Recurrent Neural Network Model for Forecasting Electricity Demand in Nigeria

K. A. Abdulsalam, O. Adegbenro, and T. O. Akinbulire
Department of Electrical and Electronics Engineering, University of Lagos, Nigeria
Email: adijakubura@yahoo.com

Abstract
This work uses modular recurrent neural network to estimate the electricity demand in Nigeria from 2015 to 2050. The network is a 2-layer multi-input, single-output model with twelve neurons trained using Levenberg-Marquardt algorithm. The data structure used for training is cell array of sequential concurrent data. The Recurrent Neural Network model was simulated as Non-linear Auto Regressive with eXogenous (NARX) model in Matlab environment and the predicted load for 2015 is about 550GWh with an expected demand increase of 7.5 % every five year.

Keywords: *energy, Levenberg-Marquardt, modular network, neurons, power*

1.0 INTRODUCTION

Electricity is a basic infrastructure for socio-economic development of modern societies and there has been an increasing quest for accurate tools for forecasting the ever-growing demand for electricity. This is so because Electrical Load Forecasting (ELF), which is the accurate prediction of both the magnitude and geographical distributions of electrical energy demand over the different periods of planning horizon, has become vital in planning and operation of electric utilities. ELF ensures a secured and efficient system of production of electricity especially in a competitive and deregulated electricity market (Oamek and English, 1984; Hippert *et al.,* 2001; Alfares and Mohammad, 2002; Mohammad *et al.,* 2002; Rafal, 2006; Fernando *et al.,* 2013; Ren *et al.*, 2016; Sami 2016). Studies have indicated that an increase of 1% in forecasting error could imply a ten million Pounds (or equivalent) increase in operating costs per year (Hippert *et al.,* 2001; Alfares and Mohammad, 2002). Thus, load forecasting is a key area of research in electrical energy with the aim of reducing such errors.

Presently, Nigeria is facing a major challenge of inadequate electricity supply. However, in a bid to solve the problem, the country has deregulated its electricity market in line with global best practise. Therefore, a great deal of statistical analysis, modelling and forecasting is required for planning operational activities ranging from generation to pricing.

There have been some earlier studies on forecasting of electrical energy needs for Nigeria, and of note is the Model for the Analysis of Energy Demand (MAED) (Sambo, 2008) in which an International Atomic Energy Agency (IAEA) energy modelling tool was used to forecast that Nigeria would need 297,900 MW of electricity in 2030. The study considered demography, socio economy indices and technology as the drivers of energy demand. It also used details of energy intensities and energy efficiency opportunities in the forecast. However, weather, an important factor in electrical energy consumption, was not considered. Also, in the disaggregated bottom up approach built on policy variables, (Ibitoye and Adenikinju, 2007) estimated demand for electricity in Nigeria based on sectoral projections in the four major end-use sectors consisting of residential, industrial, commercial and agriculture. Estimates of electricity consumption of about 35 GW in 2015 and 164 GW in 2030 were obtained. However, historical consumption data and weather variables, as determinants of electricity consumption pattern, were not considered. There is also no explicit method of how the suppressed demand is captured into the model.

Other works of consideration include the use of Feed Forward Back Propagation and Recurrent Neural Network (RNN) to predict short term load demand for the Nigeria power system (Alawode and Oyedeji, 2013). Here, the results showed that RNN gave better load forecast but the explicit forecast values were not stated. Likewise, Afolabi *et al.* (2008) used artificial neural network (ANN) to forecast electricity demand for a town (Ogbomosho) while Muhammad and Sanusi, (2012) deploy ANN for Short Term Load Forecasting (STLF) for a state (Kano State). The results of these earlier works are very divergent thereby suggesting a gap and the need for a further study into the subject taking into cognisance factors hitherto not considered in the previous studies.

Hence, the objective of this paper is to deploy modular RNN model to estimate long term electricity demand in Nigeria from 2015-2050 using factor analysis data and historical load data. The study considers the challenge of poor characterisation of electricity consumption into residential, industrial and commercial as identified by Bhattacharyya and Timilsina, (2009), and the impact of weather (temperature) as a determinant of electricity demand. Data for this study was obtained from the Power Holding Company of Nigeria, National Control Centre, Central Bank of Nigeria, World Bank database and Global Temperature database.

The Conceptual Framework

An ANN is a massively parallel-distributed processor comprising simple processing units that has a natural propensity for storing experiential knowledge and making it available for use (Haykin, 1999; Jain and Mao, 1996; Fausett, 1994). Its main features include ability to acquire knowledge through training, the use of the acquired knowledge to make decision (generalisation) and its capacity to modify its topology in response to stimuli. The use of ANN and its variants for forecasting electricity has been widely reported in literature (Hippert *et al.*, 2001; Hippert *et al.*, 2005; Tarik *et al.*, 2007; Khosravi *et al.*, 2011; Dao, 2015). The general conclusion is that RNN-based systems give higher precision and more results.

A RNN is a type of ANN with one or more feedback loops from the output or from other layers in the network. RNNs are potentially more powerful for forecasting than the Feed Forward Neural Network (FFNN) because of the ability of RNN to recognise and recall spatio-temporal patterns. The output of an RNN is a function of time while the output of a FFNN is constant (Mandic and Chambers, 2001).

A RNN is suitable for forecasting electricity demand in a developing economy because the electricity demand data are noisy and incomplete due to shortage in supply. Also, the future may not necessarily follow the past because developing economies are undergoing economic transition and increasing diffusion of technologies in the society. In addition, electricity consumption is largely influenced by human needs and societal objectives (Kumar and Prasad, 2016) which are not well defined ahead and can change with human and global development. **Figure 1** is a conceptual form of the forecasting model, which has RNN as the main forecasting technique.

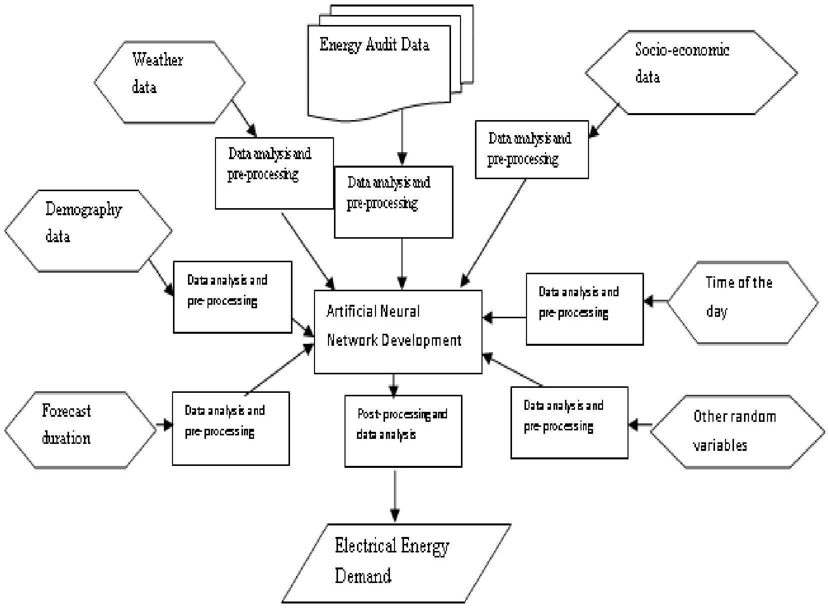


Figure 1: The Conceptual Framework of a RNN-Based Load Forecasting Model

2.0 MATERIALS AND METHOD

The following procedures were followed in developing the RNN model: normalisation of data, initialisation of weights, determination of the network structure, and training of the network.

2.1 Normalisation of Data

The data consist of the energy audit variable represented by the electric power consumption (kWh), socio-economic variables such as Gross Domestic Product (GDP) measured in local currency, GDP growth rate, inflation rate as a measure of the rate of change of the purchasing power of the income), demographic factors (total population, population annual growth rate) and measure of the impact of climate or seasonality as represented by the daily temperature readings. The data are divided into three parts in the ratio 14:3:3 for training, validating and testing respectively. Each data set is presented to the network as a matrix row of sequential concurrent data and the output as sequential data such that training is carried out on a set at an instance.

As shown in some earlier works (Haykin, 1999; Konstantious, 2002; Vladimir and Fillip, 2007; Beale *et al*., 2010), certain transformation or operations need to be carried out on the data in order to calibrate the model and parameterised it. The data preparatory steps employed include the following: outlier removal, quantity and quality checks and normalization. Normalization was performed on the data to convert input variables with different natural scales (different units of measurement) to a common scale such that the rescaled input span similar ranges of values. The data for the study are transformed to the range 0 to 1 using the steps of Eqs. (1) and (2):

For any series X_{it} such that each element $x_{it} \geq 0 \ \forall x_{it}$, let x_{it}^{min} denotes the minimum value in X_{it}, and x_{it}^{max} the maximum value in X_{it}. Therefore, the normalised value or resultant index δ, can be calculated as Eq. 1.

$$\delta = \left(x_{it} - x_{it}^{\min}\right)\left(x_{it}^{max} - x_{it}^{min}\right)^{-1} \quad (1)$$

For any series X_{it} such that $-1 \leq x_{it} < 0$,
let $\left[x_{it} + \left|x_{it}^{min}\right|\right] = \varphi_{it}$, then δ can be expressed as in Eq. 2.

$$\delta = \varphi_{it}\left(\varphi_{it}^{max} - \varphi_{it}^{min}\right)^{-1} \tag{2}$$

where φ_{it}^{max} and φ_{it}^{min} are defined as the maximum and minimum values of the data respectively. The normalised data are then tested for non-linearity using curve fitting tools in MATLAB in order to establish the appropriateness of ANN for forecasting the electrical load.

2.2 Initialisation of Weights

All the weights were initialised using the ranged randomisation of -1 to +1 in the first instance and then the Nguyen-Widrow technique was applied to improve the weights between the input layer and the hidden layer (Nguyen and Widrow, 1990; Hearton, 2011). The Nguyen-Widrow technique uses Eq. 3 to calculate beta *β*- a limiting function, in order to establish the range of the problem

$$\beta = 0.7h^{\frac{1}{i}} \tag{3}$$

where h is the number of hidden neurons in the first layer and i is the number of input neurons. Then each hidden neuron is assigned to the range of the problem. The essence of calculating *β* is to assign each hidden neuron to the range of the problems. The Euclidean norm *m* of all inputs to the current hidden neuron is obtained as in Eq. 4.

$$m = \sum_{i=0}^{i<w_{max}} {w_i}^2 \tag{4}$$

And the weights *(w)* are subsequently adjusted using Eq. 5.

$$w_{t+1} = \frac{\beta w_t}{m} \tag{5}$$

2.3 Determination of Network Structure

Determining the size of Neural Network entails finding the number of hidden units in the network, optimum weight vector and weight initialisation for the neurons and optimum network configuration required to approximate a function that best describes the available data to a high accuracy (Baun and Haussler, 1989; Murata *et al.,* 1994; Lappas, 2007).

The concept of capacity, measured as the maximum number of dichotomies that can be induced on *m* inputs, has been shown to hold for the problem of valid generalisation for arbitrary learning problem. It is related to the number of training sample and the formula for the upper bound size *S,* which is the computational unit defined as in Eq. 6.

$$S = 8 * \sqrt{\frac{2^n}{n}} \tag{6}$$

where *n* is the number of bits required to enumerate all existing training data and is defined by expression in Eq. 7.

$$n = \log[S_L] \tag{7}$$

where, S_L is the number of the existing training data, which is 240 items of data spanning a period of 40 years from 1970 to 2009 comprising historical data of generation, percent population growth, temperature, interest rate, inflation and GDP.

Modular Network for Forecasting

Electrical energy demand is the aggregate of electrical energy required by the different sectors of the economy comprising residential, commercial, and industrial. Electrical energy demand by various sectors can be determined and then aggregated to derive the total demand using modular systems.

Therefore, the input decomposition modular ANN architecture, which is a parallel combination of modules that aggregate by addition to give the total energy consumed is illustrated in **Figure 2** and expressed as $\sum_{k=1}^{N} M_k(z)$

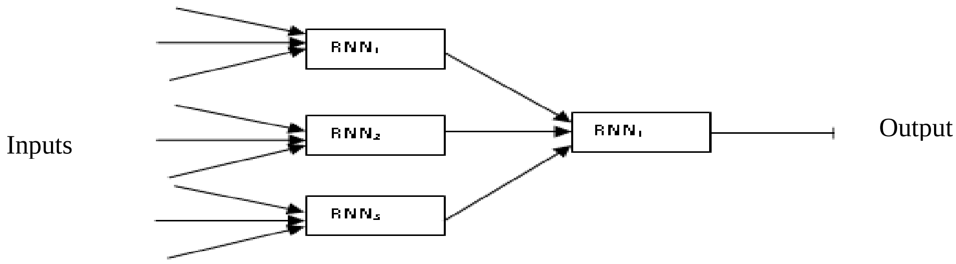


Figure 2: Modular ANN for Electrical Energy Demand

Where RNN_1, RNN_2, RNN_3, and RNN_T respectively represent electricity consumption in residential, commercial, industrial sectors and total aggregate electricity in all sectors (Mandic and Chambers, 2001) and M_k is the product of weight and activation function in each layer.

2.4 Network Training

Training is the art of learning by the neural network. Levenberg-Marquardt (LM) algorithm is employed in training algorithm development because of its robustness since it is a hybrid technique. The update rule of LM algorithm is expressed in Eq. 8 (Hao and Bogdam, 2011).

$$w_{k+1} = w_k - (J_k^T J + \mu I)^{-1} J_k e_k \tag{8}$$

where *J* is the Jacobian matrix, μ is the learning step and *I* is an identity matrix.

In using the Levenberg-Marquardt algorithm to implement ANN training, two problems are solved: calculation of the Jacobian matrix and training algorithm design. The Jacobian matrix is calculated by using traditional back propagation computation in first order algorithms in which the back propagation process is repeated for every output separately in order to obtain consecutive rows of the Jacobian matrix.

2.5 MATLAB model Simulation and Training

A MATLAB training simulation was carried out on the proposed RNN model using the Non-linear Auto Regressive with eXogenous (NARX) model in MATLAB. The NARX predicts the future values of a series from the previous output values and previous input values. It is a recurrent dynamic network with feedback connections enclosing several layers of the network based on the linear ARX model commonly used for time series modelling. This is the most suitable model for predicting electricity demand value because the present demand is a function of previous demand, previous and present inputs. The multi input single output (MISO) network is trained using multi-dimensional input data and the system responds by producing a scalar output. The scalar output is then de-normalised to obtain the aggregate electricity demand.

The NARX model uses Levenberg-Marquardt algorithm for training and it is defined by Eq. 9.

$$y(t) = f\left(y(t-1), y(t-2), ..y\left(t-n_y\right), x(t-1), x(t-2), ... x(t-n_u)\right) \quad (9)$$

in MATLAB environment which is equivalent to the neural network architecture for prediction of the form (Eq. 10).

$$y(k) = F\big(y(k-1), \dots, y(k-N), x(k-1), \dots, x(k-M)\big) \quad (10)$$

3.0 RESULTS AND DISCUSSION

3.1 Optimal Network Size

The number of neurons *S* is approximately equal to twelve. S_L is the number of training data and it is equal to 240 as expressed in Eqs. 6 and 7.

3.2 Network Initialisation Parameters for Training Development

The weight and the initial values of the Neurons used in this work are presented in **Tables 1 and 2.**

Table 1: Weight Values of the Neurons

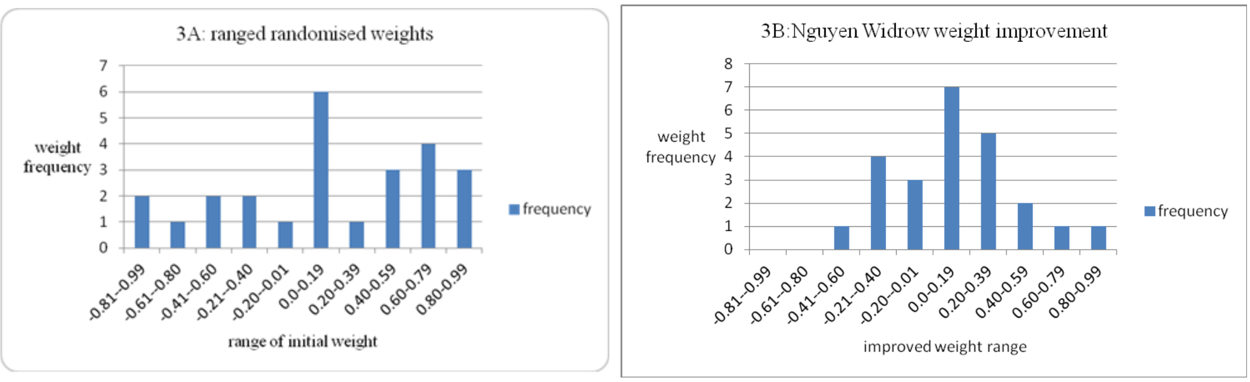
Initial weight (ranged randomisation)	Enhanced weight (Nguyen Widrow)
0.72	0.25
-0.38	-0.13
-0.73	-0.25
-0.95	-0.33
0.69	0.24
-0.31	-0.11
-0.81	-0.28
-0.57	-0.40
-0.61	-0.11
0.46	0.32
0.47	0.33
0.19	0.13
0.21	0.15
0.59	-0.14
-0.41	0.99
0.99	0
0	0
0	0
0	0

Table 2: Bias values

Initial bias	Enhanced bias
0.94	0.33
0.84	0.59
0.63	0.63

Other parameters:
β = 0.77; (hidden norm_1) m = 2.22 and (hidden norm_2) m = 1.10

Figure 3 shows the weight distribution of the RNN around zero. **Figure 3A** used random generator which was later improved with Nguyen-Widrow technique as shown in **Figure 3B**. The random generator produced even distribution of numbers around zero while the Nguyen Widrow modifies the weights such that there is a large distribution around zero.



A: using random number generator **B: when A is improved with Nguyen-Widrow technique**
Figure 3: Weight distribution around zero

3.3 MATLAB Model Simulation and Training

A modular dynamic network defined in **Table 3** is deployed and trained in MATLAB environment.

Table 3: Elements of the RNN Modular Models in MATLAB

Type of network	Multi input-single output NARX RNN with output feed back		
	Residential	Commercial	Industrial
Number of Layers	2	2	2
Number of Neurons in hidden layer	10	12	10
Number of inputs	5	6	6
Number of neurons in the output layer	1	1	1
Activation function	Sigmoid	Sigmoid	Sigmoid
Data division (random) for training, testing and validation	14:3:3	14:3:3	14:3:3
Training algorithm	Levenberg Marquardt	Levenberg Marquardt	Levenberg Marquardt
Data structure	Matrix row	Matrix row	Matrix row

Table 4 shows the results of the models. The Mean Square Error (MSE) values are approximately zero. This shows that the difference between the target and the output is insignificant. Also, the coefficient of correlation (R) values are approximately 1 to indicate a close correlation between the output value and target. The values of MSE and R show that the model is adequate. In **Figures 4 to 6,** plots A, B and C respectively show the validation performance of the training, the autocorrelation of error and the response of output elements for residential, commercial and industrial modules. The performance plots indicate that the validation and test data have the same characteristics and there is no increment in both, hence, there is no overfitting. Error Autocorrelation plots show that all errors fall within the confidence limit, except the one at zero lag indicating that the model is adequate. Time series response of plots C displays the inputs, targets and error vs. time to indicate which time points were selected for training, testing and validation. The forecasted value is presented in **Table 5**.

Table 4: Results of MATLAB Training

	Residential		Commercial		Industrial	
	MSE	R	MSE	R	MSE	R
Training	8.6147e-4	9.8913e-1	2.1651e-4	9.9635e-1	5.8489e-3	9.1512e-1
Validation	5.7537e-3	9.5395e-1	5.0831e-4	9.9522e-1	1.3266e-3	9.7942e-1
Testing	1.0179e-3	9.0934e-1	1.8767e-3	9.7881e-1	2.4757e-3	9.7448e-1

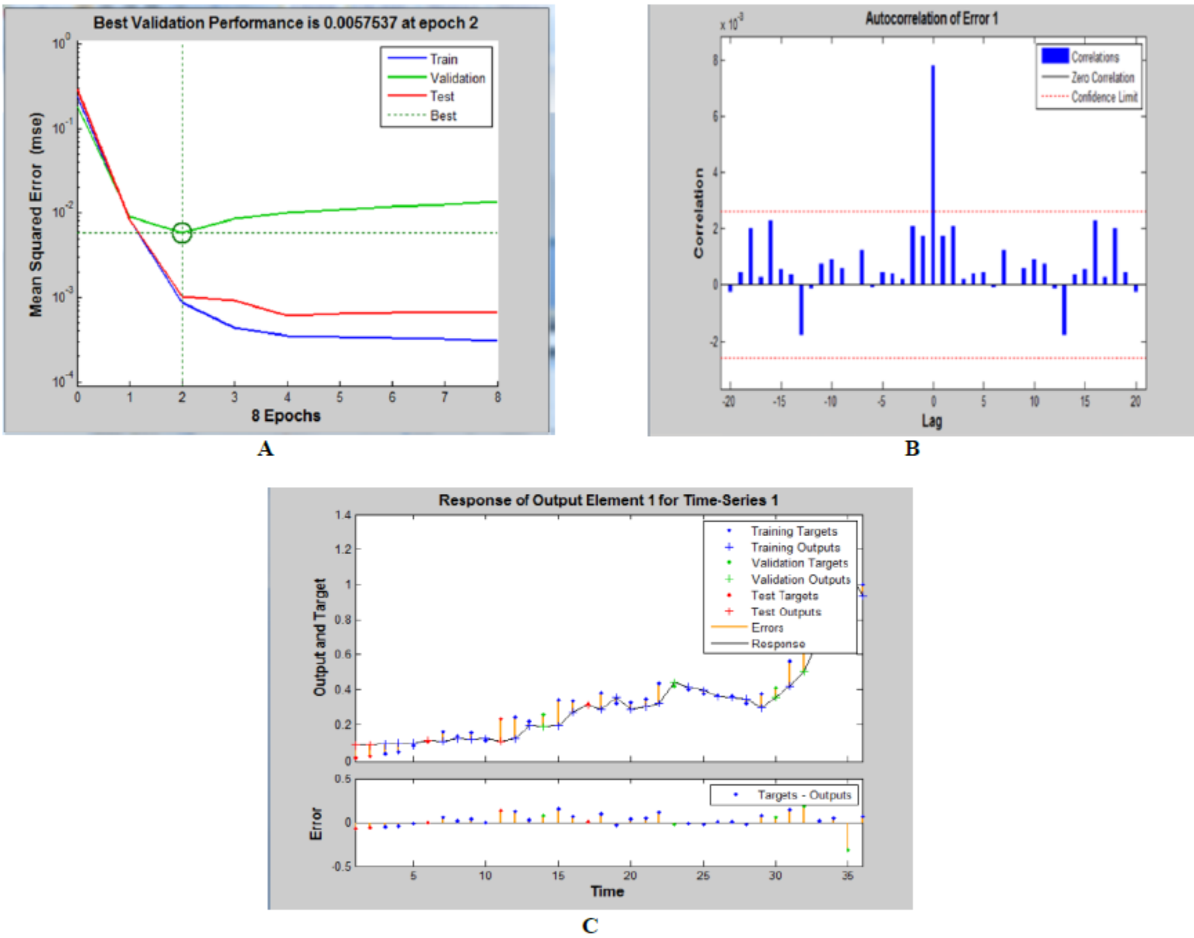


Figure 4: Plot Results for Residential Module

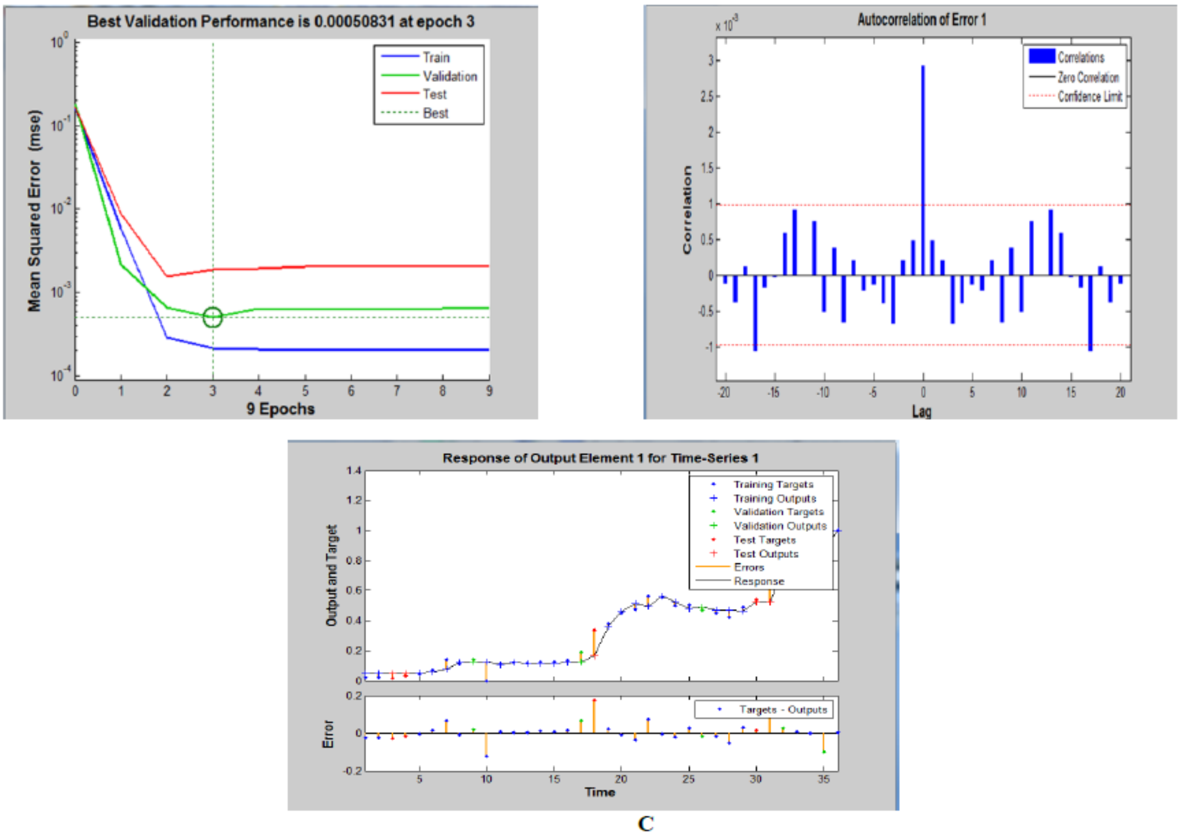
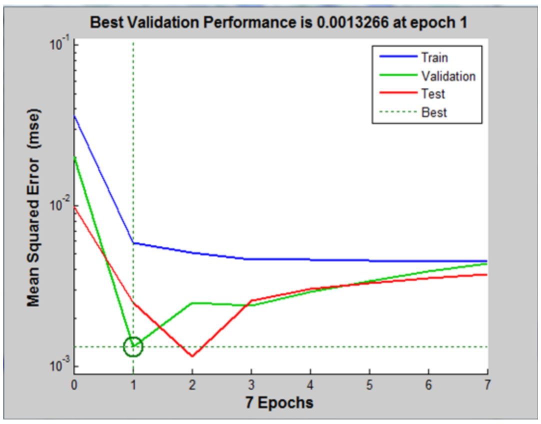
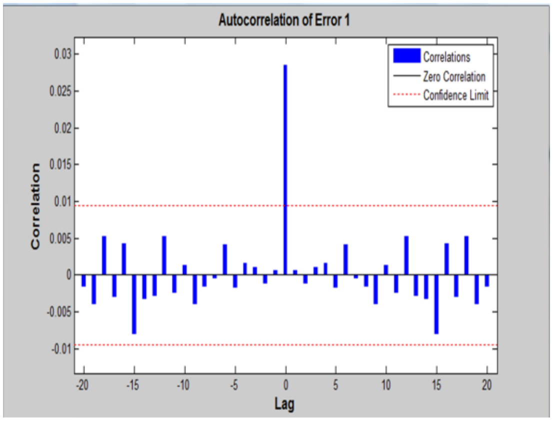


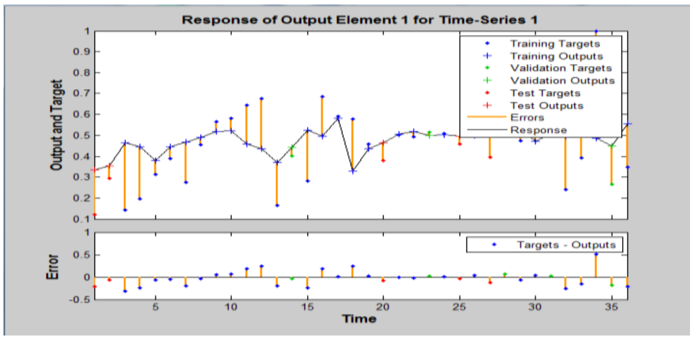
Figure 5: Plot Results for Commercial Module



A



B



C

Figure 6: Plot Results for Industrial Module

Table 5: The Predicted Results

Year	Forecast Electrical Energy Value (GWhr)	Equivalent Power (GW)
2015	548737.3	63
2020	597811.2	68
2025	651845.1	74
2030	711516.1	81
2035	771160.8	88
2040	826519.5	94
2045	878343.6	100
2050	927476.4	106

4.0 CONCLUSION

Improving the supply of electricity to meet targeted socio-economic objectives requires planning. However, a key element in planning is forecasting of electricity demand which include identifying when, where and what amount is needed. This paper has demonstrated the use and adequacy of modular recurrent neural networks for predicting electricity demand. The study used climatic, demographic and economic factors as significant inputs in developing neural network architecture to forecast electricity demand in Nigeria. The plot analysis of model simulation in MATLAB environment indicate model adequacy with a MSE of approximate zero. The predicted load is ~ 550,000 GWH in 2015, and an expected demand increase of approximate 7.5 % every five year.

REFERENCES

Afolabi, A. O., Olatunji, B.O. and. Ajayi, A. O (2008). Electricity Load Forecasting Using Artificial Neural Networks, *Journal of Engineering and Applied Sciences*, 3**(2)** Medwell Online Journal.

Agboola, O.P., (2011). Independent Power Producer (IPP) Participation Solution to Nigeria Power Generation problem, *Proceedings of the World Congress of Engineering*, III, London.

Alawode, K.O. and Oyedeji, M.O. (2013). Comparison of Neural Network models for Load Forecasting in Nigeria Power Systems, IJERT.

Alfares, H.K. and Mohammad, N. (2002). Electric Load Forecasting: Literature Survey and Classification Methods, *International Journal of System Science,* 33**(1):** 23-34.

Baum, E. B. and Haussler, D, (1989). What Size Net Gives Valid Generalisation, *Neural Computation*, *MIT Press,* 1**(1):** 151-160.

Beale, Mark H., Martin, T. H. and Howard B. D, (2010). Neural Network Toolbox™ 7 User's Guide', *Maths Works* Inc.

Bhattacharyya, S.C. and Timilsina, G.R. (2009). Energy Demand Models for Policy Formulation- A Comparative Study of Energy Demand Models, Policy Research Working Paper 4866, The World Bank Development Research Group Environment and Energy Team *http://econs.worldbank.org*

Dao, J. (2015). Study on Short Term Load Forecasting Method Based on the PSO and SVM Model, *International Journal of Control and Automation*, 8**(8):** 181-188.

Fernando, M., Juan, J. C., Monica Millan-Giraldo, Abderrahim Sellami, Pablo Escandell-Montero, Jose M. Martınez-Martınez and Emilio Soria-Olivas (2013). Machine Learning Techniques for Short-Term Electric Power Demand Prediction, *ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. Bruges (Belgium).

Haykin, S. (1999). Neural Networks: A Comprehensive Foundation, Prentice Hall Int'l, 2nd ed, 1-897.

Heaton, J. (2011). Introduction to Mathematics of Neural Networks, Heaton Research Inc. USA.

Hippert, S.H., Pedreira, C.E. and Souza, R.C. (2001). Neural Networks for Short Term Load Forecasting: A Review and Evaluation', *IEEE Transactions on Power Systems*, 16**(1):** 44-55.

Hippert, S.H., Bunn, D.W. and Souza, R.C. (2005). Large Neural Networks for Electricity Load Forecasting: Are they over fitted, *International Journal of Forecasting* 21, 425-434.

Ibitoye, F.I. and Adenikinju, A. (2007). Future Demand for Electricity in Nigeria, *Elsevier Applied Energy* (84), 492-504.

Khosravi, A., Nahavandi, S. and Creighton, D. (2011). Short Term Load Forecasting using Interval Type-2 Fuzzy Logic, *IEEE International Conference on Fuzzy Systems*, Taipei: 502-508.

Konstantinos, I D. (2002). Neural Networks and Principal Component Analysis, in Handbook of Neural Network Signal Processing, Ed. Yu Hen Hu, Electrical Engineering and Applied Signal Processing Series CRC Press USA, 211-248.

Kumar, R. C. and Prasad, T. N. (2016). Performance Comparison of Short Term Load Forecasting Techniques, *International Journal of Grid and Distributed Computing*, 9**(4):**.287-302

Lappas, G. (2007). Estimating the Size of Neural Network from the Number of Available Training Data, in Lectures Notes in Computer Science, Ed. Joaquim Marques de sa Luis A. Alexandre and Wlodzislaw DuchDanilo P. Mandic, ANN ICANN, *17TH*

International Conference Porto, Portugal, Proceeding Part 1. Springer Berlin Heidelberg, 68-77.

Mandic, D.P. and Chambers, J.A. (2001). Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures, and Stability, *John Wiley & Sons, England*.

Mohammad, S., Hatim Y. and Zuyi, Li (2002). Market Operations in Electric Power Systems: Forecasting, Scheduling and Risk Management, *John Wiley & Sons, Inc., Publication.*

Muhammad, B. and Sanusi, S. A. (2012). Short Term Load Forecasting using Artificial Neural Network, *Proceeding of the International Multi-conference of Engineers and Computer Scientists IMECS* 2012 vol 1, Hong kong.

Murata, N., Shuji Yoshizawa and Shun-ichi Amari (1994). Network Information Criterion-Determining the Number of Hidden Units for an Artificial Neural Network Model, *IEEE Transaction on Neural Networks*, 5**(6):** 865-872.

Nguyen, D. and Widrow, B. (1990). Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights, *Proceedings of the International Joint Conference on Neural Networks,* 3: 21-26.

Oamek, G.E. and English, B.C. (1984). A Review of Load Forecasting Methodologies, http://www.card.iastate.edu/publications/DBS/PDFfles/84w p5.pdf. Accessed 22/2/2010

Okoro, OI and Chikuni, E (2007). Power Sector Reforms in Nigeria: Opportunities and Challenges, *Journal of Energy in Southern Africa,* 18**(3):** 52-57.

Rafal, W. (2006). Modelling and Forecasting Electricity Loads and Prices, *John Wiley and sons UK.*

Ren, Y., Suganthan, P.N., Srikanth, N. and Gehan, A. (2016). Random Vector Functional Link Network for Short-Term Electricity Load Demand Forecasting, *Information Sciences,* 367–368**(1):** 1078-1093.

Tarik, R., Huang, B.Q., Kechadi, M-T and Gleeson, B. (2007). Autoregressive Recurrent Neural Network approach to electricity load forecast, WASET, No1.

Sambo, A.S. (2008). Matching Electric Supply with Demand in Nigeria, *Int'l Association for Energy Economics*, 4th Quarter: 32-38.

Sam,i E (2016). Electric Load Forecasting using Regularized Extreme Learning Machines, *International Journal of Industrial Electronics and Electrical Engineering*, 4**(6):** 119-122.

Vladimir, C. and Fillip, M. (2007). Learning from Data: Concepts, Theory and Methods', *John Wiley and sons New Jersey*, 2nd ed.