[Gupta, 4(4): April 2017]
DOI-10.5281/zenodo.545687

ISSN 2348 – 8034
Impact Factor- 4.022

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

MODELLING AND LOAD FORECASTING USING MULTIPLE LINEAR REGRESSION AND CURVE FITTING METHOD

Vikas Gupta*1 & Seema Pal2

*1PG Student, Department of Electrical Engineering, Jabalpur Engineering College, Jabalpur (M.P.), India

2Assistant Professor, Department of Electrical Engineering, Jabalpur Engineering College, Jabalpur (M.P.), India

ABSTRACT

This article provides a way of predicting one week ahead load forecasting using the data of Madhya Pradesh Poorva Kshetra Vidyut Vitaran Company Ltd. (MPPKVVCL). The data is used from Jan 2015 to Dec 2015 and the result for 1 January 2016 to 7 January 2016 is obtained . Method used for the load forecasting is Multiple Linear Regression Analysis and Curve Fitting .Various parameter such as temperature, humidity, dayofweek, weekofmonth, HourofDay etc are used to make the analysis more comparative and accurate. Model for both mentioned methods is developed in MATLAB. Result obtained will be compared between the two used methods. A large variation is seen on the festival and week days. 1 January 2015 historical data shows large variation and hence result shows large variation of 9.22% obtained from Curve Fitting method and variation of 15.53% is obtained from Multiple Linear regression for 1 Jan 2016 (New Year Day) and for rest of the days i.e. from 2 January 2016 to 7 January 2016 the error hardly goes beyond 2.65%.

Keywords: *Load forecasting, Multiple Linear Regression, Curve Fitting, MAPE (Mean Absolute Percentage Error), MATLAB*

I. INTRODUCTION

Load forecasting is one of the important tasks in Power system. Accurate load forecast of the power system leads to improved security and efficiency in electric power supply. The short term load forecasting with the data of MPPPKVVCL consist of the following two steps. First, the dependencies of the load demand on different parameter are studied. Next, Soft technique of load forecasting is used to forecast the one week ahead data. Different techniques of load forecasting have been studied [1]. This paper concentrates on the Regression analysis and curve fitting methods to forecast the load demand [2][3] . Short time load forecasting aimed at forecasting the load from one hour to seven days which is important for scheduling and operation of power systems. SLTF helps in energy management systems [4] which is required for the monitoring the electrical load system remotely. Load forecasting is not only important for the generation and transmission of power but also for the market operators so that economic use of the power can be done [5].

II. DATA ANALYSIS

For modelling the weather component, data obtained from the local weather forecasting centre Jabalpur, Madhya Pradesh region and Accuweather. Load data is collected from Madhya Pradesh Poorva Kshetra Vidyut Vitaran Company Ltd**.** (MPPKVVCL). The data is used from Jan 2015 to Dec 2015 and the result for 1 January 2016 to 7 January 2016 is obtained from MPPKVVCL. Due to wrong measurements and other human errors, some out-of-range values were observed in the historical load data as obtained from the weather department and MPPKVVCL . Corrections were made to such outlier values by replacing them with the average of both the preceding and succeeding values in the series. Principal Component Analysis (PCA) of the data was then carried out using MATLAB® functions “prepca” and “trapca”[6][7].

III. REGRESSION ANALYSIS

A regression with two or more explanatory variables is called a multiple regression [8]. The goal of multiple linear regressions (MLR) is to model the relationship between the explanatory and response variables. Simple Regression uses the modelling of the mean response as a straight line whereas in Multiple Regression analysis for load forecasting uses the technique of weighted least-square estimation. Based on this analysis, the statistical relationship between total load and weather conditions as well as the day type influences can be calculated. Also Multiple Linear Regression (MLR) is as modelled as a function of several explanatory variables.

Thus regression is the one of most widely used statistical techniques. The load is totally dependent on the temperature, humidity and day type parameters hence regression is used to obtain the relationship between load and these parameters.

$P_{forecasted} = P_{avg} + P_1$

MAPE (mean absolute percentage error) of power

MAPE = $\sum \{ (P_A - P_{forecasted})/P_A \} * 100$

Where

P_A =Actual power of forecasted power

$P_1 = \{R. \sigma p.(T_f - T_{avg}).(H_f - H_{avg}).(D_f - D_{avg})\}/\{ \sigma T.\sigma H.\sigma D \}$

$P_{forecasted}$ = forecasted day power

P_{avg} = Average power of the previous year data.

R=Coefficient of correlation of load power with temperature, humidity & daytype of the previous year data.

σp, σT, σH, σD = Standard deviation of power, temperature, humidity & daytype of previous year data

T_f, H_f, D_f = Forecasted day temperature, humidity and daytype

T_{avg}, H_{avg}, D_{avg} = Average temperature humidity and daytype [15].

IV. Curve Fitting

Curve fitting is the process of approximating or fitting of a series of data points with curves. Curve fitting is the process of constructing a curve, or mathematical function that has the best fit to a series of data points, possibly subject to constraints [10] .Curve fitting is used for performing exploratory data analysis , pre process and post process of data. A variety of post-processing methods for plotting, interpolation, and extrapolation; estimating confidence intervals; and calculating integrals and derivatives has been used for[11].Most popular curve fitting models are:

1. **Linear:** This curve is representing linear dependencies (include constant) between two variables.

$y = a.x + b$, where b is significant of value on y-axis for x=0 and a is gradient.

2. **Exponentials:** Usually is one-term and a two-term exponential model .Exponentials are often used when the rate of change of a quantity is proportional to the initial amount of the quantity. If the coefficient associated with e is negative, then y represents exponential decay. If the coefficient is positive, then y represents exponential growth.

3. **Fourier Series:** The Fourier series is a sum of sine and cosine functions that is used to describe a periodic signal. It is represented in either the trigonometric form or the exponential form. The toolbox provides the trigonometric

Fourier series form shown below

$$y = a_0 + \sum^n a_i.\cos(n.\ \omega.\ x) + b_i\ .\sin(n.\ \omega.\ x)$$

where a_0 is any "DC" offset of data (i=0), n is number of harmonics and ω represents frequency. Usually is speak about Fourier analysis and the fast Fourier transform[12].

Developed package uses the curve fitting (CF) technique to obtain a forecasting polynomial curve based on criterion of minimum quadratic error using the peaks of loads [12].Curve fit techniques for load forecasting is usually based on a multiple regression fit to past historical annual peak loads for each small area. The following polynomial is typical of such approaches [13] [14].

1(t) = a + a t-1 + a2t + a t2 + a4t3 = f k(t)

Where l.(t) estimate of load in cell j for year l'(t) = peak load for year t for cell j.
l1(h) = "horizon year" load level estimate for cell there are J small areas, or cells j - 1 J,there are T years load history present Year (last year of historical period. Usually, f (t) is determined by fitting to both recent load history and a horizon year load estimate.

V. RESULT AND DISCUSSION

Model is developed in MATLAB for both Multiple Linear Regression model and Curve Fitting Model for load forecasting. Both the soft techniques are applied day wise for a week. Table 1 shows the variation of actual load, Regression forecasted load and Curve Fitting forecasted load in Megawatt for 7 days. Here Day 1 stands for 1st January 2016(New Year) Friday; Day 2 stands for 2nd January 2016 Saturday and so on. Also Table 1 contains the comparison between the MAPE for Regression Model and Curve Fitting Model. A large error i.e. 15.53 MAPE in regression analysis and 9.538 MAPE in Curve fitting, are resulted in the forecasted value on Day 1 as the Day 1 is festive day (New Year) owing to 1st January. The overall error from the MLR is 3.706% and from that of the curve fitting is 3.249%. It is seen that the Curve Fitting on an average resulted in less error than that of the Regression analysis. Large variation on Day 1 in regression is due to the unpredictable variation in load historical data and less error in case of curve fitting as compared to MLR is because curve fitting involves interpolation where an exact fit to the historical data is done.

Table No 1: Shows comparative analysis of Multiple Linear Regression and Curve Fitting

Serial No	Actual Load (MW)	Regression(MW)	Curve Fitting (MW)	MAPE Regression(%)	MAPE Curve fitting(%)
Day 1	2564	2170	2322	15.53	9.538
Day2	2459	2475	2507	1.647	2.542
Day3	2463	2491	2516	1.707	2.365
Day4	2456	2457	2477	1.768	2.303
Day5	2522	2501	2520	1.994	2.454
Day6	2459	2477	2488	1.803	2.130
Day7	2565	2531	2541	1.488	1.409

 ISSN 2348 – 8034
Impact Factor- 4.022

Table No 2: Shows Average MAPE for Multiple Linear Regression and Curve Fitting

Serial No	MAPE Regression(%)	MAPE Curve fitting(%)
Day 1	15.53	9.538
Day2	1.647	2.542
Day3	1.707	2.365
Day4	1.768	2.303
Day5	1.994	2.454
Day6	1.803	2.130
Day7	1.488	1.409
AVERAGE MAPE	3.706	3.249

Furthermore, Graphs are obtained for the forecasted loads. Figure 1 shows the actual load and forecasted load obtained from MLR. Figure 2 shows the actual load and Forecasted load using Curve Fitting Method .As discussed above the huge variation in load on Day 1 can be seen in both the graphs. Figure 3 shows the comparison graph of Actual load, Forecasted load both by MLR and Curve Fitting Method for a comparative study.

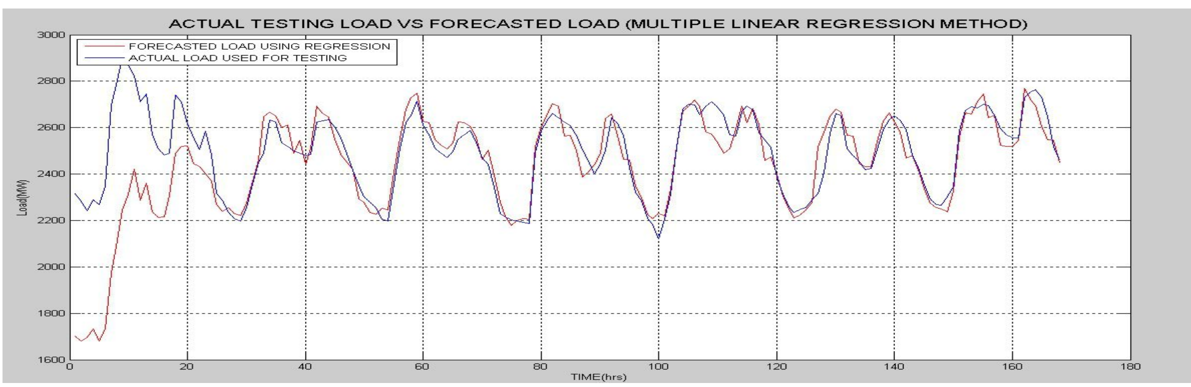

Figure 1: Variation of Actual Load and MLR Forecasted Load

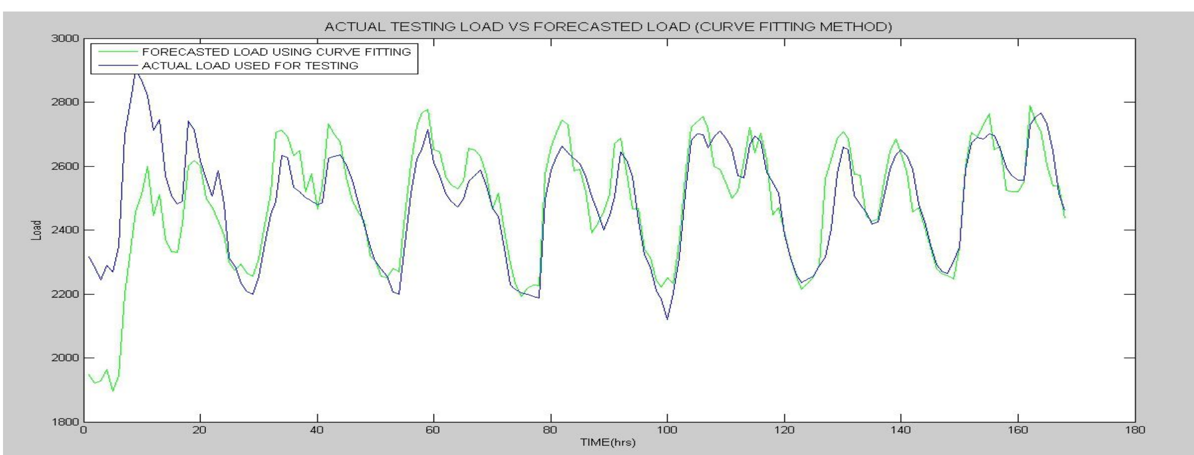

Figure 2: Variation of Actual Load and Curve Fitting Forecasted Load

ISSN 2348 – 8034
Impact Factor- 4.022

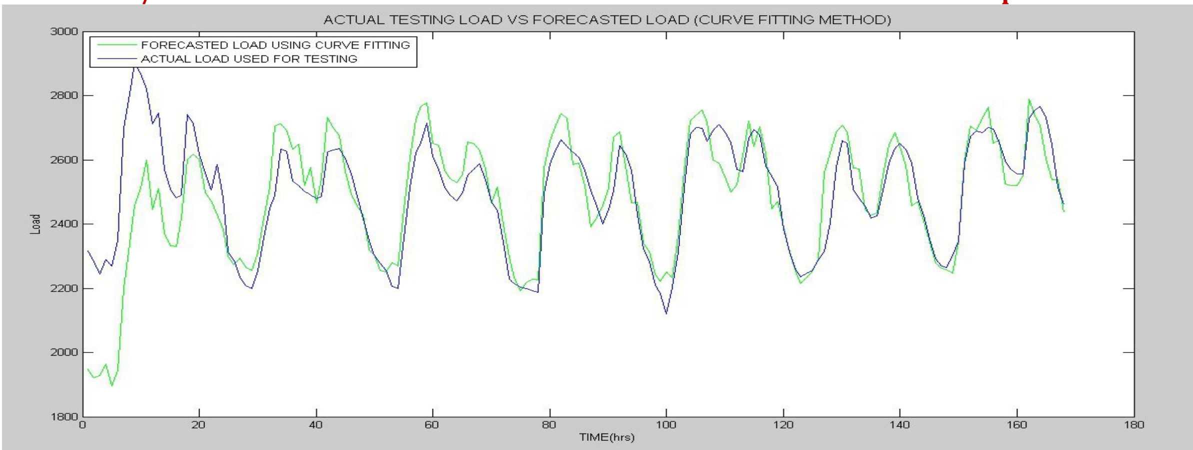

Figure 3: Variation of Actual Load and Both MLR & Curve Fitting Forecasted Load

VI. CONCLUSION

Regression Analysis and Curve Fitting fits good for the Short term Load forecasting. This paper has presented a study in the area of electric load forecasting with prediction times of the order of day and up to one week using soft techniques. STLF plays an important role in proper functioning of the power system operation as the principal driving element for all daily and Weekly operations scheduling. The result shows the soft technique based model for the dependency of a number of features such as the type of model, the data requirements, parameter identification and load forecasting needs. The paper also discussed the variation in the result of load forecasting with the impact of parameter and various practical consideration associated with the development of an STLF model and forecasting algorithm for use in a control centre environment.

Further work can be done with the inclusion of holiday and special days. As seen in the result the festive day, for example, on 1 January the forecasted result shows large error so some more optimization techniques such as Particle Swarm Optimization, Hybrid Genetic algorithm can be used in further study.

REFERENCES

1. *ISSN 2224-610X (Paper) ISSN 2225-0603 (Online)Vol.3, No.3, 2013-National Conference on Emerging Trends in Electrical, Instrumentation & Communication Engineering*
2. *lbrahim Moghram, Saifur Rahrnan, Member IEEE, IEEE Transactions on Power Systems, Vol. 4, No. 4, October 1989*
3. *H. Lee Willis, Member, IEEE R. W. Powell, Senior Member, IEEE D. L. Wall, Member, IEEE, IEEE Transactions on Power Apparatus and Systems, Vol. PAS-103, No. 5, May 1984*
4. *Ja.Sruthi, R.L.Helen Catherine, A Review on Electrical Load Forecasting in Energy Management, IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 3, March 2015.*
5. *Shu-Xia Yang; and N. Li, "Power Demand Forecast Based on Optimized Neural Networks by Improved Genetic Algorithm," Proceedings of IEEE International Conference on Machine Learning and Cybernetics,pp. 2877-2881, Aug. 200*
6. *s. K. Tong ,s. M. Shahidehpour z. Ouyang, a heuristic short-term unit commitment, 0885- 8950/91/m00-121wl.w q 1991 IEEE*
7. *Documentation " Principal component analysis of raw data" MATLAB 13.0*
8. *Applied Regression Analysis, 3rd Edition Norman R. Draper, Harry Smith ISBN: 978-0-471-17082-2*
9. *Book : G. Mbamalu, M. Elhawary, Load forecasting via no.1, pp.343-348, 1993*
10. *K. Sengupta , P. Burman , A curve fitting problem and its application in modeling objects in monocular image sequences, IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 24, Issue:5 May 2002)*
11. *Mark Campanelli, Benjamin Duck ,Keith Emery Quantifying and reducing curve-fitting uncertainty in Isc, Photovoltaic Specialist Conference (PVSC), 2015 IEEE 42nd*

12. *Arnika Soni ,A.K. Sharma, Load Prediction Using Curve Fitting Algorithm, International Journal of Engineering Science Invention ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726*
13. *H. Lee Willis, Member, IEEE R. W. Powell, Senior Member, IEEE D. L. Wall, Member, IEEE, IEEE Transactions on Power Apparatus and Systems, Vol. PAS-103, No. 5, May 1984*
14. *H.L.Willis,Member,IEEE, A. E. Schauer, J.E. D. Northcote Green, Member, IEEE, IEEE Transactions on Power Apparatus and Systems, Vol. PAS-102, No. 4, April*
15. *M.Babita Jain, Manoj kumar Nigam, Prem Chand TIwary "Curve Fitting And Regression Line Method Based Seasonal, Short Term Load Forecasting " 2012 World Congress on Information Technology, 2012.*