Scientific
Research

Isolated Area Load Forecasting Using Linear Regression Analysis: Practical Approach

Md. Apel Mahmud

School of Engineering & Information Technology, The University of New South Wales at Australian Defence Force Academy, Northcott Drive, Campbell, Australia
E-mail: *Md.Mahmud@student.adfa.edu.au*

Abstract

This paper presents an analysis to forecast the loads of an isolated area where the history of load is not available or the history may not represent the realistic demand of electricity. The analysis is done through linear regression and based on the identification of factors on which electrical load growth depends. To determine the identification factors, areas are selected whose histories of load growth rate known and the load growth deciding factors are similar to those of the isolated area. The proposed analysis is applied to an isolated area of Bangladesh, called Swandip where a past history of electrical load demand is not available and also there is no possibility of connecting the area with the main land grid system.

1. Introduction

In generation expansion planning as well as in distribution planning, load forecasting is an essential step. The importance of accurate forecast in planning is that it ensures the availability of supply of electricity as well as providing the means of avoiding over and under utilization of generating capacity and making the best possible use of capacity. Obviously, errors in forecasting can lead to bad planning which will be costly. Too high forecast lead to more plants than are required which will be an unnecessary capital expenditure. Too low forecast prevents optimum economic growth and lead to the installation of many costly and expensive-to-run generators. These costs will finally be borne by the consumers.

Different techniques have been implemented by researchers to solve the load forecasting task. However, two techniques are widely used, namely; regression and time series. The regression technique as proposed in [1,2] is based on is based on finding the functional relationship between weather components and the load demand. Therefore, the load is affected by the weather components that were used in the regression. The time series technique [3] is also a type of regression which takes a load pattern as a signal in a time series and forecasts the future load. In other words, the future load is only a function of the previous loads. Therefore, it is essential

to know the historic data of the area that will be forecasted. In [4], a load forecasting method used fuzzy logics is proposed where different factors of a day such as day's minimum temperature, day's maximum temperature, season, day capacity, rain, daylight intensity (cloudy) are considered as input. But practically, the amount of electricity consumed by a community depends on several factors such population, adult literacy, per capita income, and so on.

Linear regression technique is widely used in the area of load forecasting from the very early stage of power system planning and expansion. A regression analysis has been used in [5], to show the relationship between summer weather and summer loads.

In many countries, especially in developing ones, the major portion of the population live in remote and isolated areas. In term of electrical energy use, isolated areas are those areas which cannot be connected with the main electrical grid system due to economical or technical reasons. The economical and technical incapability of connecting isolated areas with the main grid may evolve from the distance between the grid and the isolated area or inconvenient connecting zones. The load forecasting of an isolated area is nicely described in [1-3]. But the main limitation of these works is that the load histories are assumed as known which is not practical. Recently, optimal scheduling of a renewable micro-grid in an iso-

lated load area using mixed-integer linear programming is proposed in [6], where different aspect of generation scheduling are considered to meet the demand of the community.

The aim of this paper is to forecast the load of an isolated area where there is no indication about the past history of loads. The analysis is done based on the linear regression method and other identification factors on which load growth depends. Identifications factors are calculated from the factors of an area whose load history is known as well as the selected area has the similar characteristic to that of an isolated area. This concept is also practically implemented on an isolated area of Bangladesh, called, Swandip whose load growth is similar to that of another area near author's own village.

The organization of the paper is as follows. Section 2 represents the methodology with all identification factors that are considered for load forecasting. Practical implementations of the theory and results are shown in Section 3. Finally, the paper is concluded with the new findings and further recommendation in Section 4.

2. Isolated Area Load Forecasting: Linear Regression Analysis

In this section, a linear regression analysis is adopted to develop a method of load forecasting. The method (LRA) starts with the identification of factors on which load growth depends. These factors may be different for different types of loads. Usually for an isolated area electric loads are domestic, commercial, industrial and irrigation.

2.1. Domestic Loads

The domestic load may be a function of population and standard of living of people. The variation of standard of living is caused by per capita income and adult literacy rate. All these factors are time varying quantities. The domestic load, L_D may then be expressed as,

$$L_D(t) = f_1\left(P(t), L_R(t), P_I(t)\right) \tag{1}$$

where, $P(t)$ = Population at time t, $L_R(t)$ = Adult literacy rate at time t, $P_I(t)$ = Per capita income at time t.

2.2. Industrial Loads

Industrial load may depend on the per capita income, inland communication in per unit area of total area, distance from the local town, literacy rate and agricultural land in per unit area of total area. This industrial load L_I can be calculated as,

$$L_I(t) = f_2\left(P_I(t), R_L(t), D_T(t), A_L(t)\right) \tag{2}$$

where, $R_L(t)$ = Inland communication length in per unit area at time t, $D_T(t)$ = Distance from local town, $A_L(t)$ = Agricultural land in percent of total area at time t.

In some isolated areas, instead of inland communication, communication across the sea may be the major type of communication. In that case the communication must be included the sea route length also.

2.3. Commercial Loads

The commercial load mainly depends on per capita income, inland communication in per unit area and distance from local town. The commercial load, L_C may then be expressed as,

$$L_C(t) = f_3\left(P_I(t), R_L(t), D_T(t)\right) \tag{3}$$

2.4. Irrigation Loads

The load for irrigation mainly depends on agricultural land in per unit area of total area and per capita income. This industrial load L_{IR} can be calculated as,

$$L_{IR}(t) = f_4\left(A_L(t), P_I(t)\right) \tag{4}$$

The total electrical load demand, $L(t)$ in an isolated area is the sum of the above four loads. That is,

$$L(t) = L_D(t) + L_I(t) + L_C(t) + L_{IR}(t) \tag{5}$$

Therefore, the load of an isolated area can be expressed as (6)错误！未定义书签。,

$$L(t) = f\left(P(t), L_R(t), P_I(t), R_L(t), A_L(t), D_T(t)\right) \tag{6}$$

Although Equation (6) expresses that the load is a function of six time dependent variables. However, all variables will not contribute equally to the generation of load. Let X_1, X_2, X_3, X_4, X_5, and X_6 represent the weighting factors by which each time varying factor $P(t)$, $L_R(t)$, $P_I(t)$, $R_L(t)$, $A_L(t)$, and $D_T(t)$ respectively contributes towards the load growth. The weighting factors, $[X]$ are also random in nature. They may vary with different areas. Now the load can be expressed as,

$$\left[L(t)\right] = \left[X\right] \times \begin{bmatrix} P(t) \\ L_R(t) \\ P_I(t) \\ R_L(t) \\ A_L(t) \\ D_T(t) \end{bmatrix} \tag{7}$$

From Equation (7), the weighted factors $[X]$ can be calculated as,

$$[X] = [L(t)]^{-1} \times \begin{bmatrix} P(t) \\ L_R(t) \\ P_I(t) \\ R_L(t) \\ A_L(t) \\ D_T(t) \end{bmatrix} \quad (8)$$

To calculate, the value of these weighted factor, the past history of the considered area need to be known. One more option is that one can consider an area whose behaviour is similar to that of the isolated area. This is what, is the main contribution of this work. This will be clarified through the practical implementation of the proposed method.

3. Practical Implementation and Results

For the purpose of load forecasting, the data is collected from the isolated area, Swandip and from a known area whose characteristics is similar to that of Swandip. In the district of Naogaon, Dhamuirhat Police Station, which is located on the bank of a river called, Jamuna Branch whose characteristics is very close to that of Swandip where most of the people lead their life by catching fish. The data are collected from the Naogaon Palli Bidhyut Samity which is an independent co-operative society of Rural Electrification Board (REB), Bangladesh and from the office of the Dhamuirhat Police Station under which the village is located. All other data are directly collected from the Swandip.

The data collected from the above mentioned sources can best be shown by **Table 1**.

Here, POP = Population (1000), L_R = Adult Literacy Rate (%), P_I = Per Capita Income (%), A_L = Agricultural Land (% of Total area), R_L = Road Length or Inland communication length in per unit area (km/km^2), D_T = Distance from Local Town (km), MD = Maximum Demand, September, AD = Average Demand, September, * to be calculated.

Now, the weighted factors need to be calculated for maximum demand as well as for average demand. The calculated weighted factors on different months of 2008 for average demand are shown in **Table 2**.

Again, the calculated weighted factors on different months of 2008 for maximum demand are shown in **Table 3**.

Tables 2 and **3** show that the weighting factors evaluated for different load deciding variables are not the same.

Table 2. Weighting factors X of load growth deciding variables for average demand.

Months of 2008	X_1	X_2	X_3	X_4	X_5	X_6
Jan	0.690	196.899	0.034	145.98	8.467	2.646
Feb	0.679	193.798	0.033	143.68	8.333	2.604
Mar	0.679	193.798	0.033	143.68	8.333	2.604
Apr	0.684	195.349	0.034	144.83	8.400	2.625
May	0.679	193.798	0.033	143.68	8.333	2.604
June	0.674	192.248	0.034	142.53	8.267	2.583
July	0.668	190.697	0.033	141.38	8.200	2.563
Aug	0.668	190.697	0.033	141.38	8.200	2.563
Sep	0.663	189.147	0.033	140.23	8.133	2.541
Oct	0.668	190.697	0.033	141.38	8.200	2.563
Nov	0.663	189.147	0.033	140.23	8.133	2.541
Dec	0.666	190.047	0.033	140.90	8.172	2.554

Table 3. Weighting factors X of load growth deciding variables for maximum demand.

Months of 2008	X_1	X_2	X_3	X_4	X_5	X_6
Jan	0.820	232.56	0.040	172.40	10.0	3.13
Feb	0.800	229.46	0.040	170.11	9.87	3.08
Mar	0.800	229.46	0.040	170.11	9.87	3.08
Apr	0.810	231.00	0.040	171.26	9.93	3.10
May	0.790	227.90	0.040	168.96	9.80	3.06
June	0.790	226.36	0.039	167.81	9.73	3.04
July	0.788	224.80	0.039	166.67	9.67	3.02
Aug	0.788	224.80	0.039	166.67	9.67	3.02
Sep	0.780	223.26	0.038	165.51	9.60	3.00
Oct	0.788	224.80	0.039	166.67	9.67	3.02
Nov	0.780	223.26	0.038	165.51	9.60	3.00
Dec	0.770	221.70	0.038	164.37	9.53	2.98

Table 1. Collected data.

Factors	Known Area	Unknown Area
POP(1000)	184	400
L_R (%)	64.5	58
P_I (TK)	3700	4000
A_L (%)	87	60
R_L (km/km^2)	15	10
D_T (km)	48	57
MD (MW)	150	*
AD (MW)	127	*

This clearly indicates that the load growths do not depend equally on all factors. This is, because some factors are very sensitive to the development of maximum demand or average demand while the others are not.

Again, by using the values of weighting factors, the value of maximum demand and average demand can be calculated. For this, the value of average demand and maximum demand for different months of a year is **Figures 1** and **2**, respectively.
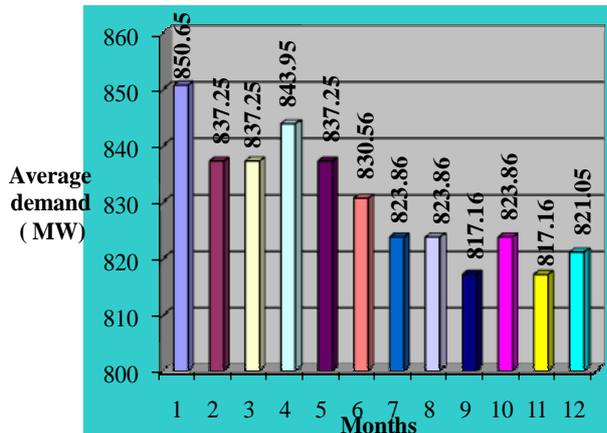

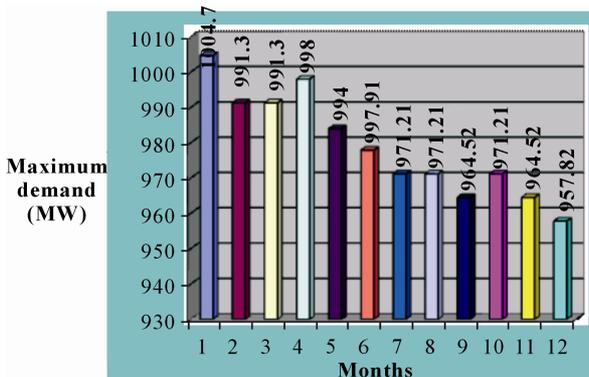**Figure 1. Forecasted average demand (MW) of an isolated area, Swandip.**

Figure 2. Forecasted maximum demand (MW) of an isolated area, Swandip.

Figures 1 and **2** show the different amount of forecasted demand of the considered area for different month of the year, 2008.

4. Conclusions

This work mainly focuses the load forecasting only for an isolated area Swandip in Bangladesh which helps an electric utility to make important decisions including decisions on purchasing and generating electric power, load switching, and infrastructure development. But this mathematical model is valid for all area to forecast electrical load. The further work will include the application of the proposed method to any area.

5. References

[1] D. Park, M. Al-Sharkawi, R. Marks, A. Atlas and M. Damborg, "Electric Load Forecasting Using an Artificial Neural Network," *IEEE Transactions on Power Systems*, Vol. 6, No. 2, 1991, pp. 442-449. doi:10.1109/59.76685

[2] A. D. Papalxopoulos and T. C. Hiterbeg, "A Regression-Based Approach to Short-Term Load Forecasting," *IEEE Transactions on Power Systems*, Vol. 5, No. 4, 1990, pp. 1535-1547. doi:10.1109/59.99410

[3] G. Gross and F. D. Galianan, "Short-Term Load Forecasting," *Proceedings of the IEEE*, Vol. 75, No. 12, 1987, pp. 1558-1572. doi:10.1109/PROC.1987.13927

[4] S. Sachdeva and C. M. Verma, "Load Forecasting Using Fuzzy Methods," *Proceeding of Joint International Conference on Power System Technology and IEEE Power India Conference*, New Delhi, 12-15 October 2008, pp. 1-4.

[5] G. T. Heinemann, D. A. Nordman and E. C. Plant, "The Relationship between Summer Weather and Summer Loads," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-85, No. 11, 1966, pp. 1144-1154. doi:10.1109/TPAS.1966.291535

[6] H. Morais, P. K. P. Faria, Z. A. Vale and H. M. Khodr, "Optimal Scheduling of a Renewable Micro-Grid in an Isolated Load Area Using Mixed-Integer Linear Programming," *Renewable Energy*, Vol. 35, No. 1, 2010, pp. 151-156. doi:10.1016/j.renene.2009.02.031