Wavelet-based feature extraction using continuous wavelet transform for power quality disturbances classification

Fabrício Ely Gossler [1]*, **Marco Aparecido Queiroz Duarte** [1]

[1] *State University of Mato Grosso do Sul (UEMS), MS 306 Road, km 06, 79543-899, Cassilândia, MS, Brazil.*
** Corresponding author. E-mail: fabricio.ely@uems.br*

ABSTRACT. Wavelet analysis is well established as an effective signal processing tool for detecting and analysing power quality disturbances. Several studies have employed the fast wavelet transform to construct feature vectors used in the power quality analysis. This paper presents a new wavelet-based feature extraction method for power quality disturbances classification using the continuous wavelet transform. A set of small scales is used for this purpose, and only the energy of each decomposed signal is used to construct the feature vector for classification purposes. The chosen scales are those representing the natural frequency and its harmonics. The proposed method is applied to voltage event signals generated by a frequently used mathematical model in the specialized literature for power quality disturbances analysis. Experiments with both clean and noisy signals demonstrate that classification accuracy approaching 100% can be achieved using only two scales of the continuous wavelet transform. Some wavelets are tested, and those with an appropriate number of vanishing moments presents better results. The results are compared with those obtained from other wavelet-based methods employing the fast wavelet transform. The effectiveness of the proposed wavelet-based feature extraction method is evaluated with several classifiers, consistently yielding high accuracy.

Key words: p*ower quality, disturbance classification, voltage events, feature extraction, continuous wavelet transform.*

DOI: https://doi.org/10.33837/msj.v8i1.1732

Received: July 25th, 2025. Accepted December 11st, 2025. Published: December 23rd, 2025.
Associate editor: Herbert J. Dias

INTRODUCTION

Power quality disturbances (PQD) are among the most frequent phenomena in electrical power systems. The increasing presence of non-linear loads and modern control equipment has raised the demand for high power quality (PQ) in generation, transmission, and distribution systems. PQDs can cause substantial losses across various sectors, particularly in large public or industrial facilities. Disturbances affect the interaction between producers and consumers, leading to energy inefficiency, constrained generation and consumption, malfunction or damage of sensitive equipment, and misoperation of industrial processes (Chawda et al., 2020).

Voltage events represent one of the most significant types of PQD. Based on magnitude, voltage events are classified as swells, sags, or interruptions; according to their duration, they are categorized as instantaneous, momentary, or temporary (IEEE Std 1159-2009). Harmonic distortion is another voltage-related disturbance that produces waveform deviations. PQ analysis from voltage waveform monitoring involves two main stages: detection and classification. Detection identifies the occurrence and location of voltage deviations, while classification determines the type of disturbance.

Several signal-processing techniques have been employed for PQD detection and classification (Caicedo et al., 2023), including the Fourier Transform (FT), the Short-Time Fourier Transform (STFT), and the Wavelet Transform (WT). WT provides simultaneous time–frequency localization, enabling multiresolution analysis and a more precise characterization of PQD signals than FT or STFT. It offers high time–frequency resolution for transient detection, efficient signal decomposition, and adaptability to signal features. However, it is computationally demanding for high-resolution or long-duration data, and its performance strongly depends on the selected wavelet, making wavelet choice a critical design parameter.

Signal processing techniques for PQ analysis are applied at multiple stages: data preparation, preprocessing, transformation, feature extraction,

detection, classification, and characterization. Feature extraction is especially critical, as it condenses the original signal into a lower-dimensional representation, while retaining relevant information. Effective feature extraction allows accurate classification of disturbances, since even the most advanced classifiers fail if features are poorly selected (Uyara et al., 2008; Markovska and Taskovski, 2018; Sengur et al., 2007).

Several WT-based feature extraction methods have been proposed. Early applications relied on visual inspection of the time-frequency plane for PQ analysis (Santoso et al., 1994, 1996; Pillay and Bhattacharjee, 1996). Barros et al. (2012) reviewed WT applications in voltage event detection and identified three main challenges: wavelet selection, efficient feature extraction, and reduction of computational time. Although previous studies have not fully overcome these challenges, several strategies have been proposed, such as selecting wavelets that minimize computational cost (He and Starzyk, 2006).

The Fast Wavelet Transform (FWT) has been widely used to extract wavelet features for PQD classification (Abdel-Galil et al., 2004; He and Starzyk, 2006; Uyara et al., 2008; Markovska and Taškovski, 2018). FWT, the algorithm for computing the discrete wavelet transform (DWT), is defined as a two-channel filter bank with downsampling (Mallat, 1989). In PQ analysis, FWT is generally preferred over the continuous wavelet transform (CWT) because it reduces sample sets at each decomposition level, while CWT, based on inner products, is redundant and computationally heavier for multiple scales (Grossmann and Morlet, 1984). Nevertheless, efficient CWT-based PQD analysis methods have been developed, including scalogram generation for classification (Huang et al., 1998; Angrisani et al., 2001; Lin and Wang, 2006; Tse et al., 2010; Salles and Ribeiro, 2023).

The algebraic model for seven voltage events introduced by Abdel-Galil et al. (2004) has served as a foundation for many PQ analysis methods combined with FWT (He and Starzyk, 2006; Uyara et al., 2008; Markovska and Taškovski, 2018 Eristi, Uçar and Demir, 2010; Decanini et al., 2011; Saini and Beniwal, 2017). In these works, features are typically constructed from the energy of wavelet coefficients at each decomposition level, and classification is performed using methods such as decision trees, support vector machines, and random forests (He and Starzyk, 2006; Uyara et al., 2008; Markovska and Taškovski, 2018).

Recently, hybrid approaches combining wavelet analysis with machine learning and deep learning have achieved near-perfect PQD classification. Examples include empirical mode decomposition with optimized extreme learning machines (Samanta et al., 2022), visual attention mechanisms with feed-forward neural networks (Zhang et al., 2022), integer factor approximation methods (Akmaz, 2022), adaptive empirical wavelet transform (Khetarpal et al., 2024), convolutional and recurrent neural networks (Cai et al., 2023), improved adaptive S-transform (Jiang et al., 2024a), and multi-channel feature fusion with CWT (Jiang et al., 2024b). These studies highlight the evolution of PQD classification and motivate the development of efficient, interpretable, and computationally light wavelet-based techniques.

Although effective, FWT-based methods present limitations such as trade-offs between decomposition depth and feature resolution, and mismatches between dyadic scales and signal harmonics. This work introduces a CWT-based feature extraction method for PQD classification that mitigates these issues. By selecting a minimal set of scales corresponding to physically meaningful frequencies, the method achieves high classification accuracy with low computational cost. Wavelet representations are constructed from this optimized scale set using the CWT scale–frequency relation, decoupling feature extraction from the decomposition-level constraints of FWT. Feature vectors are derived from the energy of CWT coefficients, attaining near-100% classification accuracy with only two scales when a suitable number of vanishing moments is used. The approach is validated on noisy signals and benchmarked against prior studies (Abdel-Galil et al., 2004; He and Starzyk, 2006; Uyara et al., 2008; Markovska and Taškovski, 2018), employing multiple classifiers to evaluate robustness.

CONTINUOUS WAVELET TRANSFORM

Wavelet analysis can reveal various signal characteristics, such as trends, singularities, discontinuities in higher derivatives, and self-similarity. Similar to the Fourier analysis, the wavelet analysis involves the expansion of signals into a set of basic functions. A family of functions of the form

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right), \tag{1}$$

generated from a single finite-energy function $\psi(t)$ in the time domain by dilations $(s > 0)$ and translation τ operations, has been widely used in various signal-processing applications. These functions are called wavelets and satisfy the admissibility condition given by:

$$c_\psi = \int_0^{+\infty} \frac{|\Psi(\Omega)|^2}{\Omega} \mathrm{d}\Omega < +\infty, \tag{2}$$

where $\Psi(\Omega)$ represents the wavelet function FT (Calderon, 1964; Grossmann and Morlet, 1984).

From (1), the CWT of a continuous-time finite-energy signal $x(t)$ is given as (Grossmann and Morlet, 1984):

$$W_x^{\psi}(s,\tau) = \int_{-\infty}^{+\infty} x(t)\psi_{s,\tau}(t)dt, \quad (3)$$

where $\psi(t)$ is a real wavelet function. The CWT coefficients $W_x^{\psi}(s,\tau)$ measure signal fluctuations around position τ and in the frequency band corresponding to the scale factor s. When s increases, $\psi(t)$ is expanded, and its frequency content moves to the lower frequency bands. Decreasing s implies the compression of $\psi(t)$, and its frequency content moves to the higher bands.

The information concerning the frequency content in Hz of the signal is provided by a relative frequency f_s estimated from

$$f_s = \frac{f_c}{s\,\Delta}, \quad (4)$$

where f_c is the wavelet center frequency (in Hz), Δ is the signal sampling period (in seconds) and s denotes the decomposition scale (Abry, 1997). This relationship is fundamental for designing our feature extraction method, since it allows us to select scales that directly correspond to specific frequency components of interest in the power signal, such as the fundamental frequency and its harmonics.

An important property of wavelets is the number of vanishing moments, which quantifies their ability to measure the local regularity of a signal. A $\psi(t)$ wavelet has N vanishing moments if condition (5) is verified:

$$\int_{-\infty}^{+\infty} t^k \psi(t)\,dt, \qquad k = 0,1,\ldots,N-1, \quad (5)$$

from which follows that any wavelet has at least one vanishing moment (Meyer, 1993). A wavelet with N vanishing moments is orthogonal to polynomials of degree up to $N-1$. It means it is sensitive to signal features of high-order differentiability and can compactly represent signals with sharp transitions or singularities. For PQD classification, a small number of vanishing moments is necessary to accurately capture and represent the transient characteristics of disturbances.

The behavior of the CWT coefficients depends on the chosen wavelet. Each wavelet can also be characterized by its symmetry and support size. Another important wavelet characteristic refers to orthogonality, see (Daubechies, 1988). Several types of wavelet functions are commonly applied in CWT-based analyses. The most widely used families in the specialized literature include Morlet, Meyer, Gaussian, Mexican Hat, Haar, Daubechies, Symlets, and Coiflets (Daubechies, 1988). It is common to use short names for some wavelets. For example, db*N* and sym*N* are the short name for Daubechies and Symlets wavelets, respectively. In these cases, *N* indicates the number of vanishing moments and a $2N-1$ is the support length. Generally, Coiflets wavelets are denoted by coif*N*, where, in this case, 2*N* is the number of vanishing moments and the support length is $6N-1$. In order to facilitate the notation, gaus*N* will be used to indicate Gaussian wavelet with *N* vanishing moments.

PROPOSED METHODOLOGY FOR WAVELET FEATURE EXTRACTION

The construction of feature vectors is the fundamental process in PQD classification Seven voltage events are considered, based on the mathematical model of (Abdel-Galil et al., 2004): normal (C1), swell (C2), sag (C3), outage (C4), harmonic (C5), swell with harmonic (C6) and sag with harmonic (C7). Normal is an undisturbed sinusoidal condition. S*well* corresponds to a temporary increase in voltage magnitude, while *sag* represents a temporary decrease in voltage. An *outage* refers to the complete loss of voltage supply, and *harmonic distortion* denotes waveform deformation due to the presence of harmonic components. The combined cases — *swell with harmonic* and *sag with harmonic* — involve simultaneous variations in voltage magnitude and waveform distortion. Figure 1 illustrates the waveforms corresponding to normal operation and the disturbances investigated in this work.

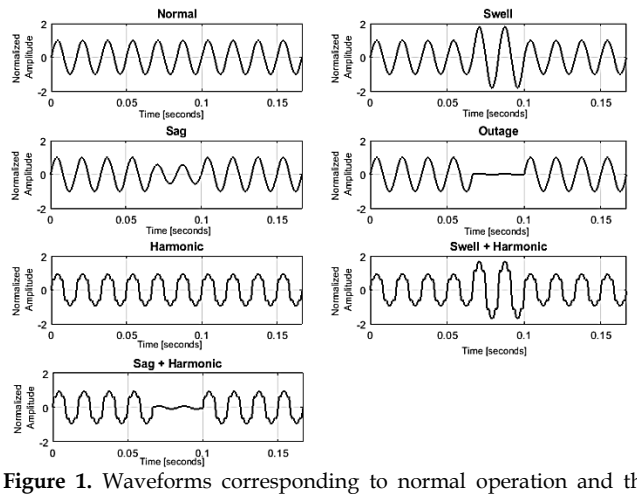

Figure 1. Waveforms corresponding to normal operation and the disturbances investigated in this work.

Figure 2 shows absolute normalized CWT coefficients for each particular voltage event signal class evaluated by Morlet wavelet for $1 \leq s \leq 170$. Note that, for the considered scales, the energy of the coefficients is more concentrate in scales $s > 120$. Such scales represent, in certain

sense, the natural frequency component oscillatory of the signals. For voltage signals with harmonics, it is possible to note a small energy contribution at intermediary scales (the interval $30 < s < 80$).

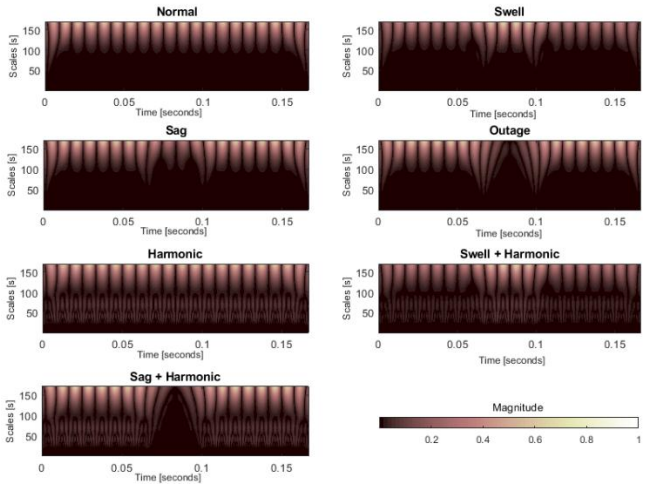

Figure 2. Scalograms for each PQD considered using the Morlet wavelet for $1 \leq s \leq 170$, where it is possible to verify the concentration for energy according to the scales.

Distinct patterns can be qualitatively observed for each voltage event in the scalograms shown in Figure 2. For this reason, we can explore such differences to construct wavelet feature vectors to represent each disturbance. In fact, a drawback in the use of CWT is that it can increase the computational cost when many scales are considered in the decomposition process. However, since CWT is a redundant transform, wavelet feature vectors can be constructed using an appropriately limited set of scales. To illustrate wavelet representations at specific scales, Figure 3 presents the CWT coefficients at $s_1 = 208$, $s_2 = 69.33$ and $s_3 = 41.60$. These scales were obtained using the scale-frequency relation described in (4), and they represent the natural frequency (s_1) and its two first harmonics (s_2 and s_3, respectively). These scales highlight the main differences among the PQD signals.

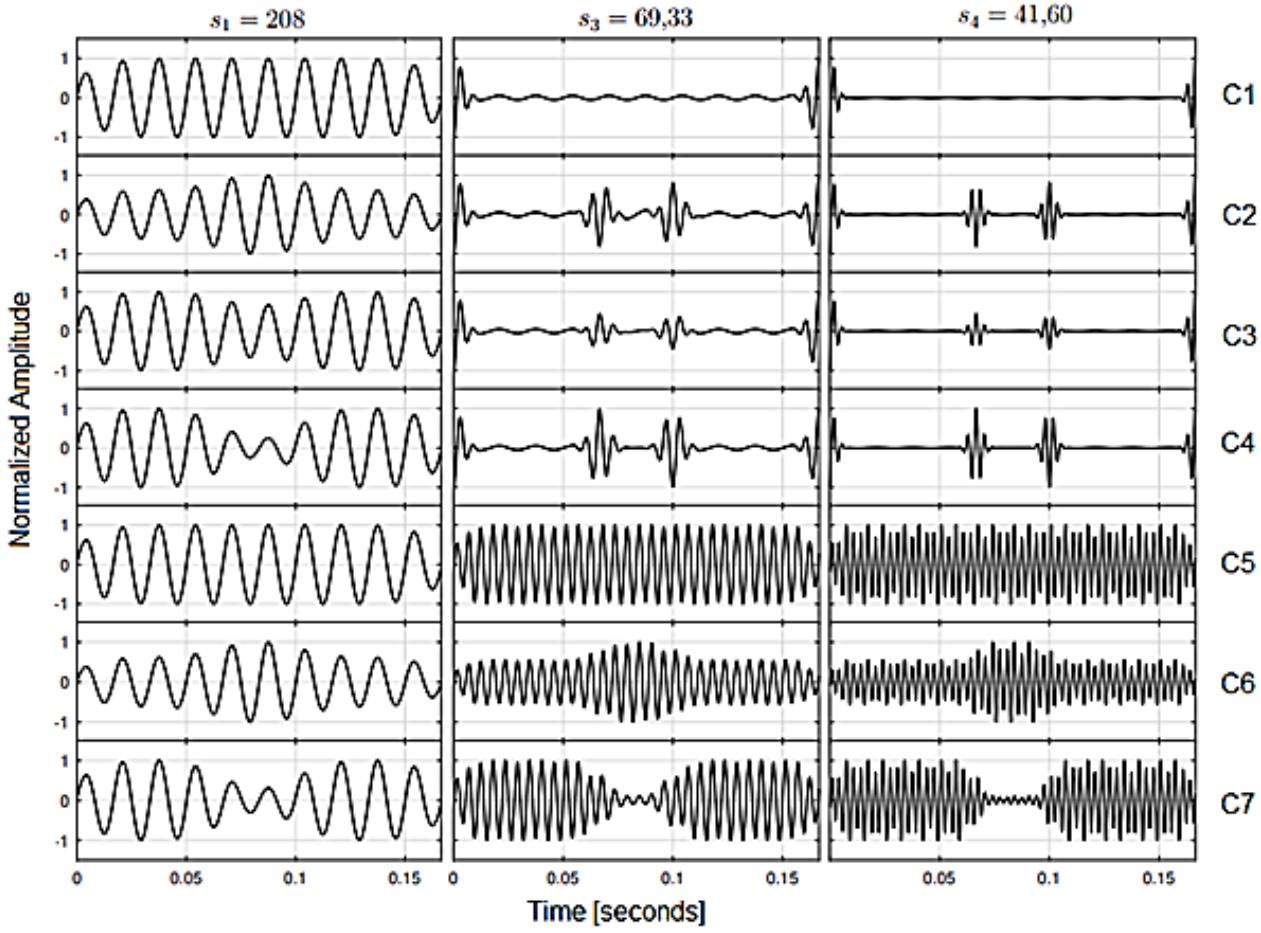

Figure 3. PQD CWT using Morlet wavelet at scales $s_1 = 208$ (left column), $s_2 = 69.33$ (middle column) and $s_3 = 41.60$ (right column). Each row plot represents PQD CWT coefficients, normalized for a better comparasion.

Inspection of the CWT-coefficient plots in Figure 3 reveals clear patterns. At scale s_1, corresponding to the fundamental oscillatory frequency, the wavelet representations of C1, C2, C3, and C4 closely resemble the time-domain waveform. The representations of C5, C6, and C7 are similar to C1, C2, and C4, respectively. Conversely, at scale s_2, larger differences appear in PQ signals containing harmonic components. For $W_x^{\psi}(s_3, \tau)$ results, the behaviour is similar to those presented by $W_x^{\psi}(s_2, \tau)$.

The rationale behind the scale-selection strategy is to shift from a dense, computationally expensive analysis across all scales to a targeted evaluation at physically meaningful frequencies. PQD are deviations in the fundamental sinusoidal waveform (e.g., magnitude changes, transients, and additional harmonic components). Therefore, the most discriminative information for classifying these events is likely concentrated around the fundamental frequency of the power system (e.g., 60 Hz) and its characteristic harmonics (e.g., 120 Hz, 180 Hz, 240 Hz, etc.). By focusing the CWT analysis specifically on the scales that correspond to these key frequencies, we efficiently capture the essential features of the disturbance while drastically reducing the data dimensionality and computational burden.

Based on these considerations, an appropriate set of scales is selected for computing the CWT in the proposed wavelet feature extraction approach. The scales chosen are those that represent the natural frequency and its harmonics. In this way, the set of scales

$$\varphi = \{s_1, s_2, \dots, s_{h+1}\}, \tag{6}$$

where h is the number of harmonics (dimensionless) used to build the wavelet feature extraction, is determined using the scale-frequency relation:

$$s_k = \frac{f_c}{k\,\Delta\,f_N}, \qquad k = 1,2,\dots,h+1, \tag{7}$$

f_c is the wavelet center frequency (in Hz), T is the sampling period (in seconds), f_N is the natural frequency (60 Hz), k is the harmonic order (dimensionless) and s_k is the scale factor (dimensionless) corresponding to the k-th harmonic component.

PQD energy is typically concentrated near the fundamental and harmonic frequencies, these scales provide the most discriminative information for classification. By focusing the CWT analysis on these pre-determined and physically meaningful scales, ensuring energy concentration in the resulting coefficients, there is a significant reduction on computational cost compared to a full-scale decomposition, with the enhancement of the discriminative power of the extracted features for classification.

Among several possible metrics to characterize the wavelet coefficients, this work employs the energy of the CWT coefficients. In (Abdel-Galil et al., 2004) and (He and Starzyk, 2006) only the energy of wavelet coefficients was also used. Thus, with the obtained CWT coefficients from each $x(t)$ PQ signal, $E_x(s_k)$ energy is computed at each scaling decomposition:

$$E_x(s_k) = \sum_{\tau} \left| W_x^{\psi}(s_k, \tau) \right|^2. \tag{8}$$

After the energy vector computation, $E_x(s_k)$ is normalized by $E_{xn}(s_1)$ which represents the normal operation energy vector. In this way, the proposed wavelet feature extraction is defined as (8):

$$F_x = [E_x(s_1)E_x(s_2)E_x(s_3)\ \dots.\ E_x(s_{h+1})] \,/E_{xn}(s_1) \tag{9}$$

Wavelet feature vector normalization is a common strategy in the PQD classification process. For example, in (Abdel-Galil et al., 2004) the wavelet feature vector is normalized by a coefficient that also represents the normal operation energy. The complete process is illustrated in the flowchart of Figure 4 and consists of the following steps:

Step 1 - Signal Acquisition and Segmentation: The PQD signal $x(t)$ is acquired with a sampling period Δ, corresponding to the fundamental power frequency f_N. The signal must be segmented into multiple complete cycles of the fundamental frequency. The specific number of cycles can be adjusted based on the application requirements, trading off between temporal resolution and feature capture capability.

Step 2 - Wavelet Selection: The process begins with the selection of an appropriate mother wavelet $\psi(t)$.

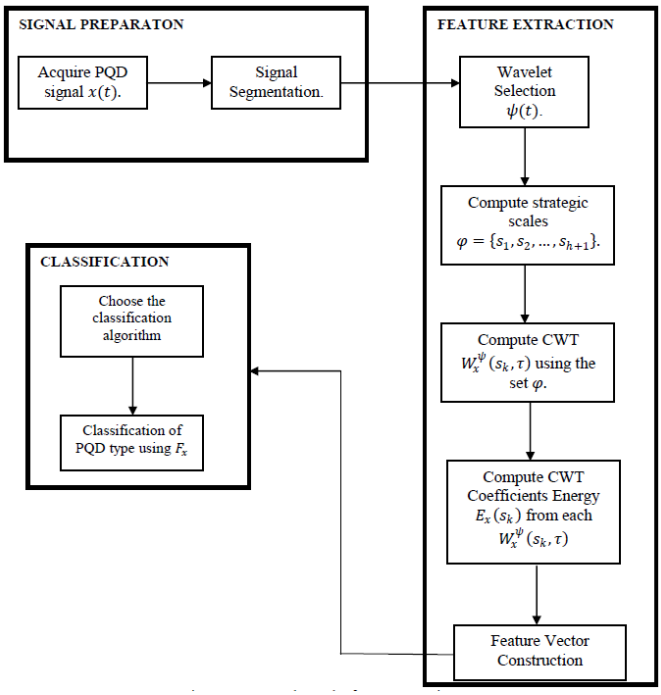

Figure 4. Proposed wavelet feature extraction.

Step 3 - Scale Selection: The scale set φ is determined using (7). This approach ensures that we capture the fundamental frequency component (s_1) and its most relevant harmonics (s_2,s_3,…).

Step 4 - CWT Coefficient Computation: For each scale s_k in the predetermined set φ, we compute the CWT coefficients $W_x^{\psi}(s_k, \tau)$. It generates a set of time-scale representations that highlights the signal's behavior at the specific frequency bands of interest, avoiding the computational burden of a full-scale analysis.

Step 5 - CWT coefficients Energy: From the CWT coefficients $W_x^{\psi}(s_k, \tau)$, the energy content is computed to obtain $E_x(s_1), E_x(s_2), E_x(s_3), \cdots, \mathrm{E_x(s_{h+1})}$. Such valuesquantify the signal strength distributed across the fundamental frequency and its harmonic components.

Step 6 - Feature Vector Construction: The feature vector F_x is assembled using the energy values from all selected scales and normalized by the energy at the fundamental scale $E_{xn}(s_1)$ obtained from a reference normal operating condition signal. This normalization step enhances robustness against system-wide amplitude variations and facilitates meaningful comparison across different disturbance types.

Step 7 – Classification: The final normalized feature vector F_x serves as input to the classification algorithm, which identifies the specific type of PQD present in the original signal.

EXPERIMENTAL AND COMPARATIVE CLASSIFICATION RESULTS

Before any PQD classification process, it is necessary to detect such disturbances in the voltage systems. In this paper, the proposed methodology is applied in the feature extraction stage and, therefore, it is assumed that the PQD detection has already been achieved.

4.1. Data generation

Signal datasets were generated following the mathematical models in Abdel-Galil et al. (2004). For each PQ class, 200 cases were produced for training and 200 for testing. Table 1 lists the signal models and their control parameters ($u(t)$ is the unit step function). Signals were sampled at 256 samples per cycle with a nominal frequency of 60 Hz; 10 cycles were used per signal, yielding 2560 samples (sampling frequency = 2.56 kHz). Parameter ranges for amplitude and duration ($\alpha, \beta, \gamma, \alpha_1$, etc.) were varied across realistic intervals to produce representative variability consistent with the cited literature. The chosen discretization follows prior studies (Abdel-Galil et al., 2004; He and Starzyk, 2006; Uyara et al., 2008; Markovska and Taškovski, 2018). Note that earlier works sometimes use different nominal frequencies (50 Hz in He and Starzyk, Uyara et al., and Markovska and Taškovski) or different cycle counts (16 cycles in Uyara et al., 2008).

To analyze classification results from smaller wavelet feature vectors, four different types of wavelet feature vectors will be explored:

- $F1_x = [E_x(s_1)]/E_{xn}(s_1)$;
- $F2_x = [E_x(s_1) \mid E_x(s_2)]/E_{xn}(s_1)$;
- $F3_x = [E_x(s_1) \mid E_x(s_2) \mid E_x(s_3)]/E_{xn}(s_1)$;
- $F4_x = [E_x(s_1) \mid E_x(s_2) \mid E_x(s_3) \mid E_x(s_4)]/E_{xn}(s_1)$.

$F1_x$ is built considering only CWT decomposition at scale s_1, which corresponds to f_N wavelet representation. $F2_x$ is composed by the CWT coefficients energy from scales s_1 and s_2, i.e., considering the first harmonic. The wavelet features $F3_x$ and $F4_x$ are respectively referent to the set of scales $\{s_1,s_2,s_3\}$ and $\{s_1,s_2,s_3,s_4\}$. In this way, we consider a scenario involving up to the third harmonic. Therefore, it will be possible to determine what set of scales implies efficient results, considering precision and computational cost factors. Some of the wavelet families discussed in Section 2 will be used to decompose each type of PQ class (C1 to C7). Therefore, different wavelets will be considered in order to determine those with the best results.

Table 1. Mathematical definitions for the used PQ disturbances. All time parameters $t, t_1, and\ t_2$ are in seconds (s). Frequency parameters f_N and ω are in Hz and rad/s, respectively. Amplitude parameters (α, β, γ, α_1, etc.) are dimensionless ratios relative to the nominal voltage

PQ Disturbance	Model	Controlling Parameters
Normal (C1)	$v_1(t) = sin(\omega t)$	$f_N = 60, \quad \omega = 2\pi f_N$
Swell (C2)	$v_2(t) = \left[1 + \alpha\left(u(t-t_1) - u(t-t_2)\right)\right]v_1(t)$	$0.1 \leq \alpha \leq 0.8$
Sag (C3)	$v_3(t) = \left[1 - \beta\left(u(t-t_1) - u(t-t_2)\right)\right]v_1(t)$	$0.1 \leq \beta \leq 0.9$
Outage (C4)	$v_4(t) = \left[1 - \gamma\left(u(t-t_1) - u(t-t_2)\right)\right]v_1(t)$	$0.9 < \gamma \leq 1.0$
Harmonic (C5)	$v_5(t) = \alpha_1 sin(\omega t) + \alpha_3 sin(3\omega t) + \alpha_5 sin(5\omega t) + \alpha_7 sin(7\omega t)$	$0.05 \leq \alpha_3, \alpha_5, \alpha_7 \leq 0.15$
Swell + Harmonic (C6)	$v_6(t) = v_2(t)v_5(t)$	$\sum \alpha_i^2 = 1$
Sag + Harmonic (C7)	$v_7(t) = v_3(t)v_5(t)$	$T \leq t_2 - t_1 \leq 9T$

.

4.2 Classification results using Decision tree

A decision-tree classifier (C4.5, as in Abdel-Galil et al. (2004)) was employed for the initial analysis. Table 2 reports overall accuracies obtained with C4.5 across the different wavelet families and feature vector types ($F1_x$ – $F4_x$). A 10-fold cross-validation protocol was used to ensure robustness; statistical indicators (mean, standard deviation, and 95% confidence intervals) were computed across repeated runs to assess significance.

Table 2. Correct classification results in terms of percentage using different wavelet families. Bold values indicate accuracies higher than 99%

Wavelet	$F1_x$	$F2_x$	$F3_x$	$F4_x$	Wavelet	$F1_x$	$F2_x$	$F3_x$	$F4_x$
db2	70.86	93.64	**99.29**	**99.00**	sym2	70.86	93.64	**99.29**	**99.00**
db3	71.14	90.29	**99.43**	**99.36**	sym3	71.14	90.29	**99.43**	**99.36**
db4	70.57	97.64	**99.29**	**99.36**	sym4	70.50	96.79	**99.21**	**99.29**
db5	70.93	**99.43**	**99.29**	**99.29**	sym5	70.93	**99.29**	**99.29**	**99.29**
db6	70.64	97.98	98.86	98.86	sym6	70.71	97.92	98.93	98.92
db7	71.07	**99.14**	96.43	96.43	sym7	70.93	98.50	**99.21**	**99.29**
db8	70.78	**99.36**	**99.36**	**99.36**	sym8	70.64	**99.14**	**99.14**	**99.14**
db9	70.82	**99.43**	**99.36**	**99.36**	sym9	70.71	**99.36**	**99.29**	**99.29**
db10	70.71	**99.36**	**99.36**	**99.36**	sym10	70.71	**99.29**	**99.29**	**99.29**
db20	70.71	**99.36**	**99.36**	**99.43**	sym20	70.43	**99.36**	**99.36**	**99.43**
coif1	71.07	84.29	**99.21**	**99.36**	gaus1	71.64	90.07	97.43	98.14
coif2	70.50	96.50	**99.29**	**99.29**	gaus2	70.71	92.36	**99.07**	**99.14**
coif3	71.00	98.50	98.50	98.57	gaus3	70.93	93.79	**99.43**	**99.43**
coif4	70.86	**99.21**	98.93	98.93	gaus4	70.21	92.79	**99.14**	**99.14**
coif5	70.79	**99.14**	**99.14**	**99.14**	gaus5	69.79	98.79	99.00	99.00
Haar	71.07	73.21	92.36	97.29	gaus6	70.36	96.43	98.79	98.79
Meyer	71.36	**99.14**	**99.14**	**99.14**	gaus7	70.14	**99.43**	**99.57**	**99.57**
Morlet	69.21	**99.21**	**99.21**	**99.43**	gaus8	69.64	**99.43**	**99.43**	**99.43**

Independently of wavelet function, the classification results obtained from $F1_x$ are always close to 70%. On the other hand, except for Haar and coif1 wavelets, $F2_x$ presents better results with accuracy superior to 90%. In fact, the precision obtained from $F2_x$ is very higher for wavelets with greater number of vanishing moments. In this case, except for sym7, wavelets containing more than 6 vanishing moments presented accuracy superior to 99%. However, as it can be seen in Table 2, db5 and sym5 wavelets also achieve such precision. For the $F3_x$ and $F4_x$ feature vectors, most wavelets achieved accuracies above 99%, except for db7, which showed a noticeable decrease. A general improvement is observed for wavelets with fewer vanishing moments. Nevertheless, wavelets with exactly six vanishing moments (db6, sym6, and coif3) and those with only one vanishing moment (Haar and gaus1) did not exceed 99% accuracy. The coif4 wavelet also exhibited slightly lower performance under these configurations.

From Table 2, it can be noted that gaus7 presents the better overall accuracy, 99.57%, using F4x wavelet feature. However, it was also possible to obtain a precision very close to gaus7 (accuracy equal to 99.43%), using $F2_x$ and db5, db9, gaus7 and gaus8 wavelets. It suggests that $F2_x$ wavelet feature extraction is the most efficient for PQD classification, since it implies more reduction in computational cost than $F3_x$ and $F4_x$.

In addition to the results presented in Table 2, a detailed statistical evaluation of the classifier performance was conducted for the best-performing wavelets. Table 3 summarizes the metrics extracted directly from the results, including mean accuracy, Kappa statistic, absolute and relative errors, and root mean squared errors. The experiments were performed using *Cross-validation* mode (10 folds) to ensure reproducibility and robustness of the models. Except for $F1_x$, the proposed wavelet feature vectors achieved accuracies greater than 99%, with Kappa values exceeding 0.99, indicating excellent agreement between predicted and true classes. Mean absolute errors (MAE) and Root Mean Squared Error (RMSE) values were below 0.004 and 0.045, respectively, for the best-performing wavelets (db5, gaus7, gaus8), showing low variance and high reliability. The Relative Absolute Error (RAE) and Root Relative Squared Error (RRSE) (RAE < 1.6%, RRSE < 13%) further confirm model consistency.

Reported execution times were negligible (less than 1 second per model on the test platform). Paired t-tests ($p > 0.05$) indicate no statistically significant difference among the top three wavelets (db5, gaus7, gaus8), corroborating the robustness of the approach. In summary, the inclusion of statistical indicators, cross-validation, and parameter sensitivity confirms the stability and generalization capability of the proposed wavelet-based feature extraction methodology.

4.3 Performance comparisons

Table 4 compares the proposed method (db5 with F2x) against four prior FWT-based approaches (Abdel-Galil et al., 2004; He and Starzyk, 2006; Uyara et al., 2008; Markovska and Taškovski, 2018). Both per-class and overall accuracies are reported. The proposed CWT-based feature extraction attains higher overall accuracy (99.43%) than the compared methods and superior or comparable per-class accuracies across most disturbance types. Figure 5 shows the performance comparison charts from the results in Table 4.

Decomposition depth influences computational cost: deeper decompositions require more operations. For example, Abdel-Galil et al. (2004) used Haar wavelets with 11 decomposition levels, while He and Starzyk and Uyara et al. used fewer levels. In contrast, the proposed approach attains high accuracy using only two selected CWT scales, resulting in a substantially smaller feature vector. Thus, the proposed wavelet feature vector has a smaller size than those presented in the compared works.

Table 3. Statistical results for the best-performing wavelets

Wavelet	Feature Vector	Mean Accuracy (%)	Kappa	MAE	RMSE	RAE (%)	RRSE (%)
Meyer	$F1_x$	71.79	0.6708	0.0888	0.2107	36.26	60.22
db5	$F2_x$	99.64	0.9958	0.0020	0.0317	0.82	9.05
gaus7	$F2_x$	99.29	0.9917	0.0039	0.0439	1.58	12.55
gaus8	$F2_x$	99.43	0.9933	0.0031	0.0396	1.28	11.32
gaus7	$F3_x$	99.64	0.9958	0.0020	0.0315	0.81	8.99
gaus7	$F4_x$	99.64	0.9958	0.0020	0.0315	0.81	8.99

Table 4. Performance comparison of correct classification results in terms of percentage. Bold values indicate the highest accuracy for each method and disturbance type.

PQ class	Method 1: Abdel-Galil et al., 2004	Method 2: He and Starzyk, 2006	Method 3: Uyara et al., 2008	Method 4: Markovska and Taškovski, 2018	Proposed Method
C1	**100**	**100**	**100**	**100**	**100**
C2	97	**100**	96.50	94.50	**100**
C3	76.50	87	88	**100**	98.50
C4	90	80.50	85.50	**100**	98
C5	**100**	**100**	**100**	99	**100**
C6	98	**100**	**100**	**100**	99.50
C7	71.50	97	**100**	99	**100**
Overall	90.4	94.93	95.71	98.93	**99.43**

For FWT applications, the wavelet filter length increases as the wavelet's number of vanishing moments also increases. Wavelets with higher support size imply more computational cost. Different wavelet filters were also used in (He and Starzyk, 2006) and (Uyara et al., 2008) to test the classification systems performance. In (He and Starzyk, 2006), it was shown that there is no statistically significant difference in PQ classification performance for different wavelet filters. Thus, in this case, Haar wavelet filter is best suited for small computational cost. In (Uyara et al., 2008), biorthogonal wavelet filters presented slightly better accuracy than db4. Nevertheless, authors indicated db4 or db5 wavelet filters since they also have advantages in terms of computational cost. Daubechies wavelet filters were also used in (Markovska and Taskovski, 2018). As seen in Table 2, among the wavelets with best results, db5 and sym5 have fewer vanishing moments and consequently smaller support sizes. Therefore, according to smaller computational cost and higher accuracy, db5 and sym5 wavelets are more indicated to build the proposed wavelet feature vector.

4.4. Classification results under different noise conditions

To assess noise robustness, additive white Gaussian noise was applied to the PQ signals at SNR levels of 20, 30, 40 and 50 dB. Table 5 reports overall accuracies for db5 and sym5 using the $F2_x$ feature; results from Uyara et al. (2008) are included for comparison. As expected, classification accuracy increases with SNR: even at 20 dB, db5 and sym5 achieved accuracies above 96%. In (He and Starzyk, 2006), PQ noisy signals were also considered and the classification accuracy was qualitatively presented. In such paper, the accuracy varies from 90% to 94% for SNR values from 20 dB to 50 dB.

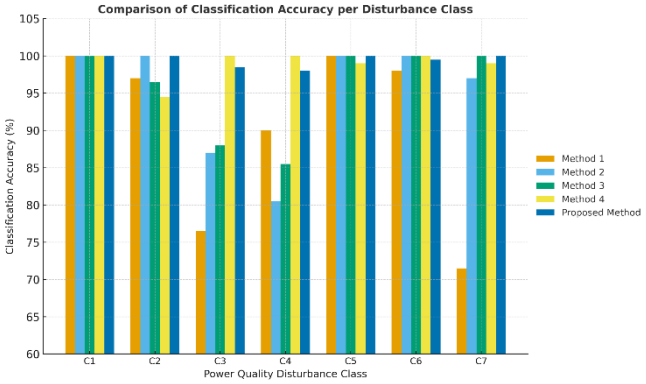

Figure 5. Comparison of classification accuracy per voltage disturbance class among previously published methods and the proposed approach

Table 5. Percentage of correct classification results from PQD under different noise conditions. The highest values are highlighted in bold.

Wavelet	20 dB	30 dB	40 dB	50 dB
db5	96.36	99.07	**99.07**	**99.50**
sym5	**96.71**	**99.14**	99.00	99.36
(Uyara et al., 2008)	89.92	91.85	93.64	95.14

Comparing the results from Tables 2 and 5 it is noted that the results obtained from noisy signals are very close to those presented for pure signals. For 20 dB noisy signals, the accuracy is superior to 96% for db5 and sym5 wavelets, which is also greater than those without noise shown in (Abdel-Galil et al., 2004; He and Starzyk, 2006; Uyara et al., 2008). Therefore, the proposed wavelet feature extraction has a robust anti-noise performance and high classification success in noisy condition can be still attained. Figures 6-(a)

and 6-(b) show the two-dimensional projections for wavelet feature F2$_x$ considering pure and noisy 20 dB PQ signals, respectively. Note that, data are efficiently separated using this wavelet feature even if noisy signals are considered.

4.5. Classification results from different classifiers

To verify generality across classifiers, Bayes Net (Koski and Noble, 2011), Multilayer Perceptron (Haykin, 1998) with backpropagation (Bishop, 1995), IBk (Aha et al., 1991), KStar (Cleary and Trigg, 1995) and Random Forest (Breiman, 2001) were evaluated using the db5 + F2$_x$ feature vector. Table 6 presents classification accuracies for each classifier under noise levels of none, 20, 30, 40 and 50 dB. Overall, instance-based (IBk) and Random Forest classifiers produced the highest accuracies, while Bayes Net and KStar exhibited slightly lower performance in noisy conditions. Thus, it is possible to verify that the proposed wavelet feature method is also efficient considering different classification systems. Table 7 reports confusion matrices for all classifiers at 20 dB SNR (200 cases per class). The matrices indicate dominant diagonal entries and limited off-diagonal misclassifications, confirming robust discrimination among the seven PQD classes under moderate noise.

4.6. Discussion, limitations, and future directions

The proposed CWT-based feature extraction addresses several gaps in PQD classification. First, it reconciles high accuracy with low computational complexity by using an extremely compact feature vector (two dimensions in the F2$_x$ configuration), facilitating real-time and embedded implementations. Second, targeted scale selection decouples feature extraction from FWT decomposition-level constraints and aligns features with physically meaningful frequencies. Third, the energy-based features yield interpretable representations compared with many black-box deep-learning methods.

Limitations should be acknowledged. Classification performance may deteriorate for signals with substantial frequency deviations outside the predefined scales or for disturbances dominated by non-harmonic components; in such cases, adding scales or employing adaptive scale selection may improve robustness. The method also assumes stable sampling and accurate time synchronization, which can be challenging in some field conditions.

Future work will investigate adaptive and data-driven scale selection strategies, hardware-level implementations for embedded monitoring, and validation on field-acquired voltage data to confirm practical performance.

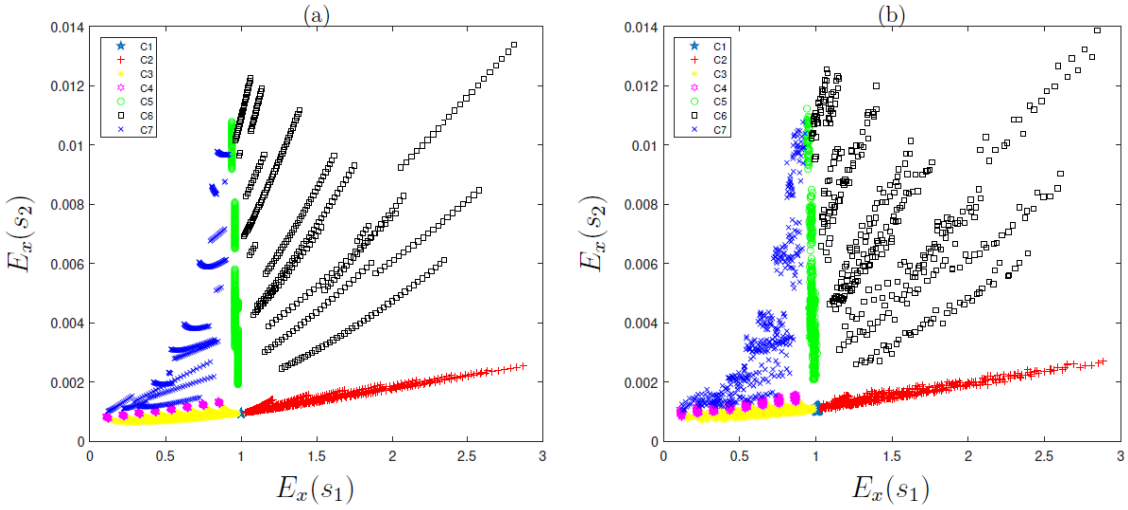

Figure 6. Two-dimensional projections for wavelet feature F2x. (a) pure signals. (b) noisy 20 dB PQ signals

Table 6. Correct classification results in terms of percentage considering different classification systems. The highest values are highlighted in bold.

Classifiers	no noise	20 dB	30 dB	40 dB	50 dB
Bayes net	96.43	93.64	96.21	96.21	96.21
Multilayer perceptron	99.07	96.57	99.00	99.00	99.07
IBk	**99.64**	97.57	99.29	**99.43**	**99.64**
KStar	94.57	92.93	94.43	94.57	94.64
Random forest	**99.64**	**97.71**	**99.64**	99.36	**99.64**

Table 7. Confusion matrices for all classifiers under 20 dB noise condition.

		C1	C2	C3	C4	C5	C6	C7
Bayes net	C1	195	2	3	0	0	0	0
	C2	0	200	0	0	0	0	0
	C3	0	0	157	25	0	0	18
	C4	0	0	4	183	0	0	13
	C5	0	0	0	0	196	3	1
	C6	0	0	0	0	5	195	0
	C7	0	0	2	13	0	0	185
	Overall	93.64%						
		C1	C2	C3	C4	C5	C6	C7
Multilayer perceptron	C1	200	0	0	0	0	0	0
	C2	3	197	0	0	0	0	0
	C3	2	0	186	12	0	0	0
	C4	0	0	4	182	0	0	1
	C5	0	0	0	0	200	0	0
	C6	0	0	0	0	2	198	0
	C7	0	0	0	10	1	0	189
	Overall	96.57%						
		C1	C2	C3	C4	C5	C6	C7
IBk	C1	199	0	1	0	0	0	0
	C2	0	200	0	0	0	0	0
	C3	0	0	189	11	0	0	0
	C4	0	0	13	187	0	0	1
	C5	0	0	0	0	199	1	0
	C6	0	0	0	0	2	198	0
	C7	0	0	2	2	2	0	194
	Overall	97.57%						
		C1	C2	C3	C4	C5	C6	C7
KStar	C1	200	0	0	0	0	0	0
	C2	16	184	0	0	0	0	0
	C3	14	0	159	27	0	0	0
	C4	0	0	1	199	0	0	1
	C5	0	0	0	0	200	0	0
	C6	0	0	0	0	11	189	0
	C7	0	0	0	25	5	0	170
	Overall	92.93%						
		C1	C2	C3	C4	C5	C6	C7
Random forest	C1	197	0	3	0	0	0	0
	C2	0	200	0	0	0	0	0
	C3	0	0	190	10	0	0	0
	C4	0	0	10	190	0	0	1
	C5	0	0	0	0	200	0	0
	C6	0	0	0	0	1	199	0
	C7	0	0	0	0	0	0	200
	Overall	97.71%						

CONCLUSION

This study proposed a compact and efficient approach for classifying power-quality disturbances based on the energy of CWT coefficients. The method employs only two decomposition scales representing the fundamental frequency and the first harmonic. Using this two-dimensional feature vector, high classification accuracy was achieved with very low computational cost. Extensive evaluations across multiple mother wavelets and classifiers identified db5 and sym5 combined with the $F2_x$ feature as the most effective configurations, achieving accuracies approaching 100% for both clean and noisy signals. Comparative analysis with previously reported FWT–based methods confirmed the superior performance and robustness of the proposed technique, even under 20 dB noise conditions. These results indicate that the

CWT-energy feature extraction provides a reliable and interpretable alternative for automatic PQD detection and classification in electric-power systems.

REFERENCES

Abdel-Galil, T. K. A., Kamel, M., Youssef, A. M., El-Saadany, E. F. E., & Salama, M. M. A. (2004). Power quality disturbance classification using the inductive inference approach. *IEEE Transactions on Power Delivery, 19*(4), 1812–1818.

Abry, P. (1997). *Ondelettes et turbulences: multirésolutions, algorithmes de décomposition, invariance d'échelle et signaux de pression*. Diderot Multimédia.

Aha, D. W., Kibler, D., & Albert, M. K. (1991). Instance-based learning algorithms. *Machine Learning, 6*(1), 37–66.

Akmaz, D. (2022). A new signal processing approach for classification of power quality disturbances. *Digital Signal Processing, 130*, 103701.

Angrisani, L., Daponte, P., & D'Apuzzo. (2001). Wavelet network-based detection and classification of transients. *IEEE Transactions on Instrumentation and Measurement, 50*(5), 1425–1435.

Barros, J., Diego, R. I., & de Apráiz, M. (2012). Applications of wavelets in electric power quality: Voltage events. *Electric Power Systems Research, 88*, 130–136.

Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford University Press.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Cai, J., Zhang, K., & Jiang, H. (2023). Power quality disturbance classification based on parallel fusion of CNN and GRU. *Energies, 16*(10), 4029.

Caicedo, J. E., Agudelo-Martínez, D., Rivas-Trujillo, E., & Meyer, J. (2023). A systematic review of real-time detection and classification of power quality disturbances. *Protection and Control of Modern Power Systems, 8*, 1–37.

Calderon, A. P. (1964). Intermediate spaces and interpolation: The complex method. *Studia Mathematica, 24*, 113–190.

Chawda, G. S., Shaik, A. G., Shaik, M., Padmanaban, S., Holm-Nielsen, J. B., Mahela, O. P., & Kaliannan, P. (2020). Comprehensive review on detection and classification of power quality disturbances in utility grid with renewable energy penetration. *IEEE Access, 8*, 146807–146830.

Cleary, J. G., & Trigg, L. E. (1995). K*: An instance-based learner using an entropic distance measure. In *Machine Learning Proceedings 1995* (pp. 108–114). Elsevier.

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics, 41*(7), 909–996.

Decanini, J. G., Tonelli-Neto, M. S., Malange, F. C., & Minussi, C. R. (2011). Detection and classification of voltage disturbances using a fuzzy-ARTMAP-wavelet network. *Electric Power Systems Research, 81*(10), 2057–2065.

Grossmann, A., & Morlet, J. (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis, 15*(4), 723–736.

Haykin, S. (1998). *Neural networks: A comprehensive foundation* (2nd ed.). Prentice Hall.

He, H., & Starzyk, J. A. (2006). A self-organizing learning array system for power quality classification based on wavelet transform. *IEEE Transactions on Power Delivery, 21*(1), 286–295.

Eristi, H., Uçar, A., & Demir, Y. (2010). Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines. *Electric Power Systems Research, 80*(7), 743–752.

Huang, S. J., Hsieh, C. T., & Huang, C. L. (1998). Application of wavelets to classify power system disturbances. *Electric Power Systems Research, 47*(2), 87–93.

IEEE Standards Association. (2009). *IEEE recommended practice for monitoring electric power quality* (IEEE Std 1159-2009). IEEE, Inc., NY, USA.

Jiang, J., Wu, H., Zhong, C., Cai, Y., & Song, H. (2024). A novel methodology for microgrid power quality disturbance classification using URPM-CWT and multi-channel feature fusion. *IEEE Access, 12*, 35597–35611.

Jiang, Z., Wang, Y., Li, Y., & Cao, H. (2024). A new method for recognition and classification of power quality disturbances based on IAST and RF. *Electric Power Systems Research, 226*, 109939. https://doi.org/10.1016/j.epsr.2024.109939

Khetarpal, P., Nagpal, N., Alhelou, H. H., Siano, P., & Al-Numay, M. (2024). Noisy and non-stationary power quality disturbance classification based on adaptive segmentation empirical wavelet transform and support vector machine. *Computers and Electrical Engineering, 118*, 109346.

Koski, T., & Noble, J. (2011). *Bayesian networks: An introduction*. Wiley.

Lin, C. H., & Wang, C. H. (2006). Adaptive wavelet networks for power quality detection and discrimination in a power system. *IEEE Transactions on Power Delivery, 21*(3), 1106–1113.

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

Markovska, M., & Taškovski, D. (2018). Efficient feature extraction and classification of power quality disturbances. *Journal of Electrical Engineering and Information Technologies, 3*(1), 13–20.

Meyer, Y. (1993). *Wavelets: Algorithms and applications*. Society for Industrial and Applied Mathematics.

Pillay, P., & Bhattacharjee, A. (1996). Application of wavelets to model short-term power system disturbances. *IEEE Transactions on Power Systems, 11*(4), 2031–2037.

Saini, M. K., & Beniwal, R. K. (2017). Optimum fractionally delayed wavelet design for PQ event detection and classification. *International Transactions on Electrical Energy Systems*, 1–15.

Salles, R. S., & Ribeiro, P. F. (2023). The use of deep learning and 2-D wavelet scalograms for power quality disturbances classification. *Electric Power Systems Research, 214*, 108834.

Samanta, I. S., Rout, P. K., Swain, K., Cherukuri, M., & Mishra, S. (2022). Power quality events recognition using enhanced empirical mode decomposition and optimized extreme learning machine. *Computers and Electrical Engineering, 100*, 107926.

Santoso, S., Powers, E. J., & Grady, W. M. (1994). Electric power quality disturbance detection using wavelet transform analysis. In *Proceedings of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*.

Santoso, S., Powers, E. J., Grady, W. M., & Hofmann, P. (1996). Power quality assessment via wavelet transform analysis. *IEEE Transactions on Power Delivery, 11*(2), 924–930.

Sengur, A., Turkoglu, I., & Ince, M. C. (2007). Wavelet packet neural networks for texture classification. *Expert Systems with Applications, 32*(2), 527–533.

Tse, N. C. F., Zhou, L., & Lai, L. L. (2010). Wavelet-based algorithm for power quality analysis. *European Transactions on Electrical Power, 20*(7), 952–964.

Uyara, M., Yildirim, S., & Gencoglu, M. (2008). An effective wavelet-based feature extraction method for classification of power quality disturbance signals. *Electric Power Systems Research, 78*(10), 1747–1755.

Zhang, Y., Zhang, Y., & Zhou, X. (2022). Classification of power quality disturbances using visual attention mechanism and feed-forward neural network. *Measurement, 188*, 110390.

To cite this paper, use:

Gossler, F.E. & Duarte, M.A.Q. (2025). Wavelet-based feature extraction using continuous Wavelet transform for power quality disturbances classification. *Multi-Science Journal*, 8(1): 52-63. DOI: https://doi.org/10.33837/msj.v8i1.1732