POLITECNICO MILANO 1863
Civil Environmental and Land Management Engineering

Vibration-based Structural Health Monitoring by using machine learning

TESI DI LAUREA MAGISTRALE IN CIVIL ENGINEERING FOR RISK MITIGATION

Author: **Mohammad Omidi Mamaghani**

Student ID: 971548
Advisor: Prof. Alberto Corigliano
Co-advisor: Prof. Stefano Mariani
Dr. Alireza Entezami
Academic Year: 2022-23

Abstract

Structural health monitoring (SHM) is an important topic of research in the civil, mechanical, and aerospace engineering fields. Vibration-based methods utilizing machine learning have become increasingly popular in the field of structural health monitoring due to their ability to perform pattern matching and their potential for online monitoring. These methods have proven to be effective in assessing the health of structures and detecting any potential damage through vibration data. In this research, we aim to develop a vibration-based structural health monitoring system using machine learning that can accurately and efficiently detect and diagnose damage in structures. For this purpose, the Autoencoder method used to remove the noise and environmental effect from the vibration data will be employed. The distance-based clustering was implemented and optimized by the Gaussian mixture model (GMM). This system will leverage the capabilities of machine learning algorithms to analyze vibration data collected from structures and determine if any anomalies or damage are present. Also, localization of the damage was performed on the I-40 bridge by using a hybrid machine learning approach. Using the calibrated Finite element calibrated by Particle swarm optimization (PSO) successfully detects and localizes the damage.

Keywords: Structural health monitoring, Gaussian mixture model, Particle swarm optimization, machine learning, Autoencoder

Abstract in lingua italiana

Il monitoraggio strutturale è un argomento di ricerca importante nei campi dell'ingegneria civile, meccanica e aerospaziale. I metodi basati sulle vibrazioni che utilizzano l'apprendimento automatico sono diventati sempre più popolari nel campo del monitoraggio strutturale grazie al loro potenziale per il monitoraggio online. Questi metodi si sono dimostrati efficaci nella valutazione della salute delle strutture e nel rilevamento di eventuali danni attraverso i dati di vibrazione. In questa ricerca, miriamo a sviluppare un sistema di monitoraggio strutturale basato su vibrazioni utilizzando l'apprendimento automatico, in grado di rilevare e diagnosticare con precisione ed efficienza i danni nelle strutture. A questo scopo, verrà utilizzato il metodo Autoencoder per rimuovere il rumore e l'effetto ambientale dai dati di vibrazione. Il raggruppamento basato sulla distanza è stato implementato e ottimizzato dal modello di miscela Gaussiana (GMM). Questo sistema sfrutterà le capacità degli algoritmi di apprendimento automatico per analizzare i dati di vibrazione raccolti dalle strutture e determinare se sono presenti anomalie o danni. Inoltre, la localizzazione del danno effettuata sul ponte I-40 utilizzando un approccio ibrido di apprendimento automatico. Utilizzando l'elemento finito calibrato dal Particle swarm optimization (PSO) rileva e localizza con successo il danno.

Key-words: Structural health monitoring, Gaussian mixture model, Particle swarm optimization , machine learning , Autoencoder

Contents

I. Introduction

Structural health monitoring has emerged as an important field in civil, mechanical, and aerospace engineering.[1] It focuses on assessing the health of structures and systems by detecting any potential damage through vibration data. This approach has gained significant attention due to its ability to provide real-time monitoring, early detection of structural anomalies, and prediction of possible failures. Structural health monitoring systems offer the possibility to detect damage in large civil infrastructures accurately and immediately, ensuring structural integrity and safety. Damage in a structure can occur due to various factors such as changes in geometrical configuration or boundary conditions, as well as material degradation caused by cracks in concrete or loose bolts.[2] To address this challenge, vibration-based structural damage detection has received considerable attention in the past two decades. Vibration-based structural damage detection involves analyzing the dynamic response of a structure to identify any changes or anomalies that could indicate damage. Vibration data collected from structural health monitoring systems provide valuable information about the behavior of structures under different conditions. Structural health monitoring systems allow engineers to monitor structural behavior and detect damage at an early stage in order to prevent catastrophic failures.

II. Aim and scope

The overall aim of this research is to develop a vibration-based structural health monitoring system that utilizes machine learning algorithms for accurate damage detection and classification. For this purpose, vibration data from structures will be collected and analyzed to identify patterns and anomalies that indicate the presence of damage. in this research bridge damage detection on two case studies, Z24 bridge and KW51 performed. Different machine learning techniques were performed to remove the effect of the environmental variability change on the sampled data. The specific scope of this research includes:

- Collecting vibration data from structures and extracting relevant features using signal processing techniques.
- Applying machine learning algorithms, such as neural networks or support vector machines, to train and classify the collected vibration data based on their patterns and anomalies. Using these trained models, the vibration data from a structure can be analyzed in real-time to detect and classify any damage that may have occurred. - Comparing the performance of different machine learning algorithms in terms of their accuracy and efficiency in detecting and classifying structural damage.[3]

In the end, a hybrid approach was performed on I-40 bridge as a case study, combining both data-driven and model-driven techniques for more accurate and reliable damage detection. By using the experimental data, a baseline for the healthy condition of the structure was created. By using this valuable data, the calibration of the finite element model became feasible. The damage introduced to the model, evaluation of the damage presence and location of the damage obtained. The results of this research demonstrate the potential of machine learning algorithms in the field of vibration-based structural health monitoring for accurate damage detection and classification.

However, with the increasing amount of vibration data collected from these systems, it has become challenging to develop more efficient and robust detection algorithms. The use of machine learning techniques in vibration-based structural health monitoring has shown promise in addressing these challenges. Machine learning refers to the field of study that focuses on developing algorithms and statistical models that allow computers to learn from data without being explicitly programmed. By leveraging the power of machine learning, engineers can analyze large amounts of vibration data and extract meaningful patterns or features that can be used to detect and diagnose structural damage.[4] Machine learning algorithms can be trained on a dataset of vibration data from both healthy and damaged structures, allowing them to learn the characteristics and patterns associated with structural damage.

These algorithms can then be applied to real-time data from a structural health monitoring system to detect anomalies and predict possible failures. By continuously monitoring the vibrations of a structure, machine learning algorithms can identify any deviations from normal behavior and alert engineers to potential issues.[5] This allows for early detection of structural anomalies, enabling timely intervention and maintenance. Furthermore, the integration of machine learning algorithms with structural health monitoring systems provides a more accurate and efficient approach to damage detection.

Chapter one

1.Structural Health Monitoring

1.1 Introduction to Structural Health Monitoring

Structural Health Monitoring is an essential system that integrates engineering structures with information technology.

This integration allows for the continuous monitoring of structural integrity, thereby potentially preventing financial and non-financial losses. In recent years, SHM systems have been widely deployed in various engineering structures, including bridges and airplanes [6]. Structural Health Monitoring is a system that combines advanced sensor technology and intelligent algorithms to monitor the health condition of structures. It enables the detection and monitoring of structures in a dynamic and real-time manner, providing relevant information about their health status. With the rapid development of information technology, advanced detection and monitoring methods have been continuously studied and updated to ensure accurate and real-time assessment of structural health. Structural Health Monitoring plays a vital role in today's world, where the integrity of engineering structures is crucial. The primary goal of Structural Health Monitoring is to ensure the continuous monitoring and assessment of the structural integrity of engineering structures. [7] By continuously monitoring structures, Structural Health Monitoring systems can detect any damage or anomalies in real time. This enables prompt action to be taken to mitigate potential risks and prevent catastrophic failures. Structural Health Monitoring is an emerging technology that aims to continuously monitor the structural integrity of engineering structures [3]. The ultrasonic pulse-echo technique is a cost-effective strategy commonly used in Structural Health Monitoring [8]. Furthermore, Structural Health Monitoring systems utilize advanced detection and monitoring methods that have been rapidly evolving as a result of advancements in information technology. These methods allow for the dynamic and real-time assessment of structural health, providing immediate warnings when significant changes occur. Overall, Structural Health Monitoring is a crucial tool in ensuring the safety and reliability of critical structures [9]. By continuously monitoring the health state and structural integrity of engineering structures, Structural Health Monitoring systems play a significant role in minimizing potential risks and increasing the lifespan of structures [10].

1.2 Fundamental Concepts of Structural Health Monitoring

Fundamental concepts of Structural Health Monitoring include the integration of engineering structures with information technology, continuous monitoring of structural integrity, and the use of advanced sensor technology and intelligent algorithms. The integration of engineering structures with information technology allows for the collection and analysis of data in real time. This integration enables the continuous monitoring of the structural health, allowing for early detection of any damage or anomalies. Continuous monitoring of structural integrity is a key aspect of Structural Health Monitoring [11]. This involves the implementation of sensors and monitoring systems that are capable of continuously collecting data on structural behavior, performance, and condition. These sensors can be embedded within the structure or placed externally to monitor various parameters such as strain, displacement, temperature, and vibration. The use of advanced sensor technology and intelligent algorithms is another fundamental concept of Structural Health Monitoring. This involves the development and implementation of sensors that are capable of accurately measuring and capturing data on structural behavior. These sensors are equipped with advanced algorithms that enable the analysis and interpretation of the collected data, providing valuable insights into structural health. Structural health monitoring in civil engineering has traditionally relied on the measurement of physical parameters such as strain, acceleration, speed, displacement, and torsion. Recently, there has been a growing interest in utilizing machine learning techniques for structural health monitoring [12]. SHM systems that utilize machine learning techniques can effectively analyze large amounts of data, identify patterns, and make accurate predictions about the health state of structures.

1.3 Level of SHM

There are different levels of structural health monitoring depending on the extent and complexity of the monitoring system. At the basic level, structural health monitoring involves installing sensors on a structure to collect vibration data. This data can be used to monitor the overall structural behavior and identify any abnormalities.

At a more advanced level, machine learning techniques can be utilized to analyze the collected vibration data and identify specific patterns or features that are indicative of structural damage.

structural health monitoring can be classified into three different levels:[13]

Level 1: Early damage detection

Level 2: Damage localization

Level 3: Damage quantification

Level 1 of structural health monitoring focuses on early damage detection. Machine learning algorithms can analyze vibration data to identify any deviations from normal behavior and detect the presence of damage. By comparing current vibration patterns to those in a healthy state, machine learning models can recognize changes that may indicate structural issues. Level 2 involves damage localization, where the goal is to determine the exact location of the damage within a structure. Machine learning techniques can be applied to identify specific patterns or features in the vibration data that are associated with damage at different locations. By analyzing these patterns, machine learning models can accurately pinpoint the areas where damage has occurred. Finally, level 3 of structural health monitoring deals with damage quantification.

1.4 Benefits of Structural Health Monitoring

Structural Health Monitoring offers a range of benefits in various fields, particularly in civil engineering and aerospace. In civil engineering, Structural Health Monitoring plays a crucial role in ensuring the safety and reliability of critical structures such as bridges, dams, and buildings [1]. By continuously monitoring the structural health, any potential damage or anomalies can be detected early on. This enables proactive maintenance and repair, reducing the risk of catastrophic failures and ensuring the longevity of these structures.[14] Moreover, Structural Health Monitoring also helps in optimizing maintenance and repair activities by identifying specific areas that require attention, thus reducing operating and maintenance costs.

In the aerospace industry, structural health monitoring is equally important. It allows for the continuous monitoring of aircraft structures, detecting any damage or degradation that may occur during flight. By detecting these issues early on, necessary repairs or replacements can be made to ensure the safety of passengers and crew. Machine learning techniques, such as artificial neural networks and support vector machines, have been widely applied in structural health monitoring to analyze the complex data collected from sensors and accurately detect structural anomalies or potential failures [15]. Through the use of machine learning algorithms, structural health monitoring systems can process and analyze the vast amount of vibration data collected from sensors in real time, allowing for quick and accurate detection of any structural anomalies [16]. The use of machine learning techniques in vibration-based structural health monitoring has shown promise in addressing the challenges associated with developing more efficient and robust detection algorithms [17]. Machine learning algorithms can analyze the patterns and trends in vibration data, allowing for early detection of structural damage. Furthermore, machine learning algorithms can also be used to predict the remaining useful life of a structure, allowing for better planning of maintenance and repair activities. By leveraging machine learning techniques, structural health monitoring systems can become more intelligent and adaptive. They can learn from the historical data collected, identify patterns and trends, and make predictions about future structural behavior.

Overall, the use of machine learning in vibration-based structural health monitoring holds great potential for accurately and efficiently detecting and diagnosing damage in civil structures [4].

This can greatly contribute to ensuring the safety and longevity of structures, reducing risks of catastrophic failures, and optimizing maintenance and repair strategies.

1.5 Definition of damage in the field of the SHM

In the field of Structural Health Monitoring, damage refers to any alteration or degradation in the structural integrity, performance, or behavior of a structure that compromises its safety, functionality, or durability.[12]

This can include but is not limited to cracks, corrosion, deformations, fatigue, material degradation, and any other changes that may impair the structural capacity or pose a risk to the structure's stability and functionality. By monitoring and detecting such damage at an early stage, engineers can assess the severity of the issue and take appropriate measures to prevent catastrophic failures. Stiffness reduction, changes in natural frequencies, and variations in modal parameters are often indicators of damage and can be monitored using vibration-based structural health monitoring systems.

Therefore, structural health monitoring systems aim to detect and quantify these changes by collecting vibration data from sensors installed on the structure [5].

1.6 Motivation for Developing SHM Technology

The development of structural health monitoring technology is primarily motivated by the need to ensure the safety and reliability of civil structures. Civil structures such as bridges, buildings, and dams are subject to various external factors such as environmental conditions, aging, and fatigue. These factors can lead to structural deterioration and potential failure if not properly monitored and addressed. [18] Structural health monitoring technology aims to provide continuous and real-time information about the condition of a structure, allowing for early detection of any damage or anomalies [19]. This early detection enables engineers to take proactive measures to prevent catastrophic failures and extend the life of the structure.

In addition to safety, the economic implications of structural damage and failure are also significant. The costs associated with repairing or replacing a damaged structure, as well as the potential loss of life and economic disruption caused by structural failures, can be substantial.[10] By implementing structural health monitoring technology, the costs associated with maintenance and repair can be optimized through the early detection of damage.

1.7 Data acquisition

To implement machine learning techniques in structural health monitoring, a substantial amount of data is required.[19] This data typically includes vibration measurements from sensors placed on or near the structure of interest. These sensors capture the dynamic response of the structure, allowing for the analysis and detection of any anomalies.[20] The data acquisition process should be carefully designed and implemented to ensure the accuracy and reliability of the collected data. Additionally, the sensors used for data acquisition should be properly calibrated and positioned to capture the most relevant information about the structural behavior. Feature extraction and selection are also important steps in the data acquisition process.

Data preprocessing is an essential step in structural health monitoring using machine learning algorithms.

1.7.1 Data Acquisition for Bridge Structures

For bridge structures, the data acquisition process involves placing multiple sensors strategically to capture vibrations from different locations of the structure. These sensors are typically placed on the bridge deck, piers, and abutments.[21] The placement of these sensors is crucial, as it ensures that the collected data represents the structural behavior accurately. Furthermore, the sensors used in data acquisition for bridge structures must be able to accurately measure vibrations under various conditions. This includes monitoring vibrations caused by traffic load, wind, earthquakes, and possible damage or deterioration of the structure over time.[22] Data acquisition for bridge structures requires careful consideration of sensor placement and the ability to measure vibrations under various conditions.

1.7.2 Data normalization

Before feeding the data into machine learning algorithms, it is important to normalize the data. Environmental and operational conditions can vary, leading to variations in the recorded vibration data.[23] Normalizing the data involves scaling and standardizing the values to a common range or distribution.

This ensures that the data from different sensors or different periods can be compared and analyzed effectively.

Normalization also allows for better comparison and analysis of different features within the data.

1.7.3 Data Cleansing

Another important aspect of data preprocessing in structural health monitoring using machine learning algorithms is data cleansing. Data cleansing involves the identification and removal of any errors or outliers within the dataset.[24] These errors or outliers may arise due to sensor malfunctions, human errors in data collection, or other factors. Data cleansing is crucial as these errors or outliers can have a significant impact on the accuracy and reliability of the subsequent analysis.

1.7.4 Feature extraction

Feature extraction is a critical step in structural health monitoring using machine learning algorithms.[25] Feature extraction involves selecting and transforming the raw vibration data into a set of relevant features that capture the essential information about the structural behavior.

1.8 Structural Health Monitoring by machine learning

Machine learning techniques offer great potential in improving the efficiency and accuracy of structural health monitoring systems.[26] By utilizing machine learning algorithms, these systems can analyze large amounts of vibration data and identify patterns or anomalies that may indicate damage or structural deterioration. These algorithms can learn from historical data and adapt to new information, allowing for real-time monitoring and early detection of potential issues.[27] Machine learning algorithms have shown significant promise in pattern recognition and classification tasks, making them well-suited for analyzing complex vibration data.[28] While traditional structural health monitoring methods primarily rely on manual analysis and interpretation of vibration data, machine learning algorithms can automate the process and provide more accurate and efficient results.

Machine learning techniques have shown promise in addressing the challenges posed by the increasing amount of vibration data collected from structural health monitoring systems [29]. Machine learning algorithms can effectively analyze large volumes of vibration data and identify patterns or anomalies that may indicate structural damage. In addition to the traditional physical parameter measurements, machine learning techniques can also utilize data from other sources such as temperature sensors, moisture sensors, and even visual data from cameras or drones. The use of machine learning in structural health monitoring offers several advantages [30]. Firstly, machine learning algorithms can handle large amounts of data efficiently and effectively. They have the ability to analyze complex patterns and detect subtle changes in the data that may indicate damage or anomalies.[31] This allows for early detection of structural issues and timely intervention, mitigating the risk of catastrophic failures. Additionally, machine learning algorithms can continuously learn and improve their performance over time. By continuously analyzing and learning from new data, machine learning algorithms can adapt and refine their detection capabilities, improving the accuracy and efficiency of structural health monitoring systems in the long run [32].

1.9 Applications of Machine Learning Techniques in Structural Health Monitoring

One specific application of machine learning techniques in structural health monitoring is the detection and prediction of bridge failures.

Bridges play a crucial role in transportation infrastructure, and their failure can lead to significant economic losses and potentially endanger human lives.[33] Traditional methods of bridge inspection and monitoring rely on visual inspections, manual measurements, and periodic assessments. These methods are often time-consuming, and expensive, and may not provide real-time information on the structural integrity of the bridge. Machine learning techniques offer a more efficient and accurate approach to bridge health monitoring.

By analyzing sensor data from various sources such as strain gauges, accelerometers, and temperature sensors, machine learning algorithms can identify patterns and anomalies that may indicate structural damage or deterioration in a bridge.[34] For example, changes in strain patterns or abnormal vibration levels can be indicative of cracks, corrosion, or other forms of damage. Machine learning algorithms can learn from historical data and observations to develop models that can accurately detect these patterns and anomalies in real-time.

This enables engineers to take corrective actions promptly, such as performing targeted inspections, repairs, or load restrictions, before any catastrophic failures occur. Machine learning algorithms can also aid in the assessment of bridge performance and remaining service life. By continuously monitoring the vibrations and stresses on a bridge, machine learning algorithms can track the deterioration process and predict when maintenance or repairs are needed. Moreover, machine learning techniques can also be applied to other types of structures such as buildings, dams, and tunnels. Machine learning algorithms can analyze the data generated by sensors installed on these structures to identify patterns and anomalies that may indicate structural damage or deterioration [35].

This information can then be used to make informed decisions regarding maintenance, repairs, or even the need for structural upgrades. The use of machine learning in structural health monitoring offers several advantages over traditional methods [22].

Firstly, machine learning algorithms can process large amounts of data quickly and accurately, allowing for real-time monitoring of structural health. This enables early detection of anomalies and potential failures, leading to timely intervention and maintenance [22]. Secondly, machine learning algorithms can adapt and improve over time. As more data is collected and analyzed, machine learning algorithms can continuously update their models to improve accuracy and reliability. Furthermore, machine learning algorithms can also handle complex and non-linear relationships between variables, which may be difficult for traditional analysis methods to capture. In addition, machine learning algorithms have the potential to integrate with other systems, such as flood early-warning systems.

This integration can enhance the effectiveness of both structural health monitoring and flood early-warning systems by providing a comprehensive understanding of potential risks to infrastructure during flood events. Structural health monitoring in civil engineering using machine learning algorithms has the potential to revolutionize the field. It allows for more accurate and efficient monitoring of structural behavior, enabling early detection of anomalies and timely intervention [36].

1.10 Advancements in Structural Health Monitoring Technology

Advancements in Structural Health Monitoring technology have been driven by the need for more accurate and reliable methods of monitoring structural behavior and detecting damage at an early stage. [18]

One of the major advancements in recent years has been the incorporation of machine learning techniques into vibration-based structural health monitoring systems [21]. Machine learning algorithms, such as artificial neural networks and support vector machines, have shown great potential in analyzing large volumes of vibration data to detect and classify damage patterns accurately and efficiently.

These algorithms are capable of learning from the data collected from sensors and can identify subtle changes in structural behavior that may indicate the presence of damage. By training these algorithms on a large dataset of vibration data from healthy and damaged structures, they can learn to recognize patterns associated with different types and severities of structural damage.

The use of machine learning algorithms in structural health monitoring allows for real-time monitoring and early detection of structural anomalies, enabling engineers to take proactive measures to address potential issues before they escalate into major failures [22]. With the increasing amount of vibration data collected from structural health monitoring systems, traditional manual analysis methods have become time-consuming and prone to error. Machine learning techniques offer a more efficient and robust approach to analyzing this data by automating the detection and classification process. By utilizing machine learning algorithms, structural health monitoring systems can effectively analyze large volumes of vibration data and identify specific patterns or features that are indicative of structural damage at different levels of severity [28].

Sensors development has also played a vital role in the advancements of structural health monitoring technology. Sensors used in structural health monitoring systems have become more sophisticated and reliable over the years. Advanced sensors are now capable of capturing high-resolution data with more accuracy and precision, allowing for more detailed analysis of structural behavior.

Nondestructive evaluation (NDE) techniques have also been incorporated into structural health monitoring systems to complement the vibration-based approach. These techniques, such as ultrasonic testing and infrared thermography, provide additional information about the internal conditions of a structure and can help verify the presence and extent of structural damage detected through vibration analysis. [30]

However, despite the benefits and advancements of using machine learning in vibration-based structural health monitoring systems, there are some opposing arguments to consider.

Firstly, there is a concern regarding the accuracy and reliability of machine learning algorithms in detecting structural damage. While these algorithms are capable of learning from data and identifying patterns, their effectiveness largely depends on the quality and representativeness of the training dataset. If the dataset used to train the algorithm does not adequately capture all possible scenarios or types of damage, it may result in false positives or negatives when applied to real-time data. Additionally, machine learning algorithms can be sensitive to outliers or anomalies in the input data, which can also impact their accuracy in detecting structural damage.

1.11 Challenges and Solutions in Structural Health Monitoring

One of the key challenges in implementing vibration-based structural health monitoring with machine learning algorithms is the availability of labeled data [1]. Labeled data, which is essential for supervised machine learning algorithms, can be limited and time-consuming to collect. To overcome this limitation, unsupervised machine learning techniques have been utilized. These techniques focus on outlier analysis to detect the presence of damage [25].

Another challenge in vibration-based structural health monitoring is the feature extraction process. The feature extraction process in vibration-based structural health monitoring often requires manual intervention and expertise prior to damage classification. This manual intervention can be time-consuming and may introduce human error, limiting the efficiency and accuracy of the monitoring system. In order to address these challenges, researchers have been exploring various solutions. One approach is the use of advanced signal processing techniques to automate the feature extraction process. These techniques involve using algorithms to analyze the vibration data and extract relevant features automatically, reducing the need for manual intervention. Another solution is the use of machine learning algorithms that are specifically designed for feature extraction. These algorithms can learn to recognize patterns and extract relevant features directly from the raw vibration data, eliminating the need for human intervention in the feature extraction process. Overall, machine learning algorithms have revolutionized the field of structural health monitoring in civil engineering.

1.12 Conclusion: The Importance of Structural Health Monitoring

In conclusion, structural health monitoring is crucial in civil engineering to ensure the safety and integrity of structures. The implementation of machine learning algorithms in vibration-based structural health monitoring has shown great potential in addressing challenges such as limited labeled data and manual intervention in the feature extraction process. Unsupervised machine learning techniques and advanced signal processing algorithms have been used to overcome these limitations, allowing for more efficient and accurate detection of structural damage. While there are still obstacles to be overcome, machine learning has significantly enhanced the field of structural health monitoring and has paved the way for further advancements in ensuring the safety of our infrastructure. Continuous research and development in this area will lead to even more sophisticated and reliable techniques for structural health monitoring, ultimately contributing to the longevity and resilience of our civil infrastructure. In today's rapidly changing world, the significance of accurate and timely weather forecasts cannot be overstated.

Chapter 2

Machine learning approach

2.1 Introduction to Machine Learning in Structural Health Monitoring

Structural health monitoring is a crucial process in civil, aerospace, and mechanical engineering industries aimed at detecting damage in structures using sensing data [21]. Recently, there has been a growing interest in utilizing machine learning techniques for damage detection in structural health monitoring. Machine learning, a subfield of artificial intelligence, has shown great potential in various applications due to its ability to make accurate predictions and identify patterns from large datasets. With advancements in sensing technology, it has become more feasible to develop an approach for damage detection based on the information gathered from sensor networks mounted on structures. The integration of machine learning algorithms with structural health monitoring systems offers the potential for more accurate and efficient damage detection [37]. Machine learning algorithms can analyze and interpret the vast amount of data collected by sensors in real-time, allowing for continuous monitoring and early detection of damage. By leveraging machine learning techniques, structural health monitoring systems can not only identify damage but also provide valuable insights into the type, extent, and prognosis of the damage. Machine learning algorithms can be trained on historical data that includes both normal and damaged structural behavior. These algorithms learn the patterns and characteristics associated with different types of damage, allowing them to make accurate predictions when new data is received. Machine learning algorithms have the potential to outperform traditional approaches in damage detection due to their ability to handle complex data and adapt to changing conditions. Machine learning algorithms can effectively handle the increasing amount of vibration data collected from structural health monitoring systems, providing a more efficient and robust approach to damage detection [28]. The integration of machine learning algorithms with structural health monitoring systems enhances the accuracy and effectiveness of damage detection [38]. Furthermore, machine learning algorithms can also assist in classifying and categorizing the detected damage, enabling engineers to prioritize and plan for appropriate repairs or maintenance activities based on the severity and urgency of the detected issues. However, there are several challenges in implementing machine learning algorithms for vibration-based structural health monitoring [2]. One challenge is the need for labeled data, which can be difficult to obtain in the field of structural health monitoring.

Another challenge is the need for continuous model updating and adaptation to account for changing structural conditions.[21] Additionally, there may be limitations in the availability and quality of sensor data, as well as challenges in selecting appropriate machine learning algorithms and parameters that are suitable for the specific context of structural health monitoring.

These challenges highlight the need for further research and development to overcome these obstacles and fully harness the potential of machine learning in vibration-based structural health monitoring [39]. Overall, the use of machine learning techniques in vibration-based structural health monitoring holds great promise for improving damage detection and prediction .

2.2 Understanding Damage Detection Techniques

In vibration-based structural health monitoring, machine learning algorithms play a crucial role in detecting and predicting damage. These algorithms utilize labeled data to learn patterns and characteristics associated with structural damage.

However, the availability of labeled data is often limited in the field of structural health monitoring [30]. This presents a challenge for machine learning algorithms, as they heavily rely on labeled data to accurately detect and classify damage. To overcome the scarcity of labeled data, unsupervised machine learning techniques can be employed. These techniques utilize outlier analysis to identify anomalies in the vibration data that may indicate the presence of damage, without needing prior knowledge or labeled data. This approach allows for the detection of damage without relying on pre-labeled data, making it more adaptable to various structural conditions and scenarios. Another challenge in vibration-based structural health monitoring is the feature extraction process.[39] The feature extraction process in vibration-based structural health monitoring typically requires manual intervention to identify relevant features that can be used for damage classification. This manual intervention can be time-consuming and subjective, as different experts may have varying opinions on which features are most indicative of damage. To address this challenge, researchers have been exploring automated feature extraction methods using machine learning algorithms. These methods aim to automatically identify relevant features from raw vibration data, reducing the need for manual intervention and increasing the efficiency of the damage detection process. Machine learning algorithms have shown promise in overcoming challenges in vibration-based structural health monitoring. For example, these algorithms are capable of analyzing patterns in vibration data that are associated with damage at different locations [29]. By analyzing these patterns, machine learning models can accurately pinpoint the areas where damage has occurred [40]. Furthermore, machine learning algorithms can adapt to varying structural conditions and scenarios, making them more robust and efficient in detecting damage. Moreover, machine learning algorithms have the ability to continuously monitor the vibrations of a structure in real-time [41]. This allows for early detection of damage and immediate response, minimizing the risk of structural failures and potential hazards. In summary, machine learning algorithms have the potential to overcome challenges in vibration-based structural health monitoring. These algorithms can utilize unsupervised techniques such as outlier analysis to detect anomalies in vibration data, even without prior knowledge or labeled data.

2.3 Application of Machine Learning in Damage Detection

Machine learning algorithms have been widely used in damage detection and structural health monitoring, particularly in the field of vibration-based analysis. These algorithms offer several advantages over traditional methods, such as manual feature extraction and subjective interpretation. Firstly, machine learning algorithms can automate the feature extraction process in structural health monitoring.

This automation eliminates the need for manual intervention and reduces the potential for human error, leading to more accurate damage classification. Secondly, machine learning algorithms can learn from the raw vibration data and extract relevant features directly. This eliminates the need for handcrafted intervention and allows for a more objective and data-driven approach to damage detection. Additionally, machine learning algorithms have the ability to continuously learn and adapt to changing structural conditions.

This adaptability makes them suitable for long-term monitoring and enables the detection of gradual or progressive damage that may not be apparent through traditional methods. Furthermore, machine learning algorithms can analyze large amounts of data in real time, allowing for early detection and prediction of possible failures. By continuously monitoring the vibrations of a structure in real-time, machine learning algorithms can identify any deviations from normal behavior and alert engineers to potential issues [42]. Moreover, machine learning algorithms have the capability to learn the characteristics and patterns associated with structural damage. These algorithms can then be applied to real-time data from a structural health monitoring system to detect anomalies and predict possible failures [22].

This capability of machine learning algorithms allows for the early detection of structural anomalies, enabling timely intervention and maintenance [30]. Despite the advantages of machine learning in damage detection and structural health monitoring, there are still several challenges that need to be addressed. One major challenge is the availability of training data.

As an example, The integration of structural health monitoring with flood early-warning systems is a promising field; however, it requires sufficient training data specific to flood conditions. Furthermore, although machine learning algorithms have shown great potential in automating the damage detection process, they are not yet fully automated. Human perception is difficult to replicate through vibration or vision-based machine learning algorithms, and there is still a need for human intervention and expertise in interpreting the results and making decisions based on the detected damage. Overall, machine learning algorithms have revolutionized the field of damage detection in structural health monitoring.

2.4 Different types of Machine learning

Several types of machine learning algorithms can be used for damage detection in structural health monitoring. One of the most commonly used types is supervised learning, where the algorithm is trained on labeled data to recognize patterns and make predictions. Another type is unsupervised learning, where the algorithm learns to identify patterns and anomalies in the data without any prior labeling. In addition to these, reinforcement learning can also be employed in structural health monitoring for damage detection.

2.4.1 Supervised learning

Supervised learning algorithms have been widely used for damage detection in structural health monitoring. These algorithms learn from labeled data, where each data point is associated with a specific class or label indicating the presence or absence of damage. By training on this labeled data, the algorithm learns to recognize patterns and make accurate predictions on unseen data. Supervised learning in structural health monitoring typically involves collecting sensor data from various structural components, such as bridges or dikes, and labeling the data based on whether the damage is present or not. The labeled data is then used to train the supervised learning algorithm, which can then be applied to new sensor data to detect damage patterns.[11]

Some common supervised learning algorithms used for damage detection in structural health monitoring include support vector machines, decision trees, random forests, and neural networks.

2.4.2 Unsupervised learning

Unsupervised learning algorithms are also utilized in the field of damage detection for structural health monitoring.

These algorithms do not require labeled data for training, but instead, they learn to identify patterns and anomalies in the data without any prior knowledge of the damage labels. By analyzing the sensor data, unsupervised learning algorithms can identify patterns that may indicate structural damage. For example, outlier analysis is a common method used in unsupervised learning for damage detection. The unsupervised learning algorithms can detect data points that deviate significantly from the majority of the data, indicating potential damage or anomalies in the structural health monitoring system.[11]

One advantage of unsupervised learning in structural health monitoring is that it does not require labeled data, making it more practical and cost-effective for large-scale monitoring of structures. In addition, unsupervised learning techniques can be used for low-level damage detection, where the damage may not be easily observable or distinguishable. However, unsupervised learning algorithms may have a lower accuracy compared to supervised learning in terms of damage detection. In some cases, supervised machine learning is more advantageous in structural health monitoring for damage detection. Supervised machine learning algorithms have the advantage of higher accuracy in detecting damage and more potential for identifying specific types of damage.[12] In contrast, unsupervised machine learning algorithms may be more favorable in practical monitoring scenarios of large structures. Machine learning techniques, such as supervised and unsupervised learning algorithms, have been extensively utilized in the field of structural health monitoring for damage detection.

2.4.3 reinforcement learning

Reinforcement learning is another machine learning technique that has the potential to be used in structural health monitoring for damage detection. Reinforcement learning is a type of machine learning that involves an agent interacting with an environment and making sequential decisions in order to maximize a reward.[11] In the context of structural health monitoring, reinforcement learning can be used to train an agent to make decisions based on sensor data in order to maximize the overall health and safety of a structure. Reinforcement learning has the advantage of being able to learn from experience and adapt its decision-making strategy over time. Furthermore, reinforcement learning algorithms can handle complex and dynamic environments, making them suitable for the unpredictable nature of structural health monitoring. Machine learning is an effective tool in structural health monitoring for damage detection, as it allows the system to learn patterns and detect anomalies based on available data.

2.5 machine learning methods

Supervised and unsupervised machine learning algorithms are commonly employed in structural health monitoring for damage detection. Supervised learning algorithms are used when training data is available from both undamaged and damaged structures. These algorithms, such as statistical process control or MLP neural networks, are trained using labeled data to learn the patterns associated with damage and identify specific types of damage.

On the other hand, unsupervised learning algorithms are utilized when data from damaged structures is rare or nonexistent. These algorithms, which include clustering and anomaly detection techniques, do not require labeled data for training. Instead, they analyze the patterns and similarities within the data to detect any anomalies or deviations from normal behavior. The choice between supervised and unsupervised machine learning algorithms depends on the availability of labeled data and the specific requirements of the structural health monitoring system.

decision tree classification algorithms, support vector machines, and neural networks are commonly used in damage detection for structural health monitoring. Clustering algorithms can be used for unsupervised learning in the absence of labeled data.

2.5.1 K-means Clustering

One example of an unsupervised machine learning algorithm that has been used for damage detection in structural health monitoring is K-means clustering. K-means clustering is a popular method for grouping data points into clusters based on similarity.

This method has been applied to detect damage in structures by analyzing sensor data and identifying clusters that deviate from the normal behavior of the structure. These deviating clusters are indicative of potential damage or anomalies within the structure.

2.5.2 Hierarchical Clustering

Hierarchical clustering is another unsupervised machine-learning algorithm that has been utilized in structural health monitoring for damage detection. Hierarchical clustering is a method that creates a hierarchy of clusters in a dataset.

This algorithm has been used to analyze sensor data and find clusters that exhibit similar patterns or behaviors. These clusters can help identify regions of the structure that may be experiencing damage or degradation.

2.5.3 Fuzzy clustering

Another unsupervised learning algorithm that has been employed for damage detection in structural health monitoring is fuzzy clustering. Fuzzy clustering is a method that allows for the assignment of data points to multiple clusters with varying degrees of membership.

This algorithm has been used to analyze sensor data and assign them to different clusters based on their similarity. The utilization of fuzzy clustering in structural health monitoring allows for a more flexible and nuanced analysis of the data, considering the possibility that a data point may belong to multiple clusters simultaneously.

2.5.4 Principal Component Analysis (PCA) unsupervised learning

Principal Component Analysis is a dimensionality reduction technique that can also be used for damage detection in structural health monitoring. PCA is a method that aims to find the directions (principal components) in which the data varies the most. These principal components can be interpreted as new variables that capture the most important information from the original data.

In the context of structural health monitoring, PCA can be applied to sensor data to reduce dimensionality and identify the most significant features that are indicative of damage or anomalies in the structure.

2.5.5 Hierarchical Clustering unsupervised learning

Hierarchical clustering is a method that creates a hierarchy of clusters in a dataset.

It is a popular unsupervised learning algorithm that has been utilized in structural health monitoring for damage detection. The hierarchical clustering algorithm works by iteratively merging or splitting clusters to form a hierarchy. At each step, the algorithm determines the similarity between clusters based on a specified distance metric. Hierarchical clustering has the advantage of requiring fewer monitoring stations and fewer samples compared to other clustering approaches [12]. Hierarchical clustering can be used in structural health monitoring to establish a correlation between damages and structural states. Hierarchical clustering can be particularly useful when dealing with structural health monitoring data, as it allows for the detection of patterns and similarities within the data without prior supervision or labeling of the data.

Chapter 3

3.Vibration-based damage detection

3.1 Understanding Time Series Acceleration Data

Time series acceleration data refers to a collection of measurements taken over time that captures the rate at which an object's velocity is changing. These measurements are crucial in structural health monitoring as they provide valuable information about the dynamic behavior of a structure.[12] By analyzing time series acceleration data, researchers and engineers can gain insights into the structural integrity of buildings, bridges, and other infrastructure. This information can then be used to detect and evaluate any damage or abnormalities that may be present in the structure. One of the main advantages of using time series acceleration data in structural health monitoring is its ability to capture high-frequency dynamic responses. This is particularly important as high-frequency responses can indicate the presence of damage or potential structural issues that may not be apparent through other types of measurements.

Time series acceleration data is typically obtained using accelerometers, which are sensors that directly measure the acceleration response of structures. Accelerometers are widely used in structural health monitoring practices due to their ability to provide accurate and reliable measurements. However, it's worth noting that accelerometers are less effective in capturing low-frequency responses. While accelerometers are excellent for capturing high-frequency dynamic responses, they may not be as sensitive to low-frequency responses. The use of displacement sensors, on the other hand, can obtain more sensitive low-frequency dynamic measurements [43]. However, displacement sensors have relatively few applications in structural health monitoring due to the inconvenience of installation and monitoring. To overcome the limitations of accelerometers in capturing low-frequency responses and to make use of displacement sensors for their sensitivity, researchers have explored different techniques to extract low-frequency information from time series acceleration data.

One technique commonly used is frequency domain analysis. Frequency domain analysis involves transforming the time series acceleration data into frequency components. This process allows researchers to identify the dominant frequencies present in the data, which can provide valuable information about structural behavior.

By analyzing the dominant frequencies in the time series acceleration data, researchers can detect any shifts or changes in these frequencies, which could indicate damage or abnormalities in the structure.

Another technique that has gained attention in the field of structural health monitoring is time-frequency analysis. Time-frequency analysis involves analyzing the time-varying characteristics of the frequency components in a signal.

This technique allows researchers to explore changes in the frequency content of the time series acceleration data over different time intervals. By performing time-frequency analysis on the time series acceleration data, researchers can identify frequency variations that may indicate damage or structural changes. Time series analysis is another powerful tool used to obtain frequency information in structural health monitoring.

Time series analysis involves decomposing the vibration signals from structures into fundamental basis functions, which are then used to characterize the vibration response [2]. These basis functions can be transformed into the frequency domain to obtain information about the dominant frequencies present in the data. Furthermore, time series analysis allows for the detection of trends and patterns in the data that can help identify changes or anomalies in structural behavior. By utilizing time series acceleration data and applying techniques such as frequency domain analysis, time-frequency analysis, and time series analysis, researchers can effectively obtain frequency information in structural health monitoring.

3.2 Time series data analysis methods

Time series analysis methods involve the statistical modeling of the measured response signals in the time domain.[15], [16] Time series analysis is ideally suited for Structural Health Monitoring (SHM) as it is possible to model the entire time history of a structure and update models with incoming real-time data. Additionally, time series analysis allows for the detection of anomalies or deviations from normal behavior, which can be indicative of structural damage. Vibration-based structural health monitoring is a widely studied and applied technique for detecting damage in structures.[17]

Vibration-based structural health monitoring utilizes the information extracted from the vibrations of a structure to make assessments about its health condition, including the presence, location, and severity of the damage.

3.3 Different types of the time series data

There are different types of time series data that can be collected for vibration-based structural health monitoring. These include stationary time series data, which have constant statistical properties over time, and non-stationary time series data, which exhibit changing statistical properties. Stationary time series data is often obtained from structures under ambient excitation, where the excitation source does not change significantly over time.[18] Non-stationary time series data, on the other hand, is usually obtained from structures subjected to varying load conditions or external disturbances. [19]

3.4 Acquisition of time-Frequency data from acceleration data

To acquire time-frequency data from acceleration data to obtain frequency information from time series acceleration data, researchers employ various techniques such as frequency domain analysis, time-frequency analysis, and time series analysis. Frequency domain analysis involves transforming the time series acceleration data into the frequency domain using techniques like Fourier analysis. Fast Fourier Transform (FFT) is commonly used to calculate the power spectral density energy, which provides information about the frequency composition of the signals [11]. FFT equation is given by:

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) \cdot e^{-j\frac{2\pi}{N}kn} \tag{1}$$

where X(k) represents the frequency domain representation of the time series acceleration data, x(n), is the time-domain signal, f is the frequency component, and j is the imaginary unit.

3.5 Role of Frequency in Structural Damage Detection

The analysis of frequency information in structural health monitoring plays a crucial role in detecting and identifying possible damage or changes in the integrity of structures. By monitoring the frequency content of vibrations, engineers can detect any shifts or anomalies in the structural behavior that may indicate damage. These frequency shifts can be indicative of structural defects such as cracks, looseness in joints, corrosion, or other forms of damage.[11] By comparing the frequency profiles of the monitored structure to a reference or baseline frequency profile, engineers can identify any deviations that may indicate damage. Furthermore, the frequency information obtained from time series acceleration data can also help in assessing structural performance and detecting early signs of deterioration. In the case of time-varying processes, such as in offshore structures, it is crucial to understand the variation of frequencies over time. Time-frequency analysis methods, which describe the frequency and energy intensity of a signal at different times, are commonly used in structural health monitoring to capture the dynamic behavior of structures and mechanical systems.[25] By utilizing time series acceleration data and employing frequency domain analysis techniques, engineers can extract valuable information about the structural health of a system. At present, the commonly used data processing methods of structural health monitoring mainly include time-domain analysis, frequency-domain analysis, and wavelet packet-based energy methods. Among these methods, frequency-domain analysis plays a significant role in detecting damage within structures.

3.6 Using Frequency for Damage Identification in Structures

The frequency content of a structure's vibrations can provide crucial information for identifying and detecting damage. Traditional time-domain and frequency-domain analysis methods may not be directly applicable to non-stationary or nonlinear vibrations caused by external loads such as earthquakes and hurricanes.[11] However, advancements in time-frequency analysis techniques have facilitated the extraction of frequency information from non-stationary signals, enabling a more accurate assessment of the structural condition and damage detection. By utilizing time series acceleration data, engineers can analyze the frequency content of the structural vibrations and track any changes that may indicate damage. Moreover, time-frequency analysis allows for the characterization of frequency components that vary over time, providing insights into the dynamic behavior of structures. By examining the frequency components present in the time series acceleration data, engineers can identify deviations from the expected behavior of a structure and determine if these deviations are indicative of damage or deterioration.

In conclusion, the use of time series acceleration data in conjunction with frequency domain analysis techniques is crucial for effective structural health monitoring and damage detection.

Chapter 4

4. Vibration based damage detection on bridge Z24 and KW51

4.1 Z24 bridge properties

To illustrate the application of vibration-based structural health monitoring, we will consider the case study of the Z24 bridge. The Z24 bridge is a large-scale civil structure that experiences regular vehicular loads. Vibration data from the bridge can be used to monitor its health condition and detect any potential damage.[44]

The Z24 bridge was in the canton Bern near Solothurn, Switzerland. It was part of the road connection between the villages of Koppigen and Utzenstorf, overpassing the A1 highway between Bern and Zürich. It was a classical post-tensioned concrete two-cell box-girder bridge with a main span of 30 m and two side spans of 14 meters. The bridge was built as a freestanding frame with the approaches backfilled later. Both abutments consisted of triple concrete columns connected with concrete hinges to the girder. Both intermediate supports were concrete piers clamped into the girder. An extension of the bridge girder at the approaches provided a sliding slab. All supports were rotated with respect to the longitudinal axis, which yielded a skew bridge. The bridge, which dated from 1963, was demolished at the end of 1998 because a new railway adjacent to the highway required a new bridge with a larger side span.[44,45,46]

In figure 4.1 the cross section of the bridge is illustrated. In addition, the location of the sensors is depicted in figure 4.2.

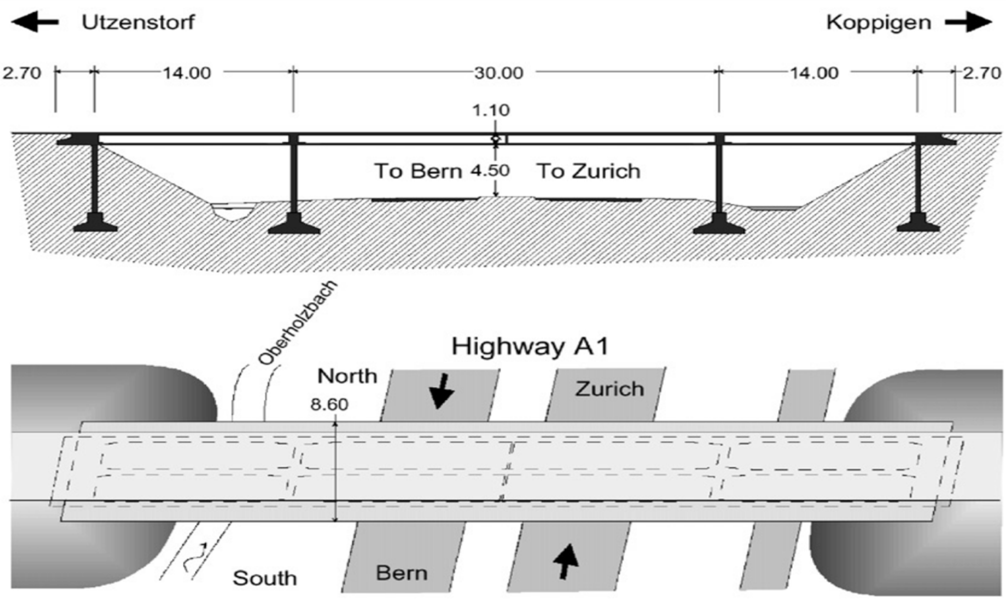


Figure 4.1: Front view (top) and top view (bottom) of the Z24 Bridge.

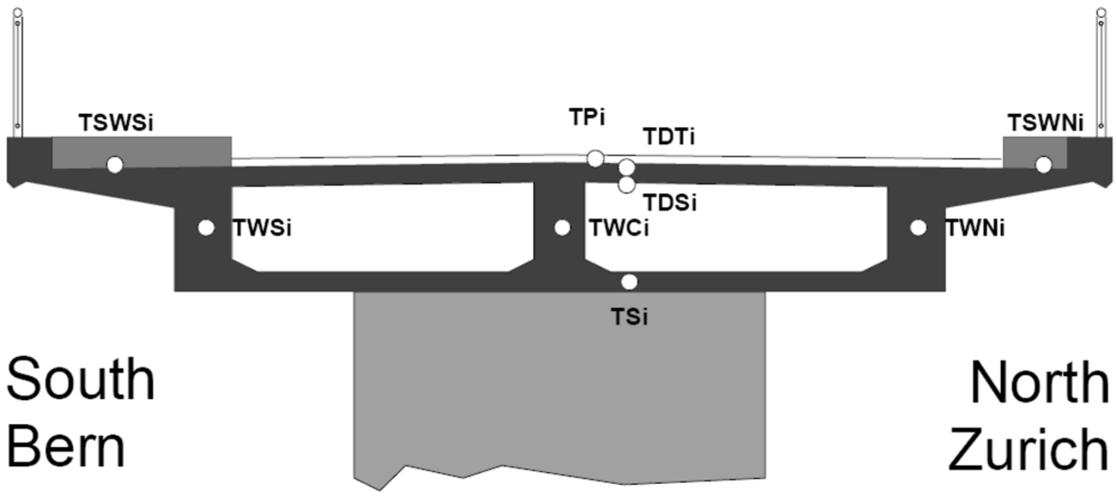


Figure 4.2: Cross section of the girder, showing the locations where the temperature Was measured

4.2 Temperature's Impact on Bridge Stiffness

Z24 bridge monitored the dynamic response of the structure and environmental conditions for a period of 12 months. During this monitoring period, one of the key environmental conditions that were measured was temperature. The measurements included air temperature and deck temperature, which provided valuable insights into the effect of temperature on the bridge's stiffness.

Young's modulus decreases rapidly up to a temperature of around 400 K, and then decreases more slowly or becomes constant at higher temperatures.[44]

The study conducted on the Z24 bridge revealed that temperature variations have a significant impact on the stiffness of the structure. The temperature measured during the monitoring period ranged from -10°C to 35°C, covering a wide range of seasonal temperatures. [45] The results showed that as the temperature decreased below 0°C, there was a noticeable change in the natural frequency of the bridge. This change in natural frequency indicates a variation in the stiffness of the bridge due to temperature fluctuations.

4.3 long-term progressive damage test

Monitoring the bridge took almost one year. Over a month, and just before the total dismantling of the bridge, a series of progressive damage tests were conducted. The procedure of the damage test is described in table 4.1. The practical importance of these assessments was validated by ensuring their relevance to the bridge's safety and by simulating common forms of damage. Table 4.1 provides a comprehensive catalog of all the progressive damage tests executed. The bridge was exposed to both a forced and ambient operational vibration test before and after each damage scenario was applied.

Date (1998)	Scenario
4-Aug	Undamaged condition
9-Aug	Installation of pier settlement system
10-Aug	Lowering of pier, 20 mm
12-Aug	Lowering of pier, 40 mm
17-Aug	Lowering of pier, 80 mm
18-Aug	Lowering of pier, 95 mm
19-Aug	Lifting of pier, tilt of foundation
20-Aug	New reference condition
25-Aug	Spalling of concrete at soffit, 12 m2
26-Aug	Spalling of concrete at soffit, 24 m2
27-Aug	Landslide of 1 m at abutment
31-Aug	Failure of concrete hinge
2-Sep	Failure of 2 anchor heads
3-Sep	Failure of 4 anchor heads
7-Sep	Rupture of 2 out of 16 tendons
8-Sep	Rupture of 4 out of 16 tendons
9-Sep	Rupture of 6 out of 16 tendons

Table 4.1: Test procedure on Z24 bridge

4.4 Time-series frequency of Z24 bridge

signals were generated using an inverse fast Fourier transform (FFT) algorithm, resulting in a fairly flat force spectrum between 3 and 30 Hz. The results of the frequency are shown in figure 4.3.

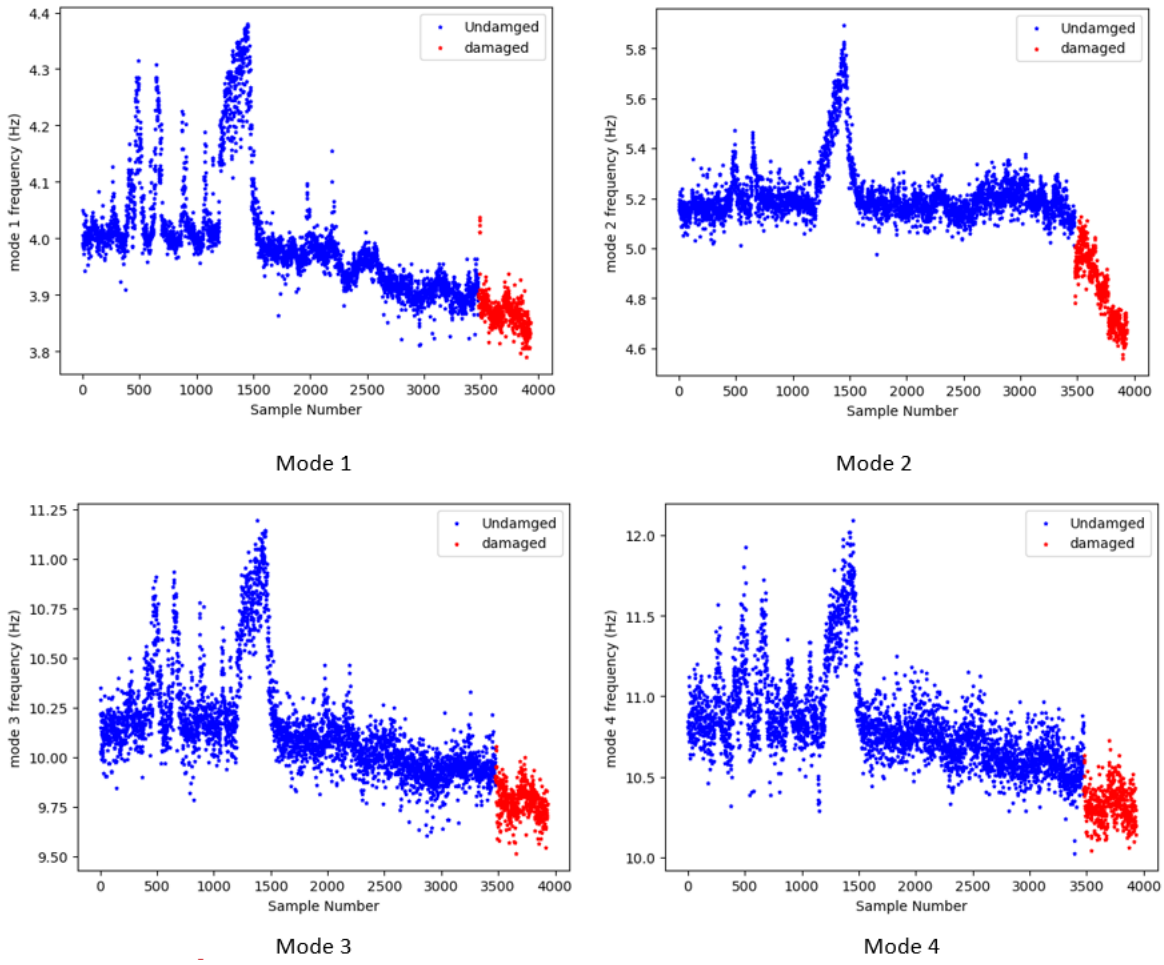


Figure 4.3: time-series frequency of the Z24 bridge

As mentioned before, temperature change can affect the frequency of the bridge based on the stiffness change of the material. In this case, changing in the Asphalt layer stiffness caused a fluctuation on time-series frequency. Figure 4.4 shows the highly affected frequencies by the temperature changes for one year monitoring. Area highlighted by yellow color is where the temperature effect is high on stiffness.

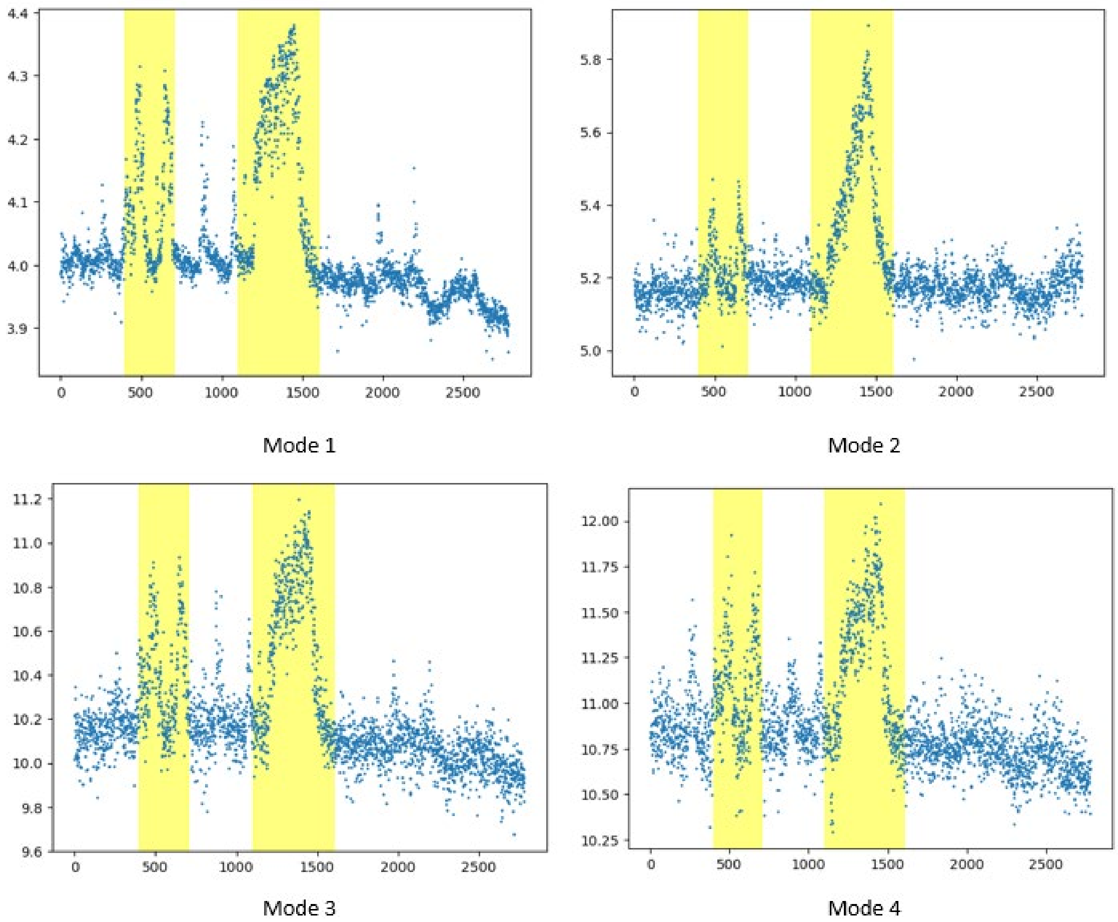


Figure 4.4: Effect of the temperature on time-series frequency

Similarly, the comparison between the different modes of the bridge frequency are shown in figures 4.5-4.8.

The initial four frequency modes are accessible, with each mode comprising 3932 observations. As is evident from the frequency mode graphs, there is a clear relationship between temperature and the bridge's natural frequencies. Fluctuation in the sample range of 400 to 600 and also 1200 to 1400 is evident. To remove the effect of the environment two possible methods are suggested that describe in detail.

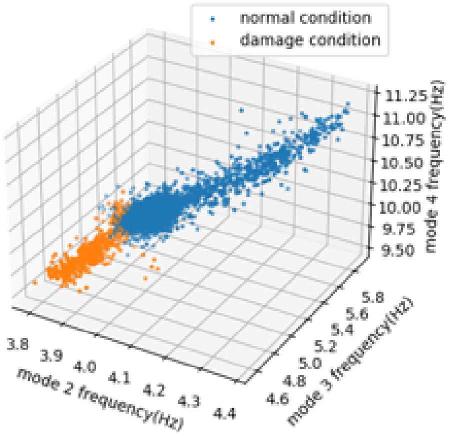


Figure 4 .5: relation between three first modes

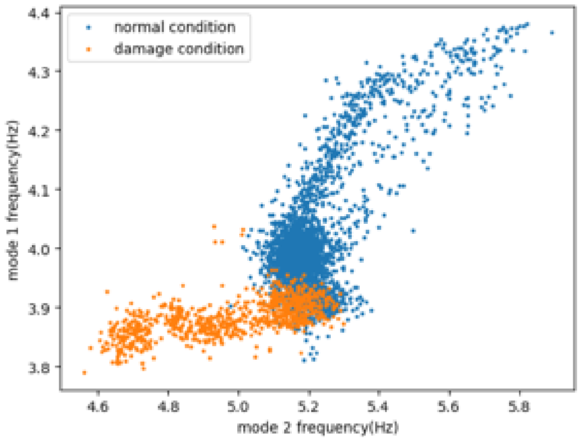


Figure 4.6: comparison between mode 1 and mode 2

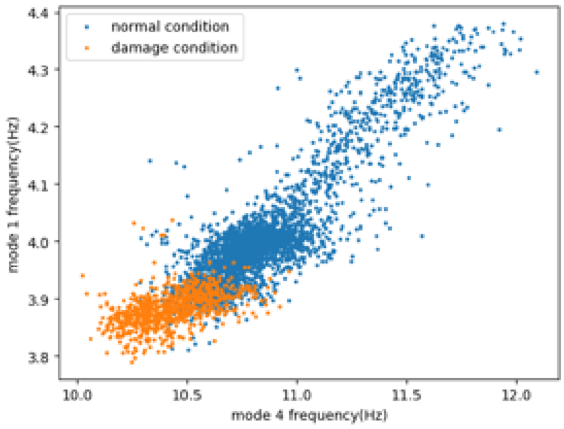


Figure 4.7: comparison between mode 1 and mode 4

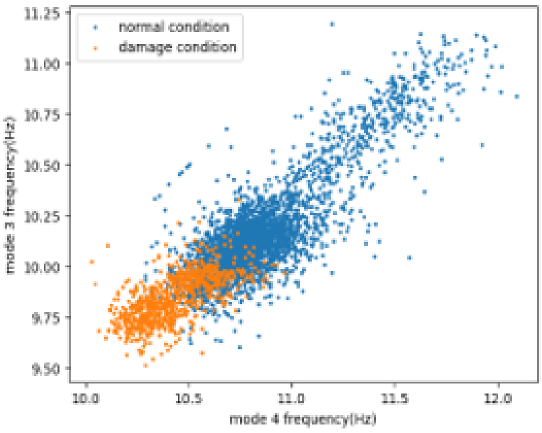


Figure 4.8: comparison between mode 3 and mode 4

4.5 Data preparation

The dataset of the Z24 bridge consists of a total of 3932 samples, which include four different modes. The first 3475 samples represent the normal condition of the Z24 bridge, while the remaining 457 samples correspond to the damaged state. The data set is divided into training and testing parts. First 80 % of the undamaged state of the structure was allocated to the training dataset $X \in \mathbb{R}^{(2780*4)}$ and the rest of the data was used as the testing in dataset $Z \in \mathbb{R}^{(1152,*4)}$.

Once the training and testing data sets were established, a methodology was implemented to detect early damage in the Z24 bridge.

The methodology used for early damage detection in the Z24 bridge involved several steps. First, to remove the temperature effect from the data, a preprocessing procedure was applied. Two methods were used for preprocessing the data to remove the noise and temperature effect on the bridge, DBSCAN and Autoencoder. The results of the damage detection then compared.

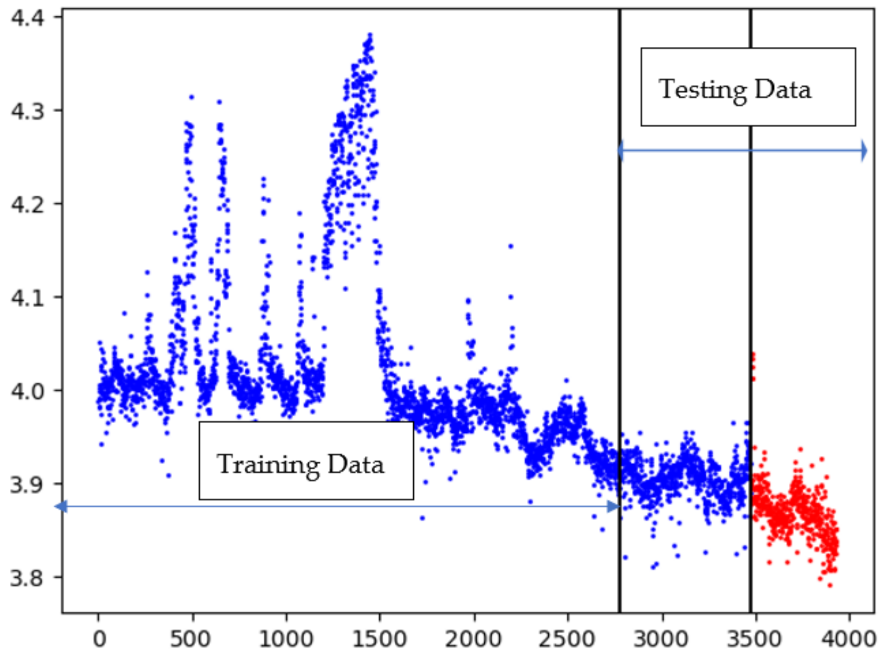


Figure 4.9: Training and Testing data division

	mode 1	mode 2	mode 3	mode 4
count	3932.000000	3932.000000	3932.000000	3932.000000
mean	3.986413	5.168640	10.109683	10.774179
std	0.105750	0.165479	0.270506	0.314407
min	3.790092	4.560032	9.513241	10.024860
25%	3.914156	5.139031	9.944156	10.585973
50%	3.975596	5.177900	10.073293	10.747572
75%	4.011745	5.220038	10.194632	10.899004
max	4.379953	5.892526	11.195249	12.091832

Table 4.2: specific statistic of the frequency-time series

4.6 Outlier detection by using DBSCAN

Density-Based Spatial Clustering of Applications with Noise is a popular and effective clustering algorithm used in data mining and machine learning. It was first introduced in 1996 by Martin Ester.[46] DBSCAN is particularly useful for discovering clusters in spatial datasets, where the proximity between data points plays a crucial role. DBSCAN differs from traditional clustering algorithms, such as k-means or hierarchical clustering, in that it does not require the number of clusters to be specified prior to running the algorithm.[47] Instead, DBSCAN identifies clusters based on the density of data points in the dataset. The algorithm operates by defining a neighborhood around each data point and determining whether the density within that neighborhood meets certain criteria.[32] Specifically, DBSCAN requires two parameters to be specified: epsilon (ε) and minPts. Epsilon (ε) defines the radius within which a data point is considered a neighbor of another data point. In Appendix B the implementation of DBSCAN in python environment is available.

This parameter determines the distance threshold for clustering and plays a crucial role in identifying clusters. The minPts parameter specifies the minimum number of data points that must be within a neighborhood in order for a point to be considered a core point. DBSCAN proceeds by iteratively expanding clusters starting from core points.[48] During the algorithm, a core point is defined as a data point that has at least minPts neighbors within its ε neighborhood. Once a core point is identified, DBSCAN recursively expands the cluster by adding all reachable points within ε distance. The DBSCAN algorithm can detect clusters of arbitrary shapes and sizes in the dataset.[49]

One of the advantages of DBSCAN is its ability to handle datasets with noise or outliers. These are data points that do not belong to any cluster and may be scattered throughout the dataset. DBSCAN can classify these noisy points as outliers and exclude them from the identified clusters, effectively improving the quality of clustering results by focusing on the essential patterns in the data.[34,50]

The algorithm of the DBSCAN was implemented on the Z24 bridge frequency data. The result shows in figure 4.10. The outlier data detected in the frequency data were successfully classified as noise points by the DBSCAN algorithm.

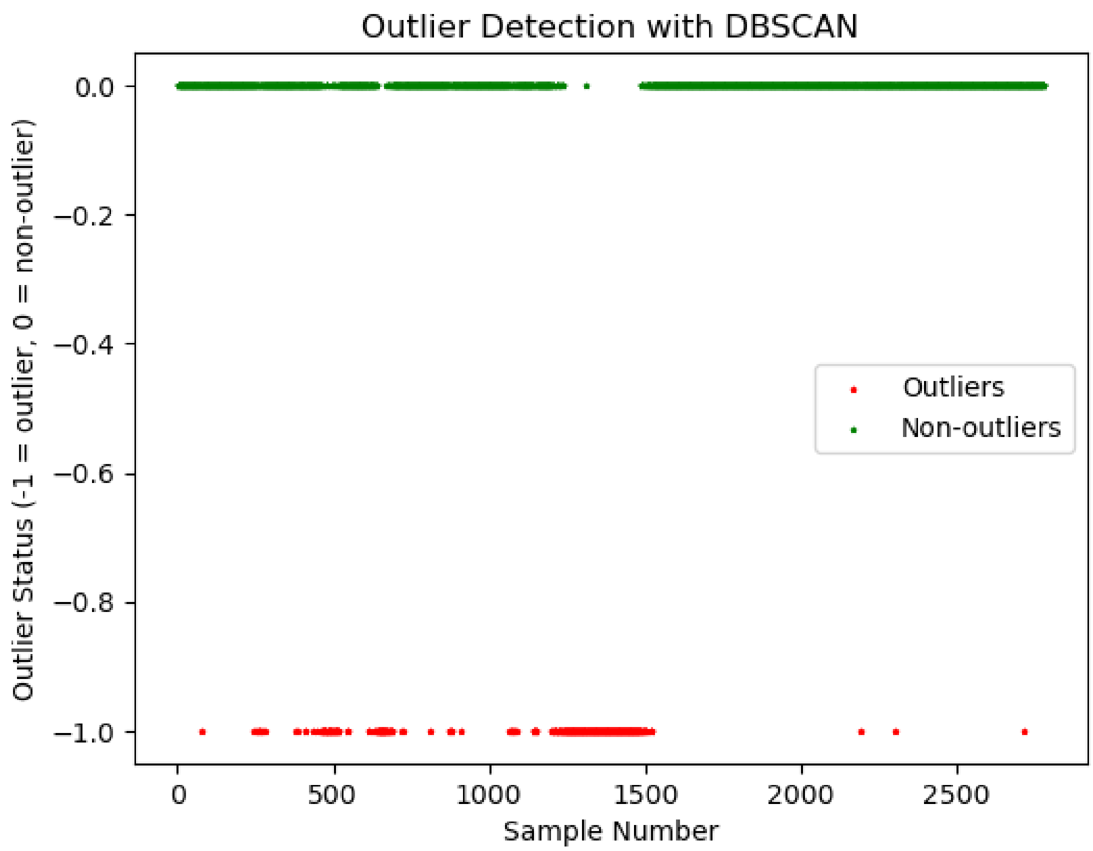


Figure 4.10: outlier data detection by using DBSCAN method

Upon examination of the timeline of severe environmental effects, it is evident that the outlier data has been correctly identified. Therefore, to obtain more reliable data, it would be prudent to remove the aforementioned outlier data. To be more specific, we can observe that the number of samples with an outlier value of -1 is 438. These outlier samples, which are clearly distinct from The DBSCAN algorithm considers these points as noise or outliers due to their dissimilarity with the majority of the data points by applying the DBSCAN algorithm, we can accurately identify and classify these outlier points as noise by removing these noise points, we can ensure that the subsequent analysis and interpretation of the data are not biased by outliers.

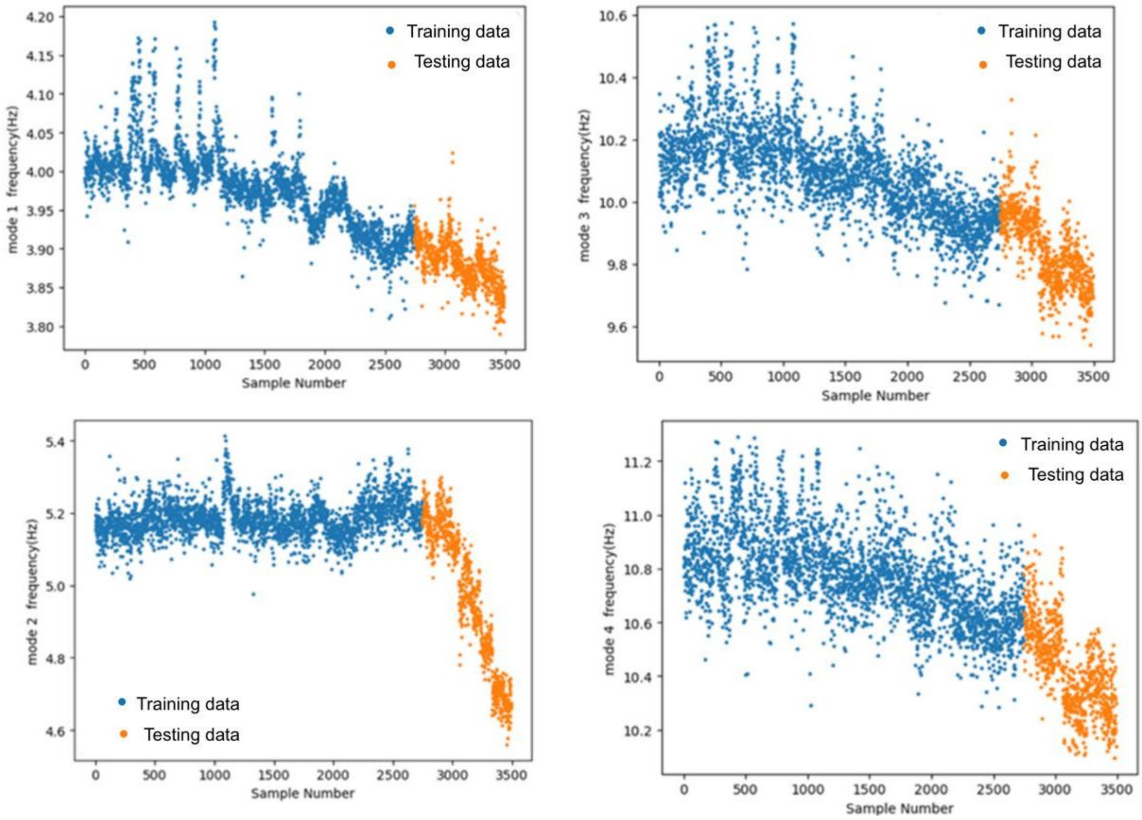


Figure 4.11: Remove the outlier data

4.7 Autoencoder outlier detection

To further enhance the detection and removal the effect of outliers in the dataset, an autoencoder outlier detection method can be applied. Autoencoder outlier detection is a technique that uses an autoencoder neural network to identify and remove outliers from a dataset.[51]

The autoencoder model is trained to reconstruct the input data, and during this process, it learns to effectively capture the underlying patterns and structures of the data while disregarding the outliers.[52] After training the autoencoder model, the reconstruction error is calculated for each data point. Points with high reconstruction errors are likely to be outliers, as they cannot be accurately reconstructed by the model. By setting a threshold for the reconstruction error, we can classify these points as outliers and exclude them from further analysis.

Another advantage of using the autoencoder outlier detection method is that it can handle datasets with complex and non-linear relationships. This is particularly useful for time series data, as it can capture the temporal dependencies and identify outliers that may arise from unexplained phenomena.[53]

The process of applying the autoencoder outlier detection method involves several steps. First, the dataset is divided into a training set and a test set. The autoencoder model is then trained on the training set, where it learns to reconstruct the input data accurately.[54] Next, the trained autoencoder is used to reconstruct the data from the test set. The reconstruction errors for each data point in the test set are computed by comparing the original data with its reconstructed counterpart. Based on the premise that an autoencoder trained on normal data will yield low reconstruction error for similar distribution, points with high reconstruction errors can be identified as potential outliers.

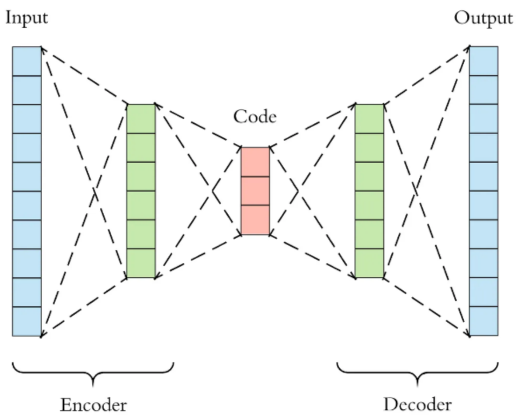


Figure 4.12: Autoencoder architecture

The mathematical equations that describe the autoencoder can be written as follows:

$$\text{1.Encoder: } \varnothing: X \rightarrow F \tag{2}$$

where x is the input data, f is the encoder function, and z is the compressed latent representation of x.

$$\text{2.Decoder: } y: F \rightarrow X \tag{3}$$

where g is the decoder function, and x' is the reconstructed output

$$\text{3.Loss function: } \iota(x, x') = ||x - x'||^2 = ||x - \sigma'(W'(\sigma(Wx + b))^2 + b'||^2 \tag{4}$$

$$z = \sigma(Wx + h) \tag{5}$$

$$x' = \sigma'(W'z + b') \tag{6}$$

The process of applying the autoencoder outlier detection method involves several steps. First, the dataset is divided into a training set and a test set. The autoencoder model is then trained on the training set, where it learns to reconstruct the input data accurately. Next, the trained autoencoder is used to reconstruct the data from the test set. The reconstruction errors for each data point in the test set are computed by comparing the original data with its reconstructed counterpart. Based on the premise that an autoencoder trained on normal data will yield low reconstruction error for similar distribution, points with high reconstruction errors can be identified as potential outliers.

ReLU (Rectified Linear Unit) activation function is commonly used in autoencoder models for outlier detection. This function helps in preserving the linearity of the autoencoder model while also providing non-linearity to capture complex patterns in the data.[55]

The ReLU activation function, given by f(x) = max(0, x), is commonly used in autoencoder models for outlier detection tasks.

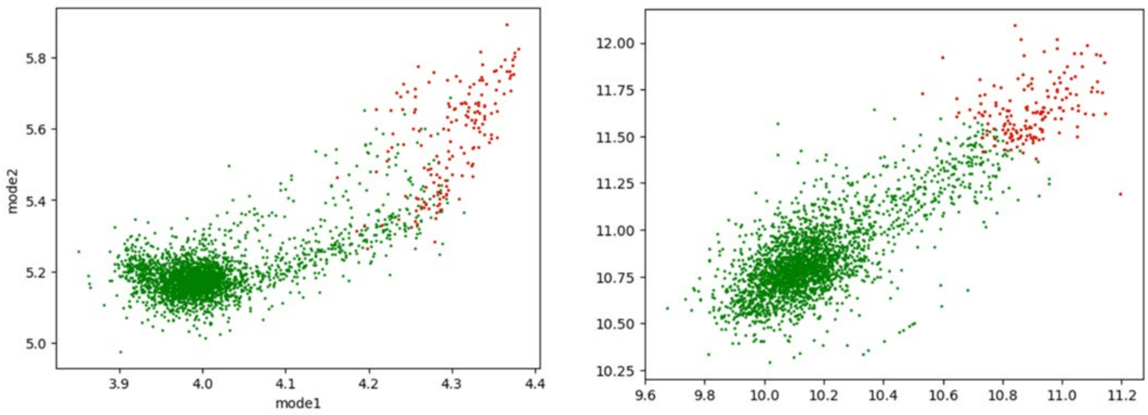


Figure 4.13: (a) mode 1 and mode 2 relation (b) mode 3 and 4 relation

4.8 Clustering data

The K-means clustering method was executed on the Z24 bridge data, by using the results of the autoencoder. The optimal cluster count was determined by implementing the gap statistic method. Gap statistics were used to determine the number of clusters in the data set. By this approach, the optimal number of clusters was determined. Figure (4.14) shows the gap statistics results of the Z24 bridge dataset, indicating the optimal number of clusters.

Compute Clustering for Different Values of K: Perform a k-means clustering for different values of k and for each k, calculate the sum of the squared distances between each data point and its cluster centroid. This sum of squared distances is often denoted as Wk.

Compute Clustering for Reference Datasets: For each reference dataset, perform k-means clustering and calculate Wk for each one (Wkb).[56]

After obtaining the values of Wk for different values of k, compute the gap statistic for each value of k. The gap statistic is calculated as follows :

$$\text{Gap}(k) = \text{average}\,(\log(\text{Wkb})) - \log(\text{Wk}) \tag{7}$$

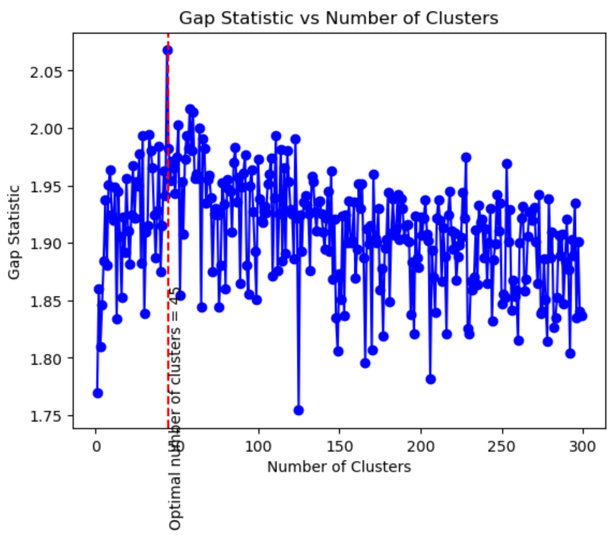


Figure 4.14: Gap statistic value for different number of cluster

GMM clustering is another method that can be used to determine the appropriate number of clusters. GMM, or Gaussian Mixture Model, clustering is a probabilistic model that assumes the data points are generated from a mixture of multiple Gaussian distributions.

Gaussian Mixture Models are a type of machine-learning method that is widely used for clustering and probabilistic modeling tasks. These models assume that the data points are generated from a mixture of Gaussian distributions. The main idea behind Gaussian Mixture Models is to represent the distribution of data points as a combination of multiple normal distributions[57]. This allows for capturing the complex underlying structure of the data by representing it as a combination of simpler components.

Gaussian Mixture Models have a simple formulation and are widely used in various applications such as databases, computer vision, and machine learning. The simple formulation and the widespread use of Gaussian Mixture Models make them an attractive choice for various applications. Gaussian Mixture Models have deep connections with clustering algorithms. Clustering refers to the process of grouping similar data points together based on their characteristics. Gaussian Mixture Models can be used to perform soft clustering, where data points are assigned probabilities indicating their membership in each cluster.

This ability to perform soft clustering is beneficial when dealing with complex data that may not necessarily belong to a single cluster. Such models have been particularly successful in applications where the underlying data distribution is multimodal or has overlapping clusters.[58]

The use of Gaussian Mixture Models in machine learning has greatly contributed to the field's ability to model and analyze complex data. Gaussian Mixture Models have applications in signal estimation, denoising, and image processing. The GMM equation represents the sum of weighted Gaussian components, where each component is represented by a mean vector and covariance matrix. These means and covariance matrices determine the shape, size, and orientation of each Gaussian component.

Bayesian Information Criterion (BIC) is commonly used to estimate the number of components, or clusters, in a Gaussian Mixture Model.[22]

The Bayesian Information Criterion is a commonly used method for estimating the number of components or clusters in a Gaussian Mixture Model.

It is a measure of the trade-off between model complexity and goodness-of-fit to the data. By evaluating the BIC for different numbers of components, one can determine the optimal number of clusters that best balance model complexity and data fit. The BIC is calculated using the log-likelihood of the data, penalized by a term that accounts for the number of parameters in the model. The availability of such an estimation method is crucial in practice, as it helps machine learning practitioners select the appropriate number of clusters for a given dataset without relying on subjective judgments or trial and error.

The BIC's solid theoretical support is one of its strongest points, as it provides a principled way to determine the optimal number of clusters. The BIC equation for evaluating the number of clusters in a Gaussian Mixture Model is:

$$BIC = -2 * \text{log-likelihood} + k * \log(n) \tag{8}$$

where log-likelihood is the logarithm of the likelihood of the data given the model, k is the number of parameters in the model, and n is the number of data points.

AIC is another criterion that can be used to estimate the number of clusters in a Gaussian Mixture Model.

In addition to the BIC, another criterion that can be used to estimate the number of clusters in a Gaussian Mixture Model is the Akaike Information Criterion. AIC, similar to BIC, is a measure that incorporates both the goodness-of-fit of the model and its complexity. However, AIC places less penalty on model complexity compared to BIC.

The Akaike Information Criterion is calculated using the log-likelihood of the data, also penalized by a term that accounts for the number of parameters in the model. The AIC equation for evaluating the number of clusters in a Gaussian Mixture Model is:

$$AIC = -2 * \text{log-likelihood} + 2 * k \tag{9}$$

where log-likelihood is the logarithm of the likelihood of the data given the model and k is the number of parameters in the model. Both AIC and BIC provide valuable insights into the trade-off between model complexity and data fit when estimating the number of clusters in a Gaussian Mixture Model.[59]

MDL is another criterion that can be used to estimate the number of clusters in a Gaussian Mixture Model. MDL, or Minimum Description Length, is a criterion that balances the complexity of the model with its ability to explain the data. The MDL criterion aims to find the model that provides the most concise description of the data.

It achieves this by minimizing the sum of two terms: the code length required to describe the data using the model, and the code length required to describe the model itself.

By minimizing this combined code length, MDL selects the number of clusters that achieves the most efficient and accurate representation of the data. The equation describes MDL is :

$$MDL = \text{log-likelihood} - (0.5 * k * \log(n)) \quad (10)$$

where log-likelihood is the logarithm of the likelihood of the data given the model, k is the number of parameters in the model, and n is the number of data points.

CV is another approach that can be used to estimate the number of clusters in a Gaussian Mixture Model. CV, or Cross-Validation, is a technique that assesses the performance of a model by splitting the dataset into training and validation subsets. The model is then trained on the training subset and evaluated using the validation subset. This process is repeated multiple times, with different subsets used for training and validation. The performance of the model is then averaged over these iterations to obtain an estimate of its generalization capability.[60]

Cross-Validation can be used to estimate the number of clusters in a Gaussian Mixture Model by evaluating the model's performance on different subsets of the data and selecting the number of clusters that consistently yield the best performance.

The CV equation for evaluating the number of clusters in a Gaussian Mixture Model is as follows:

$$CV = 1/n * \Sigma(\text{log-likelihood}) \quad (11)$$

where n is the number of data points and log-likelihood is the logarithm of the likelihood of the data given the model. In conclusion, the determination of the number of clusters in a Gaussian Mixture Model can be achieved using various methods such as the Minimum Description Length criterion and Cross-Validation. These approaches aim to strike a balance between model complexity and data explanation, ensuring that the chosen number of clusters provides an accurate representation of the data using a Gaussian Mixture Model.

Optimal number of components (BIC)	5
Optimal number of components (MDL)	49
Optimal number of components (CV)	3
Optimal number of components (AIC)	36

Table 4.3: Optimal number of component

4.9 proposed damage index

The proposed damage index is distance based and relies on the concept of Density-Based Spatial Clustering. K-means clustering is a popular clustering algorithm that aims to partition data points into clusters by minimizing the within-cluster sum of squares.

K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid). Distance-based damage index calculated by the Euclidean distance between each data point and the centroid of its cluster can provide insights into the outliers. The equation of the proposed damage index can be written as:

$$\text{Distance} = \sqrt{\sum(xi - ci)^2} \qquad (12)$$

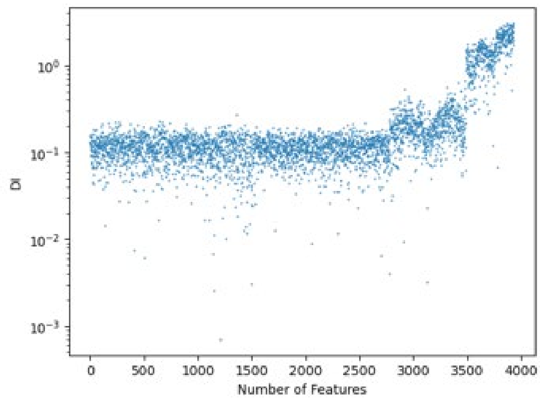
Damage index is defined based on the assumption that damage will lead to a reduction in local density and an increase in the minimum distance between samples. This is because damage may cause the structure to become more rigid, resulting in a larger separation between adjacent nodes and a reduction in the number of nodes in the vicinity of the damaged area.

The damage index can be calculated using the following equation:

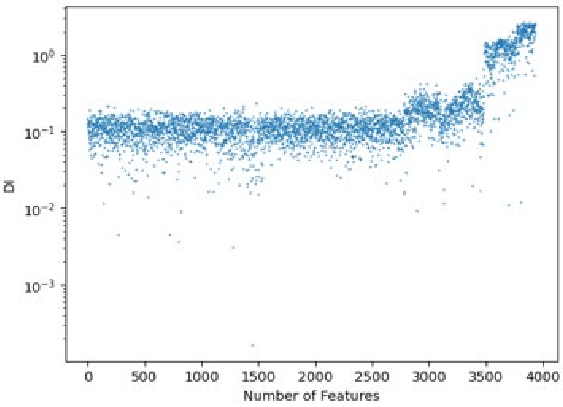
Damage Index = Local Density * Minimum Distance

Figure (4.15) shows the damage Index by different criteria for the Z24 bridge dataset.

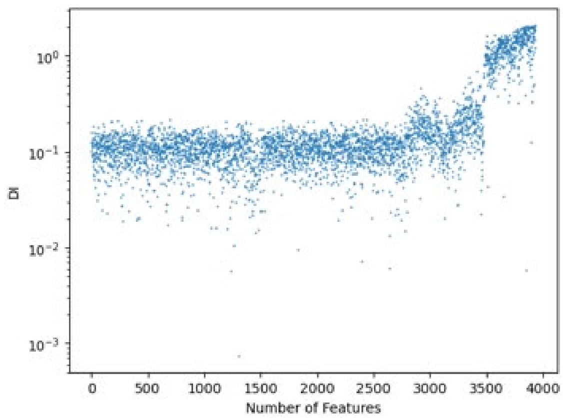
by applying the GEV distribution modeling, a threshold boundary for decision making was determined.



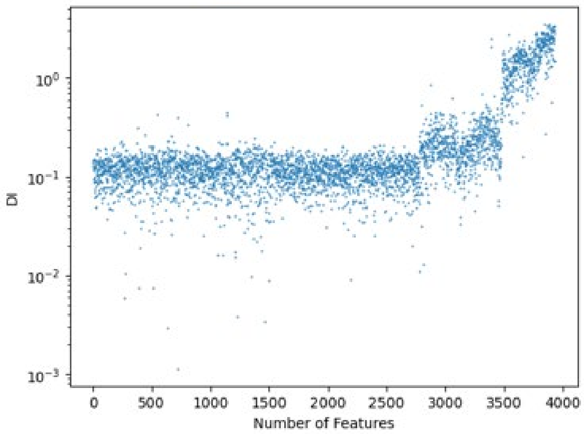
(a) -> AIC



(b) -> BIC



(c) -> CV



(d) -> MDL

Figure 4.15: damage Index by different criteria

4.10 Setting the threshold value

Setting the threshold value for the damage index is crucial for decision-making. One approach to setting the threshold value is by utilizing a statistical method such as Extreme Value Theory, Generalized Extreme Value, and Gaussian Mixture Model.

.

Generalized Extreme Value (GEV) distribution and threshold estimation can be used to determine a reliable threshold limit.[41] The threshold value can be set by fitting the Generalized Extreme Value distribution to the data above a certain threshold.

The GEV distribution is defined by three parameters: location parameter μ, scale parameter σ, and shape parameter ξ. The location parameter shifts the distribution to the left or right, the scale parameter controls the spread of the distribution, and the shape parameter controls the tail behavior of the distribution.

The probability density function (PDF) of the GEV distribution is given by:

$$f = (x, \mu, \sigma, \varepsilon) = \left(\frac{1}{\sigma}\right)((1 + \varepsilon\left(\frac{x-\mu}{\sigma}\right))^{(-\left(\frac{1}{\varepsilon}+1\right)))} e^{(-\left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}))} \quad (13)$$

by applying the GEV distribution modeling, a threshold boundary for decision-making was determined.

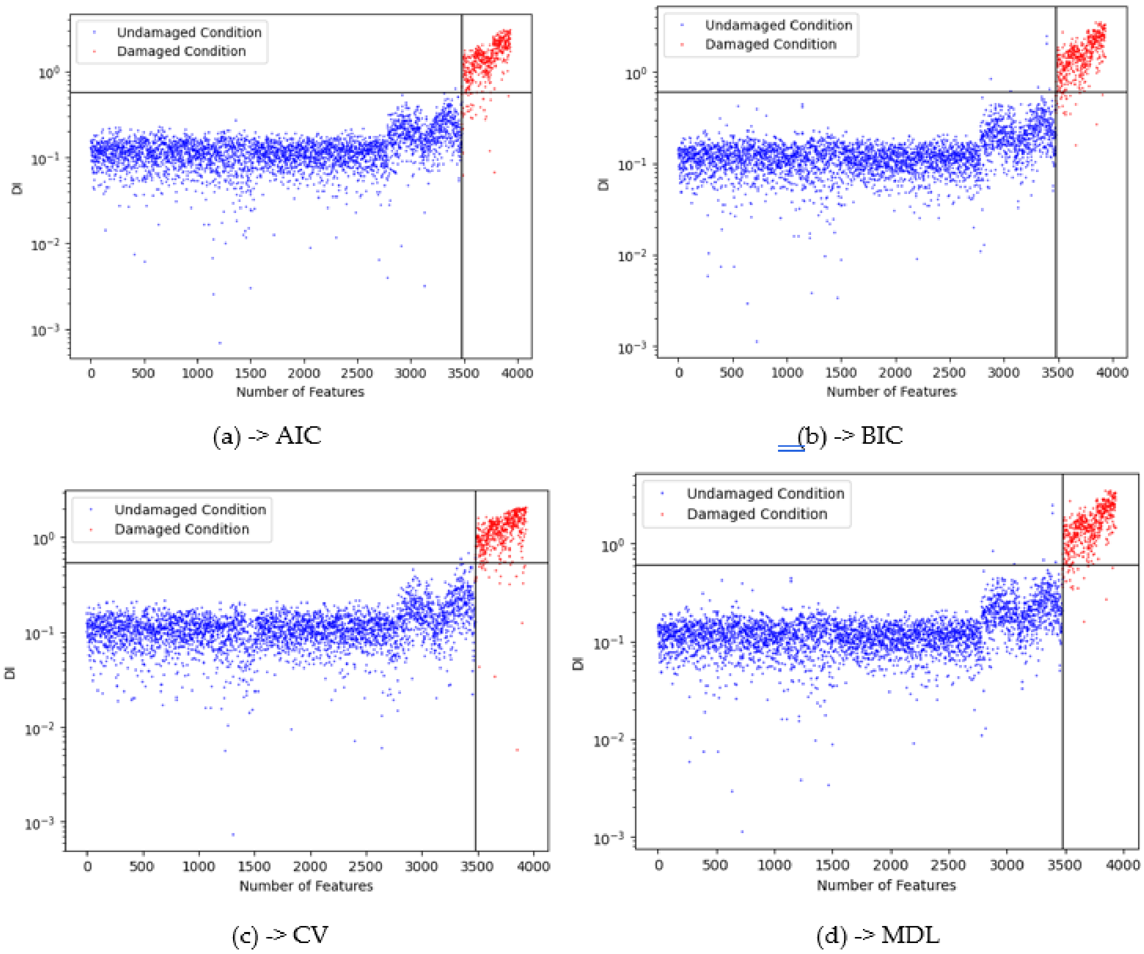


Figure 4.16: unsupervised damage detection results on Z24 bridge

Figure (4.17) shows the best unsupervised damage detection results by MDL criteria bridge on Z24. Table 4.4 summarized the efficiency of the proposed method to deal environmental effect on structure.

Error	CV	MDL	BIC	AIC
False positive	3 (0.08%)	5 (0.16%)	4(0.11%)	2(0.05%)
False negative	14 (3.03%)	11(2.3%)	19(4.11%)	16(3.43%)

Table 4.4: summarized the efficiency of the proposed method

4.11 Verification of the proposed method on KW51 bridge

In this section, the best efficient method regarding vibration-based damage detection in the previous section will be applied to a case study involving the KW51 Bridge.

4.11.1 KW51 bridge details

The KW51 bridge is a railway bridge located on railway line L36N between Leuven and Brussels, Belgium. It enables the crossing of the canal Leuven-Mechelen. The bridge is of the bowstring type and has a length of 115 m and a width of 12.4 m.[62]

A 15-month monitoring campaign was conducted on the KW51 bridge between 2018 and 2019 to collect data for vibration-based damage detection.

During this monitoring campaign, various parameters were recorded, including the acceleration response, strains in the deck and gauges of the rails, displacement at the bearings, and environmental data. The acceleration responses were processed using Operational Modal Analysis to obtain the first 14 natural frequencies of the structure. This comprehensive data collection allowed for a detailed analysis of the bridge's dynamic behavior and provided valuable insights into its structural health.

Figure 4.17: KW51 Bridge

4.11.2 Available Data

The frequency of the KW51 bridge is available as a result of the 15-month monitoring campaign conducted between 2018 and 2019. In addition to the frequency data, other parameters such as acceleration response, strains in the deck and gauges of the rails, and displacement at the bearings were also recorded.[62] With this rich dataset, the vibration-based damage detection method can be applied to identify potential damages in the KW51 bridge.

4.11.3 Damage detection approach

A total of 3130 natural frequency data points are available for the KW51 bridge in four modes. The time series frequency is shown in Figure (4.18) The highlighted part in the plot denotes the temperature effect on the frequency data. As it is considerable, the effect of the weather condition on the KW51 bridge is lower than on the Z24 bridge.

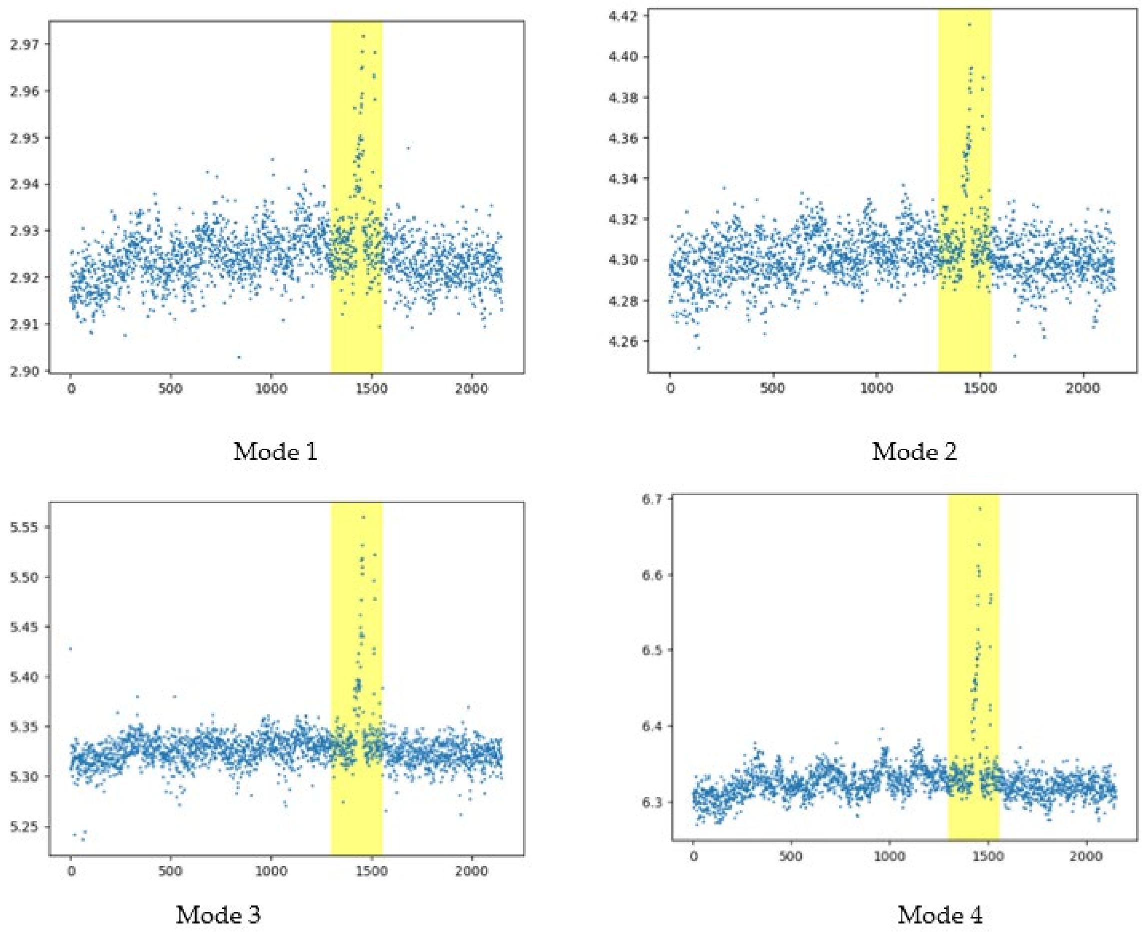


Figure 4.18: time-series frequency of the KW51 bridge in mode1, mode2,mode3,mode 4

In order to detect potential damages in the KW51 bridge, the vibration-based damage detection approach will be applied. The outlier data was detected as described in section 4.7 by using the Autoencoder method. the results of the outlier detection are shown in Figure 4.19 by comparing the two modes of frequency for better illustration.

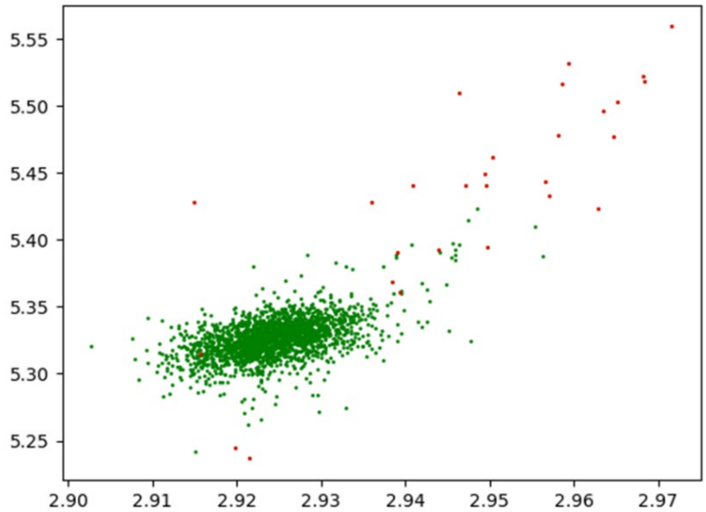


Figure 4.19: Autoencoder result by comparing mode1 and mode 3

The results show that the vibration-based damage detection approach successfully identified outliers in the natural frequency data of the KW51 bridge.

The clustering machine learning algorithm was then applied to further analyze the data. The data set is divided into training and testing parts. First 80 % of the undamaged state of the structure was allocated to the training dataset $X \in \mathbb{R}^{(2150*4)}$ and the rest of the data was used as the testing in dataset $Z \in \mathbb{R}^{(978*4)}$.

using the same methodology as performed to detect damage in the Z24 Bridge, a GMM approach was developed to optimally cluster the data of the KW51 bridge.

The damage index is proposed as the multiplication of the density and Euclidean distance from the closest cluster centroid.

Figure(4.20) shows the damage index plot for the KW51 bridge. The damage index plot for the KW51 bridge clearly indicates areas of potential damage.

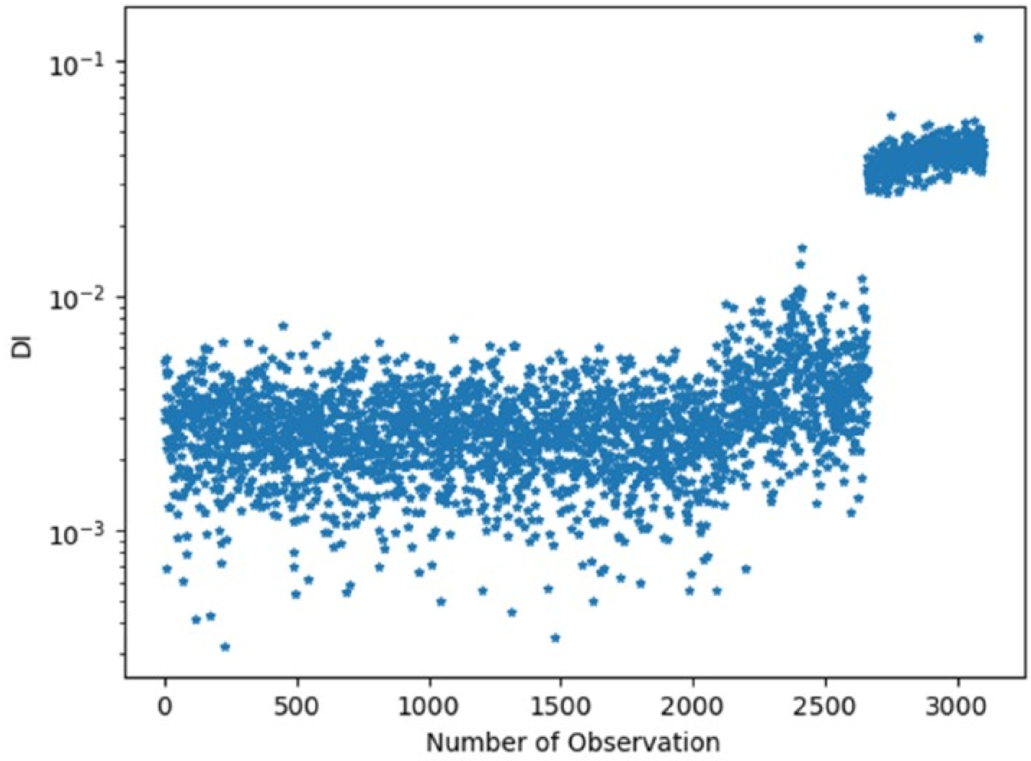
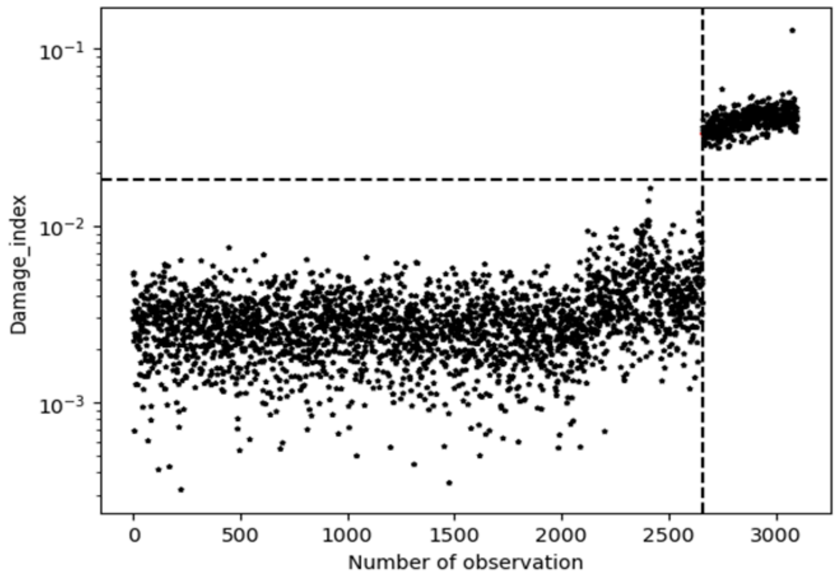


Figure 4.20: Damage index of KW51 bridge samples



False positive	1
False negative	0

Figure 4.21: vibration-based damage detection result on KW51 bridge

Chapter 5

Damage Detection using hybrid approach

Hybrid approaches are gaining popularity in the field of damage detection due to their ability to combine multiple techniques and algorithms, resulting in more accurate and reliable results.[63]

These hybrid approaches leverage both signal processing methods and artificial intelligence techniques to effectively identify and locate structural damage.

By integrating advanced signal processing methods, such as wavelet transforms or Fourier analysis, with artificial intelligence techniques like neural networks or genetic algorithms, hybrid approaches can extract meaningful information from sensor measurements and effectively detect damage. Moreover, hybrid approaches also have the advantage of combining the best aspects of model-based and data-driven techniques . For example, one study proposed a hybrid approach for damage state prediction by combining a crack growth model with regression techniques.[64] This approach not only predicted damage states accurately but also allowed the model to be updated as new measurements became available. This demonstrated the superiority of the hybrid approach compared to solely relying on either data-driven or model-based methods . Furthermore, the use of hybrid sensing approaches, which combine experimental data with results from physics-based multiscale models, has proven to be highly effective in damage detection.[65] By incorporating both data-driven and model-based techniques, hybrid approaches address the limitations of individual methods used in damage detection.

Hybrid approaches in damage detection also utilize machine learning algorithms to improve the accuracy of classification. In large-scale civil engineering structures, where the data provided by structural health monitoring systems may be limited, hybrid approaches can provide more robust and reliable results.[66] One approach involves using a finite element model to simulate scenarios that were not recorded during monitoring sessions, such as the response of an undamaged structure under operational and environmental variability. This simulated data is then combined with actual sensor measurements and fed into an MLA for improved damage classification. By integrating both the FE model and machine learning algorithms, the hybrid approach can overcome the limitations of individual methods and enhance the performance of damage detection in structural health monitoring. Additionally, it has been observed that training data should cover the full range of operational and environmental variability for successful damage detection using machine learning algorithms [66]. By considering both the physical behavior of structures and the data-driven capabilities of machine learning algorithms, hybrid approaches in damage detection offer a comprehensive solution for structural health monitoring.

In conclusion, the use of hybrid approaches in structural health monitoring for damage detection offers a more accurate and reliable solution compared to relying solely on either data-driven or model-based methods. These approaches combine the strengths of both techniques to overcome their individual limitations.[67] Hybrid approaches have been shown to be superior in predicting damage states by incorporating crack growth models with regression techniques, allowing for continuous updating as new measurements become available. Furthermore, the integration of machine learning algorithms into hybrid approaches improves classification accuracy by utilizing finite element models to simulate scenarios that were not recorded during monitoring sessions. By considering both physical behavior and data-driven capabilities, hybrid approaches provide a comprehensive solution for damage detection in structural health monitoring.

5.1 Finite element calibration

Finite element calibration is an essential step in the hybrid approach for damage detection. It involves adjusting the parameters of a finite element model to accurately represent the behavior of a structure under different loading conditions. This calibration ensures that the simulated data generated by the finite element model closely matches the actual response of the structure, improving the accuracy of damage detection algorithms. The hybrid approach for damage detection also utilizes machine learning algorithms to improve classification accuracy. These algorithms, such as supervised learning or unsupervised learning techniques, can analyze the enriched data obtained from the finite element model and identify patterns or anomalies that indicate the presence of damage. Additionally, the hybrid approach incorporates advanced signal processing techniques to analyze the sensor measurements and extract relevant information.

These techniques, such as wavelet transforms or Fourier analysis, can effectively identify damage-induced signals from the measured data, even in the presence of noise or other environmental variability. By combining the strengths of both model-based and data-driven techniques, hybrid approaches in damage detection can overcome some of the limitations associated with individual methods.

The calibration of a finite element model is crucial in accurately representing the behavior of a structure and improving the accuracy of damage detection algorithms. Calibrating the finite element model involves adjusting its parameters, such as stiffness and mass distribution properties, to ensure that it accurately reflects the response of the structure under different loading conditions.[66] Calibration is typically performed by comparing the simulated response of the finite element model to measured data from the structure, such as modal parameters obtained through vibration testing or structural deformations measured through point cloud modeling. Through a process of iterative adjustments to the model's parameters, the calibration aims to minimize the discrepancies between the simulated response and the measured data. This calibration process allows for more accurate predictions of the structure's behavior and improves the reliability of damage detection algorithms. Finite element model calibration is a crucial step in accurately representing the behavior of a structure and enhancing the precision of damage detection algorithms. Proper calibration of the finite element model ensures that it accurately captures the dynamic behavior and response of a structure under different loading conditions. Additionally, the calibration process helps in reducing uncertainties associated with material properties and boundary conditions.[65] By calibrating the finite element model, the accuracy of damage detection algorithms can be significantly improved. The calibration of a finite element model is an important and necessary step in accurately representing the behavior of a structure and improving the precision of damage detection algorithms.

It ensures that the simulated response of the model aligns closely with the measured data from the structure, such as modal parameters or structural deformations.

This alignment allows for more accurate predictions of the structure's behavior and enhances the reliability of damage detection algorithms. Moreover, finite element model calibration helps in reducing uncertainties associated with material properties and boundary conditions, which further improves the accuracy of the model. The process of calibrating a finite element model involves iteratively adjusting its parameters to minimize discrepancies between simulated and measured responses. This iterative process ensures that the model accurately represents the dynamic behavior of the structure under different loading conditions. Overall, finite element model calibration plays a crucial role in enhancing our understanding of structural behavior and improving our ability to detect and assess damage in civil structures.

5.2 I-40 Bridge

One notable application of the hybrid approach in damage detection and finite element calibration is the I-40 Bridge.

The I-40 bridge is located along Interstate Highway 40 across the Rio Grande River in Albuquerque, New Mexico, USA. It features twin spans with separate highways for each traffic direction. This structure was constructed from a concrete deck which is 13.3 meters wide and 0.178 meters thick. The deck is supported by two steel plate girders and three steel stringers.[65]

A section of the bridge that was used in the experimental study consists of three spans with a total length of 129.5 meters. The first and third spans are of equal lengths, 39.9 meters each, and the middle span is 49.7 meters long.[66]

External loads caused by traffic from the stringers are transmitted to the steel plate girders by the floor beams, with cross bracing also provided between the floor beams.

The numerical model of the I40 bridge was constructed in the MATLAB environment using shell and 3D beam elements. The materials used for the I40 bridge were concrete and steel with the modulus of elasticity, density, and the Poisson's ratio needed for the numerical modeling being identical to 24.8 GPa, 2322.6 kg/m^3, and 0.2 for concrete and 210 GPa, 7850 kg/m^3, and 0.3 for steel, respectively. The model of the concrete deck is simplified using a constant thickness of 0.2209 m, neglecting the steel rebars.[66]

The finite element model of the I40 bridge was used to simulate forced vibration tests, similar to the ones conducted on the actual bridge.

Figure (5.1) shows the reduced finite element model of the portion of the I40 bridge on which experimental modal analyses were performed.

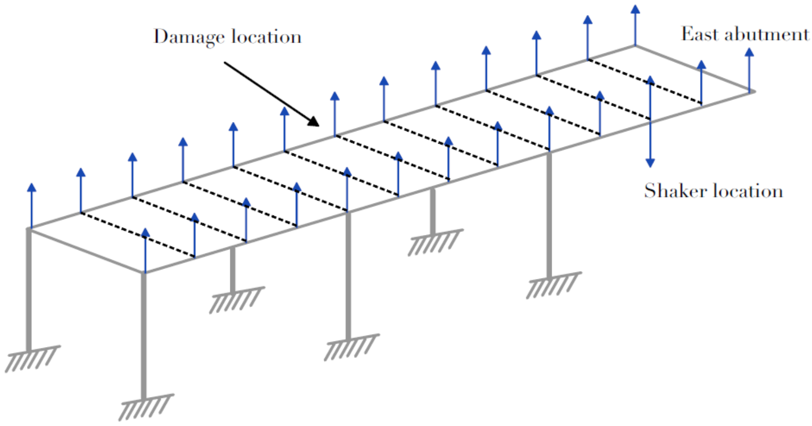


Figure 5.1: Experimental setup of the I-40 Bridge

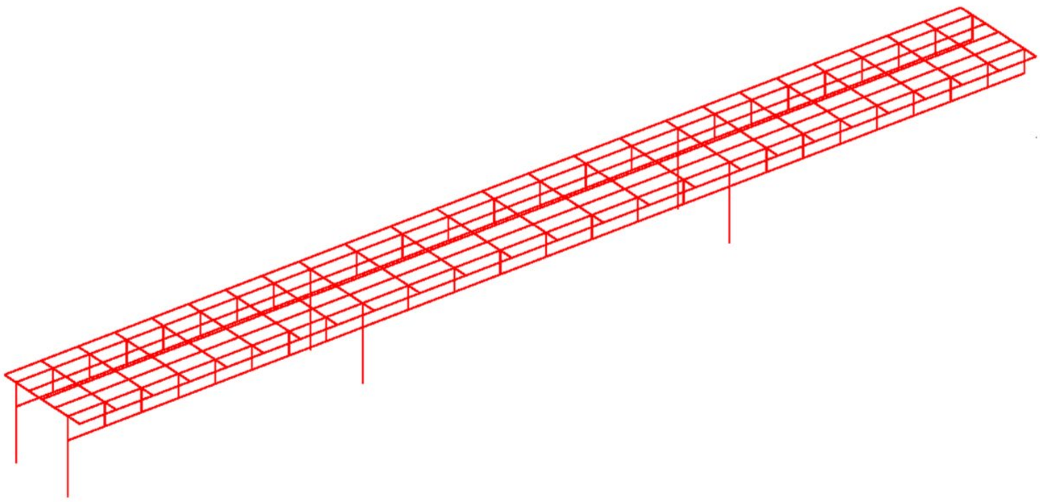


Figure 5.2: FE model built in Matlab

A total of 26 DOF of freedom in vertical directions was considered in the numerical model. Mode shapes of the structure obtained from experimental modal analysis were used as a reference for the calibration of the numerical model. The mode shapes of the I40 bridge were obtained by using the accelerometers. The mode shapes obtained from the experimental data may not be scaled due to the variability in the accuracy of the sensor readings.[65]

The introduced damage intends to simulate fatigue cracking that has been observed in plate girder bridges.[65] Four levels of damage are introduced to the middle span of the north plate girder. These different levels of damage are introduced by making various torch cuts in the web and flange of the girder, which is illustrated in Figure (5.3).

The first level consists of a 0.61m long, 10mm wide cut through the web; in the second level, this cut was extended to the bottom of the web. For the third level, the flange is cut halfway from both sides. At last, the bottom flange is completely cut to produce the fourth damage level. After each damage case, the structure is subjected to an experimental modal analysis.

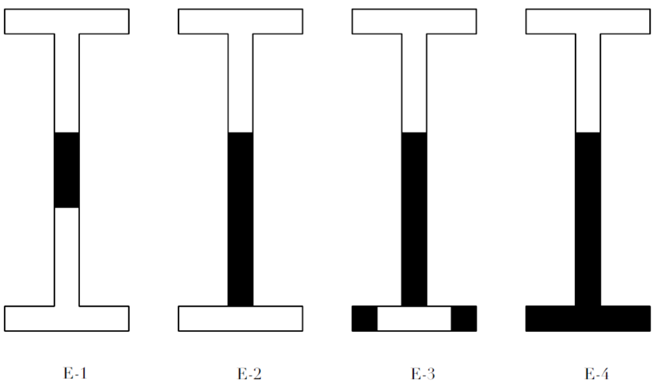


Figure 5.3: Damage scenarios introduced to the I-40 Bridge.

5.3 Damage identification method

Damage index based on the Modal stain energy: for the first damage identification algorithm, a damage index based on the modal strain energy is utilized. The damage index method evaluates the difference in modal strain energy between the undamaged and damaged structures based on the experimental data. for considering 6 modes obtained from experimental data the damage index is the sum of all 6 modes. The total length of the north girder is 420ft which is divided into 210 (2ft) segments of equal length.[66] Each segment is assigned a damage index value, calculated by summing the modal strain energy differences for each mode.

$$\beta ij = \frac{(\int_a^b [\Phi\mathrm{i}''*(\mathrm{x})]^2 dx + \int_a^L [\Phi\mathrm{i}''*(\mathrm{x})]^2 dx)}{(\int_a^b [\Phi\mathrm{i}''(\mathrm{x})]^2 dx + \int_a^L [\Phi\mathrm{i}''(\mathrm{x})]^2 dx)} \frac{\int_a^L [\Phi\mathrm{i}''(\mathrm{x})]^2 dx}{\int_a^L [\Phi\mathrm{i}''*(\mathrm{x})]^2 dx} \tag{4.1}$$

Where Φi is the mode shape of the structure, L is the length of the beam, a and b are the starting and ending point of the segment respectively.

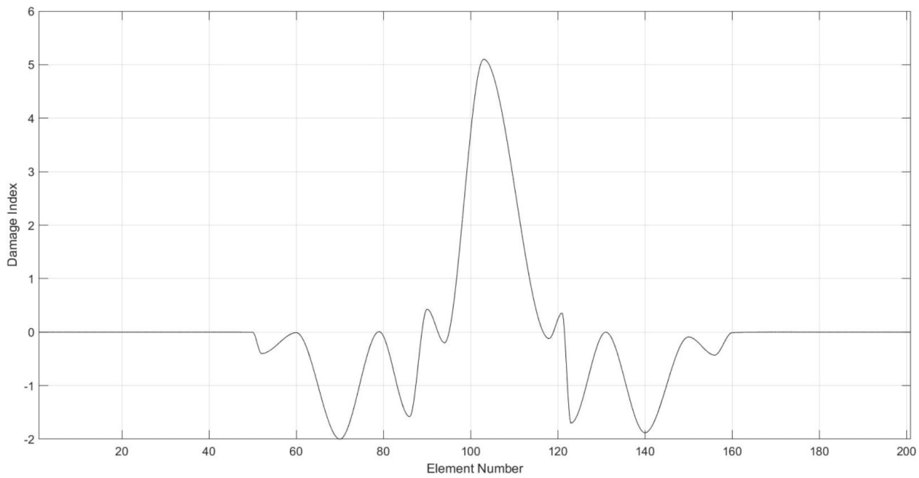


Figure 5.4: Damage index based on the modal strain energy

Change in Flexibility matrix: show that for the undamaged and damaged structure, the flexibility matrix, [F], can be approximated from the unit-mass-normalized modal data as follows: [66]

$$[F] \approx \sum_{i=1}^{n} \frac{1}{\omega i^2} \{\phi i\}\{\phi i\}^T \tag{14}$$

$$[F]^* \approx \sum_{i=1}^{n} \frac{1}{\omega i^{*2}} \{\phi i\}^*\{\phi i\}^{*T} \tag{15}$$

$$[\Delta F] = [F] - [F]^* \tag{16}$$

Where Φi is the unit mass normalized and the ωi is the natural frequency of the structure in i-th mode shape.

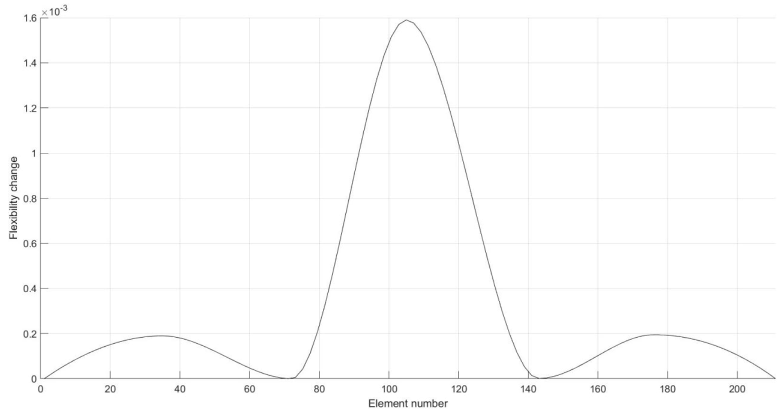


Figure 5.5: Damage index based on the flexibility change

5.4 comparing the mode shapes from experimental data and the FE model

To ensure the accuracy of the numerical model, mode shapes obtained from experimental modal analysis were used as a reference for the calibration process. Figure (5.6) shows the comparison between the mode shapes obtained from the experimental data and those predicted by the finite element model. The scaling must be done on the experimental data.[67] The scaling of the experimental data is necessary to ensure that the mode shapes obtained from both the experimental data and finite element model are comparable. To scale the experimental mode shapes the equation below used

$$\Phi^m = \nu * \psi \tag{17}$$

Where y is the experimental mode shape, Φ is the analytical mode shape and n is the factor obtains from the equations below

$$\upsilon = \frac{\Phi^{m^T} * \psi}{\psi^T * \psi} \tag{18}$$

Where ϕ is the normalized mode shape, Ψ mode shape from experimental data .Click or tap here to enter text.

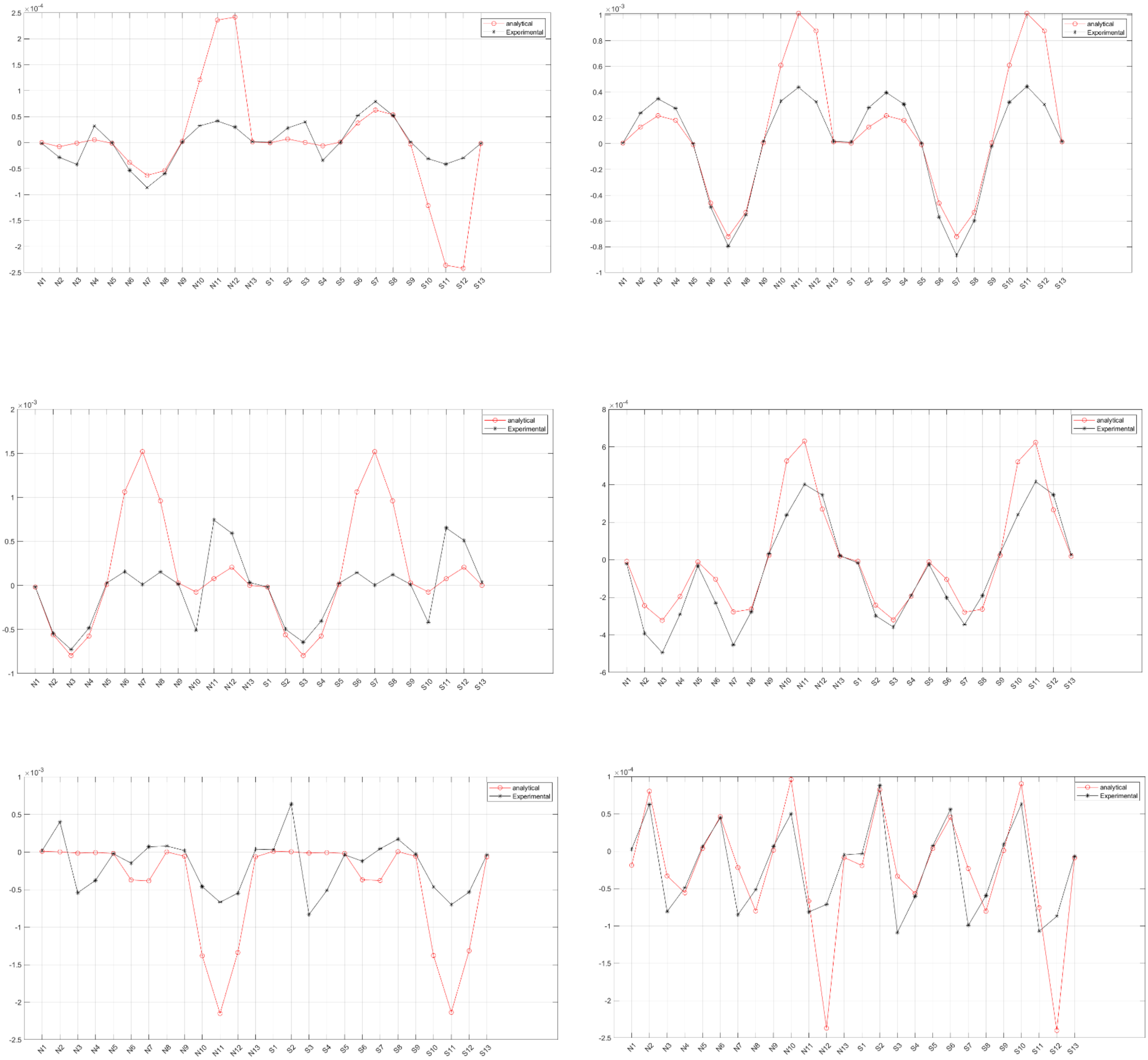


Figure 5.6: experimental mode shape comparison to analytical

The experimental mode shapes were compared to the mode shapes calculated by the finite element analysis using the Modal Assurance Criterion.

The Modal Assurance Criterion is a measure of similarity between mode shapes and is often used to compare the experimental mode shapes with the numerical mode shapes predicted by the finite element model.

This criterion quantifies the correlation between two sets of mode shapes through a scalar value ranging from 0 to 1, with higher values indicating greater similarity between the experimental and numerical mode shapes. Figure (5.7) shows the MAC value for the analytical and experimental mode shapes.

The MAC value can obtain by using the following equation: [65]

$$MAC\big((\varphi 1),(\varphi 2)\big) = \frac{|(\varphi 1^{*})(\varphi 2)|^{2}}{(|(\varphi 1^{*})(\varphi 1)|)|(\varphi 2^{*})(\varphi 2)|} \tag{19}$$

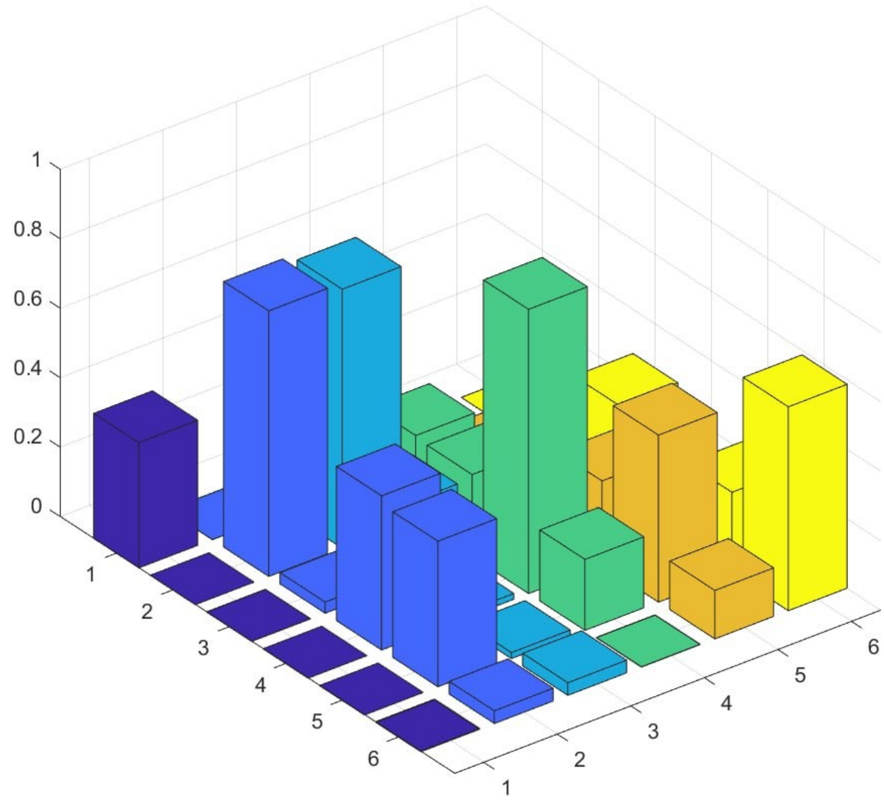


Figure 5.7: modal assurance criterion

5.5 Particle swarm optimization (PSO)

Particle Swarm Optimization is a metaheuristic optimization algorithm that was inspired by the flocking behavior of birds. It was first introduced by Kennedy and Eberhart in 1995 as a population-based optimization algorithm. The algorithm is based on the concept of a swarm, where individuals (or particles) move through the solution space in search of optimal solutions .

The algorithm works by initializing a population of particles, each representing a potential solution. These particles are then updated iteratively based on their own best-known position and the global best-known position of the swarm. During each iteration, the particles adjust their position based on two components: personal best and global best.[69] The personal best component refers to the particle's own historical best position, which represents the best solution it has found so far.

On the other hand, the global best component refers to the best position found by any particle in the swarm. In each iteration, particles adjust their velocity and position based on these two components, aiming to converge toward the optimal solution. The PSO algorithm uses a set of mathematical equations to update the velocity and position of each particle. These equations take into account factors such as the particle's current position, velocity, personal best position, and global best position. The pseudo-code for the PSO algorithm can be summarized as follows:[70]

1. Initialize a swarm of particles with random positions and velocities.
2. Evaluate the fitness of each particle's position.
3. Update the personal best position for each particle based on its fitness.
4. Update the global best position based on the personal best positions of all particles.
5. Update the velocity and position of each particle based on the personal best and global best positions.
6. Repeat steps 2-5 until a stopping criterion is met (e.g., a maximum number of iterations or convergence criteria).

One potential application of the PSO algorithm is calibrating finite element models. The PSO algorithm can be used to calibrate finite element models, which are mathematical representations of physical systems that allow engineers and scientists to simulate and analyze the behavior of these systems. By adjusting the parameters within the finite element model, such as material properties or boundary conditions, the PSO algorithm can optimize these parameters to match experimental data more accurately.[71] In the context of calibrating finite element models, the PSO algorithm can be used to find the optimal set of parameters that minimize the discrepancy between the model predictions and the actual data.

The experimental mode shapes set as the target for calibration are obtained through modal analysis techniques, such as vibration testing. The PSO algorithm can iteratively adjust the parameters of the finite element model, such as stiffness and damping coefficients, to minimize the difference between the predicted and measured mode shapes. The PSO algorithm achieves this by iteratively updating the velocity and position of each particle based on its personal best experience and the global best experience of all particles in order to find the optimal set of parameters that minimize the objective function, which measures the discrepancy between the model predictions and the target experimental data. Once the optimal set of parameters is found, the calibrated finite element model can be used for various purposes such as structural analysis and design optimization. The use of the PSO algorithm in calibrating finite element models offers several advantages. Firstly, the PSO algorithm is a powerful optimization technique that can handle complex and multidimensional parameter spaces. Secondly, the PSO algorithm is easy to implement and computationally efficient, making it suitable for solving large-scale calibration problems[71]. Thirdly, the PSO algorithm requires fewer parameter settings compared to other optimization algorithms, reducing the burden of fine-tuning and making it more user-friendly for engineers and researchers [72]. Furthermore, the PSO algorithm does not require an initial guess of the optimal parameter values, making it particularly useful in cases where prior knowledge about the system is limited. Overall, the PSO algorithm is a useful tool for calibrating finite element models.

PSO algorithm is defined as below :

Update the particle's velocity:

$$Vi^{(t+1)} = w * Vi^{t} + c1 * rand() * (PBesti - Xi^{t}) + c2 * rand() * (GBest - Xi^{t})$$

Where:

$Vi^{(t+1)}$ is the velocity of the particle i at the next time step, w is the inertia weight,

Vi^t is the current velocity of the particle i, c1 and c2 are cognitive and social scaling parameters, respectively, rand() is a random number between 0 and 1, PBest_i is the personal best position achieved by the particle i, X_i^(t) is the current position of the particle i, and GBest is the best position found by any particle in the swarm.

Update the particle's position:

$Xi^{(t+1)} = Xi^{(t)} + Vi^{(t+1)}$

Where:

$Xi^{(t+1)}$ is the new position of the particle i,

$Xi^{(t)}$ is the current position of the particle i, and

$Vi^{(t+1)}$ is the new velocity of the particle i.

The Particle Swarm Optimization algorithm is widely used as a parameter calibration method in the field of finite element modeling. Parameter alpha is defined as a factor to mass distribution and stiffness, and beta is a factor to determine the fitness of the model.

the final optimized parameter applied to the MATLAB code to obtain new mode shape for comparison.

The result of the PSO optimization shows that the calibrated finite element model accurately captures the desired behavior of the system. the MAC value of the system is in Figure 5.9, indicating a high level of modal assurance criteria between the measured and simulated mode shapes.

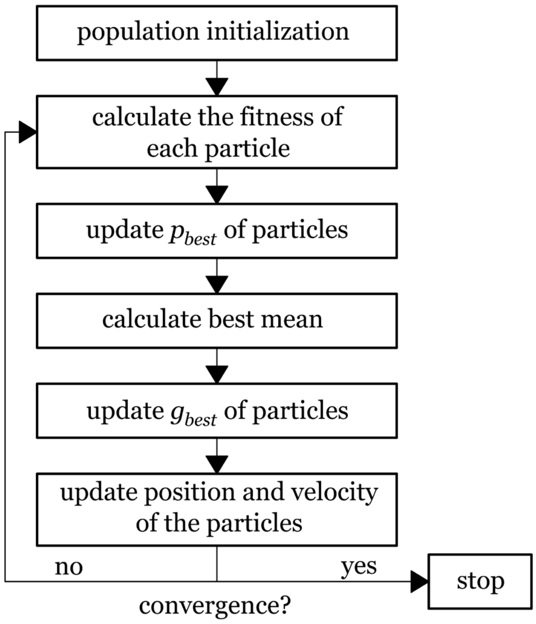


Figure 5.8: PSO algorithm flowchart

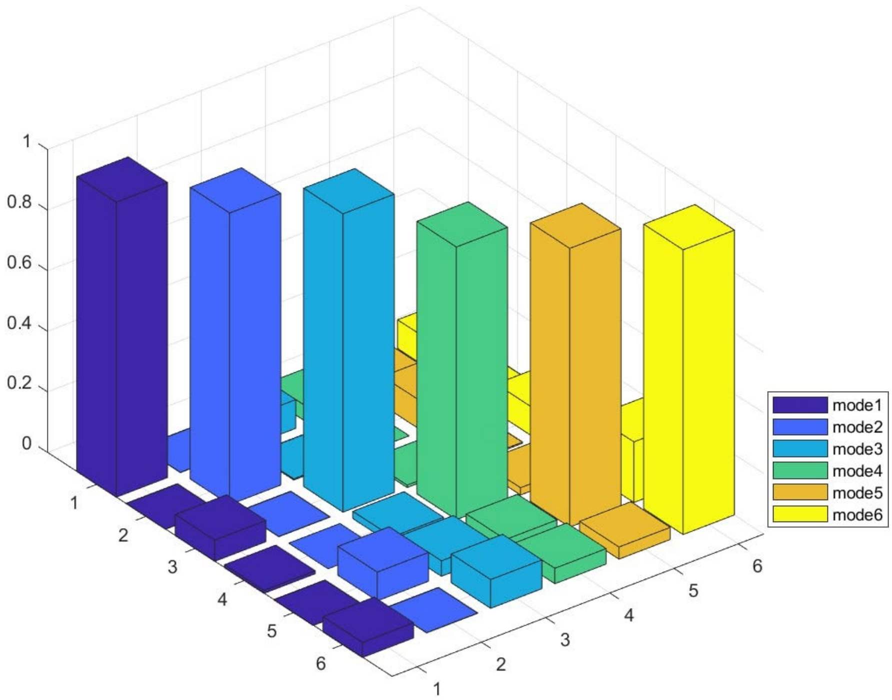


Figure 5.9: modal assurance criterion

5.6 FE model damage detection

by considering the two introduced damage indexes in the previous section, the damage detection on the north girder is performed. The damage on the bridge was introduced by reducing the stiffness of the element in the FE model. The more accurate result of the damage detection can be achieved by calibrated FE model.

The analysis of the damage detection using the calibrated finite element model demonstrates the effectiveness of the PSO algorithm in accurately identifying and localizing structural damage.

Damage introduce to the 12th element of the north girder base on the experimental data obtained. The damage index is calculated based on the modal strain energy introduce in previous section.

The damage index is shown in Figure 5.10

as it is visible the damage index increases significantly at the location of the introduced damage, indicating successful detection.

Overall, the particle swarm optimization algorithm has proven to be an effective method for calibrating finite element models and detecting structural damage. It has been widely utilized in various studies to update finite element models and accurately estimate damages in structures.

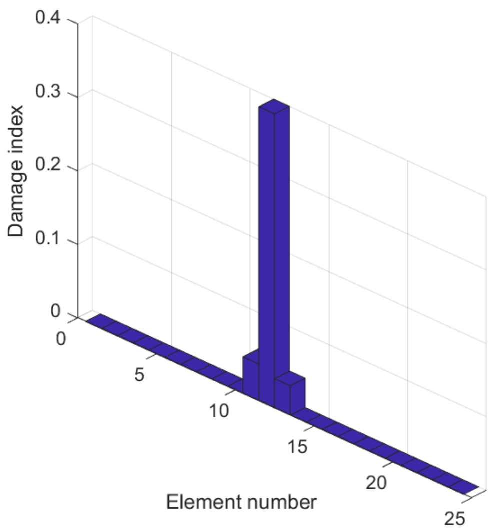


Figure 5.10: damage detection on calibrated FE model

Chapter 6

Conclusion

The overall objective of this research was to develop a vibration-based structural health monitoring system that incorporates machine learning algorithms for accurate damage detection and classification. Through the analysis of vibration data collected from structures, patterns and anomalies were identified that indicate the presence of damage. By utilizing machine learning techniques, the areas where the damage occurred could be pinpointed accurately. Furthermore, machine learning algorithms were able to assess the severity and extent of structural damage by analyzing vibration data. This research addressed several challenges in implementing vibration-based structural health monitoring with machine learning algorithms. Overall, this research has demonstrated the potential of machine learning algorithms in the field of vibration-based structural health monitoring for accurate damage detection and classification.

The research conducted on the Z24 and KW51 bridges served as validation for the effectiveness of the developed vibration-based structural health monitoring system. Time series analysis methods were also employed to analyze the measured response signals in the time domain. By utilizing machine learning algorithms, environmental variability changes in the sampled data were mitigated, allowing for more accurate damage detection and classification. Autoencoder methods were used to remove noise and environmental effects from the vibration data, while distance-based clustering optimized by the Gaussian mixture model was utilized for efficient analysis. Also, the research addressed the challenges associated with the hybrid method and FE model updating and proposed solutions to overcome these challenges.

The PSO algorithm was employed to optimize the parameters of the FE model. The MAC values were used to evaluate the accuracy of the calibrated FE model. Some factors are considered to optimize the mass distribution matrix and stiffness.

On the other hand, it is important to consider some opposing arguments to this research. One potential criticism is the reliance on machine learning algorithms for damage detection and classification. While these algorithms have shown great promise in various fields, including image recognition and natural language processing, there may be limitations when it comes to structural health monitoring. Machine learning algorithms require a large amount of data for training in order to accurately detect patterns and anomalies. However, obtaining sufficient high-quality vibration data from structures can be challenging. Factors such as limited access to structures, difficulty in collecting real-time data, and variations in environmental conditions can all affect the quality and quantity of data collected for training machine learning algorithms.

To overcome shortage of the data to train machine learning algorithms, researchers could explore the use of synthetic or simulated data to augment their training datasets. Additionally, it would be beneficial to conduct further research and investigation into the applicability of machine learning algorithms in sensor-optimized position.

Bibliography

[1] K. Al-Kodmany, "Tall Buildings, Design, and Technology: Visions for the Twenty-First Century City," Journal of Urban Technology, vol. 18, no. 3, pp. 115-140, 2011.

[2] H. Kim, Y. Kim, W. Chin and H. Hoon, "Development of Highly Efficient Construction Technologies for Super Long Span Bridge," Engineering, vol. 5, no. 8, pp. 629-636, 2013.

[3] J. Tan, K. Elbaz, Z. Wang, J. Shen and J. Chen, "Lessons Learnt from Bridge Collapse: A View of Sustainable Management," Sustainability, vol. 12, no. 3, 2020.

[4] V. Giurgiutiu, Structural Health Monitoring with Piezoelectric Wafer Active Sensors, Second Edition, 2014.

[5] K. Gkoumas, F. Marques dos Santos, M. Van Balen, A. Tsakalidis, A. Ortega, M. Grosso, A. Haq and F. Pekár, "Research and innovation in bridge maintenance, inspection and monitoring- Publications Office of the European Union, Luxembourg, 2019.

[6] P. Clemente, "Monitoring and evaluation of bridges: lessons from the Polcevera Viaduct collapse in Italy," Journal of Civil Structural Health Monitoring, 2020.

[7] J. Ko and Y. Ni, "Technology developments in structural health monitoring of large-scale bridges," Engineering Structures, vol. 27, p. 1715–1725, 2005.

[8] V. Karbhari and F. Ansari, Structural Health Monitoring of Civil Infrastructure Systems, Woodhead Publishing, 2009.

[9] P. Cawley, "Structural health monitoring: Closing the gap between research and industrial deployment," Structural Health Monitoring, vol. 17, no. 5, 2018.

[10] A. Rytter, "Vibration Based Inspection of Civil Engineering Structures," 1993.

[11] C. Farrar and K. Worden, Structural Health Monitoring - A Machine Learning Perspective, Wiley, 2013.

[12] A. Neves, I. Gonzalez, R. Karoumi and J. Leander, "Seasonal effects on novelty detection using ANNs for SHM," in 10th International Conference on Bridge Maintenance, Safety and Management (IABMAS 2020), Sapporo, Japan, 2020.

[13] A. Neves, I. González, R. Karoumi and J. Leander, "The influence of frequency content on the performance of ANN-based damage detection systems tested on numerical and experimental bridge data," Structural Health Monitoring, 2020.

[14] D. Reagan, A. Sabato and C. Niezrecki, "Feasibility of using digital image correlation for unmanned aerial vehicle structural health monitoring of bridges," Structural Health Monitoring, vol. 17, no. 5, 2018.

[15] S. Rana, S. Pichandi, R. Fangueiro and A. Correia, "A Review on Smart Self-Sensing Composite Materials for Civil Engineering Applications," AIMS Materials Science, vol. 3, no. 2, pp. 357-379, 2016.

[16] R. Lim, H. La and W. Sheng, "A Robotic Crack Inspection and Mapping System for Bridge Deck Maintenance," IEEE Transactions on Automation Science and Engineering, vol. 11, no. 2, pp. 367-378, 2014.

[17] M.-Q. Le, F. Ganet, D. Audigier, J.-F. Capsal and P.-J. Cottinet, "Printing of microstructure strain sensor for structural health monitoring," Applied Physics A, vol. 123, no. 5, 2017.

[18] Q. Mei and M. Gül, "A crowdsourcing-based methodology using smartphones for bridge health monitoring," Structural Health Monitoring, vol. 18, no. 5-6, 2018.

[19] C. R. Fava~, A. Duffey, P. A. Goldman, J. J. S. Vigil, and L. Alamos, “Uc-ooo.”

[20] V. Meruane and W. Heylen, “Structural damage assessment under varying temperature conditions,” Struct Health Monit, vol. 11, no. 3, pp. 345–357, May 2012, doi: 10.1177/1475921711419995.

[21] D. P. Rohe, “Experimental Modal Analysis using Phase Quantities from Phase-Based Motion Processing and Motion Magnification.”

[22] A. C. Neves, Structural Health Monitoring of Bridges : Data-based damage detection method using Machine Learning.

[23] V. M. Karbhari and L. S. W. Lee, “Vibration-based damage detection techniques for structural health monitoring of civil infrastructure systems,” Structural Health Monitoring of Civil Infrastructure Systems, pp. 177–212, 2009, doi: 10.1533/9781845696825.1.177.

[24] B. GÜNEŞ, “Unsupervised Learning Approach for Detection and Localization of Structural Damage using Output-only Measurements,” Academic Platform Journal of Engineering and Smart Systems, vol. 10, no. 3, pp. 149–156, Sep. 2022, doi: 10.21541/APJESS.1100238.

[25] G. Farias et al., “Automatic feature extraction in large fusion databases by using deep learning approach,” Fusion Engineering and Design, vol. 112, pp. 979–983, Nov. 2016, doi: 10.1016/J.FUSENGDES.2016.06.016.

[26] I. Mierswa, “Automatic Feature Extraction from Large Time Series,” Studies in Classification, Data Analysis, and Knowledge Organization, pp. 600–607, 2004, doi: 10.1007/3-540-28084-7_71.

[27] M. Paluszek and S. Thomas, “What Is Deep Learning?,” Practical MATLAB Deep Learning, pp. 1–24, 2020, doi: 10.1007/978-1-4842-5124-9_1.

[28] T. Zehelein, P. Werk, and M. Lienkamp, “An Evaluation of Autoencoder and Sparse Filter as Automated Feature Extraction Process for Automotive Damper Defect Diagnosis,” 2019 Fourteenth International Conference on Ecological Vehicles and Renewable Energies (EVER), May 2019, doi: 10.1109/EVER.2019.8813630.

[29] Y. An, E. Chatzi, S. H. Sim, S. Laflamme, B. Blachowski, and J. Ou, “Recent progress and future trends on damage identification methods for bridge structures,” Structural Control and Health Monitoring, vol. 26, no. 10. John Wiley and Sons Ltd, Oct. 01, 2019. doi: 10.1002/stc.2416.

[30] S. W. S. Doebling, C. R. C. Farrar, M. B. M. Prime, and D. W. D. Shevitz, “Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: A literature review,” Los Alamos National Laboratory, 1996, doi: 10.2172/249299.

[31] D. P. SOLOMATINE and R. K. PRICE, “INNOVATIVE APPROACHES TO FLOOD FORECASTING USING DATA DRIVEN AND HYBRID MODELLING,” Hydroinformatics, pp. 1639–1646, Jun. 2004, doi: 10.1142/9789812702838_0202.

[32] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, “Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks,” J Sound Vib, vol. 388, pp. 154–170, Feb. 2017, doi: 10.1016/J.JSV.2016.10.043.

[33] “Time Series Modeling for Structural Response Prediction | Semantic Scholar.” https://www.semanticscholar.org/paper/Time-Series-Modeling-for-Structural-Response-Batill-Hollkamp/ced9dd61ee349fdebf81b62405cb2dafef659f3a (accessed Jun. 26, 2023).

[34] S. Batill and J. Hollkamp, “Time Series Modeling for Structural Response Prediction,” 1988.

[35] S. Das, P. Saha, and S. K. Patro, “Vibration-based damage detection techniques used for health monitoring of structures: a review,” J Civ Struct Health Monit, vol. 6, no. 3, pp. 477–507, Jul. 2016, doi: 10.1007/S13349-016-0168-5.

[36] A. Fasana, L. Garibaldi, B. A. D. Piombo, and M. Ruzzene, “Identification of the dynamic behavior of bridges under ambient excitation using wavelet transform,” Nondestructive Evaluation of Bridges and Highways, vol. 2946, pp. 286–297, Nov. 1996, doi: 10.1117/12.259146.

[37] P. Chowdhury, P. Conrad, J. D. Bakos, and A. Downey, “Time Series Forecasting for Structures Subjected to Nonstationary Inputs,” Proceedings of ASME 2021 Conference on Smart Materials, Adaptive Structures and Intelligent Systems, SMASIS 2021, 2021, doi: 10.1115/SMASIS2021-68338.

[38] A. Teughels and G. De Roeck, “Structural damage identification of the highway bridge Z24 by FE model updating,” J Sound Vib, vol. 278, no. 3, pp. 589–610, Dec. 2004, doi: 10.1016/j.jsv.2003.10.041.

[39] M. G. Masciotta, L. F. Ramos, P. B. Lourenço, and M. Vasta, “Damage Detection on the Z24 Bridge by a Spectral-Based Dynamic Identification Technique,” Conference Proceedings of the Society for Experimental Mechanics Series, vol. 4, pp. 197–206, 2014, doi: 10.1007/978-3-319-04546-7_23.

[40] C. Krämer, C. D. Smet, and G. Roeck, “Z24 bridge damage detection tests,” 1999.

[41] E. Reynders and G. De Roeck, “Continuous Vibration Monitoring and Progressive Damage Testing on the Z24 Bridge,” Encyclopedia of Structural Health Monitoring, Jan. 2009, doi: 10.1002/9780470061626.SHM165.

[42] I. Iakovidis and K. Morfidis, “A finite element model updated by artificial neural networks to explain the behaviour of the Z24 Swiss bridge in different temperature states.,” IABSE Congress, Ghent 2021: Structural Engineering for Future Societal Needs, pp. 366–375, 2021, doi: 10.2749/GHENT.2021.0366.

[43] M. T. A. Chaudhary, “Modal analysis for damage identification: Past experiences and Swiss Z-24 Bridge,” Latin American Journal of Solids and Structures, vol. 17, no. 2, 2002, doi: 10.1590/1679-78255895.

[44] I. Iakovidis and V. Lekidis, “Z24 BRIDGE FEM UPDATE BASED ON TEMPERATURE DEPENDENT CONCRETE AND ASPHALT ELASTICITY MODULI,” 2018.

[45] C. Krämer, C. D. Smet, and G. Roeck, “Z24 bridge damage detection tests,” 1999.

[46] I. Iakovidis and V. Lekidis, “Z24 BRIDGE FEM UPDATE BASED ON TEMPERATURE DEPENDENT CONCRETE AND ASPHALT ELASTICITY MODULI,” 2018.

[47] A. Plotnikova, L. Wotherspoon, S. Beskhyroun, and Z. (Joey) Yang, “Influence of seasonal freezing on dynamic bridge characteristics using in-situ monitoring data,” Cold Reg Sci Technol, vol. 160, pp. 184–193, Apr. 2019, doi: 10.1016/J.COLDREGIONS.2019.02.006.

[48] M. A. Wahab and G. De Roeck, “EFFECT OF TEMPERATURE ON DYNAMIC SYSTEM PARAMETERS OF A HIGHWAY BRIDGE,” Structural Engineering International, vol. 7, no. 4, pp. 266–270, 1997, doi: 10.2749/101686697780494563.

[49] K. Khan, S. U. Rehman, K. Aziz, S. Fong, S. Sarasvady, and A. Vishwa, “DBSCAN: Past, present and future,” The Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2014), pp. 232–238, 2014, doi: 10.1109/ICADIWT.2014.6814687.

[50] I. Cordova and T. S. Moh, “DBSCAN on Resilient Distributed Datasets,” 2015 International Conference on High Performance Computing & Simulation (HPCS), pp. 531–540, Sep. 2015, doi: 10.1109/HPCSIM.2015.7237086.

[51] M. M. A. Patwary, D. Palsetia, A. Agrawal, W. K. Liao, F. Manne, and A. Choudhary, “A new scalable parallel DBSCAN algorithm using the disjoint-set data structure,” 2012 International Conference for High Performance Computing, Networking, Storage and Analysis, 2012, doi: 10.1109/SC.2012.9.

[52] J. Jang and H. Jiang, “DBSCAN++: Towards fast and scalable density clustering,” ArXiv, 2018.

[53] M. V. Akinin, N. V. Akinina, A. I. Taganov, and M. B. Nikiforov, “Autoencoder: Approach to the reduction of the dimension of the vector space with controlled loss of information,” Proceedings - 2015 4th Mediterranean Conference on Embedded Computing, MECO 2015 - Including ECyPS 2015, BioEMIS 2015, BioICT 2015, MECO-Student Challenge 2015, pp. 171–173, Aug. 2015, doi: 10.1109/MECO.2015.7181894.

[54] D. C. Cireşan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, “High-Performance Neural Networks for Visual Object Classification,” Feb. 2011, Accessed: Jun. 26, 2023. [Online]. Available: http://arxiv.org/abs/1102.0183

[55] Z. Hua, X. Li, Q. Qiu, and L. Zhao, “Autoencoder Network for Hyperspectral Unmixing With Adaptive Abundance Smoothing,” IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 9, pp. 1640–1644, Jul. 2021, doi: 10.1109/LGRS.2020.3005999.

[56] P. Baldi, “Autoencoders, Unsupervised Learning, and Deep Architectures,” 2011.

[57] K. Pawar and V. Z. Attar, “Assessment of Autoencoder Architectures for Data Representation,” Deep Learning: Concepts and Architectures, vol. 866, pp. 101–132, 2019, doi: 10.1007/978-3-030-31756-0_4.

[58] J. Zhai, S. Zhang, J. Chen, and Q. He, “Autoencoder and Its Various Variants,” 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 415–419, Jan. 2018, doi: 10.1109/SMC.2018.00080.

[59] S. Yue, X. Wang, and M. Wei, “Application of two-order difference to gap statistic,” Transactions of Tianjin University, vol. 14, no. 3, pp. 217–221, Jun. 2008, doi: 10.1007/S12209-008-0039-1.

[60] A. Entezami, H. Shariatmadar, and C. De Michele, “Non-parametric empirical machine learning for short-term and long-term structural health monitoring,” Struct Health Monit, vol. 21, no. 6, pp. 2700–2718, Nov. 2022, doi: 10.1177/14759217211069842.

[61] K. Maes and G. Lombaert, “Monitoring Railway Bridge KW51 Before, During, and After Retrofitting,” Journal of Bridge Engineering, vol. 26, no. 3, p. 04721001, Mar. 2021, doi: 10.1061/(ASCE)BE.1943-5592.0001668/ASSET/B7483EA6-31B5-4537-B96E-05ADD1533BFB/ASSETS/IMAGES/LARGE/BEENG-4725F38.JPG.

[62] K. Maes, L. Van Meerbeeck, E. P. B. Reynders, and G. Lombaert, "Validation of vibration-based structural health monitoring on retrofitted railway bridge KW51," Mech Syst Signal Process, vol. 165, p. 108380, Feb. 2022, doi: 10.1016/J.YMSSP.2021.108380.

[63] L. F. Fadel Miguel, R. Holdorf Lopez, and L. F. Fadel Miguel, "A hybrid approach for damage detection of structures under operational conditions," J Sound Vib, vol. 332, no. 18, pp. 4241–4260, Sep. 2013, doi: 10.1016/j.jsv.2013.03.017.

[64] M. Rezaiee-Pajand, H. Sarmadi, and A. Entezami, "A hybrid sensitivity function and Lanczos bidiagonalization-Tikhonov method for structural model updating: Application to a full-scale bridge structure," Appl Math Model, vol. 89, pp. 860–884, Jan. 2021, doi: 10.1016/j.apm.2020.07.044.

[65] C. R. Farrar, P. J. Cornwell, S. W. Doebling, and M. B. Prime, "Los Alamos Structural Health Monitoring Studies of the Alamosa Canyon and I-40 Bridges."

[66] C. R. Farrar and D. A. Jauregui, "Comparative study of damage identification algorithms applied to a bridge: II. Numerical study *," 1998.

[67] "Data Analyses from Seismic Instrumentation Installed on the I-40 Bridge Determination of When Crack Occurred on the Hernando de Soto I-40 Bridge." [Online]. Available: http://www.tn.gov/

[68] W. Heylen and T. Janter, "EXTENSIONS OF THE MODAL ASSURANCE CRITERION," J Vib Acoust, vol. 112, no. 4, pp. 468–472, 1990, doi: 10.1115/1.2930130.

[69] G. M. Nasira and S. Radhimeenakshi, "Particle Swarm Optimization- A Review on Algorithmic Rule, Application and Scope," 2014.

[70] S. Pant and S. B. Singh, "Particle swarm optimization to reliability optimization in complex system," 2011 IEEE International Conference on Quality and Reliability, pp. 211–215, 2011, doi: 10.1109/ICQR.2011.6031711.

[71] K. Parsopoulos and M. N. Vrahatis, "Particle Swarm Optimization Method for Constrained Optimization Problems," 2002.

[72] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization," Swarm Intelligence, vol. 1, pp. 33–57, 1995, doi: 10.1007/S11721-007-0002-0.

A. Appendix A

Experimental data of I-40 bridge.

I-40 Bridge Damaged and Undamaged Forced Vibration Data						
Test t16tr Undamaged Forced Vibration Global Polynomial Curve-Fit Results						
Location	Mode 1 F=2.48 Hz, x=1.06 %	Mode 2 F=2.96 Hz, x=1.29 %	Mode3 F=3.50 Hz, x=1.52 %	Mode 4 F=4.08 Hz, x=1.10 %	Mode 5 F=4.17 Hz, x=0.86 %	Mode 6 F=4.63 Hz, x=0.92 %
S1	M=257u P=184	173u 353	472u 12.9	702u 358	513u 176	414u 5.63
S2	M=6.90m P=174	5.55m 360	0.014 1.72	0.013 6.25	0.011 172	9.88m 2.90
S3	M=0.010 P=174	8.38m 1.39	0.019 1.99	0.017 6.99	0.015 173	0.013 2.99
S4	M=8.01m P=178	6.98m 4.00	0.013 3.08	994m 8.91	0.010 175	7.67m 4.04
S5	M=16.6u P=70.3	137u 165	716u 3.55	1.17m 2.70	609u 172	1.03m 1.59
S6	M=0.014 P=358	0.014 180	4.48m 172	7.72m 355	3.44m 10.8	7.49m 357
S7	M=0.023 P=3.33	0.024 183	1.04m 139	0.015 0.626	1.89m 58.4	0.014 0.09
S8	M=0.015 P=2.84	0.016 182	3.87m 5.74	9.44m 2.45	2.71m 158	8.43m 360
S9	M=456u P=353	535u 176	511u 164	1.06m 352	491u 13.8	1.08m 0.168
S10	M=8.62m P=183	8.87m 0.053	0.014 178	7.73m 350	0.012 4.50	8.61m 2.15
S11	M=0.010 P=177	0.011 356	0.020 174	0.013 349	0.017 2.41	0.014 360
S12	M=7.54m P=164	8.43m 359	0.016 177	0.011 352	0.013 5.210	0.012 1.76
S13	M=423u P=181	472u 9.33	895u 178	697u 355	918u 6.95	827u 5.18
N1	M=302u P=202	293u 150	478u 10.0	610u 351	735u 2.03	552u 193
N2	M=7.46m P=183	7.62n, 185	0.014 0.555	0.010 348	0.015 2.33	0.015 H.1
N3	M=0.010 P=183	0.011 186	0.018 0.954	0.012 347	0.019 2.57	0.018 184
N4	M=8.12m P=184	8.42m 184	0.011 360	6.58m 343	0.012 1.32	9.92m 184
NS	M=201u P=205	111u 136	652u 0.569	900u 349	869U 1.38	1.26m 181
N6	M=0.014 P=2.66	0.016 2.61	4.28m 179	7.71m 2.75	4.25m 177	9.68m 181
N7	M=0.021 P=4.04	0.024 2.32	464u 202	0.013 360	879u 143	0.017 182
N8	M=0.014 P=3.34	0.017 2.28	3.39m 0.195	6.85m 355	3.38m 5.00	0.010 182
N9	M=367u P=325	783u 347	427u 160	942u 1.73	605u 176	1.50m 179
N10	M=7.52m P=185	9.97m 182	0.012 180	0.010 10.4	0.013 182	0.011 186
N11	M=0.010 P=179	0.014 178	0.019 177	0.017 6.54	0.020 179	0.019 183
N12	M=6.99m P=178	9.54m 178	0.015 176	0.014 5.73	0.016 178	0.015 183
N13	M=434u P=188	561u 183	981u 170	1.04m 5.53	1.18m 179	1.23m 185
M is the magnitude P is the phase in degree m= 10e-3 u=10e-6						

B. Appendix B

GMM optimization algorithm

```python
import numpy as np

from sklearn.mixture import GaussianMixture
from sklearn.model_selection import GridSearchCV



# Define the range of n_components to try
n_components_range = range(1, 50)

# Method 1: BIC
bic_scores = []
for n_components in n_components_range:
    gmm = GaussianMixture(n_components=n_components)
    gmm.fit(X)
    bic_scores.append(gmm.bic(X))

# Method 2: MDL
mdl_scores = []
for n_components in n_components_range:
    gmm = GaussianMixture(n_components=n_components)
    gmm.fit(X)
    mdl_scores.append(gmm.score(X))

# Method 3: Cross-validation
param_grid = {"n_components": n_components_range}
gmm_cv = GridSearchCV(GaussianMixture(), param_grid, cv=5)
gmm_cv.fit(X)
cv_scores = gmm_cv.cv_results_["mean_test_score"]

# Method 4: AIC
aic_scores = []
for n_components in n_components_range:
    gmm = GaussianMixture(n_components=n_components)
    gmm.fit(X)
    aic_scores.append(gmm.aic(X))

# Find the optimal number of components based on different methods
optimal_bic = np.argmin(bic_scores) + 1
optimal_mdl = np.argmax(mdl_scores) + 1
```

```
optimal_cv = gmm_cv.best_params_["n_components"]
optimal_aic = np.argmin(aic_scores) + 1

print(f"Optimal number of components (BIC): {optimal_bic}")
print(f"Optimal number of components (MDL): {optimal_mdl}")
print(f"Optimal number of components (CV): {optimal_cv}")
print(f"Optimal number of components (AIC): {optimal_aic}")
```

Autoencoder Algorithm

```
from keras.layers import Input, Dense
from keras.models import Model
from sklearn.model_selection import train_test_split

# Split the data into train and validation sets
X_train, X_val = train_test_split(X, test_size=0.2, random_state=42)

# Define the input shape
input_shape = (X_train.shape[1],)
# Define the number of hidden layers and units
encoding_dim = 2
hidden_dim = 10

# Define the input layer
input_layer = Input(shape=input_shape)

# Define the encoder layers
encoded = Dense(hidden_dim, activation='relu')(input_layer)
encoded = Dense(encoding_dim, activation='relu')(encoded)

# Define the decoder layers
decoded = Dense(hidden_dim, activation='relu')(encoded)
decoded = Dense(input_shape[0], activation='linear')(decoded)
# Define the Autoencoder model
autoencoder = Model(inputs=input_layer, outputs=decoded)

# Compile the model
autoencoder.compile(optimizer='adam', loss='mse')

# Fit the model on the train set
autoencoder.fit(X_train, X_train, epochs=100, batch_size=16, verbose=0)

# Predict the reconstruction errors of the validation set
y_val_pred = autoencoder.predict(X_val)
```

```
reconstruction_errors_val = np.mean(np.power(X_val - y_val_pred, 2), axis=1)

# Find the optimal threshold value on the validation set
#percentile = 95
# Define the range of percentile values to test

percentiles = range(90, 100)

# Initialize variables to store the best F1 score and its corresponding
percentile value
best_f1 = 0
best_percentile = 0

# Iterate over the percentile values
for percentile in percentiles:
    # Find the threshold value on the validation set
    threshold = np.percentile(reconstruction_errors_val, percentile)
    # Detect the outliers in the whole dataset
    outliers = [i for i in range(len(X)) if reconstruction_errors[i] >
threshold]

    # Compute the precision and recall of the outlier detection
    true_positives = len(outliers)
    false_positives = len(X) - len(outliers)
    false_negatives = 0
    precision = true_positives / (true_positives + false_positives)
    recall = 1.0

    # Compute the F1 score
    f1 = 2 * precision * recall / (precision + recall)

    # Update the best F1 score and its corresponding percentile value
    if f1 > best_f1:
        best_f1 = f1
        best_percentile = percentile

# Print the most proper percentile value
print(f'Most proper percentile value: {best_percentile}, F1 score: {best_f1}')

# Predict the reconstruction errors of the whole dataset
y_pred = autoencoder.predict(X)
reconstruction_errors = np.mean(np.power(X - y_pred, 2), axis=1)

# Store the indices of the outlier data points
indices = [i for i in range(len(X)) if reconstruction_errors[i] > threshold]
```

GEV distribution algorithm

```
import numpy as np
from scipy.stats import genextreme
# Fit the GEV distribution to the input vector
fit_GEV = genextreme.fit(Damage_index)
# Extract the parameters of the GEV distribution
mu, sigma, k = fit_GEV
# Significance level (α)
alpha = 0.05
# Calculate the alarming threshold
x_alpha = (mu - sigma / k) * (1 - (-np.log(1 - alpha)) ** (-k))
```

List of Figures

List of Tables