Mechanical Systems and Signal Processing 221 (2024) 111727

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

# Wavelet-based and data-adaptive methods for time delay estimation in acoustic leak detection

Ndubuisi Uchendu *, Jennifer M. Muggleton, Paul R. White

*Institute of Sound and Vibration Research (ISVR), University of Southampton, Southampton, SO17 1BJ, United Kingdom*

ARTICLE INFO

Communicated by S.D. Rosa

*Keywords:*
Leakage
Time delay estimation
Multi-resolution analysis
Wavelet transform
Data-adaptive decomposition
Cross-correlation

ABSTRACT

Leakages in water distribution networks are a common issue encountered in the water industry. The acoustic cross-correlation method is generally employed for detecting and locating leaks in water pipes. One important factor that determines the effectiveness of this method is the accuracy of the time delay estimate which is usually obtained from the cross-correlation function (CCF) of two signals measured on the pipe. However, in some practical situations, accurate time delay estimate cannot be obtained using existing correlation-based time delay estimation (TDE) methods without first filtering the signals prior to calculating the CCF. Incorrect choice of filter cut-off frequencies limits the effectiveness of these methods. To deal with this issue, this paper proposes a scheme based on multi-resolution decomposition for accurately estimating time delays between leak signals. The proposed scheme first decomposes the signals at different scales using a shift-invariant wavelet transform or data-adaptive decomposition and then determines the time delay from the decomposed signals. Simulation and experimental results demonstrate the higher effectiveness of the proposed method compared to the commonly used basic and generalised cross-correlation TDE methods.

## 1. Introduction

High prevalence of leakages in water distribution networks due to pipe ageing and deterioration, poor network design and construction, inadequate corrosion protection, and mechanical damage represents a major challenge facing water companies all over the world. Water leakages have serious social, economic, and ecological consequences, including lost revenue, increased cost to consumers, and inability to meet water demand [1]. Timely detection and repair of leakages are of primal importance. A comprehensive review of different leak detection and localisation methods, including acoustic methods, infra-red thermography, ground penetration radar, and transient-based methods, can be found in [2]. Among existing methods, acoustic cross-correlation using leak noise correlators is the most commonly used method for locating leaks in water distribution networks [3]. In this method, the leak location is determined from the time delay between acoustic/vibration signals measured on either side of the suspected leak. Leak noise correlators usually estimate this time delay from the cross-correlation function (CCF) of the signals. The effectiveness of estimating the time delay from the CCF is strongly affected by the frequency region over which cross-correlation processing is carried out [4]. In most correlators, the signals are usually passed through filters to attenuate the signals in the frequency region dominated by background noise prior to computing the CCF. Incorrect choice of filter bandwidth may result in inaccurate time delay estimates, and this represents a major issue affecting the performance of correlators [5]. To address this issue, this paper presents a time delay estimation (TDE) scheme that employs a multi-resolution decomposition framework implemented with wavelet transform (WT) or

* Corresponding author.
*E-mail addresses:* n.uchendu@soton.ac.uk (N. Uchendu), jmm@isvr.soton.ac.uk (J.M. Muggleton), P.R.White@soton.ac.uk (P.R. White).

https://doi.org/10.1016/j.ymssp.2024.111727
Received 18 March 2024; Received in revised form 6 June 2024; Accepted 9 July 2024
Available online 24 July 2024

data-adaptive decomposition. This scheme avoids explicit filtering of the signals by executing the cross-correlation processing on the signals at different scales or resolutions. Numerical simulation and test data from a real water pipe are used to evaluate the performance of the TDE scheme for leak localisation. The rest of this paper is organised as follows. Section 2 describes the acoustic cross-correlation method and introduces the methods commonly used for estimating delays between leak signals. The mean square error (MSE) of the time delay estimate is derived in terms of signal and pipe properties. The principles of the TDE scheme based on multi-resolution analysis are outlined in Section 3, while the wavelet-based and the data-adaptive implementations of the scheme are discussed in Sections 4 and 5, respectively. Section 6 presents and discusses the results obtained for simulated and experimental leak signals. Finally, the main conclusions are outlined in Section 7.

## 2. Overview of acoustic cross-correlation

A leak in a water pipe generates low-frequency acoustic/vibration signals that propagate along the pipe. To locate the leak, these signals can be acquired at two access points on either side of the location of the leak, as shown in Fig. 1. The leak location is calculated as

$$d_1 = \frac{d - c \cdot \tau_{\text{peak}}}{2} \tag{1}$$

where $d_k$, $k = 1, 2$, is the distance between the leak location and the $k$th measurement point, $\tau_{\text{peak}} = (d_2 - d_1)/c$ is the time delay between the two acquired signals $x_1(t)$ and $x_2(t)$, $c$ is the propagation speed of the leak signal in the pipe, and $d = d_1 + d_2$ is the distance between the measurement points. It can be seen that accurate TDE is necessary for accurate leak localisation. The most commonly used TDE method in leak noise correlators is basic cross-correlation (BCC), in which the time delay is determined as the time lag that maximises the CCF. The CCF $R_{x_1x_2}(\tau)$ of $x_1(t)$ and $x_2(t)$ is given by

$$R_{x_1x_2}(\tau) = \frac{1}{T}\int_0^T x_1(t-\tau)x_2(t)dt = \mathcal{F}^{-1}\left\{G_{x_1x_2}(\omega)\right\} = \mathcal{F}^{-1}\left\{X_1^*(\omega)X_2(\omega)\right\} \tag{2}$$

where $X_k(\omega) = \mathcal{F}\left\{x_k(t)\right\}$; $G_{x_1x_2}(\omega)$ is the cross-power spectrum (CPS) of $x_1(t)$ and $x_2(t)$; $T$ is the measurement duration; $^*$ denotes the complex conjugate; $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ denote the Fourier transform (FT) and inverse Fourier transform (IFT), respectively. An alternative to BCC is generalised cross-correlation (GCC), in which the CPS is multiplied by a weighting function prior to performing the IFT to obtain the CCF [6]. Examples of GCC methods are phase transform (GCC-PHAT), smoothed coherence transform (GCC-SCOT), maximum likelihood (GCC-ML), and Wiener estimator. The performance of these BCC and GCC methods for acoustic leak signals was investigated in [4,7].

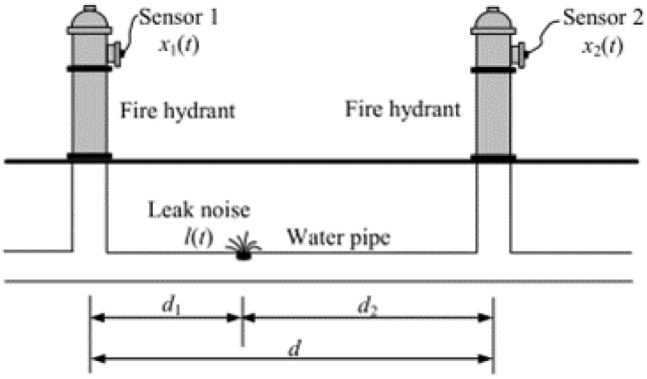


**Fig. 1.** Typical measurement set-up for the acoustic cross-correlation method.

In addition to the signals generated by the leak, acoustic/vibration signals measured on water pipes contain ambient noise and possibly interferences resulting from pipe system dynamics (for example, resonances). As noise and interferences introduce errors in the time delay estimate [8], accurate leak localisation via cross-correlation demands mitigating their effects. In leak noise correlators, this is achieved by filtering the signals prior to calculating their CCF. The aim of this filtering is to ensure that the cross-correlation processing is carried out only in the frequency region where the leak signal is significant relative to the background noise and other interferences [4,9]. As has been observed in a survey of commercial leak noise correlators [5], incorrect selection of filter bandwidth is one of the factors responsible for poor leak localisation performance. The effects of filter bandwidth, noise, and other factors on TDE accuracy can be assessed by evaluating the MSE of the time delay estimate in terms of pipe and signal properties. Referring to Fig. 1, the measured signals $x_1(t)$ and $x_2(t)$ can be represented as

$$x_1(t) = l(t) \otimes h_1(t) + n_1(t) = l_1(t) + n_1(t) \tag{3a}$$

$$x_2(t) = l(t) \otimes h_2(t) + n_2(t) = l_2(t) + n_2(t) \tag{3b}$$

where $l(t)$ is the signal generated at the leak location (i.e., the leak noise), $h_k$ is the impulse response function (IRF) that expresses the relationship between $l(t)$ and the signal at the $k$th measurement point, $l_k = l(t) \otimes h_k$ is the noise-free component of $x_k(t)$, $\otimes$

denotes linear convolution, and $n_k(t)$ is the background noise signal at the $k$th measurement point. The background noise signals are assumed to be mutually uncorrelated with each other and the leak noise. Based on an analytical model of wave propagation established in [10] for an infinitely long water pipe with no discontinuities, the frequency response function (FRF) $H(\omega, d_k) = \mathcal{F}\{h_k\}$ can be expressed as

$$H(\omega, d_k) = e^{-|\omega|\beta d_k} e^{-i|\omega|d_k/c} \tag{4}$$

where $\omega$ is radial frequency, $i = \sqrt{-1}$, and $\beta$ is the attenuation factor (a measure of the loss experienced by acoustic waves within the pipe wall). This model shows that a water pipe acts as a low-pass filter that attenuates relatively high frequencies. As stated above, leak noise correlators determine the time delay between $x_1(t)$ and $x_2(t)$ from the CCF $R_{\tilde{x}_1\tilde{x}_2}(\tau)$ of the filtered signals $\tilde{x}_1(t) = x_1(t) \otimes w_1(t)$ and $\tilde{x}_2(t) = x_2(t) \otimes w_2(t)$, where $w_1(t)$ and $w_2(t)$ are weighting functions corresponding to the filters applied to $x_1(t)$ and $x_2(t)$, respectively. If the applied filters are linear, then the MSE $\mathrm{E}\left\{\left(\hat{\tau}_{\mathrm{peak}} - \tau_{\mathrm{peak}}\right)^2\right\}$ (denoted as $\mathrm{MSE}\left(\hat{\tau}_{\mathrm{peak}}\right)$) of the unbiased time delay estimate $\hat{\tau}_{\mathrm{peak}}$ (assumed to lie in the neighbourhood of the true delay $\tau_{\mathrm{peak}}$) can be derived using the procedures in [11] as

$$\mathrm{MSE}\left(\hat{\tau}_{\mathrm{peak}}\right) = \frac{\pi}{T} \frac{\int_0^\infty \omega^2 W^2(\omega) \begin{bmatrix} G_{n_1n_1}(\omega)G_{l_2l_2}(\omega) + G_{n_2n_2}(\omega)G_{l_1l_1}(\omega) \\ +G_{n_1n_1}(\omega)G_{n_2n_2}(\omega) \end{bmatrix} d\omega}{\left[\int_0^\infty \omega^2 W(\omega) \left|G_{l_1l_2}(\omega)\right| d\omega\right]^2} \tag{5}$$

where $W(\omega) = \mathcal{F}\{w_1(t)\}^* \mathcal{F}\{w_2(t)\}$, and $\mathrm{E}\{\cdot\}$ denotes the expectation. If the same filter is applied to both signals, then $w_1(t) = w_2(t)$ and $W(\omega) = |\mathcal{F}\{w_1(t)\}|^2$. To simplify the MSE expression, $W(\omega)$ is taken as an ideal bandpass filter, i.e., $W(\omega) = 1$ in the frequency interval $\omega_1 \le \omega \le \omega_2$ and zero elsewhere. Also, the leak noise and the background noise signals are assumed to have flat spectra with $G_{ll}(\omega) = S_0$ and $G_{n_1n_1}(\omega) = G_{n_2n_2}(\omega) = N_0$. Rewriting Eq. (5) in terms of the FRFs $H(\omega, d_1)$ and $H(\omega, d_2)$ and evaluating the integrals using the relation [12]

$$\int_0^b \omega^n e^{-a\omega} d\omega = \frac{n!}{a^{n+1}} \left[1 - e^{-ab} \sum_{k=0}^{n} \frac{(ab)^k}{k!}\right], \quad a > 0 \tag{6}$$

gives the MSE as

$$\mathrm{MSE}\left(\hat{\tau}_{\mathrm{peak}}\right) = \frac{\pi}{16T} \frac{1}{\xi} \frac{(\beta d)^6}{\left(e^{-\omega_1\beta d} \sum_{m=0}^{2} \frac{(\omega_1\beta d)^m}{m!} - e^{-\omega_2\beta d} \sum_{m=0}^{2} \frac{(\omega_1\beta d)^m}{m!}\right)^2} \cdot \left[\begin{array}{l} \dfrac{\left(e^{-2\omega_1\beta d_1} \sum_{m=0}^{2} \frac{(2\omega_1\beta d_1)^m}{m!} - e^{-2\omega_2\beta d_1} \sum_{m=0}^{2} \frac{(2\omega_2\beta d_1)^m}{m!}\right)}{\left(\beta d_1\right)^3} \\ + \dfrac{\left(e^{-2\omega_1\beta d_2} \sum_{m=0}^{2} \frac{(2\omega_1\beta d_2)^m}{m!} - e^{-2\omega_2\beta d_2} \sum_{m=0}^{2} \frac{(2\omega_2\beta d_2)^m}{m!}\right)}{\left(\beta d_2\right)^3} \\ + \dfrac{4}{3} \dfrac{1}{\xi} \left(\omega_2^3 - \omega_1^3\right) \end{array}\right]. \tag{7}$$

where $\xi = \dfrac{S_0}{N_0}$. If $W(\omega)$ has a wide bandwidth such that $\omega_2 \gg \omega_1$ and $\omega_1$ is small (close to zero), then the MSE can be approximated as

$$\mathrm{MSE}\left(\hat{\tau}_{\mathrm{peak}}\right) \approx \frac{\pi}{16T} \frac{1}{\xi} \beta^3 d^6 \cdot \left[\frac{1}{d_1^3} e^{2\omega_1\beta d_2} + \frac{1}{d_2^3} e^{2\omega_1\beta d_1} + \frac{4}{3} \frac{1}{\xi} \omega_2^3 \beta^3 e^{2\omega_1\beta d}\right]. \tag{8}$$

By applying the Cauchy–Schwarz inequality [13] to the denominator of Eq. (5), the minimum value of the MSE, i.e, the Cramer–Rao lower bound (CRLB) $\tau_{\mathrm{CRLB}}$, is obtained as

$$\mathrm{MSE}\left(\hat{\tau}_{\mathrm{peak}}\right) \ge \tau_{\mathrm{CRLB}} = \frac{\pi}{T} \frac{1}{\xi} \left[\int_0^\infty \omega^2 \frac{e^{-2\omega\beta d}}{e^{-2\omega\beta d_1} + e^{-2\omega\beta d_2} + 1/\xi} d\omega\right]^{-1} \tag{9}$$

with equality if and only if $W(\omega) = K \dfrac{\xi e^{-2\omega\beta d}}{e^{-2\omega\beta d_1} + e^{-2\omega\beta d_2} + 1/\xi}$, where $K$ is an arbitrary real constant.

From the expressions in Eqs. (7)–(9), it can be observed that in addition to the cut-off frequencies of the applied filter $\omega_1$ and $\omega_2$ and the signal-to-noise ratio (SNR) of the leak noise $\xi$, the accuracy of TDE results is affected by the measurement time $T$, pipe attenuation $\beta$, inter-sensor distance $d$, and actual leak location. Large values of the attenuation factor and inter-sensor distance decrease the accuracy of the time delay estimate. The MSE is minimised when the leak is equidistant from both sensors (i.e., $d_1/d_2 = 1$) and becomes larger as the leak gets closer to any one end of the pipe. The CRLB is limited by pipe properties, measurement time, and the SNR of the leak noise. Among the factors that affect the MSE, only the measurement time and cut-off frequencies can be adjusted by the operator, so the possible measures for improving the quality of the time delay estimate in a given pipe and measurement set-up are increasing the measurement time and properly selecting $\omega_1$ and $\omega_2$ to mitigate or exclude the effects of noise and other interferences [8]. As suggested in [14], the filter bandwidth can be set as the frequency region where

the coherence of the measured signals exceeds some threshold. Proper selection of cut-off frequencies is thus difficult in situations where the spectral properties of the signals are not available *a priori* or cannot be reliably estimated. Furthermore, the coherence-based criterion is not effective when resonances are present in the signals [9]. A survey of experimental leak signals shows that background noise generally dominates at low frequencies (<5 Hz), while relatively high frequencies are severely attenuated [10]. Hence, in practice, the bandwidth of the applied filter are usually pre-set to exclude the low-frequency and the high-frequency regions [4]. Setting the cut-off frequencies in this manner without considering the pipe and signal properties may lead to inaccurate TDE results, and consequently, inaccurate leak localisation. Due to difficulties associated with properly selecting analysis frequency interval, development of approaches for estimating time delays without requiring knowledge of signal spectral properties is of practical importance. One such method proposed in the literature is the cross-cross-correlation method [15], in which the time delay is obtained as the least-squares solution of an overdetermined system of equations built from the CCF of the CCFs of each pair of measured signals. This method implicitly performs a filtering operation on the signals, approximating GCC but without requiring information about the signal spectra. However, it requires presence of at least three non-collocated sensors in the measurement set-up, thereby making it inapplicable to the acoustic leak localisation problem under consideration.

An alternative approach for estimating time delay without explicit spectral analysis or knowledge of signal and pipe properties is to carry out the cross-correlation processing in a subspace where leak signal can be separated from noise and interferences. This can be achieved via multi-resolution decomposition, which allows the separation of signals into components corresponding to different levels of details. By executing the cross-correlation processing on the multi-resolution components, it becomes possible to estimate the time delay while excluding the effects of noise and interferences. The rest of this paper concerns the principles and implementations of this TDE scheme.

## 3. Time delay estimation based on multi-resolution decomposition

Multi-resolution decomposition of a signal results in a hierarchical representation of the signal at different scales or resolutions, each of which corresponds to a frequency band [16]. The signal is separated into a series of low-frequency contents (approximations) and high-frequency contents (details) typically using a time-scale transform. By combining the approximations and details at different scales, the original signal can be recovered without loss. Multi-resolution decomposition is useful for identification and separation of signal from noise. A band-limited signal (for example, leak signals, bandwidth of which is limited by pipe attenuation properties [17]) is typically resolved at a few scales, whereas the background noise (assumed to have a wider bandwidth than the signal) often contributes to all scales. Therefore, only few resolution levels will be localised in significant part of the leak signal spectrum, while the rest of the levels will be localised in major frequency components of noise and other inferences. This leads to the idea of carrying out cross-correlation analysis on the multi-resolution components instead of the whole signals. A TDE scheme realising this idea is shown in Fig. 2. Once the leak signals are decomposed with a suitable time-scale transform, the time delay is estimated at each resolution (decomposition) level, and then the most accurate time delay estimate is selected using an appropriate criterion. In this paper, the terms ‘resolution level(s)’ and ‘decomposition level(s)’ will be used interchangeably.

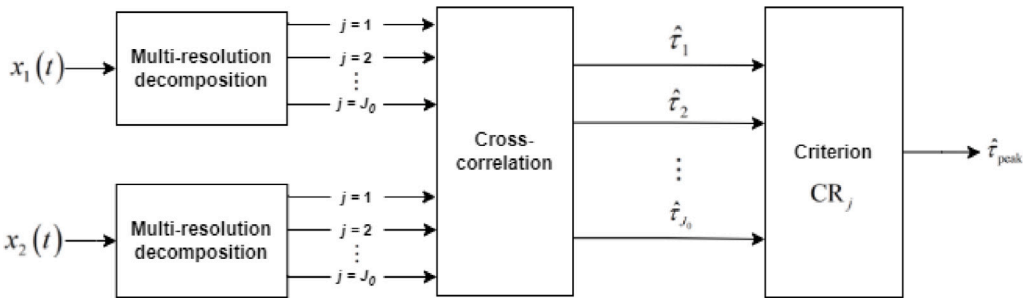


**Fig. 2.** Proposed TDE scheme. $\hat{\tau}_j$ denotes the time delay estimate obtained from the $j$th level multi-resolution components.

Suitable multi-resolution decomposition methods to be employed in the proposed TDE scheme must satisfy desirable requirements of shift invariance, low redundancy, and computational efficiency. Shift-invariance implies that any delay present in the original signal is preserved in the resolved components. In other words, a shift in the original signal manifests as an equivalent shift in the multi-resolution components at all scales. Low redundancy implies that there is very little overlap in the resolved components across scales. A non-redundant decomposition presents a compact representation of signals, which is desirable for computational efficiency in terms of time and memory requirements. Considering these requirements, two possible choices of decomposition methods for implementing the proposed TDE scheme are shift-invariant WT and data-adaptive decompositions. The former is a time-scale transform that dilates or expands a basic signal shape element (wavelet) and considers how well different regions of a signal match this prototype. The latter are methods that resolve signals into multi-resolution components without using a fixed basis function. Implementation of the TDE scheme using these decomposition methods are described in the next two sections. Also discussed are the criteria for selecting the best decomposition levels in each implementation.

## 4. Wavelet transform-based time delay estimator

Wavelet analysis relies on the characterisation of a signal by the distribution of the signal amplitude in a basis consisting of dilations and translations of a wavelet, i.e., a quickly vanishing oscillating function with good localisation in both frequency and time domains. The continuous wavelet transform (CWT) $W_\psi x$ of a signal $x(t) \in L^2(\mathbb{R})$ (the space of all finite-energy signals) with respect to a wavelet $\psi(t)$ is defined as the correlation $\langle x, \psi_{s,b} \rangle$ between $x(t)$ and $\psi_{s,b}(t)$ [18]

$$W_\psi x(s,b) = \langle x, \psi_{s,b} \rangle = \int_{-\infty}^{\infty} x(t)\psi^*_{s,b}(t)dt. \tag{10}$$

where $\psi_{s,b}(t) = |s|^{-1/2}\psi\left(\frac{t-b}{s}\right)$ corresponds to continuous scaling $s \neq 0$ and translation $b$ of $\psi(t)$. If discrete values of scaling and translation are used, then a discrete wavelet transform (DWT) is obtained [18]. Of particular interest in multi-resolution applications are orthogonal wavelets owing to their suitability for efficient decomposition and reconstruction of signals with minimal loss of information [19]. If $\psi(t)$ is an orthogonal wavelet, then for the discrete set of parameters $s_j = 2^{-j}$ and $b = n$ with $j, n \in \mathbb{Z}$ (the set of integers), the set $\{\psi_{j,n}(t)\} = \{2^{j/2}\psi(2^j(t-n))\}$ constitutes an orthonormal basis of $L^2(\mathbb{R})$ [20]. The resulting WT is known as a non-decimated dyadic DWT or maximal overlap discrete wavelet transform (MODWT). Setting $b = 2^{-j}k$ (instead of $b = n$) yields the conventional (decimated) DWT. In contrast with decimated DWT, MODWT is redundant (albeit significantly less so than CWT) and shift-invariant [21]. Hence, the MODWT represents the best choice of WT for implementing the TDE scheme.

Through the MODWT, a discrete signal $x(n)$ can be expanded as a linear combination of an orthogonal wavelet $\psi(n)$ at varying scales and translations [18]:

$$x(n) = \sum_{j=1}^{J_0} y_j(n) + R(n) = \sum_{j=1}^{J_0}\sum_{k=1}^{N} C_j(k)2^{-j/2}\psi\left(2^j(n-k)\right) + \sum_{k=1}^{N} D_{J_0}(k)2^{-j/2}\phi\left(2^j(n-k)\right) \tag{11}$$

where $J_0$ is the number of resolution levels, $N$ is the signal length, $C_j = \langle x, \psi_{j,n} \rangle$ are the wavelet coefficients at the $j$th resolution level, $y_j(n) = \sum_{k=1}^{N} C_j(k)\psi_{j,n}$ is the detail of $x(n)$ at the $j$th level, $R(n) = \sum_{k=1}^{N} D_{J_0}(k)2^{-j/2}\phi\left(2^j(n-k)\right)$ is the approximation of $x(n)$ at the $J_0$th resolution level, $D_{J_0}$ are the final-level scaling coefficients, and $\phi$ is the so-called scaling function (usually selected to satisfy a quadrature mirror filter relationship with $\psi$). The wavelet coefficients $C_j$ can be interpreted as the local residual errors between successive signal approximations at resolution levels $j$ and $j+1$. The frequency spectrum of $x(n)$ is successively divided into a high-frequency sub-band and low-frequency sub-band as the decomposition scale increases. If $F_s$ is the sampling frequency, $C_j$ and the detail $y_j$ contain the information of the signal corresponding to the frequencies in the interval $\left[2^{j-1}F_s, 2^jF_s\right]$. Thus, the MODWT resolution levels represent a set of bandpass filters. Since it is implemented using an orthogonal wavelet, MODWT preserves signal energy [20], i.e., $\|x\|^2 = \sum_{j=1}^{J_0} \|C_j\|^2 + \|D_{J_0}\|^2$, where $\|\cdot\|$ denotes the $\ell^2$-norm.

In the WT-based implementation of the TDE scheme, the CCF $R^{\psi}_{x_1x_2}(\tau, j) = R_{\tilde{x}_1\tilde{x}_2}(\tau)$ for each decomposition level $j$ is calculated as

$$R^{\psi}_{x_1x_2}(\tau, j) = \sum_{n=1}^{N} C_{j,1}(n-\tau)\, C_{j,2}(n) \tag{12}$$

where $C_{j,k} = \langle x_k, \psi_{j,n} \rangle$ are the $j$th level MODWT coefficients of the measured signal $x_k(n)$ with respect to an orthogonal wavelet $\psi$. This method of estimating the time delay from the location of the peak of $R^{\psi}_{x_1x_2}(\tau, j)$ in Eq. (12) will be termed wavelet transform cross-correlation (WTCC), and when necessary, denoted as WTCC-*wav*, where '*wav*' indicates the wavelet used. For the $j$th decomposition level, the weighting function $W(\omega)$ in Eq. (5) is given by

$$W(\omega) = \left|\mathcal{F}\left\{\psi_{j,n}(t)\right\}\right|^2 = 2^j\left|\Psi(2^{-j}\omega)\right|^2 \tag{13}$$

where $\Psi(\omega) = \mathcal{F}\{\psi(t)\}$. The bandwidth of $W(\omega)$, and consequently, the accuracy of WTCC time delay estimate will depend on the decomposition level $j$ and the wavelet $\psi(t)$. Selection of these two parameters is discussed in the next subsection.

### 4.1. Selection of wavelet function and decomposition level

In the proposed TDE scheme, the objective of multi-resolution decomposition is to separate components corresponding to frequency bands with significant leak signal content from those dominated by background noise and other interferences. As already stated, leak signals are generally characterised by low frequencies, so suitable wavelets for analysing leak signals should allow more low-frequency contents to pass. Among available choices, four well-known and readily available orthogonal wavelet families: Daubechies ('db$M$'), symlet ('sym$M$'), coiflet ('coif$M$'), and Fejer-Korovkin ('fk$M$') wavelets are considered in this paper. These wavelet families are popular and recommendable tools for analysing signals in many applications. Details about the first three wavelet families can be found in [22], while Fejer-Korovkin wavelets are described in [23]. Note that except Haar wavelet (denoted as 'haar' or 'db1'), these orthogonal wavelets cannot be expressed in closed form. In the wavelet notation, $M$ refers to the number of vanishing moments. A wavelet with $M$ vanishing moments is orthogonal to polynomials of degree $M-1$. As the number of vanishing moments grows, the greater the wavelet oscillates, and the larger its support length in the time domain and the better its frequency localisation capability [19]. Fig. 3 shows the frequency responses of different wavelets from each family at various decomposition levels. It is clear from the plot that the frequency responses are determined primarily by the complexity of the wavelets, i.e., number of vanishing moments. Wavelets with comparable level of complexity exhibit similar frequency characteristics

at the same decomposition levels. Simpler wavelets, i.e., those with smaller number of vanishing moments, are less computationally expensive and allow more low-frequency contents to pass at each decomposition level, making them more suitable for analysing leak signals.

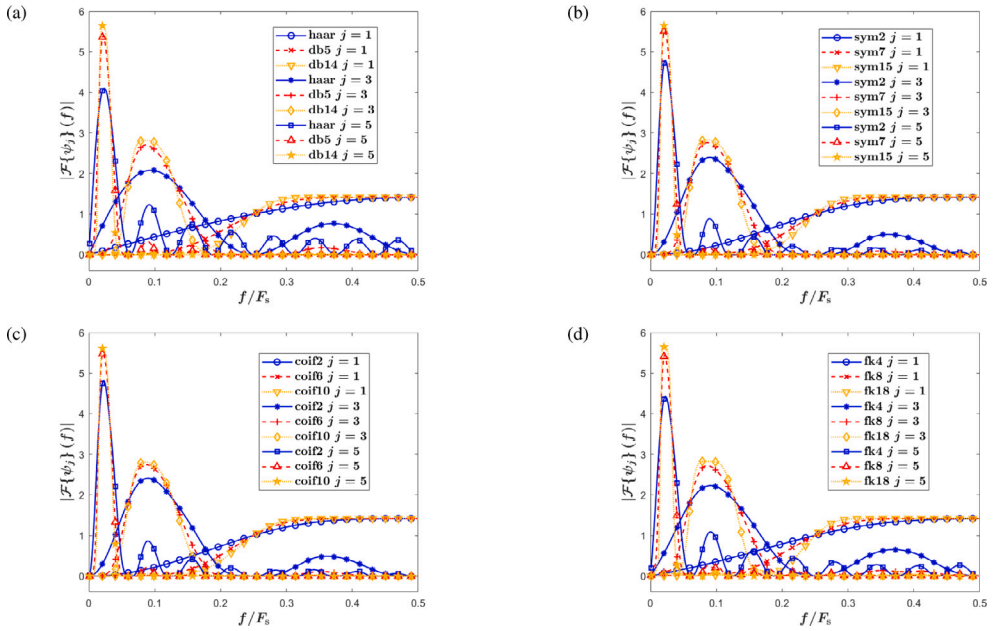


**Fig. 3.** Frequency responses of wavelet families at different decomposition levels: (a) Daubechies wavelets; (b) symlets; (c) coiflets; (d) Fejer-Korovkin wavelets. $f$ is frequency in Hertz.

Wavelet coefficients of a signal are closely related to the signal characteristics. As an energy-preserving transform, MODWT partitions signal energy by scale (or decomposition level) [20]. As already stated, the leak signal will contribute significantly to the wavelet coefficients in a few levels, while the background noise will tend to contribute to all levels. The signal energy will therefore be high at levels encompassing frequency bands where the leak signal spectrum is significant and low in the others. The best decomposition levels can be selected by employing the relative wavelet energy (RWE) criterion defined as [24]

$$\mathrm{RWE}_j = \frac{E_{j,1}}{E_{\mathrm{tot},1}} + \frac{E_{j,2}}{E_{\mathrm{tot},2}} \tag{14}$$

where $E_{j,k} = \sum_{n=1}^{N} \left|C_{j,k}(n)\right|^2$ is the energy in the $j$th decomposition level of $x_k$, and $E_{\mathrm{tot,k}} = \sum_{n=1}^{N} \left|x_k(n)\right|^2$ is the total energy in $x_k$. The best decomposition levels are those with the highest RWE. An alternative criterion can be based on the properties of the CCF at each decomposition level. The CCF resulting from a suitable decomposition level should allow for unambiguous determination of the time delay estimate. Suitability of a decomposition level can therefore be assessed using a measure of the SNR of the CCF termed peak-to-side lobe ratio (PSR), which is defined as [25]

$$\mathrm{PSR}_j = \frac{\left(R^{\psi}_{x_1x_2}(\tau_{\mathrm{peak}}, j)\right)^2}{\mathrm{var}\left\{R^{\psi}_{x_1x_2}(\tau_{\mathrm{far}}, j)\right\}} \tag{15}$$

where $\mathrm{var}\left\{R^{\psi}_{x_1x_2}(\tau_{\mathrm{far}}, j)\right\}$ denotes the variance of correlation values at *far points*, i.e., at lags $\tau_{\mathrm{far}}$ *far* from the CCF peak. If the CCF is computed in the lag interval $-n_c \le \tau \le n_c$, then the *far points* are selected to be at least $n_c/2$ samples away from $\tau_{\mathrm{peak}}$. A large PSR value implies that the CCF has a prominent main peak and low values away from this peak. Such a CCF is more likely to give an accurate time delay estimate compared to a CCF in which the main peak is of comparable height with other peaks. The PSR defined in Eq. (15) will be normalised by dividing its value by the sum of all the PSR values ($\sum_{j=1}^{J_0} \mathrm{PSR}_j$). Combining RWE and PSR yields the single criterion

$$\mathrm{CR}_j = \mathrm{RWE}_j \cdot \mathrm{PSR}_j. \tag{16}$$

for selecting the best decomposition levels. The bigger the value of $\mathrm{CR}_j$, the more suitable the $j$th decomposition level will be for TDE. By combining RWE and PSR, an appropriate decomposition level should extract high amount of energy from the leak signals being analysed, while ensuring that the output CCF gives an accurate time delay estimate. An approach that may be useful for improving TDE accuracy is to estimate the time delay from the sum of wavelet coefficients at the decomposition levels with high $\mathrm{CR}_j$ values. As the very low frequencies in leak signals are usually dominated by background noise, it is suggested to consider only decomposition levels whose lower cut-off frequencies are above 5 Hz. In other words, the maximum decomposition level considered in WTCC is set to $J_0 = \left\lfloor \log_2\left(\frac{F_s}{5}\right) - 1 \right\rfloor$, where $\lfloor\cdot\rfloor$ denotes the floor function. For example, $J_0 = 6$ for signals sampled at 1 kHz.

### 4.2. Computational complexity of WTCC

The computational costs of the steps involved in WTCC are summarised in Table 1. Counted are all basic arithmetic operations involved: addition, multiplication, and division. Each basic operation is assumed to incur the same unit cost (number of flops). The cost associated with computing orthogonal wavelet filters is assumed to be negligible. However, it should be noted that this cost could be quite significant for wavelets with high number of vanishing moments [26]. It is also assumed that all possible lags are considered in the CCF, i.e., $n_c = N-1$. Efficient CCF computation involves two fast Fourier transform (FFT) and one inverse FFT operations, each of which has cost $k_F N \log_2 N$. Computational cost of a $J_0$-level MODWT is $J_0 k_M N$ [20]. The real constants $k_F$ and $k_M$ depend on the specific FFT and MODWT algorithms used. The total cost of WTCC is approximately $\left[\left(3k_F \log_2 N + 2k_M + 8\right) N + 2\right] J_0 + 4N$. If WTCC is implemented for all possible levels, i.e., $J_0 = \lfloor \log_2 N \rfloor$, then the time complexity of WTCC in big-O notation is $\mathcal{O}\left(N(\log_2 N)^2\right)$. This is higher than the time complexity of $\mathcal{O}\left(N \log_2 N\right)$ for BCC and GCC.

**Table 1**
Computational cost of WTCC operations.

| **Step** | MODWT | RWE | PSR | CR | CCF |
|---|---|---|---|---|---|
| **Cost** | $2J_0 k_M N$ | $J_0(4N+3)+4N$ | $J_0(4N-2)$ | $J_0$ | $3J_0 k_F N \log_2 N$ |

## 5. Data-adaptive time delay estimator

Data-adaptive decompositions separate a signal $x(t)$ into a set of multi-resolution components $m_j(t)$ known as intrinsic mode functions (IMFs) and a residual $r(t)$, i.e.,

$$x(t) = \sum_{j=1}^{J_0} m_j(t) + r(t) \tag{17}$$

where $J_0$ is the number of extracted IMFs. Examples of data-adaptive decompositions are empirical mode decomposition (EMD) [27], variational mode decomposition (VMD) [28], and empirical wavelet transform (EWT) [29]. These execute signal decomposition in very different ways. Empirical mode decomposition works recursively on the time-domain signal to extract progressively lower frequency IMFs via a sifting operation, which involves the calculation of the moving average of the envelopes connecting the extrema of the signal [27]. Variational mode decomposition formulates the problem of extracting IMFs as an optimisation problem, which is solved using the alternating direction method of multipliers [30]. Empirical wavelet decomposition extracts IMFs using an adaptive wavelet filter bank constructed based on frequency content of the signal [29]. In VMD and EWT, all IMFs are extracted concurrently, and the number of IMFs has to be chosen in advance. Each VMD and EWT IMF has a compact frequency support around a central frequency, which is identified from local maxima in the signal spectrum. It should be noted that EMD is purely algorithmic and not readily amenable to mathematical analysis in contrast with VMD and EWT. All three data-adaptive decompositions can be viewed as a set of bandpass filters [31]. The centre frequencies of the filters in EMD decrease on a dyadic scale just like in MODWT [32,33]. Only EWT possesses the energy-preserving property of MODWT [34]. The data-adaptive decompositions differ in their computational efficiency. Results of empirical studies show that VMD is significantly more computationally expensive than EMD and EWT [35]. In fact, VMD takes 6 times longer than EMD to decompose the same signal in [36].

In the data-adaptive implementation of the TDE scheme, the CCF $R^{DD}_{x_1x_2}(\tau, j) = R_{\tilde{x}_1\tilde{x}_2}(\tau)$ for each decomposition level $j$ is calculated as the discrete CCF of the IMFs $m_{j,1}(t)$ and $m_{\tilde{j},2}(t)$:

$$R^{DD}_{x_1x_2}(\tau, j) = \sum_{n=1}^{N} m_{j,1}(n-\tau) m_{\tilde{j},2}(n) \tag{18}$$

where $m_{\tilde{j},2}$ denotes the IMF of $x_2(t)$ that encompasses the same frequency interval as the $j$th IMF $m_{j,1}$ of $x_1(t)$. This method of estimating the time delay from $R^{DD}_{x_1x_2}(\tau, j)$ will be denoted in this work as '*DD*-CC', where '*DD*' denotes the data-adaptive decomposition employed. Due to the dyadic nature of EMD, the IMFs used in EMD-CC correspond to the same decomposition level, i.e, $m_{\tilde{j},2} = m_{j,2}$. In the case of VMD-CC and EWT-CC, $m_{\tilde{j},2}$ has to be selected as the IMF whose centre frequency is closest to that of $m_{j,1}$. This presents no difficulty in practice as VMD and EWT centre frequencies are readily available to the user.

As in WTCC, choice of data-adaptive decomposition method can be based on computational efficiency, while selection of best decomposition level can be accomplished using PSR (or an energy-based criterion if the decomposition method preserves energy). Based on computational efficiency, EMD and EWT are better choices than VMD. For EMD-CC and VMD-CC, the criterion $\mathrm{CR}_j$ is equal to $\mathrm{PSR}_j$, whereas for EWT-CC, it is as defined in Eq. (16) with $\mathrm{RWE}_j$ given in terms of the analysis coefficients of the EWT wavelets corresponding to IMFs $m_{j,1}$ and $m_{\tilde{j},2}$. As in WTCC, the search for the best IMF in EMD-CC may be restricted to the first few decomposition levels. The TDE process can be improved by calculating the CCF from the sum of IMFs with high $\mathrm{CR}_j$ values.

The DD-CC differs from WTCC only in the decomposition step and possible absence of RWE calculation. The time cost for extracting $J_0$ IMFs using EMD is approximately $41 N_s J_0 N$, where $N_s$ denotes the number of sifting operations performed in each EMD iteration [37]. Setting $N_s$ to be constant ($\approx 10$) generally produces better results than using other stoppage criteria in the EMD algorithm [38,39]. The total computational cost of EMD-CC is therefore approximately $\left[\left(3k_F \log_2 N + 414\right) N - 1\right] J_0$. Assuming the maximum number of possible IMFs, i.e., $J_0 = \lfloor \log_2 N \rfloor$, EMD-CC has the same computational complexity $\mathcal{O}\left(N(\log_2 N)^2\right)$ as WTCC. No explicit expression for the time complexity of VMD and EWT was found in the literature. However, since EWT-CC uses a wavelet filter bank and involves RWE computation, EWT-CC will have a similar computational cost as WTCC. The relatively higher computational cost of VMD implies that VMD-CC will incur higher cost than EMD-CC and EWT-CC.

## 6. Results and discussion

In this section, the performances of the proposed WTCC and DD-CC are investigated using simulated and experimental leak signals.

### 6.1. Simulation results

Simulated leak signals were generated by filtering a white noise signal (representing the leak noise) with the FRFs $H(\omega, d_1)$ and $H(\omega, d_2)$ of a plastic pipe whose wave speed and attenuation factor are $c = 484$ m/s and $\beta = 3.6 \cdot 10^{-5}$ s/m, respectively. The distances between the leak and the sensor locations were set to $d_1 = 32.8$ m and $d_2 = 76.7$ m. These values correspond to the properties of the pipe and measurement set-up used for experimental investigation in the next subsection. The sampling rate and signal duration were set to $F_s = 1$ kHz and $T = 30$ s. Two tests were performed. In the first test, the effects of decomposition method and level on the accuracy of the time delay estimate were investigated. Gaussian noise was added to the simulated leak signals so that the SNR was 0 dB. Fig. 4 shows the theoretical root mean square error (RMSE), i.e., square root of the MSE in Eqs. (5) and (7) for the first 10 decomposition levels of wavelets from different families. Also shown are the square root of the CRLB (Eq. (9)) and the experimental RMSE calculated over $N_r = 500$ Monte Carlo simulation runs using the formula

$$\mathrm{RMSE}(\hat{\tau}_{\mathrm{peak}}) = \sqrt{\frac{1}{N_r}\sum_{k=1}^{N_r}\left(\hat{\tau}_{\mathrm{peak,k}} - \tau_{\mathrm{peak}}\right)^2} \tag{19}$$

where $\hat{\tau}_{\mathrm{peak,k}}$ is the time delay estimate from the $k$th run. For each run, different realisations of the leak noise and the added background noise were used. To achieve sub-sample accuracy, parabolic interpolation was used to calculate the time delay as [40]

$$\hat{\tau}_{\mathrm{peak}} = \frac{1}{F_s}\left(D_{\mathrm{peak}} + \frac{\theta_1 - \theta_3}{2\left(\theta_1 - 2\theta_2 + \theta_3\right)}\right) \tag{20}$$

where $D_{\mathrm{peak}}$ denotes the lag corresponding to the CCF peak, while $\theta_1$, $\theta_2$, and $\theta_3$ are the CCF values at lags $D_{\mathrm{peak}} - 1$, $D_{\mathrm{peak}}$, and $D_{\mathrm{peak}} + 1$, respectively. As anticipated in Section 4.1, the simpler the wavelet, the lower the RMSE at any decomposition level. This

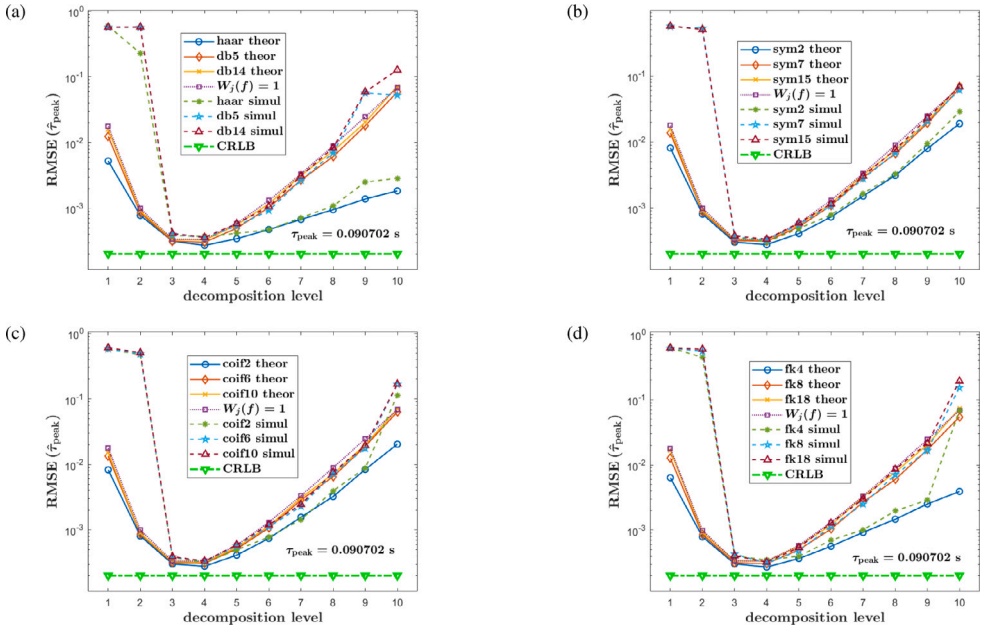


**Fig. 4.** Theoretical RMSE (indicated with 'theor') and simulation RMSE (indicated with 'simul') of the time delay estimate for different wavelet families: (a) Daubechies wavelets; (b) symlets; (c) coiflets; (d) Fejer-Korovkin wavelets.

confirms the suitability of simpler wavelets for implementing WTCC. Since more complex wavelets are characterised by 'flatter' frequency responses (see Fig. 3), which are closer to the ideal bandpass filter assumed for $W(\omega)$, the MSE expression in Eq. (7) becomes more accurate as the number of vanishing moments increases. This expression overestimates the RMSE for the simpler wavelets. The experimental RMSE agrees well with the theoretical RMSE in the decomposition levels with low RMSE. There is a large discrepancy between the theoretical and experimental RMSE at the lowest decomposition levels. One possible reason for this is that these levels encompass high-frequency regions where leak noise is severely attenuated by the pipe. An interesting observation is that all wavelets achieve smallest RMSE and gets closest to the CRLB at the exact same decomposition level. This implies that the choice of decomposition level is more critical for TDE accuracy than the type of wavelet used. The highest values of $\mathrm{RWE}_j$,

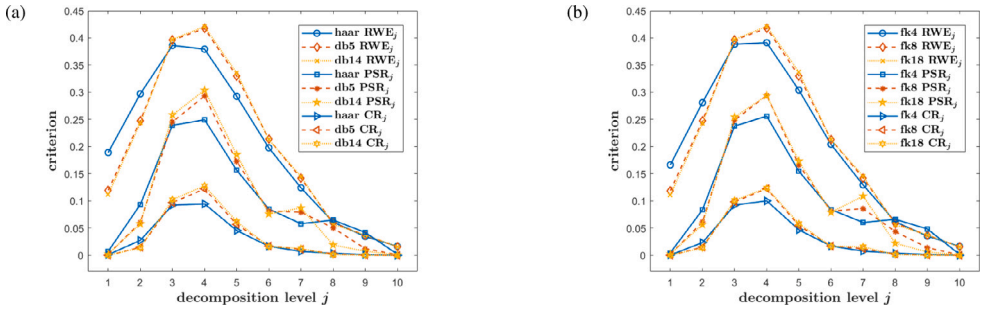


**Fig. 5.** Average values of RWE$_j$, PSR$_j$, and CR$_j$: (a) Daubechies wavelets; (b) Fejer-Korovkin wavelets.

PSR$_j$, and CR$_j$ for each wavelet occur at the decomposition levels with the lowest RMSE, as shown in Fig. 5 for Daubechies and Fejer-Korovkin wavelets. This confirms suitability of these criteria for selecting best decomposition levels.

The same observations can be made about DD-CC. Due to nonlinearity of data-adaptive decompositions, the expression for the theoretical MSE in Eq. (5) is not valid for DD-CC and thus, only the experimental RMSE is presented here. Fig. 6(a) shows the RMSE of EMD-CC, VMD-CC, and EWT-CC time delay estimates for the first 8 IMFs of the simulated leak signals. As in WTCC, the lowest RMSE is achieved by the IMFs with the highest RWE$_j$ and CR$_j$ values (Fig. 6(b)).

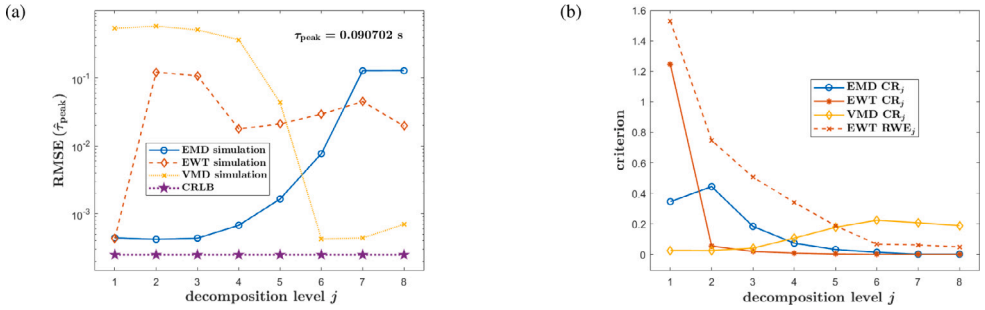


**Fig. 6.** DD-CC simulation results: (a) RMSE of the time delay estimate; (b) PSR$_j$ and CR$_j$.

The second test compares the performances of WTCC and DD-CC with those of BCC, GCC-PHAT, and GCC-ML using Monte Carlo simulations. In this case, realisations of Gaussian noise were added to the simulated leak signals to achieve SNR between −10 and 6 dB. For each SNR, 500 Monte Carlo simulations were run. The decomposition levels in WTCC and DD-CC were set to the best performing levels as determined from the results of the first test above. Fig. 7 shows the RMSE of the time delay estimate at each SNR. Regardless of the wavelet and data-adaptive decomposition used, WTCC and DD-CC achieve lower RMSE in general than BCC and GCC at low SNR. All the methods achieve similar performance when the SNR is high. These results demonstrate the superiority of the proposed multi-resolution methods for acoustic leak signals.

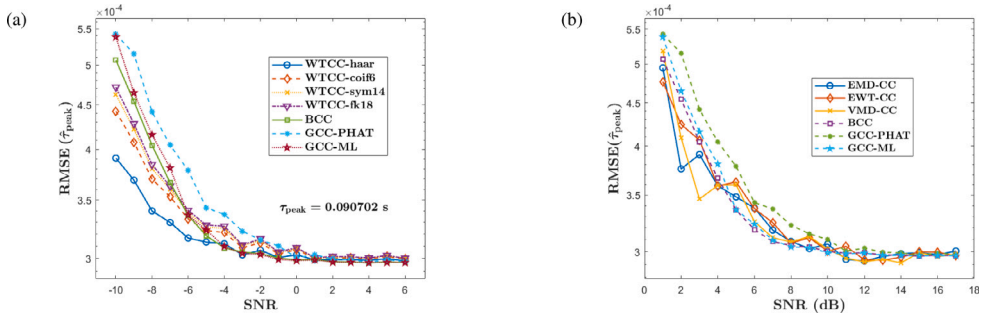


**Fig. 7.** RMSE of time delay estimates over 500 simulation runs: (a) WTCC vs GCC; (b) DD-CC vs GCC.

### 6.2. Experimental results

Signals measured on a polyvinyl chloride (PVC) pipe at a leak detection facility in Canada were used to experimentally evaluate the performance of WTCC and DD-CC. Fig. 8 shows the schematic of the test site. Detailed description of the test site and measurement procedures can be found in [41]. Leak signals generated by a leaky joint were measured using hydrophones and accelerometers installed on risers connected to two hydrants, one upstream and the other downstream of the joint. Inclusive of the

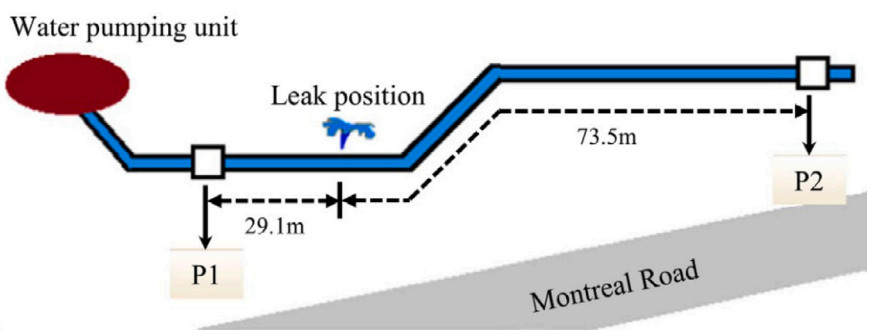


**Fig. 8.** Schematic of the experimental test site.

lengths of the downstream and upstream risers (3.7 m and 3.2 m, respectively), the distances between the leak and the measurement points are $d_1 = 32.8$ m and $d_2 = 76.7$ m, respectively. The hydrophone and accelerometer leak signals were measured for 60 and 66 s, respectively, at a sampling rate of $F_s = 500$ Hz. In previous research, the acoustic wave speed for hydrophone and accelerometer signals in the pipe was experimentally estimated to be 484 m/s and 479 m/s, respectively [42]. This gives the true delay as 90.7 ms and 91.6 ms, respectively, for the hydrophone and accelerometer signals. The time-domain resolution of the time delay estimate is $1/F_s = 2$ ms. An error of 2 ms in the time delay estimate corresponds to an absolute leak localisation error of 0.48 m.

In order to reveal the leak characteristics, spectral analysis was conducted on the signals using a 1024-point FFT, a Hann window (with 50% overlap between segments) and spectral averaging. Fig. 9 shows the CPS (magnitude and unwrapped phase) and coherence. As can be observed, the leak signals measured by the hydrophones dominate mostly in the frequency range between 10 Hz and 100 Hz, over which the CPS magnitude decreases with increasing frequency and the unwrapped phase is approximately linear. This region is also characterised by high coherence between the signals. Two resonant peaks in the CPS magnitude are observed at 56 Hz and 83 Hz, where phase shifts occur as marked by red circles. Oscillatory behaviour evident in the CPS magnitude at lower frequencies of 10–30 Hz is likely indicative of acoustic reflections in the pipe system [43]. Below 10 Hz, the CPS is erratic due to effects of noise. Signals at higher frequencies above 100 Hz become very small due to the low-pass filtering effect of the pipe. The presence of resonance and noise interference represents a situation often encountered in practice. The CPS of the accelerometer signals exhibits band-pass filtering behaviour, and the useful frequency band of the leak signals is 40–110 Hz, characterised by linear unwrapped CPS phase. In this case, there are no resonance effects.

As stated in Section 2, BCC and GCC may give inaccurate time delay estimates if the interfering noise in the leak signals is included in cross-correlation processing. Indeed, BCC, GCC-PHAT, and GCC-ML provide time delay estimates of 448, 87.6 and 87.8 ms, respectively, for the raw hydrophone signals. These estimates differ from the true delay by more than 2 ms. As in the simulation tests, sub-sample accuracy was achieved using parabolic interpolation (Eq. (20)). In the case of raw accelerometer signals, BCC, GCC-PHAT, and GCC-ML estimates are 90.9, 91.5, and 91.2 ms, respectively, which are all within 2 ms of the true delay. The proposed WTCC and DD-CC were implemented for the first 5 decomposition levels of the raw signals. Based on the values of the criterion $CR_j$ shown in Fig. 10, WTCC, EMD-CC, VMD-CC, and EWT-CC will provide most accurate time delay estimates at decomposition levels 3, 1, 1, and 2, respectively, for the hydrophone signals. In the case of the accelerometer signals, the best WTCC, EMD-CC, VMD-CC, and EWT-CC levels are 2, 1, 4, and 4, respectively. The WTCC and DD-CC time delay estimates are shown in Table 2. All estimates provided by levels where $CR_j$ is highest are within 2 ms of the true time delay. Simpler wavelets provide accurate estimates at more decomposition levels than more complex wavelets. Haar wavelet yield accurate time delay estimates even at levels 1 and 2 of the hydrophone signals, and accurate estimates at all levels of the accelerometer signals.

In order to improve the accuracy of the BCC and the GCC time delay estimates, there is need to first filter the hydrophone signals before cross-correlating the signals. Two 4th order Butterworth bandpass filters with passbands 10–50 Hz and 10–100 Hz were considered to illustrate the importance of proper filter bandwidth selection. The passband of the first filter was selected to exclude the two resonance peaks, while the passband of the second filter was selected using the coherence-based criterion proposed in [14] to encompass frequency interval where the coherence exceeds a threshold of $10^{-3}$. Comparing the CCFs obtained for the raw and filtered hydrophone signals in Fig. 11 shows that application of the filters improves the effectiveness of unambiguously estimating the time delay from the CCF. Table 3 summarises the time delay estimates obtained for raw and filtered signals using all the time delay estimators. Also shown is the absolute leak localisation error $\Delta d_1$ associated with each estimate. The WTCC and DD-CC estimates correspond to those obtained at the best levels as determined from Fig. 10. Filtering the hydrophone signals in the frequency range 10–50 Hz improves the BCC, GCC-PHAT, and GCC-ML estimates to 92.8, 91.7, and 90.1 ms, respectively. If the signals are filtered without removing the resonance peaks, BCC and GCC-ML give time delay estimates of 93.3 and 90.3 ms, respectively, while the GCC-PHAT provides an inaccurate estimate of 112 ms. Since GCC-PHAT gives equal weight to the frequency components within the whole bandwidth [4], it does not compensate for resonance effects in the cross-correlation process. All WTCC and DD-CC estimates for raw and filtered signals are accurate with small discrepancy less than the time-domain resolution of 2 ms. Although not necessary, the accelerometer signals were also passed through a 4th order Butterworth filter with passband 40–110 Hz. In this case, without and with filtering, all estimators give accurate estimates within 2 ms of the true delay.

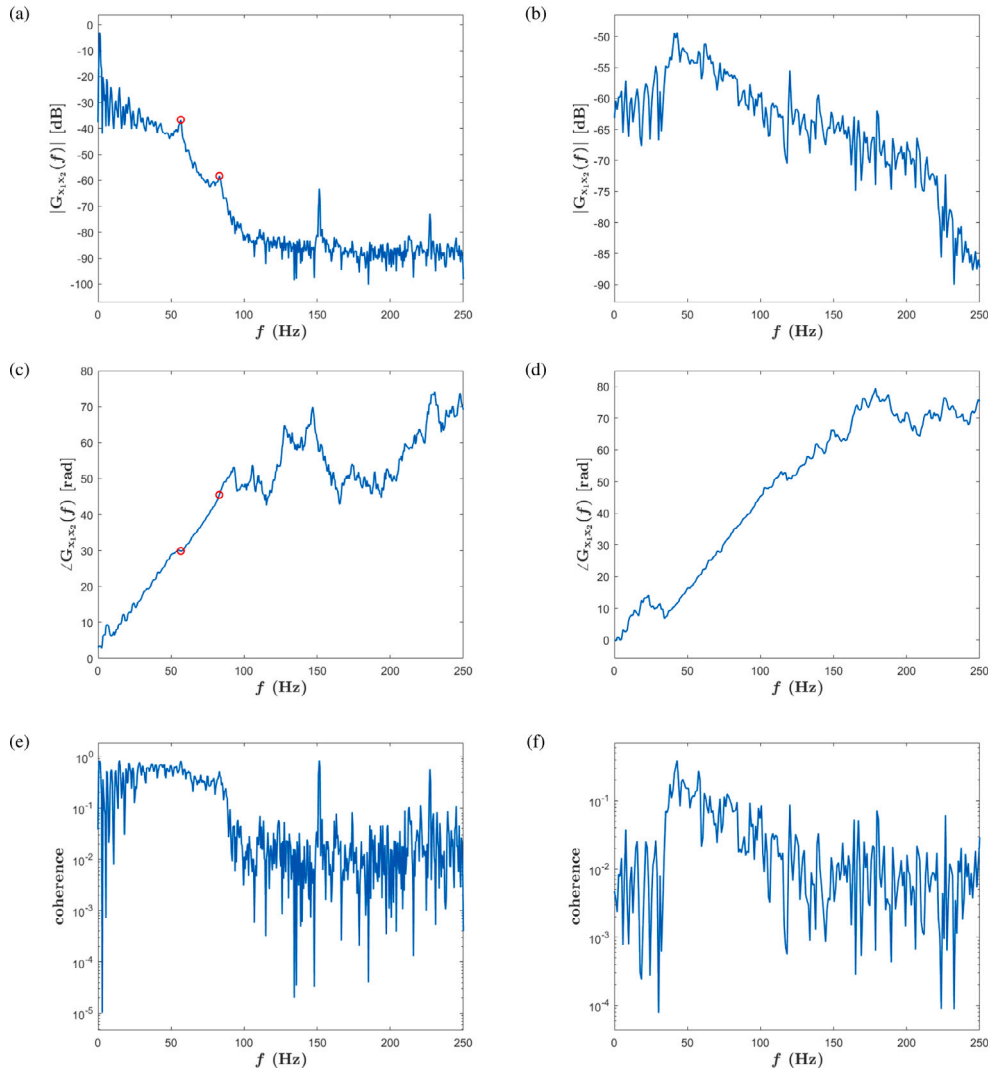


**Fig. 9.** Spectral properties of experimental signals. Hydrophone signals: (a) CPS magnitude; (c) unwrapped CPS phase; (e) coherence. Accelerometer signals: (b) CPS magnitude; (d) unwrapped CPS phase; (f) coherence.

### 6.3. Discussion

The results above demonstrate that BCC and GCC may give incorrect time delay estimates in some practical situations if leak signals are not first passed through a filter with properly selected cut-off frequencies prior to performing cross-correlation analysis. In contrast, the proposed WTCC and DD-CC are capable of accurately extracting the time delay without any pre-processing. The ability of these methods to achieve good performance when BCC and GCC fail is due to the inherent filtering associated with the multi-resolution decomposition methods employed in the TDE scheme. Accurate TDE results are achieved irrespective of the type of wavelet or data-adaptive decomposition employed as long as the decomposition level is appropriately selected using the proposed $CR_j$ criterion. Because simpler wavelets allow more useful low-frequency contents to pass, they achieve better results over more decomposition levels. One can, therefore, implement WTCC using the simplest orthogonal wavelet, Haar wavelet. Also, DD-CC can be implemented with EMD for computational efficiency.

The proposed TDE scheme has some practical benefits. The time delay can be determined directly from the measured signals without requiring spectral information of the noise and leak signals. Filtering operations are no longer required unlike in the methods currently employed in leak noise correlators. This makes the proposed scheme attractive in cases when the statistical characteristics of the signals cannot be known *a priori*. Since WTCC and DD-CC can be employed without need for any user input,

**Table 2**
WTCC and DD-CC time delay estimates for raw leak signals. Estimates that are within the time domain resolution (2 ms) of the true delay are highlighted in yellow.

| | Decomposition function/method | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $j$ | haar | sym2 | coif6 | sym11 | db14 | fk18 | EMD | EWT | VMD |
| | $\hat{\tau}_{\text{peak}}$ [ms] | | | | | | | | |
| **Hydrophone signals** | | | | | | | | | |
| **1** | 89.3 | 67.3 | 67.4 | 54.0 | −382 | 87.0 | 89.5 | 90.6 | 49.9 |
| **2** | 89.7 | 88.0 | 67.6 | 49.8 | 49.8 | 67.2 | 98.5 | 239 | 90.5 |
| **3** | 90.9 | 90.2 | 89.7 | 89.6 | 89.6 | 89.5 | 576 | 0.2 | 239 |
| **4** | −233 | 94.4 | 94.3 | 94.2 | 94.2 | 94.1 | 468 | −1900 | −584 |
| **5** | −236 | −233 | 104 | 105 | 105 | 105 | 1258 | 459 | 460 |
| **Accelerometer signals** | | | | | | | | | |
| **1** | 91.2 | 91.4 | 252 | 252 | 252 | 252 | 91.2 | 1528 | −292 |
| **2** | 91.0 | 91.1 | 91.2 | 91.2 | 91.2 | 91.2 | 90.3 | −864 | −850 |
| **3** | 90.4 | 90.3 | 90.2 | 90.2 | 90.2 | 90.2 | −308 | 90.0 | 90.2 |
| **4** | 90.9 | 91.1 | 92.3 | 93.1 | 57.1 | 56.8 | −963 | 91.6 | 91.4 |
| **5** | 90.7 | −1658 | −856 | −856 | −856 | −856 | 1408 | 90.3 | −232 |

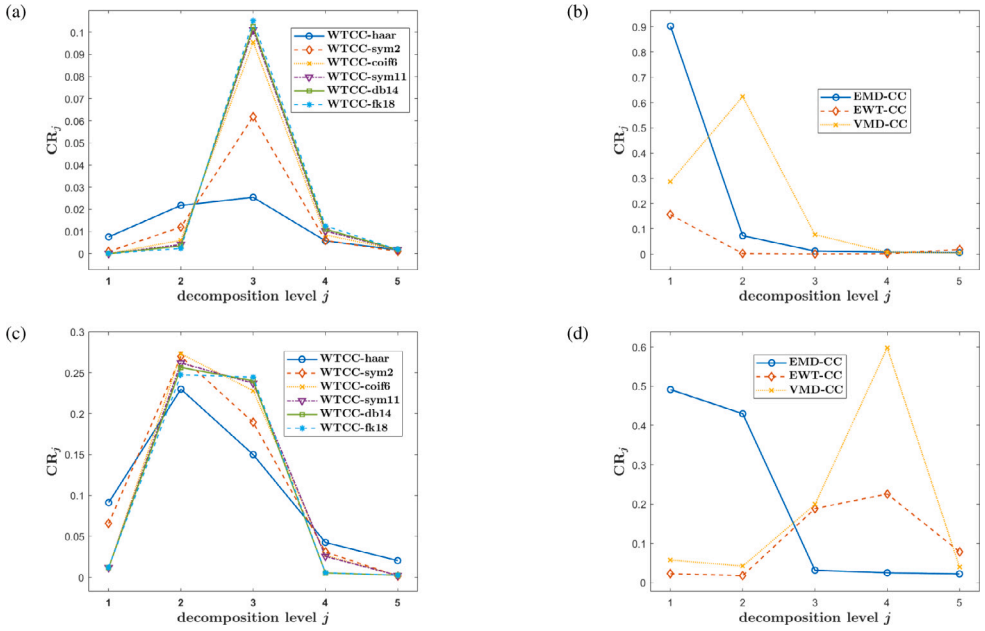


**Fig. 10.** Criterion $\text{CR}_j$. Hydrophone signals: (a) WTCC; (c) DD-CC. Accelerometer signals: (b) WTCC; (d) DD-CC.

they are suitable for use by people who lack the technical expertise required to properly set filter bandwidths. One disadvantage that can be highlighted for the proposed TDE scheme is higher computational cost compared to BCC and GCC. The computational cost of the scheme can be reduced by using simple wavelets and computationally efficient decomposition algorithms, such as the fast EMD algorithm [44,45] and successive VMD [46]. Incorporating the proposed WTCC and DD-CC into existing leak detection equipment is, therefore, worthwhile and recommended. Further experimental validation of the methods using leak signals from different water distribution networks will be considered in future works.

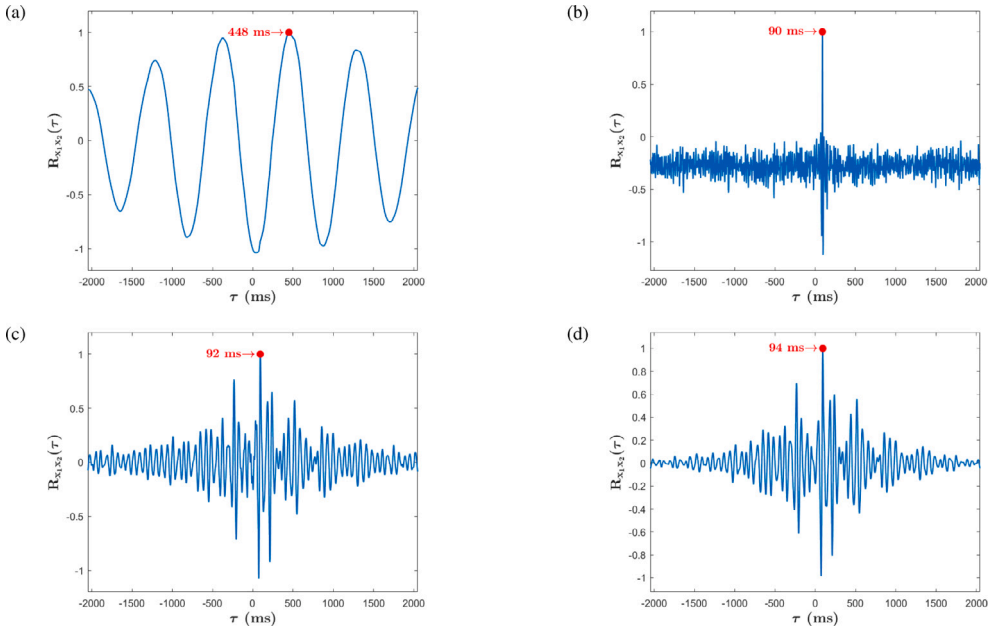


**Fig. 11.** CCFs of experimental signals. (a) Raw hydrophone signals; (b) raw accelerometer signals; (c) hydrophone signals filtered in the frequency range 10–50 Hz; (d) hydrophone signals filtered in the frequency range 10–100 Hz. Each CCF is normalised by its peak value.

**Table 3**
Time delay estimates (in ms) for raw and filtered leak signals. Values highlighted in yellow indicate estimates that differ from the true delay by less then one sample (2 ms).

| | **TDE method** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **GCC** | | | **WTCC** | | | | **DD-CC** | | | | |
| | **BCC** | **PHAT** | **ML** | **haar** | **sym2** | **coif6** | **sym11** | **db14** | **fk18** | **EMD** | **EWT** | **VMD** |
| $[f_1, f_2]$ **Hz** | $\hat{\tau}_{\text{peak}}$ [ms] $\Delta d_1$ [m] | | | | | | | | | | | |
| **Hydrophone signals** | | | | | | | | | | | | |
| **raw** | 448 | 87.6 | 87.8 | 90.8 | 90.2 | 89.7 | 89.6 | 89.6 | 89.5 | 89.5 | 90.6 | 90.5 |
| | 86.5 | 0.75 | 0.70 | 0.02 | 0.12 | 0.24 | 0.27 | 0.27 | 0.29 | 0.29 | 0.02 | 0.05 |
| **[10, 50]** | 92.7 | 91.7 | 90.1 | 91.7 | 91.1 | 90.6 | 90.6 | 90.6 | 90.5 | 91.8 | 91.8 | 91.2 |
| | 0.48 | 0.24 | 0.15 | 0.24 | 0.10 | 0.02 | 0.02 | 0.02 | 0.05 | 0.27 | 0.27 | 0.12 |
| **[5, 100]** | 93.3 | 112 | 89.3 | 90.9 | 90.2 | 89.7 | 89.6 | 89.6 | 89.5 | 89.7 | 89.5 | 91.0 |
| | 0.63 | 5.15 | 0.34 | 0.05 | 0.12 | 0.24 | 0.27 | 0.27 | 0.29 | 0.24 | 0.29 | 0.07 |
| **Accelerometer signals** | | | | | | | | | | | | |
| **raw** | 90.9 | 91.5 | 91.2 | 91.0 | 91.1 | 91.2 | 91.2 | 91.2 | 91.2 | 91.2 | 91.6 | 91.4 |
| | 0.18 | 0.04 | 0.11 | 0.16 | 0.13 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.01 | 0.07 |
| **[40, 110]** | 90.7 | 91.4 | 90.9 | 90.9 | 91.0 | 91.1 | 91.1 | 91.1 | 91.2 | 90.8 | 91.3 | 91.9 |
| | 0.23 | 0.06 | 0.18 | 0.18 | 0.16 | 0.13 | 0.13 | 0.13 | 0.11 | 0.20 | 0.08 | 0.06 |

## 7. Conclusion

This paper introduces a time delay estimation (TDE) scheme based on multi-resolution decomposition of acoustic/vibration leak signals. In the scheme, a shift-invariant wavelet transform or data-adaptive decomposition method is used to separate the signals into components corresponding to different frequency bands, and then the time delay is estimated from the components. Numerical simulations and experimental results show that the proposed scheme is capable of accurately determining time delays in situations where the basic cross-correlation (BCC) and the generalised cross-correlation (GCC) methods fail, for example, when resonances are present in the leak signals. The effectiveness of the scheme is determined primarily by the decomposition level rather than the type of multi-resolution decomposition method employed. Selection of best decomposition levels for accurate TDE results is facilitated

using a simple criterion consisting of the relative wavelet energy (RWE) and peak-to-side lobe ratio (PSR). This criterion attain a high value in decomposition levels that encompass frequency bands in which the leak signal spectrum is significant relative to the background noise and other interferences. By eliminating the need for explicit filtering of leak signals, the proposed scheme provides an attractive alternative for robust TDE in acoustic leak detection without the limitations of BCC and GCC methods.

## CRediT authorship contribution statement

**Ndubuisi Uchendu:** Investigation, Formal analysis, Validation, Software, Writing – original draft, Writing – review & editing. **Jennifer M. Muggleton:** Conceptualization, Methodology, Funding acquisition, Supervision, Project administration, Writing – review & editing. **Paul R. White:** Supervision, Project administration, Methodology, Formal analysis, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgements

The authors would like to thank Centre for Doctoral Training Sustainable Infrastructure Systems (CDT-SIS) University of Southampton, United Kingdom Water Industry Research (UKWIR), and Petroleum Technology Development Fund (PTDF) Nigeria for sponsoring this project.

## References

[1] R. Cramer, D. Shaw, R. Tulalian, P. Angelo, M. van Stuijvenberg, Detecting and correcting pipeline leaks before they become a big problem, Mar. Technol. Soc. J. 49 (1) (2015) 31–46, http://dx.doi.org/10.4031/mtsj.49.1.1.
[2] M.A. Adegboye, W.-K. Fung, A. Karnik, Recent advances in pipeline monitoring and oil leakage detection technologies: Principles and approaches, Sensors (Basel, Switzerland) 19 (11) (2019) 2548, http://dx.doi.org/10.3390/s19112548.
[3] M.J. Brennan, D.N. Chapman, P.F. Joseph, N. Metje, J.M. Muggleton, E. Rustighi, Achieving Zero Leakage by 2050: Detection and Location Methods Acoustic Leak Detection, UK Water Industry Research (UKWIR), London, United Kingdom, 2017.
[4] Y. Gao, M.J. Brennan, P.F. Joseph, A comparison of time delay estimators for the detection of leak noise signals in plastic water distribution pipes, J. Sound Vib. 292 (3–5) (2006) 552–570, http://dx.doi.org/10.1016/j.jsv.2005.08.014.
[5] S. Dray, M. Loveday, S. Tooms, T. Stephen, Leak detection on plastic pipes, in: UKWIR Technical Report, vol. 12/WM.08/47, UK Water Industry Research (UKWIR), London, UK, 2012.
[6] C. Knapp, G. Carter, The generalized correlation method for estimation of time delay, IEEE Trans. Acoust. Speech Signal Process. 24 (4) (1976) 320–327, http://dx.doi.org/10.1109/tassp.1976.1162830.
[7] N. Uchendu, J. Muggleton, E. Rustighi, P. White, Comparative study of time delay estimators for steady-state and transient acoustic leak signals, in: Proceedings of the 29th International Conference on Noise and Vibration Engineering (ISMA2020) and the 8th International Conference on Uncertainty in Structural Dynamics, USD2020, 2020, pp. 2801–2815, URL https://past.isma-isaac.be/downloads/isma2020/proceedings/Proceedings_ISMA-USD2020.pdf.
[8] F.C.L. Almeida, M.J. Brennan, P.F. Joseph, Y. Gao, A.T. Paschoalini, The effects of resonances on time delay estimation for water leak detection in plastic pipes, J. Sound Vib. 420 (2018) 315–329, http://dx.doi.org/10.1016/j.jsv.2017.06.025.
[9] F.C.L. Almeida, Improved Acoustic Methods for Leak Detection in Buried Plastic Water Distribution Pipes (Ph.D. thesis), University of Southampton, 2013.
[10] Y. Gao, M.J. Brennan, P.F. Joseph, J.M. Muggleton, O. Hunaidi, A model of the correlation function of leak noise in buried plastic pipes, J. Sound Vib. 277 (1–2) (2004) 133–148, http://dx.doi.org/10.1016/j.jsv.2003.08.045.
[11] M. Azaria, D. Hertz, Time delay estimation by generalized cross correlation methods, IEEE Trans. Acoust. Speech Signal Process. 32 (2) (1984) 280–285, http://dx.doi.org/10.1109/tassp.1984.1164314.
[12] I. Gradshteyn, I. Ryzhik, Table of Integrals, Series, and Products, seventh ed., Academic Press, USA, 2007.
[13] J.M. Steele, The Cauchy-Schwarz Master Class, Cambridge University Press, 2004, http://dx.doi.org/10.1017/cbo9780511817106.
[14] J.M. Muggleton, M.J. Brennan, Y. Gao, Determining the location of buried plastic water pipes from measurements of ground surface vibration, J. Appl. Geophys. 75 (1) (2011) 54–61, http://dx.doi.org/10.1016/j.jappgeo.2011.06.030.
[15] L. Pallotta, G. Gaetano, Accurate delay estimation for multisensor passive locating systems exploiting the cross-correlation between signals cross-correlations, IEEE Trans. Aerosp. Electron. Syst. 58 (3) (2022) 2568–2576, http://dx.doi.org/10.1109/TAES.2021.3116927.
[16] R.L. Allen, D.W. Mills, Signal Analysis: Time, Frequency, Scale, and Structure, John Wiley and Sons, 2004.
[17] O. Hunaidi, W.T. Chu, Acoustical characteristics of leak signals in plastic water distribution pipes, Appl. Acoust. 58 (3) (1999) 235–254, http://dx.doi.org/10.1016/s0003-682x(99)00013-4.
[18] S. Mallat, A Wavelet Tour of Signal Processing, third ed., Academic Press, Boston, MA, USA, 2008.
[19] I.M. Dremin, O.V. Ivanov, V.A. Nechitailo, Wavelets and their uses, Uspekhi Fizicheskih Nauk 171 (5) (2001) 465, http://dx.doi.org/10.3367/ufnr.0171.200105a.0465.
[20] D.B. Percival, A.T. Walden, Wavelet methods for time series analysis, in: Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, UK, 2000.
[21] A.P. Bradley, Shift-invariance in the discrete wavelet transform, in: Proceedings of the 7th International Conference on Digital Image Computing: Techniques and Applications, 2003, pp. 29–38.
[22] I. Daubechies, Ten lectures on wavelets, Soc. Ind. Appl. Math. (1992).
[23] M. Nielsen, On the construction and frequency localization of finite orthogonal quadrature filters, J. Approx. Theory 108 (1) (2001) 36–52.

[24] O.A. Rosso, S. Blanco, J. Yordanova, V. Kolev, A. Figliola, M. Schürmann, E. Başar, Wavelet entropy: a new tool for analysis of short duration brain electrical signals, J. Neurosci. Methods 105 (1) (2001) 65–75, http://dx.doi.org/10.1016/s0165-0270(00)00356-3.
[25] D. Casasent, G. Silbershatz, B.V.K. Vijaya Kumar, Acoustooptic matched filter correlator, Appl. Opt. 21 (13) (1982) 2356–2364, http://dx.doi.org/10.1364/AO.21.002356.
[26] T. Guo, T. Zhang, E. Lim, M. Lopez-Benitez, F. Ma, L. Yu, A review of wavelet analysis and its applications: Challenges and opportunities, IEEE Access 10 (2022) 58869–58903, http://dx.doi.org/10.1109/access.2022.3179517.
[27] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 454 (1971) (1998) 903–995, http://dx.doi.org/10.1098/rspa.1998.0193.
[28] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, IEEE Trans. Signal Process. 62 (3) (2014) 531–544, http://dx.doi.org/10.1109/tsp.2013.2288675.
[29] J. Gilles, Empirical wavelet transform, IEEE Trans. Signal Process. 61 (16) (2013) 3999–4010, http://dx.doi.org/10.1109/tsp.2013.2265222.
[30] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. 3 (1) (2011) 1–122.
[31] G. Wenjing, Z. Li, Evaluation on filter performance of variational mode decomposition and its application in separating closely spaced modes, Shock Vib. 2020 (7605683) (2020) 16, http://dx.doi.org/10.1155/2020/7605683.
[32] Z. Wu, N.E. Huang, A study of the characteristics of white noise using the empirical mode decomposition method, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 460 (2046) (2004) 1597–1611, http://dx.doi.org/10.1098/rspa.2003.1221.
[33] P. Flandrin, G. Rilling, P. Goncalves, Empirical mode decomposition as a filter bank, IEEE Signal Process. Lett. 11 (2) (2004) 112–114, http://dx.doi.org/10.1109/lsp.2003.821662.
[34] T. Huang, W. Ren, M. Lou, The orthogonal Hilbert-huang transform and its application in earthquake motion recordings analysis, in: 14th World Conference on Earthquake Engineering, 2008.
[35] M.R.A. Paterninaa, R.K. Tripathyb, A. Zamora-Mendezc, D. Dottad, Identification of electromechanical oscillatory modes based on variational mode decomposition, Electr. Power Syst. Res. 167 (2019) 71–85, http://dx.doi.org/10.1016/j.epsr.2018.10.014.
[36] W. Gu, L. Zhou, Evaluation on filter performance of variational mode decomposition and its application in separating closely spaced modes, Shock Vib. 7605683 (2020) 1–16, http://dx.doi.org/10.1155/2020/7605683.
[37] Y. Wang, C. Yeh, H.V. Young, K. Hu, M. Lo, On the computational complexity of the empirical mode decomposition algorithm, Phys. A 400 (2014) 159–167, http://dx.doi.org/10.1016/j.physa.2014.01.020.
[38] Z. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal. 1 (1) (2009) 1–41, http://dx.doi.org/10.1142/S1793536909000047.
[39] Z. Wu, N.E. Huang, On the filtering properties of the empirical mode decomposition, Adv. Adapt. Data Anal. 2 (4) (2010) 397–414, http://dx.doi.org/10.1142/S1793536910000604.
[40] I. Cespedes, Y. Huang, J. Ophir, S. Spratt, Methods for estimation of subsample time delays of digitized echo signals, Ultrason. Imaging 17 (2) (1995) 142–171.
[41] O. Hunaidi, W. Chu, A. Wang, W. Guan, Detecting leaks in plastic pipes, J. Am. Water Work. Assoc. 92 (2) (2000) 82–94, http://dx.doi.org/10.1002/j.1551-8833.2000.tb08819.x.
[42] Y. Gao, M.J. Brennan, P.F. Joseph, J.M. Muggleton, O. Hunaidi, On the selection of acoustic/vibration sensors for leak detection in plastic water pipes, J. Sound Vib. 283 (3–5) (2005) 927–941, http://dx.doi.org/10.1016/j.jsv.2004.05.004.
[43] Y. Gao, M.J. Brennan, P.F. Joseph, On the effects of reflections on time delay estimation for leak detection in buried plastic water pipes, J. Sound Vib. 325 (3) (2009) 649–663, http://dx.doi.org/10.1016/j.jsv.2009.03.037.
[44] O.O. Myakinin, V.P. Zakharov, I.A. Bratchenko, D.V. Kornilin, D.N. Artemyev, A.G. Khramov, The empirical mode decomposition algorithm via fast Fourier transform, in: Proceedings of the SPIE Optical Engineering + Applications conference, in: 9217 Applications of Digital Image Processing XXXVII, vol. 921721, 2014, http://dx.doi.org/10.1117/12.2061808.
[45] J. Zhang, F. Feng, P. Marti-Puig, C.F. Caiafa, Z. Sun, F. Duan, J. Solé-Casals, Serial-EMD: Fast empirical mode decomposition method for multi-dimensional signals based on serialization, Inform. Sci. 581 (2021) 215–232, http://dx.doi.org/10.1016/j.ins.2021.09.033.
[46] M. Nazari, S.M. Sakhaei, Successive variational mode decomposition, Signal Process. (107610) (2020) 1–10, http://dx.doi.org/10.1016/j.sigpro.2020.107610.