

Analysis of Tanzanian Energy Demand using Artificial Neural Network and Multiple Linear Regression

Baraka Kichonge
Nelson Mandela African
Institution of Science
and Technology (NM-
AIST)
P.O. Box 447 Arusha,
Tanzania

Thomas Tesha
College of Informatics
and Virtual Education
(CIVE), University of
Dodoma (UDOM),
P.O. Box 490 Dodoma,
Tanzania

Iddi S.N. Mkilaha
College of Engineering
and Technology (CoET),
University of Dar es
Salaam (UDSM),
P.O. Box 35131 Dar es
Salaam, Tanzania

Geoffrey R John
College of Engineering
and Technology (CoET),
University of Dar es
Salaam (UDSM),
P.O. Box 35131 Dar es
Salaam, Tanzania

ABSTRACT

Analysis of energy demand is of a vital concern to energy systems analysts and planners in any nation. This paper present artificial neural network-multilayer perceptron (ANN-MLP) and multiple linear regression (MLR) techniques for the analysis of energy demand in Tanzania. The techniques were employed to analyze the influence of economic, energy and environment indicators models in predicting the energy demand in Tanzania. Statistical performance indices were used to evaluate the prediction ability of economic, energy and environment indicators models using ANN-MLP and MLR techniques. Predicted responses values of ANN-MLP and MLR techniques were then compared to determine their closeness with actual data values for determining the best performing technique. The results from ANN-MLP and MLR techniques showed the best model for predicting the energy demand in Tanzania were from energy indicators as opposed to economic and environmental indicators. The ANN-MLP prediction values had a correlation coefficient (CC) of 0.9995 and mean absolute percentage error (MAPE) of 0.67% outperforming the MLR technique whose CC and MAPE values were 0.9993 and 0.83% respectively. ANN-MLP technique graphical presentation of actual against predicted values showed close relationship between actual and predicted values as opposed to the MLR technique whose predicted values deviated much from actual values. Analysis of results from both techniques conclude that ANN-MLP outperform MLR technique in predicting energy demand in Tanzania.

Keywords

ANN, absolute error, energy demand prediction, multi linear regression.

1. INTRODUCTION

Analysis and prediction of energy demand is a subject of present extensive interest among analysts of challenges in energy production and consumption. Studies have shown energy demand to be influenced by a number of indicators such as population growth, economic performance to technological developments [1, 2]. Indicators relationships to energy and their effects on futuristic energy demand have shown varied and conflicting results not only across countries but also across methodologies for the same country and have been detailed in Soyta and Sari [3] and Mozumder and Marathe [4] studies. The conflicting results calls for scholars to analytically determine the influence of key energy indicators in the futuristic energy demand of their countries. There exists few studies for Tanzania that attempted to examine the relationship between energy indicators (variables) and energy demand. These studies were limited to

the Granger causality test and autoregressive distributed lag (ARDL)-bounds testing approaches [5, 6]. The intertemporal causal relationship between energy consumption and economic growth were examined by Odhiambo [5] and found economic growth is being spurred by energy consumption. Ebohon [6] investigated energy consumption and economic growth causal directions proxied by GDP and GNP in which a simultaneous causal relationship was shown to exist. All these studies did not attempt to predict energy demand rather the link between energy consumption and economic growth.

Developing countries such as Tanzania are in the stage of improving economically where various economic policy reformation and formulation are implemented. Economic improvement will unquestionably require a proper energy demand prediction tool as energy is important aspect in realizing a sustainable development [7]. It is the absence of sufficient studies for the energy prediction and analysis tools to the influence of the energy key indicators in Tanzania the goal of this study bases. The objective was to analyze the influence of economic, energy and environmental indicators on the prediction of energy demand by the use of artificial neural network (ANN) and multiple linear regression (MLR) techniques. This is because the ANN and MLR demonstrated strong computational abilities to handle complex non-linear functions which are the characteristics possessed by energy demand indicators [8, 9]. In the last few years, many studies have applied ANNs to energy and to mention a few are solar resource potential forecasting [10], predicting global radiation [11] predictive and adaptive heating control system [12], modeling and control of combustion processes [13] and mapping of wind speed profile for energy [14]. The study results will present policy makers with an effective and accurate tool that can predict long-term energy demand.

2. PREDICTORS

2.1 Artificial Neural Networks (ANNs)

Over the past three decades much advancement has been made in developing intelligent systems that can solve problems that cannot be programmed by conventional programming approaches. This includes the artificial neural network (ANN). In fact many researchers from different scientific disciplines designed ANNs to solve a variety of problems [15]. This approach has been widely applied in solving variety of problems in pattern recognition, prediction, optimization, associative memory, and control [15]. It provides ideal environment to which the smart world can benefit in solving unpredictable and uncontrollable problems with subtle range of influencing parameters.

In fact ANN can be defined as a highly connected array of elementary processors called neurons [16]. They are network of simple processing neurons operating on their local data and communicate with other neurons [17]. The term ‘neural network’ has its origins in attempts to find the mathematical representations of information processing in biological systems [18, 19]. Indeed, the plausibility models resemblance to biological system is true with regard to the mechanism of interconnectivity of the units and their firing when a predefined threshold limit is reached; more often termed as synaptic strength in physiology. Each neuron in the network is able to receive input signals from its preceding unit, to process them and to send an output signal to the neurons after it. There are many types of neural network but this study is confined to a specific type titled Multilayer perceptron feed-forward network (MLP).

An MLP feed-forward neural network is the most widely studied type of neural network [20] and comprises of neurons (the processing units) that are ordered into layers. The layers can be put in three different types, namely input layers which receives signal from the input variables, hidden layers which processes the input fed from the predecessor layers or the input variable and the output layer which provides the desired or target output signal. Its output layer of neurons are successively connected (fully or locally) in a feed-forward fashion with no connections between units in the same layer and no feedback connections between layers [15]. Each neuron in a particular layer is connected with all neurons in the succeeding layer (

Fig 1 depicts this). Figure 2 shows the connection between the i_{th} and j_{th} neuron as characterized by the weight coefficient w_{ij} and the i_{th} neuron by the threshold coefficient ϑ_i . The weight coefficient reflects the degree of importance of the given connection in the neural network. The output value of the i_{th} neuron x_i is determined by equations (1) and (2).

$$X_i = f(\tau_i) \dots \dots \dots (1)$$

$$\tau_i = \vartheta_i + \sum_j \omega_{ij} x_j \dots \dots \dots (2)$$

Where j consists of all predecessors of the given neuron i and $f(.)$ is a transfer (activation) function which may be sigmoid, tangent or step function.

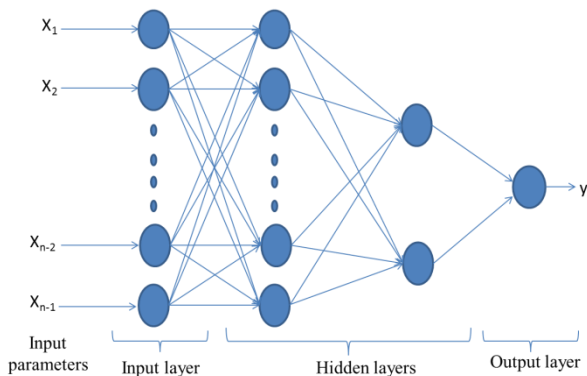


Fig 1: A multilayer feed forward neural network consisting of four layers

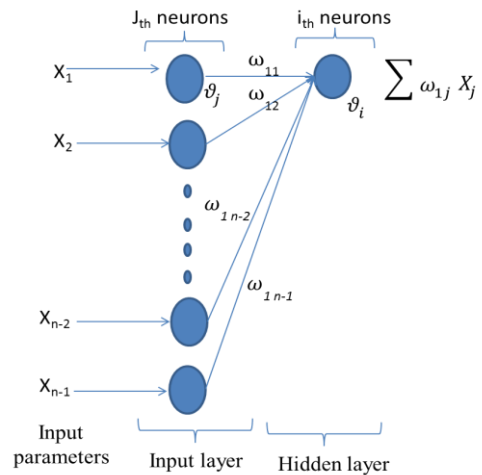


Fig 2: Connection between neurons

For the case of this research the sigmoid function of the form given in equation (3) was adopted and the summation in equation (2) is carried out over all the neurons j of the preceding layer connected to the neuron i of the current layer.

$$f(\tau) = \frac{1}{1 + \exp(-\tau)} \dots \dots \dots (3)$$

With the supervised learning process, the threshold coefficients also known as bias ϑ_i in equation (2) and weight coefficients ω_{ij} are changed to minimize the sum of the squared error between the computed and desired output values. This is done using the training data fed into the system on every neuron as information passes. The equation (4a) shows the actual minimization where x and y are the computed and desired vector for the output neurons and the summation runs over all output neurons.

$$E = \frac{1}{2} (x - y)^2 \dots \dots \dots (4a)$$

$$J(E) = \frac{1}{2} \sum_{i=1}^n (h_E(x_i) - y_i)^2 \dots \dots \dots (4b)$$

The widely used training algorithm for the MLP is the back propagation algorithm using the gradient descent applied to a sum-of-squares error function [20]. The intuition is to find how close is estimation x to the required or target y as shown in equation (4a) and update all other neurons. This can be reformulated into a cost function in equation (4b) and the network is initialized with a randomly chosen weights which is the initial guess value of E . The gradient of the error function is computed and used to correct the initial weights, this is repeatedly changed to minimize $J(E)$ until it converges to value that minimizes the Error E so that the input and output are as close as possible. With the back propagation the information update when the weight value is changed is propagated backward.

2.2 Multi Linear Regression (MLR)

Multiple linear regression (MLR) is a multivariate statistical technique that examines the relationship between a dependent variable and two or more independent variables by fitting a linear equation to observed data [21, 22]. MLR is an extension of simple linear regression analysis capable of predicting the single dependent variable using a set of known independent variables. In MLR there are p independent variables whereas the relationship amongst dependent and independent variables is given in equation (5) [22].

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + e_i \quad (5)$$

where β_0 is a constant term and β_1 to β_p represents the coefficient that relates the p independent variables. If the value of p is equaled to 1, then the equation 5 will represent a simple linear regression. MLR models have been effectively employed in the forecasting of the consumption of various commodities like electricity requirement, coal, gas and petroleum products [23, 24]. Regression analysis according to Yee [25] has been and still the most popular modeling technique in predicting energy demand.

3. METHODOLOGY

3.1 Data Collection and Preprocessing

The data collection included historical data over the period from 1990 to 2011. The data sources were from National Bureau of Statistics (NBS), World Development Indicators, International Energy Agency (IEA), Bank of Tanzania (BoT) and Tanzania Electric Supply Company Limited (TANESCO). The annual dataset used in this study for all variables included population, GDP, per capita energy use, total primary energy supply, gross national income per capita, electricity generation and green house gas emissions. The pre-processing of the data to fit in the proposed models was done. The three proposed models based on the indicators of study were economic, energy and environment indicators models. The models were proposed with the objectives of determining the influence of indicators in the prediction of energy demand.

3.2 Experimental Setup

In the experiment, two predictors were used for the study. They included the artificial neural networks (ANN) with the multilayer perceptron (MLP) architecture and the multiple linear regression (MLR). The artificial neural networks (ANN) with the multilayer perceptron (MLP) architecture is abbreviated as ANN-MLP throughout the study. The cross-validation with k -folds was also adopted for training. The training set was thus split into k chunks with $k - 1$ chunks used for the training and the remaining chunk for the validation process aimed at evaluating the model performance. In fact each of the chunk in the k splits will eventually be used as a validation against the rest. The performance measure reported by k -fold cross-validation is then the average of the values computed in the loop.

The software used for this study was Weka which is a suite of machine learning software written in Java applicable for data mining tasks [26]. Weka is composed of tools for pre-processing of data, classification, regression, clustering, association rules, and visualization [26, 27]. Weka is designed to bring a range of machine learning techniques under a common interface due to the fact that the various implementations in existence requires the data to be presented in their own format, and their own way of specifying parameters and output [28]. In fact Weka has smoothed the differences of the implementations and offers consistent method for input format, simulations and results analysis. Weka interface has made it easy enough that users need only to concern themselves with the selection of features in the data for analysis and what the output means, rather than how to use a machine learning scheme (package).

3.3 Performance Evaluation

The models' performances in both approaches were evaluated by using the following statistical parameters: correlation coefficient (CC), root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error

(MAPE) [29], root relative squared error (RRSE) and relative absolute error (RAE) [30, 31]. The values of statistical indices were derived from statistical calculation of observation in the models output predictions and are given in equations 6 - 9 [32, 33]. Selection of the best model for estimating energy demand was done considering higher correlation coefficient with the lowest root mean square error, mean absolute error and relative absolute error.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - a_i)^2} \quad (6)$$

$$RAE = \frac{\sum_{i=1}^n |P_i - a_i|}{\sum_{i=1}^n |\bar{a} - a_i|} \quad (7)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i - a_i| \quad (8)$$

$$MAPE (\%) = \frac{100}{n} \sum_{i=1}^n \frac{P_i - a_i}{P_i} \quad (9)$$

Where P_i is the actual values of P_{t+1} with $i = 1, 2, 3, 4, \dots, n$ years observations; P_i' is the average of P_{t+1} ; a_i is the predicted P_{t+1} values and n is the total observations.

4. RESULTS AND DISCUSSIONS

4.1 Training and Validation

A cross-validation technique with 10 folds was experimentally chosen for the analysis of energy demand. Data set was broken into 10 different sets of size $n/10$ also known as chunks and the training was carried out on 9 sets and testing was done on the remaining one. For each 10 experiments carried out, 9 folds were used for the training and the remaining fold for evaluation. The rotation was kept changing the evaluation fold for each iteration until each of the fold has been used for evaluation against the rest. For this case the true error was estimated by taking a mean accuracy.

4.2 Architecture Identification

The most appropriate ANN-MLP architecture was selected by considering performance indices to represent the best generalizing ability among the architectures. The first ranking experimental results for economic indicators model were from the architecture with two neuron in hidden layer (3-2-1) and CC value of 0.9983. The second and third ranking architectures for economic indicators model had CC values of 0.9983 and 0.9982 but were characterized with the highest MAE, RMSE, RAE and RRSE values as compared to the best architecture. ANN-MLP experimental results for the energy indicators model with architectures (4-4-1) showed the best accuracy as compared to other architectures that were examined. The CC values of the second and third ranking architectures had the same values as the first ranking architecture but their MAE, RMSE, RAE and RRSE values were the highest. The best results for environment indicators model were from the architecture that doubled the number of hidden neurons relative to the number of input neurons (3-6-1). The second and third ranking architectures for environment indicators model had both CC value of 0.9986 which was less than the first ranking architecture value accompanied with highest MAE, RMSE, RAE and RRSE. Results for the first ranking architecture involving economic, energy and environment indicators model are summarized in Table 1.

Table 1: ANN-MLP models performance comparison

	Economic indicators model	Energy indicators model	Environment indicators model
CC	0.9983	0.9995	0.9987
MAE	0.1331	0.0873	0.1586
RMSE	0.2108	0.1155	0.2009
RAE	3.92%	2.57%	4.67%
RRSE	5.56%	3.04%	5.29%
Architecture	4 – 2 - 1	4 – 4 - 1	3 – 6 - 1

4.3 ANN-MLP results

The economic, energy and environment indicators models' results are presented in this section to show the comparison of the predicted values against actual values for the purpose of determining the best indicators for prediction of energy demand in Tanzania based on ANN-MLP technique. The results as presented in Table 1 are the statistical parameters for performance evaluation of the models. As illustrated in Table 1, the correlation coefficient based on the energy indicators model is 0.9995 whereas the economic and environment indicators models had 0.9983 and 0.9987 respectively. The CC value of energy indicators model depicts a higher degree of correlation to energy demand as compared to economic and environment indicators model using ANN-MLP technique. The magnitude of differences between the CC values among the models is very small to determine the supremacy of the energy indicators model. Statistical performance evaluation parameters are further compared to determine the first ranking model among the three. In terms of RMSE, RAE, MAE and RRSE values, the energy indicators model values were less in comparison to economic and environment indicators model respectively. Though the results of ANN-MLP techniques are numerically close in terms of CC values, the statistical performance evaluation parameters on energy indicators model ranks the first in prediction accuracy as compared to its counterparts.

The computational of absolute percentage error (APE) using data in Table 2 and equation (10) for energy indicators model show fluctuations between 0.01% and 2.03% while in economic indicators and environment indicators model fluctuate between 0% and 4.04% and 0.27% and 4.75% respectively. The computed mean absolute percentage error (MAPE) using equation (9) for economic and environment indicators models are 0.93% and 1.19% respectively whereas in energy indicators model is 0.67%. As it can be observed the MAPE between actual and predicted values for energy indicators models are within acceptable accuracy outpacing the economic and environment indicators models. The statistical performance evaluation parameters are in favor of the energy indicators model. This is because the energy indicators perform better in comparison to its counterpart when using the ANN-MLP technique.

$$APE (\%) = \left| \frac{\text{Actual Values} - \text{Predicted Values}}{\text{Actual Values}} \right| * 100 \quad (10)$$

The graphical presentation of absolute errors deviations for economic, energy and environment indicators models using ANN-MLP technique are illustrated in Figure 3. The upper absolute errors deviation of predicted against actual values for energy indicators model is 0.214 while in economic and environment indicators models are 0.455 and 0.536 respectively. These values again confirm the energy indicators model as the best as compared to environment and economic indicators. The patterns exhibited by economic and environment indicators models have high deviations values between actual and predicted values. The absolute errors deviations of a predicted values against actual values in energy indicators model are minimal as compared to the other models. With regards to absolute errors deviations curves a conclusion is drawn that for better prediction, the energy indicators model is better in comparison to its counterparts.

4.4 Multiple Linear Regression (MLR)

Results

A summary of the MLR statistical performance results is shown in Table 3 for economic, energy and environment indicators models. The results show that the CC value for energy indicators model has a higher value as compared to economic and environment indicators models. The greater CC for energy indicators model indicates a higher correlation in predicting energy demand as opposed to economic and environment indicators using MLR technique. This implies that the prediction by energy indicators model are more accurate than that of its counterparts. The second and third are as shown in Table 3.

In Table 3 it is further shown that RMSE, RAE, MAE and RRSE for energy indicators model are correspondingly less valued as compared to the economic and environment indicators models. The lower RMSE, RAE, MAE and RRSE values as depicted by energy indicators model represents a higher accuracy in the prediction of energy demand. Using equation (10), the APE values computed from Table 4 for energy indicators model depicts fluctuations between 0.09% and 2.19% while the economic and environment indicators model fluctuates between 0.11% and 7.18% and 0.06% and 8.62% respectively. The MAPE values for economic and environment indicators models computed using equation (9) are 2.5% and 3.27% respectively while in energy indicators model is 0.83%. MAPE further shows that energy indicators model outperform its counterparts.

The graphical presentation of absolute errors deviations between actual and predicted values for all models are illustrated in Figure 4. The absolute errors values for predicted against actual values for energy indicators model are lower than the other models. The upper absolute error values deviations for energy indicators model is 0.286 while for economic and environment are 0.809 and 0.971 respectively. Looking at the three absolute errors deviations curves, the economic and environment indicators models curves exhibits the higher fluctuations over the entire dataset. It is thus drawn from the absolute errors deviations curves that for better energy prediction, the energy indicators model leads its counterparts.

Table 2: ANN-MLP technique models output comparison

YEAR	ACTUAL VALUES (MTOE)	PREDICTED VALUES (MTOE)			YEAR	ACTUAL VALUES (MTOE)	PREDICTED VALUES (MTOE)		
		Economic Indicators model	Energy Indicators model	Environment Indicators model			Economic Indicators model	Energy Indicators model	Environment Indicators model
1990	9.73	9.73	9.898	9.631	2001	14.2	14.171	14.267	14.039
1991	9.93	9.875	9.918	9.709	2002	14.92	14.653	14.944	14.548
1992	10.06	10.05	10.112	10.138	2003	15.49	15.464	15.347	15.408
1993	10.33	10.372	10.397	10.377	2004	16.2	16.286	16.202	16.332
1994	10.52	10.432	10.306	10.366	2005	17.14	17.069	17.17	16.968
1995	11.02	10.582	10.914	10.748	2006	17.81	17.883	17.803	17.858
1996	11.16	11.229	11.308	11.247	2007	18.31	18.491	18.42	18.376
1997	11.27	11.725	11.458	11.806	2008	19.1	19.134	19.029	19.019
1998	11.93	11.909	12.007	11.993	2009	19.35	19.627	19.378	19.269
1999	12.75	12.733	12.756	12.649	2010	20.04	20.095	20.05	19.961
2000	13.39	13.353	13.282	13.201	2011	20.75	20.154	20.465	20.384

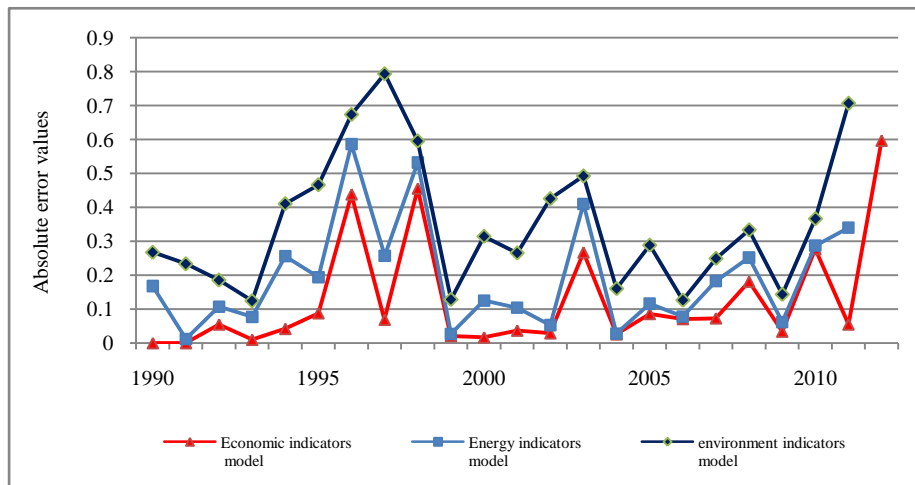


Figure 3: Absolute error values comparison among models - ANN-MLP technique

Table 3: Statistical parameters performance evaluation of models – MLR technique

	Economic indicators model	Energy indicators model	Environment indicators model
CC	0.9942	0.9993	0.9901
MAE	0.3412	0.1102	0.4431
RMSE	0.3904	0.1329	0.5123
RAE	10.06%	3.25%	13.06%
RRSE	10.29%	3.51%	13.51%

Table 4: MLR technique models output comparison

YEAR	ACTUAL VALUES (MTOE)	PREDICTED VALUES (MTOE)			YEAR	ACTUAL VALUES (MTOE)	PREDICTED VALUES (MTOE)		
		Economic Indicators model	Energy Indicators model	Environment Indicators model			Economic Indicators model	Energy Indicators model	Environment Indicators model
1990	9.73	9.163	9.88	9.013	2001	14.2	13.957	14.486	14.117
1991	9.93	9.605	10	9.387	2002	14.92	14.625	15.117	15.552
1992	10.06	10.023	9.992	9.864	2003	15.49	15.118	15.402	14.846
1993	10.33	10.397	10.194	10.278	2004	16.2	15.842	16.006	15.815
1994	10.52	10.722	10.29	10.569	2005	17.14	16.752	17.101	16.524
1995	11.02	11.163	10.945	11.409	2006	17.81	17.319	17.778	17.821
1996	11.16	11.652	11.127	11.845	2007	18.31	18.082	18.358	18.027
1997	11.27	12.079	11.23	12.241	2008	19.1	18.739	18.885	18.461
1998	11.93	12.571	12.065	12.601	2009	19.35	19.626	19.31	18.937
1999	12.75	13.059	12.901	13.009	2010	20.04	20.481	20.152	20.507
2000	13.39	13.375	13.455	13.055	2011	20.75	21.195	20.73	21.462

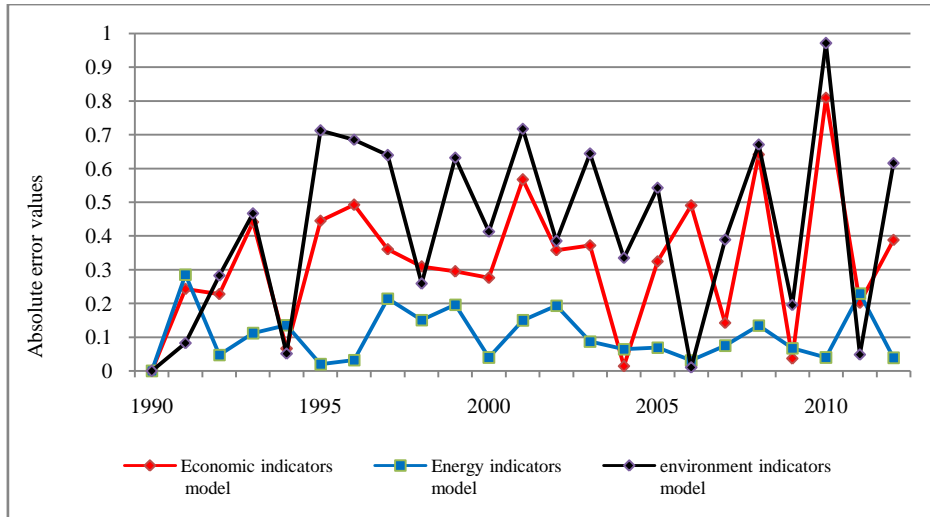


Figure 4: Absolute error values comparison among models – MLR technique

4.5 ANN-MLP and MLR Performance Comparison

The energy indicators model is shown to have a strong influence in the prediction of energy demand by outperforming the economic and environmental indicators models. Table 5 shows performance evaluation indices for energy indicators model. It is shown that the CC has greater values and MAE, RMSE, RAE, RRSE and MAPE have lesser values in the ANN-MLP in comparison to the MLR values.

Figure 5 presents the comparison between predicted values against actual values for energy indicators model in both ANN-MLP and MLR cases. The observations on the curve produced by ANN-MLP approach show that, the predicted values are close to the actual values as compared to curve produced by MLR approach. This was noted in the experimental results where RMSE, MAE, RAE, RRSE and MAPE had lower values and higher CC values. These observations show that, the ANN-MLP provides better results than MLR technique for energy demand prediction.

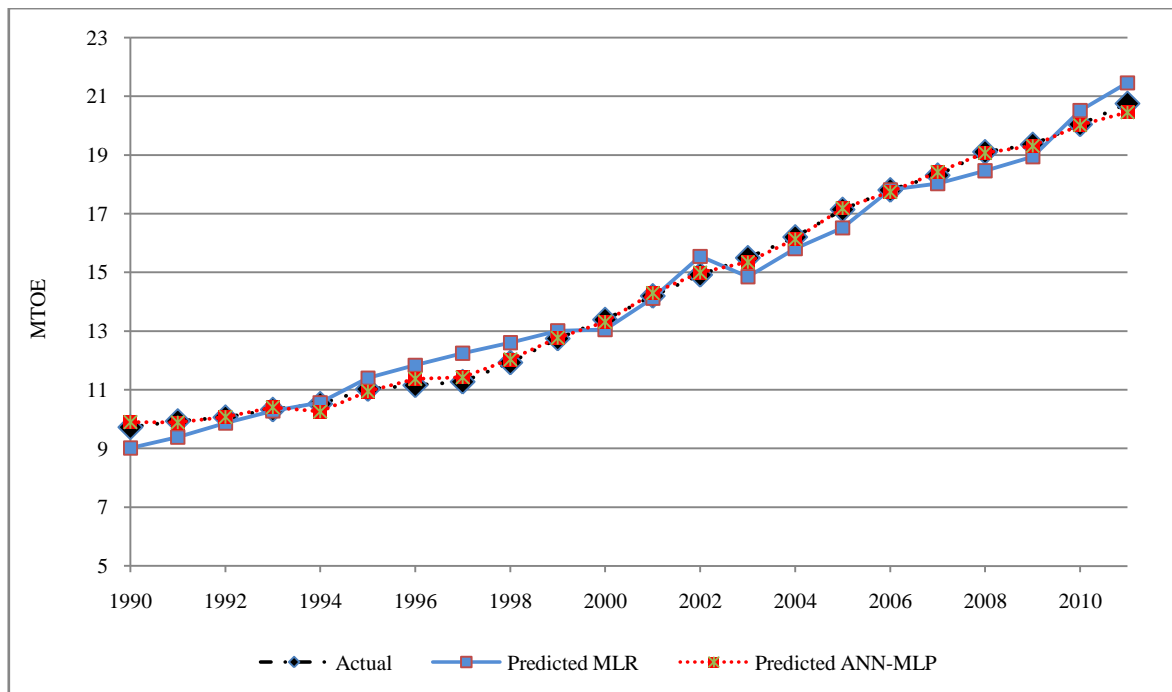


Figure 5: Comparison of actual values and predicted values for ANN-MLP and MLR techniques

Table 5: ANN-MLP and MLR performance comparison

	ANN-MLP Technique	MLR-Technique
CC	0.9995	0.9993
MAE	0.0972	0.1102
RMSE	0.1229	0.1329
RAE	2.82%	3.25%
RRSE	3.25%	3.51%
MAPE	0.67%	0.83%

5. CONCLUSION

This paper presented ANN-MLP and MLR techniques for determining an accurate prediction tool for energy demand in Tanzania using economic, energy and environment indicators models. The ANN-MLP and MLR techniques were first used to analyze separately the influence of economic, energy and environment indicators models in the energy demand of Tanzania. Then statistical performance indices were applied to evaluate the estimating ability of economic, energy and environment indicators in predicting energy demand using ANN-MLP and MLR techniques. The best performing indicators model from each techniques were then compared to determine the best energy demand prediction technique for Tanzania. Results from both ANN-MLP and MLR techniques unanimously determined that energy indicators model as the first ranking followed by economic and environment indicators models. Results for energy indicators model under ANN-MLP technique had a CC values of 0.9995 against 0.9993 for MLR technique. Comparison of the statistical performance indices showed that MAPE value of energy indicators model under ANN-MLP technique is 0.67% much better than that of MLR technique valued at 0.83%. The RMSE, MAE, RAE and RRSE values of energy indicators model under ANN-MLP technique had less numerical values as opposed to MLR technique. Additionally, ANN-MLP technique had predicted values curve close to the actual values as compared to the curve produced by MLR technique whose results deviated much from actual values. Based on the results of this study it is concluded that ANN-MLP technique outperform MLR approach in estimating energy demand for Tanzania. The study results therefore suggests energy indicators model as an accurate model in estimating energy demand of Tanzania and the ANN-MLP as the best technique in such analyses. The use of ANN-MLP technique in estimating future energy demand will endeavor government in decision making on expected energy demand for the long-run sustainable development.

6. FUTURE WORK

Further studies are needed to compare the ANN-MLP results with other algorithms for analysis of energy demand in Tanzania.

7. ACKNOWLEDGMENTS

The authors would like to extend their appreciations to The Nelson Mandela African Institution of Science and Technology (NM-AIST), Tanzania Atomic Energy Commission (TAEC), College of Engineering and Technology (CoET) of the University of Dar es Salaam, Arusha Technical College (ATC) and College of Informatics and Virtual Education (CIVE) of the University of Dodoma (UDOM) for their enabling environment that allowed successful completion of this work.

8. REFERENCES

- [1] Reister, D. B. (1987) The link between energy and GDP in developing countries, *Energy*, 12, 427-433.
- [2] Apergis, N. and Payne, J. E. (2009) Energy consumption and economic growth: Evidence from the Commonwealth of Independent States, *Energy Economics*, 31, 641-647.
- [3] Soytaş, U. and Sari, R. (2003) Energy consumption and GDP: causality relationship in G-7 countries and emerging markets, *Energy economics*, 25, 33-37.
- [4] Mozumder, P. and Marathe, A. (2007) Causality relationship between electricity consumption and GDP in Bangladesh, *Energy policy*, 35, 395-402.
- [5] Odhiambo, N. M. (2009) Energy consumption and economic growth nexus in Tanzania: An ARDL bounds testing approach, *Energy Policy*, 37, 617-622.
- [6] Ebohon, O. J. (1996) Energy, economic growth and causality in developing countries: A case study of Tanzania and Nigeria, *Energy policy*, 24, 447-453.
- [7] Vera, I. and Langlois, L. (2007) Energy indicators for sustainable development, *Energy*, 32, 875-882.
- [8] Mellit, A., Kalogirou, S. A., Hontoria, L. and Shaari, S. (2009) Artificial intelligence techniques for sizing photovoltaic systems: A review, *Renewable and Sustainable Energy Reviews*, 13, 406-419.
- [9] Mubiru, J. and Banda, E. (2008) Estimation of monthly average daily global solar irradiation using artificial neural networks, *Solar Energy*, 82, 181-187.
- [10] Sözen, A., Arcaklıoğlu, E., Özalp, M. and Çağlar, N. (2005) Forecasting based on neural network approach of solar potential in Turkey, *Renewable Energy*, 30, 1075-1090.
- [11] Azadeh, A., Maghsoudi, A. and Sohrabkhani, S. (2009) An integrated artificial neural networks approach for predicting global radiation, *Energy Conversion and Management*, 50, 1497-1505.
- [12] Morel, N., Bauer, M., El-Khoury, M. and Krauss, J. (2001) Neurobat, a predictive and adaptive heating control system using artificial neural networks, *International Journal of Solar Energy*, 21, 161-201.
- [13] Kalogirou, S. A. (2003) Artificial intelligence for the modeling and control of combustion processes: a review, *Progress in Energy and Combustion Science*, 29, 515-566.
- [14] Fadare, D. (2010) The application of artificial neural networks to mapping of wind speed profile for energy application in Nigeria, *Applied Energy*, 87, 934-942.
- [15] Jain, A. K., Mao, J. and Mohiuddin, K. (1996) Artificial neural networks: A tutorial, *Computer*, 29, 31-44.
- [16] Park, D. C., El-Sharkawi, M., Marks, R., Atlas, L. and Damborg, M. (1991) Electric load forecasting using an artificial neural network, *Power Systems, IEEE Transactions on*, 6, 442-449.
- [17] Svozil, D., Kvasnicka, V. and Pospichal, J. Í. (1997) Introduction to multi-layer feed-forward neural networks, *Chemometrics and intelligent laboratory*

- systems, 39, 43-62.
- [18] McCulloch, W. S. and Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity, *The bulletin of mathematical biophysics*, 5, 115-133.
- [19] Rosenblatt, F. (1961) Principles of neurodynamics. perceptrons and the theory of brain mechanisms (DTIC Document).
- [20] Bishop, C. M. (2006) Pattern recognition and machine learning (springer New York).
- [21] Campbell, M. J. (2001) Multiple linear regression, *Statistics at Square Two: Understanding Modern Statistical Applications in Medicine*, Second Edition, 10-31.
- [22] Tranmer, M. and Elliot, M. (2008) Multiple linear regression, *The Cathie Marsh Centre for Census and Survey Research (CCSR)*.
- [23] Sharma, D. P., Chandramohan Nair, P. and Balasubramanian, R. (2002) Demand for commercial energy in the state of Kerala, India: an econometric analysis with medium-range projections, *Energy policy*, 30, 781-791.
- [24] Bianco, V., Manca, O. and Nardini, S. (2009) Electricity consumption forecasting in Italy using linear regression models, *Energy*, 34, 1413-1421.
- [25] Yee, Y. Y. (1998) Climate and residential electricity consumption in Hong Kong, *Energy*, 23, 17-20.
- [26] Hall, M., Frank, E., Holmes, G. et al. (2009) The WEKA data mining software: an update, *ACM SIGKDD explorations newsletter*, 11, 10-18.
- [27] Witten, I. H., Frank, E., Trigg, L. E. et al. (1999) Weka: Practical machine learning tools and techniques with Java implementations.
- [28] Garner, S. R. (1995) Weka: The waikato environment for knowledge analysis, Paper presented at the Proceedings of the New Zealand computer science research students conference.
- [29] Azadeh, A., Ghaderi, S., Tarverdian, S. and Saberi, M. (2007) Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption, *Applied Mathematics and Computation*, 186, 1731-1741.
- [30] Chattefuee, S. and Hadi, A. S. (2006) Regression Analysis by Example (New Jersey, A John Wiley & Sons, Inc., Publication).
- [31] Lee, R. J. and Nicewander, W. A. (1988) Thirteen ways to look at the correlation coefficient, *The American Statistician*, 42, 59-66.
- [32] Armstrong, J. S. and Collopy, F. (1992) Error measures for generalizing about forecasting methods: Empirical comparisons, *International journal of forecasting*, 8, 69-80.
- [33] Makridakis, S. and Hibon, M. (1995) Evaluating accuracy (or error) measures (INSEAD).