A Methodology for Short Term Load Forecasting Using Fuzzy Logic and Similarity

E. Srinivas and Amit Jain, Member, IEEE

Abstract—This paper presents a methodology for the short term load forecasting problem using fuzzy logic approach. To obtain the next-day load forecast, fuzzy logic is used to modify the load curves on selected similar days. A new Euclidean norm with weight factors is proposed for the selection of similar days. The proposed methodology is illustrated through the simulation results on a typical data set.

Index Terms—Euclidean norm, Fuzzy logic approach, Short term load forecasting, Similar day method.

I. INTRODUCTION

LOAD forecasting has been an integral part in the efficient planning, operation and maintenance of a power system. Short term load forecasting is necessary for the control and scheduling operations of a power system and also acts as inputs to the power analysis functions such as load flow and contingency analysis [1]. Owing to this importance, various methods have been reported, that includes linear regression, exponential smoothing, stochastic process, ARMA models, and data mining models [2]-[7]. Of late, artificial neural networks have been widely employed for load forecasting. However, there exist large forecast errors using ANN when there are rapid fluctuations in load and temperatures [8]-[9]. In such cases, forecasting methods using fuzzy logic approach have been employed. In this paper, we propose an approach for short term load forecasting problem, using fuzzy logic approach. This approach has an advantage of dealing with the nonlinear parts of the forecasted load curves, and also has the ability to deal with the abrupt change in the weather variables such as temperature etc.

In this method, we select similar days from the previous days to the forecast day using Euclidean norm with weather variables [10]. There may be a substantial discrepancy between the load on the forecast day and that on similar days, even though the selected days are very similar to the forecast day with regard to weather and day type. Therefore, the selected similar days cannot be averaged to obtain the load forecast. To avoid this problem, the evaluation of similarity between the load on the forecast day and that on similar days is done using fuzzy logic. Many methods have been reported for the load forecast using fuzzy logic [12]–[14]. This approach evaluates the similarity using the information about the previous forecast day and previous similar days. The evaluated value represents a correction factor for the load curve on a similar day to the shape of that on the forecast day. After calculating the correction factor of load curves on similar days, the forecast load is obtained by averaging the corrected load curves on similar days. The approach suitability is verified by applying it to a typical data set.

This paper contributes to the short term load forecasting, as it shows how the forecast can be include the effect of weather variables, temperature as well as humidity, using a new Euclidean norm for the selection of the similar days and fuzzy logic approach. The paper is organized as follows: section II deals with the data analysis; section III gives the overview of the proposed forecasting method, discussing the selection of similar days using a new Euclidean norm and the fuzzy inference system is presented; section IV presents the simulation results of the proposed forecasting methodology followed by conclusions in section V.

II. VARIABLES IMPACTING THE LOAD PATTERN

The analysis on the monthly load and weather data helps in understanding the variables, which may affect load forecasting. The data analysis is carried out on data containing hourly values of load, temperature, and humidity of 7 months. In the analysis phase, the load curves are drawn and the relationship between the load and weather variables is established [11]. Also, the week and the day of the week impact on the load is obtained.

A. Load Curves

The load curve for the month of May is shown in Fig. 1. The observations from the load curves are as follows:
1. There exists weekly seasonality but the value of load scales up and down.
2. The load curves on week days are mostly similar.
3. The load curves on the weekends are similar.
4. Days are classified based on their load into 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.

\[ \text{Load Curve for the Month of May} \]

\[ \text{Fig. 1 Load Curve for the month of May} \]

**B. Variation of Load with Temperature**

Temperature is the most important weather variable that affects the load. The deviation of the temperature variable from a normal value results in a significant variation in the load. Figures 2 and 3 shows the relationship between the temperature and load. Fig 2 shows a plot between the average temperatures versus maximum demand. Fig 3 shows the plot between the average temperatures versus average demand. In the plots the dot represents the actual values and the solid line is the best fitted curve. The graphs show a positive correlation between the load and temperature for the forecast month of July i.e. demand increases as the temperature increases.

\[ \text{Plot of Average Load vs Average Temperature} \]

\[ \text{Fig 3 Plot between Average Load versus Average Temperature} \]

\[ \text{Plot of Maximum Load vs Average Temperature} \]

\[ \text{Fig 2 Plot between Maximum Load versus Average Temperature} \]

\[ \text{Plot of Maximum Load vs Average Humidity} \]

\[ \text{Fig 4 Plot between Maximum Load versus Average Humidity} \]

**C. Variation of load with Humidity**

Another weather variable that affects the load level is humidity. To study the effect of this particular weather variable on load we plot the maximum demand versus average humidity and the average demand versus average humidity graphs as shown. Fig 4 shows the plot between the average humidity versus maximum demand. Fig 5 shows the plot between the average humidity versus average demand. From the graphs it can be seen that there exists a positive correlation between load and humidity for the forecast month of July i.e. demand increases as the humidity increases.
D. Autocorrelation of load

It is known that the load at a given hour is dependent not only on the load on the previous hour but also on the load at the same hour of the previous day. Hence, it is assumed that the load curve is more or less similar to the load curve on the previous day.

III. LOAD FORECASTING USING FUZZY LOGIC

A. Similar Day Selection

In this paper, Euclidean norm with weight factors is used to evaluate the similarity between the forecast day and the searched previous days. Euclidean norm makes us understand the similarity by using the expression based on the concept of norm. Decrease in the Euclidean norm results in the better evaluation of the similar days i.e., smaller the Euclidean norm the more similar are the days to the forecast day. In general, the Euclidean norm using maximum and minimum temperatures along with the day type variable is used for the evaluation of the similar days. But, the norm using maximum and minimum temperatures is not efficient for the selection of the similar days because humidity is also an important weather variable as also shown in section II C.

In the present work, we have proposed a new Euclidean norm to account for the humidity. The new Euclidean norm uses maximum temperature, average humidity and day type with weight factors to evaluate the similarity of the searched previous days. The expression for the new Euclidean norm is as follows:

\[
EN = \sqrt{w_1 (\Delta T_{\text{max}})^2} + w_2 (\Delta H_{\text{avg}})^2 + w_3 (\Delta D)^2
\]

\[
\Delta T_{\text{max}} = T_{\text{max}} - T_{\text{max}}^p
\]

\[
\Delta H_{\text{avg}} = H_{\text{avg}} - H_{\text{avg}}^p
\]

\[
\Delta D = D - D^p
\]

Where, \( T_{\text{max}} \) and \( H_{\text{avg}} \) are the forecast day maximum temperature and average humidity respectively. Also, \( T_{\text{max}}^p \) and \( H_{\text{avg}}^p \) are the maximum temperature and average humidity of the searched previous days and \( w_1, w_2, w_3 \) are the weight factors determined by least squares method based on the regression model constructed using historical data [2]. The similar days are selected from the previous 30 days of the forecast day. The data selection is limited to account for the seasonality of the data. The day types considered for the methodology are 4(Tuesday-Friday), 3(Monday), 2(Saturday), 1(Sunday);

B. Fuzzy Inference System

The load forecasting at any given hour not only depends on the load at the previous hour but also on the load at the given hour on the previous day. Also, the Euclidean norm alone is not sufficient for the load forecast as the selected similar days for the forecast day have considerably large mean absolute percentage error (MAPE). Assuming same trends of relationships between the previous forecast day and previous similar days as that of the forecast day and its similar days, the similar days can thus be evaluated by analyzing the previous forecast day and its previous similar days.

The fuzzy inference system is used to evaluate the similarity between the previous forecast days and previous similar days resulting in correction factors, used to correct the similar days of the forecast day to obtain the load forecast. To evaluate this degree of similarity, three fuzzy input variables for the fuzzy inference system are defined [10].

\[
E_{L}^k = L_p - L_{ps}^k
\]

\[
E_{T}^k = T_p - T_{ps}^k
\]

\[
E_{H}^k = H_p - H_{ps}^k
\]

Where, \( L_p \) and \( L_{ps} \) are the average load of the previous forecast day and the previous kth similar day, \( T_p, T_{ps}, H_p, H_{ps} \) show the value corresponding to temperature and humidity respectively. \( E_{L}, E_{T}, E_{H} \) take three fuzzy set values; Low (L), Medium (M), High (H). The membership functions of the input variables and output variable are as shown in Figs 6 - 7.

The fuzzy rules for the inference system for the given fuzzy variables are based on the generalized knowledge of the effect of each variable on the load curve [11]. If the membership of \( E_L \) is \( \mu_{E_L} \), that of \( E_T \) is \( \mu_{E_T} \) and that of \( E_H \) is \( \mu_{E_H} \), the firing strength, \( \mu \), of the premise is calculated based on the min operator. The firing strength of each rule is calculated as follows:

\[
\mu_i = \min(\mu_{E_{L,i}}, \mu_{E_{T,i}}, \mu_{E_{H,i}})
\]
The membership function of an inferred fuzzy output variable is calculated using a fuzzy centroid defuzzification scheme to translate fuzzy output statements into a crisp output value, \( W_k \).

\[
W_k = \frac{\sum_{i=1}^{27} \alpha_i \mu_i^k}{\sum_{i=1}^{27} \mu_i^k}
\]

The output value is expressed by \( W_k \) which is the correction factor for the load curve on the \( k \)th similar day to the shape on the forecast day. \( W_k \) is applied to each similar day and corrects the load curve on similar days. The forecast next day load curve \( L(t) \) is then given by averaging the corrected loads on similar days.

\[
L(t) = \frac{1}{N} \sum_{k=1}^{N} (1 + W_k) L_s^k(t)
\]

Where \( L_s^k(t) \) is the power load at \( t \)’o clock on the \( k \)th corrected similar day, \( N \) is the number of similar days and \( t \) is hourly time from 1 to 24.

IV. SIMULATION RESULTS

The performance of the method for the short term load forecast is tested by using the 7 months data, from January to July of a particular data set used. The method has been simulated using the fuzzy logic toolbox available in MATLAB. Load forecasting is done for the month of July. Hence, the data of the month June has been used for the selection of similar days. The number of similar days used for the forecasting is five.

The parameters of the fuzzy membership functions are determined through the simulation of the load curve forecasting in the previous month to the forecast day. The parameters of the membership functions for the input and output variables for the next-day load curve forecasting for the month of July are as follows:

| Table 2: Parameters of the membership functions of the input variable |
|-----------------------------|-----------------------------|-----------------------------|
| (a1,a2) | (a3,a4) | (a5,a6) |
| (-1000,1000) | (-20,20) | (-20,20) |

| Table 3: Parameters of the membership functions of the output variable |
|-----------------------------|-----------------------------|-----------------------------|
| (b1,b2,b3) | (b4,b5,b6) | (b7,b8,b9) |
| (-0.3,-0.25,-0.2) | (-0.15,-0.10,-0.05) | (-0.05,0,0.05) |

The forecasted results of 4 representative days in a week are presented. These days represents four categories of classified days of week in the present methodology namely Saturday, Sunday, Monday, and Tuesday.

The humidity variations for the days to be forecasted i.e. 23rd – 27th July are as shown in Fig. 8.
The forecasted results for July 24 to July 27 are given Fig 9-12.

The forecast results deviation from the actual values are represented in the form of MAPE. Mean Absolute Percentage Error (MAPE) is defined as:

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{P_A - P_F^i}{P_A} \right| \times 100$$

$P_A$, $P_F$ are the actual and forecast values of the load. $N$ is the number of the hours of the day i.e. $N = 1, 2, \ldots, 24$

With the proposed method the MAPE error for the considered days, for which forecasted results are shown in Fig. 9-12, are calculated and these are as follows:

<table>
<thead>
<tr>
<th>Day</th>
<th>MAPE Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 July (Saturday)</td>
<td>1.07</td>
</tr>
<tr>
<td>25 July (Sunday)</td>
<td>1.37</td>
</tr>
<tr>
<td>26 July (Monday)</td>
<td>2.88</td>
</tr>
<tr>
<td>27 July (Tuesday)</td>
<td>2.10</td>
</tr>
</tbody>
</table>
V. Conclusion

This paper presented a short term load forecasting methodology using fuzzy logic, which takes into account the effect of humidity as well as temperature on load. In the method, fuzzy logic is used to correct the similar day load curves of the forecast day to obtain the load forecast. Also, a new Euclidean norm with weight factors is proposed, which is used for the selection of similar days. Fuzzy logic is used to evaluate the correction factor of the selected similar days to the forecast day using the information of the previous forecast day and the previous similar days.

To verify the forecasting ability of the proposed methodology, we performed load forecasting for the month of July in a data set of 7 months and results for four representative days of a week in the month of July are given. The results obtained from the simulation show that the proposed forecasting methodology, which proposes the use of weather variables i.e. temperature as well as humidity, gives load forecasting results with considerable accuracy, within the range of 3% MAPE. Therefore, the proposed methodology will be helpful in using more weather variables, which will certainly be better than using only temperature as the weather variable affecting the load, in short term load forecasting and hopefully provide intellectual stimulus to research community to do further research in this direction.

REFERENCES


VI. Biographies

E. Srinivas is a Master’s Student in Power Systems Research Center, International Institute of Information Technology, Hyderabad, India. He received his B. Tech degree from Sri Indu College of Engineering and Technology, Ibrahimpatan. His areas of interest include applications of computational intelligence systems and operations planning of power systems.

Amit Jain graduated from KNIT, India in Electrical Engineering. He completed his masters and Ph.D. from Indian Institute of Technology, New Delhi, India. He was working in Alstom on the power SCADA systems. He was working in Korea in 2002 as a Post-doctoral researcher in the Brain Korea 21 project team of Chungbuk National University. He was Post Doctoral Fellow of the Japan Society for the Promotion of Science (JSPS) at Tohoku University, Sendai, Japan. He also worked as a Post Doctoral Research Associate at Tohoku University, Sendai, Japan. Currently he is an Assistant Professor in IIIT, Hyderabad, India. His fields of research interest are power system real time monitoring and control, artificial intelligence applications, power system economics and electricity markets, renewable energy, reliability analysis, GIS applications, parallel processing and nanotechnology.