

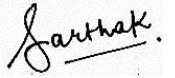
INDIAN INSTITUTE OF TECHNOLOGY DELHI
DEPARTMENT OF CIVIL ENGINEERING

CERTIFICATE PAGE

REPORT OF THE PROJECT SHORTLISTED UNDER SURA 2021

Project titled: **The Application of Machine Learning to Structural Health Monitoring**

Submitted by:

Name of the Students	Dept. / Centre	Entry no.	Contact no. & Email address	Signature of the Students
Sarthak Shrivastava	Civil Engineering	2019CE10297	+919039660358 ce1190297@iitd.ac.in	

Name and signatures of the Facilitators: Tawwaf 28/10/21 Sahil Bansal

His / Her Deptt. / Centre: Civil Engg. Deptt.

Mobile no. of the Facilitators: 2654-8414

Sahil@iitd.ac.in

Date of submission of report in IRD Unit: 29/10/2021



INDIAN INSTITUTE OF TECHNOLOGY DELHI
DEPARTMENT OF CIVIL ENGINEERING

Summer Undergraduate Research Award (SURA) 2021 Project Report

The Application of Machine Learning to Structural Health Monitoring

Sarthak Shrivastava

Entry no.: 2019CE10297

Email: ce1190297@iitd.ac.in

Supervisor:

Dr. Sahil Bansal

Assistant Professor

Department of Civil Engineering

Indian Institute of Technology, Delhi

Email: sahil@iitd.ac.in

ACKNOWLEDGEMENT

In the successful accomplishment of my project on The Application of Machine Learning to Structural Health Monitoring, I would like to thank and convey my special gratitude to Dr. Sahil Bansal, Department of Civil Engineering, IIT Delhi. Your valuable guidance and suggestions helped me in various phases of the completion of this project. I will always be thankful to you in this regard.

I would also like to thank Industrial Research & Development (IRD) Unit, IIT Delhi for giving me this opportunity to work on this project under Summer Undergraduate Research Award (SURA) 2021. The project has helped me in enhancing research skills and develop a better understanding of the subject.

Sarthak Shrivastava

Entry no.: 2019CE10297

Department of Civil Engineering

Indian Institute of Technology Delhi

Dated: October 27, 2021

TABLE OF CONTENTS

1. INTRODUCTION	1
1.1. Introduction to Machine Learning	1
1.1.1. Data-driven modelling	1
1.1.2. Data-driven methods: Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL)	1
1.1.3. Overview of Machine Learning	2
1.1.4. Types of Machine Learning techniques.....	2
1.2. Structural Health Monitoring.....	3
1.2.1. Approaches to Structural Health Monitoring	4
1.2.2. Data-driven Damage Identification	4
1.2.3. Vibration-based Damage Detection.....	5
1.3. Objective of the Report.....	6
2. MULTI-DEGREE OF FREEDOM SYSTEM.....	7
2.1. Overview of Multi-Degree of Freedom (MDOF) System	7
2.1.1. Dynamic Equation of Motion for MDOF System	7
2.1.2. Natural Vibration Modes & Frequencies.....	7
2.1.3. Damping Matrix using Caughey Series	8
2.1.4. State-Space Models	8
2.1.5. Direct Time-Integration using the Central Difference Algorithm.....	9
2.2. 8-DOF Spring-Mass System.....	9
2.2.1. Test Structure Description	10
2.3. Data Acquisition	12
2.3.1. Structural State Conditions	12
2.3.2. Excitation.....	12
2.3.3. Responses	13
3. FEATURE EXTRACTION	16
3.1. Statistical Features	16
3.1.1. Peak amplitude.....	17
3.1.2. Mean and Root-Mean-Square.....	17
3.1.3. Variance and Standard Deviation	18
3.1.4. Skewness.....	18
3.1.5. Kurtosis.....	19
3.2. Probability Density Function (PDF) & Cumulative Distribution Function (CDF) ...	19
3.2.1. Probability Density Function (PDF).....	19
3.2.2. Cumulative Distribution Function (CDF).....	20

3.3. Time-Series Analysis using Autoregressive-based Models.....	21
3.4. Features based on Modal Properties	23
3.4.1. Resonance Frequencies.....	24
3.4.2. Vibration Mode Shapes	25
3.4.3. Mode Shape Curvature	26
3.4.4. Modal Strain Energy.....	27
4. FEATURE SELECTION.....	29
4.1. Filter Methods.....	29
4.1.1. F-Statistic or F-Test	29
4.1.2. Chi-Squared Test	30
4.1.3. Feature Importance Method.....	31
4.1.4. Pearson's Correlation.....	33
4.1.5. Mutual Information Technique.....	35
5. STRUCTURAL DAMAGE CLASSIFICATION USING SUPPORT-VECTOR MACHINES (SVM)	38
5.1. Introduction to Support-Vector Machines (SVM).....	38
5.1.1. Hyperplane.....	38
5.1.2. SVM Classifier	38
5.1.3. Optimal Hyperplane and Margin	39
5.1.4. Tuning Parameters	40
5.2. Application of SVMs for Structural Damage Classification	41
5.2.1. Structural Damage Classification using all Statistical Features	41
5.2.2. Structural Damage Classification using selected Statistical Features	43
5.2.3. Structural Damage Classification using ARX Model Features	44
6. CONCLUSION.....	46
7. BIBLIOGRAPHY	48

1. INTRODUCTION

1.1. Introduction to Machine Learning

1.1.1. Data-driven modelling

Data-driven modelling and scientific discovery reflects a paradigm shift on how many scientific and engineering challenges are approached. Historically, the data that science and engineering rely on has been limited, and it was typically gathered through experiments designed to validate a hypothesis. Each experiment yielded a little amount of information. Data is now ubiquitous and widely gathered in each and every experiment. [1]

The challenges of using data-driven techniques to cater across all parts of engineering modelling and science, and the seeking to apply what has been learned from the classical method, is driving a hybrid approach in which data-driven modelling is aided with physical knowledge [2]. The objective of establishing hybrid approaches is to increase the consistency of the results by applying core principles.

1.1.2. Data-driven methods: Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL)

Artificial Intelligence (AI) – Artificial intelligence (AI) is described as a machine's ability to imitate intelligent human behaviour by employing human-inspired algorithms to approximate conventionally difficult problems [3]. Knowledge representation, reasoning, automated planning, learning, natural language processing, vision, robotics, and general intelligence are some of the core goals of AI research [4].

Machine Learning (ML) – It is a branch of artificial intelligence that tries to teach computers how to perform tasks with data without explicit programming. It uses numerical and statistical methodologies, as well as artificial neural networks, to embed learning in models. Models are created through “training” compute runs or by using them.

Deep Learning (DL) – Deep learning, also known as deep neural learning, is a subset of machine learning that employs neural networks to assess various elements with a structure similar to that of the human nervous system. It is computationally more intense than machine learning.

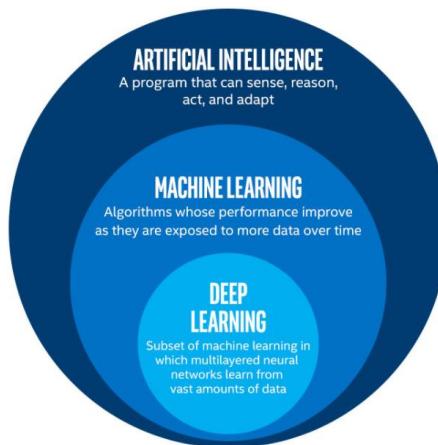


Fig. 1.1. Interrelation between AI, ML and DL.

1.1.3. Overview of Machine Learning

Machine learning (ML) is a major branch of artificial intelligence (AI) that studies, designs, and develops algorithms that can learn from data and make predictions based on that data. Machine learning refers to a computer's ability to learn without being specifically programmed. To extract information from data, ML-based models can be predictive or descriptive [3].

Despite being a branch of computer science, machine learning is distinct from conventional computational methods. Algorithms are collections of explicitly programmed instructions that computers use, to compute or solve problems in conventional computing [5]. Machine learning algorithms, on the other hand, enable computers to train on data inputs and then use statistical analysis to produce values that are within a certain range. As a result, machine learning makes it easier for computers to build models from sample data and automate decision-making processes based on data inputs [5].

Machine learning approaches are used in various disciplines; including computer science, information theory, engineering, physical and biological sciences, control computational complexity, probability and statistics, financial market, and theory and philosophy.

In 1959, Arthur Samuel defined machine learning as a “Field of study that gives computers the ability to learn without being explicitly programmed” [6]. Thomas M. Mitchell provided a widely quoted, more formal definition: “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E” [7]. Alan Turing's seminal paper [8] established a benchmark for demonstrating machine intelligence, requiring a machine to be knowledgeable and sensitive in a way that cannot be distinguished from that of a person.

1.1.4. Types of Machine Learning techniques

Tasks are usually categorised into broad categories of machine learning. These classifications are based on how learning is obtained and how the system is given input on the learning.

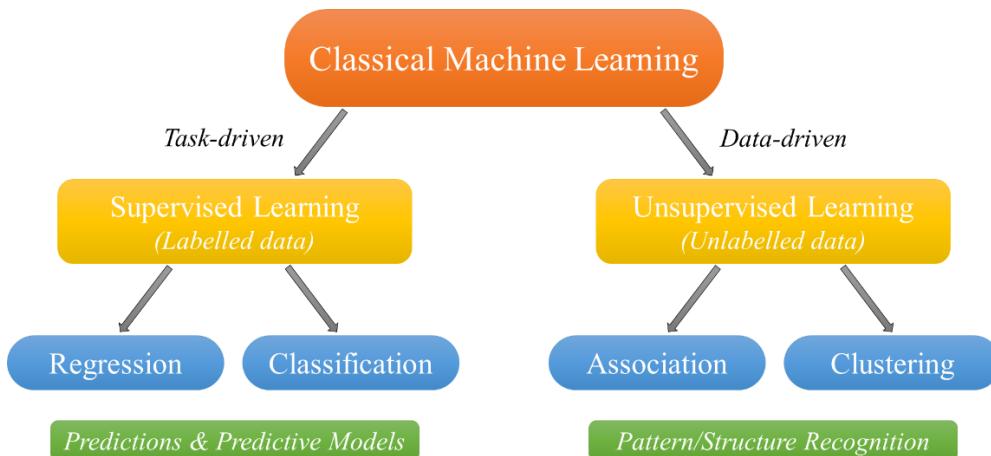


Fig. 1.2. Classification of Classical Machine Learning techniques

Supervised Learning – The computer is programmed by using well-labelled data. Labelled data is information that has already been marked with the correct response. The method of an algorithm learning from a training dataset can be compared to that of an instructor supervising the process of learning. Supervised learning techniques are mainly used for regression and classification problems.

- *Regression*: Regression predicts a single output value using training data. It is about predicting a continuous quantity. For example, predicting the price of stocks over a period.
- *Classification*: Classification means to group the output inside a class. For example, categorizing emails as important, spam, promotional, etc.

Unsupervised Learning – Since the data in unsupervised learning isn't classified, the learning algorithm is left to find correlations among the data it's given. Machine learning methods that promote unsupervised learning are particularly useful because unlabelled data is more abundant than labelled data [5]. The goal of unsupervised learning may be discovering hidden patterns within a data or learning features from raw data. In association and clustering problems, unsupervised learning methods are commonly used.

- *Association*: An association rule learning problem is where you want to discover rules that describe large portions of your data. This is about finding interesting relationship among variables in huge datasets. People who purchase a new home, for example, are more likely to purchase new furniture.
- *Clustering*: It is primarily concerned with identifying a structure or pattern in a set of uncategorized data. If natural clusters (groups) exist in your data, clustering algorithms can process it and find them. Customers could be categorised based on their purchasing power, for example.

Reinforcement Learning – A computer program interacts with a dynamic environment in which it must perform a certain goal (such as driving a vehicle), without a teacher explicitly telling it whether it has come close to its goal or not. Another example is learning to play a game by playing against an opponent [9].

Table 1.1. *Quick comparison between Supervised & Unsupervised learning*

Parameter	Supervised Learning	Unsupervised Learning
Input data	Input and output variables will be given.	Only input data will be given.
Process	Algorithms are trained using labeled data.	Algorithms are used against data which is not labeled
Computational complexity	Simpler method	Computationally complex
Accuracy of results	Highly accurate and trustworthy method.	Less accurate and trustworthy method.
Real-time learning	Learning method takes place offline.	Learning method takes place in real time.

1.2. Structural Health Monitoring

Structural Health Monitoring (SHM) is the process of implementing a damage detection and characterization strategy for engineering structures [10]. The SHM process involves the observation of structural responses over time with periodically spaced measurements, the extraction of damage sensitive features, correlation with undamaged baseline data or statistical patterns, and damage prognosis, i.e., estimating the structural facility's remaining service life [10]. This is accomplished by making it easier to detect and assess damage in a structure while it is still in use, which could adversely impact its capacity to properly and safely execute its intended purpose [11]. As a result, the challenge of damage detection is at the core of SHM, with the goal of detecting damage as soon as possible.

1.2.1. Approaches to Structural Health Monitoring

Broadly two approaches have traditionally been adopted for automated damage detection in SHM: physics-based and data-driven.

Physics-based approaches rely on the physical principles regulating structural behaviour in one form or another in order to extract relevant information about the damage and its evolution from the collected sensor data. However, the complexity of modelling complicated real-world structures, concerns of numerous sensing modalities, material and/or geometric non-linearity, and uncertainty in material characteristics, boundary conditions, and environmental/operational fluctuations are some of the reasons that make relying only on system physics impracticable in complex real-world systems. As the complexity of the underlying system grows, such an approach becomes increasingly unreliable. Because of recent advancements in information and sensor technologies, it is now possible to continuously or sporadically monitor a large number of parameters in-situ in big/complex real-world structures [11]. This promotes the adoption of a data-driven strategy for SHM, in which damage assessment is treated as a form of statistical pattern recognition problem, at least at a lower level, thus avoiding some of the major challenges associated with the physics-based approach.

1.2.2. Data-driven Damage Identification

Damage detection is a subset of the larger issue of damage identification. Damage in a structure is defined as the changes in material and/or geometric properties. The goal of a SHM system should be to gather enough information about the damage so that suitable remedial action can be performed to return the structure or system to high-quality functioning, or at the very least assure safety. In addition, for efficiency's sake, the monitor should only return the necessary information that is required [10]. It's helpful to think of the identification problem as a hierarchical structure with the framework composed of various levels as following:

- *Level 1: Detection.* The method provides a qualitative indicator that the structure may be damaged.
- *Level 2: Localization.* The method gives information about the location of the damage in the structure.
- *Level 3: Characterization.* The method classifies the type and quantifies the severity of the damage in the structure.
- *Level 4: Prediction.* The method offers information about the safety of the structure, for example estimates a residual life.

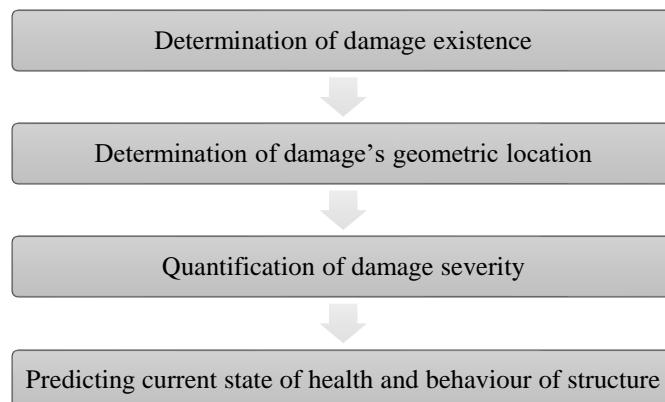


Fig. 1.3. Levels of damage detection

The procedure devised to uncover damage-related information and to support decision making in the SHM by applying data-driven methods involves:

- *Data acquisition*: to gather a large number of signals from sensors put on the structure in a variety of environmental and operational circumstances.
- *Data pre-processing*: Averaging (to improve signal-to-noise ratio), filtering (to eliminate outbound noise), and energy-normalizing (to reduce signal output unpredictability) are some of the primary signal processing techniques used to improve signal quality.
- *Feature extraction*: This is the process of extracting useful or possibly useful features from sensor inputs. This layer can be set up in two ways. To begin, statistical or signal processing techniques are utilised to calculate only one or a few features, which are then used as change detector(s) to continually monitor structural conditions. Second, once a change is recognised, advanced signal processing techniques and machine learning algorithms are used to extract more features from the data received, revealing more comprehensive characteristics. Because developing a "super feature" that is sensitive to all sorts of damage, robust to all environmental influences, and successful in all SHM levels is difficult, it is recommended to generate a large number of features to adapt the framework to diverse scenarios [12].
- *Feature selection*: To rank and select effective features from all the feature candidates generated in the feature extraction module. The selection process can be automated using a variety of machine learning methods. The result is a multi-feature vector with more discriminatory information and fewer redundant data. To improve the computational efficiency and performance of learning algorithms, feature selection is critical. Furthermore, the machine learning algorithms' effective/ineffective features can frequently provide insights into the sensor-structure system [12].
- *Pattern recognition*: To classify datasets into different categories based on the objective of a particular SHM stage: determining whether the pipe is damaged or undamaged for damage detection; determining a specific zone of interest for damage localization; and determining a type and a certain degree of severity for damage characterization.

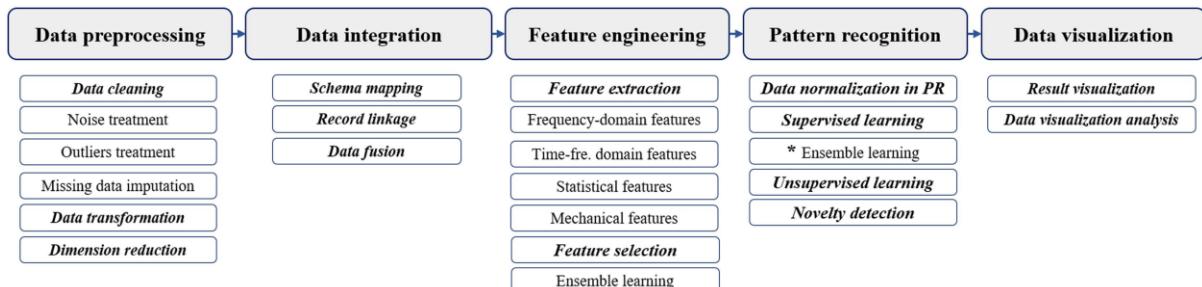


Fig. 1.4. Pipeline for data-driven analysis in SHM [13]

1.2.3. Vibration-based Damage Detection

Damage detection has a wide range of applications in practically every field of engineering research. Damage detection systems used in mechanical engineering for vibration-based health monitoring of rotating machines and machine parts are the most successful and well-established [14][15]. Such systems assess the state of a machine's bearings, gears, and shafts by measuring the machine's vibration response in terms of displacement, velocity, or acceleration [16]. Damage detection is performed by extracting a "vibration signature" from the response spectrum and then comparing the present signature to that of the reference (i.e.,

undamaged) state using a pattern recognition approach. This means that every noticeable change in a system's vibration signature can be linked to a specific form of damage. Because the types and sites of damage in such structures are well-defined, any change in the vibration signature may be easily linked to a specific type and area of damage [17].

Visual inspections are the mainstay of traditional approaches to damage diagnosis in civil structures. However, there are a number of difficulties that prevent these strategies from being used in reality [18]. Because civil structures are generally big in scale, regular inspection can be both time consuming and labour intensive. Second, because traditional systems rely on human judgement, professional and highly-trained labour is required. Third, because non-structural components and decorative coverings such as flooring, cladding, and false ceilings cover the load-carrying structural sections (i.e., footings, columns, beams, and slabs) in most situations, visual inspection is difficult. Despite this, the success of vibration-based methods in machinery health monitoring has prompted academics to apply similar techniques to civil infrastructure for SHM. As a result, numerous vibration-based structural damage detection systems have been created [19][20]. The ultimate goal of these systems is to overcome the problems associated with traditional structural damage detection methodologies by providing a systematic, viable, and consistent method of detecting, locating, and characterizing structural damage based on the vibration response of the monitored structure [21].

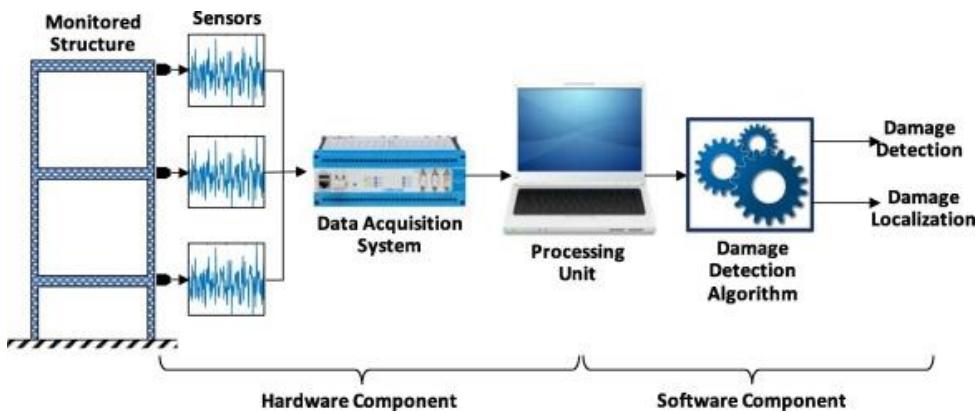


Fig. 1.5. Main components of structural damage detection systems. [21]

1.3. Objective of the Report

The objective of this study is to apply the statistical pattern recognition paradigm for SHM to data acquired from the vibration response of an 8-DOF system and to devise an approach for structural damage detection, i.e., to classify or distinguish between damaged and undamaged states, by developing machine learning models and applying on simulated data.

To investigate the applications of machine learning approaches to SHM, a literature study is done. Signal statistics and modal properties, as well as analysis of vibration response time-series using autoregressive models, are used to extract a variety of structural damage-sensitive features. Informative, discriminating, and independent features are ranked and selected from a pool of damage-sensitive features based on their effectiveness for damage detection. Supervised learning classification techniques are employed to develop predictive models for structural damage detection.

This report concludes with a general discussion and summary regarding the effectiveness of the various feature extraction and statistical modelling procedures to identify damage in the test structure.

2. MULTI-DEGREE OF FREEDOM SYSTEM

2.1. Overview of Multi-Degree of Freedom (MDOF) System

A multi-degree of freedom (MDOF) system requires more than one coordinates to describe its motion. These coordinates are called generalized coordinates when they are independent of each other and equal in number to the degree of freedom of the system.

2.1.1. Dynamic Equation of Motion for MDOF System

The Newton's second law is applied to each mass of the system in order to obtain the system's equation of motion (EOM). The system of equations of the EOM for the system's masses are linear ordinary differential equations. These equations are coupled with each other. These equations can be rearranged in matrix form. The general dynamic equation of motion for a n -degrees of freedom system is represented in Equation 2.1.1.

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{P}(t) \quad (2.1.1)$$

Where, $\ddot{\mathbf{x}}(t) \in \mathbb{R}^n$ is acceleration vector, $\dot{\mathbf{x}}(t) \in \mathbb{R}^n$ is velocity vector and $\mathbf{x}(t) \in \mathbb{R}^n$ is displacement vector. $\mathbf{P}(t) \in \mathbb{R}^n$ is the time varying force. $\mathbf{M} \in \mathbb{R}^n \times \mathbb{R}^n$ is the mass matrix, $\mathbf{C} \in \mathbb{R}^n \times \mathbb{R}^n$ is the damping matrix and $\mathbf{K} \in \mathbb{R}^n \times \mathbb{R}^n$ is the stiffness matrix.

2.1.2. Natural Vibration Modes & Frequencies

The natural frequency ω and the corresponding mode of vibration ϕ of the system can be obtained from the solution of undamped free vibration of the system (Equation 2.1.2)

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = 0 \quad (2.1.2)$$

which is given as

$$\mathbf{x}(t) = \phi e^{i\omega t} \quad (2.1.3)$$

Substituting Equation 2.1.3 in Equation 2.1.2, the following eigenvalue and eigenvector problem is obtained,

$$(\mathbf{K} - \omega^2 \mathbf{M})\phi = \mathbf{0} \quad (2.1.4)$$

The equation above represents an eigenvalue-eigenvector problem, which has as many solutions as the matrices order, n , which is also the number of degrees of freedom of the system. Each solution is a pair (ω_k, ϕ_k) .

The eigenvalues ω are called the natural vibration frequencies of the structural system, while the eigenvectors ϕ are called the vibration modes, or modal shapes. It is very important to keep in mind that the modal shapes have no prespecified scale i.e., they can be multiplied by any scale factor, α , and still remain solutions for the eigenproblem:

$$(\mathbf{K} - \omega^2 \mathbf{M})\alpha\phi = \mathbf{0} \quad (2.1.5)$$

Numerical algorithms for solving this eigenproblem are available in many environments. In Python, they are available in `scipy.linalg` module.

2.1.3. Damping Matrix using Caughey Series

Caughey and O'Kelly (1965) showed that a necessary and sufficient condition for classical normal modes to exist was for the damping matrix to satisfy

$$\mathbf{C} = \mathbf{M} \sum_{j=0}^{n-1} a_j (\mathbf{M}^{-1} \mathbf{K})^j \quad (2.1.6)$$

A method to obtain the proportional damping matrix is by using the inverse modal transformation method. From experimentally obtained modal damping factors ζ and natural frequencies, one can construct the diagonal modal damping matrix $\mathbf{C}' = \boldsymbol{\Phi}^T \mathbf{C} \boldsymbol{\Phi}$ as

$$\mathbf{C}' = 2\zeta \boldsymbol{\Omega} \quad (2.1.7)$$

From this, the damping matrix in the original coordinate can be obtained using the inverse transformation as

$$\mathbf{C} = \boldsymbol{\Phi}^{-T} \mathbf{C}' \boldsymbol{\Phi}^{-1} \quad (2.1.8)$$

where, $\zeta \in \mathbb{R}^n$ is modal damping ratios vector, $\boldsymbol{\Phi} \in \mathbb{R}^n \times \mathbb{R}^n$ is mass-normalized modal matrix such that $\boldsymbol{\Phi}^T \mathbf{M} \boldsymbol{\Phi} = \mathbf{I}$, and $\boldsymbol{\Omega} \in \mathbb{R}^n \times \mathbb{R}^n$ is the diagonal matrix of natural frequencies ω .

2.1.4. State-Space Models

The state space method largely alleviates the problem of representing complex systems with differential equations. The state space representation of a system replaces a k^{th} order differential equation with a single first order matrix differential equation.

Pre-multiplying Equation 2.1.1 with \mathbf{M}^{-1} gives

$$\mathbf{I}\ddot{\mathbf{x}}(t) + \mathbf{M}^{-1}\mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{M}^{-1}\mathbf{K}\mathbf{x}(t) = \mathbf{M}^{-1}\mathbf{P}(t) \quad (2.1.9)$$

The state-space representation of system in Equation 2.1.9, which is a system of coupled second order linear differential equations, can be given as

$$\underbrace{\begin{pmatrix} \dot{\mathbf{x}}(t) \\ \ddot{\mathbf{x}}(t) \\ \vdots \\ \dot{\mathbf{Y}}(t) \end{pmatrix}}_{\mathbf{Y}(t)} = \underbrace{\begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} \mathbf{x}(t) \\ \dot{\mathbf{x}}(t) \\ \vdots \\ \mathbf{Y}(t) \end{pmatrix}}_{\mathbf{Y}(t)} + \underbrace{\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}^{-1} \end{pmatrix}}_{\mathbf{B}} \underbrace{\begin{pmatrix} \mathbf{0} \\ \mathbf{F}(t) \end{pmatrix}}_{\mathbf{F}(t)} \quad (2.1.10)$$

or,

$$\dot{\mathbf{Y}}(t) = \mathbf{AY}(t) + \mathbf{BF}(t) \quad (2.1.11)$$

where, \mathbf{I} is an identity matrix of order n , $\mathbf{Y}(t) \in \mathbb{R}^{2n}$ is the state vector, $\mathbf{A} \in \mathbb{R}^{2n} \times \mathbb{R}^{2n}$ is state matrix (constant) and $\mathbf{B} \in \mathbb{R}^{2n} \times \mathbb{R}^{2n}$ is input matrix (constant), $\mathbf{F}(t) \in \mathbb{R}^{2n}$ is the input function, where n is the number of degrees of freedom of the MDOF system.

The system of n coupled second order linear differential equations (Equation 2.1.9) is now converted to the system of $2n$ coupled first order linear differential equations (Equation 2.1.11).

2.1.5. Direct Time-Integration using the Central Difference Algorithm

Central difference method is an explicit time-integration method. For explicit schemes the equation of motion is evaluated at the old time-step t_i , whereas implicit methods use the equation of motion at the new time step t_{i+1} .

Using the central difference formulae for first and second derivative, the expressions for velocity and acceleration in terms of displacement with time-step Δt is given as

$$\text{Velocity: } \dot{\mathbf{x}}_i = \frac{\mathbf{x}_{i+1} - \mathbf{x}_{i-1}}{2\Delta t} \quad (2.1.12)$$

$$\text{Acceleration: } \ddot{\mathbf{x}}_i = \frac{\mathbf{x}_{i+1} - 2\mathbf{x}_i + \mathbf{x}_{i-1}}{(\Delta t)^2} \quad (2.1.13)$$

At time $t = t_i$, the equilibrium condition

$$\mathbf{M}\ddot{\mathbf{x}}_i + \mathbf{C}\dot{\mathbf{x}}_i + \mathbf{K}\mathbf{x}_i = \mathbf{P}_i \quad (2.1.14)$$

Substituting Equation (2.1.12) and Equation (2.1.13) into Equation (2.1.14),

$$\underbrace{\left(\mathbf{M} + \frac{\Delta t}{2} \mathbf{C} \right)}_{\alpha} \mathbf{x}_{i+1} = (\Delta t)^2 \mathbf{P}_i - \underbrace{\left((\Delta t)^2 \mathbf{K} - 2\mathbf{M} \right)}_{\beta} \mathbf{x}_i - \underbrace{\left(\mathbf{M} - \frac{\Delta t}{2} \mathbf{C} \right)}_{\gamma} \mathbf{x}_{i-1}$$

or

$$\mathbf{x}_{i+1} = \mathbf{a}^{-1} \left((\Delta t)^2 \mathbf{P}_i - \beta \mathbf{x}_i - \gamma \mathbf{x}_{i-1} \right) \quad (2.1.15)$$

Using above Equation 2.1.15, \mathbf{x}_{i+1} can be found.

At time $t = t_0$, the initial conditions are \mathbf{x}_0 and $\dot{\mathbf{x}}_0$, from equilibrium conditions (Equation 2.1.14) $\ddot{\mathbf{x}}_0$ is obtained. From Equation 2.1.12 and Equation 2.1.13

$$\mathbf{x}_{-1} = \mathbf{x}_0 - \Delta t \dot{\mathbf{x}}_0 + \frac{(\Delta t)^2}{2} \ddot{\mathbf{x}}_0 \quad (2.1.16)$$

\mathbf{x}_{-1} is obtained from above equation in order to start the time integration.

2.2. 8-DOF Spring-Mass System

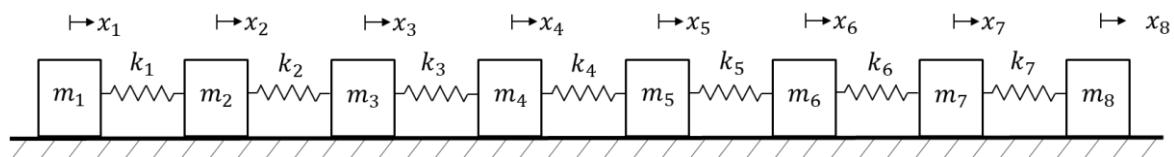


Fig. 2.1. 8-DOF spring mass system

2.2.1. Test Structure Description

The system shown in Fig. 2.1 is an 8-DOF spring mass system consisting of masses $m_1 = 0.4900 \text{ kg}$, $m_2 = m_3 = m_4 = m_5 = m_6 = m_7 = m_8 = m = 0.4652 \text{ kg}$ and springs having stiffness $k_1 = k_2 = k_3 = k_4 = k_5 = k_6 = k_7 = k = 5500 \text{ N/m}$.

Mass Matrix:

$$\mathbf{M} = \begin{pmatrix} m_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & m_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & m_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & m_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & m_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & m_6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & m_7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & m_8 \end{pmatrix}$$

Stiffness Matrix:

$$\mathbf{K} = \begin{pmatrix} k_1 & -k_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -k_1 & k_1 + k_2 & -k_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -k_2 & k_2 + k_3 & -k_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & -k_3 & k_3 + k_4 & -k_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & -k_4 & k_4 + k_5 & -k_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & -k_5 & k_5 + k_6 & -k_6 & 0 \\ 0 & 0 & 0 & 0 & 0 & -k_6 & k_6 + k_7 & -k_7 \\ 0 & 0 & 0 & 0 & 0 & 0 & -k_7 & k_7 \end{pmatrix}$$

Natural Vibration Modes & Frequencies:

On solving the eigenvalue-eigenvector problem (Equation 2.1.4), the natural frequencies ω (rad/sec) and mass-normalized modal matrix Φ are obtained. Vibration mode shapes are shown in Fig. 2.2.

$$\omega = \begin{pmatrix} 0.0000 \\ 42.1563 \\ 82.7567 \\ 120.2799 \\ 153.2848 \\ 180.4686 \\ 200.7315 \\ 213.2380 \end{pmatrix} \text{ (rad/sec)}$$

$$\Phi = \begin{pmatrix} \overbrace{0.5166}^{\varphi_1} & \overbrace{-0.7140}^{\varphi_2} & \overbrace{-0.6657}^{\varphi_3} & \overbrace{-0.5901}^{\varphi_4} & \overbrace{-0.4933}^{\varphi_5} & \overbrace{-0.3811}^{\varphi_6} & \overbrace{0.2589}^{\varphi_7} & \overbrace{-0.1308}^{\varphi_8} \\ 0.5166 & -0.6009 & -0.2595 & 0.1705 & 0.5393 & 0.7248 & -0.6706 & 0.3991 \\ 0.5166 & -0.3976 & 0.2970 & 0.7225 & 0.5001 & -0.1659 & 0.6853 & -0.6060 \\ 0.5166 & -0.1344 & 0.6814 & 0.3904 & -0.5330 & -0.5996 & -0.2944 & 0.7195 \\ 0.5166 & 0.1489 & 0.6712 & -0.4194 & -0.5068 & 0.6184 & -0.2708 & -0.7222 \\ 0.5166 & 0.4099 & 0.2721 & -0.7160 & 0.5266 & 0.1328 & 0.6757 & 0.6136 \\ 0.5166 & 0.6092 & -0.2846 & -0.1365 & 0.5135 & -0.7187 & -0.6806 & -0.4106 \\ 0.5166 & 0.7170 & -0.6764 & 0.6101 & -0.5201 & 0.4096 & 0.2826 & 0.1443 \end{pmatrix}$$

Damping Matrix:

The values of modal damping ratios are assumed to be 5% for all the modes. The damping matrix is computed using Equation 2.1.6, Equation 2.1.7 and Equation 2.1.8.

$$C = \begin{pmatrix} 4.3379 & -2.6519 & -0.6737 & -0.3427 & -0.2241 & -0.1699 & -0.1434 & -0.1323 \\ -2.6519 & 6.2308 & -2.2868 & -0.5377 & -0.2772 & -0.1890 & -0.1515 & -0.1367 \\ -0.6737 & -2.2868 & 6.3326 & -2.2386 & -0.5140 & -0.2674 & -0.1895 & -0.1626 \\ -0.3427 & -0.5377 & -2.2386 & 6.3564 & -2.2288 & -0.5146 & -0.2785 & -0.2155 \\ -0.2241 & -0.2772 & -0.5140 & -2.2288 & 6.3558 & -2.2400 & -0.5405 & -0.3313 \\ -0.1699 & -0.1890 & -0.2674 & -0.5146 & -2.2400 & 6.3299 & -2.2928 & -0.6563 \\ -0.1434 & -0.1515 & -0.1895 & -0.2785 & -0.5405 & -2.2928 & 6.2140 & -2.6178 \\ -0.1323 & -0.1367 & -0.1626 & -0.2155 & -0.3313 & -0.6563 & -2.6178 & 4.2526 \end{pmatrix}$$

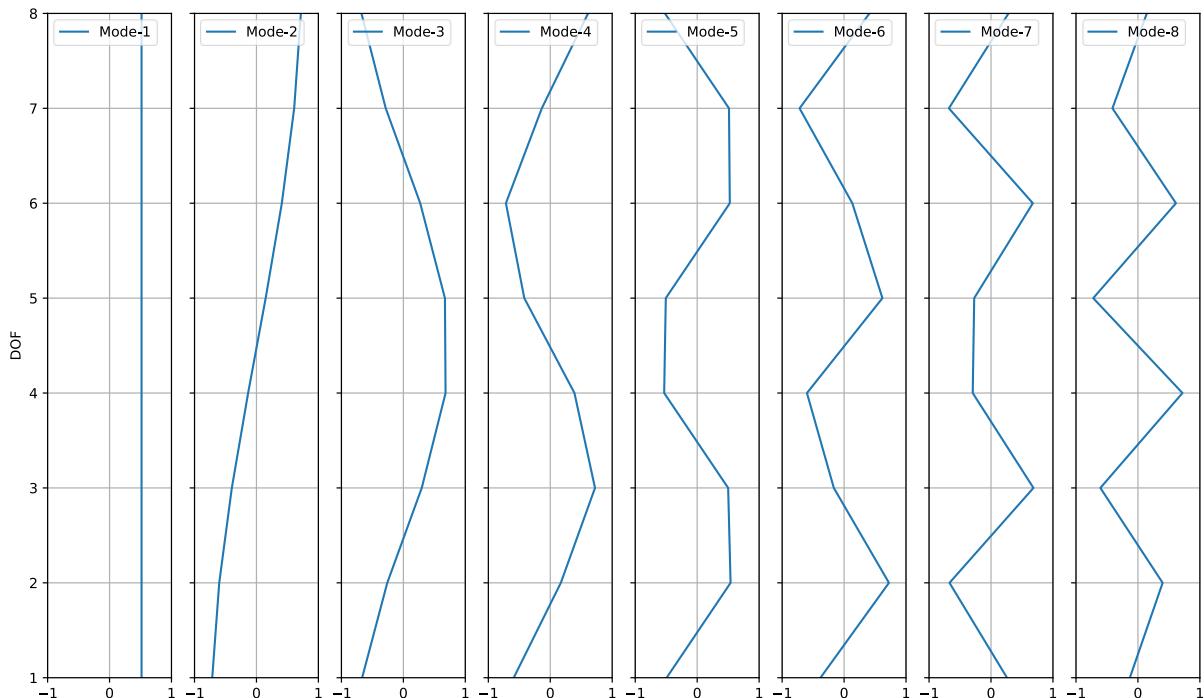


Fig. 2.2. Undamped vibration mode shapes

2.3. Data Acquisition

Force-, acceleration-, velocity- and displacement-time histories (time series or sample records) for a variety of different structural state conditions given in section 2.3.1 were computed using computational simulations. For each structural state condition, simulations were run for 100 different excitation time-histories of 100 seconds each. The acceleration, velocity and displacement responses were obtained by solving differential equations of motion computationally.

2.3.1. Structural State Conditions

Table 2.1. *Data Labels of the Structural State Conditions*

Label	State Condition	Remark
State #0	Undamaged	Baseline condition
State #1	Damaged	5% stiffness reduction in spring 3 and 5
State #2	Damaged	10% stiffness reduction in spring 3 and 5
State #3	Damaged	20% stiffness reduction in spring 3 and 5
State #4	Damaged	50% stiffness reduction in spring 3 and 5

The baseline condition is the reference or undamaged structural state and is labelled State #0. The stiffness change was introduced by reducing 3rd and 5th springs' stiffness by 5%, 10%, 20% and 50% in State #1, State #2, State #3 and State #4 respectively.

2.3.2. Excitation

The system is excited by a random white noise excitation with Gaussian distribution, zero mean and unit variance applied on mass m_1 for 100 seconds. A random force-time history is shown in Fig. 2.3.

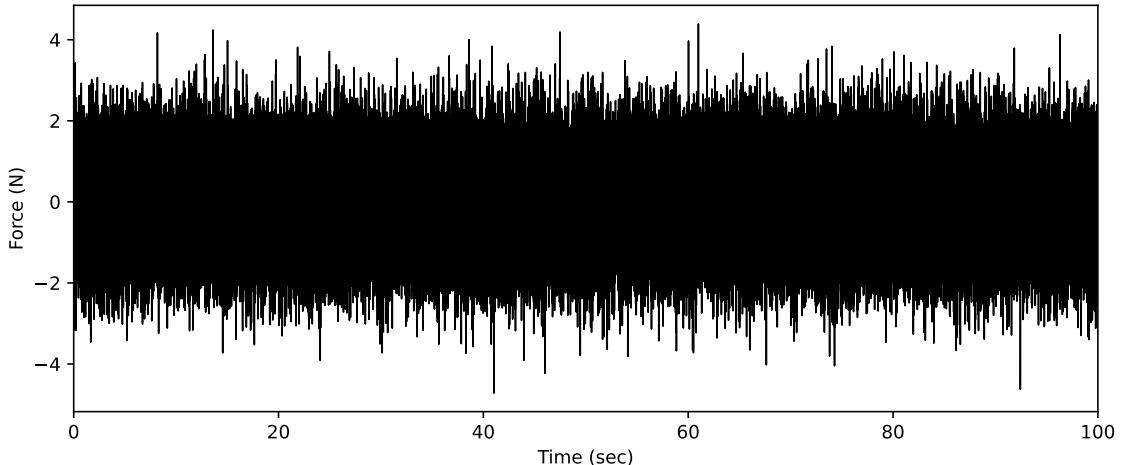
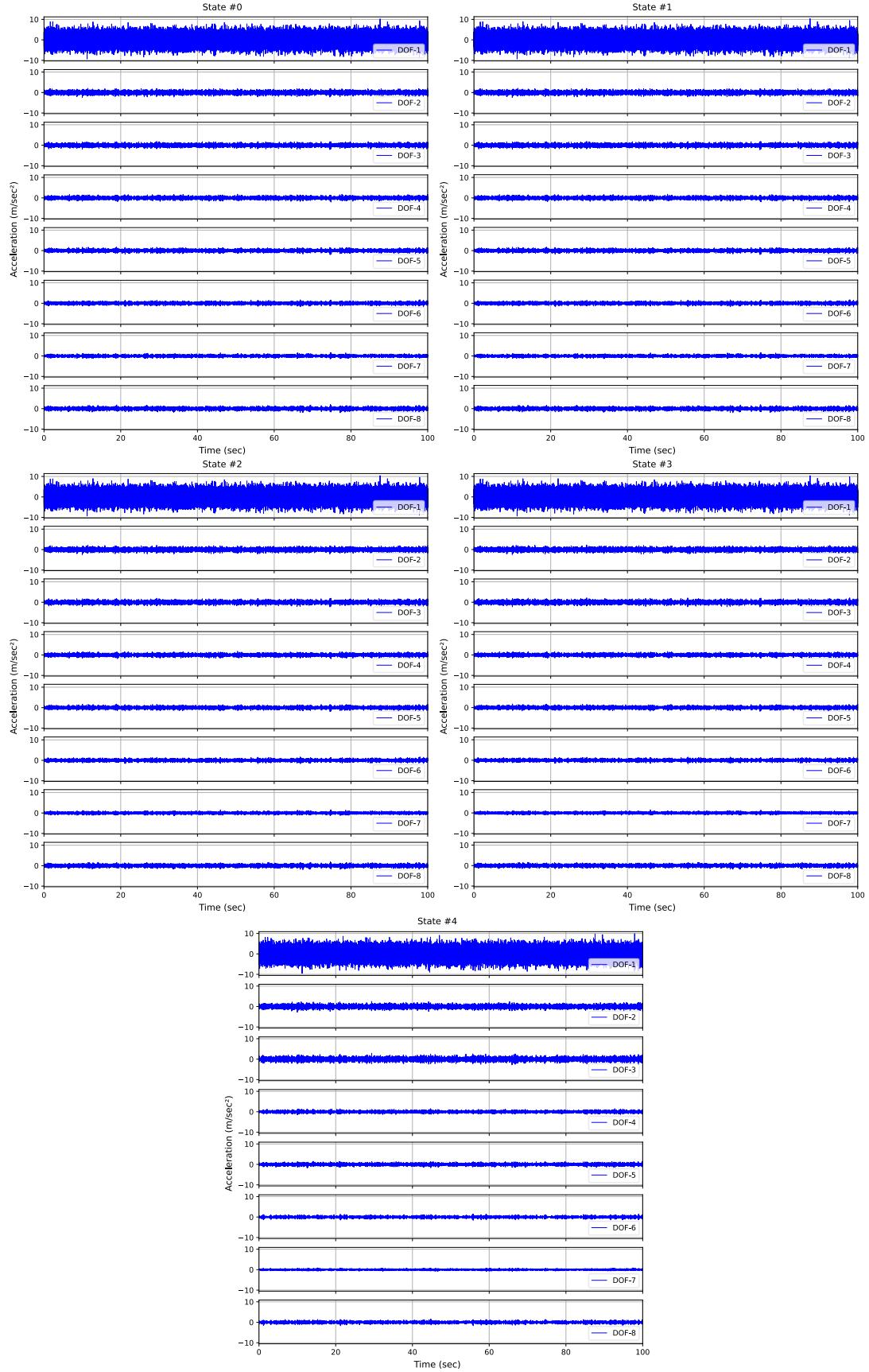


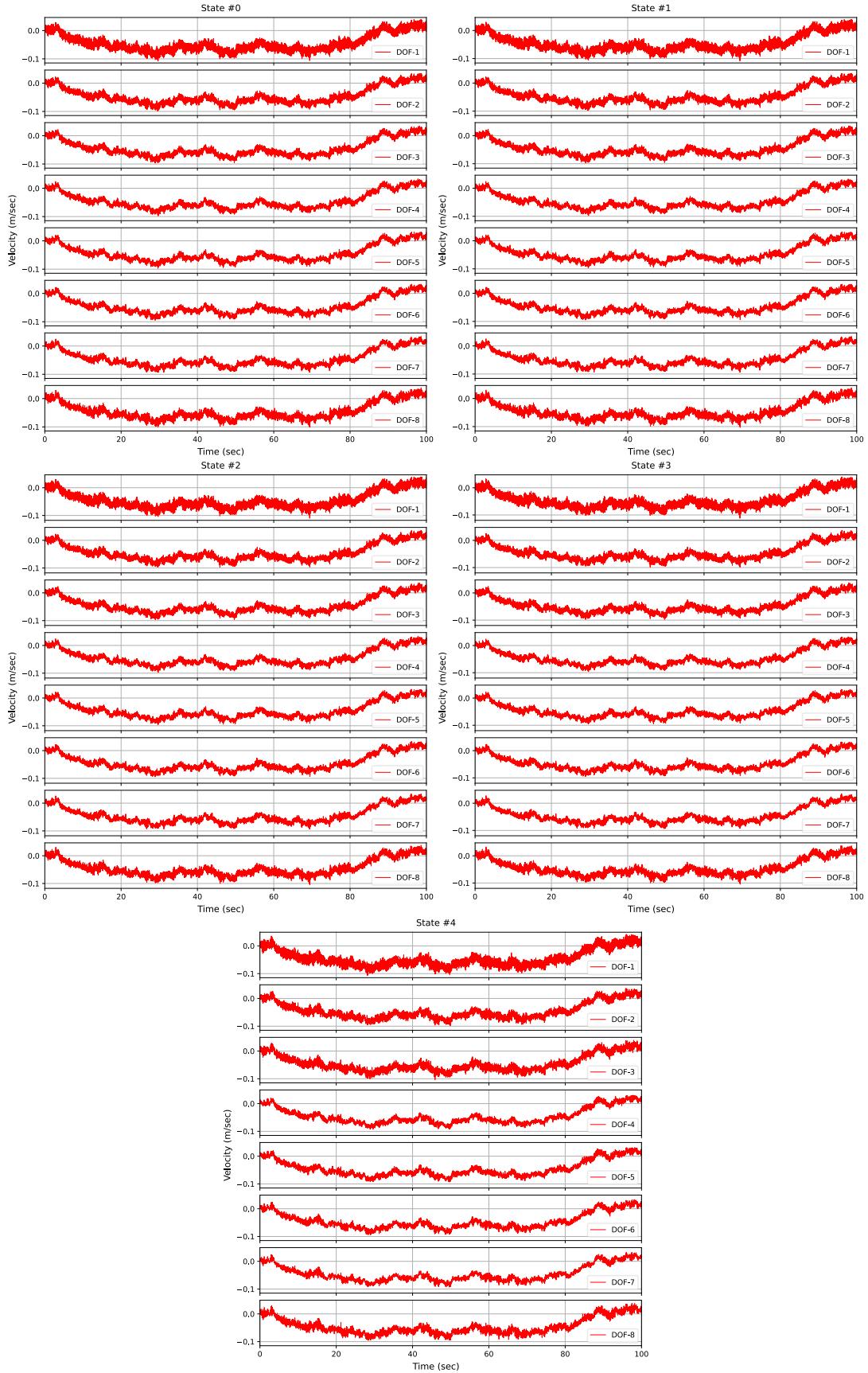
Fig. 2.3. *A random force-time history of 100 time-histories*

2.3.3. Responses

Acceleration-time responses:



Velocity-time responses:



Displacement-time responses:

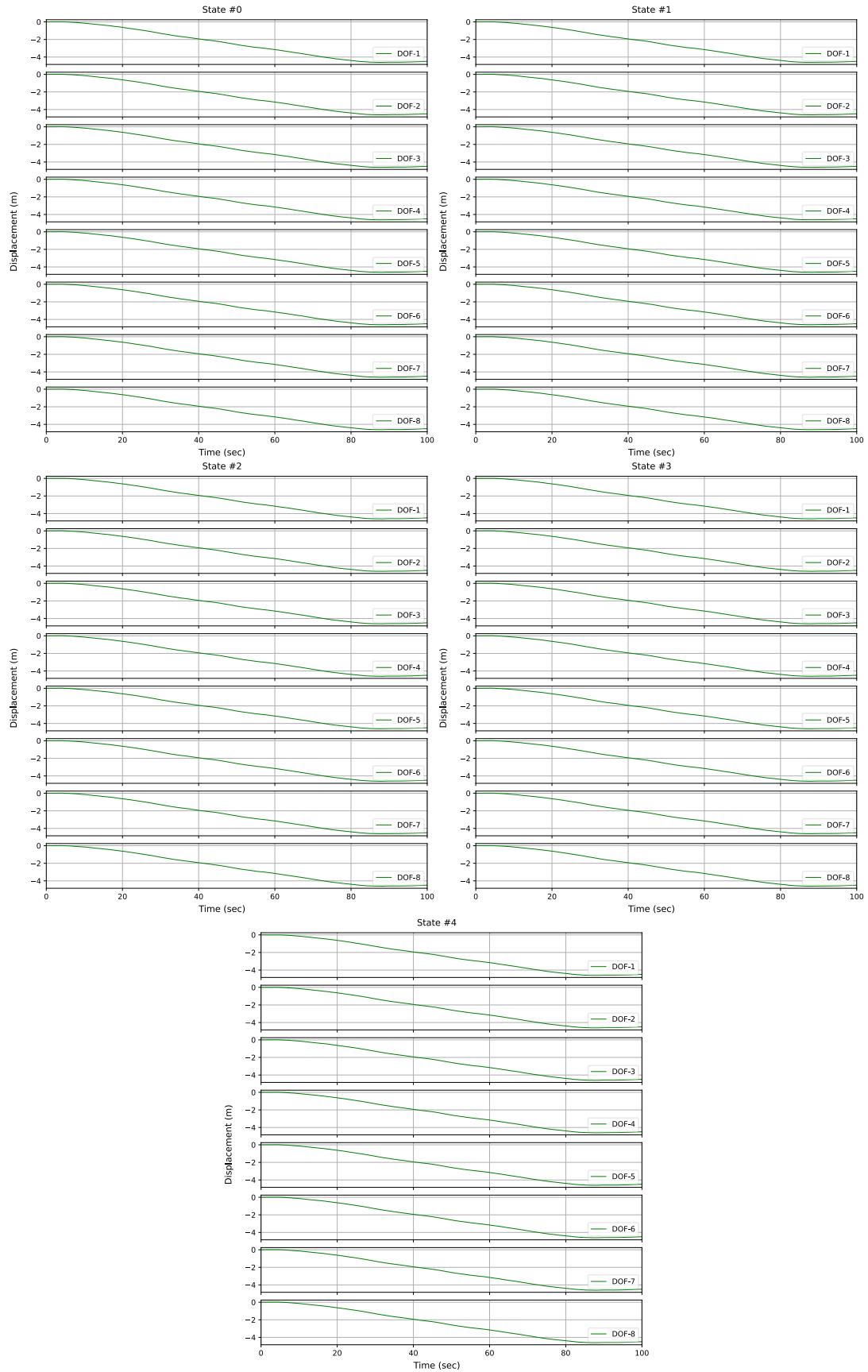


Fig. 2.4. Acceleration, velocity and displacement responses of all structural condition states.

3. FEATURE EXTRACTION

A damage-sensitive feature is some quantity extracted from the measured system response data that is used to indicate the presence of damage in a structure. A damage-sensitive feature should, in theory, change in a predictable way as the damage level rises. The majority of SHM technical literature focuses on identifying features that may accurately distinguish a damaged structure from an undamaged one [22]. These features are the quantities that the pattern recognition and machine learning algorithms will subsequently analyse in an effort to identify and quantify the damage [10]. The feature extraction process is based on fitting a model to the measured system response data, which might be physics-based or data-based. The damage-sensitive features are then the parameters of these models or the predicted errors associated with these models [23]. The ideal method for feature selection in the SHM field is to choose features that are sensitive to damage but not to operational and environmental variations [10].

3.1. Statistical Features

Many basic statistics can be used as damage-sensitive features. Examples of such statistics for an n -point discrete time series are summarised in Table 3.1.

Table 3.1. *Statistics used as damage-sensitive features [10]*

Peak amplitude (y_{peak})	$y_{peak} = \max y_i $	(3.1.1)
Mean (\bar{y})	$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$	(3.1.2)
Root-mean-square (rms)	$rms = \sqrt{\frac{1}{n} \sum_{i=1}^n y_i^2}$	(3.1.3)
Variance (σ^2) ^a	$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$	(3.1.4)
Standard deviation (σ) ^a	$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}$	(3.1.5)
Skewness (γ) ^a	$\gamma = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^3}{\sigma^3}$	(3.1.6)
Kurtosis (κ) ^a	$\kappa = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^4}{\sigma^4}$	(3.1.7)

^a Note that these expressions produce biased estimates of the statistics. However, most time series studied in SHM applications have sufficient samples such that this bias is small.

3.1.1. Peak amplitude

The peak amplitude (Equation 3.1.1) of the measured response is the first statistic to be considered. When damage reduces stiffness while the random input to the system remains unchanged, the peak amplitude of the response will typically increase. Peak amplitude of acceleration of all masses for each structural state is shown in Fig. 3.1.

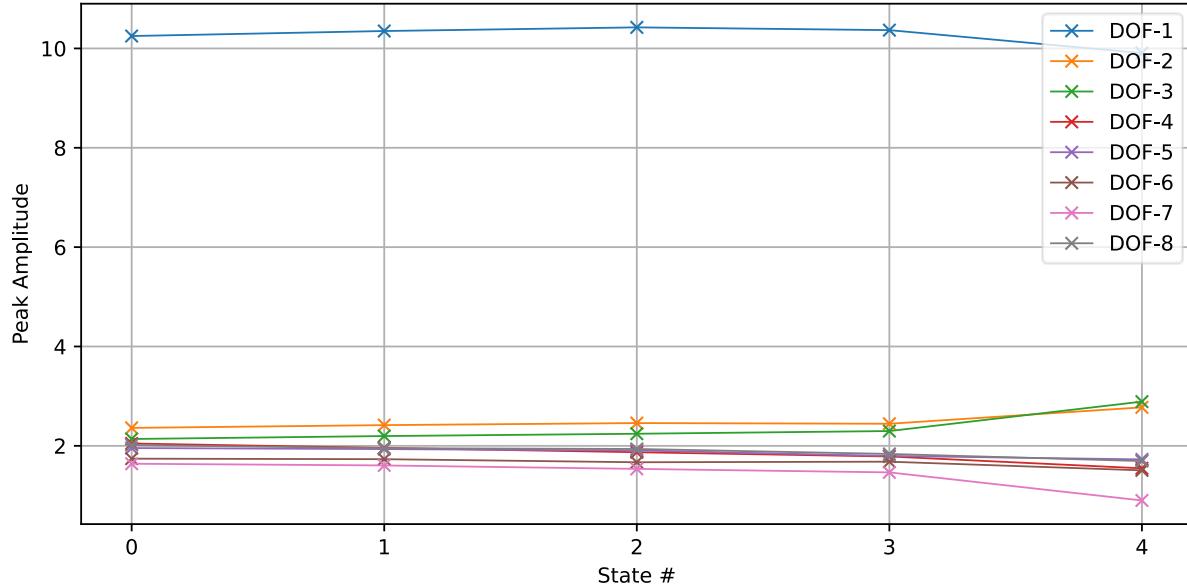


Fig. 3.1. Peak amplitude of acceleration of all masses for each structural state

3.1.2. Mean and Root-Mean-Square

The mean (Equation 3.1.2) and root-mean-square (*rms*) (Equation 3.1.3) measure the central tendency and spread of the data, respectively. The *rms* value increases with an increase in damage level. Root-mean-square of acceleration of all masses for each structural state is shown in Fig. 3.2.

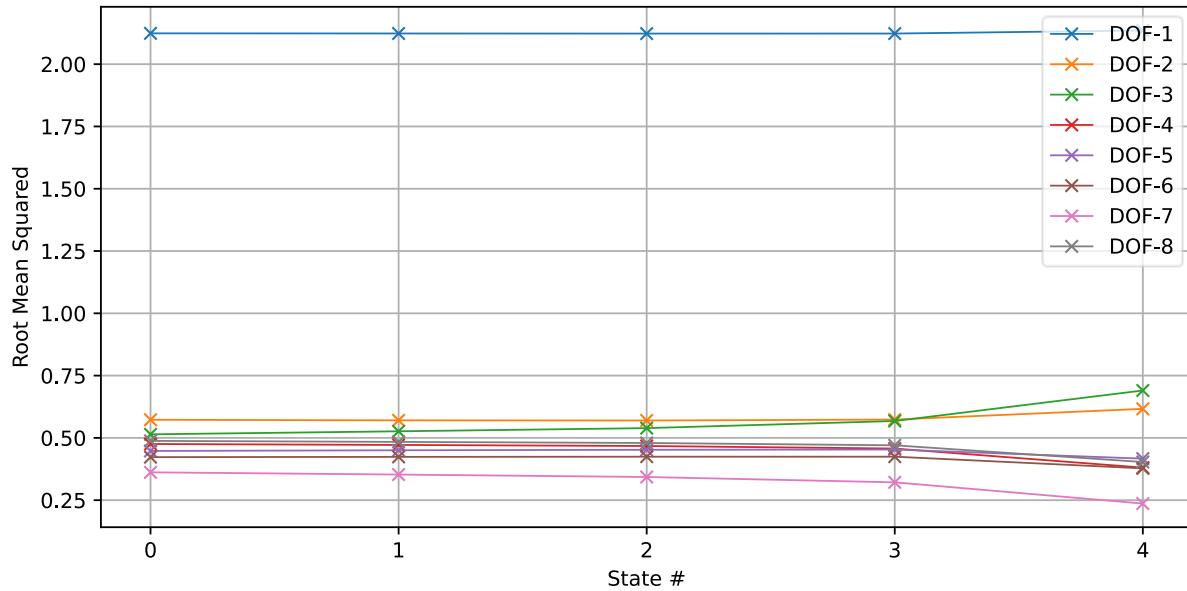


Fig. 3.2. Root-mean-square of acceleration of all masses for each structural state

3.1.3. Variance and Standard Deviation

The square root of the variance (Equation 3.1.4) is called the standard deviation (Equation 3.1.5). The variance and standard deviation measure the dispersion about the mean of the time-series amplitudes. For a fixed level of excitation, damage that reduces the stiffness of the system will, in general, cause an increase in the standard deviation of the measured dynamic response quantities such as acceleration or strain. [10]

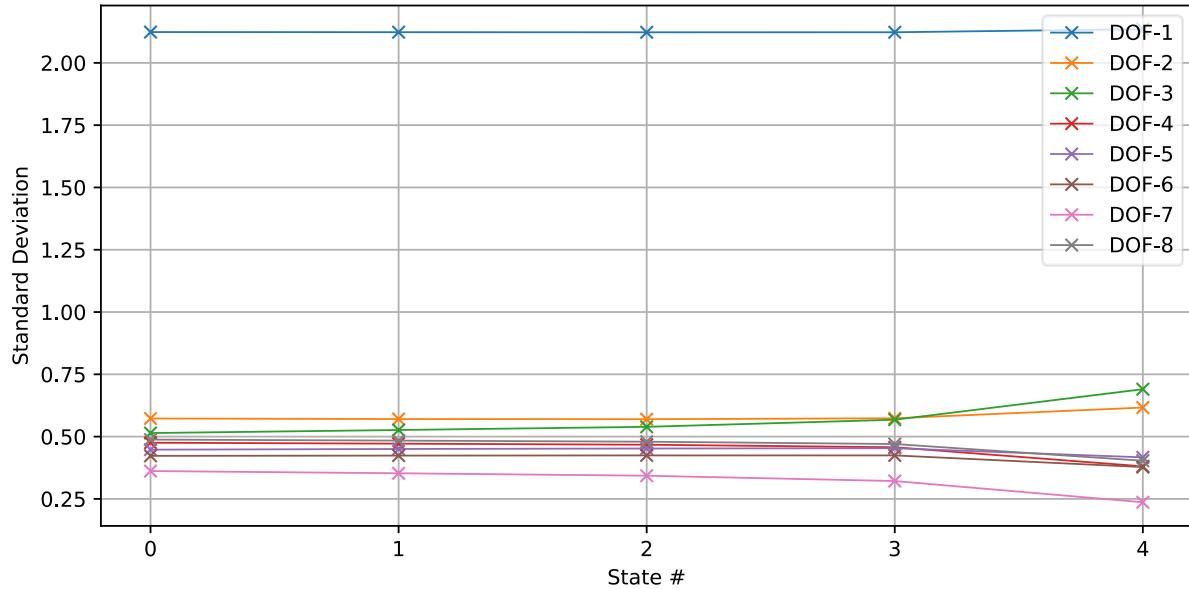


Fig. 3.3. Standard deviation for acceleration response of all masses measured for each structural state

3.1.4. Skewness

The skewness (Equation 3.1.6) is a measure of the symmetry in the distribution of a random variable. Any symmetric distribution such as the normal distribution will have a skewness value equal to zero. The skewness is sensitive to any asymmetry being introduced into an initially symmetric system.

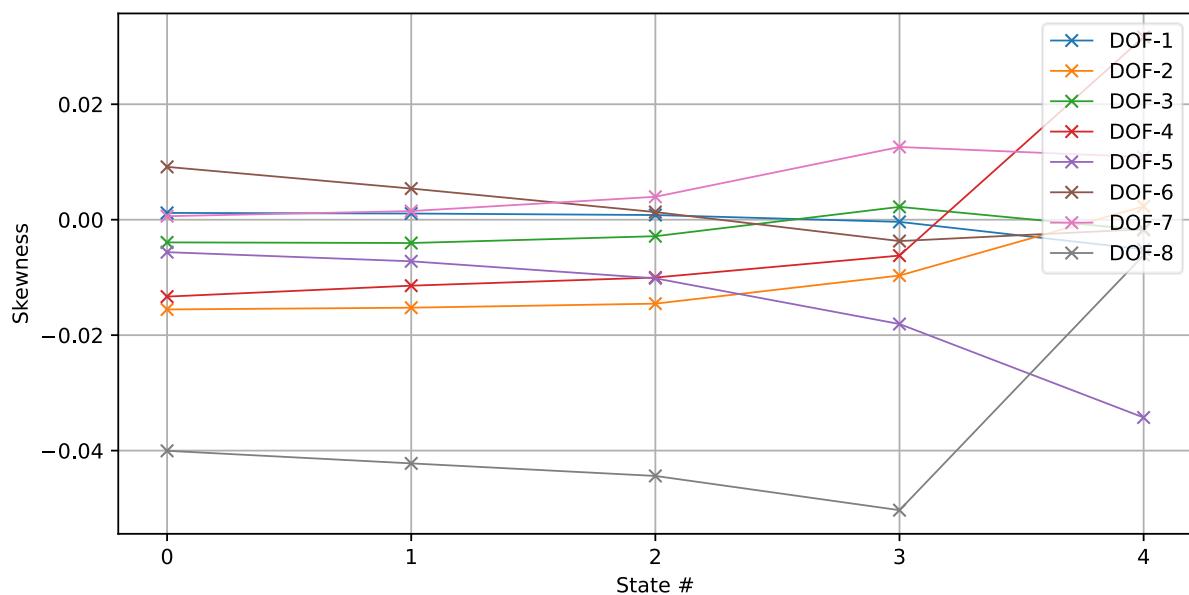


Fig. 3.4. Skewness measured for acceleration of all masses for each structural state

3.1.5. Kurtosis

The kurtosis (Equation 3.1.7) measures the peaked nature of the measured-response distribution. A kurtosis value greater than zero indicates “fat tails”, and a value less than zero indicate a peak sharper than the “normal”. Fat tails indicate the presence of outliers while a sharp peak indicates a lack of outliers [24]. Kurtosis for acceleration response of all masses measured for each structural state is shown in Fig. 3.5.

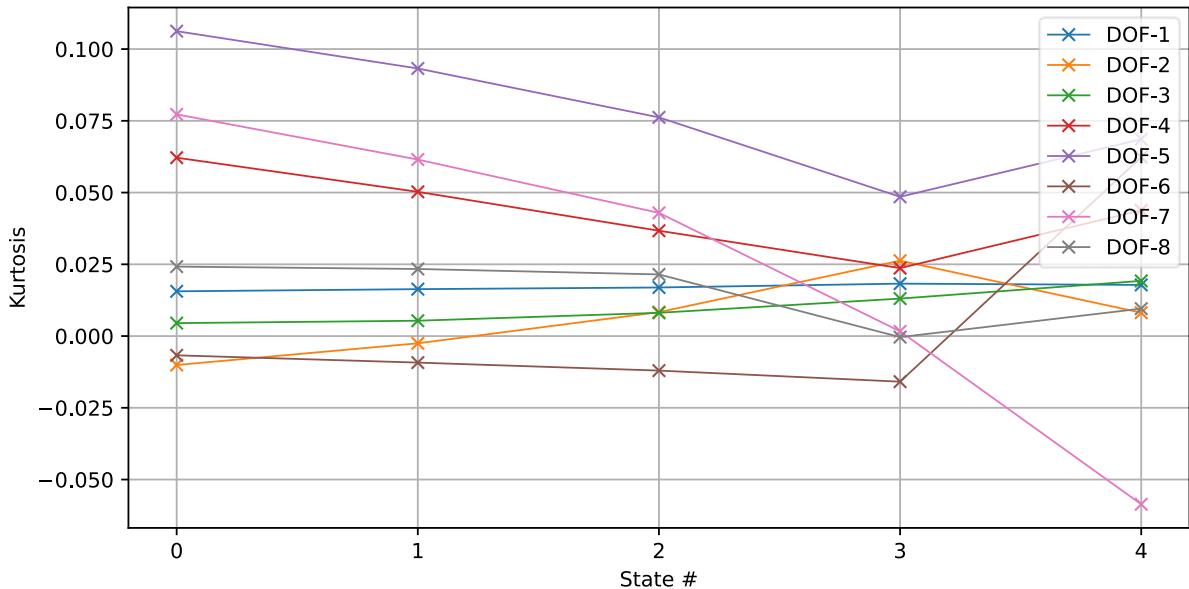


Fig. 3.5. Kurtosis measured for acceleration of all masses for each structural state

3.2. Probability Density Function (PDF) & Cumulative Distribution Function (CDF)

3.2.1. Probability Density Function (PDF)

The probability density function (PDF), denoted as p , is defined by $p(x)dx$ is the probability that random variable X takes a value between x and $x+dx$. Since for continuous distributions the probability at a single point is zero, this is often expressed in terms of an integral between two points as given in Equation 3.2.1.

$$P(X = x; a \leq x \leq b) = \int_a^b p(x)dx \quad (3.2.1)$$

$p(x)$ also satisfies the following conditions:

1. $p(x) \geq 0$, for all $x \in \mathbb{R}$
2. $p(x)$ is piecewise continuous
3. $\int_{-\infty}^{\infty} p(x)dx = 1$

For a normally distributed response data of an undamaged structure, as the level of damage increases, the PDFs deviate from a normal distribution [23].

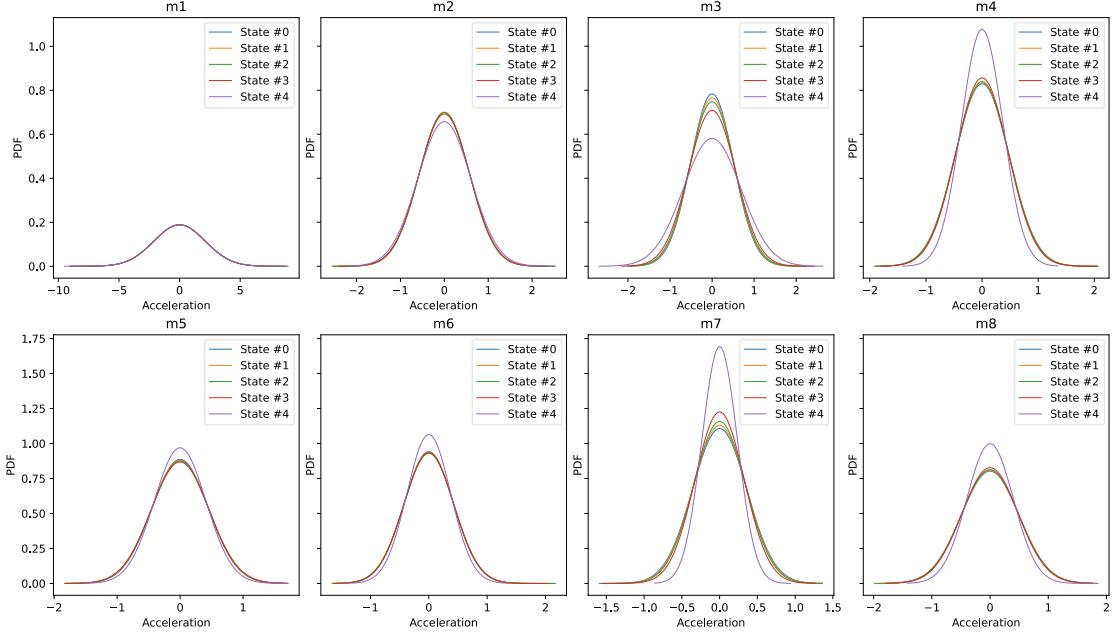


Fig. 3.6. Normal PDF plots for acceleration of all masses for each structural state

3.2.2. Cumulative Distribution Function (CDF)

The cumulative distribution function (CDF) of a random variable X is a function on the real numbers that is denoted as F and is given by

$$F(x) = P(X \leq x) = \int_{-\infty}^x p(t) dt \quad (3.2.2)$$

In other words, the CDF for a continuous random variable is found by integrating the PDF. Note that the Fundamental Theorem of Calculus implies that the PDF of a continuous random variable can be found by differentiating the CDF. This relationship between the PDF and CDF for a continuous random variable is incredibly useful.

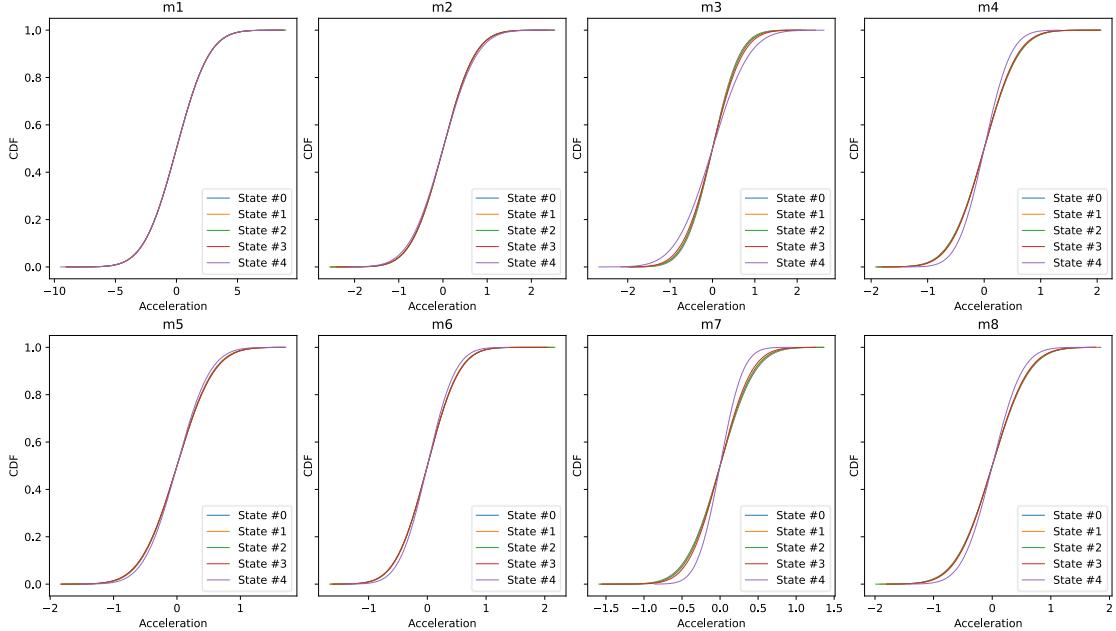


Fig. 3.7. Normal CDF plots for acceleration of all masses for each structural state

3.3. Time-Series Analysis using Autoregressive-based Models

Time-series analysis takes into account the fact that data points taken over time may have an internal structure, such as autocorrelation, trend, or seasonal variation [25]. The approach for performing SHM is to fit a time-series predictive model like autoregressive-based models to vibration response data acquired from the structure in its undamaged state. Damage-sensitive features are extracted based on residuals (the difference between the model's prediction and the observed value), model coefficients, and other model parameters, which are then utilised to check for anomalies.

Autoregressive model, denoted as AR (p), where p is the order of the model is defined as

$$y(t) = \sum_{i=1}^p \varphi_i y(t-i) + \varepsilon(t) \quad (3.3.1)$$

In above equation, $y(t)$ is discrete-time autoregressive signal, $\varepsilon(t)$ is the residual error and φ_i denotes the AR coefficient. The AR model is a kind of time series model that fit time series by establishing linear regressive equations of current value to former value, and thus the autocorrelations in time series can be considered [13]. The autoregressive moving-average model (ARMA), the autoregressive integrated moving average model (ARIMA), and the AR model with exogenous input (ARX) are all examples of AR-based models.

Another widely used time series model is autoregressive moving-average (ARMA) model given in Equation 3.3.2. ARMA model consists of two parts, an autoregressive (AR) part and a moving average (MA) part. The model is usually then referred to as the ARMA (p, q) model where p is the order of the AR and q is the order of the moving average part.

$$y(t) = \sum_{i=1}^p \varphi_i y(t-i) + \sum_{j=1}^q \theta_j \varepsilon(t-j) + \varepsilon(t) \quad (3.3.2)$$

This model is very similar in form to the AR model, but now the response is predicted from linear combination of p previous response values, and q previous moving average terms $\varepsilon(t)$, and the model is defined by the coefficients φ_i and θ_j .

An autoregressive with exogeneous input (ARX) model is similar to an AR model in structure, only there is an additional regression term on the right side of the equation for an external or exogeneous input $e(t)$ with η_j as exogeneous coefficient. An ARX model is defined as

$$y(t) = \sum_{i=1}^p \varphi_i y(t-i) + \sum_{j=0}^q \eta_j e(t-j) + \varepsilon(t) \quad (3.3.3)$$

The AR based models can be employed as a damage-sensitive feature extractor in SHM using one of two methods: (1) using the model parameters, or (2) using residual errors. The first method involves applying an AR based model to signals from both undamaged and damaged structures. The model parameters are then used as damage-sensitive features. The second method involves employing the AR based model to predict the response of data acquired from a potentially damaged structure, using parameters calculated from the baseline condition. [22]

This method is based on the assumption that damage will introduce either linear or nonlinear effects in the signal, and that the linear model developed with the baseline data will no longer

accurately predict the response of the damaged system. As a result, the residual errors caused by the damaged system would increase. [22]

Choosing the right model order is crucial when using this method. A higher-order model may wind up fitting the data's noise and hence will not generalise to other data sets, resulting in a tendency to produce erroneous positive damage indications. A low-order model, on the other hand, will not necessarily reflect the underlying physical system response. There are a variety of techniques for choosing the AR model order, such as Akaike's information criterion [26] or a partial autocorrelation function [27], that are often employed to help decide the appropriate model order.

A two-step AR-ARX model based approach is proposed in [28], an attempt is made to detect the damage regions inside a mechanical system based solely on time series analysis of vibration test data. To locate damage, the standard deviation of residual errors is calculated using a combination of AR and ARX models. Mahalanobis distance is also adopted for damage detection in an experimental research described in [29]. The Mahalanobis distance value for AR coefficients is expected to increase significantly when data from a structural state that differs from the baseline is evaluated [29]. In [24], residual series are calculated using a physically suitable vector autoregressive (VAR) model, which improves the performance of central moment statistics as damage presence and location indicators. ARMA model for the time-domain identification of the dynamic parameters of a linear MDOF structural system is investigated in [30].

In this section, an ARX model-based method of structural damage detection is applied on 8-DOF spring mass system described in section 2.2. The model is derived from linear dynamics and is expressed as an ARX model involving only (acceleration) response data. The damage sensitive feature is the standard deviation of the residual error when the reference model is applied to the observed response of an unknown state [31]. The ARX model was used instead of the usual AR model because the ARX model's coefficients have a direct relationship with structural stiffness [32]. The typical AR model, on the other hand, does not have such a correlation.

The model can be written in ARX form without the excitation term as given in (Equation 3.3.4). For derivation, refer to [31] or [32].

$$\bar{\mathbf{y}}(t) = \mathbf{A}_1\bar{\mathbf{y}}(t-1) + \mathbf{A}_2\bar{\mathbf{y}}(t-2) + \mathbf{B}_0y_j(t) + \mathbf{B}_1y_j(t-1) + \mathbf{B}_2y_j(t-2) + \boldsymbol{\varepsilon}(t) \quad (3.3.4)$$

Where, $\bar{\mathbf{y}} = \{y_1, y_2 \dots y_{j-1}, y_{j+1} \dots y_{n-1}, y_n\}^T$, $y_1 \sim y_n$ are acceleration responses at mass m_1 to m_n for n -DOF system which is excited at point j . $\mathbf{A}_1, \mathbf{A}_2, \mathbf{B}_0, \mathbf{B}_1, \mathbf{B}_2$ are the model coefficient matrices and $\boldsymbol{\varepsilon}$ is $n-1$ row residual error vector. By fitting the ARX model using a standard least-squares approach, measurements acquired from the undamaged structure can be used as reference signals to estimate the coefficient matrices of the ARX model.

In [31], it is demonstrated that comparing the standard deviation of residual errors in predicted signals for an actually measured state, $\sigma(\boldsymbol{\varepsilon}_d)$ (d stands for damaged), to the standard deviation of residual error for the reference (undamaged) state, $\sigma(\boldsymbol{\varepsilon}_0)$, is a useful feature for damage diagnosis. Larger residual errors occur at measurement locations closer to the damage in MDOF systems, allowing the damage location to be determined.

Above method is applied on 8-DOF system described in section 2.2 which is excited by a random force applied on m_1 , therefore, $j=1$ and the Equation 3.3.4 becomes

$$\bar{\mathbf{y}}(t) = \mathbf{A}_1\bar{\mathbf{y}}(t-1) + \mathbf{A}_2\bar{\mathbf{y}}(t-2) + \mathbf{B}_0y_1(t) + \mathbf{B}_1y_1(t-1) + \mathbf{B}_2y_1(t-1) + \boldsymbol{\varepsilon}(t) \quad (3.3.5)$$

Table 3.2 shows the standard deviation of residual errors for all the states described in section 2.2 at all 8 mass points.

Table 3.2. Standard deviation of residual error for damaged states of 8-DOF system

	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
State #0	1.242e-15	4.090e-12	3.925e-12	3.940e-12	4.106e-12	4.075e-12	3.933e-12	2.302e-12
State #1	1.243e-15	8.139e-06	3.893e-04	3.901e-04	3.158e-04	3.157e-04	4.770e-06	3.570e-06
State #2	1.243e-15	1.687e-05	8.017e-04	8.032e-04	6.471e-04	6.468e-04	9.869e-06	7.414e-06
State #3	1.243e-15	3.641e-05	1.690e-03	1.693e-03	1.332e-03	1.331e-03	2.113e-05	1.604e-05
State #4	1.258e-15	1.200e-04	4.831e-03	4.830e-03	3.137e-03	3.128e-03	6.680e-05	5.345e-05

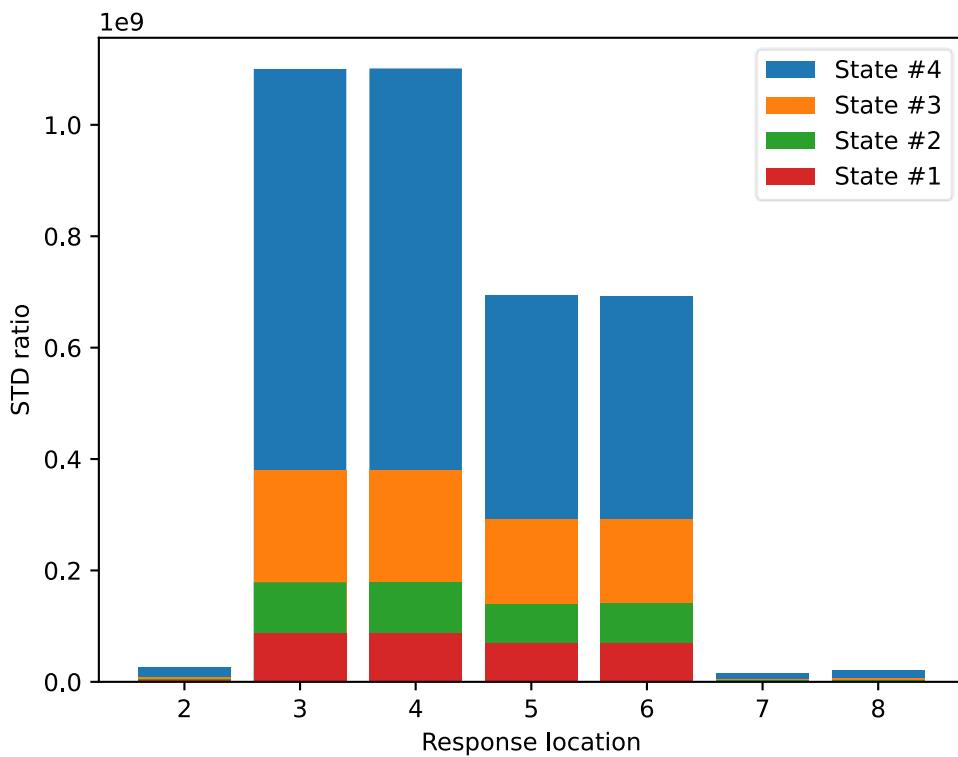


Fig. 3.8. STD ratio ($\sigma(\varepsilon_d)/\sigma(\varepsilon_0)$) of residual errors for damaged states of 8-DOF system

From Fig. 3.8, the locations of damage can be clearly identified since the STD ratio of the residual errors increase abruptly at the points near the damage. The standard deviation (STD) of the residual error, which is the difference between the measured signals from any actual state of the system and the predicated signals from the ARX model established from a reference (undamaged) state, is found to be a damage-sensitive feature.

3.4. Features based on Modal Properties

Modal properties provide an alternative approach to feature extraction where physics-based models are fitted to the measured data and the parameters of these models become the damage-sensitive features. When properly applied, features based on changes in model

parameters can potentially allow one to assess not only the existence and location of damage but also the extent of damage [10].

The underlying idea for using basic modal qualities as damage-sensitive features, such as resonance frequencies and mode shapes, is that damage will change a system's stiffness, mass, or energy dissipation characteristics. Although the concept of employing modal properties as damage-sensitive features sounds intuitive, putting it into practise faces numerous technical obstacles [10].

3.4.1. Resonance Frequencies

When dynamic loads are applied at specific frequencies, structural systems will produce a maximally magnified response. These frequencies are known as the structure's resonance frequencies, and they are determined by the mass, stiffness, and damping properties of the system, as well as the boundary conditions. A method to obtain resonance frequencies of a MDOF system is discussed in section 2.1.

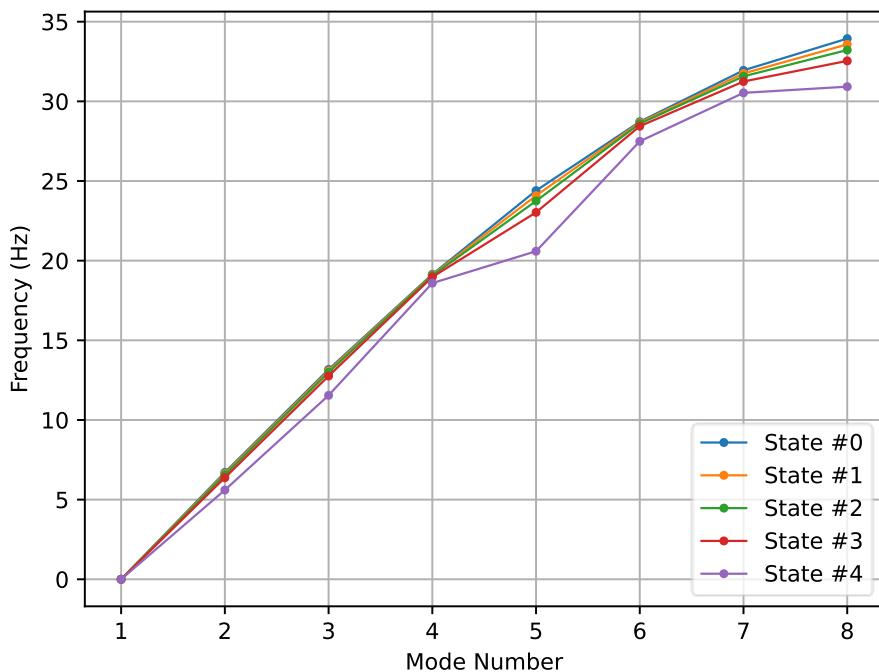


Fig. 3.9. Resonance frequencies for all modes of 8-DOF system

Resonance frequencies for all the eight modes of 8-DOF system are shown in Fig. 3.9. The deviation from undamaged state frequencies near 3rd and 5th mass can be interpreted as decrease in frequencies due to damage cause by decreasing stiffness of springs. Frequency shifts' limited sensitivity to damage necessitates either extremely precise measurements or extremely high amounts of damage for this feature to be useful as a damage indicator. Furthermore, because resonance frequencies are a global property of the structure, it's unclear if changes in this parameter may be utilised to detect more than the presence of damage. In other words, unless some type of model correlation is used, frequencies cannot convey spatial information regarding structural changes [10]. Higher modal frequencies, when the modes are coupled with local responses, are an exception to this constraint.

3.4.2. Vibration Mode Shapes

Mode shapes provide spatially distributed information about the dynamic characteristics of the structure, allowing damage to be located as well as the presence of damage to be established. When a proportionate damping linear structure is excited by a harmonic function whose frequency matches to the structure's resonance frequency, the structure deforms in a distinctive shape with the ratio of displacements for any two degrees of freedom remaining constant over time. Modes of vibration are the names given to these distinct geometries. In most circumstances, the mode form is not readily visible. The operating deflection shapes are observed and these shapes are some linear combination of the modes of vibration. If the damage alters the load path through the structure, it will most likely produce measurable changes in the lower frequency global modes of the structure. [10]

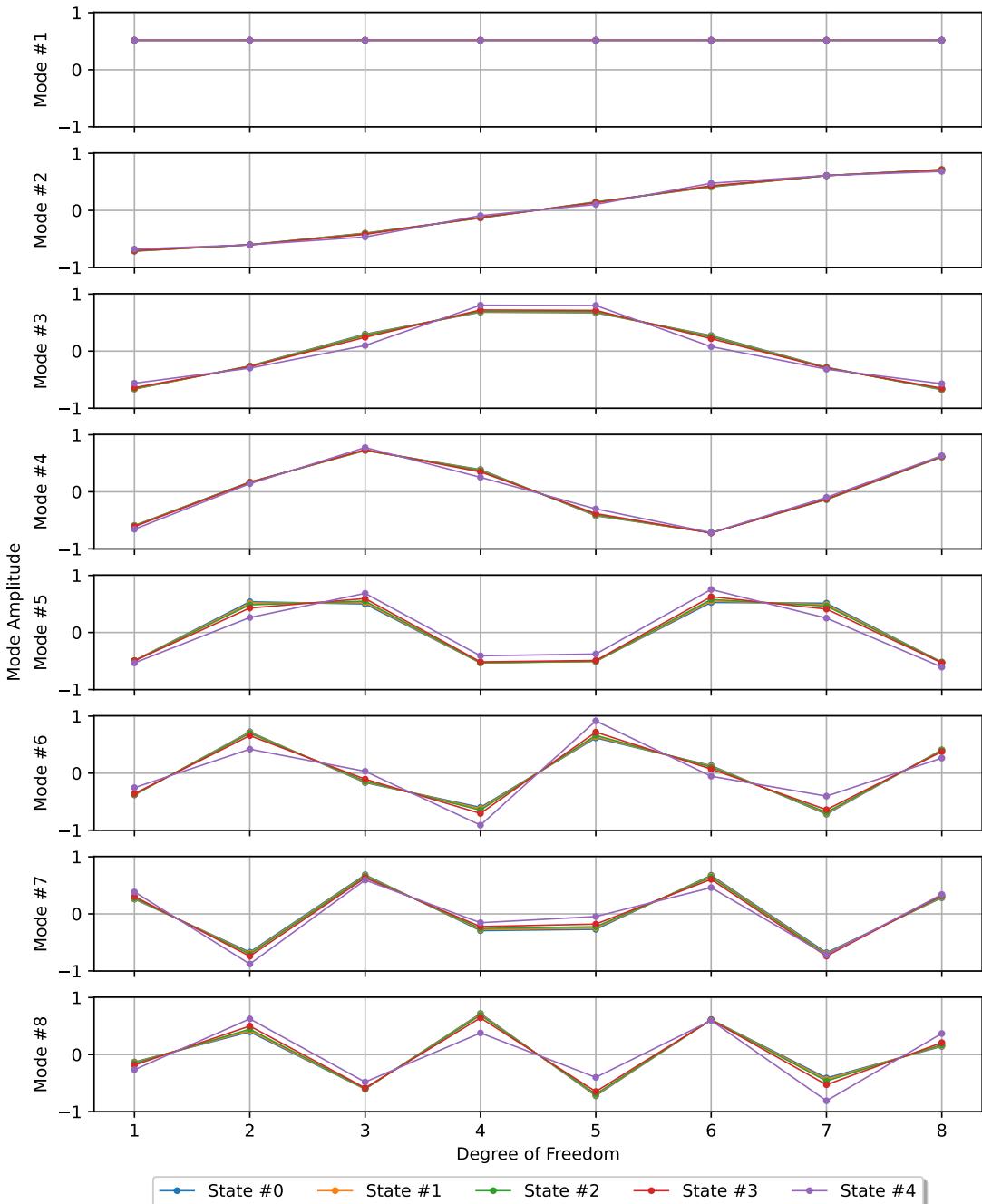


Fig 3.10. Mode shapes calculated for 8-DOF system for each structural state

The modal assurance criteria (MAC) that compares two complex modes, $\{\psi\}_x$ and $\{\psi\}_y$, is defined as

$$MAC(\{\psi\}_x, \{\psi\}_y) = \frac{|\{\psi\}_x^T \{\psi^*\}_y|^2}{(\{\psi\}_x^T \{\psi^*\}_x)(\{\psi\}_y^T \{\psi^*\}_y)} \quad (3.4.1)$$

where * indicates the complex conjugate. The MAC was developed to take advantage of the orthogonality properties of mode shapes and takes on a value between 0 and 1. A value of 1 indicates that the modes are identical up to some scalar normalisation factor and a value of 0 indicates that the two modal vectors being compared are orthogonal to each other. [10]

The matrix of MAC values in Table 3.3 compares all the modes for the 8-DOF system corresponding to the undamaged system and the system that has the 50% spring stiffness reduction introduced into spring 3 and 5, i.e., State #4. The values in this matrix's diagonal provide a comparison of a certain mode before and after damage. If there was no change in the mode shapes as a result of the damage, the MAC values matrix should be the identity matrix. Although there is a difference between the undamaged and damaged mode forms when looking at the diagonal terms in the matrix, the percentage changes in these values are substantially smaller than the individual mode shape amplitude changes. [10]

Table 3.3. *Modal assurance criteria comparing the undamaged 8-DOF system modes to modes corresponding to State #4*

Mode	1	2	3	4	5	6	7	8
1	1.0000	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000
2	0.0001	0.9929	0.0001	0.0021	0.0001	0.0020	0.0001	0.0018
3	0.0001	0.0002	0.9386	0.0000	0.0557	0.0001	0.0027	0.0000
4	0.0001	0.0035	0.0002	0.9799	0.0018	0.0125	0.0000	0.0042
5	0.0000	0.0001	0.0533	0.0005	0.8764	0.0001	0.0698	0.0009
6	0.0000	0.0013	0.0000	0.0032	0.0003	0.7851	0.0022	0.2060
7	0.0000	0.0000	0.0110	0.0003	0.0650	0.0000	0.9138	0.0123
8	0.0000	0.0036	0.0000	0.0122	0.0000	0.2017	0.0097	0.7743

3.4.3. Mode Shape Curvature

Mode shape derivatives, such as curvature, are an alternative to employing mode shapes to get spatial information regarding damage for objects that exhibit bending behaviour [33]. Any discontinuities in the mode shape induced by localised damage are amplified by the derivative process. Consider a beam cross-section at position x subjected to a bending moment $M(x)$, given the mode shapes obtained from the undamaged and damaged structure. The curvature at location x , $v''(x)$ is approximated by

$$v''(x) \approx \frac{M(x)}{EI} \quad (3.4.2)$$

where E is the modulus of elasticity and I is the cross-sectional moment of inertia. The curvature is directly proportional to the inverse of the flexural stiffness, EI, as shown by this equation. As a result, a decrease in stiffness associated with damage leads to an increase in curvature for a given moment applied to the damaged and undamaged structure [10]. The

mode shape curvature for all modes except 1st and 8th mode of 8-DOF system are shown in Fig. 3.11.

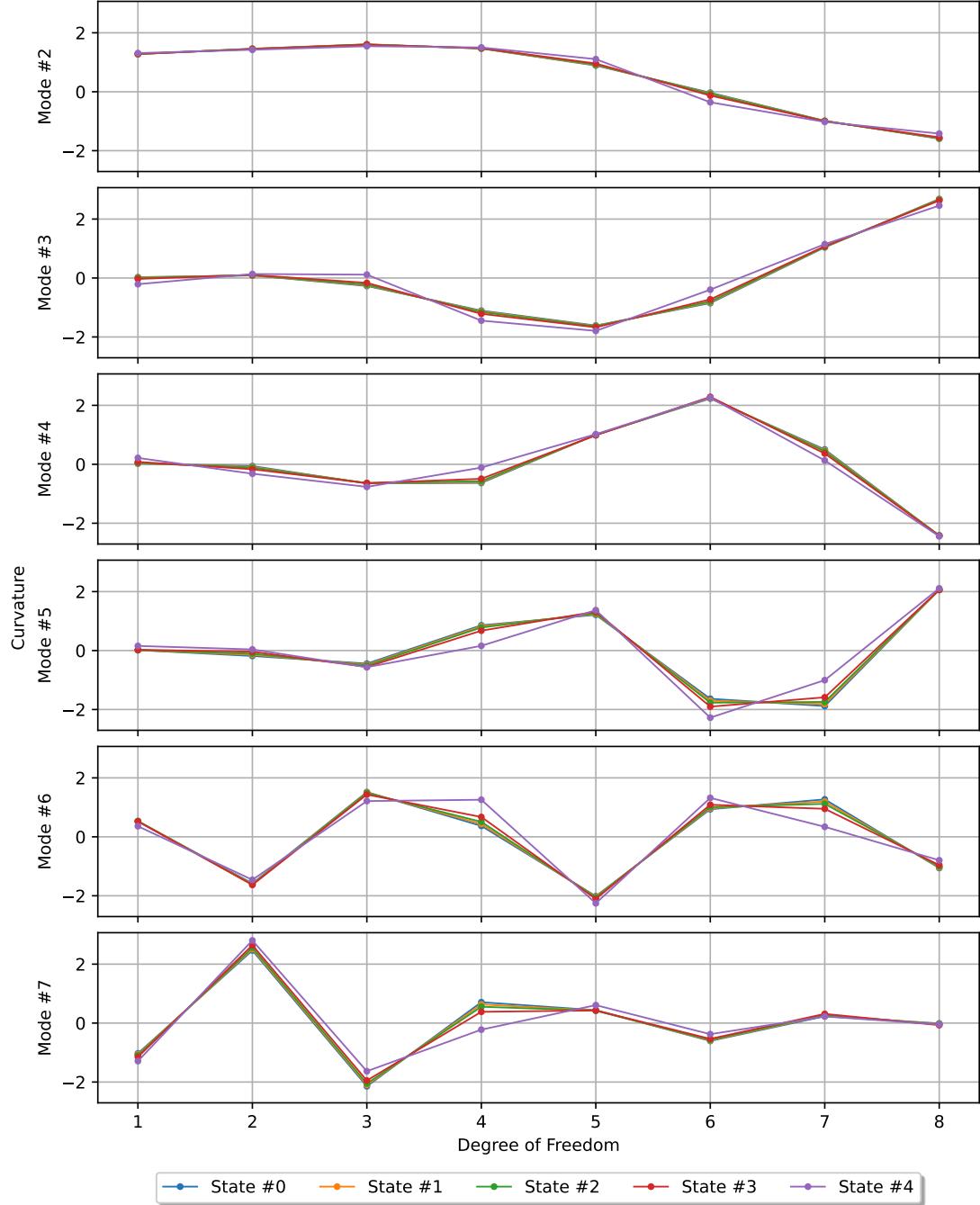


Fig 3.11. Mode shape curvature calculated for 8-DOF system for each structural state

3.4.4. Modal Strain Energy

The underlying premise is that when a structure is damaged, the distribution of strain energy stored in it changes in a more apparent way in the damaged sections. When a structural member undergoes a change in stiffness as a result of damage, which is nearly invariably a reduction, it can no longer absorb the same amount of energy as it could before the damage. When compared to the undamaged structure, this impact causes a departure from the original strain energy distribution [10]. As a result, damage can be detected and located by comparing

the strain energy distributions of the undamaged and damaged structures. The strain energy stored in the spring when the structures deforms in one of its mode shapes is defined as

$$U = \frac{1}{2}k(\Delta x)^2 \quad (3.4.3)$$

where the values of Δx are the changes in length of the springs from its undeformed state [34]. These values can be obtained from the mode shape vector by noting that the mode shape vector gives the change in the position of the DOF relative to its initial position. Fig. 3.12 shows the changes in the modal strain energies in each spring for all mode shapes of 8-DOF system correspond to all structural state conditions described in section 2.3.1.

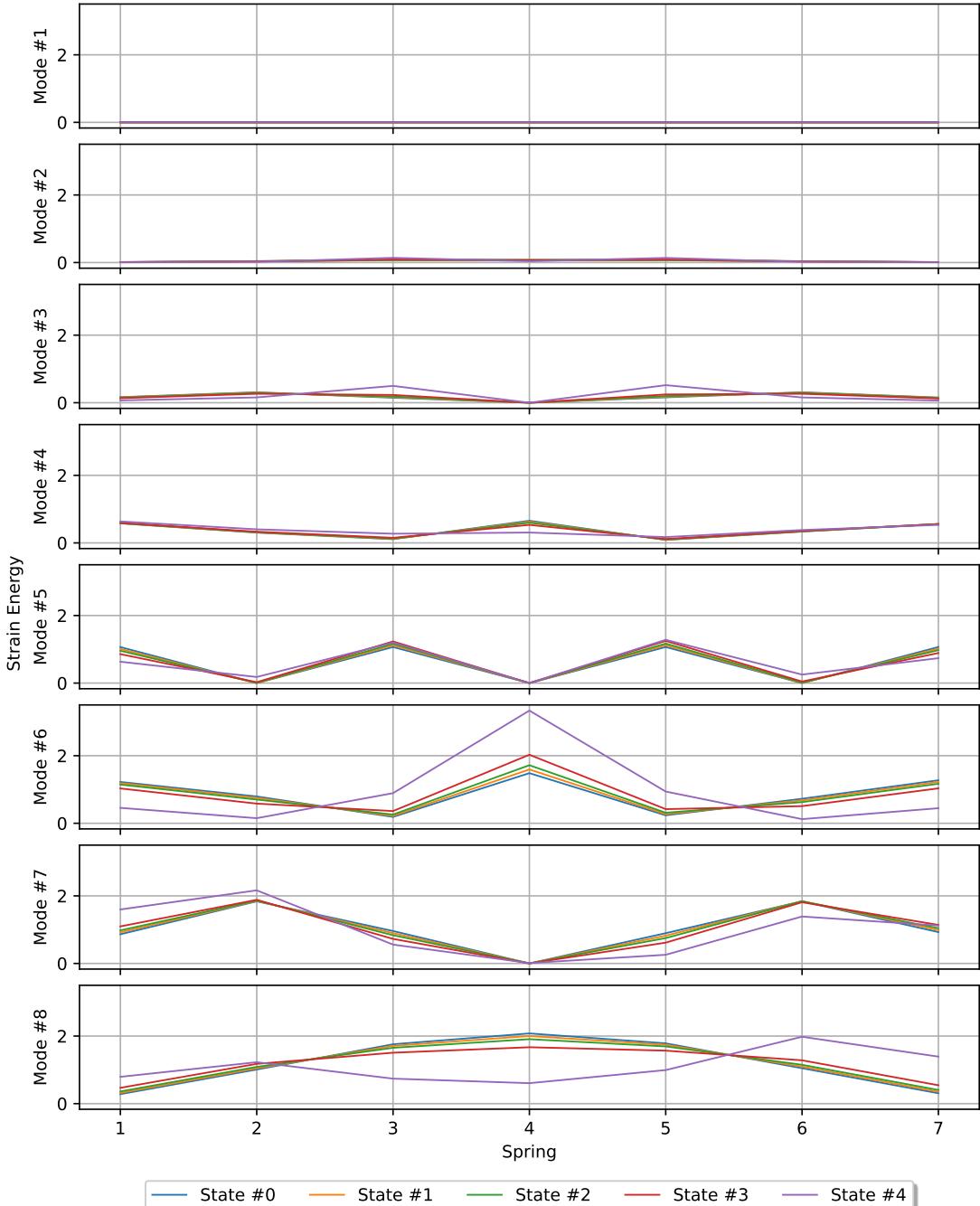


Fig 3.12. Modal strain energy calculated for 8-DOF system for each structural state

4. FEATURE SELECTION

Feature extraction procedures frequently entail numerous steps to exclude negative impacts in order to build superior features. Despite the fact that several features are available for various SHM analysis tasks, it is still difficult to assign priority for a specific problem, necessitating the use of a feature selection method. Traditional applications pick useful features by experience or sensitivity analysis, which is subjective and impractical given the enormous number of features accessible [35]. By training an ensemble of classifiers using randomly generated features for each of the classifiers, the ensemble learning framework provides an effective method for feature selection [13]. A few approaches for feature selection are discussed in this section.

The features dataset consists of the basic statistical features (peak amplitude, root-mean-square, standard deviation, skewness, kurtosis) and standard deviation ratio of residual errors derived from ARX model.

4.1. Filter Methods

Filter methods consider the relationship between features and the target variable to compute the importance of features. In this report various filter methods are used for feature selections.

4.1.1. *F-Statistic or F-Test*

An F-statistic, or F-test, is a class of statistical tests that calculate the ratio between variances values, such as the variance from two different samples or the explained and unexplained variance by a statistical test, like ANOVA. The ANOVA method is a type of F-statistic referred to here as an ANOVA F-test. ANOVA is an acronym for “analysis of variance” and is a parametric statistical hypothesis test for determining whether the means from two or more samples of data (often three or more) come from the same distribution or not.

Importantly, ANOVA is used when one variable is numeric and one is categorical, such as numerical input variables and a classification target variable in a classification task. The results of this test can be used for feature selection where those features that are independent of the target variable can be removed from the dataset.

Statistical features were used as independent variables and state of the structure, i.e., either damaged or undamaged, is used as classification target variable or dependent variable. The score for each statistical feature is shown in Fig. 4.1.

A highly correlated feature is given higher score and less correlated features are given lower score. Because root-mean-square and standard deviation scored higher than the other features, they can be chosen, but kurtosis and skewness can be ignored because to their poor scores.

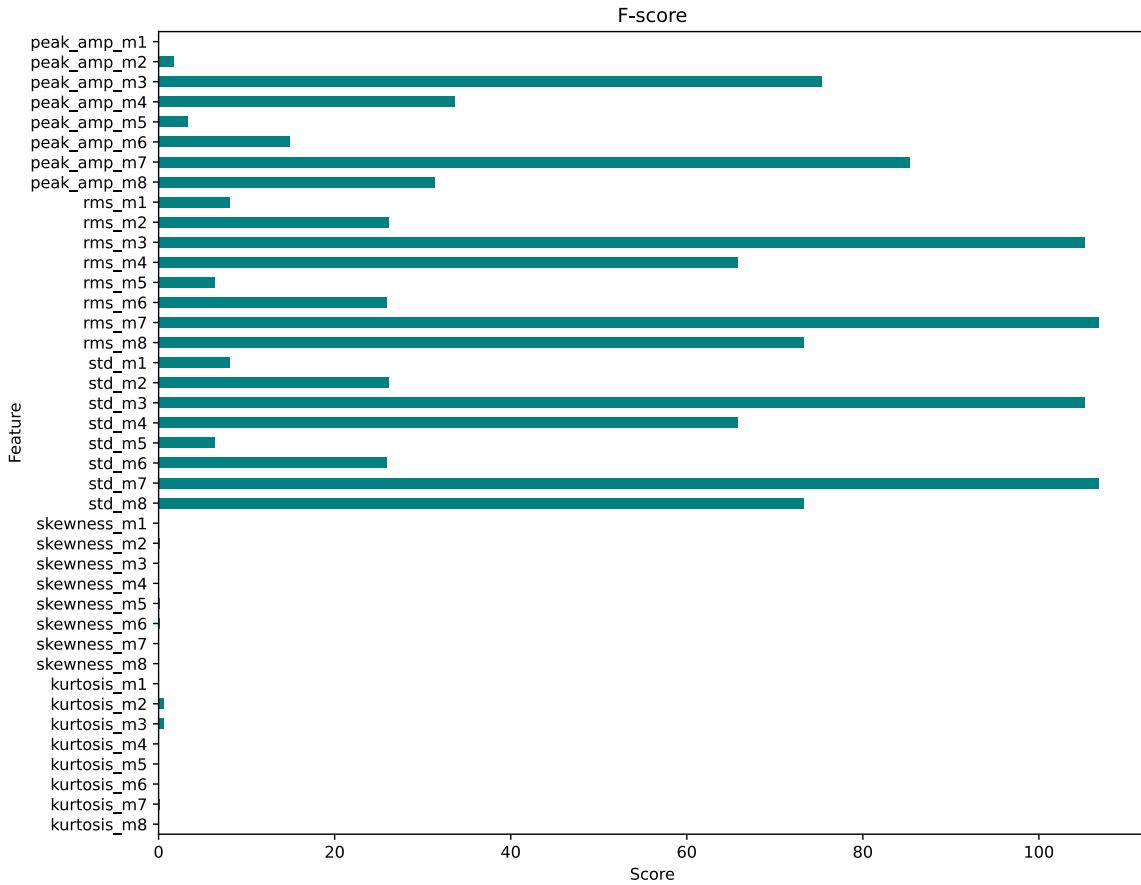


Fig. 4.1. *F-score for each statistical feature*

4.1.2. Chi-Squared Test

Mathematically, a Chi-Square test is done on two distributions to determine the level of similarity of their respective variances. In its null hypothesis, it assumes that the given distributions are independent. This test thus can be used to determine the best features for a given dataset by determining the features on which the output class label is most dependent on. For each feature in the dataset, the χ^2 is calculated and then ordered in descending order according to the χ^2 value. The higher the value of χ^2 , the more dependent the output label is on the feature and higher the importance the feature has on determining the output. The value of χ^2 is given by the expression

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad (4.1.1)$$

where O is observed frequency and E is expected frequency. Chi-Square test is done for STD ratio features extracted from ARX model and shown in Fig. 4.2.

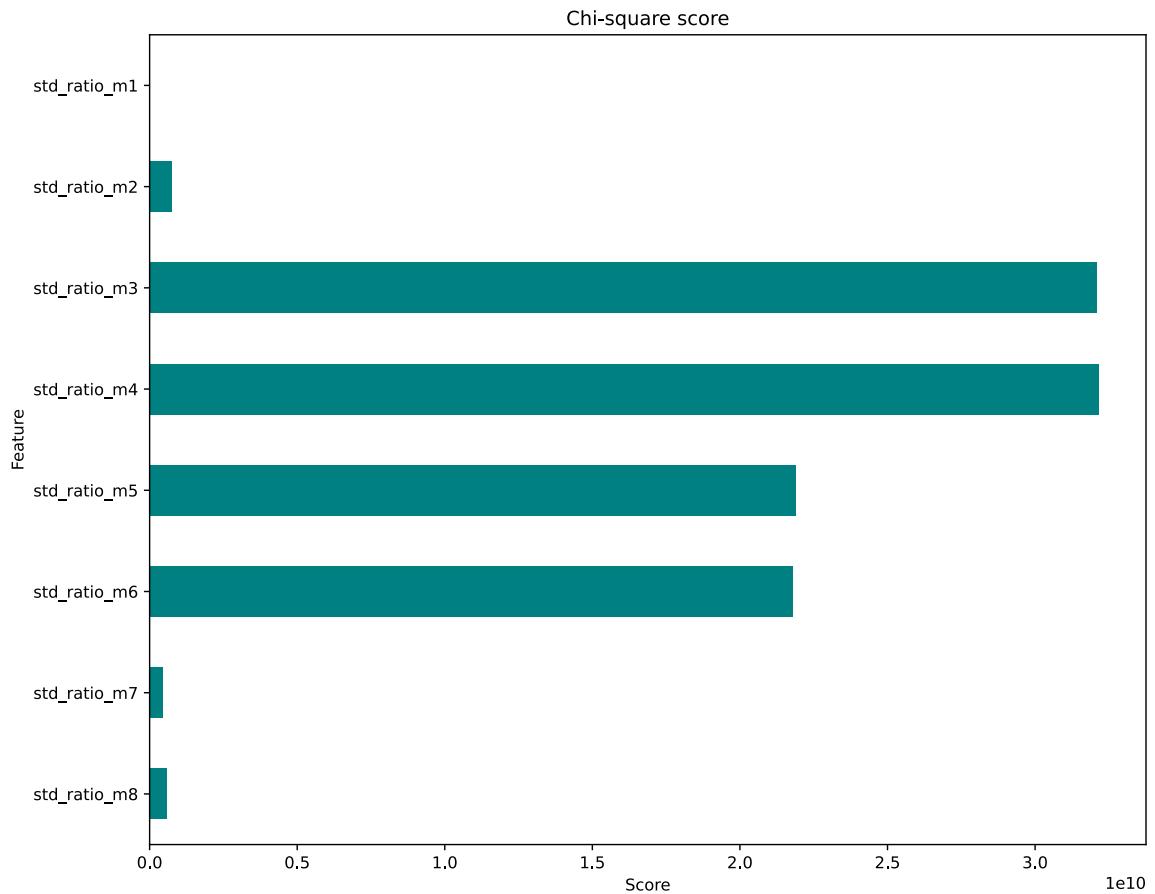


Fig. 4.2. Chi-squared score for each ARX model feature

4.1.3. Feature Importance Method

Feature importance refers to a class of techniques for assigning scores to input features to a predictive model that indicates the relative importance of each feature when making a prediction. Feature importance scores can be calculated for problems that involve predicting a numerical value, called regression, and those problems that involve predicting a class label, called classification.

There are many types and sources of feature importance scores, although popular examples include statistical correlation scores, coefficients calculated as part of linear models, decision trees, and permutation importance scores.

Feature importance scores play an important role in a predictive modelling project, including providing insight into the data, insight into the model, and the basis for dimensionality reduction and feature selection that can improve the efficiency and effectiveness of a predictive model on the problem.

Feature importance scores for statistical features, ARX model features and combination of statistical and ARX model features are shown in Fig. 4.3, Fig. 4.4 and Fig. 4.5. It may be deduced that the ARX model's STD ratio of residual errors feature is more effective for damage detection classification than statistical features.

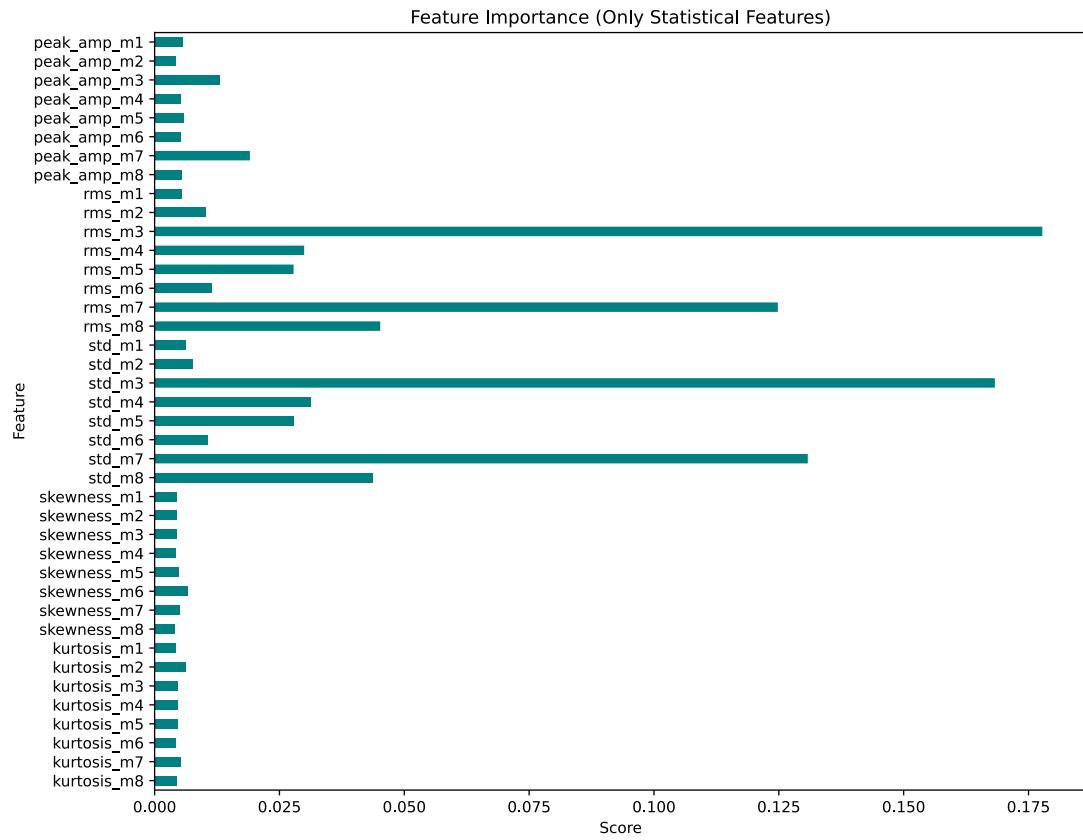


Fig. 4.3. Feature importance score for statistical features

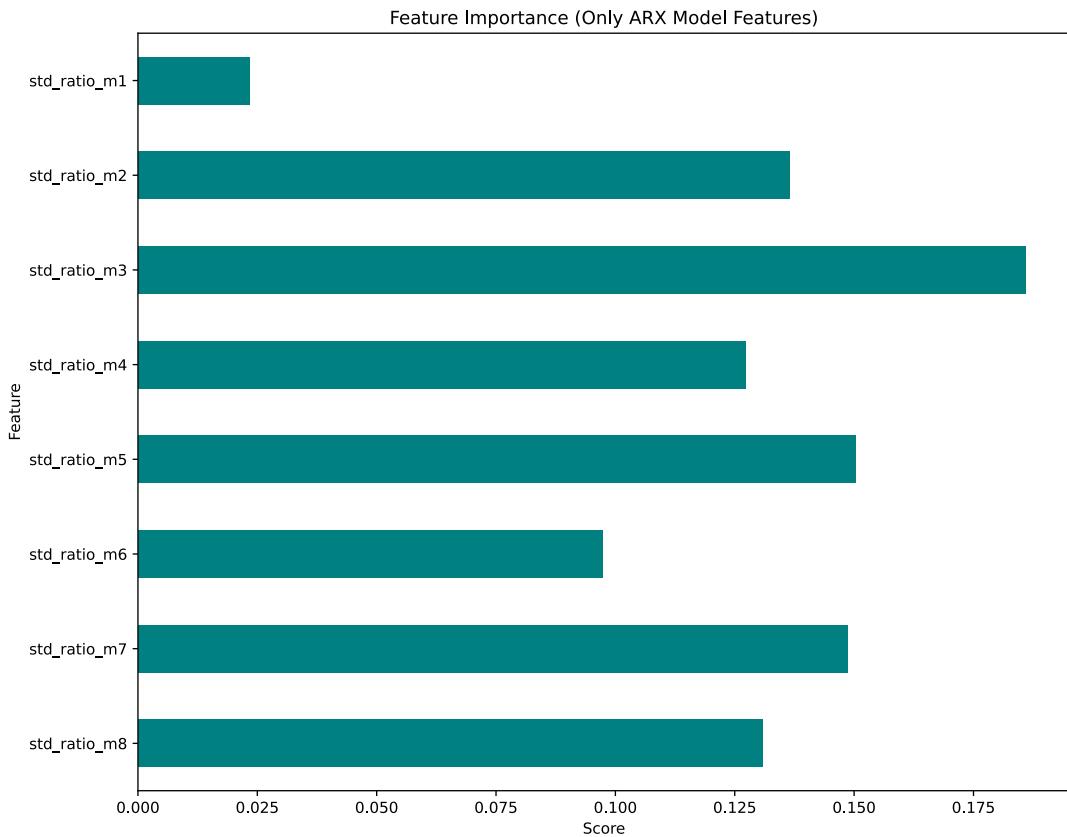


Fig. 4.4. Feature importance score for ARX model features

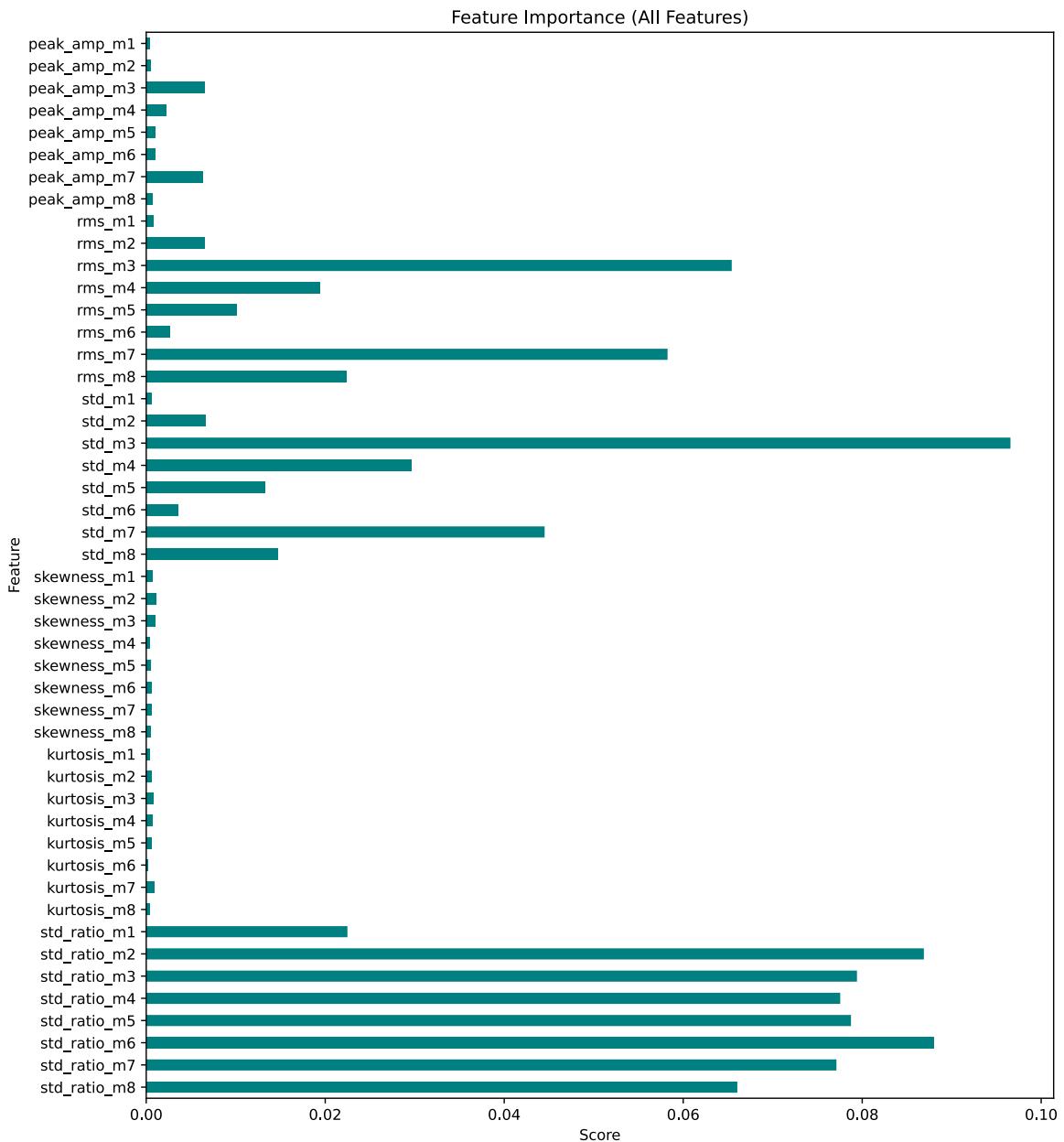


Fig. 4.5. Feature importance score for all features

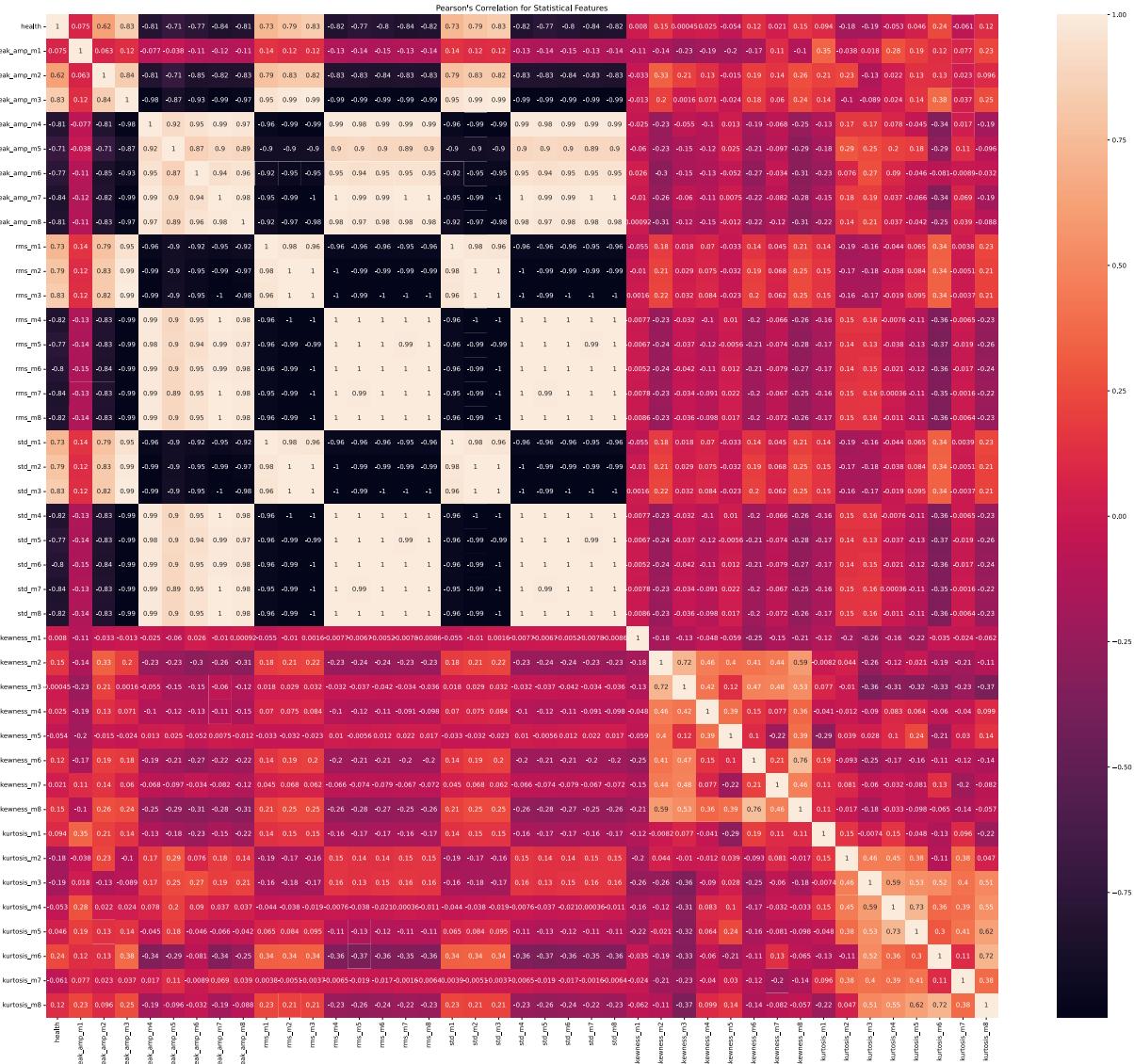
4.1.4. Pearson's Correlation

Pearson correlation is a number between -1 and 1 that indicates the extent to which two variables are linearly related. The Pearson correlation is also known as the “product moment correlation coefficient” (PMCC) or simply “correlation”. Pearson correlations are suitable only for metric variables. The correlation coefficient has values between -1 to 1. Pearson's Correlation method is used for finding the association between the continuous features and the class feature.

- A value closer to 0 implies weaker correlation (exact 0 implying no correlation)
- A value closer to 1 implies stronger positive correlation
- A value closer to -1 implies stronger negative correlation

$$R = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}} \quad (4.1.2)$$

Where, R is Pearson's correlation coefficient, \bar{x} and \bar{y} are mean of x and y variables.



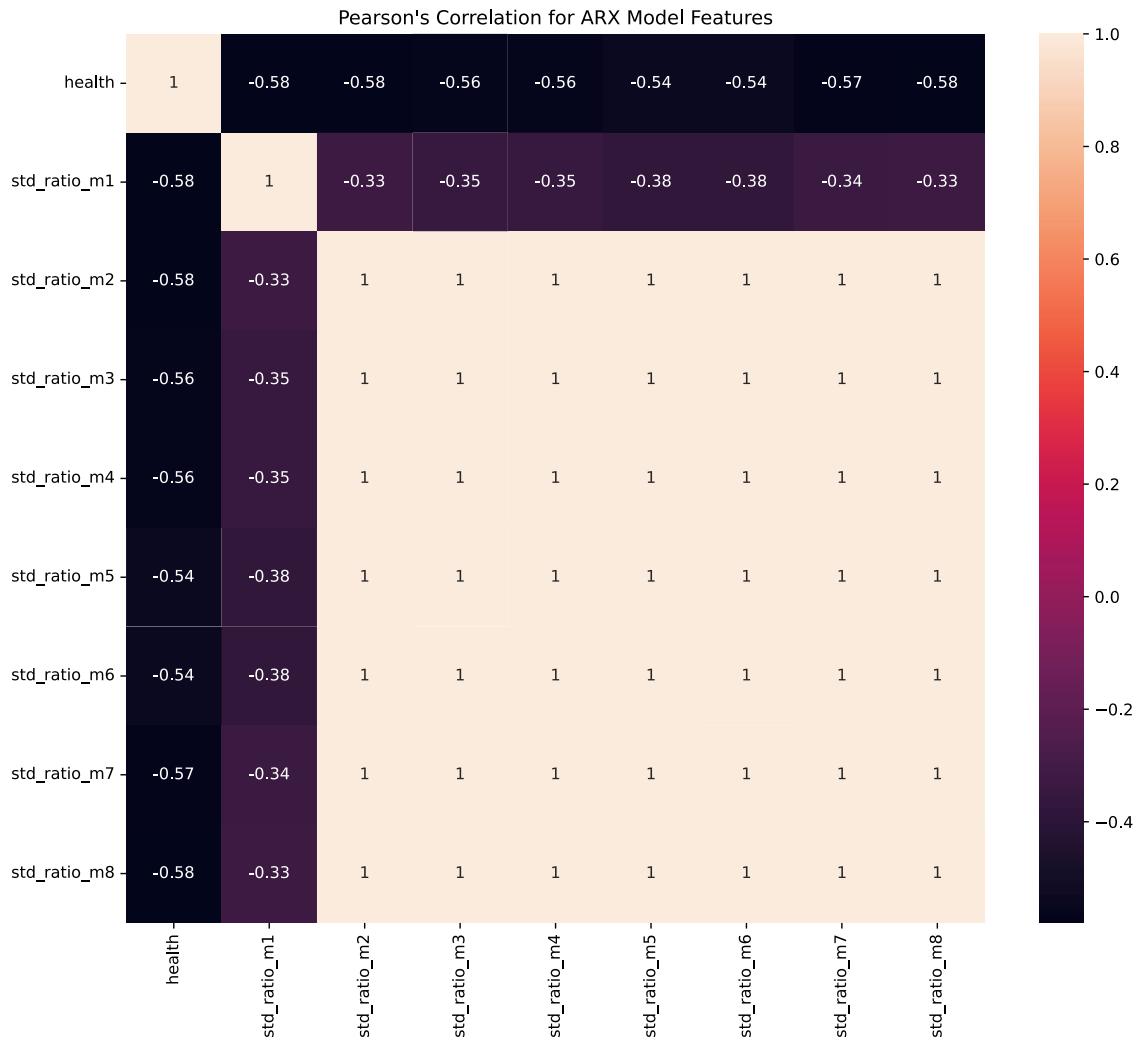


Fig. 4.7. Pearson's correlation coefficients for ARX model features

4.1.5. Mutual Information Technique

Mutual information from the field of information theory is the application of information gain to feature selection. Mutual information is calculated between two variables and measures the reduction in uncertainty for one variable given a known value of the other variable.

Mutual information is straightforward when considering the distribution of two discrete (categorical or ordinal) variables, such as categorical input and categorical output data. Nevertheless, it can be adapted for use with numerical input and categorical output. The mutual information between two random variables X and Y can be stated formally as follows

$$I(X;Y) = H(X) - H(X|Y) \quad (4.1.3)$$

where, $I(X;Y)$ is the mutual information for X and Y , $H(X)$ is the entropy for X and $H(X|Y)$ is the conditional entropy for X given Y . The result has the units of bits.

Mutual information scores for statistical features, ARX model features and combination of statistical and ARX model features are shown in Fig. 4.8, Fig. 4.9 and Fig. 4.10.

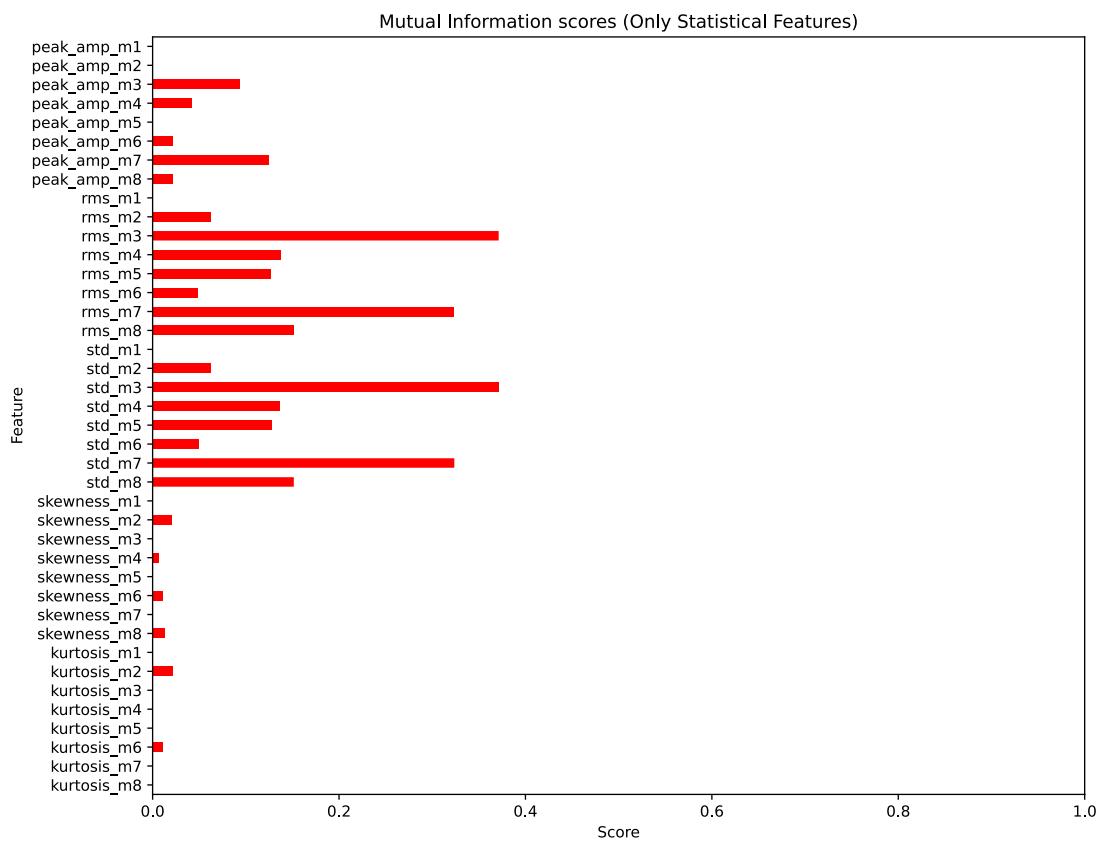


Fig. 4.8. Mutual information score for statistical features

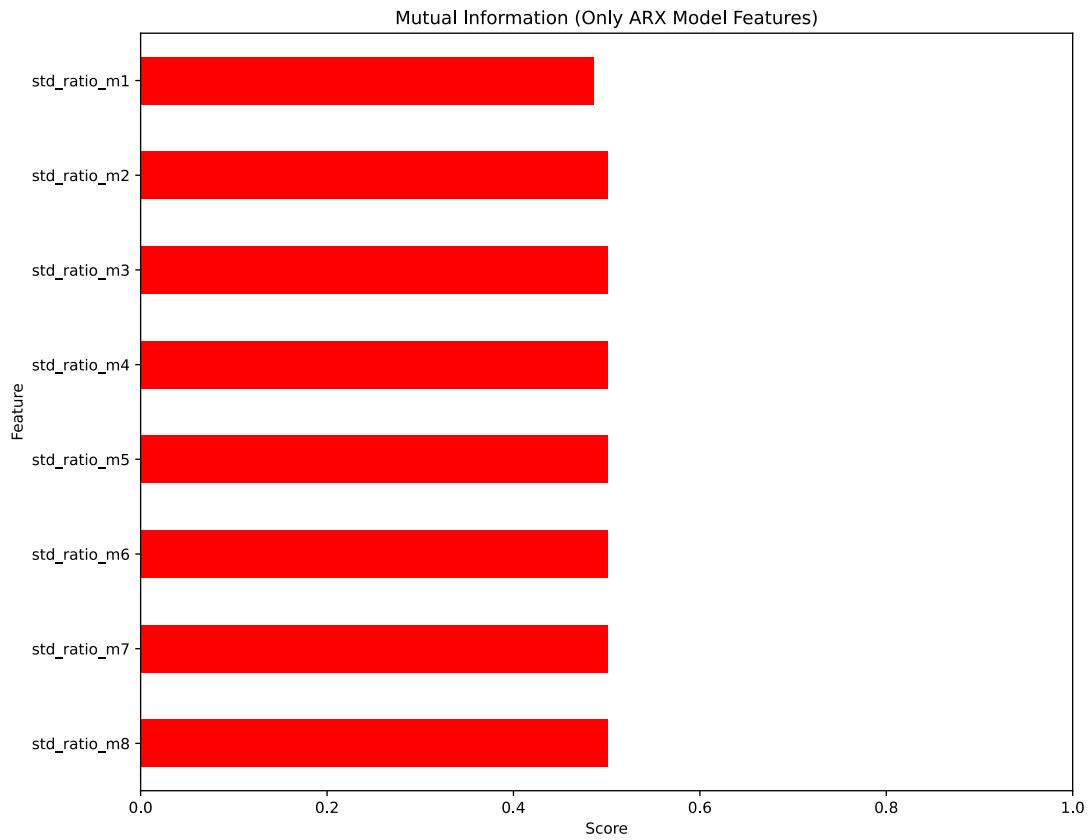


Fig. 4.9. Mutual information score for ARX model features

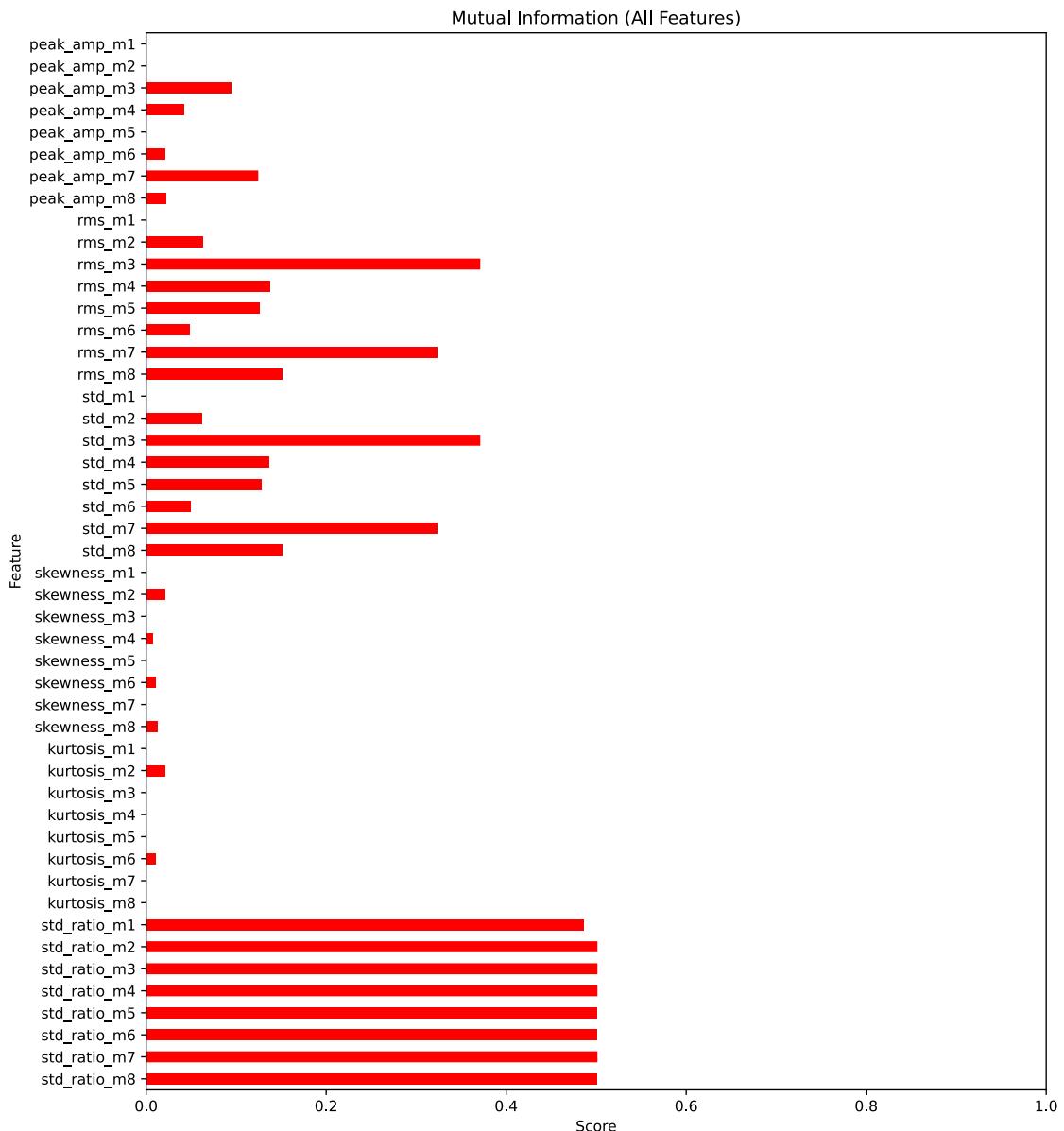


Fig. 4.10. Mutual information score for all features

As a result of the preceding discussion, the following features might be listed in decreasing order of their efficacy for structural damage classification: STD ratio of residual errors (ARX Model) > Root-mean-square \approx Standard deviation > Peak amplitude \approx Skewness \approx Kurtosis.

5. STRUCTURAL DAMAGE CLASSIFICATION USING SUPPORT-VECTOR MACHINES (SVM)

5.1. Introduction to Support-Vector Machines (SVM)

In machine learning, support-vector machines (SVM) are supervised learning models with associated learning algorithms that analyse data for classification and regression analysis. Classifying data is a common task in machine learning. Suppose some given data points each belong to one of two classes, and the goal is to decide which class a new data point will be in. In the case of support-vector machines, a data point is viewed as a D -dimensional vector, and the aim is to know whether such points can be separated with a $(D-1)$ -dimensional hyperplane. This is called a linear classifier. [36]

5.1.1. Hyperplane

Hyperplanes can be considered decision boundaries that classify data points into their respective classes in a multi-dimensional space. Data points falling on either side of the hyperplane can be attributed to different classes. In D -dimensional space, the hyperplane would always be $D-1$ operator. Equation of a linear classifier hyperplane can be given as

$$w^T X + b = 0 \quad (5.1.1)$$

where, w is the normal vector to the hyperplane, X is a data point vector.

5.1.2. SVM Classifier

The hypothesis function h is defined as

$$h(X) = \begin{cases} +1 & \text{if } w^T X + b \geq 0 \\ -1 & \text{if } w^T X + b < 0 \end{cases} \quad (5.1.2)$$

The point above or on the hyperplane will be classified as class +1, and the point below the hyperplane will be classified as class -1.

Hyperplanes H_1 and H_2 passing through the support vectors of +1 and -1 class respectively are

$$H_1 : w^T X + b = 1 \quad (5.1.3)$$

$$H_2 : w^T X + b = -1 \quad (5.1.4)$$

For any point X in the training data, let the corresponding label be Y such that $Y \in \{-1, 1\}$.

The product of a predicted and actual label would be greater than 0 (zero) on correct prediction, otherwise less than zero.

$$Y_i(w^T X_i + b) = \begin{cases} \geq 0 & \text{if correct} \\ < 0 & \text{if incorrect} \end{cases} \quad (5.1.5)$$

5.1.3. Optimal Hyperplane and Margin

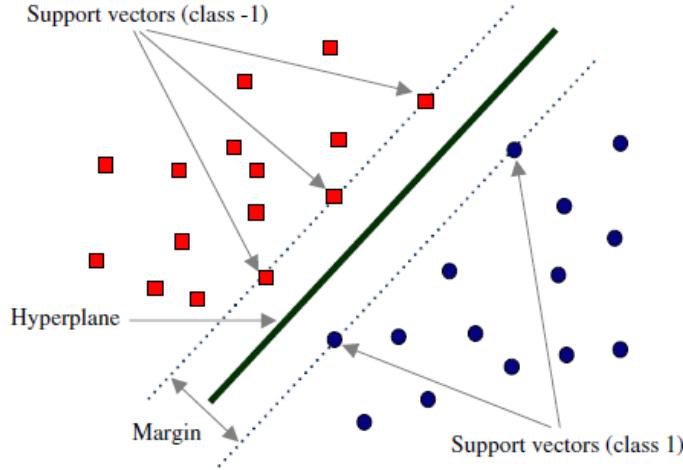


Fig. 5.1. Hyperplane, margin and support-vectors

There are many hyperplanes that might classify the data. One reasonable choice as the best hyperplane is the one that represents the largest separation, or margin, between the two classes. So, the hyperplane is chosen such that the distance from it to the nearest data point, which is called support-vector, on each side is maximized. If such a hyperplane exists, it is known as the maximum-margin hyperplane and the linear classifier it defines is known as a maximum-margin classifier; or equivalently, the perceptron of optimal stability [36]. The optimal hyperplane is one which divides the data points to maximum extent. The optimal hyperplane lies in the middle of hyperplanes H_1 and H_2 . The unit vector along w is $\frac{w}{\|w\|}$

where $\|w\|$ is the norm of w . In SVM our goal is to choose an optimal hyperplane which maximizes the margin.

Let the distance of either of the support vector from optimal hyperplane be k and X be a point on optimal hyperplane. The location of the support-vectors can be given as $X + k \frac{w}{\|w\|}$.

$$w^T(X \pm k \frac{w}{\|w\|}) + b = \pm 1$$

$$w^T X \pm k \frac{w^T w}{\|w\|} + b = \pm 1$$

$$\because w^T X + b = 0, \text{ as } X \text{ is on the optimal hyperplane}$$

$$k \frac{w^T w}{\|w\|} = 1$$

$$k \frac{\|w\|^2}{\|w\|} = 1$$

$$k = \frac{1}{\|w\|} \quad (5.1.6)$$

Since the optimal hyperplane lies in the middle, the margin width M can be given as

$$M = 2k = \frac{2}{\|w\|} \quad (5.1.7)$$

Now, to get the maximum margin width M ,

$$\max(M) = \max\left(\frac{2}{\|w\|}\right)$$

or

$$\min(\|w\|) \text{ such that } Y_i(w^T X_i + b) \geq 1 \quad (5.1.8)$$

This is a constrained minimization problem and will always lead to global minimum.

5.1.4. Tuning Parameters

Kernel

Kernel function is a method used to take data as input and transform into the required form of processing data. “Kernel” is used due to set of mathematical functions used in support-vector machine provides the window to manipulate the data. So, kernel function generally transforms the training set of data so that a non-linear decision surface is able to transformed to a linear equation in a higher number of dimension spaces. Some examples of kernel functions are linear, polynomial, sigmoid, gaussian or radial basis function (RBF). In this report, linear kernel function is used for structural damage classification problem.

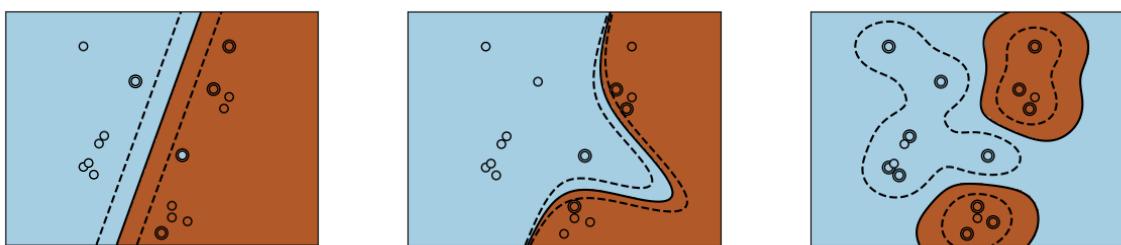


Fig. 5.2. Linear, polynomial and RBF kernel functions

Gamma

The gamma parameter defines how far the influence of a single training example reaches, with low values meaning ‘far’ and high values meaning ‘close’. In other words, with low gamma, points far away from plausible separation line are considered in calculation for the separation line whereas high gamma means the points close to plausible line are considered in calculation.

Regularization

The C parameter trades off correct classification of training examples against maximization of the decision function's margin. For larger values of C, a smaller margin will be accepted if the decision function is better at classifying all training points correctly. A lower C will encourage a larger margin, therefore a simpler decision function, at the cost of training accuracy. In other words, C behaves as a regularization parameter in the SVM. The C parameters defines the amount of violation of the margin allowed. A C=0 is no violation and results in the inflexible maximal-margin classifier described above. The larger the value of C the more violations of the hyperplane are permitted.

- The smaller the value of C, the more sensitive the algorithm is to the training data (higher variance and lower bias).
- The larger the value of C, the less sensitive the algorithm is to the training data (lower variance and higher bias).

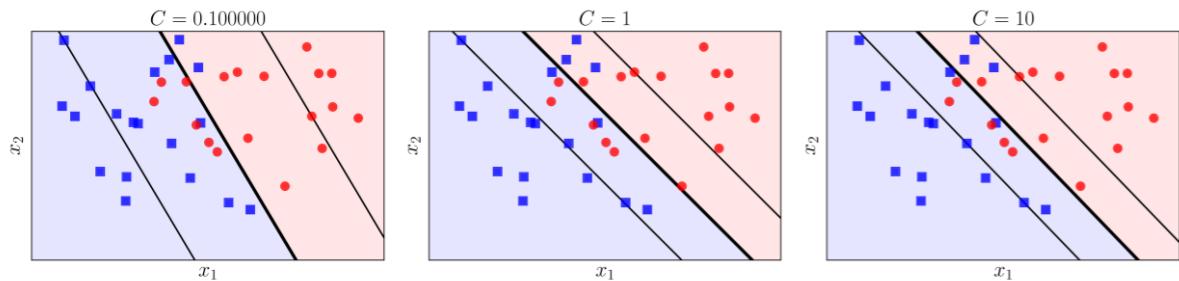


Fig. 5.3. Decrease in soft margin with increase in C value

5.2. Application of SVMs for Structural Damage Classification

For the purpose of structural damage classification using SVMs, the processed dataset is split into training and testing datasets in the ratio 70%:30% respectively. The model is trained on the training dataset and the accuracy of the model predictions on the test dataset is calculated. State #0 is labelled as “Damaged” and all other states, i.e., State #1, State #2, State #3 and State #4 are labelled as “Undamaged”. SVMs are applied for damage classification using feature datasets consisting all statistical features, selected statistical features, ARX model features individually.

5.2.1. Structural Damage Classification using all Statistical Features

SVM is applied for damage classification using the dataset consisting all statistical features described in previous chapters. Fig. 5.4 shows the variation in accuracy by changing the regularization parameter for SVM classifier. As the C value increases, the soft margin width decreases resulting in more accurately separating hyperplane but with a low margin. Based on this, C = 650 is chosen to train the SVM model which has performed reasonably well on testing dataset with an accuracy score of around 99%. Fig. 5.5 shows the heatmap of confusion matrix of SVM predictions on test dataset. The model has given one false positive result, i.e., classified one data point as damaged while its true label is undamaged, and one false negative result, i.e., classified one data point as undamaged while its true label is damaged. Table 5.1 is the classification report of SVM prediction on test dataset which contains precision, recall, f1-score, etc. that provides better insights into the prediction as compared to accuracy performance metrics.

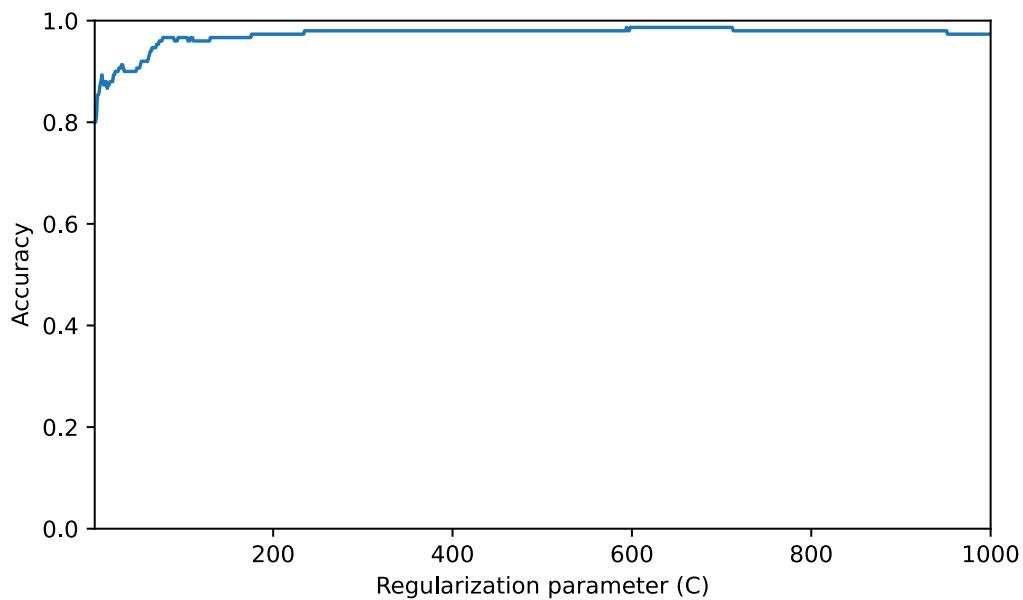


Fig. 5.4. Regularization parameter (C) vs Accuracy of SVM using all statistical features

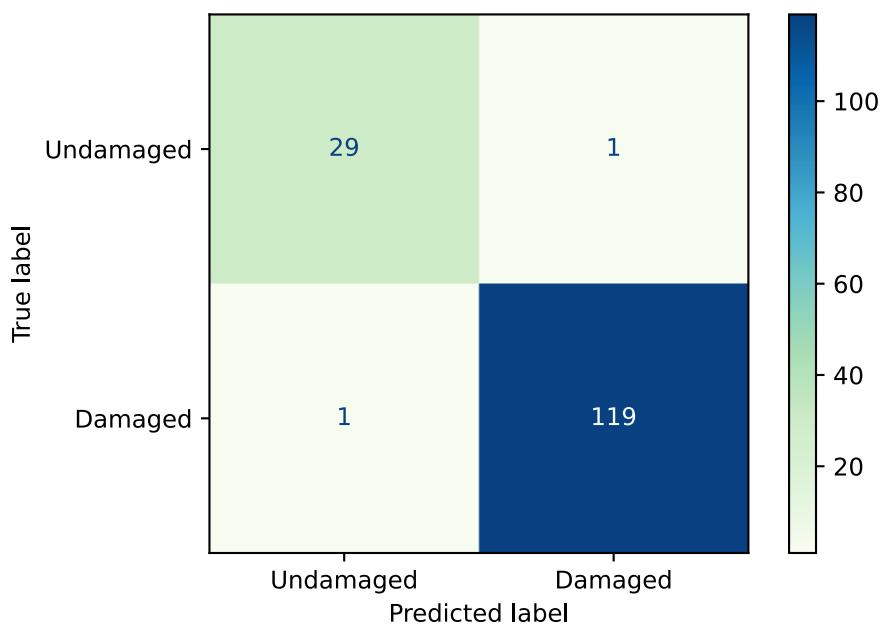


Fig. 5.5. Confusion matrix heatmap of SVM predictions on test dataset using all statistical features; Regularization parameter (C) = 650

Table 5.1. Classification report of SVM predictions on test dataset using all statistical features; Regularization parameter (C) = 650

	Precision	Recall	F1-Score	Support
Undamaged	0.97	0.97	0.97	30
Damaged	0.99	0.99	0.99	120
Accuracy			0.99	150
Macro avg.	0.98	0.98	0.98	150
Weighted avg.	0.99	0.99	0.99	150

5.2.2. Structural Damage Classification using selected Statistical Features

The dataset consisting only of selected statistical features, i.e., root-mean-square and standard deviation, using feature selection techniques described in previous chapter is used for classification using SVMs. Fig. 5.6 shows the variation in accuracy by changing the regularization parameter for SVM classifier. Better accuracy can be scored for lower values of C compared to when all statistical features were used. $C = 40$ is chosen to train the SVM model. An accuracy score of around 99% is obtained on test dataset. Fig. 5.7 shows the heatmap of confusion matrix of SVM predictions on test dataset. The model has given one false negative result. Table 5.2 is the classification report of SVM prediction on test dataset.

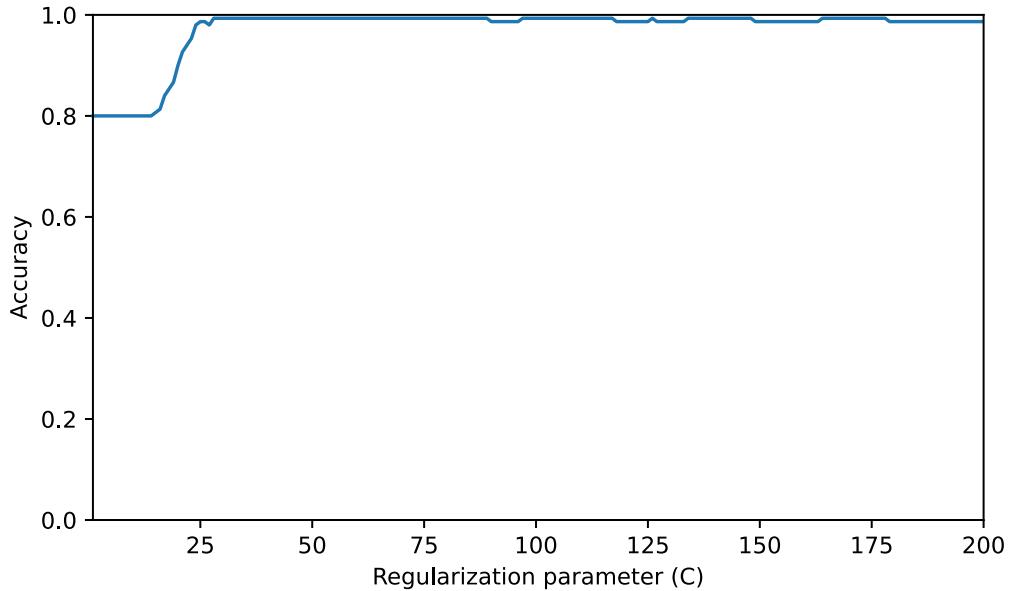


Fig. 5.6. Regularization parameter (C) vs Accuracy of SVM using selected statistical features

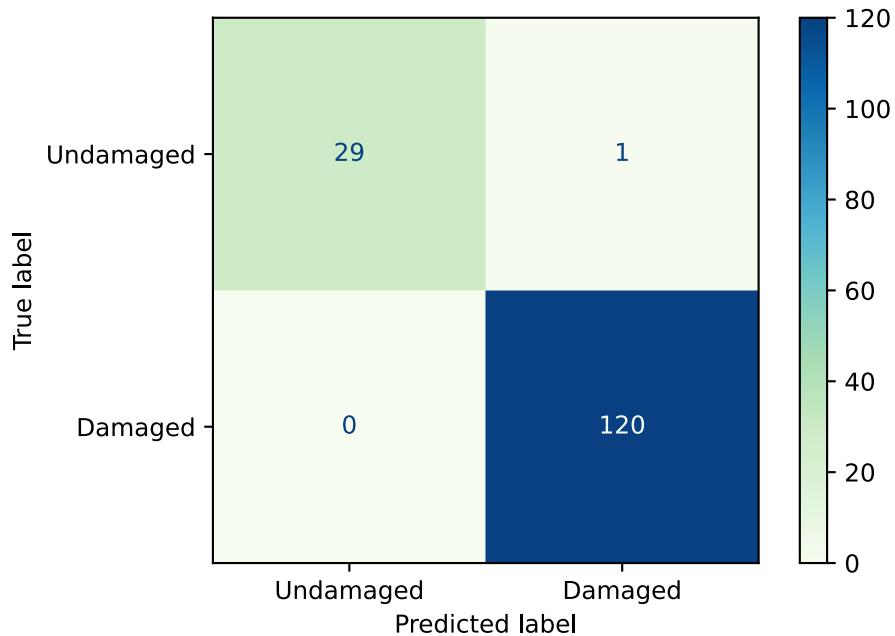


Fig. 5.7. Confusion matrix heatmap of SVM predictions on test dataset using selected statistical features; Regularization parameter (C) = 40

Table 5.2. Classification report of SVM predictions on test dataset using selected statistical features; Regularization parameter (C) = 40

	Precision	Recall	F1-Score	Support
Undamaged	1.00	0.97	0.98	30
Damaged	0.99	1.00	1.00	120
Accuracy			0.99	150
Macro avg.	1.00	0.98	0.99	150
Weighted avg.	0.99	0.99	0.99	150

5.2.3. Structural Damage Classification using ARX Model Features

The ratio of standard deviation of residual errors derived from ARX model are used as features for damage detection using SVMs. Fig. 5.8 shows accuracy vs regularization parameter plot. Very high accuracy scores are obtained for extremely low values of C of order 10^{-16} compared to when statistical features were used. To train the SVM model on training dataset, $C = 1.0 \times 10^{-15}$ is chosen. The results are remarkably accurate with an accuracy score of 100% which means all the data points on the test dataset are classified correctly as damaged or undamaged.

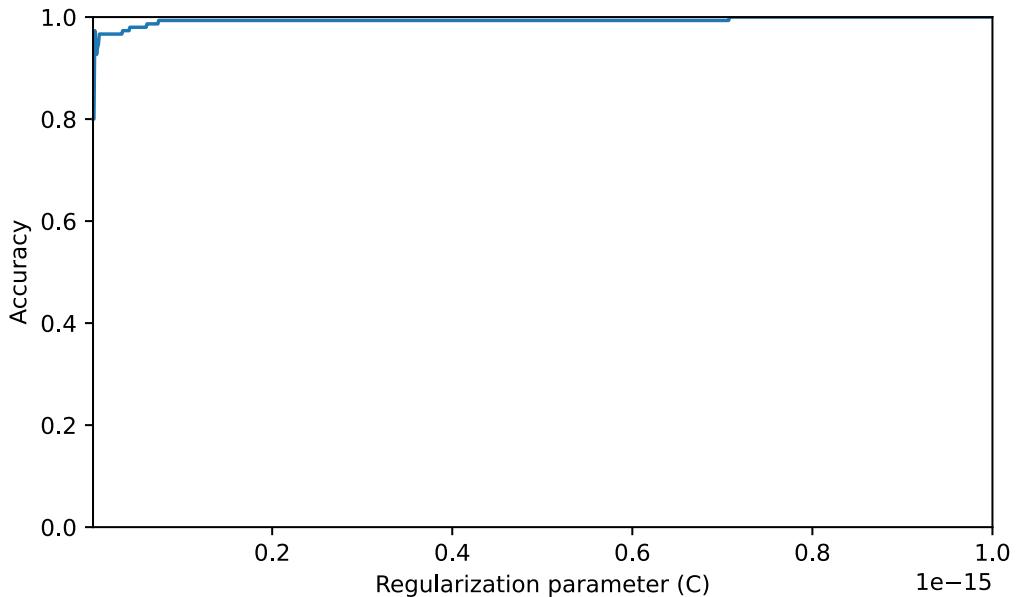


Fig. 5.8. Regularization parameter (C) vs Accuracy of SVM using ARX model features

Table 5.3. Classification report of SVM predictions on test dataset using ARX model features; Regularization parameter (C) = 1.0×10^{-15}

	Precision	Recall	F1-Score	Support
Undamaged	1.00	1.00	1.00	30
Damaged	1.00	1.00	1.00	120
Accuracy			1.00	150
Macro avg.	1.00	1.00	1.00	150
Weighted avg.	1.00	1.00	1.00	150

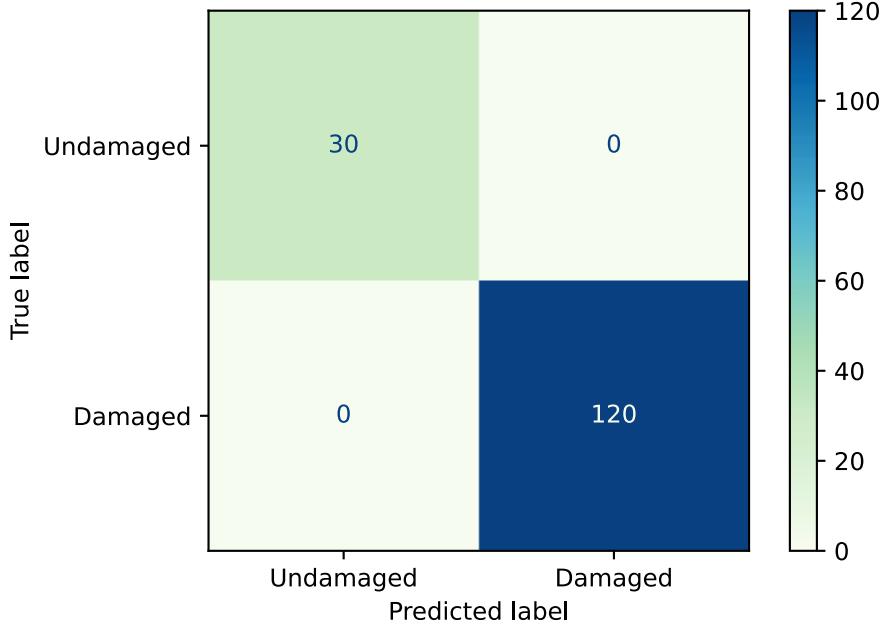


Fig. 5.9. Confusion matrix heatmap of SVM predictions on test dataset using ARX model features; Regularization parameter (C) = 1.0×10^{-15}

It's worth noting that when the effectiveness of damage-sensitive features improves, the binary classification of structural health condition, i.e., damaged or undamaged, becomes more precise. As illustrated in Fig. 5.10, where the ARX model's STD ratio at 3rd and 5th mass location feature is utilised to produce a scatter map of data points labelled damaged and undamaged, notice that undamaged labelled data points converge to point (1, 1) and both classes are separated far apart. As a result, by significantly lowering the value of the regularisation parameter, a wider margin can be gained. When ARX model features are employed instead of statistical features, a very small value of C is needed for higher accuracy, as shown in this section, because the former is a more effective damage-sensitive feature.

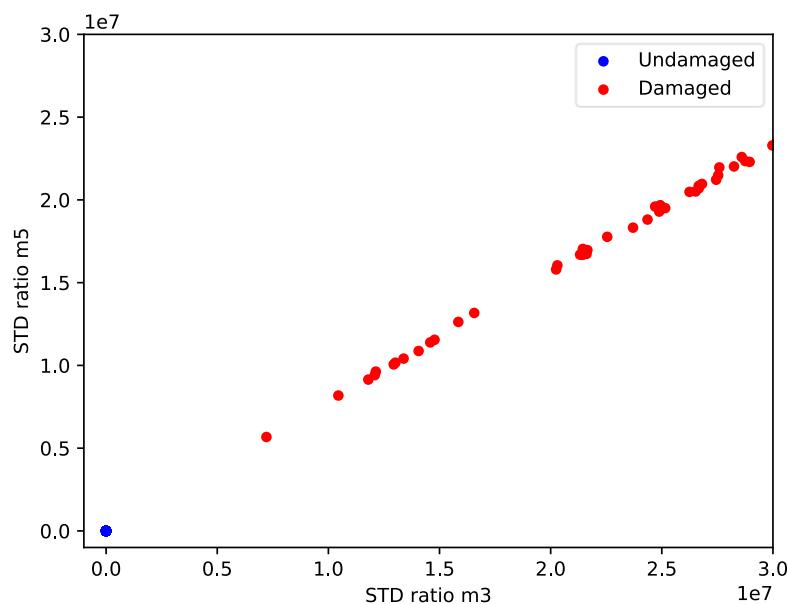


Fig. 5.10. Scatter plot using STD ratio at 3rd and 5th mass location from ARX model

6. CONCLUSION

The purpose of this research report was to apply the statistical pattern recognition paradigm for SHM to data collected from an 8-DOF system's vibration response and to devise a method for structural damage detection, i.e., to classify or distinguish between damaged and undamaged states, using machine learning models and simulated data. A thorough review of literature was done and the applications of ML methods to SHM was explored.

Vibration response, i.e., force-, acceleration-, velocity- and displacement-time histories (time series or sample records), for a variety of different structural state conditions of the structure subjected to white noise excitation were computed using computational simulations.

The acceleration-time history data was used to extract features based on signal statistics such as peak amplitude, mean, root-mean-square, standard deviation, variance, skewness, kurtosis, probability density function, and cumulative distribution function. These features tend to differ from those of an undamaged structure, indicating the existence of damage. Resonance frequencies, vibration mode forms, modal assurance criteria (MAC), mode shape curvatures, modal strain energies, and the deviation in which from the undamaged state signified the presence of damage were all extracted as well.

The main focus was given to time-series analysis which was done using autoregressive model with exogeneous input (ARX) which proved to be a potential damage-sensitive feature. The technique is based on linear dynamic equations and is expressed as an ARX model with acceleration response signals, which is a novel type of ARX model. The model coefficients are strongly related to the system's dynamic properties, allowing the creation of sensitive features for damage diagnosis. The model is also decoupled from the input excitation, which increases the method's potential robustness in real-world applications. The standard deviation (STD) of the residual error, which is the difference between the measured signals from any actual state of the system and the predicated signals from the ARX model established from a reference (undamaged) state, is found to be a damage-sensitive feature. The occurrence of damage can be detected using the standard deviation of the residual errors as a feature, and the location of damage can also be identified in an MDOF system because larger STD of the residual errors tend to occur near the actual damage locations, according to numerical simulation studies.

When the position of the response point selected as the ARX model input is close to the location of the damage, the behaviour of the proposed ARX model can be problematic. As a result, it is recommended that two independent runs of the method be performed in the diagnostic of an MDOF system, each employing two different "input" response locations. Examining the two sets of results allows for a correct detection of the damage location. It should be noted that, while the ARX model's standard deviation of residual error is sensitive to the existence of damage, it does not provide a precise indicator of the degree of damage. The tendency is similar for analogous circumstances; nevertheless, in order to quantify the degree of damage, more study will be required to expand the model's capabilities, most likely with a more appropriate damage feature. In addition, in order to deploy the model for

practical usage, the behaviour of the model in a noisy measurement environment must be examined.

Informative, discriminating, and independent features were ranked and selected from a pool of damage-sensitive features based on their effectiveness for damage detection. Filter based feature selection techniques such as mutual information, F-test, Chi-squared test, feature importance and Pearson's correlation were used and the following features were listed in decreasing order of their efficacy for structural damage classification: STD ratio of residual errors (ARX Model) >> Root-mean-square \approx Standard deviation > Peak amplitude > Skewness \approx Kurtosis.

Finally, SVM models were trained to classify structural condition states as damaged or undamaged using feature datasets consisting all statistical features, selected statistical features, ARX model features individually. An accuracy of 99% was obtained when all statistical features and selected statistical features were used to train the SVM model separately. However, the margin width was lower in case of all statistical features as compared to when selected statistical features were used. ARX model features were turned out to be most effective and useful damage sensitive feature out of all the features and the SVM model trained using these features were able to classify damaged and undamaged states with 100% accuracy. The SVM classifier had largest margin width in case of ARX model features as compared to statistical features.

A more extensive study is required to accurately identify the geometric location of the damage and quantify its severity. Future work will include application of the developed approaches on any real data rather than simulated data and the applicability, effectiveness, and efficiency of the approach will be studied through application to other structures like beams and trusses.

7. BIBLIOGRAPHY

- [1] F. J. Montáns, F. Chinesta, R. Gómez-Bombarelli, and J. N. Kutz, “Data-driven modeling and learning in science and engineering,” *Comptes Rendus - Mec.*, vol. 347, no. 11, pp. 845–855, 2019, doi: 10.1016/j.crme.2019.11.009.
- [2] A. Karpatne *et al.*, “Theory-guided Data Science: {A} New Paradigm for Scientific Discovery,” *CoRR*, vol. abs/1612.0, 2016, [Online]. Available: <http://arxiv.org/abs/1612.08544>.
- [3] H. Salehi and R. Burgueño, “Emerging artificial intelligence methods in structural engineering,” *Eng. Struct.*, vol. 171, no. May, pp. 170–189, 2018, doi: 10.1016/j.engstruct.2018.05.084.
- [4] N. J. Nilsson, “1 - Introduction,” in *Artificial Intelligence: A New Synthesis*, N. J. Nilsson, Ed. Oxford: Morgan Kaufmann, 1998, pp. 1–17.
- [5] L. Tagliaferri, “An Introduction to Machine Learning,” *Digital Ocean*, 2017. <https://www.digitalocean.com/community/tutorials/an-introduction-to-machine-learning>.
- [6] A. Samuel, “Some Studies in Machine Learning Using the Game of Checkers,” *IBM J. Res. Dev.*, vol. 3, pp. 210–229, 1959.
- [7] T. M. Mitchell, *Machine Learning*, 1st ed. USA: McGraw-Hill, Inc., 1997.
- [8] A. M. TURING, “I.—COMPUTING MACHINERY AND INTELLIGENCE,” *Mind*, vol. LIX, no. 236, pp. 433–460, 1950, doi: 10.1093/mind/LIX.236.433.
- [9] P. Dönmez, “Introduction to Machine Learning, 2nd ed., by Ethem Alpaydin. Cambridge, MA: The MIT Press2010. ISBN: 978-0-262-01243-0. \$54/£ 39.95 + 584 pages.,” *Nat. Lang. Eng.*, vol. 19, no. 2, pp. 285–288, 2013, doi: 10.1017/s1351324912000290.
- [10] C. R. Farrar and K. Worden, *Machine Learning and Statistical Pattern Recognition*. 2012.
- [11] F.-G. Yuan, S. A. Zargar, Q. Chen, and S. Wang, “Machine learning for structural health monitoring: challenges and opportunities,” vol. 1137903, no. May, p. 2, 2020, doi: 10.11117/12.2561610.
- [12] Y. Ying *et al.*, “Toward Data-Driven Structural Health Monitoring: Application of Machine Learning and Signal Processing to Damage Detection,” *J. Comput. Civ. Eng.*, vol. 27, no. 6, pp. 667–680, 2013, doi: 10.1061/(asce)cp.1943-5487.0000258.
- [13] L. Sun, Z. Shang, Y. Xia, S. Bhowmick, and S. Nagarajaiah, “Review of Bridge Structural Health Monitoring Aided by Big Data and Artificial Intelligence: From Condition Assessment to Damage Detection,” *J. Struct. Eng.*, vol. 146, no. 5, p. 04020073, 2020, doi: 10.1061/(asce)st.1943-541x.0002535.
- [14] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew, “Rotating machinery prognostics: State of the art, challenges and opportunities,” *Mech. Syst. Signal Process.*, vol. 23, no. 3, pp. 724–739, Apr. 2009, doi: 10.1016/J.YMSSP.2008.06.009.

- [15] M. S. Kan, A. C. C. Tan, and J. Mathew, “A review on prognostic techniques for non-stationary and non-linear rotating systems,” *Mech. Syst. Signal Process.*, vol. 62–63, pp. 1–20, Oct. 2015, doi: 10.1016/J.YMSSP.2015.02.016.
- [16] N. Tandon, G. S. Yadava, and K. M. Ramakrishna, “A comparison of some condition monitoring techniques for the detection of defect in induction motor ball bearings,” *Mech. Syst. Signal Process.*, vol. 21, no. 1, pp. 244–256, Jan. 2007, doi: 10.1016/J.YMSSP.2005.08.005.
- [17] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, “Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks,” *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016, doi: 10.1109/TIE.2016.2582729.
- [18] Y.-J. Cha and O. Buyukozturk, “Structural Damage Detection Using Modal Strain Energy and Hybrid Multiobjective Optimization,” *Comput. Civ. Infrastruct. Eng.*, vol. 30, no. 5, pp. 347–358, May 2015, doi: 10.1111/MICE.12122.
- [19] F. Magalhães, A. Cunha, and E. Caetano, “Vibration based structural health monitoring of an arch bridge: From automated OMA to damage detection,” *Mech. Syst. Signal Process.*, vol. 28, pp. 212–228, Apr. 2012, doi: 10.1016/J.YMSSP.2011.06.011.
- [20] M. L. Fugate, H. Sohn, and C. R. Farrar, “VIBRATION-BASED DAMAGE DETECTION USING STATISTICAL PROCESS CONTROL,” *Mech. Syst. Signal Process.*, vol. 15, no. 4, pp. 707–721, Jul. 2001, doi: 10.1006/MSSP.2000.1323.
- [21] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, and D. J. Inman, “A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications,” *Mech. Syst. Signal Process.*, vol. 147, p. 107077, 2021, doi: 10.1016/j.ymssp.2020.107077.
- [22] X. G. Hua *et al.*, “Structural health monitoring algorithm comparisons using standard data sets,” *Mech. Syst. Signal Process.*, vol. 2, no. 3, pp. 122–135, 2001, doi: 10.1115/1.1410933.
- [23] E. Figueiredo, G. Park, J. Figueiras, C. Farrar, and K. Worden, “Structural health monitoring algorithm comparisons using standard data sets,” *Struct. Heal. Monit. 2009 From Syst. Integr. to Auton. Syst. - Proc. 7th Int. Work. Struct. Heal. Monit. IWSHM 2009*, vol. 2, pp. 1320–1327, 2009.
- [24] S. G. Mattson and S. M. Pandit, “Statistical moments of autoregressive model residuals for damage localisation,” *Mech. Syst. Signal Process.*, vol. 20, no. 3, pp. 627–645, 2006, doi: 10.1016/j.ymssp.2004.08.005.
- [25] *NIST/SEMATECH e-Handbook of Statistical Methods*..
- [26] C. M. Bishop and others, *Neural networks for pattern recognition*. Oxford university press, 1995.
- [27] G. Box, G. Jenkins, and C. Reinsel, “Time series analysis: forecasting and control (third ed.,” *Time Ser. Anal. Forecast. Control Fourth Ed.*, 2013, doi: 10.1002/9781118619193.
- [28] H. Sohn and C. R. Farrar, “Time series analyses for locating damage sources in vibration systems,” *Proc. 25th Int. Conf. Noise Vib. Eng. ISMA*, vol. 836, pp. 209–213, 2000.

- [29] M. Gul and N. Catbas, “Statistical pattern recognition for Structural Health Monitoring using time series modeling: Theory and experimental verifications,” *Mech. Syst. Signal Process. - MECH SYST SIGNAL Process*, vol. 23, pp. 2192–2204, 2009, doi: 10.1016/j.ymssp.2009.02.013.
- [30] Chang-Guen Lee and Chung-Bang Yun, “Parameter identification of linear structural dynamic systems,” *Comput. Struct.*, vol. 40, no. 6, pp. 1475–1487, 1991, doi: [https://doi.org/10.1016/0045-7949\(91\)90418-L](https://doi.org/10.1016/0045-7949(91)90418-L).
- [31] Y. Lu and F. Gao, “A novel time-domain auto-regressive model for structural damage diagnosis,” *J. Sound Vib.*, vol. 283, no. 3–5, pp. 1031–1049, 2005, doi: 10.1016/j.jsv.2004.06.030.
- [32] K. Roy, B. Bhattacharya, and S. Ray-Chaudhuri, “ARX model-based damage sensitive features for structural damage localization using output-only measurements,” *J. Sound Vib.*, vol. 349, no. September, pp. 99–122, 2015, doi: 10.1016/j.jsv.2015.03.038.
- [33] A. K. Pandey, M. Biswas, M. M. Samman, A. K. Pandey, M. Biswas, and M. M. Samman, “Damage detection from changes in curvature mode shapes,” *JSV*, vol. 145, no. 2, pp. 321–332, Mar. 1991, doi: 10.1016/0022-460X(91)90595-B.
- [34] T. A. Duffey, S. W. Doebling, C. R. Farrar, W. E. Baker, and W. H. Rhee, “Vibration-Based Damage Identification in Structures Exhibiting Axial and Torsional Response,” *J. Vib. Acoust.*, vol. 123, no. 1, pp. 84–91, Jan. 2001, doi: 10.1115/1.1320445.
- [35] J. Shu, Z. Zhang, I. Gonzalez, and R. Karoumi, “The application of a damage detection method using Artificial Neural Network and train-induced vibrations on a simplified railway bridge model,” *Eng. Struct.*, vol. 52, pp. 408–421, Jul. 2013, doi: 10.1016/J.ENGSTRUCT.2013.02.031.
- [36] “Support-vector machine - Wikipedia.” https://en.wikipedia.org/wiki/Support-vector_machine (accessed Oct. 20, 2021).