Proceedings of the 17th Pan-American Conference on Soil Mechanics and Geotechnical Engineering (XVII PCSMGE), and 2nd Latin-American Regional Conference of the International Association for Engineering Geology and the Environment (IAEG), La Serena Chile, 2024.

Data-Driven Leak Detection in Water Distribution Networks

Detección de fugas basada en datos en redes de distribución de agua

Thomas Green & **Giovanni Cascante**
Department of Civil and Environmental Engineering, University of Waterloo, Canada, gcascante@uwaterloo.ca

Stanley Fong
Digital Water Solutions Inc., Canada

ABSTRACT: Water loss in water distribution networks (WDNs) is a multi-billion-dollar global issue. WDNs often develop leaks and breaks over time due to factors such as aging infrastructure, pressure transients, and operational changes. Leaks in WDNs are a critical area of concern for utilities as they contribute to water loss, increase the risk of waterborne pathogens, and pipe breakage. Additionally, they are challenging to find and repair given the subterranean nature of most WDNs. To reduce the risks and impacts of leakage, it is necessary for utilities to constantly monitor, detect, localize, and repair leaks in a timely manner. Due to the scale and complexity of municipal WDNs coupled with resource constraints at the operational level, effective leak detection approaches need to be robust, efficient, and easily deployable across networks. Currently, most leak detection methods have common deficiencies with respect to both sensing and analytical approaches that significantly limit their effectiveness in practice: vibration-based methods are not conducive in PVC networks; and the majority of analytical approaches are contingent on the availability of accurate, auxiliary information – such as network layout and pipe specifications – for detection and localization. To address these gaps, this paper focuses on tandem development across two aspects: improving the empirical understanding around hydrophones for leak detection; and further development of autonomous (i.e., self-sufficient) algorithms for real-time leak detection using machine learning and signal processing.

KEYWORDS: Leak Detection, Water Distribution Networks, Machine Learning, Signal Processing, Hydrophones.

1 INTRODUCTION

Over time, WDNs deteriorate due to factors that depend on pipe material, including internal and external corrosion, improper installation, imperfections, and leaching (of Canadian Municipalities & Research Council, 2003). Pipe deterioration eventually leads to breaks (e.g., circumferential breaks, longitudinal breaks, etc.). Water then escapes these breaks, creating leaks.

Leaks in WDNs are problematic for several reasons. They can act as entry points for pathogenic microorganisms (e.g., E. coli.), which can subsequently be ingested by consumers (Besner et al., 2011). Further, leaks cause financial losses on the order of billions of dollars each year (Frauendorfer & Liemberger, 2010; Moslehi et al., 2021).

Hydrophones are like microphones except they are designed to be used underwater. These transducers are often composed of piezo-electric material, which turn acoustic signals into electrical potential differences. Because hydrophones make measurements inside of the water, this allows them to detect planar waves, which dominate at low frequencies, and can travel hundreds of meters and still be detectable (Kafle, n.d.; Kharat et al., n.d.; Muggleton & Brennan, 2004a). As such, hydrophones can detect acoustic emissions produced by far-away leaks that may not be detectable by accelerometers (Martini et al., 2017). For these reasons, hydrophones will be the main method used to detect leaks in this paper.

In recent years, accelerometers and hydrophones have been combined with wireless sensor networks, acting as nodes that collect and transmit data to central servers; these cloud computing environments allow WDNs to be monitored 24/7. Because the number of nodes in these networks can be very large, it is important that leak detection algorithms do not consume too many spatial and computational resources.

Leak detection can be formalized as a classification problem:

$$\{(\boldsymbol{t_i}, y_i)\}_{i=1}^{s}, where\ \boldsymbol{t_i} \in \mathbb{R}^d \text{ A } y_i \in \{no\ leak, leak\} \quad (1)$$

Where $\boldsymbol{t}_i$ represents a sample (e.g., a waveform collected using a hydrophone) of size d, y_i represents the corresponding class label, and s represents the size of a dataset.

In this light, many leak detection algorithms in the literature can be described by figure 1. Signal collection involves using real-time leak detection methods (such as hydrophones or accelerometer) to collect data; representation extraction entails various pre-processing methods, such as filtering and domain transformations; feature extraction pertains to the distillation of information; and classification/detection relates to the utilization of models like standing vector machines to predict whether a leak is present.

For instance, Liu Y et al. proposed a spectral-based technique where they use empirical mode decomposition to decompose signals into a finite number of intrinsic mode functions, from which the discrete Fourier transform is applied; this can be considered as a representation extraction technique. They then use approximate entropy and PCA to extract time and frequency domain features that are then used by a standing vector machine to classify whether a signal contains leakage sound. Many other pipelines like these exist in literature (Fan et al., 2022).

Correlation-based techniques (which can also be described using figure 1) need prior knowledge about the wave speed in the

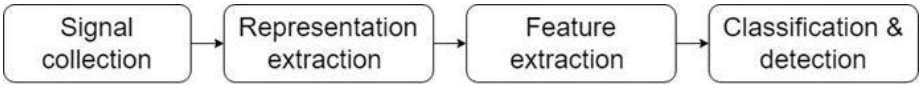

Figure 1. Flow chart describing a typical leak detection algorithm.

network in order to be effective. Wave speed depends on parameters such as those pertaining to foundation soil (e.g., density) (Muggleton & Brennan, 2004b). However, this information is often unknow, which necessitates the need for methods that do not require such.

This paper aims to complete two goals, whereby the second goal can be further broken down into two sub-goals:

1. Improve the empirical understanding of hydrophones for leak detection.
2. Develop the kernel of an autonomous algorithm for real-time leak detection using machine learning and signal processing, which is not contingent on prior-knowledge of foundation soil parameters.
    a. Compare the performance of a broad range of representation extraction and classification techniques, which consider the effects of real world WDN.
    b. Characterize a data pipeline that can be used within a larger framework that can be used to autonomously detect leaks in a cloud computing environment.

The rest of the paper is structured as follows. Section 2 provides a background on the elements of the proposed data pipeline, representation extraction, feature extraction, and data cleaning. Section 3 outlines the methodology and experimental setup that was used. Section 4 presents the results and some recommendations for future work, and section 5 concludes this paper.

2 BACKGROUND

A well-performing leak detection framework requires the use of well-studied data processing methods. This section provides an overview of such. Table 1 in the appendix describes the mathematical notation used in this section.

2.1 Representation Extraction

A signal is said to be nonstationary if its intrinsic properties vary with time (Simon Haykin, Barry Van Veen - Signals and Systems (2002, Wiley)). Hydrophone signals collected in WDNs are highly nonstationary due to processes occurring in such, and it is important that the representation extraction techniques take this behavior into consideration.

2.1.1 Short-Time Fourier Transform

Given a discrete time domain signal of length N, which can be represented by vector $\boldsymbol{t}$, the short-time Fourier transform (STFT) computes a two-dimensional time-frequency representation $\boldsymbol{X}$ (Bruce Sharpe, n.d.; Simon Haykin, Barry Van Veen - Signals and Systems (2002, Wiley)). For the remainder of this paper, when referring to a time-frequency transform, we assume it is discrete. A practical implementation of such begins by splitting up sequence $\boldsymbol{t}$ into N windows, and multiplying each by the complex conjugate of a windowing function $\boldsymbol{w}^*$ to produce a set of N window frames $\{\boldsymbol{w}_n, n = 0,1,\dots,N-1\}$, where:

$$\boldsymbol{w}_n = \boldsymbol{w}^* \odot \boldsymbol{t}_{nH:nH+L} \tag{2}$$

Here, H denotes hop size, L denotes window length, and $\odot$ represents the Hadamard product. Let matrix $\boldsymbol{W}$ represents the horizontal concatenation of all window frames:

$$\boldsymbol{W} = [\boldsymbol{w_1}, \boldsymbol{w_2}, \dots, \boldsymbol{w_{N-1}}] \tag{3}$$

The STFT $\boldsymbol{X}$ is then found by simply applying the discrete Fourier transform (DFT) to each windowed frame:

$$X_{m,n} = \sum_{k=0}^{L-1} W_{k,n}\, e^{\frac{-j2\pi mk}{L}} \quad m = 0,1,\dots,L-1 \tag{4}$$

Where $X_{m,n}$ represents the m[th] DFT coefficient of window frame n, and j represents an imaginary number. Assuming $\boldsymbol{W}$ is real, column $\boldsymbol{X}_{:,n}$ is symmetric about $m \approx L/2$; for the remainder of this paper, we assume the second half (negative frequency terms) of the DFT is omitted for each $\boldsymbol{X}_{:,n}$, $n = 0,1,\dots,N-1$. Further, vector $\hat{\boldsymbol{f}}$ (of length $\sim L/2$, excluding negative frequencies) represents the frequencies across the columns of matrix $\boldsymbol{X}$, and ranges from 0 to $\sim f_s/2$, where f_s represents the sampling frequency (hence, $\Delta f = \hat{f}_m - \hat{f}_{m-1} = f_s/L$ represents frequency resolution). Subsequently, increasing L increases frequency resolution, but decreases the temporal resolution. Of course, by Nyquist theorem, the highest frequency that can reliably be detected is $f_s/2$. In practice, the Fast-Fourier transform algorithm is used to compute the above.

Finding the modulus of each element in the STFT (i.e., $(\boldsymbol{X} \odot \bar{\boldsymbol{X}})^{\circ 1/2}$, where $\bar{\boldsymbol{X}}$ represents the element-wise complex conjugation of $\boldsymbol{X}$, and $\circ\, 1/2$ denotes the Hadamard power of degree $1/2$) produces a magnitude spectrogram $\boldsymbol{S}$. A power spectrogram is found by squaring each element of the magnitude spectrogram (i.e., $\boldsymbol{S}^{\circ 2}$). Finally, a dB spectrogram will refer to an array derived from computing $10\ log_{10}(S_{m,n}{}^2/r)$ for each element in the magnitude spectrogram $\boldsymbol{S}$, where r represents a reference (e.g., the maximum value in $\boldsymbol{S}^{\circ 2}$). Here we note that a dB spectrogram is often preferable for visualization.

2.1.2 Continuous Wavelet Transform

The continuous wavelet transform (CWT) overcomes some of the pitfalls of the STFT, as it has the ability to vary the temporal and frequency resolution (Aguiar-Conraria & Soares, 2011). The CWT $W_{x;\psi}(r,s)$ of time series $x(t)$ is computed as follows:

$$W_{x;\psi}(r,s) = \int_{-\infty}^{\infty} x(t)\psi_{c,s}{}^*(t)\, dt \tag{5}$$

Where ψ_{c,s^*} represents the complex conjugate of a daughter wavelet, which is translated in time by r and stretched by a factor of s. Daughter wavelets are the parameterizations of a mother wavelet, as defined by:

$$\psi_{c,s}(t) = \frac{1}{\sqrt{|s|}}\psi\left(\frac{t-c}{s}\right) \tag{6}$$

Similar to the STFT, taking the absolute value of a CWT produces a magnitude scalogram; taking the Hadamard power of degree 2 of such produces a power scalogram, and the application of the log can be used to help derive a dB scalogram.

2.2 Data Cleaning

In the context of classification, data cleaning is important, as it can be used to remove unwanted signal components (noise) – which are not representative of the underlying process – from the signal (Simon Haykin, Barry Van Veen -Signals and Systems (2002, Wiley)). Noise can be detrimental and increase the likelihood that a model overfits training data (Ying, 2019). Further, unwanted signal components can make it harder to visualize the data. The two main data cleaning mechanisms that will be explored in this paper are high-pass (HP) filtering and outlier removal.

2.3 Feature Extraction

A piece of information within a data representation is known as a feature (Goodfellow et al., n.d.). Features may constitute the representation itself, a subset of such, or, may be computed from such. In the context of leak detection, many features have been suggested to work well, both for time domain and frequency domain representations. In this paper, we study the following: energy, and spectral features including centroid, variance, skewness, and kurtosis. These features have been found to offer a rich and interpretable characterization of signals in a low-dimensional space (Li et al., 2018).

The computations of these features have been expressed in terms of linear algebra, such that they can easily be computed efficiently using libraries like NumPy, allowing for leak detection algorithms that use these features to be more scalable. Matrix $\boldsymbol{R}$ can take on the form of any representation type found in table 2 of the appendix (where each representation type has an associated id). The combination of signal type (raw or HP filtered), frequency-time transformation (CWT or STFT), and non-linear function (magnitude, power, or dB) constitute a representation $\boldsymbol{R}$. For instance, $\boldsymbol{R}$ could take on the form of representation 5 (raw power spectrogram), representation 12 (HP filtered dB scalogram), etc.

Energy is a common feature used in leak detection, as acoustic emissions produced by leaks produce sound. By Parseval's theorem, the energy of a finite signal can be computed either in the time domain or frequency domain. Hence, one can efficiently compute this feature across windows in $\boldsymbol{R}$ using the following (excluding negative frequencies) (Proakis & Manolakis, n.d.):

$$\boldsymbol{e} = \boldsymbol{j}^{\mathsf{T}}\boldsymbol{R}^{\circ 2} \tag{7}$$

Where $\boldsymbol{j}^{\mathsf{T}}$ represents a row vector of 1's, and $\boldsymbol{e}$ is a row vector containing the energy values of each window. As a reminder, each row in $\boldsymbol{R}$ varies in frequency, while each column in such varies temporally.

Spectral centroid is the barycenter of a spectrum. In essence, it is the expectation of frequency (as weighted by the amplitudes of each frequency), normalized by the sum of amplitudes (McAdams, 2019; Peeters, 2004; Siegrist, 2021). The values for such can be computed using:

$$\boldsymbol{c} = (\boldsymbol{j}^{\mathsf{T}}\boldsymbol{F} \odot \boldsymbol{R}) \oslash (\boldsymbol{j}^{\mathsf{T}}\boldsymbol{R}) \tag{8}$$

Where row vector $\boldsymbol{c}$ contains the centroid values across each window, array $\boldsymbol{F}$ has the same dimensionality as $\boldsymbol{R}$ and is the horizontal concatenation of column vectors $\hat{\boldsymbol{f}}$ and $\oslash$ represents the Hadamard division. Here, we note that in the case of a scalogram, $\hat{\boldsymbol{f}}$ can be found by computing the equivalent frequency of each scale that was used.

Spectral variance is the second moment of a spectrum about its mean. Accordingly, it measures the spread of a spectrum about its mean. The values of such across each window can be computed as follows:

$$\boldsymbol{v} = \boldsymbol{m_2} - \boldsymbol{c}^{\circ 2} \tag{9}$$

Where row vector $\boldsymbol{m_2}$ represents the horizontal concatenation of the 2nd central moment of each column (spectrum) in $\boldsymbol{R}$. In general, the k^{th} central moment across each window can be computed as follows:

$$\boldsymbol{m_k} = (\boldsymbol{j}^{\mathsf{T}}\boldsymbol{F}^{\circ k} \odot \boldsymbol{R}) \oslash (\boldsymbol{j}^{\mathsf{T}}\boldsymbol{R}) \tag{10}$$

As can be seen, $\boldsymbol{c} = \boldsymbol{m_1}$

Spectral skewness measures the asymmetry of a distribution about its mean, and the values of such can be computed as follows:

$$\boldsymbol{s} = (\boldsymbol{m_3} - 3\boldsymbol{c} \odot \boldsymbol{v} - \boldsymbol{c}^{\circ 3}) \oslash (\boldsymbol{v}^{\circ \frac{3}{2}}) \tag{11}$$

Finally, spectral kurtosis measures the flatness of a distribution around its mean value, and the values of such can be computed using:

$$\boldsymbol{k} = (\boldsymbol{m_4} - 4\boldsymbol{c} \odot \boldsymbol{m_3} + 6\boldsymbol{m_2} \odot \boldsymbol{c}^{\circ 2} - 3\boldsymbol{c}) \oslash (\boldsymbol{v}^{\circ 2}) \tag{12}$$

3 METHODOLOGY

The dataset used in this paper was collected by Digital Water Solution's hydrant.ai devices in a live WDN. In the corresponding experiment, the hydrophones in the hydrant.ai devices were calibrated to collect five second files every three minutes, at a sampling rate of 1994 Hz. For simplicity, we refer to these devices as "hydrants". To simulate leaks, two flow tests were conducted at a nearby fire hydrant. Figure 2 shows the hydrant and leak locations:

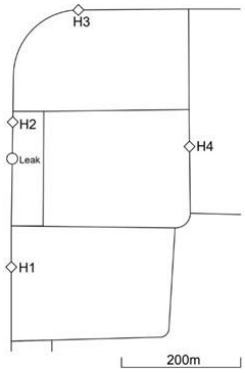

Figure 2. Map showing locations of hydrants (diamonds) and simulated leaks (circle).

Because the study took place during the daytime (between 11:00am and 2:27pm local time), these were far from ideal conditions for leak detection, as there was a lot of background noise from the live system. Both flow tests lasted about 10 minutes; the first flow test had a smaller volumetric flow rate (labeled as light leak) while the second one had a larger one (labeled as hard leak). Data collected outside of these flow tests shall be referred to as baseline data.

Figure 3 presents the experimental pipeline, which shows the various steps of representation extraction and feature extraction conducted for each hydrant and representation type. The experimental pipeline is the abstraction of what can be used as the kernel of an autonomous algorithm. Further, it was applied to each time-frequency transformation (i.e., STFT and CWT) to produce all 12 representations in table 2.

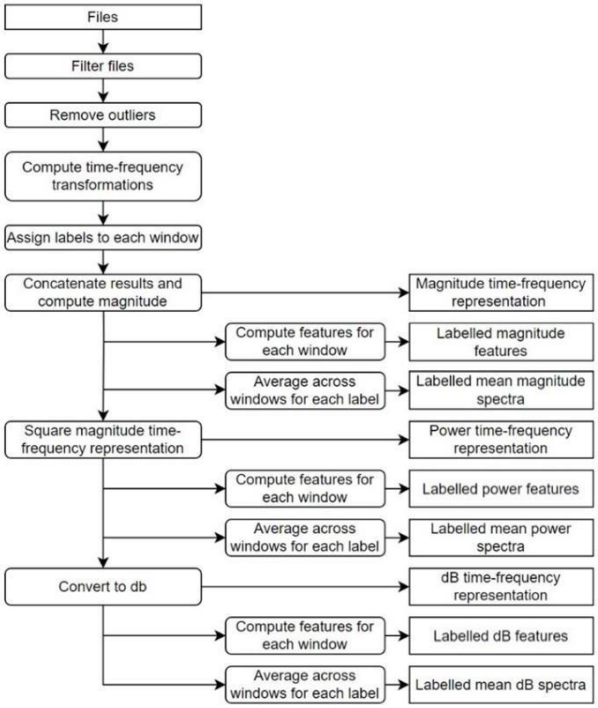

Figure 3. Flow chart describing the experimental setup.

For each hydrophone file, a sixth-order Butterworth HP filter with a critical frequency of 50 Hz was applied, thus creating two different signals, including the unfiltered (raw) signal.

The STFT was implemented using librosa (McFee, n.d.); a window length of 2 seconds (3988 samples), hop length of 0.25 seconds (498 samples), and flattop windowing function were used, and the first and last two windows of each file were discarded. Subsequently, each file produced 17 windows. The CWT was implemented using ssqueeze (John Muradeli, 2020); a Morlet mother wavelet with a mu parameter of 5 was used, which closely resembles the Gause wavelet (Ahadi & Bakhtiar, 2010).

To alleviate the memory complexity introduced by the large temporal resolution of the CWT, only one out of 480 temporal slices were kept; it is assumed that the leak itself is stationary and as such, discarding such is deemed acceptable. Like the STFT, this produced 17 windows (time slices) per file.

For each representation type, the features outlined in section 2.3 were computed across each window. Following this, the energy across each window was used to detect and remove outliers using a dynamic peak detection algorithm that is robust to non-stationarity (induced by a potential leak) (Brakel, 2014); the parameters used were a lag of 100, threshold of 2, and influence of 0.2. Each window was then labelled in accordance with the timestamps recorded during the flow tests.

The mean spectra for each representation across the different labels (like those found in figure 3) were computed using $\boldsymbol{R_l j / N}$, where $\boldsymbol{R_l}$ contains the N columns (windows) from $\boldsymbol{R}$ associated with a given label l.

Here, we note that the choice of a critical frequency of 50 Hz for the HP filter becomes clear when studying figure 3, which shows the area-normalized mean magnitude DFT across events for hydrant 2 (located near the flow tests). One can see that the baseline spectrum does not diverge substantially from the leak spectra below 50 Hz, while they do beyond 50 Hz.

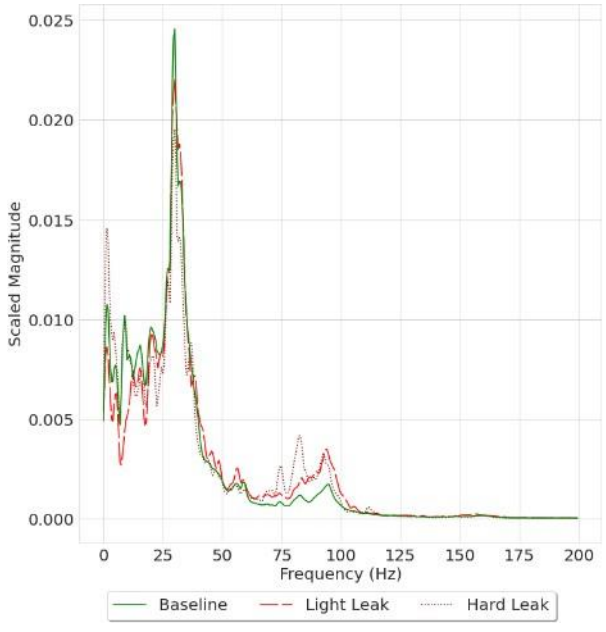

Figure 4. Mean raw STFT spectrum across events for hydrant 2.

Finally, as per figure 5, a separate dataset was created for each hydrant/representation type pair. Each dataset was further separated into a stratified 1:1 train test split. Labels were simply given a value of 0, 1, or 2 for baseline, light leak, and hard leak, respectively; these integer encodings were assigned one-hot encodings for models that required such. The scaled features were used as model inputs. For a given hydrant, there were about 300, 55, and 40 datapoints for baseline, light leak, and hard leak cases respectively. Eleven different models were then trained on each dataset's train split, before predicting the labels for the corresponding test split. The average F1 measure of each model/dataset pair was used to infer which hydrants could hear the leak and evaluate the best representation/model pair; the equation for such can be found in the appendix (Taha & Hanbury, 2015).

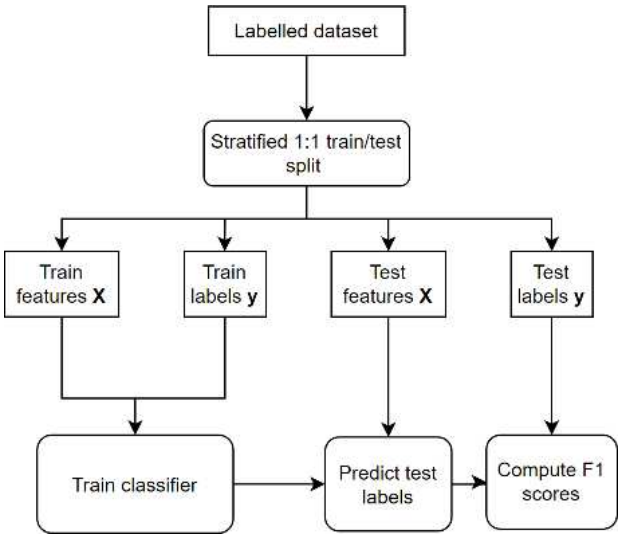

Figure 5. Flow chart describing the classification experimental setup, conducted on each combination of hydrant, representation type, and classifier.

Because each model was trained on a hydrant-to-hydrant bas is, there was no need to provide the models with informatio n about the surrounding foundation soil. Unlike correlation-b ased methods, the training process allows models to learn m anifolds that distinguish signals that contain leak signatures a nd non-leak signatures, where soil-parameters may vary there in.

4 RESULTS

Figures 6 presents outliers and inliers, as computed by the dynamic peak detection algorithm using the energy values derived from the raw magnitude spectrogram. Figure 7 presents the raw dB spectrogram and the scaled energy line plot derived from the raw magnitude spectrogram for hydrant 2. Here, we see that the leaks are not obvious when looking at the energy line plot, though, there are some spectral components just below 128Hz that are correlated with the acoustic emissions produced by the leaks.

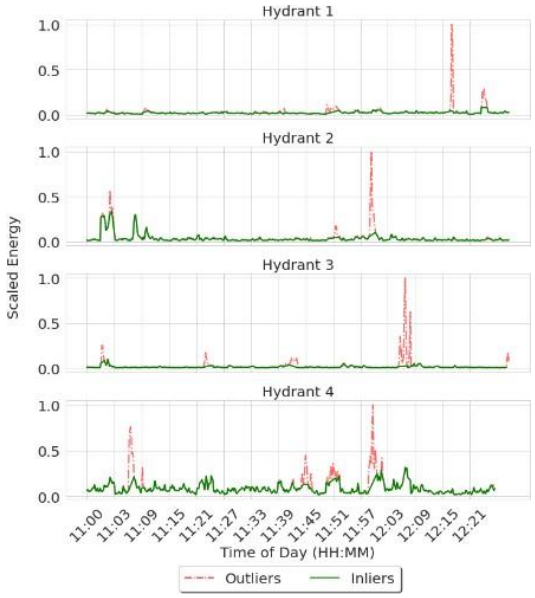

Figure 6. Outlier and inlier windows across hydrants.

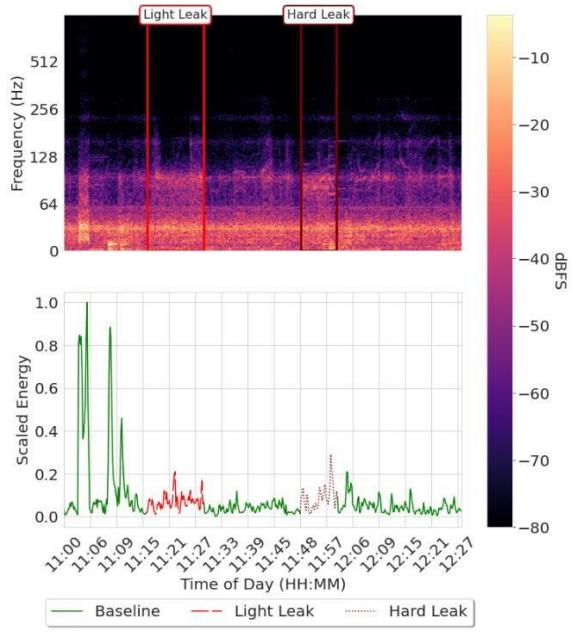

Figure 7. Raw dB Spectrogram and corresponding energy line plot.

Indeed, the presence of the leaks become much clearer w ithin the features when we turn to representations derived fr om HP filtering. Figure 8 presents the HP filtered dB scalo gram and corresponding energy line plot.

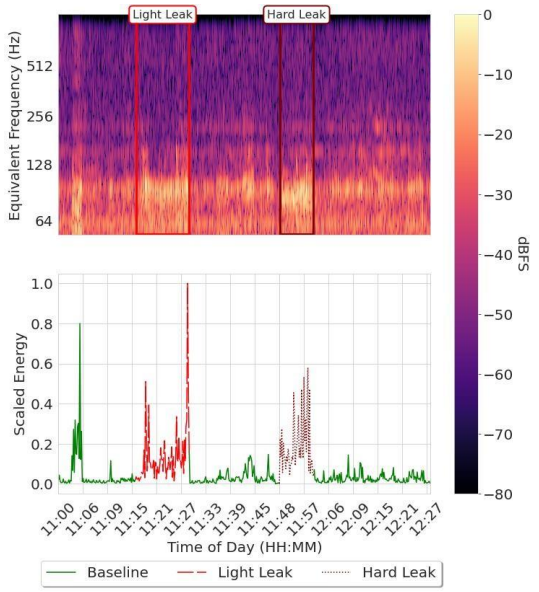

Figure 8. HP filtered dB scalogram and corresponding energy line.

As can be seen, energy beyond 50 Hz is substantially higher during both light and hard leaks. Further, the bump in energy that occurred around 11:03 is spectrally different from the leak signals; this is reflected by the other spectral features.

Figure 9 presents the kernel density estimate (KDE) plots of centroid, variance, skewness, and kurtosis values across windows, as derived from the HP filtered magnitude spectrogram for hydrant 2:

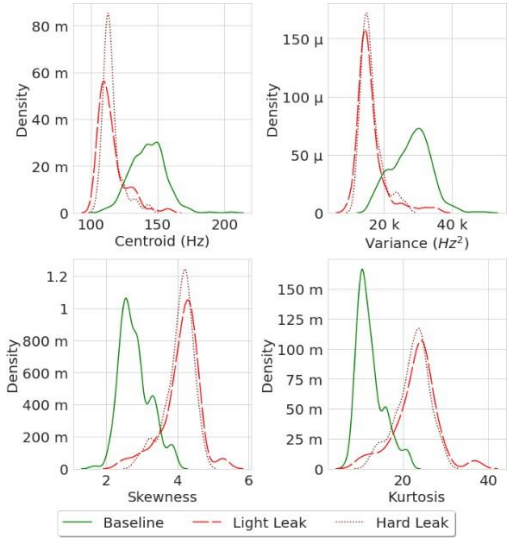

Figure 9 KDE plots of HP filtered magnitude spectrogram features across events for hydrant 2.

Here, we see that the centroid values of the HP filtered signals decrease, while the skewness values increase. This makes sense when looking at the spectrogram in figure 7, as a lot of the energy lies around 110 Hz during the leaks, which is close to the critical frequency of the HP filter. Further, the filtered leak signal is a narrower band signal than the filtered background noise, which decreases the variance values, and generally the spectra become peakier, which increases the kurtosis values.

Figure 10 shows a summary of what is seen across all 4 hydrants. Here, the values across the y-axis were derived from scaling the box-cox transform of principle component 0 across the HP filtered spectrogram features (excluding energy). Similarly, the values across the x axis were derived from the box-cox transform of scaled energy. Here, we can see that clusters form for each label across hydrants 1 and 2, while this is not the case for hydrants 3 and 4. This means that hydrants 1 and 2 are able to hear the leaks, while it is harder for hydrants 3 and 4 to do so.

These results are consistent with the classification results. Figure 11 presents the model results for the two-class problem of differentiating between baseline and leak (i.e., light leak or hard leak), while figure 12 presents that of the three-class problem of differentiating between baseline, light leak, and hard leak data.

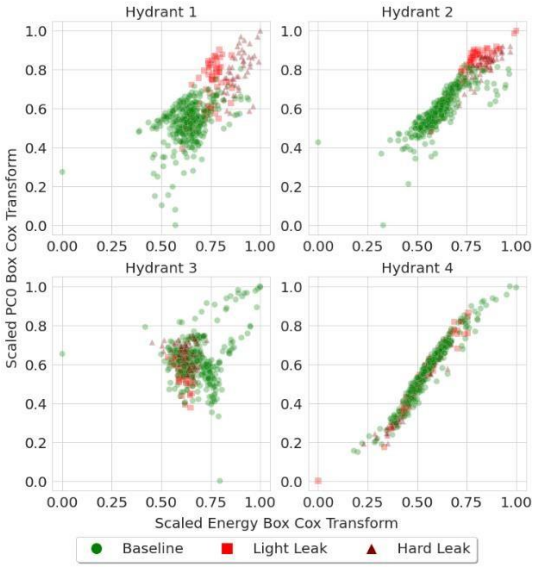

Figure 10 Scatter plots of transformed features across hydrants and events.

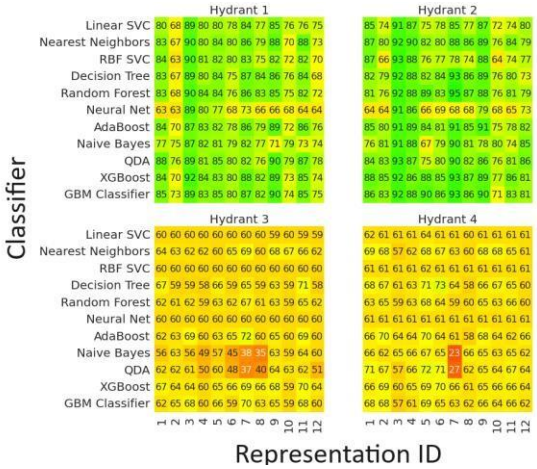

Figure 11 Heatmaps describing F1 scores across hydrant, classifier, and representation type combinations for the two-class problem.

As can be seen, the classifiers generally perform best on hydrant 2, followed by 1, 4, and 3 respectively. This is consistent with the distance of the leak from each hydrant. Further, filtering (associated with representations 3, 4, 7, 8, 11, 12) seems to enhance model results for hydrants that are closer to the leak (i.e., hydrants 1 and 2), while raw representations are better for training models on hydrants that are far away from the leaks (i.e., hydrants 3 and 4). This makes sense, as the high frequencies produced by the leak would have attenuated by the time the AEs reached far-away hydrants, where low-frequency information would be all that is left (Muggleton & Brennan, 2004a). Further, the CWT (associated with even numbered representations IDs) does not seem to offer any advantages over the STFT, assuming stationarity. Finally, the choice of the classifier seems to be less important and seems to largely depend on the representation being used.

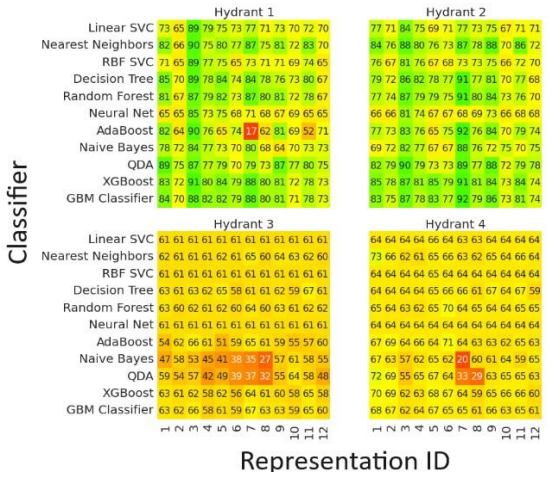

Figure 12 Heatmaps describing F1 scores across hydrant, classifier, and representation type combinations for the three-class problem.

Given these results, more research could be done to find the optimal filter design as a function of desired leak detection distance. In addition, the generalization performance of each classifier with respect to new data (from leaks outside of the train dataset) is unknown and is a key area that should be researched. Further work can also investigate the effect of including foundation soil parameters as part of input features on generalization performance; such information can be derived from measurements, or calibrated numerical models.

5 CONCLUSION

With careful consideration of properties of AE in WDNs, one can create effective leak detection algorithms by following the framework proposed in this paper. Features derived from the STFT offer a rich characterization of the spectral composition of hydrophone signals. Further, filtering can help make the variations induced by leaks within the features more distinguishable. Outlier removal that is robust to non-stationarity can help mitigate the negative effects of noisy processes in WDNs. Finally, it has been shown that many different models and features derived from various representations can achieve modest results under far from ideal conditions.

6 ACKNOWLEDGEMENTS

I would like to sincerely thank Dr. Stan Fong for being an amazing mentor, and Dr. Giovanni Cascante for his vigilance and thoughtful academic guidance.

7 REFERENCES

Aguiar-Conraria, L., & Soares, M. J. (2011). *The Continuous Wavelet Transform: A Primer*. http://www.eeg.uminho.pt/economia/nipe

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

Besner, M. C., Prévost, M., & Regli, S. (2011). Assessing the public health risk of microbial intrusion events in distribution systems: Conceptual model, available data, and challenges. In *Water Research* (Vol. 45, Issue 3, pp. 961–979). Elsevier Ltd. https://doi.org/10.1016/j.watres.2010.10.035

Brakel, J. P. G. van. (2014). *Robust peak detection algorithm using z-scores*.

Bruce Sharpe. (n.d.). *Invertibility of overlap-add processing*.

Fan, H., Tariq, S., & Zayed, T. (2022). Acoustic leak detection approaches for water pipelines. In *Automation in Construction* (Vol. 138). Elsevier B.V. https://doi.org/10.1016/j.autcon.2022.104226

Frauendorfer, R. (Of A. D. B., & Liemberger, R. (2010). *The issues and challenges of reducing non-revenue water*.

Goodfellow, I., Bengio, Y., & Courville, A. (n.d.). *Deep learning*.

John Muradeli. (2020). ssqueezepy. *GitHub. Note: Https://Github.Com/OverLordGoldDragon/Ssqueezepy/*.

Kafle, M. D. (n.d.). *Active Acoustics for Leak Detection in Water Distribution Networks*.

Kharat, D. K., Mitra, S., Akhtar, S., & Kumar, V. (n.d.). *Polymeric Piezoelectric Transducers for Hydrophone Applications*.

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: Journal of the International Measurement Confederation*, *115*, 39–44. https://doi.org/10.1016/j.measurement.2017.10.021

Martini, A., Troncossi, M., & Rivola, A. (2017). *Vibroacoustic Measurements for Detecting Water Leaks in Buried Small-Diameter Plastic Pipes*. https://doi.org/10.1061/(ASCE)

McAdams, S. (2019). *The Perceptual Representation of Timbre* (pp. 23–57). https://doi.org/10.1007/978-3-030-14832-4_2

McFee, B. C. R. D. L. D. P. E. M. M. E. B. and O. N. (n.d.). librosa: Audio and music signal analysis in python. *In Proceedings of the 14th Python in Science Conference, Pp. 18-25. 2015.*

Moslehi, I., Jalili-Ghazizadeh, M., & Yousefi-Khoshqalb, E. (2021). Developing a framework for leakage target setting in water distribution networks from an economic perspective.

Structure and Infrastructure Engineering, *17*(6), 821–837. https://doi.org/10.1080/15732479.2020.1777568

Muggleton, J. M., & Brennan, M. J. (2004a). Leak noise propagation and attenuation in submerged plastic water pipes. *Journal of Sound and Vibration*, *278*(3), 527–537. https://doi.org/10.1016/j.jsv.2003.10.052

Muggleton, J. M., & Brennan, M. J. (2004b). Leak noise propagation and attenuation in submerged plastic water pipes. *Journal of Sound and Vibration*, *278*(3), 527–537. https://doi.org/10.1016/j.jsv.2003.10.052

of Canadian Municipalities, F., & Research Council, N. (2003). *DETERIORATION AND INSPECTION OF WATER DISTRIBUTION SYSTEMS A BEST PRACTICE BY THE NATIONAL GUIDE TO SUSTAINABLE MUNICIPAL INFRASTRUCTURE*.

Peeters, G. (2004). *A large set of audio features for sound description (similarity and classification) in the CUIDADO project*.

Proakis, J. G., & Manolakis, D. G. (n.d.). *Digital Signal Processing Principles, ~ l ~ o r i t h m i , and Applications Third Edition*.

Siegrist, K. (2021, April). *Probability, Mathematical Statistics, and Stochastic Processes*. LibreTexts Statistics. https://stats.libretexts.org/Bookshelves/Probability_Theory/Probability_Mathematical_Statistics_and_Stochastic_Processes_(Siegrist)/04%3A_Expected_Value/4.04%3A_Skewness_and_Kurtosis#:~:text=The%20third%20moment%20measures%20skewness,power%20of%20the%20standard%20deviation.

Simon Haykin, Barry Van Veen - Signals and Systems (2002, Wiley) - libgen.lc (1). (n.d.).

Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Medical Imaging*, *15*(1). https://doi.org/10.1186/s12880-015-0068-x

Ying, X. (2019). An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series*, *1168*(2). https://doi.org/10.1088/1742-6596/1168/2/022022

8 APPENDIX

8.1 Notation

Table 1. Notations used in this paper and their descriptions.

Parameter	*Description*
x	Scalar
$x(t)$	Continuous Signal
$\boldsymbol{x}$	Vector
$\boldsymbol{x}_k$	Vector
x_m	Element m of vector $\boldsymbol{x}$
$\boldsymbol{X}$	Matrix
$X_{m,n}$	Element m, n of matrix $\boldsymbol{X}$
$\boldsymbol{X}_{m,:}$	Row m of matrix $\boldsymbol{X}$
$\boldsymbol{X}_{:,n}$	Column n of matrix $\boldsymbol{X}$
$\boldsymbol{x}_{i:j}$	Vector formed by indexing vector $\boldsymbol{x}$ with indexes i through j-1 (inclusive)

Table 2. Representation IDs and their corresponding descriptions.

Representation ID	*Description*
1	Raw Magnitude Spectrogram
2	Raw Magnitude Scalogram
3	HP Filtered Magnitude Spectrogram
4	HP Filtered Magnitude Scalogram
5	Raw Power Spectrogram
6	Raw Power Scalogram
7	HP Filtered Power Spectrogram
8	HP Filtered Power Scalogram
9	Raw dB Spectrogram
10	Raw dB Scalogram
11	HP Filtered dB Spectrogram
12	HP Filtered dB Scalogram

8.2 F1 Score

$$F1 = \frac{2*TP}{2*TP+FP+FN} \tag{13}$$

Where TP represents true positives, FP represents false positives, and FN represents false negatives.