On Wavelet-based Statistical Process Monitoring

Achraf Cohen, Mohamed Amine Amine Atoui

Journal Title
XX(X):2–37

Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

On Wavelet-based Statistical Process Monitoring

Achraf Cohen[1] and M. Amine Atoui[2]

Abstract

This paper presents an overview of wavelet-based techniques for statistical process monitoring. The use of wavelet has already had an effective contribution to many applications. The increase of data availability has led to the use of wavelet analysis as a tool to reduce, denoise, and process the data before using statistical models for monitoring. The most recent review paper on wavelet-based methods for process monitoring had the goal to review the findings up to 2004. In this paper, we provide a recent reference for researchers and engineers with a different focus. We focus on i) wavelet statistical properties, ii) control charts based on wavelet coefficients, iii) wavelet-based process monitoring methods within a machine learning framework. It is clear from the literature that wavelets are widely used with multivariate methods compared to univariate methods. We also found some potential research areas regarding the use of wavelet in image process monitoring and designing control charts based on wavelet statistics, and listed them in the paper.

Keywords

control charts, data-driven monitoring, multiscale methods, fault detection and diagnosis, wavelet analysis.

Introduction

Quality improvement, system performance, and safety operation have attracted much attention in recent years. These aspects can be achieved by implementing a reliable monitoring system, which includes fault detection and isolation/identification procedures that aim at determining whether or not a fault has occurred and which variables are responsible, respectively. Several approaches have been developed to tackle these aspects based on three main ideas: 1) data-driven approach, which is also referred to as statistical process monitoring, that is concerned with the collected data from processes to develop a statistical monitoring model (Atoui et al. 2019a; Yin et al. 2014; Chiang et al. 2001), 2) knowledge-based approach that is based on experts (Chiang et al. 2001), and 3) model-based approach that requires a priori physical and mathematical knowledge of the process (Isermann 2006; Chiang et al. 2001). Obviously, the best way to implement a monitoring system is to use all three approaches because any description (data, expert and physical knowledge) of the process provides useful information and reinforces system understanding. However, the cost and/or technical environment of the process may encumber practitioners and engineers to use the three approaches simultaneously. This paper focuses on data-driven approach or Statistical Process Monitoring (SPM). More particularly, wavelet-based statistical process monitoring methods.

The rise of big data technologies has contributed to process data sets available in industrial systems. These data sets are characterised by the 4 V's: Volume (from Terabytes to Zettabytes), Velocity (from Batch to

[1]Department of Mathematics and Statistics, University of West Florida, Pensacola, Florida 32514, USA.
[2]CNRS Lab-STICC, Université Bretagne-sud, France. amine.atoui@univ-ubs.fr

Corresponding author:
Achraf Cohen, Department of Mathematics and Statistics, University of West Florida, Pensacola, Florida 32514, USA.
Email: acohen@uwf.edu

Real-time), Variety (from Structured to Unstructured), and Veracity (From Noise to Uncertainty). On the other side, statistical process monitoring serves as an efficient alternative tool where other approaches (model-based, knowledge-based) may fail to provide satisfactory results, such as in complex systems where it is often tough or impossible to come up with an analytical model.

Wavelet-based techniques are often superior in performance, and that is why they are used in various modern applications such as image processing *JPEG2000* and *Wavelet Scalar Quantization* algorithm for fingerprint images developed by FBI (Bradley et al. 1993); condition monitoring (Peng and Chu 2004); chemical engineering (Reis and Saraiva 2006; Choi et al. 2008); machine transmission lines (Lebaroud and Clerc 2008; Liang et al. 1998); bearing and gearbox fault detection and diagnosis (Zarei and Poshtan 2007); wind turbines (Sun et al. 2014); biomedical analysis (Akay 1998); profiles monitoring (McGinnity et al. 2015; Nikoo and Noorossana 2013); statistics (Abramovich et al. 2000), and others (Cohen et al. 2016b,a; Kano et al. 2002b; Gao and Yan 2010; Wang 2012). In the context of statistical process monitoring, wavelet-based methods are popular for the goal of fault detection and diagnosis. Their significant advantages consist of reducing noise, extracting features, and reducing dimension (Jeong et al. 2006).

Statistical process monitoring

Traditional statistical process monitoring methods are essential to understand the variation in a process and to assess its current state (Woodall and Montgomery 2014), such as $\overline{X}$-R (Mean-Range), $\overline{X}$-S (Mean-Standard deviation), EWMA (Exponential Weighted Moving Average), CUSUM (Cumulative Sum), Multivariate EWMA (MEWMA), MCUSUM, χ^2, T^2, Q-statistic, Principal Components Analysis (PCA), Partial Least Squares (PLS), see Figure 1.

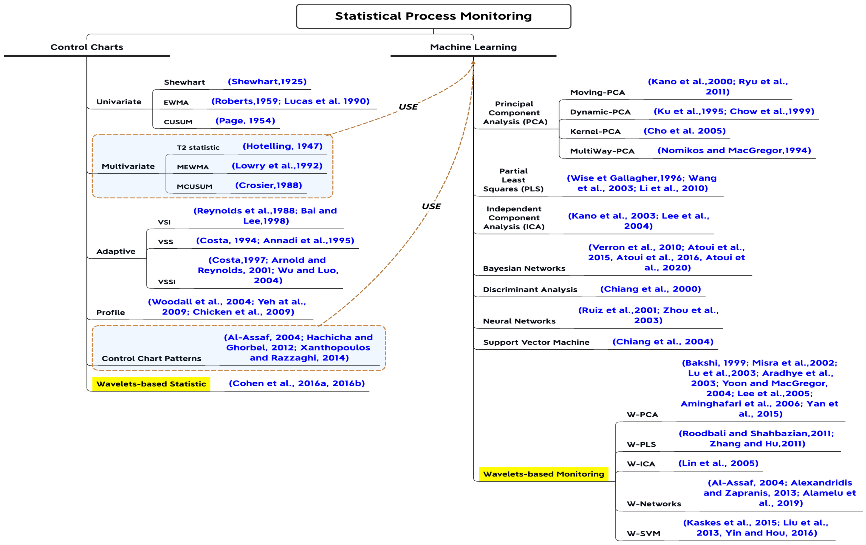

Figure 1. Statistical Process Monitoring Approaches

These methods have been used for decades and still being used because of their simplicity and efficiency to detect assignable faults in the time domain. However, these methods are not able to detect faults in a frequency domain, especially where the time domain cannot provide information about the state of the system. Several techniques in time-frequency domain have been developed and used for process monitoring, such as Linear time-frequency representation (short time Fourier transform and wavelet) and Bilinear time-frequency distribution (e.g. Wigner-Ville, Cohen class) (Feng et al. 2013).

Figure 1 shows a categorization of statistical process monitoring approaches. We distinguish the statistical process control, which mainly means control charts techniques. These are graphical tools that are often based on a plotted statistic that monitors a quality characteristic, which can be univariate, multivariate, adaptive, or profile; the control limits define the area where the control chart does not signal. The second group of techniques are machine learning/data-driven techniques, where the data are often subject to processing/feature extraction/feature selection before applying a statistical monitoring model. For example, control chart patterns (CCPs) is a different way to detect faults using machine learning techniques (Hachicha and Ghorbel 2012). Also, principal component analysis is widely present in the literature of SPM as well as its various extensions, such PCA-dynamic (dynamic process), PCA-moving (autocorrelation), PCA-Kernel (non-linear process).

Wavelet-based methods have drawn attention earlier 2000 for process monitoring in order to deal with measurements noise, autocorrelation, non-normal data, and more particularly time-frequency analysis, where the fault detection can be made in different scales or frequency. This is become a considerable subtopic in statistical process monitoring field. Sometimes called mnultiscale statistical process monitoring. However, multiscale analysis can be conducted using other methods such as empirical mode decomposition. We can distinguish techniques that use wavelet coefficients to develop a statistic to plot in a control chart, and techniques that use wavelet analysis as a preprossessing tool before using the classical technique of statistical process monitoring. For example, Multiscale Principal Component Analysis (MS-PCA) that uses wavelet with PCA to denoise data before applying PCA- Hotelling Statistics for fault detection. Kano et al. (2002b) showed that MS-PCA performs better than DISSIM (Dissimilarity) (Kano et al. (2002a)) and PCA-moving when monitoring Tennessee Eastman Process (TEP).

The scope of the paper

The most recent review paper on wavelet-based methods for process monitoring was published in 2004 by Ganesan et al. (2004). The authors reviewed the use of wavelet analysis for process monitoring and highlighted the advantages and disadvantages of multiscale methods.

More particularly, they presented the use of wavelet in process monitoring with and without process model. Thresholding techniques were also presented. The goal of this paper is to review the up-to-date findings in order to provide a recent reference for researchers and engineers, so we attempt to avoid duplication of the material listed in that paper. We focus on i) wavelet statistical properties, ii) control charts based on wavelet coefficients, iii) wavelet-based process monitoring methods within a machine learning framework.
This paper is organized as follows: Section 2 introduces wavelet and their statistical characteristics; Section 3 presents control charts using wavelets; Section 4 shows the usage of wavelet in a machine learning framework. Finally, in Section 5 conclusions and some possible research directions are presented.

Wavelet for statistical process monitoring

Wavelets were introduced by Jean Morlet in 1983. He came up with the word "wavelet" when he was working on seismic signals as a geophysicist. Afterwards, Grosmann Alex and Yves Meyer (Meyer 1993) developed the mathematical foundations of wavelets. A historical introduction to the subject of wavelet is presented by Hubbard (1998). The theory of Multi-Resolution Analysis (MRA) developed by Mallat (1989) opened the way to apply wavelet to image processing. His work also resulted in the implementation of the Fast Wavelet Transform (FWT) algorithm (Misiti et al. 2003, 1996), and then spreading out the use of wavelet in various applications. Wavelet functions are grouped by families: discrete wavelets that contain Haar (1910), Daubechies, Coiflet, Symlet and Biorthogonal (Daubechies 1992; Cohen et al. 1992); and continuous wavelets such as Morlet and Gaussian (Mallat 1999).
A suitable introduction to wavelet analysis is to compare it to Fourier analysis. In fact, Fourier transform is a decomposition that projects data

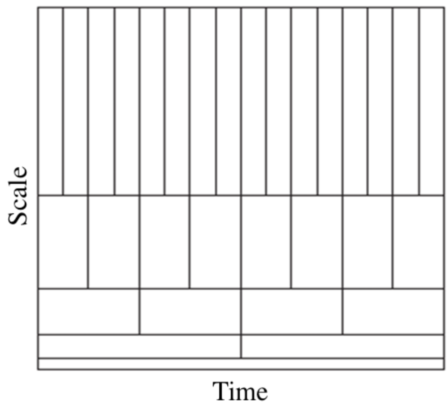

Figure 2. The multiscale representation through a wavelet transform

into a sinusoidal base, where each sinusoid corresponds to a frequency and has a coefficient weight, called Fourier coefficients. Similarly, wavelet transform projects data into a wavelet base, where the size of the window-analysis is variable and not fixed as in Fourier transform. Each window of analysis (given by stretching and shrinking of the mother wavelet) corresponds to a scale of the decomposition and contains the wavelet coefficients. The multiscale representation is given into a Time vs. Scale plan as shown in Figure 2. The data is represented into large scales or resolutions. If you look at the data with a large window (large scale), we would observe "global" features, and if we look at the data with a small window then we would observe "local" features. Wavelet analysis processes the data at different resolutions or scales. An example of Morlet wavelet is given in Figure 3 compared to a sinusoid function.

Mathematically, a wavelet is a square integrable function on Euclidean space $\mathbb{R} \times \mathbb{R}^{*+}$, usually oscillating and must satisfy some eligibility conditions (Daubechies 1992; Meyer 1993). The wavelet transformation is a way to decompose data (signal, image, video, time series, etc.) into a weighted sum of a series of bases localized in time and frequency domains.

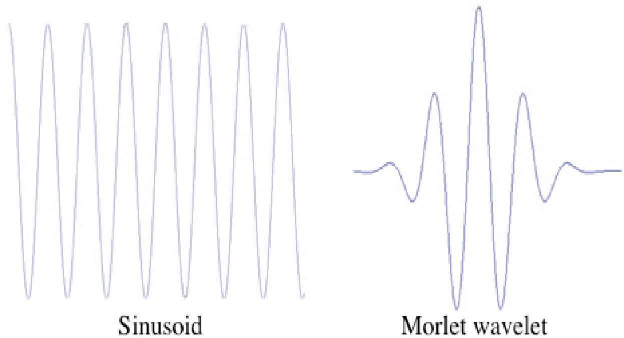

Figure 3. The Morlet wavelet and a sinusoid function

Several wavelet transforms have been proposed, such as Continuous Wavelet Transform, Discrete Wavelet Transform, Stationary Wavelet Transform, Complex Wavelet Transform and others (Mallat 1999; Gao and Yan 2010). Not all wavelet functions have an analytic expression, some are defined by a filter.

Continuous Wavelet Transform (CWT)

The continuous wavelet transform is defined as follows:

$$cwt(\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t)\psi^{*}\left(\frac{t-\tau}{s}\right)dt, \quad (1)$$

where $*$ represents the operation of complex conjugate of the mother wavelet ψ; $s \in \mathbb{R}^{+*}$ and $\tau \in \mathbb{R}$ indicate the scale and translation parameters, respectively. $x(t)$ is the data. The wavelet coefficients $cwt(\tau, s)$ are the convolution result between the data and the wavelet functions. In this transformation, the translation τ parameter is continuous and vary along the data $x(t)$. The transformed data are a function of the translation τ and the scale s parameters. The signal energy here is normalized by dividing the wavelet coefficients by $\frac{1}{\sqrt{s}}$ at each scale. An example of the CWT is given in Figure 4. The large scales correspond to low frequencies and vice versa.

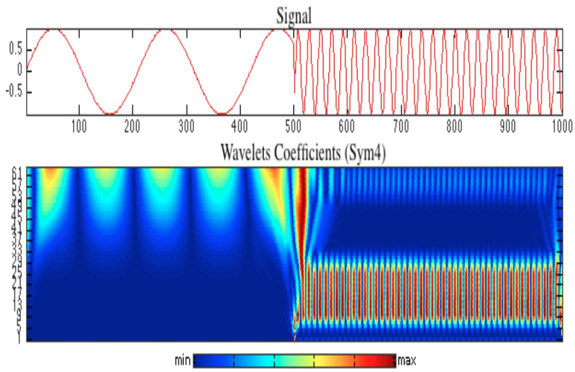

Figure 4. The Continuous Wavelet Transform

The CWT is widely used for signal processing fault detection and diagnosis such as rotary machines fault detection (e.g. bearing and gears). The vibration signals generated from these machines contain a wide range of natural and defect frequencies due to periodic behaviors of the machines. The challenge is extracting significant frequencies within a small-sized pattern for faults diagnosis. In this type of problems, the continuous wavelet coefficients has the advantage to show how well a mother wavelet correlates with a particular signal. If the continuous wavelet coefficient is large in a scale, then the signal has probably a major frequency component corresponding to that scale (Rafiee and Tse 2009). Wu and Chen (2006) proposed a fault diagnosis technique for internal engines using both acoustic and vibration data and applied Morlet wavelet. Pahon et al. (2016) developed a wavelet-based method to diagnosis a high temperature fuel cell. They used wavelet energy and entropy approaches as features. The Daubechies db4 wavelet was used because it is shown that it capture the energy of the signal. Jedliński and Jonak (2015) investigated early fault detection in a gearbox using artificial neural networks and wavelet. Eleven wavelets were examined to find the best for this application: Morlet, biorthogonal 3.1, Coiflet 3, Daubechies 4, Dmeyer, Gaussian, Haar, Mexican hat, Meyer, ReverseBior 3.1 and symlet wavelets. The most suitable type of wavelet in the investigated case was the Haar wavelet. They used the classification accuracy of their artificial neural networks to select the suitable wavelet. An extensive attention was given to the use of continuous wavelet on rotary machines (Peng and Chu 2004; Yan et al. 2014; Chen et al. 2016). Most of the published papers used CWT to extract featured data in a specific frequency band.

Discrete Wavelet Transform (DWT)

CWT is a redundant transformation since the scale and the translation parameters are changed continuously. Although the redundancy is useful in some applications such as noise reduction and feature extraction, other applications may need computational efficiency. This can be achieved by discretization of the scale s and translation τ parameters, as follows:

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k), \tag{2}$$

where $s = 2^j$ and $\tau = ks$; $j, k \in \mathbb{Z}$.

These wavelet bases are orthogonal and defined in the framework of the Multi-Resolution Analysis (MRA), which provides a multiscale decomposition using orthogonal wavelet families across filter banks (Mallat 1989), see Figure 5. The input signal is downsampled and results in two subsignals. The approximation coefficients $a_j(k)$ capture the low frequency or high scale of the input signal (using filter h_n), while the details coefficients $d_j(k)$ capture the high frequency or small scale of the input signal. In a multi-scale decomposition the process is repeated as described in Figure 5.

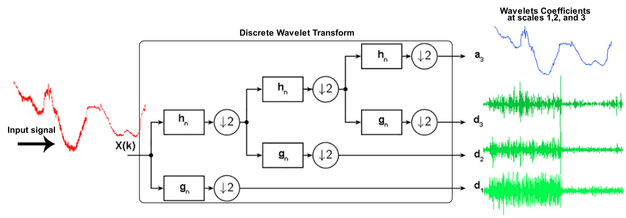

Figure 5. A Discrete Wavelet Transform through Filter Banks

The approximation at scale j is successively decomposed in two subsignals: one detail and one approximation. At each scale of decomposition j, we obtain details coefficients $d_j(k)$ until we reach the maximum level J where we obtain both approximation and details coefficients. An example of a discrete wavelet transform is given in Figure 6. The data denoted by s is decomposed into 5 levels, therefore we obtain the approximations coefficients $a_5(k)$ and the details at each scale d_1, d_2, d_3, d_4, and d_5.

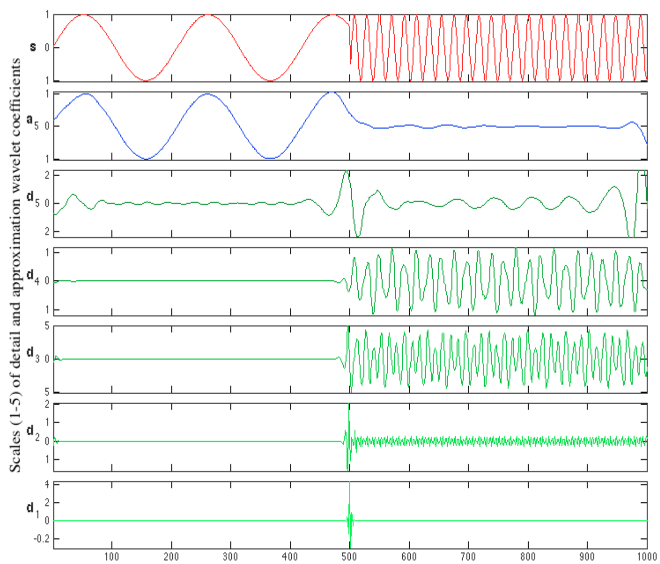

Figure 6. The Discrete Wavelet Transform applied to the signal s

The wavelet coefficients of the DWT, approximations $a_j(k)$ and details $d_j(k)$, are given as follows:

$$a_j(k) = \sum_{i=0}^{l} h[i]a_{j-1}[2k - i], \tag{3}$$

$$d_j(k) = \sum_{i=0}^{l} g[i]a_{j-1}[2k - i], \tag{4}$$

where $a_0 = x$ the original signal, j represents the decomposition scale; $k \in \mathbb{Z}$; l is the filter length; h and g are the scaling (low-pass) and wavelet (high-pass) filters, respectively.

The Discrete wavelet transforms provide parsimonious representations, which have the ability to describe data with a limited number of wavelet coefficients (Mallat 1989; Mallat and Zhong 1992). Therefore, thresholding wavelet coefficients techniques have shown an excellent performance for reducing noise in data. Several thresholds have been developed: VisuShrink (Donoho and Johnstone 1994), RiskShrink, SUREShrink (Donoho and Johnstone 1995; Donoho 1995), FirmShrink (Gao et al. 1997; Gao 1998). Andrade et al. (2016) proposed an adaptive threshold allowing the segmentation of electric signals to analyze the power quality. Kumar and Singh (2013) proposed a technique based on the DWT using Symlet wavelet for measuring outer race defect width of taper roller bearing.

Statistical properties of wavelet coefficients

There have been advances in recent years in the use of wavelet approaches for statistical modeling and applications. But a few research papers have reported the statistical characteristics of wavelet coefficients. Regarding statistical applications, the central use of wavelet analysis has been in nonparametric regression and density function estimation.

This is basically done by using the discrete wavelet transform, applying a thresholding rule (Donoho and Johnstone 1994, 1995), and then reconstructing the function using the inverse wavelet transform (Nason and Silverman 1995; Abramovich et al. 2000). Additionally, wavelet decompositions have shown good time-frequency localization (Daubechies 1990), which is a reasonable catalyst for their use in change-point problems.

The wavelet coefficients are the result of the convolution product of wavelet/scaling functions and the signal under consideration. Considering $\boldsymbol{X} = [X_1, X_2, ..., X_n]$ is a sample/signal, where X_i are independent and identically distributed random variables, which have f_{X_i} as a probability density function. The wavelet coefficients, approximations $a_j(k)$ and details $d_j(k)$ (Equations (3) and (4)), are defined as a linear combination of the random variables X_i that come from the same distribution family, which corresponds in the context of the statistical distribution to the convolutions of probability distributions f_{X_i}. Some of these convolutions have already been derived in the literature, such as the linear combination of exponential distributions (Ali and Obaidullah 1982) and $\chi 2$ chi-square distribution (Davies 1980). For additional distributions, see (Nadarajah and Kotz 2005; Johnson et al. 1994, 1995). In the case of wavelet coefficients, the prior results must first be adapted to the wavelet functions (filters) in order to derive the exact distribution of the wavelet coefficients. Various combinations could be studied for the existing different wavelets and the probability distributions of the data.

An asymptotic result is given in Ogden (1997). It is mentioned that, for normal data the approximation and details wavelet coefficients are asymptotically normal with order $O(1/n)$. Vannucci and Corradi (1999) presented some results on the covariance structure of wavelet coefficient in the case 2D wavelet transformations with a Bayesian perspective. In Ganesan et al. (2004), it is mentioned that wavelet coefficients are

Gaussian even if the data are non-normally distributed, but no references were given.

The distribution of wavelet coefficients depends on the distribution of the data of interest. For normal data, it is shown that wavelet coefficients follow a normal distribution. Assume $\boldsymbol{X} = [X_1, X_2, ..., X_n]$ is a signal, where X_i are independent and identically distributed random variables $X_i \sim \mathcal{N}(\mu_0, \sigma_0^2)$. Consider Orthonormal and Biorthogonal compactly supported wavelets (Haar, Daubechies, Symlets, Coiflets, Discrete Meyer, Biorthogonal, Reverse Biorthogonal). The multiresolution analysis applied to X provides wavelet coefficients as follows (Cohen et al. 2016a):

$$a_j(k) \sim \mathcal{N}(2^{j/2}\mu_0, \left(\sum_n h_n^2\right)^j \sigma_0^2), \tag{5}$$

$$d_j(k) \sim \mathcal{N}(0, \sum_n g_n^2 \left(\sum_n h_n^2\right)^{j-1} \sigma_0^2), \tag{6}$$

which are identically distributed random variables, and independent if the orthonormal wavelets are used, else they are slightly correlated. The wavelet coefficients are summation of normally distributed variables consequently they follow the normal distribution. Furthermore, for orthonormal wavelets families the wavelet coefficients are independent at each scale. The independence of wavelet coefficients is a consequence of the projection into orthonormal bases. This is not the case for Biorthogonal bases where the correlation coefficient can be estimated by empirical studies.

Aspects of the Wavelet-based Methods

There are three main aspects in multiscale methods that are valuable and should be taken into account while designing a wavelet-based statistical monitoring model. These include 1) the level of the decomposition, 2) the window size (e.g. a subgroup/sample in control charts), and 3) the wavelet selection. The level of the decomposition is related to the the window size with the formula $n = 2^J$, where J is the maximum level of the decomposition that can be applied to a sample of size n. In the case where the sample size does not equal to 2^J, the concept of wavelet on the interval is proposed (Meyer 1991; Cohen et al. 1993), but still not entirely satisfactory from a practical viewpoint. However, several numerical algorithms can be used to solve this problem (Strang and Nguyen 1996, Chapter 8). So we can adapt the wavelet-based method to any sample size. The issue is that by violating this condition ($n = 2^J$) correlation between wavelet coefficients can be created even if the observations are mutually independent.

The choice of the wavelet families The subject of selecting a wavelet for a given application is valuable because the choice may affect the result of wavelet transform and then the performance of the multiscale method at the end. Wavelet families have some properties, such as *symmetry*, *orthogonality*, and *compact support*. Having knowledge and understanding these properties should be helpful for selecting a candidate wavelet from the wavelet families for a specific application. For example, a *compact support* wavelet is a nonzero function only within a finite interval. This is an important property for data compression. The *orthogonality* means that the inner product of the wavelet with itself is unity, and zero with other scaled or shifted wavelets. This property is effective to decompose a signal into non-overlapping sub-frequency bands. The *symmetry* feature is useful to filtering operations. Figure 7 shows the common criteria used to select a wavelet family. A well presented description of wavelet selection measures is given in (Gao and Yan 2010, Chapter 10).

In the last decade, researchers have used different criteria to select a wavelet that is suitable for a specific application. For instance, Rafiee and

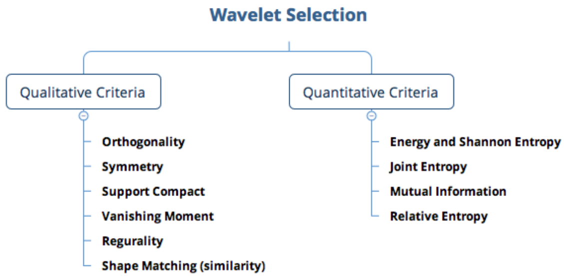

Figure 7. Wavelet Selection Criteria

Tse (2009) studied 324 candidate wavelets to find the most suitable one to fit the gearbox vibration signals. They used the variance of the continuous wavelet coefficients for each scale to select the most similar wavelet to their vibration data. The wavelet that was selected is Daubechies 44 $db44$. Kankar et al. (2011) uses Minimum Shannon Entropy criterion to select the most appropriate wavelet for fault diagnosis in rolling element bearing. The authors considered Complex Morlet. Kumar and Singh (2013) uses Symlet wavelet to measure the outer race defect size in taper roller bearing. The criteria used for the wavelet selection is the fact that is a near symmetrical/linear phase filter, which has advantage to deal with the small discontinuity present in the signal. Zhang et al. (2016) used the Symlet $sym3$ and Daubechies $db4$ wavelets in order to detect faults in the power transmission line, and the symmetrical wavelet $sym6$ is used to denoise the reflection waveform. Jedliński and Jonak (2015) used the classification accuracy of their artificial neural networks model to select the Haar wavelet as the suitable wavelet for their application. Another way to find the optimal choice of wavelet is the shape matching and correlation between Shape of the wavelet and the data being analyzed (Yan and Gao 2009).

Wavelet-based control charts

The use of wavelets analysis in control chart field has been addressed in two ways: 1) the use of the multiscale analysis to reduce noise and extract features in order to improve recognition of control chart patterns, and 2) the design of control chart statistics for detecting mean and/or variance changes.

Control Chart Patterns

Control charts have played an important role to improve product quality and to monitor processes for many decades. The process variability observed through the data results from either natural variation or unnatural variation. The goal is to discriminate between these two types of variation. The natural variation is inherent to the system. However, the unnatural variation often reflects a specific fault or a set of faults. Phase I for designing a statistical process monitoring system consists of learning and understanding the natural variation for a well-functioning system, this is named as natural pattern. This phase contains valuable information for tuning parameters and decision rules. In Phase II, a traditional control chart is used to detect faults (unnatural variation). Unfortunately, control charts often do not provide any pattern-related information to be capable to recognize different kinds of unnatural patterns (faults). These faults can be associated with a pattern-related cause. Seven common categories of control chart patterns exist: 1) natural pattern, 2) cyclic pattern, 3) upward shift, 4) downward shift, 5) upward trend, 6) downward trend, and 7) systematic pattern. These patterns can be described by some associated causes. For example, cyclic patterns can be observed in periodic rotation of operators or fluctuations in the equipment. Trend patterns concern with tool wear, operator fatigue, and equipment failure. Shift patterns often indicate an abrupt shift or change in the quality characteristic, this can describe failure sensor, introduction of new employees, and replacement of raw materials, see Figure 8. A survey related to control chart pattern recognition approaches is given in Hachicha and Ghorbel (2012).

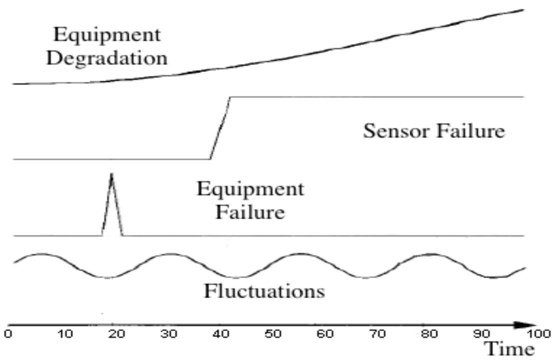

Figure 8. Example of Control Chart Patterns Without Noise (adapted from Bakshi (1998))

In real application, data collected from processes often contain noise, outliers, and they are in various scales (frequencies). Moreover, the control chart patterns or data are often non-stationary in sense that a trend pattern has low-frequency components while a shift pattern is a high-frequency component. In other words, the trend patterns need more time to be seen and a shift pattern is observed in a short time. wavelet-based methods are very useful to process these patterns in order to display their components in the time-frequency domain, which can enhance the performance of the fault detection and diagnosis.

Wang and Kuo (2007) have proposed a framework to identify six common types of control chart patterns. They particularly used a multiscale filter based on wavelet to reduce the noise. A fuzzy clustering algorithm is then adopted to discriminate patterns. They used Haar wavelet to enhance the interpretation of abrupt changes in data. Their method consists of decomposing the noisy signal into wavelet domain, then apply a threshold rules and finally reconstruct the data from the thresholded wavelet coefficients. Du et al. (2013) have studied the case of concurrent control chart patterns, in which two simultaneous patterns are considered. This situation is realistic because a process signal often contains more than one pattern, and methods that detect simultaneous faults can be valuable to fault diagnosis. The authors used the multiscale decomposition to decompose an input pattern into two single patterns using Haar wavelet transform. The multiscale decomposition allowed them to separate the concurrent control chart patterns to single patterns. Then they used a support vector machine algorithm to recognize and classify these patterns. Wavelet-based methods combined with neural network for recognition of concurrent control chart patterns was proposed by Al-Assaf (2004). Chen et al. (2007) have also suggested a method combining wavelet transform and neural network. They showed that the traditional run-rule based approach and stand alone artificial neural network approach are not capable for recognizing concurrent patterns, however incorporating wavelet decomposition allowed the recognition of concurrent patterns. Ranaee and Ebrahimzadeh (2011) proposed a method for recognition common types of control chart patterns. Their proposed approach made up of a feature extraction module based on the wavelet decomposition. Lee et al. (2012) proposed a wavelet-based distribution-free CUSUM chart for detecting shifts in the mean of a profile with noisy components. They focused on monitoring key components of the discrete wavelet transform. A dimension reduction technique was proposed based on thresholding the wavelet coefficients. A wavelet-based distribution-free tabular CUSUM chart with an adaptive thresholding has been proposed by Wang et al. (2015). Another approach that uses Haar wavelet coefficients in an SPC setting for detecting process drifts is presented here Wang et al. (2014). Their method involves the wavelet coefficients at a predetermined optimal (wavelet) level using CUSUM and EWMA control charts.

Jeong et al. (2013) used an adaptive thresholding test statistic to select wavelet coefficients adaptively according to process changes in the wavelet domain. The authors then applied an EWMA control chart on that test. The proposed control chart was applied to detect small shifts in non-linear profiles of a plasma etch process in semiconductor manufacturing. Mansouri et al. (2018) proposed a new method (WOEWMA) using wavelet with EWMA control chart applied to photovoltaic systems. The use of wavelet was in order to obtain deterministic features as well as de-correlate the data.

Wavelet thresholding is the process of cutting off some of the wavelet coefficients based on a specific threshold then reconstruct the data using the inverse of wavelet transformations. Various thresholds are proposed in the literature such as VisuShrink (Donoho and Johnstone 1994), RiskShrink, SUREShrink (Donoho and Johnstone 1995; Donoho 1995), FirmShrink (Gao et al. 1997; Gao 1998). These techniques are often used with statistical methods for monitoring processes. For instance, multiscale Principal Component Analysis (PCA) (Bakshi 1998; Aradhye et al. 2003) was proposed for dimension reduction, and it is based on combining wavelet analysis and principal component analysis. This methodology has been widely used in the literature of chemical process monitoring. It consists of applying wavelet transform on the data and then reconstruct the data after a thresholding technique is applied. Yoon and MacGregor (2004) used the multiscale PCA approach for fault detection and diagnosis, where they applied it to the Continuous Flow Stirred-Tank Reactor (CSTR) process, and they showed the usefulness of wavelet analysis to isolate the faults when prior knowledge about their frequencies is given. Sheriff et al. (2017) proposed a hybrid data-driven fault detection method where they improved the performance of the generalized likelihood ratio test chart using a moving window and wavelet analysis; see also (Kini and Madakyaru 2019; Reis and Saraiva 2006; Lee et al. 2005; Maulud et al. 2006; Aminghafari et al. 2006).

Combining wavelet analysis with partial least squares was proposed for process monitoring using the same approach we described before (Teppola and Minkkinen 2000; Lee et al. 2009; Roodbali and Shahbazian 2011; Zhang and Hu 2011; Madakyaru et al. 2016). These methodologies are used with control chart statistics such as Hotelling- T^2 and Q.

Design of Control Chart Wavelet-Statistics

There have been several advances in theory and application of wavelet as a pre-processing tool. However, less attention has been given to propose a wavelet control chart statistic for mean and/or variance changes. It is shown that by using discrete wavelet transform, the wavelet approximation coefficient can be used to monitor the process mean, and the wavelet detail coefficients can be used to control the process variability. More particularly, the Haar discrete wavelet transform is equivalent to the $\overline{X} - R$ control charts scheme (Cohen et al. 2016a).

An Illustrative Example: Weighted Wavelet Coefficients for Process Mean

In this example, we consider a window length/sample size of $n = 8$ observations and the Daubechies 2 $db2$ wavelet, and then the discrete wavelet transform is applied. Consequently, we have eight wavelet coefficients at the scale one (maximum decomposition level in this case): four approximations coefficients and four details coefficients. The aim of this example is to show the behavior of wavelet coefficients using $db2$ when a mean change is occurring. The first 39 points plotted in Fig. 9(a) consist of observations randomly generated from a normal distribution $\mathcal{N}(0, 1)$ (simulated as a in control process), and the last 41 observations consist of observations randomly generated from a normal distribution $\mathcal{N}(20, 1)$ out-of-control process with mean change.

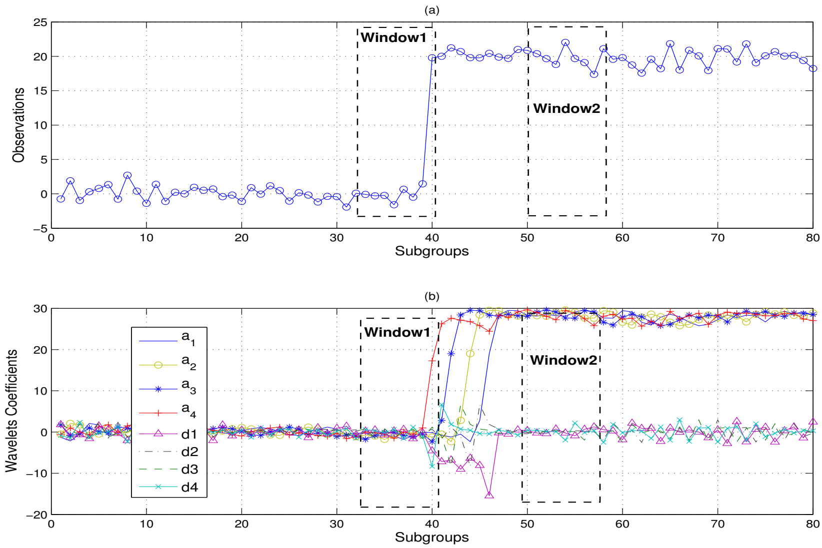

Figure 9. (a): Observations with mean change (positive shift); (b): wavelet coefficients (approximations a_i and details d_i) using $db2$ wavelet (adopted from Cohen et al. (2016b))

Figure 9(b) shows that the fourth approximation coefficient a_4 is the most sensitive to the change in the mean. As progressively the moving window enters in the region with the mean change (see Window 1 in Fig. 9(a)), the other wavelet approximation coefficients will be progressively sensitive to the mean change, and they will converge (see Window 2 in Fig. 9(b)) to $2^{j/2} \times \mu_0 = 2^{1/2} \times 20 = 28.28$. On the other side, details coefficients reveal also the mean change, but they are less sensitive than approximations coefficients. Details wavelet coefficients converge to zero, see Window2 in Fig. 9(b). However, we can note that the first detail coefficients (d_1) is the most sensitive one. We present a statistic called $OWave$ that is based on the following wavelet coefficients: a_1, a_2, a_3, a_4, d_1. It is based on weighted wavelet coefficients, as follows:

$$OWave_i = w_1 a_{1,i} + w_2 a_{2,i} + w_3 a_{3,i} + w_4 a_{4,i} + w_5 d_{1,i}, \qquad (7)$$

where i is the index of the moving window across the signal, $\sum_{i=1}^{5} w_i = 1$, $0 \leq w_{i \in \{1,2,3,4\}} \leq 1$, and $-1 \leq w_5 \leq 0$ because the detail coefficient d_1 shifts negatively when the mean increases (see Fig. 9(b)). On the other hand, the d_1 shifts positively when the mean decreases. In fact, wavelet coefficients have symmetrical behaviour with positive and negative mean shifts, then one can use the same weights w_i and symmetric control limit, in order to detect both positively and negatively shifts in the mean, more details are given in Cohen et al. (2016b).

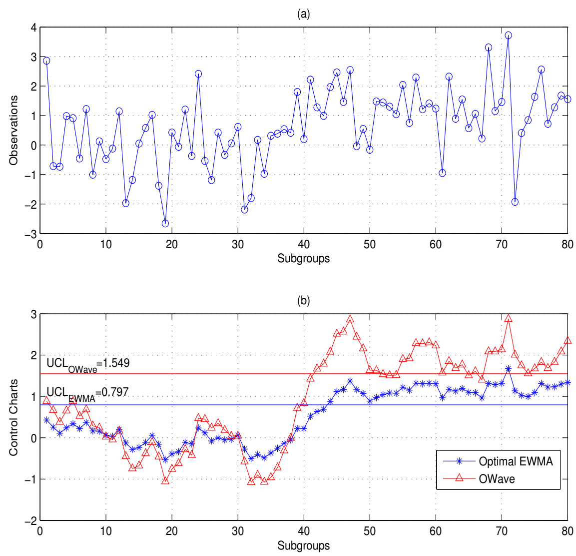

Figure 10. (a) Observations with mean change, at subgroup 40; (b) $OWave$ and Optimal EWMA ($\delta = 1$) control charts behavior

In Figure 10, the $OWave$ statistic behavior is given. The first 39 points plotted, are generated from $\mathcal{N}(0, 1)$ distribution, which are corresponding to in-control process. While the last 41 points are following $\mathcal{N}(1, 1)$ distribution corresponding to a change, $\delta = 1$, in the process mean. In Figure 10(a), a Gaussian model with mean change (δ=1) is plotted. In Figure 10(b), $OWave$ and EWMA control charts are displayed. It is shown that $OWave$ control chart perform slightly better than EWMA, CUSUM, and $\overline{X}$ in terms of Average Run Length.

Other research has been applying existing control charts to wavelet coefficients instead of the original data. Harrou et al. (2018) combined proprieties of the discrete wavelet transform and the exponentially weighted moving average control chart to appropriately detect faults in PV systems. Specifically, this approach was employed to monitor the residuals generated by a simulated model of a single-diode modeling for fault detection purposes. Similar work have been done to monitor swarm robotics systems performing a virtual visco-elastic control model for circle formation task. The proposed mechanism is applied to the uncorrelated residuals from principal component analysis model (Harrou et al. 2018), see also (Harrou et al. 2019). A distribution-free approach using a multivariate cumulative sum (CUSUM) control chart to monitor wavelet coefficients is proposed to detect location shifts Li et al. (2019).

Wavelet with Machine Learning for SPM

Machine learning has become a very important field that intersects with statistics, computer science, artificial intelligence and other engineering areas. Statistical learning has many applications in many areas of science such as monitoring complex systems. Machine learning methods include several statistical methods and generally classified as supervised or unsupervised techniques (Bishop 2006; Friedman et al. 2001).

Consider we want to monitor a system by estimating/predicting its state based on a set of features/variables. We have historical data from the system in which we observe the outcome (quantitative or categorical) and some feature measurements for a set of objects. Using these data we construct a predictive model, which enables us to predict the outcome for new objects. A good model is one that accurately predicts such an outcome. We just described above what we called supervised learning problem. It is called *supervised* because of the presence of the outcome variable to guide the learning process. In the *unsupervised* learning, we observe only the features and have no measurements of the outcome. Then the task here is rather to describe how the data are organized and clustered. Several machine learning methods are being used in statistical process monitoring context such as support vector machines, decision tree, Bayesian networks, Neural Networks, discriminant analysis, k-means, and principal component analysis with Hotelling statistics (Atoui et al. 2019b; Ge et al. 2017; Yin et al. 2014; Cohen et al. 2016a; Atoui et al. 2016, 2015).

Most statistical learning methods can be used with wavelet in an attempt to obtain better predictive performance. Wavelets are mainly used to achieve the following goals: noise reduction, creation of new features, and extraction of information in time-frequency or time-scale space. Wavelet features are useful to many process monitoring studies as wavelet can decompose the information for further analysis. One of the most usage of wavelet is to decompose the data and use the wavelet coefficients or statistics of wavelet coefficients to characterise a data set or an object (e.g. faulty or normal operating conditions). Alamelu Manghai and Jegadeeshwaran (2019) investigated the applications of wavelet for diagnosing the faults on a hydraulic brake system. They considered a list of wavelet families: Haar, Daubechies, Symlet, Coiflets, Discrete Meyer and others. They also performed a thresholding procedure to reduce the noise, and use statistics based on the de-noised data. The classification was conducted with the following techniques: decision tree, support vector machine, and neural network. It is also shown in their paper that Discrete Meyer wavelet provided the best classification accuracy across the different classification methods. Jung (2017) introduced new feature extraction technique to alleviate the high dimensionality problem of implementing multivariate statistical process monitoring when the quality characteristic is a vibration signal from bearing system. A set of multiscale wavelet scalogram features was generated to reduce the dimensionality of data, and is combined with the bootstrapping technique as nonparametric density estimation to set up an upper control limit of control chart. The simulation of a bearing system showed that the proposed method has satisfactory fault-discriminating ability without any distributional assumption.

The combination of wavelet and support vector machine has been considerably developed for the last decade (Yin and Hou 2016). It is shown that wavelet can help reduce the number of iterations when training classifiers such as SVM and Neural network. Liu et al. (2013) proposed a wavelet support vector machine technique to detect the bearing fault of electric locomotives. They showed wavelet based SVM is better than SVM in accuracy. Wavelet SVM is also well used to diagnose faults in induction motors (Keskes et al. 2013; Das et al. 2010). Most often, frequency components are hardly detected in the stator current due to its low magnitude and closeness to the supply frequency component. To overcome this drawback, the wavelet packet transform is applied to extract one parameter able to detect the fault with arbitrary working conditions and a great concern of low load cases. Different multiclass support vector machines (MSVMs) methods are evaluated with respect to accuracy, number of support vectors, and testing time. The experimental results confirm that the DAG SVMs and Symlet wavelet kernel function are fast, robust, and give the best classification accuracy of 99% (Keskes and Braham 2015).

Time-frequency strategies for process monitoring have been extensively using wavelet transformations because they can provide the time where the frequency changes. This cannot be achieved with the traditional Fourier transform. In this regard, wavelet has been the main focus to extract features in time-frequency domain. Besides the cost of a Fast Fourier

Transform is $O(n \log(n))$ and for the Fast Wavelet Transform is $O(n)$. It is worth noting that there exists other techniques for time-frequency analysis such as the Empirical Mode Decomposition (EMD). A methodology for fault detection has been proposed for Induction Motors (IMs) to detect various electrical and mechanical faults based on wavelet and support vector machine (SVM). For this, the radial, axial and tangential vibrations, and three-phase current signals are acquired from IMs having different faults. The acquired time domain signal is then transformed to timefrequency signals using continuous wavelet transform (CWT). Ten different base wavelets are used to investigate the impact of different wavelet function on the fault diagnosis of IMs. Statistical features are extracted based on the CWT, and then appropriate feature(s) are selected using the wrapper model (Gangsar and Tiwari 2019). Fault diagnosis in induction motors has been developed using wavelet features (Monfared et al. 2019; Zgarni et al. 2017; Hmida and Braham 2016).

Rato and Reis (2015) proposed a multiscale approach to deal with changes in the networked structure of process data. The authors used the sensitivity enhancing transformations to detect changes in the process, and they showed that the wavelet can be useful when the system is difficult to model. Gillis and Morsi (2017) presented a new technique based on semi-supervised machine learning and wavelet design applied to non-intrusive load monitoring.

to improve the fault diagnosis performance for rotating machinery a deep learning (DL) algorithm is proposed based on the advantages of the wavelet packet transform in vibration signal processing (the capability to extract multiscale information and more spectral distribution features) and deep convolutional neural networks (good classification performance, data-driven design and high transfer-learning ability) (Ma et al. 2019).

Some research directions

Wavelet analysis has been used for decades in statistical process monitoring. They have been extensively employed to multivariate analysis to reduce dimension, reduce noise, and extract features. There are two potential research directions that need more attention:

- *Wavelet-based statistics for control charts*: Wavelet coefficients are function of sample observations. The idea to use them as new data is not well investigated. Moreover, statistics based on wavelet coefficients have not been yet well explored. In the literature, a very small number of published papers use a statistic based on wavelet coefficients to monitor the mean and/or the variance of the process in an univarate case. Also, wavelet coefficients can be useful when data are autocorrelated (Jeske et al. 2018; Cohen et al. 2015, 2016b).
- *Image statistical control using wavelet*: Image data are become available in today's industries (Koosha et al. 2017; Megahed et al. 2011). In this context a data image is taken from a process, and using 2D wavelet transformations features are extracted and a statistic can be derived to monitor and plot on the control chart. Wavelet analysis has been widely used in image processing and we expect more papers on the use of advanced image processing using wavelet to be applied to image statistical process control (Amirkhani and Amiri 2020; Zuo et al. 2019).

Conclusions

Wavelet based methods for statistical process monitoring have been studied for decades and enormous contributions have allowed better performance for fault detection and diagnosis. They are extensively employed to achieve, in a few sentences as a conclusion, the following:

- Wavelet transformations are often used to reduce noise via threshold techniques.
- Wavelet can be used to characterize faults in time-frequency domain. This needs a better knowledge of the physics related to the faults such in bearing fault detection.
- Wavelet coefficients can be used as input data (features), instead of the original data. Another approach is to use statistic of wavelet coefficients (details or approximations) as input data. This often improves the machine learning performance algorithm.

In this paper, we introduced wavelet analysis as well as give an overview of the aspects related to its application to statistical process monitoring.

References

Abramovich, F., Bailey, T. C., and Sapatinas, T. (2000). Wavelet analysis and its statistical applications. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(1):1–29.

Akay, M. (1998). *Time Frequency and Wavelets in Biomedical Signal Processing*. IEEE press series in Biomedical Engineering.

Al-Assaf, Y. (2004). Multi-resolution wavelets analysis approach for the recognition of concurrent control chart patterns. *Quality engineering*, 17(1):11–21.

Alamelu Manghai, T. and Jegadeeshwaran, R. (2019). Vibration based brake health monitoring using wavelet features: A machine learning approach. *Journal of Vibration and Control*, 25(18):2534–2550.

Ali, M. M. and Obaidullah, M. (1982). Distribution of linear combination of exponential variates. *Communications in Statistics-Theory and Methods*, 11(13):1453–1463.

Aminghafari, M., Cheze, N., and Poggi, J.-M. (2006). Multivariate denoising using wavelets and principal component analysis. *Computational Statistics and Data Analysis*, 50(9):2381–2398.

Amirkhani, F. and Amiri, A. (2020). A novel framework for spatiotemporal monitoring and post-signal diagnosis of processes with image data. *Quality and Reliability Engineering International*, 36(2):705–735.

Andrade, L. C., Oleskovicz, M., and Fernandes, R. A. (2016). Adaptive threshold based on wavelet transform applied to the segmentation of single and combined power quality disturbances. *Applied Soft Computing*, 38:967–977.

Aradhye, H. B., Bakshi, B. R., Strauss, R. A., and Davis, J. F. (2003). Multiscale spc using wavelets: Theoretical analysis and properties. *AIChE Journal*, 49(4):939–958.

Atoui, M. A., Cohen, A., Verron, S., and Kobi, A. (2019a). A single bayesian network classifier for monitoring with unknown classes. *Engineering Applications of Artificial Intelligence*, 85:681–690.

Atoui, M. A., Cohen, A., Verron, S., and Kobi, A. (2019b). A single bayesian network classifier for monitoring with unknown classes. *Engineering Applications of Artificial Intelligence*, 85:681–690.

Atoui, M. A., Verron, S., and Kobi, A. (2015). Fault detection with conditional gaussian network. *Engineering Applications of Artificial Intelligence*, 45:473–481.

Atoui, M. A., Verron, S., and Kobi, A. (2016). A bayesian network dealing with measurements and residuals for system monitoring. *Transactions of the Institute of Measurement and Control*, 38(4):373–384.

Bakshi, B. R. (1998). Multiscale PCA with application to multivariate statistical process monitoring. *AIChE journal*, 44(7):1596–1610.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer Science+ Business Media.

Bradley, J. N., Brislawn, C. M., and Hopper, T. (1993). FBI wavelet/scalar quantization standard for gray-scale fingerprint image compression. In *Optical Engineering and Photonics in Aerospace Sensing*, pages 293–304. International Society for Optics and Photonics.

Chen, J., Li, Z., Pan, J., Chen, G., Zi, Y., Yuan, J., Chen, B., and He, Z. (2016). Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review. *Mechanical systems and signal processing*, 70:1–35.

Chen, Z., Lu, S., and Lam, S. (2007). A hybrid system for SPC concurrent pattern recognition. *Advanced Engineering Informatics*, 21(3):303–310.

Chiang, L. H., Braatz, R. D., and Russell, E. (2001). *Fault Detection and Diagnosis in Industrial Systems*. Springer.

Choi, S. W., Morris, J., and Lee, I.-B. (2008). Nonlinear multiscale modelling for fault detection and identification. *Chemical engineering science*, 63(8):2252–2266.

Cohen, A., Daubechies, I., and Feauveau, J.-C. (1992). Biorthogonal bases of compactly supported wavelets. *Communications on pure and applied mathematics*, 45(5):485–560.

Cohen, A., Daubechies, I., and Vial, P. (1993). Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 1(1):54–81.

Cohen, A., Tiplica, T., and Kobi, A. (2015). Statistical process control for AR(1) or non-gaussian processes using wavelets coefficients. In *Journal of Physics: Conference Series*, volume 659, page 012043. IOP Publishing.

Cohen, A., Tiplica, T., and Kobi, A. (2016a). Design of experiments and statistical process control using wavelets analysis. *Control Engineering Practice*, 49:129–138.

Cohen, A., Tiplica, T., and Kobi, A. (2016b). Owave control chart for monitoring the process mean. *Control Engineering Practice*, 54:223–230.

Das, S., Koley, C., Purkait, P., and Chakravorti, S. (2010). Wavelet aided SVM classifier for stator inter-turn fault monitoring in induction motors. In *IEEE PES general meeting*, pages 1–6. IEEE.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36(5):961–1005.

Daubechies, I. (1992). *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, Pa.

Davies, R. B. (1980). Algorithm as 155: The distribution of a linear combination of χ 2 random variables. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 29(3):323–333.

Donoho, D. (1995). De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627.

Donoho, D. L. and Johnstone, J. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455.

Donoho, D. L. and Johnstone, J. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *Jouranl of the American Statistical Association*, pages 1200–1224.

Du, S., Huang, D., and Lv, J. (2013). Recognition of concurrent control chart patterns using wavelet transform decomposition and multiclass support vector machines. *Computers and Industrial Engineering*, 66(4):683–695.

Feng, Z., Liang, M., and Chu, F. (2013). Recent advances in time-frequency analysis methods for machinery fault diagnosis: A review with application examples. *Mechanical Systems and Signal Processing*, 38(1):165–205.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer series in statistics New York.

Ganesan, R., Das, T. K., and Venkataraman, V. (2004). Wavelet-based multiscale statistical process monitoring: A literature review. *IIE transactions*, 36(9):787–806.

Gangsar, P. and Tiwari, R. (2019). Diagnostics of mechanical and electrical faults in induction motors using wavelet-based features of vibration and current through support vector machine algorithms for various operating conditions. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 41(2):71.

Gao, H.-Y. (1998). Wavelet shrinkage denoising using the non-negative garrote. *Journal of Computational and Graphical Statistics*, 7(4):469–488.

Gao, H.-Y., Bruce, A. G., et al. (1997). Waveshrink with firm shrinkage. *Statistica Sinica*, 7(4):855–874.

Gao, R. X. and Yan, R. (2010). *Wavelets: Theory and Applications for manufacturing*. Springer.

Ge, Z., Song, Z., Ding, S. X., and Huang, B. (2017). Data mining and analytics in the process industry: The role of machine learning. *IEEE Access*, 5:20590–20616.

Gillis, J. M. and Morsi, W. G. (2017). Non-intrusive load monitoring using semi-supervised machine learning and wavelet design. *IEEE Transactions on Smart Grid*, 8(6):2648–2655.

Haar, A. (1910). Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, 69(3):331–371.

Hachicha, W. and Ghorbel, A. (2012). A survey of control-chart pattern-recognition literature (1991–2010) based on a new conceptual classification scheme. *Computers and Industrial Engineering*, 63(1):204–222.

Harrou, F., Khaldi, B., Sun, Y., and Cherif, F. (2018). Statistical detection of faults in swarm robots under noisy conditions. In *2018 6th International Conference on Control Engineering & Information Technology (CEIT)*, pages 1–6. IEEE.

Harrou, F., Taghezouit, B., and Sun, Y. (2018). A robust monitoring technique for fault detection in grid-connected pv plants. In *2018 7th International Conference on Renewable Energy Research and Applications (ICRERA)*, pages 594–598.

Harrou, F., Taghezouit, B., and Sun, Y. (2019). Robust and flexible strategy for fault detection in grid-connected photovoltaic systems. *Energy Conversion and Management*, 180:1153 – 1166.

Hmida, M. A. and Braham, A. (2016). Arm based RSWPT implementation for embedded condition monitoring of induction motor. In *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, pages 1464–1469.

Hubbard, B. B. (1998). *The World According to Wavelets The Story of a Mathematical Technique in the Making*. Universities Press.

Isermann, R. (2006). *Fault-diagnosis systems : An Introduction from Fault Detection to Fault Tolerance*. Springer.

Jedliński, Ł. and Jonak, J. (2015). Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform. *Applied Soft Computing*, 30:636–641.

Jeong, M. K., Lu, J.-C., Huo, X., Vidakovic, B., and Chen, D. (2006). Wavelet-based data reduction techniques for process fault detection. *Technometrics*, 48(1).

Jeong, Y.-S., Kim, B., and Ko, Y.-D. (2013). Exponentially weighted moving average-based procedure with adaptive thresholding for monitoring nonlinear profiles: Monitoring of plasma etch process in semiconductor manufacturing. *Expert Systems with Applications*, 40(14):5688–5693.

Jeske, D. R., Stevens, N. T., Tartakovsky, A. G., and Wilson, J. D. (2018). Statistical methods for network surveillance. *Applied Stochastic Models in Business and Industry*, 34(4):425–445.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous univariate distributions*, volume 1. John Wiley and Sons.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous univariate distributions*, volume 2. John Wiley and Sons.

Jung, U. (2017). Nonparametric wavelet-based multivariate control chart for rotating machinery condition monitoring. In Tan, C. M. and Goh, T. N., editors, *Theory and Practice of Quality and Reliability Engineering in Asia Industry*, pages 41–56, Singapore. Springer Singapore.

Kankar, P. K., Sharma, S. C., and Harsha, S. P. (2011). Rolling element bearing fault diagnosis using wavelet transform. *Neurocomputing*, 74(10):1638–1645.

Kano, M., Hasebe, S., Hashimoto, I., and Ohno, H. (2002a). Statistical process monitoring based on dissimilarity of process data. *American Institute of Chemical Engineers. AIChE Journal*, 48(6):1231.

Kano, M., Nagao, K., Hasebe, S., Hashimoto, I., Ohno, H., Strauss, R., and Bakshi, B. R. (2002b). Comparison of multivariate statistical process monitoring methods with applications to the eastman challenge problem. *Computers and Chemical Engineering*, 26(2):161–174.

Keskes, H. and Braham, A. (2015). Recursive undecimated wavelet packet transform and DAG SVM for induction motor diagnosis. *IEEE Transactions on Industrial Informatics*, 11(5):1059–1066.

Keskes, H., Braham, A., and Lachiri, Z. (2013). Broken rotor bar diagnosis in induction machines through stationary wavelet packet transform and multiclass wavelet SVM. *Electric Power Systems Research*, 97:151–157.

Kini, K. R. and Madakyaru, M. (2019). Anomaly detection using multi-scale dynamic principal component analysis for tenneesse eastman process. In *2019 Fifth Indian Control Conference (ICC)*, pages 219–224.

Koosha, M., Noorossana, R., and Megahed, F. (2017). Statistical process monitoring via image data using wavelets. *Quality and Reliability Engineering International*, 33(8):2059–2073.

Kumar, R. and Singh, M. (2013). Outer race defect width measurement in taper roller bearing using discrete wavelet transform of vibration signal. *Measurement*, 46(1):537–545.

Lebaroud, A. and Clerc, G. (2008). Classification of induction machine faults by optimal time–frequency representations. *IEEE Transactions on Industrial Electronics*, 55(12):4290–4298.

Lee, D. S., Park, J. M., and Vanrolleghem, P. A. (2005). Adaptive multiscale principal component analysis for on-line monitoring of a sequencing batch reactor. *Journal of Biotechnology*, 116(2):195–210.

Lee, H. W., Lee, M. W., and Park, J. M. (2009). Multi-scale extension of pls algorithm for advanced on-line process monitoring. *Chemometrics and Intelligent Laboratory Systems*, 98(2):201–212.

Lee, J., Hur, Y., Kim, S.-H., and Wilson, J. R. (2012). Monitoring nonlinear profiles using a wavelet-based distribution-free CUSUM chart. *International Journal of Production Research*, 50(22):6574–6594.

Li, J., Jeske, D. R., Zhou, Y., and Zhang, X. (2019). A wavelet-based nonparametric cusum control chart for autocorrelated processes with applications to network surveillance. *Quality and Reliability Engineering International*, 35(2):644–658.

Liang, J., Elangovan, S., and Devotta, J. (1998). A wavelet multiresolution analysis approach to fault detection and classification in transmission lines. *International Journal of Electrical Power and Energy Systems*, 20(5):327–332.

Liu, Z., Cao, H., Chen, X., He, Z., and Shen, Z. (2013). Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings. *Neurocomputing*, 99:399–410.

Ma, S., Cai, W., Liu, W., Shang, Z., and Liu, G. (2019). A lighted deep convolutional neural network based fault diagnosis of rotating machinery. *Sensors*, 19(10):2381.

Madakyaru, M., Harrou, F., and Sun, Y. (2016). Improved anomaly detection using multi-scale pls and generalized likelihood ratio test. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–6. IEEE.

Mallat, S. (1999). *A wavelet tour of signal processing*. Academic press.

Mallat, S. and Zhong, S. (1992). Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (7):710–732.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693.

Mansouri, M., Al-Khazraji, A., Hajji, M., Harkat, M. F., Nounou, H., and Nounou, M. (2018). Wavelet optimized ewma for fault detection and application to photovoltaic systems. *Solar*

Energy, 167:125–136.

Maulud, A., Wang, D., and Romagnoli, J. (2006). A multi-scale orthogonal nonlinear strategy for multi-variate statistical process monitoring. *Journal of Process Control*, 16(7):671 – 683.

McGinnity, K., Chicken, E., and Pignatiello, J. J. (2015). Nonparametric changepoint estimation for sequential nonlinear profile monitoring. *Quality and Reliability Engineering International*, 31(1):57–73.

Megahed, F. M., Woodall, W. H., and Camelio, J. A. (2011). A review and perspective on control charting with image data. *Journal of Quality Technology*, 43(2):83–98.

Meyer, Y. (1991). Ondelettes sur l'intervalle. *Revista Matematica Iberoamericana*, 7(2):115–133.

Meyer, Y. (1993). *Wavelets-algorithms and applications*, volume 1. Society for Industrial and Applied Mathematics Translation.

Misiti, M., Misiti, Y., Oppenheim, G., Castanié, F., and Maître, H. (2003). *Les ondelettes et leurs applications*. Hermès science publications.

Misiti, M., Misiti, Y., Oppenheim, G., and Poggi, J. (1996). Wavelet toolbox. *The MathWorks Inc., Natick, MA*.

Monfared, O. A., Doroudi, A., and Darvishi, A. (2019). Diagnosis of rotor broken bars faults in squirrel cage induction motor using continuous wavelet transform. *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering*.

Nadarajah, S. and Kotz, S. (2005). On the linear combination of exponential and gamma random variables. *Entropy*, 7(2):161–171.

Nason, G. P. and Silverman, B. W. (1995). The stationary wavelet transform and some statistical applications. In *Wavelets and statistics*, pages 281–299. Springer.

Nikoo, M. and Noorossana, R. (2013). Phase ii monitoring of nonlinear profile variance using wavelet. *Quality and Reliability Engineering International*, 29(7):1081–1089.

Ogden, T. (1997). *Essential wavelets for statistical applications and data analysis*. Springer Science and Business Media.

Pahon, E., Steiner, N. Y., Jemei, S., Hissel, D., Péra, M., Wang, K., and Moçoteguy, P. (2016). Solid oxide fuel cell fault diagnosis and ageing estimation based on wavelet transform approach. *International Journal of Hydrogen Energy*, 41(31):13678–13687.

Peng, Z. and Chu, F. (2004). Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. *Mechanical systems and signal processing*, 18(2):199–221.

Rafiee, J. and Tse, P. (2009). Use of autocorrelation of wavelet coefficients for fault diagnosis. *Mechanical Systems and Signal Processing*, 23(5):1554–1572.

Ranaee, V. and Ebrahimzadeh, A. (2011). Control chart pattern recognition using a novel hybrid intelligent method. *Applied Soft Computing*, 11(2):2676–2686.

Rato, T. J. and Reis, M. S. (2015). Multiscale and megavariate monitoring of the process networked structure: M2NET. *Journal of Chemometrics*, 29(5):309–322.

Reis, M. S. and Saraiva, P. M. (2006). Multiscale statistical process control with multiresolution data. *AIChE journal*, 52(6):2107–2119.

Roodbali, M. S. E. and Shahbazian, M. (2011). Multi-scale PLS modeling for industrial process monitoring. *International Journal of Computer Applications*, 26(6):26–33. Published by Foundation of Computer Science, New York, USA.

Sheriff, M. Z., Mansouri, M., Karim, M. N., Nounou, H., and Nounou, M. (2017). Fault detection using multiscale PCA-based moving window GLRT. *Journal of Process Control*, 54:47 – 64.

Strang, G. and Nguyen, T. (1996). *Wavelets and filter banks*. SIAM.

Sun, H., Zi, Y., and He, Z. (2014). Wind turbine fault detection using multiwavelet denoising with the data-driven block threshold. *Applied Acoustics*, 77:122–129.

Teppola, P. and Minkkinen, P. (2000). Wavelet-PLS regression models for both exploratory data analysis and process monitoring. *Journal of Chemometrics*, 14(5-6):383–399.

Vannucci, M. and Corradi, F. (1999). Covariance structure of wavelet coefficients: theory and models in a Bayesian perspective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(4):971–986.

Wang, C.-H. and Kuo, W. (2007). Identification of control chart patterns using wavelet filtering and robust fuzzy clustering. *Journal of Intelligent Manufacturing*, 18(3):343–350.

Wang, H., Kim, S.-H., Huo, X., Hur, Y., and Wilson, J. R. (2015). Monitoring nonlinear profiles adaptively with a wavelet-based distribution-free CUSUM chart. *International Journal of Production Research*, 53(15):4648–4667.

Wang, R. (2012). *Introduction to orthogonal transforms: with applications in data processing and analysis*. Cambridge University Press.

Wang, Z., Bukkapatnam, S. T., Kumara, S. R., Kong, Z., and Katz, Z. (2014). Change detection in precision manufacturing processes under transient conditions. *CIRP Annals*, 63(1):449–452.

Woodall, W. H. and Montgomery, D. C. (2014). Some current directions in the theory and application of statistical process monitoring. *Journal of Quality Technology*, 46(1):78.

Wu, J.-D. and Chen, J.-C. (2006). Continuous wavelet transform technique for fault signal diagnosis of internal combustion engines. *NDT e International*, 39(4):304–311.

Yan, R. and Gao, R. X. (2009). Base wavelet selection for bearing vibration signal analysis. *International Journal of Wavelets, Multiresolution and Information Processing*, 7(04):411–426.

Yan, R., Gao, R. X., and Chen, X. (2014). Wavelets for fault diagnosis of rotary machines: A review with applications. *Signal Processing*, 96:1–15.

Yin, S., Ding, S. X., Xie, X., and Luo, H. (2014). A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics*, 61(11):6418–6428.

Yin, Z. and Hou, J. (2016). Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes. *Neurocomputing*, 174:643–650.

Yoon, S. and MacGregor, J. F. (2004). Principal-component analysis of multiscale data for process monitoring and fault diagnosis. *AIChE Journal*, 50(11):2891–2903.

Zarei, J. and Poshtan, J. (2007). Bearing fault detection using wavelet packet transform of induction motor stator current. *Tribology International*, 40(5):763–769.

Zgarni, S., Abid, F. B., and Braham, A. (2017). Artificial immune network for bearing fault detection of induction motor. In *5th International Conference on Control & Signal Processing (CSP-2017)*, volume 25, pages 17–20.

Zhang, J., Zhang, Y., and Guan, Y. (2016). Analysis of time-domain reflectometry combined with wavelet transform for fault detection in aircraft shielded cables. *IEEE Sensors Journal*, 16(11):4579–4586.

Zhang, Y. and Hu, Z. (2011). Multivariate process monitoring and analysis based on multi-scale KPLS. *Chemical Engineering Research and Design*, 89(12):2667–2678.

Zuo, L., He, Z., and Zhang, M. (2019). An ewma and region growing based control chart for monitoring image data. *Quality Technology & Quantitative Management*, pages 1–16.