REAL-TIME STRUCTURAL HEALTH MONITORING USING STATISTICAL METHODS

by

Emrullah Dar

B.S., Civil Engineering, Istanbul University, 2010

M.S., Earthquake Engineering, Boğaziçi University, 2015

Submitted to the Kandilli Observatory and Earthquake Research Institute
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Graduate Program in Earthquake Engineering

Boğaziçi University

2023

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor Prof. Eser Çaktı for her full support, expert guidance, understanding, and encouragement throughout my study and research. In addition, I express my appreciation to Prof. Erdal Şafak, Assoc. Prof. Ufuk Hancılar and Assist. Prof. Karin Şeşetyan for their assistance with academic research.

I would like to express my profound gratitude to my wife, Vildan Nur Dar, for her constant support and guidance throughout this process. I am also grateful to my son Ahmet Kerem Dar for his understanding and patience while I was working on my thesis. I owe a special debt of thanks to my dear uncle Muhittin Dar for his lifelong presence and encouragement. I also appreciate the diligent work of Ahmet Korkmaz on the data collection. I am fortunate to have such helpful and kind friends.

Finally, I would like to thank my mother, father, and sisters for their unconditional love and support during this study and in all my experiences.

ABSTRACT

REAL-TIME STRUCTURAL HEALTH MONITORING USING STATISTICAL METHODS

The detection of structural damage relies on understanding the long-term variation of modal parameters and their relationship to changes in atmospheric conditions. This thesis aims to address this challenge by developing a real-time algorithm for structural health monitoring systems, which are becoming increasingly important. The algorithm uses statistical models developed by analyzing four years of modal frequencies, damping ratios, and mode shapes of Hagia Sophia, a UNESCO World Heritage structure, and their correlation with atmospheric parameters such as temperature, humidity, and wind speed.

The algorithm uses four different regression models to predict the modal frequency as a function of the atmospheric conditions and selects the most suitable one. It then compares the predicted and measured frequencies to identify structural anomalies. The algorithm also employs the Modal Assurance Criterion (MAC), Coordinate Modal Assurance Criterion (COMAC), and Enhanced Coordinate Modal Assurance Criterion (ECOMAC) methods to examine the long-term variation of mode shapes. The algorithm is implemented in a user interface software called "AISHM," which displays the modal parameters and the 3-D animation of the structure in real-time. The software also has the capability to track earthquakes and analyze the structural response in real-time.

In summary, this thesis presents a comprehensive approach to real-time structural health monitoring using statistical models and advanced analysis techniques, which can have significant implications for maintaining and preserving historical structures.

ÖZET

İSTATİSTİKSEL YÖNTEMLERLE GERÇEK ZAMANLI YAPI SAĞLIĞI İZLEME

Yapısal hasarın tespiti, modal parametrelerin uzun vadeli değişimini ve bunların atmosferik koşullardaki değişikliklerle ilişkisini anlamaya dayanır. Bu tez, giderek daha önemli hale gelen yapısal sağlığı izleme sistemleri için gerçek zamanlı bir algoritma geliştirerek bu zorluğu ele almayı amaçlamaktadır. Algoritma, bir UNESCO Dünya Kültür Mirası yapısı olan Ayasofya'nın dört yıllık modal frekanslarını, sönüm oranlarını ve mod şekillerini ve bunların sıcaklık, nem ve rüzgar hızı gibi atmosferik parametrelerle korelasyonunu analiz ederek geliştirilen istatistiksel modelleri kullanmaktadır.

İlk olarak Ayasofya'nın 2013-2016 yılları arasındaki modal frekansları, modal sönümleme oranları ve mod şekilleri incelenmiştir. Bu dinamik parametreler ile sıcaklık, nem, rüzgâr hızı gibi atmosferik parametreler arasındaki korelasyon değerlendirilmiştir. Modal frekans ve atmosferik parametreler arasındaki korelasyon, doğrusal regresyon analizi kullanılarak modellenmiştir. Mod şekillerinin varyasyonu, Modal Güvence Ölçütü (MAC), Koordinat modal Güvence Ölçütü (COMAC) ve Gelişmiş Koordinat Modal Güvence Ölçütü (ECOMAC) yöntemleri kullanılarak analiz edilmiştir.

İkinci olarak, istatistiksel modellerle desteklenen gerçek zamanlı bir yapısal sağlık izleme sistemi algoritması geliştirilmiştir. Bu algoritma, yapının tahmin edilen ve ölçülen modal frekansını gerçek zamanlı olarak görüntülemektedir. Ayrıca modal sönüm oranları ve mod şekillerindeki (MAC, COMAC, ECOMAC) değişiklikler de gerçek zamanlı olarak hesaplanmaktadır. Algoritma, yapının gerçek zamanlı 3-D animasyonunu da görüntüleyen "AISHM" adlı bir kullanıcı arayüze sahip bir yazılıma dönüştürülmüştür. Bu yazılım aynı zamanda depremleri algılar ve deprem kayıtlarını analiz eder.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

f	Modal Frequency
P_{xy}	Cross Spectral Density Function
R_{xy}	Cross-Correlation Function
w	Window Function
ε	Error
ξ	Modal Damping Ratio
τ	Time index
ϕ	Mode Shape
ϕ_n	Mode shape at n^{th} window
ϕ_r	Reference mode shape
$\widetilde{\phi}$	Normalized Mode Shape
ω	Angular Frequency

LIST OF ACRONYMS/ABBREVIATIONS

3-D	Three Dimensional
AISHM	Artificial Intelligence in Structural Health Monitoring
ANN	Artificial Neural Network
API	Application Programming Interface
CPSD	Cross-Power Spectral Density
COMAC	Coordinate Modal Assurance Criterion
CWT	Continuous Wavelet Transform
DFT	Discrete Fourier Transform
DNN	Deep Neural Networ
ECOMAC	Enhanced Coordinate Modal Assurance Criterion
FDD	Frequency Domain Decomposition
HPBW	Half-Power Band Width
MAC	Modal Assurance Criterion
MGM	the General Directorate of Meteorology of Turkey
MLR	Multiple Linear Regression
OMA	Operational Modal Analysis
RT	Regression Trees
SHM	Structural Health Monitoring
SNN	Shallow Neural Network
SSI	Soil Structure Interaction
STFT	Short-Time Fourier Transform
SVD	Singular Value Decomposition
SVM	Support Vector Machines

1. INTRODUCTION

1.1. Overview

The safety and reliability of buildings, bridges, and other structures are crucial. To ensure this, Structural Health Monitoring (SHM) has become increasingly important. It involves the continuous or periodic monitoring of a structure's response to environmental loads such as wind, temperature, and seismic activity. The purpose is to detect any changes in behavior that may indicate structural damage or deterioration. The long-term variation of modal parameters and their relationship to changes in atmospheric parameters are of immense importance for the accurate detection of structural damage. Therefore, there is a growing interest in developing real-time SHM systems that can provide immediate feedback on structural performance and detect potential damage before it becomes critical.

This thesis introduces an algorithm for Structural Health Monitoring (SHM) that is based on statistical models and uses data from the Hagia Sophia, a UNESCO World Heritage Site. The dynamic parameters of this monument between 2013 and 2016, including modal frequencies, modal damping ratios, and mode shapes are investigated. The study examined the relationship between dynamic parameters and atmospheric factors, including temperature, humidity, and wind speed. Using this correlations, regression models that can predict the modal frequency of the structure based on atmospheric conditions are developed using statistical methods. This models are used to detect any potential structural issues in real-time.

The preservation of the historical and cultural heritage and its transmission to the next generations is an especially important concern to preserve the intellectual wealth of modern society. This task is quite difficult because of the complex structures and material properties of historical buildings. Therefore, the fast and reliable detection of structural anomalies in historical monuments is of crucial importance for

their preservation as part of the efforts to pass them on to the next generation. The damage and consequent repairs they have suffered in the historical process often complicate the dynamic behavior of historical structures. The study of long-term changes in modal frequencies, modal damping ratios, and mode shapes allows a more accurate interpretation of the dynamic behavior of such structures [1–8]

In historical edifices, structural health monitoring (SHM) systems are mainly implemented in mosques, cathedrals, bell towers, and minarets [1–7, 9–13] Changes in modal frequency and drift ratio are usually the two main parameters used for condition assessment and damage detection in such structures [10]. However, these parameters can provide misleading information about the health of the structure in some cases. [4] Therefore, examining changes in modal damping ratio and mode shapes increases the accuracy of damage estimates and condition assessments. Increasing the number of acceleration and atmospheric sensors allows for more accurate and faster detection of structural anomalies [3, 9].

Permanent changes in the dynamic parameters and shifts in structures are often associated with damage [14, 15]. However, these parameters are not only effected by damage but also by temperature fluctuations [1, 2, 6, 10, 16–18], humidity [7, 10, 19–21], wind speed [9, 10, 18, 22], and ground motion [2, 23]. In historical structures, an increase in temperature usually increases the frequency of linear modes and decreases the frequency of torsional modes [1, 2, 7, 13, 19, 24–26]. This is thought to be related to the opening and closing of existing cracks because of thermal expansion and contraction [1, 6]. Under freezing conditions, i.e., when the temperature falls below $+4\ ^{0}$C, the material hardens, and the rigidity increases with the consequence of an increasing modal frequency [1,7,10]. It has also been observed that in historic structures reinforced with tension bars the rise in temperature causes metal to expand, which increases the widths of existing cracks, and the modal parameters vary depending on the position and shape of the tension bars [3].

It has been found that, in some historical structures, changes in humidity ratio

may be more effective than the temperature on modal parameters [21]. In cases of insufficient waterproofing, rainwater intrusion affects the modal parameters of historic monuments [20,27]. Strong winds cause an earthquake-like effect on structures by temporarily decreasing the modal frequency, increasing the damping ratio, and changing the mode shape [9, 22]. Considering the environmental factors and their impact on structural parameters [16, 17], these must be considered for reliable system identification and damage prediction.

Historical structures form a group with quite different load-bearing systems and material properties. Their age; their level and time of exposure to different natural hazards; and their damage, repair, and restoration histories are quite different from modern buildings. Due to this complexity, atmospheric factors may have distinct impacts on historic structures compared to modern structures. Therefore, the relationship between atmospheric factors and modal parameters varies depending on the characteristics of the structure, and each structure has a different degree of sensitivity to atmospheric factors.

This thesis explores the relationship between the long-term modal parameter variations of the Hagia Sophia in Istanbul and the environmental factors such as temperature, wind speed, relative humidity and precipitation. The modal frequencies, modal damping ratios, and mode shapes are identified using continuous acceleration data recorded over the four years between 2013 and 2017. The dependencies between the modal parameters and environmental factors in this period are using statistical correlation models and approaches. To estimate modal frequency in real time, regression models such as Multiple Linear Regression (MLR) model, Regression Trees (RT) model , Support Vector Machines (SVM) model , and Artificial Neural Network (ANN) regression model used to get predictive relationships between natural frequencies and environmental parameters. Finally, a real-time structural health monitoring system software called "AISHM" was developed to monitor all modal parameters and statistical predictions in real-time.

1.2. Research Questions

This thesis is centered on addressing the following research questions:

- How do the dynamic properties of Hagia Sophia vary over time, and what is their relationship with changes in atmospheric parameters?
- Can regression models accurately estimate the modal frequency of the structure based on atmospheric parameters?
- How can changes in modal frequencies be used to detect structural anomalies in real-time?
- How do the mode shapes of Hagia Sophia vary over time, and what is the significance of these variations for structural health monitoring?
- Can the Modal Assurance Criterion, Coordinate Modal Assurance Criterion and Enhanced Coordinate Modal Assurance Criterion methods be used to analyze the long-term variation of mode shapes?
- How can a real-time structural health monitoring algorithm be developed and validated using statistical models?
- How does the developed software, "AISHM", perform in detecting earthquakes and analyzing the response of the building during seismic events?

These research questions aim to investigate the long-term variation of modal parameters of Hagia Sophia and their relationship with environmental conditions, as well as the development and validation of a real-time structural health monitoring system using statistical models. The questions also explore the significance of mode shape variations and the performance of the developed software in detecting earthquakes and

analyzing the building's response during seismic events.

1.3. Objective

The central aim of this study is to identify structural anomalies rapidly and precisely within a structure in real-time, utilizing statistical techniques. To reach this goal, it is essential to fulfill certain underlying objectives. These sub-objectives can be outlined as follows:

- To analyze the long-term variation of modal frequencies, modal damping ratios, and mode shapes of Hagia Sophia between 2013 and 2016.

- To evaluate the correlation between the dynamic parameters of Hagia Sophia and atmospheric parameters, such as temperature, humidity, and wind speed.

- To develop a statistical model to estimate the modal frequency of Hagia Sophia based on atmospheric parameters.

- To develop a real-time structural health monitoring algorithm for Hagia Sophia using statistical models, which displays the structure's predicted and measured modal frequency, modal damping ratios, and mode shapes in real-time.

- To convert the real-time algorithm into the software with a user interface called "AISHM", which also displays real-time 3-D animation of the structure and detects earthquakes.

- To validate the developed algorithm and software using field data and compare the results with existing structural health monitoring systems.

- To investigate the significance of mode shape variations using the Modal Assurance Criterion (MAC), Coordinate Modal Assurance Criterion (COMAC), and

Enhanced Coordinate Modal Assurance Criterion (ECOMAC) methods .

- To demonstrate the effectiveness of the developed algorithm and software in detecting structural anomalies and analyzing the response of the building during seismic events.

These objectives aim to achieve a comprehensive understanding of the long-term variation of modal parameters of Hagia Sophia and their relationship with environmental conditions, as well as the development and validation of a real-time structural health monitoring system using statistical models. The objectives also focus on investigating the significance of mode shape variations and demonstrating the effectiveness of the developed algorithm and software in detecting structural anomalies and analyzing the response of the building during seismic events.

1.4. Brief History and Description of the Structure and the Structural Health Monitoring System

Hagia Sophia is one of the most important structures in the world, culturally, religiously, and architecturally (Figure 1.1) [28]. It is located in Istanbul, Turkey, and constructed as an Eastern Orthodox cathedral in 537. Afterwards, in 1453, it was converted to a mosque, and then to a museum in 1935. In 2020, it was reconverted into a mosque. Hagia Sophia is one of the most important architectural masterpieces in the world and has undergone numerous modifications and renovations throughout its history. It has survived many earthquakes and atmospheric disasters due to its robust structural design.

The construction of Hagia Sophia commenced in 532 and was completed within a span of 70 months. With a dome diameter of thirty-one meters, it was the largest domed structure in the world for eight hundred years. In 558, a massive portion of its dome collapsed during an earthquake, prompting the need for reconstruction. The new dome was 6.25 meters higher than before [29, 30]. In this way, the height of the

structure reached its current height of 56 meters [23]. Hagia Sophia was inscribed on the World Heritage List in 1985 for its historical significance, which spans cultures and religions, and its architectural features.

Hagia Sophia has a similar layout to a basilica. The primary structural support system from the bottom to the top comprises four main piers, four secondary piers, and four buttress piers that support the structure on the northern and southern sides. The level at which the four main arches spring from the four main piers also marks a change in structural materials, as described in the next paragraph. The four main arches, the northern and southern ones being double arches, are supported by two semi-domes on the eastern and western sides and by the four buttress piers on the northern and southern sides. This system, along with the pendentives on the four corners, provides support for the main dome.

Underlying the Hagia Sophia are rocks estimated to be at about 1 m on the north side and about 3 m on the south side [31]. It is a masonry structure, with main and secondary piers being of stone masonry, while all arches and domes are made of brick masonry. There are significant deformations reaching 85 cm and 16 cm in the N-S and E-W directions, respectively, at the top of the main piers. The east and west arches are inclined inward by 14 and 15 cm respectively, in response to the thrust of the half-domes. The crowns of the north and south arches are 60 and 57 cm outwardly inclined, respectively. The vertical deformation at the crowns of the north and south arches, i.e. deviation from the ideal shape of an arch, is 9 cm and 18 cm, while the crowns of the east and west arches have deformations in the gravitational direction of 26 cm and 13 cm respectively [32, 33].

Figure 1.1. South-east view of Hagia Sophia.

The eastern arch of Hagia Sophia was destroyed during earthquakes in 558 and 1346, while the western arch partially collapsed in an earthquake in 989 and was subsequently rebuilt [15, 23, 34]. The long duration of acquiring strength of the mortar material used and the fleeting period between the start of construction and the inauguration led to deformations of the main piers that were already visible during the construction period. For this reason, the north and south main arches were strengthened by implementing double arches [32, 35]. Between 1573 and 1576, architect Sinan added supports around the structure. Four of these were built to support the buttress piers on the northern and southern sides to counteract their outward displacements [34].

The structural health monitoring system of Hagia Sophia (Figure 1.2) was established in August 1991 with nine three-component accelerometers operating in trigger mode [23]. This system was transformed into a real-time structural health monitoring system, transmitting data in continuous mode in 2012. One of the accelerometers is on the ground of the structure, four are on top of the main piers and four are on the main arches.

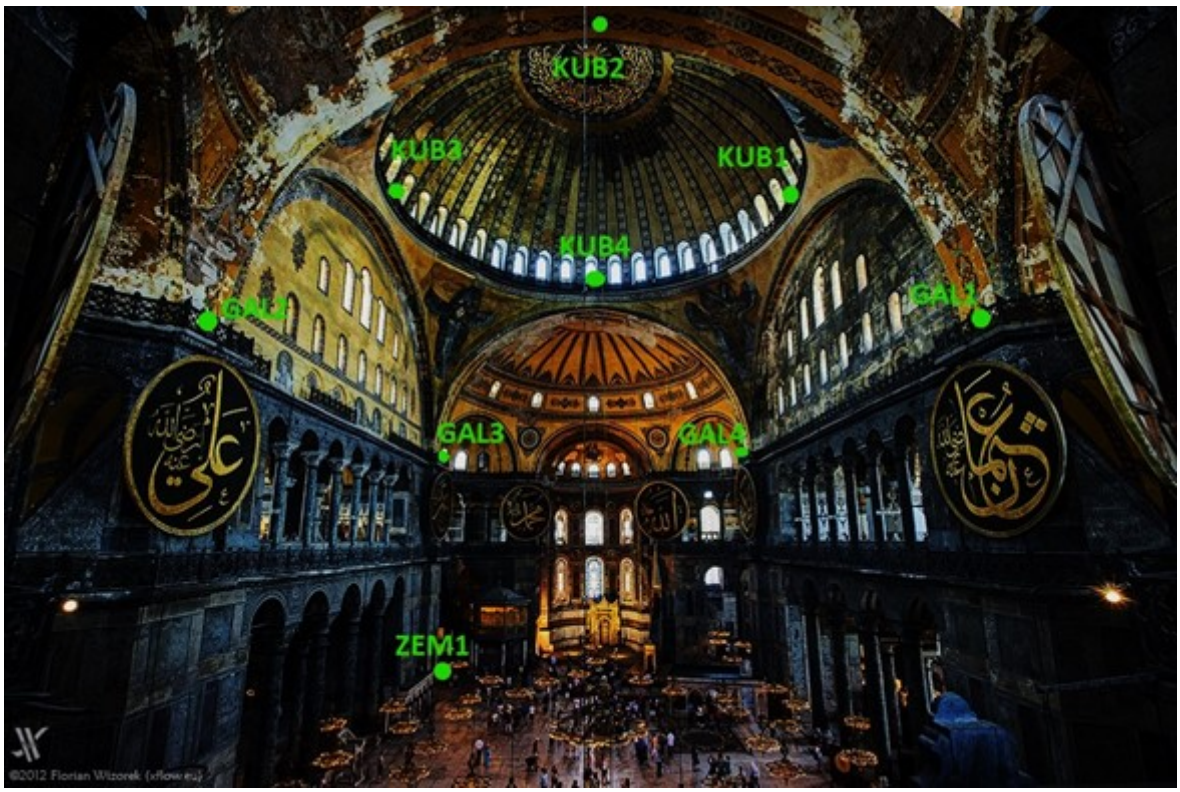

Figure 1.2. The locations of accelerometers.

The accuracy and reliability of meteorological data are essential for assessing the influence of environmental factors on structural behavior. In this study, the meteorological data was obtained from the Istanbul University Meteorological Station of the Turkish State Meteorological Service, located 1.8 km from Hagia Sophia, since the data from the meteorological stations in Hagia Sophia were not available between 2013 and 2016. The atmospheric data used in this study was assumed to be remarkably close to Hagia Sophia measurements since the Istanbul University meteorology station is close enough to Hagia Sophia.

As a result of previous research based on the analysis of accelerometer data, some anomalies regarding the dynamic response of Hagia Sophia have been identified. The main anomaly is that the lateral vibration levels of the main southwestern pier are greater than any other pier [23, 31, 36, 37]. The reason for this was hypothesized

as related to the construction material of this pier, past damages or existing cracks within the pier, or the ground conditions under the pier [23, 31]. There is more lateral movement in the east and west arches compared to other arches [23, 38–40]. Vertical displacements are larger in the eastern and western arches than in the northern and southern arches, which are comparable to their horizontal counterparts [23]. The static and dynamic response properties of Hagia Sophia were studied using finite element models [30, 32, 33, 37–44]. There was almost consensus that the eastern and/or western sides, where the semidomes meet the main dome along the main arches, are the most fragile parts of the structure [23, 31, 37, 39, 44].

1.5. Organization of Thesis

This thesis consists of six chapters. Chapter 1 presents an overview of the research problem, research questions, and objectives of the study. The chapter also provides the background and context of the study, including a brief history and description of the Hagia Sophia and the structural health monitoring system of Hagia Sophia.

Chapter 2 presents the methods of data analysis and summarizes the methodology of modal frequency estimation, modal damping estimation, mode shape estimation, and modal assurance criterion estimation.

Chapter 3 presents the results of the data analysis, including the long-term variation of modal parameters and their relationship with changes in atmospheric parameters. In this chapter, the long-term variations of the modal parameters (modal frequency, modal damping ratio, and mode shapes) are discussed.

Chapter 4 explains the regression models tested to predict modal frequencies and presents the error rates of the regression models. The chapter also discusses the criteria for choosing a suitable regression model for predicting the modal frequency in real-time.

Chapter 5 introduces the real-time structural health monitoring software called

"AISHM" and explains the features of this software. Finally, Chapter 6 summarizes and concludes the thesis.

2. DATA ANALYSIS METHODS

The four-year continuous acceleration data obtained from the SHM system of Hagia Sophia is close to 1 TB in total. Acceleration records were divided into short windows so that the time-variant parameters could be tracked. By applying modal identification algorithms in each window, modal frequency, modal damping ratio, and mode shapes were obtained. Long-term changes in modal parameters were obtained by plotting window-based modal estimations as a function of their observation time. 15-minute windows were preferred for the identification of modal frequencies, while 1-hour windows were selected for the calculation of modal damping ratios and mode shapes. Before starting the analysis, all acceleration data was calibrated, baselines were corrected, and the sampling rate was reduced from 100 Hz to 20 Hz to shorten the analysis time.

2.1. Modal Frequency Estimation

In structural health monitoring systems, estimating modal frequency is essential for understanding the dynamic behavior of a structure under different circumstances. Modal frequencies serve as a critical tool in comprehending the alterations in a structure's behavior under the impact of external stimuli, such as seismic activities. A real-time modal frequency estimation algorithm can observe the structural response and spot any deviations from the normal behavior, which can suggest damage or degradation. This facilitates early detection and prevention of further damage.

There are several approaches for estimating modal frequency in real-time, such as Fourier transform-based methods, autoregressive methods, and Kalman filter-based methods. Among these, Fourier transform-based methods are the most popular in structural health monitoring applications. The Fast Fourier Transform (FFT) is a well-established method for computing the Discrete Fourier Transform (DFT) of a signal rapidly.

Fourier transform-based techniques encompass the Continuous Wavelet Transform (CWT) [45] and the Short-Time Fourier Transform (STFT) [46]. CWT is suitable for dealing with non-stationary signals, while STFT enables adjusting the time and frequency resolutions of the signal representation. Since stationary stations were used in this study, the STFT method was preferred for Fourier transformations. The formulation of the STFT depends on whether the signal is discrete or continuous. For a continuous-time signal x(t), the STFT is defined as:

$$X(\tau, \omega) = \int_{-\infty}^{\infty} \mathrm{x(t)w(t-\tau)e^{j\omega t}} \tag{2.1}$$

where τ is the time index, ω is the angular frequency index, and w(t) is the window function. For a discrete-time signal x[n], the STFT is defined as:

$$X(\mathrm{n}, \omega) = \sum_{m=-\infty}^{\infty} \mathrm{x[n+m]w[m]e^{-j\omega m}} \tag{2.2}$$

where n is the time index, ω is the frequency index, and w[m] is the window function.

A band-pass filter was applied to the acceleration data to isolate the frequency components corresponding to the first and second modes of vibration of the structure. These modal frequencies were obtained from past studies that used different methods to analyze the dynamic response of the structure [23, 37, 42, 47, 48]. The frequency ranges of the first and second modes were defined as 1.6-1.95 Hz and 1.96-2.3 Hz, respectively, and were filtered accordingly.

2.2. Modal Damping Estimation

Modal damping is a critical parameter that affects the dynamic behavior of structures. It is essential to estimate modal damping accurately for the identification of structural damage. Several methods have been proposed for modal damping estimation, such as the Half-Power Band Width (HPBW) method, the Logarithmic Decrement method, the Peak Decay method, and the Curve Fitting method. This study employed the HPBW method to estimate the damping ratio of the structure, as this method is suitable for cases where the excitation force is unknown or ambient vibration [17, 49].

The HPBW method is simple and widely used in practice due to its ease of implementation and reliability. The main advantage of the half-power band width method is its simplicity, as it requires only basic equipment and can be performed in the field. However, the accuracy of the method may be affected by several factors, such as the signal-to-noise ratio of the measured data and the influence of higher modes. This method depends on the modal frequency and the bandwidth of that modal frequenccy and is computed as:

$$\xi = \frac{f_2 - f_1}{2f_{\text{res}}} \tag{2.3}$$

where f_{res} is the frequency corresponding to the peak amplitude on FAS. f_1 and f_2 are frequencies corresponding to a value of peak amplitude divided by the square root of two, as shown in Figure 2.1.

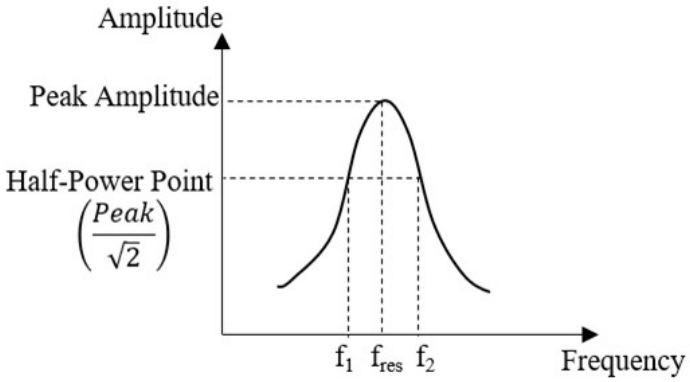

Figure 2.1. Selected frequencies for the half-power bandwidth method.

The damping ratio estimation using the Fourier amplitude spectrum (FAS) method depends on the choice of the smoothing coefficient, which controls the width of the window function applied to the spectrum. To determine the optimal smoothing coefficient for each site, a trial-and-error procedure was adopted, in which the smoothing coefficient was increased gradually until the damping ratio reached a stable value [17]. The optimal smoothing coefficient was then defined as the first value that fell within a linear trend line fitted to the damping ratio values. The same smoothing coefficient was used

for all data windows since they had the same duration.

2.3. Mode Shape Estimation

Mode shape estimation is an important aspect of structural health monitoring, as it provides valuable information on the structural behavior of a system. The mode shapes and their corresponding natural frequencies are intrinsic properties of a structure that depend on its geometry, material properties, boundary conditions, and damping. There are various methods for mode shape estimation, including experimental and numerical approaches. Experimental methods include techniques such as accelerometers, strain gauges, and displacement sensors, while numerical methods involve the use of finite element models.

One of the experimental methods for mode shape estimation is the Operational Modal Analysis (OMA) method. OMA is a method that employs output-only data, such as ambient vibration, to determine the modal properties of a structure. The OMA method can be used to estimate both the frequencies and mode shapes of a structure and is particularly useful for large and complex structures where it may not be feasible to apply external excitations. Among the various OMA techniques, the Frequency Domain Decomposition (FDD) method is a simple and robust approach that can effectively identify the mode shapes and natural frequencies of a structure [50].

Frequency Domain Decomposition (FDD) is a signal-processing technique used for modal analysis and system identification. It uses the power spectral density (PSD) of the measured signals to extract modal parameters such as modal frequencies, modal damping ratios, and mode shapes. The FDD method involves two steps: (1) decomposition of the measured response into a set of Cross-Power Spectral Density (CPSD) functions , and (2) estimation of the modal parameters from the identified frequency components. The decomposition step is typically performed using the Singular Value Decomposition (SVD) technique. The SVD technique decomposes the measured response matrix into a set of singular vectors and singular values, and the frequency

components are obtained from the singular vectors.

CPSD is the distribution of power per unit frequency and is defined as:

$$P_{xy}(\omega) = \sum_{m=-\infty}^{\infty} R_{xy}(m) e^{-j\omega m} \tag{2.4}$$

The cross-correlation sequence is defined as:

$$R_{xy}(m) = E\left\{X_{n+m} Y_n\right\} = E\left\{X_n Y_{n-m}\right\} \tag{2.5}$$

Where X_n and Y_n are jointly stationary random processes, $-\infty$ ¡ n ¡ ∞, and E is the expected value operator. The second step is the singular value decomposition of CPSD which can be defined as:

$$\text{SVD}\left(P_{xy}\right) = U_{xy} V_{xy} U_{xy}^{-T} \tag{2.6}$$

Where P_{xy} is a cross-power spectrum of each output, U is an orthogonal matrix containing singular vectors corresponding to the mode shape and V is a diagonal matrix containing singular values expressing the mode amplitudes (auto spectral value) of the corresponding modal frequency. The mode shape for a time t can be compared with the mode shape calculated for the time t+n, and in this way, the change in the long-term mode shape over time can be calculated.

2.4. Modal Assurance Criterion

Modal Assurance Criterion (MAC) is a method used to compare two sets of mode shapes, and it is commonly used in experimental modal analysis to verify the similarity between the experimentally obtained and analytically predicted mode shapes of a structure [51]. The MAC value ranges from 0 to 1, where 1 indicates a perfect match between the two sets of mode shapes. The MAC can be calculated using the following equation:

$$\text{MAC}(n) = \frac{\left|\left\{\boldsymbol{\phi}_r\right\}^T \left\{\boldsymbol{\Phi}_n\right\}\right|^2}{\left|\left\{\boldsymbol{\Phi}_r\right\}^T \left\{\boldsymbol{\phi}_r\right\}\right| \left|\left\{\boldsymbol{\Phi}_n\right\}^T \left\{\boldsymbol{\Phi}_n\right\}\right|} \tag{2.7}$$

where $\{\Phi_r\}$ is the reference mode shape and $\{\Phi_n\}$ is the mode shape at the n^{th} window.

One of the limitations of the MAC method is that it does not consider the modal scaling factor, which can result in a low MAC value even for mode shapes that are remarkably similar in shape but differ in magnitude. This limitation has led to the development of several modified versions of the MAC method, such as the Coordinate Modal Assurance Criterion (COMAC) and the Enhanced Coordinate Modal Assurance Criterion (ECOMAC), which incorporate the modal scaling factor into the comparison.

The Coordinate Modal Assurance Criterion (COMAC) is a modal assurance criterion that evaluates the similarity between two sets of mode shapes by using coordinate transformation to project the modes onto a common coordinate system [52]. COMAC compares the correlation between modal vectors without considering their spatial orientation and is preferred when a poor correlation is obtained with the MAC method. COMAC is defined as follows:

$$\mathrm{COMAC}(n) = \frac{\left\{\sum_1^N \left|\{\boldsymbol{\Phi}_r\}\,\{\boldsymbol{\Phi}_n\}\right|\right\}^2}{\left\{\sum_1^N \left|\{\boldsymbol{\Phi}_r\}\,\{\boldsymbol{\Phi}_r\}\right|\right\}^2 \left\{\sum_1^N \left|\{\boldsymbol{\Phi}_n\}\,\{\boldsymbol{\Phi}_n\}\right|\right\}^2} \tag{2.8}$$

where $\Phi_r is the r^{th}$ mode shape of the structure and $\Phi_n is the n^{th}$ analytical or experimental mode shape selected.

The ECOMAC method considers the contribution of all degrees of freedom in a structure and uses a weighting matrix to give more weight to the modes that have the highest energy content [53]. The weighting matrix is based on the diagonalization of the mass matrix and is used to normalize the mode shapes before computing the MAC values. The ECOMAC method provides a more accurate and robust measure of mode shape similarity than the MAC and COMAC methods. ECOMAC is calculated as:

$$\mathrm{ECOMAC}(n) = \frac{\left\{\sum_1^N \left\|\left\{\widetilde{\Phi_r}\right\}\left\{\widetilde{\Phi_n}\right\}\right\|\right\}^2}{2N} \tag{2.9}$$

$$\left\{\widetilde{\Phi_{j,t1}}\right\} = \{\Phi_{j,t1}\} \,/\, \|\{\Phi_{j,t1}\}\| \tag{2.10}$$

where $\widetilde{\phi}_r$ and $\widetilde{\phi}_n$ are vector normalized mode shapes. Both the Enhanced Coordinate Modal Assurance Criterion (ECOMAC) and Coordinate Modal Assurance Criterion

(COMAC) are modal assurance criterion (MAC) based methods used to quantify the similarity between two mode shapes. However, ECOMAC is an enhanced version of COMAC that considers the contributions of each coordinate to the mode shapes.

3. ENVIRONMENTAL EFFECTS ON MODAL PARAMETERS OF HAGIA SOPHIA

3.1. Modal Identification

The dynamic behavior of Hagia Sophia has been extensively studied from the past to the present, and its modal frequencies, modal damping ratios, and mode shapes are well known [23, 24]. The first modal frequency varies between 1.65 and 1.95 Hz, depending on environmental factors [17, 47]. In this mode, the structure performs in a purely linear motion in the E-W direction (Figure 3.1). The second modal frequency varies between 1.96 and 2.25 Hz and is a mode in the N-S direction. In this mode, while the structure moves in the N-S direction, the movement is contributed by oscillations in the E-W and vertical directions (Figure 3.2).

It should be noted that the N-S and E-W directions, as used in this thesis, do not correspond to the real N-S and E-W directions. It is just a preference that we and other researchers have made in the past, as it provides a significant simplicity in discussing the findings and relating them to different structural elements of Hagia Sophia. The discussion in this thesis is kept to the first two modes, to enable comprehensive analysis and discussion of all modal parameters in a relatively simpler framework.

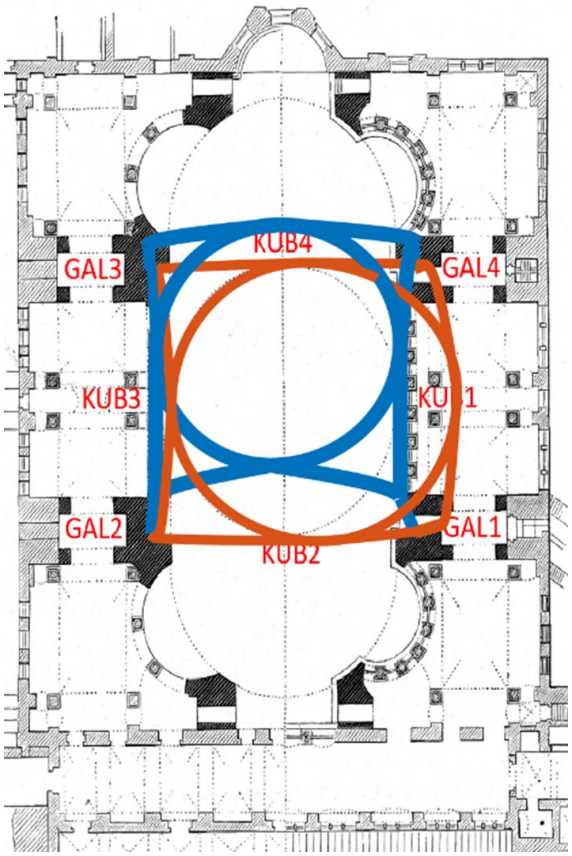

Figure 3.1. The first and second mode shapes of Hagia Sophia, top view.

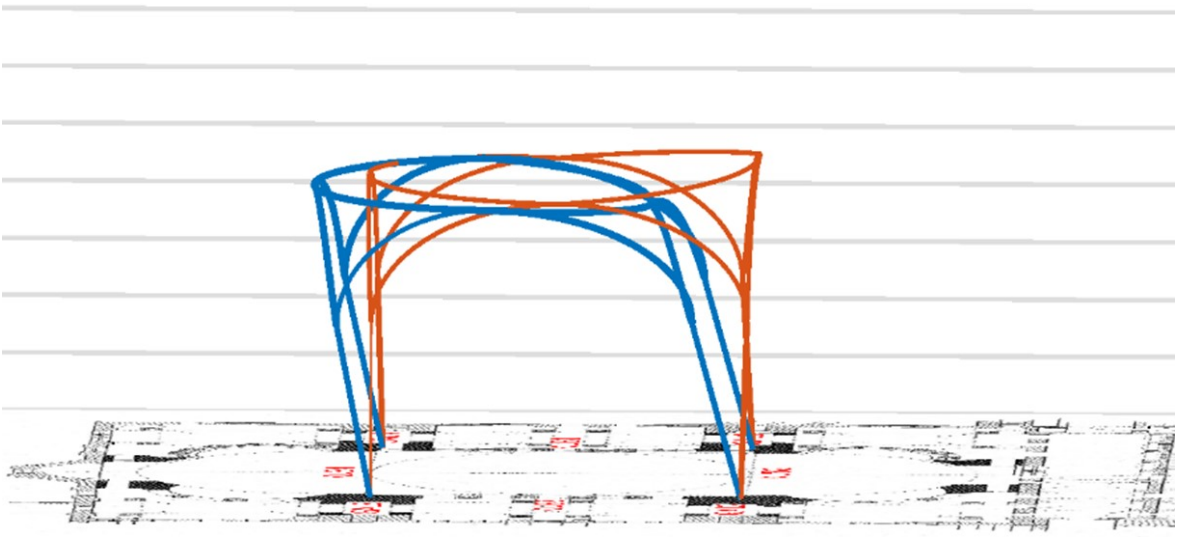

Figure 3.2. The first and second mode shapes of Hagia Sophia in vertical view.

3.2. Long-term Variation of Modal Frequency

Structures are in constant interaction with their environment, traces of which can be detected from their dynamic reactions. Ground motions and temperature are the most influential factors on dynamic parameters. Humidity, precipitation, wind speed, and air pressure are some factors that can be secondary [47]. In this section, the effects of temperature, humidity, wind speed, and precipitation on the modal frequency, modal damping, and mode shapes of Hagia Sophia between 2013 and 2017 are examined. First, the time variation of modal and atmospheric parameters was discussed. In the second part, the dependence of modal parameters on environmental factors, neglecting time as an independent variable is investigated.

Temperature change is the main atmospheric parameter affecting the modal parameters of Hagia Sophia. On the historical peninsula of Istanbul, where Hagia Sophia is located, the temperature varied from winter to summer, during four-year observation between the lowest value of -4 ^{0}C and the highest value of +38 ^{0}C (Figure 3.3). In return, the first modal frequency increased by approximately 18% from 1.65 Hz in the winter months to 1.95 Hz in the summer. Similarly, the second modal frequency increased by approximately 15% from 1.96 Hz to 2.25 Hz, as shown in Figure 3.4. The reason for this frequency increase is that the temperature increase expands the material and forces existing cracks to close [1, 6]. Therefore, it can be said that the structure becomes more rigid in the summer months. Figure 3.3 shows that the range of daily temperature variation between day and night is about 10 degrees, a range that is wider in the months from January to August than in the months leading to winter. It is interesting to note that modal frequencies exhibit a diurnal sensitivity to temperature, increasing throughout the day and decreasing overnight (Figure 3.4).

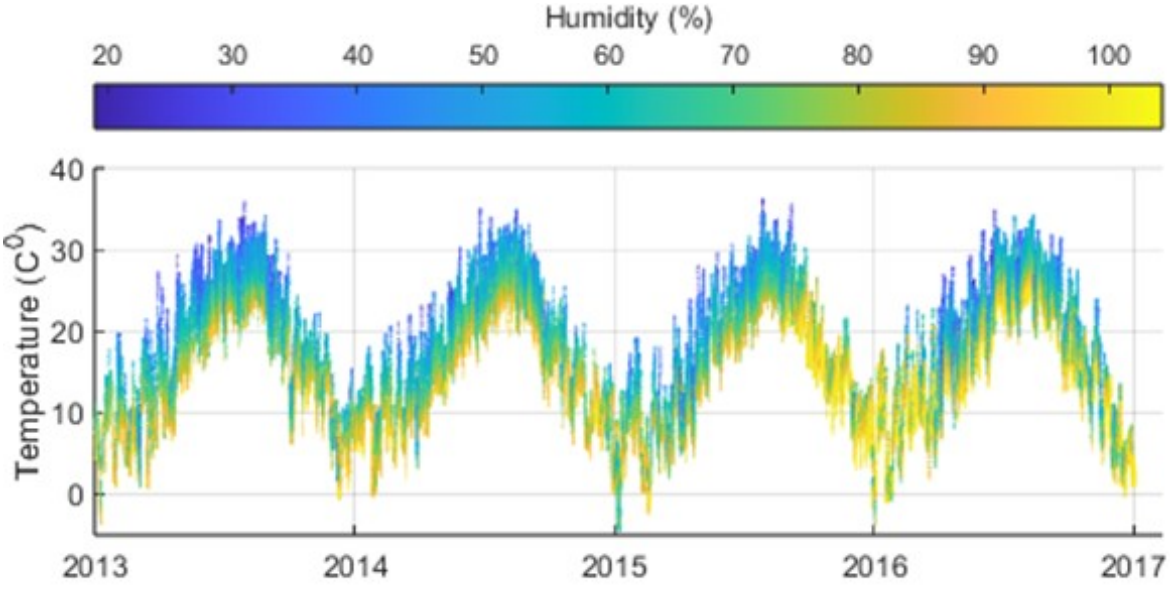

Figure 3.3. Long-term variation of air temperature and humidity between 2013 and 2017.

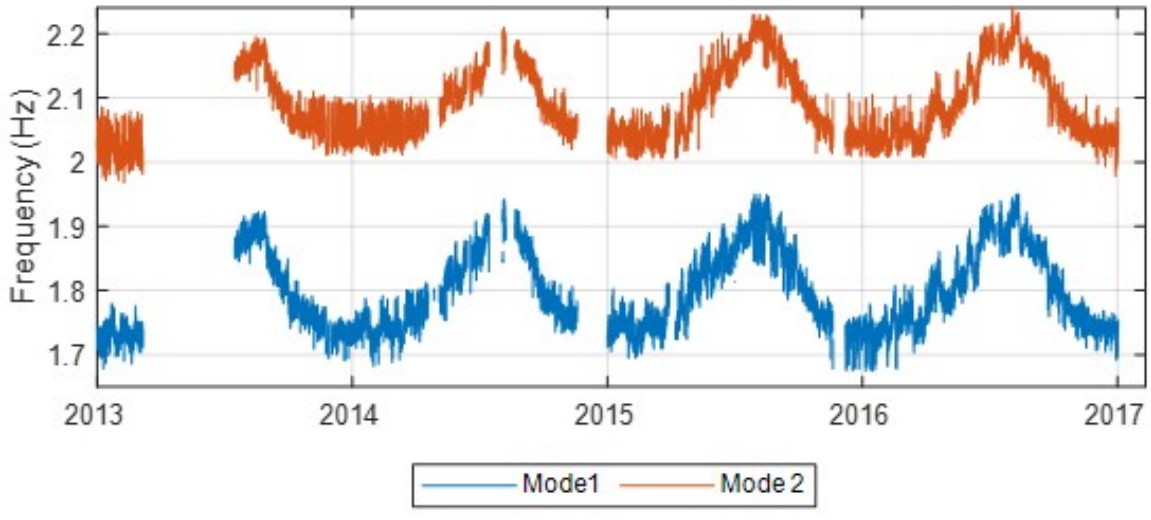

Figure 3.4. Long-term variation of first and second modal frequencies between 2013 and 2017.

It should also be noted that the SHM system registered 135 earthquakes over a four years observation period, the magnitudes of which varied between M_w 4 and 6.5. They are inherently included in our data set and therefore part of our analysis. It has already been demonstrated that increased ground accelerations force modal frequencies to drop, which recover soon after the earthquake to their pre-event values if no permanent change in the system [16,18,23]. These temporary declines go unnoticed in our analysis (Figure 3.4) showing that the decline caused by these earthquakes was well within the range of natural diurnal and seasonal frequency variations and did not cause any permanent impact on the structure.

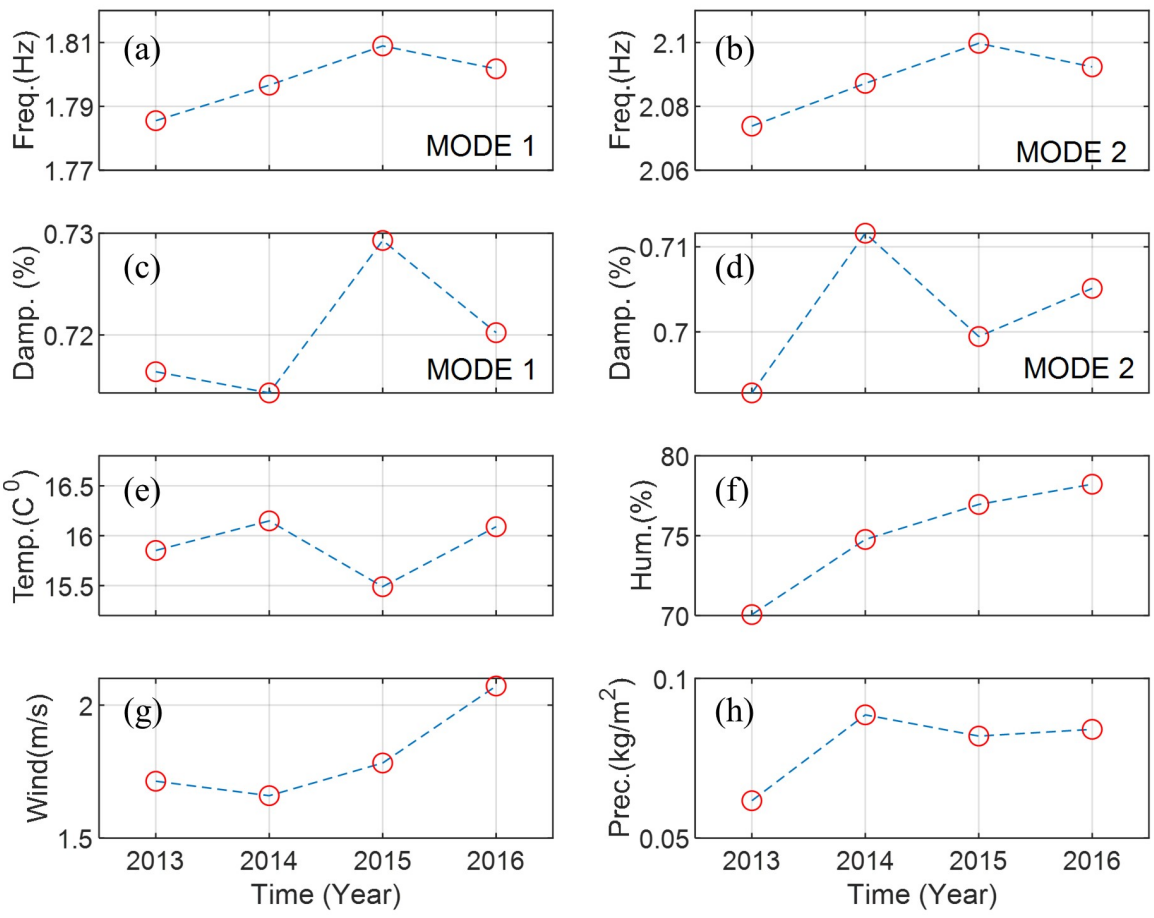

Figure 3.5. The average annual values of modal frequencies, modal damping, and atmospheric parameters.

Our results suggest that global warming, which has increased its influence in recent years, might affect Hagia Sophia as well. In these four years, although there is no significant increase in the average annual temperature (Figure 3.5 (e)); humidity and wind speed are on an increasing trend in Istanbul (Figure 3.5 (f), Figure 3.5 (g), Figure 3.5 (h). In return, the average annual modal frequencies of Hagia Sophia increased by about 0.5% each year, as seen in Figure 3.5 (a) and Figure 3.5(b). These might be early indicators of Hagia Sophia becoming more rigid with global warming. This observation might sound premature and overstretched. It needs to be verified by looking at data over longer periods. Therefore, it is an important point to make as if this is a trend and continues, then it may lead to additional deteriorations in the materials of the structure in the future [27].

This section examines how modal frequencies and damping ratios vary depending on atmospheric factors, without considering the effect of time as an independent variable. The modal frequencies of Hagia Sophia are largely controlled by the temperature. The relationship between temperature and modal frequencies can be expressed by the third-degree polynomial curve, as shown in Figure 3.6. This curve was obtained by finding the best fit using the sum of the squared estimate of errors (SSE) and R^2 scores. R^2 scores calculated for the first and second modal frequencies are 0.83 and 0.78, respectively. These results show that the first two modal frequencies are highly correlated with temperature. However, the second modal frequency is slightly more related to non-temperature factors.

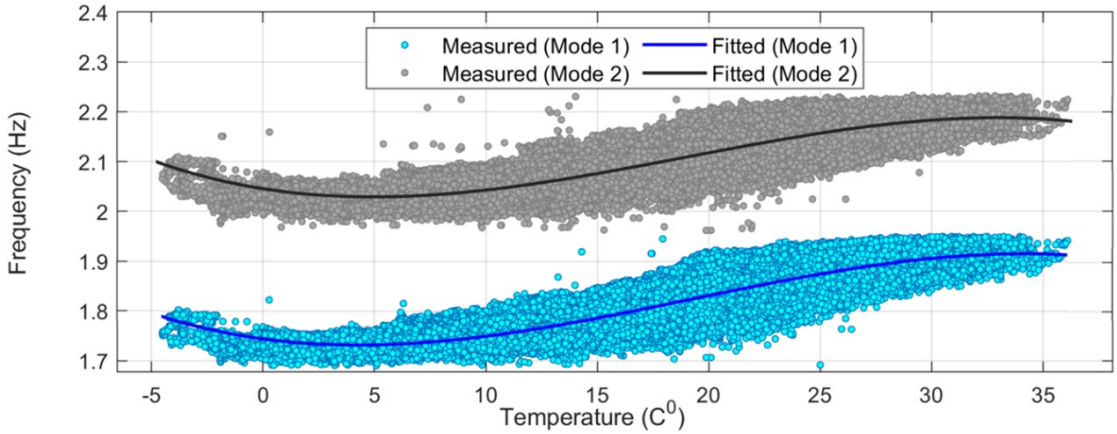

Figure 3.6. Variation of the first and second modal frequencies with temperature.

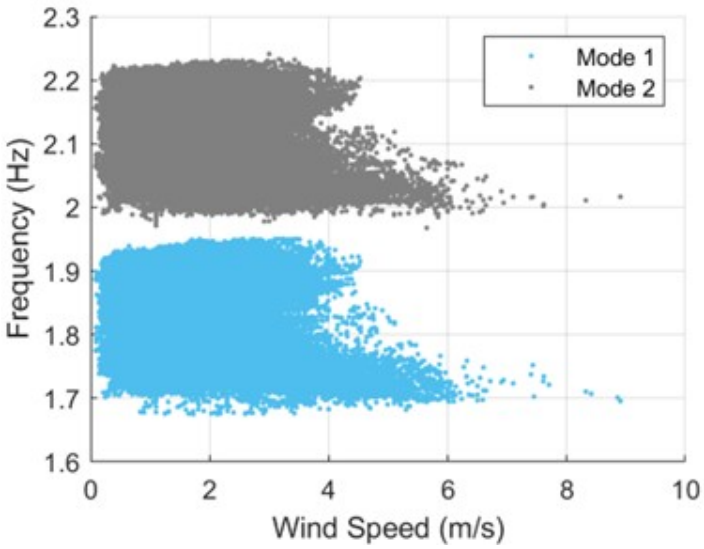

Figure 3.7. Variation of the first and second modal frequencies with wind speed.

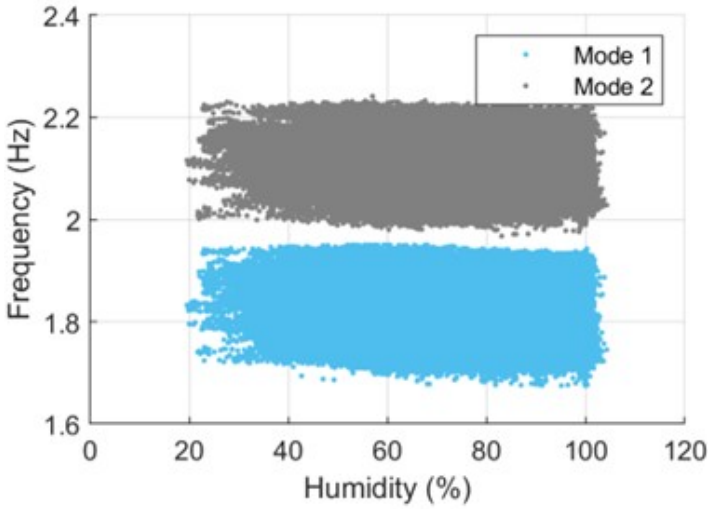

Figure 3.8. Variation of the first and second modal frequencies with humidity.

The modal frequencies exhibit distinct patterns across two temperature intervals. The first temperature range is between about -5 ^{0}C and +4 ^{0}C, while the second range is between +4 ^{0}C and +37 ^{0}C. The modal frequencies of Hagia Sophia increase as the temperature drops near 0 ^{0}C and below as presented in Figure 3.6. This increase begins at about +4 ^{0}C, where the density of the water is at its lowest, and continues up to -4 ^{0}C, which is the lowest temperature measured. In this temperature range, the first and second modal frequencies increase by approximately 4% and 3%, respectively. This is because of the icing and hardening of the water in the capillary cracks and mortar [27] below +4 ^{0}C and, as a result, the structure becomes more rigid. However, it is interesting to note that the change in both modal frequencies has an increasing trend until about 30 ^{0}C and slows down afterward.

Since wind measurements are not taken on Hagia Sophia, it is relatively difficult to discuss the correlation between wind speed and modal frequencies (Figure 3.7). However, it can be said that the first two modal frequencies tend to drop when wind speeds exceed 4 m/s. There is just a very slight trend between humidity and modal frequency (Figure 3.8). Therefore, it can be said that the causality between humidity and modal frequency is quite limited.

3.3. Effect of Modal Frequency Variation on Response Spectrum

The significant annual variation of modal frequencies in the order of 18% and 15% for the first two modes respectively, might have implications for the earthquake loading Hagia Sophia is exposed to. To gain a better understanding, the ten largest earthquakes recorded between 2012 and 2021 are selected according to the maximum absolute accelerations at ground station ZEM1. The ground station of a structural health monitoring system is generally not suitable for calculating the response spectrum due to Soil Structure Interaction (SSI), but the SSI is very low in Hagia Sophia because of the hard ground [42, 47]. Therefore, response spectra were calculated using ground station acceleration data. The properties of these earthquakes are shown in Table 3.1.

The response spectra are calculated as the geometric mean of the two horizontal components after normalizing them concerning maximum absolute acceleration. The rationale behind normalizing the records is to show the proportional change in spectral acceleration levels due to variations in modal frequencies. Response spectra are estimated for 1% and 5% damping and are shown in Figure 3.9 and Figure 3.10. The period of the first mode of the structure reaches the highest T_1=0.606 s (F_1=1.65 Hz) in winter and the lowest T_2=0.512 s (F_1=1.95 Hz) in summer depending on environmental effects. In this range, spectral accelerations of the structure increase up to 91%, and the mean spectral acceleration increases by 51% from winter to summer Table 3.1. These results show that the most unfavorable conditions for the structure in terms of spectral acceleration occur during the summer seasons due to the modal period reaching its lowest levels. In other words, spectral acceleration reaches its highest levels during extreme heat, as the modal period reaches its lowest levels.

Table 3.1. The parameters of selected earthquakes and change of spectral accelerations from winter to summer according to a 1% damping ratio.

ID	Date	Distance (km)	M_W	PGA (cm/s^2)	$Sa(T_1)$	$Sa(T_2)$	Change of Sa (%)
EQ 1	26.09.2019	66	5.8	21.6	0.92	1.32	43.8
EQ 2	24.05.2014	305	6.9	5.5	1.79	3.3	85
EQ 3	11.01.2020	66	4.7	2.3	0.31	0.48	53.6
EQ 4	24.09.2019	66	4.2	1.2	0.53	0.8	50.9
EQ 5	8.01.2013	332	5.7	1.2	5.39	5.24	-2.7
EQ 6	12.06.2017	333	6	1	3.7	5.34	44.5
EQ 7	5.02.2014	50	3.6	0.9	0.15	0.27	80.6
EQ 8	16.11.2015	27	4.4	0.9	0.67	1.15	72.9
EQ 9	14.03.2012	26	3.7	0.5	0.21	0.35	68.7
EQ 10	7.06.2012	90	5.2	0.5	1.2	2.19	82.5
Mean	-	-	-	-	1.49	2.04	37.7

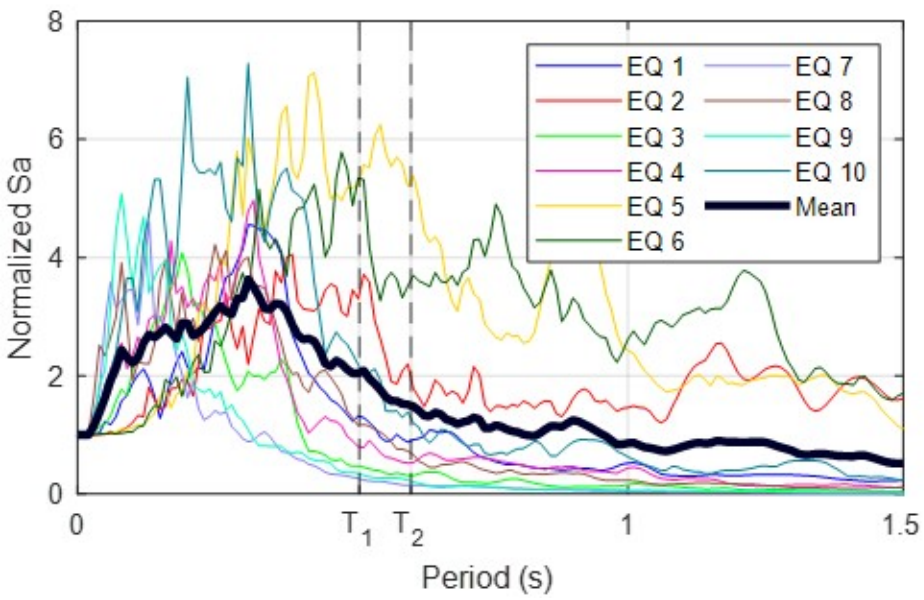

Figure 3.9. The normalized response spectra of selected earthquakes for a 1% damping ratio.

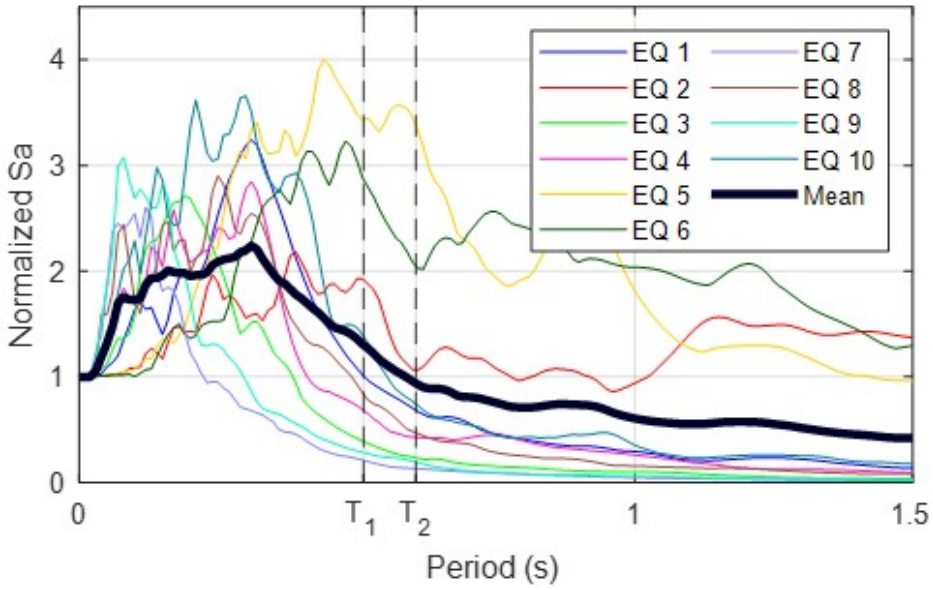

Figure 3.10. The normalized response spectra of selected earthquakes for a 5% damping ratio.

3.4. Long-term Variation of Modal Damping

The modal damping ratio is an important dynamic parameter that expresses the energy dissipation in the dynamic behavior of a structure. Permanent increases in the modal damping ratio are often associated with structural damage [54]. Therefore, studying the interaction of the modal damping ratio with environmental factors is particularly important for detecting structural anomalies.

Long-term variation of temperature and modal damping ratios between 2013 and 2017 is shown in Figure 3.11. According to the result presented in this figure, from winter to summer, the modal damping ratio gradually drops from 0.8% to 0.6%, thus decreasing by approximately 25%. No meaningful difference between the estimated damping ratios in the two orthogonal directions can be noticed. The time-dependent variation of modal damping is less pronounced than that of modal frequency. This is partly related to the fact that the estimation of modal damping includes a larger uncertainty. Yet, its distinct increase in December, January, and February and gradual

drop towards the summer months are clearly demonstrated. As already shown, within the four-year observation period, annual average humidity ratios and wind speeds are on an increasing trend in Istanbul (Figure 3.5 (f) and Figure 3.5 (g)). There is some increase in modal damping ratios in the same period, but it is not a meaningful trend (Figure 3.5 (c) and Figure 3.5 (d)).

The modal damping ratio is affected by many atmospheric parameters, unlike the modal frequency, which has a strong correlation with temperature. Figure 3.12 shows that the first and second modal damping ratios increase as the temperature decreases, but this is considered to be an indirect effect. This is because the winter months are characterized by higher rainfall and humidity, which increase the mass of the structure, and higher wind speeds, which increase the displacements. These weather conditions widen the bandwidth of the modal frequencies and the modal damping ratio. In summer, the modal frequency of the structure becomes more stable, and the bandwidth narrows with decreasing rainfall and wind speed ($\Delta f_{winter} > \Delta f_{summer}$) ; therefore, the model damping drops as illustrated indicatively in 3.15.However, as the temperature increases in summer, the Fourier amplitude of the modal first modal frequency increases ($A_{winter} < A_{summer}$). The increase in wind speed around Hagia Sophia increases the modal damping ratio by up to 45% (Figure 3.13). One of the reasons for the inverse relationship between modal frequency and modal damping can be explained as the increased humidity and precipitation increase the mass of the structure, causing frequency to drop and modal damping to increase [19]. The modal damping ratio is influenced by the humidity ratio, but this relationship is not consistent across different humidity levels. At high humidity levels, the modal damping ratio shows no correlation with the humidity ratio, whereas at low humidity levels, the modal damping ratio decreases as the humidity ratio decreases (Figure 3.14). Freezing conditions do not have a significant effect on the modal damping ratio.

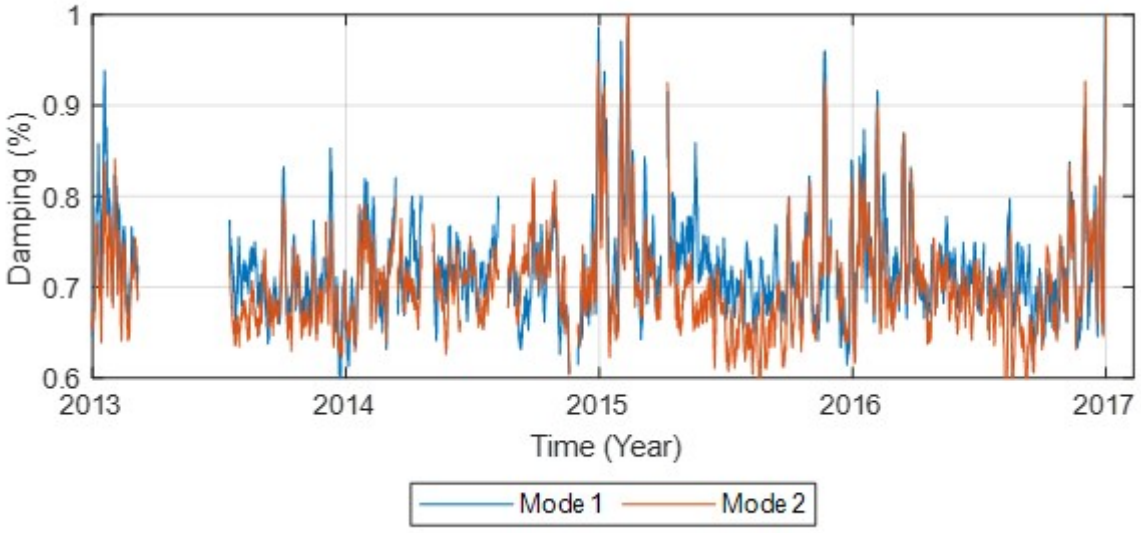

Figure 3.11. The long-term variation of first and second modal damping ratios between 2013 and 2017.

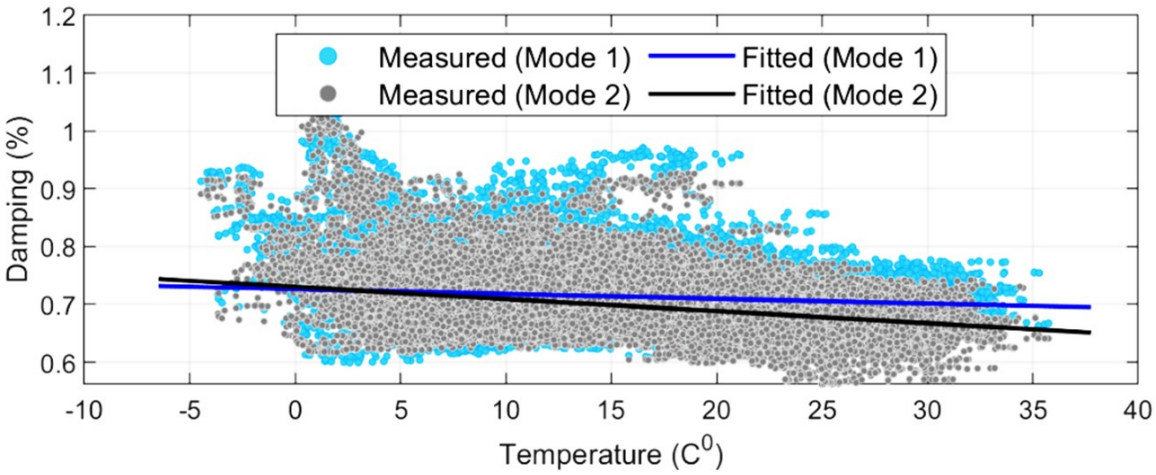

Figure 3.12. The variation of the first and second modal damping ratios with temperature.

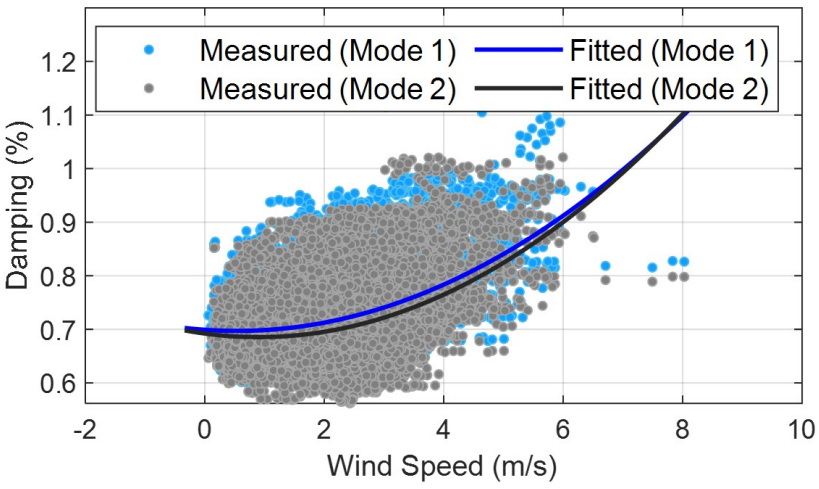

Figure 3.13. The variation of the first and second modal damping ratios with wind speed.

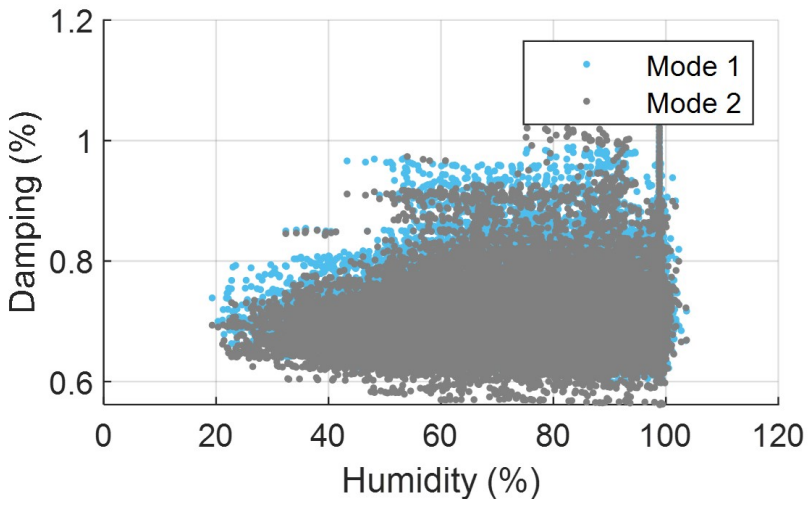

Figure 3.14. The variation of the first and second modal damping ratios with humidity.

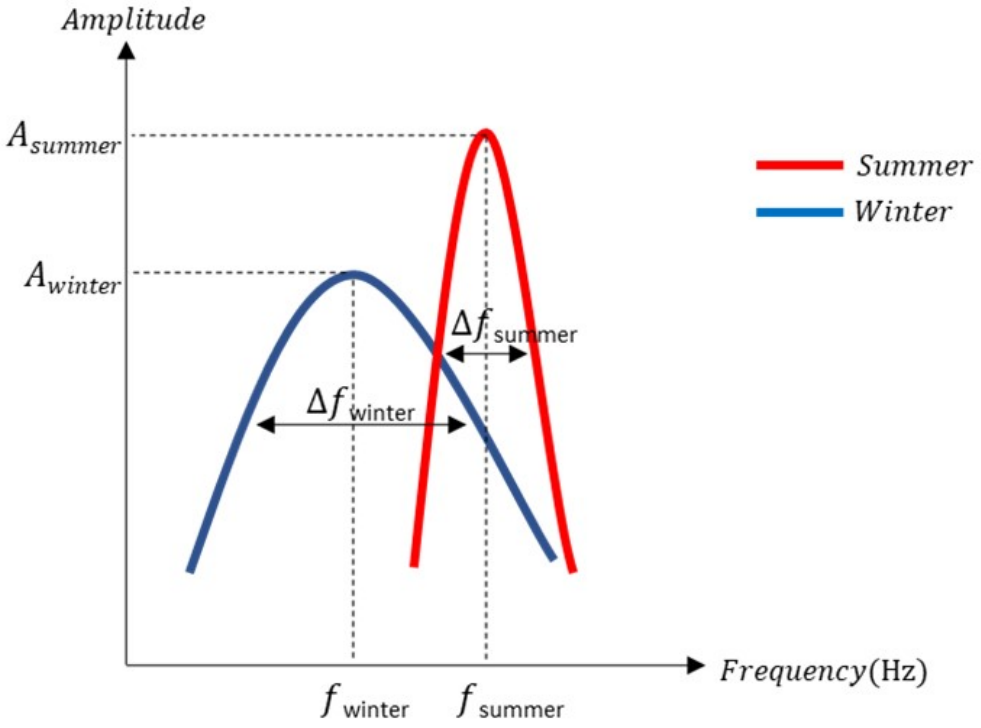

Figure 3.15. The general representation of Fourier amplitude spectra of the first mode from winter to summer.

The R^2 values obtained from the linear regression of the first and second modal damping ratios are 0.34 and 0.36, respectively (Table 3.2). Under changing atmospheric conditions, the first modal damping ratio is more stable than the second. The relationship between wind speed and modal damping ratio is best modeled by the exponential curve, as shown in Figure 3.13. The R^2 scores obtained for the second-order polynomial curves fitted to first and second-modal damping ratios are 0.42 and 0.35, respectively (Table 3.2). The relative humidity is usually a temperature-dependent parameter and varies considerably both during the day and seasonally. There is no obvious trend between relative humidity and modal damping, however, it is observed that the first and second modal damping ratios vary between 0.6 and 0.8% when the humidity is between 20% and 40% (Figure 3.14).

Correlation coefficients between modal damping, modal frequency, and atmo-

spheric parameters in both directions are shown in Figure 3.16 and Figure 3.17. The correlation coefficient "1" indicates a direct correlation and "-1" shows an inverse correlation. The highest correlation with the damping ratio is observed for precipitation, followed by wind speed, partial humidity, and atmospheric pressure. The strongest inverse correlation is observed with soil temperature, followed by air temperature and modal frequency. The correlation coefficients obtained for the first and second modal frequencies are quite close to each other.

Table 3.2. Statistical results obtained from fitted curves between atmospheric parameters and modal parameters.

	Freq-Temp		**Damping-Wind**		**Damping-Temp**	
	Mode 1	**Mode 2**	**Mode 1**	**Mode 2**	**Mode 1**	**Mode 2**
R2	0.83	0.78	0.42	0.35	0.34	0.36
RMSE	0.025	0.026	0.041	0.044	0.044	0.044

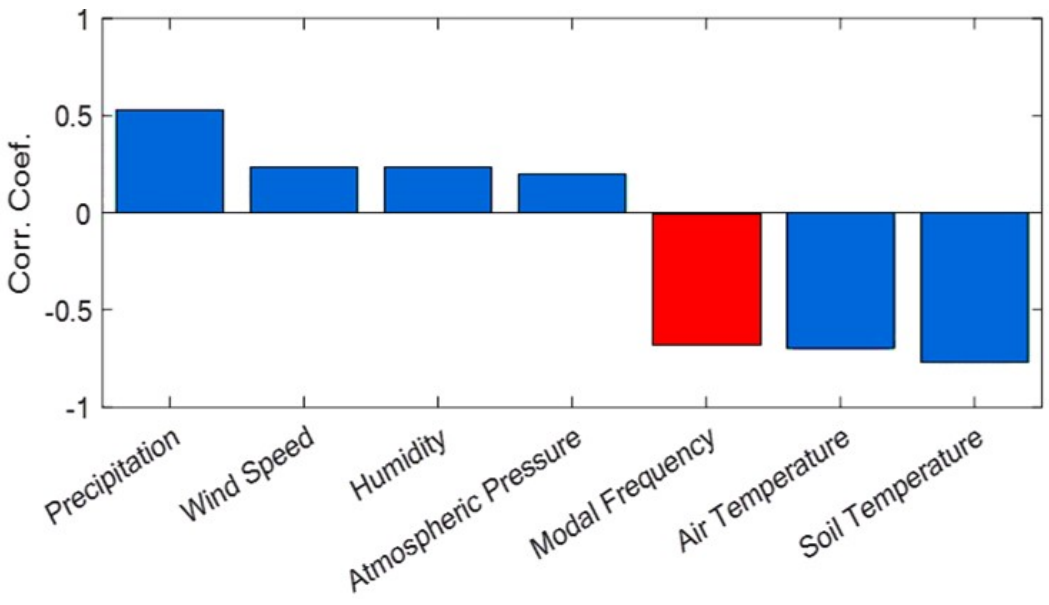

Figure 3.16. Correlation coefficients between modal damping and atmospheric factors for the first mode.

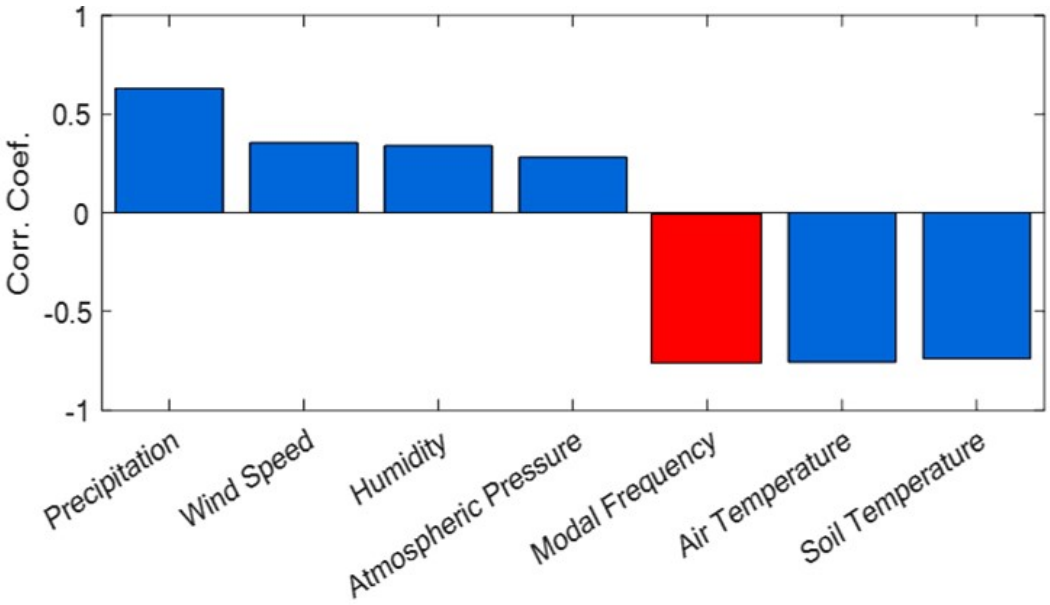

Figure 3.17. Correlation coefficients between modal damping and atmospheric factors for the second mode.

3.5. Long-term Mode Shape Variation

This section presents the analysis of the mode shape variations of Hagia Sophia during the observation period. The mode shapes were estimated for seven stations, excluding GAL4, which had a time synchronization problem, using the FDD method for each time window. Then, the MAC, COMAC, and ECOMAC methods were applied to evaluate the long-term changes in the mode shapes. The reference mode shape was assumed to be the one estimated in the time window between 00:00 and 01:00 am on January 1, 2013. The mode shapes estimated in the subsequent windows were compared with the reference mode shape to track the long-term fluctuations. Figure 3.18 and Figure 3.19 show the mode shape changes calculated using the three methods for the first and second modes, respectively. The results indicate that the first-mode shape of Hagia Sophia remained stable between 2013 and 2016, as evidenced by the high values of MAC and COMAC (above 0.995). The ECOMAC values also ranged between 0.99-1, suggesting a minor influence of seasonal effects on the mode shape.

In the second mode shape, a sudden, notable change was clear in late 2014 (Figure 3.19). This was due to a calibration error in the N-S component of station GAL3, which was corrected in late 2014. Therefore, the mode shapes estimated between 2013 and 2014 were inaccurate, and the correction resulted in a decrease of MAC and COMAC by 9% and ECOMAC by 5% compared to the erroneous mode shapes. The ECOMAC method was less sensitive to the calibration error because it is relatively robust to scaling and calibration errors.

The analysis results confirmed that the first mode shape was not influenced by environmental factors. Therefore, any permanent changes in the MAC, COMAC, and ECOMAC values of the first mode shape could indicate structural damage, unless there was a calibration error. However, the second mode shape showed some transient changes in the MAC values in some periods. This was attributed to a calibration failure and repair work on the N-S direction of the GAL 3 station.

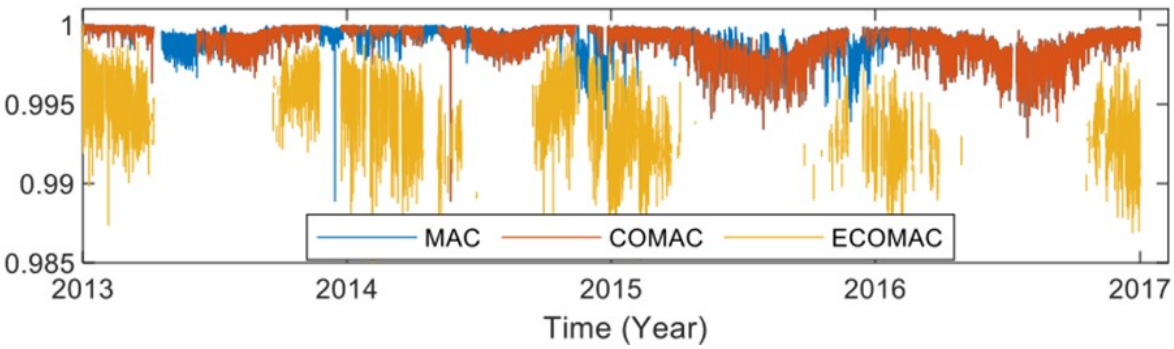

Figure 3.18. Variation of the first mode shape (MAC, COMAC and ECOMAC) between 2013-2017.

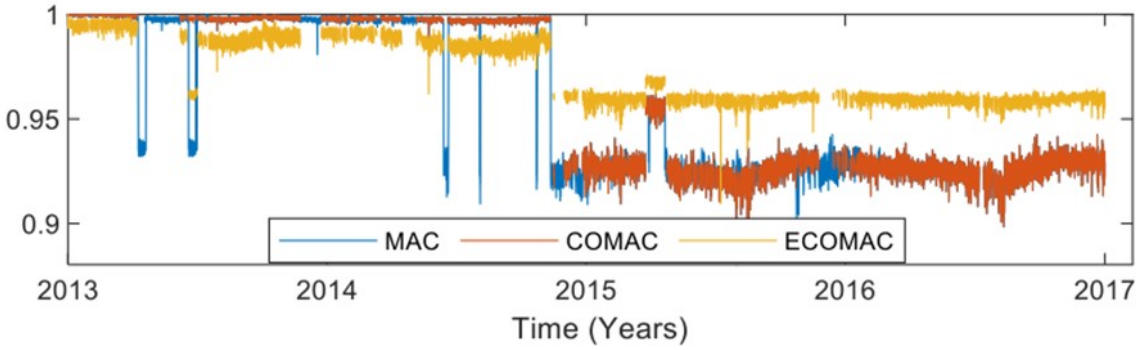

Figure 3.19. Variation of the second mode shape (MAC, COMAC and ECOMAC) between 2013-2017.

4. REGRESSION MODEL OF MODAL FREQUENCY

Structural health monitoring is a field that has benefited from the recent advances in machine learning science over the past decade. One of the main applications of machine learning in SHM is to develop predictive statistical models that can capture the correlation between modal frequency and environmental factors. However, there is no consensus on the best statistical learning method for SHM systems, and different methods may have different advantages and limitations. In this chapter, four widely used statistical regression analysis methods are applied to model the relationship between atmospheric data and the modal frequency of a historical structure. The performance and accuracy of each method are compared and discussed.

The modal frequencies of Hagia Sophia were observed to be in a cycle between 2013 and 2016. The results showed that the modal frequencies did not exhibit any significant changes that could indicate structural damage during this period. Therefore, it was assumed that the modal frequencies would remain stable unless the structure suffered any damage in the future. The main objective of the statistical modeling between modal frequency and atmospheric data was to develop a reliable and accurate method for detecting structural anomalies that could be related to damage by comparing the predicted and measured modal frequencies.

In this thesis, prediction models were created using the MATLAB Regression Learner Tool [55]. Four different regression analysis methods were employed, and their results were compared. The following is a list of these methods and their subgroups:

- Multiple Linear Regression (MLR)
  - Linear
  - Interactions Linear
  - Robust Linear
  - Stepwise Linear

- Regression Trees (RT)
  - Fine Tree
  - Medium Tree
  - Coarse Tree
- Support Vector Machines (SVM)
  - Linear
  - Quadratic
  - Cubic
- Artificial Neural Network Regression (ANN)
  - Narrow
  - Medium
  - Wide
  - Bilayered
  - Trilayered

In the developed statistical models, the input variables for the regression analysis were air temperature, relative humidity, wind speed, and precipitation, while the output variable was modal frequency. The data set consisted of 15-minute mean values of modal frequency and atmospheric data collected from 2013 to 2016. The initial three-quarters of the data (the years between 2013 and 2015) set served as the basis for training the regression models, while the final quarter (2016) was utilized to evaluate their performance.

4.1. Multiple Linear Regression

Multiple linear regression (MLR) is a widely used regression analysis method in the literature due to its simplicity, applicability, and high accuracy of estimates in linear systems. However, it is important to note that MLR has some disadvantages. For instance, it postulates that a linear relationship exists between the dependent variable and each of the independent variables. If the relationship is not linear, the model may not provide accurate results. Additionally, MLR can be sensitive to outliers and

multicollinearity, which can lead to inaccurate results.

The MLR method was examined in four subgroups: basic linear, interactions linear, robust linear, and stepwise linear. The formulation of the basic linear regression method is as stated in Eq. (4.1):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_n X_{ni} + \varepsilon_i \tag{4.1}$$

where Y_i is the i^{th} output, X_{ji} are the i^{th} value of j^{th}independent variable, n is the number of independent variables, β are regression coefficients, and ε_i is the i^{th} error [56]. Rearranging this formulation in matrix form for a modal frequency estimate based on atmospheric parameters, Eq. The formulation of (4.2) is obtained.

$$\begin{bmatrix} F_1 \\ \vdots \\ F_n \end{bmatrix} = \begin{bmatrix} 1 T_1 H_1 W_1 S_1 P_1 \\ \vdots \; \vdots \; \vdots \; \vdots \; \vdots \; \vdots \\ 1 T_n W_n H_n S_n P_n \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{4.2}$$

In this formulation, T is temperature, H is relative humidity, W is wind speed, S is soil temperature, and P is precipitation. In addition, β represents regression coefficients and ε means errors.

The second linear regression model is an interactions linear regression, which is a statistical technique that allows you to explore how the relationship between a dependent variable and one or more independent variables changes depending on the values of another independent variable. To include an interaction term in a linear regression model, you need to multiply the independent variables that are involved in the interaction and add the product term as a new independent variable in the model.

The coefficient of interaction term shows how the slope of the regression line changes across various levels of the interacting variable. The general form of interactions linear regression method with two independent variables can be expressed as follows:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{1i} X_{2i} + \varepsilon_i \tag{4.3}$$

where Y_i is the dependent output variable, X_{1i} and X_{2i} are independent input variables, β_0, β_2, β_2 are regression coefficients, and e is the error term, then β_3 is the coefficient of the interaction term $X_{1i}X_{2i}$.

The third linear regression method is robust linear regression, which is a powerful technique for modeling relationships between variables when dealing with noisy or outlier-contaminated data [57]. Outliers, which are data points that significantly differ from the expected pattern, can skew the outcomes of a regression analysis. Robust linear regression methods can minimize this effect and estimate the coefficients and error variance more reliably. It is particularly useful in fields where outliers are common, such as finance, environmental science, and engineering [58]. Some key points of robust linear regression are listed below:

- Traditional MLR regression can be significantly affected by outliers, leading to biased parameter estimates. Robust regression methods are designed to mitigate this issue by weight-reducing the impact of outliers [59].

- Robust regression uses an Iteratively reweighted Least Squares approach, which assigns weights to each data point based on their residuals. This technique demonstrates robustness against outliers due to its lower sensitivity to significant variations in minor data segments.

- The choice of weight function influences the robustness of the regression. Common weight functions include the bisquare, Huber, and Welsch functions. These functions determine how much influence each data point has on the parameter estimates [60].

The fourth and last linear regression method is stepwise linear regression, which is a statistical method that involves fitting regression models using an automatic procedure. In each step, a variable is evaluated for inclusion or exclusion from the set of explanatory variables based on a pre-specified criterion [61]. Generally, it takes the form of a series of forward, reverse, or combined F-tests or T-tests [62].

The four linear regression methods described above are used to train atmospheric parameters in order to predict the model frequencies of the structure. The Root Mean Square Error (RMSE), the coefficient of determination (R^2), training time, and prediction speed values are listed in Table 4.1 and Table 4.2. Among these methods, interaction linear regression and stepwise linear regression methods give better predictions than the other methods. Stepwise linear regression is the slowest method for training and predicting among linear regression methods.

Table 4.1. The regression test results of the first modal frequency obtained via linear regression methods.

Model Type	***RMSE***	***R^2***	***Training Time (s)***	***Prediction Speed (Obs/s)***
Linear	0.025	0.834	7	3603799
Interactions Linear	0.023	0.859	5	2408904
Robust Linear	0.025	0.839	1	3420166
Stepwise Linear	0.023	0.859	41	425373

Considering both the RMSE and the prediction speed, the most applicable linear regression method for our data is the interactions linear method. The predicted and true frequencies obtained via the interactions linear method are presented in the time domain in Figure 4.1. It is observed that the difference between the predicted and true frequencies widens at the upper and lower levels of the modal frequencies. Figure 4.2 shows the graph of the predicted frequency versus true frequency and Figure 4.3 shows the residuals in the predictions on the frequency base. These two graphs provide

Table 4.2. The regression test results of the second modal frequency obtained via linear regression methods.

Model Type	***RMSE***	R^2	***Training Time (s)***	***Prediction Speed (Obs/s)***
Linear	0.025	0.805	9	939011
Interactions Linear	0.024	0.830	6	821263
Robust Linear	0.025	0.806	5	3405903
Stepwise Linear	0.024	0.830	129	418928

evidence that the error rate increases as the frequency level deviates from the mean frequency.

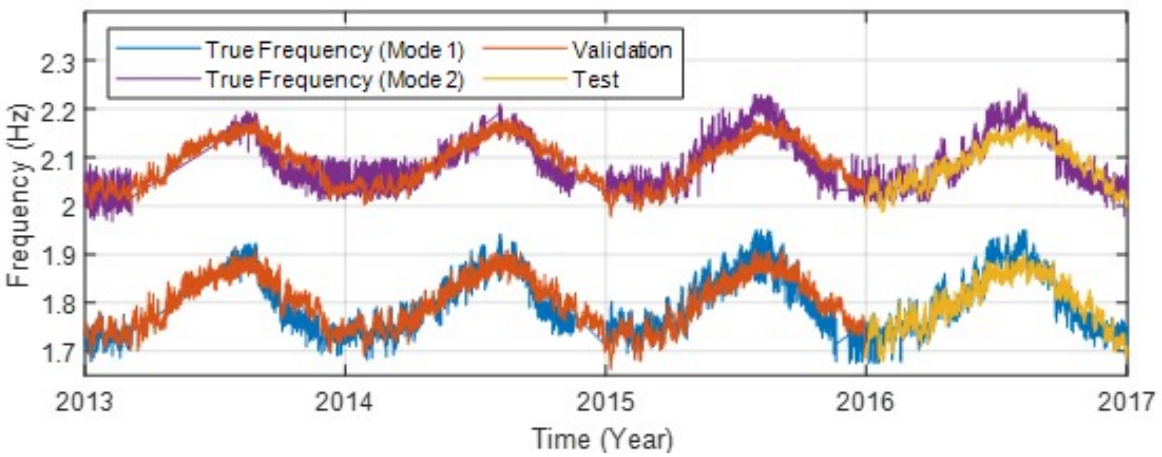

Figure 4.1. Regression results of modal frequencies using interactions MLR method.

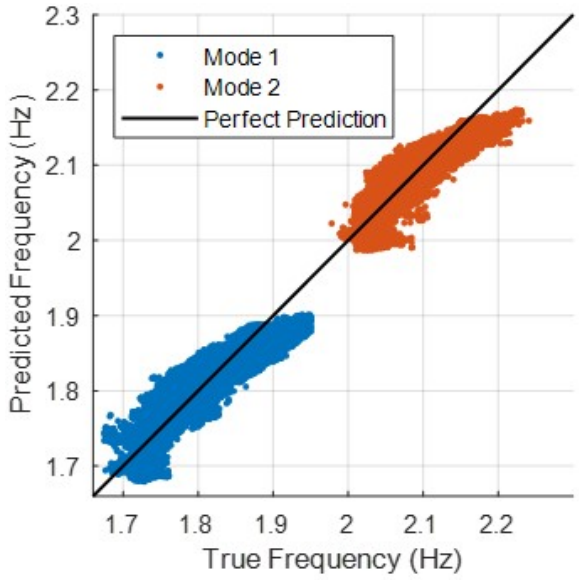

Figure 4.2. Predicted and true frequency comparison for interactions MLR method.

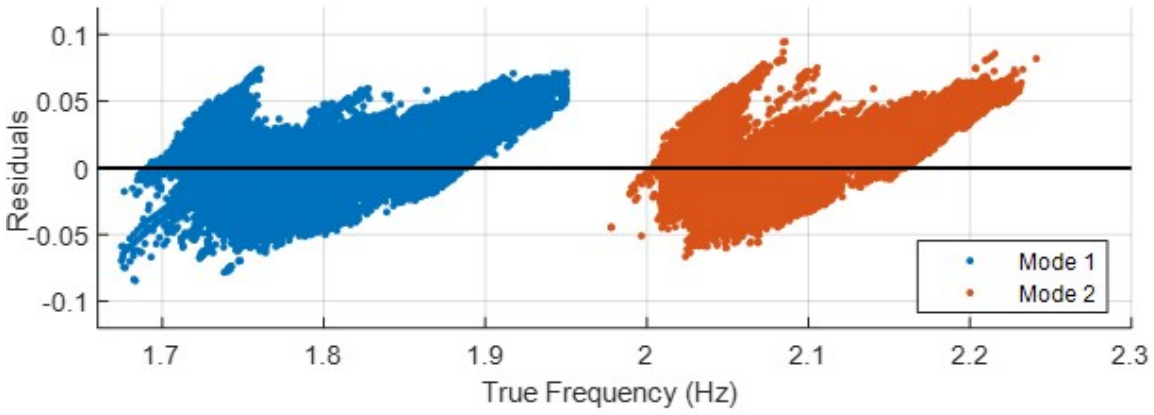

Figure 4.3. Residuals of predictions for interactions MLR method.

4.2. Regression Trees

Regression trees are a type of supervised learning method that can be used to address multiple regression problems. They provide a tree-based approximation of an unknown regression function $Y = f(x) + \varepsilon$ with $Y \in R$ and $\varepsilon \approx N(0, \sigma^2)$. The derived

models are structured as a series of logical evaluations performed on the values of any given predictor variable. In the context of decision trees, terminal nodes or leaves are used to predict the target variable Y [63].

Regression trees have several features that make them an interesting approach to several multiple regression problems. Namely, regression trees provide automatic variable selection making them highly insensitive to irrelevant variables, computational efficiency that allows addressing large problems, handling of unknown variable values, handling of both numerical and nominal predictor variables, insensitivity to predictors' scales, and interpretable models for most domains [64].

The process of regression tree analysis involves two steps. Initially, the predictor space, defined by potential feature variables, is segmented into unique, non-intersecting areas. Secondly, for each observation that falls into a region, the mean response in the training set is predicted. In principle, the algorithm can generate any shape to divide the space. However, it chooses to use high-dimensional rectangles or boxes to make data interpretation easier. The objective is to identify boxes that minimize the residual sum of squares (RSS) as follows:

$$\sum_{i=1}^{J} \sum_{i \in R_j} \left(y_i - \hat{y}_{R_j}\right)^2 \tag{4.4}$$

where R_j is the rectangle or box corresponding to the j^{th} terminal node. $\hat{y}_{R_j}$ is the predicted response associated with R_j.

The leaf size in the regression trees method is an important hyperparameter that controls the minimum number of samples required to be at a leaf node. It is used to prevent overfitting by stopping the tree from growing beyond a certain point [61]. A smaller leaf size can lead to a more complex model that is more likely to overfit the training data, while a larger leaf size can lead to a simpler model that is more likely to underfit the training data. The ideal leaf size is contingent on the specific dataset and problem and can be ascertained using methods like cross-validation [65]. In the data set used in this thesis, the fine and medium tree methods result in overfitting and

higher RMSE; Therefore, the coarse tree model: leaf size of 36, was the best performing model (Table 4.3 and Table 4.4).

Table 4.3. The regression test results of the first modal frequency obtained via regression tree methods.

Model Type	***Minimum Leaf Size***	***RMSE***	R^2	***Training Time (s)***	***Prediction Speed (Obs/s)***
Fine Tree	*4*	0.033	0.713	31	422992
Medium Tree	*12*	0.032	0.731	8	533610
Coarse Tree	*36*	0.03	0.768	2	2195214

Table 4.4. The regression test results of the second modal frequency obtained via regression tree methods.

Model Type	***Minimum Leaf Size***	***RMSE***	R^2	***Training Time (s)***	***Prediction Speed (Obs/s)***
Fine Tree	*4*	*0.032*	*0.698*	*103*	*288525*
Medium Tree	*12*	*0.031*	*0.717*	*28*	*240942*
Coarse Tree	*36*	*0.03*	*0.733*	*10*	*381202*

The predicted and true frequencies obtained via the coarse tree regression model are presented in the time domain in Figure 4.4. The graph of the predicted frequency versus true frequency is shown in Figure 4.5, while Figure 4.6 displays the residuals in the predictions on the frequency base. The graphs show that the error rate in the tree method does not change depending on the frequency change.

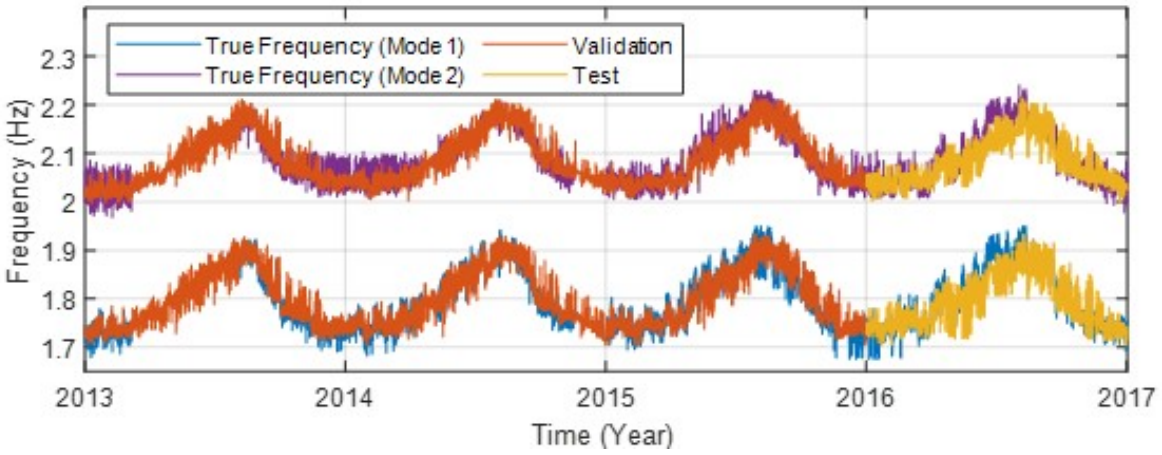

Figure 4.4. Regression results of modal frequencies using the coarse regression trees method.

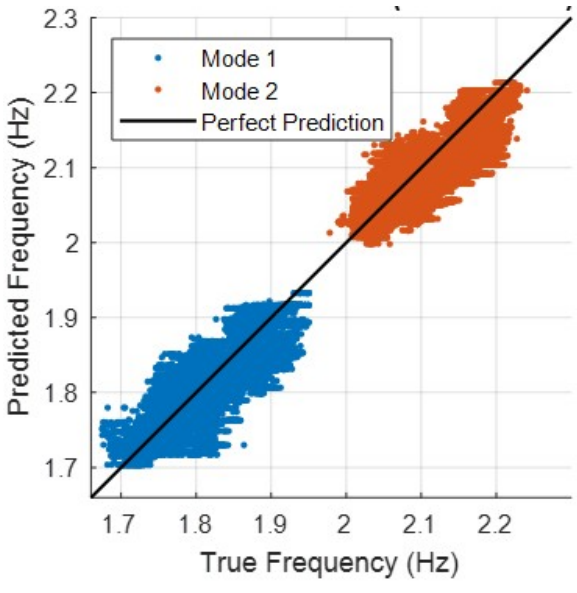

Figure 4.5. Predicted and true frequency comparison for the coarse regression trees method.

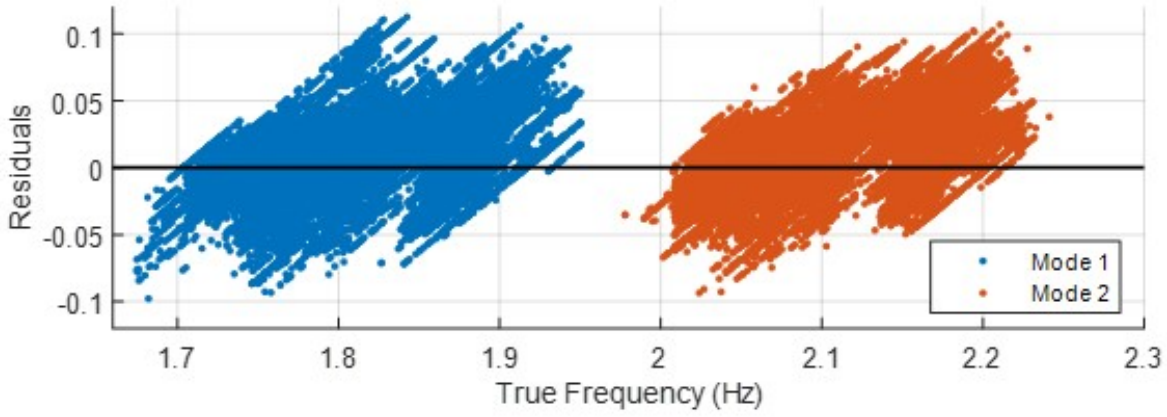

Figure 4.6. Residuals of predictions for the coarse regression trees method.

4.3. Support Vector Machines

Support Vector Machines (SVMs) algorithms are used for regression analysis as well as classification. In the context of SVM, the underlying principle involves identifying a hyperplane that optimally separates the data into distinct classes. The main aim is to identify a hyperplane that strategically maximizes the margin between two distinct classes. The margin is the shortest distance from the hyperplane to any data point. The hyperplane separates the data points into different classes [66].

A common application of SVM is to classify data that cannot be separated by a linear boundary. In such cases, kernel functions help SVMs make linear separations in higher dimensions when the data is not linearly separable [67]. The following steps provide a summary of the SVM algorithm [68]:

(i) Collect the data.
(ii) Choose a kernel function and its parameters.
(iii) Formulate the optimization problem.
(iv) Solve the optimization problem to get the optimal hyperplane.
(v) Use the hyperplane to classify new data sets.

Compared to other machine learning algorithms, SVMs have a few advantages. Even if the number of dimensions is greater than the number of samples, they are effective in high-dimensional spaces [69]. SVMs are a type of machine learning model that only needs some of the training data to create the decision function. This makes SVMs use less memory than other models. Another advantage of SVMs is that they can adapt to diverse types of data. This is because the decision function, which is the rule that determines the output of the model, can use different kernel functions. Kernel functions are mathematical functions that transform the data into a new space where it is easier to separate [66].

Linear SVMs are a type of machine learning model that works well for linearly separable data. This means that the data points can be divided into two groups by using a straight line. Non-linear SVMs are used for non-linearly separable data; such data points cannot be classified by a straight line. In this thesis, a technique known as the kernel trick that transforms the data points into a higher dimensionality where they can be separated by hyperplanes or other mathematical functions is used. The classifier used for such non-linear data is described as a non-linear SVM classifier [69]. Within the scope of this thesis, the linear, quadratic, and cubic SVM methods are tested, and results are presented in Table 4.5 and Table 4.6. Among these methods, the quadratic SVM method had the lowest RMSE value.

Table 4.5. The regression test results of the first modal frequency obtained via SVM methods.

Model Type	***RMSE***	R^2	***Training Time (s)***	***Prediction Speed (Obs/s)***
Linear	0.025	0.805	3717	9090
Quadratic	0.023	0.852	8832	17040
Cubic	0.033	0.72	9526	18452

Table 4.6. The regression test results of the second modal frequency obtained via SVM methods.

Model Type	***RMSE***	R^2	***Training Time (s)***	***Prediction Speed (Obs/s)***
Linear	0.025	0.84	2106	5991
Quadratic	0.023	0.867	6861	6640
Cubic	0.033	0.721	10179	8128

The predicted and true frequencies obtained via the quadratic SVM regression model are presented in Figure 4.7. The graph of the predicted frequency versus true frequency is shown in Figure 4.8, while Figure 4.9 displays the residuals in the predictions on the frequency base.

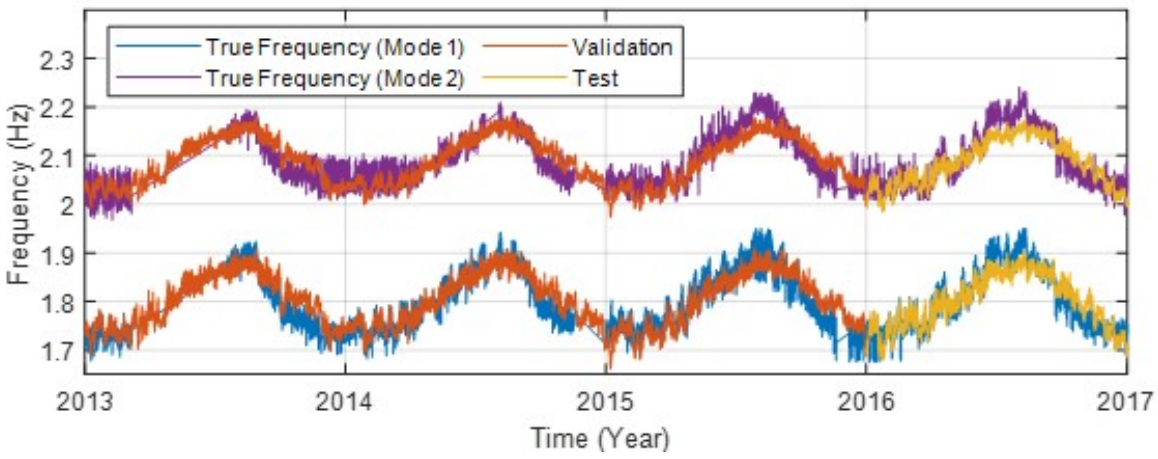

Figure 4.7. Regression results of modal frequencies using quadratic SVM method.

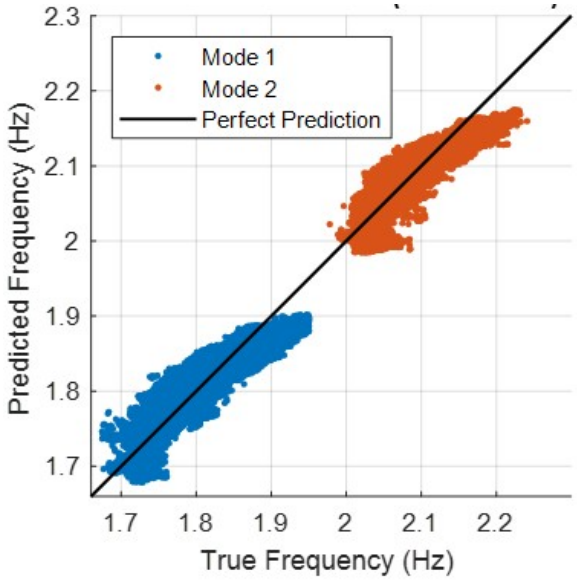

Figure 4.8. Predicted and true frequency comparison for quadratic SVM method.

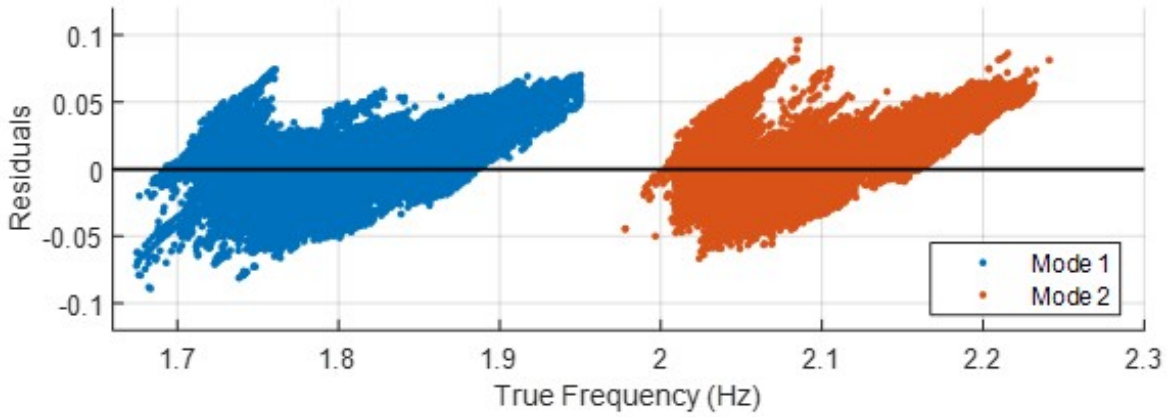

Figure 4.9. Residuals of predictions for quadratic SVM method.

4.4. Neural Network

Neural network regression is a technique that uses Artificial Neural Networks (ANNs) to estimate numerical values. These networks are comprised of interconnected units called neurons, which can perform nonlinear transformations of the input data.

ANNs adapt their internal parameters, such as weights and biases, based on input-output pairs. This adaptation allows them to estimate the relationship between input and target variables.

The most common types of neural network regression are the Shallow Neural Network (SNN) and the Deep Neural Network (DNN). A SNN has three layers: an input layer, a hidden layer, and an output layer. DNN consists of one input layer, more than one hidden layer, and one output layer. In an Artificial Neural Network, the input layer accepts data, the hidden layers execute nonlinear transformations, and the output layer generates the estimated value. The hyperparameters that require tuning for optimal performance in an Artificial Neural Network include the quantity and dimensions of hidden layers, along with the activation functions employed in each layer [70]. The general representation of SNN is shown in Figure 4.10.

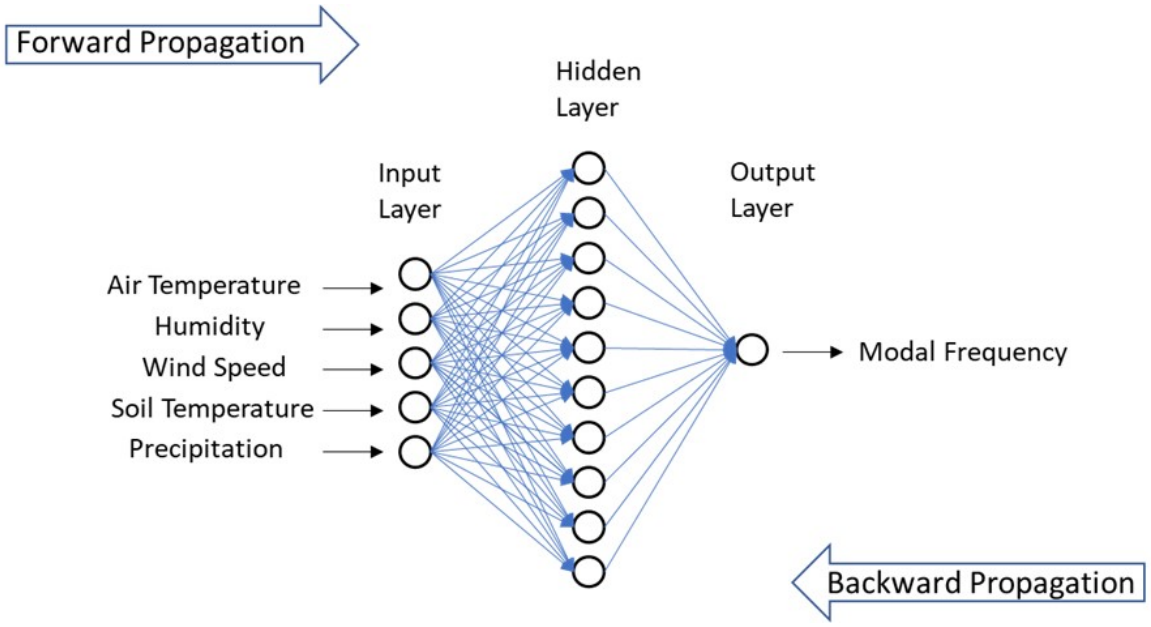

Figure 4.10. The general form of neural network Regression.

The layer size in neural networks is an important hyperparameter that controls the number of neurons in each layer. The number of neurons in a layer influences the

network's capacity for pattern learning and data generalization. A larger layer size can lead to a more complex model that is more likely to overfit the training data, while a smaller layer size can lead to a simpler model that is more likely to underfit the training data. The ideal layer size, contingent on the specific dataset and problem, can be ascertained through methods like cross-validation [71].

Within the scope of this thesis, various layer numbers and different layer sizes were tested. The test results obtained are presented in Table 4.7 and Table 4.8. According to the findings, increasing the number of layers leads to overfitting and a higher error rate. Similarly, increasing the layer size also results in overfitting and does not adequately fit the dataset. The most suitable approach is to use a neural network with a single layer and a layer size of 10 neurons (Narrow Model). Prediction results obtained using the Narrow Neural Network with single layer model are shown in Figure 4.11, Figure 4.12, and Figure 4.13.

Table 4.7. The regression test results of the first modal frequency obtained via the Neural Network method.

Model Type	**Layers**	**1^{st} Layer Size**	**2^{nd} Layer Size**	**3^{rd} Layer Size**	**RMSE**	**R^2**	**Training Time (s)**	**Prediction Speed (Obs/s)**
Narrow	1	10	-	-	0.022	0.85	908	486000
Medium	1	25	-	-	0.027	0.78	1483	359000
Wide	1	100	-	-	0.027	0.78	4617	295000
Bilayer	2	10	10	-	0.024	0.82	2904	509000
Trilayer	3	10	10	10	0.024	0.83	4434	485000

Table 4.8. The regression test results of the second modal frequency obtained via Neural Network

Model Type	**Layers**	**1 st Layer Size**	**2 nd Layer Size**	**3 rd Layer Size**	**RMSE**	**R^2**	**Training Time (s)**	**Prediction Speed (Obs/s)**
Narrow	1	10	-	-	0.022	0.86	2302	262000
Medium	1	25	-	-	0.025	0.83	2727	388000
Wide	1	100	-	-	0.029	0.78	4034	161000
Bilayer	2	10	10	-	0.024	0.84	3268	436000
Trilayer	3	10	10	10	0.028	0.79	4106	883000

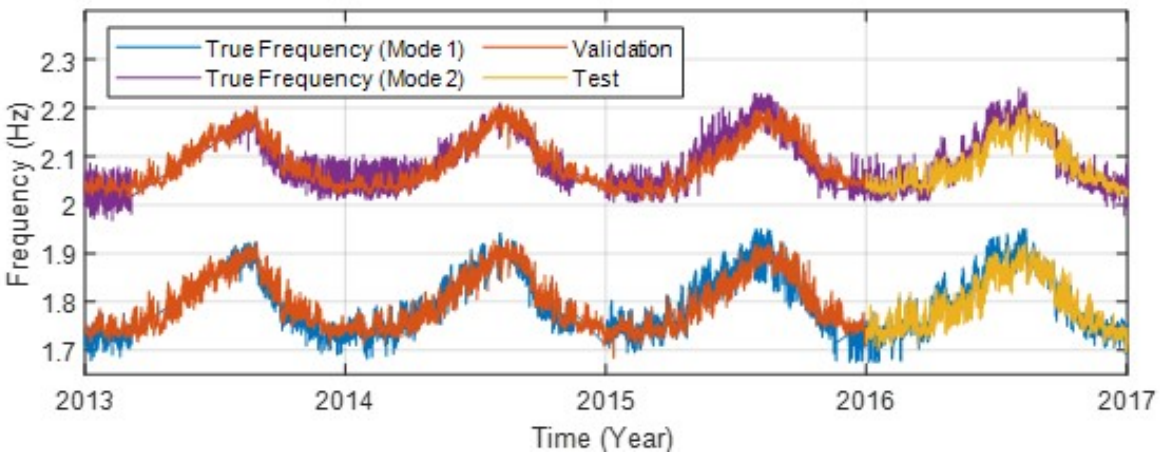

Figure 4.11. Regression results of modal frequencies using the neural network method.

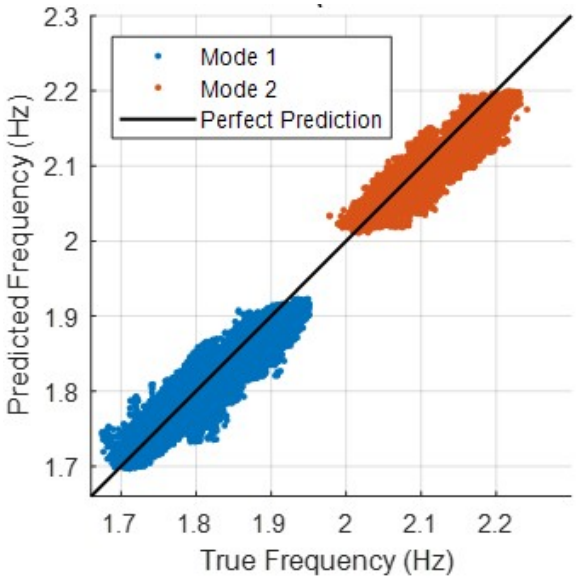

Figure 4.12. Predicted and true frequency comparison for the neural network method.

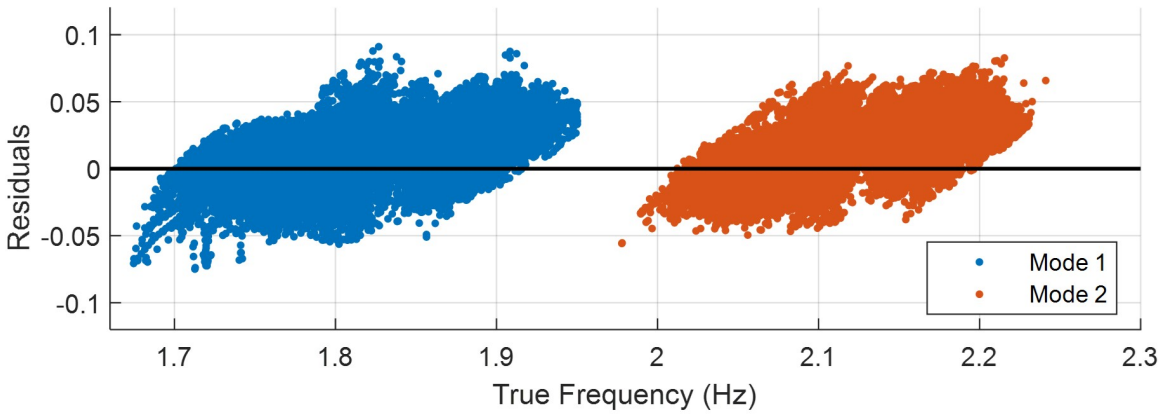

Figure 4.13. Residuals of predictions for the neural network method.

4.5. Regression Results

There exists a diverse array of methods for modeling data sets using machine learning algorithms. To determine which of these methods is most suitable for a given data set, it is essential to explore all available options and select the most appropriate

one. In this thesis, four of these methods across various variations are trained and tested. The resulting test outcomes are presented in Tables 4.9 and 4.10. Based on the findings from these tables, regression methods can be summarized as follows:

- Linear regression is the simplest and fastest method that can capture linear relationships between the features and the target variable. However, its performance may falter with complex or nonlinear data. Adding interactions can improve the performance, but also increase the complexity and computation time. Robust linear regression can handle outliers better, but it is slower than ordinary linear regression.

- Regression Trees is a nonparametric method that can handle nonlinear and categorical data. It can also provide interpretable rules for the predictions. However, it may suffer from overfitting or underfitting depending on the tree size. Fine trees have more splits and can capture more details, but they may also overfit the noise. Coarse trees have fewer splits and can generalize better, but they may also miss some key features.

- SVM is a powerful method that can handle nonlinear and high-dimensional data using different kernels. It can also provide good generalization and robustness. However, it is very computationally expensive and slow, especially for large datasets. Hyperparameters, including the kernel function and regularization parameter, necessitate meticulous adjustment. Linear kernel is the simplest and fastest, but it may not capture the nonlinear patterns. Quadratic and cubic kernels can model more complex relationships, but they may also overfit the data.

- Neural Network is a flexible and powerful method that can approximate any function and learn complex patterns from the data. It can also handle nonlinear and high-dimensional data. However, it is also very computationally intensive and slow, especially for deep networks. The network architecture and hyperparameters, including the number of layers, neurons, and the activation function,

necessitate careful calibration. Narrow networks have fewer neurons and can reduce the risk of overfitting, but they may also lack the capacity to learn the data. Wide networks have more neurons and can increase the expressive power, but they may also overfit the data and require more regularization. Medium networks are a balance between the two. Bilayered and trilayered networks have more layers and can learn more abstract and hierarchical features, but they may also suffer from the vanishing or exploding gradient problem and require more advanced optimization techniques.

Table 4.9. The regression test results of first modal frequency.

Model Type	**Preset**	**RMSE**	**R^2**	**Training Time (s)**	**Prediction Speed (obs/s)**
Linear Regression	Linear	0.025	0.834	7	3603799
	Interactions Linear	0.023	0.859	5	2408904
	Robust Linear	0.025	0.839	1	3420166
	Stepwise Linear	0.023	0.859	41	425373
Tree	Fine Tree	0.033	0.713	31	422992
	Medium Tree	0.032	0.731	8	533610
	Coarse Tree	0.03	0.768	2	2195214
SVM	Linear	0.025	0.84	2106	5991
	Quadratic	0.023	0.867	6861	6640
	Cubic	0.033	0.721	10179	8128
Neural Network	Narrow	0.022	0.868	2302	262748
	Medium	0.025	0.835	2727	388829
	Wide	0.029	0.783	4034	161822
	Bilayered	0.024	0.845	3268	436082
	Trilayered	0.028	0.798	4106	883123

Table 4.10. The regression test results of second modal frequency.

Model Type	**Preset**	**RMSE**	**R^2**	**Training Time (s)**	**Prediction Speed (obs/s)**
Linear Regression	Linear	0.025	0.805	9	939011
	Interactions Linear	0.024	0.83	6	821263
	Robust Linear	0.025	0.806	5	3405903
	Stepwise Linear	0.024	0.83	129	418928
Tree	Fine Tree	0.032	0.698	103	288525
	Medium Tree	0.031	0.717	28	240942
	Coarse Tree	0.03	0.733	10	381202
SVM	Linear	0.025	0.805	3717	9090
	Quadratic	0.023	0.852	8832	17040
	Cubic	0.033	0.72	9526	18452
Neural Network	Narrow	0.022	0.856	908	486326
	Medium	0.027	0.784	1483	359147
	Wide	0.027	0.782	4617	295543
	Bilayered	0.024	0.824	2904	509086
	Trilayered	0.024	0.831	4434	485642

5. REAL-TIME STRUCTURAL HEALTH MONITORING SOFTWARE (AISHM)

5.1. Description of AISHM

Real-time SHM is a system that involves collecting, processing, and analyzing vibration data from accelerometers placed in the structure in real-time or near real-time. These systems can enable faster and more accurate detection and localization of structural damage, as well as timely intervention and intervention in emergencies. In this thesis, a new real-time SHM software called "AISHM" is proposed and implemented on the Hagia Sophia SHM system. The meaning of AISHM is "Artificial Intelligence in Structural Health Monitoring".

AISHM is an algorithm that uses real-time acceleration data of Hagia Sophia to calculate the dynamic parameters of the structure, such as its natural frequencies, mode shapes, and damping ratios. This algorithm proposes a new method for real-time SHM systems using machine learning algorithms. Unlike existing systems that rely on sensors or visual inspections, the proposed method estimates real-time modal frequencies of the structure using data from weather stations via machine learning algorithms. In this way, the proposed method can identify any structural anomalies that may signify damage or degradation by comparing the measured modal frequencies with the predicted modal frequencies. This comparison allows the method to detect any deviations from the expected behavior of the structure, which may indicate a need for further investigation or intervention.

In addition to modal frequency estimation, AISHM also calculates and saves the results of modal frequency, modal damping ratio, and mode shapes which are the basic dynamic parameters of the structure, in real time. The algorithm also applies the MAC, COMAC, and ECOMAC methods to monitor the changes in the mode shapes over time and saves the results in txt format. Moreover, the algorithm compares the

real-time mode shapes with a predefined reference mode shape to detect any deviations. However, the algorithm may produce inaccurate results if there is a time delay problem in the devices, which affects the mode shape estimation and the MAC, COMAC, and ECOMAC values. Table 5.1 shows the inputs and outputs of the AISHM algorithm.

Table 5.1. The inputs and outputs of the AISHM algorithm.

INPUTS	**OUTPUTS**
Acceleration	Modal Frequency
Temperature	Modal Damping
Humidity	Mode Shapes
Wind Speed	MAC, COMAC, ECOMAC
Precipitation	Predicted Modal Frequency
	3-D Animation
	Time-Delay Detection
	Earthquake Detection

5.2. General Properties of AISHM Algorithm

The AISHM algorithm is developed using MATLAB 2023b, the latest release of the popular software for numerical computing and data analysis. To facilitate the interaction and visualization of the developed algorithm, a user interface was designed and implemented using the Matlab App Designer tool, which is an interactive development environment for creating desktop and web apps in Matlab. The user interface consists of various components that allow the user to adjust the parameters, run the algorithm, and display the results.

Matlab App Designer is a modern and interactive development environment for designing and programming apps in Matlab. It has many important advantages that make it preferred over other methods, such as:

- App Designer provides a fully integrated version of the Matlab Editor and a large set of interactive UI components, such as buttons, sliders, tables, charts, etc.

- App Designer facilitates app distribution either by packaging them into installer files via its toolstrip or by generating standalone desktop or web applications.

- App Designer supports 2-D and 3-D plotting, axes interactions, figure interactions, and printing.

Matlab app designer also has some important disadvantages that should be noted. Researchers who will use or develop the AISHM algorithm should be aware of these disadvantages. These disadvantages can be listed as follows:

- App Designer is a relatively new feature in Matlab, and it may have some bugs or compatibility issues with older versions of Matlab or other toolboxes. It may also change or be updated in future releases of Matlab, which could affect the functionality or appearance of AISHM.

- App Designer may have some performance issues or memory leaks when running complex or large apps. It may also take longer to load or save apps in App Designer than in GUIDE (GUI Design Environment), which is the older interactive design environment for building apps in Matlab.

Given the benefits and drawbacks of this Matlab App designer, it is evident that designing a holistic software such as AISHM as a single unit may result in various software and performance challenges. To optimize the software's performance and efficiency, the algorithm is divided into four modules. The names of modules and short explanations are listed below:

- Module 1 - Main Analyzer, this module is responsible for the main processes and analysis of AISHM. Some of the capabilities of this module are data col-

lection; calibration; preprocessing; modal frequency estimation; modal frequency prediction; modal damping estimation; mode shape estimation; MAC-COMAC-ECOMAC estimation; and saving all these outputs as text files for AISHM.

- Module 2 - Monitoring, this module reads, compiles, and visualizes the data by the analysis obtained and saved in the first module.

- Module 3 - The Time-Delay Calculator, this module is responsible for determining the time delay among stations. It visualizes the time shift in real-time by calculating the cross-correlation function of the data received in the previous hour.

- Module 4 - 3-D Animator: This module is responsible for showing real-time animation of Hagia Sophia. It reads the data in real-time, obtains the displacements, and performs the animation by moving the predefined three-dimensional structure.

- Module 5 - Earthquake Tracer: This module tracks the latest earthquakes published by the Kandilli Observatory and Earthquake Research Institute website. When it detects an earthquake that meets predefined criteria, it takes the date and time of the earthquake and extracts the data corresponding to this time interval from the acceleration data. It applies many different earthquake analyses and records the results. It also visualizes the results of these analyses.

5.3. Analyzer Module

This module is the core component of the AISHM algorithm, which performs modal analysis and estimation for structural health monitoring. It computes the modal frequency, modal damping ratio, MAC, COMAC, and ECOMAC of the structure. It also estimates the modal frequency based on the measured data. This module consists of eleven stages, as summarized in Figure 5.1. The Analyzer module receives real-time acceleration data from Hagia Sophia, which are transmitted to the Early Warning

and Emergency Response Laboratory of the Earthquake Engineering Department at Boğaziçi University. The acceleration data are stored hourly by default in GCF format. However, data transfer is internet-based and may be interrupted by internet or power outages. Therefore, detecting missing and corrupt files is a crucial step when reading GCF files. The first task of the Analyzer module is to perform this step.

GCF format is a proprietary format developed by Guralp, the company that produces the accelerometers used in Hagia Sophia. These files can only be read by specific open-source codes provided by the company. The second stage of the Analyzer module is to use these codes to read all GCF files and sort them by the stations that recorded them.

After reading the data, the Analyzer module merges them according to the dates and stations. To ensure data integrity, it assigns 0 (zero) to the missing data. The analysis results for the data assigned to zero are also 0 (zero), indicating the absence of data for that time period.

After this stage, the Analyzer module applies a two-stage correction process to the records. The first stage is baseline correction, which adjusts the records that have a baseline shift. The second stage is filtering, which removes the low-frequency waves in the records. For this purpose, a high-pass filter with a constant cutoff frequency of 0.2 Hz was applied to all records.

In order to use the acceleration records for 3-D animation, they are down-sampled by a factor of five from 100 Hz to 20 Hz using the decimate function. This is done to obtain a suitable sample rate for the animation. Since the human eye cannot perceive more than twenty frames per second, a higher sample rate would only decrease the computer's performance without improving the visual quality. In stage five, after preparing and correcting the records, the Analyzer module performs the estimation of modal parameters in real-time. For modal frequency estimation, it uses the automatic selection method of the peak of the Fourier amplitude spectrum. For this purpose, it

divides 1 hour of acceleration data into 60 segments and obtains a modal frequency for each 1-minute data segment. Before computing the Fourier amplitude spectrum, it applies a band filter to the records around the frequency value of that mode. By default, the filter ranges are 1.4-1.85 Hz for the first mode and 1.85-2.3 Hz for the second mode. The user can modify these values if needed.

In the next step, the Analyzer module estimates the modal damping ratio in real-time. It uses the half-power bandwidth method, which is based on the Fourier amplitude spectrum. Therefore, the smoothing ratio of the Fourier amplitude spectrum directly affects the calculated modal damping ratio. The optimal smoothing ratio is the lowest possible value that eliminates unrealistic peaks in the Fourier amplitude spectrum. In the Analyzer module, this ratio is set to 0.036 by default. The user can change this value if desired. To obtain more stable results for modal damping, a 5-minute window interval was chosen. It was observed that shorter window intervals were unstable and more influenced by environmental factors.

The Analyzer module then computes the mode shapes in real-time for each hourly data segment. It uses the Frequency Domain Decomposition method for mode shape estimation. It compares the obtained mode shapes with the reference mode shape using the modal assurance criterion (MAC), coordinate modal assurance criterion (COMAC), and effective coordinate modal assurance criterion (ECOMAC) methods. The reference mode shape is the average mode shape obtained from the 1-month record of January 2013. However, since the installation of the system, calibration and time drift problems have frequently occurred at GAL3 and GAL4 stations. Although these malfunctions were temporarily resolved from time to time, they recurred. For this reason, the accuracy of the mode shapes obtained in real-time is still questionable. This problem does not affect the modal frequency and modal damping estimations, but it significantly affects the mode shape estimations. Therefore, to obtain correct mode shape estimations, the faults in GAL3 and GAL4 stations must be eliminated.

The last task of this module is to estimate the modal frequency of the structure

in real-time by applying machine learning algorithms. The prediction model used in this module is the narrow-sized Neural Network Method, which has the lowest error rate among the methods compared in Chapter 5. This method requires real-time atmospheric data to make accurate modal frequency predictions. Therefore, the module collects atmospheric data from three different sources to ensure reliability and continuity. One of the sources is the official website of the General Directorate of Meteorology of Turkey (MGM), which provides meteorological data for various locations in Turkey. The other two sources are Openweather.com and tomorrow.io, which are open-source platforms that offer meteorological data for free through an Application Programming Interface (API). The module uses the web scraping technique to extract data from the MGM website, while it uses the API to access data from the open-source platforms. The module updates the atmospheric data every 30 minutes and uses the Neural Network prediction model to estimate the modal frequencies of the structure.

After completing the analysis described in the previous sections, the Analyzer module stores all the results in a text file format. This format enables straightforward data access and manipulation, along with compatibility with various software applications. The text file contains the following information: the modal frequency, the damping ratio, the mode shape, and the predicted modal frequency. The text file also includes the atmospheric data used for the prediction, such as temperature, humidity, wind speed, and precipitation. The text file is named according to the date and time of the analysis. The text file is saved in a designated folder on the computer. The screenshot of the analyzer module is displayed in Figure 5.2.

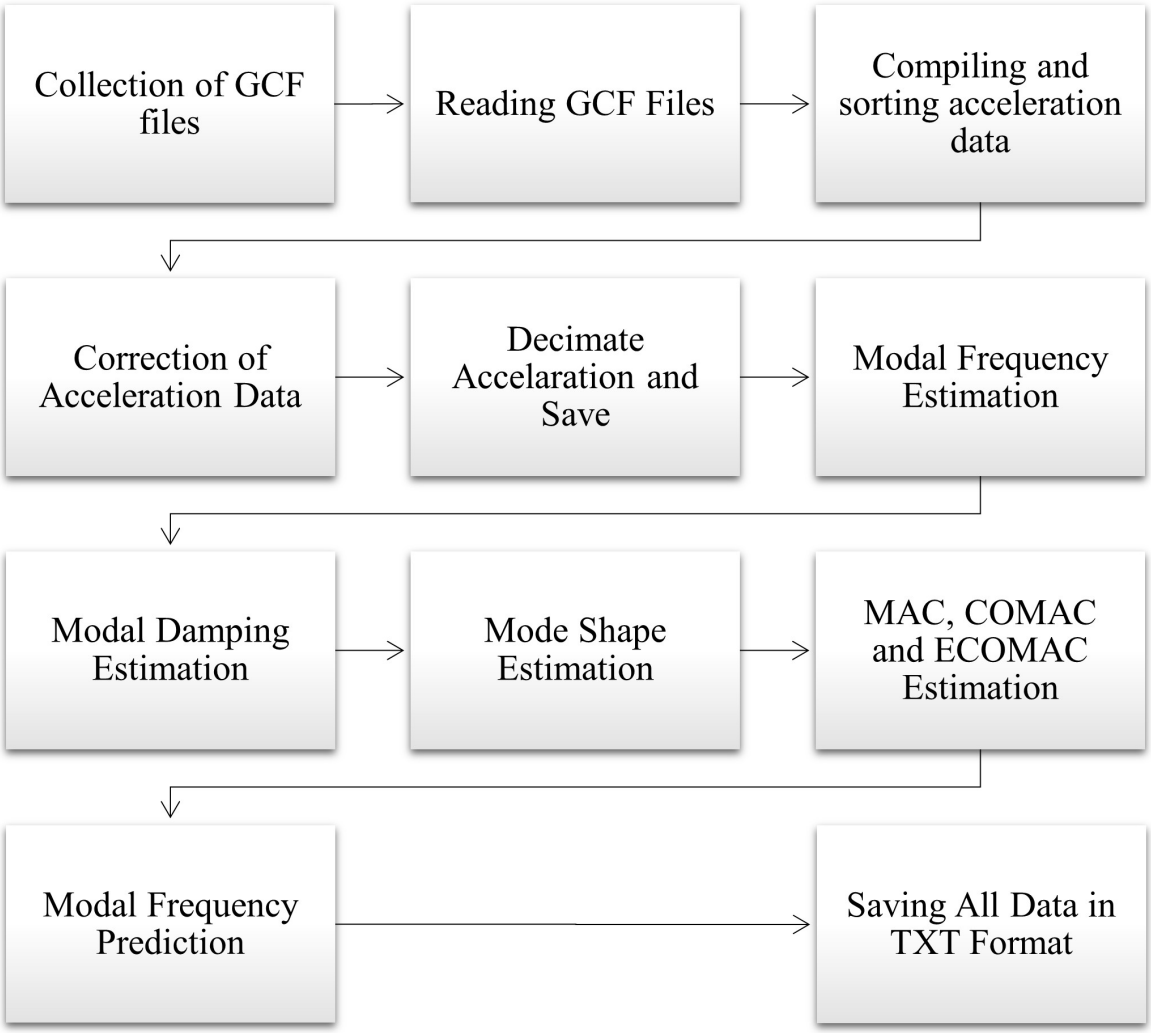

Figure 5.1. The flowchart of the analyzer algorithm.

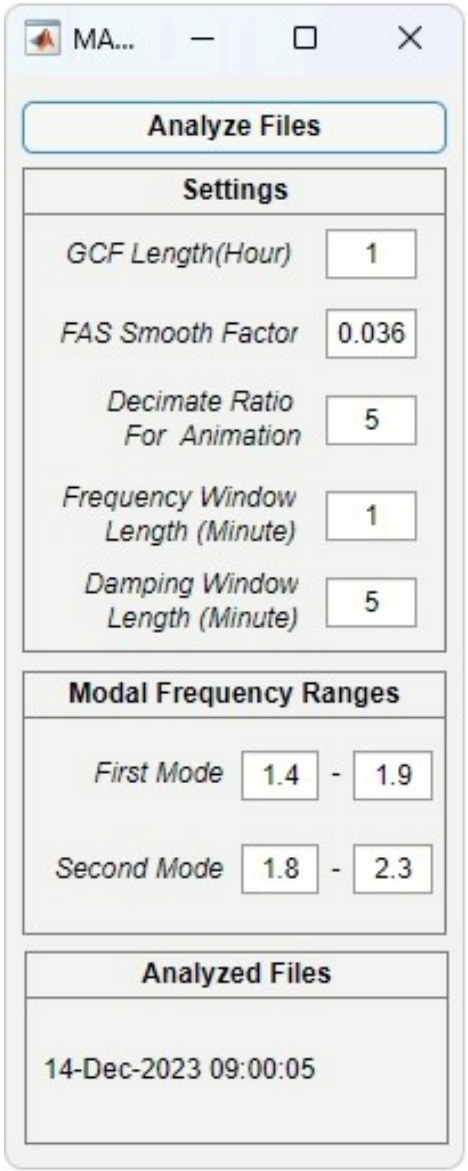

Figure 5.2. Screenshot of the "Analyzer Module".

5.4. 3-D Animation Module

3-D animation of a structure is a valuable tool in SHM that allows visualizing the structure and its behavior in real time under different loading conditions. This visualization helps engineers better understand and interpret the dynamic behavior of the structure in different modes. For example, the best way to distinguish and define horizontal and torsional modes is to examine the 3-D animation of the structure under

different modes.

Another purpose of 3-D animation is to be an additional tool to support frequency-based analyses in the detection of structural abnormalities. Structural damage often leads to permanent changes in mode shapes and displacements, which can be better understood and visualized through 3-D animations. In some cases, animations can also help identify damage localization.

3-D animation can visualize sensor malfunctions quickly and easily. SHM systems are prone to calibration failures in sensors, connection problems, power supply failures, and time drift problems. These problems can cause the system to fail or give incorrect results. 3-D animations are sensitive to sensor malfunctions and therefore have a key place in detecting such sensor malfunctions.

The Animation Module comprises six main stages. The first stage involves reading the acceleration data from the file "acc.mat", which is generated by the Analyzer module as a one-minute package every minute. The acceleration data is then filtered according to the filter range specified in the user interface. Next, the displacement data is derived by performing double integration. Subsequently, the three-dimensional model of Hagia Sophia, which is predefined, is loaded. The displacements are applied to the joint points of the three-dimensional model. Finally, the animation is created by repeating this process at a rate of twenty times per second. The steps of the whole animation process are shown in Figure 5.3. The screenshot of this module is shown in Figure 5.4.

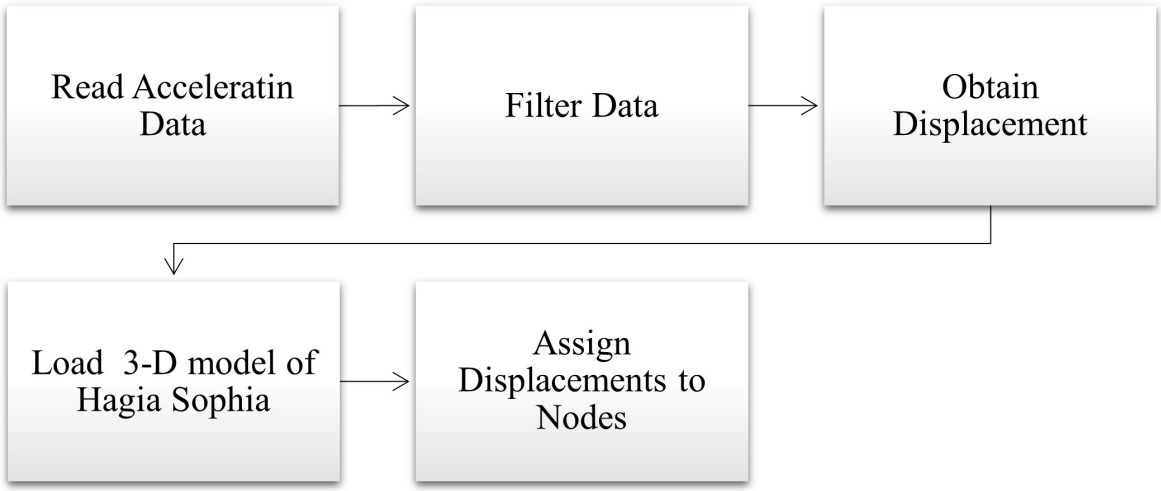

Figure 5.3. The steps of the animation process.

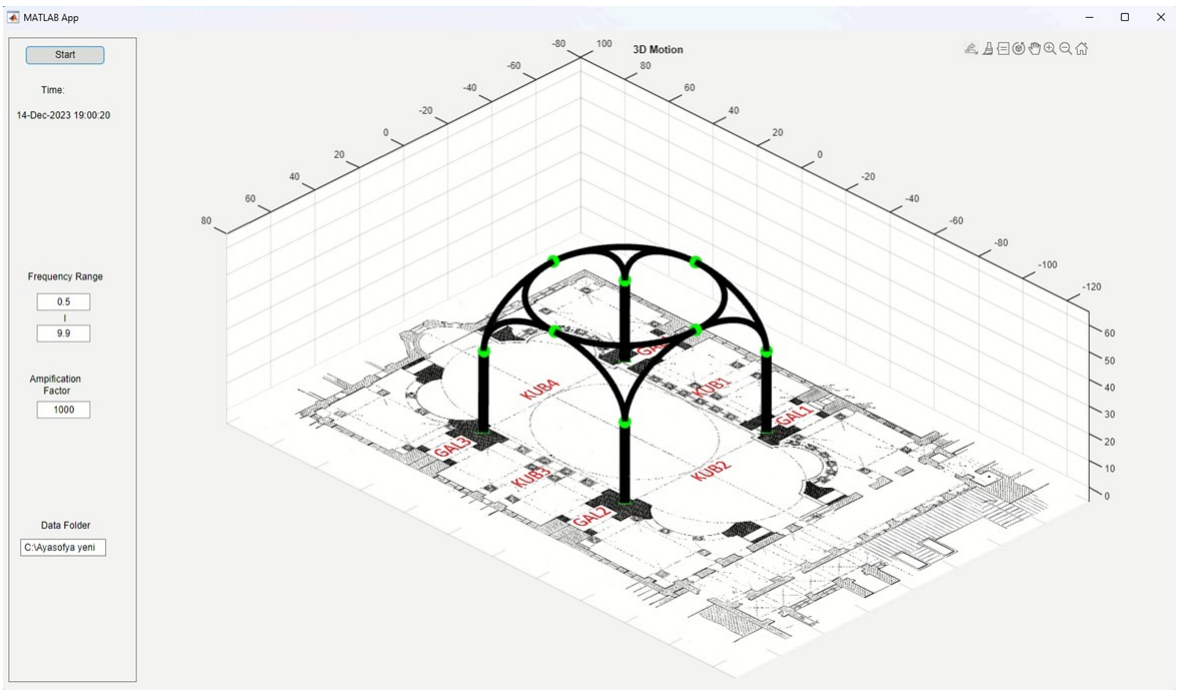

Figure 5.4. Screenshot of "3-D Animation Module".

5.5. Monitoring Module

Monitoring the modal parameters of a structure in real-time can provide valuable information for detecting structural abnormalities. Commonly monitored dynamic parameters are modal frequency, which characteristic property of a structure; modal damping ratio, which is the measure of how quickly the vibration decays; and mode shape, which is the spatial distribution of the displacement at a given frequency. Permanent changes in these dynamic parameters are often associated with damage. On the other hand, temporary changes in these parameters are generally associated with environmental effects. The proposed algorithm (AISHM) dynamically predicts the modal frequency of the input signal and eliminates the environmental effects that cause temporary frequency variations. The general flowchart of the Monitoring module is presented in Figure 5.5.

Figure 5.5. The steps of the monitoring process.

The main task of the Monitoring Module is to display the dynamic parameters, modal frequency predictions, and atmospheric parameters that are obtained and recorded in the analyzer module. The first tab of the monitoring module which is called "FAS Tab" displays the Fourier amplitude spectra and the modal frequency variation of each station. The data on modal frequencies are refreshed every minute. The screenshot of this tab is presented in Figure 5.6.

The second tab which is called the "Modal Frequency Tab" shows a comparison of the measured and predicted modal frequencies. The measured modal frequency is the average of eight stations in the same direction, represented by the blue line. The

predicted modal frequencies are calculated using atmospheric data from three different sources: MGM (black line), Openweather (red line), and Tomorrow.io (green line). The comparison of the real-time estimated and measured frequencies reveals a high degree of similarity. The predicted frequencies are refreshed every thirty minutes. The screenshot of this tab is presented in Figure 5.7.

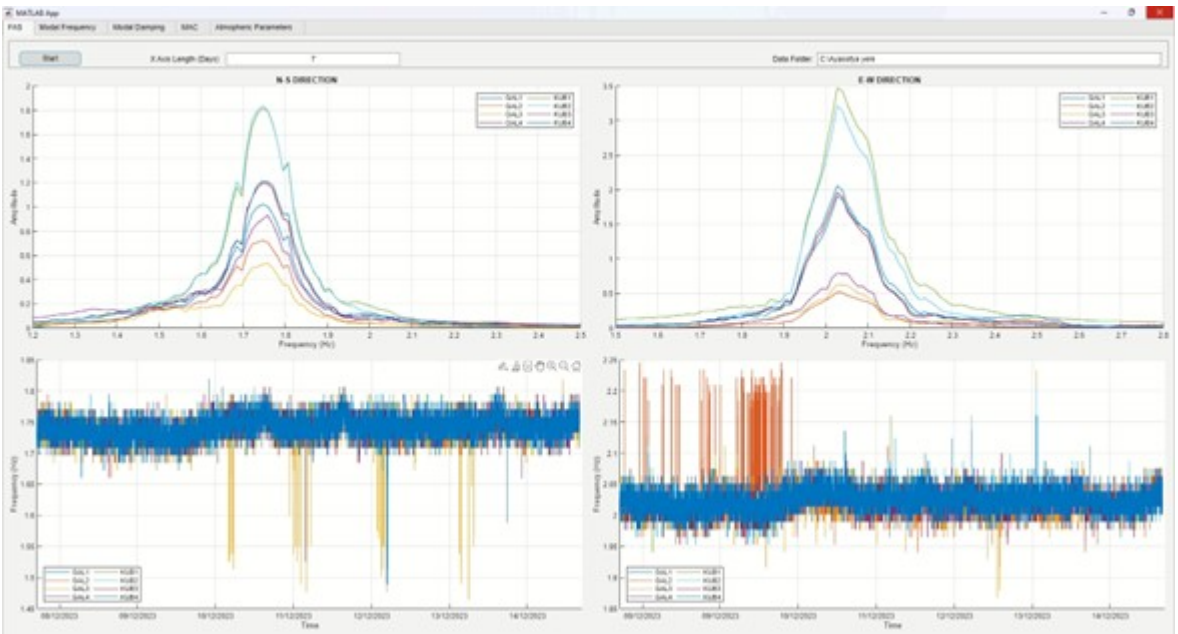

Figure 5.6. The screenshot of the "FAS" tab in the "Monitoring Module".

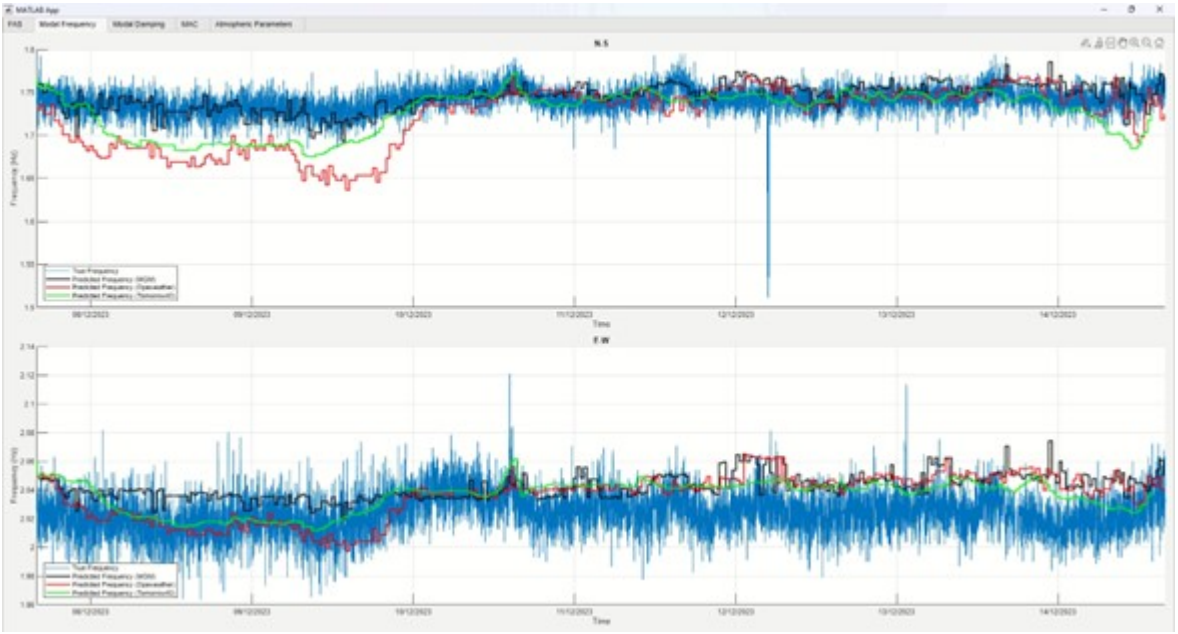

Figure 5.7. The screenshot of the "Modal Frequency" tab in the "Monitoring Module".

The third tab which is called the "Modal Damping Tab" presents the modal damping ratios for each station. In this section, modal damping ratios obtained from all stations are shown separately. The average modal damping ratio across all stations is also displayed with a thicker line. The data on modal damping ratios are refreshed every five minutes. The screenshot of this tab is presented in Figure 5.8.

The fourth tab which is named "MAC Tab" displays the variations in mode shapes. The blue line indicates MAC, the red line indicates COMAC and the yellow line indicates ECOMAC. These are measures of the correlation between real-time mode shapes and the reference mode shape. Values near one imply that the mode shape is stable, and values near zero imply that the mode shape is changing. Mode shapes depend heavily on the time synchronization at stations. Due to the ongoing time synchronization issue at the GAL3 and GAL4 stations, the MAC values are exceptionally low. However, the COMAC values, which are less affected by time synchronization, are

remarkably high. The data on mode shapes are refreshed every hour. The screenshot of this tab is presented in Figure 5.9.

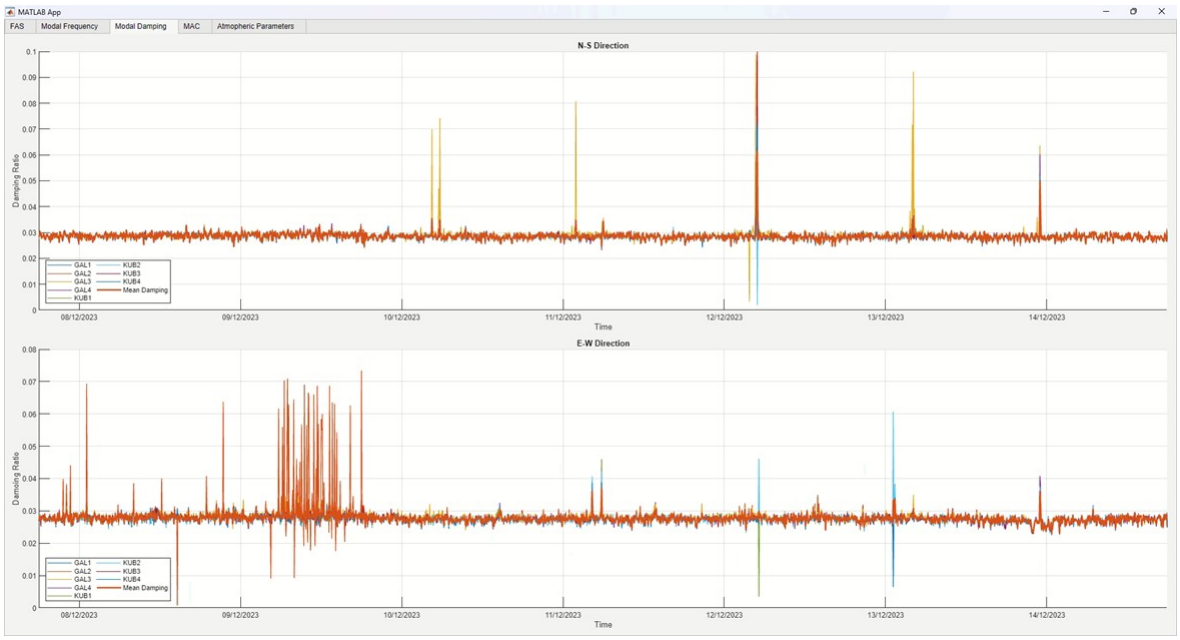

Figure 5.8. The screenshot of the "Modal Damping" tab in the "Monitoring Module".

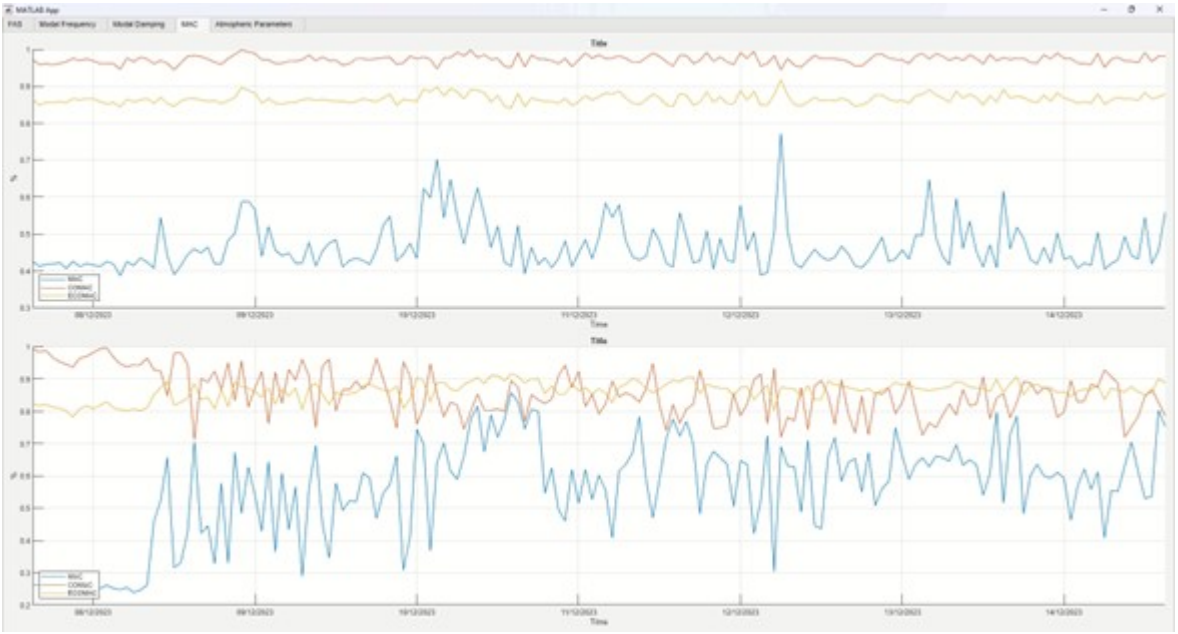

Figure 5.9. The screenshot of the "MAC" tab in the "Monitoring Module".

The fifth and final tab displays the atmospheric data from three different meteorological sources, as previously stated. This tab is named "Atmospheric Parameters Tab". The atmospheric data include Temperature, Wind Speed, Relative Humidity, and Precipitation, respectively. The atmospheric data are refreshed every 30 minutes. The screenshot of this tab is presented in Figure 5.10.

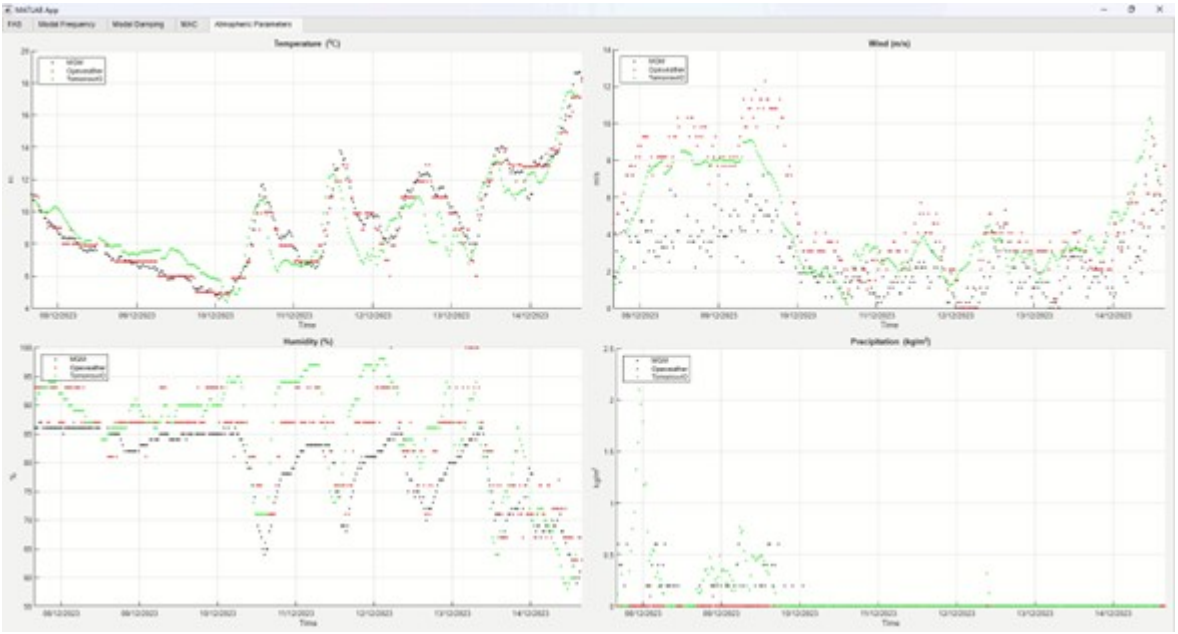

Figure 5.10. The screenshot of the "Atmospheric Parameters" tab in the "Monitoring Module".

5.6. Time-Shift Module

In structural health monitoring systems, time delay refers to the relative delay in data transmission between stations. Time delay is a crucial factor as it directly affects the accuracy of analysis results. In particular, Mode shapes cannot be calculated accurately in a system where there is no time synchronization. This may cause significant errors in damage estimates. Therefore, it is important to constantly monitor the time delay in structural health monitoring systems.

The easiest way to detect time delay in building-type structures is to analyze the movement of the structure in its first mode. The first mode of structures is generally the mode in which there is pure horizontal movement, and all stations move in the same direction at the same time. In this way, the station with a time delay problem can be easily detected as it will move in a different direction from the other stations.

In this thesis, the time shift was obtained by filtering all stations according to mode 1 and calculating the cross-correlation between them. The screenshot of this module is presented in Figure 5.11.

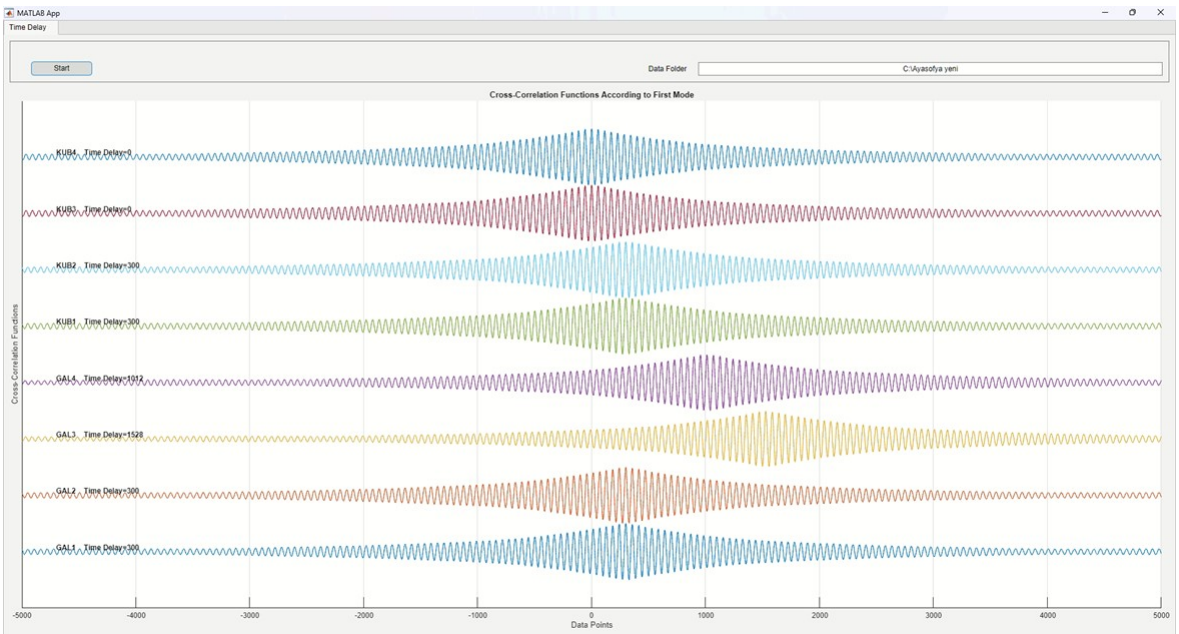

Figure 5.11. The screenshot of the "Time-Shift Module".

5.7. Earthquake Module

Earthquakes are one of the most destructive natural hazards that can induce severe damage to structures and threaten human safety. Hence, real-time detection and analysis of earthquakes is essential for structural health monitoring (SHM) systems, as it can enable rapid and precise assessment of structural damages. In this part of the AISHM algorithm, a novel approach is proposed for detecting and analyzing earthquakes in real-time and developing methods for identifying structural anomalies using the artificial neural network method.

The earthquake module consists of five main stages (Figure 5.12). The first stage is to use the web scraping method to read the latest earthquakes web page

of Kandilli Observatory and Earthquake Research Institute [72] in real-time. This web page contains the last five hundred earthquakes that occurred in Turkey and its surroundings.

The second stage is to eliminate these earthquakes with a four-stage filter according to their distance and magnitude. The filter criteria are as follows:

- Stage 1: Earthquakes with a local magnitude (M_L) more than 4 within a 50 km radius.
- Stage 2: Earthquakes with a local magnitude (M_L) more than 5 within a 150 km radius.
- Stage 3: Earthquakes with a local magnitude (M_L) of more than 6 within a 250 km radius.
- Stage 4: Earthquakes with a local magnitude (M_L) of more than 7 within a 350 km radius.

The distances were calculated by drawing a circle with the Hagia Sophia mosque as the center. The map representation of the radius of these circles is shown in Figure 5.13. These radiuses and magnitudes are predefined, and users can change these filters from the settings menu.

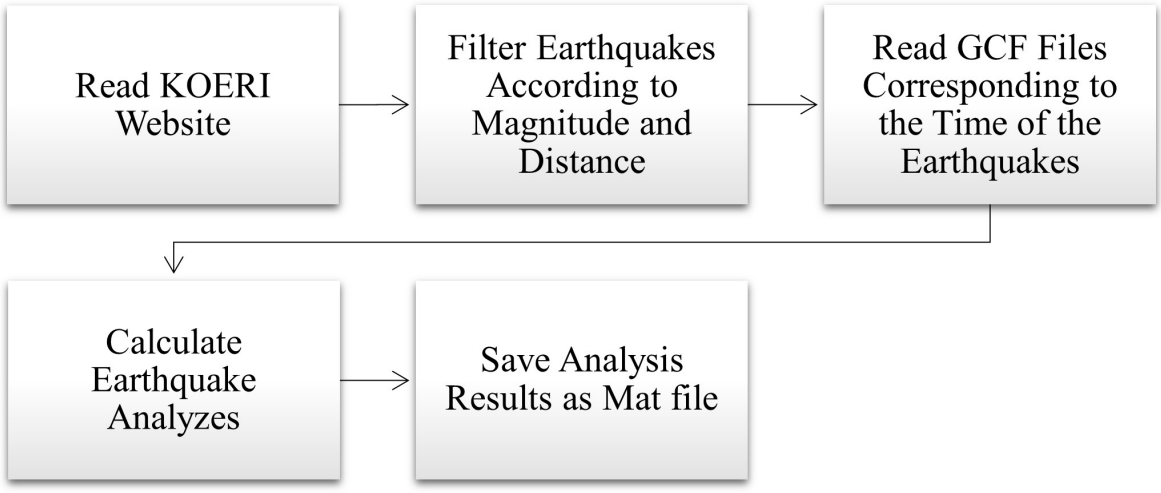

Figure 5.12. The analysis steps of "Earthquake Module".

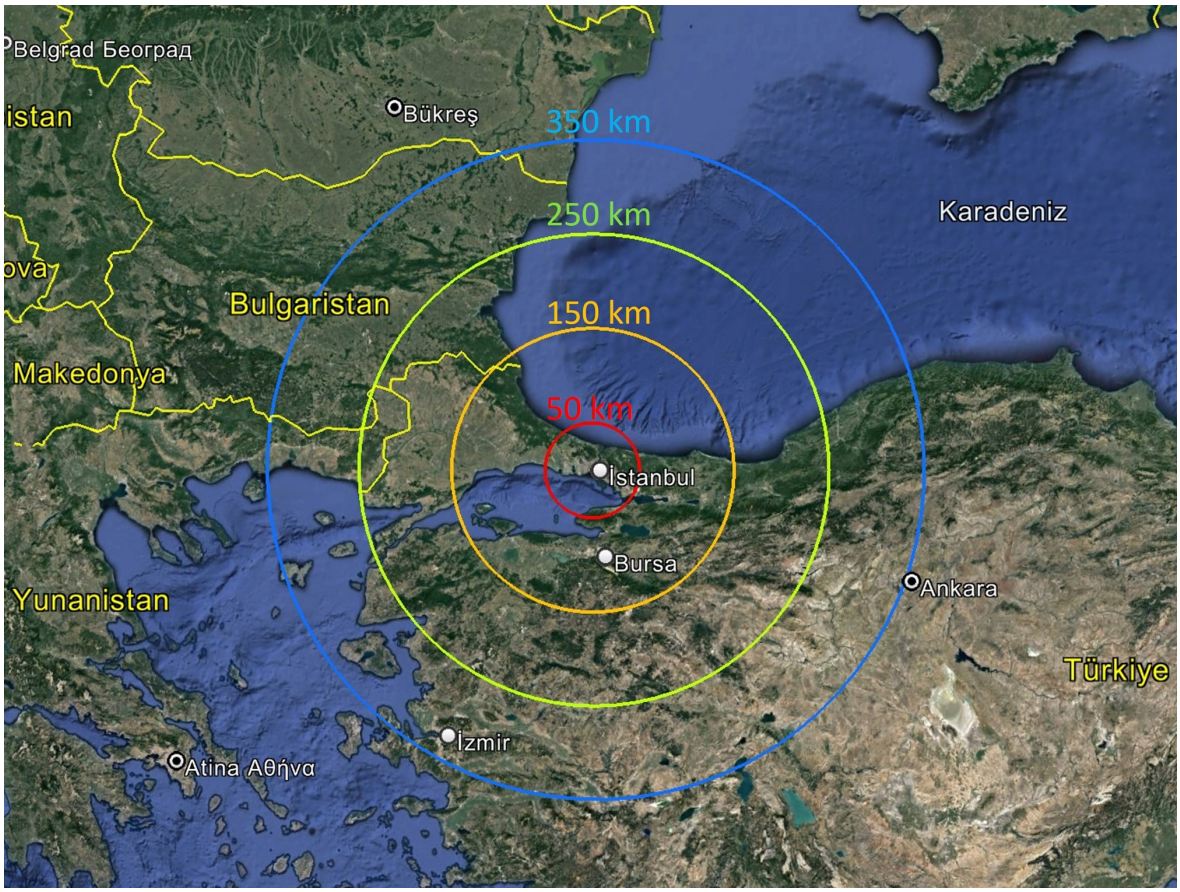

Figure 5.13. The map representation of the circles.

The third stage of the earthquake module is to detect and extract the relevant information from the selected earthquakes. The information includes the time, location, magnitude, and depth of the earthquake. The next phase is to select and extract the GCF files that correspond to the time of the earthquake from the data folder. The GCF files are then read, calibrated, and corrected. The next step is to perform some earthquake analysis processes on the corrected data. The analysis processes are:

- 3-D animation of the structure
- Display of acceleration, velocity, and displacements
- Demonstration of Fourier amplitude spectra
- Demonstration of transfer functions

- Short-time Fourier transform (Spectrogram)
- Response spectra
- Station-based acceleration, velocity, and displacement predictions

These analysis processes help to understand the impact and characteristics of the earthquake on the structure.

The Earthquake Module has a section in the lower left part where past earthquakes are listed. The user can enter the data folder address and click the load earthquakes button to display the names and dates of previously analyzed earthquakes in the earthquake list. The earthquake list also includes the earthquake analysis files recorded in real-time. The user can select an earthquake from this list and click the plot earthquake button to display the recorded information and analysis results about the earthquake. The Info tab contains some information about the earthquake, such as time, location, magnitude, depth, coordinate, and distance to the Hagia Sophia Mosque. If atmospheric data such as temperature, wind, humidity, and atmospheric pressure were recorded at the time of the earthquake, they are displayed in this tab. Using the atmospheric data, the modal frequencies were estimated and presented in this section. Additionally, recording parameters such as record length, sampling rate, horizontal peak ground acceleration, vertical peak ground acceleration, and earthquake significant duration are also displayed in this tab. The screenshot of the Info tab is presented in Figure 5.14.

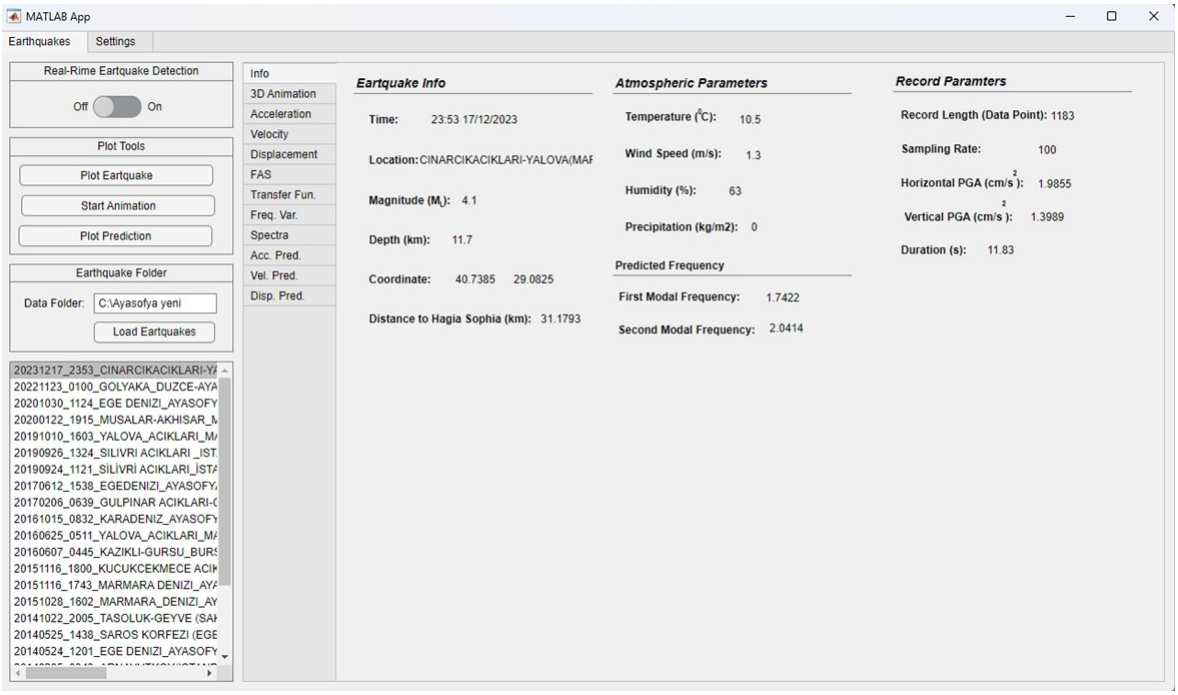

Figure 5.14. The screenshot of the "Info" tab in the "Earthquake Module".

The 3-D animation tab of the earthquake module shows the animation of the structure during the earthquake. The user needs to select the earthquake from the earthquake list and click the start animation button to view the animation. The user can also adjust the animation settings at the bottom of the animation. The settings include the animation time, magnification factor, decimation rate of the records, and filter ranges. The user can customize the animation according to their preferences with these settings. The display of the animation tab in the Earthquake Module is presented in Figure 5.15.

The acceleration, velocity, and displacement tabs display the graphical representation of the acceleration, velocity, and displacement records of the stations, respectively. These records are unfiltered raw data, which means that they have not been processed or modified in any way. The plots can help to visualize the impact of the earthquake on the stations and to compare the records of different stations. The screenshots of acceleration, velocity, and displacement tabs are displayed in Figure

5.16, Figure 5.17, and Figure 5.18, respectively.

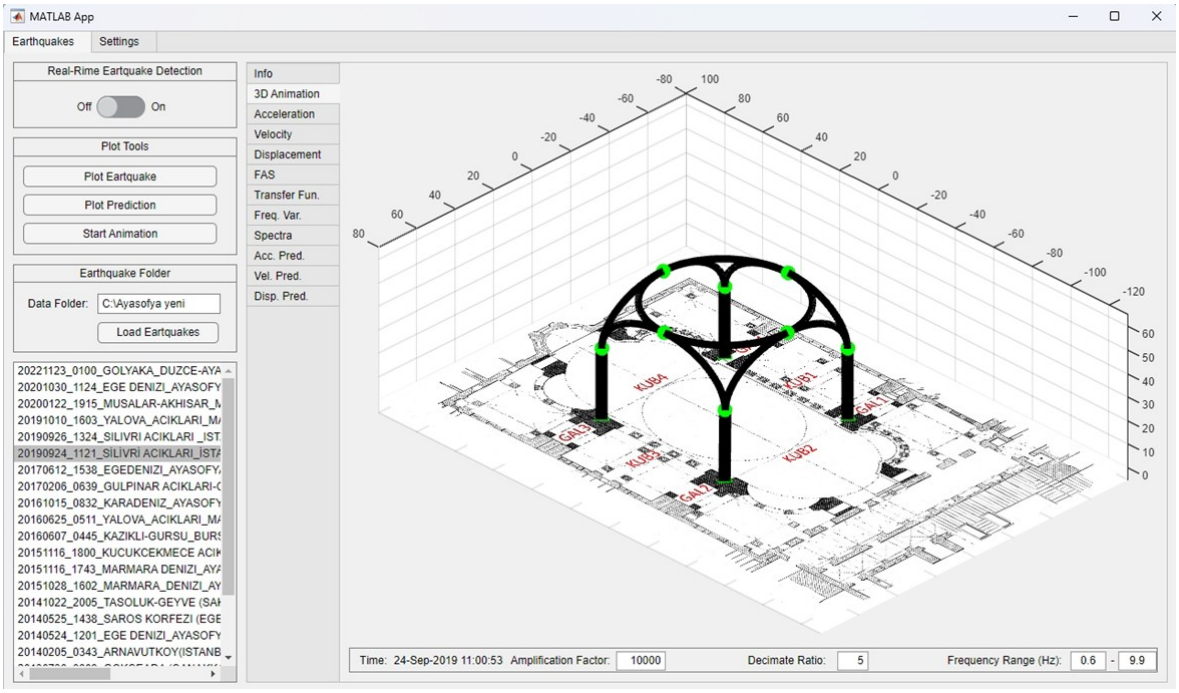

Figure 5.15. The screenshot of the "3-D Animation" tab in "Earthquake Module".

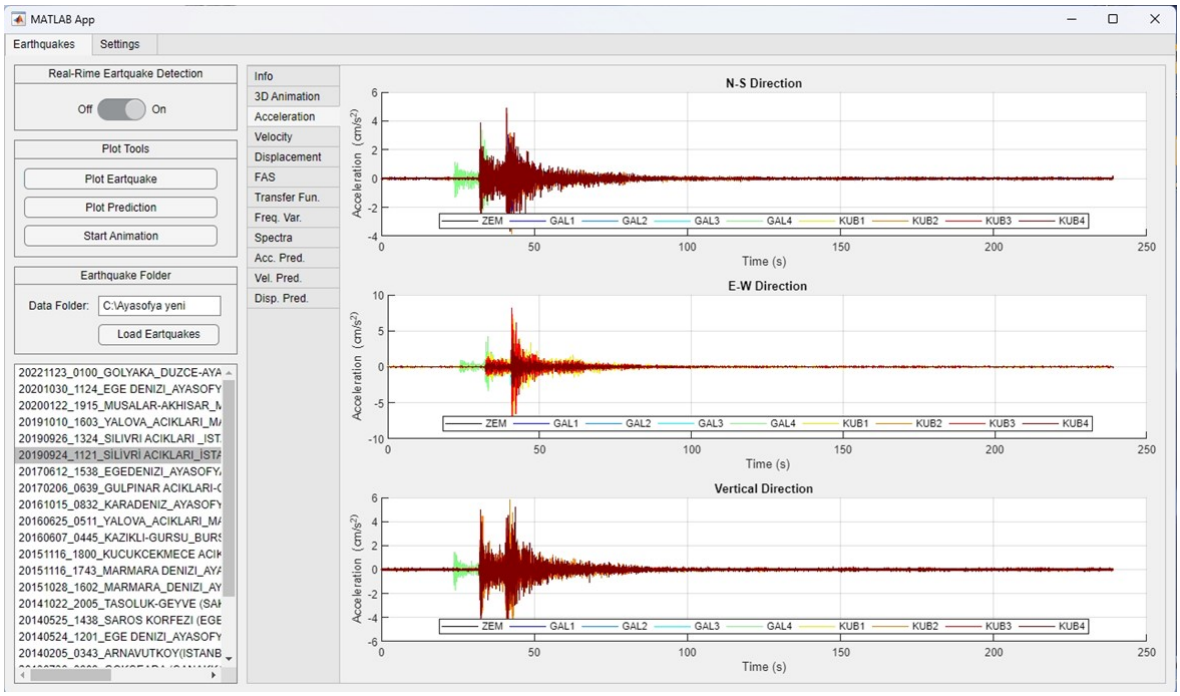

Figure 5.16. The screenshot of the "Acceleration" tab in the "Earthquake Module".

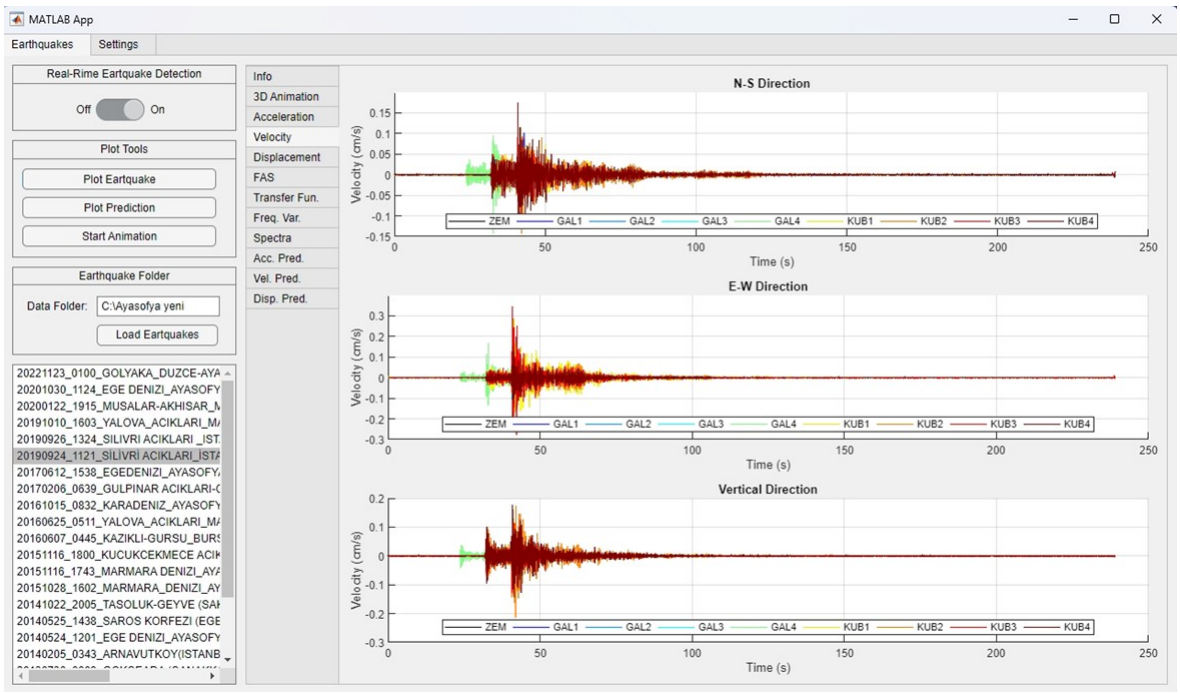

Figure 5.17. The screenshot of the "Velocity" tab in the "Earthquake Module".

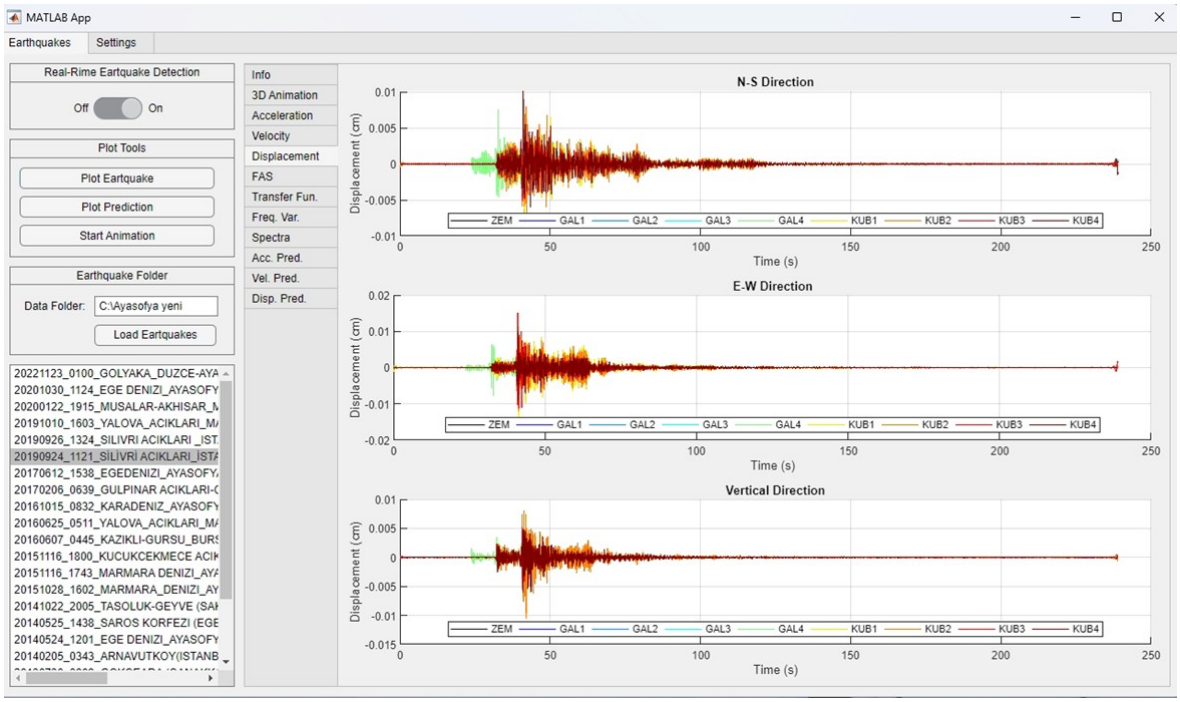

Figure 5.18. The screenshot of the "Displacement" tab in the "Earthquake Module".

The FAS and transfer function tabs show the frequency contents of the records The FAS tab presents the Fourier amplitude spectrum, which is a mathematical procedure that transforms a time-domain signal into its frequency-domain counterpart. The transfer function tab displays the spectrum free of ground effects, which is obtained by dividing the spectra of the records by the spectrum of the ground. The spectrum of the ground is the Fourier amplitude spectrum of the ground motion recorded by a reference station. The spectrum free of ground effects shows the amplification or attenuation of the records due to the structure and the soil. The screenshots of the FAS and Transfer Function tabs are presented in Figure 5.19 and Figure 5.20, respectively.

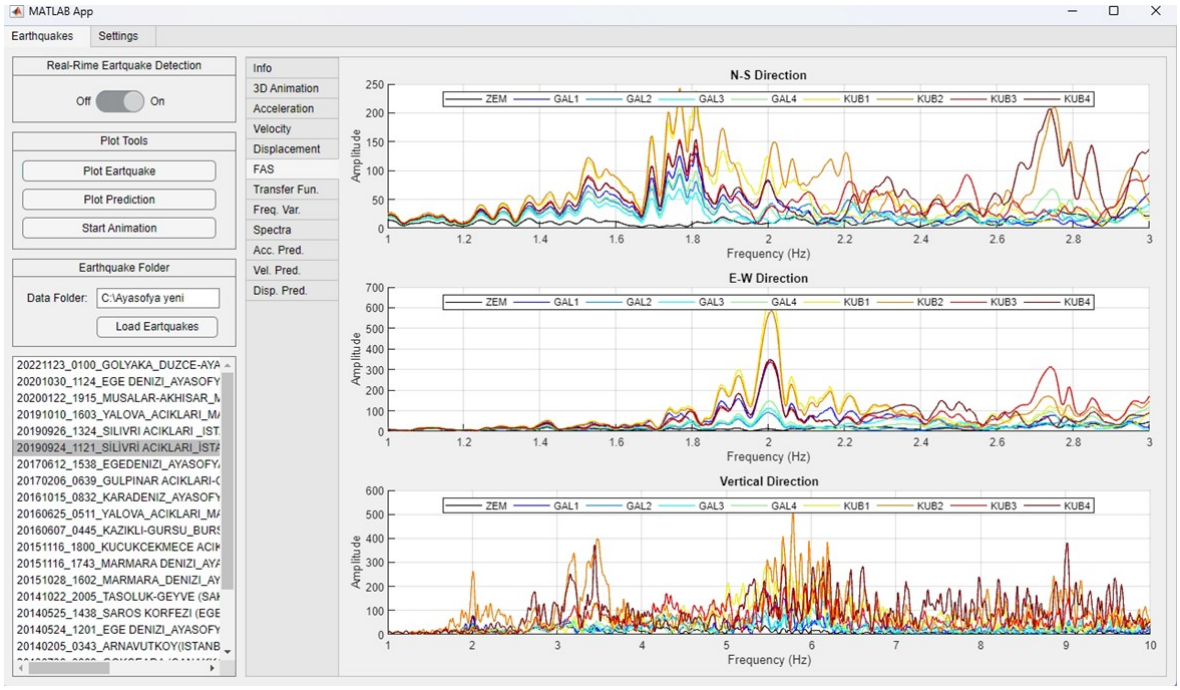

Figure 5.19. The screenshot of the "FAS" tab in the "Earthquake Module".

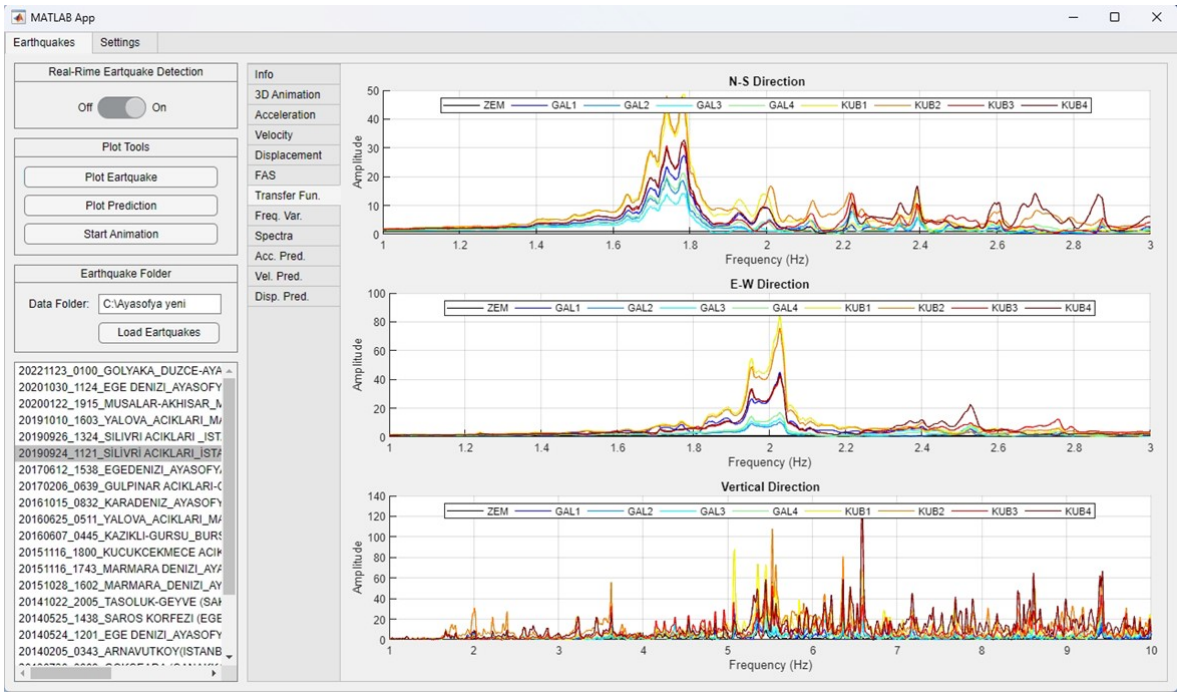

Figure 5.20. The screenshot of the "Transfer Function" tab in "Earthquake Module".

The frequency change tab shows the variation of the first two modal frequencies of the structure during the earthquake excitation. This frequency variation was obtained by the short-time Fourier transform (STFT) method [73], which is a technique to analyze the frequency content of a signal over time. The STFT method divides the signal into segments of equal length and applies a window function to each segment. The window function reduces the discontinuities at the edges of the segments and improves the frequency resolution. The Fourier transform of each windowed segment is then computed and plotted as a function of time and frequency. The result is a two-dimensional representation of the signal, also known as a spectrogram. The window width was selected as 512 data points, and the windows were calculated by overlapping 50% of the previous window. This means that each window contains 256 new data points and 256 data points from the previous window. The overlap increases the temporal resolution and reduces the aliasing effects of the windowing. The frequency change tab displays the spectrogram of the first two modal frequencies of the structure, which are the dominant frequencies that describe the dynamic behavior of the structure. The spectrogram shows how these frequencies change over time due to the earthquake excitation and the structural response. The light blue line represents the predicted frequency at the earthquake occurrence. This allows for a comparison between the predicted and the measured frequencies and an evaluation of the frequency recovery after the earthquake. The display of the frequency variation tab is shown in Figure 5.21.

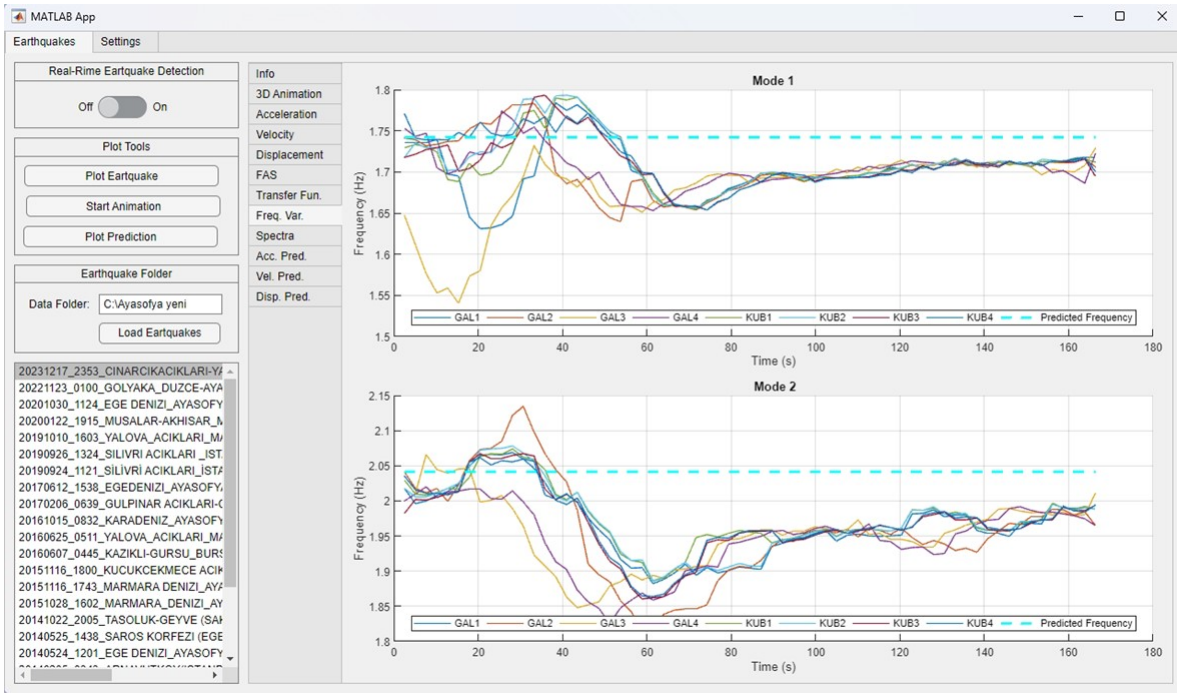

Figure 5.21. The screenshot of the "Frequency Variation" tab in the "Earthquake Module".

The Spectra tab shows the response spectrum obtained from the ground motion recorded at the ground station of the structure Ideally, a response spectrum should be calculated from a station located outside the building, because a station inside the building may involve structure-soil interaction, which can affect the accuracy of the spectrum. However, since the structure-soil interaction in Hagia Sophia is low [42], response spectra were calculated from the ground station to give an idea of the seismic response of the building. The screenshot of the Spectra tab is demonstrated in Figure 5.22.

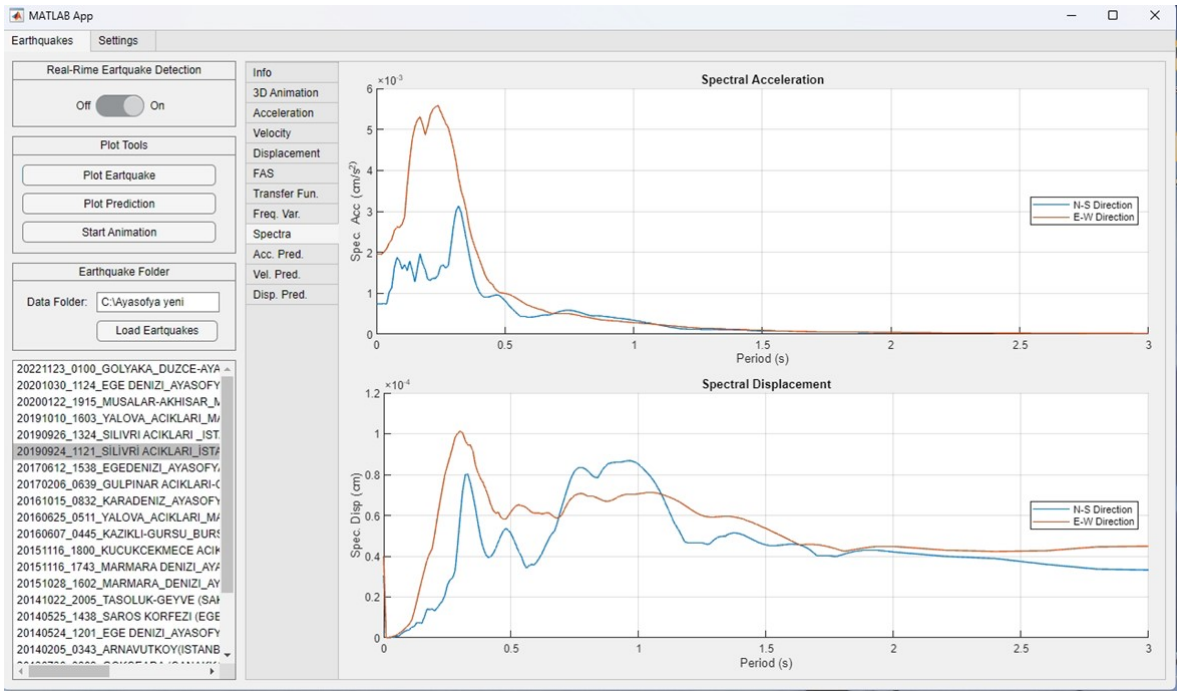

Figure 5.22. The screenshot of the "Spectra" tab in the "Earthquake Module".

The earthquake module has a tab for estimating the acceleration, velocity, and displacement of the structure. This tab is based on the assumption that there is a linear relationship between the peak ground acceleration, velocity, and displacement and the corresponding values measured at the stations on the structure. To test this assumption, 23 past earthquakes that affected the building were selected. These earthquakes were the largest ones that occurred when all the stations were operational and functioning properly. Table 5.2 shows the list of these earthquakes.

For each earthquake and each station, the peak acceleration, velocity, and displacement values were computed. Then, the linear correlation between the ground station inputs and the structural responses at the other stations was investigated. The results showed that there was a strong correlation between the ground inputs and the structural responses, especially for the displacement values. It is known that the structure did not suffer any damage during the selected earthquakes. Therefore, the linear

correlation is likely to persist unless the structure is damaged. If the measured values deviate significantly from the linear regression predictions, it may indicate a structural anomaly or damage. The displays of acceleration, velocity, and displacement prediction tabs are presented in Figure 5.23, Figure 5.24, and Figure 5.25.

Table 5.2. Selected earthquakes for acceleration, velocity, and displacement prediction.

No	*Date*	*Time*	*Latitude*	*Longitude*	*ML*	*Location*
1	14.03.2012	09:24:54	40.812	28.793	3.7	Marmara Sea
2	7.06.2012	20:54:25	40.854	27.924	5.1	Marmara Sea
3	19.10.2012	08:17:24	41.037	28.634	3.8	Esenyurt-Istanbul
4	8.01.2013	14:16:09	39.68	25.67	5.6	Aegean Sea
5	30.07.2013	05:33:07	40.303	25.79	5.3	Aegean Sea
6	5.02.2014	01:56:43	41.371	28.623	3.8	Arnavutköy-Istanbul
7	24.05.2014	09:25:01	40.304	25.28	6.5	Aegean Sea
8	25.05.2014	11:38:38	40.421	26.152	4.8	Saros Gulf
9	22.10.2014	17:11:05	40.407	30.115	4.5	Geyve-Sakarya
10	28.10.2015	16:20:02	40.822	27.764	4.5	Marmara Sea
11	16.11.2015	17:04:12	40.831	28.768	3.4	Marmara Sea
12	16.11.2015	15:45:43	40.832	28.754	4.2	Marmara Sea
13	7.06.2016	04:09:45	40.265	29.152	4.6	Gursu-Bursa
14	25.06.2016	05:40:11	40.707	29.212	4.5	Marmara Sea
15	15.10.2016	08:18:35	30.639	42.079	4.8	Black Sea
16	6.02.2017	03:51:39	39.545	26.109	5.4	Ayvacık-Çanakkale
17	12.06.2017	12:28:37	38.847	26.325	6.3	Aegean Sea
18	24.09.2019	08:00:21	40.875	28.212	4.7	Marmara Sea
19	26.09.2019	10:59:24	40.88	28.216	5.7	Marmara Sea
20	10.10.2019	16:52:03	40.695	29.26	4.1	Marmara Sea
21	22.01.2020	19:22:15	39.058	27.839	5.6	Akhisar-Manisa
22	30.10.2020	11:51:24	37.888	26.706	6.7	Aegean Sea
23	23.11.2022	01:08:14	40.831	30.998	6.1	Golyaka-Düzce

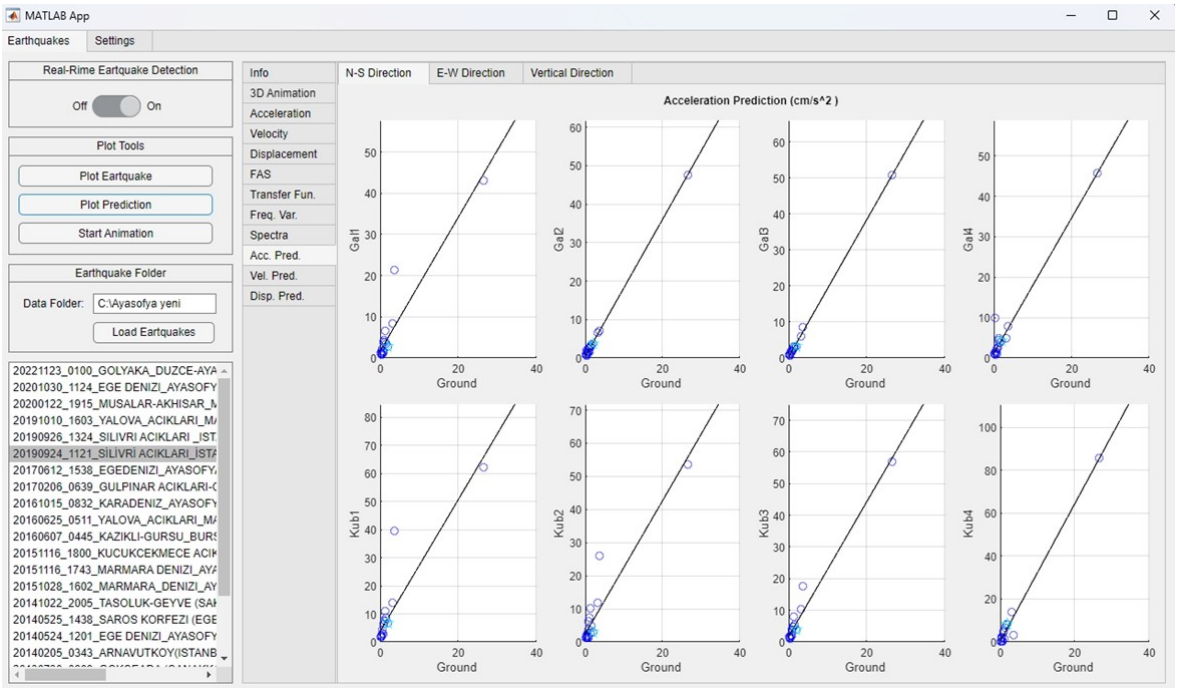

Figure 5.23. The screenshot of the "Acceleration Prediction" tab in the "Earthquake Module".

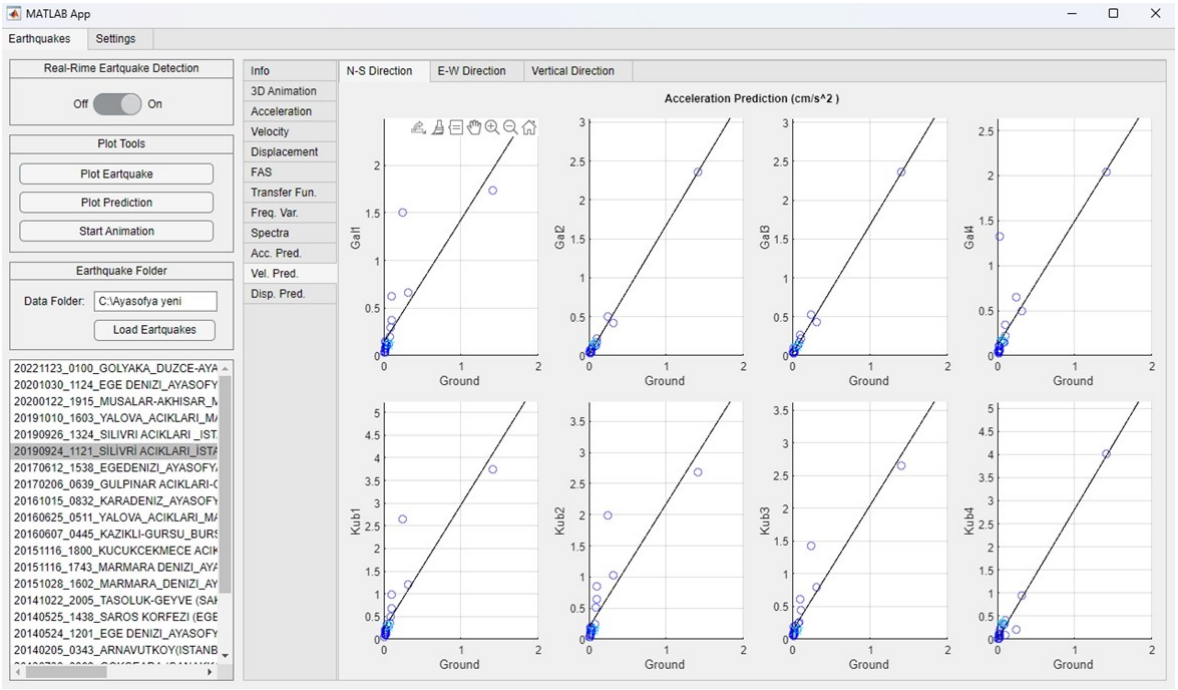

Figure 5.24. The screenshot of the "Velocity Prediction" tab in the Earthquake Module.

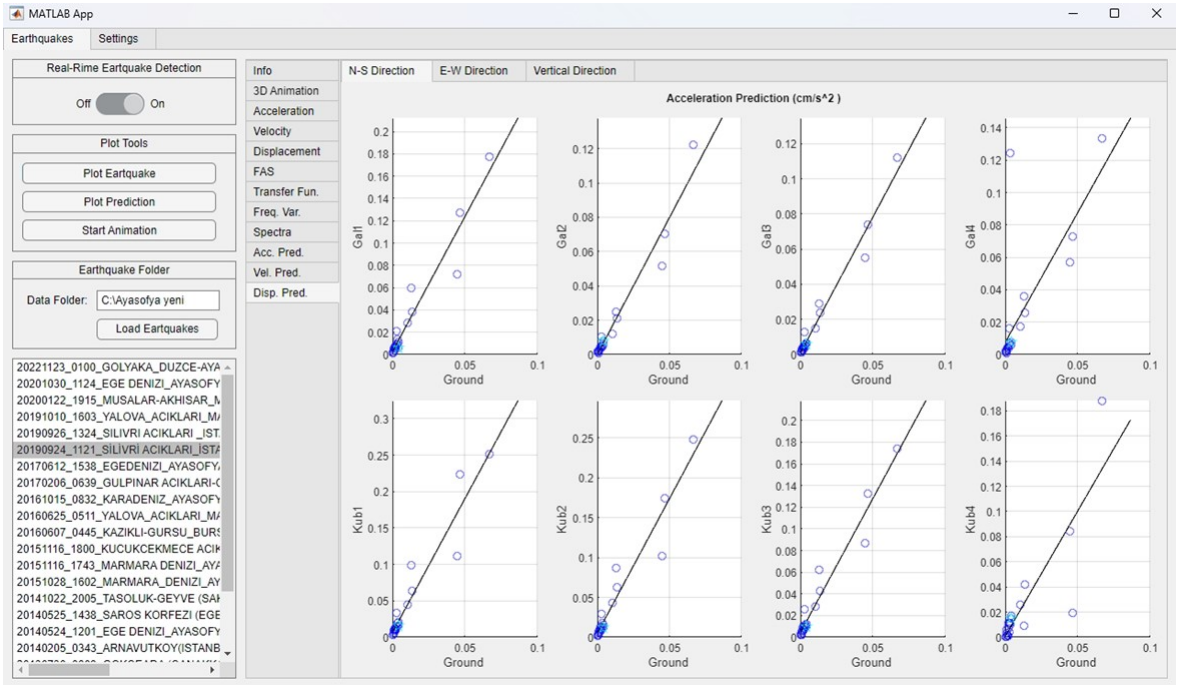

Figure 5.25. The screenshot of the "Displacement Prediction" tab in the "Earthquake Module".

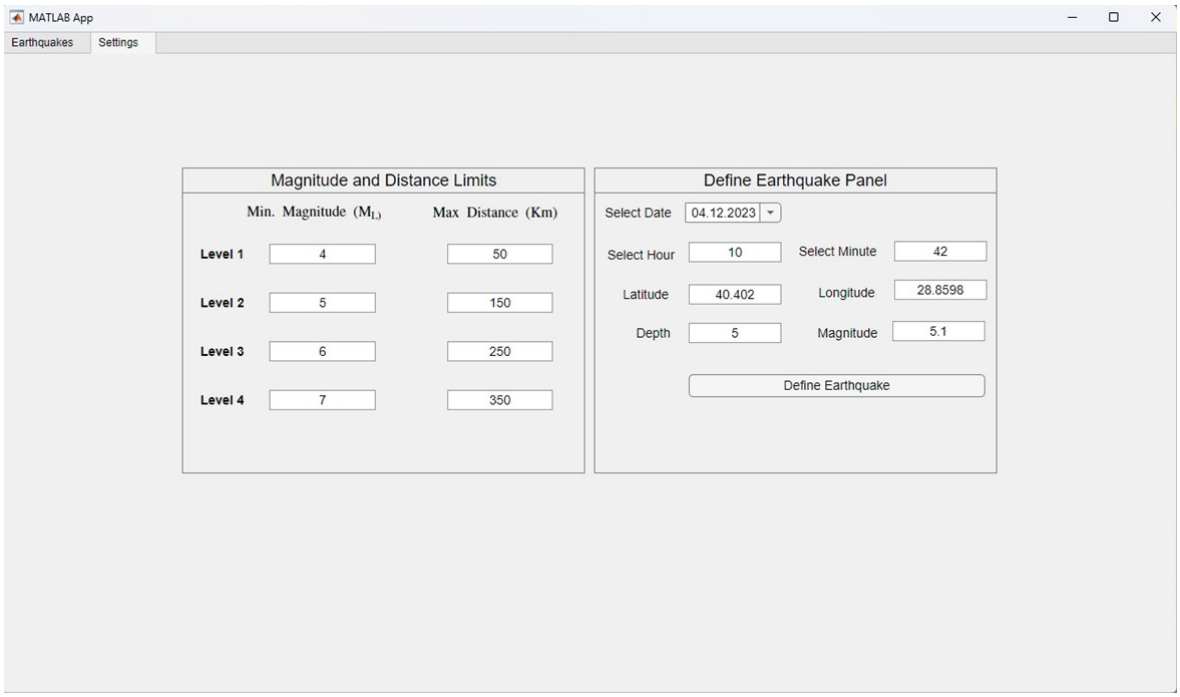

Figure 5.26. The screenshot of the "Setting" tab in the "Earthquake Module".

6. CONCLUSION

The main objective of this thesis was to develop and validate a real-time structural health monitoring (SHM) system for Hagia Sophia using statistical models and machine learning algorithms. The thesis addressed several research questions related to the long-term variation of modal parameters, the relationship between modal parameters and environmental factors, the estimation of modal frequency based on atmospheric data, the significance of mode shape variations, and the performance of the developed algorithm and software in detecting structural anomalies and earthquakes. The main findings and contributions of this thesis are summarized as follows:

- Long-term variation of modal parameters: The thesis analyzed the four-year continuous acceleration data of Hagia Sophia and obtained the modal frequencies, modal damping ratios, and mode shapes of the structure. The results showed that the modal parameters varied significantly over time and were influenced by environmental factors such as temperature, humidity, and wind speed. The thesis also found that the modal frequencies had a linear relationship with temperature and a nonlinear relationship with humidity and wind speed.

- Modal frequency prediction: The thesis developed a statistical model to predict the modal frequency of Hagia Sophia based on atmospheric data. The thesis compared four regression methods: multiple linear regression, regression trees, support vector machines, and neural networks. The results showed that the neural network method had the lowest error rate and the best performance among the methods. The thesis also found that temperature was the most important predictor variable, followed by humidity and wind speed. The thesis also demonstrated that the predicted modal frequencies could be used to detect structural anomalies in real-time by comparing them with the measured modal frequencies.

- Mode shape variation and analysis: The thesis investigated the variation of mode

shapes of Hagia Sophia over time and their significance for structural health monitoring. The thesis applied the modal assurance criterion (MAC), coordinate modal assurance criterion (COMAC), and enhanced coordinate modal assurance criterion (ECOMAC) methods to analyze the mode shape changes. The results showed that the mode shapes varied significantly over time and were affected by environmental factors and sensor malfunctions. The thesis also found that the MAC, COMAC, and ECOMAC values could be used to detect structural anomalies and damage localization by comparing them with a reference mode shape.

- Real-time structural health monitoring algorithm and software: The thesis developed and validated a real-time SHM algorithm and software for Hagia Sophia using statistical models and machine learning algorithms. The algorithm and software, called AISHM, performed modal analysis and estimation, mode shape analysis, earthquake detection and analysis, and structural anomaly detection in real-time. The algorithm and software also provided a user interface and a 3-D animation of the structure. The results showed that the algorithm and software performed well in detecting and analyzing earthquakes and structural anomalies and provided useful information and visualization for engineers and decision-makers.

- The thesis made several original and significant contributions to the field of SHM and the preservation of cultural heritage. The thesis proposed a novel method for real-time modal frequency estimation using atmospheric data and machine learning algorithms, which can improve the accuracy and reliability of SHM systems. The thesis also developed and validated a real-time SHM algorithm and software for Hagia Sophia, which is one of the most important and complex historical structures in the world. The thesis also provided a comprehensive and detailed analysis of the long-term variation of modal parameters and mode shapes of Hagia Sophia, which can enhance the understanding of the dynamic behavior and structural performance of the structure. The thesis also demonstrated the effectiveness of the developed algorithm and software in detecting and analyzing

earthquakes and structural anomalies, which can facilitate the timely intervention and prevention of further damage.

- The thesis demonstrates that historical buildings are more susceptible to atmospheric effects than modern-day, conventional structures. However, the degree of susceptibility caused by atmospheric factors varies among different historical monuments. An investigation of the degree of impact of environmental factors on the dynamic response of a historical structure should be assessed as part of studies on its dynamic response. A case study of the Hagia Sophia Mosque, one of the oldest and most iconic historical monuments in the world, reveals that its dynamic structural response parameters are affected by atmospheric factors. Structural research on this edifice should take this into consideration. Measures must be taken to protect this important historical monument from environmental effects.

The thesis also identified some limitations and challenges of the proposed method and the developed algorithm and software. The main limitations and challenges are:

- Data quality and availability: The quality and availability of the acceleration and atmospheric data are crucial for the accuracy and reliability of the SHM system. However, the data may be corrupted, missing, or delayed due to internet or power outages, sensor malfunctions, or human errors. Therefore, the SHM system should have mechanisms to detect and handle data quality and availability issues, such as data validation, data correction, data interpolation, data backup, and data synchronization.

- Model selection and tuning: The selection and tuning of the statistical models and machine learning algorithms are important for the performance and robustness of the SHM system. However, the models and algorithms may have different assumptions, advantages, and disadvantages, and may require different parameters, inputs, and outputs. Therefore, the SHM system should have criteria to select

and tune the models and algorithms appropriately, such as error rate, accuracy, precision, recall, sensitivity, specificity, and computational cost.

The thesis also suggested some directions and recommendations for future research and development. The main directions and recommendations are:

- Data fusion and integration: The integration and fusion of diverse types of data, such as acceleration, displacement, strain, temperature, humidity, wind speed, and precipitation, can improve the accuracy and reliability of the SHM system. Data fusion and integration can provide more information and insights about the structure and its environment and can reduce the uncertainty and noise in the data. Data fusion and integration can be achieved by using different techniques, such as statistical methods, machine learning methods, or sensor networks.

- Model comparison and evaluation: The comparison and evaluation of different statistical models and machine learning algorithms can improve the performance and robustness of the SHM system. Model comparison and evaluation can help to select the best model or algorithm for a given problem, and to optimize the parameters and inputs of the model or algorithm. Model comparison and evaluation can be done by using different metrics, such as error rate, accuracy, precision, recall, sensitivity, specificity, and computational cost.

- Damage characterization and quantification: The characterization and quantification of structural damage can improve the effectiveness and efficiency of the SHM system. Damage characterization and quantification can help to describe and measure the severity, extent, and location of the damage, and to assess the impact and risk of the damage on the structural performance and safety. Damage characterization and quantification can be done by using different methods, such as damage indicators, damage indices, damage features, and damage patterns.

- Software development and improvement: The development and improvement of

the SHM software can improve the usability and functionality of the SHM system. Software development and improvement can help to create and update the user interface, the 3-D animation, the data visualization, the data storage, and the data communication of the SHM system. Software development and improvement can be done by using different tools, such as Matlab, Python, Java, or C++.

In conclusion, this thesis developed and validated a real-time SHM system for Hagia Sophia using statistical models and machine learning algorithms. The thesis addressed several research questions related to the long-term variation of modal parameters, the relationship between modal parameters and environmental factors, the estimation of modal frequency based on atmospheric data, the significance of mode shape variations, and the performance of the developed algorithm and software in detecting structural anomalies and earthquakes. The thesis made several original and significant contributions to the field of SHM and the preservation of cultural heritage. The thesis also identified some limitations and challenges of the proposed method and the developed algorithm and software. The thesis also suggested some directions and recommendations for future research and development. The thesis hopes that the findings and contributions of this study can inspire and benefit other researchers and practitioners who are interested in SHM and the preservation of cultural heritage.

REFERENCES

1. Azzara, R. M., G. D. Roeck, M. Girardi, C. Padovani, D. Pellegrini and E. Reynders, "The influence of environmental parameters on the dynamic behaviour of the San Frediano bell tower in Lucca", *Engineering Structures*, Vol. 156, pp. 175 – 187, 2018.

2. Cavalagli, N., G. Comanducci and F. Ubertini, "Earthquake-Induced Damage Detection in a Monumental Masonry Bell-Tower Using Long-Term Dynamic Monitoring Data", *Journal of Earthquake Engineering*, Vol. 22, No. sup1, 2018, publisher: Taylor & Francis pages = 96–119,.

3. Ceravolo, R., A. De Marinis, M. L. Pecorelli and L. Zanotti Fragonara, "Monitoring of masonry historical constructions: 10 years of static monitoring of the world's largest oval dome", *Structural Control and Health Monitoring*, Vol. 24, No. 10, p. e1988, 2017.

4. Lorenzoni, F., M. Caldon, F. da Porto, C. Modena and T. Aoki, "Post-earthquake controls and damage detection through structural health monitoring: applications in l'Aquila", *Journal of Civil Structural Health Monitoring*, Vol. 8, No. 2, pp. 217–236, Apr. 2018.

5. Mesquita, E., A. Arêde, N. Pinto, P. Antunes and H. Varum, "Long-term monitoring of a damaged historic structure using a wireless sensor network", *Engineering Structures*, Vol. 161, Feb. 2018.

6. Saisi, A., C. Gentile and A. Ruccolo, "Continuous monitoring of a challenging heritage tower in Monza, Italy", *Journal of Civil Structural Health Monitoring*, Vol. 8, Nov. 2017.

7. Ubertini, F., G. Comanducci, N. Cavalagli, A. L. Pisello, A. L. Materazzi and

F. Cotana, "Environmental effects on natural frequencies of the San Pietro bell tower in Perugia, Italy, and their removal for structural performance assessment", *Mechanical Systems and Signal Processing*, Vol. 82, pp. 307 – 322, 2017.

8. Ubertini, F., N. Cavalagli, A. Kita and G. Comanducci, "Assessment of a monumental masonry bell-tower after 2016 Central Italy seismic sequence by long-term SHM", *Bulletin of Earthquake Engineering*, Vol. 16, No. 2, pp. 775–801, Feb. 2018.

9. Blanco, H., Y. Boffill, I. Lombillo and L. Villegas, "An integrated structural health monitoring system for determining local/global responses of historic masonry buildings", *Structural Control and Health Monitoring*, Vol. 25, No. 8, p. e2196, 2018.

10. Cabboi, A., C. Gentile and A. Saisi, "From continuous vibration monitoring to FEM-based damage assessment: Application on a stone-masonry tower", *Construction and Building Materials*, Vol. 156, pp. 252 – 265, 2017.

11. Gentile, C. and A. Saisi, "Operational modal testing of historic structures at different levels of excitation", *Construction and Building Materials*, Vol. 48, pp. 1273 – 1285, 2013.

12. Tsogka, C., E. Daskalakis, G. Comanducci and F. Ubertini, "The Stretching Method for Vibration-Based Structural Health Monitoring of Civil Structures", *Computer-Aided Civil and Infrastructure Engineering*, Vol. 32, No. 4, pp. 288–303, 2017.

13. Ubertini, F., G. Comanducci and N. Cavalagli, "Vibration-based structural health monitoring of a historic bell-tower using output-only measurements and multivariate statistical analysis", *Structural Health Monitoring*, Vol. 15, No. 4, pp. 438–457, 2016.

14. Deraemaeker, A., E. Reynders, G. De Roeck and J. Kullaa, "Vibration-based structural health monitoring using output-only measurements under changing environ-

ment", *Mechanical systems and signal processing*, Vol. 22, No. 1, pp. 34–56, 2008.

15. Plachỳ, J., J. Musílek, L. Podolka and M. Karková, "Disorders of the Building and its Remediation-Hagia Sophia, Turkey the Most the Byzantine Building", *Procedia engineering*, Vol. 161, pp. 2259–2264, 2016.

16. Çaktı, E., E. Dar and G. Uncu, "Recent Studies on Earthquake Performance Assessment of Hagia Sophia in Istanbul: Recent Developments", *Seismic Isolation, Structural Health Monitoring, and Performance Based Seismic Design in Earthquake Engineering*, pp. 195–204, Springer, 2019.

17. Dar, E. and E. Çaktı, "Dependence of Hagia Sophia's Modal Damping Ratios on Atmospheric Effects", *International Symposium on Structural Health Monitoring and Nondestructive Testing*, Saarbrücken, Germany, Oct. 2018.

18. Dar, E. and E. Çaktı, "Long-Term Mode Shape Variations of Hagia Sophia with Environmental Factors", *5th International Conference on Smart Monitoring, Assessment and Rehabilitation of Civil Structures*, Potsdam,Germany, Aug. 2019.

19. Elyamani, A., O. Caselles, P. Roca and J. Clapes, "Dynamic investigation of a large historical cathedral", *Structural Control and Health Monitoring*, Vol. 24, No. 3, p. e1885, 2017.

20. Ramos, L. F., L. Marques, P. B. Lourenço, G. D. Roeck, A. Campos-Costa and J. Roque, "Monitoring historical masonry structures with operational modal analysis: Two case studies", *Mechanical Systems and Signal Processing*, Vol. 24, No. 5, pp. 1291 – 1305, 2010.

21. Zonno, G., R. Aguilar, R. Boroschek and P. B. Lourenço, "Analysis of the long and short-term effects of temperature and humidity on the structural properties of adobe buildings using continuous monitoring", *Engineering Structures*, Vol. 196, pp. 109–299, 2019.

22. Li, H., S. Li, J. Ou and H. Li, "Modal identification of bridges under varying environmental conditions: temperature and wind effects", *Structural Control and Health Monitoring*, Vol. 17, No. 5, pp. 495–512, 2010.

23. Çaktı, E., S. Cimilli and M. Erdik, "Dynamic Response of Two Historical Monuments in Istanbul Deduced from the Recordings of Kocaeli and Düzce Earthquakes", *Bulletin of the Seismological Society of America*, Vol. 93, pp. 694–712, May 2003.

24. Çakmak, A., R. Davidson, C. Mullen and M. Erdik, "Dynamic analysis and earthquake response of Hagia Sophia", *WIT Transactions on The Built Environment*, Vol. 3, 1993.

25. Masciotta, M.-G., J. C. Roque, L. F. Ramos and P. B. Lourenço, "A multidisciplinary approach to assess the health state of heritage structures: The case study of the Church of Monastery of Jerónimos in Lisbon", *Construction and Building Materials*, Vol. 116, pp. 169–187, 2016.

26. Saisi, A., C. Gentile and M. Guidobaldi, "Post-earthquake continuous dynamic monitoring of the Gabbia Tower in Mantua, Italy", *Construction and Building Materials*, Vol. 81, Apr. 2015.

27. Moropoulou, A., B. Christaras, G. Lavas, G. Penelis, N. Zias, G. Biscontin, E. Kollias, A. Paisios, P. Theoulakis, K. Bisbikou and others, "Weathering phenomena on the Hagia Sophia Basilica Konstantinople", *WIT Transactions on The Built Environment*, Vol. 4, 1993.

28. Demirören news agency, "Milliyet", https://www.milliyet.com.tr/gundem/ayasofya-aciliyor-mu-ayasofya-ne-zaman-muze-oldu-iste-onemi-ve-tarihi-6222160, 2024, accessed: 2024-12-15.

29. Çakmak, A. S., R. M. Taylor and E. Durukal, "The structural configuration of

the first dome of Justinian's Hagia Sophia (AD 537–558): An investigation based on structural and literary analysis", *Soil dynamics and earthquake engineering*, Vol. 29, No. 4, pp. 693–698, 2009.

30. Mark, R., A. Çakmak, K. Hill and R. Davidson, "Structural analysis of Hagia Sophia: a historical perspective", *WIT Transactions on The Built Environment*, Vol. 4, 1993.

31. Erdik, M. and E. Durukal, "Use of strong motion data for the assessment of the earthquake response of historical monuments", *Proc. 11th World Conference on Earthquake Engineering*, pp. 23–28, 1996.

32. Davidson, R., "The mother of all churches: a static and dynamic structural analysis of Hagia Sophia", *Unpublished senior thesis, civil engineering and operations research department, Princeton University*, 1993.

33. Kırlangıç, A. S., "Re-evaluation of earthquake performance and strengthening alternatives of Hagia Sophia", *Graduate Program in Earthquake Engineering, Bogazici University*, 2008.

34. Mainstone, R., "The structural conservation of Hagia Sophia", *WIT Transactions on The Built Environment*, Vol. 4, 1993.

35. Livingston, R., "Materials analysis of the masonry of the Hagia Sophia Basilica, Istanbul", *WIT Transactions on The Built Environment*, Vol. 4, 1993.

36. Cappa, M., D. De Angelis, A. Pecci, L. Barba, M. Cura, G. M. Crisci, J. Blancas, H. B. Yavuz and D. Miriello, "Thermographic survey at Hagia Sophia: main arches, pendentives and tympana", *International Journal of Architectural Heritage*, Vol. 10, No. 6, pp. 726–734, 2016.

37. Erdik, M., E. Durukal, Yuzugullu, K. Beyen and U. Kadakal, "Strong-motion instrumentation of Aya Sofya and the analysis of response to an earthquake of 4.8

magnitude", *WIT Transactions on The Built Environment*, Vol. 3, 1993.

38. Almac, U., K. Schweizerhof, G. Blankenhorn, C. Duppel and F. Wenzel, "Structural behaviour of Hagia Sophia under dynamic loads", *Vienna congress on recent advances in earthquake engineering and structural dynamics, Vienna, Austria*, Vol. 475, 2013.

39. Aoki, T., S. Kato, K. Ishikawa, K. Hidaka and M. Yorulmaz, "Principle of structural restoration for Hagia Sophia dome", *WIT Transactions on The Built Environment*, Vol. 26, p. 10, 1997.

40. Özkul, T. A. and E. Kuribayashi, "Structural characteristics of Hagia Sophia: II—A finite element formulation for dynamic analysis", *Building and environment*, Vol. 42, No. 5, pp. 2100–2106, 2007.

41. Aoki, T., S. Kato and K. Ishikawa, "Structural characteristics of the dome of Hagia Sophia from measurement of micro tremor", *WIT Transactions on The Built Environment*, Vol. 4, 1993.

42. Çakmak, A., M. Natsis and C. Mullen, "Foundation effect on the dynamics of Hagia Sophia", *WIT Transactions on The Built Environment*, Vol. 17, 1995.

43. Natsis, M., "The enigma of Hagia Sophia: a dynamic structural analysis of Justinian's great church", *unpublished senior thesis, Dept. of Civil Engineering and Operations Research, Princeton University*, 1994.

44. Sahin, M. and I. Mungan, "Dynamic performance of the roof of Hagia Sophia considering cracking", *International Journal of Space Structures*, Vol. 20, No. 3, pp. 135–141, 2005.

45. Muñoz, A., R. Ertlé and M. Unser, "Continuous wavelet transform with arbitrary scales and O(N) complexity", *Signal Processing*, Vol. 82, No. 5, pp. 749–757, 2002.

46. Mitra, S. K., *Digital signal processing: a computer-based approach*, McGraw-Hill series in electrical and computer engineering, McGraw-Hill/Irwin, Boston, Mass., 2nd ed edn., 2001.

47. Dar, E., *Long-Term Dynamic Response of Hagia Sophia in Istanbul To Earthquakes and Atmospheric Conditions*, Master's Thesis, Bogazici University Kandilli Observatory and Earthquake Research Institute, 2015.

48. Durukal, E., *A Study on Structural Identification and Seismic Vulnerability Assessment Of Aya Sofya*, Master's Thesis, Bogazici University Kandilli Observatory and Earthquake Research Institute, 1992.

49. Chopra, A., *Dynamics of Structures*, Prentice-Hall international series in civil engineering and engineering mechanics, Pearson Education, 2007.

50. Brincker, R., L. Zhang and P. Andersen, "Modal identification of output-only systems using frequency domain decomposition", *Smart Materials and Structures*, Vol. 10, No. 3, pp. 441–445, Jun. 2001.

51. Pastor, M., M. Binda and T. Harčarik, "Modal Assurance Criterion", *Procedia Engineering*, Vol. 48, pp. 543–548, 2012.

52. NAJ Lieven, D. E., "Spatial correlation of mode shapes: the coordinate modal assurance criterion (COMAC)", *Proceedings of the 6th international modal analysis conference*, Florida, USA, 1988.

53. Hunt, D. L., "Application of an enhanced Coordinate Modal Assurance Criterion", *10th International Modal Analysis Conference*, Vol. 1, San Diego, USA, Jan. 1992.

54. Cao, M. S., G. G. Sha, Y. F. Gao and W. Ostachowicz, "Structural damage identification using damping: a compendium of uses and features", *Smart Materials and Structures*, Vol. 26, No. 4, p. 043001, Apr. 2017.

55. MATLAB, *MATLAB version: R2023b*, The MathWorks Inc., Natick, Massachusetts, United States, 2023, `https://www.mathworks.com`.

56. Sohil, F., M. U. Sohali and J. Shabbir, "An introduction to statistical learning with applications in R", *Statistical Theory and Related Fields*, Vol. 6, pp. 87–87, Jan. 2022.

57. Buja, A., P. Tukey and U. of Minnesota (Editors), *Computing and graphics in statistics*, No. v. 36 in IMA volumes in mathematics and its applications, Springer-Verlag, New York, 1991.

58. Holland, P. W. and R. E. Welsch, "Robust regression using iteratively reweighted least-squares", *Communications in Statistics - Theory and Methods*, Vol. 6, No. 9, pp. 813–827, Jan. 1977.

59. Huber, P. J., *Robust Statistics*, Wiley Series in Probability and Statistics, Wiley, 1 edn., Feb. 1981.

60. Street, J. O., R. J. Carroll and D. Ruppert, "A Note on Computing Robust Regression Estimates Via Iteratively Reweighted Least Squares", *The American Statistician*, Vol. 42, No. 2, p. 152, May 1988.

61. James, G., D. Witten, T. Hastie and R. Tibshirani, *An Introduction to Statistical Learning*, Vol. 103 of *Springer Texts in Statistics*, Springer New York, New York, NY, 2013.

62. Frank E. Harrell, J., *Regression Modeling Strategies*, Springer Series in Statistics, Springer International Publishing, Cham, 2015.

63. Torgo, L., "Regression Trees", C. Sammut and G. I. Webb (Editors), *Encyclopedia of Machine Learning and Data Mining*, pp. 1080–1083, Springer US, Boston, MA, 2017.

64. Breiman, L., *Classification and regression trees*, Routledge, Abingdon, 1984.

65. Hastie, T., R. Tibshirani and J. Friedman, *The Elements of Statistical Learning*, Springer Series in Statistics, Springer New York, New York, NY, 2009.

66. Montesinos López, O. A., A. Montesinos López and J. Crossa, *Multivariate statistical machine learning methods for genomic prediction*, Springer, Cham, Switzerland, 2022.

67. Adankon, M. M. and M. Cheriet, "Support Vector Machine", S. Z. Li and A. Jain (Editors), *Encyclopedia of Biometrics*, pp. 1303–1308, Springer US, Boston, MA, 2009, 10.1007/978-0-387-73003-5$_2$99.

68. Awad, M. and R. Khanna, *Efficient learning machines: theories, concepts, and applications for engineers and system designers*, The expert's voice in machine learning, Apress Open, Berkley?, 2015.

69. Cortes, C. and V. Vapnik, "Support-Vector Networks", *Machine Learning*, Vol. 20, No. 3, pp. 273–297, 1995.

70. Nielsen, M., *Neural Networks and Deep Learning*, Determination Press, 2015.

71. Goodfellow, I., Y. Bengio and A. Courville, *Deep learning*, Adaptive computation and machine learning, The MIT press, Cambridge, Mass, 2016.

72. Bogazici University Rectorship, "Regional Earthquake-Tsunami Monitoring Center", Last accessed on 2024-01-12, 2024, http://www.koeri.boun.edu.tr/scripts/lasteq.asp.

73. Oppenheim, A. V., R. W. Schafer and J. R. Buck, *Discrete-time signal processing*, Prentice Hall, Upper Saddle River, N.J, 2nd ed edn., 1999.

74. Çaktı, E. and E. Dar, "Long-term dynamic response of Hagia Sophia in Istanbul

to earthquakes and atmospheric conditions", *Structural Health Monitoring 2015*, Stanford, USA, 2015.

75. Ditommaso, R., M. Mucciarelli, S. Parolai and M. Picozzi, "Monitoring the structural dynamic response of a masonry tower: comparing classical and time-frequency analyses", *Bulletin of Earthquake Engineering*, Vol. 10, No. 4, pp. 1221–1235, Aug. 2012.

76. Gentile, C., A. Saisi and A. Cabboi, "Structural Identification of a Masonry Tower Based on Operational Modal Analysis", *International Journal of Architectural Heritage*, Vol. 9, No. 2, pp. 98–110, 2015.

77. Foti, D., M. Diaferio, N. I. Giannoccaro and M. Mongelli, "Ambient vibration testing, dynamic identification and model updating of a historic tower", *NDT & E International*, Vol. 47, pp. 88 – 95, 2012.

78. Aras, F., L. Krstevska, G. Altay and L. Tashkov, "Experimental and numerical modal analyses of a historical masonry palace", *Construction and Building Materials*, Vol. 25, No. 1, pp. 81 – 91, 2011.

79. Gentile, C., A. Saisi and A. Cabboi, "Dynamic monitoring of a masonry tower", *Proc. 8th Int. Conf. on Structural Analysis of Historical Constructions*, Vol. 24, Oct. 2012.

80. Peña, F., P. B. Lourenço, N. Mendes and D. V. Oliveira, "Numerical models for the seismic assessment of an old masonry tower", *Engineering Structures*, Vol. 32, No. 5, pp. 1466 – 1478, 2010.

81. Casarin, F. and C. Modena, "Seismic Assessment of Complex Historical Buildings: Application to Reggio Emilia Cathedral, Italy", *International Journal of Architectural Heritage*, Vol. 2, No. 3, pp. 304–327, 2008.

82. Ivorra, S. and F. J. Pallarés, "Dynamic investigations on a masonry bell tower",

Engineering Structures, Vol. 28, No. 5, pp. 660 – 667, 2006.

83. Pau, A. and F. Vestroni, "Vibration analysis and dynamic characterization of the Colosseum", *Structural Control and Health Monitoring*, Vol. 15, No. 8, pp. 1105–1121, 2008.

84. Gentile, C. and A. Saisi, "Ambient vibration testing of historic masonry towers for structural identification and damage assessment", *Construction and Building Materials*, Vol. 21, No. 6, pp. 1311 – 1321, 2007.

85. Jaishi, B., W.-X. Ren, Z.-H. Zong and P. N. Maskey, "Dynamic and seismic performance of old multi-tiered temples in Nepal", *Engineering Structures*, Vol. 25, No. 14, pp. 1827 – 1839, 2003.

86. Çakmak, A. , A. Moropoulou and C. L. Mullen, "Interdisciplinary study of dynamic behavior and earthquake response of Hagia Sophia", *Soil Dynamics and Earthquake Engineering*, Vol. 14, No. 2, pp. 125 – 133, 1995.

87. Altunışık, A., F. Okur, A. Genç, M. Günaydin and S. Adanur, "Automated Model Updating of Historical Masonry Structures Based on Ambient Vibration Measurements", *Journal of Performance of Constructed Facilities*, Vol. 32, Nov. 2017.

88. Bennati, S., L. Nardini and W. Salvatore, "Dynamic Behavior of a Medieval Masonry Bell Tower. II: Measurement and Modeling of the Tower Motion", *Journal of Structural Engineering-asce - J STRUCT ENG-ASCE*, Vol. 131, Nov. 2005.

89. Hill, K., "Hagia Sophia: static analysis of a finite element model", *Unpublished Senior thesis, Deportment of civil engineering and operations research, Princeton University*, 1991.

90. Kaya, Y. and E. Safak, "Real-time structural health monitoring and damage detection", *Topics in Dynamics of Civil Structures, Volume 4*, pp. 11–19, Springer, 2013.

91. Peeters, B. and G. De Roeck, "One-year monitoring of the Z24-Bridge: environmental effects versus damage events", *Earthquake engineering & structural dynamics*, Vol. 30, No. 2, pp. 149–171, 2001.

92. Figueiredo, E., G. Park, C. R. Farrar, K. Worden and J. Figueiras, "Machine learning algorithms for damage detection under operational and environmental variability", *Structural Health Monitoring*, Vol. 10, No. 6, pp. 559–572, 2011.

93. Reynders, E., G. Wursten and G. De Roeck, "Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification", *Structural Health Monitoring*, Vol. 13, No. 1, pp. 82–93, 2014.

94. Lorenzoni, F., F. Casarin, M. Caldon, K. Islami and C. Modena, "Uncertainty quantification in structural health monitoring: Applications on cultural heritage buildings", *Mechanical Systems and Signal Processing*, Vol. 66, pp. 268–281, 2016.

95. Sammut, C. and G. I. Webb (Editors), *Encyclopedia of machine learning and data mining*, Springer reference, Springer, New York, NY, second edition edn., 2017.