Short Term Load Forecasting using Fuzzy Adaptive Inference and Similarity

by

Amit Jain, E. Srinivas, Rasmimayee Rauta

in

Nature Biologically Inspired Computing
(NaBIC 2009)

Report No: IIIT/TR/2009/210

Centre for Power Systems
International Institute of Information Technology
Hyderabad - 500 032, INDIA
December 2009
Abstract—The main objective of short term load forecasting (STLF) is to provide load predictions for generation scheduling, economic load dispatch and security assessment at any time. Thus, STLF is needed to supply necessary information for the system management of day-to-day operations and unit commitment. This paper presents a forecasting method based on similar day approach in conjunction with fuzzy rule-based logic. To obtain the next-day load forecast, fuzzy logic is used to modify the load curves on selected similar days. A Euclidean norm considering weather variables such as ‘temperature’ and ‘humidity’ with weight factors is used for the selection of similar days. The effectiveness of the proposed approach is demonstrated on a typical load and weather data.

Keywords—Euclidean norm, fuzzy logic, optimization, short term load forecasting, similar days.

I. INTRODUCTION

Forecasting is an integral part of electric power system operations as it is the primary prerequisite for achieving the goal of optimal planning and operation of power systems. If the duration of the forecast varies from few hours to weeks, it is called as short term load forecasting. Short term load forecasting is necessary for the control and scheduling operations of a power system and also acts as input to the power analysis functions such as load flow and contingency analysis [1][2]. The load dispatcher at main dispatch center must anticipate the load pattern well in advance so as to have sufficient generation to meet the load requirements. Overestimation may cause the startup of too many generating units and lead to an unnecessary increase in the reserve and the operating costs. Underestimation of the load forecast results in failure to provide the required spinning and standby reserve and stability to the system, which may lead into collapse of the power system network. Load forecast errors can yield suboptimal unit commitment decisions.

With the recent trend of deregulation of electricity industry, STLF has gained more importance and greater challenges. In the real-time dispatch operation, forecasting error causes more electricity purchasing cost or breaking-contract penalty cost to keep the electricity supply and consumption balance. Hence accurate forecasting of the load is an essential element in power system.

The electric load is sensitive to several weather factors like temperature, humidity, and chill factor etc. Thus, the main difficulty in modeling load demand is to come up with a model that comprises only dominant factors and minimize the forecasting error. Due to the increasing pressure on the need for accurate load forecasts, numerous statistical and artificial intelligence methods have been proposed for the short-term load forecasting problem. These methods include time series, regression, stochastic process, ARMA, expert system and data mining based approaches [3]-[10]. Of late, artificial neural networks have been widely employed for load forecasting. The main reason of ANN becoming so popular lies in its ability to learn complex and nonlinear relationships that are difficult to model with conventional techniques. However, there exist large forecasting errors using ANN when there are rapid fluctuations in load and temperatures [11]-[12]. To cope with such uncertainties in the load forecasting problem, forecasting methods using fuzzy logic approach have been employed. In recent years, fuzzy set theory based approach has emerged as a complement tool to mathematical approach for solving power system problems [13]-[15]. Many methods have been reported for the load forecasting using fuzzy logic [16]-[19]. Also, hybrid methods of neural networks and fuzzy logic are also reported [20]-[23]. Expert knowledge in the form of fuzzy rule-based logic increases model reliability in cases of unusual severe weather conditions and varied holiday activities.

In this paper, we propose a fuzzy adaptive inference system for short term load forecasting problem based on the similar day approach. Recently, several methods based on similarity have been reported for the purpose of electric load forecasting. According to which, the load curve is forecasted by using the information of the days being similar to weather condition of the forecast day. In the method, we select similar days from the previous days to the forecast day using Euclidean norm with weather variables [16]. 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 the adaptive fuzzy inference system. The adaptive fuzzy inference approach updates the fuzzy membership parameters using a heuristic brute force optimization technique. 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. The fuzzy approach is not new to short term load forecasting but in this paper we have tried to utilize fuzzy inference with similar day approach where fuzzy parameters have been optimized and these adaptive fuzzy parameters helps in improving the quality of forecasted results. The approach suitability is verified by applying it to a typical data set.

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 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 AFFECTING THE LOAD PATTERN

The data analysis is carried out on data containing hourly values of load, temperature, and humidity of a typical data set used in this paper. In the analysis phase, the load curves are drawn and the relationship between the load and weather variables is established [6]. Also, the week and the day of the week effect on the load is obtained.

A. Load Curves

The seasonal effect, which dictates the relationship between weather and the load, will change with the change in the season and is accounted by limiting the previous data used for forecasting to 30 days.

The effect of the day of the week on the load is then accounted by the use of different day-types for days of the week i.e., weekday (Tuesday – Friday): 4, Monday: 3, Saturday: 2, Sunday: 1. this day of the week variable eliminates the need to model the holidays separately.

B. Variation of Load with Temperature

One of the most important weather variables that affect the load is temperature. Fig. 1 and Fig. 2 show the relationship between the load and temperature. 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.

C. Variation of Load with Humidity

Humidity is another weather variable that affects the load level. 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 in Fig. 3, which shows the plot between the average load versus average humidity.

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 somewhat similar to the load curve on the previous day.
III. SHORT TERM LOAD FORECASTING USING FUZZY INFERENCE

A. Similar Day Selection

To evaluate the similarity between the forecast day and the searched previous days, Euclidean norm with weight factors is used in the present work. 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.

We have used a Euclidean norm with the weather variables to account for the variations in temperatures as well as humidity. The 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 Euclidean norm is as follows:

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

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 [3].

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 current 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 [13].

\[ E_L^k = L_p - L_{ps} \]  
\[ E_T^k = T_p - T_{ps} \]  
\[ E_H^k = H_p - H_{ps} \]  

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 4 - 5.

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 [6]. If the membership of \( E_i \) is \( \mu_i \), that of \( E_T \) is \( \mu_T \) and that of \( E_H \) is \( \mu_H \), the firing strength, \( \mu \), of the premise is calculated based on the min operator. The firing strength of each rule is calculated as follows:

\[ \mu_i = \min(\mu_{E_L}, \mu_{E_T}, \mu_{E_H}) \]  

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 kth 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.

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

Where \( L^t_k(t) \) is the power load at tth clock on the kth corrected similar day, \( N \) is the number of similar days and t is hourly time from 1 to 24.
IV. SIMULATION RESULTS

A data set with 7 months data, from January to July is used to test the performance of the method for the short term load forecast presented in this paper. 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 obtained parameters are then optimized using a heuristic brute force technique using the RMS MAPE error in the objective function for the optimization procedure. The optimized parameters of the membership functions for the input and output variables for the next-day load curve forecasting for the month of July are given in Table I and Table II.

The forecasted results of 4 representative days in a week are presented in Fig 6-9. These days represent four categories of classified days of week in the present methodology namely Saturday, Sunday, Monday, and Friday. Forecast 1 results are with optimization of parameters and Forecast 2 results are without optimization.

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

\[
MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{P_i - \hat{P}_i}{P_i} \right| \times 100
\]

where \(P_i\) and \(\hat{P}_i\) 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. 6-9, are calculated and these are given in Table III.

<table>
<thead>
<tr>
<th>Day of Forecast</th>
<th>MAPE Error before Optimization</th>
<th>MAPE Error after Optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 July (Saturday)</td>
<td>1.07</td>
<td>0.99</td>
</tr>
<tr>
<td>25 July (Sunday)</td>
<td>1.37</td>
<td>1.54</td>
</tr>
<tr>
<td>26 July (Monday)</td>
<td>2.88</td>
<td>0.61</td>
</tr>
<tr>
<td>30 July (Friday)</td>
<td>3.34</td>
<td>0.48</td>
</tr>
</tbody>
</table>

**V. SUMMARY AND CONCLUSIONS**

A short term load forecasting methodology using fuzzy adaptive inference and similarity, which takes into account the effect of humidity as well as temperature on load, is presented in this paper. In this method, fuzzy inference is used to correct the similar day load curves of the forecast day to obtain the load forecast. Further the parameters for fuzzy inference are optimized which further improves the forecasting results. Also, a Euclidean norm with weight factors is proposed, which is used for the selection of similar days. Fuzzy adaptive inference 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.

We performed short term load forecasting for the month of July in a data set of 7 months to validate the forecasting methodology proposed in this paper 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 forecasting methodology with adaptive fuzzy parameters and similarity, which makes use of weather variables i.e. temperature as well as humidity, gives good short term load forecasting results and they are within the range of 2% MAPE. In summary, we have proposed a methodology for short term load forecasting with fuzzy adaptive inference and similarity which uses temperature as well as humidity as weather variables, which also provides a way to include more weather variables for further research in short term load forecasting process.

REFERENCES