A Parameter based New Approach for Short Term Load Forecasting using Curve-fitting and Regression line method

M. Babita Jain is Professor and Dean at Rungta Engineering College, Raipur, CG, India

Manoj Kumar Nigam, Reader in Electrical Engineering Department at RITEE Raipur (C.G.), India

Prem Chand Tiwari, is M.E. Student at RITEE Raipur (C.G.) India,

Abstract—Short term load forecasting in this paper is done by considering the sensibility of the network load to the temperature, humidity, day type parameters (THD) and previous load and also ensuring that forecasting the load with these parameters can best be done by the regression line and curve fitting methods. The analysis of the load data recognize that the load pattern is not only dependent on temperature but also dependent on humidity and day type. A new norm has been developed using the regression line concept with inclusion of special constants which hold the effect of the history data and THD parameters on the load forecast and it is used for the STLF of the test dataset of the used data set. A unique norm with a, b, c and d constants based on the history data has been proposed for the STLF using the concept of curve fitting technique. The algorithms implementing this forecasting technique have been programmed using MATLAB. The input data of each day average power, average temperature, average humidity and day type are used for prediction of power, in the case of the regression line method and the forecast previous month data and the similar month data of the previous year is used for the curve fitting method. The simulation results show the robustness and suitability of the proposed norm for the STLF as the forecasting accuracies are very good and less than 3% for almost all the day types and all the seasons. Results also indicate the curve fitting method out passes the regression technique w.r.t forecasting accuracy and hence it the best suitable method for accurate short term loads forecasting.

Index Terms—Short term load forecasting, THD (temperature, humidity, day type), curve fitting, regression line.

I. INTRODUCTION

Load forecasting is an important component for power system energy management system. Precise load forecasting helps the electric utility to make unit commitment decisions, reduce spinning reserve capacity and schedule device maintenance plan properly. Besides playing a key role in reducing the generation cost, it is also essential to the reliability of power systems. The system operators use the load forecasting result as a basis of off-line network analysis to determine if the system might be vulnerable. If so, corrective actions should be prepared, such as load shedding, power purchases and bringing peaking units on line.

With the recent trend of deregulation of electricity markets, STLF has gained more importance and greater challenges. In the market environment, precise forecasting is the basis of electrical energy trade and spot price establishment for the system to gain the minimum electricity purchasing cost. In the real-time dispatch operation, forecasting error causes more purchasing electricity cost or breaking-contract penalty cost to keep the electricity supply and consumption balance. There are also some modifications of STLF models due to this implementation of the electricity market. Weather is defined as the atmospheric condition existing over a short period in a particular location. It is often difficult to predict and it can vary significantly even over a short period. Climate also varies with time: seasonally, annually and on a decade’s basis [1]. The relationship between demand and temperature is non-linear with the demand increasing for both low and high temperature [2]. The range of the possible approaches to the forecast is to take a microscopic view of the problem and try to model the future load as a reflection of previous [3]. In the case of large variation in the temperature compared to that of the previous year, the load also changes accordingly. In such cases there would be the shortage of similar days’ data and the task of the forecasting load is very difficult [4].

II. DATA ANALYSIS

A. LOAD CURVES

For the analysis and implementation of load forecasting, data is taken from EUNITE network that was provided to participants for a competition many years ago (see acknowledgement). In the data analysis part we are going to analyze load variation with respect to day type, weather condition such as seasonal variation of load with temperature and humidity. Analyzing the monthly and yearly load curves given in Fig.1 and Fig.2 and also load variation with respect to temperature and humidity given in Fig.3 and Fig.4 the following observations are made:

- The load curve patterns of two consecutive years is similar
- The load curves of similar months of two consecutive years is also similar
- The load curves are having different pattern in weekdays and weekend days in the month.
- The load curves on the weekends are similar.
- Taking in consideration the above observations the days of the week are classified based on the following categories:
a. Normal week days (Tuesday - Friday)
b. Monday
c. Sunday
d. Saturday

Monday is accounted to be different to weekdays so as to take care for the difference in the load because of the previous day to be weekend.

B. Variation of Load with Temperature

The variation of the temperature variable results in a significant variation in the load. Fig 3 shows a plot between the maximum temperatures versus average demand. In Fig. 3 the dots represent the actual values and the solid line is the best fitted curve. The graph shows a positive correlation between the load and temperature i.e. demand increases as the temperature increases.

C. Variation of Load with Humidity

Fig 4 shows the plot between the average humidity versus average demand. From the graph it can be seen that there exists a positive correlation between load and humidity i.e. demand increases as the humidity increases.

D. Autocorrelation of Load

It is seen from the plots that the load pattern of the present year is similar to the load pattern of previous year and also the load curve of a given month is similar to the load curve of the previous years’ same month. Hence it can be considered that the load of similar month of previous year can greatly help in load forecasting along with the THD parameters.

III. SHORT TERM LOAD FORECASTING USING REGRESSION LINE

This section deals with the details of the concept and implementation of regression line method and its implementation to STLF. The relation between load and the THD parameters is defined as they have direct impact on the load as seen in the earlier section. Regression is the study of the relationship among variables, a principle purpose of which is to predict, or estimate the value of the one variable from known or assumed values of other variables related to it. For electric load forecasting regression methods are usually used to model the relationship of load consumption and other factors such as weather and day type [4]-[10].

As it is clear from the data analysis that 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. The method is basically divided into three parts as follows:

- Load variation with respect to temperature
Load variation with respect to temperature and humidity
Load variation with respect to temperature, humidity and day type.

Since it is clear from data analysis that all the three THD parameters affect the load forecast the regression line method is used to find the relation between load and all these three parameters.

Regression line (THD): The regression line equation for load forecast dependent on the temperature, humidity and day type parameters is given as follows:

\[ P_{\text{forecasted}} = P_{\text{avg}} + \frac{b}{(H_T - H_{\text{avg}}) \times (D_T - D_{\text{avg}}))} (\sigma_T, \sigma_H, \sigma_D) \]

Where \( P_{\text{forecasted}} \) = forecasted day power
\[ P_{\text{avg}} = \text{Previous year average power} \]
\[ R = \text{coefficient of relation load power with temperature and humidity of previous year} \]
\[ T_{\text{avg}}, H_{\text{avg}}, D_{\text{avg}} = \text{average temperature, humidity and day type previous year} \]
\[ \sigma_P, \sigma_T, \sigma_H, \sigma_D = \text{standard deviation of power, temperature, humidity and day type of previous year} \]

Algorithm:

Step: 1. Calculation of previous year's forecasted power, temperature, humidity, and day type.

Step: 2. Calculation of coefficient of relation R.

\[ R = \frac{(c_1 - c_2) / (c_3 + c_4 + c_5 + c_6)} {\text{where } c_1, c_2, c_3, c_4, c_5, c_6} \]

Step: 3. Taking the value of the forecasted day temperature, humidity, and day type.

Step: 4. Writing the relation between load power and the parameters.

\[ P_{\text{forecasted}} = P_{\text{avg}} + \frac{b}{(H_T - H_{\text{avg}}) \times (D_T - D_{\text{avg}}))} (\sigma_T, \sigma_H, \sigma_D) \]

Step: 5. Calculating MAPE (mean absolute percentage error) of power.

\[ MAPE = \sum_{i=1}^{N} \left| \frac{P_i - P_{\text{predicted}}}{P_i} \right| \times 100 \]

Step: 6. For next day load forecasting go step 1 to step 5.

Step: 7. Result analysis.

Step: 8. End

From the above algorithm we are easily calculating the load power. This method using the previous year load data including the temperature, humidity and the day type finding the variation of the load with respect to the parameters which is mention above.

IV. SHORT TERM LOAD FORECASTING USING CURVE FITTING

The methodology that is developed for the short term load forecasting of load using the curve fitting method would mainly focus on the variation of power with the three main parameters we have already mentioned i.e. Temperature, Humidity and the particular Day Type[11]-[14].

The first two parameters, quite evidently come under the weather changing phenomenon, but considering the time dependent variation of the load, the data that is available could not only be classified into a particular day type but it also follows a similar month pattern which implies that, for example, if we take the data of January of one particular year, there are steep chances that it is almost identical to the one we had in the same month of its previous year under a similar working environment. As from previous method similarly we are using the all three factors for the forecasting the power. It is explain in the algorithm.

Algorithm (THD):

Step 1. Write the equations between power and its parameters using curve fitting.

\[ P = a + b \times T + c \times H + d \times D \]

Step 2. Using previous year data of similar month calculate coefficient of a, b, c & d.

\[ \begin{align*}
\sum P & = a \times N + b \times \sum T + c \times \sum H + d \times \sum D \\
\sum PT & = a \times \sum T + b \times \sum T^2 + c \times \sum HT + d \times \sum DT \\
\sum PTH & = a \times \sum TH + b \times \sum TH^2 + c \times \sum H^2T + d \times \sum DTH \\
\sum PTHD & = a \times \sum THD + b \times \sum TH^2D + c \times \sum H^2TD + d \times \sum D^2TH
\end{align*} \]

Step 2: Using previous year data of similar month calculate coefficient of a, b, c & d.

Assuming the coefficients

\[ a = a_1 \]
\[ b = a_2 \]
\[ c = a_3 \]
\[ d = a_4 \]
Step 3: Coefficient substitution in the equation (3)
\[ P_{\text{forecast}} = a_1 + a_2 \cdot T_{\text{forecast}} + a_3 \cdot H_{\text{forecast}} + a_4 \cdot D_{\text{forecast}} \]

Step 4: Calculating the forecasting power of each day in present month

Step 5: Calculating MAPE (mean absolute percentage error) of power.
\[ MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{P_i - P_{\text{forecast}}}{P_i} \right| \times 100 \]

Step 6: For next month forecasting of power repeat steps 2 to 5.

Step 7: Result analysis

Step 8: End.

From the above algorithm, the forecasting of the load power with respect to the parameters are easily calculated. This method is very simple and having less complex as compared to the regression line method.

V. SUMMARY OF THE METHDOLOGY

In this part, we are discussing the summery of the methods which are used in this paper, regression line and the curve fitting are the same steps to follow and hence we draw the methodology in the same flow diagram.

VI. RESULT ANALYSIS

The result analysis of the simulation performed on the power variation with respect to its corresponding parameters clearly suggests the dependency of the load power on the main three factors that were considered in this paper, namely Temperature, Humidity, and the Day Type. To understand this, we have tabulated the individual variation of power with temperature taken separately, power with temperature and humidity taken together and finally the power variation along with day type also. Based on the tabular data of the readings, we also plotted the individual graphs each showing the variation of power with respect to its parameters separately and also when considered together. In this analysis, part separate method results are tabular form shown and the graphical analysis is given below.

Table: I
Parameters of the regression line and curve-fitting algorithms

<table>
<thead>
<tr>
<th>PARAMETERS</th>
<th>VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>P_AVERAGE</td>
<td>13053.34</td>
</tr>
<tr>
<td>R</td>
<td>0.0014</td>
</tr>
<tr>
<td>T_AVERAGE</td>
<td>48.6190</td>
</tr>
<tr>
<td>H_AVERAGE</td>
<td>38.2240</td>
</tr>
<tr>
<td>D_AVERAGE</td>
<td>3.1448</td>
</tr>
<tr>
<td>( \sigma_p )</td>
<td>13608.54</td>
</tr>
<tr>
<td>( \sigma_T )</td>
<td>17.4668</td>
</tr>
<tr>
<td>( \sigma_H )</td>
<td>19.3137</td>
</tr>
<tr>
<td>( \sigma_D )</td>
<td>1.227</td>
</tr>
<tr>
<td>a</td>
<td>11279.72</td>
</tr>
<tr>
<td>b</td>
<td>3.71783</td>
</tr>
<tr>
<td>c</td>
<td>-29.0868</td>
</tr>
<tr>
<td>d</td>
<td>760.9079</td>
</tr>
</tbody>
</table>

Table: II
Comparative STLF of Saturday, Sunday, Monday and Tuesday Load

<table>
<thead>
<tr>
<th>SN</th>
<th>DATE</th>
<th>Actual Load</th>
<th>REGRESSION Forecast Load</th>
<th>CURVEFITTING Forecast Load</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/11/97</td>
<td>11674.79167</td>
<td>12250.324</td>
<td>11720.18706</td>
</tr>
<tr>
<td>2</td>
<td>2/11/97</td>
<td>11021.91667</td>
<td>11561.25783</td>
<td>10929.62926</td>
</tr>
<tr>
<td>3</td>
<td>3/11/97</td>
<td>12754.66667</td>
<td>13080.58321</td>
<td>12495.5655</td>
</tr>
<tr>
<td>4</td>
<td>4/11/97</td>
<td>13030.375</td>
<td>13123.7481</td>
<td>13190.33612</td>
</tr>
</tbody>
</table>
Table: III
Comparative STLF MAPE Saturday, Sunday, Monday and Tuesday Load

<table>
<thead>
<tr>
<th>DATE</th>
<th>REGRESSION LINE</th>
<th>CURVEFITTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/11/97</td>
<td>4.9297</td>
<td>0.38832548</td>
</tr>
<tr>
<td>2/11/97</td>
<td>4.893351854</td>
<td>0.837308176</td>
</tr>
<tr>
<td>3/11/97</td>
<td>2.55272922</td>
<td>2.031422499</td>
</tr>
<tr>
<td>4/11/97</td>
<td>0.716580306</td>
<td>1.227601821</td>
</tr>
</tbody>
</table>

Fig. 6 Curve of the load of the Saturday, Sunday, Monday and Tuesday Load

Table: IV
Comparative STLF of Sunday, Monday, Tuesday, Wednesday, Thursday, Friday and Saturday forecasted Load

<table>
<thead>
<tr>
<th>DATE</th>
<th>ACTUAL LOAD</th>
<th>REGRESSION LINE</th>
<th>CURVEFITTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/11/97</td>
<td>11021.91667</td>
<td>11561.25783</td>
<td>10929.62926</td>
</tr>
<tr>
<td>3/11/97</td>
<td>12754.66667</td>
<td>13080.58321</td>
<td>12495.5655</td>
</tr>
<tr>
<td>4/11/97</td>
<td>13030.375</td>
<td>13123.7481</td>
<td>13190.33612</td>
</tr>
<tr>
<td>5/11/97</td>
<td>13077.625</td>
<td>13134.87086</td>
<td>13355.7966</td>
</tr>
<tr>
<td>6/11/97</td>
<td>13264.5</td>
<td>13152.66666</td>
<td>13402.35266</td>
</tr>
<tr>
<td>7/11/97</td>
<td>13129.41667</td>
<td>13079.46963</td>
<td>13343.56988</td>
</tr>
<tr>
<td>8/11/97</td>
<td>12054.45833</td>
<td>13149.77202</td>
<td>12057.19216</td>
</tr>
</tbody>
</table>

Fig. 7 MAPE of the load MAPE Saturday, Sunday, Monday and Tuesday Load

Table V
Comparative STLF MAPE of Sunday, Monday, Tuesday, Wednesday, Thursday, Friday and Saturday forecasted Load

<table>
<thead>
<tr>
<th>DATE</th>
<th>REGRESSION LINE</th>
<th>CURVEFITTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/11/97</td>
<td>4.893351854</td>
<td>0.837308176</td>
</tr>
<tr>
<td>3/11/97</td>
<td>2.55272922</td>
<td>2.031422499</td>
</tr>
<tr>
<td>4/11/97</td>
<td>0.716580306</td>
<td>1.227601821</td>
</tr>
<tr>
<td>5/11/97</td>
<td>0.437738972</td>
<td>2.127080421</td>
</tr>
<tr>
<td>6/11/97</td>
<td>0.843095478</td>
<td>1.040767906</td>
</tr>
<tr>
<td>7/11/97</td>
<td>0.380420849</td>
<td>1.63109466</td>
</tr>
<tr>
<td>8/11/97</td>
<td>9.086378284</td>
<td>0.022678997</td>
</tr>
</tbody>
</table>

From the above analysis it is clear that both the regression line and the curve fitting methods are quite suitable for the short term load forecasting giving very good forecasting accuracies with MAPE’s of most of the cases quite less than 3%. Results also indicate that the curve fitting method is better compared to the regression line method for short term load forecasting. The curve fitting method is simple and robust in comparison to the regression line method. The method yields very good results for days of all types and all around the year as the weekly result values clearly indicate.

VII. CONCLUSION

Accurate load forecasting is very important for electric utilities in a competitive environment created by the electric industry deregulation. In this paper, we have presented the regression line and curve fitting methods for short term load forecasting. The following are the conclusions derived from the proposed methods:

- In this paper, regression and curve fitting methods have strongly proved the impact of THD parameters in the STLF using the relation between input and output variable through the systematic rule. In the regression method input data is previous year’s data and for the curve fitting method data input is previous year’s similar month’s data.
- The simulation results have shown that the both the proposed methodologies are quite good and completely suitable for STLF of all types of the days and for all
months round the year giving the MAPE for most of the cases quite less than 3%.

- Results also indicate that the curve fitting method has an edge over the regression line method. The efficiency of the curve fitting method is due to the consideration of the previous years’ similar month as the training data set for the calculation of the constants. The data analysis part of Section II indicates the strong impact of the previous year similar month load on the present month. This impact of previous year similar month goes very well with the curve fitting method.

VIII. ACKNOWLEDGMENT

The authors gratefully acknowledge Mr. R. Venkatendra for providing the EUNITE Network load forecasting data, which has been used for simulation study in this paper. Author’s information about the source of data is based on the information provided by Mr. R. Venkatendra.

IX. REFERENCES


X. BIOGRAPHIES

M Babita Jain is PhD student at the International Institute of Information Technology, Hyderabad, India since 2006. The areas of interest include IT Applications to Power Systems, Load forecasting and Multi Agent Systems. She has actively involved in establishing IEEE student branch and Women in engineering affinity group activities.

Manoj Kumar Nigam is Reader in Electrical Engineering Department at RITEE Raipur (C.G.) India. He is completed his ME in Industrial systems & drives from MITS Gwalior, his area of interest is electrical drives & power electronics.

Prem chand Tiwari is the ME student at RITEE Raipur (C.G.) India since 2010. The area of interest is electronic devices and circuits and power electronics.