UACEE International Journal of Artificial Intelligence and Neural Networks - Volume 2: Issue 2 [ISSN 2250 - 3749]

A Study on Neural Network Transfer and Training Functions for Recognition of Power Quality Disturbances

Manoj Gupta
Department of Electrical Engineering
Poornima Institute of Engineering & Technology
Jaipur 302022-INDIA
manoj280332@rediffmail.com

Rajesh Kumar, R. A. Gupta
Department of Electrical Engineering
Malaviya National Institute of Technology
Jaipur 302017-INDIA
rkumar.ee@gmail.com, ragmnit@gmail.com

***Abstract*—Neural networks have been proved as an important and useful tool for solving a wide variety of practical and real-world problems. Huge research in this field alleviated in understanding and finding new and effective methods to address different problems. However, selection of apposite combination of training and transfer function for a particular problem is a cumbersome task. But, this can be ascertained through research experiences and outcomes. The objective of this work is to compare the performances of three transfer functions in tandem with fourteen training functions used for backpropagation training of neural network for recognition of power quality (PQ) disturbance signatures. The comparison is shown on the basis of Lowest MSE, number of epochs, convergence time, and accuracy. It is shown that among three transfer functions namely "logsig", "purelin", and "tansig"; the overall performance of "tansig" was superior and the accuracy of BR training function was 100 % with all the three types of transfer functions.**

***Keywords*—Artificial neural network, feedforward neural network, power quality, recognition, signature**

I. INTRODUCTION

Power quality has emerged as a topical issue in power system for utilities as well as for consumers due to increasing cost burden of poor power quality and augmented use of sensitive electronic equipments. According to IEEE standard 1159-1995 [1], the PQ disturbances include wide range of PQ phenomena namely transient (impulsive and oscillatory), short duration variations (interruption, sag and swell), power frequency variations, long duration variations (sustained under voltages and sustained over voltages) and steady state variations (harmonics, notch, flicker etc.) with time scale ranges from tens of nanoseconds to steady state. The new approaches for analysis of PQ disturbances involve application of advanced signal processing techniques predominantly wavelet transform and artificial intelligence techniques mainly neural networks [2], [3]. With the advancements in computing technology and emergence of new and powerful artificial intelligence techniques, the analysis of PQ disturbances has become a most sought after area for research with bounty of scope for innovations [4].

Analysis of power quality disturbances usually can be seen to consist of three main steps as shown in Fig. 1: a) Data generation or acquisition b) feature extraction and c) classification. Since, manual analysis is time consuming due to huge amount of data so automated procedure for analysis is preferred.

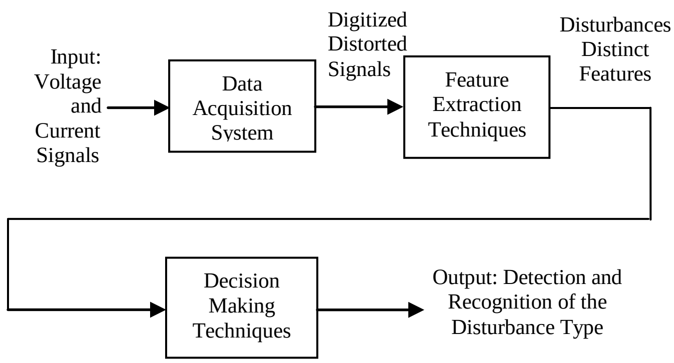


Figure 1. Principle block diagram of an automatic recognition system

PQ data is either generated using MATLAB or acquired from real system. Feature extraction recognizes and quantifies distinct and dominant features of a given disturbance. Wavelet transform based feature extraction is most suitable due to its excellent time frequency resolution [5]. As such, discrete wavelet transform is popular for signal processing due to its less computational burden along with quite a good speed but still importance of CWT cannot be ruled out for analysis of PQ disturbances CWT [6]. Disturbance recognition is final step of the analysis and is implemented through some form of artificial intelligence techniques [7]-[9]. Conventional approach is to use rule-based expert systems. More recent approach is to utilize non-linear mapping capability of artificial neural network. Neural networks dominate among all artificial intelligence techniques. This technique derives its origin from human brain. The human brain consists of neurons. These neurons send activation signals to each other thereby creating intelligent thoughts. The algorithmic version of a biological neural network (called an artificial neural network) also consists of neurons which send activation signals to each other. Hence, artificial neural networks also emulate the extraordinary functions of neurons within the brain, to build complicated machines that are programmed to solve difficult problems. They are capable to approximate a function of multiple inputs and outputs. As a corollary, neural networks can be used for a variety of data mining tasks, among which are recognition, clustering, descriptive modeling, function approximation, and time series prediction etc. Because the breadth of neural networks is very large, this work focuses on feedforward neural networks, perhaps the most common type.

Based on the work presented in reference [10], signatures of various power quality disturbances namely sag, swell, transient, harmonics, and flicker were obtained using CWT. It was corroborated that these signatures are unique in shape for a particular type of PQ disturbance and their size is proportional to the amplitude of the PQ disturbance. Hence, these signatures are used for recognition of different PQ disturbances as well as for indexing the level of PQ disturbance as high, medium, and low using the feedforward neural network [11]-[13].

Since, a number of factors influence the performance of neural networks including the complexity of the problem, the number of datasets used in training, the number of weights and biases in the network, the error goal, and the purpose of neural network like recognition or function approximation etc. Hence, there is no straightforward method available to know about the best combination of transfer and training function, in terms of convergence speed and accuracy, for a given problem. As such research efforts were made for studying the effects of transfer and training functions on the performance of neural networks for various types of problems [14]-[16]. In this work, the performance of feedforward backpropagation neural network is examined against various combinations of transfer and training functions. For this purpose, the performances of three transfer functions in tandem with fourteen training functions used for backpropagation training of neural network for recognition of power quality (PQ) disturbance signatures are investigated.

The paper is organized as follows. Section II introduces CWT based algorithm for generating the signatures of PQ disturbances. The proposed work including details of neural network model and architecture along with transfer and training functions used for recognition of PQ disturbances are presented in section III. In section IV, the simulation results are presented and effect of selection of different transfer and training functions on the performance of neural network is discussed. The conclusion is drawn in section V.

II. THE CWT BASED ALGORITHM FOR GENERATING SIGNATURES OF PQ DISTURBANCES

The Continuous Wavelet Transform (CWT) is a time-frequency representation of signals. It is convolution of a signal s(t) with a set of functions, which are generated by translations and dilations of a main function. The main function is known as the mother wavelet and the translated or dilated functions are called wavelets. Mathematically, the CWT of a signal x(t) is given by (1):

$$CWT_{\psi}\, x(a,b) = W_x\,(a,b) = \int_{-\infty}^{\infty} x(t)\Psi^{*}_{a,b}(t)dt \qquad (1)$$

Where,

$$\psi_{a,b}(t) = |a|^{1/2}\,\Psi\left(\frac{t-b}{a}\right) \qquad (2)$$

Here, “b” is the time translation and “a” is the dilation (scale) of the wavelet and both are real numbers.

As shown in Fig. 2, the CWT coefficients of pure signal are calculated and taken as reference and the CWT coefficients of PQ disturbance signal are subtracted from it, which gives the difference coefficient matrix (DCM). On careful investigation of DCM, it reveals that the scale/row wise value of coefficients of DCM follow a particular pattern for a particular PQ disturbance according to the location of the disturbance. Therefore, the coefficients of each row of DCM are summed, which gives a matrix named as the unique feature matrix (UFM).

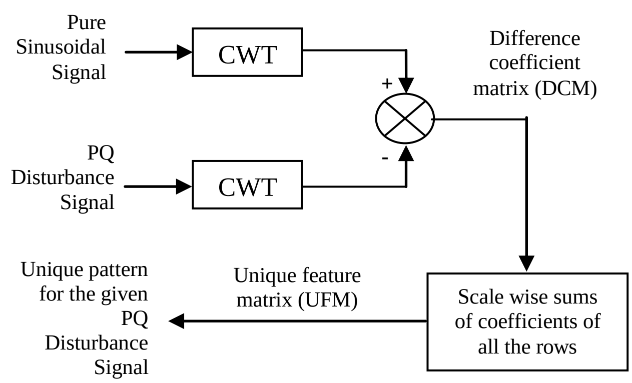


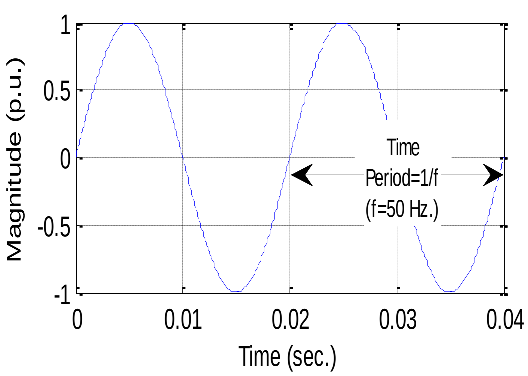
Figure 2. Algorithm for generating signatures of PQ disturbances

It is shown that the UFM posses a unique feature for a particular PQ disturbance. This feature on plotting gives unique pattern for a particular type of PQ disturbance. These unique patterns can be treated as signatures of their respective PQ disturbance. it is also observed that the size of this pattern varies in proportion to the magnitude of the disturbance but the shape remains the same for one particular type of PQ disturbance.

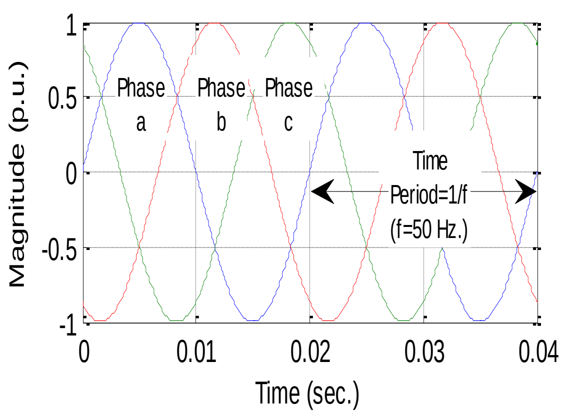
Based on the above mentioned algorithm, signature of various power quality disturbances namely sag, swell, transient, harmonics, and flicker were generated. As an example only one type of PQ disturbance i.e. sag is explained here in length, whereas other PQ signatures are obtained after replicating the same process.

Pure sinusoidal signal is generated synthetically and shown in Fig. 3 and can be visualize as matrix P of dimension of 1x512 is obtained for the pure signal. Similarly, the disturbance signals of sag with different magnitude (i.e. sag-0.2 pu, sag-0.6 pu., sag-0.8 pu. and interruption) are generated by programming in MATLAB as shown in Figs. 4(a), 4(b), 4(c) and 4(d) and accordingly matrixes of dimension of 1x512 are obtained for each of the disturbance signal of sag with different magnitude. Continuous wavelet transform (CWT) coefficient matrix of pure sinusoidal signal and each disturbance signal of sag is obtained as mentioned in algorithm. The dimensions of these matrixes are 64x512. Then, the difference coefficient matrix (DCM) is obtained by subtracting continuous wavelet transform coefficient matrix of each of the disturbance signal of sag from corresponding matrix of pure sinusoidal signal, which gives dimension of 64x512 for DCM. Thereafter, unique feature matrix (UFM) is calculated by summing coefficients of each row of DCM, which gave UFM of dimensions of 64x1. These 64 data's are plotted with respect to their row number (i.e. scales of CWT) and the patterns as shown in Figs. 4(e), 4(f), 4(g) and 4(h) are obtained for their respective disturbance signals (i.e. sag-0.2 pu, sag-0.6 pu., sag-0.8 pu. and interruption).

It is clear, that the shape of all the patterns is same as shown in Figs. 4(e), 4(f), 4(g) and 4(h). These patterns are plotted on a single graph as shown in Fig. 5 and it reveals that the shape of all the patterns is exactly same. This unique shape, as shown in Fig. 6 can be taken as the signature for the sag and interruption disturbance of PQ. Replicating the same process signature of remaining power quality disturbances namely swell, transient, harmonics, and flicker were generated as shown in Fig. 7.



(a)



(b)

Figure 3. Pure sinusoidal waveform: (a) Single-phase and, (b) Three-phase

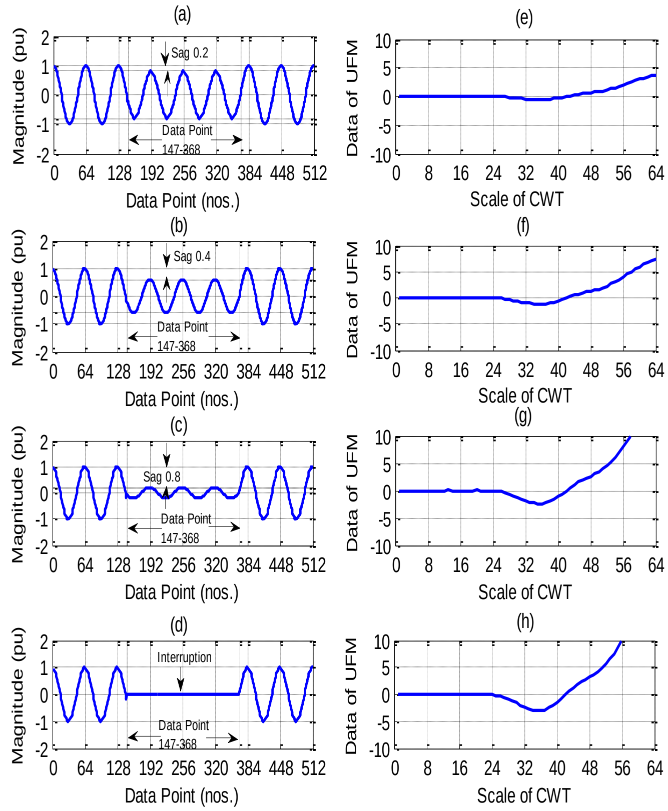


Figure 4. (a) sag - 0.2 pu, (b) sag - 0.4 pu, (c) sag - 0.6 pu, (d) interruption, (e) pattern of sag - 0.2 pu, (f) pattern of sag - 0.4 pu, (g) pattern of sag - 0.6 pu, and (h) pattern of interruption.

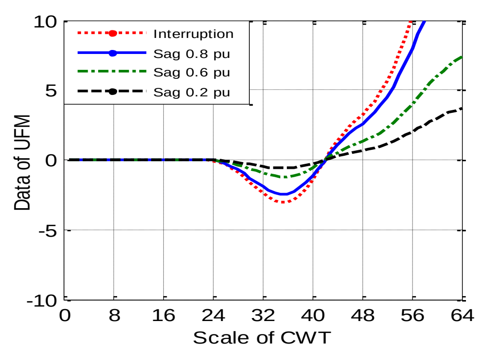


Figure 5. Patterns for interruption and sag - 0.2, 0.6 and 0.8 pu.

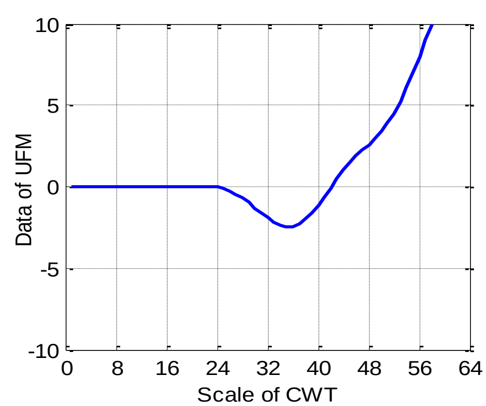


Figure 6. Signature for interruption and sag

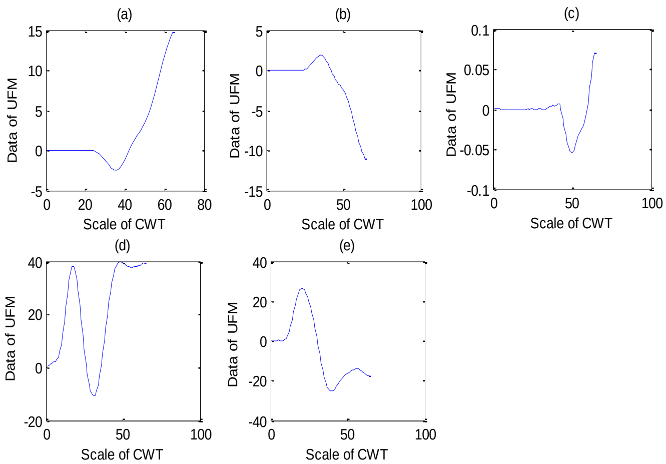


Figure 7. (a) Signature for sag and interruption, (b) signature for swell, (c) signature for transient (d) Signature for harmonics, and (e) signature for flicker.

III. The Proposed Work

A. Artificial Neural Network

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems like brain process information. The key element of this paradigm is the unique structure of the information processing system. It is composed of a large number of highly interconnected processing elements called neurons, which work in unison to solve specific problems.

Neural networks, with their extraordinary ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an “expert”, which can analyze the information given to it. There are many advantages of ANNs including their adaptive learning, self-organization of information, real time operation, and fault tolerance via redundant information coding etc. The development and profusion of high speed computers has made artificial neural network to become an increasingly popular research subject. It has been applied to almost all areas of life. The idea of neural network instigated from the most fascinating organ in the human body, the brain. The human brain consists of billions of basic units called neurons for which the structure is illustrated in Fig. 8.

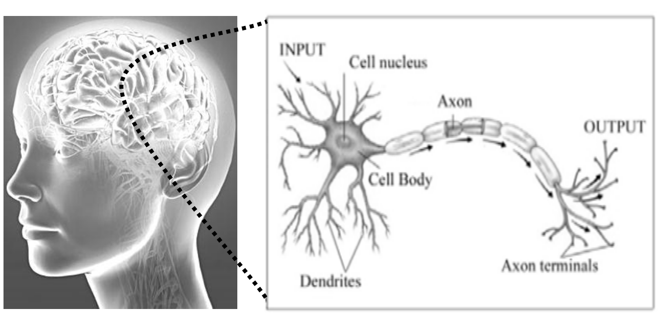


Figure 8. The structure of neuron

The neuron consists of dendrites, a cell body and an axon connecting to axon terminals. The information or inputs received by the brain are transferred into the cell body through dendrites. The cell body acts as the processing unit, where all the learned information is then transferred into outputs via axon. The muscles or other parts of the body receive the outputs for actions via the axon terminals.

B. Neural Network Model

The concept to form a mathematical model for neural network was first studied in 1943 by McCulloch and Pitts. In a simple model, one hidden layer feedforward network is considered with inputs x1,..., xi and output yk. Each input is given its own synaptic weight. The weights are then transferred into the hidden layer, which consists of a number of hidden neurons. Each neuron performs a weighted summation of the inputs and then passes a nonlinear activation function. The output of the network is given by (3).

$$y_k = \phi\left(\sum w_{ki} x_i\right) \tag{3}$$

Where,

yk = network’s output, wki = synaptic weight between output of network k and input of neuron i, vk = (∑wkixi) = activation potential of neuron i (net input), Ø = activation function, Ɵ = threshold factor.

Each neuron has an activation function associated with it (often times all neurons have the same activation function). A list of common activation functions and their names can be found in Table I.

TABLE I. Activation Functions

Name	Formula
Identity	Ø(x) = x
Sigmoid	Ø(x) = 1/(1+e-x)
Tanh	Ø(x) = (ex-e-x) / (ex + e-x)
Step	Ø(x) = -1 if x<0 = 1 if x ≥0

The threshold ‘Ɵ’ is a factor which is used in calculating the activations of the given net. Based on the value of threshold the output may be calculated, i.e. the activation function is based on the value of Ɵ. For, example, the activation functions may be,

$$y = Ø\ (net) = +1 \text{ if } net \geq Ɵ;$$
$$= -1 \text{ if } net < Ɵ$$

C. Neural Network Architecture

The behavior of the neural network depends mainly on the interaction between the different neurons. The basic architecture of neural network consists of three types of neuron layers viz. Input, Hidden, and Output.

In this work, one of the most common neural network architectures i.e. the feedforward backpropagation neural network is examined. This neural network architecture has been successfully applied to solve many problems and also extensively applied for the analysis of PQ disturbances [10]. The first term, "feedforward" depicts that in this architecture neurons are only connected foreword such that each layer of the neural network contains connections to the next layer (for example, from the input to the hidden layer), but there are no connections back. The term "backpropagation" is used to describe a form of supervised training. In supervised training method, the network is provided with both sample inputs and anticipated outputs. The anticipated outputs are compared with the actual outputs for given input and error is calculated. Using the calculated error, the backpropagation training algorithm then adjusts the weights of the various layers backwards from the output layer to the input layer.

Fig. 9 depicts this Multiple Layer Preceptorn (MLP) architecture with two layers of neurons, which is implemented in this paper. The theorem of Hornik-Stinchcombe-White states that "a neural network with two layers is sufficient to make a precise and desirable approximation of a continuous mapping marked with a finite dimensional space to another, provided the sufficiency of neurons in the hidden layers". In our case, this holds true; as only two layers give excellent results. The neural network toolbox of MATLAB provides the graphical user interface (GUI) based neural network pattern recognition tool "nntool", which is employed in this work for recognition of various signatures of PQ, as generated above. 64 neurons are taken in input layer because size of each input signature matrix is 64x1. Since, 05 types of PQ disturbances are considered for recognition purpose in this work; so, there are 05 neurons in the output layer. The number of neurons opted for hidden layer are 20. Since, we want to confirm the outputs of neural network between 0 and 1 for recognition purpose, so sigmoid transfer function at output layer is most appropriate in this case.

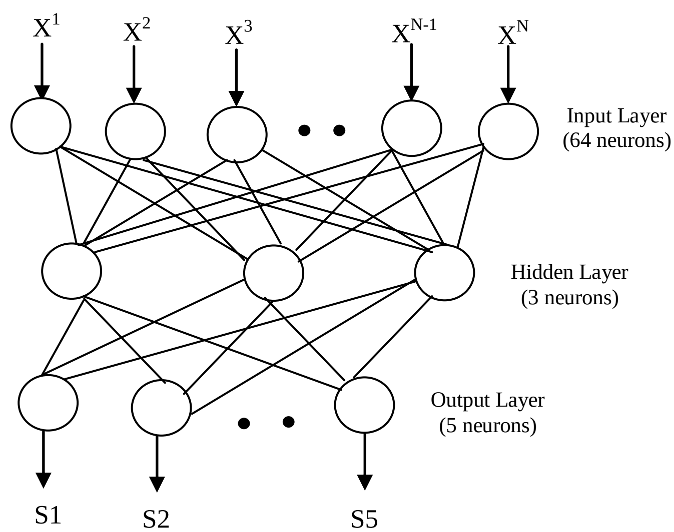


Figure 9. Architecture of the two layer feedforward neural network

D. Neural Network transfer and training functions

The Multiple Layer Preceptorn (MLP) architecture, as mentioned above with a feedforward back propagation algorithm is used in this work for recognition of signatures of PQ disturbances. The parameters investigated were the transfer and training functions. Transfer and training functions are mathematical procedures used to automatically adjust the network's weights and biases. Three transfer functions namely log-sigmoid (logsig), linear (purelin), and hyperbolic tangent sigmoid (tansig) were considered for investigation against 14 training functions as mentioned in Table II.

After the completion of the training stage, the trained networks were tested and the performance of the networks was evaluated using the criteria of Lowest MSE, number of epochs, convergence time, and accuracy.

TABLE II. THE LIST OF INVESTIGATED TRAINING ALGORITHMS AND THEIR DEFINITIONS

S. no.	Training Function	Definition of the Algorithm
1	BFG	Quasi-Newton backpropagation
2	BR	Bayesian Regularization Backpropagation
3	CGB	Conjugate Gradient Backpropagation with Powell-Beale Restarts
4	CGF	Conjugate Gradient Backpropagation with Fletcher-Reeves Update
5	CGP	Conjugate Gradient Backpropagation with Polak-Ribiére Update
6	GD	Gradient Descent backpropagation
7	GDM	Gradient Descent with Momentum backpropagation
8	GDA	Gradient Descent with Adaptive learning rate backpropagation
9	GDX	Gradient Descent with Momentum and Adaptive Learning Rate Backpropagation
10	LM	Levenberg-Marquardt backpropagation
11	OSS	One Step Secant backpropagation
12	NR	Random order incremental training with learning functions
13	NRP	Resilient backpropagation
14	NSCG	Scaled Conjugate Gradient Backpropagation

IV. RESULTS AND DISCUSSION

The formulated network was trained with 90 samples of 05 types of PQ signatures disturbances (i.e. sag, swell, transient, harmonics, and flicker) and then it was simulated for testing with 100 new samples of signatures of PQ disturbances, which were not used in training. The performance of the trained networks were investigated using the performance parameters including lowest MSE, number of epochs, convergence time, and accuracy as shown in Table III. A performance order between 1 and 5 is assigned, where "1" refers to the highest performance and the "5" refers to the lowest performance. It may be noted that accuracy below 80% is marked in grey in Table III and performance order is not assigned to it. Further, for same value of performance parameter, the same performance order is assigned.

It is clear from Table III that the overall performance of "tansig" transfer function is better than "logsig" and "purelin" transfer functions. In addition to this, it also reveals that the accuracy of "BR" training function was 100 % with all the three types of transfer functions and the accuracy of "BFG" and "NSCG" was also 100% with "tansig" and "logsig" transfer function.

TABLE III. COMPARISON OF THE PERFORMANCES OF TRANSFER AND TRAINING FUNCTIONS FOR RECOGNITION OF PQ DISTURBANCE SIGNATURES

S. no.	**Trg. Func-tion**	**Lowest MSE**	**Epo-chs**	**Time**	**Accu-racy %**	**Performance Order**			
						Lowest MSE	**Epo-chs**	**Time**	**Acc-uracy**
Transfer function: Log SIG									
1	BFG	0.00564	46	2.44	100	**1**	2	3	**1**
2	**BR**	0.68097	175	4.39	100	5	5	5	**1**
3	CGB	$7.064e^{-009}$	113	0.04	40				
4	CGF	0.13351	11	0.00	30				
5	CGP	0.81956	11	0.00	40				
6	GD	0.23948	1000	0.15	40				
7	GDM	0.15352	1000	0.15	40				
8	GDA	0.07491	129	0.02	50				
9	GDX	0.14733	79	0.01	40				
10	LM	0.03730	47	0.35	100	4	3	4	**1**
11	OSS	0.64338	18	0.01	60				
12	NR	055934	100	0.22	70				
13	NRP	0.02013	29	0.01	90	3	**1**	**1**	2
14	NSCG	0.01783	63	0.02	100	2	4	2	**1**
Transfer function: PURELIN									
1	BFG	0.27778	21	1.24	20				
2	BR	5.14130	14	0.24	100	2	**1**	2	**1**
3	CGB	0.28852	15	0.00	20				
4	CGF	0.44410	4	0.00	0				
5	CGP	0.29448	12	0.00	0				
6	GD	0.20005	26	0.00	20				
7	GDM	0.27797	1000	0.15	20				
8	GDA	0.36649	10	0.00	0				
9	GDX	0.58148	85	0.01	0				
10	LM	0.17778	24	0.11	20				
11	OSS	0.29059	45	0.01	40				
12	NR	0.72986	100	0.22	40				
13	NRP	0.16667	11	0.00	20				
14	NSCG	0.07683	80	0.02	80	**1**	2	**1**	2
Transfer function: Tan Sig									
1	BFG	7.064e - 10	43	2.34	100	**1**	2		**1**
2	BR	0.32318	140	3.37	100				**1**
3	CGB	0.32254	32	0.01	100		**1**	2	**1**
4	CGF	0.09610	19	0.00	40				
5	CGP	0.00045	104	0.03	100	2			**1**
6	GD	0.07739	1000	0.15	40				
7	GDM	0.17205	1000	0.16	60				
8	GDA	0.07343	101	0.01	60				
9	GDX	0.05239	155	0.02	100	5		3	**1**
10	LM	0.08242	70	0.45	70				
11	OSS	0.00932	67	0.02	100	3	3	3	**1**
12	NR	0.23523	100	0.22	100		5		**1**
13	NRP	0.04198	32	0.00	90	4	**1**	**1**	
14	NSCG	0.01170	70	0.01	100		4	2	**1**

V. CONCLUSION

A study for the selection of apposite combination of training and transfer function for recognition of PQ disturbance signatures is presented in this paper. The comparison of performance of feedforward back propagation neural network against three transfer functions in tandem with fourteen training functions revealed that the overall performance of “tansig” transfer function was superior among three transfer functions namely “logsig”, “purelin”, and “tansig” and the accuracy of “BR” training function was 100% with all the three types of transfer functions. This work can help researchers in the selection of transfer and training functions while addressing similar recognition problems based on neural networks.

REFERENCES

[1] IEEE Standards Board, “IEEE Std. 1159-1995, “IEEE Recommended Practice for Monitoring Electric Power Quality”, IEEE, Inc., New York, June, 1995.

[2] Mark F. McGranaghan and Surya Santoso, “Challenges and Trends in Analyses of Electric Power Quality Measurement Data”, EURASIP Journal on Advances in Signal Processing, vol. 2007.

[3] Moises V. Ribeiro, Jacques Szczupak, M. Reza Iravani, Irene Y. H. Gu, P. K. Dash, and Alexander V. Mamishev, “Emerging Signal Processing Techniques for Power Quality Applications”, EURASIP Journal on Advances in Signal Processing, vol. 2007.

[4] W. R. Anis Ibrahim, and M. M. Morcos, “Artificial intelligence and advanced mathematical tools for power quality applications: a survey,” IEEE Trans. on Power Delivery, vol. 17, no. 2, Apr. 2002, pp. 668-673.

[5] S. Santoso, E.J Powers, and W.M Grady, "Electric power quality disturbance detection using wavelet transform analysis," Proc. of the IEEE-SP Int. Symp. on Time-Frequency and Time-Scale Analysis, 1994, pp. 166-169.

[6] D. Gonzalez, J. Balcells, and J. T. Bialasiewicz, “Exploration of application of continuous wavelet transform to power quality analysis,” Proc. of the IEEE International Symposium on Industrial Electronics, 2008, pp. 2242-2246.

[7] J.D. Taboada, J. C. Cabrera, G. Ramos, and M.T. Torres, “Pattern Recognition of Phenomena Associated to Power Quality Using Neural Networks”, IEEE / PES Trans. & Distri. Con. & Exposition: Latin America, 2004, pp. 1-5.

[8] H.K. Siu, and H.W. Ngan, “Automatic Power Quality Recognition System using Wavelet Analysis,” 2004 IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies (DRPT 2004), Hong Kong, April 2004, pp. 311-315.

[9] Duarte G. Cesar, V. G .Valdomiro, and O. P. Gabriel, “Automatic Power Quality Disturbances Detection and Classification Based on Discrete Wavelet Transform and Artificial Intelligence,” 2006, IEEE, pp. 1-6.

[10] R. A. Gupta, Rajesh Kumar, Manoj Gupta, “Obtaining Patterns for Identification of Power Quality Disturbances Using Continuous Wavelet Transform”, Proc. of the International Conference & Workshop on Emerging Trends in Technology 2011 (ICWET 2011), India, vol. 1, Feb. 2011, pp. 395-399.

[11] R. A. Gupta, Rajesh Kumar, Manoj Gupta, “Continuous Wavelet Transform Applied to Generate Patterns of Phenomena Associated to Power Quality” International Journal of Computer Application, no. 4, Article 7, 2011. (in press).

[12] Manoj Gupta, Rajesh Kumar, R. A. Gupta, “A Neural Network Based System for Recognition of Power Quality Disturbances” CiiT-International Journal, Tamilnadu, India. June, 2011. (in press)

[13] Manoj Gupta, Rajesh Kumar, R. A. Gupta, “Neural Network Based Indexing and Recognition of Power Quality Disturbances” Indonesia Journal of Electrical Engineering, Indonesia, vol. 9, no. 2, August 2011. (in press).

[14] Kasthurirangan Gopalakrishnan, “Effect of Training Algorithms on Neural Networks Aided Pavement Diagnosis” International Journal of Engineering, Science and Technology Science and Technology, vol. 2, no. 2, 2010, pp. 83-92.

[15] Zulhadi Zakaria, Nor Ashidi Mat Isa, and Shahrel A. Suandi, “A Study on Neural Network Training Algorithm for Multiface Detection in Static Images” World Academy of Science, Engineering and Technology, 62, 2010, pp.170-173.

[16] Dheerendra Vikram Singh, Govind Maheshwari, and Ritu Shrivastav, “Neural Network–Comparing the Performances of the Training Functions for Predicting the Value of Specific Heat of Refrigerant in Vapor Absorption Refrigeration System” International Journal of Computer Applications, vol. 18, no. 4, March 2011, pp.1-5.