EVALUATION AND DETECTION OF LEAKS IN A LABORATORY-SCALE WATER DISTRIBUTION SYSTEM WITH ACOUSTIC, ACCELERATION, AND DYNAMIC PRESSURE SENSORS

A Dissertation

by

MOHSEN AGHASHAHI

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Margaret Katherine Banks
Co-Chair of Committee,	Kelly Brumbelow
Committee Members,	Mohsen Pourahmadi
	Radu Stoleru
Head of Department,	Robin Autenrieth

August 2021

Major Subject: Civil Engineering

ABSTRACT

Aging water distribution systems waste millions of gallons of treated water due to background leaks. In this study, a laboratory-scale water network with 7.5 m × 5 m dimensions was developed to simulate background leaks in networks with looped and branched architectures. Four types of leaks, orifice, longitudinal and circumferential cracks, and gasket, were induced in the test system to generate leak signals. Six sensors, including two hydrophones, two dynamic pressure sensors, and two accelerometers, were employed to measure testbed parameters. With induced leak rates less than thirty percent, sixteen plots and numerical features were employed to assess the leak and network changes' effects on measured data. Due to the inconsistent patterns and similar magnitudes of the plots and features, the sixteen evaluation criteria did not represent specific patterns, and the metrics' changes depended on the sensors' locations. Based on the information extent they represented to differentiate leaks and network architecture, the sensors ranked as (1) dynamic pressure sensor, (2) hydrophone, and (3) accelerometer. Hydrophone acoustic signals were employed to detect leaks using five shallow classifiers, including Support Vector Machines (SVM), one-class SVM (1CSVM), Isolation Forest (iForest), Extreme Gradient Boosting (XGBoost), and Local Outlier Factor (LOF). A wavelet transform was applied to raw signals to compute the wavelet coefficients' moduli and create a matrix. A subsampled feature matrix of the looped network was used to generate imbalanced training and test datasets with imbalanced leak and non-leak class ratios. Testing the classifiers on the looped network's imbalanced data with original features showed SVM and XGBoost ranked first in predicting leak and non-leak samples, respectively. Using the looped network's imbalanced data with reduced features showed the same algorithms' ranks but with lower F_1-scores for all algorithms. Evaluation of the branched network's acoustic imbalanced data with original and reduced dimensions

indicated more mixed data distributions, and lower F_1-scores than the looped network. The analysis of the looped network's balanced data with original and reduced features resulted in higher F_1-scores of the algorithms in detecting leaks than their counterparts using imbalanced datasets.

DEDICATION

To my family

ACKNOWLEDGEMENTS

I would like to especially thank my committee chair, Professor M. Katherine Banks, who has always supported me endlessly.

I also thank my committee members, Professor Kelley Brumbelow, Professor Mohsen Pourahmadi, Professor Radu Stoleru, and Dr. Amin Rasekh, for their guidance and support throughout this research.

Dr. Rasekh provided invaluable insights and comments about the water distribution systems' current leak detection challenges, which led to my research objectives. I much appreciated his advice.

Moreover, I am grateful for the support of Professor Lina Sela Preleman from the University of Texas at Austin, who provided essential guidance for leak detection methodologies.

Finally, I want to thank many others at Texas A&M University who made my education a joyful journey.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor M. Katherine Banks and Professor Kelly Brumbelow from the Department of Civil Engineering, Professor Mohsen Pourahmadi from the Department of Statistics, Professor Radu Stoleru from the Department of Computer Science and Engineering, and Dr. Amin Rasekh as a Texas A&M Engineering Experiment Station consultant.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

The testbed construction was made possible in part by the Texas Sea Grant Grants-in-Aid of Graduate Research Program 2019-2021 under the grant NA18-2019-E/GIA-2019-Aghash. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Texas Sea Grant College Program at Texas A&M University.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

1. INTRODUCTION

1.1. Background and Motivation

As the last chain of water supply systems, water distribution systems (WDSs) are critical infrastructures that have been designed and operated to supply safe and continuous water flow for residential, industrial, commercial, and fire prevention purposes. The American Water Works Association (AWWA, 1974) describes water networks as "including all water utility components for the distribution of finished or potable water by means of gravity storage feed or pumps though distribution pumping networks to customers or other users, including distribution equalizing storage."

Resembling the modern water networks, the Minoan society employed a network of terracotta pipe conduits buried under the Knossos palace at depths up to 3 m. Roman cities like Pompeii were equipped with some aqueducts performing as urban distribution networks delivering water from masonry water tanks to public fountains (De Feo et al., 2011). As subterranean water tunnels, Qanats have been excavated, reported from 300-1000 BC in the UAE, Oman, and Iran, to capture and store winter precipitation and distribute them throughout a year for irrigation and drinking (Manuel et al., 2017).

Emulating ancient Roman water supply systems, the first water utility in the United States (U.S.) was launched in 1652 in Boston to supply potable fire suppression water (National Research Council, 2006). In the U.S., approximately one million miles of pipelines conduct forty-two billion gallons of water across the country (American Society of Civil Engineers, 2017). According to a survey conducted by the U.S. Environmental

Protection Agency (USEPA), \$77 billion was required to maintain water pipelines and rehabilitate leaking or burst pipes between 1997 and 2017 (Selvakumar et al., 2002).

Water waste due to leaking pipes has been an ever-existing challenge across the world. In large cities of developing countries, ~40% of the water entering treatment facilities is wasted due to leaks (WHO & UNICEF, 2000). In Asia, a value of about \$9 billion is lost annually (Zhou et al., 2019). Leak rates are simplified indicators of the structural integrity state of a water distribution system. The rate is often estimated as the ratio of leak water to the total input water entering a WDS (Lee and Schwab, 2005). In North America, leak rates have been reported as 30 percent (Youcef-Toumi, 2010). According to the 2017 Infrastructure Report Card (American Society of Civil Engineers, 2017), between 14 to 18% of daily treated water is lost by leaky pipes, enough water to supply 15 million households each day. From a water-energy nexus viewpoint, leaking pipes also contribute to the waste of energy embedded in water supply systems. Based on the California Energy Commission report (2005), approximately 5% of energy use in California can be attributed to water conveyance, distribution, and treatment, where lost water due to leak could be a reason for the waste of this energy (Berger et al., 2016).

Regarding the water scarcity in the world specifically in developing countries (Lee and Schwab, 2005) and due to world population growth, leaks from pipelines have been a major issue for water utilities to resolve. Municipalities have invested in smart water networks and automated leak detection systems to address leaky pipes and as an alternative to pipeline replacements (Stephens et al., 2020).

1.2. Leak Detection Methods

Leaks in water networks are two-fold: burst and background leaks. Background leaks are very small and difficult to detect without professional devices. With a 50 m pressure head in the Torricheli equation, leaks with areas smaller than 3.4 mm^2 are background leaks (Schwaller and van Zyl, 2015). Hereafter in this document, the word "leak" refers to a background leak. Water exiting a leak hole or crack causes changes in flow, the stability of pipe or propagating sounds in pipes. For example, a leak decreases the volume of water flow passing through a pipe, alters vibrations of the pipe wall, or creates acoustic emissions that influence sound characteristics in a pipe. All of these changes leave signatures that can be detected by comparing sensor data in conditions with a leak and without a leak. Based on available technology and equipment accuracy, there are multiple leak detection methods.

Mohd-Ismail et al. (2019) classified leak detection techniques into three categories as software-based, hardware-based, and conventional methods, each of which has subcategories (Figure 1.1). In conventional methods, experts walk along the pipeline path and investigate anomalous signs, such as odor or soil moisture. Software-based techniques are used to analyze pipe and flow parameters such as flow rate, pressure, or sound contents in the pipe. Hardware-based methods, such as tracer gas injection, rely on visual observations using specialized equipment. However, neither of these methods has been solely successful in leak detection, and each has its pros and cons

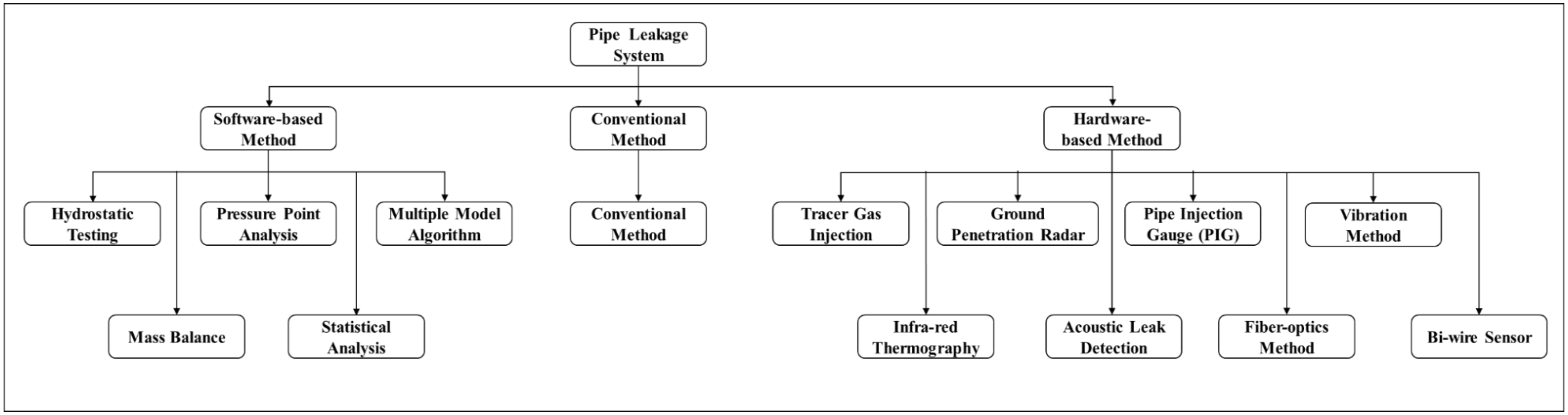

Figure 1.1. An overview of leak detection methods in water distribution systems (Mohd Ismail et al., 2019)

For instance, the pressure point analysis method, as a software-based technique, results in many false alarms with pipe pressure drop. Since conventional methods rely on personnel, they are often inaccurate for small leak detections. Using an acoustic hardware-based method, a person attempts to listen for captured sounds, often along with ambient noises, to identify an abnormal sound as a possible leak location. However, such methods are expensive, time-consuming, and inappropriate for long-range leak detections. Li et al. (2014) described software-based techniques as generally inexpensive, accurate for leak detection, and dependent on sophisticated algorithms. The authors classified the software-based technique as numerical and non-numerical model methods. The former includes an inverse transient method and frequency-domain and time-domain analyses, which require hydraulic parameters and detailed information of pipelines for simulations. While the latter are also called data-driven methods, they apply artificial intelligence (AI) algorithms to analyze sensory data. Though the non-numerical methods are data demanding and computationally expensive, with the extensive employment of the supervisory control and data acquisition systems (SCADA) in water networks operations, availability of low-cost and accurate sensors, and increasing capacity of computational machines, these methods have become promising and popular for leak detection.

Leak detection studies are performed either on real WDSs or laboratory-scale testbeds. A number of research projects has utilized actual water networks as case studies (Martini et al., 2015; Mounce et al. 2010; Brennan et al., 2018; Ma et al., 2019; Gao et al., 2018; Eliades and Polycarpou, 2012; Allen et al., 2011; Moser et al., 2018). The advantage of these studies is that leak detection experiments are conducted under real-world conditions where vital factors such as water consumption, pipe vibration, background noise, sound propagation velocity in pipes, media surrounding pipes, air temperature, and soil moisture influence parameters as they are in reality. These conditions can increase the reliability of results, mainly when sensors with low sensitivities are employed for measurements. Nevertheless, due to probable consequences of any unexpected change in a WDS operation caused by an experiment and because of security protocols, water utilities are reluctant to use their water systems in research case studies. Moreover, in some leak detection studies, controlling a parameter is necessary to investigate the effect of the parameter on the leak signature. For example, Hunaidi and Chu (1999) evaluated the effects of multiple factors, including season changes, on acoustic signals generated by leaks in a real size water testbed. Seasonal changes are often represented by water temperature and soil moisture. The authors monitored parameters of interest for months to investigate the effects of seasonal changes. Yet, water temperature and soil moisture variations could be simulated faster and easier in a laboratory scaled testbed where parameters can be changed in a controlled manner.

On the other hand, laboratory scaled testbeds provide more control over influencing factors and provides freedom to change parameters without the concern of interrupting a WDS operation and its subsequent consequences for consumers. In addition, a laboratory-scale WDS can be assembled at a low cost and in a limited area. Nonetheless, experimental results of these testbeds may be limited to less complicated components, smaller pipe dimensions, lower water pressure, and velocity, and different sound frequencies in pipes compared to those in actual WDSs. Though dimensional analysis principles and similitudes are solutions for more realistic test setups, due to the large scale of actual water networks, assembling a similitude is often impossible due to the lack of space in research facilities, especially on academic campuses. Therefore, small testbeds have played a major role in leak detection studies.

Table 1.1 lists and briefly describes laboratory-scale testbeds used in WDS leak detection studies. Except for the testbed used by Cody et al. (2018), configurations of other test systems are not close to actual conditions. For instance, water consumption has not been simulated, pipe diameters are much smaller than those used in actual WDSs, and if acoustic emission techniques are used for detection, ambient noise and backfill media are ignored.

Table 1.1. Laboratory-scale testbeds for leak detection

Author(s)	Objective(s)	Testbed Characteristics	Method(s)
Kartakis et al. (2015)	- Simulating water district metering areas (DMAs) of a water network - Leak simulations for educational purposes	- A compact designed and closed-loop testbed with transparent components with a reservoir, DMA tanks, and valves as demand nodes	- NA
Kadri et al. (2011)	- Leak detection with a wireless real-time monitoring system using a hydrophone	- A single 30 m long polypropylene pipe with 160 mm internal diameter - Leak simulated with a tap - A hydrophone in the vicinity of the leaking tap	- Visual comparison of time-domain and frequency-domain signals and labeling amplitude peaks as leaks
El-Zahab et al. (2018)	- Leak detection using accelerometers and AI algorithms	- Single ductile iron and PVC 6.5 m long pipes with 1 in. and 2 in. diameters - Accelerometers located at about every 1 m - Leaks simulated with two ball valves	- Classification with Naïve Bayes, Linear SV, and Decision Tree along with cross-validation
Karray et al. (2016)	- Leak detection with an energy-efficient wireless pressure sensor network	- A non-looped U-shaped testbed composed of 25 m polyethylene pipes with 32 mm external diameter - Two leaks simulated with garden taps - Two pressure sensors at the vicinity of the leaks	- Kalman Filter for in-node data preprocessing and linear anomaly detection

Table 1.1. Continued

Author(s)	Objective(s)	Testbed Characteristics	Method(s)
Butterfield et al. (2017)	- Quantifying leak flow rate using vibroacoustic (VA) sensors - Investigating the effects of pipe surrounding media on leak generated VA signals	- A closed-circuit oval-shaped test rig built of 140 m long, 50 mm diameter polyethylene pipe with a 3.5 kW variable speed pump - Leak induced with drilling a 1 mm diameter hole in the pipe - An accelerometer with 10 V/g sensitivity at 36 cm next to the leak	- Investigating the correlation between leak flow rate and VA emission counts, signal root mean square, peak in the magnitude of the power spectral density, and octave banding
Sadeghioon et al. (2014)	- Leak detection based on pressure changes - Development a wireless network of pressure sensors with low-power consumption	- A U-shaped test bench made of 40 mm diameter PVC pipe with a pump providing 3 bars pressure - Leak induced with a 10 mm diameter hole in the middle of the U-shaped pipe - Five pressure sensors (force-sensitive resistors sensors) mounted on pipe wall at 2 m intervals	- Visualization of time-domain pressure sensor signals and labeling pressure drops as leaks
Yazdekhasti et al. (2016)	- Leak detection by monitoring changes of correlation among vibration signals of multiple accelerometers along a pipeline	- A U-shaped 16 m long test system with 76 mm diameter PVC pipes and 1100 lit/min flow rate. - Leaks induced by ball valves in the middle of the pipe	- Comparison of a damage index with a threshold where the index is based on the difference of cross spectral-density functions of acceleration data of pipes with leak and without leak

Table 1.1. Continued

Author(s)	Objective(s)	Testbed Characteristics	Method(s)
Yazdekhasti et al. (2016) (continued)	- Formulation of an index to detect the onset and severity of leaks based on the cross-spectral density of pipe surface acceleration	- Six accelerometers mounted on the pipeline at three points where at each point one sensor is parallel to, and another is perpendicular to the leak position	- Investigating the correlation between leak flow rate and VA emission counts, signal root mean square, peak in the magnitude of the power spectral density, and octave banding
Cody et al. (2018)	- Leak detection using hydro-acoustic signals and machine learning algorithms	- A 15 m long branched testbed with 15.24 cm diameter PVC pipes - A 2.5 cm service line simulating water consumption - Three leaks induced with 0.32 cm valves located in the middle of pipe sections - A hydrophone located at the base of a fire hydrant connected to the network	- Applied singular spectrum analysis to extract leak related features from noisy signals - Used one-class support vector machine (SVM) to classify leak signals versus normal signals
Li et al. (2018)	- Leak detection in socket joints of WDSs made of ductile iron pipes using acoustic emission techniques and artificial intelligence networks	- A closed-loop system composed of two ductile iron pipes of 3 m length and 200 mm diameter connected by a socket joint - A displacement gauge mounted at the socket joint to measure the relative displacement of the joint	- Used time-domain (peak, mean, standard deviation, root mean square, crest factor, and energy) and frequency-domain (kurtosis, peak frequency, skewness, and frequency centroid) features

Table 1.1. Continued

Author(s)	Objective(s)	Testbed Characteristics	Method(s)
Li et al. (2018) (continued)	(continued)	-Two acoustic sensors attached to the pipe segment, each at one side of the joint	- Applied a shallow artificial neural network with a back-propagation algorithm for a leak versus normal signal classification

1.2.1. Leak Detection via Artificial Intelligence Algorithms

With the extensive application of AI algorithms in engineering disciplines, availability of cloud-based sensory data, and inexpensive computational resources, data-driven methods have become promising solutions for the leak detection problem in the water industry and academia. The main advantage of AI-based leak detection methods is that there is no need for a hydraulic model and detailed parameters of the network and its equipment. However, these methods often require significant amounts of data for training and should be updated regularly to reflect network data changes in predictions. Therefore, they may be slow in leak detection and result in delayed alarms in online detection applications (Li et al., 2015). Though the focus of this research is on background leaks, for a more comprehensive study, literature that applied AI to detect bursts, in addition to background leaks, will be reviewed.

Aksela et al. (2009) conducted a leak detection study where they employed an unsupervised method based on the self-organizing map (SOM) in a real WDS using flow meter readings and reported leak locations. The authors embedded a leak function, indicating the distances of reported leaks to flow meters, in the SOM to construct a map whose cells are the leak function values and show the probability of existing leaks in flowmeter measurements as input data. Once trained, the

SOM was used to predict leak functions for new flow measurements where higher values of leak functions indicated the flow data were similar to data when leaks have occurred in training. Therefore, the vicinity of that flow meter should be investigated as a potential leak area. Though this method showed promising results, it requires weekly meter data for updating and new training and cannot provide leak alarms in real-time.

Mounce et al. (2010) developed an online leak or burst detection platform at DMAs of a WDS in the United Kingdom (U.K.). The platform employed a mixture density network, as an artificial neural network (ANN) model, to predict a probability density model of the future flow profile and a fuzzy inference system to classify new flow data and compare them with the predicted flow profiles. Based on the difference between an actual new flow data and its corresponding predicted flow value, an alert would indicate a detected leak or burst. The authors also used the fuzzy system to rank the alerts by means of confidence intervals associated with each detection. Though the platform was tested successfully in an actual WDS, since the hybrid algorithm used a threshold to flag an alert based on historical data and a manual tunning, continued tunning the threshold and the ANN hypermeters require a person with ANN and domain knowledge who is rarely available in water network control rooms.

Other research on burst detection was published for several actual DMAs in the U.K. Ye and Fenner (2011) applied an adaptive Kalman filter on flow and pressure measurements to model the normal status of the measurements. Then, they used the residuals of the filter, the difference between the predicted parameter and its measured value, to predict if an abnormal water usage or pressure data is related to a burst or a newly occurred leak. Though the Kalman filter is not an ANN method and is a signal processing technique, its advantages, such as computational efficiency, quick detections, and no requirements for large training data, makes it worthy of discussion. Test results in an actual WDS indicates that the algorithm is well suited to detect sudden bursts or gradually changing leaks and less efficient for detecting long-term stable leaks. This might be because the normal data by which the model was trained included stable leaks, and the algorithm considered such leaks as a normal condition in training. Besides, the algorithm may give false alarms if a sudden industry water usage in large volumes occurs in a DMA.

Mounce et al. (2011) applied support vector regression to predict new values of flow and pressure parameters. Then, if the difference between the predicted parameter and its actual value is greater than a tolerance width, the actual observation would be labeled as a possible event. If there were enough events within a sliding window, compared to those in the training data, those events would be classified as novelties. The authors implemented the algorithm in the case study of the research conducted by Mounce et al. (2010), where an online system combining artificial neural networks and fuzzy inference system (ANN/FIS) was used to detect anomalies. The comparison of these two methods showed that the SVR algorithm could detect anomalies faster than the hybrid ANN/FIS counterpart.

Romano et al. (2014a) developed a near-real-time anomaly detection method at the DMA level using pressure and flow data. The method included wavelets for signal preprocessing, ANNs for short-term signal forecasting, statistical process control techniques for short- and long-term analysis of flow and pressure variations, and Bayesian inference systems in calculating the probability of pipe failures based on the predictions. The authors tested the methodology on a case study with several DMAs in the U.K. with simulated and real-life bursts where results indicated the method could detect events quickly and with low false alarms. Though the method performed well, the anomaly detection system had multiple parameters that needed to be set experimentally and to be updated continuously with new data. This problem was addressed by the same authors in another study, Romano et al. (2014b), where they developed a methodology to recalibrate their platform automatically. The recalibration employed an evolution algorithm optimization strategy to choose the best ANN structures and hyperparameters and an expectation-maximization algorithm for recalibrating the conditional probability tables of the Bayesian inference system. The results illustrated that the recalibration procedure could improve their previous anomaly detection system's reliability and speed. Nonetheless, the complex structures of both the initial and the optimized systems make them highly unexplainable. At the same time, the interpretability of AI algorithms has been proved to be necessary to make the algorithms applicable (Samek et al., 2019).

Similar to this study, Li et al. (2018) applied signal preprocessing and an ANN to detect a leak from the socket joint in a ductile iron pipe segment using acoustic signals. They first extracted features as representatives of the raw acoustic emission data. The features were both in the time domain (peak, mean, standard deviation, root-mean-square, crest factor, and energy) and frequency domain (skewness, kurtosis, peak frequency, and frequency centroid). The authors then selected the best features for classification. The feature selection criterion was cross-entropy that measures the distance between the probability membership of a sample in a leak and non-leak classes. Selected features were employed to train a three-layer ANN algorithm with back-propagation. Results showed that among peak, mean, kurtosis, and peak frequency as selected input features for the ANN algorithm, the combinations of those in the frequency domain gave better training performances. In the test step, the algorithm performed with the accuracy of 96.9% when mean, peak, and frequency were employed as features. Since acoustic emission data can be affected by background noise and pipe backfill media and the authors did not include these factors in their testbed, their generated data could be well separable by nature and did not mimic the complexity of realistic joint leak data, so their analysis would not be comprehensive enough for a actual water network.

Convolutional neural networks (CNNs) have been promising methods in the leak detection field. Kang et al. (2018) and Chuang et al. (2019) deployed the CNN in leak detection, where the former used one-dimensional CNN as a feature extractor and the latter applied a CNN as a classifier. Kang et al. (2019) developed the ensemble 1D-CNN-SVM model where normalized vibration data was fed to a CNN for feature extraction. Then the extracted feature vectors were used to train a fully-connected multi-layer perceptron (MLP) model, an SVM, and a combination of the MLP model and the SVM to classify a leakage. The results showed the ensemble 1D-CNN-SVM has the best performance with the area under the receiver operating characteristic curve (AUC) of 0.99.

On the other hand, Chuang et al. (2019) used mel-frequency cepstral coefficients and their first-order and second-order differences to extract features from acoustic signals. Then, the authors trained a CNN to classify new data as leak or non-leak. The performance of the algorithm on the test data had an AUC above 98%. Similar to other ANN models, CNNs have multiple hyperparameters that need tunning, which is time-consuming. The hyperparameters must be updated continuously to correspond with high computational complexities if the training data is large or massive.

A group of researchers at the University of Waterloo studied leak detection using a testbed described in Table 1.1. In the following section, AI algorithms used in the literature related to the testbed are reviewed. It is worth noting that background noise was ignored during acoustic data generation in the testbed. Cody et al. (2018) applied singular spectrum analysis (SSA) to extract leak signatures from noisy acoustic data. They employed the SSA because it is an assumption-free and a non-parametric method. Then, SVM and one-class SVM were trained and used for classification. Taking advantage of the radial basis function (RBF) as a kernel, classification results showed SVM's good performance with the AUCs in the range [0.85, 0.92] and one-class SVM with the AUC up to 0.90. However, these results may not be reproducible in actual WDSs. Due to data labeling nature in actual water networks, leak-free data might be labeled incorrectly while they are truly leaky data. Therefore, in building the one-class SVM algorithm, training data is better to include some leak data to allow for some artificial training errors and those of the algorithm. Yet, Cody et al. trained the one-class SVM regardless of this point. Harmouche and Narasimhan (2020) developed a data-driven approach to detect leaks with hidden signatures in long-term acoustic data without expert knowledge or controlled experiments. They used association rules to extract information from noisy acoustic data with small variations. Then a leak indicator was developed to capture deviation of leak data from a reference leak-free data. Their results showed the indicator could detect small leaks with high accuracy. Cody et al. (2020) presented the linear prediction technique (LP), which utilizes the LP coefficients' features representing the hidden acoustic signals. In the LP, the response of a pipe-fluid system was modeled, and was used as a basis to determine leak presence. The authors applied principal component analysis (PCA) on the LP coefficients to generate well-separated features for anomaly detection using a Gaussian mixture model (GMM). The GMM performed better with LP-PCA features compared with the time-domain features used in (Li et al., 2019). In addition to the data deficits stemmed from the testbed constraints, such as ignoring background noise, pipe backfill media, and the effects of different leak sizes, the GMM has a threshold to discern leak and leak-free data. The authors needed to adjust the threshold based on the input data. The dependency of the threshold on the data types may make it impractical for leak detection in a real WDS.

Cody and Narasimhan (2020) employed the LP-PCA and GMM to detect leaks in a Canadian network. The preliminary results show good performance of the leak detection method, specifically when short-term monitoring of the network was available. Nonetheless, the authors induced leaks by opening fire hydrants at flow rates of 25 liters per minute (LPM), 50 LPM, 100 LPM, and 200 LPM. Since the ratio of the induced leak flows to the network water consumption was not determined, one cannot guarantee that the flow rate of the detected events is either in the scale of background leaks or that of bursts. If the latter is true, then the research may not be able to detect background leaks with latent signatures in acoustic signals.

1.3. Gaps and Challenges

Leak detection in water networks depends on a number of factors, including network architecture, leak size and shape, consumption flow variations, sensor types and location, ambient noise, and materials surrounding pipes. Though many studies have investigated leak detection in actual systems or testbeds, few or none of them have approached the leak detection challenge in a generalized way so that the effects of the aforementioned factors are investigated.

Due to a reluctance of water utilities to share water network data, data availability for research and educational purposes is one of the most serious issues faced by experts and researchers in the WDS industry. Even researchers with access to real data require a facility to conduct research in a controlled manner. Developing test WDSs has been an alternative to generate the required data. Yet, most testbed characteristics are far from their actual counterparts which leads to synthetic data not reflecting real-world conditions.

AI algorithms have been extensively used and resulted in promising solutions to detect leaks in WDS. Mainly focused on ANNs, the algorithms require time-consuming adjustments to set optimum hyperparameters and network architectures. Regarding the dynamic nature of WDSs, a fitted ANN needs continuous tunning and modifications to detect leaks in continuously changing water network data. These necessary algorithmic adjustments require either personnel with AI and expert knowledge or online learning and self-optimizing algorithms. The former is rarely available in water utilities, and the latter would suffer from complexity and a lack of interpretability.

1.4. Research Objectives

This research has three main objectives. The first is to design and assemble a laboratory-scale water distribution network to generate data for leak detection studies and make them available for leak detection researchers. Though few testbeds have been developed for leak detection, they have excluded important criteria that make the collected data and analytical results valid only in specific conditions. The research testbed used in this study included ambient noise, sensor types, network architectures, and hydraulic factors to mimic real-life WDS data.

The second goal is to study leak signatures and to evaluate how different leak types and network architectures affect selected metrics. This objective provided information about leak characteristics using hydrophone, accelerometer, and dynamic pressure sensor data. We used six types of plots to investigate data without demand and sound interruptions. Ten features were also employed to numerically assess leak and network change influences.

The third main objective is to apply simple but efficient preprocessing and machine learning algorithms to detect simulated leaks using acoustic data captured by hydrophones. In the past, research carried out to develop a leak detection methodology employed data-dependent algorithms and required a cumbersome effort to build an optimum anomaly detection model. This research employed the least complicated but accurate methods with only a minimum AI and water network experience required to apply to actual distribution systems.

The specific research goals are to:

1. study characteristics of actual water networks influencing leak behaviors;
2. assess previously built testbeds for leak detection studies;
3. design and assemble a more advanced laboratory-scale testbed to generate data representative field data;
4. describe and justify the testbed design criteria and assembly procedures;
5. analyze data to learn how leak type and network changes affect leak characteristics;
6. utilize acoustic data and employ simple but accurate preprocessing and classifying algorithms to detect leaks under experimental scenarios.

1.5. Contributions of the Dissertation

The following was accomplished.

1. A laboratory-scale testbed was designed and constructed to generate data for leak detection studies. Three types of sensors, i.e., hydrophone, accelerometer, and dynamic pressure sensors, were employed. Sensors at two locations were mounted on or inside the pipes to measure the acoustic emission, vibration, and differential pressure signals, respectively. This testbed's advantage is that important factors influencing leak signatures, including leak shape and size, network architecture, consumption flow, ambient noise, transient incidents, and resonance in pipe structures, were considered in its components and experiments.
2. The testbed design and assembling procedures were clearly described, and all steps and design criteria were justified to ensure the testbed resembled actual water networks as much as possible in recognition of the constraints.
3. Representative graphs of measured signals were plotted, and visual and feature-based comparisons were performed to evaluate how leak types and network architectures

influenced leak characteristics. This research is the most comprehensive study that has assessed these changes' effects on leak signatures regarding features and sensor data to the best of our knowledge.

4. A wavelet preprocessing method was applied to extract features from acoustic signals embedding latent leak signatures. Features were employed to train and test five shallow classifiers, Support Vector Machines, Extreme Gradient Boosting, one-class Support Vector Machines, Isolation Forest, and Local Outlier Factor to discern leak and non-leak signals. The results indicated a successful performance of the leak detection methodology, particularly on balanced data.

1.6. Dissertation Outline

The main body of this dissertation consists of three sections after this one. **Section 2** describes an advanced testbed designed and built with a focus on leak detection with the presence of hydraulic and sound disturbances. The testbed components are explained, and their functions are discussed**.** Sixten plots and features were utilized to determine how the leak and network changes influenced leak characteristics, using three sensor types. **Section 3** uses hydrophone data measured from two network architectures, i.e., looped and branched networks, and employs wavelet transforms to extract features and shallow classification algorithms for leak detection. The classifier algorithms were trained and tested with datasets having different row and column structures. And finally, **Section 4** concludes the dissertation and presents avenues for future research.

1.7. References

Aksela, K., Aksela, M., & Vahala, R. (2009). Leakage detection in a real distribution network using a SOM. *Urban Water Journal*, *6*(4), 279–289. https://doi.org/10.1080/15730620802673079

Allen, M., Prels, A., Lqbal, M., Srirangarajan, S., Llm, H. B., Glrod, L., & Whittle, A. J. (2011). Real-time in-network distribution system monitoring to improve operational efficiency. *Journal - American Water Works Association*, *103*(7), 63–75. https://doi.org/10.1002/j.1551-8833.2011.tb11495.x

American Society of Civil Engineers. (2017). *American Society of Civil Engineers 2017 Infrastructure Report Card*. https://www.infrastructurereportcard.org/wp-content/uploads/2017/01/Drinking-Water-Final.pdf

Berger, M. A., Hans, L., Piscopo, K., & Sohn, M. D. (2016). *Exploring the Energy Benefits of Advanced Water Metering*. Lawrence Berkeley National Laboratory. https://eta-publications.lbl.gov/sites/default/files/lbnl-1005988.pdf

Brennan, M. J., Karimi, M., Muggleton, J. M., Almeida, F. C. L., Kroll de Lima, F., Ayala, P. C., Obata, D., Paschoalini, A. T., & Kessissoglou, N. (2018). On the effects of soil properties on leak noise propagation in plastic water distribution pipes. *Journal of Sound and Vibration*, *427*, 120–133. https://doi.org/10.1016/j.jsv.2018.03.027

Butterfield, J. D., Krynkin, A., Collins, R. P., & Beck, S. B. M. (2017). Experimental investigation into vibro-acoustic emission signal processing techniques to quantify leak flow rate in plastic water distribution pipes. *Applied Acoustics*, *119*, 146–155. https://doi.org/10.1016/j.apacoust.2017.01.002

Chuang, W.-Y., Tsai, Y.-L., & Wang, L.-H. (2019). Leak Detection in Water Distribution Pipes Based on CNN with Mel Frequency Cepstral Coefficients. *Proceedings of the 2019 3rd International Conference on Innovation in Artificial Intelligence - ICIAI 2019*, 83–86. https://doi.org/10.1145/3319921.3319926

Cody, R. A., Dey, P., & Narasimhan, S. (2020). Linear Prediction for Leak Detection in Water Distribution Networks. *Journal of Pipeline Systems Engineering and Practice*, *11*(1), 04019043. https://doi.org/10.1061/(asce)ps.1949-1204.0000415

Cody, R., Harmouche, J., & Narasimhan, S. (2018). Leak detection in water distribution pipes using singular spectrum analysis. *Urban Water Journal*, *15*(7), 636–644. https://doi.org/10.1080/1573062x.2018.1532016

Cody, R. A., & Narasimhan, S. (2020). A field implementation of linear prediction for leak-monitoring in water distribution networks. *Advanced Engineering Informatics*, *45*, 101103. https://doi.org/10.1016/j.aei.2020.101103

De Feo, G., Mays, L. W., & Angelakis, A. N. (2011). Water and Wastewater Management Technologies in the Ancient Greek and Roman Civilizations. *Treatise on Water Science*, 3–22. https://doi.org/10.1016/b978-0-444-53199-5.00071-3

Eliades, D. G., & Polycarpou, M. M. (2012). Leakage fault detection in district metered areas of water distribution systems. *Journal of Hydroinformatics*, *14*(4), 992–1005. https://doi.org/10.2166/hydro.2012.109

El-Zahab, S., Mohammed Abdelkader, E., & Zayed, T. (2018). An accelerometer-based leak detection system. *Mechanical Systems and Signal Processing*, *108*, 276–291. https://doi.org/10.1016/j.ymssp.2018.02.030

Gao, Y., Liu, Y., Ma, Y., Cheng, X., & Yang, J. (2018). Application of the differentiation process into the correlation-based leak detection in urban pipeline networks. *Mechanical Systems and Signal Processing*, *112*, 251–264. https://doi.org/10.1016/j.ymssp.2018.04.036

Harmouche, J., & Narasimhan, S. (2020). Long-Term Monitoring for Leaks in Water Distribution Networks Using Association Rules Mining. *IEEE Transactions on Industrial Informatics*, *16*(1), 258–266. https://doi.org/10.1109/tii.2019.2911064

Kadri, A., Abu-Dayya, A., Trinchero, D., & Stefanelli, R. (2011). Autonomous sensing for leakage detection in underground water pipelines. *2011 Fifth International Conference on Sensing Technology*, 639–643. https://doi.org/10.1109/icsenst.2011.6137061

Kang, J., Park, Y.-J., Lee, J., Wang, S.-H., & Eom, D.-S. (2018). Novel Leakage Detection by Ensemble CNN-SVM and Graph-Based Localization in Water Distribution Systems. *IEEE Transactions on Industrial Electronics*, *65*(5), 4279–4289. https://doi.org/10.1109/tie.2017.2764861

Karray, F., Garcia-Ortiz, A., Jmal, M. W., Obeid, A. M., & Abid, M. (2016). EARNPIPE: A Testbed for Smart Water Pipeline Monitoring Using Wireless Sensor Network. *Procedia Computer Science, 96*, 285–294. https://doi.org/10.1016/j.procs.2016.08.141

Kartakis, S., Abraham, E., & McCann, J. A. (2015). WaterBox. *Proceedings of the 1st ACM International Workshop on Cyber-Physical Systems for Smart Water Networks - CySWater'15*, 118–124. https://doi.org/10.1145/2738935.2738939

Lee, E. J., & Schwab, K. J. (2005). Deficiencies in drinking water distribution systems in developing countries. *Journal of Water and Health*, *3*(2), 109–127. https://doi.org/10.2166/wh.2005.0012

Li, R., Huang, H., Xin, K., & Tao, T. (2014). A review of methods for burst/leakage detection and location in water distribution systems. *Water Supply*, *15*(3), 429–441. https://doi.org/10.2166/ws.2014.131

Li, S., Song, Y., & Zhou, G. (2018). Leak detection of water distribution pipeline subject to failure of socket joint based on acoustic emission and pattern recognition. *Measurement*, *115*, 39–44. https://doi.org/10.1016/j.measurement.2017.10.021

Ma, Y., Gao, Y., Cui, X., Brennan, M., Almeida, F., & Yang, J. (2019). Adaptive Phase Transform Method for Pipeline Leakage Detection. *Sensors*, *19*(2), 310. https://doi.org/10.3390/s19020310

Manuel, M., Lightfoot, D., & Fattahi, M. (2017). The sustainability of ancient water control techniques in Iran: an overview. *Water History*, *10*(1), 13–30. https://doi.org/10.1007/s12685-017-0200-7

Martini, A., Troncossi, M., & Rivola, A. (2015). Automatic Leak Detection in Buried Plastic Pipes of Water Supply Networks by Means of Vibration Measurements. *Shock and Vibration*, *2015*, 1–13. https://doi.org/10.1155/2015/165304

Mohd Ismail, M. I., Dziyauddin, R. A., Ahmad Salleh, N. A., Muhammad-Sukki, F., Aini Bani, N., Mohd Izhar, M. A., & Latiff, L. A. (2019). A Review of Vibration Detection Methods Using Accelerometer Sensors for Water Pipeline Leakage. *IEEE Access*, *7*, 51965–51981. https://doi.org/10.1109/access.2019.2896302

Moser, G., Paal, S. G., & Smith, I. F. C. (2018). Leak Detection of Water Supply Networks Using Error-Domain Model Falsification. *Journal of Computing in Civil Engineering*, *32*(2), 04017077. https://doi.org/10.1061/(asce)cp.1943-5487.0000729

Mounce, S. R., Boxall, J. B., & Machell, J. (2010). Development and Verification of an Online Artificial Intelligence System for Detection of Bursts and Other Abnormal Flows. *Journal of Water Resources Planning and Management*, *136*(3), 309–318. https://doi.org/10.1061/(asce)wr.1943-5452.0000030

Mounce, S. R., Mounce, R. B., & Boxall, J. B. (2011). Novelty detection for time series data analysis in water distribution systems using support vector machines. *Journal of Hydroinformatics*, *13*(4), 672–686. https://doi.org/10.2166/hydro.2010.144

National Research Council. (2006). *Drinking Water Distribution Systems: Assessing and Reducing Risks*. National Academies Press. https://doi.org/10.17226/11728

Romano, M., Kapelan, Z., & Savić, D. A. (2014a). Automated Detection of Pipe Bursts and Other Events in Water Distribution Systems. *Journal of Water Resources Planning and Management*, *140*(4), 457–467. https://doi.org/10.1061/(asce)wr.1943-5452.0000339

Romano, M., Kapelan, Z., & Savić, D. A. (2014b). Evolutionary Algorithm and Expectation Maximization Strategies for Improved Detection of Pipe Bursts and Other Events in Water Distribution Systems. *Journal of Water Resources Planning and Management*, *140*(5), 572–584. https://doi.org/10.1061/(asce)wr.1943-5452.0000347

Sadeghioon, A., Metje, N., Chapman, D., & Anthony, C. (2014). SmartPipes: Smart Wireless Sensor Networks for Leak Detection in Water Pipelines. *Journal of Sensor and Actuator Networks*, *3*(1), 64–78. https://doi.org/10.3390/jsan3010064

Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K. (2019). *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning (Lecture Notes in Computer Science (11700))* (1st ed. 2019 ed.). Springer.

Schwaller, J., & van Zyl, J. E. (2015). Modeling the Pressure-Leakage Response of Water Distribution Systems Based on Individual Leak Behavior. *Journal of Hydraulic Engineering*, *141*(5), 04014089. https://doi.org/10.1061/(asce)hy.1943-7900.0000984

Selvakumar, A., Clark, R. M., & Sivaganesan, M. (2002). Costs for Water Supply Distribution System Rehabilitation. *Journal of Water Resources Planning and Management*, *128*(4), 303–306. https://doi.org/10.1061/(asce)0733-9496(2002)128:4(303)

Stephens, M., Gong, J., Zhang, C., Marchi, A., Dix, L., & Lambert, M. F. (2020). Leak-Before-Break Main Failure Prevention for Water Distribution Pipes Using Acoustic Smart Water Technologies: Case Study in Adelaide. *Journal of Water Resources Planning and Management*, *146*(10), 05020020. https://doi.org/10.1061/(asce)wr.1943-5452.0001266

Water-Distribution Research and Applied Development Needs. (1974). *Journal - American Water Works Association*, *66*(6), 385–390. https://doi.org/10.1002/j.1551-8833.1974.tb02035.x

WHO & UNICEF. (2000). *Global Water Supply and Sanitation Assessment 2000 Report*. Iseman Creative, Washington, DC. https://www.who.int/water_sanitation_health/monitoring/jmp2000.pdf

Yazdekhasti, S., Piratla, K. R., Atamturktur, S., & Khan, A. A. (2016). Novel vibration-based technique for detecting water pipeline leakage. *Structure and Infrastructure Engineering, 13*(6), 731–742. https://doi.org/10.1080/15732479.2016.1188318

Ye, G., & Fenner, R. A. (2011). Kalman Filtering of Hydraulic Measurements for Burst Detection in Water Distribution Systems. *Journal of Pipeline Systems Engineering and Practice*, *2*(1), 14–22. https://doi.org/10.1061/(asce)ps.1949-1204.0000070

Youcef-Toumi, K. (2010). *Analysis and Design of an In-Pipe System for Water Leak Detection* ("Master's dissertation"). Massachusetts Institute of Technology. http://hdl.handle.net/1721.1/62529

Zhou, X., Tang, Z., Xu, W., Meng, F., Chu, X., Xin, K., & Fu, G. (2019). Deep learning identifies accurate burst locations in water distribution networks. *Water Research*, *166*, 115058. https://doi.org/10.1016/j.watres.2019.115058

2. EXPERIMENT SETUP, SCENARIOS, AND OBSERVATIONS

2.1. Introduction

With increased potable water scarcity and population growth, the importance of treated water has become more evident. Domestic water demand is expected to increase dramatically between 2010 and 2050 globally, by 300% in Africa and Asia, and by 200% in Central and South America (Boretti and Rosa, 2019). On the other hand, water wastage rates due to the aging water infrastructures have provided a clarion call to address the occurring leaks and bursts. In the U.S. alone, deteriorated pipes lose 6 billion gallons of treated water each day (Allen et al., 2018).

Many techniques have been proposed for water pipe monitoring and leak detection, including software- and hardware-based methods (Li et al., 2014; Mohd-Ismail et al., 2019). A significant benchmark to evaluate these methods is the use of case studies to acquire data and test the methods. There are two ways to investigate leak detection methodologies: (1) in actual water distribution systems; (2) in laboratory-scale testbeds.

In field-scale case studies, conditions of factors affecting leaks, including water consumption flows, structural vibrations, ambient noise, media surrounding pipes, sound propagation velocity in pipes, air temperature, and soil moisture, are considered in experiments. For example, the leak flow rate has a massive weight in leak detection accuracy. Wu et al. (2010) found their pressure-dependent leak detection method useful for hydraulic conditions that occur in early day hours, in WDS designed with extra capacity, and when data loggers work closely to their accuracy boundary. Simulating these hydraulic conditions in a testbed could be very costly, if not impossible. Research using actual field-scale case studies is more reliable, and methods can be employed in final applications.

However, regarding the significant role of drinking water networks in daily life and security regulations, water companies hesitate to jeopardize their network operation due to probable disturbances that leak detection studies may cause. Additionally, research requires controlled experiments where all parameters less one should be constant. Considering the interconnectivity of parameters in actual WDSs, setting a controlled experiment seems impossible. For instance, Kang et al. (2018) suggested performance evaluation of their leak classifier is needed when leak flow varies. Changing a background leak flow in a realistic water network is impossible unless a pipe would be drilled to simulate leaks with different sizes. Yet, using a bench-scale testbed allows simulating desired alterations in one parameter when the others are unchanged. However, laboratory-scale water networks may ignore design criteria that can make research results biased and useless. Cody et al. (2018) did not mention or consider the ratio of leak flow to the total input water. If this ratio were large, their generated acoustic data would not include masked background leak data.

Except for the testbed used by Cody et al. (2018) and Harmouche and Narasimhan (2020), other studies in Table 1.1 used pipes with diameters smaller than what is often used in distribution pipelines, i.e., 15.24 cm. The smaller pipe size influences sound propagation velocity, pipe vibration and resonance intensity, and water flow regime in pipes. Moreover, ambient noise, like rotating machinery or traffic sound, has not been simulated in other testbeds. Since leak and ambient noise have frequency overlap, usually under 500 Hz (Butterfield et al., 2017b), neglecting this noise can allow leak sounds to be more pronounced in acoustic signals that leads to misleading detection algorithms.

One criterion that most leak detection testbeds have in common is inducing leaks with valves regardless of leak shapes reported in real water networks (Greyvenstein and van Zyl, 2007).

In reality, leak water jets cause flow turbulences around the leaks. This hydraulic disturbance influences recorded parameters and make acoustic signals noisier. Employing a ball valve to induce a leak does not simulate a water jet and its hydraulic effects. Additionally, relevant studies have not evaluated the influences of leak shapes that can change leak signatures in recorded signals.

Like those described in Table 1.1, reviewing the literature highlights a lack of attention to the relation between water consumption variations, transient incidents, and leak signal characteristics. Water consumption is one of the primary noise sources that makes leak detection more challenging. Water exiting a service line generates more represented sounds compared to leak sounds, especially in acoustic signals. In real water networks, leak detection studies are often carried out between midnight and early morning (Wu et al., 2010). One reason for choosing this time period is less water consumption and subsequent effects on measured signals.

Furthermore, hydraulic transients created by rapid changes in a water network component can vary flow and pressure inside a pipe (Xing and Sela, 2020). Hydraulic transient-induced changes can suppress leak signals in recorded measurements. Thus, evaluating transients in leak detection testbeds, ignored in other test systems, can result in a more comprehensive leak study.

Different parameters can provide information about leaks and the correlation of each parameter with leak signals. A variety of sensors have been employed in leak detection studies. Two most widely used sensors are (1) hydrophones that measure acoustic emissions in pipes (Kadri et al., 2011; Cody et al., 2018; Li et al., 2018); (2) accelerometers to capture pipe vibration changes when a leak occurs (El-Zahab et al., 2018; Butterfield et al., 2017; Yadekhasti et al., 2016). Pressure sensors have also been deployed in leak detection systems (Karry et al., 2016; Sadeghioon et al., 2014). Butterfield et al. (2017b) employed two types of sensors: (1) accelerometer and (2) pressure sensor. The pressure sensors were used only for monitoring purposes and did not contribute to leak detection. In this research, the application of three types of sensors will assist with the investigation of leak signatures from three perspectives where one sensor could complement the others for detection with higher accuracy.

In this study, a testbed has been designed and assembled for background leak detection research. Not only does it address the previously mentioned deficits in other setups, but also includes a broader view of leak simulation experiments and influencing factors. The feasibility of applying dimensional analysis principles in the testbed design will be evaluated, followed by a description of the setup components, dimensions, and architectures. The characteristics of induced leaks will be explained. Information about data collection devices, ambient sound, and resonance, and experiments will be discussed in depth.

2.2. Dimensional Analysis

In this section, we evaluate if, based on dimensional analysis principles, the testbed is distorted or not. Since our objective is to design a setup through which we can assess network architecture's effects on leak detection algorithms, the testbed must comprise loops and junctions to create different architectures. To simulate a real water distribution system, we employ a similitude approach to model a section of the virtual Micropolis water network (Brumbelow et al., 2007). Figure 2.1 shows the Micropolis network and two extracted loops to use for a similitude.

Since the Reynolds number determines the water regime in water distribution systems, the Reynolds number is the governing parameter for a similitude in these networks. Therefore, as a similitude, the Reynolds number (Eq. 2.1) in the model m, i.e., the testbed, and prototype (p), i.e., Micropolis loops, should be equal (Eq. 2.2).

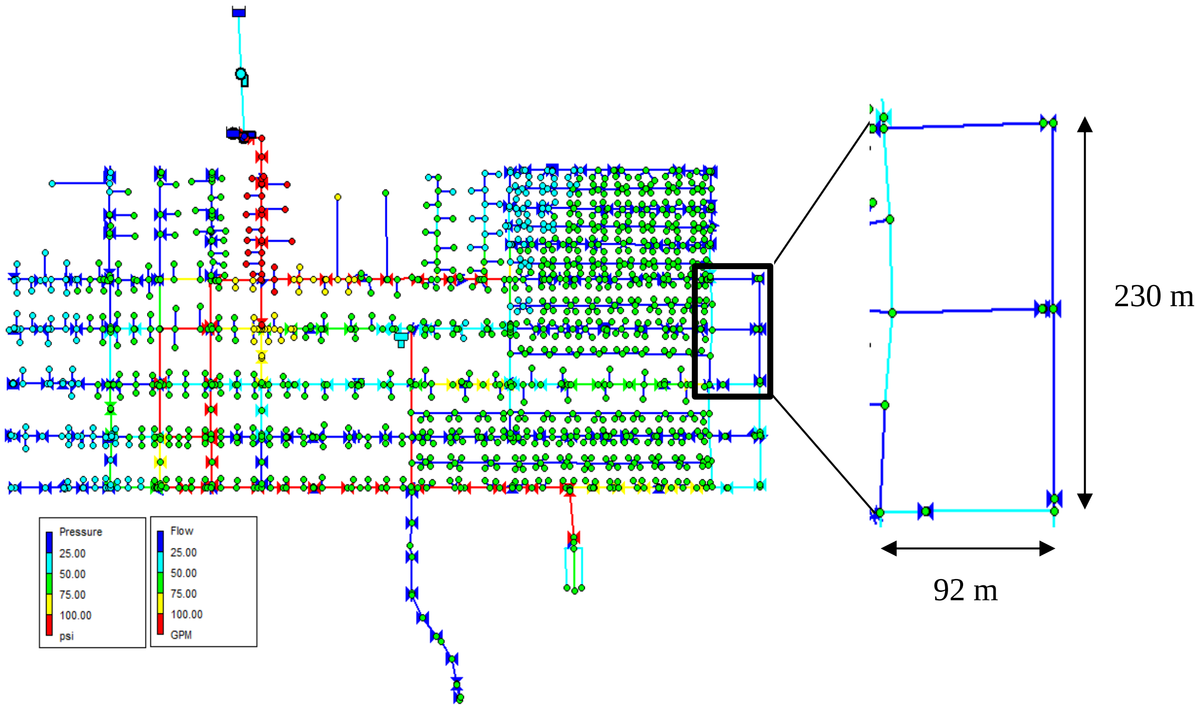

Figure 2.1. Micropolis virtual network and an extracted section for dimensional analysis

$$Re = \frac{\rho\, v\, D}{\mu} \tag{2.1}$$

$$\text{Reynolds}_m = \text{Reynolds}_p \tag{2.2}$$

Regarding the rules of similitude (Szücs 1980), since the fluid is the same in the prototype and the model, length, velocity, and time ratios of the similitude are

$$\text{Length ratio: } L_r = \frac{D_m}{D_p} \tag{2.3}$$

$$\text{Velocity ratio: } V_r = \frac{1}{L_r} \tag{2.4}$$

$$\text{Time ratio: } V_r = L_r^2 \tag{2.5}$$

The longest dimension of the testbed was 9.15 m. Regarding Eq. 2.3 and Eq. 2.4, modeling the loops in Figure 2.1 with a length of 230 m in the 9.15 m long area results in a velocity ratio of 25.14. Given the velocity ratio and using the maximum velocity of the Micropolis EPANET model, i.e. V_{max_p}, that is 0.54 (m/s), maximum velocity in the testbed, i.e. V_{max_m}, will be 13.57 (m/s). This velocity is 4.52 times larger than the maximum allowed velocity, i.e., 3 (m/s), in design guidelines (Bryan/College Station Unified Design Guidelines-Domestic Water, 2012) for water distribution systems. Therefore, building a dimensional analysis model was not possible, since the model would be distorted. Nonetheless, to make the testbed more realistic, we chose 15 cm diameter pipes, which are commonly used as the distribution pipes in water networks.

2.3. Testbed Material, Architectures, and Dimensions

The testbed is composed of two sections: (1) water supply with 2.54 cm diameter pipes (supply line); (2) water distribution with 15 cm diameter pipes (distribution section).

Pipes in the distribution section and the supply line are schedule 80 polyvinyl chloride (PVC). The distribution section is 7.3 m long and 4.9 m wide. Pipes are connected via tees, crosses, and elbows. Flanges are glued to the ends of pipes and fittings. Therefore, pipes can connect to fittings with the flanges and eight sets of bolts and nuts, and a gasket per flange connection. Figure 2.2 shows an untightened assembly of two flanges using a rubber gasket, a bolt, and a nut. Since the gasket is rubber, its flexibility helps to fill the gap between two flanges.

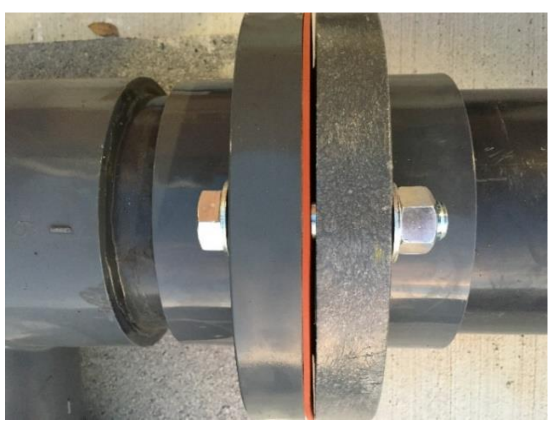

Figure 2.2. Connection of two flanges with an untightened bolt and nut

The testbed architectures are two fold: (1) looped; (2) branched. The following describes each architecture.

2.3.1. Looped Network

Figure 2.3 is a picture of the looped network. This architecture consists of seventeen pipe segments, two crosses, nine tees, four elbows, and two simulated hydrants.

Figure 2.3. An overview of the looped network

Figure 2.4 shows a schematic of the looped network illustrating its components.

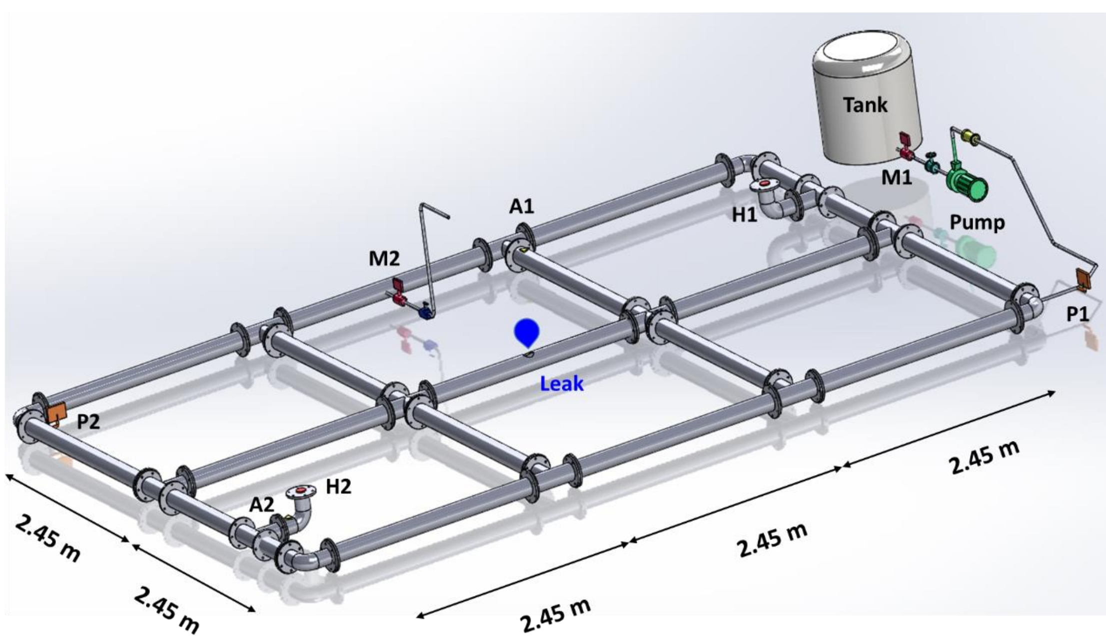

Figure 2.4. A schematic of the looped network

In Figure 2.4, H1 is hydrophone 1, H2 hydrophone 2, A1 accelerometer 1, A2 accelerometer 2, P1 pressure sensor 1, P2 pressure sensor 2, M1 meter 1 mearing total input water to the network, M2 meter 2 measuring output flow from the service valve simulating water consumption.

2.3.2. Branched Network

Figure 2.5 shows the branched network.

Figure 2.5. An overview of the branched network

Figure 2.6 shows a schematic of the branched network and its components.

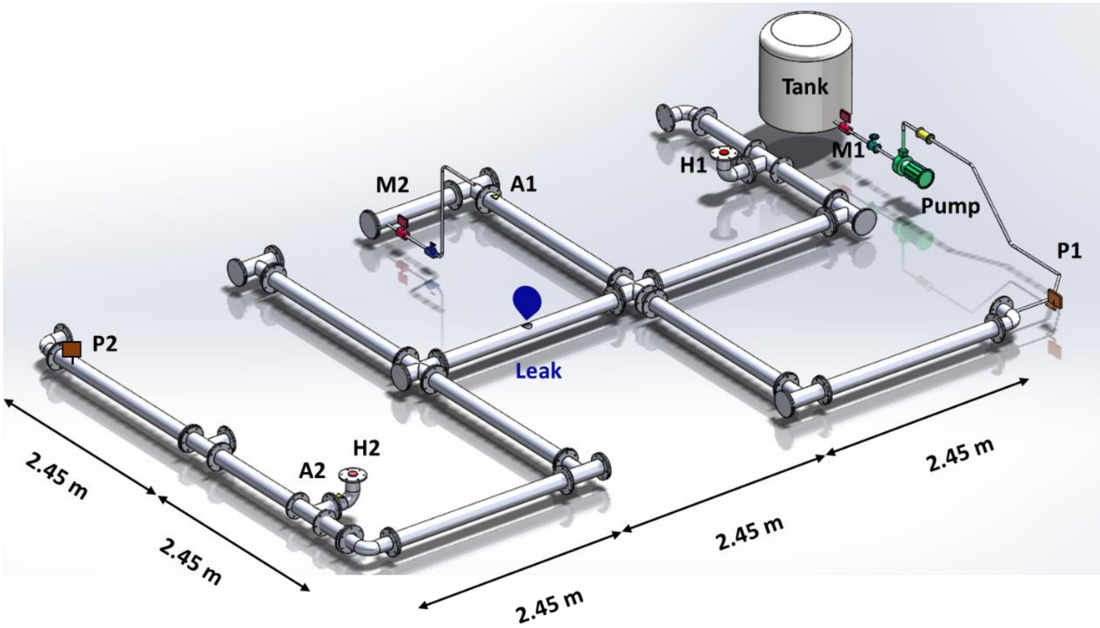

Figure 2.6. A schematic of the branched network

Regarding Figures 2.4 and 2.6, we disassembled six pipes to change the looped architecture to the branched network. This change removed loops in the looped network.

2.3.3. Water Supply Line

The supply line that includes a storage tank, a flow meter, a gate valve, a pump, and a check valve supplies water for the distribution section. Figure 2.7 shows the supply line and its components.

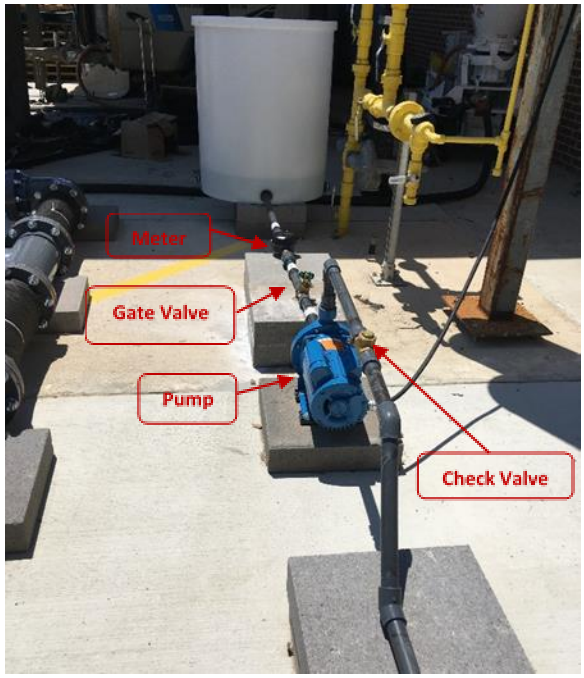

Figure 2.7. Supply line with components

We used the fewest fittings possible and employed 45-degrees elbows to decrease the minor head losses. In the following, we elaborate on the components of the supply line.

- Storage Tank

The storage tank is a plastic open-top-cylinder with a height of 92 cm and a diameter of 80 cm, filled with a water hose.

- Flow Meter

To measure input water to the distribution section, we used the 2.54 cm Neptune MACH 10 ultrasonic meter. The meter has 0.038 liters_ 0.01 U.S. gallons_ resolution, a 9-digit-reading display, and a flow direction indicator. In addition to continuously displaying instantaneous water flow, the meter displays cumulative flow every ten seconds.

- Gate Valve

The Matco brass gate valve allowed us to adjust the water flow into the distribution section and control input water volume.

- Pump

We used the Goulds 1MC1G1A0 centrifugal pump to provide enough pressure in the distribution section. It is a fixed speed pump whose information is available in Table 2.1.

Table 2.1. Pump characteristics

Suction Size	Discharge Size	Impeller Diameter	Driver	Material
3.175 cm	2.54 cm	15.56 cm	60 Hz, 2 pole, 3500 RPM, 3 HP	Cast iron

Figure 2.10 represents the pump characteristics curve. We will later calculate demand flows, which are 3 GPM and 7.5 GPM.

- Check Valve

We used a Matco brass check valve to prevent backflow from the distribution section to the pump.

2.3.4. Support Blocks

Due to the water pressure in the pipes and resulting momentums at junctions, the testbed must be constrained. We put concrete blocks under each fitting and in contact with flanges where the blocks' weight prevented the fitting from moving. With all fittings constrained, we could stablize the entire testbed. Each block was 10.16 cm × 40.64 cm × 40.64 cm with a weight of 22.68 (kg). These blocks act like the thrust blocks in real water distribution systems (Bryan/College Station Unified Design Guidelines- Domestic Water, 2012). Figures 2.3 and 2.5 show the blocks.

2.3.5. Backfill Medium

Multiple studies evaluated the influence of surrounding material on leak frequency and attenuation rates. Van Zyl (2013) reported that leak flow could fluidize surrounding media, and mobilized backfill particles can create sounds as they hit pipe walls. Though our research aims not to investigate the effect of backfill media on leak signals, to simulate a real case study, pipes were covered with two layers _5 cm_ of the Mutual NW100 Non-woven geotextile to create the damping effect of backfill material. Fox et al. (2016) and Butterfield et al. (2017a) used this method to represent a fully constrained porous media, while the former authors described that this type of geotextile fabric is a good representation of an unfluidized surrounding media. We assumed leaks mobilize soil particles only in their vicinity. Regarding the dampening effects of the sounding material, this assumption makes detecting leak signals more difficult. Therefore, we wrapped the lengths of all pipes with the fabric except for a range of 15.24 cm from leak openings. Figures 2.8(a) and 2.8(b) show how pipes without leak and with leak have been covered with the fabric.

2.3.6. Service Line and Consumption Flow

Consumption flow that is a significant source of noise in acoustic signals should be large enough so that its flow and sound dominate those of leak in a pipeline. Otherwise, leak detection would be very straightforward, and it would not need complicated algorithms. In similar research by Cody et al. (2018), though they did not mention leak and consumption flow rates, the authors used a service valve whose diameter was about eight times larger than their leak valve. They also remarked that concurrently opening a neighboring service valve with induced leaks introduces noises and makes data-driven leak detection more challenging.

Figure 2.8. (a) A pipe without leak fully wrapped with the geotextile fabric

Figure 2.8. (b) A pipe with leak and partially wrapped with the geotextile fabric

We calculated a base demand as if the distribution network supplies 100 people. Based on water network guidelines (Bryan/College Station Unified Design Guidelines- Domestic Water, 2012), design residential base demand should be 100 gallons per day per capita. Therefore, the testbed should supply 7 GPM demand for 100 people. Wu et al. (2010) reported that subsystems' step-testing to detect leaks is generally conducted between 1:00 a.m. and 5:00 a.m. to avoid customers' supply interruptions. Since customer consumption and ambient sound are low during this time, we used minimum and maximum multipliers, between 1:00 a.m. and 6:00 a.m., of pattern 2 of the Micropolise water network developed by Brumbelow et al. (2007). Using these flow multipliers and the 7 GPM base demand, the actual flow demands are shown in Table 2.2.

Table 2.2. Base demand and actual demands used as the experimental consumption flows

Time	Multiplier	Base Demand (GPM)	Actual Demand (GPM)
1:00 a.m.	0.43	7	3
6:00 a.m.	1.07	7	7.5

A 2.54 cm diameter pipe was connected to the distribution section with a saddle clamp to simulate the consumption flows. Figure 2.9 shows the service line, which includes a Neptune MACH 10 ultrasonic meter to measure consumption flow, and a globe valve to adjust the service pipe's output flow. Since the testbed's outflows are from leaks and the service line, the difference between the flow meter at the supply line and the meter at the service line determines the leak flow.

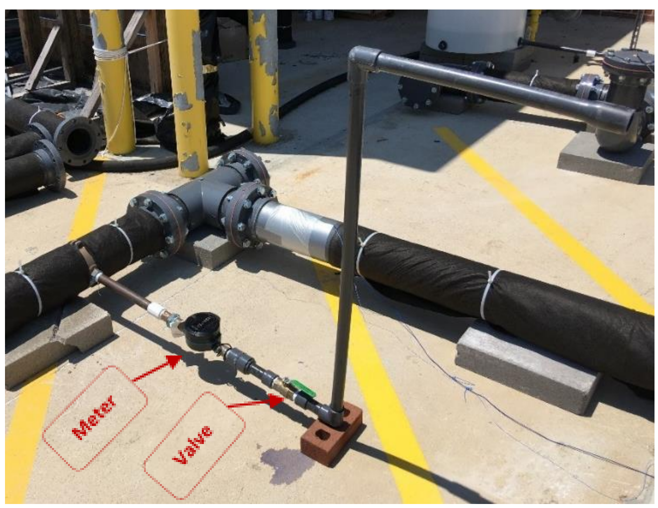

Figure 2.9. Service line with the meter and the globe valve

2.3.7. Pump and System Curves

Since the Goulds 1MC1G1A0 pump was available in our research laboratory and EPANET simulations showed it suffices for our experimental requirements, this pump was used. Curve *A* in Figure 2.10 shows the characteristics curve of the pump employed for the system design.

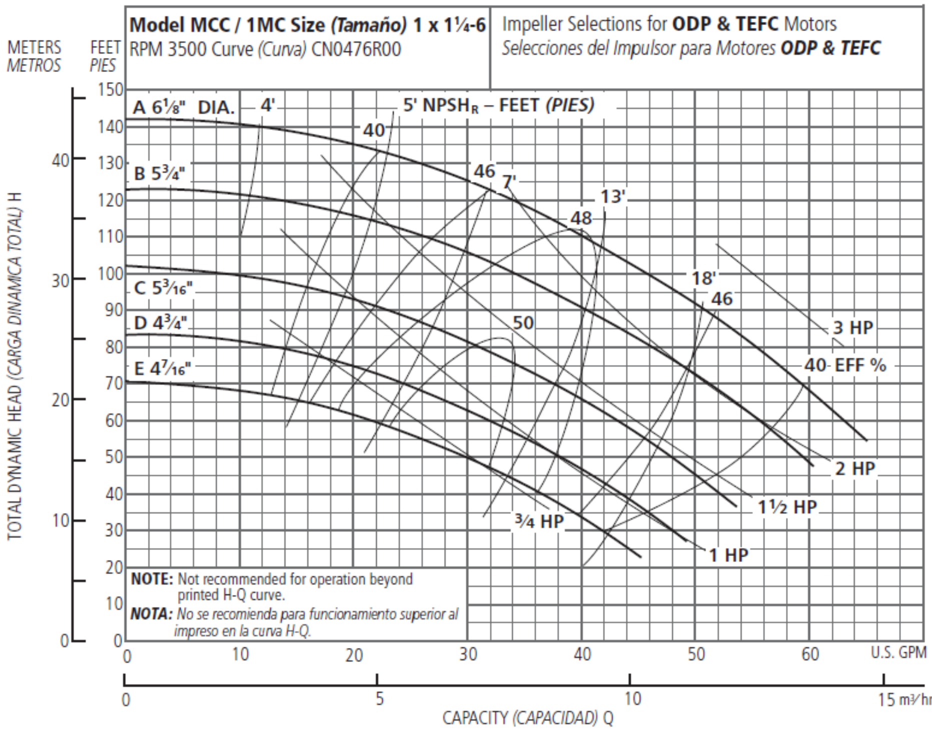

Figure 2.10. Goulds 1MC1G1A0 pump curve (Technical Brochure of MCC End Suction Centrifugal Pumps, 2018).

There are two head losses that formulate a system curve: (1) headloss caused by friction; (2) local (or minor) loss.

Eq. 2.6 gives friction headloss in feet based on the Hazen-Williams method (Walski, 2006).

$$h_f = 4.52\, C^{-1.852}\, d^{-4.8704}\, L\, q^{1.852} \quad (2.6)$$

where C = Hazen-Williams roughness coefficient, d = pipe diameter (in), L = pipe length (ft), and q = flow rate (GPM). Also, local loss can be calculated as

$$h_L = K \frac{V^2}{2g} \quad (2.7)$$

where K = minor loss coefficient, V = flow velocity in pipe (ft/s), and g = gravitational acceleration constant (ft/s^2). Regarding Figure 2.3 and Figure 2.5, the testbed's elevation difference, i.e., intercept in the system curve, is zero. Table 2.3 shows the minor loss coefficients used to determine the system curve (Rossman, 2000). We calculate the local loss for fittings located after the pump.

Table 2.3. Minor loss coefficients

Fitting	Loss Coefficient
Swing check valve, fully open	2.5
Short-radius elbow	0.9
45-degree elbow	0.4
Standard Tee - flow through branch	1.8
Standard Tee - flow through run	0.6

In the following, we calculate head losses in the supply line and distribution sections, respectively.

2.3.7.1. Supply line

2.3.7.1.1. Friction headloss

Regarding the roghness coefficient for plastic pipe, i.e., C = 150 (Rossman, 2000), d = 1 in = 2.54 cm, L = 12.21 ft = 3.72 m, the Hazen-Williams friction loss formula in the supply line (h_{fs}) is

$$h_{fs} = 4.52\, C^{-1.852}\, d^{-4.8704}\, \mathrm{L}\, q^{1.852} = 0.005\, q^{1.852} \quad (2.8)$$

where q is in GPM.

2.3.7.1.2. Minor headloss

The supply line includes a short radius elbow, a check valve, and four 45-degree elbows. Therefore, the minor loss coefficients and the local loss in the supply line (h_{Ls}) is

$$K = 0.9 + 2.5 + (4 \times 0.4) = 5 \quad (2.9)$$

$$h_{Ls} = K\,\frac{V^2}{2g} = 5\,\frac{V^2}{2g} = 5\,\frac{V^2}{64.4} = 0.078\,V^2 = 0.078\,\frac{q}{A^2}^2 = 2601.598\,q^2 \quad (2.10)$$

where the unit of q is (CFS). Eq. 2.11 shows the minor loss for q in GPM.

$$h_{Ls} = 2601.598\,q^2 = 0.013\,q^2 \quad (2.11)$$

2.3.7.2. Distribution section

2.3.7.2.1. Friction headloss

The distribution section is composed of seventeen pipes. To accurately compute friction and minor losses in the distribution section, we need to know each pipe's exact flow value and fitting. Since it is laborious to calculate the flows manually, we assumed the distribution section is a pipe with the length of (17 × 2.44) = 19.52 m, and all fittings are along this long pipe. Not only does this assumption simplify manual calculations, but it also over-estimates losses and results in the worst-case scenario. Therefore, the Hazen-Williams parameters and formula for the distribution section (h_{fd}) is

C = 150, d = 6 in = 15.24 cm, L = 136 ft = 41.45 m

$$h_{fd} = 4.52\, C^{-1.852}\, d^{-4.8704}\, L\, q^{1.852} = 0.0000093\, q^{1.852} \tag{2.12}$$

where the unit of q is GPM.

2.3.7.2.2. Minor headloss

Fittings in the distribution section are six tees, four elbows, and two crosses. For the crosses, we also use tee's coefficient. Regarding the assumption mentioned above and Table 2.3, the minor headloss formula in the distribution section (h_{Ld}) is

$$K = (6 \times 1.8) + (4 \times 0.4) + (2 \times 1.8) = 19.6 \tag{2.13}$$

$$h_{Ld} = K\, \frac{V^2}{2g} = 19.6\, \frac{V^2}{2g} = 19.6\, \frac{V^2}{64.4} = 0.304\, V^2 = 7.894\, q^2 \tag{2.14}$$

where the unit of q is (CFS). Eq. 2.15 shows the minor loss with q in GPM.

$$h_{Ld} = 0.000038\, q^2 \tag{2.15}$$

Table 2.4 summarizes the headloss formulas.

Table 2.4. Summary of headloss formulas

	Supply Line	Distribution Section
Friction Loss	$h_{fs} = 0.005\ q^{1.852}$	$h_{fd} = 0.0000093\ q^{1.852}$
Minor Loss	$h_{Ls} = 0.013\ q^2$	$h_{Ld} = 0.000038\ q^2$

The total headloss of the testbed (h_T) is the sum of all losses in Table 2.4 whose formula is:

$$h_T = 0.0050093\ q^{1.852} + 0.013038\ q^2 \tag{2.16}$$

where q is flow (GPM). We extracted the pump curve *A* in Figure 2.10 and used it to create Figure 2.11, the pump curve vs. the system curve.

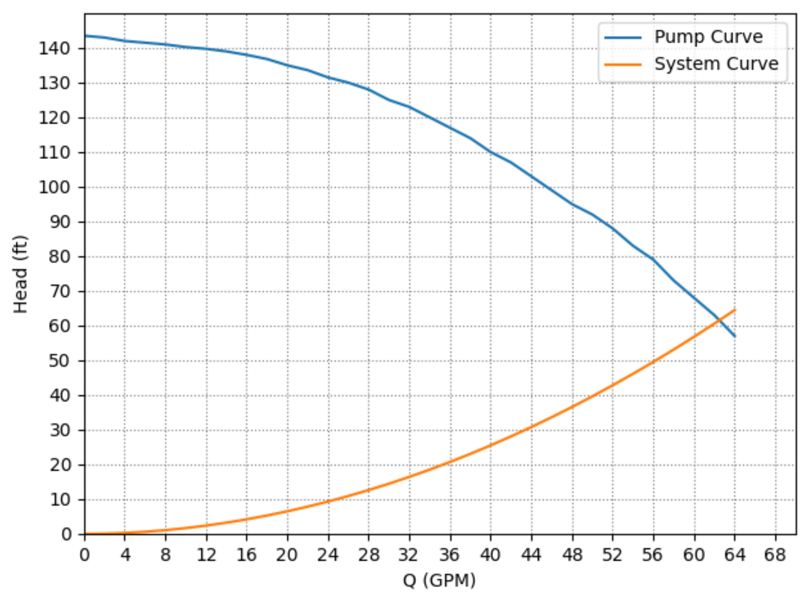

Figure 2.11. Pump and system characteristics curves of the testbed

In the next sections, we will discuss total flows in the testbed with different leaks. The highest flow will be 8.12 GPM in the looped network with an orifice leak when the demand is 7.5 GPM. By plugging $q = 8.12$ GPM in Eq. 2.16 or using Figure 2.11, the testbed's maximum headloss would be 0.34 m _1.102 ft_ or 0.48 psi. Based on Figures 2.10 and 2.11, the pump head for *Q* = 8.12 GPM is about 42.9 m_141 ft. Therefore, the actual system head for the maximum headloss would be (42.9-0.34) = 42.57 m _139.6 ft_ or 60.13 psi. This pressure shows the headloss in the testbed is negligible.

2.4. Leak Type, Size, and Flow Rate

2.4.1. Leak Type

Among all research that has used testbeds to detect leaks or evaluate factors influencing leak, none of them has considered leak types' effects on leak signature. As one of the most recent studies, Harmouche and Narasimhan (2020) induced leaks in their laboratory testbed using a 3.2 mm valve. This leak simulation does not reflect the varieties of leak types reported in the literature. Greyvenstein and van Zyl (2007) described three leak types among failed pipe segments taken from Johannesburg's water distribution system in South Africa. The leaks were orifice, longitudinal crack, and circumferential crack. Cassa et al. (2010) used these three leak types to investigate the effects of pressure on holes and cracks with numerical methods. Joint leak has also been described as another source of water loss. Covelli et al. (2015) and Stathis (1998) referenced displaced or polymerized gaskets to cause leaks at joints. This study conducted experiments with four types of leaks: (1) orifice (hole); (2) longitudinal crack; (3) circumferential crack; (4) leak at a joint gasket (gasket leak). These leaks were induced in the middle pipe between the two crosses in Figures 2.4 and 2.6.

We induced the orifice and cracks by drilling and milling the middle pipe wall, respectively. The gasket leak was also generated by loosening a flange's bolts located in the middle of the leaking pipe. It is worth noting that the leaks were induced in four different pipes, and experiments were conducted one at a time for each leak pipe. Figures 2.12-2.15 show the hole leak, longitudinal crack, circumferential crack, and gasket leak, respectively, where the leaks have been discerned with red circles.

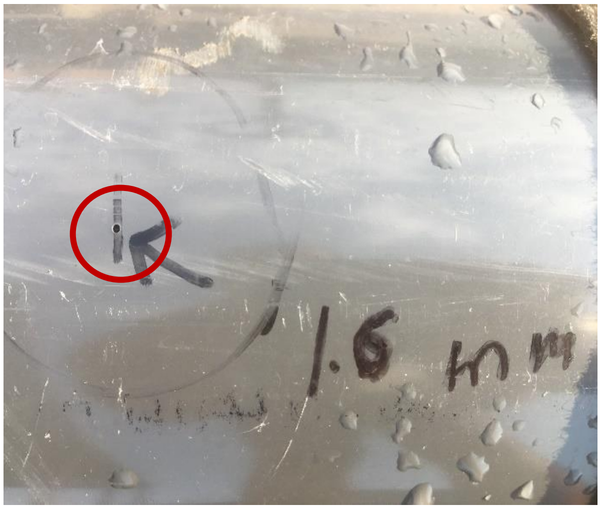

Figure 2.12. Orifice leak

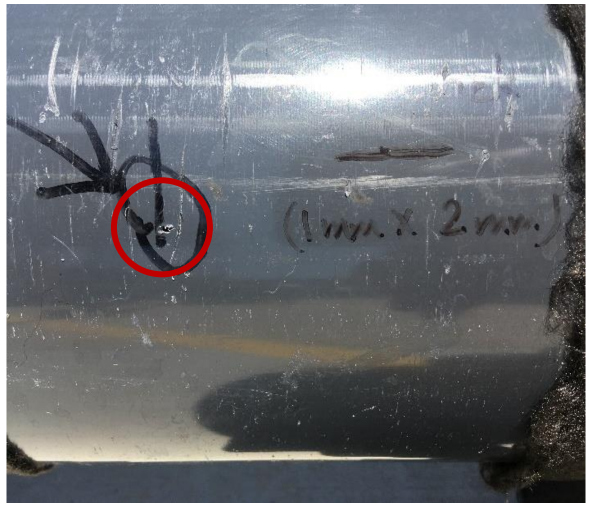

Figure 2.13. Longitudinal crack

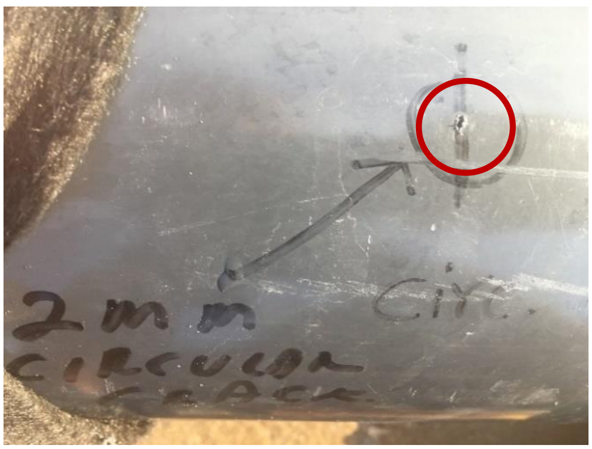

Figure 2.14. Circumferential crack

Figure 2.15. Gasket leak

2.4.2. Leak Size

Studies have referenced different leak rates in water distribution systems across the world. The leakage rate, which is a percentage of total input water to a network, typically varies from 10% to 30% in North America (El-Zahab et al., 2018; Butterfield et al., 2017b) and increases up to 60% in developing countries (Lee and Schwab, 2005) or the City of Flint in Michigan (npr, 2016).

In this research, we aim to detect background leaks that cannot be distinguished easily. Therefore, we attempted to determine the leak size such that the leak rate would be less than 30% of the input water to the testbed.

We used the Torricelli equation to determine leak sizes. Cassa et al. (2010) described the Torricelli equation, which defines the relationship between the leak flow rate Q and pressure head h, as

$$Q = C_d\ A\ \sqrt{2g}\ h^{0.5} \tag{2.17}$$

where C_d is the discharge coefficient, A the orifice area, and g the acceleration of gravity.

Greyvenstein and van Zyl (2007) represented a more general form of the Torricelli equation

$$Q = ah^{\beta} \tag{2.18}$$

where a is the leakage coefficient and β the leakage exponent.

Comparing Eq. 2.17 and Eq. 2.18, the leak coefficient a would be

$$a = C_d\ A\ \sqrt{2g} \tag{2.19}$$

Based on field studies, Farley & Trow (2003) and Farley (2007) explained that β could be greater than its value for an orifice, i.e., 0.5, and varies from 0.5 to 2.79. Greyvenstein and van Zyl (2007) conducted experiments on u-PVC, asbestos cement, and mild steel failed water pipes to investigate the failures' pressure - leakage relations. They found that the leak exponent β in u-PVC pipe changes based on leak shapes, as shown in Table 2.5.

Table 2.5. Values of the leak exponent, β, in failed u-PVC pipes (Greyvenstein and van Zyl, 2007)

Failure Type	Leak exponent for u-PVC pipe
Round hole	0.524
Longitudinal crack	1.38-1.85
Circumferential crack	0.41-0.53

As was previously discussed, the leaks should be small enough that their flow rate is less than 30% of the total input water. Moreover, the leak sizes should be large enough to be generated by machinery, such as drills and mills, available at Texas A&M University. Since one of our research goals is investigating the effect of leak shape on the leak signal, we induced leaks with different shapes but with the same areas. The smallest crack possible with our equipment for a 15.24 cm u-PVC pipe wall was a 1 mm × 2 mm crack that led to a 2 mm^2 leak area. Therefore, the longitudinal and circumferential cracks were 2 mm long and 1 mm wide. A hole leak with a 2 mm^2 area would result in a hole with a diameter of about 1.6 mm.

Figure 2.16 shows a methodology to determine leak sizes and discharge coefficients so that the leak rates are below 30%.

2.4.3. Leak Flow Rate

We used leakage exponents 0.5, 1.5, and 0.47 for the hole leak, longitudinal crack, and circumferential crack, respectively. With reference to Figure 2.16, though leak size (A) could be changed in each iteration, we maintained 2 (mm^2) leak area to make leaks as small as possible when leak detection is a more complicated task. This would allow us to change only the discharge coefficient (C_d) to reach the desired leak rate.

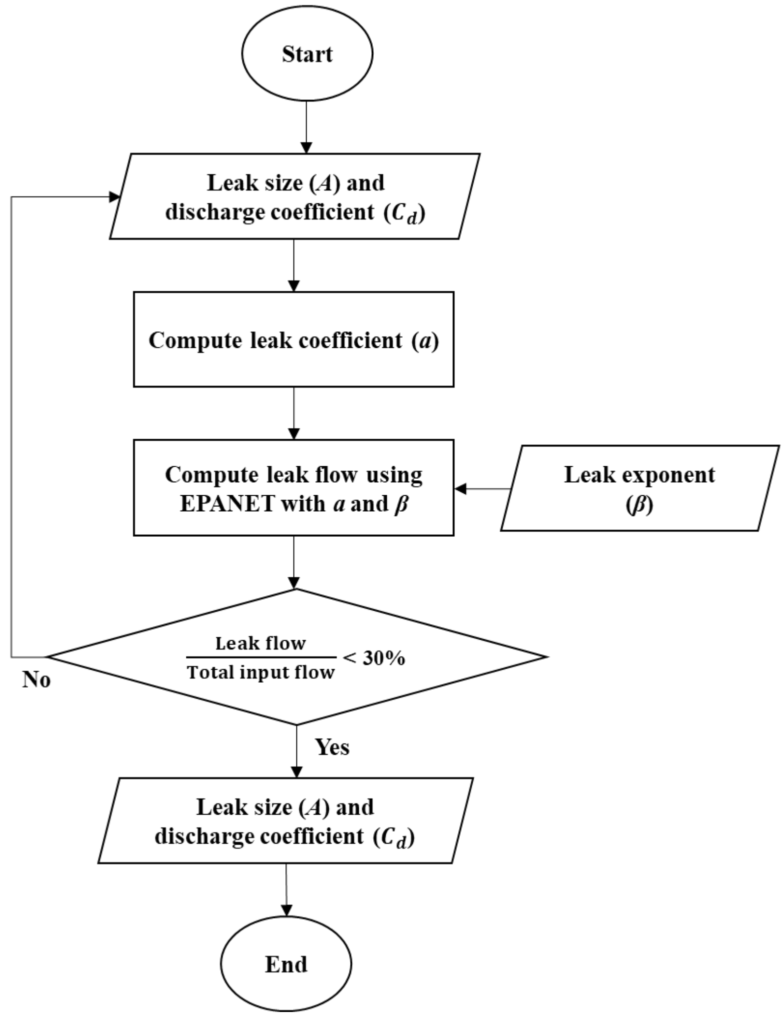

Figure 2.16. A flowchart to control with assumed leak size (A) and discharge coefficient (C_d) resulting in a leak flow rate below 30%

The section with the dashed line in Figure 2.17 shows how we calculated the leak flow rate for the hole leak in the looped network when the consumption flow rate is 3 GPM. We employed this method to calculate the leak flow rate for other leak types and service line flows in the looped and branched networks. It is worth noting that flow rates of meters 1 and 2 were determined by taking the average flow rates of both meters between t_1 and t_2 where $t_2 - t_1 = 1$ (s).

Figure 2.17 indicates how we calculated leak flow rates.

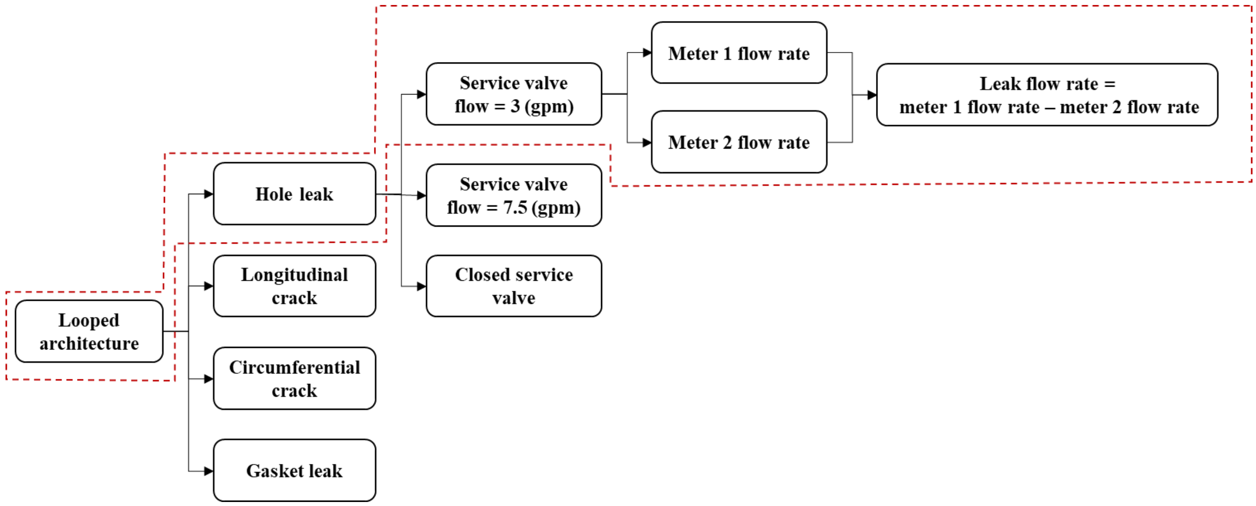

Figure 2.17. A method to calculate leak flows

Table 2.6 shows the measured leak flows for the looped and branched architectures with different leak types and demands. Values of the Meter 2 Flow in Table 2.6 are the actual demand stimulated by the 2.54 cm service line.

The key observations of Table 2.6 are:

- With an increase in demand, the leak flow rate decreases. Larger demand values cause a pressure drop in the network. Regarding Eq. 2.17, a pressure decrease leads to a smaller leak flow.
- For the hole leak, longitudinal crack, and circumferential crack, leak rates in the branched network are slightly smaller than or equivalent to those of the looped network. The difference can be due to lower network pressure caused by fewer pipes, dead ends, and less flow connectivity.

Table 2.6. Leak flows at the looped and branched networks with different leak types and demands

Architecture	Leak Type	Design Demand (GPM)	Meter 1 Flow (GPM)	Meter 2 Flow (GPM)	Leak Flow (GPM)	Rounded Leak Rate (%)
Looped	Hole Leak	0.00	0.77	0.00	0.77	-
		3.00	3.76	3.05	0.71	19
		7.50	8.12	7.51	0.61	8
	Longitudinal Crack	0.00	0.51	0.00	0.51	-
		3.00	3.51	3.10	0.41	12
		7.50	7.85	7.49	0.36	5
	Circumferential Crack	0.00	0.40	0.00	0.40	-
		3.00	3.39	3.00	0.39	12
		7.50	7.90	7.53	0.37	5
	Gasket Leak	0.00	0.84	0.00	0.84	-
		3.00	3.81	3.02	0.79	21
		7.50	8.30	7.54	0.76	9
Branched	Hole Leak	0.00	0.67	0.00	0.67	-
		3.00	3.79	3.16	0.63	17
		7.50	8.08	7.52	0.56	7
	Longitudinal Crack	0.00	0.51	0.00	0.51	-
		3.00	3.51	3.10	0.41	12
		7.50	7.85	7.55	0.30	4
	Circumferential Crack	0.00	0.40	0.00	0.40	-
		3.00	3.42	3.05	0.37	11
		7.50	7.90	7.55	0.35	4
	Gasket Leak	0.00	0.84	0.00	1.41	-
		3.00	4.20	3.08	1.12	27
		7.50	8.57	7.52	1.05	12

- Leak rates for the gasket leaks are larger than other leak types because we induced the gasket leak where we loosened the flange bolts. From a practical point of view, this method was not capable of being precisely controlled. However, these rates were below the designed 30% leak flow rate.
- Hole leak rates are greater than the leak rates in the longitudinal and circumferential cracks. This finding will be evaluated in the next section.
- The meter 2 flow rate represents demand flows. The zero-flow rate means the service valve is closed. Since we adjusted the service globe valve manually, the actual service flow rate,

i.e., meter 2 flow rate, may not precisely match designed demands. This difference introduces an inevitable error in an experimental setup, though its value is negligible.

Figures 2.18-2.21 show leaks from the hole, longitudinal crack, circumferential crack, and gasket leak, respectively.

Figure 2.18. Water jet from the orifice leak

Figure 2.19. Water jet from the longitudinal crack

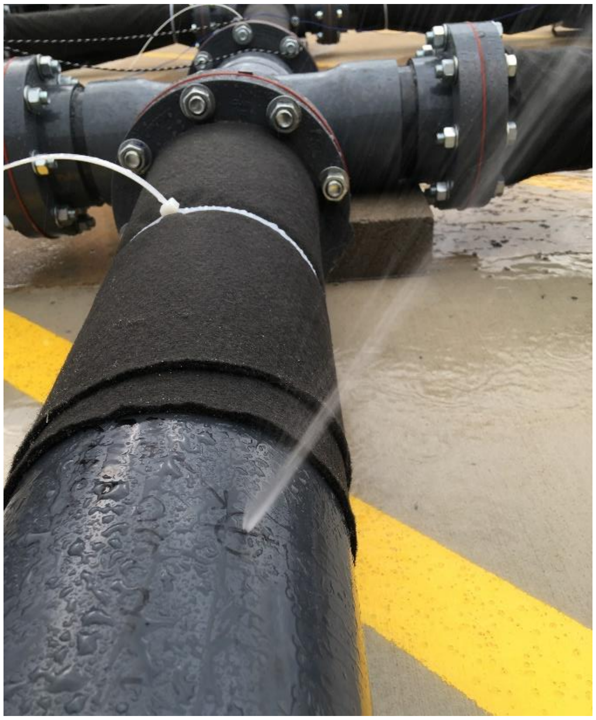

Figure 2.20. Water jet from the circumferential crack

Figure 2.21. Outflow from the gasket leak

Dissimilar leak shapes cause different shapes of water jet at the leak locations. For instance, the cross-section area of the hole water jet maintains a circular shape. In contrast, the cross-section area of the longitudinal and circumferential water jets have a rectangular shape close to the leaking crack. Then these water jets diverge from the centroid of the leak area. Figure 2.19 and Figure 2.20 show that the water jets in the longitudinal crack and the circumferential crack are parallel and perpendicular to the pipe axis, respectively.

2.4.3.1. Actual Discharge Coefficient (C_d)

Regarding the actual leak flows in Table 2.6 and the 2 (mm^2) area for the hole leak, longitudinal crack, and circumferential cracks, we can calculate the discharge coefficient in Eq. 2.17. Figure 2.22 shows a flowchart through which we could compute the discharge coefficient (C_d).

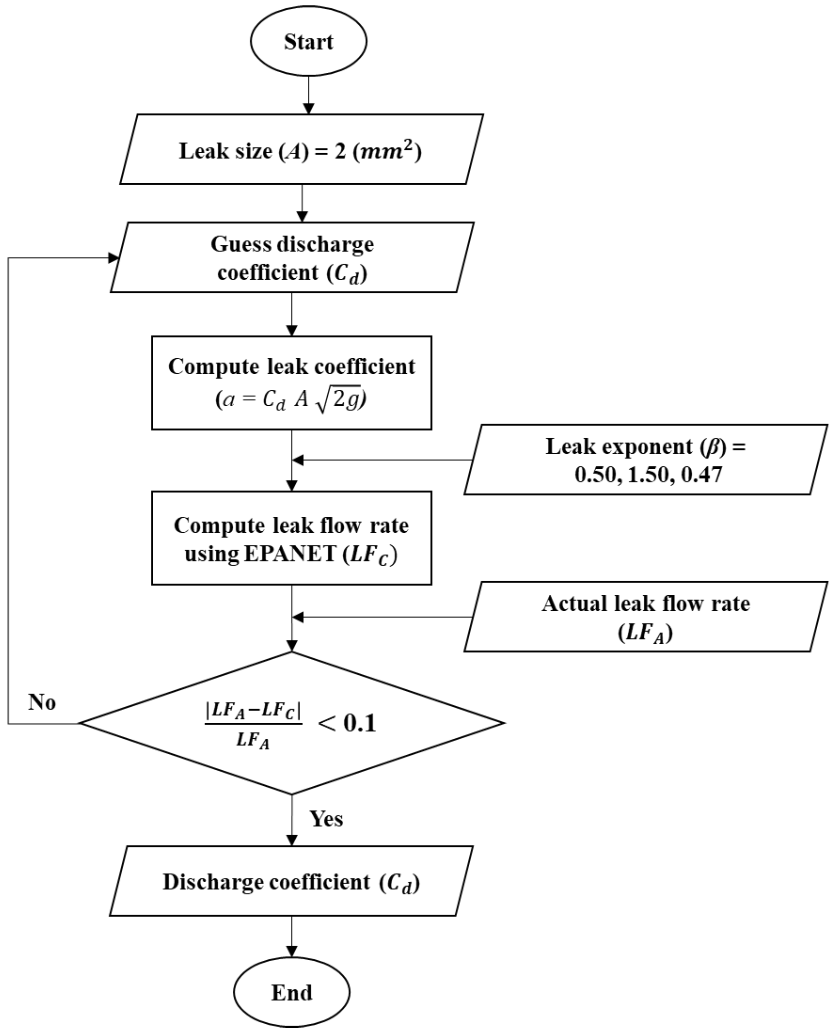

Figure 2.22. A methodology to calculate actual discharge coefficient (C_d) for each leak in both networks

It is worth noting that the actual leak flow rates are the values in Table 2.6. Table 2.7 includes the best discharge coefficients (C_d) and leak coefficients (a) based on Figure 2.22. For the gasket leak, the term leak area is not applicable. Since it is not possible to determine the leak area in the gasket leak, and the Torricelli equation is not valid, Figure 2.22 was not used to calculate the discharge coefficient. Therefore, we will employ the actual leak flow rates in Table 2.6 in future analyses of all leaks.

Table 2.7. Best discharge coefficient (C_d), leak coefficient (a), and the difference between leak flow computed with EPANET (LF_C) and the actual flow **(LF_A)**

Architecture	Leak Type	Design Demand (GPM)	Best Discharge Coefficient (C_d)	Calculated Leak Coefficient (a) ($\frac{GPM}{\sqrt{psi}}$)	EPANET Leak Flow (LF_C) (GPM)	Actual Leak Flow **(LF_A)** (GPM)	$\frac{\lvert LF_A - LF_C \rvert}{LF_A} \times 100$
Looped	Hole Leak	3	0.80	0.083	0.65	0.71	8%
		7.5			0.64	0.61	5%
	Longitudinal Crack	3	0.007	0.00082	0.39	0.41	5%
		7.5			0.38	0.36	6%
	Circumferential Crack	3	0.47	0.055	0.38	0.39	3%
		7.5			0.38	0.37	3%
Branched	Hole Leak	3	0.75	0.078	0.61	0.63	3%
		7.5			0.60	0.56	7%
	Longitudinal Crack	3	0.007	0.00082	0.39	0.41	5%
		7.5			0.38	0.30	3%
	Circumferential Crack	3	0.47	0.055	0.38	0.37	3%
		7.5			0.38	0.35	9%

Based on Table 2.7, since the difference between EPANET and actual leak flows are less than 10 percent of the actual leak flows, the chosen discharge coefficients can represent the leak characteristics. We will evaluate if the induced leaks and their flows and coefficients conform to other studies.

Concerning Eq. 2.18, to increase the leak flow, we can increase the leak coefficient (a), pressure in the pipe (h), or the leak exponent (β). Since the testbed pump is a fixed drive, we cannot increase the pressure to supply the pre-determined demands. Based on the leak types' nature, the leak exponent (β) cannot be changed. Therefore, the leak coefficient (a) is the only parameter we can alter to increase the leak flow rate.

According to Eq. 2.19, the leak coefficient (a) can be increased either by enlarging the leak area (A) or adding to the discharge coefficient (C_d). We can increase both of these parameters as long as the flow rate is less than 30% of the testbed's total input water.

Other research has used larger leak sizes. For instance, Greyvenstein and van Zyl (2007), to assess the pressure-flow relation in failed pipes, induced the leak sizes listed in Table 2.8 in a 110 mm diameter u-PVC pipe.

Table 2.8. Leak dimensions and shapes induced in (Greyvenstein and van Zyl, 2007)

Hole	Longitudinal Cracks				Circumferential Cracks			
Diameter (mm)	Thickness (mm)	Length (mm)			Thickness (mm)	Length (mm)		
12	1	50	100	150	1	90	170	270

Moreover, Franchini and Lanza (2013) have conducted experiments to generalize the Torricelli equation to calculate leak flow in pipes with different materials, diameters, and leak shapes and dimensions. They induced the leak sizes of Table 2.9 in u-PVC pipes.

Table 2.9. Leak dimensions and shapes induced in (Franchini and Lanza, 2013)

Holes			Longitudinal Cracks				
Diameter (mm)			Thickness (mm)	Length (mm)			
4	8	12	3	40	60	80	100

Table 2.10 includes the size and type of simulated leaks in the literature using laboratory-scale setups.

Table 2.10. Leak size in the literature using laboratory-scale testbeds

Author	Method to Simulate Leak	Leak Size (mm)	Setup Pipe Diameter (mm)
Harmouche and Narasimhan (2020)	Ball valve	3.2	152.4
El-Zahab et al. (2018)		25.4	50.8
Yazdekhasti et al. (2016)		25.4	76
Ismail et al. (2015)		1 and 3	25.4

None of the studies listed in Table 2.10 elaborated on the ratio of leak flow to the consumption flow when there is water consumption. Therefore, we cannot justify if their leak sizes are appropriate for such setups and demand flows. Moreover, all of the research employed a ball valve to simulate a leak that does not represent actual leak's characteristics.

We investigated whether we could double the leak dimensions that would result in an 8 (mm^2) cross-section area. Table 2.11 shows the leak rates for the hole leak, longitudinal crack, and circumferential crack with 8 (mm^2) area assuming the leak characteristics, i.e., discharge coefficients (Cd) and leakage exponent (β), remain the same. Regarding Table 2.7, since the leak coefficients (α) in the looped and the branched networks are very similar, we only assess the new leak sizes in the looped network.

Table 2.11. Leak flows and rates for 8 (mm^2) leaks

Leak Type	Dimension	Discharge Coefficient (C_d)	Leak Coefficient (a) ($\frac{GPM}{\sqrt{psi}}$)	Design Demand (GPM)	Total Input Water (GPM)	EPANET Leak Flow (GPM)	Rounded Leak Rate
Hole Leak	3.2 mm	0.80	0.378	3	5.95	2.94	50 %
				7.5	10.40	2.91	28 %
Longitudinal Crack	2 mm × 4 mm	0.007	0.0033	3	4.57	1.56	52 %
				7.5	9.01	1.52	21 %
Circumferential Crack	2 mm × 4 mm	0.47	0.221	3	4.53	1.53	34 %
				7.5	9.00	1.51	17 %

Leak rates in Table 2.11 indicate with doubling the leak dimensions, which leads to 8 (mm^2) leak areas, leak flow rates increase to above the 30% goal for 3 GPM demands.

Another parameter in the leakage coefficient (a) affecting the leak flow rate is the discharge coefficient (C_d) that is the product of the coefficient of velocity (C_v) and the coefficient of contraction (C_c) (Franchini and Lanza, 2013). The coefficient of velocity (C_v) depends on the friction between water and leak wall, and with increasing the lengths of contact of water with the wall of a leak, the coefficient becomes larger. The coefficient of contraction (C_c) depends on the leak geometry and decreases when the edges of a leak have a sharp angle. The coefficient of contraction (C_c) tends to value one when the leak has a bell shape.

In this section, we investigate leakage coefficients in the literature. Walski et al. (2004) evaluated factors to reduce leakage in a realistic water system with 200 nodes, 375 pipes, and 10.4 million gallons per day water consumption. They used two methods to model leaks using Eq. 2.18 where β = 0.5: (1) a leakage coefficient of 266 (GPM /$\sqrt{psi}$) was located in the system as a central leaking node; (2) leakage coefficients of 0.98 (GPM /$\sqrt{psi}$) were placed on each node throughout the system. Though the authors did not explain the shapes of the modeled leaks, using their leakage coefficients of 0.98 (GPM /$\sqrt{psi}$) will lead to leak rates larger than our 30% target. Wu et al. (2010) leveraged an optimization technique to locate and characterize leaks in two water distribution systems using field flow-pressure data. Table 2.12 lists the optimized leak coefficients for a water district serving more than 15 km^2 and about 3,000 properties with a hydraulic model comprised of 1,122 pipes and 841 nodes. The authors modeled leaks in the network as orifices with the leakage exponent 0.5.

Table 2.12. Optimized leak coefficients for a water network serving 3,000 households (Wu et al., 2010)

Leakage Coefficients (GPM /$\sqrt{psi}$)							
0.66	1.33	2.00	2.66	5.32	10.64	11.97	13.30

Table 2.13 shows the leak rate if we used the minimum leakage coefficient in Table 2.12, i.e., 0.66 (GPM /$\sqrt{psi}$), to model the hole leak in our testbed.

Table 2.13. Leak rates of the testbed hole leak using the minimum leakage coefficient applied in (Wu et al., 2010)

Leak Type	Leak Coefficient (a) ($\frac{GPM}{\sqrt{psi}}$)	Design Demand (GPM)	Total Input Water (GPM)	EPANET Leak Flow (GPM)	Rounded Leak Rate
Orifice Leak	0.66	3	48.57	45.56	94 %
		7.5	49.13	41.64	85 %

Since the leak rates in Table 2.13 are larger than the 30% goal, these field leakage coefficients do not apply to this our testbed.

In the following, we implement the generalized Torriccelli formula developed by Franchini and Lanza (2013) to verify the actual leak flow rates in Table 2.6 from a dimensional analysis perspective. The authors induced leaks in u-PVC and steel pipes with different sizes and applied dimensional analysis principles to examine if the Torricelli equation can represent leak flow when water pressure, pipe size, material, and leak shape change. Some of the key findings of this research are:

- In hole leaks and longitudinal linear cracks for given pipe size and material, when water pressure and leak size change, not only the initial leak area (A_0) changes, but the discharge coefficient (C_d) also varies. The authors introduced a correction factor ($\emptyset'_Q$) to adjust the product A_0 C_d in the Torricceli Equation.

- The correction factor ($Ø'_Q$) considers the effects of hydraulic and leak area elements. The former influences the coefficient of discharge (C_d), and the latter takes a count of local deformation of the initial leak area (A_0) with increasing pressure.
- Though the crack inclination is not considered in the correction factor, at the same pressure head in a u-PVC pipe, the correction factor ($Ø'_Q$) in a linear crack is larger than that of a hole leak when the leak areas are approximately equal.
- Eq. 2.20 is the generalized Torricelli formula to determine leak flow for pipes with different elastic materials, diameters, and leak dimensions and shape. The authors derived the right-hand side of Eq. 2.20 using dimensional analysis.

$$\frac{Q}{C_d\, A_0 \sqrt{2gh}} = Ø'_Q \left(\frac{h}{D}, \frac{R_{leak}}{D}, \frac{b}{a}, \frac{R_{leak}}{t}, \frac{R_{leak}\sqrt{gh}}{\upsilon}, \frac{V}{\sqrt{gh}}, \frac{\gamma h}{E}\right) \tag{2.20}$$

where, Q is leak flow, C_d discharge coefficient, A_0 initial leak flow, g acceleration of gravity, h pressure head inside the pipe, D internal pipe diameter, R_{leak} hydraulic radius of the leak, b length of the leak, a width of the leak, t pipe thickness, υ water kinematic viscosity, V water velocity inside the pipe at leak, γ specific weight of water, and E modulus of elasticity.

- Though according to White (1979), the value of the discharge coefficient (C_d) for a circular hole is 0.595 and for a linear crack is 0.61, Franchini and Lanza (2013) used C_d = 0.61 for both leak shapes.
- The authors parametrized the right-hand-side of Eq. 2.20 by conducting experiments on steel and u-PVC pipes with different sizes and different leak shapes and dimensions. They formulated the correction factor and presented Eq. 2.21 as the best generalized Torricelli formula model.

$$\frac{Q}{C_d\, A_0\, \sqrt{2gh}} = \text{Ø}'_Q\ (X_1 = \frac{h}{D}, X_2 = \frac{R_{leak}}{D}, X_3 = \frac{b}{a}, X_4 = \frac{R_{leak}}{t}, X_5 = \frac{R_{leak}\sqrt{gh}}{\upsilon}, X_6 = \frac{V}{\sqrt{gh}}, X_7 = \frac{\gamma h}{E})$$

$$\begin{aligned}
&= 2.6174\ X_2^{0.083} X_3^{0.001} X_4^{1.118} X_5^{0.145} X_6^{0.009} X_7^{0.011} \\
&- 35.0296\ X_1^{0.002} X_3^{0.001} X_5^{0.011} X_7^{0.002} \\
&+ 875.6679\ X_1^{0.042} X_2^{2.000} X_3^{1.754} X_4^{0.561} X_7^{0.605} \\
&+ 3.4805\ X_1^{0.086} X_2^{0.022} X_3^{0.042} X_6^{0.004} \\
&- 5.5281\ X_1^{0.090} X_2^{0.002} X_3^{0.062} X_4^{1.188} X_6^{0.016} X_7^{0.004} \\
&+ 34.926
\end{aligned} \tag{2.21}$$

In Table 2.14, we have calculated the correction factor $\text{Ø}'_Q$ for the hole and the cracks where the pressure at the leak is 60.12 psi_ 4.227 m_ flow velocity at the leak location is 0.015 (ft/s)_ 0.004572 (m/s), and areas of the leaks are 2 (mm^2). Since it was impossible to experimentally measure velocity and pressure at the testbed leaks, we extracted these values from the EPANET model of the looped network with the hole leak considering that other leaks or the branched network have very similar values. Due to ignoring the leak inclination in the correction factor, there is no difference between the longitudinal and circumferential cracks.

Table 2.14. Values of the correction factor ($\emptyset'_Q$) elements for the hole leak and the linear crack of the testbed

Element	Hole Leak	Crack
X_1	$\frac{4.227\ (m)}{0.1524\ (m)} = 27.74$	$\frac{4.227\ (m)}{0.1524\ (m)} = 27.74$
X_2	$\frac{0.4\times10^{-3}\ (m)}{0.1524\ (m)} = 2.62\times10^{-3}$	$\frac{0.33\times10^{-3}\ (m)}{0.1524\ (m)} = 2.187\times10^{-3}$
X_3	$\frac{1.6\times10^{-3}\ (m)}{1.6\times10^{-3}\ (m)} = 1$	$\frac{2\times10^{-3}\ (m)}{1\times10^{-3}\ (m)} = 2$
X_4	$\frac{0.4\times10^{-3}\ (m)}{10.9728\times10^{-3}\ (m)} = 36.45\times10^{-3}$	$\frac{0.33\times10^{-3}\ (m)}{10.9728\times10^{-3}\ (m)} = 30.38\times10^{-3}$
X_5	$\frac{0.4\times10^{-3}(m)\times\sqrt{9.807\ (\frac{m}{S^2})\times 4.227\ (m)}}{1.004\times10^{-6}\ (\frac{m^2}{s})*} =$ 2565.14	$\frac{0.33\times10^{-3}(m)\times\sqrt{9.807\ (\frac{m}{S^2})\times 4.227\ (m)}}{1.004\times10^{-6}\ (\frac{m^2}{s})*} =$ 2137.61
X_6	$\frac{4.572\times10^{-3}\ (\frac{m}{s})}{\sqrt{9.807\ (\frac{m}{S^2})\times 4.227\ (m)}} = 7.101\times10^{-4}$	$\frac{4.572\times10^{-3}\ (\frac{m}{s})}{\sqrt{9.807\ (\frac{m}{S^2})\times 4.227\ (m)}} = 7.101\times10^{-4}$
X_7	$\frac{9807\ (\frac{N}{m^3})\times 4.227\ (m)}{3.38\times10^9\ (\frac{N}{m^2})**} = 1.226\times10^{-5}$	$\frac{9807\ (\frac{N}{m^3})\times 4.227\ (m)}{3.38\times10^9\ (\frac{N}{m^2})**} = 1.226\times10^{-5}$

* water kinematic viscosity at 20 degrees Celsius
** PVC elastic modulus

Table 2.15 includes the terms of Eq. 2.21, using the values of Table 2.14.

Table 2.15. Correction factor for the hole leak and crack

Correction Factor ($\emptyset'_Q$)	Hole Leak	Crack
$\frac{Q}{C_d\ A_0\ \sqrt{2gh}}$	0.101747401 - 37.58361874 + 1.15133E-06 + 3.947814155 - 0.122620707 + 34.926 = 1.26932326	0.07968684 - 37.53432544 + 2.44299E-06 + 4.048312267 - 0.103131041 + 34.926 = 1.416545067

Regarding Eq. 2.19 for the hole leak and the crack, the leak flows based on the generalized Torricelli equation are according to Table 2.16.

Table 2.16. Leak flows based on the generalized Torricceli equation for the looped network

Leak Flow (GPM)	Hole Leak	Crack
Q	0.22	0.25

The calculated correction factors ($\emptyset'_Q$) and the flows in Table 2.16 indicate a larger leak flow for the crack compared to that of the hole. Comparing Table 2.16 and Table 2.6 specifies the following differences and possible reasons.

- The generalized Torricelli leak flows are smaller than the actual ones. This difference can be due to the inexact pressure and velocity values at the leaks in calculating the correction factors. These inaccurate values cause the elements X_1, X_5, X_6 and X_7 to be underestimated. Knowing the exact values of the velocity and pressure requires experimental measurements at the leak locations while we employed the values calculated by the EPANET model.
- Based on the actual values, the hole leak flows are larger than those of the cracks. However, for the generalized Torricelli leak flows, this relation is the opposite. The larger leak flows in the hole compared to that of the crack in the testbed can stem from the sharp edges of the cracks. Investigating Figures 2.12, 2.13, and 2.14 show the sharper edges of the cracks in contrast to the hole edges. The sharp edges and non-uniform cross-section areas of the cracks decrease the two components of the discharge coefficient (C_d), i.e., the coefficient of velocity (C_v) and the coefficient of contraction (C_c). C_v decreases with increasing contact of water with a leak wall, and C_c gets smaller when the leak shape deviates from a circle. Since the length of the crack wall in contact with water is longer than that of the hole, the coefficient of velocity C_v is small in cracks. Also, the linear shape of the cracks decreases the coefficient of contraction C_c. These reasons cause the discharge coefficient to be

smaller for the cracks than the hole, see Table 2.7. Moreover, the projected edges in the cracks increase friction at the leak opening and decrease pressure head, leading to smaller leak flow.

- The correction factor ignores the linear crack inclination that calculates the same leak flow for the longitudinal and circumferential cracks. However, there is a difference between the actual leak flows when the cracks' inclination changes in the testbed.

2.5. Sensor Characteristics

Three types of sensors were used in the testbed: (1) hydrophone; (2) dynamic pressure; (3) accelerometer.

2.5.1. Hydrophone

Hydrophones that measure acoustic signals have been employed for decades to detect leaks in water pipelines. Hundaidi and Chu (1999) attached hydrophones in an actual water distribution system, with sensitivity within ± 3 dB, to service connections and fire hydrants to detect leak signals. Cody et al. (2018) used two hydrophones, with -137 dB sensitivity, to detect and localize leaks in a laboratory-scale PVC pipe network with a 15.2 cm diameter.

Sounds of pumps, machinery, and vehicles in the vicinity of sensors can increase water pipes' sound frequency to 10 kHz. Also, leak signals span a broadband range of frequencies; however, due to the faster attenuation of high frequencies in pipes, low-frequency sounds represent leak signatures in practical detection methods. This agrees with the 500 Hz frequency cap for detected leaks in the laboratory and field size PVC pipes (Cody et al., 2018; Hunaidi and Chu, 1999; Hunaidi et al., 2000; El-Zahab et al., 2018; Butterfield et al., 2017b; Khuleif et al., 2012).

Cavitation, the formation or growth of air bubbles in water flow, has been named as the main source of leak sound (Papastefanou, 2011). These bubbles form or grow to large sizes when local static pressure reduces after a water jet leaves a pipe. The collapse of these bubbles generates energy and high-frequency sounds that can be measured with acoustic devices.

In this study, we employed two hydrophones with capabilities shown in Table 2.17.

Table 2.17. Characteristics of the hydrophones employed in the testbed

Brand	Model	Frequency	Sensitivity	Mounting Thread	Dimensions
Aquarian	H2c	< 100 kHz	-180dB (reference: 1V/μPa)	1/4”-18 NPT	25 mm × 58 mm

Figure 2.23 shows the Aquarian H2c hydrophone and its dimensions (H2c Hydrophone, 2020).

Figure 2.23. (a) Image of the Aquarian H2c hydrophone

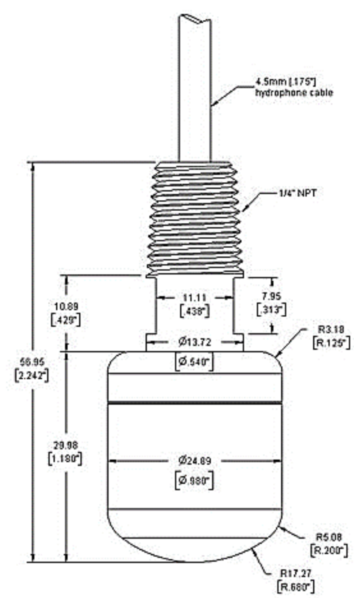

Figure 2.23 (b) Dimensions (in mm) of the Aquarian H2c hydrophone

2.5.2. Dynamic Pressure Sensor

Most of the pressure sensors employed in leak detection studies are static which are appropriate for pressure monitoring in water networks. Butterfield et al. (2017b) used static pressure transducers to solely monitor testbed pressure. Huniadi and Wang (2006) recommended using pressure sensors for large diameter PVC pipes or small leaks. Srirangarajan et al. (2012) and Zan et al. (2014) employed a network of sensor nodes, called WaterWiSe@SG, to detect and localize pipe bursts in a real water distribution system. Each sensing node included a static pressure sensor. The authors simulated bursts with opening fire hydrants that caused a 3 psi to 5 psi pressure drop in the network. They sampled and collected data at 2 kHz though Zan et al. (2014) found 250 Hz sampling frequency would be adequate for leak detection. In an experimental study, Motazedi and Beck (2017) employed static pressure sensors to detect transient waves, created by a solenoid valve, in both straight pipe and T-Junction geometries. The transient events caused pressure drops of about 1 psi. However, their testbed did not include any pressure disturbance, such as demand flow. According to Zhang et al. (2014), ordinary static pressure transmitters that measure absolute pipe pressure are not efficient in detecting small leaks. The reason is that the pressure changes originated from the small leaks have a low signal-to-noise ratio and account for a short pressure transmitter range. Therefore, when there is pressure noise, static pressure sensors have large false alarms and are not appropriate for detecting and localizing small leaks. On the other hand, Dynamic pressure sensors measure the kinetic energy, which comes from a fluid’s velocity and density. Hence, we used the PCB 102B16 pressure sensor, whose information is available in Table 2.18.

Table 2.18. Characteristics of the dynamic pressure sensors employed in the testbed

Brand	Model	Measurement Range	Sensitivity	Resolution	Maximum Static Pressure
PCB	102B16	100 psi	50 mV/psi	1 mpsi	1000 psi

Figure 2.24 shows an image and the dimensions of the PCB 102B16 pressure sensor (PCB Model 102B16, 2020) employed in the testbed.

Figure 2.24. (a) Image of the PCB 102B16 pressure sensor

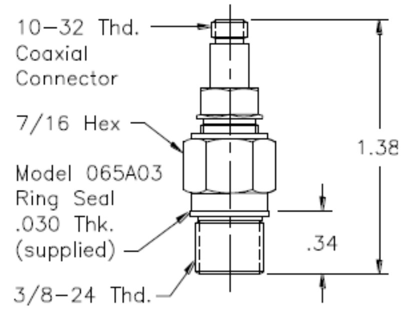

Figure 2.24. (b) Dimensions of the PCB 102B16 pressure sensor

2.5.3. Accelerometer

There are three sources of interactions between pipe and fluid: (1) friction coupling; (2) Poisson coupling; (3) junction coupling (Yazdekhasti et al. 2016). A leak in a water pipe causes a pressure transient that propagates in the water along the pipeline. When a leak happens, the leak expulsion generates a thrust force due to a water jet's momentum and a pressure difference between the water inside the pipe and air pressure. Eq. 2.22 represents the thrust force caused by a leak.

$$F_{thrust} = m\,(V_e\text{-}V_0) + (P_e\text{-}P_0)\,A_e \qquad (2.22)$$

where, m is mass of leaking water, V_e leaking water velocity, V_0 water velocity in pipe at leak location, P_e water pressure at leak location, P_0 air pressure, and A_e the cross-sectional area of the leak.

The junction coupling transfers the thrust force to the pipe's unsupported span that causes the acceleration of the leaking pipe segment. Pipe surface acceleration, $\frac{d^2y}{dt^2}$, can also be modeled as the motion differential equation of a vibrating beam (Seto, 1964):

$$\frac{d^2y}{dt^2} = -\,EI\,\frac{g}{A\gamma}\,\frac{d^4y}{dx^4} \qquad (2.23)$$

where, E is the Young's modulus of the beam, I area moment of inertia, g gravity acceleration, A pipe cross-sectional area, and γ specific weight of the beam.

Accelerometers have been used extensively to detect leaks in water distribution systems in laboratory-scale networks (Yazdekhasti et al., 2016; Yazdekhasti et al., 2017; Butterfield et al., 2017b, Marmarokopos et al., 2018; El-Zahab et al., 2018) and real size pipelines (Hunaidi and Chu 1999; Gao et al. 2005; Martini et al. 2015; Almeida et al. 2014; Almeida et al. 2018; Ma et al. 2019). In real-size water networks, accelerometers have been mounted on fire hydrants or valves (Gao et al., 2005; Almeida et al., 2014; Kang et al., 2018; Ma et al., 2019). In the case of testbeds, accelerometers are often installed in leaks' vicinity (Marmarokopos et al., 2018; Butterfield et al., 2017b; Ismail et al., 2015).

Table 2.19 describes conditions where hydrophones and accelerometers can be used effectively for leak detection.

Table 2.19. Conditions where accelerometers and hydrophones are recommended and not recommended

Author(s)	Accelerometer Recommended	Accelerometer Not Recommended	Hydrophone Recommended	Hydrophone Not Recommended
Hunaidi and Wang (2006)	- Large leaks in PVC pipes. - Any leak in metal pipes	- Large sensor to sensor distance - Small leaks - High vibration noise	- Large distance of leak with sensor - Small leaks - Low-frequency leak signals found in PVC pipes and large diameter pipes with any material	- Large diameter pipes
El-Zahab et al. (2018)	- Persistent changes in network	NA*	- Rapid transients	NA

Table 2.19. Continued

Author(s)	Accelerometer Recommended	Accelerometer Not Recommended	Hydrophone Recommended	Hydrophone Not Recommended
Almeida et al. (2014)	- Leaks with weak signals	- High vibration noise	- Large distance between sensors and leaks - Networks with many junctions	- High acoustic noise
Gao et al. (2005)	- Multiple leaks	- Leaks with low acoustic signals	- Leaks with a small signal-to-noise ratio	NA
Almeida et al. (2018)	- Large resonances caused by hydrants	NA	NA	- Large resonances caused by hydrants
Mohed Ismail et al. 2019	NA	- Large distance of leak with sensor	NA	- Large diameter pipes
Yazdekhasti et al. (2016)	- PVC pipes	NA	- Metal pipes	- PVC pipes
Yazdekhasti et al. (2017)	- Medium to small diameter PVC pipes	NA	NA	- Leaks with narrowband signals (in PVC and large diameter metallic pipelines)

NA*: Not available information

According to Table 2.19, accelerometers are appropriate for leak detection in small to medium PVC pipes, and when there is a low vibration noise. Therefore, employing accelerometers in our testbed seems reasonable.

We employed two types of accelerometers (Table 2.20).

Table 2.20. Characteristics of the accelerometers employed in the testbed

Brand	Model	Sensitivity	Measurement Range	Frequency Range
PCB	352A24	0.1 V/g	±50 g pk	1.0 to 8000 Hz
PCB	333B50	1.0 V/g	±5 g pk	0.5 to 3000 Hz

Since the intensity of leak vibrations was unknown before conducting the experiments, we first used the less sensitive accelerometer (PCB 352A24). Some initial experiments indicated the leak vibrations are in the range of a more sensitive accelerometer. Therefore, we attached the PCB 333B50 to capture vibration signals with higher accuracy. Martini et al. (2018) also employed PCB 333B50 to detect leaks in a testbed using autocorrelation analysis.

Figure 2.25 shows the images of the accelerometers (PCB Model 333B50, 2020; PCB Model 352A24/NC, 2020).

Figure 2.25. (a) Image of the PCB 333B50 accelerometer

Figure 2.25. (b) Image of the PCB 352A24 accelerometer

2.6. Sensor Localization Criteria

Sensors in leak detection have been installed at access points to pipes such as valves or fire hydrants. In this section, the reasons for choosing sensor locations in the testbed will be explained.

2.6.1. Hydrophone Location

Literature indicates that to detect leaks by hydrophones, these sensors have often been mounted at the top or at the bottom of fire hydrants (Cody et al., 2018; Hunaidi and Chu, 1999). This method of deployment is also being used in the water leak detection industry (Visenti, 2020). To simulate hydrants in the testbed, we used an erected 15.24 cm diameter pipe, with a height of 35.56 cm, whose top is closed by a blind flange and bottom is glued to an elbow connected to the distribution section. The center of the blind flange is threaded by a 1.11 cm drill and tap, and the hydrophone is mounted in the flange hole using the 0.63 cm male thread at the top section of the hydrophone body, see Figure 2.26(a). Figure 2.26(a) shows a flipped view of a blind flange with the Aquarian H2c hydrophone. Figure 2.26(b) shows the actual setup of the hydrophone and blind flange in the testbed. The yellow substance is a sealant compound to seal the mounting hole.

Figure 2.26. (a) A flipped view of a blind flange with the hydrophone

Figure 2.26. (b) Actual setup of the hydrophone in the blind flange screwed to the simulated hydrant.

Based on design guidelines (Bryan/College Station Unified Design Guidelines- Domestic Water, 2012), fire hydrants should be located close to street intersections or at specific intervals at residential districts. Therefore, we installed hydrant H1 in the middle of its main pipe, as if it is located in a residential area. Hydrant H2 was also located at one-third of the main pipe length, close to two pipes' intersection. Regarding Figures 2.4 and 2.6, the hydrants are as distant as possible and symmetrical to the bisector of testbed length. Since Hydrant H1 is closer to the service line and the pump, data recorded by hydrophone H1 is expected to be noisier.

2.6.2. Dynamic Pressure Sensor Location

Pressure sensors are often mounted on pipes at valve access points. According to design guidelines (Bryan/College Station Unified Design Guidelines- Domestic Water, 2012), valves should be located (1) one less than every leg of a cross-connection; (2) two legs of a tee connection among other locations; (3) at the end of a public line. We used two dynamic pressure sensors in the testbed. The first one is mounted on the end of the 2.54 cm diameter pipe, where water enters the distribution network. This sensor represents pressure variations before entering the systems where leaks are located. The second sensor is installed on the 15.24 cm diameter pipe located at the bottom left corner of the distribution section. This location was chosen as if it were at the end of a public line and required a valve. The sensor at this location captures pressure variations after the leaks.

The sensors are along the diagonal of the testbed rectangle, both at the same distance from the leaking pipe. However, two reasons may cause the dynamic pressures measured by the two sensors to be different: (1) sensor P1 is closer to the pump compared to sensor P2; (2) sensor P1 is mounted on a 2.54 cm diameter pipe, but sensor P2 is located on a 15.24 cm diameter pipe. Based on Bernoulli's principle, when pipe diameter changes, water pressure alters as well; and (3) sensor P2 is located at the distribution section, which includes multiple junctions. The transient effects of these junctions affect the pressure and pressure variation values at sensor P2.

To mount the pressure sensors, pipes were threaded by 0.95 cm drill and tap, and the threaded heads of the pressure sensors were screwed to the hole. A sealant tape was used to seal the mounting holes. Figures 2.27(a) and 2.27(b) show the dynamic pressure sensors P1, mounted on the 2.54 cm supply line, and the dynamic pressure sensor P2, mounted on the 15.24 cm distribution section, respectively.

Figure 2.27. (a) Dynamic pressure sensor P1 mounted on the 2.54 cm supply line

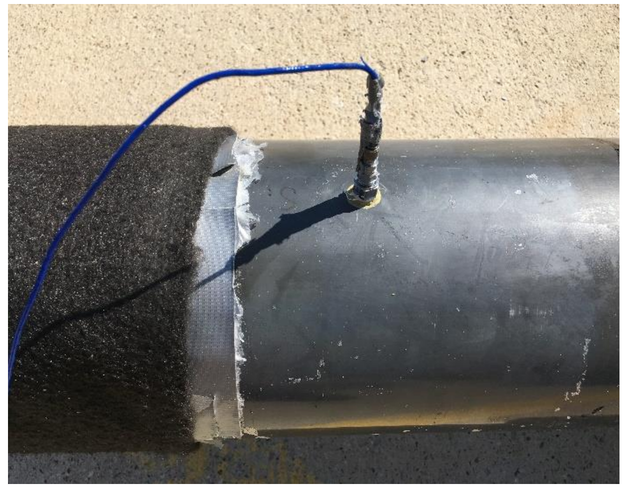

Figure 2.27. (b) Dynamic pressure sensor P2 mounted on the 15.24 cm distribution section

2.6.3. Acceleration Location

As discussed in the dynamic pressure sensors description, sensors can be installed at valve access points at either one less than every leg of a cross-connection or two legs of a tee connection. This sensor localization conforms to other leak detection research conducted on test or real size water networks. For instance, Kang et al. (2018) attached six accelerometers on pipe junctions of an actual water network in Seoul, South Korea. Using a test setup, Martini et al. (2018) installed accelerometers on customized fittings located at two valves, one at a leg of a T-joint and the other at the end of a dead-end pipe. We installed the accelerometer A1 on a leg of a tee junction and accelerometer A2 on the short pipe connected to hydrant 2, see Figure 2.4. The accelerometer locations are symmetric to leaks in both the looped and branched networks. As mentioned in the accelerometer introduction, we employed two types of accelerometers, but we will use the more sensitive one, i.e., PCB 333B50, for analysis. Figures 2.28(a) and 2.28(b) show overviews of the locations of the accelerometers A1 and A2. The red circles highlight the locations of the sensors.

Figure 2.28. (a) Overview of the accelerometer A1 location close to a T-junction

Figure 2.28. (b) Overview of the accelerometer A2 location at the bottom of hydrant 2

Figures 2.29(a) and 2.29(b) show how the accelerometers were mounted on the pipe walls.

The accelerometer PCB 333B50 is designed to be waterproofed. However, the body of accelerometer PCB 352A24 is not hermetically sealed and welded together. Therefore, we used room-temperature-vulcanizing silicone (RTV silicone) to cover the sensor and attached it to the pipe. There is a plastic seat at the bottom of the accelerometer PCB 333B50; see Figure 2.25(a), which provides the sensor with enough flat area to receive vibration signals. Due to the 15.24 cm pipe's surface curvature, we shaved the pipe surface such that enough flat area would be available for the sensor’s seat. A bonding epoxy adhesive was then used to attach the accelerometer PCB 333B50 to the flattened surface on the pipe wall.

Referred to Butterfield et al. (2017b), due to the strong coupling of the pipe wall and the wave generated by the pipe's flowing water, the radial motion of the wave is significant. Hence, the accelerometer orientation would not affect recorded vibration signals. Though the coupling mechanism may cause the leak vibration wave to propagate both radially and axially, since the leak vibration stems from a thrust force, the axial vibration that is the pipe displacement parallel to its axis is predominant (Martini et al., 2018). Thus, the accelerometer PCB 333B50 that is a uni-axial sensor can measure pipe vibrations well.

Figure 2.29. (a) Accelerometer A1 mounted on a T-junction leg

Figure 2.29. (b) Accelerometer A2 mounted on pipe connection to hydrant 2

2.7. Data Acquisition Devices

The main elements of a data acquisition system (DAQ) are (1) sensors and transducers; (2) field wiring; (3) signal conditioning; (4) data acquisition hardware; (5) computer (operating system); (6) data acquisition software (Park and Mackay, 2003).

In the previous sections, the first element of the testbed DAQ was explained. In this part, we describe other elements employed in our experiments.

The pressure sensor and accelerometer data were transmitted to acquisition hardware using low-noise coaxial cables. Hydrophones were also connected to a different acquisition device with high quality and low-noise cables. The cables were as short as possible to minimize signal noise during transmission and along the most straightforward path to the acquisition hardware.

Two NI 9234 modules were used to acquire the pressure sensors' outputs and accelerometer signals, overall six signal inputs. Table 2.21 includes information about NI 9234. NI 9234 embodies these modules: (1) signal conditioning; (2) anti-aliasing filter; (3) digitization.

Table 2.21. Characteristics of the NI 9234 module as an element of the DAQ employed in the testbed

Product Name	Channels	Signal Ranges	Input Configurations	Sample Rate	Digitization Resolution	Connectivity
NI 9234	4	±5 V	IEPE with AC Coupling, AC Coupling, DC Coupling	51.2 kS/s/ch	24 Bit	BNC plug

Figure 2.30 shows two NI 9234 modules assembled on a NI cDAQ-9178 chassis. The chassis has a power input that provides electricity for NI 9234 operation and sensors excitations, and a USB output to transfer digitized sensory data to a computer.

Since the pressure sensors and accelerometers' sampling rates are less than that of NI 9234, the sensed data would be recorded based on the sampling frequency of NI 9234, i.e., 51.2 kHz.

Figure 2.30. Two NI9234 modules assembled on a NI cDAQ-9178 chassis

To acquire acoustic signals, the hydrophones were connected to the ZOOM UAC-2 audio converter. The ZOOM UAC-2 sampled the hydrophone signals via two 24-bit/192 kHz high-resolution input channels simultaneously. The converter digitizes signals and sends them to a computer using a USB 3.0 output and transfer protocol. Figure 2.31(a) and 2.31(b) show the front and rear views of the ZOOM UAC-2 (ZOOM UAC-2 USB Audio Interface, 2020).

Figure 2.31. (a) Front view of the ZOOM UAC-2 with inputs for hydrophones H1 and H2

Figure 2.31. (b) Rear view of the ZOOM UAC-2 with a USB 3.0 output (in blue)

Audacity 2.3.3 software was employed to record the converted data on a local computer. A channel was created per hydrophone in the software, and each channel's data was saved in .RAW and .WAV formats. The .RAW data was recorded and saved on a computer with the information in Table 2.22. Audacity recorded the digitized data at a rate of 8000 Hz. To import .RAW data to

Audacity, the information in Table 2.22 can be used to convert the saved digital data to analog signals.

Table 2.22. Parameters used to save .RAW recorded data on a local computer

Data Format	Sample Rate	Channels	Byte Order	Encoding
.RAW	8000 Hz	One channel (mono) per hydrophone	Little-endian	Signed 32-bit PCM

Figure 2.32 shows recorded data of hydrophones H1 and H2 for an experimental scenario where the horizontal axis is time (s), and the vertical axis is signal amplitude.

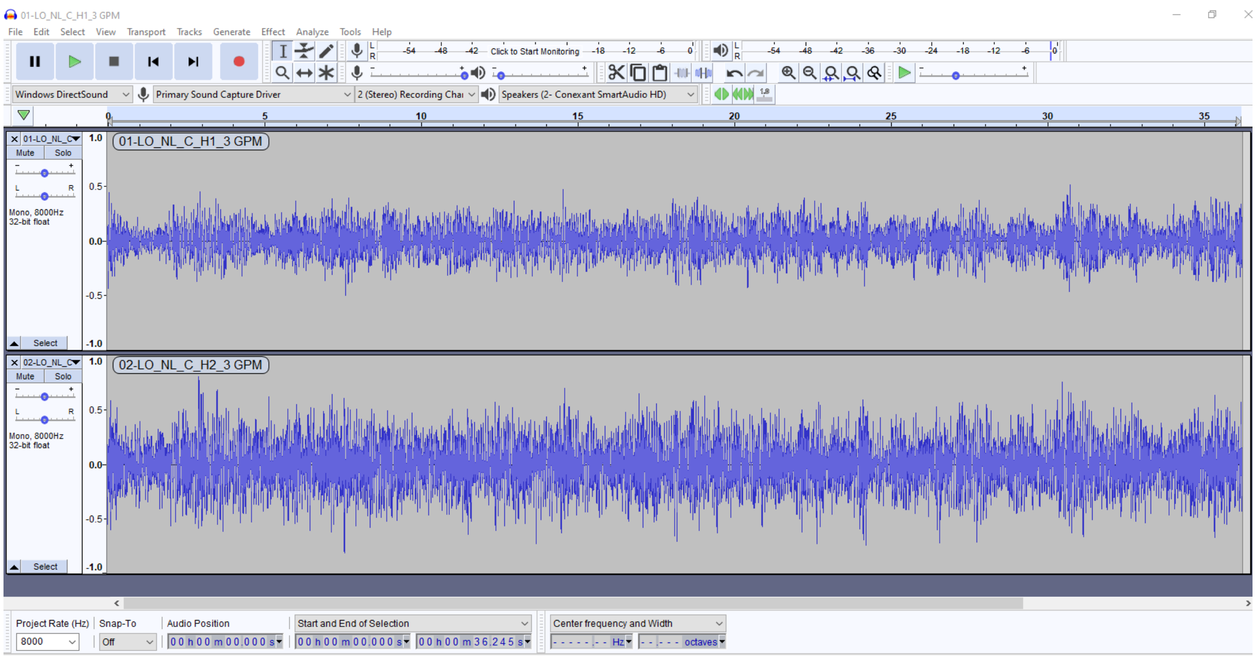

Figure 2.32. A view of Audacity visualizing data of hydrophones H1 and H2

It is worth noting that all sensors' data were recorded for 30 (s) and at rates that follow the Nyquist sampling theorem. According to the Nyquist sampling theorem, to accurately represent a time signal, the signal's sampling rate must be greater than twice the signal's maximum frequency

(Kang et al., 2018). Regarding the frequency of water leak signals that are smaller than 1000 Hz, the sensors and DAQs are selected and set so that their sampling frequencies are larger than 2000 Hz.

2.8. Ambient Sound

Background noise in sensory data has been a significant challenge in detecting leaks. Multiple research has taken into consideration the effects of ambient noise into their studies. Hunaidi and Chu (1999) measured ambient noise with a frequency between 5 and 50 (Hz) and believed that leak signals are dominated by noise for frequencies below 5 (Hz). The authors remarked underground power cables, rotating machinery, pumps, and cooling systems as the ambient noise sources. Butterfield et al. (2017b) reported the background noise, mainly generated by testbed pumps, was most dominant at frequencies less than 50 (Hz). Marmarokopos et al. (2018) also believed that external noises are at frequencies below that of consumption flow; otherwise, ambient noise can be a dominant cause of severe interferences.

In the Canadian testbed used by Cody et al. (2018), Cody et al. (2020), and Harmouche and Narasimhan (2020), though non-stationarity conditions were introduced in signals due to the fluctuations in the input flow provided by a pipe connected to city network, the authors did not take into account the effects of ambient noise as discussed above. They highlighted eliminating pump noise effects as a benefit of supplying input water directly from city pipelines that might be a simplifying factor and adversely affect sensors' data quality for leak detection.

In this testbed, background noise was generated by two sources: (1) water pump; (2) ambient noise. The water pump's noise was continually propagated throughout the network by the pipes and water filling the pipes. The ambient noise was created by (1) a traffic sound played on a

100-Watt speaker located on the leaking pipe; (2) carrying a switched-on reciprocating saw to different points of the testbed during the measurement time. Figures 2.33(a)-2.33(d) show the time series plots and spectrums of the ambient noise measured at the hydrophones H1 and H2.

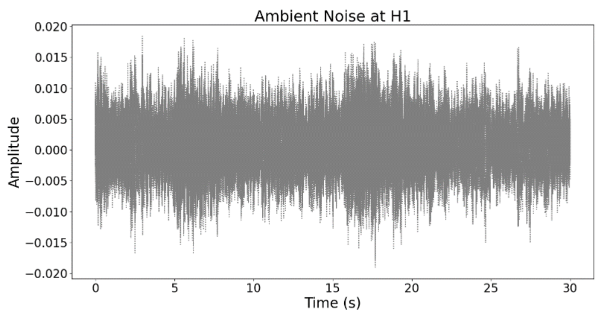

Figure 2.33. (a) Time series of ambient noise measured at H1

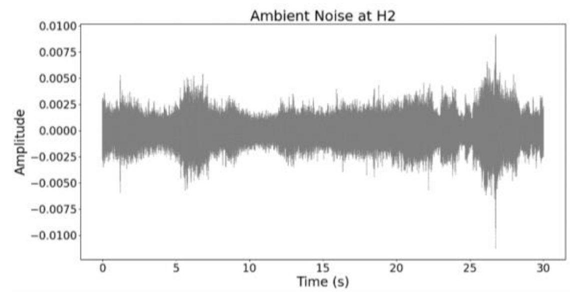

Figure 2.33. (b) Time series of ambient noise measured at H2

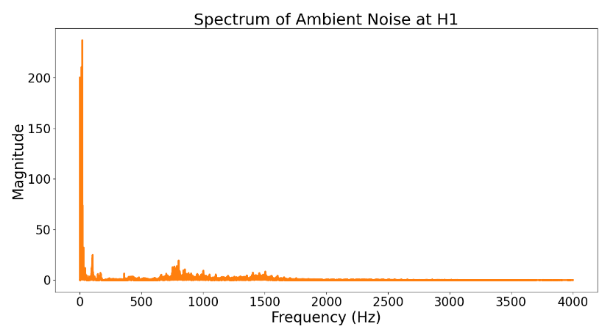

Figure 2.33. (c) Spectrum of ambient noise at H1

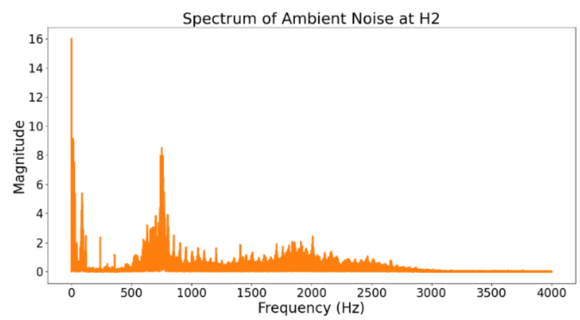

Figure 2.33. (d) Spectrum of ambient noise at H2

Comparing the above figures show that ambient noise is not the same at the two hydrophones. This difference is more aligned with actual conditions where ambient noise varies at different locations in an actual water network. Figures 2.33(a) and 2.33(c) indicate noise dominance at low frequencies of H1. The pump sound potentially generates this noise. Since H1 is closer to the water pump, the pump noise signature is more highlighted at H1 than H2 distant from the pump.

2.9. Resonance

Hunaidi and Chu (1999) observed an anomalous high amplitude at about 90 Hz of the acoustic signal measured by a hydrophone located at a testbed service connection. The authors found the resonant response of the vertical service pipe as the cause of the anomaly. They also believed that the test pipe's longitudinal resonance frequencies and soil resonance contribute to the noise dominance at frequencies below 5 Hz, where ambient noise is also present.

Gao et al. (2017) evaluated the effects of resonance on leak detection and localization methods in buried plastic pipes. According to the authors, resonance behavior in pipe systems can decrease the bandwidth over which a technique like the Basic Cross-Correlation (BCC) function can extract time delay information. Modeling the frequency response functions of resonators such as valves in leak signals, they removed adversary influences of resonators on leak signal characteristics. Such a preprocessing resulted in signals with wider bandwidth and improved shape of the cross-correlation function.

Almeida et al. (2018) investigated the efficacy of two commonly used correlation algorithms used for leak localization when pipelines' structural dynamics affect time delay estimations. The analytical and experimental modeling of a pipeline and resonators found that the Phase Transform (PHAT) method is more sensitive to resonance than the BCC function for leak localization. The authors highlighted the structural dynamics of pipe systems as the main cause of resonance in water pipelines.

In this testbed, the resonance has been created by two components in testbed: (1) the service pipe; (2) the simulated hydrants. The 2.54 cm service pipe has a height of 109.22 cm, and each of the hydrants has a height of 35.56 cm. This design makes these components resonators with at least one degree of freedom generating noise in the acoustic signal. Adding the resonance effects to the data makes acoustic data more similar to real conditions and more applicable for leak localization and detection.

2.10. Scenarios

Figure 2.34 shows the experimental scenarios performed in this study. For each network architecture, i.e., looped and branched, there are four leak types: (1) No Leak; (2) Orifice; (3) Longitudinal Crack; (4) Circumferential Crack. Per leak type, different demand and ambient sound cases created five variations: (1) Demand and Sound; (2) No Demand and No Sound; (3) No Demand and Sound; (4) Transient and Sound; (5) Transient and No Sound. In the variations, Demand and No Demand are determined if the service line has outflow, Sound and No Sound manifest whether ambient noise is present. Transient represents a condition where the flow rate abruptly changes from 7.5 GPM to 0 GPM at about the second 20. The Transient mode was created by shutting off the globe valve immediately. We employed one sensor type_ hydrophone_or three sensor types_ hydrophone, pressure transducer, and accelerometer, depending on the demand and sound variations. Per parameter, we used two sensors. These variations resulted in one hundred and forty measurements per network architecture and two hundred and eighty total measurements.

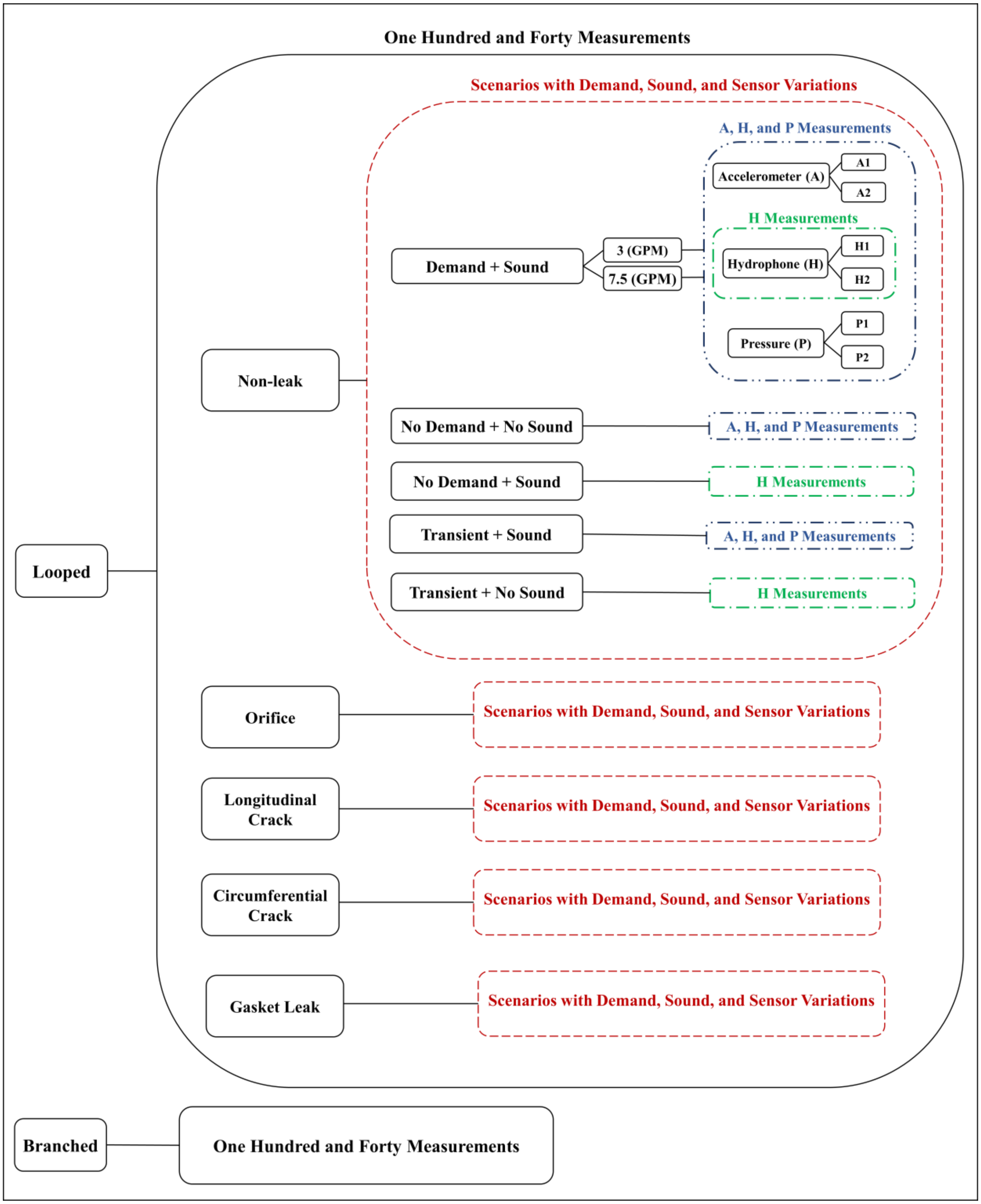

Figure 2.34. Experimental scenarios performed in this study

2.11. Experimental Design Observations

So far, the design and assembly of a laboratory scaled water distribution system were described. Of note, considering the limited room to build the testbed, building a dimensional analysis model was not possible, and the model was somewhat distorted.

The testbed was composed of two sections: (1) water supply with 2.54 cm diameter pipes (supply line); (2) water distribution with 15 cm diameter pipes (distribution section). The length and width of the distribution section are 7.35 m and 4.9 m, respectively. Pipes and fitting were connected via flanges that could be tightened with bolts and nuts.

Two network architectures were simulated: looped and branched. The looped network included eighteen 15-cm diameter pipes, and the branched network was composed of thirteen pipes with a 15-cm diameter. A middle pipe was the leak location in both networks, which could be replaced on demand.

The Goulds 1MC1G1A0 centrifugal fixed speed pump with a 43 m cut-off head was used to provide enough pressure in the distribution section. Concrete blocks were employed to fix the whole testbed and act like the thrust blocks in real water distribution systems.

To simulate the sounding material's dampening effects, all pipes were wrapped with a geotextile, except for a range of 15.24 cm from leak openings. To introduce noise in sensor data, a service valve was employed to simulate water consumption with 3 GPM and 7.5 GPM. These flows were computed regarding a base water demand for 100 people and the Micropolise virtual water network demand multipliers at 1:00 a.m. and 6:00 a.m.

Two MACH 10 ultrasonic meters were utilized to measure leak flows. One meter measured total input water to the distribution section, and the second meter measured consumption flow at the service line. The difference between these flows gave leak flow rates.

Computation of hydraulic and minor headlosses showed that the testbed's headloss was negligible, 0.34 m, compared to 42.9 m pump head at the 8.12 GPM maximum flow consumption. To simulate more realistic conditions, four types of leaks were simulated: (1) orifice (hole); (2) longitudinal crack; (3) circumferential crack; (4) leak at a joint gasket (gasket leak). These leaks were induced in the middle pipe between the two crosses in Figures 2.4 and 2.6. The orifice leak diameter was 1.6 mm, and the size of the longitudinal and circumferential cracks was 1 mm × 2 mm.

The objective in the testbed's design was inducing leaks whose flow rates are at most 30% of the total input water to the distribution section. The above leak sizes were determined based on a trial-and-error procedure, which keeps leak flow rates less than the maximum rate. Actual leak flow rates revealed the following information:

i. Leak rates for the gasket leaks are larger than other leak types. This difference stems from the way that the gasket leak was induced where flange bolts were loosened.

ii. Hole leak rates are greater than the leak rates in the longitudinal and circumferential cracks. The difference is due to dissimilar leak exponents and discharge coefficients, though leak sizes were the same.

Dissimilar leak shapes cause different shapes of water jet at the leak locations. For instance, the cross-section area of the hole water jet maintains a circular shape. In contrast, the longitudinal and circumferential water jets' cross-section area has a rectangular shape close to the leaking crack. Then these water jets diverge from the centroid of the leak area.

To evaluate if the chosen Discharge Coefficients (C_d) were reasonable, leak flows were calculated by EPANET and then were compared with measured leak flows. Since the difference between EPANET and actual leak flows are less than 10 percent of the actual leak flows, the chosen discharge coefficients can represent the leak characteristics.

Comparing the research's leak characteristics with actual leaks reported in other research shows that the induced leaks' dimensions, discharge coefficients (C_d), and leakage exponents (β) met the objective of design as long as the testbed size and flow constraints were deemed. The generalized Torriccelli formula developed by Franchini and Lanza (2013) was evaluated to verify the actual leak flow rates from a dimensional analysis perspective, which yielded the following takeaways.

i. The generalized Torricelli leak flows were smaller than the actual ones. This difference could be due to the inexact pressure and velocity values at the leaks in calculating the generalized Torriccelli formula's correction factors.

ii. Based on the actual values, the hole leak flows were larger than those of the cracks. However, for the generalized Torricelli leak flows, this relation was the opposite. The larger leak flows in the hole compared to that of the testbed's cracks, which can stem from the cracks' sharp and projected edges.

iii. The correction factor ignores the linear crack inclination that causes the same leak flows for the longitudinal and circumferential cracks. However, there is a difference between the actual leak flows when the cracks' inclination changes in the testbed.

Three types of sensors were employed in the testbed: (1) hydrophones, which measure leak and noise acoustic signals; (2) dynamic pressure sensors that measure the kinetic energy of fluids and are appropriate for capturing a short pressure transmitter range; (3) accelerometers to capture thrust forces generated by a leak water jet's momentum and a pressure difference between the water inside the pipe and air pressure.

To simulate real-world conditions, two hydrophones were mounted on two simulated fire hydrants that each was an erected 15.24 cm diameter pipe, with a height of 35.56 cm, whose top was closed by blind flanges and bottom was glued to an elbow connected to the distribution section. One hydrant was installed in the middle of its main pipe as if it is located in a residential area. Another hydrant was also located at one-third of the main pipe length, close to two pipes' intersection.

Two dynamic pressure sensors were used in the testbed. The first one was mounted on the end of the 2.54 cm diameter pipe, where water entered the distribution network. This sensor represented pressure variations before entering the systems where leaks were located. The second sensor was installed on the 15.24 cm diameter pipe located at the bottom left corner of the distribution section. This location was chosen as if it was at the end of a public line and was a valve location. The sensor at this location captured pressure variations after the leaks.

One accelerometer was mounted on a leg of a tee junction, and the second accelerometer was installed on a 15.24 cm diameter short pipe connected to a hydrant. The accelerometer locations were symmetric to leaks in both the looped and branched networks. Two NI 9234 modules were used to acquire the pressure sensors' outputs and accelerometers' signals with a 51.2 kS/s/ch sampling frequency.

The hydrophones were connected to the ZOOM UAC-2 audio converter that sampled the hydrophone signals simultaneously via two 24-bit/192 kHz high-resolution input channels.

Audacity 2.3.3 software was employed to record and save the converted acoustic data on a local computer in .RAW and .WAV formats.

To simulate ambient noise affecting real water networks, background noise was simulated by the water pump, a traffic sound played on a 100-Watt speaker located on the leaking pipe, and carrying a switched-on reciprocating saw to different points of the testbed during the measurement time. Time series plots of the ambient noise at two hydrophones showed that noise at the hydrophone close to the pump was more dominant. Moreover, frequency plots showed that the ambient noise spanned up to the 3000 Hz frequency.

In other research using real water distribution systems, anomalous high amplitudes at low frequencies were reported, which were related to the resonant response of pipe systems such as vertical service pipes, hydrants, or valves. According to Gao et al. (2017), resonance behavior could decrease the frequency bandwidth. In this testbed, two components created the resonance response: (1) a 2.54 cm diameter service pipe with a height of 109.22 cm; (2) the simulated hydrants. This design made these components resonators with at least one degree of freedom generating noise in the acoustic signal.

In this study, one hundred and forty measurements per network architecture and two hundred and twenty-four total measurements were recorded. The measurements varied based on two network architectures, four leak types, twelve demand and ambient noise scenarios, and six sensors.

2.12. Evaluation of Measurements

This subsection will evaluate leak flows, measurement plots and employ features to learn about characteristics of network architectures and leak types. In the following, ND, OL, LC, CC, and GL stand for no demand, orifice leak, longitudinal crack, circumferential crack, and gasket leak, respectively.

2.12.1. Measured Leak Flow Rates

Figures 2.35 and 2.36 show bar plots of leak flow rates in the looped and branched networks, respectively.

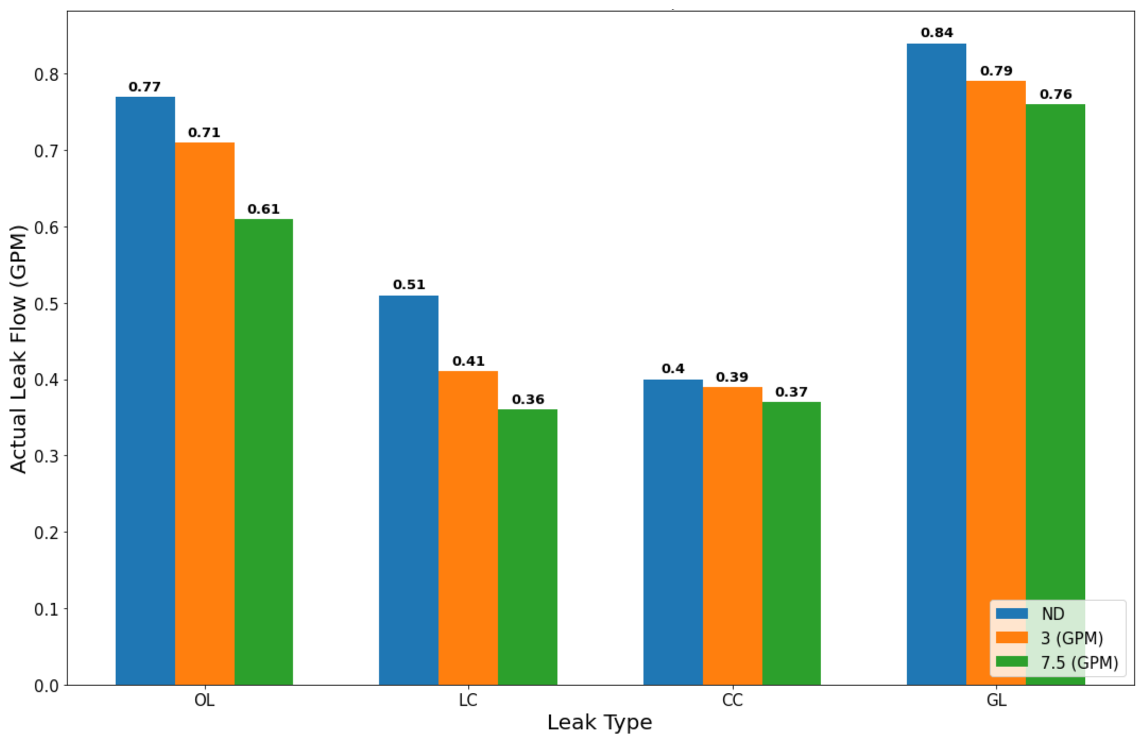

Figure 2.35. Leak flow rates (GPM) for different leaks and demand variants in the looped network

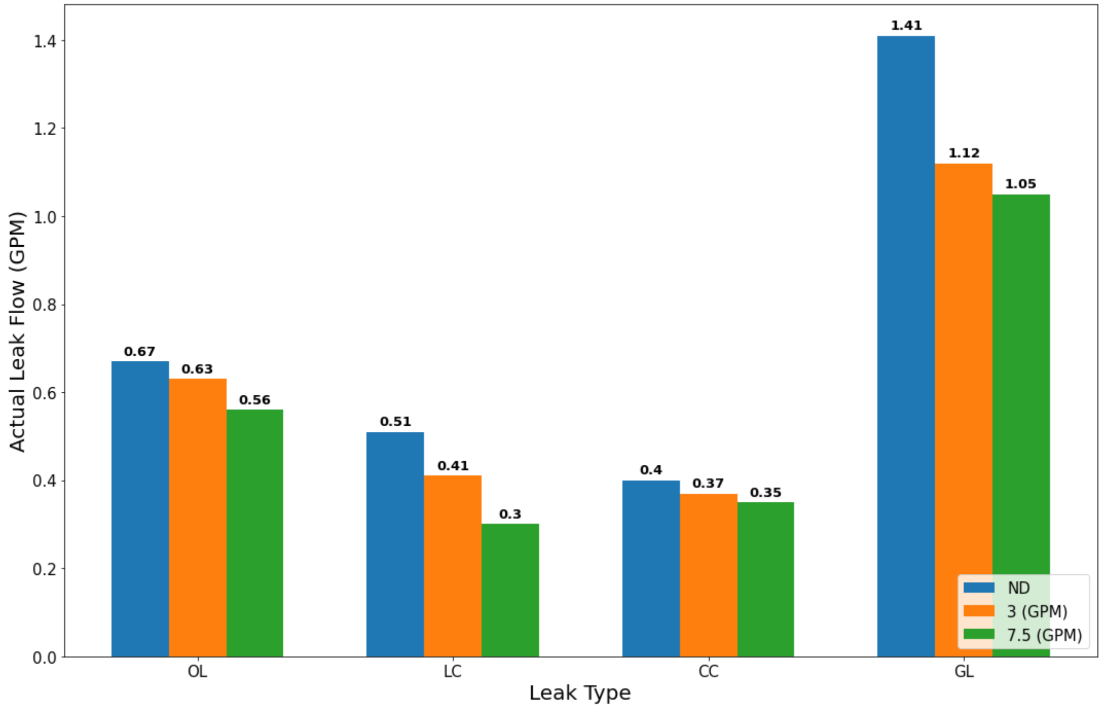

Figure 2.36. Leak flow rates (GPM) for different leaks and demand variants in the branched network

In both networks, leak flow rates decrease when demand increases. Regarding the generalized Torricelli formula, Eq. 2.20, leak flow drops when the pressure head inside a pipe reduces. Since the pump has a constant speed and regarding the pump curve in Figure 2.11, the pressure inside pipes decreases with a demand increase. Since an objective was to evaluate the effects of demand variations on leaks, we made sure to provide the experiments' design demands. Therefore, due to a pressure decrease caused by increased demands, leak flow dropped when demand flow increased.

Except for the GL, other leaks' flow rates were larger in the looped network than those of the branched network. This difference can be due to the more connectivity in the looped network, which delivers water from different paths and provides a higher pressure in the network. While in the branched network, pipes were disconnected, and pump pressure could not be distributed in the network more evenly.

2.12.2. Evaluation Criteria

Table 2.23 lists features used in other research to study leak and no leak signal characteristics. The features are in time-domain, frequency-domain, or time- and frequency-domain.

Table 2.23. List of features employed in other research to study leak and no signals

Author(s)	Sensor Type	Time-domain feature	Frequency-domain feature	Time and frequency feature
Kang et al. (2018)	Accelerometer	- Peak - Average - RMS of time-series signal	-	-
Guo et al. (2021)	Hydrophone	- RMS of signal - Mean of signal - Zero-crossing rate - Standard deviation - Mean Teager energy operator - Autocorrelation energy ratio - Energy entropy ratio	- Mean decibel (dB) of power spectral density - RMS of intrinsic mode functions - Shannon entropy of IMFs - Subband spectral entropy	- Short-time Fourier transform - Wavelet analysis - Spectrogram with different time-frequency resolutions

Table 2.23. Continued

Author(s)	Sensor Type	Time-domain feature	Frequency-domain feature	Time and frequency feature
Butterfield et al. (2018)	Accelerometer	- RMS of IMFs1-6 - Shannon entropy of IMFs 1–6 - Shannon entropy of the whole signal - RMS of the whole signal - Standard deviation - Signal power - Kurtosis - Skewness - Crest factor	- Maximum dB of power spectral density - Minimum dB of power spectral density - Fundamental frequency - Spectral flux	-
Sun et al. (2014)	Accelerometer	-	-	- Local mean decomposition envelope spectrum entropy
Sun et al. (2016)	Accelerometer	-	-	- Root mean square entropy of local mean deposition

Research in Table 2.23 employed the features to extract representative information from signals and used predictive analysis methods to detect or localize leaks or discern leak shapes. In the following subsections, plots and features were employed to evaluate (1) how the network architecture change affected leak characteristics; (2) how a leak signature varied in recorded signals when we altered a leak type. Six types of plots were used for descriptive analyses of signals with no demand, including time-series plots, Fourier transforms frequency plots, cumulative distribution plots, box plots, cross-spectral plots, and leak:noleak ratio plots. These plots were mainly for visualization purposes, so we only visually investigated signals with no demand. Also, evaluating signals free of demand interruptions is an exclusive benefit of such a testbed that is not possible in studies using actual water networks. We could focus on network architecture and leak type effects on leak signatures by investigating data without a demand.

In cross-spectral and leak:noleak ratio plots, graphs were transmitted to numbers for better comparisons. We used ten features to assess signals numerically. Table 2.24 lists the features that are in the time or frequency domains.

Table 2.24. Numerical features to evaluate measurements

Time-domain	Frequency-domain
- Mean - Standard deviation - Zero-crossing rate - Root mean square - Crest factor	- Leak detection index - Dominant frequency - Fundamental frequency - Spectral centroid - Power spectral entropy

In the following, Table 2.24 features are defined.

- Mean

Equation 2.24 indicates the mean (μ) of a signal where x_i is a signal magnitude at the time *i* and *n* is the number of samples.

$$\mu = \frac{x_1 + \ldots + x_n}{n} \tag{2.24}$$

- Standard Deviation

Equation 2.25 represents the standard deviation (SD) of a signal where x_i is a signal magnitude at the time *i*, *n* is the number of samples, and *μ* is the signal mean.

$$SD = \sqrt{\frac{\Sigma(x_i - \mu)}{n}} \tag{2.25}$$

- Zero-crossing Rate

The zero-crossing rate (ZCR) is the number of times that a signal passes the time axis and can represent frequency variations of a signal (Guo et al., 2021). Eq. 2.26 represents the ZCR formula.

$$ZCR = \frac{1}{n-1} \sum_{i=1}^{n-1} 1_{\mathbb{R}<0} \left(x_i x_{i-1} \right) \tag{2.26}$$

where x_i is a signal magnitude at the time i, n is the number of samples, and $1_{\mathbb{R}<0}$ is an indicator function.

- Root Mean Square

The root mean square (RMS) of a signal reflects the variation in the signal amplitude and the vibration energy of the signal (Sun et al., 2016). Eq. 2.27 represents the RMS formula.

$$RMS = \sqrt{\frac{1}{n} \sum_i {x_i}^2} \tag{2.27}$$

- Crest Factor

The crest factor (*C*) indicates how extreme a signal's peak is. It is the ratio of the peak value and the RMS of a signal. Eq. 2.28 represents the crest factor formula.

$$C = \frac{|x_{peak}|}{x_{rms}} \tag{2.28}$$

- Leak Detection Index

Yazdekhasti et al. (2016) and Yazdekhasti et al. (2017) devised and employed a leak detection index (LDI) to detect the onset of leaks in laboratory-scale testbeds using accelerometers. LDI is a distance-based damage index that quantifies differences in a pipeline using cross-spectral density (CSD) pipeline changes caused by a leak. The CSD can be calculated by taking the Fourier transform of two signals' correlations and represents the shared power between two signals. In other words, the CSD can capture leak-caused variations in cross-correlated power spectral densities of two different signals. The larger CSD values indicate the larger correlations between two sensor data. Based on Eq. 2.29, LDI is the normalized difference between the CSD functions of two sensors in the leak and no leak conditions.

$$LDI = \frac{\int |x(\omega)_{nl} - x(\omega)_{l}|}{\int |x(\omega)_{nl}|} \tag{2.29}$$

where, $x(\omega)_{nl}$ is the CSD of no leak condition and $x(\omega)_{l}$ is the CSD of the leak one.

- Dominant Frequency

The dominant frequency shows a periodic behavior due to one frequency and carries the maximum energy among all frequencies found in a signal's spectrum (Telg´arsky, 2013).

- Fundamental Frequency

The fundamental frequency is the smallest frequency which has a peak among all frequencies in a power spectrum (Telg´arsky, 2013). A fundamental waveform frequency is the greatest common divisor of all the frequency components in a signal to be more accurate.

- Spectral Centroid

The spectral centroid is the weighted mean of the frequencies present in a signal, determined using a Fourier transform and magnitudes as the weights.

$$\text{Spectral Centroid} = \frac{\sum_{k=1}^{N} KF[k]}{\sum_{k=1}^{N} F[k]} \tag{2.30}$$

where F [k] is the amplitude corresponding to bin k in a discrete Fourier transform spectrum.

- Power Spectral Entropy

Power spectral entropy (PSE), which is based on the Shannon entropy, is a measure of a signal energy distribution uniformity in the frequency domain. High frequency-domain entropy reflects more uniformity in a signal frequency domain energy distribution, like in a pure sinusoid. While low entropy implies less uniformity, like in white noise (Boashash, 2016). To calculate power spectral entropy, power spectral density (PSD) should be calculated using the fast Fourier transform. In the following, $p(\omega_i)$ is the PSD of a signal.

$$p(\omega_i) = \frac{1}{N} |x(\omega_i)|^2 \tag{2.31}$$

where $x(\omega_i)$ is a signal spectral amplitude at the frequency ω_i.

Then normalized PSD needs to be calculated by:

$$p_i = \frac{p(\omega_i)}{\sum_i p(\omega_i)} \tag{2.32}$$

PSE was calculated based on the Shannon entropy:

$$PSE = - \sum_{i=1}^{n} p_i \ ln(p_i) \tag{2.33}$$

where, p_i is the normalized PSD probability distribution (Helakari et al., 2019).

The next subsections will evaluate accelerometer, hydrophone, and dynamic pressure sensor data based on the described plots and features.

2.12.3. Accelerometer Measurements

In this subsection, we analyzed the data of the accelerometers A1 and A2.

2.12.3.1. Time-domain plot (for ND signals)

Figures 2.37 to 2.44 show the time-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor A1.

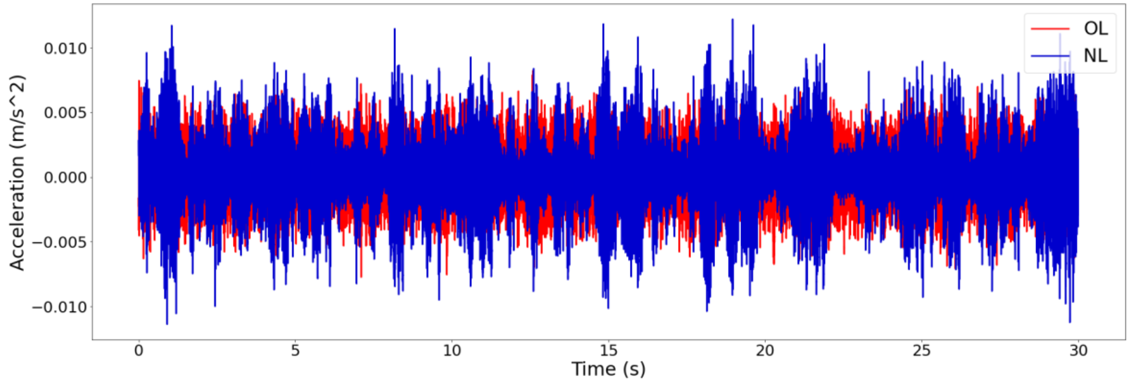

Figure 2.37. Time-domain plots of OL vs. NL signals in the looped network measured by sensor A1

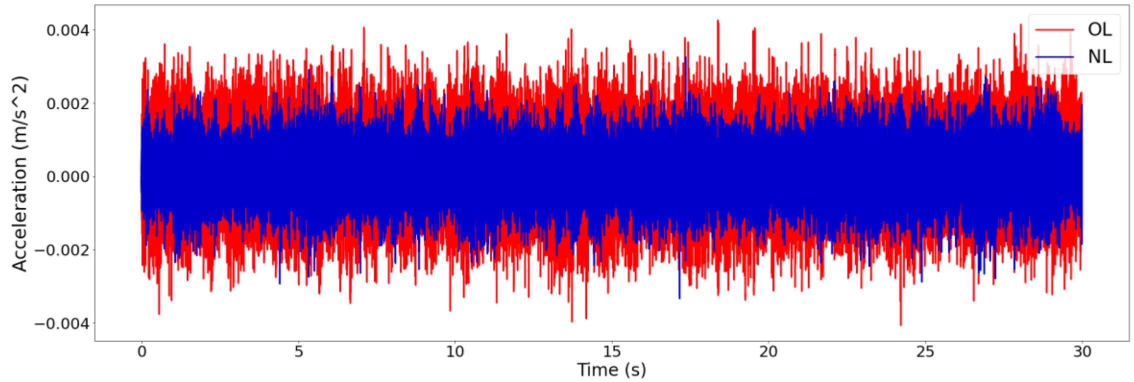

Figure 2.38. Time-domain plots of OL vs. NL signals in the branched network measured by sensor A1

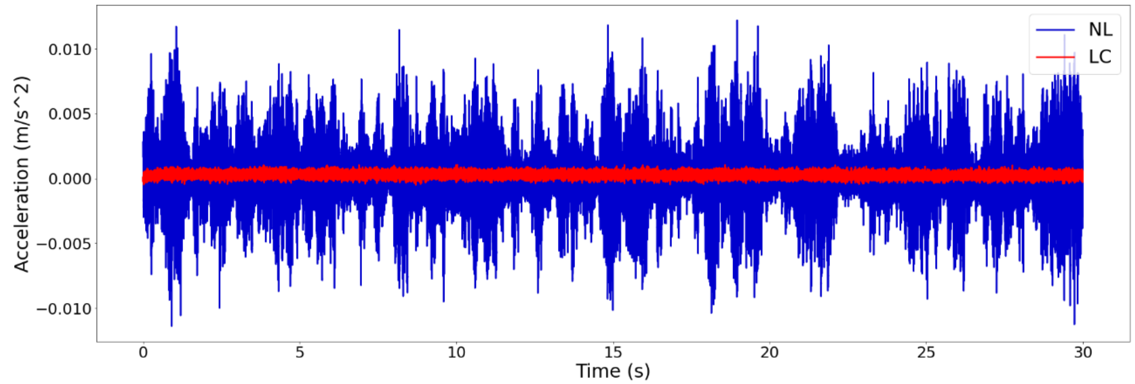

Figure 2.39. Time-domain plots of LC vs. NL signals in the looped network measured by sensor A1

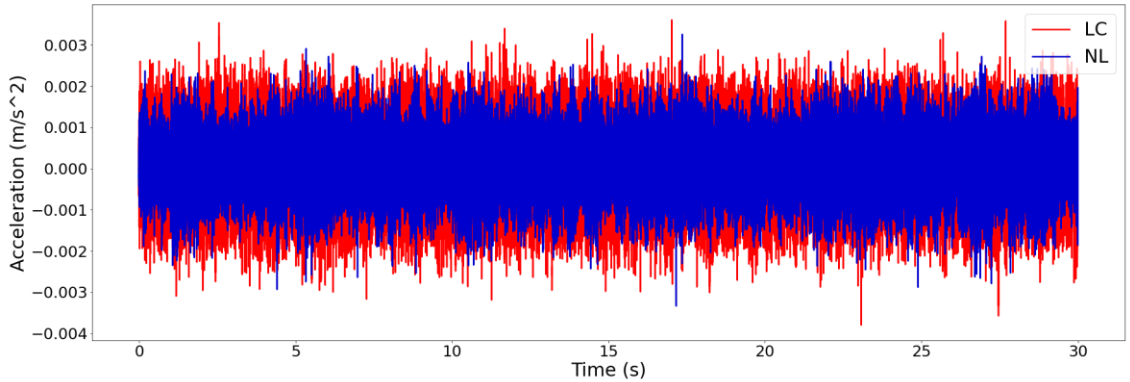

Figure 2.40. Time-domain plots of LC vs. NL signals in the branched network measured by sensor A1

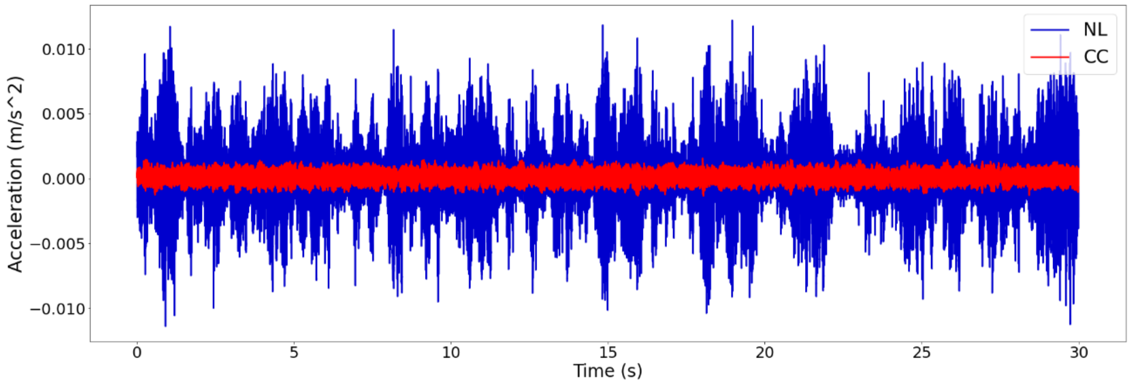

Figure 2.41. Time-domain plots of CC vs. NL signals in the looped network measured by sensor A1

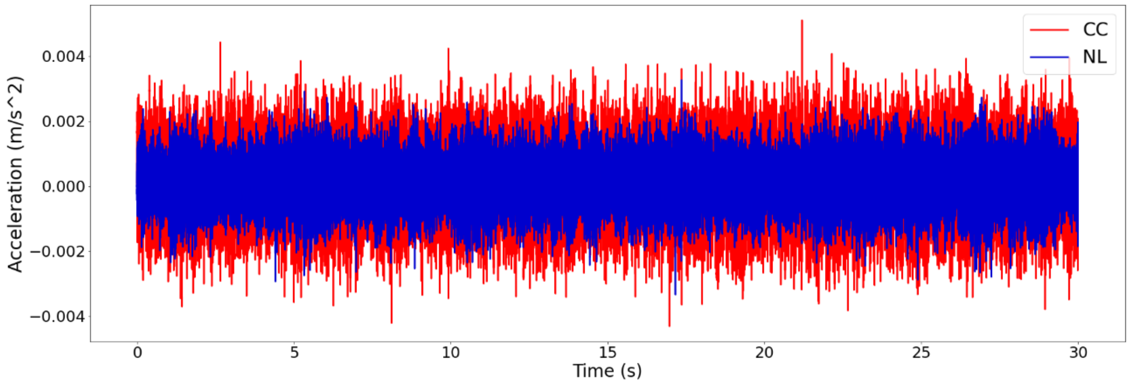

Figure 2.42. Time-domain plots of CC vs. NL signals in the branched network measured by sensor A1

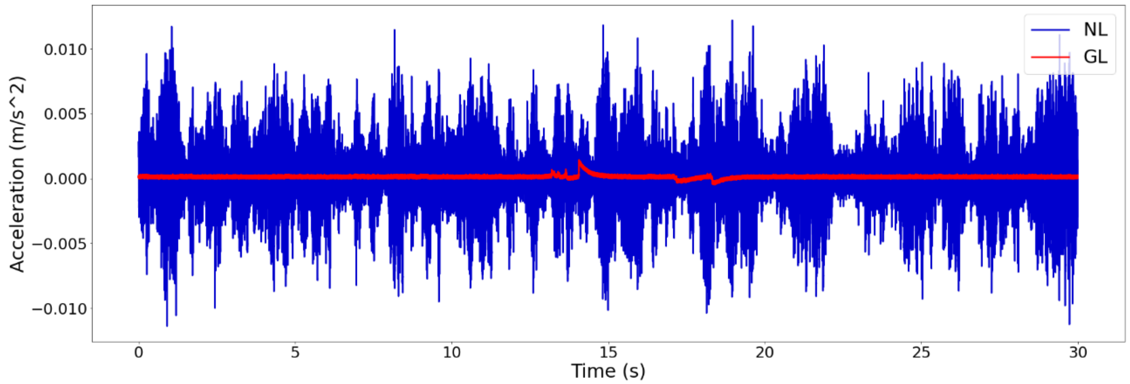

Figure 2.43. Time-domain plots of GL vs. NL signals in the looped network measured by sensor A1

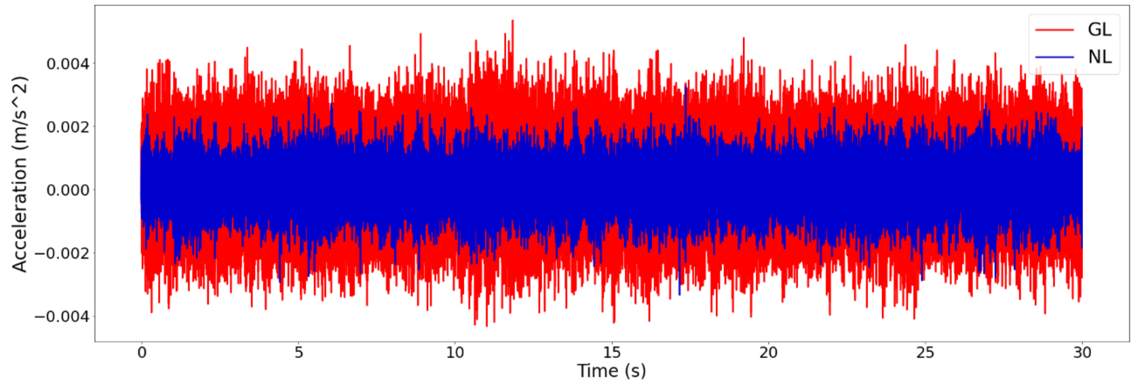

Figure 2.44. Time-domain plots of GL vs. NL signals in the branched network measured by sensor A1

Figures 2.45 to 2.52 show the time-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor A2.

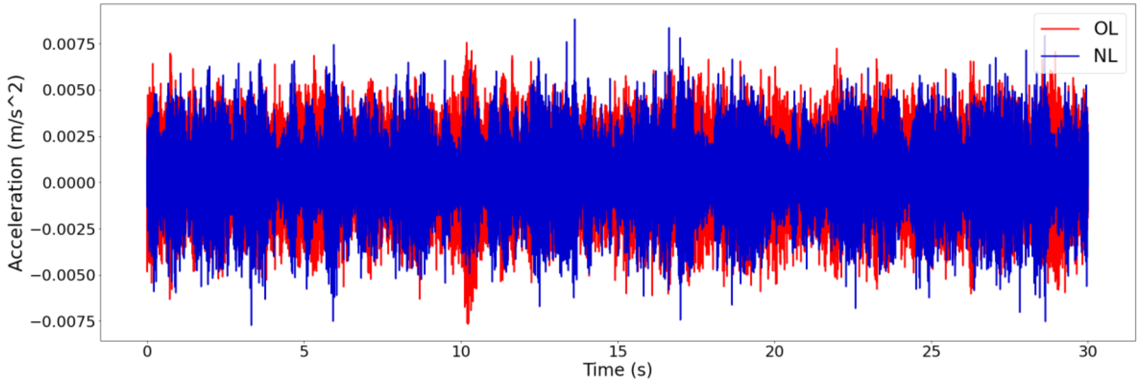

Figure 2.45. Time-domain plots of OL vs. NL signals in the looped network measured by sensor A2

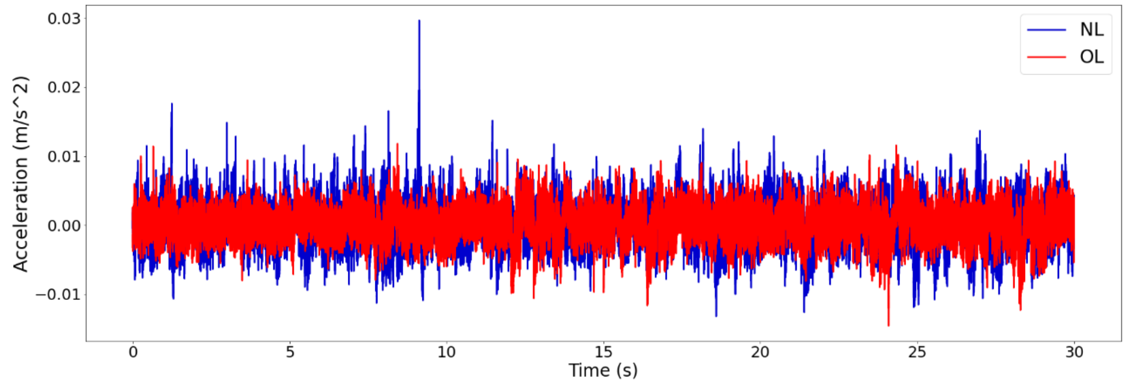

Figure 2.46. Time-domain plots of OL vs. NL signals in the branched network measured by sensor A2

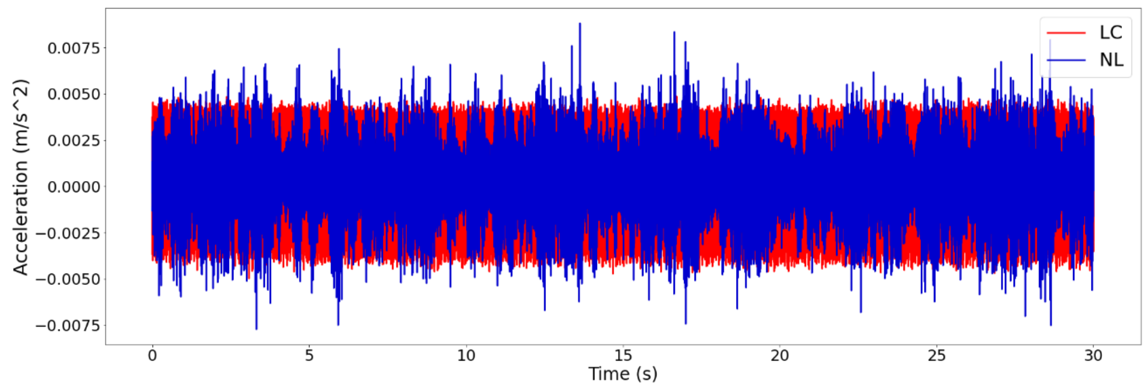

Figure 2.47. Time-domain plots of LC vs. NL signals in the looped network measured by sensor A2

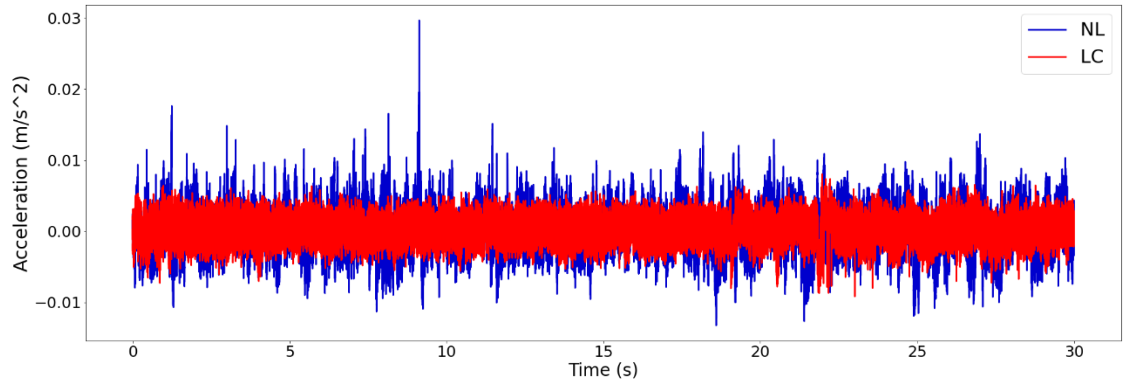

Figure 2.48. Time-domain plots of LC vs. NL signals in the branched network measured by sensor A2

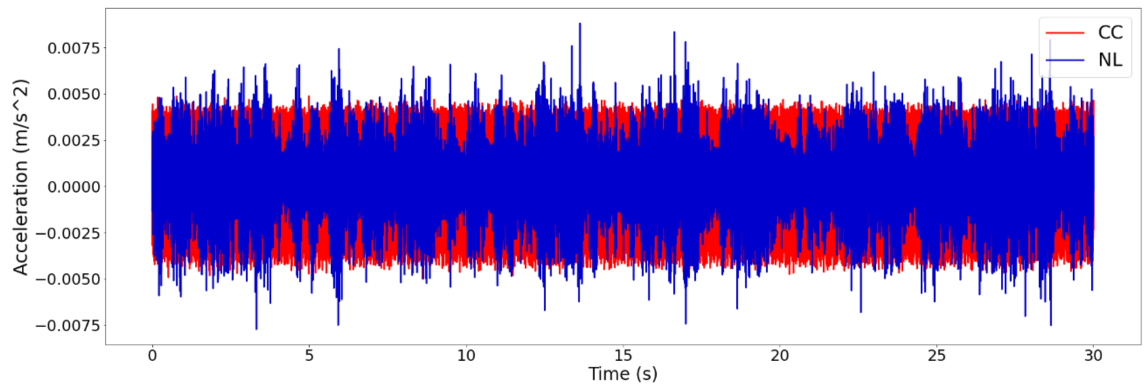

Figure 2.49. Time-domain plots of CC vs. NL signals in the looped network measured by sensor A2

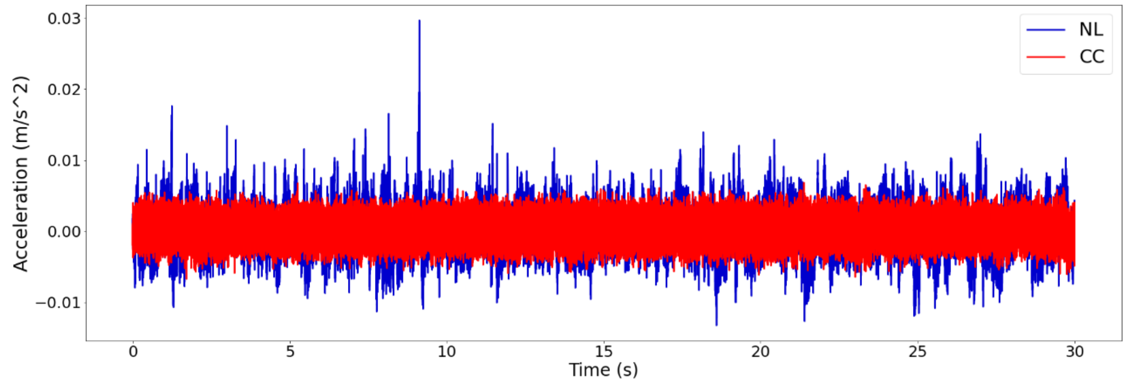

Figure 2.50. Time-domain plots of CC vs. NL signals in the branched network measured by sensor A2

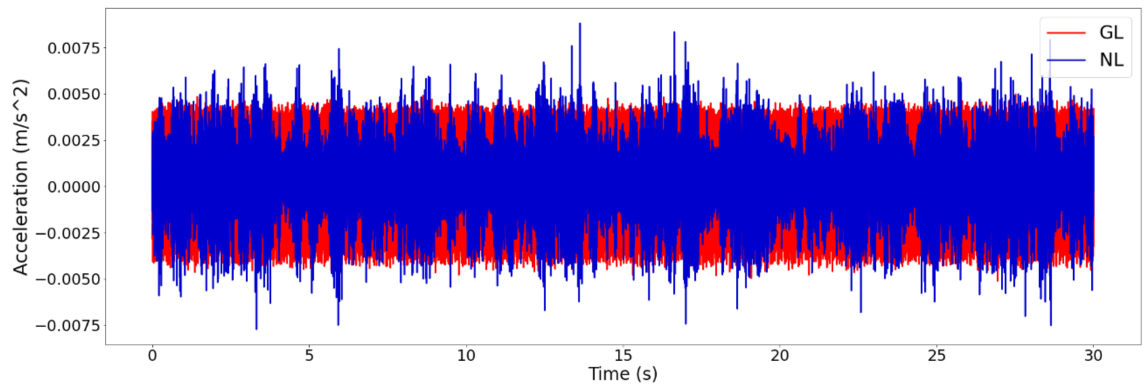

Figure 2.51. Time-domain plots of GL vs. NL signals in the looped network measured by sensor A2

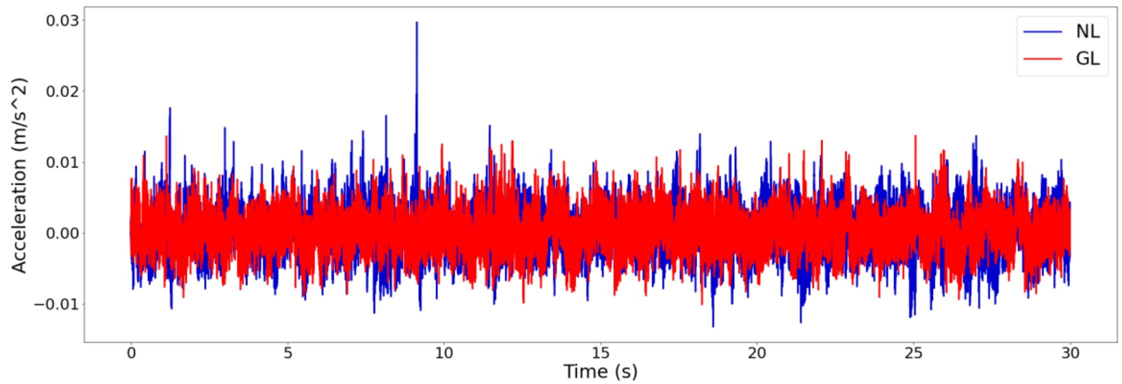

Figure 2.52. Time-domain plots of GL vs. NL signals in the branched network measured by sensor A2

Table 2.25 includes analytical information of the time-domain acceleration plots where leak and no leak signals of sensors A1 and A2 in the looped and branched networks are visually compared.

Table 2.25. Analysis of the time-domain acceleration sensor plots measured by A1 and A2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Time-domain plot (for ND signal)	NL vs. OL – A1	- NL and OL signal amplitudes are larger in the looped network than those in the branched one. - NL signal amplitudes are larger than those of OL in the looped network. - OL signal amplitudes are larger than those of NL in the branched network.

Table 2.25. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Time-domain plot (for ND signal)	NL vs. LC – A1	- NL signal amplitudes are larger in the looped network than those in the branched one. - LC signal amplitudes are smaller in the looped network than those in the branched one. - NL signal amplitudes are much larger than those of LC in the looped network. - LC signal amplitudes are larger than those of NL in the branched network. - LC signal amplitudes are close to zero in the looped network.
	NL vs. CC – A1	- NL signal amplitudes are larger in the looped network than those in the branched one. - CC signal amplitudes are smaller in the looped network than those in the branched one. - NL signal amplitudes are much larger than those of CC in the looped network. - CC signal amplitudes are larger than those of NL in the branched network. - CC signal amplitudes are close to zero in the looped network.
	NL vs. GL – A1	- NL signal amplitudes are larger in the looped network than those in the branched one. - GL signal amplitudes are much smaller in the looped network than those in the branched one. - NL signal amplitudes are much larger than those of GL in the looped network. - GL signal amplitudes are twice those of NL in the branched network. - GL signal amplitudes are close to zero in the looped network.
	All – A1	- When there is no demand, and in the looped network, NL signals' amplitudes are larger than leak signals' amplitudes; however, in the branched network, leak signals' amplitudes are larger than those of NL.
	NL vs. OL – A2	- Signal amplitudes are approximately the same in both networks. - NL signal in the branched network includes more outliers.
	NL vs. LC – A2	- NL and LC signal amplitudes are approximately the same in the looped network. - NL signal amplitudes are larger than those of LC in the branched network. - NL signal in the branched network includes more outliers. - LC signal amplitudes are more consistent than those of NL in both networks.

Table 2.25. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Time-domain plot (for ND signal)	NL vs. CC – A2	- NL and CC signal amplitudes are approximately the same in the looped network. - NL signal amplitudes are larger than those of CC in the branched network. - NL signal in the branched network includes more outliers. - CC signal amplitudes are more consistent than those of NL in both networks.
	NL vs. GL – A2	- NL and GL signal amplitudes are approximately the same in the looped network. - NL and GL signal amplitudes are approximately the same in the branched network. - NL and GL signal amplitudes in the looped network are smaller than those in the branched network. - NL signal in the branched network includes more outliers. - GL signal amplitudes are more consistent than those of NL in the looped network.
	All – A2	- When there is no demand, and in the branched network, except for the GL, NL signal amplitudes are larger than those of leak signals.

2.12.3.2. Frequency-domain plot (for ND signals)

Figures 2.53 to 2.60 show the frequency-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor A1.

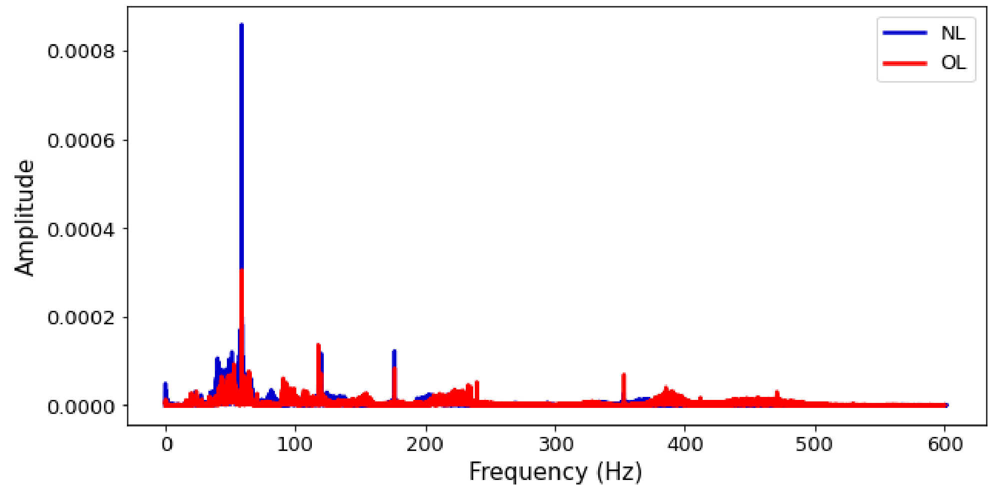

Figure 2.53. Frequency-domain plots of OL vs. NL signals in the looped network measured by sensor A1

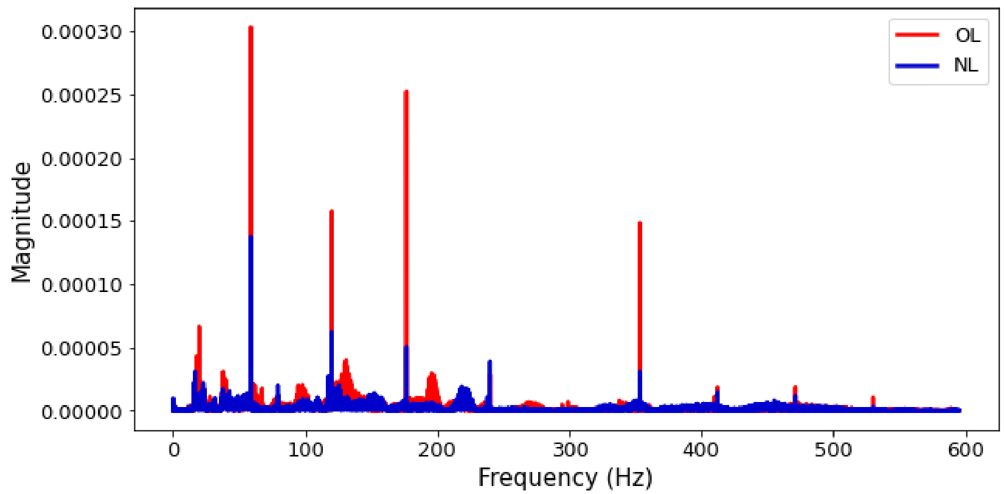

Figure 2.54. Frequency-domain plots of OL vs. NL signals in the branched network measured by sensor A1

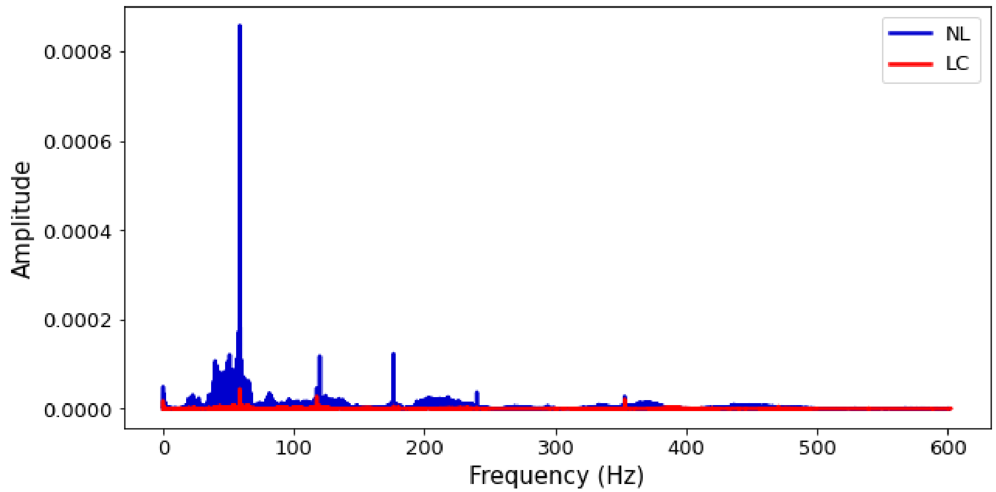

Figure 2.55. Frequency-domain plots of LC vs. NL signals in the looped network measured by sensor A1

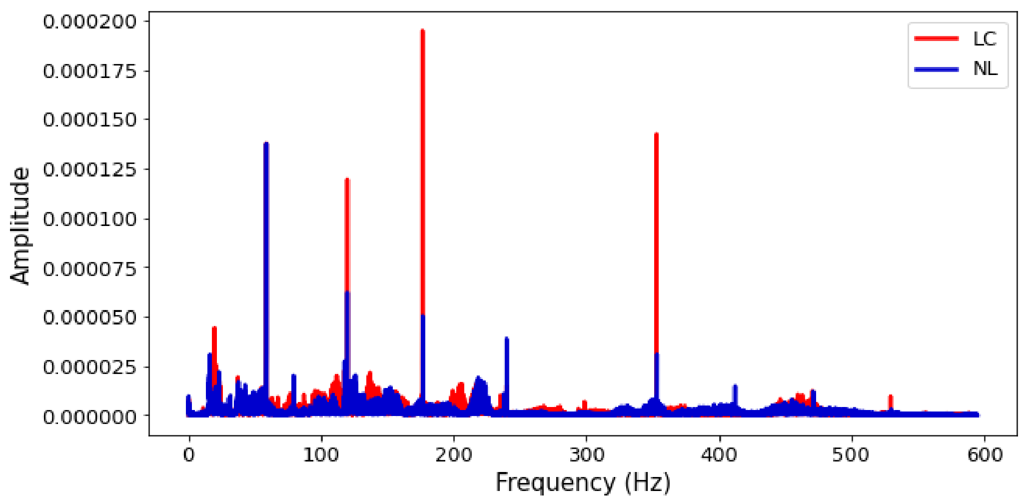

Figure 2.56. Frequency-domain plots of LC vs. NL signals in the branched network measured by sensor A1

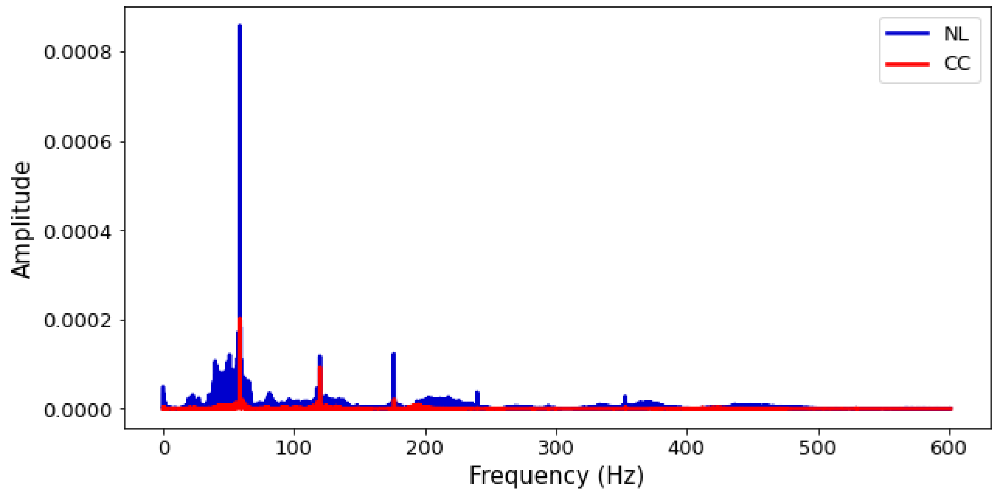

Figure 2.57. Frequency-domain plots of CC vs. NL signals in the looped network measured by sensor A1

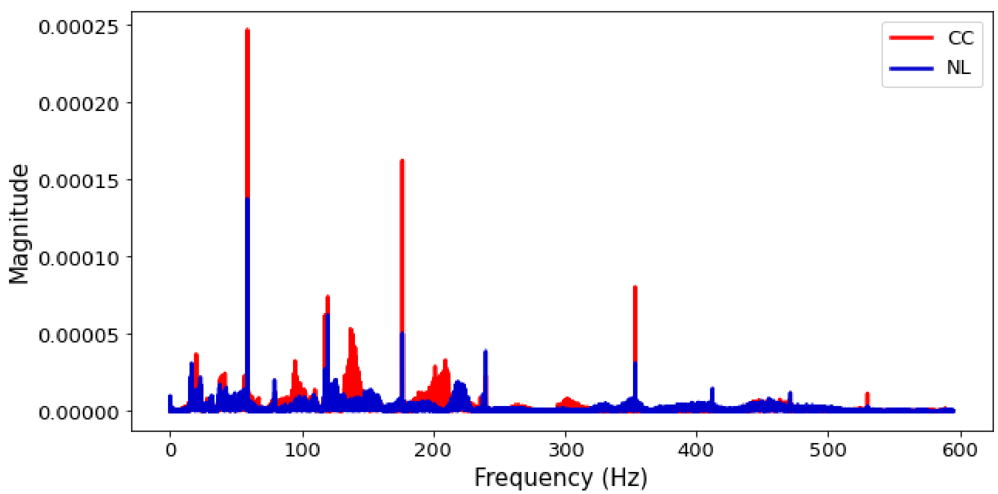

Figure 2.58. Frequency-domain plots of CC vs. NL signals in the branched network measured by sensor A1

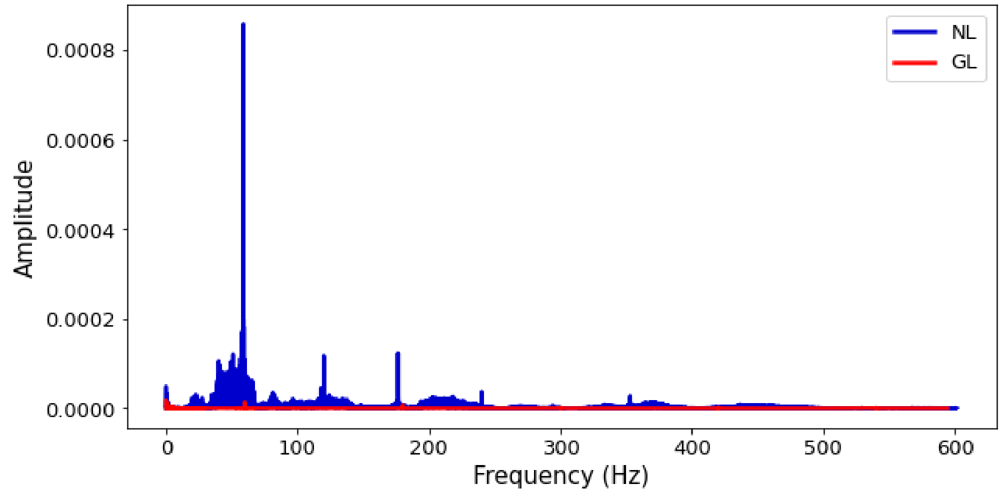

Figure 2.59. Frequency-domain plots of GL vs. NL signals in the looped network measured by sensor A1

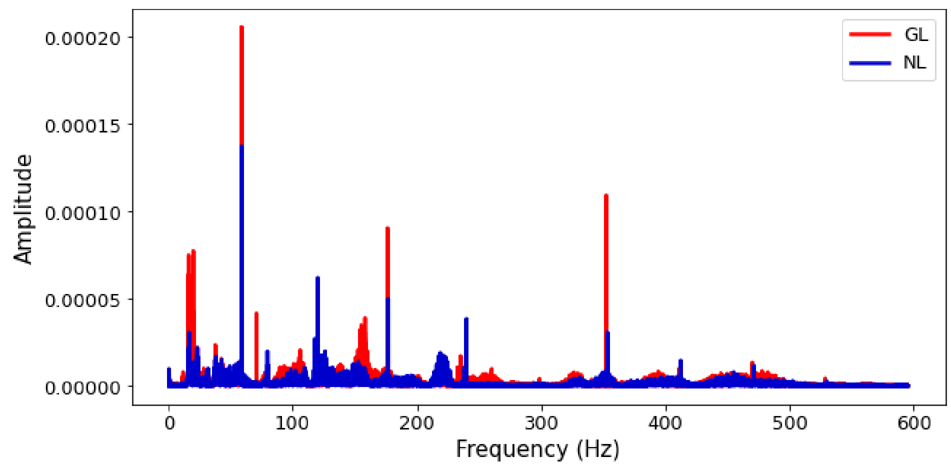

Figure 2.60. Frequency-domain plots of GL vs. NL signals in the branched network measured by sensor A1

Figures 2.61 to 2.68 show the Frequency-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor A2.

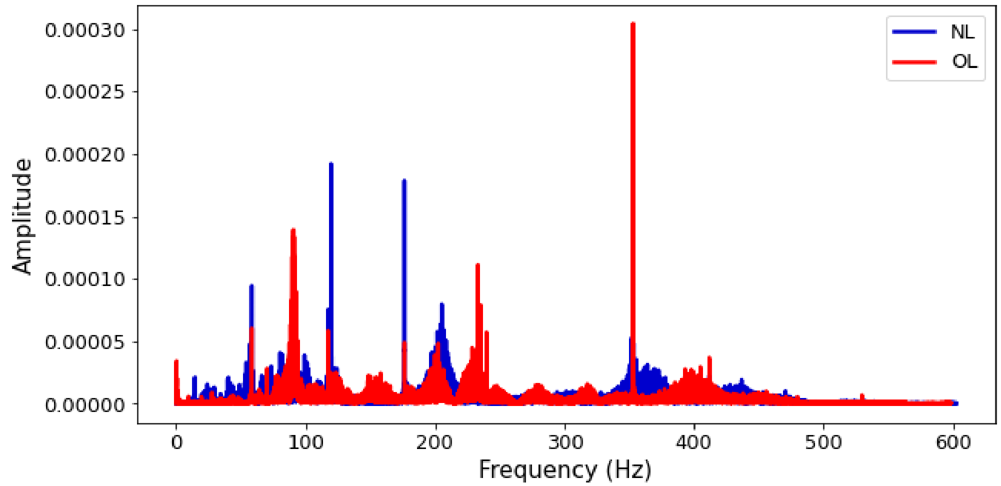

Figure 2.61. Frequency-domain plots of OL vs. NL signals in the looped network measured by sensor A2

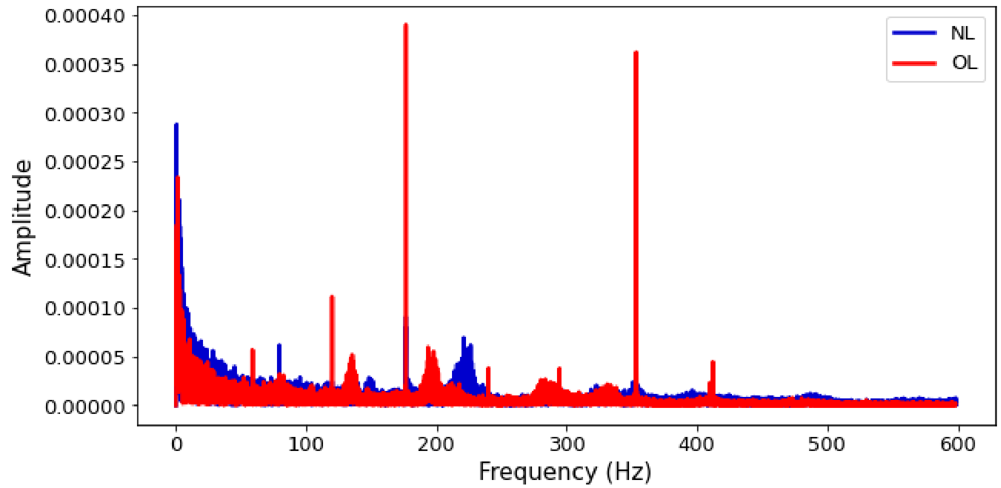

Figure 2.62. Frequency-domain plots of OL vs. NL signals in the branched network measured by sensor A2

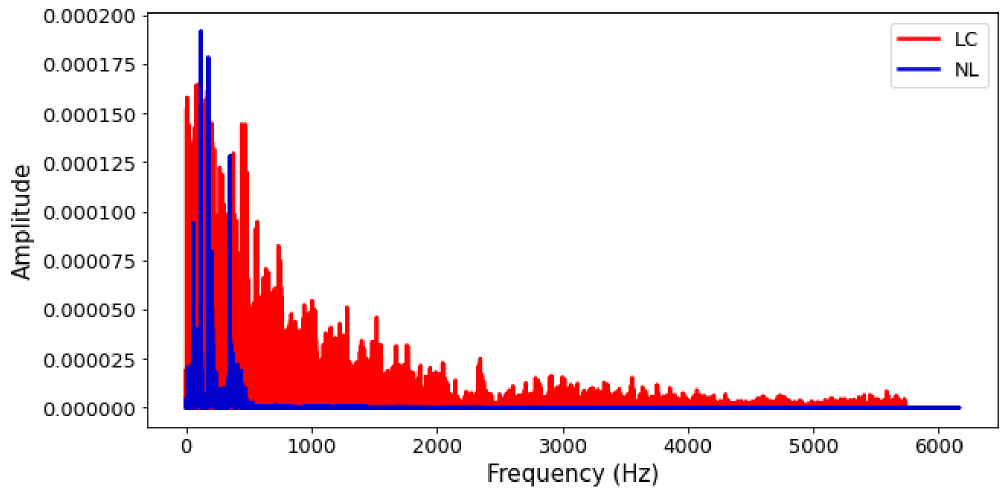

Figure 2.63. Frequency-domain plots of LC vs. NL signals in the looped network measured by sensor A2

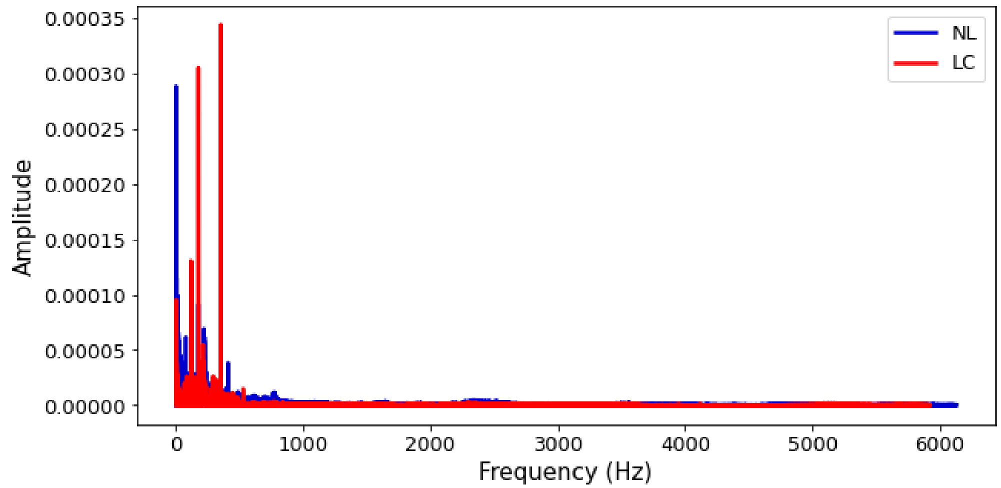

Figure 2.64. Frequency-domain plots of LC vs. NL signals in the branched network measured by sensor A2

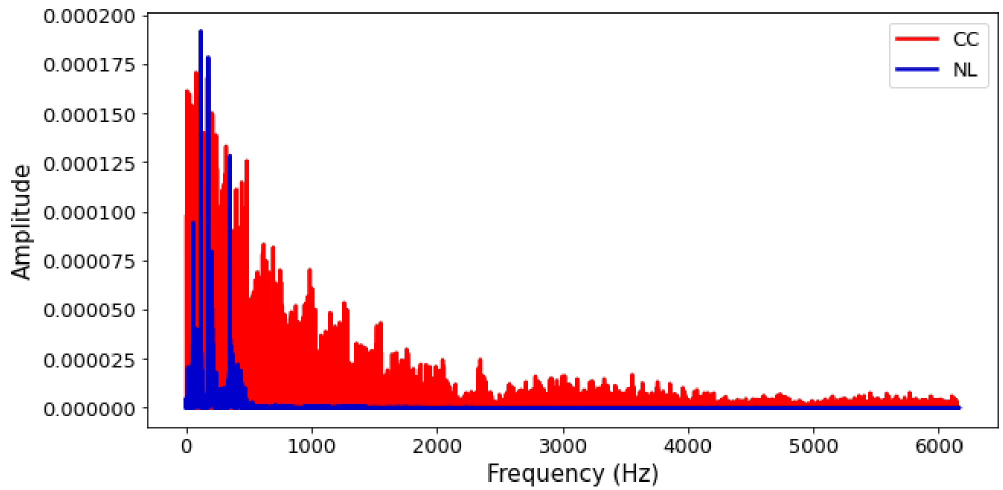

Figure 2.65. Frequency-domain plots of CC vs. NL signals in the looped network measured by sensor A2

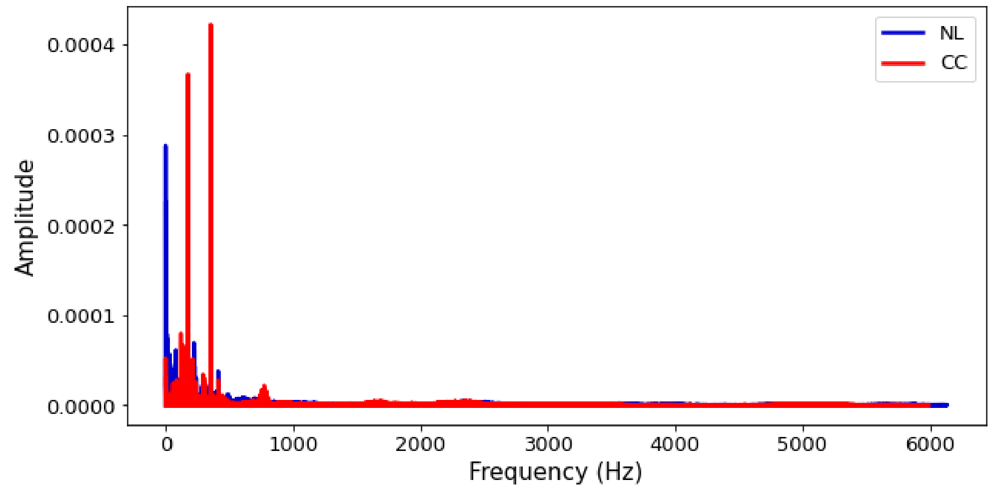

Figure 2.66. Frequency-domain plots of CC vs. NL signals in the branched network measured by sensor A2

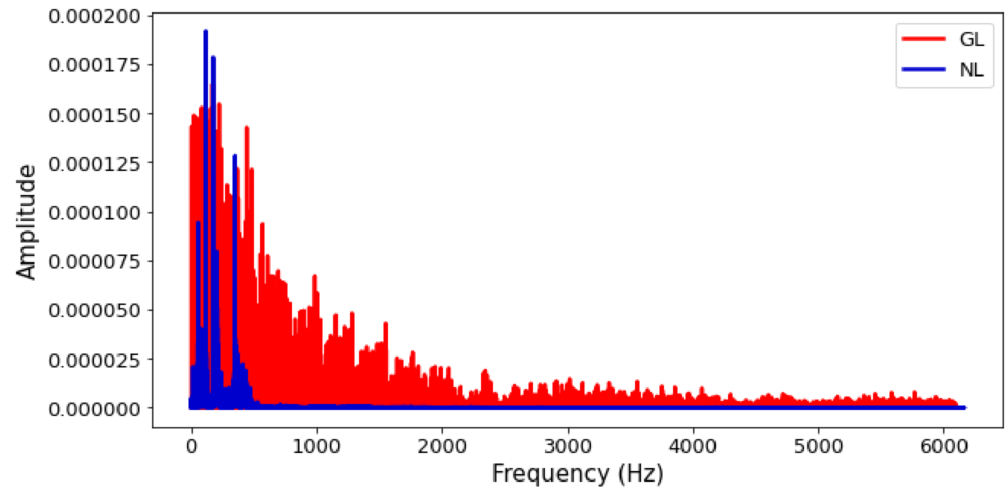

Figure 2.67. Frequency-domain plots of GL vs. NL signals in the looped network measured by sensor A2

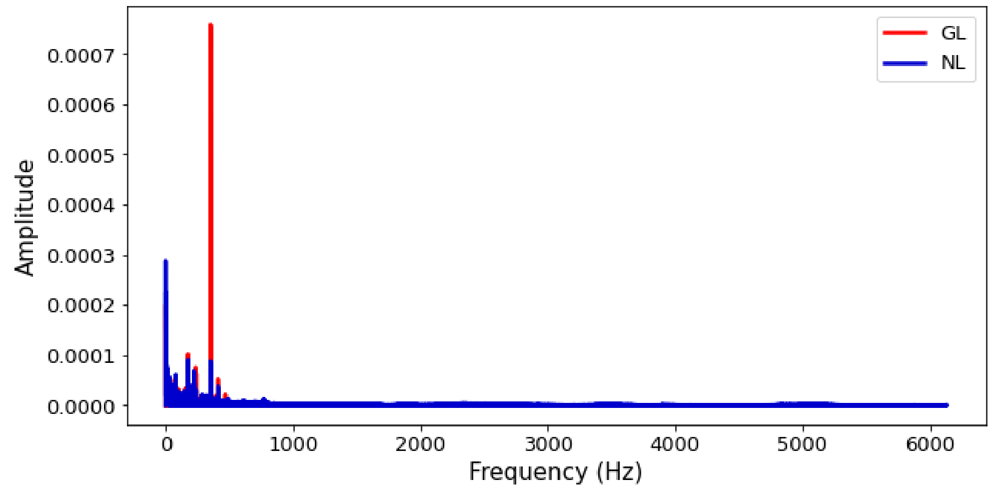

Figure 2.68. Frequency-domain plots of GL vs. NL signals in the branched network measured by sensor A2

Table 2.26 includes analytical information of the frequency-domain acceleration plots where leak and no leak signals of sensors A1 and A2 in the looped and branched networks are visually compared.

Table 2.26. Analysis of the frequency-domain acceleration sensor plots measured by A1 and A2

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Frequency-domain plot (for ND signal)	NL vs. OL – A1	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 500 Hz. - NL signal's dominant frequency has a larger amplitude in the looped network than in the branched one. - For OL signal in the looped network, frequencies with non-zero amplitudes are less than 500 Hz. - For OL signal in the branched network, frequencies with non-zero amplitudes are less than 600 Hz. - OL signal’s frequencies in the branched network are more pronounced than in the looped network. - In the looped network, frequencies of the OL signal have smaller amplitudes than the NL signal. - In the branched network, frequencies of the OL signal have larger amplitudes than the NL signal.
	NL vs. LC – A1	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 500 Hz. - NL signal's dominant frequency has a larger amplitude in the looped network than in the branched one. - For LC signal in the looped network, frequencies with non-zero amplitudes are less than 400 Hz. - For LC signal in the branched network, frequencies with non-zero amplitudes are less than 600 Hz. - LC signal’s frequencies in the branched network are much more pronounced than in the looped network. - In the looped network, frequencies of the LC signal have smaller amplitudes than the NL signal. - In the branched network, frequencies of the LC signal have larger amplitudes than the NL signal.
	NL vs. CC – A1	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 500 Hz. - NL signal's dominant frequency has a larger amplitude in the looped network than in the branched one. - For CC signal in the looped network, frequencies with non-zero amplitudes are less than 200 Hz. - For CC signal in the branched network, frequencies with non-zero amplitudes are less than 600 Hz. - In the looped network, frequencies of the CC signal have smaller amplitudes than the NL signal. - In the branched network, frequencies of the CC signal have larger amplitudes than the NL signal.

Table 2.26. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Frequency-domain plot (for ND signal)	NL vs. GL – A1	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 500 Hz. - NL signal's dominant frequency has a larger amplitude in the looped network than in the branched one. - For GL signal in the looped network, frequencies with non-zero amplitudes are less than 100 Hz. - For GL signal in the branched network, frequencies with non-zero amplitudes are less than 500 Hz. - In the looped network, frequencies of the GL signal have much smaller amplitudes than the NL signal. - In the branched network, frequencies of the GL signal have larger amplitudes than the NL signal. - In the looped network, the amplitudes of GL signal frequencies are much smaller than those of the branched network.
	All – A1	- For NL signal with no demand in both networks, frequencies with non-zero amplitudes are less than 500 Hz. - With no demand and in the looped network, frequencies of leak signals with non-zero amplitudes are less than 500 Hz. While those in the branched network are less than 600 Hz. - With no demand, leak signals' frequency caps in the looped network < leak signals' frequency caps in the branched network. - With no demand and in the looped network, order of leak signals' frequency caps with non-zero amplitudes are OL > LC > CC > GL. Excluding GL that has a different leak mechanism in this study, this order is aligned with the leaks' flow rates in Table 2.6. In Table 2.6, leaks that have a jet water flow mechanism, the order of leak flow rates are OL > LC > CC. - With no demand and in the looped network: amplitudes of the leaks' frequencies < amplitudes of NLs' frequencies. - With no demand and in the branched network: amplitudes of the leaks' frequencies > amplitudes of NLs' frequencies.

Table 2.26. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Frequency-domain plot (for ND signal)	NL vs. OL – A2	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 600 Hz. - NL signal's dominant frequency is larger in the looped network than in the branched one. - For OL signal in the looped network, frequencies with non-zero amplitudes are less than 500 Hz. - For OL signal in the branched network, frequencies with non-zero amplitudes are less than 600 Hz. - OL signal’s frequencies in the branched network are more pronounced than in the looped network. - In both networks, frequencies of the OL and NL signals have approximately the same amplitudes.
	NL vs. LC – A2	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 600 Hz. - NL signal's dominant frequency is larger in the looped network than in the branched one. - For LC signal in the looped network, frequencies with non-zero amplitudes are less than 6000 Hz. - For LC signal in the branched network, frequencies with non-zero amplitudes are less than 600 Hz. - OL signal’s frequencies in the branched network are less pronounced than in the looped network. - In the looped network, LC signal’s frequencies have larger amplitudes than NL signal. - In the branched network, NL signal’s frequencies have larger amplitudes than LC signal.
	NL vs. CC – A2	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 600 Hz. - NL signal's dominant frequency is larger in the looped network than in the branched one. - For CC signal in the looped network, frequencies with non-zero amplitudes are less than 6000 Hz. - For CC signal in the branched network, frequencies with non-zero amplitudes are less than 600 Hz. - CC signal’s frequencies in the branched network are less pronounced than in the looped network. - In the looped network, LC signal’s frequencies have larger amplitudes than NL signal. - In the branched network, NL signal’s frequencies have larger amplitudes than LC signal.

Table 2.26. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Frequency-domain plot (for ND signal)	NL vs. GL – A2	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 600 Hz. - NL signal's dominant frequency is larger in the looped network than in the branched one. - For GL signal in the looped network, frequencies with non-zero amplitudes are less than 6000 Hz. - For GL signal in the branched network, frequencies with non-zero amplitudes are less than 600 Hz. - GL signal's frequencies in the branched network are less pronounced than in the looped network. - In the looped network, LC signal's frequencies have larger amplitudes than NL signal. - In the branched network, the amplitudes of NL and GL signals' frequencies have approximately the same magnitudes.
	All – A2	- For NL signal with no demand in both networks, frequencies with non-zero amplitudes are less than 600 Hz. - With no demand and in the looped network, order of leak signals' frequency caps with non-zero amplitudes are LC = CC = GL = 6000 Hz >> OL = 600 Hz. - With no demand in the branched network, all leak signals' frequency caps are smaller than 600 Hz. - With no demand, leak signals' frequency caps in the looped network ≥ leak signals' frequency caps in the branched network.

2.12.3.3. Cumulative Distribution Plot (for ND signal)

Figures 2.69 and 2.70 show the cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped and branched networks, respectively, measured by sensor A1 with no demand. Also, Figures 2.71 and 2.72 show the cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped and branched networks, respectively, measured by sensor A2 with no demand.

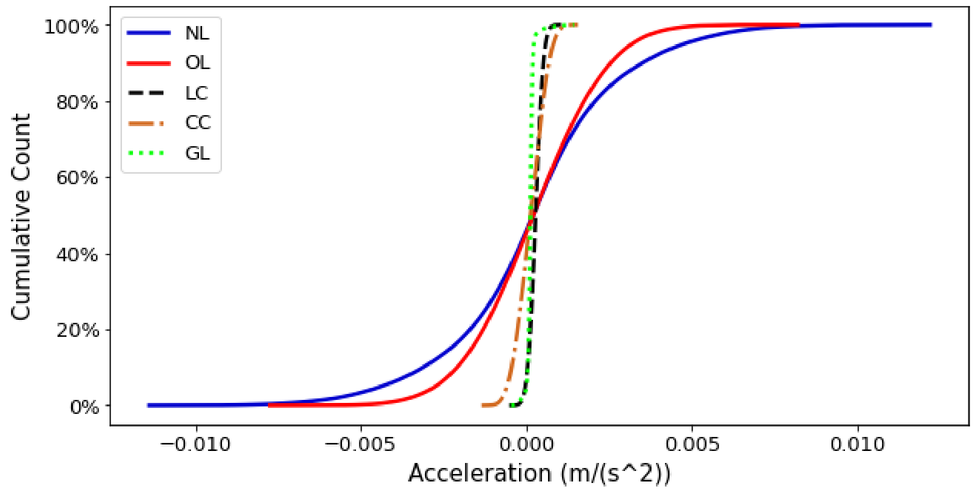

Figure 2.69. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor A1

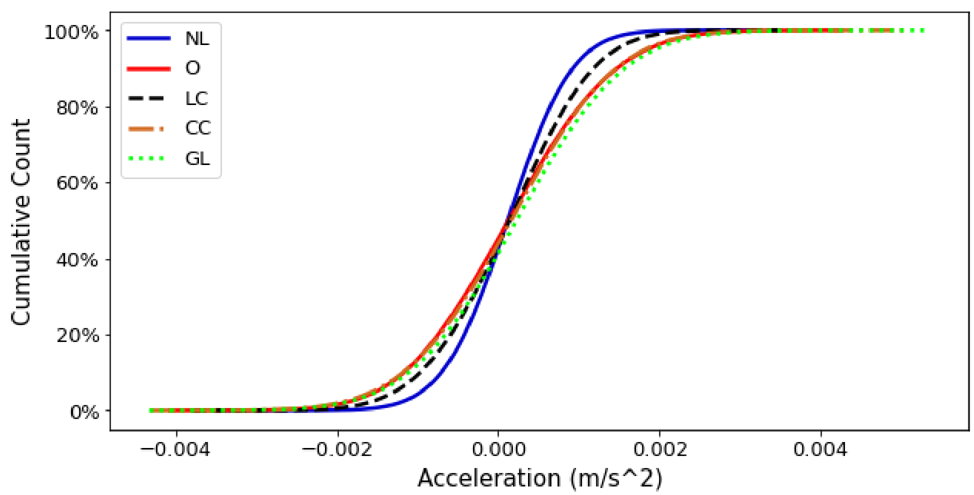

Figure 2.70. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the branched network, measured by sensor A1

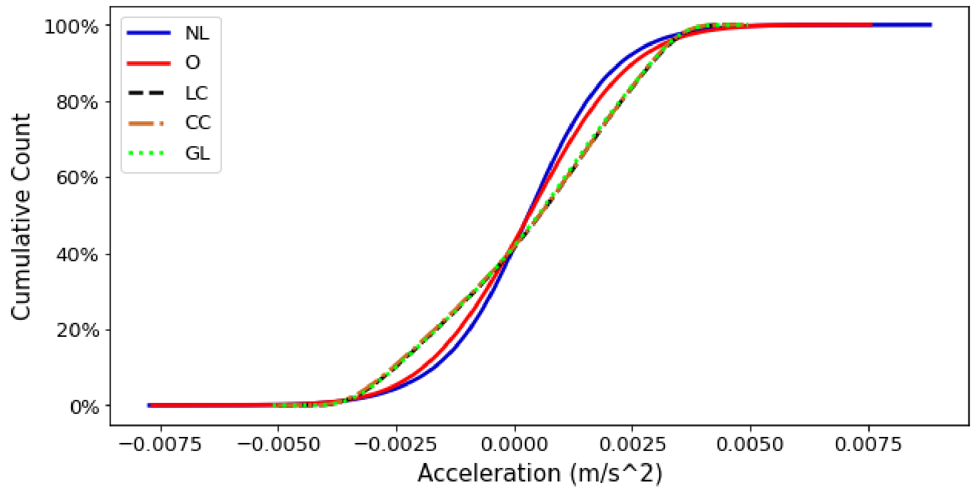

Figure 2.71. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor A2

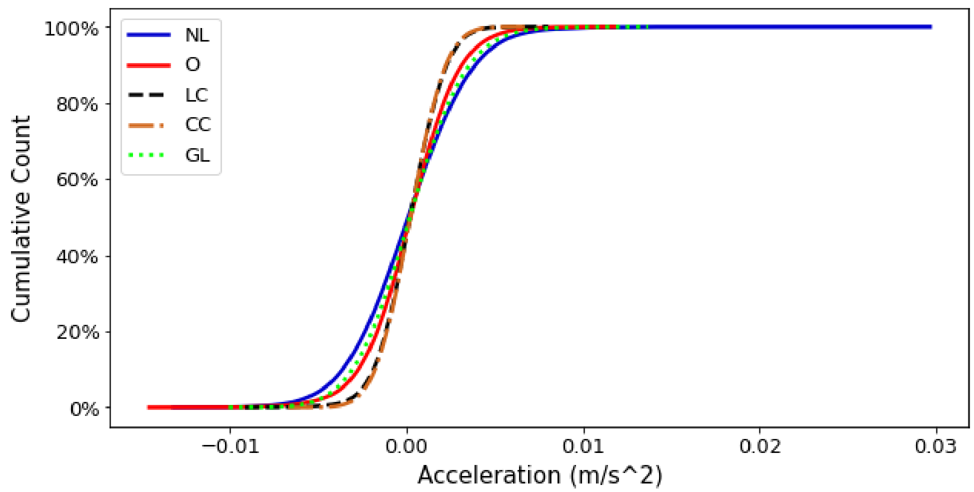

Figure 2.72. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor A2

Table 2.27 includes the analysis of the cumulative distribution plots where leak and no leak signals of sensors A1 and A2 in the looped and branched networks are visually compared.

Table 2.27. Analysis of the acceleration data's cumulative distribution plots measured by A1 and A2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Cumulative distribution plot (for ND signal)	NL – A1	- NL acceleration magnitudes are larger in the looped network than in the branched network. - NL cumulative distribution includes the widest acceleration range in the looped network. - NL acceleration magnitudes are larger than all leaks' acceleration magnitudes in the looped network. - NL acceleration magnitudes are smaller than all leaks' acceleration magnitudes in the branched network.
	OL – A1	- OL cumulative distribution includes the widest acceleration range among all leaks in the looped network. - OL acceleration magnitudes are larger than other leaks' acceleration magnitudes in the looped network. - OL acceleration magnitudes are larger in the looped network than in the branched network.
	LC – A1	- LC acceleration magnitudes are larger in the branched network than in the looped network. - LC, CC, and GL have smaller values than NL and OL in the looped network. - LC, CC, and GL acceleration magnitudes are very small and close to zero in the looped network. - Order of signal magnitude in the looped network: NL > OL > CC > LC > GL. - Order of signal magnitude in the branched network: GL > CC > OL > LC > NL.
	CC – A1	- CC acceleration magnitudes are larger in the branched network than in the looped network. - LC, CC, and GL have smaller values than NL and OL in the looped network. - LC, CC, and GL acceleration magnitudes are very small and close to zero in the looped network. - Order of signal magnitude in the looped network: NL > OL > CC > LC > GL. - Order of signal magnitude in the branched network: GL > CC > OL > LC > NL.

Table 2.27. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Cumulative distribution plot (for ND signal)	GL – A1	- GL acceleration magnitudes are larger in the branched network than in the looped network. - LC, CC, and GL have smaller values than NL and OL in the looped network. - LC, CC, and GL acceleration magnitudes are very small and close to zero in the looped network. - GL acceleration magnitudes are smaller than other leaks' acceleration magnitudes in the looped network. - GL acceleration magnitudes are larger than other leaks' acceleration magnitudes in the branched network. - Order of signal magnitude in the looped network: NL > OL > CC > LC > GL. - Order of signal magnitude in the branched network: GL > CC > OL > LC > NL.
	All – A1	- NL acceleration magnitudes are larger in the looped network than in the branched network. - Order of signal magnitude in the looped network: NL > OL > CC > LC > GL. This order conforms to the amplitude order of time-series plots for signals with no demand in the looped network. - Order of signal magnitude in the branched network: GL > CC > OL > LC > NL. This order conforms to the amplitude order of time-series plots for signals with no demand in the branched network.
	NL – A2	- Contrary to A1, NL acceleration magnitudes are larger in the branched network than in the looped network. - Contrary to A1, NL acceleration magnitudes are smaller than all leaks' acceleration magnitudes in the looped network. - Contrary to A1, NL acceleration magnitudes are larger than all leaks' acceleration magnitudes in the branched network. - Order of signal magnitude in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude in the branched network: NL > GL > OL > LC > CC.

Table 2.27. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Cumulative distribution plot (for ND signal)	OL – A2	- OL cumulative distribution includes the widest acceleration range among all leaks in the looped network. - Contrary to A1, OL acceleration magnitudes are larger in the branched network than in the looped network. - Contrary to A1, OL acceleration magnitudes are smaller than other leaks' acceleration magnitudes in the looped network. - Order of signal magnitude in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude in the branched network: NL > GL > OL > LC > CC.
	LC – A2	- Contrary to A1, LC, CC, and GL have larger values than NL and OL in the looped network. - Order of signal magnitude in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude in the branched network: NL > GL > OL > LC > CC.
	CC – A2	- Contrary to A1, LC, CC, and GL have larger values than NL and OL in the looped network. - CC acceleration magnitudes have the smallest magnitudes than NL and other leaks in the branched network. - Order of signal magnitude in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude in the branched network: NL > GL > OL > LC > CC.
	GL – A2	- GL acceleration magnitudes are larger than other leaks' acceleration magnitudes in the branched network. - Order of signal magnitude in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude in the branched network: NL > GL > OL > LC > CC.
	All – A2	- Order of signal magnitude in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude in the branched network: NL > GL > OL > LC > CC.

2.12.3.4. Box Plot (for ND signal)

Figures 2.73 and 2.74 show box plots of the NL, OL, LC, CC, and GL signals in the looped and branched networks, respectively, measured by sensor A1 with no demand. Figures 2.75 and 2.76 show the same plots for signals measured by sensor A2 with no demand.

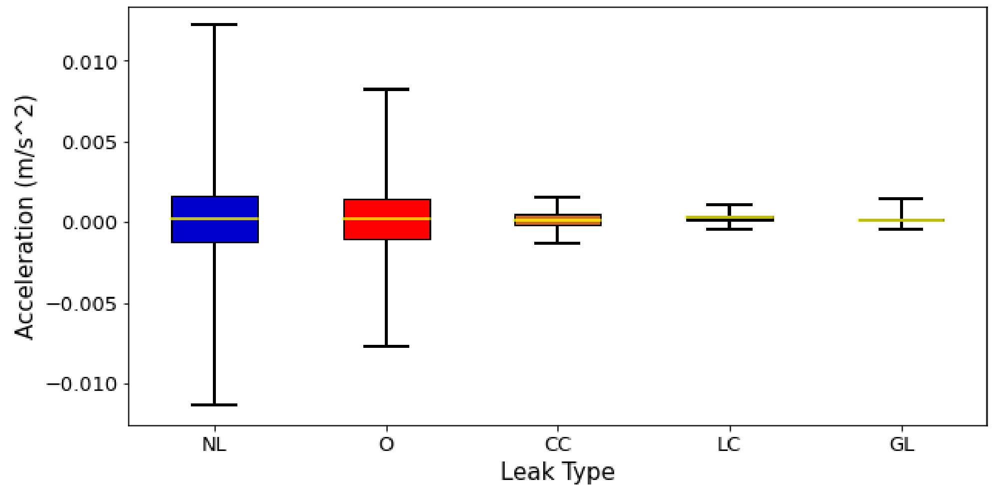

Figure 2.73. Box plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor A1

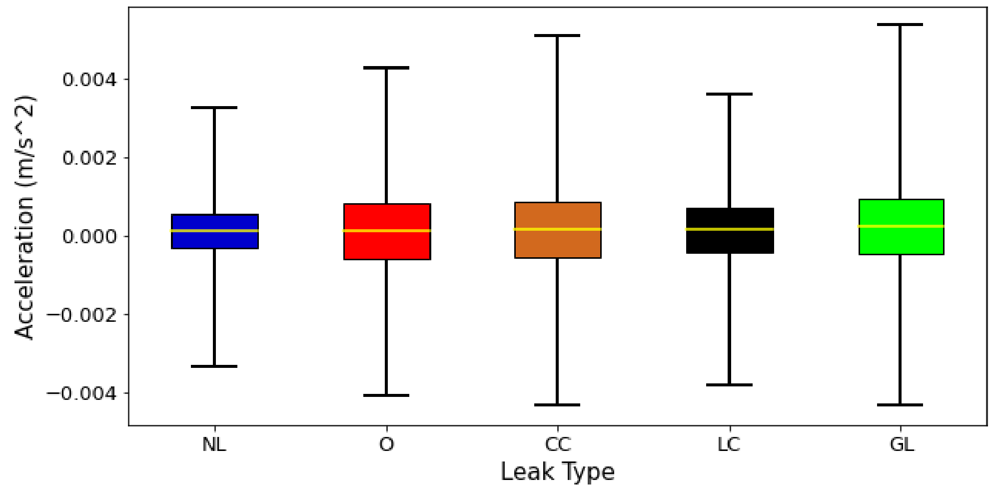

Figure 2.74. Box plots of the NL, OL, LC, CC, and GL signals in the branched network, measured by sensor A1

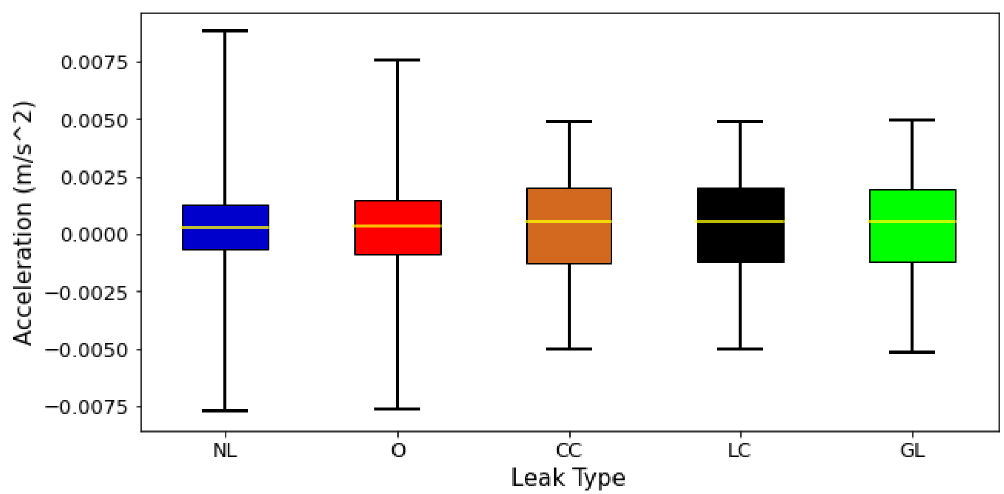

Figure 2.75. Box plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor A2

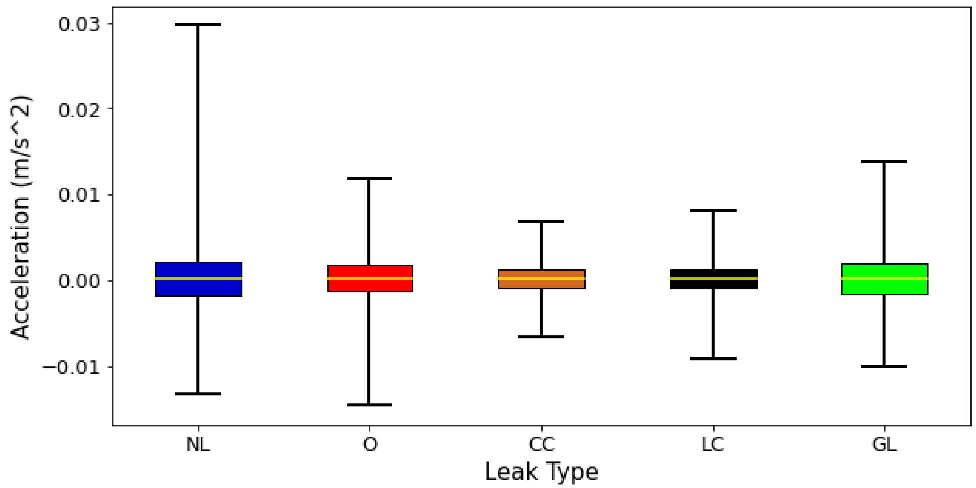

Figure 2.76. Box plots of the NL, OL, LC, CC, and GL signals in the branched network, measured by sensor A2

Table 2.28 includes the analysis of the box plots where leak and no leak signals of sensors A1 and A2 in the looped and branched networks are visually compared.

Table 2.28. Analysis of acceleration data's box plots measured by A1 and A2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Box plot (for ND signal)	NL – A1	- NL signal of the looped network includes the largest range of acceleration magnitudes compared to all signals of the looped and branched networks. - Order of signal magnitude continuum in the looped network: NL > OL > CC > LC > GL. - Order of signal magnitude continuum in the branched network: GL > CC > OL > LC > NL.
	OL – A1	- Comparison of OL signal magnitude continuum: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$ - Order of signal magnitude continuum in the looped network: NL > OL > CC > LC > GL. - Order of signal magnitude continuum in the branched network: GL > CC > OL > LC > NL.
	LC – A1	- Comparison of LC signal magnitude continuum: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$ - Order of signal magnitude continuum in the looped network: NL > OL > CC > LC > GL. - Order of signal magnitude continuum in the branched network: GL > CC > OL > LC > NL.
	CC – A1	- Comparison of CC signal magnitude continuum: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$ - Order of signal magnitude continuum in the looped network: NL > OL > CC > LC > GL. - Order of signal magnitude continuum in the branched network: GL > CC > OL > LC > NL.
	GL – A1	- Comparison of GL signal magnitude continuum: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$ - Order of signal magnitude continuum in the looped network: NL > OL > CC > LC > GL. - Order of signal magnitude continuum in the branched network: GL > CC > OL > LC > NL.
	All – A1	- NL signal of the looped network includes the largest range of acceleration magnitudes compared to all other signals of the looped and branched networks. - Order of signal magnitude continuum in the looped network: NL > OL > CC > LC > GL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot. - Order of signal magnitude continuum in the branched network: GL > CC > OL > LC > NL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot.

Table 2.28. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Box plot (for ND signal)	NL – A2	- Comparison of NL signal magnitude continuum: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$ - Order of signal magnitude continuum in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude continuum in the branched network: NL > GL > OL > LC > CC.
	OL – A2	- Comparison of NL signal magnitude continuum: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$ - Order of signal magnitude continuum in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude continuum in the branched network: NL > GL > OL > LC > CC.
	LC – A2	- Comparison of LC signal magnitude continuum: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$ - Order of signal magnitude continuum in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude continuum in the branched network: NL > GL > OL > LC > CC.
	CC – A2	- Comparison of CC signal magnitude continuum: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$ - Order of signal magnitude continuum in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude continuum in the branched network: NL > GL > OL > LC > CC.
	GL – A2	- Comparison of GL signal magnitude continuum: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$ - Order of signal magnitude continuum in the looped network: CC > LC > GL > OL > NL. - Order of signal magnitude continuum in the branched network: NL > GL > OL > LC > CC.
	All – A2	- There is no consistent pattern in the relations of the two networks' signal magnitude continuum. - Order of signal magnitude continuum in the looped network: CC > LC > GL > OL > NL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot. - Order of signal magnitude continuum in the branched network: NL > GL > OL > LC > CC. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot.

2.12.3.5. Cross Spectral Plot (for ND signal)

Figures 2.77 and 2.78 show the cross spectral plots of the NL, OL, LC, CC, and GL signals with no demands in the looped and branched networks, respectively.

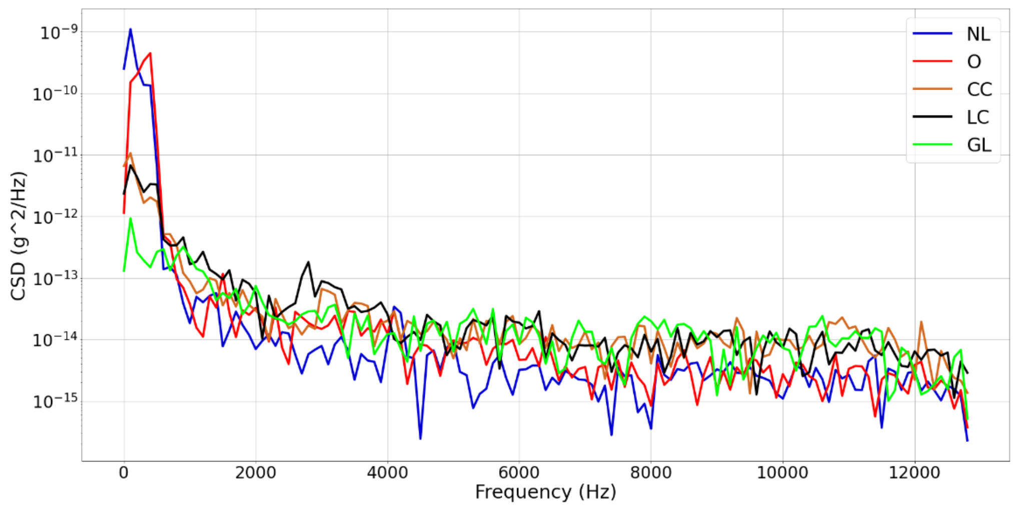

Figure 2.77. Cross spectral plots of the NL, OL, LC, CC, and GL signals in the looped network for accelerometers

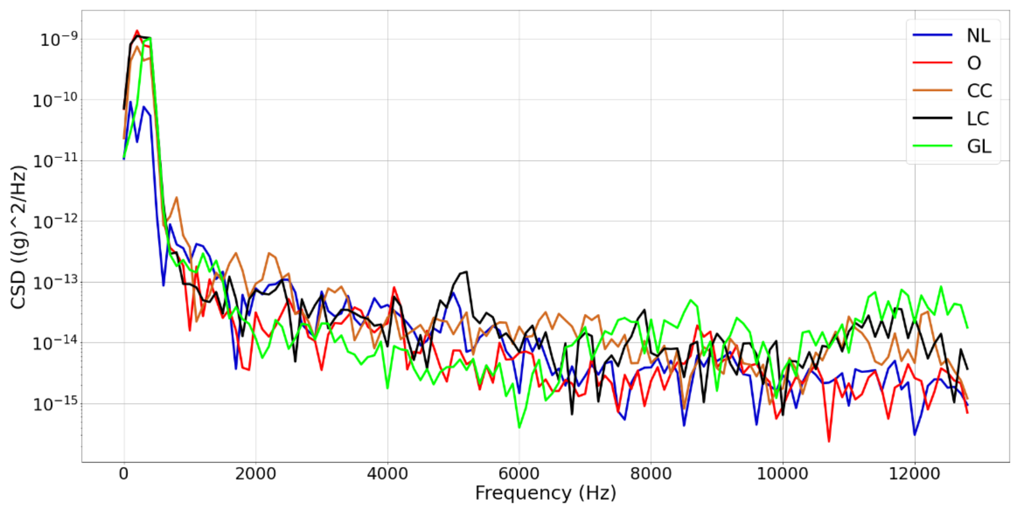

Figure 2.78. Cross spectral plots of the NL, OL, LC, CC, and GL signals in the branched network for accelerometers

Table 2.29 compares the area under the cross spectral plots of the looped and branched networks with no demand.

Table 2.29. Comparison of the area under the cross spectral plots of accelerometer data

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Area under the cross-spectral plot (for ND signal)	NL	- Comparison of NL signal's areas under the CSD plot: $NL_{lo} > NL_{br}$ - Order of areas under the CSD plot in the looped network: NL > OL > CC > LC > GL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.
	OL	- Comparison of OL signal's areas under the CSD plot: $OL_{lo} < OL_{br}$ - Order of areas under the CSD plot in the looped network: NL > OL > CC > LC > GL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.
	LC	- Comparison of LC signal's areas under the CSD plot: $LC_{lo} < LC_{br}$ - Order of areas under the CSD plot in the looped network: NL > OL > CC > LC > GL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.
	CC	- Comparison of CC signal's areas under the CSD plot: $CC_{lo} < CC_{br}$ - Order of areas under the CSD plot in the looped network: NL > OL > CC > LC > GL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.
	GL	- Comparison of GL signal's areas under the CSD plot: $GL_{lo} < GL_{br}$ - Order of areas under the CSD plot in the looped network: NL > OL > CC > LC > GL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.
	All	- With ND, there is no consistent pattern in the relations of the two networks' areas under the CSD plot. Therefore, the area under the CSD plot is not capable of capturing the effects of the network change. - Order of areas under the CSD plot in the looped network with ND: NL > OL > CC > LC > GL. - Order of areas under the CSD plot in the branched network with ND: LC > OL > CC > GL > NL.

2.12.3.6. LDI

2.12.3.6.1. Scatter Plot

Figure 2.79 shows the scatter plots of the LDI for accelerometer data measured in the looped network with 0 (GPM), 3 (GPM), and 7.5 (GPM) demand variants where the horizontal axis is the leaks' measured flow.

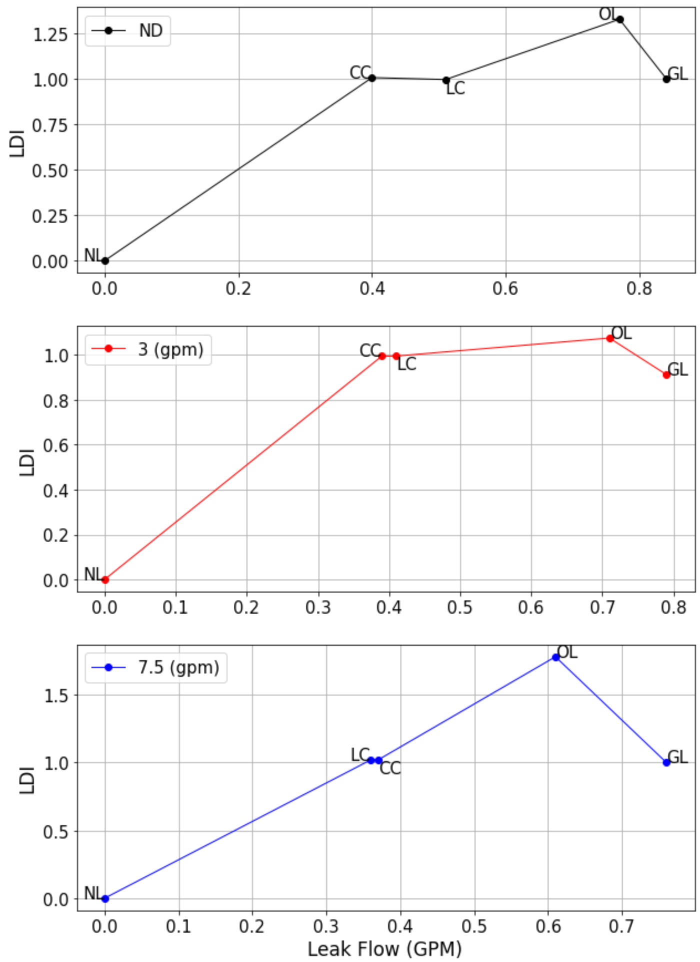

Figure 2.79. Scatter plots of the LDI for accelerometer data measured in the looped network

Figure 2.80 shows the scatter plots of the LDI for accelerometer data measured in the branched network with 0 (GPM), 3 (GPM), and 7.5 (GPM) demand variants where the horizontal axis is the leaks' measured flow.

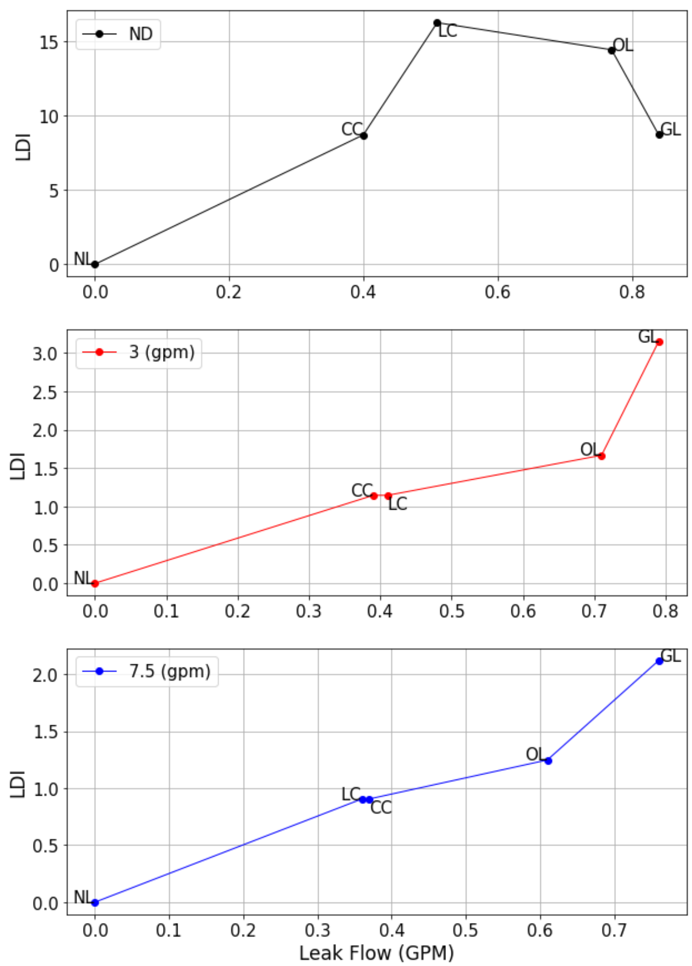

Figure 2.80. Scatter plots of the LDI for accelerometer data measured in the branched network

Table 2.30 includes an analysis of the LDI scatter plots for acceleration data of the looped and branched networks.

Table 2.30. Analysis of the LDI scatter plots for acceleration data recorded in the looped and branched networks

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
LDI (scatter plot)	All leaks and demands	- Since the LDIs of leaks are larger than those of the benchmark, i.e., NL, LDI can detect leaks. This conforms to the results of the paper of Yazdekhasti et al., 2016. However, due to the similarity of leaks' LDI, especially for CC and LC when there is a demand in both networks, and because of the inconsistent LDI magnitudes with varying demands in each network, LDI is not capable of discerning leak types. - Comparing LDI magnitudes of the looped network with a non-zero demand indicates that the LDI of OL is larger than one. Though more experiments are required, this conclusion can help distinguish OL using LDI and accelerometers in the looped network. - Comparing LDI magnitudes of the branched network with a non-zero demand indicates that the LDI of GL is above two. Though it should be evaluated with suspicion, this conclusion can help distinguish GL using LDI and accelerometers in the branched network.

2.12.3.6.2. Bar Plot

Figures 2.81 and 2.82 show the bar plots of the LDI for accelerometer data measured in the looped and branched networks, respectively, with all leak and demand variants.

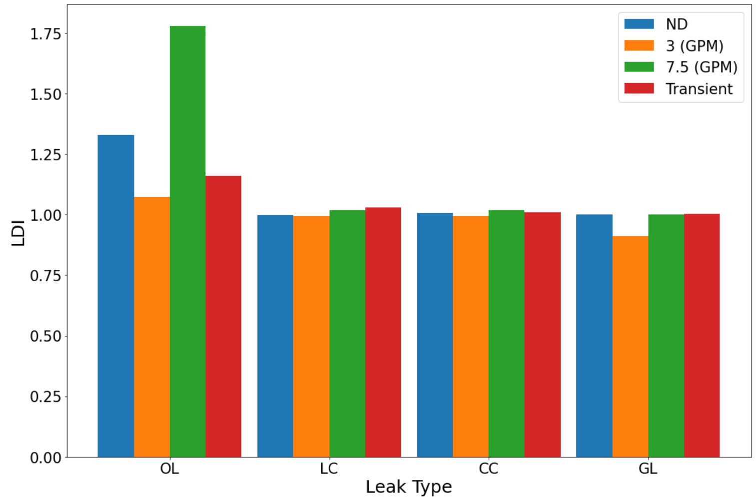

Figure 2.81. Bar plot of the LDI for accelerometer data measured in the looped network with all leak and demand variants

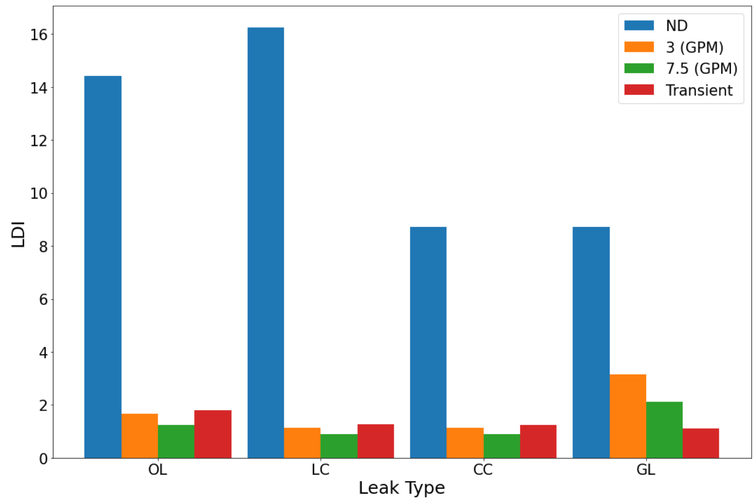

Figure 2.82. Bar plot of the LDI for accelerometer data measured in the branched network with all leak and demand variants

Table 2.31 compares the LDI bar plots of acceleration data measured in the looped and branched networks with all leak and demand variants.

Table 2.31. Analysis of acceleration data LDI bar plots measured in the looped and branched networks with all leak and demand variants

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
LDI (bar plot)	ND	- Comparison of OL signal's LDI with no demand: OL_{lo} <<< OL_{br} - Comparison of CC signal's LDI with no demand: CC_{lo} << CC_{br} - Comparison of LC signal's LDI with no demand: LC_{lo} <<< LC_{br} - Comparison of GL signal's LDI with no demand: GL_{lo} << GL_{br} - Order of LDI for signals with ND in the looped network: OL > CC > GL > LC. - Order of LDI for signals with ND in the branched network: LC > OL > GL > CC.
	3 (GPM)	- Comparison of OL signal's LDI with 3 (GPM) demand: OL_{lo} < OL_{br} - Comparison of CC signal's LDI with 3 (GPM) demand: CC_{lo} < CC_{br} - Comparison of LC signal's LDI with 3 (GPM) demand: LC_{lo} < LC_{br} - Comparison of GL signal's LDI with 3 (GPM) demand: GL_{lo} < GL_{br} - Order of LDI for signals with 3 (GPM) demand in the looped network: OL > CC = LC > GL. - Order of LDI for signals with 3 (GPM) demand in the branched network: GL > OL > LC = CC.
	7.5 (GPM)	- Comparison of OL signal's LDI with 7.5 (GPM) demand: OL_{lo} > OL_{br} - Comparison of CC signal's LDI with 7.5 (GPM) demand: CC_{lo} > CC_{br} - Comparison of LC signal's LDI with 7.5 (GPM) demand: LC_{lo} > LC_{br} - Comparison of GL signal's LDI with 7.5 (GPM) demand: GL_{lo} < GL_{br} - Order of LDI for signals with 7.5 (GPM) demand in the looped network: OL > CC = LC > GL. - Order of LDI for signals with 7.5 (GPM) demand in the branched network: GL > OL > LC = CC.

Table 2.31. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
LDI (bar plot)	Transient	- Comparison of OL signal's LDI with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$ - Comparison of CC signal's LDI with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$ - Comparison of LC signal's LDI with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$ - Comparison of GL signal's LDI with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$ - Order of LDI for signals with 7.5 (GPM) demand in the looped network: OL > LC > CC > GL. - Order of LDI for signals with 7.5 (GPM) demand in the branched network: OL > LC > CC > GL.
	All	- When there is a demand flow, LDI magnitudes for all leak types in both networks are very similar. So, the network architecture change does not affect LDI magnitudes. In other words, LDI is not capable of discerning network changes when there is a demand flow. - When there is no demand flow (ND), LDI magnitudes of all leaks in the branched network are much larger than their looped counterparts. Therefore, with ND, LDI successfully captures the effects of the network architecture changes. - With ND, LDI better-discerned leak types in the branched networks than the looped network. - Since LDI magnitudes for different leaks and non-zero demand variants are approximately the same, LDI is not capable of discerning leak types well. - In the looped network, LDI variations do not have a specific pattern given demand changes. However, in the branched network, LDI decreases when demand changes from ND to 3 (GPM) and 7.5 (GPM).

2.12.3.7. Leak:NoLeak Amplitude Plot

Figures 2.83 and 2.84 show leak:noleak amplitude plots of the accelerometer data measured by sensor A1 in the looped and branched networks, respectively, for all leak types and no demand. Figures 2.85 and 2.86 show the same plots but for sensor A2 data.

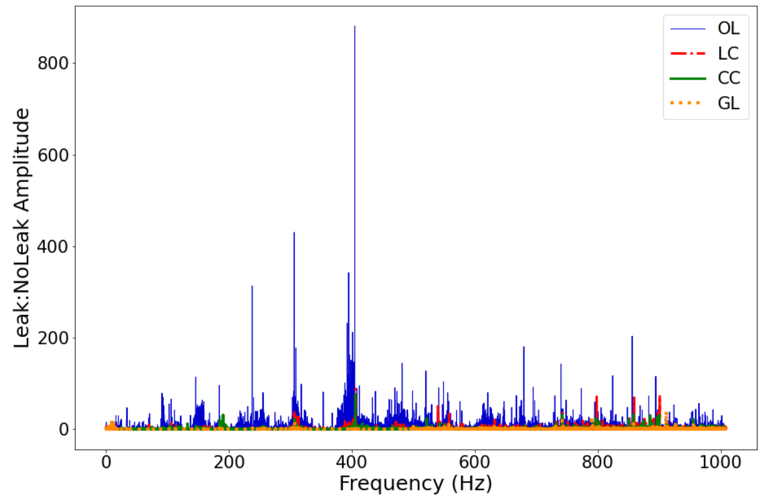

Figure 2.83. Leak:noleak amplitude plot of the accelerometer data measured by sensor A1 in the looped network with no demand

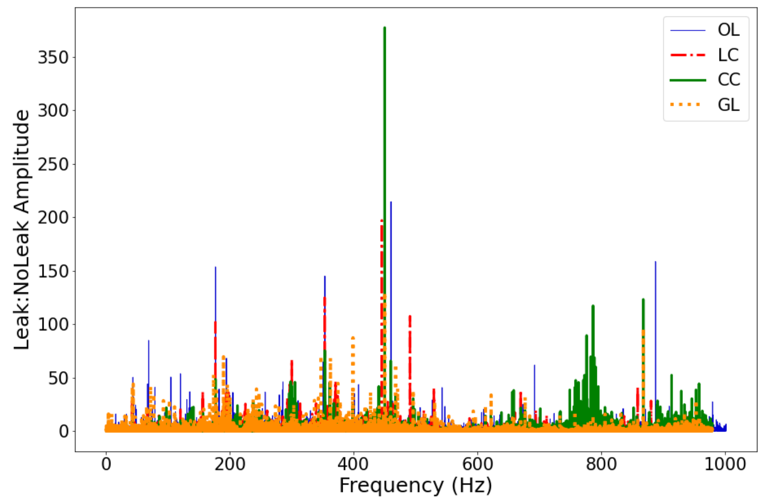

Figure 2.84. Leak:noleak amplitude plot of the accelerometer data measured by sensor A1 in the branched network with no demand

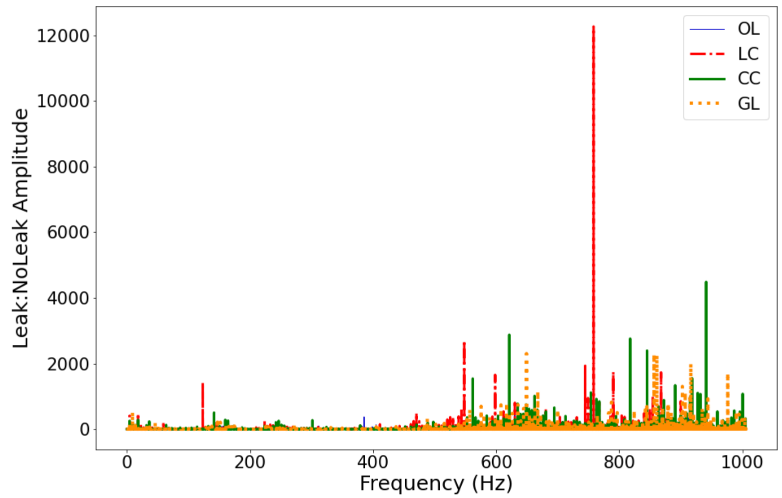

Figure 2.85. Leak:noleak amplitude plot of the accelerometer data measured by sensor A2 in the looped network with no demand

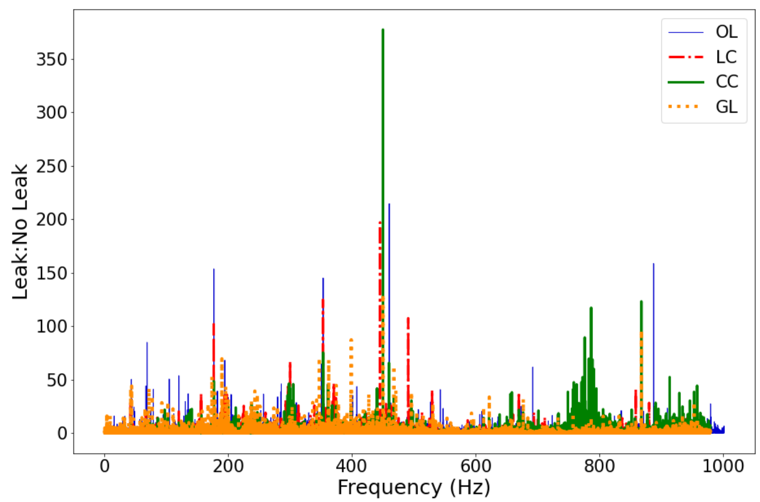

Figure 2.86. Leak:noleak amplitude plot of the accelerometer data measured by sensor A2 in the branched network with no demand

Table 2.32 compares leak:noleak amplitude plots of acceleration data measured in the looped and branched networks with all leak types and no demand.

Table 2.32. Analysis of leak:noleak amplitude plots of acceleration data measured in the looped and branched networks with all leak types and no demand

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Sum of magnitudes in the leak:noleak amplitude plot (for ND signal)	OL – A1	- Comparison of the sum of leak:noleak magnitudes for OL: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Sum of leak:noleak magnitudes for OL in the looped network is much larger than that in the branched network. - Order of sum of leak:noleak magnitudes in the looped network: OL >> LC > CC > GL. - Order of sum of leak:noleak magnitudes in the branched network: GL > CC > OL > LC.
	LC – A1	- Comparison of the sum of leak:noleak magnitudes for LC: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Order of sum of leak:noleak magnitudes in the looped network: OL >> LC > CC > GL. - Order of sum of leak:noleak magnitudes in the branched network: GL > CC > OL > LC.
	CC – A1	- Comparison of the sum of leak:noleak magnitudes for CC: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$ - Order of sum of leak:noleak magnitudes in the looped network: OL >> LC > CC > GL. - Order of sum of leak:noleak magnitudes in the branched network: GL > CC > OL > LC.
	GL – A1	- Comparison of the sum of leak:noleak magnitudes for GL: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$ - Order of sum of leak:noleak magnitudes in the looped network: OL >> LC > CC > GL. - Order of sum of leak:noleak magnitudes in the branched network: GL > CC > OL > LC.
	All – A1	- Comparing the sum of magnitudes in the leak:noleak plots of two networks with ND shows no consistent change pattern in the magnitudes when the network changes. Therefore, the sum of magnitudes in the leak:noleak plots is not capable of capturing the effects of network changes. - Order of sum of leak:noleak magnitudes in the looped network with ND: OL >> LC > CC > GL. - Order of sum of leak:noleak magnitudes in the branched network with ND: GL > CC > OL > LC.

Table 2.32. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Sum of magnitudes in the leak:noleak amplitude plot (for ND signal)	OL – A2	- Comparison of the sum of leak:noleak magnitudes for OL: $OL_{lo} > OL_{br}$. - Order of sum of leak:noleak magnitudes in the looped network: CC > LC > GL > OL. - Order of sum of leak:noleak magnitudes in the branched network: CC > GL > OL > LC.
	LC – A2	- Comparison of the sum of leak:noleak magnitudes for LC: $LC_{lo} > LC_{br}$. - Order of sum of leak:noleak magnitudes in the looped network: CC > LC > GL > OL. - Order of sum of leak:noleak magnitudes in the branched network: CC > GL > OL > LC.
	CC – A2	- Comparison of the sum of leak:noleak magnitudes for CC: $CC_{lo} > CC_{br}$. - Order of sum of leak:noleak magnitudes in the looped network: CC > LC > GL > OL. - Order of sum of leak:noleak magnitudes in the branched network: CC > GL > OL > LC.
	GL – A2	- Comparison of the sum of leak:noleak magnitudes for GL: $GL_{lo} > GL_{br}$. - Order of sum of leak:noleak magnitudes in the looped network: CC > LC > GL > OL. - Order of sum of leak:noleak magnitudes in the branched network: CC > GL > OL > LC.
	All – A2	- Contrary to sensor A1, the sum of leak:noleak magnitudes for all leaks in the looped network are larger than those of the branched network.

2.12.3.8. Dominant Frequency

Figures 2.87 and 2.88 show dominant frequency bar plots of the accelerometer data measured by sensor A1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.89 and 2.90 show the same plots but for sensor A2 data.

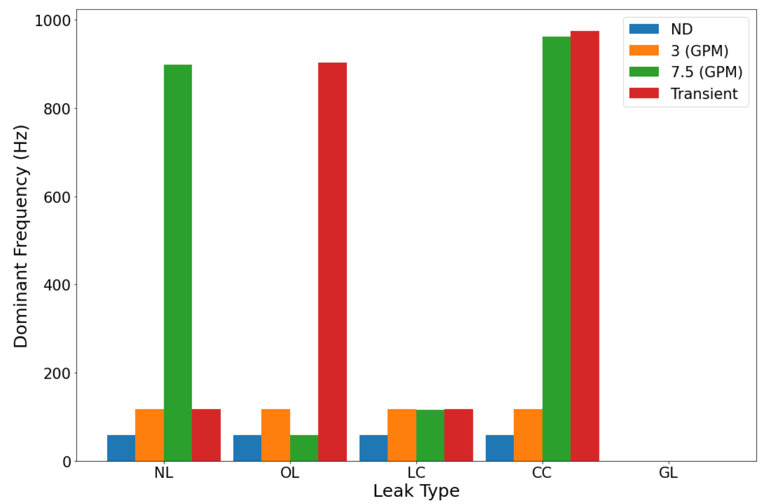

Figure 2.87. Dominant frequency bar plot of accelerometer A1 data in the looped network for all leaks and demands

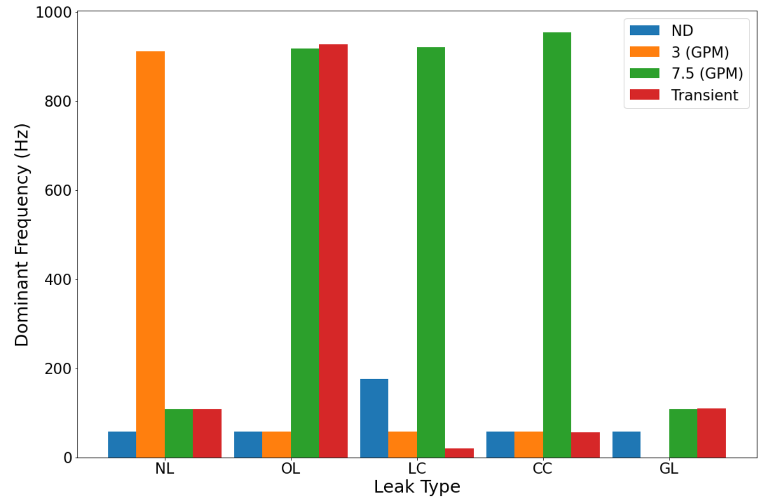

Figure 2.88. Dominant frequency bar plot of accelerometer A1 data in the branched network for all leaks and demands

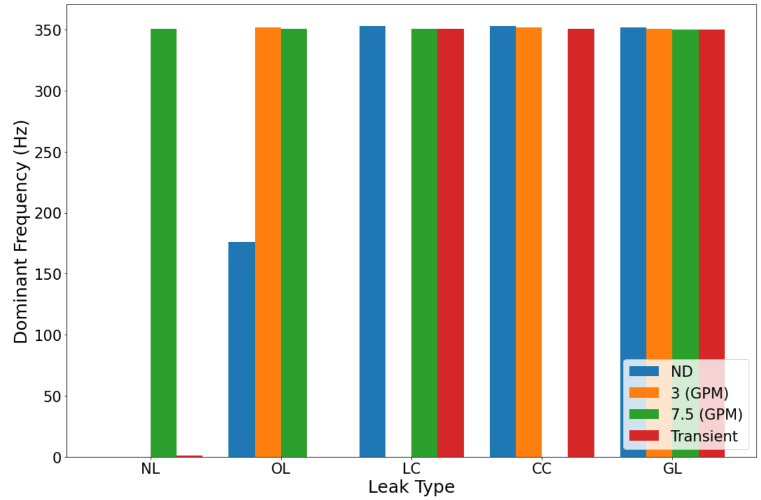

Figure 2.89. Dominant frequency bar plot of accelerometer A2 data in the looped network for all leaks and demands

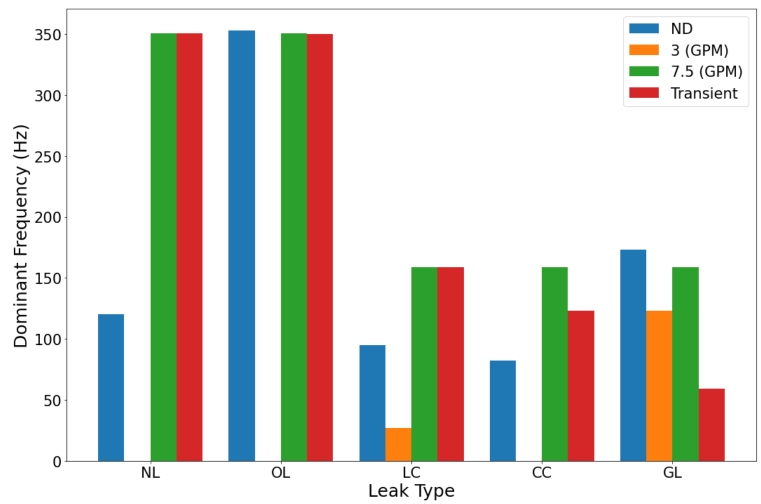

Figure 2.90. Dominant frequency bar plot of accelerometer A2 data in the branched network for all leaks and demands

Table 2.33 compares dominant frequency plots of acceleration data measured in the looped and branched networks with all leak and demand types by sensors A1 and A2.

Table 2.33. Analysis of dominant frequency plots of acceleration data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Dominant frequency	ND – A1	- Comparison of NL signal's dominant frequency with no demand: NL_{lo} = NL_{br}. - Comparison of OL signal's dominant frequency with no demand: OL_{lo} = OL_{br}. - Comparison of CC signal's dominant frequency with no demand: CC_{lo} = CC_{br}. - Comparison of LC signal's dominant frequency with no demand: LC_{lo} = LC_{br}. - Comparison of GL signal's dominant frequency with no demand: GL_{lo} << GL_{br}. - Order of dominant frequency with ND in the looped network: NL = OL = CC = LC > GL = 0. - Order of dominant frequency with ND in the branched network: NL = OL = CC = LC = GL.
	3 (GPM) – A1	- Comparison of NL signal's dominant frequency with 3 (GPM) demand: NL_{lo} << NL_{br}. - Comparison of OL signal's dominant frequency with 3 (GPM) demand: OL_{lo} > OL_{br}. - Comparison of CC signal's dominant frequency with 3 (GPM) demand: CC_{lo} > CC_{br}. - Comparison of LC signal's dominant frequency with 3 (GPM) demand: LC_{lo} > LC_{br}. - Comparison of GL signal's dominant frequency with 3 (GPM) demand: GL_{lo} = GL_{br} = 0. - Order of dominant frequency for signals with 3 (GPM) demand in the looped network: NL = OL = CC = LC > GL = 0. - Order of dominant frequency for signals with 3 (GPM) demand in the branched network: NL >> OL = CC = LC > GL = 0.

Table 2.33. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Dominant frequency	7.5 (GPM) – A1	- Comparison of NL signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{NL_{lo}}$ >> $\mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{OL_{lo}}$ << $\mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{CC_{lo}}$ > $\mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{LC_{lo}}$ << $\mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{GL_{lo}}$ < $\mathrm{GL_{br}}$. - Order of dominant frequency for signals with 7.5 (GPM) demand in the looped network: CC > NL >> LC > OL > GL = 0. - Order of dominant frequency for signals with 7.5 (GPM) demand in the branched network: CC > LC > OL >> GL > NL.
	Transient – A1	- Comparison of NL signal's dominant frequency with transient demand: $\mathrm{NL_{lo}}$ > $\mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with transient demand: $\mathrm{OL_{lo}}$ < $\mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with transient demand: $\mathrm{CC_{lo}}$ >> $\mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with transient demand: $\mathrm{LC_{lo}}$ > $\mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with transient demand: $\mathrm{GL_{lo}}$ < $\mathrm{GL_{br}}$. - Order of dominant frequency for signals with transient demand in the looped network: CC > OL >> LC = NL > GL = 0. - Order of dominant frequency for signals with transient demand in the branched network: OL >> GL > NL > LC.
	All – A1	- GL's dominant frequencies for all flow variants in the looped network are zero. - Comparing leaks' dominant frequency magnitudes of two networks indicates no consistent change pattern in the magnitudes when networks change. Therefore, dominant frequency is not capable of capturing the effects of network changes. - Since there is no consistent order of dominant frequency for signals with different demands in both networks, dominant frequency is not capable of discerning leak types consistently in both networks. - With an increase in the demand, the dominant frequencies of leaks in each network become more distinct, where with ND, leaks' dominant frequencies in each network are similar.

Table 2.33. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Dominant frequency	ND – A2	- Comparison of NL signal's dominant frequency with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's dominant frequency with no demand: $OL_{lo} >> OL_{br}$. - Comparison of CC signal's dominant frequency with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's dominant frequency with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's dominant frequency with no demand: $GL_{lo} > GL_{br}$. - Order of dominant frequency with ND in the looped network: OL > GL > NL = LC > CC. - Order of dominant frequency with ND in the branched network: LC > OL = NL = CC = GL.
	3 (GPM) – A2	- Comparison of NL signal's dominant frequency with 3 (GPM) demand: $NL_{lo} = 0 << NL_{br}$. - Comparison of OL signal's dominant frequency with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's dominant frequency with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's dominant frequency with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's dominant frequency with 3 (GPM) demand: $GL_{lo} > GL_{br} = 0$. - Order of dominant frequency for signals with 3 (GPM) demand in the looped network: GL > LC > NL = OL = CC = 0 - Order of dominant frequency for signals with 3 (GPM) demand in the branched network: NL >> OL = CC = LC > GL = 0.
	7.5 (GPM) – A2	- Comparison of NL signal's dominant frequency with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's dominant frequency with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's dominant frequency with 7.5 (GPM) demand: $CC_{lo} << CC_{br}$. - Comparison of LC signal's dominant frequency with 7.5 (GPM) demand: $LC_{lo} << LC_{br}$. - Comparison of GL signal's dominant frequency with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of dominant frequency for signals with 7.5 (GPM) demand in the looped network: NL = OL > LC = CC = GL. - Order of dominant frequency for signals with 7.5 (GPM) demand in the branched network: CC > LC > OL >> GL > NL.

Table 2.33. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Dominant frequency	Transient – A2	- Comparison of NL signal's dominant frequency with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with transient demand: $\mathrm{LC_{lo}} >> \mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of dominant frequency for signals with transient demand in the looped network: NL > OL > LC > CC > GL. - Order of dominant frequency for signals with transient demand in the branched network: OL >> GL > NL > CC > LC.
	All – A2	- Comparing leaks' dominant frequency magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, dominant frequency is not capable of capturing the effects of network changes. - Since there is no consistent order of dominant frequency for signals with different demands in both networks, dominant frequency is not capable of discerning leak types consistently in both networks.

2.12.3.9. Fundamental Frequency

Figures 2.91 and 2.92 show fundamental frequency bar plots of the accelerometer data measured by sensor A1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.93 and 2.94 show the same plots but for sensor A2 data.

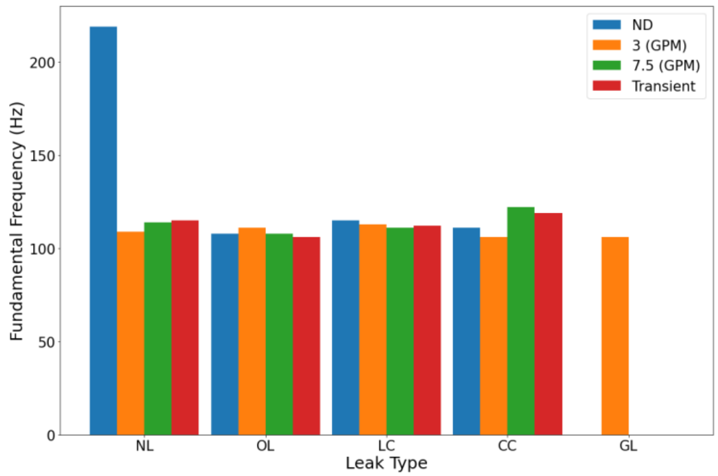

Figure 2.91. Fundamental frequency bar plot of accelerometer A1 data in the looped network for all leaks and demands

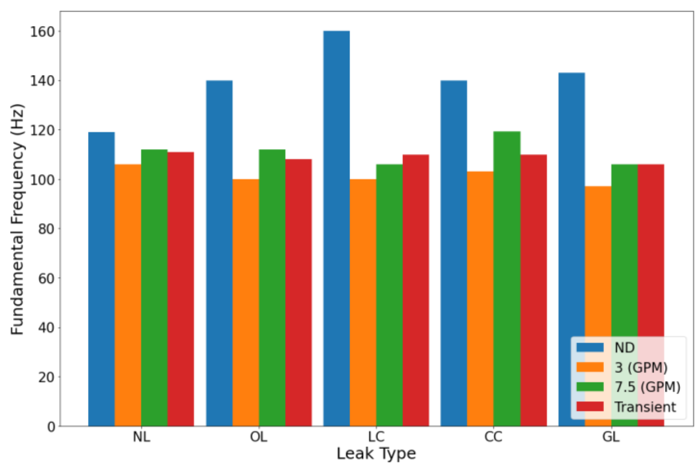

Figure 2.92. Fundamental frequency bar plot of accelerometer A1 data in the branched network for all leaks and demands

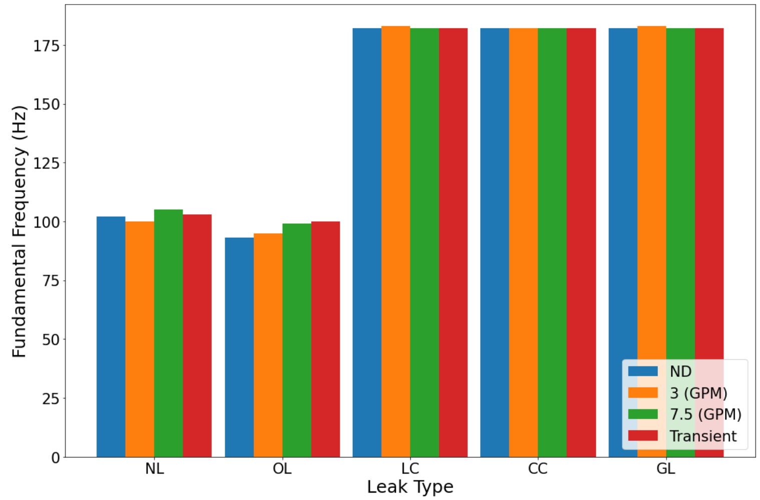

Figure 2.93. Fundamental frequency bar plot of accelerometer A2 data in the looped network for all leaks and demands

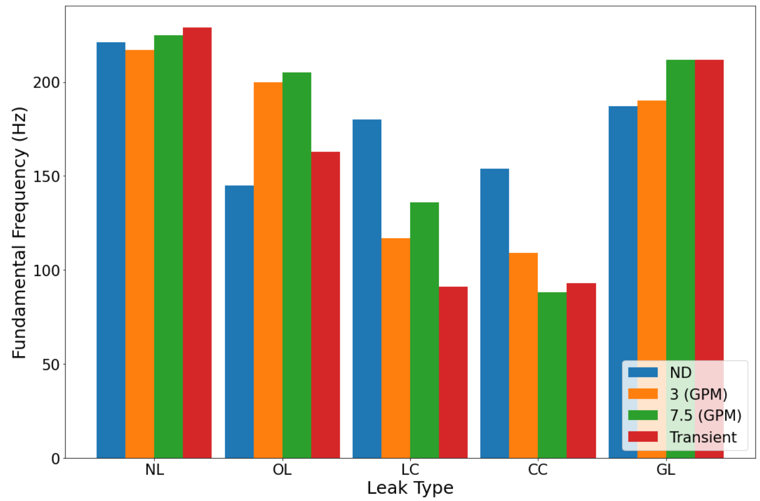

Figure 2.94. Fundamental frequency bar plot of accelerometer A2 data in the branched network for all leaks and demands

Table 2.34 compares fundamental frequency plots of acceleration data measured in the looped and branched networks with all leak and demand types by sensors A1 and A2.

Table 2.34. Analysis of fundamental frequency plots of acceleration data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Fundamental frequency	ND – A1	- Comparison of NL signal's fundamental frequency with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's fundamental frequency with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's fundamental frequency with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's fundamental frequency with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's fundamental frequency with no demand: $GL_{lo} < GL_{br}$. - Order of fundamental frequency with ND in the looped network: NL > LC > CC > OL > GL = 0. - Order of fundamental frequency with ND in the branched network: LC > GL > OL = CC > NL.
	3 (GPM) – A1	- Comparison of NL signal's fundamental frequency with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's fundamental frequency with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's fundamental frequency with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's fundamental frequency with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's fundamental frequency with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of fundamental frequency for signals with 3 (GPM) demand in the looped network: LC > OL > NL > CC = GL. - Order of fundamental frequency for signals with 3 (GPM) demand in the branched network: NL > CC > CC = OL > GL.

Table 2.34. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Fundamental frequency	7.5 (GPM) – A1	- Comparison of NL signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of fundamental frequency for signals with 7.5 (GPM) demand in the looped network: CC > NL > LC > OL > GL = 0. - Order of fundamental frequency for signals with 7.5 (GPM) demand in the branched network: CC > NL = OL > LC = GL.
	Transient – A1	- Comparison of NL signal's fundamental frequency with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of fundamental frequency for signals with transient demand in the looped network: CC > NL > LC > OL > GL = 0. - Order of fundamental frequency for signals with transient demand in the branched network: NL > CC = LC > OL > GL.

Table 2.34. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Fundamental frequency	All – A1	- In the looped network, except for NL with ND that has the largest fundamental frequency, the fundamental frequency of all leaks with different demands is similar. Therefore, the fundamental frequency cannot discern leak types in the looped network. - Comparing leaks' fundamental frequency magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, the fundamental frequency is not capable of capturing the effects of network changes. - Since there is no consistent order of fundamental frequency for signals with different demands in the branched network and due to the similar fundamental frequencies for different leaks in the looped network, the fundamental frequency is not capable of discerning leak types consistently in both networks. - When there is a demand, GL has the lowest fundamental frequency in both networks.
	ND – A2	- Comparison of NL signal's fundamental frequency with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with no demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with no demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with no demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of fundamental frequency with ND in the looped network: LC = CC = GL > NL > OL. - Order of fundamental frequency with ND in the branched network: NL > GL > LC > CC > OL.
	3 (GPM) – A2	- Comparison of NL signal's fundamental frequency with 3 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with 3 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with 3 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with 3 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of fundamental frequency for signals with 3 (GPM) demand in the looped network: LC = CC = GL > NL > OL. - Order of fundamental frequency for signals with 3 (GPM) demand in the branched network: NL > OL > GL > LC > CC.

Table 2.34. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Fundamental frequency	7.5 (GPM) – A2	- Comparison of NL signal's fundamental frequency with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's fundamental frequency with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's fundamental frequency with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's fundamental frequency with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's fundamental frequency with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of fundamental frequency for signals with 7.5 (GPM) demand in the looped network: LC = CC = GL > NL > OL. - Order of fundamental frequency for signals with 7.5 (GPM) demand in the branched network: NL > GL > OL > LC > CC.
	Transient – A2	- Comparison of NL signal's fundamental frequency with transient demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's fundamental frequency with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's fundamental frequency with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's fundamental frequency with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's fundamental frequency with transient demand: $GL_{lo} < GL_{br}$. - Order of fundamental frequency for signals with transient demand in the looped network: LC = CC = GL > NL > OL. - Order of fundamental frequency for signals with transient demand in the branched network: NL > GL > OL > CC > LC.
	All – A2	- Comparing leaks' fundamental frequency magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, the fundamental frequency is not capable of capturing the effects of network changes. - Due to the similarity of fundamental frequencies for signals with different demands in the looped network and inconsistent fundamental frequencies for signals with different demands in the branched network, the fundamental frequency is not capable of discerning leak types consistently in both networks.

2.12.3.10. Spectral Centroid

Figures 2.95 and 2.96 show spectral centroid bar plots of the accelerometer data measured by sensor A1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.97 and 2.98 show the same plots but for sensor A2 data.

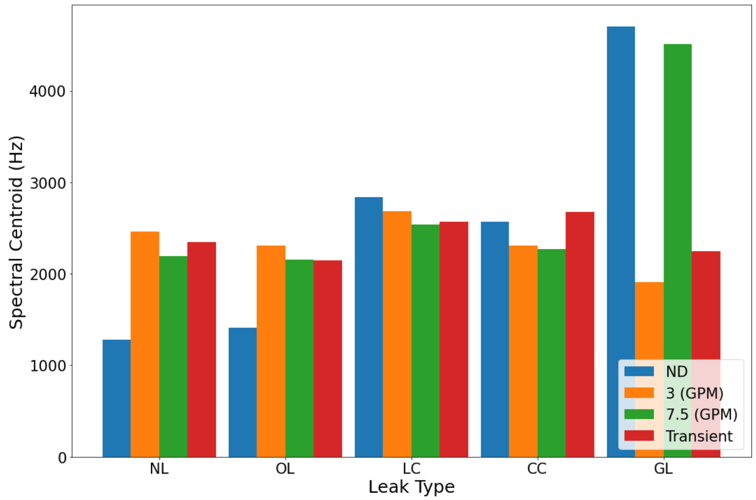

Figure 2.95. Spectral centroid bar plot of accelerometer A1 data in the looped network for all leaks and demands

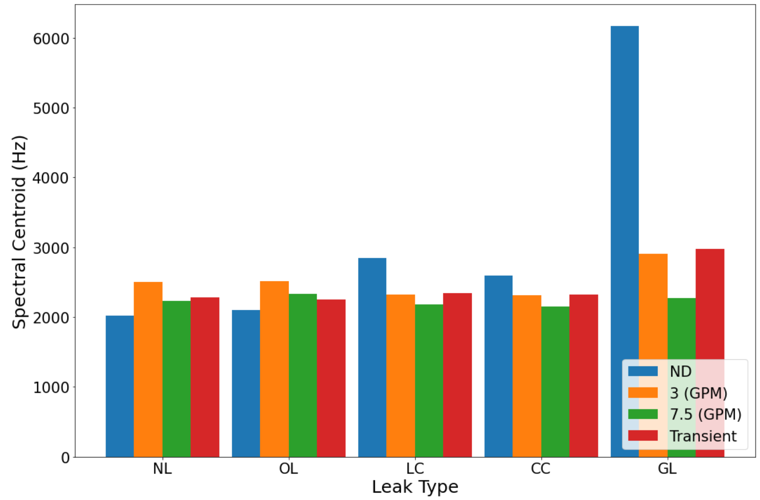

Figure 2.96. Spectral centroid bar plot of accelerometer A1 data in the branched network for all leaks and demands

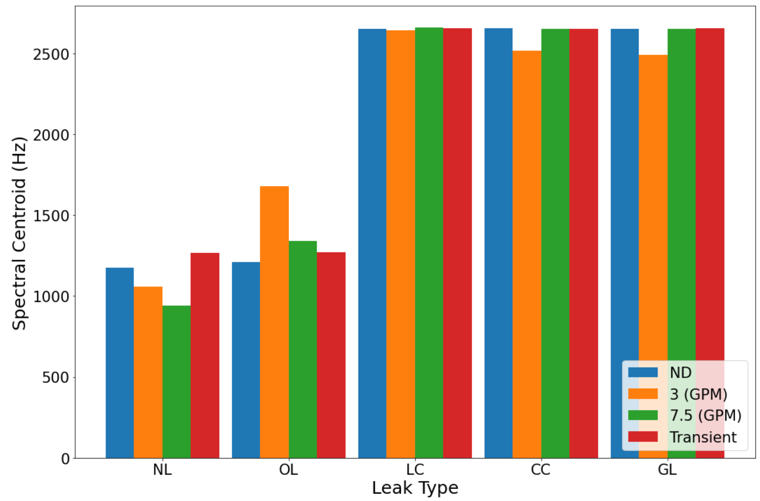

Figure 2.97. Spectral centroid bar plot of accelerometer A2 data in the looped network for all leaks and demands

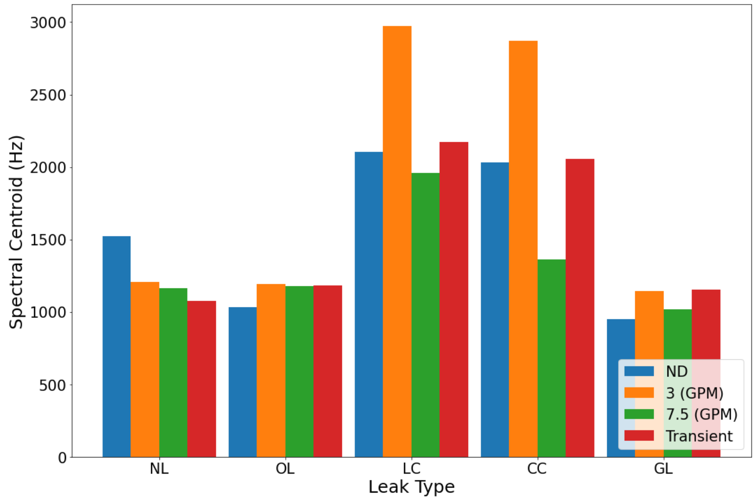

Figure 2.98. Spectral centroid bar plot of accelerometer A2 data in the branched network for all leaks and demands

Table 2.35 compares spectral centroid plots of acceleration data measured in the looped and branched networks with all leak and demand types by sensors A1 and A2.

Table 2.35. Analysis of spectral centroid plots of acceleration data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Spectral centroid	ND – A1	- Comparison of NL signal's spectral centroid with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's spectral centroid with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's spectral centroid with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's spectral centroid with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's spectral centroid with no demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid with ND in the looped network: GL > LC > CC > OL > NL. - Order of spectral centroid with ND in the branched network: GL > LC > CC > OL > NL.
	3 (GPM) – A1	- Comparison of NL signal's spectral centroid with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's spectral centroid with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's spectral centroid with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's spectral centroid with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's spectral centroid with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid for signals with 3 (GPM) demand in the looped network: LC > NL > CC > OL > GL. - Order of spectral centroid for signals with 3 (GPM) demand in the branched network: GL > OL > NL > LC > CC.

Table 2.35. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Spectral centroid	7.5 (GPM) – A1	- Comparison of NL signal's spectral centroid with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's spectral centroid with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's spectral centroid with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's spectral centroid with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's spectral centroid with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of spectral centroid for signals with 7.5 (GPM) demand in the looped network: GL > LC > CC > NL > OL. - Order of spectral centroid for signals with 7.5 (GPM) demand in the branched network: OL > GL > NL > LC > CC.
	Transient – A1	- Comparison of NL signal's spectral centroid with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's spectral centroid with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's spectral centroid with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's spectral centroid with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's spectral centroid with transient demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid for signals with transient demand in the looped network: CC > LC > NL > GL > OL. - Order of spectral centroid for signals with transient demand in the branched network: GL > LC > CC > NL > OL.
	All – A1	- Comparing leaks' spectral centroid magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, the spectral centroid is not capable of capturing the effects of network changes. - Since there is no consistent order of spectral centroid for signals with different demands in both networks, the spectral centroid cannot discern leak types consistently in both networks.

Table 2.35. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Spectral centroid	ND – A2	- Comparison of NL signal's spectral centroid with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's spectral centroid with no demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's spectral centroid with no demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's spectral centroid with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's spectral centroid with no demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of spectral centroid with ND in the looped network: CC > GL > LC > OL > NL. - Order of spectral centroid with ND in the branched network: LC > CC > NL > OL > GL.
	3 (GPM) – A2	- Comparison of NL signal's spectral centroid with 3 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's spectral centroid with 3 (GPM) demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's spectral centroid with 3 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's spectral centroid with 3 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's spectral centroid with 3 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of spectral centroid for signals with 3 (GPM) demand in the looped network: LC > CC > GL > OL > NL. - Order of spectral centroid for signals with 3 (GPM) demand in the branched network: LC > CC > NL > OL > GL.
	7.5 (GPM) – A2	- Comparison of NL signal's spectral centroid with 7.5 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's spectral centroid with 7.5 (GPM) demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's spectral centroid with 7.5 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's spectral centroid with 7.5 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's spectral centroid with 7.5 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of spectral centroid for signals with 7.5 (GPM) demand in the looped network: LC > CC > GL > OL > NL. - Order of spectral centroid for signals with 7.5 (GPM) demand in the branched network: LC > CC > NL > OL > GL.

Table 2.35. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Spectral centroid	Transient – A2	- Comparison of NL signal's spectral centroid with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's spectral centroid with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's spectral centroid with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's spectral centroid with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's spectral centroid with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of spectral centroid for signals with transient demand in the looped network: LC > GL > CC > OL > NL. - Order of spectral centroid for signals with transient demand in the branched network: LC > CC > OL > GL > NL.
	All – A2	- Comparing leaks' spectral centroid magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, the spectral centroid is not capable of capturing the effects of network changes. - Due to the similarity of the spectral centroid for signals with different demands in the looped network and inconsistent spectral centroid for signals with different demands in the branched network, the spectral centroid is not capable of discerning leak types consistently in both networks.

2.12.3.11. Power Spectral Entropy

Figures 2.99 and 2.100 show power spectral entropy bar plots of the accelerometer data measured by sensor A1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.101 and 2.102 show the same plots but for sensor A2 data.

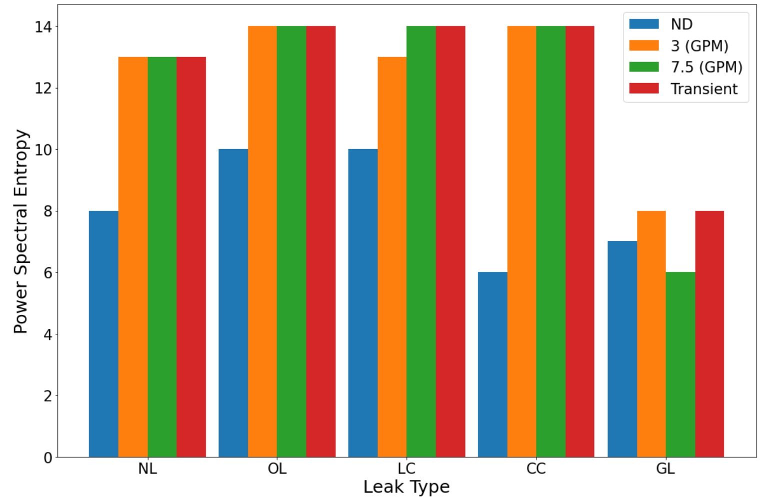

Figure 2.99. Power spectral entropy bar plot of accelerometer A1 data in the looped network for all leaks and demands

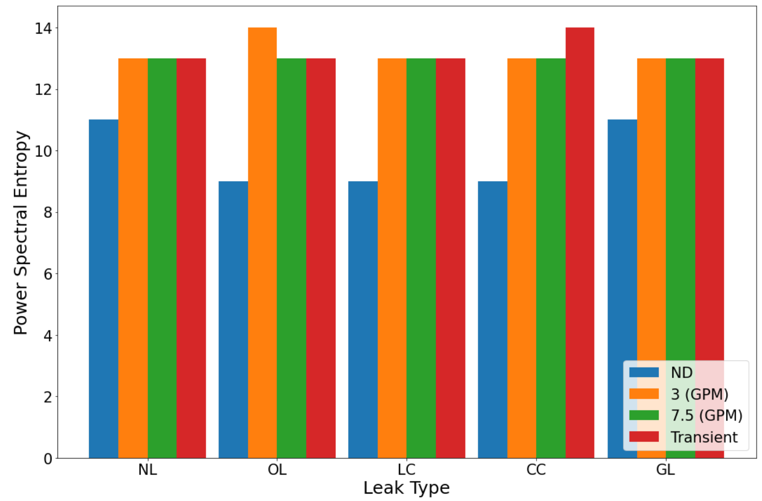

Figure 2.100. Power spectral entropy bar plot of accelerometer A1 data in the branched network for all leaks and demands

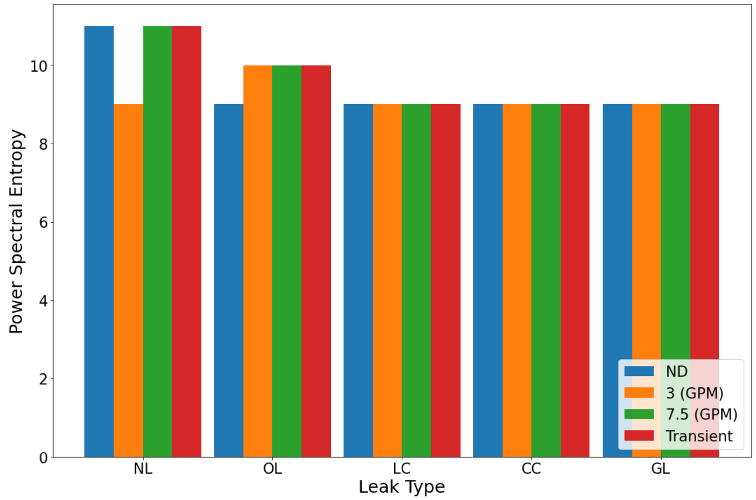

Figure 2.101. Power spectral entropy bar plot of accelerometer A2 data in the looped network for all leaks and demands

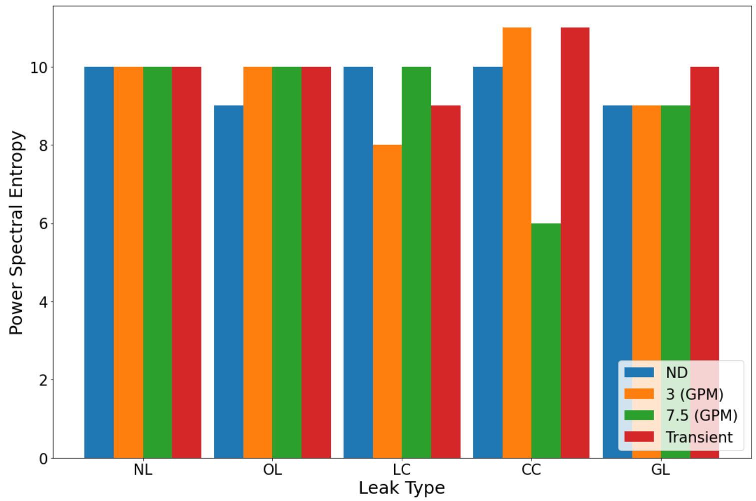

Figure 2.102. Power spectral entropy bar plot of accelerometer A2 data in the branched network for all leaks and demands

Table 2.36 compares power spectral entropy plots of acceleration data measured in the looped and branched networks with all leak and demand types by sensors A1 and A2.

Table 2.36. Analysis of power spectral entropy plots of acceleration data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Power spectral entropy	ND – A1	- Comparison of NL signal's power spectral entropy with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's power spectral entropy centroid with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's power spectral entropy centroid with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's power spectral entropy centroid with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's power spectral entropy centroid with no demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid with ND in the looped network: OL = LC > NL > GL > CC. - Order of spectral centroid with ND in the branched network: NL = GL > OL = LC = CC.
	3 (GPM) – A1	- Comparison of NL signal's power spectral entropy with 3 (GPM) demand: $NL_{lo} = NL_{br}$. - Comparison of OL signal's power spectral entropy with 3 (GPM) demand: $OL_{lo} = OL_{br}$. - Comparison of CC signal's power spectral entropy with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with 3 (GPM) demand: $LC_{lo} = LC_{br}$. - Comparison of GL signal's power spectral entropy with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of power spectral entropy for signals with 3 (GPM) demand in the looped network: OL = CC > NL = LC > GL. - Order of power spectral entropy for signals with 3 (GPM) demand in the branched network: OL > NL = LC = CC = GL.

Table 2.36. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Power spectral entropy	7.5 (GPM) – A1	- Comparison of NL signal's power spectral entropy with 7.5 (GPM) demand: $\mathrm{NL_{lo}} = \mathrm{NL_{br}}$. - Comparison of OL signal's power spectral entropy with 7.5 (GPM) demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's power spectral entropy with 7.5 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's power spectral entropy with 7.5 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's power spectral entropy with 7.5 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the looped network: OL = LC = CC > NL > GL. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the branched network: NL = OL = LC = CC = GL.
	Transient – A1	- Comparison of NL signal's power spectral entropy with transient demand: $\mathrm{NL_{lo}} = \mathrm{NL_{br}}$. - Comparison of OL signal's power spectral entropy with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's power spectral entropy with transient demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's power spectral entropy with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's power spectral entropy with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of power spectral entropy for signals with transient demand in the looped network: CC = LC = OL > NL > GL. - Order of power spectral entropy for signals with transient demand in the branched network: CC > NL = OL = LC = GL.
	All – A1	- Comparing leaks' power spectral entropy magnitudes of two networks shows either no consistent change pattern in the magnitudes or similar power spectral entropy when networks change. Therefore, power spectral entropy cannot capture the effects of network changes. - Due to the similarity of power spectral entropy magnitudes for signals with different demands in each network, power spectral entropy cannot discern leak types in both networks. However, in both networks, power spectral entropy is larger when demand is present than those with ND signals. - When demand is present, GL in both networks has the lowest power spectral entropy.

Table 2.36. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Power spectral entropy	ND – A2	- Comparison of NL signal's power spectral entropy with no demand: $\mathrm{NL_{lo}} = \mathrm{NL_{br}}$. - Comparison of OL signal's power spectral entropy with no demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's power spectral entropy with no demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's power spectral entropy with no demand: $\mathrm{LC_{lo}} = \mathrm{LC_{br}}$. - Comparison of GL signal's power spectral entropy with no demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of power spectral entropy with ND in the looped network: NL > OL = LC = CC = GL. - Order of power spectral entropy with ND in the branched network: NL = GL > OL = LC = CC.
	3 (GPM) – A2	- Comparison of NL signal's power spectral entropy with 3 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's power spectral entropy with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's power spectral entropy with 3 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's power spectral entropy with 3 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's power spectral entropy with 3 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of power spectral entropy for signals with 3 (GPM) demand in the looped network: OL > NL = LC = CC = GL. - Order of power spectral entropy for signals with 3 (GPM) demand in the branched network: OL > NL = LC = CC = GL.

Table 2.36. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Power spectral entropy	7.5 (GPM) – A2	- Comparison of NL signal's power spectral entropy with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's power spectral entropy with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's power spectral entropy with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's power spectral entropy with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's power spectral entropy with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the looped network: NL > OL > LC = CC = GL. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the branched network: NL = OL = LC = CC = GL.
	Transient – A2	- Comparison of NL signal's power spectral entropy with transient demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's power spectral entropy with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's power spectral entropy with transient demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's power spectral entropy with transient demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's power spectral entropy with transient demand: $GL_{lo} < GL_{br}$. - Order of power spectral entropy for signals with transient demand in the looped network: NL > OL > LC = CC = GL. - Order of power spectral entropy for signals with transient demand in the branched network: CC > NL = OL = LC = GL.
	All – A2	- Based on the power spectral entropy of signals recorded by sensor A2, when there is a demand, the signals' spectral entropies are larger in the branched network than those in the looped network. Therefore, by using sensor A2, power spectral entropy can capture the network change's effects. - Due to the similarity of the spectral entropies of signals with different demands in both networks, the spectral entropy is not capable of discerning leak types consistently in both networks.

2.12.3.12. Mean

Figures 2.103 and 2.104 plot the mean of the accelerometer data measured by sensor A1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.105 and 2.106 show the same plots but for sensor A2 data.

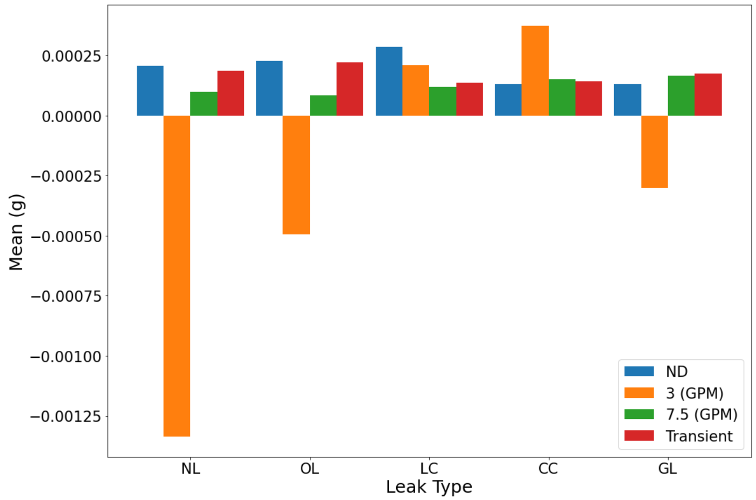

Figure 2.103. Mean plot of accelerometer A1 data in the looped network for all leaks and demands

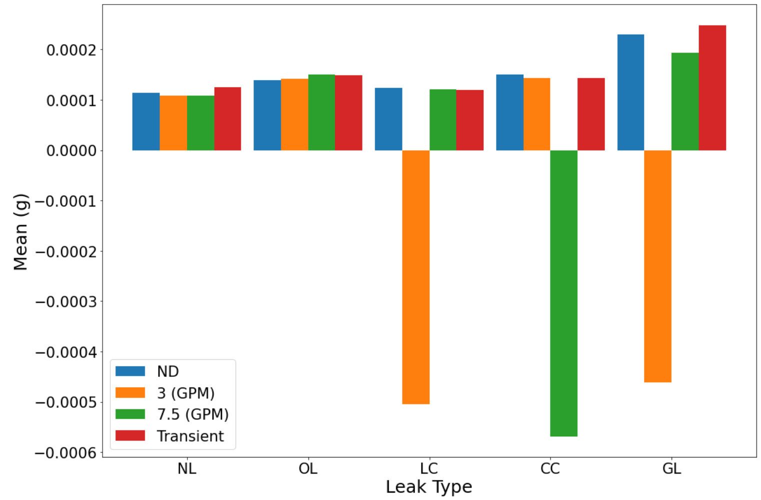

Figure 2.104. Mean plot of accelerometer A1 data in the branched network for all leaks and demands

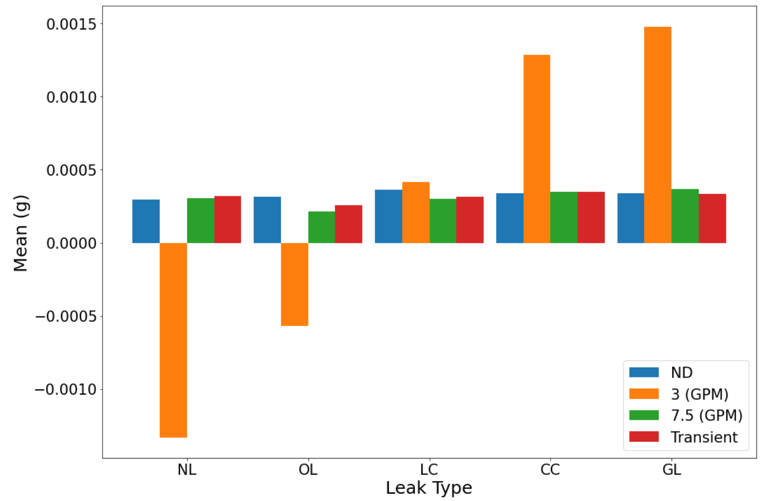

Figure 2.105. Mean plot of accelerometer A2 data in the looped network for all leaks and demands

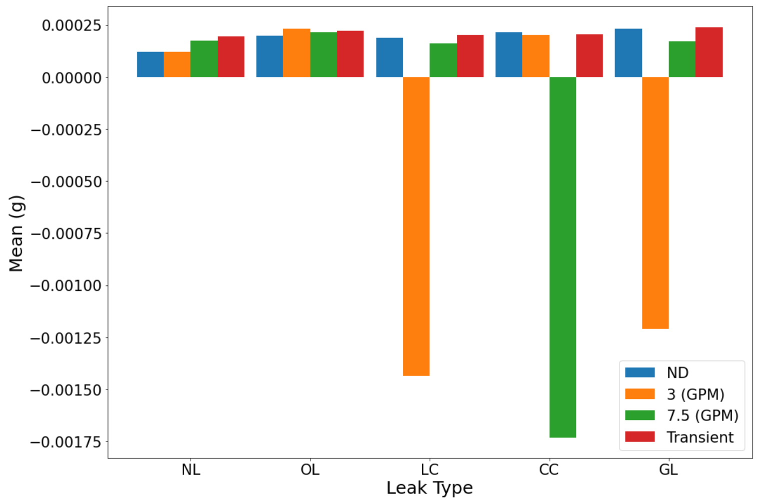

Figure 2.106. Mean plot of accelerometer A2 data in the branched network for all leaks and demands

Table 2.37 compares mean plots of acceleration data measured in the looped and branched networks with all leak and demand types by sensors A1 and A2.

Table 2.37. Analysis of mean plots of acceleration data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Mean	ND – A1	- Comparison of NL signal's mean with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's mean with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's mean with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's mean with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's mean with no demand: $GL_{lo} < GL_{br}$. - Order of mean with ND in the looped network: LC > OL > NL > CC > GL. - Order of mean with ND in the branched network: GL > CC > OL > NL > LC.
	3 (GPM) – A1	- Comparison of NL signal's mean with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's mean with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's mean with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's mean with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's mean with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of mean for signals with 3 (GPM) demand in the looped network: CC > LC > GL > OL > NL. - Order of mean for signals with 3 (GPM) demand in the branched network: OL > CC > NL > GL > LC.
	7.5 (GPM) – A1	- Comparison of NL signal's mean with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's mean with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's mean with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's mean with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's mean with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of mean for signals with 7.5 (GPM) demand in the looped network: GL > CC > LC > NL > OL. - Order of mean for signals with 7.5 (GPM) demand in the branched network: GL > OL > LC > NL > CC.

Table 2.37. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Mean	Transient – A1	- Comparison of NL signal's mean with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's mean with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's mean with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's mean with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's mean with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of mean for signals with transient demand in the looped network: OL > NL > GL > CC > LC. - Order of mean for signals with transient demand in the branched network: GL > OL > CC > NL > LC.
	All – A1	- Comparing leaks' mean magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, mean cannot capture the effects of network changes. - Due to the inconsistent order of mean for signals with different demands in both networks, mean is not capable of discerning leak types consistently in both networks. - In the looped network, with the demand increase, signals' mean for LC decreases, i.e., there is an inverse relation between demand and pipes' acceleration for LC in the looped network.
	ND – A2	- Comparison of NL signal's mean with no demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's mean with no demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's mean with no demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's mean with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's mean with no demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of mean with ND in the looped network: LC > GL > CC > OL > NL. - Order of mean with ND in the branched network: GL > CC > OL > LC > NL.

Table 2.37. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Mean	3 (GPM) – A2	- Comparison of NL signal's mean with 3 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's mean with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's mean with 3 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's mean with 3 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's mean with 3 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of mean for signals with 3 (GPM) demand in the looped network: GL > CC > LC > OL > NL. - Order of mean for signals with 3 (GPM) demand in the branched network: OL > CC > NL > GL > LC.
	7.5 (GPM) – A2	- Comparison of NL signal's mean with 7.5 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's mean with 7.5 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's mean with 7.5 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's mean with 7.5 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's mean with 7.5 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$.
	Transient – A2	- Comparison of NL signal's mean with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's mean with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's mean with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's mean with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's mean with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of mean for signals with transient demand in the looped network: CC > GL > NL > LC > OL. - Order of mean for signals with transient demand in the branched network: GL > OL > CC > OL > NL.
	All – A2	- Comparing leaks' mean magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, mean cannot capture the effects of network changes. - Due to the inconsistent order of mean for signals with different demands in both networks, mean is not capable of discerning leak types consistently in both networks.

2.12.3.13. Standard Deviation

Figures 2.107 and 2.108 show the standard deviation plots of the accelerometer data measured by sensor A1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.109 and 2.110 show the same plots but for sensor A2 data.

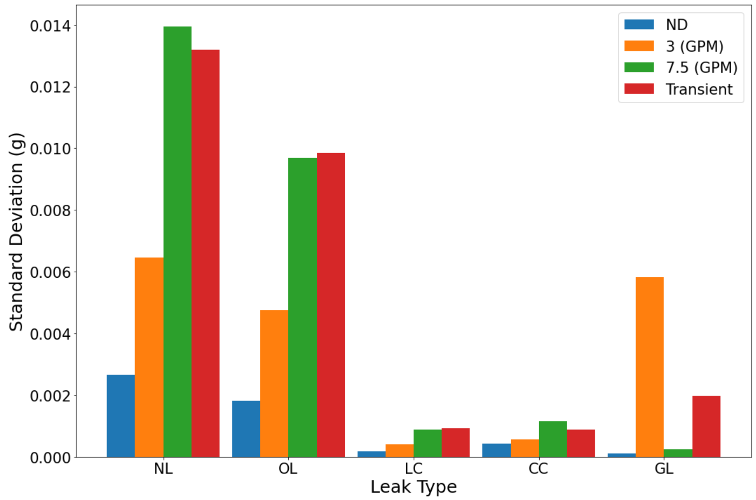

Figure 2.107. Standard deviation plot of accelerometer A1 data in the looped network for all leaks and demands

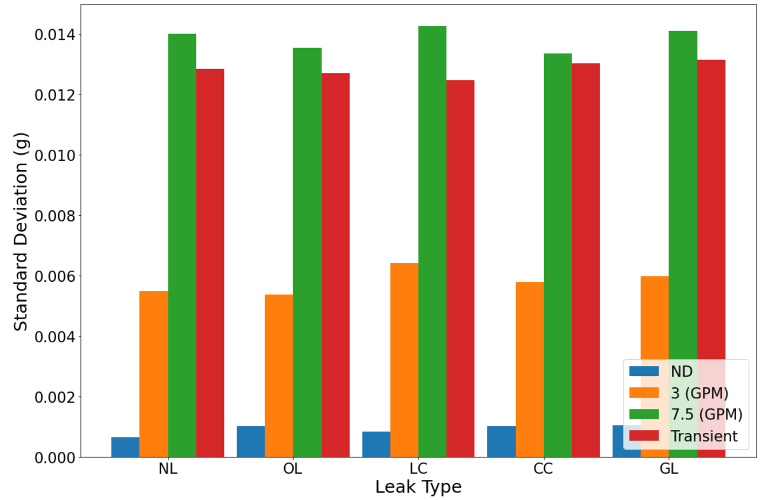

Figure 2.108. Standard deviation plot of accelerometer A1 data in the branched network for all leaks and demands

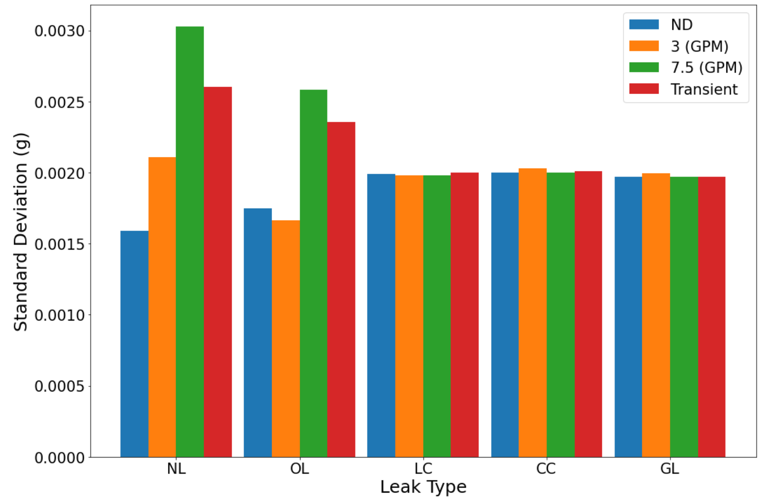

Figure 2.109. Standard deviation plot of accelerometer A2 data in the looped network for all leaks and demands

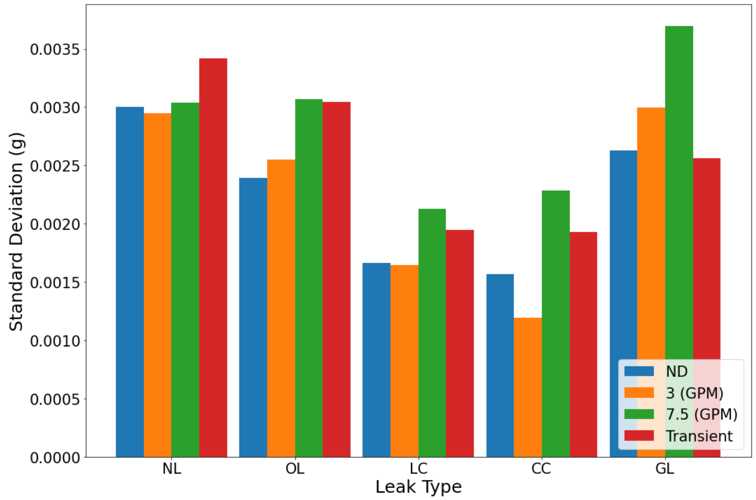

Figure 2.110. Standard deviation plot of accelerometer A2 data in the branched network for all leaks and demands

Table 2.38 compares standard deviation plots of acceleration data measured in the looped and branched networks with all leak and demand types by sensors A1 and A2.

Table 2.38. Analysis of standard deviation plots of acceleration data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Standard deviation	ND – A1	- Comparison of NL signal's standard deviation with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's standard deviation with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's standard deviation with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's standard deviation with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's standard deviation with no demand: $GL_{lo} < GL_{br}$. - Order of standard deviation with ND in the looped network: NL > OL > CC > LC > GL. - Order of standard deviation with ND in the branched network: GL > CC > OL > LC > NL.
	3 (GPM) – A1	- Comparison of NL signal's standard deviation with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's standard deviation entropy with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation entropy with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's standard deviation with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's standard deviation with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of standard deviation for signals with 3 (GPM) demand in the looped network: NL > GL > CC > OL > LC. - Order of standard deviation for signals with 3 (GPM) demand in the branched network: LC > GL > CC > NL > OL.
	7.5 (GPM) – A1	- Comparison of NL signal's standard deviation with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's standard deviation with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's standard deviation with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's standard deviation with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of standard deviation for signals with 7.5 (GPM) demand in the looped network: NL > OL > CC > LC > GL. - Order of standard deviation for signals with 7.5 (GPM) demand in the branched network: LC > GL > NL > OL > CC.

Table 2.38. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Standard deviation	Transient – A1	- Comparison of NL signal's standard deviation with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's standard deviation with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's standard deviation with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's standard deviation with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's standard deviation with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of standard deviation for signals with transient demand in the looped network: NL > OL > GL > LC > CC. - Order of standard deviation for signals with transient demand in the branched network: GL > CC > NL > OL > LC.
	All – A1	- In the branched network and for each leak type, standard deviation follows this pattern: $\mathrm{SD_{7.5\ (GPM)}} > \mathrm{SD_{transient}} > \mathrm{SD_{3\ (GPM)}} > \mathrm{SD_{ND}}$. - Comparing leaks' standard deviation magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, standard deviation cannot capture the effects of network changes. - Due to the inconsistent order of standard deviation for signals with different demands in both networks, standard deviation is not capable of discerning leak types consistently in both networks. - In the looped network, LC signals' standard deviation is the largest with all demand variants.
	ND – A2	- Comparison of NL signal's standard deviation with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's standard deviation with no demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's standard deviation with no demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's standard deviation with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's standard deviation with no demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of standard deviation with ND in the looped network: CC > LC > GL > OL > NL. - Order of standard deviation with ND in the branched network: NL > GL > OL > LC > CC.

Table 2.38. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Standard deviation	3 (GPM) – A2	- Comparison of NL signal's standard deviation with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's standard deviation with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's standard deviation with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's standard deviation with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of standard deviation for signals with 3 (GPM) demand in the looped network: NL > CC > GL > LC > OL. - Order of standard deviation for signals with 3 (GPM) demand in the branched network: GL > NL > OL > LC > CC.
	7.5 (GPM) – A2	- Comparison of NL signal's standard deviation with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's standard deviation with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's standard deviation with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's standard deviation with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of standard deviation for signals with 7.5 (GPM) demand in the looped network: NL > OL > CC > LC > GL. - Order of standard deviation for signals with 7.5 (GPM) demand in the branched network: GL > OL > NL > CC > LC.
	Transient – A2	- Comparison of NL signal's standard deviation with transient demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's standard deviation with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's standard deviation with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's standard deviation with transient demand: $GL_{lo} < GL_{br}$. - Order of standard deviation for signals with transient demand in the looped network: NL > OL > CC > LC > GL. - Order of standard deviation for signals with transient demand in the branched network: GL > OL > CC > OL > NL.

Table 2.38. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Standard deviation	All – A2	- Comparing leaks' standard deviation magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, standard deviation cannot capture the effects of network changes. - Due to the inconsistent order of standard deviation for signals with different demands in both networks, standard deviation is not capable of discerning leak types consistently in both networks.

2.12.3.14. Zero-crossing Rate

Figures 2.111 and 2.112 show the zero-crossing rate plots of the accelerometer data measured by sensor A1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.113 and 2.114 show the same plots but for sensor A2 data.

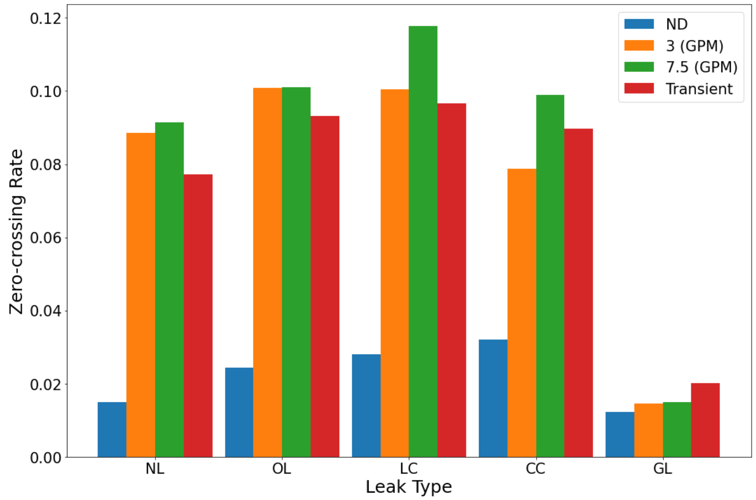

Figure 2.111. Zero-crossing rate plot of accelerometer A1 data in the looped network for all leaks and demands

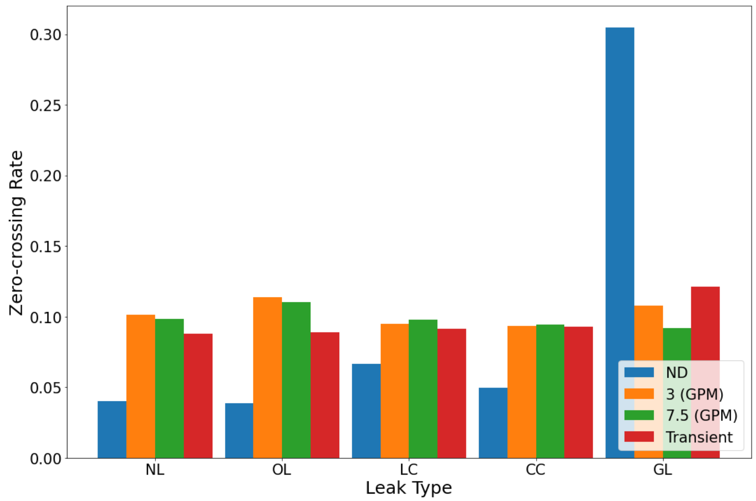

Figure 2.112. Zero-crossing rate plot of accelerometer A1 data in the branched network for all leaks and demands

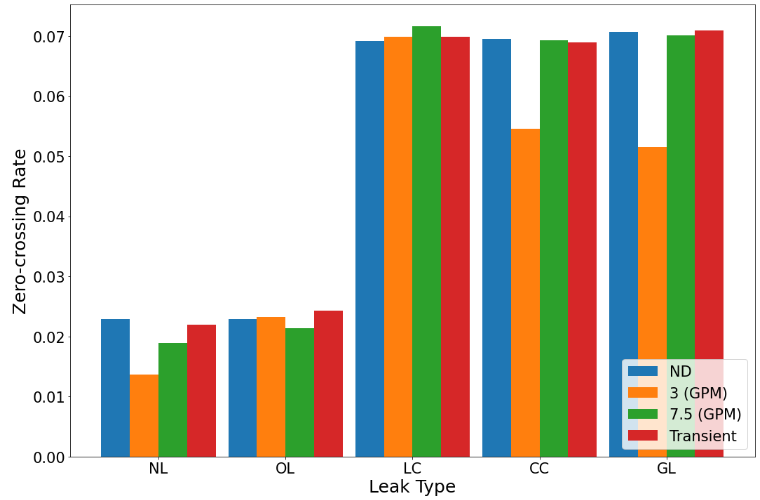

Figure 2.113. Zero-crossing rate plot of accelerometer A2 data in the looped network for all leaks and demands

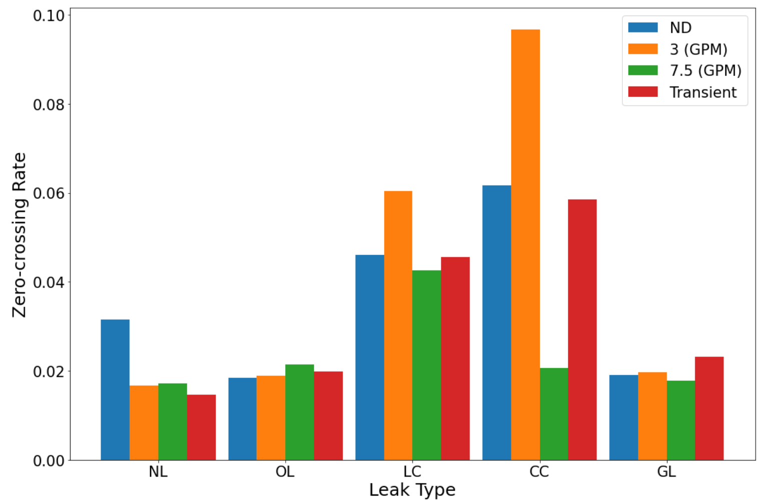

Figure 2.114. Zero-crossing rate plot of accelerometer A2 data in the branched network for all leaks and demands

Table 2.39 compares zero-crossing rate plots of acceleration data measured in the looped and branched networks with all leak and demand types by sensors A1 and A2.

Table 2.39. Analysis of zero-crossing rate plots of acceleration data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Zero-crossing rate	ND – A1	- Comparison of NL signal's zero-crossing rate with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's zero-crossing rate with no demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's zero-crossing rate with no demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's zero-crossing rate with no demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's zero-crossing rate with no demand: $\mathrm{GL_{lo}} \ll \mathrm{GL_{br}}$. - Order of zero-crossing rate with ND in the looped network: CC > LC > OL > NL > GL. - Order of zero-crossing rate with ND in the branched network: GL >> LC > CC > NL > OL.
	3 (GPM) – A1	- Comparison of NL signal's zero-crossing rate with 3 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's zero-crossing rate with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's zero-crossing rate with 3 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's zero-crossing rate with 3 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's zero-crossing rate with 3 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of zero-crossing rate for signals with 3 (GPM) demand in the looped network: OL > LC > NL > CC > GL. - Order of zero-crossing rate for signals with 3 (GPM) demand in the branched network: OL > GL > NL > LC > CC.

Table 2.39. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Zero-crossing rate	7.5 (GPM) – A1	- Comparison of NL signal's zero-crossing rate with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's zero-crossing rate with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's zero-crossing rate with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's zero-crossing rate with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's zero-crossing rate with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the looped network: LC > OL > CC > NL > GL. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the branched network: GL > CC > LC > OL > NL.
	Transient – A1	- Comparison of NL signal's zero-crossing rate with transient demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's zero-crossing rate with transient demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's zero-crossing rate with transient demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's zero-crossing rate with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's zero-crossing rate with transient demand: $GL_{lo}< GL_{br}$. - Order of zero-crossing rate for signals with transient demand in the looped network: LC > OL > CC > NL > GL. - Order of zero-crossing rate for signals with transient demand in the branched network: GL > CC > LC > OL > NL.
	All – A1	- In the looped network and with all demand types, GL has the lowest zero-crossing. - Comparing leaks' zero-crossing rate magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, zero-crossing rate cannot capture the effects of network changes. - Due to the inconsistent order of zero-crossing rate for signals with different demands in both networks, zero-crossing rate is not capable of discerning leak types consistently in both networks.

Table 2.39. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Zero-crossing rate	ND – A2	- Comparison of NL signal's zero-crossing rate with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's zero-crossing rate with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's zero-crossing rate with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's zero-crossing rate with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's zero-crossing rate with no demand: $GL_{lo} > GL_{br}$. - Order of zero-crossing rate with ND in the looped network: GL > CC > LC > NL > OL. - Order of zero-crossing rate with ND in the branched network: CC > LC > NL > GL > OL.
	3 (GPM) – A2	- Comparison of NL signal's zero-crossing rate with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's zero-crossing rate with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's zero-crossing rate with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's zero-crossing rate with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's zero-crossing rate with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of zero-crossing rate for signals with 3 (GPM) demand in the looped network: LC > CC > GL > OL > NL. - Order of zero-crossing rate for signals with 3 (GPM) demand in the branched network: CC > LC > GL > OL > NL.
	7.5 (GPM) – A2	- Comparison of NL signal's zero-crossing rate with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's zero-crossing rate with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's zero-crossing rate with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's zero-crossing rate with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's zero-crossing rate with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the looped network: LC > GL > CC > OL > NL. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the branched network: LC > OL > CC > GL > NL.

Table 2.39. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Zero-crossing rate	Transient – A2	- Comparison of NL signal's zero-crossing rate with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's zero-crossing rate with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's zero-crossing rate with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's zero-crossing rate with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's zero-crossing rate with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of zero-crossing rate for signals with transient demand in the looped network: GL > LC > CC > OL > NL. - Order of zero-crossing rate for signals with transient demand in the branched network: CC > LC > GL > OL > NL.
	All – A2	- Comparing leaks' zero-crossing rate magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, zero-crossing rate cannot capture the effects of network changes. - Due to the inconsistent order of zero-crossing rate for signals with different demands in both networks, zero-crossing rate is not capable of discerning leak types consistently in both networks. - In both networks, where demand is present, NL signals have the lowest zero-crossing rate.

2.12.3.15. RMS

Figures 2.115 and 2.116 show the RMS plots of the accelerometer data measured by sensor A1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.117 and 2.118 show the same plots but for sensor A2 data.

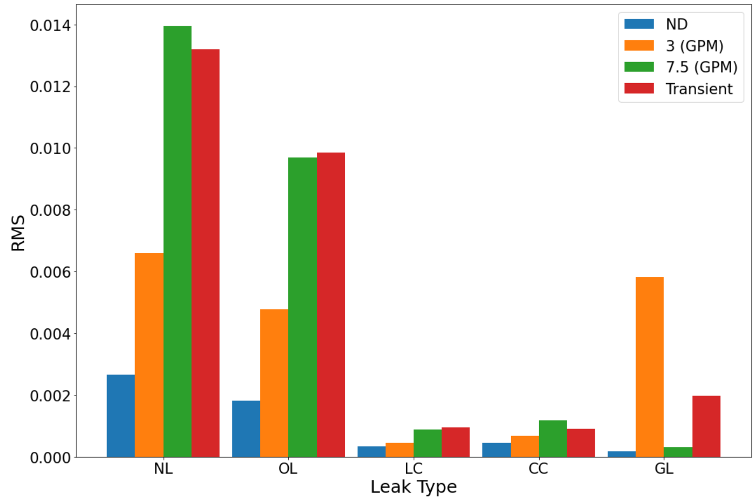

Figure 2.115. RMS plot of accelerometer A1 data in the looped network for all leaks and demands

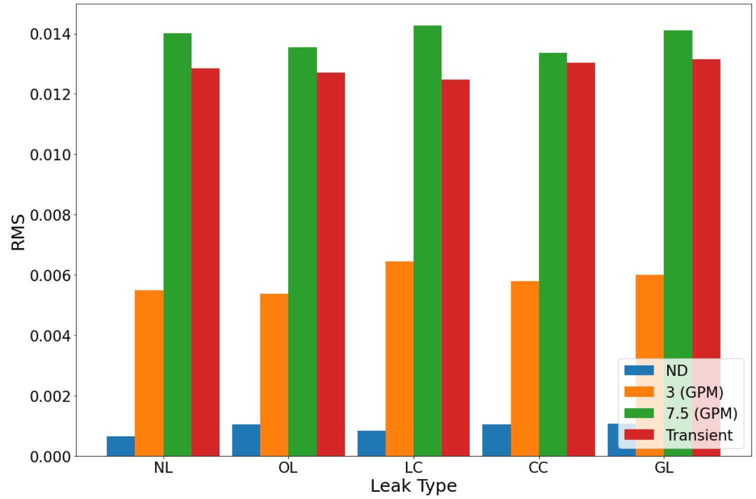

Figure 2.116. RMS plot of accelerometer A1 data in the branched network for all leaks and demands

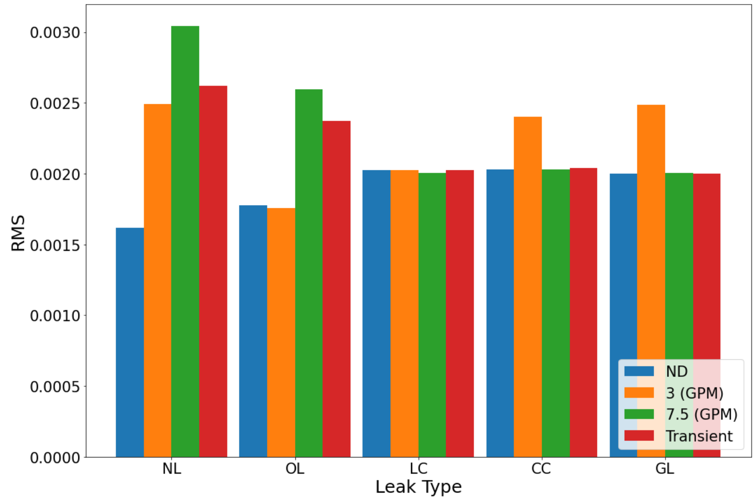

Figure 2.117. RMS plot of accelerometer A2 data in the looped network for all leaks and demands

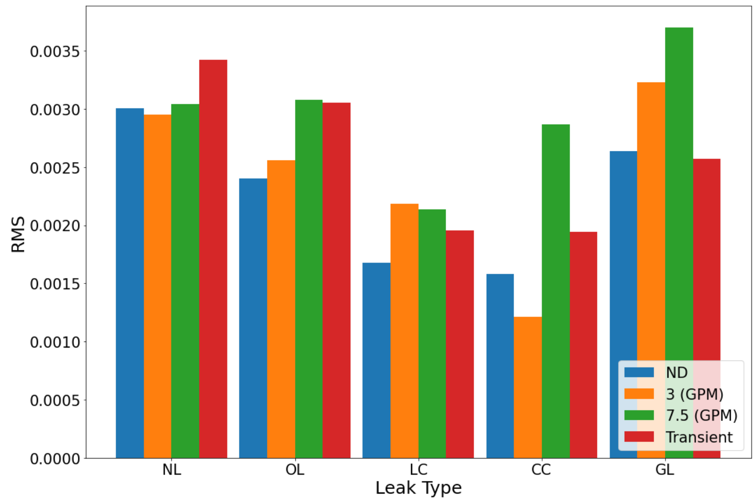

Figure 2.118. RMS plot of accelerometer A2 data in the branched network for all leaks and demands

Table 2.40 compares RMS plots of acceleration data measured in the looped and branched networks with all leak and demand types by sensors A1 and A2.

Table 2.40. Analysis of RMS plots of acceleration data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
RMS	ND – A1	- Comparison of NL signal's RMS with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's RMS with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's RMS with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's RMS with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's RMS with no demand: $GL_{lo} > GL_{br}$. - Order of RMS with ND in the looped network: NL > OL > CC > LC > GL. - Order of RMS with ND in the branched network: GL > CC > OL > LC > NL.
	3 (GPM) – A1	- Comparison of NL signal's RMS with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's RMS with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's RMS with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's RMS with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of RMS for signals with 3 (GPM) demand in the looped network: NL > GL > OL > CC > LC. - Order of RMS for signals with 3 (GPM) demand in the branched network: LC > GL > CC > NL > OL.

Table 2.40. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
RMS	7.5 (GPM) – A1	- Comparison of NL signal's RMS with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's RMS with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's RMS with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's RMS with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of RMS for signals with 7.5 (GPM) demand in the looped network: NL > OL > CC > LC > GL. - Order of RMS for signals with 7.5 (GPM) demand in the branched network: LC > GL > NL > OL > CC.
	Transient – A1	- Comparison of NL signal's RMS with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's RMS with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with transient demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's RMS with transient demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's RMS with transient demand: $GL_{lo} < GL_{br}$. - Order of RMS for signals with transient demand in the looped network: NL > OL > GL > LC > CC. - Order of RMS for signals with transient demand in the branched network: GL > CC > NL > LC > OL.
	All – A1	- In the branched network and for each leak type, RMS follows this pattern: $RMS_{7.5\ (GPM)} > RMS_{transient} > RMS_{3\ (GPM)} > RMS_{ND}$. This is the same pattern for standard deviation (SD) in the branched network and each leak type. - Comparing leaks' RMS magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, RMS cannot capture the effects of network changes. - Due to the inconsistent order of RMS for signals with different demands in both networks, RMS is not capable of discerning leak types consistently in both networks.

Table 2.40. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
RMS	ND – A2	- Comparison of NL signal's RMS with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's RMS with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's RMS with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's RMS with no demand: $GL_{lo} < GL_{br}$. - Order of RMS with ND in the looped network: CC > LC > GL > OL > NL. - Order of RMS with ND in the branched network: NL > GL > OL > LC > CC.
	3 (GPM) – A2	- Comparison of NL signal's RMS with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's RMS with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's RMS with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's RMS with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of RMS for signals with 3 (GPM) demand in the looped network: NL > GL > CC > LC > OL. - Order of RMS for signals with 3 (GPM) demand in the branched network: GL > NL > OL > LC > CC.
	7.5 (GPM) – A2	- Comparison of NL signal's RMS with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's RMS with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's RMS with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's RMS with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of RMS for signals with 7.5 (GPM) demand in the looped network: NL > OL > CC > GL > LC. - Order of RMS for signals with 7.5 (GPM) demand in the branched network: GL > OL > NL > CC > LC.

Table 2.40. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
RMS	Transient – A2	- Comparison of NL signal’s RMS with transient demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal’s RMS with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal’s RMS with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal’s RMS with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal’s RMS with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of RMS for signals with transient demand in the looped network: NL > OL > CC > LC > GL. - Order of RMS for signals with transient demand in the branched network: NL > OL > GL > LC > CC.
	All – A2	- Based on sensor A2 in the looped network, when demand is present, NL signal has the largest RMS magnitude. - Comparing leaks’ RMS magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, RMS cannot capture the effects of network changes. - Due to the inconsistent order of RMS for signals with different demands in both networks, RMS is not capable of discerning leak types consistently in both networks.

2.12.3.16. Crest Factor

Figures 2.119 and 2.120 show the crest factor plots of the accelerometer data measured by sensor A1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.121 and 2.122 show the same plots but for sensor A2 data.

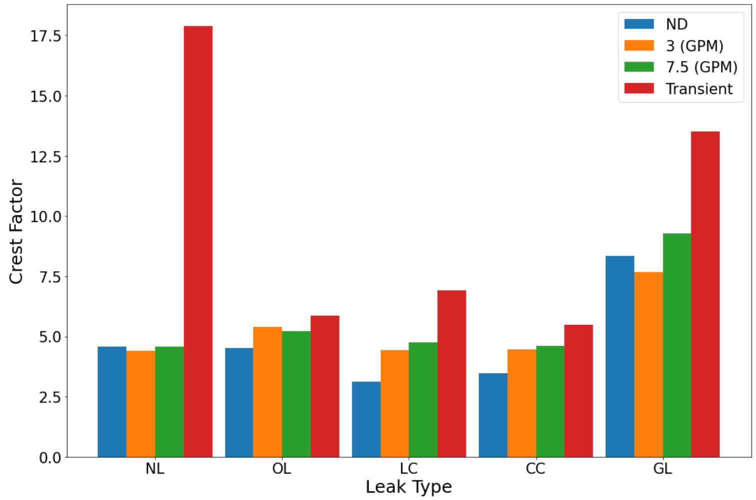

Figure 2.119. Crest factor plot of accelerometer A1 data in the looped network for all leaks and demands

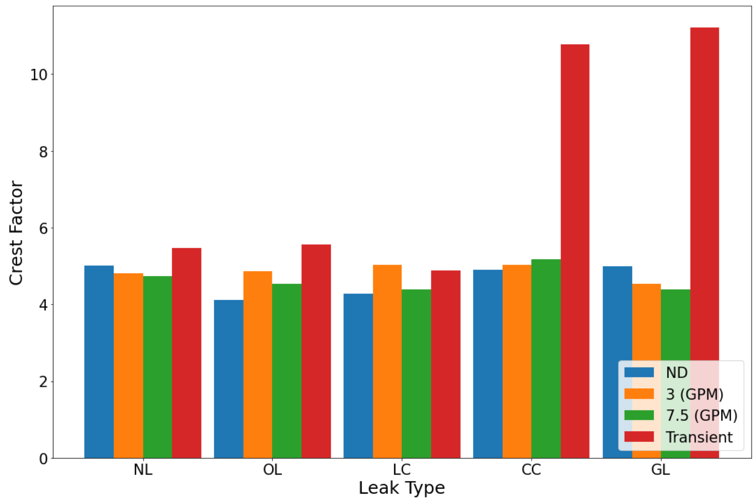

Figure 2.120. Crest factor plot of accelerometer A1 data in the branched network for all leaks and demands

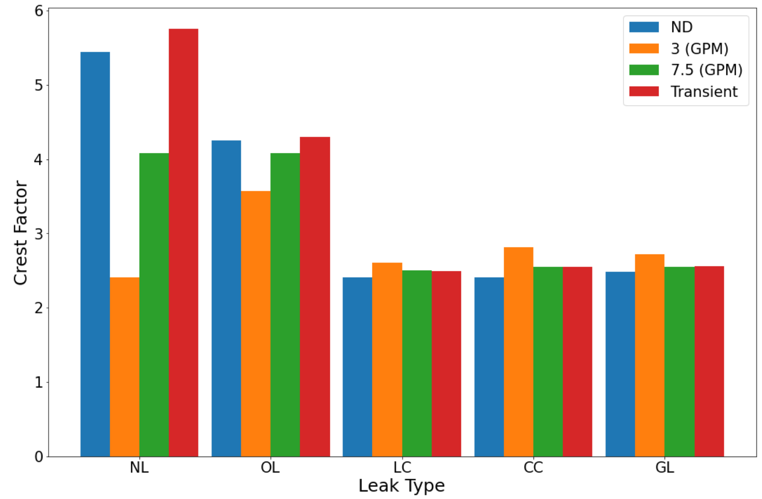

Figure 2.121. Crest factor plot of accelerometer A2 data in the looped network for all leaks and demands

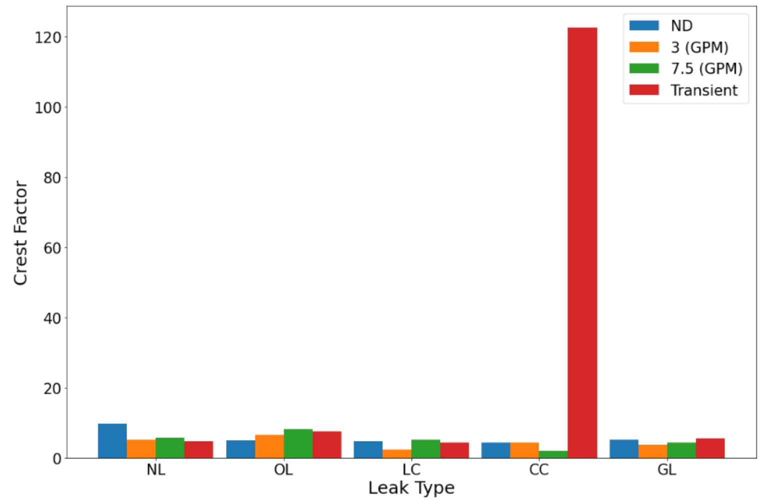

Figure 2.122. Crest factor plot of accelerometer A2 data in the branched network for all leaks and demands

Table 2.41 compares crest factor plots of acceleration data measured in the looped and branched networks with all leak and demand types by sensors A1 and A2.

Table 2.41. Analysis of crest factor plots of acceleration data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Crest factor	ND – A1	- Comparison of NL signal's crest factor with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's crest factor with no demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's crest factor with no demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's crest factor with no demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's crest factor with no demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of crest factor with ND in the looped network: GL > NL > OL > CC > LC. - Order of crest factor with ND in the branched network: NL > GL > CC > LC > OL.
	3 (GPM) – A1	- Comparison of NL signal's crest factor with 3 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's crest factor with 3 (GPM) demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's crest factor with 3 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's crest factor with 3 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's crest factor with 3 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of crest factor for signals with 3 (GPM) demand in the looped network: GL > OL > CC > NL > LC. - Order of crest factor for signals with 3 (GPM) demand in the branched network: CC > LC > OL > NL > GL.

Table 2.41. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Crest factor	7.5 (GPM) – A1	- Comparison of NL signal's crest factor with 7.5 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's crest factor with 7.5 (GPM) demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's crest factor with 7.5 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's crest factor with 7.5 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's crest factor with 7.5 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of crest factor for signals with 7.5 (GPM) demand in the looped network: GL > OL > LC > CC > NL. - Order of crest factor for signals with 7.5 (GPM) demand in the branched network: CC > NL > OL > LC > GL.
	Transient – A1	- Comparison of NL signal's crest factor with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's crest factor with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's crest factor with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's crest factor with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's crest factor with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of crest factor for signals with transient demand in the looped network: NL > GL > LC > OL > CC. - Order of crest factor for signals with transient demand in the branched network: GL > CC > OL > NL > LC.
	All – A1	- In the looped network and for each leak type, the crest factor of signal with transient demand is larger than that of signals with ND, 3 (GPM), and 7.5 (GPM) demands. - Comparing leaks' crest factor magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, the crest factor cannot capture the effects of network changes. - Though the crest factor of signal with transient demand in the looped network is dominant, the order of crest factor for signals with different demands in both networks are inconsistent. Therefore, crest factor is not capable of discerning leak types consistently in both networks.

Table 2.41. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Crest factor	ND – A2	- Comparison of NL signal's crest factor with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's crest factor with no demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's crest factor with no demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's crest factor with no demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's crest factor with no demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of crest factor with ND in the looped network: NL > OL > GL > LC > CC. - Order of crest factor with ND in the branched network: NL > GL > OL > LC > CC.
	3 (GPM) – A2	- Comparison of NL signal's crest factor with 3 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's crest factor with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's crest factor with 3 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's crest factor with 3 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's crest factor with 3 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of crest factor for signals with 3 (GPM) demand in the looped network: OL > CC > GL > LC > NL. - Order of crest factor for signals with 3 (GPM) demand in the branched network: OL > NL > CC > GL > LC.
	7.5 (GPM) – A2	- Comparison of NL signal's crest factor with 7.5 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's crest factor with 7.5 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's crest factor with 7.5 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's crest factor with 7.5 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's crest factor with 7.5 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of crest factor for signals with 7.5 (GPM) demand in the looped network: NL > OL > CC > GL > LC. - Order of crest factor for signals with 7.5 (GPM) demand in the branched network: OL > NL > LC > GL > CC.

Table 2.41. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Crest factor	Transient – A2	- Comparison of NL signal's crest factor with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's crest factor with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's crest factor with transient demand: $\mathrm{CC_{lo}} \ll \mathrm{CC_{br}}$. - Comparison of LC signal's crest factor with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's crest factor with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of crest factor for signals with transient demand in the looped network: NL > OL > GL > CC > LC. - Order of crest factor for signals with transient demand in the branched network: CC >> OL > GL > NL > LC.
	All – A2	- Comparing leaks' crest factor magnitudes of two networks shows no consistent change pattern in the magnitudes when networks change. Therefore, crest factor cannot capture the effects of network changes. - Due to the inconsistent order of crest factor for signals with different demands in both networks, crest factor is not capable of discerning leak types consistently in both networks.

2.12.3.17. Summary of Accelerometer Measurement Evaluations

Due to the inconsistent patterns and similar magnitudes of the plots and features, the sixteen evaluation criteria could not discern the leak types or the network change using the accelerometer measurements. The patterns and magnitudes of the two sensors' acceleration data were not often similar. The following are some important takeaways from the evaluations of the acceleration data.

Time-domain magnitude dominance of leak and no leak signals depends on the sensors' location. By using A1, when there is no demand and in the looped network, NL signals' amplitudes are larger than leak signals' amplitudes; however, in the branched network, leak signals’ amplitudes are larger than those of NL. On the other hand, based on A2 and in the branched network, except for the GL, NL signal amplitudes are larger than those of leak signals.

Frequency-domain plots indicate that the amplitudes of the leaks’ frequencies directly relate to the leaks’ flow rates in the looped network where frequency caps of the leak and no leak signals were the same, 500 Hz. But in the branched network, the frequency caps of leak signals are larger than those of no leak signals. Amplitudes of NL frequencies are larger than those of leaks’ frequencies in the looped network and vice versa in the branched network. A reason for the larger amplitudes of the NL frequencies in the looped network can be a decreasing vibration energy effect of a leak in the looped network with more connected pipes.

The frequency of acceleration data depends on the sensor locations. For example, based on A2, the leak signals’ frequency caps in the looped network were larger than those in the branched network. This relation is contrary to that based on A1 signals. Regarding the A1 measurements, the frequency caps of leak signals in the looped network are 500 Hz, but those frequencies reach 6000 Hz when measured by A2. Since A2 was mounted on a T-junction connected to a hydrant, the hydrant's dead-end effect increased flow turbulence in the pipe, and the fluid-pipe coupling transfers this larger turbulence to increased vibrations sensed by A2. Therefore, mounting accelerometers close to pipeline dead-ends may lead to measured vibrations that do not represent the whole network's vibrations.

Similar to the time-domain plots, the location of the accelerometers affects signals’ cumulative distribution plots. Based on A1, when there is no demand and in the looped network, NL signals’ magnitudes are larger than leak signals’ magnitudes; however, in the branched network, leak signals’ magnitudes are larger than NLs’. On the other hand, based on A2, in the looped network NL signal’s cumulative distribution plot has larger values than the leaks and vice versa in the branched network.

Similar to the time-domain and the cumulative distribution plots, the accelerometers' locations affect signals’ box plots. Based on A1, when there is no demand in the looped network, the NL magnitude continuum is larger than the leak signals’ magnitude continuum; however, in the branched network, the leak signals’ magnitude continuum is larger than NLs’.

The looped network's NL signal includes the largest range of acceleration magnitudes than all other signals of the looped and branched networks. The water pressurized by the pump transfers its stress to the pipes via the pipe-flow coupling. Assuming pipes are beams, this stress acts as a distributed load that causes vibrations in the pipeline. On the other hand, based on Eq. 2.23, a leak’s thrust force generates pipe vibrations. In the looped network without a leak, the vibration caused by water pressure is maximum. When we induced a leak, though it generated additional vibrations, the pipe vibration decreases because a leak reduces the water pressure. Therefore, one can conclude that a leak's decreasing vibration effect is more dominant than its additive vibration effect caused by the thrust force. This could be the reason for the smaller vibration magnitudes when there is a leak in the looped network with no demand. This reasoning may not be the case in actual water networks. In this research, leaks had small sizes to maintain the leak flows less than 30% of total water input. Also, the testbed’s water pressure might be less than its actual counterparts. These two factors that stem from the model distortion may cause a smaller vibration caused by leaks than that generated by water pressure in the testbed. However, since water pressure and leak sizes can be larger in actual water networks, leak vibrations can be larger than the water pressure vibrations. In the branched network with no demand, water pressure at the acceleration location might be less than that in the looped network since pipes are less connected. Therefore, leak vibrations exceeded the water pressure vibrations, leading to larger leak signals' accelerations than the NL signal.

There is no consistent pattern in the relations of the two networks' areas under the CSD plots based on accelerometer data under no demand scenario. Therefore, the areas under the CSD plots cannot capture the effects of the network change. The larger area under the CSD plot of the NL signal in the looped network is due to the similarity of spectral density magnitudes of the NL signals measured by sensors A1 and A2. For example, the NL signals' cap frequencies at A1 and A2 in the looped network were 500 Hz and 600 Hz, respectively. However, for leak signals in the looped network, different spectral contents measured by two sensors led to smaller areas under the CSD plots. One can observe the difference where the leaks' maximum frequency at A1 was 500 Hz and at A2 reached 6000 Hz. On the other hand, in the branched network, the maximum leak frequencies at two sensors were 600 Hz, indicating the signals' spectral similarity and a reason for the larger areas under the CSD plots.

Since the LDIs of leaks are larger than those of the benchmark, i.e., NL, LDI can detect leaks. This conforms to the results of the paper of Yazdekhasti et al., 2016. However, due to the similarity of leaks' LDI, especially for CC and LC when there is a demand in both networks, and because of the inconsistent LDI magnitudes with varying demands in each network, LDI cannot discerning leak types. Moreover, due to the magnitudes' similarity, LDI cannot discern network changes when there is a demand flow. On the other hand, the LDIs of ND represented by the blue bars in Figures 2.81 and 2.82 successfully capture the network architecture changes, but it is less likely that an actual water network is without water demands. LDI cannot discern leak and network changes when there is a demand due to all leaks' cross-spectral density similarities in both networks. Though generalizing the following needs more experiments, comparing the scatter plots of the LDIs indicate that (1) in the looped network with a non-zero demand, the LDI magnitude of OL is larger than one; and (2) in the branched network with a non-zero demand, the LDI magnitude of GL is above two.

In the looped network, LDIs are approximately constant when the demand changes. This indicates that the vibration correlation between the two sensors remains the same with demand variations, which stems from the looped network's pipes connectivity that causes the sensors to measure signals with more similar patterns. However, in the branched network, LDI decreases when demand changes from ND to 3 (GPM) and 7.5 (GPM). The change in the branched network's LDIs represents different vibrations measured by the sensors that can root from the less network connectivity in the branched architecture. The decrease in the LDIs corresponds to more resemblance of the CSDs of the leaks and no leak. When there is no demand, the differences between the leak and no leak CSDs are maximum, indicating fewer uniform vibrations throughout the network generated by water pressure. Since the pipes in the branched network are less connected, the pipes' water pressure is less uniform. When demand exists and increases, it balances the water pressure distributions and the network's consequent vibrations.

Based on A1, GL's dominant frequencies for all flow variants in the looped network are zero. Regarding A1 measurements, when demand is present, GL in both networks has the lowest fundamental frequency and power spectral entropy. GL has the lowest dominant and fundamental frequencies due to two reasons: (1) GL was the only leak that did not have a water jet, and its output was flowing water which created the least vibrations; (2) since GL had the largest leak flow rate, it decreased water pressure in the higher-pressure zone and caused a more balanced pressure distribution and acceleration throughout the pipeline.

Based on the power spectral entropy of signals recorded by sensor A2, when there is a demand, the signals' spectral entropies are larger in the branched network than those in the looped network. Therefore, by using sensor A2, power spectral entropy can capture the network change's effects. However, this pattern was not observed in the power spectral entropy of sensor A1 measurements.

Regarding A1 in the branched network and for each leak type, RMS follows this pattern: $\mathrm{RMS}_{7.5\,(\mathrm{GPM})} > \mathrm{RMS}_{\mathrm{transient}} > \mathrm{RMS}_{3\,(\mathrm{GPM})} > \mathrm{RMS}_{\mathrm{ND}}$. This is the same pattern for standard deviation (SD) in the branched network and each leak type.

2.12.4. Hydrophone Measurements

In this subsection, we analyzed the data of the hydrophones H1 and H2, where ND and NS stand for no demand and no background noise, respectively.

2.12.4.1. Time-domain plot (for ND and NS signals)

Figures 2.123 to 2.130 show the time-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor H1.

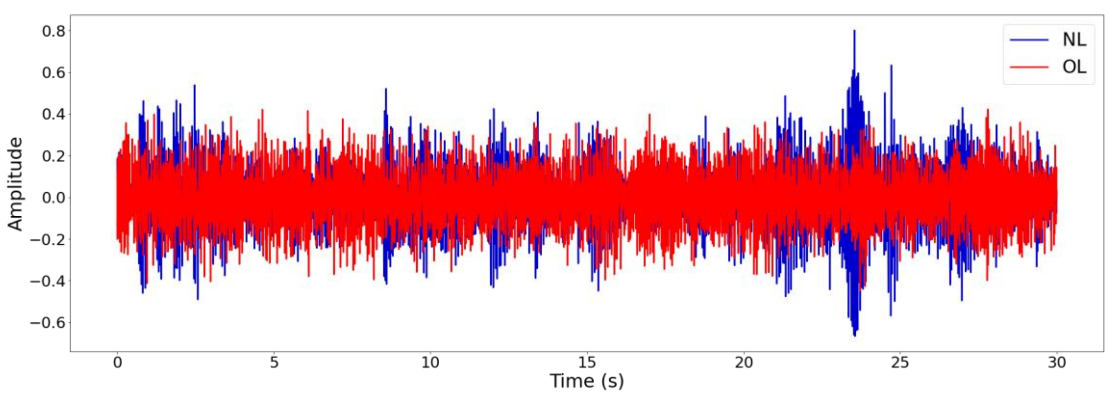

Figure 2.123. Time-domain plots of OL vs. NL signals in the looped network measured by sensor H1

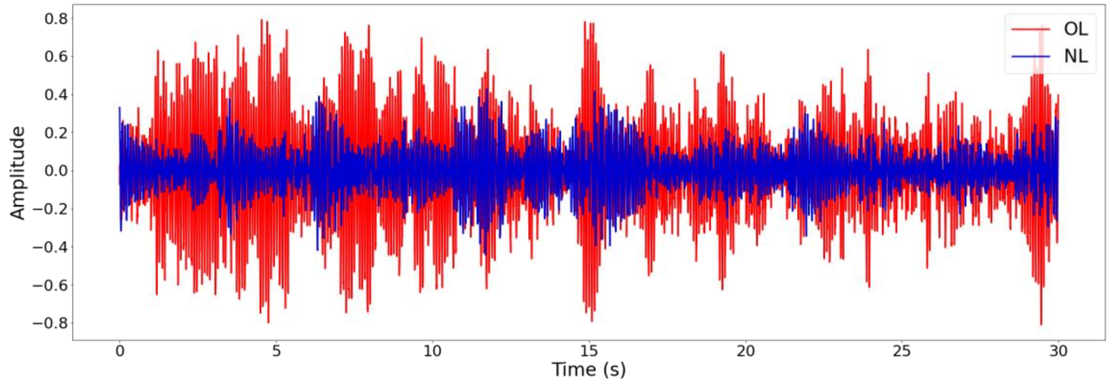

Figure 2.124. Time-domain plots of OL vs. NL signals in the branched network measured by sensor H1

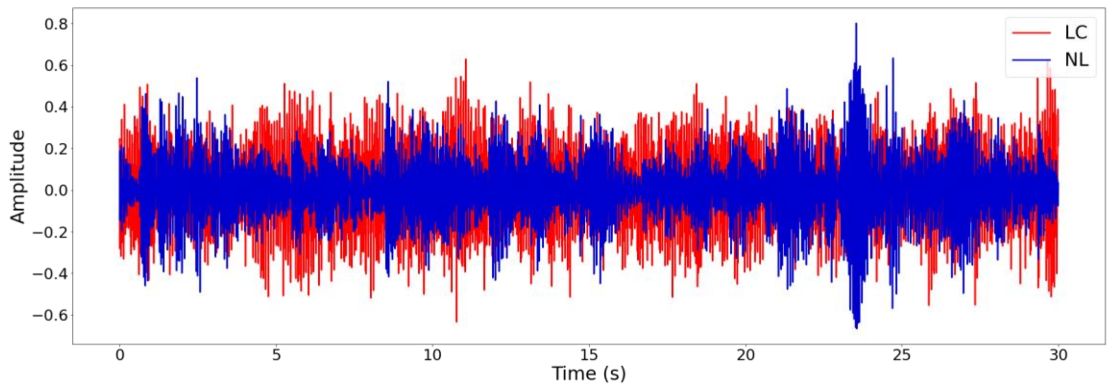

Figure 2.125. Time-domain plots of LC vs. NL signals in the looped network measured by sensor H1

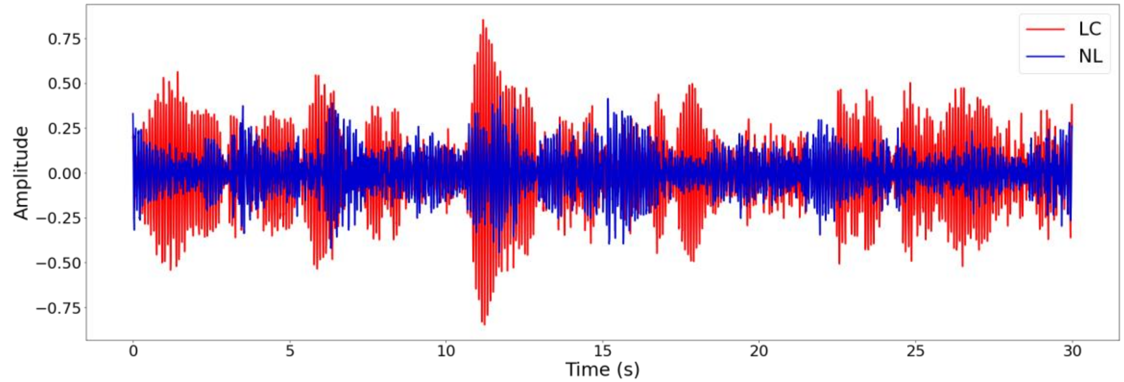

Figure 2.126. Time-domain plots of LC vs. NL signals in the branched network measured by sensor H1

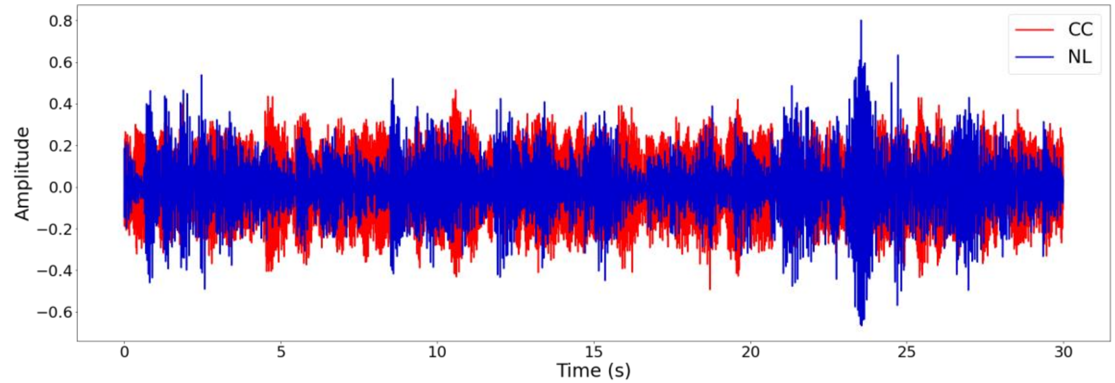

Figure 2.127. Time-domain plots of CC vs. NL signals in the looped network measured by sensor H1

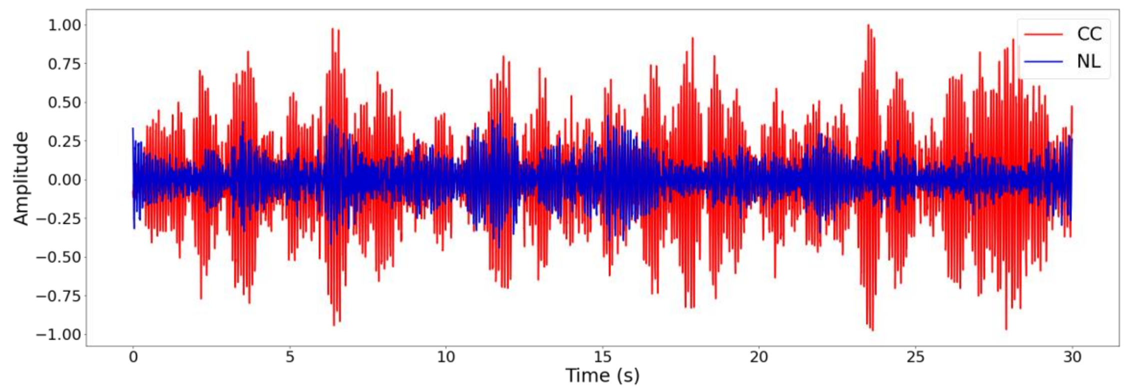

Figure 2.128. Time-domain plots of CC vs. NL signals in the branched network measured by sensor H1

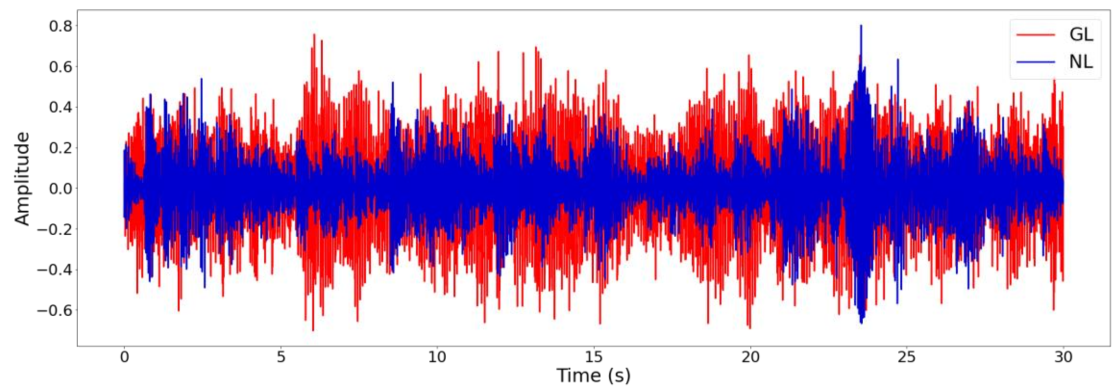

Figure 2.129. Time-domain plots of GL vs. NL signals in the looped network measured by sensor H1

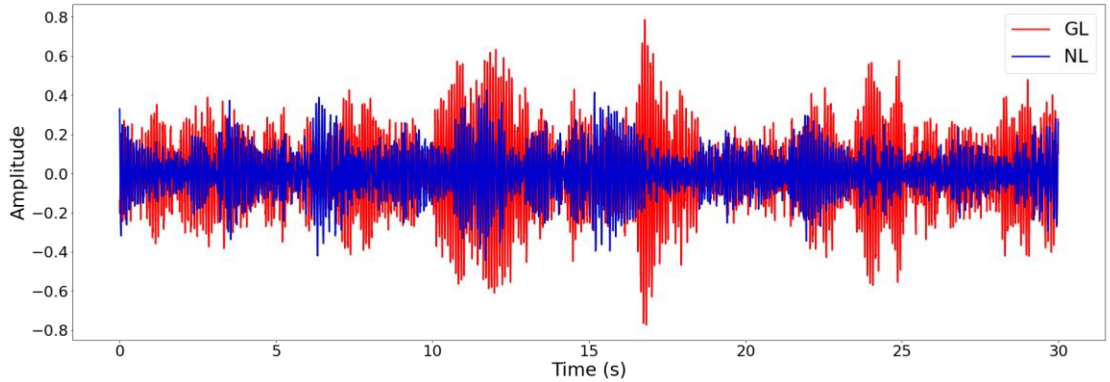

Figure 2.130. Time-domain plots of GL vs. NL signals in the branched network measured by sensor H1

Figures 2.131 to 2.138 show the time-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor H2.

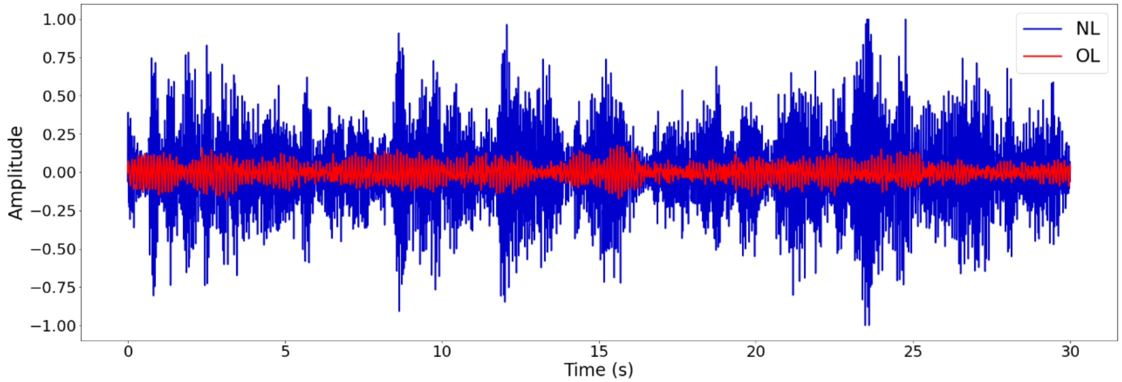

Figure 2.131. Time-domain plots of OL vs. NL signals in the looped network measured by sensor H2

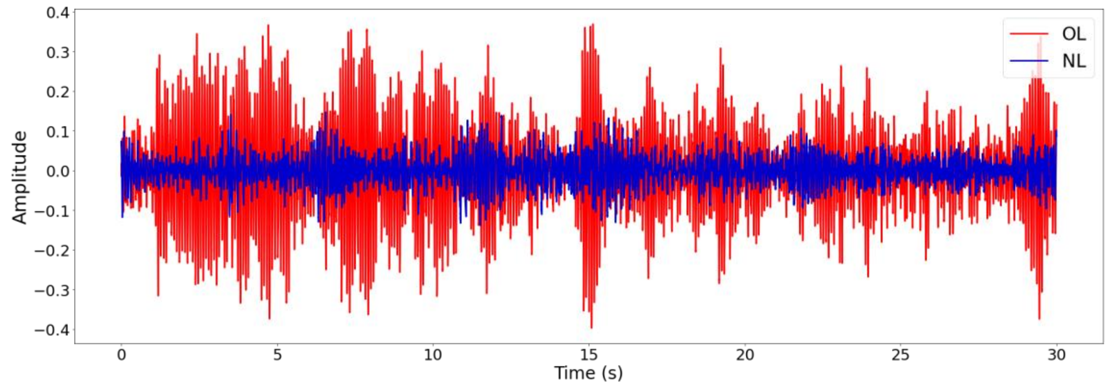

Figure 2.132. Time-domain plots of OL vs. NL signals in the branched network measured by sensor H2

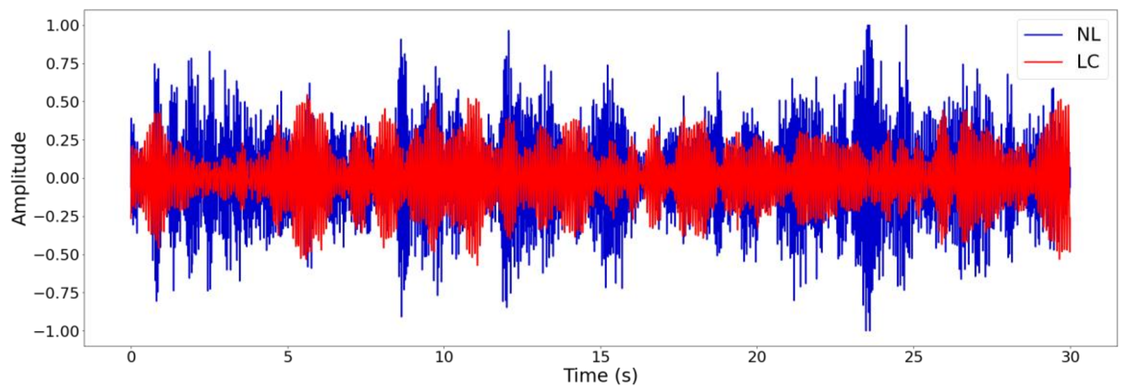

Figure 2.133. Time-domain plots of LC vs. NL signals in the looped network measured by sensor H2

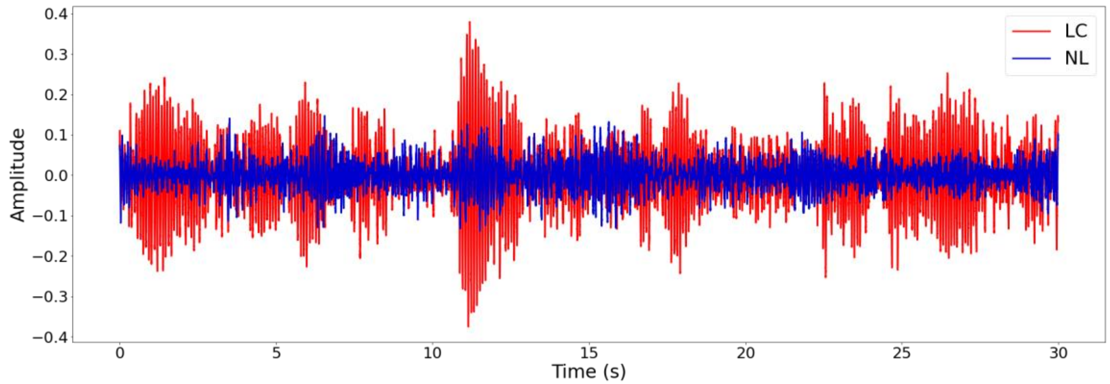

Figure 2.134. Time-domain plots of LC vs. NL signals in the branched network measured by sensor H2

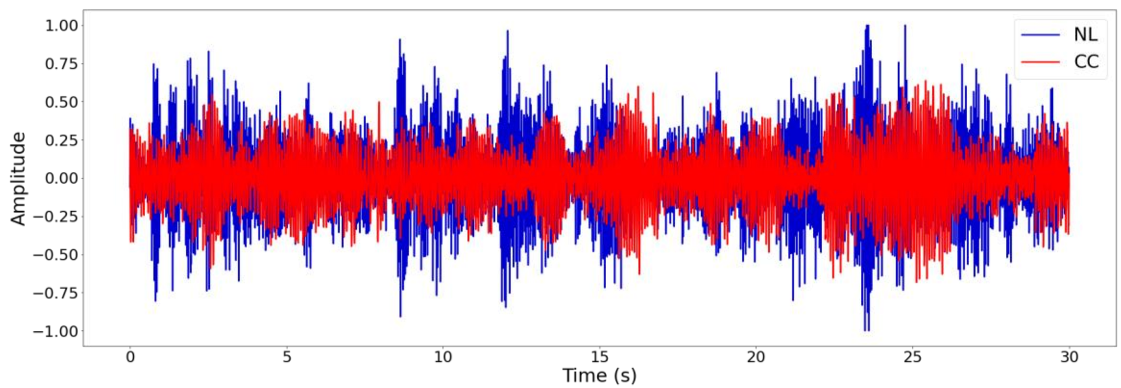

Figure 2.135. Time-domain plots of CC vs. NL signals in the looped network measured by sensor H2

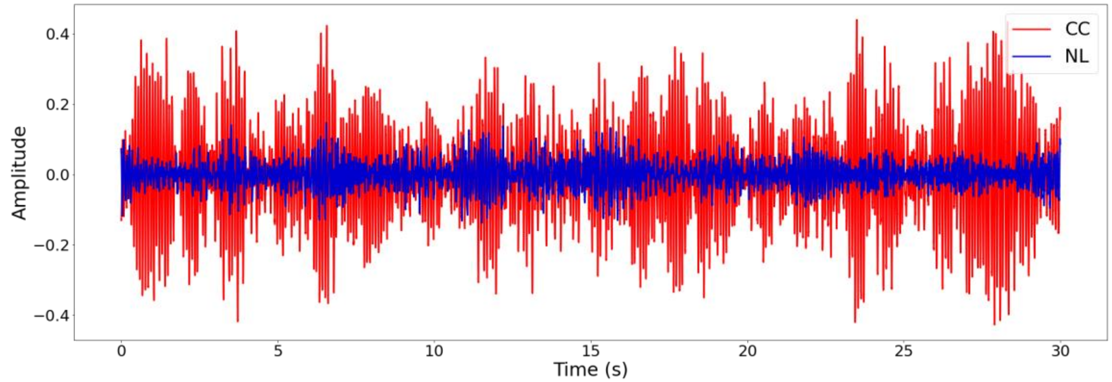

Figure 2.136. Time-domain plots of CC vs. NL signals in the branched network measured by sensor H2

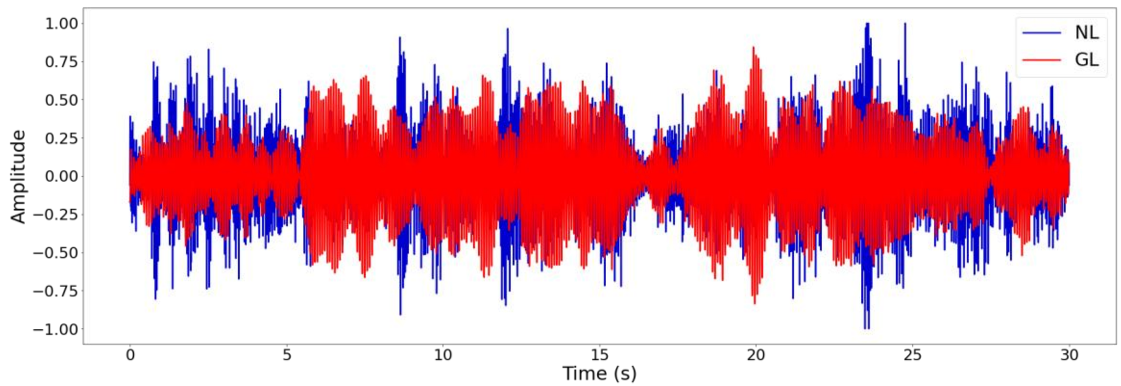

Figure 2.137. Time-domain plots of GL vs. NL signals in the looped network measured by sensor H2

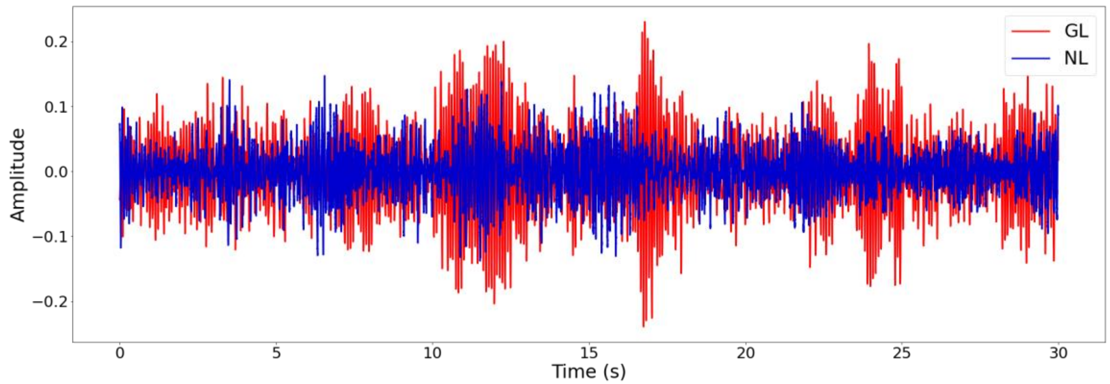

Figure 2.138. Time-domain plots of GL vs. NL signals in the branched network measured by sensor H2

Table 2.42 includes analytical information of the time-domain hydrophone plots where leak and no leak signals of sensors H1 and H2 in the looped and branched networks are visually compared. These measurements do not include background noise to highlight leak signals.

Table 2.42. Analysis of the time-domain hydrophone plots measured by H1 and H2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Time-domain plot (for ND and NS signal)	NL vs. OL – H1	- OL and NL have approximately the same amplitudes in the looped network. - Similar to A1 measurements, OL signal amplitudes are larger than those of NL in the branched network. - NL signal amplitudes are approximately the same in both networks. - The amplitudes of OL in the branched network are larger than those of the looped network.

Table 2.42. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Time-domain plot (for ND and NS signal)	NL vs. LC – H1	- Similar to A1 measurements, NL signal amplitudes are larger in the looped network than those in the branched one. - LC signal amplitudes are more uniform in the looped network than those in the branched one. - Contrary to A1 measurements, LC signal amplitudes are larger than those of NL in the looped network. - Similar to A1 measurements, LC signal amplitudes are larger than those of NL in the branched network. - While, based on A1, LC acceleration amplitudes were close to zero in the looped network, based on H1, LC acoustic amplitudes are large in the looped network.
	NL vs. CC – H1	- Similar to A1 measurements, NL signal amplitudes are larger in the looped network than those in the branched one. - CC signal amplitudes are more uniform in the looped network than those in the branched one. - Contrary to A1 measurements, CC signal amplitudes are larger than those of NL in the looped network. - Similar to A1 measurements, CC signal amplitudes are larger than those of NL in the branched network. - While, based on A1, CC acceleration amplitudes were close to zero in the looped network, based on H1, CC acoustic amplitudes are large in the looped network.
	NL vs. GL – H1	- Similar to A1 measurements, NL signal amplitudes are larger in the looped network than those in the branched one. - Contrary to A1 measurements, GL signal amplitudes are larger in the looped network than those in the branched one. - Contrary to A1 measurements, NL signal amplitudes are smaller than those of GL in the looped network. - Similar to A1 measurements, GL signal amplitudes are larger than those of NL in the branched network.

Table 2.42. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Time-domain plot (for ND and NS signal)	All – H1	- When there is no demand and no background noise, based on H1 in both networks, leak acoustic data are larger than those of NL. Based on A1, the relation of the leak and NL signal magnitudes differed in each network. - When there is no demand and no background noise, based on H1, the comparison of leak signals between the two networks did not indicate any specific pattern. - Like A1 measurements, NL signal amplitudes are larger in the looped network than those in the branched one. - While, based on A1, acceleration amplitudes of the leak signals were visually close to zero in the looped network, based on H1, acoustic amplitudes of the leak signals are large in the looped network.
	NL vs. OL – H2	- Like A1 and H1 measurements, NL signal amplitudes are larger in the looped network than those in the branched one. - Similar to H1 measurements, the amplitudes of OL in the branched network are larger than those of the looped network. - NL signal amplitudes are larger than those of OL in the looped network. - Similar to A1 and H1 measurements, OL signal amplitudes are larger than those of NL in the branched network.
	NL vs. LC – H2	- Similar to H1 measurements, NL signal amplitudes are larger in the looped network than those in the branched one. - Similar to H1 measurements, LC signal amplitudes are more uniform in the looped network than those in the branched one. - Similar to H1 measurements, NL signal amplitudes are larger than those of LC in the looped network. - Contrary to A2 but similar to H1 measurements, LC signal amplitudes are larger than those of NL in the branched network.

Table 2.42. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Time-domain plot (for ND and NS signal)	NL vs. CC – H2	- Similar to H1 measurements, NL signal amplitudes are larger in the looped network than those in the branched one. - Similar to H1 measurements, CC signal amplitudes are more uniform in the looped network than those in the branched one. - Contrary to H1 measurements, NL signal amplitudes are larger than those of CC in the looped network. - Similar to H1 measurements, CC signal amplitudes are larger than those of NL in the branched network.
	NL vs. GL – H2	- Similar to H1 and contrary to A2 measurements, NL signal amplitudes are larger in the looped network than those in the branched one. - Similar to H1 and contrary to A2 measurements, GL signal amplitudes are larger in the looped network than those in the branched one. - Contrary to H1 measurements, comparing NL and GL magnitudes does not indicate a specific pattern in the looped network. - Similar to H1 measurements, GL signal amplitudes are larger than those of NL in the branched network.
	All – H2	- When there is no demand and no background noise, based on H2 in the branched network, leak acoustic data are visually larger than those of NL. However, one cannot observe a consistent relation between leak and NL signals in the looped network. - Similar to H1 measurements, when there is no demand and no background noise, based on H2, the comparison of leak signals between the two networks did not indicate any specific pattern. - Similar to H1 measurements, NL signal amplitudes are larger in the looped network than those in the branched one.

2.12.4.2. Frequency-domain plot (for ND and NS signals)

Figures 2.139 to 2.146 show the frequency-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor H1.

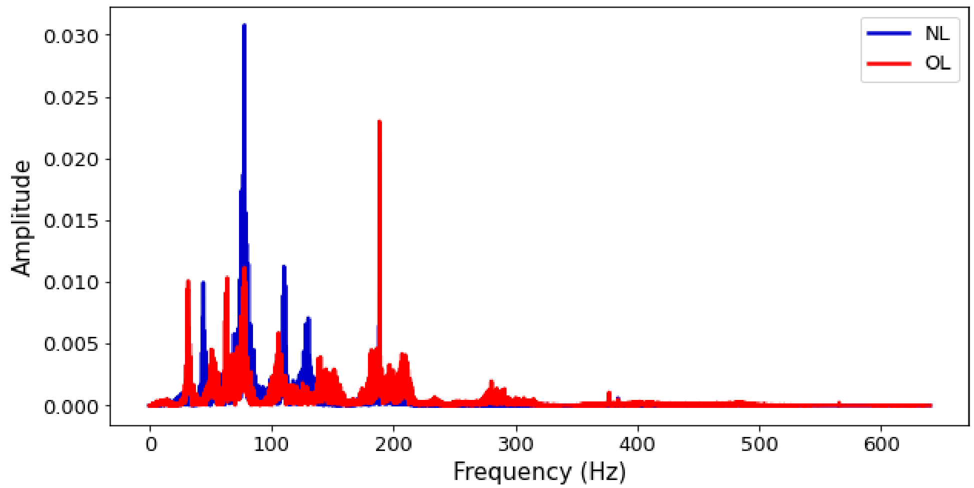

Figure 2.139. Frequency-domain plots of OL vs. NL signals in the looped network measured by sensor H1

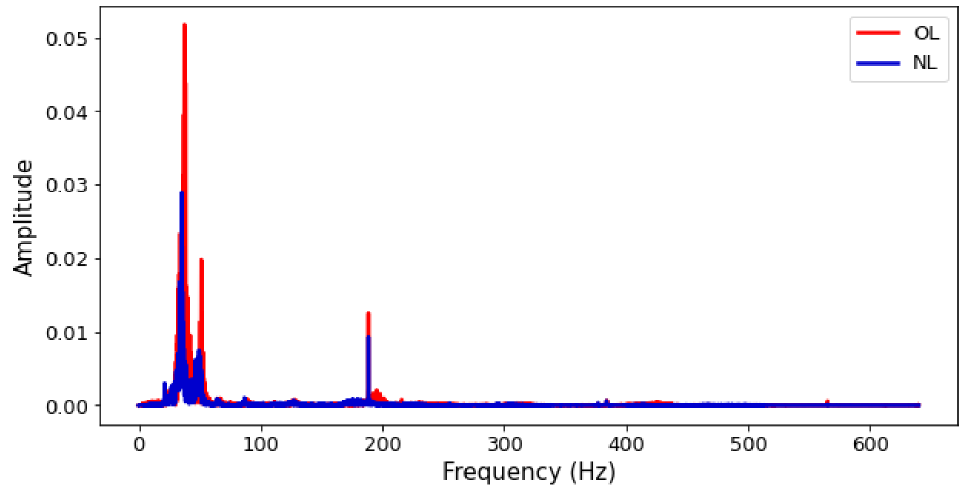

Figure 2.140. Frequency-domain plots of OL vs. NL signals in the branched network measured by sensor H1

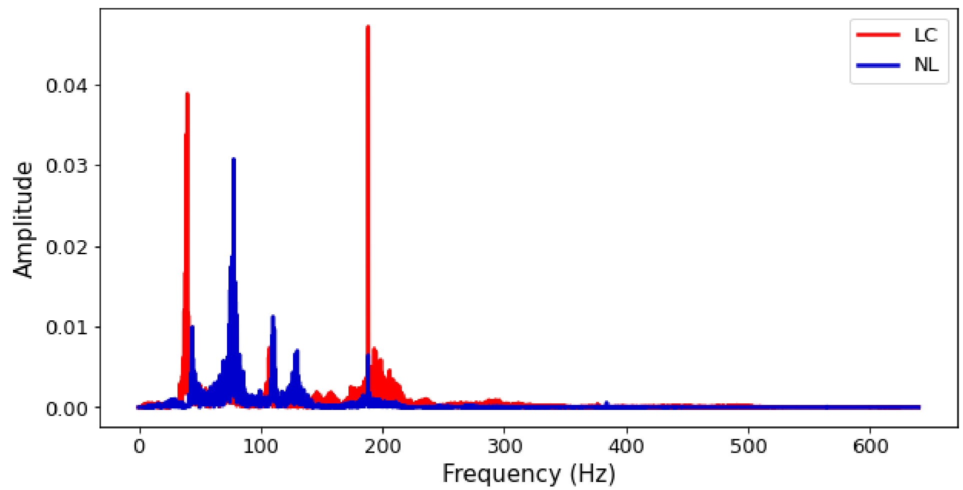

Figure 2.141. Frequency-domain plots of LC vs. NL signals in the looped network measured by sensor H1

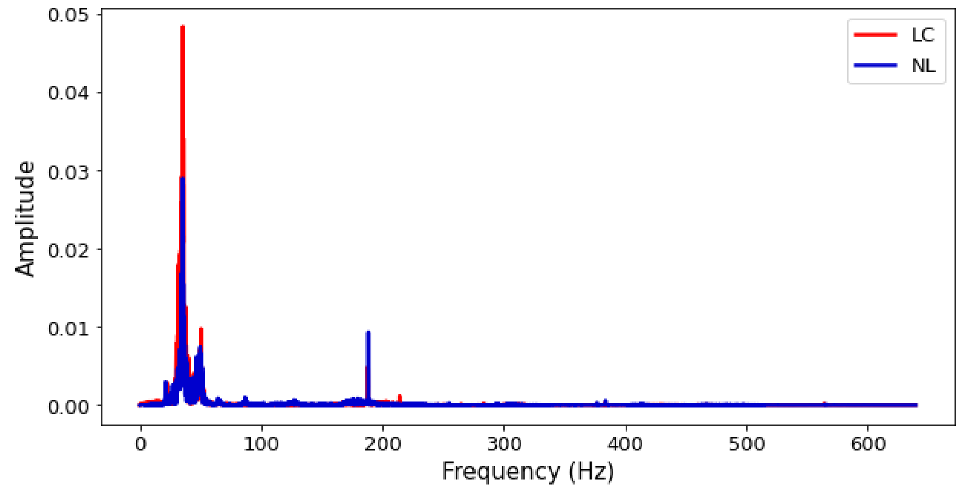

Figure 2.142. Frequency-domain plots of LC vs. NL signals in the branched network measured by sensor H1

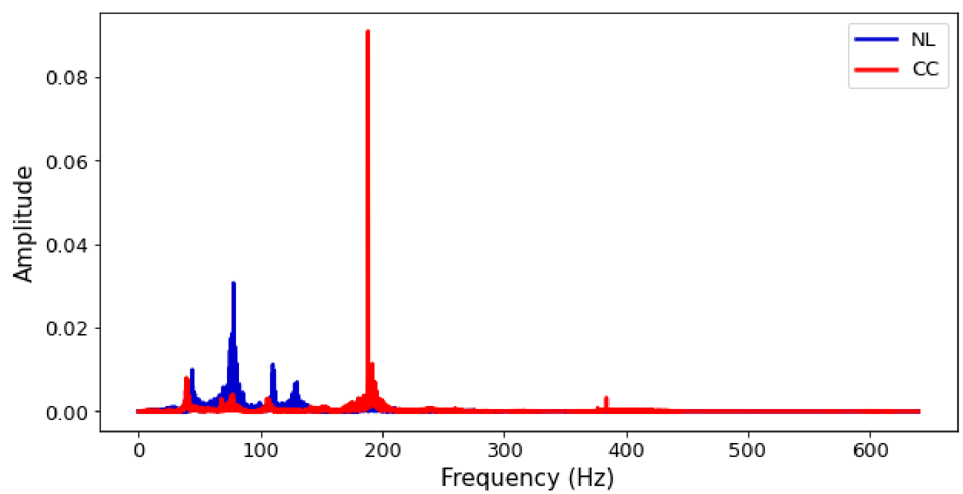

Figure 2.143. Frequency-domain plots of CC vs. NL signals in the looped network measured by sensor H1

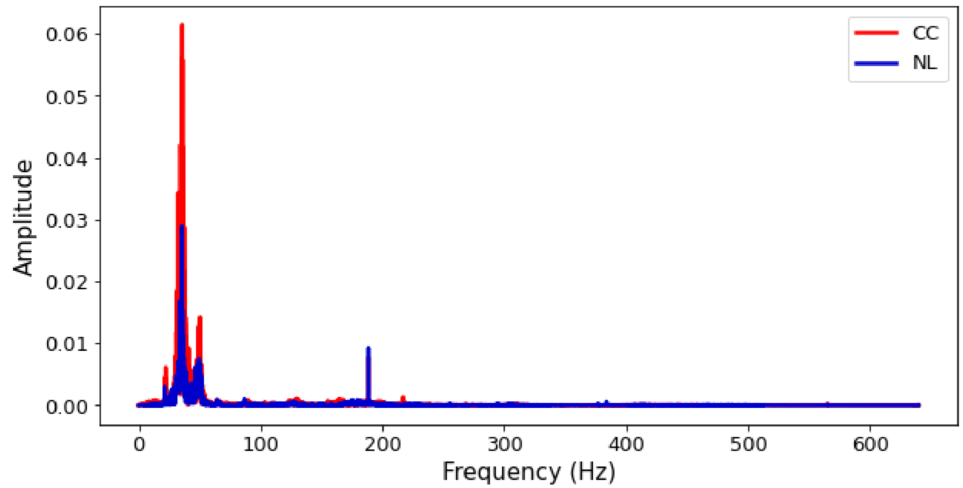

Figure 2.144. Frequency-domain plots of CC vs. NL signals in the branched network measured by sensor H1

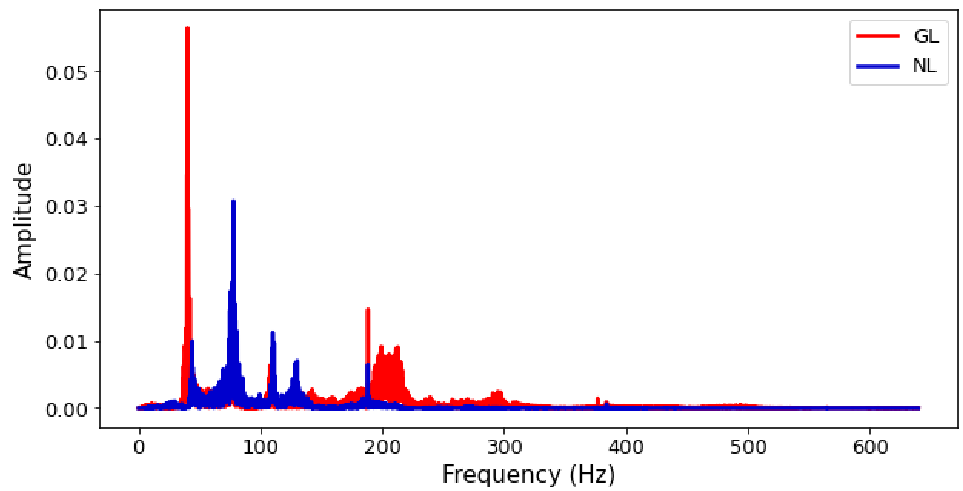

Figure 2.145. Frequency-domain plots of GL vs. NL signals in the looped network measured by sensor H1

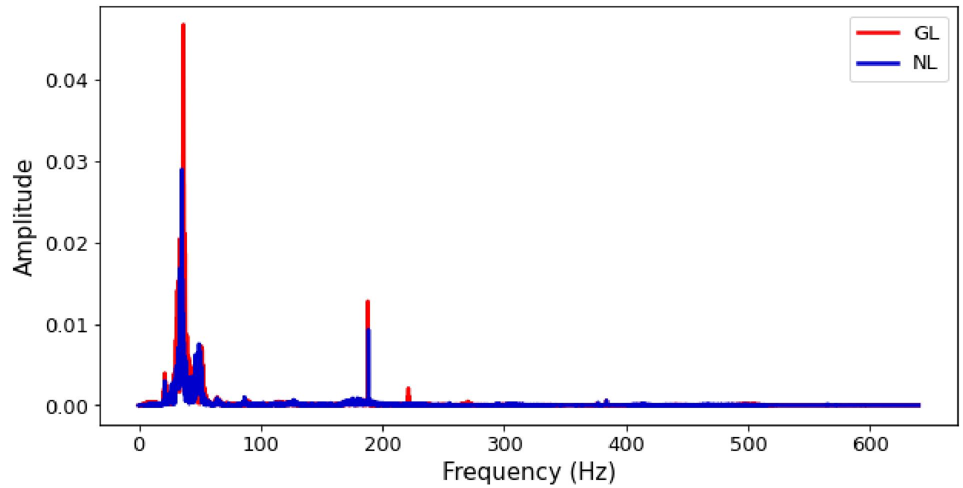

Figure 2.146. Frequency-domain plots of GL vs. NL signals in the branched network measured by sensor H1

Figures 2.147 to 2.154 show the Frequency-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor H2.

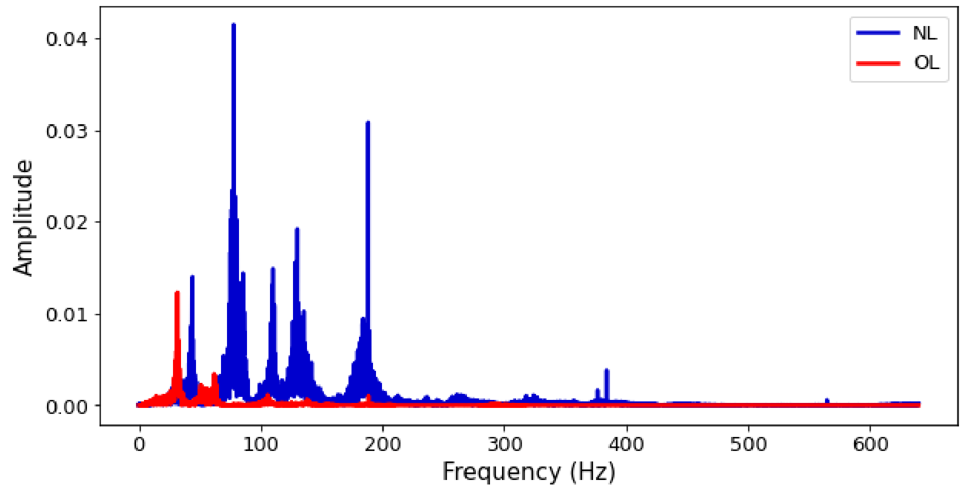

Figure 2.147. Frequency-domain plots of OL vs. NL signals in the looped network measured by sensor H2

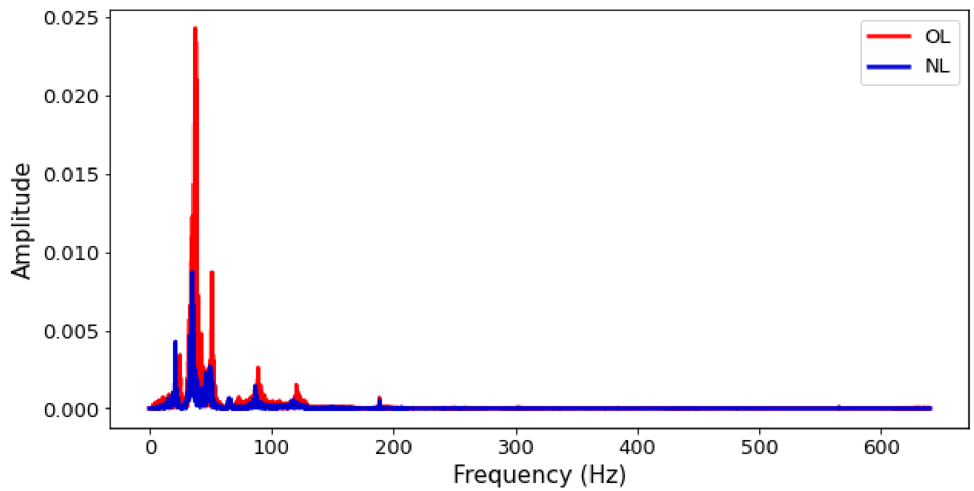

Figure 2.148. Frequency-domain plots of OL vs. NL signals in the branched network measured by sensor H2

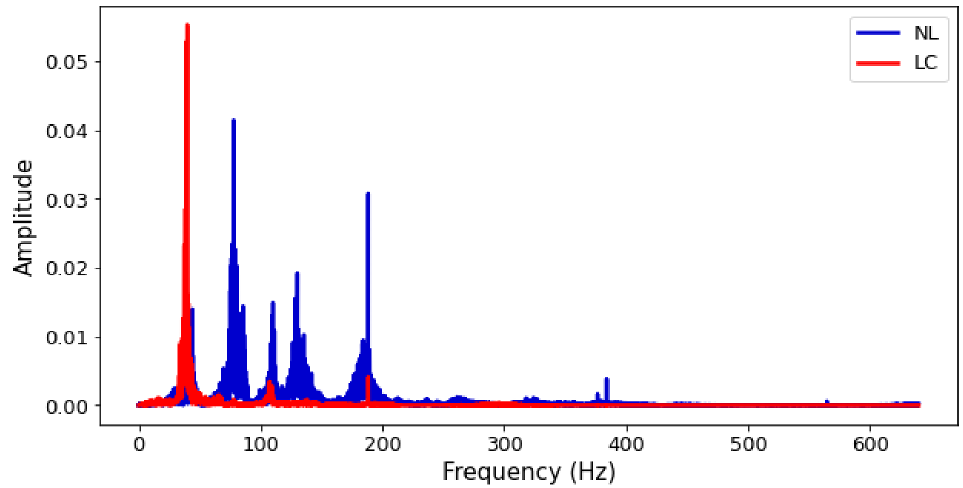

Figure 2.149. Frequency-domain plots of LC vs. NL signals in the looped network measured by sensor H2

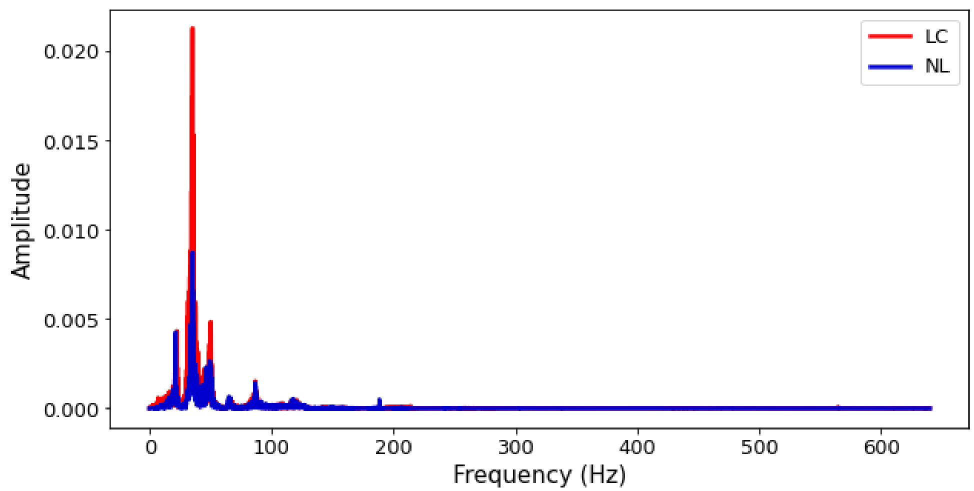

Figure 2.150. Frequency-domain plots of LC vs. NL signals in the branched network measured by sensor H2

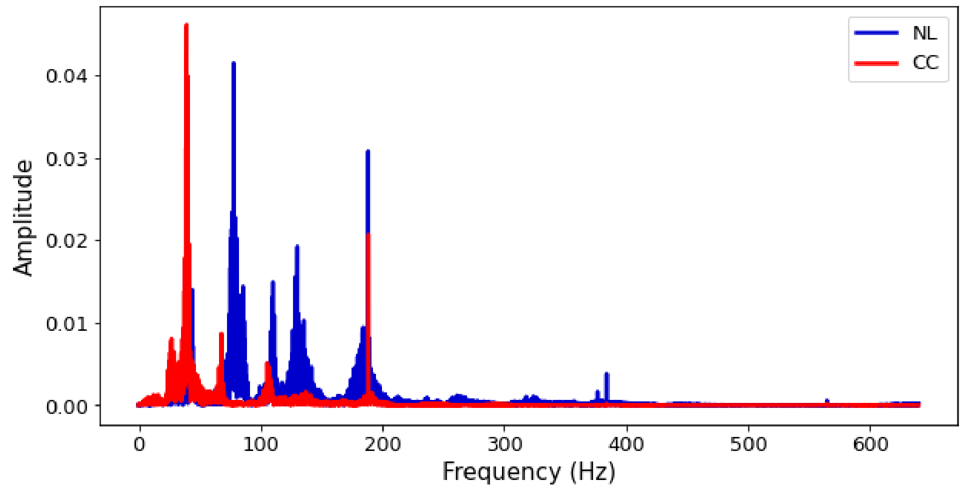

Figure 2.151. Frequency-domain plots of CC vs. NL signals in the looped network measured by sensor H2

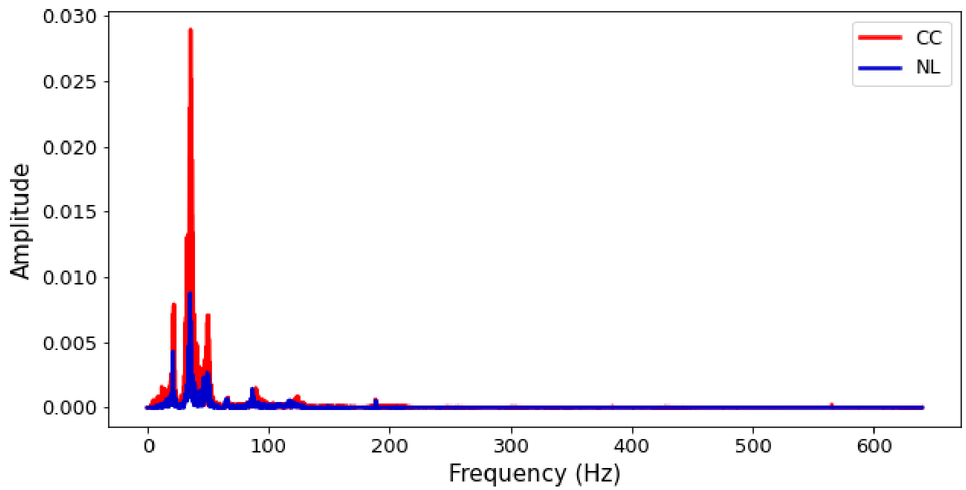

Figure 2.152. Frequency-domain plots of CC vs. NL signals in the branched network measured by sensor H2

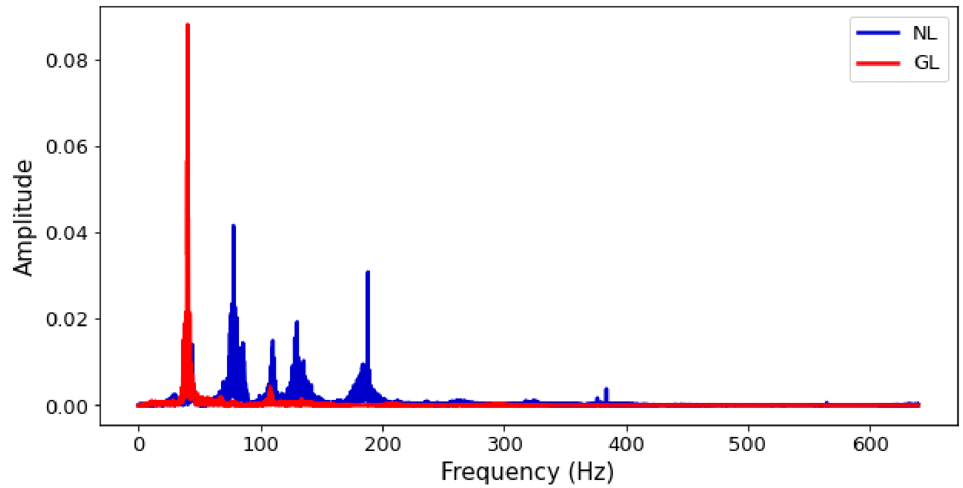

Figure 2.153. Frequency-domain plots of GL vs. NL signals in the looped network measured by sensor H2

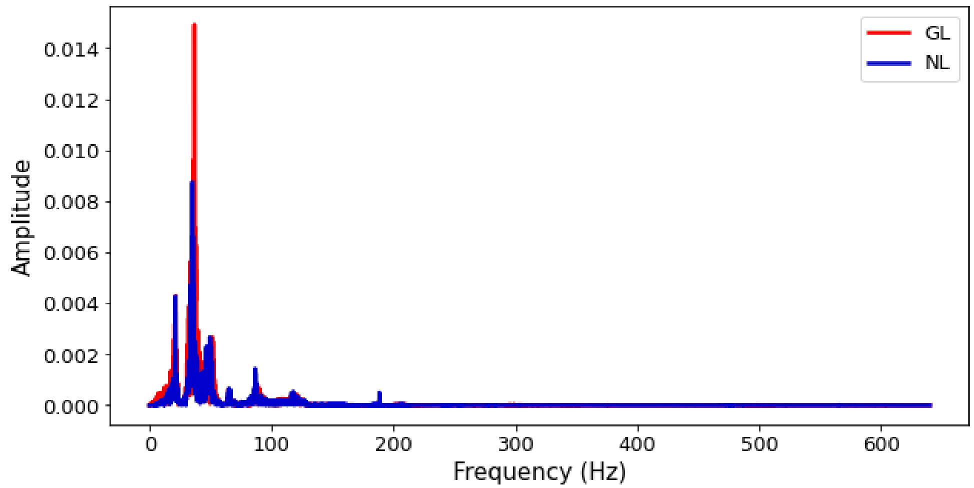

Figure 2.154. Frequency-domain plots of GL vs. NL signals in the branched network measured by sensor H2

Table 2.43 includes analytical information of the acoustic data's frequency-domain plots where leak and no leak signals of sensors H1 and H2 in the looped and branched networks are visually compared.

Table 2.43. Analysis of the frequency-domain acoustic data plots measured by H1 and H2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Frequency-domain plot (for ND and NS signal)	NL vs. OL – H1	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 400 Hz. - NL signal's dominant frequency has a slightly larger amplitude in the looped network than in the branched one. - In the looped network, OL signal's dominant frequency is larger than that of the NL signal. This difference might help detect the leak. - OL signal's dominant frequency in the branched network has a larger amplitude than that of the looped one. This difference might indicate network change effects. But the networks' dominant frequencies are approximately below 50 Hz. - Similar to A1 measurements, for OL signal in the looped network, frequencies with non-zero amplitudes are less than 500 Hz. - Similar to A1 measurements, for OL signal in the branched network, frequencies with non-zero amplitudes are less than 600 Hz. - OL signal's frequencies in the looped network are more pronounced than in the branched network. - In the branched network, OL signal's dominant frequency amplitude is larger than that of the NL signal. This difference might help detect the leak. - In the branched network, NL and leak signals have approximately the same spectral patterns. This observation can imply how similar leak and no leak signals are in the branched network.

Table 2.43. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Frequency-domain plot (for ND and NS signal)	NL vs. LC – H1	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 400 Hz. - In the looped network, LC signal’s dominant frequency is larger than that of the NL signal. This difference might help detect the leak. - LC signal’s dominant frequency in the looped network is larger than that of the branched network. This difference might indicate network change effects. But the amplitudes of the frequencies are nearly similar. - For LC signal in the looped network, frequencies with non-zero amplitudes are less than 400 Hz. - For LC signal in the branched network, frequencies with non-zero amplitudes are less than 300 Hz. - The amplitudes of the LC signal’s frequencies in the looped network are more pronounced than in the branched network. This difference might indicate network change effects. - In both networks, frequencies of the LC signal have larger amplitudes than the NL signal.
	NL vs. CC – H1	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 400 Hz. - For CC signal in the looped network, frequencies with non-zero amplitudes are less than 400 Hz. - For CC signal in the branched network, frequencies with non-zero amplitudes are less than 300 Hz. - CC signal’s dominant frequency and its amplitude in the looped network are larger than those of the branched network. These differences might indicate network change effects. - In both networks, the dominant frequency of the CC signal has a larger amplitude than the NL signal. This difference might help detect the leak. - In the looped network, LC signal’s dominant frequency is larger than that of the NL signal. This difference might help detect the leak.
	NL vs. GL – H1	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 400 Hz. - For GL signal in the looped network, frequencies with non-zero amplitudes are less than 500 Hz. - For GL signal in the branched network, frequencies with non-zero amplitudes are less than 400 Hz. - In both networks, the amplitudes of the GL signal’s dominant frequencies are larger than those of the NL signal.

Table 2.43. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Frequency-domain plot (for ND and NS signal)	All – H1	- For NL signal in both networks, frequencies with non-zero amplitudes are less than 400 Hz. - With no demand, no noise, and in the looped network, frequencies of OL and GL signals with non-zero amplitudes are less than 500 Hz, and frequencies of the CC and LC signals with non-zero amplitudes are less than 400 Hz. While, those in the branched network varied between 300 Hz and 600 Hz with a 600 Hz cap for the OL signal and a 300 Hz cap for the CC signal. - With no demand, no noise, and in the looped network, the order of leak signals' frequency caps with non-zero amplitudes are OL = GL > CC = LC. - With no demand, no noise, and in the looped network: dominant frequency of leaks > dominant frequency of NL. These differences can help detect the leaks in the looped network. - With no demand, no noise, and in the branched network: amplitude of leaks' dominant frequency > amplitude of NLs' dominant frequency. These differences can help detect the leaks in the looped network.
	NL vs. OL – H2	- For NL signal in the looped and branched networks, frequencies with non-zero amplitudes are less than 400 Hz and 200 Hz, respectively. - NL signal's dominant frequency and its amplitude are larger in the looped network than in the branched one. - The amplitude of OL signal's dominant frequency in the branched network is larger than that in the looped one. This difference can help capture the network change. - For OL signal in the looped network, frequencies with non-zero amplitudes are less than 200 Hz. - For OL signal in the branched network, frequencies with non-zero amplitudes are less than 200 Hz. - In the branched network, though the dominant frequencies of OL and NL are nearly the same, the amplitude of OL signal's dominant frequency is larger than that of NL signal. This difference can help detect the leak in the branched network.

Table 2.43. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Frequency-domain plot (for ND and NS signal)	NL vs. LC – H2	- For NL signal in the looped and branched networks, frequencies with non-zero amplitudes are less than 400 Hz, and 200 Hz, respectively. - NL signal's dominant frequency and its amplitude are larger in the looped network than in the branched one. - The amplitude of LC signal's dominant frequency in the looped network is larger than that in the branched one. This difference can help capture the network change. The dominant frequencies of both networks are nearly similar. - For LC signal in the looped network, frequencies with non-zero amplitudes are less than 200 Hz. - For LC signal in the branched network, frequencies with non-zero amplitudes are less than 200 Hz. - In both networks, the amplitude of LC signal's dominant frequency is larger than that of the NL signal. This difference can help detect the leak.
	NL vs. CC – H2	- For NL signal in the looped and branched networks, frequencies with non-zero amplitudes are less than 400 Hz, and 200 Hz, respectively. - NL signal's dominant frequency and its amplitude are larger in the looped network than in the branched one. - The amplitude of CC signal's dominant frequency in the looped network is larger than that in the branched one. This difference can help capture the network change. The dominant frequencies of both networks are nearly similar. - For CC signal in the looped network, frequencies with non-zero amplitudes are less than 200 Hz. - For CC signal in the branched network, frequencies with non-zero amplitudes are less than 200 Hz. - In both networks, the amplitude of CC signal's dominant frequency is larger than that of the NL signal. This difference can help detect the leak.

Table 2.43. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Frequency-domain plot (for ND and NS signal)	NL vs. GL – H2	- For NL signal in the looped and branched networks, frequencies with non-zero amplitudes are less than 400 Hz, and 200 Hz, respectively. - NL signal's dominant frequency and its amplitude are larger in the looped network than in the branched one. - The amplitude of GL signal's dominant frequency in the branched network is larger than that in the looped one. This difference can help capture the network change. The dominant frequencies of both networks are nearly similar. - For GL signal in the looped network, frequencies with non-zero amplitudes are less than 200 Hz. - For GL signal in the branched network, frequencies with non-zero amplitudes are less than 200 Hz. - In both networks, the amplitude of GL signal's dominant frequency is larger than that of the NL signal. This difference can help detect the leak.
	All – H2	- For NL signal in the looped and branched networks, frequencies with non-zero amplitudes are less than 400 Hz, and 200 Hz, respectively. - NL signal's dominant frequency and its amplitude are larger in the looped network than in the branched one. - With no demand, no background noise, and in the looped network, all leak signals' frequency caps with non-zero amplitudes are less than 200 Hz. - With no demand, no background noise, and in the branched network, all leak signals' frequency caps with non-zero amplitudes are less than 200 Hz. - In both networks, the amplitudes of all leak signals' dominant frequencies are larger than those of the NL signals. These differences can help detect leaks.

2.12.4.3. Cumulative Distribution Plot (for ND and NS signal)

Figures 2.155 and 2.156 show the cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped and branched networks, respectively, measured by sensor H1 with no demand and no background noise. Also, Figures 2.157 and 2.158 show the cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped and branched networks, respectively, measured by sensor H2 with no demand and background noise.

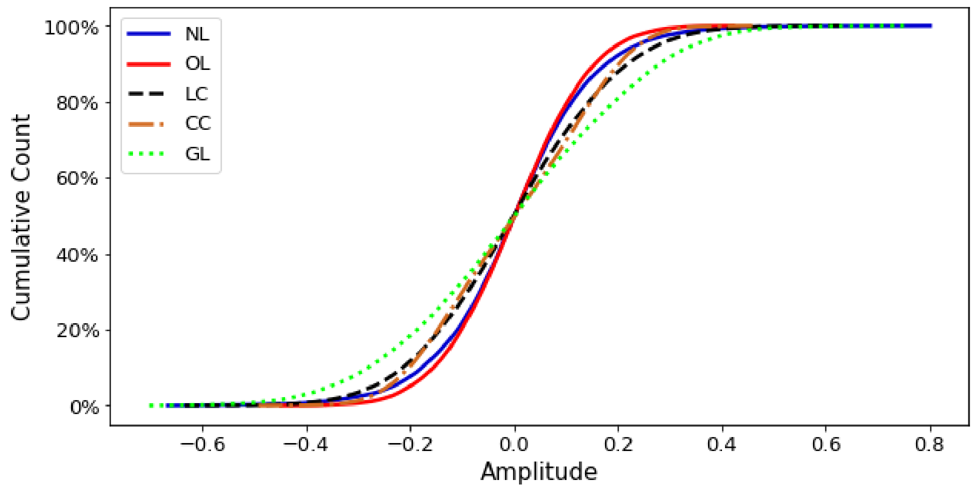

Figure 2.155. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor H1

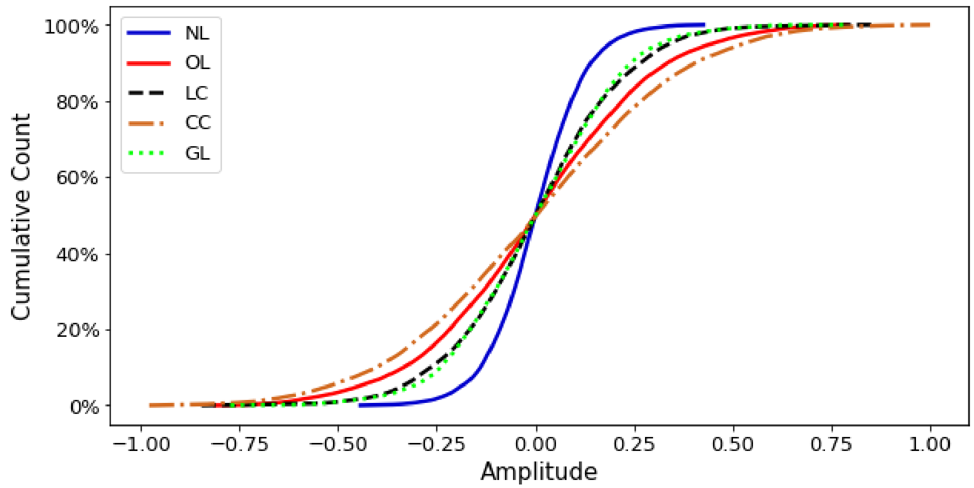

Figure 2.156. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the branched network, measured by sensor H1

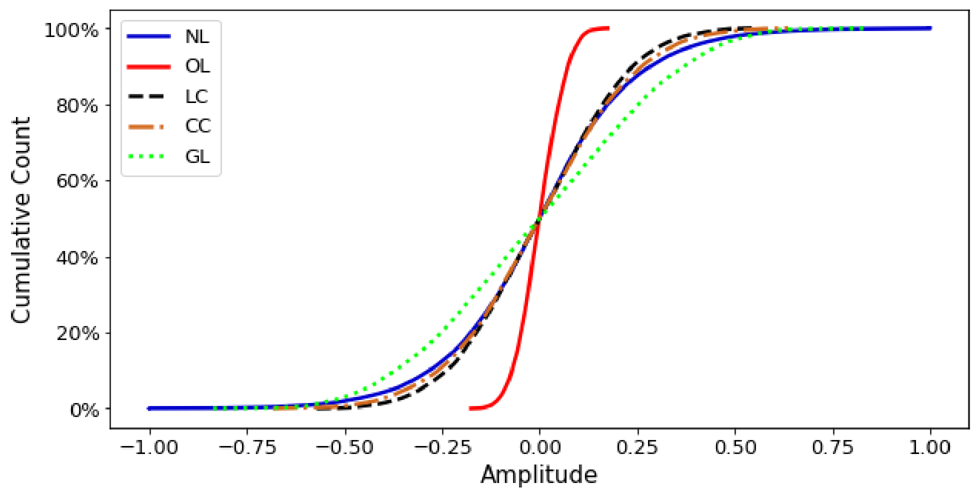

Figure 2.157. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor H2

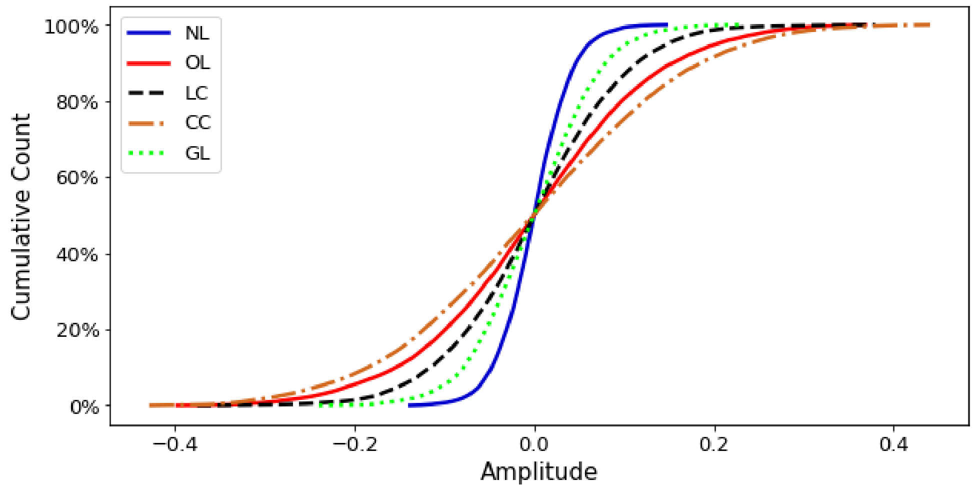

Figure 2.158. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor H2

Table 2.44 includes the analysis of the cumulative distribution plots where leak and no leak signals of sensors H1 and H2 in the looped and branched networks are visually compared.

Table 2.44. Analysis of the hydrophone data’s cumulative distribution plots measured by H1 and H2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Cumulative distribution plot (for ND and NS signal)	NL – H1	- NL hydrophone data magnitudes span a larger range in the looped network than in the branched network. - NL hydrophone data magnitudes are smaller than all leaks’ hydrophone data magnitudes in the branched network. - NL hydrophone data magnitudes in the branched network are smallest compared to all leak and NL signals of both networks. - Order of signal magnitude in the looped network: GL > CC > LC > NL > OL. - Order of signal magnitude in the branched network: CC > OL > GL > LC > NL.
	OL – H1	- OL hydrophone data magnitudes are the smallest and span the shortest range compared to other leak and NL signals of the looped network. - OL hydrophone data magnitudes of the looped network are smaller than those of the branched network. - Order of signal magnitude in the looped network: GL > CC > LC > NL > OL. - Order of signal magnitude in the branched network: CC > OL > GL > LC > NL.
	LC – H1	- LC hydrophone data magnitudes are larger in the branched network than in the looped network. - LC, and CC cumulative distribution plots approximately overlap in the looped network. - LC, and GL cumulative distribution plots approximately overlap in the branched network. - Order of signal magnitude in the looped network: GL > CC > LC > NL > OL. - Order of signal magnitude in the branched network: CC > OL > GL > LC > NL.
	CC – H1	- CC hydrophone data magnitudes are larger in the branched network than in the looped network. - LC, and CC cumulative distribution plots approximately overlap in the looped network. - CC hydrophone data magnitudes are the largest and span the widest range compared to other leak and NL signals of the branched network. - Order of signal magnitude in the looped network: GL > CC > LC > NL > OL. - Order of signal magnitude in the branched network: CC > OL > GL > LC > NL.

Table 2.44. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Cumulative distribution plot (for ND and NS signal)	GL – H1	- GL hydrophone data magnitudes are larger in the branched network than in the looped network. - GL hydrophone data magnitudes are the largest compared to other leak and NL signals of the looped network. - GL hydrophone data magnitudes are the smallest compared to other leak signals of the branched network. - Order of signal magnitude in the looped network: GL > CC > LC > NL > OL. - Order of signal magnitude in the branched network: CC > OL > GL > LC > NL.
	All – H1	- NL hydrophone data magnitudes span a larger range in the looped network than in the branched network. - NL hydrophone data magnitudes are smaller than all leaks' hydrophone data magnitudes in the branched network. - Hydrophone data magnitudes of all leaks are larger in the branched network than in the looped network. - OL hydrophone data have the smallest magnitudes, and GL hydrophone data include the largest magnitudes among all signals in the looped network. - Leak hydrophone data magnitudes are more distinct in the branched network than the looped network. - Order of signal magnitude in the looped network: GL > CC > LC > NL > OL. This order conforms to the amplitude order of time-series plots for signals with no demand and no background noise in the looped network. - Order of signal magnitude in the branched network: CC > OL > GL > LC > NL. This order conforms to the amplitude order of time-series plots for signals with no demand and no background noise in the branched network.

Table 2.44. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Cumulative distribution plot (for ND and NS signal)	NL – H2	- Similar to H1, NL hydrophone data magnitudes span a larger range in the looped network than in the branched network. - NL hydrophone data magnitudes span the largest range in the looped network. - Similar to H1, NL hydrophone data magnitudes are smaller than all leaks' hydrophone data magnitudes in the branched network. - Order of signal magnitude in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude in the branched network: CC > OL > LC > GL > NL. This order is similar to that of H1.
	OL – H2	- Similar to H1, OL hydrophone data magnitudes are the smallest and span the shortest range compared to other leak and NL signals of the looped network. - Similar to H1, OL hydrophone data magnitudes of the looped network are smaller than those of the branched network. - Order of signal magnitude in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude in the branched network: CC > OL > LC > GL > NL. This order is similar to that of H1.
	LC – H2	- Contrary to H1, LC hydrophone data magnitudes are smaller in the branched network than in the looped network. - LC, and CC cumulative distribution plots approximately overlap in the looped network. - Order of signal magnitude in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude in the branched network: CC > OL > LC > GL > NL. This order is similar to that of H1.

Table 2.44. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Cumulative distribution plot (for ND and NS signal)	CC – H2	- Contrary to H1, CC hydrophone data magnitudes are smaller in the branched network than in the looped network. - CC hydrophone data magnitudes are the largest and span the widest range compared to other leak and NL signals of the branched network. - Order of signal magnitude in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude in the branched network: CC > OL > LC > GL > NL. This order is similar to that of H1.
	GL – H2	- Contrary to H1, GL hydrophone data magnitudes are smaller in the branched network than in the looped network. - Similar to H1, GL hydrophone data magnitudes are the largest compared to other leak and NL signals of the looped network. - Similar to H1, GL hydrophone data magnitudes are the smallest compared to other leak signals of the branched network. - Order of signal magnitude in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude in the branched network: CC > OL > LC > GL > NL. This order is similar to that of H1.

Table 2.44. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Cumulative distribution plot (for ND and NS signal)	All – H2	- Similar to H1, NL hydrophone data magnitudes span a larger range in the looped network than in the branched network. - Similar to H1, NL hydrophone data magnitudes are smaller than all leaks' hydrophone data magnitudes in the branched network. - NL hydrophone data magnitudes span the largest range in the looped network - Similar to H1, leak hydrophone data magnitudes are more distinct in the branched network than the looped network. - Similar to H1, OL hydrophone data magnitudes of the looped network are smaller than those of the branched network. - Similar to H1, GL hydrophone data magnitudes are the largest compared to other leak and NL signals of the looped network. - Contrary to H1, there is no consistent relation between leak signals of the looped network and those of the branched network. - Order of signal magnitude in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude in the branched network: CC > OL > LC > GL > NL. This order is similar to that of H1.

2.12.4.4. Box Plot (for ND and NS signal)

Figures 2.159 and 2.160 show box plots of the NL, OL, LC, CC, and GL signals in the looped and branched networks, respectively, measured by sensor H1 with no demand and no background noise. Figures 2.161 and 2.162 show the same plots for signals measured by sensor H2 with no demand and no background noise.

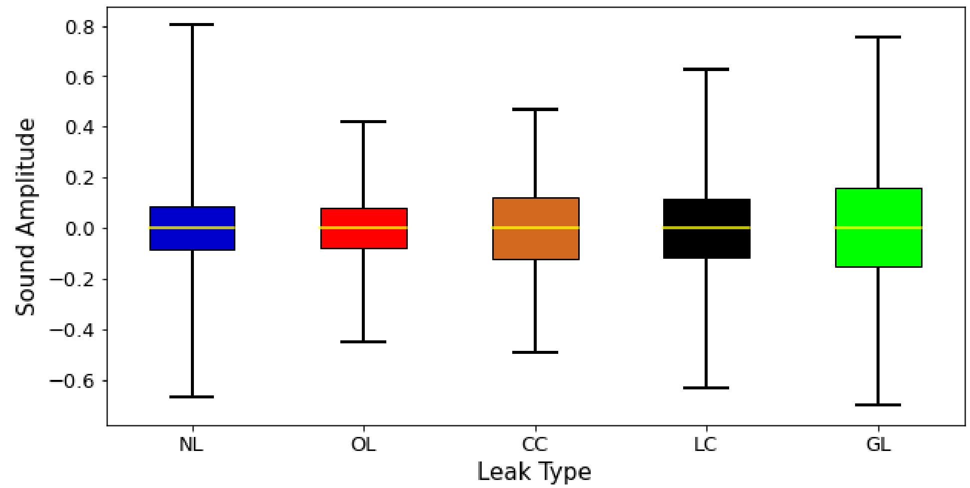

Figure 2.159. Box plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor H1

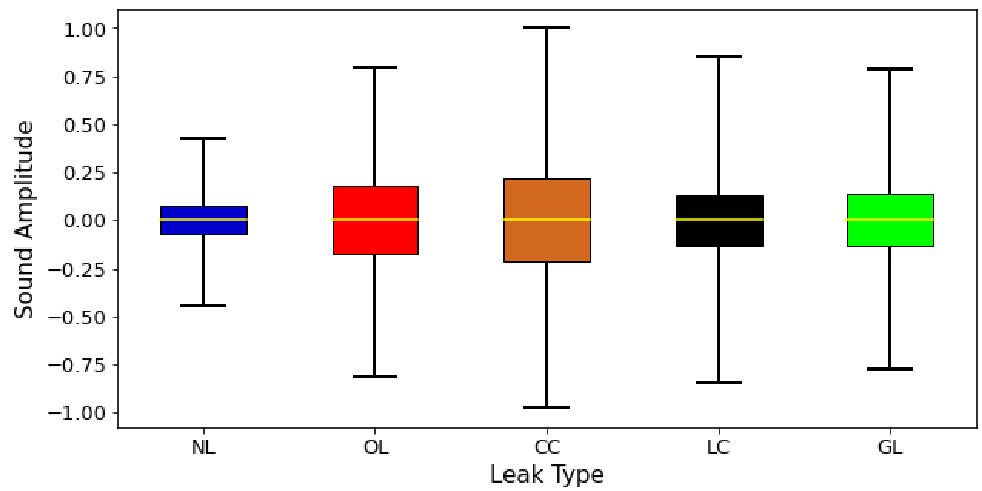

Figure 2.160. Box plots of the NL, OL, LC, CC, and GL signals in the branched network, measured by sensor H1

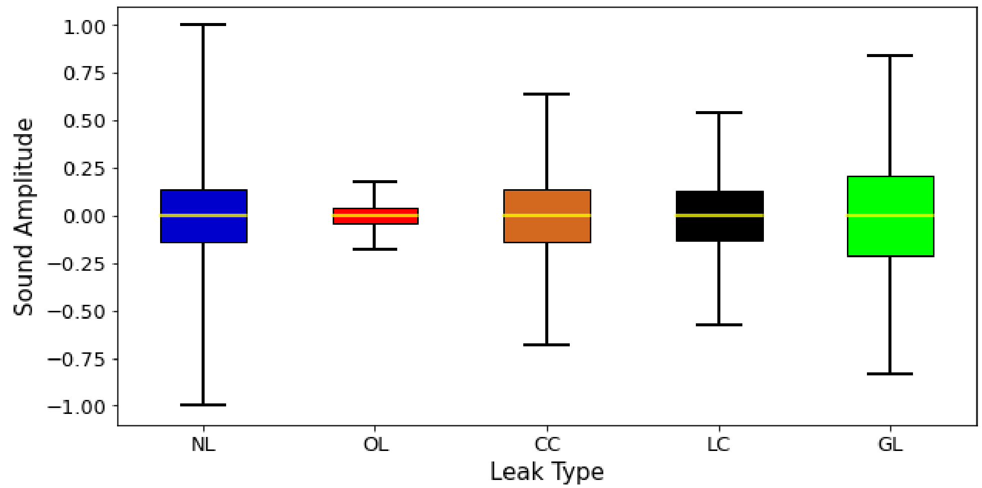

Figure 2.161. Box plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor H2

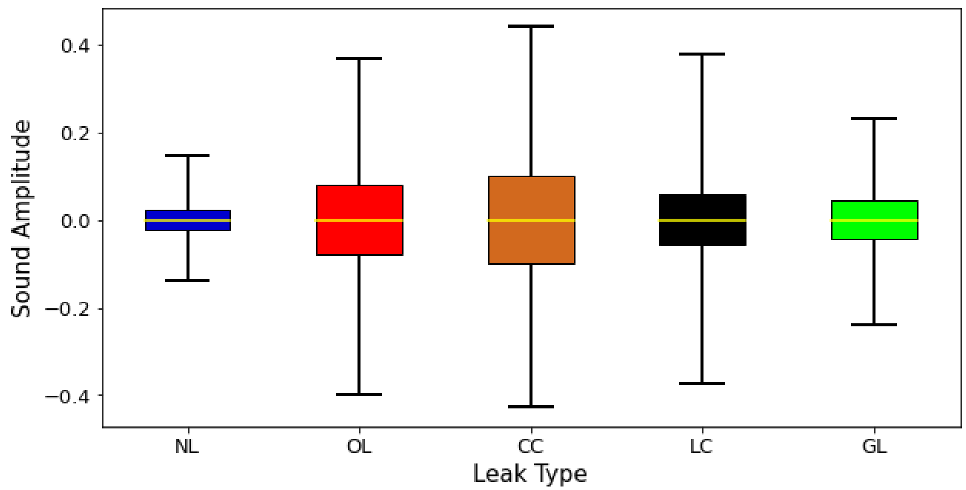

Figure 2.162. Box plots of the NL, OL, LC, CC, and GL signals in the branched network, measured by sensor H2

Table 2.45 includes the analysis of the box plots where leak and no leak signals of sensors H1 and H2 in the looped and branched networks are visually compared.

Table 2.45. Analysis of acoustic data's box plots measured by H1 and H2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Box plot (for ND and NS signal)	NL – H1	- Comparison of NL signal magnitude continuum: $NL_{lo} > NL_{br}$ - Order of signal magnitude continuum in the looped network: GL > CC > LC > NL > OL. - Order of signal magnitude continuum in the branched network: CC > OL > GL > LC > NL.
	OL – H1	- Comparison of OL signal magnitude continuum: $OL_{lo} < OL_{br}$ - Order of signal magnitude continuum in the looped network: GL > CC > LC > NL > OL. - Order of signal magnitude continuum in the branched network: CC > OL > GL > LC > NL.
	LC – H1	- Comparison of LC signal magnitude continuum: $LC_{lo} < LC_{br}$ - Order of signal magnitude continuum in the looped network: GL > CC > LC > NL > OL. - Order of signal magnitude continuum in the branched network: CC > OL > GL > LC > NL.
	CC – H1	- Comparison of CC signal magnitude continuum: $CC_{lo} < CC_{br}$ - Order of signal magnitude continuum in the looped network: GL > CC > LC > NL > OL. - Order of signal magnitude continuum in the branched network: CC > OL > GL > LC > NL.
	GL – H1	- GL signal of the looped network includes the largest range of acoustic data magnitudes compared to all other signals of the looped and the branched networks. - Comparison of GL signal magnitude continuum: $GL_{lo} > GL_{br}$ - Order of signal magnitude continuum in the looped network: GL > CC > LC > NL > OL. - Order of signal magnitude continuum in the branched network: CC > OL > GL > LC > NL.

Table 2.45. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Box plot (for ND and NS signal)	All – H1	- GL signal of the looped network includes the largest range of acoustic data magnitudes compared to all other signals of the looped and the branched networks. - There is no consistent pattern in the relations of the two networks' signal magnitude continuum. - Order of signal magnitude continuum in the looped network: GL > CC > LC > NL > OL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot. - Order of signal magnitude continuum in the branched network: CC > OL > GL > LC > NL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot. - Comparing the cracks, CC signals' magnitude continuum is larger than that of LC in both networks.
	NL – H2	- Comparison of NL signal magnitude continuum: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$ - Order of signal magnitude continuum in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude continuum in the branched network: CC > OL > LC > GL > NL.
	OL – H2	- Comparison of NL signal magnitude continuum: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$ - Order of signal magnitude continuum in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude continuum in the branched network: CC > OL > LC > GL > NL.
	LC – H2	- Comparison of LC signal magnitude continuum: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$ - Order of signal magnitude continuum in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude continuum in the branched network: CC > OL > LC > GL > NL.
	CC – H2	- Comparison of CC signal magnitude continuum: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$ - Order of signal magnitude continuum in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude continuum in the branched network: CC > OL > LC > GL > NL.
	GL – H2	- Similar to H1, GL signal of the looped network includes the largest range of acoustic data magnitudes compared to all other signals of the looped and the branched networks. - Comparison of GL signal magnitude continuum: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$ - Order of signal magnitude continuum in the looped network: GL > NL > CC > LC > OL. - Order of signal magnitude continuum in the branched network: CC > OL > LC > GL > NL.

Table 2.45. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Box plot (for ND and NS signal)	All – H2	- Similar to H1, GL signal of the looped network includes the largest range of acoustic data magnitudes compared to all other signals of the looped and the branched networks. - The relations of leak signals between two networks are similar to those relations based on sensor A2 data. - Signal magnitude continua in the looped network are larger than those of the branched network. - Order of signal magnitude continuum in the looped network: GL > NL > CC > LC > OL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot. - Order of signal magnitude continuum in the branched network: CC > OL > LC > GL > NL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot. - Comparing the cracks, CC signals' magnitude continuum is larger than that of LC in both networks.

2.12.4.5. Cross Spectral Plot (for ND and NS signal)

Figures 2.163 and 2.164 show the cross spectral plots of the NL, OL, LC, CC, and GL signals with no demands and no background noise in the looped and branched networks, respectively.

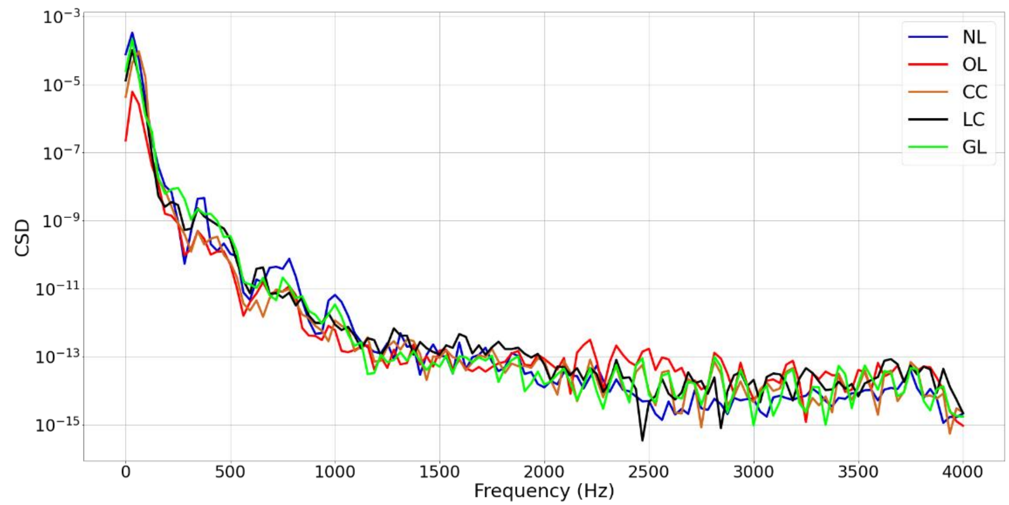

Figure 2.163. Cross spectral plots of the NL, OL, LC, CC, and GL signals in the looped network for hydrophones

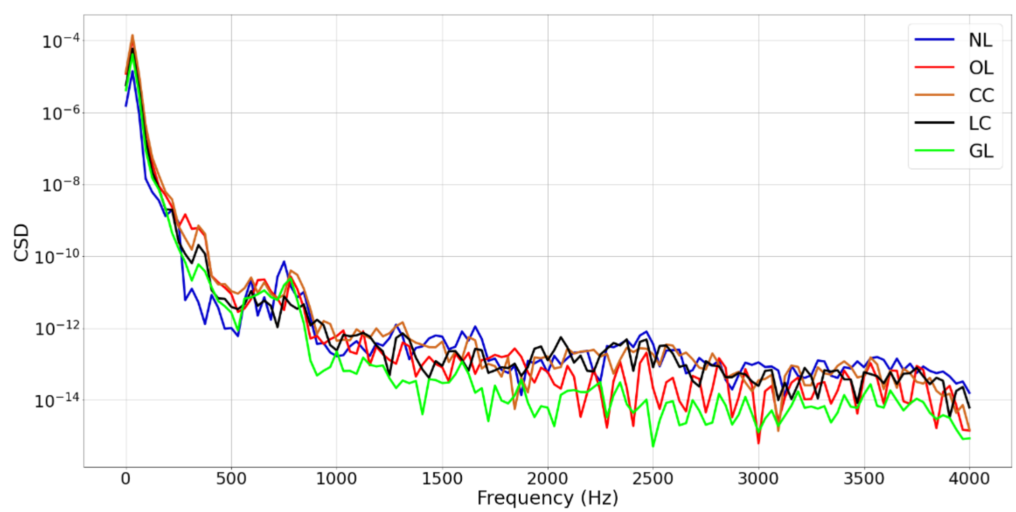

Figure 2.164. Cross spectral plots of the NL, OL, LC, CC, and GL signals in the branched network for hydrophones

Table 2.46 compares the area under the cross spectral plots of the looped and branched networks with no demand.

Table 2.46. Comparison of the area under the cross spectral plots of acoustic data

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Area under the cross-spectral plot (for ND and NS signal)	NL	- Comparison of NL signal's areas under the CSD plot: $NL_{lo} > NL_{br}$ - Order of areas under the CSD plot in the looped network: NL > GL > CC > LC > OL. - Order of areas under the CSD plot in the branched network: CC > OL > LC > GL > NL.
	OL	- Comparison of OL signal's areas under the CSD plot: $OL_{lo} < OL_{br}$ - Order of areas under the CSD plot in the looped network: NL > GL > CC > LC > OL. - Order of areas under the CSD plot in the branched network: CC > OL > LC > GL > NL.
	LC	- Comparison of LC signal's areas under the CSD plot: $LC_{lo} > LC_{br}$ - Order of areas under the CSD plot in the looped network: NL > GL > CC > LC > OL. - Order of areas under the CSD plot in the branched network: CC > OL > LC > GL > NL.
	CC	- Comparison of CC signal's areas under the CSD plot: $CC_{lo} < CC_{br}$ - Order of areas under the CSD plot in the looped network: NL > GL > CC > LC > OL. - Order of areas under the CSD plot in the branched network: CC > OL > LC > GL > NL.
	GL	- Comparison of GL signal's areas under the CSD plot: $GL_{lo} > GL_{br}$ - Order of areas under the CSD plot in the looped network: NL > GL > CC > LC > OL. - Order of areas under the CSD plot in the branched network: CC > OL > LC > GL > NL.
	All	- With ND and NS, there is no consistent pattern in the relations of the two networks' areas under the CSD plot. Therefore, the area under the CSD plot is not capable of capturing the effects of the network change. - Order of areas under the CSD plot in the looped network: NL > GL > CC > LC > OL. - Order of areas under the CSD plot in the branched network: CC > OL > LC > GL > NL.

2.12.4.6. LDI

2.12.4.6.1. Scatter Plot

Figure 2.165 shows the scatter plots of the LDI for acoustic data measured in the looped network with 0 (GPM), 3 (GPM), and 7.5 (GPM) demand variants where the horizontal axis is the leaks' measured flow.

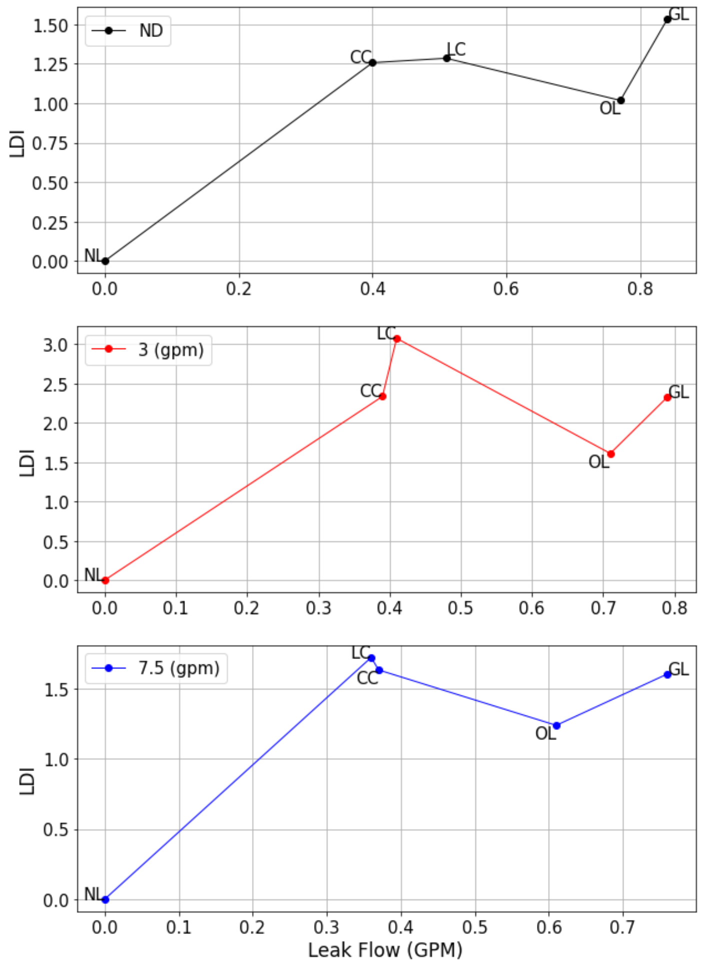

Figure 2.165. Scatter plots of the LDI for acoustic data measured in the looped network

Figure 2.166 shows the scatter plots of the LDI for acoustic data measured in the branched network with 0 (GPM), 3 (GPM), and 7.5 (GPM) demand variants and when background noise is present, where the horizontal axis is the leaks' measured flow.

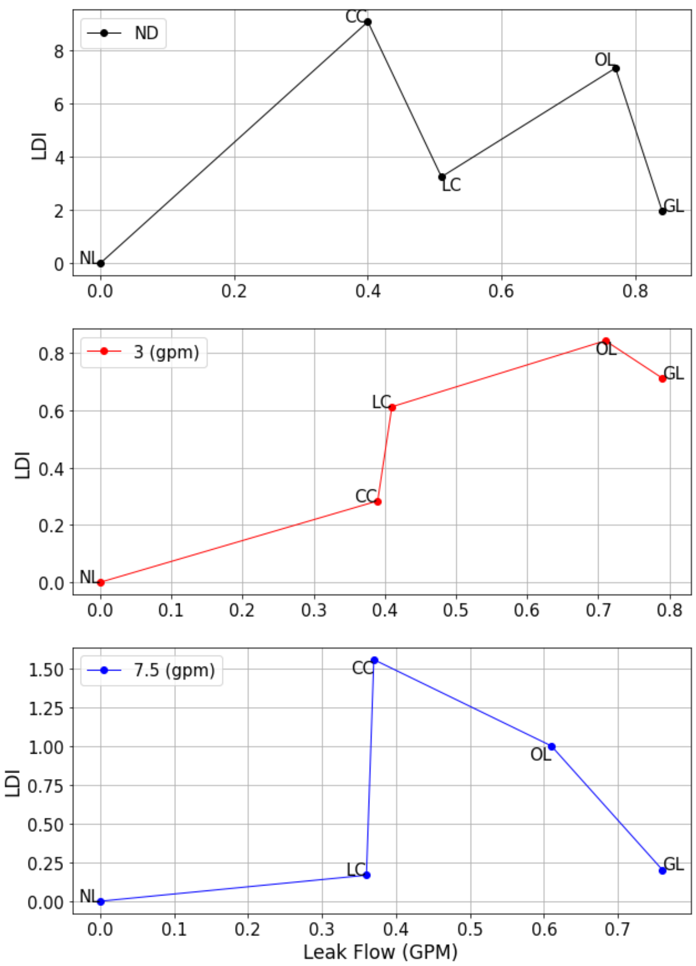

Figure 2.166. Scatter plots of the LDI for acoustic data measured in the branched network

Table 2.47 includes an analysis of the LDI scatter plots for acoustic data of the looped and branched networks.

Table 2.47. Analysis of the LDI scatter plots for acoustic data recorded in the looped and branched networks

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
LDI (scatter plot)	All leaks and demands	- Since the LDIs of leaks are larger than those of the benchmark, i.e., NL, the LDI can detect leaks. This conforms to the results of the paper of Yazdekhasti et al., 2016. - When demand exists in the looped network, the LDI has the following pattern: LC > CC > GL > OL. Despite this pattern, since the LDI magnitudes varied with the demand change, one cannot set a threshold to assign a specific LDI magnitude to a leak. Therefore, the LDI cannot discern leak types using acoustic data. - There is neither a specific pattern nor a threshold for the LDI in the branched network to distinguish leak types.

2.12.4.6.2. Bar Plot

Figures 2.167 and 2.168 show the bar plots of the LDI for acoustic data measured in the looped and branched networks, respectively, for all leak and demand variants and when background noise exists.

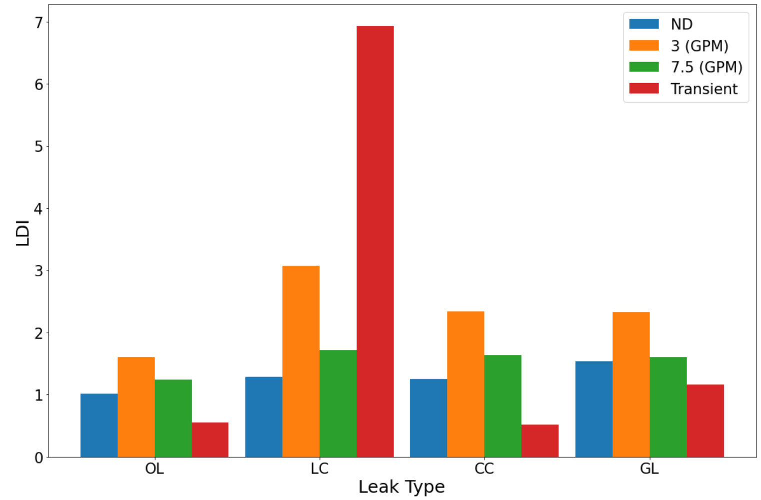

Figure 2.167. Bar plot of the LDI for acoustic data measured in the looped network with all leak and demand variants

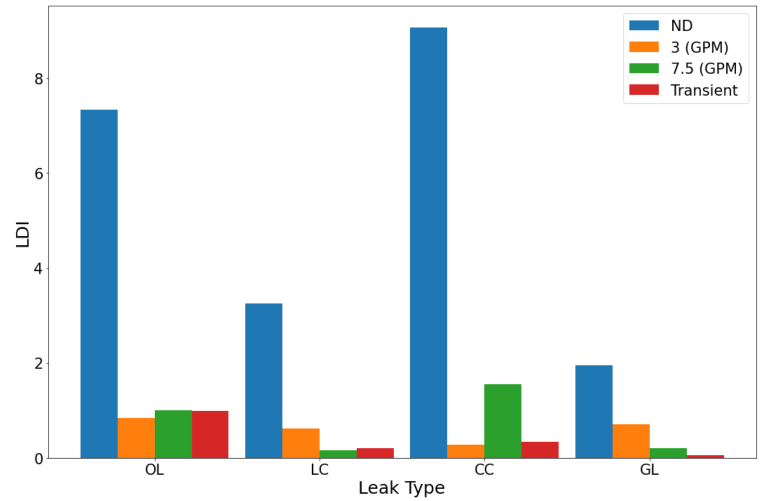

Figure 2.168. Bar plot of the LDI for acoustic data measured in the branched network with all leak and demand variants

Table 2.48 compares the LDI bar plots of acoustic data measured in the looped and branched networks with all leak and demand variants.

Table 2.48. Analysis of acoustic data LDI bar plots measured in the looped and branched networks with all leak and demand variants

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
LDI (bar plot)	ND	- Comparison of OL signal's LDI with no demand: $OL_{lo} < OL_{br}$ - Comparison of CC signal's LDI with no demand: $CC_{lo} < CC_{br}$ - Comparison of LC signal's LDI with no demand: $LC_{lo} < LC_{br}$ - Comparison of GL signal's LDI with no demand: $GL_{lo} < GL_{br}$ - Order of LDI for signals with ND in the looped network: GL > LC > CC > OL. - Order of LDI for signals with ND in the branched network: CC > OL > LC > GL
	3 (GPM)	- Comparison of OL signal's LDI with 3 (GPM) demand: $OL_{lo} > OL_{br}$ - Comparison of CC signal's LDI with 3 (GPM) demand: $CC_{lo} > CC_{br}$ - Comparison of LC signal's LDI with 3 (GPM) demand: $LC_{lo} > LC_{br}$ - Comparison of GL signal's LDI with 3 (GPM) demand: $GL_{lo} > GL_{br}$ - Order of LDI for signals with 3 (GPM) demand in the looped network: LC > CC > GL > OL. - Order of LDI for signals with 3 (GPM) demand in the branched network: OL > GL > LC > CC.
	7.5 (GPM)	- Comparison of OL signal's LDI with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$ - Comparison of CC signal's LDI with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$ - Comparison of LC signal's LDI with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$ - Comparison of GL signal's LDI with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$ - Order of LDI for signals with 7.5 (GPM) demand in the looped network: LC > CC > GL > OL. - Order of LDI for signals with 7.5 (GPM) demand in the branched network: CC > OL > GL > LC.

Table 2.48. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
LDI (bar plot)	Transient	- Comparison of OL signal's LDI with transient demand: $OL_{lo} < OL_{br}$ - Comparison of CC signal's LDI with transient demand: $CC_{lo} > CC_{br}$ - Comparison of LC signal's LDI with transient demand: $LC_{lo} > LC_{br}$ - Comparison of GL signal's LDI with transient demand: $GL_{lo} > GL_{br}$ - Order of LDI for signals with 7.5 (GPM) demand in the looped network: LC > OL > CC > GL. - Order of LDI for signals with 7.5 (GPM) demand in the branched network: OL > CC > LC > GL.
	All	- There are no specific patterns in the LDIs of the looped and branched networks. - The LDI of LC is the largest in the looped network with a non-zero demand. - Except for OL with the transient flow, the LDIs of other leaks in the looped network are larger than those of the branched one when there is a demand. - No LDI threshold could be set to discern leak types and network change effects.

2.12.4.7. Leak:NoLeak Amplitude Plot

Figures 2.169 and 2.170 show leak:noleak amplitude plots of the acoustic data measured by sensor H1 in the looped and branched networks, respectively, for all leak types, no demand, and no background noise. Figures 2.171 and 2.172 show the same plots but for sensor H2 data.

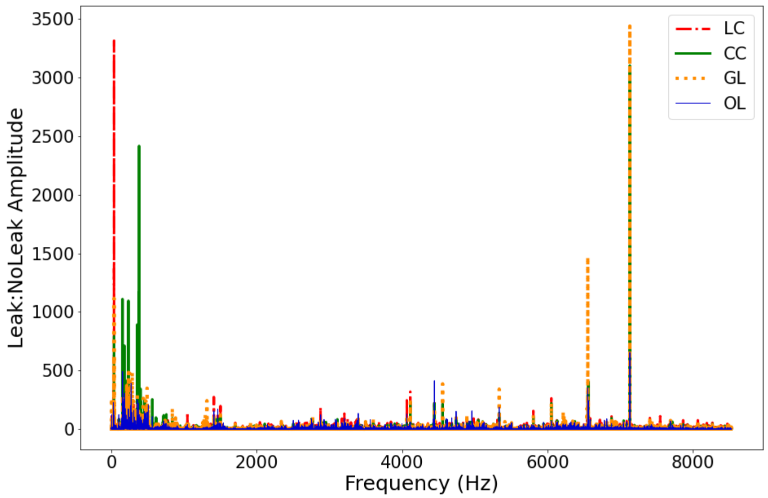

Figure 2.169. Leak:noleak amplitude plot of the acoustic data measured by sensor H1 in the looped network with no demand and no background noise

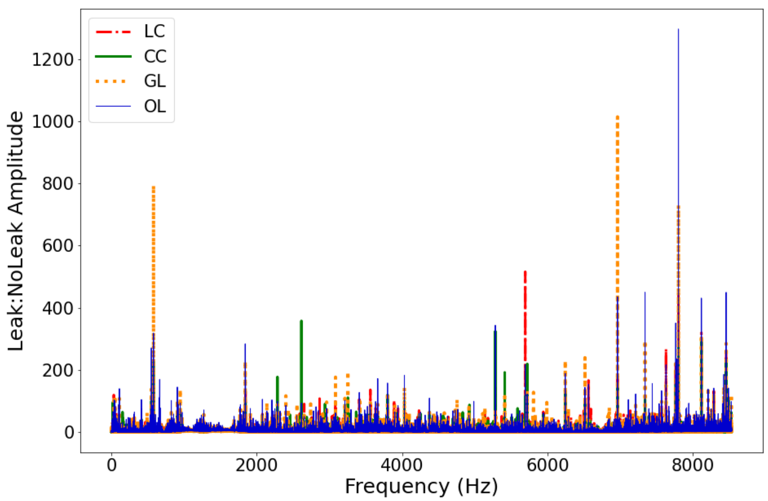

Figure 2.170. Leak:noleak amplitude plot of the acoustic data measured by sensor H1 in the branched network with no demand and no background noise

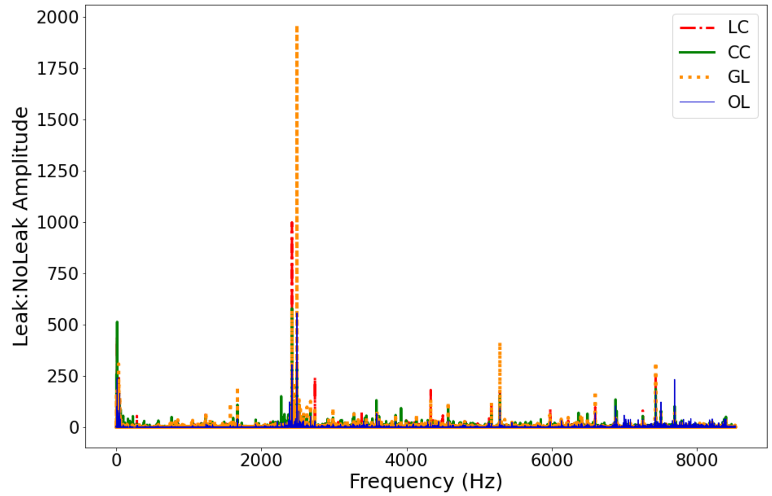

Figure 2.171. Leak:noleak amplitude plot of the acoustic data measured by sensor H2 in the looped network with no demand and no background noise

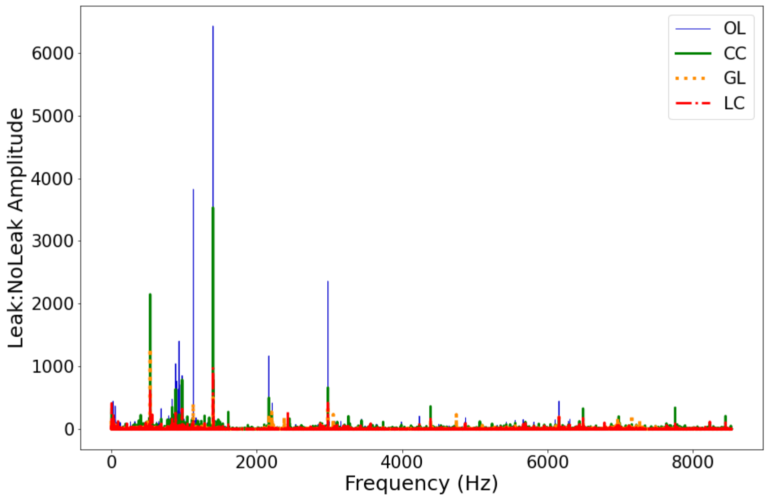

Figure 2.172. Leak:noleak amplitude plot of the acoustic data measured by sensor H2 in the branched network with no demand and no background noise

Table 2.49 compares leak:noleak amplitude plots of acoustic data measured in the looped and branched networks with all leak types, no demand, and no background noise.

Table 2.49. Analysis of leak:noleak amplitude plots of acoustic data measured in the looped and branched networks with all leak types, no demand, and no background noise

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Sum of magnitudes in the leak:noleak amplitude plot (for ND and NS signal)	OL – H1	- Comparison of the sum of leak:noleak magnitudes for OL: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Order of sum of leak:noleak magnitudes in the looped network: GL > LC > CC > OL. - Order of sum of leak:noleak magnitudes in the branched network: OL > GL > LC > CC.
	LC – H1	- Comparison of the sum of leak:noleak magnitudes for LC: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Order of sum of leak:noleak magnitudes in the looped network: GL > LC > CC > OL. - Order of sum of leak:noleak magnitudes in the branched network: OL > GL > LC > CC.
	CC – H1	- Comparison of the sum of leak:noleak magnitudes for CC: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$ - Order of sum of leak:noleak magnitudes in the looped network: GL > LC > CC > OL. - Order of sum of leak:noleak magnitudes in the branched network: OL > GL > LC > CC.
	GL – H1	- Comparison of the sum of leak:noleak magnitudes for GL: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$ - Order of sum of leak:noleak magnitudes in the looped network: GL > LC > CC > OL. - Order of sum of leak:noleak magnitudes in the branched network: OL > GL > LC > CC.
	All – H1	- Comparing the sum of magnitudes in the leak:noleak plots of two networks with ND and NS shows no consistent change pattern in the magnitudes when the network changes. Therefore, the sum of magnitudes in the leak:noleak plots is not capable of capturing the effects of network changes. - Order of sum of leak:noleak magnitudes in the looped network: GL > LC > CC > OL. - Order of sum of leak:noleak magnitudes in the branched network: OL > GL > LC > CC.

Table 2.49. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Sum of magnitudes in the leak:noleak amplitude plot (for ND and NS signal)	OL – H2	- Comparison of the sum of leak:noleak magnitudes for OL: $OL_{lo} < OL_{br}$. - Order of sum of leak:noleak magnitudes in the looped network: GL > CC > LC > OL. - Order of sum of leak:noleak magnitudes in the branched network: OL > CC > LC > GL.
	LC – H2	- Comparison of the sum of leak:noleak magnitudes for LC: $LC_{lo} < LC_{br}$. - Order of sum of leak:noleak magnitudes in the looped network: GL > CC > LC > OL. - Order of sum of leak:noleak magnitudes in the branched network: OL > CC > LC > GL.
	CC – H2	- Comparison of the sum of leak:noleak magnitudes for CC: $CC_{lo} < CC_{br}$. - Order of sum of leak:noleak magnitudes in the looped network: GL > CC > LC > OL. - Order of sum of leak:noleak magnitudes in the branched network: OL > CC > LC > GL.
	GL – H2	- Comparison of the sum of leak:noleak magnitudes for GL: $GL_{lo} < GL_{br}$. - Order of sum of leak:noleak magnitudes in the looped network: GL > CC > LC > OL. - Order of sum of leak:noleak magnitudes in the branched network: OL > CC > LC > GL.
	All – H2	- Similar to H1, GL and OL have the largest and smallest sum of leak:noleak magnitudes, respectively, in the looped network. - The sum of leak:noleak magnitudes for all leaks in the branched network are larger than those of the looped network. - Order of sum of leak:noleak magnitudes in the looped network: GL > CC > LC > OL. - Order of sum of leak:noleak magnitudes in the branched network: OL > CC > LC > GL.

2.12.4.8. Dominant Frequency

Figures 2.173 and 2.174 show dominant frequency bar plots of the acoustic data measured by sensor H1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.175 and 2.176 show the same plots but for sensor H2 data.

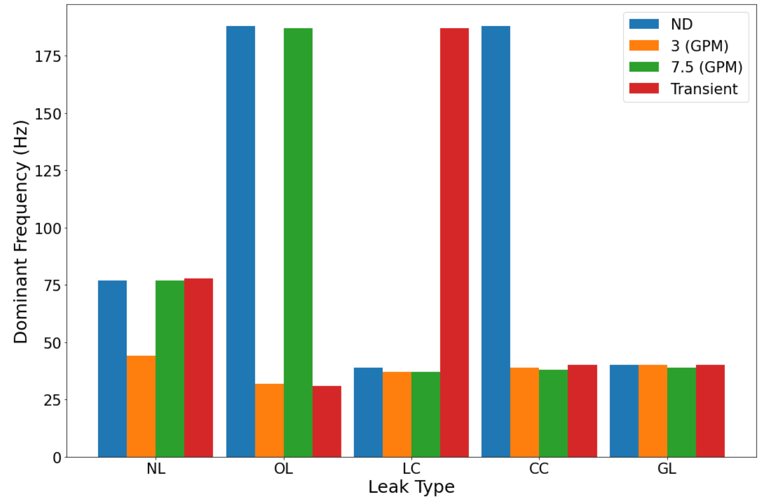

Figure 2.173. Dominant frequency bar plot of acoustic H1 data in the looped network for all leaks and demands

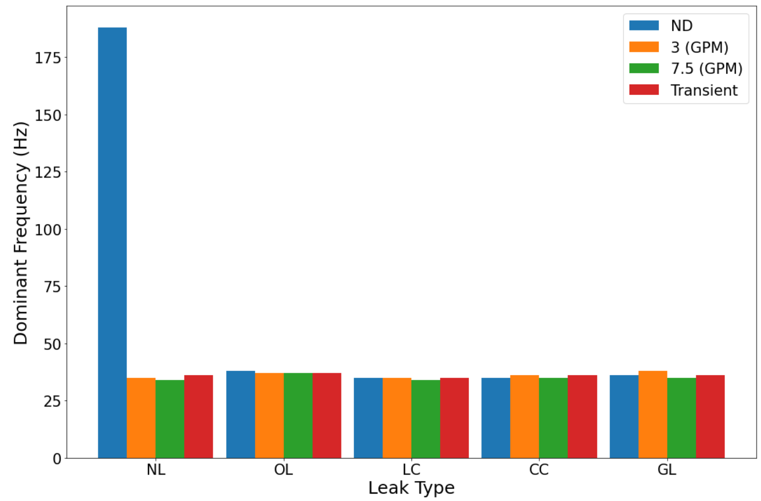

Figure 2.174. Dominant frequency bar plot of acoustic H1 data in the branched network for all leaks and demands

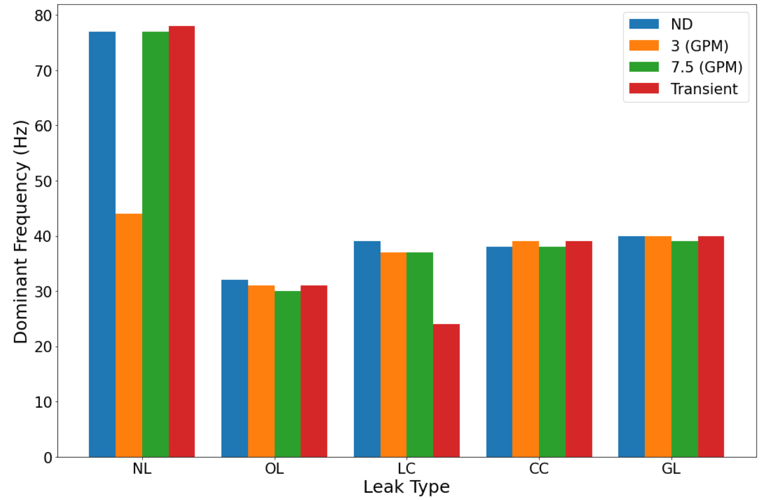

Figure 2.175. Dominant frequency bar plot of acoustic H2 data in the looped network for all leaks and demands

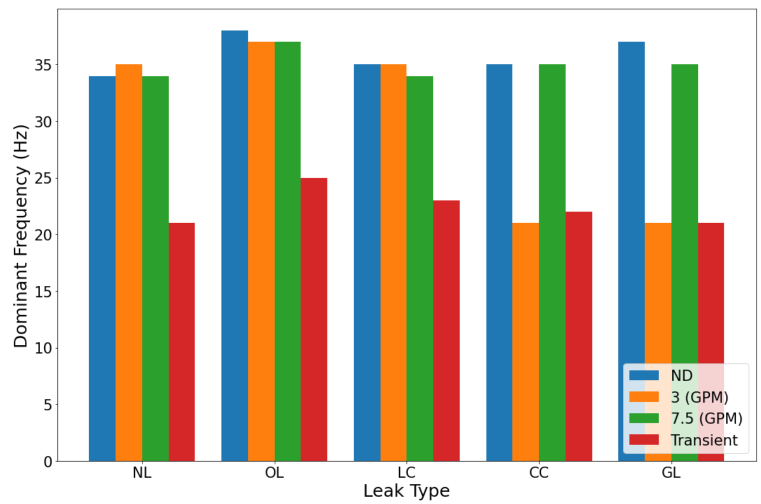

Figure 2.176. Dominant frequency bar plot of acoustic H2 data in the branched network for all leaks and demands

Table 2.50 compares dominant frequency plots of acoustic data measured in the looped and branched networks with all leak and demand types by sensors H1 and H2.

Table 2.50. Analysis of dominant frequency plots of acoustic data measured in the looped and branched networks with all leak and demand types

<table>
<tr><th>Evaluation Criterion</th><th>Evaluation Sub-criterion</th><th>Looped vs. Branched Network</th></tr>
<tr><td rowspan="2">Dominant frequency</td><td>ND – H1</td><td>- Comparison of NL signal's dominant frequency with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$.
- Comparison of OL signal's dominant frequency with no demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$.
- Comparison of CC signal's dominant frequency with no demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$.
- Comparison of LC signal's dominant frequency with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$.
- Comparison of GL signal's dominant frequency with no demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$.
- Order of dominant frequency with ND in the looped network: OL = CC > NL > GL > LC.
- Order of dominant frequency with ND in the branched network: NL > OL > GL > LC = CC.</td></tr>
<tr><td>3 (GPM) – H1</td><td>- Comparison of NL signal's dominant frequency with 3 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$.
- Comparison of OL signal's dominant frequency with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$.
- Comparison of CC signal's dominant frequency with 3 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$.
- Comparison of LC signal's dominant frequency with 3 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$.
- Comparison of GL signal's dominant frequency with 3 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$.
- Order of dominant frequency for signals with 3 (GPM) demand in the looped network: NL > GL > CC > LC > OL.
- Order of dominant frequency for signals with 3 (GPM) demand in the branched network: GL > OL > CC > LC = NL.</td></tr>
</table>

Table 2.50. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Dominant frequency	7.5 (GPM) – H1	- Comparison of NL signal's dominant frequency with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's dominant frequency with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's dominant frequency with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's dominant frequency with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's dominant frequency with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of dominant frequency for signals with 7.5 (GPM) demand in the looped network: OL > NL > GL > CC > LC. - Order of dominant frequency for signals with 7.5 (GPM) demand in the branched network: OL > CC = GL > NL = LC.
	Transient – H1	- Comparison of NL signal's dominant frequency with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's dominant frequency with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's dominant frequency with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's dominant frequency with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's dominant frequency with transient demand: $GL_{lo} > GL_{br}$. - Order of dominant frequency for signals with transient demand in the looped network: LC > NL > CC = GL > OL. - Order of dominant frequency for signals with transient demand in the branched network: OL > CC = GL = NL > LC.
	All – H1	- Comparing leaks' dominant frequency magnitudes of two networks indicates no consistent pattern in the magnitudes when networks change. Therefore, dominant frequency is not capable of capturing the effects of network changes. - Since there is no consistent order of dominant frequency for signals with different demands in both networks, dominant frequency is not capable of discerning leak types consistently in both networks.

Table 2.50. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Dominant frequency	ND – H2	- Comparison of NL signal’s dominant frequency with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal’s dominant frequency with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal’s dominant frequency with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal’s dominant frequency with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal’s dominant frequency with no demand: $GL_{lo} > GL_{br}$. - Order of dominant frequency with ND in the looped network: NL > GL > LC > CC > OL. - Order of dominant frequency with ND in the branched network: OL > GL > LC = CC > NL.
	3 (GPM) – H2	- Comparison of NL signal’s dominant frequency with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal’s dominant frequency with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal’s dominant frequency with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal’s dominant frequency with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal’s dominant frequency with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of dominant frequency for signals with 3 (GPM) demand in the looped network: NL > GL > CC > LC = OL. - Order of dominant frequency for signals with 3 (GPM) demand in the branched network: OL > NL = LC > CC = GL.
	7.5 (GPM) – H2	- Comparison of NL signal’s dominant frequency with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal’s dominant frequency with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal’s dominant frequency with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal’s dominant frequency with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal’s dominant frequency with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of dominant frequency for signals with 7.5 (GPM) demand in the looped network: NL > GL > CC > LC > OL. - Order of dominant frequency for signals with 7.5 (GPM) demand in the branched network: OL > CC = GL > LC = NL.

Table 2.50. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Dominant frequency	Transient – H2	- Comparison of NL signal's dominant frequency with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of dominant frequency for signals with transient demand in the looped network: NL > GL > CC > OL > LC. - Order of dominant frequency for signals with transient demand in the branched network: OL > LC > CC > GL = NL.
	All – H2	- GL in the looped network with a dominant frequency of 40 Hz has the largest dominant frequency compared to other leaks. - Comparing leaks' dominant frequency magnitudes of two networks indicates no consistent pattern in the magnitudes when the networks change. Therefore, dominant frequency cannot capture the effects of network changes. - Since there is no consistent order of dominant frequency for signals with different demands in both networks, dominant frequency is not capable of discerning leak types consistently in both networks.

2.12.4.9. Fundamental Frequency

Figures 2.177 and 2.178 show fundamental frequency bar plots of the acoustic data measured by sensor H1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.179 and 2.180 show the same plots but for sensor H2 data.

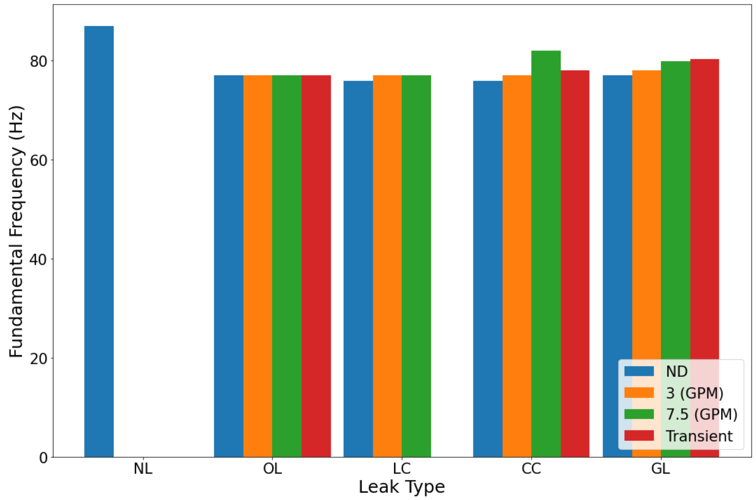

Figure 2.177. Fundamental frequency bar plot of hydrophone H1 data in the looped network for all leaks and demands

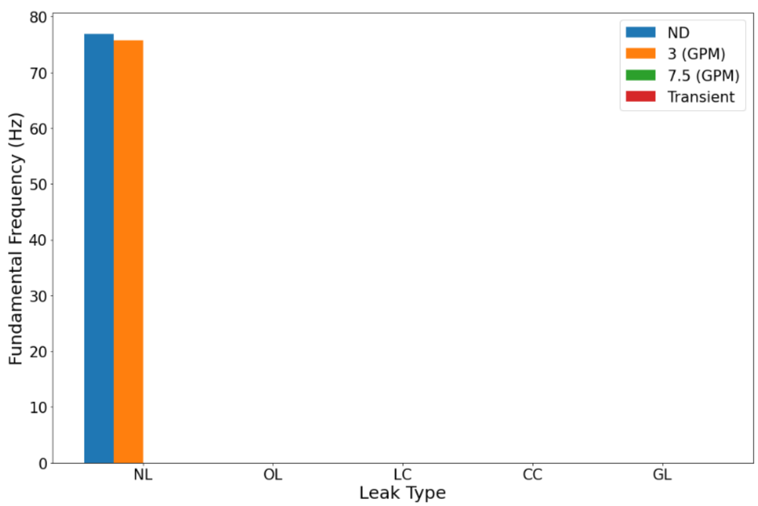

Figure 2.178. Fundamental frequency bar plot of hydrophone H1 data in the branched network for all leaks and demands

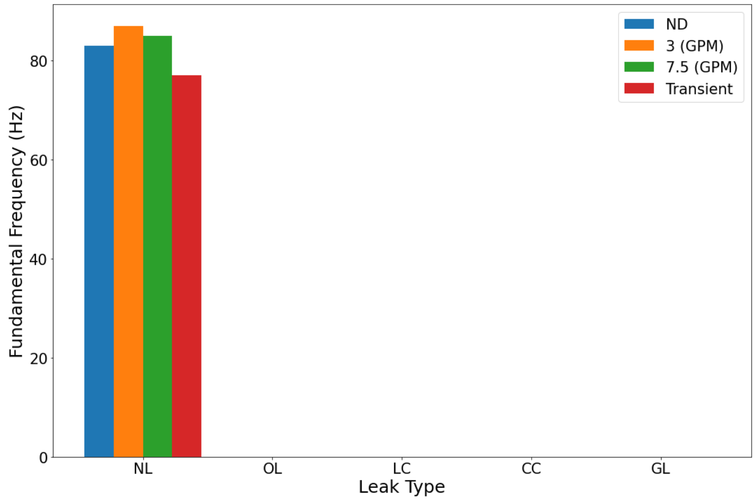

Figure 2.179. Fundamental frequency bar plot of hydrophone H2 data in the looped network for all leaks and demands

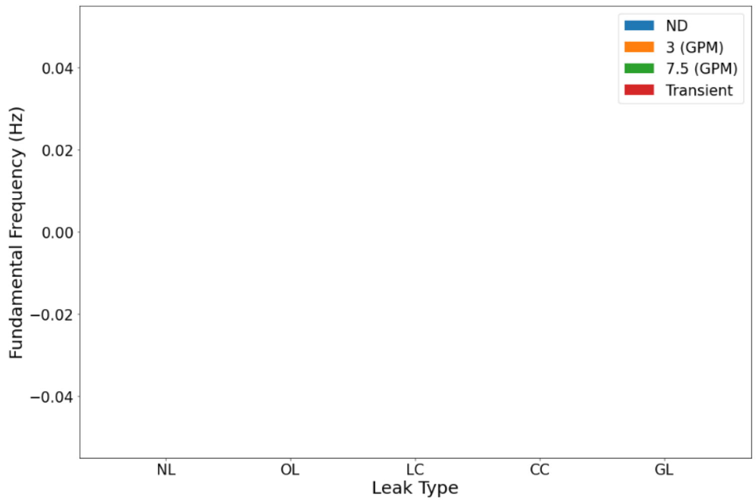

Figure 2.180. Fundamental frequency bar plot of hydrophone H2 data in the branched network for all leaks and demands

Table 2.51 compares fundamental frequency plots of acoustic data measured in the looped and branched networks with all leak and demand types by sensors H1 and H2.

Table 2.51. Analysis of fundamental frequency plots of acoustic data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Fundamental frequency	ND – H1	- Comparison of NL signal's fundamental frequency with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's fundamental frequency with no demand: $OL_{lo} > OL_{br} = 0$. - Comparison of CC signal's fundamental frequency with no demand: $CC_{lo} > CC_{br} = 0$. - Comparison of LC signal's fundamental frequency with no demand: $LC_{lo} > LC_{br} = 0$. - Comparison of GL signal's fundamental frequency with no demand: $GL_{lo} > GL_{br} = 0$. - Order of fundamental frequency with ND in the looped network: NL > OL = GL > LC = CC. - Order of fundamental frequency with ND in the branched network: NL > OL = LC = CC = GL = 0.
	3 (GPM) – H1	- Comparison of NL signal's fundamental frequency with 3 (GPM) demand: $NL_{lo} = 0 < NL_{br}$. - Comparison of OL signal's fundamental frequency with 3 (GPM) demand: $OL_{lo} > OL_{br} = 0$. - Comparison of CC signal's fundamental frequency with 3 (GPM) demand: $CC_{lo} > CC_{br} = 0$. - Comparison of LC signal's fundamental frequency with 3 (GPM) demand: $LC_{lo} > LC_{br} = 0$. - Comparison of GL signal's fundamental frequency with 3 (GPM) demand: $GL_{lo} > GL_{br} = 0$. - Order of fundamental frequency for signals with 3 (GPM) demand in the looped network: GL > OL = LC = CC > NL = 0. - Order of fundamental frequency for signals with 3 (GPM) demand in the branched network: NL > OL = LC = CC = GL = 0.

Table 2.51. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Fundamental frequency	7.5 (GPM) – H1	- Comparison of NL signal's fundamental frequency with 7.5 (GPM) demand: $NL_{lo} = NL_{br} = 0$. - Comparison of OL signal's fundamental frequency with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's fundamental frequency with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's fundamental frequency with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's fundamental frequency with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of fundamental frequency for signals with 7.5 (GPM) demand in the looped network: CC > GL > OL = LC > NL = 0. - Order of fundamental frequency for signals with 7.5 (GPM) demand in the branched network: NL = OL = LC = CC = GL = 0.
	Transient – H1	- Comparison of NL signal's fundamental frequency with transient demand: $NL_{lo} = NL_{br} = 0$. - Comparison of OL signal's fundamental frequency with transient demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's fundamental frequency with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's fundamental frequency with transient demand: $LC_{lo} = LC_{br} = 0$. - Comparison of GL signal's fundamental frequency with transient demand: $GL_{lo} > GL_{br}$. - Order of fundamental frequency for signals with transient demand in the looped network: GL > CC > OL > LC = NL = 0. - Order of fundamental frequency for signals with transient demand in the branched network: NL = OL = LC = CC = GL = 0.

Table 2.51. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Fundamental frequency	All – H1	- The fundamental frequencies of the looped network's leak signals are between 76 Hz and 80 Hz. However, the fundamental frequencies of the branched network's leak signals are all 0 Hz. - In the looped network, except for NL with ND that has the largest fundamental frequency, the fundamental frequency of all leaks with different demands is similar. Therefore, the fundamental frequency cannot discern leak types in the looped network. - Comparing leaks' fundamental frequency magnitudes of two networks shows a consistent change pattern in the magnitudes when networks change. The leaks' fundamental frequency magnitudes are all zero in the branched network, while those magnitudes are non-zero in the looped network. Therefore, the fundamental frequency can capture the effects of network changes. - Due to the similarity of fundamental frequencies for different leaks in each network, the fundamental frequency is not capable of discerning leak types in both networks.
	ND – H2	- Comparison of NL signal's fundamental frequency with no demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with no demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with no demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with no demand: $\mathrm{LC_{lo}} = \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with no demand: $\mathrm{GL_{lo}} = \mathrm{GL_{br}}$. - Order of fundamental frequency with ND in the looped network: NL > OL = LC = CC = GL = 0. - Order of fundamental frequency with ND in the branched network: NL = OL = LC = CC = GL = 0.
	3 (GPM) – H2	- Comparison of NL signal's fundamental frequency with 3 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with 3 (GPM) demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with 3 (GPM) demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with 3 (GPM) demand: $\mathrm{LC_{lo}} = \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with 3 (GPM) demand: $\mathrm{GL_{lo}} = \mathrm{GL_{br}}$. - Order of fundamental frequency with 3 (GPM) demand in the looped network: NL > OL = LC = CC = GL = 0. - Order of fundamental frequency with 3 (GPM) demand in the branched network: NL = OL = LC = CC = GL = 0.

Table 2.51. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Fundamental frequency	7.5 (GPM) – H2	- Comparison of NL signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{LC_{lo}} = \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{GL_{lo}} = \mathrm{GL_{br}}$. - Order of fundamental frequency with 7.5 (GPM) demand in the looped network: NL > OL = LC = CC = GL = 0. - Order of fundamental frequency with 7.5 (GPM) demand in the branched network: NL = OL = LC = CC = GL = 0.
	Transient – H2	- Comparison of NL signal's fundamental frequency with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with transient demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with transient demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with transient demand: $\mathrm{LC_{lo}} = \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with transient demand: $\mathrm{GL_{lo}}= \mathrm{GL_{br}}$. - Order of fundamental frequency with transient demand in the looped network: NL > OL = LC = CC = GL = 0. - Order of fundamental frequency with transient demand in the branched network: NL = OL = LC = CC = GL = 0.
	All – H2	- The fundamental frequencies of all leak and no leak signals in the branched network and all leak signals in the looped network are zero. Therefore, the fundamental frequencies of the sensor H2 data can neither discern leak types nor capture the network change effects.

2.12.4.10. Spectral Centroid

Figures 2.181 and 2.182 show spectral centroid bar plots of the acoustic data measured by sensor H1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.183 and 2.184 show the same plots but for sensor H2 data.

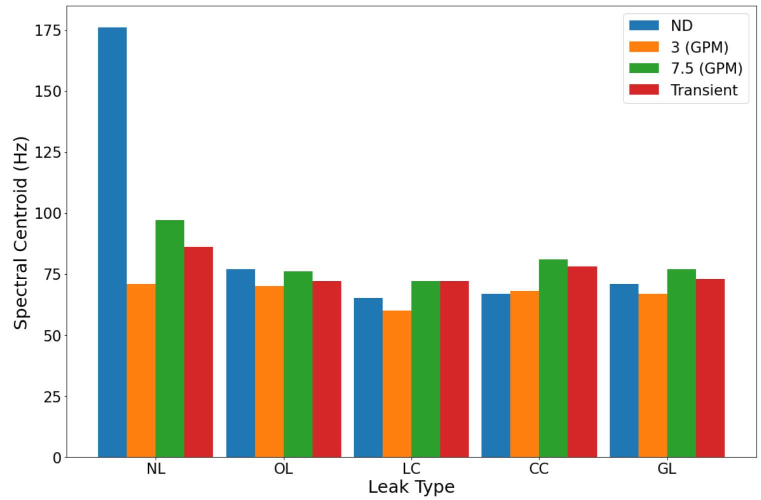

Figure 2.181. Spectral centroid bar plot of the hydrophone H1 data in the looped network for all leaks and demands

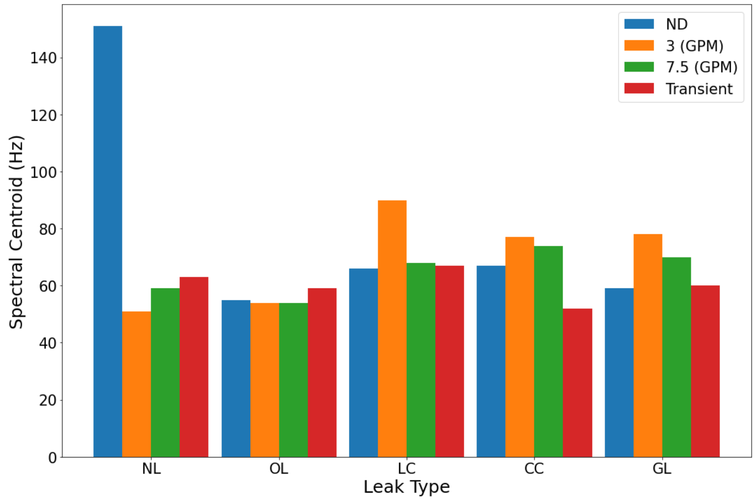

Figure 2.182. Spectral centroid bar plot of hydrophone H1 data in the branched network for all leaks and demands

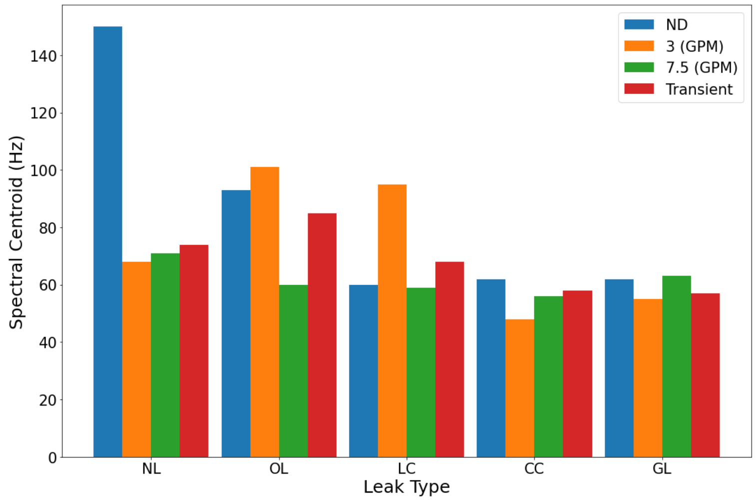

Figure 2.183. Spectral centroid bar plot of hydrophone H2 data in the looped network for all leaks and demands

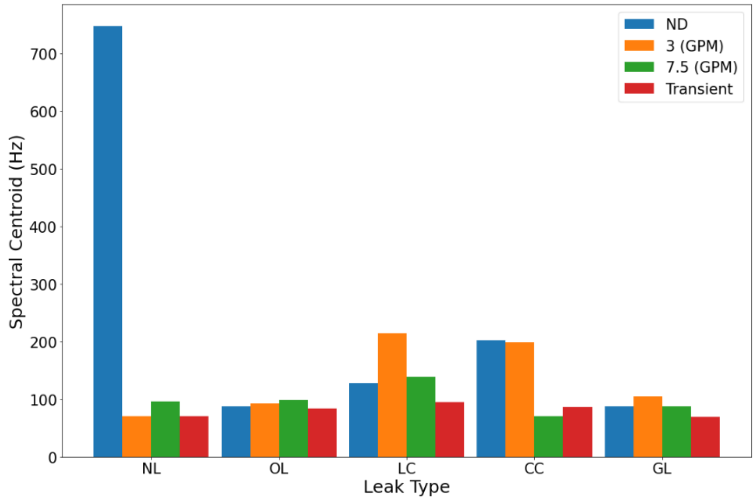

Figure 2.184. Spectral centroid bar plot of hydrophone H2 data in the branched network for all leaks and demands

Table 2.52 compares spectral centroid plots of acoustic data measured in the looped and branched networks with all leak and demand types by sensors H1 and H2.

Table 2.52. Analysis of spectral centroid plots of acoustic data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Spectral centroid	ND – H1	- Comparison of NL signal's spectral centroid with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's spectral centroid with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's spectral centroid with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's spectral centroid with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's spectral centroid with no demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid with ND in the looped network: NL > OL > CC = GL > LC. - Order of spectral centroid with ND in the branched network: NL > CC > LC > OL = GL.
	3 (GPM) – H1	- Comparison of NL signal's spectral centroid with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's spectral centroid with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's spectral centroid with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's spectral centroid with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's spectral centroid with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid for signals with 3 (GPM) demand in the looped network: OL > LC > NL > GL > CC. - Order of spectral centroid for signals with 3 (GPM) demand in the branched network: LC > CC > GL > OL > NL.

Table 2.52. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Spectral centroid	7.5 (GPM) – H1	- Comparison of NL signal's spectral centroid with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's spectral centroid with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's spectral centroid with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's spectral centroid with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's spectral centroid with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of spectral centroid for signals with 7.5 (GPM) demand in the looped network: NL > CC > GL > OL > LC. - Order of spectral centroid for signals with 7.5 (GPM) demand in the branched network: CC > GL > LC > NL > OL.
	Transient – H1	- Comparison of NL signal's spectral centroid with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's spectral centroid with transient demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's spectral centroid with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's spectral centroid with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's spectral centroid with transient demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid for signals with transient demand in the looped network: NL > CC > GL > OL = LC. - Order of spectral centroid for signals with transient demand in the branched network: LC > NL > GL > OL > CC.
	All – H1	- Comparing leaks' spectral centroid magnitudes of two networks shows no consistent change pattern in the magnitudes when the networks change. Therefore, the spectral centroid is not capable of capturing the effects of network changes. - Since there is no consistent order of spectral centroid for signals with different demands in both networks, the spectral centroid cannot discern leak types consistently in both networks.

Table 2.52. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Spectral centroid	ND – H2	- Comparison of NL signal's spectral centroid with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's spectral centroid with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's spectral centroid with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's spectral centroid with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's spectral centroid with no demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid with ND in the looped network: NL > OL > CC = GL > LC. - Order of spectral centroid with ND in the branched network: NL > CC > LC > OL = GL.
	3 (GPM) – H2	- Comparison of NL signal's spectral centroid with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's spectral centroid with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's spectral centroid with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's spectral centroid with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's spectral centroid with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid for signals with 3 (GPM) demand in the looped network: OL > LC > NL > GL > CC. - Order of spectral centroid for signals with 3 (GPM) demand in the branched network: LC > CC > GL > OL > NL.
	7.5 (GPM) – H2	- Comparison of NL signal's spectral centroid with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's spectral centroid with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's spectral centroid with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's spectral centroid with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's spectral centroid with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid for signals with 7.5 (GPM) demand in the looped network: NL > GL > OL > LC > CC. - Order of spectral centroid for signals with 7.5 (GPM) demand in the branched network: LC > OL > NL > GL > CC.

Table 2.52. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Spectral centroid	Transient – H2	- Comparison of NL signal’s spectral centroid with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal’s spectral centroid with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal’s spectral centroid with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal’s spectral centroid with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal’s spectral centroid with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of spectral centroid for signals with transient demand in the looped network: OL > NL > LC > CC > GL. - Order of spectral centroid for signals with transient demand in the branched network: LC > CC = OL > NL > GL.
	All – H2	- Comparing leaks’ spectral centroid magnitudes of two networks shows no consistent change pattern in the magnitudes when the networks change. Therefore, the spectral centroid is not capable of capturing the effects of network changes. - Due to the inconsistent spectral centroid for signals with different demands in both networks, the spectral centroid is not capable of discerning leak types consistently in both networks.

2.12.4.11. Power Spectral Entropy

Figures 2.185 and 2.186 show power spectral entropy bar plots of the acoustic data measured by sensor H1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.187 and 2.188 show the same plots but for sensor H2 data.

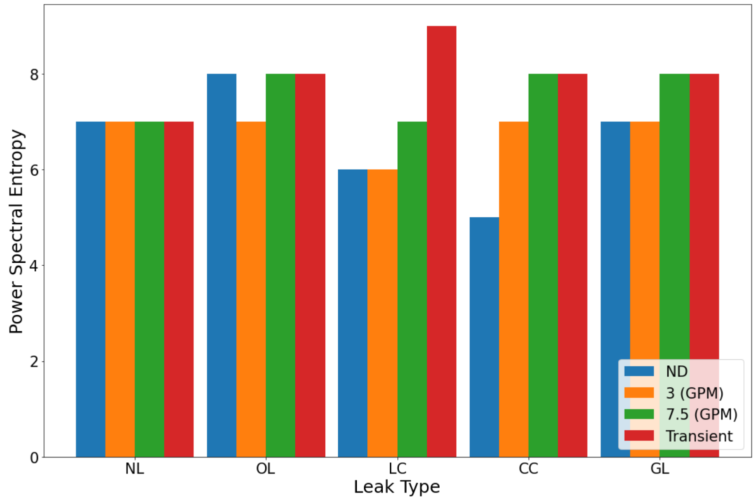

Figure 2.185. Power spectral entropy bar plot of hydrophone H1 data in the looped network for all leaks and demands

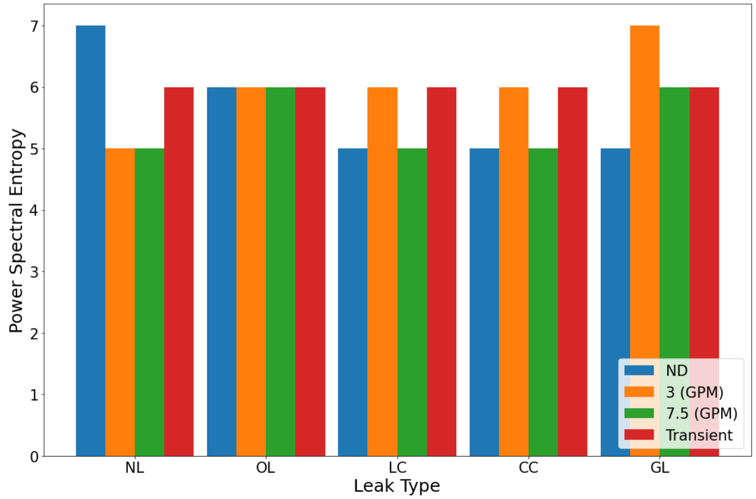

Figure 2.186. Power spectral entropy bar plot of hydrophone H1 data in the branched network for all leaks and demands

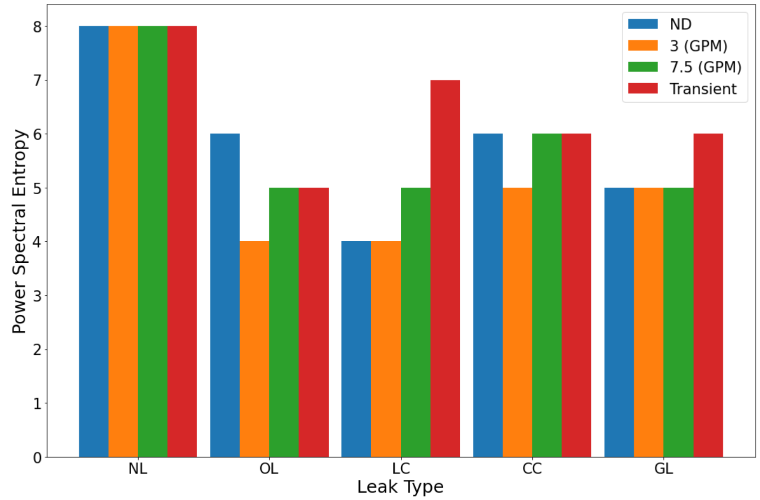

Figure 2.187. Power spectral entropy bar plot of hydrophone H2 data in the looped network for all leaks and demands

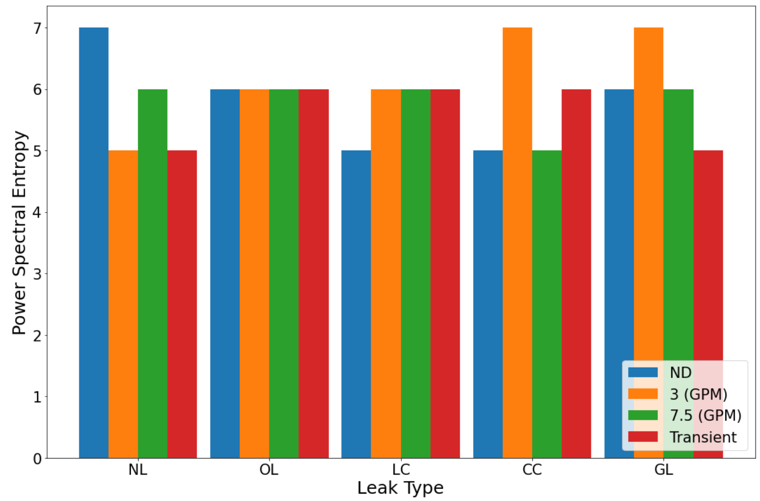

Figure 2.188. Power spectral entropy bar plot of hydrophone H2 data in the branched network for all leaks and demands

Table 2.53 compares power spectral entropy plots of acoustic data measured in the looped and branched networks with all leak and demand types by sensors H1 and H2.

Table 2.53. Analysis of power spectral entropy plots of acoustic data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Power spectral entropy	ND – H1	- Comparison of NL signal's power spectral entropy with no demand: $NL_{lo} = NL_{br}$. - Comparison of OL signal's power spectral entropy centroid with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's power spectral entropy centroid with no demand: $CC_{lo} = CC_{br}$. - Comparison of LC signal's power spectral entropy centroid with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's power spectral entropy centroid with no demand: $GL_{lo} > GL_{br}$. - Order of spectral centroid with ND in the looped network: OL > NL = GL > LC > CC. - Order of spectral centroid with ND in the branched network: NL > OL > GL = LC = CC.
	3 (GPM) – H1	- Comparison of NL signal's power spectral entropy with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's power spectral entropy with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with 3 (GPM) demand: $LC_{lo} = LC_{br}$. - Comparison of GL signal's power spectral entropy with 3 (GPM) demand: $GL_{lo} = GL_{br}$. - Order of power spectral entropy for signals with 3 (GPM) demand in the looped network: NL = OL = CC = GL > LC. - Order of power spectral entropy for signals with 3 (GPM) demand in the branched network: GL > OL = LC = CC > NL.

Table 2.53. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Power spectral entropy	7.5 (GPM) – H1	- Comparison of NL signal's power spectral entropy with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's power spectral entropy with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's power spectral entropy with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the looped network: OL = CC = GL > NL = LC. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the branched network: OL = GL > NL = LC = CC.
	Transient – H1	- Comparison of NL signal's power spectral entropy with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with transient demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's power spectral entropy with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's power spectral entropy with transient demand: $GL_{lo} > GL_{br}$. - Order of power spectral entropy for signals with transient demand in the looped network: LC > OL = CC = GL > NL. - Order of power spectral entropy for signals with transient demand in the branched network: NL = OL = LC = CC = GL = 6.
	All – H1	- Comparing leaks' power spectral entropy magnitudes of the two networks indicates that though the majority of the looped network's power spectral entropies are larger than those of the branched network, there is no consistent pattern in the relation of the networks' spectral entropy. Therefore, power spectral entropy cannot distinguish the network change. - Due to the similarity of power spectral entropy magnitudes for signals with different demands in each network, power spectral entropy cannot discern leak types in both networks.

Table 2.53. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Power spectral entropy	ND – H2	- Comparison of NL signal's power spectral entropy with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with no demand: $OL_{lo} = OL_{br}$. - Comparison of CC signal's power spectral entropy with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's power spectral entropy with no demand: $GL_{lo} < GL_{br}$. - Order of power spectral entropy with ND in the looped network: NL > OL = CC > GL > LC. - Order of power spectral entropy with ND in the branched network: NL > OL = GL > LC = CC.
	3 (GPM) – H2	- Comparison of NL signal's power spectral entropy with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's power spectral entropy with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's power spectral entropy with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's power spectral entropy with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of power spectral entropy for signals with 3 (GPM) demand in the looped network: NL > CC = GL > OL = LC. - Order of power spectral entropy for signals with 3 (GPM) demand in the branched network: CC = GL > LC = OL > NL.

Table 2.53. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Power spectral entropy	7.5 (GPM) – H2	- Comparison of NL signal's power spectral entropy with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's power spectral entropy with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's power spectral entropy with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the looped network: NL > CC > OL = LC = GL. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the branched network: NL = OL = LC = GL > CC.
	Transient – H2	- Comparison of NL signal's power spectral entropy with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's power spectral entropy with transient demand: $CC_{lo} = CC_{br}$. - Comparison of LC signal's power spectral entropy with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's power spectral entropy with transient demand: $GL_{lo} > GL_{br}$. - Order of power spectral entropy for signals with transient demand in the looped network: NL > LC > GL = CC > OL. - Order of power spectral entropy for signals with transient demand in the branched network: OL = LC = CC > NL = GL.
	All – H2	- Based on the power spectral entropy of signals recorded by H2, since there is no constant relation between the power spectral entropies of the two networks, power spectral entropy cannot identify the network change. - Due to the similarity of the spectral entropies of signals with different demands in both networks, the spectral entropy is not capable of discerning leak types consistently in both networks.

2.12.4.12. Mean

Figures 2.189 and 2.190 plot the mean of the hydrophone data measured by sensor H1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.191 and 2.192 show the same plots but for sensor H2 data.

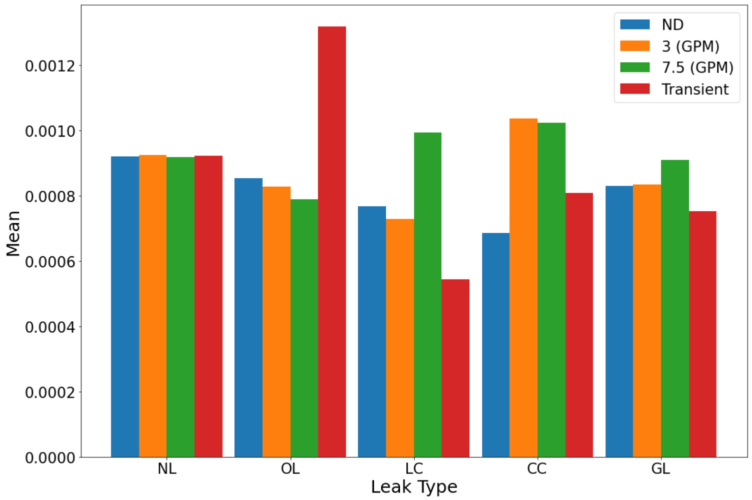

Figure 2.189. Mean plot of hydrophone H1 data in the looped network for all leaks and demands

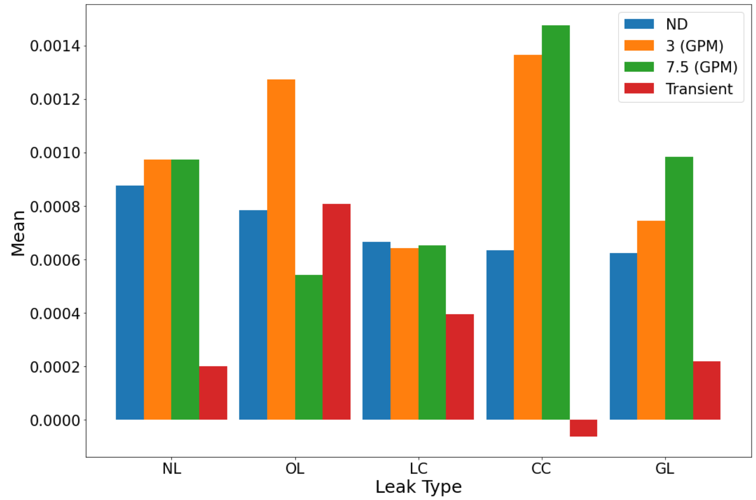

Figure 2.190. Mean plot of hydrophone H1 data in the branched network for all leaks and demands

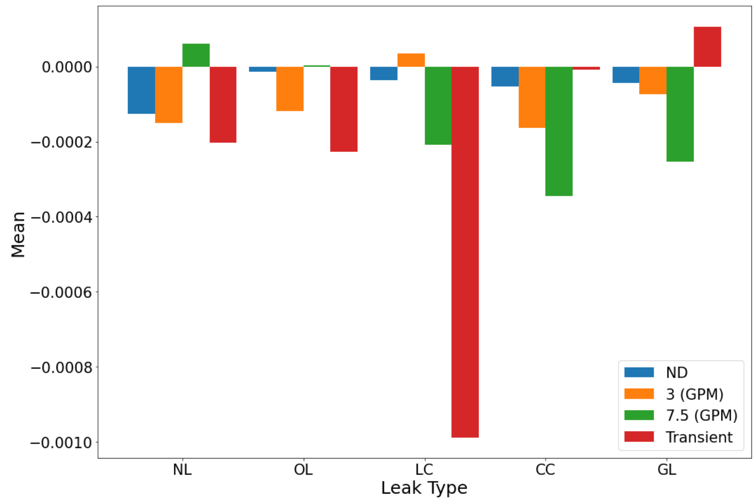

Figure 2.191. Mean plot of hydrophone H2 data in the looped network for all leaks and demands

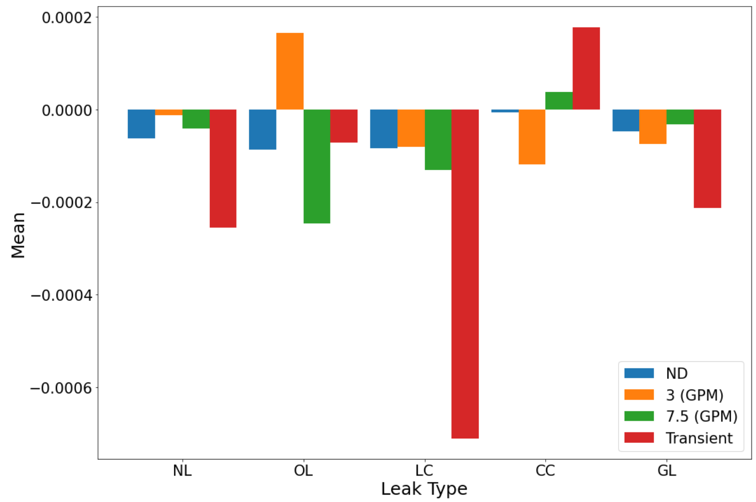

Figure 2.192. Mean plot of hydrophone H2 data in the branched network for all leaks and demands

Table 2.54 compares mean plots of acoustic data measured in the looped and branched networks with all leak and demand types by sensors H1 and H2.

Table 2.54. Analysis of mean plots of acoustic data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Mean	ND – H1	- Comparison of NL signal's mean with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's mean with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's mean with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's mean with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's mean with no demand: $GL_{lo} > GL_{br}$. - Order of mean with ND in the looped network: NL > OL > GL > LC > CC. - Order of mean with ND in the branched network: NL > OL > LC > CC > GL.
	3 (GPM) – H1	- Comparison of NL signal's mean with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's mean with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's mean with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's mean with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's mean with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of mean for signals with 3 (GPM) demand in the looped network: CC > NL > GL > OL > LC. - Order of mean for signals with 3 (GPM) demand in the branched network: CC > OL > NL > GL > LC.
	7.5 (GPM) – H1	- Comparison of NL signal's mean with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's mean with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's mean with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's mean with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's mean with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of mean for signals with 7.5 (GPM) demand in the looped network: CC > LC > NL > GL > OL. - Order of mean for signals with 7.5 (GPM) demand in the branched network: CC > GL > NL > LC > OL.

Table 2.54. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Mean	Transient – H1	- Comparison of NL signal's mean with transient demand: $\mathrm{NL}_{\mathrm{lo}} > \mathrm{NL}_{\mathrm{br}}$. - Comparison of OL signal's mean with transient demand: $\mathrm{OL}_{\mathrm{lo}} > \mathrm{OL}_{\mathrm{br}}$. - Comparison of CC signal's mean with transient demand: $\mathrm{CC}_{\mathrm{lo}} > \mathrm{CC}_{\mathrm{br}}$. - Comparison of LC signal's mean with transient demand: $\mathrm{LC}_{\mathrm{lo}} > \mathrm{LC}_{\mathrm{br}}$. - Comparison of GL signal's mean with transient demand: $\mathrm{GL}_{\mathrm{lo}} > \mathrm{GL}_{\mathrm{br}}$. - Order of mean for signals with transient demand in the looped network: OL > NL > CC > GL > LC. - Order of mean for signals with transient demand in the branched network: OL > LC > GL > NL > CC.
	All – H1	- Comparing leaks' mean magnitudes of the two networks indicates no consistent change pattern in the magnitudes when the network changes. Therefore, mean cannot identify the network's changes. - Due to the inconsistent order of mean of signals with different demands in both networks, mean is not capable of discerning leak types consistently in both networks.
	ND – H2	- Comparison of NL signal's mean with no demand: $\mathrm{NL}_{\mathrm{lo}} < \mathrm{NL}_{\mathrm{br}}$. - Comparison of OL signal's mean with no demand: $\mathrm{OL}_{\mathrm{lo}} > \mathrm{OL}_{\mathrm{br}}$. - Comparison of CC signal's mean with no demand: $\mathrm{CC}_{\mathrm{lo}} < \mathrm{CC}_{\mathrm{br}}$. - Comparison of LC signal's mean with no demand: $\mathrm{LC}_{\mathrm{lo}} > \mathrm{LC}_{\mathrm{br}}$. - Comparison of GL signal's mean with no demand: $\mathrm{GL}_{\mathrm{lo}} > \mathrm{GL}_{\mathrm{br}}$. - Order of mean with ND in the looped network: OL > LC > GL > CC > NL. - Order of mean with ND in the branched network: CC > GL > NL > LC > OL.

Table 2.54. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Mean	3 (GPM) – H2	- Comparison of NL signal's mean with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's mean with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's mean with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's mean with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's mean with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of mean for signals with 3 (GPM) demand in the looped network: LC > GL > OL > NL > CC. - Order of mean for signals with 3 (GPM) demand in the branched network: OL > NL > GL > LC > CC.
	7.5 (GPM) – H2	- Comparison of NL signal's mean with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's mean with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's mean with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's mean with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's mean with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$.
	Transient – H2	- Comparison of NL signal's mean with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's mean with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's mean with transient demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's mean with transient demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's mean with transient demand: $GL_{lo} > GL_{br}$. - Order of mean for signals with transient demand in the looped network: GL > CC > NL > OL > LC. - Order of mean for signals with transient demand in the branched network: CC > OL > GL > NL > LC.
	All – H2	- Comparing the leaks' mean magnitudes of the two networks indicates no consistent change pattern in the magnitudes when the networks change. Therefore, mean cannot identify the network change. - Due to the inconsistent order of mean of signals with different demands in both networks, mean cannot discern leak types consistently in both networks.

2.12.4.13. Standard Deviation

Figures 2.193 and 2.194 show the standard deviation plots of the acoustic data measured by sensor H1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.195 and 2.196 show the same plots but for the sensor H2 data.

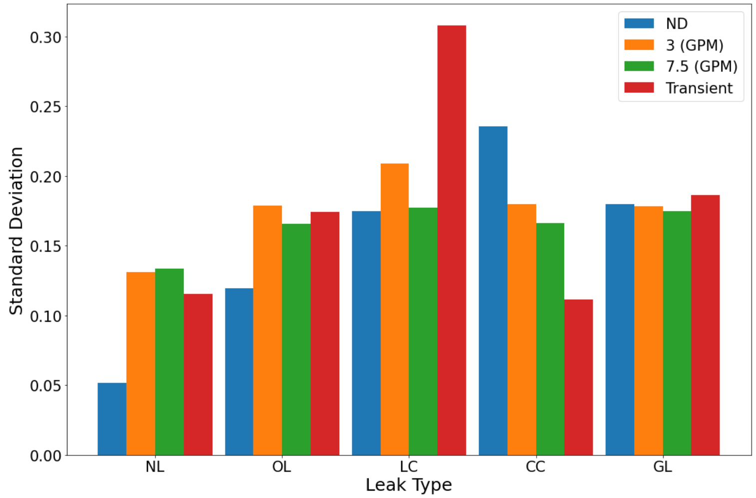

Figure 2.193. Standard deviation plot of hydrophone H1 data in the looped network for all leaks and demands

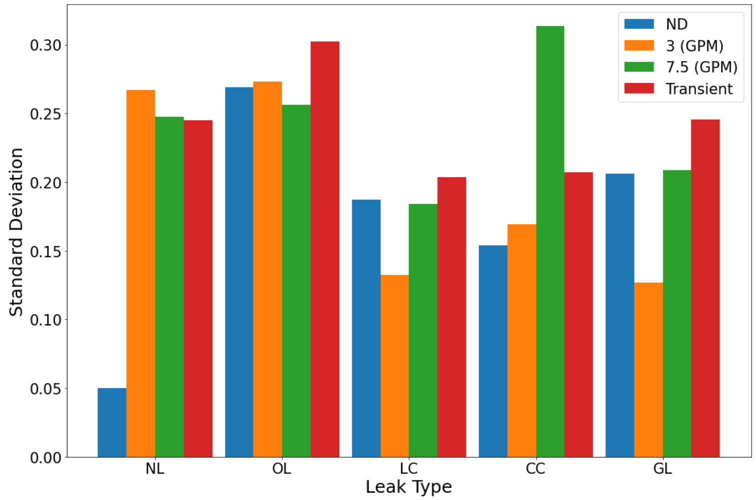

Figure 2.194. Standard deviation plot of hydrophone H1 data in the branched network for all leaks and demands

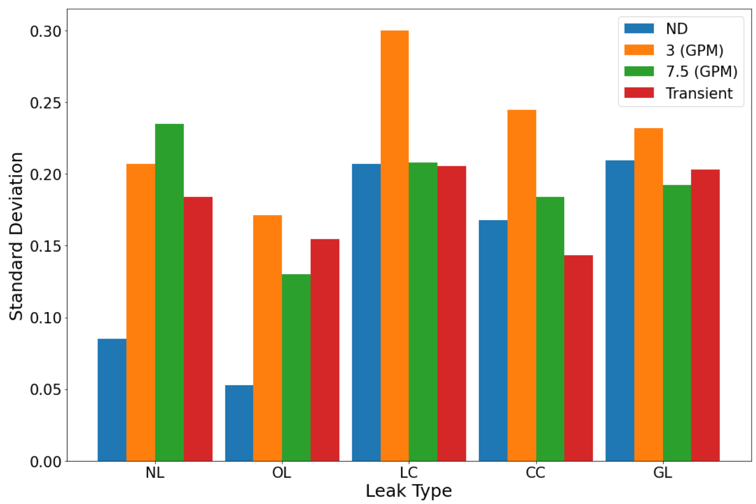

Figure 2.195. Standard deviation plot of hydrophone H2 data in the looped network for all leaks and demands

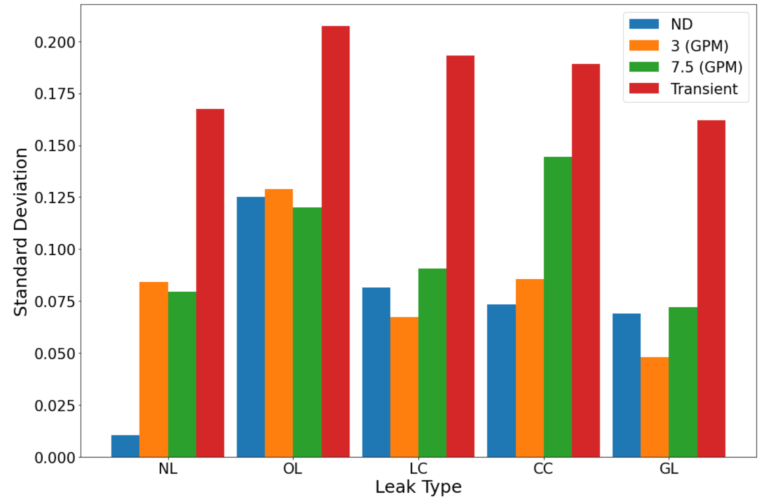

Figure 2.196. Standard deviation plot of hydrophone H2 data in the branched network for all leaks and demands

Table 2.55 compares standard deviation plots of acoustic data measured in the looped and branched networks with all leak and demand types by sensors H1 and H2.

Table 2.55. Analysis of standard deviation plots of acoustic data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Standard deviation	ND – H1	- Comparison of NL signal's standard deviation with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's standard deviation with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's standard deviation with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's standard deviation with no demand: $GL_{lo} > GL_{br}$. - Order of standard deviation with ND in the looped network: GL > LC > CC > NL > OL. - Order of standard deviation with ND in the branched network: OL > GL > LC > CC > NL.
	3 (GPM) – H1	- Comparison of NL signal's standard deviation with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's standard deviation entropy with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation entropy with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's standard deviation with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's standard deviation with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of standard deviation for signals with 3 (GPM) demand in the looped network: LC > CC > OL > GL > NL. - Order of standard deviation for signals with 3 (GPM) demand in the branched network: OL > NL > CC > LC > GL.
	7.5 (GPM) – H1	- Comparison of NL signal's standard deviation with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's standard deviation with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's standard deviation with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's standard deviation with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of standard deviation for signals with 7.5 (GPM) demand in the looped network: LC > GL > CC > OL > NL. - Order of standard deviation for signals with 7.5 (GPM) demand in the branched network: CC > OL > NL > GL > LC.

Table 2.55. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Standard deviation	Transient – H1	- Comparison of NL signal's standard deviation with transient demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's standard deviation with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's standard deviation with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's standard deviation with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's standard deviation with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of standard deviation for signals with transient demand in the looped network: LC > GL > OL > NL > CC. - Order of standard deviation for signals with transient demand in the branched network: OL > GL > NL > CC > LC.
	All – H1	- When there is a demand in the looped network, the standard deviation of LC is the largest of other signals. - Comparing leaks' standard deviation magnitudes of two networks shows no consistent change pattern in the magnitudes when the networks change. Therefore, standard deviation cannot identify the network change. - Due to the inconsistent order of standard deviation for signals with different demands in both networks, standard deviation cannot discern leak types consistently in both networks.
	ND – H2	- Comparison of NL signal's standard deviation with no demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's standard deviation with no demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's standard deviation with no demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's standard deviation with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's standard deviation with no demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of standard deviation with ND in the looped network: GL > LC > CC > NL > OL. - Order of standard deviation with ND in the branched network: OL > LC > CC > GL > NL.

Table 2.55. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Standard deviation	3 (GPM) – H2	- Comparison of NL signal's standard deviation with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's standard deviation with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's standard deviation with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's standard deviation with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's standard deviation with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of standard deviation for signals with 3 (GPM) demand in the looped network: LC > CC > GL > NL > OL. - Order of standard deviation for signals with 3 (GPM) demand in the branched network: OL > CC > NL > LC > GL.
	7.5 (GPM) – H2	- Comparison of NL signal's standard deviation with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's standard deviation with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's standard deviation with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's standard deviation with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's standard deviation with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of standard deviation for signals with 7.5 (GPM) demand in the looped network: NL > LC > GL > CC > OL. - Order of standard deviation for signals with 7.5 (GPM) demand in the branched network: CC > OL > LC > NL > GL.
	Transient – H2	- Comparison of NL signal's standard deviation with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's standard deviation with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation with transient demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's standard deviation with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's standard deviation with transient demand: $GL_{lo} > GL_{br}$. - Order of standard deviation for signals with transient demand in the looped network: LC > GL > NL > OL > CC. - Order of standard deviation for signals with transient demand in the branched network: OL > LC > CC > NL > GL.

Table 2.55. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Standard deviation	All – H2	- Comparing leaks' standard deviation magnitudes of two networks indicates no consistent change pattern in the magnitudes when the networks change. Therefore, standard deviation cannot identify the network change. - Due to the inconsistent order of standard deviation for signals with different demands in both networks, standard deviation is not capable of discerning leak types consistently in both networks.

2.12.4.14. Zero-crossing Rate

Figures 2.197 and 2.198 show the zero-crossing rate plots of the hydrophone data measured by sensor H1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.199 and 2.200 show the same plots but for sensor H2 data.

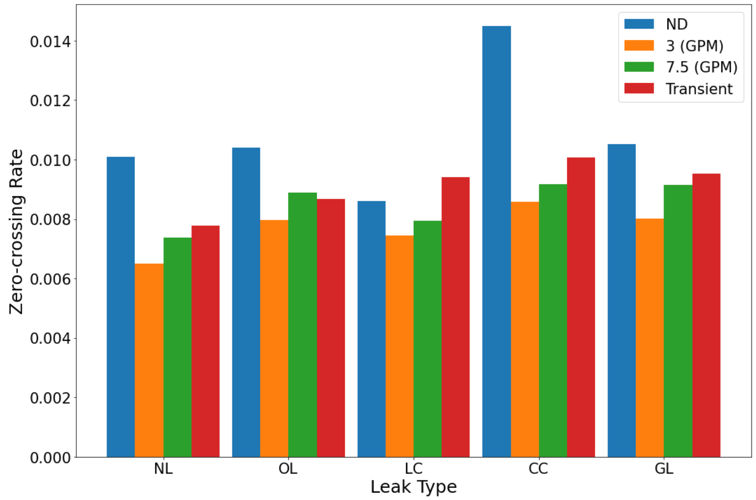

Figure 2.197. Zero-crossing rate plot of hydrophone H1 data in the looped network for all leaks and demands

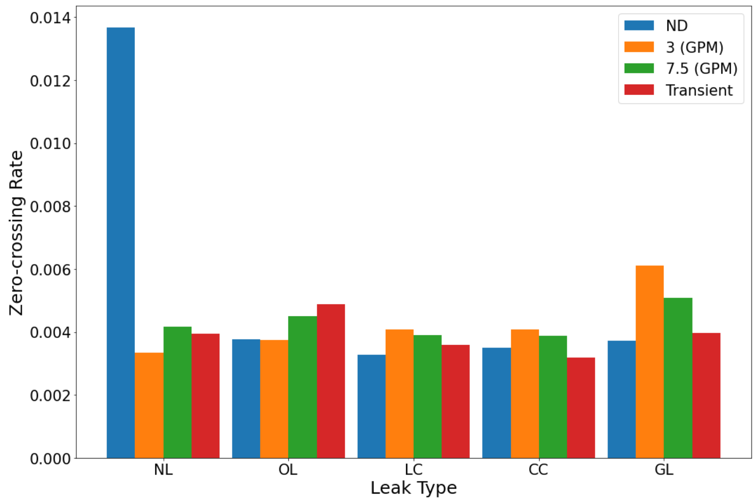

Figure 2.198. Zero-crossing rate plot of hydrophone H1 data in the branched network for all leaks and demands

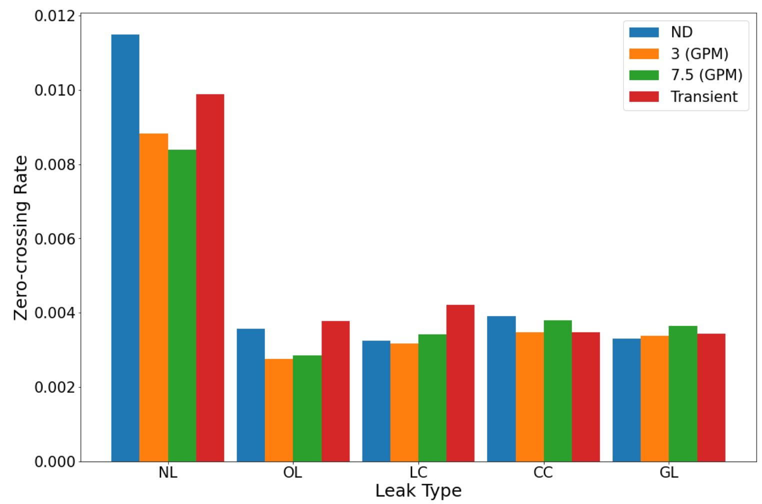

Figure 2.199. Zero-crossing rate plot of hydrophone H2 data in the looped network for all leaks and demands

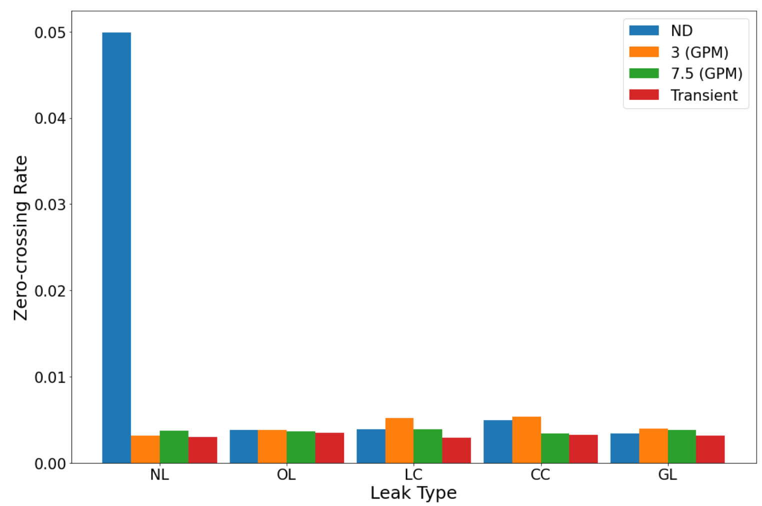

Figure 2.200. Zero-crossing rate plot of hydrophone H2 data in the branched network for all leaks and demands

Table 2.56 compares zero-crossing rate plots of acoustic data measured in the looped and branched networks with all leak and demand types by sensors H1 and H2.

Table 2.56. Analysis of zero-crossing rate plots of acoustic data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Zero-crossing rate	ND – H1	- Comparison of NL signal's zero-crossing rate with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's zero-crossing rate with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's zero-crossing rate with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's zero-crossing rate with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's zero-crossing rate with no demand: $GL_{lo} > GL_{br}$. - Order of zero-crossing rate with ND in the looped network: CC > GL > OL > NL > LC. - Order of zero-crossing rate with ND in the branched network: NL > OL > GL > CC > LC.
	3 (GPM) – H1	- Comparison of NL signal's zero-crossing rate with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's zero-crossing rate with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's zero-crossing rate with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's zero-crossing rate with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's zero-crossing rate with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of zero-crossing rate for signals with 3 (GPM) demand in the looped network: CC > GL > OL > LC > NL. - Order of zero-crossing rate for signals with 3 (GPM) demand in the branched network: GL > LC > CC > OL > NL.

Table 2.56. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Zero-crossing rate	7.5 (GPM) – H1	- Comparison of NL signal's zero-crossing rate with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's zero-crossing rate with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's zero-crossing rate with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's zero-crossing rate with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's zero-crossing rate with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the looped network: CC > GL > OL > LC > NL. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the branched network: GL > OL > NL > LC > CC.
	Transient – H1	- Comparison of NL signal's zero-crossing rate with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's zero-crossing rate with transient demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's zero-crossing rate with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's zero-crossing rate with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's zero-crossing rate with transient demand: $GL_{lo} > GL_{br}$. - Order of zero-crossing rate for signals with transient demand in the looped network: CC > GL > LC > OL > NL. - Order of zero-crossing rate for signals with transient demand in the branched network: OL > GL > NL > LC > CC.
	All – H1	- In the looped network and with all demand types, CC has the highest zero-crossing rate. - When there is a demand, the zero-crossing rate magnitudes of the looped network are larger than those of the branched one. Therefore, zero-crossing rate can help identify the network architecture change. - Due to the inconsistent order of zero-crossing rate of signals with different demands in both networks, zero-crossing rate is not capable of discerning leak types consistently in both networks.

Table 2.56. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Zero-crossing rate	ND – H2	- Comparison of NL signal's zero-crossing rate with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's zero-crossing rate with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's zero-crossing rate with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's zero-crossing rate with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's zero-crossing rate with no demand: $GL_{lo} < GL_{br}$. - Order of zero-crossing rate with ND in the looped network: NL > CC > OL > GL > LC. - Order of zero-crossing rate with ND in the branched network: NL > CC > LC > OL > GL.
	3 (GPM) – H2	- Comparison of NL signal's zero-crossing rate with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's zero-crossing rate with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's zero-crossing rate with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's zero-crossing rate with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's zero-crossing rate with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of zero-crossing rate for signals with 3 (GPM) demand in the looped network: NL > CC > GL > LC > OL. - Order of zero-crossing rate for signals with 3 (GPM) demand in the branched network: CC > LC > GL > OL > NL. This order is the same as the order of A2 data.
	7.5 (GPM) – H2	- Comparison of NL signal's zero-crossing rate with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's zero-crossing rate with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's zero-crossing rate with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's zero-crossing rate with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's zero-crossing rate with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the looped network: NL > CC > GL > LC > OL. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the branched network: LC > GL > NL > OL > CC.

Table 2.56. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Zero-crossing rate	Transient – H2	- Comparison of NL signal’s zero-crossing rate with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal’s zero-crossing rate with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal’s zero-crossing rate with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal’s zero-crossing rate with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal’s zero-crossing rate with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of zero-crossing rate for signals with transient demand in the looped network: NL > LC > OL > CC > GL. - Order of zero-crossing rate for signals with transient demand in the branched network: OL > CC > GL > NL > LC.
	All – H2	- Comparing leaks’ zero-crossing rate magnitudes of the two networks indicates no consistent change pattern in the magnitudes when the networks change. Therefore, zero-crossing rate cannot identify the network change. - Due to the inconsistent order of zero-crossing rate for signals with different demands in both networks, zero-crossing rate is not capable of discerning leak types consistently in both networks. - In the looped network, NL signal has the largest zero-crossing rate. However, based on A2, in both networks, where demand was present, NL signal had the lowest zero-crossing rate.

2.12.4.15. RMS

Figures 2.201 and 2.202 show the RMS plots of the hydrophone data measured by sensor H1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.203 and 2.204 show the same plots but for sensor H2 data.

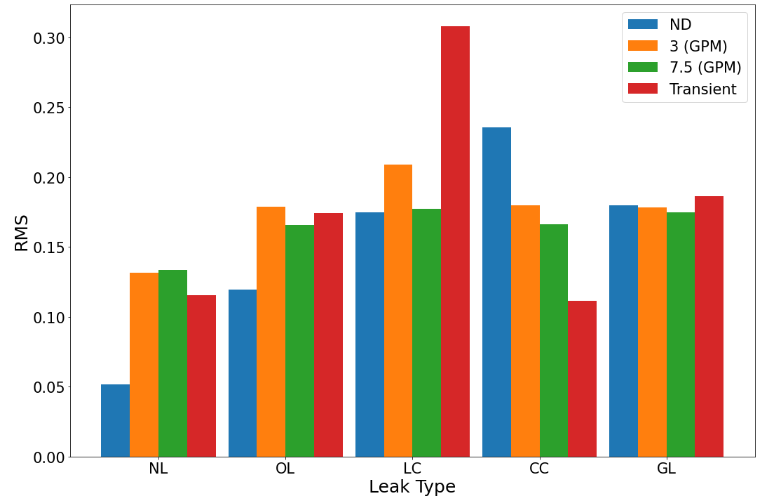

Figure 2.201. RMS plot of hydrophone H1 data in the looped network for all leaks and demands

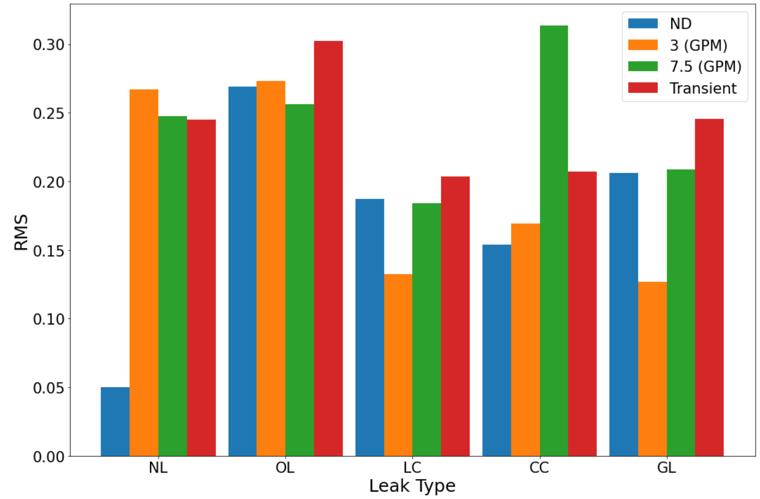

Figure 2.202. RMS plot of hydrophone H1 data in the branched network for all leaks and demands

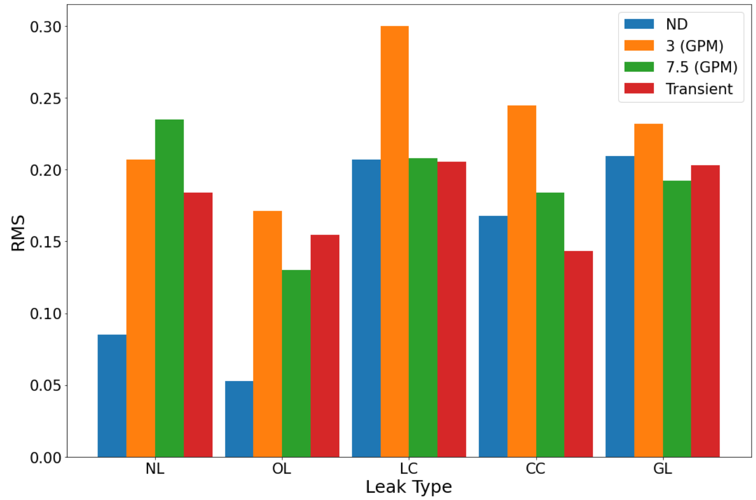

Figure 2.203. RMS plot of hydrophone H2 data in the looped network for all leaks and demands

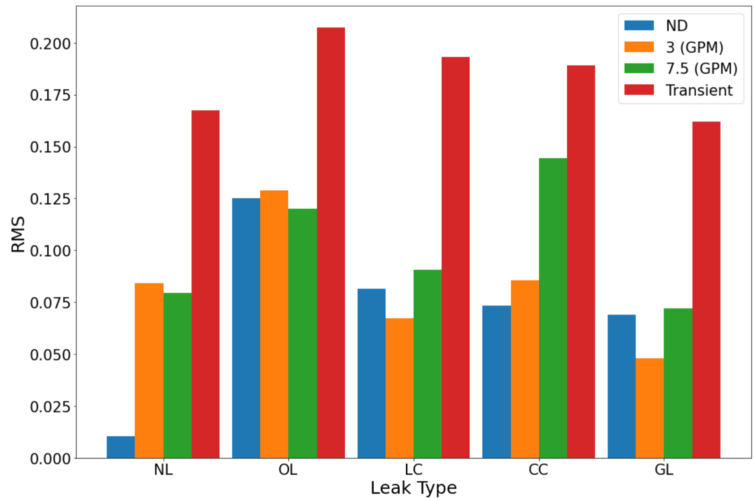

Figure 2.204. RMS plot of hydrophone H2 data in the branched network for all leaks and demands

Table 2.57 compares RMS plots of acoustic data measured in the looped and branched networks with all leak and demand types by sensors H1 and H2.

Table 2.57. Analysis of RMS plots of acoustic data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
RMS	ND – H1	- Comparison of NL signal's RMS with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's RMS with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's RMS with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's RMS with no demand: $GL_{lo} < GL_{br}$. - Order of RMS with ND in the looped network: CC > GL > LC > OL > NL. - Order of RMS with ND in the branched network: OL > GL > LC > CC > NL.
	3 (GPM) – H1	- Comparison of NL signal's RMS with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's RMS with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's RMS with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's RMS with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of RMS for signals with 3 (GPM) demand in the looped network: LC > CC > OL > GL > NL. - Order of RMS for signals with 3 (GPM) demand in the branched network: OL > NL > CC > LC > GL.

Table 2.57. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
RMS	7.5 (GPM) – H1	- Comparison of NL signal's RMS with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's RMS with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's RMS with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's RMS with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of RMS for signals with 7.5 (GPM) demand in the looped network: LC > GL > CC > OL > NL. - Order of RMS for signals with 7.5 (GPM) demand in the branched network: CC > OL > NL > GL > LC.
	Transient – H1	- Comparison of NL signal's RMS with transient demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's RMS with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with transient demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's RMS with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's RMS with transient demand: $GL_{lo} < GL_{br}$. - Order of RMS for signals with transient demand in the looped network: LC > GL > OL > NL > CC. - Order of RMS for signals with transient demand in the branched network: OL > GL > NL > CC > LC.
	All – H1	- When there is a demand in the looped network, LC has the largest RMS magnitude compared to other signals. - Comparing leaks' RMS magnitudes of the two networks indicates no consistent change pattern in the magnitudes when the network changes. Therefore, RMS cannot identify the network change. - Due to the inconsistent order of RMS for signals with different demands in both networks, RMS is not capable of discerning leak types consistently in both networks.

Table 2.57. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
RMS	ND – H2	- Comparison of NL signal's RMS with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's RMS with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's RMS with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's RMS with no demand: $GL_{lo} > GL_{br}$. - Order of RMS with ND in the looped network: GL > LC > CC > NL > OL. - Order of RMS with ND in the branched network: OL > LC > CC > GL > NL.
	3 (GPM) – H2	- Comparison of NL signal's RMS with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's RMS with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's RMS with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's RMS with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's RMS with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of RMS for signals with 3 (GPM) demand in the looped network: LC > CC > GL > NL > OL. - Order of RMS for signals with 3 (GPM) demand in the branched network: OL > CC > NL > LC > GL.
	7.5 (GPM) – H2	- Comparison of NL signal's RMS with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's RMS with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's RMS with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's RMS with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's RMS with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of RMS for signals with 7.5 (GPM) demand in the looped network: NL > LC > GL > CC > OL. - Order of RMS for signals with 7.5 (GPM) demand in the branched network: CC > OL > LC > NL > GL.

Table 2.57. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
RMS	Transient – H2	- Comparison of NL signal's RMS with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's RMS with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's RMS with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's RMS with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's RMS with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of RMS for signals with transient demand in the looped network: LC > GL > NL > OL > CC. - Order of RMS for signals with transient demand in the branched network: OL > LC > CC > NL > GL.
	All – H2	- Based on sensor H2 in the branched network, when demand is present, GL signal has the smallest RMS magnitude. - Comparing the RMS magnitudes of the two networks indicates no consistent change pattern in the magnitudes when the network changes. Therefore, RMS cannot identify the network change. - Due to the inconsistent order of RMS for signals with different demands in both networks, RMS is not capable of discerning leak types consistently in both networks.

2.12.4.16. Crest Factor

Figures 2.205 and 2.206 show the crest factor plots of the hydrophone data measured by sensor H1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.207 and 2.208 show the same plots but for sensor H2 data.

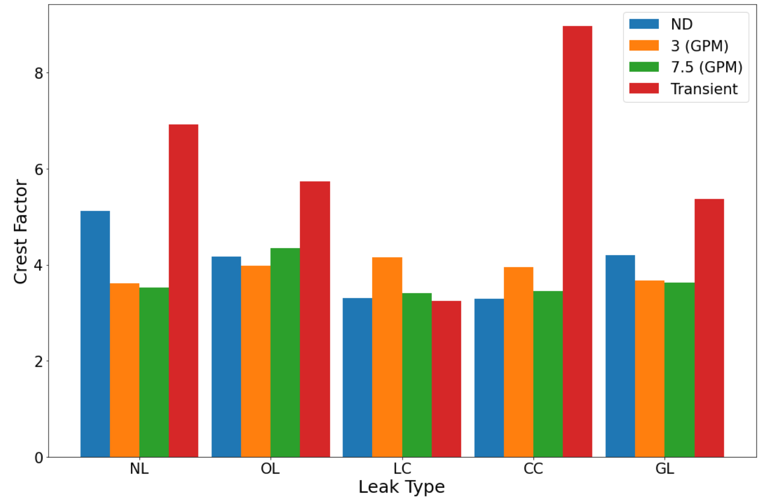

Figure 2.205. Crest factor plot of hydrophone H1 data in the looped network for all leaks and demands

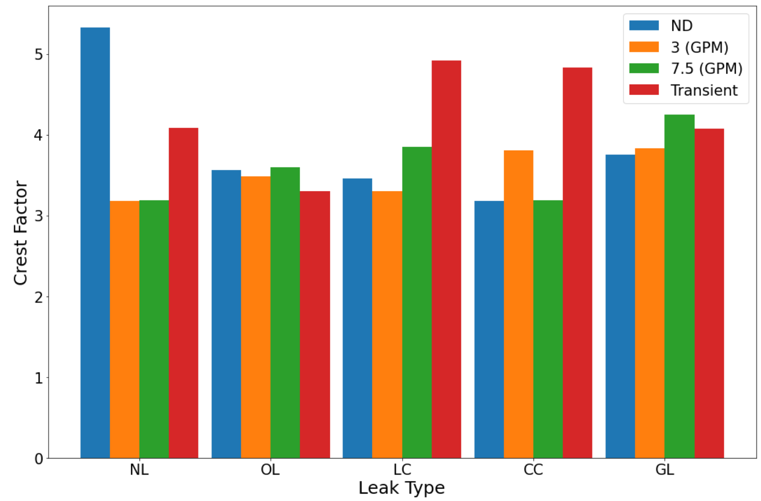

Figure 2.206. Crest factor plot of hydrophone H1 data in the branched network for all leaks and demands

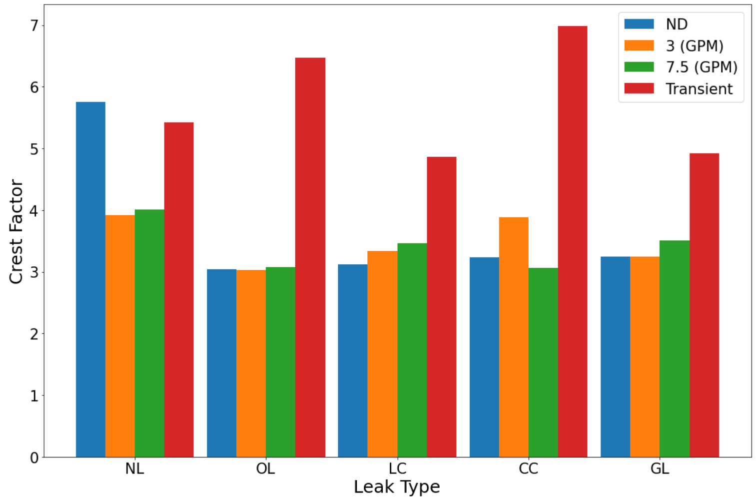

Figure 2.207. Crest factor plot of hydrophone H2 data in the looped network for all leaks and demands

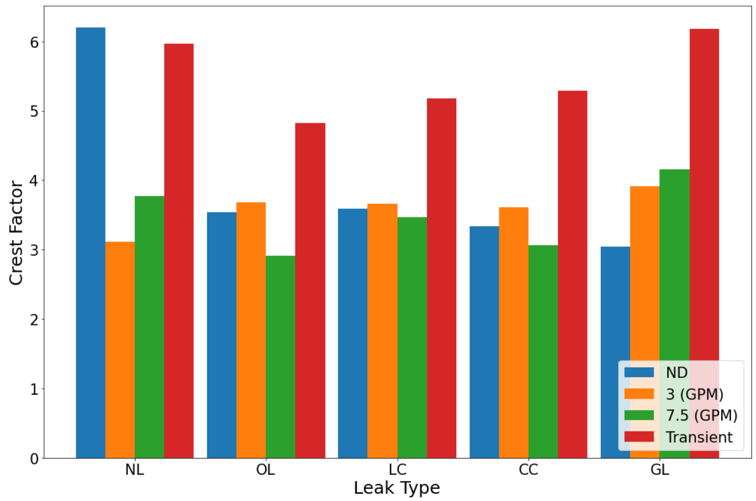

Figure 2.208. Crest factor plot of hydrophone H2 data in the branched network for all leaks and demands

Table 2.58 compares crest factor plots of acoustic data measured in the looped and branched networks with all leak and demand types by sensors H1 and H2.

Table 2.58. Analysis of crest factor plots of acoustic data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Crest factor	ND – H1	- Comparison of NL signal's crest factor with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's crest factor with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's crest factor with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's crest factor with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's crest factor with no demand: $GL_{lo} > GL_{br}$. - Order of crest factor with ND in the looped network: NL > GL > OL > LC > CC. - Order of crest factor with ND in the branched network: NL > GL > OL > LC > CC.
	3 (GPM) – H1	- Comparison of NL signal's crest factor with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's crest factor with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's crest factor with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's crest factor with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's crest factor with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of crest factor for signals with 3 (GPM) demand in the looped network: LC > OL > CC > GL > NL. - Order of crest factor for signals with 3 (GPM) demand in the branched network: GL > CC > OL > LC > NL.

Table 2.58. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Crest factor	7.5 (GPM) – H1	- Comparison of NL signal’s crest factor with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal’s crest factor with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal’s crest factor with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal’s crest factor with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal’s crest factor with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of crest factor for signals with 7.5 (GPM) demand in the looped network: OL > GL > NL > CC > LC. - Order of crest factor for signals with 7.5 (GPM) demand in the branched network: GL > LC > OL > NL > CC.
	Transient – H1	- Comparison of NL signal’s crest factor with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal’s crest factor with transient demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal’s crest factor with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal’s crest factor with transient demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal’s crest factor with transient demand: $GL_{lo} > GL_{br}$. - Order of crest factor for signals with transient demand in the looped network: CC > NL > OL > GL > LC. - Order of crest factor for signals with transient demand in the branched network: LC > CC > NL > GL > OL.
	All – H1	- Comparing leaks’ crest factor magnitudes of the two networks indicates no consistent change pattern in the magnitudes when the networks change. Therefore, crest factor cannot identify the network change. - Since the order of signals’ crest factor with different demands in both networks are inconsistent, crest factor cannot discern leak types consistently in both networks.

Table 2.58. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Crest factor	ND – H2	- Comparison of NL signal's crest factor with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's crest factor with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's crest factor with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's crest factor with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's crest factor with no demand: $GL_{lo} > GL_{br}$. - Order of crest factor with ND in the looped network: NL > GL > CC > LC > OL. - Order of crest factor with ND in the branched network: NL > LC > OL > CC > GL.
	3 (GPM) – H2	- Comparison of NL signal's crest factor with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's crest factor with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's crest factor with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's crest factor with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's crest factor with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of crest factor for signals with 3 (GPM) demand in the looped network: NL > CC > LC > GL > OL. - Order of crest factor for signals with 3 (GPM) demand in the branched network: GL > OL > LC > CC > NL.
	7.5 (GPM) – H2	- Comparison of NL signal's crest factor with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's crest factor with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's crest factor with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's crest factor with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's crest factor with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of crest factor for signals with 7.5 (GPM) demand in the looped network: NL > GL > LC > OL > CC. - Order of crest factor for signals with 7.5 (GPM) demand in the branched network: GL > NL > LC > CC > OL.

Table 2.58. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Crest factor	Transient – H2	- Comparison of NL signal's crest factor with transient demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's crest factor with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's crest factor with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's crest factor with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's crest factor with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of crest factor for signals with transient demand in the looped network: CC > OL > NL > GL > LC. - Order of crest factor for signals with transient demand in the branched network: GL > NL > CC > LC > OL.
	All – H2	- Comparing leaks' crest factor magnitudes of two networks indicates no consistent change pattern in the magnitudes when the networks change. Therefore, crest factor cannot identify the network change. - Due to the inconsistent order of crest factor for signals with different demands in both networks, crest factor is not capable of discerning leak types consistently in both networks. - When there is a demand in the branched network, GL has the largest crest factor than other signals.

2.12.4.17. Summary of Hydrophone Measurement Evaluations

The following are some important takeaways from the numerical and visual evaluations of the hydrophone data.

Like accelerometer data and based on the time-domain hydrophone plots, the relation between the leak and no leak signals depends on the hydrophones' location. When there is no demand and no background noise, and regarding H1 data in both networks and H2 measurements in the looped network, leak acoustic data are larger than those of NL. However, the H2 data of the branched network does not represent a consistent relation between the leak and NL signals. The larger leak signal amplitudes result from the acoustic disturbances caused by the leaks' water outputs propagated through the pipeline. GL has the largest amplitudes in the looped network than other leaks, which may be due to its outflow's louder sounds.

One can observe that the NL signal amplitudes are larger in the looped network than those in the branched one. Due to the discontinuity of pipes and more dead ends in the branched network, acoustic signals are more attenuated, and the hydrophones sense less intense acoustic emissions. Comparison of leak signals between the two networks does not indicate any specific magnitude pattern. However, leak signal amplitudes in the looped network are more uniform than those of the branched one. This non-uniformity can be due to the less straight paths between the leaks and the hydrophones and more barriers in the branched network, which cause more frequent signal attenuation and resonance.

Amplitudes of the H1 signals are less variable than the H2 signals. Since H1 is close to the pump, it recorded a constant background noise generated by the pump. This difference can highlight the effect of a hydrophone's location on its measurements.

Like accelerometers, hydrophone frequency caps depend on sensor locations. Based on H1, in a no demand and no noise scenario, NL signals have a maximum frequency of 400 Hz, and leak signals' frequency caps varied between 600 Hz for the OL in the looped network and 300 Hz for the CC in the branched network. In both networks and based on both hydrophones, OL and GL with larger leak flows have higher frequency caps than CC and LC with lower leak flow rates. Therefore, there is a direct relation between leak flow rates and leak frequency caps. However, regarding H2 data, NL signals' maximum frequencies are either larger than or equal to the leak signals. Given these different frequency caps, the maximum frequencies of leaks are larger than, equal to, or lower than those of the NLs. The acoustic data with no demand and no background noise in both networks reveal that the amplitudes of leaks' dominant frequencies are larger than those of the NL signals, a sign of leaks in the acoustic data. The maximum frequencies of the leak and NL signals in the looped network are larger than those of the branched network. This difference is due to the more intense signal attenuations in the branched network caused by disconnected pipes and blind flanges, which prohibit high-frequency signals reach hydrophones.

Regarding the cumulative distribution plots of signals with no demand and no background noise, leak signal magnitudes are larger than the NL ones in the branched network that is the leaks' effects. Leak signal magnitudes are more distinct in the branched network than their looped network's counterparts. Since there is only one path from the leak location to each hydrophone and regarding dissimilar sound emission characteristics of each leak, acoustic signals become differently dissipated on their ways to the hydrophones. On the other hand, due to more paths from the leak location to the looped network's hydrophones, leaks' acoustic emissions propagate more uniformly via different pipe trajectories. GL and CC have the largest magnitudes in the looped and branched networks, respectively. We observed that these leaks constituted more air in their output flows and generated more sounds which caused their larger magnitudes.

Similar to the cumulative distribution plots, box plots demonstrate GL and CC have the largest magnitudes in the looped and branched networks, respectively. As it was mentioned, these leaks' water jets constituted more air streams and emitted more intense sound signals. Though there is no consistent pattern in the relations of the two networks' signal magnitude continua based on H1 measurements, H2 data indicate larger signal magnitude continua in the looped network than those of the branched one.

There is no constant pattern in the relations of the two networks' areas under the CSD plots with no demand and no background noise. Therefore, the areas under the CSD plots cannot represent the change in the network architecture. Similar to accelerometer measurements, the larger area under the CSD plot of the NL signal in the looped network is due to the similarity of spectral density magnitudes of the NL signals at H1 and H2. For example, the NL signals' cap frequencies at H1 and H2 in the looped network were 400 Hz, while the leak signals' cap frequencies were between 500 Hz and 400 Hz at H1 and 200 Hz at H2. On the other hand, in the branched network, the leak signals' maximum frequencies at the hydrophones were more similar than those of the NL signals. For instance, the maximum frequency of CC in the branched network at H1 and H2 were 300 Hz and 200 Hz, respectively, which led to the signals' spectral similarity and a reason for the CC signal's larger area under the CSD plot in the branched network.

There were no specific patterns or thresholds in the LDIs of the looped and branched networks to discern leak types and network architectures. In the looped network with a non-zero demand, the cracks' LDIs are larger than GL and OL. The larger LDIs correspond to larger differences between the CSDs of the leak and no leak signals. The more significant LDIs of the cracks can be due to their water jets' characteristics that constitute more airflows and generate more intense sounds than OL. LDI decreased when demand changed from ND to 3 (GPM) and 7.5 (GPM) in the looped network. The decrease in the looped network's LDIs reflects more resemblance between the leaks' CSDs and no leak. When the demand increases, water velocities in the pipes become larger to deliver more water to the demand node. Evaluation of the looped network's EPANET model indicated that due to the network architecture's symmetry, the length of paths that leak sound should traverse in and against the water flow directions to reach the hydrophones are the same. Since leak acoustic signals are propagated by water in the pipes and not by PVC pipe walls, the main factor affecting leak sound propagation is water velocity. Water flows attenuate leak sounds more intensely when water velocity is higher. The more significant attenuation causes leak sounds to reach hydrophones with weaker intensities, leading to more resemblance between the CSDs of the leak and no leak signals in the looped network with higher demand.

Based on both hydrophone data, GL and OL have the largest and smallest sum of leak:noleak magnitudes, respectively, in the looped network. Regarding Figures 2.18 to 2.21, OL had a water jet with the largest height and a circular cross-section, mainly filled with water and constituted less airflow. On the other hand, GL water output had an irregular cross-section, a less uniform shape, and more sound. Therefore, we can conclude that leak signal magnitudes depend on the shape of the leak water jet or outflow rather than its water jet height. The shape of leaks' water output also affected their dominant frequencies. For example, based on H2, GL in the looped network with a dominant frequency of 40 Hz had the largest dominant frequency compared to other leaks.

The majority of features listed in Table 2.24 could not identify the network change or discern leak types except for the cases highlighted below that depend on sensor locations. Therefore, their results cannot be generalized. But they provide some insights about informative features for hydrophone data.

Similar to other parameters, information gained from fundamental frequency depends on hydrophone locations. Regarding H1 data, the looped network's leak signals' fundamental frequencies were between 76 Hz and 80 Hz, and those of the branched network's leak signals were all 0 Hz. Therefore, the fundamental frequencies of H1 measurements can help discern network changes. However, this was not the case for H2 measurements' fundamental frequencies.

Based on H1, when there was a demand, looped network's zero-crossing rate magnitudes were larger than those of the branched one. Therefore, a zero-crossing rate can help identify the network architecture change.

2.12.5. Dynamic Pressure Sensor Measurements

In this subsection, we analyzed the data of the dynamic pressure sensors P1 and P2, where ND stands for no demand.

2.12.5.1. Time-domain plot (for ND signals)

Figures 2.209 to 2.116 show the time-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor P1.

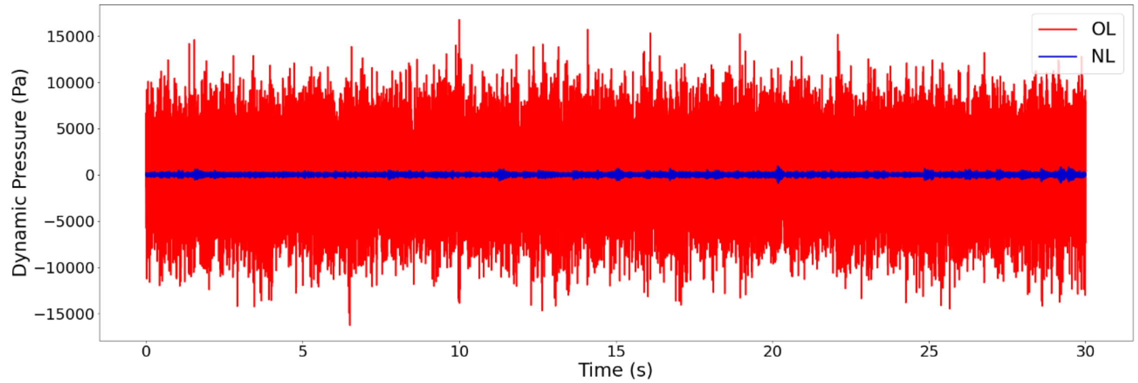

Figure 2.209. Time-domain plots of OL vs. NL signals in the looped network measured by sensor P1

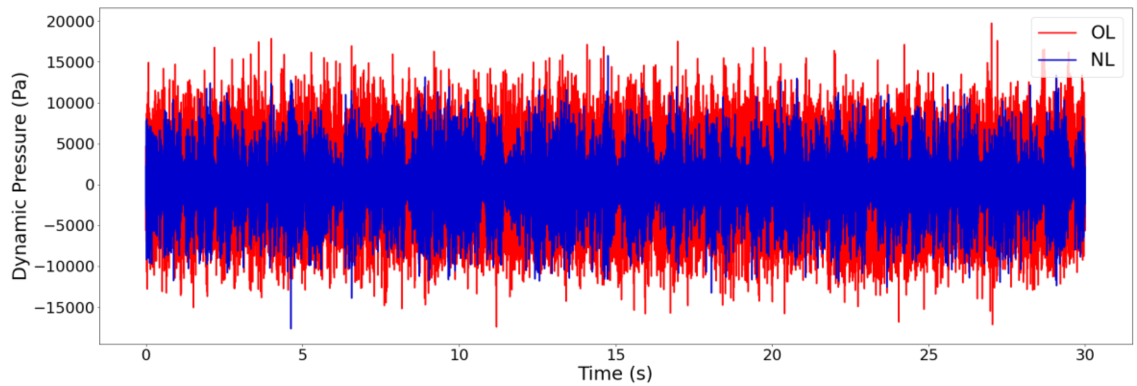

Figure 2.210. Time-domain plots of OL vs. NL signals in the branched network measured by sensor P1

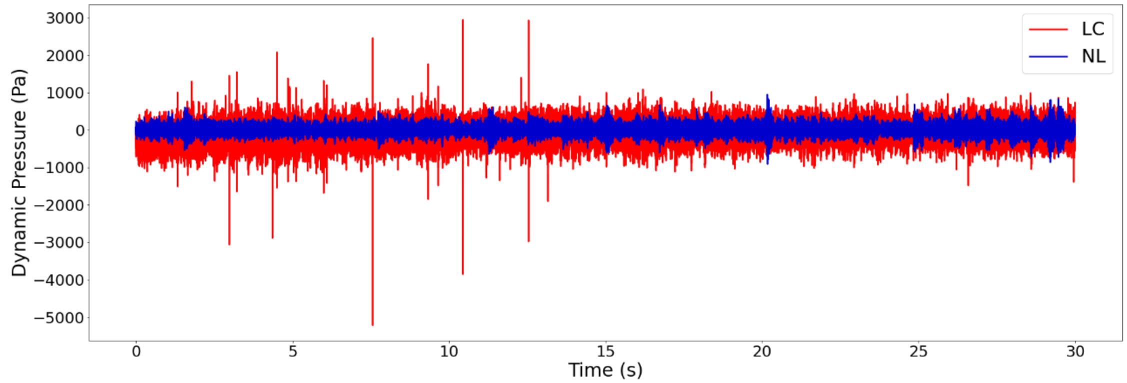

Figure 2.211. Time-domain plots of LC vs. NL signals in the looped network measured by sensor P1

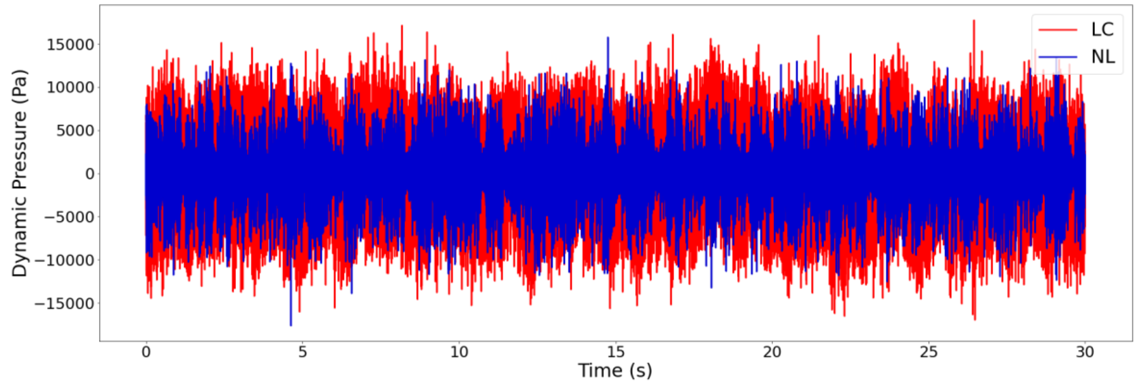

Figure 2.212. Time-domain plots of LC vs. NL signals in the branched network measured by sensor P1

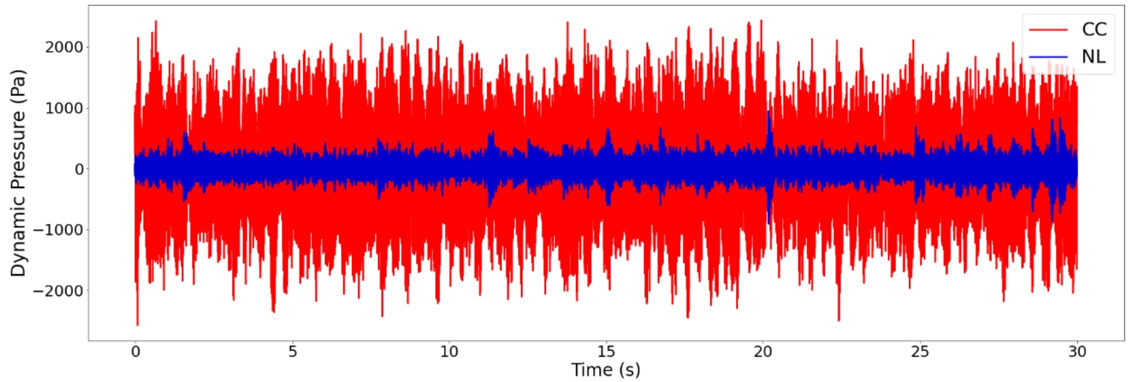

Figure 2.213. Time-domain plots of CC vs. NL signals in the looped network measured by sensor P1

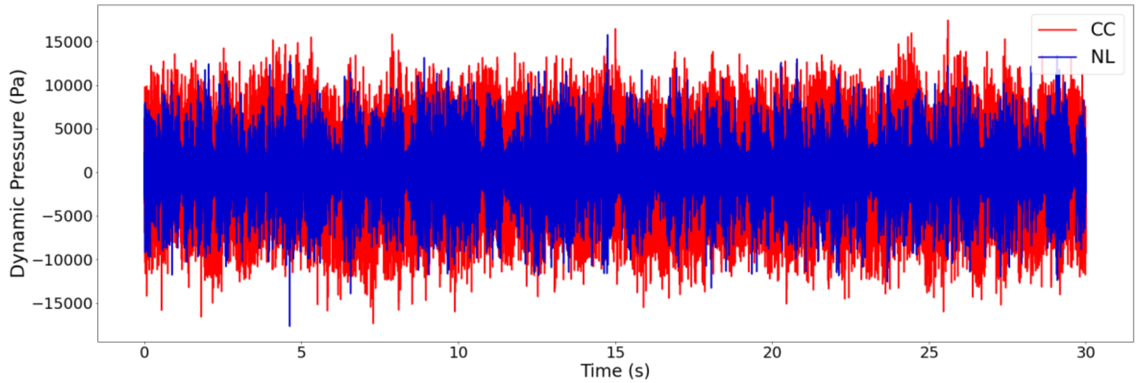

Figure 2.214. Time-domain plots of CC vs. NL signals in the branched network measured by sensor P1

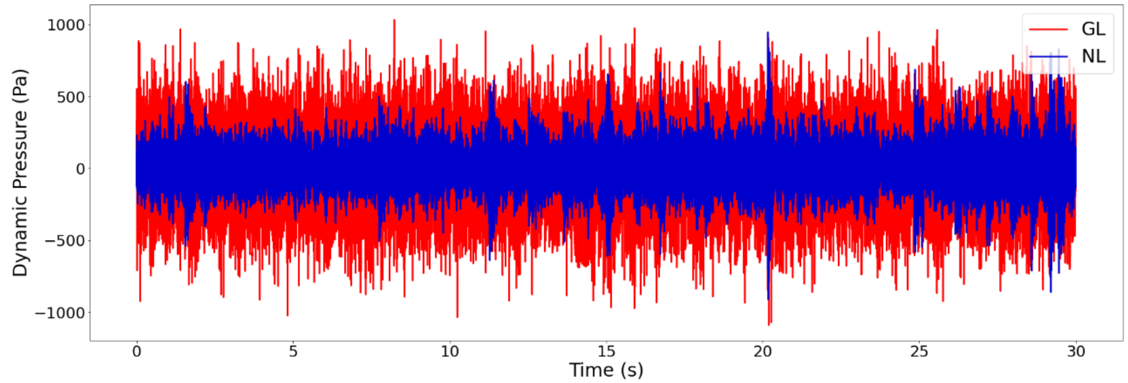

Figure 2.215. Time-domain plots of GL vs. NL signals in the looped network measured by sensor P1

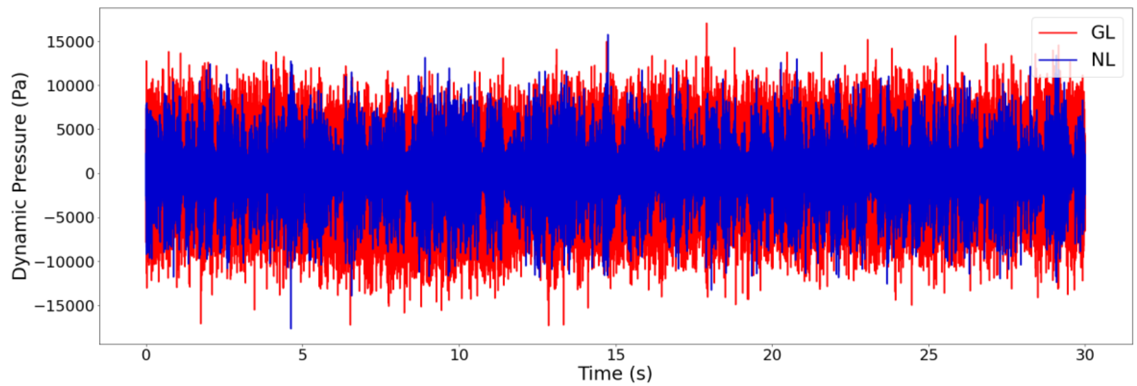

Figure 2.216. Time-domain plots of GL vs. NL signals in the branched network measured by sensor P1

Figures 2.217 to 2.224 show the time-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor P2.

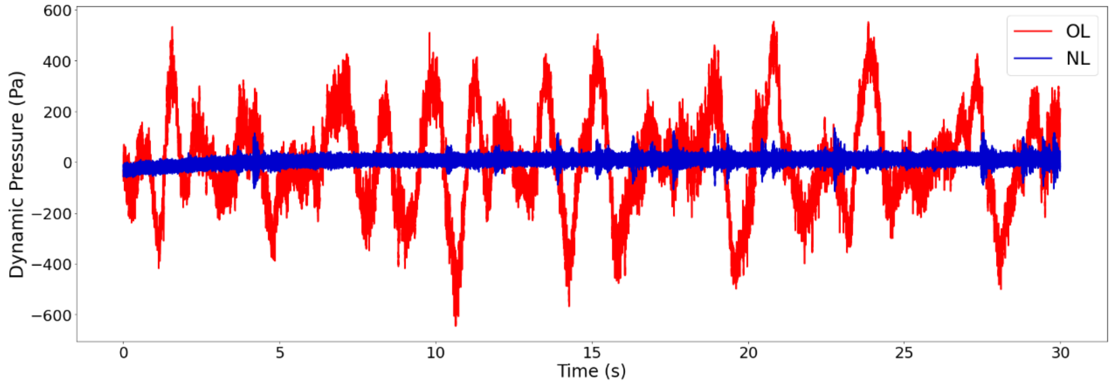

Figure 2.217. Time-domain plots of OL vs. NL signals in the looped network measured by sensor P2

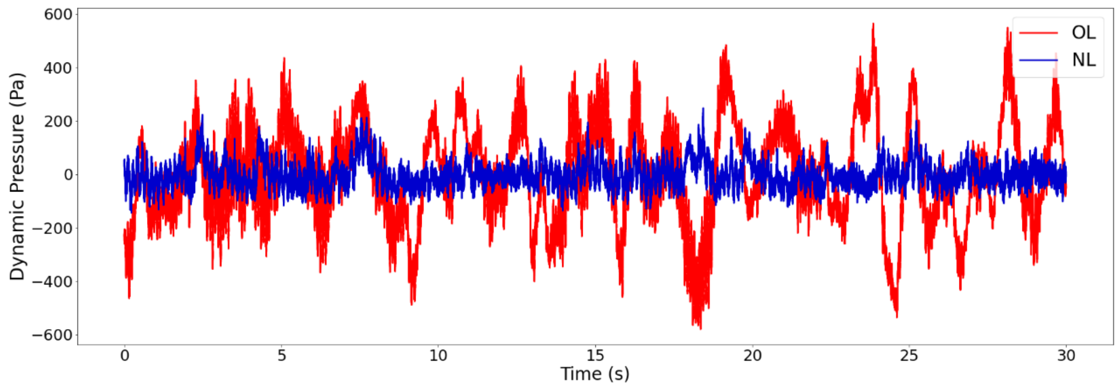

Figure 2.218. Time-domain plots of OL vs. NL signals in the branched network measured by sensor P2

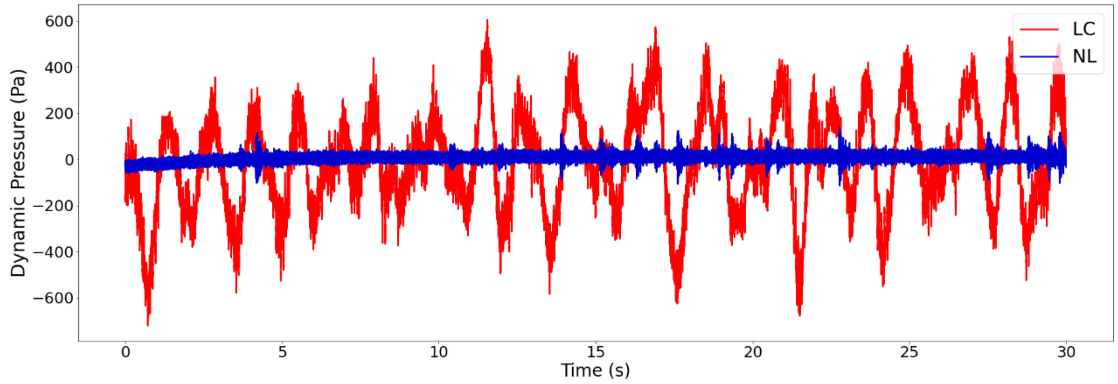

Figure 2.219. Time-domain plots of LC vs. NL signals in the looped network measured by sensor P2

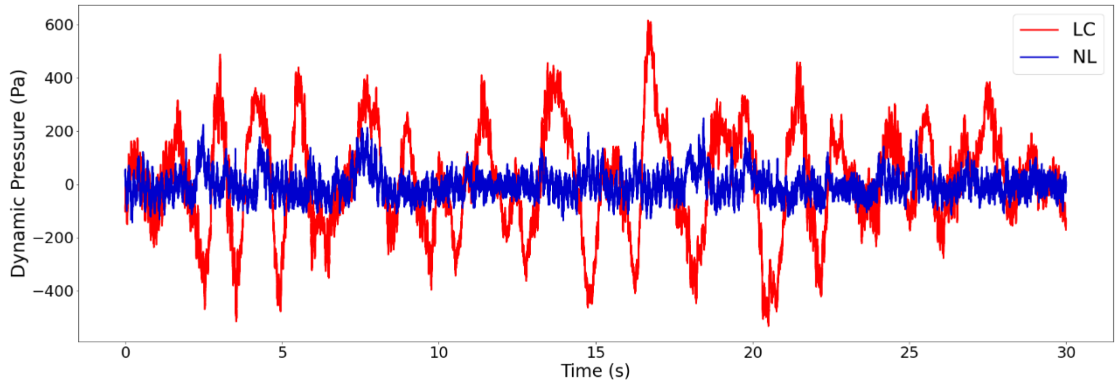

Figure 2.220. Time-domain plots of LC vs. NL signals in the branched network measured by sensor P2

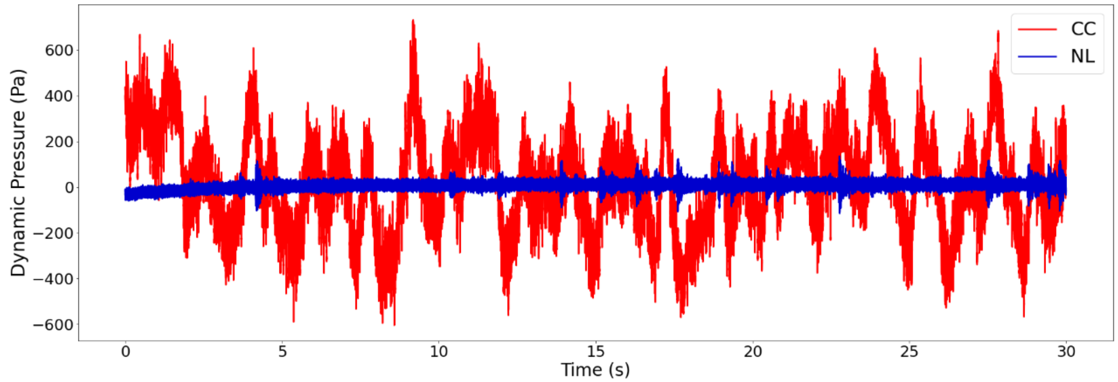

Figure 2.221. Time-domain plots of CC vs. NL signals in the looped network measured by sensor P2

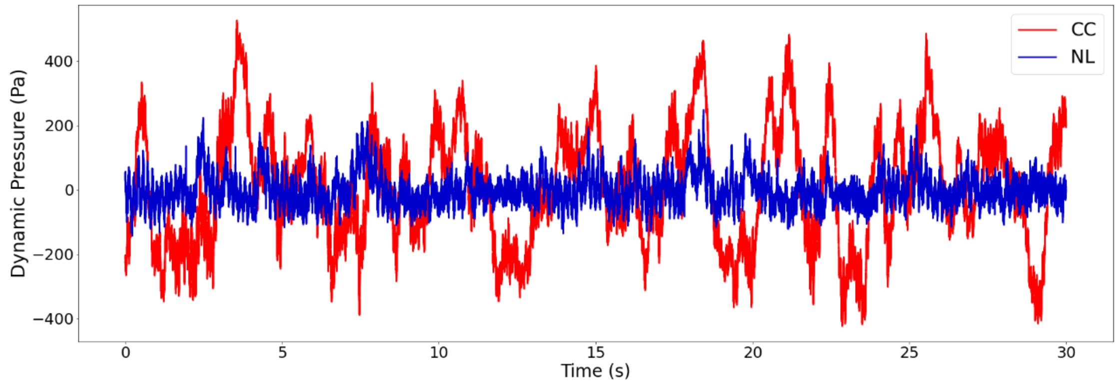

Figure 2.222. Time-domain plots of CC vs. NL signals in the branched network measured by sensor P2

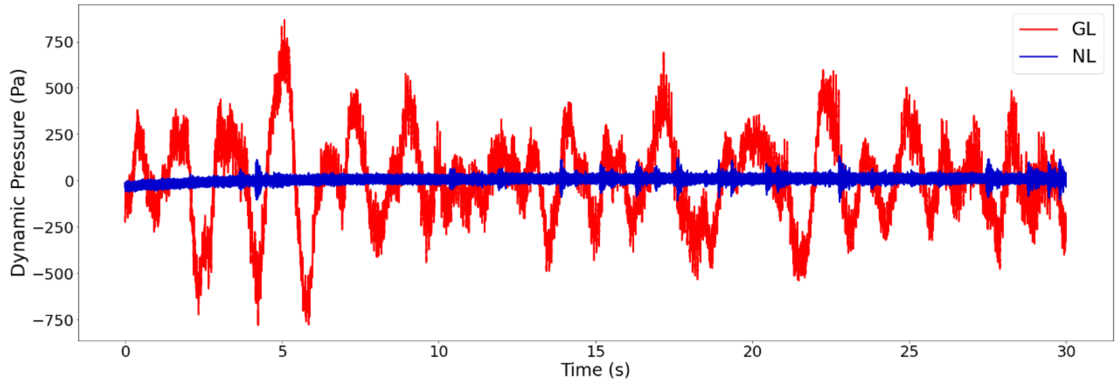

Figure 2.223. Time-domain plots of GL vs. NL signals in the looped network measured by sensor P2

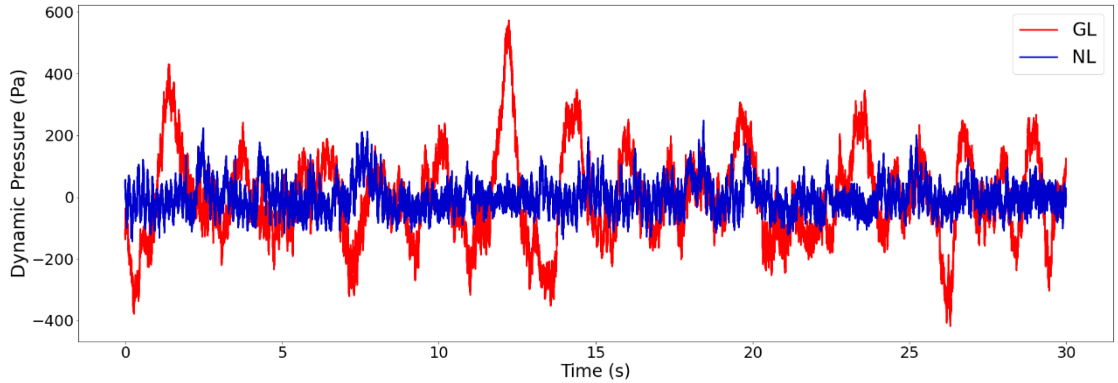

Figure 2.224. Time-domain plots of GL vs. NL signals in the branched network measured by sensor P2

Table 2.59 includes analytical information of the time-domain dynamic pressure sensor plots where leak and no leak signals of sensors P1 and P2 in the looped and branched networks are visually compared. These data are recorded under no demand scenario to highlight leak signals.

Table 2.59. Analysis of the time-domain dynamic pressure sensor plots measured by P1 and P2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Time-domain plot (for ND signal)	NL vs. OL – P1	- OL signal amplitudes are much larger than those pf NL in the looped network. - OL signal amplitudes are larger than those of NL in the branched network. - Branched network's NL signal amplitudes are larger than those in the looped network. - NL signal amplitudes in the looped network are close to zero; however, those amplitudes are large in the branched network.
	NL vs. LC – P1	- NL signal amplitudes are smaller in the looped network than those in the branched one. - LC signal amplitudes are smaller in the looped network than those in the branched one. - LC signal amplitudes have more outliers in the looped network than those in the branched one. - LC signal amplitudes are larger than those of NL in the looped network. - LC signal amplitudes are larger than those of NL in the branched network.
	NL vs. CC – P1	- CC signal amplitudes are smaller in the looped network than those in the branched one. - NL signal amplitudes are smaller in the looped network than those in the branched one. - CC signal amplitudes are much larger than those of NL in the looped network. - CC signal amplitudes are larger than those of NL in the branched network.
	NL vs. GL – P1	- Contrary to A1 and H1 measurements, NL signal amplitudes are smaller in the looped network than those in the branched one. - Similar to A1 and contrary to H1 measurements, GL signal amplitudes are smaller in the looped network than those in the branched one. - Contrary to A1 and similar to H1 measurements, NL signal amplitudes are smaller than those of GL in the looped network. - Contrary to A1 and H1 measurements, NL signal amplitudes are smaller than those of GL in the branched network.

Table 2.59. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Time-domain plot (for ND signal)	All – P1	- When there is no demand, based on P1 in both networks, the absolute values of all leak dynamic pressure data are larger than those of NL. - When there is no demand, based on P1 and except for OL, leaks' amplitudes in the looped network are smaller than those in the branched one. - When there is no demand and based on P1, NL signals' amplitudes in the looped network are smaller than those in the branched one.
	NL vs. OL – P2	- Visual comparison of OL signals did not result in a specific relation between OL signals' amplitude in the looped and branched network. - OL signal amplitudes are larger than those of the NL signal in the looped network. - OL signal amplitudes are larger than those of the NL signal in the branched network. - In both networks, OL signals are more variable and cyclic compared to NL signals. The more variable and cyclic pattern of the leak signal can help discern leaks using dynamic pressure sensors. - NL signal magnitudes are close to zero and more uniform in the looped network. However, NL signal magnitudes in the branched network are larger and more variable.
	NL vs. LC – P2	- Visual comparison of LC signals did not result in a specific relation between LC signals' amplitude in the looped and branched network. - LC signal amplitudes are larger than those of the NL signal in the looped network. - LC signal amplitudes are larger than those of the NL signal in the branched network. - In both networks, LC signals are more variable and cyclic compared to NL signals. - NL signal magnitudes are close to zero and more uniform in the looped network. However, NL signal magnitudes in the branched network are larger and more variable.

Table 2.59. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Time-domain plot (for ND signal)	NL vs. CC – P2	- CC signals of the looped network have larger absolute values than those of the branched network. - CC signal amplitudes are larger than those of the NL signal in the looped network. - CC signal amplitudes are larger than those of the NL signal in the branched network. - In both networks, CC signals are more variable and cyclic compared to NL signals. - NL signal magnitudes are close to zero and more uniform in the looped network. However, NL signal magnitudes in the branched network are larger and more variable.
	NL vs. GL – P2	- GL signals of the looped network have larger absolute values than those of the branched network. - GL signal amplitudes are larger than those of the NL signal in the looped network. - GL signal amplitudes are larger than those of the NL signal in the branched network. - In both networks, GL signals are more variable and cyclic compared to NL signals. - NL signal magnitudes are close to zero and more uniform in the looped network. However, NL signal magnitudes in the branched network are larger and more variable.
	All – P2	- When there is no demand and based on P2 in both networks, leak signals are more variable and cyclic compared to NL signals. The more variable and cyclic pattern of the leak signals can help discern leaks using dynamic pressure sensors. - When there is no demand, based on P2 in both networks, the absolute values of all leak dynamic pressure data are larger than those of NL. - When there is no demand and based on P2 in both networks, NL signal magnitudes are close to zero and more uniform in the looped network. However, NL signal magnitudes in the branched network are larger and more variable. - When there is no demand, and in both networks, P1 leak and NL signals' amplitudes are larger than those of P2. - When there is no demand in both networks, leak signals measured by P2 are more variable and cyclic than those of P1.

2.12.5.2. Frequency-domain plot (for ND signals)

Figures 2.225 to 2.232 show the frequency-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor P1.

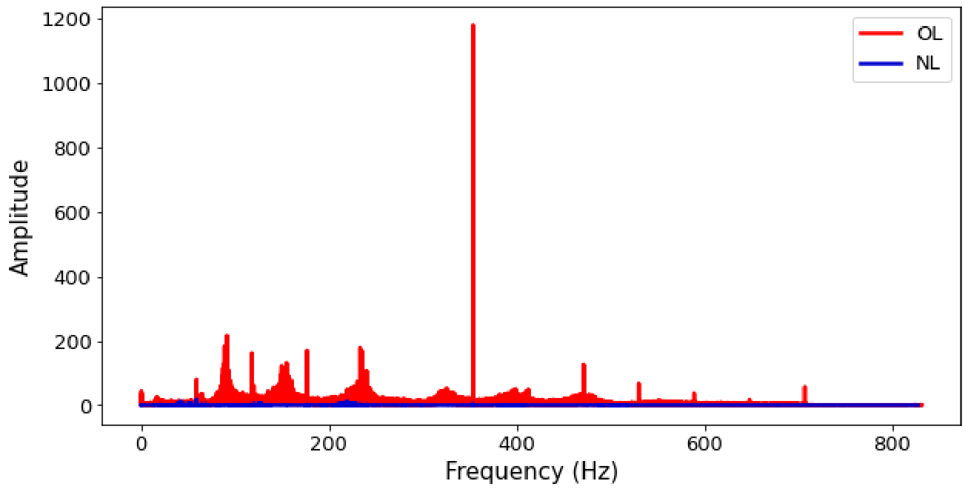

Figure 2.225. Frequency-domain plots of OL vs. NL signals in the looped network measured by sensor P1

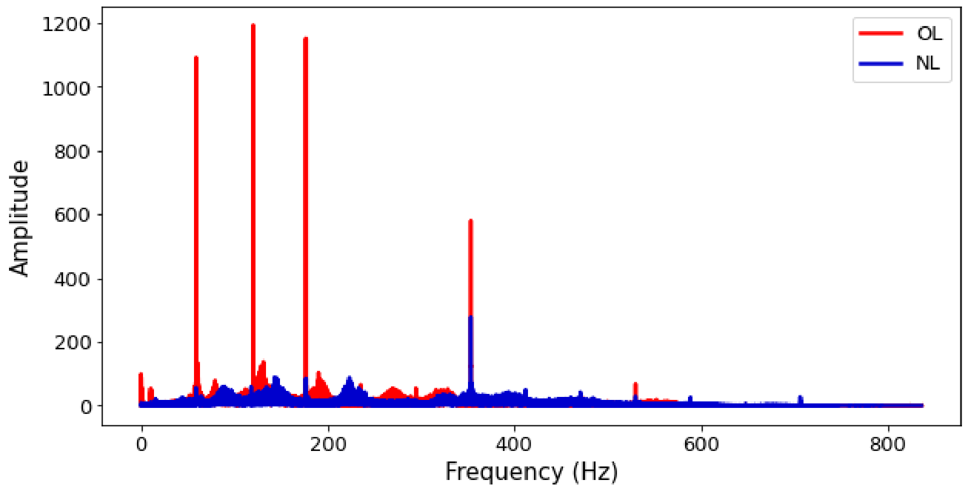

Figure 2.226. Frequency-domain plots of OL vs. NL signals in the branched network measured by sensor P1

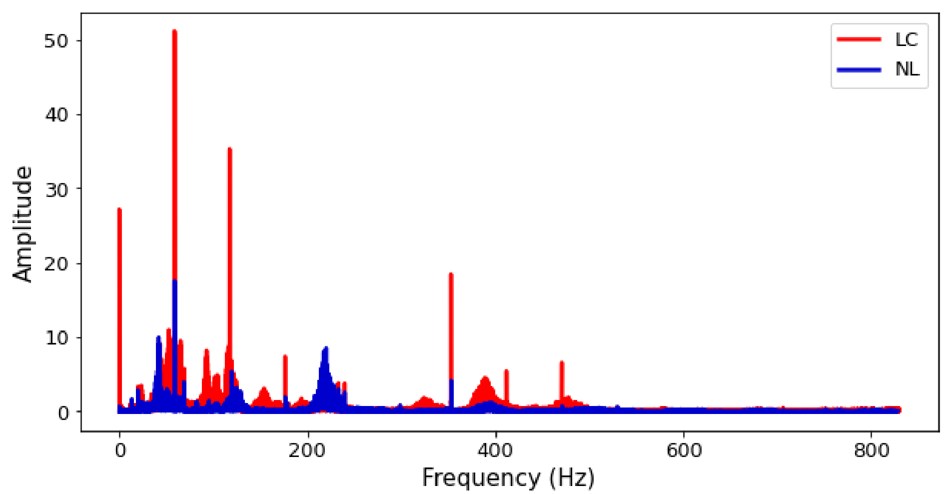

Figure 2.227. Frequency-domain plots of LC vs. NL signals in the looped network measured by sensor P1

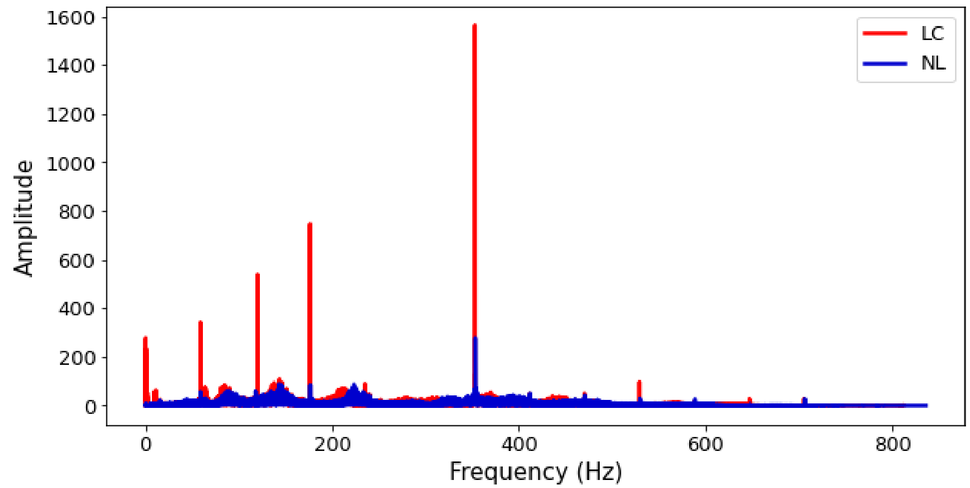

Figure 2.228. Frequency-domain plots of LC vs. NL signals in the branched network measured by sensor P1

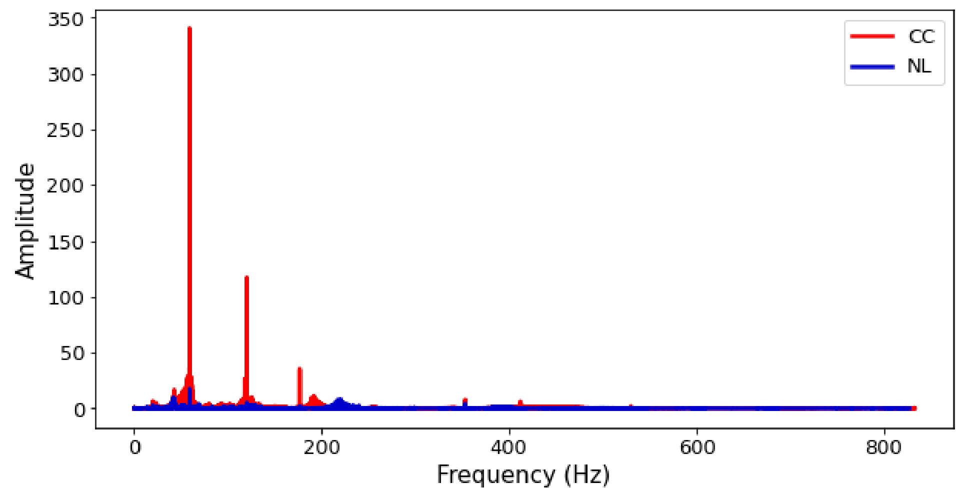

Figure 2.229. Frequency-domain plots of CC vs. NL signals in the looped network measured by sensor P1

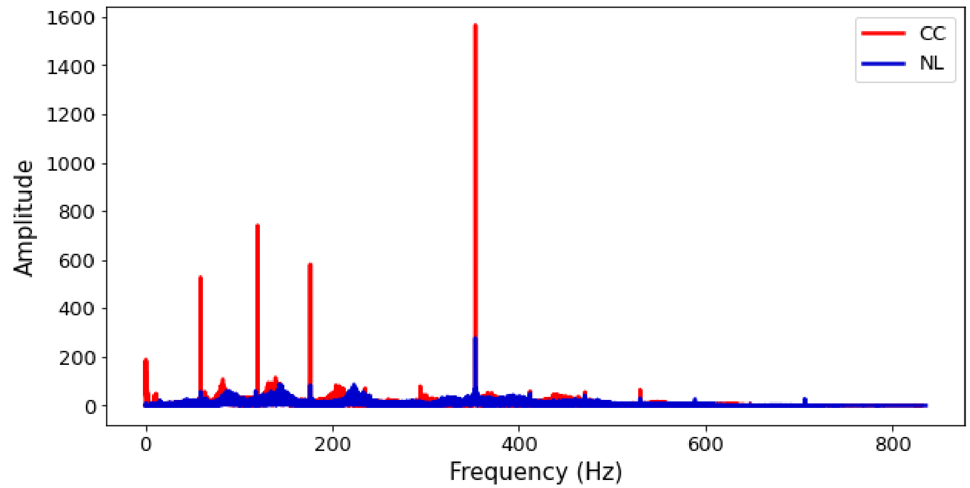

Figure 2.230. Frequency-domain plots of CC vs. NL signals in the branched network measured by sensor P1

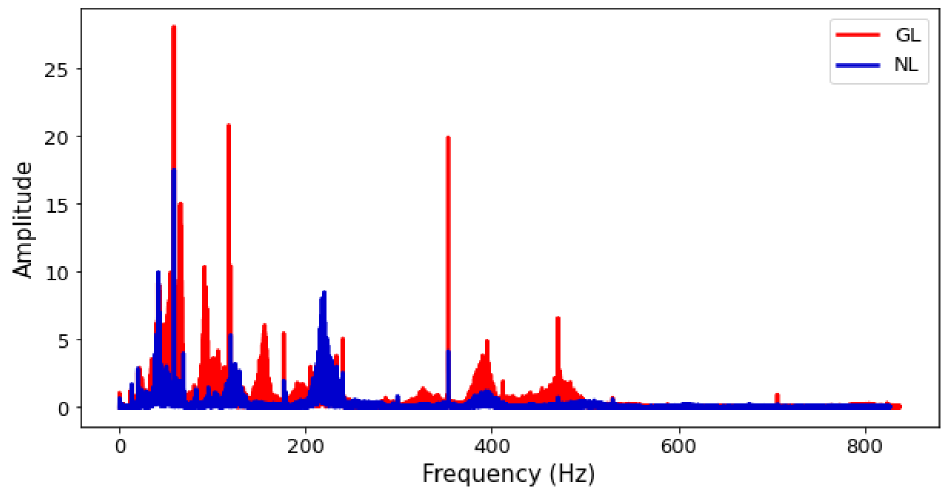

Figure 2.231. Frequency-domain plots of GL vs. NL signals in the looped network measured by sensor P1

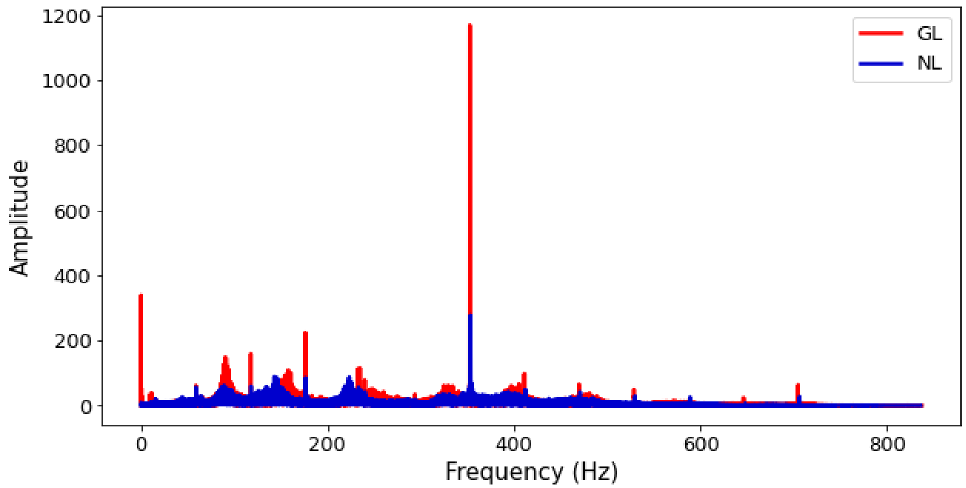

Figure 2.232. Frequency-domain plots of GL vs. NL signals in the branched network measured by sensor P1

Figures 2.233 to 2.240 show the Frequency-domain plots of the OL, LC, CC, and GL signals versus NL signal in the looped and branched networks, all measured by sensor P2.

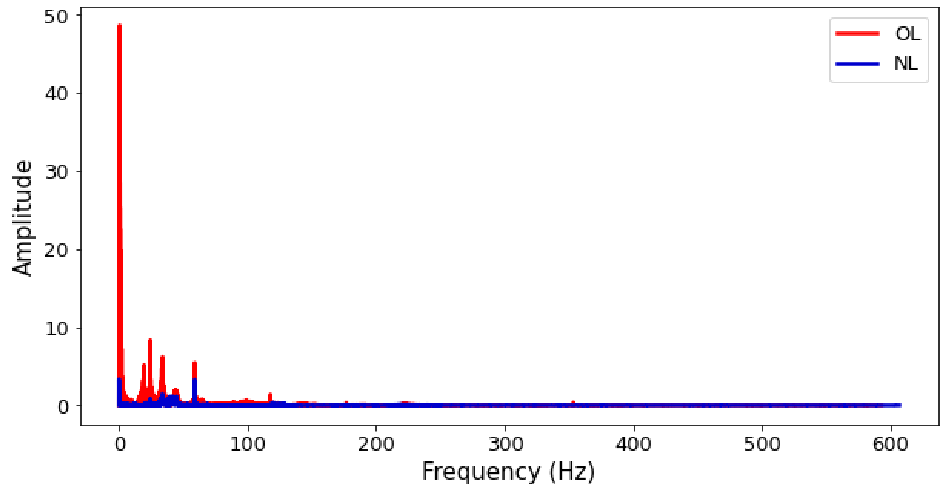

Figure 2.233. Frequency-domain plots of OL vs. NL signals in the looped network measured by sensor P2

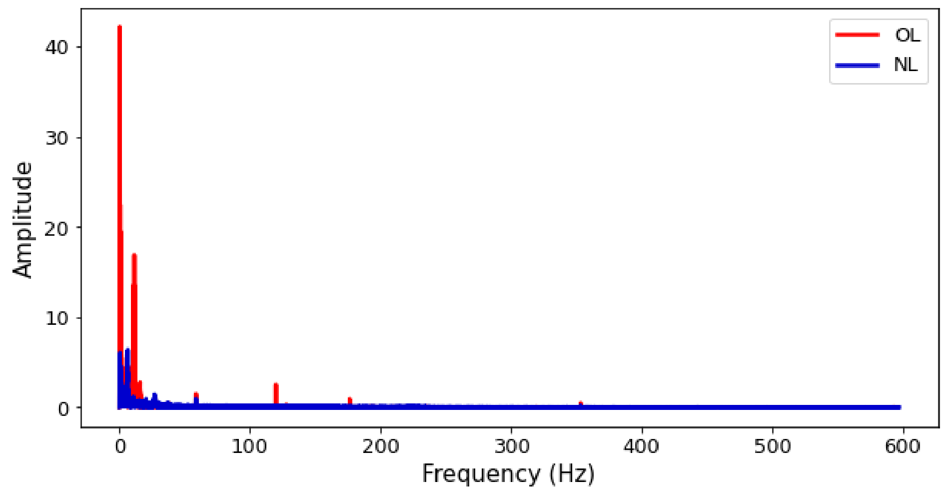

Figure 2.234. Frequency-domain plots of OL vs. NL signals in the branched network measured by sensor P2

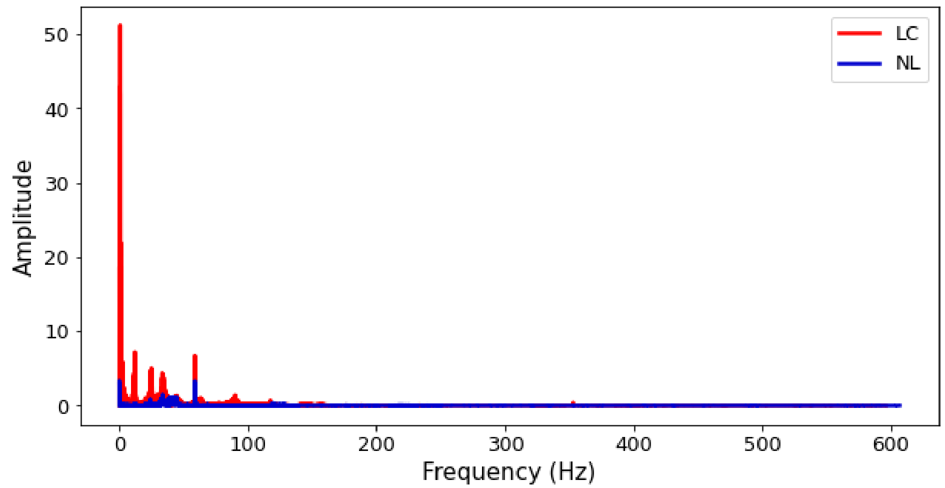

Figure 2.235. Frequency-domain plots of LC vs. NL signals in the looped network measured by sensor P2

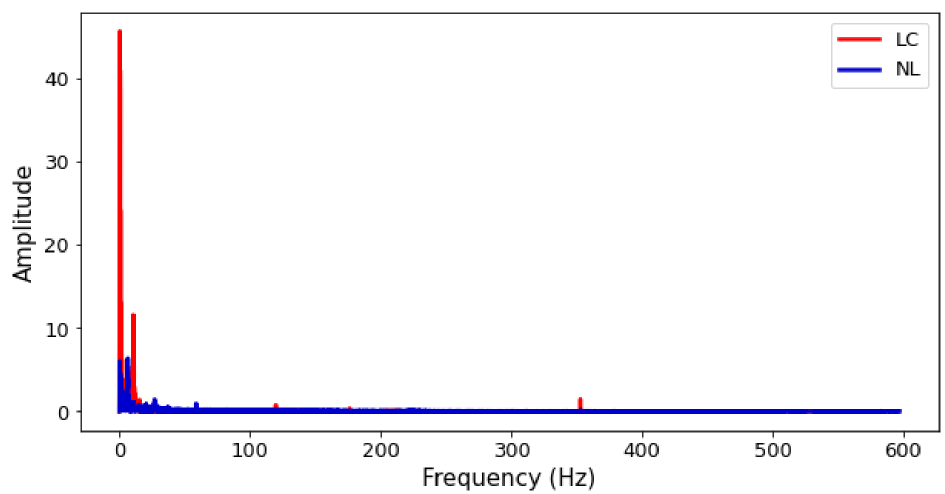

Figure 2.236. Frequency-domain plots of LC vs. NL signals in the branched network measured by sensor P2

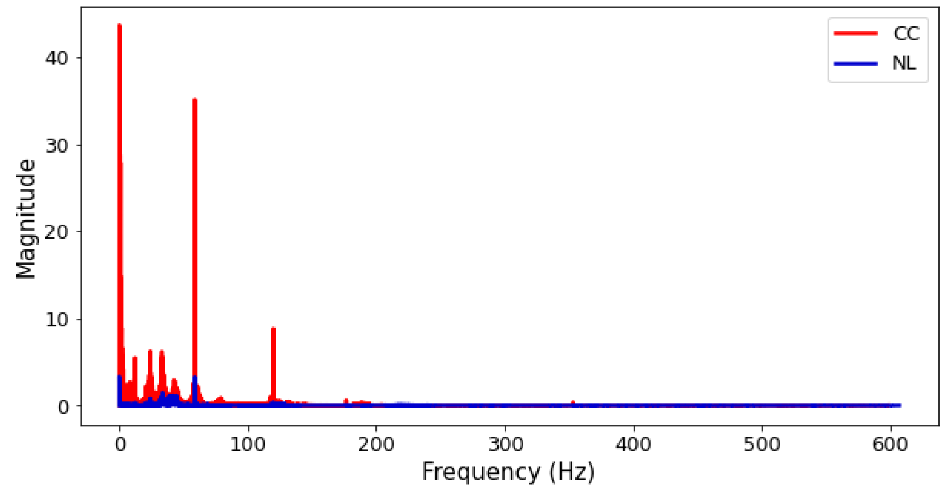

Figure 2.237. Frequency-domain plots of CC vs. NL signals in the looped network measured by sensor P2

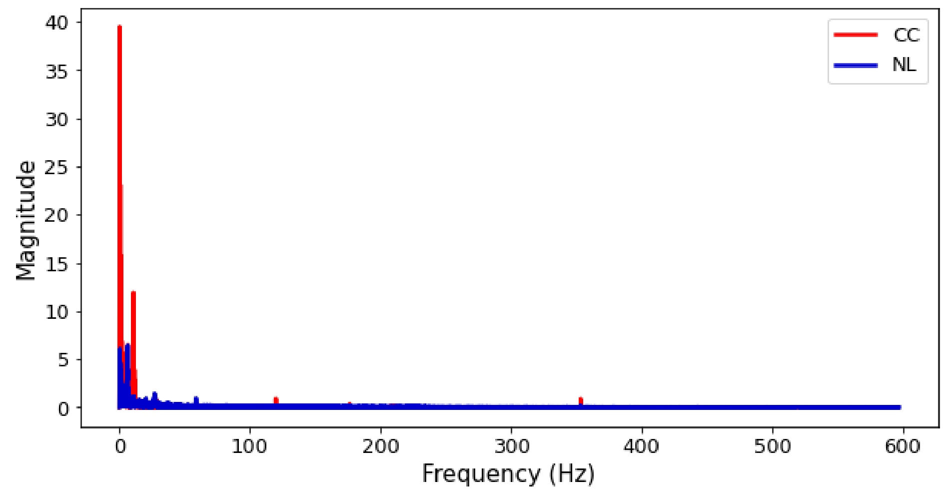

Figure 2.238. Frequency-domain plots of CC vs. NL signals in the branched network measured by sensor P2

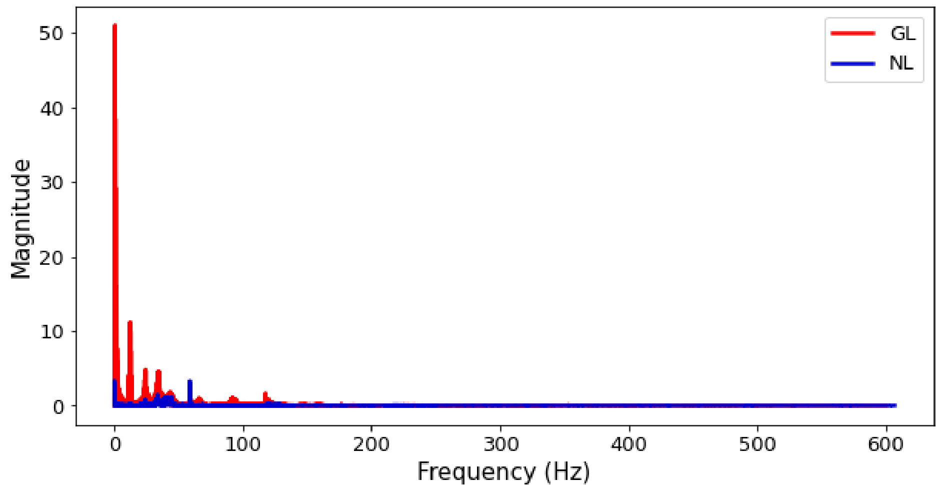

Figure 2.239. Frequency-domain plots of GL vs. NL signals in the looped network measured by sensor P2

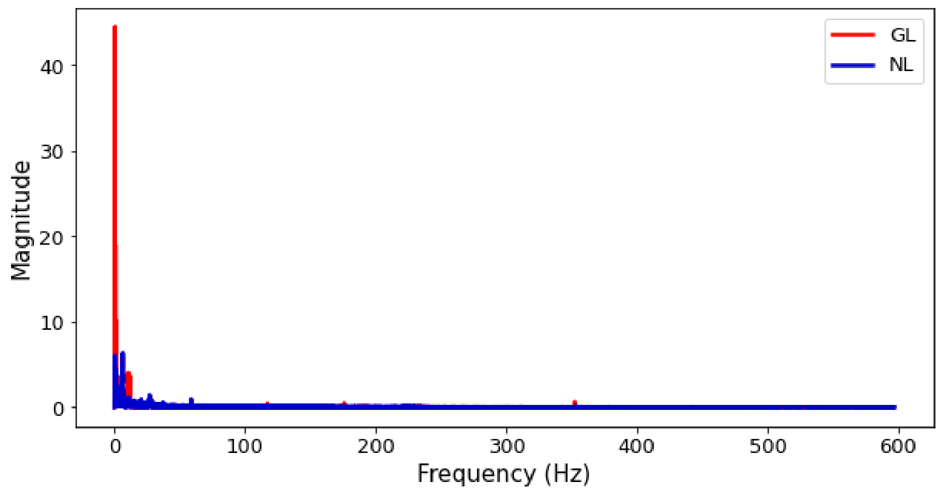

Figure 2.240. Frequency-domain plots of GL vs. NL signals in the branched network measured by sensor P2

Table 2.60 includes analytical information of the dynamic pressure data's frequency-domain plots where leak and no leak signals of sensors P1 and P2 in the looped and branched networks are visually compared.

Table 2.60. Analysis of the frequency-domain dynamic pressure data plots measured by P1 and P2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Frequency-domain plot (for ND signal)	NL vs. OL – P1	- For NL signal in the looped network, frequencies with non-zero amplitudes are less than 300 Hz. - For NL signal in the branched network, frequencies with non-zero amplitudes are less than 800 Hz. - NL signal's dominant frequency and its amplitude are smaller in the looped network than in the branched one. - Compared to the branched network, the amplitudes of NL signal's frequencies are negligible in the looped network. - OL signal's dominant frequency in the looped network is larger than that of the branched network. - OL signal's maximum frequency in the looped network is below 800 Hz. - OL signal's maximum frequency in the branched network is below 600 Hz. - OL signal's frequency amplitudes are more pronounced in the branched network than the looped one. - In the looped network, OL signal's dominant frequency is larger than that of NL signal. - In the branched network, OL signal's dominant frequency is smaller than that of NL signal. - In both networks, the amplitudes of OL signal's frequencies are more pronounced than those of NL signal. - In both networks, the amplitudes of OL signal's dominant frequencies are larger than those of NL signal. Therefore, the amplitude of dominant frequency can help detect OL leak in both networks with ND.

Table 2.60. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Frequency-domain plot (for ND signal)	NL vs. LC – P1	- LC signal’s dominant frequency and its amplitude are larger in the branched network than in the looped network. This difference may help identify the network change. - NL signal’s dominant frequency and its amplitude are larger in the branched network than in the looped network. - LC signal’s frequency cap in the looped network is below 600 Hz. - LC signal’s frequency cap in the branched network is below 800 Hz. - NL signal’s frequency cap in the looped network is below 600 Hz. - NL signal’s frequency cap in the branched network is below 800 Hz. - In both networks, dominant frequencies of LC and NL are the same. Therefore, dominant frequency cannot be a leak detection distinguishing feature. - In both networks, frequencies of the LC signals have larger amplitudes than the NL signals. - In both networks, the amplitudes of LC signal’s dominant frequencies are larger than those of NL signal. Therefore, the amplitude of dominant frequency can help detect LC leak in both networks with ND.
	NL vs. CC – P1	- CC signal’s dominant frequency and its amplitude are larger in the branched network than in the looped network. This difference may help identify the network change. - NL signal’s dominant frequency and its amplitude are larger in the branched network than in the looped network. - CC signal’s frequency cap in the looped network is below 600 Hz. - CC signal’s frequency cap in the branched network is below 800 Hz. - NL signal’s frequency cap in the looped network is below 600 Hz. - NL signal’s frequency cap in the branched network is below 800 Hz. - In both networks, dominant frequencies of CC and NL are the same. Therefore, dominant frequency cannot be a leak detection distinguishing feature. - In both networks, frequencies of the CC signals have larger amplitudes than the NL signals.

Table 2.60. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Frequency-domain plot (for ND signal)	NL vs. CC – P1	- In both networks, the amplitudes of CC signal’s dominant frequencies are larger than those of NL signal. Therefore, the amplitude of dominant frequency can help detect CC leak in both networks with ND.
	NL vs. GL – P1	- GL signal’s dominant frequency and its amplitude are larger in the branched network than in the looped network. This difference may help identify the network change. - NL signal’s dominant frequency and its amplitude are larger in the branched network than in the looped network. - GL signal’s frequency cap in the looped network is below 800 Hz which is larger than other leaks’ frequency caps in the looped network. - GL signal’s frequency cap in the branched network is below 800 Hz. - NL signal’s frequency cap in the looped network is below 600 Hz. - NL signal’s frequency cap in the branched network is below 800 Hz. - In both networks, dominant frequencies of GL and NL are the same. Therefore, dominant frequency cannot be a leak detection distinguishing feature. - In both networks, the amplitudes of GL signal’s dominant frequencies are larger than those of NL signal. Therefore, the amplitude of dominant frequency can help detect GL leak in both networks with ND.
	All – P1	- With no demand and based on P1, except for OL, other leaks’ dominant frequencies and their amplitudes are larger in the branched network than those of the looped network. - With no demand and based on P1, the NL signal’s dominant frequency and amplitude are larger in the branched network than in the looped network. - With no demand and based on P1, NL signal’s frequency cap in the looped network is below 600 Hz.

Table 2.60. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Frequency-domain plot (for ND signal)	All – P1	- With no demand and based on P1, NL signal's frequency cap in the branched network is below 800 Hz. - With no demand and based on P1, except for GL, other leak signals' non-zero frequencies in the looped network are below 600 Hz. - With no demand and based on P1, all leak signals' non-zero frequencies in the branched network are below 800 Hz. - In both networks, the amplitudes of all leak signals' dominant frequencies are larger than those of NL signal. Therefore, the amplitudes of dominant frequencies can help detect leaks in both networks with ND.
	NL vs. OL – P2	- OL signal's dominant frequency and its amplitude are approximately the same in both networks. Therefore, OL signal's dominant frequency and its amplitude cannot help identify network change. - NL signal's dominant frequency is larger in the looped network than in the branched network. - NL signal's dominant frequency amplitude is slightly larger in the branched network than in the looped network. - OL signal's frequency cap in the looped network is below 400 Hz. - OL signal's frequency cap in the branched network is below 400 Hz. - NL signal's frequency cap in the looped network is below 100 Hz. - NL signal's frequency cap in the branched network is below 100 Hz. - In both networks, dominant frequency of NL signal is larger than that of OL. Therefore, dominant frequency may be used for leak detection in both networks. - In both networks, frequencies of the OL signals have larger amplitudes than the NL signals. - In both networks, the amplitudes of OL signal's dominant frequencies are larger than those of NL signal. Therefore, the amplitude of dominant frequency can help detect OL leak in both networks with ND.

Table 2.60. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Frequency-domain plot (for ND signal)	NL vs. LC – P2	- LC signal’s dominant frequency and its amplitude are approximately the same in both networks. Therefore, LC signal’s dominant frequency and its amplitude cannot help identify network change. - NL signal’s dominant frequency is larger in the looped network than in the branched network. - NL signal’s dominant frequency amplitude is slightly larger in the branched network than in the looped network. - LC signal’s frequency cap in the looped network is below 400 Hz. - LC signal’s frequency cap in the branched network is below 400 Hz. - NL signal’s frequency cap in the looped network is below 100 Hz. - NL signal’s frequency cap in the branched network is below 100 Hz. - In both networks, dominant frequency of NL signal is larger than that of LC. Therefore, dominant frequency may be used for leak detection in both networks. - In both networks, frequencies of the LC signals have larger amplitudes than the NL signals. - In both networks, the amplitudes of LC signal’s dominant frequencies are larger than those of NL signal. Therefore, the amplitude of dominant frequency can help detect LC leak in both networks with ND.
	NL vs. CC – P2	- CC signal’s dominant frequency and its amplitude are approximately the same in both networks. Therefore, CC signal’s dominant frequency and its amplitude cannot help identify network change. - NL signal’s dominant frequency is larger in the looped network than in the branched network. - NL signal’s dominant frequency amplitude is slightly larger in the branched network than in the looped network. - CC signal’s frequency cap in the looped network is below 400 Hz. - CC signal’s frequency cap in the branched network is below 400 Hz. - NL signal’s frequency cap in the looped network is below 100 Hz. - NL signal’s frequency cap in the branched network is below 100 Hz. - In both networks, dominant frequency of NL signal is larger than that of CC. Therefore, dominant frequency may be used for leak detection in both networks.

Table 2.60. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Frequency-domain plot (for ND signal)	NL vs. CC – P2	- In both networks, frequencies of the CC signals have larger amplitudes than the NL signals. - In both networks, the amplitudes of CC signal’s dominant frequencies are larger than those of NL signal. Therefore, the amplitude of dominant frequency can help detect CC leak in both networks with ND.
	NL vs. GL – P2	- GL signal’s dominant frequency and its amplitude are approximately the same in both networks. Therefore, GL signal’s dominant frequency and its amplitude cannot help identify network change. - NL signal’s dominant frequency is larger in the looped network than in the branched network. - NL signal’s dominant frequency amplitude is slightly larger in the branched network than in the looped network. - GL signal’s frequency cap in the looped network is below 200 Hz. - GL signal’s frequency cap in the branched network is below 400 Hz. - NL signal’s frequency cap in the looped network is below 100 Hz. - NL signal’s frequency cap in the branched network is below 100 Hz. - In both networks, dominant frequency of NL signal is larger than that of GL. Therefore, dominant frequency may be used for leak detection in both networks. - In both networks, frequencies of the GL signals have larger amplitudes than the NL signals. - In both networks, the amplitudes of GL signal’s dominant frequencies are larger than those of NL signal. Therefore, the amplitude of dominant frequency can help detect GL leak in both networks with ND.
	All – P2	- With no demand and based on P2, since all leaks’ dominant frequencies and their amplitudes are approximately the same in both networks, leaks’ dominant frequencies and their amplitudes cannot identify the network change. - With no demand and based on P2, NL signal’s dominant frequency is larger in the looped network than in the branched network. - With no demand and based on P2, NL signal’s dominant frequency amplitude is slightly larger in the branched network than in the looped network. - With no demand and based on P2, NL signal’s frequency cap in both networks is below 100 Hz.

Table 2.60. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Frequency-domain plot (for ND signal)	All – P2	- With no demand and based on P2, all leak signals' non-zero frequencies in both networks are below 400 Hz. - In both networks, the amplitudes of all leak signals' dominant frequencies are larger than those of NL signal. Therefore, the amplitudes of dominant frequencies can help detect leak in both networks with ND. - In both networks, dominant frequencies of NL signals are larger than those of leak signals. Therefore, dominant frequency may be used for leak detection in both networks.

2.12.5.3. Cumulative Distribution Plot (for ND signal)

Figures 2.241 and 2.242 show the cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped and branched networks, respectively, measured by sensor P1 with no demand. Also, Figures 2.243 and 2.244 show the cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped and branched networks, respectively, measured by sensor P2 with no demand.

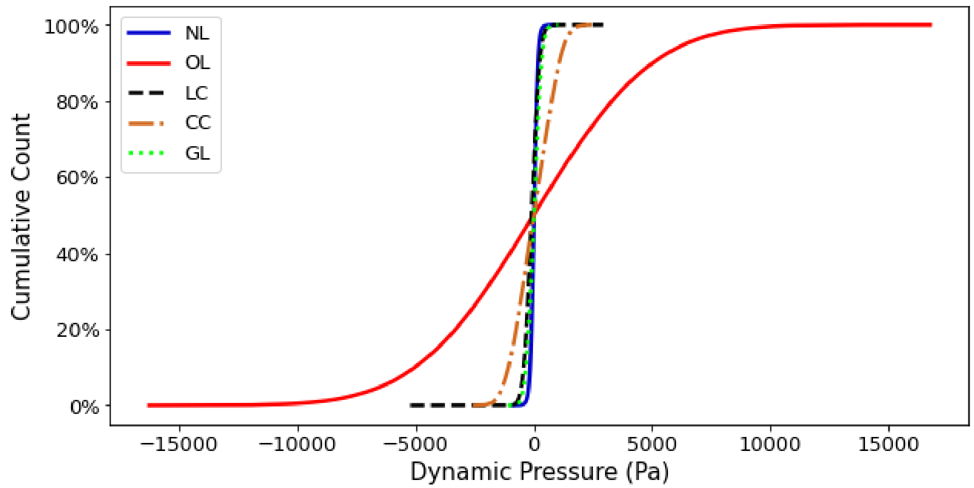

Figure 2.241. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor P1

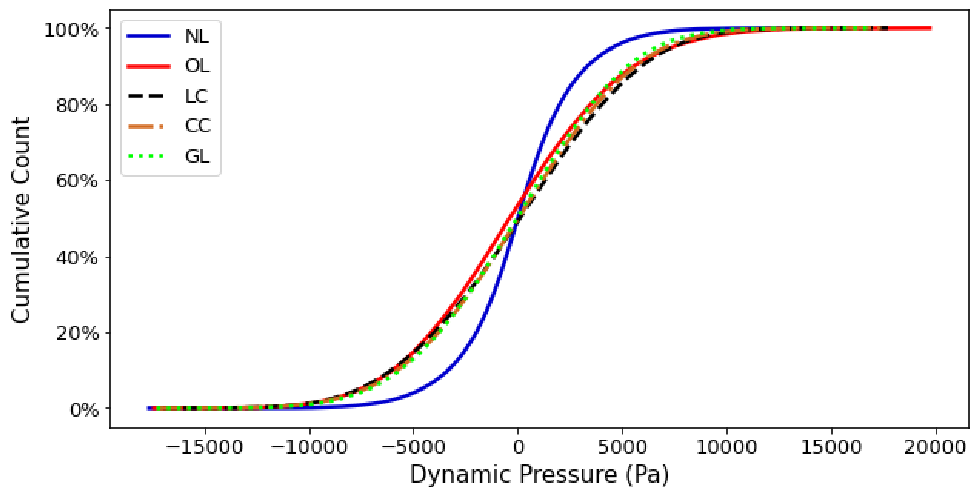

Figure 2.242. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the branched network, measured by sensor P1

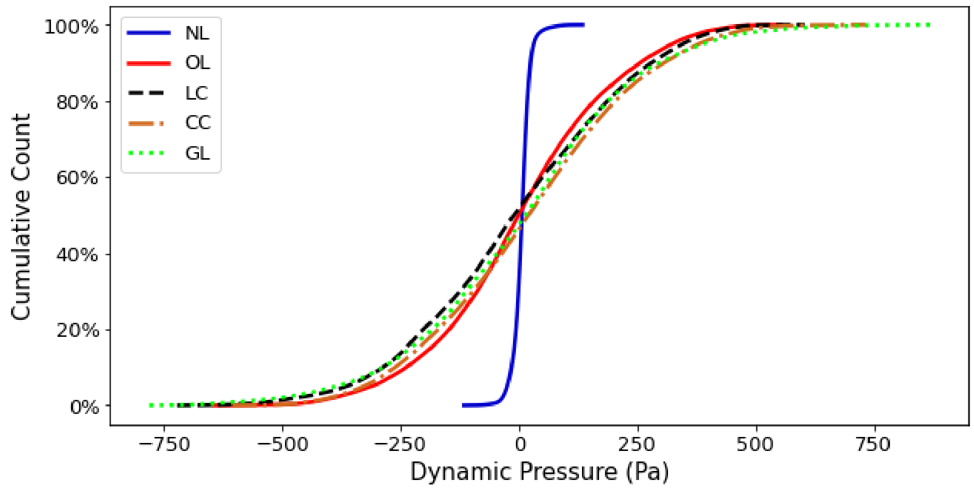

Figure 2.243. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor P2

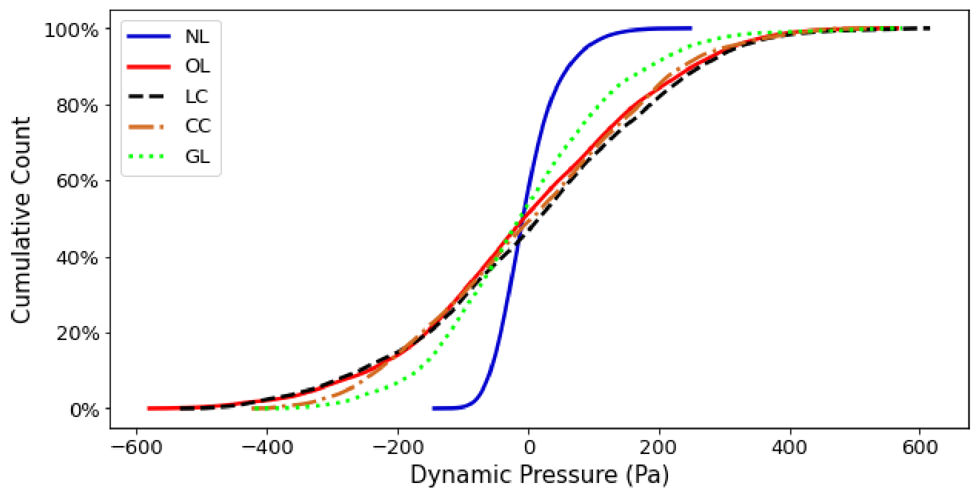

Figure 2.244. Cumulative distribution plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor P2

Table 2.61 includes the analysis of the cumulative distribution plots where leak and no leak signals of sensors P1 and P2 in the looped and branched networks are visually compared.

Table 2.61. Analysis of dynamic pressure data's cumulative distribution plots measured by P1 and P2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Cumulative distribution plot (for ND signal)	NL – P1	- NL dynamic pressure data magnitudes span a larger range in the branched network than in the looped network. - NL dynamic pressure data magnitudes are larger in the branched network than in the looped network. - NL dynamic pressure data magnitudes in the looped network are smallest compared to all leak and NL signals of both networks. - Order of signal magnitude in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude in the branched network: LC > OL > CC > GL > NL.
	OL – P1	- OL dynamic pressure data include the largest magnitudes and span the widest range of magnitudes compared to other leak and NL signals of the looped network. - Order of signal magnitude in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude in the branched network: LC > OL > CC > GL > NL.
	LC – P1	- LC dynamic pressure data magnitudes are larger in the branched network than in the looped network. - LC, and GL cumulative distribution plots approximately overlap in both networks. - Order of signal magnitude in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude in the branched network: LC > OL > CC > GL > NL.
	CC – P1	- CC dynamic pressure data magnitudes are larger in the branched network than in the looped network. - CC and LC cumulative distribution plots approximately overlap in the branched network. - Order of signal magnitude in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude in the branched network: LC > OL > CC > GL > NL.

Table 2.61. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Cumulative distribution plot (for ND signal)	GL – P1	- GL dynamic pressure data magnitudes are larger in the branched network than in the looped network. - GL, and LC cumulative distribution plots approximately overlap in both networks. - Order of signal magnitude in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude in the branched network: LC > OL > CC > GL > NL.
	All – P1	- OL dynamic pressure data include the largest magnitudes and span the widest range of magnitudes compared to other leak and NL signals of the looped network. - Except for OL, NL, and other leaks' cumulative distribution plots of the looped network are smaller than those of the branched network. - Except for OL, NL, and other leaks' cumulative distribution plots of the looped network overlap. - All leaks' cumulative distribution plots of the branched network approximately overlapped. - Order of signal magnitude in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude in the branched network: LC > OL > CC > GL > NL.
	NL – P2	- NL dynamic pressure data include the smallest magnitudes and span the narrowest range of magnitudes compared to other leak signals of both networks. - Order of signal magnitude in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude in the branched network: LC > CC > OL > GL > NL.
	OL – P2	- Contrary to P1, OL includes the smallest and span the narrowest range of dynamic pressure magnitudes compared to other leaks in the looped network. - OL signal magnitudes of the looped network are larger than those of the branched network. - Order of signal magnitude in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude in the branched network: LC > CC > OL > GL > NL.

Table 2.61. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Cumulative distribution plot (for ND signal)	LC – P2	- LC and GL cumulative distribution plots approximately overlapped in the looped network. - LC, and OL cumulative distribution plots approximately overlapped in the branched network. - LC dynamic pressure data magnitudes are larger in the looped network than the branched one. - Order of signal magnitude in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude in the branched network: LC > CC > OL > GL > NL.
	CC – P2	- CC dynamic pressure data magnitudes are larger in the looped network than the branched one. - Order of signal magnitude in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude in the branched network: LC > CC > OL > GL > NL.
	GL – P2	- GL and LC cumulative distribution plots approximately overlapped in the looped network. - GL dynamic pressure data magnitudes are larger in the looped network than the branched one. - GL dynamic pressure data magnitudes are the smallest compared to other leak signals in the branched network. - Order of signal magnitude in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude in the branched network: LC > CC > OL > GL > NL.
	All – P2	- Leak dynamic pressure data magnitudes are larger in the looped network compared to those of the branched network. - In both networks, NL dynamic pressure data include the smallest and span the narrowest range of magnitudes. - In the looped network, leak cumulative distribution plots are very similar. - GL dynamic pressure data magnitudes are the smallest compared to other leak signals in the branched network. - Order of signal magnitude in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude in the branched network: LC > CC > OL > GL > NL.

2.12.5.4. Box Plot (for ND signal)

Figures 2.245 and 2.246 show box plots of the NL, OL, LC, CC, and GL signals in the looped and branched networks, respectively, measured by sensor P1 with no demand. Figures 2.247 and 2.248 show the same plots for signals measured by sensor P2 with no demand.

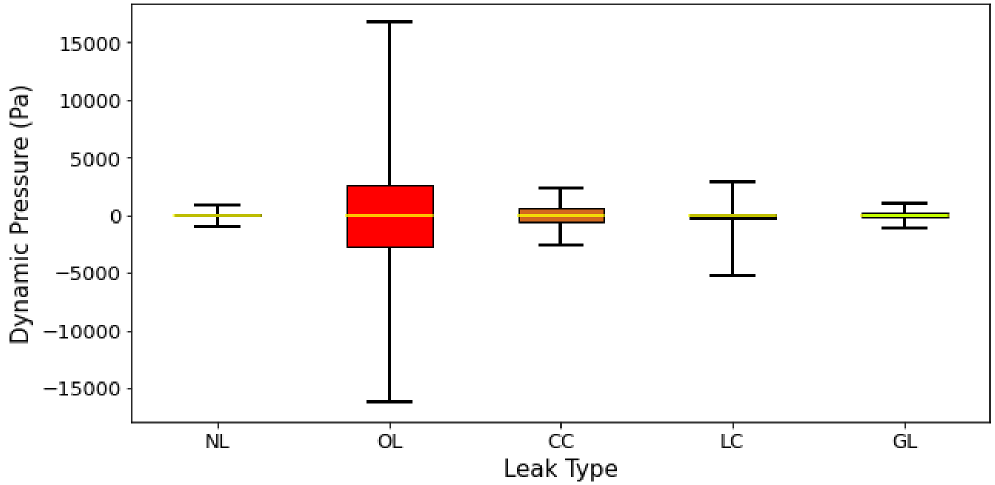

Figure 2.245. Box plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor P1

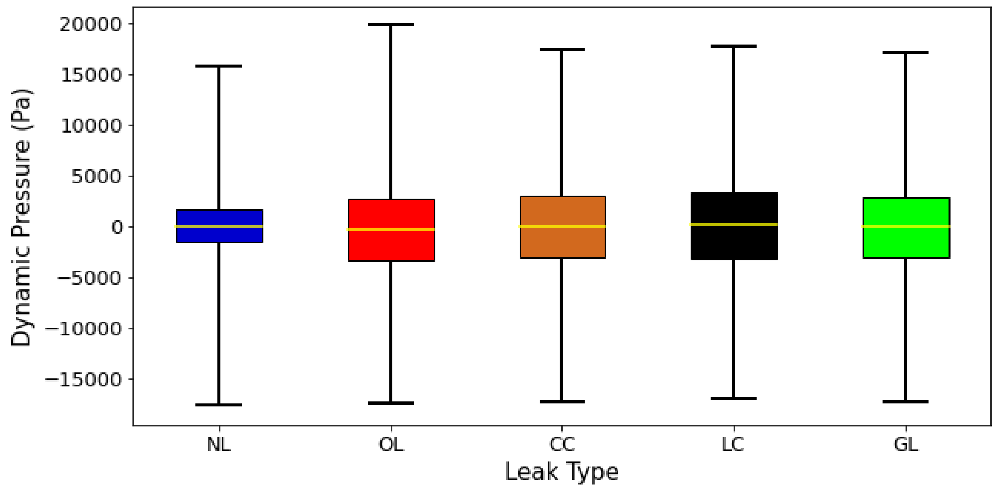

Figure 2.246. Box plots of the NL, OL, LC, CC, and GL signals in the branched network, measured by sensor P1

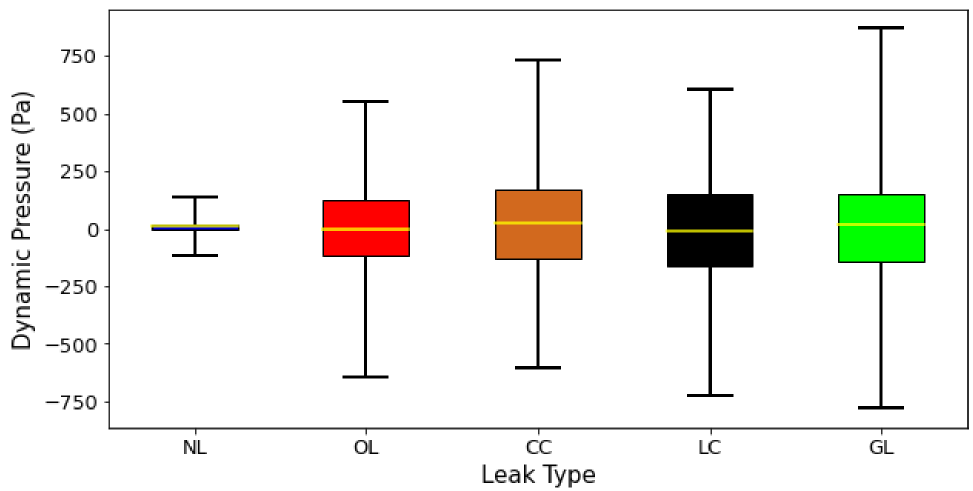

Figure 2.247. Box plots of the NL, OL, LC, CC, and GL signals in the looped network, measured by sensor P2

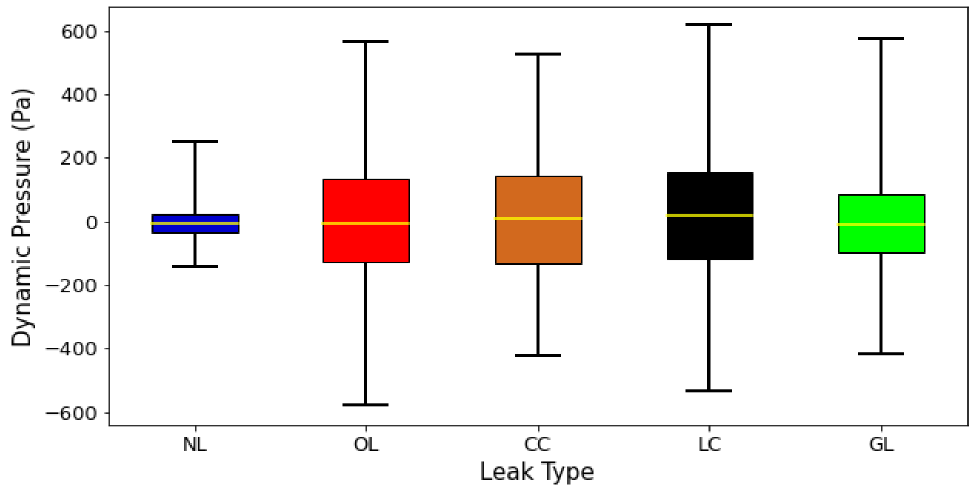

Figure 2.248. Box plots of the NL, OL, LC, CC, and GL signals in the branched network, measured by sensor P2

Table 2.62 includes the analysis of the box plots where leak and no leak signals of sensors P1 and P2 in the looped and branched networks are visually compared.

Table 2.62. Analysis of dynamic pressure data's box plots measured by P1 and P2

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Box plot (for ND signal)	NL – P1	- Comparison of NL signal magnitude continuum: $NL_{lo} < NL_{br}$ - Order of signal magnitude continuum in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude continuum in the branched network: LC > OL > CC > GL > NL.
	OL – P1	- Comparison of OL signal magnitude continuum: $OL_{lo} < OL_{br}$ - Order of signal magnitude continuum in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude continuum in the branched network: LC > OL > CC > GL > NL.

Table 2.62. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Box plot (for ND signal)	LC – P1	- LC signal of the branched network includes the largest range of dynamic pressure data magnitudes compared to all other signals of the looped and the branched networks. - Comparison of LC signal magnitude continuum: $LC_{lo} < LC_{br}$ - Order of signal magnitude continuum in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude continuum in the branched network: LC > OL > CC > GL > NL.
	CC – P1	- Comparison of CC signal magnitude continuum: $CC_{lo} < CC_{br}$ - Order of signal magnitude continuum in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude continuum in the branched network: LC > OL > CC > GL > NL.
	GL – P1	- Comparison of GL signal magnitude continuum: $GL_{lo} < GL_{br}$ - Order of signal magnitude continuum in the looped network: OL > CC > LC > GL > NL. - Order of signal magnitude continuum in the branched network: LC > OL > CC > GL > NL.
	All – P1	- LC signal of the branched network includes the largest range of dynamic pressure data magnitudes compared to all other signals of the looped and the branched networks. - The magnitude continuum of NL and leak signals of the branched network are larger than those in the looped network. Therefore, based on P1, the difference between the 1st and the 3rd quartiles of dynamic pressure magnitudes can identify network change. - Order of signal magnitude continuum in the looped network: OL > CC > LC > GL > NL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot. - Order of signal magnitude continuum in the branched network: LC > OL > CC > GL > NL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot.
	NL – P2	- Comparison of NL signal magnitude continuum: $NL_{lo} < NL_{br}$ - Order of signal magnitude continuum in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude continuum in the branched network: LC > CC > OL > GL > NL.

Table 2.62. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Box plot (for ND signal)	OL – P2	- Comparison of NL signal magnitude continuum: $OL_{lo} < OL_{br}$ - Order of signal magnitude continuum in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude continuum in the branched network: LC > CC > OL > GL > NL.
	LC – P2	- LC signal of the looped network includes the largest range of dynamic pressure data magnitudes compared to all other signals of the looped and the branched networks. - Comparison of LC signal magnitude continuum: $LC_{lo} > LC_{br}$ - Order of signal magnitude continuum in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude continuum in the branched network: LC > CC > OL > GL > NL.
	CC – P2	- Comparison of CC signal magnitude continuum: $CC_{lo} > CC_{br}$ - Order of signal magnitude continuum in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude continuum in the branched network: LC > CC > OL > GL > NL.
	GL – P2	- Comparison of GL signal magnitude continuum: $GL_{lo} > GL_{br}$ - Order of signal magnitude continuum in the looped network: LC > CC > GL > OL > NL. - Order of signal magnitude continuum in the branched network: LC > CC > OL > GL > NL.
	All – P2	- LC signal of the looped network includes the largest range of dynamic pressure data magnitudes compared to all other signals of the looped and the branched networks. Based on P1, the LC signal of the branched network had such a condition. - Contrary to P1 and based on P2, there is no specific relation between dynamic pressure magnitudes of the looped network and those of the branched network. - Order of signal magnitude continuum in the looped network: LC > CC > GL > OL > NL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot. - Order of signal magnitude continuum in the branched network: LC > CC > OL > GL > NL. This order conforms to the order of signals' magnitude continuum in the cumulative distribution plot.

2.12.5.5. Cross Spectral Plot (for ND signal)

Figures 2.249 and 2.250 show the cross spectral plots of the NL, OL, LC, CC, and GL signals with no demands in the looped and branched networks, respectively.

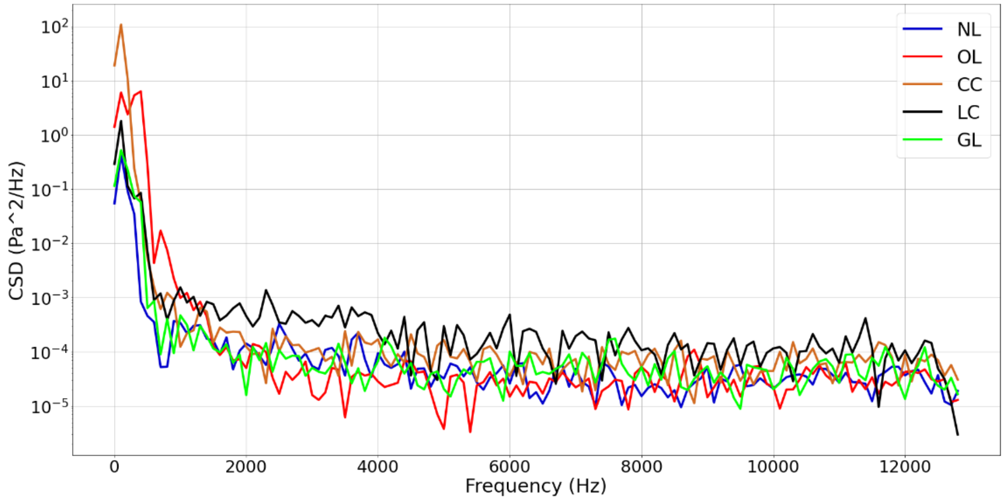

Figure 2.249. Cross spectral plots of the NL, OL, LC, CC, and GL signals in the looped network for dynamic pressure sensors

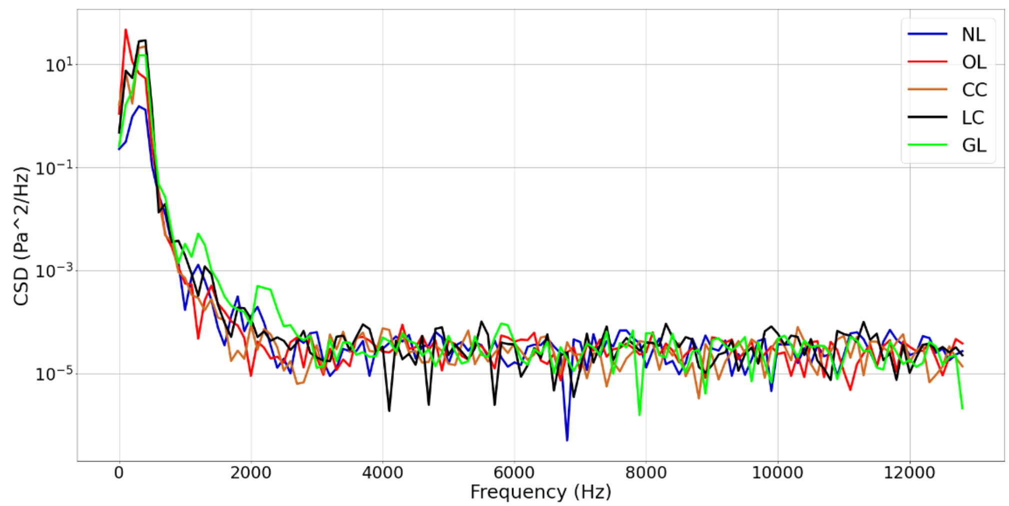

Figure 2.250. Cross spectral plots of the NL, OL, LC, CC, and GL signals in the branched network for dynamic pressure sensors

Table 2.63 compares the area under the cross spectral plots of the looped and branched networks with no demand.

Table 2.63. Comparison of the area under the cross spectral plots of dynamic pressure data

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Area under the cross-spectral plot (for ND signal)	NL	- Comparison of NL signal's areas under the CSD plot: $NL_{lo} < NL_{br}$ - Order of areas under the CSD plot in the looped network: CC > OL > LC > GL > NL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.
	OL	- Comparison of OL signal's areas under the CSD plot: $OL_{lo} < OL_{br}$ - Order of areas under the CSD plot in the looped network: CC > OL > LC > GL > NL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.

Table 2.63. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Area under the cross-spectral plot (for ND signal)	LC	- Comparison of LC signal's areas under the CSD plot: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$ - Order of areas under the CSD plot in the looped network: CC > OL > LC > GL > NL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.
	CC	- Comparison of CC signal's areas under the CSD plot: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$ - Order of areas under the CSD plot in the looped network: CC > OL > LC > GL > NL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.
	GL	- Comparison of GL signal's areas under the CSD plot: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$ - Order of areas under the CSD plot in the looped network: CC > OL > LC > GL > NL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.
	All	- With ND, there is no consistent pattern in the relations of the two networks' areas under the CSD plots. Therefore, the area under the CSD plot cannot identify the network change. - Order of areas under the CSD plot in the looped network: CC > OL > LC > GL > NL. - Order of areas under the CSD plot in the branched network: LC > OL > CC > GL > NL.

2.12.5.6. LDI

2.12.5.6.1. Scatter Plot

Figure 2.251 shows the scatter plots of the LDI for dynamic pressure data measured in the looped network with 0 (GPM), 3 (GPM), and 7.5 (GPM) demand variants, where the horizontal axis is the leaks' measured flow.

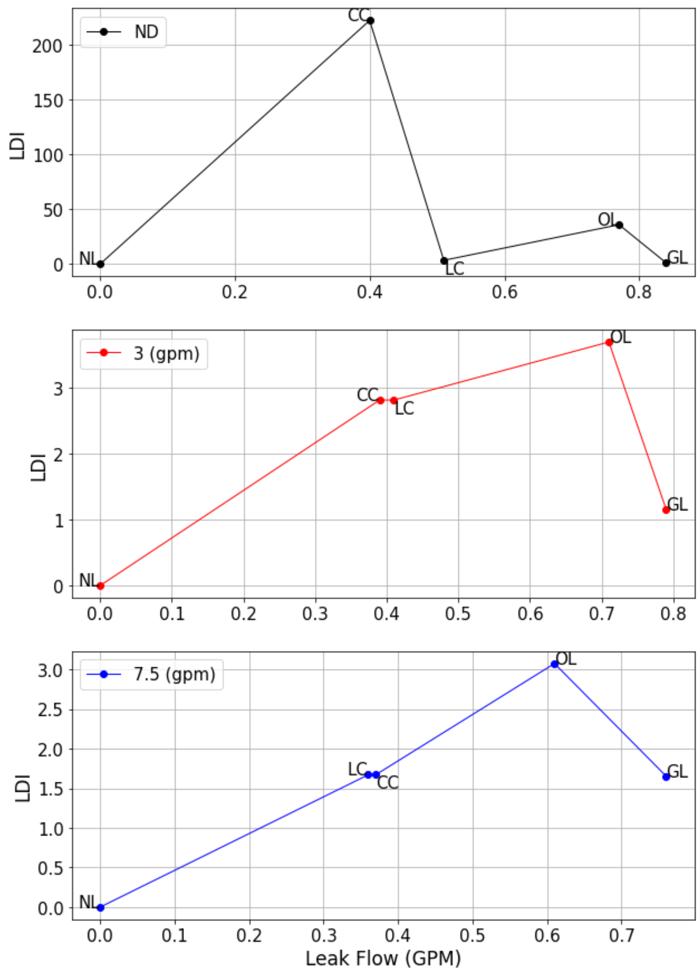

Figure 2.251. Scatter plots of the LDI for dynamic pressure data measured in the looped network

Figure 2.252 shows the scatter plots of the LDI for acoustic data measured in the branched network with 0 (GPM), 3 (GPM), and 7.5 (GPM) demand variants, where the horizontal axis is the leaks' measured flow.

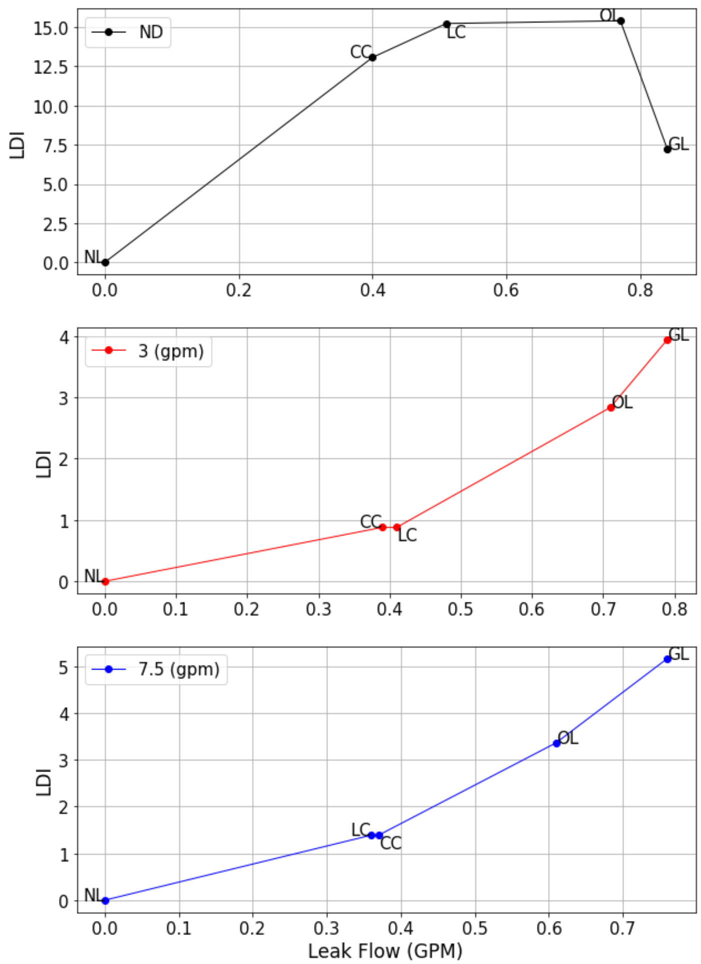

Figure 2.252. Scatter plots of the LDI for dynamic pressure data measured in the branched network

Table 2.64 includes an analysis of the LDI scatter plots for dynamic pressure data of the looped and branched networks.

Table 2.64. Analysis of the LDI scatter plots for dynamic pressure data recorded in the looped and branched networks

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
LDI (scatter plot)	All leaks and demands	- Since the LDIs of leaks are larger than those of the benchmark, i.e., NL, the LDI can detect leaks. This conforms to the results of the paper of Yazdekhasti et al., 2016. - When demand exists in the looped network, the LDI has the following pattern: OL > LC = CC > GL. - When demand exists in the branched network, the LDI has the following pattern: GL > OL > CC = LC. - Due to the similarity of LC and CC LDI magnitudes, LCI cannot discern leak types. Moreover, since the LDI magnitudes varied with the demand change, one cannot set a threshold to assign a specific LDI magnitude to a leak or a network. Therefore, the LDI cannot discern leak types or identify network change using dynamic pressure data.

2.12.5.6.2. Bar Plot

Figures 2.253 and 2.254 show the bar plots of the LDI for dynamic pressure data measured in the looped and branched networks, respectively, for all leak and demand variants.

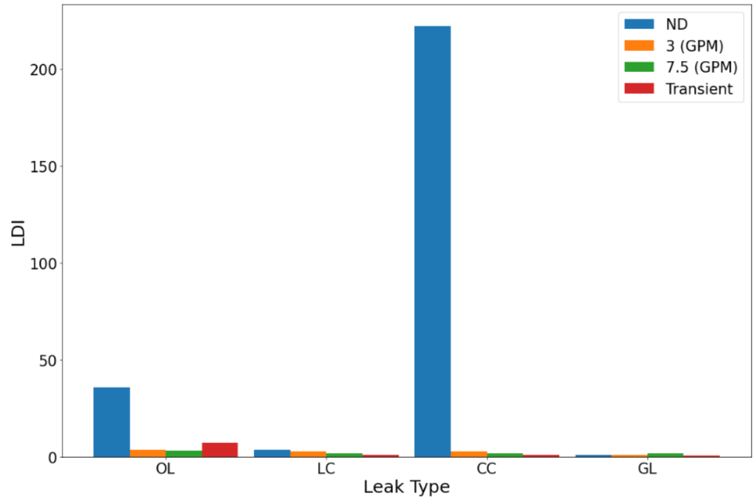

Figure 2.253. Bar plot of the LDI for dynamic pressure data measured in the looped network with all leak and demand variants

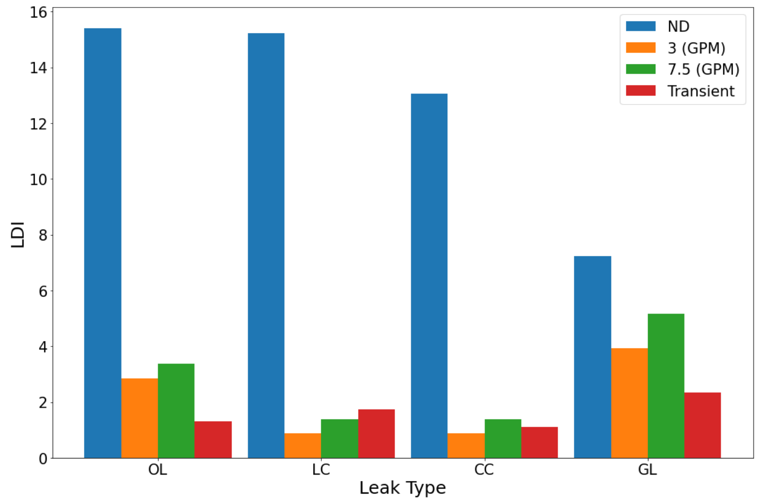

Figure 2.254. Bar plot of the LDI for dynamic pressure data measured in the branched network with all leak and demand variants

Table 2.65 compares the LDI bar plots of dynamic pressure data measured in the looped and branched networks with all leak and demand variants.

Table 2.65. Analysis of dynamic pressure data LDI bar plots measured in the looped and branched networks with all leak and demand variants

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
LDI (bar plot)	ND	- Comparison of OL signal's LDI with no demand: $OL_{lo} > OL_{br}$ - Comparison of CC signal's LDI with no demand: $CC_{lo} > CC_{br}$ - Comparison of LC signal's LDI with no demand: $LC_{lo} < LC_{br}$ - Comparison of GL signal's LDI with no demand: $GL_{lo} < GL_{br}$ - Order of LDI for signals with ND in the looped network: CC > OL > LC > GL. - Order of LDI for signals with ND in the branched network: OL > LC > CC > GL.
	3 (GPM)	- Comparison of OL signal's LDI with 3 (GPM) demand: $OL_{lo} > OL_{br}$ - Comparison of CC signal's LDI with 3 (GPM) demand: $CC_{lo} > CC_{br}$ - Comparison of LC signal's LDI with 3 (GPM) demand: $LC_{lo} > LC_{br}$ - Comparison of GL signal's LDI with 3 (GPM) demand: $GL_{lo} < GL_{br}$ - Order of LDI for signals with 3 (GPM) demand in the looped network: OL > CC = LC > GL. - Order of LDI for signals with 3 (GPM) demand in the branched network: GL > OL > LC = CC.
	7.5 (GPM)	- Comparison of OL signal's LDI with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$ - Comparison of CC signal's LDI with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$ - Comparison of LC signal's LDI with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$ - Comparison of GL signal's LDI with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$ - Order of LDI for signals with 7.5 (GPM) demand in the looped network: OL > CC = LC > GL. - Order of LDI for signals with 7.5 (GPM) demand in the branched network: GL > OL > CC = LC.

Table 2.65. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
LDI (bar plot)	Transient	- Comparison of OL signal's LDI with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$ - Comparison of CC signal's LDI with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$ - Comparison of LC signal's LDI with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$ - Comparison of GL signal's LDI with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$ - Order of LDI for signals with 7.5 (GPM) demand in the looped network: OL > CC > LC > GL. - Order of LDI for signals with 7.5 (GPM) demand in the branched network: GL > LC > OL > CC.
	All	- Due to the similarity of CC and LC LDIs of the looped and branched networks, LDI cannot discern leak types. - With a non-zero demand, OL and GL have the largest LDIs in the looped and branched networks, respectively. - Due to the inconsistent relations between the leak LDIs of both networks, LDI cannot identify the network change. - No LDI threshold could be set to discern leak types and network change effects.

2.12.5.7. Leak:NoLeak Amplitude Plot

Figures 2.255 and 2.256 show leak:noleak amplitude plots of the dynamic pressure data measured by sensor P1 in the looped and branched networks, respectively, for all leak types and no demand. Figures 2.257 and 2.258 show the same plots but for sensor P2 data.

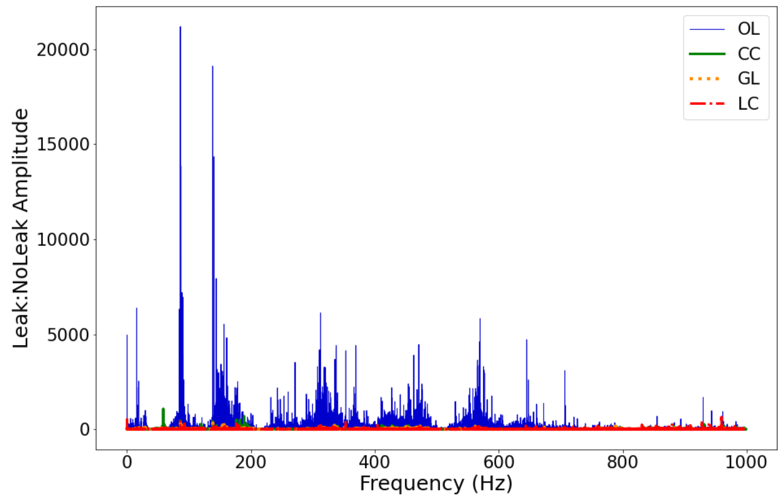

Figure 2.255. Leak:noleak amplitude plot of the dynamic pressure data measured by sensor P1 in the looped network with no demand

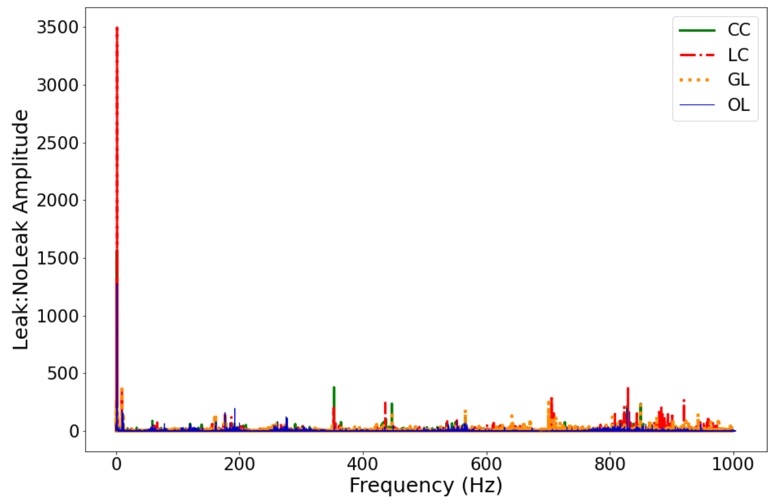

Figure 2.256. Leak:noleak amplitude plot of the dynamic pressure data measured by sensor P1 in the branched network with no demand

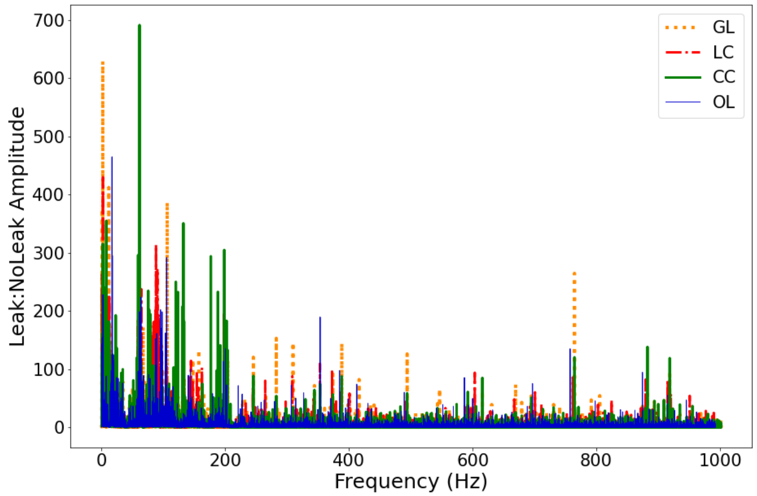

Figure 2.257. Leak:noleak amplitude plot of the dynamic pressure data measured by sensor P2 in the looped network with no demand

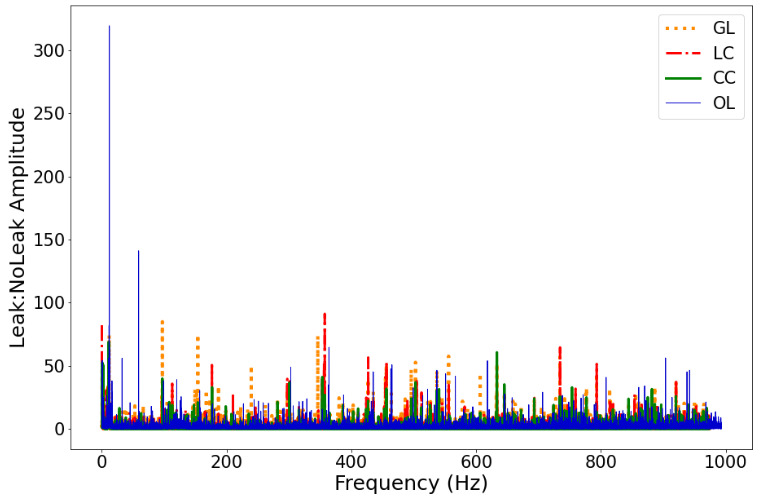

Figure 2.258. Leak:noleak amplitude plot of the dynamic pressure data measured by sensor P2 in the branched network with no demand

Table 2.66 compares leak:noleak amplitude plots of dynamic pressure data measured in the looped and branched networks with all leak types and no demand.

Table 2.66. Analysis of leak:noleak amplitude plots of dynamic pressure data measured in the looped and branched networks with all leak types and no demand

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Sum of magnitudes in the leak:noleak amplitude plot (for ND signal)	OL – P1	- Comparison of the sum of leak:noleak magnitudes for OL: $OL_{lo} > OL_{br}$. - Order of sum of leak:noleak magnitudes in the looped network: OL > LC > CC > GL. - Order of sum of leak:noleak magnitudes in the branched network: LC > GL > CC > OL.
	LC – P1	- Comparison of the sum of leak:noleak magnitudes for LC: $LC_{lo} > LC_{br}$. - Order of sum of leak:noleak magnitudes in the looped network: OL > LC > CC > GL. - Order of sum of leak:noleak magnitudes in the branched network: LC > GL > CC > OL.
	CC – P1	- Comparison of the sum of leak:noleak magnitudes for CC: $CC_{lo} > CC_{br}$ - Order of sum of leak:noleak magnitudes in the looped network: OL > LC > CC > GL. - Order of sum of leak:noleak magnitudes in the branched network: LC > GL > CC > OL.
	GL – P1	- Comparison of the sum of leak:noleak magnitudes for GL: $GL_{lo} > GL_{br}$ - Order of sum of leak:noleak magnitudes in the looped network: OL > LC > CC > GL. - Order of sum of leak:noleak magnitudes in the branched network: LC > GL > CC > OL.
	All – P1	- Comparing the sum of magnitudes in the leak:noleak plots of two networks with ND indicates that the leak:noleak ratios are larger in the looped network compared to the branched network. Therefore, based on P1 data, the sum of magnitudes in the leak:noleak plots can identify the network change when there is no demand. - Based on P1, OL in the looped network has the largest sum of leak:noleak magnitudes compared to all other leaks in both networks. - Based on P1, we did not observe any relation between the sum of magnitudes in the leak:noleak plots and leak flows. - Order of sum of leak:noleak magnitudes in the looped network: OL > LC > CC > GL. - Order of sum of leak:noleak magnitudes in the branched network: LC > GL > CC > OL.

Table 2.66. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Sum of magnitudes in the leak:noleak amplitude plot (for ND signal)	OL – P2	- Comparison of the sum of leak:noleak magnitudes for OL: $\mathrm{OL}_{\mathrm{lo}} > \mathrm{OL}_{\mathrm{br}}$. - Order of sum of leak:noleak magnitudes in the looped network: CC > GL > LC > OL. - Order of sum of leak:noleak magnitudes in the branched network: GL > OL > LC > CC.
	LC – P2	- Comparison of the sum of leak:noleak magnitudes for LC: $\mathrm{LC}_{\mathrm{lo}} > \mathrm{LC}_{\mathrm{br}}$. - Order of sum of leak:noleak magnitudes in the looped network: CC > GL > LC > OL. - Order of sum of leak:noleak magnitudes in the branched network: GL > OL > LC > CC.
	CC – P2	- Comparison of the sum of leak:noleak magnitudes for CC: $\mathrm{CC}_{\mathrm{lo}} > \mathrm{CC}_{\mathrm{br}}$ - Order of sum of leak:noleak magnitudes in the looped network: CC > GL > LC > OL. - Order of sum of leak:noleak magnitudes in the branched network: GL > OL > LC > CC.
	GL – P2	- Comparison of the sum of leak:noleak magnitudes for GL: $\mathrm{GL}_{\mathrm{lo}} > \mathrm{GL}_{\mathrm{br}}$ - Order of sum of leak:noleak magnitudes in the looped network: CC > GL > LC > OL. - Order of sum of leak:noleak magnitudes in the branched network: GL > OL > LC > CC.
	All – P2	- Comparing the sum of magnitudes in the leak:noleak plots of two networks with ND indicates that the leak:noleak ratios are larger in the looped network compared to the branched network. Therefore, based on P2 data, the sum of magnitudes in the leak:noleak plots can identify the network change. - Based on P2 in the branched network, the larger the leak flow, the larger the sum of magnitudes in the leak:noleak plots. - Order of sum of leak:noleak magnitudes in the looped network: CC > GL > LC > OL. - Order of sum of leak:noleak magnitudes in the branched network: GL > OL > LC > CC.

2.12.5.8. Dominant Frequency

Figures 2.259 and 2.260 show dominant frequency bar plots of the dynamic pressure data measured by sensor P1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.261 and 2.262 show the same plots but for sensor P2 data.

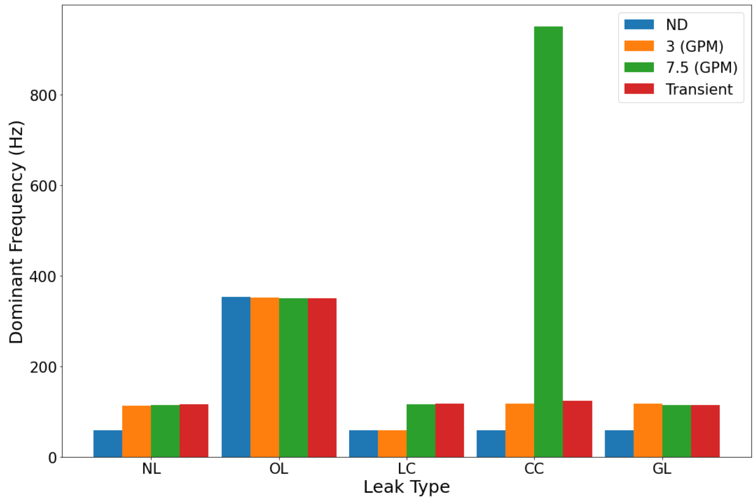

Figure 2.259. Dominant frequency bar plot of sensor P1 data in the looped network for all leaks and demands

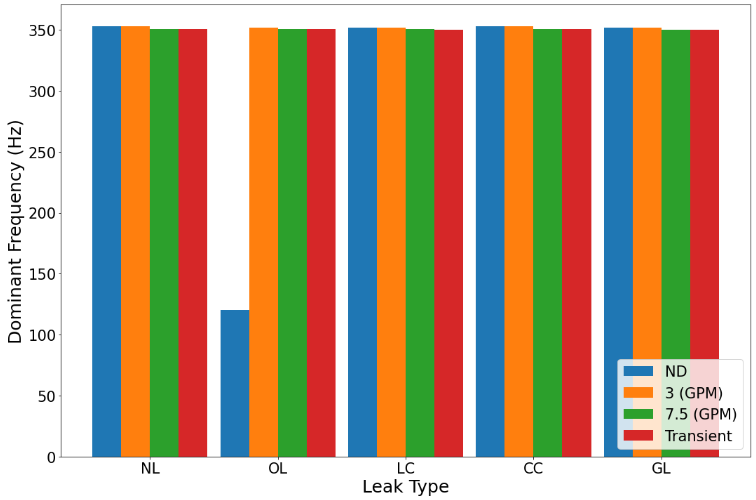

Figure 2.260. Dominant frequency bar plot of sensor P1 data in the branched network for all leaks and demands

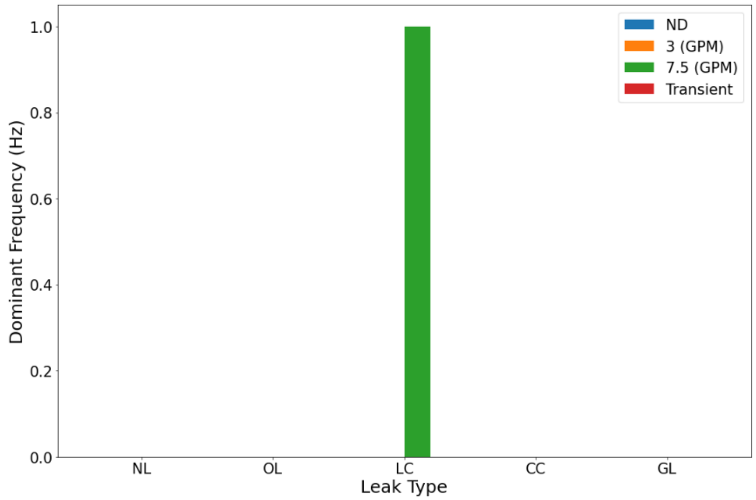

Figure 2.261. Dominant frequency bar plot of sensor P2 data in the looped network for all leaks and demands

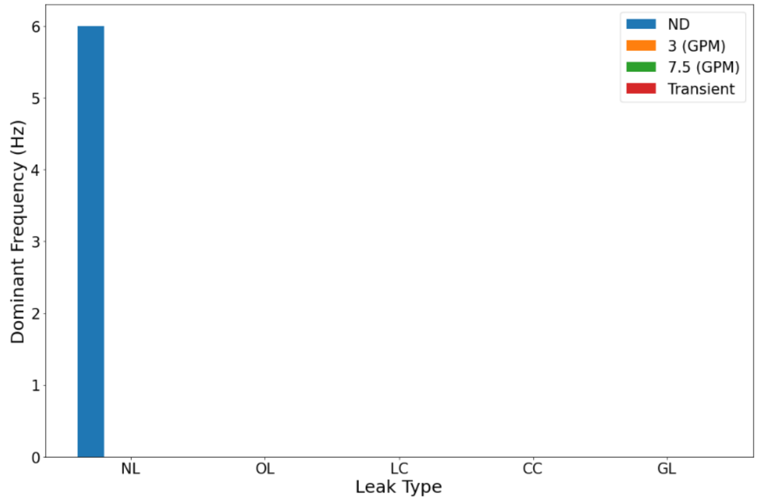

Figure 2.262. Dominant frequency bar plot of sensor P2 data in the branched network for all leaks and demands

Table 2.67 compares dominant frequency plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types by sensors P1 and P2.

Table 2.67. Analysis of dominant frequency plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Dominant frequency	ND – P1	- Comparison of NL signal's dominant frequency with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with no demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with no demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with no demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$.

Table 2.67. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Dominant frequency	ND – P1	- Comparison of GL signal's dominant frequency with no demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of dominant frequency with ND in the looped network: OL > NL = LC = CC = GL. - Order of dominant frequency with ND in the branched network: NL = CC > GL = LC > CC.
	3 (GPM) – P1	- Comparison of NL signal's dominant frequency with 3 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with 3 (GPM) demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with 3 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with 3 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with 3 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of dominant frequency for signals with 3 (GPM) demand in the looped network: OL > GL = CC > NL > LC. - Order of dominant frequency for signals with 3 (GPM) demand in the branched network: NL = CC > OL = LC = GL.
	7.5 (GPM) – P1	- Comparison of NL signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of dominant frequency for signals with 7.5 (GPM) demand in the looped network: CC > OL > LC > GL > NL. - Order of dominant frequency for signals with 7.5 (GPM) demand in the branched network: NL = OL = CC = LC > GL.

Table 2.67. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Dominant frequency	Transient – P1	- Comparison of NL signal's dominant frequency with transient demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with transient demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of dominant frequency for signals with transient demand in the looped network: OL > CC > LC > NL > GL. - Order of dominant frequency for signals with transient demand in the branched network: NL = OL = CC > LC > GL.
	All – P1	- Comparing leaks' dominant frequency magnitudes of two networks indicates no consistent pattern in the magnitudes when the network changed. Therefore, dominant frequency is not capable of identifying the network change. - Since there is no consistent order of dominant frequency for signals with different demands in both networks or due to the similarity of those magnitudes, dominant frequency is not capable of discerning leak types consistently in both networks.
	ND – P2	- Comparison of NL signal's dominant frequency with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with no demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with no demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with no demand: $\mathrm{LC_{lo}} = \mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with no demand: $\mathrm{GL_{lo}} = \mathrm{GL_{br}}$. - Order of dominant frequency with ND in the looped network: NL = OL = LC = CC = GL. - Order of dominant frequency with ND in the branched network: NL > OL = LC = CC = NL.

Table 2.67. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Dominant frequency	3 (GPM) – P2	- Comparison of NL signal's dominant frequency with 3 (GPM) demand: $\mathrm{NL_{lo}} = \mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with 3 (GPM) demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with 3 (GPM) demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with 3 (GPM) demand: $\mathrm{LC_{lo}} = \mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with 3 (GPM) demand: $\mathrm{GL_{lo}} = \mathrm{GL_{br}}$. - Order of dominant frequency for signals with 3 (GPM) demand in the looped network: NL = OL = LC = CC = GL. - Order of dominant frequency for signals with 3 (GPM) demand in the branched network: NL = OL = LC = CC = GL.
	7.5 (GPM) – P2	- Comparison of NL signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{NL_{lo}} = \mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with 7.5 (GPM) demand: $\mathrm{GL_{lo}} = \mathrm{GL_{br}}$. - Order of dominant frequency for signals with 7.5 (GPM) demand in the looped network: NL = OL = CC = GL > LC. - Order of dominant frequency for signals with 7.5 (GPM) demand in the branched network: NL = OL = LC = CC = GL.
	Transient – P2	- Comparison of NL signal's dominant frequency with transient demand: $\mathrm{NL_{lo}} = \mathrm{NL_{br}}$. - Comparison of OL signal's dominant frequency with transient demand: $\mathrm{OL_{lo}} = \mathrm{OL_{br}}$. - Comparison of CC signal's dominant frequency with transient demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's dominant frequency with transient demand: $\mathrm{LC_{lo}} = \mathrm{LC_{br}}$. - Comparison of GL signal's dominant frequency with transient demand: $\mathrm{GL_{lo}} = \mathrm{GL_{br}}$.

Table 2.67. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Dominant frequency	Transient – P2	- Order of dominant frequency for signals with transient demand in the looped network: NL = OL = LC = CC = GL. - Order of dominant frequency for signals with transient demand in the branched network: NL = OL = LC = CC = GL.
	All – P2	- The majority of the leaks' dominant frequencies in both networks are 0 Hz. - Due to the similarity of leaks' dominant frequency magnitudes of two networks, dominant frequency cannot identify the network changes. - Due to the similarity of dominant frequencies of different leak types in both networks, dominant frequency is not capable of discerning leak types in both networks.

2.12.5.9. Fundamental Frequency

Figures 2.263 and 2.264 show fundamental frequency bar plots of the dynamic pressure data measured by sensor P1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.265 and 2.266 show the same plots but for sensor P2 data.

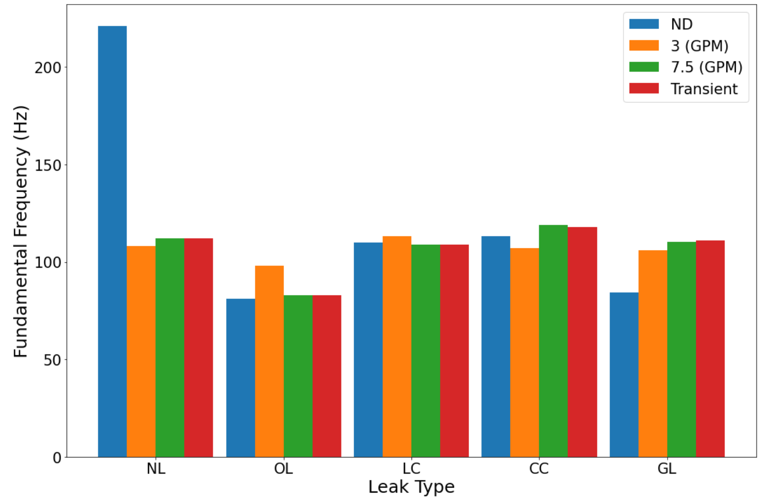

Figure 2.263. Fundamental frequency bar plot of sensor P1 data in the looped network for all leaks and demands

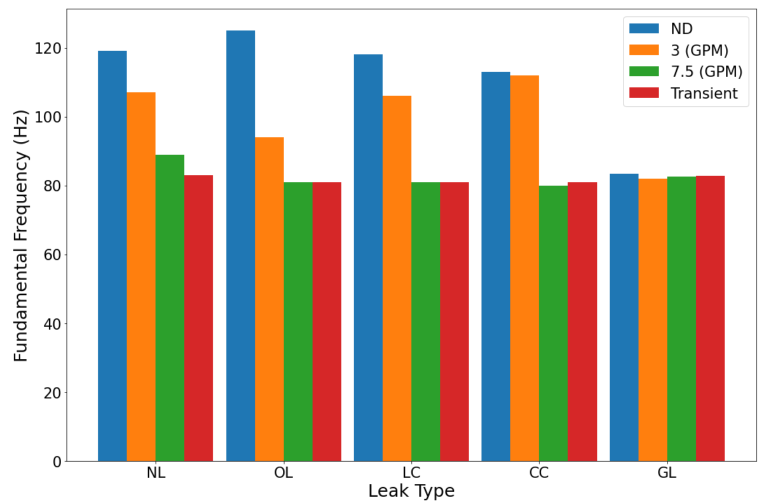

Figure 2.264. Fundamental frequency bar plot of sensor P1 data in the branched network for all leaks and demands

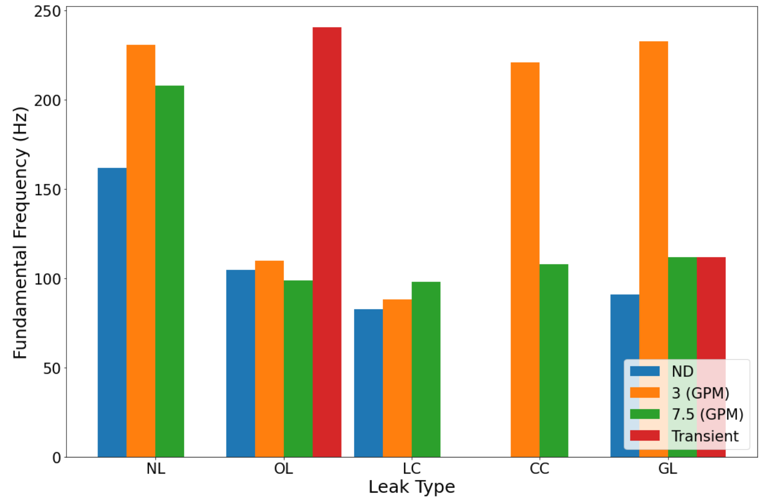

Figure 2.265. Fundamental frequency bar plot of sensor P2 data in the looped network for all leaks and demands

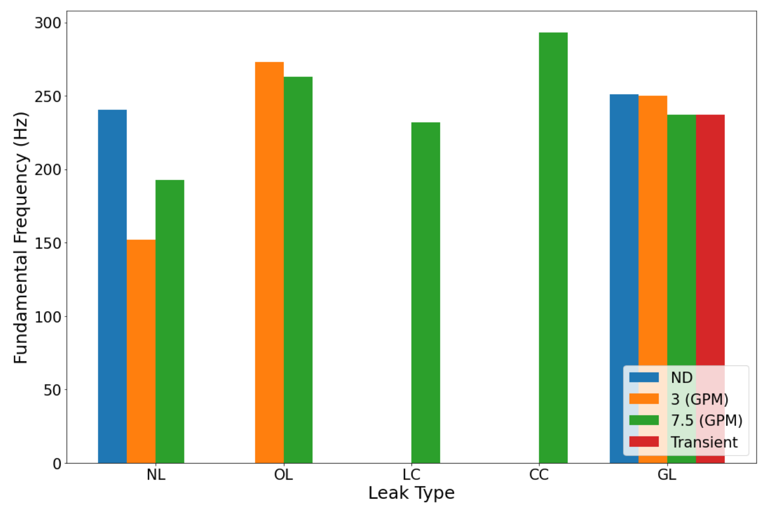

Figure 2.266. Fundamental frequency bar plot of sensor P2 data in the branched network for all leaks and demands

Table 2.68 compares fundamental frequency plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types by sensors P1 and P2.

Table 2.68. Analysis of fundamental frequency plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Fundamental frequency	ND – P1	- Comparison of NL signal's fundamental frequency with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's fundamental frequency with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's fundamental frequency with no demand: $CC_{lo} = CC_{br}$. - Comparison of LC signal's fundamental frequency with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's fundamental frequency with no demand: $GL_{lo} > GL_{br}$. - Order of fundamental frequency with ND in the looped network: NL > CC > LC > GL > OL. - Order of fundamental frequency with ND in the branched network: OL > NL > LC > CC > GL.
	3 (GPM) – P1	- Comparison of NL signal's fundamental frequency with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's fundamental frequency with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's fundamental frequency with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's fundamental frequency with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's fundamental frequency with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of fundamental frequency for signals with 3 (GPM) demand in the looped network: LC > NL > CC > GL > OL. - Order of fundamental frequency for signals with 3 (GPM) demand in the branched network: CC > NL > LC > OL > GL.

Table 2.68. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Fundamental frequency	7.5 (GPM) – P1	- Comparison of NL signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with 7.5 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of fundamental frequency for signals with 7.5 (GPM) demand in the looped network: CC > NL > GL > LC > OL. - Order of fundamental frequency for signals with 7.5 (GPM) demand in the branched network: NL > GL > OL = LC > CC.
	Transient – P1	- Comparison of NL signal's fundamental frequency with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of fundamental frequency for signals with transient demand in the looped network: CC > NL > GL > LC > OL. - Order of fundamental frequency for signals with transient demand in the branched network: NL > GL > OL = LC = CC.

Table 2.68. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Fundamental frequency	All – P1	- The fundamental frequencies of the looped network's leak signals are between 83 Hz and 119 Hz. - The fundamental frequencies of the branched network's leak signals are between 80 Hz and 112 Hz. - In the looped network, the fundamental frequency of all leaks with different demands do not have a specific pattern. Therefore, the fundamental frequency cannot discern leak types in the looped network. - In the branched network, due to the similarity of the fundamental frequencies of all leaks with different demands, the fundamental frequency cannot discern leak types in the branched network. - Comparing leaks' fundamental frequency magnitudes of two networks does not indicate a consistent change pattern in the magnitudes when the network changed. Therefore, the fundamental frequency cannot identify the network change.
	ND – P2	- Comparison of NL signal's fundamental frequency with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with no demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with no demand: $\mathrm{CC_{lo}} = \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with no demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$.
	3 (GPM) – P2	- Comparison of NL signal's fundamental frequency with 3 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's fundamental frequency with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's fundamental frequency with 3 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's fundamental frequency with 3 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's fundamental frequency with 3 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of fundamental frequency with 3 (GPM) demand in the looped network: GL > NL > CC > OL > LC. - Order of fundamental frequency with 3 (GPM) demand in the branched network: OL > GL > NL > LC = CC = 0.

Table 2.68. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Fundamental frequency	7.5 (GPM) – P2	- Comparison of NL signal's fundamental frequency with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's fundamental frequency with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's fundamental frequency with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's fundamental frequency with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's fundamental frequency with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of fundamental frequency with 7.5 (GPM) demand in the looped network: NL > GL > CC > OL > LC. - Order of fundamental frequency with 7.5 (GPM) demand in the branched network: CC > OL > GL > LC > NL.
	Transient – P2	- Comparison of NL signal's fundamental frequency with transient demand: $NL_{lo} = NL_{br}$. - Comparison of OL signal's fundamental frequency with transient demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's fundamental frequency with transient demand: $CC_{lo} = CC_{br}$. - Comparison of LC signal's fundamental frequency with transient demand: $LC_{lo} = LC_{br}$. - Comparison of GL signal's fundamental frequency with transient demand: $GL_{lo} < GL_{br}$. - Order of fundamental frequency with transient demand in the looped network: OL > GL > NL = LC = CC = 0. - Order of fundamental frequency with transient demand in the branched network: GL = NL = OL = LC = CC = 0.
	All – P2	- Due to the inconsistent fundamental frequency order of leak and no leak signals in both networks, fundamental frequency cannot discern leak types in both networks. - Comparing leak and no leak signals' fundamental frequencies of both networks indicates inconsistent patterns or similar fundamental frequency magnitudes. Therefore, fundamental frequency cannot identify the network change using P2 data.

2.12.5.10. Spectral Centroid

Figures 2.267 and 2.268 show spectral centroid bar plots of the dynamic pressure data measured by sensor P1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.269 and 2.270 show the same plots but for sensor P2 data.

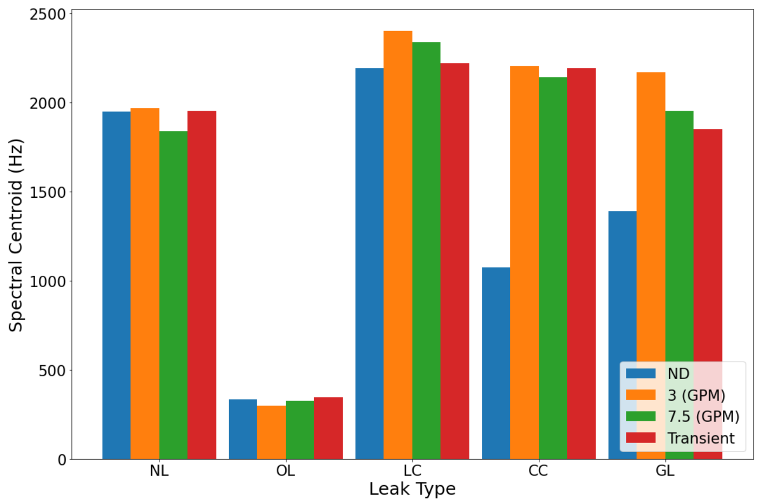

Figure 2.267. Spectral centroid bar plot of the sensor P1 data in the looped network for all leaks and demands

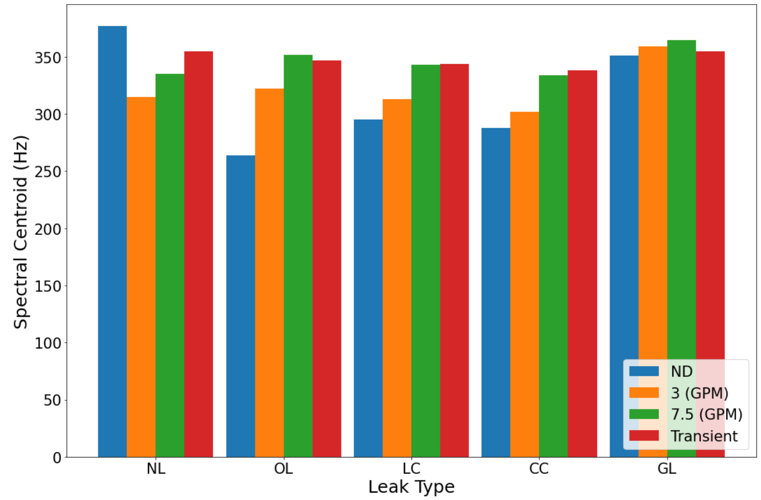

Figure 2.268. Spectral centroid bar plot of the sensor P1 data in the branched network for all leaks and demands

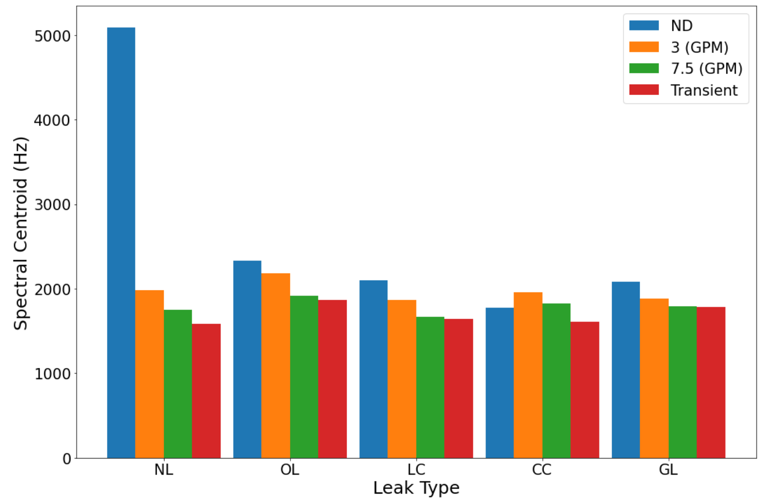

Figure 2.269. Spectral centroid bar plot of the sensor P2 data in the looped network for all leaks and demands

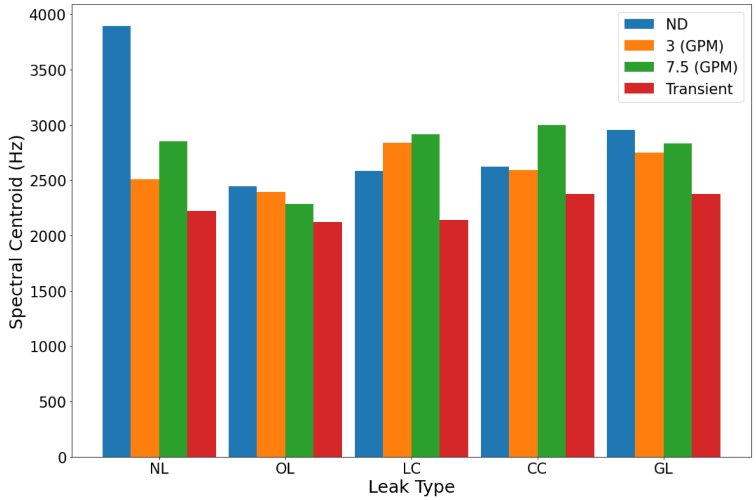

Figure 2.270. Spectral centroid bar plot of the sensor P2 data in the branched network for all leaks and demands

Table 2.69 compares spectral centroid plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types by sensors P1 and P2.

Table 2.69. Analysis of spectral centroid plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Spectral centroid	ND – P1	- Comparison of NL signal's spectral centroid with no demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's spectral centroid with no demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's spectral centroid with no demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's spectral centroid with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's spectral centroid with no demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of spectral centroid with ND in the looped network: LC > NL > GL > CC > OL. - Order of spectral centroid with ND in the branched network: NL > GL > LC > CC > OL.
	3 (GPM) – P1	- Comparison of NL signal's spectral centroid with 3 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's spectral centroid with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's spectral centroid with 3 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's spectral centroid with 3 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's spectral centroid with 3 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of spectral centroid for signals with 3 (GPM) demand in the looped network: LC > CC > GL > NL > OL. - Order of spectral centroid for signals with 3 (GPM) demand in the branched network: GL > OL > NL > LC > CC.

Table 2.69. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Spectral centroid	7.5 (GPM) – P1	- Comparison of NL signal's spectral centroid with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's spectral centroid with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's spectral centroid with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's spectral centroid with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's spectral centroid with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of spectral centroid for signals with 7.5 (GPM) demand in the looped network: LC > CC > GL > NL > OL. - Order of spectral centroid for signals with 7.5 (GPM) demand in the branched network: GL > OL > LC > NL > CC.
	Transient – P1	- Comparison of NL signal's spectral centroid with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's spectral centroid with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's spectral centroid with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's spectral centroid with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's spectral centroid with transient demand: $GL_{lo} > GL_{br}$. - Order of spectral centroid for signals with transient demand in the looped network: LC > CC > NL > GL > OL. - Order of spectral centroid for signals with transient demand in the branched network: NL = GL > OL > LC > CC.
	All – P1	- There is no consistent pattern in the relations of leak and no leak signals' spectral centroid of both networks. Therefore, the spectral centroid cannot identify the network change. - Should we ignore NL signal, leak signals' spectral centroids in the looped network have the following order: LC > CC > GL > OL. Therefore, using P1 data when demand is non-zero, spectral centroid can discern leak types in the looped network. - Similar to the looped network, if we ignore NL signal, leak signals' spectral centroids in the branched network have the following order: GL > OL > LC > CC. Therefore, using P1 data when demand is non-zero, spectral centroid can discern leak types in the branched network.

Table 2.69. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Spectral centroid	ND – P2	- Comparison of NL signal's spectral centroid with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's spectral centroid with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's spectral centroid with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's spectral centroid with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's spectral centroid with no demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid with ND in the looped network: NL > OL > LC = GL > CC. - Order of spectral centroid with ND in the branched network: NL > GL > CC > LC > OL.
	3 (GPM) – P2	- Comparison of NL signal's spectral centroid with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's spectral centroid with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's spectral centroid with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's spectral centroid with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's spectral centroid with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid for signals with 3 (GPM) demand in the looped network: OL > NL > CC > GL > LC. - Order of spectral centroid for signals with 3 (GPM) demand in the branched network: LC > GL > CC > NL > OL.
	7.5 (GPM) – P2	- Comparison of NL signal's spectral centroid with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's spectral centroid with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's spectral centroid with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's spectral centroid with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's spectral centroid with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of spectral centroid for signals with 7.5 (GPM) demand in the looped network: OL > CC > GL > NL > LC. - Order of spectral centroid for signals with 7.5 (GPM) demand in the branched network: CC > LC > NL > GL > OL.

Table 2.69. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Spectral centroid	Transient – P2	- Comparison of NL signal's spectral centroid with transient demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's spectral centroid with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's spectral centroid with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's spectral centroid with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's spectral centroid with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of spectral centroid for signals with transient demand in the looped network: OL > GL > LC > CC > NL. - Order of spectral centroid for signals with transient demand in the branched network: GL > CC > NL > LC > OL.
	All – P2	- Based on P2, when there is a demand, leak and no leak signals' spectral centroid magnitudes of the branched network are larger than those of the looped network. Therefore, the spectral centroid can identify the network change. - Since there is no consistent order of spectral centroid for signals with different demands in both networks, the spectral centroid cannot discern leak types consistently in both networks. - When there is demand in the looped network, OL has the largest spectral centroid that represents the signal's larger amplitude magnitudes at higher frequencies. On the other hand, OL spectral centroid in the branched network has the smallest magnitude. Since there are more dead ends and fewer connected pipes in the branched network, high frequency contents of the OL signal's dynamic pressure waves become more attenuated. However, in the looped network, since pipes are more connected, high frequency contents of the OL signal's dynamic pressure waves propagate more significantly due to more flow continuity.

2.12.5.11. Power Spectral Entropy

Figures 2.271 and 2.272 show power spectral entropy bar plots of the dynamic pressure data measured by sensor P1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.273 and 2.274 show the same plots but for sensor P2 data.

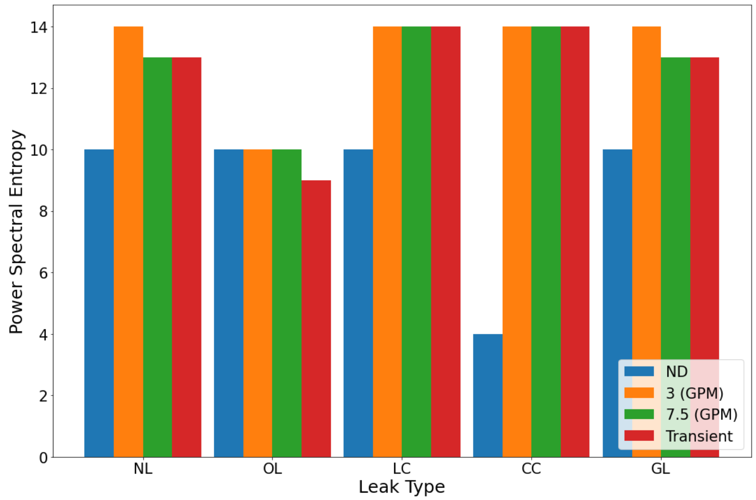

Figure 2.271. Power spectral entropy bar plot of sensor P1 data in the looped network for all leaks and demands

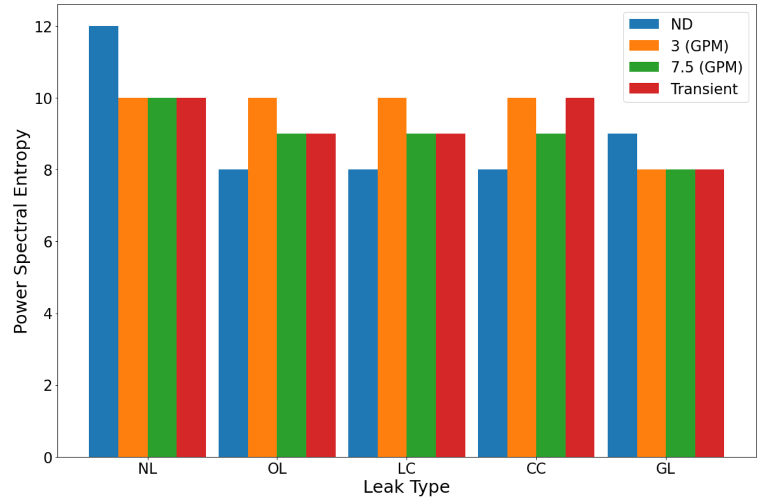

Figure 2.272. Power spectral entropy bar plot of sensor P1 data in the branched network for all leaks and demands

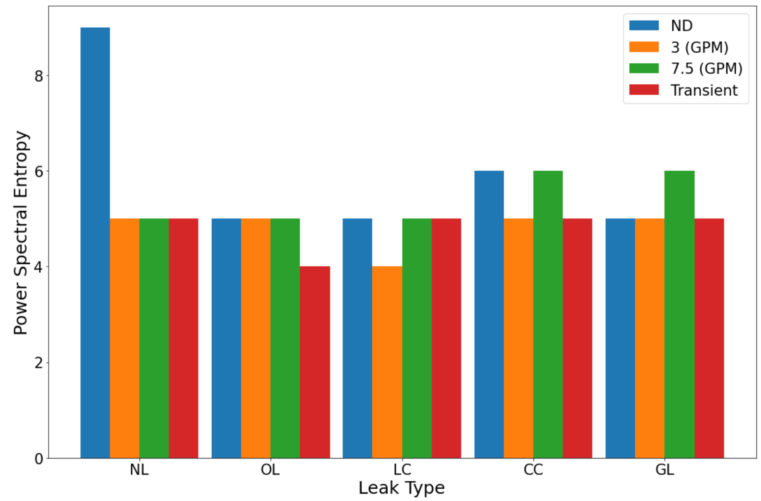

Figure 2.273. Power spectral entropy bar plot of sensor P2 data in the looped network for all leaks and demands

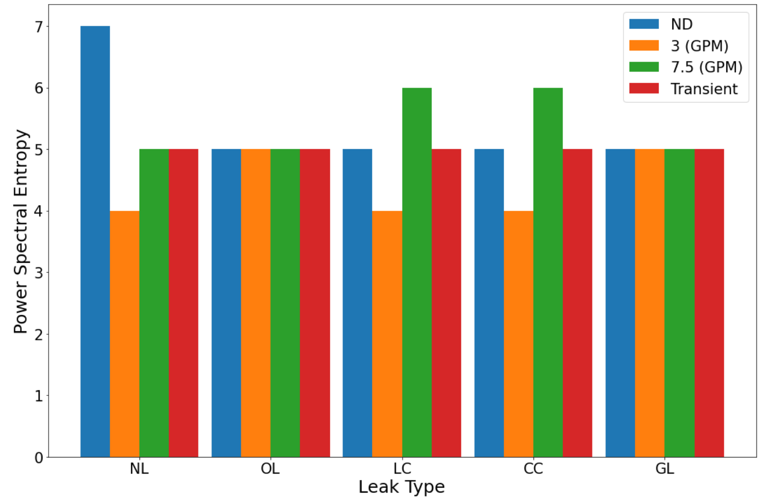

Figure 2.274. Power spectral entropy bar plot of sensor P2 data in the branched network for all leaks and demands

Table 2.70 compares power spectral entropy plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types by sensors P1 and P2.

Table 2.70. Analysis of power spectral entropy plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Power spectral entropy	ND – P1	- Comparison of NL signal's power spectral entropy with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's power spectral entropy centroid with no demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's power spectral entropy centroid with no demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's power spectral entropy centroid with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$.

Table 2.70. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Power spectral entropy	ND – P1	- Comparison of GL signal's power spectral entropy centroid with no demand: $GL_{lo} > GL_{br}$. - Order of spectral centroid with ND in the looped network: NL = OL = LC = GL > CC. - Order of spectral centroid with ND in the branched network: NL > GL > OL = LC = CC.
	3 (GPM) – P1	- Comparison of NL signal's power spectral entropy with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with 3 (GPM) demand: $OL_{lo} = OL_{br}$. - Comparison of CC signal's power spectral entropy with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's power spectral entropy with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of power spectral entropy for signals with 3 (GPM) demand in the looped network: NL = LC = CC = GL > OL. - Order of power spectral entropy for signals with 3 (GPM) demand in the branched network: NL = OL = LC = CC > GL.
	7.5 (GPM) – P1	- Comparison of NL signal's power spectral entropy with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's power spectral entropy with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's power spectral entropy with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the looped network: LC = CC > NL = GL > OL. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the branched network: NL > OL = LC = CC > GL.
	Transient – P1	- Comparison of NL signal's power spectral entropy with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with transient demand: $OL_{lo} = OL_{br}$. - Comparison of CC signal's power spectral entropy with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's power spectral entropy with transient demand: $GL_{lo} > GL_{br}$.

Table 2.70. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Power spectral entropy	Transient – P1	- Order of power spectral entropy for signals with transient demand in the looped network: LC = CC > NL = GL > OL. - Order of power spectral entropy for signals with transient demand in the branched network: NL = CC > OL = LC > GL.
	All – P1	- Comparing leaks' power spectral entropy magnitudes of the two networks indicates that since OL signal's power spectral entropy in the looped network is equal to that of the branched network, power spectral entropy cannot distinguish the network change. - Due to the similarity of power spectral entropy magnitudes for signals with different demands in each network, power spectral entropy cannot discern leak types in both networks.
	ND – P2	- Comparison of NL signal's power spectral entropy with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with no demand: $OL_{lo} = OL_{br}$. - Comparison of CC signal's power spectral entropy with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with no demand: $LC_{lo} = LC_{br}$. - Comparison of GL signal's power spectral entropy with no demand: $GL_{lo} = GL_{br}$. - Order of power spectral entropy with ND in the looped network: NL > CC > OL = LC = GL. - Order of power spectral entropy with ND in the branched network: NL > OL = LC = CC = GL.
	3 (GPM) – P2	- Comparison of NL signal's power spectral entropy with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's power spectral entropy with 3 (GPM) demand: $OL_{lo} = OL_{br}$. - Comparison of CC signal's power spectral entropy with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's power spectral entropy with 3 (GPM) demand: $LC_{lo} = LC_{br}$. - Comparison of GL signal's power spectral entropy with 3 (GPM) demand: $GL_{lo} = GL_{br}$. - Order of power spectral entropy for signals with 3 (GPM) demand in the looped network: NL = OL = CC = GL > LC. - Order of power spectral entropy for signals with 3 (GPM) demand in the branched network: OL = GL > NL = LC = CC.

Table 2.70. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Power spectral entropy	7.5 (GPM) – P2	- Comparison of NL signal's power spectral entropy with 7.5 (GPM) demand: $NL_{lo} = NL_{br}$. - Comparison of OL signal's power spectral entropy with 7.5 (GPM) demand: $OL_{lo} = OL_{br}$. - Comparison of CC signal's power spectral entropy with 7.5 (GPM) demand: $CC_{lo} = CC_{br}$. - Comparison of LC signal's power spectral entropy with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's power spectral entropy with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the looped network: GL = CC > NL = OL = LC. - Order of power spectral entropy for signals with 7.5 (GPM) demand in the branched network: LC = CC > NL = OL = GL.
	Transient – P2	- Comparison of NL signal's power spectral entropy with transient demand: $NL_{lo} = NL_{br}$. - Comparison of OL signal's power spectral entropy with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's power spectral entropy with transient demand: $CC_{lo} = CC_{br}$. - Comparison of LC signal's power spectral entropy with transient demand: $LC_{lo} = LC_{br}$. - Comparison of GL signal's power spectral entropy with transient demand: $GL_{lo} = GL_{br}$. - Order of power spectral entropy for signals with transient demand in the looped network: NL = LC = CC = GL > OL. - Order of power spectral entropy for signals with transient demand in the branched network: NL = OL = LC = CC = GL.
	All – P2	- Based on the power spectral entropy of signals recorded by P2, due to equal magnitudes of the power spectral entropies of the two networks, power spectral entropy cannot identify the network change. - Due to the similarity of the spectral entropies of signals with different demands in both networks, the spectral entropy is not capable of discerning leak types in both networks.

2.12.5.12. Mean

Figures 2.275 and 2.276 plot the mean of the dynamic pressure data measured by sensor P1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.277 and 2.278 show the same plots but for sensor P2 data.

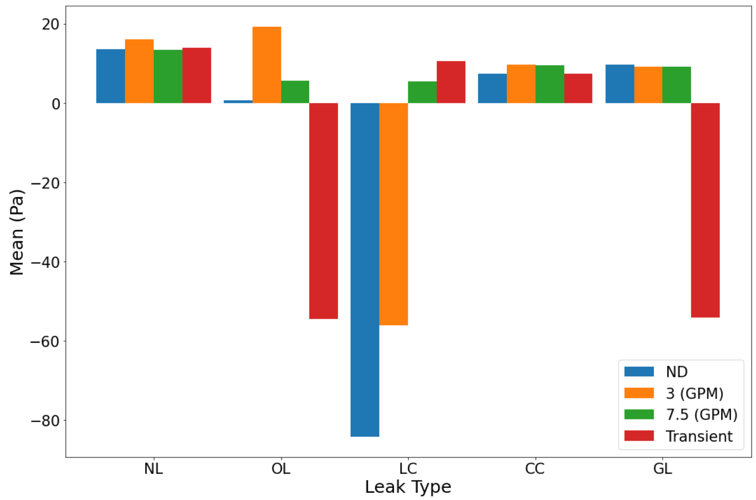

Figure 2.275. Mean plot of sensor P1 data in the looped network for all leaks and demands

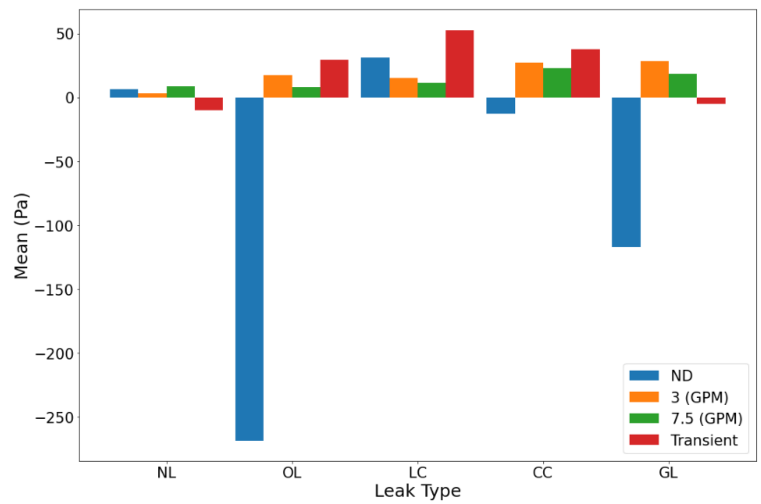

Figure 2.276. Mean plot of sensor P1 data in the branched network for all leaks and demands

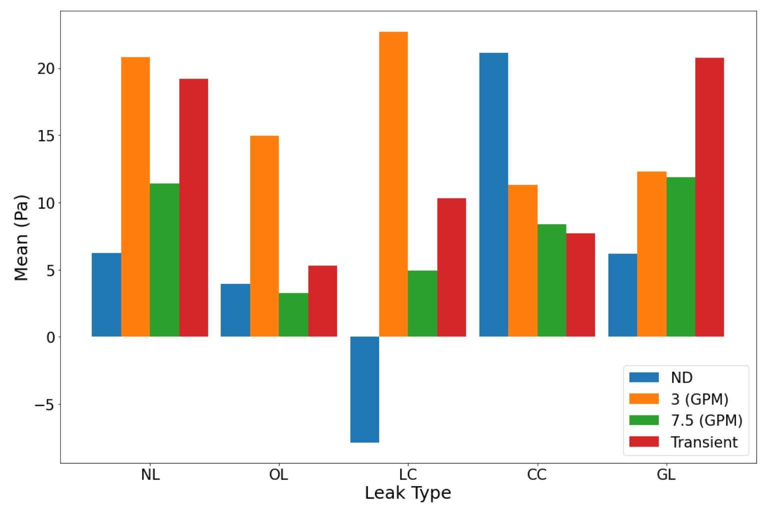

Figure 2.277. Mean plot of sensor P2 data in the looped network for all leaks and demands

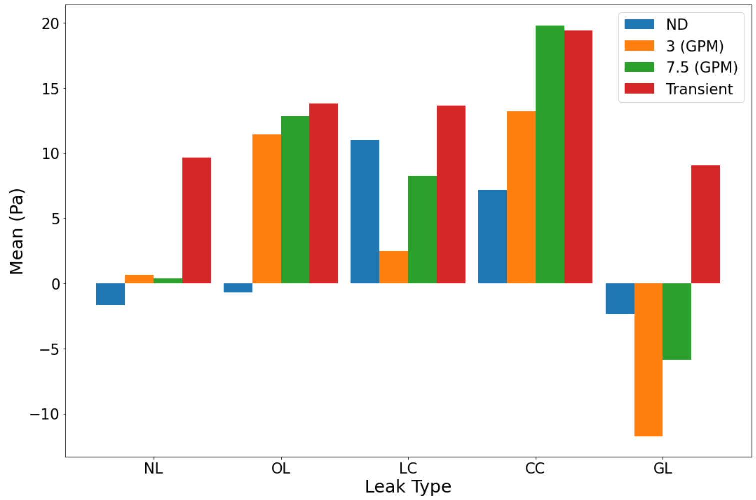

Figure 2.278. Mean plot of sensor P2 data in the branched network for all leaks and demands

Table 2.71 compares mean plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types by sensors P1 and P2.

Table 2.71. Analysis of mean plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Mean	ND – P1	- Comparison of NL signal's mean with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's mean with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's mean with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's mean with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's mean with no demand: $GL_{lo} > GL_{br}$. - Order of mean with ND in the looped network: NL > GL > CC > OL > LC. - Order of mean with ND in the branched network: LC > NL > CC > GL > OL.
	3 (GPM) – P1	- Comparison of NL signal's mean with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's mean with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's mean with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's mean with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's mean with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of mean for signals with 3 (GPM) demand in the looped network: OL > NL > CC > GL > LC. - Order of mean for signals with 3 (GPM) demand in the branched network: GL > CC > OL > LC > NL.
	7.5 (GPM) – P1	- Comparison of NL signal's mean with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's mean with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's mean with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's mean with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's mean with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of mean for signals with 7.5 (GPM) demand in the looped network: NL > CC > GL > OL > LC. - Order of mean for signals with 7.5 (GPM) demand in the branched network: CC > GL > LC > NL > OL.

Table 2.71. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Mean	Transient – P1	- Comparison of NL signal's mean with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's mean with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's mean with transient demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's mean with transient demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's mean with transient demand: $GL_{lo} < GL_{br}$. - Order of mean for signals with transient demand in the looped network: NL > LC > CC > GL > OL. - Order of mean for signals with transient demand in the branched network: LC > CC > OL > GL > NL.
	All – P1	- Comparing leaks' mean magnitudes of the two networks indicates no consistent change pattern in the magnitudes when the network changes. Therefore, mean cannot identify the network change. - Due to the inconsistent order of mean of signals with different demands in both networks, mean is not capable of discerning leak types consistently in both networks.
	ND – P2	- Comparison of NL signal's mean with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's mean with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's mean with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's mean with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's mean with no demand: $GL_{lo} > GL_{br}$. - Order of mean with ND in the looped network: CC > NL > GL > OL > LC. - Order of mean with ND in the branched network: LC > CC > OL > NL > GL.
	3 (GPM) – P2	- Comparison of NL signal's mean with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's mean with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's mean with 3 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's mean with 3 (GPM) demand: $LC_{lo} > LC_{br}$.

Table 2.71. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Mean	3 (GPM) – P2	- Comparison of GL signal's mean with 3 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of mean for signals with 3 (GPM) demand in the looped network: LC > NL > OL > GL > CC. - Order of mean for signals with 3 (GPM) demand in the branched network: CC > OL > LC > NL > GL.
	7.5 (GPM) – P2	- Comparison of NL signal's mean with 7.5 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's mean with 7.5 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's mean with 7.5 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's mean with 7.5 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's mean with 7.5 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of mean for signals with 7.5 (GPM) demand in the looped network: GL > NL > CC > LC > OL. - Order of mean for signals with 7.5 (GPM) demand in the branched network: CC > OL > LC > NL > GL.
	Transient – P2	- Comparison of NL signal's mean with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's mean with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's mean with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's mean with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's mean with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$.
	All – P2	- Comparing the leaks' mean magnitudes of the two networks indicates no consistent change pattern in the magnitudes when the network changed. Therefore, mean cannot identify the network change. - Due to the inconsistent order of mean of signals with different demands in the looped network, mean of P2 sensor data cannot discern leak types in the looped network. However, when there is a demand in the branched network, mean of dynamic pressure magnitudes have the following order: CC > OL > LC > NL > GL.

2.12.5.13. Standard Deviation

Figures 2.279 and 2.280 show the standard deviation plots of the dynamic pressure data measured by sensor P1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.281 and 2.282 show the same plots but for the sensor P2 data.

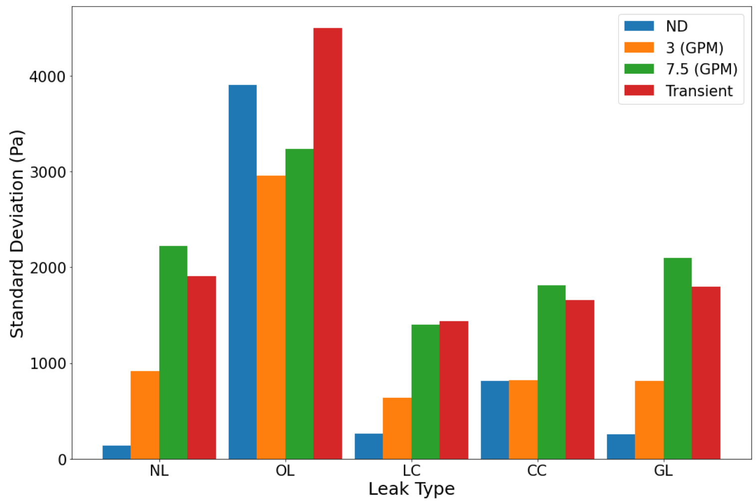

Figure 2.279. Standard deviation plot of sensor P1 data in the looped network for all leaks and demands

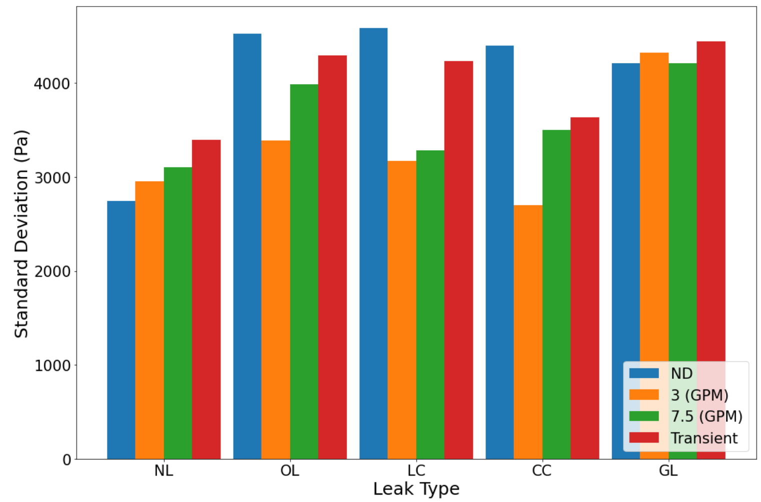

Figure 2.280. Standard deviation plot of sensor P1 data in the branched network for all leaks and demands

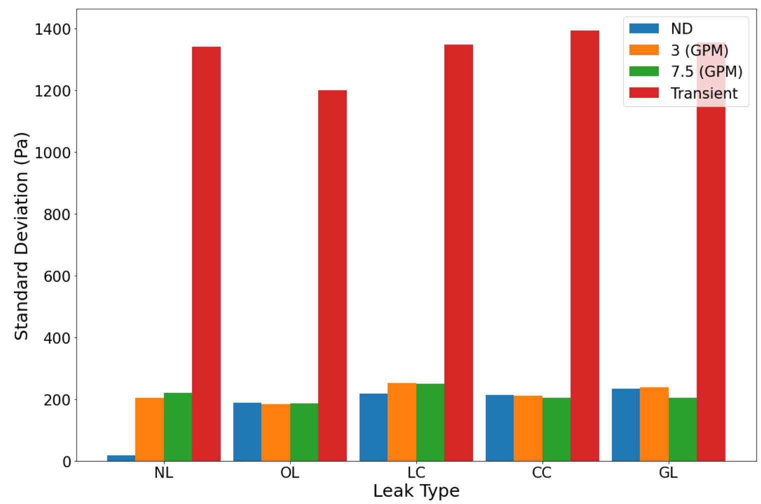

Figure 2.281. Standard deviation plot of sensor P2 data in the looped network for all leaks and demands

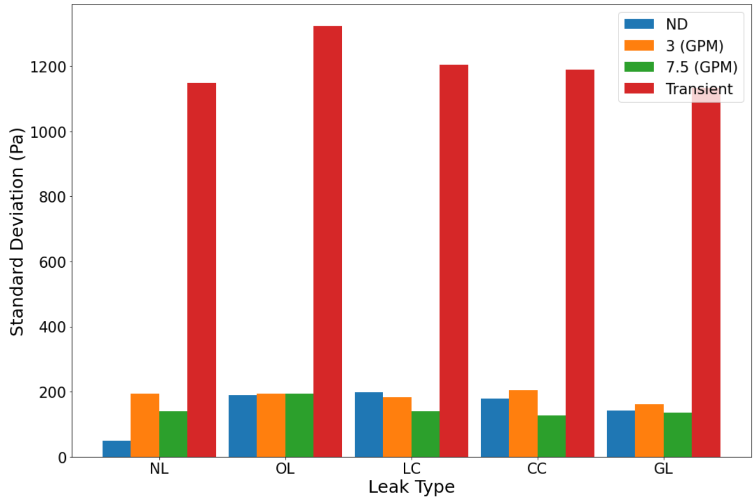

Figure 2.282. Standard deviation plot of sensor P2 data in the branched network for all leaks and demands

Table 2.72 compares standard deviation plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types by sensors P1 and P2.

Table 2.72. Analysis of standard deviation plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Standard deviation	ND – P1	- Comparison of NL signal's standard deviation with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's standard deviation with no demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's standard deviation with no demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's standard deviation with no demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's standard deviation with no demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of standard deviation with ND in the looped network: OL > CC > LC > GL > NL. - Order of standard deviation with ND in the branched network: LC > OL > CC > GL > NL.
	3 (GPM) – P1	- Comparison of NL signal's standard deviation with 3 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's standard deviation entropy with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's standard deviation entropy with 3 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's standard deviation with 3 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's standard deviation with 3 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of standard deviation for signals with 3 (GPM) demand in the looped network: OL > NL > CC > GL > LC. - Order of standard deviation for signals with 3 (GPM) demand in the branched network: GL > OL > LC > NL > CC.
	7.5 (GPM) – P1	- Comparison of NL signal's standard deviation with 7.5 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's standard deviation with 7.5 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's standard deviation with 7.5 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's standard deviation with 7.5 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's standard deviation with 7.5 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of standard deviation for signals with 7.5 (GPM) demand in the looped network: OL > NL > GL > CC > LC. - Order of standard deviation for signals with 7.5 (GPM) demand in the branched network: GL > OL > CC > LC > NL.

Table 2.72. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Standard deviation	Transient – P1	- Comparison of NL signal's standard deviation with transient demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's standard deviation with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's standard deviation with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's standard deviation with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's standard deviation with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of standard deviation for signals with transient demand in the looped network: OL > NL > GL > CC > LC. - Order of standard deviation for signals with transient demand in the branched network: GL > OL > LC > CC > NL.
	All – P1	- Except for OL with transient demand, the standard deviation of all other leak and no leak signals are larger in the branched network than those of the looped network. Therefore, the standard deviation of sensor P1 data may be used for identifying the network change. - Based on P1 and with all demand variations, OL has the largest standard deviation in the looped network. - Based on P1 and when there is a non-zero demand, GL and OL have the largest and the second largest standard deviation in the branched network. - Due to the inconsistent order of standard deviation for signals with different demands in both networks, standard deviation cannot discern leak types consistently in both networks.
	ND – P2	- Comparison of NL signal's standard deviation with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's standard deviation with no demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's standard deviation with no demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's standard deviation with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's standard deviation with no demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of standard deviation with ND in the looped network: GL > LC > CC > OL > NL. - Order of standard deviation with ND in the branched network: LC > OL > CC > GL > NL.

Table 2.72. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Standard deviation	3 (GPM) – P2	- Comparison of NL signal's standard deviation with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's standard deviation with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's standard deviation with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's standard deviation with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of standard deviation for signals with 3 (GPM) demand in the looped network: LC > GL > CC > NL > OL. - Order of standard deviation for signals with 3 (GPM) demand in the branched network: CC > NL > OL > LC > GL.
	7.5 (GPM) – P2	- Comparison of NL signal's standard deviation with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's standard deviation with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's standard deviation with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's standard deviation with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of standard deviation for signals with 7.5 (GPM) demand in the looped network: LC > NL > LC > GL > OL. - Order of standard deviation for signals with 7.5 (GPM) demand in the branched network: OL > LC > NL > GL > CC.
	Transient – P2	- Comparison of NL signal's standard deviation with transient demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's standard deviation with transient demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's standard deviation with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's standard deviation with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's standard deviation with transient demand: $GL_{lo} > GL_{br}$. - Order of standard deviation for signals with transient demand in the looped network: CC > GL > LC > NL > OL. - Order of standard deviation for signals with transient demand in the branched network: OL > LC > CC > NL > GL.

Table 2.72. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Standard deviation	All – P2	- Based on P2 data and with a non-zero demand, except for OL, the standard deviation of leak and no leak signals are larger in the looped network than those in the branched network. Therefore, the standard deviation of sensor P2 data may be used for identifying the network change. - Based on P2 data, when there is a non-zero demand, OL has the smallest standard deviation in the looped network. However, based on P1 data, OL had the largest standard deviation in the looped network. - Due to the inconsistent order of standard deviation for signals with different demands in both networks, standard deviation is not capable of discerning leak types consistently in both networks.

2.12.5.14. Zero-crossing Rate

Figures 2.283 and 2.284 show the zero-crossing rate plots of the dynamic pressure data measured by sensor P1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.285 and 2.286 show the same plots but for sensor P2 data.

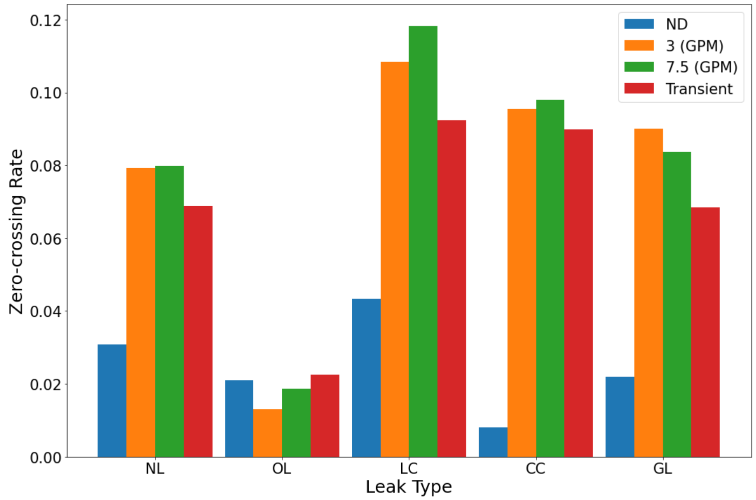

Figure 2.283. Zero-crossing rate plot of sensor P1 data in the looped network for all leaks and demands

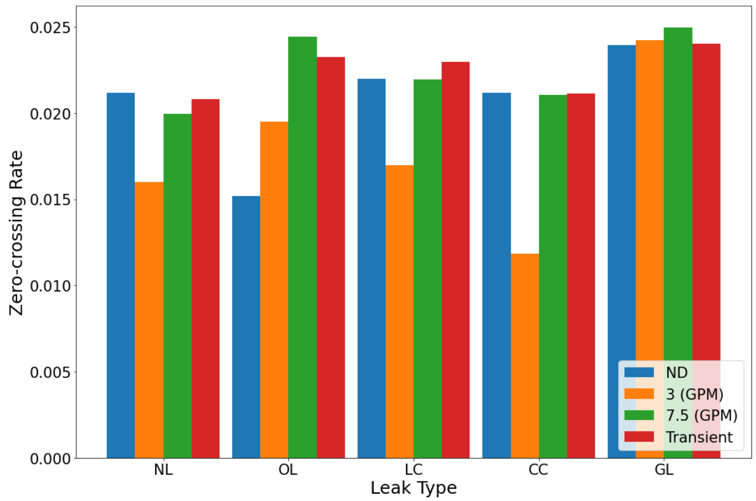

Figure 2.284. Zero-crossing rate plot of sensor P1 data in the branched network for all leaks and demands

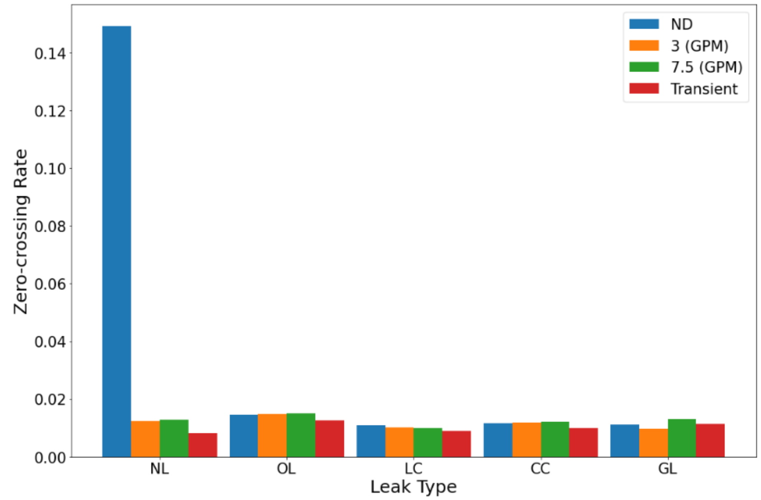

Figure 2.285. Zero-crossing rate plot of sensor P2 data in the looped network for all leaks and demands

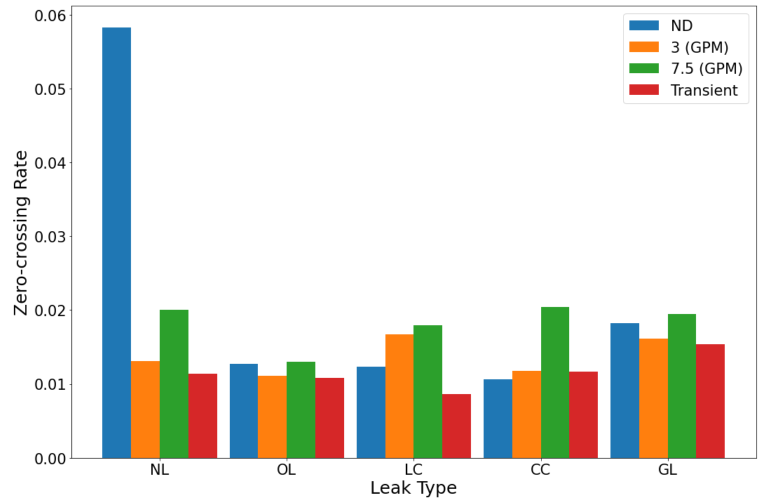

Figure 2.286. Zero-crossing rate plot of sensor P2 data in the branched network for all leaks and demands

Table 2.73 compares zero-crossing rate plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types by sensors P1 and P2.

Table 2.73. Analysis of zero-crossing rate plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Zero-crossing rate	ND – P1	- Comparison of NL signal's zero-crossing rate with no demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's zero-crossing rate with no demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's zero-crossing rate with no demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's zero-crossing rate with no demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's zero-crossing rate with no demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of zero-crossing rate with ND in the looped network: LC > NL > GL > OL > CC. - Order of zero-crossing rate with ND in the branched network: GL > LC > CC > NL > OL.
	3 (GPM) – P1	- Comparison of NL signal's zero-crossing rate with 3 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's zero-crossing rate with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's zero-crossing rate with 3 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's zero-crossing rate with 3 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's zero-crossing rate with 3 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of zero-crossing rate for signals with 3 (GPM) demand in the looped network: LC > CC > GL > NL > OL. - Order of zero-crossing rate for signals with 3 (GPM) demand in the branched network: GL > OL > LC > NL > CC.

Table 2.73. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Zero-crossing rate	7.5 (GPM) – P1	- Comparison of NL signal's zero-crossing rate with 7.5 (GPM) demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's zero-crossing rate with 7.5 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's zero-crossing rate with 7.5 (GPM) demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's zero-crossing rate with 7.5 (GPM) demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's zero-crossing rate with 7.5 (GPM) demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the looped network: LC > CC > GL > NL > OL. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the branched network: GL > OL > LC > CC > NL.
	Transient – P1	- Comparison of NL signal's zero-crossing rate with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's zero-crossing rate with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's zero-crossing rate with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's zero-crossing rate with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's zero-crossing rate with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of zero-crossing rate for signals with transient demand in the looped network: LC > CC > NL > GL > OL. - Order of zero-crossing rate for signals with transient demand in the branched network: GL > OL > LC > CC > NL.
	All – P1	- In the looped network and with all demand variations, LC has the highest zero-crossing rate. - In the branched network and with all demand variations, GL has the highest zero-crossing rate. - Based on P1 data and with a non-zero demand, except for OL, the zero-crossing rate of leak and no leak signals are larger in the looped network than those in the branched network. Therefore, the zero-crossing rate of sensor P1 data may be used for identifying the network change. This pattern was observed in the standard deviation of sensor P2 data. - Due to the inconsistent order of zero-crossing rate of signals with different demands in both networks, zero-crossing rate cannot discern leak types consistently in both networks.

Table 2.73. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Zero-crossing rate	ND – P2	- Comparison of NL signal's zero-crossing rate with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's zero-crossing rate with no demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's zero-crossing rate with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's zero-crossing rate with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's zero-crossing rate with no demand: $GL_{lo} < GL_{br}$. - Order of zero-crossing rate with ND in the looped network: NL > OL > CC > GL > LC. - Order of zero-crossing rate with ND in the branched network: NL > GL > OL > LC > CC.
	3 (GPM) – P2	- Comparison of NL signal's zero-crossing rate with 3 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's zero-crossing rate with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's zero-crossing rate with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's zero-crossing rate with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's zero-crossing rate with 3 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of zero-crossing rate for signals with 3 (GPM) demand in the looped network: OL > NL > CC > LC > GL. - Order of zero-crossing rate for signals with 3 (GPM) demand in the branched network: LC > GL > NL > CC > OL.
	7.5 (GPM) – P2	- Comparison of NL signal's zero-crossing rate with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's zero-crossing rate with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's zero-crossing rate with 7.5 (GPM) demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's zero-crossing rate with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's zero-crossing rate with 7.5 (GPM) demand: $GL_{lo} < GL_{br}$. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the looped network: OL > GL > NL > CC > LC. - Order of zero-crossing rate for signals with 7.5 (GPM) demand in the branched network: CC > NL > GL > LC > OL.

Table 2.73. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Zero-crossing rate	Transient – P2	- Comparison of NL signal's zero-crossing rate with transient demand: $\mathtt{NL_{lo}} < \mathtt{NL_{br}}$. - Comparison of OL signal's zero-crossing rate with transient demand: $\mathtt{OL_{lo}} > \mathtt{OL_{br}}$. - Comparison of CC signal's zero-crossing rate with transient demand: $\mathtt{CC_{lo}} < \mathtt{CC_{br}}$. - Comparison of LC signal's zero-crossing rate with transient demand: $\mathtt{LC_{lo}} > \mathtt{LC_{br}}$. - Comparison of GL signal's zero-crossing rate with transient demand: $\mathtt{GL_{lo}} < \mathtt{GL_{br}}$. - Order of zero-crossing rate for signals with transient demand in the looped network: OL > GL > CC > LC > NL. - Order of zero-crossing rate for signals with transient demand in the branched network: GL > CC > NL > OL > LC.
	All – P2	- Comparing leaks' zero-crossing rate magnitudes of the two networks indicates no consistent change pattern in the magnitudes when the networks change. Therefore, zero-crossing rate cannot identify the network change. - Due to the inconsistent order of zero-crossing rate for signals with different demands in both networks, zero-crossing rate cannot discern leak types consistently in both networks.

2.12.5.15. RMS

Figures 2.287 and 2.288 show the RMS plots of the dynamic pressure data measured by sensor P1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.289 and 2.290 show the same plots but for sensor P2 data.

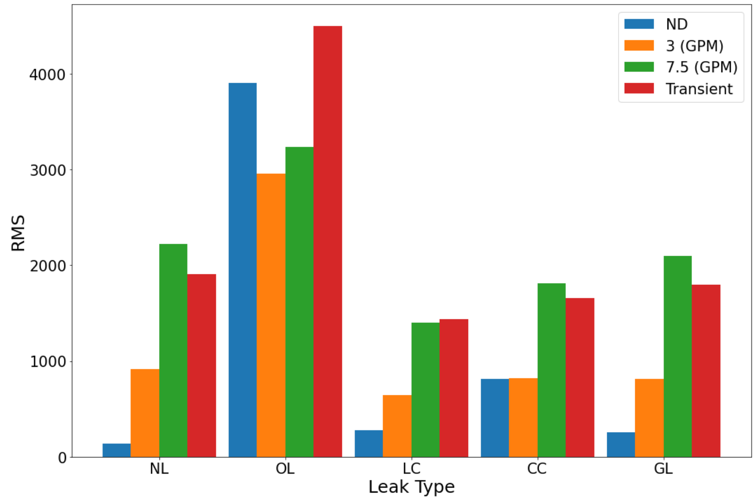

Figure 2.287. RMS plot of sensor P1 data in the looped network for all leaks and demands

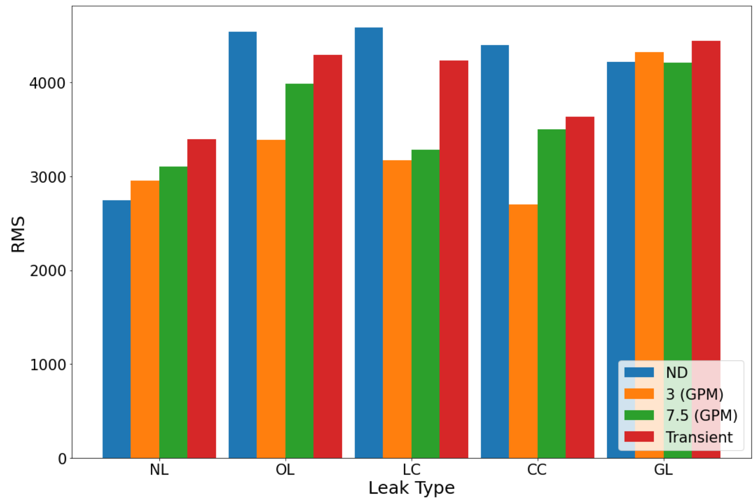

Figure 2.288. RMS plot of sensor P1 data in the branched network for all leaks and demands

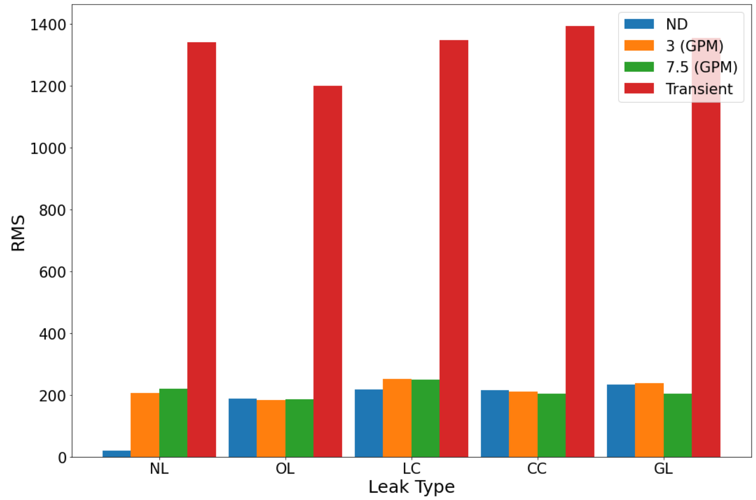

Figure 2.289. RMS plot of sensor P2 data in the looped network for all leaks and demands

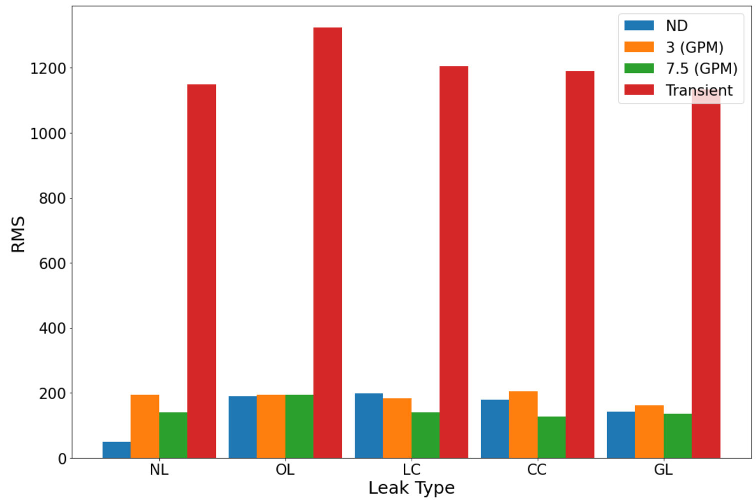

Figure 2.290. RMS plot of sensor P2 data in the branched network for all leaks and demands

Table 2.74 compares RMS plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types by sensors P1 and P2.

Table 2.74. Analysis of RMS plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
RMS	ND – P1	- Comparison of NL signal's RMS with no demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's RMS with no demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's RMS with no demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's RMS with no demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's RMS with no demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of RMS with ND in the looped network: OL > CC > LC > GL > NL. - Order of RMS with ND in the branched network: LC > OL > CC > GL > NL.
	3 (GPM) – P1	- Comparison of NL signal's RMS with 3 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's RMS with 3 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's RMS with 3 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's RMS with 3 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's RMS with 3 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of RMS for signals with 3 (GPM) demand in the looped network: OL > NL > CC > GL > LC. - Order of RMS for signals with 3 (GPM) demand in the branched network: GL > OL > LC > NL > CC.

Table 2.74. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
RMS	7.5 (GPM) – P1	- Comparison of NL signal's RMS with 7.5 (GPM) demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's RMS with 7.5 (GPM) demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's RMS with 7.5 (GPM) demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's RMS with 7.5 (GPM) demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's RMS with 7.5 (GPM) demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of RMS for signals with 7.5 (GPM) demand in the looped network: OL > NL > GL > CC > LC. - Order of RMS for signals with 7.5 (GPM) demand in the branched network: GL > OL > CC > LC > NL.
	Transient – P1	- Comparison of NL signal's RMS with transient demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's RMS with transient demand: $\mathrm{OL_{lo}} > \mathrm{OL_{br}}$. - Comparison of CC signal's RMS with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's RMS with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's RMS with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of RMS for signals with transient demand in the looped network: OL > NL > GL > CC > LC. - Order of RMS for signals with transient demand in the branched network: GL > OL > LC > CC > NL.
	All – P1	- Based on P1 data and when there is a demand, GL has the largest RMS magnitude compared to other signals in the branched network. - Based on P1 data and with all demand variants, OL has the largest RMS magnitude compared to other signals in the looped network. - Except for OL with the transient demand, the RMS of all other leak and no leak signals in the branched network are larger than those in the looped network. Therefore, by using P1 data, RMS may discern the network change. - Due to the inconsistent order of RMS for signals with different demands in both networks, RMS cannot discern leak types in both networks.

Table 2.74. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
RMS	ND – P2	- Comparison of NL signal's RMS with no demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's RMS with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's RMS with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's RMS with no demand: $GL_{lo} > GL_{br}$. - Order of RMS with ND in the looped network: GL > LC > CC > OL > NL. - Order of RMS with ND in the branched network: LC > OL > CC > GL > NL.
	3 (GPM) – P2	- Comparison of NL signal's RMS with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's RMS with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's RMS with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's RMS with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of RMS for signals with 3 (GPM) demand in the looped network: LC > GL > CC > NL > OL. - Order of RMS for signals with 3 (GPM) demand in the branched network: CC > OL > NL > LC > GL.
	7.5 (GPM) – P2	- Comparison of NL signal's RMS with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's RMS with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's RMS with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's RMS with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's RMS with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of RMS for signals with 7.5 (GPM) demand in the looped network: LC > NL > CC > GL > OL. - Order of RMS for signals with 7.5 (GPM) demand in the branched network: OL > LC > NL > GL > CC.

Table 2.74. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
RMS	Transient – P2	- Comparison of NL signal's RMS with transient demand: $\mathrm{NL_{lo}} > \mathrm{NL_{br}}$. - Comparison of OL signal's RMS with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's RMS with transient demand: $\mathrm{CC_{lo}} > \mathrm{CC_{br}}$. - Comparison of LC signal's RMS with transient demand: $\mathrm{LC_{lo}} > \mathrm{LC_{br}}$. - Comparison of GL signal's RMS with transient demand: $\mathrm{GL_{lo}} > \mathrm{GL_{br}}$. - Order of RMS for signals with transient demand in the looped network: CC > GL > LC > NL > OL. - Order of RMS for signals with transient demand in the branched network: OL > LC > CC > NL > GL.
	All – P2	- Except for OL with non-zero demands, the RMS of all other leak and no leak signals in the looped network are larger than those in the branched network. Therefore, by using P2 data, RMS may discern the network change. - Due to the inconsistent order of RMS for signals with different demands in both networks, RMS is not capable of discerning leak types consistently in both networks.

2.12.5.16. Crest Factor

Figures 2.291 and 2.292 show the crest factor plots of the dynamic pressure data measured by sensor P1 in the looped and branched networks, respectively, for all leak and demand variants. Figures 2.293 and 2.294 show the same plots but for sensor P2 data.

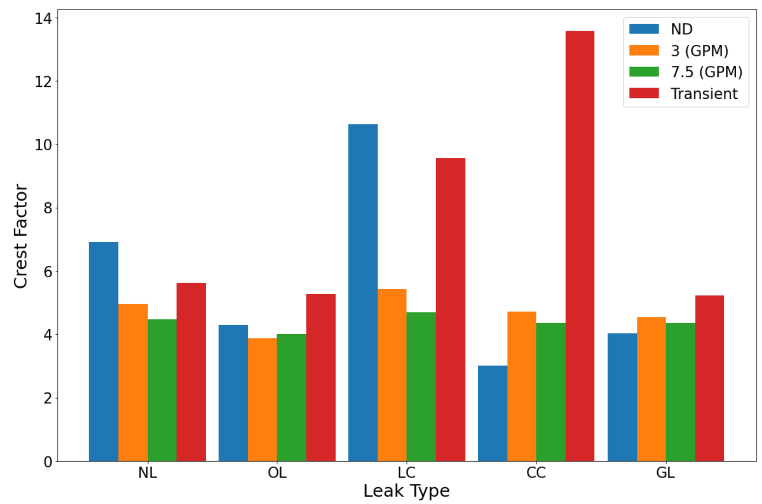

Figure 2.291. Crest factor plot of sensor P1 data in the looped network for all leaks and demands

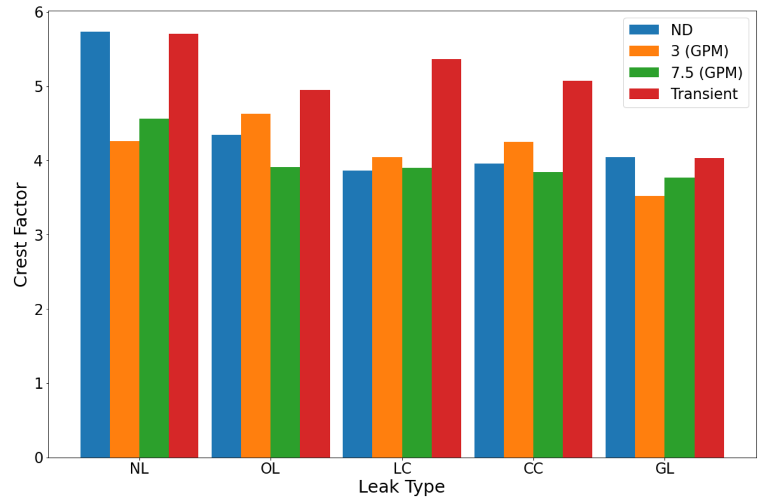

Figure 2.292. Crest factor plot of sensor P1 data in the branched network for all leaks and demands

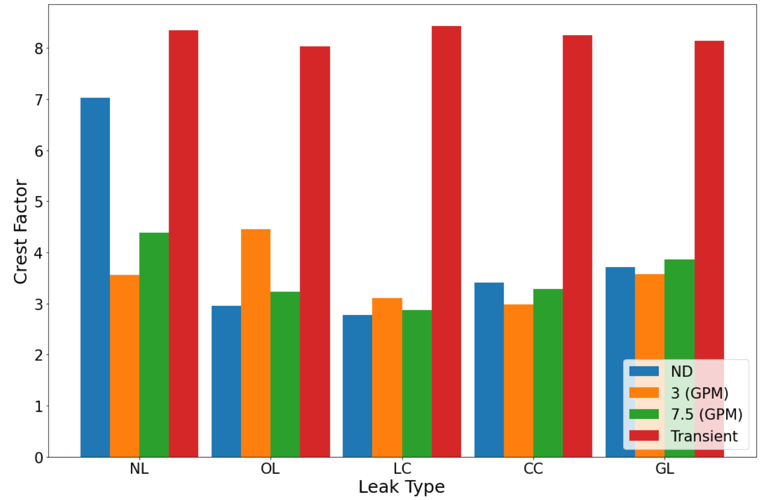

Figure 2.293. Crest factor plot of sensor P2 data in the looped network for all leaks and demands

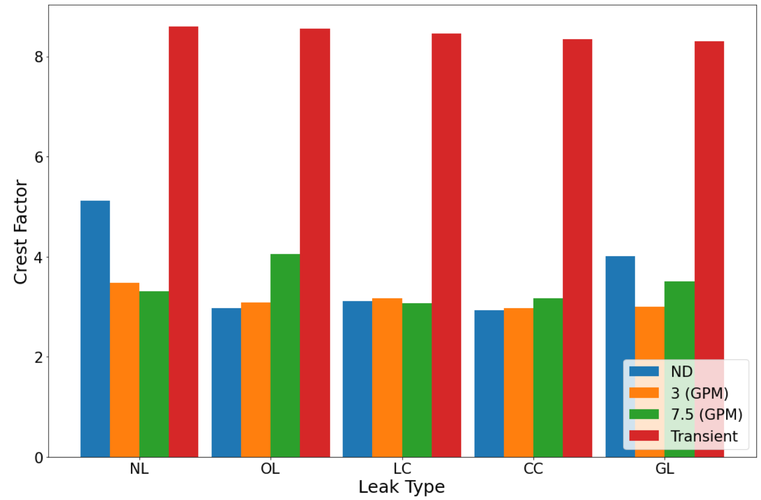

Figure 2.294. Crest factor plot of sensor P2 data in the branched network for all leaks and demands

Table 2.75 compares crest factor plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types by sensors P1 and P2.

Table 2.75. Analysis of crest factor plots of dynamic pressure data measured in the looped and branched networks with all leak and demand types

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Crest factor	ND – P1	- Comparison of NL signal's crest factor with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's crest factor with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's crest factor with no demand: $CC_{lo} < CC_{br}$. - Comparison of LC signal's crest factor with no demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's crest factor with no demand: $GL_{lo} < GL_{br}$. - Order of crest factor with ND in the looped network: LC > NL > OL > GL > CC. - Order of crest factor with ND in the branched network: NL > OL > GL > CC > LC.
	3 (GPM) – P1	- Comparison of NL signal's crest factor with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's crest factor with 3 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's crest factor with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's crest factor with 3 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's crest factor with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of crest factor for signals with 3 (GPM) demand in the looped network: LC > NL > CC > GL > OL. - Order of crest factor for signals with 3 (GPM) demand in the branched network: OL > NL > CC > LC > GL.

Table 2.75. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Crest factor	7.5 (GPM) – P1	- Comparison of NL signal's crest factor with 7.5 (GPM) demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's crest factor with 7.5 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's crest factor with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's crest factor with 7.5 (GPM) demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's crest factor with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of crest factor for signals with 7.5 (GPM) demand in the looped network: LC > NL > GL > CC > OL. - Order of crest factor for signals with 7.5 (GPM) demand in the branched network: NL > OL > LC > CC > GL.
	Transient – P1	- Comparison of NL signal's crest factor with transient demand: $NL_{lo} < NL_{br}$. - Comparison of OL signal's crest factor with transient demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's crest factor with transient demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's crest factor with transient demand: $LC_{lo} > LC_{br}$. - Comparison of GL signal's crest factor with transient demand: $GL_{lo} > GL_{br}$. - Order of crest factor for signals with transient demand in the looped network: CC > LC > NL > OL > GL. - Order of crest factor for signals with transient demand in the branched network: NL > LC > CC > OL > GL.
	All – P1	- Comparing leaks' crest factor magnitudes of the two networks indicates no consistent change pattern in the magnitudes when the network changed. Therefore, crest factor cannot identify the network change. - Since the order of signals' crest factor with different demands in both networks are inconsistent, crest factor cannot discern leak types consistently in both networks.

Table 2.75. Continued

Evaluation Criterion	**Evaluation Sub-criterion**	**Looped vs. Branched Network**
Crest factor	ND – P2	- Comparison of NL signal's crest factor with no demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's crest factor with no demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's crest factor with no demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's crest factor with no demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's crest factor with no demand: $GL_{lo} < GL_{br}$. - Order of crest factor with ND in the looped network: NL > GL > CC > OL > LC. - Order of crest factor with ND in the branched network: NL > GL > LC > OL > CC.
	3 (GPM) – P2	- Comparison of NL signal's crest factor with 3 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's crest factor with 3 (GPM) demand: $OL_{lo} > OL_{br}$. - Comparison of CC signal's crest factor with 3 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's crest factor with 3 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's crest factor with 3 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of crest factor for signals with 3 (GPM) demand in the looped network: OL > GL > OL > LC > CC. - Order of crest factor for signals with 3 (GPM) demand in the branched network: NL > LC > OL > GL > CC.
	7.5 (GPM) – P2	- Comparison of NL signal's crest factor with 7.5 (GPM) demand: $NL_{lo} > NL_{br}$. - Comparison of OL signal's crest factor with 7.5 (GPM) demand: $OL_{lo} < OL_{br}$. - Comparison of CC signal's crest factor with 7.5 (GPM) demand: $CC_{lo} > CC_{br}$. - Comparison of LC signal's crest factor with 7.5 (GPM) demand: $LC_{lo} < LC_{br}$. - Comparison of GL signal's crest factor with 7.5 (GPM) demand: $GL_{lo} > GL_{br}$. - Order of crest factor for signals with 7.5 (GPM) demand in the looped network: NL > GL > CC > OL > LC. - Order of crest factor for signals with 7.5 (GPM) demand in the branched network: OL > GL > NL > CC > LC.

Table 2.75. Continued

Evaluation Criterion	Evaluation Sub-criterion	Looped vs. Branched Network
Crest factor	Transient – P2	- Comparison of NL signal's crest factor with transient demand: $\mathrm{NL_{lo}} < \mathrm{NL_{br}}$. - Comparison of OL signal's crest factor with transient demand: $\mathrm{OL_{lo}} < \mathrm{OL_{br}}$. - Comparison of CC signal's crest factor with transient demand: $\mathrm{CC_{lo}} < \mathrm{CC_{br}}$. - Comparison of LC signal's crest factor with transient demand: $\mathrm{LC_{lo}} < \mathrm{LC_{br}}$. - Comparison of GL signal's crest factor with transient demand: $\mathrm{GL_{lo}} < \mathrm{GL_{br}}$. - Order of crest factor for signals with transient demand in the looped network: LC > NL > CC > GL > OL. - Order of crest factor for signals with transient demand in the branched network: NL > OL > LC > CC > GL.
	All – P2	- Comparing leaks' crest factor magnitudes of two networks indicates no consistent change pattern in the magnitudes when the networks change. Therefore, crest factor cannot identify the network change. - Due to the inconsistent order of crest factor for signals with different demands in both networks, crest factor is not capable of discerning leak types consistently in both networks.

2.12.5.17. Summary of Dynamic Pressure Sensor Measurement Evaluations

The following are some important takeaways from the numerical and visual evaluations of the dynamic pressure data. Figures 2.295 and 2.296 show the dynamic pressure plots of OL and no leak data with 7.5 (GPM) demand measured by sensor P1 at the looped and branched networks, respectively.

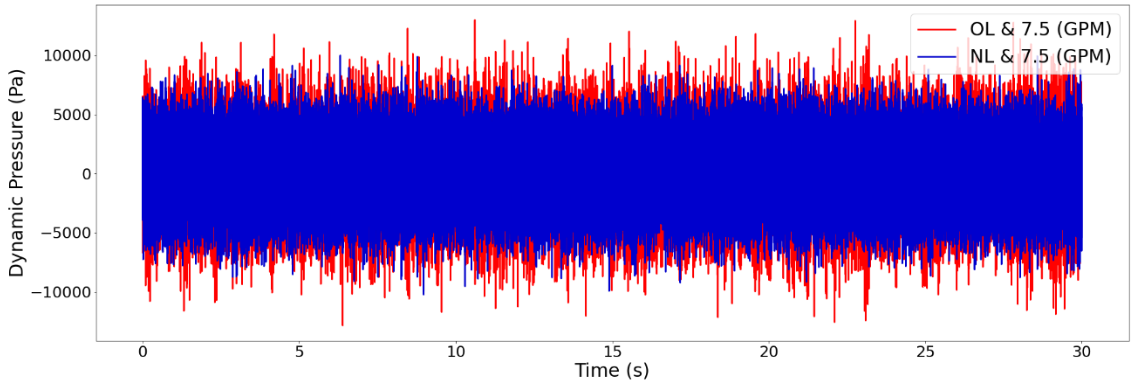

Figure 2.295. Time-domain plots of OL vs. NL signals with 7.5 (GPM) demand in the looped network measured by sensor P1

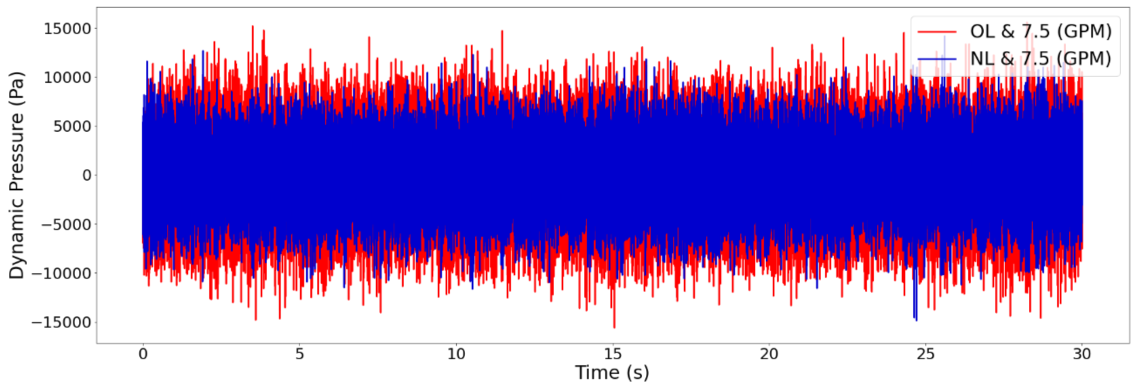

Figure 2.296. Time-domain plots of OL vs. NL signals with 7.5 (GPM) demand in the branched network measured by sensor P1

Figures 2.297 and 2.298 show the dynamic pressure plots of OL and no leak data with transient demand measured by sensor P1 at the looped and branched networks, respectively.

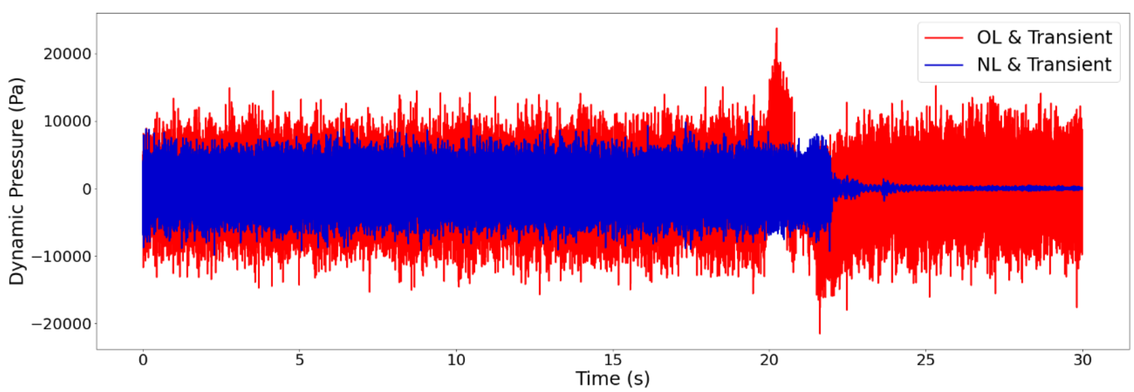

Figure 2.297. Time-domain plots of OL vs. NL signals with transient demand in the looped network measured by sensor P1

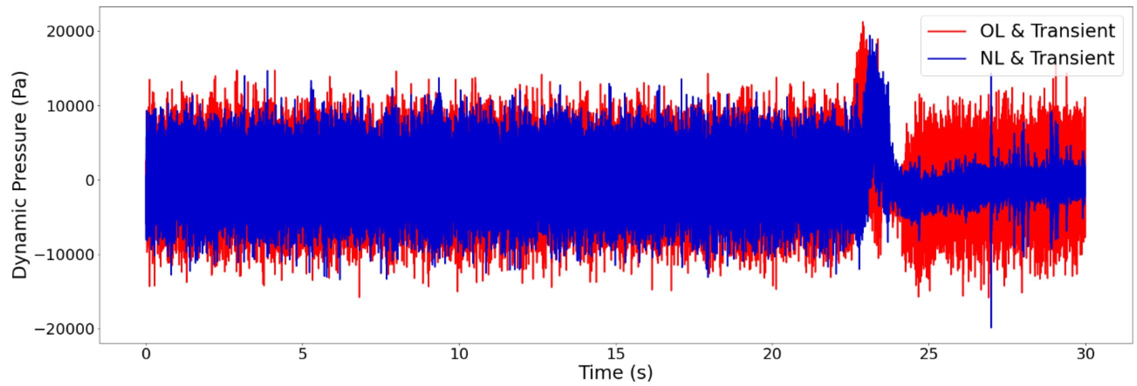

Figure 2.298. Time-domain plots of OL vs. NL signals with transient demand in the branched network measured by sensor P1

Figures 2.299 and 2.300 show the dynamic pressure plots of OL and no leak data with 7.5 (GPM) demand measured by sensor P2 at the looped and branched networks, respectively.

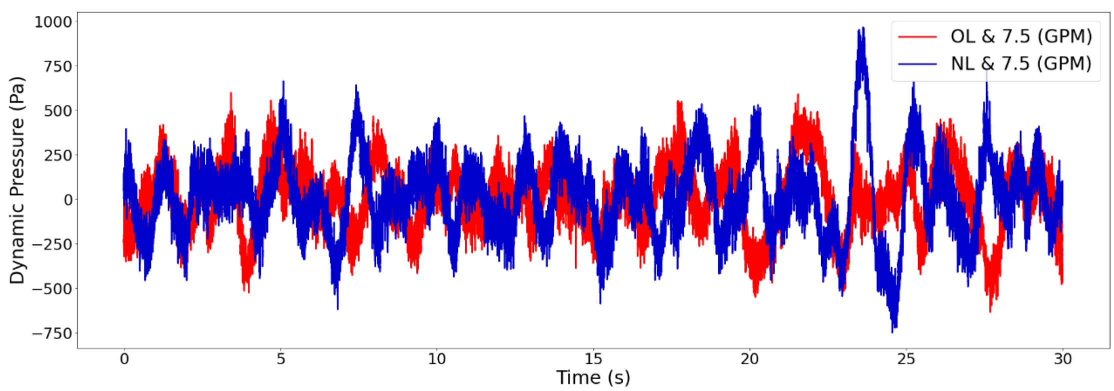

Figure 2.299. Time-domain plots of OL vs. NL signals with 7.5 (GPM) demand in the looped network measured by sensor P2

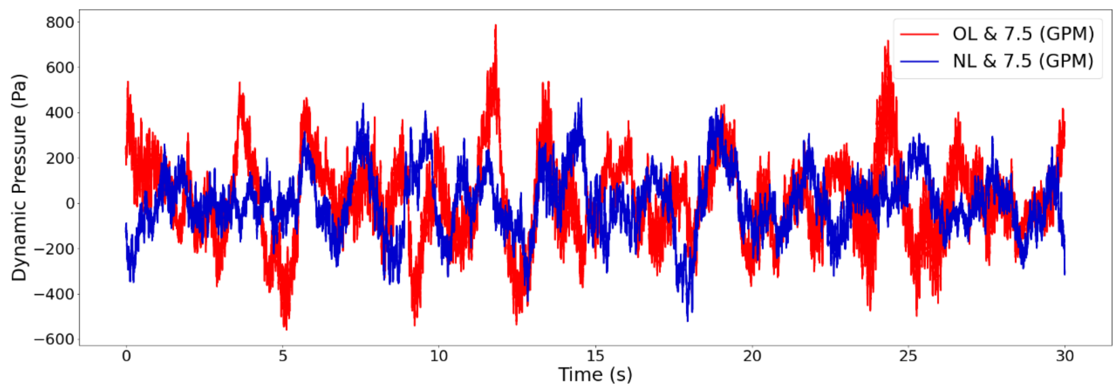

Figure 2.300. Time-domain plots of OL vs. NL signals with 7.5 (GPM) demand in the branched network measured by sensor P2

Figures 2.301 and 2.302 show the dynamic pressure plots of OL and no leak data with transient demand measured by sensor P2 at the looped and branched networks, respectively.

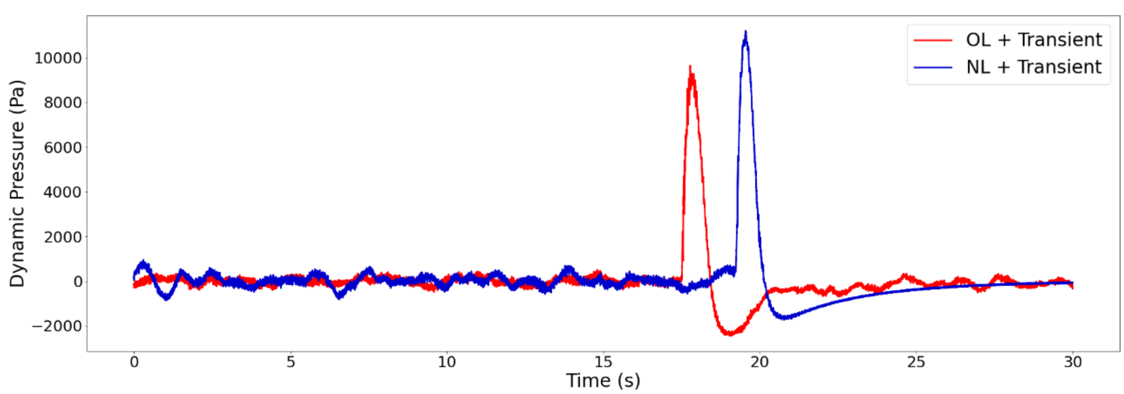

Figure 2.301. Time-domain plots of OL vs. NL signals with transient demand in the looped network measured by sensor P2

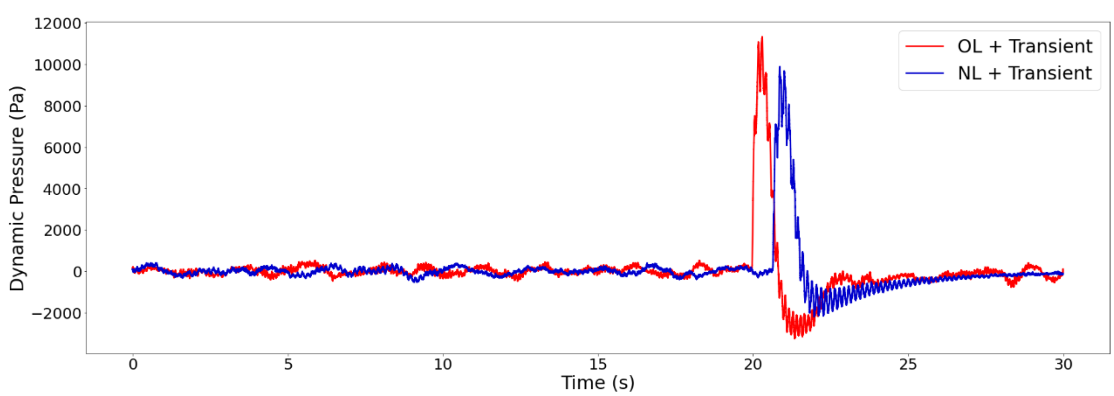

Figure 2.302. Time-domain plots of OL vs. NL signals with transient demand in the branched network measured by sensor P2

Evaluation of sensor P1 plots with zero demand indicates that all signals have larger amplitudes in the branched network than in the looped network. The time-domain plots of no leak and OL signals with 7.5 (GPM) and transient demands also follow this pattern, see Figures 2.295 to 2.298. Therefore, the absolute value of signals may help identify the network change. Moreover, based on P1 plots with zero demand, all leak amplitudes are larger than those of no leak in both networks. However, regarding Figure 2.295 where there is 7.5 (GPM) demand, since some no leak amplitudes are larger than or equal to the OL's amplitudes, we could not conclude the larger leak amplitudes.

Visualization of sensor P2 data with zero demand indicates that leak signals are more cyclic and have larger absolute values than the no leak ones. In Figures 2.301 and 302 which show OL and NL time series with transient demand, there are dynamic pressure spikes at about second 20. These spikes or the wave shapes of dynamic pressure are due to the quick service valve shut-off, which acted as a transient source in the pipeline. These wave shapes or cycles are also evident in the P2 sensor’s leak signals with no demand. Therefore, one can conclude that leaks have transient effects on pipes' dynamic pressure and represent their signatures with cyclic patterns in time series plots. However, these cyclic patterns are not only due to the leaks. Regarding Figures 2.299 and 2.300 that show the OL and NL signals with 7.5 (GPM) demand in both networks, not only do leak signals include cycles, but no leak data also have a cyclic pattern. Therefore, any water output, including leak outflow and demand flow, can cause transient effects on dynamic water pressure.

Comparing the time-domain plots of the leak and no leak data with and without a demand measured by P1 and P2 indicate the following: (1) data of sensor P1 have larger amplitudes than those of sensor P2; (2) sensor P2 data are more cyclic than those of sensor P1. Dynamic pressure measures a fluid's kinetic energy whose quantity is defined by:

$$q = \frac{1}{2}\rho u^2 \tag{2.34}$$

where q is dynamic pressure (Pa), ρ is fluid mass density ($kg.m^{-1}.s^{-2}$), and u is flow speed (m/s). P1 is mounted on a pipe with a cross-section area of one-sixth of that of P2's pipe. Therefore, dynamic pressure at P2 can be thirty-six times smaller than that at P1. In addition to the leak and service valve that cause transient effects on dynamic pressure, tees and crosses also have transient influences in the pipes. Since these transient sources are more available in pipes between the pump and P2 than P1 and the pump, P2 data are more cyclic and include more transients.

Based on the more variability of leak dynamic pressure data in the plots with no demand, we investigated if the standard deviations of leak data are larger than those of no leak data. However, based on Figures 2.279 to 2.282 that are standard deviation plots of the leak and no leak dynamic pressure data with non-zero demands, there are some cases where the standard deviations of no leak signals are larger than those of leak signals.

Similar to accelerometers and hydrophones, dynamic pressure sensor frequency contents depend on sensor locations. Based on P1 and no demand data, non-zero frequencies of NL and leak signals are below 600 Hz in the looped network and less than 800 Hz in the branched network. Due to the similarity of leak maximum frequencies in both networks, there is no relation between leak flows and their frequency caps. Based on P2 data, NL signals' maximum frequency in both networks is 100 Hz, and leak signals' maximum frequency in both networks is 400 Hz. A reason for lower maximum frequencies of the leak and no leak signals at P2 than those at P1 is the junctions between two sensors. The tees and crosses in the pipeline act as energy dissipators which dampen waves with higher frequencies and cause waves with lower frequencies to reach P2 that is more distant to the pump. When there is no demand, a similar pattern is present in the data of both dynamic pressure sensors and both networks that is the larger amplitudes of leak signals' dominant frequencies than those of NL signals. Therefore, the amplitude of the dominant frequency can be used as a feature to detect leaks, though more evaluation on data with non-zero demands is necessary.

The dynamic pressure cumulative distribution plots of leaks depend on the sensor locations. For example, based on P1, leaks' cumulative distribution plots show larger magnitudes in the branched network than those of the looped one. However, based on P2 cumulative distribution plots, leak dynamic pressure data magnitudes are larger in the looped network than those of the branched network. When there is no leak and no demand flow, the fluid velocity is the least, and subsequent dynamic pressure is minimal. This is why NL has the smallest magnitudes regarding the order of signal magnitudes in both networks using both dynamic pressure sensors.

Box plots of dynamic pressure data with no demand indicate that the order of signal magnitude continuum, calculated as the difference of the 3rd and 1st quartiles of the data, varies based on the sensor location. For example, the largest leak magnitude range in the looped network is for OL, based on P1, and LC, based on P2 measurements. Moreover, the relation of dynamic pressure magnitudes of the looped network and those of the branched network depends on the sensor location. Though P1 data represent larger leak and no leak dynamic pressure magnitudes in the branched network, there is no specific relation between those networks' magnitudes given P2 measurements. However, regarding both sensors' data, the networks' box plots suggest that the dynamic pressure magnitude ranges of leaks are larger than those of no leak signals. This is due to larger water velocities when leaks are present, while water velocity is about zero when there is no leak and no demand.

There is no constant pattern in the relations of the two networks' areas under the CSD plots of dynamic pressure data with no demand. Therefore, the areas under the CSD plots cannot identify the change in the network architecture. Unlike accelerometer and hydrophone data, NL signals have the smallest areas under the CSD plots of dynamic pressure data with no demand in both networks representing the least covariance between P1 and P2 data for NL in both networks. Comparing the blue plots in Figures 2.231 and 2.239 or those in Figures 2.232 and 2.240 indicates how dissimilar the frequency contents of NL signals are at sensors P1 and P2. On the other hand, due to the similar patterns of CC signals' frequency contents at sensors P1 and P2 in the looped network, particularly at frequencies lower than 200 Hz, see Figures 2.229 and 2.237, CC has the largest area under the CSD plot for dynamic pressure data in the looped network. The resemblance of LC signals' frequency contents at sensors P1 and P2 in the branched network, see Figures 2.228 and 2.236, cause the larger area under the CSD plot of dynamic pressure data in the branched network.

There were no specific patterns or thresholds in the LDIs of the looped and branched networks to discern leak types and network architectures using dynamic pressure sensor data. In both networks, the LDIs of CC and LC are the same, which shows these leaks' similar signatures in their cross-spectral density functions. When demand increased from 3 (GPM) to 7.5 (GPM), the LDI of leaks elevated. The similar LDI and demand patterns indicate that with the increase of demand, the difference between leaks' and no leaks' dynamic pressure data similarities at sensors get larger. Therefore, we can expect larger LDI magnitudes at zones with higher demands.

Comparing the sum of magnitudes in the leak:noleak plots of two networks with ND indicates that the leak:noleak ratios are more significant in the looped network than the branched network based on P1 and P2 sensors. Therefore, the sum of magnitudes in the leak:noleak plots can identify the network change when there is no demand. Despite the experiments with hydrophones, there is no relation between leak flow rates and the leak:noleak plots based on dynamic pressure data.

Evaluation of spectral centroids implies that spectral centroid can help discern leak types in both networks based on P1 dynamic pressure data. However, the order of leaks' spectral centroids differs based on network architecture. It is worth noting that the spectral centroid is not able to distinguish leak and no signals. Moreover, based on P2 data, when there is a demand, leak, and no leak signals' spectral centroid magnitudes of the branched network are larger than those of the looped network. Therefore, the spectral centroid can identify the network change. With a non-zero demand in the looped network, OL has the largest spectral centroid representing the signal's larger amplitude magnitudes at higher frequencies. On the other hand, OL spectral centroid in the branched network has the smallest magnitude. Since there are more dead ends and fewer connected pipes in the branched network, high-frequency contents of the OL signal's dynamic pressure waves become more attenuated. However, since pipes are more connected in the looped network, high-frequency contents of the OL signal's dynamic pressure waves propagate more significantly due to fewer obstacles in the pipeline.

Regarding the time-domain plots of dynamic pressure data, standard deviation seems a distinctive feature. However, the information it represents depends on the sensor location. For instance, based on P1, the standard deviation of the leak, except for OL, and no leak signals are larger in the branched network than those of the looped network. On the other hand, based on P2 data and with a non-zero demand, except for OL, the standard deviation of the leak and no leak signals are larger in the looped network than those in the branched network. The standard deviation of a leak type changed regarding the network architecture. For example, based on P2 data, when there is a non-zero demand, OL has the smallest standard deviation in the looped network. However, based on P1 data, OL has the largest standard deviation in the looped network. A similar pattern was observed between the zero-crossing rate and standard deviation, where based on P1, features' magnitudes are larger in the looped network than the branched network. Both of these features represent the variability of the signals.

Similar to standard deviation, RMS magnitudes depend on the sensor locations. Based on P1 data, except for OL with the transient demand, the RMS of all other leak and no leak signals in the branched network are larger than those in the looped network. This relation is contrary to that of standard deviation and zero-crossing rate where, based on P1, these metrics were larger in the looped network than the branched one. However, the RMSs of P2 data represent an opposite relation compared to the relation represented by the RMS's of P1.

2.13. Conclusion

This subsection described the testbed components, design and assembly procedures, and the experimental scenarios that resulted in two hundred and twenty-four measurements using accelerometer, hydrophone, and dynamic pressure sensors.

Evaluations of leak flows indicated that in both networks, leak flow rates decreased when demand increased, which was due to the constant head pump and pressure decrease caused by elevated demands. Due to the more connectivity and more evenly distributed water pressure in the looped network, the flow rates of orifice and crack leaks were larger in the looped network than the branched one.

All measurements were analyzed by six types of plots and ten frequency and time-domain features to evaluate (1) how the network architecture change affected leak characteristics; (2) how a change of leak type affected its signature. The plots were the signals with zero demand to focus on network architecture and leak type effects on leak signatures without demand interruptions. The numerical features were evaluated based on data with demand and noise variations.

Due to the inconsistent patterns and similar magnitudes of the plots and features, the sixteen evaluation criteria could not discern the leak types or identify the network change based on each parameter's two sensor data unanimously. However, based on the sensors' locations, some criteria could help detect leaks, discern leak types, or identify the network change.

The only feature that discerned the leak types was spectral centroids of P1 in both networks. Features that identified the network change were power spectral entropy of A2 with demand, fundamental frequency of H1 with leaks, zero-crossing rate of H1 with demand, absolute value of P1, spectral centroids of P1 with demand, and RMS and standard deviation of P1 and P2. These features and sensors indicated that based on the information extent they represented to differentiate leaks and network architecture, the sensors ranked as following: (1) dynamic pressure sensor; (2) hydrophone; (3) accelerometer. This ranking conforms to Table 2.19 studies that do not recommend accelerometers for detecting small leaks, leaks with low acoustic signals, and pipelines with long distances or many junctions between leak and sensor. However, our results do not confirm Almeida et al. (2018), Yazdekhasti et al. (2016), and Yazdekhasti et al. (2017) that suggested accelerometers when there are large resonances caused by hydrants and medium to small diameter PVC pipes. Our different results may stem from the smaller leak sizes and larger pipes than those in the mentioned studies.

The poorer performance of hydrophones than dynamic pressure sensors also follows Table 2.19 that does not recommend hydrophones when there are significant resonances caused by hydrants or leaks with narrowband signals in PVC pipes. Our findings do not align with Almeida et al. (2014) or Hunaidi and Wang (2006) that found hydrophones good fits for detecting small leaks and networks with many junctions, respectively. It is worth noting that in the former, the authors induced much larger leaks with higher sound intensities, and in the latter, an actual case study with large leaks was employed.

Hydrophone data characteristics depended on the sensor location. For instance, regarding H1 data in both networks and H2 measurements in the looped network, leak acoustic data were larger than those of NL. However, the H2 data of the branched network did not represent such a relation. The amplitudes of the H1 signals were less variable than the H2 signals. Since H1 is close to the pump, it recorded a constant background noise generated by the pump. This difference can highlight the effect of a hydrophone's location on its measurements.

Acoustic leak signal amplitudes in the looped network were more uniform than those of the branched one. This non-uniformity can be due to more barriers in the branched network, which cause more frequent signal attenuations and resonances. Due to more paths from the leak location to the looped network's hydrophones, leaks' acoustic emissions could propagate more uniformly via different pipe trajectories, which led to more similar acoustic leak signal magnitudes in the looped network.

GL and CC had the largest sound magnitudes in the looped and branched networks, respectively. Their larger magnitudes are due to more air in these leaks' outflow and their more intense sounds.

Leak data of P2 were more cyclic and more variable than P1 data due to transients. Not only do leaks have transient effects on pipes' dynamic pressure, but demand flow and junctions also caused cyclic patterns in P2 measurements. The tees and crosses also reduced dynamic pressure magnitudes at P2, which is at a farther location to the pump than P1. The dynamic pressure magnitude ranges of leaks were larger than those of no leak signals due to larger water velocities when leaks were present.

Acceleration data indicated that a leak's decreasing vibration effect was more significant than its additive impact caused by the leak thrust force, which is why the vibration magnitudes were smaller when there was a leak in the looped network with no demand compared to no-leak signals. Since leaks and water pressure can be larger in actual networks, leaks may increase accelerations throughout a pipeline.

Plots of the looped network's acceleration data without demand denoted a direct relation between leak flow rates and leaks' frequency amplitudes. There was a direct relation between leak flow rates and maximum frequencies for both networks' hydrophone data without demand and background sound. Due to the similarity of leak maximum frequencies in both networks' dynamic pressure data, there was no relation between leak flows and their frequency caps.

Similar to time-domain plots, the frequency contents of signals depended on sensor locations. Due to turbulences caused by the hydrant's blind end, the maximum frequency at A2, mounted on the hydrant connection, reached 6000 Hz, while the maximum frequency at A1 located at a tee was 500 Hz. Therefore, accelerometers close to dead-ends may measure vibrations that do not represent the whole network. The energy-dissipating impacts of tees and crosses between P1 and P2 dampened high-frequency waves and prohibited them from reaching P2. That is why we observed lower maximum frequencies of the leak and no leak signals at P2 than those at P1.

Since blind flanges and discontinuous pipes in the branched network attenuated high-frequency acoustic data, the pronounced high frequencies of the leak and NL signals in the looped network were larger than those of the branched network. This result aligns with the study by Cody et al. (2018), where junctions accounted for acoustic signal attenuations.

Though more assessment on non-zero demand data is necessary, dynamic pressure data with no demand indicated larger amplitudes of leak signals' dominant frequencies than those of NL signals. Therefore, the amplitude of dominant frequency may be a distinctive feature for leak detection.

The areas under the CSD plots of three sensors' zero demand data did not represent constant patterns in both networks and could not detect leaks, discern leak types or identify the network change.

Leak:NoLeak plots of hydrophone data without demand and sound indicated when a leak's water outflow included more airflow and had a more irregular shape, its leak:noleak ration was larger. That is why GL and OL had the largest and smallest sum of the leak:noleak magnitudes, respectively, based on acoustic data. Therefore, leak signal magnitudes depend on the shape of the leak water jet or outflow rather than its water jet height. The shape of leaks' water output also affected their dominant frequencies. For example, based on H2 data with zero demand and no sound, GL in the looped network with a dominant frequency of 40 Hz had the largest magnitude than other leaks.

Due to LC and CC LDI magnitudes' similarity, LDI irregular patterns, and magnitude variability, it could not distinguish leak types or identify the network change. Based on acceleration data, though leak LDI magnitudes became larger with leak flow increase in the branched network, this was not the case in the looped network that is not in agreement with studies of Yazdekhasti et al. (2016) and Yazdekhasti et al. (2017). This difference is due to GL's large flow rate but negligible vibration in our study, while the papers did not assess GL's LDI. Moreover, based on acoustic and dynamic pressure data, LDI did not change according to leak flow increases.

Based on dynamic pressure and acceleration data, since LDI increased with higher demands, we can expect larger LDI magnitudes at zones with higher consumptions.

GL had different characteristics than the cracks and the orifice. Not only did it have the largest leak flow rate and more intense outflow sound, but its dominant and fundamental frequencies were the lowest, based on accelerometer data. GL was the only leak where outflow was not a water jet, and water exited the leak opening with low pressure and irregular shape. Since pressure in actual networks is higher, GL flow can generate more intense sound and accelerations, leading to results different than those of our study.

Overall, neither of the sixteen plots and features could distinguish leaks and network architectures unanimously based on each parameter's two sensor data. Although RMS and standard deviation of dynamic pressure data could identify the network change, their discerned network was different based on P1 and P2. The lack of constant patterns and the incapability of features in differentiating leaks or network architecture is due to the complex characteristics of water distribution systems. Though the experiments helped study the effects of factors in a controlled manner, water turbulence, variable water pressure, and distinct pipe attenuation for different signals and frequencies caused the evaluation metrics to not represent the same information at various locations or scenarios. A solution to address these challenges is to employ statistical models, like binary or multi-label classifiers, that can learn complex data patterns and discern data categories based on a higher number of features, even though each of them is not distinctive enough.

2.14. References

Allen, M., Clark, R., Cotruvo, J. A., & Grigg, N. (2018). Drinking Water and Public Health in an Era of Aging Distribution Infrastructure. *Public Works Management & Policy*, *23*(4), 301–309. https://doi.org/10.1177/1087724x18788368

Almeida, F., Brennan, M. J., Joseph, P., Gao, Y., & Paschoalini, A. T. (2018). The effects of resonances on time delay estimation for water leak detection in plastic pipes. Journal of Sound and Vibration, 420, 315–329. https://doi.org/10.1016/j.jsv.2017.06.025

Almeida, F., Brennan, M., Joseph, P., Whitfield, S., Dray, S., & Paschoalini, A. (2014). On the Acoustic Filtering of the Pipe and Sensor in a Buried Plastic Water Pipe and its Effect on Leak Detection: An Experimental Investigation. Sensors, 14(3), 5595–5610. https://doi.org/10.3390/s140305595

Boashash, B. (2016). *Time-Frequency Signal Analysis and Processing (Second Edition)*. Academic Press. https://doi.org/10.1016/B978-0-12-398499-9.00012-1

Boretti, A., & Rosa, L. (2019). Reassessing the projections of the World Water Development Report. *Npj Clean Water*, *2*(1), 1–6. https://doi.org/10.1038/s41545-019-0039-9

Brumbelow, K., Torres, J., Guikema, S., Bristow, E., & Kanta, L. (2007). Virtual Cities for Water Distribution and Infrastructure System Research. Proceedings of the World Environmental and Water Resources Congress 2007 : Restoring Our Natural Habitat, 355–361. https://ascelibrary.org/doi/abs/10.1061/40927(243)469

Bryan/College Station Unified Design Guidelines- Domestic Water. (2012). Bcsunited. http://bcsunited.net/2012_Files/Design_Manual/Water_2012.pdf

Butterfield, J. D., Collins, R. P., Krynkin, A., & Beck, S. B. M. (2017a). Experimental Investigation into the Influence of Backfill Types on the Vibro-acoustic Characteristics of Leaks in MDPE Pipe. Procedia Engineering, 186, 311–318. https://doi.org/10.1016/j.proeng.2017.03.252

Butterfield, J. D., Krynkin, A., Collins, R. P., & Beck, S. B. M. (2017b). Experimental investigation into vibro-acoustic emission signal processing techniques to quantify leak flow rate in plastic water distribution pipes. Applied Acoustics, 119, 146–155. https://doi.org/10.1016/j.apacoust.2017.01.002

Cassa, A. M., van Zyl, J. E., & Laubscher, R. F. (2010). A numerical investigation into the effect of pressure on holes and cracks in water supply pipes. Urban Water Journal, 7(2), 109–120. https://doi.org/10.1080/15730620903447613

Cody, R., Dey, P., & Narasimhan, S. (2020). Linear Prediction for Leak Detection in Water Distribution Networks. Journal of Pipeline Systems Engineering and Practice, 11(1), 04019043. https://doi.org/10.1061/(asce)ps.1949-1204.0000415

Cody, R., Harmouche, J., & Narasimhan, S. (2018). Leak detection in water distribution pipes using singular spectrum analysis. Urban Water Journal, 15(7), 636–644. https://doi.org/10.1080/1573062x.2018.1532016

Covelli, C., Cozzolino, L., Cimorelli, L., Della Morte, R., & Pianese, D. (2015). A model to simulate leakage through joints in water distribution systems. Water Supply, 15(4), 852–863. https://doi.org/10.2166/ws.2015.043

El-Zahab, S., Mohammed Abdelkader, E., & Zayed, T. (2018). An accelerometer-based leak detection system. Mechanical Systems and Signal Processing, 108, 276–291. https://doi.org/10.1016/j.ymssp.2018.02.030

Farley, M. (2007). Losses in Water Distribution Networks: A Practitioner's Guide to Assessment, Monitoring and Control. IWA Publishing.

Farley, M., & Trow, S. (2003). Losses in Water Distribution Networks. Van Haren Publishing.

Fox, S., Collins, R., & Boxall, J. (2016). Physical investigation into the significance of ground conditions on dynamic leakage behaviour. Journal of Water Supply: Research and Technology-Aqua, 65(2), 103–115. https://doi.org/10.2166/aqua.2015.079

Franchini, M., & Lanza, L. (2013). Leakages in pipes: generalizing Torricelli's equation to deal with different elastic materials, diameters and orifice shape and dimensions. Urban Water Journal, 11(8), 678–695. https://doi.org/10.1080/1573062x.2013.868496

Gao, Y., Brennan, M. J., Joseph, P. F., Muggleton, J. M., & Hunaidi, O. (2005). On the selection of acoustic/vibration sensors for leak detection in plastic water pipes. Journal of Sound and Vibration, 283(3–5), 927–941. https://doi.org/10.1016/j.jsv.2004.05.004

Gao, Y., Brennan, M. J., Liu, Y., Almeida, F. C. L., & Joseph, P. F. (2017). Improving the Shape of the Cross-Correlation Function for Leak Detection in a Plastic Water Distribution Pipe Using Acoustic Signals. Applied Acoustics, vol. 127, pp. 24–33. Crossref, doi:10.1016/j.apacoust.2017.05.033

Greyvenstein, B., & van Zyl, J. E. (2007). An experimental investigation into the pressure - leakage relationship of some failed water pipes. Journal of Water Supply: Research and Technology-Aqua, 56(2), 117–124. https://doi.org/10.2166/aqua.2007.065

Guo, B., Song, S., Ghalambor, A., & Lin, L. R. (2014). *Offshore Pipelines: Design, Installation, and Maintenance* (2nd ed.). Gulf Professional Publishing. https://doi.org/10.1016/B978-0-12-397949-0.00030-3

Guo, G., Yu, X., Liu, S., Ma, Z., Wu, Y., Xu, X., Wang, X., Smith, K., & Wu, X. (2021). Leakage Detection in Water Distribution Systems Based on Time–Frequency Convolutional Neural Network. *Journal of Water Resources Planning and Management*, *147*(2), 04020101. https://doi.org/10.1061/(asce)wr.1943-5452.0001317

Harmouche, J., & Narasimhan, S. (2020). Long-Term Monitoring for Leaks in Water Distribution Networks Using Association Rules Mining. IEEE Transactions on Industrial Informatics, 16(1), 258–266. https://doi.org/10.1109/tii.2019.2911064

Helakari, H., Kananen, J., Huotari, N., Raitamaa, L., Tuovinen, T., Borchardt, V., Rasila, A., Raatikainen, V., Starck, T., Hautaniemi, T., Myllylä, T., Tervonen, O., Rytky, S., Keinänen, T., Korhonen, V., Kiviniemi, V., & Ansakorpi, H. (2019). Spectral entropy indicates electrophysiological and hemodynamic changes in drug-resistant epilepsy – A multimodal MREG study. *NeuroImage: Clinical*, *22*, 101763. https://doi.org/10.1016/j.nicl.2019.101763

Hunaidi, O., & Chu, W. (1999). Acoustical characteristics of leak signals in plastic water distribution pipes. Applied Acoustics, 58(3), 235–254. https://doi.org/10.1016/s0003-682x(99)00013-4

Hunaidi, O., Chu, W., Wang, A., & Guan, W. (2000). Detecting leaks in plastic pipes. Journal - American Water Works Association, 92(2), 82–94. https://doi.org/10.1002/j.1551-8833.2000.tb08819.x

Hunaidi, O., & Wang, A. (2006). A new system for locating leaks in urban water distribution pipes. Management of Environmental Quality: An International Journal, 17(4), 450–466. https://doi.org/10.1108/14777830610700928

H2c Hydrophone. (2020). Aquarian. https://www.aquarianaudio.com/h2c-hydrophone.html

Ismail, M., Dziyauddin, R. A., & Salleh, N. A. (2015). Performance evaluation of wireless accelerometer sensor for water pipeline leakage. 2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS), Langkawi, 2015, 120-125. https://doi.org/10.1109/IRIS.2015.7451598

Kadri, A., Abu-Dayya, A., Trinchero, D., & Stefanelli, R. (2011). Autonomous sensing for leakage detection in underground water pipelines. *2011 Fifth International Conference on Sensing Technology*, 639–643. https://doi.org/10.1109/icsenst.2011.6137061

Kang, J., Park, Y., Lee, J., Wang, S, & Eom, D. (2018). Novel Leakage Detection by Ensemble CNN-SVM and Graph-Based Localization in Water Distribution Systems. IEEE Transactions on Industrial Electronics, 65(5), 4279–4289. https://doi.org/10.1109/tie.2017.2764861

Karray, F., Garcia-Ortiz, A., Jmal, M. W., Obeid, A. M., & Abid, M. (2016). EARNPIPE: A Testbed for Smart Water Pipeline Monitoring Using Wireless Sensor Network. Procedia Computer Science, 96, 285–294. https://doi.org/10.1016/j.procs.2016.08.141

Khulief, Y. A., Khalifa, A., Mansour, R. B., & Habib, M. A. (2012). Acoustic Detection of Leaks in Water Pipelines Using Measurements inside Pipe. Journal of Pipeline Systems Engineering and Practice, 3(2), 47–54. https://doi.org/10.1061/(asce)ps.1949-1204.0000089

Lee, E. J., & Schwab, K. J. (2005). Deficiencies in drinking water distribution systems in developing countries. Journal of Water and Health, 3(2), 109–127. https://doi.org/10.2166/wh.2005.0012

Li, S., Song, Y., & Zhou, G. (2018). Leak detection of water distribution pipeline subject to failure of socket joint based on acoustic emission and pattern recognition. Measurement, 115, 39–44. https://doi.org/10.1016/j.measurement.2017.10.021

Li, R., Huang, H., Xin, K., & Tao, T. (2014). A review of methods for burst/leakage detection and location in water distribution systems. *Water Supply*, *15*(3), 429–441. https://doi.org/10.2166/ws.2014.131

Ma, Y., Gao, Y., Cui, X., Brennan, M., Almeida, F., & Yang, J. (2019). Adaptive Phase Transform Method for Pipeline Leakage Detection. Sensors, 19(2), 310. https://doi.org/10.3390/s19020310

Marmarokopos, K., Doukakis, D., Frantziskonis, G., & Avlonitis, M. (2018). Leak Detection in Plastic Water Supply Pipes with a High Signal-to-Noise Ratio Accelerometer. Measurement and Control, 51(1–2), 27–37. https://doi.org/10.1177/0020294018758526

Martini, A., Rivola, A., & Troncossi, M. (2018). Autocorrelation Analysis of Vibro-Acoustic Signals Measured in a Test Field for Water Leak Detection. Applied Sciences, 8(12), 2450. https://doi.org/10.3390/app8122450

Martini, A., Troncossi, M., & Rivola, A. (2015). Automatic Leak Detection in Buried Plastic Pipes of Water Supply Networks by Means of Vibration Measurements. Shock and Vibration, 2015, 1–13. https://doi.org/10.1155/2015/165304.

Mohd Ismail, M. I., Dziyauddin, R. A., Ahmad Salleh, N. A., Muhammad-Sukki, F., Aini Bani, N., Mohd Izhar, M. A., & Latiff, L. A. (2019b). A Review of Vibration Detection Methods Using Accelerometer Sensors for Water Pipeline Leakage. *IEEE Access*, *7*, 51965–51981. https://doi.org/10.1109/access.2019.2896302

Motazedi, N., & Beck, S. (2017). Leak detection using cepstrum of cross-correlation of transient pressure wave signals. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 232(15), 2723–2735. https://doi.org/10.1177/0954406217722805

National Public Radio. (2016). Flint's water system is falling apart. Fixing it could cost $100 million. Michiganradio. https://www.michiganradio.org/post/flints-water-system-falling-apart-fixing-it-could-cost-100-million

Park, J., & Mackay, S. (2003). Practical Data Acquisition for Instrumentation and Control Systems. Pages 1-12. Elsevier Gezondheidszorg. https://doi.org/10.1016/B978-0-7506-5796-9.X5000-9

Papastefanou, A. (2011). An experimental investigation of leak noise from water filled plastic pipes (Thesis). University of Southampton. https://eprints.soton.ac.uk/190853/1/P2783.pdf

PCB Model 102B16. (2020). PCB Piezotronics. https://www.pcb.com/products?m=102B16

PCB Model 333B50. (2020).). PCB Piezotronics. https://www.pcb.com/products?model=333b50

PCB Model 352A24/NC. (2020). PCB Piezotronics. https://www.pcb.com/products?m=352a24_nc

Rossman, L. A. (2000). EPANET 2 USERS MANUAL. EPANET 2 Users Manual, 1–199. https://nepis.epa.gov/Adobe/PDF/P1007WWU.pdf

Sadeghioon, A., Metje, N., Chapman, D., & Anthony, C. (2014). SmartPipes: Smart Wireless Sensor Networks for Leak Detection in Water Pipelines. Journal of Sensor and Actuator Networks, 3(1), 64–78. https://doi.org/10.3390/jsan3010064

Seto, W. W. (1964). Theory and problems of mechanical vibrations (Schaum's outline series) (1st Edition). McGraw-Hill.

Srirangarajan, S., Allen, M., Preis, A., Iqbal, M., Lim, H. B., & Whittle, A. J. (2012). Wavelet-based Burst Event Detection and Localization in Water Distribution Systems. Journal of Signal Processing Systems, 72(1), 1–16. https://doi.org/10.1007/s11265-012-0690-6

Stathis, J. A. (1998). A Model for Determining Leakage in Water Distribution Systems (Master's dissertation). Virginia Polytechnic Institute and State University. https://vtechworks.lib.vt.edu/bitstream/handle/10919/46422/ETD.PDF?sequence=1&isAllowed=y

Sun, J., Xiao, Q., Wen, J., & Wang, F. (2014). Natural gas pipeline small leakage feature extraction and recognition based on LMD envelope spectrum entropy and SVM. *Measurement*, *55*, 434–443. https://doi.org/10.1016/j.measurement.2014.05.012

Sun, J., Xiao, Q., Wen, J., & Zhang, Y. (2016). Natural gas pipeline leak aperture identification and location based on local mean decomposition analysis. *Measurement*, *79*, 147–157. https://doi.org/10.1016/j.measurement.2015.10.015

Szücs, E. (1980). Similitude and Modelling. pp. 113–117, North Holland.

Telg´arsky, R. (2013, June 1). *Dominant Frequency Extraction*. ArXiv.Org. https://arxiv.org/pdf/1306.0103.pdf

Technical Brochure of MCC End Suction Centrifugal Pumps. (2018). Goulds Water Technology. https://documentlibrary.xylemappliedwater.com/wp-content/blogs.dir/22/files/2012/07/BMCC-R5.pdf

Van Zyl, J. E., Alsaydalani, M. O. A., Clayton, C. R. I., Bird, T., & Dennis, A. (2013). Soil fluidisation outside leaks in water distribution pipes – preliminary observations. Proceedings of the Institution of Civil Engineers - Water Management, 166(10), 546–555. https://doi.org/10.1680/wama.11.00119

Visenti. (2020). LeakViewTM. https://www.visenti.com/resources/front/template/visenti/doc/1703-Visenti-LeakView_Final-Lowres.pdf

Walski, T. (2006). A history of Water distribution. Journal of American Water Works Association, 98(3), 110–121. https://doi.org/10.1002/j.1551-8833.2006.tb07611.x

Walski, T., Bezts, W., Posluszny, E. T., Weir, M., & Whitman, B. (2004). Understanding the Hydraulics of Water Distribution System Leaks. Critical Transitions in Water and Environmental Resources Management.

White, F. M. (1979). Fluid Mechanics (1st ed.). McGraw Hill.

Wu, Z. Y., Sage, P., & Turtle, D. (2010). Pressure-Dependent Leak Detection Model and Its Application to a District Water System. Journal of Water Resources Planning and Management, 136(1), 116–128. https://doi.org/10.1061/(asce)0733-9496(2010)136:1(116)

Xing, L., & Sela, L. (2020). Transient simulations in water distribution networks: TSNet python package. *Advances in Engineering Software*, *149*, 102884. https://doi.org/10.1016/j.advengsoft.2020.102884

Yazdekhasti, S., Piratla, K. R., Atamturktur, S., & Khan, A. (2017). Experimental evaluation of a vibration-based leak detection technique for water pipelines. Structure and Infrastructure Engineering, 14(1), 46–55. https://doi.org/10.1080/15732479.2017.1327544

Yazdekhasti, S., Piratla, K. R., Atamturktur, S., & Khan, A. A. (2016). Novel vibration-based technique for detecting water pipeline leakage. Structure and Infrastructure Engineering, 13(6), 731–742. https://doi.org/10.1080/15732479.2016.1188318

Zan, T. T. T., Lim, H. B., Wong, K.-J., Whittle, A. J., & Lee, B.-S. (2014). Event Detection and Localization in Urban Water Distribution Network. IEEE Sensors Journal, 14(12), 4134–4142. https://doi.org/10.1109/jsen.2014.2358842

Zhang, Y., Chen, S., Li, J., & Jin, S. (2014). Leak detection monitoring system of long-distance oil pipeline based on dynamic pressure transmitter. Measurement, 49, 382–389. https://doi.org/10.1016/j.measurement.2013.12.009

ZOOM UAC-2 USB Audio Interface. (2020). ZOOM. https://zoomcorp.com/en/us/audio-interface/audio-interfaces/uac-2/

3. ASSESSMENT OF DATA EMPLOYING MACHINE LEARNING TECHNIQUES TO DETECT LEAKS USING HYDOPHONE SENSORS

3.1. Introduction

Artificial intelligence techniques have been broadly employed to detect leaks in water distribution systems. Mounce et al. (2002) presented one of the first applications of ANNs to detect bursts using historical data. They used the mixed density network to produce probability density forecasts of DMA flow time series data and predicted the conditional probability density function of the data and compared it with their counterparts to form residuals. Using prediction confidence intervals and Fuzzy memberships for the residuals, abnormalities in a time window were considered as bursts. Possible challenges in this study were choosing the right length for the time window and threshold values in flagging the abnormalities.

Cody and Narasimhan (2020) applied a fusion of a linear prediction (LP) model for feature extraction and a multivariate gaussian mixture model (GMM), for classification, in a small section of a Canadian city's distribution network. Though their field implementation yielded promising results, three points may adversely influence their study's reproducibility, and accuracy: (1) the LP model uses a short-term Fourier transform (STFT) to extract LP coefficients as features. On the one hand, the STFT employs a uniform-frequency convolution to calculate the coefficients. On the other hand, an acoustic signal's frequency contents are not uniform and change when a leak is induced or grows. Therefore, the STFT is not the best frequency-extraction option for a signal with varying frequency contents and wavelet convolutions are better aligned with this purpose; (2) equal numbers of leak and non-leak samples were used to build a dataset and based on that, equal weights were given to the mixture models. However, in an actual water network, the number of leak samples is much less than those of non-leak. The authors could build an imbalanced dataset with more non-leak samples and give larger weights to the mixture models of leak signals to highlight them. This would enable them to take advantage of more samples in model training; (3) accuracy was used as an evaluation metric for the classifier. Regarding the imbalanced nature of the leak and non-leak data in a real water network, accuracy might be biased to the majority class, the non-leak class in this case.

Convolutional neural networks (CNNs) have become widely used to detect leaks and bursts or extract features to be directed into other classifiers. Bohorquez et al. (2020) compared the efficacy of dense neural networks and convolutional neural networks to detect leaks based on the generation of a transient event. They found a 1-dimensional convolutional neural network (1D-CNN) performed more accurately in leak detection with at least 25,000 required examples to train the model. In another paper, Bohorquez et al. (2021) designed 1D-CNNs to detect bursts simulated numerically and in a single pipeline. The authors applied CNNs since they have fewer weights to be tunned and are less vulnerable to over-fitting. Kang et al. (2019) used accelerometer data to detect leaks in a real water network where a 1D-CNN was employed as a feature extractor for an ensemble 1D-CNN-SVM classifier.

Moreover, Chuang et al. (2019) trained a CNN model to classify acoustic data as leak or non-leak, where Mel frequency cepstrum coefficients were extracted as features for the model training and test. Recently, Guo et al. (2021) studied leak detection based on time-frequency convolutional neural networks (TFCNN). They extracted spectrograms, with three different resolutions, of short-time windows sliding through vibration signals and trained a CNN for binary classification. They compared the performance of a TFCNN with other CNNs trained with time- and frequency-domain features and shallow classifiers such as decision tree, SVM, random forest, multi-layer perceptron, and XGBoost. Results showed that the best classifier was a TFCNN with an AUC of 0.99 and the worst model was a decision tree with an AUC of 0.78.

Though CNN-based models have successfully detected leaks in real case studies, they lack two benefits of shallow classifiers: (1) lots of parameters should be optimized to build these deep classifiers. For instance, CNN's used by Guo et al. (2021) needed 2,949,120 hyperparameters to be set to choose the most optimized model. Bohorquez et al. (2021) trained a 1D-CNN network by adjusting between 26,808 and 81,250 weights. And, the more hyperparameters, the more time and computational complexities. Besides, these networks are prone to overfitting, particularly in fully-connected networks. Even though locally connected CNNs can address these complexities (Kang et al., 2018), they may miss some important spatial patterns ignored by local filters; (2) large data is needed to train these networks and have them learn different patterns in data. For example, Kang et al. (2018), Bohorquez (2020), and Guo et al. (2021) composed training datasets with 94,080, 25,000, and 33,335 samples, respectively. These deep networks suffer from over-fitting in the case of small training data.

Nonetheless, shallow classifiers are less prone to overfit on a small dataset and have a low computational cost in the training phase. Pasupa and Sunhem (2016) compared shallow classifiers, SVM, and ANN with one hidden layer, with a deep neural network, CNN, to classify face images. They found that SVM and ANN using hand-crafted features performed better than a CNN that used raw or reconstructed images without any regularization. The authors concluded that a deep model that used regularization techniques, especially the dropout technique for CNN, can address an over-fitting problem in small datasets. Comparing the performance of shallow and deep models to classify image and text data, Yin et al. (2014) empirically verified that shallow classifiers are more robust in the classification of good quality images, while deep models have an encouraging performance on relatively low-quality images. They also argued that a hybrid model of both conventional and deep models works better for challenging and real recognition problems.

Among applications of conventional models, Cody et al. (2018) employed SVM and one-class SVM to classify hydrophone data for leak detection. The models performed successfully where SVM had AUCs between 0.85 and 0.92, and one-class SVM classified with an AUC of 0.90. Kampelopoulos et al. (2020) compared the performance of a decision tree (DT) and SVM models to detect leaks in noisy industrial pipelines. They used time- and frequency domain features and found that the classifiers with the following parameters discerned leak and non-leak data accurately: SVM: C=60, kernel, polynomial, degree = 7, gamma scale; DT: max_depth = 7, max_features = 7, min_samples_split = 9, min_samples_leaf = 6. Though DT had a higher accuracy, which was 0.9780, SVM performed more consistently on different combinations of datasets with an average accuracy of 0.9760.

Moreover, the two models had great performances in classifying both leak and non-leak data. Rashid et al. (2015) developed a wireless sensor network to detect leaks in oil transmission pipes using four machine learning algorithms, i.e., SVM, K-nearest neighbor (KNN), Gaussian mixture model (GMM), and Naive Bayes (NB). They extracted time- and frequency-domain features of five pressure transducers data mounted on pipes of different sizes. The authors found that SVM had the best detection performance with an accuracy of 94.5%, followed by GMM, NB, and KNN. Xiao et al. (2019) utilized SVM to detect gas leaks in a laboratory setup. Time, frequency, and wavelet domain features were extracted from acoustic signals and built training and test data. It was found that the preprocessing methods, along with SVM classifiers, could discern leak and non-leak signals by high accuracy of 99.4%. Yet, the authors did not mention their method's accuracy in predicting leak and non-leak data separately.

At least three reasons make efficient feature extraction highly crucial in a leak detection methodology: (1) datasets employed in leak detection studies, especially in real case studies, have small signal-to-noise ratios; (2) unlike anomalies caused by bursts, background leaks have latent signatures which are not easily distinguishable; (3) the efficacy and computational complexity of classification algorithms depend on the input data, specifically when data is large and high-dimensional. Therefore, feature extraction and selection play significant roles in leak detection. Extracted features used for leak detection are in four folds: (1) time-domain; (2) frequency-domain; (3) wavelet domain; (4) others such as CNN outputs or Singular Spectral Analysis (SSA) coefficients. Table 3.1 lists some research that employed these feature extraction methods.

Table 3.1. List of leak detection studies with different features as inputs to classifiers

Authors	Time domain	Frequency domain	Wavelet domain	Others
Xiao et al. (2019)	absolute mean, standard deviation (SD), crest factor, short-term energy, kurtosis, skewness	frequency centroid, frequency band, peak frequency	wavelet mean frequency, wavelet entropy	-
Cody et al. (2018)	-	-	-	SSA coefficients
Li et al. (2018)	peak, mean, SD, root mean square (RMS), crest factor, energy	kurtosis, peak frequency, skewness, frequency centroid	-	-
Guo et al. (2021)	RMS, mean, zero-crossing rate, autocorrelation energy ratio, energy entropy ratio	mean dB of power spectral density, RMS of intrinsic mode functions (IMFs), Shannon entropy of IMFs, subband spectral entropy, spectrogram	-	Mean Teager energy operator

Table 3.1. Continued

Authors	Time domain	Frequency domain	Wavelet domain	Others
Rashid et al. (2015)	energy, gradient, kurtosis, entropy, mean, variance	spectrogram, pseudospectrum, power spectral density	-	-
Harmouche and Narasimhan (2020)	-	spectral energy contents between and out of specific frequencies	-	-
Chuang et al. (2019)	-	Mel frequency cepstral coefficients	-	-
Yu and Li (2017)	mean, SD, RMS, peak, energy	frequency centroid, peak frequency, average frequency	-	-
Kang et al. (2018)	-	-	-	CNN output vectors

Wavelet transforms are popular methods to extract features in anomaly detection studies. As a structural health monitoring research, Ebrahimkhanlou et al. (2019) employed modular coefficients of continuous wavelet transforms as feature inputs to an autoencoder to detect and localize acoustic emission sources in plate-like structures. They utilized three different complex mother wavelets, including complex Morlet, 2^{nd}-order complex Gaussian, and 8^{th}-order complex Gaussian. The authors realized comparable performance for input patterns generated by the 8^{th}-order complex Gaussian and complex Morlet to characterize and localize acoustic emission on two aluminum panels. Ebrahimkhanlou and Salamone (2018) used a complex Morlet mother wavelet to train CNN and autoencoder deep learning models. The authors extracted the modulus of continuous wavelet values at frequencies of 75 (kHz), 200 (kHz), and 325 (kHz). Then, they generated vectors whose values were normalized continuous wavelet magnitudes at lower frequencies. The vectors were then inputted into the deep learning models. Sarrafi et al. (2018) and Mostavi et al. (2017) applied continuous wavelet transforms in health monitoring studies. According to the review paper conducted by Li et al. (2014), multiple research found that the wavelet transform analysis is an effective method to extract signal information to detect leaks. In different studies, Ferante and Brunone (2003) and Taghvaei (2009) applied the wavelet transform method, where their results indicated the method is capable of identifying small leaks in water pipelines. In similar research but on a different case study, Xiao et al. (2019) employed wavelet transform coefficients to extract features for an SVM classifier and found the wavelet mean frequency as one of the first three most valuable features that represent leak signature for classification. Research performed by Ahadi and Bakhtiar (2010) and Zhang et al. (2014) emphasize the efficiency of wavelet transforms to detect leaks in unsteady and dynamic signals.

In the following, subsection “Methodology” describes the wavelet transform, classification algorithms, and evaluation metrics used in this research. Section “Case Study” explains the experimental setup and modeling steps. Section “Results and Discussion” discusses performance of models and effective classifiers. And, section “Conclusion” summarizes the study and offers future work options.

3.2. Methodology

3.2.1. Wavelet Transform

A wavelet is a function with a zero average. The continuous wavelet transform (CWT) of any function f concerning a wavelet family is defined as:

$$C(f,t) \ = \ \frac{1}{\sqrt{s(f)}} \int_{-\infty}^{+\infty} r(\tau) \, \Psi^* \, (\frac{\tau - t}{s(f)}) \, d\tau \tag{3.1}$$

where f and t are the frequency and translation, or position, parameters of a wavelet transform. $\Psi^*(t)$ is a complex conjugation of $\Psi(t)$ that is a mother wavelet. In addition, $s(f)$ is the non-dimensional scale parameter of the transform (Stéphane, 2009).

The wavelet transform is a means to measure the similarity between wavelets and acoustic signals. If an acoustic signal has a dominant component within a specific range of frequencies, the continuous wavelet transform coefficients calculated at those frequencies in the time domain would be relatively larger. As a result, CWT can reveal local transient features in the time and frequency domains of a signal, which effectively extracts time-frequency features of signals. Thus, the wavelet transform, a convolution composed of different frequencies, can analyze time-varying nonstationary acoustic signals (Xiao et al., 2019). One of the advantages of the wavelet transform compared to the Fourier transform is that the former shows frequency variations of a signal over time. While the latter only presents frequency changes regardless of the time those changes have happened.

The most efficient wavelet for this study's application was found to be the complex Morlet mother wavelet (Mallat, 1999) defines as

$$\Psi_M(t) \ = \ \frac{1}{\sqrt{\pi f_b}} \, exp \, (2\pi f_c jt \ - \ \frac{t^2}{f_b}) \tag{3.2}$$

where f_c (central frequency) and f_b (bandwidth) are two user-defined, non-dimensional parameters (Ebrahimkhanlou et al., 2019). The

3.2.2. Classification Algorithms

Since this study's objective is detecting leaks, a classification algorithm is desired that would be trained with labeled or unlabeled historical data and can tag an unseen signal as leak or non-leak. The following subsections explain the classifiers employed in this research.

3.2.2.1. One-class Support Vector Machine

One-class Support Vector Machines (1CSVMs) are general semi-supervised approaches that map the original data from the input space to the feature space through different kernels. The algorithm finds a smooth plane or boundary to separate mapped data in the new feature space to normal and anomalous values. The kernel functions map the data vectors to a higher dimensional inner product space, and the boundaries that separate the two-class data in the feature space usually have nonlinear shapes in the input data space (Erfani et al., 2016; Scholkopf and Smola, 2018).

The most popular kernels used in 1CSVM are linear, polynomial, sigmoidal, and Gaussian radial basis functions (RBF). There are different formulations of one-class SVM, among which the method proposed by Schölkopf et al. (2001) will be used in this study. The authors proposed a hyper-plane based one-class SVM, where mapped vectors in the feature space are separated from the origin by a hyperplane with the largest possible margin. The vectors contained in the half-space close to the origin are anomalies. This plane-based 1CSVM is called PSVM, which is formulated as:

$$\min_{W,\xi,\rho} \frac{1}{2} \|W\|^2 + \frac{1}{\upsilon l} \sum_i \xi_i - \rho \tag{3.3}$$

$$\text{subject to } (W.\, \phi(X_i)) \geq \rho - \xi_i \,, \xi_i \geq 0 \tag{3.4}$$

where v represents an upper bound of the fraction of training errors and a lower bound of the fraction of support vectors and takes values from (0,1], and ξ_i is the distance of sample i in the margin from the margin boundaries.

3.2.2.2. Isolation Forest

Isolation Forest (iForest) is an unsupervised anomaly detection algorithm based on the idea 'anomalies are few and different,' and therefore, they are more likely to be isolated. iForest composed of multiple isolation trees, entitled iTrees finds anomalies regarding the point that anomalies need fewer partitions to be isolated, unlike normal points. Hence, anomalies have a shorter path length in their isolation trees. Some advantages of iForest are: (1) since the algorithm focuses on abnormal points, it requires small sample datasets which mainly include anomalies; (2) if anomalies are swamped or mask their presence, iForest builds a partial model by sub-sampling that addresses the problems of swamping and masking; (3) since it uses a portion of datasets, iForest has low memory and time complexities; (4) if equipped with an additional attribute selector, iForest works well when for high dimensional problems with many irrelevant attributes (Liu et al., 2008; Hariri et al., 2019; Talagala et al., 2020). Algorithms 1 and 2 show the details of training steps where more details can be found in the paper of Liu et al. (2008).

Algorithm 1: *iForest* (X; T; N)

Input: input dataset X, number of trees t, subsampling size Ψ

Output: a set of *iTrees*

1: Initialize *Forest* = {}

2: set *iTree* height h = ceiling ($log_2 N$)

3: **for** i = 1 to T **do**

4: $X' \leftarrow$ *sample* (X, N)

5: *Forest* $\leftarrow$ *Forest* $\cup$ *iTree* (X', 0, h)

6: **end for**

return *Forest*

There are two input parameters to the iForest algorithm. They are the sub-sampling size Ψ and the number of trees t. Sub-sampling size controls the training data size. When Ψ increases to the desired value, iForest detects reliably, and there is no need to increase Ψ further. An experimental study conducted by Liu et al. (2008) shows that $\Psi = 256$ is enough to perform anomaly detection across a wide range of data.

The number of tree t controls the ensemble size. Liu et al. (2008) found that lengths usually converge well before $t = 100$.

Algorithm 2: *iTree* (X, e, l)

Inputs: input data X, current tree height e, height limit l

Output: an *iTree*

```
if e ≥ l or |X| ≤ 1 then
    return exNode {Size ← |X|}
else
    let Q be a list of attributes in X
    randomly select an attribute q ∈ Q
    randomly select a split point p from max and min values of attribute q in X
    X_l ← filter (X, q < p)
    X_r ← filter (X, q ≥ p)
    return inNode {Left ← iTree (X_l , e + 1, l),
                   Right ← iTree (X_r, e + 1, l),
                   SplitAtt ← q,
```

12: *SplitValue* ← *p*}

13: **end if**

Once trained, iForest returns a collection of trees ready for the evaluation stage. Algorithm 3 includes the *PathLength* function used in the evaluation stage.

Algorithm 3: *PathLength* (x, T, e)

Inputs: an instance x, an *iTree* T, current path length e; to be initialized to zero when first called

Output: path length of x

1: **if** T is an external node *then*

2: return $e + c(T.size)$ {$c(.)$ is defined in Equation 1}

3: **end if**

4: $a \leftarrow T.splitAtt$

5: **if** $x_a < T.splitValue$ **then**

6: return *PathLength*(x, $T.left$, $e + 1$)

7: **else** {$x_a \geq T.splitValue$}

8: return *PathLength*(x, $T.right$, $e + 1$)

9: **end if**

3.2.2.3. Local Outlier Factor

Local Outlier Factor (LOF) is an unsupervised anomaly detection algorithm that measures a data point's local deviation concerning its neighbors. It generates an anomaly score for each data point by measuring the local density of a given data point to the data near it. In other words, a local density can be determined by estimating distances between data points that are k-nearest neighbors. Once local densities are calculated, one can determine which data points have similar densities and which have a lesser density than their neighbors by comparing these densities. Those with lesser densities are considered outliers (Breunig et al., 2000). The following are some notations of LOF. Given a dataset *O* and an object *p*, the basic procedure for calculating the LOF value of *p* is as follows (You et al., 2020).

1. Find the *k*-nearest neighbors (*kNN*) of *p*:

   kNN returns the set $N_k^{(p)} \subseteq O$ of size *k* such that:

   $$\forall_O \in N_k(p), \forall_q \in O, q \notin N_k(p) \Rightarrow dist(p, o) \leq dist(p, q) \quad (3.5)$$

   where *dist(p, q)* is the Euclidean distance between object *p* and object *q*.

2. Calculate the *k*-distance of *p* as follows:

   $$k - dist(p) = max\ \{dist(p, o)|o \in N_k(p)\} \quad (3.6)$$

3. For each object $o \in N_k(p)$, calculate the reachability distance of *p* w.r.t. *o* as follows:

   $$reach - dist(p, o) = max\ \{k - dist(o), dist(p, o)\} \quad (3.7)$$

4. Calculate the local reachability density of p as follows:

   $$lrd(p) = \frac{k}{\sum_{o\, \in\, N_k(p)} reach - dist(p,o)} \quad (3.8)$$

5. Calculate the LOF value of p as follows:

   $$lof(p) = \frac{\sum_{o\, \in\, N_k(p)} \frac{lrd(o)}{lrd(p)}}{k} \quad (3.9)$$

3.2.2.4. Support Vector Machine

Support vector machines (SVMs) are versions of support vector classifiers (SVCs) that can separate data with non-linear boundaries. An SVC can be formulated as the following.

maximize M (3.10)

subject to $\sum_{j=1}^{p} = 1,$ (3.11)

$y_i\,(\beta_o + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq M\,(1 - \epsilon_i),$ (3.12)

$\epsilon_i \geq 0, \sum_{i=1}^{n} \epsilon_i \leq C,$ (3.13)

where *C* is a nonnegative tuning parameter. *M* is the width of the margin, and this quantity should become as large as possible. In Eq. (9.12), $\epsilon_1, \ldots, \epsilon_n$ are slack variables that allow individual observations to be on the wrong side of the margin or the hyperplane. Once Eq. (9.10) to Eq. (9.13) are solved, a test observation $x*$ can be classified by simply determining on which side of the hyperplane, i.e. $\beta_o + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}$, it lies. That is, the test observation can be classified based on the sign of $f(x^*) = \beta_0 + \beta_1 x_1^* + \ldots + \beta_p x_p^*$.

The slack variable ϵ_i tells us where the i^{th} observation is located, relative to the hyperplane and relative to the margin. If $\epsilon_i = 0$, then the i^{th} observation is on the correct side of the margin. If $\epsilon_i > 0$ then the i^{th} observation is on the wrong side of the margin, and we say that the i^{th} observation has violated the margin. If $\epsilon_i > 1$ then it is on the wrong side of the hyperplane.

In Eq. (3.13), *C* bounds the sum of the ϵ_i's, and so it determines the number and severity of the violations to the margin (and to the hyperplane) that can be tolerated. *C* could be thought of as a budget because the *n* observations can violate the margin. If $C = 0$, then there is no budget for violations to the margin, and it must be the case that $\epsilon_1 = \ldots = \epsilon_n = 0$. For $C > 0$ no more than *C* observations can be on the wrong side of the hyperplane because if an observation is on the wrong side of the hyperplane, then $\epsilon_i > 1$, and (3.13) requires that $\sum_{i=1}^{n} \epsilon_i \leq C$. As budget *C* increases, violations to the margin will be more tolerated, and so the margin will widen. Conversely, as *C* decreases, violations to the margin will be less tolerated, and so the margin narrows.

Equation (3.12) defines a linear hyperplane for classification in a two-class setting if the two classes' boundary is linear. However, in practice and when class boundaries are non-linear, a support vector classifier or any linear classifier will perform poorly. To address non-linear boundaries between classes, feature space is enlarged by mapping to a new higher-order feature space. In the new feature space, the classes can be separated by a smooth hyperplane. In other words, the support vector machine is an extension of the support vector classifier that results from enlarging the feature space in a specific way, using kernels. It is worth noting that the solution to the support vector classifier problem, Eq. (3.10) to Eq. (3.13), involves only the observations' inner products. A kernel is a generalized and faster method to compute the observations' inner products and is a function that quantifies the similarity of two observations. There are many forms of kernel functions included, but not limited to linear, polynomial, and radial basis function (RBF). Eq. (3.14), Eq. (3.15), and Eq. (3.16) represent linear, polynomial, and RBF kernels, respectively.

$$\langle x, x' \rangle \tag{3.14}$$

$$(\gamma \langle x, x' \rangle + r)^d \tag{3.15}$$

$$\exp\left(-\gamma \parallel x - x' \parallel^2\right) \tag{3.16}$$

where γ is a positive constant, d is a polynomial degree, and r is a coefficient. These parameters are set as inputs to the algorithm and are recommended to be determined by cross-validation (James et al., 2013; Hastie et al., 2010; Bishop, 2006).

3.2.2.5. XGBoost

XGBoost, which stands for eXtreme Gradient Boosting, is a gradient tree boosting classification algorithm used extensively by the data science community (Bekkerman, 2015; Bennett et al., 2007). It is a scalable end-to-end tree boosting system that adds weak learning trees in sequence to correct the errors of the previously trained trees. In each step, a new tree minimizes the loss function while the previous trees remain unchanged. Some of the main advantages of XGBoost are its scalability, which makes it run much faster than its counterparts with a cash-aware structure, capability of handling sparse data, and appropriateness for parallel and distributed computations (Chen and Guestrin, 2016). Below explains how the objective function of XGBoost is formulated that makes it a fast and efficient classifier.

For a given data set with n examples and m features $D = \{(x_i, y_i)\}$ ($|D| = n$, $x_i \in \mathbb{R}^m$, $y_i \in \mathbb{R}$), a tree ensemble model uses K additive functions to predict the output.

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in \mathcal{F} \tag{3.17}$$

where $\mathcal{F} = \{f(x) = w_{q}(x)\}(q : \mathbb{R}^m \rightarrow T, w \in \mathbb{R}^T)$ is the space of regression trees. Here q represents the structure of each tree that maps an example to the corresponding leaf index. T is the number of leaves in the tree. Each f_k corresponds to an independent tree structure q and leaf weights w. To learn the set of functions used in the model, the following regularized objective should be minimized.

$$\Upsilon(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{3.18}$$

where $£(f) = \gamma T + \frac{1}{2}\lambda \parallel \omega \parallel^2$

Here l is a differentiable convex loss function that measures the difference between the prediction $\hat{y}_i$ and the target y_i. The second term $\mathbf{\Omega}$ penalizes the complexity of the model (i.e., the regression tree functions). The additional regularization term helps to smooth the final learnt weights to avoid over-fitting. Intuitively, the regularized objective will tend to select a model employing simple and predictive functions.

The tree ensemble model in Eq. (3.18) includes functions as parameters and cannot be optimized using traditional Euclidean space optimization methods. Instead, the model is trained in an additive manner. Let $\hat{y}_i$ be the prediction of the *i*-th instance at the *t*-th iteration, f_t is needed to minimize the following objective.

$$\Upsilon^t = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{3.19}$$

The algorithm includes some hyperparameters that are: (1) number of trees (i.e., weak learners); (2) maximum depth of trees; (3) learning rate that determines the contribution of each weak learner to the additive model (Taormina and Galelli, 2018).

3.2.3. Principal Component Analysis

Principal components analysis (PCA) is a popular approach for deriving a low-dimensional set of features from a large set of variables. PCA is a technique for reducing the dimension of an $n \times p$ data matrix X. A data matrix with lower dimensions results in capturing the combination of most representative features while time and memory complexities decrease. The first principal component direction of the data is that along which the observations vary the most. Projecting the observations onto any other line would yield projected observations with lower variance. Projecting a point onto a line simply involves finding the location on the line which is closest to the point. There is also another interpretation for PCA: the first principal component vector defines the line that is as close as possible to the data (James et al., 2013; Hastie et al., 2010).

If we use PCA for dimensionality reduction, we construct a $d \times k$–dimensional transformation matrix W that allows us to map a sample vector x onto a new k–dimensional feature subspace that has fewer dimensions than the original d–dimensional feature space:

$$\begin{aligned} &\boldsymbol{x} = [x_1, x_2, \ldots, x_d], \quad \boldsymbol{x} \in \mathbb{R}^d \\ &\downarrow \boldsymbol{xW}, \quad \boldsymbol{W} \in \mathbb{R}^{d \times k} \\ &\boldsymbol{z} = [z_1, z_2, \ldots, z_k], \quad \boldsymbol{z} \in \mathbb{R}^k \end{aligned} \tag{3.20}$$

As a result of transforming the original d-dimensional data onto this new k-dimensional subspace (typically $k \ll d$), the first principal component will have the largest possible variance, and all consequent principal components will have the largest variance given the constraint that these components are uncorrelated (orthogonal) to the other principal components — even if the input features are correlated, the resulting principal components will be mutually orthogonal (uncorrelated) (Bishop, 2006). In this research, we will use PCA to extract the most representative features for training algorithms and also employ original features. Then the performances of the algorithms will be compared regarding the influence of PCA-based reduced dimensionality.

3.2.4. Evaluation Metrics

In this study, a classifier is desired that (1) can predict all leaks (positive incidents) correctly and (2) misclassifies the fewest non-leaks (negative incidents) as leaks. Therefore, those evaluation metrics are required to emphasize correct predictions and give overall algorithm performance measures on separating both classes. To this aim, accuracy, recall, precision, and mean F_1- measure are used as evaluation metrics. These metrics are based on values calculated in a confusion matrix (Figure 3.1).

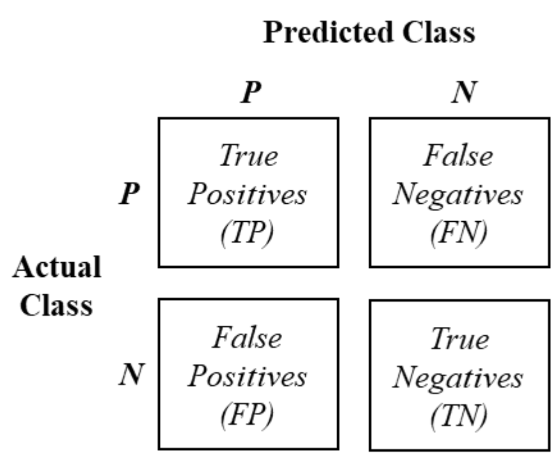

Figure 3.1. Confusion matrix for an evaluation of a classifier

Accuracy is the most common measure to evaluate a classifier and is defined as the degree of correct predictions of a model (or the percentage of misclassification errors).

$$Accuracy = \frac{tp + tn}{tp + fp + tn + fn} \quad (3.21)$$

In general, the accuracy metric measures the ratio of correct predictions over the total number of instances evaluated.

Recall is used to measure the ratio of positive classes that are correctly classified.

$$Recall = \frac{tp}{tp + fn} \quad (3.22)$$

Precision is a metric to measure the positive classes that are correctly predicted from the total predicted positive classes, either correctly or incorrectly.

$$Precion = \frac{tp}{tp + fp} \quad (3.23)$$

F_1-measure represents the harmonic mean between recall and precision values and is a metric to compare different algorithms based on the same emphasis on both recall and precision.

$$F_1 = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (3.24)$$

All of these four metrics have values between 0 and 1, where 0 and 1 represent the poorest and the best performance of an algorithm (Taormina and Galelli, 2018; Hossin and Sulaiman, 2015; Chicco and Jurman, 2020).

3.2.5. Methodology Overview

Figure 3.2 shows a flow diagram of the preprocessing, training, and test steps. NL, GL, OL, LL, and CL are no-leak, gasket leak, orifice leak, longitudinal leak, and circumferential leak, respectively. Moreover, H1 and H2 refer to hydrophone 1and hydrophone 2, CWT stands for continuous wavelet transform, and FE represents feature extraction.

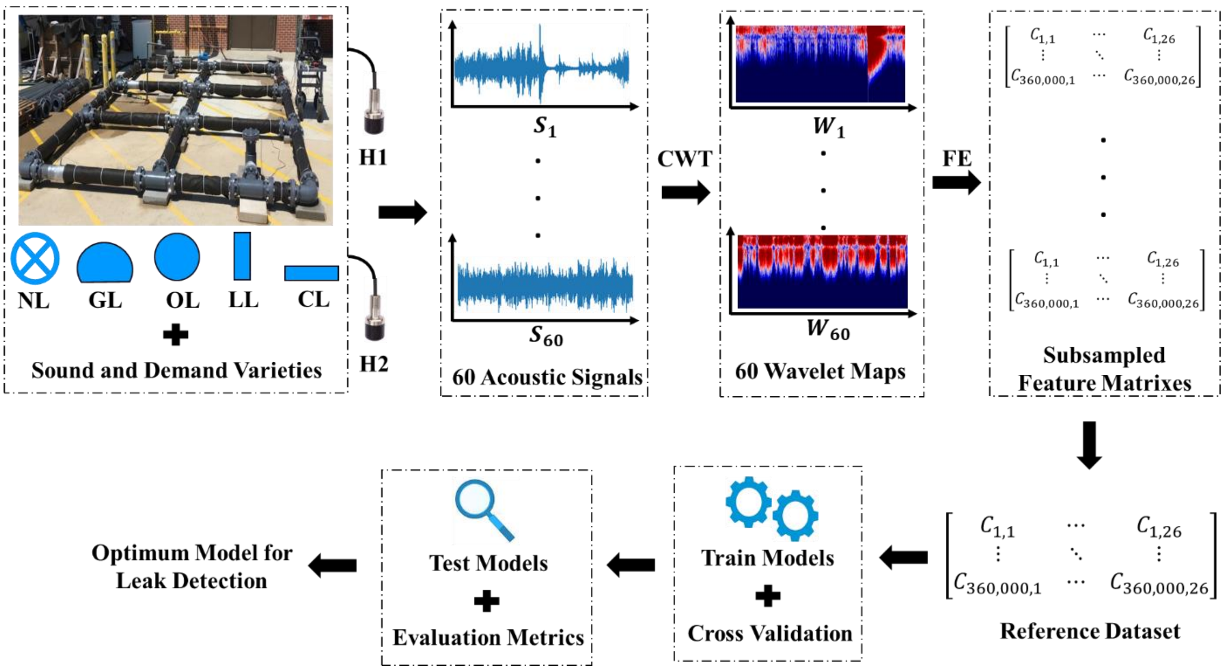

Figure 3.2. Methodology overview

Regarding Figure 3.2, a wavelet transform was utilized to extract time and frequency domain features from acoustic data acquired from H1 and H2. The features were subsampled to build a dataset for training and testing classification algorithms. Based on some evaluation metrics, algorithms with the best performances would be reported for leak detection.

3.3. Case Study

3.3.1. Experimental Setup

In this research, a laboratory scaled water distribution network with a looped network was developed, whose information is available in chapter 2 of this document. The testbed is a network composed of 15.24 cm diameter pipes as the distribution section is fed with a water supply line comprised of 2.54 cm diameter pipes. Figures 2.3 and 2.4 show real and schematic images of the network. In this section, the acoustic data of two hydrophones, H1 and H2 in Figure 2.4, are used for leak detection. The sensor data are 30 seconds long, with a sampling frequency of 8000 Hz. As mentioned in chapter 2, four types of leaks, i.e., orifice, longitudinal, circumferential, and gasket leaks were induced to simulate leak conditions. In addition, acoustic signals in a non-leak condition were used as a benchmark. According to Figure 2.34, the following scenarios were run for the looped network whose signals were applied for leak detection.

Table 3.2. Summary of scenarios used for leak detection

<table>
<tr><th>Leak Type</th><th>Scenario</th><th>Demand (GPM)</th><th>Sensor</th></tr>
<tr><td rowspan="12">Non-leak
Orifice
Longitudinal
Circumferential
Gasket</td><td rowspan="2">Demand and Sound</td><td>3.0</td><td rowspan="6">H1</td></tr>
<tr><td>7.5</td></tr>
<tr><td>No Demand and Sound</td><td rowspan="2">-</td></tr>
<tr><td>No Demand and No Sound</td></tr>
<tr><td>Transient and Sound</td><td rowspan="2">7.5 to 0.0</td></tr>
<tr><td>Transient and No Sound</td></tr>
<tr><td rowspan="2">Demand and Sound</td><td>3.0</td><td rowspan="6">H2</td></tr>
<tr><td>7.5</td></tr>
<tr><td>No Demand and Sound</td><td rowspan="2">-</td></tr>
<tr><td>No Demand and No Sound</td></tr>
<tr><td>Transient and Sound</td><td rowspan="2">7.5 to 0.0</td></tr>
<tr><td>Transient and No Sound</td></tr>
</table>

Regarding Table 3.2, twelve acoustic signals were acquired per leak type, which led to sixty measurements for the looped network. Based on the mentioned length and frequency values, each acoustic signal is composed of 240,000 samples. In the following, figures of non-leak and orifice leak hydrophone data are plotted, where ‘NL’ and ‘O’ stand for non-leak and orifice, respectively.

3.3.1.1. Scenarios with demand and sound

Figures 3.3 to 3.6 show time series of non-leak versus orifice leak acoustic signals of hydrophones H1 and H2 in the looped network with the presence of demand and ambient noise.

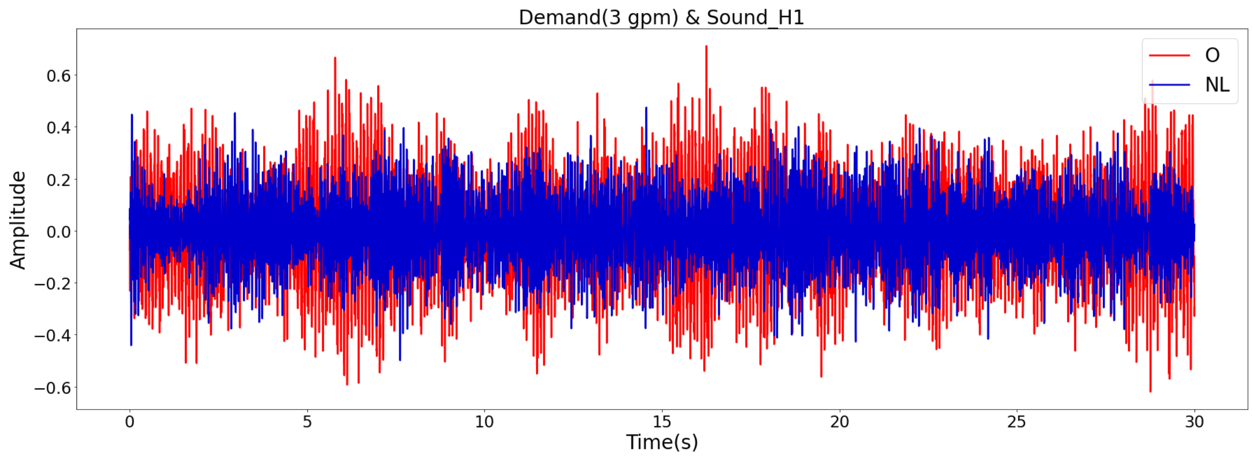

Figure 3.3. The acoustic signal measured by hydrophone H1, with 3 GPM flow at service valve and with ambient noise

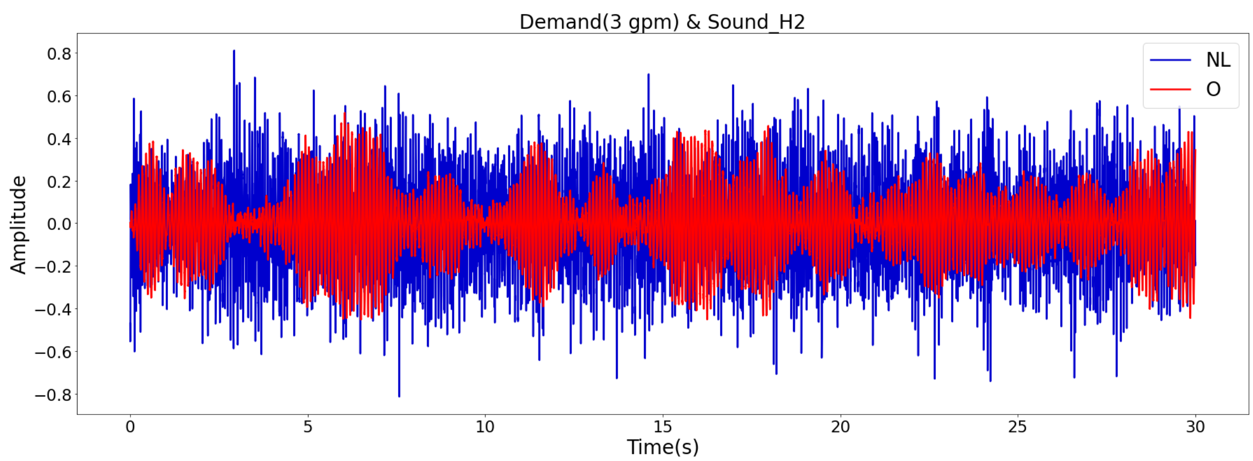

Figure 3.4. The acoustic signal measured by hydrophone H2, with 3 GPM flow at service valve and with ambient noise

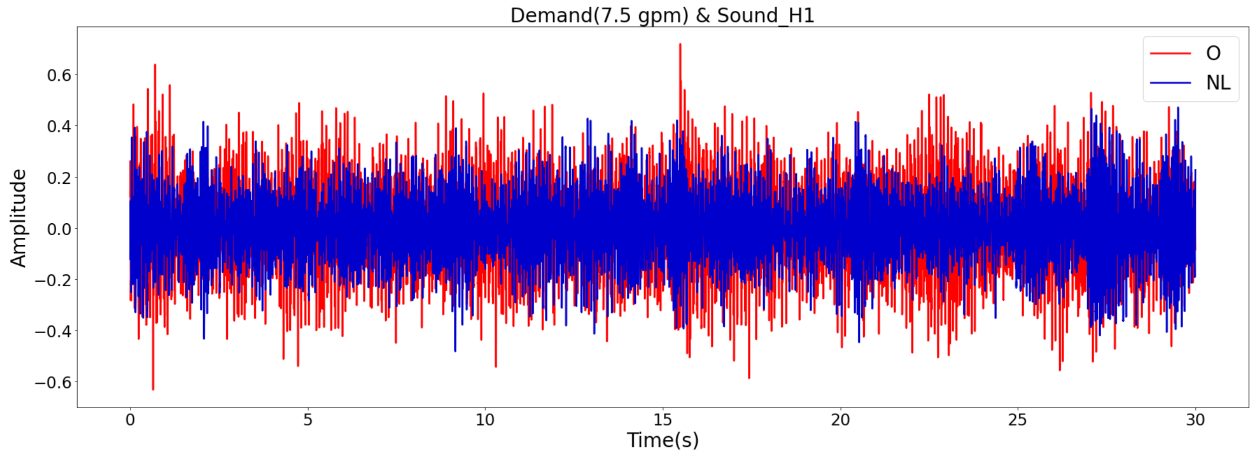

Figure 3.5. The acoustic signal measured by hydrophone H1, with 7.5 GPM flow at service valve and with ambient noise

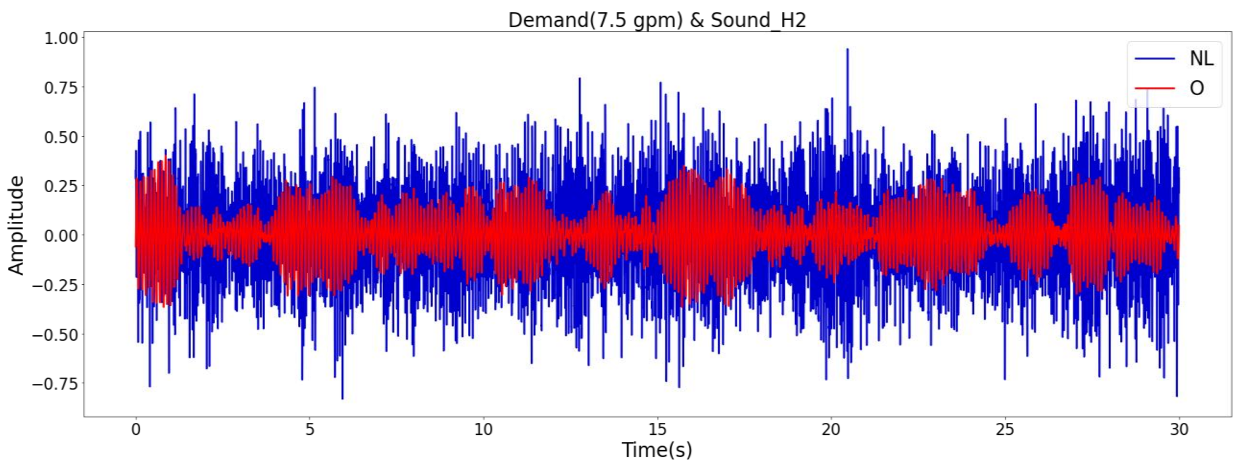

Figure 3.6. The acoustic signal measured by hydrophone H2, with 7.5 GPM flow at service valve and with ambient noise

3.3.1.2. Scenarios without demand and with sound

Figures 3.7 and 3.8 show time series of non-leak versus orifice leak acoustic signals of hydrophones H1 and H2 in the looped network without demand and with ambient noise.

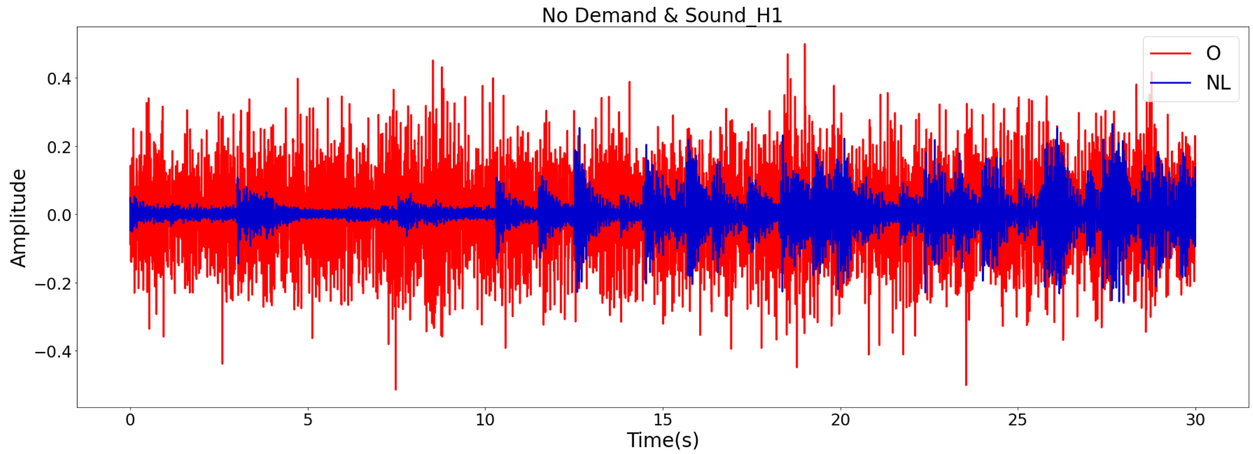

Figure 3.7. The acoustic signal measured by hydrophone H1, without flow at service valve and with ambient noise

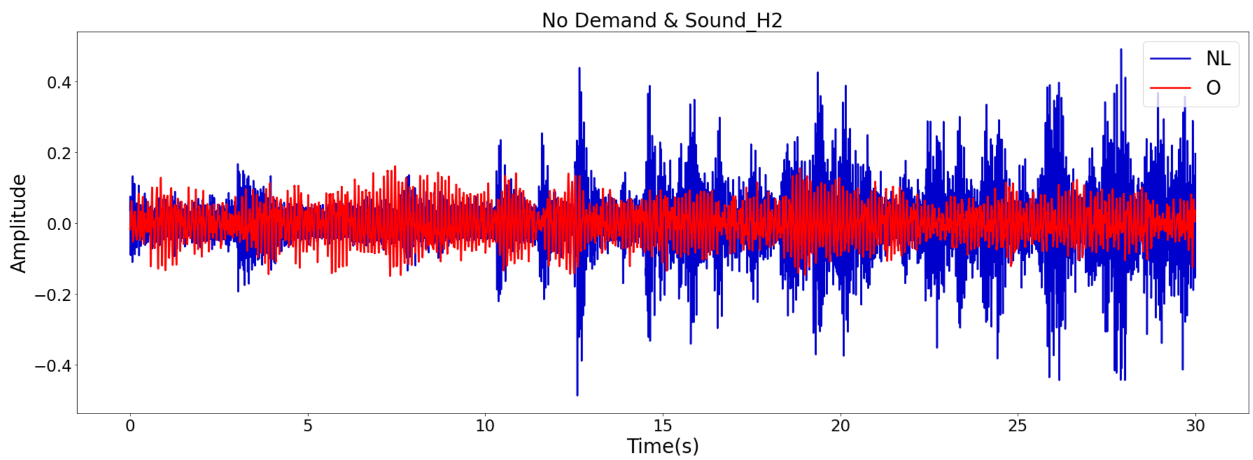

Figure 3.8. The acoustic signal measured by hydrophone H2, without flow at service valve and with ambient noise

3.3.1.3. Scenarios without demand and without sound

Figures 3.9 and 3.10 show time series of non-leak versus orifice leak acoustic signals of hydrophones H1 and H2 in the looped network without demand and without ambient noise.

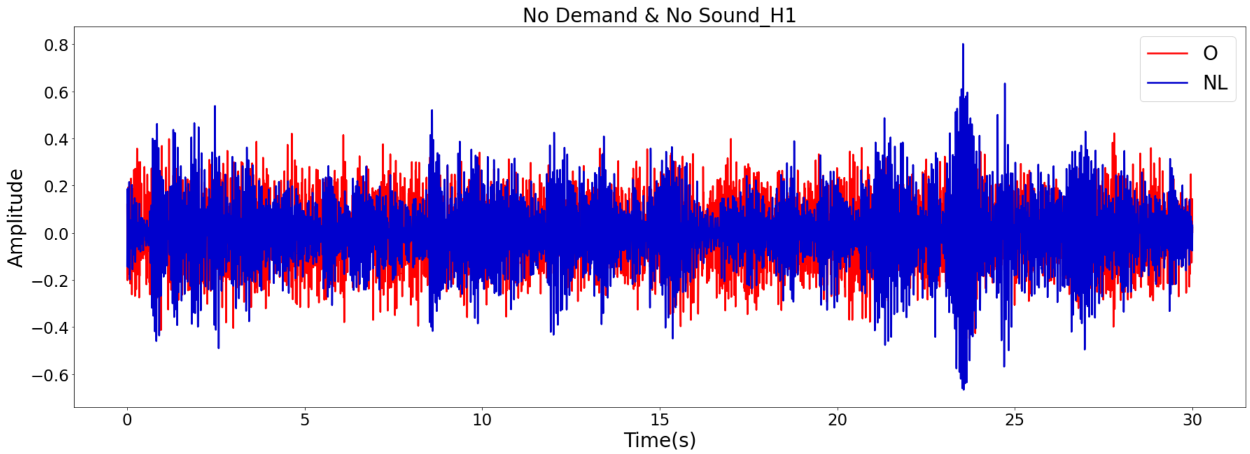

Figure 3.9. The acoustic signal measured by hydrophone H1, without flow at service valve and without ambient noise

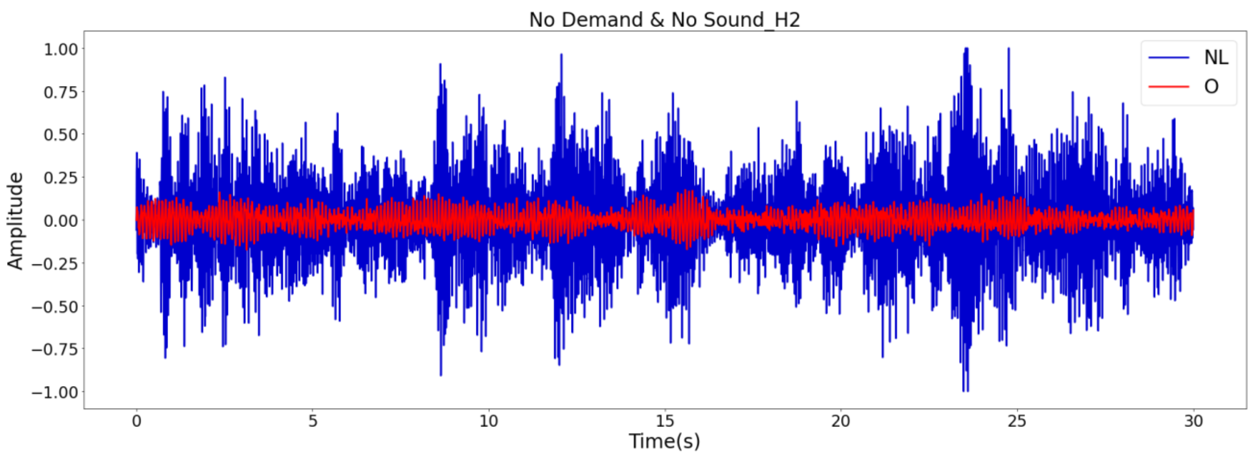

Figure 3.10. The acoustic signal measured by hydrophone H2, without flow at service valve and without ambient noise

3.3.1.4. Scenarios with the transient flow (7.5 GPM to 0 GPM) and with sound

Figures 3.11 and 3.12 show time series of non-leak versus orifice leak acoustic signals of hydrophones H1 and H2 in the looped network with transient demand flow, i.e., rapid flow change from 7.5 GPM to 0 GPM, and with ambient noise.

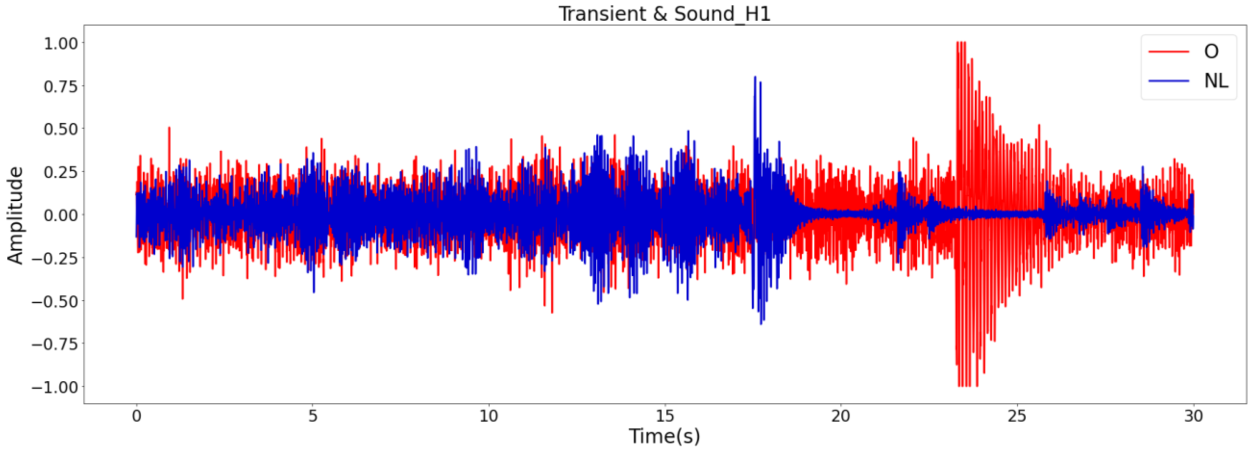

Figure 3.11. The acoustic signal measured by hydrophone H1, with the transient flow (rapid flow change from 7.5 GPM to 0 GPM) at service valve and with ambient noise

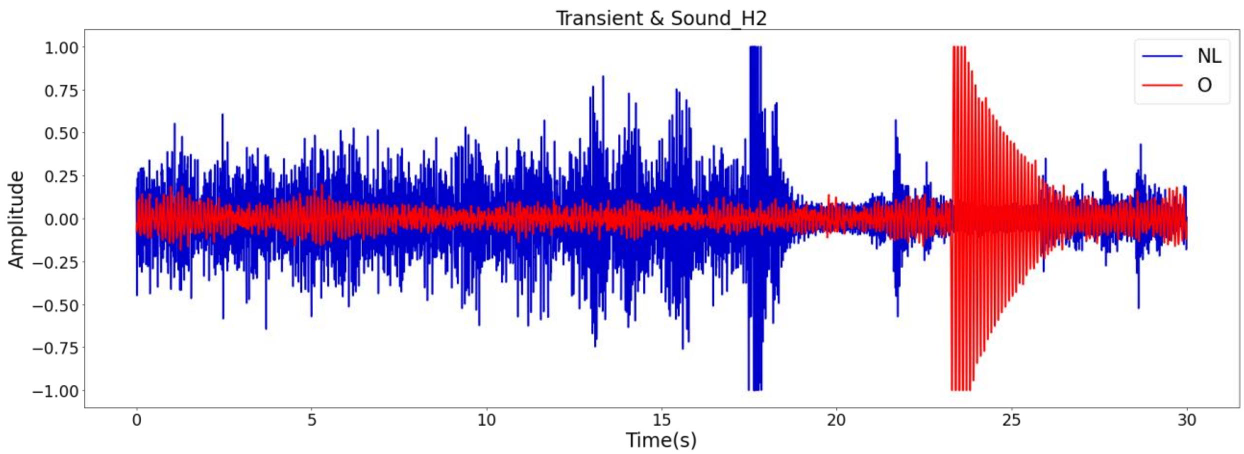

Figure 3.12. The acoustic signal measured by hydrophone H2, with the transient flow (rapid flow change from 7.5 GPM to 0 GPM) at service valve and with ambient noise

3.3.1.5. Scenarios with the transient flow (7.5 GPM to 0 GPM) and without sound

Figures 3.13 to 3.14 show the time series of non-leak versus orifice leak acoustic signals of hydrophones H1 and H2 in the looped network with transient demand flow, i.e., rapid flow change from 7.5 GPM to 0 GPM, and without ambient noise.

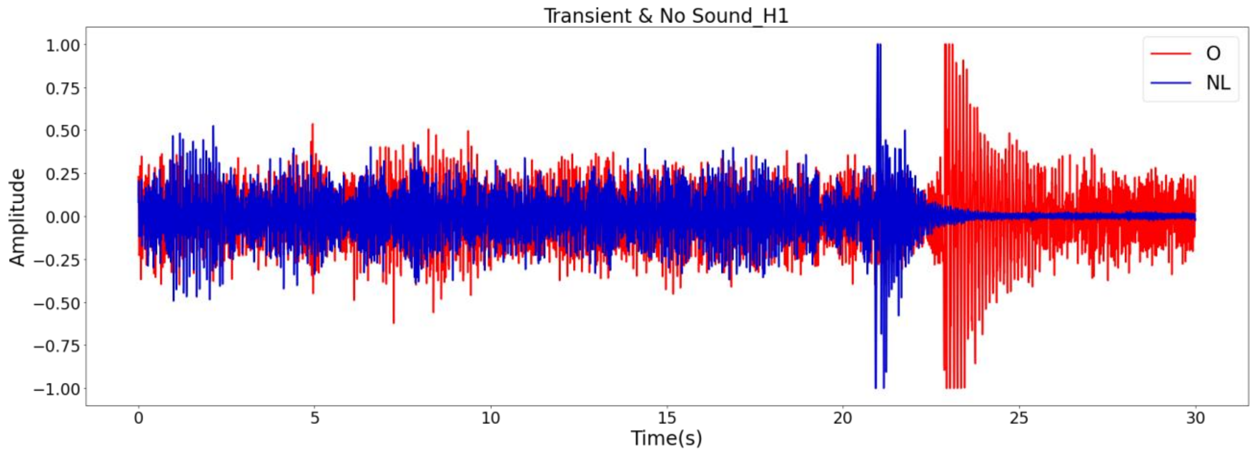

Figure 3.13. The acoustic signal measured by hydrophone H1, with the transient flow (rapid flow change from 7.5 GPM to 0 GPM) at service valve and without ambient noise

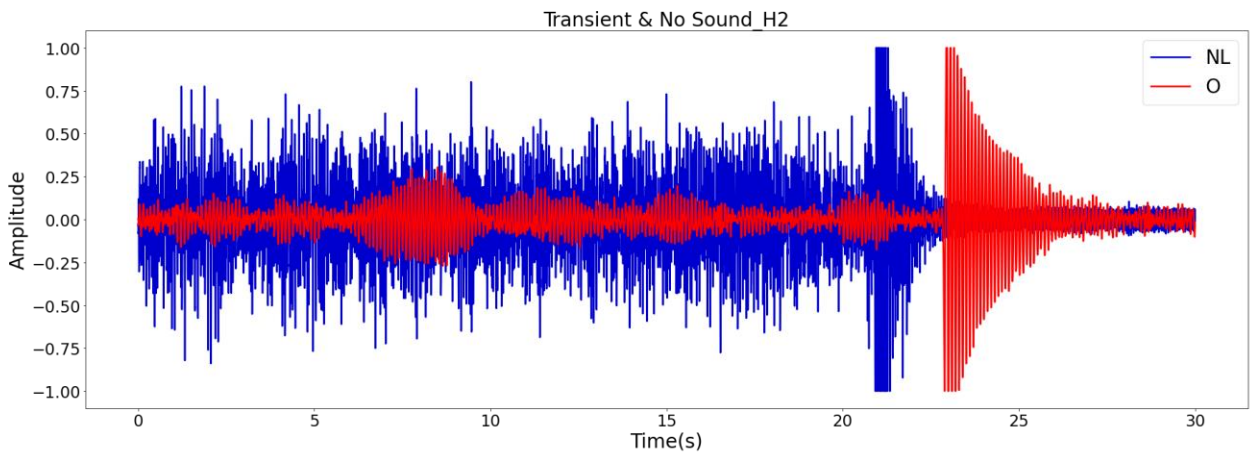

Figure 3.14. The acoustic signal measured by hydrophone H2, with the transient flow (rapid flow change from 7.5 GPM to 0 GPM) at service valve and without ambient noise

3.3.2. Extracting Features as Model Inputs

In this study, the complex Morlet mother wavelet, defined in Eq. (3.2), is used to transfer time-domain acoustic signals to a wavelet space and extract features as inputs to the classification algorithms. This type of wavelet has performed successfully in other research for feature extractions (Adamczyk et al., 2015; Montejo and Suarez, 2006; Cui et al., 2020; Soro and Lee, 2019). Two parameters determine the characteristics of a wavelet: central frequency (f_c) and bandwidth (f_b). Figure 3.14 shows the imaginary and real components of a Morlet wavelet with central frequency f_c = 2 and bandwidth f_b=14. This wavelet will be used to calculate the continuous wavelet transform defined in Eq. (3.1).

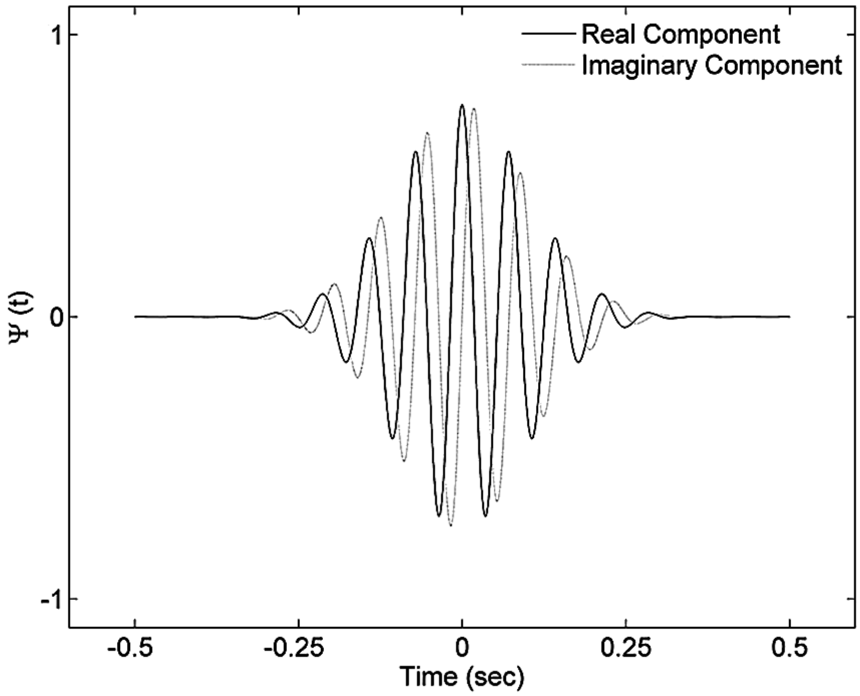

Figure 3.15. A Morlet wavelet with central frequency f_c = 2 and bandwidth f_b=14 (Reprinted from Adamczyk et al., 2015)

Regarding Figure 3.15, a mother wavelet is composed of sinusoids with different frequencies. The frequency variability of wavelets fits acoustic signals whose frequency contents change with leaks' occurrence or evolution within time. Since a wavelet transform is a method to measure a sinusoidal signal's similarity with a mother wavelet, wavelet transform coefficients indicate how similar a signal is with the selected mother wavelet.

In this study, a Morlet wavelet is required that meets the following conditions:

1- Since leak signals are concealed with the flow and ambient noises, the wavelet should have a large enough resolution such that leak and noise frequencies can be discerned well.

2- The wavelet's frequency contents need to highlight latent leak signals that are prevalent in frequencies less than 1000 Hz (Cody et al., 2018; Butterfield et al., 2017), but dominated by noises.

3- The wavelet's frequencies should conform to transient events that include a big range of frequencies.

4- The resolution of wavelets is better to be as low as possible to decrease time and computational complexities while calculating a signal's wavelet transform.

In the following, wavelet maps show how these four conditions were deemed to choose the best Morlet wavelet. In the below figures, the horizontal axis represents a signal's time. The vertical axis is the frequency contents of a sinusoidal acoustic signal. The colors show the modulus of wavelet coefficients. The wavelet coefficients are complex numbers that result from the application of Eq. (3.1) to a time-domain acoustic signal. According to the legends, the wavelet coefficient modulus magnitudes range from -6 to 0, where 0, corresponding to the red color, and -6, showing the dark blue color, represent the largest magnitude and the smallest magnitudes of the wavelet coefficient moduli, respectively.

Figures 3.16 and 3.17 are wavelet maps of acoustic signals measured by hydrophones H1 and H2, respectively, with the presence of the orifice leak in the looped network, a 3 GPM consumption flow and ambient noise. These figures are generated using a complex Morlet wavelet with f_b = 0.5 and f_c = 1.

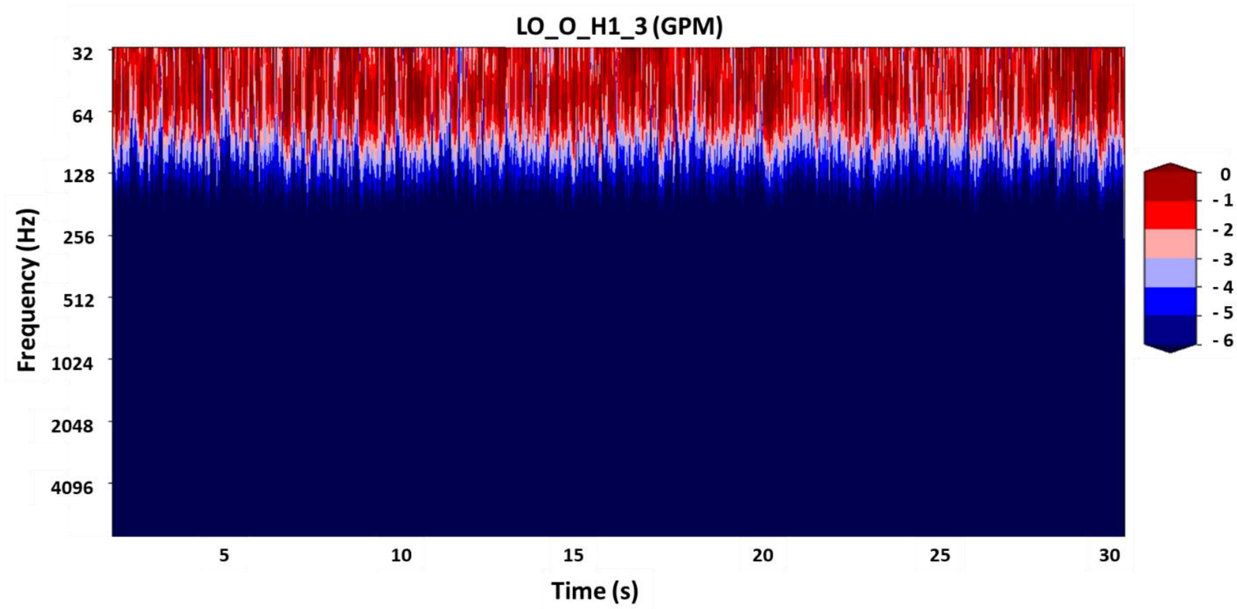

Figure 3.16. Wavelet map (f_b = 0.5 and f_c = 1) of the acoustic signal measured by H1 in the looped network with the orifice leak, a 3 GPM consumption flow, and ambient noises

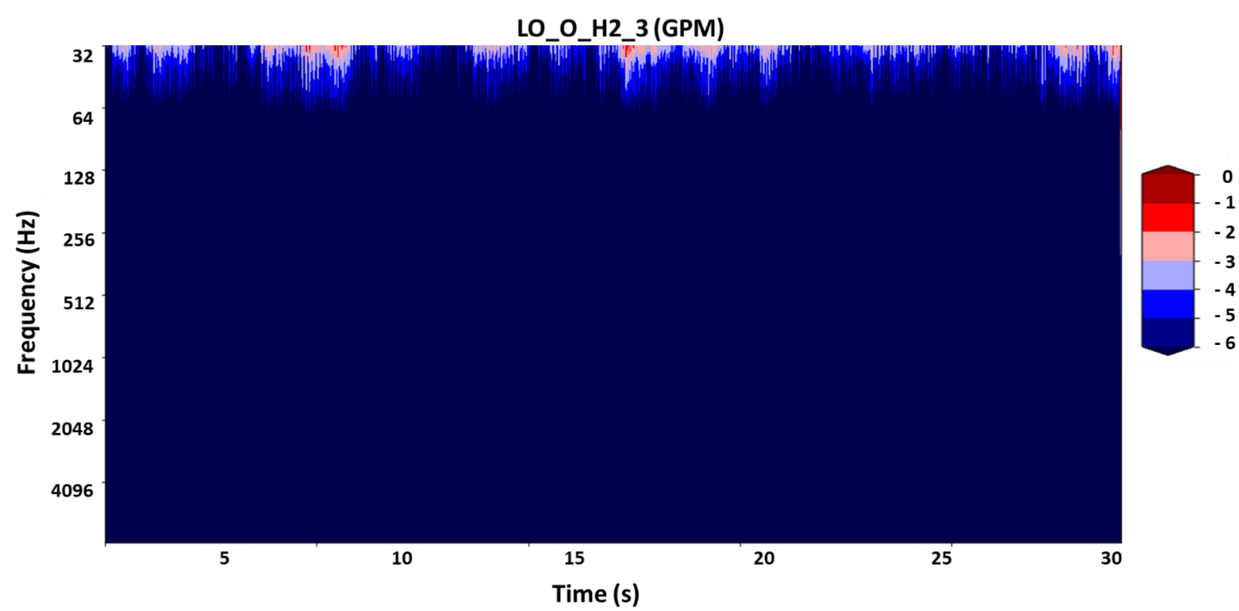

Figure 3.17. Wavelet map (f_b = 0.5 and f_c = 1) of the acoustic signal measured by H2 in the looped network with the orifice leak, a 3 GPM consumption flow, and ambient noises

Regarding Figure 3.17, one can observe that the wavelet with f_b = 0.5 and f_c = 1 is not able to well extract the waveform signal measured at hydrophone H2, which would result in uncaptured leak signals at low frequencies. A wavelet with much smaller central frequency might be a solution. Figures 3.18 and 3.19 show wavelet maps of acoustic signals recorded at H1 and H2 in the looped network with the orifice leak, where there is a transient consumption flow and an ambient noise. The Morlet wavelet has f_b = 0.5 and f_c = 0.125.

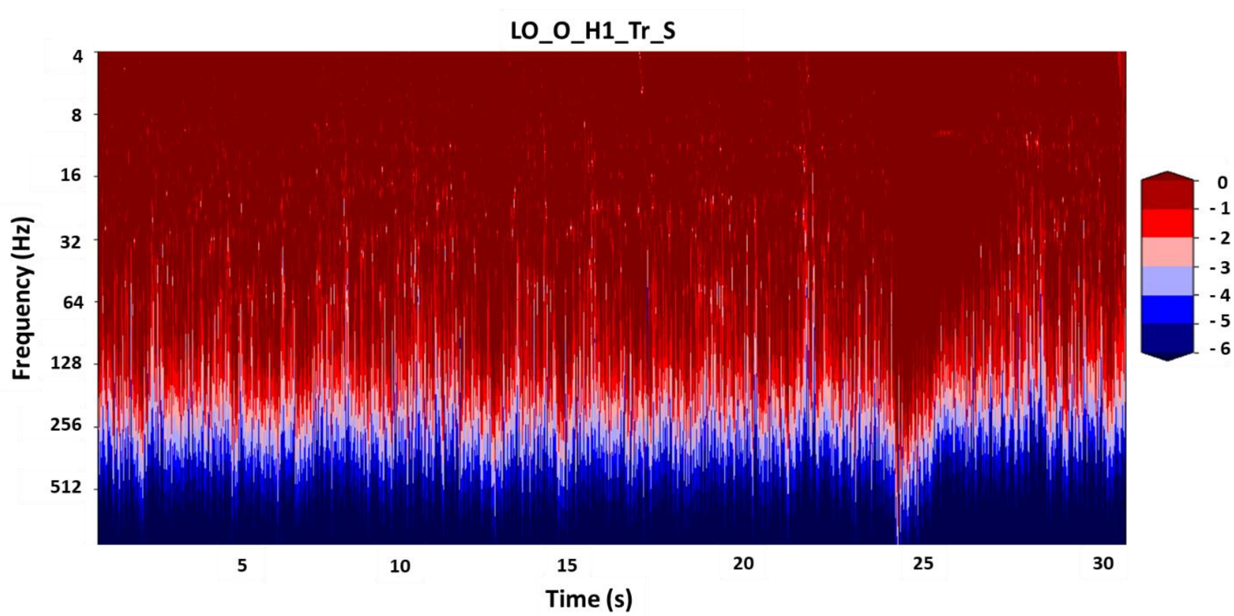

Figure 3.18. Wavelet map (f_b = 0.5 and f_c = 0.125) of the acoustic signal measured by H1 in the looped network with the orifice leak, a transient (7.5 GPM to 0 GPM) consumption flow, and ambient noises

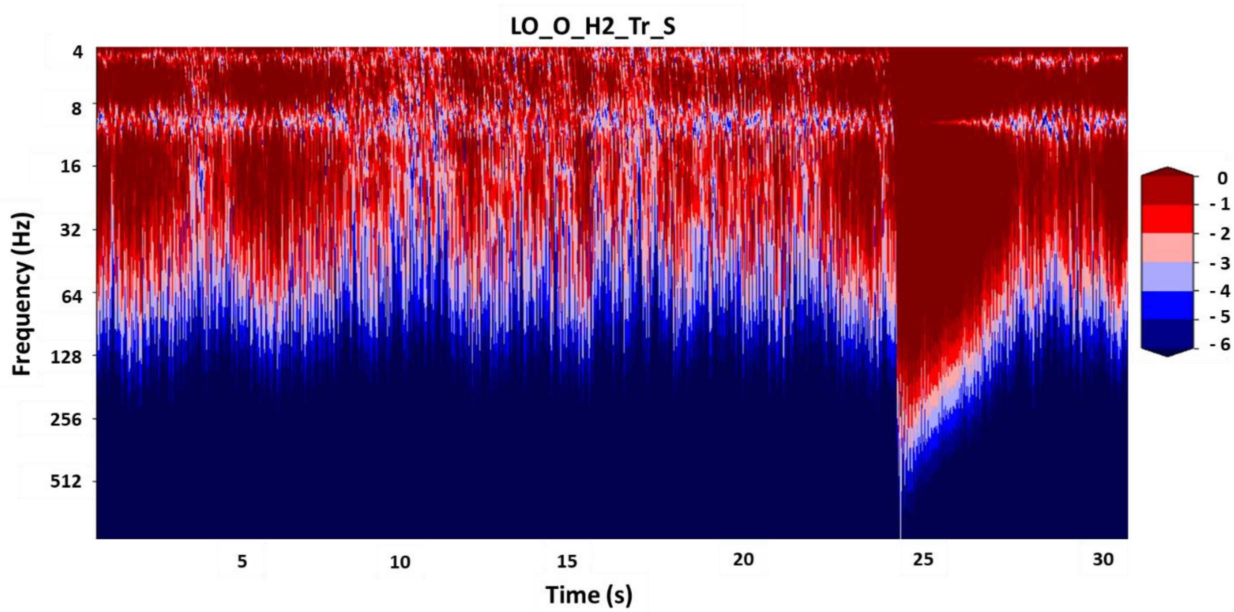

Figure 3.19. Wavelet map (f_b = 0.5 and f_c = 0.125) of the acoustic signal measured by H2 in the looped network with the orifice leak, a transient (7.5 GPM to 0 GPM) consumption flow, and ambient noises

Based on Figures 3.18 and 3.19, though a smaller central frequency presents more detailed frequency contents, it suffers from two drawbacks: (1) a very small central frequency focuses on low-frequency contents and does not extract wavelet coefficients at higher frequencies. Comparing the vertical axes of Figures 3.17 and 3.19, one can observe the latter has calculated wavelet coefficients of a smaller frequency range; (2) a wavelet with a very small central frequency is incapable of extracting wavelet coefficients of transient incidents which include a long continuum of frequencies. For example, at the time close to the second 25 in Figures 3.18 and 3.19, coefficient waves hit the maximum frequency. This limit might prevent a transform from well representing the wavelet transform coefficients at higher frequencies. Moreover, computing the wavelet maps of the Morlet wavelet using f_c = 0.125 was computationally expensive where it took about 10 minutes to plot each of Figures 3.18 and 3.19. The high computational overhead significantly slows down the feature extraction procedure that would increase the leak detection time complexity.

With trials and errors, the complex Morlet wavelet with f_b = 0.5 and f_c = 0.5 was found to be the most appropriate mother wavelet for all signals with different leak types. Figures 3.20 and 3.21 show the wavelet transforms of two different acoustic signals at hydrophone H2 of the looped network with an orifice leak. Figure 3.20 corresponds to a signal that was measured with disturbances caused by a 7.5 GPM consumption flow and ambient sounds. Therefore, for an accurate classification, enough resolution of wavelet coefficients at low frequencies are desired to capture discriminative feature information. In addition, Figure 3.21 is a wavelet map of a signal including a transient incident and ambient noises. Hence, the wavelet transform should be able to well represent wavelet coefficients at high frequencies where the transient event is dominant.

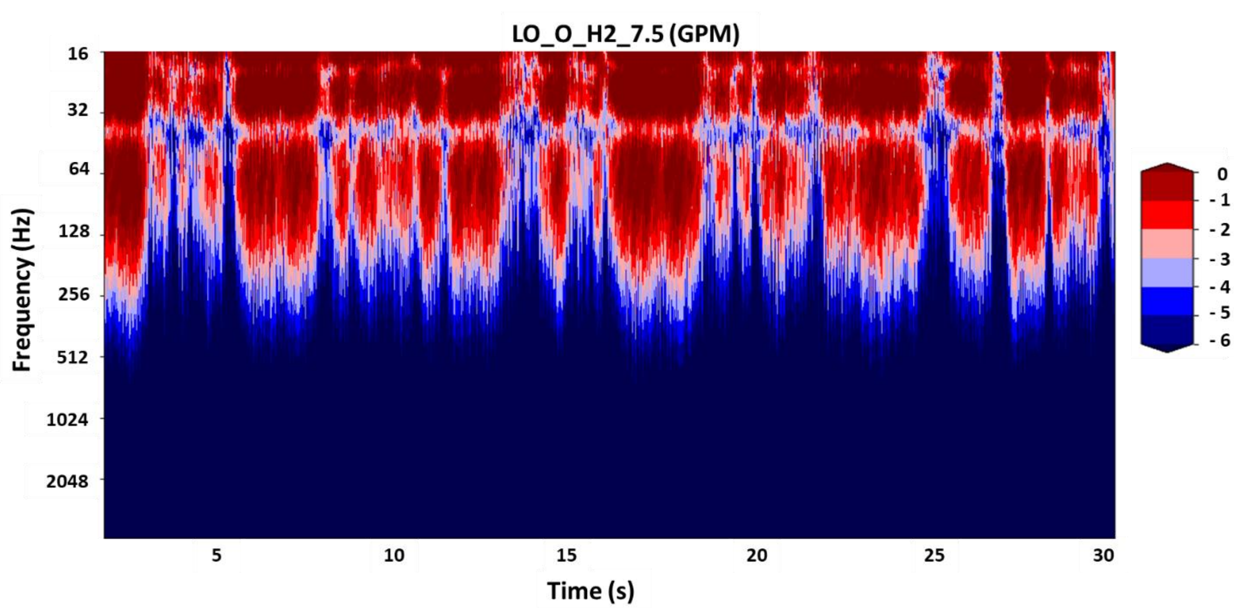

Figure 3.20. Wavelet map (f_b = 0.5 and f_c = 0.5) of the acoustic signal measured by H2 in the looped network with the orifice leak, a 7.5 GPM consumption flow and ambient noises

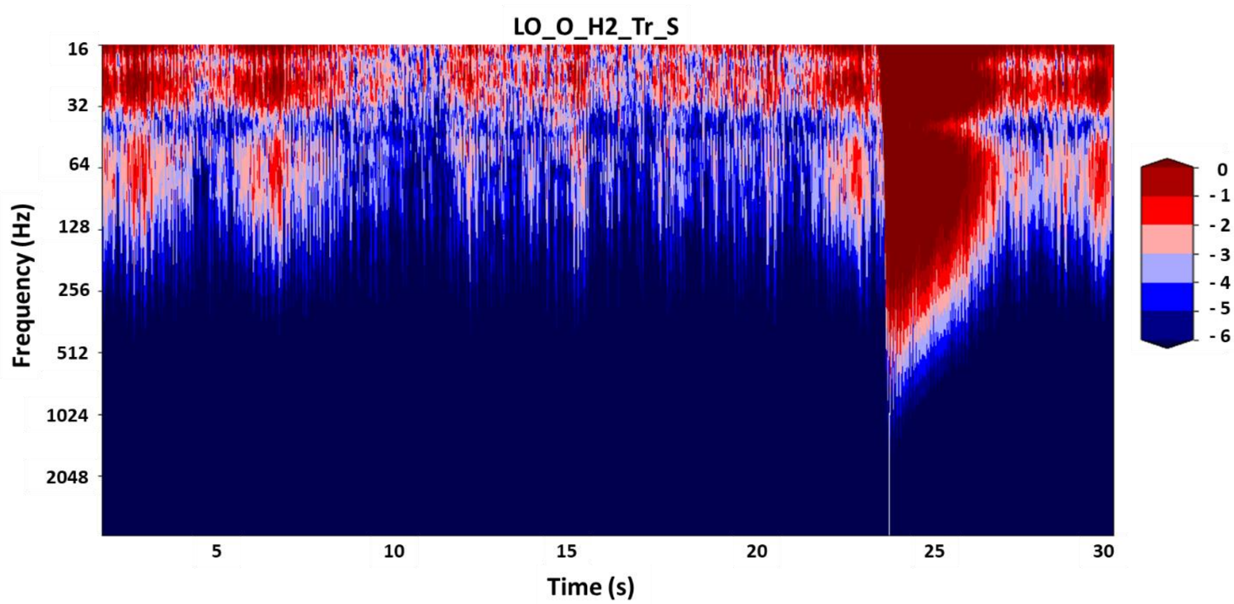

Figure 3.21. Wavelet map (f_b = 0.5 and f_c = 0.5) of the acoustic signal measured by H2 in the looped network with the orifice leak, a transient (7.5 GPM to 0 GPM) consumption flow, and ambient noises

Comparing Figures 3.19 and 3.21 shows a Morlet wavelet with f_b = 0.5 and f_c = 0.5 can satisfy the previously four mentioned conditions. Moreover, the following figures visually indicate that the wavelet transform can discern time-frequency differences of acoustic signals with the presence and absence of a leak. Figures 3.22 and 3.23 are wavelet maps of the acoustic data measured at hydrophone H1 of the looped network with 7.5 GPM consumption flows and ambient noises. The former and the latter figures represent conditions without and with an orifice leak, respectively. One can observe that the wavelet map of the signal with an orifice leak, Figure 3.23, is denser at frequencies under 256 Hz while the color of the wavelet coefficient magnitudes of the signal without a leak, Figure 3.22, is less unform at those frequencies. This difference can stem from the presence of leak signals in Figure 3.23 that stand out with large wavelet coefficient magnitudes and make magnitude colors brighter.

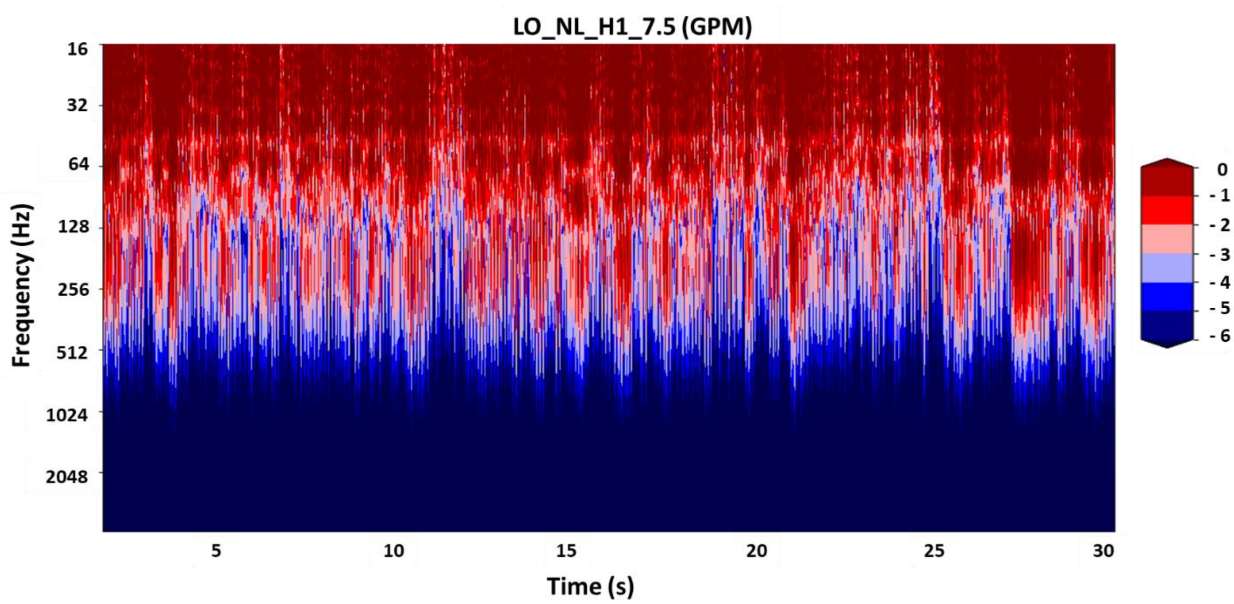

Figure 3.22. Wavelet map of the acoustic signal measured at hydrophone H1 of the looped network without a leak, with a 7.5 GPM consumption flow and ambient noises

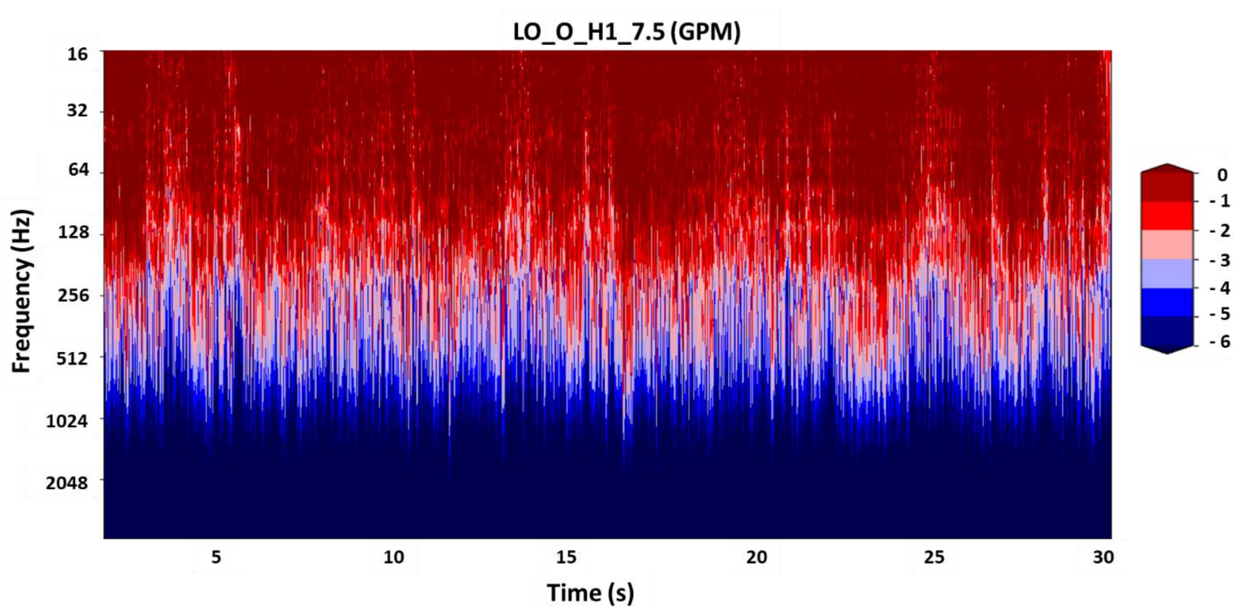

Figure 3.23. Wavelet map of the acoustic signal measured at hydrophone H1 of the looped network with an orifice leak, a 7.5 GPM consumption flows, and ambient noises

Noises generated by consumption flows and ambient sounds introduce undesired disturbances in acoustic signals that make the leak detection task more challenging. Bandpass filters have been employed in different papers to remove noises and extract frequency bandpass of interest relevant to leaks. Table 3.3 lists filters used as a preprocessing technique in different studies.

Table 3.3. Characteristics of filters as preprocessing techniques employed in leak detection studies

Author(s)	Filter Details
Butterfield et al., 2017	Used Hanning window and 10^{th} order Butterworth filters to remove signals greater than 1000 Hz.
Hunaidi and Chu, 1999	Employed 200 Hz cut-off frequency to remove noises and a bandpass filter, between 15 and 100 Hz, before a cross-correlation analysis using Fourier transform.
Almeida et al., 2014	Geophone and accelerometer data were passed through bandpass filters with lower and upper limits set to 10 Hz and 150 Hz, respectively.
Gao et al, 2005	Used fourth order Butterworth filters where the lower and upper cut-off frequencies were set at 10 and 50 Hz for hydrophone-measured signals, and 30 and 140 Hz for accelerometer-measured signals.
Martini et al., 2015	A 200-600 Hz bandpass filter increased leak detection method via using signal power spectrum.
Kang et al., 2018	Firstly, used an antialiasing filter with a 1-kHz cutoff frequency to limit the high-frequency components, then applied a bandpass filter ranging from 100 to 800 Hz to extract the frequency band of interest caused by leaks.
Harmouche and Narasimhan, 2020	Employed a high-pass filter to remove the low-frequency components (less than 2 Hz)
Guo et al., 2021	Used a bandpass filter with a frequency response ranging from 100 to 2,000 Hz was used to extract the frequency band of interest caused by leaks.

Regarding Table 3.3, Butterfield et al. (2017), Kang et al. (2018), and Guo et al. (2021) used bandpass filters with large upper cut-offs, between 1000 Hz to 2000 Hz, to remove high-frequency noises. While Hunaidi and Chu (1999), Gao et al. (2005), and Almeida et al. (2014) employed bandpass filters to extract signals with low frequencies, approximately between 10 Hz and 200 Hz, Harmouche and Narasimhan (2020) focused on high-frequency signals by employing a high-pass filter to remove low-frequency components. In the following, it will be evaluated if using a filter can prepare better signals for feature extractions and classification algorithms.

Regarding the Butterworth bandpass filter's popularity, a 5th order Butterworth filter with 30 Hz and 2000 Hz cut-off frequencies is applied to the acoustic signal measured at hydrophone H2 in the looped network with the orifice leak, without a consumption flow and with the absence of ambient noises. Figure 3.24 shows the wavelet map of a filtered signal that has been generated by using a Morlet wavelet with f_b = 0.5 and f_c = 0.5. In addition, Figure 3.25 shows the wavelet map of the same signal but without being filtered, generated by a similar Morlet wavelet.

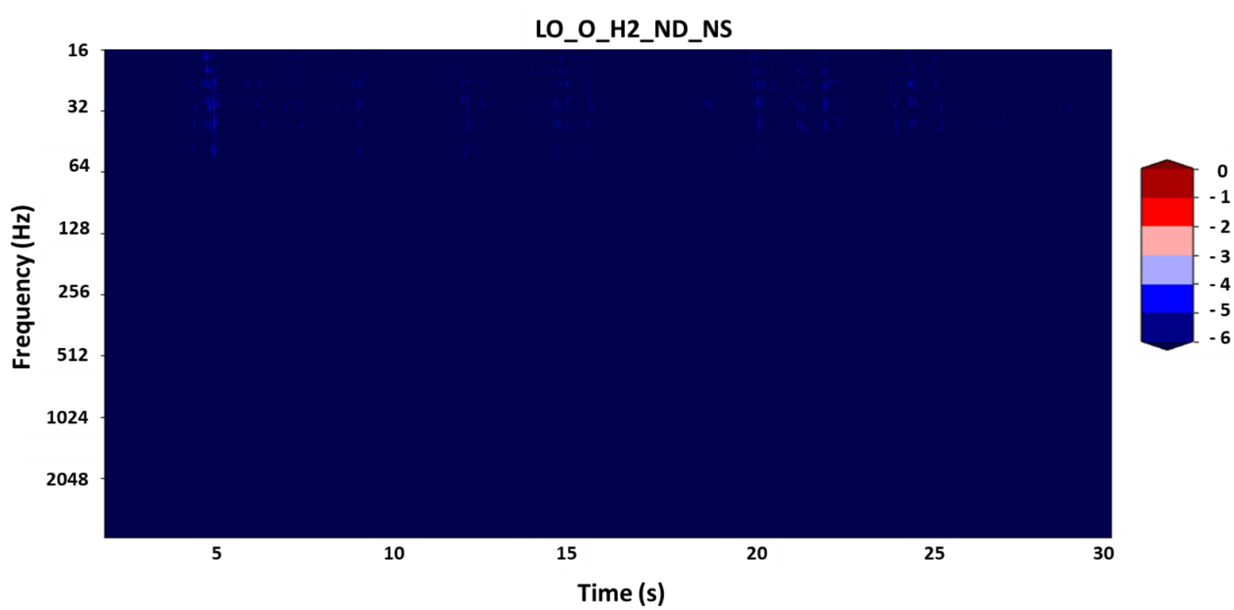

Figure 3.24. Wavelet map of an acoustic signal filtered by a 5th order Butterworth filter with 30 Hz and 2000 Hz cut-off frequencies where the signal is measured at hydrophone H2 in the looped network with the orifice leak, without a consumption flow and with the absence of ambient noises

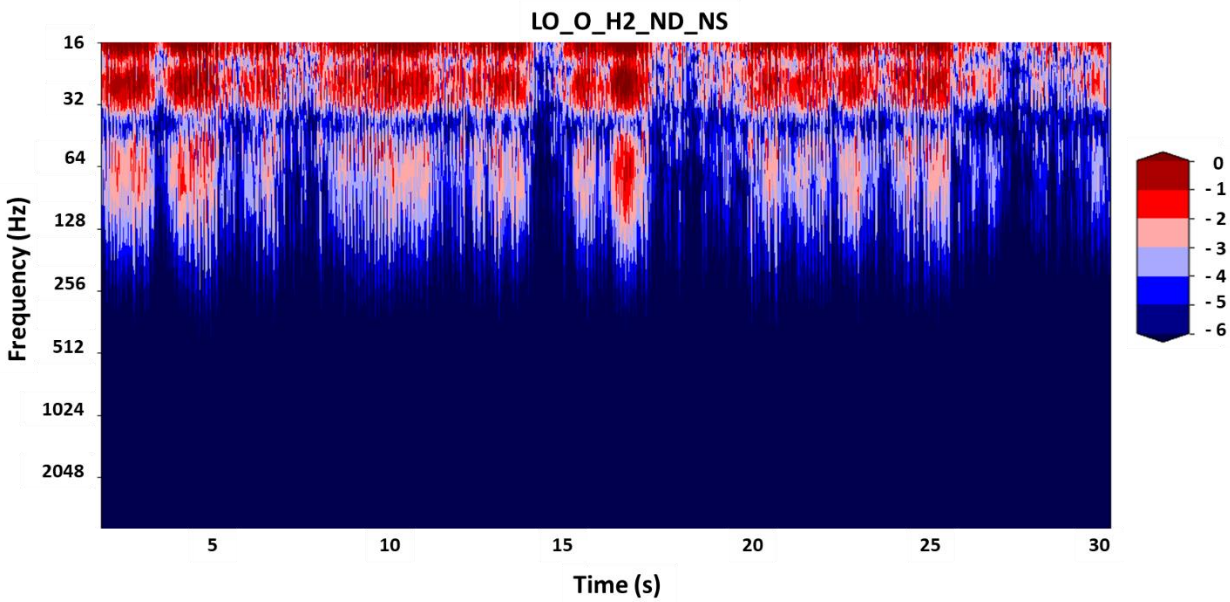

Figure 3.25. Wavelet map of an unfiltered acoustic signal measured at hydrophone H2 in the looped network with the orifice leak, without a consumption flow and with the absence of ambient noises

Comparing Figures 3.24 and 3.25 indicate that the filter has drastically decreased the magnitudes of the signal's wavelet transform coefficient. Though these filters are expected to remove undesired signal contents under or above the bandpass cut-offs, Figure 3.24 shows that the applied filter has removed the majority wavelet coefficients. It is worth emphasizing the signal of Figures 3.24 and 3.25 does not include any consumption flow or environmental noises. Therefore, it is reasonable to assume that the signal contents and the wavelet map in Figure 3.25 mostly represent leak signals, compared to other signals that include noises. As a result, regarding Figure 3.24, the filter has removed the signal contents that are out of bandpass cutoffs and are 'undesired', yet they can embed frequencies that represent the orifice leak signatures. This reason may be convincing enough to mention that applying bandpass filters on this research's acoustic signals may cause the risk of eliminating leak frequencies.

As a recap the objective of applying the wavelet transform was extracting well representing features from acoustic signals as inputs for classification algorithms. The following explains how feature matrices are built based on the wavelet maps.

To plot the wavelet maps, the PyWavelet package, an open-source wavelet transform software for Python programming language developed by Lee et al. (2019), was used to calculate continuous wavelet transforms (CWTs). The package uses a function named “pywt” that takes a signal as data, non-dimensional scales, a mother wavelet characteristic, and a signal’s period as inputs. The function returns frequencies corresponding to the scales and complex numbers of the CWT at the frequencies (Lee et al., 2019). The function uses another function titled “scale2frequency” to convert scales to frequencies using Eq. (3.25) (Lee et al., 2019).

$$Frequency = scale2frequency(wavelet, scale)/sp \tag{3.25}$$

where *sp* is the sampling period of the input signal, 0.00125 (s) $= {}^{1}/_{8000\ (Hz)}$ in his case study.

In this research, the scale range of (1, 513) showed to provide enough resolution and to be computationally efficient for computing wavelet transforms of the looped network acoustic signals. With that being said, magnitudes of the previous wavelet maps are moduli of the wavelet transform coefficients computed at frequencies corresponding to the scale values in (1, 513). Based on the (1, 513) scale range and the Morlet wavelet with f_b = 0.5 and f_c = 0.5, the “pywt” function returns wavelet transform coefficients for frequencies between 7.8125 Hz to 4000 Hz. The 7.8125 Hz to 4000 Hz frequency range would be enough for a leak detection study since research have reported the frequency of leak signals are smaller than 1000 Hz (Cody et al, 2018; Hunaidi and Chu, 1999; Kange et al., 2018; Butterfield et al. 2017).

The reason that the vertical axis labels of Figures 3.20 to 3.25 are within the frequencies 16 Hz and 2048 Hz is the visualization memory limits of the machine used for plotting. However, the "pywt" function employed to plot these figures returns 512 wavelet coefficients for frequencies between 7.8125 Hz and 4000 Hz. The number 512 follows the length of scale range (1, 513) determined by a user, i.e., one frequency per scale.

Regarding the 30 s long acoustic signals and their 8000 Hz sampling frequency, each signal includes 240,000 samples. Therefore, based on the mentioned signal's length and period, scale range, and the Morlet wavelet characteristics, the CWT generates a matrix with a dimension of [512 × 240,000] whose elements are moduli of the complex wavelet coefficients. This results in a feature matrix with 122,880,000 elements for each signal. In the next subsection, it will be described that the coefficients' moduli of 60 signals will be used to build feature matrices to train and test machine learning algorithms. If the wavelet coefficients of all frequencies from 60 signals will be used, a total of (7,372,800,000 = 60 × 122,880,000) elements should be calculated to build training and test datasets for machine learning algorithms. Such extensive calculations need very large computational resources and many hours. Therefore, it is necessary to decrease the number of elements at each signal's feature matrix.

Speculating the previous wavelet maps show that there are regions where the maps' colors do not change. This means constant or very similar values at different rows or columns of a signal's feature matrix. Therefore, we subsampled the feature matrix of each signal. Those elements that show larger magnitudes of wavelet coefficients were sampled with more elements and few of those elements that have negligible values are kept. Therefore, each signal's new feature matrix has fewer rows and columns. Since leak signals are more probable to exist in low frequencies, more elements at low frequency regions were sampled. In contrary, fewer rows and columns were kept for the new feature matrix. Since the subsampling frequency at different rows were not similar, the elements of the new feature matrix that did not have values were filled with zero.

Figures 3.26 and 3.27 show examples of an original and a new subsampled feature matrix extracted from an acoustic signal, where $C_{i,j}$ is a feature matrix element for a frequency *i* and a feature *j*.

$$
\begin{array}{c}
7.8125\text{ (Hz)} \\
7.8275\text{ (Hz)} \\
8\text{ (Hz)} \\
\vdots \\
1536\text{ (Hz)} \\
2048\text{ (Hz)} \\
4000\text{ (Hz)}
\end{array}
\begin{bmatrix}
C_{1,1} & C_{1,2} & \cdots & C_{1,239,999} & C_{1,240,000} \\
C_{2,1} & C_{2,2} & \cdots & C_{2,239,999} & C_{2,240,000} \\
C_{3,1} & C_{3,2} & \cdots & C_{3,239,999} & C_{3,240,000} \\
\vdots & \vdots & \vdots & \vdots & \\
C_{510,1} & C_{510,2} & \cdots & C_{510,239,999} & C_{510,240,000} \\
C_{511,1} & C_{511,2} & \cdots & C_{511,239,999} & C_{511,240,000} \\
C_{512,1} & C_{512,2} & \cdots & C_{512,239,999} & C_{512,240,000}
\end{bmatrix}
$$

Figure 3.26. The original feature matrix of a signal with 512 rows and 240,000 columns

$$
\begin{array}{c}
8\text{ (Hz)} \\
10\text{ (Hz)} \\
12\text{ (Hz)} \\
\vdots \\
1536\text{ (Hz)} \\
2048\text{ (Hz)} \\
4000\text{ (Hz)}
\end{array}
\begin{bmatrix}
C_{1,1} & C_{1,2} & \cdots & C_{1,29,999} & C_{1,30,000} \\
C_{2,1} & C_{2,2} & \cdots & C_{2,29,999} & C_{2,30,000} \\
C_{3,1} & C_{3,2} & \cdots & C_{3,29,999} & C_{3,30,000} \\
\vdots & \vdots & \vdots & \vdots & \\
C_{24,1} & C_{24,2} & \cdots & C_{24,29,999} & C_{24,30,000} \\
C_{26,1} & C_{25,2} & \cdots & C_{25,29,999} & C_{25,30,000} \\
C_{26,1} & C_{26,2} & \cdots & C_{26,29,999} & C_{26,30,000}
\end{bmatrix}
$$

Figure 3.27. The subsampled feature matrix of a signal with 26 rows and 30,000 columns

Due to the different subsampling rates at each selected frequency, the number of non-zero feature elements decreases with an increase in frequencies. For example, all 30,0000 feature elements at the lowest frequency, i.e., 8 Hz, are non-zero values. However, only 120 feature elements at the maximum selected frequency, i.e., 4000 Hz, have non-zero values, and the remaining elements, i.e., the elements between 120 and 30,000, were filled with zero. Figure 3.28 shows the real structure of a subsampled feature matrix.

$$
\begin{array}{c}
8\ (\mathrm{Hz}) \\ 10\ (\mathrm{Hz}) \\ 12\ (\mathrm{Hz}) \\ \cdot \\ \cdot \\ \cdot \\ 1536\ (\mathrm{Hz}) \\ 2048\ (\mathrm{Hz}) \\ 4000\ (\mathrm{Hz})
\end{array}
\begin{bmatrix}
C_{1,1} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & C_{1,30,000} \\
C_{2,1} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & C_{2,26,250} & 0 \\
C_{3,1} & \cdots & \cdots & \cdots & \cdots & \cdots & C_{3,22,500} & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
C_{24,1} & \cdots & \cdots & \cdots & C_{24,330} & 0 & 0 & 0 & 0 \\
C_{25,1} & \cdots & \cdots & C_{25,240} & 0 & 0 & 0 & 0 & 0 \\
C_{26,1} & \cdots & C_{26,120} & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

Figure 3.28. The real structure of a subsampled feature matrix based on the required samples at each selected frequency

It is worth noting that the required number of elements at each frequency determined the subsampling details. For instance, since 30,000 samples were desired at the 8 (Hz) frequency and regarding the total 240,000 samples per frequency in the original matrix, 30,000 values were chosen from the total 240,000 samples with intervals of 8. This sampling interval for the frequency 4000 (Hz), whose required number of samples was 120, is 2000. With the subsampling, the number of elements in a signal's new feature matrix is (780,000 = 26 × 30,000) with many zero-value elements, which is a 99% size decrease compared to the original feature matrix with 122,880,000 elements.

Regarding the point that leak signals have been reported to happen at lower frequencies mainly, more emphasis has been put on the subsampled feature matrices' lower frequencies. Table 3.4 lists the selected frequencies to generate a feature matrix for each signal. Evaluating Table 3.4 shows that 17 frequencies out of the 26 selected frequencies (about 65%) have values less than 150 Hz.

Table 3.4. Selected frequencies used in a subsampled feature matrix for each signal

No.	Frequency (Hz)
1	8
2	10
3	12
4	14
5	16
6	20
7	24
8	28
9	32
10	40
11	48
12	56
13	64
14	80
15	96
16	112
17	128
18	192
19	256
20	384
21	512
22	768
23	1024
24	1536
25	2048
26	4000

To build training and test datasets for algorithm training, a reference feature matrix for each leak type should be generated, including a subsampled feature matrix per signal. Table 3.2

summarizes the leak types and signals used for creating the training and test datasets. Regarding the frequencies and the structure of the subsampled feature matrix explained earlier, a reference feature matrix for each leak type with twelve signals has the format of Table 3.5.

Table 3.5. Feature matrix generated from twelve signals for each leak type

Frequency (Hz)	S_1	…	S_{12}	Total Features per Frequency
8	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
10	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
12	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
14	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
16	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
20	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
24	$C_{1,1}, \dots, C_{1.30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
28	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
32	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
40	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
48	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
56	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
64	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
80	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
96	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
112	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
128	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
192	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
256	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
384	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
512	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
768	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
1024	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{1,230,000}$	360,000
1536	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
2048	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
4000	$C_{1,1}, \dots, C_{1,30,000}$	…	$C_{12,1}, \dots, C_{12,30,000}$	360,000
			Total Features	9,360,000

In Table 3.5, $C_{i,j}$ is the j^{th} wavelet coefficient's modulus of signal *i* at a selected frequency and S_1,

…, S_{12}, indicate signals 1 to 12 recorded for each leak type. Yet, regarding the real structure of a subsampled feature matrix for each signal, see Figure 3.28, the non-zero elements of a feature matrix for each leak type is according to Table 3.6.

Table 3.6. Non-zero elements of a feature matrix generated for each leak type

Frequency (Hz)	S_1	…	S_{12}	Total Non-zero Features per Frequency
8	$C_{1,1},\dots,C_{1,30,000}$	…	$C_{12,1},\dots,C_{12,30,000}$	360,000
10	$C_{1,1},\dots,C_{1,26,250}$	…	$C_{12,1},\dots,C_{12,26250}$	315,000
12	$C_{1,1},\dots,C_{1,22,500}$	…	$C_{12,1},\dots,C_{12,22,500}$	270,000
14	$C_{1,1},\dots,C_{1,18750}$	…	$C_{12,1},\dots,C_{12,18750}$	225,000
16	$C_{1,1},\dots,C_{1,15,000}$	…	$C_{12,1},\dots,C_{12,15,000}$	180,000
20	$C_{1,1},\dots,C_{1,13110}$	…	$C_{12,1},\dots,C_{12,13110}$	157,320
24	$C_{1,1},\dots,C_{1.11250}$	…	$C_{12,1},\dots,C_{12,11250}$	135000
28	$C_{1,1},\dots,C_{1,9360}$	…	$C_{12,1},\dots,C_{12,9360}$	112,320
32	$C_{1,1},\dots,C_{1,7500}$	…	$C_{12,1},\dots,C_{12,7500}$	90,000
40	$C_{1,1},\dots,C_{1,6570}$	…	$C_{12,1},\dots,C_{12,6570}$	78,840
48	$C_{1,1},\dots,C_{1,5610}$	…	$C_{12,1},\dots,C_{12,5610}$	67,320
56	$C_{1,1},\dots,C_{1,4680}$	…	$C_{12,1},\dots,C_{12,4680}$	56,160
64	$C_{1,1},\dots,C_{1,3750}$	…	$C_{12,1},\dots,C_{12,3750}$	45,000
80	$C_{1,1},\dots,C_{1,3270}$	…	$C_{12,1},\dots,C_{12,3270}$	39,240
96	$C_{1,1},\dots,C_{1,2,820}$	…	$C_{12,1},\dots,C_{12,2,820}$	33,840
112	$C_{1,1},\dots,C_{1,2340}$	…	$C_{12,1},\dots,C_{12,2340}$	28,080
128	$C_{1,1},\dots,C_{1,1860}$	…	$C_{12,1},\dots,C_{12,1860}$	22,320
192	$C_{1,1},\dots,C_{1,1,650}$	…	$C_{12,1},\dots,C_{12,1,650}$	19,800
256	$C_{1,1},\dots,C_{1,1,410}$	…	$C_{12,1},\dots,C_{12,1,410}$	16,920
384	$C_{1,1},\dots,C_{1,1,170}$	…	$C_{12,1},\dots,C_{12,1,170}$	14,040
512	$C_{1,1},\dots,C_{1,930}$	…	$C_{12,1},\dots,C_{12,930}$	11,160
768	$C_{1,1},\dots,C_{1,690}$	…	$C_{12,1},\dots,C_{12,690}$	8,280
1024	$C_{1,1},\dots,C_{1,480}$	…	$C_{12,1},\dots,C_{12,480}$	5,760
1536	$C_{1,1},\dots,C_{1,360}$	…	$C_{12,1},\dots,C_{12,360}$	4,320
2048	$C_{1,1},\dots,C_{1,240}$	…	$C_{12,1},\dots,C_{12,240}$	2,880
4000	$C_{1,1},\dots,C_{1,120}$	…	$C_{12,1},\dots,C_{12,120}$	1,440
			Total Non-zero Features	2,300,040

Nonetheless, feature matrices like Table 3.5, which are filled with zero paddings, will be used to build training and test datasets. In the next subsection, the characteristics of training and test data will be explained. Moreover, parameters of the classification algorithms will be discussed.

3.3.3. Development of Datasets and Classification Models

3.3.3.1. Datasets

To prepare a finalized data with which the algorithms could be trained and tested, the feature matrix of each leak type shown in Table 3.5 was transposed. Therefore, each leak type's transposed feature matrix will have a shape of [360,000 × 26] shown in Figure 3.29. This step was performed for non-leak, orifice, longitudinal, circumferential, and gasket leak types. Then, the following steps were performed to build training and test datasets.

Then, 80% of the data generated from the non-leak feature matrix (288,000 = 360,000 × 0.8) was randomly chosen to build a basis for training data, and the remaining 20% data, i.e., 72,000 data, was selected as a basis for test data. This procedure was performed by applying the "train_test_split" module of the Scikit-learn Application Programming Interface (API) (Pedregosa et al., 2011) on non-leak signals. It is worth noting that all of the signals were from the looped network. To make training data more realistic, some leak data were introduced to the non-leak training data. This idea conforms to the real case studies where due to human errors or undetected small leaks, some leak signals may be incorrectly labeled as non-leak. To this aim, 5% anomalies were added to the non-leak training data, labeled as non-leak, which resulted in (14,400 = 288,000 × 0.05) abnormalizes. Twenty percent anomalies, labeled as leak, of the 288,000 initial training samples, i.e., 56,600 = 288,000 × 0.20 samples, were added to the training data. The 72,000 anomalies, i.e., 57,600 samples labeled as leak and 14,400 labeled as non-leak, were equally selected from the orifice, longitudinal circumferential, and gasket leak data. Then the 72,000 anomalies were added to the 288,000 initial training samples.

$$
\begin{array}{cc}
 & \begin{array}{ccccc} 8\text{ (Hz)} & 10\text{ (Hz)} & \cdots & 2048\text{ (Hz)} & 4000\text{ (Hz)} \end{array} \\
\begin{array}{c} S_1 \\ \\ S_2 \\ \\ \vdots \\ \\ S_{11} \\ \\ S_{12} \end{array} &
\begin{bmatrix}
\begin{matrix} C_{1,1} \\ , \\ \vdots \\ , \\ C_{1,30,000} \end{matrix} & \begin{matrix} C_{1,1} \\ , \\ \vdots \\ , \\ C_{1,30,000} \end{matrix} & \ddots & \begin{matrix} C_{1,1} \\ , \\ \vdots \\ , \\ C_{1,30,000} \end{matrix} & \begin{matrix} C_{1,1} \\ , \\ \vdots \\ , \\ C_{1,30,000} \end{matrix} \\
\begin{matrix} C_{2,1} \\ , \\ \vdots \\ , \\ C_{2,30,000} \end{matrix} & \begin{matrix} C_{2,1} \\ , \\ \vdots \\ , \\ C_{2,30,000} \end{matrix} & \ddots & \begin{matrix} C_{2,1} \\ , \\ \vdots \\ , \\ C_{2,30,000} \end{matrix} & \begin{matrix} C_{2,1} \\ , \\ \vdots \\ , \\ C_{2,30,000} \end{matrix} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\begin{matrix} C_{11,1} \\ , \\ \vdots \\ , \\ C_{11,30,000} \end{matrix} & \begin{matrix} C_{11,1} \\ , \\ \vdots \\ , \\ C_{11,30,000} \end{matrix} & \ddots & \begin{matrix} C_{11,1} \\ , \\ \vdots \\ , \\ C_{11,30,000} \end{matrix} & \begin{matrix} C_{11,1} \\ , \\ \vdots \\ , \\ C_{11,30,000} \end{matrix} \\
\begin{matrix} C_{12,1} \\ , \\ \vdots \\ , \\ C_{12,30,000} \end{matrix} & \begin{matrix} C_{12,1} \\ , \\ \vdots \\ , \\ C_{12,30,000} \end{matrix} & \ddots & \begin{matrix} C_{12,1} \\ , \\ \vdots \\ , \\ C_{12,30,000} \end{matrix} & \begin{matrix} C_{12,1} \\ , \\ \vdots \\ , \\ C_{12,30,000} \end{matrix}
\end{bmatrix}
\end{array}
$$

Figure 3.29. Transposed matrix for a leak type with twelve signals used to build the datasets

Therefore, the final training data had 360,000 = 288,000 (normal) + 57,600 (abnormal) + 14,400 (mislabeled) samples and formed a matrix with the shape of [360,000 × 26] whose leak and non-leak rows were randomly distributed. To build the test data, the 72,000 non-leak data randomly selected for testing was used as a basis test dataset. To test how trained algorithm performs on normal and abnormal unseen data, 20% anomalies, labeled as leak, and 5% mislabeled anomalies, labeled as non-leak, were added to the basis test dataset.

With concatenating the 72,000 normal, 14,400 abnormal data, and 3600 mislabeled anomalies, the test dataset had 90,000 = 72,000 (normal) + 14,400 (abnormal) + 3,600 (mislabeled) samples and formed a matrix with the shape of [90,000 × 26] with a shuffled row arrangement. Since SVM and 1CSVM algorithms assume that their input data are in a standard range, both datasets were 'hard' normalized by mapping each feature's min and max values to 0 and 1. Then, a label column as the 27th column was added to both train and test datasets with labels -1 and 1 corresponding to outlier and inlier labels, respectively.

Figures 3.30 and 3.31 show the parallel plots of the training and test datasets for the looped network, which visualize individual observations at each feature (frequency). In these plots, each data, respectively.

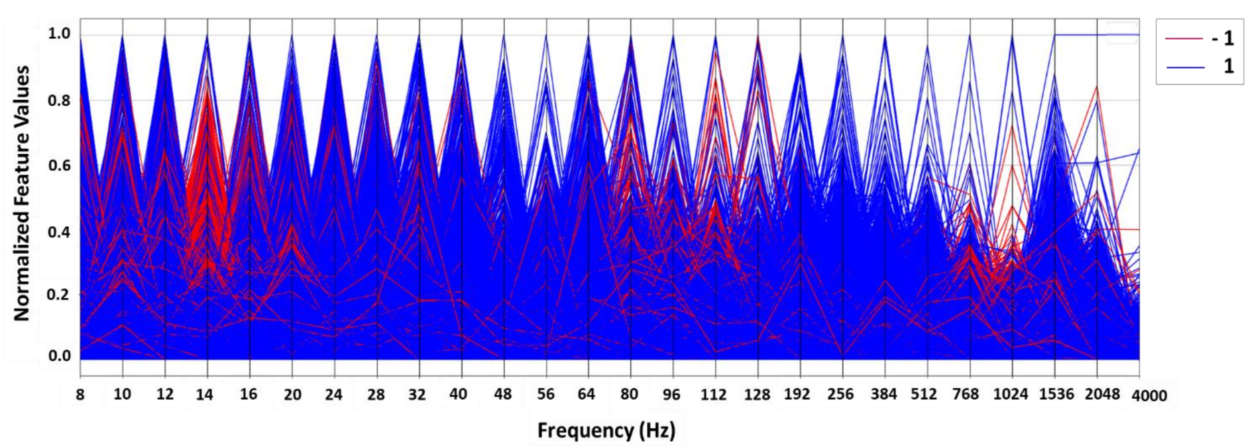

Figure 3.30. A parallel plot of training data for the looped network

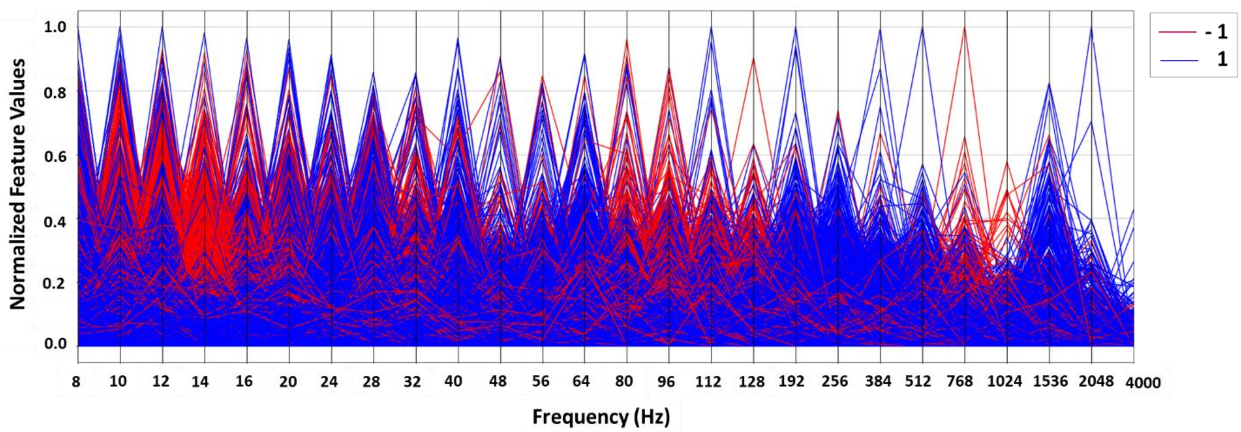

Figure 3.31. A parallel plot of test data for the looped network

Regarding Figure 3.30, the majority of abnormal data, that are leak signals, have frequencies less than 150 Hz and are mainly focused at the frequency of 14 Hz, and some are located between 64 Hz and 128 Hz. Some ensembles of red lines between frequencies 512 Hz and 1536 Hz could be relevant to leak signals' transient events. The wavelet map of Figure 3.21 that includes a transient event of the orifice leak shows the large magnitudes of leak signals with frequencies higher than 512 Hz. Besides, due to the samples' random distribution of training and test datasets, lots of outlier data, red lines, could have been masked by inliers, blue lines, in Figures 3.30 and 3.31. Hence, the observable clusters of red lines show how dominant the outliers are at that those frequencies. Moreover, the locations of the red lines in the parallel plots justify why using a bandpass filter could have removed leak signals and make data biased.

Figure 3.31 shows that the test data has signal contents approximately similar to those of the training data with a prevalence of leak data at frequencies under 150 Hz, ensembles of outliers at low frequencies such as 14 Hz and at the higher frequencies like in the range of 512 Hz and 1536 Hz. However, though the ratios of leak data to non-leak data are the same in both training and test datasets, the test data's parallel plot has some differences compared to its training counterpart. For instance, the test data includes leak data more pronounced in the frequency ranges of (9 Hz, 20 Hz) and (64 Hz, 128 Hz). Additionally, the blue lines, i.e., non-leak data, are less dense in the low frequencies range, under 150 Hz, than the training data's parallel plot. However, these points might be a few differences in the training and test datasets highlighted based on visual evaluations. Having dissimilar test data gives the opportunity of testing the classification algorithms on new data with unseen patterns. In the following, the classification algorithms and their performance on the looped network data will be discussed.

3.3.3.2. Classification Models

The classification task was performed utilizing the Scikit-learn 0.23.2 API, a tool for machine learning analyses in Python (Pedregosa et al., 2011). The five shallow classifiers used in this study, 1CSVM, iForest, LOF, SVM, and XGBoost, include parameters that needed to be adjusted for the best performance. Table 3.7 includes important parameters and the numbers or options used for them. Other input parameters were set according to default values of the Scikit-learn 0.23.2.

Table 3.7. Parameters to be optimized for each classification model

Model	Parameters
1CSVM	*kernel* = {linear, poly, rbf} ν = {10^{-12}, 10^{-6}, 10^{-1}} *d* = {2, 6, 10} γ = {2×10^{-13}, 2, 2×10^{3}}
iForest	*n_estimators* = {2, 4, 6, 8, 10, 15, 20, 50} *contamination* = 0.2
LOF	*n_neighbors* = {5, 10, 15, 20, 50, 100} *contamination* = 0.2
SVM	*kernel* = {linear, poly, rbf} *C* = {10^{-4}, 10^{-2}, 1, 10, 100} *d* = {2, 6, 10}
XGBoost	*n_estimators* = {10, 100, 500, 1000} *max_depth* = {5, 8, 10, 15, 50, 100, 500, 1000}

In Table 3.7, v is an upper bound of the fraction of training errors and a lower bound of the fraction of support vectors, d is the degree of the polynomial kernel function, γ is the coefficient of the RBF kernel function, *n_estimators* in iForest is the number of trees used to isolate anomalies, *contamination* in iForest and LOF is the percentage of errors or abnormalities in a training data, C in SVM is a penalty constant for slack variables, and in XGBoost, *n_estimators* is the number of trees (i.e., weak learners) and *max_depth* is the maximum depth of trees.

The classification models were trained with two types of datasets: (1) original training data with 26 features and (2) training data with reduced dimensions using PCA. Also, 10-fold stratified cross-validation was employed to avoid overfitting and evaluating each model's performance on randomly selected folds of training data.

3.3.3.2.1. Shallow Classifiers with Original Input Data

The following tables include the optimized parameters and evaluation metric results of the best 1CSVM, iForest, LOF, SVM, and XGBoost models tested on the test dataset. All of the models were trained and tested with data having 26 features of the looped network. It is worth noting that the classes '-1' and '1' refer to leak and non-leak data, respectively. Regarding the leak class's importance in leak detection, the best model was chosen based on the highest F_1-score of the leak class and the Weighted Average, even though the overall F_1-score of the selected model is not the largest. Moreover, the 'Support' refers to the number of samples employed in each model's optimization algorithm for each class or for all data. For example, the 'Support =14,400' in Table 3.8 means all of the 14,400 leak data, out of 90,000 samples in the test data, have been used to test the 1CSVM model trained with the parameters in Table 3.8. Also, the 'Weighted Average' indicates that an evaluation metric has been calculated considering each label's proportion in the dataset.

Table 3.8. The best 1CSVM with a linear kernel, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **1CSVM**					
Kernel	v				
linear	10^{-12}				
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.16	0.88	**0.27**	14400
	1	0.78	0.08	0.15	75600
Weighted Average		0.68	0.22	**0.17**	90000
Accuracy				0.22	
F_1-score				0.15	

Table 3.9. The best 1CSVM with a polynomial kernel, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **1CSVM**					
Kernel	v		*d*		
poly	10^{-12}		10		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.12	0.43	**0.19**	14400
	1	0.78	0.40	0.53	75600
Weighted Average		0.67	0.41	**0.47**	90000
Accuracy				0.41	
F_1-score				0.53	

Table 3.10. The best 1CSVM with an RBF kernel, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **1CSVM**					
Kernel		ν		γ	
rbf		10^{-1}		2	
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.30	0.20	**0.24**	14400
	1	0.85	0.91	0.88	75600
Weighted Average		0.76	0.79	**0.77**	90000
Accuracy				0.79	
F_1-score				0.88	

Table 3.11. The best iForest, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **iForest**					
n_estimators			*contamination*		
20			0.2		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.24	0.31	**0.27**	14400
	1	0.85	0.80	0.83	75600
Weighted Average		0.75	0.72	**0.73**	90000
Accuracy				0.72	
F_1-score				0.83	

Table 3.12. The best LOF, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **LOF**					
n_neighbors			*contamination*		
50			0.2		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.25	0.34	**0.29**	14400
	1	0.86	0.80	0.83	75600
Weighted Average		0.76	0.73	**0.74**	90000
Accuracy				0.73	
F_1-score				0.83	

Table 3.13. The best SVM with a linear kernel, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **SVM**					
Kernel	*C*				
linear	10^{-2}				
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.41	0.32	**0.36**	14400
	1	0.87	0.91	0.89	75600
Weighted Average		0.79	0.81	**0.80**	90000
Accuracy				0.81	
F_1-score				0.89	

Table 3.14. The best SVM with a polynomial kernel, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **SVM**					
Kernel	*C*		*d*		
poly	1		2		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.40	0.44	**0.42**	14400
	1	0.89	0.78	0.88	75600
Weighted Average		0.81	0.80	**0.80**	90000
Accuracy				0.80	
F_1-score				0.88	

Table 3.15. The best SVM with an RBF kernel, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **SVM**					
Kernel	*C*				
rbf	1				
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.34	0.65	**0.44**	14400
	1	0.91	0.74	0.82	75600
Weighted Average		0.82	0.73	**0.76**	90000
Accuracy				0.73	
F_1-score				0.82	

Table 3.16. The best XGBoost, its optimized parameters, and evaluation metrics for the looped network

<table>
<tr><td colspan="6">Optimized Parameters for XGBoost</td></tr>
<tr><td colspan="3">n_estimators</td><td colspan="3">max_depth</td></tr>
<tr><td colspan="3">500</td><td colspan="3">8</td></tr>
<tr><td colspan="6">Evaluation Metrics</td></tr>
<tr><td colspan="2"></td><td>Precision</td><td>Recall</td><td>F₁-score</td><td>Support</td></tr>
<tr><td rowspan="2">Class</td><td>-1</td><td>0.89</td><td>0.13</td><td>0.23</td><td>14400</td></tr>
<tr><td>1</td><td>0.85</td><td>1.00</td><td>0.92</td><td>75600</td></tr>
<tr><td colspan="2">Weighted Average</td><td>0.86</td><td>0.85</td><td>0.80</td><td rowspan="3">90000</td></tr>
<tr><td colspan="4">Accuracy</td><td>0.85</td></tr>
<tr><td colspan="4">F₁-score</td><td>0.92</td></tr>
</table>

3.3.3.2.2. Shallow Classifiers with Reduced-dimension Input Data

Figure 3.32 shows the distribution of the training data at each feature. Regarding Figure 3.32, the values of most of the features, approximately at frequencies above 40 (Hz), are zero or very small. This can indicate that the roles of features with higher frequencies in training the classifiers are negligible. Moreover, having features with a majority of zero values might refer to a sparsity of the training data that could affect the classification models' optimization procedures. Figure 3.33 also depicts the distribution of the test data at each feature. Like the training data, most non-zero values of the test data are under the 40 Hz feature. This resemblance indicates that a classifier would be tested on new data that has approximately similar feature patterns.

A heatmap of the training data, Figure 3.34, has been plotted to evaluate how dependent the features are. Based on the training data's heatmap, features are not strongly correlated, and some extents of dependencies exist among frequencies with close values. For instance, the largest correlation coefficient is 0.65, which corresponds to the dependency of the frequencies 80 Hz and 96 Hz, while distant features, like 8 Hz and 4000 Hz, are completely uncorrelated. The lack of correlations can indicate that the features are not redundant and removing some features might adversely affect the classification models' performance.

The reason for the prevalence of zero values at features larger than 150 Hz is the absence of acoustic signals at higher frequencies, which leads to zero values of wavelet coefficients.

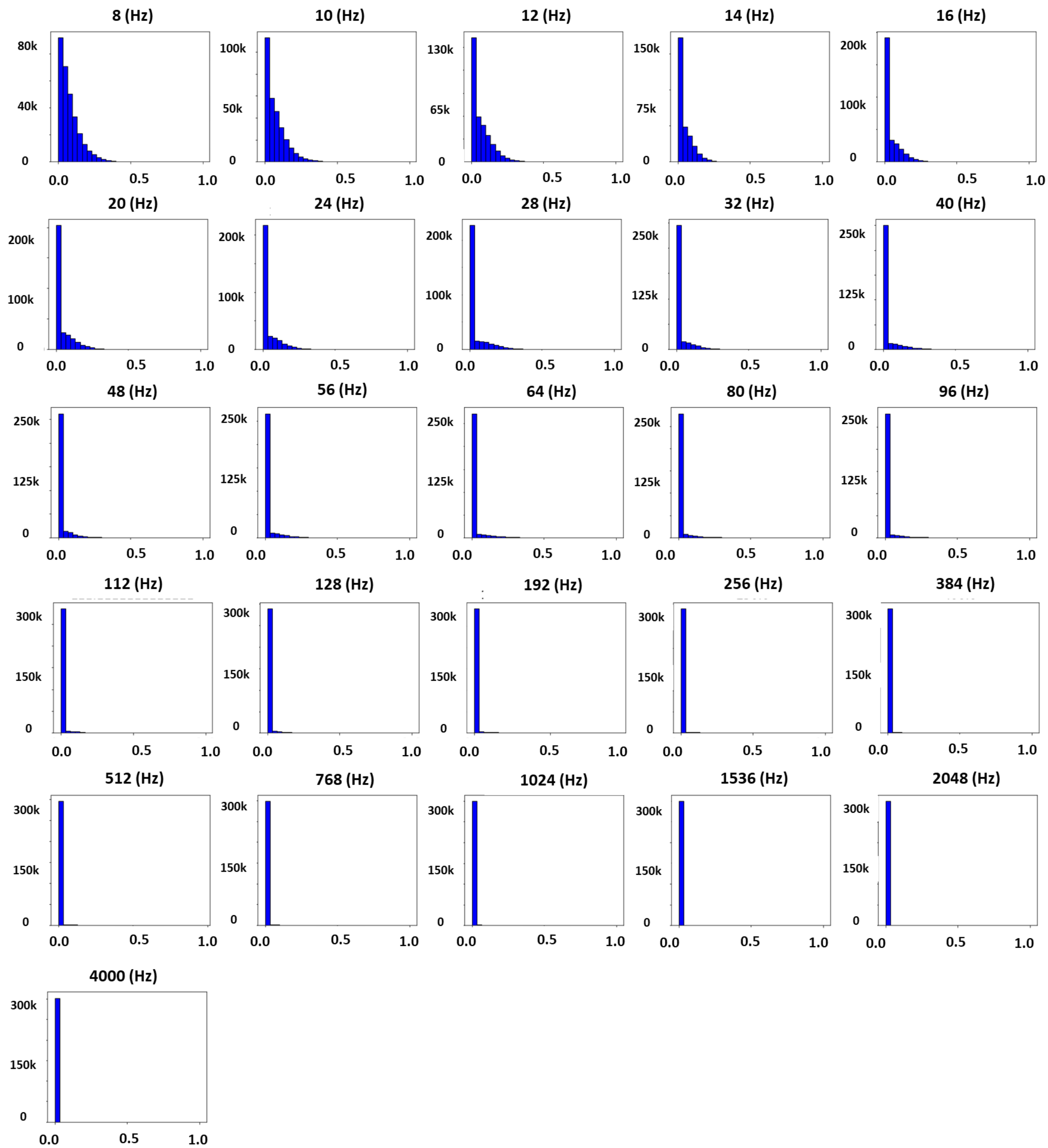

Figure 3.32. Histogram of the training data at each feature for the looped network

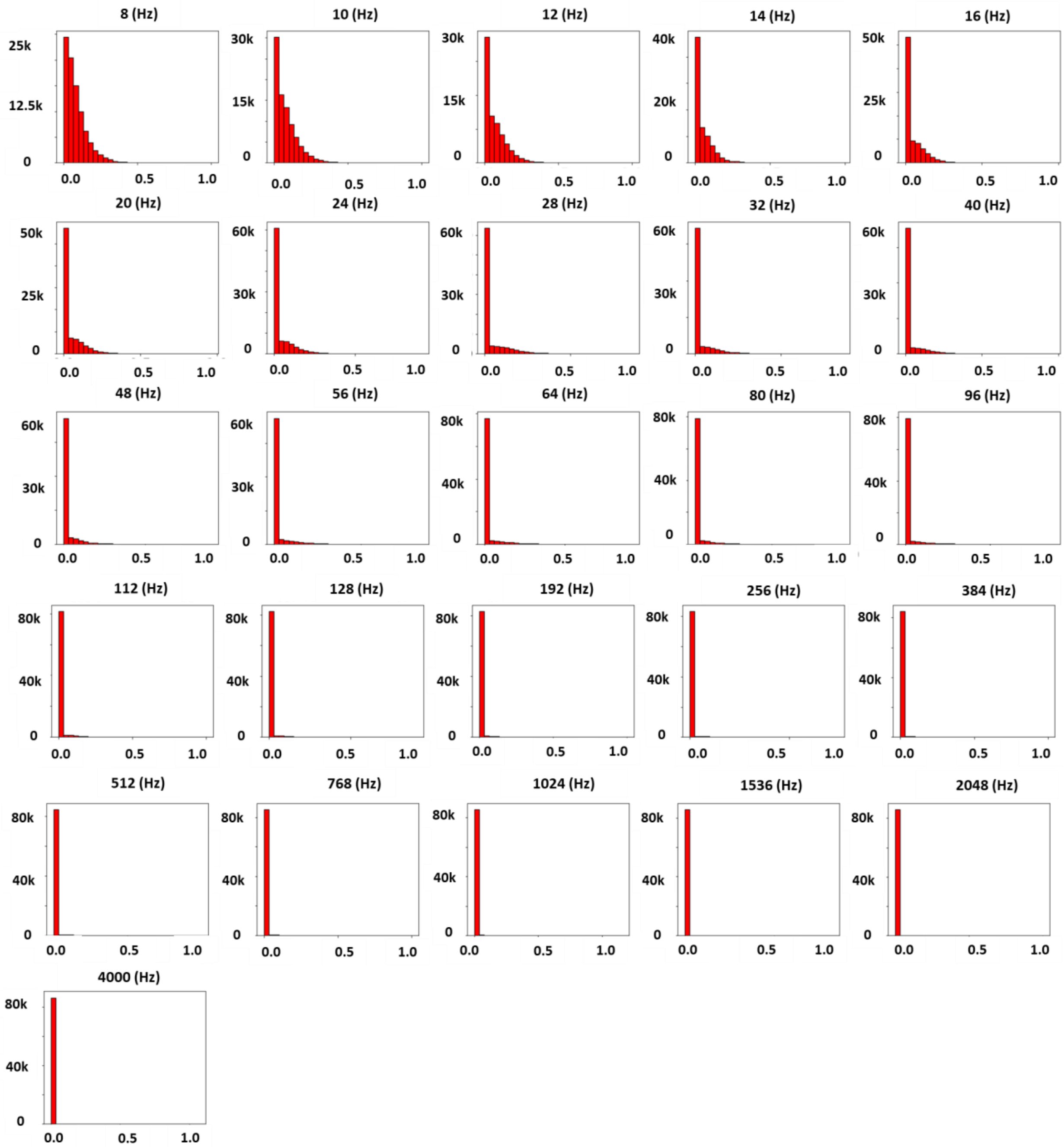

Figure 3.33. Histogram of the test data at each feature for the looped network

One of the challenges in the leak detection procedure was the long time required to train the shallow classifiers, though a computational resource described in Table 3.17 was utilized. For instance, training the 1CSVM model with an RBF kernel, γ = 2×10^3 and ν = 10^{-1}on data with all

26 features took about 05:11:02 hours. However, training the same model with the same machine but on two-dimensional data required 00:55:23 hours. Therefore, models were trained and tested on data with lower dimensions, and their performances were compared.

Table 3.17. Hardware overview of the computing cluster used for the machine learning analyses

System Name	Terra (an Intel x86-64 Linux cluster)
Processor Type	Intel Xeon E5-2680 v4 2.40GHz 14-core
Number of Used Nodes	1
Number of Cores per Used Node	28
Memory per Used Node	128 GB DDR4, 2400 MHz
Accelerator(s)	1 NVIDIA K80 Accelerator

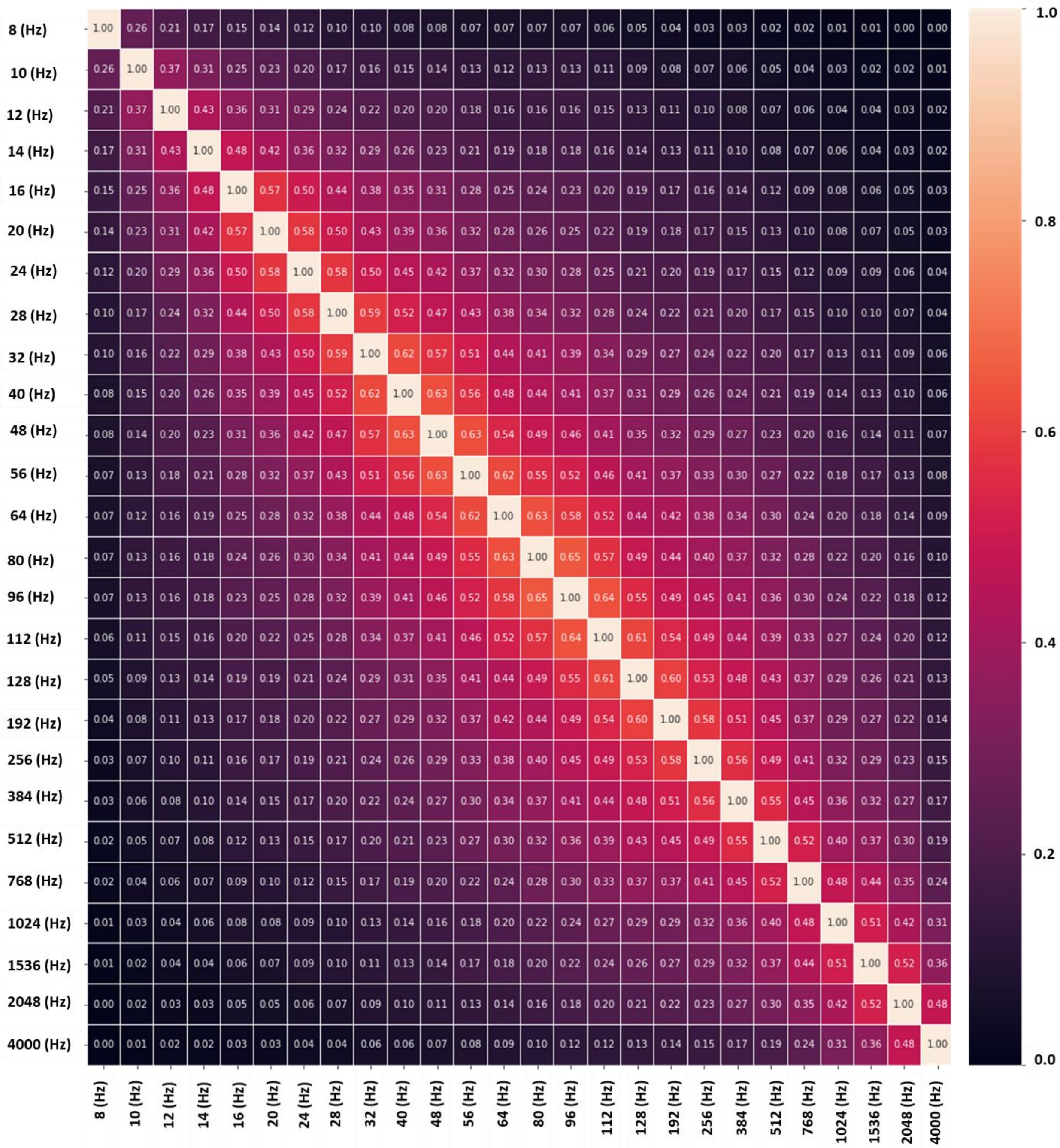

Figure 3.34. Heatmap of training data for the looped network

The following tables include the best results of the five models on the test data with their optimized parameters and evaluation metric results, where the algorithms were fed with reduced-dimension data using PCA (0.95). PCA (0.95) indicates that only those principal components were employed which capture 95% of training data's variance. Each table is titled with 'model name-PCA' that indicates the 'model name' classifier is trained with low dimension data utilizing PCA(0.95).

Table 3.18. The best 1CSVM-PCA with a linear kernel, its optimized parameters, and evaluation metrics using for the looped network

<table>
<tr><td colspan="6">Optimized Parameters for 1CSVM-PCA</td></tr>
<tr><td>Kernel</td><td colspan="5">N</td></tr>
<tr><td>linear</td><td colspan="5">10^{-1}</td></tr>
<tr><td colspan="6">Evaluation Metrics</td></tr>
<tr><td colspan="2"></td><td>Precision</td><td>Recall</td><td>F_1-score</td><td>Support</td></tr>
<tr><td rowspan="2">Class</td><td>-1</td><td>0.18</td><td>0.93</td><td>0.30</td><td>14400</td></tr>
<tr><td>1</td><td>0.91</td><td>0.13</td><td>0.23</td><td>75600</td></tr>
<tr><td colspan="2">Weighted Average</td><td>0.79</td><td>0.27</td><td>0.24</td><td rowspan="3">90000</td></tr>
<tr><td colspan="4">Accuracy</td><td>0.27</td></tr>
<tr><td colspan="4">F_1-score</td><td>0.23</td></tr>
</table>

Table 3.19. The best 1CSVM-PCA with a polynomial kernel, its optimized parameters, and evaluation metrics for the looped network

<table>
<tr><td colspan="6">Optimized Parameters for 1CSVM-PCA</td></tr>
<tr><td>Kernel</td><td colspan="3">v</td><td colspan="2">d</td></tr>
<tr><td>poly</td><td colspan="3">10^{-12}</td><td colspan="2">2</td></tr>
<tr><td colspan="6">Evaluation Metrics</td></tr>
<tr><td colspan="2"></td><td>Precision</td><td>Recall</td><td>F_1-score</td><td>Support</td></tr>
<tr><td rowspan="2">Class</td><td>-1</td><td>0.16</td><td>0.83</td><td>0.27</td><td>14400</td></tr>
<tr><td>1</td><td>0.78</td><td>0.13</td><td>0.22</td><td>75600</td></tr>
<tr><td colspan="2">Weighted Average</td><td>0.68</td><td>0.24</td><td>0.22</td><td rowspan="3">90000</td></tr>
<tr><td colspan="4">Accuracy</td><td>0.24</td></tr>
<tr><td colspan="4">F_1-score</td><td>0.22</td></tr>
</table>

Table 3.20. The best 1CSVM-PCA with an RBF kernel, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **1CSVM-PCA**					
Kernel		ν		γ	
rbf		10^{-1}		2×10^{3}	
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.25	0.24	**0.25**	14400
	1	0.85	0.85	0.85	75600
Weighted Average		0.75	0.75	**0.75**	90000
Accuracy				0.75	
F_1-score				0.85	

Table 3.21. The best iForest-PCA, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **iForest-PCA**					
n_estimators			*contamination*		
8			0.2		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.28	0.37	**0.31**	14400
	1	0.86	0.81	0.83	75600
Weighted Average		0.77	0.73	**0.75**	90000
Accuracy				0.73	
F_1-score				0.83	

Table 3.22. The best LOF-PCA, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **LOF-PCA**					
n_neighbors			*contamination*		
20			0.2		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.17	0.29	**0.22**	14400
	1	0.84	0.72	0.77	75600
Weighted Average		0.73	0.65	**0.68**	90000
Accuracy				0.65	
F_1-score				0.77	

Table 3.23. The best SVM-PCA with a linear kernel, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **SVM-PCA**					
Kernel	*C*				
linear	1				
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.26	0.53	**0.35**	14400
	1	0.88	0.69	0.78	75600
Weighted Average		0.78	0.67	**0.70**	90000
Accuracy				0.67	
F_1-score				0.78	

Table 3.24. The best SVM-PCA with a polynomial kernel, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **SVM-PCA**					
Kernel	*C*		*d*		
poly	10		2		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.26	0.53	**0.35**	14400
	1	0.88	0.69	0.78	75600
Weighted Average		0.78	0.67	**0.70**	90000
Accuracy				0.67	
F_1-score				0.78	

Table 3.25. The best SVM-PCA with an RBF kernel, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **SVM-PCA**					
Kernel	*C*				
rbf	10				
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.25	0.65	**0.36**	14400
	1	0.90	0.62	0.73	75600
Weighted Average		0.79	0.62	**0.67**	90000
Accuracy				0.62	
F_1-score				0.73	

Table 3.26. The best XGBoost-PCA, its optimized parameters, and evaluation metrics for the looped network

Optimized Parameters for **XGBoost-PCA**					
n_estimators			*max_depth*		
1000			100		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.26	0.08	**0.12**	14400
	1	0.84	0.96	0.89	75600
Weighted Average		0.74	0.81	**0.76**	90000
Accuracy				0.81	
F_1-score				0.89	

3.4. Results and Discussion

3.4.1. Performance Evaluation of Classifiers

Tables 3.27 and 3.28 rank the best classifiers trained and tested with data having all and reduced dimensions, respectively. Since the Weighted Average F_1-score considers F_1-score of both classes based on their proportional data, this metric was selected to rank the algorithms. The Class '-1' F_1-score was regarded as the second criterion for the ranking should the Weighted Average F_1-score was the same for different models. It is worth noting that only the kernels that resulted in the best 1CSVM and SVM models were selected in the ranking tables.

Table 3.27. Ranked performance of classification models on test data with all features for the looped network

Model	Weighted Average F_1-score	Class '-1' F_1-score	Class '1' F_1-score	Accuracy
SVM (poly)	0.80	0.42	0.88	0.80
XGBoost	0.80	0.23	0.92	0.85
1CSVM (rbf)	0.77	0.24	0.88	0.79
LOF	0.74	0.29	0.83	0.73
iForest	0.73	0.27	0.83	0.72

Table 3.28. Ranked performance of classification models on test data with reduced dimensions for the looped network

Model	Weighted Average F_1-score	Class '-1' F_1-score	Class '1' F_1-score	Accuracy
XGBoost-PCA	0.76	0.12	0.89	0.81
iForest-PCA	0.75	0.31	0.83	0.73
1CSVM-PCA (rbf)	0.75	0.25	0.85	0.75
SVM-PCA (linear or poly)	0.70	0.35	0.78	0.67
LOF-PCA	0.68	0.22	0.77	0.65

Regarding Tables 3.23 and 3.24, the model SVM-PCA had the same performance with the linear and poly kernels. Therefore, the SVM-PCA with both of these kernels were considered in the ranking of Table 3.28. In the following, the models listed in Tables 3.27 and 3.28 are compared, and some points are highlighted for leak detection in the looped network case study.

Figures 3.35 and 3.36 show the distributions of the training and test data for the looped network, respectively, for the first three features, i.e., 8 Hz, 10 Hz, and 12 Hz.

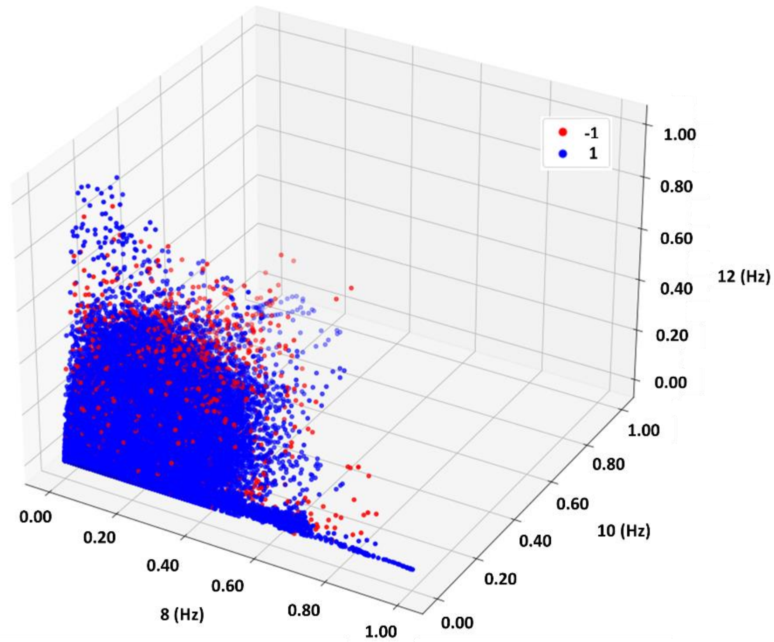

Figure 3.35. Distribution of the first three features of training dataset for the looped network

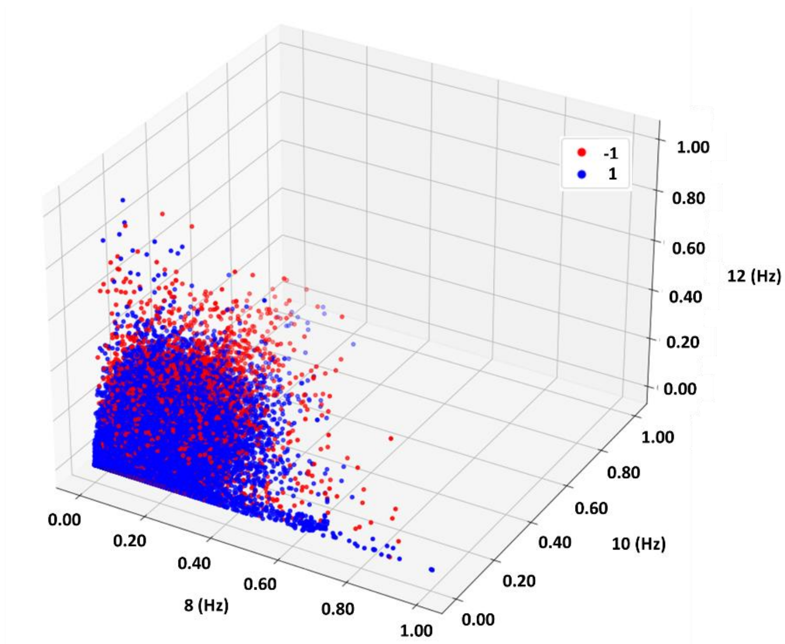

Figure 3.36. Distribution of the first three features of test dataset for the looped network

Though both figures show the data distributions at only three features, it can be seen that leak and non-leak samples are very mixed and are not easily separable. This inseparability could cause the SVM with a 2nd-degree polynomial kernel to be the best classifier for the test data with all features.

The low rank of LOF in Table 3.27 can be due to the following reason. LOF compares the local density of a point to the average local density of that point's *K* neighbors, see Eq. (3.9). Since leak and non-leak data are mixed and close, the LOF of a leak example resembles non-leak data. In other words, the close distances of the leak and non-leak examples would locate leak data in the neighborhood of non-leak data.

The iForest ranked last among the classifiers on data with full features. It could be due to the following reasons.

i. The mechanism of iForest is isolating abnormal data based on feature values. Since leak and non-leak data have many similar values, especially at low frequencies, iForest could not find features whose values are much different for leak and non-leak data.

ii. Another reason for this algorithm's performance could be the low value set for the Contamination hyperparameter. Results showed that increasing this value from 0.2 to 0.7 could slightly improve the algorithm's performance in predicting leak data, though it decreased the F_1-score for non-leak data prediction.

Nonetheless, regarding the "Class -1 F_1-score", iForest ranked 3rd and performed better than 1CSVM and XGBoost.

Though XGBoost uses regularization to prevent overfitting, since classes are mixed, see Figure 3.35, weak leaner trees train for a long time to reach an appropriate termination pint.

Therefore, the boosting algorithm would fit the wrong class, which causes a misclassification. On the other hand, when classes are well separated since growing trees' termination points would happen earlier, a boosting algorithm trains enough and avoids overfitting. This could explain why XGBoost has the best performance on reduced-dimensional data where the two classes are not mixed.

Although XGBoost ranked 1st and 2nd on data with reduced dimensions and full dimensions, respectively, it has the worst performance in predicting the leak data and vice versa. This difference might indicate how XGBoost is sensitive to class imbalances, and it gives a better result for the majority class.

A reason for the better performance of XGBoost on the majority class might be the more samples of non-leak training data available for the algorithm to build more numbers of weak learners and decrease training errors using more gradients.

Contrary to other studies where a dimension reduction with PCA improved their data analysis results (Erfani et al., 2016; Abokifa et al., 2019; Jollife and Cadima, 2016), PCA(0.95) did not increase the efficiency of classifiers in this study. For example, PCA caused the Weighted Average F_1-score to decrease from 0.80 to 0.76 when the best models' performances are considered with using and without using dimension reduction. One reason could be the features' uncorrelated nature, while PCA might have removed useful information by reducing the number of attributes.

PCA will be beneficial if features are highly correlated.

Though this study and that of Cody et al. (2018) have many common characteristics, There is a difference between the best classifiers. For example, in the paper of Cody et al. (2018), 1CSVM with RBF kernel resulted in a high-detection accuracy, while it is not the case in this study where 1CSVM ranked the 3rd among the five classifiers. This difference can indicate that the testbed's acoustic data has more complexities and leak data are not easily detectable even by means of using a semi-supervised classifier that takes advantage of a nonlinear kernel. Moreover, this difference can imply that the testbed's main design objective, i.e., generating data simulating more realistic conditions compared to its counterparts, has been achieved.

XGBoost performed as the best and the 2nd best model for the reduced-dimension and all-dimension test data. This conformed to the papers of Gou et al. (2021) and Taormina and Galelli (2018) when they used shallow classifiers. There could be two reasons for this model's superiority: (1) XGBoost is an ensemble of weak learners that combines multiple base learners' prediction and generates one overall prediction for each input. This makes the model capable of learning more complex relationships among the features and classes in the training data; (2) The model implements binary decision trees as base learners. They are highly efficient at learning nonlinear relations between features and targets.

Neither of the algorithms had a Class '-1' F_1-score greater than 0.5, which means they are less successful at detecting leaks than their performance in predicting non-leak data. This poor result could stem from three reasons: (1) the relation among features and classes are that complex that even ensemble learners or classifiers with nonlinear kernels, such as XGBoost or SVM with a polynomial kernel, are not able to learn those patterns; (2) leak and non-leak data generated by the testbed are very similar such that they are not easily discernable with shallow classifiers. This point acknowledges that the generated leak data is in the range of background leaks, not bursts, and is well-aligned with the ideas of using the data for other leak detection and localization studies.

For both test datasets with full and reduced-dimension features, 1CSVM with an RBF kernel ranked 3rd among the five classifiers. This is in contrast to the study of Cody et al. (2018). Two reasons might cause the performance of the 1CSVM in this study:

i. Considering Figures 3.35 and 3.36, many leak data are adjacent to the non-leak data. Hence, the algorithm learns many leak samples as normal ones, which creates learning errors.

ii. The wavelet transform cannot extract features well, so leak and non-leak features are not separable enough.

iii. The introduction of five percent leak data into the training data could have misled the 1CSVM algorithm, which expects completely normal training data.

Although deep learning algorithms could be time and computationally expensive and depend on many hyperparameters, they seem to be promising solutions to detect leaks with more accurate results, especially in predicting leak samples. Examples of such successful studies are those conducted by Chandy et al. (2019) and Taormina et al. (2018).

3.4.2. Influence of Network Architecture

Acoustic datasets were acquired from two hydrophones in the branched network, shown in Figures 2.5 and 2.6. Then features extraction, training, and test steps were performed for the five classifiers 1CSVM, iForest, LOF, SVM, and XGBoost, following the methodology in Figure 3.2, while the network architecture was changed to the branched.

Figures 3.37 shows the branched network's training data's parallel plot.

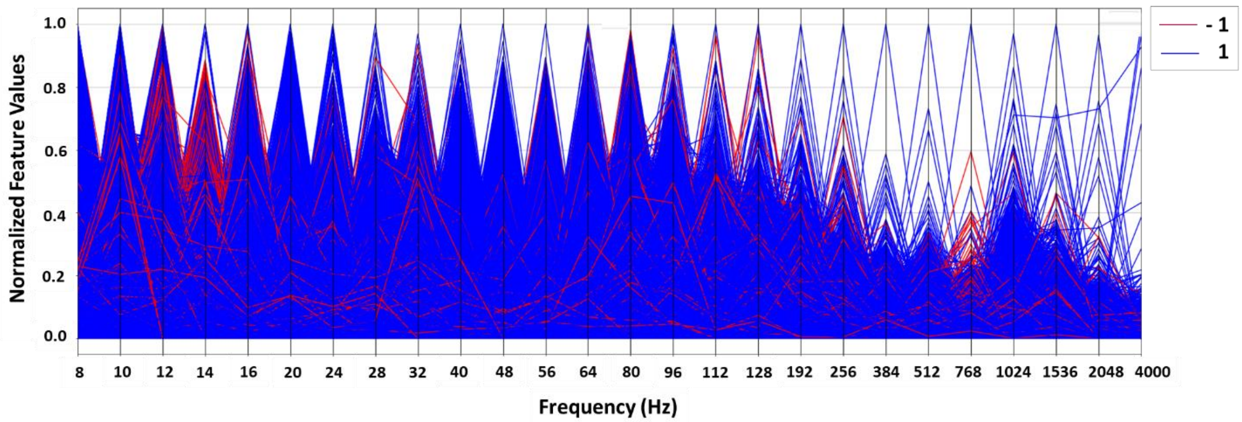

Figure 3.37. A parallel plot of training data for the branched network

Though the non-leak data might mask them, ensembles of leak data, red lines, are mainly located at frequencies 14 Hz and 768 Hz. But the majority of the outliers span between 8 Hz and 16 Hz, 96 Hz and 384 Hz, and 512 Hz and 1024 Hz. Compared to the same plot of the looped network, Figure 3.31, leak data are less pronounced in the branched network, while in the looped network's training data parallel plot, leak data are very dominant at frequencies such as 14 Hz and the higher frequencies like in the ranges of 80 Hz and 128 Hz, and between 512 Hz and 1536 Hz. Also, non-leak samples of the branched network are denser at lower frequencies such as 56 Hz and 96 Hz and less dense at frequencies higher than 256 Hz. These differences can stem from at least two reasons: (1) fewer pipes in the branched architecture prevent leak data, with lower frequencies, from reaching the hydrophones; (2) less connectivity and existing more blind flanges in the branched network cause more signal attenuations, specifically at higher frequencies that are prone to be dampened faster.

Figures 3.38 shows the branched network's test data's parallel plot.

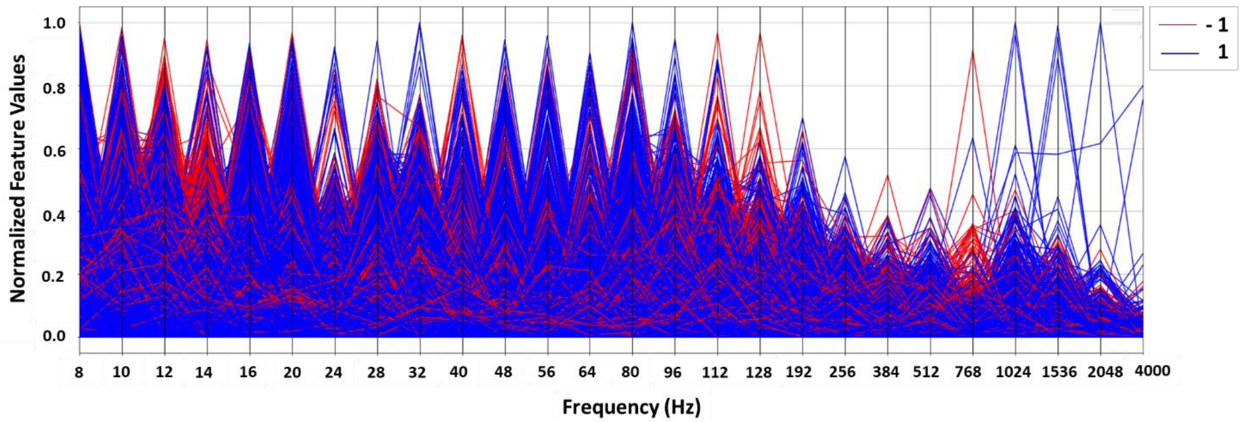

Figure 3.38. A parallel plot of test data for the branched network

Comparing the test data parallel plots of the branched network and the looped network, one could see the differences mentioned above, which are less pronounced leak data at the branched network due to few pipes and less network connectivity, and smaller feature values at higher frequencies due to a more severe attenuation caused by more dead-end pipes.

The classifiers in Table 3.7 were employed to compare the effects of the network change on classification results.

Tables 3.29 to 3.37 show the best results of the classifiers trained and tested on data with all 26 features. The hyperparameters in Table 3.7 have been used for tuning the algorithms' parameters.

Table 3.29. The best 1CSVM with a linear kernel, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **1CSVM**					
Kernel	N				
linear	10^{-1}				
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.02	0.01	**0.01**	14400
	1	0.82	0.90	0.86	75600
Weighted Average				**0.72**	90000
Accuracy				0.75	
F_1-score				0.85	

Table 3.30. The best 1CSVM with a polynomial kernel, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **1CSVM**					
Kernel	N		d		
poly	10^{-6}		6		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.08	0.15	**0.11**	14400
	1	0.80	0.67	0.72	75600
Weighted Average		0.68	0.58	**0.62**	90000
Accuracy				0.58	
F_1-score				0.72	

Table 3.31. The best 1CSVM with an RBF kernel, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **1CSVM**					
Kernel	ν			γ	
rbf	10^{-1}			2×10^{3}	
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.23	0.74	**0.35**	14400
	1	0.90	0.50	0.64	75600
Weighted Average		0.79	0.54	**0.59**	90000
Accuracy				0.54	
F_1-score				0.64	

Table 3.32. The best iForest, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **iForest**					
n_estimators			*contamination*		
20			0.2		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.21	0.27	**0.24**	14400
	1	0.85	0.80	0.82	75600
Weighted Average		0.74	0.71	**0.73**	90000
Accuracy				0.71	
F_1-score				0.82	

Table 3.33. The best LOF, its optimized parameters, and evaluation metrics for the branched network

<table>
<tr><td colspan="6">Optimized Parameters for LOF</td></tr>
<tr><td colspan="3">n_neighbors</td><td colspan="3">contamination</td></tr>
<tr><td colspan="3">5</td><td colspan="3">0.2</td></tr>
<tr><td colspan="6">Evaluation Metrics</td></tr>
<tr><td colspan="2"></td><td>Precision</td><td>Recall</td><td>F_1-score</td><td>Support</td></tr>
<tr><td rowspan="2">Class</td><td>-1</td><td>0.34</td><td>0.58</td><td>0.43</td><td>14400</td></tr>
<tr><td>1</td><td>0.90</td><td>0.77</td><td>0.83</td><td>75600</td></tr>
<tr><td colspan="2">Weighted Average</td><td>0.81</td><td>0.74</td><td>0.76</td><td rowspan="3">90000</td></tr>
<tr><td colspan="4">Accuracy</td><td>0.74</td></tr>
<tr><td colspan="4">F_1-score</td><td>0.83</td></tr>
</table>

Table 3.34. The best SVM with a linear kernel, its optimized parameters, and evaluation metrics for the branched network

<table>
<tr><td colspan="6">Optimized Parameters for SVM</td></tr>
<tr><td>Kernel</td><td colspan="5">C</td></tr>
<tr><td>linear</td><td colspan="5">10^{-2}, 1, 10, 100</td></tr>
<tr><td colspan="6">Evaluation Metrics</td></tr>
<tr><td colspan="2"></td><td>Precision</td><td>Recall</td><td>F_1-score</td><td>Support</td></tr>
<tr><td rowspan="2">Class</td><td>-1</td><td>0.24</td><td>0.42</td><td>0.31</td><td>14400</td></tr>
<tr><td>1</td><td>0.86</td><td>0.73</td><td>0.79</td><td>75600</td></tr>
<tr><td colspan="2">Weighted Average</td><td></td><td></td><td>0.71</td><td rowspan="3">90000</td></tr>
<tr><td colspan="4">Accuracy</td><td>0.68</td></tr>
<tr><td colspan="4">F_1-score</td><td>0.79</td></tr>
</table>

Table 3.35. The best SVM with a polynomial kernel, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **SVM**					
Kernel	*C*		*d*		
poly	1, 100		2		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.26	0.25	**0.26**	14400
	1	0.85	0.86	0.86	75600
Weighted Average				**0.76**	90000
Accuracy				0.76	
F_1-score				0.76	

Table 3.36. The best SVM with an RBF kernel, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **SVM**					
Kernel	*C*				
rbf	100				
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.32	0.65	**0.43**	14400
	1	0.91	0.73	0.81	75600
Weighted Average		0.81	0.72	**0.75**	90000
Accuracy				0.72	
F_1-score				0.81	

Table 3.37. The best XGBoost, its optimized parameters, and evaluation metrics for the branched network

<table>
<tr><td colspan="6">Optimized Parameters for XGBoost</td></tr>
<tr><td colspan="3">n_estimators</td><td colspan="3">max_depth</td></tr>
<tr><td colspan="3">1000</td><td colspan="3">500</td></tr>
<tr><td colspan="6">Evaluation Metrics</td></tr>
<tr><td colspan="2"></td><td>Precision</td><td>Recall</td><td>F_1-score</td><td>Support</td></tr>
<tr><td rowspan="2">Class</td><td>-1</td><td>0.53</td><td>0.16</td><td>0.25</td><td>14400</td></tr>
<tr><td>1</td><td>0.85</td><td>0.97</td><td>0.91</td><td>75600</td></tr>
<tr><td colspan="2">Weighted Average</td><td>0.80</td><td>0.84</td><td>0.80</td><td rowspan="3">90000</td></tr>
<tr><td colspan="4">Accuracy</td><td>0.84</td></tr>
<tr><td colspan="4">F_1-score</td><td>0.91</td></tr>
</table>

Figure 3.39 shows the distribution of the branched network’s training data at each feature. At frequencies above 16 Hz, zero values become very dominant. Comparing the looped and branched networks' training data histograms implies that the branched network’s training data includes more zero values even at low frequencies below 40 Hz. This could be due to weaker signals captured by the hydrophones of the branched network. Figure 3.40 also depicts the distribution of the test data at each feature for the branched network. Compared to the looped network's test data histograms, the branched network's test data has fewer non-zero values, particularly at low frequencies where leak data are expected to exist. This point also acknowledges that the branched network's low-frequency features embed fewer leak data than those of the looped network.

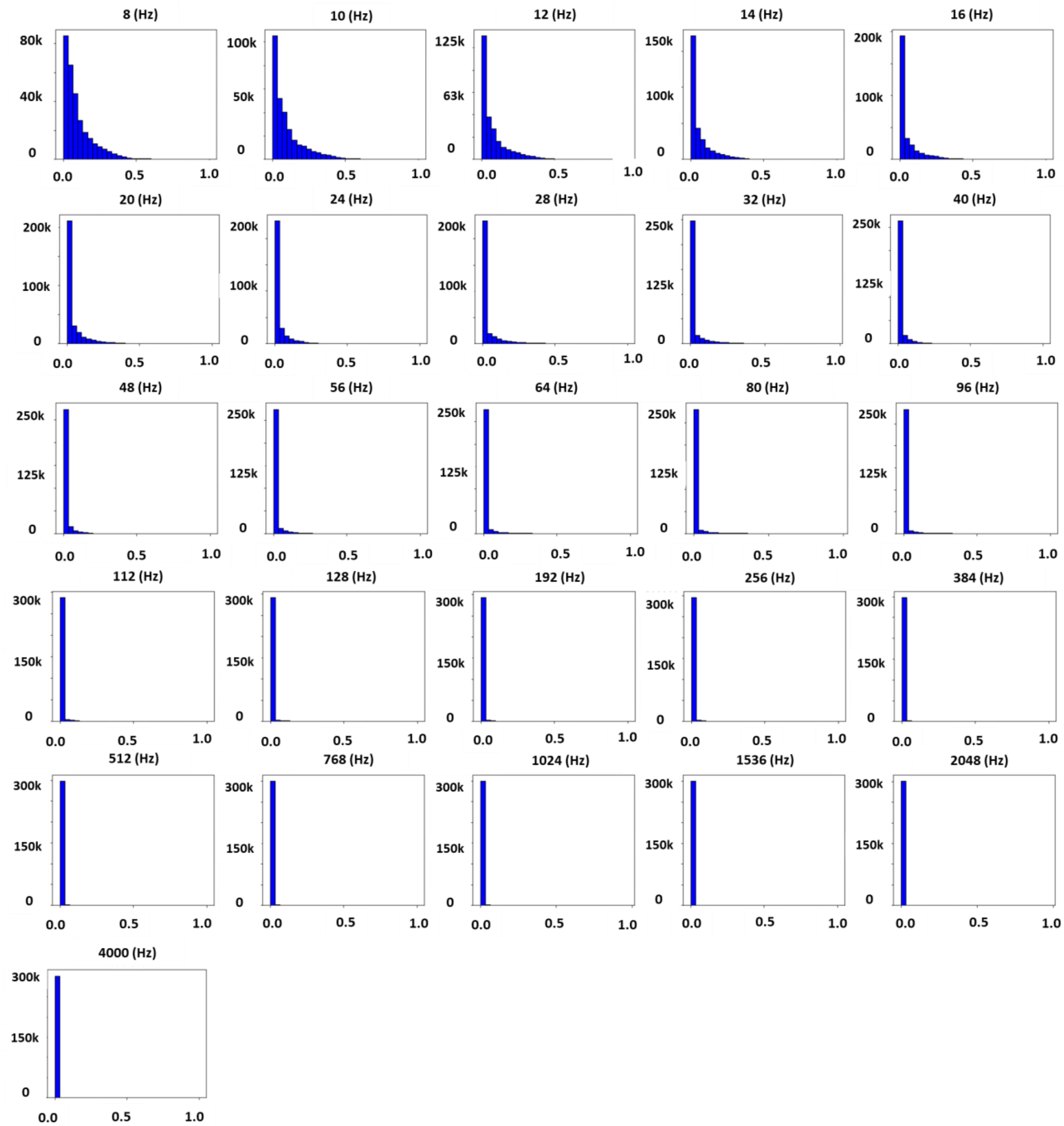

Figure 3.39. Histogram of the training data at each feature for the branched network

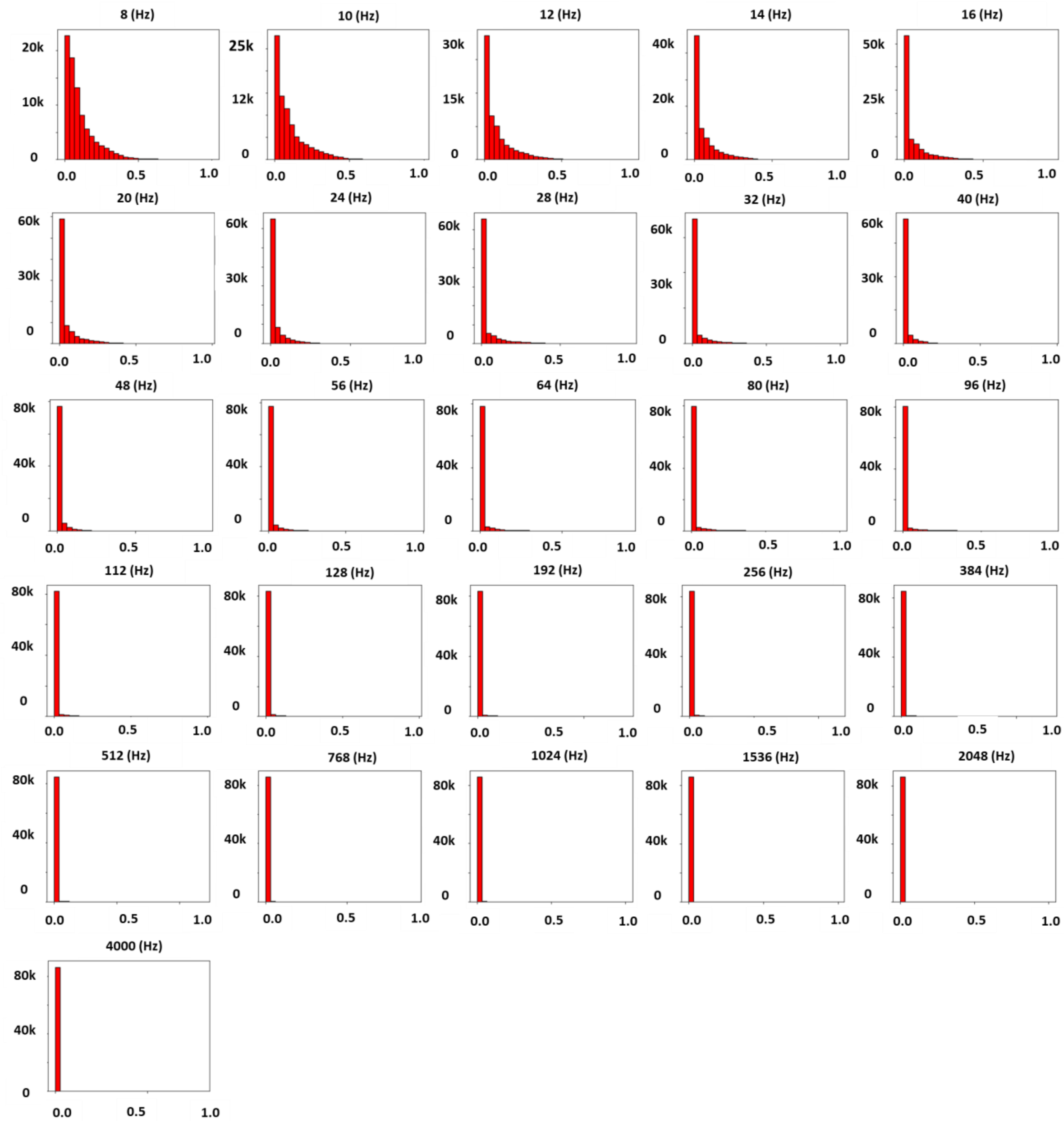

Figure 3.40. Histogram of the test data at each feature for the branched network

To figure out how the branched network's features are correlated and if a dimension reduction helps with classification, a heatmap of training data for the branched network is plotted in Figure 3.41.

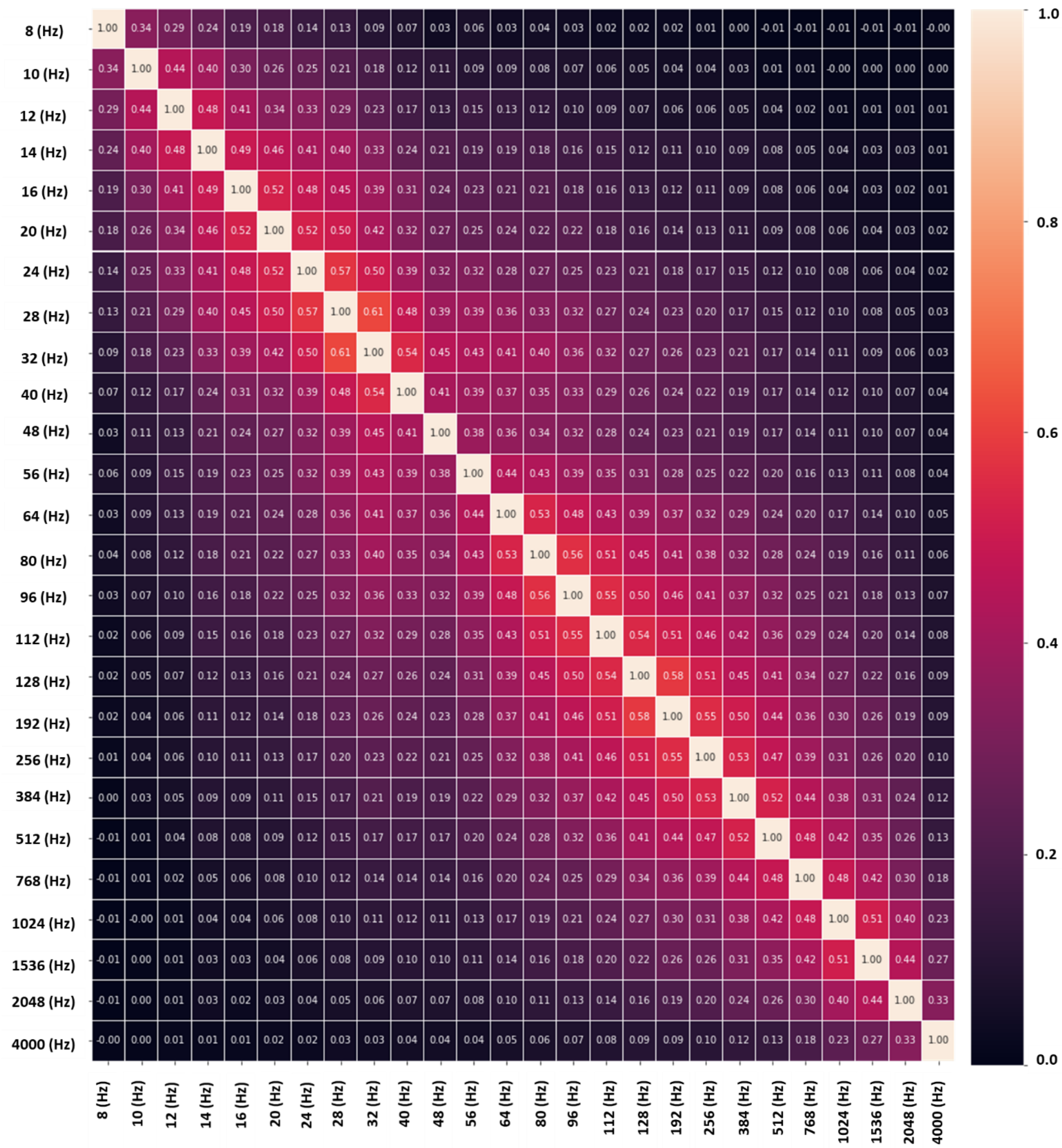

Figure 3.41. Heatmap of training data for the branched network

Comparing the looped and branched networks' heatmaps indicates that non-sequential features of the looped network are more correlated. Due to the more physical connectivity in the looped network, acoustic data captured by the two hydrophones have more overlapping information, leading to more similar contents at distinct features.

To evaluate if a dimension reduction improves classification results, PCA(0.95) was applied to the training and test data, and the algorithms were trained and tested with lower dimension data. Tables 3.38 to 3.46 include classification results for the classification algorithm using lower dimension data.

Table 3.38. The best 1CSVM-PCA with a linear kernel, its optimized parameters, and evaluation metrics using for the branched network

Optimized Parameters for **1CSVM-PCA**					
Kernel	*v*				
linear	10^{-1}				
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.22	0.57	**0.32**	14400
	1	0.88	0.61	0.72	75600
Weighted Average		0.77	0.60	**0.65**	90000
Accuracy				0.60	
F_1-score				0.72	

Table 3.39. The best 1CSVM-PCA with a polynomial kernel, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **1CSVM-PCA**					
Kernel	*v*		*d*		
poly	10^{-6}		6		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.20	0.30	**0.24**	14400
	1	0.84	0.76	0.80	75600
Weighted Average		0.74	0.68	**0.70**	90000
Accuracy				0.68	
F_1-score				0.80	

Table 3.40. The best 1CSVM-PCA with an RBF kernel, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **1CSVM-PCA**					
Kernel		ν		γ	
rbf		10^{-1}		$2{\times}10^{3}$	
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.23	0.28	**0.25**	14400
	1	0.85	0.81	0.83	75600
Weighted Average		0.75	0.72	**0.73**	90000
Accuracy				0.72	
F_1-score				0.83	

Table 3.41. The best iForest-PCA, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **iForest-PCA**					
n_estimators			*contamination*		
8			0.2		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.23	0.29	**0.26**	14400
	1	0.85	0.81	0.83	75600
Weighted Average		0.75	0.72	**0.73**	90000
Accuracy				0.72	
F_1-score				0.83	

Table 3.42. The best LOF-PCA, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **LOF-PCA**					
n_neighbors			*contamination*		
10			0.2		
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.18	0.31	**0.23**	14400
	1	0.84	0.71	0.77	75600
Weighted Average		0.73	0.64	**0.68**	90000
Accuracy				0.64	
F_1-score				0.77	

Table 3.43. The best SVM-PCA with a linear kernel, its optimized parameters, and evaluation metrics for the branched network

Optimized Parameters for **SVM-PCA**					
Kernel	*C*				
linear	1				
Evaluation Metrics					
		Precision	Recall	F_1-score	Support
Class	**-1**	0.23	0.45	**0.31**	14400
	1	0.86	0.70	0.78	75600
Weighted Average		0.76	0.66	**0.70**	90000
Accuracy				0.66	
F_1-score				0.78	

Table 3.44. The best SVM-PCA with a polynomial kernel, its optimized parameters, and evaluation metrics for the branched network

<table>
<tr><td colspan="6">Optimized Parameters for SVM-PCA</td></tr>
<tr><td>Kernel</td><td colspan="2">C</td><td colspan="3">d</td></tr>
<tr><td>poly</td><td colspan="2">10^{-2}, 1, 10, 100</td><td colspan="3">2</td></tr>
<tr><td colspan="6">Evaluation Metrics</td></tr>
<tr><td colspan="2"></td><td>Precision</td><td>Recall</td><td>F_1-score</td><td>Support</td></tr>
<tr><td rowspan="2">Class</td><td>-1</td><td>0.20</td><td>0.06</td><td>0.09</td><td>14400</td></tr>
<tr><td>1</td><td>0.84</td><td>0.95</td><td>0.89</td><td>75600</td></tr>
<tr><td colspan="2">Weighted Average</td><td>0.73</td><td>0.81</td><td>0.76</td><td rowspan="3">90000</td></tr>
<tr><td colspan="4">Accuracy</td><td>0.81</td></tr>
<tr><td colspan="4">F_1-score</td><td>0.59</td></tr>
</table>

Table 3.45. The best SVM-PCA with an RBF kernel, its optimized parameters, and evaluation metrics for the branched network

<table>
<tr><td colspan="6">Optimized Parameters for SVM-PCA</td></tr>
<tr><td>Kernel</td><td colspan="5">C</td></tr>
<tr><td>rbf</td><td colspan="5">100</td></tr>
<tr><td colspan="6">Evaluation Metrics</td></tr>
<tr><td colspan="2"></td><td>Precision</td><td>Recall</td><td>F_1-score</td><td>Support</td></tr>
<tr><td rowspan="2">Class</td><td>-1</td><td>0.23</td><td>0.77</td><td>0.36</td><td>14400</td></tr>
<tr><td>1</td><td>0.91</td><td>0.49</td><td>0.64</td><td>75600</td></tr>
<tr><td colspan="2">Weighted Average</td><td>0.80</td><td>0.54</td><td>0.59</td><td rowspan="3">90000</td></tr>
<tr><td colspan="4">Accuracy</td><td>0.54</td></tr>
<tr><td colspan="4">F_1-score</td><td>0.64</td></tr>
</table>

Table 3.46. The best XGBoost-PCA, its optimized parameters, and evaluation metrics for the branched network

<table>
<tr><td colspan="6">Optimized Parameters for XGBoost-PCA</td></tr>
<tr><td colspan="3">n_estimators</td><td colspan="3">max_depth</td></tr>
<tr><td colspan="3">1000</td><td colspan="3">100, 500, 1000</td></tr>
<tr><td colspan="6">Evaluation Metrics</td></tr>
<tr><td colspan="2"></td><td>Precision</td><td>Recall</td><td>F₁-score</td><td>Support</td></tr>
<tr><td rowspan="2">Class</td><td>-1</td><td>0.25</td><td>0.07</td><td>0.11</td><td>14400</td></tr>
<tr><td>1</td><td>0.84</td><td>0.96</td><td>0.89</td><td>75600</td></tr>
<tr><td colspan="2">Weighted Average</td><td>0.74</td><td>0.81</td><td>0.76</td><td rowspan="3">90000</td></tr>
<tr><td colspan="4">Accuracy</td><td>0.81</td></tr>
<tr><td colspan="4">F₁-score</td><td>0.89</td></tr>
</table>

Tables 3.47 and 3.48 rank the best classifiers trained and tested with data having all and reduced dimensions for the branched network. Since the Weighted Average F_1-score considers F_1-score of both classes based on their proportional data, this metric was selected to rank the algorithms. The Class '-1' F_1-score was regarded as the second criterion for the ranking should the Weighted Average F_1-score is the same for different models. It is worth noting that only the kernels that resulted in the best 1CSVM and SVM models were selected in the ranking tables.

Table 3.47. Ranked performance of classification models on test data with all features for the branched network

Model	Weighted Average F_1-score	Class '-1' F_1-score	Class '1' F_1-score	Accuracy
XGBoost	0.80	0.25	0.91	0.84
LOF	0.76	0.43	0.83	0.74
SVM (rbf)	0.75	0.43	0.81	0.72
iForest	0.73	0.24	0.82	0.71
1CSVM (rbf)	0.59	0.35	0.64	0.54

Table 3.48. Ranked performance of classification models on test data with reduced dimensions for the branched network

Model	Weighted Average F_1-score	Class '-1' F_1-score	Class '1' F_1-score	Accuracy
XGBoost-PCA	0.76	0.11	0.89	0.81
iForest-PCA	0.73	0.26	0.83	0.72
1CSVM-PCA (rbf)	0.73	0.25	0.83	0.72
LOF-PCA	0.68	0.23	0.77	0.64
SVM-PCA (rbf)	0.59	0.36	0.64	0.54

In the following, Tables 3.47 and 3.48 results will be discussed, along with some explanatory figures.

Figures 3.42 and 3.43 show the distributions of the training and test data for the branched network, respectively, at the first three features, i.e., 8 Hz, 10 Hz, and 12 Hz.

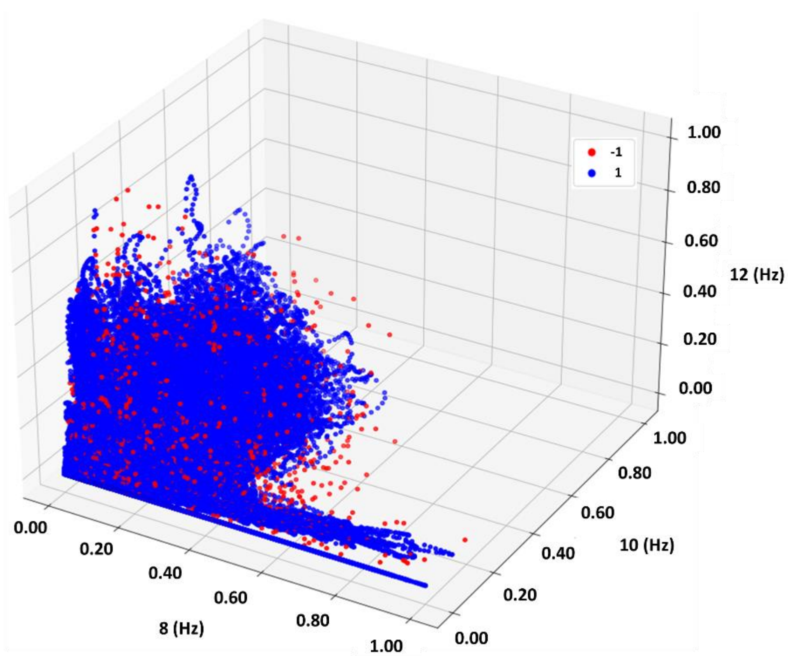

Figure 3.42. Distribution of the first three features of training dataset for the branched network

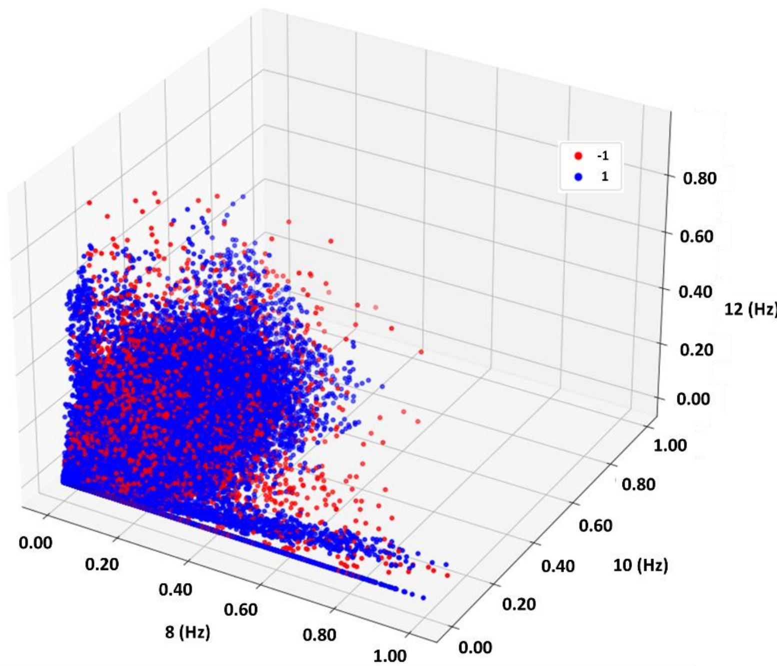

Figure 3.43. Distribution of the first three features of test dataset for the branched network

Both training and test data of the branched networks are mixed. Therefore, none of the classifiers have been able to discern leak and non-leak data well.

Though XGBoost has the highest "Weighted Average F_1-score" on full and reduced dimension test data, it performs poorly to detect leak samples. A reason for the better performance of XGBoost on non-leak data could be the availability of more training non-leak data for the algorithm that allows it to learn better and decrease its training errors.

On test data with all dimensions, both LOF and SVM (rbf) algorithms had the best performance in detecting leak data. LOF performed better since it has a local approach to find data points similar to the point of interest, which helps the algorithm overcome the overall mixed distribution of data with full dimensions. Also, SVM (rbf) had better leak detection since it is a supervised algorithm and takes advantage of knowing leak classes in the training step, and also the RBF non-linear kernel with the large budget for slack variables, i.e., C=100, helps the algorithm to form a hyperplane that separates leak and non-leak data as the best as possible.

iForest has the most inadequate performance in leak data detection. It could be due to the leak and non-leak samples' closeness that makes many leak data seem as inliers.

1CSVM (rbf) ranked 2nd and last in detecting the leak and non-leak samples, respectively. Since 1CSVM only uses non-leak data for training, there might be some non-leak samples in the test data with which the algorithm has not faced in the training step. Therefore, the algorithm misses those previously unseen samples in the test step.

Similar to the looped network, contrary to the paper of Cody et al. (2018) where one-class SVM worked satisfactorily, 1CSVM ranked 5th with original feature data and 3rd with reduced dimension data. It is worth noting that if the "Class '-1' F_1-score" was considered in ranking the classifiers, 1CSVM would rank 3rd for original and reduced dimension datasets.

XGBoost ranked 1st for original and reduced dimension datasets of the branched network. Its superior rank is due to XGBoost's good performance in predicting the majority class, i.e., non-leak, while the algorithm had an unsatisfactory performance on leak data. This point implies that XGBoost highly depends on the number of training samples in each class.

Should the "Class '-1' F_1-score" is selected in ranking the classifiers, LOF and SVM(rbf) rank 1st on the data with original features, and SVM-PCA(rbf) performs the best for the reduced dimension data. Both SVM(rbf) and SVM-PCA(rbf) take advantage of an RBF kernel with slack variables budget C=100, which shows a highly nonlinear hyperplane is required to separate leak and non-leak data.

Tables 3.49 and 3.50 summarize the F_1-score metrics for the looped and branched networks on data with original and reduced dimensions.

Table 3.49. Summary of F_1-score metrics for the looped network

Network	Feature Dimension	Metric	Value
Looped	Original	Mean of Weighted Average F_1-score	0.77
		Mean of Class '-1' F_1-score	0.29
		Mean of Class '1' F_1-score	0.87
	Reduced	Mean of Weighted Average F_1-score	0.73
		Mean of Class '-1' F_1-score	0.25
		Mean of Class '1' F_1-score	0.82

Table 3.50. Summary of F_1-score metrics for the branched network

Network	Feature Dimension	Metric	Value
Branched	Original	Mean of Weighted Average F_1-score	0.73
		Mean of Class '-1' F_1-score	0.34
		Mean of Class '1' F_1-score	0.80
	Reduced	Mean of Weighted Average F_1-score	0.70
		Mean of Class '-1' F_1-score	0.24
		Mean of Class '1' F_1-score	0.79

In the following, the metrics of the two networks are compared.

i. The Mean of Weighted Average F1-score and the Mean of Class '1' F1-score for the looped network are larger than those for the branched network on both data feature types. This difference can be due to the effect of the majority class on the algorithms' performance.

ii. The Mean of Class '-1' F_1-score for the looped network is smaller than that of the branched network on data with original features. Nonetheless, this score for the looped network is larger than that of the branched network on data with reduced features.

iii. SVM had the best performance in detecting the leak class on both feature types of the two networks. However, the branched network's SVM kernels were more nonlinear (rbf kernels) than those for the looped network (polynomial or linear kernels). This difference could stem from a less class separability in the branched network.

iv. XGBoost had the best performance in detecting the non-leak class on both feature types of the two networks. As it was previously mentioned, the superiority of XGBoost for the non-leak data is due to the larger data numbers of the non-leak samples, which allows the algorithm to decrease its training error.

v. XGBoost had a least satisfactory performance in leak class detection than other algorithms. This point might stem from the sensitivity of XGboost to the number of training data.

vi. 1CSVM performed mediocre for all data types and networks, while the algorithm detected leaks in the paper of Cody et al. (2018). In addition to a less separable data distribution of this study, another reason for the 1CSVM's lower rank could be the five percent mislabeled abnormality in the training data. Yet, 1CSVM is supposed to be trained with entirely normal samples. The reason for introducing the five percent masked abnormal training data is simulating real conditions where leak samples might be mislabeled as non-leak.

3.4.3. Influence of Class Ratios in Datasets

In the previous analyses, training and test datasets were composed of imbalanced proportions of classes. However, a problem with imbalanced datasets is that conventional classification learning algorithms are often biased towards the majority class. Therefore, the minority class's misclassification would be higher (Lopez et al., 2013; Patel et al., 2020). This subsection evaluates how the classifiers perform if the training and test datasets include equal ratios of the leak and non-leak classes. Therefore, an original dataset was formed from similar numbers of the leak and non-leak data. Then 80% and 20% percent of the original data were randomly selected for training and test datasets, respectively.

Figure 3.44 shows the parallel plot of the looped network's balanced training data.

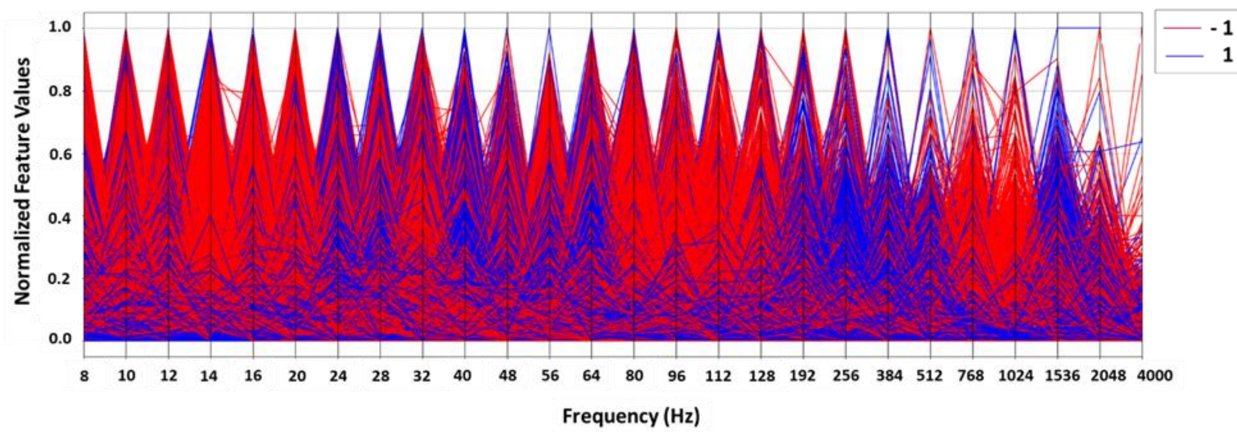

Figure 3.44. A parallel plot of the looped network's balanced training data

Figure 3.44 shows that in the looped network's balanced data, the leak samples become pronounced at most features, particularly at those features dominant in the parallel plot of the looped network's balanced data, see Figure 3.30. Like the imbalanced data's plot, the features where the leak data are highlighted are at frequencies below 150 Hz and between 512 Hz and 1536 Hz. The leak data at higher frequencies could be due to the pump and ambient noises.

Figure 3.45 represents the parallel plot of the looped network's balanced test data.

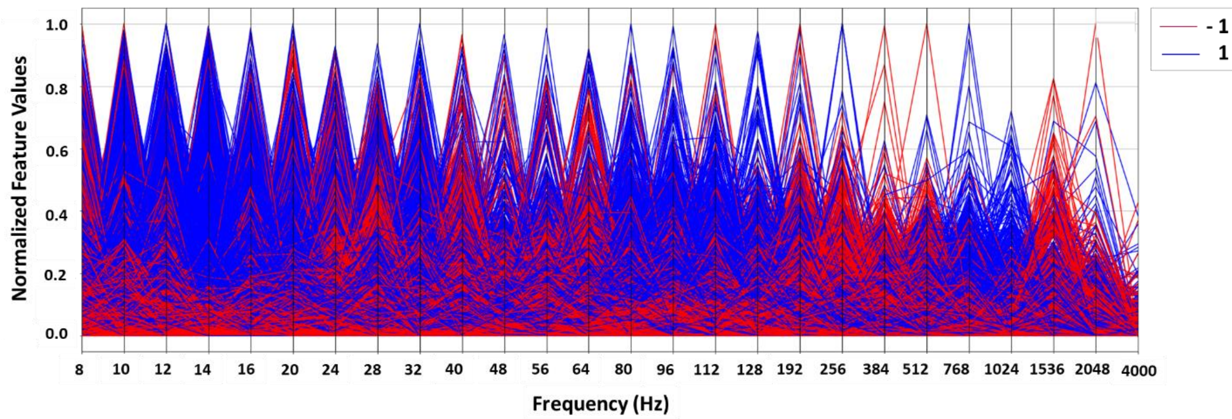

Figure 3.45. A parallel plot of the looped network's balanced test data

Compared to its Imbalanced counterpart, Figure 3.45 shows that the looped network's balanced test data's leak samples are more distributed at different features. Additionally, there are more outliers at higher frequencies that were absent in the plot of imbalanced data. Two other reasons could cause the differences between the parallel plots of the balanced and imbalanced data:

i. The new leak samples added to the balanced data have different frequencies than those in the imbalanced data.

ii. A source data matrix was first generated randomly from the leak and non-leak matrixes with the same ratios of classes. Then the source matrix was shuffled, and the training and test datasets were created randomly. The randomness in the data selection could cause dissimilar leak samples at different frequencies.

Figure 3.46 shows the first three features of balanced training data. Similar to its imbalanced counterpart, data classes are very mixed and with more leak data visually visible. One can see that many non-leak data have close-to-zero values at 10 Hz frequency.

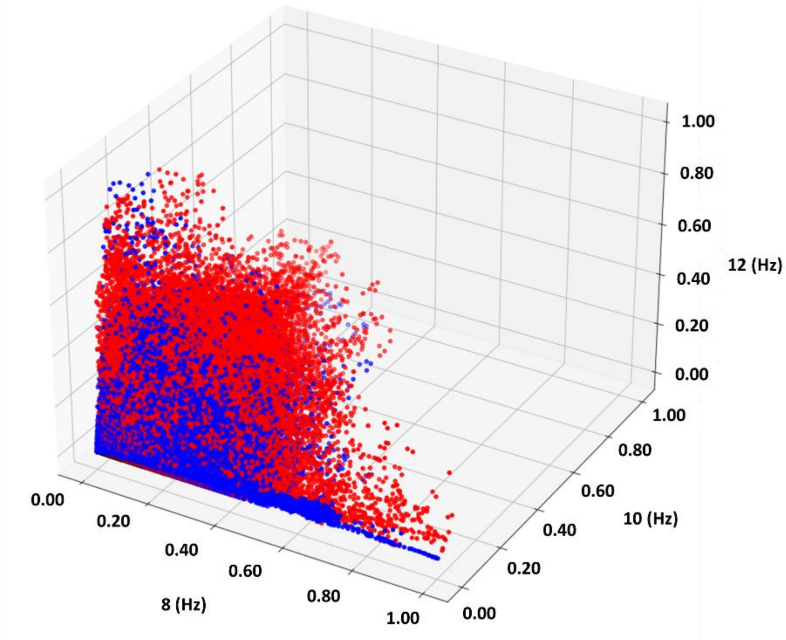

Figure 3.46. Distribution of the first three features of the balanced training dataset for the looped network

Figure 3.47 shows the first three features of balanced test data.

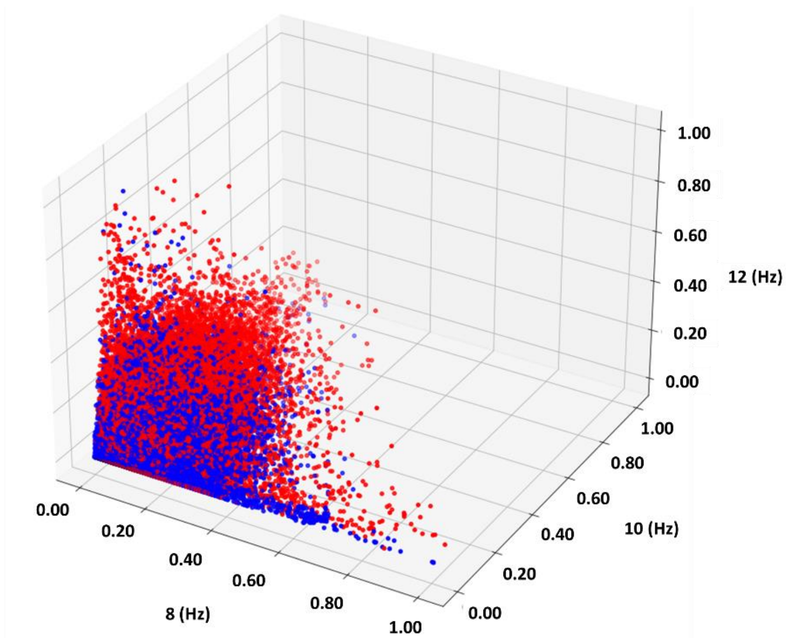

Figure 3.47. Distribution of the first three features of the balanced test dataset for the looped network

Figure 3.47 shows the first three features of balanced test data. Based on the first three features, leak samples have larger non-zero values at feature 12 Hz than those of non-leak.

Tables 3.51 and 3.52 show the five algorithms' results on balanced datasets with the same hyperparameters in Tables 3.27 and 3.28. It is worth noting that Tables 3.27 and 3.28 represent the algorithms' results on imbalanced datasets of the looped network. The number of training data for each leak class in the balanced dataset is 72,000 samples.

Table 3.51. Ranked performance of classification models, with the same algorithm hyperparameters in Table 3.27, on balanced test data with all features for the looped Network

Model	Weighted Average F_1-score	Class '-1' F_1-score	Class '1' F_1-score	Accuracy
XGBoost	0.76	0.77	0.74	0.76
SVM (poly)	0.64	0.56	0.71	0.65
iForest	0.48	0.32	0.63	0.52
1CSVM (rbf)	0.44	0.21	0.66	0.53
LOF	0.43	0.26	0.60	0.48

Table 3.52. Ranked performance of classification models, with the same algorithm hyperparameters in Table 3.28, on balanced test data with reduced features for the looped Network

Model	Weighted Average F_1-score	Class '-1' F_1-score	Class '1' F_1-score	Accuracy
SVM-PCA (linear or poly)	0.61	0.57	0.65	0.61
XGBoost-PCA	0.59	0.59	0.59	0.59
iForest-PCA	0.52	0.38	0.66	0.56
1CSVM-PCA (rbf)	0.48	0.30	0.65	0.54
LOF-PCA	0.47	0.36	0.58	0.50

Comparing Tables 3.51 and 3.52 with Tables 3.27 and 3.28 shows though the algorithms' hyperparameters on balanced data and their counterparts on imbalanced data are the same for the looped network, the algorithms did not rank the same due to the difference in the class ratios.

Figure 3.48 shows the effects of the dataset's class ratios on the five classifiers based on the Weighted Average F_1-score with original features.

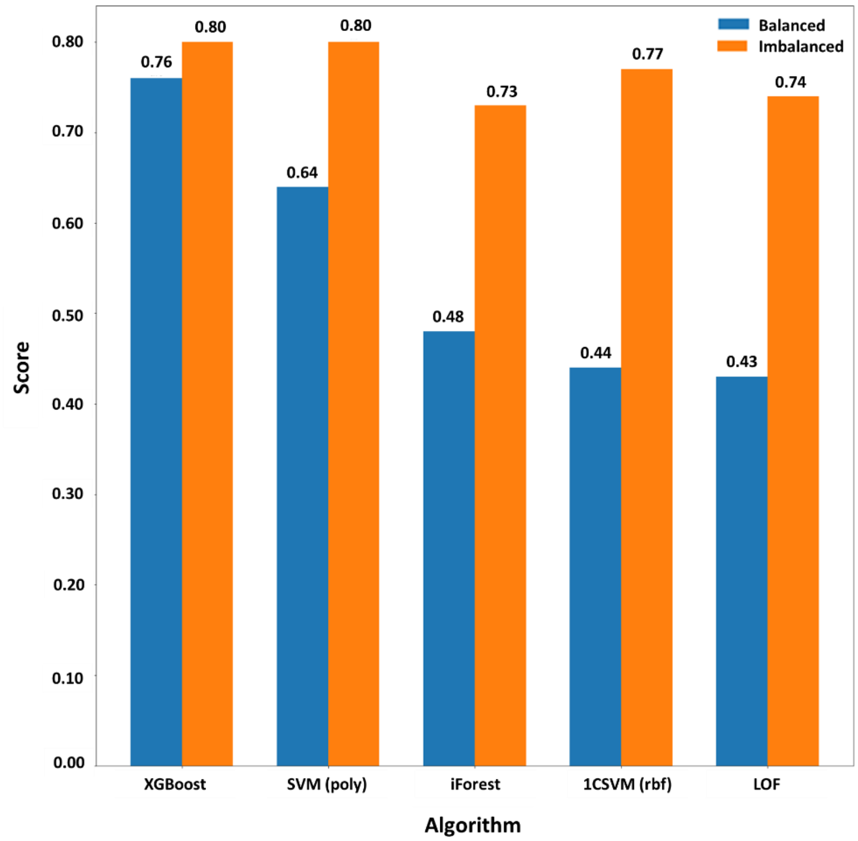

Figure 3.48. Weighted Average F_1-score of the classifiers on balanced and imbalanced datasets with original features of the looped network

Considering the Weighted Average F_1-score, XGBoost, and SVM (poly) have the best performances on balanced and imbalanced datasets, which could be due to SVM's supervised and XGBoost's boosting nature. All classifiers on imbalanced data have higher Weighted Average F_1-scores than on the balanced data. This difference is due to the bias of the algorithms to the majority class in the imbalanced data.

Figure 3.49 represents the Class '-1' F_1-score of the classifiers using balanced and imbalanced data.

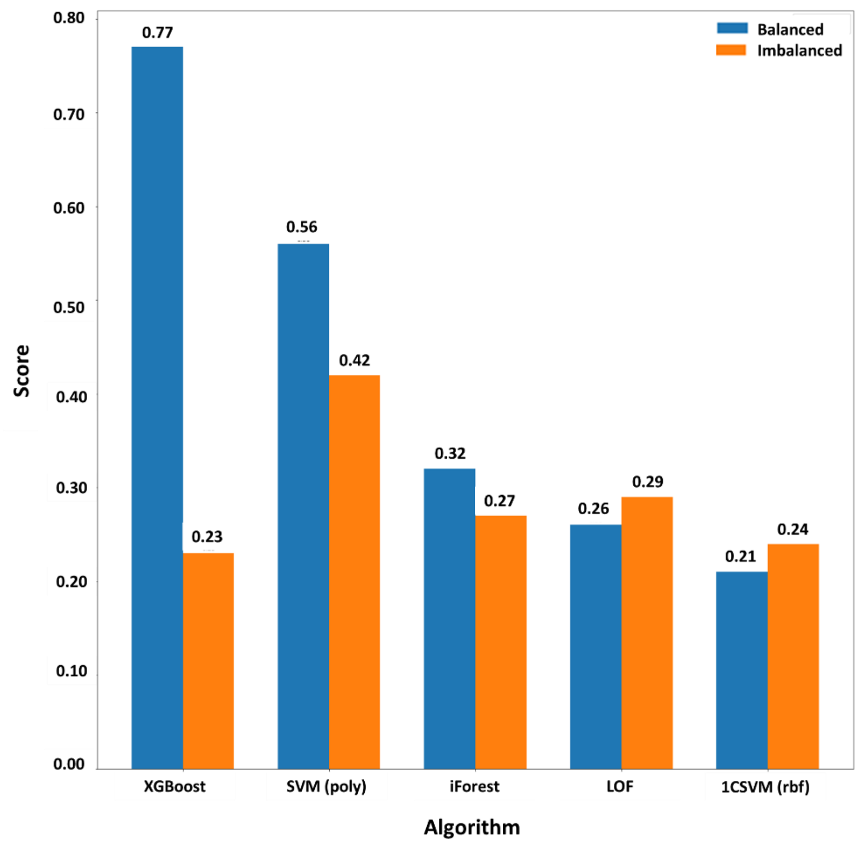

Figure 3.49. Class '-1' F_1-score of the classifiers on balanced and imbalanced datasets with original features of the looped network

One could see the much better performance of XGBoost when applied to balanced data. This better result is due to the higher number of leak data available for XGBoost training that allowed it to fit more leak samples and decrease its training error. Besides, SVM and iForest gave better results using the balanced data. One reason for their better performance is that they see more abnormal data during training. Interestingly, LOF had a lower Class '-1' F_1-score on the balanced data. Reasons for the lower score of the LOF with more leak data could be as the following. In LOF, the sample class is determined based on its distance to k-neighbor samples' classes. Added leak samples may be in the region of non-leak data. Therefore, due to new leak samples' vicinity to non-leak ones, the algorithm might have labeled new leak samples as non-leak. Since decreasing the value of the *n_neighbors* parameter, which represents *k*-neighbor points, would determine a sample's class based on closer samples, it could address this problem. In other words, the *n_neighbors* parameter might be too large, so that it makes the model complex and vulnerable to overfitting when more training data is added. Additionally, since LOF does not employ regularization, introducing more training samples could cause overfitting, which leads to a lower score in the test step. Also, 1CSVM had a lower Class '-1' F_1-score on the balanced dataset. This could be due to the similarity of new abnormal samples to normal data with which 1-CSVM was tested and these normal-resubliming samples, though they are abnormal, increased false negatives.

Figure 3.50 shows the Class '1' F_1-score of the classifiers on balanced and imbalanced datasets with original features

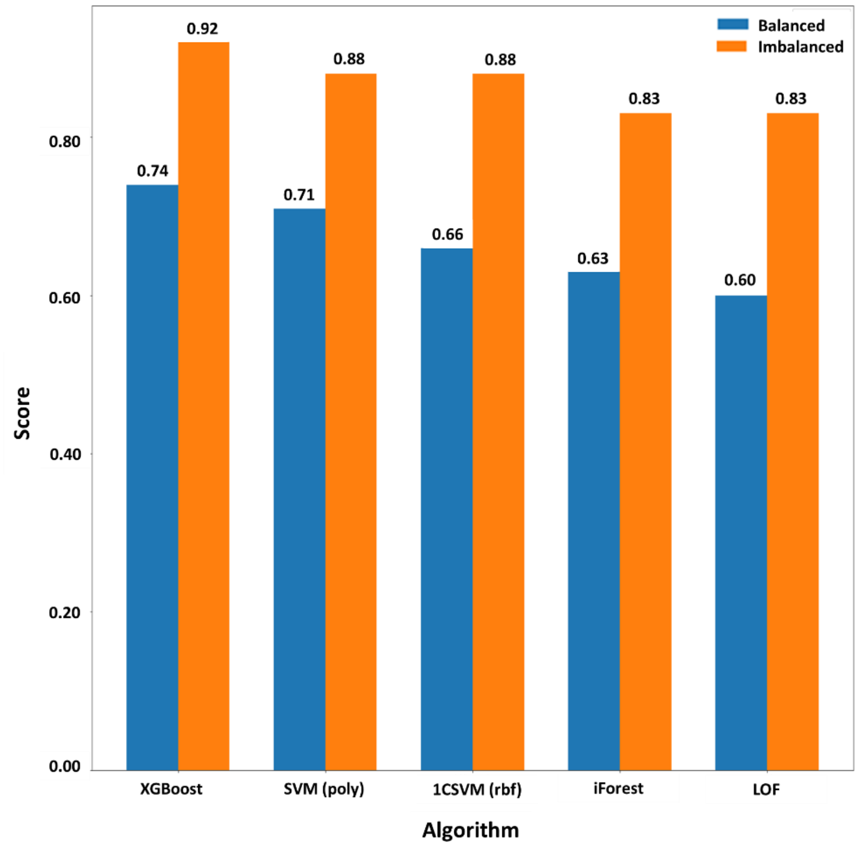

Figure 3.50. Class ‘1’ F_1-score of the classifiers on balanced and imbalanced datasets with original features of the looped network

Based on Figure 3.50, all algorithms have a lower Class ‘1’ F_1-score on balanced data. Regarding the balanced leak and non-leak data's mixed distributions, see Figures 3.46 and 3.47, algorithms would consider added leak data as non-leak ones, which can cause a less accurate pattern recognition during training and more mislabeled data in the test step. The Class ‘1’ F_1-score decrease is more significant for unsupervised algorithms, iForest and LOF, than the supervised ones, including XGBoost and SVM. The largest drop of the score is for the 1CSVM with the value of 0.22 due to more abnormal samples in the balanced training data.

On the balanced data, supervised algorithms, like XGBoost and SVM, and even 1CSVM as a semi-supervised algorithm performed better than their unsupervised counterparts in classifying non-leak samples. The supervised and semi-supervised algorithms' superiority was also the case on the imbalanced data.

Figure 3.51 represents the classifiers' Weighted Average F_1-score on reduced feature data using PCA.

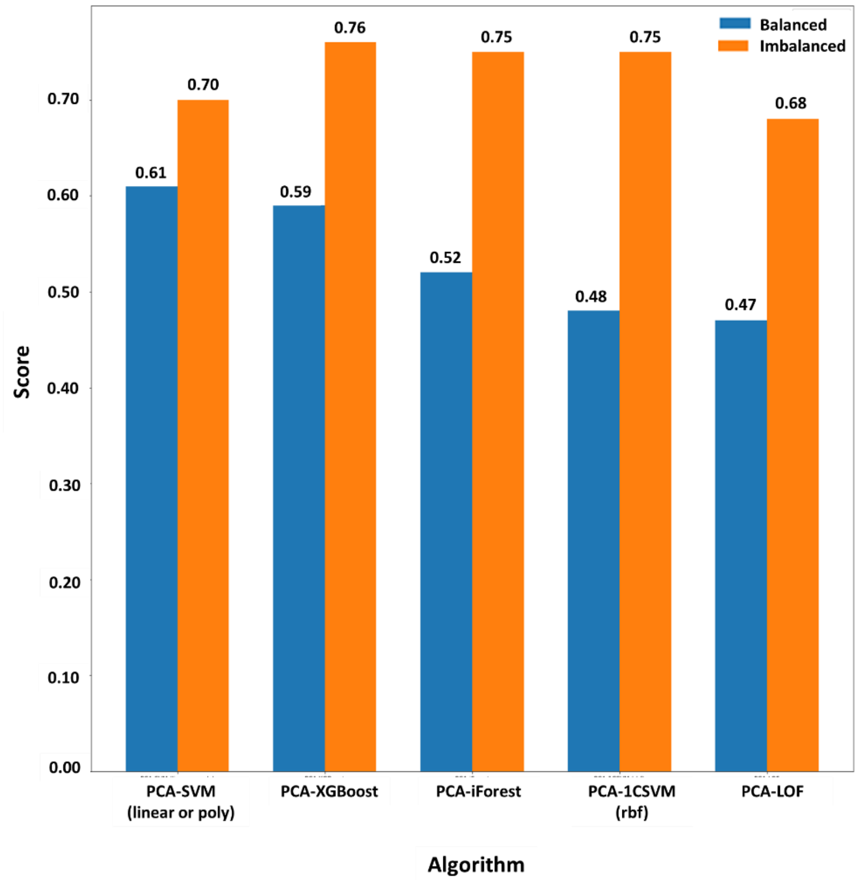

Figure 3.51. Weighted Average F_1-score of the classifiers on balanced and imbalanced datasets with reduced features of the looped network

The classifiers' Weighted Average F1-score decreases on balanced data stem from the drops in the Class '1' F_1-scores, which will be discussed in Figure 3.53.

Figure 3.52 represents the classifiers' Class '-1' F1-scores on balanced and imbalanced datasets with reduced features.

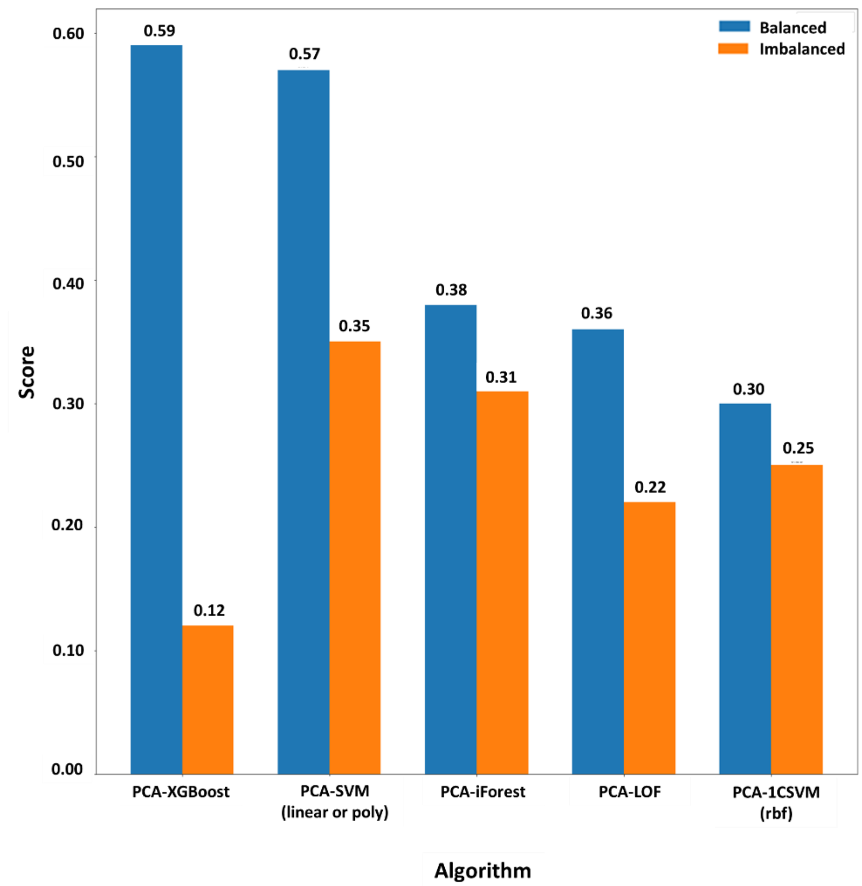

Figure 3.52. Class '-1' F_1-score of the classifiers on balanced and imbalanced datasets with reduced features of the looped network

Similar to data with original features, balancing classes in data with reduced features makes supervised algorithms more accurate than the unsupervised ones. For instance, the Class '-1' F_1-score of XGBoost and SVM become approximately five times and 1.6 times larger when balanced data are used. The reason for the better performance of the classifiers on balanced data is the availability of more leak data in training. Comparing Figures 3.49 and 3.52 indicates the following points.

i. The algorithms rank the same in predicting leak samples using balanced original and reduced dimension data.

ii. Class ‘-1’ F_1-scores increase more significantly on balanced data with reduced features than the balanced data with original features. This difference may imply that PCA makes data more separable when there are equal numbers of data from two classes.

iii. Except for XGBoost, dimension reduction increases the Class ‘-1’ F_1-scores of the algorithms when applied to balanced data.

iv. PCA adversely affects XGBoost’s Class ‘-1’ F_1-score on both balanced and imbalanced datasets. Since XGBoost is a correlation robust algorithm, reducing features may have eliminated useful information in removed features.

Figure 3.53 represents the Class ‘1’ F_1-score of the classifiers on balanced and imbalanced datasets with reduced features. Based on Figure 3.53, iForest slightly outperforms in predicting non-leak samples on balanced data with reduced dimensions. Similar to data with original features, SVM successfully predicts non-leak samples on balanced data. The Class ‘1’ F_1-scores of all classifiers have dropped when applied to balanced data with reduced features. It could be due to the leak and non-leak samples' similarity, especially at low frequencies, making algorithms mislabel data.

Comparing the Class ‘1’ F_1-scores of balanced data with original features versus those with reduced features shows PCA's adverse effect on the algorithms’ performance. This means the feature reduction eliminated features whose non-leak data included valuable information.

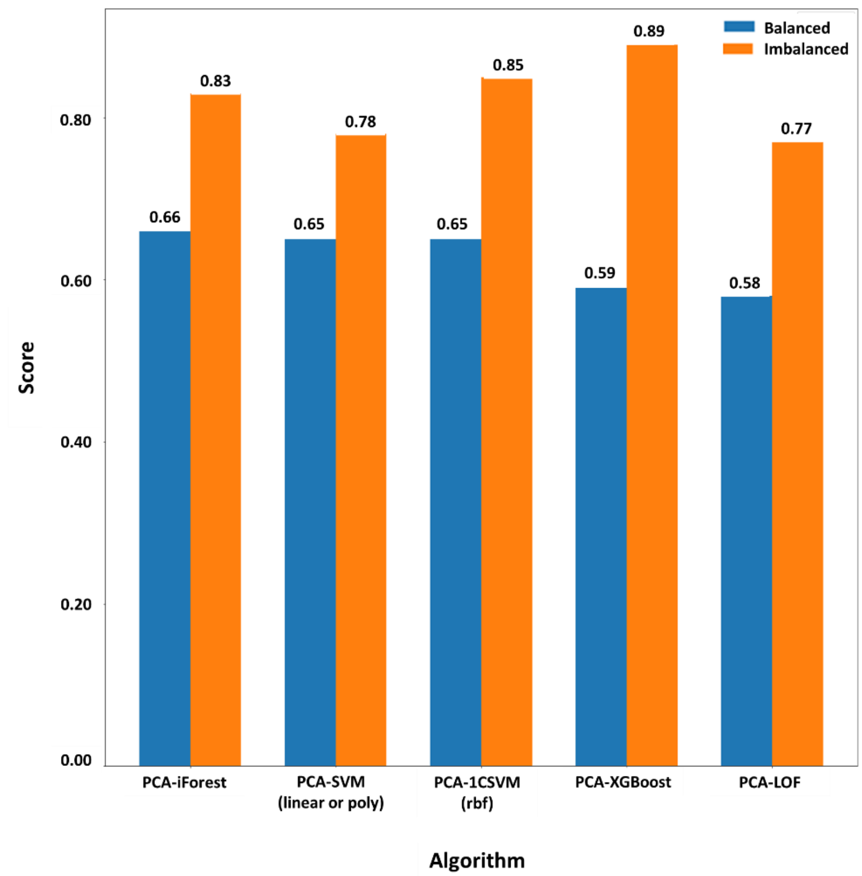

Figure 3.53. Class ‘1’ F_1-score of the classifiers on balanced and imbalanced datasets with reduced features of the looped network

3.5. Conclusions

In this section, acoustic data of the testbed’s two hydrophones and five shallow classifiers, XGBoost, SVM, 1CSVM, LOF, and iForest, were employed to evaluate three objectives: (1) performance of classifiers on imbalanced data of the looped network; (2) influence of the network architecture on classification results; (3) influence of datasets’ class ratios on classifiers. In the following, the results of these subsections are summarized.

i. Performance of classifiers on imbalanced data of the looped network

Parallel plots of training and test data of the looped network show the magnitude of most of the features, approximately at frequencies above 40 (Hz), are zero or very small.

The distributions of training and test data with all features are mixed, and leak and non-leak samples are not easily separable at low-frequency features.

SVM with C=1 and a 2^{nd}-degree polynomial kernel had the best performance in detecting leak samples in data with all features.

Though XGBoost had the highest Weighted Average F_1-score on data with original features, it ranked the last in detecting leak signals and the first in predicting non-leak data. This difference in performance shows XGBoost could be biased to the majority class.

Histograms and heatmaps showed weak correlations among twenty-six original features.

XGBoost-PCA had the highest Weighted Average F_1-score on looped network data with reduced features due to its good performance in predicting non-leak data.

Considering the Weighted Average F_1-score in selecting the best classifiers, SVM-PCA with C=10 and a 2^{nd}-degree polynomial kernel outperformed in predicting leak data.

PCA decreased the Weighted Average F_1-score of all classifiers and specifically adversely affected the Class '-1' F_1-score of the supervised algorithms, SVM and XGBoost.

Neither of the algorithms had a Class '-1' F_1-score greater than 0.5, which means they are less successful at detecting leaks than their performance in predicting non-leak data.

1CSVM was not the best classifier due to the low separability of two classes and leak samples in training data.

ii. Influence of the network architecture on classification results

Parallel plots of the branched network's training and test data indicated that most of the leak samples span between 8 Hz and 16 Hz, 96 Hz and 384 Hz, and 512 Hz and 1024 Hz.

Compared to the same plot of the looped network, leak data are less pronounced in the branched network due to few pipes and less network connectivity, and smaller feature values at higher frequencies due to a more severe attenuation caused by more dead-end pipes.

Non-leak samples of the branched network are denser at lower frequencies such as 56 Hz and 96 Hz and less dense at frequencies higher than 256 Hz.

Both training and test data of the branched network are mixed, and none of the classifiers were able to discern leak and non-leak data well.

Like the looped network, XGBoost has the best performance in predicting non-leak signals and a very low rank in detecting leak samples.

On the branched network's test data with all dimensions, both LOF and SVM (rbf) with C=100 algorithms had the best performance in detecting leak data.

Comparing SVM classifiers, the best SVM of the branched network had a more nonlinear kernel and a larger slack variable budget, RBF kernel and C=100, compared to those of the looped network, polynomial kernel, and C=10, which shows branched network's data are less separable.

iForest ranked last in leak data detection. It can be due to the leak and non-leak samples' closeness that makes many leak data seem as inliers.

Like the looped network, reducing features decreases the F_1-scores in the branched network. The mean F_1-scores for the branched network's data with reduced and original features were smaller than those for the looped network. This difference could be due to a more complex pattern of the branched network's data.

XGBoost had the best performance in detecting the non-leak class on both feature types of the two networks. This superiority of XGBoost for the non-leak data is due to the larger data numbers of the non-leak samples, which allows the algorithm to decrease its training error.

iii. Influence of datasets' class ratios on classifiers

Increasing the number of leak samples in the training and test data so that both leak and non-leak classes equally contribute to the data indicated the following results.

Though the algorithms on balanced and imbalanced data have similar hyperparameters for the looped network, they do not rank the same when balanced data is used for training and testing.

All classifiers on imbalanced data with original features have higher Weighted Average F_1-scores than on the balanced data. This difference is due to the bias of the algorithms to the majority class in the imbalanced data.

Based on the Class '-1' F_1-score, XGBoost outperformed when applied to balanced data with original features. This better result is due to the higher number of leak data available for XGBoost training that allowed it to fit more leak samples and decrease its training error.

SVM and iForest also gave better results using the balanced data. One reason for their better results is that they see more abnormal data during training.

LOF had a lower Class '-1' F_1-score on the balanced data with original features due to the similarity of the leak and non-leak data in a local neighborhood. Decreasing the *n_neighbors* parameter could improve LOF's performance.

1CSVM had a lower Class '-1' F_1-score on the balanced dataset. This is due to more abnormal samples in the balanced training data than those in the imbalanced data, while the algorithm only needs normal data during training.

All algorithms have a lower Class '1' F_1-score on balanced data with original features. Regarding the leak and non-leak data's mixed distributions, algorithms would consider added leak data as non-leak ones, which can cause a less accurate pattern recognition during training and more inaccurately labeled data in the test step.

The Class ‘1’ F_1-score decrease is more significant for unsupervised algorithms, iForest and LOF, than the supervised ones, including XGBoost and SVM.

On balanced data with original features, supervised algorithms, like XGBoost and SVM, and even 1CSVM as a semi-supervised algorithm performed better than their unsupervised counterparts in classifying non-leak samples. The supervised and semi-supervised algorithms’ superiority was also the case on the imbalanced data with original features.

Similar to data with original features, balancing classes in data with reduced features makes supervised algorithms more accurate than the unsupervised ones.

Based on the Class ‘-1’ F_1-score, balanced data with reduced features improved the algorithms’ performance more significantly than the balanced data with original features. This difference may imply that PCA makes data more separable when there are equal numbers of data from two classes.

PCA adversely affects XGBoost’s Class ‘-1’ F_1-score on both balanced and imbalanced datasets. Since XGBoost is a correlation robust algorithm, reducing features may have eliminated useful information in features projected on new dimensions.

Regarding the Class ‘1’ F_1-score of the classifiers on balanced and imbalanced datasets with reduced features, iForest slightly outperforms in predicting non-leak samples on balanced data with reduced dimensions.

The Class ‘1’ F_1-scores of all classifiers have dropped when applied to balanced data with reduced features. It could be due to the leak and non-leak samples' similarity, especially at low frequencies, making algorithms mislabel data.

Overall, based on wavelet coefficient magnitudes extracted from acoustic data, leak signals are less pronounced in the branched network than in the looped network. More balanced data increases classifiers’ performance in detecting leaks, especially for tree-based algorithms like XGBoost. When leak and non-leak data are mixed and classes are balanced, two solutions can improve classification results: (1) using tree-based methods such XGBoost and random forest; (2) reducing feature dimension methods such as PCA. The former uses the most distinctive features to classify new data, regardless of feature collinearity, and the latter generates more informative features through projecting original features on new axes.

A reason for the classifiers' poor performance in predicting new leak cases, even with balanced data, can be the mixed distribution of classes that either originate from the complex nature of the raw acoustic signals or the wavelet coefficient transform method’ inefficiency in feature extraction. If the latter is the case, other feature extraction such as those in voice recognition research can improve classification results. Since the wavelet transform feature extraction method was time-consuming and made algorithm training time-complex, this feature extraction method does not fit a real-time leak detection platform that depends on highlighting anomalies in high-resolution time scales and needs fast preprocessing procedures.

3.6. References

Abokifa, A. A., Haddad, K., Lo, C., & Biswas, P. (2019). Real-Time Identification of Cyber-Physical Attacks on Water Distribution Systems via Machine Learning–Based Anomaly Detection Techniques. *Journal of Water Resources Planning and Management*, *145*(1), 04018089. https://doi.org/10.1061/(asce)wr.1943-5452.0001023

Adamczyk, M., Genzel, L., Dresler, M., Steiger, A., & Friess, E. (2015). Automatic Sleep Spindle Detection and Genetic Influence Estimation Using Continuous Wavelet Transform. *Frontiers in Human Neuroscience*, *9*, 1–20. https://doi.org/10.3389/fnhum.2015.00624

Ahadi, M., & Bakhtiar, M. S. (2010). Leak detection in water-filled plastic pipes through the application of tuned wavelet transforms to Acoustic Emission signals. *Applied Acoustics*, *71*(7), 634–639. https://doi.org/10.1016/j.apacoust.2010.02.006

Almeida, F., Brennan, M., Joseph, P., Whitfield, S., Dray, S., & Paschoalini, A. (2014). On the Acoustic Filtering of the Pipe and Sensor in a Buried Plastic Water Pipe and its Effect on Leak Detection: An Experimental Investigation. *Sensors*, *14*(3), 5595–5610. https://doi.org/10.3390/s140305595

Bekkerman, R. (2015, August). *The Present and the Future of the KDD Cup Competition: an Outsider's Perspective*. https://www.linkedin.com/pulse/present-future-kdd-cup-competition-outsiders-ron-bekkerman/

Bennett, J., Elkan, C., Liu, B., Smyth, P., & Tikk, D. (2007). KDD Cup and workshop 2007. *ACM SIGKDD Explorations Newsletter*, *9*(2), 51–52. https://doi.org/10.1145/1345448.1345459

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer.

Bohorquez, J., Alexander, B., Simpson, A. R., & Lambert, M. F. (2020). Leak Detection and Topology Identification in Pipelines Using Fluid Transients and Artificial Neural Networks. *Journal of Water Resources Planning and Management*, *146*(6), 04020040. https://doi.org/10.1061/(asce)wr.1943-5452.0001187

Bohorquez, J., Simpson, A. R., Lambert, M. F., & Alexander, B. (2021). Merging Fluid Transient Waves and Artificial Neural Networks for Burst Detection and Identification in Pipelines. *Journal of Water Resources Planning and Management*, *147*(1), 04020097. https://doi.org/10.1061/(asce)wr.1943-5452.0001296

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: Identifying Density-Based Local Outliers. *ACM SIGMOD Record*, *29*(2), 93–104. https://doi.org/10.1145/335191.335388

Butterfield, J. D., Krynkin, A., Collins, R. P., & Beck, S. B. M. (2017). Experimental investigation into vibro-acoustic emission signal processing techniques to quantify leak flow rate in plastic water distribution pipes. *Applied Acoustics*, *119*, 146–155. https://doi.org/10.1016/j.apacoust.2017.01.002

Chandy, S. E., Rasekh, A., Barker, Z. A., & Shafiee, M. E. (2019). Cyberattack Detection Using Deep Generative Models with Variational Inference. *Journal of Water Resources Planning and Management*, *145*(2), 04018093. https://doi.org/10.1061/(asce)wr.1943-5452.0001007

Chen, T., & Guestrin, C. (2016). XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 498–511. https://doi.org/10.1145/2939672.2939785

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*(1), 38–51. https://doi.org/10.1186/s12864-019-6413-7

Chuang, W.-Y., Tsai, Y.-L., & Wang, L.-H. (2019). Leak Detection in Water Distribution Pipes Based on CNN with Mel Frequency Cepstral Coefficients. *Proceedings of the 2019 3rd International Conference on Innovation in Artificial Intelligence - ICIAI 2019*, 83–86. https://doi.org/10.1145/3319921.3319926

Cody, R. A., & Narasimhan, S. (2020). A field implementation of linear prediction for leak-monitoring in water distribution networks. *Advanced Engineering Informatics*, *45*, 101103. https://doi.org/10.1016/j.aei.2020.101103

Cody, R., Harmouche, J., & Narasimhan, S. (2018). Leak detection in water distribution pipes using singular spectrum analysis. *Urban Water Journal*, *15*(7), 636–644. https://doi.org/10.1080/1573062x.2018.1532016

Cui, Z., Gao, Y., Hu, J., Tian, S., & Cheng, J. (2020). LOS/NLOS Identification for Indoor UWB Positioning Based on Morlet Wavelet Transform and Convolutional Neural Networks. *IEEE Communications Letters*, 1. https://doi.org/10.1109/lcomm.2020.3039251

Ebrahimkhanlou, A., Dubuc, B., & Salamone, S. (2019). A generalizable deep learning framework for localizing and characterizing acoustic emission sources in riveted metallic panels. *Mechanical Systems and Signal Processing*, *130*, 248–272. https://doi.org/10.1016/j.ymssp.2019.04.050

Ebrahimkhanlou, A., & Salamone, S. (2018). Single-Sensor Acoustic Emission Source Localization in Plate-Like Structures Using Deep Learning. *Aerospace*, *5*(2), 50. https://doi.org/10.3390/aerospace5020050

Erfani, S. M., Rajasegarar, S., Karunasekera, S., & Leckie, C. (2016). High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. *Pattern Recognition*, *58*, 121–134. https://doi.org/10.1016/j.patcog.2016.03.028

Ferrante, M., & Brunone, B. (2003). Pipe system diagnosis and leak detection by unsteady-state tests. 2. Wavelet analysis. *Advances in Water Resources*, *26*(1), 107–116. https://doi.org/10.1016/s0309-1708(02)00102-1

Gao, Y., Brennan, M. J., Joseph, P. F., Muggleton, J. M., & Hunaidi, O. (2005). On the selection of acoustic/vibration sensors for leak detection in plastic water pipes. *Journal of Sound and Vibration*, *283*(3–5), 927–941. https://doi.org/10.1016/j.jsv.2004.05.004

Guo, G., Yu, X., Liu, S., Ma, Z., Wu, Y., Xu, X., Wang, X., Smith, K., & Wu, X. (2021). Leakage Detection in Water Distribution Systems Based on Time–Frequency Convolutional Neural Network. *Journal of Water Resources Planning and Management*, *147*(2), 04020101. https://doi.org/10.1061/(asce)wr.1943-5452.0001317

Hariri, S., Carrasco Kind, M., & Brunner, R. J. (2019). Extended Isolation Forest. *IEEE Transactions on Knowledge and Data Engineering*, 1. https://doi.org/10.1109/tkde.2019.2947676

Harmouche, J., & Narasimhan, S. (2020). Long-Term Monitoring for Leaks in Water Distribution Networks Using Association Rules Mining. *IEEE Transactions on Industrial Informatics*, *16*(1), 258–266. https://doi.org/10.1109/tii.2019.2911064

Hastie, T., Tibshirani, R., & Friedman, J. (2020). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics)* (2nd ed.). Springer.

Hossin, M., & Sulaiman, M. N. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*, *5*(2), 01–11. https://doi.org/10.5121/ijdkp.2015.5201

Hunaidi, O., & Chu, W. T. (1999). Acoustical characteristics of leak signals in plastic water distribution pipes. *Applied Acoustics*, *58*(3), 235–254. https://doi.org/10.1016/s0003-682x(99)00013-4

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R (Springer Texts in Statistics)* (1st ed. 2013, Corr. 7th printing 2017 ed.). Springer.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Kampelopoulos, D., Papastavrou, G. N., Kousiopoulos, G. P., Karagiorgos, N., Goudos, S. K., & Nikolaidis, S. (2020). Machine Learning Model Comparison for Leak Detection in Noisy Industrial Pipelines. *2020 9th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, 18–22. https://doi.org/10.1109/mocast49295.2020.9200261

Kang, J., Park, Y.-J., Lee, J., Wang, S.-H., & Eom, D.-S. (2018). Novel Leakage Detection by Ensemble CNN-SVM and Graph-Based Localization in Water Distribution Systems. *IEEE Transactions on Industrial Electronics*, *65*(5), 4279–4289. https://doi.org/10.1109/tie.2017.2764861

Lee, G., Gommers, R., Waselewski, F., Wohlfahrt, K., & O'Leary, A. (2019). PyWavelets: A Python package for wavelet analysis. *Journal of Open Source Software*, *4*(36), 1237. https://doi.org/10.21105/joss.01237

Li, R., Huang, H., Xin, K., & Tao, T. (2014). A review of methods for burst/leakage detection and location in water distribution systems. *Water Supply*, *15*(3), 429–441. https://doi.org/10.2166/ws.2014.131

Li, S., Song, Y., & Zhou, G. (2018). Leak detection of water distribution pipeline subject to failure of socket joint based on acoustic emission and pattern recognition. *Measurement*, *115*, 39–44. https://doi.org/10.1016/j.measurement.2017.10.021

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. *2008 Eighth IEEE International Conference on Data Mining*, 413–422. https://doi.org/10.1109/icdm.2008.17

López, V., Fernández, A., García, S., Palade, V., & Herrera, F. (2013). An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Information Sciences*, *250*, 113–141. https://doi.org/10.1016/j.ins.2013.07.007

Mallat, S. (1999). *A Wavelet Tour of Signal Processing: The Sparse Way* (2nd Edition). Academic Press, London, United Kingdom.

Martini, A., Troncossi, M., & Rivola, A. (2015). Automatic Leak Detection in Buried Plastic Pipes of Water Supply Networks by Means of Vibration Measurements. *Shock and Vibration*, *2015*, 1–13. https://doi.org/10.1155/2015/165304

Montejo, L. A., & Suarez, L. E. (2006). Wavelet-based Identification of Site Frequencies from Earthquake Records. *Journal of Earthquake Engineering*, *10*(4), 565–594. https://doi.org/10.1080/13632460609350610

Mostavi, A., Kamali, N., Tehrani, N., Chi, S.-W., Ozevin, D., & Indacochea, J. E. (2017). Wavelet based harmonics decomposition of ultrasonic signal in assessment of plastic strain in aluminum. *Measurement*, *106*, 66–78. https://doi.org/10.1016/j.measurement.2017.04.013

Mounce, S. R., Day, A. J., Wood, A. S., Khan, A., Widdop, P. D., & Machell, J. (2002). A neural network approach to burst detection. *Water Science and Technology*, *45*(4–5), 237–246. https://doi.org/10.2166/wst.2002.0595

Pasupa, K., & Sunhem, W. (2016, October). A comparison between shallow and deep architecture classifiers on small dataset. *2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE)*. https://doi.org/10.1109/iciteed.2016.7863293

Patel, H., Singh Rajput, D., Thippa Reddy, G., Iwendi, C., Kashif Bashir, A., & Jo, O. (2020). A review on classification of imbalanced data for wireless sensor networks. *International Journal of Distributed Sensor Networks*, *16*(4), 155014772091640. https://doi.org/10.1177/1550147720916404

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

Rashid, S., Akram, U., & Khan, S. A. (2015). WML: Wireless Sensor Network based Machine Learning for Leakage Detection and Size Estimation. *Procedia Computer Science*, *63*, 171–176. https://doi.org/10.1016/j.procs.2015.08.329

Sarrafi, A., Mao, Z., Niezrecki, C., & Poozesh, P. (2018). Vibration-based damage detection in wind turbine blades using Phase-based Motion Estimation and motion magnification. *Journal of Sound and Vibration*, *421*, 300–318. https://doi.org/10.1016/j.jsv.2018.01.050

Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the Support of a High-Dimensional Distribution. *Neural Computation*, *13*(7), 1443–1471. https://doi.org/10.1162/089976601750264965

Scholkopf, B., & Smola, A. J. (2018). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning series)* (Reprint ed.). The MIT Press.

Soro, B., & Lee, C. (2019). A Wavelet Scattering Feature Extraction Approach for Deep Neural Network Based Indoor Fingerprinting Localization. *Sensors*, *19*(8), 1790. https://doi.org/10.3390/s19081790

Stéphane, M. (2009). *A Wavelet Tour of Signal Processing* (Third Edition). Academic Press. https://doi.org/10.1016/B978-0-12-374370-1.00020-3

Taghvaei, S. M. (2009). Fast Leak Detection and Location of Water Pipelines Based Using Orthogonal Wavelet Transform. *ICPTT 2009*, 186–192. https://doi.org/10.1061/41073(361)19

Talagala, P. D., Hyndman, R. J., & Smith-Miles, K. (2020). Anomaly Detection in High-Dimensional Data. *Journal of Computational and Graphical Statistics*, 1–15. https://doi.org/10.1080/10618600.2020.1807997

Taormina, R., & Galelli, S. (2018). Deep-Learning Approach to the Detection and Localization of Cyber-Physical Attacks on Water Distribution Systems. *Journal of Water Resources Planning and Management*, *144*(10), 04018065. https://doi.org/10.1061/(asce)wr.1943-5452.0000983

Taormina, R., Galelli, S., Tippenhauer, N. O., Salomons, E., Ostfeld, A., Eliades, D. G., Aghashahi, M., Sundararajan, R., Pourahmadi, M., Banks, M. K., Brentan, B. M., Campbell, E., Lima, G., Manzi, D., Ayala-Cabrera, D., Herrera, M., Montalvo, I., Izquierdo, J., Luvizotto, E., … Ohar, Z. (2018). Battle of the Attack Detection Algorithms: Disclosing Cyber Attacks on Water Distribution Networks. *Journal of Water Resources Planning and Management*, *144*(8), 04018048. https://doi.org/10.1061/(asce)wr.1943-5452.0000969

Xiao, R., Hu, Q., & Li, J. (2019). Leak detection of gas pipelines using acoustic signals based on wavelet transform and Support Vector Machine. *Measurement*, *146*, 479–489. https://doi.org/10.1016/j.measurement.2019.06.050

You, L., Peng, Q., Xiong, Z., He, D., Qiu, M., & Zhang, X. (2020). Integrating aspect analysis and local outlier factor for intelligent review spam detection. *Future Generation Computer Systems*, *102*, 163–172. https://doi.org/10.1016/j.future.2019.07.044

Yu, L., & Li, S. Z. (2017). Acoustic emission (AE) based small leak detection of galvanized steel pipe due to loosening of screw thread connection. *Applied Acoustics*, *120*, 85–89. https://doi.org/10.1016/j.apacoust.2017.01.016

Zhang, Y., Chen, S., Li, J., & Jin, S. (2014). Leak detection monitoring system of long distance oil pipeline based on dynamic pressure transmitter. *Measurement*, *49*, 382–389. https://doi.org/10.1016/j.measurement.2013.12.009

4. CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

4.1. Summary of the Dissertation

To decrease water loss in aging water networks, leakage that may account for water waste up to fifty percent of produced water should be detected and controlled. There have been multiple methods to detect leaks that utilize hardware and software methods or combinations, which have shown promising solutions for leak detection. Employing sensors and analyzing their data with statistical algorithms have become successful methods for leak detection. In addition to developing accurate and economic sensing devices, the availability of inexpensive computational resources has made sensory data analysis more feasible.

This study's research objective was to design and assemble a laboratory-scale water distribution system where leaks could be induced, the effects of leak types and network changes on leak characteristics could be assessed, and sensor data could be analyzed to detect leaks. Towards these goals, this dissertation describes a research testbed to evaluate leak characteristics and detect leaks using acoustic data and shallow classifiers.

4.2. Laboratory Scale Water Distribution System

Section 2 described the design and assembly of a laboratory-scale water distribution system. Regarding limited space availability, the testbed was distorted from a dimensional point of view. The testbed had looped and branched architectures and was 7.35 m long and 4.9 m wide. The test bench was composed of a supply line and a distribution section where three types of sensors, hydrophone, dynamic pressure sensor, and accelerometer, measured system changes caused by induced leaks. Four types of leaks were induced in a pipe section, including orifice leak, longitudinal and circumferential cracks, and gasket leak. The orifice leak and longitudinal and circumferential cracks had a 2 mm^2 cross-sectional area. Leak flow rates differed based on leak characteristics. Two consumption flows were generated to simulate actual conditions and create consumption noises in the sensor data.

The testbed met expectations by inducing leaks whose flows were less than 30 percent of the total input water to the testbed. Comparing the research's leak characteristics with leaks reported in other research showed that the induced leaks' characteristics met the design's objective considering the testbed size and flow and pressure constraints.

All sensor data were analyzed by six types of plots and ten features to evaluate (1) how the network architecture change affected leak characteristics; (2) how a change of leak type affected its signature. Due to the inconsistent patterns and similar magnitudes of the plots and features, the sixteen evaluation criteria could not discern the leak types or identify the network. However, based on the sensor's locations, some criteria could help detect leaks, distinguish leak types, or identify the network change.

The only feature that discerned the leak types was spectral centroids of P1 in both networks. Features that identified the network change were power spectral entropy of A2, the fundamental frequency and zero-crossing rate of H1, the absolute value of P1, spectral centroids of P1 in both networks, and RMS and standard deviation of P1 and P2. These features and sensors indicated that based on the information extent they represented to differentiate leaks and network architecture, the sensors ranked as following: (1) dynamic pressure sensor; (2) hydrophone; (3) accelerometer. The analyses indicated that leak signatures in the measured data depend on sensor location.

Leak:NoLeak plots of hydrophone data without demand and sound indicated when a leak's water outflow included more airflow and had a more irregular shape, its leak:noleak ration was larger. The shape of leaks' water output also affected their dominant frequencies. For example, based on H2 data with zero demand and no sound, GL in the looped network with a dominant frequency of 40 Hz had the largest dominant frequency amplitude than other leaks.

Based on acceleration data, though leak LDI magnitudes became larger with leak flow increase in the branched network, this was not the case in the looped network that is not in agreement with studies of Yazdekhasti et al. (2016) and Yazdekhasti et al. (2017) where LDI increased with higher leak flow rates. Based on dynamic pressure and acceleration data, since LDI increased with higher demands, we can expect larger LDI magnitudes at zones with higher consumptions.

4.3. Leak Detection via Shallow Classifiers and Using Hydrophone Data

Section 3 elaborated on applying the wavelet transform to extract features from hydrophones' acoustic data and employing them to predict leaks using five shallow classifiers, XGboost, SVM, 1CSVM, iForest, and LOF.

A complex Morlet mother wavelet was applied to sixty leak and non-leak signals to extract features. The moduli of the wavelet coefficients at selected frequencies were calculated as a large-scale feature matrix. Since there was redundant information in the original feature matrix, features were subsampled at frequencies where the wavelet coefficients' moduli varied significantly. It was shown that bandpass filters could eliminate useful feature information, particularly at low frequencies where leak signals were present. The subsampled feature matrix was employed to detect leaks via the classifiers. The data matrix was split into 80% for training and 20% for the test. Twenty percent abnormal data was added to the training data, whose majority were non-leak samples to simulate real conditions. Moreover, the test data included 20% leak and 80% non-leak samples.

Parallel plots of the looped network's training and test data showed most abnormal data, that were leak signals, have frequencies less than 150 Hz and are mainly focused at the frequency of 14 Hz, and some are located between 64 Hz and 128 Hz. Moreover, distribution plots of the looped network's training and test data at 8 Hz, 10 Hz, and 12 Hz frequencies showed leak and non-leak samples were mixed.

SVM with a 2nd-degree polynomial kernel had the best performance in predicting leak signals on the looped network's imbalanced data with original features. On the other hand, XGBoost ranked 1st in predicting non-leak samples. A dimension reduction with PCA was applied to the training and test data. Running the classifiers on the reduced dimensional data gave lower F_{1}-scores. These lower scores could stem from weak correlations among features, where reducing features could remove useful information. A visual feature evaluation using heatmaps also confirmed this reason.

To evaluate how a different network architecture could affect the detection methodologies, the mentioned techniques and algorithms were employed on the branched network's acoustic data with the same class ratios. Results showed that leak signature was less pronounced in the branched networks' acoustic data, and detecting leaks required algorithms with more nonlinearity. The parallel plots of the branched network's training and test data represented ensembles of leak data mainly located at frequencies 14 Hz and 768 Hz. Like the Lopped network, training and test data of the branched network were mixed, and none of the classifiers were able to discern leak and non-leak data well. XGBoost had the best performance in predicting non-leak signals and the worst results in detecting leak samples. On the branched network's test data with all dimensions, both LOF and SVM (rbf) with C=100 algorithms had the best performance in predicting leak samples. Comparing SVM classifiers, the best SVM of the branched network had a more nonlinear kernel and a larger slack variable budget, RBF kernel and *C*=100, compared to those of the looped network, polynomial kernel, and *C*=10. Like the looped network, reducing features decreased the F_1-scores in the branched network. The mean F_1-scores for the branched network's data with reduced and original features were smaller than those for the looped network.

Making the number of leak and non-leak samples equal in training and test data showed that SVM and iForest gave better results using the balanced data than the imbalanced data. Moreover, XGBoost outperformed in predicting leak samples when applied to balanced data with original features. All algorithms performed worse in predicting non-leak samples on balanced data with original features. Regarding the leak and non-leak data's mixed distributions, algorithms would consider added leak data as non-leak ones, which could cause a less accurate pattern recognition during training and more mislabeled data in the test step. The Class '1' F_1-score decrease was more significant for unsupervised algorithms, iForest and LOF, than the supervised ones, including XGBoost and SVM.

Additionally, on balanced data with original features, supervised algorithms, like XGBoost and SVM, and even 1CSVM as a semi-supervised algorithm performed better than their unsupervised counterparts in classifying non-leak samples. The supervised and semi-supervised algorithms' superiority was also the case on the imbalanced data with original features. Balanced data with reduced features improved the algorithm's performance more significantly than the balanced data with original features. This difference might imply that PCA made data more separable when there are equal numbers of data from two classes. PCA adversely affected XGBoost's Class '-1' F_1-score on both balanced and imbalanced datasets. Since XGBoost is a correlation robust algorithm, reducing features could have eliminated useful information in removed features. The algorithms' capability to predict non-leak samples dropped when the classifiers applied to balanced data with reduced features. It could be due to the leak and non-leak samples' similarity, especially at low frequencies, making algorithms mislabel data.

4.4. Suggestions for Future Research

4.4.1. Improvements in the Testbed

EPANET simulations showed that flow velocity in the distribution section was minimal. For instance, the distribution section's velocity changed between 0 m/s and 0.003 m/s when the orifice leak was induced. Moreover, the maximum height of water jets at leak locations was about 3 m, which was for the orifice leak. These points indicate that the pump was not powerful enough to supply enough pressure at pipes. Employing a more powerful pump can increase water velocity and pressure and make these parameters resemble more actual conditions.

Since the dynamic pressure sensors represented pressure variations in pipes, there was no information about the testbed's absolute pressure. Installing absolute pressure gauges in the supply line and distribution section could provide data about real-time pressure in the network and whether water pressure adjustments were required for a more actual water network. As it was noted, the testbed was distorted from a dimensional analysis viewpoint. Should a larger area be available, with dimensions of 77 m and 30 m, the Micropolis virtual network's extracted section in Figure 2.1 could be simulated by an undistorted testbed. Such a large-scale testbed would provide parameters representing a more realistic water network, especially for pressure and acceleration variables.

4.4.2. Improvements in Leak Detection with Acoustic Data

The shallow classifiers were not very successful at predicting leak samples. At the best condition, XGBoost had an F_1-score of 0.77 in predicting leak data on the looped network's balanced data with original features. This score was even less than 0.50 for all algorithms on the looped and branched networks' imbalanced data with original and reduced dimensions. The unsatisfactory performance of the algorithms could be due to the following reasons.

i. The wavelet transform has not been an effective method to extract features to make leak and non-leak samples more separable. Other features such as those of voice recognition field, like Mel frequency cepstral coefficients (MFCC), linear prediction coefficients (LPC), linear prediction cepstral coefficients (LPCC), linear spectral frequencies (LSF), and perceptual linear prediction (PLP) could provide useful information for classification. Moreover, spectrogram analysis employed in image processing would be another option to extract more informative features.

ii. Though deep learning algorithms are time and memory expensive and need tunning many hyperparameters, they have proven to be promising alternatives for classifying data with complicated patterns and few abnormal data. CNN or variational autoencoders algorithms have successfully learned latent relations in datasets with the least dependence on preprocessing steps.

iii. Due to the subsampling methodology in the feature extraction step, many elements of the training and test matrixes had zero values. For example, more than 70 percent of the looped network's imbalanced training data had zero values. This prevalence of zero values would make the data matrixes sparse, making convergence problems for algorithm optimization. The sparsity could be a reason for the long training time for algorithms, particularly those using nonlinear kernels. Using less sparse training and test datasets and algorithms with more robustness to sparsity could address the shallow classifiers' poor performance in predicting leak samples.

4.4.3. Leak Detection Using Dynamic Pressure Sensor and Accelerometer Data

One of the advantages of this study's testbed is applying three types of sensors that captured the system's changes. In this dissertation, only acoustic data were analyzed for leak detection. Section 1 showed the success of other leak detection research that merely employed pressure and acceleration data. Therefore, stand-alone analyses of the dynamic pressure and acceleration recordings could provide helpful information. These measurements could still be used as complementary to fill the gaps in leak detection via acoustic signals.

4.4.4. Leak Localization

A recent active research field is localizing leak location using in-network sensors with temporal and spatial algorithms. This objective would help water utilities to spot background leaks and fix them before they develop into bursts. Three factors make this testbed suitable for leak localization: (1) sensor varieties; (2) distributed locations of the sensors; (3) possibility of changing network architectures using flanged connections. These characteristics make it easy to create different leaks, induce arrival time delays, and capture signals at various locations and differential time.

4.4.5. Sensor Performance in Leak Detection

Due to the availability of three types of data, one can evaluate which sensor may be more successful in leak detection. For acoustic, dynamic pressure, and acceleration data, the time and frequency-domain features in section 2 can be extracted and input to a tree-based classifier such as XGBoost or random forest with predefined tunning parameters. The F1-score can measure the classifiers' performance for leak data.