimhan, 2018), and electrical currents (Vanli, 2014) are used, though this approach is still nascent in the fields of civil and mechanical infrastructure.

Failures can stem from internal factors like aging and material quality or external ones

such as environmental conditions including vibrations and humidity. These failures are typically categorized into two types: predictable gradual failures, which are monitored through condition indicators, and unpredictable sudden failures, which occur without prior indication (Moubray, 2001). Degradation, the performance and reliability decline in assets, is depicted in models as occurring until a critical failure threshold is reached. Degradation can be either natural, driven by aging, or forced by external loads increasing over time (Wireman, 2004). Prognostic approaches, which predict the future reliability of systems, vary widely (Ma, 2007). These include physics-based, knowledge-based, data-driven and hybrid approaches.

1. **Model-based** prognostics employ dynamic mathematical models, either physics-based like crack growth models for critical systems (Kulkarni and Achenbach, 2008) or statistical models such as Kalman/particle filtering which use Bayesian estimation to predict fault evolution (Chelidze and Cusumano, 2004; Vachtsevanos et al., 2006).

2. **Knowledge-based** approaches address problems typically resolved by human experts and are increasingly used in fault prognostics without requiring explicit models, often incorporating expert systems and fuzzy logic to handle real-world uncertainties (Jardine, 2002; Wang and Zhang, 2008).

3. **Data-driven** approaches leverage statistical and machine learning techniques from pattern recognition to predict system behaviour. These methods range from multivariate statistical methods to more complex models like Neural Networks (NNs) and Hidden Markov Models (HMMs), which are prevalent in advanced prognostics (Jardine, 2002).

4. **Hybrid approaches** use combination physics based approaches with data-driven based methods to correctly estimate the degradation where modelling through individual approaches fails.

These approaches are pivotal in dealing with the complex dynamics of system health and reliability, providing a robust framework for the ongoing assessment and management of engineered systems. The classification of prognosis approaches is provided in Figure 2.2

2.4.2 Physics based prognosis methods

Physics-based prognosis in structural engineering relies on mechanistic models capturing degradation processes like fracture, creep, and corrosion. These models leverage physical knowledge to predict system behavior under new conditions. Despite challenges in representing complex and probabilistic real-world systems, physics-based models offer valuable insights, as exemplified by the Paris crack growth law. However, computational intensity and rigorous validation are essential for their practical application. These models

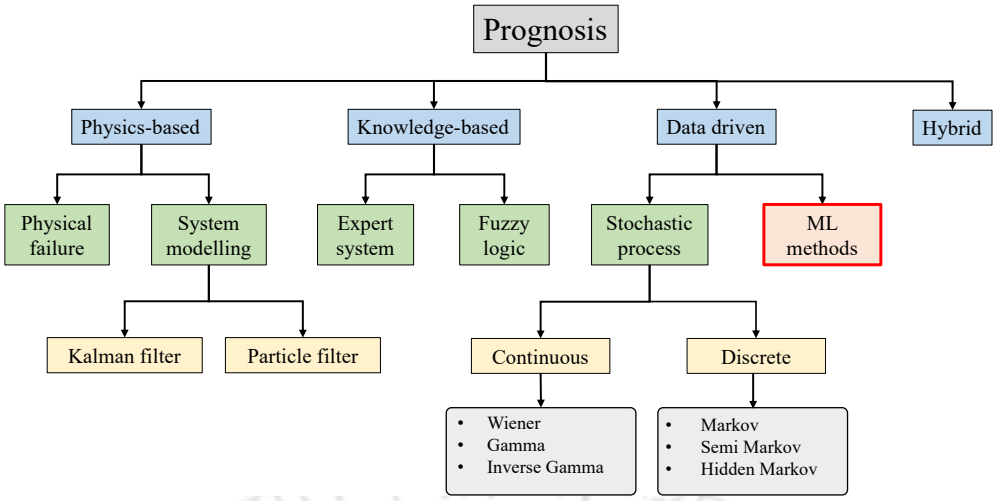


Figure 2.2: Illustration of classification of prognosis methodologies for structural health monitoring

excel in predicting new scenarios based on physical understanding, outperforming data-driven approaches. Continuous refinement improves accuracy and addresses performance issues.

Huston (1994) emphasized the complexities in predicting fatigue life in composite materials due to various failure mechanisms. Huston's experimental approach involved tension fatigue tests on carbon fiber reinforced epoxy, where the correlation of residual strength and stiffness models with cumulative damage predictions provided a methodological groundwork for future studies.

Luo et al. (2003) propose a comprehensive framework for degradation modeling and prognosis using physics-based approaches, which illustrates a procedural approach to prognostics. Their framework includes modeling system degradation with a set of variables and a degradation function, evolving system dynamics through Monte Carlo simulations, and tracking degradation with an interacting model to estimate the remaining life based on current damage states and anticipated future usage. Transitioning from empirical to computational models, Lee et al. (1999) demonstrated the potential of machine learning in modeling constant-stress fatigue behavior in carbon-fiber composites. While their approach excelled in constant-stress conditions, they acknowledged the challenges and additional efforts required to extend these models to variable-stress conditions, thereby setting a direction for further research in this area. Further, exploiting computational methods, Roshan Rammohan and colleagues adopted hidden Markov models to predict the remaining service life of structures based on damage diagnosis observations. This approach highlighted the adaptability of stochastic models in capturing damage accumulation and propagation, an essential aspect of structural health monitoring (Rammohan and Taha, 2005).

The year 2002 marked significant contributions by Kamiński, who explored both deterministic and stochastic approaches in fatigue analysis of composite materials using the Stochastic Finite Element Method (SFEM). This method, along with homogenization and reliability tools, offered a sophisticated framework for addressing the random nature of fatigue processes (Kamiński, 2002). Similarly, Fish and Yu (2002) introduced a multiscale fatigue analysis model employing internal state variables and non-local damage theory to model fatigue damage evolution in brittle composite materials. Their approach not only addressed the multiscale damage effects but also improved the numerical integration methods used in fatigue predictions. The application of technology in fatigue monitoring saw a novel approach by Connor et al. (2010), who utilized wireless sensor networks along with rainflow counting methods, paving the way for more integrated and real-time monitoring systems. He et al. (2012) further contributed to this field by integrating dynamic analysis with fatigue damage prognosis in bridges, demonstrating a concurrent methodology that enhances the predictive maintenance strategies.

Chiachıo et al. (2013) developed a prognostic framework that balanced the release of strain energies from competing damage modes in composites, a crucial factor in predicting the damage evolution in fiber-reinforced composites under fatigue loads. This work was part of a broader effort to refine the prognostic frameworks and tailor them to specific material behaviors under stress. Lastly, Tang et al. (2016) applied evidence theory for uncertainty quantification in concrete fatigue life prediction, integrating empirical formulas and Paris law models to enhance the reliability of fatigue life predictions in concrete structures. Palanisamy et al. (2022) presented a technique for the real time diagnosis and prognosis through observing Lamb wave in adhesively bonded joints, presents in structural members. In reviewing deterioration modeling, Biondini and Frangopol (2016*a*) emphasize the impact of environmental factors such as corrosion, which in steel structures and concrete infrastructures affected by chloride-induced corrosion, can significantly alter the predicted lifespan and structural integrity. They highlight that the corrosion rate in reinforcing steel follows a diffusion process described by Fick's laws, suggesting that even in highly stochastic conditions, physics-based models can provide valuable predictive insights when appropriately tailored and applied.

2.4.3 Data driven based prognosis methods

Typically, the degradation signals collected from comparable systems under identical environmental and testing conditions exhibit considerable variation in their trajectories. This variability is often attributed to measurement errors, differences between units, and intrinsic deterioration mechanisms. Such discrepancies pose challenges for purely deterministic physics-based models, which may struggle to fully account for these uncertainties in their predictions of remaining useful life. In contrast, data-driven models

do not depend on physical theories but rather utilize statistical principles alongside observed data to develop models that can accommodate the variability observed in degradation paths. In subsequent sections, we will explore statistical approaches, hybrid methodologies, and the role of artificial intelligence in enhancing these predictive models.

Statistical and stochastic approaches

Liu and Mahadevan (2005) proposed damage accumulation model for fatigue life prediction aligns well with experimental observations and a methodology for probabilistic fatigue life prediction and reliability calculation is presented. The methodology includes a new damage accumulation model for fatigue life prediction, based on the ply level and using a multiaxial damage index, as well as predicting fatigue life and comparing results with experimental data.

Mohanty (2010) presented thesis that demonstrates improved SHM and probabilistic modeling for fatigue damage estimation in structures, integrating SHM and adaptive prognosis models for aerospace structural components. Baraldi et al. (2010) highlighted how uncertainties from randomness and incomplete knowledge affect prognostic results, including a case study on stochastic fatigue crack growth model with uncertainties.

Sankararaman et al. (2011) detailed a methodology that includes uncertainty quantification in fatigue crack growth prognosis for complex structures, incorporating finite element analysis, multi-axial loading, and Gaussian processes to identify critical uncertainties.

Gobbato et al. (2012) presents a probabilistic methodology for predicting the remaining service life of adhesively bonded joints in composite aircraft structures, including assessing damage, updating joint probability distribution functions, and computing reliability indices for the composite wing structure. The methodology involves the use of non-destructive evaluation techniques, recursive Bayesian inference, probabilistic models for aerodynamic loads, and damage evolution to predict the remaining service life of adhesively bonded joints in composite aircraft structures.

Daigle and Goebel (2012) further enhanced model-based prognostics by integrating concurrent damage progression processes, utilizing a particle filter for state-parameter estimation and developing a variance control algorithm for uncertainty management. J. Chiachío et al. (2014) introduced a model-based prognostics framework incorporating Bayesian particle filtering to predict fatigue damage evolution and reliability in composites.

Dimitrov et al. (2015) developed a probabilistic model for estimating the fatigue life of laminated composite plates based on lamina-level input data, using Bayesian inference to estimate model parameters and allowing for the prediction of fatigue properties for a wide range of laminate configurations. The paper detailed the modeling techniques, parameter estimation procedure, and provided an illustrative application. The methodology involves

developing a probabilistic model using lamina-level input data, with model parameters estimated through Bayesian inference. The study describes the modeling techniques and parameter estimation procedure.

Peng et al. (2015) introduces a stiffness degradation model for composite fatigue life prediction, conducts in-situ fatigue testing with piezoelectric sensors, and validates the prognosis framework using experimental data. The methodology involves combining Bayesian inference, piezoelectric sensor measurements, a mechanical stiffness degradation model, growth rate-based stiffness degradation model, Bayesian updating-based fatigue life prediction method, in-situ composites fatigue testing with piezoelectric sensors, signal processing techniques for damage diagnosis features extraction, and integration of stiffness degradation into the Bayesian inference framework for remaining useful life prediction.

Chiachío et al. (2015) include the importance of continuous health assessment using SHM technology, successful demonstration of the prognostics approach using SHM data, and the effectiveness in predicting the remaining useful life with uncertainty management. The methodology involves a reliability-based prediction framework for estimating the remaining useful life of composite materials under fatigue degradation, utilizing Bayesian filtering to estimate degradation phenomena and incorporating information from multi-scale damage models and measurements. The estimation of remaining useful life is derived as a probability from time-dependent reliability predictions.

Eleftheroglou and Loutas (2016) emphasizes the importance of diagnostic and prognostic tools in health monitoring, presents a procedure for estimating optimal model parameters, and allows for diagnostics and prognostics using a stochastic model. The methodology involved modeling damage evolution in composites as a doubly stochastic hidden Markov process based on acoustic emission data from fatigue testing, with a procedure for estimating optimal model parameters and applying them to unseen data through cross-validation.

Eleftheroglou et al. (2016) utilized a Non-Homogenous Hidden Semi Markov model to describe the composite material's damage evolution under fatigue loading. The effectiveness of the prognosis based on the model's estimations compared to actual remaining useful life (RUL) was discussed. The methodology involves utilizing a Non-Homogenous Hidden Semi Markov model with SHM data from strain measurements obtained through the DIC technique for fatigue testing.

Francesco et al. (2020) found that Bayesian models, particularly multi-level models, outperformed existing models in BS 7910, emphasizing the need for sophisticated models to address multi-variate uncertainty and dependencies in fatigue crack growth assessments. The methodology involves establishing Bayesian models, demonstrating Bayesian predictive modeling, and using Hamiltonian Monte Carlo with Markov Chains

for model fitting.

Pugalenthi et al. (2021) developed a particle-filter-based online prognostic framework for predicting crack evolution in composite laminates due to fatigue. This study optimized a bimodal model for damage prediction in Glass-Fiber Reinforced Polymer (GFRP) laminates, enhancing safety by accurately forecasting damage progression and estimating remaining useful life.

Apart from the above discussed literature, a brief summary of the methodologies in the past for the prognosis using statistical approaches are provided in Table 2.1

Table 2.1: Statistical approaches for the prognosis of the engineering systems

Methodology	Application	References
Gamma process	Two-phase degradation	Prakash and Narasimhan (2018)
Model updating	Crack in beam girder	Zárate et al. (2012)
Multi-model	Stiffness degradation in GFRP	Banerjee et al. (2019)
Anomaly detection	Rolled element bearing	Qi et al. (2024)
Smooth particle filter	RUL of lithium-ion batteries	Greer and Bashir (2023)
Bayesian and EM	Path dependent degradation	Si et al. (2013)
Gaussian process	Fatigue growth prognosis	Ling and Mahadevan (2012)
NSOGRO model	Prognosis through guided wave	Chen et al. (2019)
Weiner process	Fatigue crack width estimation	Yan et al. (2021)
Particle filter	Fatigue of aluminium structures	Li et al. (2022)
Stochastic correlation	RUL of a complex system	Wu et al. (2023)
MCDAN	RUL of bearings	Dong et al. (2024)
AE and DIC	RUL of composite structures	Eleftheroglou et al. (2018)
Mixture proposal PF	Hole edge cracked structure	Chen et al. (2021)
FIDES	Service life of bearings	Chen et al. (2022)
Wiener process	Degradation process modelling	Gao et al. (2020)
WPD and MoG-HMM	Prognostics of bearing	Tobon-Mejia et al. (2012)
Particle filter	CFRP cantilever beams	Cristiani et al. (2021)
Particle filter	Crack monitoring	Chen et al. (2020*b*)
Kalman filtering	Fatigue prediction in truss	Papadimitriou et al. (2011)

Hybrid approaches

Liu and Mahadevan (2010) discusses probabilistic models for fatigue behavior, reviews existing methods, and demonstrates a probabilistic life prediction method under variable amplitude loading. The methodology involves using probabilistic models for fatigue life prediction, reviewing existing methods, quantifying uncertainties in loading and material properties, utilizing the Karhunen-Loeve expansion technique, and applying probabilistic methods for solving reliability problems.

Xiang and Liu (2010) discussed the use of probabilistic fatigue crack growth prediction methodology for real-time damage prognosis, highlighting the transformation of random loading to an equivalent constant amplitude loading process for efficiency.

Wang et al. (2012) aims to develop a generic framework for structural health prognostics applicable to various engineered systems, consisting of four core elements, and its effectiveness is demonstrated through two case studies. The methodology used in the study involves a generic framework for structural health prognostics with four core elements: a generic health index system, a sparse Bayes learning technique for offline learning, a similarity-based interpolation for online prediction, and an uncertainty propagation map for uncertainty management.

Guan et al. (2013) developed a maximum entropy-based framework for probabilistic fatigue damage prognosis that outperforms classical Bayesian algorithms in life prediction confidence bounds.

Li et al. (2013) utilizes sensor data for updating a degradation model, employs Bayesian theorem and MCMC for parameter calculations, and uses Bayesian updating for improved prognosis and RUL estimation. The methodology involves using continuously collected sensor data to update parameters of a stochastic structural degradation model, and employing Bayesian theorem and Markov Chain Monte Carlo (MCMC) sampling to calculate posterior distributions of stochastic parameters.

Shiri et al. (2015) found that the modified model improved the accuracy of fatigue life prediction for composites under two-stage loading, while fatigue life dispersion reduced the reliability of predicted remaining fatigue cycles. The methodology involved assessing a stiffness-based model for fatigue damage simulation, developing a modified model, and conducting an uncertainty analysis using the Monte Carlo approach to evaluate reliability.

Chen and Liu (2021) introduced the probabilistic physics-guided neural network (PPgNN) for estimating probabilistic fatigue S-N curves with physics guidance. This innovative approach enhances the accuracy and consistency of fatigue data analysis by encoding known physics constraints into machine learning models, providing a flexible and robust framework for uncertainty quantification.

Apart from the above discussed literature, a brief summary of the methodologies in the past for the prognosis using statistical approaches are provided in Table 2.2

Table 2.2: Hybrid approaches for the prognosis of the engineering systems

Methodology	Application	References
UKF and LS-SVM	RUL of Lithium-ion batteries	Li et al. (2021)
GRU plus HMM	Battery health prediction	Lin et al. (2023)
ConvDenseNet and EM	Rolled bearing	Ding et al. (2023)
Physics plus ML	Fatigue crack propagation	Neerukatti et al. (2018)
ARMA plus Markov model	Fatigue damage prognosis	Ling et al. (2011)

Artificial intelligence based approaches

Vassilopoulos et al. (2007) presented that artificial neural networks are effective in modeling fatigue life of composite materials under constant loading patterns, reducing the amount of experimental data needed for analysis by 40-50% and potentially decreasing costs and time for establishing S-N curves. The methodology involved utilizing artificial neural networks to model fatigue life of multidirectional composite laminates under different loading patterns, validating the model's accuracy with experimental data from two material systems.

Jiang and Mahadevan (2008) found that deep learning, specifically long short-term memory (LSTM), is effective in predicting the remaining useful life of manufacturing machines. LSTM outperforms traditional machine learning algorithms in revealing degradation patterns. The proposed LSTM-based approach is verified to be effective using NASA's C-MAPSS dataset. The methodology involves proposing an approach based on the Long Short-Term Memory (LSTM) network for tracking system degradation and predicting the Remaining Useful Life (RUL), with evaluation using NASA's C-MAPSS dataset.

Mao et al. (2018) demonstrated a significant improvement in predictive accuracy and numerical stability with the proposed method when applied to bearing datasets. The methodology involves proposing a new criterion for state division, extracting deep features using a convolutional neural network and Hilbert–Huang transform, and utilizing a long short-term memory neural network for prediction modeling.

Mohamed (2018) explored the application of supervised machine learning models (ANN, GPR, SVMR) to predict the remaining useful life of components in industrial settings, noting that neural networks excel in RUL estimation for slow-speed bearings.

Li et al. (2019) emphasizes the importance of accurate evaluation of machine degradation, introduces a deep learning-based method for RUL prediction of bearings, and demonstrates its effectiveness and superiority through experiments. The methodology involves a novel intelligent RUL prediction method based on deep learning, exploring time-frequency domain information for prognostics, and implementing multi-scale feature extraction using convolutional neural networks.

Kraus and Feuerriegel (2019) developed a structured-effect neural network for predicting remaining useful life, estimating parameters using variational Bayesian inferences, and comparing the model's performance with traditional machine learning and recurrent neural networks.

Elasha et al. (2019) explored the use of machine learning to predict the Remaining Useful Life (RUL) of wind turbine gearbox bearings. Their study focused on analyzing vibration measurements to accurately estimate the remaining life and define bearing failure stages.

Maher and Danouj (2020) include the importance of deep learning in establishing optimal predictive maintenance policies, the significant improvements in deep learning made by researchers, the impact of deep learning on upgrading companies in the Industry 4.0 revolution, the effectiveness of deep learning in predicting the residual useful life of machines, the flexibility and continuous improvement opportunities provided by deep learning, the benefits of transfer learning in improving neural network performance, and the potential role of cloud and fog computing in scaling up deep learning for predictive maintenance. The methodology involves a review of various deep learning methods used for predictive maintenance, including RBM, DBN ensemble, CNN-based regressor, semi supervised learning, RNN, LSTM, transfer learning, unsupervised techniques, and domain adaptation. These methods are applied to estimate the Remaining Useful Life (RUL) of machines and components.

Elenchezhian et al. (2021) found that Artificial intelligence is widely used in various aspects of composite materials. The review analyzes the gaps and uncertainties in the development of AI models for structural health monitoring and prognostics of polymer matrix composites. The paper discusses the pipeline for real-time prognostics, hybrid approaches, uncertainty quantification, and reliability standards for complex composite materials. The methodology involves a comprehensive review of AI models developed for SHM and prognostics health management of polymer matrix composites, along with discussing real-time prognostics pipeline, hybrid approaches, uncertainty quantification, and reliability standards.

Zhang et al. (2021) found that deep learning model exhibited superior prediction accuracy and generalization ability compared to conventional machine learning models. The methodology involved proposing a general life prediction method, integrating data, and applying conventional and deep learning models for life predictions. Wen et al. (2021) found that accurate RUL prediction can reduce downtime maintenance and improve economic benefits. A stochastic degradation model is used to derive reliability characteristics, and a nonlinear data fusion method based on Genetic Programming is proposed to construct a superior composite Health Indicator. The methodology involves using a stochastic degradation model, developing desirable properties of the Health Indicator, and employing a nonlinear data fusion method based on Genetic Programming to predict Remaining Useful Life for complex systems with multiple CM signals.

2.5 Advanced machine learning based prognosis

Generally, numerous reviews in the literature (Jardine et al., 2006; Sikorska et al., 2011; An et al., 2015; Lei et al., 2018) outline the application of AI in predictive health management (PHM) as encompassing three key phases: (i) data collection and processing; (ii) algorithm development, training, and validation; and (iii) prediction of Remaining Useful Life (RUL)

and making maintenance decisions. Lim (2014) presented a thesis on Lamb wave approach for fatigue damage prognosis, developing a model for predicting fatigue crack growth. Kim et al. (2017) describe the fundamentals of the prognosis and various physics-based and data driven methodologies for the health monitoring systems to predict damage or degradation behavior in prognostics. Manuel (2013) monitored structural performance to update damage models leads to more accurate fatigue damage predictions that consider uncertainty. The methodology involves updating a damage model based on monitored structural responses to make accurate predictions while considering uncertainty.

Table 2.3: Advanced machine learning techiques based on SVM, GPR, and BNN for the prognosis of the engineering systems

	Methodology	**Application**	**References**
SVM	Recurrent LSSVR	RUL of bearing	Ling et al. (2011)
	Reliability assessment	Dumbbell-shaped cylinder	Liu et al. (2022)
	Bayesian OC-SVM	Bearing fault prognosis	Fong and Narasimhan (2021)
GPR	ML assisted GPR	Fatigue in turbine discs	Hu et al. (2020)
	Kernel-based GPR	RC parameter estimation	Rayjada et al. (2023)
	Acoustic emissions	Aeronautical structures	Galanopoulos et al. (2023)
	Simple GPR	Metallic materials	Gao, Wang, Xu, Wang, Yan and Wang (2022)
	BPNN-GPR	Multi axial fatigue	Gao et al. (2023)
	Bayesian plus GPR	Creep in nuclear structures	Baraldi et al. (2015)
	KL divergence plus GPR	RUL of bearing	Kumaraswamidhas and Laha (2021)
	Deep GPR	Battery health management	Tagade et al. (2020)
	GPR	Fault grid in AC microgrid	Srivastava and Parida (2022)
	GPR	Aeronautical structures	Aye and Heyns (2017)
	PI Gaussian process	Crack in plates	Jones et al. (2023)
	Adaptive ensemble GPR	Bearing fault	Hou and Peng (2023)
	Deep GPR	Battery prognosis	Tagade et al. (2020)
	LSTM-GPR and EKF-GPR	PHM of fuel cell	Wang et al. (2022)
	GPR	Lithium-ion batteries	Liu and Chen (2019)
BNN	Bayesian transfer learning	Fault prognosis	Zhu et al. (2023)
	Multi-variate PCA	Prognosis	Fallahdizcheh and Wang (2022)
	Bayesian DL	Bearing degradation	Zhu et al. (2022)
	BLKAN	Remainig useful life	Wang et al. (2023)
	PI-Bayesian RNN	PHM in aerospace industry	Fernández Salas et al. (2024)
	Bayesian DL	Complex systems	Kim and Liu (2021)
	Bayesian DL	Unceratinty quantifications	Li et al. (2020)
	Variational BNN	Miter gates	Vega and Todd (2022)

Further, a dynamic data-driven hierarchical Bayesian degradation model is proposed to estimate the damage growth trend efficiently, validated through systematic case analyzes (Yuhang et al., 2017). The methodology involves a dynamic data-driven hierarchical Bayesian degradation model that integrates a physical finite element model and a

data-driven Bayesian framework to predict structural damage growth. Systematic case analyzes are performed to validate the method. Wang and Rocková (2020) explored sparse deep learning uncertainty quantification using Bayesian non-parametrics, applying Bernstein-von Mises theorems for valid Bayesian credible regions. Martti and John (n.d.) introduced data analysis methods for aging components in nuclear power plants, demonstrating prognosis models that utilize machine learning to detect trends and clusters in component data. Hristos and Georgia (2022) provided an extensive review of probabilistic forecasting and prediction techniques in machine learning. This review covers concepts, methods, and metrics for assessing probabilistic predictions, enhancing predictive uncertainty estimation and discussing challenges in the research field. Lastly, a detailed summary of prognosis works through advanced machine learning techniques are provided in Table 2.3.

2.6 Gaps in the literature

For the damage detection and prognosis of highly reliable systems, a large amount of research, as discussed in the previous sections, has been published in recent years, which broadly focuses on damage detection, degradation modeling, change point detection and RUL estimation. However, several important issues such as an integrated framework for damage detection and prognosis, use of probabilistic ML framework, remain at best only partially addressed in the literature. Further, literatures lack integration of ML with probability for damage detection in civil engineering domain. A summary of the gap areas in the existing research on damage detection and prognosis is described below, which forms the motivation for current thesis.

1. **Limited diagnosis of complex system in civil engineering:** The system identification in the literature, especially related to civil engineering is focussed on simpler system and if the complexity arises, the research shifts to the finite element model updating of the diagnosis. However, it is tedious and need high computational requirements. Thus, there should be system identification frameworks for the diagnosis of complex systems.

2. **Limited integration of probabilistic damage detection and prognosis:** The literature review showed that most of the machine learning work which is focused on damage detection is deterministic in nature and doesn't account the uncertainties present in engineering systems. Research using condition information for damage prognosis of civil and mechanical infrastructure is lacking and there exist only few studies where damage detection results are extended to RUL estimation. Thus, the integrated framework which undertakes probabilistic diagnosis and prognosis systematically is still lacking in the literature.

3. **Validation of models in real-world applications:** Many studies provide models that are either not fully integrated into a unified SHM system or lack comprehensive validation across a range of real-world scenarios. Further, the studies using machine learning techniques finds application primarily related to the bearing and battery degradation. Few researches exist using ML techniques on the fatigue prognosis of joints and a very little for civil engineering dynamical systems like isolation system. Thus, there is a need to develop unified probabilistic SHM framework which is extensively validated through numerical case studies, controlled laboratory experiments, and practical applications, especially related to fatigue degradation in composites and degrading complex system in civil engineering domain.

2.7 Specific objectives

The specific research objectives are:

1. **To develop an integrated probabilistic framework for prognosis:** This objective focuses on the development of comprehensive framework that integrates Gaussian Process Regression (GPR) with a robust solution to address extrapolation challenges, ensuring accurate and reliable predictions beyond the training data range. Adaptive mean functions will be incorporated to enhance the model's ability to capture evolving degradation patterns, while Bayesian change point detection will enable the framework to dynamically adapt to shifts in system behavior. This framework will be optimized using a global convergence method to mitigate the risk of premature optimization, ensuring the exploration of the entire parameter space for optimal performance.

2. **To develop and evaluate uncertainty-quantified diagnosis:** This objective encompasses the development of uncertainty aware degradation indicator through Bayesian inference based system identification and damage estimation. This objective focuses on the critical assessment and selection of appropriate diagnosis methodology depending on the case studies, particularly Bayesian filters and Bayesian Neural Networks are implemented to effectively quantify and propagate uncertainties in structural health indicators. The goal is to develop and extract a reliable degradation indicator for the prognosis for complex engineering applications.

3. **To validate the framework through real-world case studies:** This objective involves applying the proposed probabilistic framework to real-world scenarios, specifically the prognosis of elastomeric rubber bearings and fatigue crack monitoring in lap joints. The aim is to demonstrate the framework's practical applicability, effectiveness in predicting RUL, and ability to account for uncertainties in different engineering domains.

Further, the ultimate goal is to provide a holistic and accurate assessment of the remaining useful life (RUL) of deteriorating structures, explicitly accounting for various sources of uncertainty and facilitating real-time adaptation to changing conditions

Chapter 3

Uncertainty quantification in surrogate modelling for prognosis

3.1 Introduction

In the realm of Structural Health Monitoring (SHM), accurately and reliably predicting a structure's condition is paramount. Surrogate models, which serve as computationally efficient approximations of complex systems, have become invaluable tools in this domain. However, the inherent uncertainties in engineering systems, arising from material variability, operational conditions, and measurement errors necessitate a shift towards uncertainty-aware surrogate modeling. This chapter delves into the concept of uncertainty quantification (UQ) and its integration with surrogate models, with a particular focus on Gaussian Process Regression (GPR). We will discuss the various sources of uncertainty in engineering systems and how they can be accounted for within the surrogate modeling framework. Additionally, we will discuss the details of GPR, highlighting its strengths and weaknesses as a surrogate model for prognosis. By addressing the challenges associated with GPR, we aim to lay the groundwork for developing robust and reliable prognosis models that can capture and quantify the uncertainties inherent in real-world structural health monitoring applications.

3.2 Uncertainties in engineering systems

3.2.1 Sources of uncertainties

The treatment of uncertainty is essential in diagnosis and prognostics. The unpredictable nature of future loading and operating conditions further amplifies these uncertainties, impacting the reliability of damage indices, degradation indicators, estimated model parameters, and ultimately, the prediction of remaining useful life. A comprehensive understanding of these diverse uncertainty sources is essential for developing robust probabilistic machine-learning models. Two broad categories of uncertainty are recognized: aleatory and epistemic as shown in Figure 3.1. The details about the uncertainties can be found in the following literature. (Kiureghian and Ditlevsen, 2009;

Hüllermeier and Waegeman, 2021)

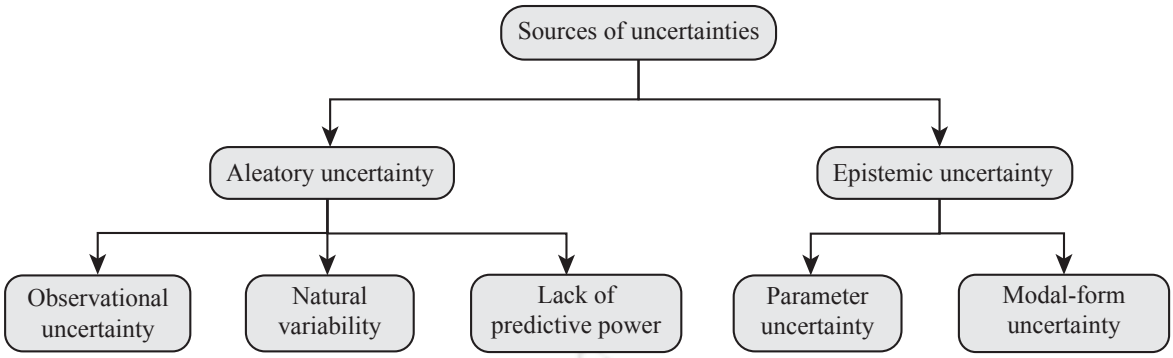


Figure 3.1: Illustration of different sources of uncertainties and their classification

Aleatory uncertainty arises from inherent randomness in the physical world or measurement processes. This includes physical uncertainty stemming from variability in material properties due to manufacturing or environmental factors, measurement uncertainty due to sensor limitations or noise, and process uncertainty resulting from the complex interactions within the system itself, or unpredictable external factors like random loads or disturbances.

Epistemic uncertainty, on the other hand, originates from limitations in our knowledge or modeling assumptions. This encompasses uncertainties due to simplified or incomplete models, gaps in our understanding of system behavior, and data sparsity. The choice of model structure significantly influences the degree of uncertainty, while limited, incomplete, or biased datasets can hinder a model's ability to generalize or accurately predict behavior outside its training scope.

Consider a model with a training dataset (X, Y), parameterized by a set of weights, $\Omega = [\omega_1, \omega_2, \omega_3, \ldots, \omega_n]$. Here, X and Y represent the independent and dependent variables, respectively. Epistemic uncertainty, which arises from the limitations of the accessible training dataset, manifests itself in the variability of the weights Ω and is reflected in the conditional probability $p(\Omega \mid X, Y)$. If $p(\Omega)$ is the prior of the weights, then the epistemic uncertainty can be quantified using Bayes' theorem as follows:

$$p(\Omega \mid X, Y) = \frac{p(Y \mid X, \Omega)p(\Omega)}{p(Y \mid X)} \tag{3.1}$$

where $p(\Omega)$ represents the prior knowledge about the unknown function weights. Through training, the posterior of the function weights $p(\Omega \mid X, Y)$ is estimated, which has a narrower distribution interval. Further, aleatoric uncertainty, which arises due to noisy measurements, directly affects the prediction outcome. In probabilistic terms, with a test input x and test output y, it can be expressed as $p(y \mid x, \Omega)$. Thus, both epistemic and aleatoric uncertainties

can be effectively incorporated into a single Bayesian framework as follows:

$$p(y \mid x, X, Y) = \int_{\Omega} p(y \mid x, \Omega) p(\Omega \mid X, Y)\, d\Omega \tag{3.2}$$

This equation estimates the predictive uncertainty in the output result y for any new sample x. It takes into account the historical data and associated uncertainties. A similar analogy can be understood using the law of total variance and variance accumulation theory, which predicts the variance of the output y as:

$$\text{Var}(y) = \underbrace{\text{Var}[\mathbb{E}(y \mid x)]}_{\text{Epistemic uncertainty}} + \underbrace{\mathbb{E}[\text{Var}(y \mid x)]}_{\text{Aleatoric uncertainty}} \tag{3.3}$$

This formulation helps in comprehensively capturing both types of uncertainties within a unified framework, facilitating more reliable and robust predictions.

3.2.2 Bayesian framework for uncertainty quantification

The probabilistic approach to modeling leverages probability theory to articulate all forms of uncertainty. This theory serves as the foundational mathematical language for representing and handling uncertainty, much like calculus is used to manage rates of change. Employing a probabilistic approach to modeling is fundamentally straightforward. It utilizes probability distributions to depict all uncertain, unobserved elements within a model, including structural, parametric, and noise-related uncertainties, and their connections to the data. Using basic principles of probability theory, this approach allows for the inference of unobserved quantities based on observed data. The learning process involves the transformation of prior probability distributions, defined prior to data acquisition, into posterior distributions, refined post-data observation. This process essentially updates our beliefs based on new evidence and is a key aspect of Bayesian learning. In probability theory, two fundamental rules are the sum rule and the product rule. The sum rule is expressed as $P(x) = \sum_{y \in Y} P(x, y)$, which illustrates how to compute the marginal probability of x by summing (or integrating, in the case of continuous variables) the joint probabilities over all possible values of y. Here, $P(x)$ represents the probability of observing x, and $P(x, y)$ denotes the joint probability of observing both x and y simultaneously. The product rule, on the other hand, states $P(x, y) = P(x)P(y \mid x)$, which describes the joint probability as the product of the marginal probability of x and the conditional probability of y given x. In this formulation, $P(y \mid x)$ is the probability of observing y conditioned on the value of x. From these two rules, Bayes' rule can be derived, providing a powerful method for updating the probability estimate for a hypothesis as more evidence or information becomes available.

Bayes' rule is formulated as:

$$P(y \mid x) = \frac{P(x \mid y)P(y)}{P(x)} = \frac{P(x \mid y)P(y)}{\sum_{y \in Y} P(x, y)} \tag{3.4}$$

This equation calculates the conditional probability of y given x by dividing the product of the conditional probability of x given y and the prior probability of y by the marginal probability of x. Bayes' rule thus updates our understanding of the likelihood of y based on new evidence.

In machine learning, we can adapt probability theory to specific contexts. Here, we replace x with D for observed data, y with θ for unknown parameters, and condition all terms on m, the class of probabilistic models. This leads to:

$$P(\theta \mid D, m) = \frac{P(D \mid \theta, m)P(\theta \mid m)}{P(D \mid m)} \tag{3.5}$$

In this equation, $P(D \mid \theta, m)$ is the likelihood, measuring the probability of the observed data given the parameters in model m, $P(\theta \mid m)$ is the prior probability of the parameters, reflecting assumptions before observing the data and $P(\theta \mid D, m)$ is the posterior probability after observing the data, representing our updated belief about the parameters. The training of machine learning model involves updating the prior belief about the parameters in light of new data, transforming it into a posterior belief. This posterior becomes the new prior for future analysis. To predict new, unseen data D_{test}:

$$P(D_{\text{test}} \mid D, m) = \int P(D_{\text{test}} \mid \theta, D, m)P(\theta \mid D, m)\, d\theta \tag{3.6}$$

This integrates over all possible parameter values, weighted by their posterior probability, to predict the likelihood of the new data. We can also compare different models by computing the posterior probability of a model m given the data D:

$$P(m \mid D) = \frac{P(D \mid m)P(m)}{P(D)} \tag{3.7}$$

$$P(D \mid m) = \int P(D \mid \theta, m)P(\theta \mid m)\, d\theta \tag{3.8}$$

Here, $P(D \mid m)$ is the marginal likelihood or model evidence, which integrates the likelihood of the data across all possible parameter values, weighted by their prior probabilities. This measure favors simpler models, penalizing those that fit the data only through complex or unlikely parameter configurations.

Bayesian learning applies these probability principles to the learning process from data. It

aids in making predictions and decisions under uncertainty, providing a rigorous means to quantify uncertainties and deepen our understanding of the model's behavior through probabilistic insights. This approach offers clarity on which parts of the model are uncertain and the reasons behind this uncertainty. Moreover, a critical challenge in machine learning (ML) for prognosis lies in the performance of models when faced with input samples from regions characterized by low signal-to-noise ratios or out-of-distribution (OOD) samples. These scenarios, where inputs deviate significantly from the training data distribution, often result in inaccurate predictions. This challenge is particularly relevant in Gaussian Process Regression (GPR) models, where extrapolation beyond the training data range can lead to overly optimistic uncertainty estimates, further worsen the issue. Accurate and appropriate uncertainty quantification in ML models, including GPR, is essential for robust decision-making. By providing a measure of confidence in their predictions, uncertainty-aware models can automatically flag high-uncertainty samples, indicating potential inaccuracies. Ideally, well-calibrated models should exhibit high uncertainty when their predictions are likely to be incorrect. This ability to pinpoint "difficult-to-predict" samples allows for targeted scrutiny, potentially involving domain experts, and ultimately leads to more informed and reliable decisions in critical applications like structural health monitoring.

This chapter aims to provide a comprehensive introduction to the surrogate methods for estimating the predictive uncertainty of data-driven ML models. Primarily, the data-driven ML models used in the thesis are Gaussian Process Regression (GPR) for prognosis and Bayesian Neural Networks (BNN) for feature mapping. The detailed background and challenges of the GPR are discussed in the subsequent sections. The methodology for addressing the issues with GPR related to extrapolation and optimization are proposed and discussed.

3.3 Gaussian process regression (GPR)

Gaussian Process Regression (GPR) offers a robust and flexible framework for supervised learning tasks, encompassing both regression and classification problems. While classification focuses on discrete class labels, regression deals with continuous output variables (Rasmussen and Williams, 2006). GPR models can be derived from various perspectives, notably the weight-space view, rooted in Bayesian principles, and the function-space view, which considers GPs as distributions over functions. In the function-space view, a Gaussian Process (GP) is defined as a collection of random variables, where any finite subset follows a Gaussian distribution. Mathematically, for any set $\mathcal{X}$, a GP on $\mathcal{X}$ is a set of random variables $f(x), x \in \mathcal{X}$ such that any finite collection of these variables $(f(x_1), ..., f(x_n))$ forms a multivariate Gaussian distribution. Given that a Gaussian distribution is fully characterized by its mean vector and covariance matrix, a GP

is similarly defined by a mean function μ and a covariance function $k\left(x, x'\right)$. Consequently, we can express a GP as $f\left(x\right) \sim GP\left(\mu\left(x\right), k\left(x, x'\right)\right)$:

$$\begin{aligned} \mu(x) &= \mathbb{E}\left[f\left(x\right)\right] \\ k\left(x, x'\right) &= \mathbb{E}\left[\left(f\left(x\right) - \mu\left(x\right)\right)\left(f\left(x'\right) - \mu\left(x'\right)\right)^{T}\right] \end{aligned} \tag{3.9}$$

Now we have a general regression model $y = f(\boldsymbol{x}) + \varepsilon$, where $f(\boldsymbol{x}) \sim \mathcal{GP}(\mu, k)$ and $\varepsilon \sim \mathcal{N}\left(0, \sigma_n^2\right)$. Given n pairs of observations $\left\{\left(\boldsymbol{x}_i, y_i\right)\right\}_{i=1}^{n}, \mathbf{x}_i \in \mathbb{R}^p, y_i \in \mathbb{R}$, it yields that $\left[f\left(\boldsymbol{x}_1\right), \ldots, f\left(\boldsymbol{x}_n\right)\right]$ follow a multivariate Gaussian distribution

$$\left[f\left(\boldsymbol{x}_1\right), f\left(\boldsymbol{x}_2\right), \ldots, f\left(\boldsymbol{x}_n\right)\right]^{T} \sim \mathcal{N}(\boldsymbol{\mu}, K) \tag{3.10}$$

where $\boldsymbol{\mu} = \left[\mu\left(\boldsymbol{x}_1\right), \ldots, \mu\left(\boldsymbol{x}_n\right)\right]^{T}$ is the mean vector and K is the $n \times n$ covariance matrix of which the (i, j)-th element $K_{ij} = k\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)$. In order to predict $f_* = f(X_*)$ at the test locations $X_* = \left[\boldsymbol{z}_1, \cdots, \boldsymbol{z}_m\right]^{T}$, the joint distribution of the training observations $\boldsymbol{y}$ and the predictive targets f_* are given by

$$\begin{bmatrix} \boldsymbol{y} \\ f_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}(X) \\ \boldsymbol{\mu}(X_*) \end{bmatrix}, \begin{bmatrix} K(X, X) + \sigma_n^2 \mathbf{I} & K(X_*, X)^{T} \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right) \tag{3.11}$$

where $\boldsymbol{\mu}(X) = \boldsymbol{\mu}, \boldsymbol{\mu}(X_*) = \left[\mu\left(\boldsymbol{z}_1\right), \ldots, \mu\left(\boldsymbol{z}_m\right)\right]^{T}, K(X, X) = K, K(X_*, X)$ is an $m \times n$ matrix of which the (i, j)-th element $\left[K(X_*, X)\right]_{ij} = k(\boldsymbol{z}_i, \boldsymbol{x}_j)$, and $K(X_*, X_*)$ is an $m \times m$ matrix with the (i, j)-th element $\left[K(X_*, X_*)\right]_{ij} = k(\boldsymbol{z}_i, \boldsymbol{z}_j)$. Thus, predictive distribution is

$$p\left(f_* \mid X, \boldsymbol{y}, X_*\right) = \mathcal{N}(\hat{\boldsymbol{\mu}}, \hat{\Sigma}) \tag{3.12}$$

$$\begin{aligned} \hat{\boldsymbol{\mu}} &= K(X_*, X)^{T}(K(X, X) + \sigma_n^2 \mathbf{I})^{-1}(\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{X})) + \boldsymbol{\mu}(\boldsymbol{X_*}) \\ \hat{\Sigma} &= K(X_*, X_*) - K(X_*, X)^{T}(K(X, X) + \sigma_n^2 \mathbf{I})^{-1} K(X_*, X) \end{aligned} \tag{3.13}$$

Taking noise part into consideration, the predictive distribution of targets y_* given the training set and the test locations are finally written by

$$p(y_* \mid X, y, X_*) = \mathcal{N}(\hat{\boldsymbol{\mu}}, \hat{\Sigma} + \sigma_n^2 \mathbf{I}) \tag{3.14}$$

3.3.1 Kernel function

The kernel $k(\cdot, \cdot)$ is pivotal in determining the predictive mean and variance in Gaussian Process Regression (GPR). Kernel contains our assumptions about the function we aim to learn and define the similarity between data points. Thus, the choice of kernel significantly influences the performance of a GPR model, much like how activation functions and learning

rates impact the outcomes of neural networks (Mercer, 1909). A symmetric $n \times n$ matrix C is considered positive semi-definite (PSD) if, for any non-zero column vector $\lambda \in \mathbb{R}^n$, the condition $\lambda^T C\lambda \geq 0$ holds. Before discussing several useful kernels, we introduce the concept of positive semi-definite kernels.

A positive semi-definite kernel (also called covariance function) on $\mathcal{X}$ is a function $k : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}, s.t. \forall n \in \mathbb{N}, \forall x_1, \ldots, x_n \in \mathcal{X}$, the matrix C is positive semi-definite, where $C_{ij} = k(x_i, x_j)$. For example, $\mathcal{X} = \mathbb{R}, k(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T\boldsymbol{y}$, hence $C = \boldsymbol{x}\boldsymbol{x}^T$. Letting $\boldsymbol{a} \in \mathbb{R}$, then $\boldsymbol{a}^T\boldsymbol{C}\boldsymbol{a} = \boldsymbol{a}^T\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{a} = \left(\boldsymbol{a}\boldsymbol{x}^T\right)^2 \geq 0$. Therefore, the bivariate function $k(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T\boldsymbol{y}$ is a PSD kernel. A general method of reproducing a proper kernel is introduced in the following and more details can be found in (Mercer, 1909; Minh et al., 2006). Moreover, a function $k : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ can be written as $k(x, y) = \langle \Phi(x), \Phi(y) \rangle$, where $\Phi(x)$ is a feature map:$x \mapsto \Phi(x) \in \mathbb{H}$ (Hilbert space) and $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{H}$, if and only if $k(x, y)$ is PSD kernel. Some commonly-used kernels are listed as follows.

1. **Squared exponential (SE):** The most widely-used kernel in GPR is Squared Exponential (SE), which is defined as

$$k_{SE}(x, x') = s_f^2 \exp\left(-\frac{\| x - x' \|^2}{2\ell^2}\right) \tag{3.15}$$

   where $\| \cdot \|$ is $L^2 - norm$ (Euclidean norm), s_f^2 is the signal variance and is also considered as an output-scale amplitude and the parameter ℓ is the input (length or time) scale (Roberts et al., 2013).
2. **Linear (Lin):** Another widely used kernel is linear kernel which is defined as

$$k_L(x, x') = s_f^2(x - c)(x' - c) \tag{3.16}$$

   where s_f^2 is the signal variance and is also considered as an output-scale amplitude and the parameter c is the bias.
3. **Periodic (PER):** It is used to model functions which exhibit a periodic pattern and is written as follows:

$$k_PER(x, x') = k_SE(\varpi(x), \varpi(x')) = s_f^2 \exp(-(2sin^2(\pi\ ((x - x'))/p))/\ell^2) \tag{3.17}$$

   where $\varpi(x) = [\sin(\pi x/p), \cos(\pi x/p)]T$ and p is the period.
4. **Rational quadratic (RQ):** The rational quadratic (RQ) kernel is given by

$$k_{RQ}(x, x') = s_f^2(1 + \frac{(x - x')^2}{(\alpha\ell^2)})^(- \alpha), \tag{3.18}$$

   where α is known as the index. Rasmussen and Williams (2006) show that this is equivalent to a scale mixture of SE kernels with different length scales, the latter

distributed according to a Beta distribution with parameters α and ℓ^{-2}. This gives variations with a range of time scales, the distribution peaking around ℓ but extending to significantly longer period (but remaining rather smooth). When $\alpha \to \infty$, the RQ kernel reduces to the SE kernel with length scale ℓ.

5. **Matern functions:** The Matérn class of covariance functions is defined by

$$k_M(x, x) = s_f^2 \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(2\sqrt{\nu}\frac{|\, x - x' \,|}{\lambda}\right) \mathbb{B}_\nu \left(2\sqrt{\nu}\frac{|\, x - x' \,|}{\lambda}\right) \tag{3.19}$$

where s_f is the output scale, λ is the input scale, $\Gamma(\cdot)$ is the standard Gamma function and $\mathbb{B}(\cdot)$ is the modified Bessel function of second order. The additional hyperparameter ν controls the degree of differentiability of the resultant functions modelled by a GP with a Matern covariance function, such that they are only $(\nu + 1/2)$ times differentiable. As $\nu \to \infty$, so the functions become infinitely differentiable and the Matérn kernel becomes the SE one. Taking $\nu = 1/2$ gives the exponential kernel

$$k(x, x') = s_f^2 \exp\left(-\frac{|\, x - x' \,|}{\lambda}\right) \tag{3.20}$$

Stationary and Non-stationary: The Squared Exponential (SE) and Periodic kernels are stationary, meaning their value depends only on the difference $x - x'$. This implies that the probability of observing a particular dataset remains constant even if all the x values are shifted by the same amount. In contrast, the Linear kernel is non-stationary, meaning that the corresponding GP model will yield different predictions if the data is shifted while keeping the kernel parameters fixed.

Combining kernels

For various types of structures, it is feasible to construct a custom kernel tailored to specific desired properties. Hence, the importance of combining kernel functions arises. There are primarily two ways of combining kernels: addition and multiplication

$$\begin{aligned} k_a + k_b &= k_a(x, x') + k_b(x, x') \\ k_a \times k_b &= k_a(x, x') \times k_b(x, x') \end{aligned} \tag{3.21}$$

Multiplying two positive-definite kernels together always results in another positive- definite kernel. Working with kernels, rather than the parametric form of the function itself, allows us to express high-level properties of functions that do not necessarily have a simple parametric form. The resultant kernel properties is dependent on its parent kernels as follows:

1. **Polynomial regression.** Multiplying (T) linear kernels produces a prior on polynomials of degree (T)
2. **Functions with increasing mean amplitude.** Multiplication by a linear kernel results

in the marginal standard deviation of the modeled function growing linearly away from the location specified by the kernel parameter.

3. **Locally periodic functions.** For univariate data, multiplying a periodic kernel by the Squared Exponential (SE) kernel transforms global trend into local trend.

In this particular thesis, we primary work with the linear and square exponential kernel function. Further, for the prognosis or forecasting the time series, the extrapolation is needed. The product-kernel model is more flexible, and so remains uncertain about the function away from the data while additive kernel structure sometimes allows us to make predictions far from the training data.

3.3.2 Mean or basis function

The choice of mean function in a Gaussian Process (GP) model significantly impacts its performance and should be considered carefully (Roberts et al., 2013; Chen, 2017). While a zero-offset mean function is often used for simplicity, particularly when data can be centralized to meet the zero-mean assumption, this may not always be the optimal choice. The influence of the mean function becomes more pronounced in regions far from the training data, especially in forecasting scenarios where long-term predictions are more sensitive to the chosen mean function than short-term ones. For instance, consider a dataset generated by an exponential decay function with added noise: $y = f + \epsilon$, with $f = \exp(-0.2x)$ where $\varepsilon \sim \mathcal{N}(0, 0.02)$ and the seven points of x are equally spaced in [0,3]. In this case, the observations are generated by an exponential decay function with noise. Here, a suitable mean function could take the form: $\mu(x) = a\exp(-bx)$, where a, b are unknown hyperparameters to be estimated alongside the kernel hyperparameters.

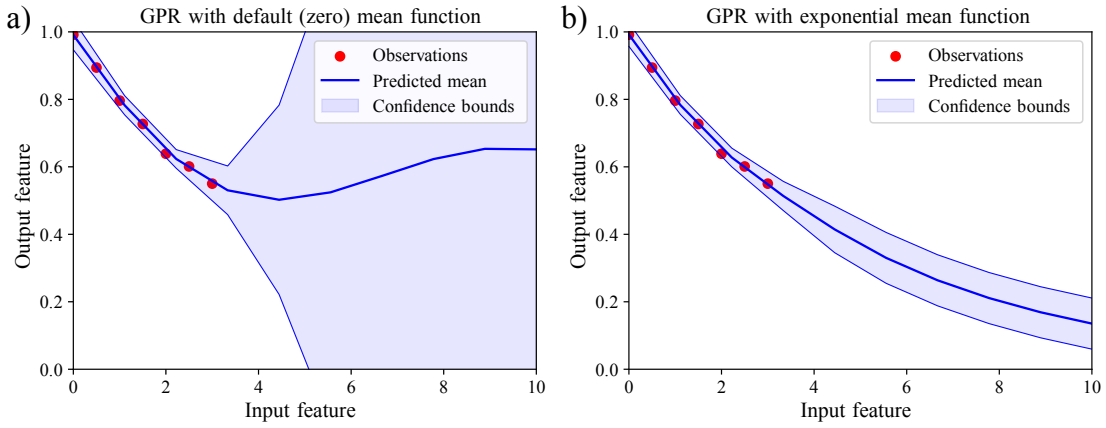


Figure 3.2: (a) It uses the default (zero) mean function while (b) used exponential decay mean function for prediction

While GPR demonstrates excellent performance on training data, long-term predictions can be heavily influenced by the choice of mean function. Figure 3.2(b) illustrates this phenomenon, where two GPR models share the same kernel function but differ in their

mean functions. Despite having only seven data points, the model with an exponential decay mean function (with unknown hyperparameters) clearly dominates the long-term forecasts. This example underscores the potential of incorporating domain-specific knowledge through an appropriate mean function to significantly improve GPR's predictive performance. However, identifying the most suitable mean function remains a non-trivial task and requires careful consideration.

3.3.3 Model parameter estimation

In previous sections, we explored the construction of Gaussian process regression (GPR) models with a specified kernel and zero mean function. We saw that predictive means and variances can be obtained once the undetermined hyperparameters are learned from the data. Following a Bayesian approach, we must define a prior distribution over these hyperparameters and integrate them out to make predictions. In other words, our goal is to determine the posterior distribution of the hyperparameters, given the observed data. Thus, we need to estimate:

$$p(y_* \mid X_*, \mathcal{D}) = \int p(y_* \mid X_*, \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathcal{D}) d\boldsymbol{\theta} \tag{3.22}$$

where y_* is the sum of f_* and the noise and $\boldsymbol{\theta} = \{\theta_1, \theta_2, \cdots\}$ denotes the set of all hyperparameters. However, obtaining an analytical solution for this integral is typically infeasible. To address this challenge, two primary methods are employed (MacKay et al., 1998): (1) Evaluate posterior directly through Monte Carlo methods, and (2) maximizing the marginal likelihood. The first approach involves approximating the integral using the most plausible hyperparameter values $\boldsymbol{\theta}_R$, expressed as $p(y_* \mid X_*, \mathcal{D}) = p(y_* \mid X_*, \mathcal{D}, \boldsymbol{\theta}_R)$. Although Monte Carlo methods can perform GPR without estimating hyperparameters, they are computationally intensive. Therefore, approximating the integral with the most plausible hyperparameter values is often preferred. When the most plausible hyperparameters value refers to the most probable value, this approach is known as maximum marginal likelihood. The marginal likelihood for a noisy regression, denoted as $p(\boldsymbol{y} \mid X, \boldsymbol{\theta})$, can be computed using:

$$p(\boldsymbol{y} \mid X, \boldsymbol{\theta}) = \int p(\boldsymbol{y} \mid f, X, \boldsymbol{\theta}) p(f \mid X, \boldsymbol{\theta}) df \tag{3.23}$$

In GPR models, the prior, likelihood and marginal likelihood are Gaussian and given by:

$$\begin{aligned} p(f \mid X, \boldsymbol{\theta}) &= \mathcal{N}(\mathbf{0}, K), \\ p(\boldsymbol{y} \mid f, X, \boldsymbol{\theta}) &= \mathcal{N}\left(f, \sigma_n^2 \mathbf{I}\right). \end{aligned} \tag{3.24}$$

$$p(\boldsymbol{y} \mid X, \boldsymbol{\theta}) = \int \mathcal{N}\left(f, \sigma_n^2 I\right) \mathcal{N}(\mathbf{0}, K) df = \mathcal{N}\left(\mathbf{0}, K + \sigma_n^2 I\right) = \mathcal{N}\left(\mathbf{0}, \Sigma_\theta\right) \tag{3.25}$$

where we denote $\Sigma_\theta = K_\theta + \sigma_n^2\mathbf{I} = K + \sigma_n^2\mathbf{I}$ since $\boldsymbol{\theta}$ is involved in the covariance matrix K. The commonly used negative log marginal likelihood is denoted by

$$\mathcal{L} = -\log p(\boldsymbol{y} \mid X, \boldsymbol{\theta}) = \frac{1}{2}\boldsymbol{y}^T\Sigma_\theta^{-1}\boldsymbol{y} + \frac{1}{2}\log\det\Sigma_\theta + \frac{n}{2}\log 2\pi. \tag{3.26}$$

The partial derivatives of negative log marginal likelihood with respect to the hyperparameters are given by

$$\frac{\partial\mathcal{L}}{\partial\theta_i} = \frac{1}{2}\,tr(\Sigma_\theta^{-1}\frac{\partial\Sigma_\theta}{\partial\theta_i}) - \frac{1}{2}\boldsymbol{y}^T\Sigma_\theta^{-1}\frac{\partial\Sigma_\theta}{\partial\theta_i}\Sigma_\theta^{-1}\boldsymbol{y} \tag{3.27}$$

In fact, the noisy GPR models can also be considered as the noise-free regression with a noisy kernel,

$$y = f,\ f \sim \mathcal{GP}\left(0, k'\right), \tag{3.28}$$

where $k' = k'\left(x_i, x_j\right) = k\left(x_i, x_j\right) + \delta_{ij}\sigma_n^2$ and $\delta_{ij} = 1$ if $i = j$, otherwise $\delta_{ij} = 0$. The marginal likelihood is regarded as the Gaussian prior multiplying the identity likelihood, and of course, it remains Gaussian. It is the same expression as the noisy regression model since

$$p(\boldsymbol{y} \mid X, \boldsymbol{\theta}) = p(f \mid X, \boldsymbol{\theta}) = \mathcal{N}\left(\mathbf{0}, K'\right) = \mathcal{N}\left(\mathbf{0}, K + \sigma_n^2 I\right) = \mathcal{N}\left(\mathbf{0}, \Sigma_\theta\right) \tag{3.29}$$

where $K' = K'(X, X) = K(X, X) + \sigma_n^2\mathbf{I}$. The third equality is due to the definition of the kernel and the covariance matrix. In this case, the estimation of hyperparameters must contain the noise level and then the negative log marginal likelihood is rewritten as $\mathcal{L}\left(\boldsymbol{\theta}, \sigma_n^2\right)$. The partial derivatives of negative log marginal likelihood with respect to σ_n^2 are given by

$$\begin{aligned}
\frac{\partial}{\partial\theta_i}\mathcal{L}\left(\boldsymbol{\theta}, \sigma_n^2\right) &= \frac{1}{2}tr\left(\Sigma_\theta^{-1}\frac{\partial\Sigma_\theta}{\partial\theta_i}\right) - \frac{1}{2}\boldsymbol{y}^T\Sigma_\theta^{-1}\frac{\partial\Sigma_\theta}{\partial\theta_i}\Sigma_\theta^{-1}\boldsymbol{y} \\
\frac{\partial}{\partial\sigma_n^2}\mathcal{L}\left(\boldsymbol{\theta}, \sigma_n^2\right) &= \frac{1}{2}tr\left(\Sigma_\theta^{-1}\right) - \frac{1}{2}\boldsymbol{y}^T\Sigma_\theta^{-1}\Sigma_\theta^{-1}\boldsymbol{y}
\end{aligned} \tag{3.30}$$

3.3.4 Advantages of the GPR

Gaussian Processes (GPs) offer several compelling advantages that make them particularly well-suited for constructing regression models:

1. **Analytic inference:** Given a kernel function and some observations, the predictive posterior distribution can be computed exactly in closed form. This is a rare and valuable property for nonparametric models.
2. **Expressivity:** By selecting an appropriate covariance function, GPs can embody a wide range of modeling assumptions, allowing for flexible and expressive model design.

3. **Integration over hypotheses:** The ability of a GP posterior, given a fixed kernel, to integrate exactly over a broad spectrum of hypotheses helps mitigate overfitting. Unlike neural networks, which require the estimation of many parameters, GPs involve relatively few parameters, reducing the need for complex optimization or regularization techniques.
4. **Model selection:** The capacity to integrate over all hypotheses also allows for the computation of the marginal likelihood of the data given a model. This provides a principled method for comparing different models.
5. **Closed-form predictive distribution:** The predictive distribution of a GP at a set of test points is simply a multivariate Gaussian distribution. This feature facilitates the composition of GPs with other models or decision-making processes.
6. **Easy to analyze:** Although restricting to a limited model class might seem unsatisfactory compared to attempting inference in the set of all computable functions, simple models can serve as well-understood building blocks for constructing more sophisticated models.

For instance, linear models, despite their limited scope, are simple, easy to analyze, and can be integrated into other models or procedures. Gaussian Processes can be viewed as an extension of linear models that retain these desirable properties while offering enhanced flexibility and capability.

3.3.5 Challenges and limitations of the GPR

While Gaussian Process Regression (GPR) offers many advantages, it also presents several challenges:

1. **Local convergence:** The marginal likelihood function for many kernels is not convex with respect to the hyperparameters, leading optimization algorithms to often converge to local optima rather than the global optimum, which might provide better results. Consequently, the optimized hyperparameters obtained through maximum likelihood estimation and the performance of GPR may heavily depend on the initial values of the optimization algorithm (Duvenaud et al., 2013).

2. **Computation cost:** Evaluating the gradient of the log likelihood necessitates the computation of the inverse of a matrix, which incurs a computational cost of $O(n^3)$. As a result, computing gradients is time-consuming for large datasets, making GPR computationally intensive for large-scale problems (MacKay et al., 1998).

3. **Non-Gaussian likelihood:** The predictive distribution of a standard GP model is Gaussian. However, in some cases, non-Gaussian predictive likelihoods are desirable, such as for robustness to outliers or for classification tasks. Utilizing non-Gaussian likelihoods requires approximate inference, which can be challenging to implement

effectively and out of the scope of this work.

4. **Proper selection of kernel:** The flexibility of GP models raises the critical question of which kernel to use for a given problem. Choosing an appropriate kernel is akin to learning a useful representation of the input. Although kernel parameters can be automatically set by maximizing the marginal likelihood, historically, the choice of the kernel's parametric form required human expertise .

5. **Extrapolation:** GPR can be unreliable when extrapolating beyond the range of the training data. The product-kernel model, while flexible, may remain uncertain about the function away from the data. In contrast, additive kernel structures sometimes allow predictions far from the training data. This makes GPR's uncertainty estimates potentially overly optimistic when extrapolating.

An important method known as cross-validation is effective for selecting a kernel and estimating parameters, besides Monte Carlo methods and maximum marginal likelihood estimation. Cross-validation compares various models and chooses the one with the least error, often outperforming maximum likelihood estimation when the kernel is improperly specified (Bachoc, 2013). However, in our study, we adopt the maximum marginal likelihood approach but enhances the optimization through entropy-assisted genetic algorithm for evaluating the maximum likelihood. This global optimization method uses Rényi's entropy to diversify the initial population of the genetic algorithm, preventing premature convergence. Detailed discussions on this method are provided in the following sections.

Further, to visualize the performance of GPR for interpolation and extrapolation using the squared exponential (SE) kernel, consider two sets of parameters for the GPR model: (i) suboptimal parameters: $s_f = 2$, $l = 2$ with noise $\sigma = 0.2$, and (ii) optimal parameters: $s_f = 4$, $l = 4$ with noise $\sigma = 0.05$. The results for both interpolation and extrapolation with these parameter sets are illustrated in Figure 3.3, showing (a) suboptimal GPR interpolation, (b) optimal GPR interpolation, (c) suboptimal GPR extrapolation, and (d) optimal GPR extrapolation using different hyperparameter values.

The most popular choice for the kernel function in Gaussian Process Regression is the Squared Exponential (SE) kernel. However, the SE kernel has a strong tendency to produce forecasts that revert to the mean function as you move away from the training data points, as shown in Figure 3.3. This behavior occurs because the SE kernel's covariance decays exponentially with distance, assigning less weight to distant points. To address this limitation, this thesis primarily employs a product of linear and squared exponential kernels for degradation prognosis. This combined approach leverages the strengths of both kernels to improve prediction accuracy.

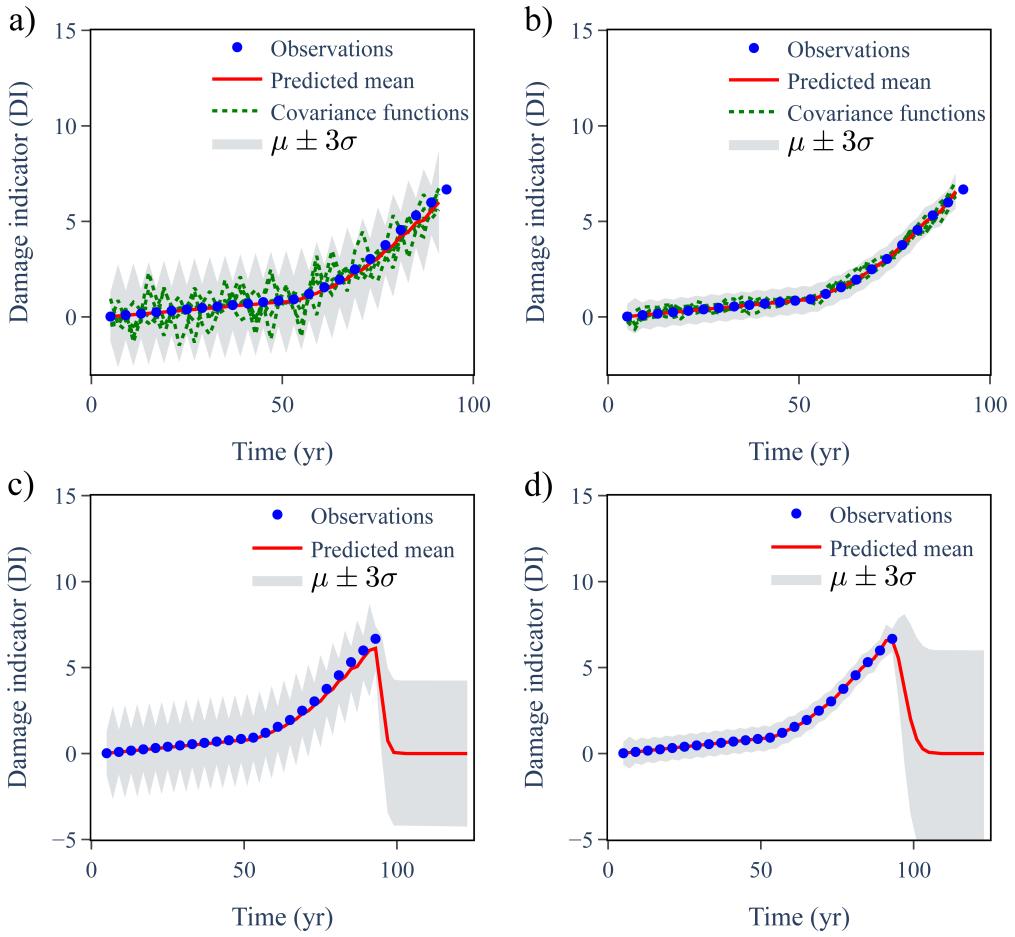


Figure 3.3: Illustration of the working mechanism of GPR (a) suboptimal GPR interpolation, (b) suboptimal GPR extrapolation, (c) suboptimal GPR extrapolation, and (d) optimal GPR extrapolation.

In the field of prognosis and remaining useful life (RUL) estimation, Meng et al. (2022) highlight the challenges of direct GPR extrapolation. They utilized a compound kernel function based on the Matérn covariance class, specifically Ma3 + Ma5, for their comparative study. Their results indicate that the prediction performance of direct GPR is significantly constrained by the amount of available degradation data, even when using advanced kernel functions. This underscores the necessity for adaptive mean functions to enhance extrapolation capabilities.

3.4 Addressing GPR challenges

For the extrapolation, the strong requirement is the adoption of mean or basis function. The formulation of incorporating the mean fuunction is discussed in this section. Consider the prior function $g(\boldsymbol{x}) = f(\boldsymbol{x}) + \boldsymbol{h}(\boldsymbol{x})^{\top}\beta$ where $f(\boldsymbol{x}) \sim \mathcal{GP}(0, k(\boldsymbol{x}, \boldsymbol{x'}))$ is a zero mean Gaussian process, $\boldsymbol{h}(\boldsymbol{x})$ are a set of fixed explicit basis functions that transform the original feature vector into a new feature vector, and β are basis function coefficients. Here, linear and quadratic polynomial basis functions are selected depending on the observations. The

basis function is given as follows:

$$h(x) = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \text{ or } \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \tag{3.31}$$

The explicit basis function coefficients, say β can be optimized along with the hyperparameters of covariance function. Moreover, if the prior distribution of β is assumed be Gaussian, $\beta \sim \mathcal{N}(\boldsymbol{b}, B)$, where $\boldsymbol{b}$ and B are the parameters, then the prior function is written as follows:

$$g(\boldsymbol{x}) \sim \mathcal{GP}\left(\boldsymbol{h}(\boldsymbol{x})^\top \boldsymbol{b}, k(\boldsymbol{x}, \boldsymbol{x'}) + \boldsymbol{h}(\boldsymbol{x})^\top B \boldsymbol{h}(\boldsymbol{x'})\right) \tag{3.32}$$

Thus, the posterior function after fitting the priors on observed data is written as:

$$\begin{aligned} \bar{\boldsymbol{g}}(X_*) &= H_*^\top \bar{\beta} + K_* K_y^{-1}\left(\boldsymbol{y} - H\bar{\beta}\right) = \bar{\boldsymbol{f}}(X_*) + R^\top \bar{\beta}, \\ \operatorname{cov}(\boldsymbol{g}_*) &= \operatorname{cov}(\boldsymbol{f}_*) + R^\top \left(B^{-1} + H K_y^{-1} H^\top\right)^{-1} R \end{aligned} \tag{3.33}$$

where, $\bar{\beta} = \left(B^{-1} + H K_y^{-1} H^\top\right)^{-1}\left(H K_y^{-1}\mathbf{y} + B^{-1}\boldsymbol{b}\right)$, and $R = H_* - H K_y^{-1} K_*$. The $*$ represents the test data set.

The challenge exist here is about proper selection of the mean function $\boldsymbol{h}(\boldsymbol{x})$. Thus, in this work, a set of basis functions are tested on the observations and the function with the best goodness of fit and increasing trend is selected as the mean function. The next challenge of local and poor convergence is addressed in this thesis through proposal of entropy assisted genetic algorithm for the optimization of hyperparameters while evaluation the maximum log likelihood which is discussed below. In this study, the GPR that uses basis function is termed as Basis GPR or BGPR.

3.4.1 Proposed entropy-assisted genetic algorithm (EGA)

The conjugate gradient method, while effective for optimizing Gaussian Process Regression (GPR) hyperparameters, can be sensitive to initial values and may converge to local optima if these values are not carefully chosen (Fang et al., 2018). Conventional GPR optimization often involves a two-step process: first, maximizing the log-likelihood function with respect to basis parameters (β) for assumed hyperparameters, and then maximizing the β-profiled log-likelihood over the hyperparameters to obtain their estimates. This process, however, can be complex and may lead to suboptimal solutions due to the limitations of the conjugate gradient method.

To address these challenges, this thesis proposes the use of a genetic algorithm (GA) for optimizing Basis GPR (BGPR) hyperparameters. GAs offer several advantages over traditional methods, including the ability to search a larger domain of hyperparameters, handle discretization effectively, and provide global optimum solutions. However, classical GAs can also suffer from premature convergence and stagnation due to improper population initialization (Jiacheng and Lei, 2020).

To overcome this limitation, we introduce an entropy-assisted genetic algorithm (EGA) for BGPR hyperparameter optimization. EGA leverages information entropy to diversify the initial population, promoting a broader exploration of the search space and reducing the risk of premature convergence. The specific advantages of EGA for BGPR optimization include:

1. **Enhanced exploration of the hyperparameter space:** EGA's ability to diversify the initial population allows for a more thorough search, increasing the chances of finding the global optimum.

2. **Effective handling of discretization:** EGA can naturally handle discrete hyperparameters, which are common in GPR models.

3. **Leveraging information entropy:** EGA utilizes information entropy as a measure of population diversity, guiding the selection and reproduction processes to maintain a balance between exploration and exploitation.

4. **Global optimization:** By mitigating premature convergence, EGA increases the likelihood of finding the globally optimal hyperparameters.

5. **Efficiency:** EGA can be computationally efficient, especially when combined with parallel computing techniques.

By integrating EGA with BGPR, we create a powerful and robust framework for prognosis, capable of handling complex datasets, addressing extrapolation challenges, and quantifying uncertainties effectively. The entropy-assisted genetic algorithm (EGA) leverages the concept of information entropy to enhance population diversity, a crucial factor in achieving global optimization. In information theory, entropy quantifies the uncertainty or randomness associated with selecting information. It can also be interpreted as the average amount of information conveyed by a variable. Within the context of EGA, a non-uniform population distribution implies a lower entropy value, as the statistical average of information content across the probability space is reduced (Shannon, 1948; Rényi, 1961). By incorporating entropy as a guiding principle, EGA ensures a more diverse initial population, promoting a broader exploration of the search space and reducing the risk of premature convergence to

local optima. The Rényi entropy of order α , where $0 < \alpha < \infty$ and $\alpha \neq 1$, is defined as

$$\mathrm{H}_{\alpha}(X) = \frac{1}{1-\alpha} \log \left(\sum_{i=1}^{n} p_i^{\alpha} \right) \tag{3.34}$$

Further, for $\alpha = 0, 1, \infty$,it is defined as

$$\mathrm{H}_{\alpha}(X) = \lim_{\gamma \to \alpha} \mathrm{H}_{\gamma}(X) \tag{3.35}$$

Here, X is a discrete random variable referring with possible outcomes in the set $\mathcal{A} = \{x_1, x_2, \ldots, x_n\}$ and corresponding probabilities $p_i \doteq \Pr(X = x_i)$ for $i = 1, \ldots, n$. Thus, in the case of genetic algorithm, the X refers to the population generated out of the various possible populations. The process of initialization of population and complete algorithm of EGA can be summarized as follows:

Step 1: Initialization of population: To initiate the optimization process, hyperparameters are encoded into chromosomes, each composed of genes that the genetic algorithm (GA) can manipulate. The population represents a set of candidate solutions, with each chromosome's encoded content serving as the genotype – a potential feasible solution of hyperparameters. In this work, we adopt binary encoding, converting each hyperparameter into a bitstring of zeros and ones. The bitstring length (e.g., 4, 8, 16 bits) depends on the parameter's complexity and value range. However, a uniformly distributed population, crucial for maximizing diversity and avoiding premature convergence, requires maximum entropy. Achieving this balance involves selecting an optimal bitstring size. While larger bitstring offer finer granularity, they increase computational cost. Conversely, smaller bitstring may lead to lower entropy and premature convergence.To address this trade-off, we employ a novel approach: generating the initial population using an optimum bitstring size that maximizes Rényi's entropy. **Thus, the bitstring size that provide maximum Rényi's entropy is considered as the optimal size**. This approach ensures sufficient diversity while maintaining computational feasibility, ultimately improving the GA's effectiveness in finding the global optimum. Thus, to uniformly distribute the initial population within the solution space, avoid centralized clustering, and enhance the diversity of the initial population, information entropy can be used for initialization. The process is outlined as follows and is illustrated in Figure 3.4:

1. **Set the critical entropy value** S_0: Determine the threshold entropy value to ensure adequate diversity.
2. **Generate the first chromosome randomly**: Create the first individual in the chromosome domain randomly.
3. **Iteratively generate and evaluate chromosomes**:

- Generate a new chromosome.
- Calculate the entropy value S of this chromosome relative to the existing population.
- If $S > S_0$, accept the chromosome; otherwise, reject it and generate a new one.
- Repeat the process until $S > S_0$ is met.

4. **Repeat until the population is complete**: Continue steps 3 and 4 until the initial population reaches the desired size.

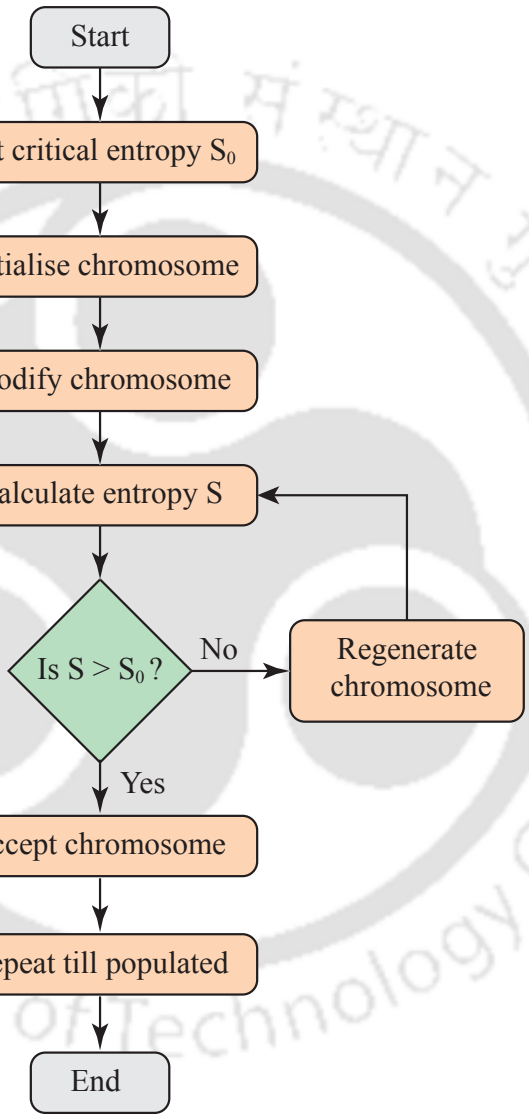


Figure 3.4: Illustration of working mechanism of initialization of population through maximizing entropy

Step 2: Fitness value: For each candidate solution (set of hyperparameters) within the population, the corresponding GPR model is evaluated, and a fitness value is calculated. In this context, the fitness value is defined as the maximum likelihood achieved by the GPR algorithm when using the given set of hyperparameters. The genetic algorithm's role is to iteratively refine the population, searching for the optimal set of hyperparameters that

maximizes this likelihood function and, consequently, the model's overall performance.

Step 3: Selection of parents: The process of natural selection, as mirrored in a genetic algorithm, begins by identifying the fittest individuals within a population. These individuals are then selected as parents to produce offspring, inheriting their characteristics. Higher parental fitness generally results in superior offspring, increasing their likelihood of survival and further reproduction. In the context of hyperparameter optimization using a genetic algorithm, the candidate solutions with the highest fitness values (i.e., those producing GPR models with the highest likelihood) are chosen as parents. Offspring are generated through crossover (combining genetic material from two parents) and mutation (randomly altering genes), while elitism ensures the preservation of the best individuals across generations. This iterative process continues, evolving the population towards optimal solutions over successive generations.

Step 4: Crossover operation, mutation, and elitism: In the crossover operation, a random crossover point is selected for each pair of parent chromosomes. For instance, in a 16-bit string, this point could be the 10th bit. The gene values up to the crossover point are then exchanged between the parents, resulting in two new offspring. Mutation, on the other hand, introduces random alterations to some gene values within the parents. This process ensures genetic diversity and helps prevent the algorithm from getting trapped in local optima. Elitism further enhances diversity by randomly introducing a foreign candidate into the population, though its ratio is kept low to maintain a reasonable convergence rate. The crossover probability is typically set high to encourage the exploration of new combinations of genetic material, while the mutation probability is kept relatively low to avoid excessive disruption of promising solutions. This balance between exploration and exploitation promotes the efficient discovery of optimal hyperparameters for the GPR model.

Step 5: Obtaining final solution: The genetic optimization process terminates when either the specified number of generations has been reached or the population converges, indicating that further improvement is unlikely. The chromosome from the final generation that exhibits the highest fitness value is then selected as the optimal solution, representing the most suitable set of hyperparameters for the GPR model.

By incorporating entropy as a guiding principle, EGA ensures a more diverse initial population, promoting a broader exploration of the search space and reducing the risk of premature convergence to local optima. Figure 3.5 shows the comparision of initial population distribution generated in EGA with the tradition GA. To verify the performance of the EGA algorithm, a set of well-known base test functions are selected to test it in this paper. This set of benchmark functions includes six single-peak functions and six multi-peak functions. The single-peak functions F1-F6 have only one globally optimal solution, so they are often used to evaluate the algorithm's local exploitation ability. In

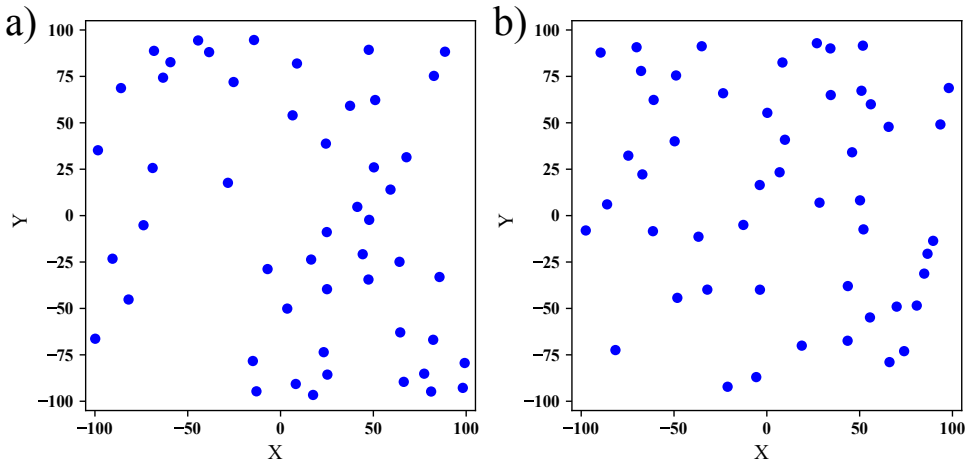


Figure 3.5: A comparision of population generated through (a) randomly initialised , and (b) entropy assisted initialised

addition to the global optimal solution, the multi-peak functions F7-F12 have multiple local optimal solutions, which are used to evaluate the algorithm's ability to explore globally and avoid falling into local optima. A 3D visualization of these functions is given in Figure 3.6. Further, Table 3.1 shows the results of gradient based method, GA, and EGA based optimization of benchmark functions.

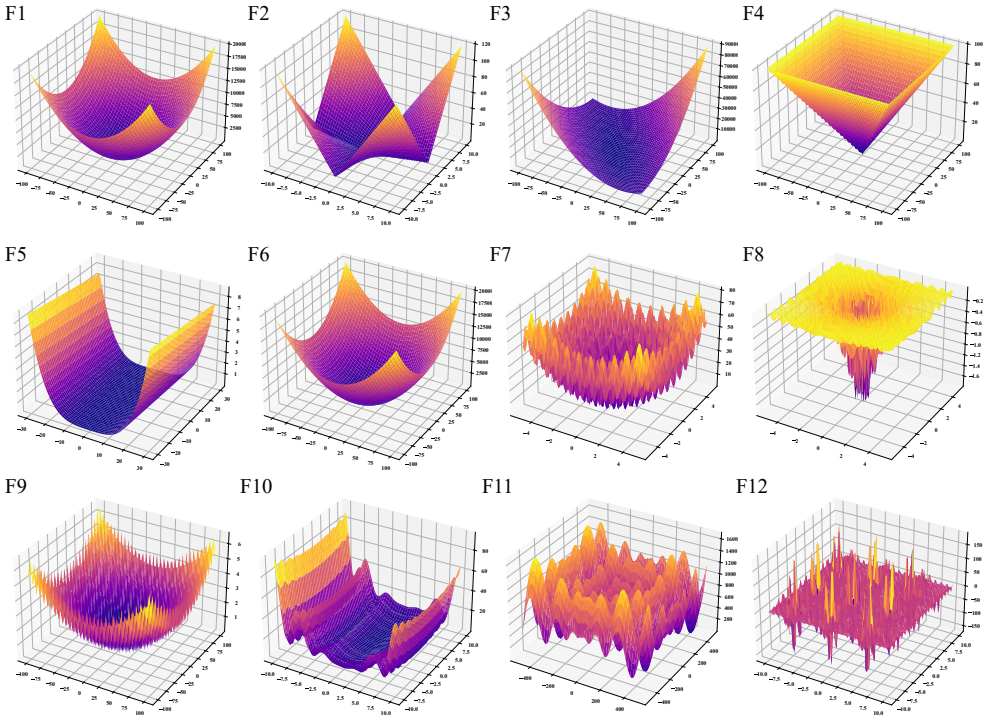


Figure 3.6: 3D visualization of the benchmark functions for optimization

It can be clearly seen that EGA results are better. Thus, leveraging the EGA for optimization of prognosis surrogate model i.e., Gaussian process regression, we achieve

Table 3.1: The results of different approaches for the optimization of benchmark function F1 -F12

	Function name	Optimum	Gradient-based	GA	EGA
F1	Sphere function	0	6.25E-15	5.16E-24	2.56E-110
F2	Schwefel 2.22 function	0	1.41E-08	3.83E-17	9.81E-45
F3	Sum Squares function	0	1.43E-12	2.28E-16	1.06E-24
F4	Infinity Norm function	0	6.06E-10	4.91E-15	1.99E-18
F5	Rosenbrock function	0	0.792	0.205	0.012
F6	Shifted Sphere function	0	4.99E-17	1.05E-25	3.08E-33
F7	Rastrigin function	0	3.97	2.03E-03	4.03E-08
F8	Drop-Wave function	-1	-0.16	-1.76	-1.16
F9	Griewank function	0	0.717	0.01	0.009
F10	Levy function	0	1.16	9.48E-15	1.49E-32
F11	Schwefel function	0	415.03	1.18E-02	2.54E-05
F12	Shubert function	186.73	-12.76	-132.52	-186.73

global convergence. Further, the effectiveness and accuracy of the EGA algorithm in prognosis of deteriorating systems will be further demonstrated through real-world applications in subsequent chapters. These applications will highlight the versatility of the EGA-BGPR framework in addressing diverse degradation scenarios. It is important to note that the development of a degradation index is a case-specific process, depending on the nature of the problem and the available data. In this thesis, we utilize the Bhattacharyya distance as a statistical measure for damage quantification, a two-stage constrained unscented Kalman filter for identifying parameters in complex systems, and a Bayesian neural network for mapping signals to crack widths. Each of these algorithms is strategically chosen based on its suitability for the specific case study.

Given its relevance as a surrogate model in this thesis, a brief overview of Bayesian neural networks (BNNs) is presented in this chapter. BNNs offer a powerful approach to uncertainty quantification in machine learning, making them particularly valuable for prognosis tasks where reliable uncertainty estimates are crucial.

3.5 Bayesian neural network

The objective of Bayesian Neural Networks (BNNs) is to estimate the posterior parameter distributions $P(\boldsymbol{\theta} \mid \boldsymbol{y})$ of a set of network parameters $\boldsymbol{\theta}$. This estimation is performed without any prior knowledge of the underlying phenomenon, aiming to align the statistics of the neural network's output $f(\boldsymbol{x})$ with the observed data $\boldsymbol{y}$. We employ a fully-connected feed-forward neural network with multiple hidden layers to approximate the output $f(\boldsymbol{x})$ from the input $\boldsymbol{x}$. The hidden variables of the k-th hidden layer are denoted as $\boldsymbol{z}^k$, and the neural network architecture can be summarized as follows (illustrated here with two hidden

layers):

$$\boldsymbol{z}_0 = \boldsymbol{t} \tag{3.36}$$

$$\boldsymbol{z}_1 = \sigma(\boldsymbol{W}_1 \boldsymbol{z}_0 + \boldsymbol{b}_1) \tag{3.37}$$

$$\boldsymbol{z}_2 = \sigma(\boldsymbol{W}_2 \boldsymbol{z}_1 + \boldsymbol{b}_2) \tag{3.38}$$

$$\boldsymbol{z}_3 = (\boldsymbol{W}_3 \boldsymbol{z}_2 + \boldsymbol{b}_3) \tag{3.39}$$

The output of the final layer, $\boldsymbol{z}_3$, approximates the true solution $\boldsymbol{x}$. Here, $\sigma(\cdot)$ represents the non-linear activation function (chosen as the hyperbolic tangent function, $\tanh(\cdot)$), while $\boldsymbol{W}_k$ and $\boldsymbol{b}_k$ denote the network weights and biases of the k-th layer. To estimate the posterior probability distribution of the network parameters $\boldsymbol{\theta}$, we utilize Bayes' theorem:

$$P(\boldsymbol{\theta} \mid \boldsymbol{y}) = \frac{P(\boldsymbol{y} \mid \boldsymbol{\theta}) P(\boldsymbol{\theta})}{P(\boldsymbol{y})} \tag{3.40}$$

In this equation, $P(\boldsymbol{y} \mid \boldsymbol{\theta})$ is the likelihood (conditional probability of the data given fixed parameters), $P(\boldsymbol{\theta})$ is the prior (probability distribution of the parameters), $P(\boldsymbol{y})$ is the marginal likelihood (evidence), and $P(\boldsymbol{\theta} \mid \mathbf{y})$ is the posterior (conditional probability of the parameters given the data). Here, the likelihood $P(\boldsymbol{y} \mid \boldsymbol{\theta})$ is modeled as the product of individual likelihoods $p_i(y \mid \boldsymbol{\theta})$:

$$P(\boldsymbol{y} \mid \boldsymbol{\theta}) = \prod_{i=0}^{n} p_i(y \mid \boldsymbol{\theta}) \text{ with } p_i(y \mid \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\| y - x(t_i) \|^2}{2\sigma^2} \right) \tag{3.41}$$

Each $p_i(y \mid \boldsymbol{\theta})$ follows a normal distribution $\mathcal{N}(\mu, \sigma)$, where μ is the model output $x(t_i)$ and σ is a fixed standard deviation. For the prior distributions $P(\boldsymbol{\theta})$, we assume independent, weakly informative normal distributions with zero means for both the network weights $\boldsymbol{W}_k$ and biases $\boldsymbol{b}_k$.

The Bayesian approach to neural networks offers a powerful framework for addressing epistemic uncertainty, stemming from limited training data and reflecting our lack of complete knowledge about the surrogate model. Unlike traditional artificial neural networks with fixed weights and biases, Bayesian neural networks (BNNs) focus on learning weight distributions. This involves defining prior and posterior distributions for the network's parameters, which are updated during training, enabling the model to adapt and refine its understanding of the underlying relationships. Moreover, BNNs extend their capabilities to address aleatoric uncertainty, which arises from inherent randomness or noise in the data. By producing a distribution of outputs rather than a single point estimate, BNNs, also known as probabilistic BNNs (PBNNs), capture both epistemic and aleatoric uncertainty. The fundamental architecture of PBNNs, illustrated in Figure 3.7,

offers several advantages:

1. **Robustness to overfitting:** PBNNs mitigate overfitting by incorporating Bayesian inference, making them effective even with sparse datasets.

2. **Comprehensive uncertainty quantification:** By addressing both epistemic and aleatoric uncertainty, PBNNs provide a more holistic understanding of the uncertainties associated with predictions, leading to increased model reliability.

3. **Variance decomposition:** PBNNs produce output distributions whose variance is governed by the law of total variance and variance accumulation theory. This allows for the decomposition of output variance, attributing it to different sources of uncertainty, thus providing deeper insights into the model's behavior.

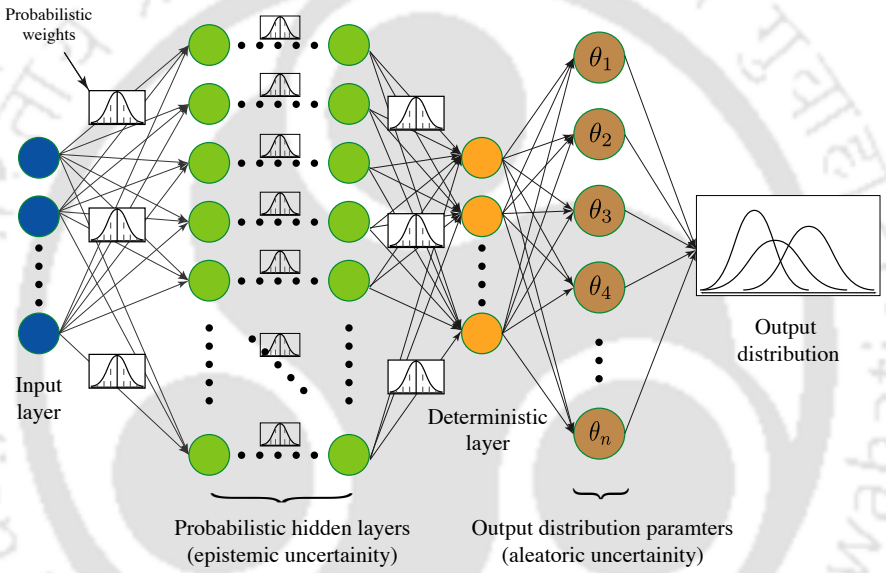


Figure 3.7: Basic mechanism of PBNN

Now, extending the formulation to the probabilistic Bayesian neural network, the weights $\Omega = \{\boldsymbol{W_1}, \boldsymbol{W_2}, , \ldots, \boldsymbol{W_n}\}$ are considered to be probabilistic having Gaussian prior, which represents epistemic uncertainty. Further, in PBNNs, an immediate layer to output features the parameters of the output distribution and thereby accounts for the aleatoric uncertainty and is represented by parameter θ as shown in the aleatoric uncertainty layer in Figure 3.7. Thus, for the PBNNs, the probability of the output can be written as

$$p(y \mid x, \Omega) = \int p(y; \theta)\, p(\theta \mid x, \Omega)\, d\theta \tag{3.42}$$

where Ω denotes random model weights. When the input features X are known and the corresponding labels Y are available, the posterior $p(\Omega \mid X, Y)$ is written as

$$p(\Omega \mid X, Y) = \frac{\int p(Y \mid \theta) p(\theta \mid X, \Omega) p(\Omega) d\theta}{\int p(Y \mid \theta) p(\theta \mid X, \Omega) p(\Omega) d\theta d\Omega} \tag{3.43}$$

having, $p(\theta \mid x, \Omega) = p(\theta \mid f_{\Omega}(x))$. The function $f_{\Omega}(x)$ is just like the weighted sum as obtained in case of ANN. The difference here is that the weights are now Gaussian in nature. Each component of weights Ω is considered as independent Gaussian distribution with zero mean, and the w_l and b have the variances $\sigma_{w,l}$ and $\sigma_{b,l}$, respectively, for the number of neurons in the hidden layers $l = 0, 1, \ldots, L$. For such a nontrivial model, even if the evidence has been computed, directly sampling the posterior become difficult if there is a large number of hidden layers associated with the PBNNs. Thus, the posterior is evaluated through Markov chain Monte Carlo (MCMC) methods (Hastings, 1970) if the model is shallow, or variational inference (Blei et al., 2017) which approximates the posterior through assumed guide distribution. This thesis incorporates the Hamiltonian Monte Carlo method and stochastic variance inference method whole details are provided in Appendix A and B.

3.6 Summary and conclusions

This chapter provides a foundational understanding of probabilistic machine learning, emphasizing the integration of probability theory with machine learning algorithms for robust uncertainty quantification. The chapter then underscores the importance of Bayesian machine learning frameworks, specifically Gaussian Process Regression (GPR). GPR is shown to excel in providing detailed uncertainty measures alongside predictions. The chapter elaborates on the GPR background, the impact of kernel and mean functions, and the optimization process. In the present study, GPR is used as a surrogate model for prognosis, with limitations and challenges discussed. The local convergence issue of GPR is addressed through the introduction of an entropy-assisted genetic algorithm (EGA). The detailed application of EGA in optimizing the GPR will be discussed in the next chapter while developing the framework for prognosis.

Furthermore, Bayesian Neural Networks (BNNs) are introduced for uncertainty quantification in complex, high-dimensional data. This chapter extends BNNs to the probabilistic Bayesian neural network, explicitly separating aleatoric and epistemic uncertainties. These techniques, by explicitly modeling and managing uncertainty, form the basis for the advanced probabilistic structural health monitoring (SHM) frameworks developed and presented in subsequent chapters.

Chapter 4

An integrated probabilistic methodology for the prognosis of multi-stage degradation using Gaussian process regression

4.1 Overview

This chapter introduces a integrated probabilistic methodology for the prognosis of systems exhibiting multi-stage degradation. The framework leverages machine learning appraoch known as Gaussian Process Regression (GPR) as its core modeling technique, addressing key challenges such as hyperparameter optimization and extrapolation issues. An innovative Entropy-Assisted Genetic Algorithm (EGA) is employed to optimize GPR hyperparameters, promoting global convergence and mitigating premature optimization. To enhance GPR's predictive capabilities for evolving degradation patterns, an adaptive mean/basis function is incorporated within the BGPR (basis GPR) model. Additionally, a Bayesian approach is implemented for change point detection, enabling the framework to dynamically adapt to shifts in system behavior. The effectiveness and robustness of this integrated methodology are demonstrated through numerical case studies showcasing its potential for accurate and reliable remaining useful life (RUL) prediction in engineering applications.

4.2 Background

The past few decades have witnessed significant advancements in condition-based monitoring (CBM) of dynamic structural systems, especially for those safeguarding structures against seismic events (Guo et al., 2019; Lei et al., 2020). The seismic control devices, such as rubber bearings offer both vertical support and horizontal flexibility yet remain susceptible to fatigue-induced delamination under ambient vibrations and crack development due to seismic excitations (Mars and Fatemi, 2005; Saintier et al., 2006). Thus, continuous monitoring of these critical components is essential for post-earthquake structural integrity assessments. Further, the data-driven methods for structural system diagnosis rely on either statistical analysis or extraction of system parameters from

measurements (Yang and Ma, 2003). Various approaches utilizing autoregressive models, SVM, neural networks, and EMD have been used to develop damage indicators based on surrogate measures (Sohn and Farrar, 2001; Nair et al., 2006; Hou et al., 2020). For system diagnosis, methodologies encompassing autoregressive modeling, Bayesian filters, and machine learning techniques have been extensively explored (Salehi and Burgueño, 2018; Yang et al., 2006; Ajavakom et al., 2008). However, structural systems prognosis remains an area demanding extensive research.

While prognostic methods can be physics-model-based, data-driven, or hybrid (Hu et al., 2019), this work focuses on data-driven algorithms that rely on empirical data and patterns. These techniques leverage probabilistic methods, and stochastic modeling to predict damage progression and estimate RUL. Notable reviews of statistical, data-driven methods for RUL estimation exist in following literature (Si et al., 2011). Gaussian process regression (GPR) offers a compelling surrogate modeling tool due to its flexibility, ability to incorporate uncertainty, and high-precision modeling of nonlinear relationships (Rasmussen and Williams, 2006). However, traditional GPR implementations often use a time-invariant or non-informative zero mean function, limiting their suitability for extrapolation. Further, optimizing GPR hyperparameters through the conventional conjugate gradient method suffers from susceptibility to local minima as discussed in Chapter 3 (Fang et al., 2018; Ouyang and Zou, 2021). This work addresses the aforementioned limitations and proposes the following innovations for enhanced damage prognosis:

1. **Basis GPR (BGPR):** Introduction of informative mean/basis functions to improve GPR performance, particularly for extrapolation. Thus, the use of mean/basis function in the Gaussian process regression will rise to the naming as basis GPR (BGPR). These models require the training of the appropriate posterior means through training samples as well as the proper selection of kernel function (Roman et al., 2021).
2. **Entropy-assisted genetic algorithm (EGA):** This hybrid algorithm combines Ranyi's information entropy with GA to diversify the initial population, thereby mitigating premature convergence and enhancing global optimization of BGPR hyperparameters.
3. **Bayesian change-point detection:** A Bayesian approach to identify shifts in degradation patterns through the likelihood function, facilitating the identification of change points for updating the mean function of the GPR.
4. **Adaptive mean function:** The mean function of the GPR function is adapted with the occurrence of change point. This adaption helps the surrogate model to extrapolate the prediction considering the change in degradation pattern.

This chapter explores damage prognosis by employing BGPR optimized through EGA.

Importantly, a data-driven change-point detection method is integrated to update the BGPR model. The effectiveness of proposed framework of optimizing BGPR through EGA considering the change point identification is validated on numerical case studies.

4.3 Methodology

The overall methodology for the prognosis in this study involves the Gaussian process regression with basis function, optimized through entropy assisted genetic algorithm. To carry out the prognosis, we need to first identify the damage or features through indirect observations and develop the degradation index. Next step starts with modeling degradation as a stochastic Gaussian process. Thus, the surrogate model of degradation is developed through Gaussian process regression (GPR). In GPR, observations are modeled using a combination of a mean function and a covariance function (also known as a kernel function). While a zero mean function is often used for various tasks, relying solely on the kernel function can significantly degrade the model's performance during extrapolation. Therefore, for prognosis problems that involve extrapolation, incorporating a mean function is crucial. The details about mathematical formulations of the GPR and explanations of the limitations for extrapolation is discussed in Chapter 3. Further, detailed explanation of the entropy assisted genetic algorithm is already discussed while addressing the challenges of the GPR. In the current study, this mean function is either explicitly trained as well as fine-tuned alongside the kernel hyperparameters. There are two approaches for incorporating the mean function (1) Normalization, and (2) Explicit basis function.

1. **Normalization:** Select the appropriate mean function, subtract the prior mean vector from the observations to obtain the normalized dataset, and simulate this dataset through the zero mean Gaussian process regression (GPR). Once the model gets trained and predicts the result, the same mean function is added back to the normalized prediction to obtain the actual predictions.

2. **Explicit basis function:** The GPR formulation is carried out using the basis function referred to as BGPR, and the basis function parameters are trained along with the GPR hyperparameters.

In this study, the several prior mean functions are pre-trained on the dataset and the optimum mean function is selected for basis function GPR modelling. Next, this pre-trained mean function is further optimized along with the GPR hyperparameters through entropy assisted genetic algorithm (EGA). The details of the EGA are already presented in previous chapter. Here, selection of the prior mean function and the formulations about Bayesian change point identification is discussed.

4.3.1 Selection of prior mean function and identification of change point

As surrogate modelling is performed through basis GPR (BGPR) optimized through EGA, it is important that framework should adapt the mean function on occurrence of change point (when trajectory changes) for accurate estimation. The basis functions can be linear, polynomial of different degree or exponential depending on the history of the data. Further, complete lifespan dataset of degradation index is not available for the pre-training of prior mean/basis function. Rather, the degradation index is evaluated till the current time and the dataset used for pre-training constitute the data points from the latest change point location to the current time. This pre-training helped in carrying out proper choice of the mean function as well as degree of the polynomial mean function. Thus, this adaptive mean function changes when the change point occurred and thus helps in prognosis through observing correct recent trend.

This chapter introduces a novel method to identify the occurrence of change point through observing changes in the parameters of stochastic Gaussian process. The degradation in this thesis is modelled using the Gaussian process regression which indirectly assumes that all the observation points are the part of the Gaussian distributions. Thus, the change in two subsequent values of degradation index can be simulated through Gaussian random number. The degradation index might possess a single change point or multiple change points depending on the history of the system. The distribution of the Gaussian random number is given as:

$$N(\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_j-\mu)^2}{2\sigma^2}\right) \tag{4.1}$$

It is important to mention here that as the degradation pattern changes, it can be assumed that the difference between successive damage index values, i.e., ξ_{jt} will also change its distribution parameters. Suppose the degradation pattern changes at τ and the last time of measurement is T , then

$$\xi_{jt} = DI_{j(t+t_0)} - DI_{jt} \sim N(\mu_0,\sigma_1) t_0 < t \leq \tau + 1 \tag{4.2}$$

$$\xi_{jt} = DI_{j(t+t_0)} - DI_{jt} \sim N(\mu_0,\sigma_2) \tau + 1 < t \leq T \tag{4.3}$$

where ξ_{ij} is the increment in damage index DI from time t to $t + t_0$ and t_0 time interval between two successive values of degradation index. Here, it is worth noting that the mean of the distribution is kept constant before and after the degradation pattern changes. With known variations $\xi_{j1}, \xi_{j2}, \ldots, \xi_{j(T-1)}$, the likelihood function of parameters σ_2 and τ can be obtained from above equation and written as follows:

$$L(\sigma_2,\tau) = \prod_{t=1}^{\tau} \frac{1}{\sqrt{2\pi{\sigma_1}^2}} \exp\left(-\frac{(\xi_{jt}-\mu_0)^2}{2{\sigma_1}^2}\right) \times \prod_{t=\tau+1}^{T-1} \frac{1}{\sqrt{2\pi{\sigma_2}^2}} \exp\left(-\frac{(\xi_{jt}-\mu_0)^2}{2{\sigma_2}^2}\right) \tag{4.4}$$

Taking log on both side, we will obtain the following log likelihood expression

$$\begin{aligned}\text{logL}\left(\sigma_2,\tau\right) = &-\sum_{\text{t=1}}^{\tau}\frac{1}{2}\ln\left(2\pi\right)-\sum_{t=1}^{\tau}\frac{1}{2}\ln\left({\sigma_1}^2\right)-\sum_{t=1}^{\tau}\frac{\left(\xi_{jt}-\mu_0\right)^2}{2{\sigma_1}^2}\\ &-\sum_{t=\tau+1}^{T-1}\frac{1}{2}\ln\left(2\pi\right)-\sum_{t=\tau+1}^{T-1}\frac{1}{2}\ln\left({\sigma_2}^2\right)-\sum_{t=\tau+1}^{T-1}\frac{\left(\xi_{jt}-\mu_0\right)^2}{2{\sigma_2}^2}\end{aligned} \tag{4.5}$$

The maximum likelihood is achieved by differentiating the log likelihood with respect to parameter σ_2 and equating the terms to zero. This will provide the value of σ_2 as follows:

$${\sigma_2}^2 = \frac{1}{T-\tau-1}\sum_{t=\tau+1}^{T-1}\left(\xi_{jt}-\mu_0\right)^2 \tag{4.6}$$

4.3.2 Computation of likelihood for non-Gaussian distributions

Till now, we have assumed that the successive change in degradation ξ_{jt} is following the Gaussian distribution with constant mean and change in variance is observed when change point occurs. Here, we will discuss the cases where ξ_{jt} does not follow the Gaussian distribution after the change points occurs. In such a case, the obtaining the closed form solution is difficult, however, the combination with exponential distribution and lognormal distribution provides closed form results.

When ξ_{jt} follows exponential distribution after change point τ

The distribution of the exponential random number is given as:

$$f(x,\lambda) = \begin{cases}\lambda e^{-\lambda x} & \text{if } x \geq 0\\ 0 & \text{if } x < 0\end{cases} \tag{4.7}$$

Here, it is worth noting that the mean of the distribution is kept constant before and after the degradation pattern changes ,i.e., $\mu_1 = \mu_2 =$ Constant. With known statistical relationship between exponential distribution and mean & variance of the data $\mu_2 = 1/\lambda, \sigma_2 = 1/\lambda$, the likelihood function of parameters σ, λ, τ can be obtained from above equation and written as follows:

$$L(\sigma,\lambda,\tau) = \prod_{t=1}^{\tau}\frac{1}{\sqrt{2\pi{\sigma_1}^2}}\exp\left(-\frac{\left(\xi_{jt}-\mu_1\right)^2}{2{\sigma_1}^2}\right)\times\prod_{t=\tau+1}^{T-1}\lambda e^{-\lambda\xi_{jt}} \tag{4.8}$$

Taking log on both side, we will obtain the following log likelihood expression

$$\begin{aligned}\text{logL}(\sigma_1,\lambda,\tau) = &-\sum_{\text{t=1}}^{\tau}\frac{1}{2}\ln\left(2\pi\right)-\sum_{t=1}^{\tau}\frac{1}{2}\ln\left({\sigma_1}^2\right)\\ &-\sum_{t=1}^{\tau}\frac{\left(\xi_j-\mu_1\right)^2}{2{\sigma_1}^2}-\sum_{t=\tau+1}^{T-1}ln(\lambda)-\sum_{t=\tau+1}^{T-1}\lambda\xi_{jt}\end{aligned} \tag{4.9}$$

$$\begin{aligned}\log L(\sigma_1,\sigma_2,\tau) = &-\sum_{t=1}^{\tau}\frac{1}{2}\ln(2\pi)-\sum_{t=1}^{\tau}\frac{1}{2}\ln\left({\sigma_1}^2\right)\\ &-\sum_{t=1}^{\tau}\frac{(\xi_j-\mu_1)^2}{2{\sigma_1}^2}-\sum_{t=\tau+1}^{T-1}ln\left(\frac{1}{\sigma_2}\right)-\sum_{t=\tau+1}^{T-1}\frac{\xi_{jt}}{\sigma_2}\end{aligned} \tag{4.10}$$

The maximum likelihood is achieved by differentiating the log likelihood with respect to parameter σ_2 and equating the terms to zero. This will provide the value of σ_2 as follows:

$$\sigma_2 = \frac{1}{T-\tau-1}\sum_{t=\tau+1}^{T-1}\xi_{jt} \tag{4.11}$$

When ξ_{jt} follows lognormal distribution after change point τ

The distribution of the lognormal random number is given as:

$$f(\lambda,\eta) = \frac{1}{\xi_{jt}\eta\sqrt{2\pi}}\exp\left(-\frac{(\ln(\xi_{jt})-\lambda)^2}{2\eta^2}\right) \tag{4.12}$$

Here, it is worth noting that the mean of the distribution is kept constant before and after the degradation pattern changes ,i.e., $\mu_1 = \mu_2 =$ Constant. With known statistical relationship between lognormal distribution and mean & variance of the data $\lambda_2 = \ln\left(\frac{\mu_2^2}{\sqrt{\mu_2^2+\sigma_2^2}}\right)$, and $\eta_2^2 = \ln\left(1+\frac{\sigma_2^2}{\mu_2^2}\right)$, the likelihood function of parameters $\sigma, \lambda, \eta, \tau$ can be obtained from above equation and written as follows:

$$\begin{aligned}L(\sigma,\lambda,\tau) = &\prod_{t=1}^{\tau}\frac{1}{\sqrt{2\pi{\sigma_1}^2}}\exp\left(-\frac{(\xi_{jt}-\mu_1)^2}{2{\sigma_1}^2}\right)\\ &\times\prod_{t=\tau+1}^{T-1}\frac{1}{\xi_{jt}\eta\sqrt{2\pi}}\exp\left(-\frac{(\ln(\xi_{jt})-\lambda)^2}{2\eta^2}\right)\end{aligned} \tag{4.13}$$

Taking log on both side and approximating the $\eta_2 \approx \frac{\sigma}{\mu}$, , we will obtain the following log likelihood expression

$$\begin{aligned}\log L(\mu_1,\sigma_1,\lambda_2,\eta_2,\tau) = &-\sum_{t=1}^{\tau}\frac{1}{2}\ln(2\pi)-\sum_{t=1}^{\tau}\frac{1}{2}\ln\left({\sigma_1}^2\right)-\sum_{t=1}^{\tau}\frac{(\xi_j-\mu_1)^2}{2{\sigma_1}^2}\\ &-\sum_{t=\tau+1}^{T-1}\frac{1}{2}\ln(2\pi)-\sum_{t=\tau+1}^{T-1}\frac{1}{2}\ln(\xi_{jt}\eta_2)\\ &-\sum_{t=\tau+1}^{T-1}\frac{(\ln(\xi_{jt})-\lambda_2)^2}{2\eta_2}\end{aligned} \tag{4.14}$$

$$\begin{aligned}\log \mathrm{L}(\mu_1, \sigma_1, \mu_2, \sigma_2, \tau) = &-\sum_{t=1}^{\tau} \frac{1}{2} \ln(2\pi) - \sum_{t=1}^{\tau} \frac{1}{2} \ln\left({\sigma_1}^2\right) - \sum_{t=1}^{\tau} \frac{(\xi_j - \mu_1)^2}{2{\sigma_1}^2} \\ &- \sum_{t=\tau+1}^{T-1} \frac{1}{2} \ln(2\pi) - \sum_{t=\tau+1}^{T-1} \frac{1}{2} \ln\left(\frac{\xi_{jt}\sigma_2}{\mu_1}\right) \\ &- \sum_{t=\tau+1}^{T-1} \frac{\mu_2}{2\sigma_2} \left(\ln(\xi_{jt}) - \ln(\mu_2) - \frac{\sigma_2^2}{2\mu_2^2}\right)^2\end{aligned} \tag{4.15}$$

The maximum likelihood is achieved by differentiating the log likelihood with respect to parameter σ_2 and equating the terms to zero. The solution of the following equation will provide the value of σ_2:

$$-\sum_{t=\tau+1}^{T-1} \frac{1}{\sigma_2} - \frac{\partial}{\partial \sigma_2} \sum_{t=\tau+1}^{T-1} \frac{\mu_2}{2\sigma_2} \left(\ln(\xi_{jt}) - \ln(\mu_2) - \frac{\sigma_2^2}{2\mu_2^2}\right)^2 \tag{4.16}$$

Through proper choice of posterior mean, we can reduce the above expression as follows:

$$\sigma_2 = \frac{1}{T - \tau - 1} \sum_{t=\tau+1}^{T-1} \left(\ln(\xi_{jt})\right)^2 \tag{4.17}$$

Once σ_2 is estimated, this is substituted to calculate the profile log-likelihood value $\mathcal{L}(\sigma_2, t)$ for any time t. The process is repeated for all $t (1 < t < T)$ and profile log-likelihoods $\mathcal{L}(\sigma_2, t)$; corresponding times t are stored in an array. Finally, the value of t that maximizes $\mathcal{L}(\sigma_2, t)$ is selected as the optimum change-point location $\hat{\tau}$. Mathematically, it can be written

$$\hat{\tau} = \arg\max \log \{\mathrm{L}(\sigma_2, \mathrm{t})\} \tag{4.18}$$

In certain scenarios, when the distribution becomes increasingly complex, estimating the likelihood becomes computationally intractable, necessitating sampling approaches that may not be computationally efficient. In such cases, while potentially less reliable, alternative data-driven methods can be employed for change point detection. One such approach involves monitoring sudden changes in autoregression coefficients. For the scope of this thesis, we restrict our discussion to scenarios where the degradation difference adheres to a Gaussian distribution both before and after the change point. Empirical observations indicate that a shift in the degradation pattern often manifests as a sharp change in the likelihood curve, with the local maximum of this curve serving as a suitable approximation of the change point. Analyzing different sections of the likelihood curve can reveal insights into the nature of the various degradation patterns. To illustrate this concept, consider a degradation curve that is both continuously increasing and bilinear. Figure 4.1 showcases the identification of the change point and the corresponding degradation patterns using the aforementioned approach.

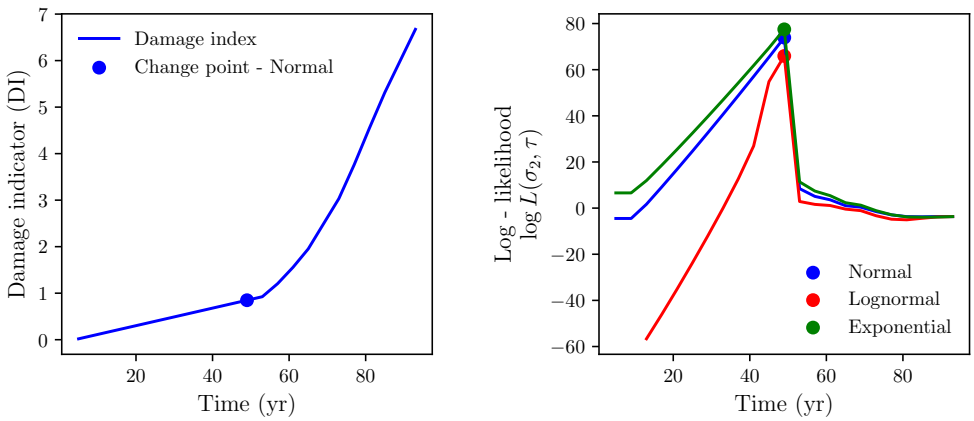


Figure 4.1: Demonstration of change point identification (a) Damage indicator having single change point, (b) log-likelihood profile showing the change point due to presence of local maximum

Now, as the best basis function is selected after occurrence of change point, the parameters of basis functions are optimized along with the hyperparameters of the basis function GPR (BGPR) through entropy assisted genetic algorithm. Based on the process described above, this study has compared the results of two approached in detail, namely,

1. BGPR: Gaussian process regression with basis function optimized using conventional methods.

2. EGA-BGPR: Gaussian process regression with basis function optimized using entropy assisted genetic algorithm.

Moreover, the results with conventional GPR (mean function trained on dataset prior optimization) are also presented where-ever needed. The details of the complete workflow for the identification of damage and change point along with extrapolation and prognosis through EGA-BGPR is presented in Figure 4.2.

4.4 Proposed algorithm

The following step-by-step procedure outlines the proposed framework for adapting mean functions and conducting prognosis with EGA-BGPR:

Step 1: Development of damage index

- Sensor data collection: The relevant sensor and simulated data (e.g., vibration, strains, energy dissipated) are collected from the monitored system.
- Feature extraction: The signal processing techniques like empirical mode decomposition (EMD), time varying auto regression (TVAR) are applied to extract meaningful features that represent the system's health condition.
- Damage index calculation: The statistical Bhattacharyya distance helps in quantifying the extent of damage using extracted features called the damage index (DI).

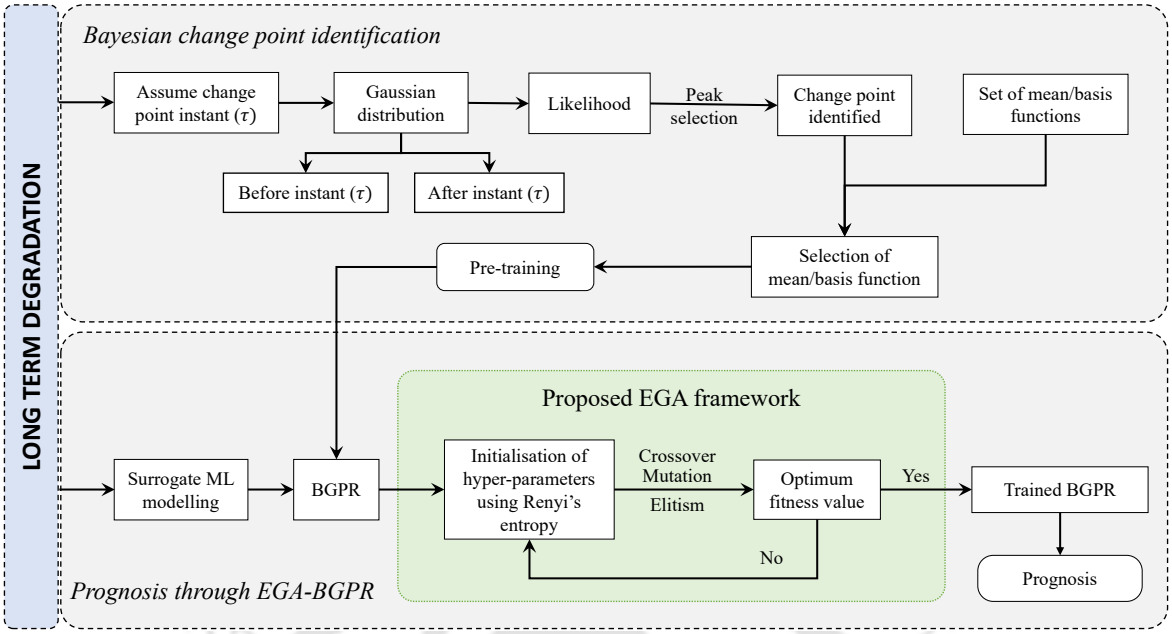


Figure 4.2: Illustrative methodology to develop the damage index and carrying out prognosis through EGA-BGPR to predict the remaining useful life.

Step 2: Initialization of adaptive mean/basis function

- Pre-training: Select a suitable mean/basis function for BGPR (e.g., linear, polynomial, exponential) and pre-train its coefficients using the initial observations of the damage index. This captures the initial degradation trend.
- Optimization using EGA: Utilize the EGA to optimize the hyperparameters (e.g., kernel parameters, noise variance) of the BGPR model as well as fine tune the coefficient of mean/basis function. EGA's ability to avoid local minima ensures a more robust optimization process.

Step 3: Bayesian change point monitoring

- Online Monitoring: There is continuous monitoring of the incoming damage index data using a Bayesian change point detection algorithm. This algorithm detects significant shifts in the damage pattern, signalling potential occurrence of change point.

Step 4: Updation of adaptive mean function (Triggered by change point)

When a change point is detected, the mean function is updated to reflect the new degradation trend. This involves:

- Selecting a new mean function form (e.g., switching from linear to polynomial)
- Retraining the mean function coefficients using the data after the change point
- Re-optimizing the BGPR hyperparameters and fine-tuning of mean function coefficients using EGA

Step 5: RUL prediction and extrapolation

- BGPR prediction: The trained BGPR model with updated mean function is used to predict the future evolution of the damage index.
- Extrapolation: The predicted damage index trajectory beyond the observed data is achieved to estimate the RUL. This involves selecting an appropriate threshold for the damage index that corresponds to failure.
- RUL Estimation: Calculate the RUL as the time remaining until the predicted damage index reaches the failure threshold. The mean and variance of the estimated RUL will deduce the probability of failure of system.

Step 6: Continuous monitoring and update

- Repeat: The damage index is continue monitoring and the steps 3-5 are repeated whenever a new change point is detected. This ensures that the model remains adaptive to the evolving multistep degradation behaviour.

4.5 Numerical case study

The chapter presents a numerical case study in which a hysteresis system is modelled through uniaxial Bouc-Wen model. The Bouc-Wen model is very common under structural health monitoring for modelling various types of non-linear hysteresis systems like beam column joints, soil structure interaction, seismic protection devices called isolators and dampers. In present case, the seismic isolator is simulated for unidirectional excitation through this phenomenological model which has the restoring force given by: $f_r(t) = \alpha k u(t) + (1-\alpha) Q_y z(t)$, where, u is the input displacement, $k = F_y/u_y$ is the elastic stiffness, F_y is the yield force, u_y is the yield displacement, Q_y is the characteristic strength which is kept equal to $0.05 \times mass \times g$, g is the gravitational constant and is the ratio of post elastic stiffness to the elastic stiffness. It can also be considered as the ratio of linearity to the non-linearity in the system. The evolution parameter $z\,(t)$ is known as hysteresis displacement and is given by:

$$\begin{aligned} \dot{z} = \frac{1}{D_x} \Big(& A\dot{u}\,(t) - \beta \left|\dot{u}\,(t)\right| \left|z\,(t)\right|^{n-1} z\,(t) \\ & - \gamma \dot{u}\,(t) \left|z\,(t)\right|^{n} \Big) \end{aligned} \tag{4.19}$$

where A, β, γ and n are the hysteresis shape parameters and $D_x = Q_y/k$ is the yield displacement (Wen, 1976; Bouc, 1967). Thus, equation of motion with input excitations $\ddot{u}_g$ is written as:

$$m\ddot{u} + c\dot{u} + \alpha k u + (1-\alpha) Q_y z = -m\ddot{u}_g \tag{4.20}$$

Consider an isolation system with initial elastic stiffness $k_i = 1392000\mathrm{N/m}$ that varies with

the assumed degradation law for the long-term degradation as follows:

$$k = \begin{cases} 1.00k_i \times \exp(-0.05 \times 10^{-2} \times T) & 0 < T < 25 \\ 0.98k_i \times \exp(-0.20 \times 10^{-2} \times T) & 25 \leq T < 75 \\ 0.94k_i \times \exp(-0.50 \times 10^{-2} \times T) & 75 \leq T < 100 \end{cases} \tag{4.21}$$

Where T is in years. A visual interpretation of Bouc-Wen hysteresis model for simulating isolator is shown in Figure 5.19 (a). Further, the variation of the stiffness over the 100 years can be visualized from Figure 5.19 (b). The mass is 6800 kg with a damping of 3740 Ns/m. The values of the parameters are $A = 1, \alpha = 0.15, \beta = 1.1, \gamma = 0.2, n = 2$.

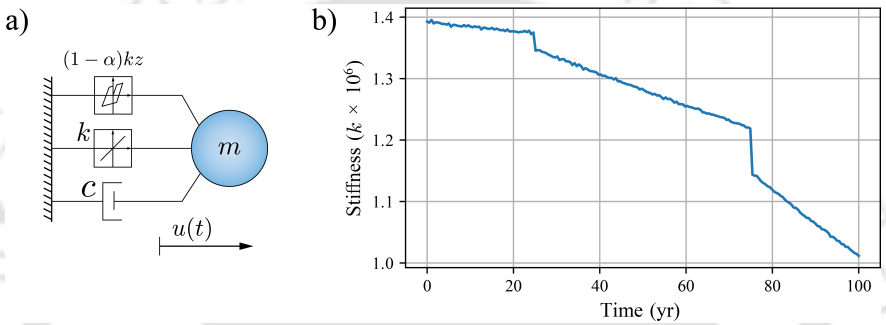


Figure 4.3: (a) Visual representation of the Bouc-Wen model, (b) Long-term stiffness degradation considered in the numerical example.

In this numerical study, the isolation system with degrading stiffness is excited with a 20 second low amplitude ambient noise every year and the system response are obtained. Thus, a total of 100 simulation were carried out, and damage index is developed. The typical details of the input excitation and the output responses at any time point is shown in Figure 4.4.

4.5.1 Development of damage index

The restoring force response function is processes to develop the damage index. Before proceeding, the very first step is to filter the response and remove the unwanted frequencies and noise from the signal. There are several methods to remove noise from signal like principal orthogonal decomposition (POD), spectral analysis, empirical mode decomposition (EMD) and more. The EMD has the practical advantage that it can work one single sensor measurement data. It decomposes the various components of the different frequencies of signal to provide a mono-component signal (Huang et al., 1998). The extracted mono-components are called Intrinsic mode functions (IMFs). The fundamental of EMD is that it filters out functions for a time domain signal, which forms a complete and orthogonal basis for the original signal locally. The response signal $u(t)$ is

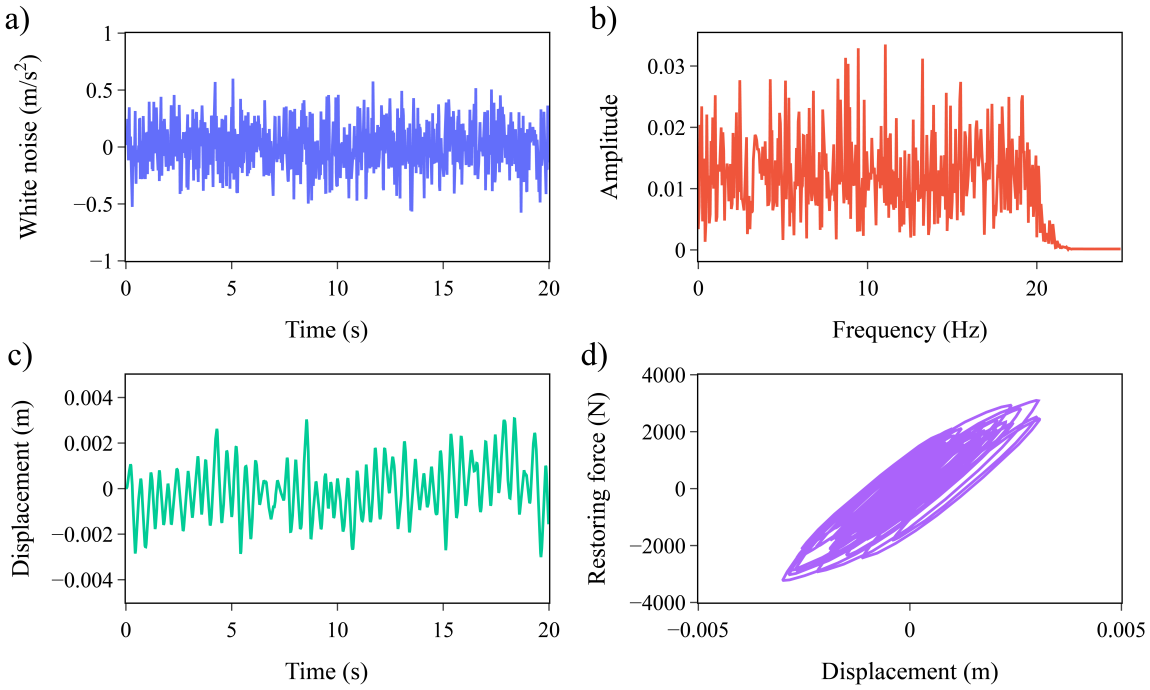


Figure 4.4: Sample system responses (a) white noise excitation input, (b) frequency bandwidth of white noise, (c) displacement response, and (d) hysteresis response.

thus represented as a summation of true signal $s(t)$ and white noise $w(t)$.

$$u\left(t\right) = s\left(t\right) + w\left(t\right) \tag{4.22}$$

$$u\left(t\right) = \sum_{i=1}^{j} \mathrm{IMF}_i + \mathrm{Residual} \tag{4.23}$$

Where, j is the total number of IMFs. The IMF that has the maximum correlation with the original signal is selected as dominant IMF and further forwarded for time varying auto regression (TVAR) modelling. The leverage TVAR is that it provides a cluster of the coefficients that undergo significant change due to various amplitude of damage (Mosavi et al., 2012; Carden and Fanning, 2004). Further, due to varying coefficients, it is suitable for non-stationary signals. Let the dominant IMF be represented as $y(t)$ and $n \in [1; N]$ is the discrete time index with N number of sample points. It is important to analyze the signal for the appropriate model order because higher model order cause high computation cost and lower model order will lead loss of important features. Here, it is selected as one for which the curve of embedding dimension with respect to normalized singular value comes close to zero (Shin et al., 2003). Thus, the TVAR model with the model order p is represented as

$$y\left(n\right) = \sum_{k=1}^{p} a_k y\left(n-k\right) + v\left(n\right) \tag{4.24}$$

where $v\left(n\right)$ is the noise or error term, and a_k are the time-varying coefficients estimated

through adaptive Kalman filter. The coefficients of the TVAR model are represented in the state space equations as follows:

$$a(n) = a(n-1) + w(n) \tag{4.25}$$

$$y(n) = C(n)\, a(n) + v(n) \tag{4.26}$$

where, $a(n) = [a_1(n), a_2(n), \ldots, a_p(n)]$ is the TVAR coefficients vector, $w(n)$ is the process noise with covariance σ_w^2, $v(n)$ is the measurement noise with covariance σ_v^2 , $C(n) = [y(n-1), \ldots, y(n-p)]$ is the system measurement matrix with discrete n steps. The Kalman gain calculation and update equation comprises prediction and correction step which is not presented here for brevity. Once, the modeling is complete, the coefficient matrix is obtained as: $a(N \mid N) = [a_1(N \mid N), \ldots, a_p(N \mid N)]$. These coefficients matrix will form the cluster in n-dimensional subspace. The variations in the shape, orientation, spread, of the coefficient cluster will signify the change in the system. These variations do not essentially imply the existence of damage unless their magnitude is quantified using damage-sensitive features like statistical distance in this case. As the stiffness degrades and damage evolve in the system, the states of these coefficient clusters will exhibit the topological and statistical changes. There might be a possibility that the damage evolution is slow, and it leads to slight migrations of the are clusters leading to the associated overlapping and interweaving of the feature vectors. This intervention might give poor results. Therefore, a filtering technique called local outlier factor (LOF) is used here to remove the anomalies and segregate the clusters (Breunig et al., 2000). The results before and after removing outliers from TVAR coefficients are presented in Figure 4.5.

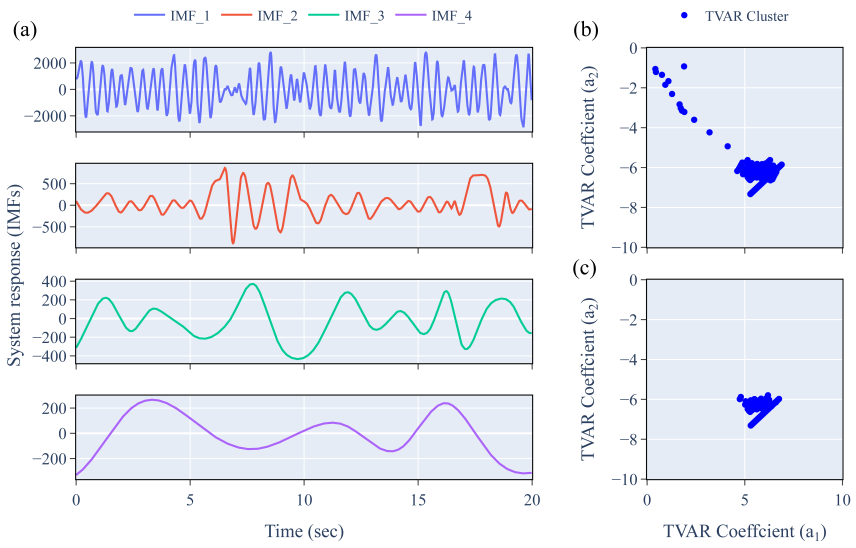


Figure 4.5: (a) IMFs of the filtered response, (b) TVAR coefficient cluster before removal of outlier, and (c) TVAR coefficient cluster after removal of outlier.

The TVAR clusters are prepared for the 100 year and outliers are removed. The different clusters of all years are combined presented in Figure 4.6(a). The very first year cluster is considered as of healthy state and the Bhattacharyya distance of all other successive clusters are calculated with respect to the health state cluster. The Bhattacharyya distance (BD) is the statistical measure of the difference between the distributions of two classes. The present research deals with evaluation of Bhattacharyya distance to quantify the separation of two distinct coefficient clusters. These coefficient clusters are of multidimensional subspace, for a multidimensional BD, the mean and variance are replaced by mean vector and covariance matrices (Guorong et al., 1996).

$$B_D(\beta_1,\beta_2) = \frac{1}{2}\ln\left\{\frac{\left|\frac{1}{2}\left(COV_{\beta_1}+COV_{\beta_2}\right)\right|}{\sqrt{|COV_{\beta_1}||COV_{\beta_2}|}}\right\} + \frac{1}{4}(\mu_{\beta_1}-\mu_{\beta_2})^T\left[COV_{\beta_1}+COV_{\beta_2}\right]^{-1}(\mu_{\beta_1}-\mu_{\beta_2}) \tag{4.27}$$

where COV_{β_1} and COV_{β_2} are the covariance and μ_{β_1} and μ_{β_2} are mean of distribution β_1 and β_2 respectively. This statistical distance is referred here as damage index and is presented in Figure 4.6 (b). It can be observed that there are several change points and multiple damage patterns over the year due to the mentioned stiffness degradation. To choose the correct mean function for the BGPR model and prognosis, it is important to identify the change point that segregate the data into two parts. The observations after change point are selected for the pre-training of the mean function. The results of the change point identification and selection of mean function are discussed in subsequent sections.

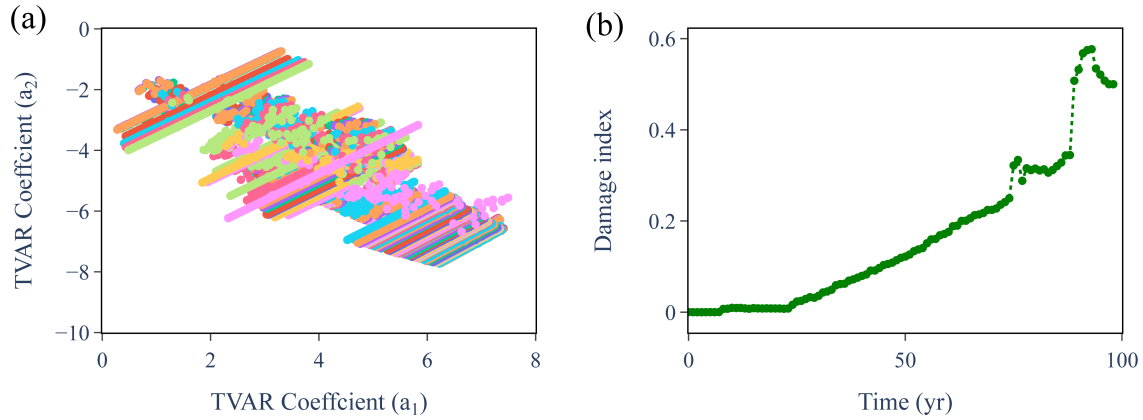


Figure 4.6: (a) TVAR coefficient clusters of all different years , and (b) Bhattacharyya distance as damage index.

4.5.2 Surrogate modelling, change point detection and prognosis

Figure 4.6 (b) demonstrate the Bhattacharyya distance as the damage index that is modelled through BGPR for the prognosis at difference time interval. In a long term prospective, this damage index is referred as degradation indicator. It has changed the

trajectory multiple time depending on the magnitude of damage occurred in the structure. The local fluctuations are attributed to various uncertainties associated with the development of damage index. Thus, the points before the change of trajectory are irrelevant for the training of BGPR. It is important to note here that the data points between the current time instant and the previous instant of change points will be considered for the selection and pre-training of mean function. This same dataset will further be considered for the training of hyperparameters of the BGPR through entropy assisted genetic algorithm (EGA-BGPR). Therefore, it is important to identify the change point accurately.

In this case study, the successive change in the degradation indicator is considered to be a part of Gaussian distribution. Thus, the Gaussian distribution changes as the change point occurs referring here as prior and posterior distribution. However, for simplicity, it is assumed that the mean of the distribution will not change due to change point and the mean of the posterior distribution is kept equal to the prior distribution. In this numerical example, for demonstration purpose, the log-likelihood profile is derived for the complete set of damage index and can be visualized in Figure 4.7 along with the damage index. It can be clearly observed that every local maximum of the log likelihood function is represented as a change point.

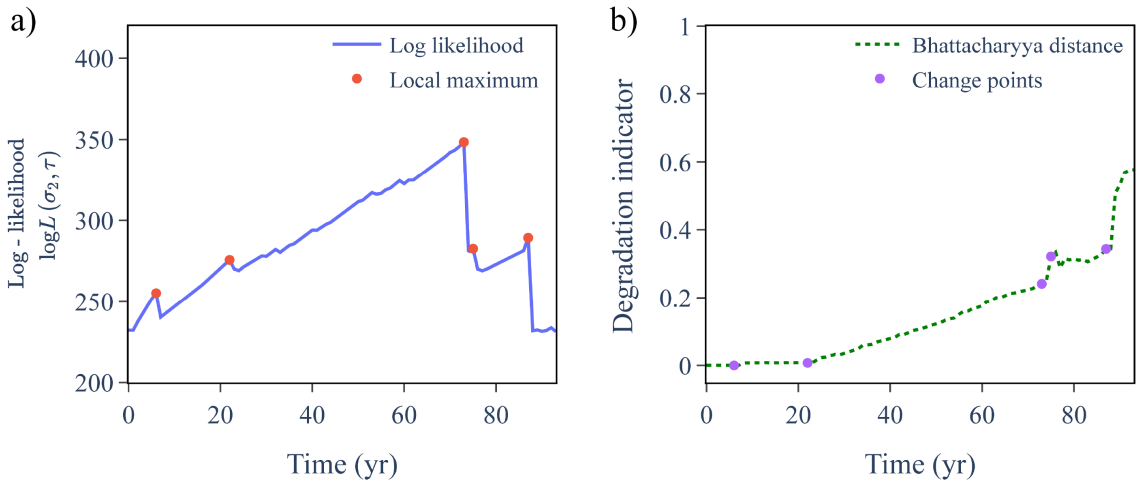


Figure 4.7: (a) log-likelihood profile showing the change point due to presence of various local maxima (b) degradation indicator showing multiple change points

Through careful analysis of the above curve, the years of occurrence of change points are 6, 22,73,75, 87 years. There are local maxima in the log-likelihood function between 22 and 73 years. However, these minima are likely insignificant due to the small changes in log-likelihood value. These can be attributed to minor fluctuations in the curve rather than representing true change points. Let us select the few-time instants for which the prognosis will be carried out, and the proposed algorithm will be validated. These time instants are, say 20th year, 60th year, 85th year and 93rd year. For, each time instant, the log likelihood

profile is checked, and change point is identified. The observations post change points are used for the pre-training of mean function. The pre-trained mean function coefficients are further fine-tuned along with the hyperparameters of the BGPR model through entropy assisted genetic algorithm (EGA). Thus, the Log-likelihood function and observations 20th year, 60th year, 85th year and 93rd year of the prior training are presented in Figure 4.8.

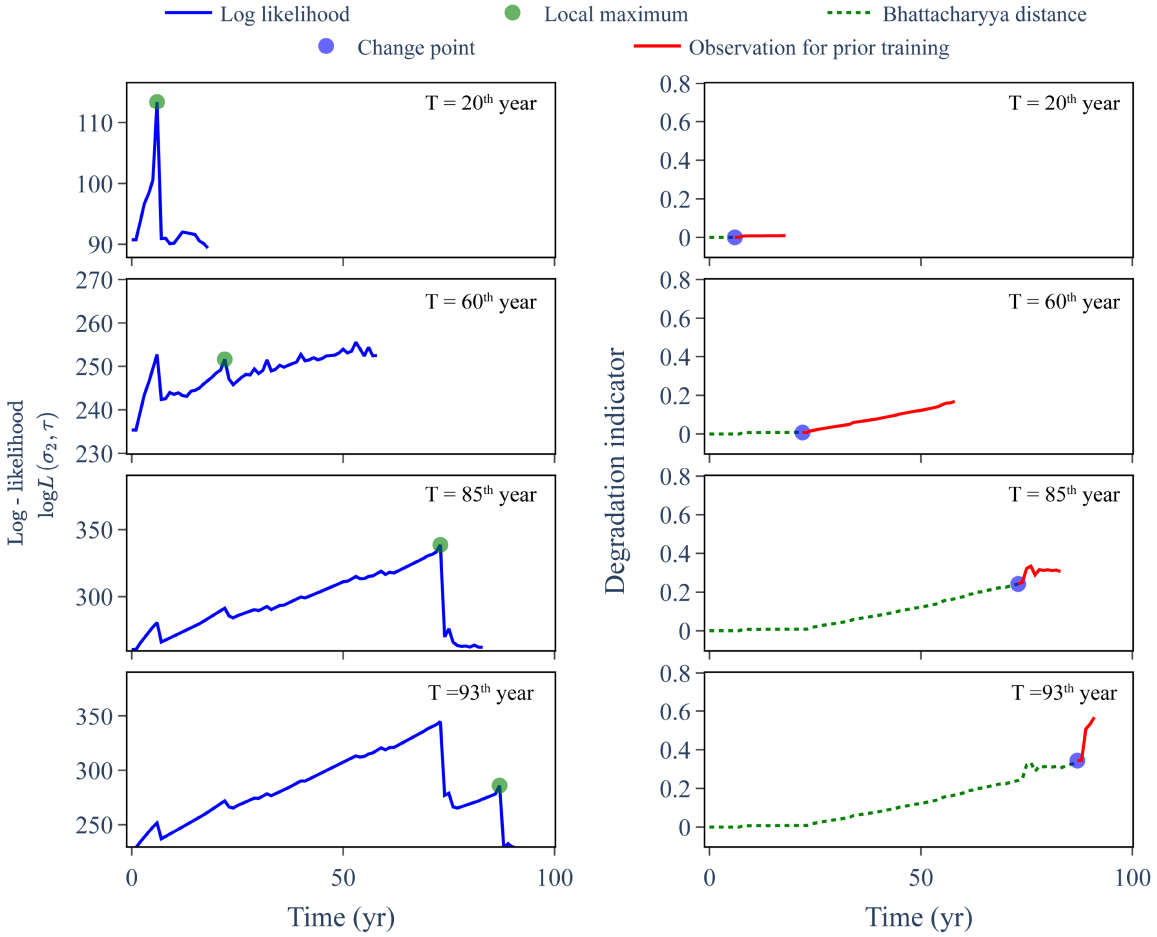


Figure 4.8: Log likelihood profiles and change points observed at different time instant of prognosis

As demonstrated earlier, GPR performance in extrapolation cases significantly deteriorates without an optimized mean function. This necessitates incorporating a mean function for accurate forecasting, leading to the development of basis function Gaussian Process Regression (BGPR). Identifying the appropriate form of the mean function is crucial for BGPR's effectiveness. To achieve this, we fit various candidate mean functions to the dataset before optimization. The degree and parameters of the basis function are then selected based on these preliminary fits and subsequently optimized alongside the BGPR hyperparameters using the Entropy-Assisted Genetic Algorithm (EGA). Table 4.1 provides details on the considered mean functions and their corresponding coefficients.

Therefore, the selection of mean functions determines the appropriate degree of the polynomial basis function for both the conventionally optimized BGPR and the EGA-BGPR

Table 4.1: Details of the obtained mean/basis function after observing change point for application in BGPR, optimized through EGA.

Time instant	Degree	Prior mean function	Constant
20^{th} year	Linear	$y = a + bx$	$a = -0.00012,\ b = 0.00056$
60^{th} year	Quadratic	$y = a + bx + cx^2$	$a = -0.05960,\ b = 0.00268$ $c = 0.00002$
85^{th} year	Cubic	$y = a + bx + cx^2 + dx^3$	$a = -245.40019,\ b = 9.29621$ $c = -0.11715,\ d = 0.00049$
93^{rd} year	Linear	$y = a + bx$	$a = -5.05735,\ b = 0.06152$

approaches. BGPR serves as the surrogate model for the damage index, enabling prognosis based on this index. Table 4.2 reveals that both conventionally optimized BGPR and EGA-BGPR achieve high likelihood values. While the results from both methods are comparable, EGA demonstrates notable improvements ranging from approximately 5% to 40% in various cases.

Table 4.2: Improvement achieved using entropy assisted genetic (EGA) algorithm in maximum likelihood value during optimization of BGPR

Time instant	Degree of polynomial	Maximum likelihood value BGPR	EGA-BGPR
20^{th} year	Linear	48.302	50.871
60^{th} year	Quadratic	93.849	134.271
85^{th} year	Cubic	13.974	15.514
93^{rd} year	Linear	0.002	0.679

The prediction results are provided in Figure 4.9 and Figure 4.10. It can be well observed that with the help of mean function, extrapolation of the GPR can be possible and a well-defined curve for the prognosis can be obtained. Furthermore, in the conventional GPR, as test points move farther than the training points, the variance tends to be appeared as zero because it is calculated on normalized data. This creates an issue and thus prediction appears to be as mean function only, in case of conventional GPR. In the BGPR, the variance appears to be very high when optimized conventionally, however, EGA works far better in reducing the variance and provides optimum results.

Before proceeding to the RUL estimation, the effectiveness of use of Renyi's entropy in EGA-BGPR is presented. The initialization of the population is the crucial step in the process of genetic algorithm. There are possibilities that the population itself doesn't contain the solution or not uniformly generated so that GA can search for it. In such an circumstance, the genetic algorithm stacks and return to the local maxima rather than converging globally. Further, during encoding process, if unnecessary large bitstring are used for the encoding,

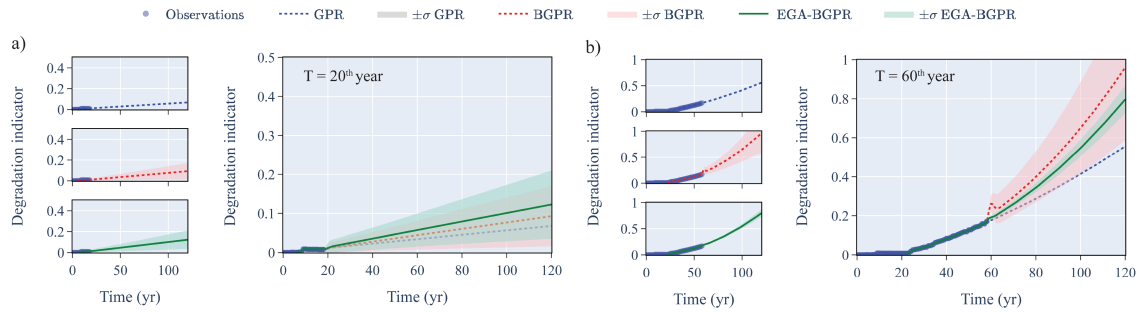


Figure 4.9: Prediction through GPR, basis-GPR, EGA-BGPR at different time instant (a) 20th year (Linear), and (b) 60th year (Quadratic)

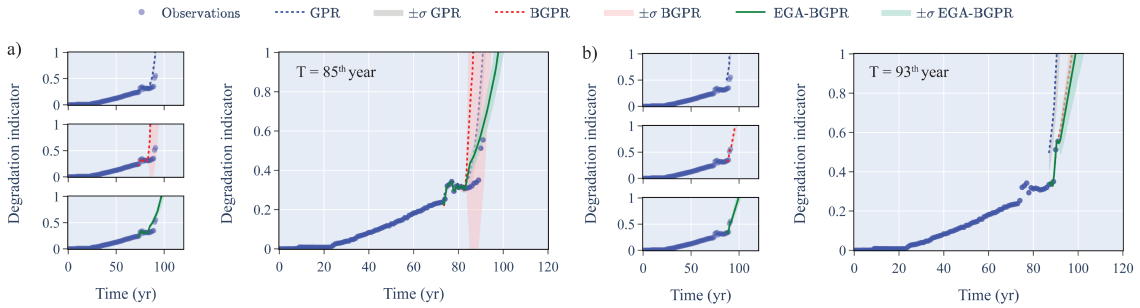


Figure 4.10: Prediction through GPR, basis-GPR, EGA-BGPR at different time instant (a) 85th year (Cubic), and (b) 93rd year (Linear)

they can give solutions but are computationally exhaustive. Thus, to protect the genetic algorithm to fall for local minima/maxima, the initial population is generated by maximizing the Renyi entropy. In this work, binary encoding is carried out in genetic algorithm for the hyperparameters, the size of the bitstring defines the data and its corresponding entropy, the higher the size of bitstring, the better is the resolution of data and more computational time is required. Here, four different types of bitstring sizes corresponding to different Renyi's entropy are selected, i.e., 4 bits, 8 bits, 12 bits and 16 bits and population is generated for each size. The results in Figure 4.11 shows that the maximum fitness value is achieved 16 bitstring in case (a) and 12 bitstring in case (b). It also shows that effect of selection of bitstring on the local and global convergence of genetic algorithm.

4.5.3 Efficacy of the proposed EGA-BGPR over conventional BGPR

To validate the efficacy of the proposed EGA-BGPR method compared to conventional BGPR, we considered six distinct prognosis scenarios: (a) 20 years, (b) 40 years, (c) 60 years, (d) 85 years (with training data up to the change point at 75 years), (e) 85 years (with training data up to 60 years), and (f) 93 years. For consistency, a quadratic mean function was used in all cases. We examined the effect of entropy maximization on BGPR optimization through EGA by considering four different bitstring lengths (4, 8, 12, and 16 bits). The results were compared with conventional BGPR using three metrics: fitness value (log-likelihood), RMSE, and relative percentage error. Figure 4.12 illustrates the prognosis results for all six cases,

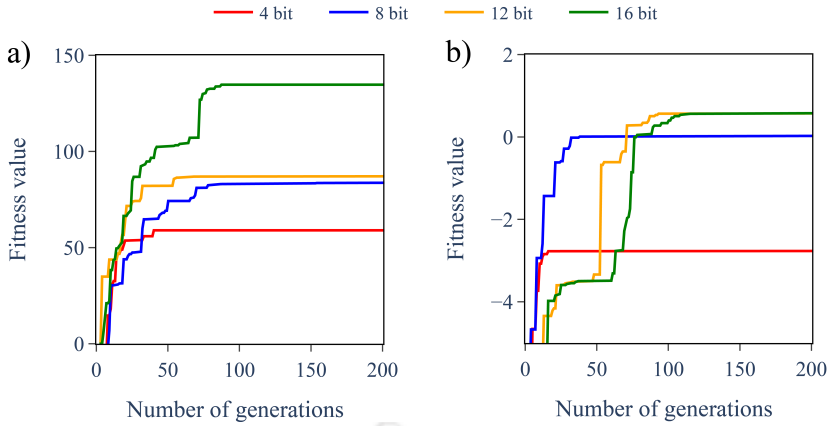


Figure 4.11: Fitness value obtained through EGA using different bitstring size for (a) 60th year (Quadratic), and (b) 85th year (Cubic)

while Table 6.1 provides a detailed comparison between EGA-BGPR and conventionally optimized BGPR. Moreover, the RMSE metrics is calculated by comparing the prognosis results for next 20 years with the ground truth. The EGA-BGPR consistently achieves higher log-likelihood values (indicating better fitness) and lower RMSE values (indicating more accurate predictions) than conventional optimized BGPR model. Furthermore, the relative percentage error is consistently lower for EGA-BGPR, underscoring its superior performance in both optimization and variance reduction.

Table 4.3: Comparison of the results for numerical case study, obtained through entropy assisted genetic algorithm in terms of fitness value (maximum likelihood value achieved), rmse error and relative percentage error

Metrics	Algorithm	Case (a) 20 years	Case (b) 40 years	Case (c) 60 years	Case (d) 85 years	Case (e) 85 years	Case (f) 93 years
Fitness value	BGPR	48.302	55.168	93.849	50.422	48.122	26.645
	EGA-BGPR	50.871	58.415	134.271	50.862	64.414	40.317
RMSE	BGPR	0.0096	0.0079	0.007	0.0257	0.0257	0.1925
	EGA-BGPR	0.0092	0.0042	0.0048	0.0256	0.0212	0.1463
Relative error (%)	BGPR	9.577	22.204	6.910	11.316	11.293	5.336
	EGA-BGPR	8.170	11.932	4.588	11.293	7.111	3.509

Further, it can be seen from Figure 4.12 that the best results are coming for the EGA-BGPR algorithm when encoding is done in high bits which corresponds to higher entropy. For example, in case (a), the BGPR along with 4 bit and 8 bit EGA optimization fails to provide a satisfactory quadratic curve for prognosis. However, when maximum entropy achieved for 16 bit, the results are found to be satisfactory. The similar observations can be found for the other cases as well. In case (b) and (c), the BGPR results are close to the 16 bit EGA-BGPR results , however, when compared with ground truth for next 20 years (refer Table 4.5), EGA-BGPR results are found to be more promising. In case (d) and (e) where

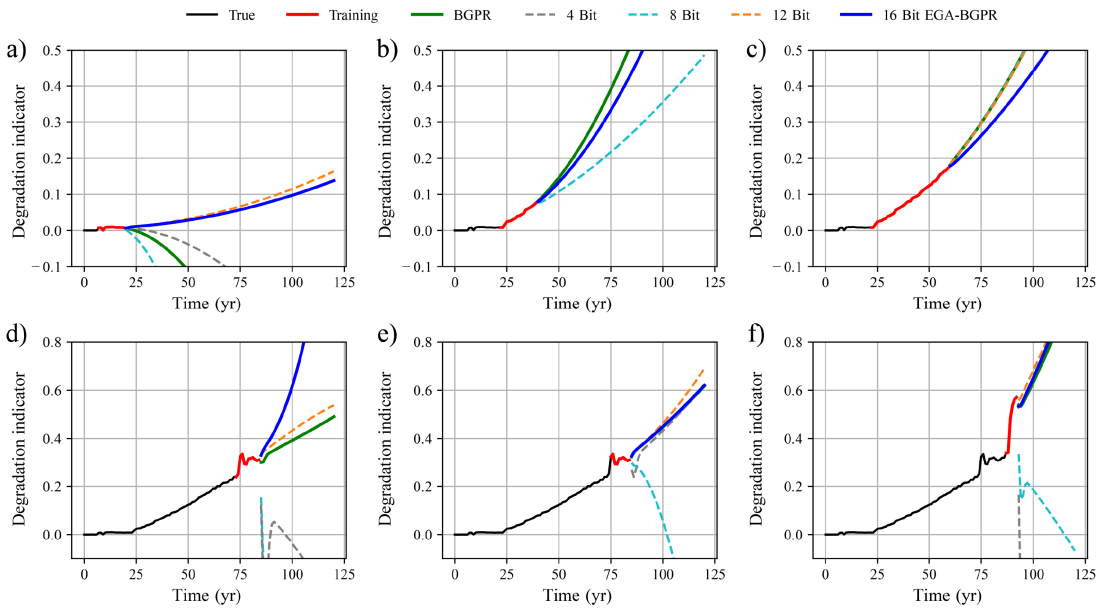


Figure 4.12: Comparison of prognosis achieved through EGA-GPR using different sizes of bit string along with conventional BGPR approach for 6 different cases

training of BGPR is challenging due to sparse training set, the perfect solution is given by cubic polynomial as provided earlier. However, for comparison, we restrict here with quadratic basis function. The results of EGA for case (d) and (e) through quadratic basis function are found to be more accurate. Similarly, trend is observed for case (f). Moreover, the performance of EGA-GPR for different sizes of bit string for the creation of encoded population can be analyzed by observing the maximum fitness value achieved for the BGPR surrogate model through EGA as given in Table 4.4. Similarly, the comparison of RMSE values for EGA-GPR having different sizes of bitstring for the creation of encoded population with the conventional approach of BGPR is presented in Table 4.5

Table 4.4: Comaprison of fitness value (maximum likelihood value) achieved for EGA-GPR for different sizes of bit string along with conventional BGPR approach

Algorithm		Case (a) 20 years	Case (b) 40 years	Case (c) 60 years	Case (d) 85 years	Case (e) 85 years	Case (f) 93 years
EGA-BGPR	4 Bit	40.224	55.162	55.922	27.738	25.618	1.6935
	8 Bit	47.521	55.192	83.626	34.004	40.317	15.239
	12 Bit	48.263	55.261	87.215	46.510	56.677	20.762
	16 Bit	50.871	58.415	134.271	50.862	64.414	40.317
BGPR		48.302	55.168	93.849	50.422	48.122	26.645

4.5.4 Calculation of remaining useful life (RUL)

The final step is to use the predicted mean and variance vector for estimating the remaining useful life (RUL). For this purpose, the results of the proposed EGA-BGPR methods are utilized. RUL is defined as the time span till the system fails or crosses a certain threshold

Table 4.5: Comparison of RMSE values for EGA-GPR having different sizes of bit string along with conventional BGPR approach

Algorithm		Case (a) 20 years	Case (b) 40 years	Case (c) 60 years	Case (d) 85 years	Case (e) 85 years	Case (f) 93 years
EGA-BGPR	4 Bit	0.0131	0.0082	0.0047	0.0346	0.0172	0.3043
	8 Bit	0.0284	0.0041	0.0049	0.0752	0.0231	0.0321
	12 Bit	0.0093	0.0083	0.0060	0.0266	0.118	0.0232
	16 Bit	0.0092	0.0079	0.0047	0.0256	0.0201	0.1092
BGPR		0.0096	0.0082	0.0072	0.0257	0.0301	0.1463

to become non-operational. Let the damage index DI_n after 'n' number of cycles, then the surrogate model for the fatigue damage index is represented as $DI_n = h(n;\theta)$, $h(\cdot)$ is the functional form used to simulate the model DI_n after n number of cycles. The parameter θ is used to describe the model. The remaining useful life is the path-dependent process, after k number of cycles is defines as L_k provided with the observation time history $DI_{1,}, DI_2, DI_3, \ldots\ldots, DI_k$, given by Si et al. (2013)

$$\text{RUL} = L_k = \inf\{l : DI_{k+l} > \xi_T | DI_{1,}, DI_2, DI_3, \ldots\ldots, DI_k < \xi_T\} \tag{4.28}$$

In the simple form, if the μ_{k+l} is the predicted mean and σ_{k+l} is the predicted variance at any time instant, then the RUL at that time instant with threshold value ξ_T is given by the cumulative density function of normal distribution as follows:

$$L_k = \Phi\left(\frac{\mu_{k+l} - \xi_T}{\sigma_{k+l}}\right) \tag{4.29}$$

The estimated RUL with the failure probability for the current numerical study at different time instant is presented in Figure 4.13.

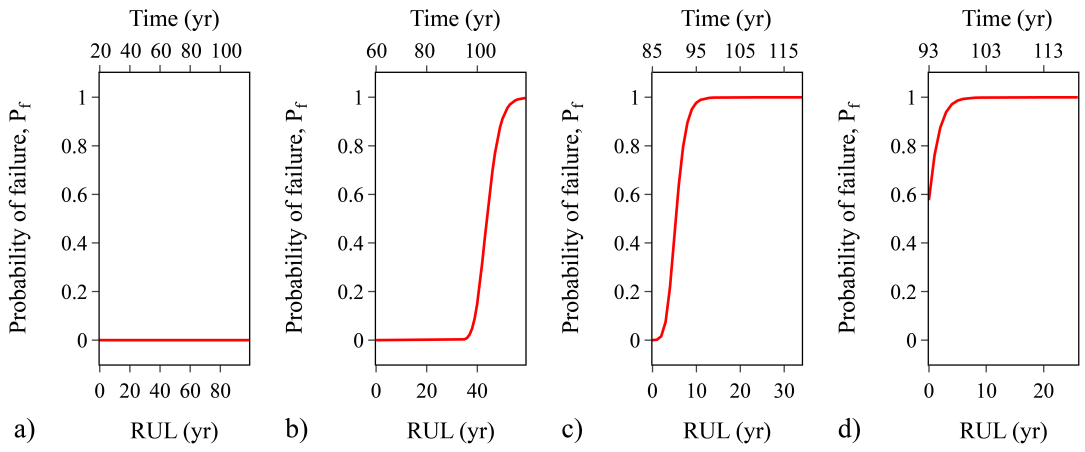


Figure 4.13: Estimation of failure probability with RUL at different time instant (a) 20th year (Linear), (b) 60th year (Quadratic), (c) 85th year (Cubic), and (d) 93rd year (Linear)

4.6 Conclusion

This chapter presents a robust, data-driven framework for system prognosis by integrating Gaussian Process Regression (GPR) with an Entropy-Assisted Genetic Algorithm (EGA). The key contributions can be highlighted as follows:

1. **Enhanced hyperparameter optimization:** By exploiting Renyi's entropy, EGA effectively overcomes the limitations of local minima in traditional GPR hyperparameter tuning, leading to superior global optimization.
2. **Accurate extrapolation and adaptability:** Using basis functions and an ensemble-based approach for mean function determination, the proposed EGA-BGPR framework achieves reliable extrapolation, critical for prognosis applications. Furthermore, the embedded change-point detection enables timely re-evaluation of hyperparameters and model adaptation, ensuring accurate predictions even as the system's degradation trajectory evolves.
3. **Validation and superiority:** The numerical case study demonstrate the framework's effectiveness and its advantage over traditional methods, as evidenced by metrics like maximum likelihood, RMSE, and relative percentage error.

While the EGA-BGPR framework has demonstrated significant potential for prognosis, its full capabilities can be better realized when applied to larger, more complex datasets and the associated computational challenges. To showcase the framework's adaptability and robustness in real-world scenarios, the subsequent chapters will delve into two distinct case studies: a fatigue experiment on elastomeric rubber bearings and the monitoring of crack width in aluminum lap joints.

Each case study presents unique challenges in developing reliable damage indices and identifying relevant system parameters from noisy or limited data, a common occurrence in practical SHM applications. To address these challenges, tailored diagnostic methodologies will be introduced, specifically designed to extract meaningful information from the available data. These methodologies will then be seamlessly integrated with the EGA-BGPR framework to provide accurate and reliable prognosis, ultimately demonstrating the framework's versatility and effectiveness in diverse engineering contexts.

Chapter 5

Two-stage constrained unscented Kalman filtering for system identification and prognosis: A case study on elastomeric rubber bearing

5.1 Overview

This chapter investigates the application of a two-stage constrained unscented Kalman filtering (CUKF) methodology for system identification followed with the prognosis of degradation of elastomeric rubber bearings through proposed framework. Utilizing experimental data from a bearing subjected to varying operational conditions, this chapter proposes the two-stage CUKF to account the coupling behavior with limited measurements and accurately identify the bearing's dynamic characteristics. Subsequently, these identified parameters are integrated into the EGA-BGPR framework to carry out the prognosis. This case study serves as a practical demonstration of the integrated framework's effectiveness in a real-world scenario, showcasing its potential for enhancing the reliability and longevity of critical engineering components.

5.2 Background

The presence of unsymmetrical stiffness distribution in many buildings leads to excessive deformation under seismic loads due to torsional coupling of translational and rotational behavior (Basu et al., 2012; Chopra, 2007; Fajfar, 2000). To reduce the seismic forces in order to overcome the adverse influence of unsymmetrical plan irregularity, seismically isolated structures are designed (Constantinou et al., 2007). These buildings often uses high damping rubber bearings (HDRBs) for isolation and energy dissipation (Naeim and Kelly, 1999; Constantinou et al., 2007). However, HDRB performance degrades over time due to factors such as heating and loading rate, and their behavior is fundamentally affected by bidirectional excitations (Kitayama and Constantinou, 2019; Constantinou et al., 2007; Kalpakidis and Constantinou, 2009; Yamamoto et al., 2012; Kato et al., 2015). Various studies have examined the dynamic response and torsional effects within base-isolated structures (Nagarajaiah et al., 1991, 1993; Nagarajaiah and Xiaohong, 2000).However, evaluating the

deterioration of base-isolated systems at the component level under bidirectional seismic excitation remains a complex challenge. The key challenges and chapter focus is presented as follows:

1. **Modeling complexity:** Capturing the true behavior of systems under bidirectional excitations demands models that can accurately represent hysteresis phenomena. The widely used Bouc-Wen model has been extended for biaxial behavior, allowing for more realistic simulations (Bouc, 1967; Wen, 1976; Baber and Noori, 1985; Foliente, 1995; Ismail et al., 2009; Kalkan and Graizer, 2007).

2. **System identification:** Estimating the governing parameters of biaxial hysteresis models adds complexity due to the interaction between orthogonal axes. This chapter focuses on the Constraint Unscented Kalman Filter (CUKF), a Bayesian filter that provides a robust framework for parameter identification in systems with coupled behavior (Arulampalam et al., 2002; Chen et al., 2003; Burkhart, 2019; Li and Wang, 2021; Calabrese et al., 2018; Yang and Ma, 2003; Sen and Bhattacharya, 2017; Gupta and Hauser, 2007).

3. **Deterioration detection:** The integration of deterioration and fatigue-crack growth models into the SHM framework is explored, enabling the identification of progressive damage within the base-isolated system (Tripura et al., 2020; Pamwani and Shelke, 2018; Yang and Nagarajaiah, 2014; Nagarajaiah and Yang, 2017)

4. **Prognosis:** The extracted system parameters are integrated into the proposed EGA-BGPR framework for the prognosis and predicting the degradation with increase in number of fatigue cycles.

This chapter adopts the phenomenological Bouc-Wen model for its accuracy and wide acceptance, extending it to a biaxial form. Morevoer, the suitability of modelling the seismic bearing using Bouc-Wen model is explored and discussed in Appendix D. This study proposes a system identification framework based on Bayesian filters, which have proven effective in recursive estimation problems (Arulampalam et al., 2002; Chen et al., 2003). Unscented Kalman filter (UKF) is widely used to estimate the uniaxial Bouc-Wen model parameters for different hysteresis systems in seismic control applications (Yang et al., 2006; Li and Wang, 2021; Calabrese et al., 2018; Hassani et al., 2014). The UKF generates sigma points which captures the mean and covariance of the posterior distribution of the system states. The operations are performed on sigma points and then transformed back for updating the states. Currently, to avoid the redundant mathematical solution, and achieving the solutions consistent with system physics, the variation of Kalman filter called constrain Kalman filter is used. Application of constraint EKF and constraint UKF (CUKF) for the uniaxial Bouc-Wen model can be found in the following

literature (Li and Wang, 2021; Calabrese et al., 2018; Yang and Ma, 2003; Sen and Bhattacharya, 2017; Gupta and Hauser, 2007). This chapter specifically focuses on the Constraint Unscented Kalman Filter (CUKF) for robust estimation of Bouc-Wen parameters in both deteriorating and non-deteriorating systems with complex biaxial behavior. After estimation of the system parameter, it is further processed to develop the degradation indicator which be forecasted using the presented EGA-BGPR framework.

5.3 Constraint unscented Kalman filter (CUKF)

The improved version of unscented Kalman Filter (UKF) for handling different physical constraints and restricting the estimated states within the bounded region for achieving faster convergence is constraint unscented Kalman filter (CUKF). This is more suitable for the highly nonlinear system like coupled biaxial Bouc-Wen to get the consistent solution as there can be redundant solutions of estimated parameters mathematically. The problem with UKF is that it can generate the sigma points outside the defined bound that are very far away from the true value leading to delay in convergence or no convergence. The details of the UKF can be found in the Appendix C. Further, the violation of bounds can occur in correction step even after efficient prediction. There are several ways to handle constraints in the Kalman filter. Here, we have implemented unscented recursive nonlinear dynamics data reconciliation (URNDDR) for generating the sigma points within bounds (Vachhani et al., 2005). The basic state-space model of nonlinear dynamics system equations is composed of two equations: (1) process model, and (2) measurement model. Since there is always some difference in actual process and analytical model, this error is modelled using process noise $\boldsymbol{w}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q}_k)$. Similarly, we have observation noise $\boldsymbol{v}_{k+1} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R}_{k+1})$

$$\begin{aligned} \boldsymbol{x}_{k+1} &= f(\boldsymbol{x}_k) + \boldsymbol{w}_k \\ \boldsymbol{y}_{k+1} &= h(\boldsymbol{x}_{k+1}) + \boldsymbol{v}_{k+1} \end{aligned} \tag{5.1}$$

Where, f and h are the state prediction function and observation function respectively. The state variable $\boldsymbol{x}_k \in \mathcal{N}$ is assumed to be normally distributed represented as $\boldsymbol{x}_k \sim \mathcal{N}(\hat{\boldsymbol{x}}_k, \boldsymbol{P}_k)$. We also suppose that the initial state, $\hat{\boldsymbol{x}}_0$, is known with corresponding initial error covariance matrix $\boldsymbol{P}_0$.

5.3.1 Process of selecting sigma points in CUKF

The lower and upper bounds of the state vector is given by $\boldsymbol{x}_L$ and $\boldsymbol{x}_U$ respectively such that $\boldsymbol{x}_L < \boldsymbol{x} < \boldsymbol{x}_U$. In UKF, all step sizes are equal to $\sqrt{L+\kappa}$, where, L is the length of state vector and κ is the scaling parameter. While in CUKF, the step size along the selected sigma point is given by τ_i^k which is not equal for all. The step size is reduced to keep sigma points within constraints (Mandela et al., 2012). The directions of selected sigma points are given

by $s_i^j = \left(\sqrt{\boldsymbol{P}_{k|k}}\right)_i$ and $s_{L+i}^j = -\left(\sqrt{\boldsymbol{P}_{k|k}}\right)_i$. The step size reduction is performed as follows:

$$\begin{aligned}
\tau_i^k &= \min(\sqrt{n+\kappa}, \tau_1^k, \tau_2^k) \\
\tau_1^k &= \min(\infty, ((x_U)_j - (\widehat{x}_{k|k})_j)/(s_i^k)_j : (s_i^k)_j > 0) \\
\tau_2^k &= \min(\infty, ((x_L)_j - (\widehat{x}_{k|k})_j)/(s_i^k)_j : (s_i^k)_j < 0)
\end{aligned} \tag{5.2}$$

As mentioned, the sigma points selected in CUKF are not equidistant from the $\widehat{\boldsymbol{x}}_{k|k}$. This results in unequal weights as well. However, the all weights sum up to unity, and weights corresponding to different sigma points are as follows: $W_i = a\tau_i + b$ where,

$$\begin{aligned}
a &= \frac{2\kappa - 1}{2(L+\kappa)(S_0 - (2L-1))(\sqrt{L+k}))} \\
b &= \frac{1}{2(L+k)} - \frac{2\kappa - 1}{2\sqrt{L+\kappa}\left(S_0 - (2L+1)\left(\sqrt{L+\kappa}\right)\right)} \\
S_0 &= \sum_{i=1}^{2n} \tau_i
\end{aligned} \tag{5.3}$$

5.3.2 Calculating Kalman gain within correction step in CUKF

The CUKF solves the explicit optimization problem for each sigma point $(\mathcal{X})$. The optimization formulation subjected to the bound constraint $\boldsymbol{\mathcal{X}}_L < \boldsymbol{\mathcal{X}}_{k+1|k+1,i} < \boldsymbol{\mathcal{X}}_U$ is given below

$$\min_{\mathcal{X}_{k+1|k+1,i}} \left\{ \begin{array}{l} (\mathcal{X}_{k+1|k+1,i} - \chi_{k+1|k,i})^T (\boldsymbol{P}_{k+1|k})^{-1} (\mathcal{X}_{k+1|k+1,i} - \mathcal{X}_{k+1|k,i}) \\ +(y_{k+1} - h(\mathcal{X}_{k+1|k+1,i}))^T (\boldsymbol{R}_{k+1|k})^{-1} (y_{k+1} - h(\mathcal{X}_{k+1|k+1,i})) \end{array} \right\} \tag{5.4}$$

The solution of the problem will lead to an expression similar to the Kalman gain $(\boldsymbol{K})$ for linear updating of the state matrix as follows:

$$\boldsymbol{\mathcal{X}}_{k+1|k+1,i} = \boldsymbol{\mathcal{X}}_{k+1|k,i} + \boldsymbol{K}_{k+1}(y_{k+1} - \boldsymbol{Y}_{k+1|k,i}) \tag{5.5}$$

Where, $\boldsymbol{Y}_{k+1|k,i}$ is the predicted value. The posterior update of state vector $\widehat{x}_{k+1|k+1}$ and its covariance matrix is calculated with the following equations.

$$\widehat{\boldsymbol{x}}_{k+1|k+1} = \sum_{i=0}^{2n} W_{k,i} \boldsymbol{\mathcal{X}}_{k+1|k+1,i} \tag{5.6}$$

$$\begin{aligned}
\boldsymbol{P}_{k+1|k+1} = \sum W_{k,i} (\boldsymbol{\mathcal{X}}_{k+1|k+1,i} - \hat{\boldsymbol{x}}_{k+1|k+1})(\boldsymbol{\mathcal{X}}_{k+1|k+1,i} - \hat{\boldsymbol{x}}_{k+1|k+1})^T \\
+ \boldsymbol{Q}_k + \boldsymbol{K}_{k+1} \boldsymbol{R}_{k+1} \boldsymbol{K}_{k+1}^T
\end{aligned} \tag{5.7}$$

Where, $W_{k,i}$ is the weight associated with i^{th} sigma point.

5.4 Procedure for long term condition monitoring

The current chapter presents the suitable efficacy of CUKF in identification of system hysteresis parameters which are generally secular in nature and predicting the system

dynamics for non-deteriorating and deteriorating case for the biaxial coupled hysteresis system. For every healthy system, the knowledge of dynamics i.e., stiffness and damping are the known prior. Here, the biaxial Bouc-Wen model is used to resemble the true behaviour of isolator under biaxial excitation and its hysteresis shape parameters are estimated using CUKF. Once, the hysteresis parameters are captured which are being secular in nature and are not going to be changed. Thus, for the damaged states and for even further healthy states, the system dynamic parameters are estimated by keeping the secular parameters constant as identified earlier. Further, in Kalman filter, when the number of states to be estimated get increased with only limited measurements, the accuracy gets hampered tremendously and therefore system dynamics and hysteresis parameters are not to be identified simultaneously. The issues are:

1. **Computational complexity:** This complexity arises whenever the model contains a lot many predictive variables, and a lot of observations at each time. In both the cases. the prediction and update of the covariance matrix become computationally unfavourable. When the system dynamics and hysteresis parameters are estimated together, the quantity of predictive variables become large and thereby causing convergence issues.
2. **Overconfidence:** The efficiency of the Kalman filters depends on the comparable variance of observations and predicted values. As the number of states increases with only limited observations, there is substantial difference arises in variances, results in total divergence of updated states from the observations and results in poor estimation.

Here, two-stage Kalman filter is proposed for avoiding the two common issues of the Kalman filters based on the past literature works (Kim et al., 2006, 2007; Yi'nan et al., 2021). Considering the issues and the fact that the hysteresis parameters are not going to change, they have been first identified and then utilised for the estimation of system dynamics as presented in Figure 5.1

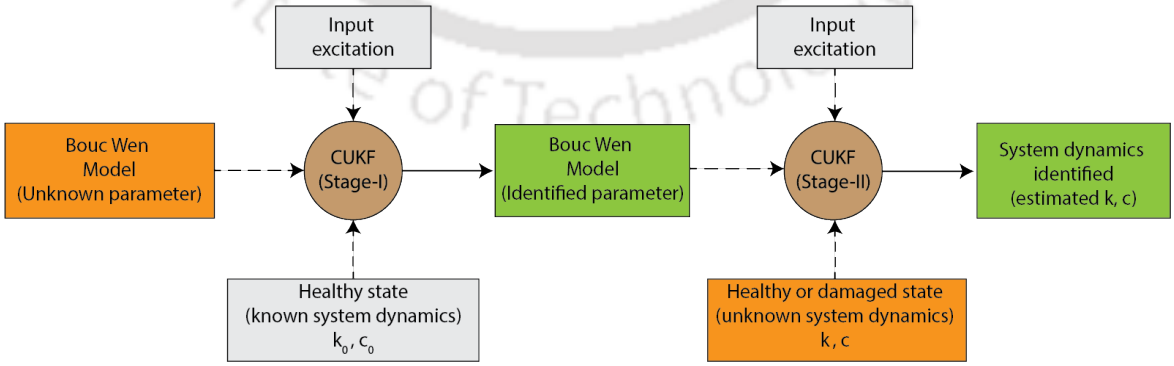


Figure 5.1: Sequence of CUKF used to identify the hysteresis parameters and system dynamics

For long term condition monitoring, two scenarios are considered (1) estimation of stiffness degradation during an extreme event and (2) estimation of stiffness using ambient

vibration for periodic monitoring. Here, the vision is to capture the stiffness deterioration as well as carrying out post-event and periodic maintenance of the structure through unified framework of condition monitoring for systems experiencing bidirectional excitation. Thus, the amplitude of the acceleration responses is monitored and then different actions are taken for two scenarios. In scenario (1), as the amplitude of the measurement response exceeds the threshold limit, the algorithm identifies it as the seismic event and will trigger the CUKF for the event duration. The CUKF estimates the degrading stiffness through proper adaptation of process noise. After the event, the algorithm provides a choice to pass the acceleration data for post-event monitoring and re-estimation of stiffness for damaged state. Similarly, for scenario (ii), i.e., the amplitude will not exceed the threshold, and the system experiences ambient vibrations, the framework provides a choice for the condition monitoring of the system based on the ambient vibration and assessing the stiffness or otherwise, keep it idle for reducing computational load. The framework for long term condition monitoring of the system is illustrated in Figure 5.2. Through condition monitoring of system, we gain insights into how its stiffness and effective damping change over time. These estimated parameters are then used in a prognosis algorithm as detailed in the later part of this chapter to forecast potential future degradation.

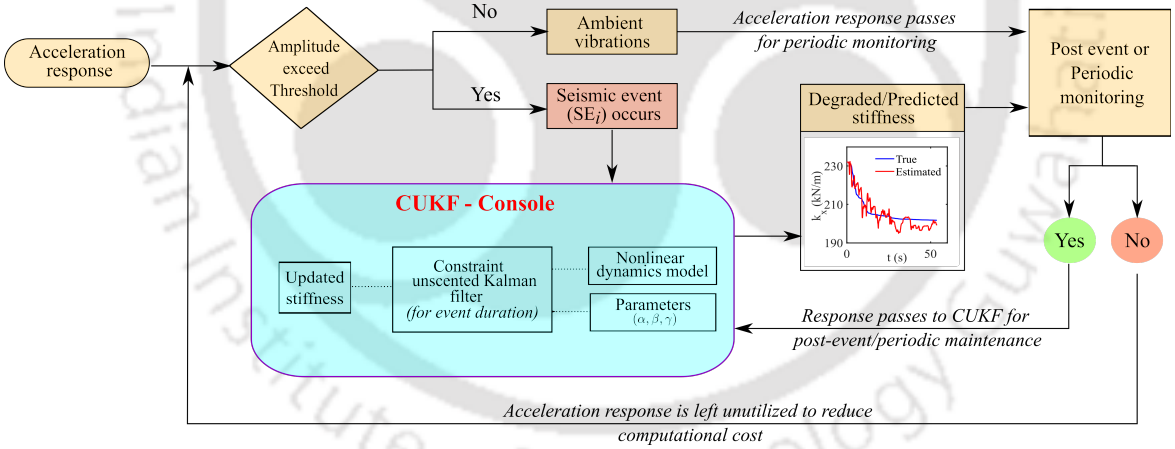


Figure 5.2: Flow chart illustrating the algorithm for real-time monitoring of system

5.5 Numerical case study of rubber bearing

The hysteresis behaviour of high damping rubber bearing exhibits bidirectional effects under large shear strain. The exposure to bi-directional seismic ground motions would cause uneven damage due to torsion or through characteristic behaviour of bearing (Kalkan and Graizer, 2007; Yamamoto et al., 2012). Such effects cannot visualise in uniaxial model as it does not capture the effects in the orthogonal direction and therefore, hysteresis modelling is carried out using the bi-axial formulation of the popular Bouc-Wen model. The restoring

force given by Bouc Wen model is written as

$$\boldsymbol{f_r} = c\dot{\boldsymbol{u}} + \alpha k_i \boldsymbol{u} + (1-\alpha)k_i \boldsymbol{z} \tag{5.8}$$

where, for the isolation system, c represents viscous damping coefficient, $\boldsymbol{u}$ is the displacement vector, the elastic stiffness $k_i = F_y/u_y$ .The F_y and u_y is the yield force and displacement respectively. The parameter $\alpha = k_f/k_i$ is the ratio of post-yield stiffness (k_f) to pre-yield stiffness and the term $(1-\alpha)\, k_i z\,(t)$ represents hysteretic component with hysteresis evolution parameter $z\,(t)$ and is given by:

$$\dot{z} = A\dot{u}\,(t) - \beta\,|\dot{u}\,(t)|\,|z\,(t)|^{n-1} z\,(t) - \ \gamma\dot{u}\,(t)\,|z\,(t)|^{n} \tag{5.9}$$

where A, β, γ and n are the hysteresis shape parameters. This formulation is extended to two orthogonal directions using biaxial Bouc-Wen model. The formulation of state space equations and results of parameter/state estimation of the numerical study are presented together in respective sections.

5.5.1 Seismic rubber bearing under bidirectional excitation

Biaxial Bouc-Wen model

A visual interpretation of bi-axial hysteresis model for simulating isolator moving in two orthogonal direction is shown in Figure 5.3. Consider a system having lumped mass, m , stiffness k_x, k_y , and viscous damping coefficients c_x, c_y , is excited using input ground acceleration vector $\ddot{\boldsymbol{u}}_{gx}, \ddot{\boldsymbol{u}}_{gy}, \ddot{\boldsymbol{u}}_{gv}$. The displacement response of the system is given by $\boldsymbol{u}_{bx}, \boldsymbol{u}_{by}$ where, subscript 'b' represents the base storey. The dynamic equations of motion in two orthogonal directions considering P-delta effects are written as follows:

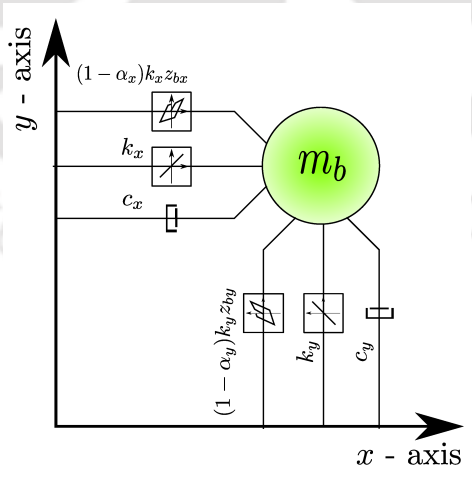


Figure 5.3: Schematic view of biaxial hysteresis mode

$$\begin{aligned} m\ddot{\boldsymbol{u}}_{bx} + c_x\dot{\boldsymbol{u}}_{bx} + \alpha_x k_x \boldsymbol{u}_{bx} + (1-\alpha_x)k_x \boldsymbol{z}_{bx} &= -m\ddot{\boldsymbol{u}}_{gx} - m(\ddot{\boldsymbol{u}}_{gv} - g)\boldsymbol{u}_{bx}/h \\ m\ddot{\boldsymbol{u}}_{by} + c_y\dot{\boldsymbol{u}}_{by} + \alpha_y k_y \boldsymbol{u}_{by} + (1-\alpha_y)k_y \boldsymbol{z}_{by} &= -m\ddot{\boldsymbol{u}}_{gy} - m(\ddot{\boldsymbol{u}}_{gv} - g)\boldsymbol{u}_{by}/h \end{aligned} \tag{5.10}$$

Here, α_x, α_y are the ratios of inelastic to elastic stiffness, and $\boldsymbol{z}_{bx}$,$\boldsymbol{z}_{by}$ are the hysteresis evolution parameters in x and y directions. The terms $m(\ddot{\boldsymbol{u}}_{gv} - g)\boldsymbol{u}_{bx}/h, m(\ddot{\boldsymbol{u}}_{gv} - g)\boldsymbol{u}_{by}/h$ represent the force introduced in the system from gravity loads. Here, $\ddot{\boldsymbol{u}}_{\boldsymbol{gv}}$ is the vertical ground motion, h is the constant gravitational acceleration, and is the height of the storey.

The dynamic equation can be written in normalized form by defining dimensionless parameters as $\boldsymbol{\upsilon}_{bx} = \boldsymbol{u}_{bx}/D_{Yx}, \boldsymbol{\varphi}_x = \boldsymbol{z}_{bx}/D_{Yx}, \boldsymbol{\upsilon}_{by} = \boldsymbol{u}_{by}/D_{Yy}, \boldsymbol{\varphi}_y = \boldsymbol{z}_{by}/D_{Yy}$ where, D_{Yx}, D_{Yy} are the yield displacements and is given by:

$$\begin{aligned}\ddot{\boldsymbol{\upsilon}}_{bx} + 2\xi_x\omega_x\dot{\boldsymbol{\upsilon}}_{bx} + \alpha_x\omega_x^2\boldsymbol{\upsilon}_{bx} + (1-\alpha_x)\omega_x^2\boldsymbol{\varphi}_x &= -\ddot{\boldsymbol{u}}_{gx}D_{Yx} - (\ddot{\boldsymbol{u}}_{gv}/g - 1)\theta_x\omega_x^2\boldsymbol{\upsilon}_x \\ \ddot{\boldsymbol{\upsilon}}_{by} + 2\xi_y\omega_y\dot{\boldsymbol{\upsilon}}_{by} + \alpha_y\omega_y^2\boldsymbol{\upsilon}_{by} + (1-\alpha_y)\omega_y^2\boldsymbol{\varphi}_y &= -\ddot{\boldsymbol{u}}_{gy}D_{Yy} - (\ddot{\boldsymbol{u}}_{gv}/g - 1)\theta_y\omega_y^2\boldsymbol{\upsilon}_y\end{aligned} \tag{5.11}$$

Where, ω_x, ω_y are the natural frequency and $\xi_x = c_x/(2m\omega_x), \xi_y = c_y/(2m\omega_y)$, are the damping ratios. The parameters $\theta_x = mg/k_xh, \theta_y = mg/k_yh$ are the stability factors (MacRae, 1994; Humar et al., 2006). In the presented numerical example, the energy-based calculation of degradation is not considered and therefore, the normalized nonlinear hysteresis displacement $\boldsymbol{\varphi}_x$,$\boldsymbol{\varphi}_y$ considering is expressed as:

$$\begin{aligned}\dot{\boldsymbol{\varphi}}_x &= \dot{\boldsymbol{\upsilon}}_{bx} - \boldsymbol{\varphi}_x I \\ \dot{\boldsymbol{\varphi}}_y &= \dot{\boldsymbol{\upsilon}}_{by} - \boldsymbol{\varphi}_y I\end{aligned} \tag{5.12}$$

The seismic bearing, like high damping rubber bearing, shows different characteristics under bidirectional excitations (Kato et al., 2015; Yamamoto et al., 2012). Thus, an essential term in the Equation 5.12 differentiates the model from the uniaxial one and denotes the interaction behaviour among both axes and is expressed as follows:

$$I = |\dot{\boldsymbol{\upsilon}}_{bx}|\,|\boldsymbol{\varphi}_x|^{n_x-1}\left[\beta_x + \gamma_x \mathrm{sgn}(\dot{\boldsymbol{\upsilon}}_{bx}\boldsymbol{\varphi}_x)\right] + |\dot{\boldsymbol{\upsilon}}_{by}|\,|\boldsymbol{\varphi}_y|^{n_y-1}\left[\beta_y + \gamma_y \mathrm{sgn}(\dot{\boldsymbol{\upsilon}}_{by}\boldsymbol{\varphi}_y)\right] \tag{5.13}$$

where, $\mathrm{sgn}(\cdot)$ is the signum function, $\beta_x, \gamma_x, \beta_y, \gamma_y$ are the Bouc-Wen hysteresis shape parameters

State space model formulation for estimating parameters

The process model equation $\boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k) + \boldsymbol{w}_k$ where, f is the state prediction function and $\boldsymbol{w}_k$ is the process noise vector, the state space vector $\boldsymbol{x}$ is formed for the estimation of the parameters in Stage 1 through CUKF are given by:

$$x = \{x_1, x_2, x_3, \ldots, x_{12}\}^T = \{\upsilon_{bx}, \dot{\upsilon}_{bx}, \upsilon_{by}, \dot{\upsilon}_{by}, \varphi_x, \varphi_y, \alpha_x, \alpha_y, \beta_x, \beta_y, \gamma_x, \gamma_y\}^T \tag{5.14}$$

and state space equations can be written as shown in Equation 5.15. Here, $I = |x_2|\,|x_5|^{n_x-1}\left[x_9 + x_{11}\mathrm{sgn}(x_2x_5)\right] + |x_4|\,|x_6|^{n_y-1}\left[x_{10} + x_{12}\mathrm{sgn}(x_4x_6)\right]$, and u_{gx}, u_{gy}, u_{gv} are the input ground acceleration. The time derivatives of the biaxial Bouc-Wen parameters are equal to zero because the parameters are assumed to be constant. Since, there is

no-deterioration, the assumed uncertainty in zero and the process noise is a null vector.

$$\{\dot{x}\} = \begin{Bmatrix} \dot{v}_{bx} \\ \ddot{v}_{bx} \\ \dot{v}_{by} \\ \ddot{v}_{by} \\ \dot{\varphi}_x \\ \dot{\varphi}_y \\ \dot{\alpha}_x \\ \dot{\alpha}_y \\ \dot{\beta}_x \\ \dot{\beta}_y \\ \dot{\gamma}_x \\ \dot{\gamma}_y \end{Bmatrix} = \begin{Bmatrix} x_2 \\ -2\xi_x\omega_x x_2 - \omega_x^2 x_8 x_1 - (1-x_8)\omega_x^2 x_5 \\ \quad -u_{gx}/D_{yx} + (1-u_{gv}/g)\,\theta_x\omega_x x_1 \\ x_4 \\ -2\xi_y\omega_y x_4 - \omega_y^2 x_9 x_3 - (1-x_9)\omega_y^2 x_6 \\ \quad -u_{gy}/D_{yy} + (1-u_{gv}/g)\,\theta_y\omega_y x_3 \\ x_2 - x_5 I \\ x_4 - x_6 I \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{5.15}$$

The measurement equation $\boldsymbol{y}_{k+1} = h(\boldsymbol{x}_{k+1}) + \boldsymbol{v}_{k+1}$ is sum of observation function and measurement noise $\boldsymbol{v}$. The observations are obtained from the accelerometer, however, numerically, it can be obtained using following equation.

$$\begin{aligned} \{y\} = \begin{Bmatrix} \ddot{v}_{bx} \\ \ddot{v}_{by} \end{Bmatrix} &= \begin{Bmatrix} -2\xi_x\omega_x x_2 - \omega_x^2 x_7 x_1 - (1-x_7)\omega_x^2 x_5 \\ -2\xi_y\omega_y x_4 - \omega_y^2 x_8 x_3 - (1-x_8)\omega_y^2 x_6 \end{Bmatrix} \\ &+ \begin{Bmatrix} u_{gx}/D_{yx} - (1-u_{gv}/g)\,\theta_x\omega_x x_1 \\ u_{gy}/D_{yy} - (1-u_{gv}/g)\,\theta_y\omega_y x_3 \end{Bmatrix} + v \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} \end{aligned} \tag{5.16}$$

Here, v is the additive measurement noise.

State space model formulation for estimating system dynamics

The hysteresis behaviour of the system is quite a secular phenomenon and the parameters will not change with time. Thus, the parameters estimated in Stage 1 will be used to estimate system dynamics in Stage 2. The state space vector $\boldsymbol{x}$ is formed for estimating system dynamics in Stage 2 through CUKF is given by

$$x = \{x_1, x_2, x_3, \ldots, x_{10}\}^T = \left\{v_{bx}, \dot{v}_{bx}, v_{by}, \dot{v}_{by}, \dot{\varphi}_x, \dot{\varphi}_y, \dot{k}_x, \dot{k}_y, \dot{c}_x, \dot{c}_y\right\}^T \tag{5.17}$$

and state space equations can be written as follows:

$$\{\dot{x}\} = \begin{Bmatrix} \dot{v}_{bx} \\ \ddot{v}_{bx} \\ \dot{v}_{by} \\ \ddot{v}_{by} \\ \dot{\varphi}_x \\ \dot{\varphi}_y \\ \dot{k}_x \\ \dot{k}_y \\ \dot{c}_x \\ \dot{c}_y \end{Bmatrix} = \begin{Bmatrix} x_2 \\ -(x_9/m)x_2 - (x_7/m)\alpha_x x_1 - (1-\alpha_x)(x_7/m)x_5 \\ \qquad -u_{gx}/D_{yx} + (1 - u_{gv}/g)\,\theta_x(x_7/m)x_1 \\ x_4 \\ -(x_{10}/m)x_4 - (x_8/m)x_3 - (1-\alpha_y)(x_8/m)x_6 \\ \qquad -u_{gy}/D_{yy} + (1 - u_{gv}/g)\,\theta_y(x_8/m)x_3 \\ x_2 - x_5 I \\ x_4 - x_6 I \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{5.18}$$

where, $I = |x_2|\,|x_5|^{n_x-1}\left[\beta_x + \gamma_x \mathrm{sgn}(x_2 x_5)\right] + |x_4|\,|x_6|^{n_y-1}\left[\beta_y + \gamma_y \mathrm{sgn}(x_4 x_6)\right]$, and u_{gx}, u_{gy}, u_{gv} are the input ground acceleration. Since, there is no-deterioration, the assumed uncertainty in zero and the process noise is a null vector as well as the time derivatives of stiffness and damping are equal to zero. The measurement equation $\boldsymbol{y}_{k+1} = h(\boldsymbol{x}_{k+1}) + \boldsymbol{v}_{k+1}$ will remain same as for the stage 1.

Results for estimation of hysteresis parameters and system dynamics

For the estimation of hysteresis system parameters, the UKF and CUKF are applied to the simulated system. The typical bound constraint is applied on the hysteresis shape parameters for the CUKF, and the range of the parameter β is [0.5 1.5] and, γ is [-0.3 0.5] (Ma et al., 2004). The systems are excited under Kobe earthquake and acceleration response is used for the estimation of parameters and system dynamics through CUKF. The bearing used in the numerical example is simulated with the bi-axial hysteresis parameters as $\alpha_x = 0.15, \alpha_y = 0.2, \beta_x = 1.1, \beta_y = 1.4, \gamma_x = 0.2, \gamma_y = 0.5$. The mass m, stiffness k, and damping c for the system are taken as 6800 kg, 1392 kN/m, 3740 Ns/m, respectively. It can be clearly observed that the accuracy of the CUKF is much higher, it also shows faster convergence than unconstraint UKF. The results for the parameters and hysteresis curve estimation using stage 1 CUKF in both X and Y directions are provided in Figure 5.4 and Figure 5.5. The estimation of the interaction behaviour among both the axis of the system is presented in Figure 5.6.

Once, the hysteresis parameters are captured, system dynamics are identified keeping the secular parameters constant. The results of the stiffness and damping estimation are presented in Figure 5.7

The graphical comparison of the final estimated values of parameters and system dynamics

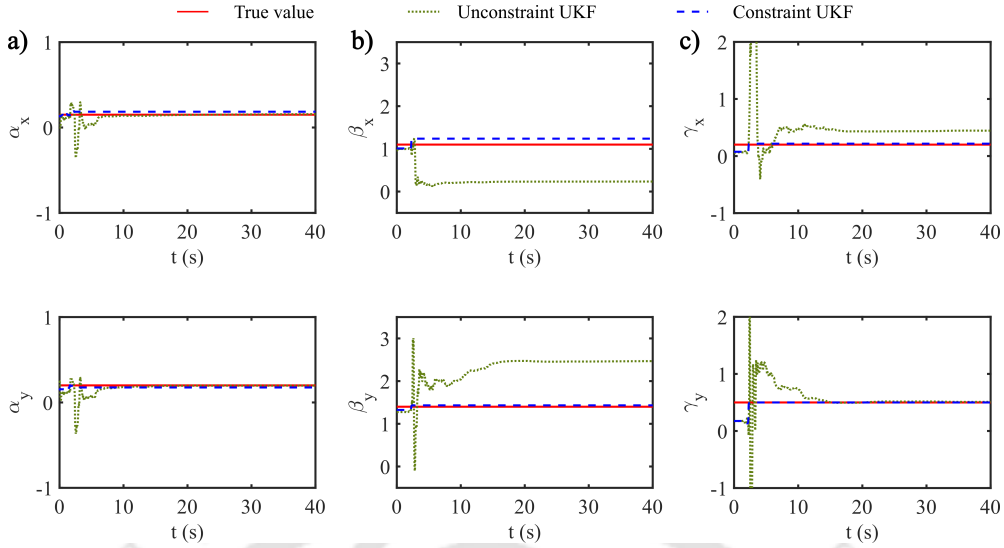


Figure 5.4: Estimation of biaxial Bouc Wen parameters, a) stiffness ratio (α) , b) shape parameter (β), and c) shape parameter (γ) in the X-Y direction

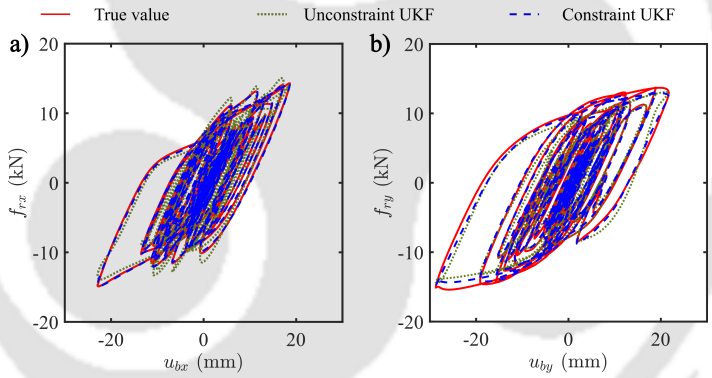


Figure 5.5: Estimation of hysteresis behavior along (a) X-axis, and (b) Y-axis for BBW model

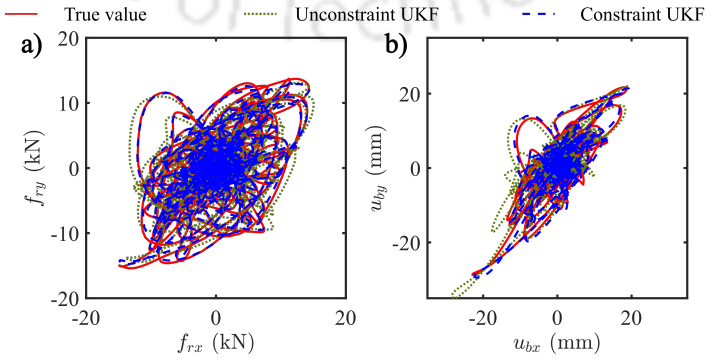


Figure 5.6: Estimation of (a) Force interaction, and (b) Displacement interaction for BBW model

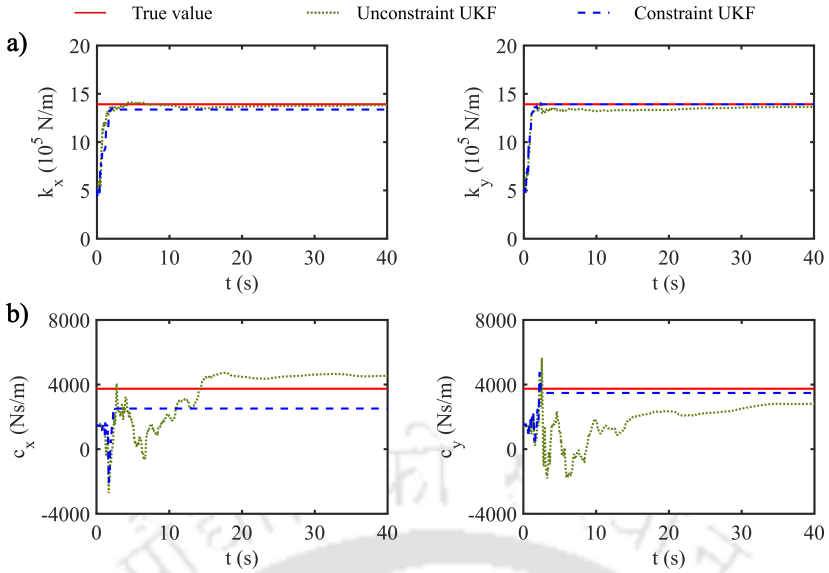


Figure 5.7: Estimation of a) stiffness, and b) damping of the LRB system along X and Y axes

using the CUKF and UKF methods is already shown. To further investigate the influence of noise levels, the relative percentage error and convergence steps are calculated and presented in Table 5.1 and Table 5.2. Overall, an observation can be made that CUKF performs better for the noise levels from 5% to 15% in estimating the stiffness and damping parameters. Further, CUKF has faster convergence at each level of noise.

Table 5.1: Relative percentage error values for the different noise levels

Noise level	Type of Kalman filter	Parameters									
		α_x	α_y	β_x	β_y	γ_x	γ_y	k_x	k_y	c_x	c_y
5 %	UKF	37.0	30.0	78.7	76.3	122.1	2.1	0.6	2.0	21.8	23.8
	CUKF	18.8	26.5	12.6	2.3	8.3	0.0	3.9	0.1	32.7	6.9
10 %	UKF	32.7	26.1	73.5	98.9	176.9	88.2	6.6	12.6	268.0	302.3
	CUKF	24.5	15.3	14.2	7.1	1.4	4.8	6.3	7.8	50.4	10.2
15 %	UKF	29.6	23.5	74.1	96.6	182.9	81.6	0.9	3.6	43.0	18.0
	CUKF	11.0	15.3	4.1	1.1	34.7	0.1	10.5	25.2	10.7	218.6

Table 5.2: Relative number of steps to converge Kalman filter

Noise level	Type of Kalman filter	Parameters									
		α_x	α_y	β_x	β_y	γ_x	γ_y	k_x	k_y	c_x	c_y
5%	UKF	777	719	746	1656	1542	1652	1138	1012	1766	2062
	CUKF	209	157	254	231	228	249	194	223	232	264
10%	UKF	672	672	663	1541	1337	1518	1077	1194	1855	17800
	CUKF	230	363	786	445	406	918	300	403	585	4391
15%	UKF	618	631	627	1581	1580	1806	1031	2106	1561	1515
	CUKF	215	234	265	319	261	292	247	251	309	493

5.5.2 Description and estimation of fatigue degradation for condition monitoring

In this section, the fatigue degradation based on Paris law is introduced in the Bouc-Wen model to simulate the accelerated fatigue degradation in the form of stiffness deterioration. The degradation of stiffness in both directions for the true response in the numerical example is modelled using the process described in subsequent subsection.

Fatigue based degradation model used for simulating stiffness degradation

The cyclic loads applied during the seismic events and ambient vibrations causes fatigue deterioration that might be slow or fast depending on the condition of structure. This fatigue deterioration manifest itself in the form of stiffness and strength degradation. For an initial crack length of A_0 , consider $x(t), A(t)$, $D(t)$ be the isolator displacement, crack size, and degradation due to fatigue. The degradation $D(t) = f(A(t))$ is depending on the propagation of the crack size $A(t)$. The numerical model used in the current work considers the fatigue deterioration only in the form of stiffness. The stiffness of the isolation system $k(t)$ is related to the degradation $D(t)$ with nonlinear mapping function q as given by (Sobczyk and Trebicki, 2000)

$$q(D(t)) = \alpha_1 + \alpha_2 \exp(-\alpha_3 D^{-\alpha_4}) \tag{5.19}$$

where, α_i are constants. According to the Paris–Erdogan equation (Paris and Erdogan, 1963), the rate of change of fatigue crack length with the number of cycles N for a crack size A is $dA/dN = C(\Delta K)^m$, where C, m are the constants and ΔK is the stress intensity factor. Sobczyk and Trebicki (2000) define some degradation measure $D = \psi(A)/\psi^*$, where $\psi(A)$ is a nonlinear transformation of crack size and for the stress range ΔS_Y , the differential of degradation measure dD is $(1/\psi^*)\, C_1(\Delta S_Y)^m dN$ where C_1 is the constant. Further, $dN \approx \mu dt$ and for a lightly damped system and narrowband response, the degradation $D(t)$ can be written in the state space form as expressed as given by Equation 5.20.

$$\dot{D}(t) = \frac{dD}{dt} = C_1 \left[Y_1^2(t) + Y_2^2(t) \right]^{m/2} \tag{5.20}$$

Here, $Y_1 = y_1$ and $Y_2 = \dot{y}_1$ are the state space vectors representing the isolator displacement and velocity, C_1 is the constant. Therefore, the expression of stiffness in defining restoring force is given as $k_i = qk_{i0}$, where, $(i = x, y)$ is the nonlinear function of degradation as given in Equation 5.19 and k_{i0} is the initial stiffness in X and Y direction for the isolator. The fatigue model parameters are selected in order to simulate an accelerated fatigue environment for stiffness degradation. The value of the constants a_1, α_2, α_3, and α_4 in the nonlinear function are taken as 0.4, 0.6, 0.1 and 1 respectively. For the accelerated fatigue, the value of the degradation parameters C_1, m for in both axes in the Paris law formulation are taken as 0.005 and 2.

State space model formulation for estimating degrading system dynamics

The hysteresis behaviour of the system is quite a secular phenomenon and the parameters will not change with time. Thus, the parameters estimated in Stage 1 will be used to estimate system dynamics in Stage 2. However, to estimate the continuously deteriorating stiffness under accelerated fatigue, the state space formulation will be modified and now Equation 5.20 will be introduced. The state space vector $\boldsymbol{x}$ is formed for estimating deteriorating system dynamics through CUKF is given by and state space equations can be written as Equation 5.22.

$$x = \{x_1, x_2, x_3, \ldots, x_{12}\}^T = \{\upsilon_{bx}, \dot{\upsilon}_{bx}, \upsilon_{by}, \dot{\upsilon}_{by}, \dot{\varphi}_x, \dot{\varphi}_y, D_x, D_y, C_{1x}, C_{1y}, m_x, m_y\}^T \tag{5.21}$$

Here, $I = |x_2|\,|x_5|^{n_x-1}\left[\beta_x + \gamma_x \mathrm{sgn}(x_2 x_5)\right] + |x_4|\,|x_6|^{n_y-1}\left[\beta_y + \gamma_y \mathrm{sgn}(x_4 x_6)\right]$, and u_{gx}, u_{gy}, u_{gv} are the input ground acceleration. Thus, the degradation is incorporated in state space equations for identifying the deteriorating stiffness with time. Further, the approach for real-time condition monitoring of the deteriorating systems is already discussed and results are presented in subsequent section.

$$\{\dot{x}\} = \begin{Bmatrix} \dot{\upsilon}_{bx} \\ \ddot{\upsilon}_{bx} \\ \dot{\upsilon}_{by} \\ \ddot{\upsilon}_{by} \\ \dot{\varphi}_x \\ \dot{\varphi}_y \\ \dot{D}_x \\ \dot{D}_y \\ \dot{C}_{1x} \\ \dot{C}_{1y} \\ \dot{m}_x \\ \dot{m}_y \end{Bmatrix} = \begin{Bmatrix} x_2 \\ -(c_x/m)x_2 - (q_x k_x/m)\alpha_x x_1 - (1-\alpha_x)(q_x k_x/m)x_5 \\ -u_{gx}/D_{yx} + (1 - u_{gv}/g)\,\theta_x (q_x k_x/m)x_1 \\ x_4 \\ -(x_{10}/m)x_4 - (q_y k_y/m)x_3 - (1-\alpha_y)(q_y k_y/m)x_6 \\ -u_{gy}/D_{yy} + (1 - u_{gv}/g)\,\theta_y (q_y k_y/m)x_3 \\ x_2 - x_5 I \\ x_4 - x_6 I \\ x_9\left[x_1^2 + x_2^2\right]^{x_{11}/2} \\ x_{10}\left[x_1^2 + x_2^2\right]^{x_{12}/2} \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{5.22}$$

Results for estimation of degradation parameters and system dynamics

The systems are excited under Kobe earthquake and acceleration response is used for the estimation of parameters and system dynamics through CUKF. It can be observed that the accuracy of the CUKF is good and shows faster convergence. The results for the degradation parameters estimation using CUKF in both X and Y directions are provided in Figure 5.8. The estimation of the degradation state and degrading stiffness among both the axis of the system is presented in Figure 5.9.

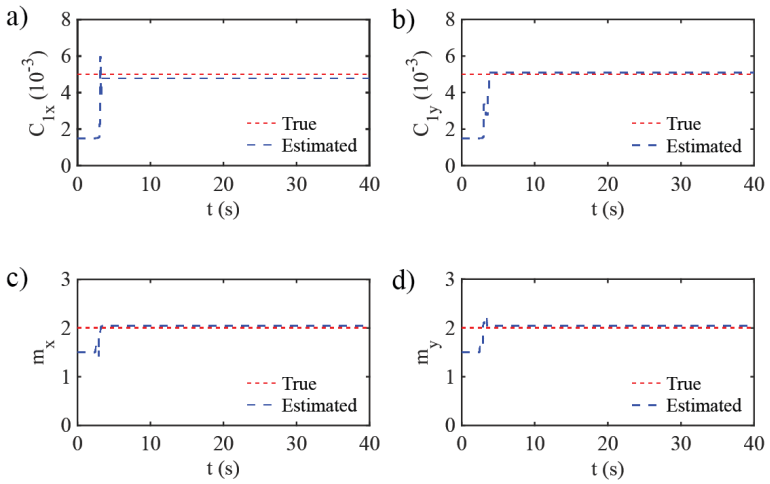


Figure 5.8: Estimation of degradation parameters in X and Y axis

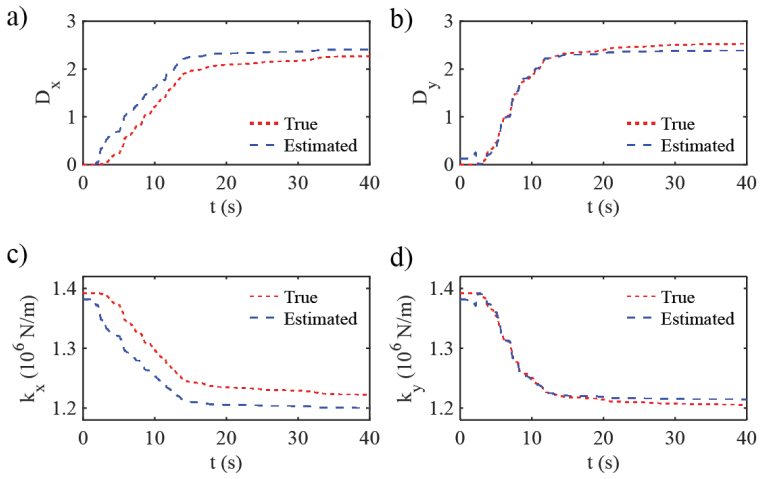


Figure 5.9: Estimation of degradation state and degrading stiffness in X and Y axis

5.6 Experimental study on elastometic rubber bearing

In order to validate the efficacy of two-stage CUKF framework in estimating the parameters of Biaxial Bouc- Wen model, an experiment is conducted on the elastomeric rubber bearing in the Structural Laboratory, IIT Guwahati. The details of specimen are provided in Table 5.3. A graphics representing the geometric properties is also presented and is shown in Figure 5.10.

5.6.1 Experiment setup and test program

In order to investigate the degradation in the elastomeric bearing, the test setup is developed with the combination of two actuators (make: MTS USA) which act on the specimen simultaneously to input the horizontal cyclic displacement under the constant vertical load. The details of the actuator capacity are provided in Table 5.4. A steel block of

Table 5.3: Description of elastomeric rubber bearing used in the study

Item	Details
Size of specimen	300 mm x 300 mm
Number of elastomer layers	8
Thickness of single elastomeric layer	10 mm
Rubber cover thickness	5.5 mm
Total height of elastomer	118 mm
Number of steel shims	9
Thickness of single steel shim	3 mm
Size of steel shims	288 mm x 288 mm
Aspect ratio	1
Shear modulus of elastomer	0.7

required dimension has been put over the specimen. The bearing is in contact with lower surfaces of the steel block and placed over the steel plate which is fixed at the required height and connected to the strong floor. The schematic-diagram for the experimental setup is shown in Figure 5.11. The specimens are placed at the angle of 30 degrees and 45 degrees under nominal axial load of 100 kN corresponding to axial pressure of 1.2 MPa for accounting the bidirectional effects as the excitation is provided in one direction only. The input displacement and output force are then decoupled in two orthogonal axes through coordinate transformation for proper application of Biaxial Bouc-Wen model. A constant vertical load is applied using a vertically placed servo-hydraulic actuator on the top surface of the steel block. The details of the experiment setup and the loading protocol are shown in Figure 5.12.

Table 5.4: Description of elastomeric rubber bearing used in the study

Actuator	Stroke	Force (kN)
Actuator 1 - Vertical (MTS)	±500	1000 Compression 800 Tension
Actuator 2 - Horizontal (MTS)	±250	250 Compression 250 Tension

In the experimental study, the output is the hysteresis force with displacement input, thus there is slight change in the state space formulation and it is now written as $\{\dot{x}\} = \left\{\dot{\varphi}_x, \dot{\varphi}_y, \dot{\alpha}_x, \dot{\alpha}_y, \dot{\beta}_x, \dot{\beta}_y, \dot{\gamma}_x, \dot{\gamma}_y\right\} = \{u_x - x_1 I, u_y - x_2 I, 0, 0, 0, 0, 0, 0\}$ where, $I = |u_x|\,|x_1|^{n_x-1}\left[x_5 + x_7 \mathrm{sgn}(u_x x_1)\right] + |u_y|\,|x_2|^{n_y-1}\left[x_6 + x_8 \mathrm{sgn}(u_x x_2)\right]$, and u_x, u_y are the components of the input displacement

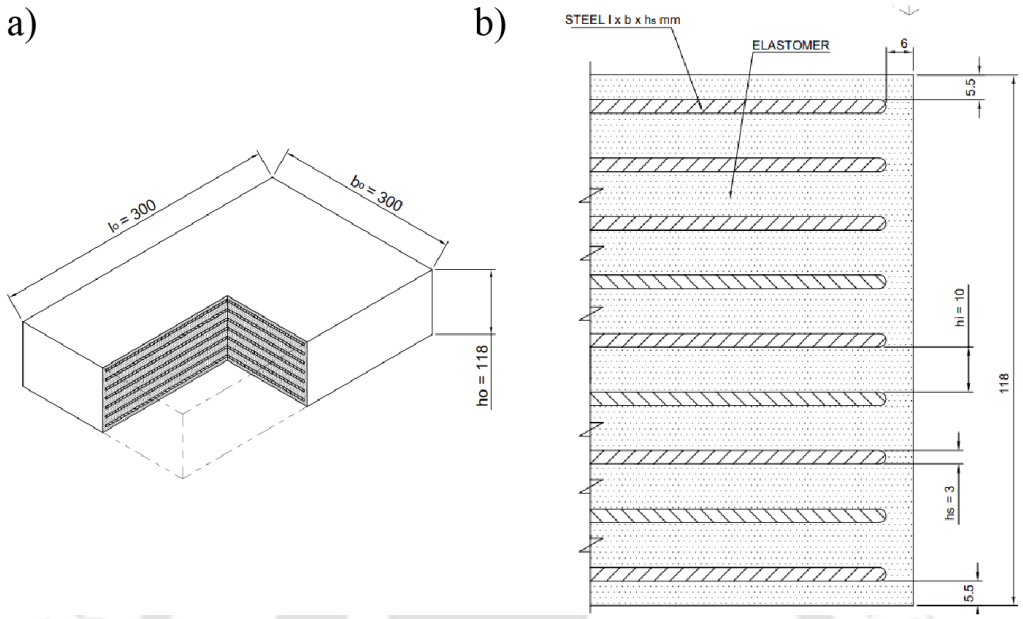


Figure 5.10: Specimen cross-section showing the stacking of steel shims and elastomeric rubber

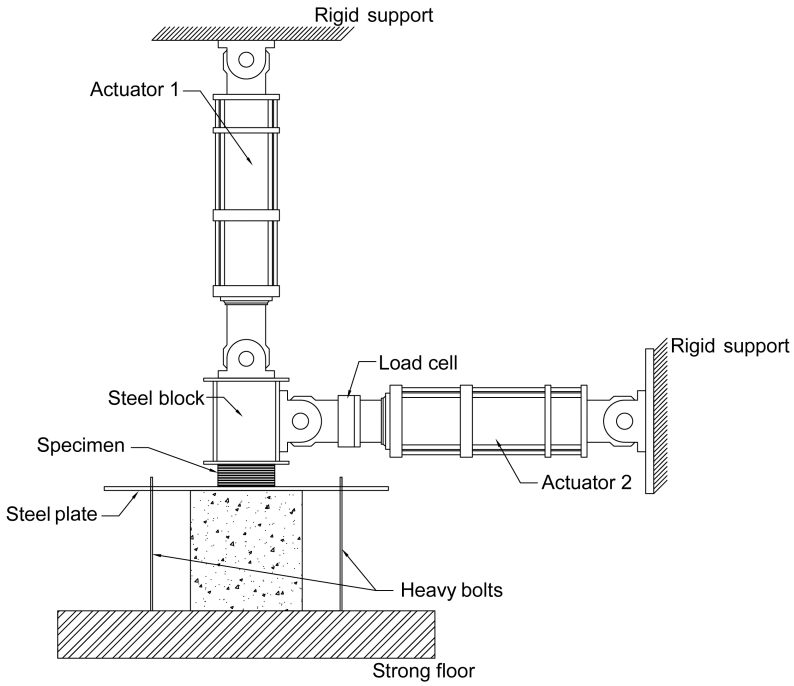


Figure 5.11: Schematic diagram showing the actuator representation

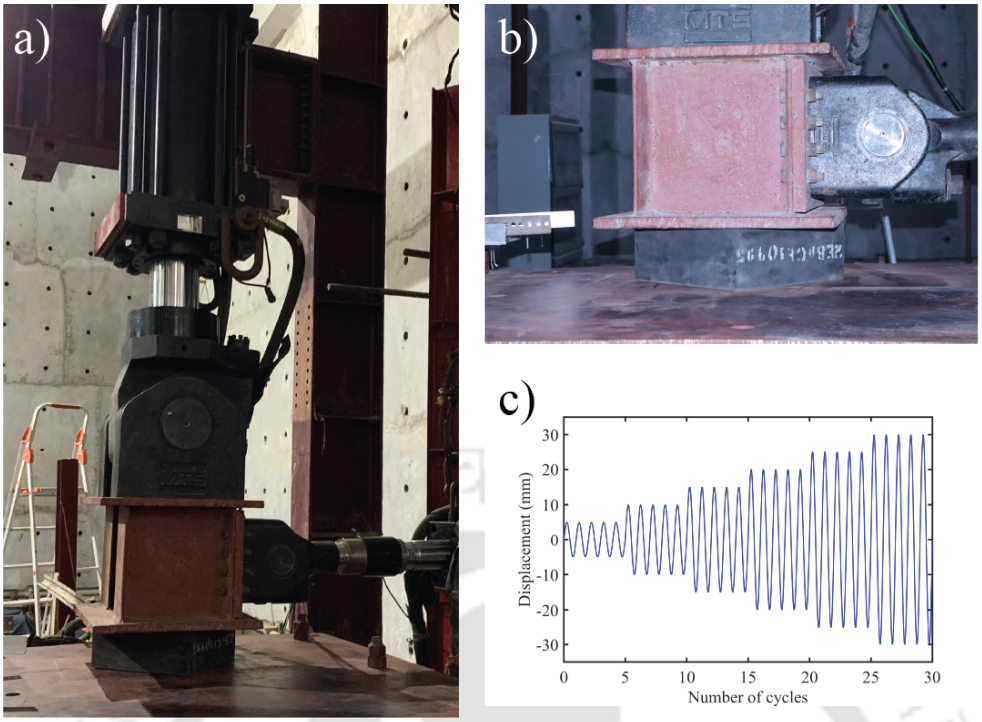


Figure 5.12: Experimental setup (a) full view, (b) close view and (c) loading protocol

5.6.2 Results for estimation of hysteresis parameters for the experimental study

For the estimation of hysteresis system parameters, the CUKF are applied and the estimated force through biaxial Bouc-Wen model is compared with the experimental hysteresis force output in order to predict the hysteresis parameters in both the directions. The two specimens with orientation of 30 degrees and 45 degrees with the horizontal actuators are used to validate the accuracy of CUKF for the estimation of biaxial Bouc Wen model. The results for the hysteresis curve estimation using CUKF in both X and Y directions for the specimen at 30 degree and 45 degrees orientation are provided in Figure 5.13 and Figure 5.15 respectively. The residual error plot in the estimation of the restoring force in both X and Y directions for the specimen at 30 degree and 45 degrees orientation are provided in Figure 5.14 and Figure 5.16 respectively. Further, the values of the estimated parameters in both X and Y directions for the specimen at 30 degrees and 45 degrees orientation are provided in Figure 5.17 and Figure 5.18 respectively. The experimental study shows that CUKF has robust capabilities in system identification of biaxial coupled hysteresis system. The residual error plot in force estimation shows the accuracy of CUKF algorithm.

5.6.3 Development of damage index for long term prognosis

The experimental study is now extented to analyze the fatigue behavior under cyclic loading. The typical service design loads for elastomeric bearings used for seismic

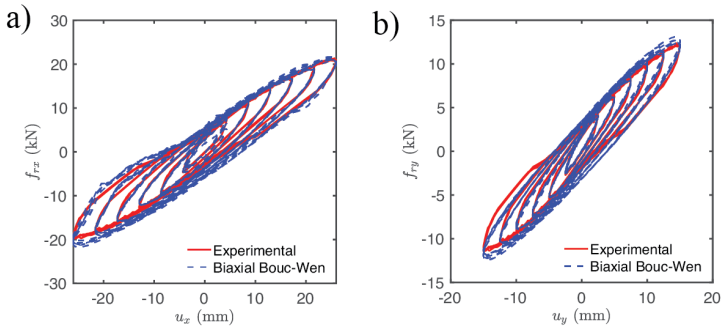


Figure 5.13: Hysteresis estimation for the specimen at 30 degree orientation (a) X- axis, and (b) Y-axis

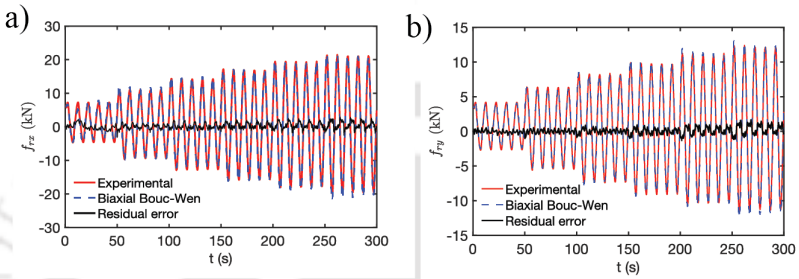


Figure 5.14: Estimated force with residual error for specimen at 30 degree orientation (a) X- axis, and (b) Y-axis

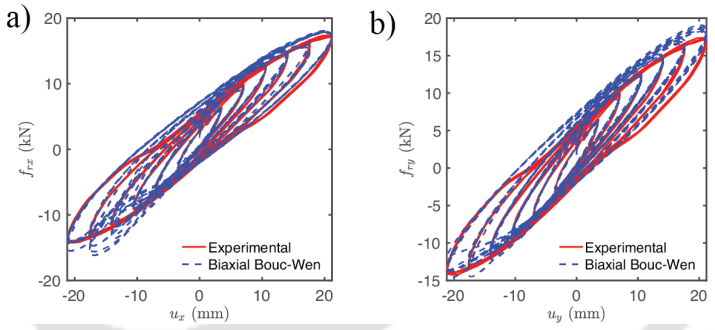


Figure 5.15: Hysteresis estimation for the specimen at 45 degree orientation (a) X- axis, and (b) Y-axis

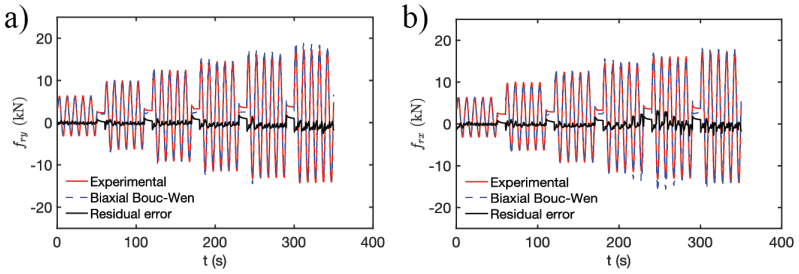


Figure 5.16: Estimated force with residual error for specimen at 45 degree orientation (a) X- axis, and (b) Y-axis

isolation of structures vary in the range of 3 MPa to 7 MPa. Thus, a nominal axial load of 100 kN corresponding to axial pressure of 1.2 MPa was used. Fatigue is critically explained using curves that define the stress range as a function of the number of cycles (i.e., S-N

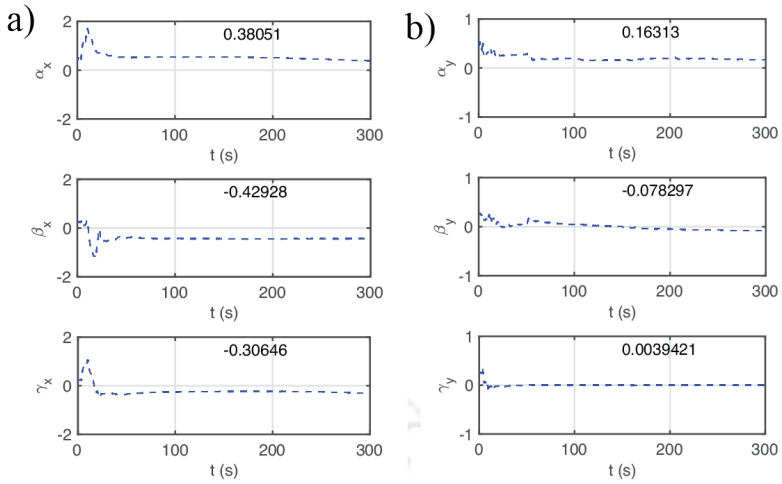


Figure 5.17: Estimated biaxial Bouc Wen parameters for the specimen at 30 degree orientation (a) X-axis, and (b) Y-axis

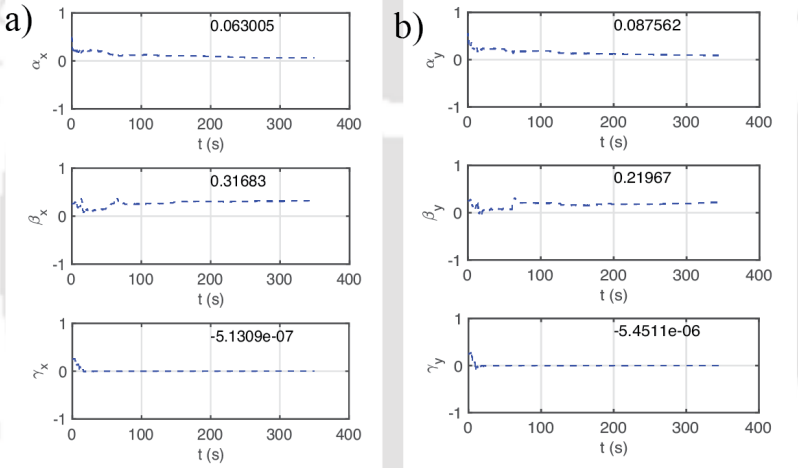


Figure 5.18: Estimated biaxial Bouc-Wen parameters for the specimen at 45 degree orientation (a) X-axis, and (b) Y-axis

curves). However, it is inadequate to obtain the dataset for the S-N due to time and specimen constraint. The test program was designed to form a matrix of critical parameters like energy dissipated per cycle, with respect to both fatigue fracture mechanics and structure health monitoring, that can be carried forward for the prognosis. The test has been performed using sine wave displacement input for three different sets of input frequencies i.e., 0.2 Hz, 0.4 Hz, 0.6 Hz at two different levels of shear strains i.e., 18% and 20% in a successive way. Thus, for each of the excitation frequency, we have 2000 cycles of 18% shear strain and 3000 cycles of 20% shear strain. The frequency will increase after completion of 5000 cycles. This procedure is followed to finish test faster and counts the effects of the frequencies and strain on the degradation. This variation also helps us to generate multiple change points in degradation which further utilized while

demonstrating the efficacy of proposed EGA-BGPR framework. In order to leverage the advantages of the two-stage CUKF framework, we have selected few cycles after 100 cycles to estimate the biaxial -BoucWen hysteresis parameters and system dynamics. Thus, we obtained the hysteresis loop and corresponding stiffness at each 100 cycles of fatigue degradation as shown in Figure 5.19.

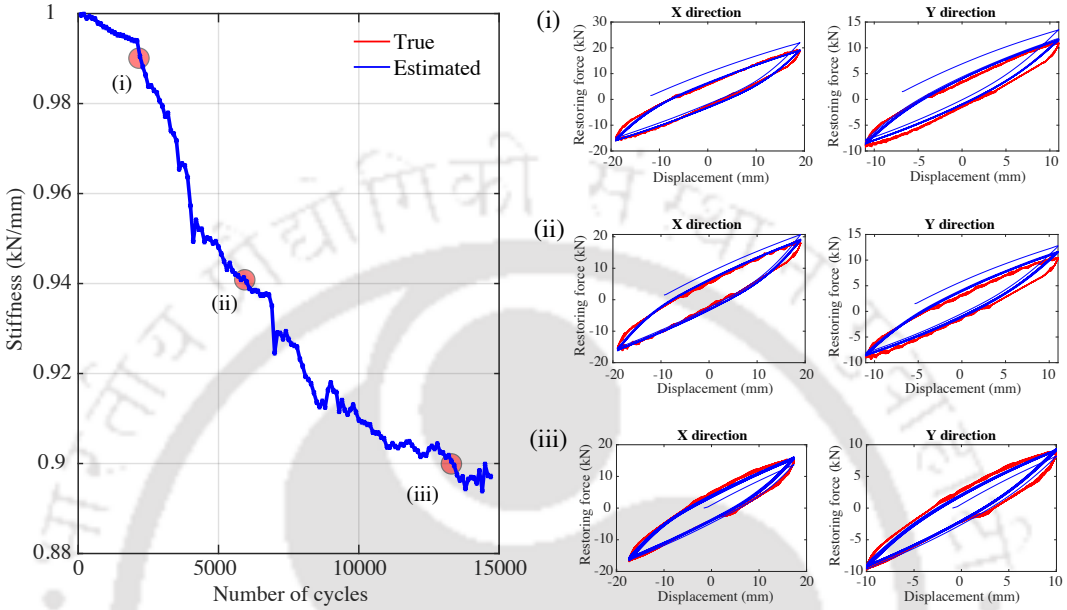


Figure 5.19: Stiffness degradation estimated through CUKF over a long duration

These estimated hysteresis loops are used to evaluate the energy dissipated at each cycles which is further used to develop the damage index. The damage index in the case of fatigue is defined as the cumulative sum of the damage coefficient which is the damage associated with each cycle. The concept for the development of the damage index has been taken from the work of J.B. Mander which defined the damage as the ratio of the maximum stress at the load reversal point in the successive cycle with respect to the first cycle (Mander et al., 1994). Further, the damage associated with the fatigue is proportional to the amplitude of shear strain and inversely proportional to loading frequency up to 20Hz Guennec et al. (2013). Hence, the damage index for the study of specimens tested under varied frequency and shear strain for fatigue is defined as:

$$\text{Damageindex} = \sum_{i=1}^{N} \left(\frac{\tau_i}{\tau_0}\right)^p \left(\frac{f_0}{f_i}\right)^q \frac{\varepsilon_k^i}{\varepsilon_0^i} \tag{5.23}$$

Where f_0 is the referenced input frequency and τ_0 is the referenced shear strain and the $\varepsilon_k^i/\varepsilon_0^i$ is the energy ratio in which ε_0^i is the energy dissipated during first cycle and ε_k^i is the energy dissipated in next successive cycles for a particular input frequency f_i and shear strain τ_i . The term p and q are the modifiers which can vary from system to system. In this case, the modifiers are assumed to be 1 and the referenced input frequency and shear

strain are 0.2 Hz and 18%. The plot of the damage index with the number of cycles and the log-likelihood function is presented in Figure 5.20. It can be clearly observed that there are three change points and each one is associated with the local maximum of the log likelihood function. Further, for each time instant of prognosis, the log likelihood profile is checked, and change point is identified, and accordingly adaptive mean function is optimized. The Log-likelihood function and observations of the prior training are presented in Figure 5.21.

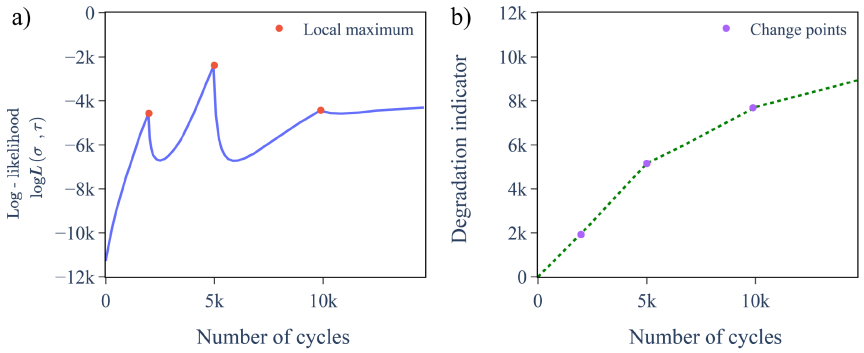


Figure 5.20: (a) log-likelihood profile showing the change point due to presence of various local maximum (b) damage index

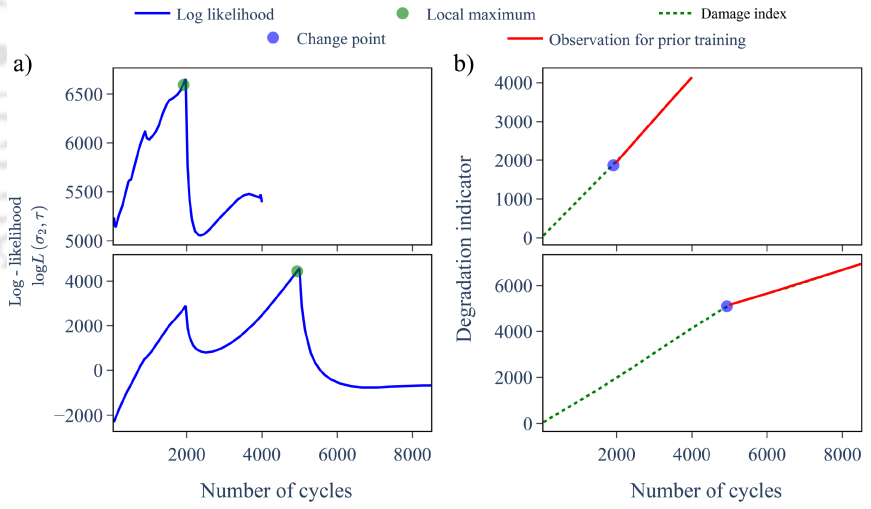


Figure 5.21: (a) Log likelihood profiles and (b) change points in the damage index observed at different time instant of prognosis

The Gaussian process regression is used for the surrogate modelling of the degradation indicator and prognosis has been carried out based on the degradation. As mentioned earlier, different cases like BGPR, and EGA-BGPR are used as the surrogate model and prediction is carried out. For the experimental data, the damage index looks quite simpler and linear basis functions are used in case of BGPR and EGA-BGPR. The similar results of prognosis are obtained as in case of numerical example presented in previous chapter and the maximum value of likelihood value achieved for the basis-GPR through EGA is better. The result of prognosis for the test instant of 4000 cycles is presented in Figure 5.22

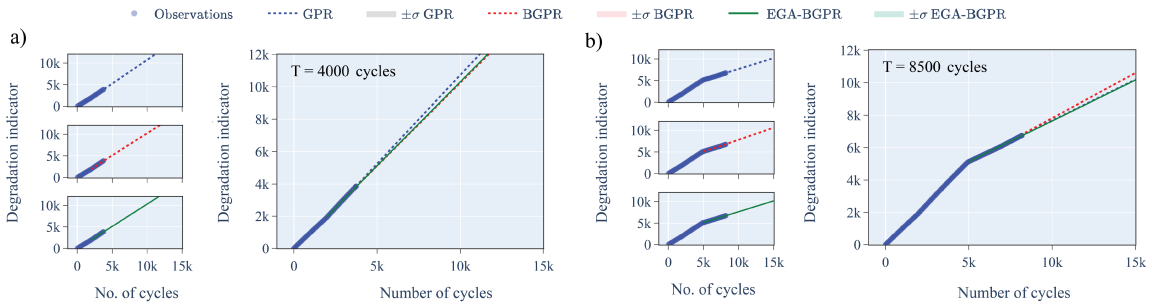


Figure 5.22: Prediction through GPR, basis-GPR, EGA-BGPR at (a) 4000 cycles, and (b) 8500 cycles

Further, the initialization of the population is the crucial step in the process of genetic algorithm. Here, four bitstring sizes are selected for study purpose, i.e., 4 Bits, 8 Bits, 12 Bits and 16 Bits to generate the population. Thus, four cases are studies for the local and global convergence of genetic algorithm. The population generated using 4 bitstring which has the minimum entropy is not able to provide global convergence and therefore should be rejected. The results in Figure 5.23 shows that the maximum fitness value is achieved 16 bitstring in case (a) and 16 bitstring in case (b). Moreover, the results for the 12 bitstring are also closer. Further, the comparison of the prognosis results with ground truth for next 2000 cycles are carried out on the basis of fitness value, rmse error and relative percentage error, presented in Table 5.5. The metrics obtained for the experimental study in Table 5.5 supports the effectiveness of the proposed framework of EGA-BGPR for the prognosis over other conventional BGPR framework.

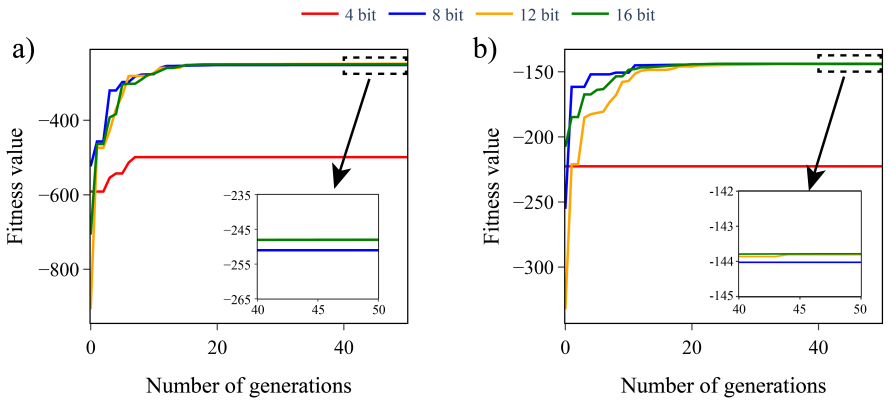


Figure 5.23: Fitness value obtained through EGA using different bitstring sizes for (a) 4000 cycles, and (b) 8500 cycles

Table 5.5: Comparison of the results for experimental study, obtained through entropy assisted genetic algorithm in terms of fitness value (maximum likelihood value achieved), rmse error and relative percentage error

Metrics	Algorithm	Case (a) 4000 cycles	Case (b) 8500 cycles	Case (c) 13000 cycles
Fitness value	BGPR	-393.428	-481.910	-140.228
	EGA-GPR	-247.927	-143.775	-111.470
RMSE	BGPR	1.153	1.012	0.381
	EGA-GPR	0.996	0.643	0.287
Relative error (%)	BGPR	16.70	20.34	9.80
	EGA-GPR	14.22	13.11	7.4

5.7 Conclusion

This chapter demonstrates the efficacy and robustness of the proposed two-stage Constraint Unscented Kalman Filter (CUKF) approach for system identification in both non-deteriorating and deteriorating systems with complex biaxial hysteresis behavior. In non-deteriorating systems, numerical case study highlight the CUKF's ability to accurately estimate hysteresis parameters and system dynamics, outperforming the unconstrained UKF by effectively utilizing bound constraints on model parameters.

For deteriorating systems, the framework successfully captures stiffness degradation due to fatigue crack growth. By employing a coupled response degradation model, the two-stage approach accurately estimates degrading parameters in complex, interacting biaxial hysteresis systems. The experimental validation, where the two-stage CUKF reliably handled responses from a base isolator system at various orientations, further confirms the algorithm's robustness and strengthens its potential for real-world applications of structural health monitoring in systems with coupled biaxial hysteresis.

Having established the two-stage CUKF's capability for accurate system identification, even in complex and deteriorating scenarios, the next logical step is to leverage this information for robust and reliable prognosis. The estimated hysteresis loops are utilized to extract energy dissipated and formulation of the damage index. This long term degradation indicator is further modelled using GPR for the prognosis using proposed EGA-BGPR framework. The results presented highlights the efficiency of the proposed framework in prognosis of the real-world structures considering the uncertainties involved in the process as well as measurement. The next chapter discussed another application to the proposed framework on the aluminium lap joint specimen.

Chapter 6

Monitoring and prognosis of fatigue crack width in lap joints using acoustic waves

6.1 Overview

This chapter presents a comprehensive experimental case study on the monitoring and prognosis of fatigue crack growth in lap joints. Here. we investigate the effectiveness of acoustic wave-based features for detecting and quantifying crack width in the current state through Bayesian neural network. Leveraging the integrated probabilistic framework developed in Chapter 4 for prognosis, the EGA-BGPR methodology is employed to predict crack width evolution of the lap joints.

6.2 Background

The study of damage quantification and its subsequent effects on structural integrity and remaining useful life (RUL) estimation has gained significant attention in recent years within the field of non-destructive testing (NDT) and prognostics health management (PHM) of engineering systems. Data-driven techniques analyze signals from both acoustic sensors and ambient vibrations, providing valuable insights into underlying structural damage (Mutlib et al., 2016; Farrar and Worden, 2012*b*). These approaches are founded on the principle that even unobservable damage mechanisms cause measurable changes in structural properties such as stiffness, damping, or crack width. By tracking these changes through system parameters (Yang and Ma, 2003; Yang et al., 2006), signal processing (Sohn and Farrar, 2001; Nair et al., 2006), and machine learning techniques (Rezaeianjouybari and Shang, 2020; Xu et al., 2023; Salehi and Burgueño, 2018; Ochella et al., 2022; Prakash et al., 2021), a detailed assessment of structural health is possible. Guided wave technologies, particularly using piezoelectric sensors, show great promise for localized damage detection. The low-cost of these sensors and high-efficiency make them ideal for structural health monitoring (SHM), especially in aircraft (Giurgiutiu et al., 2002; Ihn and Chang, 2004). Lamb waves, guided waves in plate-like structures, offer several advantages including strong penetration, low attenuation when propagating long distances, omnidirectional spread, and sensitivity to

various forms of damage (Kim et al., 2010; Kudela et al., 2008; Farrar and Worden, 2007; Giurgiutiu, 2007). Damage detection with Lamb waves focuses on how discontinuities alter wave characteristics (Viktorov, 1967; Lu et al., 2008). However, analyzing the complex interaction of Lamb waves with damage or boundaries poses significant challenges, making the extraction of crack-related features a non-trivial task.

While several methods exist for Lamb-wave-based damage detection (Wang et al., 2004; Fromme and Sayir, 2002; Clézio et al., 2002; Ihn and Chang, 2004), robust quantification of uncertainties inherent in sensor measurements is crucial for reliable diagnostics and prognostics. These uncertainties can arise from various sources. Bayesian neural networks (BNNs) have emerged as powerful tools for modeling uncertainty and preventing overfitting in complex deep neural networks (DNNs) (Kendall and Gal, 2017; Bishop, 2006).

However, traditional BNNs lacked a systematic framework for integrating both aleatoric and epistemic uncertainties, hindering their full potential. Combining aleatoric uncertainty modeling with BNNs gives rise to probabilistic Bayesian neural networks (PBNNs), which provide distributions instead of single output predictions. This chapter introduces a framework which integrates the maximal overlap discrete wavelet transform (MODWT) and probabilistic Bayesian neural network (PBNN) for crack size quantification. The advantages of the proposed framework over other deterministic data-driven machine learning approaches are as follows:

1. It is probabilistic in nature; hence it provides a distribution of the estimated crack width with uncertainty quantification, which inherits the influence of uncertainty in measurement.

2. It is a data-driven technique that takes advantage of the few measurements from a field experiment in conjunction with Bayesian inference to train the network as compared to the other data hungry neural network architectures. This makes the algorithm ease of use in situations of data sparsity as well as it prevents network architecture from overfitting.

The unseen damage detection data from fatigue testing on a realistic lap-joint component is used as a validation to verify the effectiveness of the proposed method. This research uses the dataset available online as a part of 11th Annual Conference of the PHM Society 2019 Scottsdale, AZ. (Peng et al., 2015; He et al., 2013). The proposed framework on this dataset shows excellent capabilities to identify the probabilistic crack width accurately. Lastly, this study extends itself to the prognosis of the fatigue crack growth through the proposed EGA-BGPR algorithm.

6.3 Methodology

This chapter introduces an automated framework using maximal optimal discrete wavelet transform and probabilistic Bayesian neural networks to account for the measurement and process uncertainty in the experiment and thus, maps the Lamb wave signatures to estimate the probabilistic crack width in the aluminum lap-joint. The proposed methodology is applied to specimens consisting of two aluminum panels, which are joined by rivets and subjected to fatigue. These panels are examined in parallel through Lamb wave for the development of crack width. Thus, we have set of acoustic response for both healthy/baseline and degraded condition. The details of the complete datasets are provided in Experiment section and can be found in following literature (He et al., 2013; Peng et al., 2015; Yang et al., 2016). The very first challenge here is that baseline signal varying with each specimen depending on the quality of lap joint, dimensions and thickness of panels. This issue is resolved through residual feature reconstruction which involves decomposition of the baseline and damaged responses into different components, selecting the dominant components and calculating the change in the dominant feature space. This decomposition is achieved through empirical mode decomposition of both healthy and damaged condition Lamb wave signals. The details are presented below:

6.3.1 Residual feature reconstruction using empirical mode decomposition (EMD)

The acquired Lamb wave responses contain the feature that depicts crack width condition of the lap joint along with noise content. Thus, the signal needs to be filtered for noise removal and extraction of necessary component. The empirical mode decomposition (EMD) is one of the filtering techniques that decomposes the components of the various different frequencies of signal to provide a single-component signal (Huang et al., 1998). The extracted singular-components are called Intrinsic mode functions (IMFs). Thus, each IMF constitute an important characteristic of the signal. The necessary IMFs should be selected to extract the dominant components of the signal. The Lamb wave response $x(t)$ is thus represented as a summation of true signal $s(t)$ and white noise $w(t)$.

$$x(t) = s(t) + w(t) \tag{6.1}$$

The fundamental of EMD is that it filters out functions for a time domain signal, which forms a complete and orthogonal basis for the original signal locally. In general, the neighboring components could have data sections that carry the same frequency at different time durations. However, locally, any two components should be orthogonal for all practical purposes for the extraction of IMF. Thus, the resultant signal is called Intrinsic mode functions (IMFs).

$$x\left(t\right) = \sum_{i=1}^{m} \mathrm{IMF}_i + \mathrm{Residual} \tag{6.2}$$

c Where, m is the total number of IMFs. The dominant IMFs are selected, which has the maximum correlation with the original response. Thus, we obtained the dominant IMFs or say, modal features for both the baseline and damaged condition. Finally, the change in IMFs is observed and residual feature reconstruction is carried out to obtain a filtered residual which contains the properties of both baseline as well as damaged condition responses. This residual feature response is now examined for both temporal and spectral analysis.

Since, the residual feature response has both frequency and temporal characteristics that shows variations with development of crack width. Therefore, it is important to identify a feature indicator that contains both characteristics. One such transformations is Fourier transform that performs time-frequency analysis simultaneously, however, it is limited to the stationary signal. Considering uncertainties and abrupt changes due to crack width, non-stationarity is present in the output response. The improved version of Fourier transform is wavelet transform which takes care of non-stationarity and perform better locally over the signal. It can be continuous and discrete in nature. The discrete wavelet transform (DWT) decomposes the original signal into approximation and detailed coefficients through filtering and down sampling. It uses Mallat's algorithm for the practical applications, which consists of low-pass filter (scaling function) and a high-pass filter (wavelet function). Thus, the signal is decomposed into multiple levels, and each level is characterized through its approximation and detailed coefficients. However, it has certain limitations with the signal length as DWT is incapable of handling shift-invariance and has narrow frequency resolution, thus restricting its application to the fixed signal length with an integer multiple of a power of two (Cornish et al., 2006; Percival and Walden, 2000; Nason, 2008). Here, an improved version of DWT called maximal overlap discrete wavelet transform (MODWT) is introduced for the training feature extraction. The training features will be used to train the probabilistic deep learning algorithm which is discussed in subsequent sections. The complete details for the development of training feature dataset are discussed in next section.

6.3.2 Maximal overlap discrete wavelet transform (MODWT)

The residual feature response is transformed into wavelet coefficients through maximal overlap discrete wavelet transform. An important characteristic of MODWT is that it can be properly defined for arbitrary signal length and achieves redundancy through an oversampled representation which enables more accurate statistical analysis (Percival and Walden, 2000). The MODWT divides the frequency band of the residual feature response into scaling and wavelet coefficients using low- and high-pass filters, that is, scaling and wavelet filters. Let $X = X_t$ where $t = 0, \ldots, T-1$, be the residual feature response obtained through experiment, and the wavelet and scaling filters at the j^{th} decomposition level are denoted as $\tilde{h}_{jl}$ and $\tilde{g}_{jl}$, respectively. By appying Mallat's algorithm, the scaling

and wavelet coefficients are calculated as follows:

$$\widetilde{W}_{j,t} = \sum_{l=0}^{T-1} \tilde{h}_{j,l} X_{t-l \bmod T}, \quad \widetilde{V}_{j,t} = \sum_{l=0}^{T-1} \tilde{g}_{j,l} X_{t-l \bmod T} \tag{6.3}$$

where $j = 1, 2, ..., J_0$ is the level of wavelet decomposition, and 'mod' denotes the remainder of dividing two numbers. Here, $\left\{\tilde{h}_{j,l}\right\}$, and $\{\tilde{g}_{j,l}\}$ are the j^{th} level MODWT high-pass filter $\left(\tilde{h}_{j,l} \equiv h_{j,l}/2^{j/2}\right)$ and low-pass filter $\left(\tilde{g}_{j,l} \equiv g_{j,l}/2^{j/2}\right)$ and J_0 is the highest decomposition level. The wavelet coefficient vectors $\widetilde{\boldsymbol{W}}_j$ and scaling coefficients vectors $\widetilde{\boldsymbol{V}}_{\text{j}}$ of MODWT are written as:

$$\widetilde{\boldsymbol{W}}_j = \left\{\widetilde{W}_{j,0}, \widetilde{W}_{j,1}, \ldots, \widetilde{W}_{j,T-1}\right\}, j = 1, 2, \ldots, J_0 \tag{6.4}$$

$$\widetilde{\boldsymbol{V}}_{\text{j}} = \left\{\widetilde{V}_{\text{j},0}, \widetilde{V}_{j,1}, \ldots, \widetilde{V}_{j,T-1}\right\}, j = 1, 2, \ldots, J_0, \tag{6.5}$$

where $\widetilde{\boldsymbol{V}}_{\text{j}}$ and $\widetilde{\boldsymbol{W}}_j$ are related to the smallest and highest frequency components of the original signal. The scaling coefficients are input to the high and low pass filters to obtain the next level of decomposition. Thus, we get the series of wavelet coefficients which are bind together to form the training feature matrix. The residual feature response are decomposed to yield the wavelet and scaling coefficients, then these wavelet coefficients that contains the temporal and spectral features will form the training feature matrix for the probabilistic deep learning framework. This training feature matrix contains the crack width dependent acoustic properties. Figure 6.1 presents the pictorial representation of wavelet decomposition and development of training feature matrix.

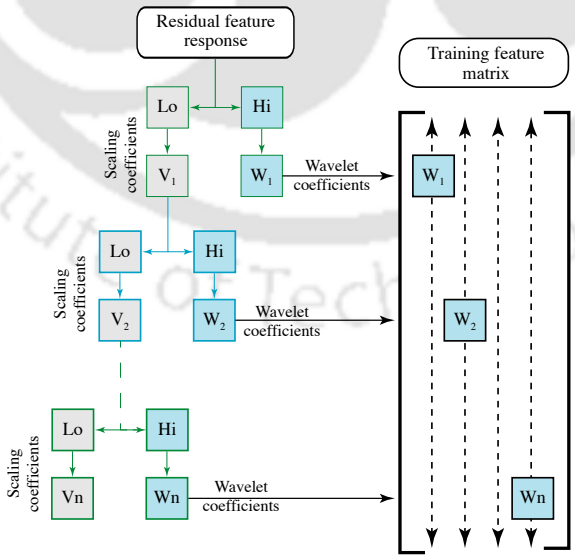


Figure 6.1: Basic mechanism of MODWT and preparation of feature matrix

Further, the inverse MODWT uses the selected wavelet and scaling coefficients can be used to reconstruct the desired signal. In the current work, it is not needed to reconstruct the signal

back, as the wavelet coefficients will serve as the training feature matrix which will be used for the training of machine learning surrogate model. It is important to note that diverse structural geometries and varying working conditions make the reliable acquisition and interpretation of Lamb wave responses difficult. Moreover, it introduces the aleatoric and epistemic uncertainty in the dataset. Thus, the machine learning model should be capable of defining this uncertainty in the parameters to account for measurement and process noise. This study utilizes the probabilistic Bayesian neural network (PBNN) to map the training feature matrix with crack width. This PBNN is capable of not only handling uncertainty, rather it can separately provide the variance due to epistemic as well as aleatoric causes. The details of the probabilistic deep learning model, methodology for capturing the uncertainty, and optimization methods are discussed in the next section.

6.3.3 Neural network based surrogate model

There are various uncertainties involved in the process, from measuring the acoustic response to obtaining the training feature matrix. The surrogate model which maps the training feature matrix with the crack width should take care of this stochasticity in measurement as well as in process. To begin with, we will first introduce the basics of feed forward artificial neural network, then its limitations and need for the Bayesian deep learning is discussed.

Feed forward artificial neural network

The artificial neural network (ANN) comprises the input layer, hidden layers, and output layer. Each layer constitutes a different number of neuron units depending on the input features and desired output. These neuron units are called nodes and are parameterized by deterministic weights (w) and biases (b), therefore not capturing the stochastic relation between input features and output (Wu and Feng, 2018). In ANN, the input feature (x) is first multiplied by its weights prior reaching to the hidden layer node. The weighted sum $(f_{w,b})$ of the weighted parameters collected at the nodes is computed. The weighted sum can be defined as

$$f_{w,b} = wx + b \tag{6.6}$$

Subsequently, the weighted sum is transmitted to the next layer using the activation function and finally to the output layer as the target data. The activation function $\phi(\cdot)$ can be hyperbolic tangent or sigmoid function which related the output as:

$$u = \phi(f_{w,b}) \tag{6.7}$$

The weights and bias of the ANN are optimized through a backpropagation algorithm in which the model loss is propagated back and is minimized using the gradient descent

method. The loss function $((\mathcal{L}\ (w,b))$ for this is given by

$$\mathcal{L}\ (w,b) = \sum_{k=1}^{N} \|u - \widetilde{u}\| \tag{6.8}$$

where N is the total number of training data points available, u is the observed response and $\widetilde{u}$ is the estimated response. Even though ANN architectures are quite capable of learning complex engineering phenomena, there exist crucial issues related to interpretability, generalizability, overfitting, and demands for significant data. These factors limit their utility and computational tractability for data-driven models. It is to be noted that for the current work, the data collected from the experiment is very sparse. Moreover, ANN is quite incapable of predicting variation for the stochastic input. Thus, the direct application of data-driven ANN will lead to overfitting and thus is inconceivable to get adequate results.

Sources of uncertainties and their quantification

Under the acoustic sensing technology, challenges arise when Lamb waves interact with damage or boundaries, making it complex to extract features associated with cracks due to mixed mode. To address such a challenge, researcher proposes the damage index that directly maps the signal characteristics to the associate anomaly. For example, Ihn proposed a damage index defined as the ratio of monitoring signal energy to baseline signal energy (Ihn and Chang, 2004). However, limitations in the accessible monitoring data and inherent shortcomings of the sensor technology introduce two types of uncertainties: aleatoric uncertainty, reflecting the noise pollution in data collection and transmission, and epistemic uncertainty, reflecting the ignorance of model property. A step-by-step procedure for accounting the uncertainties through probabilistic transformation is provided in Chapter 3. In this work, we have sparse training samples in the dataset due to limited experiments. Hence, considering the limitations of the artificial neural network, and importance of the uncertainties arises due to limited dataset and measurement errors, this study uses probabilistic Bayesian neural network for mapping the training feature matrix to the crack width in the aluminium lap joint.

Probabilistic Bayesian neural network

Bayesian neural networks (BNNs) address the challenge of epistemic uncertainty (uncertainty about the model itself) arising from limited training data. Unlike traditional neural networks with fixed weights and biases, BNNs treat these parameters as probability distributions. Prior distributions are defined, and training involves refining these distributions to capture model uncertainty. To incorporate aleatoric uncertainty (inherent data noise), probabilistic Bayesian neural networks (PBNNs) further augment the output layer to model a distribution, rather than a single prediction. The key advantages are (1) robustness to overfitting, (2) enhanced uncertainty modelling, and (3) output variance

analysis. The details can be found in the Chapter 3. To apply PBNNs, Gaussian priors are assumed for the weights, representing epistemic uncertainty. Since the posterior distribution cannot be evaluated directly for complex models, methods like Hamiltonian Monte Carlo (HMC) or variational inference are used to approximate it. The details about HMC and VI can be found in Appendix A, and B. This approach allows PBNNs to quantify uncertainties, making them well-suited for scenarios where data is limited and decision-making requires risk assessment. In this work , the training feature matrix is mapped to the crack width through PBNNs. Once, the PBNN model get trained, it will predict the mean estimation of the crack width for test specimen as well as the prediction distribution. Thus, we have output mean $\left(\mu_y^*\right)$ and total output variance $\sigma_{total}^2\left(y^*\right)$ as final prediction. The detailed step-by-step procedure for generating the output results are presented in next section.

6.3.4 Proposed probabilistic deep learning algorithm

The overarching goal of this work is to leverage the capabilities of MODWT and PBNNs to estimate the probabilistic crack width under the condition of sparse and noisy data and then leverage the EGA-BGPR framework for prognosis. The proposed approach is driven by the requirement for an efficient parsimonious model that evades overfitting, utilize the time-frequency variation and balances out the uncertainties arise due to the limited data and noisy measurement. Unlike deterministic deep learning, the proposed probabilistic deep learning framework is stochastic and is designed particularly to handle a ubiquitous class of sparse data perturbed by noise and uncertainty. The proposed deep learning framework is shown in Figure 6.2.

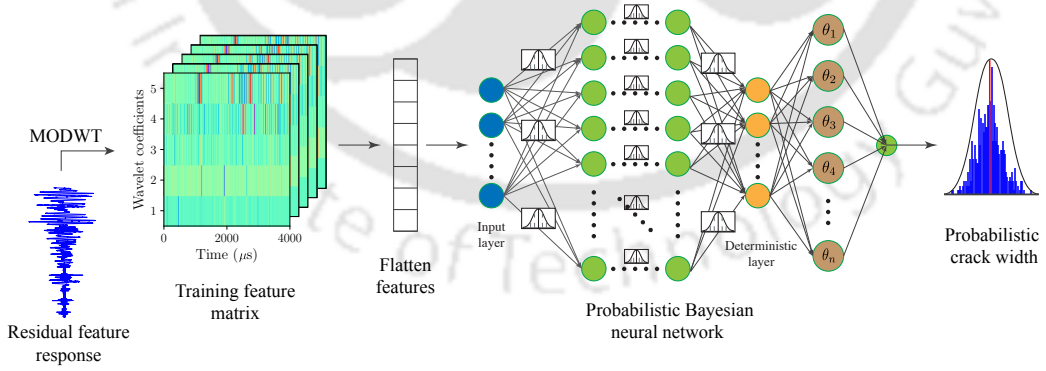


Figure 6.2: Illustration of probabilistic deep learning model

Here, the objective is to estimate the crack width in the aluminium lap joint subjected to fatigue and the corresponding uncertainty arises in the real-life scenario. The presented approach employs MODWT and PBNNs in a single framework. The proposed deep learning framework is proposed to handle the challenges inherit due to interaction of Lamb waves with the interfaces generating mix mode, thereby very difficult to extract feature directly

with some physical meaning. This framework leverage the maximal overlap discrete wavelet transform for the feature extraction and is directly mapped to the crack width through probabilistic Bayesian neural network to account epistemic and aleatoric uncertainties.

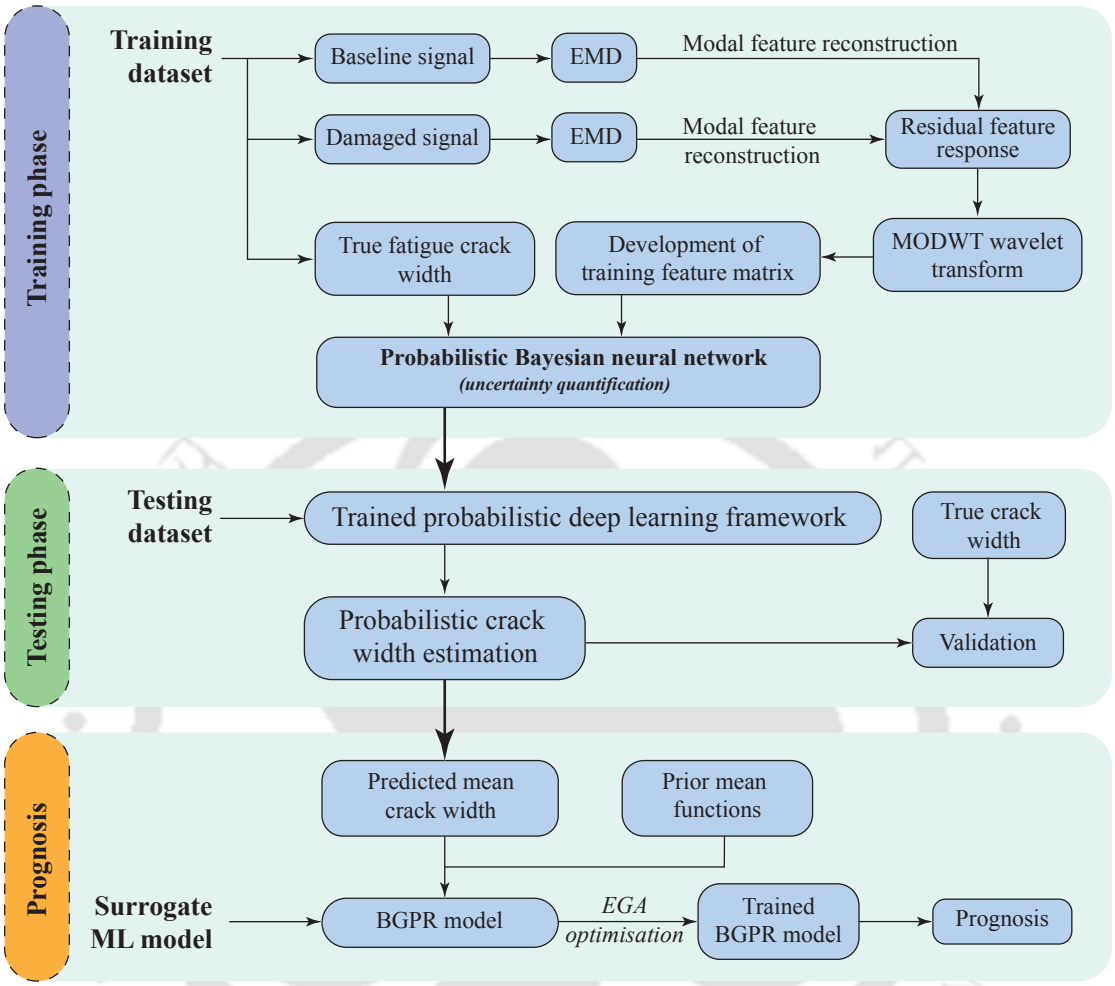


Figure 6.3: Flow chart depicting the step-by-step procedure for the estimation of probabilistic crack width and prognosis of the crack width using probabilistic deep learning

The acoustic response obtained through experiment is cleaned and MODWT is performed to extract the wavelet coefficients. Five level decomposition is implemented here because adding more levels will increase computational cost and reducing level will destroy features. These wavelet coefficients forms the training feature matrix, which captures the time frequency snapshot information related to the crack width in aluminium lap joint. This information is also amalgamated with noise which is taken care through PBNNs. The conventional data-driven ANN is not competent in estimating the crack width from the sparse, noisy data collected from field/ real experiments. Further, the ANN fails as the data available is sparse. Moreover, neural networks have considerable potential for overfitting data and has no place for the uncertainty quantification. The PBNNs uses Bayesian inference for the uncertainty quantification as well as it prevents neural network architecture from overfitting. The uncertainties inherit in the framework help in estimation

of output distribution instead of a point estimate. Finally, output is sampled using posterior inference models. Thus, we have both mean estimate of crack width and variance of the prediction. This mean estimate is further used for the prognosis. The overall framework consists of three main steps: (1) training of the probabilistic deep learning model, (2) testing of the probabilistic deep learning model and predicting probabilistic crack width estimates, and the last step is to (3) carrying out prognosis of crack width using EGA-BGPR approach. The complete framework is illustrated in the Figure 6.3.

6.4 Experimental dataset

In the current work, the fatigue dataset containing total 8 specimens (naming from T1 to T8) out of which observations from five specimens (T1 to T6 except T4) are used for the training of the framework and remaining T4 specimens is used for validating and testing the framework. The data from specimen T7 and T8 are not used in this study. A hydraulic material testing machine working at 5Hz at room temperature conducts the fatigue testing. Both constant amplitude loading and variable amplitude loading cases are studied. It is worth mentioning that specimen T8 has tested under fatigue cycles with varied amplitude while all other specimens are tested under constant cycles. During the testing, the optical microscope is used to identify the location of the crack appearance. Further, this variability will only affect the prognosis because probabilistic crack width estimation is dependent only on the acoustic responses rather than fatigue cycles. This data set consists of measurements from piezoelectric (PZT) sensors that are mounted on aluminum specimens that were tested to fatigue conditions. The failure will occur due to formation of cracks. Figure 6.4 illustrated the sensor mounted on the aluminium lap joint specimen subjected to the fatigue cycles. The

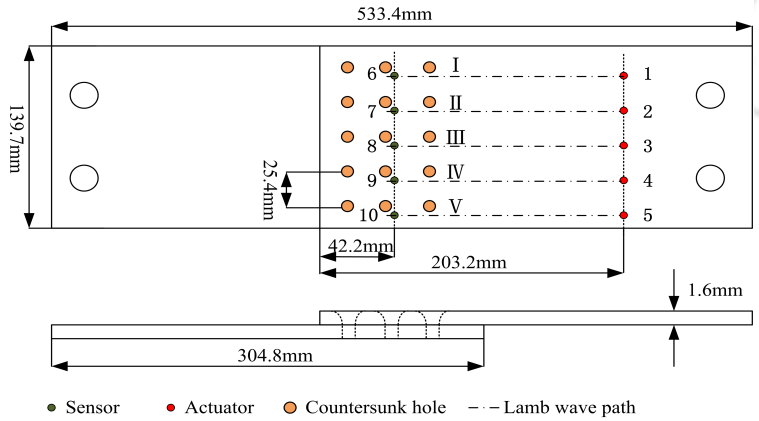


Figure 6.4: Illustration of aluminum lap-joint used for fatigue testing and condition monitoring (Peng et al., 2015; Liu and Peng, 2019)

specimens consisting of two aluminum panels, and they are joined by three rivet rows by five rivets. The riveted panels are made of 1.6mm-thickness aircraft-grade 2024-T3 aluminum sheets. Surface-bonded piezoelectric (PZT) ceramic wafers are used as actuators to generate Lamb waves as well as sensors to receive Lamb signal. The pitch-catch configuration is

used for Lamb wave damage detection. According to the existing experimental data, the crack usually initiates at the countersunk hole in the first row, it will connect these holes and finally break the specimen. Therefore, only the countersunk holes in the first row are chosen for inspection. The geometry of the lap joint component and sensor network configuration is shown in Figure 6.4. The more details regarding the experimental setup, testing protocols can be found in the following literature (Peng et al., 2015). Further, the stress amplitude for the constant cycles and variable cycles is presented in Figure 6.5.

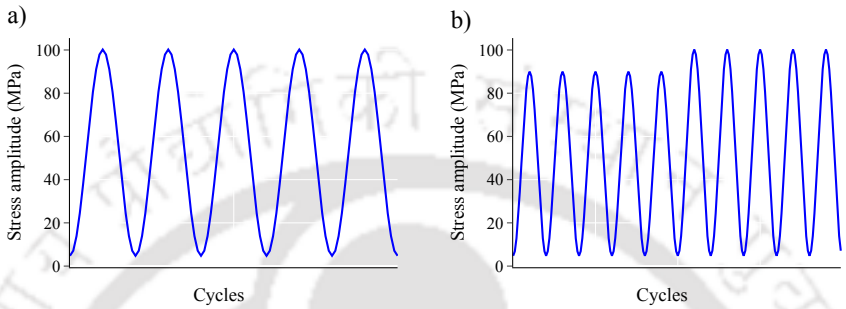


Figure 6.5: Illustration of constant stress amplitude cycles (4.77 MPa to 100.21 MPa) and varying stress amplitude cycles with change point at 90 MPa

For the specimen tested under stress cycles presented above, an acoustic Lamb wave was sent and received at regular intervals to monitor the condition of the crack size. Simultaneously, the ground truth of the crack width is measured optically. Here, we will present the results from the training specimens as well as test specimens. The acoustic responses observed for the training specimen T2 is shown in Figure 6.6.

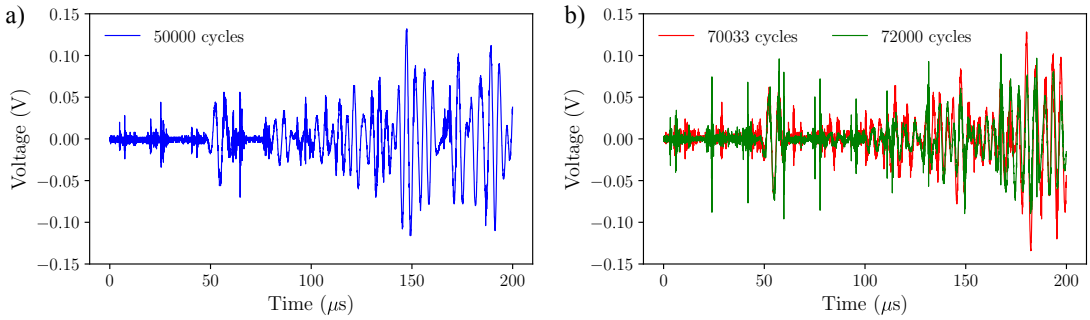


Figure 6.6: Illustration of time domain responses of the Lamb wave for the training specimen T2 for (a) uncracked condition and (b) cracked condition

The baseline signal should have a fatigue cycle equal to zero. However, from the available dataset, the signal that has the crack width of zero is referred to here as the baseline signal. From Figure 6.6, it can be seen that the baseline signal has undergone changes due to the appearance of a crack on the surface. It further transforms with the increment in the crack width. Thus, this change is important to train the probabilistic deep learning framework and predict the probabilistic crack width estimate. Moreover, the obtained crack

width estimate will update the mean function of the proposed EGA-BGPR framework and prognosis is achieved. The next section will discuss the results at various stages of the proposed framework.

6.5 Results for probabilistic crack width estimation

6.5.1 Signal processing and development of training feature matrix

The acoustic wave responses obtained during the experiment are cleaned and two sets are created. The responses of the specimens T1 to T6 are considered for the training of the proposed probabilistic deep learning framework and the specimens T7 and T8 are used as the evaluation/test set. The obtained responses for the different specimens for healthy state and damaged state conditions in the frequency domain are presented in Figure 6.7. It can be seen that there is a distinct baseline acoustic response for each specimen giving rise to the complexity in model training. Thus, instead of using a signal for the model training directly, the modal feature extraction is carried out for both the baseline and damaged response. This model features are extracted through empirical mode decomposition. The EMD decomposed the signal into separate mono components for the different frequencies. The dominant components from all extracted intrinsic mode decomposition (IMFs) are selected using cross-correlation with the original acoustic response. The dominant components are nothing but the model feature for the acoustic response. Similarly, the model features are extracted from damaged specimen conditions. It is important to note here that neither model feature of healthy or damaged state response is sufficient enough for developing the proper mapping model to the crack width. Therefore, the dominant IMFs of baseline and damaged state acoustic response are selected for the model feature reconstruction, and then a residual the calculated to observe the model changes. Figure 6.8 and Figure 6.9 illustrate the normalized correlation of the IMFs and the dominant IMFs of the baseline and damaged response for specimen T1.

Further, out of the three IMFs of the residual feature response, it can be seen that high amplitudes are observed only in a single IMF, which states that it has mono-components that observe significant change due to the development of the crack width. Therefore, residual feature response can be directly related to the crack width in the aluminum lap joint. This residual feature response is now used as the primary input to the probabilistic deep learning framework. It contains both time-dependent as well as frequency-dependent components which inherently show some correlation with the crack width. A common approach to evaluate the time-frequency analysis is a short-term Fourier transform (STFT). However, it is limited by a fixed resolution, and the width of the windowing function used in the STFT affects how the signal is represented. The choice of window width determines whether the representation emphasizes the ability to distinguish closely spaced

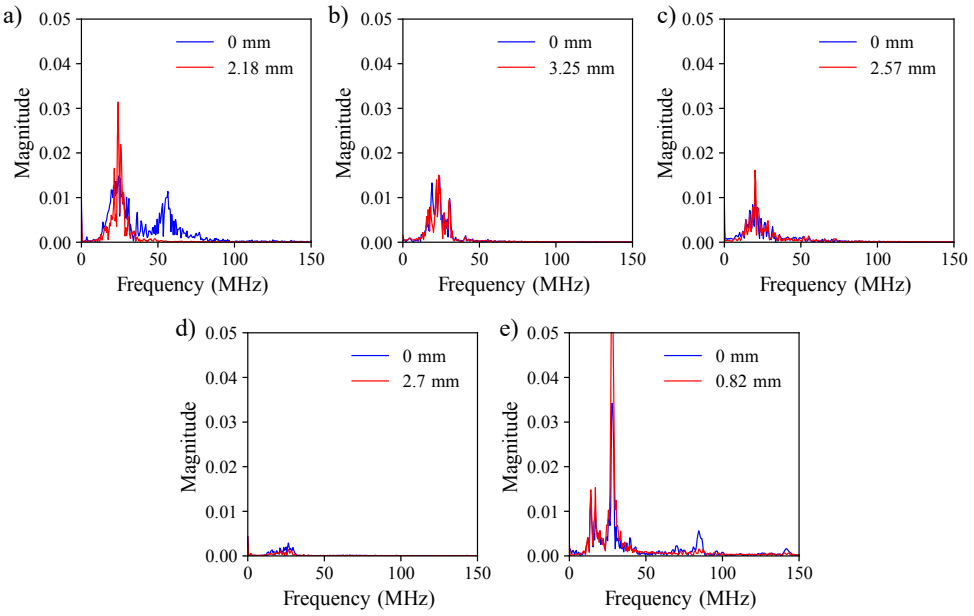


Figure 6.7: Illustration of frequency domain responses of the Lamb wave for all training specimens from (a) T1, (b) T2, (c) T3, (d) T5, (e) T6 for baseline condition (0 mm crack width) and damaged condition (crack width is provided in the respective subfigure)

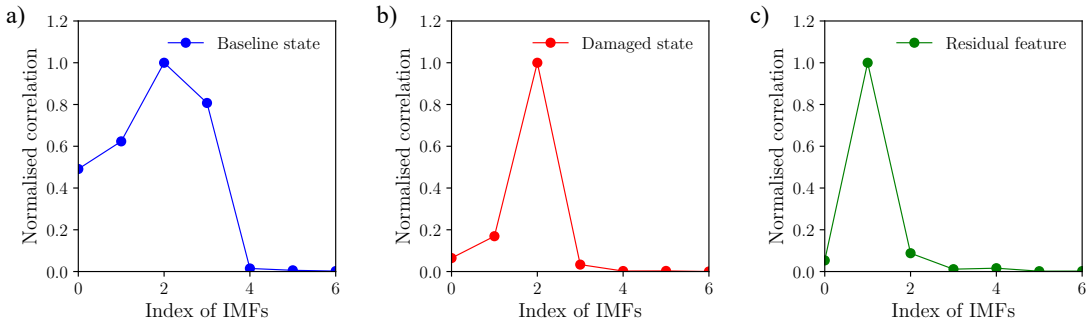


Figure 6.8: Illustration of normalized correlation of the IMFs with the original response of the training specimen T1 for (a) baseline condition, (b) damaged condition, and (c) residual feature

frequency components and capture the timing of frequency changes. This gives rise to the multi-resolution analysis through wavelet transform. This work introduces maximal overlap discrete wavelet transform (MODWT) for the training feature extraction from the residual response. A five-level decomposition is carried out using wavelet transform and the wavelet coefficients at each level constitute the training feature matrix. The wavelet spectrum of the five-level decomposition for specimen T1 at different levels of damage i.e., at various crack widths is presented in Figure 6.10. It can be observed that there are minor changes in the wavelet spectrum as the crack width increases. Now, this wavelet spectrum is passed through a flattened layer which constitutes the direct input to the probabilistic Bayesian neural network.

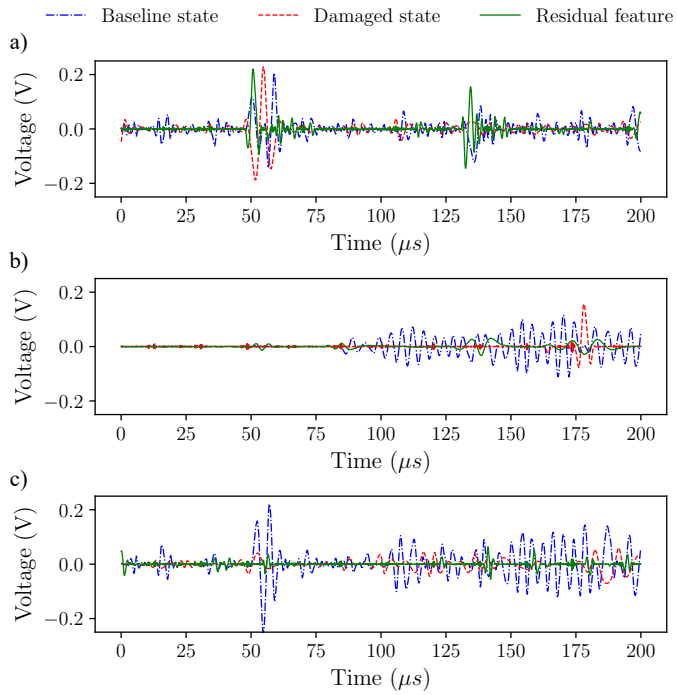


Figure 6.9: Illustration of time domain responses of the Lamb wave for the training specimen T1 for (a) uncracked condition and (b) cracked condition

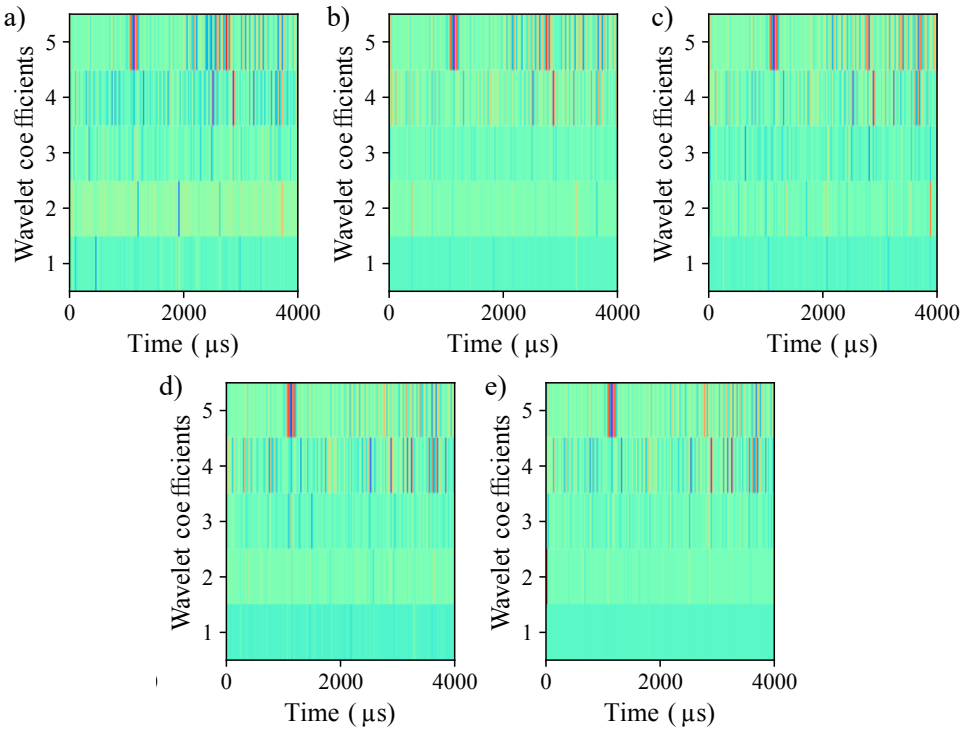


Figure 6.10: Illustration of wavelet decomposition upto five levels for the residual feature response for the training specimen T1 at various crack widths, (a) 2.18 mm, (b) 2.76 mm, (c) 3.51 mm, (d) 4.51 mm, (e) 4.90 mm, (f) 7.46 mm

6.5.2 Development and training of the Probabilistic Bayesian neural network

This chapter proposed a shallow probabilistic Bayesian neural network for the mapping of the training feature matrix with the crack width. It constitutes the input layer with 20000 neurons followed by hidden layers having 10 neurons each. The hidden layer is followed by the aleatoric uncertainty layer and lastly, it has one neuron in the output layer for the mean prediction. However, sampling from the Gaussian distribution is used for the variance prediction. The neural network uses an activation function after the input layer to normalize the input parameters. The activation function used in the study is tangent of hyperbolic function. It should be noted that all the weights and biases are stochastic, and these are simulated using a Gaussian random variable with a mean equal to zero and a standard deviation of 10. For the optimization and training of the PBNN, two approaches are explored in this study, (i) HMC sampling and (ii) Variational inference (VI) method. HMC sampling is the direct approach for the posterior evaluation while VI uses an approximation posterior (mean-field guide function) to simulate the actual posterior. An auto-diagonal normal mean field guide function is used for the variational inference method. The model loss obtained through VI is shown in Figure 6.11.

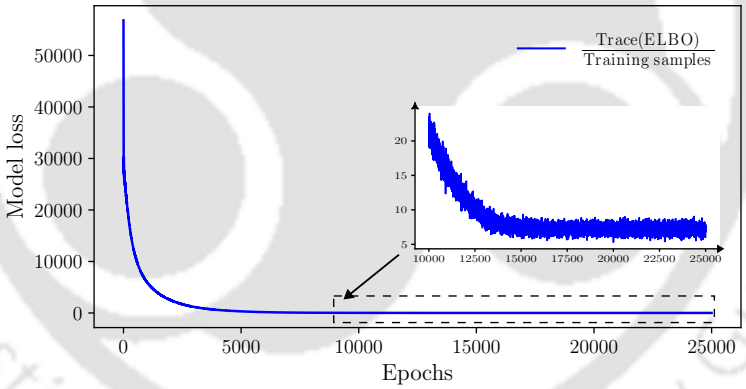


Figure 6.11: Illustration of model loss for the stochastic variational inference method while training model using auto-diagonal Gaussian mean field guide function

Apart from VI, PBNN is trained using the HMC method. It can be seen that model loss calculating through VI is not able to be optimized properly and we have obtained variance in the order of 100. Further, the accuracy obtained from the HMC sampling is more than 97.5 %. A comparison of the optimization results achieved with each neural network throughout the simulations is represented in Table 6.1. It can be seen that the results from the HMC sampling are more accurate as compared to VI optimization. Thus, this work presented the results from the HMC optimization and sampling. The model accuracy of the PBNN is illustrated in Figure 6.12 for the different specimens at various crack widths.

Table 6.1: Comparison of the statistics obtained through 50 simulations for the PBNN model using HMC sampler and VI with Gaussian mean field guide function

Metrics	Training points		Test points	
	HMC	VI	HMC	VI
RMSE	0.805	1.628	0.585	1.917
Q1	2.613	3.778	2.180	3.703
Q3	4.917	3.9287	3.733	3.894
IQR	2.303	0.150	1.552	0.190

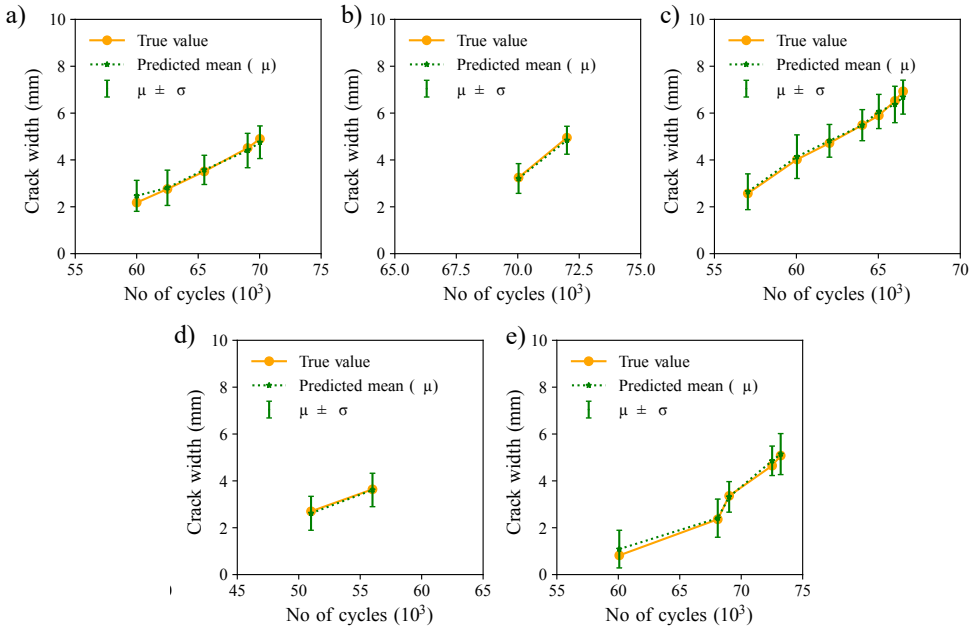


Figure 6.12: Training accuracy plots depicting the model accuracy obtained from the HMC sampling method for different specimens from (a) T1, (b) T2, (c) T3, (d) T5, (e) T6

Figure 6.13 demonstrates the prediction distribution for the particular estimated crack width from each training specimen. As shown in Figure 6.12 and 6.13, the proposed PBNNs framework has shown accuracy and efficiency in estimating crack width in aluminum lap joints while accounting for the various epistemic and aleatoric uncertainties associated with the microcrack density, stiffness reduction, and measurements.

6.5.3 Prediction of the unknown crack width and prognosis

The trained model is now applied to the test dataset to predict an estimate of the crack growth for the unknown specimen T4. Figure 6.14 display the predicted mean and uncertainty distributions for the specimens T4 at different time instant. The presented probabilistic deep learning framework is implemented successfully on the experimental fatigue damage data taken under laboratory-controlled conditions, Further, it is showing high variance depicting the presence of uncertainty due to real-life conditions, making the adoption of both deterministic physics-based models and traditional feed-forward networks unfeasible Sriramula and Chryssanthopoulos (2009). The physics-based models

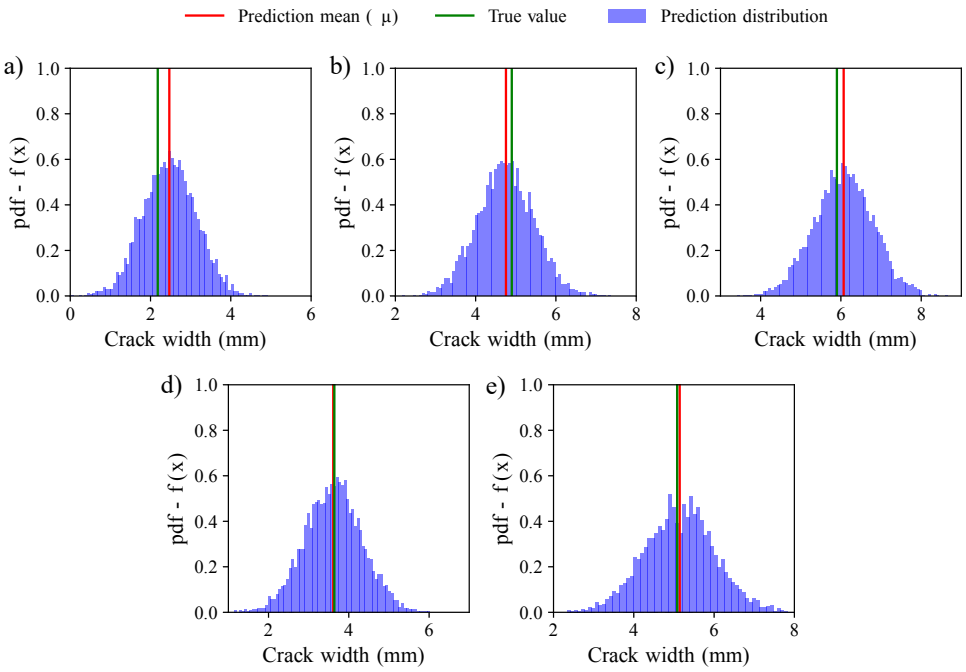


Figure 6.13: Predicted distribution for the particular crack width obtained from the HMC sampling method for different specimens from (a) T1, (b) T2, (c) T3, (d) T5, (e) T6

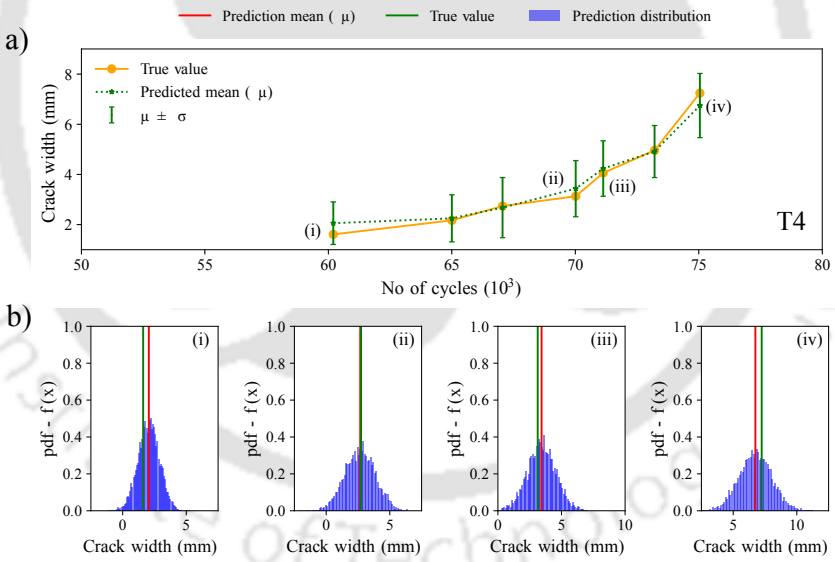


Figure 6.14: Testing prediction results from the HMC sampling method for different unknown specimens (a) T7, and (b) T8

have proved efficiency and predictability only in low-scale structures and highly controlled environments. They are deterministic and do not consider the uncertainty inherent in fatigue damage of composite materials, even in laboratory conditions Further, ANN overfits in case of data sparsity. From the engineering aspect, failing the determine the uncertainty will leads to high safety factors and thus lead to poor utilization. Hence, the proposed probabilistic deep learning framework can be useful for on-board structural

health monitoring systems where damage data is collected and simultaneously predictions with quantified uncertainty can be made during the operation. PBNNs do not overfit the model data and address both types of uncertainty which enhances the model's robustness and provides a more comprehensive understanding of the uncertainties associated with predictions. Hence, the proposed data-driven method allows for the scalability to complex structures and environments, while the uncertainty in the required estimate is quantified.

Results for prognosis using estimated crack width

The probabilistic deep learning framework provides the mean estimate of the crack width and corresponding associated uncertainty. For the prognosis of the crack growth, this study utilizes proposed EGA-BGPR methodology where the mean function are updated according to the mean estimate of the crack width. As the Bayesian neural network returns the distribution of the prediction, we consider the mean estimates of the crack width and this predicted mean of the crack width is plotted against the number of cycles as shown in Figure 6.15

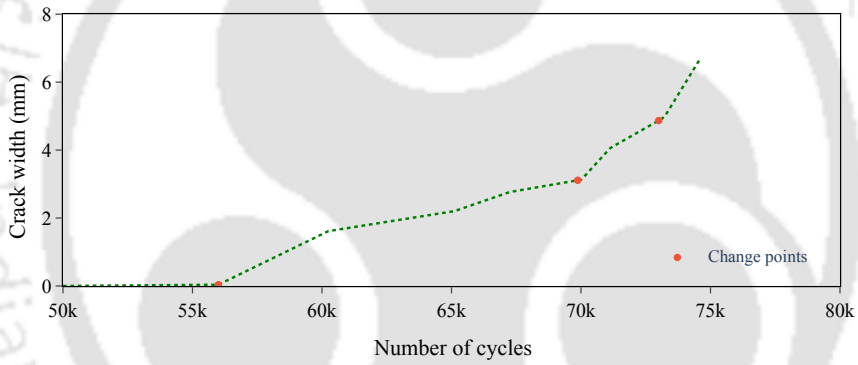


Figure 6.15: Predicted distribution results from the HMC sampling method for different unknown specimens (a) T7, and (b) T8, for the first crack width in series

A log likelihood function is evaluated for identification of change points. The Log-likelihood function and observations of the prior training are presented in Figure 6.16. It can be clearly observed that there are change points and each one is associated with the local maximum of the log likelihood function. Further, for each time instant of prognosis, the log likelihood profile is checked, and change point is identified, and accordingly prior mean function is optimized, and degree of basis function is selected. The Gaussian process regression is used for the surrogate modelling of the degradation indicator and prognosis has been carried out based on the damage index. The result of prognosis for the time instant of 66500 cycles and 73000 cycles are presented for the specimen T4 in Figure 6.17 Figure 6.17 shows the uncertainty bounds obtained during the prognosis. It can also be seen that the mean prediction trend gets closer to the actual crack width values provided in the testing dataset, which indicates the effectiveness of the proposed framework in estimating prognostic uncertainties.

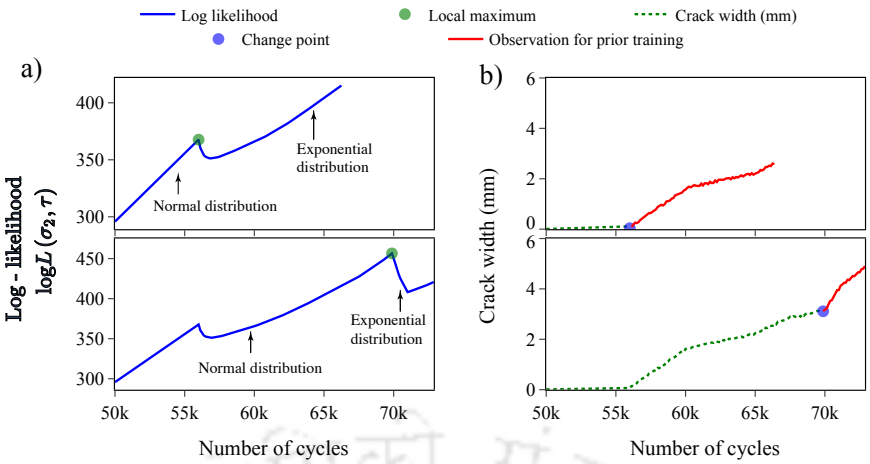


Figure 6.16: Estimation of posterior distribution of Paris law parameters using estimated mean crack width though HMC sampling method

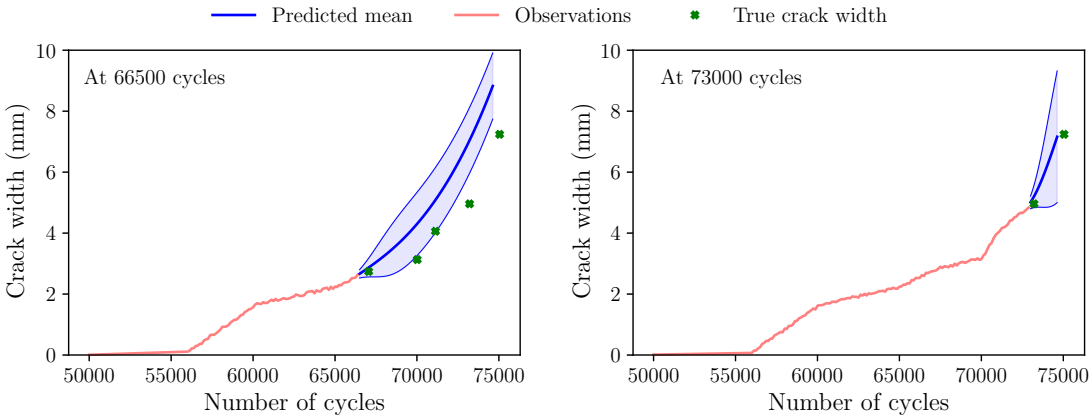


Figure 6.17: Estimation of prognosis using proposed EGA-BGPR for the unknown specimens T4 at 66500 and 73000 cycles. The mean curve follows are correct trend of the true crack width, potentially highlighting the robustness of proposed framework

6.6 Conclusions

This chapter discusses a practical case study of accurate fatigue crack width estimation in aircraft structures, especially aluminium lap joint and extends it to the data driven based prognosis under EGA-BGPR approach. The framework addresses following key areas with limited research:

1. Directly linking Lamb wave signatures to crack width: The model learns this relationship automatically using an automated approach that uses wavelet feature extractor (MODWT), eliminating the need for manual feature engineering.

2. Incorporating uncertainty using probabilistic deep learning: The proposed framework employs a probabilistic Bayesian neural network to quantify both aleatoric (data-driven) and epistemic (model-driven) uncertainties, providing more robust and reliable crack width estimates with confidence bounds. In this particular

case, the results obtained from HMC optimization are found to be more reliable and needs further investigation.

3. Adapting the mean function of the EGA-BGPR framework based on estimated mean crack width for the prognosis.

The obtained results on unseen in-situ data from fatigue experiments on aluminum lap-joint specimens demonstrates the efficacy of proposed framework and paving the way for practical implementation in real-world scenarios. Further, it opens avenues for investigating the applicability of proposed framework to other materials and geometries, potentially leading to more reliable fatigue life prognosis and enhanced aircraft safety.

Chapter 7

Conclusion and future scope

The thesis presents an integrated probabilistic machine learning approach for effective diagnosis and reliable prognosis using measured condition information for degrading civil and aviation infrastructure applications. By comprehensively accounting for measurement noise, modeling uncertainties, and operational variability, the proposed approach provides valuable insights for informed decision-making. A probabilistic ML tool known as Gaussian process regression is employed as the surrogate model for degradation indicator to perform prognosis. To address the challenges of extrapolation and local optima in the GPR model, adaptive mean/basis functions and an entropy-assisted genetic algorithm (EGA) are integrated. Comparative analysis with traditional conjugate gradient optimization demonstrates the superior performance of the EGA-optimized GPR. The feasibility of the proposed methodology in real-world applications is substantiated through practical implementation. Furthermore, the developed diagnostic methods provide case-specific relationships between indirect measurements and degradation indicators. This chapter summarizes the key research contributions, their implications, and potential avenues for future work.

7.1 Significant contributions

The significant contributions of the thesis work are as follows:

1. An integrated probabilistic machine learning framework has been developed and validated for damage detection, prognosis, and remaining useful life (RUL) estimation of components where direct degradation is unobservable. This framework employs a multi-stage degradation modeling approach, where degradation indicators are derived from either statistical measures or system parameters. A change point in the degradation path is identified to shift the adaptive mean function with the current trajectory for extrapolation of GPR model. The parameters of the adaptive mean function and surrogate model are simultaneously evaluated while prognosis and RUL estimation.

2. An information theory based approach on Renyi's entropy assisted genetic algorithm

(EGA) is proposed to optimize Gaussian Process Regression (GPR) addressing the challenges of local convergence and extrapolation. This approach helps the EGA to generate sufficiently diverse population, leading to a higher likelihood of achieving global convergence, as opposed to traditional optimization techniques like the conjugate gradient method which usually converges to local minima. The advantage of this approach is that it results in reasonably good estimates of hyperparameters without showing any premature convergence.

3. The diagnosis methodology developed in this dissertation are general in nature. A damage index based on Bhattacharyya distance using indirect measures is developed while discussing numerical case study. For the diagnosis of real life case studies, two-stage constraint unscented Kalman filter is proposed for evaluating the system parameters in the complex systems like elastomeric rubber bearing exposed to bidirectional excitation. This two-stage framework is capable of handling overconfidence issues while evaluating large variance matrix due to high number of states. Additionally, a probabilistic Bayesian neural network (PBNN) framework is proposed for aviation infrastructure diagnosis, leveraging wavelet transform for automated feature extraction and linking Lamb wave signatures to crack width estimation considering the inherent uncertainties.

4. The proposed method has been successfully applied to elastomeric bearing degradation and the development of crack width in lap joints under fatigue loading. For both applications, the parameters of the degradation surrogate model are estimated using experimental data, adaptive mean functions are re-evaluated when change points occurs and the RUL distribution is updated with monitored condition data.

5. The methodological advancements contribute to a more reliable framework for predicting remaining useful life of degrading systems and significantly enhance prognostic accuracy.

7.2 Conclusions

The conclusions are summarized as follows:

1. The numerical examples and case studies underscore the critical role of integrating prior knowledge in pre-training the adaptive mean function for the proposed framework's optimal performance. When effectively combined with degradation data, this approach yields more accurate and efficient prognosis.

2. The analysis reveals that utilizing more degradation data post-change point leads to timely fault detection, improved model hyperparameter estimates, variance reduction and

reasonably accurate RUL predictions.

3. From the comparison study of entropy assisted genetic optimization with conjugate gradient method for maximum likelihood evaluation using numerical degradation data demonstrate the superior performance of proposed optimization. While significant improvements have been achieved, further enhancements can be expected with more complex trajectories.

4. The case study demonstrated that the proposed degradation modeling and prognosis framework can successfully be applied to various real-life applications such as fatigue degradation in elastomeric rubber bearing and aluminium lap joint. However, developing suitable surrogate measures can be challenging and often requires extensive data preprocessing and advanced techniques. For instance, identifying system parameters for elastomeric rubber bearing deterioration necessitates a two-stage constrained unscented Kalman filter, while mapping crack width to Lamb waves demands wavelet transforms and a probabilistic Bayesian neural network.

7.3 Recommendations for future work

Several potential extensions for applying this research to condition-based maintenance of critical infrastructure using condition monitoring data are outlined below:

1. Degradation signals are often corrupted by measurement noise, introducing non-stationarity that complicates change detection and degradation modeling. A rigorous investigation into the applicability of GPR for surrogate modeling under increasing non-stationarity is necessary to assess the scalability of the proposed approach.

2. While the thesis demonstrates the potential of adaptive mean functions within GPR frameworks for extrapolation and prognosis using EGA-GPR, the limitations of this approach in handling complex time series are evident. Extrapolation, especially with uncertainty quantification, remains a challenging task in machine learning. Future research should focus on developing novel methods to enhance the extrapolation capabilities of machine learning frameworks tailored for prognosis.

3. The proposed algorithm's performance relies on the appropriate selection of an adaptive mean function. The current approach using a set of functions can be refined. Despite the closed-form solutions provided by GPR, developing explainable AI (XAI) techniques specifically for probabilistic models and complex feature spaces is essential. This includes visualizing prediction uncertainties, conducting sensitivity analyses, and tracing the reasoning behind model outputs.

4. The two-stage constraint unscented Kalman filter and other probabilistic models

presented in this thesis offer advantages in handling sparse datasets. However, their computational complexity, particularly with large datasets under Bayesian frameworks, is a significant challenge. Future research should explore methods like model simplification and efficient optimization algorithms to address these computational demands.

5. For modeling aluminum lap joint degradation, employing more sophisticated Lamb wave features and comparing their performance to existing features is recommended.
6. The limited experimental data for structural engineering applications restricts a comprehensive evaluation of the proposed method. A more extensive dataset and diverse application examples are necessary to fully understand the method's performance, including the identification of effective features for damage detection and prognosis in the civil engineering domain.

List of publication

Publications related to doctoral thesis work

Journal articles:

1. **Ojha, S.**, Kalimullah, N.M. and Shelke, A., 2022. Application of constrained unscented Kalman filter (CUKF) for system identification of coupled hysteresis under bidirectional excitation. *Structural Control and Health Monitoring, Wiley*, 29(12), p.e3115.
2. **Ojha, S.**, and Shelke, A., 2024. Probabilistic deep learning approach for fatigue crack width estimation and prognosis in lap joint using acoustic waves *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, ASME*, pp.1-25.
3. **Ojha, S.**, and Shelke, A., 2025. Estimation of the Remaining Useful Life of Axially Loaded Members Through Nested Least-Square Support Vector Regression. Journal of Earthquake Engineering, 29(1), pp.243-264.
4. **Ojha, S.**, and Shelke, A., Diagnosis and prognosis of multi-stage degradation through gaussian process and entropy-assisted genetic algorithm. **(Under review)**

Conference proceedings and book chapters:

1. **Ojha, S.** and Shelke, A., 2022, November. Influence of Biaxial Interaction in Bouc-Wen Hysteresis Model Used for Modeling High Damping Rubber Bearing Under Bidirectional Excitation. In 17^{th} *Symposium in Earthquake Engineering* (pp. 139-150). Singapore: Springer Nature Singapore.
2. **Ojha, S.**, Pamwani, L.G. and Shelke, A., 2020, March. Damage Detection in Base-Isolated Steel Structure Using Singular Spectral Analysis. In *Indian Structural Steel Conference* (pp. 387-401). Singapore: Springer Nature Singapore.
3. **Ojha, S.**, Shelke, A. and Habib. A., 2023. Least-Square Support Vector Regression for the Prognosis of the Deteriorating Structure Under the Seismic Excitations Using Autoregressive Model. *14th International Workshop on Structural Health Monitoring (IWSHM)*, Stanford University, CA.
4. **Ojha, S.**, Shelke, A., Tiwari, S.B., Santhosh, B., Thomas, S. and Habib, A., 2023, July. Damage Localization in Plates Using Energy of Acoustic Emission through Gaussian

Process Regression. In *50th Annual Review of Progress in Quantitative Nondestructive Evaluation*, Austin, TX (Vol. 87202, p. V001T09A006). American Society of Mechanical Engineers.

Others publications apart from doctoral thesis work

Journal articles:

1. **Ojha, S**, Jangid, N., Shelke, A., Habib, A., Probabilistic damage localization in plates through multi-output GPR using acoustic Lamb waves. *Measurement, Elsevier*
2. Rajeev, A., Pamwani, L., **Ojha, S** and Shelke, A., 2023. Adaptive autoregressive modelling based structural health monitoring of RC beam-column joint subjected to shock loading. *Structural Health Monitoring, SAGE*, 22(2), pp.1049-1068.

Appendix A

Hamiltonian Monte Carlo (HMC)

Hamiltonian Monte Carlo (HMC) is an efficient Markov Chain Monte Carlo (MCMC) method that leverages Hamiltonian dynamics and incorporates Metropolis-Hasting algorithms for sampling from continuous distributions. In MCMC methods, the goal is to construct a Markov chain, represented by a sequence of random samples (S_i), where each sample probabilistically depends only on the preceding one (S_{i-1}), and the overall sequence follows a desired distribution. Unlike standard sampling approaches such as rejection or inversion sampling, most MCMC algorithms, including HMC, require an initial burn-in period for the Markov chain to reach the desired distribution. Among MCMC methods, the Metropolis-Hastings algorithm stands out because it doesn't demand knowledge about the exact probability distribution $P(\Omega \mid X, Y)$ but only requires a function proportional to that distribution, such as a Bayesian posterior distribution. However, the spread of the proposal distribution in the Metropolis-Hastings algorithm needs careful tuning to avoid a high rejection rate or excessive autocorrelation in the samples. While there is no general method for tweaking these parameters, a strategy is required to obtain the new proposed sample. Additionally, successive samples generated by MCMC may be correlated, necessitating the generation of a large set of samples, followed by subsampling to obtain approximately independent samples from the target distribution. However, storing the final collection of samples after training is computationally expensive for many deep-learning models. Despite these challenges, MCMC methods, particularly the Metropolis-Hastings algorithm, are widely regarded as some of the best and most popular solutions for sampling exact posterior distributions in Bayesian statistics. The HMC algorithm does the job as it uses Hamiltonian dynamics to propose new samples, providing an alternative and potentially more efficient approach compared to traditional Metropolis-Hastings methods. In PBNNs, the target posterior distribution of the weights $\Omega = w_1, w_2, \ldots, w_n$ for the training set is given by the joint probabilities $p(\Omega \mid X, Y)$. Considering, the auxiliary momentum variable r for each model variable Ω, HMC will generate sample independently of the joint density function as

follows

$$p(\Omega, r) \propto \exp\left\{\mathcal{L}(\Omega) - \frac{1}{2} r^T r\right\} \tag{A.1}$$

Algorithm 1 Hamiltonian Monte Carlo

Given $w^0, \epsilon, L, \mathcal{L}, M$
for $m = 1$ to M **do**
    Sample $r^0 \sim \mathcal{N}(0, I)$.
    Set $w^m \leftarrow w^{m-1}, w \leftarrow w_{m-1}, r \leftarrow r_0$.
    for $i = 1$ to L **do**
        Set $w, r \leftarrow$ Leapfrog (w, r, ϵ).
    end for
    Metropolis-Hastings Step:
    With probability $\alpha = \min\{1, \frac{\exp\{\mathcal{L}(\tilde{w}) - \frac{1}{2}\tilde{r}\cdot\tilde{r}\}}{\exp\{\mathcal{L}(w^{m-1}) - \frac{1}{2} r^0 \cdot r^0\}}\}$
    set $w^m \leftarrow \tilde{w}, r^m \leftarrow -\tilde{r}$.
end for

function Leapfrog (w, r, ϵ)
Set $\tilde{r} \leftarrow r + (\epsilon/2)\nabla_w \mathcal{L}(w)$.
Set $\tilde{w} \leftarrow w + \epsilon\tilde{r}$
Set $\tilde{r} \leftarrow r + (\epsilon/2)\nabla_w \mathcal{L}(w)$
return $\tilde{w}, \tilde{r}$

The working mechanism of the HMC is provided in Algorithm 1. Here, L is the length of the sample space, M is the total number of iteration, $\mathcal{L}$ is the logarithm of the joint density of the trainable weights w. Further, in this thesis, all the weights of the PBNNs is kept trainable and the No-U-Turn sampling technique is implemented (Hoffman et al., 2014), which allows the hyperparameters of the algorithm to be automatically tweaked instead of manually setting them.

Appendix B

Variational Inference

Due to large amount of parameters, and non-linear activation function between the variational hidden layers, it is impossible to track the close form of the posterior of weights i.e., $p(\Omega \mid X, Y)$. Thus, the approximate posterior distribution, parameterized by ϕ, called variational distribution $q_\phi(\Omega)$ is developed using stochastic variational inference. Unlike $p(\Omega \mid X, Y), q_\phi(\Omega)$ can be transformed into closed form, which reduces the inference difficulty. In this work, for the complex model like deep PBNNs, the mean field Gaussian distribution is considered as approximate posterior variational distribution. The values of the parameters ϕ are then learned such that the variational distribution $q_\phi(\Omega)$ is as close as possible to the exact posterior $P(\Omega \mid X, Y)$. The measure of closeness that is commonly used is the Kullback-Leibler divergence (KL-divergence) (Kullback and Leibler, 1951), which measures the difference between probability distributions based on Shannon's information theory Shannon (1948). The KL divergence between $q_\phi(\Omega)$ and $p(\Omega \mid X, Y)$ is defined as

$$D_{KL}\left(q_\phi(\Omega) \mid p(\Omega \mid X, Y)\right) = -\int q_\phi(\Omega) \ln \frac{p(\Omega \mid Y, X)}{q_\phi(\Omega)} d\Omega \tag{B.1}$$

Accordingly, if we can find a $q_\phi(\Omega)$ that minimizes the $D_{KL}\left(q_\phi(\Omega) \mid p(\Omega \mid X, Y)\right)$, the predictive distribution can be approximated by

$$q(y \mid x, X, Y) = \int p(y; \theta) p(\theta \mid x, \Omega) q_\phi(\Omega) d\theta d\Omega \tag{B.2}$$

Thus, to optimize the KL divergence, we employ the Adam optimizer to train ϕ. The detailed algorithm is given in Algorithm 2.

Algorithm 2 Variational inference

1: Require: an initial state for ϕ
2: **for** $k = 1, 2 \ldots N$ **do**
3: Sample $z^{(j)} \sim \mathcal{N}(0, I)$
4: $\mathbf{\Omega}^{(j)} \leftarrow \mu + \sigma \cdot z^{(j)}, j = 1, 2 \ldots N_z$,
5: $L(\phi) \leftarrow \frac{1}{N_z} \sum_{j=1}^{N_z} \left[D_{KL}\left(q_\phi(\Omega) \mid p(\Omega)\right)\right]$
6: Update ϕ with gradient $\nabla_\phi L(\phi)$ using Adam optimizer.
7: **end for**

The algorithm is repeated for every parameter till the KL divergence gets optimized.

Appendix C

Unscented Kalman filter

The Bayesian inference methodology utilized in Kalman filtering consists of three principal stages: (1) generating a prediction based on initial conditions and a mathematical model, (2) incorporating measurements to enhance the confidence of this prediction, and (3) deriving a corrected optimal estimate by synthesizing both the prediction and the measurements, thereby refining the output. Unscented transformation is developed by Julier and Uhlmann (2004) for approximating the joint distributions of random variables x and y describe as $\mathbf{x} \sim N(m, P)$, $\mathbf{y} = \mathrm{g}(\mathbf{x})$. Here, $\mathbf{x}$ is the random variable with normal distribution having to mean equal to m and variance matrix equal to P , and $g(\cdot)$ is the nonlinear function. The Unscented Kalman Filter (UKF) surpasses the Extended Kalman Filter (EKF) for complex, nonlinear systems. Unlike the EKF, it avoids linearization errors by using a sampling approach. The UKF employs carefully chosen "sigma points" to approximate the state distribution, capturing both mean and covariance accurately. These sigma points are then propagated through the nonlinear system, leading to a more accurate estimation of the posterior mean and covariance. Let us consider the nonlinear state-space dynamical model with the discrete measurement at a regular interval of Δt having,

$$\text{State equation:} \qquad \mathbf{x}_{k+1} = \mathbf{x}_k + \int_{(k)\Delta t}^{(k+1)\Delta t} \mathrm{f}(\mathbf{x}(\tau), u_k) d\tau + w_k \tag{C.1}$$

$$\text{Measurement equation:} \qquad \mathbf{y}_{k+1} = \mathrm{g}(\mathbf{x}_{k+1}) + v_{k+1} \tag{C.2}$$

where w_k and v_{k+1} are the Gaussian white noise referred as process noise and measurement noise respectively with covariance matrix Q_k and R_{k+1} . Further, $\mathbf{x} \in R^n$ and $y \in R^m$ are system state and measurement vector. The nonlinear functions $\mathbf{f}$ and $\mathbf{g}$ represents dynamics which transform the distribution of random variable as shown in Figure C.1. Lastly, $\mathbf{u}_k$ is the input vector, and k is the time step. For calculating the statistics of the $\mathbf{y}$, UKF generates a set of symmetric $(2L + 1)$ sigma vectors $\chi_{k|k,_i}$ as shown in Equation C.5 with the associated

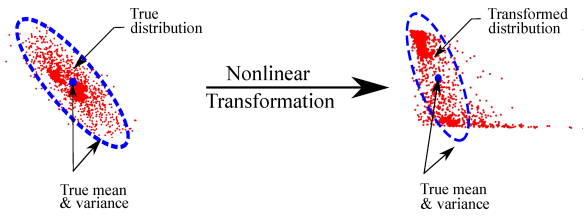


Figure C.1: Non-linear transformation of distribution

weights $\mathbf{W}$.

$$\chi_{k|k,0} = \widehat{\mathbf{x}}_{k|k} \qquad ,\mathbf{W}_0 = \frac{\kappa}{L+\kappa} \tag{C.3}$$

$$\chi_{k|k,i} = \widehat{\mathbf{x}}_{k|k} + \sqrt{(L+\kappa)(\mathbf{P_{k|k}})_i} \qquad ,\mathbf{W}_i = \frac{1}{2(L+\kappa)} \tag{C.4}$$

$$\chi_{k|k,i+L} = \widehat{\mathbf{x}}_{k|k} - \sqrt{(L+\kappa)(\mathbf{P_{k|k}})_i} \qquad ,\mathbf{W}_{i+L} = \frac{1}{2(L+\kappa)} \tag{C.5}$$

Here, κ is the tuning parameter indirectly determines the spread of sigma points about the mean point or filtered state estimates $\widehat{x}_{k|k}$. The term $\sqrt{(L+\kappa)(\mathbf{P_{k|k}})_i}$ represents the i^{th} column of matrix square root of $\mathbf{P}_{k|k}$, which can be computed using lower triangular Cholesky factorization. Further, W_i is the weight associated with the i^{th} column. Provided the sum of weights W_i equal to unity due to symmetry, the weighted covariance matrix $\mathbf{P}_{k|k}$ is given by:

$$\mathbf{P}_{k|k} = \sum_{i=0}^{2L} W_i(\chi_{k|k,i} - \widehat{\mathbf{x}}_{k|k})(\chi_{k|k,i} - \widehat{\mathbf{x}}_{k|k})^T \tag{C.6}$$

The sigma points vector χ is propagated through the non-linear system state space model and predicted sigma points vector χ^* is obtained as follows:

$$\chi^*_{k+1|k,i} = \chi_{k|k,i} + \int_{k\Delta t}^{(k+1)\Delta t} f(x(\tau), u_k)d\tau \tag{C.7}$$

The predicted state estimate $\widehat{\mathbf{x}}_{k+1|k}$ and its covariance matrix $\mathbf{P}_{k+1|k}$ are calculated from the predicted sigma points vector.

$$\widehat{\mathbf{x}}_{k+1|k} = \sum_{i=0}^{2L} W_i\chi^*_{k+1|k,i} \tag{C.8}$$

$$\mathbf{P}_{k+1|k} = \sum_{i=0}^{2L} W_i(\chi^*_{k+1|k,i} - \widehat{\mathbf{x}}_{k+1|k})(\chi^*_{k+1|k,i} - \widehat{\mathbf{x}}_{k+1|k})^T + \mathbf{Q}_k \tag{C.9}$$

The predicted measurements vector $\mathbf{Y}_{k+1|k}$ are obtained by propagating the predicted sigma points vector through non-linear measurement state space model.

$$\mathbf{Y}_{k+1|k,i} = \mathrm{g}(\chi^*_{k+1|k,i}, u_{k+1}) \tag{C.10}$$

By making use of predicted measurement vector, the predicted measurement mean vector $\mathbf{y}_{k+1}$, predicted covariance matrix of the measurement $\mathbf{P}_{vv}$, and cross covariance of the state and the measurement $\mathbf{P}_{xv}$ are as follows:

$$\widehat{\mathbf{y}}_{k+1} = \sum_{i=0}^{2L} W_i \mathbf{Y}_{k+1,i} \tag{C.11}$$

$$\mathbf{P}_{vv,k+1} = \sum_{i=0}^{2L} W_i (\mathbf{Y}_{k+1,i} - \widehat{\mathbf{y}}_{k+1})(\mathbf{Y}_{k+1,i} - \widehat{\mathbf{y}}_{k+1})^T + \mathbf{R}_{k+1} \tag{C.12}$$

$$\mathbf{P}_{xv,k+1} = \sum_{i=0}^{2L} W_i (\chi_{k+1|k,i} - \widehat{\mathbf{x}}_{k+1|k})(\mathbf{Y}_{k+1,i} - \widehat{\mathbf{y}}_{k+1})^T \tag{C.13}$$

The process of generation of sigma points, their propagation through non-linear model, and finally predicting the posterior mean and covariance is illustrated in Figure C.2.

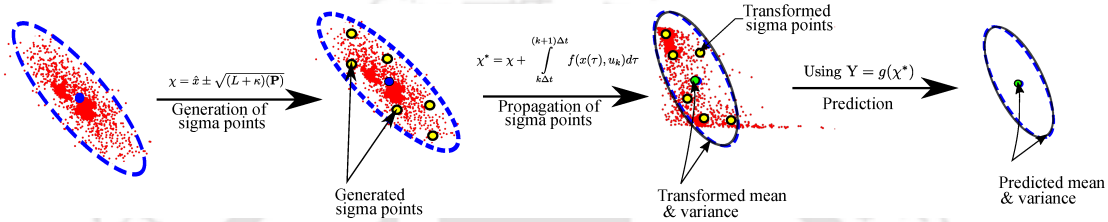


Figure C.2: Process of unscented Kalman filter for predicting posterior estimates

The Kalman gain matrix is calculated by multiplying the cross-covariance matrix $\mathbf{P}_{vv}$ with the inverse of variance matrix $\mathbf{P}_w$ as follows:

$$\mathbf{K}_{k+1} = \mathbf{P}_{xv,k+1} \mathbf{P}_{vv,k+1}^{-1} \tag{C.14}$$

Family of Kalman filter uses the linear update equation for the state estimates. The updated state estimates and the error covariance matrix in the updated estimates are as follows:

$$\begin{aligned} \widehat{\mathbf{x}}_{k+1|k+1} &= \widehat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}(\mathbf{y}_{k+1} - \widehat{\mathbf{y}}_{k+1}) \\ \mathbf{P}_{k+1|k+1} &= \mathbf{P}_{k+1|k} - \mathbf{K}_{k+1} \mathbf{P}_{vv,k+1} \mathbf{K}_{k+1}^T \end{aligned} \tag{C.15}$$

Appendix D

Modelling of rubber bearing

Among seismic bearings, high damping rubber bearing (HDRB) exhibits significant bidirectional effects under large shear strain (Yamamoto et al., 2012). The suitability of the biaxial Bouc-Wen (BBW) model should be checked before using it for application purpose of CUKF. Hence, it is compared with uniaxial Bouc-Wen (UBW) model. The three sets of model parameter as mentioned in Table D.1 are adopted to compare BBW with UBW model. The mass m, stiffness k, and damping c for the LRB system are taken as 6800 kg, 1392 kN/m, 3740 Ns/m, respectively. The hysteresis responses for BBW and UBW for LRB are obtained for the Kobe earthquake, and the same is presented in Figure D.1.

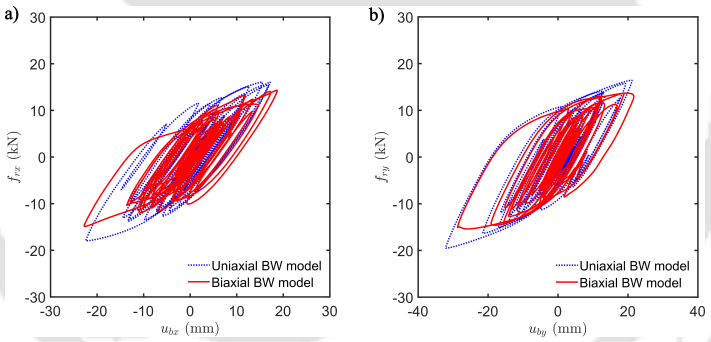


Figure D.1: Comparison of hysteresis of UBW with BBW model for the LRB a) X-axis and b) Y-axis

Table D.1: Hysteresis system parameters of LRB, NRB, and HDRB

Parameters	LRB		NRB		HDRB	
	X	Y	X	Y	X	Y
α	0.15	0.2	0.29	0.31	0.23	0.26
β	1.1	1.4	1.15	0.7	1	1.3
γ	0.2	0.5	-0.1	-0.4	-0.3	-0.3
D_y (mm)	13	13	12.7	12.7	12	12

It can be clearly observed that there are significant discrepancies in the responses which have arisen due to mutual interaction among the principal axes, therefore, it is prudent to model the biaxial behaviour of the structural components. The maximum change is observed in case of HDRB both in terms of displacement and energy dissipation.

Table D.2: Comparison of peak displacement and energy dissipation for seismic isolators considering UBW and BBW model

Isolator type	Displacement (mm)				Energy (kNm)			
	UBW		bBW		UBW		UBW	
	X	Y	X	Y	X	Y	X	Y
LRB	17.18	21.18	18.67	21.68	1.66	2.37	1.72	2.40
NRB	85.93	144.36	88.92	140.60	38.74	102.97	35.06	83.98
HDRB	63.94	109.70	72.27	107.91	47.22	80.90	34.50	74.85

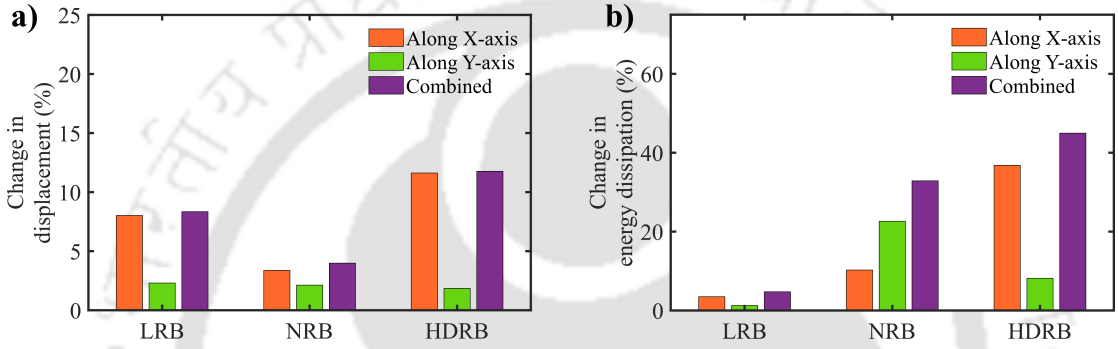


Figure D.2: Comparision of UBW and BBW model by observing the change in (a) displacement and, (b) energy dissipation

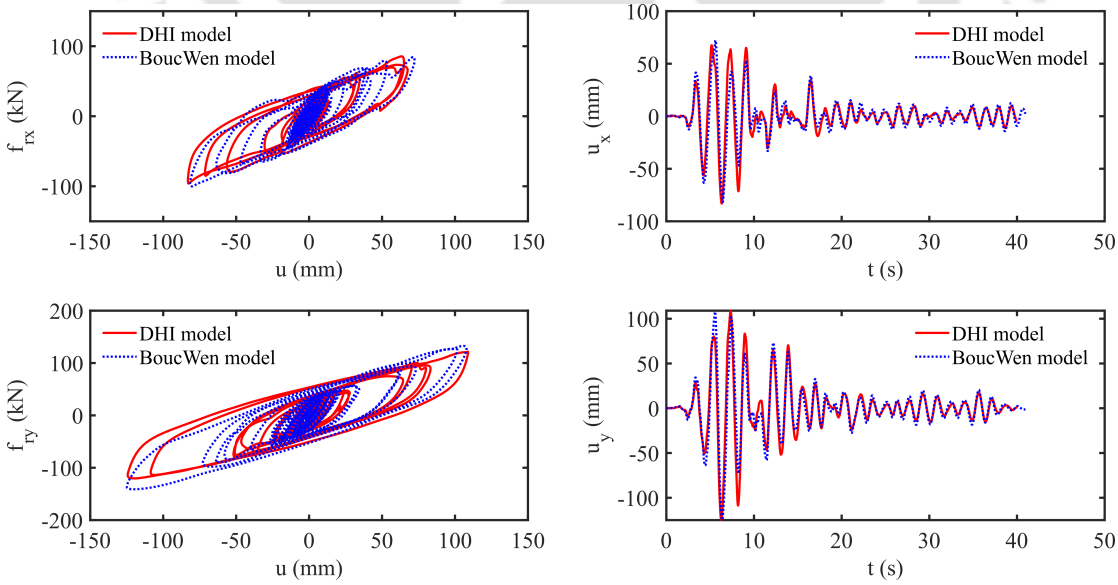


Figure D.3: Comparison of estimated BBW model with the DHI model for HDRB (a) Hysteresis in X direction, (b) Displacement in X direction, (c) Hysteresis in Y Direction, and (d) Displacement in Y Direction

Table D.3: Summary of the simulated and estimated results for the various types of rubber bearings

Type of seismic bearing		LRB		NRB		HDRB	
		X	Y	X	Y	X	Y
Energy dissipated	Simulated system, (kNm)	1.72	24.00	29.80	86.30	34.40	74.10
	Estimated BBW system, (kNm)	1.72	24.00	32.50	79.10	34.50	74.80
	Error, (%)	-0.27	-0.15	-8.86	8.38	-0.33	-0.96
Peak displacement	Simulated system, (mm)	18.65	21.72	89.30	150.03	67.56	108.99
	Estimated BBW system, (mm)	19.33	21.42	93.97	139.17	72.27	107.91
	Error, (%)	-3.65	1.37	-5.22	7.24	-6.98	0.99

This clearly indicates that the BBW model is capturing the interaction behaviour properly and is suitable for the modelling the isolators. To compare the difference in the behaviour of UBW and BBW, the peak displacement and energy dissipation for the LRB, NRB, and HDRB are calculated and presented in Table D.2. The change in displacement and energy dissipation for different isolators when BBW is compared with UBW model is presented in Figure D.2. Moreover, the estimated response of the BBW model is compared with the 'Deformation-History Integral Type (DHI)' model for high damping rubber bearing (HDRB). The hysteresis and displacement response of BBW model resemble close relation with the DHI model and are presented in Figure D.3 . The numerical comparison of estimated BBW model with the simulated system is presented in Table D.3. It can be observed that the energy dissipation and the peak displacement from the simulated and estimated BBW model are comparable. Therefore, the presented biaxial Bouc-Wen model is well suited for the simulation of seismic isolators.

Bibliography

Abdullah and Haldar (2017), 'Structural damage prognosis of three-dimensional large structural systems', *Structure and Infrastructure Engineering* **13**(12), 1596–1608.

Abe, M., Yoshida, J. and Fujino, Y. (2004), 'Multiaxial behaviors of laminated rubber bearings and their modeling. i: Experimental study', *Journal of Structural Engineering* **130**(8), 1119–1132.

Ajavakom, N., Ng, C. and Ma, F. (2008), 'Performance of nonlinear degrading structures: Identification, validation, and prediction', *Computers & Structures* **86**(7-8), 652–662.

Ali, J. B., Chebel-Morello, B., Saidi, L., Malinowski, S. and Fnaiech, F. (2015), 'Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network', *Mechanical Systems and Signal Processing* **56**, 150–172.

Allemang, R. J. (1982), A correlation coefficient for modal vector analysis, *in* 'Proc. of the 1st IMAC', pp. 110–116.

Amer, A. and Kopsaftopoulos, F. P. (2022), 'Statistical guided-waves-based structural health monitoring via stochastic non-parametric time series models', *Structural Health Monitoring* **21**(3), 1139–1166.

An, D., Kim, N. H. and Choi, J.-H. (2015), 'Practical options for selecting data-driven or physics-based prognostics algorithms with reviews', *Reliability Engineering & System Safety* **133**, 223–236.

Arulampalam, M. S., Maskell, S., Gordon, N. and Clapp, T. (2002), 'A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking', *IEEE Transactions on signal processing* **50**(2), 174–188.

Ay, A. M., Khoo, S. and Wang, Y. (2019), 'Probability distribution of decay rate: a statistical time-domain damping parameter for structural damage identification', *Structural Health Monitoring* **18**(1), 66–86.

Aye, S. A. and Heyns, P. S. (2017), '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**, 485–498.

Baber, T. T. and Noori, M. N. (1985), 'Random vibration of degrading, pinching systems', *Journal of Engineering Mechanics* **111**(8), 1010–1026.

Bachoc, F. (2013), 'Cross validation and maximum likelihood estimations of

hyper-parameters of gaussian processes with model misspecification', *Computational Statistics & Data Analysis* **66**, 55–69.

Balageas, D., Fritzen, C.-P. and Güemes, A. (2010), *Structural health monitoring*, Vol. 90, John Wiley & Sons.

Banerjee, P., Palanisamy, R. P., Udpa, L., Haq, M. and Deng, Y. (2019), 'Prognosis of fatigue induced stiffness degradation in GFRPs using multi-modal NDE data', *Composite Structures* **229**, 111424.

Baraldi, Irina, C. P. and Z., E. (2010), 'Methods of Uncertainty Analysis in Prognostics', *International journal of performability engineering* **6**(4), 303.

Baraldi, P., Mangili, F. and Zio, E. (2015), 'A prognostics approach to nuclear component degradation modeling based on Gaussian Process Regression', *Progress in Nuclear Energy* **78**, 141–154.

Barbosh, M., Singh, P. and Sadhu, A. (2020), 'Empirical mode decomposition and its variants: A review with applications in structural health monitoring', *Smart Materials and Structures* **29**(9), 093001.

Barile, C., Casavola, C., Pappalettera, G., Pappalettere, C. and Vimalathithan, P. K. (2020), 'Detection of Damage in CFRP by Wavelet Packet Transform and Empirical Mode Decomposition: an Hybrid Approach', *Applied Composite Materials* **27**(5), 641–655.

Basu, D., Whittaker, A. S. and Constantinou, M. C. (2012), 'Estimating rotational components of ground motion using data recorded at a single station', *Journal of engineering mechanics* **138**(9), 1141–1156.

Beck, J. and Katafygiotis, L. (1992), 'Probabilistic System Identification and Health Monitoring of Structures'.

Beck, J. L. and Au, S.-K. (2002), 'Bayesian updating of structural models and reliability using markov chain monte carlo simulation', *Journal of engineering mechanics* **128**(4), 380–391.

Beck, J. L. and Huang, Y. (2017), Bayesian uncertainty quantification and sparse Bayesian learning for model updating in SHM, *in* 'Proceedings of the Joint COST TU1402 - COST TU1406 - IABSE WC1 Workshop: The Value of Structural Health Monitoring for the reliable Bridge Management', University of Zagreb Faculty of Civil Engineering.

Biondini, F. and Frangopol, D. M. (2016*a*), 'Life-cycle performance of deteriorating structural systems under uncertainty', *Journal of Structural Engineering* **142**(9), F4016001.

Biondini, F. and Frangopol, D. M. (2016*b*), 'Life-Cycle Performance of Deteriorating Structural Systems under Uncertainty: Review', *Journal of Structural Engineering* **142**(9), F4016001.

Bishop, C. (2006), 'Pattern recognition and machine learning', *Springer* **2**, 5–43.

Blei, D. M., Kucukelbir, A. and McAuliffe, J. D. (2017), 'Variational inference: A review for statisticians', *Journal of the American statistical Association* **112**(518), 859–877.

Bornn, L., Farrar, C. R., Park, G. and Farinholt, K. (2009), 'Structural health monitoring with autoregressive support vector machines'.

Bouc, R. (1967), Forced vibrations of mechanical systems with hysteresis, *in* 'Proc. of the Fourth Conference on Nonlinear Oscillations, Prague, 1967'.

Brehm, M. and Deraemaeker, A. (2014), 'Assessment of damage indicators for structural health monitoring using a probabilistic framework'.

Breunig, M. M., Kriegel, H.-P., Ng, R. T. and Sander, J. (2000), Lof: identifying density-based local outliers, *in* 'Proceedings of the 2000 ACM SIGMOD international conference on Management of data', pp. 93–104.

Brownjohn, J. M. (2007), 'Structural health monitoring of civil infrastructure', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **365**(1851), 589–622.

Brownjohn, J. M. W., De Stefano, A., Xu, Y.-L., Wenzel, H. and Aktan, A. E. (2011), 'Vibration-based monitoring of civil infrastructure: challenges and successes', *Journal of Civil Structural Health Monitoring* **1**(3-4), 79–95.

Bull, L., Rogers, T., Wickramarachchi, C., Cross, E., Worden, K. and Dervilis, N. (2019), 'Probabilistic active learning: An online framework for structural health monitoring', *Mechanical Systems and Signal Processing* **134**, 106294.

Burkhart, M. C. (2019), *A discriminative approach to Bayesian filtering with applications to human neural decoding*, ProQuest Dissertations Publishing.

Burman, A., Maity, D., Sreedeep, S. and Gogoi, I. (2011), 'Long-term influence of concrete degradation on dam–foundation interaction', *International Journal of Computational Methods* **8**(03), 397–423.

Cai, G. and Mahadevan, S. (2018), 'Big Data Analytics in Uncertainty Quantification: Application to Structural Diagnosis and Prognosis', *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* **4**(1), 04018003.

Calabrese, A., Strano, S. and Terzo, M. (2018), 'Adaptive constrained unscented kalman filtering for real-time nonlinear structural system identification', *Structural Control and Health Monitoring* **25**(2), e2084.

Calvi, G. M., Moratti, M., O'Reilly, G. J., Scattarreggia, N., Monteiro, R., Malomo, D., Calvi, P. M. and Pinho, R. (2019), 'Once upon a time in italy: The tale of the morandi bridge', *Structural Engineering International* **29**(2), 198–217.

Cangru, J. (2009), 'Research on Damage Identification of Truss Structure Based on BP-PSO'.

Cao, M. S., Sha, G. G., Gao, Y. F. and Ostachowicz, W. (2017), 'Structural damage identification using damping: a compendium of uses and features', *Smart Materials and structures* **26**(4), 043001.

Capecchi, D., Ciambella, J., Pau, A. and Vestroni, F. (2016), 'Damage identification in a

parabolic arch by means of natural frequencies, modal shapes and curvatures', *Meccanica* **51**(11), 2847–2859.

Carden, E. P. and Fanning, P. (2004), 'Vibration based condition monitoring: a review', *Structural health monitoring* **3**(4), 355–377.

Chandra, R., Singh, S. P. and Gupta, K. (1999), 'Damping studies in fiber-reinforced composites–a review', *Composite structures* **46**(1), 41–51.

Chang, C.-C. and Chen, L.-W. (2005), 'Detection of the location and size of cracks in the multiple cracked beam by spatial wavelet based approach', *Mechanical Systems and Signal Processing* **19**(1), 139–155.

Chelidze, D. and Cusumano, J. P. (2004), 'A dynamical systems approach to failure prognosis', *J. Vib. Acoust.* **126**(1), 2–8.

Chen, C., Li, B., Guo, J., Liu, Z., Qi, B. and Hua, C. (2022), 'Bearing life prediction method based on the improved FIDES reliability model', *Reliability Engineering & System Safety* **227**, 108746.

Chen, J. and Liu, Y. (2021), 'Probabilistic physics-guided machine learning for fatigue data analysis', *Expert Systems with Applications* **168**, 114316.

Chen, J., Yuan, S. and Jin, X. (2019), 'On-line prognosis of fatigue cracking via a regularized particle filter and guided wave monitoring', *Mechanical Systems and Signal Processing* **131**, 1–17.

Chen, J., Yuan, S., Sbarufatti, C. and Jin, X. (2021), 'Dual crack growth prognosis by using a mixture proposal particle filter and on-line crack monitoring', *Reliability Engineering & System Safety* **215**, 107758.

Chen, J., Yuan, S. and Wang, H. (2020*a*), 'On-line updating gaussian process measurement model for crack prognosis using the particle filter', *Mechanical Systems and Signal Processing* **140**, 106646.

Chen, J., Yuan, S. and Wang, H. (2020*b*), 'On-line updating Gaussian process measurement model for crack prognosis using the particle filter', *Mechanical Systems and Signal Processing* **140**, 106646.

Chen, Z. (2017), Gaussian process regression methods and extensions for stock market prediction, PhD thesis, University of Leicester.

Chen, Z. et al. (2003), 'Bayesian filtering: From kalman filters to particle filters, and beyond', *Statistics* **182**(1), 1–69.

Chiachıo, J., Chiachıo, M., Saxena, A., Rus, G. and Goebel, K. (2013), An energy-based prognostics framework to predict fatigue damage evolution in composites, *in* 'Proceedings of the annual conference of the prognostics and health management society', Vol. 1, pp. 363–371.

Chiachío, J., Chiachío, M., Saxena, A., Sankararaman, S., Rus, G. and Goebel, K. (2015),

'Bayesian model selection and parameter estimation for fatigue damage progression models in composites', *International Journal of Fatigue* **70**, 361–373.

Ching, J. and Beck, J. L. (2004), 'Bayesian Analysis of the Phase II IASC–ASCE Structural Health Monitoring Experimental Benchmark Data', *Journal of Engineering Mechanics* **130**(10), 1233–1244.

Chopra, A. K. (2007), *Dynamics of structures*, Pearson Education India.

Ciampoli, M., Sibilio, E. and Beck, J. (2008), *Structural health monitoring by Bayesian updating*, Taylor & Francis, pp. 275–291.

Clézio, E., Castaings, M. and Hosten, B. (2002), 'The interaction of the s0 Lamb mode with vertical cracks in an aluminium plate', *Ultrasonics* **40**(1-8), 187–192.

Connor, S., Kim, J., Lynch, J. P., Law, K. H. and Salvino, L. (2010), Fatigue life monitoring of metallic structures by decentralized rainflow counting embedded in a wireless sensor network, *in* 'Smart Materials, Adaptive Structures and Intelligent Systems', Vol. 44168, pp. 751–759.

Constantinou, M. C., Whittaker, A., Kalpakidis, Y., Fenz, D. and Warn, G. P. (2007), 'Performance of seismic isolation hardware under service and seismic loading', *Technical Rep. No. MCEER-07* **12**.

Cornish, C. R., Bretherton, C. S. and Percival, D. B. (2006), 'Maximal overlap wavelet statistical analysis with application to atmospheric turbulence', *Boundary-Layer Meteorology* **119**, 339–374.

Cornwell, P., Doebling, S. W. and Farrar, C. R. (1999), 'Application of the strain energy damage detection method to plate-like structures', *Journal of sound and vibration* **224**(2), 359–374.

Coverley, P. and Staszewski, W. (2003), 'Impact damage location in composite structures using optimized sensor triangulation procedure', *Smart materials and structures* **12**(5), 795.

Cristiani, D., Sbarufatti, C. and Giglio, M. (2021), 'Damage diagnosis and prognosis in composite double cantilever beam coupons by particle filtering and surrogate modelling', *Structural Health Monitoring* **20**(3), 1030–1050.

Daigle, M. J. and Goebel, K. (2012), 'Model-based prognostics with concurrent damage progression processes', *IEEE Transactions on Systems, man, and cybernetics: systems* **43**(3), 535–546.

Davíð (2019), 'Quantifying uncertainty in structural condition with Bayesian deep learning'.

Dimitrov, N., Der Kiureghian, A. and Berggreen, C. (2015), 'Bayesian inference model for fatigue life of laminated composites', *Journal of Composite Materials* **50**(2), 131–143.

Ding, W., Li, J., Mao, W., Meng, Z. and Shen, Z. (2023), 'Rolling bearing remaining useful life prediction based on dilated causal convolutional DenseNet and an exponential model', *Reliability Engineering & System Safety* **232**, 109072.

Doebling, S. W., Farrar, C. R. and Prime, M. B. (1998), 'A summary review of vibration-based damage identification methods', *Shock and vibration digest* **30**(2), 91–105.

Doebling, S. W., Farrar, C. R., Prime, M. B. and Shevitz, D. W. (1996), 'Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review'.

Dong, X., Zhang, C., Liu, H., Wang, D. and Wang, T. (2024), 'A multi-constrained domain adaptation network for remaining useful life prediction of bearings', *Mechanical Systems and Signal Processing* **206**, 110900.

Duffey, T. A., Doebling, S. W., Farrar, C. R., Baker, W. E. and Rhee, W. H. (2001), 'Vibration-based damage identification in structures exhibiting axial and torsional response', *Journal of Vibration Acoustics* **123**(1), 84–91.

Duvenaud, D., Lloyd, J., Grosse, R., Tenenbaum, J. and Zoubin, G. (2013), Structure discovery in nonparametric regression through compositional kernel search, *in* 'International Conference on Machine Learning', PMLR, pp. 1166–1174.

Elasha, Suliman Shanbr, Xiaochuan Li, David and David (2019), 'Prognosis of a Wind Turbine Gearbox Bearing Using Supervised Machine Learning.', *Sensors* **19**(14), 3092.

Eleftheroglou, N. and Loutas, T. (2016), 'Fatigue damage diagnostics and prognostics of composites utilizing structural health monitoring data and stochastic processes', *Structural Health Monitoring* **15**(4), 473–488.

Eleftheroglou, N., Zarouchas, D., Loutas, T., Alderliesten, R. and Benedictus, R. (2018), 'Structural health monitoring data fusion for in-situ life prognosis of composite structures', *Reliability Engineering & System Safety* **178**, 40–54.

Eleftheroglou, N., Zarouchas, D., Loutas, T., Alderliesten, R. C. and Benedictus, R. (2016), 'Online Remaining Fatigue Life Prognosis for Composite Materials Based on Strain Data and Stochastic Modeling', *Key Engineering Materials* **713**, 34–37.

Elenchezhian, M. R. P., Vadlamudi, V., Raihan, R., Reifsnider, K. and Reifsnider, E. (2021), 'Artificial intelligence in real-time diagnostics and prognostics of composite materials and its uncertainties—a review', *Smart Materials and Structures* **30**(8), 083001.

Ellingwood, B. R. (2005), 'Risk-informed condition assessment of civil infrastructure: state of practice and research issues', *Structure and infrastructure engineering* **1**(1), 7–18.

Enright, M. P. and Frangopol, D. M. (1998), 'Probabilistic analysis of resistance degradation of reinforced concrete bridge beams under corrosion', *Engineering structures* **20**(11), 960–971.

Fajfar, P. (2000), 'A nonlinear analysis method for performance-based seismic design', *Earthquake spectra* **16**(3), 573–592.

Fallahdizcheh, A. and Wang, C. (2022), 'Transfer learning of degradation modeling and prognosis based on multivariate functional analysis with heterogeneous sampling rates', *Reliability engineering & system safety* **223**, 108448.

Fang, D., Zhang, X., Yu, Q., Jin, T. C. and Tian, L. (2018), 'A novel method for carbon dioxide emission forecasting based on improved gaussian processes regression', *Journal of cleaner production* **173**, 143–150.

Fang, X., Luo, H. and Tang, J. (2005), 'Structural damage detection using neural network with learning rate improvement', *Computers & structures* **83**(25-26), 2150–2161.

Farrar, C. R. and Worden, K. (2007), 'An introduction to structural health monitoring', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **365**(1851), 303–315.

Farrar, C. R. and Worden, K. (2012*a*), *Structural health monitoring: a machine learning perspective*, John Wiley & Sons.

Farrar, C. R. and Worden, K. (2012*b*), *Structural health monitoring: a machine learning perspective*, John Wiley & Sons.

Fernández Salas, J., Chiachío Ruano, J., Barros, J., Chiachío Ruano, M. and Kulkarni, C. S. (2024), 'Physics-guided recurrent neural network trained with approximate Bayesian computation: A case study on structural response prognostics'.

Figueiredo, E., Park, G., Farrar, C. R., Worden, K. and Figueiras, J. (2011), 'Machine learning algorithms for damage detection under operational and environmental variability', *Structural Health Monitoring* **10**(6), 559–572.

Fish, J. and Yu, Q. (2002), 'Computational mechanics of fatigue and life predictions for composite materials and structures', *Computer Methods in Applied Mechanics and Engineering* **191**(43), 4827–4849.

Foliente, G. C. (1995), 'Hysteresis modeling of wood joints and structural systems', *Journal of Structural Engineering* **121**(6), 1013–1022.

Fong, S. and Narasimhan, S. (2021), 'An unsupervised bayesian oc-svm approach for early degradation detection, thresholding, and fault prediction in machinery monitoring', *IEEE Transactions on Instrumentation and Measurement* **71**, 1–11.

Francesco, D. D., Chryssanthopoulos, M., Faber, M. H. and Bharadwaj, U. (2020), 'Consistent and coherent treatment of uncertainties and dependencies in fatigue crack growth calculations using multi-level Bayesian models', *Reliability Engineering & System Safety* **204**, 107117.

Frangopol, D. M., Kong, J. S. and Gharaibeh, E. S. (2001), 'Reliability-Based Life-Cycle Management of Highway Bridges', *Journal of Computing in Civil Engineering* **15**(1), 27–34.

Friswell, M. I. (2007), 'Damage identification using inverse methods', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **365**(1851), 393–410.

Fromme, P. and Sayir, M. B. (2002), 'Detection of cracks at rivet holes using guided waves', *Ultrasonics* **40**(1-8), 199–203.

Fukada, S., Kajikawa, Y., Hayashi, H., Yoshikawa, M. and Sanuki, Y. (1998), Vibration characteristics of highway bridge with isolators and jointless system under moving vehicles, *in* 'Proc., 5th Int. Conf. on Short and Medium Span Bridges', pp. 1007–1017.

Galanopoulos, G., Milanoski, D., Eleftheroglou, N., Broer, A., Zarouchas, D. and Loutas, T.

(2023), 'Acoustic emission-based remaining useful life prognosis of aeronautical structures subjected to compressive fatigue loading', *Engineering Structures* **290**, 116391.

Gao, H., Cui, L. and Dong, Q. (2020), 'Reliability modeling for a two-phase degradation system with a change point based on a Wiener process', *Reliability Engineering & System Safety* **193**, 106601.

Gao, J., Wang, C., Xu, Z., Wang, J., Yan, S. and Wang, Z. (2022), 'Gaussian process regression based remaining fatigue life prediction for metallic materials under two-step loading', *International Journal of Fatigue* **158**, 106730.

Gao, J., Wang, J., Xu, Z., Wang, C. and Yan, S. (2023), 'Multiaxial fatigue prediction and uncertainty quantification based on back propagation neural network and Gaussian process regression', *International Journal of Fatigue* **168**, 107361.

Gao et al. (2022), 'Strength and stiffness degradation modeling and fatigue life prediction of composite materials based on a unified fatigue damage model', *Engineering Failure Analysis* **137**, 106290.

Ghahramani, Z. (2015*a*), 'Probabilistic machine learning and artificial intelligence', *Nature* **521**(7553), 452–459.

Ghahramani, Z. (2015*b*), 'Probabilistic machine learning and artificial intelligence', *Nature* **521**(7553), 452–459. Publisher: Nature Publishing Group UK London.

Giurgiutiu, V. (2007), *Structural health monitoring: with piezoelectric wafer active sensors*, Elsevier.

Giurgiutiu, V., Zagrai, A. and Jing Bao, J. (2002), 'Piezoelectric wafer embedded active sensors for aging aircraft structural health monitoring', *Structural Health Monitoring* **1**(1), 41–61.

Gobbato, M., Conte, J. P., Kosmatka, J. B. and Farrar, C. R. (2012), 'A reliability-based framework for fatigue damage prognosis of composite aircraft structures', *Probabilistic Engineering Mechanics* **29**, 176–188.

Goodfellow, I., Bengio, Y. and Courville, A. (2016), *Deep learning*, MIT press.

Gordan, M., Sabbagh-Yazdi, S.-R., Ismail, Z., Ghaedi, K., Carroll, P., McCrum, D. and Samali, B. (2022), 'State-of-the-art review on advancements of data mining in structural health monitoring', *Measurement* **193**, 110939.

Gorjian, N., Ma, L., Mittinty, M., Yarlagadda, P. and Sun, Y. (2010), A review on degradation models in reliability analysis, *in* D. Kiritsis, C. Emmanouilidis, A. Koronios and J. Mathew, eds, 'Engineering Asset Lifecycle Management', Springer London, London, pp. 369–384.

Gorlov, A. (1984), 'Disaster of the i-95 mianus river bridge. where could lateral vibration come from ?'.

Greer, M. and Bashir, I. (2023), 'Physics-based model informed smooth particle filter for remaining useful life prediction of lithium-ion battery', *Measurement* **214**, 112838.

Guan, X., He, J., Jha, R. and Liu, Y. (2013), *Structure Reliability and Response Prognostics under Uncertainty Using Bayesian Analysis and Analytical Approximations*, IGI Global, pp. 358–375.

Guennec, B., Ueno, A., Sakai, T., Takanashi, M. and Itabashi, Y. (2013), Effect of loading frequency in fatigue properties and micro-plasticity behavior of jis s15c low carbon steel, *in* 'Proceedings of the 13th International Conference on Fracture, Beijing, China', pp. 16–21.

Guo, J., Li, Z. and Li, M. (2019), 'A review on prognostics methods for engineering systems', *IEEE Transactions on Reliability* **69**(3), 1110–1129.

Guorong, X., Peiqi, C. and Minhui, W. (1996), Bhattacharyya distance feature selection, *in* 'Proceedings of 13th International Conference on Pattern Recognition', Vol. 2, IEEE, pp. 195–199.

Gupta, N. and Hauser, R. (2007), 'Kalman filtering with equality and inequality state constraints', *arXiv preprint arXiv:0709.2791* .

Hassani, S. and Shadan, F. (2022), 'Using incomplete FRF measurements for damage detection of structures with closely-spaced eigenvalues', *Measurement* **188**, 110388.

Hassani, V., Tjahjowidodo, T. and Do, T. N. (2014), 'A survey on hysteresis modeling, identification and control', *Mechanical systems and signal processing* **49**(1-2), 209–233.

Hastings, W. K. (1970), 'Monte carlo sampling methods using markov chains and their applications'.

Haynes, C. and Todd, M. (2012), Bayesian probabilistic modeling for damage assessment in a bolted frame, *in* T. Kundu, ed., 'SPIE Proceedings', SPIE.

He, J., Guan, X., Peng, T., Liu, Y., Saxena, A., Celaya, J. and Goebel, K. (2013), 'A multi-feature integration method for fatigue crack detection and crack length estimation in riveted lap joints using Lamb waves', *Smart Materials and Structures* **22**(10), 105007.

He, J., Lu, Z. and Liu, Y. (2012), 'New method for concurrent dynamic analysis and fatigue damage prognosis of bridges', *Journal of Bridge Engineering* **17**(3), 396–408.

He, M., Zhang, Z. and Ramakrishnan, K. R. (2018), 'Delamination identification for FRP composites with emphasis on frequency-based vibration monitoring-a review', *Structural Durability & Health Monitoring* **12**(4), 213.

Hein, H. and Feklistova, L. (2011), 'Computationally efficient delamination detection in composite beams using Haar wavelets', *Mechanical Systems and Signal Processing* **25**(6), 2257–2270.

Hendricks, W. R. (1991), The aloha airlines accident—a new era for aging aircraft, *in* 'Structural integrity of aging airplanes', Springer, pp. 153–165.

Hoffman, M. D., Gelman, A. et al. (2014), 'The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo.', *J. Mach. Learn. Res.* **15**(1), 1593–1623.

Hou, B., Wang, D., Wang, Y., Yan, T., Peng, Z. and Tsui, K.-L. (2020), 'Adaptive weighted signal preprocessing technique for machine health monitoring', *IEEE Transactions on Instrumentation and Measurement* **70**, 1–11.

Hou, W. and Peng, Y. (2023), 'Adaptive ensemble gaussian process regression-driven

degradation prognosis with applications to bearing degradation', *Reliability Engineering & System Safety* **239**, 109479.

Hristos, T. and Georgia, P. (2022), 'A review of probabilistic forecasting and prediction with machine learning', *arXiv.org* **abs/2209.08307**.

Hu, D., Su, X., Liu, X., Mao, J., Shan, X. and Wang, R. (2020), 'Bayesian-based probabilistic fatigue crack growth evaluation combined with machine-learning-assisted GPR', *Engineering Fracture Mechanics* **229**, 106933.

Hu, Y., Liu, S., Lu, H. and Zhang, H. (2019), 'Remaining useful life model and assessment of mechanical products: a brief review and a note on the state space model method', *Chinese Journal of Mechanical Engineering* **32**, 1–20.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C. and Liu, H. H. (1998), 'The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis', *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences* **454**(1971), 903–995.

Huang, Y. and Beck, J. L. (2013), Novel Sparse Bayesian Learning for Structural Health Monitoring Using Incomplete Modal Data, *in* 'Computing in Civil Engineering', American Society of Civil Engineers.

Hüllermeier, E. and Waegeman, W. (2021), 'Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods', *Machine learning* **110**(3), 457–506.

Humar, J., Mahgoub, M. and Ghorbanie-Asl, M. (2006), 'Effect of second-order forces on seismic response', *Canadian Journal of Civil Engineering* **33**(6), 692–706.

Hussein, A. and Haldar, A. (2015), 'Structural health assessment at a local level using minimum information', *Engineering Structures* **88**, 100–110.

Huston, R. (1994), 'Fatigue life prediction in composites', *International Journal of Pressure Vessels and Piping* **59**(1-3), 131–140.

Ihn, J.-B. and Chang, F.-K. (2004), 'Detection and monitoring of hidden fatigue crack growth using a built-in piezoelectric sensor/actuator network: I. diagnostics', *Smart materials and structures* **13**(3), 609.

Ismail, M., Ikhouane, F. and Rodellar, J. (2009), 'The hysteresis bouc-wen model, a survey', *Archives of computational methods in engineering* **16**, 161–188.

J. Chiachío, M. Chiachío, A. Saxena, G. Rus and K. Goebel (2014), 'A model-based prognostics framework to predict fatigue damage evolution and reliability in composites'.

Jardine, A. K. (2002), Optimizing condition based maintenance decisions, *in* 'Annual Reliability and Maintainability Symposium. 2002 Proceedings (Cat. No. 02CH37318)', IEEE, pp. 90–97.

Jardine, A. K., Lin, D. and Banjevic, D. (2006), 'A review on machinery diagnostics and prognostics implementing condition-based maintenance', *Mechanical systems and signal processing* **20**(7), 1483–1510.

Jiacheng, L. and Lei, L. (2020), 'A hybrid genetic algorithm based on information entropy and game theory', *Ieee Access* **8**, 36602–36611.

Jiang, X. and Mahadevan, S. (2008), 'Bayesian wavelet methodology for structural damage detection', *Structural Control and Health Monitoring* **15**(7), 974–991.

Jones, M. R., Rogers, T. J. and Cross, E. J. (2023), 'Constraining Gaussian processes for physics-informed acoustic emission mapping', *Mechanical Systems and Signal Processing* **188**, 109984.

Julier, S. J. and Uhlmann, J. K. (2004), 'Unscented filtering and nonlinear estimation', *Proceedings of the IEEE* **92**(3), 401–422.

Kalkan, E. and Graizer, V. (2007), 'Coupled tilt and translational ground motion response spectra', *Journal of Structural Engineering* **133**(5), 609–619.

Kalpakidis, I. V. and Constantinou, M. C. (2009), 'Effects of heating on the behavior of lead-rubber bearings. i: Theory', *Journal of Structural Engineering* **135**(12), 1440–1449.

Kamal, A. M. and Taha, I. M. (2010), 'Vibration damping behavior of fiber reinforced composites: A review', *Key Engineering Materials* **425**, 179–194. Publisher: Trans Tech Publ.

Kamiński, M. (2002), 'On probabilistic fatigue models for composite materials', *International Journal of Fatigue* **24**(2-4), 477–495.

Kato, H., Mori, T., Murota, N. and Kikuchi, M. (2015), 'Analytical model for elastoplastic and creep-like behavior of high-damping rubber bearings', *Journal of Structural Engineering* **141**(9), 04014213.

Kendall, A. and Gal, Y. (2017), 'What uncertainties do we need in bayesian deep learning for computer vision?', *Advances in neural information processing systems* **30**.

Kim, C.-W., Zhang, Y., Wang, Z., Oshima, Y. and Morita, T. (2018), 'Long-term bridge health monitoring and performance assessment based on a Bayesian approach', *Structure and Infrastructure Engineering* **14**(7), 883–894.

Kim, K. H., Lee, J. G. and Park, C. G. (2006), 'Adaptive two-stage kalman filter in the presence of unknown random bias', *International Journal of Adaptive Control and Signal Processing* **20**(7), 305–319.

Kim, K. H., Lee, J. G. and Park, C. G. (2007), 'The stability analysis of the adaptive two-stage kalman filter', *International Journal of Adaptive Control and Signal Processing* **21**(10), 856–870.

Kim, M. and Liu, K. (2021), 'A Bayesian deep learning framework for interval estimation of remaining useful life in complex systems by incorporating general degradation characteristics', *IISE Transactions* **53**(3), 326–340.

Kim, N.-H., An, D. and Choi, J.-H. (2017), 'Prognostics and health management of engineering systems', *Switzerland: Springer International Publishing* .

Kim, S. B., Lee, C. G., Hong, J.-W., Park, H. W. and Sohn, H. (2010), 'Applications of an

instantaneous damage detection technique to plates with additional complexities', *Journal of Nondestructive Evaluation* **29**, 189–205.

Kiranyaz, S., Ince, T. and Gabbouj, M. (2015), 'Real-time patient-specific ECG classification by 1-D convolutional neural networks', *IEEE transactions on biomedical engineering* **63**(3), 664–675.

Kitayama, S. and Constantinou, M. C. (2019), 'Probabilistic seismic performance assessment of seismically isolated buildings designed by the procedures of asce/sei 7 and other enhanced criteria', *Engineering Structures* **179**, 566–582.

Kiureghian, A. and Ditlevsen, O. (2009), 'Aleatory or epistemic? does it matter?', *Structural safety* **31**(2), 105–112.

Kraus, M. and Feuerriegel, S. (2019), 'Forecasting remaining useful life: Interpretable deep learning approach via variational Bayesian inferences', *Decision Support Systems* **125**, 113100.

Krauthammer, T. (1999), 'Blast-resistant structural concrete and steel connections', *International journal of impact engineering* **22**(9-10), 887–910.

Kubat, M. (2017), *An Introduction to Machine Learning*, Springer International Publishing, Cham.

Kudela, P., Ostachowicz, W. and Żak, A. (2008), 'Damage detection in composite plates with embedded pzt transducers', *Mechanical Systems and Signal Processing* **22**(6), 1327–1335.

Kulkarni, S. and Achenbach, J. (2008), 'Structural Health Monitoring and Damage Prognosis in Fatigue', *Structural Health Monitoring* **7**(1), 37–49.

Kullback, S. and Leibler, R. A. (1951), 'On information and sufficiency', *The annals of mathematical statistics* **22**(1), 79–86.

Kumar, S. and Padture, N. P. (2018), 'Materials in the aircraft industry', *Metallurgical design and industry: Prehistory to the space age* pp. 271–346.

Kumaraswamidhas, L. A. and Laha, S. K. (2021), 'Bearing degradation assessment and remaining useful life estimation based on Kullback-Leibler divergence and Gaussian processes regression', *Measurement* **174**, 108948.

Kwon, D., Kim, H., Kim, J., Suh, S. C., Kim, I. and Kim, K. J. (2019), 'A survey of deep learning-based network anomaly detection', *Cluster Computing* **22**(S1), 949–961.

Lee, J., Almond, D. and Harris, B. (1999), 'The use of neural networks for the prediction of fatigue lives of composite materials', *Composites Part A: Applied Science and Manufacturing* **30**(10), 1159–1169.

Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L. and Siegel, D. (2014), 'Prognostics and health management design for rotary machinery systems—reviews, methodology and applications', *Mechanical systems and signal processing* **42**(1-2), 314–334.

Lei, Y., Li, N., Guo, L., Li, N., Yan, T. and Lin, J. (2018), 'Machinery health prognostics: A

systematic review from data acquisition to RUL prediction', *Mechanical systems and signal processing* **104**, 799–834.

Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N. and Nandi, A. K. (2020), 'Applications of machine learning to machine fault diagnosis: A review and roadmap', *Mechanical Systems and Signal Processing* **138**, 106587.

Li, D. and Wang, Y. (2021), 'Parameter identification of a differentiable bouc-wen model using constrained extended kalman filter', *Structural Health Monitoring* **20**(1), 360–378.

Li, G., Yang, L., Lee, C.-G., Wang, X. and Rong, M. (2020), 'A Bayesian deep learning RUL framework integrating epistemic and aleatoric uncertainties', *IEEE Transactions on Industrial Electronics* **68**(9), 8829–8841.

Li, L. and Chakik, M. (2021), 'A review of corrosion in aircraft structures and graphene-based sensors for advanced corrosion monitoring', *Sensors* **21**(9), 2908.

Li, T., Lomazzi, L., Cadini, F., Sbarufatti, C., Chen, J. and Yuan, S. (2022), 'Numerical simulation-aided particle filter-based damage prognosis using Lamb waves', *Mechanical Systems and Signal Processing* **178**, 109326.

Li, X., Ma, Y. and Zhu, J. (2021), 'An online dual filters RUL prediction method of lithium-ion battery based on unscented particle filter and least squares support vector machine', *Measurement* **184**, 109935.

Li, X., Zhang, W. and Ding, Q. (2019), 'Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction', *Reliability Engineering & System Safety* **182**, 208–218.

Li, Z. and Zhang, Y. (2014), 'Extreme value theory-based structural health prognosis method using reduced sensor data', *Structure and Infrastructure Engineering* **10**(8), 988–997.

Li, Z., Zhang, Y. and Wang, C. (2013), 'A sensor-driven structural health prognosis procedure considering sensor performance degradation', *Structure and Infrastructure Engineering* **9**(8), 764–776.

Liew, K. M. and Wang, Q. (1998), 'Application of Wavelet Theory for Crack Identification in Structures', *Journal of Engineering Mechanics* **124**(2), 152–157.

Lifshitz, J. and Rotem, A. (1969), 'Determination of Reinforcement Unbonding of Composites by a Vibration Technique', *Journal of Composite Materials* **3**(3), 412–423.

Lim (2014), 'A Lamb wave approach for fatigue damage prognosis with piezoelectric transducers'.

Limongelli, M. P. (2010), 'Frequency response function interpolation for damage detection under changing environment', *Mechanical Systems and Signal Processing* **24**(8), 2898–2913.

Lin, M., You, Y., Wang, W. and Wu, J. (2023), 'Battery health prognosis with gated recurrent unit neural networks and hidden Markov model considering uncertainty quantification', *Reliability Engineering & System Safety* **230**, 108978.

Ling, Y. and Mahadevan, S. (2012), 'Integration of structural health monitoring and fatigue damage prognosis', *Mechanical Systems and Signal Processing* **28**, 89–104.

Ling, Y., Shantz, C., Mahadevan, S. and Sankararaman, S. (2011), 'Stochastic prediction of fatigue loading using real-time monitoring data', *International Journal of Fatigue* **33**(7), 868–879.

Liu, A., Wang, L., Bornn, L. and Farrar, C. (2018), 'Robust structural health monitoring under environmental and operational uncertainty with switching state-space autoregressive models', *Structural Health Monitoring* **18**(2), 435–453.

Liu, F. and Yan, L. (2018), 'Study of stiffness and bearing capacity degradation of reinforced concrete beams under constant-amplitude fatigue', *Plos one* **13**(3), e0192797.

Liu, J. and Chen, Z. (2019), 'Remaining useful life prediction of lithium-ion batteries based on health indicator and Gaussian process regression model', *Ieee Access* **7**, 39474–39484.

Liu, X., Shangguan, W.-B. and Zhao, X. (2022), 'Probabilistic fatigue life prediction model of natural rubber components based on the expanded sample data', *International Journal of Fatigue* **163**, 107034.

Liu, Y. and Mahadevan, S. (2005), 'Probabilistic fatigue life prediction of multidirectional composite laminates', *Composite Structures* **69**(1), 11–19.

Liu, Y. and Mahadevan, S. (2010), *Probabilistic fatigue life prediction of composite materials*, Elsevier, pp. 220–248.

Liu, Y. and Peng, T. (2019), 'Fatigue crack growth in aluminum lap joint'.
URL: https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/

Lombaert, B. Moaveni, Xianfei He and J. Conte (2009), 'Damage Identification of a Seven-Story Reinforced Concrete S hear Wall Building Using Bayesian Model Updating'.

Lu, C. J. and Meeker, W. O. (1993), 'Using Degradation Measures to Estimate a Time-to-Failure Distribution', *Technometrics* **35**(2), 161–174.

Lu, Q., Ren, G. and Zhao, Y. (2002), 'Multiple damage location with flexibility curvature and relative frequency change for beam structures', *Journal of Sound and vibration* **253**(5), 1101–1114.

Lu, Y., Ye, L., Su, Z. and Yang, C. (2008), 'Quantitative assessment of through-thickness crack size based on Lamb wave scattering in aluminium plates', *NDT & e International* **41**(1), 59–68.

Luengo, M., Kolios, A. and Wang, L. (2016), 'Structural health monitoring of offshore wind turbines: A review through the Statistical Pattern Recognition Paradigm', *Renewable and Sustainable Energy Reviews* **64**, 91–105.

Luo, J., Bixby, A., Pattipati, K., Qiao, L., Kawamoto, M. and Chigusa, S. (2003), An interacting multiple model approach to model-based prognostics, *in* 'SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics.

Conference Theme-System Security and Assurance (Cat. No. 03CH37483)', Vol. 1, IEEE, pp. 189–194.

M. Dirbaz (2013), 'A BAYESIAN UPDATING APPROACH IN STRUCTURAL HEALTH MONITORING FOR DAMAGE DETECTION AND ASSESSMENT'.

Ma, F., Zhang, H., Bockstedte, A., Foliente, G. C. and Paevere, P. (2004), 'Parameter analysis of the differential model of hysteresis', *J. Appl. Mech.* **71**(3), 342–349.

Ma, L. (2007), Condition monitoring in engineering asset management, *in* 'Proceedings of APVC2007 12th Asia Pacific Vibration Conference', JSME, pp. 1–16.

MacKay, D. J. et al. (1998), 'Introduction to gaussian processes', *NATO ASI series F computer and systems sciences* **168**, 133–166.

MacRae, G. A. (1994), 'P-δ effects on single-degree-of-freedom structures in earthquakes', *Earthquake spectra* **10**(3), 539–568.

Mahadevan, S., Jiang, X. and Guratzsch, R. F. (2008), 'Probabilistic Approaches to Sensor Layout Design, Data Processing, and Damage Detection'.

Maher, Y. and Danouj, B. (2020), 'Survey on Deep Learning applied to predictive maintenance', *International Journal of Electrical and Computer Engineering (IJECE)* **10**(6), 5592.

Mandela, R., Kuppuraj, V., Rengaswamy, R. and Narasimhan, S. (2012), 'Constrained unscented recursive estimator for nonlinear dynamic systems', *Journal of Process Control* **22**(4), 718–728.

Mander, J. B., Panthaki, F. and Kasalanati, A. (1994), 'Low-cycle fatigue behavior of reinforcing steel', *Journal of Materials in Civil Engineering* **6**(4), 453–468.

Manson, G., Worden, K., Holford, K. and Pullin, R. (2001), 'Visualisation and Dimension Reduction of Acoustic Emission Data for Damage Detection', *Journal of Intelligent Materials Systems and Structures* **12**(8), 529–536.

Manuel (2013), 'Fatigue prognosis in composites: a bayesian framework'.

Mao, W., He, J., Tang, J. and Li, Y. (2018), 'Predicting remaining useful life of rolling bearings based on deep feature representation and long short-term memory neural network', *Advances in Mechanical Engineering* **10**(12), 168781401881718.

Mars, W. V. and Fatemi, A. (2005), 'Multiaxial fatigue of rubber: Part i: equivalence criteria and theoretical aspects', *Fatigue & Fracture of Engineering Materials & Structures* **28**(6), 515–522.

Martti, S. and John, E. H. (n.d.), 'Machine-Learning Methods in Prognosis of Ageing Phenomena in Nuclear Power Plant Components', *International Journal of Computing* **20**(1).

Mata, J. (2011), 'Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models', *Engineering structures* **33**(3), 903–910.

Mechbal, N., Uribe, J. S. and Rébillat, M. (2015), 'A probabilistic multi-class classifier for structural health monitoring', *Mechanical Systems and Signal Processing* **60-61**, 106–123.

Meng, J., Yue, M. and Diallo, D. (2022), 'A degradation empirical-model-free battery end-of-life prediction framework based on gaussian process regression and kalman filter', *IEEE Transactions on Transportation Electrification* .

Mercer, J. (1909), 'Xvi. functions of positive and negative type, and their connection the theory of integral equations', *Philosophical transactions of the royal society of London. Series A, containing papers of a mathematical or physical character* **209**(441-458), 415–446.

Meruane, V., Aichele, D., Ruiz, R. and López Droguett, E. (2021), 'A deep learning framework for damage assessment of composite sandwich structures', *Shock and Vibration* **2021**, 1–12. Publisher: Hindawi Limited.

Milillo, P., Giardina, G., Perissin, D., Milillo, G., Coletta, A. and Terranova, C. (2019), 'Pre-collapse space geodetic observations of critical infrastructure: The morandi bridge, genoa, italy', *Remote Sensing* **11**(12), 1403.

Minh, H. Q., Niyogi, P. and Yao, Y. (2006), Mercer's theorem, feature maps, and smoothing, *in* 'International Conference on Computational Learning Theory', Springer, pp. 154–168.

Mohamed, E. S. S. (2018), 'Prognosis of Bearing Acoustic Emission Signals Using Supervised Machine Learning', *IEEE Transactions on Industrial Electronics* **65**(7), 5864–5871.

Mohanty, S. (2010), *Structural health monitoring and condition based fatigue damage prognosis of complex metallic structures*, Arizona State University.

Morgese, M., Ansari, F., Domaneschi, M. and Cimellaro, G. P. (2020), 'Post-collapse analysis of morandi's polcevera viaduct in genoa italy', *Journal of Civil Structural Health Monitoring* **10**, 69–85.

Mosavi, A. A., Dickey, D., Seracino, R. and Rizkalla, S. (2012), 'Identifying damage locations under ambient vibrations utilizing vector autoregressive models and mahalanobis distances', *Mechanical systems and signal processing* **26**, 254–267.

Moubray, J. (2001), *Reliability-centered maintenance*, Industrial Press Inc.

Moughty, J. J. and Casas, J. R. (2017), 'A state of the art review of modal-based damage detection in bridges: Development, challenges, and solutions', *Applied Sciences* **7**(5), 510.

Mukhopadhyay, S., Luş, H. and Betti, R. (2016), 'Probabilistic Structural Health Assessment with Identified Physical Parameters from Incomplete Measurements', *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* **2**(3).

Mustafa, S., Matsumoto, Y. and Yamaguchi, H. (2018), 'Vibration-Based Health Monitoring of an Existing Truss Bridge Using Energy-Based Damping Evaluation', *Journal of Bridge Engineering* **23**(1), 04017114.

Mutlib, N. K., Baharom, S. B., El-Shafie, A. and Nuawi, M. Z. (2016), 'Ultrasonic health monitoring in structural engineering: buildings and bridges', *Structural Control and Health Monitoring* **23**(3), 409–422.

Müller, A. C. and Guido, S. (2016), *Introduction to machine learning with Python: a guide for data scientists*, " O'Reilly Media, Inc.".

Naeim, F. and Kelly, J. M. (1999), *Design of seismic isolated structures: from theory to practice*, John Wiley & Sons.

Nagarajaiah, S., Reinhorn, A. M. and Constantinou, M. C. (1991), 'Nonlinear dynamic analysis of 3-d-base-isolated structures', *Journal of Structural Engineering* **117**(7), 2035–2054.

Nagarajaiah, S., Reinhorn, A. M. and Constantinou, M. C. (1993), 'Torsion in base-isolated structures with elastomeric isolation systems', *Journal of structural engineering* **119**(10), 2932–2951.

Nagarajaiah, S. and Xiaohong, S. (2000), 'Response of base-isolated usc hospital building in northridge earthquake', *Journal of structural engineering* **126**(10), 1177–1186.

Nagarajaiah, S. and Yang, Y. (2017), 'Modeling and harnessing sparse and low-rank data structure: a new paradigm for structural dynamics, identification, damage detection, and health monitoring', *Structural Control and Health Monitoring* **24**(1), e1851.

Nair, K. K., Kiremidjian, A. S. and Law, K. H. (2006), 'Time series-based damage detection and localization algorithm with application to the asce benchmark structure', *Journal of Sound and Vibration* **291**(1-2), 349–368.

Nason, G. P. (2008), *Wavelet methods in statistics with R*, Springer.

Neerukatti, R. K., Chattopadhyay, A., Iyyer, N. and Phan, N. (2018), 'A hybrid prognosis model for predicting fatigue crack propagation under biaxial in-phase and out-of-phase loading', *Structural Health Monitoring* **17**(4), 888–901.

Nelson, W. B. (2009), *Accelerated testing: statistical models, test plans, and data analysis*, John Wiley & Sons.

Nikulin, M., Limnios, N., Balakrishnan, N., Kahle, W. and Huber-Carol, C., eds (2010), *Advances in Degradation Modeling: Applications to Reliability, Survival Analysis, and Finance*, Birkhäuser Boston, Boston, MA.

Ochella, S., Shafiee, M. and Dinmohammadi, F. (2022), 'Artificial intelligence in prognostics and health management of engineering systems', *Engineering Applications of Artificial Intelligence* **108**, 104552.

Omenzetter, P. and Brownjohn, J. M. W. (2006), 'Application of time series analysis for bridge monitoring', *Smart Materials and Structures* **15**(1), 129.

Ouyang, Z.-L. and Zou, Z.-J. (2021), 'Nonparametric modeling of ship maneuvering motion based on gaussian process regression optimized by genetic algorithm', *Ocean Engineering* **238**, 109699.

Palanisamy, R. P., Banerjee, P., Haq, M. and Deng, Y. (2022), 'Diagnosis and prognosis of fatigue damage in adhesively bonded joints using ultrasound non-destructive evaluation', *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems* **5**(4), 041008.

Pamwani, L. and Shelke, A. (2018), 'Damage detection using dissimilarity in phase space topology of dynamic response of structure subjected to shock wave loading',

Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems **1**(4), 041004–041004.

Pandey, A. K., Biswas, M. and Samman, M. M. (1991), 'Damage detection from changes in curvature mode shapes', *Journal of sound and vibration* **145**(2), 321–332.

Pandey, M. D. (1998), 'Probabilistic models for condition assessment of oil and gas pipelines', *Ndt & E International* **31**(5), 349–358.

Pandey, M. D., Lu, D. and Komljenovic, D. (2009), The impact of probabilistic modelling on predicting the remaining life of pipes in nuclear plants, *in* 'International Conference on Nuclear Engineering', Vol. 435-12, pp. 503–511.

Papadimitriou, C., Fritzen, C.-P., Kraemer, P. and Ntotsios, E. (2011), 'Fatigue predictions in entire body of metallic structures from a limited number of vibration sensors using Kalman filtering', *Structural Control and Health Monitoring* **18**(5), 554–573.

Paris, P. and Erdogan, F. (1963), 'A critical analysis of crack propagation laws'.

Park, B., Sohn, H., Olson, S. E., DeSimio, M. P., Brown, K. S. and Derriso, M. M. (2012), 'Impact localization in complex structures using laser-based time reversal', *Structural Health Monitoring* **11**(5), 577–588.

Peng, T., He, J., Xiang, Y., Liu, Y., Saxena, A., Celaya, J. and Goebel, K. (2015), 'Probabilistic fatigue damage prognosis of lap joint using bayesian updating', *Journal of Intelligent Material Systems and Structures* **26**(8), 965–979.

Percival, D. B. and Walden, A. T. (2000), *Wavelet methods for time series analysis*, Vol. 4, Cambridge university press.

Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M. P., Shyu, M.-L., Chen, S.-C. and Iyengar, S. S. (2019), 'A Survey on Deep Learning: Algorithms, Techniques, and Applications', *ACM Computing Surveys* **51**(5), 1–36.

Prakash, G. and Narasimhan, S. (2018), 'Bayesian two-phase gamma process model for damage detection and prognosis', *Journal of Engineering Mechanics* **144**(2), 04017158.

Prakash, G., Yuan, X.-X., Hazra, B. and Mizutani, D. (2021), 'Toward a big data-based approach: A review on degradation models for prognosis of critical infrastructure', *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems* **4**(2), 021005.

Pugalenthi, K., Trung Duong, P. L., Doh, J., Hussain, S., Jhon, M. H. and Raghavan, N. (2021), 'Online prognosis of bimodal crack evolution for fatigue life prediction of composite laminates using particle filters', *Applied Sciences* **11**(13), 6046.

Qi, J., Zhu, R., Liu, C., Mauricio, A. and Gryllias, K. (2024), 'Anomaly detection and multi-step estimation based remaining useful life prediction for rolling element bearings', *Mechanical Systems and Signal Processing* **206**, 110910.

Rabiei, M. (2011), *A Bayesian framework for structural health management using acoustic emission monitoring and periodic inspections*, University of Maryland, College Park.

Rabiei, M. and Modarres, M. (2013), 'A recursive Bayesian framework for structural health management using online monitoring and periodic inspections', *Reliability Engineering & System Safety* **112**, 154–164.

Rammohan, R. and Taha, M. (2005), Exploratory Investigations for Intelligent Damage Prognosis using Hidden Markov Models, *in* '2005 IEEE International Conference on Systems, Man and Cybernetics', IEEE.

Rasmussen, C. and Williams, C. (2006), 'Gaussian processes for machine learning.,(mit press: Cambridge, ma)'.

Rayjada, S. P., Raghunandan, M. and Ghosh, J. (2023), 'Machine learning-based RC beam-column model parameter estimation and uncertainty quantification for seismic fragility assessment', *Engineering Structures* **278**, 115111.

Rényi, A. (1961), On measures of entropy and information, *in* 'Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics', Vol. 4, University of California Press, pp. 547–562.

Rezaeianjouybari, B. and Shang, Y. (2020), 'Deep learning for prognostics and health management: State of the art, challenges, and opportunities', *Measurement* **163**, 107929.

Roberts, S., Osborne, M., Ebden, M., Reece, S., Gibson, N. and Aigrain, S. (2013), 'Gaussian processes for time-series modelling', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **371**(1984), 20110550.

Roman, I., Santana, R., Mendiburu, A. and Lozano, J. A. (2021), 'Evolving gaussian process kernels from elementary mathematical expressions for time series extrapolation', *Neurocomputing* **462**, 426–439.

Saintier, N., Cailletaud, G. and Piques, R. (2006), 'Multiaxial fatigue life prediction for a natural rubber', *International Journal of Fatigue* **28**(5-6), 530–539.

Saito, T., Mase, S. and Morita, K. (2005), 'A probabilistic approach to structural damage estimation', *Structural Control and Health Monitoring* **12**(3-4), 283–299.

Sajedi, S. O. and Liang, X. (2021), 'Uncertainty-assisted deep vision structural health monitoring', *Computer-Aided Civil and Infrastructure Engineering* **36**(2), 126–142.

Sakin, R. and Ay, I. (2008), 'Statistical analysis of bending fatigue life data using weibull distribution in glass-fiber reinforced polyester composites', *Materials & Design* **29**(6), 1170–1181.

Salawu, O. S. and Williams, C. (1995), 'Bridge Assessment Using Forced-Vibration Testing', *Journal of Structural Engineering* **121**(2), 161–173.

Salehi, H. and Burgueño, R. (2018), 'Emerging artificial intelligence methods in structural engineering', *Engineering structures* **171**, 170–189.

Sankararaman, S. and Mahadevan, S. (2011), 'Uncertainty quantification in structural damage diagnosis', *Structural Control and Health Monitoring* **18**(8), 807–824.

Sankararaman, You Ling, Christopher Shantz and Sankaran Mahadevan (2011), 'Uncertainty Quantification in Fatigue Crack Growth Prognosis', **2**(1).

Santos, A., Figueiredo, E., Silva, M. F. M., Sales, C. S. and Costa, J. (2016), 'Machine learning algorithms for damage detection: Kernel-based approaches', *Journal of Sound and Vibration* **363**, 584–599.

Sen, S. and Bhattacharya, B. (2017), 'Online structural damage identification technique using constrained dual extended kalman filter', *Structural Control and Health Monitoring* **24**(9), e1961.

Shannon, C. E. (1948), 'A mathematical theory of communication', *The Bell system technical journal* **27**(3), 379–423.

Shi, Z. Y., Law, S. S. and Zhang, L. M. (2000), 'Structural Damage Detection from Modal Strain Energy Change', *Journal of Engineering Mechanics* **126**(12), 1216–1223.

Shih, H. W., Thambiratnam, D. P. and Chan, T. H. T. (2009), 'Vibration based structural damage detection in flexural members using multi-criteria approach', *Journal of sound and vibration* **323**(3-5), 645–661.

Shin, K., Feraday, S., Harris, C., Brennan, M. and Oh, J.-E. (2003), 'Optimal autoregressive modelling of a measured noisy deterministic signal using singular-value decomposition', *Mechanical Systems and Signal Processing* **17**(2), 423–432.

Shiri, S., Pourgol-Mohammad, M. and Yazdani, M. (2015), 'Prediction of remaining fatigue cycles in composite materials under uncertainty', *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg* **2**(1).

Si, X.-S., Wang, W., Chen, M.-Y., Hu, C.-H. and Zhou, D.-H. (2013), 'A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution', *European Journal of Operational Research* **226**(1), 53–66.

Si, X.-S., Wang, W., Hu, C.-H. and Zhou, D.-H. (2011), 'Remaining useful life estimation–a review on the statistical data driven approaches', *European journal of operational research* **213**(1), 1–14.

Sikorska, J. Z., Hodkiewicz, M. and Ma, L. (2011), 'Prognostic modelling options for remaining useful life estimation by industry', *Mechanical systems and signal processing* **25**(5), 1803–1836.

Simoen, E., Moaveni, B., Conte, J. P. and Lombaert, G. (2013), 'Uncertainty Quantification in the Assessment of Progressive Damage in a 7-Story Full-Scale Building Slice', *Journal of Engineering Mechanics* **139**(12), 1818–1830.

Singh, R., Park, J. H. and Atluri, S. N. (1994), 'Growth of multiple cracks and their linkup in a fuselage lap joint', *AIAA Journal* **32**(11), 2260–2268.

Siringoringo, D. M. and Fujino, Y. (2008), 'System identification of suspension bridge from ambient vibration response', *Engineering Structures* **30**(2), 462–477.

Sobczyk, K. and Trebicki, J. (2000), 'Stochastic dynamics with fatigue-induced stiffness degradation', *Probabilistic Engineering Mechanics* **15**(1), 91–99.

Sohn, H. (1999), *A Bayesian probabilistic approach to damage detection for civil structures*, Stanford University.

Sohn, H., Czarnecki, J. A. and Farrar, C. R. (2000), 'Structural Health Monitoring Using Statistical Process Control', *Journal of Structural Engineering* **126**(11), 1356–1363.

Sohn, H. and Farrar, C. R. (2001), 'Damage diagnosis using time series analysis of vibration signals', *Smart materials and structures* **10**(3), 446.

Sohn, H., Farrar, C. R., Hemez, F. M., Shunk, D. D., Stinemates, D. W., Nadler, B. R. and Czarnecki, J. J. (2003), 'A review of structural health monitoring literature: 1996–2001', *Los Alamos National Laboratory, USA* **1**, 16.

Sriramula, S. and Chryssanthopoulos, M. K. (2009), 'Quantification of uncertainty modelling in stochastic analysis of frp composites', *Composites Part A: Applied Science and Manufacturing* **40**(11), 1673–1684.

Srivastava, A. and Parida, S. K. (2022), 'Data driven approach for fault detection and Gaussian process regression based location prognosis in smart AC microgrid', *Electric Power Systems Research* **208**, 107889.

Staszewski, W., Boller, C. and Tomlinson, G. R. (2004), *Health monitoring of aerospace structures: smart sensor technologies and signal processing*, John Wiley & Sons.

Sun, Y., Liu, S., Li, R., Ye, Z., Kang, Y. and Chen, S. (2015), 'A new magnetic flux leakage sensor based on open magnetizing method and its on-line automated structural health monitoring methodology', *Structural Health Monitoring* **14**(6), 583–603.

Suresh, S. (1998), *Fatigue of materials*, Cambridge university press.

Tagade, P., Hariharan, K. S., Ramachandran, S., Khandelwal, A., Naha, A., Kolake, S. M. and Han, S. H. (2020), 'Deep Gaussian process regression for lithium-ion battery health prognosis and degradation mode diagnosis', *Journal of Power Sources* **445**, 227281.

Tang, H., Li, D., Chen, W. and Xue, S. (2016), Uncertainty quantification using evidence theory in concrete fatigue damage prognosis, *in* '2016 IEEE International Conference on Prognostics and Health Management (ICPHM)', IEEE, pp. 1–7.

Tavares, S. and Castro, P. (2017), 'An overview of fatigue in aircraft structures', *Fatigue & Fracture of Engineering Materials & Structures* **40**(10), 1510–1529.

Tee, K. F., Cai, Y. and Chen, H.-P. (2012), 'Structural damage detection using quantile regression', *Journal of Civil Structural Health Monitoring* **3**(1), 19–31.

Tehrani, H., Bakhshi, A. and Yang, T. T. (2021), 'Online probabilistic model class selection and joint estimation of structures for post-disaster monitoring', *Journal of Vibration and Control* **27**(15-16), 1860–1878.

Tibaduiza, D.-A., Torres-Arredondo, M.-A., Mujica, L. E., Rodellar, J. and Fritzen, C.-P. (2013), 'A study of two unsupervised data driven statistical methodologies for detecting

and classifying damages in structural health monitoring', *Mechanical Systems and Signal Processing* **41**(1-2), 467–484.

Tobon-Mejia, D. A., Medjaher, K., Zerhouni, N. and Tripot, G. (2012), 'A data-driven failure prognostics method based on mixture of Gaussians hidden Markov models', *IEEE Transactions on reliability* **61**(2), 491–503.

Tripura, T., Bhowmik, B., Pakrashi, V. and Hazra, B. (2020), 'Real-time damage detection of degrading systems', *Structural Health Monitoring* **19**(3), 810–837.

Vachhani, P., Rengaswamy, R., Gangwal, V. and Narasimhan, S. (2005), 'Recursive estimation in constrained nonlinear dynamical systems', *AIChE Journal* **51**(3), 946–959.

Vachtsevanos, G., Lewis, F., Roemer, M., Hess, A. and Wu, B. (2006), *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*, 1 edn, Wiley.

Vanik, M. W., Beck, J. L. and Au, S. K. (2000), 'Bayesian Probabilistic Approach to Structural Health Monitoring', *Journal of Engineering Mechanics* **126**(7), 738–745.

Vanli, O. A. (2014), 'A Failure Time Prediction Method for Condition-Based Maintenance', *Quality Engineering* **26**(3), 335–349.

Vassilopoulos, A., Georgopoulos, E. and Dionysopoulos, V. (2007), 'Artificial neural networks in spectrum fatigue life prediction of composite materials', *International Journal of Fatigue* **29**(1), 20–29.

Vega, M. A. and Todd, M. D. (2020), 'A variational Bayesian neural network for structural health monitoring and cost-informed decision-making in miter gates', *Structural Health Monitoring* **21**(1), 4–18.

Vega, M. A. and Todd, M. D. (2022), 'A variational Bayesian neural network for structural health monitoring and cost-informed decision-making in miter gates', *Structural Health Monitoring* **21**(1), 4–18.

Viktorov, I. A. (1967), 'Rayleigh and Lamb waves', *Rayleigh and Lamb Waves* p. 33.

Wang, C. H., Rose, J. T. and Chang, F.-K. (2004), 'A synthetic time-reversal imaging method for structural health monitoring', *Smart materials and structures* **13**(2), 415.

Wang, L., Cao, H., Ye, Z. and Xu, H. (2023), 'Bayesian large-kernel attention network for bearing remaining useful life prediction and uncertainty quantification', *Reliability Engineering & System Safety* **238**, 109421.

Wang, P., Youn, B. D. and Hu, C. (2012), 'A generic probabilistic framework for structural health prognostics and uncertainty management', *Mechanical Systems and Signal Processing* **28**, 622–637.

Wang, T., Zhou, H. and Zhu, C. (2022), 'A short-term and long-term prognostic method for PEM fuel cells based on Gaussian process regression', *Energies* **15**(13), 4844.

Wang, V. Z., Pease, T. and Robinson, S. (2016), 'Statistical Damage Prognosis for In-Service Civil Structures against Hazards: Formulations and Applications', *Journal of Engineering Mechanics* **142**(3), 04015090.

Wang, W. and Zhang, W. (2008), 'An asset residual life prediction model based on expert judgments', *European Journal of operational research* **188**(2), 496–505.

Wang, Y. and Rocková, V. (2020), Uncertainty quantification for sparse deep learning, *in* 'International Conference on Artificial Intelligence and Statistics', PMLR, pp. 298–308.

Wei, F. and Pizhong, Q. (2011), 'Vibration-based Damage Identification Methods: A Review and Comparative Study', *Structural Health Monitoring* **10**(1), 83–111.

Wen, P., Zhao, S., Chen, S. and Li, Y. (2021), 'A generalized remaining useful life prediction method for complex systems based on composite health indicator', *Reliability Engineering & System Safety* **205**, 107241.

Wen, Y.-K. (1976), 'Method for random vibration of hysteretic systems', *Journal of the engineering mechanics division* **102**(2), 249–263.

Whittle, A. and Davies, R. (2006), Nicoll highway collapse: evaluation of geotechnical factors affecting design of excavation support system, *in* 'International conference on deep excavations', Vol. 28, p. 30.

Wireman, T. (2004), *Total productive maintenance*, Industrial Press Inc.

Worden, K. and Lane, A. J. (2001), 'Damage identification using support vector machines', *Smart materials and structures* **10**(3), 540.

Worden, K. and Manson, G. (2006), 'The application of machine learning to structural health monitoring', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **365**(1851), 515–537.

Wu, B., Shi, H., Zeng, J., Zhang, X. and Wang, Z. (2023), 'Remaining useful life prediction for complex systems with multiple indicators of stochastic correlation considering random shocks', *Mechanical Systems and Signal Processing* **204**, 110767.

Wu, J., Xu, X., Liu, C., Deng, C. and Shao, X. (2021), 'Lamb wave-based damage detection of composite structures using deep convolutional neural network and continuous wavelet transform', *Composite Structures* **276**, 114590.

Wu, Y.-c. and Feng, J.-w. (2018), 'Development and application of artificial neural network', *Wireless Personal Communications* **102**, 1645–1656.

Wu, Z. and Huang, N. E. (2009), 'Ensemble empirical mode decomposition: A noise-assisted data analysis method', *Advances in Adaptive Data Analysis* **01**(01), 1–41.

Xiang, Y. and Liu, Y. (2010), Efficient probabilistic methods for real-time fatigue damage prognosis, *in* 'Annual Conference of the PHM Society', Vol. 2.

Xu, Y., Kohtz, S., Boakye, J., Gardoni, P. and Wang, P. (2023), 'Physics-informed machine learning for reliability and systems safety applications: State of the art and challenges', *Reliability Engineering & System Safety* **230**, 108900.

Xu, Z. and Saleh, J. H. (2021), 'Machine learning for reliability engineering and safety applications: Review of current status and future opportunities', *Reliability Engineering & System Safety* **211**, 107530.

Yamamoto, M., Minewaki, S., Yoneda, H. and Higashino, M. (2012), 'Nonlinear behavior of high-damping rubber bearings under horizontal bidirectional loading: full-scale tests and analytical modeling', *Earthquake Engineering & Structural Dynamics* **41**(13), 1845–1860.

Yan, B., Ma, X., Huang, G. and Zhao, Y. (2021), 'Two-stage physics-based Wiener process models for online RUL prediction in field vibration data', *Mechanical Systems and Signal Processing* **152**, 107378.

Yang, J., He, J., Guan, X., Wang, D., Chen, H., Zhang, W. and Liu, Y. (2016), 'A probabilistic crack size quantification method using in-situ Lamb wave test and bayesian updating', *Mechanical Systems and Signal Processing* **78**, 118–133.

Yang, J. N., Lin, S., Huang, H. and Zhou, L. (2006), 'An adaptive extended kalman filter for structural damage identification', *Structural Control and Health Monitoring: The Official Journal of the International Association for Structural Control and Monitoring and of the European Association for the Control of Structures* **13**(4), 849–867.

Yang, Y. and Ma, F. (2003), 'Constrained kalman filter for nonlinear structural identification', *Journal of Vibration and Control* **9**(12), 1343–1357.

Yang, Y. and Nagarajaiah, S. (2014), 'Blind identification of damage in time-varying systems using independent component analysis with wavelet transform', *mechanical systems and signal processing* **47**(1-2), 3–20.

Yegnanarayana, B. (2009), *Artificial neural networks*, PHI Learning Pvt. Ltd.

Yi'nan, Z., Taolin, L., Zhiyun, Z. and Wen, Y. (2021), 'State of charge estimation of lithium ion battery based on parallel kalman filter', *Energy storage science and technology* **10**(6), 2352.

Yuan, X.-X., Pandey, M. D. and Bickel, G. A. (2008), 'A probabilistic model of wall thinning in CANDU feeders due to flow-accelerated corrosion', *Nuclear Engineering and Design* **238**(1), 16–24.

Yuhang, Shuai, Q., Zhou, S. and Tang (2017), 'Prognosis of structural damage growth via integration of physical model prediction and bayesian estimation', *IEEE Transactions on Reliability* **66**(3), 700–711.

Zhang, K., Li, H., Duan, Z. and Law, S. (2011), 'A probabilistic damage identification approach for structures with uncertainties under unknown input', *Mechanical Systems and Signal Processing* **25**(4), 1126–1145.

Zhang, Q.-H., NI, Y.-Q. and ZHOU, L. (2019), A Bayesian Probabilistic Approach for Damage Detection of a Population of Nominally Identical Structures: Application to Railway Wheel Condition Assessment, *in* 'Structural Health Monitoring 2019', DEStech Publications, Inc.

Zhang, X.-C., Gong, J.-G. and Xuan, F.-Z. (2021), 'A deep learning based life prediction method for components under creep, fatigue and creep-fatigue conditions', *International Journal of Fatigue* **148**, 106236.

Zhao, Y., Noori, M., Altabey, W. A., Ghiasi, R. and Wu, Z. (2018), 'Deep learning-based damage, load and support identification for a composite pipeline by extracting modal macro strains from dynamic excitations', *Applied Sciences* **8**(12), 2564.

Zheng, W. and Yu, Y. (2013), 'Bayesian Probabilistic Framework for Damage Identification of Steel Truss Bridges under Joint Uncertainties', *Advances in Civil Engineering* **2013**, 1–13.

Zhong, J., Gardoni, P. and Rosowsky, D. (2010), 'Stiffness degradation and time to cracking of cover concrete in reinforced concrete structures subject to corrosion', *Journal of engineering mechanics* **136**(2), 209–219.

Zhong, S. and Oyadiji, S. O. (2007), 'Crack detection in simply supported beams without baseline modal parameters by stationary wavelet transform', *Mechanical Systems and Signal Processing* **21**(4), 1853–1884.

Zhu, R., Chen, Y., Peng, W. and Ye, Z.-S. (2022), 'Bayesian deep-learning for RUL prediction: An active learning perspective', *Reliability Engineering & System Safety* **228**, 108758.

Zhu, R., Peng, W., Wang, D. and Huang, C.-G. (2023), 'Bayesian transfer learning with active querying for intelligent cross-machine fault prognosis under limited data', *Mechanical Systems and Signal Processing* **183**, 109628.

Zárate, B. A., Caicedo, J. M., Yu, J. and Ziehl, P. (2012), 'Bayesian model updating and prognosis of fatigue crack growth', *Engineering Structures* **45**, 53–61.