Analysis of Tanzanian Biomass Consumption Using Artificial Neural Network

Thomas Tesha1* and Baraka Kichonge2
1College of Informatics and Virtual Education (CIVE), University of Dodoma, Tanzania
2Mechanical Engineering Department, Arusha Technical College (ATC), Tanzania

Abstract

The growing biomass consumption in developing countries context is being driven by a mixture of concerns over energy security, sustainable development and the climate change mitigation. The development of the comprehensive, sustainable and efficient biomass energy sector policies, strategies and investments requires proper biomass utilization and planning which in fact has not yet received the attention it deserves in the developing countries policies. This paper aims are twofold; one being to demonstrate the practicability of the application of artificial neural network multilayer perceptron (ANN-MLP) in the analysis of the biomass energy consumption and two, to identify the demographic and economic indicators which works better in the analysis and prediction of biomass consumption in Tanzania. Three models made up of Tanzania rural, Tanzania urban and Tanzania population with the addition of economic indicators were formulated for the analysis. The ANN-MLP has shown promising results with the statistical correlation coefficient of 0.9972 indicating that it can be used for practical analysis and prediction of biomass energy consumption. Furthermore the results show the use of Tanzania population model in the analysis and prediction of biomass consumption gives better results in comparison to the Tanzania rural and Tanzania urban population models individually.

Keywords: Artificial neural network multilayer perceptron; Biomass consumption; Energy planning; Energy policy

Introduction

The future energy needs are linked to the production and consumption of goods and services in an effort to combat poverty by ensuring sustainable development and the climate change mitigation in the developing countries like Tanzania. The biomass being the renewable energy, is among the major sources of energy in many developing countries as well as in parts of non-industrialized world [1,2]. Biomass rank the fourth among world’s energy resources after coal, oil and natural gas and is the largest renewable energy option that can be used to produce diverse forms of energy [3,4]. Biomass worldwide usage for thermal applications is attributed to its renewability, versatility and widespread across the globe as compared to other energy resources [4]. Biomass has a substantial impact on the environment and socio-economic development and is currently the most significant energy option for rural and urban population of Tanzania [5]. Currently, the environmental concerns, put biomass resources as an attractive future energy for Tanzania and the world at large. The country’s estimated annual demand of biomass is 40 million cubic metres although the sustainable biomass yield is 24.3 million cubic metres giving about an annual deficit of 15.7 million m3 equivalent to over 392,000 hectares [6]. The main sources of biomass in the country includes natural forests, plantation forests and agriculture wastes [5,7]. The most affected area in un-sustainable harvesting of biomass for charcoal and fuel-wood is natural forest [8].

Although defined to be mostly harmless emissions, clean energy and reduces landfills, the development of the comprehensive, sustainable and efficient biomass energy sector policies, strategies and investments requires biomass energy planning which in fact has not yet received the attention it deserves in the developing countries policy levels like Tanzania. Furthermore, the un-sustainable use of biomass has impacted a severe strain on its resources, which in turn has led to desertification and deforestation of many parts of the country [8,9]. In reality, addressing these two observation, the biomass energy planning process needs strong tools to predict consumption based on the current and past information to establish trends of the future consumption pattern. This in essence will enhance sustainable and friendly usage of biomass resources in the country while postulating future demands. In fact this will contribute to the enhancement of the long term sustainable and ensured energy sector policies, strategies and investments in the management, harvesting and supply of the biomass for utilization without making damage to the environment.

For this reason, a thorough analysis of the past consumption pattern of biomass with selected indicators is therefore important to enable determination of better and efficient tools for the prediction of the consumption. A limited number of studies in Tanzania have been carried out exclusively at the household level to explore biomass consumption. In this study, an attempt is made to analyze the influence of various demographic and economic indicators in the consumption prediction of the Tanzanian biomass. The machine learning approach explicitly artificial neural network multilayer perceptron (ANN-MLP) was adopted for the analysis. The choice of ANN-MLP was based on the fact that it has shown good performance in various energy analysis tasks. The successful applications of the ANN in general for energy and other thermal applications includes forecasting renewable energy consumption [10,11]; energy demand analysis [12,13]; forecasting net energy consumption [14] and energy and exergy analysis of refrigeration, air-conditioning and heat pump systems [15]. Findings in this study are expected to endeavor biomass energy planners in Tanzania at the policy level to make better decision based on country’s current status.
and usage patterns thus drawing baselines for future consumption expectation. Moreover, it will create awareness and sensitize efforts on realizing the sustainable usage of biomass in the country.

A brief overview of biomass energy use

Primary energy supply in the country is dominated by biomass in the form of fuel wood and charcoal accounting for approximately 90% of the total consumption [16-18]. The total sustainable energy potential of biomass from natural forests, agricultural wastes and plantation forests is approximated at 12 MTOE the largest share being from natural forests which are state owned [7,19]. It is estimated that more than 80% of energy from biomass is consumed in rural areas households in the form of fuel wood whereas the urban households consumption is dominated by charcoal [20,21]. The annual consumption of charcoal mostly for urban consumption was estimated at 1.2 million tonnes [22,23].

Urban households heavily rely on charcoal is credited to its high calorific values of 30MJ/kg than fuel wood 15MJ/kg with additional of its easy of transport and handling [9,24]. The use of charcoal in urban areas has been increasing replacing kerosene as illustrated in Figure 1. This might reflect the increase in the price of petroleum products in the international market affecting the local energy market in terms of consumption behavior [25]. Though the extent of the charcoal dominance varies across rural and urban areas with the consumption pattern in Dar es Salaam substantially differ from other urban areas. It has been shown in [9] study that biomass in the form of charcoal and firewood will continue to dominate the primary energy supply of Tanzania in both short and long term perspectives.

Machine learning

Machine learning is a field of artificial intelligence (AI) that involves studying how to automatically learn to make models that can provide intuitively reasonable results (predictions) based on past observations or experience from the environment. Environment maybe real-time observable pattern from a working agent or historical pattern extracted from an agent. The intention is to give computer the autonomy to learn from the environment and therefore act on them rather than explicitly writing program to solve such problems. This autonomy is acquired by the learning process where for a case of supervised learning the pattern (data) is presented as the training set to the learning algorithm (LA). The LA acquires knowledge or build model which is used to infer pattern from new data to make the predictions. Of a great interest in this study is the Artificial neural network.

**Artificial neural network (ANN)**

The Artificial Neural Networks represent a type of computing that is based on the way that the brain performs computations; and are good at fitting non-linear functions and recognizing patterns [26]. In fact they consist of a number of interconnected processing elements commonly referred to as neurons or nodes. Various types of ANN have been invented but the Multi-layer Perceptron neural network will be of the interest for the case of this study.

**The artificial neural network multi-layer perceptron (ANN-MLP)**

The ANN-MLP are the class of networks consisting of multiple layers of computational units, usually interconnected in a feed-forward way representing the nonlinear mapping between input vector and the output vector. Each neuron in one layer has direct connections to the neurons of the subsequent layer as depicted in Figures 2 and 3. In fact ANN-MLP has proven useful to a wide variety of applications as they provide excellent generalization performance on a wide range of problems. The advantage of the ANN-MLP in modeling is that they are proven to be universal approximators [27-29]. This means they are capable of approximating any measurable function to any desired degree of accuracy [29]. Each neuron in the network operates by taking the sum of its weighted inputs and passing the result through a nonlinear function (activation function) such as sigmoid, tangent etc. that means if \( f^p \) is the neuron in the current layer and \( f^p \) is the neuron in the preceding layer, the output of \( f^p \) neuron will take the form:

\[
X_i^p = f\left( \sum_j \omega_{ij} x_j^p \right)
\]  

(12)

Where \( f \) consists of all predecessors of the given neuron \( i \) and \( f(\cdot) \) is a transfer (activation) function.

In Figure 2, the input layer receives the variables to be analyzed and or predicted and the inner once perform computations to acquire knowledge of the pattern. In this architecture two inner layers are presented and the last layer in that order gives the output (predicted results and for this case the predicted biomass energy) on testing or in the real application.

The ANN-MLP uses the back propagation algorithm as a learning mechanism. It is the learning mechanism which has been widely studied and used for neural networks learning processes since its inception [30]. In fact the back propagation looks for the minimum of the error function in the weight space using the method of gradient descent.
also known as steepest descent [27,30]. These weights are modeled in the activation function. There many kind activation functions which have been proposed. Their common property is differentiability. All activation function must be differentiable as a requirement to back propagation. The combination of weights which minimizes the error function is considered to be a solution of the learning problem.

In biomass utilization indicator, the sigmoid activation function in (13) with the plot Figure 4 was adopted and the value of \( y \) for the case of biomass utilization indicctor was as shown in (14) as can be noted in (11)

\[
x = \frac{1}{1 + e^{-y}}
\]

(13)

\[
y_i = \theta_i + \sum_j \omega_{ij} x_j
\]

(14)

Methodology

The methodology involved data collection and preprocessing, experimental setup and performance evaluation. These are discussed in the in the subsection 2.1, 2.2 and 2.3 respectively.

Data collection and preprocessing

The historical data on the indicators which includes population (rural and urban), GDP and household biomass consumption from 1990 to 2011 were composed from National Bureau of Statistics of Tanzania (NBS), World Development Indicators and International Energy Agency (IEA). The choice of urban and rural population indicators was due to different levels of influence they have shown in biomass consumption. The most common source of fuel for thermal applications in rural population is biomass in the form of fuel wood whereas for urban is biomass in the form of charcoal. The rate of increase in the consumption of biomass is high replacing other fuels for urban population [25].

The GDP indicator was selected as a representative of economic growth as energy determines the economic growth of a country and its standards of living [31]. The influence in biomass consumption as an alternative energy forms is determined by the economic growth of a country and individuals [25]. The ‘Year’ input represents the calendar year. This was included because different years tend to have different pattern of household biomass consumption, for instance certain specific calendar year in different decades tend to receive very few rainfall, some receive normal average rainfall while others tend to receive high rainfall. This can be observed with the inter-annual rainfall variation. In fact this is in particular the case for the degree of temperature which raises the need for energy in either cooling or heating etc. and this behavioral pattern usage has to be learned by the ANN-MLP.

Three models formulated for the analysis of Tanzanian biomass consumption using artificial neural network. First model was the Tanzania urban population which had the urban population, household biomass consumption and GDP indicators. The second model was the Tanzania rural population which had the rural population, household biomass consumption and GDP indicators whereas the third model was the Tanzania population which was comprised of Tanzania population, household biomass consumption and GDP indicators. The three models were proposed with the intentions of determining the influence of selected indicators in the prediction of biomass consumption.

Cross-validation

In the experiment, the ANN-MLP was used for the study in which the sigmoid activation function was chosen and this is because it is widely studied and applied for modeling data and train with the back propagating approach. Data for all the experiments were cross-validated using \( k \)-folds cross-validation (CV). The idea was to split the data into \( k \) disjoint and equally sized subsets.

The validation was done on a single subset and training was done using the union of the remaining \( k-1 \) subsets. This procedure was repeated \( k \) times, each time with a different subset for validation. The intention was to allow for the large data in the dataset to be used for training and all cases appear for the validation cases (testing). For this case the true error was estimated as the average error rate.

Performance evaluation

The models’ performances in both approaches were evaluated by calculating the following statistical parameters: correlation coefficient [32] root mean squared error (RMSE) as in (15), mean absolute error (MAE) as in (16), relative absolute error (RAE) as in (17) and root relative squared error as in (18). The values of statistical indices were
derived from statistical calculation of observation in the models output predictions and are given in equations 15–18 [33,34]. 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 [12].

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - a_i)^2}
\]

(15)

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |P_i - a_i|
\]

(16)

\[
RAE = \frac{\sum_{i=1}^{n} |P_i - a_i|}{\sum_{i=1}^{n} |a_i|}
\]

(17)

\[
R^2 = 1 - \frac{\sum_{i=1}^{n} (P_i - a_i)^2}{\sum_{i=1}^{n} (P_i - P_{i1})^2}
\]

(18)

where \( P_i \) is the actual value of \( P_{i1} \) with \( i = 1, 2, 3, 4, \ldots, n \) years observations; \( P_{i1} \) is the average of \( P_{i1} \); \( a_i \) is the predicted \( P_{i1} \) values and \( n \) is the total observations.

**Result and Discussion**

**ANN-MLP Architecture Identification**

Various ANN-MLP architectures were generated and tested for urban, rural and the entire population models. The back propagation algorithm (BP) was used to adjust the learning procedure. The resulting best ten ANN-MLP architectures for each of the three models and their statistical parameters are as shown in Tables 1-3. In Table 1, the ANN-MLP architecture which gave optimal results was the one with identification (id.) number VII with only one hidden layer consisting of three neurons. The statistical parameters values for the best architecture in Tanzania urban population model are CC (0.997), MAE (128.2), RMSE (161.7), RAE (6.997%) and RRSE (7.9%).

Likewise, the ANN-MLP architecture which gave optimal results with the Tanzania rural population model as depicted in Table 2 is the one with identification (id.) number IV. The architecture consists of only one hidden layer with 2 neurons interconnected with weights from the input layer. In fact when testing the Tanzania rural population model with this architecture, the model gave the statistical parameters values: CC (0.997), MAE (128.9), RMSE (159.7), RAE (7.03%) and RRSE (7.81%).

The experimental results with the Tanzania population model shows the ANN-MLP architecture with the identification (id.) number VII consisting of only one hidden layer which had three neurons gave optimal results in comparison to its counterpart architectures. This is depicted in Table 3 and actually the Table shows the identified architecture gave the statistical parameters values: CC (0.9972), MAE (128.451), RMSE (191.37), RAE (7.01%) and RRSE (7.71%).

As a matter of fact, each model has shown promising and interesting performance with only one hidden layer. The experiment results shown in Tables 1-3 shows that adding more hidden layers has no significant improvement rather deterioration of the models’ performance.

The comparison of the statistical parameters for the ANN-MLP architectures which gave better results among the three models is depicted in Table 4. The CC values of Tanzania Urban and Rural population models are the lowest as compared to that of Tanzania population model. In general, the minimum values of MAE, RMSE, RAE and RRSE were obtained from ANN-MLP in Tanzania population model. Furthermore, Tanzania population model had relatively low errors when compared to the errors in the Tanzania urban and rural models. In this case, the result shows the Tanzania population model outperform its counterparts in predicting biomass consumption.

<table>
<thead>
<tr>
<th>ANN-MLP architecture id.</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of neurons in hidden layer 1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Number of neurons in hidden layer 2</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Correlation Coefficient (CC)</td>
<td>0.9964</td>
<td>0.9964</td>
<td>0.9968</td>
<td>0.996</td>
<td>0.9966</td>
<td>0.996</td>
<td>0.997</td>
<td>0.996</td>
<td>0.9965</td>
<td>0.9957</td>
</tr>
<tr>
<td>Mean absolute error (MAE)</td>
<td>143.67</td>
<td>143.62</td>
<td>131.48</td>
<td>145.4</td>
<td>144.7076</td>
<td>165.7</td>
<td>128.2</td>
<td>154.4</td>
<td>140.6848</td>
<td>157.7</td>
</tr>
<tr>
<td>Root mean squared error (RMSE)</td>
<td>175.94</td>
<td>175.45</td>
<td>165.79</td>
<td>179</td>
<td>174.4688</td>
<td>201.2</td>
<td>161.7</td>
<td>187.7</td>
<td>173.1374</td>
<td>192.79</td>
</tr>
<tr>
<td>Relative absolute error (RAE)</td>
<td>7.85%</td>
<td>7.83%</td>
<td>7.17%</td>
<td>7.93%</td>
<td>7.8827%</td>
<td>9.04%</td>
<td>6.99%</td>
<td>8.42%</td>
<td>7.67%</td>
<td>8.61%</td>
</tr>
<tr>
<td>Root relative squared error (RRSE)</td>
<td>8.60%</td>
<td>8.58%</td>
<td>8.10%</td>
<td>8.75%</td>
<td>8.529%</td>
<td>9.83%</td>
<td>7.90%</td>
<td>9.18%</td>
<td>8.46%</td>
<td>9.42%</td>
</tr>
</tbody>
</table>

**Table 1**: ANN-MLP architecture identification for Tanzania urban population model.

<table>
<thead>
<tr>
<th>ANN-MLP architecture id.</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of neurons in hidden layer 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Number of neurons in hidden layer 2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Correlation Coefficient (CC)</td>
<td>0.9968</td>
<td>0.9962</td>
<td>0.9967</td>
<td>0.997</td>
<td>0.9963</td>
<td>0.996</td>
<td>0.996</td>
<td>0.997</td>
<td>0.9963</td>
<td>0.9963</td>
</tr>
<tr>
<td>Mean absolute error (MAE)</td>
<td>141.33</td>
<td>149.44</td>
<td>141.81</td>
<td>128.9</td>
<td>145.0862</td>
<td>163.1</td>
<td>160.3</td>
<td>129.4</td>
<td>151.6215</td>
<td>150.74</td>
</tr>
<tr>
<td>Root mean squared error (RMSE)</td>
<td>168.34</td>
<td>185.79</td>
<td>169.81</td>
<td>159.7</td>
<td>178.4261</td>
<td>191</td>
<td>189.2</td>
<td>157.6</td>
<td>180.7447</td>
<td>179.74</td>
</tr>
<tr>
<td>Relative absolute error (RAE)</td>
<td>7.71%</td>
<td>8.15%</td>
<td>7.73%</td>
<td>7.03%</td>
<td>7.91%</td>
<td>8.89%</td>
<td>8.74%</td>
<td>7.06%</td>
<td>8.84%</td>
<td>8.19%</td>
</tr>
<tr>
<td>Root relative squared error (RRSE)</td>
<td>8.23%</td>
<td>9.08%</td>
<td>8.30%</td>
<td>7.81%</td>
<td>8.72%</td>
<td>9.34%</td>
<td>9.25%</td>
<td>7.70%</td>
<td>8.84%</td>
<td>8.79%</td>
</tr>
</tbody>
</table>

**Table 2**: ANN-MLP architecture identification for Tanzania rural population model.
Based on the discussions above, the optimal architectural model identified is therefore Tanzania population model and is shown in Figure 5. The optimal model has three inputs, one hidden layer and the output layer. The output for this case was the household biomass consumption in MTOE (Million Tonnes of Oil Equivalent) and the input were Year, GDP and the Tanzania population.

Analysis of absolute errors deviations for the three models is depicted in Figure 6. It appears that there is a cyclic variation in the models predictions year by year but it is evident that the absolute errors deviations depicted by Tanzania population model are the minimum among the three models. The maximum absolute error deviation exhibited by Tanzania population model is 0.27314 in 1997 against 0.308617 and 0.286515 in 2000 and 1997 for rural and urban models respectively.

Prediction curves comparison among the models is depicted in Figure 7. The predicted curves by all three models lie closer to each other than any one of them to the actual. The small statistical values variation between the models as depicted in Table 4 supports the results depicted in Figure 7. Despite this fact, the generalization performance exhibited by Tanzanian population showed better prediction values as compared to its counterparts. The superior performance of Tanzania population model is likely influenced by the fact that actual input data were represented by both rural and urban data. As a result the rural or urban categorization was not

<table>
<thead>
<tr>
<th>ANN-MLP architecture id. Number</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of neurons in hidden layer 1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Number of neurons in hidden layer 2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Correlation Coefficient (CC)</td>
<td>0.9969</td>
<td>0.9965</td>
<td>0.997</td>
<td>0.9965</td>
<td>0.9959</td>
<td>0.9959</td>
<td>0.9972</td>
<td>0.9963</td>
<td>0.9957</td>
<td>0.9961</td>
</tr>
<tr>
<td>Mean absolute error (MAE)</td>
<td>141.0639</td>
<td>142.8868</td>
<td>129.1508</td>
<td>142.274</td>
<td>162.46</td>
<td>160.95</td>
<td>128.451</td>
<td>150.596</td>
<td>163.9014</td>
<td>152.2951</td>
</tr>
<tr>
<td>Root mean squared error (RMSE)</td>
<td>167.2642</td>
<td>175.2562</td>
<td>160.1707</td>
<td>174.672</td>
<td>191.96</td>
<td>191.37</td>
<td>157.68</td>
<td>180.126</td>
<td>196.391</td>
<td>182.9918</td>
</tr>
<tr>
<td>Relative absolute error (RAE)</td>
<td>7.69%</td>
<td>7.79%</td>
<td>7.04%</td>
<td>7.76%</td>
<td>8.86%</td>
<td>8.78%</td>
<td>7.01%</td>
<td>8.21%</td>
<td>8.94%</td>
<td>8.31%</td>
</tr>
<tr>
<td>Root relative squared error (RRSE)</td>
<td>8.18%</td>
<td>8.57%</td>
<td>7.83%</td>
<td>8.54%</td>
<td>9.38%</td>
<td>9.36%</td>
<td>7.71%</td>
<td>8.81%</td>
<td>9.55%</td>
<td>8.95%</td>
</tr>
<tr>
<td>Statistical Parameters</td>
<td>Tanzania Urban Population Model</td>
<td>Tanzania Rural Population Model</td>
<td>Tanzania Population Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation Coefficient (CC)</td>
<td>0.997</td>
<td>0.997</td>
<td>0.9972</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean absolute error (MAE)</td>
<td>128.2</td>
<td>128.9</td>
<td>128.451</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root mean squared error (RMSE)</td>
<td>161.7</td>
<td>159.7</td>
<td>157.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative absolute error (RAE)</td>
<td>6.99%</td>
<td>7.03%</td>
<td>7.01%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root relative squared error (RRSE)</td>
<td>7.90%</td>
<td>7.81%</td>
<td>7.71%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: ANN-MLP architecture identification for Tanzania population model.

Table 4: Statistical parameters for the best ANN-MLP model.
Figure 6: The absolute errors deviations among model.

Figure 7: Prediction comparisons among models.

The use of machine learning approach specifically ANN-MLP gives promising results for the prediction and prediction of the biomass consumption pattern in Tanzania.

Conclusions

This paper has discussed the adoption of machine learning approach (ANN-MLP) in an attempt to analyze the influence of various demographic and economic indicators in the prediction of Tanzanian biomass consumption. The analysis has shown that:

- The use of machine learning approach specifically ANN-MLP gives promising results for the analysis and prediction of the biomass consumption in Tanzania.
- The use of Tanzania population in the prediction of biomass consumption gives better results as compared to the use of rural or urban population separately.

Despite the good generalization performance of Tanzania population, the statistical values variation between the models was small. The difference is most likely influenced by the behavioral combined effect of the usage pattern from both the rural and urban together and for this fact, the Tanzania population was identified as a better indicator in analysis of the biomass consumption in Tanzania.

Future Works

This study results showed slightly close results in all the models. The adoption of other selected machine learning techniques for the analysis of the proposed models in this study is of great interest so as to compare the results precisely.

References

of Tanzania, Dar es Salaam.


34. Makridakis S, Hibon M (1995) Evaluating accuracy (or error) measures. INSEAD.