

Vibration-based Damage Detection and Localization in Pipelines Using Data Analysis

Mohammadsajad Salimi

A Thesis
in
The Department
of
Building, Civil and Environmental Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Applied Science (Civil
Engineering) at Concordia University
Montreal, Quebec, Canada

December 2024

© Mohammadsajad Salimi, 2024

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared
By: Mohammadsajad Salimi

Entitled: Vibration-based Damage Detection and Localization in Pipelines Using Data
Analysis

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science

complies with the regulations of the University and meets the accepted standards with respect to
originality and quality.

Signed by the final Examining Committee:

_____ Chair
Dr. G. Hafeez

_____ Examiner
Dr. A. Bhowmick

_____ Examiner
Dr. S. Narayanswamy

_____ Supervisor
Dr. A. Bagchi

Approved by _____
Chair of Department or Graduate Program Director

_____ 2024 _____
Dean of Faculty

ABSTRACT

Vibration-based Damage Detection and Localization in Pipelines Using Data Analysis

Mohammadsajad Salimi

Pipelines are important indicators of modern infrastructure, which allow for easy transportation of significant resources like gas, water, and petroleum over long distances. They often cross very sensitive environments, and any damage to them poses severe and irreversible consequences for marine ecosystems. Also, global warming brings deterioration in infrastructure, which calls for urgent needs for advanced monitoring and maintenance solutions. Their challenges bring about the development of efficient and practical systems to meet the current and future demands of industry. This research focuses on the prediction of pipeline behavior through the vibration signals as a primary method for detection, location, and showing the difference in the extent of damage. Methods like this provide critical details about pipe structural integrity and hence notify early stages of a possible problem well before failure occurs. In this work, pipeline conditions were simulated with numerical modeling using the ANSYS software and then validated with experimental data. Features extracted from sensors were specifically velocity and acceleration. Analysis was done by Principal Component Analysis (PCA), which tries to diffuse data complexity and emphasizes only the most significant variations, where the first principal component carries the most critical information about pipeline conditions and is used as our desirable feature. Independent Component Analysis (ICA) as a method for finding statistically independent components for refining the data is used for detection phase. Application of ICA in this area helps maximize the detection rate for anomalies such as corrosion or structural damage. Then, it combines Mahalanobis distance and K-nearest neighbor methods to accurately localize damage in the pipeline. Initial results using data reveal that the algorithm works well, hence may be applied to real life.

Acknowledgments

I would like to express my gratitude to everyone who contributed to my academic achievements. First and foremost, I am deeply thankful to my research supervisor, Dr. Ashutosh Bagchi, for his patient guidance and unwavering encouragement throughout the invaluable journey of developing this research. I also wish to extend my heartfelt appreciation to my wife, parents, and brother, whose love and understanding have been my greatest sources of support. Without their unwavering belief in me, reaching this level of success would not have been possible.

Table of Contents

Chapter 1: Introduction	1
1.1 Background	1
1.2 Objective of research	4
1.3 Thesis structure	4
2.1 Introduction	7
2.2 Methods of infrastructure health monitoring	8
2.2.1 Local Methods	9
2.2.2 Vibration-based methods	10
2.2.2.1 Model-based approach	11
2.2.2.2 Model-free approach	19
2.3 Scope of Current Work	20
2.4 summary	23
Chapter 3: Methodology	24
3.1 Introduction	24
3.2 Description of Experimental model (Benchmark):	25
3.3 Numerical model	26
3.3.1 Model Description and Boundary Conditions	26
3.3.2 Loading condition	27
3.3.3 Mesh Sensitivity	28
3.4 Dynamic Simulation	29
3.4.1 Modal Analysis	30
3.4.2 Transient Analysis	31
3.5 Data Acquisition	32
3.6 Algorithm for damage detection	33
3.6.1 Processing of Data	33
3.6.2 Principal Component Analysis for detection	34
3.6.3 Independent Component Analysis	37
3.6.3.1 FastICA Method	39
3.7 Algorithm for Damage Localization	41

3.7.1 Principal Component Analysis for Localization	42
3.7.2 K- nearest neighbor	43
3.7.3 Mahalanobis Distance.....	45
3.7.4 Local Mahalanobis Distance	45
Summary.....	47
Chapter 4: Results and Discussion	48
4.1. Introduction:	48
4.2 First Scenario for Detection and Localization.....	48
4.2.1 Data Acquisition	49
4.2.2 Damage detection	50
4.2.2.1 Calculating PCA.....	51
4.2.2.2 Calculating ICA	53
4.2.3 Damage Localization Through LMD	55
4.3 Second Scenario for Localization	58
4.3.1 Graphs for Local Mahalanobis Distance	58
4.4 Third Scenario for Detection	61
4.4.1 Graphs for ICA	62
4.5 Fourth scenario for sensitivity	64
4.5.1 Graphs for Local Mahalanobis distance.....	65
4.6 Fifth scenario for sensitivity investigation	66
4.7 Sixth scenario for size of hole investigation	68
4.7.1 Comparison between model with hole and model without hole	68
4.7.2 Mathematical and statistical investigation	69
Summary.....	72
Chapter 5: Conclusion and Future work.....	73
5.1 Conclusion	73
5.2 Future works	75
References	76
Appendices.....	84
Appendix A:.....	84
A.1 PCA (Principal Component Analysis) calculation	84

A.2 ICA (Independent Component Analysis) calculation	86
A.3 LMD (Local Mahalanobis Distance) calculation	88
Appendix B:.....	91
B.1.1 Data related to baseline pipe (without hole):	91
B.1.1.1 Acceleration for intact pipe	91
B.1.1.2 Velocity for the intact pipe.....	94
B.1.2 Data related to pipe with hole (hole is located 4.4m from the beginning of pipe):.....	97
B.1.2.1 Acceleration for pipe with hole	97
B.1.2.2 Velocity for pipe with hole	100
C. Endurance Limit Factors :	103

Chapter 1: Introduction

1.1 Background

Infrastructure forms the backbone of economic development and quality of life in any given society. Pipelines are the major subset of infrastructures and have been one of the most reliable, effective, and safe ways to transport energy resources across long distances. Alberta is an example to be mentioned because it is an industrially developed area having a considerable volume of pipeline infrastructure. Alberta has a powerful pipeline network stretching over 446,092 kilometers with an average annual growth of approximately 1% for five years (Alberta Energy Regulator, 2023, 2024). These are made up of different materials. Of the close to 388,000 kilometres of pipeline in Alberta, approximately 85% is steel and 15% non-metallic, predominantly polyethylene, fibreglass, and composites that resist corrosion at a better rate. The majority of the former are less than 6 inches in diameter and carry oil and gas products from individual wells to various processing facilities. Only 7% of the total pipeline network that comes under Alberta Energy Regulator is large diameter pipelines with a diameter of 12 inches or more.

By comparing each of the three annual reports in sequence, from the Alberta Energy Regulator's annual expenses in the field of energy regulation, it came to CAD 203,753,000 in 2021, CAD

221,629,000 in 2022, and CAD 223,496,000 in 2023, up to CAD 243,568,000 in 2024. There should be a good response for the reason of bearing such a huge amount, which is needed for regulating pipelines, especially over the last two years when expenses were increasingly growing (Alberta Energy Regulator, 2022, 2023, 2024).

Apart from the above-mentioned pipeline regulation costs, deterioration of the present infrastructure, including pipelines, is being accelerated by persisting problems on civil engineering structures caused by global issues like climate change and steel corrosion.

Recently, with great steps of advance in sensing technology, it has adopted new practical methods, making maintenance process much more efficient. This makes it now possible to keep track of infrastructure assets almost in real time. In real time the amount of information produced by sensors is huge, with the number of measurements a sensor can gather per hour growing into millions. So, the infrastructure health can be much more precisely gauged than ever, and much quicker responses to possible problems, like premature deterioration, structural damage, can be obtained accordingly (Chen & Ni, 2018).

While the volume of data brings its own set of issues, large volumes of data facilitate much better decision-making around when to maintain or repair something in order to reduce risk and extend the lives of an asset. Therefore, through the review of operational data by means of data analysis, a company can achieve the forecast of points of failure, timely repair, and thereby optimization of the repair and maintenance strategy. The prognosis of major failures decreases and extends pipeline reliability for longer periods by this proactive approach (Gulgec et al., 2017).

Nowadays, some new methods like Machine Learning can analyze data as part of Artificial Intelligence; it may work towards creating algorithms with the objective of giving computers the

capacity for far superior data manipulation compared to previous techniques. By using these algorithms, a machine can generate an automatic knowledge base from data. The more data they process, the better they become at making predictions and improving results (Avci et al.,2021).

Fundamentally, there are a few broad categories of applied machine learning techniques in health assessment: supervised and unsupervised learning. Supervised learning is focused on finding the mapping between the input data and the corresponding output. For infrastructure monitoring, however, it relies on the balanced data from undamaged and damaged structures, which is hard to collect. The methods for unsupervised learning aim at identifying patterns within the input data without any prior knowledge regarding the target outputs. These methods can make predictions on new, unseen data based on knowledge that is acquired during training (Santos et al., 2016).

It has drawn increasing attention in recent years considering some certain characteristics (multivariate features) at the same time by developing some data fusion methods. These data analysis techniques try to combine certain information as selected features from various sensor sources with the purpose of enhancing the diagnostic processes' accuracy and effectiveness, or sometimes it could be a method for extracting desirable features out of selected features (Wu et al., 2018).

In the context of the above, this work aims at developing a practical framework by considering the size of the pipeline (the focus is given to the pipes with a diameter less than 6 inches due to practical usage based on annual reports which have explained already) which leverages data-fusion and data-separation techniques and uses them as desirable features for detection and localization goals. The developed system will go beyond pipeline health monitoring on mere damage detection and localization but will also support companies in their further decision-making.

1.2 Objective of research

The main objective of the present research is to develop a data-driven damage detection technique for pipelines using vibration data. Machine learning (ML) algorithms such as Artificial Neural Network (ANN) technologies are ideal for developing a data-driven model. To achieve the above objective, the following sub-goals and tasks have been identified.

- (1) Conduct a state-of-the-art review on the health monitoring of infrastructure, mainly infrastructures and particularly pipelines. For assessment, investigation would be done on traditional methods as well as new approaches using different methods including artificial intelligence.
- (2) Develop a surrogate model using ANSYS software, which would be verified through experimental results for opting selected features in order to extract desirable features.
- (3) Identify a suitable algorithm for analyzing the vibration data by leveraging the advantages of data fusion and data separation techniques to perform multivariate feature analysis.
- (4) Develop a suitable framework to provide for industrial use, focusing on detecting, localizing, and differentiating the extent of damage in a pipeline.

1.3 Thesis structure

Chapter 1:

The first chapter consists of the introduction to the research. In this chapter, an effort is made to discuss the reasons for conducting the research, as well as its significance.

Chapter 2:

This chapter presents a literature review of topics related to this research. These literature reviews vary in scope, including articles on machine learning and other related topics such as methods of data analysis and data fusion.

Chapter 3:

This chapter elaborates at great length on methodologies adopted for the research, while step-by-step development of the finite element model is discussed. This includes a step-by-step explanation of how the model is made and used in the implementation. The next segment presents the process for data extraction from the model, considering essential steps to extract relevant data for analysis. In the final part, the framework of the anomaly detection and localization carried out is reviewed by describing techniques and processes applied in the identification and location of defect within the pipe.

Chapter 4:

In this chapter, the introduced frameworks for localization and detection are considered in five different scenarios. The first scenario considers a detailed analysis of the processes of localization and detection. Different steps involved in this scenario are elaborated in detail. For the process of detection, the steps involved in data fusion and data separation are elaborated, which lead to the detection of defects through the presented framework. The localization stage mainly presents the procedure of data fusion and the transformation of chosen data to desirable data. From the framework presented, the localization is performed using local Mahalanobis distance analysis. In the second case, similar to Case 1, the localization process is investigated and verified by presenting preliminary graphs with regard to LMD (Local Mahalanobis Distance) and the process of extracting greatest peak are rendered. The finding of defects in Scenario 3 is done on a pipe

containing two holes at different locations with variable sizes. In Scenario 4, the proposed approach has performed localization by increasing the distance between sensors in order to study the sensitivity of the approach. Mathematical models are introduced in Scenario 5 in order to differentiate between holes of various sizes.

Chapter 5:

In this chapter, the summary of the results is discussed, while the efficiency of the framework is reviewed. Also, some suggestions for future works is rendered.

Chapter 6:

In this chapter, an appendix which is related to necessary data are rendered.

Chapter 2: Literature Review

2.1 Introduction

Pipelines create a critical way of modern infrastructure; in many ways, they become vital for the survival of world economies and security. Since they usually act as conduits for key resources, their failure or hindrance often tends to be serious. Therefore, it is also a focus area for economic and industrial planners (Salimi and Bagchi., 2024).

The current chapter is supposed to provide, based on reviewed literature, a critical review of vibration-based monitoring of the health of infrastructural elements. Attention has shifted to this method since it works effectively in monitoring and evaluating pipeline condition. By analyzing the behavior due to vibration, early damage or even wear is possible to detect as failure, which sustains pipeline reliability and safety. It reviews past research, technologies, and applications of vibration monitoring. Hence, it sheds light on methods being applied in reality to avoid major failures and increase the life span of a pipeline.

2.2 Methods of infrastructure health monitoring

This overview of the available literature is categorized into two distinct classes (Avci et al., 2020). First, termed as local method, encompasses conventional methods that typically rely upon a high degree of expertise from a seasoned technician. The techniques either constitute close observation or a test the structure, and they have seen extensive application over many years and in those situations where structural or surface damage is evident. The second category, termed global method, involves all vibration-based methods in front of the growing understanding of the importance of vibrations in structural health. These methods make use of the analysis of vibration responses of structures for detection and assessment of damage and in many cases these methods are so similar to the human nervous systems (refer to Figure 1). Vibration-based methods have received considerable popularity because of the potential monitoring of a structure's internal health without direct access or invasive inspections. In the following sections, we will explore the different methods within these two categories in various types of infrastructure.

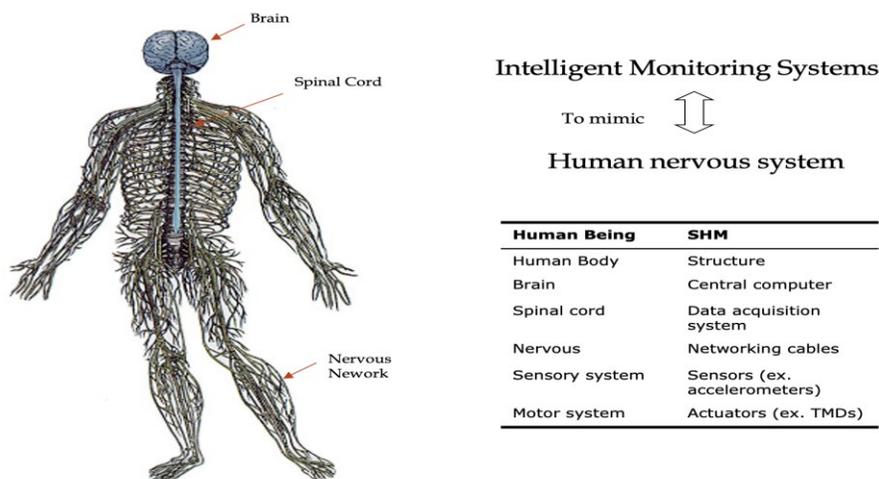


Figure 2.1. similarity between SHM and nervous system source: <https://civilresearchgroup.ulusofofona.pt/research/structural-health-monitoring>

2.2.1 Local Methods

Most techniques falling into this category would be called non-destructive testing (NDT). As previously mentioned, the Local methods seek to examine the damaged areas on a smaller scale. They also require the availability of qualified personnel in order to carry out the tests effectively and accurately.

These methods also include Ultrasonic Testing, Infrared Thermography, Radiographic Testing, Magnetic Flux Leakage, Magnetic Particle Testing, Liquid Penetrant Tests (PT), Ground Penetrating Radar (GPR), Leak Testing and Visual Testing (VT), among the techniques used for sectional area testing of infrastructure (Rens et al., 1997; Sophian et al., 2001; Chang and Liu, 2003; Dwivedi et al., 2018). Those who would like to go deeper, an array of references and texts involving extensive detail are suggested.

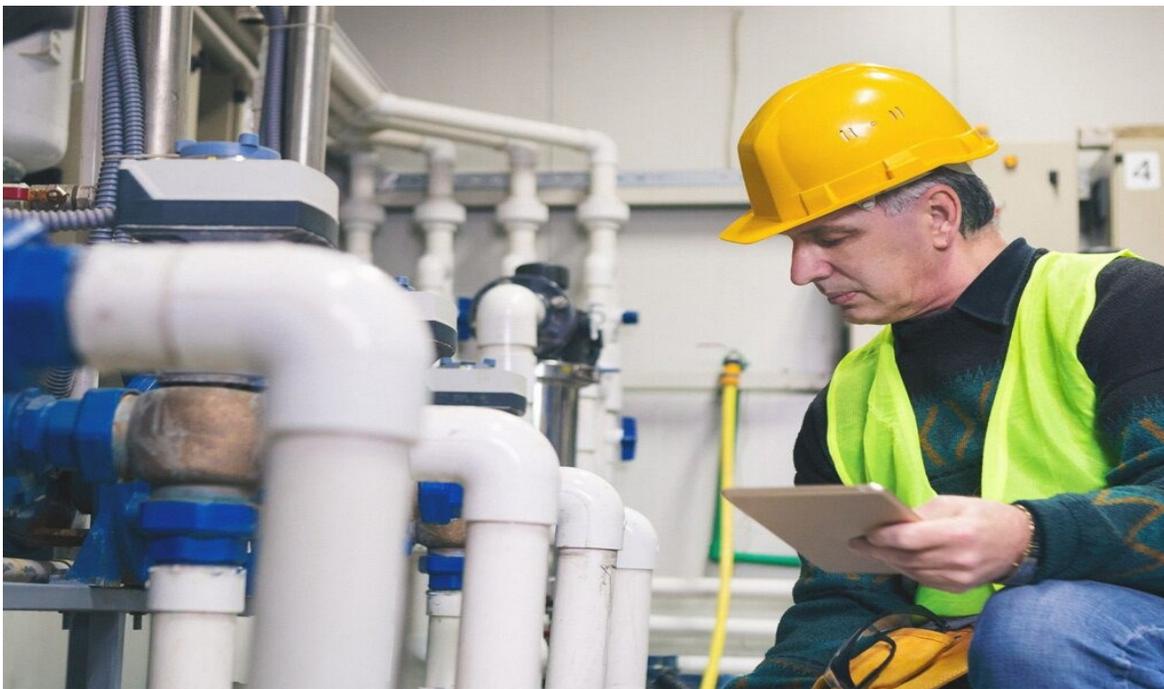


Figure 2. 2. visual test for pipes_ sources: <https://www.fbrcontrol.com/en/controllo-visivo/>

2.2.2 Vibration-based methods

Starting to think about the vibration methods for damage detection, the historical background goes back to the beginning of the 1800s, when railway workers moved around and checked the condition of a train wheel by hitting it with a hammer and listening to the succeeding sound. Quite a modest approach that gave them a qualitative judgment whether the wheel was damaged or in good condition. Development of vibration methods for detecting damage was only possible in the 1980s due to improvements in computing sensing technologies and analysis the data. The latter technological improvements allowed more exact and data-driven approaches in determining structural damage (Simoen et al., 2015).

Some of the features that have made vibration-based damage detection more feasible and also more reliable compared to the other methods, especially traditional ones include (Chen and Ni, 2018):

- **Early detection of damage:** In the vibration-based damage detection approach, the structure expresses different behavior throughout each stage of damage. Even minor wear or corrosion of a structure causes detectable changes either in the pattern or characteristics of the vibration like frequencies, mode shapes and damping, thus allowing early identification of any problems that may occur (Humar et al., 2006).
- **No need for prior knowledge of damage location:** By having vibration data through data acquisition and selecting certain features and analyzing them, the process is going to be completed. Therefore, exactly where the damage will occur may not be known in advance.

- **Portable and easy-to-use equipment:** The equipment used for most vibration-based techniques is portable, easy to operate in terms of recording data, and can be quickly installed onto the structure being monitored.

Various classifications have been proposed under infrastructure health monitoring category from different sources. In this study, based on the focus of the research which integrates experimental parts, numerical parts and any other methods, two broad classes are adopted: model-based approach and model-free approaches (Wang et al., 2023).

Most of the work done in vibration-based infrastructure health monitoring is presented in the following sections. Reviewing the types of models, type of damage, and methods for obtaining results in different studies conducted with a focus on vibration analysis is the main theme of this review.

2.2.2.1 Model-based approach

The model-based approach uses some stem concepts, such as finite element analysis, for simulating the infrastructure (to build surrogate models in many cases) and extract some desired parameters useful for assessing structural health. While this indeed provides the user with greater freedom and flexibility to extract desirable features that facilitate more refined analyses and optimization, based on specific needs, this is heavily dependent on creating accurate models, which are able to predict the structure's behavior in various conditions. The next section shall present the papers classified in this category, in chronological order, so as to preserve the timeline of developments.

(Yeung and Smith, 2005) applied some damage detection techniques in bridges; the focus was to make use of neural networks for pattern recognition of vibration data. The test structure used is a suspension bridge, already over 100 years old, and simulated **damage includes that of corrosion or loosening joints (reduced stiffness)**. A finite element model generates natural frequency response data under moving traffic loads, and this data is used to train two unsupervised neural networks for damage detection.

Based on the framework developed by (Lee et al., 2005), a multi-girder bridge model was analyzed, with damage relating to **loss of stiffness due to changes in the bending rigidity** of structural elements. Modal data, such as mode shape differences between intact and damaged states, are used in this approach to train neural networks. Numerical simulations and real experimental tests on the bridge structure will generate the data. In this, a back-propagation neural network classifier is used for damage identification.

The methodology for damage detection in beams having fixed boundary condition using finite element model was developed by (Pawar et al., 2006). The type of damage being studied here is that of **stiffness reduction that may occur in beams**. In this methodology, the use of spatial Fourier analysis of natural frequencies and mode shapes of the beam is made for quantifying the changes arising due to damage and subsequently making use of neural network for classification purpose.

(Yuen and Lam, 2006) improved the damage detection techniques applied to the five-story building models with the motivation of identifying damages in forms such as **reduction in inter-story stiffness at different levels**. The developed technique uses a Bayesian probabilistic approach together with ANNs for the identification and classification of the type of damage based on natural frequencies and mode shapes.

A method was developed by (Lee and Kim, 2007) for structural damage detection using frequency domain analysis, with a focus on steel girder damages caused by **loss of stiffness due to cracking or loosening connections**. The structure analyzed in this research is a bridge model with steel girders, while the kind of damage studied involves saw-cut cracks and loosening of connection. Thus, a hybrid methodology is developed coupling experimental testing with finite elements simulations. dynamic response data, frequency response functions, are collected for the intact and damaged structure. The classification algorithm adopted in this work is a NN that processes, preferably, frequency response data to detect any damage in the bridge girders.

(Mehrjoo et al., 2008) proposed an efficient method for damage estimation in the joints of two different truss bridges including Louisville Bridge in the U.S.A using back-propagation Artificial Neural Networks. The nature of the damage considered in this study is **fatigue damage in the joints** of truss structures, which is usually not easily detectable by conventional inspections. The modal data including natural frequencies and mode shapes are used as input to pinpoint damage in this study.

(Gonzalez and Zapico, 2008) attempted to improve seismic damage identification in buildings, but they limited their field of study to the steel moment-frame structures that are most vulnerable to **stiffness decreases** due to beam and joint damage caused by seismic events. The method proposed neural networks to identify damage, the model requiring, as input for estimating changes in structural behavior, modal data such as natural frequencies and mode shapes. Data for the study come from finite element simulations of a 5-story office building under seismic conditions.

(Lam and Ng., 2008) proposed a structural damage detection model using an intelligent hybrid ANN system for diagnosing the occurrence of damage in structures. The model under question is from a 10-story shear building, where the primary challenges faced are identifying the **stiffness**

reduction caused due to the presence of the damage at each story. According to the methodology, the simulation is carried out by executing a modal analysis with the aid of simulated data through the GRNNFA model, a general regression neural network combined with fuzzy ART (FA), for detection and localization of damage.

(Bakhary et al., 2010) presented an approach to damage detection in reinforced concrete slabs, focusing on detecting a **reduction in stiffness arising from cracks** due to loads applied to the structure. The structure considered in the analysis was a two-span concrete slab with controlled increasing loads that simulate various levels of simulated damage. Modal data, including natural frequencies and mode shapes obtained with finite element simulations and experimental testing, were considered as an input for the methodology, able to estimate changes occurred in structural behavior because of loadings. A multistage ANN classifier was employed for detection.

(Jiang et al., 2011) developed a method for structural damage detection. The model analyzed in the study is a 7-degree-of-freedom building and a 12-story reinforced concrete frame. The types of damage considered are the **stiffness reduction at various structural components** to simulate damage patterns for extracting modal parameters. The methodology that was mainly adopted in this research work is the immobilization of rough set data and probabilistic neural network (PNN) for multi-sensor data fusion.

(Jiang et al., 2011) improved structural damage detection based on a two-stage approach that combines fuzzy neural networks with data fusion techniques. A seven-story shear beam building serves as a model for analysis. The type of damage studied involves **stiffness reduction at different story levels** to simulate the deterioration of structures. Their methodology includes the use of modal parameters from vibration responses as inputs into models of FNNs. These parameters undergo data fusion to refine the results of the damage detection.

(Cury and Crémona, 2012) have proposed an identification methodology for monitoring two different models: a railway bridge in France and a simply supported beam with damage scenarios based on **stiffness reductions simulating real structural modifications**, considering the application of vibration data and modal properties - natural frequencies and mode shapes—transformed into symbolic data, combined with Bayesian Decision Trees, Neural Networks, and Support Vector Machines for classification.

(Goh et al., 2013) review the methods that have been in use for structural health monitoring in order to carry out analysis on a two-span continuous reinforced concrete slab with simulated damage scenarios made using **stiffness reduction in some sections** to represent actual structural damage. The approach here is the use of a multi-stage ANN model to first estimate unmeasured modal data and then detect and quantify the damage using this estimated data. This approach also uses probability tools to consider uncertainties in the measured data.

(Betti et al., 2014) discussed the efficiency of combination techniques using ANNs and GAs in structural damage detection of a three-story steel frame. Progressive damage was simulated by **cutting a steel column above the first story which took away some of the element stiffness and created structural weaknesses**. The methodology involves an experimental layout with an instrumented steel frame using accelerometers to capture vibrational data in terms of natural frequencies with mode shapes under ambient vibration. Later, ANN is to be trained so as to assess the modal properties, natural frequencies and mode shapes, and for the determination of error functions by optimization using a Genetic Algorithm (GA).

The work by (Zhou et al., 2014) was dedicated to enhancing damage localization techniques for cable-supported bridges; identification of structural damage is dealt with by applying frequency data of mode. The structures investigated represent two famous cable-supported bridges: Tsing Ma

Bridge and Ting Kau Bridge, which have simulated cases of damage in the form of **a reduction of stiffness in deck units, stay cables, and bearings**. The method that has been presented will create 3D finite element models of these bridges, analyzing the change in modal frequencies due to damage introduced. The classifier acts within a Bayesian probabilistic framework that detects and localizes damage using the Probabilistic Neural Network classifier.

(Hakim et al., 2015) conducted a study related to the detection of structural damage associated with beam-like structures using artificial neural network (ANN) as a classifier. **The stiffness reduction due to cracks** comes into focus. For analysis, the model made use of a steel I-beam structure. Types of damage analyzed include cracks at several locations. The experimental works are combined with FEA in order to generate desirable data comprising natural frequencies and mode shapes.

Critical elements of the monitoring process, such as the identification of the natural frequency, modal analysis, and updating of finite element models, were conducted in a study by (Valinejadshoubi et al. 2016) through a case study approach. They acknowledged the effectiveness of M-FEM software for model updating through an iterative approach. A concrete beam was taken as a test object on which **acceleration** data allowed determining its natural frequency in two states, at pre- installation and post-installation of concrete blocks, by applying FFT analysis. Further, they also conducted modal analysis and finite element models to evaluate how structural damage could be identified in the beam.

(El Mountassir et al., 2018) propose a new technique to reduce the impact of environmental and operational condition changes by using sparse estimation of monitoring signals. This approach involved comparing current signals, via sparse estimation, to a baseline database of signals from a healthy pipeline, with the estimation error serving as a damage index. It was applied to signals

collected from a small-scale pipeline in laboratory conditions. Tests with ultrasonic guided waves were conducted for data collection.

(Lyapin et al., 2018) focused on identifying structural damages by monitoring variations in dynamic characteristics, eigenfrequencies and mode shapes, **introduced by defects**. The finite element modeling of changes in significant vibration modes allows for the identification of targeted damage and its location using a limited number of sensors. This investigation supports that when the mode shapes are well-defined with a sufficient number of measurement points, damage detection relies on factors such as damage severity, the number of trials performed to average the mode shapes, sensor placement relative to the damage, and measurement accuracy.

(Houdek et al., 2022) propose a new approach to the problem of impact detection in three-dimensional structures using the wavelet transform. The method uses an array of redundant sensors capturing the vibrations coming from impacts; it adopts Continuous Wavelet Transform with optimal partitioning of intervals to achieve accurate time-of-arrival estimates.

(Mousavi et al., 2022) present a general review of vibration feature extraction using signal processing techniques. Various signal processing techniques that offer a view on works that use the process of feature extraction are organized into two main classes in a systematic way: time-domain approaches and frequency-domain approaches. This will be relevant to enable the recognition of changes or damages in dynamic systems by means of vibration signals, which can be useful for detecting damage..

(Xie et al., 2023) performs acoustic emission tests to identify fatigue damage **in oil and gas pipelines**, even under harsh conditions of measured noise onto offshore platforms. The authors introduce the pipeline acoustic signal collection and later inject the signals with measured field

noise to simulate real-life conditions. Empirical mode decomposition and the probabilistic neural network are then applied for identification of noise from actual damage signals. The time difference method serves for preliminary positioning, refined with backpropagation neural networks for adjustment of results.

Li et al. (2023) aimed to demonstrate how the **presence and angular positions of defects** affect the detection of defects based on the application of damage indices. Numerical models of laminated pipeline structures with different defect scenarios were designed in such a way that the defects' effects were analyzed for the damage indices. The models used in this section are a steel pipe, rigid polyurethane foam, and high-density polyethylene. The procedures consisted of signal averaging, a level 5 wavelet packet transform, and arranging computed damage index values in matrix format.

(Alves and Cury, 2023) rendered an approach using filter-type feature selection. The workflow encompasses a multilevel feature extraction in time, frequency, and quefreny domains, combined with a subsequent unsupervised infinite feature selection technique for the actual filtering out of irrelevant redundancy from the feature set. The core of the proposed technique is an outlier analysis-based percentile intervals of filtered features, computing a damage-sensitive index that allows for the detection of anomalies relative to a healthy structure state.

(Wang et al., 2023) have proposed a new enhanced damage detection methodology in infrastructures using data fusion and machine learning. They developed a framework that takes into consideration three major items of data, namely acceleration, deflection, and bending moment measured by sensors mounted on a simply supported bridge and a continuous bridge. Principal Component Analysis and Mahalanobis distance were applied to clean the data and extract the

features related to damage. Obtained features are then fed into a Deep Convolutional Autoencoder that classifies the normal and damaged conditions.

2.2.2.2 Model-free approach

The model-free approach applies to those detecting cases of damage through other means, such as through signal processing or statistical techniques, machine learning, among others. No comprehensive model of the structure or finite element model making is involved in this approach. Instead, the identification of anomalies and possible damage is done by applying data-driven methods based on signals or data gathered from the structure. We present these studies from this category next in chronological order.

Damage detection by stiffness reduction due to **the removal of braces** is pursued for a four-story, two-bay steel frame structure by (Lam et al., 2008). This study features the ASCE benchmark model in which damage scenarios, including partial and full removal of braces, result in **stiffness reduction** at specific locations. It performs a combination of vibration analysis through dynamic response data generated under ambient and shaker excitation, extracting modal parameters such as natural frequencies and mode shapes. A Bayesian Artificial Neural Network classifier is then applied for the identification of damage.

(Bao et al., 2013) proposed Autoregressive Moving Average (ARMA) model-based detection in subsea pipeline systems using signal and statistical methodologies to analyze data. First, this method preprocesses the vibration data to reduce the influence of variable loading conditions by partitioning, normalization, and the autocorrelation function, enhancing the quality of the signal and reducing the effects of noise. A good ARMA model is developed based on a Partial Autocorrelation Function analysis wherein the autoregressive parameters are considered to be

vectors representing the features of damage. A defined damage indicator corresponding to the Mahalanobis distance between ARMA models is applied for identification.

(Ching-Tai Ng., 2014) attempts to investigate multi-story steel frame structural damage detection, particularly the **stiffness reduction in the braces caused by the damages**. A finite element model of a four-story steel frame was used in conducting research on various simulated damage conditions (Phase II IASC-ASCE SHM benchmark structure). The methodologies involved vibration feature collection in the form of natural frequencies and mode shapes from both undamaged and damaged structure states, which were processed with a Bayesian-designed artificial neural network to identify patterns of stiffness reduction.

(Yang and zhao, 2022) proposes a leak detection and localization approach (not vibration based) in **pipelines** by incorporating a variant of artificial intelligence called BiLSTM, which can learn patterns in data sequences backward and forward. This will help solve the problem associated with traditional systems, which might get confused between normal changes in **pressure and leaks**. This works through pattern analysis of pressure changes versus time in order to differentiate an actual leak and normal fluctuations due to, for example, pumps working. These systems are trained on real pipeline data for accuracy and tested on simulated and actual pipeline data.

2.3 Scope of Current Work

Before discussing the scope of the current work, it is important to note that the nature of damage can be either intentional, due to deliberate actions like explosion or terrorist attack, or non-intentional, such as damage caused by corrosion. Furthermore, depending on its nature, damage can take various shapes, such as an elliptical form.

In this project, a circular hole with a size of 1 mm on top of the pipe (+ 90 degree) is considered as the representative damage for a basic model. This choice was made to focus on data analysis while minimizing the size of the defect for ensuring a more comprehensive assessment.

Based on above explanation, most of the research which rendered in Sections 2.2.2.1 and 2.2.2.2 has used natural frequencies and modal shapes as desirable features (the most frequency), with stiffness reduction as an indicator for damage which is highlighted. This project, by taking a different approach toward these desirable features, justifies itself and even makes it necessary. It is relevant to note that both processes in this project, detection and localization, will be carried out solely through signal processing.

After data acquisition which will be explained in next chapter, to extract unique data, various techniques will have to be applied. Among them, the process of data fusion for generating unique data by combining different data sources is used which has gained considerable attention over recent years (Wang et al., 2023).

The Principal Component Analysis (PCA) usually is a method of unsupervised learning, although it can be applied to fuse the information in any other way. It cleans the data in two ways: dimensionality reduction and retaining the variance of data. In other words, the principal component analysis is one of the statistical methods that may convert a set of selected features onto a set of principal components while retaining the maximum variance of data (Tharwat, 2016).

The purposes of using PCA for data fusion in the infrastructure project are two-fold. One is to do whitening (Hyvärinen et al. 2001), which will be elaborated more in the section for Independent Component Analysis, and the other is to apply data fusion on the selected features to obtain a desirable feature, then use it as an input to perform anomaly localization.

For the final part, data which is combined by means of data fusion can be understood using distance metrics for classification or any other goals (refer to Figure 2.3). Among many distance metrics, such as Euclidean distance, Mahalanobis distance, and Minkowski distance, the Mahalanobis distance is used in this project.

The Mahalanobis distance metric has a similar conceptual meaning to the Euclidean distance in algebra (Ghorbani, 2019). However, the two differ in that whereas Euclidean distance defines the distance between two points, Mahalanobis distance defines the variable distance between points and a set of points. It does this by combining two main capabilities: accounting for the correlation between variables and accounting for the differences in variances. It tends to put a little weight on the highly correlated variable as well as the variable with high variance; hence all the characteristics are weighed equally.

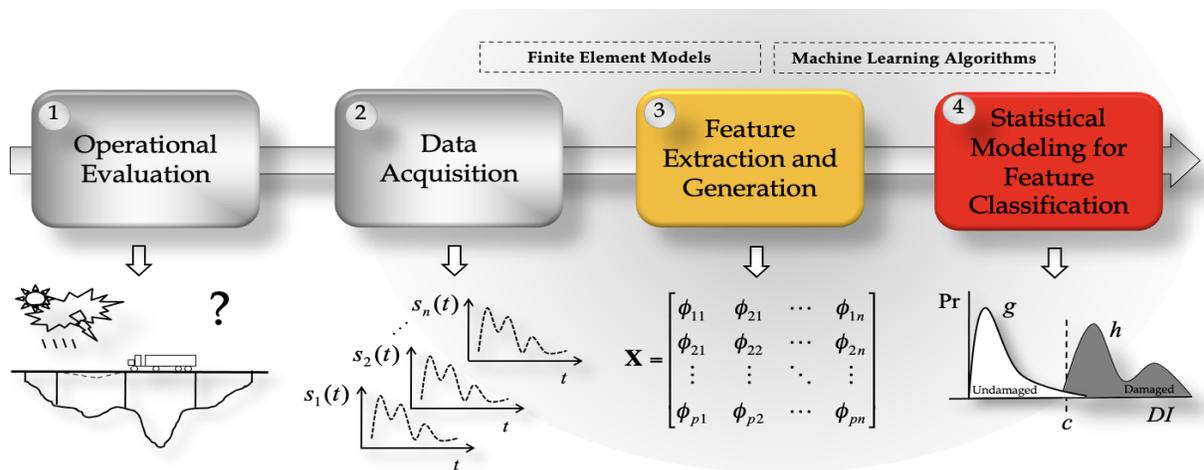


Figure 2. 3. Health Monitoring steps source: <https://civilresearchgroup.ulusofona.pt/research/structural-health-monitoring>

2.4 summary

This chapter presented an overview of major works conducted in the field of vibration-based SHM. The studies have been reviewed with regards to the type of damage analysis, structure under investigation, and methodology used for detecting the damage. It is certain from this review that stiffness reduction has been the most commonly used damage indicator in the majority of studies. The methodology of damage detection depended largely on the type of structure analyzed, and a great portion of the works reviewed fell into the category of the model-based approach. The main reasons why the model-based methods were preferred include their flexibility in feature extraction through the creation of a surrogate model. Guided by this insight, a model-based approach has been used in the development of the surrogate model and feature extraction in the current project. Therefore, in this context, an effective solution is to be proposed which could minimize the cost and consumption of time for damage localization. The next few chapters will introduce the method used along with the validation through a verified surrogate model experimentally.

Chapter 3: Methodology

3.1 Introduction

The objective of this chapter is to give an overview of the developed package for damage detection and damage localization. As discussed in Chapter 2 under the literature review, several studies have been undertaken on damage detection hinged on the model-based and model-free approaches outlined therein. Each of the methods presented has tackled damage detection and localization in a certain way, depending on the initial assumptions made during the selection of desirable features. These features are important for the identification and location of damage in a structure and are selected according to the approach that will be used to perform the analysis.

What sets different studies apart is the selection of desirable features. In many cases, this aspect has led researchers to generate data specifically to identify these features. By creating a mathematical model (surrogate model), they are able to utilize the capabilities of the model to extract the required data. This makes the extraction of features more focused, powerful and economic, thus identifying the required insights in damage detection and analysis. The mathematical models can then be validated and may be calibrated with small-scale laboratory models. In this way, verification of numerical (surrogate) models will provide real conditions and enhance their reliability for further analysis and predictions in damage detection. Below is the package developed for damage detection and localization. This chapter elaborates on every step in

the proposed package, letter by letter, to ensure that everything is explained, and ambiguities are removed. It aims to give a complete understanding of the workflow in damage detection and

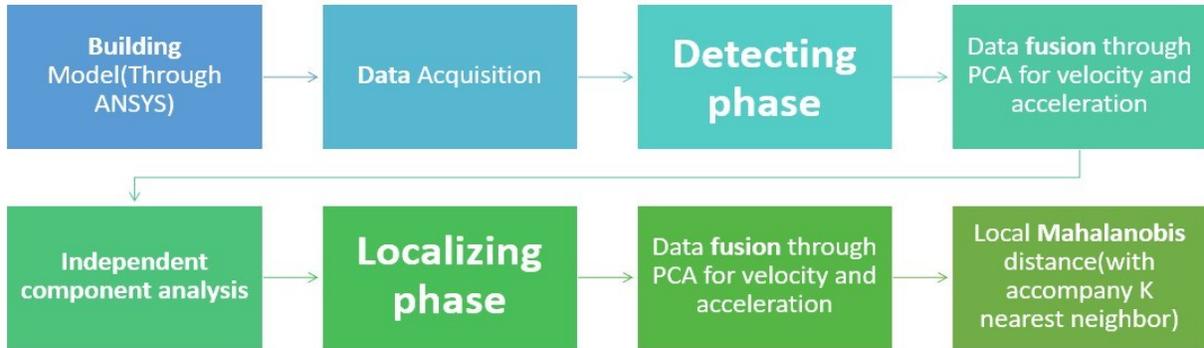


Figure 3.1. Proposed package for detecting and localizing phase

3.2 Description of Experimental model (Benchmark):

This project incorporates experimentally obtained data and numerical modeling to study certain phenomena. The experimental part is based on the research reported in (Khazeli , 2018), serving as a benchmark for this project. In the above experiment, the pipe was segmented into ten equal parts. An accelerometer was mounted on the pipe, and impacts were delivered by a hammer at eleven points, including the pipe ends. Each point received five impacts, with results averaged to avoid potential noise interference. Data was transferred and stored using a data acquisition system (NI 4310). The process of collecting was complemented with (PCB-type) sensors and hammers, along with noise-suppressing cables to ensure data integrity.

Experimental work relies on modal tests conducted on a 3-meter-long pipe to determine its natural frequencies, allowing for the extraction of frequency and modal shape—crucial elements for validating the accuracy of the numerical model. More details on the experiment can be found in (Khazeli, 2018).

3.3 Numerical model

Following the steps outlined in Figure 3.1, a numerical base model is built using ANSYS Workbench to match the experimental model for verification purposes, ensuring that results are comparable to the experimental data. This section provides a detailed description of each step involved in building the numerical model, following the necessary sequence for accurate simulation and comparison with laboratory results.

3.3.1 Model Description and Boundary Conditions

A finite element simulation was conducted on a pipeline characterized by a circular cross-section 3000 mm in length, with an inner radius of 25.14 mm, and an outer radius of 30.15 mm. Other characteristics of the pipeline are shown in Table 3.1.

Table 3.1. Characteristics of material

Characteristics	Values	Unit
E (Elasticity Modulus)	2e+11	pa
μ (Poisson Ratio)	0.3	–
G (Shear Modulus)	7.7e+10	pa
Coefficient of thermal expansion	1.2e-05	1/C0

In this simulation, the target is to consider a boundary condition in such a way that numerical model closely simulated to experimental description as much as possible. Based on matching the frequency and modal shape with the laboratory conditions, two hinge supports are used for both sides of the pipe. Numerical modeling assumes homogeneous and isotropic pipe material with

elastic behavior in the present analysis. This implies that the material properties are the same everywhere for the pipe, and stimulation (analysis) will be performed within the elastic limit.

3.3.2 Loading condition

For the simulation of model, the behavior of the pipe is evaluated due to impact loading. Based on an experimental benchmark presented in (Khazeli, 2018), the calculation of impact loading takes into consideration the maximum permissible stress (Nisbett et al. 2015), adjusted by factors related to the endurance limit to incorporate various practical aspects like load, surface state, temperature, etc. (Refer to Appendix C). This impact loading equal to 300 N is applied at a fixed distance from the beginning of the pipe (in the models examined for damage detection and damage localization, a distance of 15 cm has been considered.). This kind of impulse aims to mimic the hammer test previously discussed. The whole period when the dynamic behavior of the pipe under the effect of an impact is investigated is 0.1 seconds, it should be mentioned that the dynamic load is applied at the time interval of 0.01 seconds and the time step is considered 0.001 seconds.

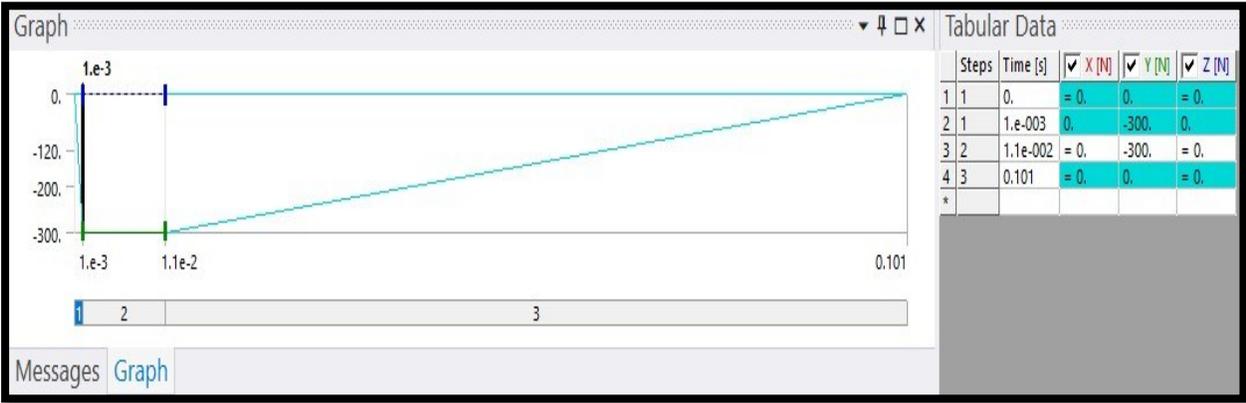
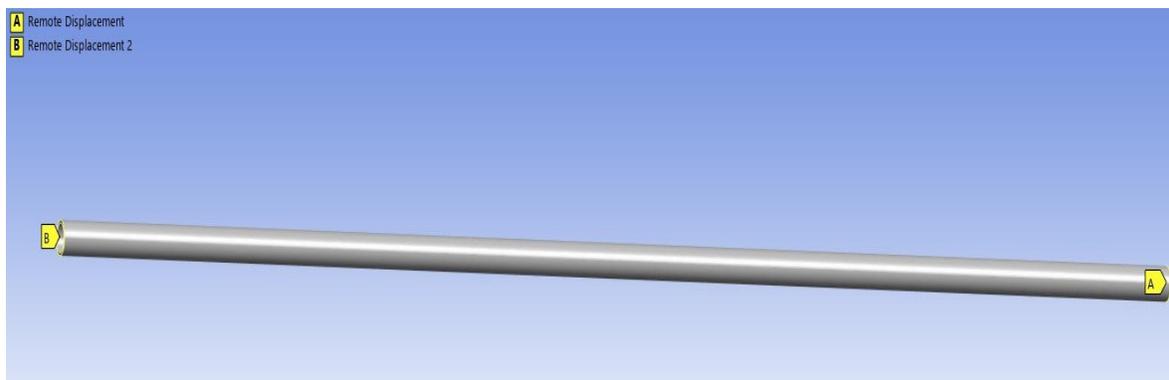


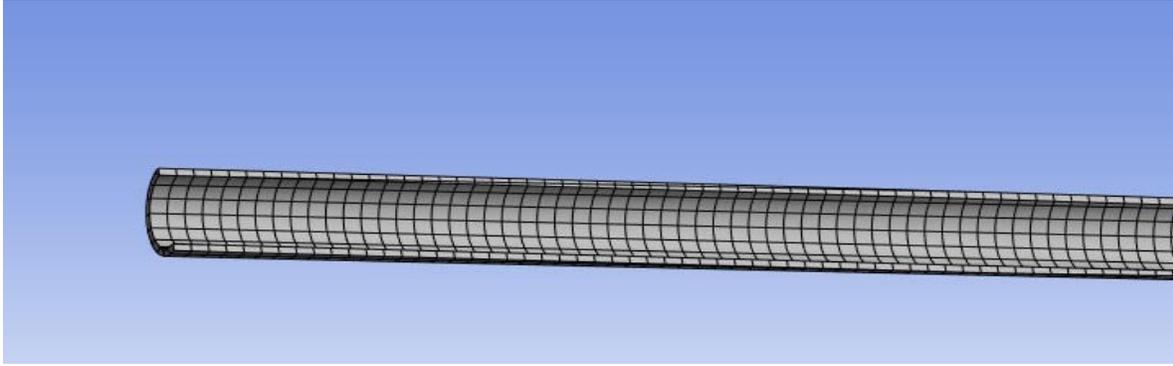
Figure 3.2. The impact loading which is applied 15 cm from the start of pipe (transient Analysis step)

3.3.3 Mesh Sensitivity

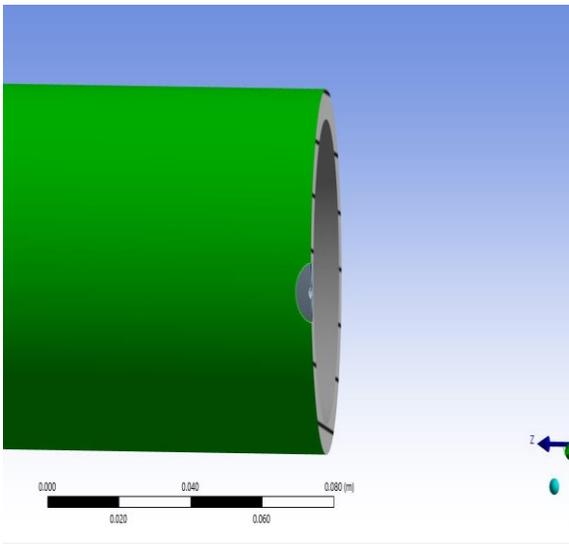
The most technical part of the simulation is related to how to mesh. Different techniques are examined through trial-and-error methods. For the pipe that is intact and modeled to match the experimental setup, a default mesh with moderate size was used during the simulation process. For the basic pipeline model, the number of nodes and elements are 36595 and 5586 respectively, also size of elements are 1 cm. It is worthy to say that each hole in a model is located on the top of the pipeline (+90 degrees) and for models with holes, three different mesh techniques are used. The first technique is meshing with elements of fine, moderate, and coarse size, different models are built and due to the limitation of the number of nodes, elements, and proper accuracy, the moderate size is applied for the first step. The second step pertains to using a multi-zone technique with (Tetra) mesh type and (Hexa) mapped type for the zone with a hole because without applying a correct technique for the mentioned zone, different sorts of errors are reported through software. The last step refers to how to merge different zones, and for this step, meshes were edited with the refinement technique, these different zones are merged through the Node merge technique and different sorts of error messages are modified. (Refer to Figure 3.3)



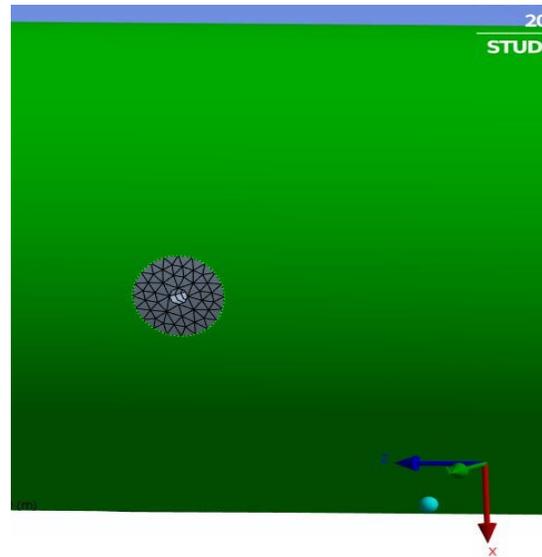
(a)



(b)



(c)



(d)

Figure 3.3. Examples of different views for the pipe (a) Considering two hinged supports at both ends based on remote displacement technic. (b) Section view of a typical mesh grid for pipe with no hole. (c) and (d) Multizone technic for the pipe with hole (basic model)

3.4 Dynamic Simulation

In this modeling, different methods for dynamic simulation were tested. Finally, two methods were used for modeling because of selecting desirable features, including Modal analysis and transient analysis. Modal analysis is done to extract different mode shapes and frequencies that are

necessary for verification purposes. Transient analysis is done to simulate impact loading (For all the models which explained in the next chapters impact loading is applied at the fixed distance of 15 cm from the beginning of the pipe). The total simulation time for investigating transient analysis is 0.1 sec, divided into three stages which will be explained in the forthcoming sections.

3.4.1 Modal Analysis

Modal Analysis has been done for confirmation purposes and for conforming numerical model with experimental results. Based on Table 3.2, considering the most suitable support conditions for initiating a close state amongst numerical models with experimental work, support is modeled through the remote displacement technique in ANSYS Workbench. In the following, the frequency of four different mode shapes for two different states, including results of experimental work and numerical simulation are shown in Table 3.2.

Table 3.2. Frequency (HZ) of mode for models

Mode	Experiment	with support	Error
Mode1	40	38.6	2.6%
Mode2	110.9	105.7	4.7%
Mode3	216.6	205.6	3.4%
Mode4	365.2	336.5	7.8%

According to Table 3.2, frequency values and mode shapes are shown for modes 1 to 4. It is worthy to mention that all the steps for detecting and localizing damage are set based on mode 1 as dominant mode, other Modes are not participated for detecting and localizing purpose because of different reasons, like preventing to put sensors at node point.

Based on Table 3.2, It could be seen that for mode 1 as the dominant mode, only 2.6% error has been calculated for the numerical model in comparison with experimental result (More than 97% precision). Also, for mode 2, the calculated frequency error in comparison with experimental result is limited to 4.7 %. Moreover, for mode 3 and mode 4 the frequency error is computed up to 3.4% and 7.8% in order (even up to 93% precision for mode 4 as the lowest precision). By considering mode 1 as our target, it could be seen that there is a good correlation between simulation results and experimental states.

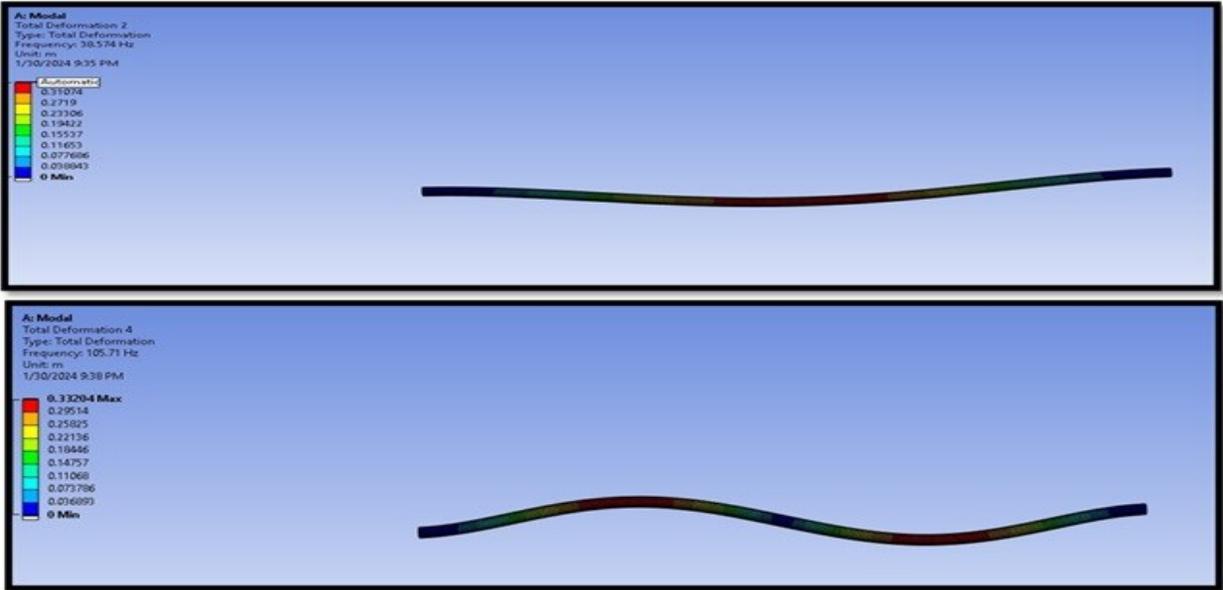


Figure 3.4. The first and second mode for the pipe with hinged support

3.4.2 Transient Analysis

For the simulation of dynamic load (impact load), the transient analysis technic is applied, the analysis setting is set in three distinct stages with the chosen time step as 0.001 s, and based on Figure 3.2 the impact load is considered 300 N in the time interval between 1e-3 seconds and 1.1e-2 seconds.

3.5 Data Acquisition

Different models of pipes are numerically simulated in this work. In the initial stages of modeling done by the software, tries to correlate the numerical model developed as much as possible with the experimental model taken as the reference. The results reported in Table 3.2 depict this agreement quite well, where quite a good consistency between numerical simulations and experimental data is well appreciated.

Table 3.2 reflects nearly a 2% error in the frequency of the first mode, which we are going to base our numerical modeling on from the experimental model. This corresponds to an accuracy of 98%. This minor deviation lies within the realizable and acceptable errors, as induced by the capacity and also the inherent limitations of the used ANSYS software for conducting the simulations. For localization and detection of damage, an undamaged or an intact pipe model has been used as a baseline model.

Given the capabilities of ANSYS, it is possible to extract a wide range of data from both the baseline and damaged models. At the first glance, the extracted data, representing the structural responses, may appear quite similar. Therefore, selecting the right desirable features for analysis becomes a crucial and essential task.

For this project, various types of data were extracted, including acceleration, velocity, displacement, and strain, etc. Ultimately, after simulating a total of 36 various models based on factors like location of holes, distance between sensors and applying other methods for extracting desirable features like wavelet transform, cross correlation technics for measuring the lag distance and etc., the use of two key parameters—velocity and acceleration—was prioritized. These

parameters were selected based on their effectiveness in distinguishing between the responses of the baseline and damaged models (supervised learning method).

3.6 Algorithm for damage detection

There are important aspects that have been examined in this part regarding different works related to vibration-based structural health monitoring, including but not limited to how damage is represented within the structure and the selection of desirable features. It seems that, by definition, damage detection is largely a problem of understanding how the data for response from a damaged structure behaves in comparison to that from an undamaged structure. The subsequent sections describe how the processing of the data is performed, with some concepts and formulations related to providing more details on how this might be accomplished in damage detection.

3.6.1 Processing of Data

Selection of desirable characteristics to detect and locate damage is a matter of prime importance, as inappropriate selection could render the entire analysis ineffective. The feature characteristics must be different in nature, but their behaviors must show similarities to have accurate and reliable results. Two of the primary features that are subsequently applied to this project are velocity and acceleration. Both are inextricably linked, really, because instantaneous acceleration is defined as the velocity change over time. Here, the goal is to tie these two concepts together—velocity and its rate of change—to create data that effectively models the properties of each in providing a greater understanding of the dynamic performance of a system.

For this purpose, two types of signals are gathered from different points on the pipe, equally distant from one another. The defect-free pipe and the pipe containing a defect provide 12 signals for both velocity and acceleration. All such signals are different from one another, each reflecting pipelines velocity or acceleration regarding the specific point of monitoring. The key objective at this stage involves the processing of signals from the time series collected from various measurement points along the pipe.

For this purpose, two different techniques are considered during the detection phase: the first technique is based on the PCA method (Shlens, 2005), which relies on data fusion and investigates the first principal component.

The ICA method (Hyvärinen and Oja, 2000) is oriented towards data separation to complete it. In the localization step, which aims to identify defects by processing the signals between two consecutive sensors, both PCA and LMD (Local Mahalanobis Distance) methods are used. Each of these methods is described in more detail in the following sections.

3.6.2 Principal Component Analysis for detection

At this stage, the data extracted is organized into a matrix form in an Excel file. This matrix is a 101x12 matrix; the 101 number represents the number of rows—the number of samples of that data—and the 12 columns correspond to the 12 sensors collecting that data. In other words, it would be divided further into 6 columns for the velocity data and another 6 column for the acceleration data.

Including velocity and acceleration as raw data within a single matrix can also allow for the data to be correlated with one another, given that they are different physical quantities. The reason this

might be the case is that both sets of data, while different, are interrelated in such a way that one is a derivative of the other—that is, acceleration being the change in velocity over time. The nature of the data is different in this regard; hence, careful management of their relationships during the analysis is needed. Therefore, the aim here would be to remove the potential correlation that might exist among the data and reduce the dimensionality using the PCA technique.

The PCA belongs to this class of unsupervised learning. This reduces dimensionality with a method that the important knowledge from data should be retained.

The most salient feature of PCA is to capture most of the variance of the data. Put differently, PCA represents one of the statistical approaches in which a set of features selected can be linearly transformed to a set of principal components like PC1, PC2, in such a way that maximum possible variance is preserved (Tharwat, 2016). Figure 3.4. regarding first and second principal component and the mathematical relations associated with PCA are rendered below.

The first step in calculating PCA is to perform standardization. Since velocity and acceleration are two quantities with different units, it is crucial to ensure that they contribute equally to the analysis. Therefore, scaling is done using the standard score (Z-Score) formula, which standardizes the data. This ensures that each feature is treated equally during the PCA process, so Z-Score is as follow:

$$Z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (3.1)$$

In the equation (3.1), x_{ij} is shown for the value of the j^{th} variable for the i -th observation. The mean of the j -th variable is denoted by μ_j , and its standard deviation is represented by σ_j .

In the second step a covariance matrix is calculated, the purpose is to determine how the variables in the input dataset deviate from their mean values and how they relate to each other.

Essentially, it helps identify if there are any correlations between the variables. Often, variables may be strongly correlated, leading to redundant information.

$$C = \frac{1}{m-1} \sum (z_{ia} - \bar{z})(z_{ib} - \bar{z}) \quad (3.2)$$

Here, z_{ia} and z_{ib} are shown the variables located in the a -th and b -th columns of the i -th row of the matrix $z_{m \times n}$ and \bar{z} is the mean of matrix.

$$\bar{z} = \frac{1}{m} \sum_{i=1}^m z_i \quad (3.3)$$

In the equation (3.4) and (3.5), V and λ represent the eigenvectors and eigenvalues of the covariance matrix C , respectively.

The process of finding these eigenvectors and eigenvalues is carried out through eigen decomposition.

Since all elements in equation (3.4) are in matrix form, the eigen decomposition (equation (3.5)) is expressed using matrix operations to decompose the covariance matrix into its corresponding eigenvalues and eigenvectors.

$$Cv = \lambda v \quad (3.4)$$

$$C = v\lambda v^T \quad (3.5)$$

The method then orders the eigenvectors and their corresponding eigenvalues in decreasing magnitude of the eigenvalue. A new matrix, w , is then created by selecting t eigenvectors that have the highest eigenvalues, with $t < n$. Projected data—defined as principal component P —is then

achieved through a dot product between the matrix of eigenvectors w and the original data matrix z , as shown below:

$$P = w \cdot z \quad (3.6)$$

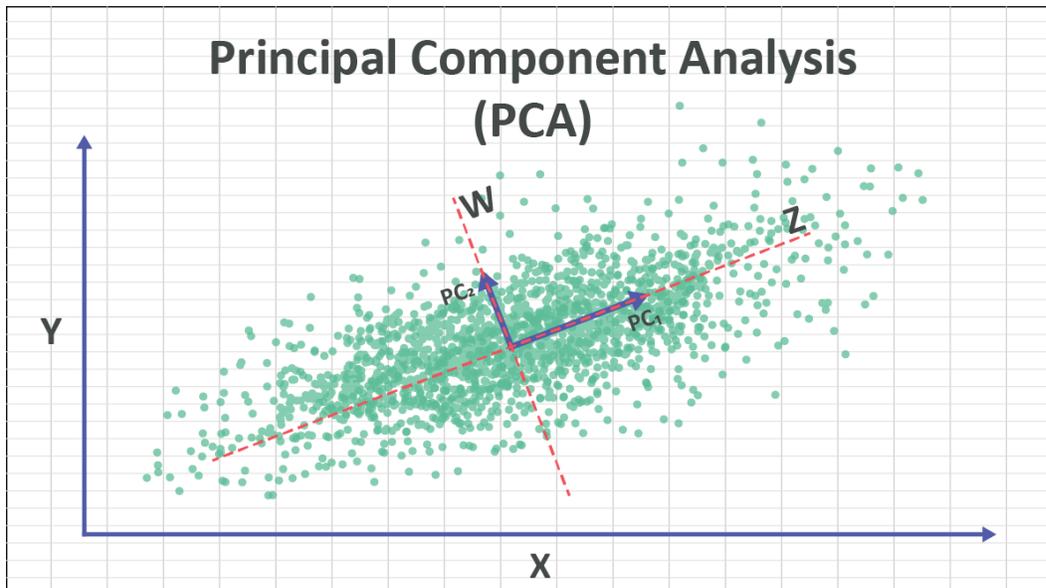


Figure 3.5. Principal components source: <https://numxl.com/blogs/principal-component-analysis-pca-101/>

3.6.3 Independent Component Analysis

Here, the rationale behind the detection process using ICA in pipelines will be explained. We begin, for an introduction, with the well-known cocktail party example. This is a generally well-known example in signal processing, especially in source separation. The example shall try to illustrate how ICA will go about isolating individual signals from a mixture and therefore will be useful in identifying and separating various sources (Hyvärinen and Oja, 2000). Imagine there are some people talking at a party, and at the same time, there are some microphones (same number) capturing their voices. Each microphone picks up a combination of sounds (signals). The power

of using ICA lies in its ability to separate these combined signals into distinct, individual signals. This way, ICA can determine which signal corresponds to each person, effectively isolating their voices from the mix. In Figure 3.6, the overall functioning of ICA is presented.

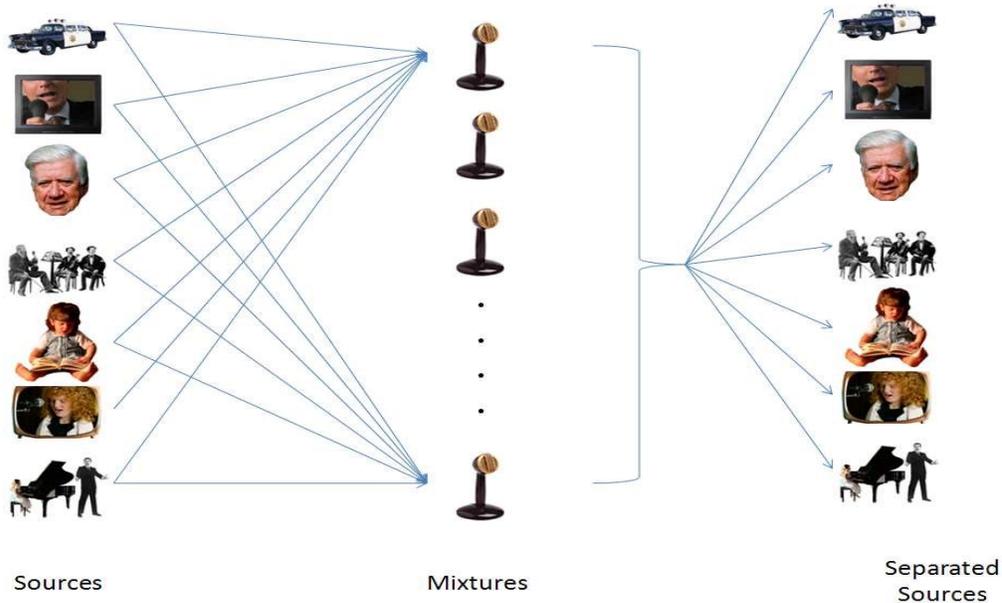


Figure 3.6. Function of ICA_ source: <https://onionesquereality.wordpress.com/tag/cocktail-party-problem/>

Here we explore the ICA technique to separate the signals that were previously combined using the PCA technique. This means that, through ICA, the process of separation is applied to multivariate signals (velocity and acceleration) to produce a set of components that are highly independent from one another and by selecting an appropriate feature—in this case, the greatest peak method—the independent components of the damaged pipe are compared with those of the baseline condition.

We, below, offer brief descriptions of typical first assumptions and other relationships that provide the essential mathematics behind FastICA.

3.6.3.1 FastICA Method

There are, in principle, two methods to maximize the statistical independence of the estimated outputs: either by minimizing the mutual information between the components or by maximizing their non-Gaussianity. The so-called FastICA algorithm follows the latter principle while considering the selfsame non-Gaussianity characteristics in order to render the components as independent as possible (Hyvärinen and Oja, 2000).

The preprocessing is composed of two important processes, namely centering and whitening. In the centering step as the first step, the data is translated to have a zero mean, which is easier to work with. According to Equation (3.7), let x be the observed data and \bar{x} be the centered data obtained by subtracting the mean.

$$\bar{x} = x - \text{mean}(x) \quad (3.7)$$

In the whitening part as the second step, the centered data is transformed to make the components uncorrelated and ensure they each have unit variance (as velocity and acceleration are correlated to each other the process of whitening through PCA is necessary for removing this correlation). This transformation is achieved by performing eigenvalue decomposition on the covariance matrix of the data.

$$\text{cov}(\bar{x}) = EDE^T \quad (3.8)$$

In Equation (3.8), E represents the matrix of eigenvectors, while D is a diagonal matrix containing the eigenvalues. The data is then whitened (z), which is derived from the transformation using these matrices.

$$\mathbf{z} = \mathbf{D}^{(-1/2)} \mathbf{E}^T \bar{\mathbf{x}} \quad (3.9)$$

For the following step which is Choosing a suitable function and initialize weight vectors, The key idea behind ICA is to find a transformation W such that the components of Equation (3.10) are as non-Gaussian as possible. One way to measure non-Gaussianity is to use the kurtosis or negentropy while here negentropy is used. For a random variable y ,

$$\mathbf{s} = \mathbf{W} \cdot \mathbf{z} \quad (3.10)$$

$$J(y) = H(y_{\text{gauss}}) - H(y) \quad (3.11)$$

where H is the entropy and y_{gauss} is a Gaussian variable with the same covariance as y . The higher the value of negentropy, the higher the non-Gaussianity. In this step of weight-update in a fixed-point algorithm, FastICA uses a fixed-point iteration scheme so that it can maximize non-Gaussianity. This iteration can be written as:

$$\omega^+ = \mathbb{E}\{\mathbf{z}g(\omega^T \mathbf{z})\} - \mathbb{E}\{g'(\omega^T \mathbf{z})\}\omega \quad (3.12)$$

Where ω is the weight vector, g is a nonlinear function, often chosen as Equations (3.13) or (3.14) and g' represents the derivative of g .

$$g(u) = \tanh(u) \quad (3.13)$$

$$g(u) = u^3 \quad (3.14)$$

In this step the weight vectors should be Normalized and orthogonalized, that is, after each update, the weight vector ω is normalized.

$$\omega = \frac{\omega}{\|\omega\|} \quad (3.15)$$

The process should be Repeated until the independent components are identified. To ensure that each weight vector finds a distinct independent component, they need to be orthogonal to each other. Once the process converges, the weight vectors are orthogonalized using the Gram-Schmidt method.

$$\omega_i = \omega_i - \sum_{j=1}^{i-1} (\omega_i^T \omega_j) \omega_j \quad (3.16)$$

Once the weight matrix W is obtained, the independent components s can be extracted by:

$$s = W.z \quad (3.17)$$

3.7 Algorithm for Damage Localization

After completing the damage detection process, the next step is localization or determining the possible location of the damage. This part, like the previous one, involves considering two models: one of the damaged pipe and the other of a healthy (intact) pipe.

Each model is equipped with sensors placed at equal intervals, and these sensors, or elements, are responsible for recording acceleration and velocity data. Acceleration and velocity signals are extracted from both the damaged pipe (with a hole) and the intact pipe, and the data analysis process begins to determine the location of the damage between the sensors.

The core method of data analysis in this section is Mahalanobis distance. Based on Figure 3.7, unlike the Euclidean distance, which considers a relationship between two single points about the correlation between them, Mahalanobis distance is a statistical concept that calculates the relationship between several points and one distribution or a set of points. Another advantage of

Mahalanobis distance is that it can process multi-dimensional signals. This aspect of flexibility makes it suitable for the analysis of complex data where different types of measurements are concerned. Normally, a spike in Mahalanobis distance is marked as an anomaly; that means a point has a high difference from the normal pattern and can show some abnormality or irregularity in the data.

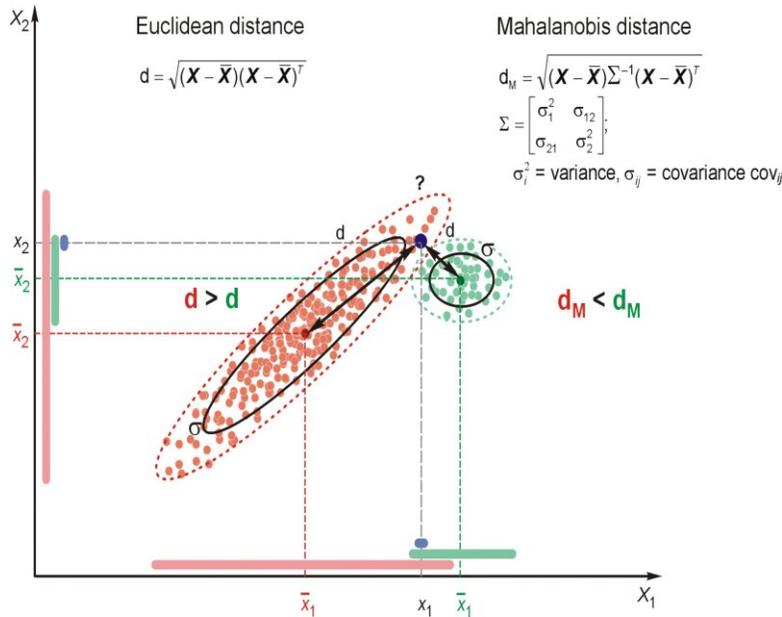


Figure 3.7. Mahalanobis distance and Euclidean distance in one frame_ source: <https://ouzhang.rbind.io/2020/11/16/outliers-part4/>

3.7.1 Principal Component Analysis for Localization

As it was discussed previously, the choice of appropriate features for damage localization matters. In this project, velocity and acceleration were chosen as the desirable features because they are interrelated parameters: acceleration is defined in terms of velocity and vice-versa. This relation makes them very powerful in picking out the key characteristics required for good localization.

Acceleration is the instantaneous rate of change of velocity at any point in time, so in this case we do data fusion for two variables: velocity and its instantaneous rate of change. Performing PCA

for data fusion of these two variables serves the purpose not only because of dimensionality reduction but also because it produces a new dataset that concentrates and retains information of highest importance, namely the maximum variance in this context. The mathematical equations related to PCA were presented in Section 3.6.2, so they are not included in this section.

3.7.2 K- nearest neighbor

k-NN is a nonparametric statistical method that can be applied for regression and classification. In other words, in parametric methods, the model is defined by a set of parameters, such as weights, which are estimated during training and subsequently used for decision-making. Nonparametric models do not depend on predefined parameters like weights. Therefore, such methods are applicable even for small-sized datasets.

Non-parametric methods are instance-based. In non-parametric methods, the model relies upon labeled examples. Such labeled examples then serve as the basis for decisions. Because of this reason, such techniques are also known as instance-based learners. In instance-based learning for a variable x , the process of decision-making involves selecting k of the nearest instances and finding out which class recurs again and again. This recurring class is assigned to the variable.

One important advantage of the k-NN algorithm is that it is nonlinear, meaning that the decision boundary given by k-NN can be very complex and is not confined to a linear state. This technique strongly relies on the choice of the values of k , for which usually odd numbers are preferred in order to avoid any ties at the time of classification. To define the decision boundary, a neighborhood is assigned to every point such that all the points within that neighborhood have the given point as their nearest neighbor. In short, Voronoi tessellation is done, and with an increment

in k , the decision boundary between different classes, represented by the Voronoi tessellation, becomes smoother.

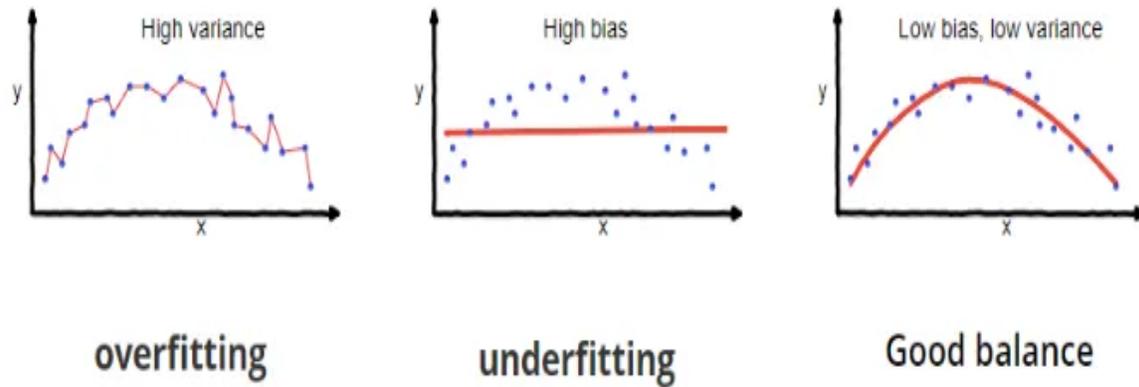


Figure 3.8. Different state of fitting based on k value_ source: <https://medium.com/30-days-of-machine-learning/day-3-k-nearest-neighbors-and-bias-variance-tradeoff-75f84d515bdb>

3.7.3 Mahalanobis Distance

The goal of this section is to present the mathematical equations related to the Mahalanobis distance in a concise manner. The general formula for Mahalanobis distance is shown in Equation (3.18) below, where x represents a multivariate vector with a mean of μ and Σ as in Equation (3.19) is the covariance matrix (Ghorbani, 2019).

$$D_m(x) = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)} \quad (3.18)$$

$$\Sigma = \begin{bmatrix} \text{Var}(X) & \cdots & \text{Cov}(X, Z) \\ \vdots & \ddots & \vdots \\ \text{Cov}(Z, X) & \cdots & \text{Var}(Z) \end{bmatrix} \quad (3.19)$$

$$\text{Var}(x) = \sum (x_i - x_{ave})^2 / (n - 1) \quad (3.20)$$

$$\text{Cov}(X, Y) = \sum (X_i - X_{ave}) * (Y_i - Y_{ave}) / (n - 1) \quad (3.21)$$

As could be seen in Equation (3.19), the Σ includes some element which based on Equation (3.20) and (3.21) are Variance and Covariance.

3.7.4 Local Mahalanobis Distance

Unlike the standard Mahalanobis distance that relies on a global covariance matrix with the aim of measuring the sensitivity of data, in Local Mahalanobis Distance, LMD primarily focuses its attention on analyzing sensitivity in a local manner. In this approach, local statistical techniques such as k-nearest neighbors used throughout this project estimate the sensitivity of data on a local

basis rather than on a global one. This is an important approach because, instead of taking the entire dataset for examination and a one-time calculation of the Mahalanobis distance, datasets will be split into smaller subsets. Then the Mahalanobis distance computation is performed multiple times on these subsets to allow a more local and fine-grained evaluation of sensitivities within the data. This processing of data leads to a better sensibility of the anomalies detection. In order to compute LMD (Local Mahalanobis Distance) as in Eq. 22, local mean as in Equation (3.22) and local covariance matrix as in Equation (3.23) are calculated utilizing the computational power of k-nearest neighbors (k-NN) based on the explanation stated in section 3.7.2 as follows.

$$\mu_{loc} = 1/k \sum_{x_i \in N_k} (x_i) \quad (3.22)$$

$$\Sigma_{local}(x) = 1/(k - 1) \sum_{x_i \in N_k} (x_i - \mu_{local}) (x_i - \mu_{local})^T \quad (3.23)$$

$$D_{lm}(x) = \sqrt{((x - \mu_{local})^T \Sigma_{local}' (x - \mu_{local}))} \quad (3.24)$$

Summary

This chapter has presented the details of the steps taken for damage detection and localization. First, the experimental model was presented as the benchmark along with presenting the characteristics of modes 1 through 4. Then, the construction of the numerical model was explained in order to make feature extractions that are desirable.

Such steps involve modal analysis to extract mode shapes for the calibration of the numerical model with the experimental setup and also transient analysis to get the desirable features. Following the construction of the numerical model and identification of desirable features, a process of data separation and data fusion is presented to execute the detection and localization tasks together with the mathematical equations involved.

Chapter 4: Results and Discussion

4.1. Introduction:

This chapter presents the analysis of five different pipeline scenarios. Later, the performance of the proposed methods for damage detection and localization is evaluated using data extracted from ANSYS software, considering various scenarios based on time history of desirable features, including the effect of changing the hole location, distance between sensors, hole size, and the impact of two simultaneous holes on the pipeline. In the next part, after explanation of scenario, some sample outputs of one of the models are shown for illustrative purposes before providing the results on damage detection and localization.

4.2 First Scenario for Detection and Localization

The model specifications will be explained first in order to explain the first scenario. This first scenario, therefore, explains these specifications in detail, while most of the repetitive details are omitted in the ensuing scenarios to avoid redundancy. Two 6-meter pipes with identical characteristics are modeled in this section. The first model is a 6-meter pipe without any holes - intact and serving as the baseline. Another model is also a 6-meter pipe, but it includes a hole with a radius of 1 mm that is positioned 4.4 meters from the starting point of the pipe. It is considered that the pipe will be subjected to a dynamic load of 300 N, placed 15 cm from the beginning of the pipe. In conformity with Figure 3.2 this load zooms during a time interval starting from $1e-3$ sec up to $1.1e-2$ seconds. The model shall be studied in the period of 0 to 0.1 seconds with the time

step of 0.001 seconds. Figure 4.1 shows that there are six acceleration-measuring sensors and also six velocity-measuring sensors distributed along the pipe at equal distances.

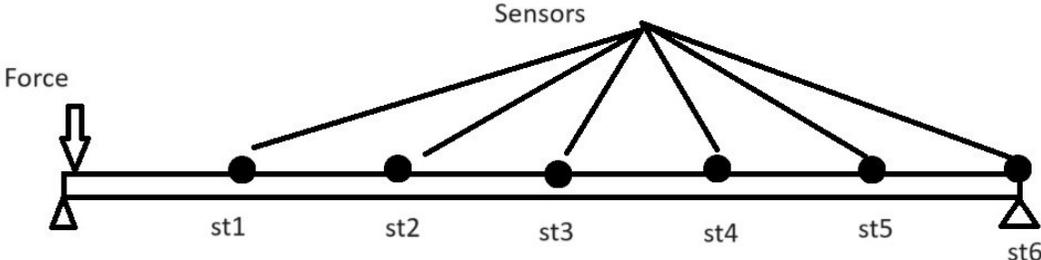
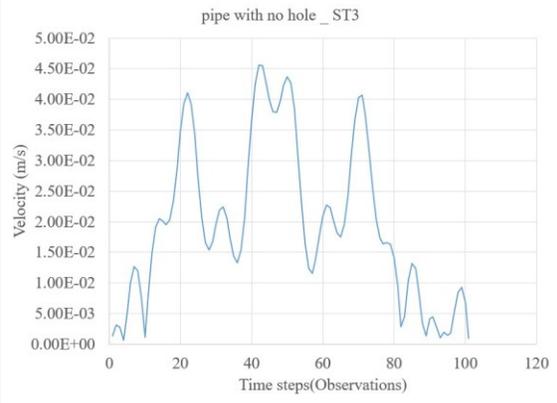


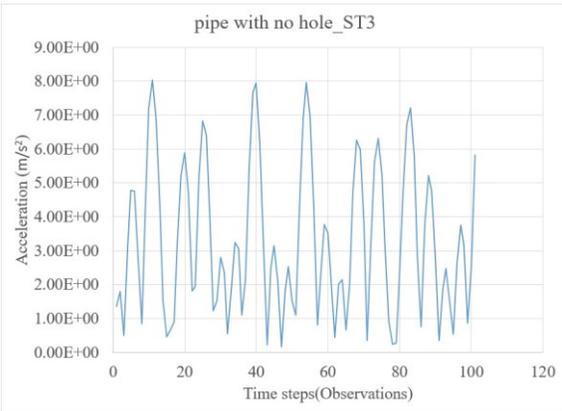
Figure 4.1. Schematic figure of the intact pipe with sensors and the force

4.2.1 Data Acquisition

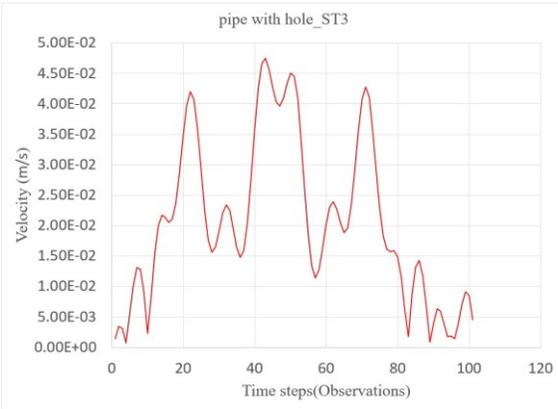
As has been discussed in a previous section, each numerical model has 12 sensors which catch the data about velocity and acceleration. Before diverting into the topics of damage detection some examples of raw data are given in the following figures. These figures illustrate acceleration and velocity of a sensor that is located 3 meters away from the beginning of the pipe. Because of the great volume of the data, here only a sample of them is shown here, and the rest are brought in the appendix.



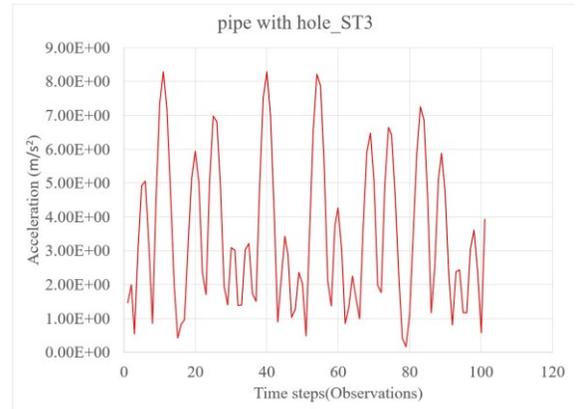
(a)



(b)



(c)



(d)

Figure 4.2. Examples of time series data regarding acceleration and velocity at 3m from the beginning of the pipe (a) and (b) acceleration and velocity for pipe with no hole. (c) and (d) acceleration and velocity for the pipe with hole which located 4.4m from the beginning of the pipe

Examples of raw acceleration and velocity time series data extracted from a sensor located 3 meters from the start of the pipe are shown in Figure 4.2. From these examples and considering that there is a hole of radius 1 mm, which is 4.4 meters from the start of the pipe, the acceleration and velocity graphs show very negligible noticeable differences between intact and damaged states. Both sets of data are very much alike.

4.2.2 Damage detection

As explained in Chapter 3, the process of damage detection involves a two-step process. This stage covers the analysis of data extracted from the software, following the procedure of data fusion through Principal Component Analysis (PCA) on the velocity time history and acceleration time history extracted from the software. Each pipe has 12 signals—6 signals obtained from velocity and 6 signals obtained from acceleration. The reduction through data fusion has reduced these signals to 6 signals. In this way, the characteristics of the data may be preserved with the reduction

of complexity in data. For it selects the first principal component in which the variance is effectively guided, which is very important for the detection of anomalous behavior.

The goal of the next step is to transform the 6 principal components into 6 maximally independent components. In Chapter 3, when considering ICA for the first time, two of the most salient features regarding the 6 principal components were that they are statistically independent and that they have non-Gaussian properties. Therefore, data separation is conducted in order to enhance those independent features.

The other reason for applying Principal Component Analysis in this project is that Independent Component Analysis requires pre-processing in order to remove the correlation—a process well-known as whitening. It has always been the case that PCA is one of the most common methods applied for whitening. Hence, detection here is meant to extract the desirable features in the most efficient methodology. Therefore, the greatest peak was chosen as the desirable feature technique, as it normally represents arrival or conversion points in wave modes within complex signals.

4.2.2.1 Calculating PCA

At this stage, Principal Component Analysis (PCA) is separately conducted for the damaged pipe and the intact pipe, where velocity and acceleration signals are combined at each station. Based on the data matrix provided in the appendix, the velocity and acceleration signals coming from sensors positioned at distances of 1, 2, 3, 4, 5, and 6 meters are combined. Figures (4.3) and (4.4) shows samples of the projected data onto the first principal component (since the first principal component (PC1) represents more than 90% of the variance in both velocity and acceleration, other components especially (PC2) are not considered in this stage). In the sequel, the principal

components extracted from each location are arranged in a matrix to carry on the Independent Component Analysis (ICA).

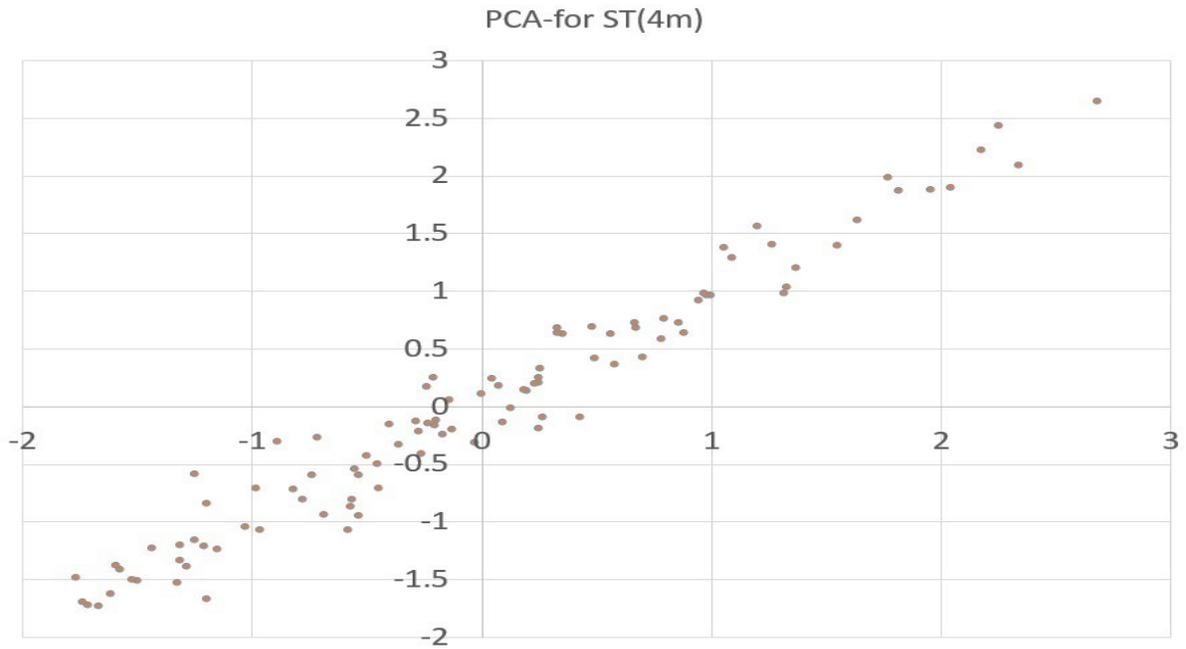


Figure 4.3. PC1 and PC2 for ST(4) for both no hole and with hole data

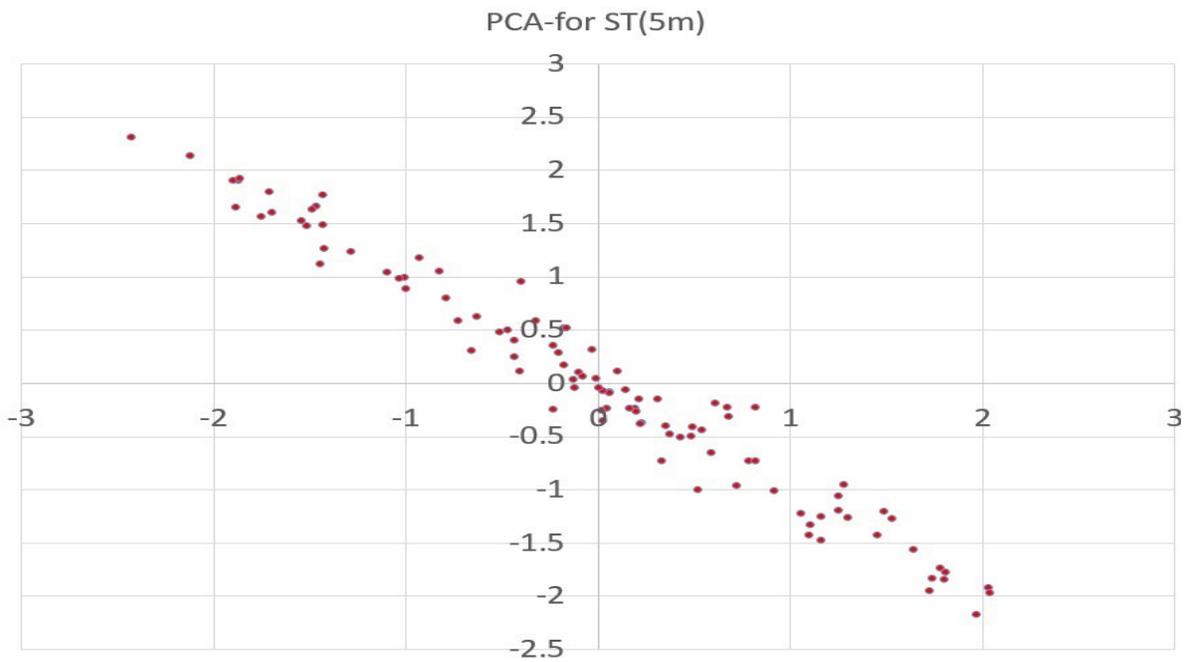


Figure 4.4. PC1 and PC2 for ST(5) for both no hole and with hole data

First, it should be mentioned that since velocity and acceleration are of different dimensions, the data was normalized by performing z-score scaling as described in chapter 3. Also, as it has already been described, the first principal component was dominating, the plot of PC1 against time is provided in Figure 4.5.

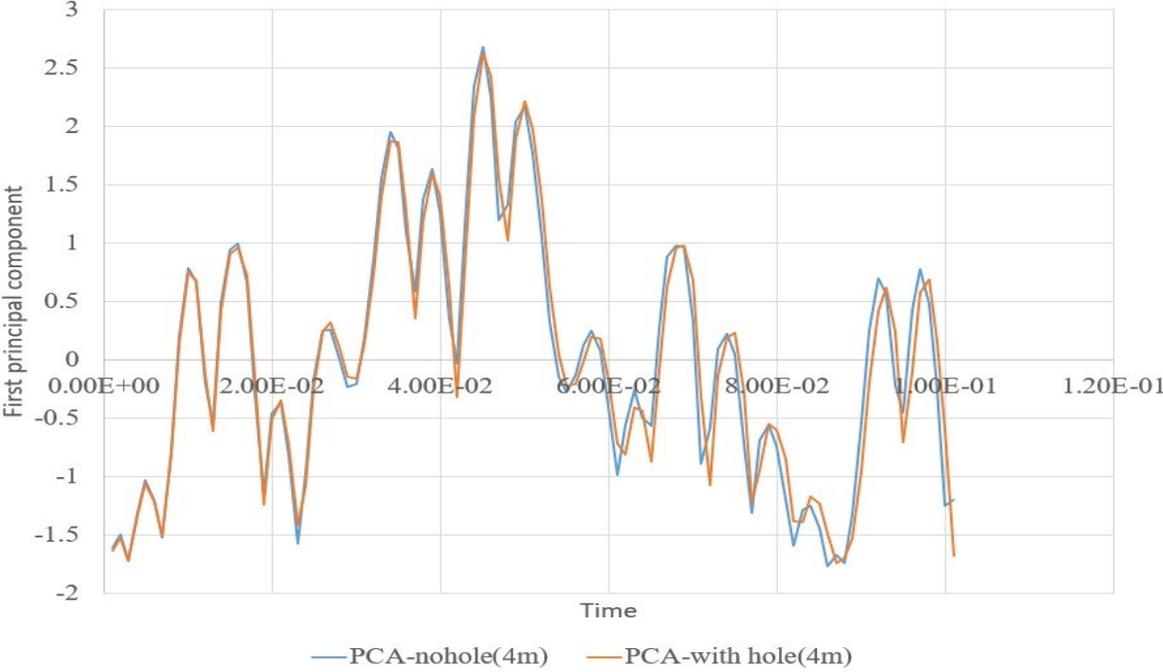


Figure 4.5. PC1_time for ST(4) for both no hole and with hole data

4.2.2.2 Calculating ICA

As mentioned in section 4.2.2, in this stage, the Independent Component Analysis (ICA) method will be applied to each of the first principal components retrieved in the previous stage. According to the discussions presented in Chapter 3, the maximize non-Gaussianity approach will be adopted to extract the independent components using FastICA since it would produce maximally independent components. Since all pipes, which were either damaged or intact, are with 6 principal

components each, the same number, i.e., independent components, will be produced using this technique as below Figure.

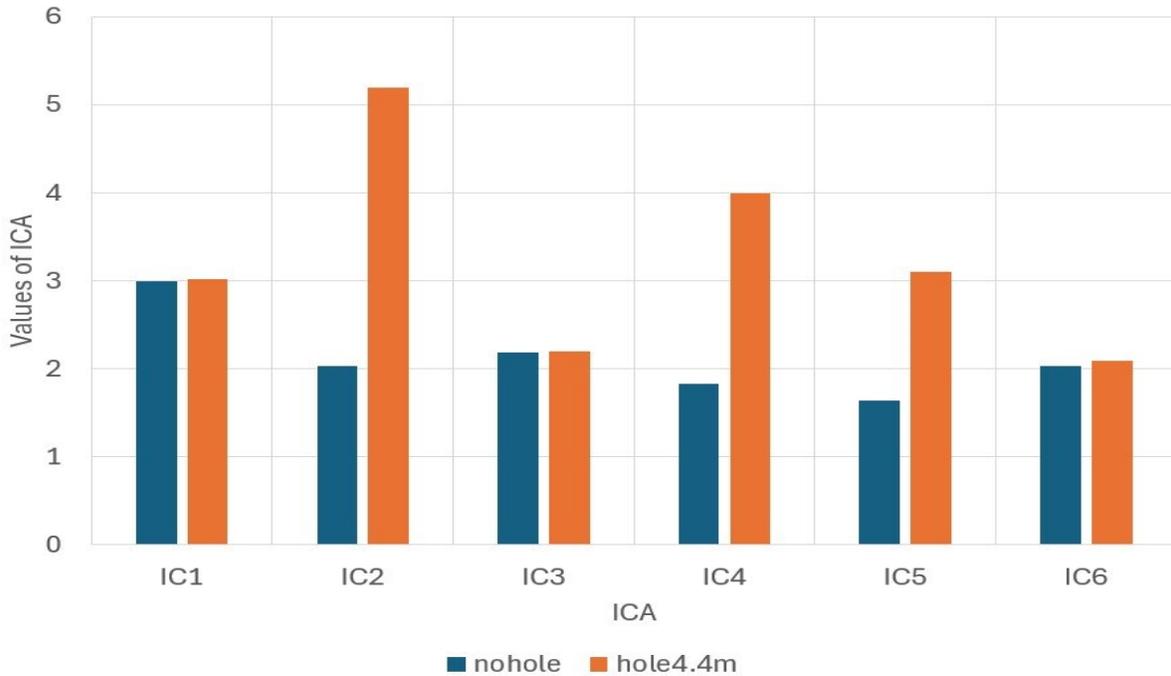


Figure 4.6. Independent components for both pipe with hole and without hole

Looking at Figure 4.6, components IC2, IC4, and IC5 show a large percentage of difference from the pipe with a hole to the pipe without. For instance, the value of IC2 is about 5 for the pipe with a hole, while for the intact pipe, it is 2—a difference of over 100%. While other components such as IC1, IC3, and IC6 show approximately similar values from both the damaged pipe and the intact one. For instance, IC6 shows an almost identical value of 2 for either condition.

Let's delve into another aspect, the two pipes represent, respectively, two people who have different voice tones in explanations given in using the cocktail party analogy in Chapter 3. Using two different receivers, a mix of both individual sounds is recorded. The power of Independent

Component Analysis (ICA) is to associate sounds with the right person. In this regard, if the healthy pipe is taken as our reference, any disagreement from the independent components of the healthy pipe will indicate a defect or damage, thus making damage detection in the pipe possible.

This finding might suggest that some ICs are insensitive to the presence of a hole, as they tend to give very similar values regardless of the pipe's state. To further verify the sensitivity of these ICs, a pipe with two different holes will be investigated in the succeeding sections.

4.2.3 Damage Localization Through LMD

This section calculates the local Mahalanobis distance (LMD) for each of the first principal components derived in Section 4.2.2.1. The concept of using the first principal components from both the intact pipe and pipe with a hole is to further refine the data by fusing the velocity and acceleration which are our desirable features. As discussed in the prior chapter, whereas Mahalanobis distance relies on the usage of a global mean and global covariance, Local Mahalanobis distance uses local mean and local covariance, which enhances the sensitivity significantly. To this end, the k-NN algorithm assigns each data point in the dataset to its nearest neighbors. For this identified nearest neighbor, the local mean and local covariance are calculated, which are used in the computation of the local Mahalanobis distance. Therefore, Mahalanobis distance is computed from every data point from the local distribution. Taking advantage of k-NN, Mahalanobis distance can be computed several times for each data point in the dataset, with an advisable k value around 5 to 6 percent of the data points, according to some references (Sarmadi et al., 2020).

In reference to Equation (3.21) in Chapter 3, regarding the covariance matrix, the feature that makes the Mahalanobis distance far more powerful than the Euclidean distance is its ability to

make measurements of the dispersion of data relative to their mean as well as how two variables or data points vary in relation to each other. This capability is made possible with a covariance matrix since it embodies both variance and covariance in one matrix.

The strategy followed for localization on this project intends to localize the defect between two consecutive sensors with aid from signal processing. To this effect, the proposed method for the interpretation of the results involves collating every three consecutive sensors into a single group.

Before going deep into the methodology of sensor grouping and result interpretation, it has to be underlined that in each of the intact and defective pipes there are 12 sensors measuring velocity and acceleration. For each station, the acceleration and velocity data have been fused using Principal Component Analysis (PCA) in order to retrieve the first principal component. So, for each defective pipe, the six first principal components will be compared with the six first principal components of the intact pipe using Local Mahalanobis Distance in the computation process.

After calculating the Local Mahalanobis Distance corresponding to each sensor and the baseline, the greatest peak method is applied in order to extract the peak value from the Local Mahalanobis Distance index. This method assigns a particular numerical value to a sensor because of its peak distance index.

It is the stage where peak values obtained for each sensor are grouped. According to sensors ST(1) through ST(6), four different groups are defined as shown in Figure 4.7. The group 1 includes ST(1), ST(2), and ST(3); Group 2 includes ST(2), ST(3), and ST(4); Group 3 includes ST(3), ST(4), and ST(5); and Group 4 includes ST(4), ST(5), and ST(6).

This is where, as already said, the highest peak output of each of the signals should be considered as the desired feature in this step. Maximum values in a signal indicate the arrival or transformation

points within a given signal because some peaks appear to get affected differently compared to other peaks. In Group 4 of Figure 4.7, three consecutive phases are visible: the build-up phase, the spike phase, and the decline phase.

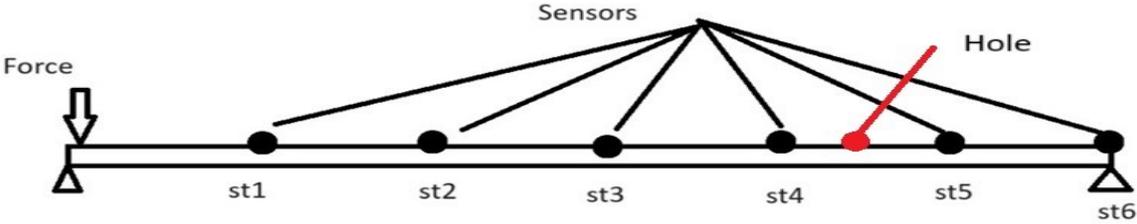
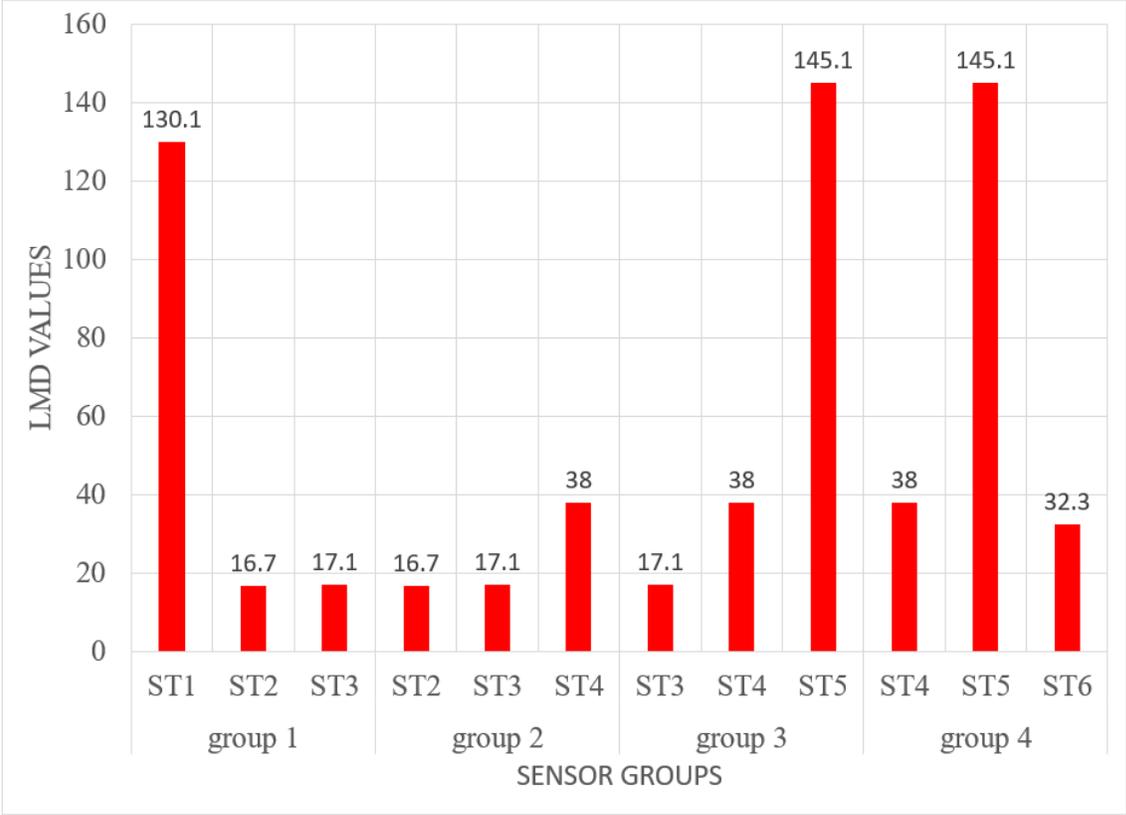


Figure 4.7. LMD for both pipe with hole and without hole

It is worth mentioning that different approaches for structuring the extracted data were tried. Finally, inspired by a classic earthquake sequence with three phases, build-up, spike, and decline, the groups were analyzed. The method proposed in this project predicts the defect between the

build-up and spike phase. Figure 4.7 shows that Group 4 behaves coherently according to the proposed methodology. This group has ST(5) representing the sensor at Station 5m, giving an LMD index of 145.1 for the spike phase observed over the period of time. Furthermore, its predecessor and successor give LMD indices of 38 and 32.3 for the build-up and decline phases, respectively, fitting well in these three defined phases. In Groups 1 and 3, however, the three defined phases are not observed: during the decline phase of Group 3, the maximum value is reached, while for Group 1, it happens during the build-up phase. Looking at the modeled pipe, it is clearly seen that the location of the hole is at 4.4 meters from the starting of pipe, exactly between the build-up and spike phases between sensors ST(4) and ST(5), proving the efficiency of the proposed method.

4.3 Second Scenario for Localization

As in Section 4.2, all model specifications—length, material, mode shapes, dynamic load application, and sensor placement—are identical except that this section has the hole location moved 3.4 meters from the start of the pipe. This scenario aims to revalidate the proposed method for damage localization, adding the feature of presenting LMD graphs over time so that a clearer view of the behavior of the LMD index and its effectiveness regarding the defect detection can be noticed.

4.3.1 Graphs for Local Mahalanobis Distance

In this section, model set up is the same as in Scenario 1. In both pipes under study, one intact and the other with a hole located 3.4 meters away from the start, sensors are placed in an identical arrangement to that used in Scenario 1. Under the action of a dynamic load of 300 N applied at 0.01 seconds, velocity and acceleration are measured in a timeframe of 0 to 0.101 seconds with a time step of 0.001 seconds. Firstly, data fusion is done through Principal Component Analysis,

showing how the first principal components can be calculated for each sensor. It captures the maximum invariance of the data, so calculation includes the computation of the covariance matrix, computation of eigenvectors and eigenvalues, projection of the data onto these eigenvectors to operate the principal components, and selection of the principal components based on the computed eigenvalues. The LMD values of sensors ST(3), ST(4), and ST(5) for the given time window are plotted in Figures 4.8 and Figure 4.9. Figures in each, show the LMD values in the selected time window by two successive sensors, and their maximum peak is monitored for each sensor. For instance, ST(3)'s maximum magnitude is 31.5, for ST(4) it is 135.4, and for ST(5) it is 43.8. The power of LMD is in amplifying the tiniest difference between the first principal components corresponding to the damaged pipe and the intact pipe. These graphs are intended to illustrate this amplification—as best as possible—driven principally by the covariance matrix behavior. This magnification can considerably enhance defect detection by emphasizing minute differences.

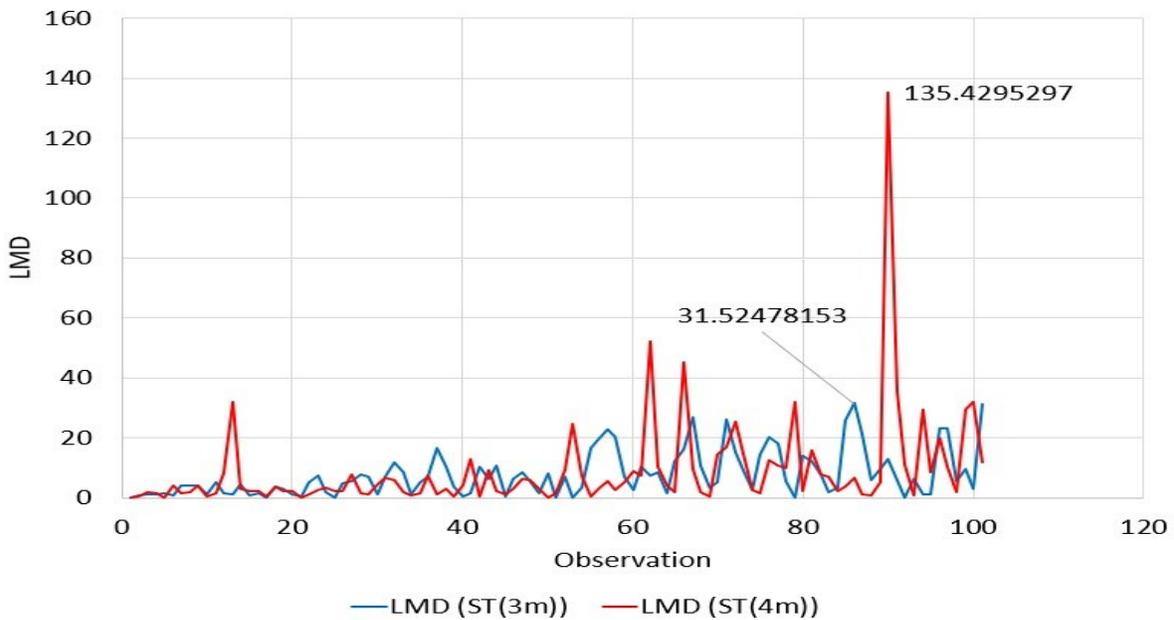


Figure 4.8. LMD for both pipe with hole and without hole

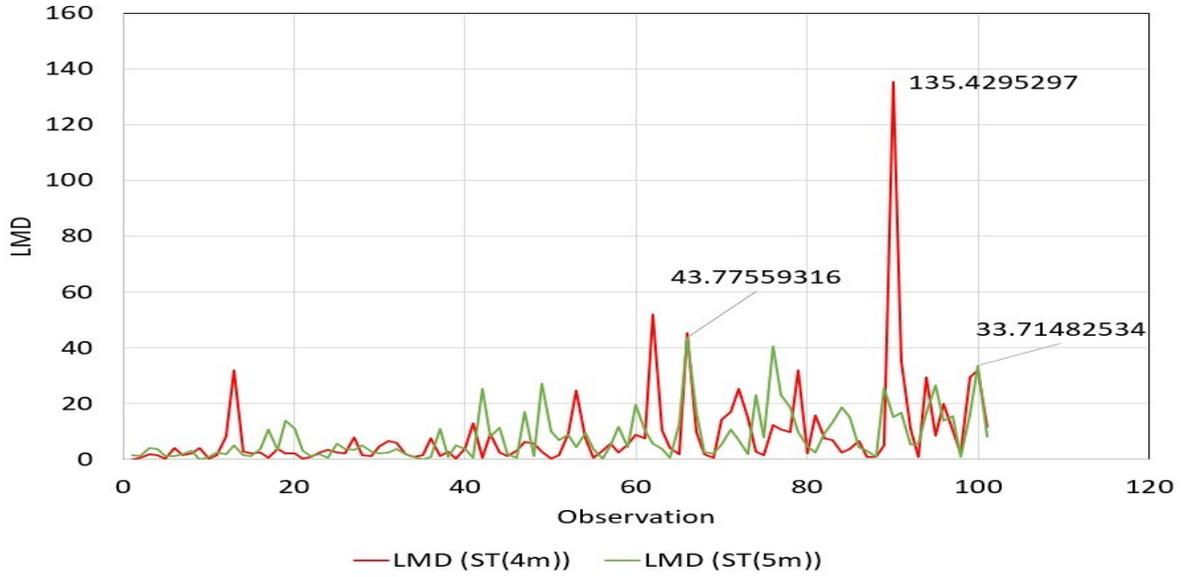


Figure 4.9. LMD for both pipe with hole and without hole

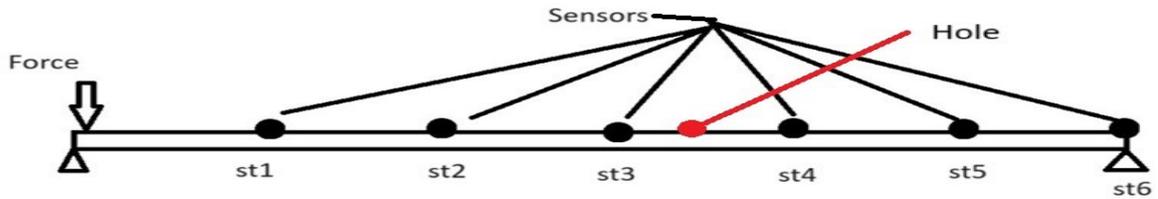
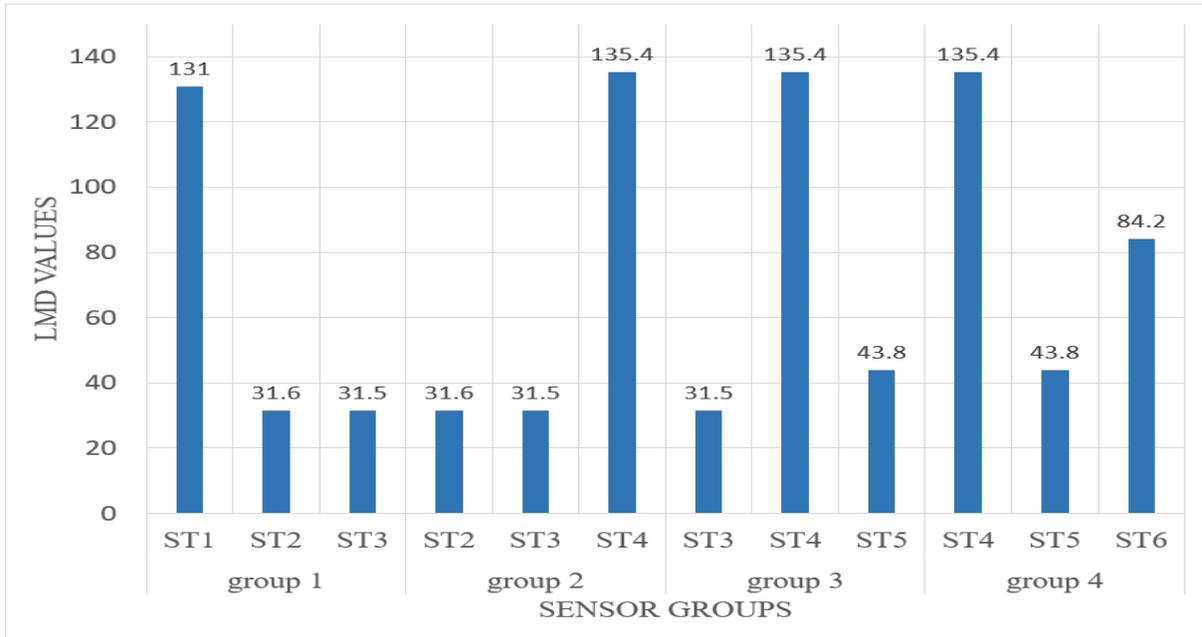


Figure 4.10. LMD for both pipe with hole and without hole

In Figure 4.10, the sensors_ from ST(1) to ST(6)_ are arranged in sets of three. The procedure followed here is in accordance with that in Section 4.2.3, where every column corresponds to the greatest peak LMD. Considering Groups 1 to 4 and examining the buildup, spike, and decline phases using the approach suggested in this research, Group 3 reflects what is expected. Specifically, for Group 3, the LMD value is 31.5 in the buildup phase, 135.4 in the spike phase, and 43.8 in the decline phase. Using this method, a defect is likely located between the buildup and spike phases. Since the hole is 3.4 meters from the beginning, this method accurately predicts the defect location for this model.

This trend is not followed by the rest of the groups. For example, in Group 1 and Group 4, the peak values of LMD are 131 and 135.4, respectively, and lie in the buildup phase. In Group 2, the peak value of 135.4 lies in the decline phase, which does not agree with the sequence of phases as proposed.

4.4 Third Scenario for Detection

In this section, the sensitivity of independent components to two simultaneous holes will be assessed, as previously discussed in Section 4.2.2.2. Model specifications remain consistent with Scenarios 1 and 2, with the only change being the addition of two holes: a 1 mm hole at 2.4 meters and a 1.5 mm hole at 4.3 meters from the pipe's start.

A 300 N dynamic load is applied, and desirable features, specifically velocity and acceleration, are selected from Stations 1 through 6. To analyze these selected data, data fusion is conducted using Principal Component Analysis (PCA) for dimensionality reduction and feature extraction. As noted earlier, PCA is employed as a preprocessing step to eliminate signal correlation, also known as signal whitening.

For defect detection at this stage, data separation is carried out using FastICA to extract independent components, with a focus on maximizing non-Gaussianity. A key difference between ICA and PCA is that, unlike PCA—where the number of inputs will be reduced in the outputs (e.g., reducing 12 selected features to 6 for each pipe in the previous section)—ICA retains an equal number of inputs and outputs.

4.4.1 Graphs for ICA

For this section, we have 6 inputs (PC1 to PC6) for each pipe, so, through ICA 6 outputs will be extracted as desirable features (IC1 to IC6). An important aspect for using ICA is that if the independent components (ICs) of two pipes (damaged or intact) are similar, it indicates that both pipes are likely in a similar condition. On the other hand, if the ICs are different, by using one pipe as a baseline, the defect in the other pipe will be detected.

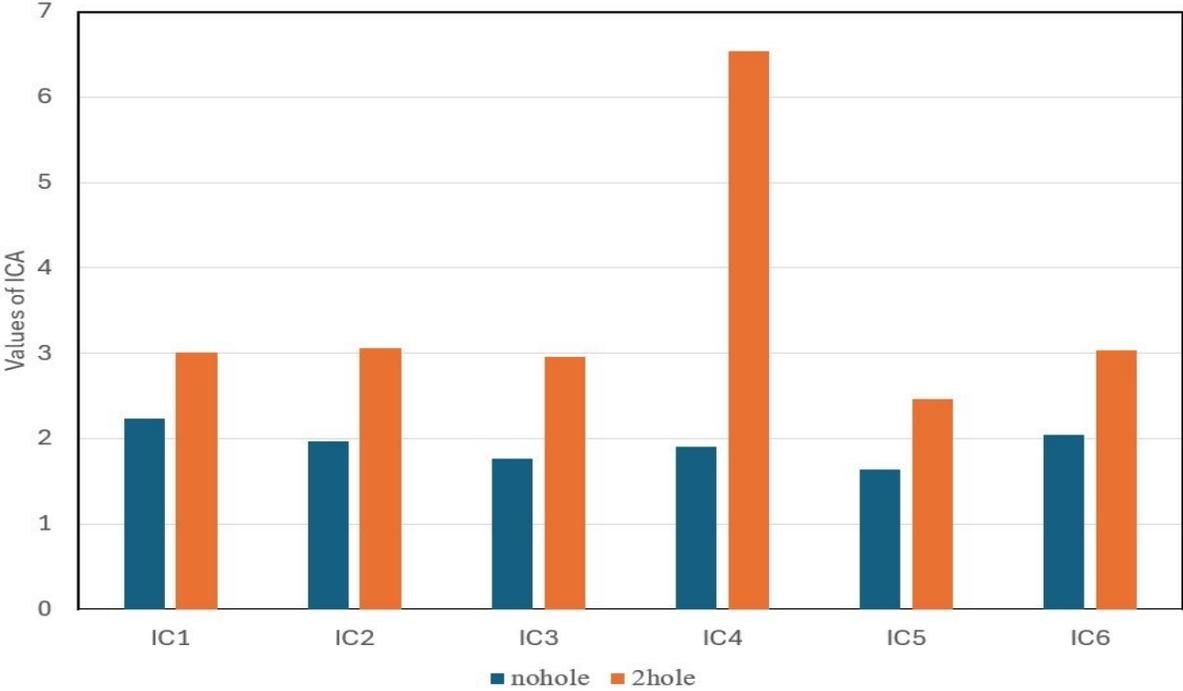


Figure 4.11. ICA for both pipe with hole and without hole

According to Figure 4.11, where the blue columns represent the intact pipe as a baseline, increasing the number of holes to two clearly highlight differences in the values of IC1 through IC6 between the two pipes, indicating that all components are sensitive to the presence of defects. For instance, IC1, IC2, IC3, IC5, and IC6 shows a 50% difference between the intact and defective pipes, while IC4 shows the difference for more than 100%. Also, in this scenario LMD values for pipelines with two holes are also investigated and as observed from Figure 4.12 the pattern of start, spike and decline phases could not be detected from this graph, which means that due to different reasons like the signal interference damage localizing could not be confirmed.

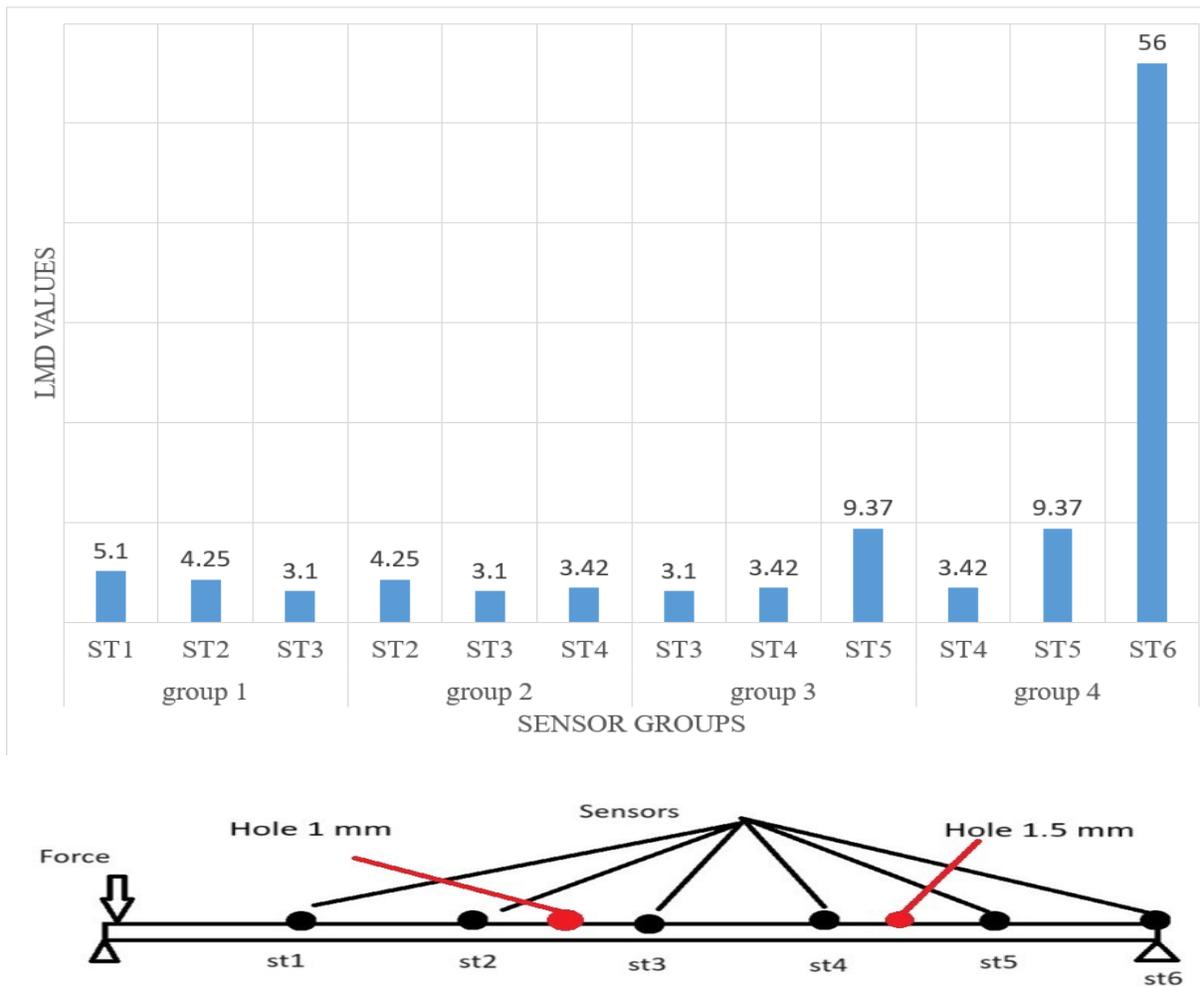


Figure 4.12. LMD for both pipe with hole and without hole

4.5 Fourth scenario for sensitivity

This section is devoted to the analysis of the performance of the proposed method in defect localization by considering the increase of distances between sensors on the pipeline. Consequently, by increasing the sensor spacing by 50%, the sensors are placed successively at 1.5 m, 3 m, 4.5 m, and 6 m intervals. All the other conditions are the same, for example, the characteristics of the pipe and the application of dynamic load. Acceleration and velocity sensors will be attached for both pipes at ST(1.5), ST(3), ST(4.5), and ST(6). One pipe is intact, free of holes, used as a reference; the other has a 1 mm hole, which is located 3 meters from the start of the pipe. Of particular interest in this setup is not only to study the effect of increasing sensor spacing but also to verify the performance of the proposed methodology in detecting defects at the location of the sensor. At present, the selected desirable features pertain to the four sensors of both velocity and acceleration components; hence, there are eight features for each pipe. For localization proposed here, PCA is adopted in order to conduct data fusion with dimensionality reduction. Hence, each pipe now has four first principal components. Using the LMD technique, a pair of the two first principal components at each station, one from the baseline pipe and one from the target pipe, are compared. The highest peak method represents the result of this comparison at each station as one value. Hence, for localization, there are four measurements generated at the four sensors. From Figure 4.13, it can be observed that the four sensors are split into two groups: group 1 consists of ST(1.5), ST(3), and ST(4.5), and group 2 is made up of ST(3), ST(4.5), and ST(6). It is observable from Figure 4.13 that Group 2 reflects all three consecutive phases: build-up, spike, and decline, with values of 6.9, 14.7, and 8.1, respectively. The proposed method estimates the position of the hole to fall between the build-up and spike phases, which also satisfies the assumptions made for the problem. In Group 1, the peak value of 14.7 falls in the decline phase; hence, it is not suitable for our study.

4.5.1 Graphs for Local Mahalanobis distance

Based on the presence of sensor in 3 m from the start of the pipe, which is a special state, the values of LMD are shown below (Fig. 4.13).

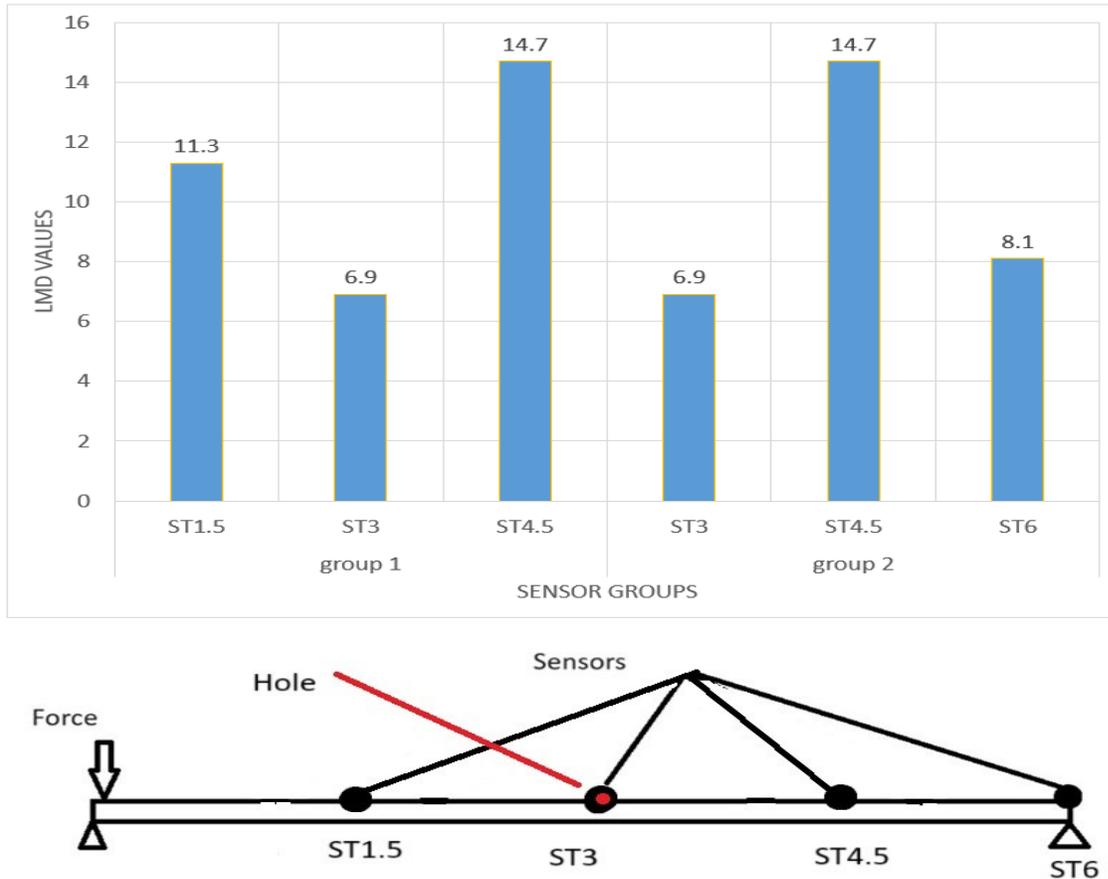


Figure 4.13. LMD for both pipe with hole and without hole

Comparison of the LMD values in Figure 4.13 and those in Figure 4.10 and Figure 4.7 shows that the LMD indices in these latter figures are higher. For instance, in Figure 4.10, where Group 3 was selected, the spike phase is at index 135.4. In Figure 4.7, the spike phase is at index 145.1 for the selection of Group 4. Increasing the distances between sensors results in a smaller value of LMD, but this does not impact prediction accuracy.

4.6 Fifth scenario for sensitivity investigation

In this section, the effect of increasing in diameter of the pipe is investigated. In this case, the inner and outer diameters are 37.71 mm and 42.22 mm, respectively (50% increase from the previous cases), while other specifications of pipeline remain similar to previous scenarios. In this part, one pipe of 6 m length was used as baseline while the other one has a hole of 1 mm which is located 3 m from the start of the pipe.

Based on this scenario and through several steps like extracting features like velocity and acceleration, the process of data fusion is done and 12 first principal components from both scenarios acted as input for computing ICA. After converting multivariate signal into maximally independent signal through ICA, both ICs (from both pipe with hole and without hole) are compared with each other (applying ICA in supervised learning style).

By referring to Figure 4.14, small mismatch between the components (except IC5) are detected, hence the detection is confirmed, but there is a point which needs more explanation and refers to the lower sensitivity of ICs by increasing the size of the diameter and having a constant impact load of 300 N.

After detecting a hole, localizing the hole between two sensors through analyzing data should be done. In this regard, 12 first principal components from both pipelines should be used as the input for computing local Mahalanobis distance.

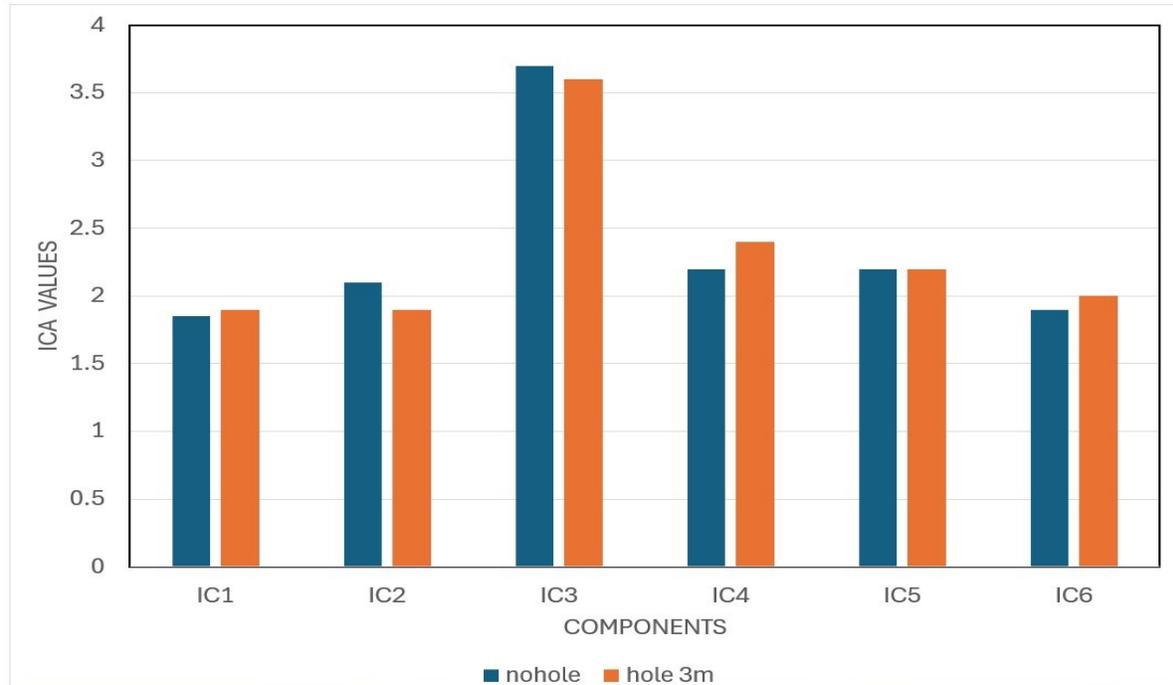


Figure 4.14. ICA for both pipe with hole and without hole

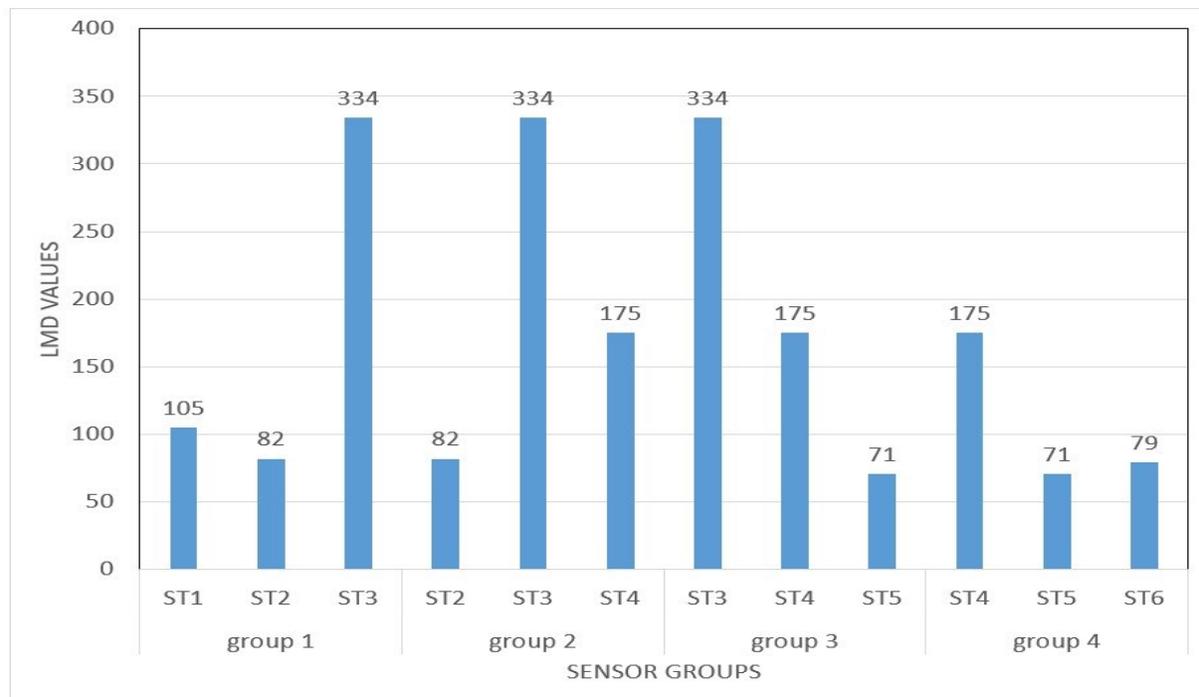


Figure 4.15. LMD for both pipe with hole and without hole

By referring to Figure 4.15 and based on the analysis of results in group of three, desirable group includes three different phases which are start, spike and decline. In this regard, group 2 conformed with the definition and location of hole estimated between ST2 (start) and ST3 (spike). Based on this definition, localizing damage between two sensors is confirmed.

4.7 Sixth scenario for size of hole investigation

This section aims to evaluate the performance of statistical models in distinguishing pipe behavior when two different hole sizes, namely 1 mm and 2 mm, are present. Since the approach in modeling is somewhat different here, the models will be described in detail. Here, the length of the pipes is 3 m and the inner and outer radius are 25.14 mm and 30.15 mm, respectively. The dynamic loading is as for the above sections: an impact load of 300 N is applied 15 cm from the start of the pipe in a 0.01 sec timeframe. The behavior of the pipes is then analyzed over a 0.1sec.

In this section, the acceleration is measured by only one sensor placed at the end of the pipe. First is an intact model serving as a baseline. In addition to the model serving as the baseline, 10 diverse models of pipes with a 1 mm hole and 10 diverse models with a 2 mm hole are built. Each step, the holes in the created models are 30 cm apart.

4.7.1 Comparison between model with hole and model without hole

The behaviors of the pipe due to dynamic loadings for two different hole sizes, 1 and 2 mm, can be visualized by graph provided via Total Acceleration technique which are different from each other, based on Figure 4.16. It is observable from mentioned Figures that with increasing the hole size, the amount of acceleration under the impact load effect has a tangible decrease—about one unit in the maximum state. Its causes could be different, and one of them refers to the effect of damping. It is worthy to say, the signal, while traveling through the structure, may continuously

lose energy due to certain reasons such as damping. Notably, the amplitude of damping tends to increase with the increase of damage severity which is interpreted as the hole size in our modeling (Cao et al., 2017).

Another subject which is considered in our simulation refers to neglecting the noise effect which we don't have it in our simulation through ANSYS.

The raw data for 20 different models which is only acceleration will be extracted and analyzed by using mathematical models in the next stage.

4.7.2 Mathematical and statistical investigation

In this regard, mathematical and statistical models are used to summarize the results into graphs form for better comprehension. Based on research (Ying et al. 2013), six different mathematical models have been used in analyzing extracted signals on this project as follow:

- **Skewness:** The degree of probability distribution takes a measure with the use of Skewness, and the value of the respective Skewness is interpreted as normal distribution ($SK = 0$), left side with more weight ($SK > 0$), right side with more weight ($SK < 0$). (refer to Figure 4.18)
- **curve length:** For time-scale changes using the curve length of the signal is a good solution. An abrupt change in the curve length of a signal may relate to several reasons, such as an abrupt change in the modal amplitude. (refer to Figure 4.19)
- **Pearson product-moment correlation:** The Pearson product-moment correlation coefficient is another way of assessing the relationship between two variables. This statistic varies between -1 and 1. (refer to Figure 4.20)

- **Kurtosis:** Kurtosis is a function used in measuring the tailedness of the distribution. It comprises mesokurtic as medium kurtosis, platykurtic as low kurtosis, and leptokurtic as high kurtosis, moving up to more liability to anomaly. (refer to Figure 4.21)
- **L2 norm:** the vector from a certain source is calculated by using the data analysis method of the L2 norm. The related figures are presented in the following. (refer to Figure 4.17). All relevant mathematical equations are provided in the appendix.

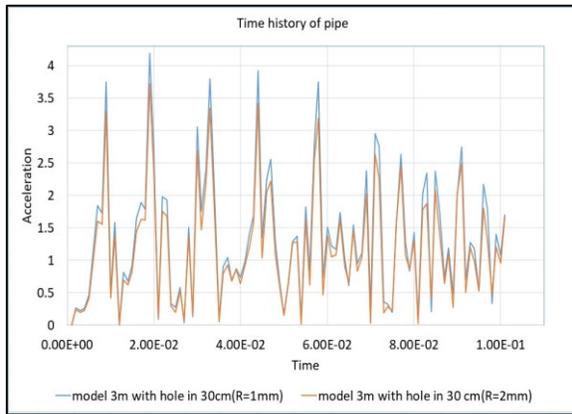


Figure 4.16. Time history response of total acceleration

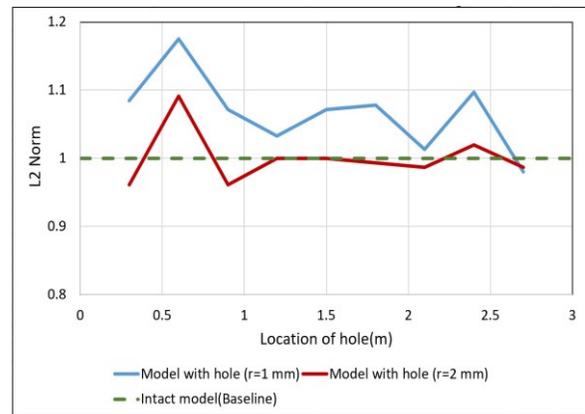


Figure 4.17. L2 Norm of the data

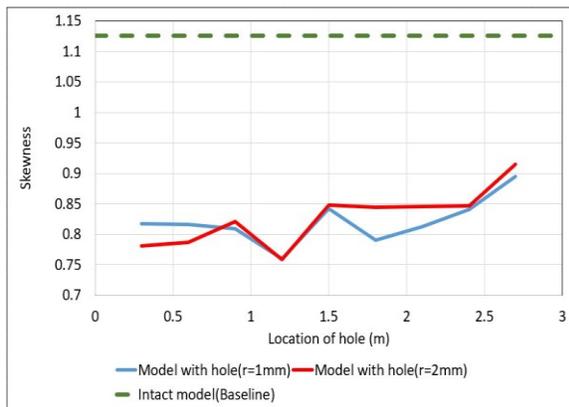


Figure 4.18. Skewness of the data

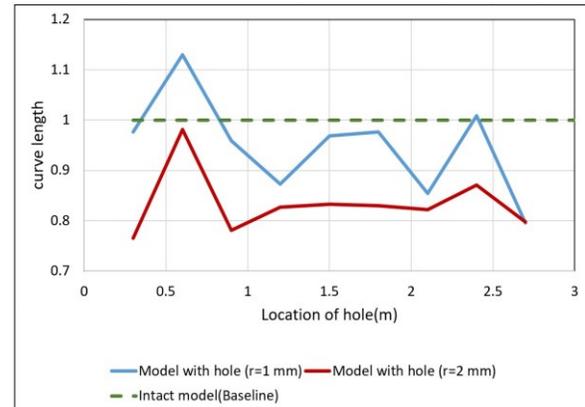


Figure 4.19. Curve length of the data

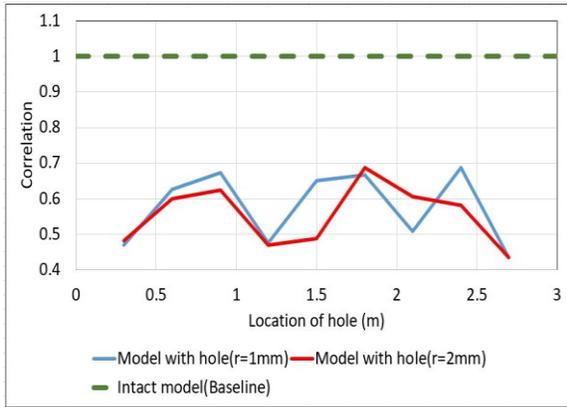


Figure 4.20. Correlation of the data

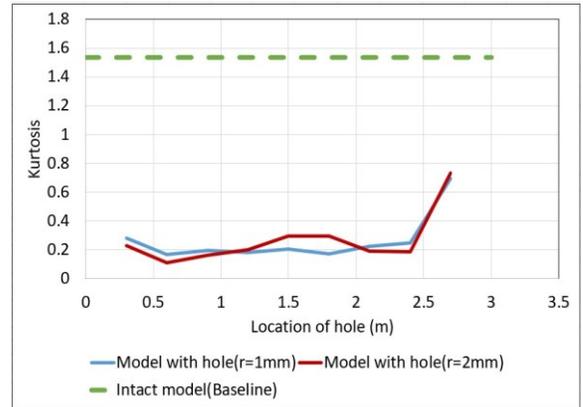


Figure 4.21. Kurtosis of the data

The curve length and L2-Norm index diagrams are quite different in gravity, as shown in Figures (4.19) and (4.17), respectively, and in terms of values of curve length and L2Norm, the pipe with a 2 mm hole is even lower than that with a 1 mm hole. If we want to investigate regarding the reason, and by considering that the graphs extracted from time acceleration, from Figure 4.16 as the time acceleration signal, with the increase in the hole size, the transmission acceleration decreases and this decrease shows it self as reduction about 1 unit at the peak point ($t = 0.02$ sec). The remaining indices, such as skewness, correlation, and kurtosis (refer to Figures 4.18, 4.20, and 4.21), also point out differences between types of damages but not as clear as L2-Norm and CL. However, the most significant feature of these graphs is distinguishing between pipes with and without holes.

Figure 4.18 shows that the skewness index ranges from 1.1 to 1.15 when the pipe has no hole (green line). Figure 4.20 indicates the correlation index is one for an intact pipe, while it falls between 0.4 and 0.7 when the pipe has a hole. Figure 4.21 shows that the kurtosis index ranges from 1.4 to 1.6 for an intact pipe and drops between 0 and 0.8 when the pipe has a hole.

Summary

This chapter presented the package we developed for localization and detection over different scenarios. Each of these scenarios was introduced with an objective to be able to research the different sides of the topic. While Scenario 1 and Scenario 2 were about localization and detection, respectively, Scenario 3 looked into the performance analysis of ICA on a pipe with two simultaneous holes, and Scenario 4 and 5 focused on the case of increased distance between sensors and diameter change (increase) for comparing the sensitivity. It is to be noticed that in all these scenarios (except of considering two holes at the same time), the detection and localization method developed here works effectively. Lastly, mathematical models have been developed and used to analyze the effect of hole size. From that, some of the statistical relationships are able to determine different hole sizes.

Chapter 5: Conclusion and Future work

5.1 Conclusion

In this project, we present the statistical framework that copes with the challenge of detecting and localizing the damage on a 6-meter pipeline. Our strategy for addressing the target of this thesis which is localizing damage only through data analysis, is based on the integration of data fusion techniques on selected features by an innovative application of Independent Component Analysis for detection scope. We also engage a Mahalanobis Distance with a nearest neighbor approach to deeply analyze multiple signals. Pieced together, this comprehensive methodology subsumes into a thorough statistical package that is suited quite effectively for the identification and localization of damage. We confirm this analysis with the implementation of a surrogate model from data taken in a prior experiment as a benchmark. A small defect is introduced, represented by a little hole, to simulate real conditions. This defect was the object of study and allowed us to analyze the behavior of the pipeline in different situations.

Acceleration and velocity data were collected for all of pipeline scenarios as part of our data analysis. First principal components were computed based on the gathered datasets, which are very important in describing maximum variance in the data and as a crucial step toward enhancing the model sensitivity. In the following some main results are rendered:

- Feature selection is one of the main steps in any developed model for the purpose of detection and localization. In this project, velocity and acceleration were selected as main

features after several experiments in a trial-and-error manner. The main reason behind such a choice is that even though these two are different concepts, they are correlated to each other in such a way that by combining them through PCA, desirable feature (first principal component) could be obtained.

- ICA indeed became very effective in such monitoring of pipeline conditions and damage detection. The effectiveness mainly lies in the processing and extraction of the independent components of signals. Since ICA needs a pre-processing stage to remove the correlations between data, using PCA in the whitening process is very helpful.
- Since most of the work in the K-Nearest Neighbor algorithm, including training and testing, is done at test time, selection of the value of k becomes crucial. A small value of k will suffer from high variance—overfitting, while a large value of k may cause underfitting. Hence, after trial and error, the value of k has been restricted to 5 in this project which is almost 5% of the data for a typical sensor and preferably k should be odd number.
- Curve length and L2 norm index have the potential to indicate differences in damage severity. Specifically, as the size of the hole increases, the values of these metrics tend to decrease.
- Increasing the distance between sensors results in LMD reduction but predicting hole location is correct.

5.2 Future works

The current study examines concepts such as detection and localization in pipelines by presenting a comprehensive package. The purpose of this section, as a future work suggestion, is to expand upon the current study. The following points are proposed to continue this research:

- Expanding the current model in a way that brings it as close as possible to real-world conditions, for example, by considering a real project with actual boundary conditions, so that factors such as environmental effects are incorporated into the modeling as much as possible.
- In respect of such a potentiality of the ICA, further consideration on several models, each representing the type and multiplicity of damage, is recommended. Further, the ICA components in extracting desirable features are recommended to be used, using the curve length and L2 norm capabilities. Also, the issue of models with different holes and physical characteristics can be explored.
- As in this research, the location of hole on the top of the pipe (+ 90 degrees) is investigated, examining hole location on the side of pipeline will be suggested.
- Considering other types of hole shaped like elliptical shape and investigating the pipeline behavior is suggested.
- Decision-making, as the final stage of health monitoring in consideration of the continuity of the monitoring process, becomes highly significant when projects involve several kilometers of pipeline length. Methods designed for this purpose should be employed; applying reinforcement learning and also game theory in this field is suggested for future works.

References

- Alberta Energy Regulator. (2024). *2023/24 Annual Report*. Alberta Energy Regulator. Retrieved from <https://www.aer.ca/protecting-what-matters/reporting-on-our-progress/annual-report>
- Alberta Energy Regulator. (2023). *2022/23 Annual Report*. Alberta Energy Regulator. Retrieved from <https://www.aer.ca/protecting-what-matters/reporting-on-our-progress/annual-report>
- Alberta Energy Regulator. (2022). *2021/22 Annual Report*. Alberta Energy Regulator. Retrieved from <https://www.aer.ca/protecting-what-matters/reporting-on-our-progress/annual-report>
- Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., & Inman, D. (2021). A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. *Mechanical Systems and Signal Processing*, *147*, 107077. <https://doi.org/10.1016/j.ymssp.2020.107077>
- Alves, V., & Cury, A. (2023). An automated vibration-based structural damage localization strategy using filter-type feature selection. *Mechanical Systems and Signal Processing*, *190*, 110145. <https://doi.org/10.1016/j.ymssp.2023.110145>
- Bakhary, N., Hao, H., & Deeks, A. J. (2010). Sub-structuring technique for damage detection using statistical multi-stage artificial neural network. *Advances in Structural Engineering*, *13*(1), 95–110.
- Bao, C., Hao, H., & Li, Z.-X. (2013). Integrated ARMA model method for damage detection of subsea pipeline system. *Engineering Structures*, *48*, 176–192. <https://doi.org/10.1016/j.engstruct.2012.09.033>

Betti, M., Facchini, L., & Biagini, P. (2014). Damage detection on a three-storey steel frame using artificial neural networks and genetic algorithms. *Meccanica*, 50(4), 875–886. <https://doi.org/10.1007/s11012-014-0085-9>

Cury, A., & Crémona, C. (2012). Pattern recognition of structural behaviors based on learning algorithms and symbolic data concepts. *Structural Control and Health Monitoring*, 19(2), 161–186. <https://doi.org/10.1002/stc.412>

Chang, P. C., & Liu, S. C. (2003). Recent research in non-destructive evaluation of civil infrastructure. *Journal of Materials in Civil Engineering*, 15(3), 298. [https://doi.org/10.1061/\(ASCE\)0899-1561\(2003\)15:3\(298\)](https://doi.org/10.1061/(ASCE)0899-1561(2003)15:3(298))

Chen, H.-P., & Ni, Y.-Q. (2018). *Structural health monitoring of large civil engineering structures*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119166641>

Cunningham, P., & Delany, S. J. (2021). *k-Nearest neighbour classifiers - A tutorial*. *ACM Computing Surveys*, 54(6), Article 128. <https://doi.org/10.1145/3459665>

Dwivedi, S. K., Vishwakarma, M., & Soni, P. A. (2018). Advances and research on non-destructive testing: A review. *Materials Today: Proceedings*. <https://doi.org/10.1016/j.matpr.2017.11.620>

El Mountassir, M., Yaacoubi, S., Mourot, G., & Maquin, D. (2018). Damage detection and localization in pipelines under non-stationary environment variation using sparse estimation of monitoring signals. International Symposium on Structural Health Monitoring and Nondestructive Testing 4-5 Oct 2018, Saarbrücken – Germany. *e-Journal of Nondestructive Testing* Vol. 23(12).

- González, M. P., & Zapico, J. L. (2008). Seismic damage identification in buildings using neural networks and modal data. *Computers and Structures*, *86*(3-5), 416–426. <https://doi.org/10.1016/j.compstruc.2007.02.021>
- Goh, L. D., Bakhary, N., Rahman, A. A., & Ahmad, B. H. (2013). Prediction of unmeasured mode shape using artificial neural network for damage detection. *Jurnal Teknologi (Sciences & Engineering)*, *61*(1), 57–66. <https://doi.org/10.11113/jt.v61.1624>
- Ghorbani, H. (2019). Mahalanobis distance and its application for detecting multivariate outliers. *Facta Universitatis Series: Mathematics and Informatics*, *34*(3), 583. <https://doi.org/10.22190/FUMI1903583G>
- Gulgec, N., Shahidi, G., Matarazzo, T., & Pakzad, S. (2017). Current challenges with big data analytics in structural health monitoring. In *Structural health monitoring & damage detection* (pp. 79–84). Springer. https://doi.org/10.1007/978-3-319-54109-9_9
- Hakim, S. J. S., Abdul Razak, H., & Ravanfar, S. A. (2015). Fault diagnosis on beam-like structures from modal parameters using artificial neural networks. *Measurement*, *76*, 45–61. <https://doi.org/10.1016/j.measurement.2015.08.021>
- Hyvärinen, A., & Oja, E. (2000). *Independent component analysis: A tutorial*. Helsinki University of Technology, Laboratory of Computer and Information Science.
- Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent Component Analysis*. John Wiley & Sons, Inc.

- Houdek, V., Smolík, L., & Kubin, Z. (2022). Impact point localization in three-dimensional structures using wavelet transform. *Mechanical Systems and Signal Processing*, 179, 109365. <https://doi.org/10.1016/j.ymssp.2022.109365>
- Humar, J., Bagchi, A., & Xu, H. (2006). Performance of vibration-based techniques for the identification of structural damage. *Structural Health Monitoring*, 5(3), 215–241. <https://doi.org/10.1177/1475921706067738>
- Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent component analysis*. Wiley. <https://doi.org/10.1002/0471221317>
- Jiang, S.F., Zhang, C.-M., & Yao, J. (2011). Eigen-level data fusion model by integrating rough set and probabilistic neural network for structural damage detection. *Advances in Structural Engineering*, 14(2), 333–349. <https://doi.org/10.1260/1369-4332.14.2.333>
- Jiang, S. F., Zhang, C. M., & Zhang, S. (2011). Two-stage structural damage detection using fuzzy neural networks and data fusion techniques. *Expert Systems with Applications*, 38(1), 511–519. <https://doi.org/10.1016/j.eswa.2010.06.093>
- Khazaeli, S. (2018). *Intelligent one-point damage localization of an isotropic surface pipeline using Gaussian process regression* (Master's thesis). Concordia University, Montreal, Quebec, Canada.
- Lam, H. F., & Ng, C. T. (2008). The selection of pattern features for structural damage detection using an extended Bayesian ANN algorithm. *Engineering Structures*, 30(10), 2762–2770. <https://doi.org/10.1016/j.engstruct.2008.03.012>

- Lyapin, A., & Shatilov, Y. (2018). Vibration-based damage detection of steel pipeline systems. In *Book Title* (pp. 63–72). Springer. https://doi.org/10.1007/978-3-319-56579-8_5
- Li, Y., Qu, L., & Qi, B. (2023). Simulation study on axial location identification of damage in layered pipeline structures based on damage index. *Applied Sciences*, *13*, 8850. <https://doi.org/10.3390/app13158850>
- Lee, J., & Kim, S. (2007). Structural damage detection in the frequency domain using neural networks. *Journal of Intelligent Material Systems and Structures*, *18*(8), 785–792. <https://doi.org/10.1177/1045389X06073640>
- Lee, J. J., Lee, J. W., Yi, J. H., Yun, C. B., & Jung, H. Y. (2005). Neural networks-based damage detection for bridges considering errors in baseline finite element models. *Journal of Sound and Vibration*, *280*(3–5), 555–578. <https://doi.org/10.1016/j.jsv.2004.01.003>
- Mehrjoo, M., Khaji, N., Moharrami, H., & Bahreininejad, A. (2008). Damage detection of truss bridge joints using artificial neural networks. *Expert Systems with Applications*, *35*(3), 1122–1131. <https://doi.org/10.1016/j.eswa.2007.08.008>
- Mousavi, A., Zhang, C., Masri, S., & Gholipour, G. (2022). Vibration feature extraction using signal processing techniques for structural health monitoring: A review. *Mechanical Systems and Signal Processing*, *177*, 109175. <https://doi.org/10.1016/j.ymsp.2022.109175>
- Ng, C. (2014). Application of Bayesian-designed artificial neural networks in Phase II structural health monitoring benchmark studies. *Australian Journal of Structural Engineering*, *15*(1), 27–37. <https://doi.org/10.7158/S12-042.2014.15.1>

Nisbett, J. K., Budynas, R. G., & Shigley, J. E. (2015). *Shigley's mechanical engineering design* (11th ed.). McGraw Hill.

Pawar, P. M., Reddy, K. V., & Ganguli, R. (2006). Damage detection in beams using spatial Fourier analysis and neural networks. *Journal of Intelligent Material Systems and Structures*, 18(4), 347–359. <https://doi.org/10.1177/1045389X06066292>

Rens, K. L., Wipf, T. J., & Klaiber, F. W. (1997). Review of non-destructive evaluation techniques of civil infrastructure. *Journal of Performance of Constructed Facilities*, 11(4), 152. [https://doi.org/10.1061/\(ASCE\)0887-3828\(1997\)11:4\(152\)](https://doi.org/10.1061/(ASCE)0887-3828(1997)11:4(152))

Salimi, M., & Bagchi, A. (2024). Damage detection of pipelines using numerical analysis. *GeoMontréal 2024, the 77th Canadian Geotechnical Conference and the 16th Joint CGS/IAH-CNC Groundwater Conference*, Montreal, Quebec, Canada.

Sarmadi, H., Entezami, A., Razavi, B., & Yuen, K.-V. (2020). Ensemble learning-based structural health monitoring by Mahalanobis distance metrics. *Structural Control and Health Monitoring*, 28. <https://doi.org/10.1002/stc.2663>

Santos, A., Figueiredo, E., Silva, M., Sales, C., & Costa, J. (2015). Machine learning algorithms for damage detection: Kernel-based approaches. *Journal of Sound and Vibration*, 363, Article ID 363. <https://doi.org/10.1016/j.jsv.2015.11.008>

Shlens, J. (2005). *A tutorial on principal component analysis*. Systems Neurobiology Laboratory, Salk Institute for Biological Studies.

- Simoen, E., De Roeck, G., & Lombaert, G. (2015). Dealing with uncertainty in model updating for damage assessment: A review. *Mechanical Systems and Signal Processing*, 56–57, 123–149. <https://doi.org/10.1016/j.ymssp.2014.11.001>
- Sophian, A., Tian, G. Y., Taylor, D., & Rudlin, J. (2001). Electromagnetic and eddy current NDT: A review. *Insight Non-Destructive Testing and Condition Monitoring*.
- Tharwat, A. (2016). Principal component analysis - A tutorial. *International Journal of Applied Pattern Recognition*, 3(3), 197. <https://doi.org/10.1504/IJAPR.2016.079733>
- Valinejadshoubi, M., Bagchi, A., & Moselhi, O. (2016). Structural health monitoring of buildings and infrastructure. *World Academy of Science, Engineering and Technology - International Journal of Civil and Environmental Engineering*, 10, 731–738.
- Wang, H., Barone, G., & Smith, A. (2023). A novel multi-level data fusion and anomaly detection approach for infrastructure damage identification and localisation. *Engineering Structures*, 292, 116473. <https://doi.org/10.1016/j.engstruct.2023.116473>
- Wu, R. T., & Jahanshahi, M. R. (2018). Data fusion approaches for structural health monitoring and system identification: Past, present, and future. *Structural Health Monitoring*. <https://doi.org/10.1177/1475921718798769>
- Xie, Y., Gao, C., Wang, P., Zhou, L., Zhang, C., & Qu, X. (2023). Research on vibration fatigue damage locations of offshore oil and gas pipelines based on the GA-improved BP neural network. *Shock and Vibration*, 2023, 1–18. <https://doi.org/10.1155/2023/2530651>
- Yang, L., & Zhao, Q. (2022). A BiLSTM based pipeline leak detection and disturbance assisted localization method. *IEEE Sensors Journal*, 190, 110145.

Ying, Y., Garrett, J., Oppenheim, I., Soibelman, L., Harley, J., Shi, J., & Jin, Y. (2013). Toward data-driven structural health monitoring: Application of machine learning and signal processing to damage detection. *Journal of Computing in Civil Engineering*, 27(6), 667–680. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000258](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000258).

Yuen, K.-V., & Lam, H.-F. (2006). On the complexity of artificial neural networks for smart structures monitoring. *Engineering Structures*, 28(7), 977–984. <https://doi.org/10.1016/j.engstruct.2005.11.002>

Yeung, W. T., & Smith, J. W. (2005). Damage detection in bridges using neural networks for pattern recognition of vibration signatures. *Engineering Structures*, 27(5), 685–698. <https://doi.org/10.1016/j.engstruct.2004.12.006>

Zhou, X. T., Ni, Y. Q., & Zhang, F. L. (2014). Damage localization of cable-supported bridges using modal frequency data and probabilistic neural network. *Mathematical Problems in Engineering*.

Appendices

Appendix A:

Python code:

In this part python codes are provided in three parts, the first part related to calculation of principal component analysis, the second part related to independent component analysis and the last one related to local Mahalanobis distance.

A.1 PCA (Principal Component Analysis) calculation

```
#PCA calculation-pipe 6m
import pandas as pd
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
# Load the Data
file_path = '/content/model6m+12sensor - 2hole.xlsx'
xls = pd.ExcelFile(file_path)
# Load both sheets
nohole_df = pd.read_excel(xls, 'nohole')
hole3m_df = pd.read_excel(xls, 'hole3m')
# Defining groups
nohole_groups = {
    'group1': ['Acc(1m)', 'Vel(1m)'],
    'group2': ['Acc(2m)', 'Vel(2m)'],
    'group3': ['Acc(3m)', 'Vel(3m)'],
    'group4': ['Acc(4m)', 'Vel(4m)'],
    'group5': ['Acc(5m)', 'Vel(5m)'],
```

```

    'group6': ['Acc(6m)', 'Vel(6m)']}]
hole3m_groups = {
    'group1': ['acc(1m)', 'vel(1m)'],
    'group2': ['acc(2m)', 'vel(2m)'],
    'group3': ['acc(3m)', 'vel(3m)'],
    'group4': ['acc(4m)', 'vel(4m)'],
    'group5': ['acc(5m)', 'vel(5m)'],
    'group6': ['acc(6m)', 'vel(6m)']}]
# Applying PCA Function
def apply_pca_to_groups(df, groups):
    pca_results = {}
    for group_name, columns in groups.items():
        scaler = StandardScaler()
        standardized_data = scaler.fit_transform(df[columns])
        pca = PCA(n_components=1)
        principal_components = pca.fit_transform(standardized_data)
        pca_results[group_name] = pd.DataFrame(data=principal_components,
        columns=[f'{group_name} Principal Component 1'])
    return pca_results
# Apply PCA to groups
nohole_pca_results = apply_pca_to_groups(nohole_df, nohole_groups)
hole3m_pca_results = apply_pca_to_groups(hole3m_df, hole3m_groups)
# Combine
output_file_path = 'pca_results_groups.xlsx'
with pd.ExcelWriter(output_file_path) as writer:
    for group_name, pca_df in nohole_pca_results.items():
        noh3m_result_df = pd.concat([nohole_df[nohole_groups[group_name]], pca_df], axis=1)
        noh3m_result_df.to_excel(writer, sheet_name=f'nohole_{group_name}', index=False)
    for group_name, pca_df in hole3m_pca_results.items():
        hole3m_result_df = pd.concat([hole3m_df[hole3m_groups[group_name]], pca_df], axis=1)

```

```

hole3m_result_df.to_excel(writer, sheet_name=f'hole3m_{group_name}', index=False)
# variance ratios for each group
explained_variance_nohole = {group_name:
PCA(n_components=1).fit(StandardScaler().fit_transform(nohole_df[nohole_groups[group_name]])).explained_variance_ratio_ for group_name in nohole_groups.keys()}
explained_variance_hole3m = {group_name:
PCA(n_components=1).fit(StandardScaler().fit_transform(hole3m_df[hole3m_groups[group_name]])).explained_variance_ratio_ for group_name in hole3m_groups.keys()}
explained_variance_nohole, explained_variance_hole3m

```

A.2 ICA (Independent Component Analysis) calculation

The chosen approach to solve this problem is with the use of FastICA via scikit-learn (source: <https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.FastICA.html>).

```

#ICA calculation _pipe6m_hole4.4m
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import FastICA
import matplotlib.pyplot as plt
# Loading data
file_path = '/content/PCA_vel_acc_6m_2 hole.xlsx'
data = pd.read_excel(file_path, sheet_name=None)
# nohole sheet as baseline
nohole_data = data['nohole']
# hole3m sheet for comparing
hole3m_data = data['hole3m']

```

```

# Columns
nohole_columns = ['Acc(6m)', 'Acc(5m)', 'Acc(4m)', 'Acc(3m)', 'Acc(2m)', 'Acc(1m)']
hole3m_columns = ['acc(6m)', 'acc(5m)', 'acc(4m)', 'acc(3m)', 'acc(2m)', 'acc(1m)']

# Standardizing data
scaler_nohole = StandardScaler()
scaler_hole3m = StandardScaler()

# Standardizing for nohole
nohole_scaled = scaler_nohole.fit_transform(nohole_data[nohole_columns])

# Standardizing for hole3m
hole3m_scaled = scaler_hole3m.fit_transform(hole3m_data[hole3m_columns])

# Apply ICA
ica = FastICA(n_components=len(nohole_columns), max_iter=2000, tol=0.01)

# Fit the ICA to nohole
nohole_ica = ica.fit_transform(nohole_scaled)
hole3m_ica = ica.transform(hole3m_scaled)

# DataFrames for ICA
nohole_ica_df = pd.DataFrame(nohole_ica, columns=[f'IC{i+1}' for i in
range(nohole_ica.shape[1])])
hole3m_ica_df = pd.DataFrame(hole3m_ica, columns=[f'IC{i+1}' for i in
range(hole3m_ica.shape[1])])

# Saving ICA results
nohole_ica_df.to_csv('nohole_ica_results.csv', index=False)
hole3m_ica_df.to_csv('hole3m_ica_results.csv', index=False)

```

A.3 LMD (Local Mahalanobis Distance) calculation

```
##LMD calculation for vel+acc+PCA+ pipe 6m_2 hole
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler
from sklearn.neighbors import NearestNeighbors
import matplotlib.pyplot as plt

# Loading data
file_path = '/content/PCA_vel_acc_6m_2 hole.xlsx'
data = pd.read_excel(file_path, sheet_name=None)

# nohole as baseline
nohole_data = data['nohole']

# hole3m to compare
hole3m_data = data['hole3m']

# column names
print("Nohole columns:", nohole_data.columns)
print("Hole3m columns:", hole3m_data.columns)

# Columns for use
columns_nohole = ['Acc(6m)', 'Acc(5m)', 'Acc(4m)', 'Acc(3m)', 'Acc(2m)', 'Acc(1m)']
columns_hole3m = ['acc(6m)', 'acc(5m)', 'acc(4m)', 'acc(3m)', 'acc(2m)', 'acc(1m)']

# Standardizeing data
scaler = StandardScaler()
nohole_scaled = scaler.fit_transform(nohole_data[columns_nohole])
hole3m_scaled = scaler.fit_transform(hole3m_data[columns_hole3m])

# DataFrame
nohole_df = pd.DataFrame(nohole_scaled, columns=columns_nohole)
hole3m_df = pd.DataFrame(hole3m_scaled, columns=columns_nohole) # Ensuring the same
column names

# Local Regions and Calculate Local Mahalanobis Distance
def calculate_local_mahalanobis(nohole_col, hole3m_col, k=5):
```

```

nbrs = NearestNeighbors(n_neighbors=k).fit(nohole_col.values.reshape(-1, 1))
distances, indices = nbrs.kneighbors(nohole_col.values.reshape(-1, 1))
local_means = []
local_vars = []
for idx in indices:
    neighborhood = nohole_col.iloc[idx].values
    local_mean = neighborhood.mean()
    local_var = np.var(neighborhood) # Use variance instead of covariance for 1D
    local_means.append(local_mean)
    local_vars.append(local_var)
local_means = np.array(local_means)
local_vars = np.array(local_vars)
distances = []
for i in range(len(hole3m_col)):
    diff = hole3m_col.iloc[i] - local_means[i]
    if local_vars[i] == 0: # Avoid division by zero
        md = 0
    else:
        md = np.sqrt(diff ** 2 / local_vars[i])
    distances.append(md)
return distances
# plot Local Mahalanobis Distance
for column in columns_nohole:
    lmd = calculate_local_mahalanobis(nohole_df[column], hole3m_df[column])
    lmd_df = pd.DataFrame(lmd, columns=[f'Local Mahalanobis Distance ({column})'])
    plt.figure(figsize=(10, 6))
    plt.plot(lmd_df)
    plt.title(f'Local Mahalanobis Distance for {column}')
    plt.xlabel('Observation')
    plt.ylabel('Local Mahalanobis Distance')

```

```
plt.show()
# Save LMD
lmd_df.to_csv(f'local_mahalanobis_distances_{column}.csv', index=False)
```

Appendix B:

B.1 First scenario

In this part the data related to the first scenario are rendered

B.1.1 Data related to baseline pipe (without hole):

B.1.1.1 Acceleration for intact pipe

Acc(6m)	Acc(5m)	Acc(4m)	Acc(3m)	Acc(2m)	Acc(1m)
1.13E-01	0.35027	7.65E-01	1.35E+00	2.44E+00	5.83E+00
1.63E-01	0.44394	0.9753	1.79E+00	3.67E+00	1.06E+01
1.33E-02	0.17603	0.36784	5.07E-01	6.59E-01	8.12E+00
0.15374	0.74555	1.71E+00	3.06E+00	3.31E+00	4.82E+00
0.25374	0.91162	2.29E+00	4.79E+00	6.70E+00	1.60E+00
0.22918	4.51E-01	1.5371	4.75E+00	8.52E+00	1.13E+00
7.64E-02	0.7081	0.65691	2.74E+00	8.31E+00	2.35E+00
0.14915	1.8677	3.3032	8.42E-01	6.38E+00	2.41E+00
0.35345	2.6467	5.6097	4.49E+00	3.59E+00	1.69E+00
0.44106	2.4608	6.539	7.20E+00	1.01E+00	8.58E-01
3.55E-01	1.0764	5.4629	8.04E+00	8.35E-01	0.35662
0.10245	1.5197	2.4857	6.86E+00	1.02E+00	3.36E-01
0.24342	4.3384	1.6771	4.31E+00	1.79E-01	7.82E-01
0.56666	6.8415	5.5279	1.54E+00	9.80E-01	1.40E+00
0.75486	8.1566	7.9783	4.61E-01	1.57E+00	1.72E+00
7.40E-01	7.8042	8.2703	6.65E-01	1.11E+00	1.49E+00
5.23E-01	5.7907	6.4018	9.07E-01	3.63E-01	7.76E-01
0.17462	2.6048	3.1116	3.18E+00	2.15E+00	1.28E-01
1.93E-01	0.96632	0.51676	5.21E+00	3.49E+00	6.75E-01
4.68E-01	4.0682	3.0343	5.90E+00	3.58E+00	5.70E-01
5.79E-01	6.1183	3.8217	4.70E+00	2.15E+00	5.41E-01
5.20E-01	6.9188	2.7967	1.81E+00	5.29E-01	1.86E+00
0.34665	6.6868	0.59333	1.95E+00	3.27E+00	3.21E+00
0.15102	5.9552	2.1741	5.17E+00	5.17E+00	3.88E+00
2.08E-02	5.1837	4.0701	6.83E+00	5.29E+00	3.37E+00
4.49E-03	4.6916	4.7288	6.41E+00	3.34E+00	1.60E+00
9.34E-02	4.5354	4.1833	4.24E+00	3.21E-01	1.08E+00
0.22922	4.5214	3.0018	1.23E+00	4.29E+00	3.81E+00
0.33157	4.3644	2.0179	1.53E+00	7.52E+00	5.82E+00
3.36E-01	3.7462	1.862	2.81E+00	8.90E+00	6.41E+00
2.22E-01	2.6193	2.6949	2.35E+00	8.02E+00	5.31E+00

2.60E-02	1.2496	4.1644	5.54E-01	5.25E+00	2.77E+00
0.17417	0.24097	5.5297	1.82E+00	1.58E+00	5.36E-01
0.29277	0.67967	5.947	3.24E+00	1.80E+00	3.75E+00
0.27299	0.46449	4.8687	3.05E+00	3.76E+00	6.13E+00
0.11358	0.72961	2.3564	1.10E+00	3.92E+00	7.26E+00
0.12706	1.9849	0.92457	2.16E+00	2.56E+00	7.16E+00
0.35177	3.0091	3.8409	5.46E+00	5.55E-01	6.23E+00
0.46253	3.1331	5.3792	7.66E+00	1.70E+00	5.01E+00
0.40016	2.0472	4.8973	7.94E+00	2.62E+00	3.91E+00
0.16989	0.31066	2.4784	6.21E+00	2.30E+00	3.13E+00
0.158	2.8805	1.1598	3.15E+00	1.09E+00	2.58E+00
0.47297	5.3522	4.6174	2.23E-01	2.69E-01	2.03E+00
6.65E-01	6.8037	6.7922	2.51E+00	1.02E+00	1.14E+00
6.67E-01	6.7515	6.9134	3.16E+00	7.19E-01	3.69E-01
0.4768	5.173	4.9363	2.12E+00	6.78E-01	2.04E+00
0.16078	2.5097	1.5364	1.78E-01	2.20E+00	3.72E+00
1.75E-01	0.52268	2.2145	1.83E+00	3.17E+00	4.87E+00
0.42294	3.0732	5.0702	2.53E+00	2.86E+00	5.06E+00
5.16E-01	4.6252	6.3165	1.53E+00	9.82E-01	4.21E+00
4.47E-01	4.9715	5.8879	1.10E+00	2.09E+00	2.62E+00
0.26804	4.3078	4.2883	4.27E+00	5.39E+00	9.72E-01
6.67E-02	3.1052	2.4038	6.90E+00	7.79E+00	4.63E-01
7.51E-02	1.8351	1.1039	7.96E+00	8.36E+00	5.52E-01
0.1122	0.77684	0.79697	6.98E+00	6.79E+00	1.00E+00
5.42E-02	6.91E-02	1.4175	4.28E+00	3.53E+00	2.98E+00
4.29E-02	0.73982	2.5268	8.12E-01	4.43E-01	5.26E+00
0.1036	1.6156	3.368	2.20E+00	3.70E+00	7.01E+00
6.91E-02	2.8364	3.2792	3.78E+00	5.29E+00	7.52E+00
7.44E-02	4.3746	2.1195	3.52E+00	4.77E+00	6.48E+00
0.28668	5.946	0.36113	1.81E+00	2.44E+00	4.12E+00
4.88E-01	7.1433	1.4534	4.52E-01	8.25E-01	1.07E+00
5.94E-01	7.4546	2.0847	2.01E+00	3.78E+00	1.98E+00
0.54703	6.6405	1.2064	2.15E+00	5.49E+00	4.22E+00
0.34918	4.8247	1.1415	6.72E-01	5.52E+00	5.29E+00
6.07E-02	2.5249	4.254	2.05E+00	4.14E+00	5.23E+00
0.21895	0.53079	7.0585	4.71E+00	2.05E+00	4.39E+00
0.38905	1.0677	8.4687	6.27E+00	4.00E-01	3.30E+00
3.87E-01	1.1563	7.8334	5.98E+00	1.10E+00	2.30E+00
0.21298	0.27724	5.1997	3.78E+00	1.02E+00	1.57E+00
6.73E-02	1.7722	1.3203	3.52E-01	2.32E-01	1.05E+00
3.48E-01	3.5061	2.6092	3.17E+00	5.71E-01	5.17E-01
5.23E-01	4.4761	5.38	5.63E+00	6.79E-01	4.67E-01
0.52395	4.1829	6.2071	6.32E+00	6.17E-01	1.71E+00
0.34743	2.5672	4.9991	5.23E+00	2.48E+00	3.18E+00
5.35E-02	8.18E-02	2.3645	3.06E+00	4.65E+00	4.48E+00

0.25816	2.8313	0.64571	9.14E-01	6.16E+00	5.06E+00
0.48721	5.225	2.9484	2.50E-01	6.27E+00	4.50E+00
0.56995	6.6575	3.8507	2.72E-01	4.73E+00	2.75E+00
5.01E-01	6.9585	3.2602	2.20E+00	1.90E+00	1.66E-01
0.33089	6.316	1.6648	4.72E+00	1.27E+00	2.66E+00
0.14159	5.1845	0.50635	6.71E+00	3.61E+00	4.90E+00
1.05E-02	3.9641	1.6242	7.21E+00	4.21E+00	5.96E+00
2.18E-02	2.9166	1.8393	5.80E+00	2.70E+00	5.63E+00
3.34E-02	2.0798	1.1537	2.83E+00	5.10E-01	4.15E+00
1.21E-01	1.3111	1.55E-01	7.52E-01	4.42E+00	2.21E+00
0.1673	0.38849	0.51224	3.75E+00	7.78E+00	6.53E-01
0.1188	1.0566	0.33121	5.22E+00	9.54E+00	5.59E-01
3.49E-02	2.6656	1.4314	4.77E+00	9.23E+00	4.71E-01
0.25085	4.2633	3.3781	2.81E+00	7.13E+00	1.79E+00
0.44883	5.3705	5.0851	3.50E-01	4.06E+00	3.89E+00
0.54332	5.4884	5.6748	1.85E+00	1.15E+00	5.95E+00
0.48017	4.3865	4.6389	2.47E+00	7.14E-01	7.31E+00
0.26171	2.2106	2.0185	1.51E+00	1.12E+00	7.61E+00
5.00E-02	0.52088	1.5033	5.40E-01	3.33E-01	6.89E+00
0.35415	3.0669	4.7949	2.64E+00	1.27E+00	5.53E+00
0.54797	4.71	6.7869	3.76E+00	2.45E+00	3.98E+00
0.56595	5.0056	6.7848	3.17E+00	2.77E+00	2.57E+00
4.05E-01	3.9408	4.7932	8.61E-01	2.09E+00	1.46E+00
0.12577	1.9408	1.5314	2.53E+00	8.20E-01	5.30E-01
0.16965	0.33676	1.9589	5.83E+00	4.11E-01	4.66E-01

B.1.1.2 Velocity for the intact pipe

Vel(6m)	Vel(5m)	Vel(4m)	Vel(3m)	Vel(2m)	Vel(1m)
1.13E-04	3.50E-04	7.65E-04	1.35E-03	2.44E-03	5.83E-03
2.76E-04	7.94E-04	1.74E-03	3.14E-03	6.11E-03	1.64E-02
2.89E-04	6.45E-04	1.40E-03	2.69E-03	6.75E-03	2.46E-02
1.36E-04	2.44E-04	4.52E-04	6.26E-04	3.56E-03	2.93E-02
1.18E-04	1.04E-03	2.62E-03	5.27E-03	3.50E-03	3.08E-02
3.47E-04	1.46E-03	4.14E-03	9.95E-03	1.20E-02	2.98E-02
4.24E-04	8.73E-04	3.62E-03	1.27E-02	2.02E-02	2.76E-02
2.75E-04	1.28E-03	6.84E-04	1.20E-02	2.66E-02	2.53E-02
7.88E-05	3.77E-03	5.50E-03	7.61E-03	3.01E-02	2.36E-02
5.20E-04	6.18E-03	1.20E-02	1.10E-03	3.09E-02	2.28E-02
8.75E-04	7.22E-03	1.73E-02	7.98E-03	3.02E-02	2.25E-02
9.78E-04	5.81E-03	1.98E-02	1.48E-02	2.92E-02	2.22E-02
7.34E-04	1.64E-03	1.82E-02	1.91E-02	2.91E-02	2.14E-02
1.67E-04	5.63E-03	1.27E-02	2.05E-02	3.01E-02	2.01E-02
5.87E-04	1.36E-02	4.77E-03	2.01E-02	3.16E-02	1.84E-02
1.33E-03	2.14E-02	3.79E-03	1.95E-02	3.27E-02	1.70E-02
1.85E-03	2.72E-02	1.02E-02	2.03E-02	3.23E-02	1.62E-02
2.02E-03	2.98E-02	1.33E-02	2.34E-02	3.02E-02	1.63E-02
1.83E-03	2.88E-02	1.29E-02	2.86E-02	2.67E-02	1.68E-02
1.36E-03	2.47E-02	9.93E-03	3.45E-02	2.32E-02	1.72E-02
7.85E-04	1.87E-02	6.20E-03	3.92E-02	2.12E-02	1.68E-02
2.66E-04	1.18E-02	3.50E-03	4.10E-02	2.17E-02	1.49E-02
8.09E-05	5.15E-03	2.91E-03	3.92E-02	2.50E-02	1.17E-02
2.32E-04	8.93E-04	4.77E-03	3.40E-02	3.01E-02	7.98E-03
2.53E-04	5.99E-03	8.64E-03	2.73E-02	3.53E-02	4.80E-03
2.57E-04	1.07E-02	1.33E-02	2.09E-02	3.86E-02	3.43E-03
3.50E-04	1.52E-02	1.75E-02	1.66E-02	3.83E-02	4.47E-03
5.80E-04	1.97E-02	2.05E-02	1.54E-02	3.40E-02	8.12E-03
9.11E-04	2.41E-02	2.25E-02	1.68E-02	2.65E-02	1.38E-02
1.25E-03	2.78E-02	2.43E-02	1.95E-02	1.76E-02	2.00E-02
1.47E-03	3.03E-02	2.69E-02	2.19E-02	9.62E-03	2.53E-02
1.49E-03	3.16E-02	3.11E-02	2.24E-02	4.40E-03	2.80E-02
1.32E-03	3.17E-02	3.66E-02	2.06E-02	2.86E-03	2.75E-02
1.03E-03	3.12E-02	4.25E-02	1.74E-02	4.66E-03	2.37E-02
7.54E-04	3.09E-02	4.74E-02	1.44E-02	8.41E-03	1.77E-02
6.41E-04	3.16E-02	4.98E-02	1.33E-02	1.23E-02	1.05E-02
7.68E-04	3.36E-02	4.88E-02	1.54E-02	1.48E-02	3.41E-03
1.12E-03	3.65E-02	4.50E-02	2.07E-02	1.52E-02	3.07E-03
1.58E-03	3.95E-02	3.96E-02	2.83E-02	1.37E-02	7.98E-03
1.98E-03	4.14E-02	3.47E-02	3.62E-02	1.12E-02	1.18E-02
2.15E-03	4.12E-02	3.23E-02	4.25E-02	9.11E-03	1.49E-02
1.99E-03	3.83E-02	3.34E-02	4.56E-02	8.09E-03	1.75E-02

1.52E-03	3.31E-02	3.81E-02	4.55E-02	8.33E-03	1.95E-02
8.56E-04	2.64E-02	4.49E-02	4.30E-02	9.23E-03	2.06E-02
1.89E-04	1.98E-02	5.18E-02	4.00E-02	9.80E-03	2.03E-02
2.88E-04	1.48E-02	5.67E-02	3.79E-02	9.13E-03	1.83E-02
4.48E-04	1.23E-02	5.82E-02	3.79E-02	6.94E-03	1.46E-02
2.74E-04	1.27E-02	5.61E-02	3.97E-02	3.80E-03	9.71E-03
1.49E-04	1.56E-02	5.10E-02	4.22E-02	1.20E-03	4.73E-03
6.65E-04	2.01E-02	4.47E-02	4.37E-02	9.38E-04	7.25E-04
1.11E-03	2.50E-02	3.88E-02	4.26E-02	2.20E-03	2.21E-03
1.38E-03	2.93E-02	3.45E-02	3.84E-02	7.42E-03	3.06E-03
1.45E-03	3.24E-02	3.22E-02	3.15E-02	1.51E-02	2.82E-03
1.37E-03	3.42E-02	3.13E-02	2.35E-02	2.34E-02	2.67E-03
1.26E-03	3.50E-02	3.07E-02	1.66E-02	3.01E-02	3.67E-03
1.21E-03	3.49E-02	2.93E-02	1.24E-02	3.36E-02	6.61E-03
1.25E-03	3.42E-02	2.68E-02	1.16E-02	3.32E-02	1.18E-02
1.35E-03	3.26E-02	2.34E-02	1.38E-02	2.95E-02	1.87E-02
1.42E-03	2.98E-02	2.01E-02	1.75E-02	2.42E-02	2.61E-02
1.35E-03	2.54E-02	1.80E-02	2.10E-02	1.95E-02	3.24E-02
1.06E-03	1.94E-02	1.78E-02	2.28E-02	1.70E-02	3.66E-02
5.71E-04	1.23E-02	1.93E-02	2.23E-02	1.79E-02	3.76E-02
2.27E-05	5.00E-03	2.14E-02	2.03E-02	2.16E-02	3.57E-02
5.70E-04	1.90E-03	2.26E-02	1.82E-02	2.71E-02	3.16E-02
9.19E-04	6.65E-03	2.14E-02	1.75E-02	3.26E-02	2.64E-02
9.80E-04	9.13E-03	1.72E-02	1.96E-02	3.67E-02	2.12E-02
7.61E-04	9.45E-03	1.02E-02	2.43E-02	3.86E-02	1.68E-02
3.72E-04	8.48E-03	1.88E-03	3.05E-02	3.87E-02	1.35E-02
1.53E-05	7.54E-03	6.35E-03	3.65E-02	3.78E-02	1.12E-02
2.28E-04	7.63E-03	1.15E-02	4.03E-02	3.69E-02	9.67E-03
1.61E-04	9.40E-03	1.28E-02	4.07E-02	3.67E-02	8.63E-03
1.87E-04	1.28E-02	1.02E-02	3.75E-02	3.72E-02	8.17E-03
7.10E-04	1.72E-02	4.91E-03	3.19E-02	3.77E-02	8.61E-03
1.23E-03	2.13E-02	1.82E-03	2.55E-02	3.72E-02	1.03E-02
1.58E-03	2.38E-02	6.64E-03	2.03E-02	3.47E-02	1.35E-02
1.64E-03	2.38E-02	8.99E-03	1.73E-02	3.01E-02	1.79E-02
1.38E-03	2.10E-02	8.36E-03	1.64E-02	2.40E-02	2.29E-02
8.90E-04	1.58E-02	5.44E-03	1.66E-02	1.77E-02	2.72E-02
3.20E-04	9.16E-03	1.69E-03	1.63E-02	1.30E-02	3.00E-02
1.81E-04	2.28E-03	1.95E-03	1.41E-02	1.11E-02	3.01E-02
5.12E-04	4.17E-03	3.43E-03	9.46E-03	1.24E-02	2.74E-02
6.54E-04	9.31E-03	3.17E-03	2.84E-03	1.60E-02	2.26E-02
6.64E-04	1.32E-02	1.81E-03	4.55E-03	2.02E-02	1.67E-02
6.42E-04	1.61E-02	9.51E-04	1.03E-02	2.29E-02	1.12E-02
6.76E-04	1.82E-02	1.77E-03	1.32E-02	2.24E-02	7.14E-03
7.96E-04	1.95E-02	1.86E-03	1.24E-02	1.80E-02	4.93E-03
9.64E-04	1.98E-02	1.36E-03	8.67E-03	1.02E-02	4.35E-03

1.08E-03	1.88E-02	1.10E-03	3.45E-03	6.61E-04	4.55E-03
1.05E-03	1.62E-02	2.28E-03	1.33E-03	8.58E-03	4.45E-03
7.97E-04	1.19E-02	5.59E-03	4.13E-03	1.57E-02	2.80E-03
3.48E-04	6.53E-03	1.07E-02	4.40E-03	1.98E-02	1.55E-03
1.96E-04	1.22E-03	1.64E-02	2.71E-03	2.09E-02	7.30E-03
6.76E-04	3.41E-03	2.10E-02	1.04E-03	2.02E-02	1.45E-02
9.37E-04	5.61E-03	2.30E-02	1.97E-03	1.91E-02	2.21E-02
8.87E-04	5.09E-03	2.15E-02	1.47E-03	1.89E-02	2.89E-02
5.33E-04	2.05E-03	1.67E-02	1.81E-03	2.01E-02	3.45E-02
1.47E-05	2.80E-03	9.91E-03	5.38E-03	2.26E-02	3.84E-02
5.81E-04	7.73E-03	3.15E-03	8.48E-03	2.53E-02	4.10E-02
9.86E-04	1.17E-02	1.72E-03	9.29E-03	2.74E-02	4.25E-02
1.11E-03	1.36E-02	3.20E-03	6.76E-03	2.82E-02	4.30E-02
9.42E-04	1.33E-02	1.29E-03	9.69E-04	2.79E-02	4.25E-02

B.1.2 Data related to pipe with hole (hole is located 4.4m from the beginning of pipe):

B.1.2.1 Acceleration for pipe with hole

acc(6m)	acc(5m)	acc(4m)	acc(3m)	acc(2m)	acc(1m)
0.11389	0.49475	7.73E-01	1.46E+00	2.50E+00	6.28E+00
0.16536	0.65248	0.9921	1.99E+00	3.81E+00	1.14E+01
1.48E-02	0.16402	0.38787	5.48E-01	8.18E-01	8.62E+00
0.15365	0.94546	1.7013	3.07E+00	3.57E+00	5.16E+00
0.25608	1.2513	2.3059	4.92E+00	7.02E+00	1.93E+00
0.23419	0.75488	1.5987	5.05E+00	8.85E+00	1.54E+00
8.28E-02	0.61983	0.76043	3.17E+00	8.59E+00	2.84E+00
0.14447	2.0274	3.2947	8.58E-01	6.71E+00	2.80E+00
0.3543	3.1251	5.548	4.56E+00	4.02E+00	1.89E+00
0.45028	3.1331	6.5923	7.35E+00	1.42E+00	1.08E+00
0.3727	1.7068	5.6656	8.29E+00	1.12E+00	5.45E-01
0.12391	1.244	2.8187	7.18E+00	1.34E+00	4.69E-01
0.22578	4.3942	1.5065	4.74E+00	4.15E-01	8.28E-01
0.56179	7.4056	5.341	2.03E+00	8.88E-01	1.53E+00
0.76929	9.1245	7.9002	4.31E-01	1.66E+00	2.00E+00
7.74E-01	8.9436	8.4032	8.36E-01	1.35E+00	1.82E+00
5.71E-01	6.8136	6.7727	9.66E-01	2.63E-01	1.02E+00
0.22382	3.2597	3.647	3.12E+00	2.01E+00	1.29E-01
0.15716	0.80943	0.22495	5.13E+00	3.57E+00	8.57E-01
0.45676	4.3925	2.8195	5.94E+00	3.93E+00	9.08E-01
0.59631	6.7615	3.8267	5.03E+00	2.70E+00	5.41E-01
0.55985	7.643	3.0672	2.34E+00	3.05E-01	1.74E+00
0.39551	7.2662	1.0788	1.71E+00	2.95E+00	3.23E+00
0.19211	6.3267	1.943	5.08E+00	5.07E+00	4.23E+00
4.13E-02	5.3461	4.0299	6.98E+00	5.61E+00	3.97E+00
8.63E-04	4.7085	4.9279	6.81E+00	4.01E+00	2.27E+00
7.14E-02	4.5255	4.5641	4.88E+00	6.54E-01	7.79E-01
0.20367	4.6003	3.422	1.95E+00	3.81E+00	3.66E+00
0.31835	4.6284	2.3569	1.40E+00	7.24E+00	6.14E+00
0.34498	4.168	2.0557	3.10E+00	9.00E+00	7.13E+00
0.25343	3.0382	2.6579	3.01E+00	8.55E+00	6.26E+00
6.85E-02	1.476	3.9335	1.38E+00	6.09E+00	3.68E+00
0.13925	0.19865	5.2811	1.40E+00	2.49E+00	8.82E-02
0.28288	1.0684	5.9586	3.03E+00	1.29E+00	3.58E+00
0.29643	1.0261	5.2408	3.21E+00	3.61E+00	6.42E+00
0.16476	0.29686	3.0388	1.74E+00	4.18E+00	7.88E+00
6.68E-02	1.6814	0.22141	1.50E+00	3.23E+00	7.90E+00
0.30746	3.1575	3.2141	4.92E+00	1.30E+00	6.87E+00

0.45569	3.7602	5.1763	7.50E+00	1.77E+00	5.68E+00
0.43931	2.9869	5.2413	8.29E+00	3.04E+00	4.46E+00
0.24569	0.83603	3.2953	7.00E+00	2.98E+00	3.48E+00
7.06E-02	2.2629	0.34096	4.17E+00	1.80E+00	2.74E+00
0.40596	5.2601	3.8187	8.99E-01	2.05E-01	2.19E+00
0.64557	7.3496	6.3831	2.19E+00	1.00E+00	1.51E+00
0.70549	7.8011	7.1079	3.43E+00	1.16E+00	3.01E-01
0.5637	6.4254	5.7165	2.88E+00	4.05E-01	1.76E+00
0.2677	3.5974	2.6777	1.04E+00	1.78E+00	3.50E+00
8.40E-02	0.23719	1.1833	1.27E+00	2.91E+00	5.04E+00
0.3789	3.0658	4.4245	2.36E+00	3.07E+00	5.64E+00
5.32E-01	5.176	6.2335	2.03E+00	1.72E+00	5.06E+00
0.51375	5.8536	6.3855	4.80E-01	1.14E+00	3.46E+00
0.35908	5.2187	5.1955	3.58E+00	4.49E+00	1.69E+00
1.47E-01	3.7868	3.3603	6.55E+00	7.42E+00	5.23E-01
3.38E-02	2.242	1.7541	8.21E+00	8.76E+00	9.49E-01
0.11993	0.91718	0.98552	7.88E+00	7.95E+00	8.17E-01
9.93E-02	0.14979	1.1554	5.60E+00	5.14E+00	2.42E+00
1.18E-02	0.7116	2.0013	2.12E+00	1.21E+00	4.58E+00
7.10E-02	1.4273	3.0699	1.37E+00	2.71E+00	6.90E+00
7.97E-02	2.473	3.4656	3.73E+00	5.02E+00	8.07E+00
2.01E-02	3.9487	2.7959	4.27E+00	5.33E+00	7.57E+00
0.2101	5.664	1.2453	3.06E+00	3.59E+00	5.45E+00
4.25E-01	7.2368	0.82771	8.50E-01	5.73E-01	2.21E+00
0.57742	8.1039	1.8187	1.34E+00	2.89E+00	1.27E+00
0.59589	7.7555	1.5513	2.25E+00	5.23E+00	4.13E+00
0.45535	6.121	0.21484	1.60E+00	5.95E+00	5.76E+00
0.19157	3.5836	3.0074	9.88E-01	5.20E+00	6.01E+00
0.1094	1.0629	6.0283	3.75E+00	3.34E+00	5.17E+00
0.34247	1.2318	8.1188	5.91E+00	1.26E+00	3.98E+00
0.42352	1.9934	8.4071	6.48E+00	1.23E+00	2.80E+00
0.32247	1.2769	6.6096	5.05E+00	1.63E+00	1.80E+00
7.62E-02	0.75212	3.1632	1.98E+00	9.84E-01	1.05E+00
0.22384	2.8857	0.91378	1.76E+00	1.17E-01	5.08E-01
4.66E-01	4.6358	4.3753	4.96E+00	8.22E-01	4.31E-01
0.55724	5.1147	6.1997	6.65E+00	5.78E-01	1.49E+00
0.46178	3.9981	5.9419	6.42E+00	1.76E+00	2.81E+00
0.20913	1.4997	3.8776	4.65E+00	3.78E+00	4.19E+00
0.11613	1.7251	0.87483	2.27E+00	5.60E+00	5.31E+00
0.4076	4.8007	1.9507	4.11E-01	6.40E+00	5.39E+00
0.57701	6.9668	3.67E+00	1.74E-01	5.72E+00	4.15E+00
0.58564	7.8058	3.935	1.10E+00	3.53E+00	1.71E+00
0.45528	7.353	2.8677	3.45E+00	3.99E-01	1.40E+00
0.25521	6.0603	1.1272	5.87E+00	2.55E+00	4.32E+00
7.11E-02	4.5441	1.1828	7.26E+00	4.20E+00	6.25E+00

3.09E-02	3.1578	2.0589	6.86E+00	3.81E+00	6.69E+00
3.02E-02	2.112	1.8936	4.59E+00	1.37E+00	5.62E+00
4.23E-02	1.3357	0.9875	1.18E+00	2.43E+00	3.56E+00
0.12082	0.65458	0.23229	2.54E+00	6.38E+00	1.75E+00
0.13629	0.52119	0.41063	5.12E+00	9.24E+00	6.22E-01
4.91E-02	1.9593	0.86978	5.87E+00	1.01E+01	8.00E-01
0.1306	3.6455	2.3914	4.73E+00	8.93E+00	1.17E+00
0.34364	5.2348	4.1639	2.32E+00	6.15E+00	2.91E+00
0.50473	6.096	5.3482	8.09E-01	2.85E+00	5.11E+00
0.53735	5.71	5.2611	2.38E+00	1.75E-01	7.17E+00
0.40708	3.9422	3.6232	2.44E+00	1.26E+00	8.19E+00
1.39E-01	1.1161	0.63149	1.16E+00	1.24E+00	7.94E+00
0.18673	2.0543	2.9007	1.17E+00	4.17E-01	6.65E+00
0.46492	4.6805	5.7359	3.04E+00	1.99E+00	5.03E+00
0.60182	6.0103	7.0119	3.62E+00	3.05E+00	3.36E+00
5.52E-01	5.6894	6.2392	2.40E+00	3.04E+00	1.89E+00
0.3345	3.8932	3.7281	5.85E-01	2.00E+00	6.86E-01
2.87E-02	1.3034	0.36684	3.93E+00	7.57E-01	3.64E-01

B.1.2.2 Velocity for pipe with hole

vel(6m)	vel(5m)	vel(4m)	vel(3m)	vel(2m)	vel(1m)
1.14E-04	4.95E-04	7.73E-04	1.46E-03	2.50E-03	6.28E-03
2.79E-04	1.15E-03	1.76E-03	3.45E-03	6.31E-03	1.77E-02
2.94E-04	1.01E-03	1.44E-03	3.12E-03	7.11E-03	2.63E-02
1.40E-04	1.64E-04	5.78E-04	7.71E-04	4.09E-03	3.12E-02
1.16E-04	1.20E-03	2.58E-03	5.42E-03	4.45E-03	3.24E-02
3.50E-04	1.94E-03	4.17E-03	1.01E-02	1.32E-02	3.14E-02
4.33E-04	1.46E-03	3.74E-03	1.31E-02	2.18E-02	2.94E-02
2.88E-04	9.86E-04	7.99E-04	1.29E-02	2.83E-02	2.70E-02
6.61E-05	3.82E-03	5.74E-03	8.93E-03	3.18E-02	2.52E-02
5.16E-04	6.85E-03	1.21E-02	2.26E-03	3.24E-02	2.41E-02
8.89E-04	8.54E-03	1.74E-02	8.33E-03	3.14E-02	2.36E-02
1.01E-03	7.52E-03	2.00E-02	1.55E-02	3.03E-02	2.32E-02
7.87E-04	3.17E-03	1.88E-02	2.01E-02	3.02E-02	2.23E-02
2.25E-04	4.84E-03	1.37E-02	2.17E-02	3.10E-02	2.10E-02
5.44E-04	1.36E-02	5.96E-03	2.14E-02	3.24E-02	1.94E-02
1.32E-03	2.24E-02	3.35E-03	2.06E-02	3.35E-02	1.80E-02
1.89E-03	2.92E-02	1.01E-02	2.10E-02	3.35E-02	1.71E-02
2.11E-03	3.24E-02	1.37E-02	2.38E-02	3.15E-02	1.70E-02
1.96E-03	3.16E-02	1.39E-02	2.88E-02	2.80E-02	1.74E-02
1.50E-03	2.72E-02	1.13E-02	3.47E-02	2.46E-02	1.77E-02
9.03E-04	2.06E-02	7.60E-03	3.97E-02	2.25E-02	1.72E-02
3.43E-04	1.31E-02	4.60E-03	4.21E-02	2.27E-02	1.56E-02
5.26E-05	5.91E-03	3.53E-03	4.07E-02	2.57E-02	1.25E-02
2.45E-04	6.55E-04	4.84E-03	3.60E-02	3.06E-02	8.91E-03
2.86E-04	5.77E-03	8.25E-03	2.92E-02	3.58E-02	5.74E-03
2.86E-04	1.05E-02	1.28E-02	2.25E-02	3.93E-02	4.19E-03
3.58E-04	1.50E-02	1.74E-02	1.76E-02	3.96E-02	4.97E-03
5.61E-04	1.96E-02	2.08E-02	1.57E-02	3.58E-02	8.28E-03
8.80E-04	2.41E-02	2.31E-02	1.65E-02	2.86E-02	1.37E-02
1.22E-03	2.82E-02	2.49E-02	1.91E-02	1.97E-02	1.99E-02
1.48E-03	3.11E-02	2.74E-02	2.21E-02	1.12E-02	2.57E-02
1.55E-03	3.26E-02	3.12E-02	2.35E-02	5.25E-03	2.93E-02
1.41E-03	3.27E-02	3.65E-02	2.25E-02	2.99E-03	2.94E-02
1.12E-03	3.19E-02	4.24E-02	1.97E-02	4.27E-03	2.59E-02
8.28E-04	3.12E-02	4.77E-02	1.65E-02	7.87E-03	1.96E-02
6.63E-04	3.14E-02	5.07E-02	1.48E-02	1.19E-02	1.23E-02
7.30E-04	3.31E-02	5.06E-02	1.59E-02	1.47E-02	4.77E-03
1.04E-03	3.61E-02	4.74E-02	2.04E-02	1.56E-02	3.23E-03
1.49E-03	3.95E-02	4.22E-02	2.76E-02	1.48E-02	8.55E-03
1.93E-03	4.22E-02	3.70E-02	3.56E-02	1.27E-02	1.25E-02

2.18E-03	4.28E-02	3.38E-02	4.24E-02	1.05E-02	1.57E-02
2.11E-03	4.05E-02	3.41E-02	4.66E-02	9.18E-03	1.85E-02
1.70E-03	3.54E-02	3.79E-02	4.75E-02	8.97E-03	2.05E-02
1.06E-03	2.84E-02	4.43E-02	4.57E-02	9.59E-03	2.19E-02
3.51E-04	2.09E-02	5.14E-02	4.28E-02	1.02E-02	2.21E-02
2.13E-04	1.48E-02	5.71E-02	4.04E-02	9.83E-03	2.05E-02
4.81E-04	1.13E-02	5.98E-02	3.97E-02	8.05E-03	1.70E-02
3.97E-04	1.11E-02	5.87E-02	4.10E-02	5.16E-03	1.20E-02
1.79E-05	1.39E-02	5.44E-02	4.33E-02	2.28E-03	6.66E-03
5.14E-04	1.87E-02	4.83E-02	4.50E-02	1.46E-03	2.10E-03
1.03E-03	2.43E-02	4.19E-02	4.46E-02	1.95E-03	2.19E-03
1.39E-03	2.94E-02	3.67E-02	4.10E-02	6.14E-03	3.60E-03
1.53E-03	3.32E-02	3.33E-02	3.45E-02	1.32E-02	3.34E-03
1.50E-03	3.55E-02	3.21E-02	2.63E-02	2.14E-02	3.32E-03
1.38E-03	3.64E-02	3.16E-02	1.88E-02	2.88E-02	4.08E-03
1.28E-03	3.63E-02	3.07E-02	1.35E-02	3.36E-02	6.50E-03
1.27E-03	3.56E-02	2.88E-02	1.14E-02	3.48E-02	1.10E-02
1.34E-03	3.42E-02	2.57E-02	1.27E-02	3.22E-02	1.74E-02
1.42E-03	3.17E-02	2.22E-02	1.62E-02	2.72E-02	2.48E-02
1.40E-03	2.78E-02	1.94E-02	2.01E-02	2.19E-02	3.16E-02
1.19E-03	2.21E-02	1.84E-02	2.31E-02	1.83E-02	3.69E-02
7.64E-04	1.50E-02	1.92E-02	2.40E-02	1.79E-02	3.91E-02
1.87E-04	7.03E-03	2.10E-02	2.27E-02	2.08E-02	3.79E-02
4.09E-04	1.05E-03	2.26E-02	2.04E-02	2.60E-02	3.44E-02
8.64E-04	7.11E-03	2.24E-02	1.88E-02	3.19E-02	2.93E-02
1.06E-03	1.07E-02	1.94E-02	1.97E-02	3.68E-02	2.37E-02
9.46E-04	1.15E-02	1.35E-02	2.34E-02	3.96E-02	1.87E-02
6.04E-04	1.03E-02	5.59E-03	2.93E-02	4.06E-02	1.47E-02
1.80E-04	8.63E-03	3.71E-03	3.58E-02	4.02E-02	1.19E-02
1.42E-04	7.69E-03	1.03E-02	4.08E-02	3.92E-02	1.01E-02
2.18E-04	8.44E-03	1.34E-02	4.28E-02	3.85E-02	9.08E-03
5.81E-06	1.12E-02	1.25E-02	4.11E-02	3.85E-02	8.69E-03
4.71E-04	1.56E-02	8.17E-03	3.61E-02	3.87E-02	9.10E-03
1.03E-03	2.04E-02	2.09E-03	2.94E-02	3.85E-02	1.06E-02
1.49E-03	2.44E-02	5.03E-03	2.30E-02	3.68E-02	1.34E-02
1.70E-03	2.58E-02	8.80E-03	1.84E-02	3.34E-02	1.75E-02
1.58E-03	2.41E-02	9.67E-03	1.61E-02	2.81E-02	2.23E-02
1.18E-03	1.93E-02	7.75E-03	1.58E-02	2.18E-02	2.70E-02
5.98E-04	1.24E-02	4.11E-03	1.59E-02	1.61E-02	3.06E-02
1.28E-05	4.77E-03	8.49E-04	1.49E-02	1.26E-02	3.23E-02
4.42E-04	2.81E-03	3.20E-03	1.17E-02	1.22E-02	3.09E-02
6.98E-04	8.83E-03	3.83E-03	6.19E-03	1.47E-02	2.66E-02
7.69E-04	1.32E-02	3.24E-03	1.80E-03	1.89E-02	2.10E-02
7.38E-04	1.62E-02	1.82E-03	8.60E-03	2.27E-02	1.50E-02
7.08E-04	1.84E-02	1.58E-03	1.32E-02	2.41E-02	9.80E-03

7.50E-04	1.97E-02	2.43E-03	1.43E-02	2.17E-02	6.26E-03
8.71E-04	2.03E-02	2.40E-03	1.17E-02	1.53E-02	4.51E-03
1.01E-03	1.98E-02	2.02E-03	6.61E-03	6.05E-03	4.02E-03
1.06E-03	1.80E-02	2.31E-03	8.51E-04	4.15E-03	4.30E-03
9.26E-04	1.43E-02	4.26E-03	4.00E-03	1.30E-02	3.91E-03
5.82E-04	9.11E-03	8.41E-03	6.32E-03	1.92E-02	1.46E-03
7.73E-05	3.16E-03	1.38E-02	6.02E-03	2.20E-02	5.26E-03
4.60E-04	2.72E-03	1.90E-02	3.97E-03	2.21E-02	1.20E-02
8.67E-04	6.65E-03	2.25E-02	1.79E-03	2.09E-02	1.97E-02
1.01E-03	7.77E-03	2.30E-02	1.85E-03	1.99E-02	2.73E-02
8.20E-04	5.71E-03	2.01E-02	1.42E-03	2.03E-02	3.39E-02
3.55E-04	1.06E-03	1.44E-02	3.85E-03	2.23E-02	3.89E-02
2.47E-04	5.02E-03	7.48E-03	7.03E-03	2.53E-02	4.23E-02
7.99E-04	1.07E-02	1.28E-03	9.10E-03	2.83E-02	4.42E-02
1.13E-03	1.46E-02	2.61E-03	8.52E-03	3.03E-02	4.48E-02
1.16E-03	1.58E-02	2.80E-03	4.60E-03	3.08E-02	4.45E-02

C. Endurance Limit Factors :

Endurance limit (S_e) which is modified force by considering some modification factors are rendered in the following (Nisbett et al. (2015)):

$$S_e = S_e' \cdot C_s \cdot C_z \cdot C_l \cdot C_t \cdot C_r \cdot C_m$$

Based on the S_e formulation, the definition of different components are rendered as:

S_e' : Unmodified endurance limit

Surface Factor (C_s): Accounts for surface roughness.

Size Factor (C_z): Adjusts for the size of the component

Load Factor (C_l): Reflects the type of loading (e.g., axial, bending, torsion)

Temperature Factor (C_t): Accounts for operating temperatures

Reliability Factor (C_r): Higher reliability results less allowable stress.

Miscellaneous Factors (C_m): Account for other specific conditions like corrosion