# PERFORMANCE RANKING OF ARTIFICIAL NEURAL NETWORK LEARNING ALGORITHMS IN SOLAR RADIATION FORECAST

D. A. Fadare*, T. Olugasa and A. Falana

*Department of Mechanical Engineering, Faculty of Technology, University of Ibadan, Ibadan, Nigeria*

## ABSTRACT

Artificial Neural Networks (ANNs) are a promising alternative to conventional tools in modeling and prediction of complex and non-linear parameters. However, the selection of appropriate network parameters for optimum performance pose application challenges In this study, the modeling and predictive performances of six backpropagation learning algorithms: Levenberg-Marquardt (LM), BFGS Quasi-Newton (BFG), Resilient Backpropagation (RP), Fletcher-Powell Conjugate Gradient (CGF), Variable Learning Rate Backpropagation (GDX) and Bayesian Reglarization (BR) in solar radiation forecast were investigated.

Multilayer perceptron (MPL) neural network with five, ten and one neuron(s) in the input, hidden and output layers, respectively was designed with MATLAB® neural network toolkit and trained with the six learning algorithms using the daily global solar radiation data of Ibadan (Lat. $7.4^0$ N; Long. $3.9^0$ E; Alt. 227.2m), Nigeria. The network performance was ranked based on the number of iterations required for convergence, and coefficient of correlation (*r-value*), mean square error (MSE) and mean absolute percentage error (MAPE) between the actual and predicted values of the training and testing datasets. Results showed that the LM and BR learning algorithms are the two best algorithms to be considered for use in modeling and forecasting of solar radiation data.

***Keywords***: Artificial neural network; Learning algorithm; Performance evaluation, Modeling and forecasting; Solar radiation.

## 1.0 INTRODUCTION

Artificial Neural Networks (ANNs) are a promising alternative to conventional tools in modeling and prediction of complex and non-linear system variables such as pattern recognition and function approximation. They are information processing paradigms inspired by biological nervous systems. ANNs, like human-being, can learn by example, associate data and recall information. As learning in biological systems involve adjustments of synaptic connections that exist between the neurons, so are ANNs made up of simple processing units which are linked by adjustable weight connections to form structures that are able to learn relationships between sets of variables. After being sufficiently trained, ANNs can perform predictions at very high speed (Mellit et al, 2006). They are able to deal with non-linear problems such as multivariate time series prediction. However, the accuracy of the model is a function of the network parameters utilised. Hence, the selection of appropriate network parameters for optimum network performance poses great challenges in application of ANN models.

* Correspondence author. Tel.: +234 802 3838 593; E-mail address: fadareda@yahoo.com (D. A. Fadare)

The network designer chooses the network parameters such as topology, performance function, learning algorithm, learning rate, etc based mainly on previous experience or on iterative processes.

Solar radiation forecast constitutes an essential component of weather forecast and a major design input parameter in solar energy application systems. For efficient conversion and utilization of solar energy resource, accurate, detailed and timely knowledge of the available solar radiation is required. Different ANN models with wide range of network structure and learning algorithm have been developed by many researchers and applied with varying degrees of success to modeling and forecasting of solar radiation data (Mellit et al., 2006; Elizondo et al., 1996; Al-Alawi and Al-Hinai, 1998; Fadare, 2009; Fadare and Olugasa 2009).

In particular, Fadare and Olugasa (2009) reported the effect of network structure such as type and number of neurons in the input and hidden layer, and number of layers in the hidden layer on the network performance for time series forecast of daily solar radiation data. The effect of network parameters on the performance of neural network in forecasting of electricity demand has been reported by Azadeh and Behshtipour (2008), while the effect of network learning algorithm on network performance for modeling and prediction of a wide range of engineering and medical data has been reported by Demuth and Beale (2000). The results of their study showed that networks trained with different learning algorithms performed better for different datasets depending on the nature of the data. Thus, indicating thatthe selection of the best learning algorithm for network training is data specific.

Some of the commonly used algorithms for function approximation model are: Hebb's rule, Hopfield law, delta rule, Gradient Descent rule, Kohonen's learning law and the backpropagation algorithm. The backpropagation learning algorithms are the most used for training the multilayer perceptron (MLP) networks. It has been shown to perform adequately in many applications (Gardner and Dorling, 1998; Foresee and Hagan, 1997). Some examples of backpropagation algorithms are: Batch Gradient (GD), One-step-secant Algorithm (OOS), Levenberg-Marquardt (LM), BFGS Quasi-Newton (BFG), Bayesian Regularization (BR), Resilient Backpropagation (RP), Fletcher-Powell Conjugate Gradient (CGF), Variable Learning Rate Backpropagation (GDX), etc (Demuth and Beale, 2000; MacKay, 1992a; 1992b).

The backpropagation learning algorithms are supervised learning method, implemented based on the Delta rule (Gill et al., 1981). The summary of the backpropagation technique is as follows: A training sample is presented to the network; the network output is compared to the desired output from that sample; what the output should have been, and a scaling factor of how much lower or higher the output must be adjusted to match the desired output- the local error for each neuron is calculated and the weights of each neuron is adjusted to lower the local error. There exist a considerable volume of literature on the effect of learning algorithm on network performance for modeling and prediction of wide range of datasets. However, it appears that no study has reported the effect of learning algorithm on the network performance for solar radiation forecast. The aim of this study was to investigate the effect of six backpropagation learning algorithms on modeling and predictive performances of MLP network in solar radiation forecast with a view to achieving error minimisation.

**MATERIALS AND METHODS**

The basic steps adopted in the study are: (1) Selection of optimum network structure; (2) Design of the neural network model; (3) Collection of sample dataset; (4) Pre-processing and partitioning of data; (5) Training of the selected neural network with different learning algorithms; and (6) Testing of the neural network performance.

**Selection of optimum network structure**

The selection of the optimum network structure was based the previous study of the authors on the effect of network structure on the network performance for forecasting of solar radiation data (Fadare and Olugasa, 2009). Based on this study, the optimum network structure with 5 neurons in the input layer, 10 neurons in the single hidden layer, and 1 neuron in the output layer was selected. For this structure, daily solar radiation values for 5 previous days (*t-5, t-4, t-3, t-2 and t-1*) were used as input parameters to forecast the value for the current day (*t*). Hyperbolic tangent sigmoid transfer function '*tansig*' was used in the hidden layer, while linear transfer function '*purelin*' was used in the output layer.

**Design of the ANN model**

Neural Network Toolbox for MATLAB® was used to design the model. The structure of the MPL network model is shown in Figure 1.

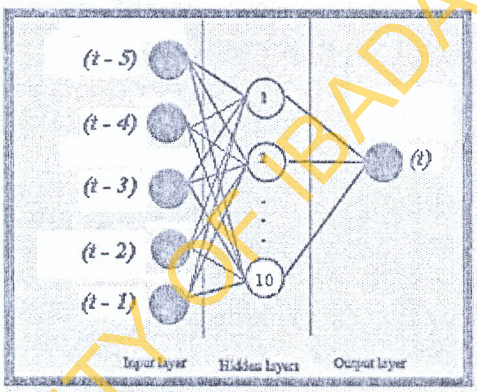


Fig. 1: Structure of the selected MLP network model

**Collection of sample dataset**

As case study, daily global solar radiation data for Ibadan (Lat. $7.43^0$ N; Long. $3.9^0$ E; Alt. 227.2m), Nigeria, for the period of 1984 to 2007 (24 years) obtained the meteorological ground station was used for training and testing the performance of the network.

**Pre-processing and partitioning of data**

The values of the dataset were normalized to range between 0 and 1. The normalized dataset was then partitioned into two subsets: training and testing datasets. The dataset for the period of 1984 to 2006 (23 years) was used for training the network, while the dataset for the 2007 was used as the testing dataset. The training dataset was further sub-divided randomly into two subsets: training (75%) and validation (25%) dataset.

**Training of the neural network model**

The network model was trained with six different backpropagation learning algorithms: Levenberg-Marquardt (LM), BFGS Quasi-Newton (BFG), Resilient Backpropagation (RP), Fletcher-Powell Conjugate Gradient (CGF), Variable Learning Rate Backpropagation (GDX) and Bayesian Reglarization (BR). The learning rate was set at 0.8, performance function was set as MSE with a threshold equal to 0.001 and maximum number of iteration was set at 1000. The training dataset was used for computing the gradient and updating the network weights and biases, the validation dataset was used to check the network from being over trained. The error on the validation dataset was monitored during the training process. The validation error will normally decrease during the initial phase of training, as does the training dataset error. However, when the network begins to over fit the data, the error on the validation set will typically begin to rise. When the validation error increases for a specified number of iterations, the training is stopped, and the weights and biases at the minimum of the validation error are returned. The networks with the different learning algorithms were trained 30 times with different randomly selected initial weights and biases.

**Testing of the ANN model**

The accuracy of the network model trained with the different learning algorithms for prediction of the testing dataset (solar radiation data for 2007) was based on the mean, maximum, minimum and standard deviation of the number of iterations required for convergence during the training process, and Pearson's coefficient of correlation (*r-value*), mean square error (MSE) and mean absolute percentage error (MAPE) between the actual and predicted values of the training and testing dataset of 30 different training sections. The performance functions (*r-value*, MSE and MAPE) were computed by equations 1-3 (Oyawale, 2006):

$$r = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{n\sigma_x\sigma_y} \quad (1)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2 \quad (2)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_i-Y_i}{X_i}\right| \quad (3)$$

Where, $X_i$ and $Y_i$ are the actual and predicted values, $X$ and $Y$ are the means of the actual and ANN predicted values, $\sigma_x$ and $\sigma_y$ are the standard deviations of the actual and predicted values and n is the number observations.

## RESULT AND DISCUSSION

The performance of the network trained with the six learning algorithms for 30 training trials was determined based on the number of iterations required for convergence during the training process, and coefficient of correlation (*r-value*), mean square error (MSE) and mean absolute percentage error (MAPE) between the actual and predicted values of the training and testing datasets. The mean, maximum, minimum and standard deviation of the number of iterations required for convergence during the training process is presented in Table 1. The number of iterations required for convergence is a measure of the speed and computer time required for training the network. Learning algorithm with small number of iterations is generally preferred due to short computation time required in training the network.

As shown in Table 1, the LM algorithm was ranked the fastest learning algorithm with the smallest mean number of iterations of 10.75, followed by other algorithms in the order: BR (15.75), RP (19.50), BFG (21.75), CGF (26.25) and GDX (79.00). In this case, LM algorithm was 7.35 times faster in convergence than GDX algorithm. The capability of LM algorithm to reach desired error in short time frame in modeling of wide range of data has been reported by many researchers (Demuth and Beale, 2000; Azadeh and Behshtipour, 2008; Mellit et al., 2006; Elizondo et al., 1996; Al-Alawi and Al-Hinai, 1998; Fadare, 2009; 2010).

Table 1: Performance ranking of the learning algorithms based on the number of iterations.

| Learning Algorithm | Performance ranking | No. of iterations | | | |
|---|---|---|---|---|---|
| | | Mean | Minimum | Maximum | Standard Deviation |
| BFG | 4 | 21.75 | 14 | 36 | 10.15 |
| BR | 2 | 15.75 | 8 | 20 | 5.32 |
| CGF | 5 | 26.25 | 14 | 42 | 12.76 |
| GDX | 6 | 79.00 | 16 | 100 | 42.00 |
| LM | 1 | 10.75 | 9 | 13 | 2.06 |
| RP | 3 | 19.50 | 16 | 25 | 3.87 |

The comparison between the predicted values for different learning algorithms and actual values of solar radiation for testing dataset is shown in Figures 2. The mean, maximum, minimum and standard deviation of the correlation (*r-value*) between actual and ANN predicted values based on the testing dataset is given in Table 2. The *r-value* based on the testing dataset is the measure of the generalization capability of the network to predict new dataset that is not used in training the network.

Thus, network with high *r-value* (for testing dataset) has high predictive accuracy than network with low value. Table 2 shows that BR algorithm had the highest *r-value* (for testing dataset) with mean value of 0.79, followed by other algorithm in the order: LM (0.70), BFG (0.63), RP (0.57), CGF (0.50) and GDX (0.33). Therefore, the network trained with BR algorithm has the highest generalization or forecasting capability for solar radiation data compared to other learning algorithms investigated.

Table 2: Performance ranking of the learning algorithms based on the coefficient of correlation (*r-value*) for the testing dataset

| Learning Algorithm | Performance ranking | *r-value (testing dataset)* | | | |
|---|---|---|---|---|---|
| | | Mean | Minimum | Maximum | Standard Deviation |
| BFG | 3 | 0.63 | 0.41 | 0.77 | 0.16 |
| BR | 1 | 0.79 | 0.79 | 0.79 | 0.00 |
| CGF | 5 | 0.50 | 0.08 | 0.73 | 0.30 |
| GDX | 6 | 0.33 | -0.01 | 0.73 | 0.25 |
| LM | 2 | 0.70 | 0.66 | 0.73 | 0.14 |
| RP | 4 | 0.57 | 0.15 | 0.77 | 0.29 |

The ranking of the different learning algorithms based on the coefficient of correlation (*r-value*) of the actual and predicted values of the training dataset is given in Table 3. As shown in Table 3, LM algorithm had the highest mean *r-value* of 0.71, while the GDX algorithm had the lowest value of 0.28. The *r-value* for the training dataset is the measure of the network ability to learn relationships between input/output dataset used in training the network. Hence, the network trained with the LM learning algorithm showed the highest capability of modeling the relationships between the input and output variables in solar radiation forecast (Table 3) and also the fastest converging training algorithm (Table 1).

Tables 4 and 5 show the ranking of the learning algorithms based on the mean square error (MSE) for the training and testing datasets, respectively. Based on the training dataset,

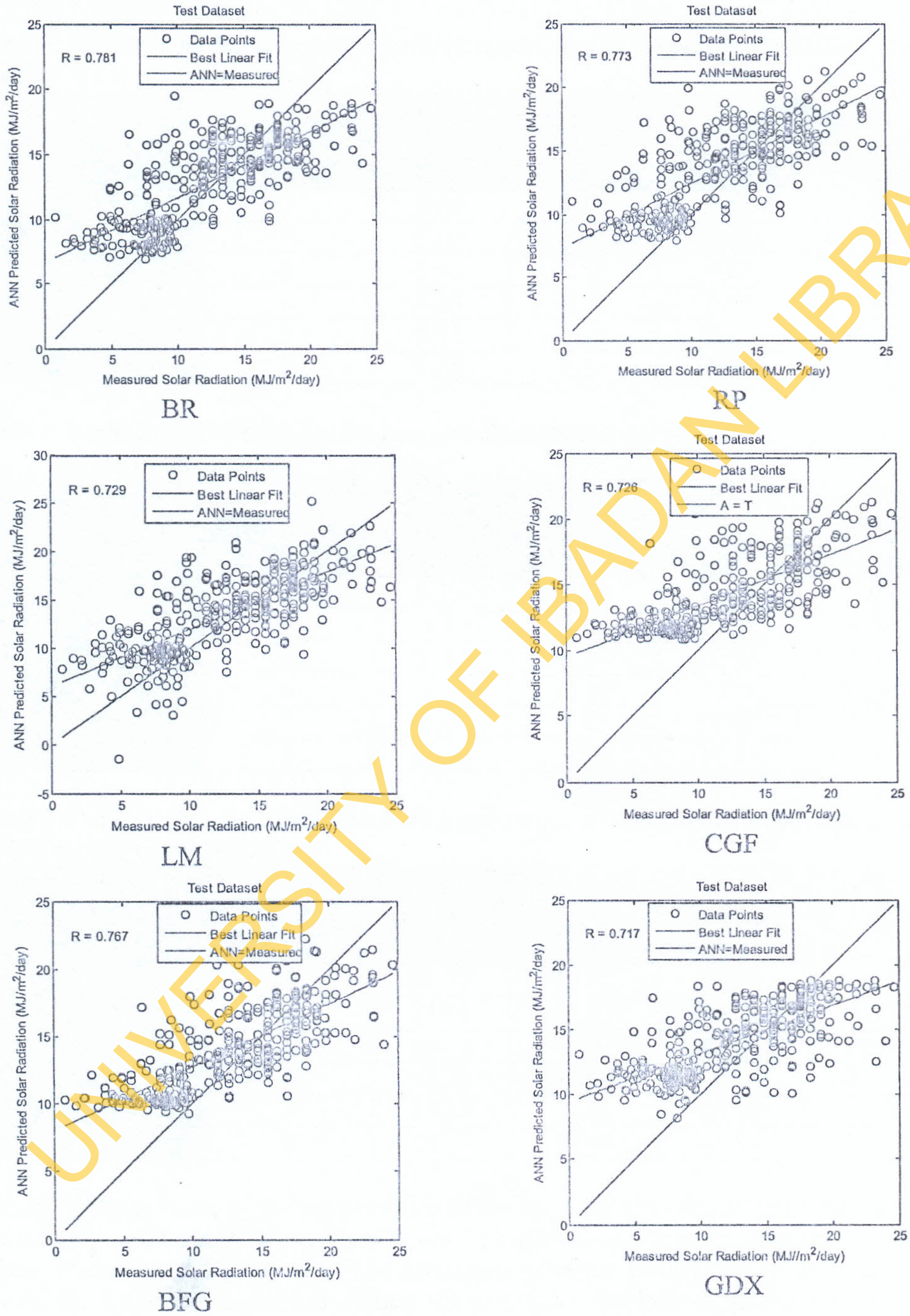


Figure 2: Comparison of ANN prediction values with different learning algorithm and actual solar radiation for testing dataset

the network trained with LM algorithm was ranked the first the lowest average MSE of 8.36, while the network trained with GDX algorithm was ranked the last with the highest value of 28.81 (Table 4). For the testing dataset (Table 5), the network trained with BR algorithm gave the lowest average MSE (13.06), while the network trained with GDX algorithm gave the highest value of 48.98.

Table 3: Performance ranking of the learning algorithms based on the correlation coefficient (r-value) for the training dataset

| Learning Algorithm | Performance ranking | *r-value (training dataset)* | | | |
|---|---|---|---|---|---|
| | | Mean | Minimum | Maximum | Standard Deviation |
| BFG | 2 | 0.64 | 0.57 | 0.68 | 0.05 |
| BR | 3 | 0.62 | 0.62 | 0.62 | 0.00 |
| CGF | 4 | 0.61 | 0.51 | 0.68 | 0.08 |
| GDX | 6 | 0.28 | -0.53 | 0.58 | 0.54 |
| LM | 1 | 0.71 | 0.65 | 0.75 | 0.04 |
| RP | 5 | 0.58 | 0.54 | 0.65 | 0.05 |

Table 4: Performance ranking of the learning algorithms based on the mean square error (MSE) for the training dataset

| Learning algorithm | Performance ranking | *MSE (training dataset)* | | | |
|---|---|---|---|---|---|
| | | Mean | Minimum | Maximum | Standard Deviation |
| BFG | 2 | 9.79 | 8.94 | 11.17 | 1.03 |
| BR | 4 | 11.10 | 10.77 | 11.79 | 0.46 |
| CGF | 3 | 10.31 | 8.90 | 12.24 | 1.49 |
| GDX | 6 | 28.81 | 11.26 | 79.94 | 34.10 |
| LM | 1 | 8.36 | 7.67 | 9.95 | 1.07 |
| RP | 5 | 11.15 | 9.70 | 11.84 | 0.99 |

Table 5: Performance ranking of the learning algorithms based on the mean square error (MSE) for the testing dataset

| Learning algorithm | Performance ranking | *MSE (testing dataset)* | | | |
|---|---|---|---|---|---|
| | | Mean | Minimum | Maximum | Standard Deviation |
| BFG | 3 | 20.10 | 12.49 | 33.00 | 8.31 |
| BR | 1 | 13.06 | 12.13 | 15.02 | 1.35 |
| CGF | 5 | 24.67 | 14.64 | 41.06 | 11.96 |
| GDX | 6 | 48.98 | 16.45 | 117.29 | 46.56 |
| LM | 2 | 17.26 | 13.41 | 23.06 | 4.53 |
| RP | 4 | 21.40 | 12.46 | 40.31 | 12.82 |



Similarly, the ranking of the performance of the learning algorithms based on the MAPE for the training and testing datasets are shown in Tables 6 and 7, respectively. As shown in the tables, the network trained with LM algorithm gave the lowest MAPE for training dataset (Table 6), while BR algorithm-trained network gave the lowest MAPE for testing dataset (Table 7). The network trained with GDX algorithm has the highest MAPE for both training and testing datasets. The LM algorithm has the lowest MSE and MAPE for the training dataset. Thus, indicating the high modelling accuracy of the network trained with LM algorithm, while the network trained with BR algorithm with the lowest MSE and MAPE for testing dataset showed a high predictive accuracy in solar radiation forecast.

Table 6: Performance ranking of the learning algorithms based on the mean absolute percentage error (MAPE) for the training dataset

| Learning algorithm | Performance ranking | *MAPE (training dataset)* | | | |
|---|---|---|---|---|---|
| | | Mean | Minimum | Maximum | Standard Deviation |
| BFG | 2 | 12.99 | 12.30 | 14.01 | 0.76 |
| BR | 5 | 15.29 | 14.83 | 16.16 | 0.59 |
| CGF | 3 | 13.53 | 12.42 | 15.32 | 1.30 |
| GDX | 6 | 22.12 | 13.57 | 44.79 | 15.14 |
| LM | 1 | 12.29 | 11.33 | 13.47 | 0.93 |
| RP | 4 | 14.33 | 13.12 | 15.34 | 0.94 |

Table 7: Performance ranking of the learning algorithms based on the mean absolute percentage error (MAPE) for the testing dataset

| Learning algorithm | Performance ranking | *MAPE (testing dataset)* | | | |
|---|---|---|---|---|---|
| | | Mean | Minimum | Maximum | Standard Deviation |
| BFG | 3 | 20.12 | 17.89 | 23.02 | 2.35 |
| BR | 1 | 18.61 | 17.92 | 19.93 | 0.91 |
| CGF | 5 | 21.60 | 18.37 | 25.56 | 3.39 |
| GDX | 6 | 30.33 | 19.41 | 54.14 | 16.08 |
| LM | 2 | 20.08 | 18.57 | 21.63 | 1.58 |
| RP | 4 | 21.20 | 17.25 | 27.13 | 4.31 |

## CONCLUSION

In this study, the modeling and predictive performances of a multilayer percetron artificial neural network trained with six backpropagation learning algorithms for solar radiation forecast were ranked based on the highest r-value and lowest MSE and MAPE values. The solar radiation data for Ibadan, Nigeria was used as case study. The network trained with the LM algorithm has the fastest speed of convergence during the training process, highest r-value and lowest values of MSE and MAPE for the training dataset.

Thus, indicating the high accuracy of the LM algorithms in modeling the solar radiation data. Based on the testing dataset, the BR algorithm-trained network has the highest r-value and the corresponding lowest values of MSE and MAPE. These showed the superiority of the BR algorithm in predicting or forecasting of the solar radiation data compared to other algorithms investigated in this study. Hence, in terms of speed and accuracy, the LM and BR learning algorithms are the two best algorithms to be considered for use in modeling and forecasting of solar radiation data.

## REFERENCES

Al-Alami, S.M. and H.A. Al-Hinai (1998) An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation. Renewable Energy, 14(1-4):199-204.

Azadeh, A. and B. Behshtipour (2008) The effect of neural network parameters on the performance of neural network forecasting. IEEE International Conference on Industrial Informatics (INDIN), 1498-1505.

Demuth, H. and M. Beale (2000) Neural network toolbox manual. Mathworks Inc. USA. 850pp.

Elizondo, D., G. Hoogenboom and R. Mc Clendon (1996) Development of neural network to predict daily solar radiation. Agric. Forest Meteorol. 71:115-32.

Fadare, D.A. and T.T. Olugasa (2009) An artificial neural network model for forecasting daily global solar radiation in Ibadan, Nigeria. Global Journal of Engineering and Technology, 2(2):123–131.

Fadare, D.A. (2009) Modelling of solar energy potential in Nigeria using an artificial neural network model. Applied Energy, 86(9):1410-1422.

Fadare, D.A. (2010) The application of artificial neural networks to mapping of wind speed profile for energy application in Nigeria. Applied Energy, 87(3):934-942.

Foresee, F.D. and M.T. Hagan (1997) Gauss-Newton approximation to Bayesian learning. Proceedings of the 1997 International Joint Conference on Neural Networks.

Gardner, M.W. and S.R. Dorling (1998) Artificial neural networks (the multilayer perceptron) - a review of applications in the atmospheric sciences. Atmospheric Environment 32 (14/15):2627-2636.

Gill, P.E., W. Murray, and M.H. Wright (1981) Practical Optimization, Academic Press, London.

MacKay, D.J.C. (1992a) Bayesian interpolation. Neural Computation, 4(3): 415-447.

MacKay, D.J.C. (1992b) A practical Bayesian framework for backpropagation networks. Neural Computation, 4(3): 448-472.

Mellit, A., M. Benghanen and S.A. Kalogirou (2006) An adaptive wavelet network model for forecasting daily total solar radiation. Applied Energy, 83:705-722.

Oyawale, F.A. (2006) Statistical methods- an introduction. International Publishers Ltd, Nigeria, 213pp.