Application of Multilayer Perceptron Neural Network Model for Predicting Industrial Sector's Energy Consumption

Oludolapo A. Olanrewaju
Industrial Engineering Department
Durban University of Technology
Durban, South Africa
oludolapoo@dut.ac.za

Abstract

Roles played by industries in promoting economic growth makes them the highest consumer of energy. The energy used by industries has proved difficult to forecast due to the nature of the independent variables. In recent years, artificial neural network (ANN) with multilayer perceptron (MLP) backpropagation function has been broadly utilized in the machine learning area. Most noticeable is in the field of forecasting. The MLP-ANN is applied to forecast South Africa's energy consumption of five subsectors (basic chemicals, non-metallic minerals, basic iron and steel, basic non-ferrous metals and other manufacturing), with input variables in the form of activity, structure and intensity whereas energy consumed is the output from 1970 to 2016. In contrast with regression model, the results demonstrated a higher accuracy of industrial energy consumption in terms of statistical measures of performance.

Keywords
Industries, energy consumption, ANN, regression, forecast.

1. Introduction

Roles played by industries in promoting economic growth makes them the highest consumer of energy. Industries utilize electricity to put in the working state the following: industry motors and machines, lights, computers/ laptops, workplace tools and for facility heating, cooling, and ventilation equipment among other things. The development of any economy is not a long way from her energy and material assets (Kant and Sangwan 2015). The different classifications of energy resources include fossil fuels, renewable and nuclear resources (Demirbas 2000). Considering the best classification of energy resources to utilize ought to be based on economic, social, environmental and security reasons (Azadeh, Babazadeh et al. 2013). Despite the fact that renewable energy emerges from the classifications, there are numerous hindrances keeping its regular use. Numerous reasons can be calculated in for preparation of how energy consumption should be modelled in the future. Extremely huge among the reasons are the constraints of energy resources and unremitting increment in energy consumption trend.

Estimating both renewable and non-renewable energy use are fundamental for both local and industrial use. These estimates can be essential in future energy decision plans (Ermis, Midilli et al. 2007). The use of energy cannot be left out in the future energy plans (Ermis, Midilli et al. 2007). The significance of such estimation is depicted in Figure 1 below. A successful planning for future energy use will include understanding the relationship between energy and the factors responsible for the consumption, which eventually prompts its forecast. These predictive tools are generally referred to as energy models. Policies on energy can be effortlessly guided using energy models (Kialashaki and Reisel 2014).
Different energy systems can without much of a stretch be forecasted using numerical models regardless of the dependable fluctuating parameters (Kialashaki and Reisel 2014). The capacity of numerical models to examine the past and inform the future makes it pertinent to energy modelling (Kialashaki and Reisel 2014). Conventional statistical methods like the regression analysis have proved helpful to benchmark and to predict. Machine learning

(ML) like the artificial neural networks (ANNs) have gained the upper hand in this regard. The achievement of ANN can be ascribed to the following unique features – capacity to learn, self-adaptability, fault tolerance, adaptability and capacity to respond in a real time manner (Martellotta, Ayr et al. 2017).

Energy consumed by the industrial sectors has turned out to be among the hardest to investigate, model and forecast (Kialashaki and Reisel 2014). The present study aims to estimate South African industrial sector's energy consumption for the purpose of projecting the future. ANN with MLP technique will be compared alongside the conventional regression model. The article is structured as follows. Section 2 reviews few studies that focused on the application of ANN in energy forecast investigation. Section 3 details the data and methodology for this investigation. Section 4 discusses the results and section 5 concludes with future research direction.

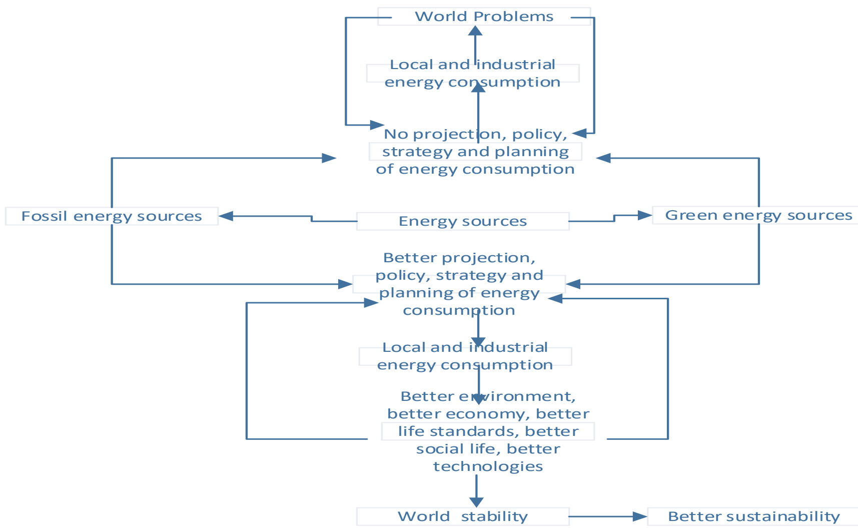

Figure 1. Benefits of forecasting energy use and planned programs. First half is without planning, second half is with planning (Ermis, Midilli et al. 2007).

2. Literature on ANN

A productive use of energy is credited to limiting energy loss because of prediction inaccuracies of energy consumption (Mate, Peral et al. 2016). Amongst the literature study that focussed on successful energy prediction through ANN techniques include (Martellotta, Ayr et al. 2017), (Jassian, Lu et al. 2018), (Kant and Sangwan 2015), (Zeng, Zeng et al. 2017), (Azadeh, Babazadeh et al. 2013), (Kialashaki and Reisel 2014), (Rossi, Velazquez et al. 2014), (Biswas, Robinson et al. 2016) and (Ermis, Midilli et al. 2007). In the use of ANN to predict the hourly energy consumption for heating of a simulated building, ANN gave a great outcome as approved by the regression coefficient (Martellotta, Ayr et al. 2017). Artificial neural network was employed in the assessment of energy consumed as well as the CO2 emitted of off-highway trucks (Jassian, Lu et al. 2018). The application likewise brought about uncovering the input factors that yielded huge contribution in the environmental effect of the trucks. To regulate the energy efficiency in equipment, ANN was employed in the model prediction of energy consumed by machines (Kant and Sangwan 2015). The resulting values were almost identical to the values observed from experiments approving the significance of ANN.

A hybrid application in the form of back propagation neural network with differential algorithm was successfully applied to predict a multifactor-influenced energy consumption (Zeng, Zeng et al. 2017). The differential algorithm was utilized to locate the appropriate initial connection weights and thresholds to permit the back propagation neural network to accomplish a progressively exact forecast. Considering the most essential factors – carbon dioxide emission, nitrogen oxide emission, carbon mono oxide emission, gas price, oil price and GDP in the prediction of renewable energy resources for the purpose of planning was effectively done utilizing ANN (Azadeh, Babazadeh et al. 2013). The investigation likewise contrasted the ANN result with those acquired by various regression and fuzzy-regression models, ANN stood out. The industrial sector of US energy consumption was effectively modelled using ANN between 1980 and 2012 and projected from 2013 to 2030. Contrasted with multiple regression analysis, ANN demonstrated superiority when coefficient of correlation was reflected (Kialashaki and Reisel 2014). To determine energy consumption baseline production in a Combined Heat and Power (CHP) plant, ANN showed satisfactory robustness. As commendable as ANN was in the investigation, a specific setback was the susceptibility to differences in the operational factors outside the limit trained by for the model (Rossi, Velazquez et al. 2014).

In the prediction of energy consumed by residential building (TxAIRE Research houses), neural network model was employed (Biswas, Robinson et al. 2016). The investigation focussed on energy consumed by the house and heat pump. An exceptionally acceptable outcome result was achieved when contrasted with statistical analysis and previous literature. Vital variables prompting to energy consumed in the residential sectors like influence of behaviour and occupants performances were however absent as the simulated house was unoccupied. The prediction of the world green energy consumption utilizing ANN while considering world coal consumption, world oil consumption, world natural gas consumption and world primary energy consumption over time as inputs was successful (Ermis, Midilli et al. 2007). Before the world green energy consumption was predicted, each of the inputs were predicted separately while considering two inputs – the year and the inputs for the world green energy. ANN demonstrated its worth in the investigation.

3. Data and Methodology

3.1 Data

Energy consumed through the independent variables in the form of activity, structure and intensity within the industries of South Africa between 1970 and 2016 from five industrial subsectors is estimated. These subsectors are manufacturing, basic non-ferrous metals, basic iron and steel, non-metallic minerals and basic chemicals. The dependent and independent variables were achieved through decomposing GDP and energy data supplied by Quantec, a consultancy that provides financial and economic data in South Africa. Figure 2 depicts the decomposed data.

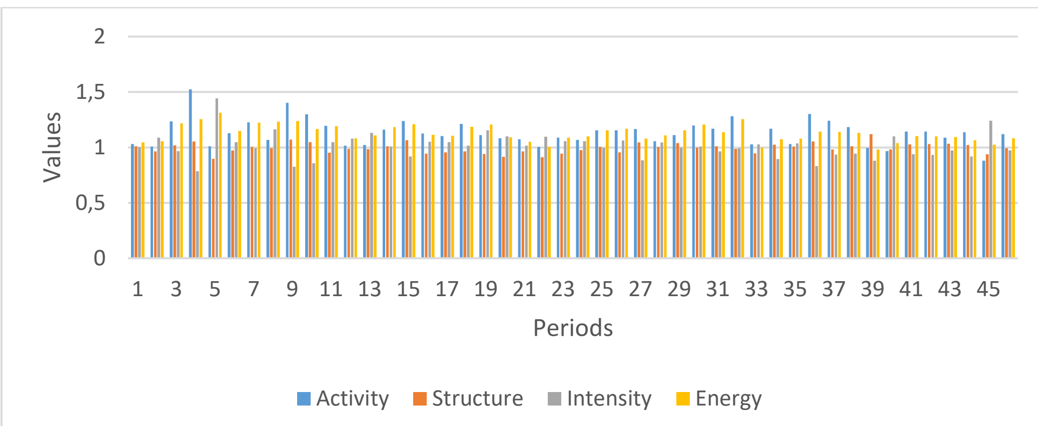

Figure 2. Data for the study

3.2 Methodology

3.2.1 ANN

ANN is a data handling framework that shares certain execution attributes for all intents and purpose with biological neural networks (Kialashaki and Reisel 2014). The computing nature of ANN model is comprised of three layers with diverse neurons within each of the layers. Within the ANN, the neurons are associated collectively to shape a system which mirrors a natural sensory structure (Kant and Sangwan 2015). The three layers are input, hidden and output layers. The layers are additionally linked to one another in a way, to the point that every neuron in one layer is associated with other neurons in the following layer (Kant and Sangwan 2015). A neural network is trained to achieve a specific capacity by altering the connection (weight) values between neurons (Kant and Sangwan 2015). The preparation of the weight values for the input and hidden units is completed through the allotting of initial weight and bias values for the simple reason that, within each input pattern, all things considered, the net input to each of the hidden units will be within the space of which the hidden neuron will learn more freely (Kialashaki and Reisel 2014) ANNs consist of qualities that tend to flawless results when applied to learn either linear or nonlinear mapping (Azadeh, Babazadeh et al. 2013). What was learnt will be transferred back in a repeated manner to the model to modify the weight values between layers to decrease the differences between the forecasted and actual data (Zeng, Zeng et al. 2017).

Neural networks have been established as an overview of numerical model of human comprehension dependent on the hypothesis stated below (Fausett 1994):

- Data processing takes place at several simple units known as neurons
- Signals are distributed among neurons over connection links
- Every connection link has as connected weight which increases the signal transferred
- Every neuron indicates an activation function to its net input to decide its output.

Exhaustive introduction to ANN modelling can be found among the referenced works. The computational formula is given below (Hsu and Chen 2003):

$$\mathrm{Y_j} = \mathrm{f}(\textstyle\sum_{\mathrm{i}} \mathrm{w_{ij}} \mathrm{X_{ij}}), \quad (1)$$

$\mathrm{Y_j}$ represents the output at node j, $f(.)$ refers to transfer function, $\mathrm{w_{ij}}$ the connection weight between nodes j and i whereas $\mathrm{X_i}$ refers to the input signal at node i. The backpropagation process for this study is the gradient descent. The process endeavours to enhance the neural network's performance by lessening the aggregate difference by varying the weight values at its gradient. The backpropagation depends on a steepest descent strategy with a momentum weight, that is the bias, this ascertains the change in weight change for the specified neuron. From the studies of (Huang, Hwang et al. 2002, Hsu and Chen 2003): $\Delta \mathrm{w_{ij}^p(n)}$ indicates the synaptic weight interfacing output neuron i to input neuron j in the pth layer during iteration. The alteration $\Delta \mathrm{w_{ij}^p(n)}$ to $\mathrm{w_{ij}^p(n)}$ is stated as

$$\Delta \mathrm{w_{ij}^p(n)} = -\eta(\mathrm{n}) \frac{\partial \mathrm{E(n)}}{\partial \mathrm{w_{ij}^p}}, \quad (2)$$

η(n) represents learning rate parameter. Through the utilization of the chain rule of differentiation, the weights of the network with the backpropagation learning rule are improved by employing the formulae below:

$$\Delta \mathrm{w_{ij}^p(n)} = \eta(\mathrm{n})\, \partial_{\mathrm{j}}^{\mathrm{p}}(\mathrm{n})\, \mathrm{X_i^{p-1}(n) m(n)} \Delta \mathrm{w_{ij}^p(n-1)}, \quad (3)$$

$$\Delta \mathrm{w_{ij}^p(n+1)} = \mathrm{w_{ij}^p(n)} + \Delta \mathrm{w_{ij}^p(n)}, \quad (4)$$

with $\partial_{\mathrm{j}}^{\mathrm{p}}(\mathrm{n})$ representing the $\mathrm{n^{th}}$ error signal at the $\mathrm{j^{th}}$ neuron of the $\mathrm{p^{th}}$ layer, $\mathrm{X_i^{p-1}(n)}$ refers to the output signal of i neuron from the below layer and m representing momentum factor.

3.2.2 Regression

Regression technique is an established and most used techniques for analysing reliance of a measure through a group of independent factors (Kialashaki and Reisel 2014). Regression analysis has its various benefits which include simple usage, easy interpretation, possibility of adjustment over the transformed variables, and the act of rational, assuming the hypothesis of normality, homoscedasticity and inter-correlation between the error and the variables responsible for the prediction (Pino-Mejías, Perez-Fargallo et al. 2017). In this study, decomposed factors represent the independent variables whereas the energy consumption represents the dependent variable. Regression equation for this study takes the form $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3$, and error $e_1$is equivalent to $y - \hat{y}$, which is the difference between the measured and predicted. With b_0, b_1, b_2 and b_3 representing the regression coefficients whereas x_1, x_2 and x_3 are the activity, structure and intensity (decomposed variables). The regression formula resolves for the coefficients through the reduction of the sum of the squares of the variations in the data from the process (least-square fit). The computation results in the following least-square fit below:

$y = 0.8858x_1 + 0.8839x_2 + 0.9336x_3$ with coefficient of correlation to be 0.8389.

3.2.3 Designing the artificial neural network

The ANN designed for this study was actualized through Weka (Waikato Environment for Knowledge Analysis) 3.6 version. For the learning parameters of ANN, network training function used was the MLP backpropagation. Sigmoid and linear activation functions were used. A trial and error method was applied to estimate the appropriate amount hidden neurons, in this case, five neurons was preferred after trying between 1 and 10 hidden neurons. Values of 0.3 and 0.2 were the learning and momentum rate utilized for this investigation. The training time and validation threshold customized for this design was 500 epochs and 20. The test mode was based on a 10-fold cross validation. The training, testing and validation data were randomly selected. Data was from 1970/71 to 2015/16 with activity, structure and intensity as inputs whereas the energy consumption served as the output. Figure 3 depicts the network design. Assessing ANN performance, conventional measurements like correlation of coefficient, mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE) and root relative squared error (RRSE) were used.

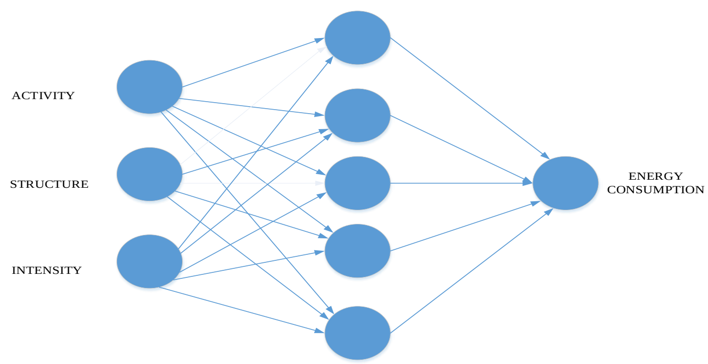

Figure 3. Architecture of proposed ANN

4. Results

Decomposition results leading to the factors responsible for industrial energy use in South Africa from year 1970/71 to 2015/16 of manufacturing, basic non-ferrous metals, basic iron and steel, non-metallic minerals and basic chemicals were considered in this investigation. The performance of ANN, as depicted in Figure 4 is truly remarkable. Correlation of coefficient, mean absolute error, root mean squared error, relative absolute error and the root relative squared error in the comparison Table 1 with regression analysis is as remarkable as comparing Figure 4 to Figure 5. The Table 1 proved ANN to be a better predictive model compared to regression model. As Figure 3 illustrates, commendable pact was achieved between the actual and forecasted values using the MLP-ANN model. The proposed MLP-ANN model was able to mimic the consumed energy throughout the period of study.

Table 1: Statistical measures of performance

	REG	MLP
Correlation coefficient	0.8389	0.9914
Mean absolute error (MAE)	0.0236	0.0063
Root mean squared error (RMSE)	0.0406	0.0105
Relative absolute error (RAE)	36.99%	9.87%
Root relative squared error (RRSE)	52.90%	13.72%

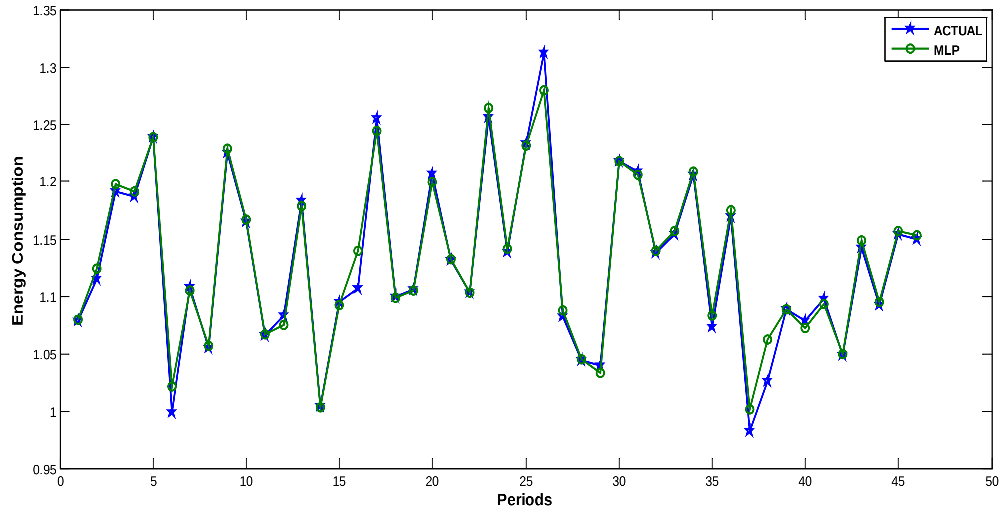

Figure 4. Visual inspection of Actual versus MLP Prediction

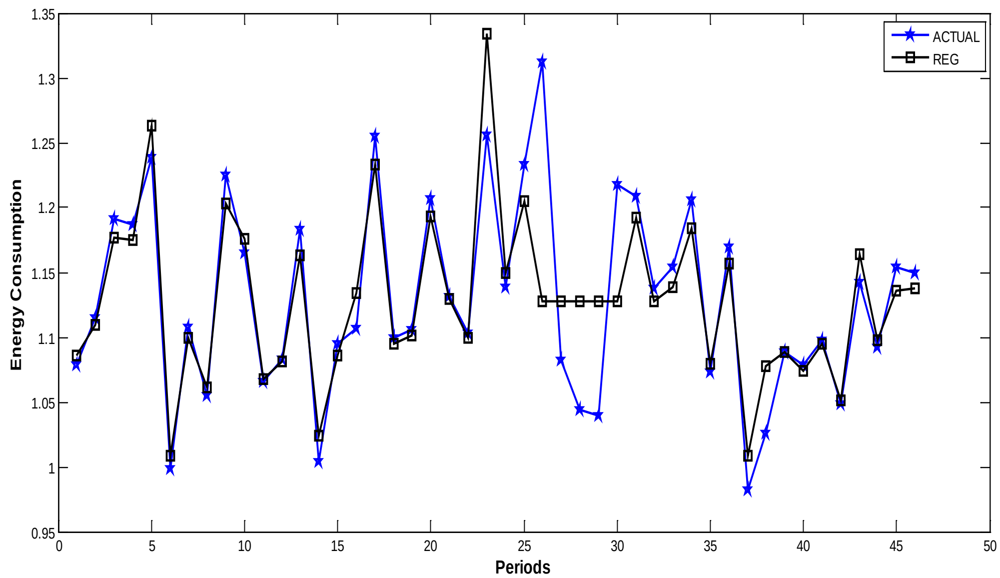

Figure 5. Visual inspection of Actual versus Regression (REG) Prediction

5. Conclusion

The applied MLP-ANN to forecast the yearly energy consumption from five industrial subsectors demonstrated an exceptionally viable result. Apart from the performance exhibited by the ANN, which can be attributed to the extensive number of information, constantly offered promising results, more remarkable results were achieved through statistical measures of performance. Reference to the ANN's use of three input variables, the relative errors resulted in extreme low values when compared to those of the regression model. This implies that ANN can give a worthy prediction of industrial energy consumption. The regression model, however, obtained good performance as evidenced in the correlation coefficient of more than 50%. It is envisaged that the results during the study can open opportunities on the road to future projection regarding industrial energy consumption.

Accurate energy demand plays an essential role for planning programs as depicted in Figure 1 above. Forecasting by the developed MLP-ANN, owing to its high precision, can be utilized for energy efficiency management decisions.

With such prevalence illustrated, this model can likewise serve to schedule energy demand forecasting in the industrial environment. Further studies will consider other ML techniques which include Support Vector Machines (SVMs) and Radial Basis Functions (RBFs) to contrast with MLP-ANN in the energy prediction. Such comparison will give light to how best to plan when energy use in the industrial sector is concerned.

References

Azadeh, A., R. Babazadeh and S. M. Asadzadeh (2013). "Optimum estimation and forecasting of renewable energy consumption by artificial neural networks." Renewable and Sustainable Energy Reviews 27: 605-612.

Biswas, R. M. A., M. D. Robinson and N. Fumo (2016). "Prediction of residential building energy consumption: A neural network approach." Energy 117: 84-92.

Demirbas, A. (2000). "Recent advances in biomass conversion technologies." Environmental Science and Technology 6: 19-40.

Ermis, K., A. Midilli, I. Dinier and M. A. Rosen (2007). "Artificial neural network analysis of world green energy use " Energy Policy 35: 1731-1743.

Fausett, L. (1994). Fundamentals of neural networks, architectures, algorithm, and application. Englewood Cliffs, Prentice Hall.

Hsu, C.-C. and C.-Y. Chen (2003). "Regional load forecasting in Taiwan-applications of artificial neural networks." Energy conversion and Management 44: 1941-1949.

Huang, H., R. Hwang and J. Hsieh (2002). "A new artificial intelligent peak power load forecaster based on non-fixed neural networks." Electric Power Energy Syst 24: 245-250.

Jassian, H. S. H., W. Lu and T. Olofsson (2018). "Assessing energy consumption and carbon dioxide emissions of off highway trucks in earthwork operations: An artificial neural network model." Journal of Cleaner Production 198: 364-380.

Kant, G. and K. S. Sangwan (2015). "Predicting Modelling for Energy Consumption in Machining using Artificial Neural Network." Procedia CIRP 37: 205-210.

Kialashaki, A. and J. R. Reisel (2014). "Development and validation of artificial neural network models of the energy demand in the industrial sector of the United States." Energy 76: 749-760.

Martellotta, F., U. Ayr, P. Stefanizzi, A. Sacchetti and G. Riganti (2017). "On the use of artificial neural networks to model household energy consumptions." Energy Procedia 126: 250-257.

Mate, A., J. Peral, A. Ferrandez, D. Gil and J. Trujillo (2016). "A hybrid integrated architecture for energy consumption prediction." Future Generation Computer Systems 63: 131-147.

Pino-Mejías, R., A. Perez-Fargallo, C. Rubio-Bellido and J. A. Pulido-Arcas (2017). "Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and CO2 emissions." Energy 118: 24-36.

Rossi, F., D. Velazquez, I. Mondero and F. Biscarri (2014). "Artificial neural networks and physical modeling for determination of baseline consumption of CHP plants." Expert Systems with Applications 41: 4658-4669.

Zeng, Y.-R., Y. Zeng, B. Choi and L. Wang (2017). "Multifactor-influenced energy consumption forecasting using enhanced back-propagation neural network." Energy 127: 381-396.

Biography

Oludolapo A. Olanrewaju is a Senior Lecturer, and Postgraduate coordinator in Industrial Engineering at the Durban University of Technology, Durban, South Africa. He earned B.Sc. in Electrical Engineering and Masters in Industrial Engineering from the University of Ibadan, Nigeria and PhD in Industrial Engineering from Tshwane University of Technology, Pretoria, South Africa. He has published journals, book chapters and conference papers. His research interests include energy and greenhouse gases, analysis, renewable energy, data management and analysis, optimization and artificial intelligence. He is member of SAIE.