Electricity Demand Forecasting in Nigeria using Time Series Model.

A.A. Mati, M.Eng.1*, B.G. Gajoga, Ph.D.2, B. Jimoh, Ph.D.2, A. Adegbeye, Ph.D.2, and D.D. Dajab, Ph.D.2

1Centre for Energy Research and Training, Ahmadu Bello University, Zaria, Nigeria.
2Department of Electrical Engineering, Ahmadu Bello University, Zaria, Nigeria.

*E-mail: aamati2003@yahoo.com
Telephone: +234 802 373 7965

ABSTRACT

Demand forecasts are generally required for the expansion, controlling and scheduling of power systems. The forecasts help in determining the optimal mix of generating capacities and which devices to operate in a given period, so as to minimize costs and secure demand even when local failures may occur in the system. Here, we employ a multiple regression time series modeling approach to model the electricity demand in Nigeria using previous data of 1970 to 2005. The model uses the electricity consumption and percentage connectivity to the national grid as the main regressors. Due to the incoherent trend observed in the historical data of 1994 to 2003, an innovative progressive reconstruction technique was employed for the trend restoration. The results of the final forecast obtained using the model proved promising up to 2015.

(Keywords: time series model, multiple regression models, electricity demand forecast)

INTRODUCTION

According to NIST/ITL (1997), time series is generally an ordered sequence of values of a variable at equally spaced time intervals. Electricity demand (or consumption) recorded over a period of time at fixed interval is a typical time series modeling problem, which is generally employed for forecasting. Sometimes it is common practice to employ a white noise sequence as input to a linear filter whose output is the power system for predicting the load time series as in Baharudin (2007). This approach involves more complex computations and the results are not any better than the time series models. On the other hand, simple regression time series models are often sufficient with promising results. Using this framework for a couple of approaches ranging from the uni-variate to multivariate, AR, ARMA and ARIMA models have been developed. Within the univariate framework, seasonal ARIMA and state space models are among the earlier set of models employed for short term forecasts. The latter became very attractive in the 1980’s because of the computational efficiency of the Kalman filter proposed by Campo and Ruiz (1987) and good performance is still reported in more recent work of Infield and Hill (1998).

ARIMA modeling has been used by many (Liang and Smith, 1987; Darbelley and Slama, 2000; Abraham and Nath, 2001) as a sophisticated benchmark for evaluating alternative proposals. However, Simpler methods such as non-parametric regression (Charytoniuk, et al., 1998), time series models and general exponential smoothing (Christiaanse, 1971), are always attractive due to the small number of parameters involved, which make them easy to implement. Recently, multiple regression and time series models have been used by Ching-Lai, et al., (2005), Obok, et al., (2008) and Parkpoom and Harrison, (2006) to forecast monthly electricity demand based on weather variables, gross domestic product, and population growth.

Time series models are generally classified as top-down models and they represent the relationship between energy consumption with time. One drawback of the top-down models is that they cannot explicitly represent technologies and, use highly aggregated data in addition to disregarding
the technically most efficient technologies, thus underestimating the potential for energy improvements (Nakata, 2004). The Bottom-up models on the other hand are based upon engineering technology models. They utilize system engineering approach by integrating detailed description of technologies used for energy consumption and generation (Wei, et al., 2006). This approach consists of classifying energy demand according to its major end uses through projections of economic activity by sector, population, urbanization, electrification, etc., as determinants of energy demand as in Sathaye (1995). An ideal case would be when the stock of appliances and their usage patterns in households are known and details about the composition of the load are valued, as in Capasso et al., (2005). The accuracy of bottom-up models depends very much on the availability of grass-root level consumption details, which is often not available for most developing countries.

Here, we employ the multiple regression time series modeling approach to model the electricity demand in Nigeria using previous data of 1970 to 2005 shown in Table 1 (CBN, 2005). The model uses electricity consumption and percentage connectivity as the regressors. The main consideration in the model formulation is that within the period of 1970 to somewhat late 80’s, electricity supply has been fairly constant without much interruption throughout the country. However during this period, only about 22% to 25% of the population has access to electricity. This informs the choice of the two variables in the model formulation.

METHODOLOGICAL FRAMEWORK

Moving Average Estimation: The historical data shows two visible trends; one starting from 1970 to 1993 and the other starting from 1994 to 2003, with the former being relatively coherent and the later being highly incoherent. A moving average (MA) smoothening process was employed for the coherent part, while the progressive reconstruction technique was employed for the incoherent part. In the MA process of Mohram and Rahman (1989), the current value of time series $X_t$ is expressed linearly in terms of current and previous values of a white noise series $\varepsilon_t, \varepsilon_{t-1}, \ldots \ldots$. This noise is constructed from the forecast errors or residuals of the recorded consumption data in Table 1.

For an MA of order q (ie. MA(q)), this model can be written as;

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \quad (1)$$

The MA trend for the 1970 to 1993 period evaluated using the model is as shown in Figure 1. This trend was therefore used to forecast (reconstruct) the time series for 1994 to 2005 trend.

![Figure 1: MA Smoothened Plot of the Historical Data.](image)

Multiple Linear Regression Model Estimation: The multiple linear regression trend models are generally expressed as;

$$Y_i = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_k X_k + \varepsilon \quad (2)$$

For a bi-variate least squares linear trend model, the parameters are defined as;

- $Y_i$: Annual Consumption in MW
- $X_i$: Affecting factors (Annual increment and Connectivity)
- $b_0$: Constant parameter of the model with respect to the affecting factors
- $b_1$: Coefficient of annual increase in consumption
- $b_2$: Coefficient of consumption due to increase in connectivity
### Table 1: Historical Electricity Consumption in Nigeria.

<table>
<thead>
<tr>
<th>Year</th>
<th>Industrial Consumption (MW)</th>
<th>Services Consumption (MW)</th>
<th>Residential Consumption (MW)</th>
<th>Total Consumption (MW)</th>
<th>% Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>91.4</td>
<td>0</td>
<td>53.9</td>
<td>145.3</td>
<td>22</td>
</tr>
<tr>
<td>1971</td>
<td>114.9</td>
<td>0</td>
<td>66.2</td>
<td>181.1</td>
<td>22</td>
</tr>
<tr>
<td>1972</td>
<td>138.2</td>
<td>0</td>
<td>72.9</td>
<td>211.1</td>
<td>22</td>
</tr>
<tr>
<td>1973</td>
<td>146.1</td>
<td>0</td>
<td>86.6</td>
<td>232.7</td>
<td>22</td>
</tr>
<tr>
<td>1974</td>
<td>163.2</td>
<td>0</td>
<td>103</td>
<td>266.2</td>
<td>25</td>
</tr>
<tr>
<td>1975</td>
<td>200.4</td>
<td>0</td>
<td>118.3</td>
<td>318.7</td>
<td>25</td>
</tr>
<tr>
<td>1976</td>
<td>214.6</td>
<td>0</td>
<td>155.2</td>
<td>369.8</td>
<td>25</td>
</tr>
<tr>
<td>1977</td>
<td>253</td>
<td>0</td>
<td>182.7</td>
<td>435.7</td>
<td>25</td>
</tr>
<tr>
<td>1978</td>
<td>157.7</td>
<td>93.5</td>
<td>253.2</td>
<td>504.4</td>
<td>25</td>
</tr>
<tr>
<td>1979</td>
<td>160.3</td>
<td>77.9</td>
<td>221.9</td>
<td>460.1</td>
<td>30</td>
</tr>
<tr>
<td>1980</td>
<td>199.7</td>
<td>94.1</td>
<td>243.1</td>
<td>536.9</td>
<td>30</td>
</tr>
<tr>
<td>1981</td>
<td>121</td>
<td>21.3</td>
<td>193.6</td>
<td>335.9</td>
<td>30</td>
</tr>
<tr>
<td>1982</td>
<td>262</td>
<td>79.1</td>
<td>344.5</td>
<td>685.6</td>
<td>30</td>
</tr>
<tr>
<td>1983</td>
<td>254.4</td>
<td>84.3</td>
<td>358</td>
<td>696.7</td>
<td>30</td>
</tr>
<tr>
<td>1984</td>
<td>217.2</td>
<td>81.7</td>
<td>326.6</td>
<td>625.5</td>
<td>30</td>
</tr>
<tr>
<td>1985</td>
<td>259.8</td>
<td>85.6</td>
<td>372</td>
<td>717.4</td>
<td>30</td>
</tr>
<tr>
<td>1986</td>
<td>280.5</td>
<td>84.7</td>
<td>476.6</td>
<td>841.8</td>
<td>30</td>
</tr>
<tr>
<td>1987</td>
<td>294.1</td>
<td>90.2</td>
<td>468.6</td>
<td>852.9</td>
<td>30</td>
</tr>
<tr>
<td>1988</td>
<td>291.1</td>
<td>118.6</td>
<td>443.8</td>
<td>863.5</td>
<td>30</td>
</tr>
<tr>
<td>1989</td>
<td>257.9</td>
<td>195.3</td>
<td>523.6</td>
<td>976.8</td>
<td>30</td>
</tr>
<tr>
<td>1990</td>
<td>230.1</td>
<td>217.6</td>
<td>450.8</td>
<td>898.5</td>
<td>33</td>
</tr>
<tr>
<td>1991</td>
<td>253.7</td>
<td>254.1</td>
<td>459.3</td>
<td>946.6</td>
<td>33</td>
</tr>
<tr>
<td>1992</td>
<td>245.3</td>
<td>266.1</td>
<td>481.6</td>
<td>993</td>
<td>33</td>
</tr>
<tr>
<td>1993</td>
<td>237.4</td>
<td>311.6</td>
<td>532.4</td>
<td>1141.4</td>
<td>33</td>
</tr>
<tr>
<td>1994</td>
<td>233.3</td>
<td>306.7</td>
<td>575</td>
<td>1115</td>
<td>33</td>
</tr>
<tr>
<td>1995</td>
<td>218.7</td>
<td>279.6</td>
<td>552.6</td>
<td>1050.9</td>
<td>35</td>
</tr>
<tr>
<td>1996</td>
<td>235.3</td>
<td>280</td>
<td>518</td>
<td>1033.3</td>
<td>35</td>
</tr>
<tr>
<td>1997</td>
<td>236.8</td>
<td>264.5</td>
<td>508.3</td>
<td>1009.6</td>
<td>35</td>
</tr>
<tr>
<td>1998</td>
<td>218.9</td>
<td>253.9</td>
<td>500</td>
<td>972.8</td>
<td>35</td>
</tr>
<tr>
<td>1999</td>
<td>191.8</td>
<td>236.8</td>
<td>455.1</td>
<td>883.7</td>
<td>35</td>
</tr>
<tr>
<td>2000</td>
<td>223.8</td>
<td>274.7</td>
<td>518.8</td>
<td>1017.3</td>
<td>38</td>
</tr>
<tr>
<td>2001</td>
<td>241.9</td>
<td>296.3</td>
<td>564.5</td>
<td>1104.7</td>
<td>38</td>
</tr>
<tr>
<td>2002</td>
<td>146.2</td>
<td>372.6</td>
<td>752.8</td>
<td>1271.6</td>
<td>38</td>
</tr>
<tr>
<td>2003</td>
<td>196</td>
<td>417.9</td>
<td>905.6</td>
<td>1515.5</td>
<td>40</td>
</tr>
<tr>
<td>2004</td>
<td>398</td>
<td>489.3</td>
<td>938.5</td>
<td>1825.8</td>
<td>40</td>
</tr>
<tr>
<td>2005</td>
<td>182.3</td>
<td>486.6</td>
<td>1194.3</td>
<td>1873.1</td>
<td>40</td>
</tr>
</tbody>
</table>

The parameters $b_i$ are estimated from the observations of $Y_i$ and $X_i$. Assuming that the error term $\varepsilon$ has a mean value equal to zero and a constant variance (i.e. 50% chance of being positive and negative, respectively), it could be omitted in calculating the parameters as in Bowerman, et al., 2005. Thus, for this case the model reduces to:

$$Y = b_0 + b_1X_1 + b_2X_2$$  \hspace{1cm} (3)

Then the least square estimates technique, which minimizes the sum of squared residuals (SSE) is employed to obtain the parameters [19]. Thus:

$$B = \begin{bmatrix} b_0 & b_1 & b_2 \end{bmatrix} = (X^T X)^{-1} X^T Y$$  \hspace{1cm} (4)

Where, $Y$ and $X$ are the following column vector and matrix:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$ and

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix}$$
Substituting the values of the actual (MA) consumption for y and the corresponding years for x, where n is 22, the parameters were evaluated.

Thus;

\[ b_0 = 280.6144; \]
\[ b_1 = 45.682; \]
\[ b_2 = -728.75 \]

Hence the bi-variate time series model for the electricity consumption from 1970 to 1993 is given by;

\[ y = 280.61 + 45.68x_1 - 728.75x_2 \quad (5) \]

This model in Equation (5) shows good fit to the historical data as shown in Figure 2.

**RESULTS AND ANALYSIS**

Simulating the reconstructed models in Equations (6) through (10), and comparing with the historical consumption yields the graph in Figure 3. The progressive improvements of the trend as year 2004 is approached can be clearly seen.

**Model Reconstruction:** From the historical data and our knowledge of electricity supply in Nigeria, it is clear that the supply is able to meet the demand from 1970 up to around 1993 when there was sudden dip in supply. Therefore, subsequent years met with gross undersupply and despite the limited connectivity to the national grid, electricity supply could not meet the available demand. Thus, the 1994 to 2005 historical consumption data cannot be used to predict future trend with acceptable accuracy. We therefore, employ a method of progressive reconstruction of 1994 to 2005 trend until a reasonably accurate prediction is obtained.

The first stage of reconstruction uses the trend obtained in Equation (5) to estimate the model parameters of the entire trend 1970 to 2005. The resulting model is given by;

\[ Y = 285.36 + 49.13 X_1 - 1084.89 X_2 \quad (6) \]

Similarly, the second stage reconstruction uses the trend obtained in Equation (6) to estimate the model parameters. We progressively used the trend of the reconstructed model to estimate the model parameters until an optimum model was arrived at after the 5th reconstruction. As no appreciable improvement in the trend was noticed after the 5th reconstructed model, this model constitutes the final model for our projection of 1994 to 2005. The reconstructed time series models from the second to the final model is as shown in Equation (7) through (10).

\[ Y = 358.503 + 52.501 X_1 - 1477.85 X_2 \quad (7) \]
\[ Y = 497.467 + 57.3773 X_1 - 2156.52 X_2 \quad (8) \]
\[ Y = 570.837 + 60.108 X_1 - 2524.26 X_2 \quad (9) \]
\[ Y = 605.258 + 61.970 X_1 - 2717.74 X_2 \quad (10) \]

Thus, our final model parameters were estimated using Equation (10) with trend values of 1985 to 2005, resulting in the model shown in Equation (11).

\[ Y = 149.149 + 58.279 X_1 - 1184.31 X_2 \quad (11) \]
Next, the final model depicted in Equation (11) was simulated and compare with the historical trend from 1985 to 2005 as shown in Figure 4. It could be seen that the predictions proved promising up to 2015, after which the trends begins to somewhat saturate and therefore not reliable.

The standard root mean square error (s) has been evaluated using the expression;

$$s = \sqrt{\frac{MSE}{n-(k+1)}} \quad (12)$$

where,

$$MSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = (4.4E-12)^2 \quad (13)$$

and;

$$y_i \quad : \text{actual} ; \quad \hat{y} \quad : \text{forecasted} ; \quad n = 21 ; \quad k = 2$$

Thus, $s = 2.444E-13$, which is very negligible!

Using this model, the forecasted value for year 2003 consumption was evaluated as 1656.9MW. This value, according to NEPA, (2002) is the average billed consumption, which does not take into account transmission to billing losses and other losses, which averaged to about 35% to 40%. Taking this into consideration therefore, brings the average demand to 2319.7MW. Using a load factor of 75%, which is the recorded average load factor for the period, the peak demand becomes 3092.8MW. This compares favorably with the recorded peak demand of 3171.3MW for the same year. However, it should be noted that this values does not take into account the suppressed consumption of the Country that is not supplied by the national grid, which has been estimated to 2051MW as in ECN and IAEA (2005).

With this suppressed demand, the peak demand would be 5222.3MW by 2003. The total installed capacity of the Country as at that year was about 5600MW, which means that the demand can only be met if all plants were operating at 95% capacity factor.

**CONCLUSION**

The modeling of the Country’s electricity demand using multiple regression time series modeling approach has been carried. The model uses the historical electricity consumption data of 1970 to 2005 and percentage connectivity to the national grid as the main regressors. The modeling process involved data smoothing using Moving Average (MA) technique for the coherent trend starting from 1970 to 1993, which was used to estimate the primary regression model. This model was subsequently used as a basis for reconstructing the incoherent trend observed from 1994 to 2003 by employing an innovative progressive reconstruction technique. This produces an optimal trend after five cycles of reconstruction. The final model parameters were
therefore, estimated using this trend starting from 1985 to 2005.

The forecast of the electricity consumption obtained from the model simulation were fairly accurate from 1985 up to 2015. Like most time series models, the forecast after 15 years proved to be inaccurate in view of the uncertainty of future trend and saturation of the series. This paper therefore, concludes that time series models alone may not be adequate for long term electricity demand forecast of this nature. An innovative approach to recommend would be to exploit a combination of the time series and the end-use modeling strategy, thereby complementing the draw back of each of the techniques.

REFERENCES


ABOUT THE AUTHORS

A.A. Mati, is a Research Fellow at the Centre for Energy Research and Training, Ahmadu Bello University, Zaria. He holds an M.Eng. degree in electrical engineering and is currently pursuing his Ph.D. research in long range electricity expansion modeling. His research interest is on electric power system planning and general energy systems modeling. He also lectures Energy Studies in the Department of Building Technology, Ahmadu Bello University, Zaria - Nigeria.

B.G. Bajoga, is a professor of electrical engineering at the Department of Electrical Engineering, Ahmadu Bello University, Zaria. He has supervised several dozen Ph.D. research works in and around the University. His broad area of interest include modeling and simulation, power line communication, mitigation of environmental effects of power system expansion and general systems engineering.

B. Jimoh, is a senior lecturer at the Department of Electrical Engineering of Ahmadu Bello University, Zaria. He holds a Ph.D. in electrical power systems. His research interests are in the areas of electric power system planning and general power system modeling.

B.A. Adegboyce is a senior lecturer at the Department of Electrical Engineering of Ahmadu Bello University, Zaria. He holds a Ph.D. in electrical power systems. His research interests are in the area of power line communication, reliability engineering and general system engineering.

D. D. Dajab is a senior lecturer at the Department of Electrical Engineering of Ahmadu Bello University, Zaria. He holds a Ph.D. in communications engineering. His research interests are in the areas of communication system planning and power line communication.

SUGGESTED CITATION