Short Term Load Demand Forecasting for Transnet Port Terminal (TPT) in East London using Artificial Neural Networks

Mncedisi S Figlan, Elisha D. Markus

Department of Electrical and Electronics Engineering, Central University of Technology, Free State, South Africa.

mfiglan@gmail.com; emarkus@cut.ac.za

Abstract **- The stochastic nature of energy consumption patterns in a port varies daily and weekly. The port must manage their load to save electricity costs, plan future infrastructure development and consolidate the power utility's optimal power generation efforts. This requires accurate Short-Term Load Forecast (STLF) for quality, quantity and efficiency. Recently Artificial Neural Networks (ANNs) have been proposed for short-term load forecasting (STLF). Here we aim to create a more exact method of short-term load forecasting using the nonlinear, autoregressive, multi-variable exogenous input (NARX) ANNs. We propose a new network architecture: an open-loop training neural network with real load and weather information, followed by a closed-loop network to produce a prediction with the predicted load as its feedback data. The proposed method utilizes its own performance to improve precision, essentially implementing a load feedback loop that is less dependent on external data. In our proposed system, mean absolute percentage errors were reached in the forecast range of 1%: 30% better on the average error by feedforward ANNs. Our test used actual reticulation network load data from Transnet Port Terminal in East London and climate statistics to predict port terminal load for a week in advance.**

Keywords - Load forecasting, Neural network, training, energy consumption

I. INTRODUCTION

As the power system networks grow steadily and their complexity increases, many elements have played an important role in the generation, demand and administration of electrical power [1]. For simple operations of a power plant, forecasting is intended to help planners make strategic decisions regarding unit engagement, hydrothermal coordination, interchange evaluations and security assessments etc. Most of the forecast models are using statistical methods or artificial intelligence algorithms, such as regression, neural networks, fuzzy logic, and expert systems [2-5]. ANNs are successful in generalization and mastering non-linear relationships between variables, ANN-based strategies are frequently preferred for STLF problems [6].

A. Overview of Load Forecasting Techniques

There are two types of load forecasting models that exists. Time of day models and dynamic models. Time of day consist of values that are predicted for each hour for the duration of forecasting and uses an approach that is not dynamic. The second model is a dynamic model which understands the load's latest behavior not limited by the function of the time of the day. Factors such as weather, time, customer class etc. contain a function which differs that can be represented mathematically by the load given by:

$$\mathcal{L} = L_n + L_w + L_s + L_r \tag{1}$$

Where $\mathcal{L}$ is the total estimated load, L_n represents the normal part of the load, L_w represents the weather sensitive part of the load, L_s is a special event component that creates a substantial deviation from the usual load pattern, and L_r is a complete random term, the noise. The representation of multiplicative approach can be formulated as.

$$\mathcal{L} = L_n \bullet F_w \bullet F_s \bullet F_r \tag{2}$$

Where L_n is the normal (base) load, and the correction factors F_w, and F_r are positive numbers that can increase or decrease the overall load. These corrections are based on current weather (F_w), special events (F_s) and random fluctuation (F_r) which include factors such as electricity pricing or load growth factor.

B. Classification of Forecasting Methods

The approaches used for load forecasting are the same day approach, models of regression, time series, neural networks, expert systems, fuzzy logic, statistical learning algorithms. Such methods can be categorized according to their statistical analytical levels in the forecast models.

- Regression model [7]
- Time series [8]
- Time-of-day models [[9]
- Similar-day Approach [10]
- Stochastic time series models [11]
- Intelligent system, based models (i.e. ANNs and GA [12]

II. RELATED WORK

The following reviewed papers illustrate a variety of solutions for load forecasting related problems using different methods, particularly for short-term load forecasting, thus a universal distinction regarding shortcomings of various techniques has been drawn and the proposed approach could be a superior attempt.

Different researchers have used different methods to address a load forecasting. For example, Sharif and Taylor, 2005 presented a multi-layer feed-forward neural network model with the aim to compare the forecasting accuracy of a time-series and an ANN-based model [13]. The ANN-based model gave reasonable results. [14] evaluated the impact of electricity prices in a load forecasting model. This assessment would typically be suitable for areas where sudden electricity tariff increases are experienced as it greatly affects the forecasting accuracy. In [15], a supervised neural network – based model was used to forecast the load in the Nigerian power system. The study however did not consider the influences due to weather conditions, thus the accuracy could be improved. Furthermore, [16] have discussed the effect of the temperature on the load trend using an integrated ANN-based method. The study concluded that the integrated model resulted in less error of prediction. Among other weather variables, only the temperature was incorporated in the model, thus a consideration of other factors would greatly improve the result. Other works by [17] presented a feed forward and feedback multi-context artificial neural network (FFFB – MCANN) as a practical approach for load forecasting. They have proposed the use of the rate values rather than the absolute to produce better accuracy.

III. METHODOLOGY AND RESEARCH DESIGN

A. Investigation Methods

Data collection is a complicated exercise, so it is necessary to define a combination of research methodologies. Quantitative research method is used in this work as a tool for obtaining the required data. Some data are obtained through questionnaires, personal interviews, and equipment (power meters) for numerical data measurements, among other research investigation techniques.

Questionnaires – typically have been designed for energy end - users, i.e. Grain Elevator, Millwright Workshop, Office building, etc. to provide information about the pattern of electricity use.

Personal interviews – This arrangement is ideal for interacting with system operators, maintenance teams and meteorology officers. In order to determine critical monitoring points or priority areas, such link structures are necessary.

Modelling – Selected input data are carefully interpreted and standardized to avoid over - fitting and redundancy before being introduced as model input.

B. Sampling Methodology

The selected method does not need to be sampled. However, it is necessary to validate the forecasts by comparing the model output with the actual load data. The measurement points in the Port Terminal ring network must also be considered.

C. Data Collection, Normalization, Bad Data Detection, Treatment and Architecture

East London Port Terminal historic load data has been obtained using measurements (i.e. using Landis+gyr E 650 meters) and weather data from the weather services of a local weather office. The weather-related data are essential for the model. The collected data were then analyzed, interpreted, normalized and assigned to the model in a simplified manner.

TABLE I. SUMMARY OF DATA COLLECTED

Data Type	**Data Collected**			
	Data Source	***Period***	***Unit***	***Duration***
Depart-mental Load	Landis gyr meter	30 minutes	Kw	7 months
Total Load	Main subsattion	30 minutes	Mw	7 months
Weather	SA Weather	30 minutes	Temp. Rain Wind Humidity	7 months

TABLE II. ANN ARCHICTURE

ANN Architecture	
Input Nodes Hidden	One node for each input variable
Layers Hidden	1
Nodes Output	Equal to number of input nodes
Nodes	Equal to size of forecasting horizon (1 node)
Interconnection	Full
Activation Function	Sigmoid Function $S(t) = \frac{1}{1+e^{-t}}$ (3)
Learning Algorithm	Levenberg-Marquardt backpropagation

TABLE III. SAMPLE RAW DATA-EXCEL FORMAT

Time	Load (Mw)	Time	Load (Mw)	Time	Load (Mw)
0:30:00	659	8:30:00	829	16:30:00	833
1:00:00	586	9:00:00	862	17:00:00	833
1:30:00	586	9:30:00	862	17:30:00	833
2:00:00	571	10:00:00	862	18:00:00	833
2:30:00	571	10:30:00	862	18:30:00	862
3:00:00	571	11:00:00	862	19:00:00	1017
3:30:00	571	11:30:00	845	19:30:00	1032
4:00:00	571	12:00:00	840	20:00:00	1032
4:30:00	586	12:30:00	846	20:30:00	1032
5:00:00	610	13:00:00	833	21:00:00	1012
5:30:00	663	13:30:00	823	21:30:00	961
6:00:00	748	14:00:00	813	22:00:00	896
6:30:00	824	14:30:00	823	22:30:00	821
7:00:00	840	15:00:00	833	23:00:00	751
7:30:00	824	15:30:00	833	23:30:00	682
8:00:00	824	16:00:00	833	0:00:00	640

D. Proposed Framework, NARX Model

The dynamics of the ANN using NARX can be described by its input-output relationship [18]:

$$y(t) = F[x(t), x(t-\Delta t), \ldots, x(t-n\Delta t), y(t), y(t-\Delta t), \ldots, y(t-m\Delta t)] \quad (4)$$

where n is the number of time delay steps in the input, m is the number of time delays on the feedback(output) and F is typically a nonlinear function. Notice that in Equation (4), in addition to the exogenous variables x, we incorporate the lagged output y. The time and weather variables are exogenous inputs *x-t*, and the load y_t is the endogenous input. The network is trained using the actual values of the load y_i and then used in closed-loop by feeding back each one-hour step prediction of the load y_i to produce the next 24-hour forecast. The network is trained using the 365 previous consecutive days of data in open-loop using a Levenberg– Marquardt backpropagation algorithm. The construction of the proposed neural network starts with the structure of a feedforward perceptron network in order to learn the behavior of the output (target) y at time t (y_t), by using inputs y_t , and modeled as a nonlinear functional form of a regression model for y (output layer):

$$y_t = \Phi[\beta_o + \textstyle\sum_{i=1}^{q} \beta_i\, h_{it}] \quad (5)$$

where (hidden layer)

$$h_{it} = \Psi[\gamma_{i0} + \textstyle\sum_{j=1}^{n} \gamma_{ij}\, x_{jt}] \quad (6)$$

Φ is the activation function for the output, which is $\Phi(x) = x$, the linear function; Ψ is the activation function for the hidden neurons; in our case, the logistic function of the form:

$$\Psi(t) = \frac{1}{1 + e^{-t}} \quad (7)$$

which is used to flatten or limit the neural weights; β_o is the output bias; β_i are the output layer weights; γ_{i0} is the input bias; and γ_{ij} are the weights of the input layer. i is the subindex of the q neurons, and j is the subindex of the n inputs. Combining Equations (4) and (6), we have:

$$y_t = \Phi\{\beta_o + \textstyle\sum_{i=1}^{q} \beta_i\, \Psi[\gamma_{i0} + \sum_{j=1}^{n} \gamma_{ij}\, x_{jt}]\} \quad (8)$$

Then, we add the dynamic term, an auto-regression on the output in order to describe a recurrent network where the hidden layers are described by:

$$h_{it} = \Psi[\gamma_{i0} + \textstyle\sum_{j=1}^{n} \gamma_{ij}\, x_{jt} + \sum_{r=1}^{q} \delta_{ir}\, h_{r,t-1}] \quad (9)$$

where δ_{ir} is the weight of the delayed $h_{r,t-1}$ feedback term. By replacing (10) into (6), we obtain:

$$y_t = \Phi\{\beta_o + \textstyle\sum_{i=1}^{q} \beta_i\, \Psi[\gamma_{i0} + \sum_{j=1}^{n} \gamma_{ij}\, x_{jt} + \sum_{r=1}^{q} \delta_{ir}\, h_{r,t-1}]\} \quad (10)$$

Equation (10) represents network dynamics: past output values and multiple inputs. Our model, however, accounts for only one hidden neural layer to date. We must extend the description to N layers by adding index l and the multi-dimensional nature of the *t* outputs by adding index k to yield:

$$y_t^k = \Phi\{\beta_o^k + \textstyle\sum_{l=1}^{N} \sum_{i=1}^{q} \beta_o^k\, \Psi[\gamma_{io}^l + \sum_{j=1}^{n} \gamma_{ij}^l\, x_{jt} + \sum_{r=1}^{q} \delta_{ir}\, h_{r,t-1}]\} \quad (11)$$

$$k = 1 \cdots \tau$$

The NARX Neural Network implemented in this research is described in Equation (7). Open and closed loop networks are the same except for the value of the delayed output is obtained. The open-loop network obtains the value of y from known past values of the output, and therefore, it is a regular input to the network; and the closed-loop network obtains the value from the predicted value of the output. An example of the NARX implementation can be found in [19].

IV. RESULTS

The model for load forecasting was trained using MATLAB®, *ver.* 9.0 environment using Neural Network Toolbox, *ver.* 9.0. The NARX forecast was generated in closed-loop, i.e., the network was trained in open-loop by using known values of the load at the East London harbour from Wednesday 1st to the 7th of August 2018, then the first hourly load forecast value is calculated with the trained network, and it is fed back to the input in order to obtain the second value, and so on. The open-loop fit of the NARX model is shown in Figure 1.

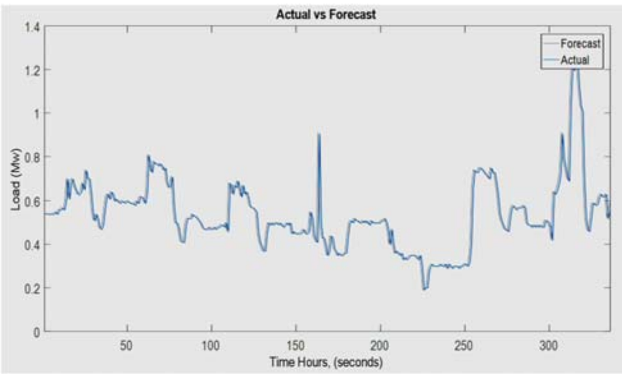


Fig. 1. Open-loop fit of the NARX ANN model for one week.

The training with the backpropagation algorithm of Levenberg-Marquardt converged after less than 12 epochs and demonstrated stability (no increase after convergence) and no overshoot (no increase before converging), as shown in Figure 2.

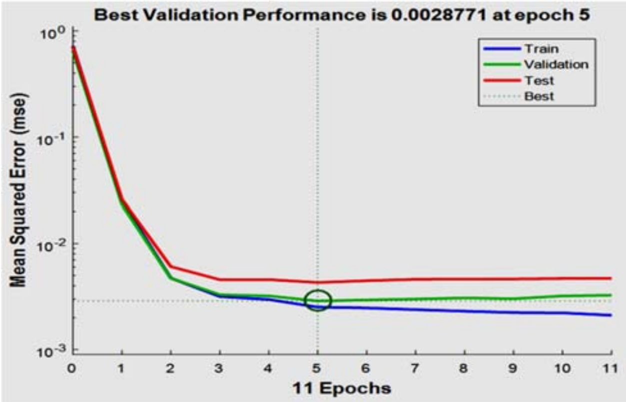


Fig. 2. Training performance graph.

In addition, the error histogram for the fit set is shown in Figure 3.

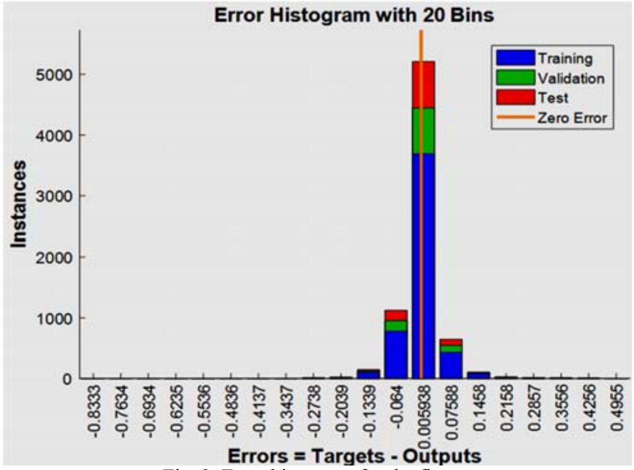


Fig. 3. Error histogram for the fit set.

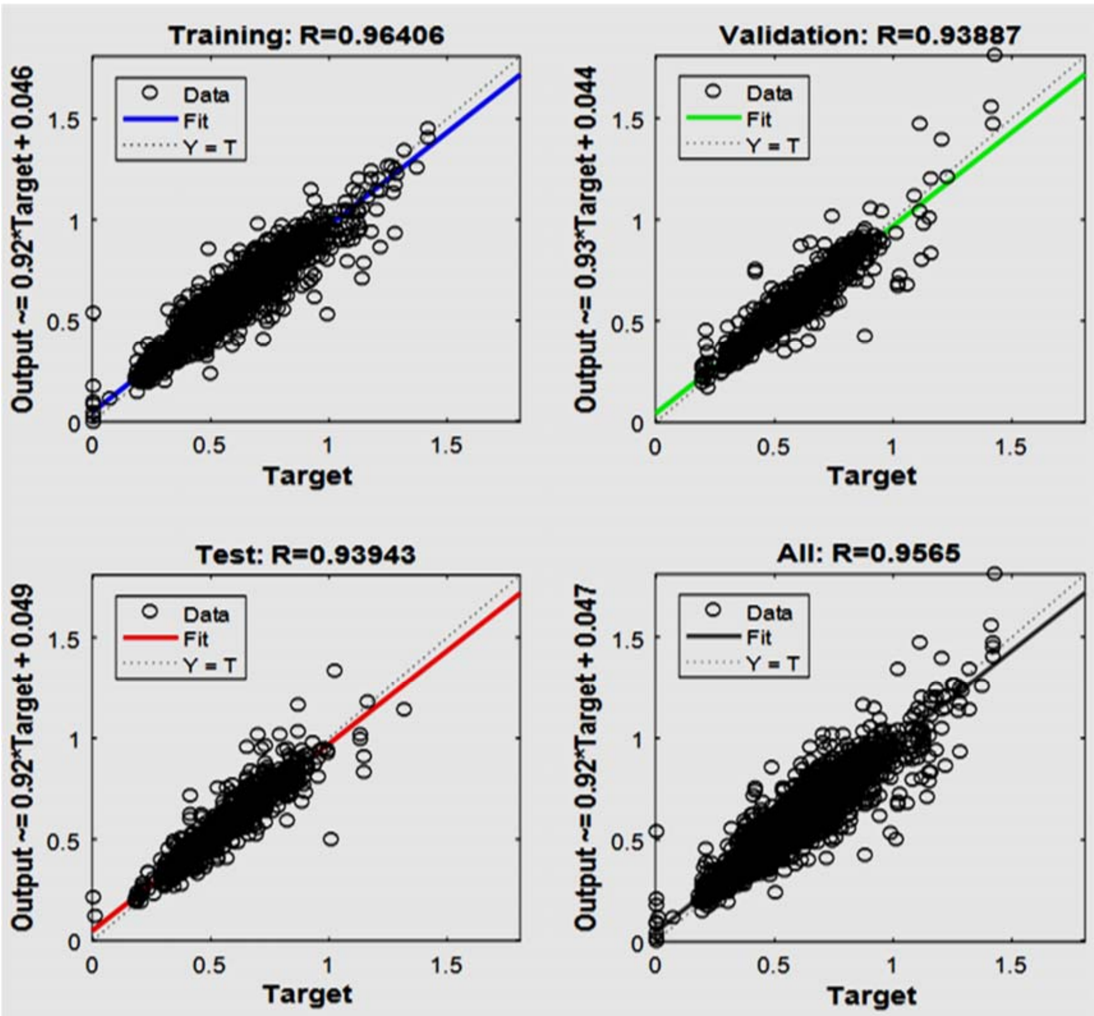
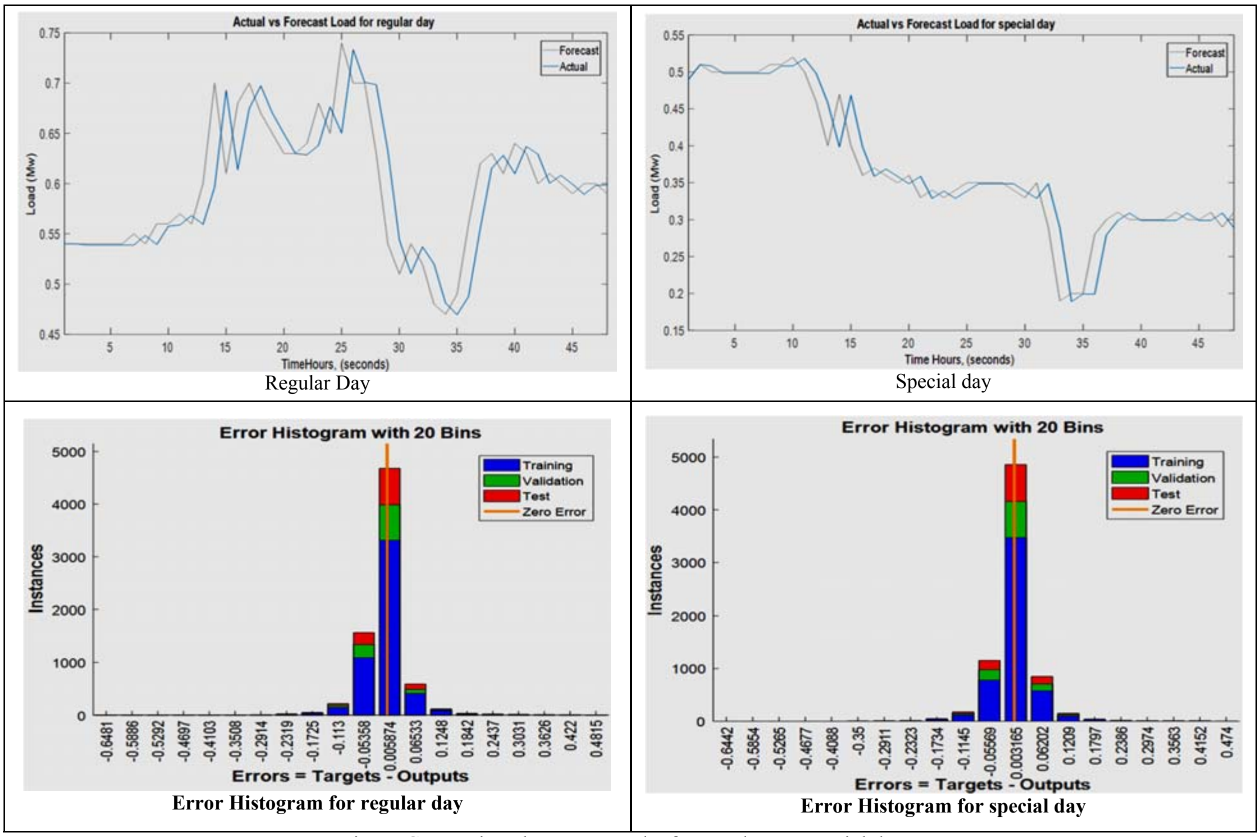


Fig. 4. Error regression on forecast vs. actual.



Regular Day

Special day

Error Histogram for regular day

Error Histogram for special day

Fig .5. Comparison between results for regular vs. special days.

V. DISCUSSION

The algorithm was run under many different conditions and configurations. The following items have been taken into consideration:

1. Various input periods (both in terms of time and duration): this is to include and exclude the effects of holidays, working and non-working days, seasons, etc. ANN NARX works well when sufficient cyclical data is available for training.
2. The number of neurons in the hidden layer in the model: simulations were run from 1–30 neurons in the hidden layer. Typically, the forecast accuracy peaks at around 20 neurons.
3. Number of delayed inputs and feedback: simulations ranging from 1 h–48 h delays have been performed. The 24-hour delay for inputs worked best.
4. The redundancy on the input data: reinforcement of training data with the most recent available data. It was found that feeding redundant data for the last week and the last 24 h worked best.

The NARX model generally provides a much better forecast throughout the time series. More study is required to evaluate the effect on the mean error and the point error of the number of hidden neurons. A comparison of the results for regular days (Tuesday to Friday) versus special days (Mondays, weekends and holidays) is shown in Figure 5. The results are slightly different, if you look on a special day for both figures indicate less consumption of electricity because most people are off as this special day is Sunday.

VI. CONCLUSION

A novel implementation of a nonlinear autoregressive neural network with exogenous input (NARX) has been presented. The approach shows that an ANN can be trained in open-loop by using all of the available endogenous and exogenous inputs. Electric load has been used as the endogenous variable, and time and weather are used as exogenous inputs. The network is a recursive ANN with connections between the output, hidden and input layers. The network was trained using a Levenberg–Marquardt backpropagation algorithm. Once the neural weights are calculated, the load input is disconnected, and the predicted (forecast) value of the output is fed back to the input. It isolates the network and reduces the need for retraining to produce each performance instance (predicted load) and increases the precision of the conventional ANN network to a forecast MAPE error of less than 1%. The accuracy in the forecast is very important, especially for borderline cases in which a plant start up may be delayed for an hour or even avoided altogether given a reliable forecast.

ABBREVIATION

The following abbreviations are used in this manuscript:

ANN Artificial Neural Network
MAPE Mean Absolute Percent Error
NARX Nonlinear Auto Regressive model with eXogenous input
STLF Short-term Load Forecasting

REFERENCE

[1] M. EL-Naggar, and A. AL-Rumaih (2005)"*Electric Load Forecasting using Genetic Based Algorithm*

[2] *Markus, Elisha Didam, O. U. Okereke, and John T. Agee. "Predicting telephone traffic congestion using multi layer feedforward neural networks." Advanced Materials Research. Vol. 367. Trans Tech Publications Ltd, 2012.*

[3] *Koyana, Ntombikayise, Elisha Didam Markus, and Adnan M. Abu-Mahfouz. "A Survey of Network and Intelligent air pollution monitoring in South Africa." 2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC). IEEE, 2019.*

[4] *Ntshako, Neo, et al. "Potable Water Quality Monitoring: A Survey of Intelligent Techniques." 2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC). IEEE, 2019.*

[5] E.A. Feinberg and D. Genethliou (2005) *"Load forecasting In: Applied Mathematics forRestructured Electric Power Systems": Optimization, Control, and Computational Intelligence*,J.H. Chow et al. (eds.), Springer

[6] Ping-Feng Pai (2006) *"Hybrid ellipsoidal fuzzy systems in forecasting regional electricity loads" Energy Conversion and Management*, *Volume 47, Issues 15-16*, *September 2006*, *Pages 2283-2289*

[7] Hayati, M. Short term load forecasting using artificial neural networks for the west of Iran. J. Appl. Sci. **2007**, 7, 1582–1588.

[8] S.Alvisi, G. Mascellani, M. Franchini and A. Bardossy, "*Water level forecasting through fuzzy logic and artificial neural network approaches"*, Hydrology and Earth System Science Discussions, 2, pp1107 – 1145, June 2005

[9] Cevik, Hasan H., and Mehmet unkas. "A Fuzzy Logic Based ShortTerm Load Forecast for the Holidays." International Journal of Machine Learning and Computing 6.1 (2016): 57.

[10] James, M. 2002. "Back-Propagation for the Uninitiated". http://www.generation5.org

[11] Eugene, A.F. and Dora, G. 2004. *Load Forecasting*. 269-285. State University of New York: New York. The Mathworks, Inc. 2004. "Back-Propagation (Neural Network Toolbox)". http://www.mathworks.com

[12] Koseleva, N.; Ropaite, G. Big data in building energy efficiency: Understanding of big data and main challenges. Procedia Eng. **2017**, 172, 544–549. [CrossRef

[13] S.S. Sharif and J.H. Taylor (2005) *"Short-term Load Forecasting by Feed Forward Neural Networks"*, Proc. IEEE ASME First Internal Energy Conference (IEC), Al Ain, United Arab Emirate, *http://www.ee.unb.ca*

[14] Hong Chen, Caludiom A. Canicares, and Ajit Singh, (2002) *"ANN-based Short-term load forecasting in Electricity Markets" Proc. IEEE Power Engineering Society Winter Meeting.*

[15] G.A.Adepoju, S.O.A Ogunjuyigbe, and K.O Alawode *"Application of Neural Network to Load Forecasting in Nigerian Electrical Power System*" The Pacific Journal of Science and Technology, Volume 8, Number 1, May 2007.

[16] B. Satish, K. S. Swarup, S. Srinivas and A. Hanumantha Rao (2004) "*Effect of temperature on short term load forecasting using an integrated ANN*" *Electric Power Systems Research*, *Volume 72, Issue 1*, *15 November 2004*, *Pages 95-101*

[17] T, Rashid and T. Kechadi (2005) *"A practical approach for electricity load forecasting"* Proceedings of World Academy of Science, Engineering and Technology, Volume 5, *http://www.waset.org*

[18] Badri, A.; Ameli, Z.; Birjandi, A.M. Application of Artificial Neural Networks and Fuzzy logic Methods for Short Term Load Forecasting. Energy Procedia **2012**, 14, 1883–1888.

[19] Bennett, C.; Stewart, R.A.; Lu, J. Autoregressive with exogenous variables and neural network short-term load forecast models for residential low voltage distribution networks. Energies 2014, 7, 2938–2960.