A step towards the live identification of pipe obstructions with the use of passive acoustic emission and supervised machine learning

Hefft, Daniel Ingo ; Alberini, Federico

DOI:
10.1016/j.biosystemseng.2019.12.015

Document Version
Peer reviewed version

Citation for published version (Harvard):
Hefft, DI & Alberini, F 2020, 'A step towards the live identification of pipe obstructions with the use of passive acoustic emission and supervised machine learning', *Biosystems Engineering*, vol. 191, pp. 48-59. https://doi.org/10.1016/j.biosystemseng.2019.12.015

A step towards the live identification of pipe obstructions in channels with the use of passive acoustic emission and supervised machine learning

Daniel Ingo Hefft[1], *Federico Alberini*[1*]

[1] *School of Chemical Engineering, University of Birmingham, Edgbaston, B15 2TT, UK*

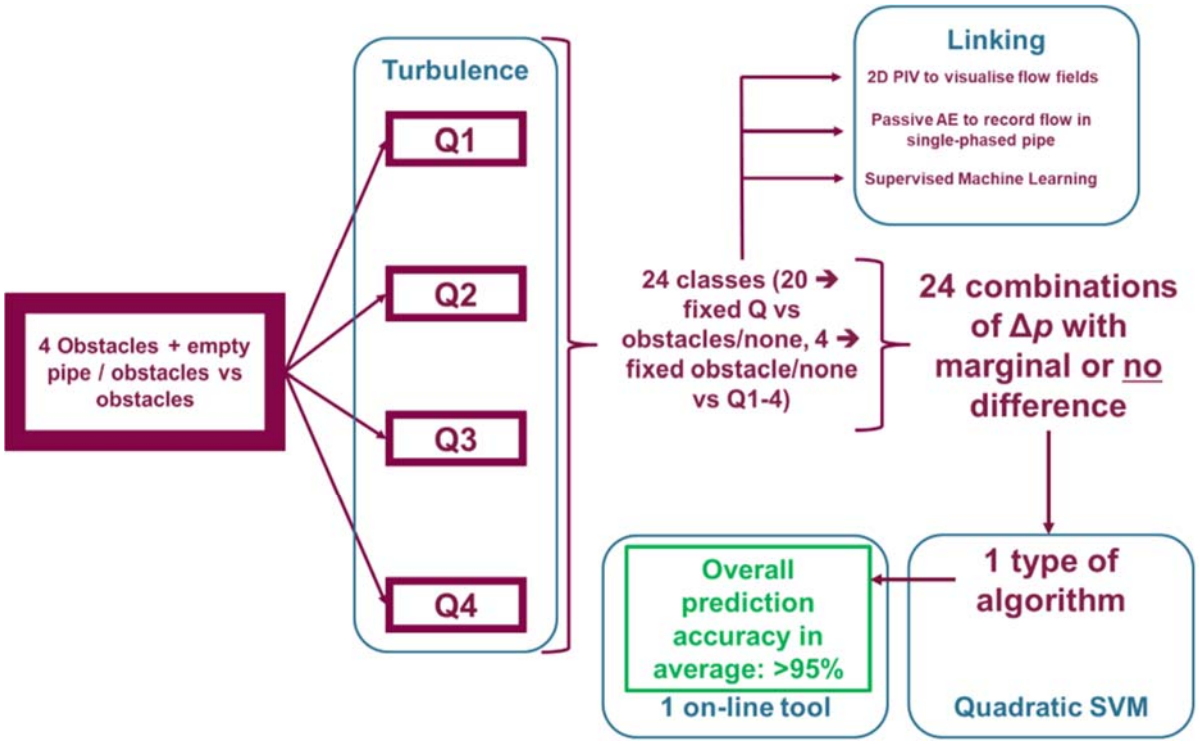
**corresponding author: f.alberini@bham.ac.uk*

Abstract

A single passive acoustic emission sensor is used to collect signals coming from an obstructed pipe in a water recirculation system. Four geometrically different obstructions are under investigation. The flow field of water around each obstruction is visualised with the use of 2D Particle Image Velocimetry to identify the different flow features. In parallel, the acoustic emission signals have been acquired by locating a piezoelectric sensor on the outer wall of the pipe at the tip of the obstruction. The acoustic emission signals are then pre-processed and the frequency domain is extracted for 100 recordings in each case. Signals are processed further by using Principle Component Analysis and a matrix is created for Supervised Machine Learning algorithms. This methodology is applied over a range of four flow rates, all in fully developed turbulent flow.

Results show that different obstructions generate different acoustic signals and flow fields, which reflect the different flow fields observed with PIV. The average velocity and amplitude of the acoustic signals are increasing in magnitude with an increase in flow rate. The machine-learning algorithm with highest prediction values is quadratic SVM with predictions in the area of 95% accuracy or above. This makes the combination of machine learning and a single passive acoustic sensor a viable option to predict pipe obstructions and the type of obstruction. This may lead to a useful application for urban water supply or sewage system as well as agricultural practice for field irrigation or the detection of nozzle blockages.

Graphical Abstract



Keywords

Obstruction; hydrology; pipe blockage; machine learning; Particle Image Velocimetry; acoustics

Research Highlights

- Online tool to predict blockages in pipes.
- The use of a single acoustic emission sensor.
- Combination of supervised machine learning and passive acoustics.
- Prediction accuracies of 95% and more.

Symbols and abbreviations

Roman

A_{circle}	Area of the pipe
b	A scalar in the hyperplane equation. A physical quantity that does not change with changes in position or orientation
D	Diameter
d_{pipe}	Internal diameter of the pipe
H	Height
L	Length
L_p	Length of the interrogation area
v_{max}	maximum velocity in the pipe
$v_{superficial}$	Superficial velocity in the pipe
$\dot{V}$	Volumetric flow rate
W	Normal vector to the hyperplane (W, b)
X	Position vector
Y	Peak Intensity

Greek

Δt	Time difference of laser pulses
η_{fluid}	Dynamic viscosity
ρ	Density

Acronyms

~~AE~~	~~Acoustic emission~~
AI	Artificial Intelligence
AVG%	arithmetic mean percentages
CCD	Charge-coupled device
CLA	Classification Learner Application
FFT	Fast Fourier Transform
H	Height
ID	Internal Diameter

k-NN	k-nearest neighbour, a machine learning classifier
IoT	Internet of things
Nd:Yag	Neodymium-doped yttrium aluminium garnet
ML	Machine Learning
PIV	Particle Image Velocimetry
SVM	Support-Vector Machine

1. Introduction

Water is the most precious resource on the planet and even subject to several local conflicts on the globe (Gleick, 2014; Lohmar, Wang, Rozelle, Huang, & Dawe, 2003; Treszkai, 2018). Agriculture is by distance the largest sector in water consumption, by using up 60% to 90% of the total water available from human use (Thenkabail, Hanjra, Dheeravath, & Gumma, 2017). Also with projections of population increase and climate change will make this resource scarcer than ever before.

That makes it necessary to create effective engineering solutions to limit losses and detect faulty supply systems. Blocked water pipes will have a negative impact on field irrigation and livestock. Also, recent trends towards intelligent precision farming makes it necessary to develop online tools to check farming systems, improve farming practice and increase sustainability. The so-called internet of things (IoT) or algorithm driven systems (machine learning, artificial intelligence (AI)) support this trend (The Economist Newspaper Limited, 2016).

This paper explores the combination of passive acoustic emission sensing ~~(AE)~~ and supervised machine learning to create prediction algorithms for different pipe obstructions and flow rates.

This is a first approach to combine a patent pending mechanical design, a single passive acoustic emission sensor and supervised machine learning to detect changes of the fluid (not the structure (i.e. the pipe wall)) to detect blockages in a unbranched pipe system.

The flow field is visualised by using 2D Particle Image Velocimetry (PIV) and the pressure drop is measured for each obstacle.

For the measurements of acoustic emissions, there are two sensor systems available. Such sensors can either be *active*, meaning they are based on an emitter-receiver system, or *passive*. Where *active* acoustic emission sensors measure the change of introduced acoustic waves (sometimes referred as slot waves) over the distance from the emitter to the receiver, *passive* acoustic emission sensors only detect the acoustic emission that the process emits itself (Boyd & Varley, 2001). Hence, *active* acoustic emission is based on an active energy input whereas *passive* acoustic emission simply measures the energy release.

Although *active* acoustic emission sensors are more present in research and in industry they face several challenges, which include loss of signals due to obstructions, bubbles, distance or simply not reaching the necessary penetration depth (Borenstein & Koren, 1988; Hauptmann, Hoppe, & Püttmer, 2002). Overall, *active* acoustic emission sensors prove to give good predictions on factors such as flow rate, degree of gassing or solid content. This technology works well for Newtonian and non-Newtonian fluids (Kotzé, Ricci, Birkhofer, & Wiklund, 2016; Rahman, Håkansson, & Wiklund, 2015).

There have been several attempts to make use of acoustic emission data to gain a greater understanding of pipe flow (Joseph D. Butterfield, Collins, Krynkin, & Beck, 2017; Hou, Hunt, & Williams, 1999; Li, Song, & Zhou, 2018). However, work has mainly been focussed on active acoustic emission (Borodina, Zaitsev, & Teplykh, 2018; Das, Das, & Mazumdar, 2013), multiphase flow (Hou et al., 1999; O'Keefe, Maron, Felix, van der Spek, & Rothman, 2010) or water (Khulief & Khalifa, 2012; Li et al., 2018; Martini, Troncossi, & Rivola, 2016).

Passive acoustic emission sensors are used for leak detection in water pipes by employing in-pipe hydrophones (Chatzigeorgiou, Youcef-Toumi, & Ben-Mansour, 2015; Khulief & Khalifa, 2012) or by aiming to recognise acoustic patterns based on the signals of a series of sensors (Li et al., 2018). Other work on passive acoustic emission sensing has mainly focused on

multiphase systems (Finfer et al., 2015; Hou et al., 1999; O'Keefe et al., 2010). However, no solutions have been created so far on enclosed, fully flooded and single-phased pipe systems. This also includes creating a new system to use passive acoustic emission sensing as a mean to get information about fluid flow fields and as a tool for process control.

In addition, all know passive acoustic emission technologies are based on clamp on solutions and at least 2 sensors (J.D. Butterfield, Krynkin, Collins, & Beck, 2017). The solution proposed is only based on the application of a single passive acoustic emission device that is capable to predict blockages upstream. The technology looks at changes to the acoustic information of the fluid (by implementing a T shaped pin) and not like traditional approaches to structural changes of the pipe. To overcome issues around attenuation (Martini, Rivola, & Troncossi, 2018; Pavić, 1992) the T shaped pin is made from metal, however, can easily be incorporated in other materials. Other technologies to overcome attenuation problems can be overcome by computational solutions and pattern recognition technologies, however, do require additional computational power (Sause, Gribov, Unwin, & Horn, 2012). Another advantage of the technology compared to hydrophones is that the sensor is placed on the outer wall of the pipe. Hydrophones are in-pipe solutions, which makes it difficult to maintain or replace such equipment. A hydrophone are microphones that have been waterproofed which makes them suitable to be immersed in fluids. By positioning the hydrophone close to a suspected leak or by applying those inside a pipe system the distinctive sound of water escaping can be detected. Hydrophones can additionally be used to listen for leaks in pressurised pipe systems that run close to the wall (Khalifa, Chatzigeorgiou, Youcef-Toumi, Khulief, & Ben-Mansour, 2010) and work for plastic piping well (Hunaidi & Chu, 1999).

A key challenge making use of such data is the sheer volume of incoming data points. To make use of such complex and high-volume data machine learning can be used.

The term *machine learning* (ML) applies to artificial systems that can generate knowledge based on experience. Such systems learn on given, and to the problem representative, examples. It is the system's task to create logical paths or identify logical patterns to the given problem.

In general, *machine learning* can either solve classification or regression problems. In the first case *machine learning* aims to classify data into classes, such as fault classification. Regression problems are those where the system will be asked to provide future predictions of a continuous variable (e.g. temperature trends) (Kubat, 2015; Moraru, Pesko, Porcius, Fortuna, & Mladenic, 2010).

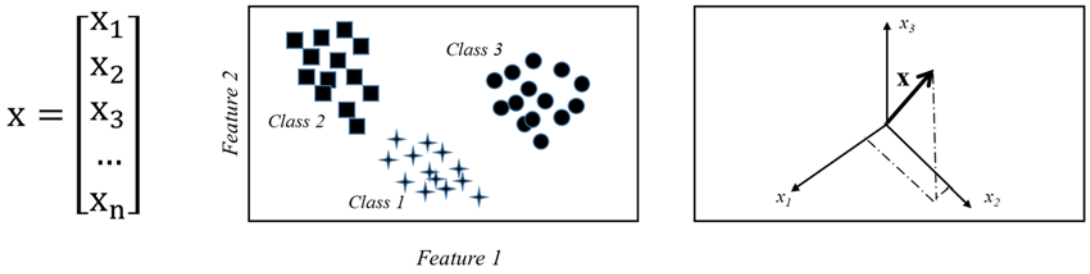
In the end of this process, the system will create a general algorithm/ system that is capable to solve problems to previously unseen datasets. The system/ algorithm can also evaluate unknown datasets (learning transfer) or fail to learn unknown data (overfitting) (Koza, Bennett, Andre, & Keane, 1996; Reitmaier, 2015).

Based on the methodology ML approaches a problem and makes use of datasets three types of ML learning algorithms can be distinguished; namely *supervised*, *unsupervised* and *reinforced* learning (Fischer, 1999).

The majority of machine learner employed in industry (between 80 to 90%) are of supervised and unsupervised nature with a dominance towards supervised machine learning (Ge, Song, Ding, & Huang, 2017).

The way the system learns is by giving it labelled examples, hence the answer to the question is already known. The system will try to build a logical path with the help of algorithms (Quadratic SVM, Decision Trees etc.). This will be done with only a selected proportion of

dataset, whereas the remaining parts are used to optimise the algorithm and finally test it to an unknown dataset, with a common split being 60% for training, 20% for optimisation and 20% for testing (Chong, Abraham, & Paprzycki, 2004; Wuest, Weimer, Irgens, & Thoben, 2016). Results are projected as observations, which are dependent on the number of features. An observation is a specific number set that characterises a class (Kotsiantis, Zaharakis, & Pintelas, 2007). Hence, data can be seen as points in an *n*-dimensional space. The number of attributes *n* are directly related to the class observed from experience or the experiments. In order to achieve a better classification between data classes, each class will be plotted in another dimensional space, the so-called feature space (Guan et al., 2019; Masud et al., 2010). *Figure ~~2~~ 1* illustrates the presentation of machine learning data.



Feature Vector | 2-dimensional scatter projection (component space) | 3-dimensional feature space projection

Figure 1 Schematic projections of machine learning data, adopted from (Liu & Wang, 2017).

2. Materials and Method

2.1 Water recirculation system and pipe obstructions

A water recirculation system fed by a 40 L water tank and is powered by a centrifugal pump (Alfa Laval, Sweden) of the type I KA-5 132SSS1. The internal diameter *ID* of the pipework is 25.4 mm with a 120 mm in length pipe segment from stainless steel, which is used for the ~~AE~~ data acquisition of passive acoustic signals. A schematic drawing of the test rig is given below (*Figure ~~1~~2*).

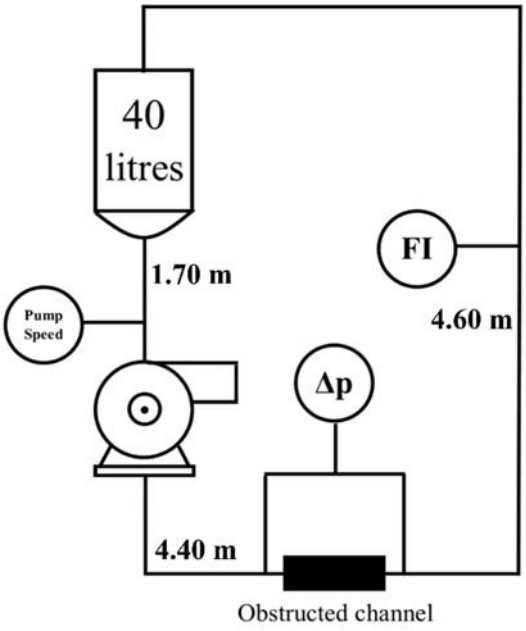


Figure ~~1~~ 2 Experimental setup (schematic).

All four in-pipe obstructions are designed in AutoCad 2018 (Autodesk Inc., USA), and extruded by a FlashForge Dreamer 3D printer (Zhejiang Flashforge 3D Technology Co., Ltd., China). Each obstruction can be slotted into the pipe and has a length of 1.5 times the inner diameter (38.1 mm). The four geometric obstructions are a wall-leaning cone, cross (four wedges meeting in the central focal point), three triangular aligned holes and a semicircle. An IMO iDrive AC Inverter Drive control unit (IMO Precision Controls Ltd, United Kingdom) is used to set the four flow rates of 1300, 3000, 4530 and 6350 l h^{-1}. The pressure drop has been measured by using a TPI 665L Digital Manometer (Test Products International Europe Ltd., UK).

All four ~~obstacles~~ in-pipe obstructions under investigation are presented in *Figure ~~2~~3*.

Figure ~~3~~1 3D animation of extruded obstacles to be slotted as obstructions into the pipe. ID=25.4 mm, H=38.1 mm, drawings are not to scale. Obstruction types are "cone", "three holes", "semicircle" and "cross" (top left to right, bottom left to right).

The free area of each obstruction is:

Cone: $1.287*10^{-4}$ m.

Three holes: $2.83*10^{-5}$ m.

Semicircle: $2.53*10^{-4}$ m.

Cross: $2.53*10^{-4}$ m.

The geometries chosen represent either potential deposit types (cone type for slight build-up, semi circular for half blocked pipe and three holes for a full blockage with perforations delivering partial flow) and the potential design of s spray nozzle (cross).

2.2 2D Particle Image Velocimetry

PIV is an established technology in fluid dynamics to visualise flow fields by a combining laser technology, a camera and seeding particles. PIV is a non-intrusive technology that delivers time-resolved velocity information and flow maps on a nearly instantaneous basis, however, is still mainly exclusively used within academia (Adrian, 2005). A pulsed laser beam is bend into a light slab that is fired onto a translucent chamber. The fluid within the chamber contains seeding particles that are assumed to behave as the fluid within the chamber. The laser pulse delay is referred as Δt. By the help of cross-correlation algorithms, the particle displacement (X, Y direction) and velocity can be determined.

The time delay Δt between each frame depends on the maximum distance a particle could travel within the interrogation window. The formula Δt for is given below and follows *Adrian* (1986):

$$\Delta t < 0.25 * \frac{S*L_p}{v_s} \quad (1)$$

with S being the magnification factor, L_p is the length of the interrogation area in pixels v_s for the superficial velocity in the pipe.

The superficial velocity v_s, a hypothetical flow velocity calculated which does not take local differences into account, in the pipe has been calculated by the division of the volumetric flow rate $\dot{V}$ by the area A_{circle} of the channel:

$$v_s = \frac{\dot{V}}{A_{circle}} \quad (2)$$

The Reynolds number in the pipe has been calculated by using

$$Re = \frac{v * d_{pipe} * \rho_{fluid}}{\eta_{fluid}} \quad (3)$$

with d_{pipe} being the internal pipe diameter and η_{fluid} being the dynamic viscosity of water.

For the flow field visualisation a TSI 2D PIV system (TSI Inc., USA) has been used. A dual head green 532 nm Nd:Yag laser (Litron Nano PIV,UK) pulsing at 7 Hz has been formed to light sheet. The laser is mounted and levelled 500 mm above the acrylic chamber, with the light sheet follows the length of the chamber. The system is synchronised to a single TSI Power View 4 megapixels (2048 x 2048 pixels) 12-bit CCD camera that can be controlled in its X-Y-Z direction by a remotely controllable frame construction. The CCD camera is connected to another synchroniser (TSI 610035) which is attached to a desktop PC (DELL Inc., USA). The 2D PIV system is controlled by TSI Insight 4G software. Since fully developed turbulent flow is apparent, 2x500 images are taken. The area is set to 32 x 32 pixels under the application of a recursive *Nyquist* grid. The resolution has been adjusted for each experiment as changeover between obstructions make readjustments necessary. Captured images are combined, *Fourier* transformed, averaged and debugged in order to determine the final averaged flow field. The generated flow field is visualised with Tecplot 360 software (Constellation Software, Canada). The seeding particles used are silver coated, hollow glass spheres (D= 10µm, ρ= 1400 kg m^{-3}) (Dantec Dynamics, Denmark).

2.3 Passive Acoustic Emission

The acoustic emissions are detected with a piezoelectric VS375-M sensor (Vallen Systeme GmbH, Germany). The sensor is linked to an AEP5H preamplifier (Vallen Systeme GmbH, Germany) along with a DCPL2 decoupling unit (Vallen Systeme GmbH, Germany), a PicoScope 5000 Series oscilloscope (Pico Technology Ltd, UK) and a personal computer using PicoScope version 6.13.15 software (Pico Technology Ltd, UK). One hundred recordings, the so-called buffers, each of a length of 500 ms, a resolution of 16-bit and an amplitude of maximum ±1 V are taken. The sampling number is set to 600 kS to ensure that the sampling frequency of 1.2 MHz is at least twice the number of the resonance frequency to meet Shannon's theorem (Shannon, 1948). The choice of 500 ms is justified as a 10 s buffer is divided that long until the lowest recording time is reached where Fast Fourier Transform (FFT) spectra are still visually similar. The sensor has been placed on the outer side of the pipe wall and sits on top of a circular shaped pinhead, located on the tip of the obstruction. This means that sensor and fluid are never getting into direct contact, which brings hygienic as well as maintenance benefits (patent application GB1909291.5).

2.3.1 Data Processing

Recorded acoustic emission are converted from the psdata file into mat files to make them available for Matlab R2018a (MathWorks Inc, USA). All data are processed on a computing system with 8-gigabyte memory. In a first step, each buffer gets assigned with a corresponding

label, describing the experimental condition (e.g. flow rate 1, obstacle "cross"). Once assigned the data is cleared from environmental noise. This is regarded to be frequencies below 4 kHz, meeting literature reporting environmental noise being in the range between 2-6 kHz and below (Chang et al., 2011; Forrest, 1994).

In a consecutive step, *Discrete Fourier Transform* is applied to transform the time-bound signals into the frequency domain (FFT spectrum). ~~In addition,~~When converted into Matlab, positive and negative infinitive values appear, representing over range values. These values are filtered and replaced by the value ±1 or for those values that are underneath the detection limit of the sensor the value zero is inserted.

The last reduction step is the selection of only the 5,000 FFT values with the largest relative variance. To get the relative variance the mean and the variance for each column (frequency) has to be found by using the following formulae:

$$V_j = \frac{1}{N-1} * \sum_{i=1}^{N} \left| A_{ij} - \mu_j \right|^2 \quad (4)$$

$$\mu_j = \frac{1}{N} * \sum_{i=1}^{N} A_{ij} \quad (5)$$

$$V_{Relative,j} = \frac{\frac{1}{N-1} * \sum_{i=1}^{N} \left| A_{ij} - \mu_j \right|^2}{\frac{1}{N} * \sum_{i=1}^{N} A_{ij}} \quad (6)$$

with A_{ij} being the magnitude of the FFT transform of the frequency j for the i-th row which corresponds to the i-th buffer.

The relative variance $V_{Relative}$ was chosen instead of the variance as this way absolute values can be weighted on the arithmetic mean values. This means in consequence that small values with a relative small absolute change in value but big relative change become considered.

2.4 Supervised Machine Learning

Support-Vector machines (SVM) are a classifier type within machine learning. The basic concept of SVM is the establishment of hyperplanes, seeking to create separation between classes. A hyperplane (for problems higher than 2-dimensional or binary classification) is the n-dimensional space that creates the boundary of data points for different clusters that belong to different classes.

The general equation for a hyperplane for the case of linear separability is given in *Equation 7*:

$$W^T * X + b = 0 \quad (7)$$

where $W = [w_1\ w_2\ \cdots\ w_m]^T$ and b is a scalar.

W represents the normal vector to the hyperplane (*W*, *b*) and *X* being the position vector of the points lying on the plane.

The classes are separated by their value being either positive (+) or negative (-) measured based upon their distance from the hyperplane (Dagher, 2008).

This idea of the hyperplane dates back to the mid-1930s (Fisher, 1936), however, a first appearance within the background of artificial networks was not done until 1958 (Rosenblatt, 1958). These initial concepts were further developed by *Vapnik & Chervonenkis* (1974) and lead to today's SVMs.

Noble & Street (2006) state four key concepts behind every SVM. Those are

- The separation of the hyperplane,
- The maximum margin hyperplane,
- The soft margin, and
- The kernel function.

To make the data available for the Matlab R2018a Classification Learner Application (CLA) (MathWorks Inc, USA) feature scaling and mean normalisation was applied to the reduced FFT spectrum. The necessity of this operation is to have features with a comparable range of values.

Principal Component Analysis (15 components) is enabled for the CLA. The goal of the PCA is to project the data points of the matrix into an *n*-dimensional subspace in such a way that as little information as possible is lost and existing redundancy is summarised in the form of correlation in the data points. This enhances the supervised ML performance, contributing to a better support vector delimitation. The dataset for the Machine Learning is further split into in a training (60%), optimisation (20%) and a test dataset (20%). The test dataset is given to the supervised classifiers (e.g. k-NN, Decision Tree, SVM) whilst the second dataset is not fed into the CLA to evaluate the accuracy of the algorithms.

Figure ~~3~~ 4 shows a summary of the steps from the raw data acquisition to a supervised machine learning prediction:

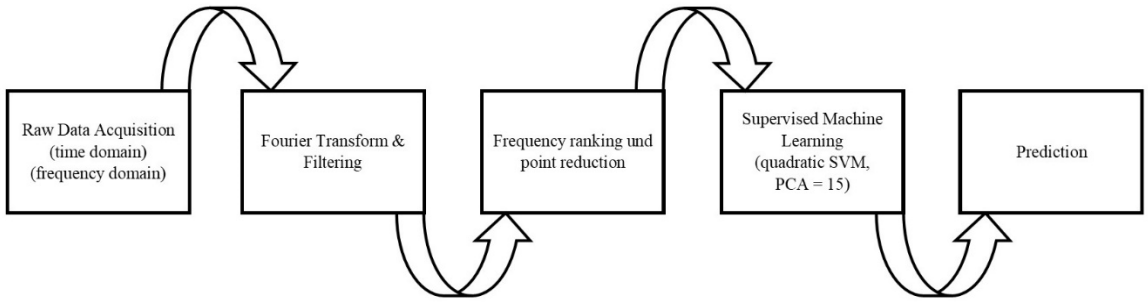


Figure ~~2~~ 4 Transformation schematic from the raw data acquisition to a prediction.

3. Results and discussion

3.1 2D Particle Image Velocimetry

2D PIV has been used to visualise the flow around four geometrically different in-pipe obstruction on fixed flow rates (1'300, 3'000, 4'530 and 6'350 l h^{-1}). All pictures represent fully developed turbulent flow, with flow running from right to left (abscissa). Selected pictures represents the time-averaged data over the 500 image pairs. The images acquired are presented in *Figures 4 5 and ~~5~~ 6*:

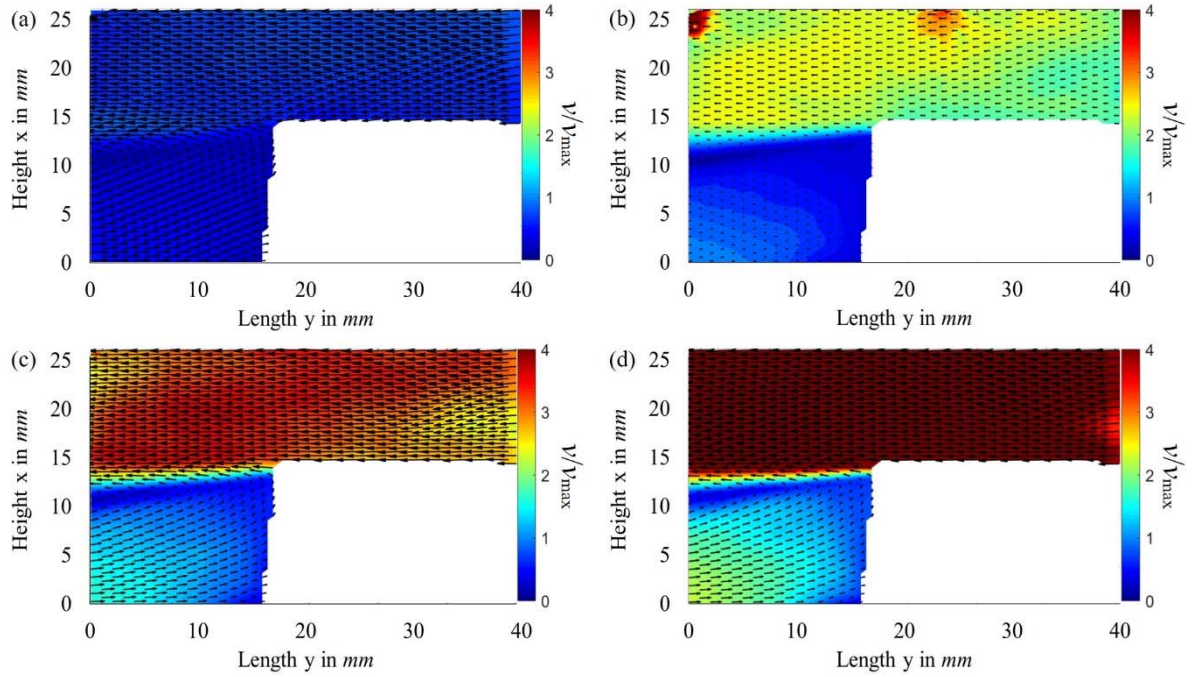


Figure ~~3~~ 5 Exemplary contour plot of 2D PIV results for one obstruction (semicircle) over four flow rates. The Reynolds numbers for the highest point of restriction are (a) Re ≈36,251, (b) Re ≈83,701, (c) Re ≈126,195 and (d) Re ≈176,182. Blanked parts show the in-pipe obstruction.

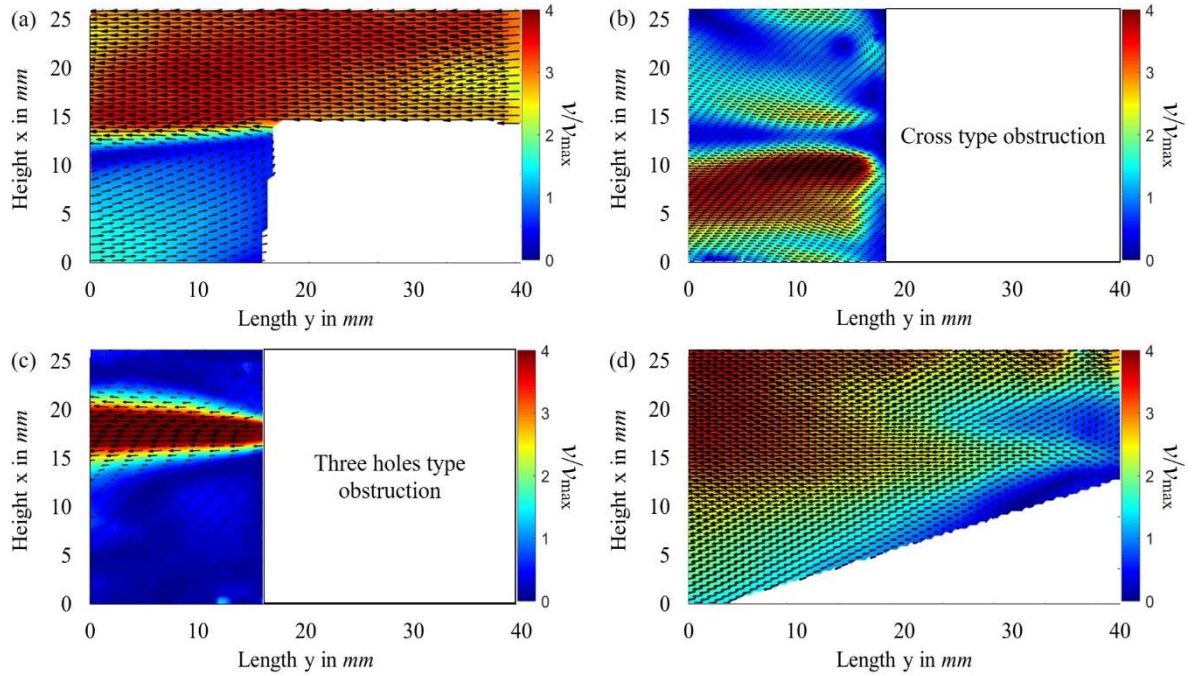


Figure ~~4~~ 6 Exemplary contour plot of 2D PIV results for one flow rate (4,530 lh^{-1}) against all obstruction types. The Reynolds numbers for the highest point of restriction are (a) Re ≈126,195, (b) Re ≈126,195, (c) Re ≈267,489 and (d) Re ≈252,390.

With increasing flow rates increases in velocity magnitude can be observed (*Figure ~~4~~ 5 (a)-(d)*).

Looking at the semicircular obstruction, a drag up effect is apparent along the *semicircle*'s tip (*Figure ~~4~~ 5 (a)-(d)*).

For the *cone* type obstruction (*Figure ~~5~~ 6 (d)*) preferential flow is present following the open space above the cone. When inserting the *cross* type obstruction (*Figure ~~5~~ 6 (b)*) a butterfly wing-like structure appears, provoked by swirls at the direct exit of the obstruction. In *Figure ~~5~~ 6 (c)* jetting is present for the three holes obstruction. The jet opens up in width with increasing flow rates.

3.2 Pressure Drop Measurements

The pressure drop has been measured 5 mm before and 120 mm behind the obstruction. Plotting the pressure drop Δp against the Reynolds number for the unobstructed pipe as per equation 3 the following figure can be obtained (*Figure ~~7~~6*).

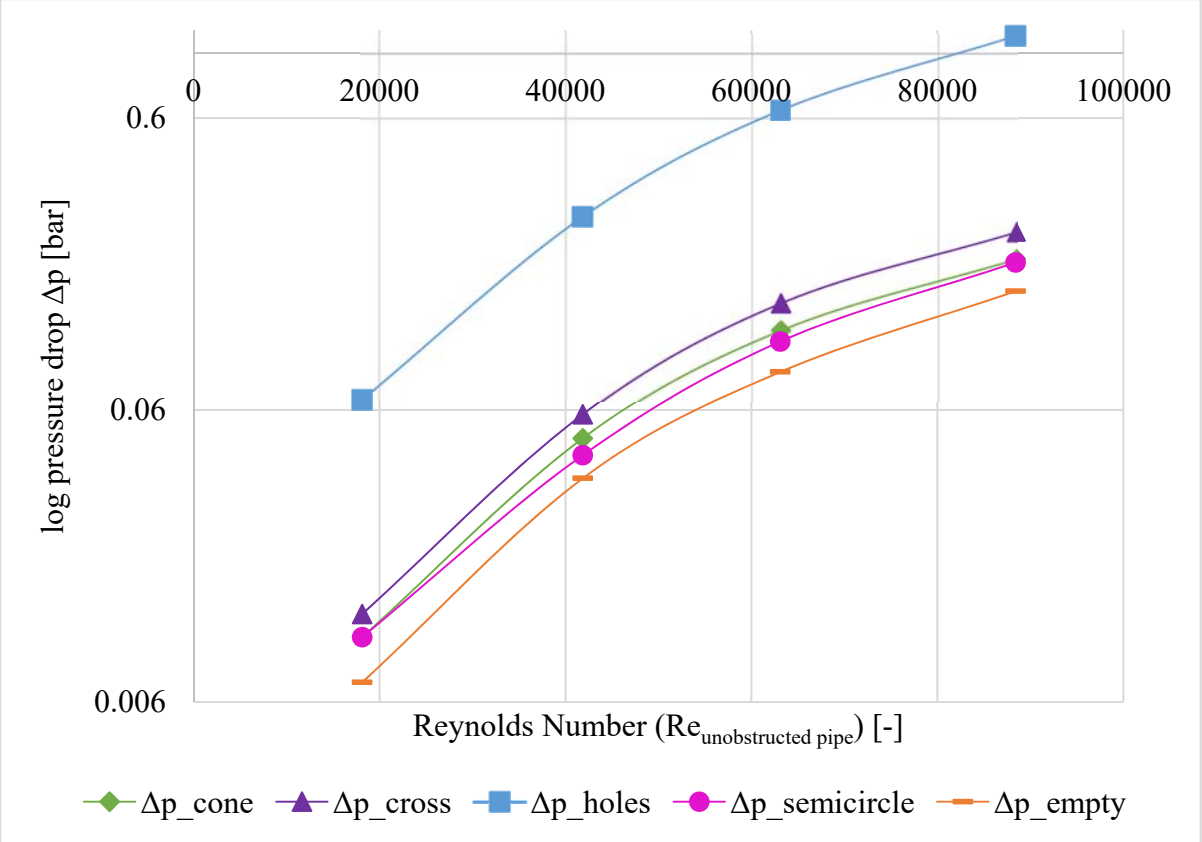


Figure ~~5~~ 7 Pressure drop Δp of different obstacles against Re as determined for the unobstructed pipe.

The obstruction type with the three holes shows a very different pressure drop to the other three obstructions. This will be explained with the high degree of restriction along each hole (D= 2 mm). Overall, all curves are in a parallel arrangement. For Re≈18,125 the semicircle and the cone share a value at 0.01 bar. Also for Re≈88,406 the data points are nearly congruent as they differ only by 0.007 bar. This shows that it is challenging to distinguish between different runs purely based on pressure drop measurements.

3.3 Pressure Drop Measurements versus Acoustics

As per 3.~~1~~2 it is challenging to distinguish between different experimental runs when looking at the pressure drop values. However, when linking the pressure drop to acoustic response, distinguishable responses can be retrieved. Even for the case where obstructions share the same pressure drop of 0.001 bar, it is possible to get different obstructions (*Figure ~~7~~8*).

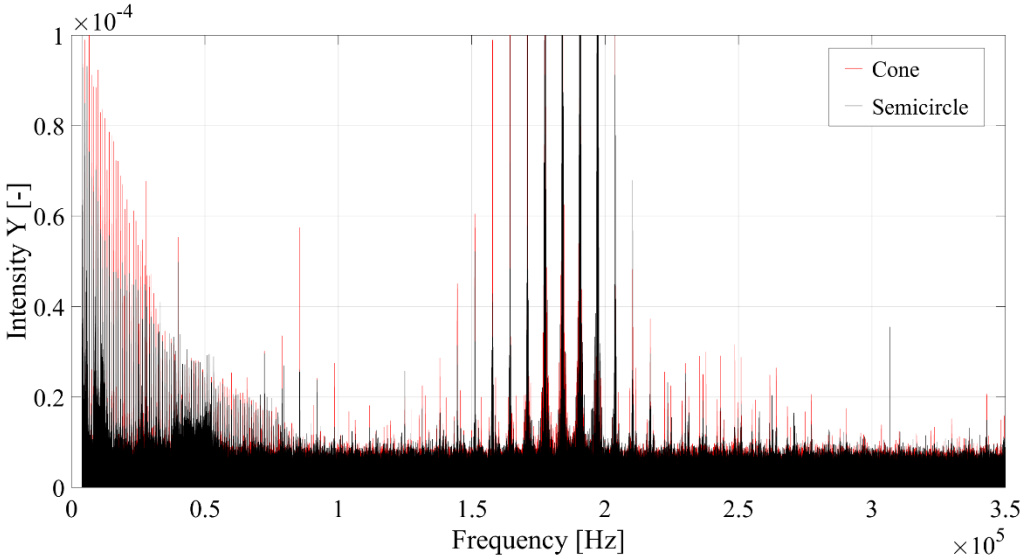


Figure ~~7~~ 8 Cropped and from environmental noise cleared FFT spectra for different obstructions (cone (red) versus semicircle (black)) but the same pressure drop Δp (0.01 bar).

This results in the advantage that the sensing technique is more sensitive than conventional methods. In addition, the detection of these pressure drops can consequently be used to predict leakages in a very early stage.

3.4 Acoustic Emission Signals

The signal collected from a piezoelectric passive acoustic emission sensor has been collected. After the removal of the environmental noise, the discrete Fourier Transform was applied to convert the signal from the time-domain into a frequency-domain. These acoustic "fingerprints" are presented in *Figures ~~8~~ 9 and ~~9~~10*.

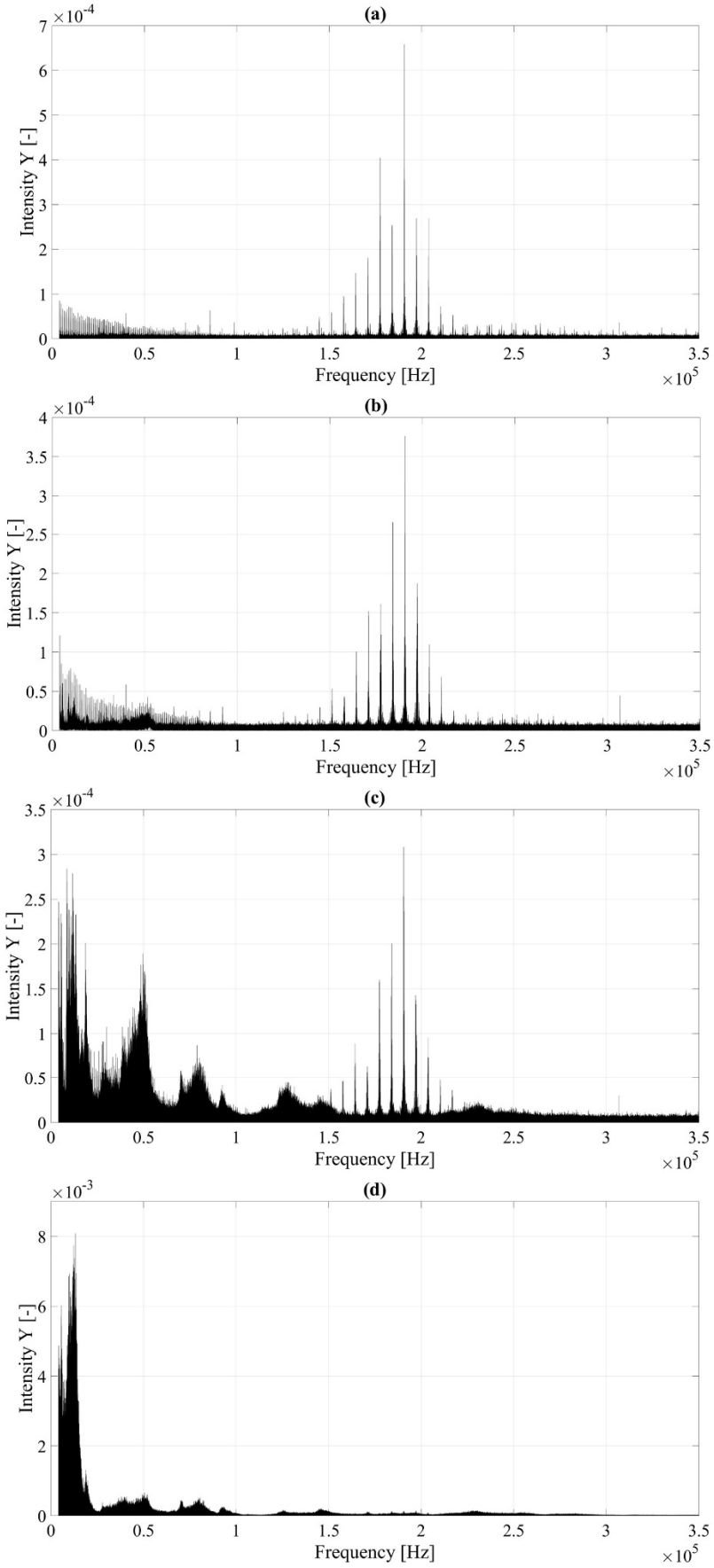


Figure ~~8~~9 Acoustic fingerprints for the obstruction type semicircle across all flow rates, with (a) Re ≈36,251, (b) Re ≈83,701, (c) Re ≈126,195 and (d) Re ≈176,182.

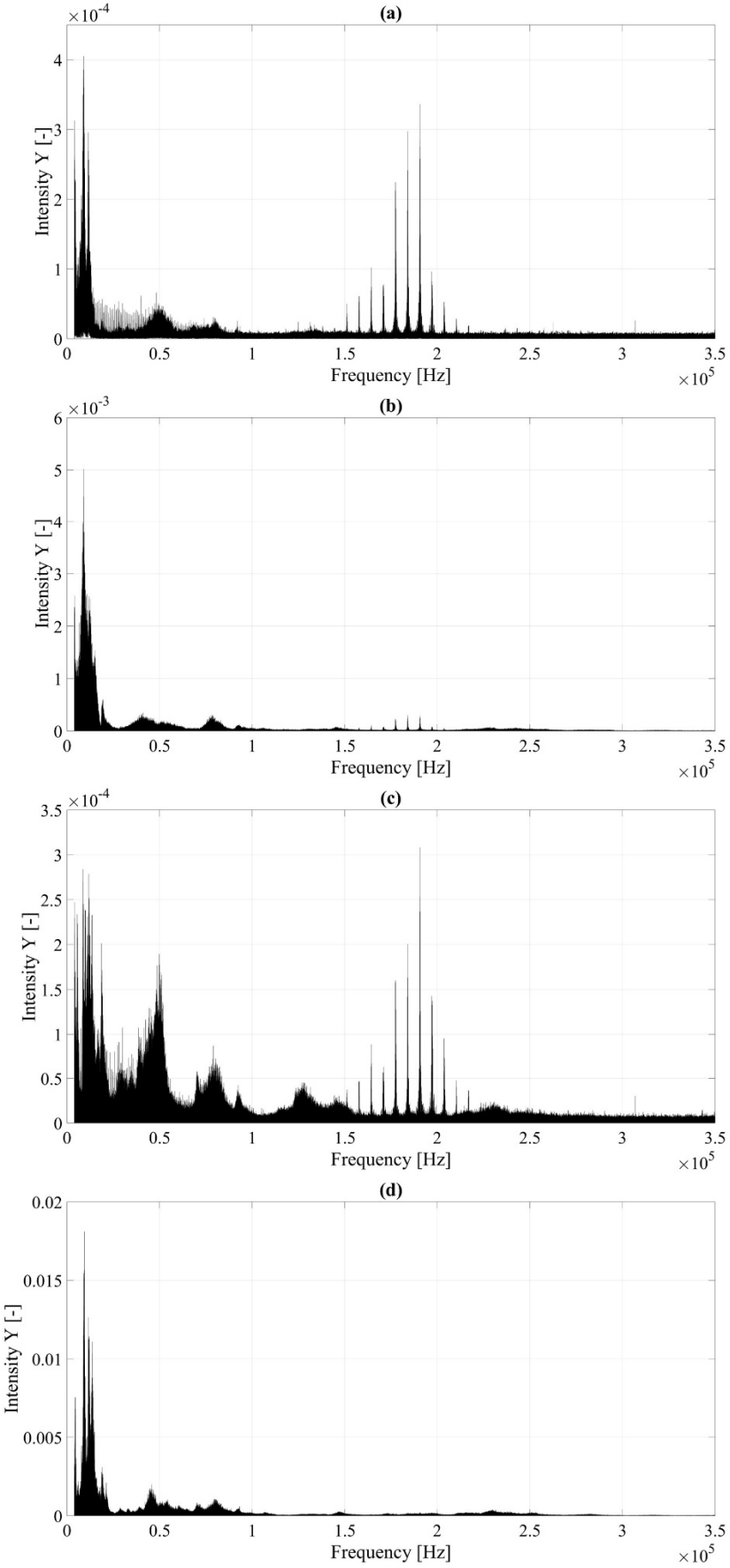


Figure ~~6~~ 10 Acoustic fingerprints for all obstruction types((a) cross, (b) cone, (c) Semicircle, (d) holes and (e) unobstructed) against a fixed flow rate (4,530 lh^{-1}).

Increases in flow rate come with increases in the peak intensity *Y*. In addition, an increase in flow rate leads to the formation of wider mountains and more defined shapes. A peak shift towards the lower frequencies (left side) is characteristic for increases in flow rate (*Figure ~~8~~9*). Towards the side of the higher frequencies (right) a periodic pattern of spike and valley can be observed, that remains throughout flow rate increases.

Figure ~~9~~10 shows that for each obstruction type on a fixed flow rate of 4,530 lh^{-1} unique acoustic "fingerprints" can be retrieved as well. Building this into a database this sensing technique cannot just only be used to detect blockages in a pipe system, but also the type of obstruction or degree of obstruction. This can be used to determine the point of servicing or cleaning the system.

3.5 Supervised Machine Learning

A total of 22 different machine learning algorithms have been tried, with quadratic SVM delivering best prediction accuracies. Quadratic SVM also has shown to be capable to give highly accurate predictions with only a limited number of input data points. The figure below represents a selection of a few of the algorithms tested and their performance against the number of features (Figure ~~10~~11).

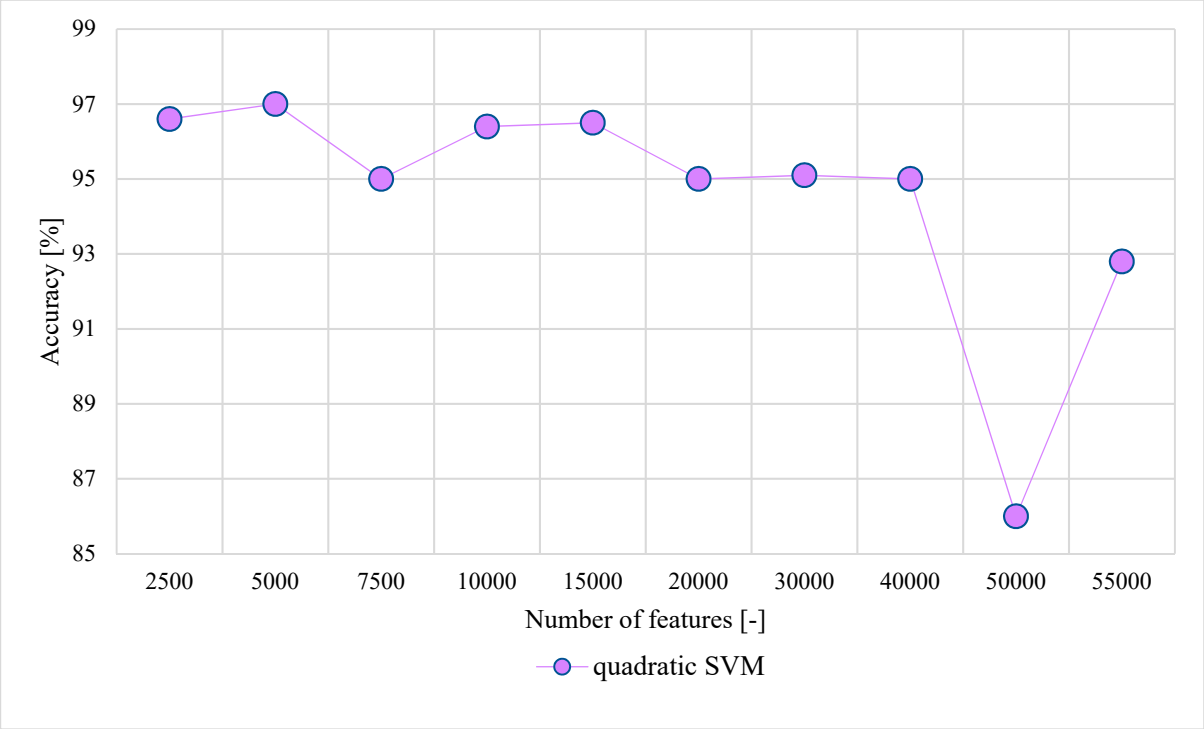


Figure ~~7~~ 11 Performance plot for quadratic SVM learner against the number of features (PCA 15 components enabled).

Amongst all 22 algorithms, Sub Space KNN performed the poorest in terms of prediction accuracies and consistency of predictions. The prediction accuracies for the machine learning on unseen datasets are presented in Table 1:

Table 1 Presentation of Machine Learning predictions on previously unseen data sets.

Case	Prediction Accuracy water run 1	Prediction Accuracy water run 2 (repeat 1)	Prediction Accuracy water run 3 (repeat 2)	Prediction Accuracy water run 4 (repeat 3)	AVG%
Cross vs all Q	95	98	97	97	96.75
Cone vs all Q	96	96	95	98	96.25
Semi vs all Q	94	95	96	96	95.25
Holes vs all Q	99	98	97	97	97.75
Empty vs all Q	95	95	95	96	95.25
Obstruction vs Obstruction	96	96	95	96	95.75

The results show that machine learning is a powerful tool to predict obstructions and flow in single-phased and fully flooded water channels, which delivers instant results. This possibility to use passive acoustic emission on such as system is not reported in literature previously. It shows that results are consistent over several repetitive runs for the same settings. All results are in average in the region of 95% or above. In general a case describes the condition machine learning has been tested upon. For example, the first case "Cross vs Q" means that machine learning has been given the labels for all flow rates on the fixed obstruction type "cross". After training the algorithm machine and having applied PCA, the system will deliver a prediction based on unknown data. In terms of output machine learning provides a confusion matrix where the systems prediction is plotted against its true class. This result describes the accuracy prediction percentage. Each percentage is the output of a test of 100 test sets. This equates to 100,000 individual test data points for 1 individual class (e.g. Q1 and cross).

The table (Table 1) shows that when comparing different obstructions against a fixed flow rate can be confidently distinguished as well as each obstruction against its different flow rate. The same applies when investigating the case of the unobstructed pipe. Machine learning also delivers very consistent predictions when looking into all cases investigated against each other. The range of the poorest performing prediction and the best prediction in its average deviate by only 2.5%. Within runs for the same cases (e.g. obstruction cross against run 1 to 4) the range is in the 3%. Hence, data is repeatable which is of high importance when aiming for an industrial solution to build a surveillance for irrigation systems.

4. Conclusions

A single passive acoustic emission sensor has been placed on a water channel. This channel has been obstructed by different geometries, aiming to represent blockages through build up and a spray nozzle. The different geometries lead to different flow patterns as supported by 2D PIV. When analysing the acoustic signals, higher flow rates lead to higher peak values and different obstructions to distinguishable different FFT spectra. Even marginally different or same pressure drops still deliver distinguishable different acoustic responses. This technology also makes it possible for the first time (in reported literature) to measure such features in a single phased, fully flooded channel. In addition this work shows that the proposed technology is also capable to deliver very different "acoustic fingerprints" when the pressure drop is marginal. This makes it a highly sensitive technology that might be applied in the future to detect upstream blockages at a very early stage.

Highest prediction accuracy levels in supervised ML are attained when applying the quadratic SVM learner. For all cases under investigation accuracy levels are in the region of 95% or above with all experiments having been repeated four times and a hundred buffers for each case.

The combination of the sensor, pipe device and machine learning can be used as an in situ measuring tool for pipe blockages and may be implemented as a prediction tool for water supply system, spray systems etc.

Future research will focus on the refining of the algorithms and create a user interface. Further, scale-up must be considered. Other work shall focus on branched systems and the maximum detection distance of the sensor system. Also, in field applications on irrigation and spraying systems shall be investigated. The application for other materials must also be studied.

Funding Note

This work was supported by the School of Chemical Engineering, University of Birmingham [EPSRC grant number 2112012].

Acknowledgment

The authors wish to express their gratitude for the support received from Philip W Harris and Robert W Sharpe in the engineering workshop at the University of Birmingham. Further thanks goes to Dr Thomas Mills providing necessary training and enabling access to the 3D printing system. For technical support, the authors wish to acknowledge Matteo Antonelli and Giuseppe Forte.

Declaration of interest

None

References

Adrian, R. J. (1986). Image shifting technique to resolve directional ambiguity in double-pulsed velocity. *Applied Optics*, *25*(21), 3855–3858. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18235708

Adrian, R. J. (2005). Twenty years of particle image velocimetry. *Experiments in Fluids*, *39*(2), 159–169. https://doi.org/10.1007/s00348-005-0991-7

Borenstein, J., & Koren, Y. (1988). Obstacle avoidance with ultrasonic sensors. *IEEE Journal on Robotics and Automation*, *4*(2), 213–218. https://doi.org/10.1109/56.2085

Borodina, I. A., Zaitsev, B. D., & Teplykh, A. A. (2018). The influence of viscous and conducting liquid on the characteristics of the slot acoustic wave. *Ultrasonics*, *82*, 39–43. https://doi.org/10.1016/j.ultras.2017.07.011

Boyd, J. W. R., & Varley, J. (2001). The uses of passive measurement of acoustic emissions from chemical engineering processes. *Chemical Engineering Science*, *56*(5), 1749–1767. https://doi.org/10.1016/S0009-2509(00)00540-6

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

Butterfield, Joseph D., Collins, R. P., Krynkin, A., & Beck, S. B. M. (2017). 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

Chang, T.-Y., Liu, C.-S., Huang, K.-H., Chen, R.-Y., Lai, J.-S., & Bao, B.-Y. (2011). High-frequency hearing loss, occupational noise exposure and hypertension: a cross-sectional study in male workers. *Environmental Health*, *10*(1), 35. https://doi.org/10.1186/1476-069X-10-35

Chatzigeorgiou, D., Youcef-Toumi, K., & Ben-Mansour, R. (2015). Design of a Novel In-Pipe Reliable Leak Detector. *IEEE/ASME Transactions on Mechatronics*, *20*(2), 824–833. https://doi.org/10.1109/TMECH.2014.2308145

Chong, M., Abraham, A., & Paprzycki, M. (2004). Traffic accident data mining using machine learning paradigms. *Fourth International Conference on Intelligent Systems Design and Applications (ISDA'04), Hungary*, 415–420.

Dagher, I. (2008). Quadratic kernel-free non-linear support vector machine. *Journal of Global Optimization*, *41*(1), 15–30. https://doi.org/10.1007/s10898-007-9162-0

Das, S., Das, R., & Mazumdar, A. (2013). Circulation characteristics of horseshoe vortex in scour region around circular piers. *Water Science and Engineering*, *6*(1), 59–77. https://doi.org/https://doi.org/10.3882/j.issn.1674-2370.2013.01.005

Finfer, D., Parker, T. R., Mahue, V., Amir, M., Farhadiroushan, M., & Shatalin, S. (2015). Non-intrusive Multiple Zone Distributed Acoustic Sensor Flow Metering. *SPE Annual Technical Conference and Exhibition*, 1–9. https://doi.org/10.2118/174916-MS

Fischer, P. (1999). *Algorithmisches Lernen* (1st ed.). https://doi.org/10.1007/978-3-663-11956-2

449 Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of*
450 *Eugenics*, *7*(2), 179–188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x

451 Forrest, T. G. (1994). From Sender to Receiver: Propagation and Environmental Effects on
452 Acoustic Signals. *American Zoologist*, *34*(6), 644–654.
453 https://doi.org/10.1093/icb/34.6.644

454 Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data Mining and Analytics in the Process
455 Industry: The Role of Machine Learning. *IEEE Access*, *5*, 20590–20616.
456 https://doi.org/10.1109/ACCESS.2017.2756872

457 Gleick, P. H. (2014). Water, Drought, Climate Change, and Conflict in Syria. *Weather,*
458 *Climate, and Society*, *6*(3), 331–340. https://doi.org/10.1175/WCAS-D-13-00059.1

459 Guan, R., Wang, X., Marchese, M., Yang, M. Q., Liang, Y., & Yang, C. (2019). Feature
460 space learning model. *Journal of Ambient Intelligence and Humanized Computing*,
461 *10*(5), 2029–2040. https://doi.org/10.1007/s12652-018-0805-4

462 Hauptmann, P., Hoppe, N., & Püttmer, A. (2002). Application of ultrasonic sensors in the
463 process industry. *Measurement Science and Technology*, *13*(8), 201.
464 https://doi.org/10.1088/0957-0233/13/8/201

465 Hou, R., Hunt, A., & Williams, R. . (1999). Acoustic monitoring of pipeline flows:
466 particulate slurries. *Powder Technology*, *106*(1–2), 30–36.
467 https://doi.org/10.1016/S0032-5910(99)00051-0

468 Hunaidi, O., & Chu, W. T. (1999). Acoustical characteristics of leak signals in plastic water
469 distribution pipes. *Applied Acoustics*, *58*(3), 235–254. https://doi.org/10.1016/S0003-
470 682X(99)00013-4

471 Khalifa, A. E., Chatzigeorgiou, D. M., Youcef-Toumi, K., Khulief, Y. A., & Ben-Mansour,
472 R. (2010). Quantifying Acoustic and Pressure Sensing for In-Pipe Leak Detection.
473 *Volume 13: Sound, Vibration and Design*, 489–495.
474 https://doi.org/10.1115/IMECE2010-40056

475 Khulief, Y. A., & Khalifa, A. (2012). On the In-Pipe Measurements of Acoustic Signature of
476 Leaks in Water Pipelines. *Volume 12: Vibration, Acoustics and Wave Propagation*, 429.
477 https://doi.org/10.1115/IMECE2012-86731

478 Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2007). *Supervised Machine Learning : A*
479 *Review of Classification Techniques General Issues of Supervised Learning Algorithms*.
480 *31*, 501–520.

481 Kotzé, R., Ricci, S., Birkhofer, B., & Wiklund, J. (2016). Performance tests of a new non-
482 invasive sensor unit and ultrasound electronics. *Flow Measurement and Instrumentation*,
483 *48*, 104–111. https://doi.org/10.1016/j.flowmeasinst.2015.08.013

484 Koza, J. R., Bennett, F. H., Andre, D., & Keane, M. A. (1996). Automated Design of Both
485 the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. In
486 *Artificial Intelligence in Design '96* (pp. 151–170). https://doi.org/10.1007/978-94-009-
487 0279-4_9

488 Kubat, M. (2015). *An Introduction to Machine Learning*. https://doi.org/10.1007/978-3-319-
489 20010-1

490 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, C., & Wang, W. (2017). Contextual Regression: An Accurate and Conveniently Interpretable Nonlinear Model for Mining Discovery from Scientific Data. *BioRxiv*. https://doi.org/10.1101/210997

Lohmar, B., Wang, J., Rozelle, S., Huang, J., & Dawe, D. (2003). *China's Agricultural Water Policy Reforms: Increasing Investment, Resolving Conflicts, and Revising Incentives* (p. 27). p. 27.

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. (2016). Leak Detection in Water-Filled Small-Diameter Polyethylene Pipes by Means of Acoustic Emission Measurements. *Applied Sciences*, *7*(1), 2. https://doi.org/10.3390/app7010002

Masud, M. M., Chen, Q., Gao, J., Khan, L., Han, J., & Thuraisingham, B. (2010). Classification and Novel Class Detection of Data Streams in a Dynamic Feature Space. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (6322nd ed., pp. 337–352). https://doi.org/10.1007/978-3-642-15883-4_22

Moraru, A., Pesko, M., Porcius, M., Fortuna, C., & Mladenic, D. (2010). Using Machine Learning on Sensor Data. *Journal of Computing and Information Technology*, *18*(4), 341. https://doi.org/10.2498/cit.1001913

Noble, W. S., & Street, P. (2006). *What is a support vector machine ? 24*(12), 1565–1567.

O'Keefe, C. V., Maron, R., Felix, J., van der Spek, A. M., & Rothman, P. (2010). Non-invasive passive array technology for improved flow measurements of slurries and entrained air. In CiDRA Holdings (Ed.), *The 4th International Platinum Conference: Platinum in transition 'Boom or Bust'* (pp. 21–30). Retrieved from https://www.saimm.co.za/Conferences/Pt2010/021-030_OKeefe.pdf

Pavić, G. (1992). Vibroacoustical energy flow through straight pipes. *Journal of Sound and Vibration*, *154*(3), 411–429. https://doi.org/10.1016/0022-460X(92)90776-T

Rahman, M., Håkansson, U., & Wiklund, J. (2015). In-line rheological measurements of cement grouts: Effects of water/cement ratio and hydration. *Tunnelling and Underground Space Technology*, *45*, 34–42. https://doi.org/10.1016/j.tust.2014.09.003

Reitmaier, T. (2015). *Aktives Lernen für Klassifikationsprobleme unter der Nutzung von Strukturinformationen* (Band 5 von). Kassel: Kassel University Press GmbH.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386–408. https://doi.org/10.1037/h0042519

Sause, M. G. R., Gribov, A., Unwin, A. R., & Horn, S. (2012). Pattern recognition approach to identify natural clusters of acoustic emission signals. *Pattern Recognition Letters*, *33*(1), 17–23. https://doi.org/10.1016/j.patrec.2011.09.018

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

The Economist Newspaper Limited. (2016, September). Precision agriculture: TV dinners. *The Economist*, 3.

Thenkabail, P. S., Hanjra, M. A., Dheeravath, V., & Gumma, M. (2017). Global Croplands and Their Water Use from Remote Sensing and Nonremote Sensing Perspectives. In Q. Weng (Ed.), *Advances in Environmental Remote Sensing: Sensors, Algorithms, and Applications* (1st ed., pp. 383–421). https://doi.org/10.1201/b10599

Treszkai, Á. (2018). General Overview of the Consequences of African Water Conflicts. *UKH Journal of Social Sciences*, *2*(2), 11–18. https://doi.org/10.25079/ukhjss.v2n2y2018.pp11-18

Vapnik, V. N., & Chervonenkis, A. Y. (1974). *Teoriya raspoznavaniya obrazov: Statisticheskie problemy obucheniya [Theory of pattern recognition: Statistical problems of learning]*. Moscow.

Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research*, *4*(1), 23–45. https://doi.org/10.1080/21693277.2016.1192517