Flow Measurement and Instrumentation, Volume 68, August 2019, Article number 101568
DOI: 10.1016/j.flowmeasinst.2019.05.002

Identification of Gas-Liquid Flow Regimes Using a Non-intrusive Doppler Ultrasonic Sensor and Virtual Flow Regime Maps

Somtochukwu Godfrey Nnabuife[a], Karl Ezra S. Pilario[a,b], Liyun Lao[a,*], Yi Cao[c], Mahmood Shafiee[a]
[a] Energy and Power, Cranfield University, Cranfield, United Kingdom
[b] Department of Chemical Engineering, University of the Philippines Diliman, Republic of the Philippines
[c] College of Chemical and Biological Engineering, Zhejiang University, People's Republic of China
*Corresponding author, email address: l.lao@cranfield.ac.uk

Abstract

The accurate prediction of flow regimes is vital for the analysis of behaviour and operation of gas/liquid two-phase systems in industrial processes. This paper investigates the feasibility of a non-radioactive and non-intrusive method for the objective identification of two-phase gas/liquid flow regimes using a Doppler ultrasonic sensor and machine learning approaches. The experimental data is acquired from a 16.2-m long S-shaped riser, connected to a 40-m horizontal pipe with an internal diameter of 50.4 mm. The tests cover the bubbly, slug, churn and annular flow regimes. The power spectral density (PSD) method is applied to the flow modulated ultrasound signals in order to extract frequency-domain features of the two-phase flow. Principal Component Analysis (PCA) is then used to reduce the dimensionality of the data so as to enable visualisation in the form of a virtual flow regime map. Finally, a support vector machine (SVM) is deployed to develop an objective classifier in the reduced space. The classifier attained 85.7% accuracy on training samples and 84.6% accuracy on test samples. Our approach has shown the success of the ultrasound sensor, PCA-SVM, and virtual flow regime maps for objective two-phase flow regime classification on pipeline-riser systems, which is beneficial to operators in industrial practice. The use of a non-radioactive and non-intrusive sensor also makes it more favorable than other existing techniques.

Keywords: Doppler ultrasound, Support vector machine (SVM), Probability density function (PDF), Principal component analysis (PCA), S-shaped riser

1 Introduction

Two-phase gas-liquid flow is encountered frequently in industrial operations such as nuclear power plant steam generators, boilers, chemical reactors and petroleum transportation (Julia and Hibiki, 2011). The different types of interfacial structures between different phases of fluids, known as *multiphase flow regimes*, can be geometrically complex and varying. The flow can be steady or unsteady, turbulent or laminar, gas/liquid segregated or mixed. Gas can flow within the liquid as bubbles or liquid can flow within the gas as droplets (Falcone et al., 2002).

The governing flow regime is influenced by many parameters such as gas/liquid superficial velocities, gas/liquid densities, gas/liquid surface tension, gas/liquid viscosities, pipe diameter,

and pipe inclination (Thorn et al., 2012). Traditionally, *flow regime maps* are used to illustrate the dependency of the flow regime on two quantities, which are usually the superficial gas and liquid flow rates (Falcone et al., 2009). Yet, characterising and measuring two-phase flow is still challenging due to its inherently complex nature. Thus, the problem of flow regime identification remains relevant.

Flow regime identification methods can either be subjective techniques (direct observation) or objective techniques (scientific or indirect determination) (Rouhani and Sohal, 1983). Subjective or direct techniques involve the operator visually interpreting an image of the flow to classify it into a flow regime. Objective or indirect determination is a two-part process. The operator must first utilise a suitable experimental methodology to measure flow parameter features correctly and then analyse the flow features objectively to categorise the flow regime (Juliá et al., 2008).

Currently, gas/liquid two-phase flow regime identification is mainly accomplished by subjective means such as direct visual observation and via cameras (Peddu et al., 2017). Hence, the accurate classification of flow regimes is yet to be standardised, and it mostly depends upon the interpretation of individual visual views, which can lead to inconsistency in flow regime identification due to human subjectivity. The main drawback of visual observations is that the pictures are often confusing and challenging to interpret, in particular when handling high flow velocities even with high-speed cameras. Moreover, flow channels are often opaque, so flow identification by visual means is impossible (Barnea et al., 1980). Although there are numerous flow regime identification approaches already studied for two-phase gas/liquid flow, industrial acceptance remains challenging. Subjective techniques cannot facilitate industrial automation where many important decisions depend on the governing flow regime.

Significant efforts have already been made to develop flow regime identification using objective methodologies. Several research studies have used a phase distribution measurement approach. One of these methods is the use of invasive-point sensors such as pitot tubes, fibre-optic or

electrical probes and hot-wire anemometers (Barnea, 1987). The major drawback to these methods is that the sensors disturb the flow fields during the measurement of void or pressure fluctuations (Dyakowski, 1996). Hence, non-invasive means must be deployed to differentiate the boundaries between diverse flow regimes.

Objective flow regime identification using a clamped-on, non-invasive sensor is of great interest in many industries. Non-invasive methods are highly attractive as they eliminate the need for immersion of instrumentation in the flow. Jones and Delhaye (1976) investigated and summarised different measuring methods applied to a two-phase flow of which few are employed directly to characterise the flow regime. For instance, Barnea et al. (1980) used an enhanced electrical conductance probe in two-phase near horizontal, horizontal and vertical flows to identify flow regime.

Among the non-invasive sensors, radiation attenuation methods are more widely used in many industrial applications due to their reliability. Jones and Zuber (1975) studied an X-ray void measurement system for vertical two-phase flow in a rectangular channel; Salgado et al. (2010) achieved flow regime identification using gamma-ray pulse height distributions (PHDS) and artificial neural networks (ANNs); Blaney and Yeung (2008) analysed probability distributions using a self-organising feature map and gamma densitometer data for multiphase flow regime identification; Sunde et al. (2005) proposed an enhanced method, which compares the visualisation of the intensity of gamma ray measurements at every flow condition. Generally, radiation attenuation methods based on gamma rays, X-rays and neutrons are already established online measurement systems. When compared with each other, the gamma densitometer has merits, such as high penetration and cost-effectiveness (Chaouki et al., 1997). However, the major drawback to these methods is their radioactive nature, which is hazardous. The need to increase the gamma source strength with an increase in density or the pipe wall thickness requires increased radiation protection and hence minimises its portability.

Chakraborty et al. (2009) presented a novel ultrasonic method for two-phase flow void fraction measurement using an ultrasonic sensor and two signal processing techniques established on the time series analysis approach: the logical signal space partitioning and symbolic filtering. Although the theory on symbolic dynamic filtering was established, identification using pulse-echo mode is not a full classification method of flow regimes, but rather of flow patterns (Jha et al., 2012). It was noted that more research needs to be carried out on experimental and computational work before applying the method in the industry. Another drawback to this method is that the set-up is invasive even though the ultrasonic method itself is non-intrusive. As an extension of the work by Chakraborty et al., (2009), Jha et al. (2012) presented the concept of implementing ultrasonic pulse echoes in a clamped-on set-up in connection with symbolic dynamic filtering for deployment in the industries.

Regardless of the prospect of using ultrasonic pulse-echo for flow regime identification, the method is based on computational models. The computational models apply a set of non-linear equations which are frequently simplified for flow regime identification. In practice, the simplified equations are difficult to implement since the knowledge of various flow parameters is required, such as pipe thickness and pipe diameter. The accuracy of these equations is also compromised when flow parameters deteriorate with time (Meribout et al., 2010). In addition, the ultrasound pulse-echo method is limited by the maximum velocity that it can measure due to the Nyquist criterion (Evans and McDicken, 2000).

Doppler ultrasonic sensors can also achieve non-invasive flow velocity measurement. This technique is ubiquitous in the medical field. The method utilises the shift in frequency due to flow velocities to predict the flow regime (Übeyli and Güler, 2005). The applicability of continuous wave ultrasonic Doppler (CWUD) in two-phase flow velocity measurement was investigated by Kouam et al. (2003). They suggested the use of frequency resolution methods to resolve the issue

of the presence of coloured noise in velocity measurement, which otherwise poses a severe problem to the classical frequency estimators.

In this paper, a non-intrusive and non-radioactive method for the objective identification of two-phase gas/liquid flow regime is proposed using output signals from a commercial CWUD flow metering device, and the machine learning (ML) approaches. The method is based on the assumption that the gas liquid flow patterns will have their unique signatures on the output signals of the CWUD device when subjected to the two-phase flows. On the other hand ML solutions to objective flow regime identification have already been proposed, such as (Xie et al., 2004; Hanus et al., 2017; Wang and Zhang, 2009; Trafalis et al., 2005). In this work, to better facilitate the applicability to industrial practice, principal components analysis (PCA) is used to visualise the information from intrinsic flow regime features in 2-dimensional space. To this end, a mapping is created so that in 2-dimensions, the mapped samples can be found clustered according to their respective flow regimes. Support Vector Machine (SVM) is then applied to the samples in the 2-dimensional space to create boundaries between the clusters. This leads to a *virtual flow regime map* that serves as a visual aid to human operators for objective flow regime identification (Eyo et al., 2019). In summary, the main contributions of this paper are as follows: (i) we explore the feasibility of visualizing frequency-domain features from ultrasonic Doppler signals in a 2D virtual flow regime map; and (ii) we make the first known effort towards the applicability of continuous wave Doppler ultrasound and the SVM to objectively identify flow regime in an S-shaped riser. By using safer and more advanced techniques for two-phase flow measurement and instrumentation, industries can enhance production, achieve better process performance, and hence, have economic advantages.

This paper is organised as follows: Section 2 presents the sensor principle and the algorithm for CWUD. In Section 3, the experimental method used in this study is described. Signal analysis

using ML approaches are also discussed. In Section 4, the results and discussion of the analysed data are presented. Finally, conclusions and future work are given in Section 5.

2 Measurement Sensor and Algorithm

The Doppler shift (or Doppler Effect) is the frequency variation of an acoustic wave when movement exists between the acoustic receiver and the source, where the change in frequency is proportional to the acoustic source velocity (Weinstein, 1982). Thus, the velocity of the acoustic source is obtained by calculating the frequency shift between the acoustic receiver and the source (see Fig. 1(b)). In the ultrasonic Doppler flowmeter, illustrated in Fig. 1(a), a fixed-frequency acoustic beam is released continuously from the transducer into the flow. The beam is then reflected by the moving scatterers in the fluid, which could be bubbles in the flow (Chivers and Hill, 1975). Another ultrasonic transducer receives the scattered acoustic beam so that the velocity of the fluid can be estimated with the frequency shift based on the Doppler Effect. The mathematical description could be found elsewhere, like e.g Evans et al (1989), but is included here for completeness.

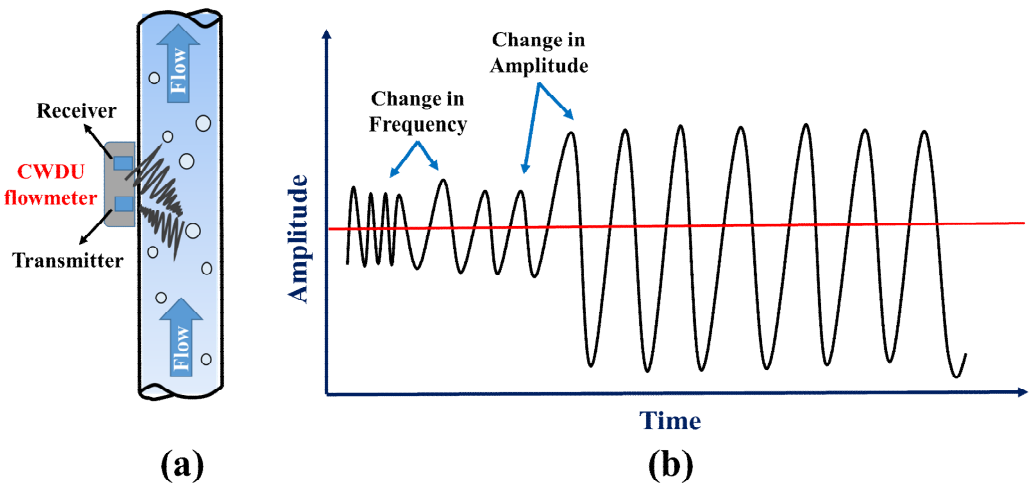

Figure 1: Ultrasound Doppler principle (Meire and Farrant, 1995)

First, assume that the signal transmitted is

$$x_t(t) = \varepsilon_t \cos(\omega_t t) \tag{1}$$

and that the corresponding received signal from one of the scatterers is

$$x_r(t) = \varepsilon_r \cos(\{\omega_t + \omega_d\}t + \theta_1) \tag{2}$$

where $\omega_t = 2\pi f_t$ is the angular frequency of the transmitted signal, $\omega_d = 2\pi f_d$ is the amount of shift in the angular frequency, and θ_1 is the phase shift based on the scatterer distance between the receiver and the transducer (Evans et al., 1989).

Multiplying the two signals electronically results in:

$$x_t(t)x_r(t) = \varepsilon_t \varepsilon_r \cos(\omega_t t) \cos(\{\omega_t + \omega_d\}t + \theta_1) \tag{3}$$

$$x_t(t)x_r(t) = \frac{\varepsilon_t \varepsilon_r}{2} [\cos(\omega_d t + \theta_1) + \cos(\{2\omega_t + \omega_d\}t + \theta_1)]. \tag{4}$$

The resulting signal is then low-pass filtered to remove the $2f_t$ source frequency, leaving only the desired Doppler signal (Evans et al., 1989):

$$x_d(t) = \frac{\varepsilon_t \varepsilon_r}{2} \cos(\omega_d t + \theta_1). \tag{5}$$

Additional signal processing may be needed since the received signal has reflected an ultrasound of amplitude greater than the signal backscattered from the moving scatterers. Finally, the relationship between the Doppler shift f_d and the velocity of the scatterer can be described as follows (Sanderson and Yeung, 2002):

$$f_d = 2f_t \frac{v}{c} \cos\theta \tag{6}$$

where f_d is the Doppler frequency shift, f_t is the transmitted ultrasound frequency, v is the flow velocity average, and θ is the angle between the flow velocity and the ultrasound beam.

The continuous-wave ultrasonic Doppler (CWUD) used in this work is a DFM-2, a commercial non-invasive flowmeter developed by United Automation Ltd, Southport, U.K., using 500 kHz ultrasonic transducers. The flow meter's output signal ranges 0-5 volt for flow velocities 0 – 6.1 m/s. In this study the analogue signal was sampled by using NI PCI-6040E, a NI data acquisition system. To achieve a suitable bond between the external conduit surface and the sensor, a glycerine gel was applied to avoid air cavities trapped between the sensor and the conduit surface.

For this study, it is worth to note that the CWUD device often give "velocity" readings fluctuating from negative to positive values. Those readings do not necessarily reflect the true velocities of the gas-liquid flows. Instead, they suggest a very complex interfacial structure in the two-phase flow.

3 Test Rig and Experimental Procedure

3.1 Two-phase flow test rig set-up

The experiment was carried out on a 2-inch S-shaped riser of the three-phase flow loop at Cranfield University oil and gas centre. The 2-inch flow loop is made up of a 40-m horizontal pipeline, 5.5-m vertical lower section, 1.5-m down-comer, 5.7-m vertical upper section and 3.5-m topside section. This test rig is operated using the DeltaV (Fieldbus based supervisory, control and data acquisition) software provided by Emerson Process Management. The schematic diagram of the test rig is presented in Figure 2. The air used was supplied from a bank of two compressors connected in parallel. When both compressors are run in parallel, a maximum air flow rate of 1410 m^3/hr FAD at 7 bar can be supplied. The air from the two compressors accumulates in an 8-m^3 capacity receiver to reduce the pressure fluctuation from the compressor. Air from the receiver passes through a bank of three filters (coarse, medium, fine) and then through a cooler where debris and condensates present in the air are stripped from the air before it enters the flow meters. The water flow rate was supplied from a 12.5-m^3 capacity water tank. The water was supplied to the flow loop by two multistage Grundfos CR90-5 pumps. The water pump has a duty of 100 m^3/hr at 10 bar. The speed control is achieved using frequency variable

inverters. The water pumps are operated remotely using DeltaV, a fieldbus based supervisory, control, and data acquisition software (SCADA). The water flow rate was metered by a 1-inch Rosemount 8742 magnetic flow meter (up to 7.36 l/s) and 3-inch Foxboro CFT50 Coriolis meter (up to 30 kg/s).

After the experiment, air and water were separated in an 11.12-m^3 horizontal three-phase gravity separator. After the separation in the three-phase separator and cleaning, the air was exhausted into the atmosphere while water from the three-phase separator entered a 1.6-m^3 coalescer, where the water is further cleaned before returning to the storage tank.

The 2-inch S-shaped flow loop test facility used in this experiment has a 54.8-mm internal diameter, 40-m length and 1.5-m downcomer. The 2-inch S-shaped flow loop test section has a transparent pipe for flow regime observation. The air flow-rate was adjusted by controlling the valves through the DeltaV to achieve the desired flow regime.

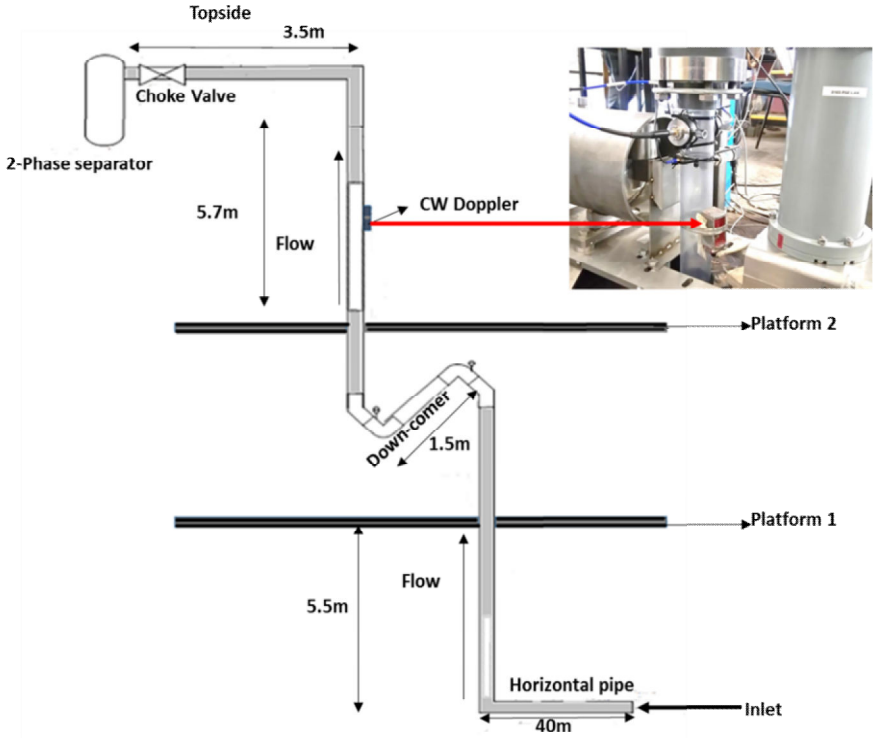

Figure 2: Schematic diagram of S-shape rig

A clamp-on non-intrusive CWUD transducer with an excitation voltage of ±10V, operating at a carrier frequency of 500 kHz was attached to the top-side of the S-shaped riser as illustrated in Fig. 2. The ultrasound beam incident angle was 58° with respect to flow direction on the S-shaped riser. It is essential to place the ultrasonic sensor on the pipe at least 10 diameters away from tees, valves, and bends to prevent measurement errors from cavitation, swirls and turbulent eddies. A gel coupling agent was applied between the pipe wall and the Doppler transducer to make the ultrasound energy transmission easier. The electronics of the CWUD flow meter was adapted to record the voltage signals converted from the Doppler frequency shifts, for further analysis (see Fig. 3).

Ultimately, the process variable being measured by the ultrasound Doppler is average flow velocity. Based on the pipe scale and flow velocity range, it was estimated that the value of the flow velocity fluctuates at a frequency no more than 2 kHz. Hence, in the LabVIEW data

acquisition system, a sampling frequency of 10 kHz is appropriate with respect to the Nyquist criterion, since this is five times the estimated upper limit frequency of the flow velocity fluctuations.

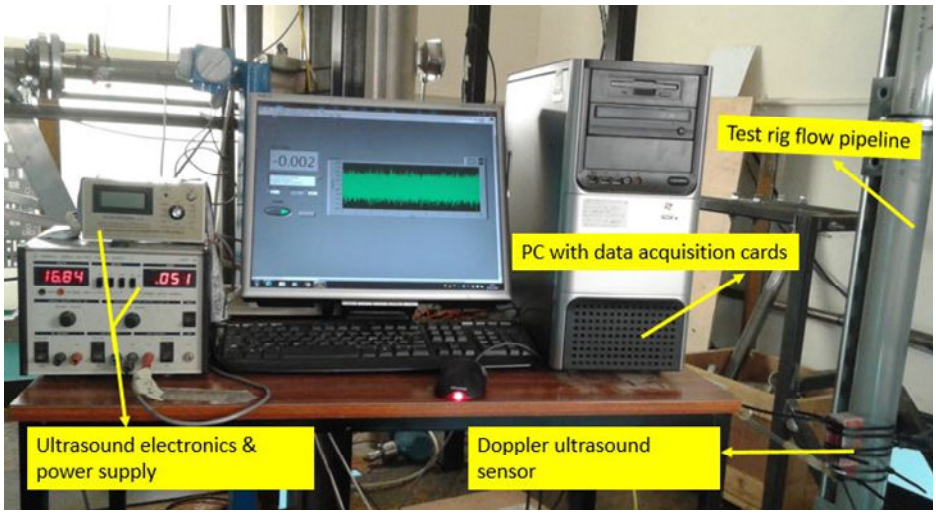

Figure 3: Doppler ultrasonic sensor and its auxiliary instruments

3.2 Flow regime classification methodology

3.2.1 Feature extraction from ultrasonic Doppler signals

Feature extraction is the most crucial step for any flow regime identification method. This step aims to find any information from the measurement data that can be used to best distinguish among flow regimes. For high-frequency data, such as the ultrasonic Doppler signals, features can be extracted either from the time domain or frequency domain. In this work, the widely used frequency-domain power spectral density (PSD) features are adopted (de Kerret et al., 2017).

Given a stationary discrete-time signal $x(n)$, the power spectral density function $P_x(f)$ of this signal is defined as the Fourier transform of the autocorrelation sequence $R_x(k)$ (Xie et al., 2004):

$$P_x(f) = \sum_{k=-\infty}^{\infty} R_x(k) \exp\left(-2\pi i k \frac{f}{f_s}\right) \tag{7}$$

where f_s is the sampling frequency. Since the signal is only measured on a finite interval $[0, \dots, N-1]$, Welch's method is adopted to obtain the PSD, which is given as

$$\hat{P}_x(f) = \sum_{k=-N+1}^{N-1} \hat{R}_x(k) \exp\left(-2\pi i k \frac{f}{f_s}\right) \tag{8}$$

where the autocorrelation is (Xie et al., 2004):

$$\hat{R}_x(k) = \frac{1}{N} \sum_{n=0}^{N-1-k} x(n+k)x(n). \tag{9}$$

Using Welch's method in the same way Abbagoni and Yeung (2016), the PSD features were analysed from each sample of ultrasonic Doppler signals at various gas-liquid flow rates as presented in Fig. 4. A total of 130 data samples of different superficial gas and liquid velocities were recorded. Different flow regime labels were assigned to each data sample by visual observation, and ambient temperature conditions were recorded at the same time. Each data sample acquired consists of Doppler frequency shift signals recorded for a period of 900s. The data set was sub-divided into 70% for training (91 samples) and 30% for testing (39 samples). With a sampling frequency of 10 kHz, a Hanning window with a length 1,024 and a 75% overlap were used in the Welch method.

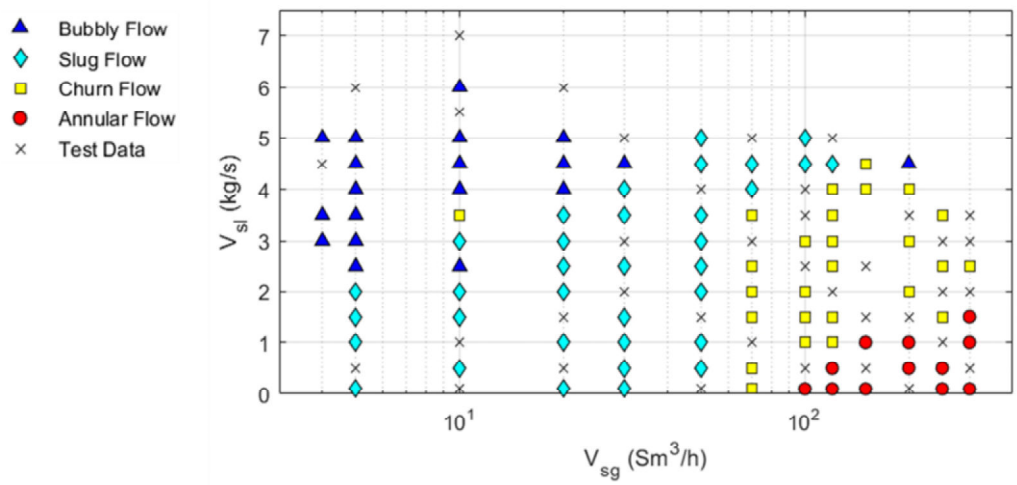

Figure 4: Gas and liquid flow rates of all samples from the experiment

Typical power spectral estimates from each flow regime are presented in Fig. 5. The relevant frequency spectrum ranges from 0 to 1200 Hz. In this range, the PSD spectrum is distinct in each flow regime. To obtain the actual features that can distinguish between the flow regimes, bands of a length of 120 Hz were taken from the power spectrum, and, following Abbagoni and Yeung, (2016) the mean PSD was computed on each band. Also, the maximum peak of the PSD, the mean weighted frequency $\bar{f}$ of the spectral power, and the variance of the spectral power equation σ_f^2 were computed for each sample. The last two are computed as

$$\bar{f} = \frac{\sum_i f_i P_x(f_i)}{\sum_i P_x(f_i)} \tag{10}$$

$$\sigma_f^2 = \frac{\sum_i \left(f_i - \bar{f}\right)^2 P_x(f_i)}{\sum_i P_x(f_i)}. \tag{11}$$

In total, 13 features are obtained from the power spectrum of ultrasound signals: the mean PSD for each of the 10 frequency bands, the maximum peak of the PSD, $\bar{f}$, and σ_f^2. This approach is commonly used to distinguish each flow regime using features in the frequency-domain (Abbagoni and Yeung, 2016; Drahoš and Čermák, 1989). Our work takes the further step of taking

these features and visualising them in 2-dimensional space, before the flow regime is classified by an efficient pattern recognition technique.

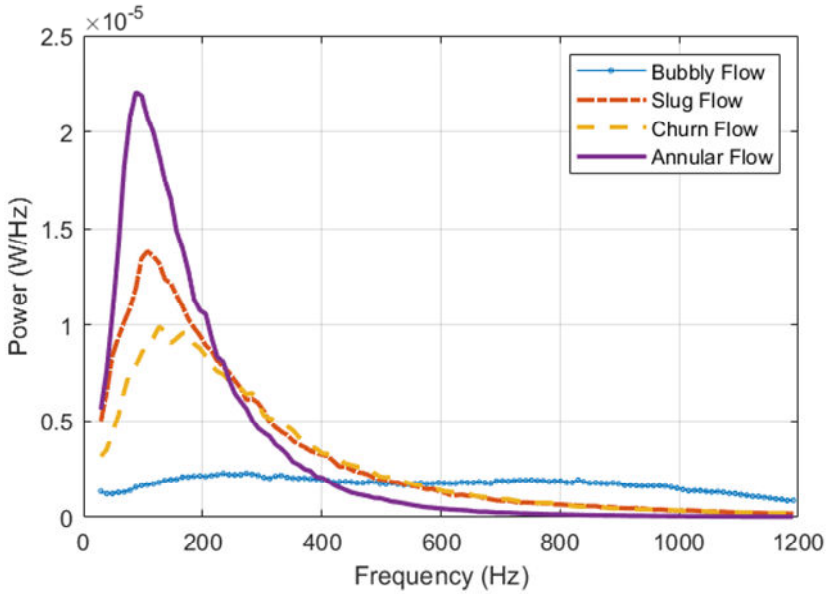

Figure 5: Typical power spectra of each flow regime

3.2.2 Dimensionality Reduction for Visualization

In unsupervised machine learning, dimensionality reduction is a family of methods used to express the same information from a high-dimensional data set using only a few dimensions. In the previous subsection, the information as expressed in 13 features (dimensions) was taken from ultrasound Doppler signals for flow regime identification. Here, the same information is to be retained using only two dimensions by performing a dimensionality reduction method. Since this step is unsupervised, the information about the flow regime labels of the samples is not used yet. Nonetheless, the benefit of reducing the data to two dimensions is the ability to *visualise* the information in a 2-dimensional space. This leads to the realisation of a *virtual flow regime map* completely from ultrasound Doppler data.

In our work, principal components analysis (PCA) is used for linear dimensionality reduction, which is by far the most popular (Van Der Maaten et al., 2009). In PCA, the information is

retained in a set of latent variables that are linear combinations of the original set of features. The PCA algorithm is outlined as follows.

Given an M-dimensional data set of N samples, $\boldsymbol{x}_i \in \Re^M, i = 1,2,\dots N$, PCA proceeds by first normalising the data to zero mean and unit variance, yielding $\bar{\boldsymbol{X}} \in \Re^{M\times N}$. The sample covariance matrix of this data set is computed as

$$\boldsymbol{\Sigma_{xx}} = \frac{1}{N-1}\bar{\boldsymbol{X}}^T\bar{\boldsymbol{X}} \in \Re^{N\times N}. \tag{12}$$

The eigenvalue decomposition of the covariance matrix can be written as

$$\boldsymbol{\Sigma_{xx}} = \boldsymbol{V\Lambda V^T} \tag{13}$$

where $\boldsymbol{V} \in \Re^{N\times N}$ is the matrix of eigenvectors on each column and $\boldsymbol{\Lambda} \in \Re^{\mathrm{N}\times\mathrm{N}}$ is the diagonal matrix of decreasing eigenvalues. The columns of matrix $\boldsymbol{V}$ represent the principal directions that successively explain the maximum variance in the data, while the eigenvalues in $\boldsymbol{\Lambda}$ are scaling factors equivalent to the data variance values themselves. The projection matrix is given by:

$$\boldsymbol{P} = \boldsymbol{V\Lambda^{-1/2}} \in \boldsymbol{\Re^{N\times N}}. \tag{14}$$

By using only the first two columns of $\boldsymbol{P}$, denoted as matrix $\boldsymbol{P}_2$, the dimensionality of the data is reduced to two while preserving as much information possible. The projections are then applied to the covariance matrix to obtain latent variables $\boldsymbol{L}$ as

$$\boldsymbol{L} = \boldsymbol{P_2^T\Sigma_{xx}} \in \Re^{2\times N}. \tag{15}$$

After the application of Eq. (15), each training sample is now represented by every column of $\boldsymbol{L}$, which has two features that can be plotted in a 2-dimensional space. A machine learning technique for classification can then be used to create decision boundaries objectively between the samples in 2-dimensional space.

3.2.3 Support Vector Machine for Classification

This paper proposes the use of a support vector machine (SVM) for objectively classifying flow regimes in an S-shape riser using 2-dimensional features from the ultrasonic Doppler data.

Cortes and Vapnik, (1995) originally proposed the SVM for binary classification. Given N data samples of features $\boldsymbol{x}_i \in \Re^M$ each belonging to either of two classes, labelled $y_i \in \{+1, -1\}$, the aim of binary classification is to learn a mapping function that can be used to predict the unknown class of a new sample. SVM solves this by searching for a *linear* separating hyperplane in the M-dimensional feature space that maximises the margin of separation between samples from each opposing class. This separating hyperplane can then serve as a decision boundary between classes. To achieve nonlinear separations, *kernels* can be used to first transform the original feature space using nonlinear projections, prior to seeking the separating hyperplane (Cristianini and Shawe-Taylor, 2014). The idea of "maximum margin of separation" is the logic offered by the SVM approach, which replaces the human subjectivity in flow regime classification. Hence, using SVM, an objective flow regime classifier can be developed.

More specifically, the dual formulation of kernel binary SVM classification is posed as the following convex quadratic programming problem (Cristianini and Shawe-Taylor, 2014):

$$\max \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j K(\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{16}$$

$$\text{subject to} \quad \sum_{i=1}^{N} y_i \alpha_i = 0,$$
$$0 \leq \alpha_i \leq C, \quad i = 1, \ldots, N$$

where $\boldsymbol{x}_i \in \Re^M, i = 1,2, \ldots, N$ is the N training samples with M features, α_i are Lagrange multipliers, $K(\cdot\,,\cdot)$ is a kernel function, $y_i \in \{+1, -1\}$ are the known labels for each sample, that is positive or negative, and C is a regularisation parameter. To project the data into the kernel feature space, the widely used radial basis kernel function is adopted:

$$K(\boldsymbol{x}, \boldsymbol{x}') = \exp\left(\frac{-\|\boldsymbol{x} - \boldsymbol{x}'\|^2}{k_w}\right) \tag{17}$$

where k_w is the kernel width. The advantage of SVM over other pattern recognition models is that the solution to Eq. (16) is unique and can be calculated efficiently. On the other hand, ANNs require an iterative gradient descent solution, which may converge to local minima. Our application area has no issue with large data sets since the number of training samples is only in the order of 10^2 and the number of features is in the order of 10. This setting is ideal for an SVM solution. Once the problem in Eq. (16) is solved, the optimal values α_i^* are obtained, wherein the ith training samples $\boldsymbol{x}_i$ that correspond to $\alpha_i^* > 0$ are deemed *support vectors*. Support vectors participate in creating the boundaries between two classes, defined by the decision function:

$$f(\boldsymbol{x}) = \sum_{i \in SV} y_i \alpha_i^* K(\boldsymbol{x}_i, \boldsymbol{x}) + b^* \tag{18}$$

where SV is the set of support vectors and b^* is a bias term calculated so that $y_i f(\boldsymbol{x}_i) = 1$ for any i with $0 < \alpha_i^* < C$ (Cristianini and Shawe-Taylor, 2014). For any test sample $\boldsymbol{x}$, the function $y = \text{sign}(f(\boldsymbol{x}))$ outputs either +1 or -1 to signify if the sample belongs to the positive or the negative class. Accordingly, the exact boundary between the two classes consists of the points $\boldsymbol{x}$ where the SVM decision becomes indifferent, that is $\text{sign}\big(f(\boldsymbol{x})\big) = 0$.

In the case of the experiment, samples belong to one of the four classes: (1) Bubbly Flow, (2) Slug Flow, (3) Churn Flow, or (4) Annular Flow. Thus, multi-class SVM needs to be implemented. Various strategies for multi-class classification have been proposed, such as one-against-one and one-against-rest. Here, an efficient one-against-one strategy proposed by Platt et al., (2000) called DAGSVM, is adopted. It has been reported that DAGSVM retains the accuracy offered by other approaches, but it is faster to train and evaluate (Platt et al., 2000). Previously, DAGSVM has been adopted in the objective identification of two-phase flow regimes using electrical capacitance data (Wang and Zhang, 2009).

In this work, DAGSVM is employed by training six binary classifiers, one for each possible pair of distinct flow regime classes, e.g. 1-vs-2, 1-vs-3, 1-vs-4, 2-vs-3, and so on. Training a binary classifier involves solving for a decision boundary between two classes in the form of Eq. (18). The classifiers are then arranged in a decision directed acyclic graph (DDAG) as presented in Fig. 6. The DDAG structure is key to the efficiency of DAGSVM, which makes it advantageous over other multi-class classification strategies.

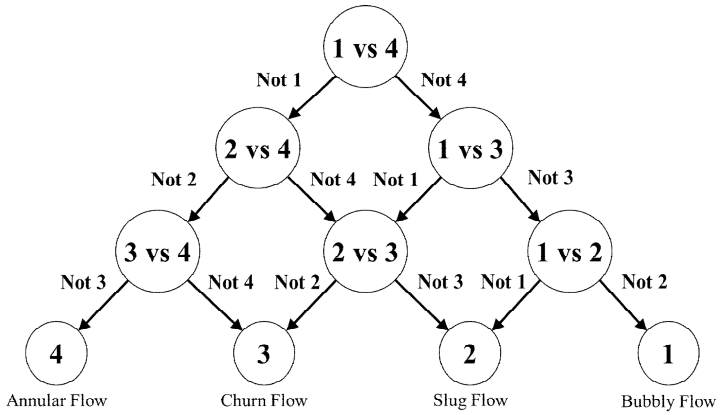

Figure 6: DDAG for Multi-class SVM classification

In the DDAG, any new incoming sample goes through the decision at each node, always starting from the 1-vs-4 node. Each node represents the binary decision of which class the sample is definitely excluded from, for example, the 1-vs-4 node classifies the sample as either “Not 1” or “Not 4”. The branch corresponding to the decision of the current SVM classifier is then traversed. As the downward traversal progresses, the sample is continuously being classified at every node visited by eliminating the excluded class, until only a single class is retained. At this point, the bottom of the DDAG is reached and the sample has been associated to a single flow regime. The implementation of DAGSVM used in this work is available online in Pilario, (2018).

By taking the 2-dimensional data from ultrasound Doppler signals after dimensionality reduction, with the flow regime labels of each sample, the DAGSVM is used to create exact boundaries

between the flow regimes. This completes the virtual flow regime map for use in objective flow regime classification. The summary of the methodology is given in Fig. 7.

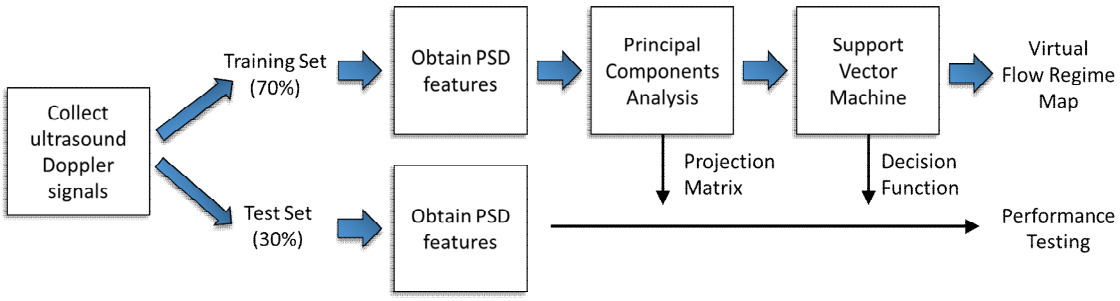

Figure 7: Proposed methodology for objective flow regime identification

4 Results and Discussions

Flow regime data samples were acquired experimentally from the S-shape riser system described in Section 3.1, and preliminary results are presented in Fig. 5. In this section, we proceed with the proposed approach for data analysis and discussion of the results.

4.1 PCA Visualisation

Figure 8 presents the resulting 2D visualisation of the ultrasound Doppler data for each training sample after applying PCA. The benefit of PCA visualization can be demonstrated by establishing the relationship between Fig. 8 and Fig. 4. These figures are similar in that they both represent a map where every point location is associated with a distinct gas-liquid flow rate value pair.

In Fig. 8, Annular Flow samples (high gas flow rate and low liquid flow rate) are found at the lower left corner of the map, while Bubbly Flow samples (low gas flow rate and high liquid flow rate) are found at the right and upper right corners of the PCA map. The Slug Flow and Churn Flow samples are found in a specific order in the middle region. The direction of increasing gas flow rate can be elucidated from right to left in the PCA map, while the direction of increasing

liquid flow rate results from bottom to top. These directions correspond to the axes of the flow regime map in Fig. 4. Because of this relationship, PCA was able to discover the gas-liquid flow rate information of every sample using only the PSD features obtained from the ultrasound Doppler experiment. Thus, PCA can arrange the Doppler data meaningfully in 2D space, further enabling the construction of a virtual flow regime map.

However, there is no clear gap or boundary between the samples from different flow regimes in the PCA map. By comparing it with Fig. 4, it was found that these samples lie mostly in the transition regions. Hence, a soft margin SVM can be used to establish the boundary between the various flow regimes, by setting the value of C to be less than ∞. By further varying the SVM parameters, k_w and C, one can control the complexity of the boundaries between flow regimes. This investigation is carried out in Section 4.2.

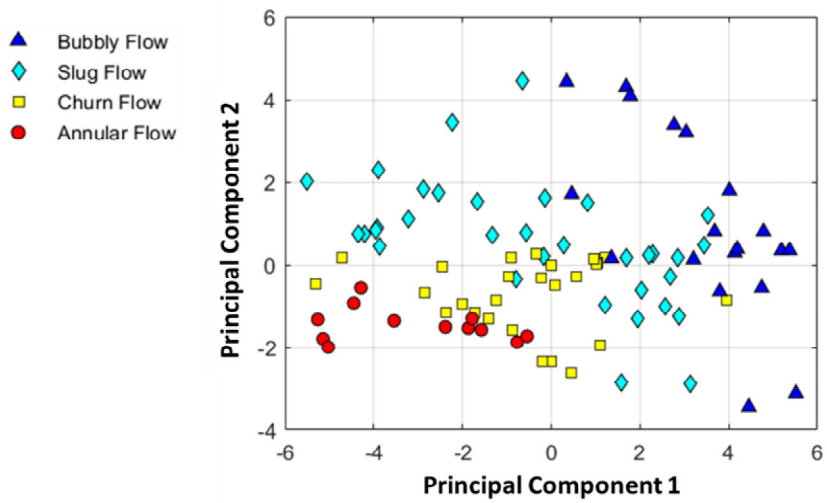

Figure 8: PCA visualisation of training samples of ultrasound Doppler signals

Another benefit of PCA visualisation of data is the detection of outliers. The consistency of the human expert in labelling flow regimes may be impeded by certain factors, leading to the presence of outlier data samples. One obvious case is the sample at 10 Sm^3/hr air flow rate and 3.5 kg/s water flow rate, which was observed to exhibit churn flow (see Fig. 4), yet which is found between

bubbly and slug flow regime samples. In the PCA visualisation of ultrasound Doppler data (see Fig. 8), this specific data point lies at a position near (4.0, -1.0) on the 2D map, also between bubbly and slug flow samples. Hence, this data point is considered an outlier. Other outliers confirmed in the same way in our training dataset include those at air-water flow rates of 200 Sm^3/h, 4.5 kg/s and 10 Sm^3/h, 2.5 kg/s. Ultimately, the degree to which these outliers are tolerated by the subsequent SVM classification is dictated by setting appropriate parameters of k_w and C.

4.2 SVM classification

Figure 9 is a sample flow regime map for $k_w = 7$ and $C = 100$ where the samples of both the training and test data are superimposed. On this map, the background colours denote the results from the SVM classification, for example, SVM identified the Slug Flow regime for every point location in the L-shaped region between Bubbly and Churn Flow regime. Superimposed on this map are the training data samples (circles) and test data samples (triangles). By noting the mismatch between the sample colours and background colours, the training and test data classification accuracies are found to be 85.7% and 84.6% respectively. Without counting the confirmed outliers in the training samples, the accuracy in the training data is 88.6%. These results depict the capability of the SVM approach for objective classification of two-phase flow regime based on Doppler ultrasound data.

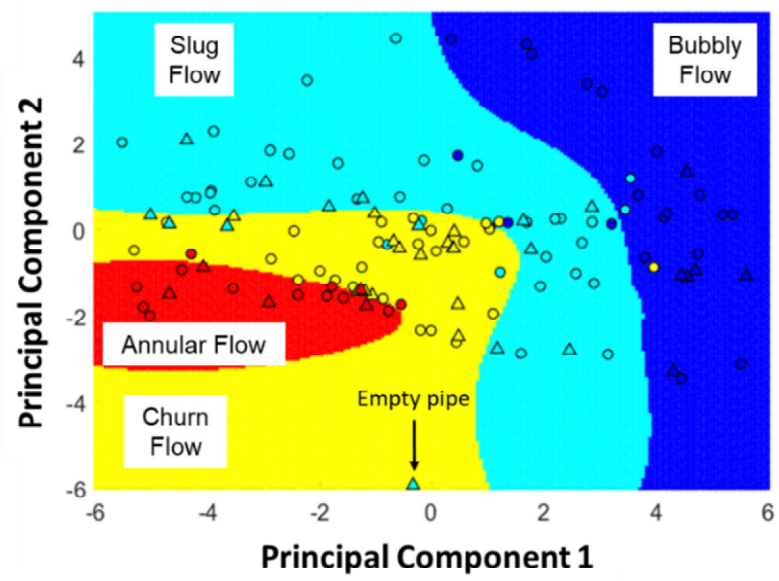

Figure 9: Virtual flow regime map using SVM at $\boldsymbol{k_w = 7}$ and $\boldsymbol{C = 100}$. The background colours denote the SVM identification results. A mismatched sample colour and background colour indicates misclassification. Legend: Circles – training samples; triangles – test samples

For completeness, a case of empty pipe was included in the test samples. Interestingly, this sample was identified as churn flow by the SVM. Although this result is unexpected, the fact that the PCA mapping placed the empty-pipe ultrasound Doppler signal data at the bottom of the map validates that the direction of increasing liquid flow rate occurs upward in the PCA visualisation.

4.3 SVM Performance at Different Parameters

The list of misclassified samples in Fig. 9 is presented in Table 1. Some disparities in using the proposed identification method were observed. In particular, the objective classifier identifies some Slug Flow samples as Churn Flows samples. This misclassification is related to the parameter choice issues in the SVM objective classifier. Various settings for k_w and C give varying classification performance. The SVM accuracy over a grid of parameter values, such as $k_w \in \{1,3,\dots,9\}$ and $C \in \{1,10,\dots,10^4\}$, is presented in Fig. 10.

Table 2: List of Misclassified Samples in Fig. 9

Misclassified Training Samples (13 out of 91)				
Vsg (Sm^3/h)	Vsl (kg/s)	SVM Classification	Actual Classification	Outlier?
10	2.5	Slug Flow	Bubbly Flow	Y
10	3.5	Bubbly Flow	Churn Flow	Y
20	2	Churn Flow	Slug Flow	N
20	4	Slug Flow	Bubbly Flow	N
30	4	Bubbly Flow	Slug Flow	N
50	2	Churn Flow	Slug Flow	N
50	3	Churn Flow	Slug Flow	N
50	4.5	Bubbly Flow	Slug Flow	N
120	1	Annular Flow	Churn Flow	N
120	4	Slug Flow	Churn Flow	N
200	4.5	Slug Flow	Bubbly Flow	Y
300	0.1	Churn Flow	Annular Flow	N
300	1.5	Churn Flow	Annular Flow	N
Misclassified Test Samples (6 out of 39)				
Vsg (Sm^3/h)	Vsl (kg/s)	SVM Classification	Actual Classification	
5	0.5	Churn Flow	Slug Flow	
20	0.5	Churn Flow	Slug Flow	
30	2	Churn Flow	Slug Flow	
50	0.1	Churn Flow	Slug Flow	
300	0.5	Churn Flow	Annular Flow	
0	0	Churn Flow	Empty Pipe	

Accurate classification of training data can be obtained by adjusting k_w and C towards the direction of overfitting (lower k_w and higher C). However, overfitting demonstrates poor generalisations of unseen test data. Concisely, overfitting makes the classification biased towards the training samples. On the other hand, at high k_w and low C, under-fitting occurs. In the case of under-fitting, the boundaries tend towards linearity at the expense of higher misclassification rates. In general, the only way to increase the level of confidence with the resulting flow regime map is to validate it against as numerous unseen test data samples as possible. With only the available data, the choice of $k_w = 7$ and $C = 100$ already provides useful results for objective flow regime classification, while striking a balance between overfitting and underfitting.

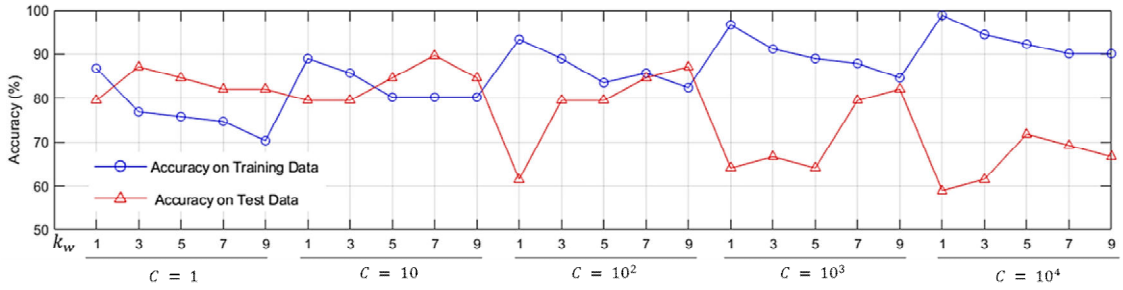

Figure 10: Accuracy of SVM classification at various parameter settings of k_w and C

With the virtual flow regime map at hand, further analysis on the flow regime transitions and uncertainties can be performed. More importantly, online objective flow regime identification can be developed from the approach proposed in this work. Using a continuous feed of ultrasound Doppler-based flow velocity information, PCA-SVM can automatically visualise the frequency-domain features and classify the flow regime at every sampling instant. Hence, the proposed approach has broad potential for industrial applications.

5 Conclusion and future works

In this paper, the necessity of objective, non-invasive and non-intrusive measurement methods for flow regime identification in industrial practice is highlighted. Specifically, this work proposes the use of non-invasive clamp-on continuous wave ultrasound Doppler (CWUD) and machine learning approaches for objective two-phase gas/liquid flow regime identification. From the ultrasonic signals, Power Spectral Density (PSD) features were extracted and subjected to principal components analysis (PCA) to project the data in 2-dimensional space. A multi-class support vector machine (SVM) classifier is trained to establish exact boundaries between the flow regimes in the reduced data space. In the end, the objective classifier accuracy for both the training and testing data samples was 85.7% and 84.6% respectively. More importantly, the generation of virtual flow regime maps provided useful data visualisations of the Doppler signals, which can

aid in detecting outliers and explain the decisions made by the SVM classifier. These results justify the suitability of our approach for flow regime identification in industrial practice.

To improve this work, the proposed approach must be tested against many other test rigs and configurations to determine if the generated virtual flow regime maps are indeed capable of visualising flow regime patterns from the CWUD data. In addition, the feature extraction and dimensionality reduction steps are deemed the most important steps in the entire procedure. Many other techniques for these steps must be tested to see if various samples from different flow regimes can be clearly separated.

Further research can also be done in applying the proposed approach to examine two-phase water-oil flow, mostly to address the necessity of clamp-on non-invasive ultrasonic flow monitoring for oil well testing.

Acknowledgement

The authors would like to thank the DOST-ERDT Faculty Development Fund of the Republic of the Philippines for supporting this work, as well as the Process Systems Engineering Laboratory at Cranfield University.

References

Abbagoni, B.M., Yeung, H., 2016. Non-invasive classification of gas-liquid two-phase horizontal flow regimes using an ultrasonic Doppler sensor and a neural network. Meas. Sci. Technol. 27. https://doi.org/10.1088/0957-0233/27/8/084002

Barnea, D., 1987. A Unified Model for Predicting Flow-Pattern Transitions for the Whole Range of Pipe Inclinations. Int. J. Multiph. Flow 13, 1–12.

Barnea, D., Shoham, O., Taitel, Y., 1980. Flow pattern characterization in two phase flow by electrical conductance probe. Int. J. Multiph. Flow 6, 387–397.

Blaney, S., Yeung, H., 2008. Investigation of the exploitation of a fast-sampling single gamma

densitometer and pattern recognition to resolve the superficial phase velocities and liquid phase water cut of vertically upward multiphase flows. Flow Meas. Instrum. 19, 57–66. https://doi.org/10.1016/j.flowmeasinst.2007.09.002

Chakraborty, S., Keller, E., Talley, J., Srivastav, A., Ray, A., Kim, S., 2009. Void fraction measurement in two-phase flow processes via symbolic dynamic filtering of ultrasonic signals. Meas. Sci. Technol. 20, 023001. https://doi.org/10.1088/0957-0233/20/2/023001

Chaouki, J., Larachi, F., Dudukovic, M.P., 1997. Non-invasive Monitoring of Multiphase Flows, Elsevier.

Chivers, R.C., Hill, C.R., 1975. A spectral approach to ultrasonic scattering from human tissue: methods, objectives and backscattering measurements. Phys. Med. Biol. 20, 799–815.

Cortes, C., Vapnik, V., 1995. Support-Vector Networks. Mach. Learn. 20, 273–297. https://doi.org/10.1023/A:1022627411411

Cristianini, N., Shawe-Taylor, J., 2014. Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press.

de Kerret, F., Béguin, C., Etienne, S., 2017. Two-phase flow pattern identification in a tube bundle based on void fraction and pressure measurements, with emphasis on churn flow. Int. J. Multiph. Flow 94, 94–106. https://doi.org/10.1016/j.ijmultiphaseflow.2017.04.013

Drahoš, J., Čermák, J., 1989. Diagnostics of gas—liquid flow patterns in chemical engineering systems. Chem. Eng. Process. Process Intensif. 26, 147–164. https://doi.org/http://dx.doi.org/10.1016/0255-2701(89)90007-X

Dyakowski, T., 1996. Process tomography applied to multi-phase flow measurement. Meas. Sci. Technol. 7, 343–353.

Evans, D., McDicken, W., Skidmore, R., Woodcock, J., 1989. Doppler ultrasound: Physics, instrumentation, and clinical applications. John Wiley & Sons, Inc.

Evans, D.H., McDicken, N.W., 2000. Doppler Ultrasound: Physics, Instrumentation, and Signal Processing.

Eyo, E.N., Pilario, K.E.S., Lao, L., Falcone, G., 2019. Development of a Real-Time Objective Gas – Liquid Flow Regime Identifier Using Kernel Methods. IEEE Trans. Cybern. PP, 1–11. https://doi.org/10.1109/TCYB.2019.2910257

Falcone, G., Hewitt, G., Alimonti, C., 2009. Multiphase Flow Metering: Principles and Applications, Elsevier.

Falcone, G., Hewitt, G.F., Alimonti, C., Harrison, B., 2002. Multiphase Flow Metering: Current Trends and Future Developments. J. Pet. Technol. 54, 77–84. https://doi.org/10.2118/74689-JPT

Hanus, R., Zych, M., Petryka, L., Świsulski, D., Strzępowicz, A., 2017. Application of ANN and PCA to two-phase flow evaluation using radioisotopes. EPJ Web Conf. 143, 2–5. https://doi.org/10.1051/epjconf/201714302033

Jha, D.K., Ray, A., Mukherjee, K., Chakraborty, S., 2012. Classification of Two-Phase Flow Patterns by Ultrasonic Sensing. J. Dyn. Syst. Meas. Control 135, 024503. https://doi.org/10.1115/1.4007555

Jones, O.C., Delhaye, J.M., 1976. Transient and statistical measurement techniques for two-phase flows: A critical review. Int. J. Multiph. Flow 3, 89–116. https://doi.org/10.1016/0301-9322(76)90001-X

Jones, O.C., Zuber, N., 1975. The interrelation between void fraction fluctuations and flow patterns in two-phase flow. Int. J. Multiph. Flow 2, 273–306.

Julia, J.E., Hibiki, T., 2011. Flow regime transition criteria for two-phase flow in a vertical annulus. Int. J. Heat Fluid Flow 32, 993–1004. https://doi.org/10.1016/j.ijheatfluidflow.2011.06.001

Juliá, J.E., Liu, Y., Paranjape, S., Ishii, M., 2008. Upward vertical two-phase flow local flow regime identification using neural network techniques. Nucl. Eng. Des. 238, 156–169. https://doi.org/10.1016/j.nucengdes.2007.05.005

Kouam, D., Girault, J., Remenieras, J., 2003. High Resolution Processing Techniques for Ultrasound Doppler Velocimetry in the Presence of Colored Noise . Part II : Multiplephase Pipe-Flow Velocity Measurement. IEEE 50, 267–78.

Meire, H.B., Farrant, P., 1995. Basic Ultrasound.

Meribout, M., Al-Rawahi, N.Z., Al-Naamany, A.M., Al-Bimani, A., Al-Busaidi, K., Meribout, A., 2010. A multisensor intelligent device for real-time multiphase flow metering in oil fields. IEEE Trans. Instrum. Meas. 59, 1507–1519. https://doi.org/10.1109/TIM.2009.2028210

Peddu, A., Chakraborty, S., Kr. Das, P., 2017. Visualization and flow regime identification of downward air-water flow through a 12 mm diameter vertical tube using image analysis. Int. J. Multiph. Flow 100, 1–15. https://doi.org/10.1016/j.ijmultiphaseflow.2017.11.016

Pilario, K.E., 2018. Binary and Multi-class SVM [WWW Document]. MATLAB Cent. File Exch. URL https://uk.mathworks.com/matlabcentral/fileexchange/65232-binary-and-multi-class-svm (accessed 6.8.18).

Platt, J., Cristianini, N., Shawe-Taylor, J., 2000. Large Margin DAGs for Multiclass Classification. Int. Conf. Neural Inf. Process. Syst. 547–553. https://doi.org/10.1.1.158.4557

Rouhani, S.Z., Sohal, M.S., 1983. Two-phase flow patterns: A review of research results. Prog. Nucl. Energy 11, 219–259. https://doi.org/10.1016/0149-1970(83)90012-4

Salgado, C.M., Pereira, C.M.N.A., Schirru, R., Brandão, L.E.B., 2010. Flow regime identification and volume fraction prediction in multiphase flows by means of gamma-ray attenuation and artificial neural networks. Prog. Nucl. Energy 52, 555–562. https://doi.org/10.1016/j.pnucene.2010.02.001

Sanderson, M.L., Yeung, H., 2002. Guidelines for the use of ultrasonic non-invasive metering techniques. Flow Meas. Instrum. 13, 125–142. https://doi.org/10.1016/S0955-5986(02)00043-2

Sunde, C., Avdic, S., Pázsit, I., 2005. Classification of two-phase flow regimes via image analysis and a neuro-wavelet approach. Prog. Nucl. Energy 46, 348–358. https://doi.org/10.1016/j.pnucene.2005.03.015

Thorn, R., Johansen, G. a, Hjertaker, B.T., 2012. Three-phase flow measurement in the petroleum industry. Meas. Sci. Technol. 24, 012003. https://doi.org/10.1088/0957-0233/24/1/012003

Trafalis, T.B., Oladunni, O., Papavassiliou, D. V., 2005. Two-Phase Flow Regime Identification with a Multiclassification Support Vector Machine (SVM) Model. Ind. Eng. Chem. Res. 44, 4414–4426. https://doi.org/10.1021/ie048973l

Übeyli, E.D., Güler, I., 2005. Improving medical diagnostic accuracy of ultrasound Doppler signals by combining neural network models. Comput. Biol. Med. 35, 533–554. https://doi.org/10.1016/j.compbiomed.2004.03.006

Van Der Maaten, L.J.P., Postma, E.O., Van Den Herik, H.J., 2009. Dimensionality Reduction: A Comparative Review. J. Mach. Learn. Res. 10, 1–41. https://doi.org/10.1080/13506280444000102

Wang, H.X., Zhang, L.F., 2009. Identification of two-phase flow regimes based on support vector machine and electrical capacitance tomography. Meas. Sci. Technol. 20. https://doi.org/10.1088/0957-0233/20/11/114007

Weinstein, E., 1982. Measurement of the Differential Doppler Shift. IEEE Trans. Acoust. 30, 112–117. https://doi.org/10.1109/TASSP.1982.1163849

Xie, T., Ghiaasiaan, S.M., Karrila, S., 2004. Artificial neural network approach for flow regime classification in gas-liquid-fiber flows based on frequency domain analysis of pressure signals. Chem. Eng. Sci. 59, 2241–2251. https://doi.org/10.1016/j.ces.2004.02.017