A Methodology for Damage Detection Using Unsupervised Learning in the Field of Structural Health Monitoring

Based on Gaussian Mixture Modeling

Þorvaldur Kári Vilhjálmsson

Delft University of Technology

TU Delft

A Methodology for Damage Detection Using Unsupervised Learning in the Field of Structural Health Monitoring

Based on Gaussian Mixture Modeling

by

Þorvaldur Kári Vilhjálmsson

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Monday June 26, 2023 at 10:30 AM.

Student Name	Student Number
Þ.K. Vilhjálmsson	5360455

Chair:	Dr. E. Lourens
Daily Supervisor:	Dr. H. Wang
Committee Member:	Dr. K. van Dalen
Faculty:	Faculty of Civil Engineering, Delft

Cover: The 1915 Çanakkale Bridge. Image taken by the author.

TU Delft

Preface

With the conclusion of my thesis, my journey at the Technical University of Delft comes to an end. My time here has had its incredible ups and downs. I moved to Delft during the middle of a pandemic and dealt with the loss of three of my grandparents and my beautiful dog during my stay. This proved to be an extremely difficult time for me but despite all the hardships, I am happy and I am proud.

This thesis is dedicated to my family, and especially to those that I have lost.

My dear grandmother Karen, who was the sweetest and most kind soul. She always wanted me to become a doctor.

My dear grandmother Sidda, a true role model that always gave me a reason to smile.

My dear grandfather Halldór, who always took the best care of his loved ones and showed true grit throughout his life. He always wanted me to move back to Iceland during the middle of my studies.

My dear golden-retriever Týra, the best dog a family could ask for. She unlocked a side of me that taught me unconditional love and loyalty, and showed me the joy and happiness that can be found in even the simplest things in life.

All of my family and grandparents gave me incredible support and strength throughout my time here. I am forever grateful.

I would also like to express my sincere gratitude to my graduation committee, for your time, effort, and support during my thesis project. Your guidance, advice, and feedback have been invaluable to me throughout this journey.

I appreciate the commitment you have shown to my work, providing critical feedback, and offering constructive suggestions. Your contributions have helped me to refine my ideas and present my research in the best possible light.

I feel truly blessed to have had the opportunity to work with such an accomplished and dedicated committee.

Þorvaldur Kári Vilhjálmsson
Delft, June 2023

"Whichever wine was within, it was decidedly not identical to its neighbors. On the contrary, the contents of the bottle in his hand was the product of a history as unique and complex as that of a nation, or a man. In its color, aroma, and taste, it would certainly express the idiosyncratic geology and prevailing climate of its home terrain. But in addition, it would express all the natural phenomena of its vintage. In a sip, it would evoke the timing of that winter's thaw, the extent of that summer's rain, the prevailing winds, and the frequency of clouds. Yes, a bottle of wine was the ultimate distillation of time and place; a poetic expression of individuality itself."

— Amor Towles, *A Gentleman in Moscow*

Abstract

This thesis aims to investigate the feasibility of developing a successful unsupervised Structural Health Monitoring (SHM) methodology to detect damage in structures, specifically bridges. Detecting damage, especially in its earliest stages, is challenging, thus prompting the need for robust and effective methods. The success of such a methodology could lead to timely inspections and interventions, resulting in significant economic benefits and preventing further damage, including potential failure of the structure in use.

The approach involves a literature review to establish relevant background knowledge and useful concepts. From this, a methodology is developed utilizing unsupervised machine learning, specifically Gaussian Mixture Models (GMM), to identify abnormal behavior indicative of structural damage.

A Finite Element Method (FEM) model of a simple bridge is created and monitored over a three-year period, serving as a testing ground for the methodology and a primary source for data generation. Temperature data and its effects on the natural frequencies of the bridge model are used to establish a baseline for normal or healthy behavior. Synthetic damage, such as settlement and stiffness reduction, is then introduced to the model to create anomalies or abnormal behavior. The developed methodology is tested using three case studies, each with varying types of synthetic damage. By using both the healthy and unhealthy data generated from the model, the healthy behavior of the bridge is captured using GMM. The model then progressively incorporates unhealthy data into the proposed anomaly detection algorithm. The algorithm evaluates the likelihood of each incoming data point of belonging within the healthy distribution, resulting in data points being classified as either healthy or flagged as abnormal.

The case studies presented in this research underscore the efficacy of the proposed anomaly detection approach. In scenarios involving sudden or abrupt damage, the algorithm swiftly and accurately labels abnormal points. For gradual damage scenarios, such as settlement, the algorithm consistently identifies abnormal points, with the rate of abnormal point detection accelerating over time. This detection rate is contrasted with the rate of erroneous abnormal point labeling when processing an exclusively healthy data set through the anomaly detection algorithm. This comparison reveals a higher rate of abnormal point identification when actual damage is present, affirming the effectiveness of the unsupervised SHM methodology in pinpointing abnormal behavior within the modeled bridge structure.

Contents

List of Figures

List of Tables

List of Algorithms

Nomenclature

Abbreviations

Abbreviation	Definition
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
FEM	Finite Element Method
GMM	Gaussian Mixture Model
HDBSCAN	Hierarchical Density-Based Spatial Clustering of Applications with Noise
kPCA	Kernel Principal Component Analysis
LDA	Linear Discriminant Analysis
OPTICS	Ordering Points To Identify Applications with Noise
PCA	Principal Component Analysis
PC	Principal Component
PLS	Partial Least Squares
SHM	Structural Health Monitoring
SVD	Singular Value Decomposition

1

Introduction

Civil infrastructure, comprising buildings, bridges, dams, roads, highways, and many other structures, constitutes an indispensable element of modern society. Its crucial role lies not only in facilitating economic growth and societal development but also in providing a safe and secure environment for human habitation. Given the perpetual presence of external forces, particularly the environment, civil infrastructure plays a vital role in safeguarding societies from the devastating impacts of natural disasters and other hazards. Civil infrastructure is therefore subject to continual wear and tear, with its lifespan often shortened by factors such as exposure to the elements and daily usage. As a result, regular inspections, maintenance, and repairs are critical to ensuring the safety and longevity of these structures. Failure to do so can result in the escalation of seemingly minor damages into significant safety concerns or even catastrophic events, posing a risk to human lives and society as a whole.

In order to combat the potential degradation of civil infrastructure due to external forces and aging, Structural Health Monitoring (SHM) has emerged as a valuable tool to better understand the behavior of monitored structures. By continuously monitoring the structural condition, SHM enables engineers and maintenance personnel to detect early signs of deterioration, allowing for prompt and targeted repairs and maintenance. This not only helps to prevent small damages from escalating into larger safety risks or catastrophic failures but also ensures that the infrastructure remains safe and reliable for continued use. Elaborating on SHM, it can be said that it is the process of identifying damage in aerospace, civil and mechanical engineering infrastructure. This process involves the regular observation of a structure or mechanical system over time using measurements, extracting damage-sensitive features from these measurements, and using statistical analysis to determine the current state of the system's health. The output of this process is updated information on the structure's ability to continue performing its intended function over time, considering the aging and damage accumulation that occurs in operational environments [1].

One major limitation of SHM is the frequent need to rely on data from damaged conditions, which can pose significant challenges and costs. There are two possible approaches for acquiring such data: through physics-based modeling or experimental testing. However, accurately modeling complex structural geometries or material compositions can be difficult, while experiments can be expensive and limited by the physical size of scaled specimens or components. Additionally, the effectiveness of damage detection may be limited to the specific types of damage considered, leaving novel, unconsidered types of damage undetected [2].

In light of these challenges, there is increasing interest in applying unsupervised learning to SHM tasks. Unsupervised learning is a type of machine learning where algorithms discern patterns and structures in a data set without the need for labeled or predetermined target outcomes. Rather than focusing on specific outputs, unsupervised learning algorithms aim to discover inherent patterns, groupings, or relationships within the data itself. Techniques commonly used in unsupervised learning encompass clustering, dimensionality reduction, and anomaly detection.

These approaches are particularly beneficial in SHM as they can eliminate the need for damaged condition data. By dispensing with this requirement, unsupervised learning provides a notable advantage, as obtaining damaged condition data can be extremely challenging and hard to come by.

This thesis, therefore, seeks to investigate the potential of unsupervised learning in the field of SHM, providing valuable insights and underscoring its possible advantages.

1.1. Problem Statement

The goal of this thesis is to develop a framework or methodology that employs unsupervised learning techniques within the field of SHM. This framework aims to detect damages at the earliest possible stage. The success of such a methodology could lead to timely inspections and interventions, resulting in significant economic benefits and preventing further damage, including potential failure of the structure in use. However, several challenges are associated with the implementation of unsupervised learning in a civil engineering context, particularly in SHM.

One of the primary challenges is accounting for the operational and environmental variability that can significantly influence the behavior and performance of civil engineering structures. Factors such as temperature fluctuations, humidity, wind loads, and other environmental conditions can cause changes in the structural response that may be unrelated to damage. Similarly, variations in operational loads, such as traffic patterns or construction activities, can also impact the structure's performance. Distinguishing between these influences and actual damage-induced changes presents a significant challenge for unsupervised learning algorithms. In this thesis, one of the goals is to address one of these environmental variability factors in such a way that it becomes a valuable asset instead of a limitation, thereby contributing to the effectiveness of the proposed framework.

1.2. Research Question

After having discussed motivations for using unsupervised learning in the context of SHM and providing an overview of the scope of this thesis, the research question can be formulated as follows:

> *"How to develop a methodology using unsupervised learning techniques that can effectively detect early-stage damages in structures under environmental variability?"*

Detecting damage, especially in its earliest stages, is challenging. Therefore, the study will determine whether an unsupervised learning method can effectively identify abnormal behavior. To aid in answering the above research question, a research methodology will be created and followed.

1.3. Research Methodology & Thesis Structure

After introducing SHM, unsupervised learning, and defining the research question, this section outlines the methodology for conducting the research. The subsequent chapter, the Literature Review, will explore critical areas pertinent to the proposed methodology. These areas encompass temperature effects on bridges, damage detection in SHM, machine learning, unsupervised learning, and clustering, all of which provide the necessary background for developing the unsupervised anomaly detection methodology.

Using the information learned by the literature review, an unsupervised learning methodology for anomaly detection will be formulated and applied. This application process will entail the creation of a FEM model and the execution of three distinct case studies, each concentrating on a different damage scenario - settlement, stiffness reduction, and a combination of the two. The FEM model will supply the necessary data for conducting the case studies. The anomaly detection results from these case studies will be presented, analyzed, and discussed, ultimately leading to a conclusion regarding the effectiveness of the proposed anomaly detection method being drawn. In addition to this recommendations for future work will be addressed and highlighted.

2

Literature Review

This chapter delves into the background knowledge relevant to the objective of developing an unsupervised technique for detecting structural damage. Specifically, the methodology will be applied to a FEM of a bridge. The impact of temperature on concrete and concrete bridges will be examined to establish a correlation between temperature and natural frequencies in bridge structures. Additionally, a brief section will be dedicated to various types of structural damage that will be simulated during the application of the developed methodology. Furthermore, the concept of SHM is introduced, which is closely related to damage detection. This section offers a concise overview of SHM and presents common approaches utilized in the field, including both supervised and unsupervised methods. The rationale for choosing an unsupervised learning approach is highlighted. To better comprehend the difference between supervised and unsupervised learning, the topic of machine learning is introduced. This section also introduces several key unsupervised methods.

2.1. Temperature Effects

The daily variations in temperature can significantly affect the behavior and performance of bridges, and it is essential to consider these fluctuations when monitoring and computing the natural frequencies of these structures. Concrete bridges are subject to thermal expansion and contraction due to temperature changes, leading to variations in the stiffness and damping of the structure. These changes can have a significant impact on the natural frequencies of the bridge, which is a critical parameter in its dynamic behavior and response to loads. Additionally, variations in temperature can alter the elastic modulus and shear modulus of the material, which are key factors in determining a bridge's vibration behavior. Changes in temperature, whether daily or seasonal, can also impact the stiffness of expansion joints or supports, thereby altering the constraints and causing shifts in the vibration characteristics [3]. Furthermore, fluctuations in temperature create substantial thermal stresses or stress redistribution within a structure, potentially affecting the vibration characteristics of bridges [4]. Studies have shown that temperature is a major contributor to modal variability, with changes in modal frequencies due to temperature reaching 5-10% for highway bridges. This is often greater than the frequency changes resulting from structural damage or degradation [5]. Keeping this in mind, with change in the structural dynamic characteristics, the accuracy of damage identification using vibration-based methods is greatly affected [6]. By considering daily variations in temperature, engineers can gain a more accurate understanding of the natural frequency of the bridge and its dynamic behavior, allowing for more effective monitoring and maintenance.

The impact of temperature variations on the vibration behavior of long-span bridges is complex. Typically, temperature distribution and changes within a bridge are non-uniform and vary over time, causing asynchronous changes in physical parameters among different structural elements [6]. Figure 2.1 perfectly displays the complicated manner of temperature in bridges. It can be seen how countless parameters play a role in the actual temperature inside the concrete. Some of these parameters can be listed as:

- Shape of cross-section
- Solar radiation & inclination
- Wind speed
- Air temperature
- Thermal irradiation
- Convection

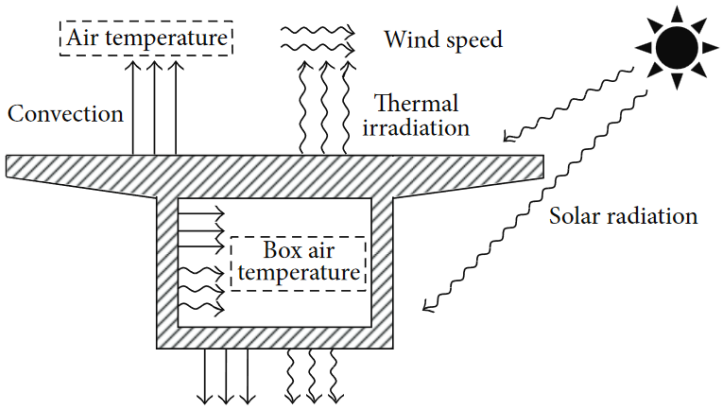


Figure 2.1: Typical cross-section of a concrete bridge showing the heat exchange between the boundary and the environment [6].

This thesis concentrates on creating a methodology for real-time anomaly detection using unsupervised learning, so a simplified approach will be employed regarding the consideration of temperature effects in the constructed FEM model. This approach will take into account the influence of temperature on both the elastic modulus and the size of the structure, i.e. the structural members in the model.

According to [7], the impact of temperature on the elastic modulus of concrete is much greater than on a reinforced elastic modulus. When considering a simply supported beam, concrete is the primary material, so the effect of temperature on the reinforcing bar is disregarded. The calculation of the effect of temperature on the elastic modulus of the simply supported beam is performed using the formula provided in the European Concrete Specification CEB-FIP Model Code 2010 [8]:

$$E_T = E_{20°\mathrm{C}} \left[1 - \vartheta_E (T - 20)\right] \tag{2.1}$$

where,

E_T : Elastic modulus [MPa] of concrete at temperature T,
T : Temperature of concrete [°C],
$E_{20°\mathrm{C}}$: Elastic modulus [MPa] of concrete at 20°C,
ϑ_E : Temperature coefficient of elastic modulus, with a value of 0.003.

As it relates to the structure size, [7] asserts that the effect of temperature on the dimensions of a simply supported beam primarily manifests as a change in the beam's length and the modification of the section's moment of inertia, which is related to the material properties. For a rectangular equal cross section of a simply supported beam, the change in beam length can be calculated using the following formula:

$$l_1 = (1 + \alpha \Delta T)\, l \tag{2.2}$$

where,

l_1 : Beam length after temperature change,
α : Linear expansion coefficient with reinforcement of 1.2×10^{-6} and concrete of 1.0×10^{-5},
ΔT : Change in temperature,
l : Initial beam length.

The change in beam width b and beam height h can be obtained by:

$$b_1 = (1 + \nu \alpha \Delta T)\, b \tag{2.3}$$

$$h_1 = (1 + \nu \alpha \Delta T)\, h \tag{2.4}$$

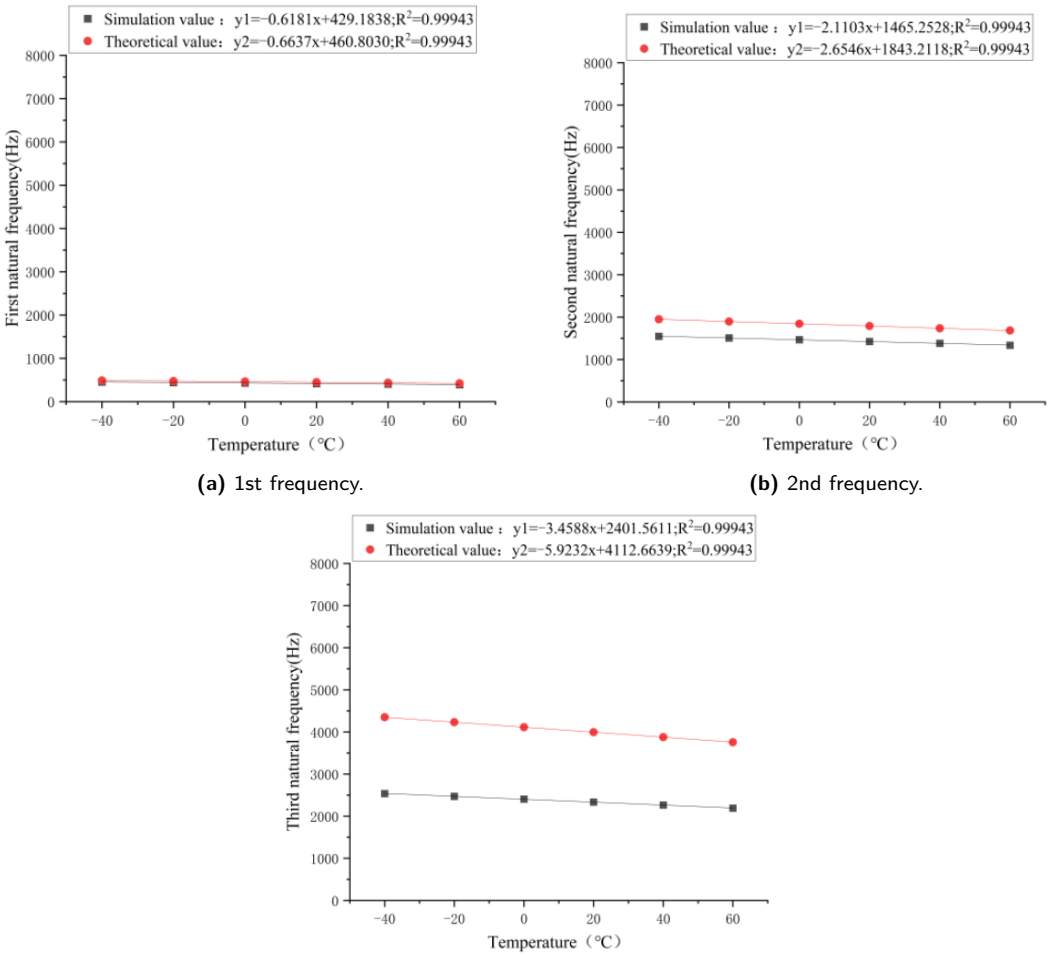
where,

ν : Poisson's ratio (concrete $\nu = 0.2$ and reinforced concrete $\nu = 0.31$).

At last, by using Equations 2.3 and 2.4 when computing the moment of inertia of the section, one obtains the variation of the moment of inertia as,

$$I = \frac{(1 + \nu \alpha \Delta T)^4 b h^3}{12} \tag{2.5}$$

In the study conducted by [7], it was found that for every 1°C rise in temperature from -40°C to 60°C, the natural frequency of their simply supported concrete beam decreased by 0.148%. These results were found after having used an acceleration sensor and dropping a steel ball onto the testing beam from three different landing points in order to obtain three waveforms. The testing of the beams in this experiment was done at -40°C, -20°C, 0 °C, 20°C, 40°C, and 60°C.



(a) 1st frequency.

(b) 2nd frequency.

(c) 3rd frequency.

Figure 2.2: Results from [7]. Relationship between temperature and the first three frequencies.

From Figure 2.2, a clear linear relationship between the temperature and the natural frequencies can be seen in both the experimentally obtained values and the theoretically obtained values. However, as is stated by [7], the derived theoretical natural frequencies of the simply supported beam was only applicable to the 1^{st} frequency, not higher-order frequencies. It is therefore safe to conclude that potential future research in relation to this thesis can certainly include a more complex approach when including temperature effects. This would mean taking into account the impact of uneven temperature distribution on the natural vibration characteristics.

Having established a straightforward linear relationship between temperature and natural frequency, it is noteworthy to observe a frequent phenomenon in bridge monitoring. An example of this phenomenon appeared after continuous monitoring of the Z24 bridge, a classical post-tensioned concrete box girder bridge located in Switzerland. The goal of this monitoring system, operating from November 11, 1997 to September 11, 1998, was to gather both environmental and vibration information. With this monitoring conducted by [9], a bilinear behavior was observed between the frequency and the temperature. This behavior is seen in Figure 2.3, where a sharp increase in the frequency occurs as the measured temperature goes below 0°C.

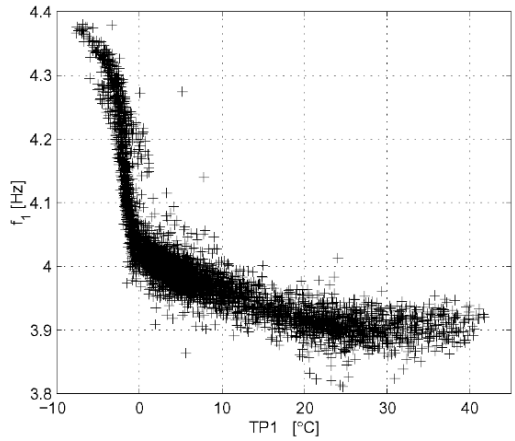


Figure 2.3: Bilinear behavior observed around the freezing point. Relationship between temperature and the natural frequency [9].

In [10], it is demonstrated that the bilinear behavior is due to the asphalt layer. While during warm periods, the asphalt does not have any impact, it significantly contributes to the bridge's stiffness during cold periods. [11] confirms that the stiffness of asphalt has been found to significantly change near the freezing point. To further reinforce the occurrence of this phenomena, the monitoring by [12] of the Dowling Hall Footbridge located on the campus of Tufts University, shows similar results. Figure 2.4 shows the first six natural frequencies versus the western abutment temperature T_{C1}. For all six modes, a knee-like observation can be made around the freezing point. This observation is also attributed to the asphalt layer on the bridge.

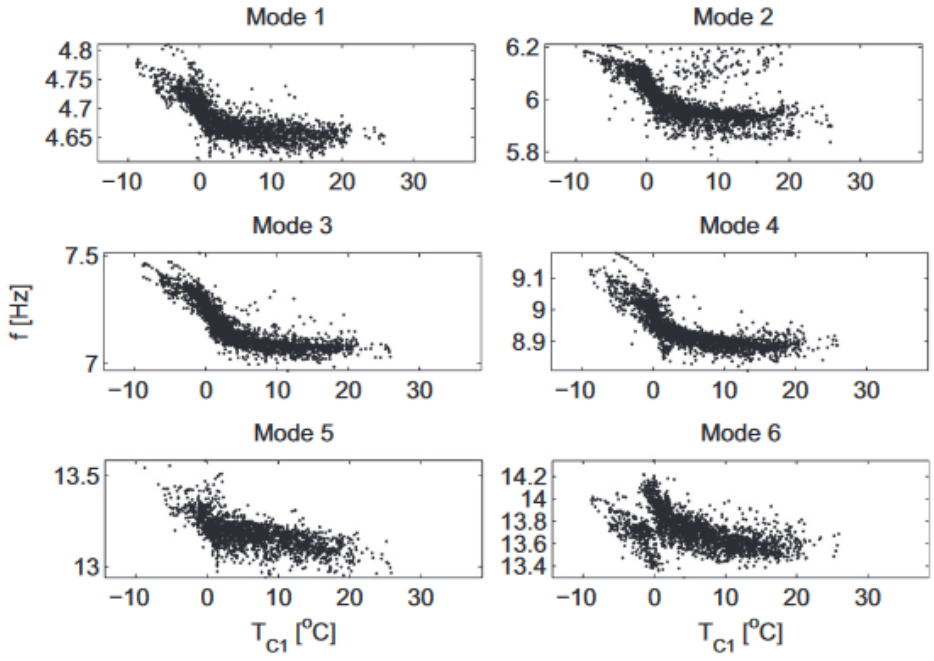


Figure 2.4: Bilinear behavior observed around the freezing point. Relationship between temperature and the natural frequency [12].

While studying the seasonal effects on the stiffness properties of a ballasted railway bridge in Sweden, [13] observes the same bilinear relationship, as can be seen in Figure 2.5. The figure aims to illustrate the shift from warm to cold periods and from cold to warm periods.

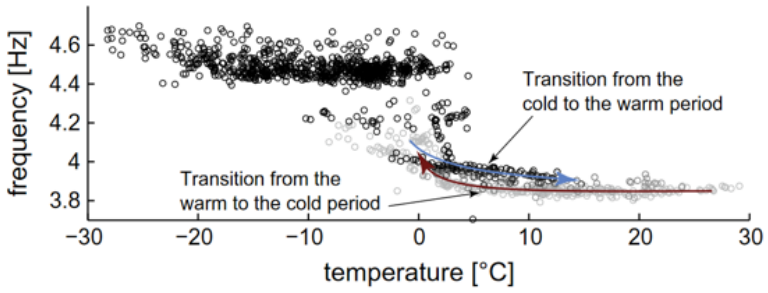


Figure 2.5: Bilinear behavior observed around the freezing point. Relationship between temperature and the natural frequency [13].

In conclusion, incorporating temperature effects into a finite element model presents a significant challenge, but it is a worthwhile endeavor. As previously stated, it is advisable to keep things simple and adopt the Equations 2.1-2.5 for the constructed FEM model used in this thesis. However, when working with real data, the use of synthetic data is no longer a concern, as it is not employed in that situation. Finally, the utilization of temperature information is crucial for the current project. By taking into account temperature effects, a correlation between temperature and natural frequency can be established, thereby providing a benchmark for what is considered normal behavior in the modeled structure.

2.2. Structural Damage in Bridges

Most engineering structures are subjected to loads which can lead to damages or cracks in areas where excessive stress occurs. Cracks in structural members, such as beams, can result in changes in stiffness that primarily depend on the crack's location and depth. These cracks can alter the physical properties of a structure and affect its dynamic response behavior [14]. Studies, like the one carried out by [14], demonstrate that cracks in beams result in a decrease in natural frequencies, particularly when the crack is located where the bending moment is at its peak. In a study conducted by [15], it was shown that relatively simple beam models with a limited number of degrees of freedom can effectively simulate the effects of an open crack. Despite the complex geometry of cracks, [15] argues that for low frequency vibrations, only a significant reduction in stiffness is necessary to accurately represent its effects. As shown in the study, more complicating approaches can be adopted.

In the created bridge model for this project, structural damage is purposefully incorporated to the model as a way to give rise to the detection of anomalies. Three different scenarios are implemented into the model:

- Settlement of bridge pier
- Reduction of stiffness in the bridge deck
- Combination of settlement and reduction of stiffness

Regarding the incorporation of damage in the FEM model developed in this thesis, a simple approach is favored. To simulate the settlement of one of the bridge piers, one of the supports will be linearly displaced vertically over the course of a year, and new natural frequencies will be continuously calculated to detect any abnormal behavior. To simulate cracking, a straightforward method is adopted that involves reducing the stiffness of an element at the point where the bending moment is highest. Additionally, a combination of these two damage scenarios will also be explored.

2.3. Damage Detection in SHM

Structural health monitoring (SHM) is the process of identifying damage in civil engineering infrastructure. It involves observing a structure or mechanical system over time with periodically spaced measurements, extracting damage-sensitive features from these measurements, and statistically analyzing these features to determine the current state of the system's health. With this in mind and the aim to identify damage in infrastructure, damage has been defined as changes that negatively affect a system's current or future performance, including changes to material and geometric properties, boundary conditions, and system connectivity. However, the term "damage" does not imply a total loss of system functionality, but rather that the system is no longer operating optimally. As damage grows, it reaches a point where it negatively affects the system's operation to an unacceptable level, which is referred to as failure. Damage can accumulate gradually over long periods or occur suddenly due to scheduled or unscheduled events [1]. As structures inevitably degrade over time, it is of paramount importance to utilize SHM methods to make sure that they are up to all necessary safety and service standards to keep its users safe from catastrophic failures and events.

Within SHM lies vibration-based methods. Vibration-based techniques can be broadly classified into two categories: model-based and data-driven. In the case of model-based SHM methods, they are generally implemented using system identification and model updating approaches [16]. Methods for model updating involve using physics-based models of structures, which are calibrated using real measurements obtained from the monitored structure. By detecting, localizing, and quantifying deviations in the updated system parameters, these methods can identify damage. Updating the parameters can be achieved through various approaches, such as sensitivity-based updating [17] and iterative optimization updating [18, 19]. Bayesian model updating methods are another option, where system parameters are represented as probability distributions based on information gathered from measurements of the structure [2, 20]. Using model-based SHM for long-term monitoring of structures can be expensive. Finite element model updating for large-scale structures can be computationally demanding, which could make it challenging for real-time SHM applications [21]. In addition, model updating methods are prone to errors that arise from modeling, including complex joint behavior, complex materials, and boundary conditions [22].

Data-driven SHM techniques utilize statistical learning methods that extract damage-sensitive features from structural measurement data, often from acceleration time series. The statistical learning

approach can be supervised or unsupervised. In supervised learning, both undamaged and damaged condition data are used for training, which can provide information about the location, type, and extent of damage. However, obtaining damaged condition data can be challenging and costly. Unsupervised learning is a more practical option for SHM, as it relies on novelty detection methods that are trained on data extracted from the structure under normal conditions only. This eliminates the need for damaged condition data and makes it more feasible to acquire training data from field measurements [2, 23].

Unsupervised learning methods offer several advantages over traditional approaches, including lower costs and less complexity. In the next section, machine learning will be discussed in greater detail, including supervised and unsupervised learning. Concepts such as principal component analysis and clustering, which are examples of unsupervised learning, will also be introduced and explored.

2.4. Machine Learning

Machine learning is a subfield of artificial intelligence that deals with the development of algorithms and statistical models that allow computers to perform tasks without explicit instructions. Several definitions arise when the topic of machine learning is brought up. One definition, given by [24], is when "a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E". In other terms, machine learning involves training systems on large data sets, enabling them to make predictions or decisions based on the patterns they have learned from the data. The versatility of machine learning allows it to be utilized in various fields, including image recognition, speech recognition, natural language processing, and predictive analytics. In particular, its application in Structural Health Monitoring (SHM) is relevant to the current thesis.

Machine learning can generally be classified into three categories: supervised, unsupervised, and reinforcement learning. In Figure 2.6, [25] shows how machine learning can be divided into these three main groups. From the figure it is seen how supervised learning includes classification and regression methods, while unsupervised learning deals with clustering. The training data is represented by colored points and triangles, and new data that can be predicted by the trained model is represented by yellow stars.

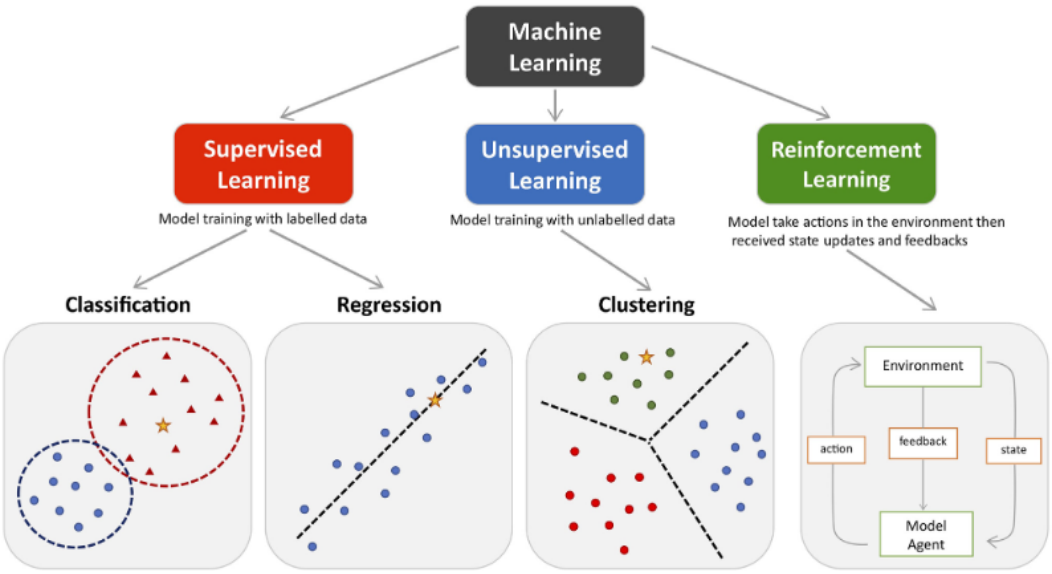


Figure 2.6: The main types of machine learning [25].

A key distinction can be made between supervised and unsupervised learning. When dealing with supervised learning, the algorithm is trained on a labeled data set. This is not the case for unsupervised learning. Labeled data is a type of data that includes both input variables and corresponding output variables, while unlabeled data does not include corresponding output variables. Figure 2.7, illustrates the difference between unlabeled and labeled data using a data set consisting of images of dogs and

cats. Here the input variables are the images of dogs and cats and the corresponding output variables are the labels, or tags, associated with each image. It can be seen how for the case of the labeled data set on the far left, each image has been labeled correctly, either as a *dog* or a *cat*. In addition to this, a separate labeled data set can be seen with the same input variables, however in this case the data set has the corresponding weight of each animal as its output variable. Comparing these two data sets to the data set marked as unlabeled, the input variables remain unchanged but there are no corresponding output variables.

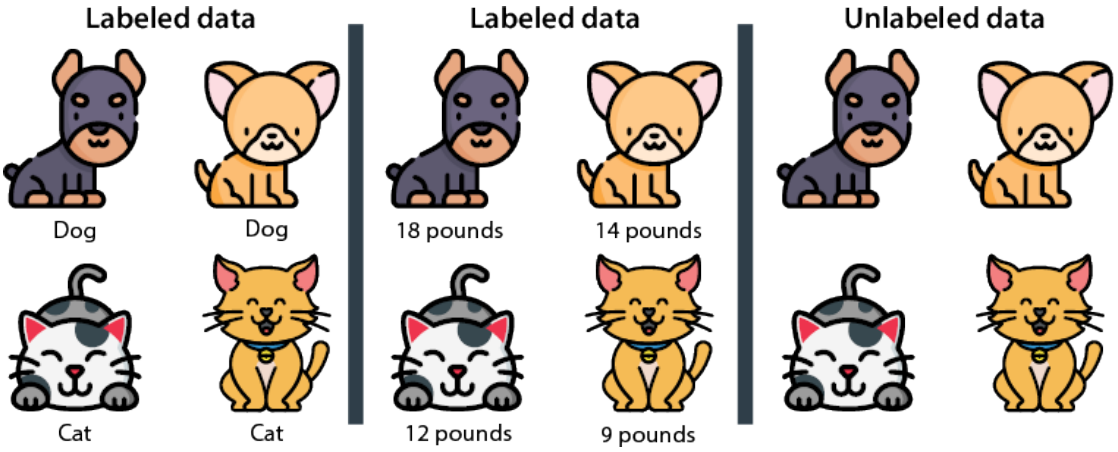


Figure 2.7: Example of labeled data and unlabeled data [26].

Tying this information together with the objective of this thesis, the choice between using supervised learning versus unsupervised learning is clear. In the context of structural health monitoring, supervised learning requires extensive data from the sensing system and physical models. This means that both damaged and undamaged information is available. Since this data is seldom available, unsupervised learning is the more appropriate choice. This requires creating an initial baseline assuming normal conditions of the bridge or structure. Then, the algorithm detects deviations from this baseline by analyzing new data from the sensing system [27, 28, 29].

2.5. Unsupervised Learning

Having determined that unsupervised learning will be employed for detecting damage in the analyzed bridge model in this project, it is of high importance to delve into and introduce an important unsupervised method and approach that will be utilized to achieve this objective. One of the methods being Principal Component Analysis (PCA) and the other approach being anomaly detection through clustering. These techniques play a critical role in various applications such as image compression, fraud detection, and fault detection. However, they can also be employed in the realm of SHM, which is the present undertaking.

2.5.1. Dimensionality Reduction

Dimensionality reduction is a critical process in data mining and machine learning, particularly when dealing with high-dimensional data. High-dimensional data, where the number of features can exceed the number of independent samples, is a common occurrence encountered by statisticians in both academic and industrial settings [30]. Dimensionality reduction techniques aim to minimize data dimensionality and expose the inherent data structure while forfeiting minimal accuracy and information. These techniques can be either linear or non-linear. Linear methods transform data into a lower-dimensional feature space through linear combinations of the original variables. Conversely, non-linear methods come into play when the initial data space encompasses non-linear relationships and structures. Non-linear dimensionality reduction techniques, including kernel PCA (kPCA), manifold learning methods, and autoencoders, generally create a lower-dimensional data representation while preserving inter-data point distances.

Linear techniques can be classified as either supervised or unsupervised. Unsupervised methods like PCA focus solely on the input feature matrix to discern patterns. In contrast, supervised methods

such as Partial Least Squares (PLS) and Linear Discriminant Analysis (LDA) take both features and responses into account. Several linear supervised techniques include LDA, PLS, and active subspace methods, while unsupervised linear methods include PCA, kPCA, and Factor Analysis. These techniques train on unlabeled data and consider only input features when seeking lower-dimensional representations [31].

As this thesis deals with linear data and employs unsupervised learning, PCA is the method of choice for dimensionality reduction in this context. Techniques such as PCA, greatly simplify the learning and training of regression and classification models for extracting patterns in the data. This is achieved by transforming the original variables into a new set of variables, thereby reducing the number of features while preserving the most important and sensitive ones. This process not only simplifies the data analysis but also helps in identifying patterns in the data. The overall goal of dimensionality reduction methods is to enhance the accuracy and efficiency of data mining by reducing the data set and increasing data quality [31].

When a structure is monitored for a period of one or two years, a vast amount of data is collected. However, analyzing and making sense of this large and complex data can be a daunting task, especially when many data features are present. As previously stated, PCA aims to simplify large and complex data sets by reducing the number of variables while retaining as much variation as possible in the data. This is accomplished through the transformation of the original variables into a new set of uncorrelated variables called *principal components*. The first few principal components are ordered in such a way that they capture most of the variation in all the original variables. The computation of these principal components involves solving an eigenvalue-eigenvector problem for a semi-definite symmetric matrix [32].

The monitoring of structures involves the use of tens or hundreds of different sensors to collect data. These sensors generate single or multi-feature data continuously with varying sampling rates. Extended monitoring, lasting several years, can result in the generation of a massive amount of data. However, not all of this data is necessarily used when it comes to the analysis of it. Additionally, as it pertains to the increasing dimensionality of the features, external environmental and operational factors play an important role [23]. Using PCA for dimensionality reduction would be highly beneficial in this scenario. PCA reduces the number of features while preserving the most important and sensitive features. It also retains the ability to identify patterns in the data [27].

In the field of SHM, PCA has been established as a valuable and essential tool. This thesis involves generating vasts amounts of data with numerous features over an extended period of time. Incorporating PCA into the project's methodology will be crucial. The methodology chapter will dig deeper into PCA, covering its theoretical interpretation and implementation.

2.5.2. Clustering

Clustering is a popular unsupervised learning approach which is often related to the task of anomaly detection. The goal is to identify data points that deviate from the normal patterns in a data set. This technique is used in various fields of work and is especially useful in the field of civil engineering as it relates to structural health monitoring.

One of the frequently used clustering techniques in anomaly detection is k-means. Other popular clustering methods include Gaussian Mixture Model (GMM), Spectral Clustering, and Density-Based Spatial Clustering of Applications with Noise (DBSCAN). These techniques group data points into clusters based on their similarities using different algorithms and are able to detect anomalies by recognizing data points that do not belong to any of the formed clusters. In the following sections, these clustering methods are explored further.

K-means

The k-means algorithm is a widely used, iterative method in unsupervised learning that partitions n objects into k clusters. The objective is to have data points within a cluster be as similar as possible while being as dissimilar as possible to data points in other clusters. In Figure 2.8, [33] displays a flowchart of the algorithm. k-means works by first randomly initializing a user specified k number of centroids, and then iteratively reassigning data points to the closest centroid and updating the centroid position until convergence.

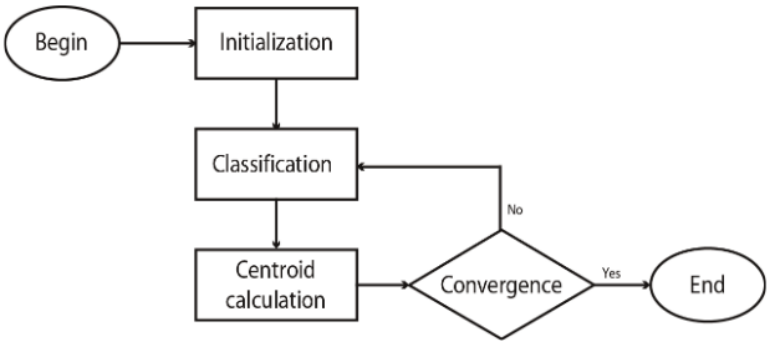


Figure 2.8: Standard k-means algorithm [33].

A pseudocode of the clustering algorithm is also supplied by [33], which gives a more comprehensive, step-by-step illustration of the algorithm's operation. In the first initialization step, N represents the set of objects to be clustered. It is defined as $N := \{x_1, x_2, ..., x_n\}$, where $x_i \in \mathbb{R}^d$ for $i = 1, 2, ..., n$ and $d \geq 1$ is the number of dimensions. The algorithm will attempt to divide the objects in N into k clusters, where k is specified as an input to the algorithm. M represents the set of centroids and is defined as $M := \{\mu_1, \mu_2, ..., \mu_k\}$, where $\mu_1, \mu_2..., \mu_k$ are the centroids for each of the k clusters. The k-means algorithm starts by randomly selecting k objects from N as the initial centroids, and then adjusts the centroids iteratively during the classification and centroid calculation steps, until the algorithm converges. The final positions of the centroids will determine the boundaries of the clusters.

Algorithm 1: Pseudocode of the k-means algorithm

```
// Input
N := x_1, ..., x_n;
k: number of clusters;
// Initialization
Randomly select k objects from N and set them as initial centroids;
M := μ_1, ..., μ_k;
// Classification
for x_i ∈ N do
    Calculate the Euclidean distance from x_i to each centroid;
    Assign object x_i to the closest centroid μ_k;
end
// Centroid calculation
for μ_k ∈ M do
    Recalculate the centroid as the mean of all objects assigned to μ_k;
end
// Convergence
repeat
    if the sum of squared distances of objects to their assigned centroids decreases by less
      than a predefined threshold then
        Stop the algorithm;
    else
        Go to Classification;
    end
until Stop the algorithm;
```

K-means is a partitioning technique and is well-suited for data that is spherical and equally sized. Partitioning techniques group objects based on the proximity of the objects to one another. With this in mind, k-means performs best when clusters are separated by clear, distinct boundaries and each cluster has roughly the same number of points. When the data has clusters with different shapes, sizes, or densities, or when the clusters are not separated by distinct boundaries, k-means may not perform as well. In simpler terms, the algorithm struggles when it comes to uncover clusters of complex shapes. Additionally, k-means is sensitive to the initial choice of centroids, so if the initial

choice is poor, the final solution may not be ideal. In these cases, alternative clustering algorithms such as Spectral Clustering, DBSCAN, or GMM, may be more appropriate [34].

Spectral Clustering

Spectral Clustering is a well-regarded and effective technique for clustering high-dimensional data. Its foundation is rooted in linear algebra, graph theory, and eigenvalue decomposition, which allows for the data clustering problem to be transformed into an eigenvector problem. The eigenvectors of a suitable matrix representation of the data are then related to the cluster assignments. The steps of Spectral Clustering are seen in Algorithm 2. In general, the procedure of the algorithm can be summarized as [35, 36]:

1. **Similarity Matrix Generation:**
   The initial stage involves generating a similarity matrix that depicts the relationships between the data points. Typically, the similarity matrix is represented as a matrix of pairwise distances between the data points, where higher distances indicate lower similarity.

2. **Normalized Laplacian Matrix Calculation:**
   The next step is to compute the normalized Laplacian matrix, which is a graph representation of the data. The Laplacian is calculated as $\mathbf{L} = \mathbf{D} - \mathbf{W}$, where $\mathbf{D}$ represents a diagonal matrix of node degrees and $\mathbf{W}$ is the weighted adjacency matrix.

3. **Eigenvector Factorization:**
   The third stage involves performing eigenvector factorization on the Laplacian matrix. This involves breaking down the matrix into its eigenvectors and eigenvalues. The eigenvectors of the Laplacian are equivalent to the cluster assignments, and they serve as the feature vectors for the clustering process.

4. K**-means Partitioning:**
   The final step involves performing k-means partitioning on the eigenvectors, which serve as the feature vectors for the clustering process. The k-means algorithm divides the data into k clusters based on the similarities between the feature vectors.

Algorithm 2: Pseudocode of the Spectral Clustering [35, 36]

Data: Similarity matrix $\mathbf{S} \in \mathbb{R}^{n \times n}$, number of clusters k
Result: Clusters $A_1, A_2, \dots, A_k$
Construct a similarity graph with weighted adjacency matrix $\mathbf{W}$;
Compute the normalized Laplacian matrix $\mathbf{L}$;
Compute the first k generalized eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ of the generalized eigenproblem $\mathbf{Lu} = \lambda \mathbf{Du}$;
Let $\mathbf{U} \in \mathbb{R}^{n \times k}$ be the matrix containing the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ as columns;
for $i = 1, 2, \dots, n$ **do**
  Let $\mathbf{y}_i \in \mathbb{R}^k$ be the vector corresponding to the i-th row of $\mathbf{U}$
end
Cluster the points $(\mathbf{y}_i)_{i=1}^n$ in $\mathbb{R}^k$ with the k-means algorithm into clusters $C_1, C_2, \dots, C_k$;
Define $A_i = j \mid \mathbf{y}_j \in C_i$ for $i = 1, 2, \dots, k$;
return the clusters $A_1, A_2, \dots, A_k$;

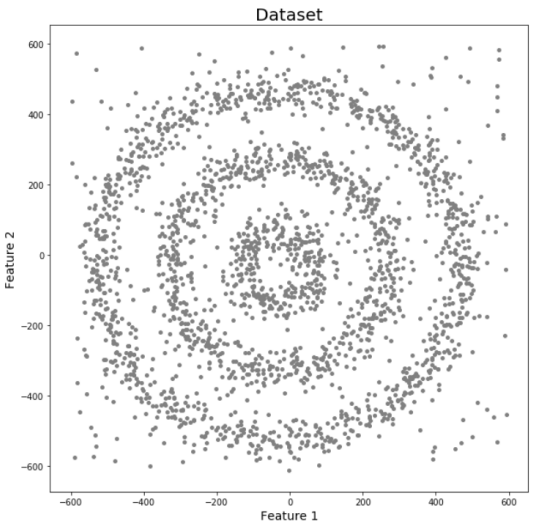
The Spectral Clustering method can perform exceptionally well in certain circumstances. However, it has limitations in terms of scalability. The algorithm involves computing eigenvectors, which can be a resource-intensive calculation, especially when handling large matrices [34].

DBSCAN

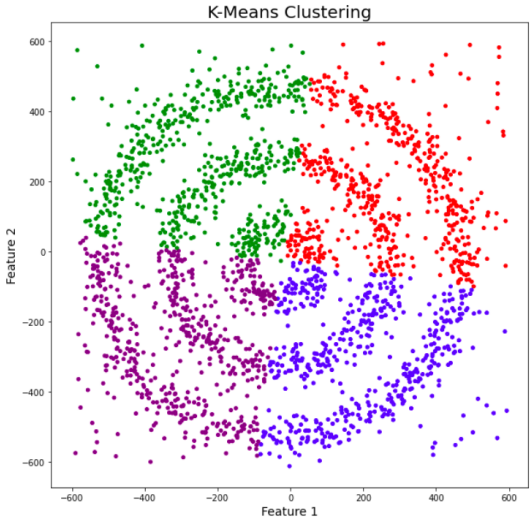
There are various clustering methods to choose from, many of which excel in identifying clusters with spherical or elliptical shapes. As previously discussed, methods such as k-means and GMM can struggle in identifying clusters with more complex shapes. Spectral Clustering is a good choice for more intricate shapes, but its performance is limited by its computational cost and may not be suitable for large datasets. In this thesis, a data set of three years of monitoring data from a modeled bridge

is used, resulting in a massive data set that poses potential challenges in terms of computational time and cost.

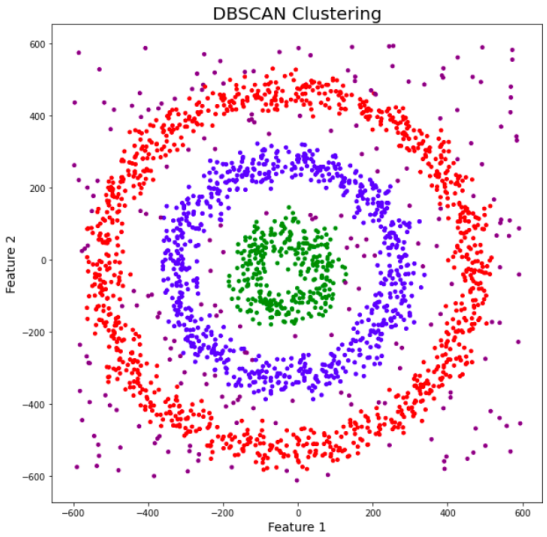
However, another method that excels in finding arbitrary-shaped clusters is DBSCAN. DBSCAN is a density-based clustering method that models clusters as dense regions in the data space, separated by sparse regions. This approach allows for the discovery of non-spherical clusters [34]. Monitoring data from structures can often be high-dimensional and have shapes that are not simple or spherical in nature. [37] showcases the strengths of the DBSCAN method by comparing it with other clustering methods such as k-means, using a data set of points densely present in the form of concentric circles. Figure 2.9 shows the unlabeled data set as well as the results of the two clustering methods. It is evident that DBSCAN performs better than k-means with this particular data set, managing to label three distinct clusters in a way that a human would. DBSCAN also successfully labels outliers, which are data points that do not belong to any specific cluster and are colored in purple. In comparison, k-means identifies four clusters, which are not labeled correctly.



(a) Unlabeled data.



(b) Clustering by k-means.



(c) Clustering by DBSCAN.

Figure 2.9: Clustering performed on unlabeled data set (a), by means of k-means (b) and DBSCAN (c) [37].

DBSCAN is a straightforward and efficient method that demonstrates key concepts crucial for any density-based clustering technique. The method defines three different types of points and requires only two user-defined parameters. By utilizing these parameters and a simple distance function,

DBSCAN is able to go from point to point and label each point either as a core point, border point, or a noise point. Inside of a cluster, there are only two kinds of points, a core point or a border point. In order for an arbitrary point p to be labeled as a core point, it must satisfy the user-defined parameters `minPts` and `eps`. That is, point p, with the help of a distance function, must satisfy the specified reachability condition set by `eps`, as well as fulfil the density threshold defined by `minPts` [38, 34]. To better understand these definitions and parameters they are further explained below:

- **`eps`:** Often referred to as the eps-neighborhood of a point p, [38] defines this parameter by $N_{\texttt{eps}}(p) = \{q \in D \mid dist(p,q) \leq \texttt{eps}\}$. Where, D is a database of points, and q is a point similar to p. The specified value of `eps` defines the radius of point p.
- **`minPts`:** Defines a limit for the number of neighbors required for a core point to exist. This parameter can be looked at as a density threshold when designating areas or regions as densely populated.
- **Core points:** A data point p is a core point if the number of neighboring points within a given distance of `eps` is exceeded. Core points are considered to be the dense regions of a cluster and satisfy the condition $|N_{\texttt{eps}}(q)| \geq \texttt{minPts}$.
- **Border points:** A data point that falls within $N_{\texttt{eps}}(p)$ of another core point but has fewer than the specified number of `minPts` of points within its defined distance `eps` from it.
- **Noise points:** If a data point does not satisfy the conditions of core points nor border points, it is classified and labeled as a noise point (outlier).

To help visualize how DBSCAN works, [39] provides Figure 2.10 as a way to depict the principles of how data points go through the process of satisfying the above definitions. In the figure, it can be seen how $\texttt{minPts} = 4$ as well as how each circle represents the defined radius of `eps`. It can be deduced from the image that the A labeled points are considered as core points, B and C as border points, and the lone blue point N as a noise/outlier point.

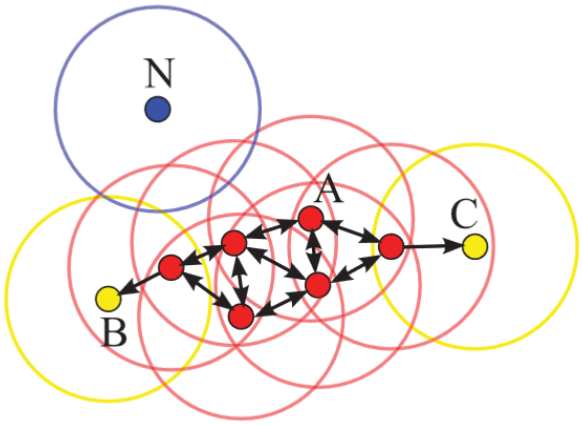


Figure 2.10: Illustration of DBSCAN [39].

Furthermore, Figure 2.10 illustrates some important definitions laid out in [38], relating to *direct density-reachability*, *density-reachability*, and *density-connectivity*. The definitions of these terms are given by [38] as:

- **Direct Density-Reachability:** A point p is *directly density-reachable* from a point q with respect to `eps`, and `minPts` if
  1. $p \in N_{\texttt{eps}}(q)$ and
  2. $|N_{\texttt{eps}}(q)| \geq \texttt{minPts}$ (core point condition).

- **Density-Reachability:** A point p is *density-reachable* from a point q with respect to eps and `minPts` if there is a chain of points $p_1, p_2, ...p_n$, with $p_1 = q$ and $p_n = p$ such that p_{i+1} is directly density-reachable from p_i.
- **Density-Connectivity:** A point p is *density-connected* to a point q with respect to eps and `minPts` if there is a point o such that both, p and q are density-reachable from o with respect to eps and `minPts`

With these terms available, the definition for a cluster and noise can be made. [38] defines these as:

- **Cluster:** Let D be a database of points. A *cluster* C with respect to eps and `minPts` is a non-empty subset of D satisfying the following conditions:
  1. $\forall\, p, q$: if $p \in C$ and q is density-reachable from p with respect to eps and `minPts`, then $q \in C$ (maximality).
  2. $\forall\, p, q \in C : p$ is density-connected to q with respect to eps and `minPts` (connectivity).
- **Noise:** Let $C_1, C_2, ..., C_k$ be the clusters of the database D with respect to parameters eps_i and $\texttt{minPts}_i$, $i = 1, 2, ..., k$. Then we define the *noise* as the set of points in the database D not belonging to any cluster C_i, i.e. $\text{noise} = \{p \in D \mid \forall\, i : p \notin C_i\}$.

With all of this in mind, the DBSCAN method is shown in Algorithm 3.

Algorithm 3: Pseudocode of the DBSCAN algorithm [34]

```
Data: Data points D, radius eps, minimum number of points minPts
Result: Clusters of data points labeled
foreach point p in D do
    if p is not visited then
        mark p as visited
        N = {q ∈ D | dist(p, q) ≤ eps}
        if |N| < minPts then
            mark p as noise
        end
        else
            create a new cluster
            add all points in N to the new cluster
            foreach point q in N do
                mark q as visited
                M = {r ∈ D | dist(q, r) ≤ eps}
                if |M| ≥ minPts then
                    add all points in M to N
                end
            end
        end
    end
end
```

GMM

GMM, also known as Gaussian Mixture Model, is a commonly used probabilistic model for clustering. The basic concept is to view each cluster in a data set as a Gaussian distribution, with the assumption that each data point originates from one of these Gaussian distributions. The model then determines the parameters of these distributions, such as the means and covariance matrices, to maximize the probability of the data given the model. When a data set contains two dominant clusters, a single Gaussian distribution may not accurately capture its structure. However, combining two Gaussians results in a more precise representation of the data [40].

In Figure 2.11, a basic data set is used as an example to demonstrate the use of Gaussian distributions in GMM. In the figure, blue curves show the contours of constant probability density. On the left, a single Gaussian distribution is fitted to the data using the maximum likelihood method. It can be observed that this distribution fails to capture the two dominant clusters in the data and places significant probability mass in the central region between the clusters where the data is sparse. On the right, the distribution is represented by a linear combination of two Gaussian distributions, which have been fitted to the data using the maximum likelihood method. This representation provides a better characterization of the data structure by accounting for the presence of two distinct clusters.

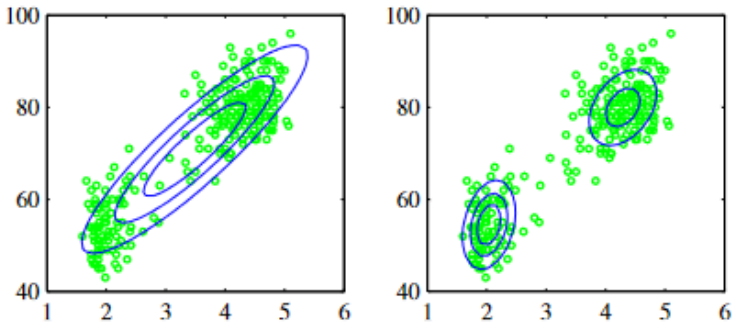


Figure 2.11: A plot of an example data set. On the left side, a single Gaussian distribution is applied to the data, while on the right, a linear superposition of two Gaussian distributions has been used to fit the data [40].

Figure 2.12 illustrates an example of 500 data points drawn from a mixture of 3 Gaussian distributions, as shown in Figure 2.13. Three different figures are present, (a) depicts the complete ground truth of the data set, (b) shows the unlabeled data, and (c) displays the values of the responsibilities associated with each data point. The responsibility values act as a measure of how much each component of the model is "responsible" for explaining a particular data point. These values are used to assign each data point to a cluster, with the cluster being defined as the mixture component with the highest responsibility value for that data point. Plotting the responsibilities results in a heatmap, revealing the extent to which each component is responsible for each data point. Higher values on the heatmap indicate a more significant correlation between the component and the data point. These plots help identify well-separated clusters of data points that have a strong association with a particular component, as well as ambiguous data points that may belong to multiple components.

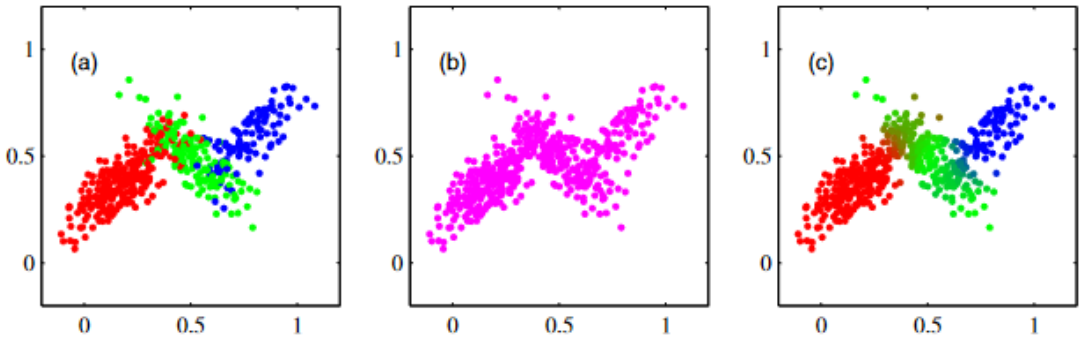


Figure 2.12: Example of 500 data points drawn from the mixture of 3 Gaussians [40].

Figure 2.13 depicts an illustration of a mixture of 3 Gaussian distributions in a two-dimensional space. Of the three sub-figures, (a) shows the contours of constant density for each of the mixture components, represented by red, blue, and green, with the values of the mixing coefficients displayed below each component, while (b) presents the contours of the marginal probability density, $p(x)$, of the mixture distribution, and (c) displays a surface plot of the distribution $p(x)$. More information relating to these figures and the theory behind mixtures of Gaussians can be found in [40].

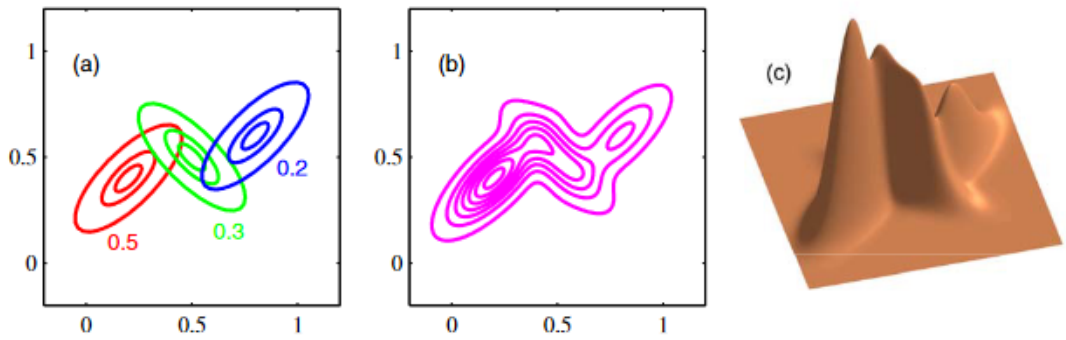


Figure 2.13: Illustration of a mixture of 3 Gaussians in a two-dimensional space [40].

A general pseudocode for GMM is seen in Algorithm 4 below. The algorithm starts by initializing the parameters of the Gaussian distributions that make up the mixture models, such as the means μ, covariance matrices Σ, and mixing coefficients π. The algorithm then alternates between two steps until convergence: the expectation (E) step and the maximization (M) step.

- **E-step**
  The responsibilities $\gamma_{n,k}$ for each data point x_n and mixture component k are computed using the current estimates of μ, Σ, and π. This step computes the probability that each data point belongs to each cluster.
- **M-step**
  The parameters μ, Σ, and π are updated using the responsibilities $\gamma_{n,k}$. This step maximizes the likelihood of the mixture model given the current responsibilities.

The algorithm continues to alternate between the E-step and the M-step until the change in the parameters Δ is less than a pre-defined tolerance ε. At this point, the algorithm has converged and the final clusters and their parameters are returned. If the maximum number of iterations T_{max} is reached before convergence, the algorithm will return the best results found so far [40].

Algorithm 4: Pseudocode of the GMM algorithm

Data: Data set X, number of clusters K, maximum number of iterations T_{max}, tolerance ε
Result: Clusters and their parameters
`// Initialization`
Initialize the parameters of the Gaussian distributions, such as means μ, covariance matrices Σ, and mixing coefficients π;
for $t = 1$ *to* T_{max} **do**
  `// E-step`
  Compute the responsibilities $\gamma_{n,k}$ for each data point x_n and each mixture component k using the current estimates of μ, Σ, and π;
  `// M-step`
  Update the parameters μ, Σ, and π using the responsibilities $\gamma_{n,k}$;
  `// Check for convergence`
  if $\Delta < \varepsilon$ **then**
    break;
  end
end
return the final clusters and their parameters

GMM is a well-known approach for clustering points in $\mathbb{R}^n$. However, it has some limitations, such as the requirement to make the assumption that each cluster's density is Gaussian in order to use parametric density estimators. Furthermore, the log-likelihood function of GMM can have multiple local minima, which can result in the need for multiple restarts of iterative algorithms to reach a satisfactory solution [41]. Due to these limitations, GMM may not be the optimal choice for clustering certain shapes, particularly those that are non-Gaussian or non-elliptical.

Nevertheless, GMM has been chosen as the preferred clustering method for developing an anomaly detection methodology in the context of this thesis. This choice is motivated by GMM's exceptional capability to model the distribution of healthy data points, thereby facilitating anomaly detection based on the likelihood of new data belonging to the healthy distribution or not. A detailed explanation of the proposed anomaly detection procedure is provided later in the report.

2.6. Conclusion

In this chapter, an overview of the relevant literature and key concepts for detecting anomalies in SHM in an unsupervised manner was provided. The significance of taking into account the impact of temperature on the natural frequency of bridges was emphasized, as well as simple damage scenarios that will be modeled in the developed FEM model. The use of PCA was introduced and its potential importance was discussed. Moreover, four different clustering algorithms were explored, making it possible to choose the most suitable algorithm for a specific task.

3

Methodology

3.1. General Framework

A methodology is now to be constructed in order to bring about the possibility of anomaly detection. The steps that must be taken to achieve the goal of this project can be summarized in a workflow diagram. The workflow is a rough estimate of the procedure and is shown in Figure 3.1. The goal is to outline a method that can reveal whether or not new incoming monitoring data of a bridge can be considered healthy or unhealthy. In other words, by defining what normal behavior of a specific bridge is, the method allows for the detection of abnormal behavior of said bridge as incoming data is measured in real-time. This makes the problem unsupervised in nature, as the data used to determine normal behavior is unlabeled.

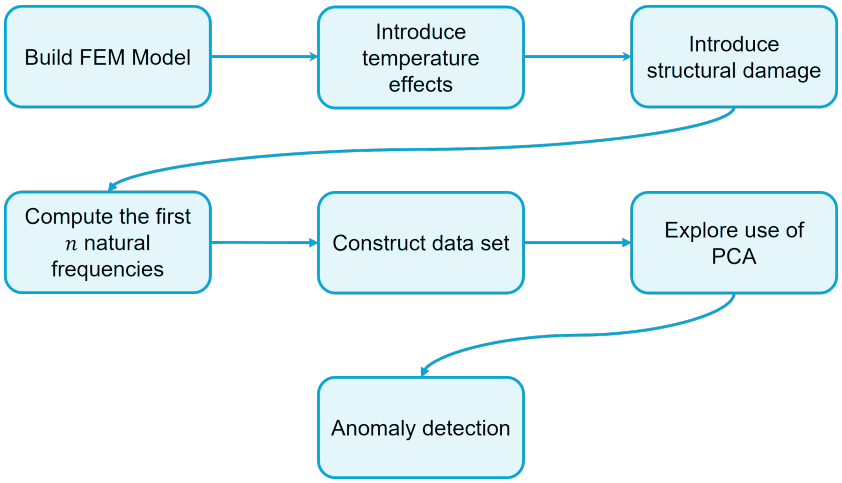


Figure 3.1: Workflow diagram.

The framework is based on going from a simple FEM model to detecting abnormal behavior in the analyzed bridge model. After having developed a concrete method that achieves the sought after goal, the method can consequently be transferred for case studies of bridges or different types of monitored structures.

3.2. Data Generation & Modeling

Data generation can be attributed to the first few operations in the workflow defined in Figure 3.1. For any kind of anomaly detection, it is necessary to have data relating to the monitored structure. In a real-life scenario, sensors or accelerometers are placed throughout the structure to be monitored. These devices provide useful information that allow, for example, the natural frequencies of the monitored structure to be found. With no available monitoring data to make use of, a FEM model of

the structure can be prepared. With a FEM model, essential measurement data can be synthetically produced. This becomes the first important step out of many, to developing a methodology that can be applied to detect abnormal behavior of bridges using unsupervised machine learning. The unsupervised learning aspect of this methodology is highly important, as it eliminates the need for damaged condition data which is a challenge in and of itself to come by and acquire. By eliminating the need for damaged condition data, monitoring becomes more cost-effective and less complex.

In this thesis, a simple FEM model is created to represent a two-dimensional bridge. This model is made with the aid of the *feastruct* package in `Python`. The geometry of the bridge can be seen in Figure 3.2. The bridge consists of three spans and two piers. The smaller two spans on the left-hand side and the right-hand side are both 2.5 meters in length while the middle span is 5 meters long. The two piers have a height of 2.5 meters.

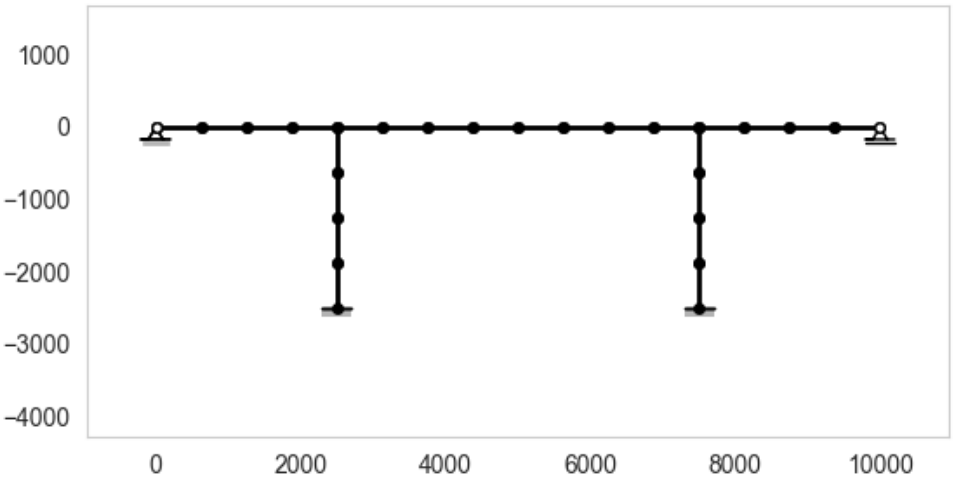


Figure 3.2: Geometry of the bridge to be analyzed.

In total there are 24 Euler-Bernoulli beam elements and 25 nodes connecting all of them. The left hand-side of the bridge deck is simply supported, that is, displacement is restricted in the horizontal and vertical directions. On the other side of the bridge, a roller support is present, where the displacement is only restricted in the vertical direction. The bottoms of each pier are fixed to the ground, fully restricting rotation and displacement of those nodes. In this FEM model, the section is assumed to be rectangular, with dimensions of 225 mm for the base and 300 mm for the height. The material is assumed to be concrete, and an elastic modulus of 30,000 MPa is used.

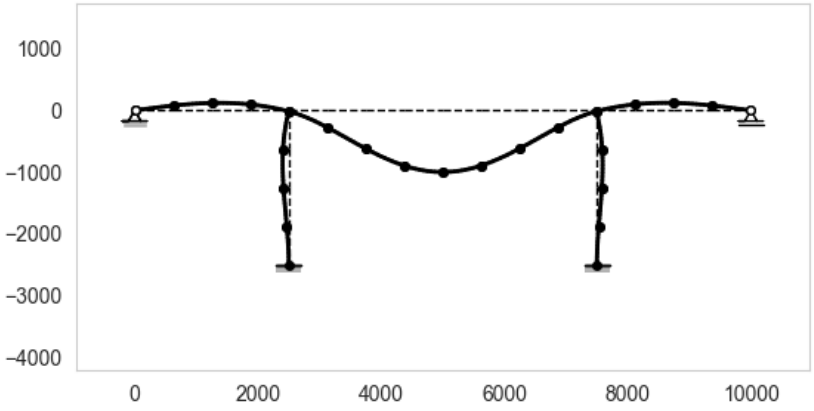


Figure 3.3: Natural frequency $f_1 \approx 34.65$ Hz and its mode shape.

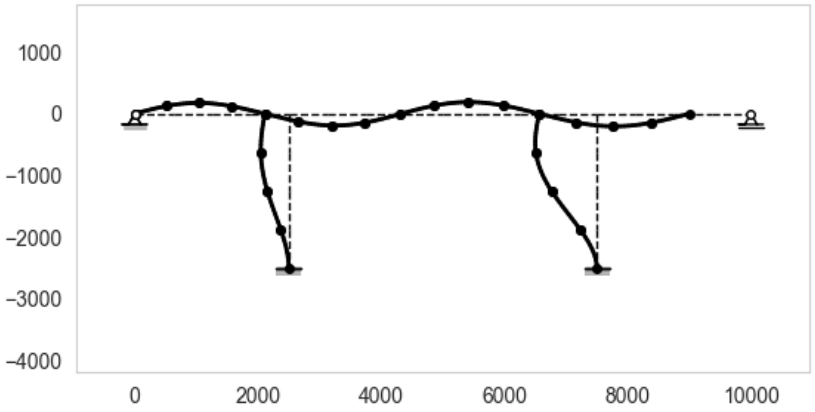


Figure 3.4: Natural frequency $f_2 \approx 82.04$ Hz and its mode shape.

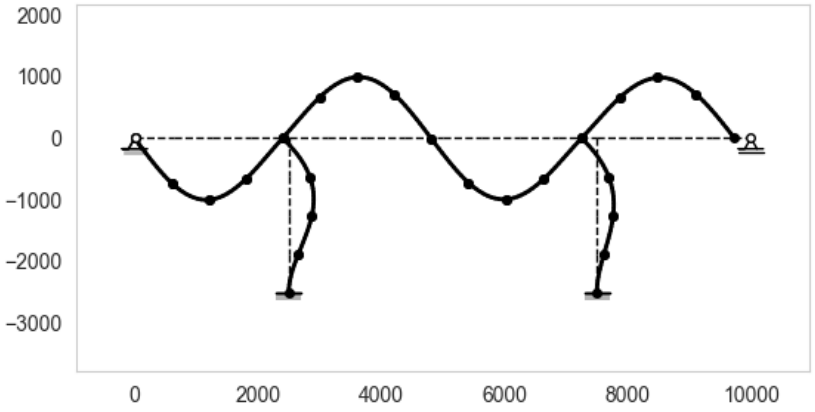


Figure 3.5: Natural frequency $f_3 \approx 85.34$ and its mode shape.

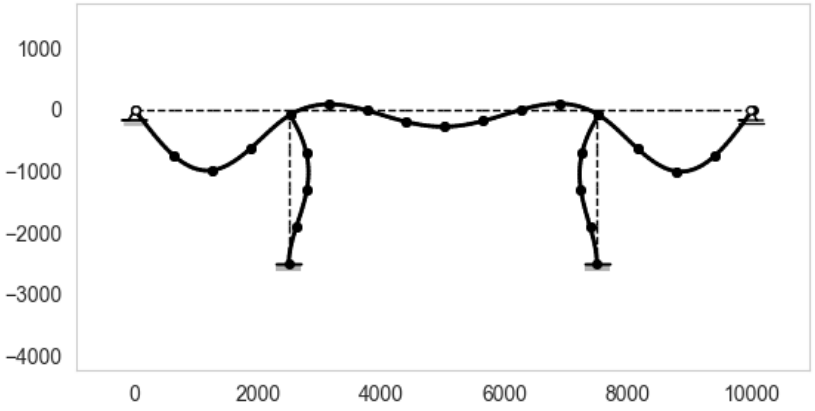


Figure 3.6: Natural frequency $f_4 \approx 97.29$ Hz and its mode shape.

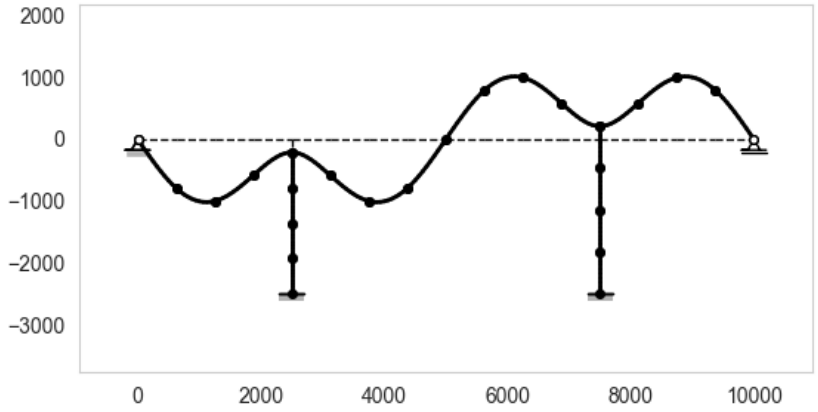


Figure 3.7: Natural frequency $f_5 \approx 110.61$ Hz and its mode shape.

As a part of the FEM model, it is important to introduce temperature effects to the model. By introducing the effects of temperature on the modeled bridge, it allows for a relationship between temperature and the natural frequencies to be established. This relationship is critical because it allows for the normal behavior in the modeled bridge to be determined. By comparing new data to existing data, deviations from the established relationship between temperature and natural frequencies can potentially be observed. This step is crucial for progressing further with the methodology and detecting anomalies. This relationship is seen in the Literature Review chapter and is implemented into the FEM model. Establishing such a relationship, allows for the behavior of the modelled bridge to be analyzed over time. This is done by importing real temperature data over a specific time span. Computing the first n natural frequencies over the desired time span will lay out the 'training'* data to be used.

In addition to the temperature effects, a type of structural damage ought to be introduced to the analyzed bridge model. This structural damage could be a settlement of one of the supports over a specific time span. This time span should be in addition to the one used to generate the original healthy data, as introducing structural damage to the model will result in the obtention of unhealthy data when computing the first n natural frequencies.

After computing the first n natural frequencies over time for healthy conditions of the modeled bridge as well as for abnormal conditions, a data set is consequently constructed for both states. One data set represents the healthy state which is used to define a behavioral baseline for the bridge and the other data set to represent the damaged state of the bridge. This data set will prove useful when it comes to detecting abnormal behavior of the analyzed bridge. An example of such a data set is presented in Table 3.1, which is derived from the first case study using the methodology outlined in this chapter. In this study, the first five natural frequencies are computed over a period of three years. Settlement (δ_s) is introduced to the model at the beginning of the third year and linearly increases from 0 mm to 500 mm during the last year, resulting in anomalies being introduced to the data set. Therefore, the first two years are considered healthy data while the third year is classified as unhealthy. It is important to note however, that the column containing the settlement information is disregarded and treated as if it were not there.

*In this context, the term 'training' data is used to refer to the data that is used to establish a baseline for normal or healthy behavior of the monitored structure, even though the term is typically associated with supervised machine learning rather than unsupervised learning.

Table 3.1: Constructed data set.

n	T [°C]	f_1 [Hz]	f_2 [Hz]	f_3 [Hz]	f_4 [Hz]	f_5 [Hz]	δ_s [mm]
1	1.129	35.62	84.33	87.73	100.01	113.71	0
2	0.616	35.64	84.41	87.77	100.02	113.69	0
3	-0.279	35.67	84.46	87.92	100.26	113.93	0
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
13139	2.635	35.12	82.34	87.10	100.61	112.22	499.89
13140	1.785	35.44	81.18	87.39	99.55	112.80	500.00

3.3. Data Preprocessing

It is of high importance to make sure that all data is preprocessed appropriately before it subjected to either PCA or clustering. This preprocessing is both necessary for the original healthy data as well as the incoming anomalous data. Before PCA or clustering is performed on the original data set, the data needs to be scaled. With the help of *scikit-learn*'s `StandardScaler` object in `Python`, the data features are standardized by removing the mean and scaling to unit variance. This is carried out using Equation 3.1,

$$Z = \frac{x - \mu}{\sigma} \tag{3.1}$$

with Z being the standard score of a sample x, μ is the mean, and σ is the standard deviation of the samples. Exploring normalization of data prior to or after standardization can be desirable. Normalization scales the variables between the range of -1 and 1, and can be advantageous in some cases. However, in this scenario, only scaling the data sets is necessary. As monitoring data of bridges deals with features that are not necessarily at the same scale, e.g. temperature [°C] and natural frequencies [Hz], standardization removes the units of measurements and normalizes the data to a common scale. This is precisely why preprocessing data is crucially important as it helps to ensure that the data is in a suitable format for further analysis. This allows for more accurate results from any subsequent analysis of the data.

Using the process and procedure of PCA outlined in the following section, it is realized that projecting incoming data onto the already generated PCA space is possible by performing PCA a second time, for the incoming anomalous data. First, the original data is projected onto a new subspace of lower dimensionality, as determined by the number of components specified when initializing the PCA. Next, using the already fitted 'healthy' PCA, the same PCA transformation is applied to the incoming data. This projects the incoming data onto the same lower-dimensional subspace as the 'healthy' data [42]. With this projection, outlier/anomaly detection can be conducted as a next step. Prior to this projection, the anomalous data set undergoes the same preprocessing as the original data set.

3.4. Principal Component Analysis

To better understand or explore what, if anything, occurs when structural damage is introduced to the model, PCA will be performed on the healthy data set. By performing PCA, linear dimensionality reduction takes place while capturing the variance in the data set and preserving the most important information that it holds. PCA can be particularly useful for large data sets as it can help to reduce dimensionality However, it is important to determine whether or not PCA is appropriate for the specific data set being analyzed, as there may be a significant loss of information when reducing to a lower dimension, which could ultimately impact subsequent analysis in a negative way.

PCA is especially useful for large data sets, as it enables one to interpret the data more easily. As one can imagine, datasets consisting of thousands of entries and hundreds of features are as often

as not, difficult to make sense of or interpret. Performing PCA on such a database, will reduce the dimensionality of the dataset, while maintaining as much statistical information as possible. This statistical information can be related to the variability. Preserving the statistical information or variability of a dataset is achieved by identifying new variables, known as principal components (PCs), which are linear functions of the original variables. These new variables successively maximize variance and are uncorrelated with one another. The process of finding these variables, or principal components, is done by solving an eigenvalue/eigenvector problem [43].

Different methods are available for performing PCA. There are two methods that are considered to be more common than others, one being related to using EVD (Eigenvalue Decomposition) and the other, using SVD (Singular Value Decomposition). The procedure for the EVD method and finding the PCs is summarized in the following steps:

1. Acquire data set and organize it into a matrix $\mathbf{X}$
2. Compute the mean-centered matrix $\mathbf{X_c}$
3. Compute the covariance matrix $\mathrm{Cov}(\mathbf{X_c})$
4. Compute the eigenvalue decomposition of $\mathrm{Cov}(\mathbf{X_c})$
5. Obtain the principal components

The principal components are found after having computed the eigenvectors from the covariance matrix and ordering the eigenvalues in decreasing value. By rearranging the eigenvalues, the corresponding matrix containing the eigenvectors are also sorted by decreasing eigenvalues. The column vectors of this matrix, the eigenvectors, are now ordered in terms of significance. A higher eigenvalue represents more significance or importance when it comes to information loss. That is, by choosing to ignore the eigenvectors with the largest eigenvalue would be extremely consequential as much information related to the original data set would be lost. However, by ignoring the principal components, or the eigenvectors with the smallest eigenvalues, one can minimize the information loss and perform a successful dimensionality reduction of the original data set [44, 45].

Similarly, the principal components can be obtained or computed using SVD. The SVD theorem is stated as the following in [46]:

Theorem. Let $\mathbf{X} \in \mathbb{R}^{m\times n}$. Then there exist orthogonal matrices $\mathbf{U} \in \mathbb{R}^{m\times m}$ and $\mathbf{V} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{X} = \mathbf{U\Sigma V}^T \tag{3.2}$$

where

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{S} & 0 \\ 0 & 0 \end{bmatrix}$$

and $\mathbf{S} = \mathrm{diag}(\sigma_1, ..., \sigma_r)$ with

$$\sigma_1 \geq ... \geq \sigma_r > 0.$$

This factorization proves useful when it comes to PCA, as the principal directions for the matrix representing the data set, $\mathbf{X}$ are contained in the columns of $\mathbf{V}$ or the rows of $\mathbf{V}^T$. To eliminate one direction, the last column of $\mathbf{V}$ can, for example, be removed. To project the original data points from $\mathbf{X}$ onto the remaining principal directions, one can perform the matrix multiplication $\mathbf{XV}$, after having removed any unwanted principal directions. Note that since SVD ensures that the principal directions in $\mathbf{V}$ have a magnitude of one, this projection is simply a matrix multiplication. As described by [47], the procedure for computing the PCs is summarized in the following steps:

1. Acquire data set and organize it into a matrix $\mathbf{X}$
2. Compute the mean-centered matrix $\mathbf{X_c}$
3. Use the SVD to find $\mathbf{X_c} = \mathbf{U\Sigma V}^T$
4. Keep the first k columns of $\mathbf{U\Sigma}$, and the first k rows of $\mathbf{V}^T$
5. Compute $\mathbf{X_c} \approx \mathbf{U\Sigma V}^T$ with the reduced $\mathbf{U\Sigma}$ and $\mathbf{V}^T$

With these steps, the data set $\mathbf{X}$ with a dimensionality of $n \times d$ is reduced to $n \times k$ (where $k < d$). The principal components of $\mathbf{X}$ are found as the columns of

$$\begin{aligned} \mathbf{X} &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \\ \mathbf{X}\mathbf{V} &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{V} \\ \mathbf{X}\mathbf{V} &= \mathbf{U}\mathbf{\Sigma} \qquad \text{Since } \mathbf{V}^T\mathbf{V} = \mathbf{I}\,. \end{aligned}$$

To illustrate the effects of PCA, Figure 3.8 shows how a two dimensional data set is reduced to a one dimensional one. The blue points of the original data set are projected onto the first principal component, in this case $\mathbf{u_1}$. These projected points are marked as red.

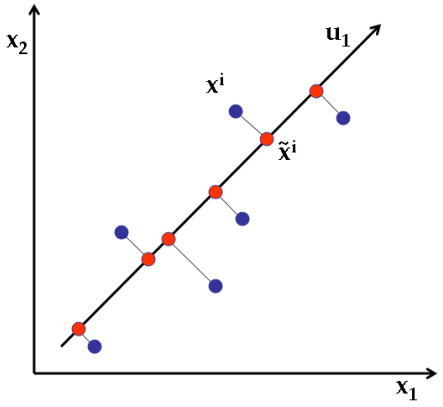


Figure 3.8: Example of PCA: 1-D projection (red) of 2-D points (blue) in the original space [48].

With the help of Figure 3.9, the goals of PCA are better demonstrated. It can be seen how PCA seeks to identify a set of directions that maximizes the variance of the data and minimizes the residuals. By maximizing the variance, the residuals that are associated with information loss and projection error are minimized.

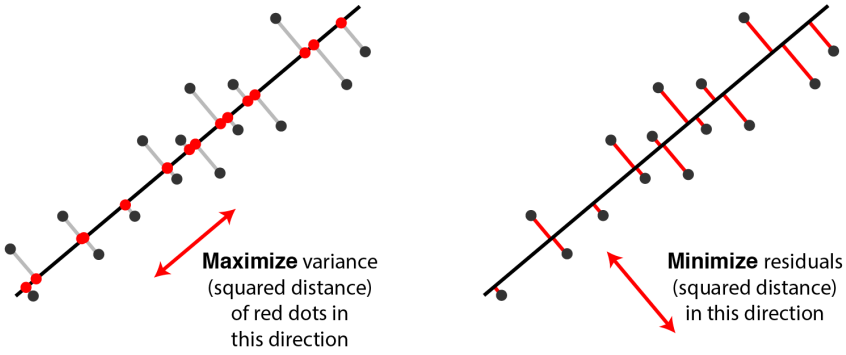


Figure 3.9: Two equivalent views of principal component analysis [49].

3.5. Anomaly Detection

The detection of unusual patterns in data is referred to as anomaly detection. These patterns deviate from the expected behavior and can be called anomalies, outliers, noise, discordant observations, aberrations, or contaminants depending on the field of application. Anomaly detection is widely used in various fields such as credit card fraud detection, insurance or healthcare fraud detection, cyber-security intrusion detection, fault detection in safety-critical systems, and military surveillance for identifying enemy activities [50]. Another significant and pertinent area where anomaly detection

is used, is the field of structural health monitoring (SHM). SHM is highly relevant to this thesis, and thus a focus will be placed on this field in relation to anomaly detection. In Figures 3.10 and 3.11, two examples of anomaly detection have been graphically illustrated.

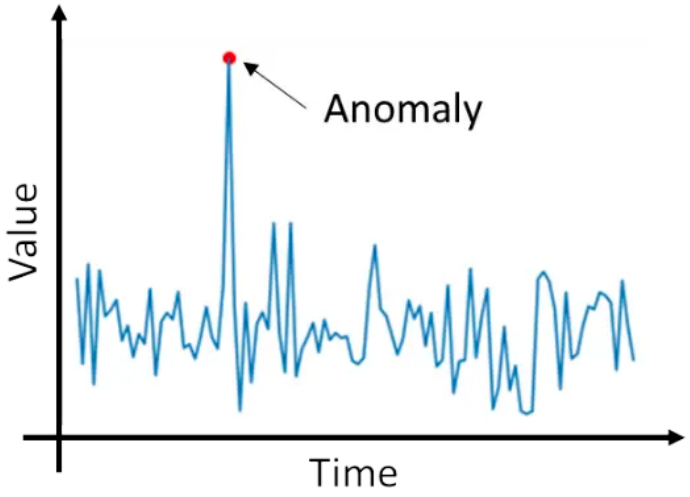


Figure 3.10: Example of anomaly detection [51].

In the above figure, a normal and established trend of values can be observed over time. However, an anomaly has been detected in the graphed data as a peak value appears approximately at one-third's of the time axis. In comparison to this anomaly detection, is the figure below. In Figure 3.11, the majority of the graphed data points are considered normal and are represented by N_1 and N_2, while the minority, considered abnormal, are represented by O_1, O_2, and O_3.

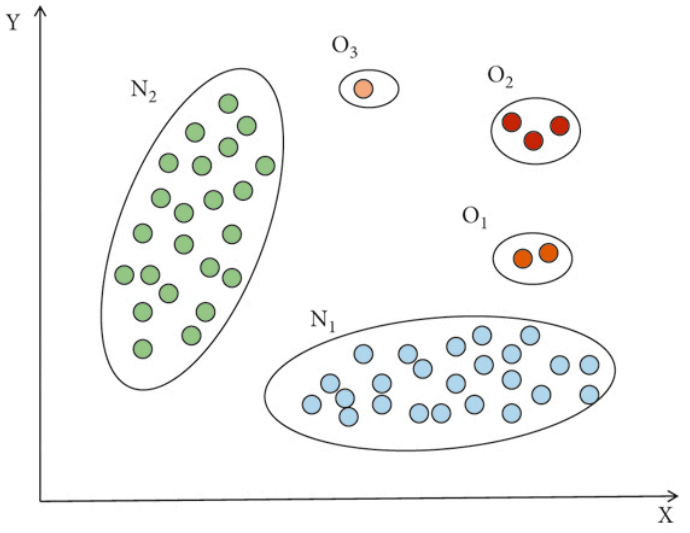


Figure 3.11: Example of anomaly detection using clustering [52].

3.6. Clustering

Cluster analysis involves researching methods and algorithms for grouping objects based on their intrinsic characteristics or similarity. Unlike classification or discriminant analysis, cluster analysis does not rely on pre-existing category labels. This lack of prior identification makes clustering an unsupervised learning technique, whereas classification and discriminant analysis are supervised. The goal of clustering is to discover patterns in data, making it an exploratory method [53]. As was mentioned previously, the monitoring data that is dealt with is unlabeled in nature. This makes clustering a suitable choice as a next step when it comes to anomaly detection. Clustering provides the possibility to detect and identify outliers, which are values that are distinct from the rest of the groups (see Figure 3.11), as if they were created by a different process or mechanism [34]. Relating this to the problem at hand, this mechanism could could very well be associated with structural damage of the bridge being analyzed or monitored. With this consideration in mind, a clustering method must now be chosen for the anomaly detection of the monitoring data.

When it comes to choosing a clustering method, it is important to realize that the choice of method depends on the characteristics of the data and the requirements of the problem at hand. Therefore, it might prove useful to try multiple methods and evaluate the performance of each one to ultimately find the best one. In Table 3.2, four different types of clustering methods are listed along with the general characteristics associated with each one. After determining the characteristics of the data set being used, certain algorithms are chosen to investigate further. It is important to note, all data sets are unique in their own way and different methods are more successful than others. Figure 3.12 shows seven different clusters all with their own unique shape, density, and pattern. One can imagine that not one single clustering technique can label each cluster correctly [53]. Emphasis must therefore be put on testing and comparing different methods when it comes to clustering a data set, as each data set has its own characteristics.

Table 3.2: Overview of clustering methods [34].

Method	**General Characteristics**
Partitioning methods	– Find mutually exclusive clusters of spherical shape – Distance-based – May use mean or medoid (etc.) to represent cluster center – Effective for small- to medium-size data sets
Hierarchical methods	– Clustering is a hierarchical decomposition (i.e., multiple levels) – Cannot correct erroneous merges or splits – May incorporate other techniques like micro-clustering or consider objects as "linkages"
Density-based methods	– Can find arbitrary shaped clusters – Clusters are dense regions of objects in space that are separated by low-density regions – Cluster density: Each point must have a minimum number of points within its "neighborhood" – May filter out outliers
Grid-based methods	– Use a multi-resolution grid data structure – Fast processing time (typically independent of the number of data objects, yet dependent on grid size)

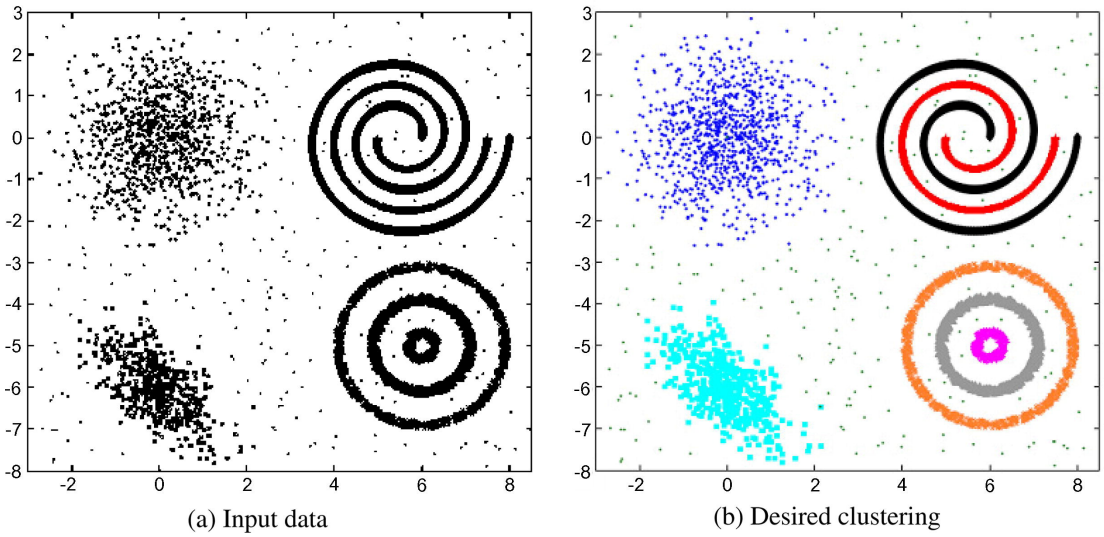


Figure 3.12: A diverse example of clusters and clustering. Seven unlabeled clusters in (a) and labeled in (b) [53].

In this thesis, the method or algorithm of choise is GMM. This method proved to be the most successful in terms of clustering the monitoring data as well as labeling outliers after having compared it with other algorithms such as, k-means, spectral clustering, and DBSCAN. In the following sub-chapter, GMM will be further introduced while other algorithms that could have been used for other data sets are discussed in greater detail in the literature review chapter.

3.6.1. GMM-Based Clustering

As has been discussed and introduced in the Literature Review, GMM is a clustering method that assumes each cluster in a data set can be represented by a Gaussian distribution. The model considers that each data point comes from one of these Gaussian distributions and estimates their parameters, such as the mean and covariance matrices, to maximize the probability of the data given the model. This approach is suitable when a data set has multiple clusters, and a single Gaussian distribution may not adequately capture its structure. By combining multiple Gaussians, the model provides a more accurate representation of the data, making it a widely used probabilistic model for clustering [40].

As highlighted in the Literature Review and later in the Results chapter, the relationship between the temperature and computed natural frequencies is linear (assuming that the effects of asphalt are excluded). Considering the addition of noise, it can be envisioned that the distribution of data points for a single frequency can be modeled using a single Gaussian distribution. Based on this idea, it is proposed to represent the healthy data (i.e., the first two years of data) using Gaussian distribution(s) and to define a boundary for this distribution to identify incoming abnormal data that either belongs to the healthy distribution cluster or an abnormal behavior cluster. This approach is aimed at distinguishing between normal and abnormal behavior and is an important step in the proposed anomaly detection method.

Once the healthy data distribution has been modeled, the log-likelihood of each data point in both the healthy data set and the entire data set is computed. The log-likelihood serves as a measure of how well a data point fits the Gaussian distribution learned by the GMM. Based on the log-likelihoods of the healthy data points, a threshold is defined. The data points in the entire data set are then classified as either abnormal or healthy depending on their log-likelihoods. If the log-likelihood of a data point is lower than the defined threshold, it is considered abnormal; otherwise, it is considered to belong to the healthy distribution. The threshold is determined by selecting a percentile of the log-likelihoods that captures most, if not all, of the points belonging to the initial healthy data distribution.

To elaborate further on the steps of the proposed GMM-based anomaly detection procedure, the following steps in the procedure are:

1. **Standardization:** The data preprocessing section has already produced the standardized data

set. Consequently, step 1 has already been completed.

2. **GMM:** The GMM models the data as a mixture of K Gaussian distributions. The probability density function of the GMM for a data point $\mathbf{x}$ can be written as:

$$p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\mathbf{x} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\right)$$

where:

- π_k is the mixing coefficient for the k-th Gaussian, with the constraint $\sum_{k=1}^{K} \pi_k = 1$*
- $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is the Gaussian distribution with mean $\boldsymbol{\mu}_k$ and covariance matrix $\boldsymbol{\Sigma}_k$ for the k-th component.
- K is the number of components in the GMM.
  - The Gaussian distribution can be calculated as:

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{1}{(2\pi)^{d/2}|\boldsymbol{\Sigma}_k|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\mathbf{x} - \boldsymbol{\mu}_k)\right),$$

    where d is the number of features in the data set.

The GMM is fitted to the standardized healthy data set X'_{h}, by estimating the parameters $\pi_k, \boldsymbol{\mu}_k$, and $\boldsymbol{\Sigma}_k$ using the Expectation-Maximization (EM) algorithm (see Algorithm 4).

3. **Log-likelihood:** Using the probability function of the GMM, for each data point $\mathbf{x}_i$ in both the healthy data set and the entire data set, the log-likelihood is computed as:

$$\ell = \log\left(p\left(\mathbf{x}_i\right)\right)$$

The log-likelihood measures how well a data point fits the probability distribution learned by the GMM. It is calculated as the logarithm of the probability that the data point belongs to the GMM. A higher log-likelihood means the data point is more likely to be part of the modeled distribution, while a lower log-likelihood suggests the data point is less likely to belong to the distribution.

4. **Threshold:** A threshold τ is set based on the log-likelihoods of the healthy data set. The threshold is set to the value corresponding to the φ percentile of the log-likelihoods in the healthy data set:

$$\tau_\varphi = \operatorname{percentile}\left\{\ell\left(X'_{\mathrm{h}}\right), \varphi\right\}$$

A low threshold value or percentile φ corresponds to setting a larger boundary for the healthy data point distribution. This value varies for different data sets, and it is recommended to choose it in such a way that it groups the healthy data points together effectively. This ensures that incoming healthy data points are not mislabeled. Using a high percentile, the number of false positives may potentially rise (healthy data points incorrectly identified as abnormal).

5. **Labeling:** For each data point $\mathbf{x}_i$ in the entire data set, it is labeled as abnormal (1) or healthy (0) based on its log-likelihood $\ell(\mathbf{x}_i)$:

$$\operatorname{label}\left(\mathbf{x}_i\right) = \begin{cases} 1 & \text{if } \ell\left(\mathbf{x}_i\right) < \tau_\varphi \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

*The constraint $\sum_{k=1}^{K} \pi_k = 1$ is necessary because the mixing coefficients π_k represent the probabilities that a data point is drawn from the corresponding Gaussian component in the Gaussian Mixture Model. Since a data point must come from one of the K Gaussian components, the sum of the probabilities across all components must equal 1. This constraint ensures that the GMM forms a valid probability distribution over the data points.

3.7. Validation

Once the clustering process is finished, validating the results can be performed as the final step. One of the key advantages of using synthetic data is that it allows for easy validation of the clustering outcome. Even though the data is being treated as unlabeled, it is possible to verify the accuracy of the labeling since the data has been intentionally manipulated to reflect the gradual or sudden deterioration of the bridge's structure. This makes it possible to check whether the data points have been correctly labeled. Furthermore, conducting the anomaly detection on a purely healthy data set and comparing the results to those obtained when abnormalities are present enables a determination of whether the anomaly detection is functioning effectively or not.

3.8. Conclusion

In summary, the methodology presented in this chapter consists of the following steps:

1. Build a FEM model of the structure
2. Introduce structural damage
3. Compute the first n natural frequencies over a period of k years
4. Construct a data set based on the computed natural frequencies
5. Perform PCA on the healthy data set and assess its effectiveness
6. Clustering to identify anomalies
    a. If PCA proves beneficial, perform clustering using the PCA results
    b. Otherwise, perform clustering directly on the data to identify anomalies without using PCA

This approach offers a comprehensive framework for detecting abnormal behavior in bridge monitoring data. By simulating structural damage, analyzing the resulting data, and identifying any unusual patterns through clustering techniques, this method allows for efficient and accurate detection of potential issues in bridge structures. This can, in turn, aid in timely maintenance and ensure the continued safety and functionality of such critical infrastructure.

4

Results

After outlining a methodology for detecting abnormal behavior in monitored structures, three case studies will be conducted and their results will be presented. First, the application of the methodology procedure will be explained in detail. Then, using the previously introduced FEM modeled bridge, three different case studies will be considered to simulate various damages in the modeled bridge, including settlement, reduced stiffness, and a combination of both. These case studies will help evaluate the performance of the unsupervised outlier detection method.

The bridge's performance will be evaluated over a three-year period, taking into account the effects of temperature on its natural frequencies. The temperature data used in the study is collected from Amsterdam, Netherlands, from the start of 2017 to the end of 2019 [54]. The data is recorded every two hours throughout the three years. Additionally, various types of structural damage are introduced to the model at the start of the third year. These damages are not present during the first two years and will gradually increase in magnitude and severity over the course of the final year. These damages are considered as anomalies in the study of the bridge's behavior, as they deviate from the normal, undamaged state.

With this in mind, the introduction of structural damage can be viewed as a form of an anomaly when studying the behavior of the bridge. The first two years of the bridge's performance are considered to be normal, as there is no damage present. This serves as a baseline for what typical, healthy behavior of the bridge looks like. Once the damage is introduced, it creates an anomaly that must be identified. The objective is to detect any deviation from the normal behavior previously observed, caused by the structural damage. The damage types that will be implemented to the FEM model are the following:

1. Settlement of pier
2. Loss of stiffness
3. Combination of pier settlement and loss of stiffness

For each damage type, a case study will be conducted where the previously defined methodology is applied and put to the test.

4.1. Case Study I: Settlement

The first case study that is performed involves the addition of settlement as a form a structural damage. This settlement is applied to the right pier support and starts at the beginning of the third year. The settlement (δ) starts at 0 mm and increases linearly to the final value of 500 mm over the course of the year. The progression of this damage can be seen in Figure 4.1, which shows the bridge's geometry halfway through the third and final year. The final settlement value of 500 mm equates to 10% of the length of the mid-span. This means that the settlement increases by 41.67 mm each month of the third year.

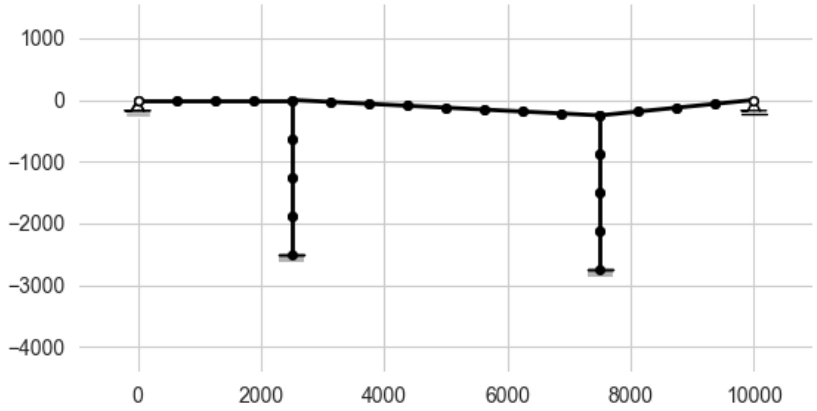


Figure 4.1: Introduced settlement ($\delta = 250$ mm) to the right pier.

4.1.1. Relationship Between Temperature and Frequency

After having computed the natural frequencies over time while taking into account the effects of temperature, one can observe the relationship between frequency and temperature. In Figure 4.2, the first natural frequency, f_1 has been plotted over time. It is important to note that noise has been artificially added to the computed and plotted first natural frequency. On the same graph, the temperature can also be observed from the year 2017 until the end of year 2019.

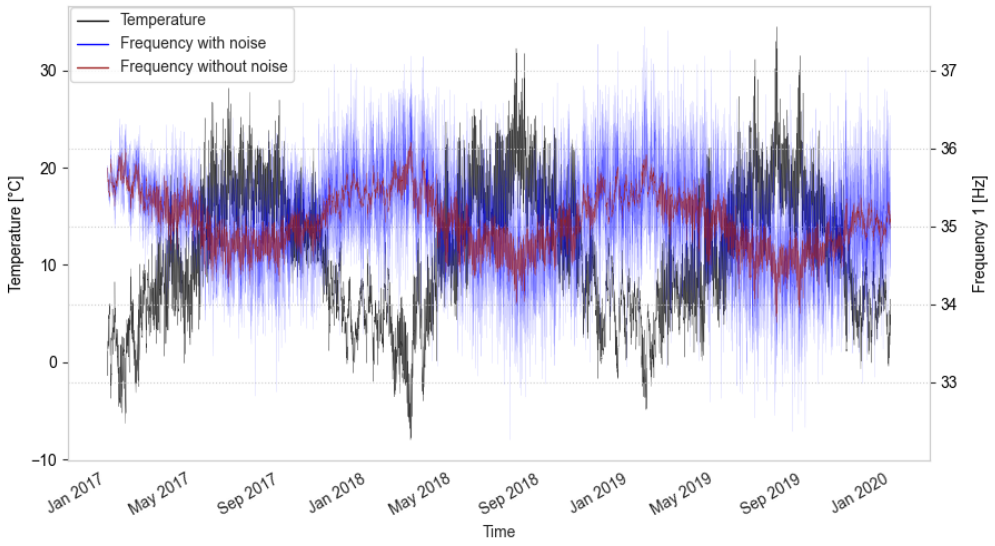
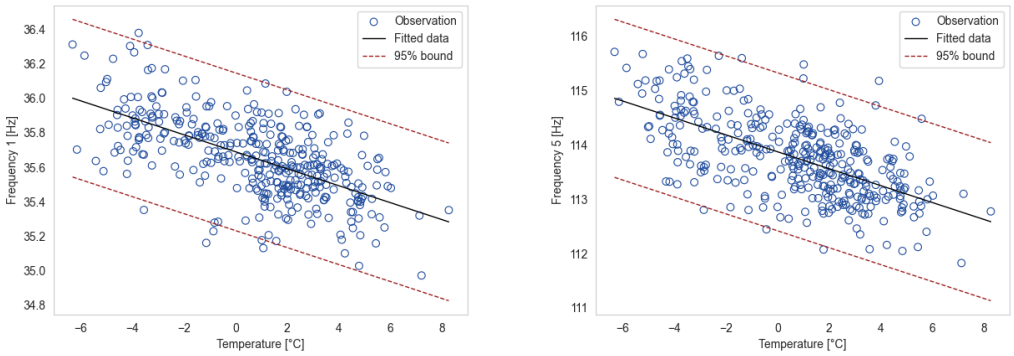


Figure 4.2: Frequency and temperature plotted over time.

To simulate the measurement errors that typically occur when monitoring bridges, a Gaussian-distributed error is added to the computed natural frequencies.* This is done for all presented case studies in this thesis. This noise is introduced in the time domain and is applied to each frequency separately. For each computation, the standard deviations of the natural frequencies are calculated. The standard deviations are then multiplied by 1.96/2 (a Z score value of 1.96 corresponds to the 95%

*Although the natural frequencies are computed in this thesis, it is important to keep in mind that when it comes to the natural frequencies of real structures they would need to be identified as they are not directly measured. This is a step which is neglected due to the fact that real SHM data is not being used in this research.

confidence interval for a Gaussian distribution). This multiplication is used as a scaling parameter when generating random values of normally distributed noise that is added to the calculated original frequency value. By choosing to multiply the standard deviation by $1.96/2$, the noise that is introduced reflects what is seen in real scenarios.[†] Figure 4.3 shows the relationship between the calculated first and fifth natural frequencies and the temperature. These computed frequencies include the artificially generated noise. From the figure, it can be observed that most data points or observations fall within their 95% bounds after fitting a linear regression model to the data. The temperature measurements used in Figure 4.3 come from the first month of the year 2017.



(a) First natural frequency vs. temperature. **(b)** Fifth natural frequency vs. temperature.

Figure 4.3: Relation of natural frequencies to temperature of the FEM modeled bridge.

When only considering a single day of temperature measurements, similarities can also be observed with what is seen in [55] for the Tsing Ma Bridge. Figure 4.4 shows this relationship for a single day at the end of March 2017.

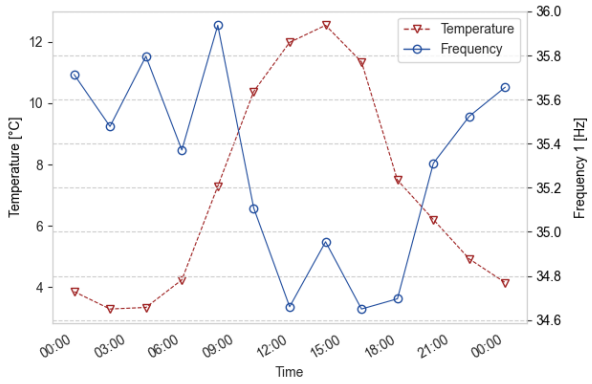


Figure 4.4: Relationship between the first natural frequency and the temperature (single day).

After having validated the accuracy of the incorporated artificial noise and computing the first five natural frequencies over the three years, a data set can be formed. This data set was previously shown in the Methodology chapter, as an example of what a constructed data set might look like (Table 3.1). Figures 4.5 and 4.6, allows for the visualization of the data. The two-dimensional plot of

[†]In [55], the relationship of natural frequencies and temperature is shown for the Tsing Ma Bridge. From this relationship a 95% error bound is observed. To replicate noise similar to what is generally seen, it is desirable that artificial noise that is added to the computed natural frequency lands most commonly within a 5% range of the computed value. Therefore in combination with the computed standard deviations, a Z score value of $1.96/2$ is defined. Initially, using 1.96 resulted in excessive noise, so it was divided by 2, which then better resembled the noise pattern seen in [55].

the frequency versus the temperature shows a linear relationship between the two, similar to what was found in [7], when a comparison is made with Figure 2.2. As the temperature increases, the natural frequency that is obtained will consequently decrease. No clear relationship between the frequencies and the introduced settlement can be observed.

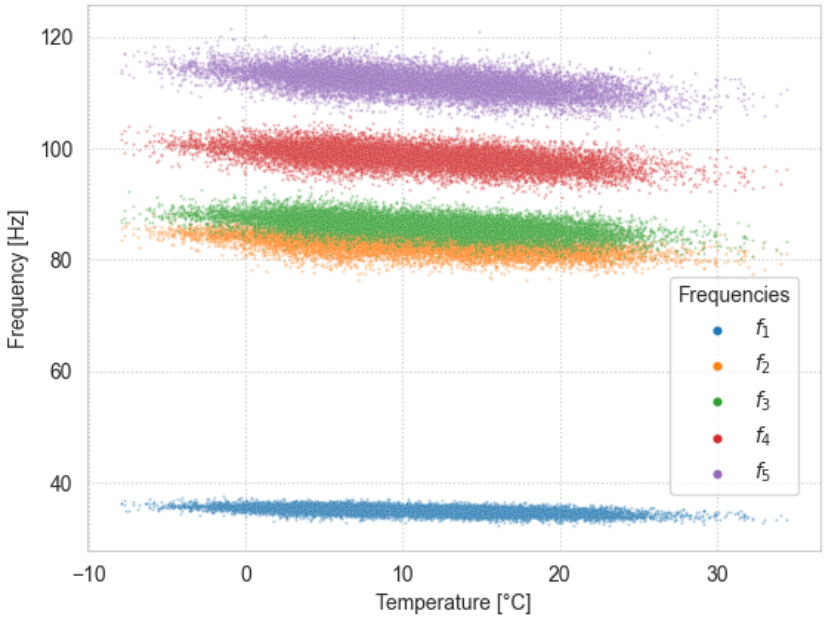


Figure 4.5: Frequency plotted against the temperature.

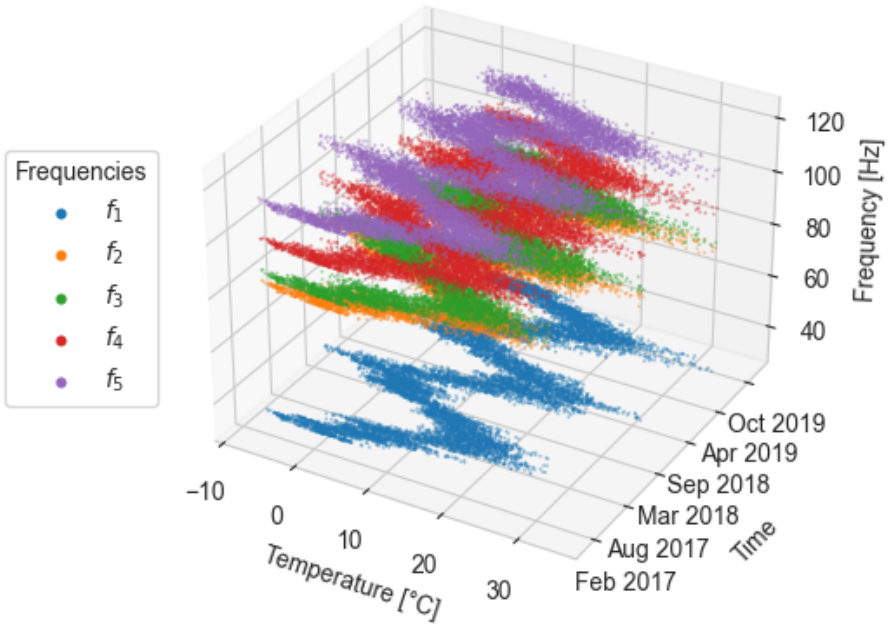


Figure 4.6: Frequency plotted against temperature and time.

4.1.2. Principal Component Analysis

Following the general framework that was outlined in Chapter 3.1, the next step to endeavor on is to explore the effectiveness of dimensionality reduction on the constructed data set with the use of PCA. The current data set from Table 3.1 has a dimensionality of six. The dimensions being the temperature T, and the first five frequencies, f_1, f_2, f_3, f_4, and f_5. Here the column representing the settlement δ_s has been dropped, as it is important to realize that in real conditions, one would not be aware that settlement is taking place. This point addresses the ultimate goal of attempting to detect behavioral changes, in other words anomalies, that point towards settlement or other types of structural damage taking place, without being aware of their presence.

The data set is first imported into `Python` and split up into what is known to be the undamaged state of data as well as the damaged state of the data. Next, using the obtained healthy and unhealthy data sets, data related to time and settlement, are dropped. With this, the data sets are appropriately preprocessed before performing PCA.

By performing PCA, it is possible to generate a scree plot, which is a graphical representation showing the explained variance of each component. This plot indicates which components are most significant in the data set. The explained variance of the five principal components generated through PCA is illustrated in Figure 4.7. The figure highlights that each principal component holds a substantial amount of explained variance, implying that every component is relatively important. Based on the scree plot, it can be inferred that performing dimensionality reduction may not be beneficial since discarding any component would result in significant loss of information.

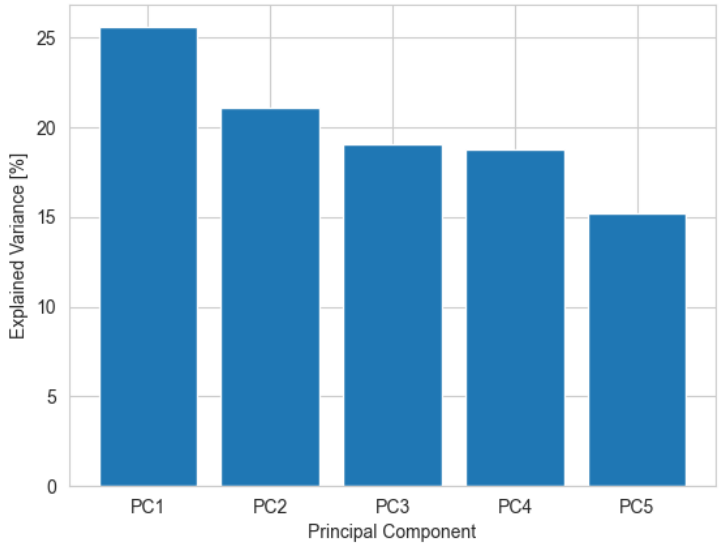


Figure 4.7: PCA scree plot for a purely healthy data set.

Based on the results of the preliminary PCA analysis, it is clear that utilizing PCA for anomaly detection would not be advantageous. Therefore, it is recommended to directly apply unsupervised clustering on the data set. However, moving forward with PCA might prove to be beneficial for other cases/data sets.

4.1.3. Clustering

Prior to finalizing the anomaly detection method for this research, multiple strategies were explored and evaluated for their performance. The investigated methods included:

- PCA and DBSCAN
  - This approach involved performing PCA on the data, followed by DBSCAN for outlier detection.

- Single Frequency - GMM
  - This method only uses one of the five obtained natural frequencies, models the healthy distribution for that frequency, and defines a threshold for incoming data.
- First Five Frequencies - GMM
  - This method uses all five obtained natural frequencies and is the method selected for the anomaly detection in this thesis.
- Single Frequency - PCA and GMM
  - This approach only uses one of the five obtained natural frequencies, performs PCA on the data, models the healthy distribution for that frequency, and defines a threshold for incoming data.
- First Five Frequencies - PCA and GMM
  - This method uses all five obtained natural frequencies, performs PCA on the data, and models the healthy distribution for each frequency using GMM.

The below figures show how the methods performed. From the figures it is evident that the superior method is indeed the one that includes all five natural frequencies to model the healthy distribution using GMM, the same one that has been used in this project.

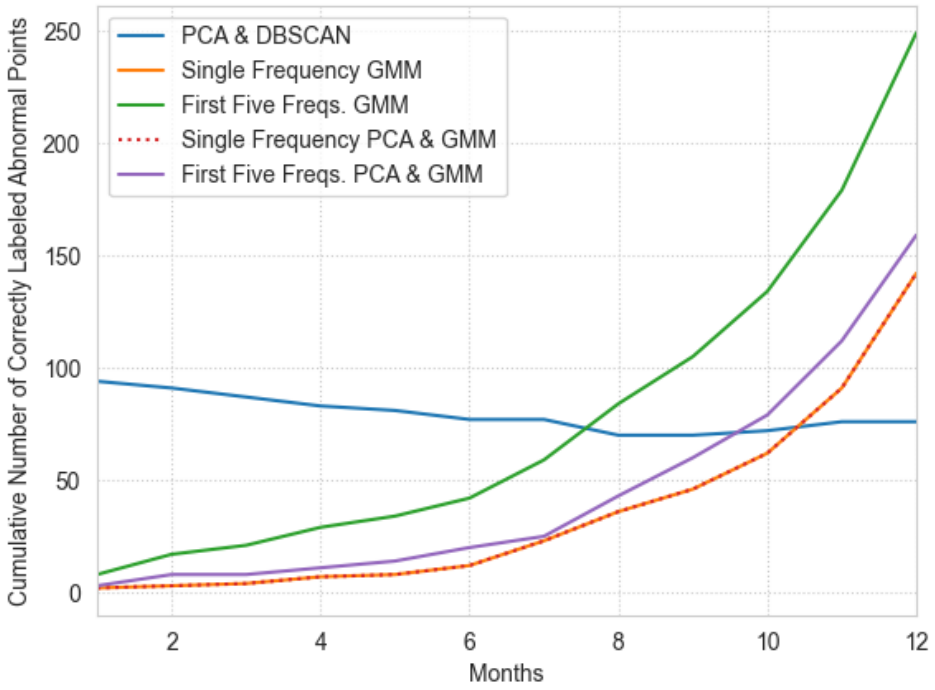


Figure 4.8: Comparison of methods - 500 mm settlement.

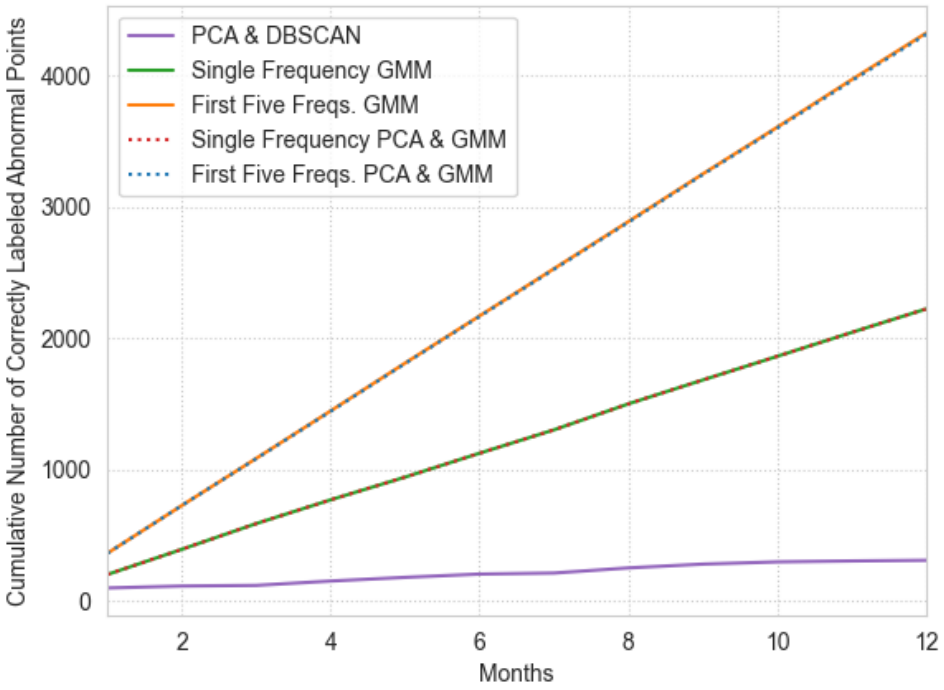


Figure 4.9: Comparison of methods - 10% Sudden loss of stiffness.

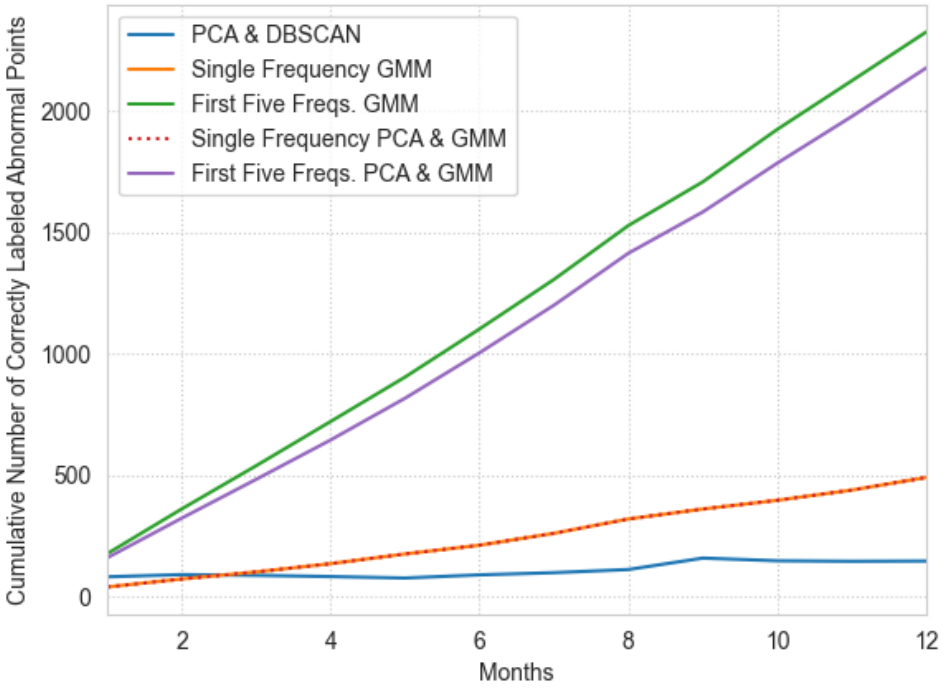


Figure 4.10: Comparison of methods - 5% Sudden loss of stiffness.

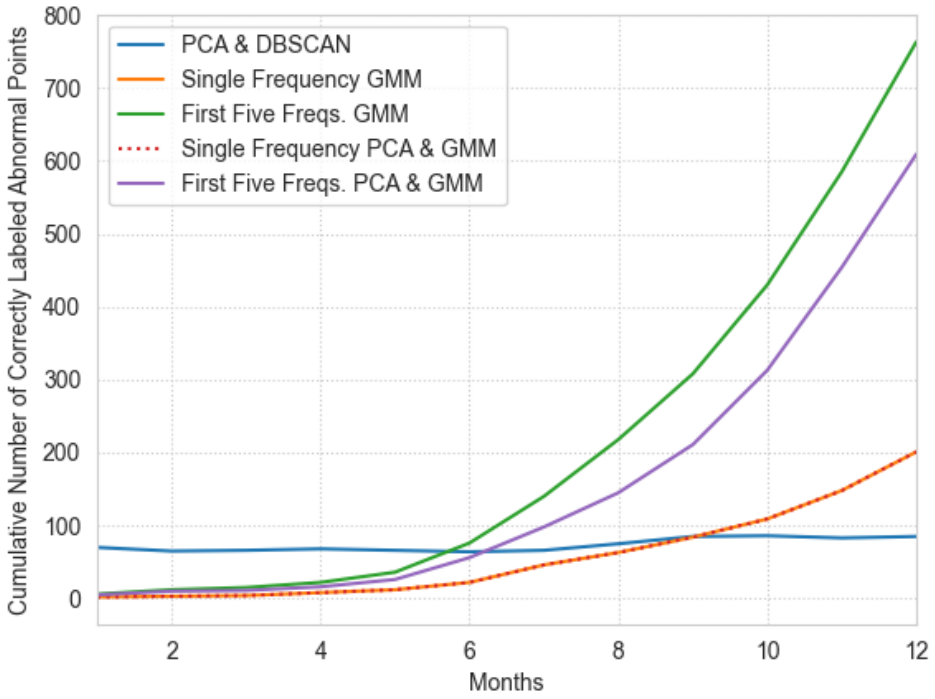


Figure 4.11: Comparison of methods - 10% gradual stiffness reduction.

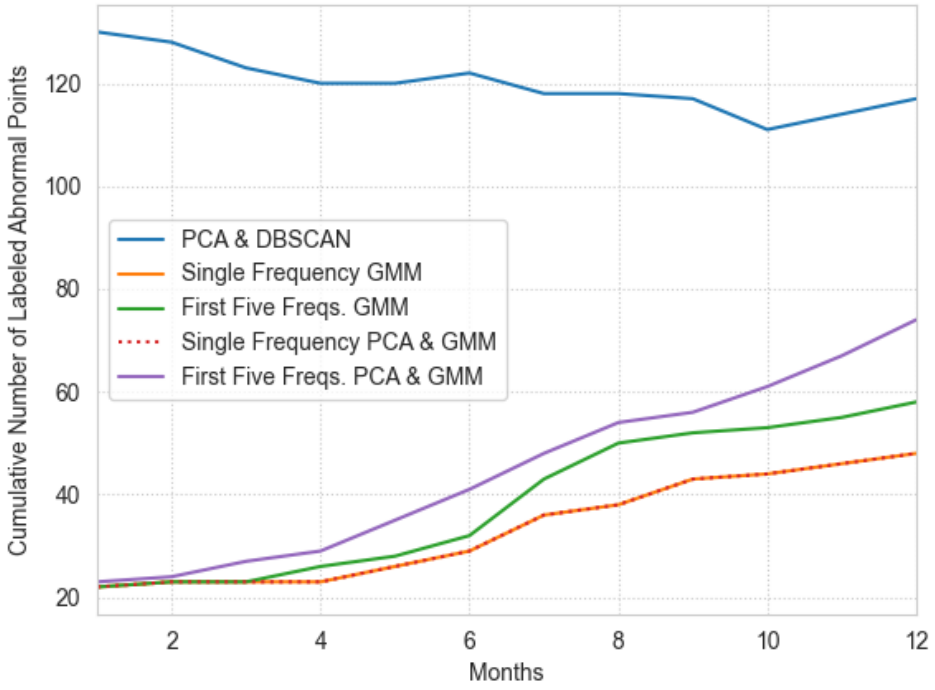


Figure 4.12: Comparison of methods - Anomaly detection on a healthy data set.

After having gone through experimentation with various clustering algorithms, including those discussed in the Literature Review as well as the above investigated methods, GMM clustering was found to be the optimal algorithm for the task at hand. By using a GMM-based anomaly detection method, the distribution of the healthy data can be modeled. By establishing a model for the healthy data distribution, it enables one to define a baseline distribution for the structure's normal behavior. Subsequently, setting a threshold makes it possible to determine the likelihood of a new incoming data point belonging to the healthy distribution or not.

GMM-Based Anomaly Detection

Using the understanding of GMM clustering that was introduced in the Literature Review, a procedure for detecting anomalies can be formed. This thesis proposes a procedure that works by fitting a GMM model with a user-specified number of components to the healthy data and then using it to identify abnormal data points in the entire data set (healthy and abnormal data combined).

The proposed procedure takes advantage of the GMM clustering approach to identify data points that deviate from the expected distribution of the healthy data. By fitting the GMM model to the healthy data, the procedure can capture the underlying structure of the data and then use it to detect anomalies or outliers in the entire data set. Below is an elaboration on the proposed procedure, which can be broken down into the following steps:

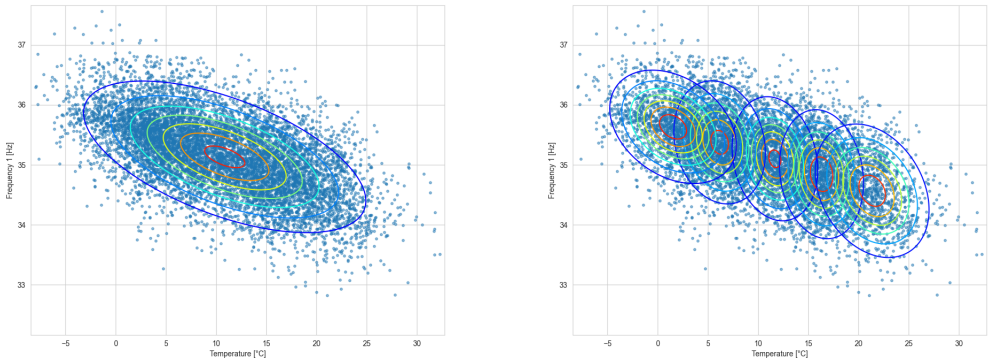
1. The data set is divided into healthy and abnormal subsets. First two years healthy and all data after belongs to the abnormal subset.
2. The healthy data points are used to calculate scaling parameters (mean and standard deviation) and apply this scaling to make the data have a mean of 0 and a standard deviation of 1. This scaling is applied to both the healthy and the entire data set
3. A GMM model with a user-specified number of Gaussians or components is trained using the standardized healthy data.
4. Log-likelihoods for each data point in both the healthy data set and the entire data set are computed. Log-likelihood serves as a measure of how well a data point fits or belongs to the Gaussian distributions learned by the model.
5. A threshold is set based on the log-likelihoods of the healthy data points. In this case, the threshold is set to the value corresponding to the 0.25 percentile of the log-likelihoods in the healthy data set. After experimenting with various threshold values, a percentile value of 0.25 was ultimately chosen as it effectively captured most true healthy data points within the same distribution. Selecting a higher threshold value could lead to mislabeling true healthy data points as abnormal, particularly those located in less dense areas near the distribution's perimeter. The 0.25 threshold value helps with minimizing the mislabeling of true healthy points as abnormal points.
6. The data points in the entire data set are labeled as abnormal or healthy based on their log-likelihoods. If the log-likelihood of a data point is lower than the threshold, it is considered abnormal, otherwise healthy.

The idea behind this approach is that the GMM learns the underlying structure of the healthy data and should assign high log-likelihoods to data points that closely resemble the healthy data. Data points that deviate significantly from the healthy data will have lower log-likelihoods and are considered anomalies. By comparing the log-likelihood of each data point to the threshold, the algorithm can detect abnormal data points that might indicate potential issues or failures in the structural system being monitored. In the Methodology chapter, detailed steps for the proposed procedure are provided. By following these steps, one can gain a deeper understanding of the proposed procedure and its underlying principles. With this in mind and having revisited and summarized the proposed procedure, it is now possible to present its results.

Clustering Results

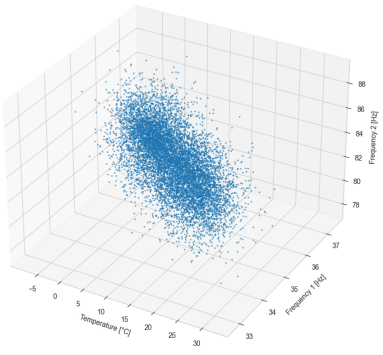
The proposed GMM-based anomaly detection procedure is applied in monthly intervals to monitor the data for abnormal behavior. After establishing the baseline distribution for the healthy data, incoming data is processed through the procedure in monthly batches until the abnormal behavior reaches its peak (in this case, 500 mm of settlement of the right pier).

Figure 4.5 suggests that a single Gaussian distribution is sufficient to model the distribution of one of the natural frequencies. However, the proposed procedure allows for flexibility in the number of Gaussian distributions used to model the distribution of the healthy data. As a way to investigate the effect of the number of Gaussians on the labeling of abnormal data points, the method is applied with 1, 5, 25, and 125 Gaussians. By varying the number of Gaussians, the proposed procedure can explore the trade-off between model complexity and performance in accurately labeling abnormal data points. The results of this analysis can help determine the optimal number of Gaussians to use for anomaly detection. To better visualize the GMM models the distribution of the healthy data, Figure 4.13 shows the data (of the first natural frequency) being modeled using 1 Gaussian and using 5 Gaussians. Additionally, Figure 4.14 presents a three-dimensional graph featuring the addition of the 2nd frequency, depicted from three distinct viewpoints. This figure helps to visualize the distribution better and come to a conclusion of how many Gaussians are appropriate to use.

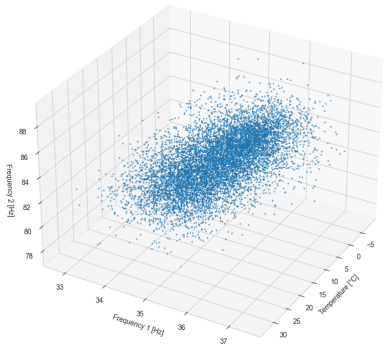


(a) Healthy distribution modeled using 1 Gaussian. **(b)** Healthy distribution modeled using 5 Gaussians.

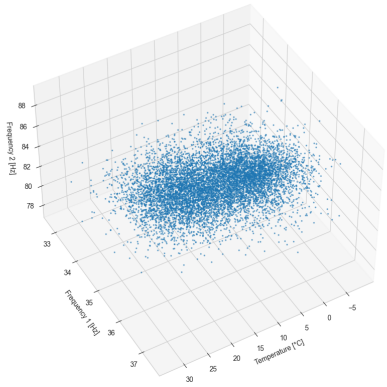
Figure 4.13: Healthy data point distribution of the first natural frequency modeled using GMM.



(a) 3D graph of T, f_1, and f_2. Viewpoint 1.



(b) 3D graph of T, f_1, and f_2. Viewpoint 2.

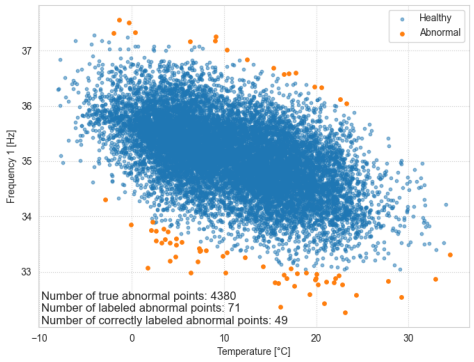


(c) 3D graph of T, f_1, and f_2. Viewpoint 3.

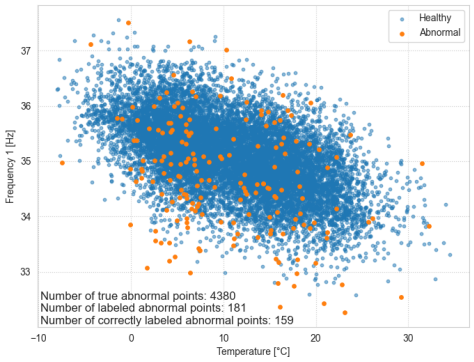
Figure 4.14: Plotting temperature, frequency 1, and frequency 2 on the same graph with different viewpoints.

The analysis of the healthy data suggests that the distribution follows a linear trend, indicating that using 1 Gaussian distribution for the GMM is likely sufficient for modeling the distribution of the healthy data.

As a quick comparison of whether to move forward with using a single frequency for the anomaly detection algorithm versus using all five, Figure 4.15 provides the results for cumulative number of labeled abnormal points after 12 months of gradual settlement (500 mm).



(a) Results for using only the first natural frequency.



(b) Results for using all five natural frequencies.

Figure 4.15: Anomaly detection results for different number of frequencies used after 12 months of settlement.

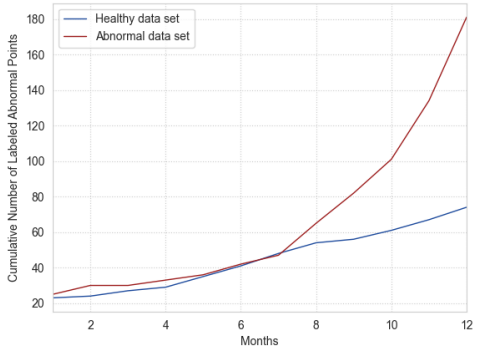
As demonstrated by the figure, using only the first frequency in the anomaly detection algorithm creates a prominent boundary around the healthy data point distribution. Such a clear boundary does

not emerge when all five natural frequencies are incorporated into the algorithm. The incorporation of all five frequencies allows the model to label abnormal points that fall within the healthy data distribution. This strategy results in a higher number of labeled abnormal points, specifically, 159 abnormal points are accurately labeled, compared to 49 when only the 1st frequency is used. This disparity can be attributed to the transition of the data going from univariate to multivariate, revealing a richer data set that enables the model to more accurately differentiate between healthy and abnormal data points. With this in mind, all five frequencies will be used for all upcoming clustering of the data.

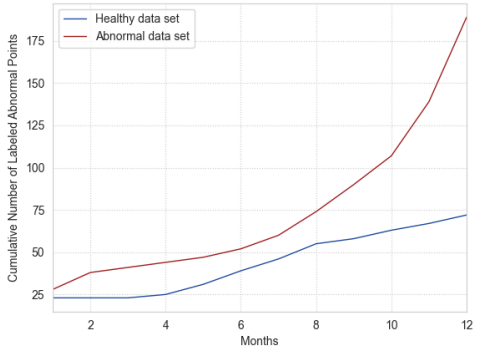
Figure 4.16 presents a comparison of the effectiveness of the anomaly detection procedure when using different numbers of Gaussians to model the distribution of the healthy data points. Each sub-figure shows the results of performing anomaly detection on a purely healthy data set as well as on the true abnormal data set. The figures show the cumulative number of abnormal points labeled across the 12 months of data.

The sub-figures demonstrate that the proposed procedure labels abnormal points at a similar rate for approximately the first 6 months of incoming data for both the healthy data set and the abnormal data set, regardless of the number of Gaussians used. However, it is worth noting that using 5 and 25 Gaussians labels more abnormal points throughout the entire year compared to the number of abnormal points labeled using a healthy data set.

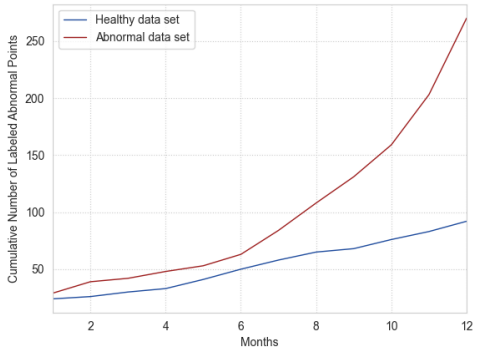
Figure 4.17 presents the results of the anomaly detection procedure after 1 month, 6 months, and 12 months of abnormal data using 1 Gaussian distribution to model the healthy distribution. The figures on the left show the data points labeled as abnormal by the proposed procedure, while the figures on the right provide a comparison by showing the actual abnormal points in the data set (i.e., the ground truth). For all of the monthly clustering results, please refer to Appendix A.



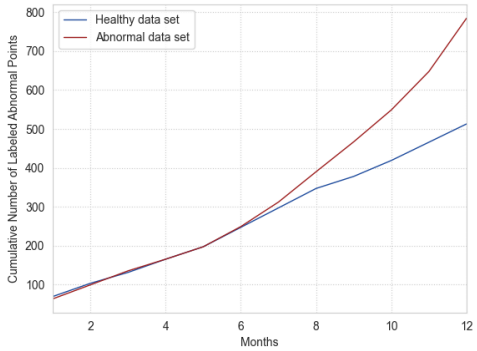
(a) Anomaly detection results using 1 Gaussian.



(b) Anomaly detection results using 5 Gaussians.

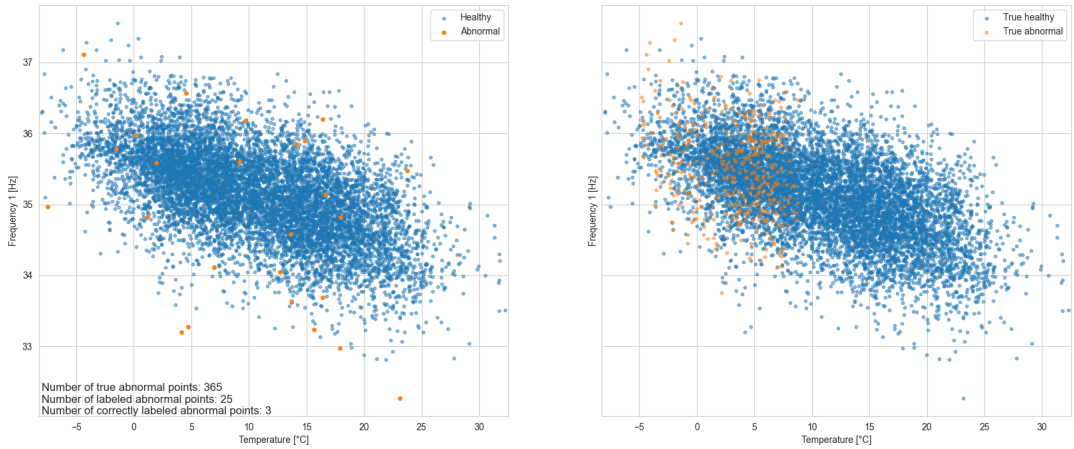


(c) Anomaly detection results using 25 Gaussians.

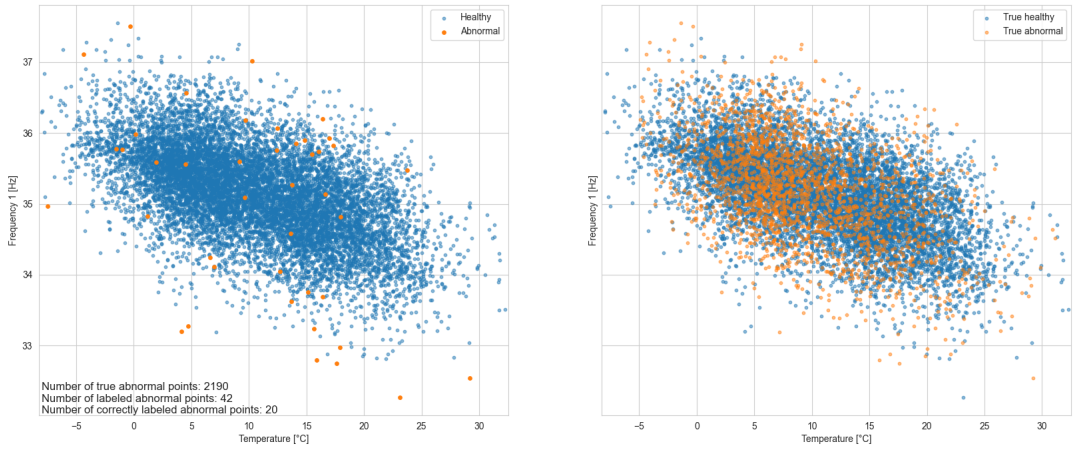


(d) Anomaly detection results using 125 Gaussians.

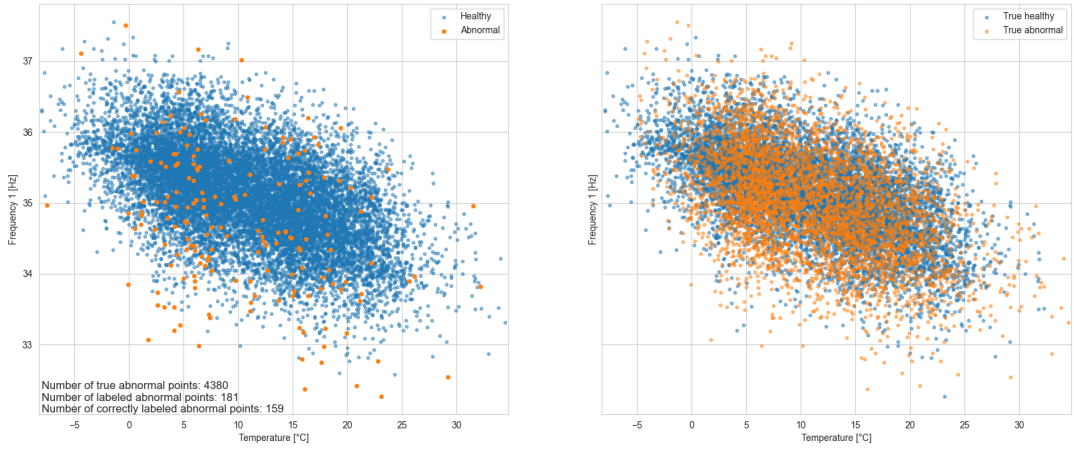
Figure 4.16: Anomaly detection results using different numbers of Gaussians.



(a) Anomaly detection results after 1 month of gradual settlement ($\delta = 41.67$ mm).



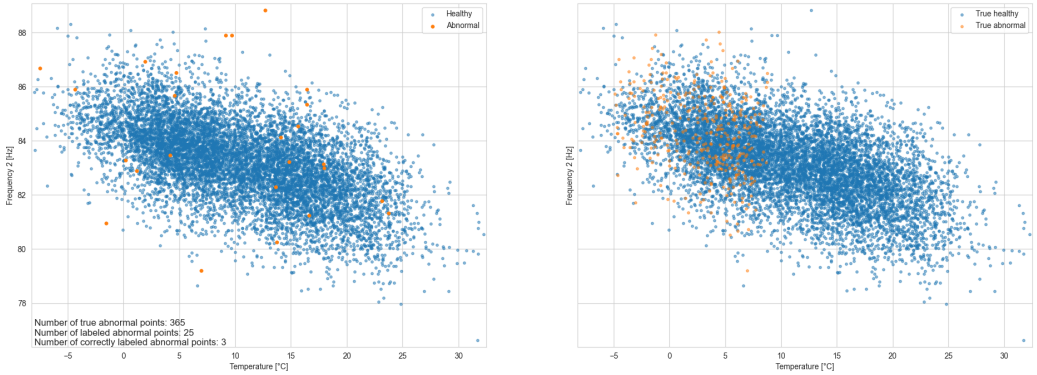
(b) Anomaly detection results after 6 months of gradual settlement ($\delta = 250$ mm).



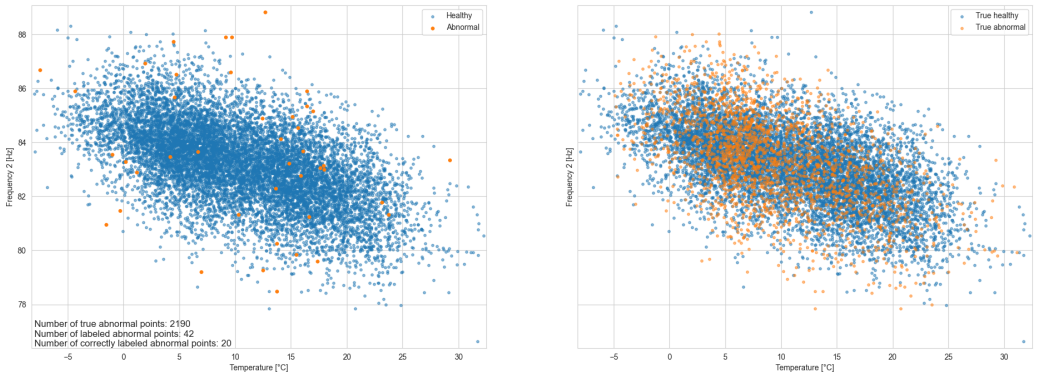
(c) Anomaly detection results after 12 months of gradual settlement ($\delta = 500$ mm).

Figure 4.17: Anomaly detection results for gradual settlement (1st frequency).

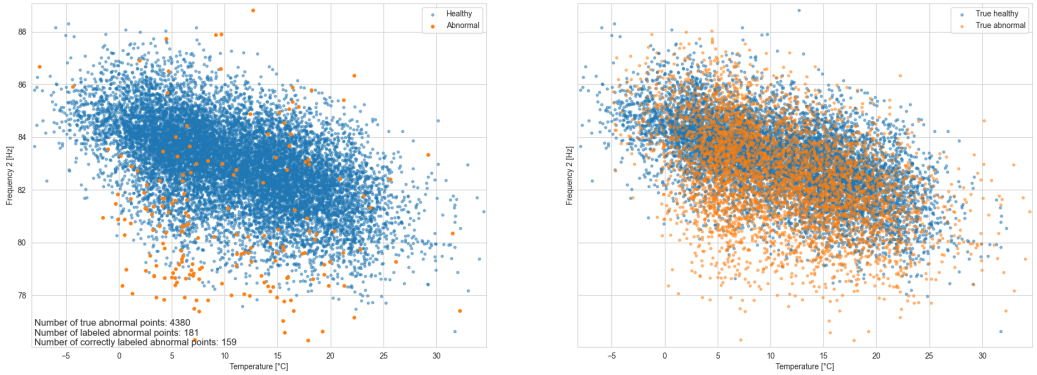
The mode shape for the second frequency (see Figure 3.4) suggests that this frequency would exhibit a higher sensitivity to a settlement in one of the piers. Consequently, Figure 4.18 offers a comparative perspective on the anomaly detection results previously discussed for the first frequency. Despite the different frequency under consideration, the number of labeled abnormal points remains consistent across all months. This consistency is a consequence of the anomaly detection process, which utilizes all five natural frequencies and treats each data point as a multivariate observation within the GMM. Therefore, this figure provides an alternative visualization of the labeled abnormal points, rather than a different analytical outcome. A closer look at the second frequency underscores its heightened sensitivity, as evident in the more pronounced deviation of the true abnormal points from the healthy distribution.



(a) Anomaly detection results after 1 month of gradual settlement ($\delta = 41.67$ mm).



(b) Anomaly detection results after 6 months of gradual settlement ($\delta = 250$ mm).



(c) Anomaly detection results after 12 months of gradual settlement ($\delta = 500$ mm).

Figure 4.18: Anomaly detection results for gradual settlement (2nd frequency).

The results seen in Figures 4.16 and 4.17 demonstrate how the GMM-based anomaly detection procedure can identify outliers in the initial months of new data, and as more months are included, additional outliers are detected at a more effective rate. This provides valuable insight to the user that there may be abnormal behavior in the modeled bridge. While the proposed anomaly detection may not correctly label every data point, it can effectively alert analysts of potential issues.

Table 4.1 shows the percentage of outliers that were correctly labeled as damaged by the GMM-based anomaly detection algorithm (for the case of using 1 Gaussian). The table includes the number of abnormal months, the corresponding true number of unhealthy points, and the percentage of

correctly labeled damaged points. As expected, the accuracy of the algorithm increases with the number of abnormal months, reaching a maximum of 3.63% when the settlement has reached its peak of 500 mm. It is worth noting that the first 6 months of abnormal data label a similar percentage of abnormal points correctly. From the 7th month onwards, the percentage of correctly labeled outliers increases each month, consistent with the results shown in Figure 4.16, where the rate of correctly labeling anomalies increases after the first six months.

Table 4.1: Percentage of abnormal points that are correctly labeled (for the case of using 1 Gaussian).

No. of Abnormal Months	True No. of Abnormal Points	Correctly Labeled as Damaged [%]
1	365	0.82
2	730	1.10
3	1095	0.73
4	1460	0.75
5	1825	0.77
6	2190	0.91
7	2555	0.98
8	2920	1.47
9	3285	1.83
10	3650	2.16
11	4015	2.79
12	4380	3.63

Revisiting Figure 4.16a, a baseline rate for labeling abnormal data can be established. This rate provides a standard for comparing the rate of abnormal point labeling. Figure 4.20 clearly identifies the moment at which abnormal behavior can be reliably detected. As observed in Figure 4.16a, the rate of abnormal point labeling stays constant for both data sets during the initial months. However, as time progresses, a notable increase in the anomaly detection rate for the abnormal data set becomes evident, indicating a point at which abnormal behavior can be confidently flagged.

In Figure 4.19, the first two months are used to establish a trend line representing the healthy data. This trend line encapsulates the expected rate of abnormal point labeling under normal conditions (cumulative). Thus, any significant deviation from this trend can be interpreted as a reliable indicator of abnormal behavior. In the specific context of the modeled bridge settlement, it can be confidently asserted that abnormal behavior becomes discernible approximately 6 months after the onset. This corresponds to a right pier settlement of 250 mm in the FEM modeled bridge.

Similarly, Figure 4.20 displays the same results, but the labeling in this figure is not cumulative. Moreover, a horizontal line representing a threshold is included. When the number of labeled points exceeds this line, it serves as an indicator of abnormal behavior. In this context, the threshold is determined as the average value of the number of labeled abnormal points during the anomaly detection process applied to a healthy data set. The healthy distribution is modeled after 1.5 years of monitoring, and the remaining 6 months of the second year are utilized for anomaly detection of the healthy data set, during which the average is calculated. It is important to note that this threshold can be defined differently based on specific requirements. The provided threshold serves as an illustrative example of what it could resemble.

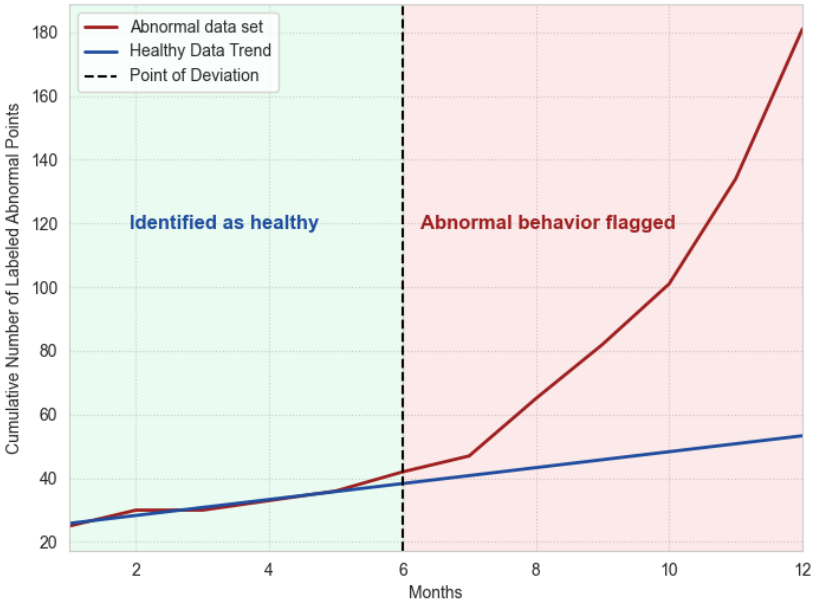


Figure 4.19: Flagging of abnormal behavior using the abnormal data set of 500 mm gradual settlement.

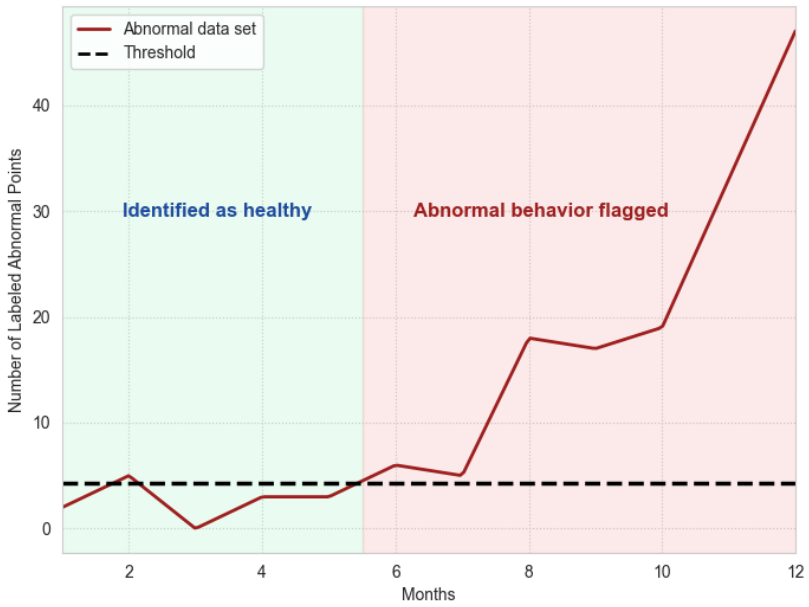


Figure 4.20: Flagging of abnormal behavior using the abnormal data set of 500 mm gradual settlement.

4.2. Case Study II: Stiffness Reduction

This case study examines the effectiveness of anomaly detection in identifying the loss of stiffness in the mid-span of a modelled bridge and the top region of each pier, where the bending moment is at its maximum. Such loss of stiffness, often due to cracking, is analyzed through three scenarios. These scenarios are as follows:

- Gradual 10% reduction in stiffness
- Sudden 10% reduction in stiffness
- Sudden 5% reduction in stiffness

All three scenarios involve the loss of stiffness in the members highlighted in red in Figure 4.21. However, unlike the figure, the mesh is refined in these cases to keep the loss of stiffness localized to a smaller element length.

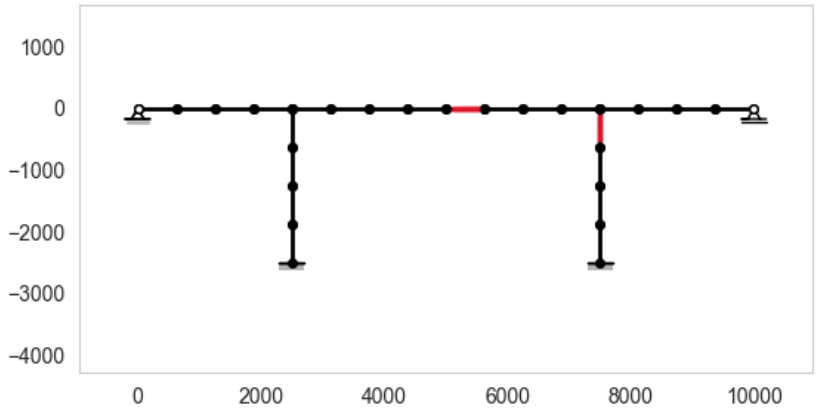


Figure 4.21: Loss of stiffness introduced in the red members.

4.2.1. Gradual 10% Loss in Stiffness

This scenario of the case study will incorporate a gradual 10% reduction in stiffness in two locations inside the model. This loss in stiffness takes place over the entirety of the year 2019 (the final and 3rd year) and the reduction in stiffness decreases linearly over this time.

Relationship Between Temperature and Frequency

The current case study follows the same approach as the previous one, where natural frequencies are computed over time using temperature data from three years. However, in this case, a linear and gradual decrease in stiffness occurs in certain members during the third year. The graph in Figure 4.22 displays the changes in the first natural frequency, f_1, over the same three-year period, while also showing the corresponding temperature variations from 2017 to 2019. Despite the gradual loss of stiffness in the specified members, there is no significant or noticeable alteration in the first frequency that can be observed from the beginning of 2019.

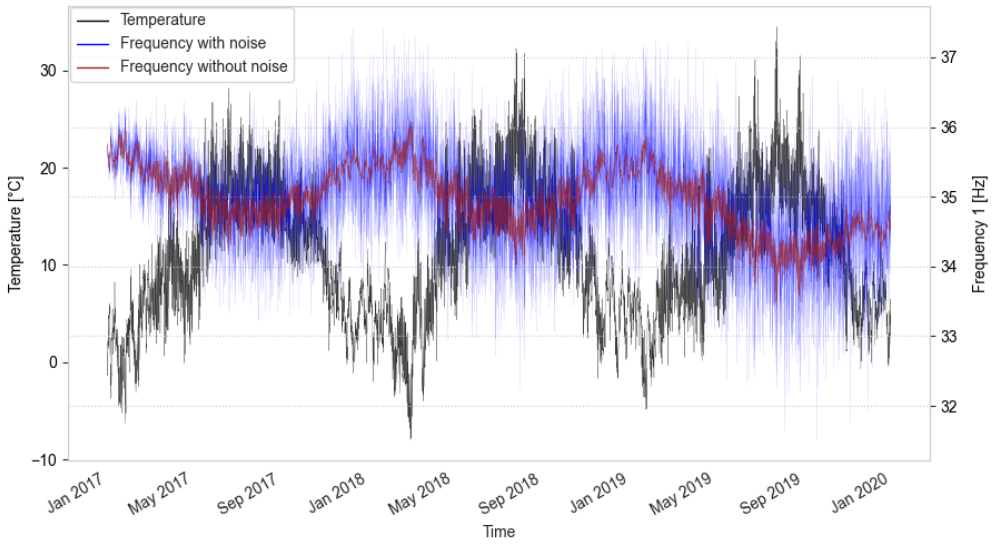


Figure 4.22: Frequency and temperature plotted over time.

A data set is created from the natural frequency calculations, similar to the one presented in Table 3.1. Two- and three-dimensional plots of the data set are shown in Figures 4.23 and 4.24, respectively. However, these plots do not reveal any obvious anomalies in the data. Based on the absence of noticeable irregularities in the data set, it is recommended to perform the same anomaly detection procedure used in the previous case study. Although the possibility of using PCA was highlighted in the previous case study, it was shown that it will not prove advantageous in this case. Therefore, GMM-based anomaly detection will be performed directly in this case study.

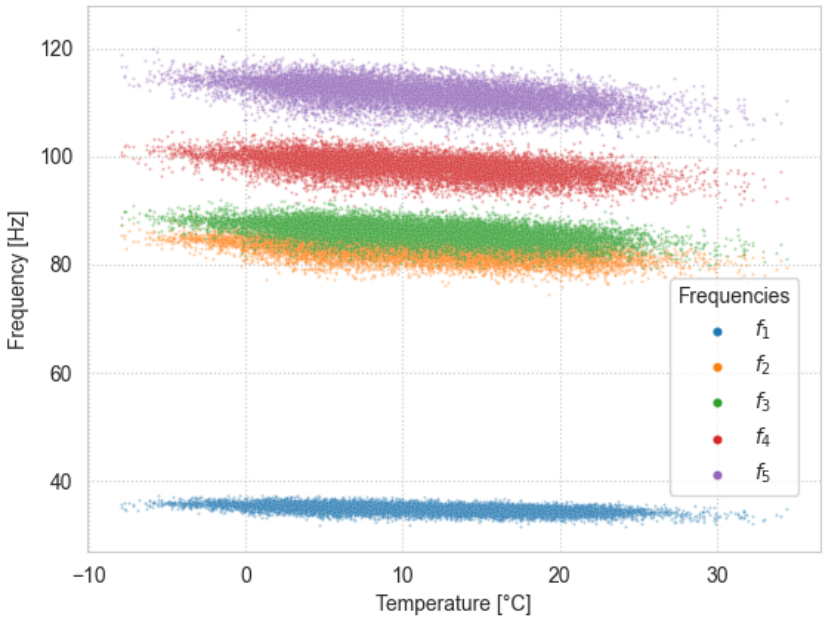


Figure 4.23: Frequency plotted against the temperature.

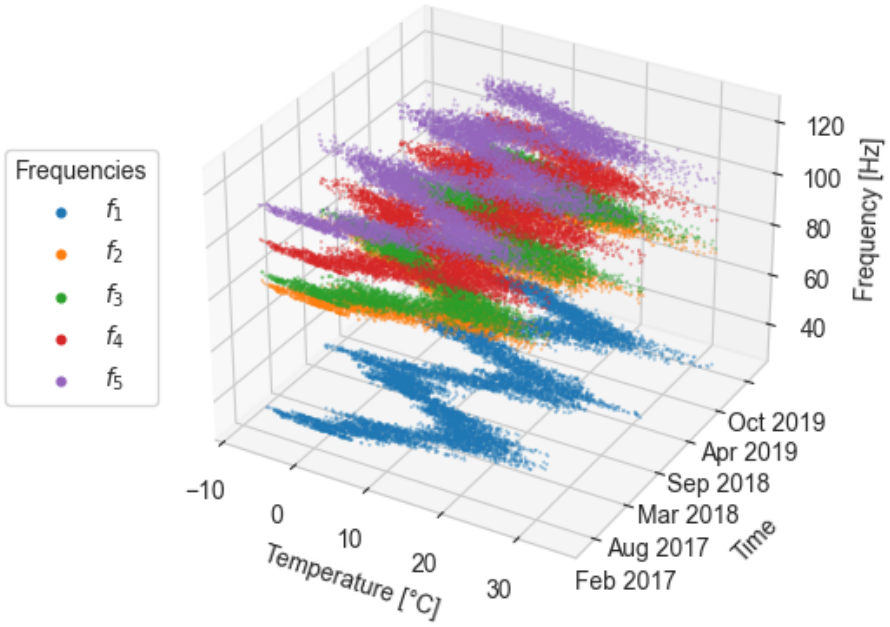


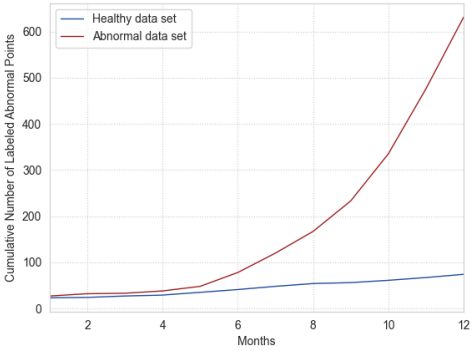
Figure 4.24: Frequency plotted against temperature and time.

Clustering Results

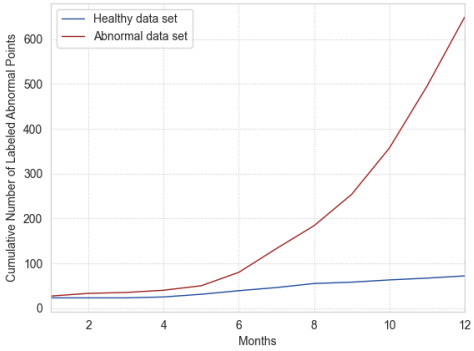
For Case Study I, the distribution of healthy data points was modeled using GMM with 1, 5, 25, and 125 Gaussian distributions. The same procedure is followed for the current case of gradual stiffness loss in two members of the FEM modeled bridge, where the proposed GMM-based anomaly detection algorithm yields promising results. Figure 4.25 compares the effectiveness of the anomaly detection procedure when using different numbers of Gaussians to model the distribution of healthy data points. Each sub-figure displays the results of performing anomaly detection using one year of additional healthy data points versus using a year's worth of anomalous data points.

The sub-figures demonstrate that the proposed procedure identifies abnormal data points at a consistent rate for the first 3-4 months of incoming data for both the healthy and abnormal data sets, regardless of the number of Gaussians used. However, after three months of receiving abnormal data, the rate of labeling abnormal points increases significantly for all variations of the number of Gaussians used. The mislabeling of abnormal points using a purely healthy data set remains at a steady and constant rate throughout the 12 months of anomaly detection.

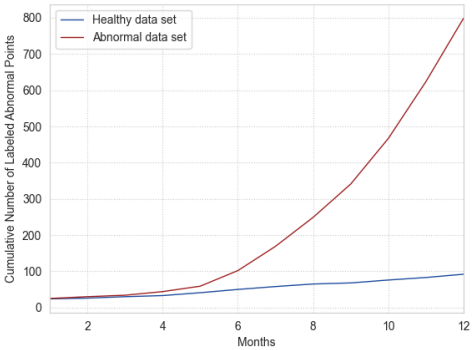
Figure 4.26 illustrates the results of the anomaly detection procedure after 1 month, 6 months, and 12 months of abnormal data using 1 Gaussian distribution to model the healthy distribution. The figures on the left display the data points labeled as abnormal by the proposed procedure, while those on the right provide a comparison by showing the actual abnormal points in the data set (i.e., the ground truth). At the end of 12 months of anomaly detection, the algorithm correctly labels a total of 609 anomalous data points out of the 4380 present.



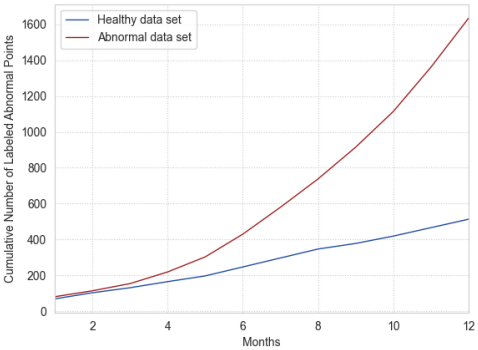
(a) Anomaly detection results using 1 Gaussian.



(b) Anomaly detection results using 5 Gaussians.

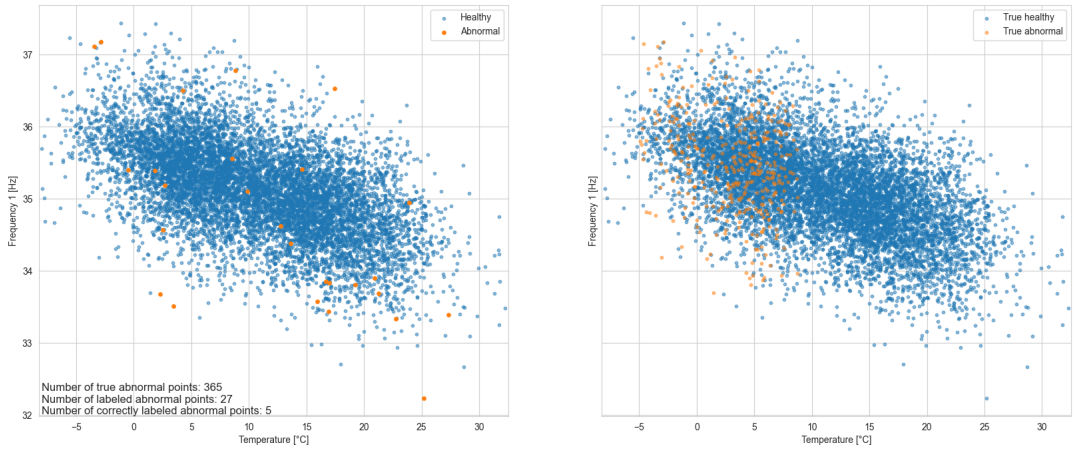


(c) Anomaly detection results using 25 Gaussians.

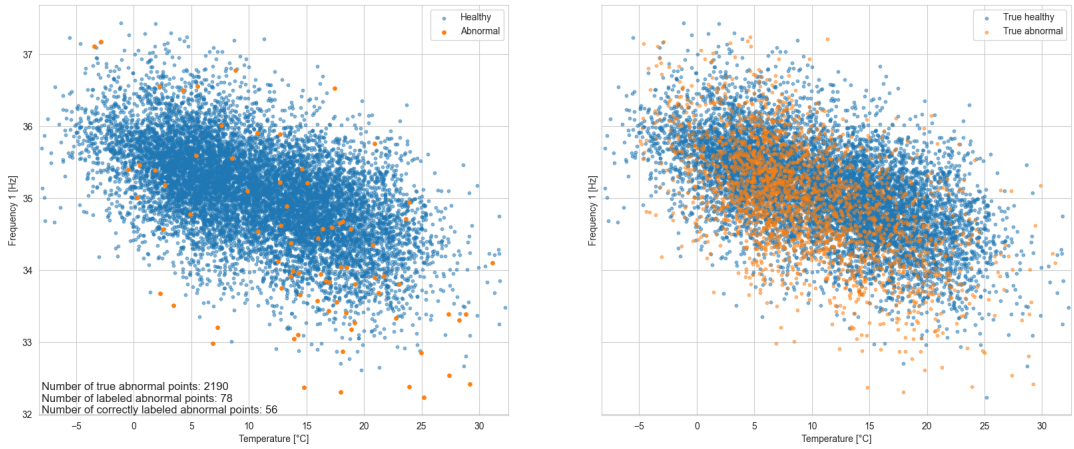


(d) Anomaly detection results using 125 Gaussians.

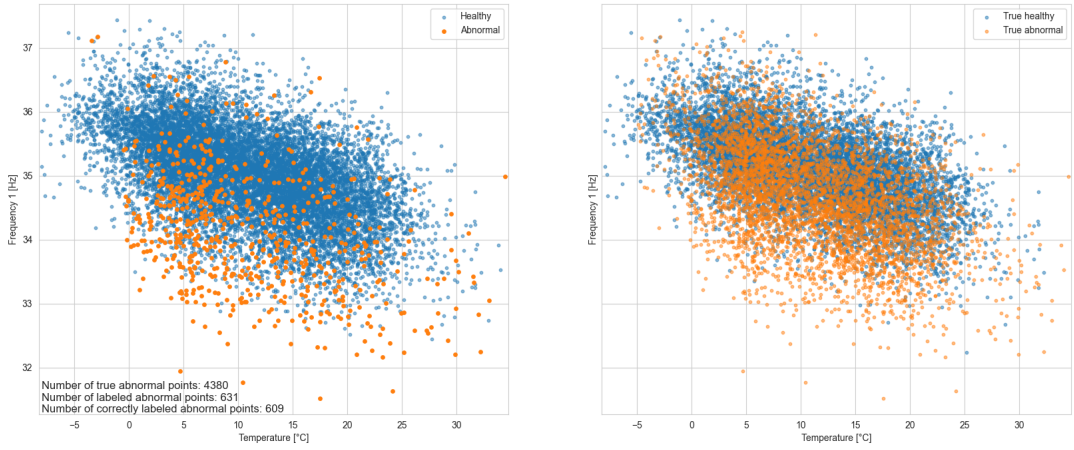
Figure 4.25: Anomaly detection results using different numbers of Gaussians.



(a) Anomaly detection results after 1 month of gradual stiffness reduction.



(b) Anomaly detection results after 6 months of gradual stiffness reduction.



(c) Anomaly detection results after 12 months of gradual stiffness reduction.

Figure 4.26: Anomaly detection results for gradual stiffness loss.

Results obtained in this current case study, as seen in Figures 4.25 and 4.26, are similar to those of the first case study, demonstrating how the GMM-based anomaly detection procedure can identify outliers in the initial months of new data. As additional months are included, the algorithm detects additional outliers at a more effective rate, providing valuable insight to the user of potential abnormal behavior in the modeled bridge.

Table 4.2 displays the percentage of outliers correctly labeled as damaged by the GMM-based anomaly detection algorithm (when using 1 Gaussian). The table presents the number of abnormal months, the corresponding true number of unhealthy points, and the percentage of correctly labeled damaged points. As expected, the algorithm's accuracy increases with the number of abnormal months, reaching a maximum of 13.90% when the stiffness loss reaches its peak of 10%. Notably, the first 4-5 months of abnormal data label a similar percentage of abnormal points correctly. From the 5-6th month onwards, the percentage of correctly labeled outliers increases each month, consistent with the results shown in Figure 4.25, where the rate of correctly labeling anomalies increases after the first 3-4 months.

Table 4.2: Percentage of abnormal points that are correctly labeled (for the case of using 1 Gaussian).

No. of Abnormal Months	True No. of Abnormal Points	Correctly Labeled as Damaged [%]
1	365	1.37
2	730	1.37
3	1095	1.00
4	1460	1.10
5	1825	1.42
6	2190	2.56
7	2555	3.84
8	2920	4.97
9	3285	6.42
10	3650	8.58
11	4015	11.31
12	4380	13.90

To conclude the case study of gradual stiffness loss, it is found that abnormal behavior can be identified approximately 5 months after the onset of stiffness reduction (Figure 4.27). This corresponds to a 4.2% loss of stiffness in the two selected members of the FEM modeled bridge when abnormal behavior is detected. Similar to the previous case study involving settlement of the right pier, the first two months are used to establish a healthy data trend line, which serves as a basis for comparing the rate of abnormal point labeling to that of the trend line. Additionally, Figure 4.28 shows the same results, albeit with non-cumulative labeling. In this figure, the same threshold mentioned in Case Study I is utilized. This threshold allows for the identification of abnormal behavior at a similar time frame, typically around 4-5 months after anomalies are introduced into the data set.

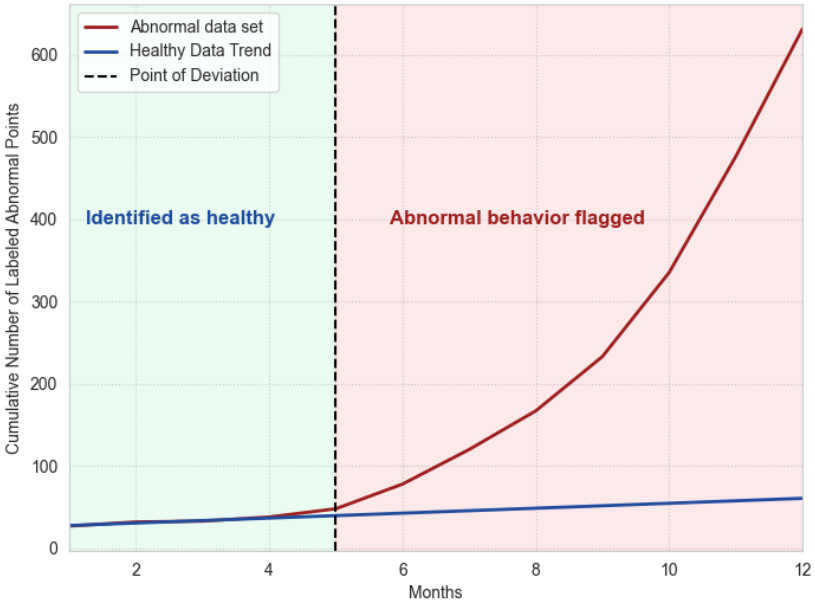


Figure 4.27: Flagging of abnormal behavior using the abnormal data set of 10% gradual stiffness loss.

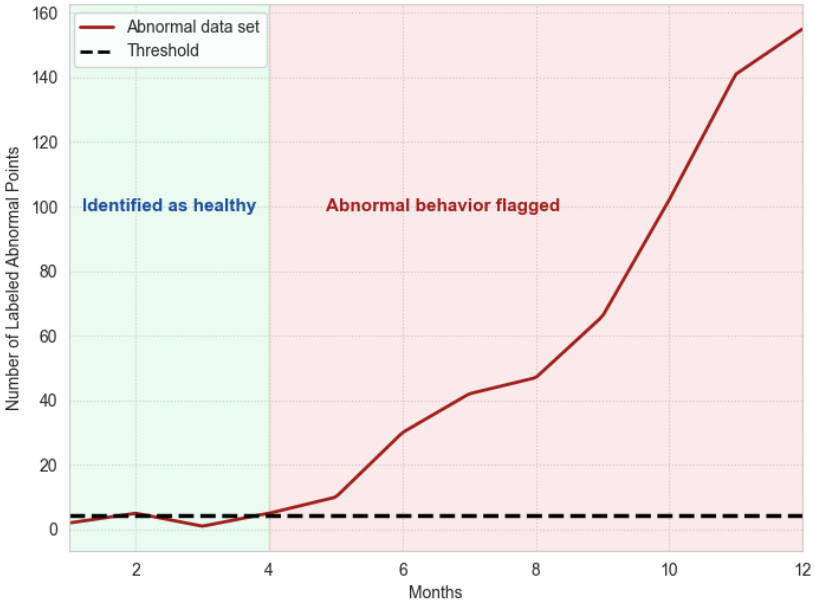


Figure 4.28: Flagging of abnormal behavior using the abnormal data set of 10% gradual stiffness loss.

4.2.2. Sudden 10% Loss in Stiffness

In contrast to Case Study I and the scenario in this case study with gradual stiffness loss, this scenario involves sudden damage occurring at the onset of the third year. Despite this difference, the same methodology is employed, with the only variation being the damage type.

Relationship Between Temperature and Frequency

In Figure 4.29, one can observe a clear shift in frequency occurring at the beginning of the third year. This shift indicates that an abnormal event has taken place and affected the system.

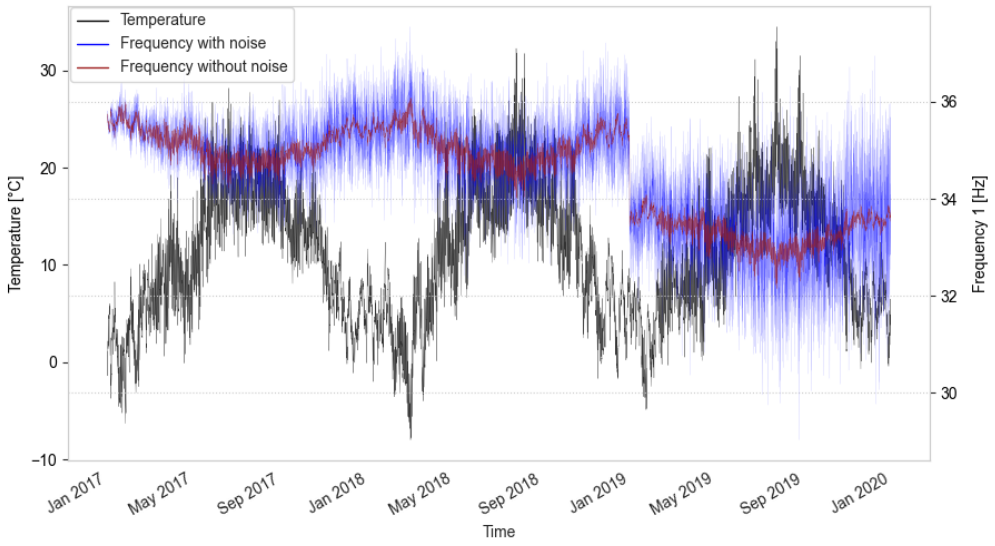


Figure 4.29: Frequency and temperature plotted over time.

While observing the frequency shift in Figures 4.29 and 4.30, it is evident that the analyzed bridge requires inspection and investigation. The shift in frequencies clearly indicates that some event has taken place, causing damage. However, when plotting the relationship between temperature and natural frequencies in three dimensions, as seen in Figure 4.31, this shift in natural frequencies is not easily observed. The sudden loss of stiffness resulting in a shift in natural frequencies raises the question of whether clustering is necessary. However, it is of interest to put the developed anomaly detection algorithm to the test to see its performance in identifying and flagging abnormalities in such a scenario.

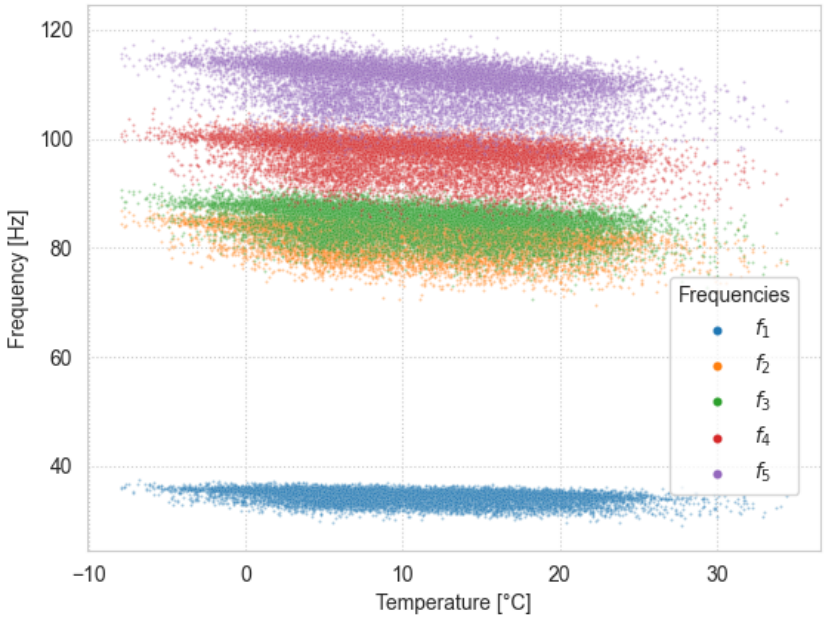


Figure 4.30: Frequency plotted against the temperature.

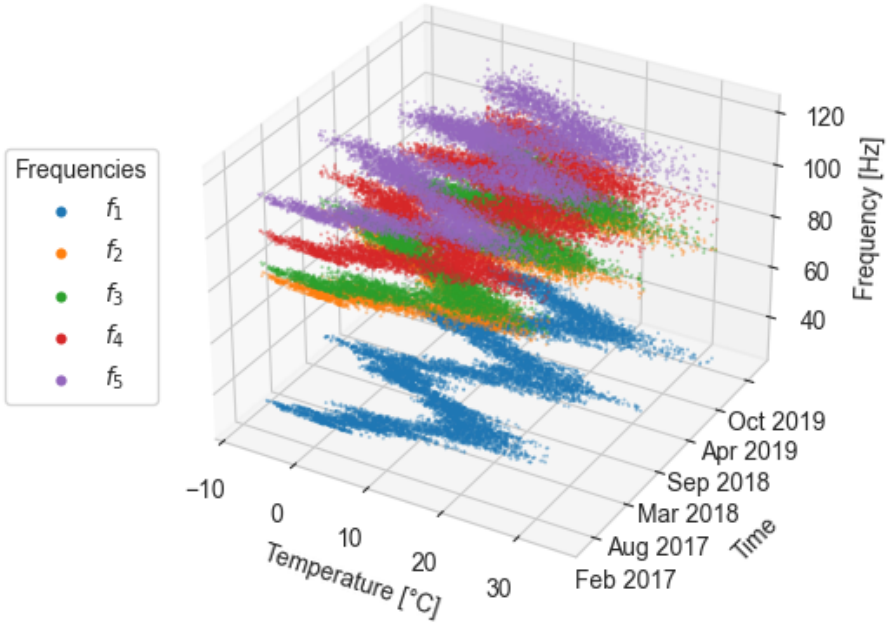


Figure 4.31: Frequency plotted against temperature and time.

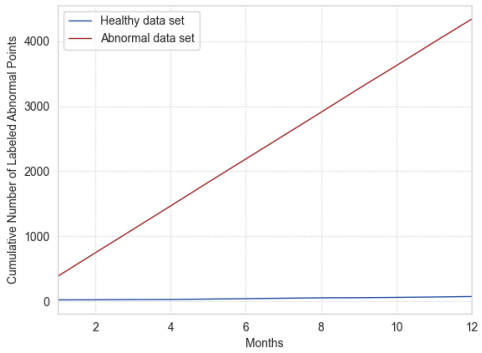
Clustering Results

From Figure 4.30, shifts in all frequencies are clearly visible. As such, it is expected that performing the GMM-based anomaly detection, which essentially defines a boundary and threshold for the healthy data distribution, will be able to identify most of the points that have shifted outside of the healthy distribution and label them as abnormal. This is indeed the case, as is shown in both Figures 4.32 and 4.33. Similar to before, Figure 4.32 illustrates the effectiveness of the anomaly detection procedure when using different numbers of Gaussians to model the distribution of healthy data points. Each

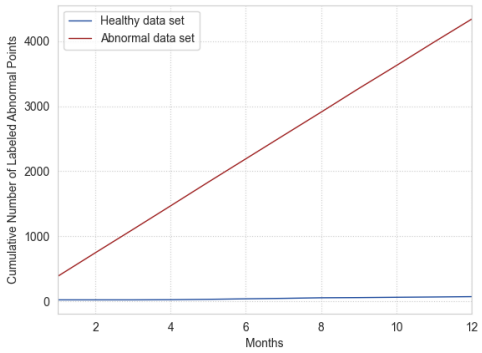
sub-figure shows the results of performing anomaly detection using one year of additional healthy data points versus using a year's worth of anomalous data points.

Anomalies are instantaneously detected and labeled at a significantly higher rate from the moment that the sudden loss of stiffness occurs, compared to when a healthy data set is put through anomaly detection. This rate of detecting abnormal points remains steady throughout the entire 12 months. Additionally, it is observed that there is no benefit in increasing the number of Gaussians used to model the healthy data point distribution. For this damage case, it is evident that using one Gaussian is sufficient, as there are minimal gains available when modeling the healthy distribution with more than one Gaussian distribution.

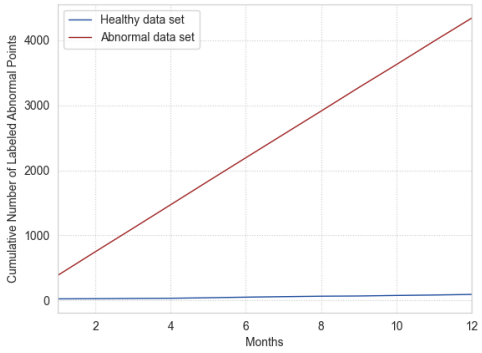
Figure 4.33 presents the results of the anomaly detection procedure after 1 month, 6 months, and 12 months of abnormal data using 1 Gaussian distribution to model the healthy distribution. The figures on the left show the data points labeled as abnormal by the proposed procedure, while those on the right provide a comparison by displaying the actual abnormal points in the data set. For each month, the labeled abnormal points are almost identical to the ground truth. This figure demonstrates the algorithm's exceptional performance in detecting anomalies for the damage scenario of 10% sudden stiffness reduction.



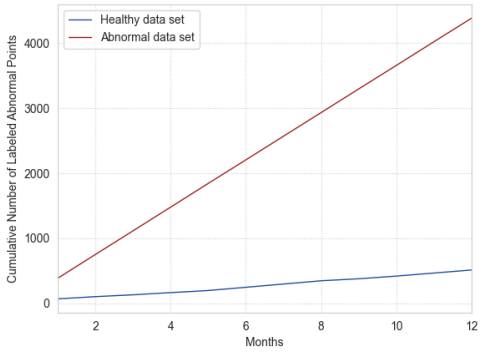
(a) Anomaly detection results using 1 Gaussian.



(b) Anomaly detection results using 5 Gaussians.

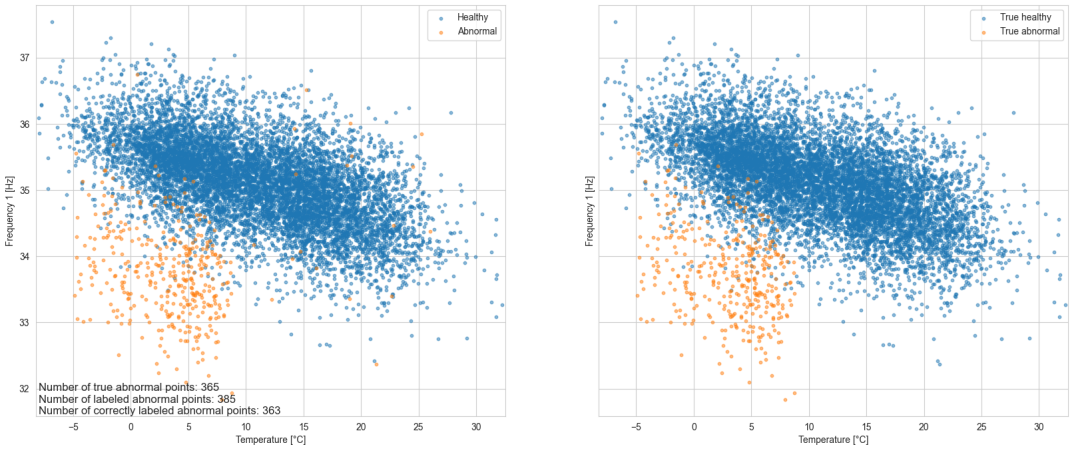


(c) Anomaly detection results using 25 Gaussians.

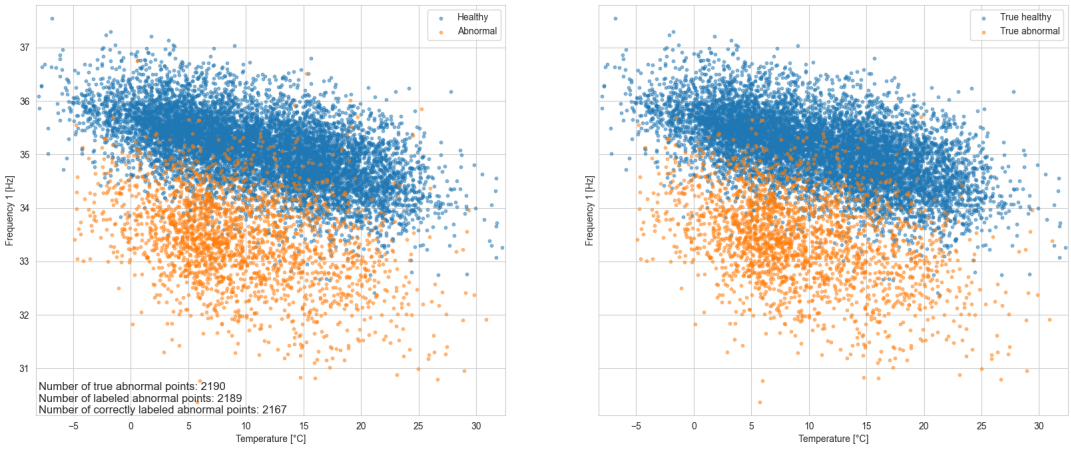


(d) Anomaly detection results using 125 Gaussians.

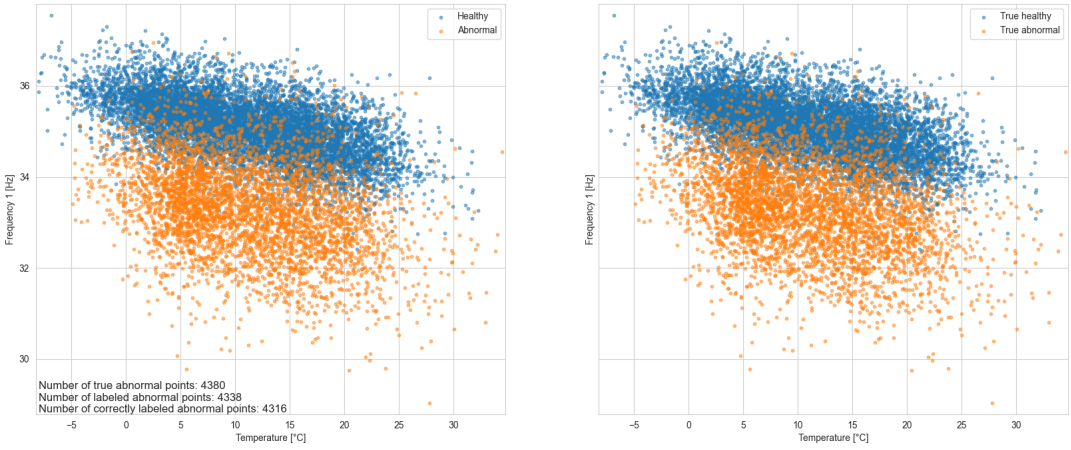
Figure 4.32: Anomaly detection results using different numbers of Gaussians.



(a) Anomaly detection results after 1 month of 10% sudden stiffness reduction occurring.



(b) Anomaly detection results after 6 months of 10% sudden stiffness reduction occurring.



(c) Anomaly detection results after 12 months of 10% sudden stiffness reduction occurring.

Figure 4.33: Anomaly detection results for sudden 10% stiffness loss.

To further validate the strength of the anomaly detection for this damage scenario, Table 4.3 shows the percentage of correctly labeled damaged points. From the 1st month to the 12th month, the success rate of accurately labeling abnormal points is consistently above 98.50%. This demonstrates the high level of accuracy and reliability of the proposed anomaly detection algorithm in detecting sudden stiffness reduction of 10% for the FEM modeled bridge.

Table 4.3: Percentage of abnormal points that are correctly labeled (for the case of using 1 Gaussian).

No. of Abnormal Months	True No. of Abnormal Points	Correctly Labeled as Damaged [%]
1	365	99.45
2	730	99.32
3	1095	99.00
4	1460	98.97
5	1825	99.07
6	2190	98.95
7	2555	98.86
8	2920	98.80
9	3285	98.87
10	3650	98.68
11	4015	98.68
12	4380	98.54

4.2.3. Sudden 5% Loss in Stiffness

To further evaluate the capability of the GMM-based anomaly detection method, a sudden 5% loss in stiffness is also examined. This scenario is more subtle compared to the sudden 10% loss of stiffness and therefore will be more interesting to determine if the method performs as well as before.

Relationship Between Temperature and Frequency

When plotting the relationship between the temperature and the natural frequencies, a clear difference is observed when the sudden stiffness reduction amounts to 10%. However, for the case of 5% sudden stiffness reduction, it is more challenging to discern any apparent shift in the natural frequencies, as seen in Figures 4.34 to 4.36. Therefore, it is recommended to perform the GMM-based anomaly detection.

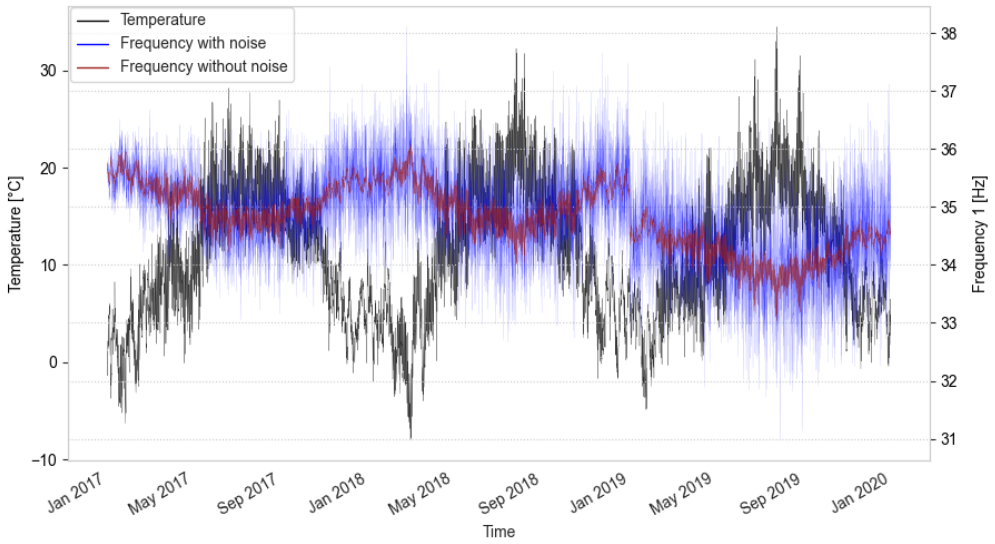


Figure 4.34: Frequency and temperature plotted over time.

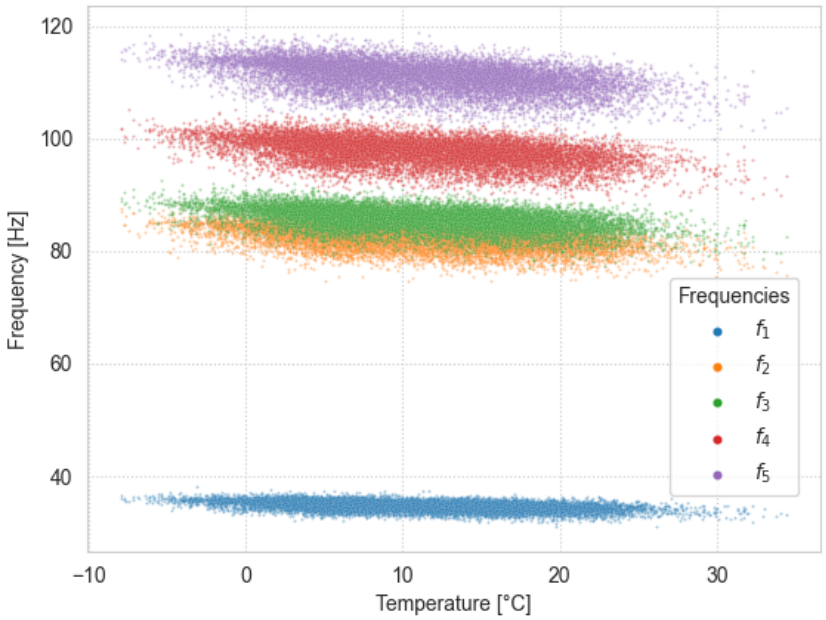


Figure 4.35: Frequency plotted against the temperature.

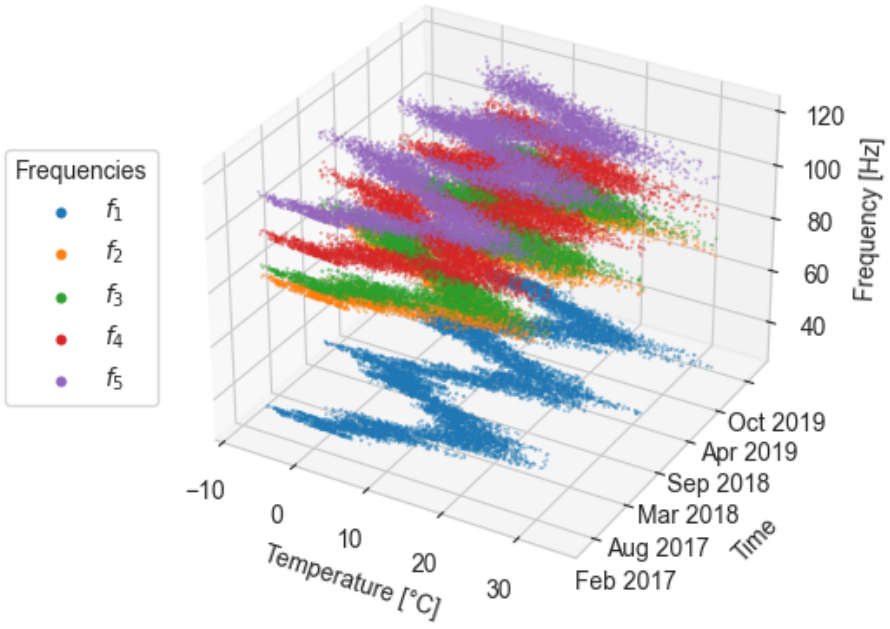
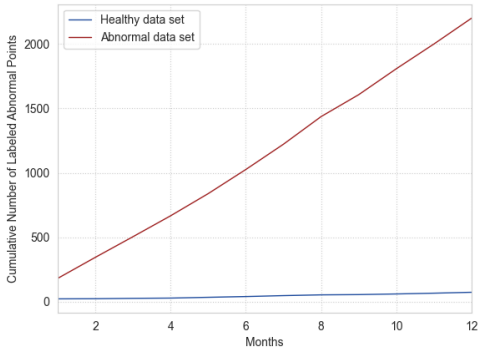


Figure 4.36: Frequency plotted against temperature and time.

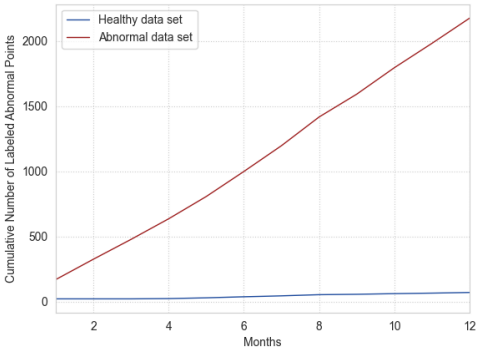
Clustering Results

Performing anomaly detection for the damage scenario of 5% sudden reduction in stiffness of two members in the FEM modeled bridge yields promising results. Figure 4.37 demonstrates that for 1, 5, 25, and 125 Gaussians used to model the healthy distribution, labeling of abnormal points is highly successful. As in previous cases, the figure also compares the results from performing the anomaly detection on a healthy data set instead of an abnormal one. The rate at which abnormal points are labeled compared to that of using a healthy data set is significantly higher, indicating the effectiveness of the anomaly detection. Anomalies are instantaneously detected and labeled from the exact moment that the sudden stiffness reduction takes place. Similar to the more severe case of 10% sudden stiffness loss, using a single Gaussian to model the distribution of the healthy data points is sufficient, and employing more Gaussians is unnecessary.

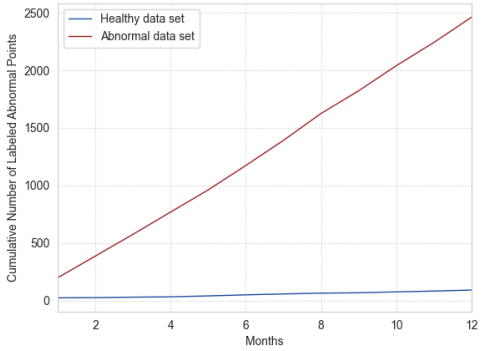
Figure 4.38 depicts the results of the anomaly detection procedure after 1 month, 6 months, and 12 months of abnormal data using 1 Gaussian distribution to model the healthy distribution. The figures on the left display the data points labeled as abnormal by the proposed procedure, while those on the right provide a comparison by showing the actual abnormal points in the data set. For each month, the labeled abnormal points exhibit the same shape as the ground truth, although in a less dense manner. This figure highlights the impressive performance of the algorithm in detecting anomalies for the damage scenario of 5% sudden stiffness reduction.



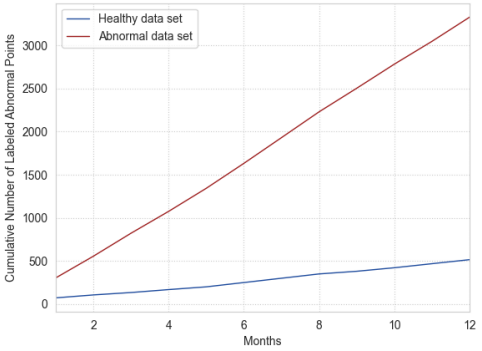
(a) Anomaly detection results using 1 Gaussian.



(b) Anomaly detection results using 5 Gaussians.

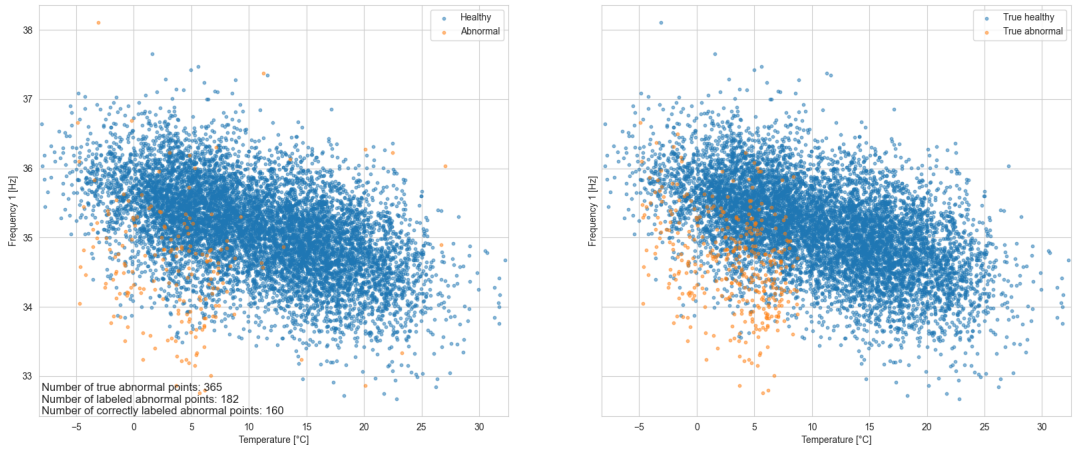


(c) Anomaly detection results using 25 Gaussians.

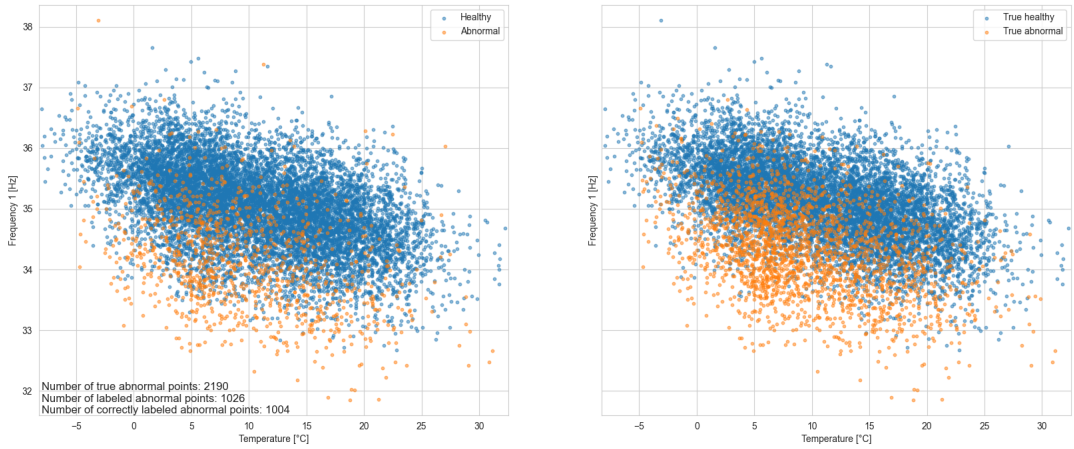


(d) Anomaly detection results using 125 Gaussians.

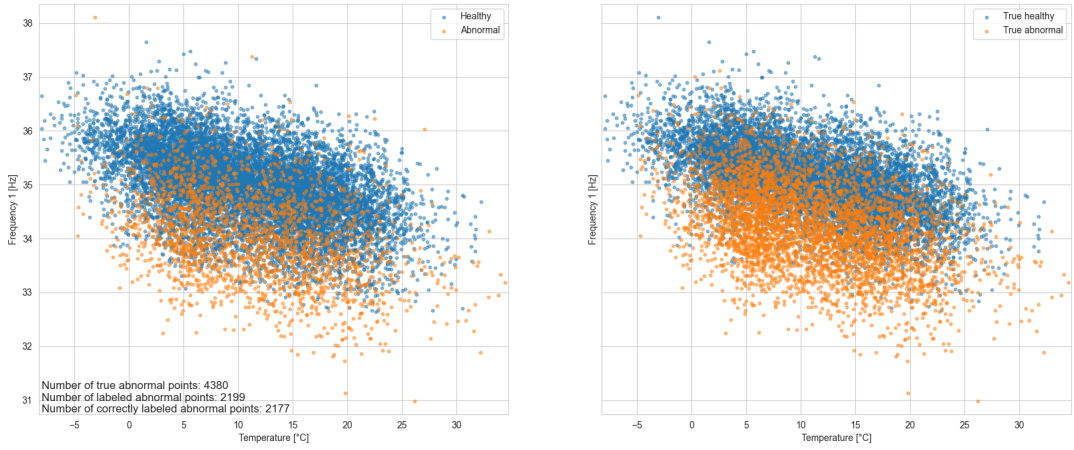
Figure 4.37: Anomaly detection results using different numbers of Gaussians.



(a) Anomaly detection results after 1 month of 5% sudden stiffness reduction occurring.



(b) Anomaly detection results after 6 months of 5% sudden stiffness reduction occurring.



(c) Anomaly detection results after 12 months of 5% sudden stiffness reduction occurring.

Figure 4.38: Anomaly detection results for sudden 5% stiffness loss.

Table 4.4 presents the percentage of correctly labeled damaged points for the sudden 5% stiffness reduction scenario. The proposed anomaly detection algorithm consistently labels abnormal points with high accuracy, ranging from 43.84 to 49.70% of abnormal points being correctly labeled throughout the 12 months. These results demonstrate the reliability and effectiveness of the algorithm in detecting sudden stiffness reduction of 5% for the FEM modeled bridge.

Table 4.4: Percentage of abnormal points that are correctly labeled (for the case of using 1 Gaussian).

No. of Abnormal Months	True No. of Abnormal Points	Correctly Labeled as Damaged [%]
1	365	43.84
2	730	44.25
3	1095	44.11
4	1460	44.18
5	1825	44.77
6	2190	45.84
7	2555	46.97
8	2920	48.42
9	3285	48.22
10	3650	48.90
11	4015	49.24
12	4380	49.70

4.3. Case Study III: Stiffness Reduction & Settlement

The last of the damage scenario to be examined involves a combination of the first two case studies. Specifically, it entails introducing settlement to one of the supports while decreasing the stiffness of two elements in the FEM model. However, this case study considers a less severe level of settlement and stiffness reduction. It will be divided into two damage scenarios:

- Combination of settlement ($\delta = 150$ mm) and gradual 3% reduction in stiffness
- Combination of settlement ($\delta = 150$ mm) and sudden 3% reduction in stiffness

Figure 4.39 highlights the elements where a reduction of the stiffness takes place as well as showing the geometry of the bridge after the settlement reaches its maximum value of 150 mm.

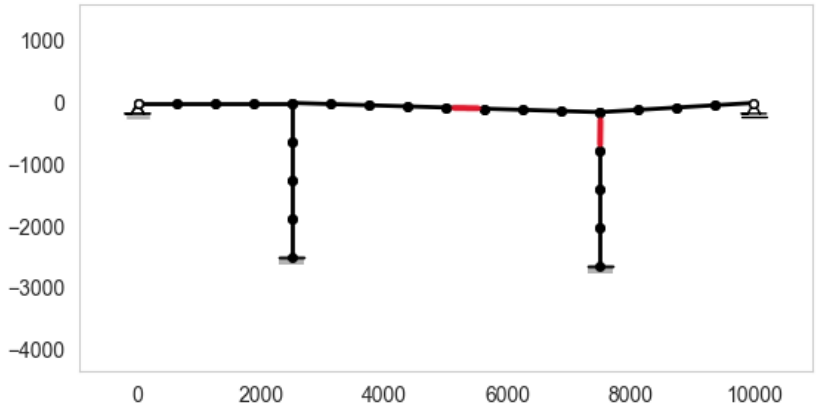


Figure 4.39: Geometry of the bridge after a maximum settlement of 150 mm. The red elements indicate that they are subject to a loss of stiffness.

4.3.1. Combination of Settlement and Gradual 3% Loss in Stiffness

The case study will analyze the gradual stiffness loss of the two highlighted members shown in Figure 4.39, which totals 3%, as well as the gradual settlement of one of the supports occurring at the same rate as the stiffness reduction. It is anticipated that the presence of two concurrent damage scenarios will assist the anomaly detection method in identifying abnormal data points more rapidly.

Relationship Between Temperature and Frequency

Based on the graphs presented in Figures 4.40-4.42, the FEM modeled bridge seems to be in good health throughout the last year of data. However, the absence of visible damage does not rule out the possibility of hidden abnormalities. Therefore, using the GMM-based anomaly detection method is recommended for this damage scenario.

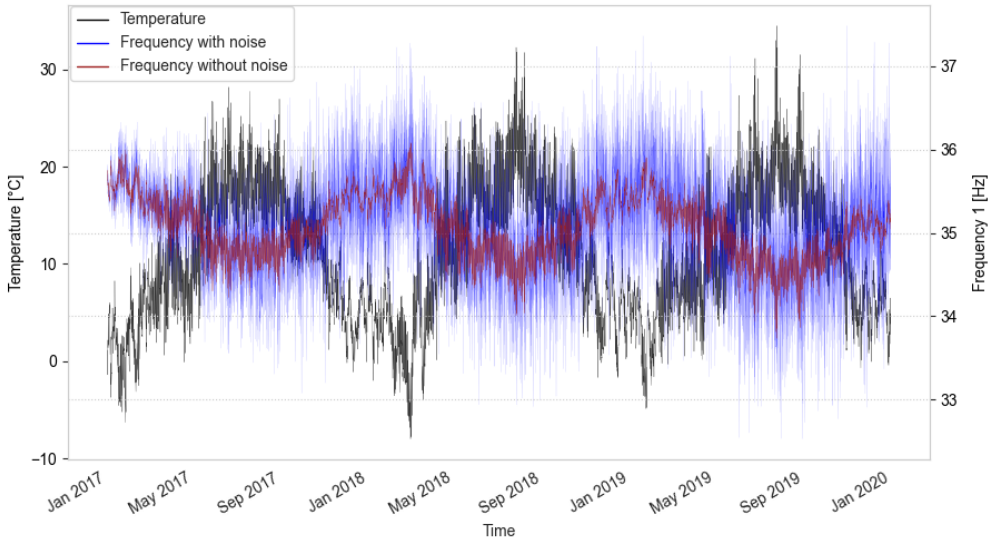


Figure 4.40: Frequency and temperature plotted over time.

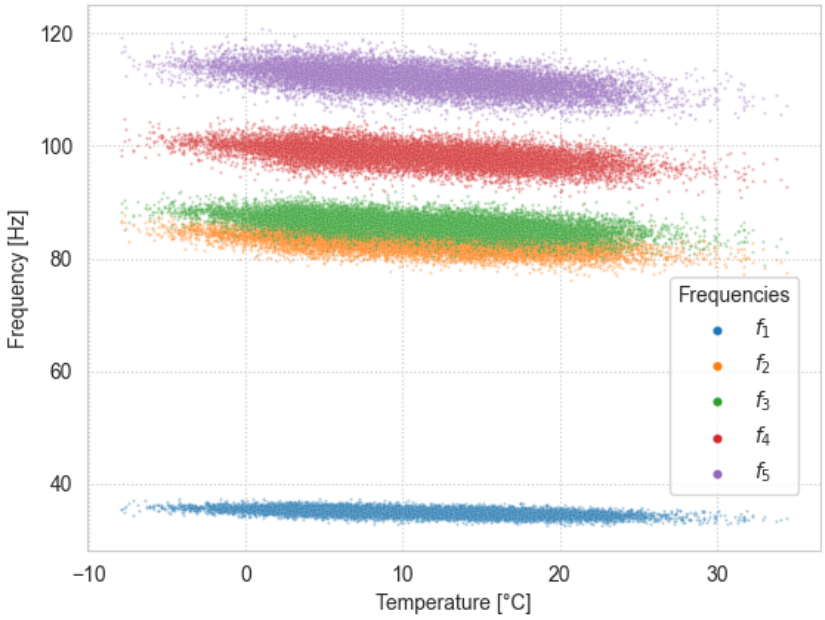


Figure 4.41: Frequency plotted against the temperature.

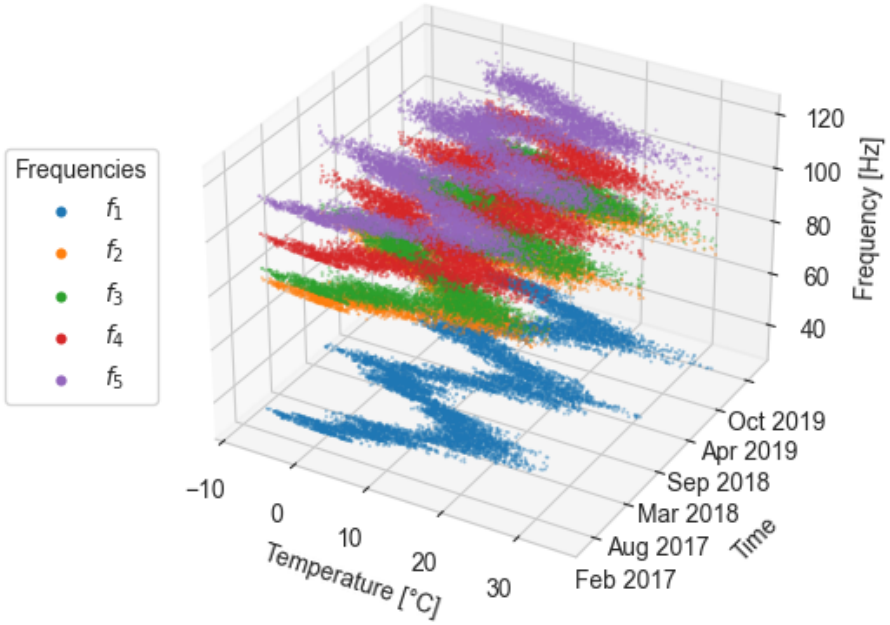


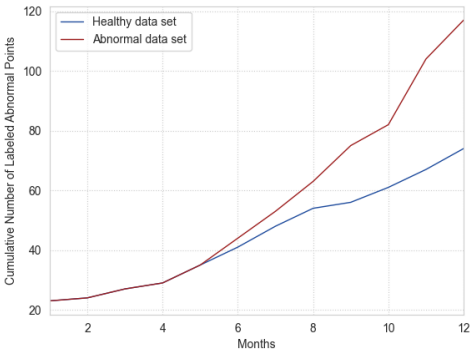
Figure 4.42: Frequency plotted against temperature and time.

Clustering Results

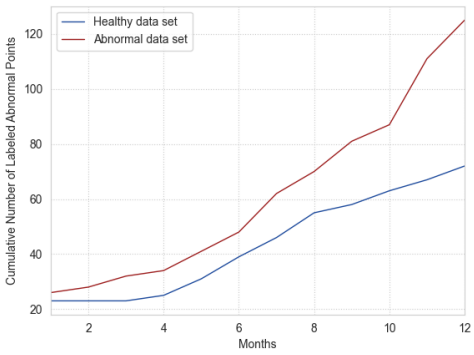
After performing anomaly detection for the gradual settlement and stiffness reduction scenario using different numbers of Gaussians, Figure 4.43 shows that abnormal data points are initially labeled at the same rate as when using a healthy data set for approximately the first 3.5 months of incoming data. After this point, the abnormal data set starts to deviate from the healthy data set and abnormal points are labeled at a higher rate. At this point, settlement has reached approximately 45 mm and the stiffness of the two selected members has reduced by 0.875%. Increasing the number of Gaussians

from 1 to 5, as shown in Figure 4.16b, can lead to labeling more abnormal points than when using only 1 Gaussian for the first 3.5 months. However, it is debatable whether this provides any advantage.

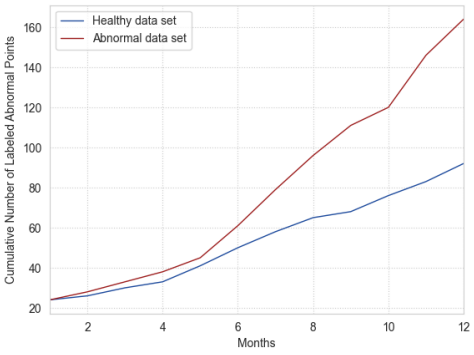
The effectiveness of the GMM-based anomaly detection algorithm for the damage scenario involving settlement and gradual stiffness reduction is demonstrated in Figure 4.44. The figure presents the results of the algorithm after 1 month, 6 months, and 12 months of abnormal data using 1 Gaussian distribution to model the healthy data set. The figures on the left depict the data points labeled as abnormal by the algorithm, while those on the right show the actual abnormal points in the data set. As the number of abnormal months increases, the number of labeled abnormal points also increases, but the increase is not as significant and noticeable as in the damage scenarios involving sudden stiffness loss.



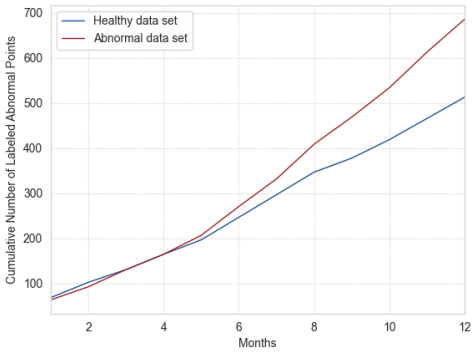
(a) Anomaly detection results using 1 Gaussian.



(b) Anomaly detection results using 5 Gaussians.

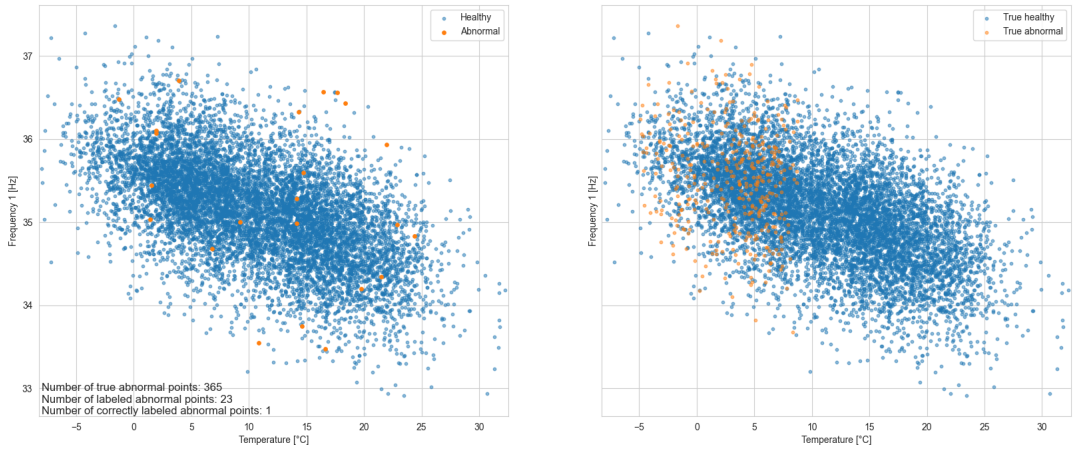


(c) Anomaly detection results using 25 Gaussians.

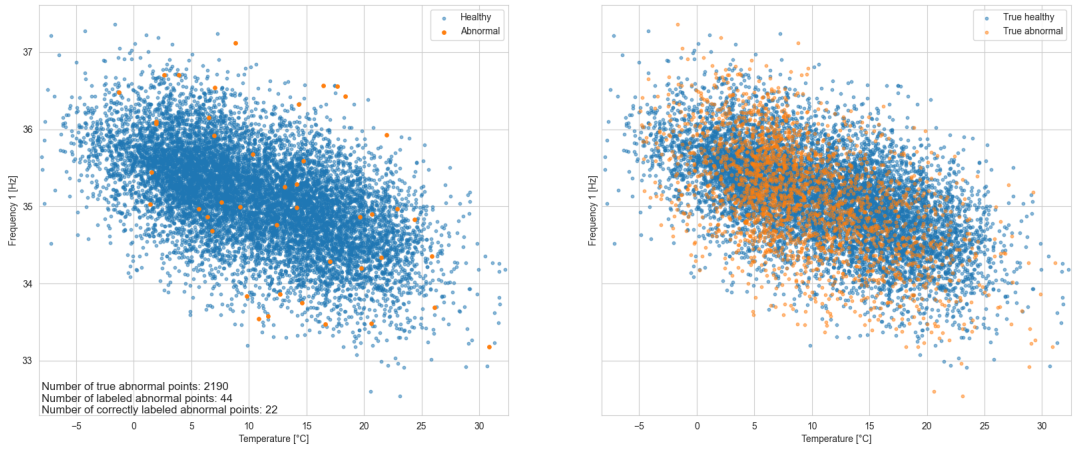


(d) Anomaly detection results using 125 Gaussians.

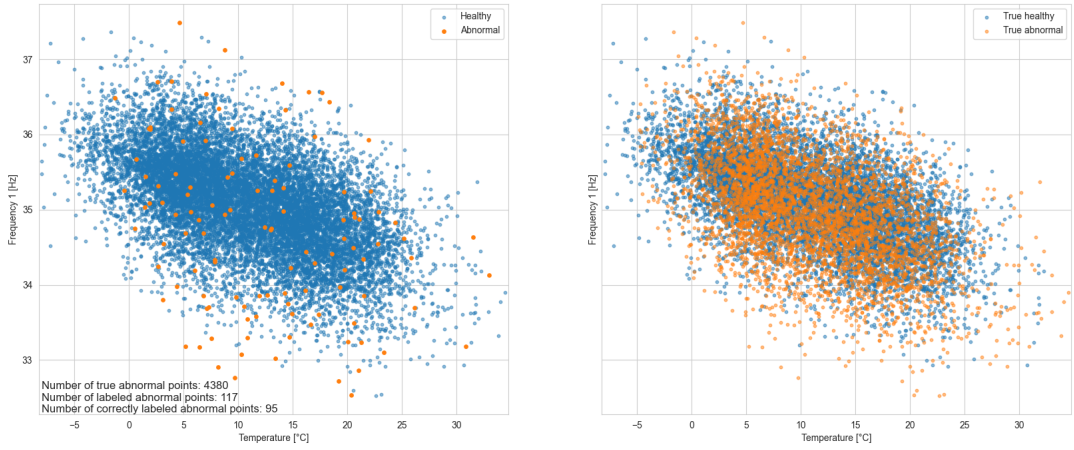
Figure 4.43: Anomaly detection results using different numbers of Gaussians.



(a) Anomaly detection results after 1 month of gradual settlement and gradual loss of stiffness.



(b) Anomaly detection results after 6 months of gradual settlement and gradual loss of stiffness.



(c) Anomaly detection results after 12 months of gradual settlement and gradual loss of stiffness.

Figure 4.44: Anomaly detection results for 150 mm settlement and gradual 3% stiffness loss.

Table 4.5 provides numerical results for the performance of the GMM-based anomaly detection method using a single Gaussian. As observed from the previous figures, the rate at which abnormal points are labeled remains stable for the first 3-4 months, and this is also reflected in the table. The percentage of correctly labeled abnormal points does not exceed 0.5% until the 5th month, after which it consistently rises. These results, along with the previous figures, indicate that even a small percentage of correctly labeled abnormal points is sufficient to conclude that the analyzed model has deviated from its normal behavior.

Table 4.5: Percentage of abnormal points that are correctly labeled (for the case of using 1 Gaussian).

No. of Abnormal Months	True No. of Abnormal Points	Correctly Labeled as Damaged [%]
1	365	0.27
2	730	0.27
3	1095	0.46
4	1460	0.48
5	1825	0.71
6	2190	1.00
7	2555	1.21
8	2920	1.40
9	3285	1.61
10	3650	2.19
11	4015	2.04
12	4380	2.17

Around the 5-month mark since anomaly detection began, it becomes possible to reliably flag abnormal behavior for the scenario involving a 3% gradual stiffness reduction and 150 mm settlement. Figure 4.45 illustrates that the number of labeled abnormal points (non-cumulative) exceeds the threshold for healthy behavior around this time frame. Subsequently, the number of points identified as abnormal consistently remains above this threshold, with a peak occurring at the 11-month mark.

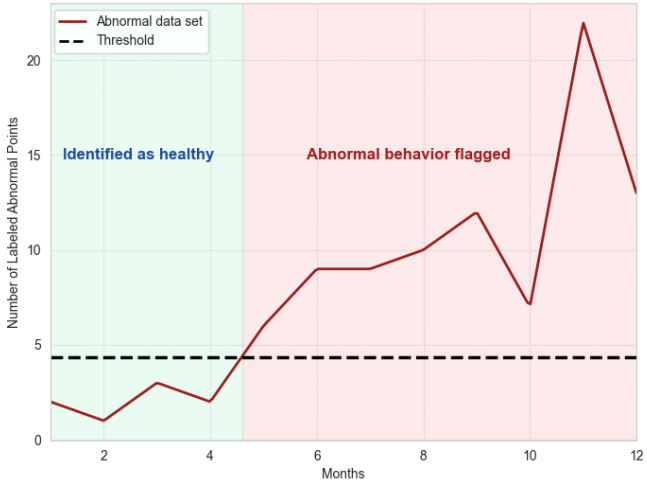


Figure 4.45: Flagging of abnormal behavior using the abnormal data set of 3% gradual stiffness loss and 150 mm settlement.

4.3.2. Combination of Settlement and Sudden 3% Loss in Stiffness

In the upcoming scenario, the settlement remains unchanged, but there is a abrupt loss of stiffness instead of it being gradual over time. The analysis of the combination of these two damage cases will be conducted using the same methodology and procedure.

Relationship Between Temperature and Frequency

From graphing the relationship between the temperature and the natural frequencies (Figures 4.46-4.48, no apparent shift in the frequencies can be observed. This highlights the need to conduct clustering through the GMM-based anomaly detection procedure.

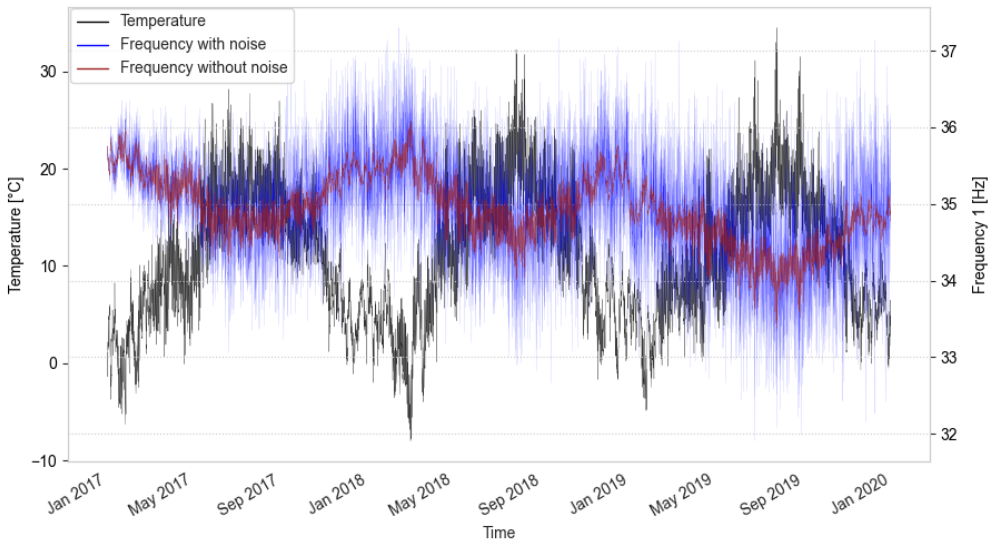


Figure 4.46: Frequency and temperature plotted over time.

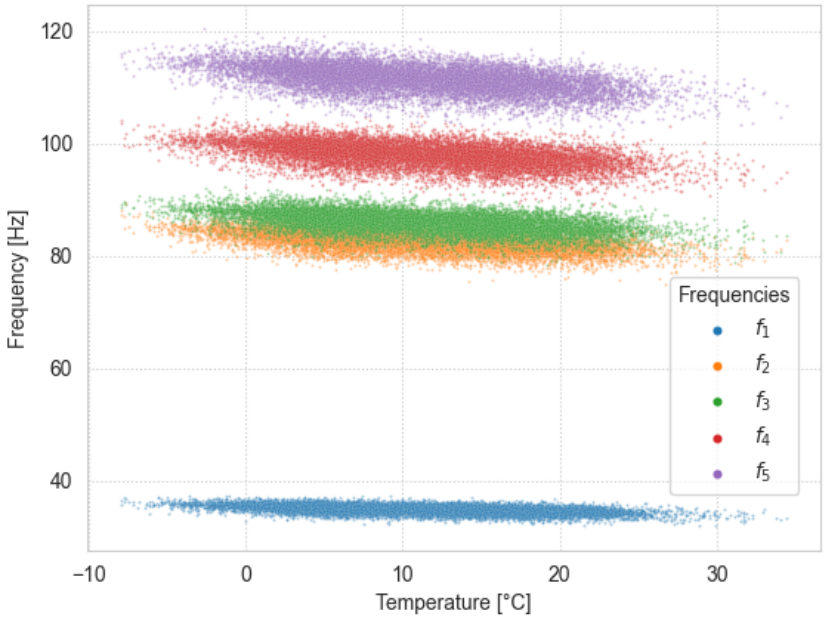


Figure 4.47: Frequency plotted against the temperature.

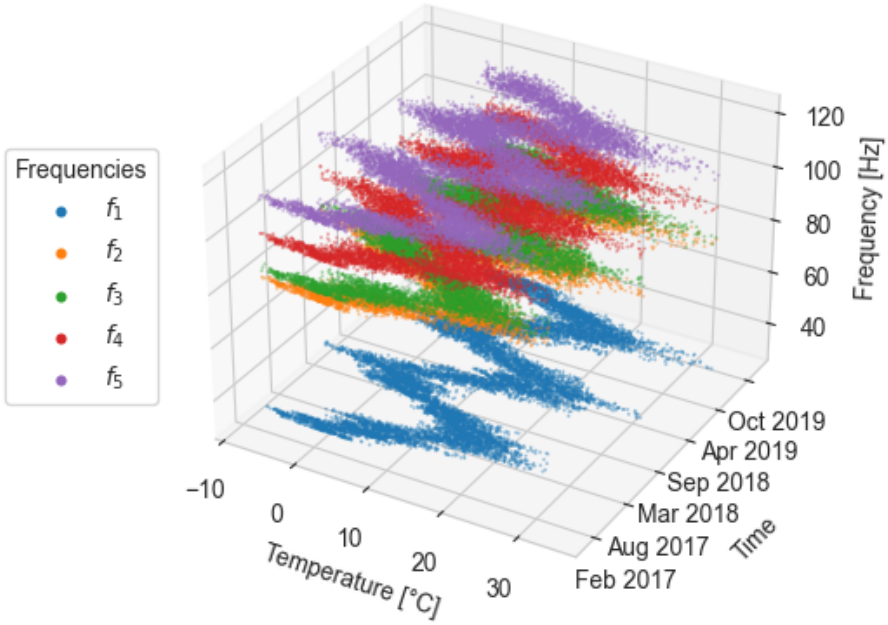


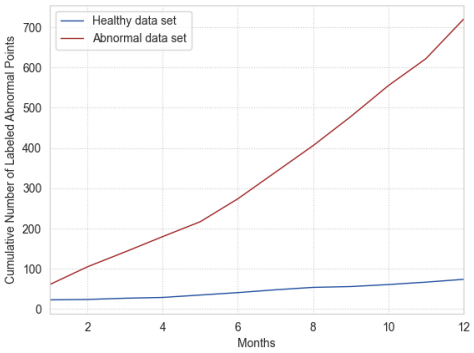
Figure 4.48: Frequency plotted against temperature and time.

Clustering Results

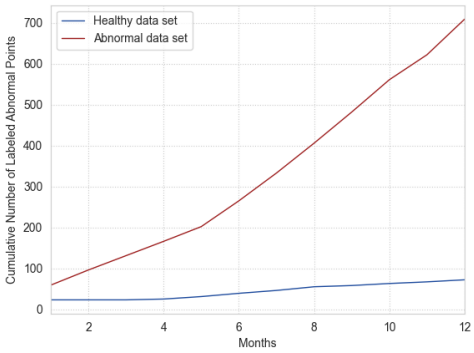
The GMM-based anomaly detection method performs well in detecting the damage scenario of a 5% sudden stiffness reduction and gradual settlement over time. Figure 4.49 shows the results of the algorithm using 1, 5, 25, and 125 Gaussians to model the healthy data distribution. The figures compare the anomaly detection performance between the healthy and abnormal data sets. The algorithm successfully identifies abnormal data points from the beginning of the damage occurrence, regardless of the number of Gaussians used. The detection rate is significantly higher than that of

using the healthy data set for anomaly detection, demonstrating the effectiveness of the proposed GMM-based anomaly detection method.

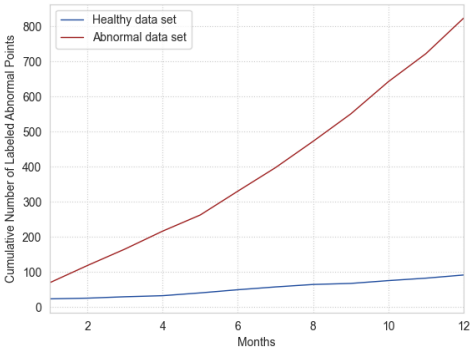
Figure 4.50 presents the performance of the GMM-based anomaly detection algorithm results after 1 month, 6 months, and 12 months of abnormal data using 1 Gaussian distribution to model the healthy distribution. The figures on the left display the data points labeled as abnormal by the proposed procedure, while those on the right provide a comparison by showing the true abnormal points in the data set. The algorithm is able to identify a significant number of abnormal points after just one month of abnormal data, with 61 points labeled as anomalous, of which 39 are correctly labeled. As the number of abnormal months increases, so does the number of labeled abnormal points, with 274 points labeled as abnormal after six months, of which 252 are correctly labeled. After a full year of abnormal data, 720 data points are labeled as abnormal, with 698 correctly labeled, demonstrating the effectiveness and reliability of the proposed anomaly detection algorithm for this particular damage scenario.



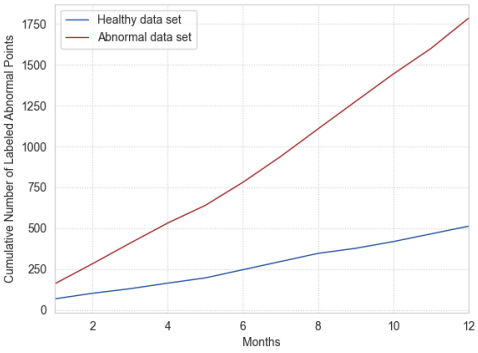
(a) Anomaly detection results using 1 Gaussian.



(b) Anomaly detection results using 5 Gaussians.

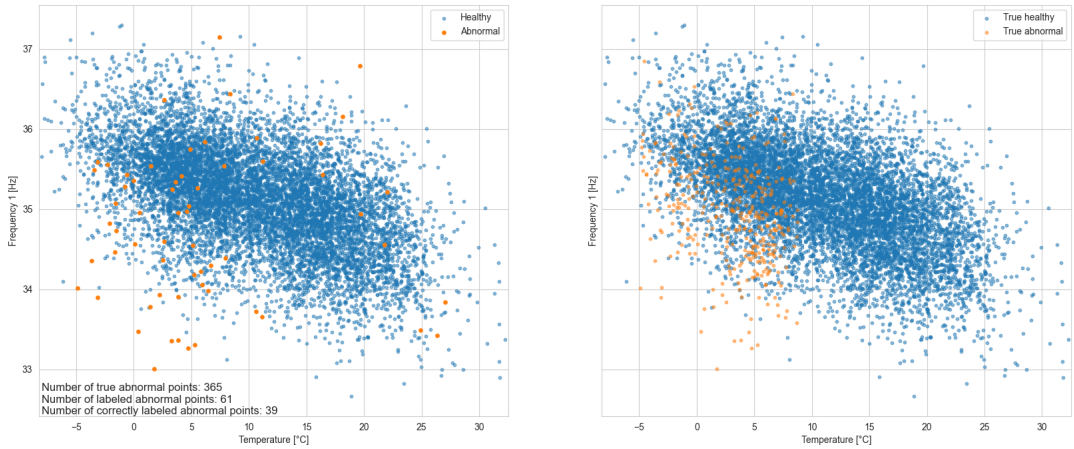


(c) Anomaly detection results using 25 Gaussians.

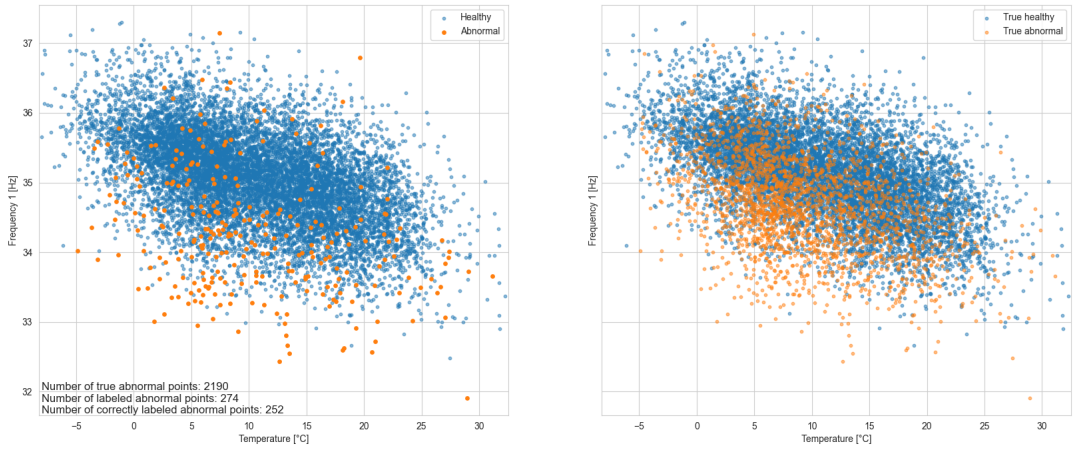


(d) Anomaly detection results using 125 Gaussians.

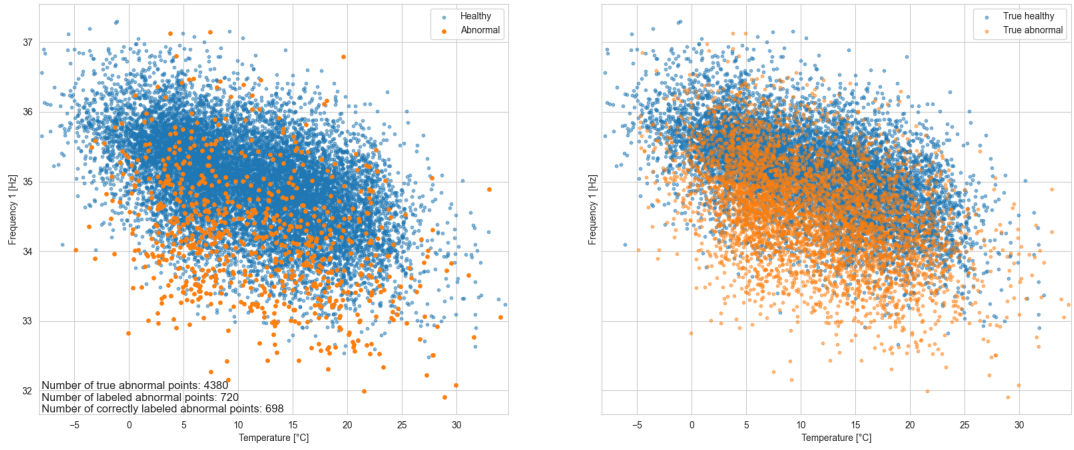
Figure 4.49: Anomaly detection results using different numbers of Gaussians.



(a) Anomaly detection results after 1 month of gradual settlement and sudden loss of stiffness.



(b) Anomaly detection results after 6 months of gradual settlement and sudden loss of stiffness.



(c) Anomaly detection results after 12 months of gradual settlement and sudden loss of stiffness.

Figure 4.50: Anomaly detection results for 150 mm settlement and sudden 3% stiffness loss.

Table 4.6 displays the percentage of correctly labeled damaged points for the damage scenario of 3% sudden stiffness reduction and gradual settlement over time. The proposed GMM-based anomaly detection algorithm accurately labels abnormal points with high precision and consistency. From the first month to the 12th month, the algorithm correctly labels abnormal points ranging from 10.68-15.94%. These results further demonstrate the effectiveness and reliability of the proposed algorithm in detecting the damage in the FEM modeled bridge, starting from the very moment when the damage is introduced.

Table 4.6: Percentage of abnormal points that are correctly labeled (for the case of using 1 Gaussian).

No. of Abnormal Months	True No. of Abnormal Points	Correctly Labeled as Damaged [%]
1	365	10.68
2	730	11.37
3	1095	10.96
4	1460	10.82
5	1825	10.68
6	2190	11.51
7	2555	12.45
8	2920	13.15
9	3285	13.88
10	3650	14.60
11	4015	14.94
12	4380	15.94

4.4. Conclusion

Three case studies were conducted using the methodology developed in Chapter 3. Prior to performing anomaly detection for each case study, the FEM model of the analyzed bridge was presented and introduced. The model was integrated with temperature data from Amsterdam spanning from the beginning of 2017 until the end of 2019, which helped establish the relationship between the natural frequencies of the structure and the temperature. This relationship played a critical role in determining the normal behavior of the model. The three case studies were designed to focus on different damage scenarios to the modeled bridge and they affect the performance on the developed method for anomaly detection. The scenarios being:

- Case Study I: Settlement
- Case Study II: Stiffness Reduction
- Case Study III: Stiffness Reduction & Settlement

The first case study served as a foundation for the subsequent case studies by detailing the procedure for performing anomaly detection and presenting its results. In the first study, the effectiveness of incorporating PCA in the anomaly detection procedure was explored. It was found that each principal component contained a significant amount of information. It was therefore ultimately decided to abandon the idea of employing PCA but rather directly perform GMM-based anomaly detection after preprocessing the data appropriately. The results of all case studies showed no visible abnormal behavior in the relationship between temperature measurements and computed natural frequencies, except for the case of 10% sudden stiffness reduction, where a clear shift in natural frequencies was observed.

For all case studies, the GMM-based anomaly detection method proved effective, particularly for sudden stiffness reduction scenarios of 10%, 5%, and 3%. The algorithm also successfully labeled abnormal points in scenarios with gradual damage over time, the scenarios being settlement and

gradual stiffness reduction. However, in these cases, abnormal point labeling remained at the same rate as mislabeling of abnormal points when performing anomaly detection on a healthy data set for the first few months. It was only after this initial period that the labeling of abnormal points began to deviate from the mislabeling of the healthy data set, leading to the conclusion of abnormal behavior in the structure being analyzed. The GMM-based anomaly detection algorithm proved reliable and timely in detecting abnormal behavior in the analyzed structure for all case studies.

5

Discussion & Conclusion

The development of a methodology using unsupervised SHM to detect anomalies in structures has proven to be valuable and worthwhile. It involved exploring and studying necessary background knowledge from various literature sources, which allowed for the presentation and implementation of a general framework. The methodology was applied and tested on three case studies, which involved the use of a FEM model of a simple bridge and three years of temperature data collected from Amsterdam. With this FEM model and the downloaded temperature data, a relationship between the temperature and the natural frequencies of the bridge model was established. Using this relationship a baseline for healthy behavior of the bridge was defined. This allowed incoming data, in this case anomalous data, to be tested for detection using a clustering algorithm.

The case studies conducted showcased the potential of the developed methodology for damage detection. The effectiveness of the proposed algorithm is particularly evident in gradual and, more notably, abrupt damage scenarios. Detecting subtle damages is crucial and highlights the need for such a method. In these cases, behavioral changes in the data can easily go unnoticed, as they are virtually undetectable to the naked eye. Applying the methodology enables data analysts or engineers to uncover these abnormal behavioral changes in the data. In contrast, severe sudden damage scenarios, such as a localized 10% or greater reduction in bridge stiffness, are often visible early on when analyzing the data. This was observed in Case Study II, where a shift in natural frequencies occurred. In such instances, abnormal behavior or damage might be detected even before performing any anomaly detection.

The true advantage of this method is its ability to operate without the need for acquiring damaged data. Obtaining damaged data is challenging in real-world situations, making methods that require training on such data impractical. This underscores the importance of using unsupervised learning techniques in the SHM field. By relying solely on real-time monitoring data of the structure, the need for complex, economically and computationally expensive techniques can be eliminated. Another significant strength of the proposed GMM-based anomaly detection is the ability for the user to define the number of Gaussian distributions required to model the healthy distribution. This implies that if the relationship between temperature and natural frequencies is bilinear, two or more Gaussians could likely model and establish an accurate baseline for the healthy behavior of the monitored structure.

In the case studies conducted in this thesis, it was observed that the success of anomaly detection was independent of the number of Gaussians employed. By increasing the number of Gaussians to model the healthy distribution, the number of abnormal points labeled, as well as their correct labeling, increased. However, when using the same number of Gaussians and performing anomaly detection on a healthy data set, the rate of clustering labels for abnormal points remained consistent for both data sets during the initial few months. Following this period, the anomaly detection rate for the abnormal data set increased significantly compared to the healthy data set. This indicates the importance of establishing a healthy rate for labeling abnormal data, as it can help prevent the incorrect conclusion that abnormal behavior is present in the analyzed structure. By defining this rate, one can confidently conclude that abnormal behavior is present when the rate deviates from its established norm. Figures 4.19 and 4.27 demonstrated the point at which abnormal behavior can be safely identified for two of the considered damage scenarios.

For the gradual settlement and gradual stiffness loss scenarios, the first two months of data were used to establish a healthy data trend line. This trend line represented the normal and expected rate of abnormal point labeling. Consequently, if the labeling of abnormal points would begin to deviate from this trend, abnormal behavior can be safely flagged, prompting further investigation or visual inspection of the monitored structure. In the case of settlement, abnormal behavior was seen to be confidently identified at approximately 6 months after its onset. This corresponded to a settlement of 250 mm for the right pier in the FEM modeled bridge. For the gradual stiffness loss scenario, abnormal behavior was seen to be safely flagged around 5 months after the stiffness begins to decrease gradually. This corresponded to a 4.2% loss of stiffness in the two selected members of the FEM modeled bridge at the time when abnormal behavior is identified. It is worth noting that the threshold for deviation from the norm can be adjusted, allowing for abnormal behavior to be flagged even earlier. This means that in the case that the threshold is lowered, a smaller deviation from the established normal rate of labeling abnormal points would be sufficient to alert the analyst or engineer to the presence of abnormal behavior.

In terms of future work, there is significant potential for improvement. The FEM model used in this study can be refined further by incorporating more realistic representations of joints and modeling the bridge in three dimensions. These improvements would lead to a more accurate depiction of the bridge's behavior. Although temperature effects on structures are highly complex, the treatment of temperature in this thesis was relatively simplistic. Efforts to integrate temperature effects in a more comprehensive and realistic manner could be valuable, despite the challenges involved. Additionally, modeling an asphalt layer on the bridge can be considered to account for bilinear behavior in the relationship between temperature and natural frequencies around the freezing point. However, it is important to note that while modeling is crucial, the primary focus of this thesis was on the development and application of an unsupervised SHM method. To further enhance the methodology's credibility, testing it with real-world data, minimizing or even eliminating structural modeling, would be highly beneficial. Such testing would enable the collection of various measurements and features, allowing for diverse evaluations of the algorithm, not limited to natural frequencies alone. This could also provide an opportunity for PCA to play a valuable role in detecting abnormal behavior.

Regarding clustering, different data shapes may emerge when dealing with various structures and data sets. Consequently, it is essential to test multiple clustering algorithms to choose the most suitable one for damage detection. The algorithms discussed in this thesis do not cover all possible approaches for detecting anomalies. Exploring other options such as OPTICS (Ordering Points To Identify the Clustering Structure) and HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) could prove promising for anomaly detection and warrant further investigation. Additionally, using probabilistic PCA for data sets like the one in this thesis could be highly beneficial. More frequent monitoring with additional healthy data points can also be considered, such as by generating data points at hourly intervals instead of two-hour intervals. Bootstrapping and the Monte Carlo method could be helpful for generating more data points, although this may increase computational costs. With more healthy data points, a more accurate distribution for healthy data can be modeled, enhancing anomaly detection accuracy. Investigating whether filtering out noise before defining a healthy distribution baseline would be advantageous is another area to explore. Assessing the impact of reducing the number of obtained natural frequencies on the results is also worth considering for future work. Ultimately, experimental and real-life work would be the ideal direction for future research related to this unsupervised SHM methodology.

References

[1] Charles R. Farrar and Keith Worden. "An introduction to structural health monitoring". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365.1851 (Feb. 2007), pp. 303–315. ISSN: 1364503X. DOI: 10.1098/RSTA.2006.1928. URL: https://royalsocietypublishing.org/.

[2] Kareem A. Eltouny and Xiao Liang. "Bayesian-optimized unsupervised learning approach for structural damage detection". In: *Computer-Aided Civil and Infrastructure Engineering* 36.10 (Oct. 2021), pp. 1249–1269. ISSN: 1093-9687. DOI: 10.1111/mice.12680.

[3] A B Siddique, B F Sparling, and L D Wegner. "Assessment of vibration-based damage detection for an integral abutment bridge 1". In: (2007). DOI: 10.1139/L07-023.

[4] Guang Dong Zhou and Ting Hua Yi. *A summary review of correlations between temperatures and vibration properties of long-span bridges*. 2014. DOI: 10.1155/2014/638209.

[5] J. M. Ko and Y. Q. Ni. "Technology developments in structural health monitoring of large-scale bridges". In: *Engineering Structures* 27.12 SPEC. ISS. (2005), pp. 1715–1725. ISSN: 01410296. DOI: 10.1016/j.engstruct.2005.02.021.

[6] Guang Dong Zhou and Ting Hua Yi. "Thermal load in large-scale bridges: A state-of-the-art review". In: *International Journal of Distributed Sensor Networks* 2013 (2013). ISSN: 15501477. DOI: 10.1155/2013/217983.

[7] Yanxia Cai et al. "Influence of Temperature on the Natural Vibration Characteristics of Simply Supported Reinforced Concrete Beam". In: (2021). DOI: 10.3390/s21124242. URL: https://doi.org/10.3390/s21124242.

[8] *CEB-FIP Model Code 2010. InFib Special Activity Group 5*. Lausanne, Switzerland, 2010.

[9] Bart Peeters, Johan Maeck, and Guido De Roeck. "Vibration-based damage detection in civil engineering: Excitation sources and temperature effects". In: *Smart Materials and Structures* 10.3 (June 2001), pp. 518–527. ISSN: 09641726. DOI: 10.1088/0964-1726/10/3/314.

[10] Bart Peeters et al. *Article in Proceedings of SPIE-The International Society One year monitoring of the z24-bridge: environmental influences versus damage events*. Tech. rep. 2000, pp. 1570–1576.

[11] Dennis Keíth Wøtson. "Seasonal Vaniation in Nlaterial Pnoperties of a Flexible Pavement". In: (1996).

[12] Peter Moser and Babak Moaveni. "Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge". In: *Mechanical Systems and Signal Processing* 25.7 (Oct. 2011), pp. 2336–2357. ISSN: 08883270. DOI: 10.1016/J.YMSSP.2011.03.005.

[13] Ignacio Gonzales, Mahir Ülker-Kaustell, and Raid Karoumi. "Seasonal effects on the stiffness properties of a ballasted railway bridge". In: *Engineering Structures* 57 (Dec. 2013), pp. 63–72. ISSN: 01410296. DOI: 10.1016/j.engstruct.2013.09.010.

[14] Prathamesh M Jagdale and M A Chakrabarti. "Free Vibration Analysis of Cracked Beam". In: *Journal of Engineering Research and Applications www.ijera.com* 3 (), pp. 1172–1176. URL: www.ijera.com.

[15] M I Friswell and J E T Penny. "Crack Modeling for Structural Health Monitoring". In: (). DOI: 10.1106/145792102028836.

[16] M. I. Friswell and J. E. Mottershead. "Finite Element Model Updating in Structural Dynamics". In: Solid Mechanics and its Applications 38 (1995). DOI: 10.1007/978-94-015-8508-8. URL: http://link.springer.com/10.1007/978-94-015-8508-8.

[17] A. Teughels and G. De Roeck. "Structural damage identification of the highway bridge Z24 by FE model updating". In: *Journal of Sound and Vibration* 278.3 (Dec. 2004), pp. 589–610. ISSN: 0022460X. DOI: 10.1016/J.JSV.2003.10.041.

[18] Pelin Gundes Bakir, Edwin Reynders, and Guido De Roeck. "Sensitivity-based finite element model updating using constrained optimization with a trust region algorithm". In: *Journal of Sound and Vibration* 305.1-2 (Aug. 2007), pp. 211–225. ISSN: 10958568. DOI: 10.1016/J.JSV.2007.03.044.

[19] Bijaya Jaishi and Wei-Xin Ren. "Structural Finite Element Model Updating Using Ambient Vibration Test Results". In: *Journal of Structural Engineering* 131.4 (Apr. 2005), pp. 617–628. ISSN: 0733-9445. DOI: 10.1061/(ASCE)0733-9445(2005)131:4(617).

[20] J. L. Beck and L. S. Katafygiotis. "Updating Models and Their Uncertainties. I: Bayesian Statistical Framework". In: *Journal of Engineering Mechanics* 124.4 (Apr. 1998), pp. 455–461. ISSN: 0733-9399. DOI: 10.1061/(ASCE)0733-9399(1998)124:4(455).

[21] Tshilidzi Marwala. *Finite-element-model Updating Using Computional Intelligence Techniques*. London: Springer London, 2010. ISBN: 978-1-84996-322-0. DOI: 10.1007/978-1-84996-323-7.

[22] Masoud Sanayei et al. "Significance of Modeling Error in Structural Parameter Estimation". In: *Computer-Aided Civil and Infrastructure Engineering* 16.1 (Jan. 2001), pp. 12–27. ISSN: 1093-9687. DOI: 10.1111/0885-9507.00210.

[23] Charles R. Farrar and Keith Worden. *Structural Health Monitoring: A Machine Learning Perspective*. 1st ed. Wiley, 2012. ISBN: 9781119994336. DOI: 10.1002/9781118443118.

[24] Tom M. Mitchell. *Machine Learning*. 1st ed. McGraw-Hill Education, Mar. 1997.

[25] Junjie Peng et al. *Machine Learning Techniques for Personalised Medicine Approaches in Immune-Mediated Chronic Inflammatory Diseases: Applications and Challenges*. Sept. 2021. DOI: 10.3389/fphar.2021.720694.

[26] Luis G. Serrano. *Grokking Machine Learning*. URL: https://livebook.manning.com/book/grokking-machine-learning/chapter-2/36.

[27] Arman Malekloo et al. "Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights". In: 21.4 (2019), pp. 1906–1955. DOI: 10.1177/14759217211036880.

[28] Piervincenzo Rizzo et al. "An unsupervised learning algorithm for fatigue crack detection in waveguides". In: (2009). DOI: 10.1088/0964-1726/18/2/025016.

[29] Thanh T X Tran and Ekin Ozer. "Automated and Model-Free Bridge Damage Indicators with Simultaneous Multiparameter Modal Anomaly Detection". In: (). DOI: 10.3390/s20174752. URL: www.mdpi.com/journal/sensors.

[30] Naveen Naidu Narisetty. "Bayesian model selection for high-dimensional data". In: *Handbook of Statistics*. Vol. 43. Elsevier B.V., Jan. 2020, pp. 207–248. ISBN: 9780444642110. DOI: 10.1016/bs.host.2019.08.001.

[31] Chun Kit, Jeffery Hou, and · Kamran Behdinan. "Dimensionality Reduction in Surrogate Modeling: A Review of Combined Methods". In: 7 (2022), pp. 402–427. DOI: 10.1007/s41019-022-00193-5. URL: https://doi.org/10.1007/s41019-022-00193-5.

[32] I.T. Jolliffe. *Principal Component Analysis*. 2nd ed. New York, NY: Springer-Verlag, 2002. ISBN: 0-387-95442-2. DOI: 10.1007/b98835.

[33] Joaquín Pérez-Ortega et al. "The K-Means Algorithm Evolution". In: *Introduction to Data Science and Machine Learning*. IntechOpen, Mar. 2020. DOI: 10.5772/intechopen.85447.

[34] Jiawei Han, Micheline Kamber, and Jian Pei. *Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)*. 3rd ed. 2011.

[35] Ulrike Von Luxburg. "A tutorial on spectral clustering". In: *Statistics and Computing* 17.4 (Dec. 2007), pp. 395–416. ISSN: 09603174. DOI: 10.1007/s11222-007-9033-z.

[36] Jianbo Shi and Jitendra Malik. "Normalized cuts and image segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.8 (2000), pp. 888–905. ISSN: 01628828. DOI: 10.1109/34.868688.

[37] Abhishek Sharma. *How Does DBSCAN Clustering Work? | DBSCAN Clustering for ML*. URL: https://www.analyticsvidhya.com/blog/2020/09/how-dbscan-clustering-works/.

[38] Martin Ester et al. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". In: (1996). URL: www.aaai.org.

[39] Erich Schubert et al. "9 DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN". In: *DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. ACM Trans. Database Syst* 42.3 (2017). DOI: 10.1145/3068335. URL: https://doi.org/10.1145/3068335.

[40] M Jordan, J Kleinberg, and B Schölkopf. "Pattern Recognition and Machine Learning". In: ().

[41] Andrew Y Ng and Michael I Jordan. "On Spectral Clustering: Analysis and an algorithm". In: ().

[42] F. Pedragosa et al. "Scikit-learn: Machine learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[43] Ian T. Jollife and Jorge Cadima. *Principal component analysis: A review and recent developments*. Apr. 2016. DOI: 10.1098/rsta.2015.0202.

[44] Tarek Elgamal and Mohamed Hefeeda. "Analysis of PCA Algorithms in Distributed Environments". In: (2015).

[45] Lindsay I Smith. "A tutorial on Principal Components Analysis Chapter 1". In: (2002).

[46] V. Klema and A. Laub. "The singular value decomposition: Its computation and some applications". In: *IEEE Transactions on Automatic Control* 25.2 (1980), pp. 164–176. ISSN: 0018-9286. URL: https://www.academia.edu/6369464/The_Singular_Value_Decomposition_Its_Computation_and_Some_Applications.

[47] *19.2. PCA using the Singular Value Decomposition — Principles and Techniques of Data Science*. URL: https://www.samlau.me/test-textbook/ch/19/pca_svd.html.

[48] *CIS520 Machine Learning | Lectures / PCA*. URL: https://alliance.seas.upenn.edu/~cis520/dynamic/2022/wiki/index.php?n=Lectures.PCA.

[49] *Everything you did and didn't know about PCA · Its Neuronal*. URL: http://alexhwilliams.info/itsneuronalblog/2016/03/27/pca/#f3b.

[50] Nadipuram R. Prasad, Salvador Almanza-Garcia, and Thomas T. Lu. "Anomaly detection". In: *Computers, Materials and Continua* 14.1 (2009), pp. 1–22. ISSN: 15462218. DOI: 10.1145/1541880.1541882.

[51] *Anomaly Detection: Industrial Asset Insights Without Historical Data | Engineering.com*. URL: https://www.engineering.com/story/anomaly-detection-industrial-asset-insights-without-historical-data.

[52] Chuanlei Zhang et al. "Unsupervised Anomaly Detection Based on Deep Autoencoding and Clustering". In: (2021). DOI: 10.1155/2021/7389943. URL: https://doi.org/10.1155/2021/7389943.

[53] Anil K. Jain. "Data clustering: 50 years beyond K-means". In: *Pattern Recognition Letters* 31.8 (June 2010), pp. 651–666. ISSN: 01678655. DOI: 10.1016/J.PATREC.2009.09.011.

[54] *Data Platform – Open Power System Data*. URL: https://data.open-power-system-data.org/weather_data/2020-09-16.

[55] Yong Xia et al. *Temperature effect on vibration properties of civil structures: A literature review and case studies*. May 2012. DOI: 10.1007/s13349-011-0015-7.

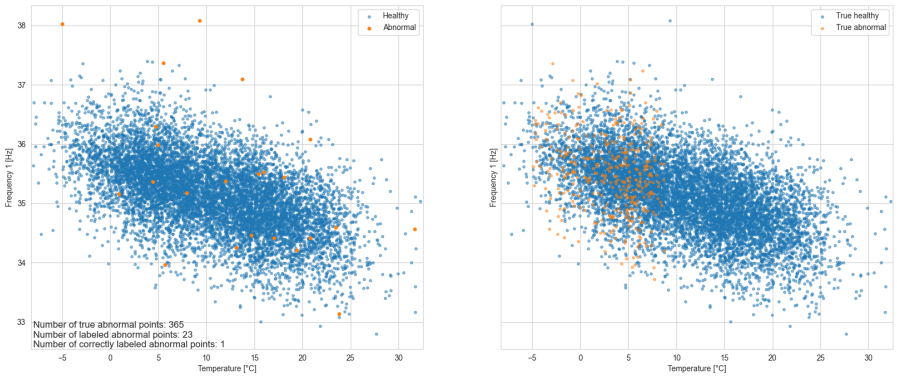
A

Results - Case Studies

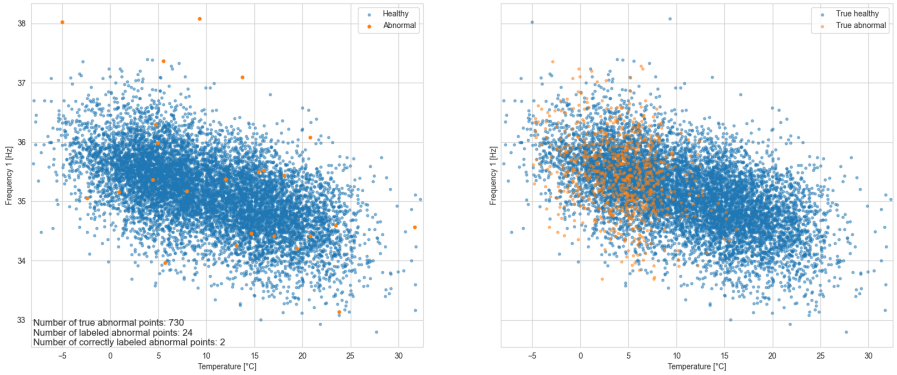
This section of the annex presents the clustering results for each month of all case studies, as well as the results of the anomaly detection performed on a healthy data set.

A.1. Anomaly Detection: Healthy Data Set

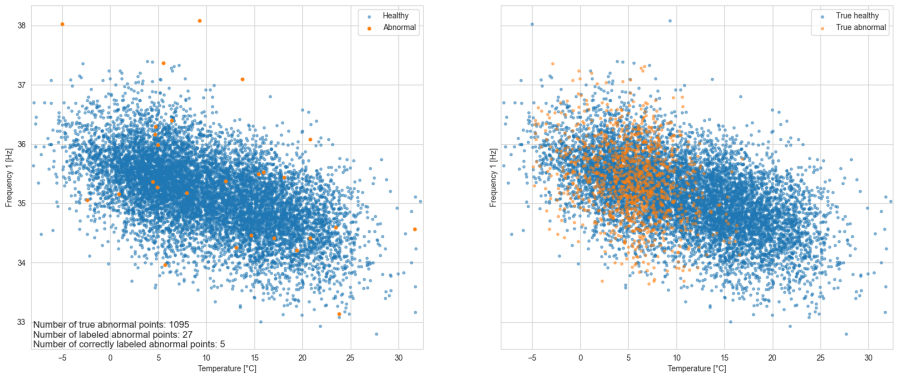
In order to better understand the performance of the proposed anomaly detection method, a test was conducted on a completely healthy data set. The results of this test are presented below.



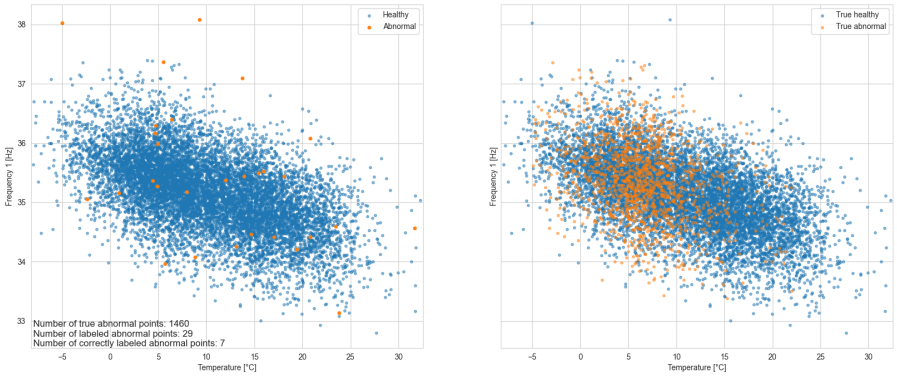
GMM-based anomaly detection conducted on a healthy data set - Months: 1



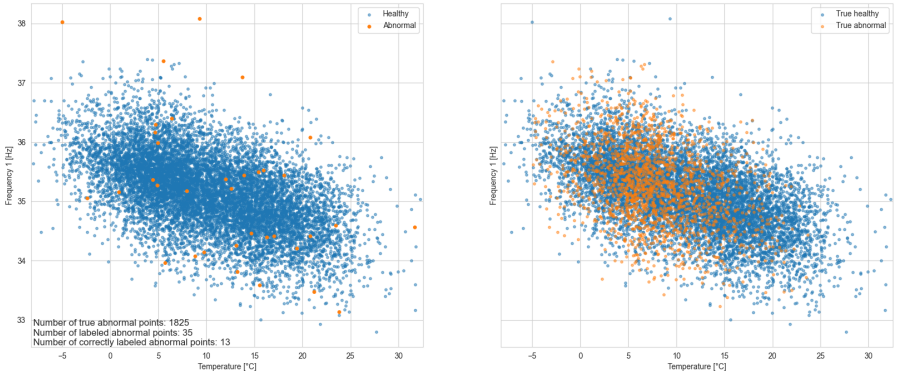
GMM-based anomaly detection conducted on a healthy data set - Months: 2



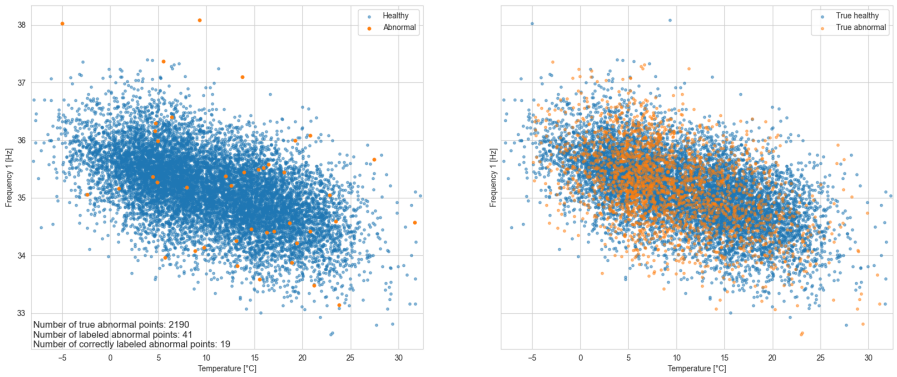
GMM-based anomaly detection conducted on a healthy data set - Months: 3



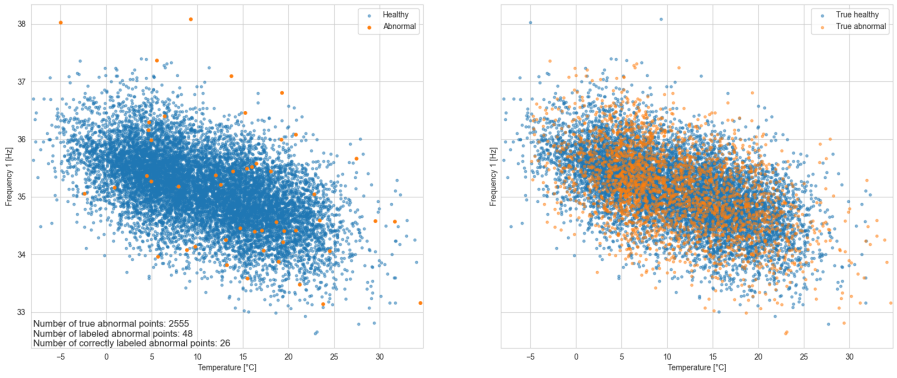
GMM-based anomaly detection conducted on a healthy data set - Months: 4



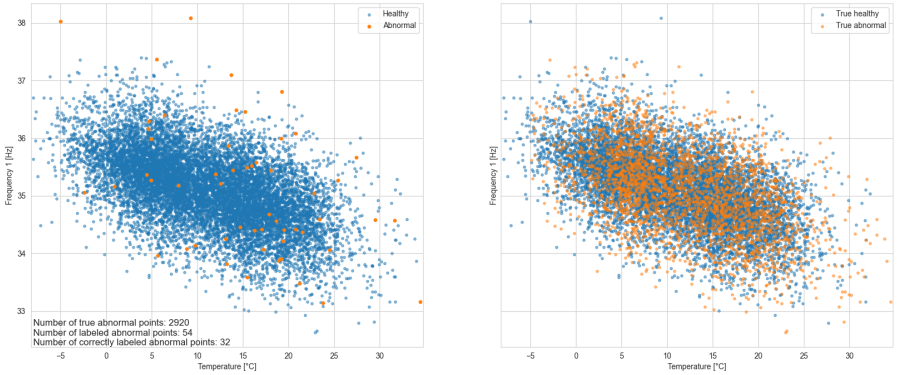
GMM-based anomaly detection conducted on a healthy data set - Months: 5



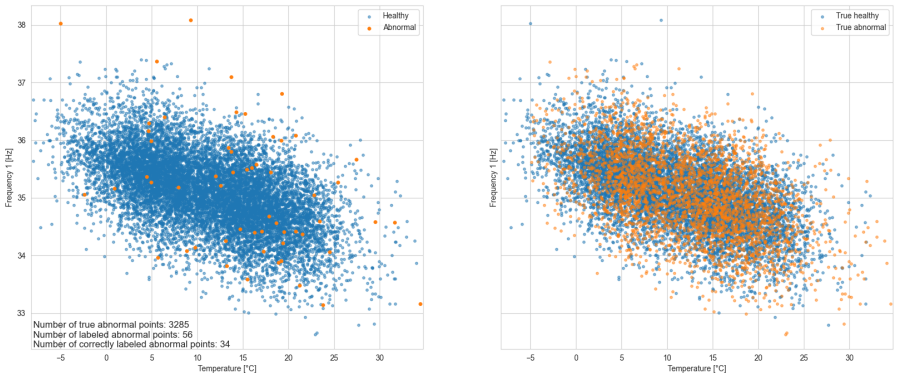
GMM-based anomaly detection conducted on a healthy data set - Months: 6



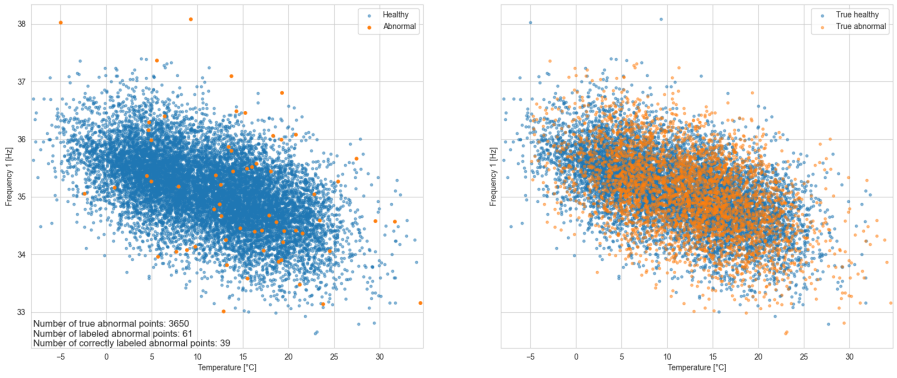
GMM-based anomaly detection conducted on a healthy data set - Months: 7



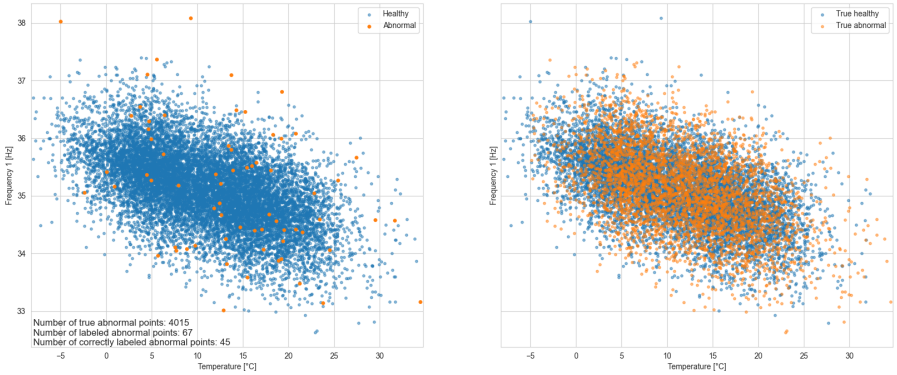
GMM-based anomaly detection conducted on a healthy data set - Months: 8



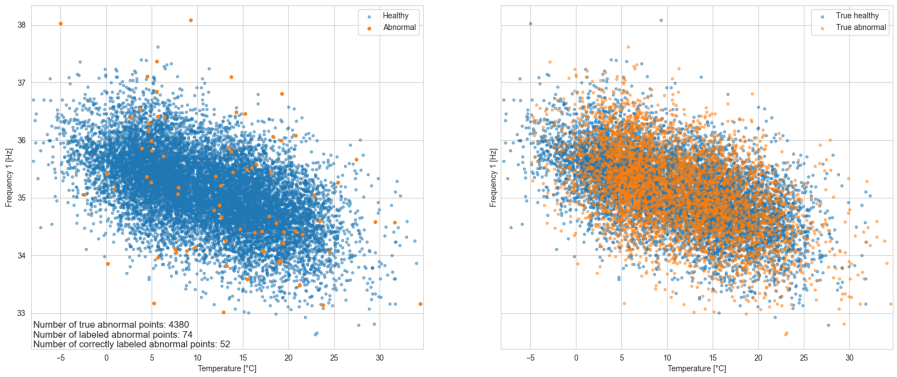
GMM-based anomaly detection conducted on a healthy data set - Months: 9



GMM-based anomaly detection conducted on a healthy data set - Months: 10

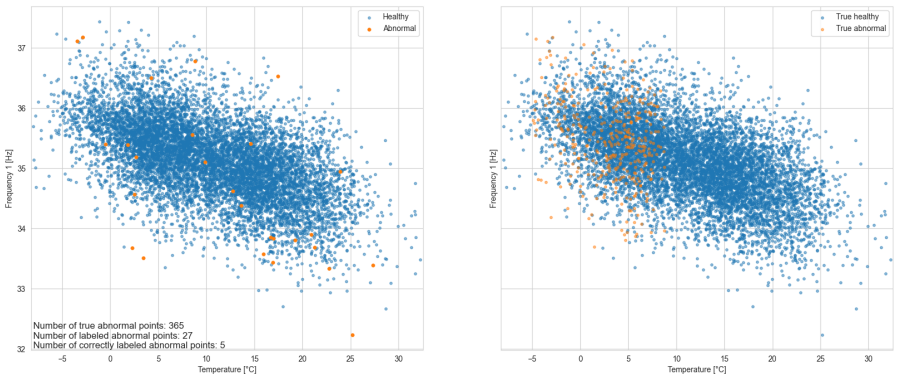


GMM-based anomaly detection conducted on a healthy data set - Months: 11

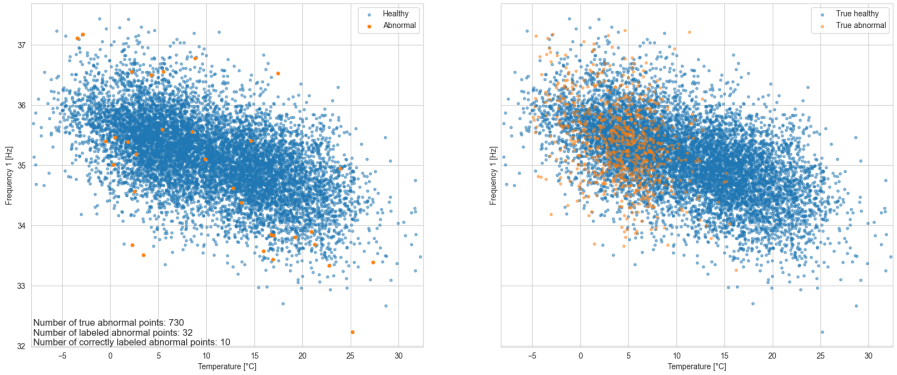


GMM-based anomaly detection conducted on a healthy data set - Months: 12

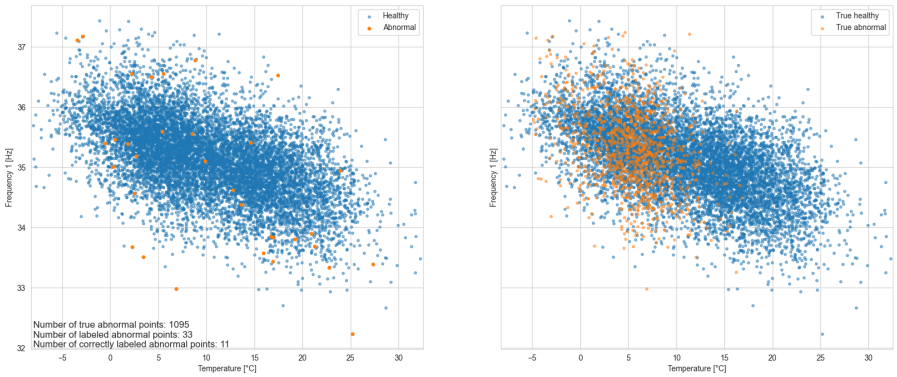
A.2. Case Study I: Settlement



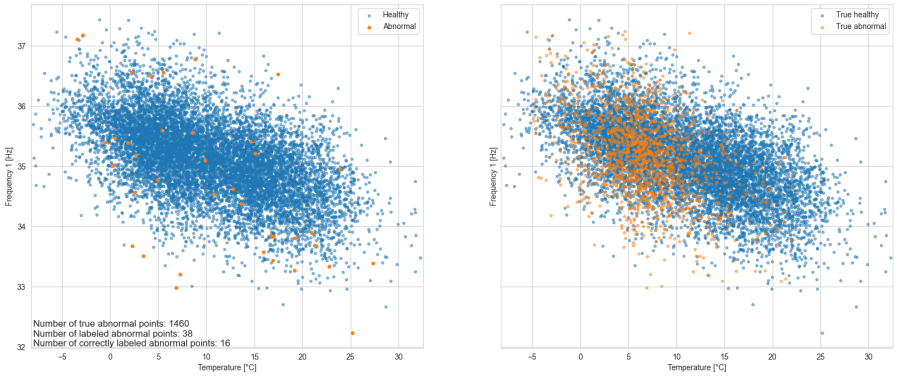
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 1



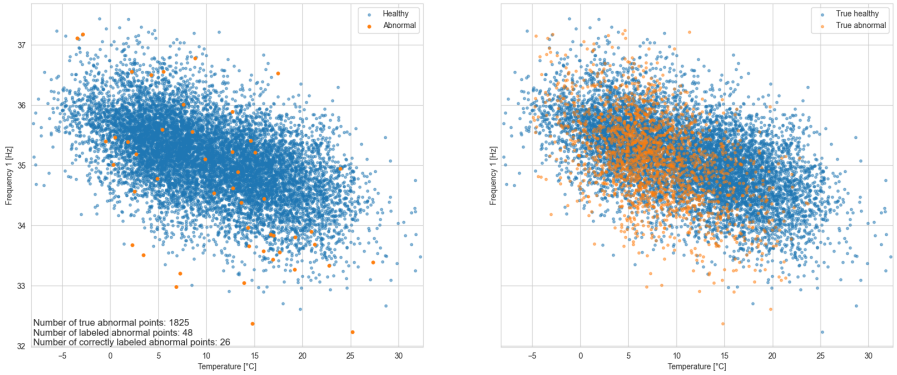
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 2



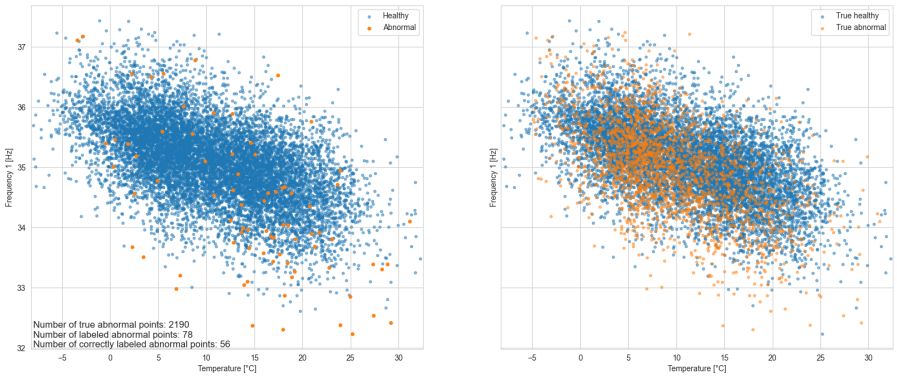
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 3



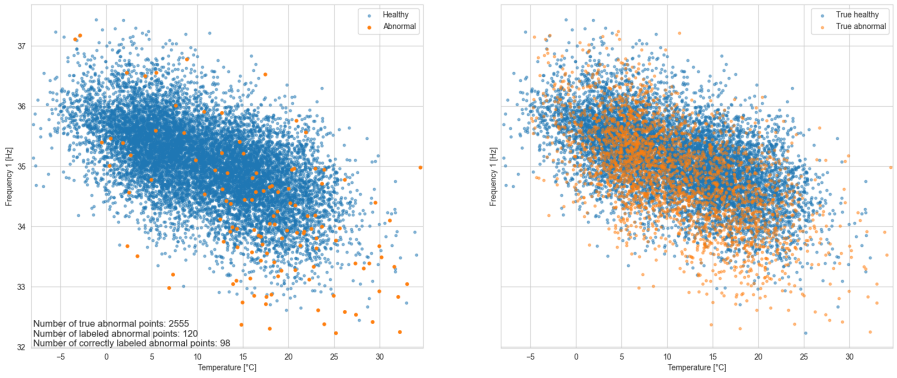
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 4



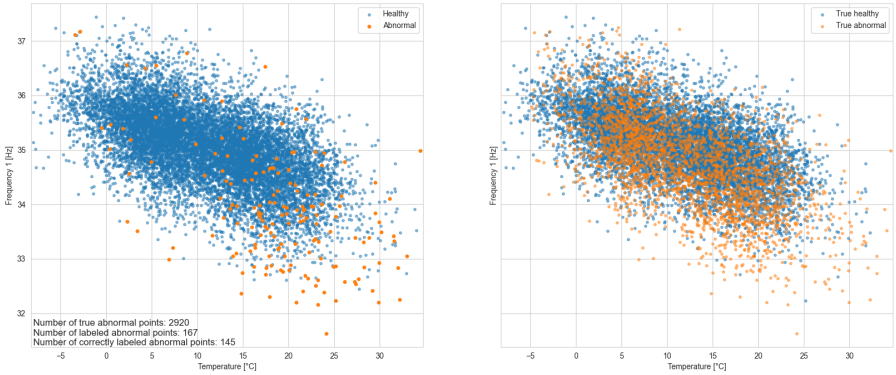
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 5



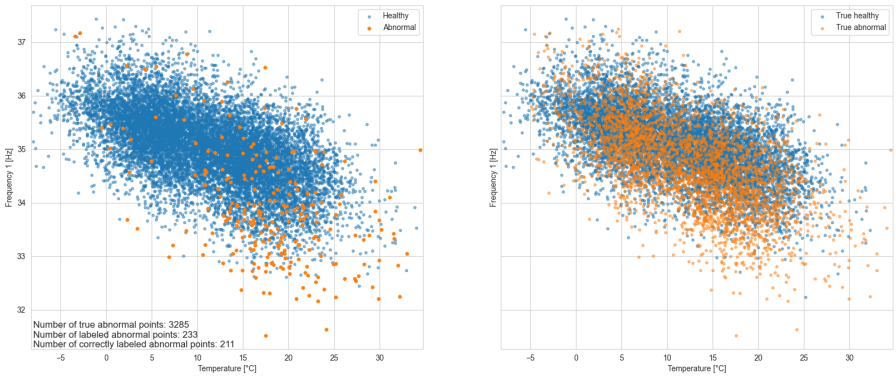
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 6



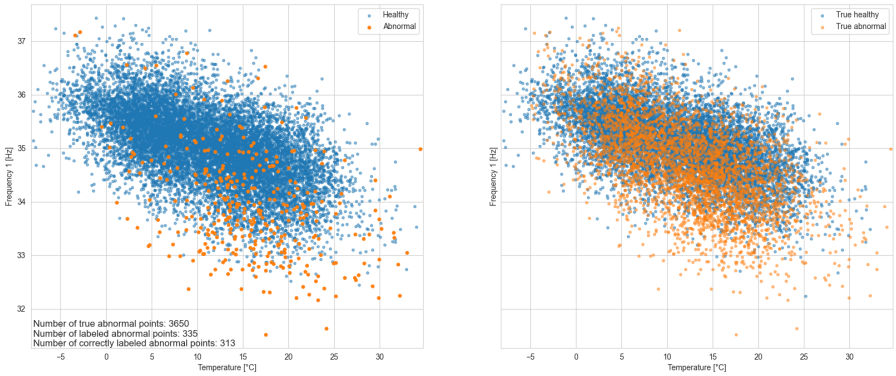
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 7



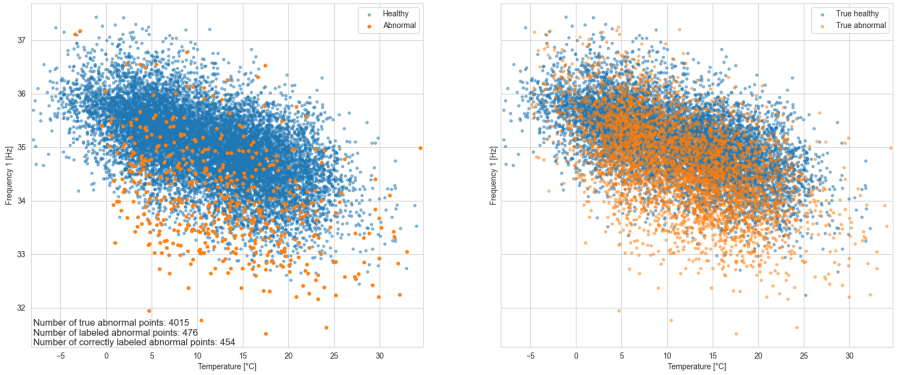
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 8



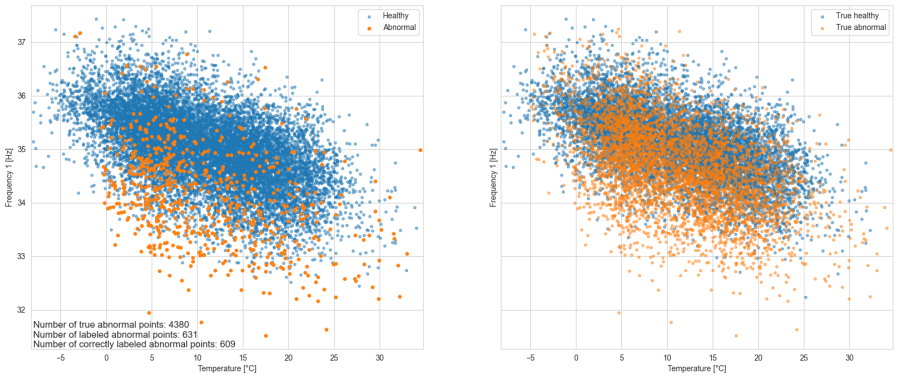
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 9



GMM-based anomaly detection results for 500 mm gradual settlement - Months: 10



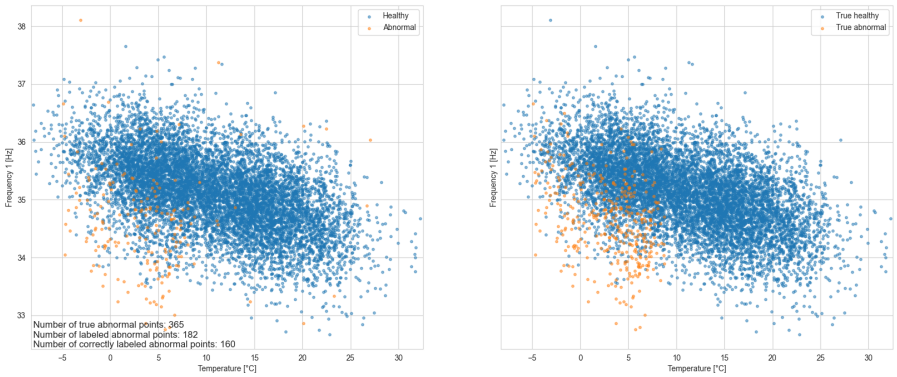
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 11



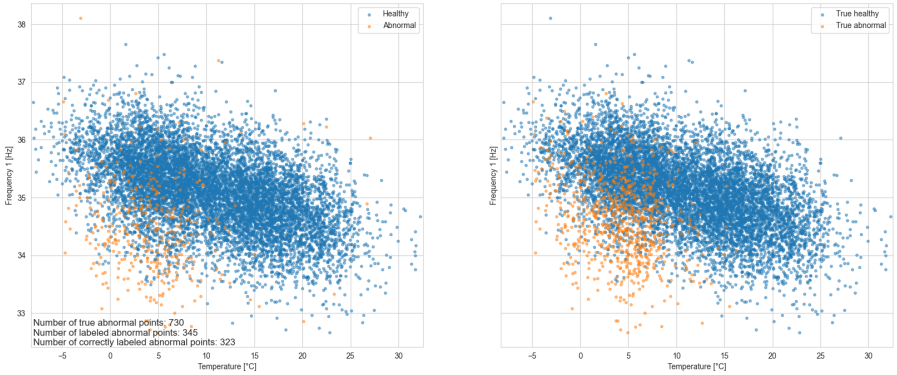
GMM-based anomaly detection results for 500 mm gradual settlement - Months: 12

A.3. Case Study II: Stiffness Reduction

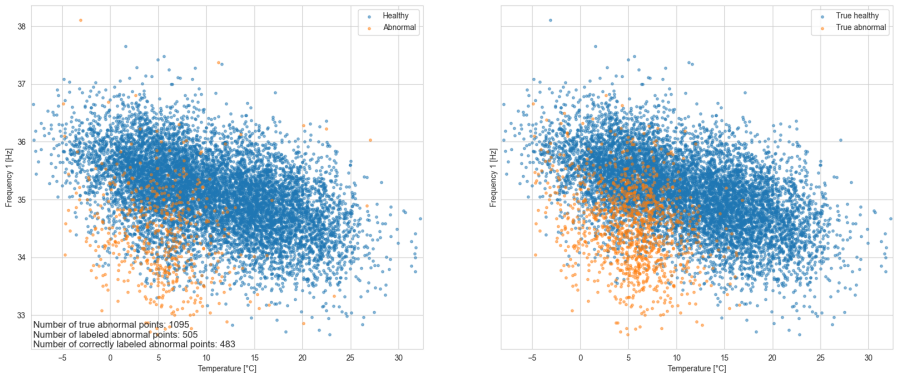
A.3.1. 10% Gradual Stiffness Loss



GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 1

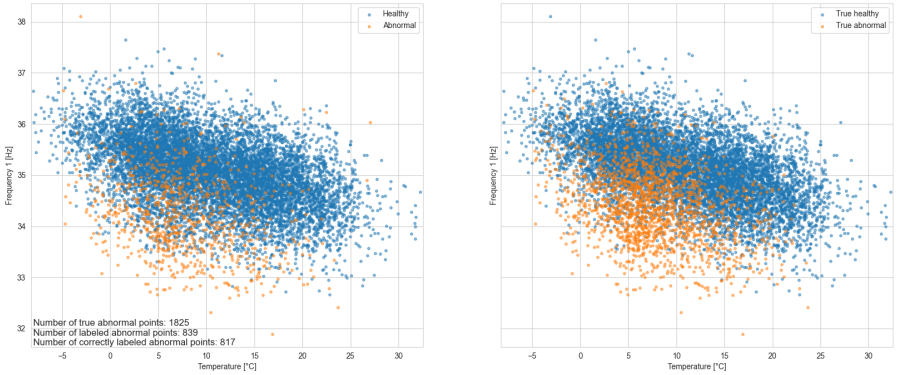


GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 2

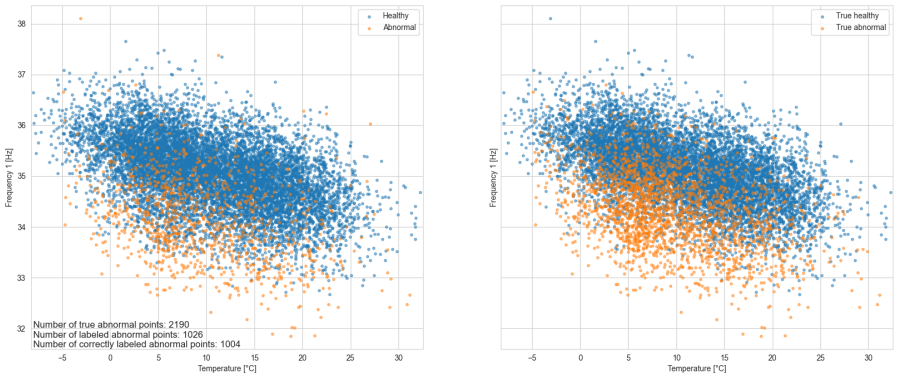


GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 3

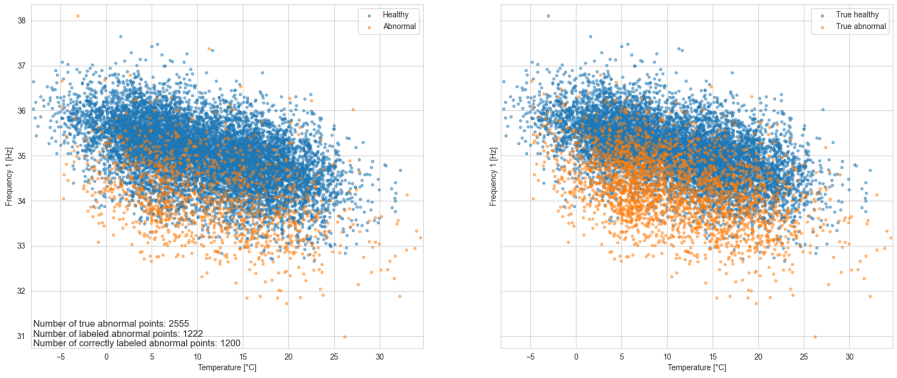
GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 4



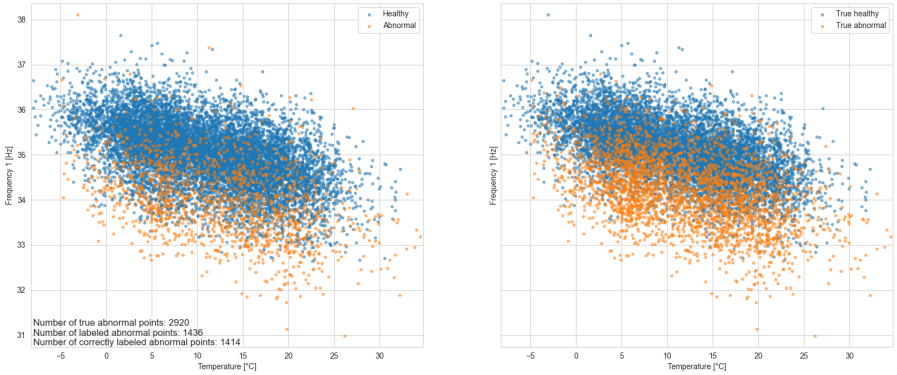
GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 5



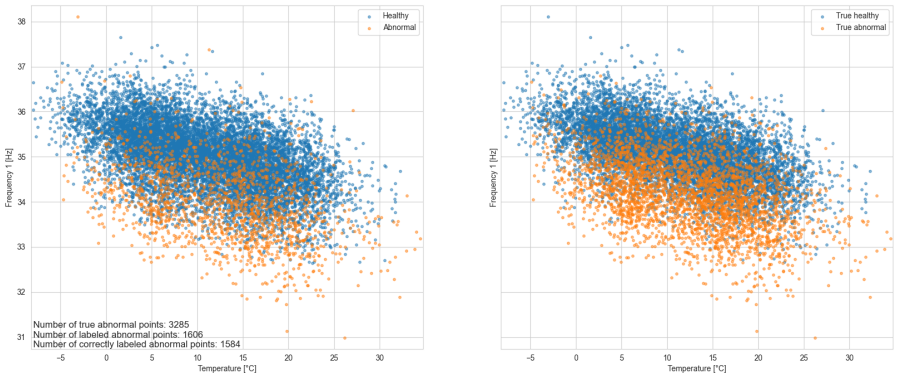
GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 6



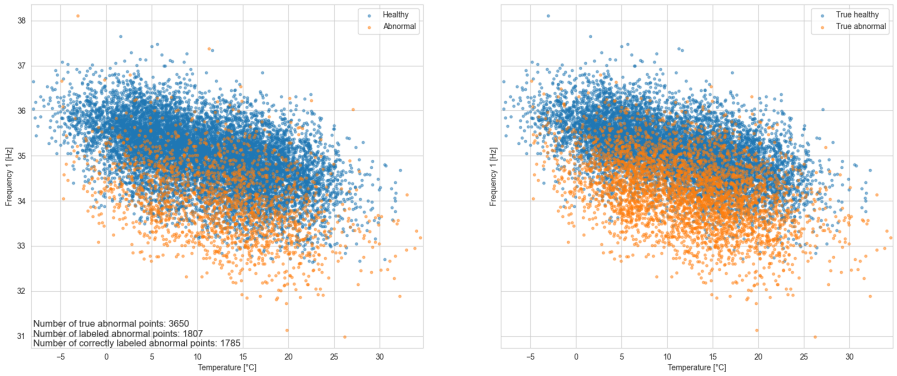
GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 7



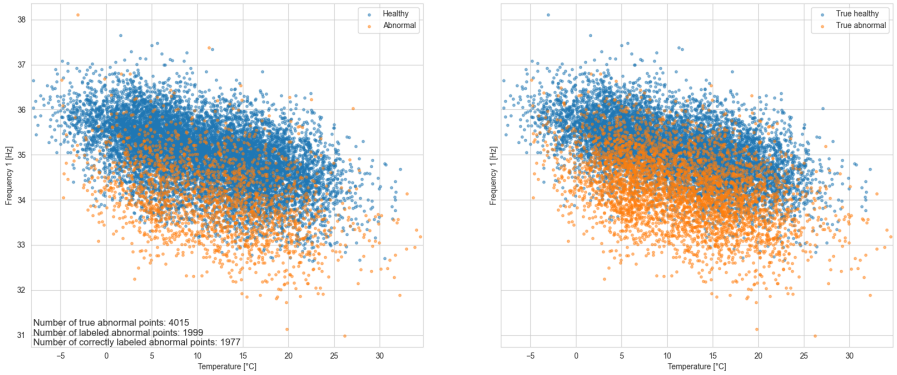
GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 8



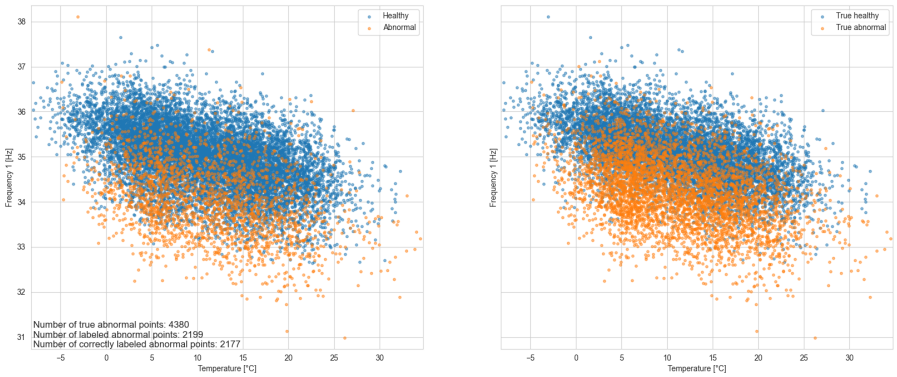
GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 9



GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 10

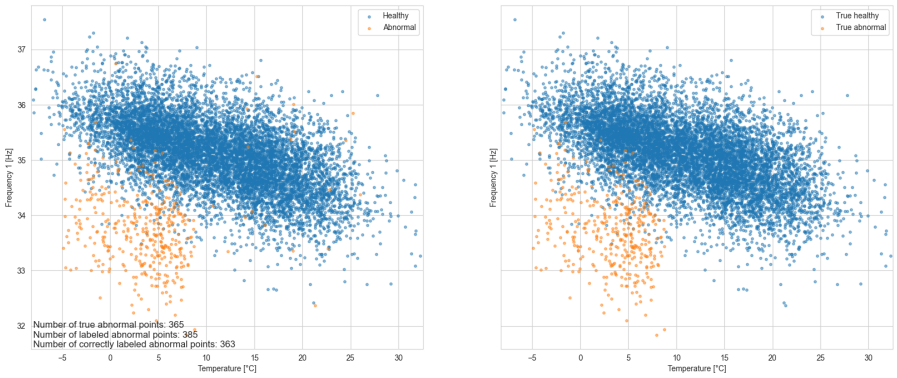


GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 11

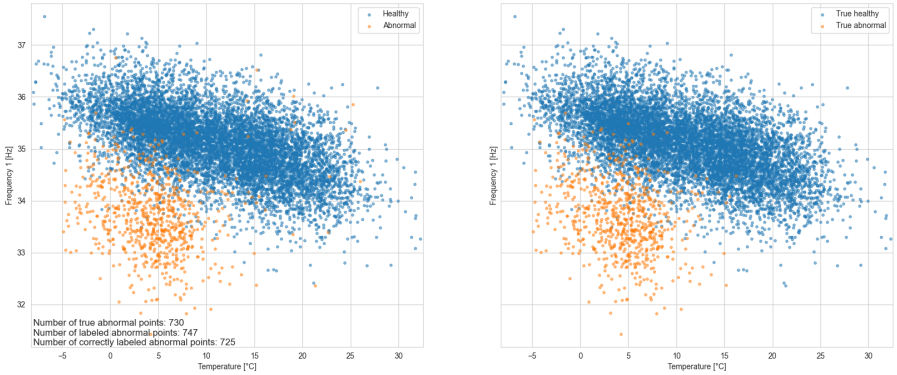


GMM-based anomaly detection results for 10% gradual stiffness reduction - Months: 12

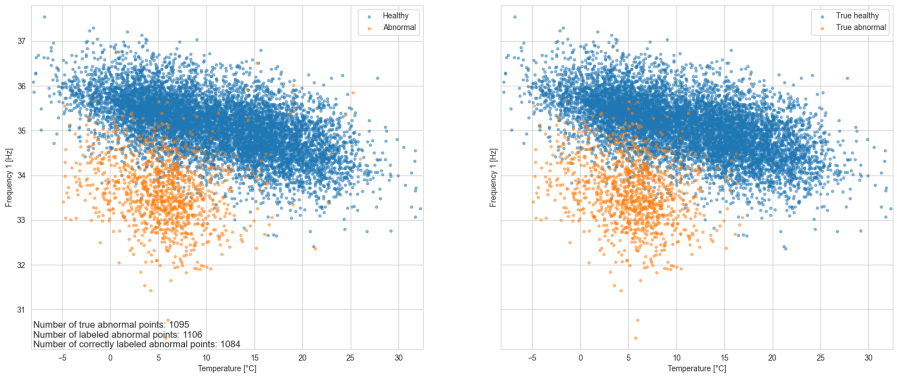
A.3.2. 10% Sudden Stiffness Loss



GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 1

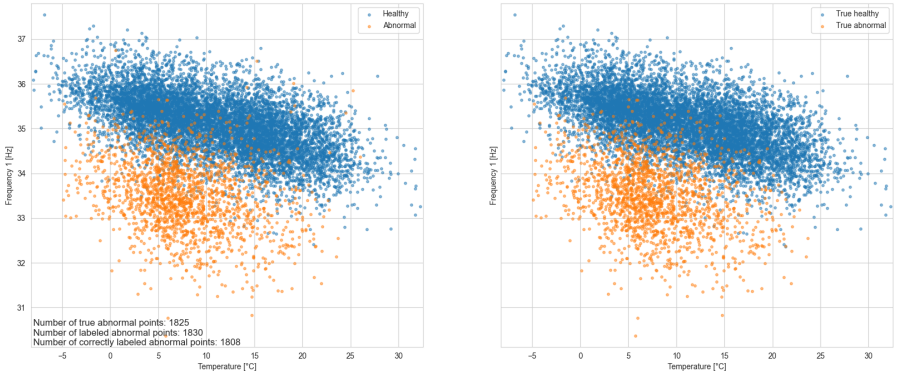


GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 2

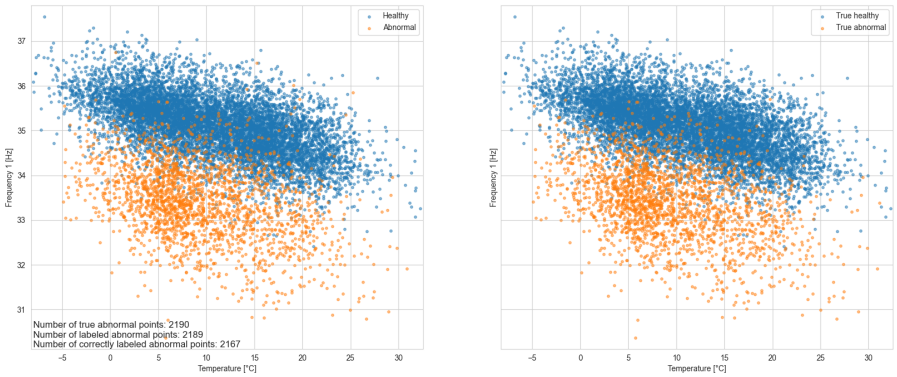


GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 3

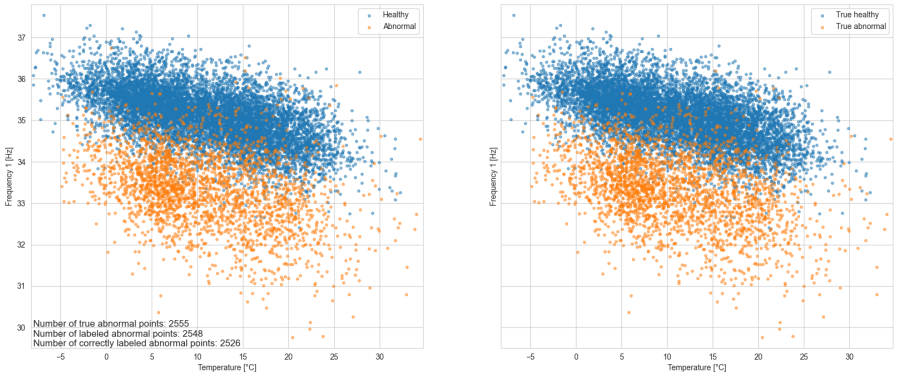
GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 4



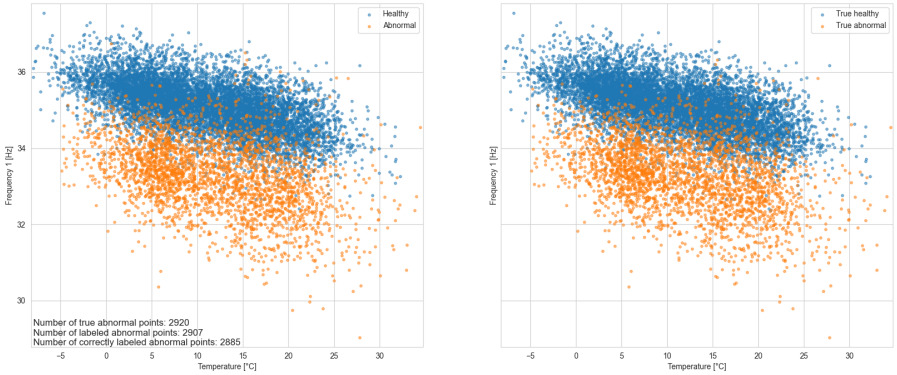
GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 5



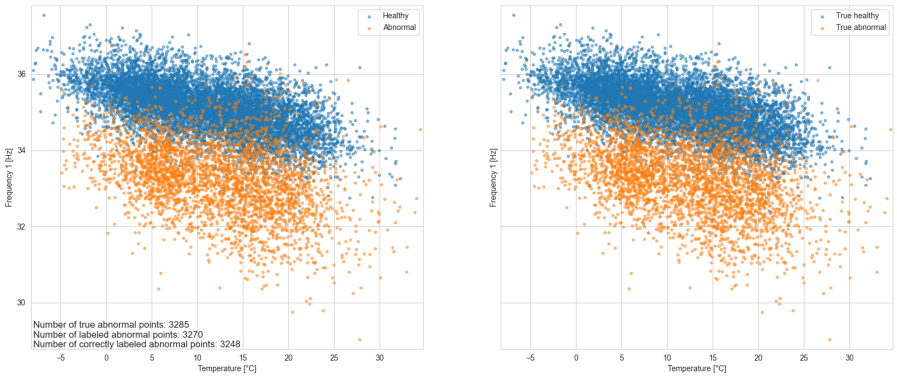
GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 6



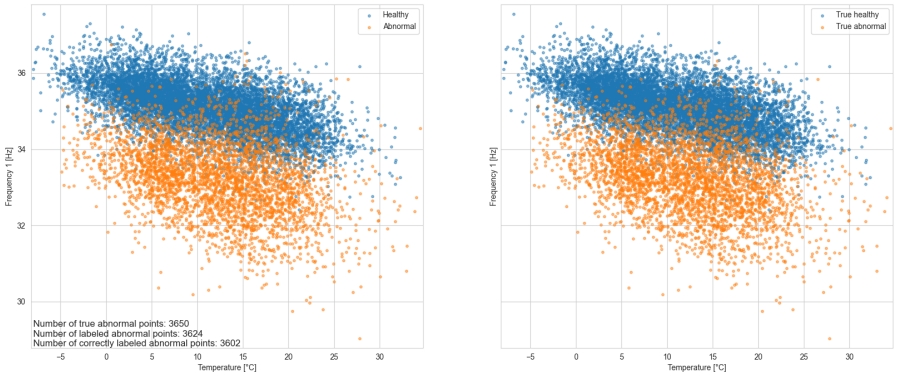
GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 7



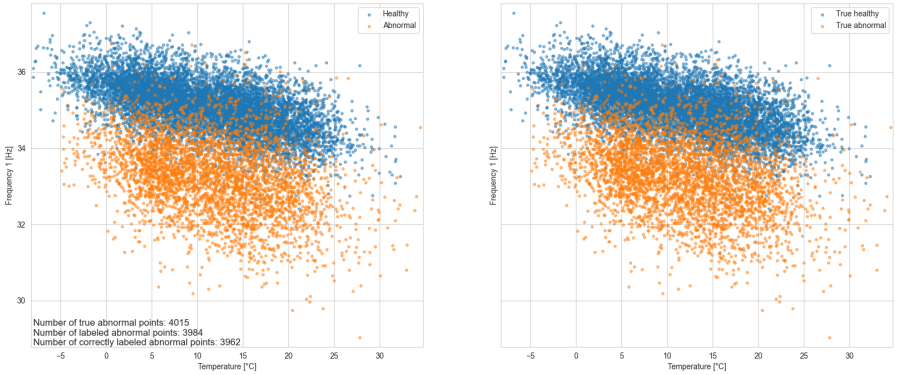
GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 8



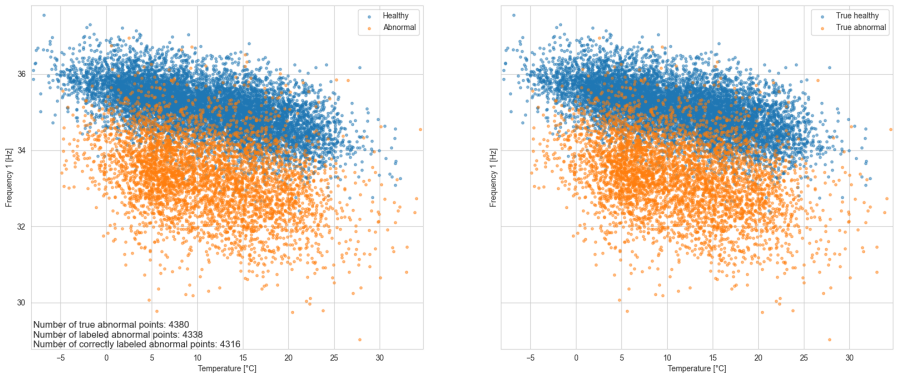
GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 9



GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 10

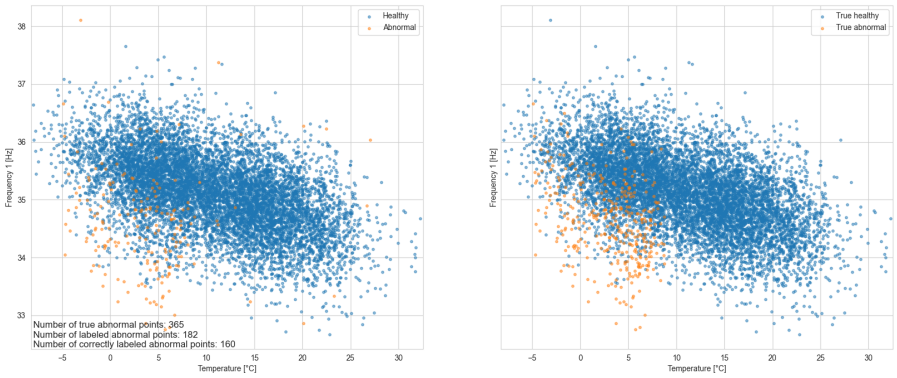


GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 11

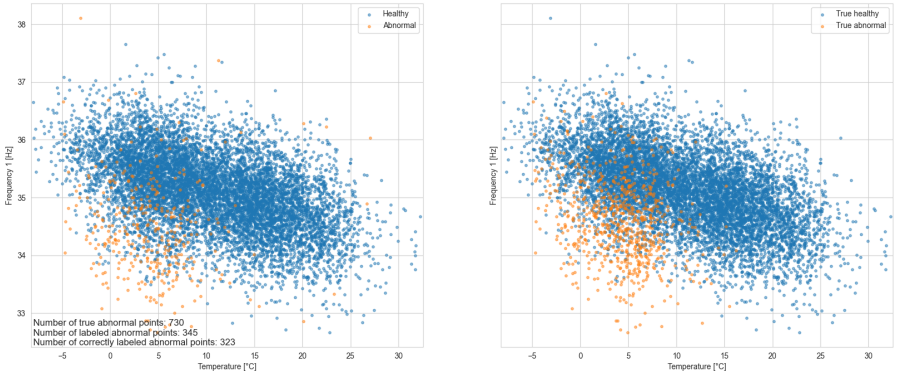


GMM-based anomaly detection results for 10% sudden stiffness reduction - Months: 12

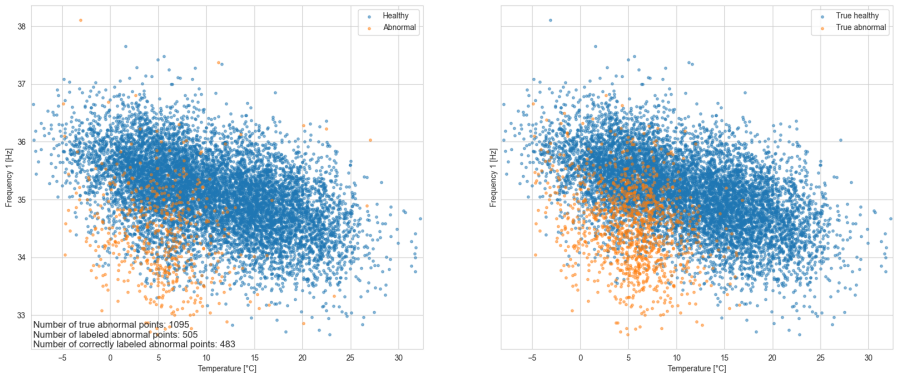
A.3.3. 5% Sudden Stiffness Loss



GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 1

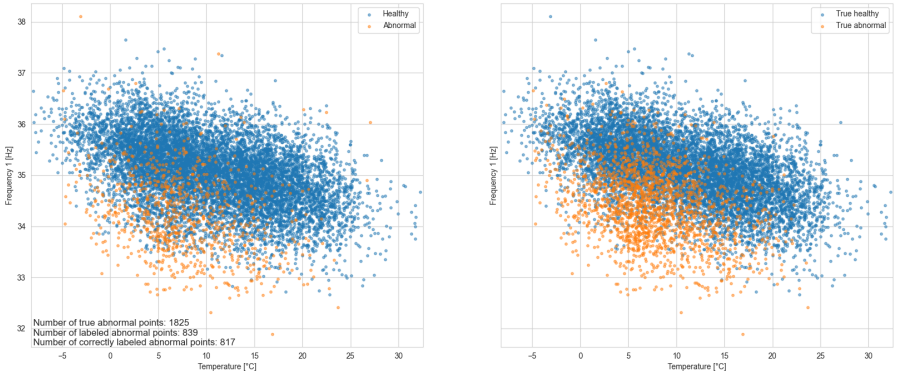


GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 2

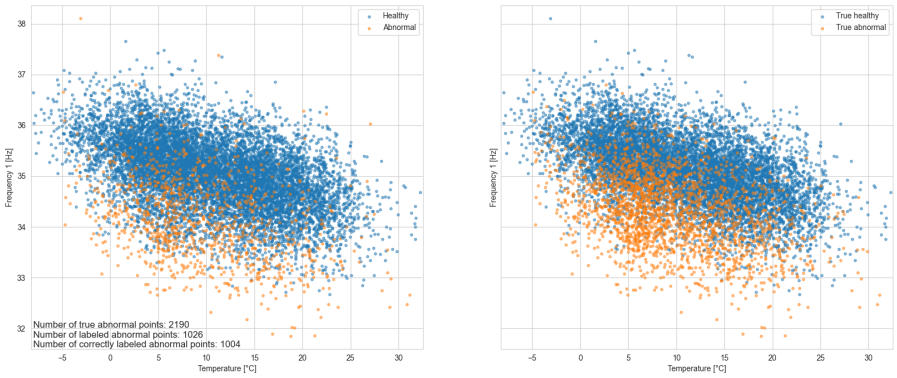


GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 3

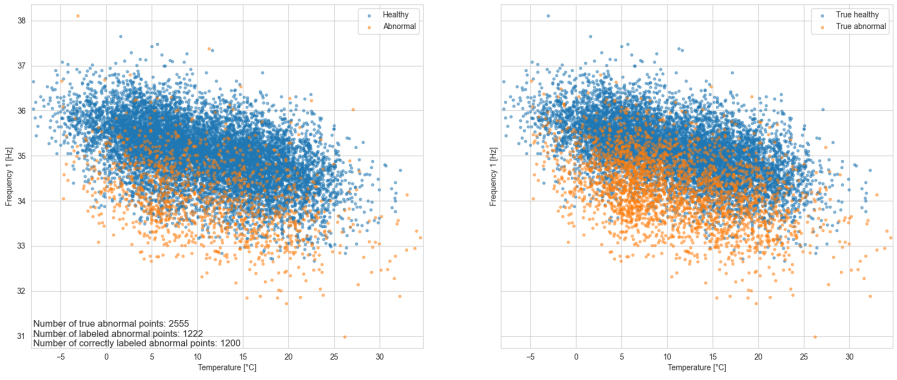
GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 4



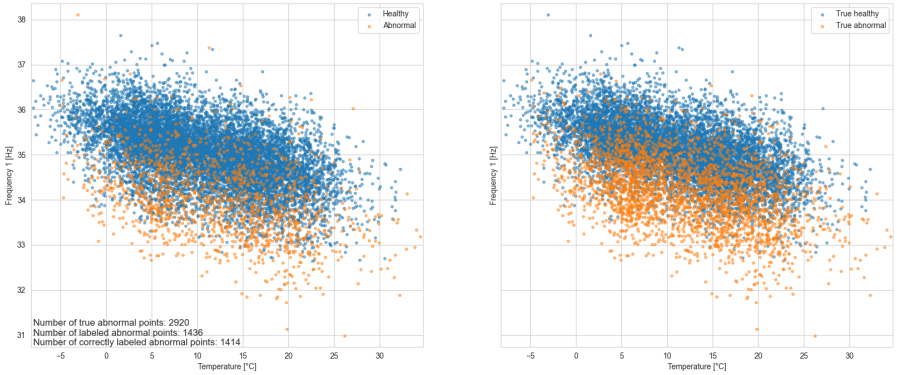
GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 5



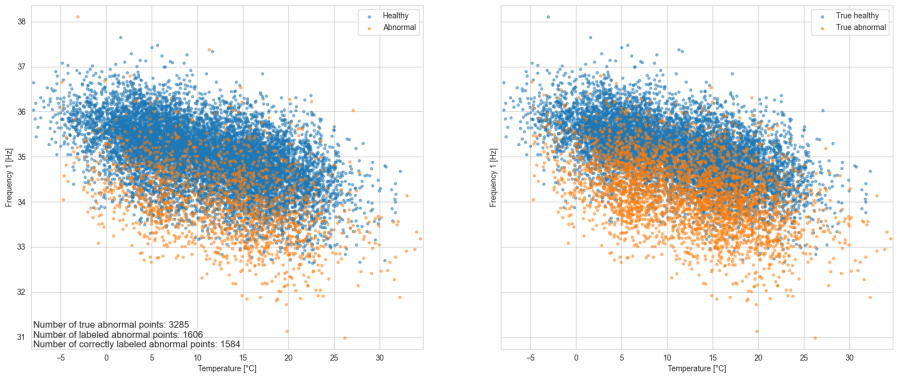
GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 6



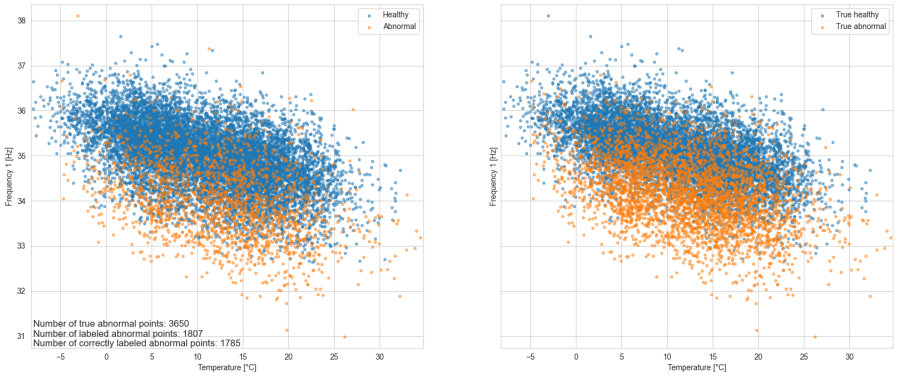
GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 7



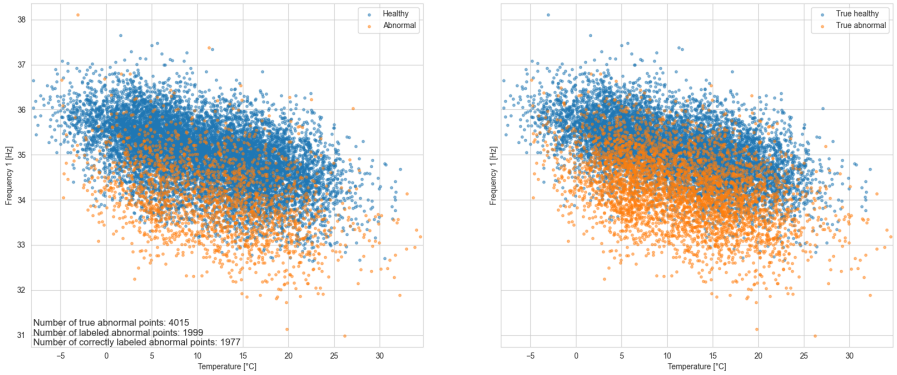
GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 8



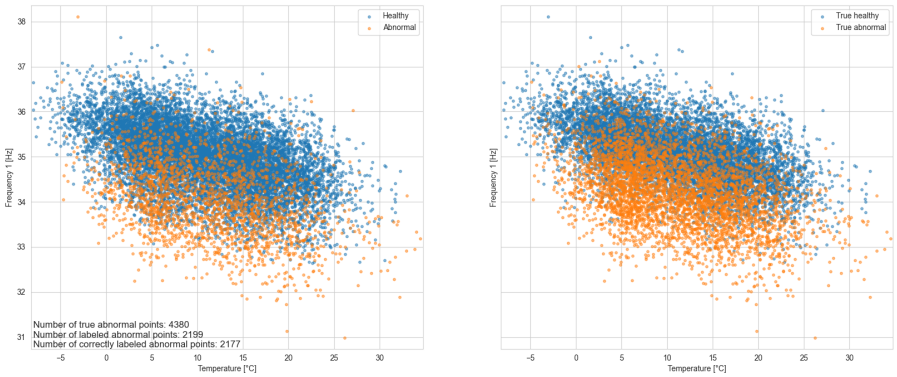
GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 9



GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 10



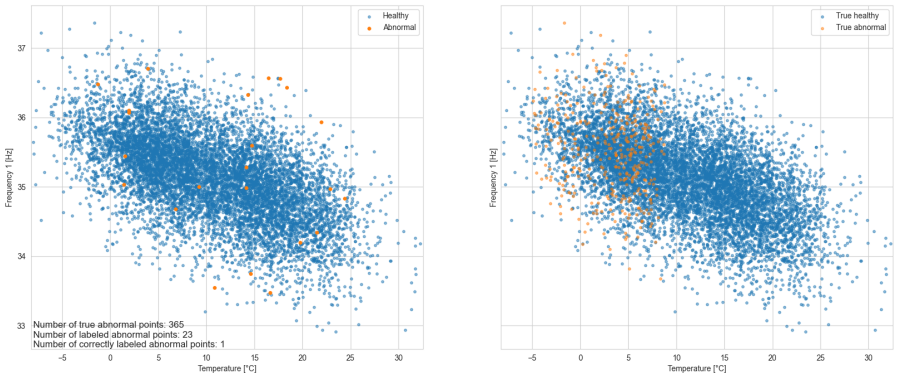
GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 11



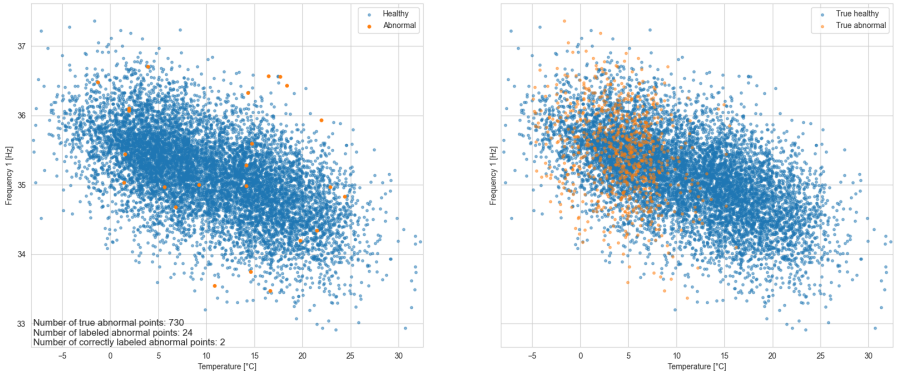
GMM-based anomaly detection results for 5% sudden stiffness reduction - Months: 12

A.4. Case Study III: Stiffness Reduction & Settlement

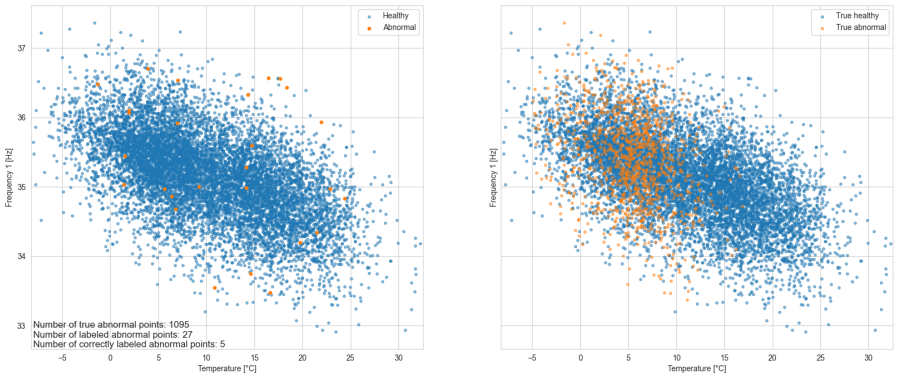
A.4.1. 150 mm Gradual Settlement and 3% Gradual Stiffness Loss



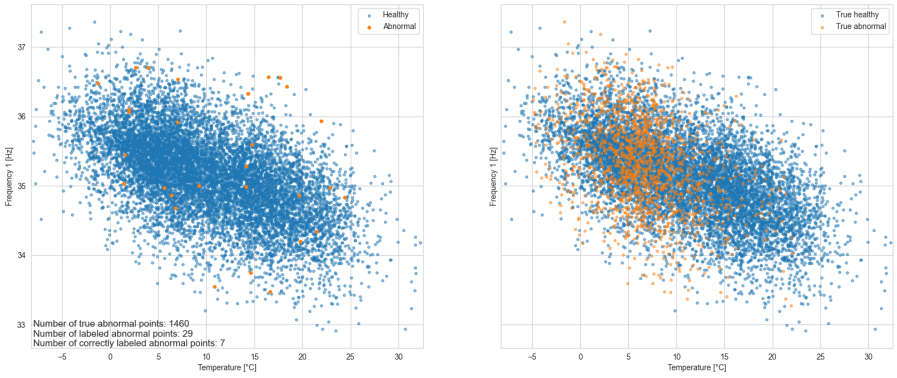
GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 1



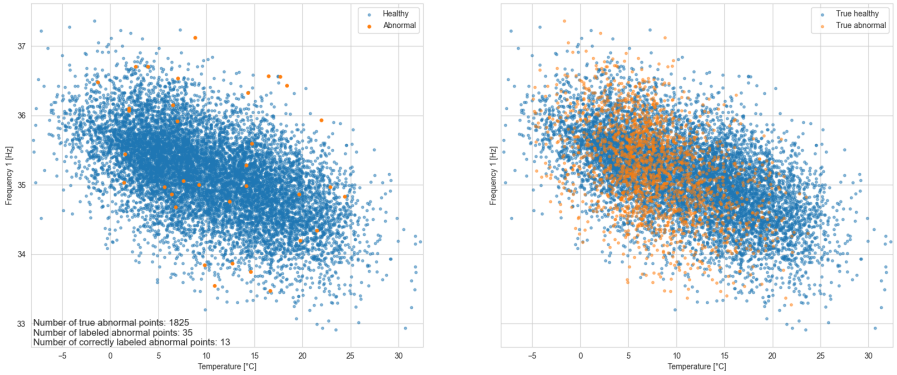
GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 2



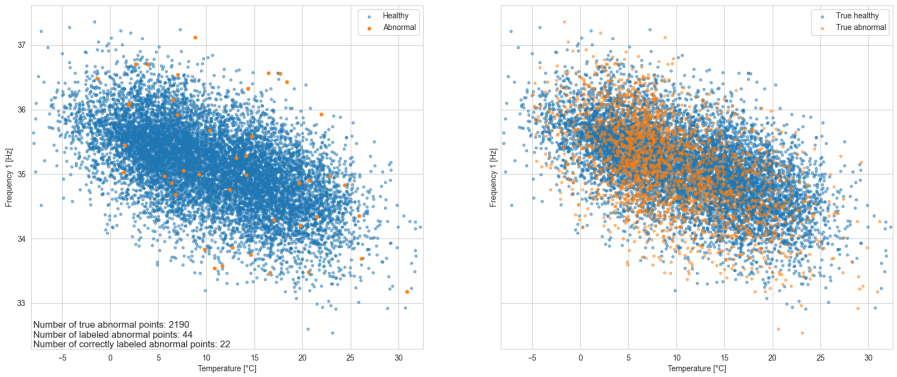
GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 3



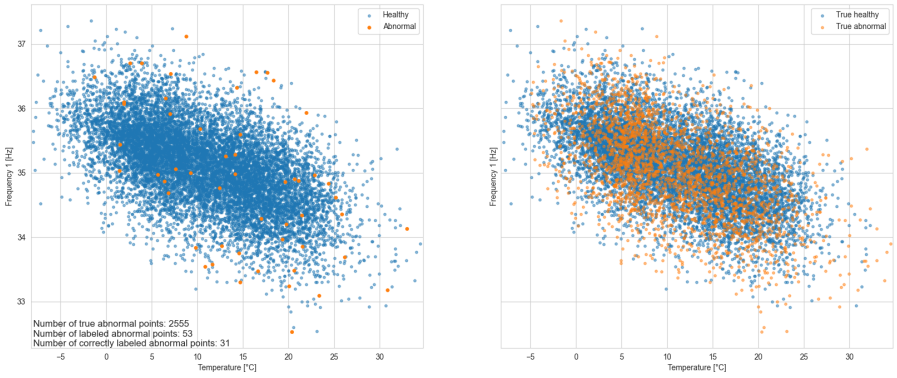
GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 4



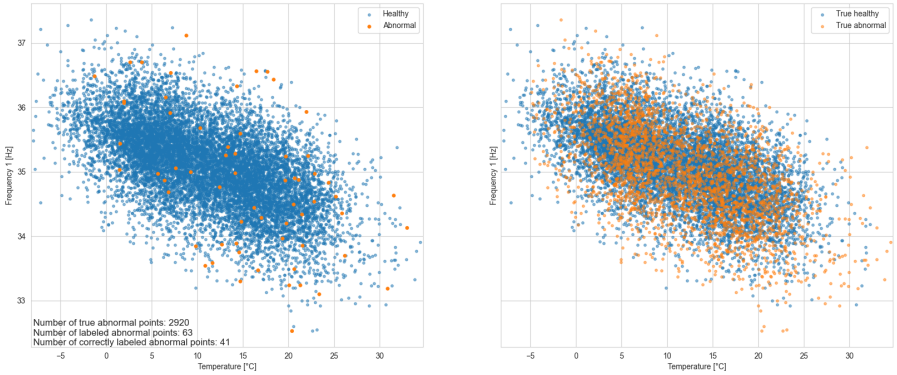
GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 5



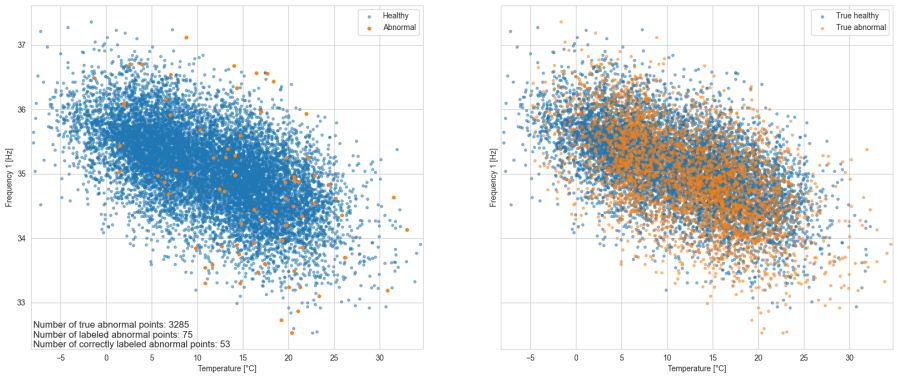
GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 6



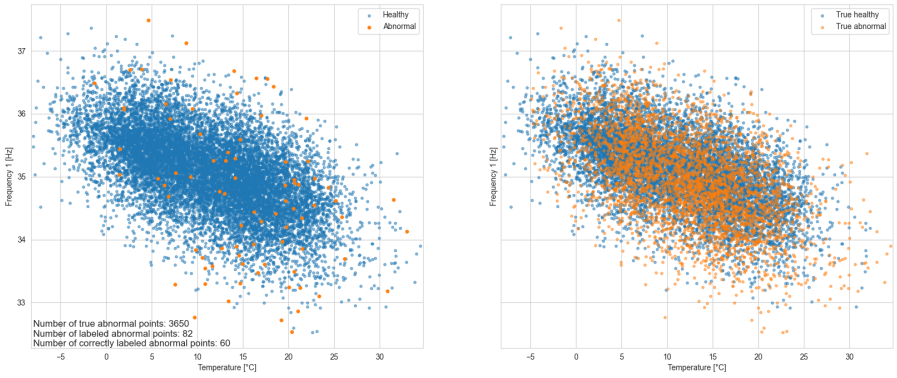
GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 7



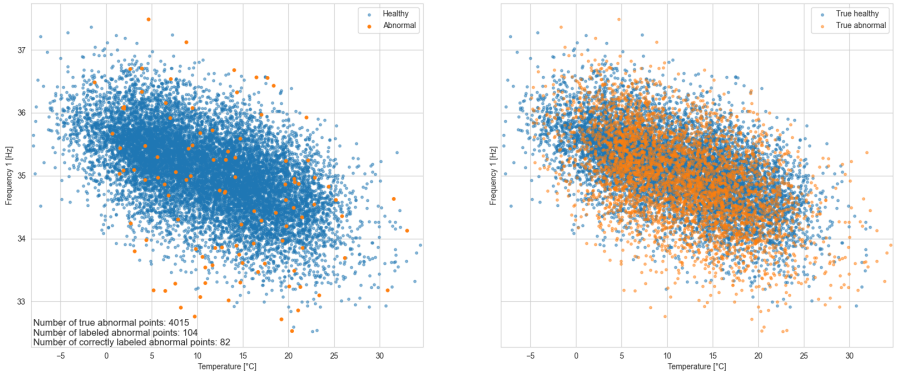
GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 8



GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 9



GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 10

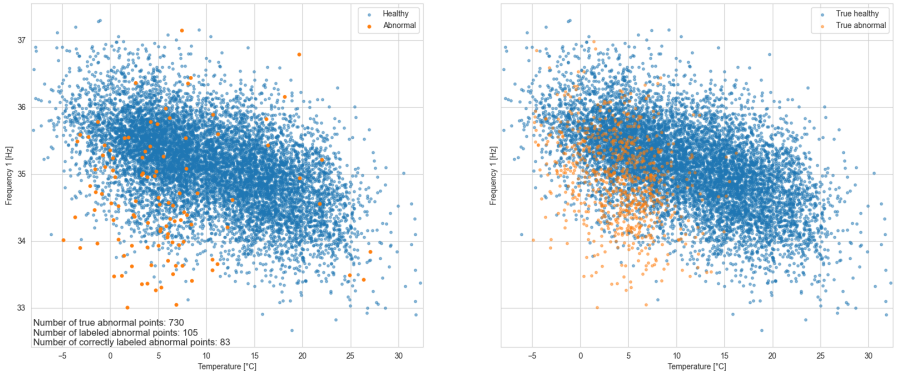


GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 11

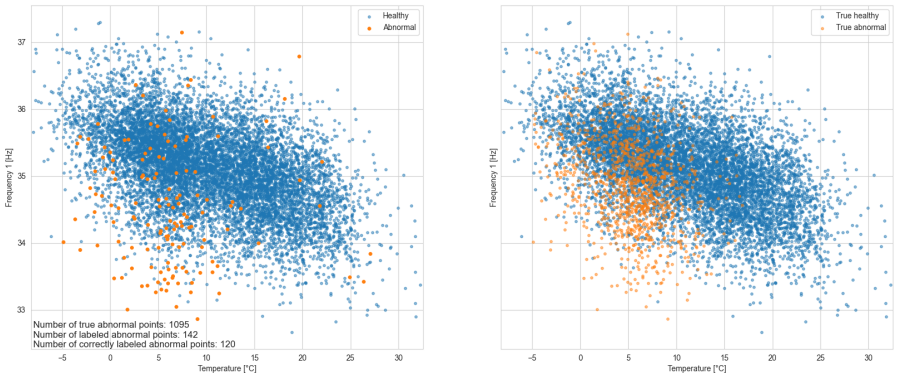
GMM-based anomaly detection results for 150 mm gradual settlement and 3% gradual stiffness reduction - Months: 12

A.4.2. 150 mm Gradual Settlement and 3% Sudden Stiffness Loss

GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 1

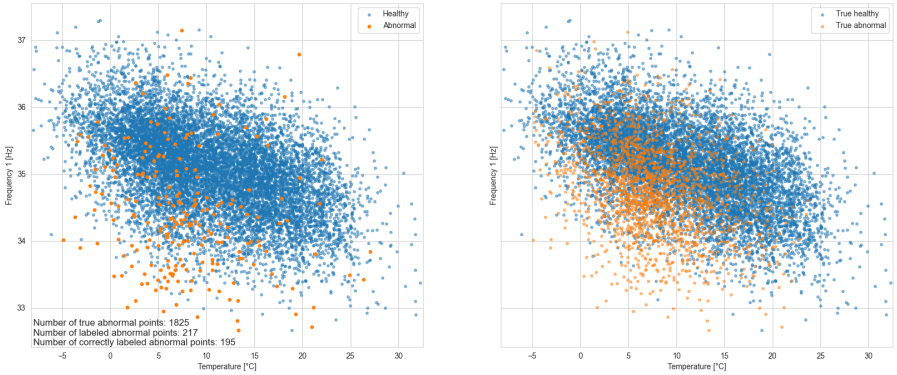


GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 2

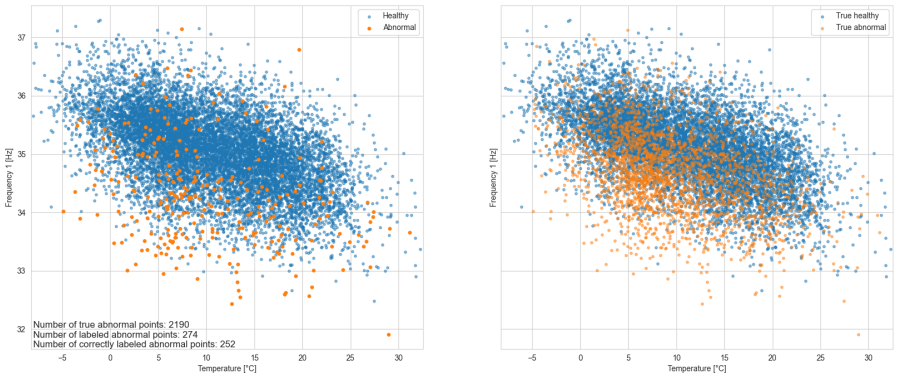


GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 3

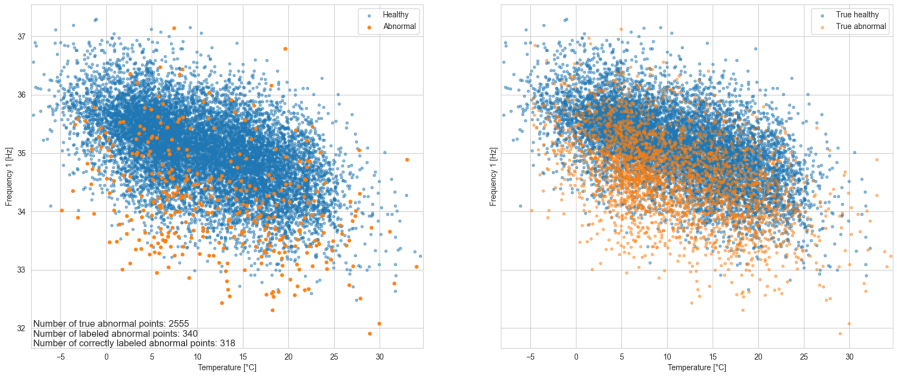
GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 4



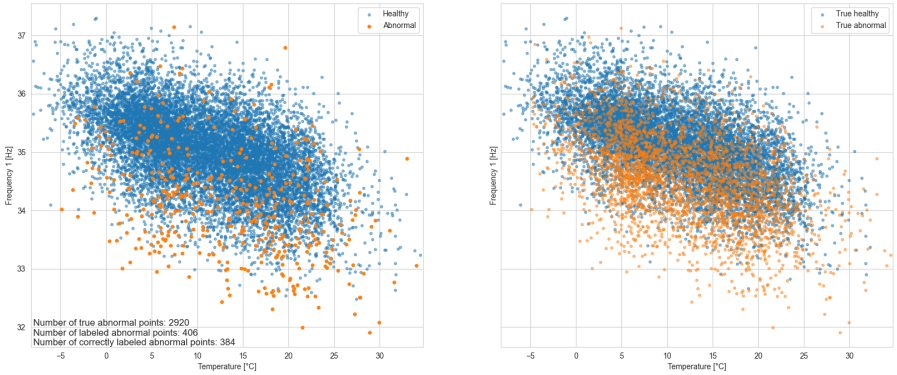
GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 5



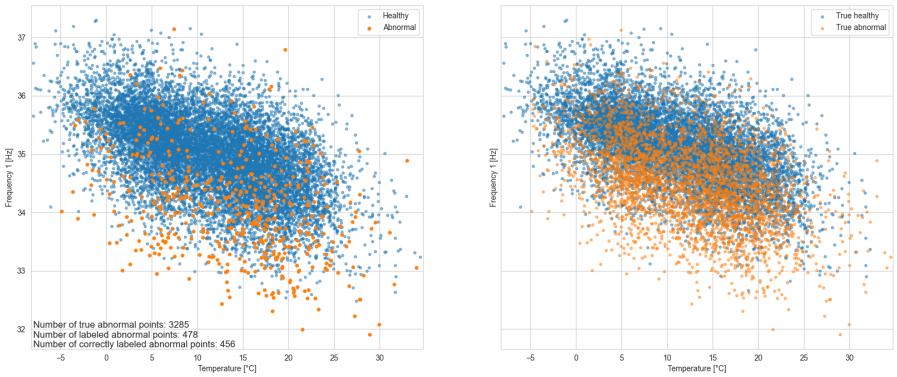
GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 6



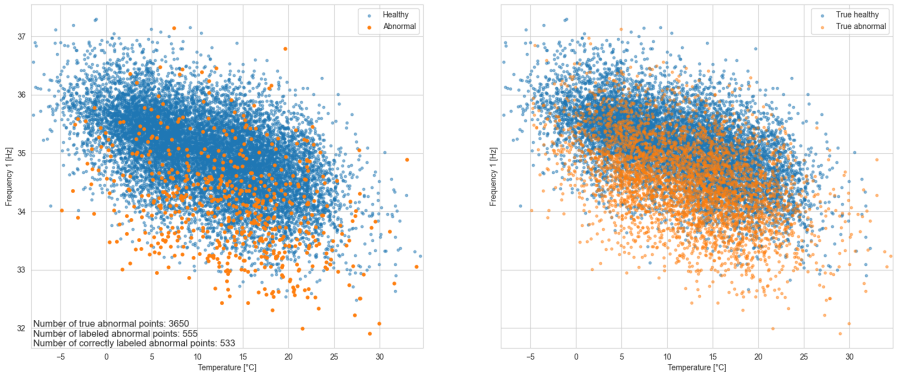
GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 7



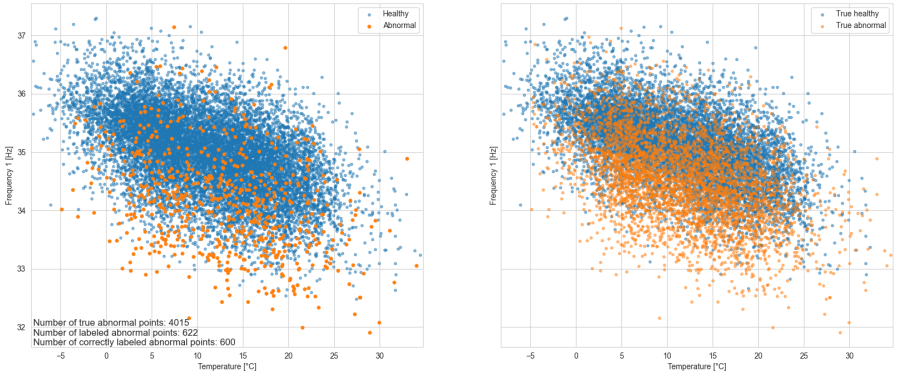
GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 8



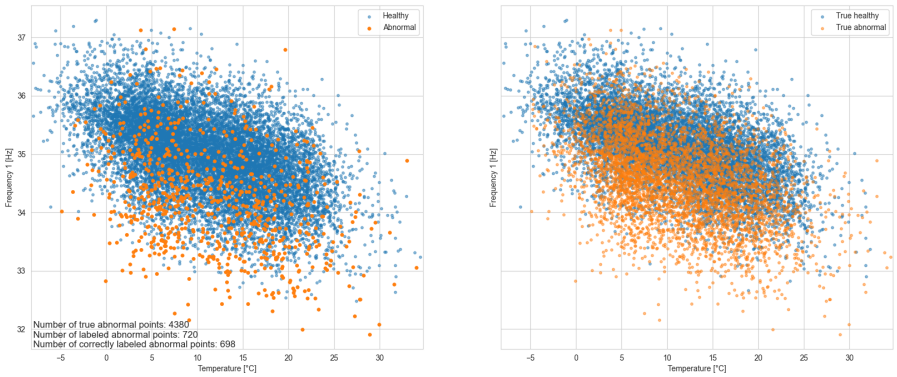
GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 9



GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 10



GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 11



GMM-based anomaly detection results for 150 mm gradual settlement and 3% Sudden stiffness reduction - Months: 12