Combined Novel Approach of DWT and Feedforward MLP-RBF Network for the Classification of Power Signal Waveform Distortion

Aslam P. Memon1*, M. Aslam Uqaili1, Zubair A. Memon1, A. Asif Ali2, Ahsan Zafar2

1Department of Electrical Engineering, Mehran University of Engineering, & Technology, Jamshoro, Sindh, Pakistan; aslam@quest.edu.pk
2Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Sindh, Pakistan

Abstract
Power Quality (PQ) has become a major concern owing to its increased use of sensitive electronic equipment. In order to improve PQ problems, the detection and classification of PQ Disturbances (PQDs) must be carried out first. This paper presents a simple software based technique for detection and classification of PQDs by time-frequency analysis of Wavelet Transform (WT) as features extraction and Artificial Neural Network (ANN) as classifier. This approach detects and classifies the types of Waveform Distortion (WFD) problems of PQDs selecting suitable feature extraction with statistical parameters, as an input of feedforward Radial Basis Function (RBF) and Multilayer Perceptron (MLP). This methodology shows applicability, simplicity, and accuracy proving as promising tool for the automatic detection and classification of WFD of EPQ problems.

Keywords: Discrete Wavelet Transform, Feedforward Radial Basis Function and Multilayer Perceptron, Multiresolution Analysis, Waveform Distortion

1. Introduction
Any power problem manifested in voltage, current, or frequency deviations that result in failure or disoperation of customer equipment and system itself is termed as Electrical Power Quality (EPQ) problem. In electrical energy systems, voltages and especially currents become very irregular due to the increasing popularity of power electronics and other non-linear loads. More in particular, the power supplies for Information Technology (IT) equipment and high efficiency lighting, inverters and Adjustable Frequency Devices (AFD) are the main sources of Power Quality Disturbances (PQDs)\(^1\)-\(^3\). Hence PQ has become a major concern owing to increased use of sensitive electronic equipment and the ultimate reason that we are interested in power quality for its economic value. In order to mitigate or improve EPQ problems, the detection and classification of PQ disturbances must be carried out first. It is also the fact that PQDs vary in a wide range of time-frequency domain, which make automatic detection of PQDs problems often very difficult and elusive to diagnose. Hence one of the most important issues in power quality problems nowadays is how to detect these disturbance waveforms automatically in an efficient manner\(^2\)-\(^4\).

1.1 Types of Electrical Power Quality Disturbances (EPQDs)
In an electrical power system, there are various kinds of power quality disturbances. IEEE and IEC have defined power quality disturbances into seven categories\(^1\)-\(^4\): Transients, Short Duration Voltage Variations (SDVV), Long Duration Voltage Variations (LDVV), Voltage Imbalance, Waveform Distortion, Power Frequency Variations and Voltage Fluctuations. Some disturbances come from the supply network, whereas others are produced by the load itself\(^5\).
Waveform distortion is focused in this paper, it would be better to review briefly this important type of power quality disturbances.

1.1.1 Waveform Distortion (WFD)

A steady-state deviation from a sine wave of power frequency is called waveform distortion\(^6\). There are five primary types of waveform distortions: DC offset harmonics, interharmonics, notching, and electric noise. Table 1 shows categories and characteristics of power system electromagnetic phenomena of WFD. A Fourier series is usually used to analyze the nonsinusoidal waveform.

i) DC Offset: The presence of a DC current and/or voltage component in an AC system is called DC offset\(^7\). Main causes of DC offset in power systems are: Employment of rectifiers and other electronic switching devices, and geomagnetic disturbances\(^1\)–\(^6\),\(^8\) causing GICs.

The main detrimental effects of DC offset in alternating networks are: Half-cycle saturation of transformer core\(^2\), Generation of even harmonics\(^7\) in addition to odd harmonics\(^2\). Additional heating in appliances leading to a decrease of the lifetime of transformers\(^3\)–\(^5\),\(^6\)–\(^8\), rotating machines, and electromagnetic devices, and Electrolytic erosion of grounding electrodes and other connectors.

ii) Harmonics: Harmonics are sinusoidal voltages or currents with frequencies that are integer multiples of the power system (fundamental) frequency (usually, \(f = 50\) or \(60\) Hz). For example, the frequency of the \(n\)th harmonic is \((nf)\). Periodic non-sinusoidal waveforms can be subjected to Fourier series and can be decomposed into the sum of fundamental component and harmonics. Main sources of harmonics in power systems are industrial nonlinear loads, such as power electronic equipment, for example, drives rectifiers, inverters, or loads generating electric arcs, for example, arc furnaces, welding machines, and lighting. Residential loads with switch-mode power supplies such as television sets, computers, and fluorescent and energy-saving lamps\(^1\)–\(^6\),\(^8\). Some detrimental effects of harmonics are: maloperation of control devices, additional losses in capacitors, transformers, and rotating machines, additional noise from motors and other apparatus, telephone interference, and causing parallel and series resonance frequencies (due to the power factor correction capacitor and cable capacitance), resulting in voltage amplification even at a remote location from the distorting load.

Recommended solutions to reduce and control harmonics are applications of high-pulse rectification, passive, active, and hybrid filters, and custom power devices such as active-power line conditioners (APLCs) and unified power quality conditioners (UPQCs)\(^7\).

iii) Interharmonics: The frequency of interharmonics are not integer multiples of the fundamental frequency. Interharmonics appear as discrete frequencies or as a band spectrum. Main sources of interharmonic waveforms are static frequency converters, cycloconverters, induction motors, arcing devices, and computers. Interharmonics cause flicker, low-frequency torques\(^1\)–\(^2\),\(^6\)–\(^8\), additional temperature rise in induction machines\(^3\),\(^4\), and malfunctioning of protective (under-frequency) relays\(^1\)–\(^5\).

Interharmonics have been included in a number of guidelines such as the IEC 61000-4-7\(^6\),\(^8\) and the IEEE-519. However, many important related issues, such as the range of frequencies, should be addressed in revised guidelines.

iv) Notching: A periodic voltage disturbance caused by normal operation of power electronics devices when current is commutated from one phase to another is termed notching. Notching tends to occur continuously and can be characterized through the harmonic spectrum of the affected voltage. The frequency components can be quite high and may not be able to describe with measurement equipment used for harmonic analysis. VFD, arc welders and light dimmers are the main cause of notching. The effects of notching include the halt or loss of the data of the system\(^2\).

---

Table 1. Categories and characteristics of power system electromagnetic phenomena

<table>
<thead>
<tr>
<th>Categories of WFD</th>
<th>Typical contents</th>
<th>Typical Duration</th>
<th>Typical Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC offset</td>
<td>Steady state</td>
<td>0–0.1%</td>
<td></td>
</tr>
<tr>
<td>Harmonics</td>
<td>0–100th harmonic</td>
<td>Steady state</td>
<td>0–20%</td>
</tr>
<tr>
<td>Interharmonics</td>
<td>0–6 kHz</td>
<td>Steady state</td>
<td>0–2%</td>
</tr>
<tr>
<td>Notching</td>
<td>Broadband</td>
<td>Steady state</td>
<td>0–1%</td>
</tr>
<tr>
<td>Noise</td>
<td>Broadband</td>
<td>Steady state</td>
<td>0–1%</td>
</tr>
</tbody>
</table>
v) **Electric Noise:** Electric noise is defined as unwanted electrical signals with broadband spectral content lower than 200 kHz\(^6\)–\(^8\) superimposed on the power system voltage or current in phase conductors, or found on neutral conductors or signal lines. Electric noise may result from faulty connections in transmission or distribution systems, arc furnaces, electrical furnaces, power electronic devices, control circuits, welding equipment, loads with solid-state rectifiers, improper grounding, turning off capacitor banks, adjustable-speed drives, corona, and Broadband Power Line (BPL) communication circuits. The problem can be mitigated by using filters, line conditioners, and dedicated lines or transformers. Electric noise impacts electronic devices such as microcomputers and programmable controllers.

### 1.2 Digital Signal Processing (DSP) Techniques Used In PQ Disturbances

Various signal processing techniques used in power quality disturbances field are briefly discussed\(^2\)–\(^6\) with their drawbacks and in last the advantages of wavelet transform with solid reasoning have been discussed. This proposed methodology is the extension of author’s work published in his paper\(^2\).

In 1994, the use of wavelets was proposed to study power systems non-stationary harmonics distortion\(^2\). This technique is used to decompose the signal in different frequency sub-bands and study separately its characteristics. Finally, the STFT (Short Time Fourier Transform) is commonly known as a sliding window version of the FFT, which has shown better results in terms of resolution and frequency selectivity. However, STFT has a fixed frequency resolution for all frequencies, and has shown to be more suitable for harmonic analysis of voltage disturbances than binary tree filters or wavelets when is applied to study voltage dip\(^2,5,7\)–\(^9,11\).

Traditionally, the Fourier transform permits mapping signals from time domain to frequency domain by decomposing the signals into several frequency components. This technique is criticized in that the time information of transients is totally lost, although the accuracy of frequency components is high. Fourier transform does not fit the analysis of transients owing to the non-stationary property of its signals in both time and frequency domains. Wavelet transform generally offers this facility\(^12\)–\(^20\).

Latest advances in electrical power quality mitigation techniques are based on extraction of disturbances data instead of traditional methods. Hence time-frequency analysis is more suitable to detect disturbances from data. PQ disturbances also vary in a wide range of time and frequency, and Wavelet Transformation (WT) has unique ability to examine the signal in time and frequency ranges at the same time which makes WT a best suited tool for power quality disturbances\(^2,9\)–\(^11\).

Detailed studies of literature survey\(^2,9\)–\(^20\) indicate that several researches have been presented proposing the implementation of WT or pattern recognition artificial intelligence or both for the detection and classification of EPQDs types\(^2\)–\(^5\).

PQ disturbances have been defined into seven categories. The most focusing point of research is on few types of disturbance. The important type of PQ disturbances “waveform distortion” have not been fully addressed or focused. DC offset, harmonics, interharmonics, notching and noise are five major disturbances\(^2\).

This paper presents a simple software based technique for detection and classification of WFD of PQDs by time-frequency analysis of Discrete Wavelet Transform (DWT) with Multiresolution Analysis (MRA) and artificial neural network (MLP-RBF) as classifier.

This approach with proposed time frequency, feature extraction with statistical parameters as feature vector as input to FFNN (MLP-RBF) as classifier proposed methodology can detect and classify EPQ problems of WFD proving a promising tool for the automatic detection and classification of EPQDs problems with an efficient manner.

### 1.3 Wavelet Transform (WT)

#### 1.3.1 Introduction

WT (time-frequency) analysis is more suitable to detect disturbances from data for feature extraction which the latest techniques of DSP. PQ disturbances vary in a wide range of time and frequency, and WT has unique ability to examine the signal in time and frequency ranges at the same time which makes WT a best suited tool for power quality disturbances\(^21\)–\(^22\).

The wavelet transform represents signal as a sum of wavelets at different locations (positions) and scales (duration). The wavelet coefficients work as weights of the wavelets to represent the signal at these locations and scales.
Types: The wavelet transform can be accomplished three different ways.

The Continuous Wavelet Transform (CWT) where one obtains the surface of the wavelet coefficients, for different values of scaling and translation factors. It maps a function of a continuous variable into a function of two continuous variables.

The second transform is known as the Wavelet Series (WS) which maps a function of continuous variables into a sequence of coefficients. The third type of wavelet transform is the Discrete Wavelet (DWT), which is used to decompose a discretized signal into different resolution levels. It maps a sequence of numbers into a different sequence of numbers.

Mathematics: In wavelet theory known as the mathematics tool, defining a model for non-stationary signals $X(t)$ are decomposed by a set of small wave components called wavelet. Every wavelet is created by scaling and translation operations in the functions called mother wavelet. Firstly this transformation is expressed by Continuous Wavelet Transform (CWT) as:

$$W_\psi(a, b) = \frac{1}{\sqrt{a}} \int x(t) \psi^*(\frac{t - b}{a}) dt$$

(1)

where, $a$ and $b$ are scale and translation real numbers, $a \neq 0$ and $\psi^*$ is complex conjugate of $\psi$ and $W_\psi(a, b)$ are the wavelet coefficients. Secondly for computer implementations Discrete Wavelet Transform (DWT) will be utilized as:

$$W_\psi(m, n) = \frac{1}{\sqrt{a_0^n}} \sum_{k=-\infty}^{\infty} x(k) \left( \frac{k - a_0^n b_0}{a_0^n} \right)$$

(2)

Where $a = a_0^n$, $b = a_0^n b_0$, $m$ and $n$ are the integer numbers provided $a_0$ 1 and $b_0 \neq 0$. Due to this process redundancy of continuous form must be eliminated hence $a_0$ and $b_0$ be selected as to from orthogonal basis by satisfying the condition as $a_0 = 2$ and $b_0 = 1$. This requirement invites us to use Multiresolution Analysis (MRA), which is also known as Multiresolution Wavelet Method (MWM). In this method original signal $X(t)$ is decomposed into different scales resolutions and the mother wavelet function

$$\psi(t) = 2 \sum_{n=0}^{\infty} d_n \phi(2t - n)$$

is chosen with function

$$\phi(t) = 2 \sum_{n=0}^{\infty} c_n \phi(2t - n)$$

known as scaling function, where $d_n$ and $c_n$ are squared sum able sequences.

1.4 Discrete Wavelet Transform (DWT) and Multiresolution Analysis (MRA)

In Multiresolution Analysis (MRA), wavelet functions and scaling of functions are used as building blocks to decompose and reconstruct the signal at different resolution levels. The wavelet functions will generate the detail version of the decomposed signal and the scaling function will generate the approximated version of the decomposed signal.

MRA refers to the procedures to obtain low-pass approximations and high-pass details from the original signal. An approximation contains the general trend of the original signal while a detail embodies the high-frequency contents of the original signal. Approximations and details are obtained through a succession of convolution processes. The original signal is divided into different scales of resolution, rather than different frequencies, as in the case of Fourier analysis. The maximum number of wavelet decomposition levels is determined by the length of the original signal and the level of detail required. Details and approximations of the original signal are obtained by passing it through a filter bank, which consists of low and high-pass filters. A low pass filter removes the high frequency components, while the high pass filter picks out the high-frequency contents in the signal being analyzed.

1.4.1 Suitable Mother Wavelet

Selection of a suitable mother wavelet for power system transient signal is an art, instead of developing algorithms of wavelets for different problems. At the lowest scale like scale 1, the mother wavelet is most localized in time and oscillates most rapidly within a very short period of time. As the wavelet goes to higher scales, the analyzing wavelets become less localized in time and oscillate less due to the dilation nature of the wavelet transform analysis. As a result of higher scale signal decomposition, fast and short transient disturbances will be detected at lower scales, whereas slow and long transient disturbances will be detected at higher scales.

Choice of mother wavelets plays a significant role in detecting various types of power quality disturbances, especially when considering small scale signal decompositions. For fast and short transient disturbances, Daub4 (called as Daubechies 4) and Daub6 wavelets are better, while for slow and long transient disturbances, Daub8 and Daub10 are particularly good. At the lowest scale like
scale 1, the mother wavelet is most localized in time and oscillates most rapidly within a very short period of time. As the wavelet goes to higher scales, the analyzing wavelets become less localized in time and oscillate less due to the dilation nature of the wavelet transform analysis\(^{2–4,21–25}\).

1.5 Artificial Intelligence (AI)

Controlling complex systems highly non-linear has shown to be very difficult using conventional control theory. The artificial intelligence with its natural language has proven to be useful in these cases as it deals with uncertainties, what brings it closer to human being logic thought. Artificial neural network has been chosen from artificial intelligence.

1.5.1 Artificial Neural Networks (ANNs)

ANN and its applications in electrical power system problems is not a new topic of research, because it has been suggested in many research areas with fast growing interest.

Literatures indicate that artificial neural network is swiftly drawing the attentions and recognition amongst the power system researchers. They are enormously useful in the area of electrical engineering within few years\(^{26–30}\).

It is called neural network because it is a network of interconnected elements. These elements were inspired from the studies of biological nervous systems. In other words, neural networks are an attempt at creating machines that work in a similar way to the human brain by building these machines using components that behave like biological neurons. The function of neural networks is to produce an output pattern when presented with an input pattern. ANNs can easily handle complicated problems and can identify and learn correlated patterns between sets of input data and corresponding target values. After training, these networks can be used to predict the outcome from new input data\(^{26–27}\).

As ANNs are universal function approximators, they are capable of approximating any continuous nonlinear functions to arbitrary accuracy. The other advantages inherent in ANNs are their robustness, parallel architecture and fault tolerant capability\(^{26}\).

Many ANN architectures have been proposed. These can be roughly divided into three large categories: feedforward (multilayer) neural networks, feedback neural networks and cellular neural networks. This paper is solely concerned with the feedforward neural networks.

Feedforward neural networks can be divided into two major classes: the Multilayer Perceptron (MLP) and Radial Basis Function (RBF) neural networks.

1.6 Feature Extraction

Power quality disturbance recognition is often a problem, because it involves a broad range of disturbance or classes from low frequency dc offsets to high frequency transients or low duration impulse to steady state events. Therefore, decision boundaries of disturbance feature may overlap. So selection of the feature vector of scale coefficient has a great importance.

Feature extraction is the key for pattern recognition so that it is the most important component of designing the intelligent system based on pattern recognition since even the best classifier will perform poorly if the features are not chosen well. A feature extractor should reduce the pattern vector (i.e., the original waveform) to a lower dimension, which contains most of the useful information from the original vector\(^{2–4,31}\).

Wavelet Transform is suitable for feature extraction. Its properties, like limited effective time duration, band pass spectrum, waveform similar to disturbance and orthogonality, allow locating information in time and frequency domains. Thus, it is possible to obtain high correlation when PQ disturbances occur and decompose these events into different components without energy aliasing. With this information it is possible to identify the difference of disturbances and some patterns can also be selected using this decomposition\(^{2–4,31}\).

The input of the neural network is a preprocessed signal. In this case, PQ disturbance signal in the time domain is transformed into the wavelet domain before applying as input to the neural network.

The wavelet coefficients obtained by wavelet and multiresolution analysis to disturbance signals have effective feature information. Most useful features must be first extracted from those coefficients to reduce the dimension of feature vectors, in order to more effectively recognize the type of PQ disturbances.

To avoid the adverse influence of noise a de-noising procedure based on DWT is performed. This carried out by discarding the noise related DWT coefficients which are lower than a threshold level. The reconstructed signal is almost free of noise and has the same energy content. To enhance the diagnosis procedure preprocessing of the event is used to extract characteristic information (feature vector).
For this methodology it consists of, a five level WT decomposition using Db4 which is found sufficient to explore minimum significant information in different frequency band.

The statistical parameter of DWT coefficients such as standard deviation, mean and maximum absolute value of the various scale levels can be a representation of event signal energy band to aid in its classification.\(^{12-33}\)

2. The Proposed Methodology

The projected methodology for the detection and classification of waveform distortion of power system is proposed on following stages:

1. Generation of data
2. Feature extraction
3. Data normalization
4. Network training

2.1 Data Generation

The equations for the WFD PQDs signals are developed taking the parameters varied as described within the ranges by the IEEE 1159.\(^{2,4-5}\)

Data normalization: The higher raw input data can suppress the influence of smaller ones; hence to avoid this, the raw data is normalized before the process of FFNN classification and the data is normalized with following formula:

\[
d_N = \frac{(d - d_{\text{min}}) \times \text{range}}{(d_{\text{max}} - d_{\text{min}})} + \text{starting value}
\]

Where \(d_N\) is the normalized value and \(d_{\text{max}}\) and \(d_{\text{min}}\) are the minimum and maximum values of \(d^{2,4-5,24-25}\).

2.2 Feature Extraction

The features of the input data are extracted with the help of DWT-MRA with standard statistical parameters transforming time signal into time-frequency domain.

To achieve feature vector (characteristic information) the diagnosis procedure called preprocessing of the event is used which consists of 6-level WT-MRA decomposition with Db4 as mother wavelet is sufficient to investigate significant information in different frequency band.\(^{2,4-5,27}\)

The WFD signals are sampled at the rate of 10 kHz in this methodology, d1, d3 and d5 detail and a5 approximation coefficients levels are sufficient for this work which simply, and efficiently, provide useful data for disturbances of WFD signals feature vector. Standard statistical parameters selected are: signal power, standard deviation of first, third, fifth details coefficients (d1, d3, d5), maximum absolute of first, third, fifth details levels, and means value of approximate fifth level (a5), for low frequency events such as dc offset). In this way the whole feature vector will produce 8 elements which is an optimum vector length for this range of neural network classifier. This proposed computed feature vector is reseated as the input to FFNN (RBF-MLP) to train as classifier for the detection of WFDs problems.

The output of processing unit consists of a vector (one element for each class). To make a decision about the disturbance type and to provide a level of confidence for the decision, we used a simple threshold levels for each events. The selected thresholds depend on number of train data, learning error and reliability of classifier. Patterns that are out of all thresholds specified as unknown events.

2.3 Network Training:

2.3.1 Classification Algorithm

ANN can be applied to wide variety of problems, such as storing and recalling data or patterns, classifying patterns, performing general mappings from input patterns to output patterns, grouping similar patterns. Training a network by i) selected algorithms of feedforward NN (back-propagation and orthogonal least-squares (OLS) as the input training pattern, ii) the adjustment of the weights and bias during the learning process, and iii) the ANN weights are adapted in order to create the desired output vectors are the main stages for classification purposes. After training the weights and bias they can be used as classifier data for WFD power system problems.

Figure 1 shows the flow chart of proposed methodology showing the combined wavelet transformation (DWT-MRA and db 4 as mother wavelet) and neural network (RBF-MLP) approach for the detection and classification of WFD power quality disturbance problems.

2.3.2 Design of Classification Algorithm

(i) Design methodology of MLP network

First layer has weights coming from the input. Each subsequent layer has a weight coming from the previous layer. The last layer is the network output. The weights and biases are initialized and adapted with a specified learning
function. Training is done with specified hyperbolic tangent sigmoid transfer function. Performance is measured according to the specified performance function.  

The feedforward network architectures have been selected for this work. Iteration type loop uses Levenberg-Marquardt algorithm which is the fastest training algorithm for networks of moderate size and possesses very simple and an efficient Matlab implementation, occupying very less memory. Mean squared error (mse) is selected as performance function and network is stored with a command “generate simulation net”.  

A large data set may cause the learning process to be very slow; hence large data set is not essential for successful training. An appropriate subset of data for training can improve the speed of training without having a serious effect on the performance.  

With the help of Levenberg-Marquardt algorithm into the standard back-propagation learning algorithm, the training time of an MLP network can be improved significantly.

There is no straightforward rule of choosing an appropriate number of hidden layer neurons for an optimal performance, which is a big disadvantage of MLP networks. This number is chosen by trial and error methods, starting with two or three neurons, and then increasing the number gradually, until satisfactory performance is achieved. This method of MLP is laborious and time-consuming task. The weights and biases have been initialized and adapted with a specified learning function and training with specified hyperbolic tangent sigmoid transfer function. In the last performance is measured according to the specified performance function. Table 2 describes MLP architecture and training parameters.

(ii) Design methodology of RBF network

The potentiality and applicability of radial basis networks for this application is being investigated, because they have distinctive properties of simple network structure, efficient learning way, and best approximation, which make radial basis networks more powerful tool than the other types of neural networks.

The radial basis networks consist of three utterly different layers. The input layer or first layer consists of a number of units fastened to the input vector. The units constituted by hidden layer or second layer have an overall response function, mostly a Gaussian function. The function of each class is computed by third layer.

Diversity of different algorithms has been proposed in the latest literature for choosing the proper radial basis network centres. We prefer the universal approximator orthogonal least-squares (OLS) algorithm for this research work.

OLS learning algorithm generates radial basis network, which have a hidden layer, smaller than that of radial basis network with arbitrarily chosen centres.

<table>
<thead>
<tr>
<th>architectures</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>the number of layers</td>
<td>Input: 09, hidden: 07, output: 4</td>
</tr>
<tr>
<td>the number of neuron on the layers</td>
<td>Random</td>
</tr>
<tr>
<td>activation functions</td>
<td>Tangent sigmoid</td>
</tr>
<tr>
<td>training parameters</td>
<td>Levenberg–Marquardt back-propagation (BP) supervised learning algorithm</td>
</tr>
<tr>
<td>learning rule</td>
<td>0.67</td>
</tr>
<tr>
<td>epochs</td>
<td>15000</td>
</tr>
<tr>
<td>mean-squared error</td>
<td>$1e^{-06}$</td>
</tr>
</tbody>
</table>
Radial basis networks are used to fairly accurate function. They include neurons to the second or hidden layer awaiting it meets the precise Mean Square Error Goal (MSEG)\textsuperscript{26–28,35}.

For this application two-layer network has been created. The input or first layer takes radial basis transfer function neurons which calculates its weighted inputs and its net input with net product. The second layer takes linear transfer function neurons and calculates its weighted inputs with dot product and its net inputs with net sum. First and second both layers have biases.

At first no neuron is available in radial basis transfer function and FF network architectures with two-layer structures are generated.

The following steps are repeated until the network’s mean square error falls below the goal\textsuperscript{27–29,34–37}.

1. The network is simulated
2. The input vector with the greatest error is found
3. A radial basis transfer function neuron is added with weights equal to that vector
4. The linear transfer function layer weights are redesigned to minimize error

In this case feedforward network architectures are created with two-layer structure.

31 (Thirty-one) neurons are selected in the first hidden layer with activation function of radial basis transfer function.

In the second layer 1 (one) neuron or node with linear transfer function is trained.

Spread constant is 1.75 and error goal is 0.00001.

3. Application of Methodology and Simulation Results

3.1 Application of DWT with MRA

For the efficient computational analysis, decomposition at level 6, with db4 (Daubechies 4 as mother wavelet) and MRA algorithm of DWT is proposed for this work.

Higher level scaling of wavelet is more suitable but taking the benefit of computational efficiency and practical consideration, six level decomposition with Db4 wavelet is observed sufficient to get significant information\textsuperscript{2–4}. In first step the signals are generated at 6 periods (0.12 seconds) with sampling rate of 10KHZ, where DWT algorithm is applied with the help of Simulink 7.0, Wavelet Blocksets 4.1 and DSP Toolbox 6.6 of MATLAB.

Table 3. Approximate coefficients (a\textsubscript{1−6}) and their frequency bands

<table>
<thead>
<tr>
<th>a\textsubscript{i} level</th>
<th>0–2500</th>
</tr>
</thead>
<tbody>
<tr>
<td>a\textsubscript{1} level</td>
<td>0–2500</td>
</tr>
<tr>
<td>a\textsubscript{2} level</td>
<td>0–1250</td>
</tr>
<tr>
<td>a\textsubscript{3} level</td>
<td>0–625</td>
</tr>
<tr>
<td>a\textsubscript{4} level</td>
<td>0–312.5</td>
</tr>
<tr>
<td>a\textsubscript{5} level</td>
<td>0–156</td>
</tr>
<tr>
<td>a\textsubscript{6} level</td>
<td>0–78.125</td>
</tr>
</tbody>
</table>

Table 4. Detail coefficients (d\textsubscript{1−6}) and their frequency bands

<table>
<thead>
<tr>
<th>d\textsubscript{i} level</th>
<th>2500–5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>d\textsubscript{1} level</td>
<td>1250–2500</td>
</tr>
<tr>
<td>d\textsubscript{2} level</td>
<td>625–1250</td>
</tr>
<tr>
<td>d\textsubscript{3} level</td>
<td>312.5–625</td>
</tr>
<tr>
<td>d\textsubscript{4} level</td>
<td>156.25–312.65</td>
</tr>
<tr>
<td>d\textsubscript{5} level</td>
<td>78.125–156.25</td>
</tr>
</tbody>
</table>

In second step for the classification, Neural Network Toolbox 7.0.2 has been utilized. The frequency bands of approximate and detail coefficients obtained in first step as are shown in Table 3 and 4.

It can be seen that DWT coefficient at the interval of disturbance are much higher than other times. First scale detail coefficient signal (d\textsubscript{1}) includes the highest frequency band (2500–5000Hz) detects the disturbance instantaneously. The high scale approximate coefficient signal (a\textsubscript{6}) detects low frequency band (0–78.125) of signal in sixth scale. The accuracy of disturbance time localization decrease as scale increases. This decomposition gives time and frequency information of signal with good resolution. It is also visualized that the noise affects almost all scales especially at lower scales. This property makes time detection of disturbances accurate at lower scales and frequency at high scales.

Figure 2 shows detail coefficients d\textsubscript{1−6} and approximate coefficient a\textsubscript{6} at level 6 decomposition. The detail coefficients of WT at low scales show notch disturbance instantaneously, because higher frequency components are observed at these low scales at first look. Figure 3 shows the DWT detail coefficients at the instance of disturbances are higher than other times as shown in Table 1. At high scales DWT approximate coefficients are lower as shown in Table 2. This proves that as the scale increase the accuracy.
of disturbances time resolution decreases, and frequency resolutions increase. Hence good time localization is at low scale and good frequency localization at high scales.

DWT detail coefficients at low scales show noise signals having high frequency bands $d_1$ and $d_2$ in Figure 4. It is observed that noise on lower scales; make time localization of disturbances difficult as shown in $d_1$ and $d_2$. As scale increases noise removes and periodic notches are clearly detected.

DWT detail coefficients at low scales show high and fast transients disturbances. As harmonics are slow and continuous disturbances hence those are detected at high scales as shown in Figure 5.

Figure 6 shows DWT detail coefficients at low and high scales with fast transients disturbances. As
harmonics are slow and continuous disturbances hence those are detected at high scales. From $d_5$ and $d_6$ it is clear that these disturbances are not multiple integers of fundamental components, they are expected as interharmonics with harmonics. In Figure 7 only noise was added which was removed by WT filtration ($d_1$ to $d_4$).

DWT Wavelet detail coefficients at low scales show shift of DC offset instantaneously, and frequency resolution is better at high scales shown in Figure 8 and in Figure 9 detail coefficients at low scales show noise in signal having high frequency bands ($d_1$ to $d_4$). High scale WT coefficients, shows removal of noise from signal and becomes pure sine wave signal.

4. Conclusions

An algorithm for detection and classification of WFDs problems of PQDs has been proposed and successfully verified in a very simple and an efficient way using Matlab.

A time-frequency technique of DWT-MRA with Daubechies-4 for analysis of WFDs and feature extraction with statistical parameters for MLP-RBF classifier is proposed in this paper. Using Matlab/Simulink/WT/NNToolboxes, the information of signals gathered from detail, approximate coefficients and scaling, properties of WT techniques the detection and feature extraction data collection becomes easier task because WT is a powerful tool to analyze WFD PQ disturbances having time-frequency information required simultaneously.

It has been concluded that FFNNs are highly applicable and suitable as classifier for waveform distortion problems of PQDs. The input of the FFNN (MLP-RBF) neural network is a preprocessed technique of signal processing. In this methodology, WFD PQD signals in the time domain is transformed into the wavelet domain with statistical parameters before applying as an input feature vectors to the neural network for classification purposes. This stage of feature extraction method is the key for pattern recognition. A feature extractor always reduces the huge data of original waveform into lower and the most useful data, which is the prime requirement of pattern recognition.
The overall classification accuracy of MLP as classifier is 93.36% with 125 samples tested and 96.40% with 250 samples tested respectively. In case of RBF as classifier the accuracy is 94.4% and 96.8% with same samples as in case of MLP and shown in Tables 5 and 6.

Table 5. Shows the evaluations performance of developed model of MLP FFNN, with its classification results during testing

<table>
<thead>
<tr>
<th>WFDs</th>
<th>Samples</th>
<th>Identified</th>
<th>Unidentified</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>25</td>
<td>24</td>
<td>01</td>
</tr>
<tr>
<td>S2</td>
<td>25</td>
<td>23</td>
<td>02</td>
</tr>
<tr>
<td>S3</td>
<td>25</td>
<td>22</td>
<td>03</td>
</tr>
<tr>
<td>S4</td>
<td>25</td>
<td>24</td>
<td>01</td>
</tr>
<tr>
<td>S5</td>
<td>25</td>
<td>24</td>
<td>01</td>
</tr>
<tr>
<td>05</td>
<td>125</td>
<td>117</td>
<td>08</td>
</tr>
</tbody>
</table>

Overall accuracy 93.36%

Table 6. Shows the evaluations performance of developed model of RBF FFNN, with its classification results during testing

<table>
<thead>
<tr>
<th>WFDs</th>
<th>Samples</th>
<th>Identified</th>
<th>Unidentified</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>50</td>
<td>49</td>
<td>01</td>
</tr>
<tr>
<td>S2</td>
<td>50</td>
<td>47</td>
<td>03</td>
</tr>
<tr>
<td>S3</td>
<td>50</td>
<td>47</td>
<td>03</td>
</tr>
<tr>
<td>S4</td>
<td>50</td>
<td>49</td>
<td>01</td>
</tr>
<tr>
<td>S5</td>
<td>50</td>
<td>49</td>
<td>01</td>
</tr>
<tr>
<td>05</td>
<td>250</td>
<td>241</td>
<td>09</td>
</tr>
</tbody>
</table>

Overall accuracy 96.40%

The most important conclusion of this investigation is that both the architectures of FFNN (MPL and RBF networks) are suitable for this application as classifier and can produce robust performance. But it has been observed that the performance of RBF NNs is better than MLP networks as classifier. RBF NNs possess another property of the characteristics in terms of the compact network structure and best approximation as compared with MLP NNs.

Faster training time property with OLS algorithm, which selects an appropriate number of the RBF from input data, avoids the bothering problem of selecting an optimal number of hidden layer neurons automatically. But RBF networks require more hidden layer neurons than MLP networks for the solution of the same problem.

5. Acknowledgement

The authors would like to offer their sincere thanks to higher authorities of Mehran U.E.T Jamshoro and Quaid-e-Awam U.E.S.T Nawabshah, for continuous encouragement and time to time facilities when required during this research work, which is an attempt for PhD project.

6. References


34. Memon AP. Artificial neural network applications in electrical alternator excitation systems [M.Phil Thesis]. Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan; 2002.
