MODIFIED NARX NETWORK FOR LOW VOLTAGE CONSUMER LOAD PREDICTION

Ayodele Isqeel ABDULLATEEF[1], Momoh-jimoh Eyiomika SALAMI[2], Ismaila Bayo TIJANI[3], Mosud Taiwo AJALA[4]
[1]Department of Electrical and Electronics Engineering, University of Ilorin – Nigeria.
[2]Department of Electrical and Electronics Engineering, Elizade University – Nigeria.
[3]Departmen of Electronics/Electrical, Abu Dhabi Men's College – UAE.
[4]Department of Industrial Maintenance Engineering, Yabatech. Lagos - Nigeria
*e-mail: abd_lateef.aii@unilorin.edu.ng

Keywords: –
Consumer load prediction, Data acquisition, Differential genetic, Algorithm evolution, Nonlinear autoregressive with eXogenous input.

Article History: –
Received: January, 2019.
Reviewed: April, 2019
Accepted: August, 2019
Published: September, 2019

ABSTRACT

This paper is a continuation of our previous work on Nonlinear autoregressive with eXogenous input (NARX) for load prediction. Application of NARX network in real-time might be difficult since the tapped delay was selected by trial and error, leading to nine hidden neurons which makes the network complex. This NARX network was modified based on Genetic Algorithm (GA) and Differential Evolution (DE) resulting in a new model coded NARX-DE-GA. GA and DE search for the number of hidden neurons and tapped delay automatically. The NARX-DE-GA was used to predict the consumer load using one month energy data with 8928 data points. The results show that NARX- DE -GA outperformed the NARX network. The training mean square error (MSE) value for NARX- DE -GA is 0.0253 while validation is 0.0612. These values are slightly higher when compared with NARX network in previous study which are 0.0225 and 0.0533 respectively. However, the network structure which is one input and output tapped delay, and one hidden neuron is simple and applicable in real time.

1. INTRODUCTION

Load prediction is essential to the management of power system in that it facilitates the estimation of future electricity demand based on past records. Besides, it assists the utilities in making decisions on operational planning such as routine maintenance, dispatch of generators, fuel management, and infrastructure development [1]. Inaccurate load prediction increases the operational cost of the utility company, most notably under a deregulated market environment [2]. Load prediction could be long term, medium term, short term load and very short term., Very short-term load forecasting (VSTLF) has been proposed in areas that need a forecast for a short time leads such as real-time consumer load prediction which can be used for monitoring individual consumer. This has led to its application for the prediction of load consumption from a few minutes to several hours [3]. Historical load data, weather data, temperature and seasonal data are the common input variables in load forecasting. However, instead of modeling the relationships between these variables, as in other types of load prediction, VSTLF extrapolates the recently observed load pattern to the nearest future [4] as it relies on real-time data to predict future energy demand.

Usually, data used in the prediction are cumulative consumer data which fail to reveal individual energy consumption pattern needed to monitor individual consumer connected to the power network [5]. Individual consumer data is crucial for load prediction as well as monitoring consumer's activities (such as electricity theft) on the distribution network. Several techniques on short-time load prediction have been reported. They include statistical method such as multiple linear regression, exponential smoothing, time series, state-space, and Kalman filter techniques. A very short-term load prediction between 10 and 30 minutes ahead of using minute-by-minute British electricity demand load data has been proposed [6]. A seasonal autoregressive moving average (SARMA) for home peak load was proposed by [7]. The study noted that the ability to predict the stochastic activities of the consumer and routines are more significant for home load prediction. Koprinska et al.[1] proposed a weekday based prediction model for electricity load forecasting using autocorrelation feature selection and machine learning algorithm. A strategy for developing a very short-term load prediction suggested by Trudnowski, et al. [8] uses slow and fast Kalman estimators.

Furthermore, artificial intelligent paradigm such as an artificial neural network (ANN), fuzzy logic (FL), wavelet neural network, and knowledge-based expert systems for VSTLF have been proposed [9]–[12]. Yang et al. [13] proposed a fuzzy neural system (FNS) for very-short-term electric load prediction based on chaotic dynamics reconstruction technique. A hybrid model of the similar day and neural network for load forecast proposed by Fok and Vai [14] considered hourly weather information which, is not often considered in other VSTLF literature as one of the input variables.

Furthermore, nonlinear autoregressive network with eXogenous inputs (NARX) has been used in system identification and time series prediction [15]–[17]. Aside from the fact that NARX neural network can handle nonlinear data, its learning ability with gradient descent algorithms has been shown to be effective and superior to any other recurrent network [18]. NARX network generalizes better and converges faster than other networks [19]. Consumer load prediction based on NARX network was proposed in our previous paper [20]. However, it was observed that the real-time application of the proposed network for consumer load prediction might be difficult. Two tap delay and nine hidden neurons were selected by trial and error. This made the network to be complex. Hence, there is a need to modify the NARX optimally for improved performance. In this paper, NARX network was modified such that the number of tap delay and hidden neurons were selected optimally using DE and GA respectively. This results in a simple and accurate network that can be used in real time to predict individual consumer's load consumption. This paper is divided into five sections. Section II presents the methodology which involves data acquisition and modified NARX network. Section III presents the results and discussion. Section IV describes the comparison between NARX and NARX-DE-GA networks results. Section V provides the conclusi0on.

2. METHODOLOGY

2.1. Data acquisition and processing

The detail explanation of the data acquisition procedure can be seen in [20]. Fig. 1 illustrated the block diagram of the data experimental set up which includes consumer load prototype where different household equipment were connected. Current sensor A (ACS785, Allegro MicroSystem Inc., USA), is connected to the live line to acquire current consumed by the loads. BNC-2110 and PCI 6420E card are programmable LABVIEW hardwares used for data conversion and logging to the computer.

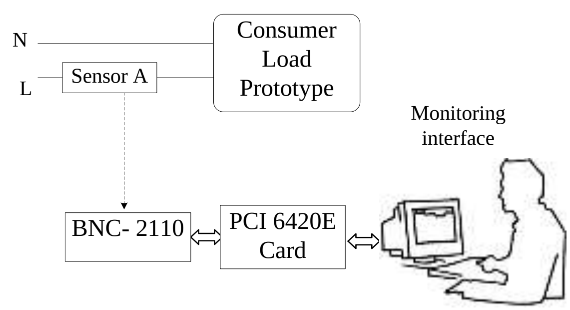


Fig. 1. Block Diagram of Experimental Set Up

Fig. 2 illustrate the process of the consumer energy conversion.

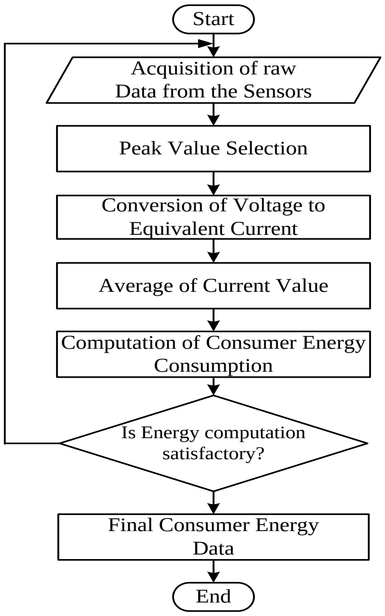


Fig. 2. Consumer energy conversion.

2.2. NARX Network Architecture

Fig. 3 shows a typical NARX network. It is formed by the input tapped delay which slides over the input signal, and the output tapped delay which slides over the output signal. It is different from another recurrent neural network in that its feedback network comes from output neuron only and not from hidden neurons.

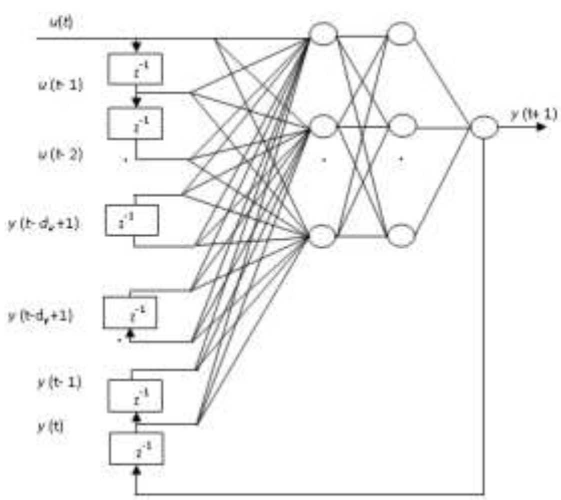


Fig. 3. NARX Architecture with Input *du* and Output *dy* Tapped Delays

Moreover, it could be designed as a feed-forward time-delay neural network (Without the feedback loop of delayed outputs) when applied to time series prediction.

NARX model can be expressed mathematically as

$$y(t+1) = f[y(t), y(t-1), \ldots y(t-d+1); u(t-1), \ldots u(t-d+1)] \quad (1)$$

Concisely it can be written as

$$y(t+1) = f[y(t); u(t)] \quad (2)$$

where $u(t) \in \Re$ and $y(t) \in \Re$ represent the input and output of the model at time *t*, du ≥ 1 and dy ≥ 1, du ≤ dy, are the input memory and output memory orders respectively, which is analogous to the order of an autoregressive model. The function *f* is a nonlinear mapping function which is approximated by using multi-layer perceptron neural network (MLP). The computational capabilities of NARX recurrent network as compared to other recurrent neural networks have been evaluated [21] and likened to a tuning machine.

2.3. Differential Evolution

Differential evolution (DE), proposed by Storn and Price [22] is undoubtedly one of the most potent stochastic real-parameter optimization algorithms. Its operation is through similar computational steps as employed by population-based standard stochastic global optimizers which include mutation, recombination, and selection. However, unlike other evolutionary algorithms, a mutation in DE involves the addition of a weighted difference between two population vectors (target vector) to a third vector and consequently produces the donor vector. While its recombination between the target vectors and the donor vector produces the trial vectors (offspring). Furthermore, the selection process is basically a one-to-one greedy selection between the trial vector and the target vector. The conceptual and algorithmic simplicity, ability to handle non-differentiable, nonlinear and multimodal objective functions, excellent convergence characteristics and robustness of DE has made it an efficient and popular technique for real-valued parameter optimization [23]–[25]. Thus, DE has been applied in several studies to solve optimization problems [26]–[29]. The essential operation of DE algorithms is shown schematically in Fig.4.

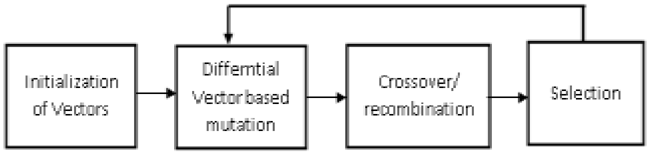


Fig. 4. Basic Operation of DE Algorithms

2.4. Genetic Algorithms

Genetic algorithms (GA) are search algorithms that are based on the concept of natural selection that is the survival of the fittest and natural genetics. GA, as an optimization technique, introduced by Holland [30], solves the optimization problem by either maximizing or minimizing the objective function using genetic operators such as selection, reproduction, crossover, and mutation. Unlike other search methods that search in one direction and sometimes end up in local minimal, GA searches the entire solutions space and works with the coding of the parameter set, rather than the parameter values itself. It also uses objective function information without any gradient information [31]. The basic principles of a GA are the formulation of the initial population, evaluation of the objective function, finding the fitness function, and applying genetic operation such as reproduction, crossover, and mutations as illustrated in Fig. 5.

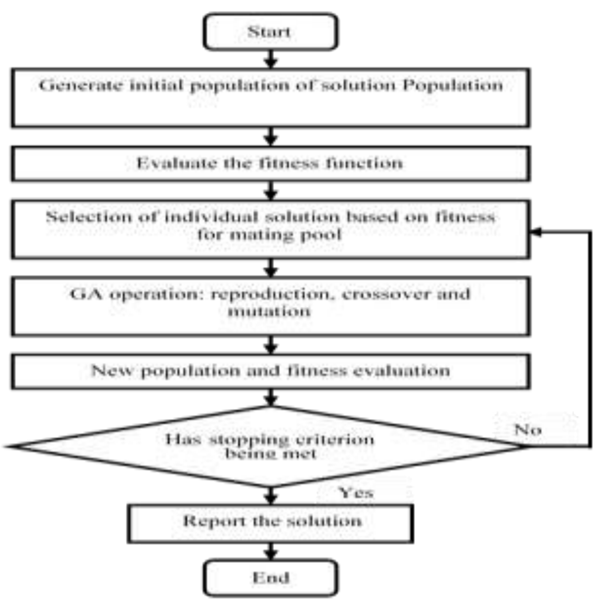


Fig. 5. Flow Chart of Genetic Algorithm

2.5. NARX- DE -GA

The modified NARX coded as NARX-DE-GA was used to predict the consumer load. Genetic algorithm (GA) and differential evolution (DE) search for the number of the hidden neurons and tapped delay automatically. Thereby optimizes the network for better performance. The detailed algorithm formulation had been presented in [32].

The following procedures are executed by the algorithm. For a set of input and output data pairs of a system to be modeled, the objective of the DE in stage one of the optimization is to search for a set of network architecture parameters defined as β, which minimizes the objective function. β is defined as

$$\beta := [TD, G] \tag{3}$$

where

TD and G are the tapped delays and hidden neuron respectively.

$$_{\min} f(\beta) = 0.5 * (\delta(\mathrm{W}) + \lambda_c) \tag{4}$$

where λ_c is the percentage of network connection used which indicate the complexity of the network and is expressed as:

$$\lambda_c = \frac{\text{total network used}}{\text{total possible weight}} \times 100 \tag{5}$$

λ_c is also the value of objective function returned by the GA sub-algorithm. GA algorithm searches for a set of network weights, W, that minimizes the objective function, which is defined as prediction error, i.e., the difference between the actual value and predicted value for a given data, length, *N*, that is

$$_{\min} \delta(\mathrm{W}) = \frac{1}{N} \sum_{i=1}^{N} e^2 \tag{6}$$

where $e = y(n) - y(n+1)$

The overall flow chart of the NARX-DE- GA algorithm is shown in Fig. 6.

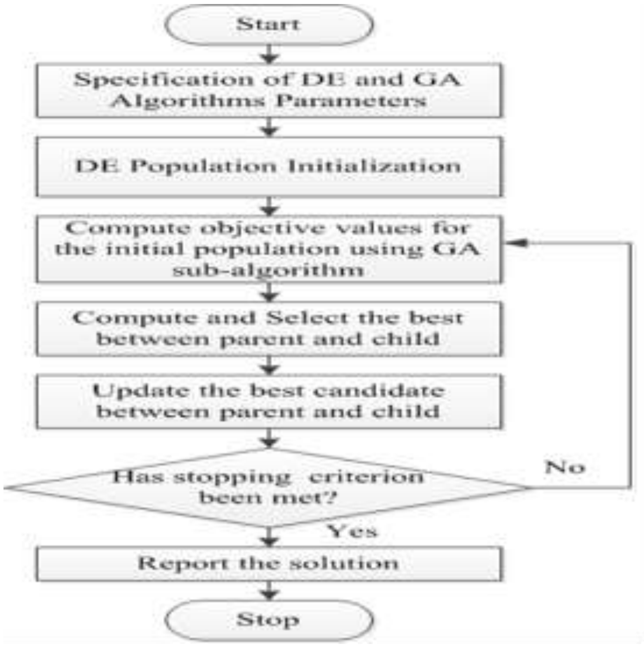


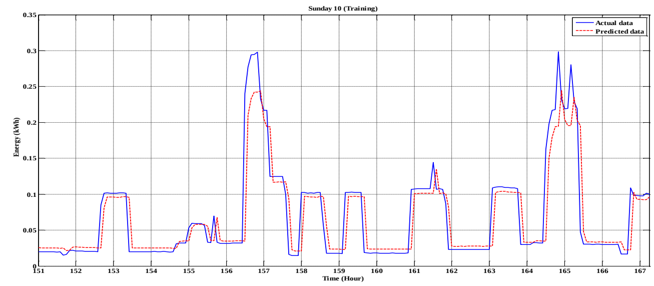
Fig. 6. Flow Chart of the Developed NARX-GA-DE Algorithm

This hybrid algorithm, subsequently refers to as NARX-GA-DE network algorithm, is written in MATLAB (2009a, The MathWorks). This facilitates the integration of the MATLAB inbuilt GA and NARX network algorithm in the hybrid optimization algorithm process

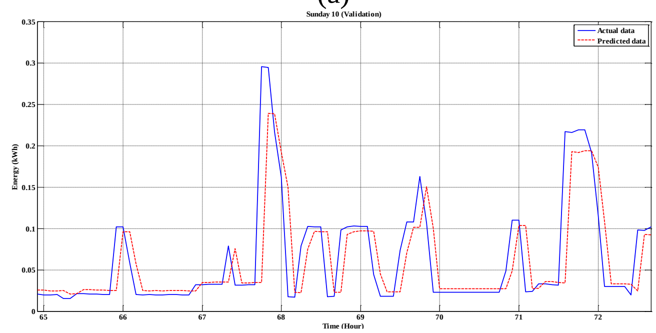
3. Results and Discussion

The algorithm was applied to the same sets of data that was used in the convention NARX [20] for the prediction of consumer load. The dimension D, of the problem, is three that is, the input tapped delays, the output tapped delays and the hidden neuron. The optimization parameters used for both GA and DE are, population size (30), maximum generation size (20); maximum network size, which includes the tapped delay TD = 10, and hidden neuron, G = 10. The DE mutation factor, (0.75) and cross over constant, (0.25), GA mutation rate, 0.01, probability crossover, 0.8.

The result shows that one input and output tapped delay, one hidden neuron, as well as one output neuron, is the best architecture for the optimal network structure for predicting the consumer load while the total weight is 5. Fig. 7a to Fig. 13a show the prediction obtained for one week using the NARX-GA-DE while the Fig. 7b to Fig. 13b indicate the prediction when validation data was used.

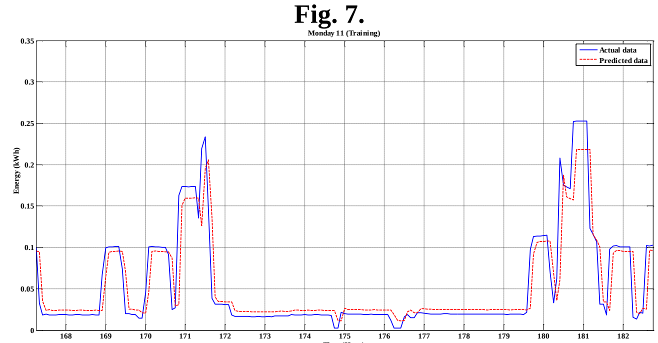


(a)

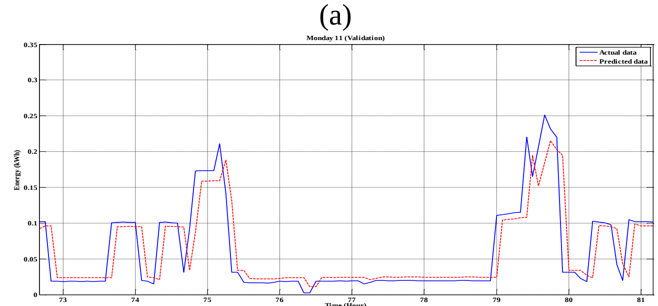


(b)

Fig. 7.

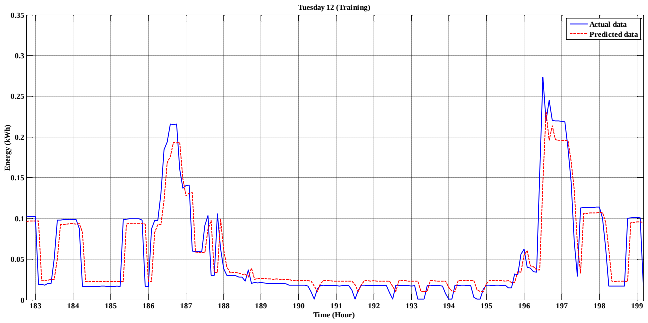


(a)

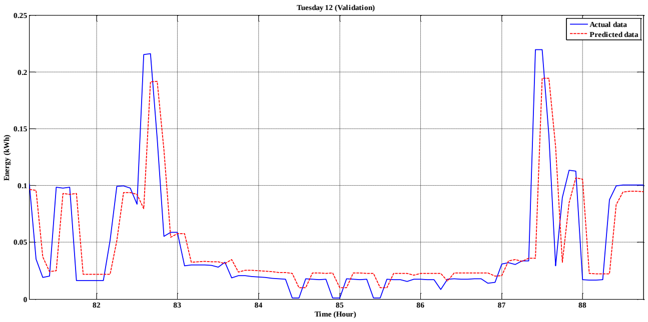


(b)

Fig. 8.

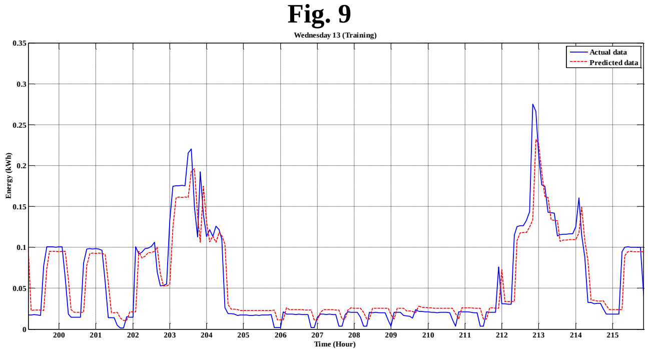


(a)

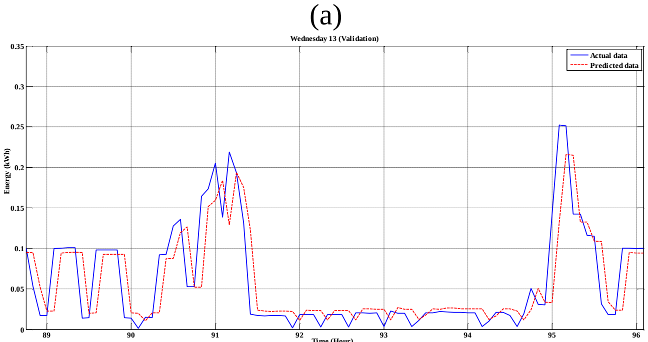


(b)

Fig. 9

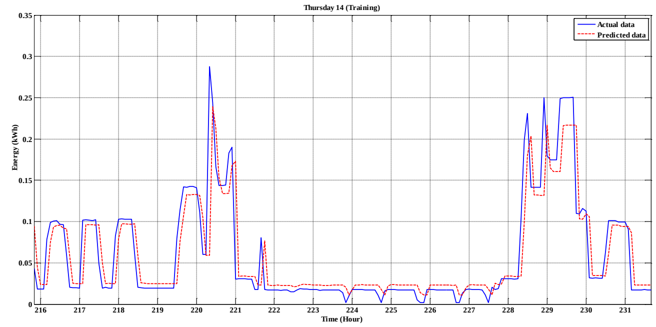


(a)

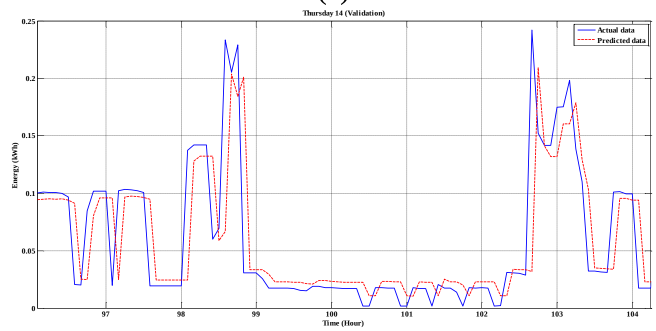


(b)

Fig. 10.

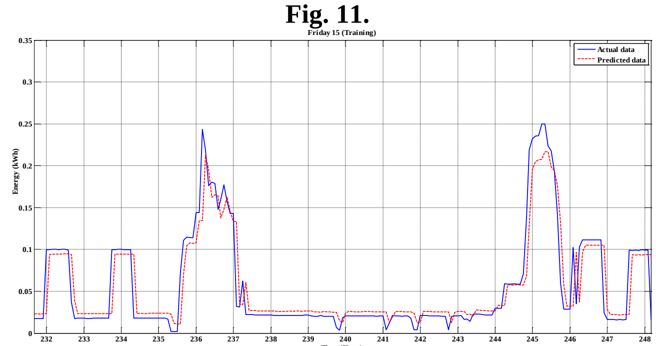


(a)

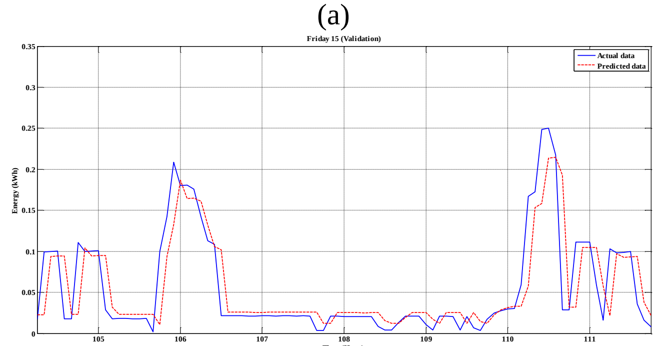


(b)

Fig. 11.

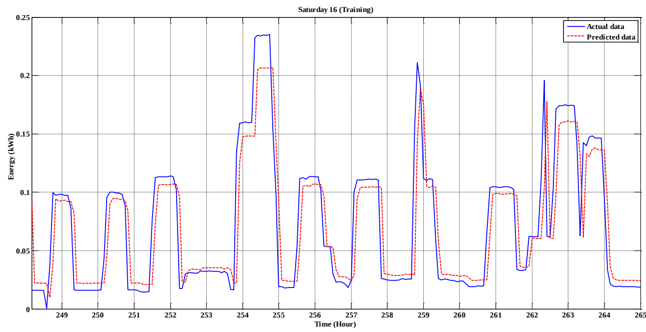


(a)

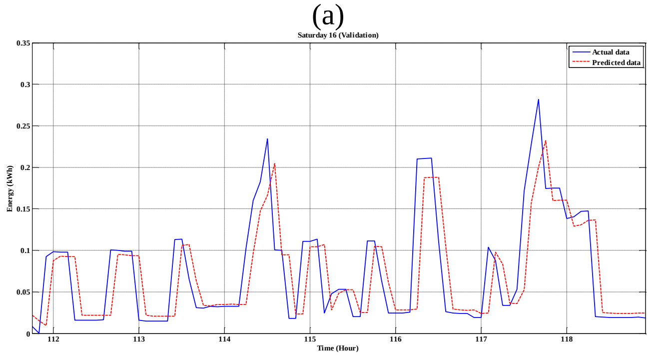


(b)

Fig. 12



(a)



(b)

Fig. 13.

Table 5.1 illustrates the breakdown of the results. The prediction error is 0.0253 while the validation error stands at 0.0612 meaning the average error is 0.0432 and the network size is 5.

Table 1. NARX-GA-DE Prediction result

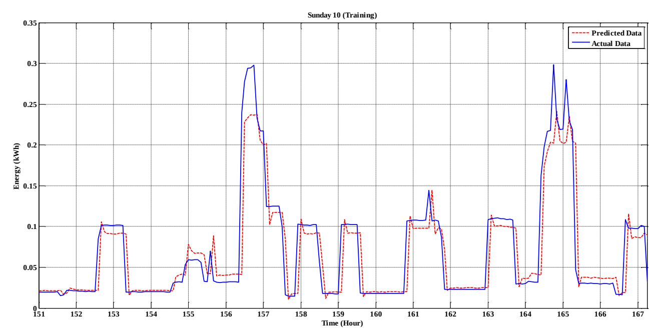
Network size	Training MSE	Testing MSE	Network weight
[1 1 1]	0.0252	0.0612	5

4. COMPARISON BETWEEN NARX AND NARX-GA-DE NETWORKS FOR CONSUMER LOAD PREDICTIONS.

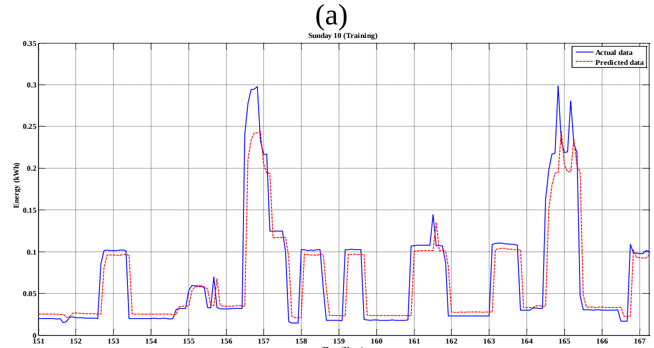
The NARX-GA-DE algorithm outperformed the conventional NARX network when used to predict the consumer load. The former MSE is slightly higher than the latter; the training MSE value for NARX-GA-DE is 0.0252 while its validation MSE is 0.0612. The MSE for the conventional NARX is 0.0225 for training and 0.0533 for validation, respectively. Besides, the network structure of the developed algorithm is simple and performed better with one input and output tapped delay and one hidden neuron, as shown in Fig. 15, and Fig. 16 for a typical weekend and weekday load prediction. The simplicity of its structure is a huge advantage over the convectional NARX network where the hidden neuron is nine, and the input and output tapped delays is two. Moreover, the network selection of the modified NARX is automatic, while conventional NARX network selection involves trial and error technique.

For instance, it is observed that the predicted data could not follow the actual data perfectly, as shown throughout the results. For clarity, two of such graphs are repeated in Fig. 15, and Fig. 16 where noise is observed most specifically between 152 hours and 156 hours, from Fig. 15a, and between 200 and 203 in Figure 16a. However, within the same period, the NARAX-GA-DE predicted data was able to follow the actual data without any noise, as illustrated in Fig. 15(b) and Fig. 16(b).

In general, the NARX-GA-DE algorithm performance was better and free from noise. Moreover, considering that the system will be applied in real-time to predict the consumer load, a simple network structure is much more preferred to a complex network structure.

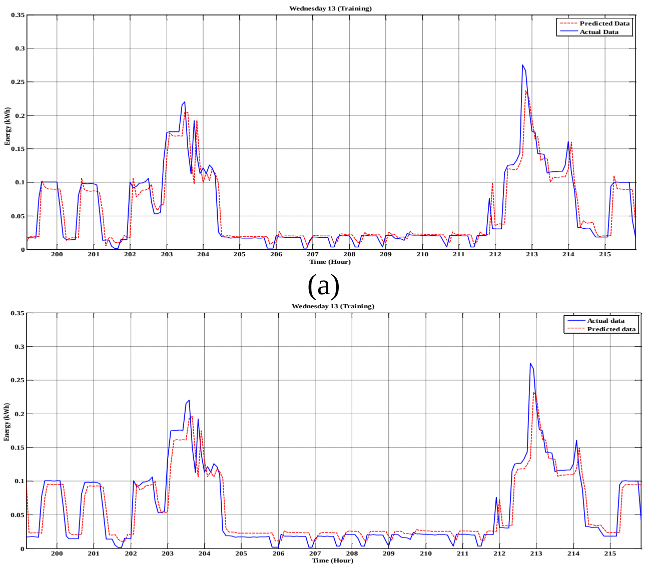


(a)



(b)

Fig. 14.



(a)

(b)

Fig. 15.

5. Conclusion

NARX network has been modified and coded NARX-GA-DE network to predict individual consumer load in this paper. NARX GA-DE architecture has one hidden neuron, and one input and output tapped delays. Energy consumption data acquired from consumer load prototype for one month which consists 8928 data points was used and 5-minute step ahead load prediction was achieved. The paper overcomes the shortcoming of our previous study in terms of the simplicity of the architecture and proposed real-time application. This will facilitate real-time monitoring of individual consumer activities connected to the power distribution network.

REFERENCES

[1] I. Koprinska, M. Rana, and V. Agelidis, "Electricity Load Forecasting: A Weekday-Based Approach," in *Artificial Neural Networks and Machine Learning – ICANN 2012*, vol. 7553, A. Villa, W. Duch, P. Érdi, F. Masulli, and G. Palm, Eds. Springer Berlin Heidelberg, 2012, pp. 33–41.

[2] A. Ghanbari, E. Hadavandi, and S. Abbasian-Naghneh, "An Intelligent ACO-SA Approach for Short Term Electricity Load Prediction," in *Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence*, vol. 6216, D.-S. Huang, X. Zhang, C. Reyes García, and L. Zhang, Eds. Springer Berlin Heidelberg, 2010, pp. 623–633.

[3] H. Chen, K. Huang, and L. Chang, "Application of neural networks for very short-term load forecasting in power systems," in *Advances in Neural Networks–ISNN 2005*, W. Jun, L. Xiao-Feng, and Y. Zhang, Eds. Berlin Heidelberg: Springer, 2005, pp. 628–633.

[4] W. Charytoniuk and M.-S. Chen, "Very short-term load forecasting using artificial neural networks," *Power Syst. IEEE Trans.*, vol. 15, no. 1, pp. 263–268, 2000.

[5] J. V Paatero and P. D. Lund, "A model for generating household electricity load profiles," *Int. J. energy Res.*, vol. 30, no. 5, pp. 273–290, 2006.

[6] J. W. Taylor, "An evaluation of methods for very short-term load forecasting using minute-by-minute British data," *Int. J. Forecast.*, vol. 24, no. 4, pp. 645–658, Oct. 2008.

[7] R. P. Singh, P. X. Gao, and D. J. Lizotte, "On hourly home peak load prediction," in *Third International Conference on Smart Grid Communications* , 2012, pp. 163–168.

[8] D. J. Trudnowski, S. Member, W. L. Mcreynolds, J. M. Johnson, and A. Member, "Real-Time Very Short-Term Load Prediction for Power-System Automatic Generation Control," vol. 9, no. 2, pp. 254–260, 2001.

[9] C. O. Adika and L. Wang, "Short term energy consumption prediction using bio-inspired fuzzy systems," in *North American Power Symposium (NAPS)*, 2012, pp. 1–6.

[10] N. I. Sapankevych and R. Sankar, "Time series prediction using support vector machines: a survey," *IEEE Comput. Intell. Mag.*, vol. 4, no. 2, pp. 24–38, 2009.

[11] W. Yan, "Toward automatic time-series forecasting using neural networks," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 23, no. 7, pp. 1028–1039, 2012.

[12] E. Diaconescu, "The use of NARX neural networks to predict chaotic time series," *WSEAS Trans. Comput. Res.*, vol. 3, no. 3, pp. 182–191, 2008.

[13] H. Y. Yang, H. Ye, G. Wang, J. Khan, and T. Hu, "Fuzzy neural very-short-term load

forecasting based on chaotic dynamics reconstruction," *Chaos, Solitons & Fractals*, vol. 29, no. 2, pp. 462–469, 2006.

[14] C. Y. Fok and M. I. Vai, "Very Short Term Load Forecasting for Macau Power System," in *Intelligent Computing Technology*, H. De-Shuang, J. Changjun, B. Vitoantonio, and C. F. Juan, Eds. Berlin Heidelberg: Springer, 2012, pp. 538–546.

[15] A. Andalib and F. Atry, "Multi-step ahead forecasts for electricity prices using NARX: a new approach, a critical analysis of one-step ahead forecasts," *Energy Convers. Manag.*, vol. 50, no. 3, pp. 739–747, 2009.

[16] Y.-W. Lee and T.-L. Chang, "Application of NARX neural networks in thermal dynamics identification of a pulsating heat pipe," *Energy Convers. Manag.*, vol. 50, no. 4, pp. 1069–1078, 2009.

[17] S. Çoruh, F. Geyikçi, E. Kılıç, and U. Çoruh, "The use of NARX neural network for modeling of adsorption of zinc ions using activated almond shell as a potential biosorbent," *Bioresour. Technol.*, 2013.

[18] L. Tsungnan, B. G. Horne, P. Tino, and C. L. Giles, "Learning long-term dependencies in NARX recurrent neural networks," *Neural Networks, IEEE Trans.*, vol. 7, no. 6, pp. 1329–1338, 1996.

[19] Y. Gao and M. J. Er, "NARMAX time series model prediction: feedforward and recurrent fuzzy neural network approaches," *Fuzzy sets Syst.*, vol. 150, no. 2, pp. 331–350, 2005.

[20] A. I. Abdullateef, E.-J. Salami, and B. I. Tijani, "Consumer Load Prediction Based on NARX for Electricity Theft Detection," in *International Conference on Computer and Communication Engineering (ICCCE)*, 2016, pp. 294–299.

[21] H. T. Siegelmann, B. G. Horne, and C. L. Giles, "Computational capabilities of recurrent NARX neural networks," *Syst. Man, Cybern. Part B Cybern. IEEE Trans.*, vol. 27, no. 2, pp. 208–215, 1997.

[22] R. Storn and K. Price, "Differential Evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces. International Computer Science Institute, Berkeley," CA, 1995, Tech. Rep. TR-95–012, 1995.

[23] S. Das and P. N. Suganthan, "Differential Evolution: A Survey of the State-of-the-Art," *Evol. Comput. IEEE Trans.*, vol. 15, no. 1, pp. 4–31, 2011.

[24] V. Kenneth, "An introduction to differential evolution, New ideas in optimization." McGraw-Hill Ltd., UK, Maidenhead, UK, 1999.

[25] R. Storn and K. Price, "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," *J. Glob. Optim.*, vol. 11, no. 4, pp. 341–359, 1997.

[26] Q. Anyong, "Dynamic differential evolution strategy and applications in electromagnetic inverse scattering problems," *Geosci. Remote Sensing, IEEE Trans.*, vol. 44, no. 1, pp. 116–125, 2006.

[27] F. Neri and E. Mininno, "Memetic Compact Differential Evolution for Cartesian Robot Control," *Comput. Intell. Mag. IEEE*, vol. 5, no. 2, pp. 54–65, 2010.

[28] N. Noman and H. Iba, "Inferring gene regulatory networks using differential evolution with local search heuristics," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 4, no. 4, pp. 634–647, 2007.

[29] C.-H. Sun, C.-C. Chiu, C.-L. Li, and C.-H. Huang, "Time Domain Image Reconstruction for Homogenous Dielectric Objects by Dynamic Differential Evolution," *Electromagnetics*, vol. 30, no. 4, pp. 309–323, 2010.

[30] J. H. Holland, *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. U Michigan Press, 1975.

[31] T. V Mathew, "Genetic algorithm," *Indian Inst. Technol. Bombay, Mumbai*, 2012.

[32] M. J. E. Salami, I. B. Tijani, A. I. Abdullateef, and M. A. Aibinu, "Hybrid algorithm for NARX network parameters' determination using differential evolution and genetic algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 53, no. 1, p. 12062, 2013.