



Title	Abnormal Sound Detection in Pipes Using a Wireless Microphone and Machine Learning
Author(s)	Notani, Kota; Hayashi, Takahiro; Mori, Naoki
Citation	Materials Transactions. 2022, 63(12), p. 1622-1630
Version Type	VoR
URL	<a href="https://hdl.handle.net/11094/89937">https://hdl.handle.net/11094/89937</a>
rights	© 2022 The Japanese Society for Non-Destructive Inspection
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

# Abnormal Sound Detection in Pipes Using a Wireless Microphone and Machine Learning

Kota Notani, Takahiro Hayashi\* and Naoki Mori

Graduate School of Engineering, Osaka University, Suita 565-0871, Japan

Abnormal sound detection using a one-class support vector machine (OCSVM) and a principal component analysis (PCA) is proposed aiming to stable and objective inspection without skilled plant inspectors. For measurement of acoustic signals, we developed a compact microphone unit that can work in sound detection, signal transmission, and power supply, wirelessly. Six signal parameters were extracted as features from filtered and segmented acoustic signals. Using the features standardized and reduced in dimensionality by PCA, an anomaly detection model using OCSVM is built to detect abnormal sounds. The proposed method is verified by acoustic diagnosis of sound waves leaking from pipeworks with running water. Diagnostic accuracies were evaluated for artificial abnormal sounds with different types of burst waves output from a piezoelectric element attached to the pipe and Pencil Lead Break sound in water flowing background noise. Burst wave changes could be detected in almost all patterns, and the diagnostic accuracy was 100% for the Pencil Lead Break sound.

[doi:10.2320/matertrans.MT-I2022001]

(Received May 16, 2022; Accepted September 12, 2022; Published October 28, 2022)

**Keywords:** nondestructive inspection, pipes, abnormal sound detection, machine learning, wireless microphone

## 1. Introduction

The maintenance of the countless pipes in a plant is an extremely heavy burden for plant users. In particular, pipe inspection at high altitudes, high temperatures, and high radiation levels, which seems to be unproductive work, requires a great deal of work, such as shutting down plant operations, setting up scaffolding, and careful inspection.

One of the authors has been studying a remote inspection technique for piping in plants to obtain defect images inside metal pipes by measuring elastic waves that propagate through the pipes and leak into the air when they are irradiated by a laser.<sup>1–5)</sup> By using a laser Doppler vibrometer to receive the elastic waves, it is possible to make measurements remotely, which is expected to make inspection work very efficient. Furthermore, it has been shown that a MEMS microphone used in smartphones can be used as a device to detect elastic waves leaking from the pipe surface.<sup>6)</sup> This enables stable detection of elastic waves even when there is oscillation, such as in outdoor pipes. Because this MEMS microphone unit is small, inexpensive, and can be used wirelessly by transmitting acoustic signals via Bluetooth, it is expected to have a wide range of applications in addition to the defect imaging described above. The purpose of this study is to improve this microphone device to enable wireless power supply and data transmission, and to investigate the possibility of detecting abnormal sounds appearing in pipes, with a view to expanding plant maintenance technology using the MEMS microphone unit.

In factories and plants, workers unconsciously sense their surroundings from the vision, sound, heat, and smell to determine if there are any abnormalities in their daily work. Among them, auditory information is as easy for humans to perceive as visual information, and we often detect abnormal sounds based on our experience in the workplace. In recent years, because the performance of audio microphones has

been dramatically improved by MEMS technology, sound as well as visual information has become reasonable means for diagnosis of factories and plants. However, even if sound can be detected with high sensitivity, it is difficult to distinguish small abnormal sound emitted by particular equipment and machines from various noise and useful sounds in the surrounding environment, and thus inspectors who are familiar with the site is needed to diagnose the presence of abnormalities. On the other hand, the number of skilled inspectors is limited and they cannot judge all sounds recorded from a wide range of factories. Furthermore, the inspection results are also affected by variations of inspectors, physical condition, and the work environment due to the subjective human sense, making it difficult to obtain stable and objective results. In addition, transferring the skill of distinguishing abnormal sounds is not easy because of individual differences of sound detection.

Therefore, machine learning for abnormal sound detection is expected to be an alternative way of the skilled personnel. Machine learning is an algorithm that analyzes and learns patterns in training data to classify and predict unknown test data. Recent advances in computers have made it possible to collect and process large amounts of data, and machine learning is used not only in academia but also in a wide range of practical fields such as medical, finance, marketing, and engineering. The problem of detecting anomalies in data, such as detecting abnormal sounds among many acoustic signals, is called anomaly detection, which includes two types of techniques in machine learning: supervised learning and unsupervised learning.<sup>7)</sup> Supervised learning builds an anomaly detection model based on known normal and abnormal data, which leads highly accurate classification and prediction. Typical algorithms include support vector machine (SVM), k-nearest neighbor (kNN), logistic regression (LR), etc. Pipeline leak detection, bearing failure diagnosis, and wind turbine blade damage detection have been proposed using these algorithms.<sup>8–14)</sup> However, supervised learning is difficult to detect unexpected abnormal sounds that cannot be prepared in advance. In contrast,

\*Corresponding author, E-mail: hayashi@mech.eng.osaka-u.ac.jp

unsupervised learning does not require preparing anomaly data and builds an anomaly detection model from data measured in appropriate operations. Typical algorithms include the Gaussian mixture model (GMM), self-organizing map (SOM), one-class support vector machine (OCSVM), etc. As with supervised learning, many studies on anomaly detection using unsupervised learning have been published.<sup>15–19</sup> In addition, both supervised and unsupervised algorithms exist for a machine learning called neural network (NN). In recent years, anomaly detection using advanced forms of NN such as autoencoder (AE), recurrent neural network (RNN), and convolutional neural network (CNN) has been actively studied.<sup>20–22</sup>

In this study, we develop an abnormal sound detection using OCSVM,<sup>23</sup> an unsupervised learning non-hierarchical clustering method, intending to detect unexpected abnormal sounds. OCSVM requires less processing power than NN and can be used to detect abnormal sounds. Before building the model, principal component analysis (PCA)<sup>24</sup> is used to reduce the dimensions of the data based on the cumulative contribution ratio. To verify the effectiveness of abnormal sound detection by machine learning with the wireless microphones and OCSVM, acoustic diagnosis of sound waves leaking from pipeworks with running water is performed. In addition to the sound of water flowing through pipeworks, burst waves are incident from a piezoelectric element attached to the pipe, and by changing the parameters of the burst waves, various types of acoustic data are prepared and used for acoustic diagnosis. Furthermore, abnormal sound detection is also performed for the sound of a mechanical pencil lead pressed and broken on the pipe surface (Pencil Lead Break).

## 2. Development of a Wireless Microphone Unit

In the previous study by one of the authors about defect imaging in plate-like structures using a scanning laser source technique, the acoustic waves leaking from the flat plate were measured by a small microphone unit.<sup>6</sup> The microphone unit is a palm-sized prototype that consists of two MEMS microphones, amplifiers, and a battery. By connecting a commercially available Bluetooth transmitter for audio to the acoustic signal output terminal, the measured waveforms can be transmitted wirelessly. However, the battery needed to be recharged for each test, which implies we cannot avoid laborious work of collecting the unit periodically.

In this study, a new microphone unit that can be remotely powered by a solar cell is used (Fig. 1). This microphone unit measures approximately  $45 \times 45 \times 20 \text{ mm}^3$  and is equipped

with two MEMS microphones (Knowles, SPU0410LR5H) as acoustic sensors. External monocrystalline solar cells (Anysolar Ltd., SM101K12L) can be used to remotely supply power and charge the built-in battery (Data Power Technology, DPT502535). In addition, a Bluetooth 3.0 module (Silicon Labs, WT32I-E) is built in to transmit the measured acoustic signals, allowing the device to be placed anywhere without the hassle of wiring, within the range of irradiation to the solar cells and Bluetooth communication. This microphone unit requires 150 mW during measurements. On the other hand, two solar cells shown above were used in this study, and they output totally about 500 mW under the Sun. In the laboratory tests, two blue semiconductor lasers with a measured output power of about 4.4 W was used as the light source to confirm device activation and acoustic signal measurement from a distance more than 10 m. The internal battery required to realize stable power supply and stable signal measurements. The communication profile is the Advanced Audio Distribution Profile (A2DP), which is mainly used for audio playback with wireless earphones or headphones. The MEMS microphones can measure up to about 100 kHz, but because data is transmitted using the Sub Band Codec (SBC) compression method, the sample rate of the acoustic signal is limited to 44.1 kHz.

## 3. Acoustic Diagnosis Procedure

In this study, an anomaly detection model is constructed by OCSVM with training data of acoustic signals, and test sound data is classified into normal or abnormal data using the model. The acoustic signals used for the training and test data are measured by microphones. The acoustic diagnosis procedure is summarized in Fig. 2.

As a preprocessing of the acoustic signals, band-pass filtering with different cutoff frequencies is performed first to eliminate noise and to ensure diversity in the features to be extracted later. Hereafter, letting the number of band-pass filters and the Nyquist frequency of the acoustic signal be  $n_f$  and  $f_{nyq}$ , respectively, two band-pass filtering techniques with different methods of cutoff frequency determination are considered; one is uniform bandwidth of the pass band, and the other is non-uniform bandwidth with the cutoff frequencies of Mel filter bank. Since Mel filter bank refers to the human auditory perception and widely used in speech recognition and other applications, anomaly detection using the filter is also investigated in this study.

In the uniform bandwidth, band-pass filters with a constant pass band width  $f_{bw}$  are equally spaced from DC to  $f_{nyq}$ . The filter spacing  $\Delta f$  is expressed as

$$\Delta f = \frac{f_{nyq} - f_{bw}}{n_f - 1}. \quad (1)$$

In the non-uniform bandwidth, the cutoff frequency of the band-pass filter is determined with reference to a Mel filter bank. A Mel filter bank is a set of triangular windows in the frequency domain, in which the edges of triangular windows are located at the constant interval  $\Delta m$  on the Mel scale. In this study, these window-edges are set to the cutoff frequencies. The conversion between the frequency  $f$  Hz and



Fig. 1 Microphone unit.

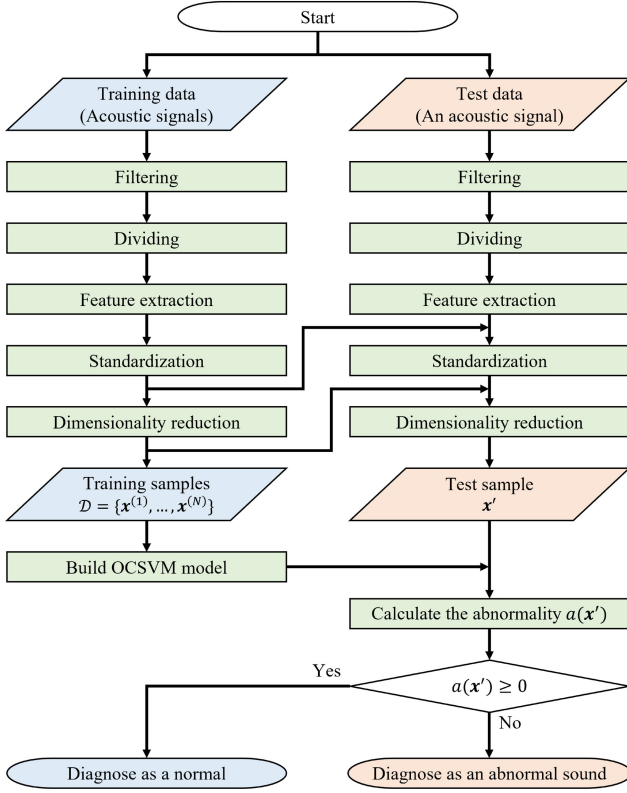


Fig. 2 Flow of acoustic diagnosis.

the Mel scale  $m$  mel is expressed by the following eqs. (2) and (3),<sup>25)</sup> and  $\Delta m$  can be calculated by eq. (4).

$$m \text{ [mel]} = \mathcal{M}(f) = 2595 \log_{10} \left( 1 + \frac{f \text{ [Hz]}}{700} \right), \quad (2)$$

$$f \text{ [Hz]} = \mathcal{M}^{-1}(m) = 700 \left( 10^{\frac{m \text{ [mel]}}{2595}} - 1 \right), \quad (3)$$

Table 1 Passbands of filters /Hz.

	Uniform bandwidth		Non-uniform bandwidth	
	Lower cutoff	Higher cutoff	Lower cutoff	Higher cutoff
1	0	$f_{bw}$	0	$\mathcal{M}^{-1}(2\Delta m)$
2	$\Delta f$	$\Delta f + f_{bw}$	$\mathcal{M}^{-1}(\Delta m)$	$\mathcal{M}^{-1}(3\Delta m)$
3	$2\Delta f$	$2\Delta f + f_{bw}$	$\mathcal{M}^{-1}(2\Delta m)$	$\mathcal{M}^{-1}(4\Delta m)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n_f - 1$	$(n_f - 2)\Delta f$	$(n_f - 2)\Delta f + f_{bw}$	$\mathcal{M}^{-1}((n_f - 2)\Delta m)$	$\mathcal{M}^{-1}(n_f \Delta m)$
$n_f$	$(n_f - 1)\Delta f$	$f_{nyq}$	$\mathcal{M}^{-1}((n_f - 1)\Delta m)$	$f_{nyq}$

$$\Delta m \text{ [mel]} = \frac{\mathcal{M}(f_{nyq})}{n_f + 1}, \quad (4)$$

where  $\mathcal{M}$  denotes a conversion function from  $f$  Hz to  $m$  mel. Table 1 shows the cutoff frequencies of  $n_f$  Butterworth band-pass filters determined in the uniform and non-uniform bandwidth.

The filtered acoustic signal of time length  $t_L$  is segmented by cutting out with rectangular windows that have a time width of  $t_w$  and slide by  $t_w/3$ , as in Fig. 3, in which typical signals detected with the microphone unit for water flow and burst waves were shown. The number of divisions  $n_d$  (integer) of the acoustic signal is expressed as

$$n_d = \frac{3t_L}{t_w} - 2. \quad (5)$$

Then, the number of samples increases  $n_d$  times.

Six signal parameters shown in Table 2 are extracted as features from each segmented time-domain waveform and its frequency-domain spectrum obtained by Fast Fourier Transform (FFT). To remove discontinuities at the edges of the segmented waveforms, a Hanning window is applied before FFT. Table 2 shows the features to be extracted, where  $x_i$ ,

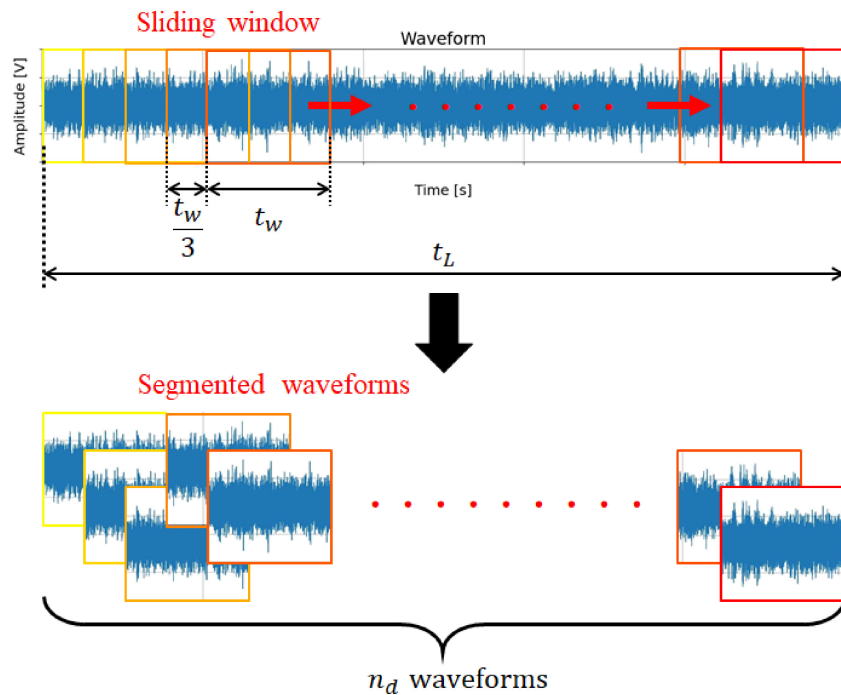
Fig. 3 Division into  $n_d$  waveforms with the sliding window at time duration of  $t_w$  from an acoustic signal.

Table 2 Equation of features in time domain and frequency domain.

Domain	Feature	Equation
Time	RMS	$F_1 = \sqrt{\frac{1}{m} \sum_{i=1}^m x_i^2}$
	Crest factor	$F_2 = \frac{\max_{i \in \mathbb{m}}  x_i }{F_1}$
Frequency	Peak	$F_3 = \max_{i \in \mathbb{M}} X_i$
	Peak frequency	$F_4 = i \text{ s.t. } X_i = F_3$
	Median frequency	$F_5 = \arg \min_{k \in \mathbb{M}} \left( \sum_{i=1}^k X_i - \sum_{j=k+1}^M X_j \right)$
	Relative peak	$F_6 = \max_{i \in \mathbb{M}} \frac{X_i}{\bar{X}_i}$

$X_i$  and  $\bar{X}_i$  are the values at the  $i$ -th sample point of the waveform, frequency spectrum of the acoustic signal, and average frequency spectrum of training data.  $m$  and  $M$  are the numbers of sample points of the waveform and spectrum, and  $\mathbb{m}$  and  $\mathbb{M}$  denote the set of natural numbers from 1 to  $m$  and  $M$ , respectively.  $\mathbb{M}'$  is the set of indexes of the sample points of the spectrum within the passband of the filter. Although it is not clear whether the six features in Table 2 are sufficient, since the time position of the signal is separated by the cutout of waveforms and the frequency is separated by the filter, features related to signal intensity and shape of wave were employed.

Because six signal parameters are extracted as features for each segmented and filtered data, the number of components in a feature vector for one segmented sample waveform is  $6n_f$ . The numbers of segmented waveforms and measured waveforms for training are  $n_d$  and  $N$ , respectively, then the number of samples becomes  $Nn_d$ . Therefore, the feature data is expressed by a  $6n_f \times Nn_d$  matrix. Since each of the  $6n_f$  features has a different mean and variance, each row of the feature data is standardized to have mean 0 and variance 1. After standardization, dimension is reduced by PCA so that a cumulative contribution ratio becomes 90%. This greatly reduces the computational load and allows a small computer to determine even large amounts of acoustic data in the inspection site.

Based on the training data obtained in the above procedure, an anomaly detection model using OCSVM with a Gaussian kernel is built. The built model is then used to diagnose the acoustic signals in the test data. Since  $n_d$  segmented waveforms are obtained per an acoustic signal, the number of calculated abnormality degrees  $\alpha$  is also  $n_d$ . The average of these values is considered as the abnormality degree of the acoustic signal, and if the value is negative, the acoustic signal is judged to be an anomaly.

Detection rate and recall are used to calculate diagnostic accuracy. Detection rate is the probability that the test data

measured under the same conditions as the training data was correctly judged to be normal. Recall is the probability that the test data having abnormal sounds was correctly judged to be anomaly. In this study, the product of detection rate and recall is defined as the diagnostic accuracy.

#### 4. Acoustic Diagnosis Experiment of Pipeworks

The acoustic diagnosis was performed according to the procedure stated previously, and the accuracy was examined. Assuming that abnormal sound is detected in the pipeworks during operation of the facility, a microphone unit described below was used for detecting abnormal sound in leakage sound from pipeworks that are located away from the inspectors in wireless measurement. A wide variety of abnormal sounds are considered in inspection site, such as gas leakage, mechanical failure, sudden pressure drop in pressure vessel and pipes, and other unexpected noise. In these laboratory tests, various burst waves from a piezo-electric element attached to the pipe were generated as abnormal and normal sounds because they are precisely changeable, and Pencil Lead Break sounds were also used as pseudo abnormal sounds to confirm whether they could be detected as abnormal sounds.

##### 4.1 Experimental method

Figure 4 shows the mock-up pipework in the laboratory with an aluminum alloy straight pipe one meter long connecting to water supply. The aforementioned microphone unit was placed on an aluminum straight pipe with the outer diameter of 114.3 mm and thickness of 6.0 mm to measure acoustic signals, where microphone holes were facing to the pipe surface, and therefore most of detected sounds were leaking sounds from the small area of pipe surface under the microphone holes after propagation through the pipe. As shown in Fig. 4, the microphone unit can be remotely powered by solar cells, but a 5 V DC power supply was connected in these laboratory tests. The acoustic signals are measured in 10 s with 44.1 kHz sample rate and 16 bits. During the measurement, burst waves generated by a function generator (NF Corporation, WF1944) were output

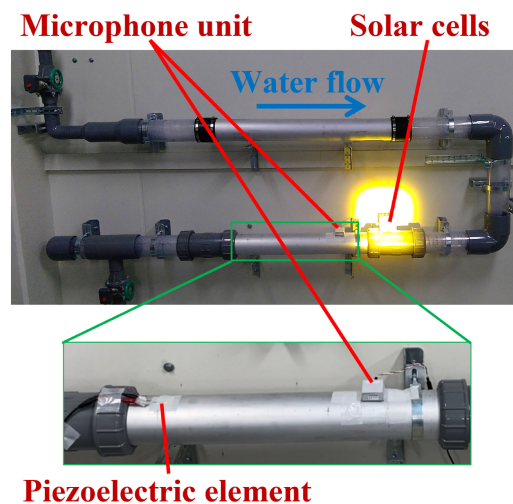


Fig. 4 Mock-up pipework and experimental equipment.



Table 3 Parameters of burst waves and the number of measurements.

Data label	Frequency / kHz	Amplitude / V	Interval / ms	Cycles / times	Number of measurements
BURST NONE		0			150
BURST 1	8	5	300	50	150
BURST 2	7	5	300	50	150
BURST 3	9	5	300	50	150
BURST 4	8	5	200	50	150
BURST 5	8	5	2000	50	150
BURST 6	8	5	300	20	150
BURST 7	8	5	300	100	150

from a piezoelectric element attached to the aluminum alloy part, with 8 different burst wave parameters. The burst wave parameters include frequency, amplitude, interval, and number of cycles, as summarized in Table 3. Among the measured data, only BURST NONE is the acoustic data of water flow in the pipework measured without burst wave output from the piezoelectric element. BURST 1–7 contain the burst signals from the piezoelectric element with slightly different parameters as written in Table 3.

The acoustic diagnosis was performed as follows, considering the water flow sound as background noise and the change of burst wave emitted from the piezoelectric element as normal and abnormal sound. First, 80% of each measured data (120 acoustic signals) was used as training data, and an anomaly detection model using OCSVM was built from features extracted from the training data. Then, the remaining 20% of the data (30 acoustic signals) and different types of measurement data were used as test data, whose abnormality was determined from the model built from the training data. In addition, we investigated the effect on the diagnostic results when changing the parameters such as the number of band-pass filters  $n_f$ , the division time width  $t_w$ , and the determination methodologies of the cutoff frequency in

the feature extraction process. LabVIEW (National Instruments) was used to process the acoustic diagnosis. The settings of OCSVM were manually adjusted by observing the results for two parameters,  $\nu$  and  $\sigma$ , which determine the percentage of outliers in the training data and the constant of the Gaussian kernel, respectively.

In order to investigate more practical abnormal sound detection, sound of Pencil Lead Break, which is a standard technique to produce an artificial acoustic emission source, was measured. As an acoustic signal with an abnormal sound, 50 measurements were made with the Pencil Lead Break as well as the background noise of water. To confirm whether the Pencil Lead Break sound can be judged as an abnormal sound, an acoustic diagnosis was performed using an anomaly detection model built with BURST NONE as training data, with the remaining acoustic signal of BURST NONE and Pencil Lead Break sound as test data.

## 4.2 Experimental results and discussions

### 4.2.1 Uniform bandwidth with $n_f = 20$ , $f_{bw} = 10$ kHz, and $t_w = 1$ s

For the uniform bandwidth with the number of band-pass filters  $n_f = 20$ , the passband width  $f_{bw} = 10$  kHz, and the division time width  $t_w = 1$  s, Table 4 shows the results of acoustic diagnosis for all patterns when 8 types of measurement data were used as both training data and test data, respectively. The numbers in Table 4 represent the diagnostics accuracy, which is the product of detection rate and recall. For example, when BURST 1 is used as training data and BURST 2 is test data, the diagnostic accuracy was 100%. This means that the model built with the training data of 120 acoustic signals measured when the burst wave with the parameters of BURST 1 is output, as shown in Table 3, can judge all 30 acoustic signals of remaining BURST 1 as normal data, and all 150 acoustic signals of the burst wave with the parameter of BURST 2 as an anomaly. The diagnostic accuracy was 100% for almost all patterns in Table 4, indicating that acoustic signals of the same type as the training data were normal and acoustic signals with different burst wave parameters from the training data were

Table 4 Diagnostic accuracy of section 4.2.1 ( $n_f = 20$ ,  $f_{bw} = 10$  kHz,  $t_w = 1$  s,  $\nu = 0.085$ ,  $\sigma = 0.0302$ ).

Accuracy %		Test data							
		BURST NONE	BURST 1	BURST 2	BURST 3	BURST 4	BURST 5	BURST 6	BURST 7
Training data	BURST NONE		100	100	100	100	100	100	100
	BURST 1	100		100	100	100	100	100	99
	BURST 2	100	100		100	100	100	100	100
	BURST 3	100	100	100		100	100	100	100
	BURST 4	100	100	100	100		100	100	100
	BURST 5	1	100	100	100	100		33	100
	BURST 6	100	100	100	100	100	100		100
	BURST 7	100	100	100	100	100	100	100	

detected as abnormal sounds. However, the accuracy of test data of BURST NONE and BURST 6 were 1% and 33%, respectively, for the training data of BURST 5.

Now, we consider the reasons why the acoustic signals of BURST NONE and BURST 6 could not be detected as abnormal sounds for the training data of BURST 5. The reason why the model built using BURST 5 as training data did not detect BURST NONE as an abnormal noise is evident from the relationship between the time width of the segmented waveform and the interval of the burst wave. The interval of BURST 5 shown in Table 4 is 2000 ms ( $= 2$  s), whereas the division time width is  $t_w = 1$  s. Therefore, the segmented waveforms in the acoustic signals of BURST 5 sometimes contain a burst wave and sometimes not. The segmented waveforms without a burst wave can be regarded as the same as the segmented waveform of BURST NONE. Consequently, the BURST NONE in test data is sometimes misjudged with the BURST 5. It is also obvious that the model built with BURST 5 as the training data seems to diagnose normal for the acoustic signal of BURST NONE, and the accuracy of 1% is a reasonable result. In this pattern, the accuracy can be improved by setting a larger value for the division time width  $t_w$  so that all segmented waveforms contain burst waves.

The low accuracy for BURST 5 of training data and BURST 6 of test data is considered from the viewpoint of principal components. As a result of the PCA, 21 principal components were selected based on the cumulative contribution ratio. The distributions of 21 principal components were extracted from the training and test data as shown in Fig. 5(a). For comparison, Fig. 5(b) shows the distributions of the principal components when BURST 6 is used as training data and BURST 5 as test data. The colors

in Fig. 5 represent the density of the distribution, with red representing high density and blue low density. The left distributions in Figs. 5(a) and (b) for training data show the characteristic of PCA that the variation of the principal components decreases as their number increases. Because the first few principal components of the test data have wide distributions, only the first and second principal components with large variations were extracted to visualize the distribution of their features as shown in Fig. 6. The blue circle in the left figures and red crosses in the right figures show training data and test data, respectively. In Fig. 6(a), most of the blue circles are enclosed by the black dotted line representing the dominant area of the training data, and the enclosed area also includes the red crosses. Because OCSVM establishes class boundaries so that most of the training data is normal, it could not judge the test data of BURST 6 in the enclosed area as an anomaly. This is because segmented waveforms in BURST 5 include the data both with and without the burst wave, resulting in a wider distribution of features. In contrast, Fig. 6(b) shows that training data shown in blue circles with an enclosed line and test data in red crosses are not in the same region, although some area is overlapped. In this case, the test data was successfully judged to be anomaly because the abnormality degree of the data out of the enclosed area greatly affect the abnormality degree of the entire acoustic signal.

#### 4.2.2 Uniform bandwidth with $n_f = 20$ , $f_{bw} = 10$ kHz, and $t_w = 2$ s

Table 5 is the diagnostic accuracy when the number of band-pass filters  $n_f$  and passband width  $f_{bw}$  are the same as in section 4.2.1 and the division time width  $t_w$  is 2 s. Table 5 shows that the accuracy improved significantly to 100% when the training data was BURST 5, and the test data was

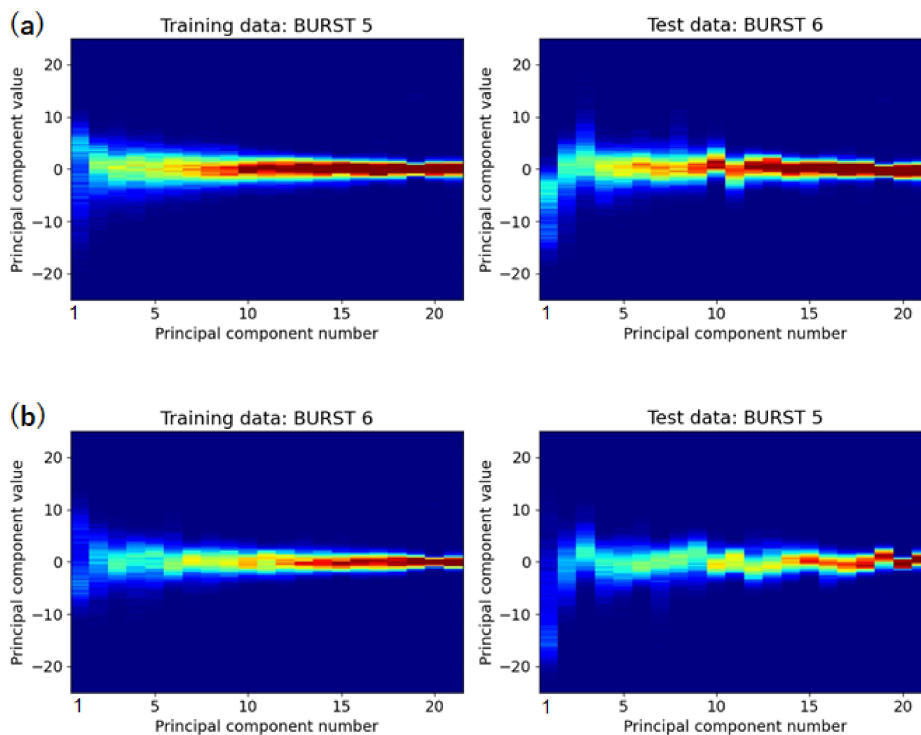


Fig. 5 Distributions of the principal components of BURST 5 and BURST 6: (a) training data is BURST 5 and (b) training data is BURST 6.

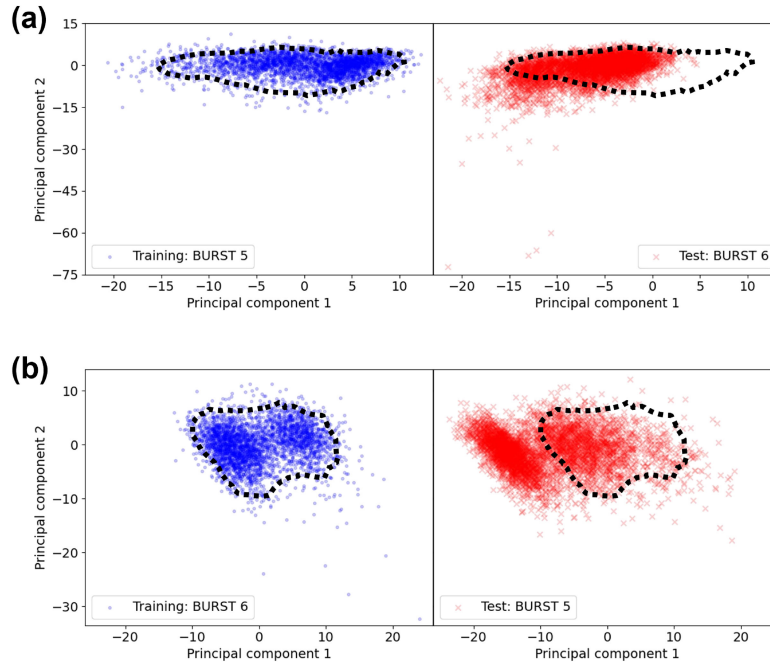


Fig. 6 Distributions of the first and second principal components of BURST 5 and BURST 6: (a) training data is BURST 5 and (b) training data is BURST 6.

Table 5 Diagnostic accuracy of section 4.2.2 ( $n_f = 20$ ,  $f_{bw} = 10$  kHz,  $t_w = 2$  s,  $\nu = 0.02$ ,  $\sigma = 0.013$ ).

Accuracy %		Test data							
		BURST NONE	BURST 1	BURST 2	BURST 3	BURST 4	BURST 5	BURST 6	BURST 7
Training data	BURST NONE		100	100	100	100	100	100	100
	BURST 1	100		100	100	100	100	99	7
	BURST 2	100	100		100	100	100	100	100
	BURST 3	100	100	100		100	100	100	100
	BURST 4	100	100	100	100		100	100	100
	BURST 5	100	100	100	100	100		41	100
	BURST 6	100	100	100	100	100	87		100
	BURST 7	100	100	100	100	100	100	100	

BURST NONE. The time width of the segmented waveform has become sufficiently long compared to the interval of the burst wave so that burst waves are now present in all segmented waveforms, which brought this improvement. However, the accuracy did not improve much in BURST 5 as training data and BURST 6 as test data. The accuracy in BURST 1 as training data and BURST 7 in test data was also very poor at 7%, but the accuracy in BURST 7 as training data and BURST 1 as test data was 100%, which is because the distribution of BURST 1 included the smaller distribution of BURST 7.

Thus, even if the diagnostic accuracy can be improved for a particular pattern by adjusting the parameters of model building, it may affect the other patterns. The OCSVM used in this study is an unsupervised learning algorithm that does

not include class labels in the training data, and since the algorithm detects outliers in the data, it often fails to detect anomalies close to the distribution of the training data. If the characteristics of the data can be quantified in a higher dimension and a more comprehensive manner through more creative preprocessing of acoustic signals and extracted features, the accuracy of certain patterns may be improved with keeping high accuracy for almost all patterns.

#### 4.2.3 Uniform bandwidth with $n_f = 50$ , $f_{bw} = 10$ kHz, and $t_w = 1$ s

Table 6 is the diagnostic accuracy when the number of band-pass filters is set to 50, while the passband width and division time width are the same as in section 4.2.1. Table 6 shows that the model using BURST 5 as training data is inaccurate in diagnosing BURST NONE and BURST 6,



Table 6 Diagnostic accuracy of section 4.2.3 ( $n_f = 50$ ,  $f_{bw} = 10$  kHz,  $t_w = 1$  s,  $\nu = 0.094$ ,  $\sigma = 0.0137$ ).

Accuracy %		Test data							
		BURST NONE	BURST 1	BURST 2	BURST 3	BURST 4	BURST 5	BURST 6	BURST 7
Training data	BURST NONE		100	100	100	100	100	100	100
	BURST 1	97		97	97	97	97	97	94
	BURST 2	97	97		97	97	97	97	97
	BURST 3	100	100	100		100	100	100	100
	BURST 4	100	100	100	100		100	100	100
	BURST 5	3	100	100	100	100		37	100
	BURST 6	100	100	100	100	100	100		100
	BURST 7	100	100	100	100	100	100	100	

Table 7 Diagnostic accuracy of section 4.2.4 ( $n_f = 20$ ,  $t_w = 1$  s,  $\nu = 0.08$ ,  $\sigma = 0.0228$ ).

Accuracy %		Test data							
		BURST NONE	BURST 1	BURST 2	BURST 3	BURST 4	BURST 5	BURST 6	BURST 7
Training data	BURST NONE		100	100	100	100	99	100	100
	BURST 1	100		100	100	0	99	21	0
	BURST 2	100	100		100	100	100	100	100
	BURST 3	100	100	100		100	100	100	100
	BURST 4	100	100	100	100		100	100	100
	BURST 5	3	100	100	100	100		31	100
	BURST 6	100	97	100	100	100	84		100
	BURST 7	100	100	100	100	100	100	100	

which is like in section 4.2.1 shown in Table 4. In the training data of BURST 1 or BURST 2, although the diagnostic accuracy was 97% or 94% for all test data, the slightly lower detection rate is not a big issue considering the failsafe concept.

#### 4.2.4 Non-uniform bandwidth with $n_f = 20$ and $t_w = 1$ s

Table 7 is the diagnostic accuracy when the number of band-pass filters and the division time width is the same as in section 4.2.1 and non-uniform bandwidth with the cutoff frequencies in reference to Mel filter bank as shown in Table 1 is used. Compared to the results in uniform bandwidth of  $f_{bw} = 10$  kHz shown in Table 4, the accuracy deteriorated when BURST 1 was used as training data. Considering the frequency range used in the burst waves as shown in Table 3, it can be predicted that the frequency range between 7 and 9 kHz affect the results in diagnosis. However, the number of band-pass filters that include the range between 7 and 9 kHz in the passband among the 20 filters is 15 for the uniform bandwidth, but only 3 for the non-

uniform bandwidth. In other words, the number of features considered effective for detecting abnormal sound is expected to differ 5 times even before PCA is performed and this is the reason why the accuracy deteriorated when the non-uniform bandwidth was used.

#### 4.2.5 Diagnosis of Pencil Lead Break sound

Using the model built by BURST NONE, the diagnosis was performed by the test data of Pencil Lead Break sound with the prospect of more practical cases. The diagnostic accuracy was 100% using the BURST NONE model with all four parameters written in sections 4.2.1–4.2.4.

In this chapter, the effectiveness of the OCSVM and the MEMS microphone unit developed in this study was verified by anomaly detection in a pipe with water flow noise and various burst waves from the piezoelectric elements and Pencil Lead Break sound. The training model obtained in this study cannot be used as it is for various noises in a factory. However, the process presented here is effective in determining a noise diagnosis model.

## 5. Conclusions

This study discussed acoustic diagnosis of sounds leaking from pipeworks with water flow, as an example of abnormal sound detection. For remote measurement of acoustic signals, a palm-sized microphone unit with two MEMS microphones, a battery, and a Bluetooth transmitter was fabricated. It can be remotely powered by external solar cells and can be placed anywhere within range of the irradiation to the solar cells and Bluetooth communication. Abnormal sound is detected using machine learning in which an anomaly detection model is built from the signals recorded by the microphone unit. The pseudo abnormal sounds to be detected were different types of burst waves output from a piezoelectric element attached to the pipe, and Pencil Lead Break sound. PCA and OCSVM were used for data dimensionality reduction and anomaly detection model building.

Nine types of acoustic data were measured, including seven types of burst waves with different parameters output from the piezoelectric element and a Pencil Lead Break sound. Four different model building parameters were used to demonstrate diagnostic accuracy, with the best accuracy being achieved when the number of band-pass filters was 20, the passband width was 10 kHz, and the division time width was 1 s. Burst wave changes could be detected in almost all patterns, giving a diagnostic accuracy of 100% for the Pencil Lead Break sound, which leads the conclusion that the acoustic diagnosis procedure used in this study can detect abnormal sound in pipeworks, and is expected to detect a wide variety of abnormal sounds that occur in actual workplaces. In that sense, this study will become important as a bridge between academia and plant maintenance personnel.

## Acknowledgments

This work was supported by Japan Society for the Promotion of Science KAKENHI [grant number 21H01573].

## REFERENCES

- 1) T. Hayashi, M. Murase and M.N. Salim: *J. Acoust. Soc. Am.* **126** (2009) 1101–1106.
- 2) T. Hayashi and M. Fukuyama: *J. Acoust. Soc. Am.* **140** (2016) 2427–2436.
- 3) T. Hayashi: *Appl. Phys. Lett.* **108** (2016) 081901.
- 4) T. Hayashi: *NDT&E Int.* **85** (2017) 53–62.
- 5) T. Hayashi and K. Ishihara: *Ultrasonics* **77** (2017) 47–53.
- 6) A. Maeda and T. Hayashi: *Mater. Trans.* **59** (2018) 320–323.
- 7) S. Omar, A. Ngadi and H.H. Jebur: *Int. J. Comput. Appl.* **79**(2) (2013) 33–41.
- 8) R.P. da Cruz, F.V. da Silva and A.M.F. Fileti: *Clean Technol. Environ. Policy* **22** (2020) 627–638.
- 9) B. Ahn, J. Kim and B. Choi: *Eng. Fract. Mech.* **210** (2019) 381–392.
- 10) V.G. Salunkhe and R.G. Desavale: *J. Nondestruct. Eval. Diagnostics Progn. Eng. Syst.* **4** (2021) 031004.
- 11) L. Xu, Q. Wang, I. Ivrišsimtziš and S. Li: *J. Nondestruct. Eval. Diagnostics Progn. Eng. Syst.* **5** (2022) 011004.
- 12) D. Goyal, S.S. Dhami and B.S. Pabla: *J. Nondestruct. Eval. Diagnostics Progn. Eng. Syst.* **4** (2021) 021006.
- 13) A. Sharma: *J. Nondestruct. Eval. Diagnostics Progn. Eng. Syst.* **5** (2022) 031004.
- 14) T. Regan, C. Beale and M. Inalpolat: *J. Vib. Acoust.* **139** (2017) 061010.
- 15) Y. Hong, M. Kim, H. Lee, J.J. Park and D. Lee: *IEEE Trans. Instrum. Meas.* **68** (2019) 4746–4755.
- 16) L. Li, R.J. Hansman, R. Palacios and R. Welsch: *Transp. Res., Part C Emerg. Technol.* **64** (2016) 45–57.
- 17) J. Venskys, P. Treigys, J. Bernatavičienė, G. Tamulevičius and V. Medvedev: *Sensors* **19** (2019) 3782.
- 18) Z. Wang, Y. Fu, C. Song, P. Zeng and L. Qiao: *IEEE Access* **7** (2019) 181580–181588.
- 19) Y. Aoki, F. Ehara and Y. Muto: 8th IIAE Int. Conf. Ind. Appl. Eng., (2020) pp. 17–20.
- 20) L. Basora, X. Olive and T. Dubot: *Aerospace* **6** (2019) 117.
- 21) Y. Koizumi, S. Saito, H. Uematsu, Y. Kawachi and N. Harada: *IEEE/ACM Trans. Audio, Speech: Lang. Process.* **27** (2019) 212–224.
- 22) Z. Tang, Z. Chen, Y. Bao and H. Li: *Struct. Control Heal. Monit.* **26** (2019) e2296.
- 23) D.M.J. Tax and R.P.W. Duin: *Mach. Learn.* **54** (2004) 45–66.
- 24) J. Kang and A.K. Patterson: *Principal Component Analysis*, (Springer-Verlag, New York, 2002).
- 25) D. O'Shaughnessy: *Speech Communication Human and Machine*, (Addison-Wesley, Boston, 1987) p. 150.