PREDICTION OF INDIA’S ELECTRICITY DEMAND USING ANFIS

S. Saravanan¹, S. Kannan² and C. Thangaraj³

¹Department of Electrical and Electronics Engineering, Kalasalingam University, India
E-mail: s.saravanan@klu.ac.in

²Department of Electrical and Electronics Engineering, Ramco Institute of Technology, India
E-mail: kannaneeps@gmail.com

³Vignan University, India
E-mail: thangaraj.vc@gmail.com

Abstract

This study aims to provide an accurate and realistic prediction model for electricity demand using population, imports, exports, gross domestic product (GDP) and per capita gross domestic product (GDP) data for India. Four different models were used for different combinations of the above five input variables and the effect of input variables on the estimation of electricity demand has been demonstrated. In order to train the network 29 years data and to test the network 9 years data have been used. The future electricity demand for a period of 8 years from 2013 to 2020 has been predicted. The performance of the ANFIS technique is proved to be better than Multiple Linear Regression (MLR) and Artificial Neural Networks (ANN).

Keywords:
ANFIS, ANN, Exports, GDP, GNI, Imports, Load Forecasting, MLR

1. INTRODUCTION

Electricity, as a resource of energy, with its ever growing role in world economy, and its application in production and consumption, has gained special consideration. It is also one of the most significant sources for social and economic development of all nations. Any power system development starts with the forecast of future demand. Electricity demand increases due to growing population, higher per capita consumption, rapid development of industrial and commercial sectors, higher Gross Domestic Product (GDP) growth, increasing living standards and structural change in the economy. Further government policies concerning the energy sector development in the world energy markets will play a key role in the future patterns of energy production and consumption [1]. These factors make forecasting of electricity demand a relatively difficult task.

Load forecasting is very important for the reliable and economical operation of the power system. Modeling and prediction of electricity consumption play a vital role in developed and developing countries for policy makers and related organizations [2]. The long term forecasting (5 years to 20 years) is used to determine the capacity of generation, transmission or distribution system additions and the type of facilities required in transmission expansion planning. Hence, the time lines and accuracy of long term forecasting have significant effects on power system planning to construct new power generation plants and transmission facilities to meet the power demand in the future [3].

Several methods have been developed to perform accurate long term load forecasting. These methods are normally classified into two categories: conventional approaches and techniques based on Artificial Intelligence. Estimation of electricity demand is usually dependent on many socio-economic factors such as population, imports, exports, GDP, growth rate and available energy resources. Considering all parameters in energy demand modeling is a difficult task, as it requires much detailed study and also more data, for which many of the data are unavailable [4]. The underestimation of the demand would lead to potential outages that are devastating to life and economy, whereas the overestimation would lead to unnecessary idle capacity which means wasted financial resources [3, 5]. Therefore, it is very much essential to model the electricity demand with good accuracy. Also it is better to use models that can handle nonlinearities among variables as the expected nature of energy consumption data is nonlinear.

In the following section, a brief description of the literature survey about the solution approaches to the problem is given. In section 3, the concept of techniques including Multiple Linear Regression (MLR), Artificial Neural Networks (ANN) and Adaptive Neuro Fuzzy Inference System (ANFIS), data preprocessing, performance measure and testing are given. In section 4, Electricity demand forecasting model, which is developed for India, is given and results obtained by the ANFIS approach is compared with MLR and ANN and future projections are presented. Finally, the study is concluded in section 5.

2. LITERATURE SURVEY

Electricity demand modeling is becoming popular among practitioners and academicians concerned with problems of energy production and consumption [1]. Since 1980s, ANN methods received a great deal of attention and were proposed as powerful computational tools to solve the complex non-linear problem [6]. In recent years, much research has been conducted on the application of ANN to forecast long term electricity/energy demand/consumption forecasting [7-26] due to its ability to learn and construct nonlinear mapping through a set of input and output variables. The population and GDP were used as input variables to predict Iran’s annual electricity load using ANN [7]. The population, GDP, import and export were used as input variables to predict South Korea’s energy demand using ANN [8]. The population, imports, exports and Gross National Product (GNP), were used to estimate the Turkey’s electricity consumption using ANN [9]. The yearly ambient temperature, installed power capacity, yearly per resident electricity consumption, and GDP were used to predict long term energy consumption for Greek using ANN [10]. The Population, GDP, stock index, revenue from exporting industrial products and electricity consumption were
used to forecast the electricity demand in Thailand using ANN [11]. The population, GDP, average price of electricity and was
developed to forecast the electricity demand in New Zealand [12].
The population, GDP, GNP, number of households, number of
air-conditioners, amount of CO₂ pollution, index of industrial
production, oil price, amount of energy consumption, electricity
price, average temperature and maximum electric power of the
previous year were used for prediction of peak electric loads in
Japan using ANN [13].

However, ANN method has several inherent drawbacks. i.e.,
over-fitting, training, local minima, difficult determination
of network architecture, and poor generalizing performance, remain
unsolved and impede the application of the ANN approach into
practice [9].

ANFIS shows very good learning and prediction capabilities,
which makes it an efficient tool to deal with encountered
uncertainties in any system. The ANFIS is a combination of
ANN and Fuzzy system, have the benefit of two models and is
selected instead of ANN. This method covers advantage of
Fuzzy logic and ANN in same structure. The application of
ANFIS to long term electricity demand forecasting requires less
computational efforts and provides good accuracy [27].

3. TECHNIQUES

MLR, ANN and ANFIS are used to predict the long term
electricity consumption in India. Four different models which
include various combinations of input variables are selected for
analysis. The electricity consumption is the predicted output
variable for all the four models.

3.1 DATA PREPROCESSING

Data normalization is carried out only for ANN and ANFIS.
Normalization (scaling) of data within a uniform range (e.g., 0-
1) is essential (i) to prevent larger numbers from overriding
smaller ones, and (ii) to prevent premature saturation of hidden
nodes, which impedes the learning process [28]. This is
especially true when actual input data take large values. There
are different normalization algorithms which are Min-Max
Normalization, Z-Score Normalization and Sigmoid Normalization.
The Min-Max normalization scales the numbers in a data set to improve the accuracy of the subsequent numeric calculations [29] and many authors used this method. Hence Min-Max Normalization is used here.

3.2 PERFORMANCE MEASURE AND TESTING

Performance measure is a way to compare different
techniques. The following performance criterions are used: Root
Mean Square Error (RMSE), Mean Absolute Percentage Error
(MAPE) and Mean Bias Error (MBE) to each model. They are
expressed as follows:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^{n} (A_i - P_i)^2}
\]

(1)

\[
MAPE = \left(\frac{1}{n} \sum_{i=0}^{n} \left|\frac{A_i - P_i}{A_i}\right|\right) \times 100
\]

(2)

\[
MBE = \frac{1}{n} \sum_{i=0}^{n} (A_i - P_i)
\]

(3)

where, \(P_i, A_i\) are the predicted and actual values and ‘n’ is the
total number of predictions.

MAPE and RMSE measure the residual errors, which gives
a global idea of the difference between the predicted and actual
values. MBE indicates if the predicted data are over/under
estimated. All methods, except MAPE have scaled output.
MAPE method is the most suitable method to estimate the
relative error because input data used for the model estimation,
preprocessed data and raw data have different scales [30, 31].
Hence, MAPE method is used here.

3.3 MULTIPLE LINEAR REGRESSION(MLR)

MLR analysis is a technique used for modeling and analysis of
numerical data. It is an analysis of values of a dependent
(output) variable based on the values of one or more independent
(input) variables. The four models used different input variables
into consideration are as follows:

Model 1: \(y_1 = a_1x_1 + b_1x_2 + d_1x_4 + c_1 + e_1\)  
(4)

Model 2: \(y_2 = a_2x_1 + c_2x_2 + d_2x_4 + e_2\)  
(5)

Model 3: \(y_3 = a_3x_1 + b_3x_2 + e_3x_4 + d_3 + e_3\)  
(6)

Model 4: \(y_4 = a_4x_1 + d_4x_2 + f_4x_4 + e_4\)  
(7)

where, \(y\) represents the predicted electricity consumption: \(a_1; a_2; a_3; a_4; c_1; c_2; c_3; c_4; b_1; b_2; b_3; b_4; d_1; d_2; d_3; d_4\) and \(e_1\) values are regression coefficient. The \(x\) values represents the five independent (input) variables (i.e., \(x_1=\) population, \(x_2=\) import, \(x_3=\) exports \(x_4=\) per capita GDP and \(x_5=\) per capita GNI). The input parameters \((x_1-x_4)\) are estimated to
give a best fit of the data. The error term \(e\) \((e_1 \ldots e_4)\) represents
unexplained variation in the dependent variable and treated as a
random variable. Typically, the best fit is evaluated using the
least squares method. Although these models are simply a linear
and additive association of explanatory variables, they have been
extensively used with satisfactory results.

3.4 ANN

Artificial Neural Networks (ANNs) are computational
models inspired by the architecture of human brain. The ANN
is a tool of great importance in forecasting, classification
and pattern recognition. The feed forward multilayer perceptron
model with the back propagation techniques is the most popular
for prediction. This model generally consists of three layers such
as: input layer, hidden layer, and output layer. Studies show that
one hidden layer is sufficient for most forecasting problems [16].
A network can be trained to perform a particular function by
adjusting the values of the connections between elements.
Commonly, Neural networks are trained (i.e. weights are
adjusted), so that a particular input leads to a specific target
output. One of the most important parameters of a multilayer
perceptron for forecasting purposes is the number in the hidden
layer that determines the efficiency of ANN [16]. Determining
the number of hidden neurons is a critical task. The number of
hidden neurons depends mostly on the specific application of
interest and involves “trial and error” [17].

986
3.5 ANFIS

Fuzzy logic is a system that can be applied to transform linguistic variables to mathematical and computational structure for many purposes. But Fuzzy systems do not have good ability to learn and adapt to changing conditions. Combination of ANN and Fuzzy logic methods can overcome this drawback. A combined ANN and Fuzzy logic system helps researchers to choose and design parameters of Fuzzy logic inferences [29]. The structure of ANFIS is shown in Fig. 1.

![Fig.1. Structure of ANFIS](image)

ANFIS was proposed by Jang in the year 1993. ANFIS modeling refers to the method of applying various learning techniques developed in the Neural network literature to a Fuzzy Inference System (FIS) [29]. ANFIS shows very good learning and prediction capabilities, which makes it an efficient tool to deal with encountered uncertainties in any system. ANFIS can serve as a basis for constructing a set of fuzzy ‘IF-THEN’ rules with appropriate membership function to generate the stipulated input-output pairs. The membership functions are tuned to the input-output data. ANFIS is about taking an initial FIS and tuning it with a Back propagation algorithm based on the collection of input-output data [33]. FIS is the main core of ANFIS. FIS is based on expertise expressed in terms of ‘IF-THEN’ rules and can thus be employed to predict the behavior of many uncertain systems[34]. FIS advantage is that it does not require knowledge of the underlying physical process as a pre-condition for its application [35]. A FIS generally consist of five components; including fuzzification interface, rule base, database, decision making unit, and defuzzification interface. FIS type selection is one of the main steps for ANFIS development. Different methods exist for developing the FIS.

In this work, we used first order Takagi–Sugeno FIS with two fuzzy rules. It is suited for modeling nonlinear system. It is also used due to its computational efficiency and is proving to be more suitable for developing a systematic approach to generate FIS from given input-output data [29].

4. RESULTS AND DISCUSSIONS

Correlation coefficients between electricity consumption and input variables were analyzed to evaluate the influence of each variable and given in Table.1. These coefficients provide a measure of the linear relation between two variables. These input variables were selected through the calculation of correlation coefficients.

Table 1. Correlation coefficient between the independent variables and electricity consumption

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>7.134</td>
<td>6.617</td>
<td>7.027</td>
<td>6.294</td>
</tr>
<tr>
<td>MBE</td>
<td>-2.182</td>
<td>-0.714</td>
<td>-1.560</td>
<td>-4.391</td>
</tr>
<tr>
<td>RMSE</td>
<td>1.477</td>
<td>0.845</td>
<td>1.249</td>
<td>2.095</td>
</tr>
</tbody>
</table>

In ANN several network structures were tested to find most suitable model. The trial and error procedure was followed for deciding the optimal number of hidden neurons. The hidden neurons varied from 3 to 15 and the results in terms of MAPE are given in Table.3. It is clear that the best architecture (Model 1) consisted of a three layer network with three neurons in the input and eleven neurons in the hidden layer and one neuron in the output layer. The scaled conjugate gradient (‘trainscg’) method is used for training with the back propagation technique. The Fletcher-Reeves Restarts (‘traincgf’) method is not used for further analysis as it gave higher error values than ‘trainscg’ for the same set of training and testing values. Among the four models, the best fitting model in terms of MAPE is Model 4. The model has better MAPE for the forecasted electricity demands than MLR.

Table 3. MAPE for four models with respect to number of hidden neurons

<table>
<thead>
<tr>
<th>Hidden Layer Neuron (HLN)</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.1917</td>
<td>2.8667</td>
<td>2.2875</td>
<td>3.934</td>
</tr>
<tr>
<td>4</td>
<td>3.0263</td>
<td>3.2727</td>
<td>2.9175</td>
<td>4.3761</td>
</tr>
<tr>
<td>5</td>
<td>3.0376</td>
<td>2.8308</td>
<td>1.6382</td>
<td>4.319</td>
</tr>
<tr>
<td>6</td>
<td>3.7236</td>
<td>2.8309</td>
<td>3.5491</td>
<td>3.3321</td>
</tr>
<tr>
<td>7</td>
<td>3.2111</td>
<td>2.1536</td>
<td>2.726</td>
<td>4.7881</td>
</tr>
</tbody>
</table>
In ANFIS, trial and error procedure is followed for deciding the optimal number of membership function (MF) and results are given in Table 4. The MFs varied from 2, 2, 2 to 4, 4, 4 for three variable models (Model 1, Model 2 and Model 4) and 2, 2, 2 to 4, 4, 4, 4 for four variables model (Model 3). Out of all the combination Table 4 shows these MFs whose MAPE is less than 2.1% for all models. Those MFs having corresponding MAPE greater than 2.1% are neglected. It is difficult to distinguish between the accuracies of the models that were selected, but the numerical values shows that Model 1 with the ‘gbell’ of output MFs type, 3, 2, 2 number of MFs for input, gives the best result for all the analysis. Table 5 appears to have similar trend between actual and predicted data though there exists a discrepancy. The ANFIS model 1 also has much better MAPE (0.9215) for the predicted electricity demands than MLR and ANN.

Table 4. ANFIS Test Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean Absolute Percentage Error (MAPE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>NMF* MAPE</td>
</tr>
<tr>
<td>2,4,2</td>
<td>1.4085</td>
</tr>
<tr>
<td>3,2,2</td>
<td>0.9215</td>
</tr>
<tr>
<td>3,2,4</td>
<td>1.8403</td>
</tr>
<tr>
<td>4,2,2</td>
<td>1.8107</td>
</tr>
<tr>
<td>4,2,3</td>
<td>1.6397</td>
</tr>
</tbody>
</table>

NMF* Number of membership function

The details of simulation results for MLR, ANN and ANFIS model with minimum MAPE and comparison between actual and predicted electricity demand results are displayed in the Table 5. MLR have wider discrepancy between actual and predicted electricity consumption data; ANN and ANFIS appears to have similar trend between actual and predicted data through there exists a discrepancy. ANN performed better than MLR, whereas ANFIS performance is superior to ANN and MLR. The units of actual and forecasted (predicted) electricity demand data are given in Billion kilo Watt hour (bkWh).

Table 5. Electricity consumption Actual Vs Forecasted

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual data in bkWh</th>
<th>Output in bkWh</th>
<th>Predicted electricity consumption in bkWh</th>
</tr>
</thead>
<tbody>
<tr>
<td>1976</td>
<td>66,639</td>
<td>66,750</td>
<td>67,548</td>
</tr>
<tr>
<td>1980</td>
<td>82,367</td>
<td>92,107</td>
<td>82,152</td>
</tr>
<tr>
<td>1981</td>
<td>90,245</td>
<td>98,883</td>
<td>89,952</td>
</tr>
<tr>
<td>1991</td>
<td>207,645</td>
<td>188,047</td>
<td>204,526</td>
</tr>
<tr>
<td>2001</td>
<td>322,459</td>
<td>347,760</td>
<td>338,495</td>
</tr>
<tr>
<td>2003</td>
<td>360,937</td>
<td>392,210</td>
<td>360,213</td>
</tr>
<tr>
<td>2005</td>
<td>411,887</td>
<td>430,712</td>
<td>409,634</td>
</tr>
<tr>
<td>2007</td>
<td>501,977</td>
<td>504,225</td>
<td>496,476</td>
</tr>
<tr>
<td>2010</td>
<td>694,392</td>
<td>665,766</td>
<td>714,253</td>
</tr>
</tbody>
</table>

From the results in Table 5, the actual electricity demand in the year 1981 was 90,245 bkWh and in the year 2001 was 322,459 bkWh. The electricity consumption was calculated, using MLR was 98.88 bkWh, and 347.76 bkWh for the years 1981 and 2001 respectively. The MAPE, RMSE and MBE are 6.295, 2.307 and -5.303. The electricity demand was calculated, using ANN was 89.952 bkWh, and 338.495 bkWh for the year 1981 and 2001 respectively. The MAPE, RMSE and MBE are 1.159, 1.657 and -2.745. The electricity demand was calculated, using ANFIS was 89.722 bkWh, and 329.59 bkWh for the year 1981 and 2001 respectively. The MAPE, RMSE and MBE are 0.9215, 1.076 and 1.157. The results for the remaining years are also presented in Table 5. ANFIS technique produced better results than the MLR and ANN. Future electricity demand for the year 2013 to 2020 was calculated with estimated Population, Import and per capita GDP.

From the predicted results it is clear that the Model 1 of ANFIS (0.9215) and ANN (1.459) with the input variables population, Import and per capita GDP gave the least MAPE when compared to MLR. Even though the correlation coefficient between the electricity consumption and each of the input variables such as population, per capita GDP, per capita GNI (Model 4) is better, the prediction results in terms of MAPE error was not obtained better. This implies that the better correlation coefficient between electricity consumption and each of the input variables does not ensure the better result for prediction. Its results are not promising; hence the prediction is carried out in Model 1 with ANN and ANFIS. With the validated ANN and ANFIS model, India’s electricity consumption is forecasted up to 2020. The prediction is done between the years 2013 and 2020 and the results are displayed in the Table 6 numerically.

Table 6. Estimated Population, Import and Per capita GDP

<table>
<thead>
<tr>
<th>Year</th>
<th>Population in Billion</th>
<th>Import in Billion Rupee</th>
<th>Per capita GDP in Billion Rupee</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>1.27463</td>
<td>287.8185</td>
<td>525.61</td>
</tr>
<tr>
<td>2014</td>
<td>1.29211</td>
<td>323.284</td>
<td>559.96</td>
</tr>
<tr>
<td>2015</td>
<td>1.30773</td>
<td>360.873</td>
<td>595.96</td>
</tr>
<tr>
<td>2016</td>
<td>1.32420</td>
<td>400.592</td>
<td>633.64</td>
</tr>
<tr>
<td>2017</td>
<td>1.34062</td>
<td>442.442</td>
<td>672.98</td>
</tr>
<tr>
<td>2018</td>
<td>1.35698</td>
<td>486.145</td>
<td>713.98</td>
</tr>
<tr>
<td>2019</td>
<td>1.37329</td>
<td>532.515</td>
<td>756.66</td>
</tr>
<tr>
<td>2020</td>
<td>1.38955</td>
<td>580.748</td>
<td>800.99</td>
</tr>
</tbody>
</table>

Table 7. Future Electricity Consumption (bkWh) of India

<table>
<thead>
<tr>
<th>Year</th>
<th>Predicted electricity consumption in bkWh</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>ANN 872.6 ANFIS 924.0</td>
</tr>
<tr>
<td>2014</td>
<td>958.1 1003.8</td>
</tr>
<tr>
<td>2015</td>
<td>1051.1 1082.6</td>
</tr>
<tr>
<td>2016</td>
<td>1150.9 1159.5</td>
</tr>
<tr>
<td>2017</td>
<td>1256.5 1233.9</td>
</tr>
</tbody>
</table>
In ANN the prediction of electricity consumption is carried out with same number of hidden neurons and the results are given in Table.7. In ANFIS, to predict the electricity consumption, the Model 1 which gives the best results is used. The predicted results are given in Table.7 numerically.

The Central Electricity Authority, Government of India in its reports on 18th electric power survey of India predicted the electricity demand in the year 2018-2019 as 1291.86 bkWh and 2019-20 as 1388.816 bkWh. But the electricity demand calculated using ANFIS was 1305.6 bkWh and 1374.4 bkWh for the year 2018 and 2020 respectively. In comparing the results by ANFIS is very closer to the results obtained by the Government of India.

### Table 7: Predicted electricity consumption of India

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual</th>
<th>Predicted by ANFIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>1366.2</td>
<td>1305.6</td>
</tr>
<tr>
<td>2019</td>
<td>1477.7</td>
<td>1374.4</td>
</tr>
<tr>
<td>2020</td>
<td>1588.5</td>
<td>1440.4</td>
</tr>
</tbody>
</table>

5. CONCLUSION

Modeling and forecasting of electricity demand has a significant importance in developing sustainable energy policies. Accurate prediction of electricity consumption is vital, when demand grows faster. Estimation of electricity demand in India with four different models based on MLR, ANN and ANFIS is suggested. The input variables are selected through calculation of correlation coefficient. The unique feature of ANFIS technique is ideal for complex and uncertain data as it combines both ANN and Fuzzy systems. This hybrid approach can be used to predict future electricity consumption with greater accuracy compared with MLR and ANN. The proposed model of ANFIS had excellent forecasting capacity than the MLR and ANN technique in terms of MAPE with 0.92.

The electricity consumption in India between 1975 and 2020 in Model 1 are based on Population, Import, and Per Capita GDP data. The policy makers may use the forecasts by ANFIS approach to plan the new investments. The proposed model predicted the electricity consumption better than the MLR, ANN and the other three ANFIS models. The predicted electricity consumption increases approximately 8.04% annually from 2013 to 2020. Finally, in this study, it is predicted that the future electricity consumption of India would vary between 1440.4 and 1588.5 bkWh in the year 2020.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the management of Kalasalingam University, Krishnankoil, Tamil Nadu, India for their constant support and encouragement extended during this research. They are also thankful to the Department of Science and Technology, New Delhi, for its support through the project: SR/S4/MS/364/06.

REFERENCES


