Flow Regime Identification for Air Valves Failure Evaluation in Water Pipelines Using Pressure Data

Haixing Liu[1], Yan Zhu[2], Shengwei Pei[1], Dragan Savić[3,4], Guangtao Fu[4], Chi Zhang[1], Yixing Yuan[2], Jinsong Zhang[5]

[1] School of Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China

[2] School of Municipal and Environmental Engineering, Harbin Institute of Technology, Harbin, China

[3] KWR Watercycle Research Institute, Nieuwegein, Netherlands

[4] College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

[5] Shenzhen Water Group, Shenzhen, China

Corresponding author: Chi Zhang

Abstract: Air valve failure can cause air accumulation and result in a loss of carrying capacity, pipe vibration and even in some situations a catastrophic failure of water transmission pipelines. Air is most likely to accumulate in downward sloping pipes, leading to flow regime transition in these pipes. The flow regime identification can be used for fault diagnosis of air valves, but has received little attention in previous research. This paper develops a flow regime identification method that is based on support vector machines (SVMs) to evaluate the operational state of air valves in freshwater/potable pipelines using pressure signals. The laboratory experiments are set up to collect pressure data with respect to the four common flow regimes: bubbly flow, plug flow, blow-back flow and stratified flow. Two SVMs are constructed to identify bubbly and plug flows and validated based on the collected pressure data. The results demonstrate that pressure signals can be used for identifying flow regimes that represent the operational state (functioning or malfunctioning) of air valves. Among several signal features, Power Spectral Density and Short-Zero Crossing Rate are found to be the best indictors to classify flow regimes by SVMs. The sampling rate and time of pressure signals have significant influence on the performance of SVM classification. With optimal SVM features and pressure sampling parameters the identification accuracies exceeded 93% in the test cases. The findings of this study show that the SVM flow regime identification is a promising methodology for fault diagnosis of air valve failure in water pipelines.

Keywords: Water transmission Pipeline; Air valve; Flow regime identification; Support vector machine; Fault diagnosis

1 Introduction

Freshwater resources are unevenly distributed temporally and spatially and particularly mismatching with urbanization development, thus leading to severe imbalance between supply and demand. Many long-distance water transmission pipelines were built to address this type of problem, such as in the case of the Snowy Mountains Hydro-electric Scheme in Australia (Bergmann, 1999), the Great Lakes Basin Water Diversion (Becker and Easter, 1995), the Central Valley Project in the USA (Mariño and Loaiciga, 1985), and South to North Water Diversion Project in China (Barnett et al., 2015; Yu et al., 2018). In the pressurized water transmission pipeline, air valves are a common component which is used to deaerate the pipe (Meng et al, 2016; Pothof and Clemens, 2012). The existence of air in water pipelines will not only decrease water conveyance capacity and increase head losses (Escarameia, 2007; Lubbers, 2007), but could also lead to pipe vibration due to the pressure fluctuation of air-water mixing flows. Worse still, it may even cause pipe to burst during the hydraulic transient process (Pothof and Clemens, 2010). Air is most likely to accumulate in the downward sloping water pipe, which can usually be removed by flowing water or discharged through air valves (Escarameia, 2007). If the water velocity is lower than the "clearing velocity" (Kalinske and Bliss, 1943; Kent, 1952; Wisner et al, 1975), the air cannot be completely taken away by hydraulic actions, therefore, air valves are essential for removing air in this situation. Air valves often break down due to blockage in the vent hole or valve-stem rupture caused by pipe vibration (Ramezani et al, 2015). However, air valves that are normally located at the high spots of water conveyance pipelines are inconvenient to inspect and maintain (Pothof and Clemens, 2012). Moreover, air valve failures cannot be easily detected and rectified, affecting the performance of the pipeline.

Fault diagnosis of air valves located in mechanical systems (e.g., diesel, compressor) is often performed by various pattern recognition algorithms based on the analysis of acoustic and/or vibration signals (Pichler et al, 2011; Qin et al, 2012; Verma et al, 2011). However, acoustic and vibration signals are not suitable for fault diagnosis of air valves in water conveyance pipeline due to poor working conditions (e.g. humid underground, limited power and communication, and external disturbance) (Stephens et al., 2004; Schwaller and van Zyl, 2015). However, pressure signals have been successfully employed in valve fault diagnosis in mechanical systems (Feng et al, 2011). In the downward sloping water pipes, when the air valves are out of service, the air accumulation can lead to flow regime transition. This in turn can lead to the change in the dynamic behavior of the valve and pressure fluctuations in the air-water flow. Moreover, pressure signals in the water pipeline are commonly available, therefore, this study investigates their use to identify flow regime changes and detect the operational state of air valves in downward sloping water pipes.

Flow regime identification has been widely studied in the literature. Numerous investigators have applied various types of instruments to collect different data (e.g., flow image, void fraction or differential pressure) for flow regime identification (Arvoh et al, 2012; Lee et al, 2008a,b; Roshani et al, 2015; Salgado et al, 2010). The features, e.g., frequency, stochasticity, fractals or chaotic time series characteristics, are extracted from the data in order to improve the identification accuracy for a specific application (Cai et al, 1994; Elperin and Klochko, 2002; Franca et al, 1991; Sun et al, 2013; Vince and Lahey, 1982).

Moreover, identification of signal features (pattern recognition) can utilize the classification methods, i.e., distinguishing the features of anomaly events from a spectrum of event sets. (Mi et al, 2001; Roshani et al, 2015; Tan et al, 2007). Support vector machines (SVM) and artificial neural networks (ANN) are probably the most commonly used pattern recognition methods. SVM are less prone to being trapped in a local minimum and require less data during the training process, but most importantly can overcome the key weakness of ANN, which require an appropriate structure to be selected and optimized for the problem at hand (Yang et al., 2017). In the water sector, SVM has been used for water quality classification as an early warning tool. For example, a least square support vector machine (LS-SVM) was combined with fuzzy clustering to estimate water quality failures in water distribution networks (Aydogdu and Firat, 2015; Modaresi and Araghinejad, 2014). Moreover, SVM is used for anomaly detection in water distribution systems based on the pressure and flow signals (Mounce et al, 2011). In addition to classification applications, SVM is also applied in precipitation and runoff predictions (Ahmadi et al, 2015; Bray and Han, 2004). These applications prove that SVM has a strong capability as a classification tool. In this study, SVM is used for flow regime identification in water pipes, whereby the flow regime changes could be used for diagnosing air valve operation states.

This paper aims to evaluate the operational state of air valve in the downward sloping water pipe based on pressure data using SVM classification. Laboratory experiments are set up to collect the pressure data. Then the data are used to train and validate SVM models. Bubbly flow (including quasi-pure water state) or plug flow can be classified amongst all flow regimes through the SVMs. According to the analysis of experimental data and the SVM-based results, the following aspects are addressed by: (i) analysis of the optimal time-frequency characteristics of pressure signals corresponding to different flow regimes in downward sloping water pipes, (ii) identification accuracy of SVM models using the different features extracted from pressure data, and (iii) parameter analysis of SVM input data (e.g. optimal sampling rate and time)? This paper presents a novel methodology for flow regime identification in water pipe systems. Flow regime in a water pipe can then directly be linked to the operational state of the air valve, thus performing air valve fault diagnosis and ensuring the safety of water pipelines. The findings of this study can also provide potential guidance for parameter estimation of the SVM classification model.

2 Experimental facility

A single-pipe transmission system is set up to collect the pressure signal data under the various flow conditions. As shown in Fig. 1, the experimental set-up is based on a circulating water system pressurized by a pump. All pipes are made of plexiglas, and the total length of the pipeline is about 80 m. The inner and outer diameters of the pipes are 90 mm and 110 mm, respectively. To simulate air-water two-phase flow in downward slope, the air was injected into the upstream of the pipeline by an air compressor. The digital signals of pressure sensor and ultrasonic flowmeter are collected by a multiple-channel data acquisition card (Advantech USB-4711A).

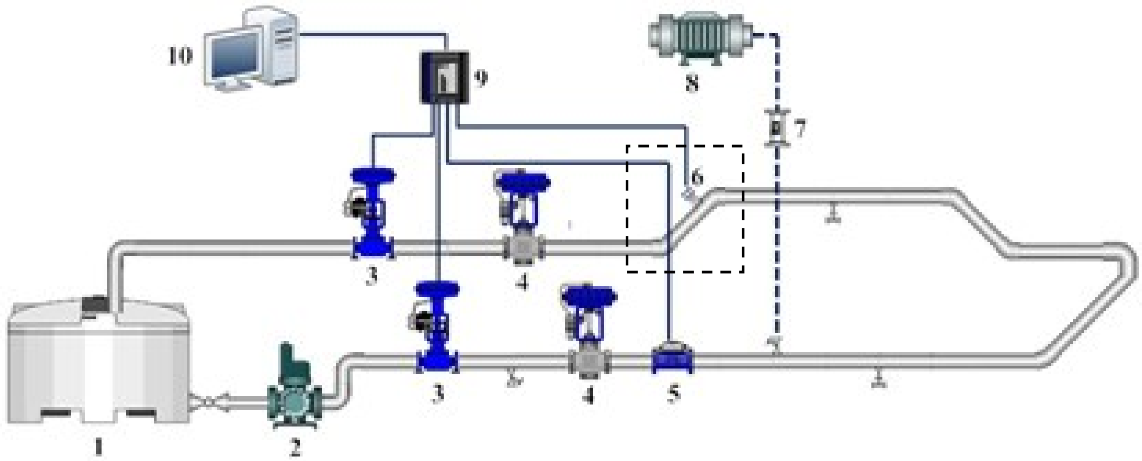


Fig. 1 Diagram of the test rig used in this study (test section is shown in the dashed box) - 1: water tank; 2: pump; 3: electric valve; 4: pneumatic butterfly valve; 5: ultrasonic flowmeter; 6: pressure sensor; 7: gas rotameter; 8: air compressor; 9: data acquisition instrument; 10: computer

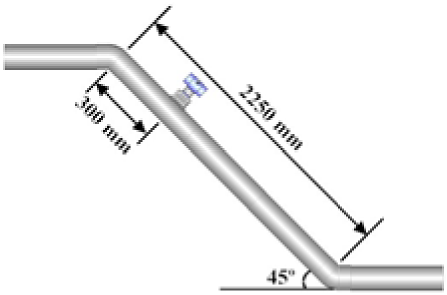


Fig. 2 A detail of the test section with the pressure sensor

The tested pipe section in downward slope is illustrated in Fig. 2. The length of the test section is 2,250 mm, and the pressure sensor is deployed at 300 mm away from the upstream elbow. The measurement range of the pressure sensor is 0~200 kPa with an accuracy of 0.2%. The experimental procedure includes the following aspects:

1) The planned water velocity range included 16 different values (0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6 and 2.7 m/s), which were controlled by the electric valve near the pump. The exact value of the velocity was then measured by the ultrasonic flowmeter.

2) The air flow measured by the gas rotameter was set to 8 discrete values (0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5 and 4.0 m^3/h), which was controlled by the outlet valve at the air compressor.

3) A total of 128 cases were tested during the experiment based on the orthogonal combination of different water velocities and air flow values.

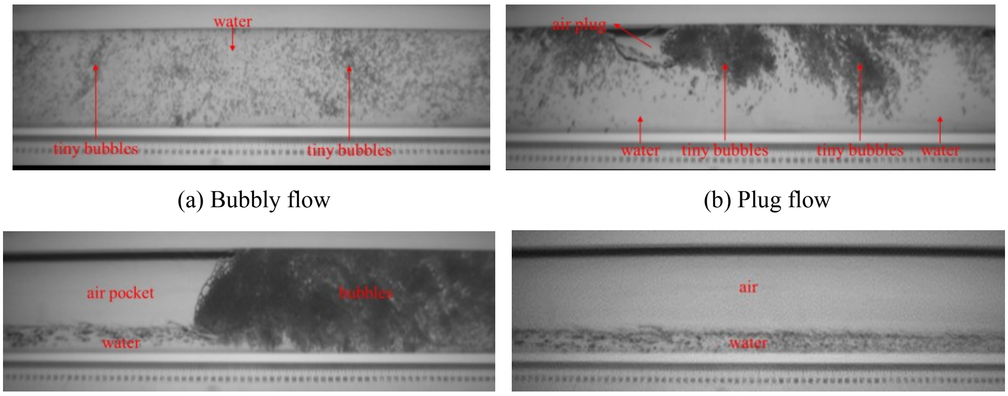
4) The pressure signals were sampled at a frequency of 1 kHz and the sampling time was set to 20 s.

5) The images of air-water flow regime at the test section were recorded by a high-speed camera from a side view.

3 Methodology

3.1 Description of Flow Regimes

It is known that four flow regimes may appear in the downward sloping water pipes (Pothof and Clemens, 2011). Those can be seen in the images of obtained in the tests, as shown in Fig. 3. It should be noted that the camera lens is placed in parallel with the longitudinal pipe axis, thus the pipes in the images appear to be horizontal.



(a) Bubbly flow

(b) Plug flow

(c) Blow-back flow

(d) Stratified flow

Fig. 3 Typical flow regime images

There is always a small amount of dissolved air in the water, which can be assumed as a state of "quasi-pure water" (Zhu et al, 2018). Although the amount of dissolved air is very small, the effect of local resistance by the elbow will aggregate it into tiny bubbles that are dispersed in pipes (Barnea, 1986). Therefore, bubbly flow (Fig. 3a) can be considered the normal flow regime in the downward sloping water pipes. When the water velocity is lower than "clearing velocity", small bubbles will accumulate to create plug flow (Fig. 3b), and the air plugs will finally be discharged out of the pipe if the air valve is working normally. However, if the air valve is not working properly, the air plugs will gradually coalesce into a large air pocket at the top of the slope, which is called "blow-back flow" as shown in Fig. 3c. As the large air pocket continues to expand until occupying the entire slope, stratified flow occurs with the water flowing beneath the air (Fig. 3d). According to the flow regime analysis, the occurrence of blow-back flow or stratified flow in downward sloping pipes indicates that the air valve may be malfunctioning.

3.2 SVM-based Flow Regime Identification

Classification is a process of classifying data points into specific groups (or classes). SVM is one of the powerful methods for data classification. The training process of a SVM classifier involves finding the best hyperplane that divides the two classes with the given data points. If it is necessary to divide the data points into n classes, n-1 hyperplanes should be constructed by the classifier. The original maximum-margin hyperplane algorithm proposed by Vapnik (1963) is used to build a linear classifier and can only solve linearly classification problems. In this paper, we used the maximum-margin method, proposed by Boser et al. (1992), to deal with the nonlinear classification problems by a spectrum of linear decision functions separating hyperplanes in a transformed high-dimensional feature space. The transformation algorithm is similar to the mapping calculation except dot product which is replaced by a nonlinear kernel function (Boser et al., 1992). For a specific nonlinear classification problem, an appropriate kernel function is crucial for the classification performance, and therefore the Radial Basis Function (RBF) (Ring and Eskofier, 2016) is used as kernel functions through several trials.

The flow regimes of blow-back or stratified flows are the most severe situations of air accumulation in pipelines, which indicates air-valve is likely to fail completely. Although plug flow regime shows the less volume of air accumulation than blow-back and stratified flows, the frequent occurrence of plug flow indicates potential drawbacks for air exhausting (e.g., the insufficient number of air valves). Therefore, the four flow regimes are hierarchically classified into three categories for improving the identification accuracy, since both the blow-back and stratified flows are considered in the same category. Fig. 4 shows the flowchart of the flow regime identification. As can be seen in Fig. 4, two SVM models have been constructed to identify the different flow regimes: 1) SVM-1, which is used to distinguish bubbly flow (including quasi-pure water state) from the other three flow regimes; and 2)

SVM-2, which is used to distinguish plug flow from blow-back and stratified flows. The detailed training and testing process are introduced in the following section.

According to Fig. 4, pressure signals are collected with low frequency (twice a day, for example) at normal conditions. The sampling frequency (i.e. sampling times per minute) will be conducted and maintained continuously when the results of SVM-1classification demonstrates bubbly flow. If the results of SVM-1 are not identified as the quasi-pure water state or bubbly flow, the features abstracted from pressure signals will be passed as input to SVM-2. When the SVM-2 classification results in the plug flow classification, a “Warning” signal will be given and then the sampling frequency of pressure signals should be increased in order to strength the monitoring of the transition process from plug flow to blow-back and stratified flows. Once the results of SVM-2 classification does not show plug flow, an “Alarm” signal will be launched, which indicates the potential failure of air valve.

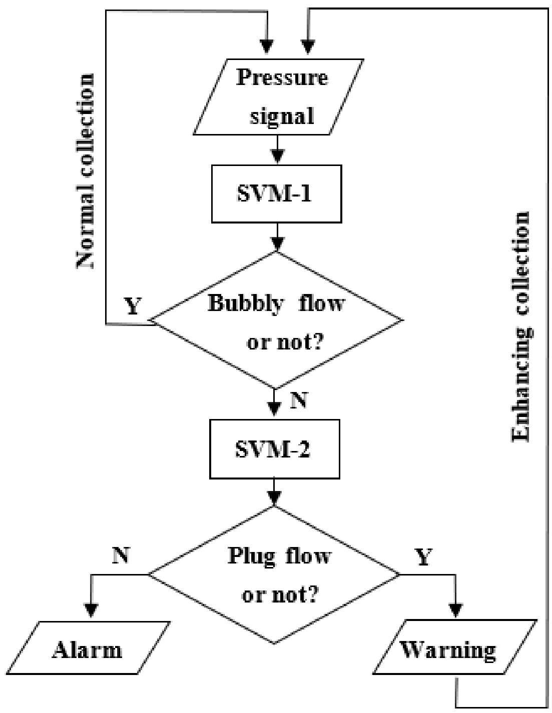


Fig. 4 Flowchart of the evaluation of air valve exhaust

3.3 SVM Training and Testing

Among the entire 128 cases, two cases were not used for SVM analysis due to the faulty data acquisition card (one for blow-back flow and the other for stratified flow). Among the remaining 126 cases, there are 69, 34, 18 and 5 cases for bubbly flow, plug flow, blow-back flow and stratified flow, respectively. These were determined by analyzing flow regime photos (see Fig. 3). SVM-1 is trained based on 94 cases, and the remaining 32 cases are used to test the identification accuracy of SVM-1. SVM-2 is trained based on 42 cases exhibiting the three characteristic flow regimes (plug flow, blow-back flow and stratified flow). The remaining 15 are used to test the identification accuracy of SVM-2. The SVM training and testing in this study are performed on the Matlab R2011b platform, and the procedures of SVMs training and testing can be described as follows:

1) Pressure data are preprocessed by downsampling, which can change the sampling rate of the original pressure signal.

2) In order to make the model training and the obtained testing accuracy more reliable, it is necessary to enlarge the data volume for both SVM training and testing. This is done by dividing data series based on the same interval. Assuming that the total length of one pressure signal series is L, the length of one sample is L', and the interval between the adjacent samples is I, the number of samples (N) in a pressure signal series can be calculated by:

$$N = \lceil (L - L')/I \rceil + 1 \tag{1}$$

Fig. 5 shows the example of how samples are created from one pressure data series in this study. As can be seen in Fig. 5, the total data volume of one pressure signal [P_i] is 10. The length of one sample is 5.

The interval between two adjacent samples is 1. In Fig. 5, the first sample is from P_1 to P_5, and the last sample is from P_6 to P_{10}. Hence the total number of samples is 6, which is equal to the result calculated by Eq. (1).

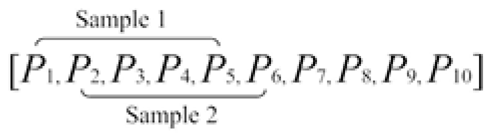


Fig. 5 Samples for an experimental process

3) Every training and testing sample is labeled with 1 or 0, which indicates whether it belongs to a given category.

4) The SVMs are trained based on the features extracted from the training samples. The pressure signals with respect to different flow regimes have significant differences in terms of three types of features, i.e., pressure fluctuation, periodicity and frequency distribution. The features and their types are listed in Table 1.

Table 1. Features for flow regime identification

Feature type	Feature name	Abbreviation
Pressure Fluctuation	Variance	V
	Short-time Zero-crossing Rate	SZR
Periodicity	Autocorrelation Coefficient	AC
Frequency Distribution	Hilbert-Huang Transform	HHT
	Power Spectrum Density	PSD

In Table 1, Variance (V) can reflect the amplitudes of pressure signals for different flow regimes. Short-time Zero-crossing Rate (SZR) is the rate at which the pressure signal changes from positive to negative within a short period of time after subtracting their mean values. Autocorrelation Coefficient (AC) is a feature vector which consists of 51 autocorrelation coefficients, since each test lasts 2 seconds including 51 samples. Hibert-Huang Transform (HHT) is an algorithm that decomposes a signal into various components for obtaining the instantaneous frequency (Ding et al, 2007). The frequency band in the range of 0 to 10 Hz is evenly divided into 5 components. Each frequency band is 2 Hz here. The remaining frequency range (> 10 Hz) belongs to the 6th component. Six eigenvalues (i.e., components) therefore exist in the feature vector of HHT. With respect to Power Spectrum Density (PSD), the pressure signal value is normalized by subtracting the mean value to eliminate the interference of the average pressure. Then the Nyquist rate (using half of the sampling rate, 500 Hz) is equally divided into 128 frequency bands, where the average size of frequency band (i.e. frequency interval) is about 3.9 Hz. The 129 eigenvalues are included in the feature vector of PSD. Both HHT and PSD are the frequency-domain features. Since the feature extraction of HHT is more time-consuming than PSD, the number of frequency bands in the HHT feature is less than that of PSD. The frequency bands of HHT and PSD are set differently which will be discussed in 4.1.2.

5) The accuracy of the SVM classification for flow regime identification is investigated using the experimental samples. The RBF is used as the kernel function in the SVM model and the parameters of SVM are estimated by the Sequential Minimal Optimization (SMO) method during each training (Platt, 1998).

4. Results and Discussion

4.1 Time-frequency Analysis of Experimental Data

Flow regimes in water pipes are different from those in other media (e.g., oil-gas transportation, chemical pipelines or nuclear reactors) (Crawford et al, 1985). Therefore, the time-frequency characteristics of pressure signals of different flow regimes need to be analyzed for flow regime identification.

4.1.1 Time-domain Characteristics of Pressure Signal

Four typical cases are selected to show the time-series pressure signals of different flow regimes in Fig. 6. q_a refers to the air flow and v_w represents the water velocity. As shown in Fig. 6a, the pressure fluctuation of bubbly flow is relatively stable. Because there is little amount of air in bubbly flow, the interaction between air and water phases is not strong. In the plug flow case (Fig. 6b), the air plug moves up and down due to its volume change or deformation, and thus the pressure signal of the plug flow shows occasional pressure drops. In Fig. 6c, the characteristic of pressure fluctuation of blow-back flow is similar to that of plug flow, but the changes in the pressure drop are more pronounced. The greater volume of air has been distributed throughout the downward slope with respect to stratified flow, and the air-water interface intermittently flows back into the slope. Hence, the pressure of stratified flow fluctuates periodically, as shown in Fig. 6d. Fig. 6 also demonstrates that the amplitudes of pressure fluctuation of the stratified flow is the largest one, while the second largest is the blow-back flow. In order to further confirm this result, the variance analysis of all the pressure signals is conducted to express the amplitude of pressure changes.

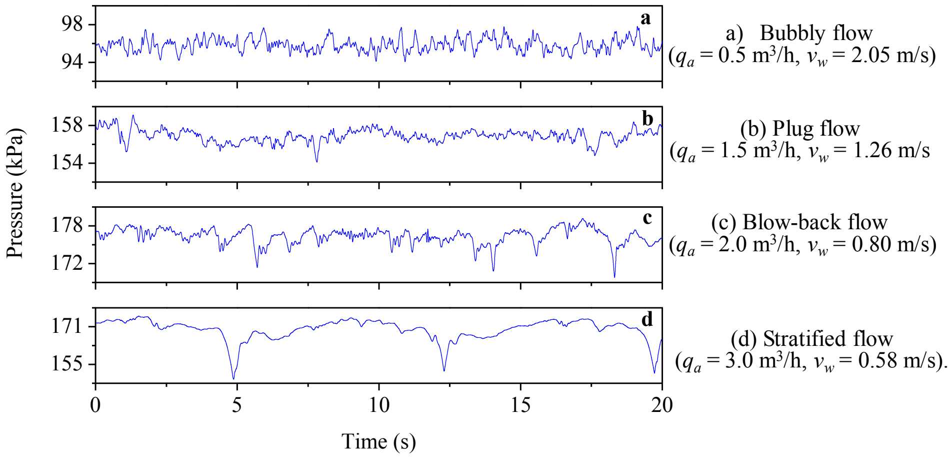


Fig.6 Time-domain pressure signals of different flow regimes

Fig. 7 shows the variances of all pressure signals collected in the tests. As shown in Fig. 7, the variances of all the pressure signals with respect to different flow regimes can be ranked in a descending order as follows: stratified flow, blow-back flow, plug flow and bubbly flow. That is because the transition in flow regime is closely related to the change of the air fraction, and it is generally accepted that bubbly flow, plug flow, blow-back flow and stratified flow appear sequentially in the downward sloping water pipe as the air fraction increases (Pothof and Clemens, 2011). Moreover, the pressure fluctuation of air-water flow will become increasingly intense with the increase of air content when the air fraction is lower than 50% (Riverin et al, 2006). According to the test conditions, the maximum air fraction in our tests was 25%. Thus, the amplitude of pressure fluctuation in stratified flow is the largest among the four regimes, which is followed by blow-back flow, plug flow and bubbly flow in order. As a result, the variance of pressure signal (in Table 1) can be used as an index (feature which can be input into the SVM) for flow regime identification.

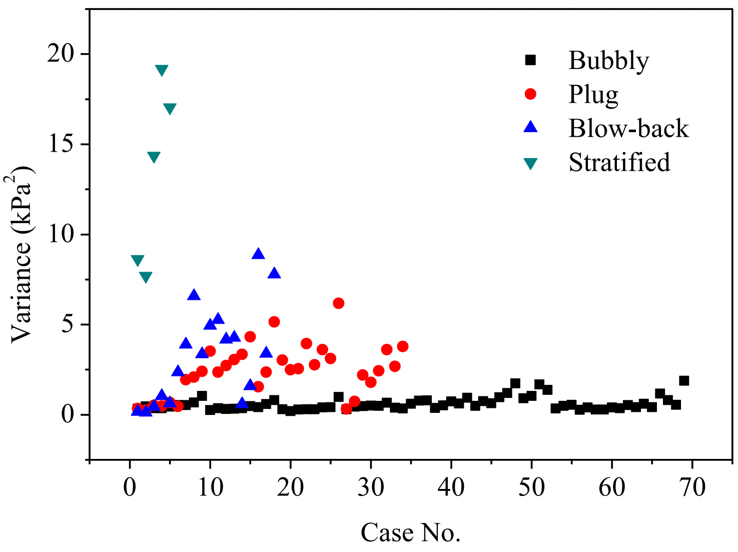


Fig. 7 Variances of all the pressure signals

4.1.2 Frequency-domain Characteristics of Pressure Signal

Fig. 8 shows the results of Fast Fourier Transform (FFT) on the pressure fluctuation signals according to the four typical cases in Fig. 6. The pressure fluctuation signals can be obtained by subtracting the mean values from the original pressure signal values.

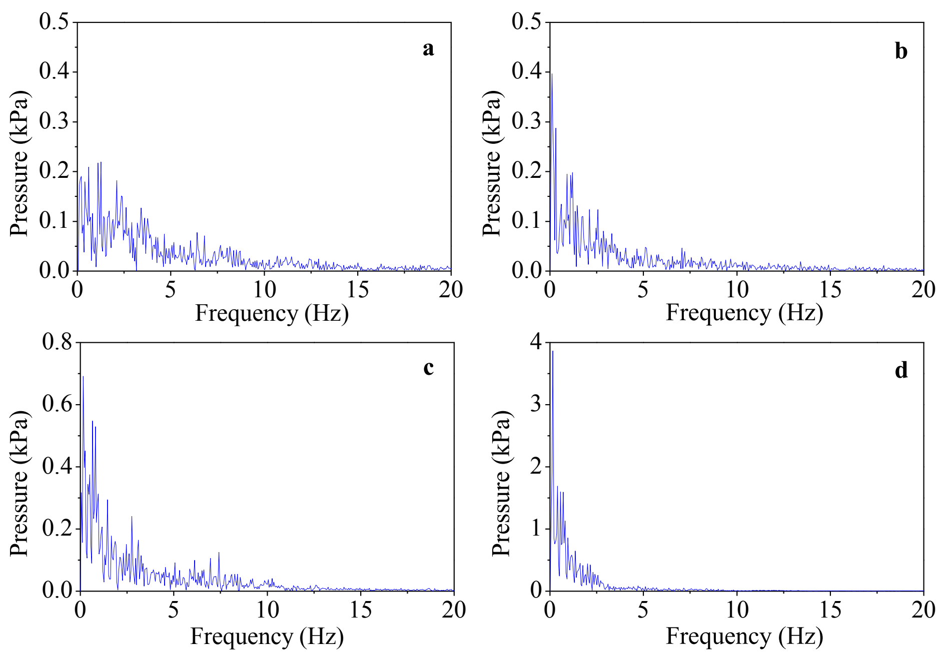


Fig. 8 Fourier spectra of pressure signals for different flow regimes (a) bubbly flow, (b) plug flow, (c) blow-back flow, (d) stratified flow.

As shown in Fig. 8, the frequency-domain distributions of four typical cases are mainly shown within 10 Hz, therefore the frequency-domain analysis on pressure signals should focus on this frequency range when the frequency-domain feature is used for flow regime identification. The frequency features of HHT and PSD used in this study are set with the frequency bands of 2 Hz and 3.9 Hz, respectively. Moreover, the frequency-domain distributions of pressure fluctuation signals show the single-peak characteristic in Fig. 8 except for bubbly flow (Fig. 8a), which indicates the pressure fluctuations of plug flow, blow-back flow and stratified flow have more significant periodicity in time-domain. It can also be seen in Fig. 8 that the single-peak characteristic becomes increasingly obvious and the pressure value of the peak increases when comparing from Fig. 8b to Fig. 8d. The main frequency components (i.e. pressure peak) are 0.2 KPa, 0.4 KPa, 0.7 KPa, 3.7 KPa in Figs 8a-8d, respectively. It indicates the periodical characteristic of the pressure signals based on the main frequency components becomes more significant as the air fraction increases. Consequently, the feature AC representing periodicity is used for flow regime identification (shown in Table 1).

4.2 SVM-based Flow Regime Identification

4.2.1 Effect of Pressure Signal Features on Identification Accuracy

The identification accuracy is defined as the ratio of the number of samples that are correctly classified into the specific flow regime by the SVM model to the total number of samples. Before studying the effect of pressure signal features on identification accuracy, the original pressure signals are first downsampled to 200 Hz, therefore, the length of one set of pressure signal (L) is 4000. The sample length (L') is set to 1000, which means the sampling time (the ratio of sample length to sampling rate) is 5 s. The interval between the adjacent samples (I) is 200, so every set of pressure signal can obtain 16 samples for SVM training or testing. The training data mentioned in Section 3.3 are used to train the SVM models, while the testing data are used to investigate the identification accuracy in SVM models. The identification accuracies of different features and their combinations for SVM-1 and SVM-2 are listed in Table 2.

Table 2. Flow identification accuracies corresponding to different features

Feature	SVM-1	SVM-2	Feature	SVM-1	SVM-2
SZR	89.6%	80.4%	SZR+V	89.8%	80.9%
V	82.5%	70.7%	AC+ PSD	79.0%	77.3%
AC	87.1%	73.3%	V+PSD	89.2%	87.1%
HHT	73.5%	83.1%	V+AC	88.1%	74.2%
PSD	89.2%	87.6%	SZR+AC+PSD	79.0%	76.0%
SZR+PSD	88.5%	85.8%	SZR+V+PSD	88.5%	85.8%
HHT+PSD	81.7%	73.8%	V+AC+PSD	78.8%	76.9%
SZR+AC	89.2%	83.1%	SZR+V+AC+PSD	78.8%	75.6%

As shown in Table 2, the comparative results show that SZR combined with V provide only a marginal improvement over using only SZR for SVM-1 classification, while the PSD is the optimal feature for SVM-2. Therefore, SZR and PSD are the best individual features for SVM-1 and SVM-2, respectively. The features related to HHT perform worse than other parameters in accuracy comparisons for both SVM-1 and SVM-2, and HHT calculation is computationally demanding. Thus HHT is not recommended to use for flow regime identification. As can be seen from Table 2, when comparing the identification accuracy of AC and AC-related feature combinations, SVM-1 is superior to SVM-2.

In order to explain the reason why SZR and PSD have better performance on identifying flow regimes as input features, the SZR values derived from all the original pressure signals in all cases are shown in Fig. 9 and the PSD distributions of four typical cases in Fig. 6 are demonstrated in Fig. 10. In general, the SZR values can be ranked in term of the four flow regimes in descending order in Fig. 9: bubbly flow, plug flow, blow-back flow and stratified flow. The SZR values in a few cases of bubbly flow are mixed with those of plug flow, and thus the identification accuracy in SVM-1 that is only used to classify bubbly flow cannot reach the higher values (the highest accuracy is 89.6%). SVM-2 is used to classify plug flow which can mix both bubbly flow and blow-back flow in Fig. 9, and thus the identification value of SZR decreases further (equal to 80.4%) when only SZR is used as input feature in SVM-2. Additionally, it is distinct between stratified and blow-back flows associated with SZR values, and however they cannot be an indication of flow regime transition in SVM due to the limited experimental samples.

Fig. 10 shows the PSD results of the pressure signals according to the four typical cases from Fig. 6. The PSD feature is a vector that consists of 129 eigenvalues based on the fixed frequency bands. The PSD distribution of stratified flow in the high frequency range (>50 Hz) shows periodical changes and distinguishes from other three flow regimes. The key difference between plug and blow-back flows in PSD distribution lies in the low frequency range (< 50 Hz). The bubble flow shows relative values of PSD when the frequency bands are larger than 100 Hz. Therefore, it indicates the PSD as input feature in SVM models would perform well. Although only four cases of different flow regimes are chosen here to show the difference in PSD distributions, other cases of PSD have also been investigated in this study.

The results show the similar rules associated with PSD. Based on the above analysis, SZR and PSD are effective as input features for flow regime identification in SVM models, and the two features will be used for the sampling parameters study in the following section.

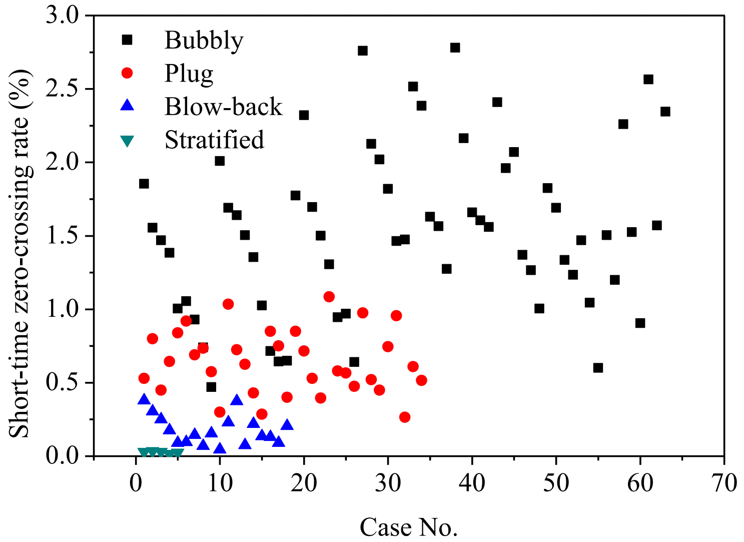


Fig. 9 Short-time zero-crossing rates (SZR) for all cases

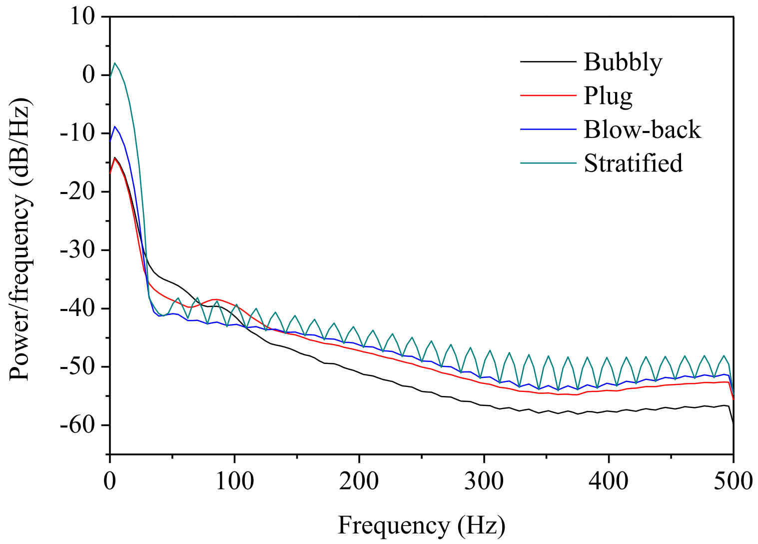
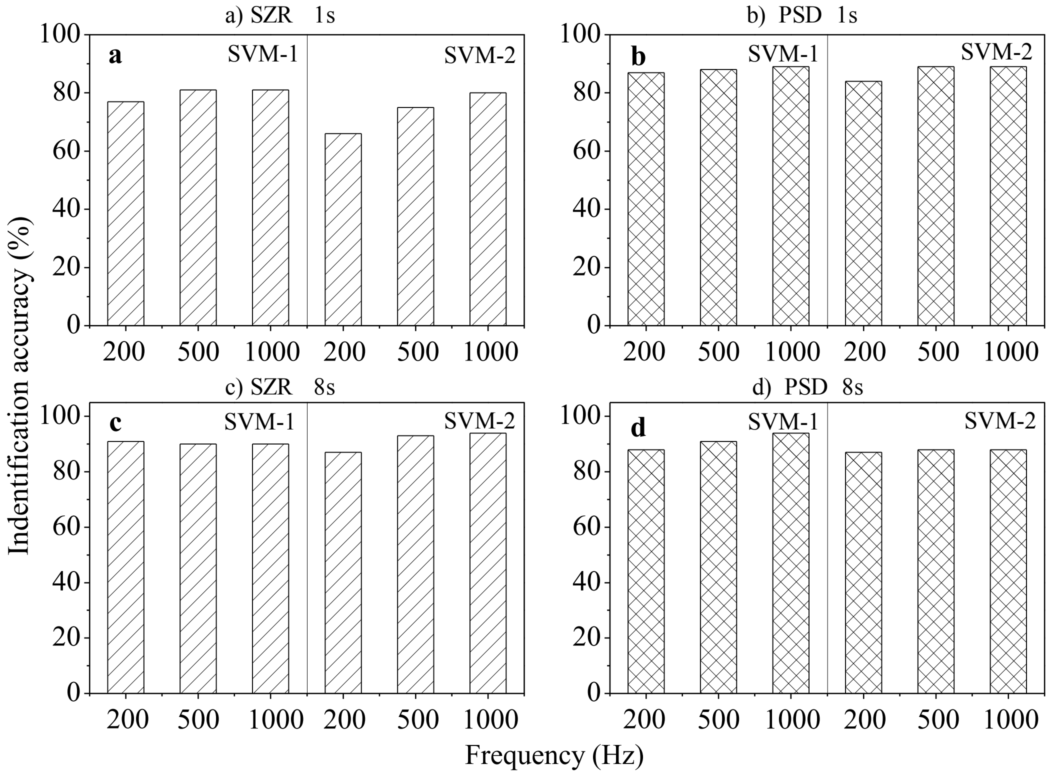


Fig. 10 PSD results of pressure signals for four typical cases

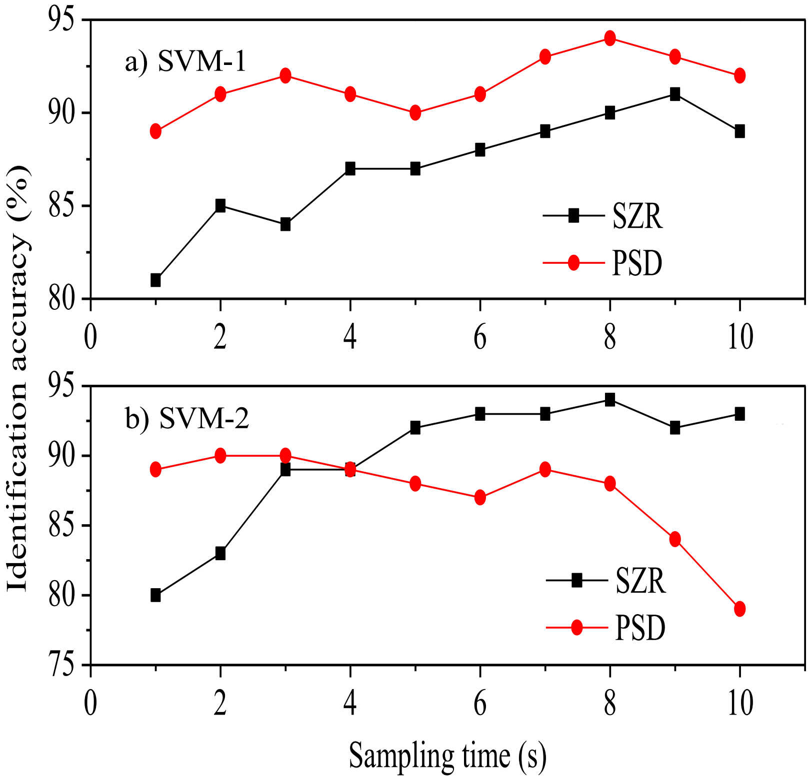
4.2.2 Impact of Sampling Parameters on Identification Accuracy

The sampling rate and sampling time can be adjusted by downsampling and changing the sample length (L'), respectively. With the features of SZR and PSD, the impact of sampling rate on identification accuracy in SVM models is investigated when sampling time (1 s or 8 s, for example) is given. In order to make a relatively large difference among all the samples, the overlapping proportion between the adjacent samples should be as small as possible, therefore the interval between the adjacent samples in this study is set to 1,000. Fig. 11 shows the effect of sampling rate on the flow regime identification accuracy. The identification accuracy is generally improved (i.e. accuracy value increases consecutively) with the increase in sampling rate. In some cases, sampling rate does not show significant impact on the identification accuracy (SVM-1 in Fig. 11c). In Fig. 11c, the identification accuracy increases from 87% to 94% for SVM-2 with the increase in sampling rate. In Fig. 11a, the identification accuracy has been improved from 66% to 80% when SZR is used as the feature. It demonstrates that the identification accuracy of SZR is more sensitive to the sampling rate than that of PSD, because higher sampling rates can amplify distinctions of the SZR values among different flow regimes. Therefore, the sampling rate of 1 kHz (the maximum value) is recommended for the applications of flow regime identification.



354

355 Fig. 11 Impact of sampling rate on identification accuracy



356

357 Fig. 12 Impact of sampling time on identification accuracy

In order to study the impact of sampling time on identification accuracy, the sampling rate is first fixed to 1 kHz, and the results with different features (SZR and PSD) are shown in Fig. 12. In Fig. 12a, the identification accuracies for both SZR and PSD show the rising trends with the increase of the sampling time in SVM-1, however the identification accuracy of PSD is better than that of SZR. Moreover, as mentioned before, the SZR values of some cases in bubbly flow are mixed with those in plug flow. Hence the feature SZR is not the best feature for SVM-1 since SVM-1 is built to classify bubbly flow from all flow regimes. As shown in Fig. 12b, the identification accuracy of SZR in SVM-2 ascends as the sampling time increases, however the PSD shows the opposite trend. As can be seen in Table 1, SZR is a time-domain feature that can reflect the pressure fluctuation of air-water flow, and SZR can distinguish plug flow from blow-back flow and stratified flow according to Fig. 9; while PSD can reflect the frequency-domain distribution of the pressure signal, which can clearly differentiate bubbly flow from the other three flow regimes by referring to Fig. 8 and Fig. 10. Therefore, it is concluded that PSD is the optimal feature for SVM-1, and SZR is the optimal feature for SVM-2 under various sampling time conditions. With the optimal features and sampling parameters (sampling rate: 1 kHz; sampling time: 8 s), the results of Fig. 12 show that the best identification accuracies for SVM-1 and SVM-2 are 94.3% and 93.9%, respectively.

4.3 Discussion

In water transmission pipelines, pressure signals can be obtained with a lower cost compared with void fraction (Arvoh et al, 2012; Lee et al, 2008b; Roshani et al, 2015; Salgado et al, 2010). If differential pressure signal is adopted as the indicator to evaluate the operation state of air valves in water pipes, it will be difficult to match an appropriate interval between the pressure taps, and therefore differential pressure signal cannot accurately predict flow regimes (Sun, 2005). Accordingly, pressure signal is the best indicator used to identify the operational state of air valves in water pipes.

The energy consumption of pressure signal collection with high frequency (> 100 Hz) can be ignored in the air-water two-phase flows monitoring (Elperin and Klochko, 2002; Sun, 2005). The sampling rate of 1 kHz is suggested to be used in practical applications, since the collected pressure signals will contain the essential information which could amplify the distinctions in SZR and PSD for different flow regimes.

According to the previous study by Lee et al. (2008a), instantaneous flow regime identification was applied to practical applications indicating that the sampling time of 1 s could identify the flow regimes successfully in the downward sloping pipe using the cross-sectional void fraction indicator. However, in this study the sampling time of 8 seconds is found to give the best identification accuracy using the pressure signal indicator. The possible reason is that only longer sampling time can reflect the dynamic characteristics of air-water pressure fluctuation in the flow regime transition.

Due to the limitation of the experimental facility, the training data cannot cover all kinds of pipe conditions and operational cases. The results of flow regime identification can be extended by more experimental data. Moreover, the void fraction signal can be used in the same methodology, if void fraction is measured via air valves. In that case, the better results of flow regime identification may be obtained by integrating pressure and void fraction measurements.

5 Conclusions

This paper proposed a novel SVM method to identify flow regimes using pressure signals. Flow regimes are closely related to air fraction in water pipes, and thus indicate operation state (functioning or malfunctioning) of air valves. An experiment facility involving a transmission pipeline is set up with the testing section that includes the pressure gauge on a segment of downward sloping pipe. Several features including Variance (V), Short-time Zero-crossing Rate (SZR), Autocorrelation Coefficient (AC), Hilbert-Huang Transform (HHT), Power Spectrum Density (PSD), are used to extract the information that represents fluctuations, periodicity and frequency distributions from the collected pressure signals. The combination of the features as inputs variables to the SVM models is investigated. Two SVM models are set up for classifying four flow regimes, and the parameters in the SVM classification are examined for improving the identification accuracy. The key conclusions are drawn:

1) For the analysis of collected pressure signals, the amplitude of pressure fluctuations for different flow regimes can be sorted into a descending ordered list as follows: stratified flow, blow-back flow, plug flow and bubbly flow. The periodicity of pressure fluctuations becomes more obvious as the air fraction increases, and the periodical characteristics of pressure signals are distinct among different flow regimes. In the analysis of frequency-domain distributions of pressure signals, the fluctuation frequency of four typical flow regimes occurs within 10 Hz. The four flow regimes show a strong variation trend in terms of pressure fluctuation.

2) SVM-1 is used to classify bubbly flow and then SVM-2 distinguish plug flow from blow-back and stratified flows. PSD and SZR are found to be the best features for the SVM classification through

the combination analysis. The SVMs perform well for flow regime identification using the suitable features as input variables.

3) The sampling parameters have a significant impact on the performance of SVM for identifying flow regimes. Based on the experimental data, the pressure sampling rate should be set relatively high (maximum 1 kHz) and the long sampling time is recommended to improve the identification accuracy of SVMs (e.g., 8 s in this study). With the best SVM features (i.e., PSD and SZR) and the optimal sampling parameters, the identification accuracy of SVM-1 and SVM-2 can reach 94.3% and 93.9%, respectively.

Data and method in this paper have been collected and tested based on an in-door experimental facility. The practical application of this flow regime identification method should be updated with the pressure data to train the SVM models. In the future, some field tests of the method should be done to validate the method. The high frequency pressure sensors which are used in the experiments should be adequately installed in the real pipelines to carry out this flow regime identification approach. Other sensors (e.g. acoustic or vibration sensors) are promising and can be used to extend this study.

Acknowledgments: This work was financially supported by the National Natural Science Foundation of China (Grant No. 91647201, 91747102, 51579027, 51708086). The experiment is conducted at Harbin Institute of Technology, where the staff have given a great support for the data collection. The Shenzhen Water group also provides partially financial support. The authors would like to thank the editors and three anonymous reviewers for the constructive comments which have substantially improved the quality of the paper.

References

Ahmadi A, Han D, Lafdani EK, Moridi A (2015) Input selection for long-lead precipitation prediction using large-scale climate variables: a case study. Journal of Hydroinformatics 17:114-129

Arvoh BK, Hoffmann R, Valle A, Halstensen M (2012) Estimation of volume fraction and flow regime identification in inclined pipes based on gamma measurements and multivariate calibration. Journal of Chemometrics 26:425-434

Aydogdu M, Firat M (2015) Estimation of failure rate in water distribution network using fuzzy clustering and LS-SVM methods. Water Resources Management 29:1575-1590

Barnea, D. (1986) Transition from annular flow and from dispersed bubble flow—unified models for the whole range of pipe inclinations. International journal of multiphase flow 12(5), 733-744.

Barnett, J., S. Rogers, M. Webber, B. Finlayson and M. Wang (2015). "Sustainability: transfer project cannot meet China's water needs." Nature News 527(7578): 295.

Becker, N., Easter, K.W., 1995. Water diversions in the great lakes basin analyzed in a game theory framework. Water Resour. Manag. 9, 221 – 242.

Bergmann, M., 1999. The Snowy Mountains Hydro-electric Scheme: How Did it Manage Without an EIA? Australian National University.

Bray M, Han D (2004) Identification of support vector machines for runoff modelling. Journal of Hydroinformatics 6:265-280

Boser B E, Guyon I M, Vapnik V N (1992). A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory. ACM, 144-152

Cai S, Toral H, Qiu J, Archer JS (1994) Neural network based objective flow regime identification in air - water two phase flow. The Canadian Journal of Chemical Engineering 72:440-445

Crawford T, Weinberger C, Weisman J (1985) Two-phase flow patterns and void fractions in downward flow Part I: steady-state flow patterns. International journal of multiphase flow 11:761-782

Ding H, Huang Z, Song Z, Yan Y (2007) Hilbert–Huang transform based signal analysis for the characterization of gas–liquid two-phase flow. Flow measurement and instrumentation 18:37-46

Elperin T, Klochko M (2002) Flow regime identification in a two-phase flow using wavelet transform Experiments in Fluids 32:674-682

Escarameia M Investigating hydraulic removal of air from water pipelines. In: Proceedings of the Institution of Civil Engineers-Water Management, 2007. vol 1. Thomas Telford Ltd, pp 25-34

Feng K, Jiang Z, He W, Ma B (2011) A recognition and novelty detection approach based on Curvelet transform, nonlinear PCA and SVM with application to indicator diagram diagnosis. Expert Systems with Applications 38:12721-12729

Franca F, Acikgoz M, Lahey R, Clausse A (1991) The use of fractal techniques for flow regime identification. International Journal of Multiphase Flow 17:545-552

Kalinske A, Bliss PH (1943) Removal of air from pipe lines by flowing water. Proceedings of the American Society of Civil Engineers (ASCE) 13:3

Kent JC (1952) The entrainment of air by water flowing in circular conduits with downgrade slopes. University of California, Lee JY, Ishii M, Kim NS (2008a) Instantaneous and objective flow regime identification method for the vertical upward and downward co-current two-phase flow. International Journal of Heat and Mass Transfer 51:3442-3459

Lee JY, Kim NS, Ishii M (2008b) Flow regime identification using chaotic characteristics of two-phase flow. Nuclear Engineering and Design 238:945-957

Lubbers CL (2007) On gas pockets in wastewater pressure mains and their effect on hydraulic performance vol 11. IOS Press,

Mariño, M.A., Loaiciga, H.A., 1985. Quadratic model for reservoir management: application to the Central Valley Project. Water Resour. Res. 21, 631 – 641.

Meng F, Fu G, Butler D (2016) Water quality permitting: From end-of-pipe to operational strategies Water research 101:114-126

Mi Y, Ishii M, Tsoukalas L (2001) Flow regime identification methodology with neural networks and two-phase flow models. Nuclear Engineering and Design 204:87-100

Modaresi F, Araghinejad S (2014) A comparative assessment of support vector machines, probabilistic neural networks, and K-nearest neighbor algorithms for water quality classification. Water resources management 28:4095-4111

Mounce SR, Mounce RB, Boxall JB (2011) Novelty detection for time series data analysis in water distribution systems using support vector machines. Journal of hydroinformatics 13:672-686

Pichler K, Schrems A, Buchegger T, Huschenbett M, Pichler M Fault detection in reciprocating compressor valves for steady-state load conditions. In: Signal Processing and Information Technology (ISSPIT), 2011 IEEE International Symposium on, 2011. IEEE, pp 224-229

Platt, John. (1998) Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines Advances in Kernel Methods-Support Vector Learning. 208.

Pothof I, Clemens F (2010) On elongated air pockets in downward sloping pipes. Journal of Hydraulic Research 48:499-503

Pothof I, Clemens F (2011) Experimental study of air–water flow in downward sloping pipes. International journal of multiphase flow 37:278-292

Pothof I, Clemens F (2012) Air pocket removal from downward sloping pipes. IAHR/IWA Joint Committee Urban Drainage, Working Group on Data and Models

Qin Q, Jiang Z-N, Feng K, He W (2012) A novel scheme for fault detection of reciprocating compressor valves based on basis pursuit, wave matching and support vector machine. Measurement 45:897-908

Ramezani, L., Karney, B. and Malekpour, A. (2015) The Challenge of Air Valves: A Selective Critical Literature Review. Journal of Water Resources Planning and Management 141(10), 04015017

Ring, M. and Eskofier, B.M. (2016) An approximation of the Gaussian RBF kernel for efficient classification with SVMs. Pattern Recognition Letters 84, 107-113

Riverin J, De Langre E, Pettigrew M (2006) Fluctuating forces caused by internal two-phase flow on bends and tees. Journal of sound and vibration 298:1088-1098

Roshani G, Nazemi E, Feghhi S, Setayeshi S (2015) Flow regime identification and void fraction prediction in two-phase flows based on gamma ray attenuation. Measurement 62:25-32

Salgado CM, Pereira CM, Schirru R, Brandão LE (2010) Flow regime identification and volume fraction prediction in multiphase flows by means of gamma-ray attenuation and artificial neural networks. Progress in Nuclear Energy 52:555-562

Schwaller, J and van Zyl, JE (2015) Modeling the Pressure-Leakage Response of Water Distribution Systems Based on Individual Leak Behavior. Journal of hydraulic engineering 141(5), 04014089

Stephens, M., Lambert, M., Simpson, A., Vítkovský, J. and Nixon, J. (2004) Field Tests for Leakage, Air Pocket, and Discrete Blockage Detection Using Inverse Transient Analysis in Water Distribution Pipes. Critical Transitions in Water and Environmental Resources Management.

Sun B (2005) The Identification Method of Gas-Liquid Two-Phase Flow Regime Based on Wavelet and Chaos Theory. North China Electric Power University

Sun Z, Shao S, Gong H (2013) Gas–liquid Flow Pattern Recognition Based on Wavelet Packet Energy Entropy of Vortex-induced Pressure Fluctuation Measurement. Science review 13:83-88

Tan C, Dong F, Wu M (2007) Identification of gas/liquid two-phase flow regime through ERT-based measurement and feature extraction. Flow Measurement and Instrumentation 18:255-261

Vapnik V (1963) Pattern recognition using generalized portrait method Automation and remote control 24: 774-780.

Verma NK, Kadambari J, Abhijit B, Tanu S, Salour A Finding sensitive sensor positions under faulty condition of reciprocating air compressors. In: Recent Advances in Intelligent Computational Systems (RAICS), 2011 IEEE, 2011. IEEE, pp 242-246

Vince M, Lahey R (1982) On the development of an objective flow regime indicator. International Journal of Multiphase Flow 8:93-124

Wisner PE, Kouwen N, Mohsen FN (1975) Removal of air from water lines by hydraulic means. Journal of the Hydraulics Division 101:243-257

Yang, T., Asanjan, A.A., Welles, E., Gao, X., Sorooshian, S. and Liu, X. (2017) Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information. Water Resources Research 53(4), 2786-2812

Yu, M., C. Wang, Y. Liu, G. Olsson and C. Wang (2018). Sustainability of mega water diversion projects: Experience and lessons from China. Science of The Total Environment 619-620: 721-731.

Zhu, Y., Duan, H.F., Li, F., Wu, C.G., Yuan, Y.X. and Shi, Z.F. (2018) Experimental and numerical study on transient air–water mixing flows in viscoelastic pipes. Journal of Hydraulic Research 56(6), 877-887.