

APPLICATION OF LEAST SQUARES SUPPORT VECTOR MACHINES AND ACOUSTIC MEASUREMENTS FOR LEAK FLOW RATE PREDICTION

Joseph D. Butterfield^{1*}, Viviana Meruane², Richard P. Collins³ Gregory Meyers⁴, and Stephen B.M. Beck¹

¹Department of Mechanical Engineering, University of Sheffield, Sheffield, South Yorkshire, S13 JD, UK.

²Department of Mechanical Engineering, Universidad de Chile, Beaucheff 851, Santiago, Chile.

³Department of Civil and Structural Engineering, University of Sheffield, Sheffield, South Yorkshire, S13 JD, UK.

⁴College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter, EX4 4QF, UK.

* Corresponding author: jbutterfield1@sheffield.ac.uk

ABSTRACT

Leakage from water distribution pipes represent a huge issue worldwide with economic. Leaks are normally found by placing sensors either side of the leak recording the leaks acoustic emission as it discharges the leak hole. As the leak noise is intrinsic to the leak, it contains information which could provide information about the leak, including the leak flow rate. Any tool which can accurately determine the leak's flow rate could be used by water industry practitioners in order to prioritise leakage repair activities by repairing the higher leak flow rates first. This will result in economic savings through reduced water lost and better allocation of company resources. This paper demonstrates a small element of research undertaken at the University of Sheffield in collaboration with several UK water companies. The aim of the research is to develop a tool in order to predict leak flow rate using acoustic emission sensors. The research uses Least Squares-Support Vector Machines in order to predict leak flow rate in MDPE pipe using high quality data from a unique experimental pipe rig. The results demonstrate that there is sufficient information within the leak's acoustic emission signal in order to predict leak flow rate. Therefore the research represented in this paper presents a tool which can be used by water industry practioners to prioritise leak repair.

Keywords: Leakage, Pipeline, Acoustic Emission.

1 BACKGROUND

Leakage from Water Distribution Systems (WDS) represent a huge problem worldwide due to the economic and environmental effects [1]. Leakage levels in developing countries can be as high as 50% [2], whereas in newer, well maintained WDS leakage levels can be as low as 5%.

A number of methods exist in order to reduce leakage levels. One of the most common methods is that of leak noise correlation. As water discharges the leak hole it produces an acoustic emission. Accelerometers or Hydrophones can then be placed either side of the leak at some distance away and the signal analysed in order to find the location of the leak (Figure 1). The leak is found using the following equation:

$$L_1 = \frac{d - c\tau_{delay}}{2} \tag{1}$$

where d is the distance between the two accelerometers or hydrophones and c is the wavespeed of the leak signal. τ_{delay} is the lag in arrival time between the two acoustic emission sensors, calculated from the peak in the cross correlation function:

$$R_{x_1, x_2} = E[x_1(t)x_2(t + \tau_{delay})], \tag{2}$$

where $E[\cdot]$ is the expectation operator. Leak noise correlation has been shown to be highly successful on metallic pipes but has been relatively unsuccessful on plastic pipe due to the rapid attenuation of the leak signal [2]. Moreover, it has also been shown by a number of authors that a number of variables influence the leak signal, including pipe material [3], [4], backfill [5], [6], leak shape [4] and leak flow rate [7].

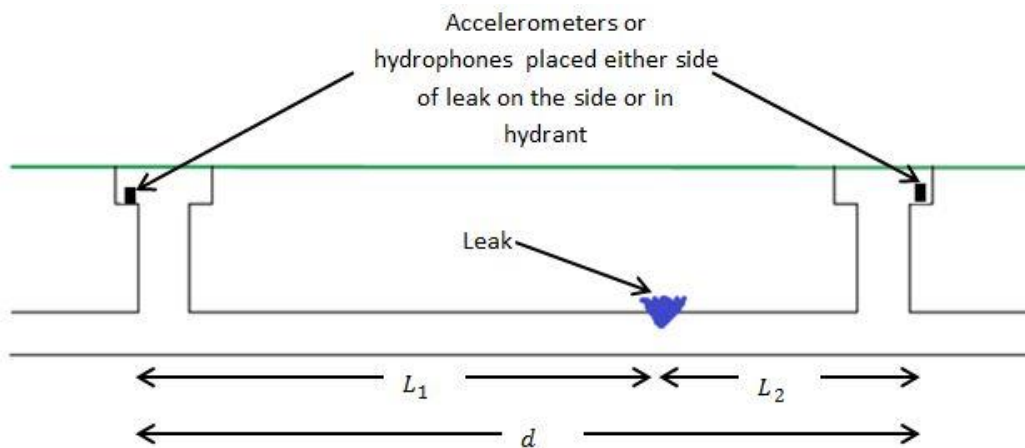


Figure 1: Identifying the presence of a leak. Figure reproduced from [5]

UK water companies are under increased pressure from customers and the water economic regulator (Ofwat) and customers to reduce leakage levels. Ofwat incentives water companies to reduce leakage with economic incentives and also to achieve their Sustainable Economic Level of Leakage (SELL). The SELL mechanism provides a trade-off between the cost of fixing a leak (pipe repair, road excavation etc.) with the cost of the water lost. Therefore, in some cases it is cheaper not to repair a leak. As the cost of leakage is directly linked to a leak flow rate, water companies can achieve their SELL through knowledge of their leak flow rate. However, no tool exists to accurately predict leak flow rate in WDS despite the benefits of doing so.

The prediction of leak flow rate in WDS will provide authors with valuable tool in order to prioritise leak repair. A few attempts have been tried in the gas industry but there is no technology actively used in the water industry despite the obvious benefits. Butterfield et al. [7] compared a variety of data driven methods in order to predict leak flow rate. They found that the signal Root Mean Square (RMS) had a strong correlation with leak flow rate. Signal RMS was also shown to be useful predictor of leak flow rate through leaky gas valves by Kaewwaewnoi et al. [8] and Chen et al. [9]. The RMS is of signal with samples, $x[0], x[1], \dots, x[N - 1]$ is given by:

$$RMS = \left(\frac{1}{2} \sum_{n=0}^{N-1} x[n]^2 \right)^{0.5} \quad (3)$$

Such data driven methods rely on the comparison of new data to existing data and therefore in order to make accurate predictions a large data set is required [10]. The advantage of such data driven methods is that they are intrinsic to the leak and therefore represent information about the leaks flow rate, leak shape, surrounding backfill etc. However, as such a large training set is required, it would be difficult to gain a training set big enough to demonstrate the wide ranging conditions found in WDS.

Of greater value would be a method that can collect data and learn from the data set that will adapt to the wide ranging conditions found in a real WDS. A method from machine learning may provide useful tool in this regard. Sun et al. [11] presented a study whereby a number of ‘features’ were extracted from a leak’s acoustic signal from leaky gas pipes in order to classify the aperture of round hole leaks. These features were used as inputs in to a Support Vector Machine model. They found time-frequency features to be useful for leak aperture recognition. However, the study by Sun et al. [11] only classified round holes which only represent a small proportion of the leak shapes found on real WDS and moreover took place on gas pipes.

Least Squares-Support Vector Machines (LS-SVM) models have been shown to be successful in predictive problems in a wide variety of disciplines, including fault detection in non-linear systems [12], cracks in images [13] to financial time series prediction [14]. LS-SVM was originally developed by Suykens et al [15] and it is advised that the reader refers to Suykens et al. [15] for the associated detail on the algorithm. The overall aim of this paper is to therefore establish whether enough information exists within the leak acoustic signal in order to predict leak flow rate.

2 METHODOLOGY

2.1 Experimental methods

The research conducted herein utilised a novel Leaks in ViscoElastics (LIVE) pipe rig at the University of Sheffield, UK, which was specifically designed for the purpose of this study. A schematic of the pipe rig can be found in Figure 2. The pipe rig consists of a 26 m long, 63 mm outer diameter Medium Density Polyethylene pipe loop. Water is pumped from an upstream reservoir (0.95 m³ volume) by a 3.5 kW (Wilo, Burton upon Trent, UK) variable speed pump set at 15 rpm. System pressure is measured by two pressure sensors (Gems Plainville 2200) downstream and upstream of the leak. A removable 5.5 m ‘test section’ is situated in the middle of the pipe rig. This removable section allows for the alteration of leak shapes and sizes. 3 leak shapes consisting of a number of sizes were drilled in to these test sections (further details about the leak shapes and sizes are provided in Table 1). A separate test section was used for each test. System pressure was varied in order to alter leak flow rate by adjusting the downstream gate valve. The removable test section passes through a rectangular box measuring 0.5 x 0.5 x 0.5 m whereby the backfill type can be alternated. 3 different backfill types were measured for all leak shapes, including 5-12 mm diameter pea gravel, geotextile fabric and submerging the pipe in water. The backfill types represent idealised versions of backfills found on real distribution networks.

Table 1. Leak shapes measured. All shapes were recorded at 5 different leak flow rates under 3 different backfill types.

Leak Shape	Width	Length	Diameter	Depths	Area
Horizontal Slit	2 mm	5 mm			10 mm ²
	2 mm	8 mm			16 mm ²
	2 mm	12 mm			24 mm ²
	2 mm	16.5 mm			33 mm ²
Round Hole			3.5 mm		9.6 mm ²
			4.5 mm		15.9 mm ²
			5.5 mm		23.8 mm ²
			6.5 mm		33.2 mm ²
Electrofusion Joint	4 mm			4 mm	16 mm ²
	8 mm			4 mm	32 mm ²

The leak noise was measured using 2 hydrophones Bruel and Kjaer type 8103, 50 x 9.5 mm), located approximately 2.25 m from the leak. The hydrophones were set at a sample rate of 2500 Hz. Hydrophones were connected to a current source unit (Dyrtran Instruments type 4102C) and then connected to a CCLD conditioning amplifier via a 6m integral cable. 20 samples were taken at each leak flow rate for each leak shape and backfill type.

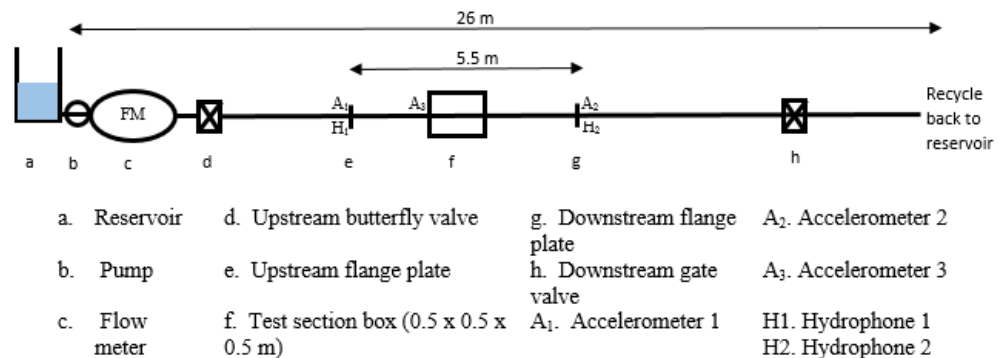


Figure 2: Schematic of the LIVE pipe at the University of Sheffield, UK. Reproduced from Butterfield et al. [5]

2.2 Signal processing methods

Signals gathered from the hydrophones were imported in to Matlab and initially bandpass filtered using a 4th Order Butterworth bandpass filter, at set points <10 Hz and >1000 Hz. Vital to the success of the model is the accurate derivation of a number of signal features. 24 features representing time domain, frequency domain and time-frequency domain were extracted from the leak’s acoustic emission signal and used and imported in to the LS-SVM model. A list of the features can be found in Table 2. Redundant features were removed from the model by identifying the features most relevant to the model using the greedy forward search algorithm.

Table 2. List of features used as inputs in to the LS-SVM mode.

Feature No.	Name	Feature No.	Name
1-6	RMS of IMFs1-6	18	Standard deviation
7-12	Shannon Entropy of IMFs 1-6	19	Signal power
13	Shannon Entropy of whole signal	20	Fundamental frequency
14	RMS of whole signal	21	Spectral flux
15	Mean dB of PSD	22	Kurtosis
16	Maximum dB of PSD	23	Skewness
17	Minimum dB of PSD	24	Crest factor

3 RESULTS

3.1 Performance of the leak flow rate prediction model

The model output of the LS-SVM model presented in Figure 3 includes round holes, longitudinal slits and electrofusion joints, all buried in either gravel, geotextile fabric or submerged in water. The individual shapes are colour coded in the model in order to establish if there is an effect on leak shape. The colour coding in Figure 3 demonstrates there is little trend between accuracy and leak shape. However, a visual inspection it appears that electrofusion joints tend to predict more accurately compared to the other leak shapes, whilst longitudinal slits and round holes tend to have a slightly larger spread of data. The associated Root Mean Square Error of this model output is 3.19 l/min.

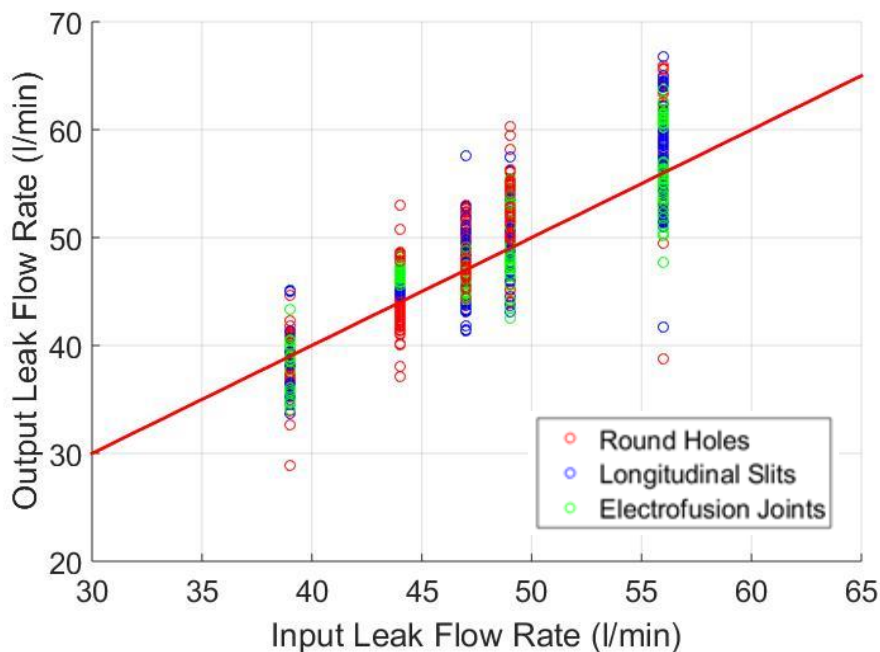


Figure 3: LS-SVM output using hydrophone data for leak flow rate prediction. Individual colours denote the predictive results for each leak shape.

Figure 4 demonstrates the accuracy in prediction ($\pm 10\%$). Both figures clearly show that it is possible to predict leak flow rate, despite shape and backfill type to $>78\%$. Of the flow rates studied, Figure 4 suggests that the model performed better at lower leak flow rates and worst at the highest leak flow rates.

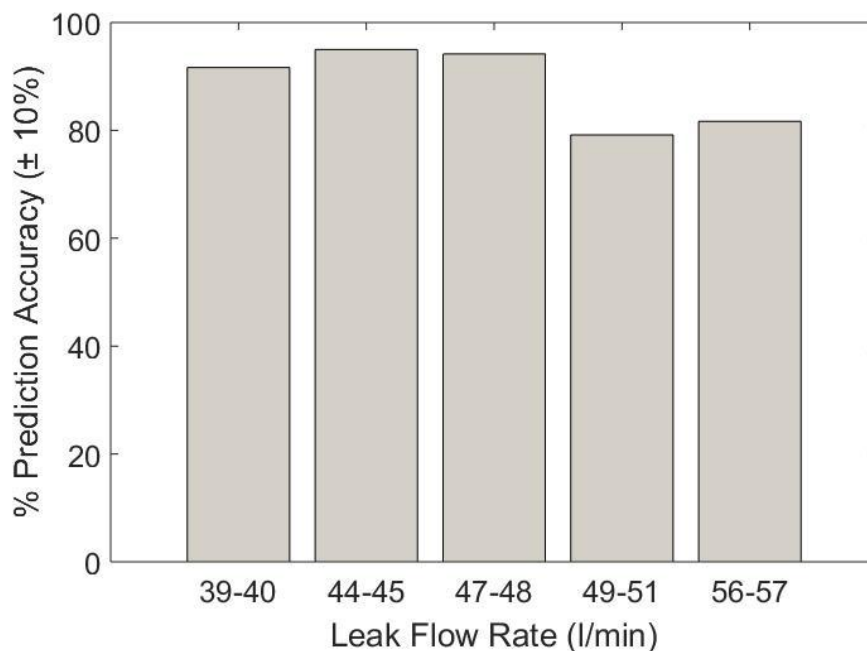


Figure 4: Accuracy in the prediction of leak flow rate found from the LS-SVM model ($\pm 10\%$).

3.2 Sensitivity of the model to leak shape and backfill

The influence of leak shape on the prediction accuracy was evaluated by training on the model on one leak shape and tested on the other leak shapes. In this case, only the features relating to the individual leak shape is used as an input to the model. The results presented in Figure 5 demonstrate the model being trained on one leak shape and tested on another. The results clearly demonstrate that the best results will always come from knowing the leak shape prior to predicting leak flow rate, which is shown in all the figures. The maximum prediction accuracy came from training on electrofusion joints and testing on electrofusion joints at the lowest leak flow rate achieving 96% accuracy (Figure 5c). On average, electrofusion joints had the greatest prediction accuracy when comparing all results, however this only came from training and testing on electrofusion joints (92% on average).

By training on one shape and testing on an alternative, this appeared to impact the model negatively, reducing prediction accuracy, always dropping to $<50\%$. This is true for all cases, apart from training on round holes and testing on longitudinal slits which was able to achieve 56% prediction accuracy (Figure 5b). The worst results came from training on longitudinal slits and testing on electrofusion joints, which achieved an average of 32.8% prediction accuracy across all leak flow

rates. This appears to demonstrate that the leak shape has a significant impact on the leak signal, and this influence was enough in order to have a negative influence on the prediction accuracy.

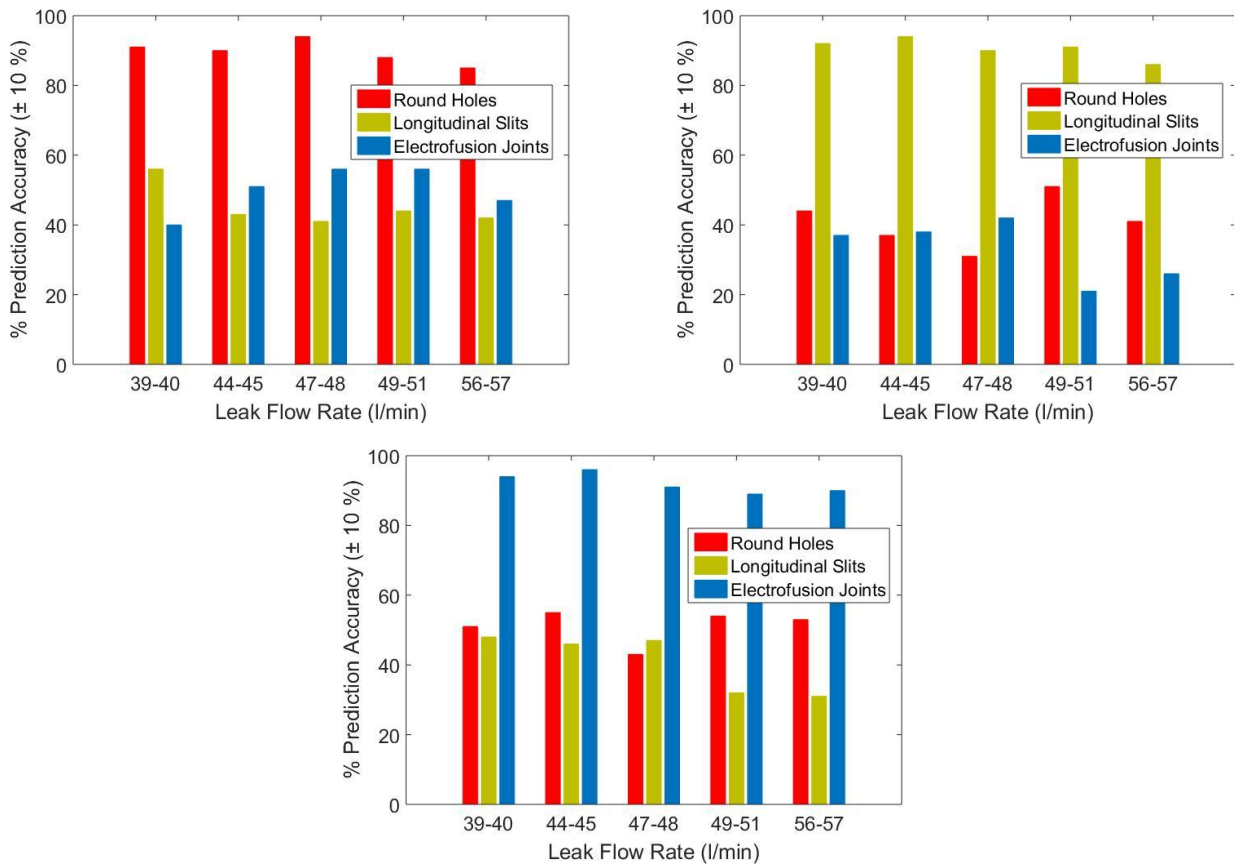


Figure 5: Sensitivity of model to changes in backfill type: (a) trained on round holes and tested on alternate leak shapes; (b) trained on longitudinal slits and tested on alternate leak shapes; and (c) trained on electrofusion joints and tested on alternate leak shapes.

The sensitivity of the model to backfill type was assessed by training on one backfill type and testing on an alternative, in a similar fashion to the above. The results for these are shown in Figure 6. The best results would always be achieved by training and testing on the same backfill type. This was shown especially in the case of gravel and submerged backfill, which when trained and tested on gravel was able to achieve an average of >90% prediction accuracy across all leak flow rates. In the case of the submerged, accuracy was slightly less 85.2% across all leak flow rates when training and testing on the submerged data. Training on one backfill type and testing on an alternative backfill type resulted in a reduction in prediction accuracy. The most negatively impacted seemed to be from training on submerged and then testing on geotextile fabric backfill.

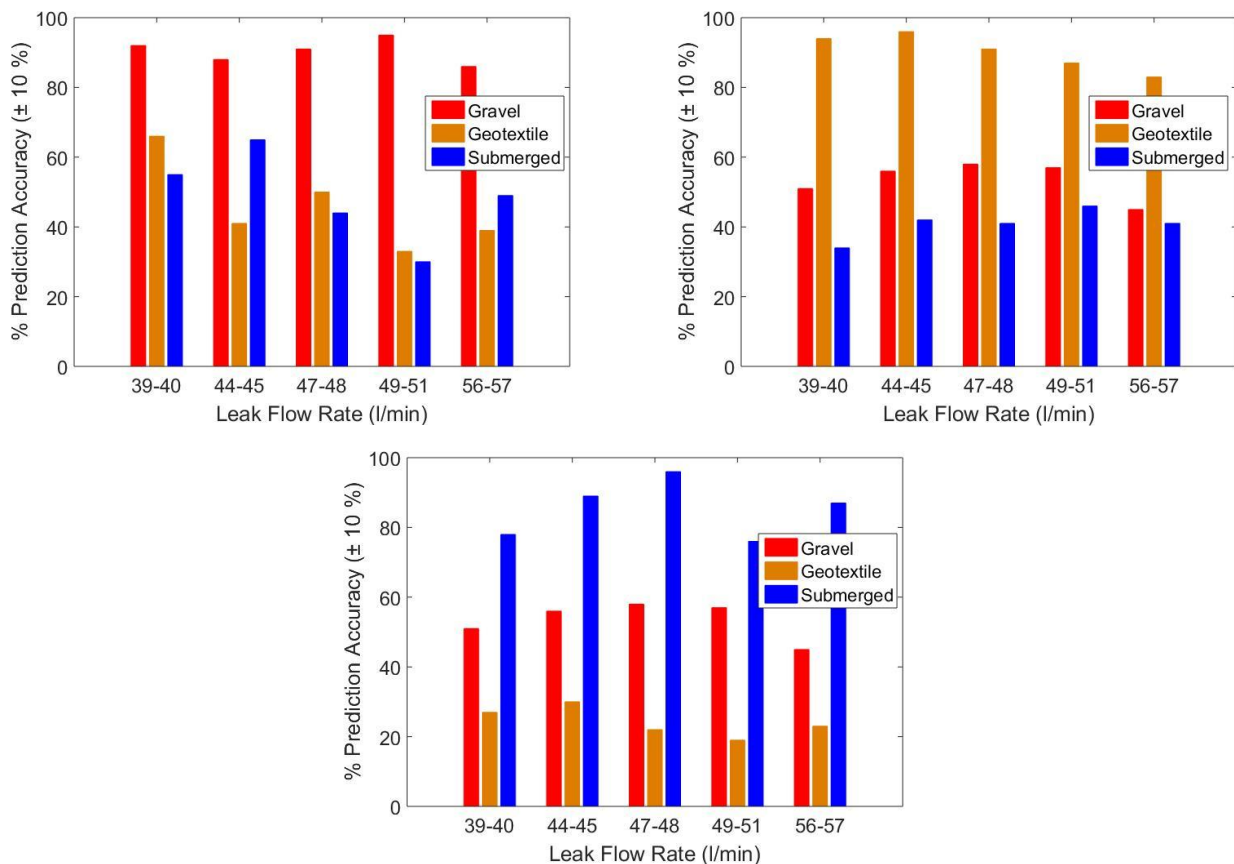


Figure 6: Sensitivity of model to changes in backfill type: (a) trained on gravel and tested on alternate backfill; (b) trained on geotextile and tested on alternate backfill; and (c) trained on submerged and tested on alternate backfill.

4 CONCLUSIONS

The research presented in this paper presents a method of predicting leak flow rate using solely the leaks acoustic emission signal. It was found that the leak flow rate can be predicted to an accuracy of >78% independent of leak shape and backfill type. It was also found that the model was highly sensitive to changes in leak shape and backfill type. The results presented in this paper are of interest to water industry practitioners as the model provides them with a tool to prioritise leak repair based on the leak flow rate.

Acknowledgements

The authors would like to thank Northumbrian Water, Severn Trent Water, Thames Water Utilities, Scottish Water and the EPSRC – United Kingdom under grant number EP/G037094/1 for their funding and help with this research.

References

- [1] A. F. Colombo and B. W. Karney, “Energy and Costs of Leaky Pipes: Toward Comprehensive Picture,” *J. Water Resour. Plan. Manag.*, vol. 128, no. 6, pp. 441–450, 2002.
- [2] F. Almeida, M. Brennan, P. Joseph, S. Whitfield, S. Dray, and A. Paschoalini, “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*, vol. 14, no. 3, pp. 5595–5610, 2014.
- [3] J. D. Butterfield, R. P. Collins, and S. B. M. Beck, “Influence of pipe material on the transmission of leak vibro-acoustic signals in water distribution systems,” in *Proceedings 8th International Conference on Acoustic Emission*, 2016.
- [4] M. Pal, “Leak detection and location in polyethylene pipes,” 2008.
- [5] J. D. Butterfield, R. P. Collins, A. Krynkin, and S. B. M. Beck, “Experimental investigation into the influence of backfill types on the vibro-acoustic characteristics of leaks in MDPE pipe,” *Procedia Eng.*, pp. 1–9, 2016.
- [6] J. M. Muggleton and M. J. Brennan, “Leak noise propagation and attenuation in submerged plastic water pipes,” *J. Sound Vib.*, vol. 278, no. 3, pp. 527–537, 2004.
- [7] J. D. Butterfield, A. Krynkin, R. P. Collins, and S. B. M. Beck, “Experimental investigation into vibro-acoustic emission signal processing techniques to quantify leak flow rate in plastic water distribution pipes,” *Appl. Acoust.*, 2017.
- [8] W. Kaewwaewnoi, A. Prateepasen, and P. Kaewtrakulpong, “Investigation of the Relationship Between Internal Fluid Leakage through a Valve and the Acoustic Emission Generated from the Leakage,” *Meas. J. Int. Meas. Confed.*, vol. 43, no. 2, pp. 274–282, 2010.
- [9] P. Chen, P. S. K. Chua, and G. H. Lim, “A study of hydraulic seal integrity,” *Mech. Syst. Signal Process.*, vol. 21, no. 2, pp. 1115–1126, 2007.
- [10] M. Meland, E. Thornhill, N.F. Lunde, E. Rasmussen, “Quantification of valve leakage rates,” *AIChE J.*, vol. 58, no. 4, pp. 1181–1193, 2011.
- [11] J. Sun, Q. Xiao, J. Wen, and Y. Zhang, “Natural gas pipeline leak aperture identification and location based on local mean decomposition analysis,” *Meas. J. Int. Meas. Confed.*, vol. 79, pp. 147–157, 2016.
- [12] T. Khawaja and G. J. Vachtsevanos, “A novel Bayesian least squares support vector machine based anomaly detector for fault diagnosis,” *Proc. Annu. Conf. Progn. Heal. Manag. Soc.*, pp. 1–8, 2009.
- [13] Y. Gao and Y. Yang, “Crack image detection based on LS-SVM optimized by PSO,” *Control Conf. (CCC), 2013 32nd Chinese*, vol. 1, no. 1, pp. 3602–3606, 2013.
- [14] T. Van Gestel, J. A. K. Suykens, D. Baestaens, A. Lambrechts, G. Lanckriet, B. Vandaele, B. De Moor, and J. Vandewalle, “Financial Time Series Prediction Using Least Squares Support Vector Machines Within the Evidence Framework,” *IEEE Trans. Neural Networks*, vol. 12, no. 4, pp. 809–821, 2001.
- [15] J. A. . Suykens and J. Vandewalle, “Least squares support vector machine classifiers,” *Neural processing letters*, vol. 9, no. 3. pp. 293–300, 1999.