



Comparison of signal processing methods considering their optimal parameters using synthetic signals in a heat exchanger network simulation

Émilie Thibault^a, Francis Lebreux Désilets^a, Bruno Poulin^b, Moncef Chioua^a, Paul Stuart^{a,*}

^a Ecole Polytechnique, Montreal, Canada
^b CanmetENERGY, Varennes, Canada



ARTICLE INFO

Keywords:

Noise reduction
 Performance analysis
 EWMA
 Kalman filter
 Short time Fourier transform
 Wavelet transform

ABSTRACT

Plant sensor data contain errors that can hamper process analysis and decision-making. Those dataset are not used to their full potential due to the complexity of their processing. This paper addresses these challenges by comparing popular data processing techniques based on their ability to process sensor data, all that while using optimal parameters. The latter are obtained for all approaches using an algorithm that performs a parametric sweep. The performance of Kalman filter, exponential weighted moving average filter, short-time Fourier transform, and wavelet transform to process synthetic flow and temperature signals from a heat exchanger network simulation is quantified given two criteria: signal-to-noise ratio (SNR) and root mean square error (RMSE). It is found that most of the time, wavelet transform showed the highest RMSE reduction and SNR improvement; the wavelet transform can effectively filter signals from distinct variables from a heat exchanger network simulation when optimal parameters are selected.

1. Introduction

In process industry, several variables are measured and recorded automatically by sensors for process control, process optimization and/or economic evaluation. Sensor data present generally several types and sources of error. This is unavoidable because data acquisition systems have inherent sources of error including sensor uncertainty, signal transmission losses, fluctuations in the electrical network, incorrect calibration of measuring instruments, and deficiencies in sensor quality/installation/maintenance. There are two main types of errors: random, and systematic (gross error or bias). Readers are referred to (Bagajewicz, 2000; Bhat and Saraf, 2004; Crowe, 1996; Mah and Tamhane, 1982; Martini et al., 2014; Narasimhan and Jordache, 2000; Sun et al., 2011) for a comprehensive survey of gross error detection. Some types of measurement error, especially bias, can lead to significant deterioration in a plant's performance, optimization and control which can cause unprofitable or unsafe operation. As part of this paper, methods for random error (noise) reduction to obtain precise data for further analysis such as steady-state detection, data reconciliation, process monitoring, optimization, building a predictive model or process control are evaluated. These methods are essential in the case of these applications; depending on the subsequent application, more precision and accuracy

might be required. For instance, noise reduction methods mitigate the impact of variability sources and measurement noise on industrial control systems, which would otherwise propagate, and sometimes amplify throughout the process, leading to consequent impacts on process efficiency, products quality, and safety (Rendall and Reis, 2014). Using a suitable and tuned denoising filter can directly and significantly affect the consistency of the end products. Therefore, it is crucial to implement such a filter to establish high-quality production standards and increase efficiency. A review of industrial applications where noise was removed is presented in Thibault et al. (2023).

Industrial process data are generally corrupted with different forms of noise, and extracting the process state or relevant information from recorded noisy data, i.e., minimizing the presence of noise components in measured signals with maximum retention of the underlying process information, before performing any analysis is an essential preliminary step – often referred to as data cleaning, cleansing, filtering or denoising (Ganesan et al., 2004). Signal denoising strongly affects the outcome of higher-level tasks, and impact processes and product variability (Rendall and Reis, 2014). Reducing random errors increases the quality and added-value of the data collected and enables detection of the steady state, which is required for many decision-making processes. Signal processing is used widely in industrial processes because the right

* Corresponding author.

E-mail address: paul.stuart@polymtl.ca (P. Stuart).

Table 1

Comparison of approaches considering parameters selection.

Reference	Signal type	Techniques compared	Performance evaluation metric	Parameters available	Parameter selection justification
Ebrahimzadeh et al. (2015)	Electro-cardiogram	Least mean square (LMS), Normalized LMS (NLMS), Block LMS, Recursive least squares (RLS), Unbiased and normalized adaptive noise reduction	Mean squared error, % of noise reduction, computational complexity, stability	Partially	None
Subramanian et al. (2013)	Electro-cardiogram	Empirical mode decomposition (EMD), Butterworth high pass filter (BHPF) + EMD, BHPF+EMD+WT	SNR	Partially	None
Ganji et al. (2021)	Audio	Discrete Wavelet Transform (DWT), Linear Predictive Coding, Line Spectral Frequencies, Cepstral, Principal Component Analysis, and Power Spectral Density	Support vector machines and k nearest neighbor	No	None
Krishna et Yadav (2016)	Recorded speech	Wiener filter, adaptive filter algorithms (LMS, NLMS and RLS)	Step size, mean and variance of noise, mean square error, SNR, speed, number of iterations, complexity, convergence rate, stability	No	None
Guney et al. (2019)	Photo-acoustic	FIR low-pass, band-pass filters and DWT-based filters	RMSE, SNR, contrast-to-noise ratio	Yes	Yes
Huang et al. (2008)	Speech	Wiener filter, subspace method, spatial-temporal prediction approach	SNR, noise reduction factor	Yes	None
Sraith et Jabrane (2021)	Electro-cardiogram	Decomposition method (empirical mode decomposition (EMD), ensemble EMD, DWT, Stationary Wavelet Transform) followed by local means filtering	SNR improvement, signal-to-noise and distortion ratio, approximate entropy, fuzzy entropy	Yes	Yes
Moosavi et al. (2018)	Downhole gauges pressure data	Wavelet transform, regression-based smoothers, autoregressive smoothing methods	Mean square error	No	None
Zych et al. (2018)	Radioisotope signals	Infinite Impulse Response Butterworth fourth-order band-pass filter, signal spectrum filtering, DWT, Nadaraya-Watson kernel estimator	SNR	Partially	None
Bolaers et al. (2011)	Vibratory signals	Self adaptive noise cancellation, wavelet denoising, spectral subtraction (STFT)	Statistical indicators (kurtosis, crest factor)	Partially	Yes
Rendall et Reis (2014)	Synthetic and simulated signals	Moving average, Online multiscale filter, EWMA, Holt filter, Butterworth, Chebyshev, and elliptic filters, wavelet transform, Kalman filter	RMSE	Partially	Yes
This paper	Simulated HEN Industrial signals	EWMA, KF, STFT, DWT	RMSE, SNR	Yes	Yes

data will enable quick identification of equipment maintenance needs, leakage in the process, or faulty equipment, troubleshooting to solve operational problems, and improvements in process control and decision-making. Therefore, rigorous signal processing is a critical step in performing these tasks properly.

A variety of data processing methods reducing unwanted high-frequency components, i.e., noise, in signals are available, such as finite-impulse response filters (moving average filter, polynomial filter) and infinite-impulse response filters (exponential filter, Kalman filter) (Narasimhan and Jordache, 2000). In addition to filters, time-frequency domain methods, e.g., the short-time Fourier transform, and the wavelet transform, can be used. All those approaches will not necessarily present the same efficiency, especially when it comes to filtering process sensor signals which presents a wide range of frequency components. As part of this paper, four methods are considered: Kalman Filter (KF), Exponential Weighted Moving Average (EWMA) filter, short-time Fourier transform (STFT), and Wavelet transform (WT). These methods are selected because of their numerous applications, their industrial applicability, and their respective advantages (described in Section 3). The performance of the aforementioned signal processing methods has been compared in the literature to identify the most adequate technique for various applications (Dong et al., 2009; Dyason et al., 2017; Ibrahim et al., 2018; Perez and Barros, 2006; Pérez and Barros, 2012; Wang and Veluvolu, 2017; Yinfeng et al., 2008; Zhao et al., 2018). While the aforementioned publications are interesting contributions, the authors, in the case of this work, are interested in comparing filtering techniques for one targeted application, i.e., noise reduction.

An important aspect to consider is the fact that the performance of these data processing techniques depends on the selected parameters.

These are values that must be supplied exogenously in order to use the data processing technique. Therefore, when comparing the performance of aforementioned filtering techniques, it is essential to make sure that suitable parameters are being used.

Focusing solely on noise reduction application, comparisons between signal processing techniques (including others not considered here) with and without the parameter selection have been actively studied. Denoising filters are used in various signal processing activities, e.g., image processing, speech, spectra, sensors, in numerous fields of science such as medicine, communication, and industrial processes. Table 1 presents some papers where noise reduction performance is under consideration. The objective is to present what has been done in terms of denoising filters comparison and highlight the consideration of parameter selection in each paper. The table addresses what type of signal is assessed, which techniques are compared, what metrics were used to evaluate the different approaches and lastly if the parameters of each techniques were mentioned and if their choice were justified (if there is an explanation for the parameters selection). Concomitantly, several papers discuss ways to select (manually or automatically) the values of the parameters and the ease of tuning them for Kalman Filter (Saho and Masugi, 2015; Yuen et al., 2007), EWMA Filter (Everett, 2011), Short-Time Fourier Transform and Wavelet Transform (Abbaszadeh and Alipour, 2017; Cao et al., 2022; El-Dahshan, 2011; Hasan et al., 2011; Katunin and Przystalka, 2015; Liu et al., 2011; Shafri and Yusof, 2009; Zhou et al., 2009).

Among the publications presented in Table 1, none specifically focuses on process data, i.e. data collected from process sensors (flow, temperature, level, etc.), except Rendall et Reis (2014). First, they employed a collection of different well-established base synthetic signals

presenting a variety of features that have been used in the literature, namely, ‘HeaviSine’, ‘Doppler’, ‘Blocks’, and ‘Bumps’. Second, they explored the reactor temperature denoising. This signal is generated by the simulation of dynamical systems, namely, a continuous stirred tank reactor with nonlinear reaction kinetics and feedback control.

To the best of the authors knowledge, studies that evaluate the performance of the KF, EWMA, STFT and WT to reduce noise in a systematic and objective way for industrial process data from sensors while considering the selection of their optimal parameters are lacking in the literature. In this work, the technique’s performance are compared after adequate tuning of each techniques’ parameters.

Table 1 showcases that papers comparing signal processing techniques do not always disclose the selected approach’s parameters (Ganji et al., 2021; Hird & McDermid, 2009; Krishna and Yadav, 2016; Moosavi et al., 2018), often times, information about the parameters selection cannot be identified in the papers (Ebrahizadeh et al., 2015; Huang et al., 2008; Subramanian et al., 2013; Zych et al., 2018). However, when they are available, they are selected rather arbitrarily, or the selection method is not explicit. In fact, in most publications, the parameters are just mentioned without any explanation regarding the choice of signal processing parameters, and, in some others, the selection is trivial or succinct. Most authors did not look for the most adequate parameters.

In this work, the performance of different denoising filters is compared based on the RMSE and the SNR by using a variety of simulated signals that contains realistic features and represent the diversity of situations and events likely to be found in practice, i.e., set point change, impact of seasons, different dynamics superimposed, corrupted with noise of varying magnitudes. Therefore, the proposed comparison scenario is closer to the reality of industrial process operations. This paper contemplated single-scale and multi-scale, data-driven and model-based as well as time and frequency domain filtering approaches. The latter classifies techniques according to whether the filter is designed and tuned in the frequency or time domain. To develop filters, some rely solely on data collected from the process, i.e. data-driven approaches, whereas other use a priori knowledge about the systems generating the signals to be filtered, i.e., model-based approaches. Lastly, the design and implementation of filters can consider time or frequency information of a single granularity, i.e. single-scale techniques, or different scales of time and frequency, i.e., multi-scale.

Due to the crucial role of filtering, a comprehensive and meticulous comparative analysis of existing denoising techniques is undertaken. The outcome of this study will provide practical and effective recommendations to aid practitioners in selecting the most appropriate approach for their specific circumstances. Hence, the objective of this paper is to assess the optimal parameters of each signal processing method, to compare which approach shows the best performance towards synthetic process sensor data and to elaborate on how their parameters are generally established. In this regard, **Section 2** of this paper presents the different parameters of various signal processing techniques and **Section 3** addresses the pros and cons of each. **Section 4** elaborates on the methodology for comparing those techniques based on their performance in reducing errors in process data considering their optimal parameters. Then, **Section 5** reports the comparison of the signal processing methods by firstly constructing and characterizing synthetic noisy signals, assessing their denoising performance and discussing their optimal parameters establishment. **Section 6** gives a conclusion of the present work.

2. Signal processing methods: parameters

The method-specific parameters of KF, EWMA, STFT and WT are described below. They are the input parameters, the information required to use the various signal processing techniques.

2.1. Kalman filter

The Kalman filter (KF) is a recursive linear estimation algorithm with the goal, given a sequence of noisy observations, of estimating the state of a system such that the estimation error is minimized. The KF is a powerful tool for estimating the states of a linear system in the presence of noise. It minimizes the variance of the estimation error (the mean square error of the estimated states). This filter consists of a prediction phase and an update phase. Information about KF can be found in several references (Dash and Chilukuri, 2004; Dyason et al., 2017; Ibrahim et al., 2018; Kim and Bang, 2018; Labbe Jr, 2020; Narayan et al., 2013; Park et al., 2019). A brief description of the KF parameters is provided here.

The system is described by the following state-space representation:

$$x_t = F_t x_{t-1} + B_t u_t + w_t \quad (1)$$

$$y_t = H_t x_t + v_t \quad (2)$$

where F_t is the state transition matrix, x_{t-1} is the previous state, B_t is the input matrix, u_t is the control vector, w_t is the process noise vector, y_t is an observation vector, H_t is the observation matrix, x_t is the state, and v_t is an observation noise vector.

Prediction phase of the filter: In this first phase, a prediction (estimation) of the state $\hat{x}_{t|t-1}$ and the state estimate covariance matrix $P_{t|t-1}$ are calculated at time t as follows:

$$\hat{x}_{t|t-1} = F_t \hat{x}_{t-1|t-1} + B_t u_t \quad (3)$$

$$P_{t|t-1} = F_t P_{t-1|t-1} F_t^T + Q_t \quad (4)$$

where Q_t is the process or system noise (co)variance matrix. The covariance matrix must be initialized at the beginning of the process.

Update phase of the filter: In the second phase, the measurement residual vector \tilde{z}_t and the residual variance matrix S_t are calculated as follows:

$$\tilde{z}_t = y_t - H_t \hat{x}_{t|t-1} \quad (5)$$

$$S_t = H_t P_{t|t-1} H_t^T + R_t \quad (6)$$

where R_t is the measurement noise (co)variance matrix. Eqs. (5) and 6 give the updated state estimation vector

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t \tilde{z}_t \quad (7)$$

and the updated estimate variance matrix

$$P_{t|t} = (I - K_t H_t) P_{t|t-1} \quad (8)$$

where K_t is the optimal Kalman gain that minimizes the mean square error $E[(x_t - \hat{x}_{t|t})^2]$:

$$K_t = P_{t|t-1} H_t^T S_t^{-1} \quad (9)$$

The Kalman filter parameters represent the distribution of noise in the system. The first one is the process noise covariance matrix. Process noise is associated with noise in the states. This value indicates, to some extent, the nonlinearity of the state-space model, and the larger it is, the more nonlinearity the model shows (Dong et al., 2009). Typically, process noise causes signal drift, and the result is related to the system dynamics. The covariance matrix Q_t expresses how much variance and covariance exist in the process. The matrix diagonal contains the variance of each state variable, and the off-diagonal elements contain the covariances between various state variables (Labbe Jr, 2020). If the state parameters are independent, then the matrix has only a diagonal. Setting the process noise to 0 tells the KF that the process model is perfect (Labbe Jr, 2020). The second parameter is the measurement noise covariance matrix. Measurement noise comes from a sensor and

represents the noise characteristic of the measurement device. The measurement noise matrix R_t models the noise from sensors as a covariance matrix (Labbe Jr, 2020), which varies accordingly to sensor sensitivity. The covariance matrix contains the variance of each measurement and is calculated from the sensor accuracy. This is obtained through the standard deviation of the measured values using the signal of interest.

If Q_t is smaller than R_t , K_t tends to be low, then the prediction is more likely to carry more significance because the system's uncertainty is inherently low. Conversely, when Q_t is relatively larger than R_t , the estimate will rely more on the measured value (Rendall and Reis, 2014).

2.2. EWMA filter

The exponentially weighted moving average filter (Evans, 2019; Nounou and Bakshi, 2000; Perry, 2010; Smith, T. H. & Boning, 1997), a popular infinite impulse response (IIR) filter (recursive filter), is used to smooth time-series data by exponentially averaging incoming reading with all previous measurements. This technique is often employed to obtain an average of samples at the time they are "arriving" to the control system, i.e., it is used on-line. The EWMA is used for its noise filtering capabilities, it cancels the effects of noise at the output. The noise superimposed to the input is assumed Gaussian, with zero mean value and a known variance. This low pass filter reduces high frequency components. The EWMA filter makes it possible to specify the weight of the last sample acquired versus the previous filtered value using a smoothing factor α . The average is "moving" in the way that it is computed each time a new sample is obtained. EWMA is mathematically defined as follows:

$$S_t = \alpha x(t) + (1 - \alpha)S_{t-1} \quad (10)$$

where α is the smoothing factor and represents the weight, it defines the cut-off frequency above which features are eliminated; $x(t)$ is the current input value; and $S(t)$ is the filtered output at time step t . This recursive formula saves memory and makes fast calculation.

The smoothing factor, α , ranges from 0 to 1; the higher the value, the greater is the impact of the last reading, allowing the filter to follow closely sudden changes. Hence, less smoothing is done, and more weight is put on this last entry. On the other hand, a small value of α means more weight on the average, meaning that more smoothing is done and the last entry has little impact; the filter is less responsive to new measurements (Evans, 2019). Hence, a value of one results in no filtering (the output is only the current observation), and a value of zero does not take new measurements into account. The value of the smoothing parameter is typically correlated with measurement precision or accuracy.

2.3. Short-time Fourier transform

The Fourier transform (FT) gives the characteristics of a signal in the frequency domain. However, time information is lost because the transformed signal has a frequency on the x-axis and an amplitude on the y-axis. This would not be an issue if the signal under evaluation are stationary (constant frequency content). The short-time Fourier transform (STFT) is used to process time-variant frequency signals. STFT carries out the FT on a short period, accounting for the time component. This makes STFT adequate for non-stationary signals. Information about the short-time Fourier transform can be found in several references (Kehtarnavaz, 2008; Manhertz et al., 2016; Rahimi et al., 2007; Selesnick, 2009; Smith III, 2011; Zabidi et al., 2012).

The STFT is defined as

$$X_L[k] = \sum_{t=0}^{N-1} x[t] w[t - LH] e^{-j2\pi nk/N} \text{ for } L = 0, 1, \dots \quad (11)$$

where $x[t]$ is the time-discrete signal, $w[t]$ is the analysis window, t is the time-domain index, L is the frame number, and H is the hop (step) size, or the time separation between adjacent frames.

STFT is an effective signal processing technique that can decompose a time-domain signal into its time-frequency components. It divides a whole signal into shorter data frames of equal length by multiplying it by an analysis window function. Consequently, the frequency content is stored in time windows, where each of them is then analyzed using the fast Fourier transform. This procedure is called the overlap-add method (Bahoura, 2019). STFT reveals the frequency content of the signal at each time point (Zabidi et al., 2012).

STFT is performing iterations over L , hence skipping through time with L , and jumping to the adjacent window, i.e., going to the next time interval by hop size. The STFT is usually computed using overlapping windows, where the adjacent segments share some overlapped data instead of just being end-to-end. The purpose of using overlapping frames is to avoid spectral leakage, achieve better time locality, and obtain smoother results (Smith III, 2011). The output of the STFT is a sequence of spectra having the same size.

The first parameter used in STFT is the sampling rate; it corresponds to the number of data points per second. A higher sampling rate results in more samples and therefore more flexibility in choosing the window size, the FFT size, and the hop size. In the case of this paper, the discrete synthetic signals (Section 4.1) are sampled every minute, and consequently the sampling rate is a fixed value.

The second parameter is the Fast Fourier transform (FFT) size (or length); it is displayed as a number of bins in the analysis window, rather than as an absolute value. Hence, it defines the number of equal bins used to divide the window, which is typically a power of two. The number of bins is proportional to the number of samples in the window. It expresses the number of frequency bands into which the analysis window will be cut (Diatkine, 2011). A bin is a spectrum sample that defines the frequency resolution of the window (Diatkine, 2011). The FFT size is independent of window length (size), but it must be greater than or equal to the latter.

Third, the window function (Diatkine, 2011; Mitra, 2000; Smith, 2011) is a real function that is symmetric around the origin. Window functions are defined and distinguished by their main lobe (peak at the center) width and their highest side lobe (small lobe next to the main lobe), i.e., peak side-lobe, amplitude relative to the main lobe. The many windows available include Blackman, Chebyshev, Gaussian, Hamming, Hann, Rectangular, Triangular, and others. For each window function, there is a compromise with respect to these two parameters. The trade-off is that the wider the main lobe, the lower must be the peak side-lobe amplitude, and vice-versa. The compromise is between spectral resolution and spectral leakage. A window with a smaller side lobe yields less spectral leakage, but has increased main-lobe width, which limits spectral resolution. Increasing the window length for a given accepted leakage level will increase spectral resolution, but decrease the time domain resolution. To minimize the impact of the latter side effect, adjacent windows are overlapped. The window length is an important parameter that specifies the frame length of the window analysis function. Its value must be greater than one, and it represents the number of samples.

The width of the main lobe is primarily determined by the window length, i.e., increasing its value will decrease the main-lobe width. The window length (Diatkine, 2011; Mitra, 2000; Smith, 2011) determines the resolution in both the time and frequency domains, i.e., the spectro-temporal resolution of the signal; a good time resolution requires a short window, whereas a good frequency resolution requires a long window. It is not possible to have at the same time both high time resolution and high frequency. This phenomenon is known as the Gabor limit, or the Heisenberg-Gabor limit, or Heisenberg's Uncertainty principle. In the same way that there is a trade-off in window choice, there is also one in window size.

Window functions can be described as having high resolution, but



Fig. 1. Hop size, inspired by MathWorks (2023).

low dynamic range, or as having high dynamic range, but low resolution. The latter can distinguish features with different frequencies and amplitudes, whereas the former can distinguish components with similar amplitudes and frequencies. High-dynamic-range windows have a wide main lobe and very low amplitude peak side lobes and are used in wideband applications. The rectangular window – the simplest function – shows the narrowest main-lobe width and the highest-peak side lobe, resulting in what is described as high resolution. The Gaussian, Blackman-Harris, and Kaiser-Bessel windows have a relatively large dynamic range. On the dynamic range spectrum, between these window functions, there are moderate windows such as the Hamming and Hann windows.

Window functions are used in signal analysis, i.e., as an analysis window, as well as in signal reconstruction, i.e., as a synthesis window. The two do not have to be the same, but in this paper, they are.

Finally, the hop size (Diatkine, 2011; Mitra, 2000; Smith, J. O., 2011) corresponds to the number of samples between each time frame, i.e., the window step. The window is moved forward in time by the step size of H samples. In STFT, the windows are not juxtaposed successively, but overlap to improve the precision of the analysis. They are overlapped by a certain overlap length. Overlapping affects the calculation precision, improves the temporal resolution, but may increase the calculation time. Hence, the hop size is the difference between the window length and the overlap length of two consecutive frames (segments), as shown in Fig. 1. Typical values of the overlap length are one-half, two-thirds, or three-quarters of window length.

2.4. Wavelet transform

The wavelet transform provides the characteristics of non-stationary signals in the time-frequency domain. The WT is defined as the projection of a signal $x(t)$ onto the scaled wavelet, i.e., the convolution of $x(t)$

and $\psi_{\tau,a}(t)$. The wavelet is an oscillatory window function of finite length, i.e., compactly supported, and defined by translation and dilation:

$$\psi_{\tau,a}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-\tau}{a}\right) \quad (12)$$

The continuous wavelet transform is expressed as follows:

$$W(\tau, a) = \langle x(t), \psi_{\tau,a}(t) \rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi^*\left(\frac{t-\tau}{a}\right) dt \quad (13)$$

where τ is the shift parameter used to control the translation, a is the scale parameter, $\psi(t)$ is the mother wavelet function, and $*$ indicates complex conjugation. When the values of τ and a are continuous, the WT is called continuous, and when they are discrete, it is called discrete. DWT suppresses the redundancy of the continuous WT while conserving the reversibility property.

A wavelet transform decomposes a signal into several levels by using a set of filters, i.e., low-pass and high-pass filters, to produce the approximate and detailed coefficients respectively. These filters enable the signal to be represented by a sum of approximations (scaling functions $\Phi_{j,i}$) that form the smoothed signal and a sum of details (noise) represented by wavelet functions $\psi_{j,k}$ at different levels. The levels are used to display the characteristics of the signal over time and at various frequencies (Bakshi, Bhavik R., 1999):

$$x(t) = x_0 = \sum_{i \in I_0} c_{0,i} \Phi_{0,i} = \sum_{i \in I_1} c_{1,i} \Phi_{1,i} + \sum_{k \in K_1} d_{1,k} \psi_{1,k} \quad (14)$$

$$= \sum_{i \in I_2} c_{2,i} \Phi_{2,i} + \left(\sum_{k \in K_1} d_{1,k} \psi_{1,k} + \sum_{k \in K_2} d_{2,k} \psi_{2,k} \right) \quad (15)$$

$$= \sum_{i \in I_j} c_{j,i} \Phi_{j,i} + \left(\sum_{k \in K_j} d_{j,k} \psi_{j,k} + \sum_{L=1}^j \sum_{k \in K_L} d_{L,k} \psi_{L,k} \right) \quad (16)$$

$$= \sum_{i \in I_j} c_{j,i} \Phi_{j,i} + \left(\sum_{j=1}^J \sum_{k \in K_L} d_{j,k} \psi_{j,k} \right) \quad (17)$$

The first term, $\sum_{i \in I_j} c_{j,i} \Phi_{j,i}$, corresponds to the smooth signal that

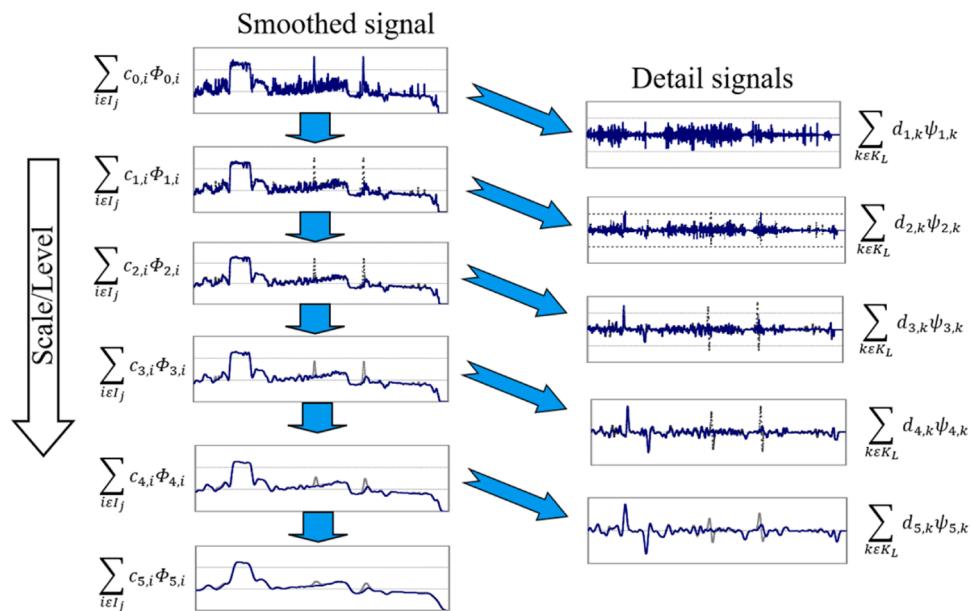


Fig. 2. Decomposition of a signal into smoothed and detailed signals, inspired by (Korbel et al., 2014).

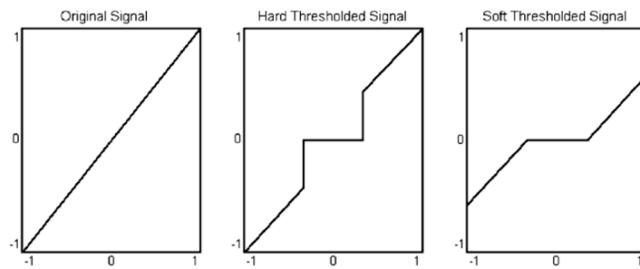


Fig. 3. Soft and Hard Thresholding.

contains the low-frequency elements of the original signal present at level j (Fig. 2, left). The second term, $\sum_{j=1}^J \sum_{k \in K_j} d_{j,k} \psi_{j,k}$, represents the detailed signal containing the high-frequency elements (noise) of the original signal at this level (Fig. 2, right) (Jiang et al., 2003; Korbel et al., 2014). Therefore, the difference between the smooth signal of level 1 ($s = 1$) and the detailed signal of level 2 ($s = 2$) gives the smooth signal at level 2, and so forth (Bakshi, B. R. and Stephanopoulos, 1994). The multi-level representation makes it possible to adequately identify signal characteristics located at different times and places, thus reducing noise in the data (Nounou and Bakshi, 1999). Information about wavelet transforms can be found in several references (Bakshi, 1999; Daubechies, 1992a, 1992b; Doymaz et al., 2001; Fugal, 2009; Mallat, 2009; Rioul and Vetterli, 1991).

The quality of DWT denoising is strongly affected by several critical parameters, including the mother wavelet function type, the threshold rule, the decomposition level, and the threshold estimate. First, the mother wavelet function is usually selected from families that somewhat resemble the shape of the desired signal. Many types of wavelet families exist (supported, orthogonal, biorthogonal, etc.), and they contain wavelet functions of different orders: Daubechies (Haar, db2...db45), Biorthogonal (bior1.1...bior1.5, bior2.2...bior2.8, bior3.1...bior3.9), Coiflets (coif1...coif5), Symlets (sym2...sym45), Morlet, Meyer, Splines, Gaussian, etc.

Second, for the threshold rule, either soft or hard thresholding can be used (Fig. 3). For a given threshold λ and value of the wavelet coefficient d , hard thresholding sets to zero those coefficients with absolute values lower than the selected threshold and maintains the same value for coefficients that exceed it:

$$D^H(d|\lambda) = \begin{cases} 0, & |d| \leq \lambda \\ d, & |d| > \lambda \end{cases} \quad (18)$$

Soft thresholding, which is an extension of hard thresholding, acts the same way for values lower than the threshold, but shrinks the coefficient values by the value of the threshold limit:

$$D^S(d|\lambda) = \begin{cases} 0, & |d| \leq \lambda \\ d - \lambda, & |d| > \lambda \\ d + \lambda, & |d| < -\lambda \end{cases} \quad (19)$$

Soft thresholding, also called wavelet shrinkage, takes both positive and negative coefficient values, and shrinks them towards zero, whereas hard thresholding either keeps or removes coefficients. The latter may result in larger variance in the signal after reconstruction but can better represent peaks and discontinuities. The former, on the other hand, has a larger bias, but gives good visual quality of filtering.

Third, the scale factor – the decomposition level, is related to the error frequency, i.e., the frequency correlates with the scale. Small scale values correspond to high-frequency information (compressed signals), i.e., rapid changes in the detailed signal, whereas high values correspond to low-frequency components (stretched-out signals) associated with a non-detailed global view of the signal. Scaling means either dilatation or compression of the signal.

Finally, the fourth parameter is the thresholding selection rule – the

threshold estimate. The threshold value is generally defined based on the standard deviation of the error (noise) amplitude (El-Dahshan, 2011). Selecting a proper threshold value is essential for effective denoising. Some thresholds are: Rigrsure – a threshold selected using the Stein's Unbiased Risk Estimate (SURE) principle; Universal – a threshold equal to universal $\sigma_j \sqrt{2 \log N}$; and Minimax – a threshold equal to max MSE. The threshold is selected according to the nature of the noise. The Universal threshold reduces noise more efficiently, whereas the other two are more conservative and are more convenient when small details of the signal lie near the noise range (MathWorks, 2020).

3. Analysis of signal filtering techniques

This section presents some advantages and drawback for each signal filtering technique and discusses how their parameters are selected.

First, the EWMA filter is widely used in process industries, its application is very straightforward (Belitsky, 2001; Harrison and Stuart, 2011). A positive aspect of this approach is that its recursive nature makes it easy to implement, and it requires low computational effort (reduced memory and fast calculation); EWMA filter consumes significantly less memory and works faster than most techniques. On the downside, enough knowledge on the process data is required to set the smoothing factor α – the weight of the last entering value in the data set. Some tests might be required to find the most suitable value according to the application.

Second, the Kalman Filter is theoretically an optimal estimator; it can correct itself between the prediction and update phases, and provides high-quality-estimation with a minimum variance of the estimator error. It will accurately estimate the states affected by external disturbance. A major advantage of the Kalman Filter is its ability to provide the estimate with an indicator of the filter accuracy, i.e. the prediction error covariance (Harkat et al., 2016). Of all existing filters, Kalman is the only one that minimizes the covariance of the estimation error (Simon, 2001). Also, compared to other filtering methods, its advantage lies in its capacity to effectively integrate input and output information in the form of a model with process measurements. Kalman Filter lean to good results and will outperform data-driven techniques when a reasonable model of the system is provided. On the other hand, this filter needs as an input the process noise variance as well as the observation noise variance. Selecting those values require great knowledge of the process and the sensors (sensor reaction time, sensor bandwidth).

Additionally, to use the KF, the state transition matrix needs to be established. To do so, a perfect model is rarely available, especially in process industries. Equations that compose the model (mass and energy balances) are never exact due to the many assumptions behind them. To the best of the authors knowledge, no one ever used the KF as a noise reducer on process sensor readings. What comes closest to this are publications related to the estimation of states that are not measurable or not measured – one of principal objective of the filter – for reactors based on a model (MacGregor et al., 1986).

Moreover, at a plant, process data are generated by nonlinear systems, and their dynamics are not very well described by a linear state space model. Consequently, the Kalman Filter will not have an optimal performance if the estimated model can not capture the process dynamics efficiently.

In order to be fair with the three other processing technique discussed here, which have no clue about the process model (mass and energy balances) of the HEN, Eqs. (1), (3) and (4) need to be adjusted. Firstly, there is no state transition ($F_t = 1$). Thus, it is not a matrix, it is a scalar value; the Kalman filter is used in a univariate way – it is processing one variable at the time. Also, the system does not have any input control vector (u_t), hence B_t , the control matrix, is null. The assumption here is that the only input in the system is the process noise. Therefore, the (co)variance matrix update (Eq. (4)) depends greatly on Q_t ; the states vary due to the process noise variance matrix, the signals

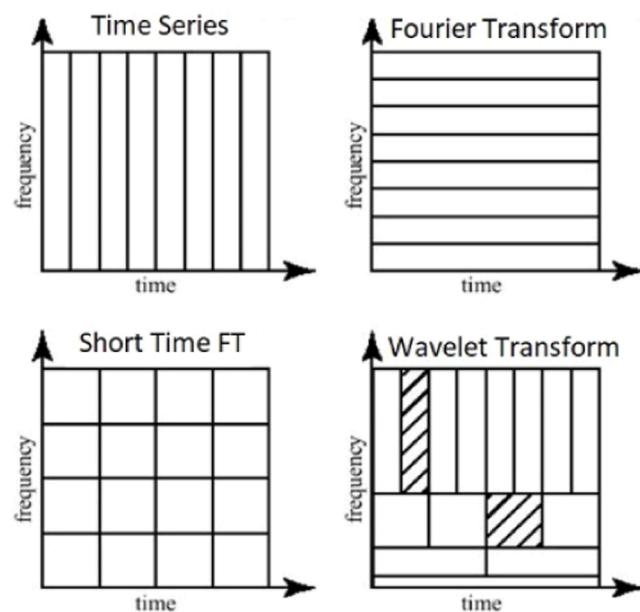


Fig. 4. STFT Fixed-Window (constant aspect ratio) and WT Variable-Window (Ganesan et al., 2004).

are filtered according to this. Hence, the implementation of the KF approach for reducing noise in simulated sensor readings does not consider the model, it only depends on the measurements quality. In other words, the part regarding the model dynamics is ignored, the algorithm focuses on the data and noise aspects. Otherwise, considering the model information in the KF technique would automatically introduce a bias since the other approaches don't have access to this information.

Conventional linear filters are effective in reducing random errors above or below a certain frequency (out-band). These are simple methods that are easy to use and implement. However, their action are limited for in-band signals since the data is represented for only one frequency level. In other words, these filters are single scale in nature as the basis functions have a fixed time-frequency localization (Nounou and Bakshi, 2000). Indeed, when they are applied to in-band signals, they also remove the signal of interest. Consequently, linear filters must find a balance between efficient removal of temporally global random error and accurate representation of temporally localized characteristics retained in the signal (Nounou and Bakshi, 2000).

Sensor measurement data are well known to be non-stationary and multiscale given the variety of events occurring in a process (they often present features in different regions of the time-frequency domain), and thus are not effectively described by single-scale approaches developed in the time or frequency domain, they are less suited for these analysis. In fact, it is quite common for measured signals to exhibit a diverse range of dynamical and stochastic characteristics. These characteristics can arise from factors such as sensor noise, high-frequency interferences, long and short-term process variability, and rapid transitions (Rendall and Reis, 2014). These complexities make it difficult to accurately describe the signals using single-scale approaches. Such approaches often require compromising certain features in favor of others, thereby limiting their ability to filter and describe the signals effectively. Industrial plant data require denoising filters able to pave the time-frequency plane to capture their multiscale features. Thus, the adoption of either one of the two approaches mentioned above, i.e. EWMA and KF, should consider the nature of the signals to be denoised in order to check if they are dominated by single-scale characteristics or present multiscale dynamical patterns.

This problem is partially overcome with Fourier Transform (FT), which allows frequency analysis of data. On the other hand, it is

inefficient in extracting temporal information from signal characteristics (Fugal, 2009). Moreover, few spectral methods rely on the implicit fundamental assumption of signals being periodic and stationary, which is not the case for signals from process sensors (Ganesan et al., 2004). Indeed, FT is not suitable for analyzing non-stationary signals and especially those containing sudden changes. For non-stationary signals, the FT takes the average of the frequency components over the whole duration of the signal (Ganesan et al., 2004). However, these downsides can be overcome with STFT that is an excellent noise reduction tool for non-stationary signals. STFT, in addition to a good frequency resolution, allows evaluating at what time a particular frequency event occurs in the signal.

One limitation of STFT is that the resolution remains constant over time, i.e., STFT uses a single, fixed analysis window, which means that the width of the time and frequency window is constant throughout the entire STFT process (Fig. 4). Hence the time-frequency resolution is quite a limitation mainly due to using a fixed window size; once it is fixed, that window is the same for all frequencies. It is difficult to know which size would be adequate for the signal; a wide window size would have good frequency resolution but poor time resolution and vice versa for a too narrow window. The latter is not adequate for low frequency events and the former is not required for high-frequency components. Thus, the accuracy of the filtering method is greatly limited by the shape of the window. Therefore, STFT is not as well suited for signals with a wide range of frequencies and for locating singularities or short time events of high frequency (Fugal, 2009; Ganesan et al., 2004). This aspect make it difficult to get proper characteristics from the signals (Park et al., 2016). Henceforth, a major disadvantage of the STFT is the resolution trade-off between time and frequency.

With WT, the notion of scale is introduced as an alternative to frequency, leading to a so-called *time-scale representation* – the scale factor is defined as the inverse of frequency. It is the equivalent of the time-frequency plane used in the STFT. In contrast to the STFT, which uses a single analysis window, WT uses windows of varying sizes adjusted to capture the signal important characteristics: narrow windows for high frequencies and broader windows for low frequencies (Fig. 11). As a result, WT achieves an ideal balance of time-scale resolution (it has more flexibility in time and scale resolution). Also, compared to STFT, it is better at revealing short-time changes or characteristics in signals occurring at high frequency (carrying out local analysis) and describing local features.

The wavelet Transform is known to be the most common, powerful, and successful technique for non-stationary signal denoising (Alyasseri et al., 2020). DWT is the most generally used time-frequency filtering method (Ben Slama et al., 2018). By looking at a signal at various scales and analyzing it with different resolutions (given its inherent multi-resolution capabilities), WT gets very efficient at detecting and representing elements in the signal (e.g., noise, abnormalities, features, and trends). This approach can remove unwanted features that are superimposed on the signal of interest through frequency decomposition and the use of statistically based thresholding, without impacting it. In other words, the implementation of wavelet transform results in minimum distortion of the filtered data (WT is able to decompose complex signals and then reconstruct them without losing information relevant to the process) and helps to extract features that are associated only with sensor dynamics (Upadhyaya et al., 2014).

This method provides excellent time-frequency localized information and can deliver a multiresolution representation of a signal for multiscale events examination. It also provides an excellent means to analyze nonstationary and autocorrelated non-Gaussian data since it generates stationary and independently and identically distributed Gaussian wavelet coefficients irrespective of the input data distribution, even if it is non-stationary. Hence, the WT analysis gives the opportunity of applying statistical procedures that require Gaussian distributed data to the wavelet decomposed data.

Multiscale methods are characterized by their rapid adaptation to

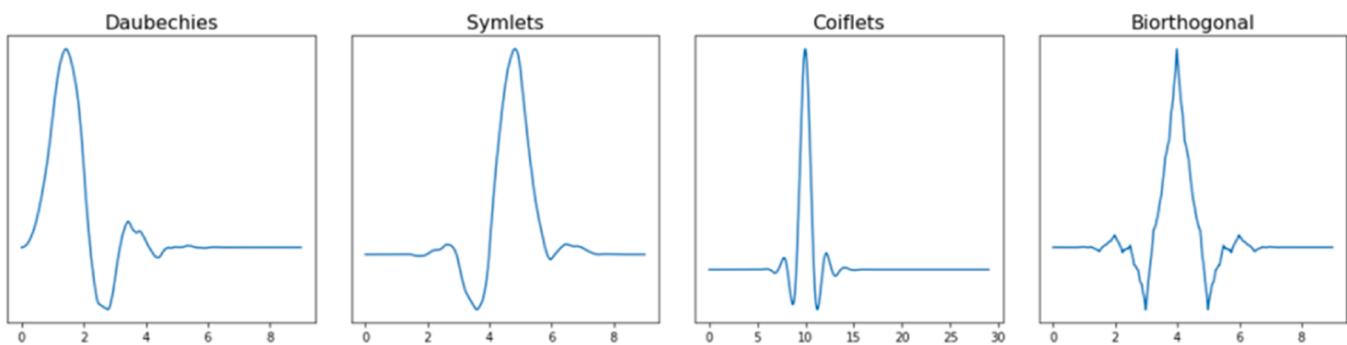


Fig. 5. Examples of Wavelets Functions, inspired by (Taspinar, 2018).

Table 2

Ranges of commonly used method-specific denoising parameters for KF, EWMA, STFT and WT.

Methods	Method-specific parameters	Range or value
Kalman filter	Measurement noise R_t	Standard deviation of the noise
	Process noise Q_t	$R_t/1000, R_t/100, R_t/10, R_t$ (Ma'arif et al., 2020)
Exponential weighted moving average	Smoothing factor α	0.05 to 0.95 (with a step of 0.05)
	Window Type	Blackman; Chebyshev; Gaussian; Hamming; Hann; Rectangular; Triangular
Short time Fourier transform	Window Size	Hour; Day; Week; Month
	FFT length	Twice the window size
Wavelet transform	Sampling rate	One point per minute
	Overlap length	1%; 10%; 25%; 50%; 75% overlap of the window size
	Thresholding value (multiplier of the original signal's std dev)	1 to 10 (with a step of 1)
	Wavelet function	Daubechies (db2-db15); Symlet (sym2-sym15); Coiflet (coif1-coif5); Biorthogonal (bior1.1-bior6.8)
	Scale (Decomposition level)	1 to 10
	Thresholding function	Soft threshold; Hard threshold
	Threshold selection rule	Rigsure; Universal; Minimax

local signal features, thanks to their multiresolution structure and parallel processing of information within various frequency bands. This grants them the ability to detect patterns across the entire frequency domain with equal effectiveness. Unlike single-scale methods, which are limited by their design compromise towards certain signal features, i.e., smooth trends (more intensive denoising) or abrupt changes (less denoised signals) (Rendall and Reis, 2014).

Additionally, signals with sharp changes might be better analyzed with wavelet functions than with a smooth sinusoid provided by STFT. Wavelets are used in a broad range of application fields because it can process signals of various shapes thanks to the different families of wavelets. Indeed, STFT decomposes signals into sines and cosines whereas WT offers many different basis functions for decomposition that have various shapes (Fig. 5). Hence, these basis functions can fit to various signal's characteristics, and it gives the WT the ability to denoise unusual signals.

4. Methodology for comparing signal processing methods

This paper proposes to compare the denoising performance of four signal processing methods on synthetic signals (with known errors) from

a heat exchanger network (HEN) simulation, i.e., assess their performance in adequately removing errors. The performance analysis is carried out using two criteria: the root-mean-square error and the signal-to-noise ratio between processed signal and the signal of interest. Therefore, the RMSE and SNR characterize in an unambiguous way the various approaches performance based on the values of the signal of interest for each noisy signal to be denoised. Furthermore, this paper considers the selection of the optimal parameters of each method. The adjustable parameters of all filtering techniques are tuned for optimal performance to ensure that the filters being analyzed are subjected to comparably fair conditions. The parameters tuning is based on the minimization of the error between the signal filtered by the processing technique and the signal of interest (presented in Fig. 7) using a parametric sweep (grid search). The different values of the parameters for each filtering technique (see Section 2) are presented in Table 2. Thus, the methodology is divided in the tuning stage, i.e., identifying the parameters that minimizes RMSE, and with the filters tuned in an optimal way, the methods are then compared based on two metrics in the testing stage (assessment the estimation accuracy of the underlying signal of interest for all technique).

Regarding the Kalman Filter parameters (a covariance matrix that reflects the uncertainty of the measurements and another one that reflects the uncertainty on the model), when experimenting with various values for the measurement noise and process noise, their ratio is what matter the most. In other words, when tuning a Kalman filter, rather than adjusting separately Q and R, it is their ratio that matters, it is related to the trust in the model versus the trust in the measurements. The measurement noise is assumed to be the measurement variance of the signal of interest. The different ratios for the process noise used in this work are taken from (Ma'arif et al., 2020).

A survey done with the pulp and paper sector regarding data analysis practices found that common analyses in industry are done on a daily, weekly or monthly basis in order to take various meaningful decisions (Thibault et al., 2023). That served as the basis for the STFT window size selection.

5. Comparison of signal processing methods for error detection

First, the synthetic signals are obtained through a simulation (heat exchanger network in this case) to which fictional random errors are added. Next, the optimal parameters are identified for each processing technique using a parametric sweep that keeps the parameters that minimizes the error between the filtered signal and the signal of interest, i.e., the RMSE. Then, the optimal parameters are used to process the synthetic signals and appraise the performance of the Kalman filter, EWMA filter, short-time Fourier transform, and wavelet transform using quantitative criteria, i.e., RMSE and SNR. Lastly, the approaches are compared qualitatively in the discussion.

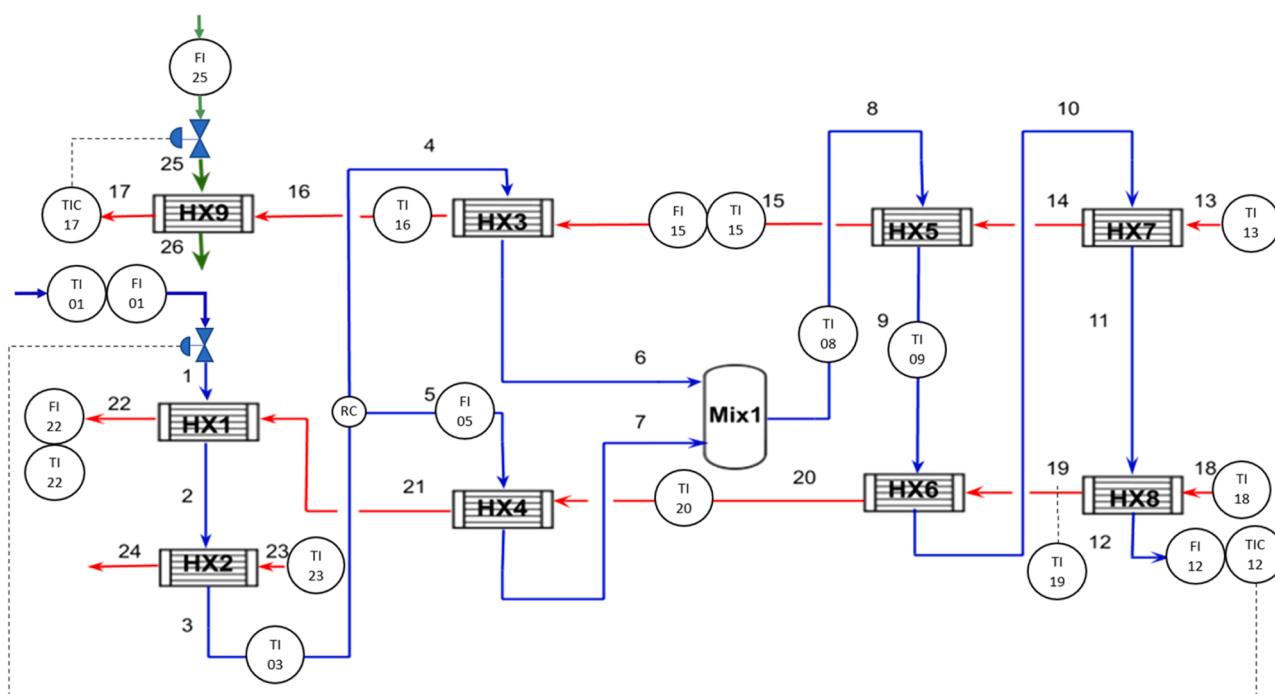


Fig. 6. Heat exchanger network simulation: flow and temperature signals.

5.1. Simulation and synthetic signal

Industrial process signals are complex, this reflects the complex nature of the systems that generate them. These signals are often non-stationary and non-periodic. Therefore, the synthetic signals used in this work mimic processes dynamics likely to be found in real-world industrial applications.

The synthetic (simulated) signals are obtained from the simulation of a heat exchanger network, inspired by a crude oil preheating train, developed in VBA Excel based on the number of transfer units (NTU) approach (Fig. 6) (Bergman et al., 2011). Flow and temperature sensors are not available on all streams, as it would be the case in a real industrial process. This simulation provides six flow signals and 14 temperature signals. The blue streams are water, whereas the red streams have the specific heat capacity (C_p) of oil.

Out of the twenty (20) sensors, two (2) are part of a control loop and 18 are indicators. The temperature controller TIC12 adjusts the opening of a control valve on stream 1 according to a temperature set point on stream 12. Additionally, stream 25 is cooling stream 16, it has the Cp of air, and its flow is controlled according to the temperature of stream 17. Thus, TIC17 adjusts the opening of a control valve on stream 25 to control the temperature of stream 17. Lastly, the split ratio between stream 4 and 5 can also be controlled. Fig. 6 shows the ratio controller (RC) at the split.

The simulation signals are generated considering the rate of change by using a superimposed high frequency random function over seasonal and operational variations (trend of a process). Small high frequency

random values are used to demonstrate the fast variability that a signal can have. The latter are generated uniformly around the trend. Those small random values are normally distributed given the fact that the hypothesis stating that random errors are following a gaussian/normal distribution is generally accepted, it is a good approximation of many real-world environments (Park et al., 2019). Random errors are added considering the amplitude of the measurement to provide realistic and interesting process signals.

The simulation provided a year of data sampled every minute. These simulated signals aim to be as close as possible to real signals without random errors; they are representative of true and realistic process signals. For that reason, sequential calculation method has been used to simulate the noise.

To obtain a good representation of time-series data from process sensors, understand the process dynamics and mimic real process signals, a setpoint change is applied to the stream 12 temperature controller (TIC12) (Fig. 7). This allows to see how the other process variables react to this change and thus evaluate which filtering approach is robust enough to reduce errors notwithstanding the setpoint change; that is, to assess if the approaches are able to adjust to this change, reduce random error while estimating the signal's trend.

Therefore, the HEN simulates normal continuous operations, i.e., (no tube leaks, fouling, sudden variation of the split ratio, heat exchanger by-pass), with small random variations as well as a setpoint change. In an actual industrial process system, there are, among other things, grade changes, adjustments in production and seasonal variation. Seasonal variation is considered in the simulation, as presented in Fig. 8. The

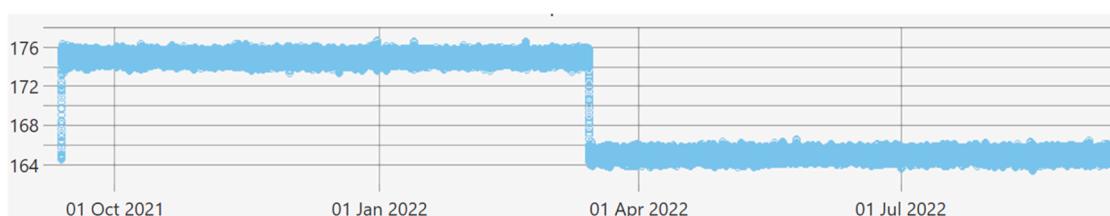


Fig. 7. Change in the final temperature setpoint (March 14th) – stream 12.

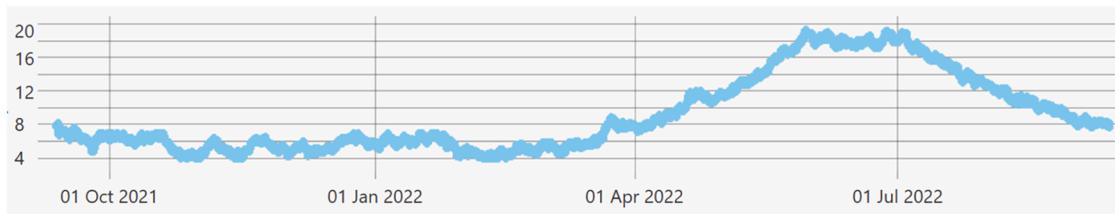


Fig. 8. Temperature on stream 1 exhibit seasonal variation.

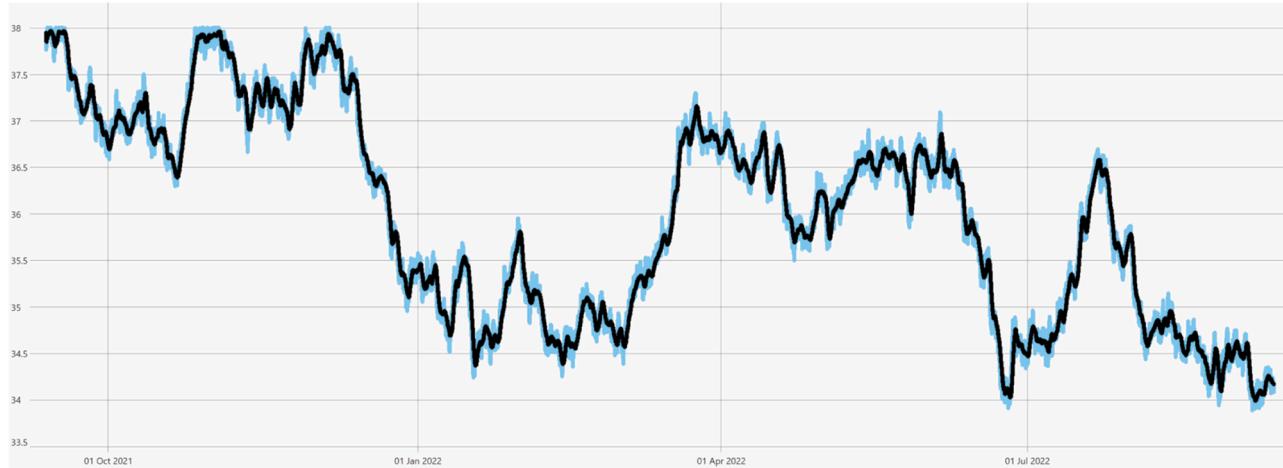


Fig. 9. Raw flow measurement signal (in blue) and the signal of interest (in black) – stream 15.

temperature of stream 1 is increasing during summer and goes back around 8 °C afterwards.

Lastly, Fig. 9 presents an example of a raw flow measurement signal in blue and the useful signal (signal of interest) in black. The signal of interest aims to represent the signal trend as accurately as possible, reducing the effects of measurement and process errors, whereas the raw signal is the one containing noise.

5.2. Denoising performance evaluation – synthetic signals from a heat exchanger network simulation

In this section, the filtering performance of four signal processing methods (Kalman Filter, EWMA Filter, STFT and Wavelet Transform) are assessed and compared using the root mean square error and the signal-to-noise ratio, two easily interpretable criteria and the most used performance criteria in signal processing (Ben Slama et al., 2018; Muppalla et al., 2017). These denoising performance metrics are evaluated systematically before and after signal processing for each method, and the difference is analyzed.

RMSE is the average distance (in absolute value), measured along a vertical line, between every pair of data points in two signals. It measures how close each processed signal is to its respective noiseless version, i.e., the quality of the filtered signal. The values are always non-negative, they have the same units as the signal (they are directly interpretable in terms of measurement units), and the closer they are to zero, the better, i.e., the smaller the RMSE, the more similar the processed signal is to the original noiseless data. The objective is to minimize the error between the processed signal and the original signal. The deviations between the real (original) signals and the measured noisy ones, as well as between the real signals the processed ones, are given respectively by the following equations:

$$RMSE_{before} = \sqrt{\frac{1}{N} \sum_{i=1}^N [x(i) - \bar{x}(i)]^2}; \quad RMSE_{result} = \sqrt{\frac{1}{N} \sum_{i=1}^N [x(i) - \hat{x}(i)]^2} \quad (20)$$

where $\bar{x}(i)$ is the synthetic noisy signal, $\hat{x}(i)$ is the denoised signal, $x(i)$ is the original (reference) signal, and N is the number of samples in the signal.

This criterion is used to assess the effectiveness of a processing method by grading the difference between $RMSE_{before}$ and $RMSE_{result}$. The former should be larger than the latter. Then this difference is divided by $RMSE_{before}$ to evaluate the $RMSE_{reduction}$. The larger the value, the greater is the reduction, and the more effective is the algorithm. A 100% reduction would mean that there is no more error and that the filtered signal is identical to the original one:

$$RMSE_{difference} = RMSE_{before} - RMSE_{result} \quad (21)$$

$$RMSE_{reduction} [\%] = \frac{RMSE_{difference}}{RMSE_{before}} \times 100 \quad (22)$$

The noise reduction efficiency is evaluated using the signal-to-noise ratio. SNR is defined as the ratio of signal power to noise power, often expressed in decibels, as presented below:

$$SNR_{db} = 10 \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right) \quad (23)$$

SNR compares the magnitude of a desired signal to the level of background noise. This characteristic reflects the dynamic of the signal. A ratio greater than 0 dB indicates more signal than noise, so the higher the SNR, i.e., the lower the noise content, the better is the signal quality.

The SNR before denoising is defined as

$$SNR_{before} (dB) = 10 \log_{10} \frac{\sum_{i=1}^N \bar{x}^2(i)}{\sum_{i=1}^N [\bar{x}(i) - x(i)]^2} \quad (24)$$

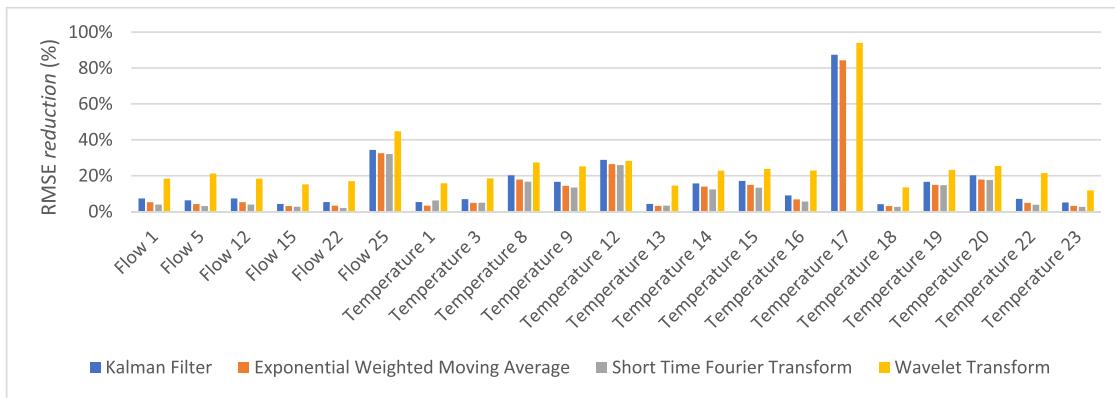


Fig. 10. RMSE reduction for synthetic signals for KF, EWMA, STFT and WT.

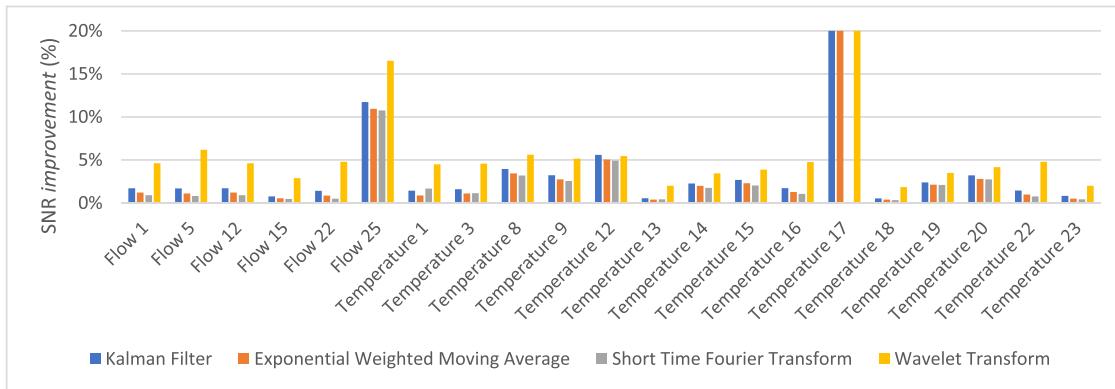


Fig. 11. SNR improvement for synthetic signals for KF, STFT, WT and EWMA.

The SNR after denoising is defined as

$$SNR_{result}(dB) = 10 \log_{10} \frac{\sum_{i=1}^N \hat{x}^2(i)}{\sum_{i=1}^N [\hat{x}(i) - x(i)]^2} \quad (25)$$

SNR_{before} should be less than SNR_{result} to define the effectiveness of denoising techniques. $SNR_{improvement}$ is used to assess the noise reduction efficiency. The higher the value, the better is the result.

$$SNR_{difference}(dB) = SNR_{results}(dB) - SNR_{before}(dB) \quad (26)$$

$$SNR_{improvement}[\%] = \frac{SNR_{difference}}{SNR_{before}} \times 100 \quad (27)$$

The results of denoising by Kalman filter, EWMA filter, STFT, and WT are compared by giving the RMSE and SNR parameters. The ideal signal processing technique would be the one with the highest value for both

criteria, i.e., showing the greatest SNR improvement and the greatest RMSE reduction. The assessment of these metrics is only viable and can be made in a rigorous way because of synthetic signals and simulated noise; therefore it is possible to know exactly the underlying signal, i.e., the signal of interest. Figs. 10 and 11 present the RMSE and SNR results, denoted respectively as $RMSE_{reduction}$ and $SNR_{improvement}$ for all four filtering methods for all temperature and flow signals.

Figs. 10, and 11 present respectively the comparison of different signal processing techniques in terms of RMSE reduction and SNR improvement in percent when processing synthetics temperature and flow signals from a HEN simulation. WT approach dominates in term of RMSE reduction, it is the best-performing method for most of the signals, except for temperature 12 where WT is almost as performant as the other techniques (28.19% for WT compared to 28.83% for KF). This signal presented in Fig. 7 show the change in the final temperature setpoint. It seems that WT approach is not as performant as the KF for setpoint variation in signals. In regard to the SNR improvement, WT is the preferred approach for all signals, except temperature 12 where the SNR improvement is 5.60% for KF and 5.45% for WT. STFT appears to be the least efficient techniques towards RMSE reduction as well as SNR improvement for most signals. An analysis is proposed in the following section.

RMSE reduction varied over a wide range, from null values for STFT to almost a 100% for WT. This means that signal processing technique performance varied widely, and the same is true for SNR improvement. Temperature 17 signal shows high values of SNR improvement in Fig. 11 for KF (37.50%), EWMA (33.58%) and WT (50.95%), however there is no improvement when the STFT is considered. This approach was not able to reduce noise from a signal that present rapid variations of low amplitude. Finally, both criteria led to the same conclusions: WT is the most efficient for Gaussian noise reduction for simulated HEN signals

Table 3
RMSE and SNR averages with their standard deviations for all approaches.

	Sensors	KF	EWMA	STFT	WT
Average RMSE	Flow	10.80% ± 11.60%	8.91% ± 11.58%	7.93% ± 11.84%	22.46% ± 11.09%
	Temperature	17.62% ± 20.66%	15.55% ± 20.32%	9.52% ± 7.33%	25.88% ± 19.52%
	All	15.67% ± 18.51%	13.66% ± 18.22%	9.07% ± 8.56%	24.91% ± 17.32%
Average SNR	Flow	3.17% ± 4.21%	2.65% ± 4.07%	2.40% ± 4.10%	6.60% ± 4.97%
	Temperature	4.60% ± 9.20%	3.97% ± 8.29%	1.68% ± 1.32%	7.11% ± 12.19%
	All	4.19% ± 8.01%	3.60% ± 7.26%	1.88% ± 2.35%	6.97% ± 10.50%

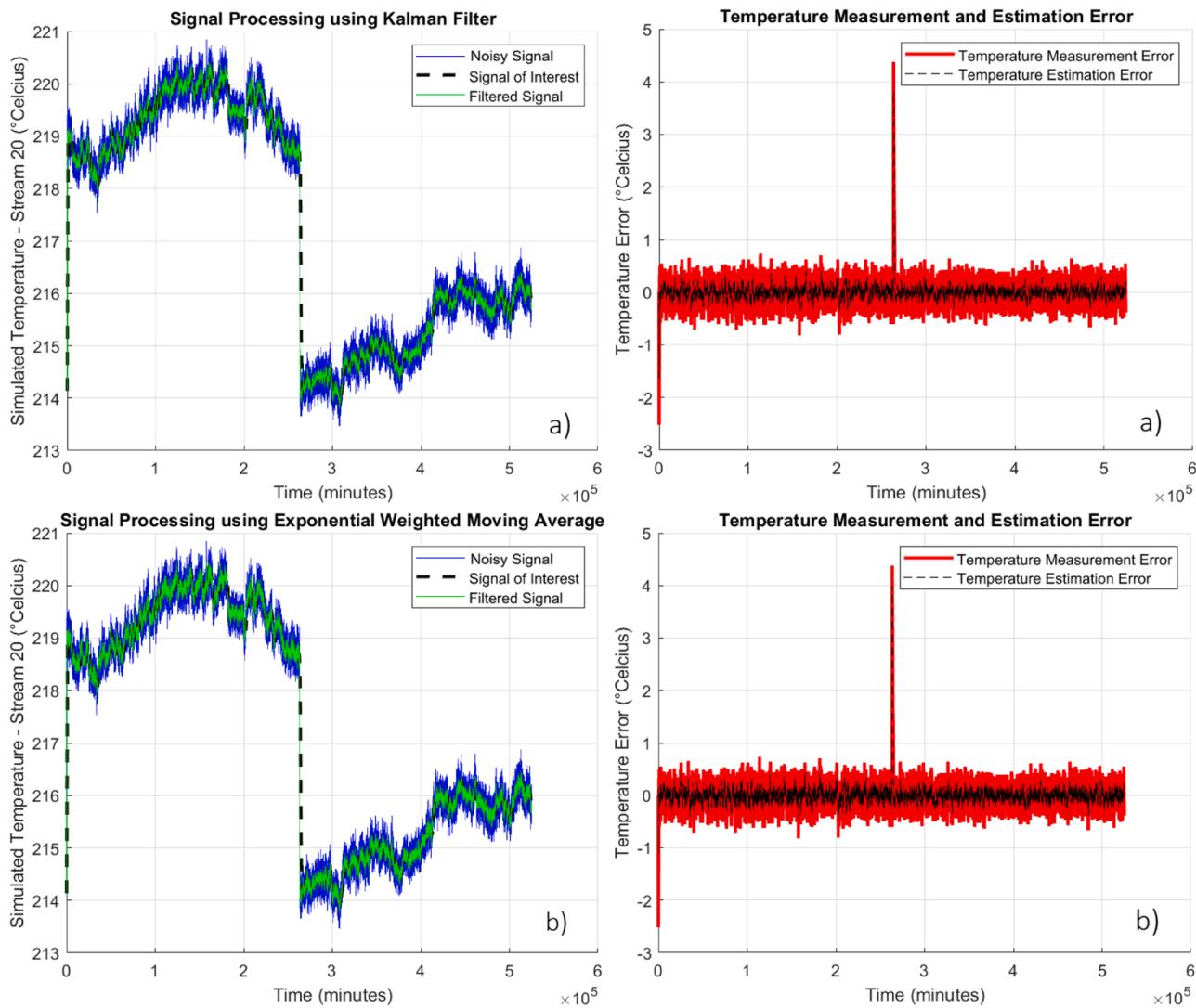


Fig. 12. Denoising results of synthetic signal temperature 20 using a) KF, b) EWMA, c) STFT, and d) WT.

when compared with KF, EWMA, and STFT. Denoising with WT secured the lowest RMSE and the highest SNR most of the time, hence the largest difference from the original signal.

All processing techniques presented higher RMSE reduction than SNR improvement values, which means that they all reduced the difference between the original signals and those of interest, but the signals are still noisy. It is not possible to perfectly reconstruct a noise-free signal, but a proper filter will give the trend in the data rather than its instantaneous variation. Therefore, in this case, the filtered signals must resemble the signal of interest.

The two criteria are always in agreement; a signal that has a high value in RMSE reduction also has a high value in SNR improvement. WT showed its effectiveness by performing almost equally well for both metrics (RMSE and SNR) regardless of the signal; the RMSE reduction is about 25% and SNR improvement around 7%. Table 3 represents averaged values of RMSE reduction and SNR improvement, along with their standard deviations, for all signals (flows, temperatures, and both). The results unambiguously show that the WT are globally significantly better than other approaches according to the RMSE reduction and SNR improvement; WT exhibits better denoising performance. In fact, WT shows on average 9% more RMSE reduction and 3% more SNR improvement when compared to KF, EWMA and STFT when considering all signals.

Therefore, signals obtained from systems exhibiting complex dynamics and non-linearity present significant and predominant multi-scale features, such as signal from a HEN in a plant. WT is found to be the most effective method for achieving the best denoising performance for these signals, as illustrated in the scenario presented in this work.

Fig. 12 exhibits, on the left side, the original TI20 signal (blue), its signal of interest (black), and the filtered signal (green) for all methods and on the right side, both the measurement error (red) and the estimation error (black) of the respective method. The former evaluates the difference or distance between the original signal and the signal of interest, and the latter is the difference between the original and filtered signal. As can be observed, the signal filtered using wavelet transform is able to follow more closely the trend of the underlying signal of interest compared to other approaches.

Fig. 12a) shows a greatly decreased level of noise and a smaller error than what is originally in the signal. The KF presents a considerable improvement – a RMSE reduction of 20.17% and a SNR improvement of 3.21%. Even if the results presented in Fig. 12b) are quite similar as those obtained with KF, the denoising is not as good as the one obtained using KF – EWMA presents a RMSE reduction of 17.82% and a SNR improvement of 2.80%. Fig. 12c) presents a poor execution of filtering with STFT; there is almost no difference from the error before filtering (RMSE reduction and SNR improvement are respectively at 17.58% and

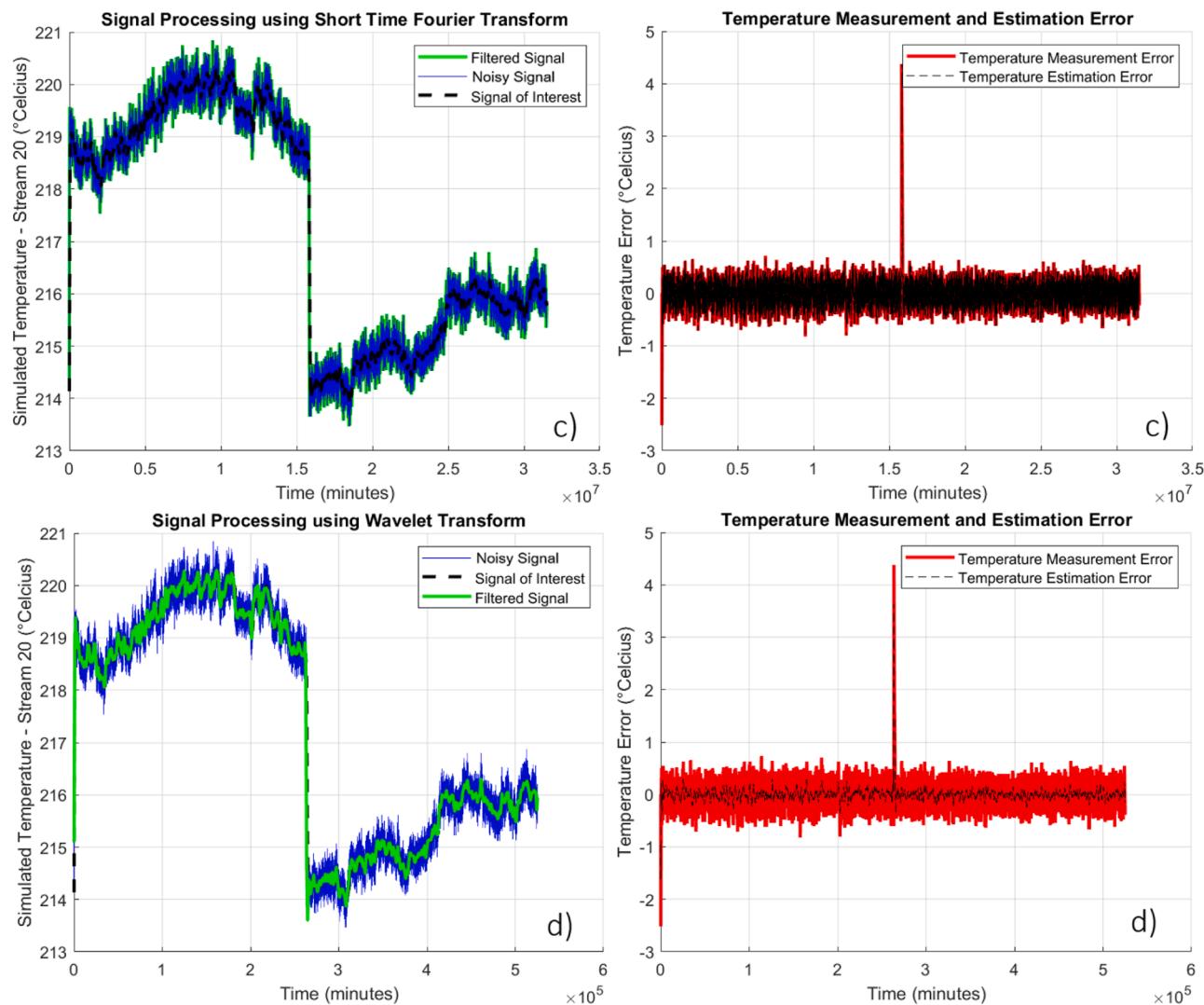


Fig. 12. (continued).

2.75%. In Fig. 12d), almost no noise remains; the signal of interest is practically superimposed with the filtered signal. WT offers the lowest percentage of error relative to the signal of interest: RMSE reduction of 25.39% and a SNR improvement of 4.17%. The improvement of the estimation error is clearly visible and differs from all the other approaches.

6. Discussion

Following the comparison analysis of four processing technique, the identification of optimal parameters – those that could minimize the RMSE – is presented in Table 4.

For the KF, the ratio to obtain the optimal process noise value is identical for all experiments; it is R_t divided by 1000. The algorithm selected the smallest value as the optimal one. Moreover, a high process noise value implies a higher state uncertainty (Cortés et al., 2022). However, on the contrary, a lower value suggest a higher confidence on the process or state model, i.e., both are well known (Cortés et al., 2022). Nevertheless, the measurement noise values also have physical meaning as they represent the variance of the measurement.

With the short-time Fourier transform, when comparing various overlap lengths, FFT lengths and window type, the results showed that the optimal values differ from one signal to another. In general, more overlap – smaller hop size – will give more analysis points and hence

smoother results over time, but the computational expense is proportionately greater. Moreover, to achieve a smooth signal, the frequency domain information does not need to be very specific. Hence, when looking for a window length that results in the loss of spectral resolution, a narrow window is preferred. However, a longer window enables better trend identification. Therefore, for the window size (FFT length), a compromise must be found between the time and the frequency resolution. Lastly, the original signal standard deviation multiplicator threshold is always at its higher, and most of the time, at its highest value. The filter is trying to truncate noise as much as possible.

For the wavelet transforms, different mother wavelet functions are tested. Because orthogonal wavelets such as the Coiflet, Daubechies and Symlet families are the best for denoising signals, these wavelet functions are tested for various orders (coif1-coif5, sym2 to 15 and db2 to db15). The biorthogonal wavelet function (bior1.1-bior6.8) is also tested. Decomposition levels (scale) from 1 to 10 are considered. The latter must not be too low, in which case high frequency components might hide the trend, but also, it must not over-smooth the signal (hence, not too high of a scale). The optimal values for the wavelet function and the scale are obtained for all signal by minimizing the RMSE. It is found that Daubechies and Symlet are the two selected function and decomposition level 10 is chosen for all signals. This value is coherent with the fact that the signals are corrupted with high frequency components that needed to be reduced. Therefore, the highest scale is required to reach

Table 4

Method-specific parameters values for KF, STFT, WT, and EWMA filter for all signals.

Signals	Signal processing approaches optimal parameters										
	Kalman filter		Short time Fourier transform				Wavelet transform			EWMA	
	Measurement noise R_t^*	Process noise Q_t	Window type	FFT length	Overlap length (% of the window size)	Threshold value	Wavelet Function	Decomp. level	Threshold selection rule		
F 1	1.69E-01	1.69E-04	rectwin	Week	10%	8	db4	10	UT	Soft	0.05
F 5	1.19E-01	1.19E-04	chebwin	Day	10%	9	db4	10	UT	Soft	0.05
F 12	1.69E-01	1.69E-04	rectwin	Week	10%	8	db4	10	UT	Soft	0.05
F 15	5.55E-03	5.55E-06	chebwin	Month	25%	10	sym2	10	MM	Soft	0.05
F 22	3.98E-01	3.98E-04	chebwin	Week	10%	10	db4	10	MM	Soft	0.05
F 25	2.88E-01	2.88E-04	gausswin	Day	10%	10	db4	10	UT	Soft	0.05
T 1	1.98E-02	1.98E-05	hamming	Hour	1%	10	sym2	10	UT	Soft	0.05
T 3	3.77E-02	3.77E-05	chebwin	Week	25%	10	sym3	10	MM	Soft	0.05
T 8	6.78E-02	6.78E-05	chebwin	Day	10%	10	db12	10	UT	Soft	0.05
T 9	9.81E-02	9.81E-05	triang	Week	25%	10	db13	10	UT	Soft	0.05
T 12	1.03E-01	1.03E-04	rectwin	Hour	75%	10	db12	10	UT	Soft	0.05
T 13	4.78E-03	4.78E-06	chebwin	Week	1%	10	sym3	10	MM	Soft	0.05
T 14	9.83E-03	9.83E-06	chebwin	Day	1%	10	db12	10	UT	Soft	0.05
T 15	2.11E-02	2.11E-05	rectwin	Hour	75%	10	db12	10	UT	Soft	0.05
T 16	7.36E-02	7.36E-05	hamming	Month	75%	10	db4	10	UT	Soft	0.05
T 17	3.53E-02	3.53E-05	hamming	Day	10%	10	db7	10	MM	Soft	0.05
T 18	5.37E-03	5.37E-06	gausswin	Hour	25%	10	sym3	10	MM	Soft	0.05
T 19	9.35E-03	9.35E-06	hamming	Hour	75%	10	db12	10	UT	Soft	0.05
T 20	2.37E-02	2.37E-05	rectwin	Hour	75%	10	db12	10	UT	Soft	0.05
T 22	2.88E-01	2.88E-04	gausswin	Day	10%	10	sym3	10	UT	Soft	0.05
T 23	5.91E-03	5.91E-06	chebwin	Hour	1%	10	sym2	10	MM	Soft	0.05

F = Flow, T = Temperature. * °C for T sensors and ton/h for Flow sensors. UT = UniversalThreshold, MM = Minimax.

the signal of interest. The optimal thresholding selection rule is either minimax or universal threshold. The former is more conservative and is more convenient when small details of the signal lie near the noise range whereas the latter is known to reduce the noise more effectively (MathWorks, 2020). Finally, the scales highly affected by noise are processed by soft thresholding. This thresholding function led to the best results for all signals. This may be explained by the fact that hard thresholding results in larger variance after reconstruction and can very well represent peaks in the signal, but since the signal's trend is what matter here (signal of interest), this is not applicable.

Lastly, for the exponential weighted moving average, the smoothing factor is all the time at its smallest value (0.05). A value near zero means that the filter is less responsive to new incoming measurements, hence there is a greater smoothing effect, more weight is put on the average.

The success of a signal processing method depends on the optimal configuration of its control parameters. Working with synthetic simulated signals with known errors makes possible the identification of optimal parameters values. However, in a real-life context, the selection of these parameters is much more complex. In the case of KF, setting the state equations requires a good knowledge of the process, and this is hard because processes are complex. Additionally, the process noise is unknown, and the measurement noise is hard to determine, it varies according to sensor sensitivity and accuracy. Therefore, for a real process, obtaining optimal result would be much harder because the process is more complex and assessing the state equation and both matrices is tedious. In the case of EWMA, the value of the smoothing parameter is typically correlated with measurement precision or accuracy. This parameter is not set according to the signal's appearance or characteristics but rather on the confidence in data.

However, because measurement signals are generally characterized by noise at several frequencies, these filters force a compromise between the quality of improvement over a long period (signal distortion) and of local signal characteristics. Linear filters are therefore not effective in filtering signals containing several characteristics located at different positions in the time and/or frequency domains (Bakshi and Stephanopoulos, 1994; Nounou and Bakshi, 1999). Problems with linear filter efficiency have led to the development of multi-stage filtration methods,

such as WT. Moreover, WT makes it possible to shift the wavelet function to align it with events to reduce noise efficiently.

Wavelet transforms parameters are selected according to the signal's characteristics. The mother wavelet function must match the shape of the signal, the scale and thresholding parameters must be chosen in order to retain only the signal's trend or important component and reduce high frequency elements. It could be interesting to assess the relation between signals characteristics and the WT parameters used to process them in specific applications (as signals are always processed according to a targeted application).

STFT parameters are also selected according to the frequency content of the signal, however this approach did not perform very well on either criteria, there is less error in the signals, but noise is still very much present and Figs. 10–12c corroborates that. In fact, STFT is usually reported as a good approach for non-stationary signals analysis purposes, but, for filtering HEN flow and temperature signals, this method is not recommended in face of the results obtained. Single-scale methods such as the Kalman Filter tend to present superior performances in this context, even if the underlying signal of interest presents multiscale dynamics.

7. Conclusion

This article proposes a comparison of four signal processing approaches. Synthetic signals containing random errors are created and processed one at the time with KF, EWMA, STFT, and WT. Because the errors are known, it is possible to assess their optimal parameters and compare the methods to find the best-performing one. The performance of these denoising algorithms is evaluated using two criteria: RMSE and SNR.

As a result, single scale filtering methods (EWMA and KF) are less efficient for simultaneously filter non-stationary measured signals containing several characteristics/multiple features located at different positions in time-frequency domain and accurately represent these characteristics and features, compared to the Wavelet transform. Additionally, short time Fourier transform is found to be the most inadequate approach to reduce error in heat exchanger network simulated signals.

Compared to other signal processing approaches, WT is found to be performant towards noise reduction for industrial process signals from a HEN simulation and adequate for improving their SNR and reducing the RMSE. In fact, WT enabled on average 9% more RMSE reduction and 3% more SNR improvement than KF, EWMA and STFT when considering all signals. It is a robust signal processing technique that can remove errors from signals while retaining useful information; it retains the variability of the original signal, which reflects the true changes and characteristics of the process. WT stand out for its denoising performance for simulated HEN industrial signals, and for the way its parameters (mother function, scale, threshold, etc.) are selected. In this paper, the HEN simulated normal continuous operations only, however, tube leaks, fouling, sudden variation of the split ratio, heat exchanger by-pass could be considered. Further research could be conducted on the 4 approaches by adding different types of faults in the signals.

Furthermore, this work used heat exchanger network simulation signals as example of industrial signals. It would be interesting to replicate the methodology employed in this paper to other cases study as the conclusions are limited to this case study, they are case-specific. The results stating that WT stand out for its denoising performance compared to EWMA, KF and STFT cannot be generalized for all industrial process signals. However, the results represent a relevant contribution to the growing body of knowledge in the field of industrial process signal noise reduction. In order to draw conclusions on how generalizable the results are and hence to prove that WT is the most performant approach for industrial signal denoising, more industrial simulations are required. The methodology could be extrapolated to signals coming from other industrial process sectors that have similar behavior, e.g., boilers, evaporators, steam generators. Since WT were robust for HEN signals, one might expect similar results with these departments. Moreover, the literature (Ganji et al., 2021; Guney et al., 2019; Moosavi et al., 2018; Rendall and Reis, 2014; Sraith and Jabrane, 2021; Subramanian et al., 2013; Zych et al., 2018) showed that WT is adequate for many types of signals, especially signals that have characteristics that cover various frequency ranges. Considering that WT involves a sophisticated implementation considering its tuning parameters, the idea of assessing the relation between the signal's characteristics and the signal processing technique's parameters to observe how beneficial it could be for processing a wide range of signals in a systematic and efficient way could be further addressed in a following paper.

Rectified data will serve as the basis for many subsequent decisions and enable better control of plant operations and decision-making. Signal processing increases the efficiency of tasks related to the operation of chemical processes: configuration of control systems, detection and diagnosis of faults directly related to process operations (leaks or deposits), and planning and scheduling of process operations and maintenance. Because process data contain errors that may be detrimental, it is critical to remove them before making any important decisions.

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 gratefully acknowledge the financial support of the CRSNG and further thank CanmetENERGY for the technical support

using the EXPLORE software.

References

- Abbaszadeh, P., Alipour, A., 2017. Development of a coupled wavelet transform and evolutionary Levenberg-Marquardt neural networks for hydrological process modeling. *Comput. Intell.* 34, 1–25. <https://doi.org/10.1111/coin.12124>.
- Alyasser, Z.A.A., Khader, A.T., Al-Betar, M.A., Abasi, A.K., Makhadmeh, S.N., 2020. EEG signals denoising using optimal wavelet transform hybridized with efficient metaheuristic methods. *IEEE Access* 8, 10584–10605. <https://doi.org/10.1109/ACCESS.2019.2962658>.
- Bagajewicz, M., 2000. A brief review of recent developments in data reconciliation and gross error detection/estimation. *Latin Am. Appl. Res.* 30 (4), 335–342.
- Bahoura, M., 2019. Efficient FPGA-Based Architecture of the Overlap-Add Method for Short-Time Fourier Analysis/Synthesis 8 (12), 1533. <https://doi.org/10.3390/electronics8121533>.
- Bakshi, B.R., 1999. Multiscale analysis and modeling using wavelets. *J. Chemom.* 13 (3–4), 415–434. [https://doi.org/10.1002/\(SICI\)1099-128X\(199905/08\)13:3/43.0.CO;2-8](https://doi.org/10.1002/(SICI)1099-128X(199905/08)13:3/43.0.CO;2-8).
- Bakshi, B.R., Stephanopoulos, G., 1994. Representation of process trends—III. Multiscale extraction of trends from process data. *Comput. Chem. Eng.* 18 (4), 267–302. [https://doi.org/10.1016/0098-1354\(94\)85028-3](https://doi.org/10.1016/0098-1354(94)85028-3).
- Belitsky, E., 2001. Maintenance by Numbers. *Plant Engineering and Maintenance Magazine*. November.
- Ben Slama, A., Lentka, L., Mouelhi, A., Diouani, M.F., Sayadi, M., Smulko, J., 2018. Application of statistical features and multilayer neural network to automatic diagnosis of arrhythmia by ECG signals. In: *Metrology and Measurement Systems*, 25. <https://doi.org/10.24425/118163>.
- Bergman, T.L., Lavine, A., Incropera, F.P., DeWitt, D.P., 2011. *Introduction to Heat Transfer*. Wiley.
- Bhat, S.A., Saraf, D.N., 2004. Steady-State Identification, Gross Error Detection, and Data Reconciliation for Industrial Process Units. *Ind. Eng. Chem. Res.* 43 (15), 4323–4336. <https://doi.org/10.1021/ie030563u>.
- Bolaers, F., Olivier, C., Estocq, P., Chiementin, X., Dron, J.P., 2011. Comparison of denoising methods for the early detection of fatigue bearing defects by vibratory analysis. *J. Vibrat. Control* 17, 1983–1993. <https://doi.org/10.1177/1077546309348853>.
- Cao, Z., Jia, H., Zhao, T., Fu, Y., Wang, Z., 2022. An adaptive self-organizing migration algorithm for parameter optimization of wavelet transformation. *Math. Prob. Eng.* 2022, 6289215. <https://doi.org/10.1155/2022/6289215>.
- Cortés, I., van der Merwe, J.R., Lohan, E.S., Nurmi, J., Felber, W., 2022. Performance Evaluation of Adaptive Tracking Techniques with Direct-State Kalman Filter. *Sensors (Basel)* 22, 22. <https://doi.org/10.3390/s22020420>.
- Crowe, C.M., 1996. Data reconciliation — Progress and challenges. *J. Process Control* 6 (2), 89–98. [https://doi.org/10.1016/0959-1524\(96\)00012-1](https://doi.org/10.1016/0959-1524(96)00012-1).
- Dash, P.K., Chilukuri, M.V., 2004. Hybrid S-transform and Kalman filtering approach for detection and measurement of short duration disturbances in power networks. *IEEE Trans. Instrum. Meas.* 53 (2), 588–596. <https://doi.org/10.1109/TIM.2003.820486>.
- Daubechies, I. (1992a). The What, Why, and How of Wavelets. Dans *Ten Lectures on Wavelets* (p. 1–16).
- Daubechies, I., 1992b. 3. Discrete Wavelet Transforms: Frames. Dans *Ten Lectures on Wavelets* 53–105. <https://doi.org/10.1137/1.9781611970104.ch3>.
- Diatkine, C., 2011. Signal Analysis Introduction - FFT Parameters. Support IRCAM. <http://support.ircam.fr/docs/AudioSculpt/3.0/co/FFT%20Params.html>.
- Dong, Y., Li, Y., Xiao, M., Lai, M., 2009. Unscented Kalman filter for time varying spectral analysis of earthquake ground motions. *Appl. Math. Model.* 33 (1), 398–412. <https://doi.org/10.1016/j.apm.2007.11.020>.
- Doymaz, F., Bakhtazad, A., Romagnoli, J.A., Palazoglu, A., 2001. Wavelet-based robust filtering of process data. *Comput. Chem. Eng.* 25 (11), 1549–1559. [https://doi.org/10.1016/S0098-1354\(01\)00718-9](https://doi.org/10.1016/S0098-1354(01)00718-9).
- Dyason, W., Van Niekerk, T.I., Phillips, R., & Stopforth, R. (2017). *Performance evaluation and comparison of filters for real time embedded system applications*. 28th Annual Symposium of the Pattern Recognition Association of South Africa and the 10th Robotics and Mechatronics International Conference, PRASA-RobMech 2017, November 29, 2017 - December 1, 2017, Bloemfontein, South Africa (vol. 2018-January, p. 242–248). [10.1109/RoboMech.2017.8261155](https://doi.org/10.1109/RoboMech.2017.8261155).
- Ebrahimpour, E., Pooyan, M., Jahani, S., Bijar, A., Setarehdan, S., 2015. ECG signals noise removal: selection and optimization of the best adaptive filtering algorithm based on various algorithms comparison. *Biomed. Eng. Appl. Basis Commun.* 27, 1–13. <https://doi.org/10.4015/S1016237215500386>.
- El-Dahshan, E.-S.A., 2011. Genetic algorithm and wavelet hybrid scheme for ECG signal denoising. *Telecommun. Syst.* 46 (3), 209–215. <https://doi.org/10.1007/s11235-010-9286-2>.
- Evans, W.C., 2019. Exponentially-weighted moving average. GeoGebra. <https://www.geogebra.org/m/th88mqrm>.
- Everett, J. (2011). *The exponentially weighted moving average applied to the control and monitoring of varying sample sizes* (vol. 51).
- Fugal, L.D., 2009. *Conceptual Wavelets in Digital Signal Processing*, 1e éd. Space & Signals Technical Publishing.
- Ganesan, R., Das, T.K., Venkataraman, V., 2004. Wavelet-based multiscale statistical process monitoring: a literature review. *IIE Trans.* 36 (9), 787–806. <https://doi.org/10.1080/07408170490473060>.
- Ganji, M.R., Ghelmani, A., Golroo, A., Sheikhzadeh, H., 2021. A brief review on the application of sound in pavement monitoring and comparison of tire/road noise

- processing methods for pavement macrotexture assessment. *Arch. Comput. Methods Eng.* 28 (4), 2977–3000. <https://doi.org/10.1007/s11831-020-09484-4>.
- Guney, G., Uluc, N., Demirkiran, A., Aytac-Kipergil, E., Unlu, M.B., Birgul, O., 2019. Comparison of noise reduction methods in photoacoustic microscopy. *Comput. Biol. Med.* 109, 333–341. <https://doi.org/10.1016/j.combiomed.2019.04.035>.
- Harkat, S., Boukharouba, K., Douaoui, A.E.K., 2016. Multi-site modeling and prediction of annual and monthly precipitation in the watershed of Cheliff (Algeria). *Desalination Water Treat.* 57 (13), 5959–5970. <https://doi.org/10.1080/19443994.2014.956798>.
- Harrison, R.P., & Stuart, P.R. (2011). *Data pre-processing techniques for multivariate analysis to treat industrial operating data for retrofit design*.
- Hasan, A.M., Samsudin, K., Ramli, A.R., 2011. Intelligently tuned wavelet parameters for GPS/INS error estimation. *Int. J. Autom. Comput.* 8 (4), 411–420. <https://doi.org/10.1007/s11633-011-0598-9>.
- Huang, Y., Benesty, J., Chen, J., 2008. Analysis and comparison of multichannel noise reduction methods in a common framework. *IEEE Trans. Audio Speech Lang. Process.* 16 (5), 957–968. <https://doi.org/10.1109/TASL.2008.921754>.
- Ibrahim, R.K., Watson, S.J., Djurović, S., Crabtree, C.J., 2018. An effective approach for rotor electrical asymmetry detection in wind turbine DFIGs. *IEEE Trans. Ind. Electr.* 65 (11), 8872–8881. <https://doi.org/10.1109/TIE.2018.2811373>.
- Jiang, T., Chen, B., He, X., Stuart, P., 2003. Application of steady-state detection method based on wavelet transform. *Comput. Chem. Eng.* 27 (4), 569–578. [https://doi.org/10.1016/S0098-1354\(02\)00235-1](https://doi.org/10.1016/S0098-1354(02)00235-1).
- Katunin, A., Przystalka, P., 2015. Automated wavelet-based damage identification in sandwich structures using modal curvatures. *J. Vibroeng.* 17, 2977–2986.
- Kehtarnavaz, N. (2008). Frequency Domain Processing. Dans N. Kehtarnavaz (éd.), *Digital Signal Processing System Design* (2e éd., p. 175–196). Academic Press.
- Kim, Y., Bang, H., 2018. Introduction to Kalman Filter and Its Applications. Dans *Kalman Filter*. InTechOpen.
- Korbel, M., Bellec, S., Jiang, T., Stuart, P., 2014. Steady state identification for on-line data reconciliation based on wavelet transform and filtering. *Comput. Chem. Eng.* 63, 206–218. <https://doi.org/10.1016/j.compchemeng.2014.02.003>.
- Krishna, B., Yadav, G., 2016. Performance comparison of different variable filters for noise cancellation in real-time environment. *Int. J. Signal Process. Image Process. Pattern Recognit.* 9, 107–126. <https://doi.org/10.14257/ij sip.2016.9.2.10>.
- Labbe Jr, R.R. (2020). *Kalman and Bayesian Filters in Python*.
- Liu, W., Chang, Q.R., Guo, M., Xing, D.X., Yuan, Y.S., 2011. Extraction of first derivative spectrum features of soil organic matter via wavelet de-noising. *Guang Pu Xue Yu Guang Pu Fen Xi* 31 (1), 100–104.
- Ma'arif, A., Iswanto, Nuryono, A., Alfiyan, R., 2020. Kalman Filter for noise reducer on sensor readings. *Signal Image Process. Lett.* 1, 11–22. <https://doi.org/10.31763/simple.v1i2.2>.
- MacGregor, J.F., Kozub, D.J., Penlidis, A., Hamielec, A.E., 1986. State estimation for polymerization reactors. *IFAC Proc. Volumes* 19 (15), 147–152. [https://doi.org/10.1016/S1474-6670\(17\)59414-8](https://doi.org/10.1016/S1474-6670(17)59414-8).
- Mah, R.S.H., Tamhane, A.C., 1982. Detection of gross errors in process data. *AIChE Journal* 28 (5), 828–830. <https://doi.org/10.1002/aic.690280519>.
- Mallat, S.G., 2009. *A Wavelet Tour of Signal processing: the Sparse way*, 3e éd. Elsevier /Academic Press.
- Manhertz, G., Modok, D., Á, B, 2016. Evaluation of short-time fourier-transformation spectrograms derived from the vibration measurement of internal-combustion engines. In: 2016 IEEE International Power Electronics and Motion Control Conference (PEMC), pp. 812–817. <https://doi.org/10.1109/EPEPEMC.2016.7752098>.
- Martini, A., Coco, D., Sorice, A., Traverso, A., Levorato, P., 2014. Gross Error Detection Based on Serial Elimination: Applications to an Industrial Gas Turbine, Volume 3A. Coal, Biomass and Alternative Fuels; Cycle Innovations; Electric Power; Industrial and Cogeneration. <https://doi.org/10.1115/gt2014-26746>.
- MathWorks, 2020. Threshold Selection For Denoising. MathWorks. <https://www.mathworks.com/help/wavelet/ref/thselect.html>.
- MathWorks, 2023. Short-Time Fourier Transform. MathWorks. <https://www.mathworks.com/help/dsp/ref/dsp.stft.html>.
- Mitra, S.K.K., 2000. *Digital Signal Processing: A Computer-Based Approach*. McGraw-Hill Higher Education.
- Moosavi, S.R., Qajar, J., Riazi, M., 2018. A comparison of methods for denoising of well test pressure data. *J. Petrol. Explor. Prod. Technol.* 8 (4), 1519–1534. <https://doi.org/10.1007/s13202-017-0427-y>.
- Muppalla, V., Suraj, N.S.S.K., Reddy, V.Y.S., Suman, D., 2017. Performance evaluation of different denoising techniques for physiological signals. In: 2017 14th IEEE India Council International Conference (INDICON), pp. 1–6. <https://doi.org/10.1109/INDICON.2017.8487739>.
- Narasimhan, S., Jordache, C., 2000. *Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data*. Gulf Professional Publishing.
- Narayan, K., Mahesh, B., Andreas, S., 2013. *An Introduction to Kalman Filtering With MATLAB Examples*. Morgan & Claypool.
- Nounou, M.N., Bakshi, B.R., 1999. On-line multiscale filtering of random and gross errors without process models. *AIChE J.* 45 (5), 1041–1058. <https://doi.org/10.1002/ai ch.690450513>.
- Nounou, M.N., & Bakshi, B.R. (2000). Chapter 5 - Multiscale methods for denoising and compression. Dans B. Walczak (éd.), *Data Handling in Science and Technology* (vol. 22, p. 119–150). Elsevier.
- Park, S.M., Lee, T.J., Sim, K.B., 2016. Heuristic feature extraction method for BCI with harmony search and discrete wavelet transform. *Int. J. Control. Autom. Syst.* 14 (6), 1582–1587. <https://doi.org/10.1007/s12555-016-0031-9>.
- Park, S., Gil, M.S., Im, H., Moon, Y.S., 2019. Measurement noise recommendation for efficient Kalman filtering over a large amount of sensor data. *Sensors (Basel)* 19 (5), 1168. <https://doi.org/10.3390/s19051168>.
- Perez, E., Barros, J., 2006. Voltage event detection and characterization methods: a comparative study. In: 2006 IEEE/PES Transmission & Distribution Conference and Exposition: Latin America, pp. 1–6. <https://doi.org/10.1109/TDCLA.2006.311552>.
- Pérez, E., Barros, J., 2012. Application of advanced digital signal processing tools for analysis of voltage events in power systems. *Int. J. Electr. Eng. Educ.* 46 (3), 211–224. <https://doi.org/10.7227/ijeee.46.3.1>.
- Perry, M. (2010). The exponentially weighted moving average. Dans.
- Rahimi, F., Bustami, A., Hanif, M., Md Saad, M.H., Jailani, M., Mohd Nor, J., . . . Inayatullah, O. (2007). The application of short time fourier transform and image processing techniques to detect human heart abnormalities.
- Rendall, R., Reis, M., 2014. A comparison study of single-scale and multiscale approaches for data-driven and model-based online denoising. *Qual. Reliab. Eng. Int.* 30 <https://doi.org/10.1002/qre.1709>.
- Rioult, O., Vetterli, M., 1991. Wavelets and signal processing. *IEEE Signal Process. Mag.* 8 (4), 14–38. <https://doi.org/10.1109/79.91217>.
- Saho, K., Masughi, M., 2015. Automatic parameter setting method for an accurate Kalman Filter tracker using an analytical steady-state performance index. *IEEE Access* 3, 1. <https://doi.org/10.1109/ACCESS.2015.2486766>.
- Selesnick, I., 2009. The short-time Fourier transform and speech denoising. *Connexions*. <http://cnx.org/content/m32294/1.1/>.
- Shafri, H., Yusof, R., 2009. Determination of optimal wavelet denoising parameters for red edge feature extraction from hyperspectral data. *J. Appl. Remote Sens.* 3 <https://doi.org/10.1117/1.3155804>.
- Simon, D. (2001). Kalman Filtering. *Embedded Systems Programming*.
- Smith III, J.O., 2011a. *Spectral Audio Signal Processing*. Stanford University.
- Smith, J.O., 2011b. Spectral Audio Signal Processing. Center for Computer Research in Music and Acoustics (CCRMA). <https://ccrma.stanford.edu/~jos/sasp/>.
- Smith, T.H., Boning, D.S., 1997. A self-tuning EWMA controller utilizing artificial neural network function approximation techniques. *IEEE Trans. Compon. Packag. Manufact. Technol. Part C* 20 (2), 121–132. <https://doi.org/10.1109/3476.622882>.
- Sraith, M., Jabrane, Y., 2021. A denoising performance comparison based on ECG Signal Decomposition and local means filtering. *Biomed. Signal Process. Control* 69, 102903. <https://doi.org/10.1016/j.bspc.2021.102903>.
- Subramanian, B., Ramasamy, A., Joseva, A., 2013. Performance comparison of hybrid empirical mode decomposition based techniques for electrocardiogram denoising. *Int. J. Simul. Syst. Technol.* 14, 1–8. <https://doi.org/10.5013/IJSSST.a.14.01.01>.
- Sun, S., Huang, D., Gong, Y., 2011. Gross error detection and data reconciliation using historical data. *Procedia Eng.* 15, 55–59. <https://doi.org/10.1016/j.proeng.2011.08.012>.
- Taspinar, A., 2018. A Guide For Using the Wavelet Transform in Machine Learning. <https://attaspinar.com/2018/12/21/a-guide-for-using-the-wavelet-transform-in-machine-learning/>.
- Thibault, É., Chioua, M., McKay, M., Korbel, M., Patience, G., Stuart, P., 2023. Experimental methods in chemical engineering: data processing and data usage in decision-making. *Canad. J. Chem. Eng.* <https://doi.org/10.1002/cje.25014>.
- Upadhyaya, B.R., Mehta, C., Bayram, D., 2014. Integration of time series modeling and wavelet transform for monitoring nuclear plant sensors. *IEEE Trans. Nucl. Sci.* 61 (5), 2628–2635. <https://doi.org/10.1109/TNS.2014.2341035>.
- Wang, Y., Veluvolu, K.C., 2017. Time-frequency analysis of non-stationary biological signals with sparse linear regression based fourier linear combiner. *Sensors (Basel)* 17 (6), 1386. <https://doi.org/10.3390/s17061386>.
- Yinfeng, D., Yingmin, L., Mingkui, X., Ming, L., 2008. Analysis of earthquake ground motions using an improved Hilbert–Huang transform. *Soil Dyn. Earthquake Eng.* 28 (1), 7–19. <https://doi.org/10.1016/j.soildyn.2007.05.002>.
- Yuen, K.V., Hoi, K.I., Mok, K.M., 2007. Selection of noise parameters for Kalman filter. *Earthquake Eng. Eng. Vibrat.* 6 (1), 49. <https://doi.org/10.1007/s11803-007-0659-9>.
- Zabidi, A., Mansor, W., Lee, Y.K., Fadzal, C.W.N.F.C.W., 2012. In: 2012 International Conference on System Engineering and Technology (ICSET), pp. 1–4. <https://doi.org/10.1109/ICSEngT.2012.6339284>.
- Zhao, Z., Särkkä, S., Rad, A.B., 2018. Spectro-temporal ECG analysis for atrial fibrillation detection. In: 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1–6. <https://doi.org/10.1109/MLSP.2018.8517085>.
- Zhou, D., Wang, Q.J., Tian, Q.J., Lin, Q.Z., Fu, W.X., 2009. Wavelet analysis and its application in denoising the spectrum of hyperspectral image. *Guang Pu Xue Yu Guang Pu Fen Xi* 29 (7), 1941–1945.
- Zych, M., Hanus, R., Wilk, B., Petryka, L., Świsłuski, D., 2018. Comparison of noise reduction methods in radiometric correlation measurements of two-phase liquid-gas flows. *Measurement* 129, 288–295. <https://doi.org/10.1016/j.measurement.2018.07.035>.