Performance Comparison of Short Term Load Forecasting Techniques

Kumar Reddy Cheepati¹* and T. Nageswara Prasad²

*Asst. Professor, Dept. of E.E.E, S.V.E.C, Tirupathi, A.P, India
Professor & Head, Dept. of E.E.E, S.V.E.C, Tirupathi, A.P, India
Kumarreddy98@gmail.com, np_thunga@yahoo.com

Abstract

Load forecasting plays a major role in planning and operation of a power system. Many techniques are available in the literature among these neural networks, linear multiple regression, curve fitting and averaging models are the most popular because these models give accurate solutions with very less tolerable Least Mean Absolute Percent Error (MAPE). In this paper a comparative study was made between these forecasting models and it was found that when compared to the four independent models, the averaging model i.e. combination of Curve Fitting, Regression Trees & Neural Network gives less MAPE. MATLAB programming results validates that averaging model gives better performance than individual models.

Keywords: Artificial Neural Network (ANN), Mean Absolute Percent Error (MAPE), Linear Multiple Regression, Regression Trees, averaging, Load Forecasting

1. Introduction

Electric load demand is a function of weather variables and human social activities, industrial activities as well as community developmental level and is difficult to forecast [1]. Statistical techniques and Expert system techniques have failed to adequately address this issue. The daily operation and planning activities of an electric utility requires the prediction of electricity demand of its customers. In general, the required load forecasts can be categorized into short-term, mid-term, and long-term forecasts [2]. The short-term forecasts refer to hourly prediction of the load for a lead time ranging from one hour to several days out. The mid-term forecasts can either be hourly or peak load forecasts for a forecast horizon of one to several months ahead which refer to Scheduling of fuel purchases, Load flow studies or contingency analysis, and planning for energy, while the long-term forecasts refer to forecasts made for one to several years in the future. The quality of short-term hourly load forecasts has a significant impact on the economic operation of the electric utility since decisions such as economic scheduling of generating capacity, transactions such as ATC (Available Transmission Capacity) are based on these forecasts and they have significant economic consequences [4].

The need for accurate load forecasting will increase in the future because of the dramatic changes occurring in the structure of the utility industry due to deregulation and competition. This environment compels the utilities to operate at the highest possible efficiency, which, as indicated above, requires accurate load forecasts. Studies reported indicate that there is a 3-7% increase of electric load per year for many years. The increase of load depends on population growth, local area development, industrial expansion etc. The taxonomy of load forecasting can be considered as Spatial forecasting & Temporal forecasting. Forecasting future load distribution in a particular region, such as a country, a state, or the whole country is called spatial forecasting. Temporal forecasting is deals with forecasting load for a specific supplier or collection of consumers in future hours, days, months, or even years. The temporal forecasting can be
broadly divided into 4 types – long term, medium term, short term and very short term [3]. Typically the short term load forecast covers a period of one week. The forecast calculates the estimated load for each hour of the day, the daily peak load and the daily/weekly required energy generation [3]. The forecasted data is used for:

- Unit commitment (Selection of generators in operation, start-up/shut down of generation to minimize operation cost) [4].
- Hydro scheduling to optimize water release from reservoirs.
- Hydro-Thermal co-ordination to determine the least cost operation mode (optimum mix)
- Interchange scheduling & energy purchase
- Transmission line loading
- Power system security assessment (load flow & transient stability studies)

Forecasting can be defined as the mechanism that uses the historical data to determine the direction of future trends. In the field of power system, the electrical load forecasting is used to predict the future power demand of consumers. Future power demand is estimated on the basis of the historical load data. The forecast model processes the exogenous relation of the data provided and consequently anticipates the future load demand. Figure 1 illustrates the block diagram of processes involved on how to design a load forecast model that predicts the future load demand.

The previous load is divided into Monday, Tuesday, Wednesday, Thursday, Friday, Saturday and Sunday loads. The load has two distinct patterns i.e. week day and week end patterns. The week day includes Tuesday, Wednesday, Thursday, Friday and the week end includes Saturday, Sunday and Monday.

1.1. Techniques for Load Forecasting

Various techniques are available; Intelligence techniques include Expert Systems, Fuzzy inference and Fuzzy – neural models. For decades the problem of improving the accuracy of load forecasts has been an important topic of research. Different types of load forecasting methodologies have their own advantages. Load forecasting techniques such as regression analysis, statistical methods, artificial neural networks, genetic algorithm, fuzzy logic etc., Jenkins transfer functions [11], ARMAX models [12], optimization techniques [13], non-parametric regression [14], structural models and curve – fitting procedures [15]. The most popular ones are linear regression ones models [6-8], threshold auto-regressive models[9], methods based on Kalman-filtering[10], Box –. Artificial Time series models (load is modelled as a function of its past observed values) [5], multiplicative auto-regressive models [6], dynamic linear or non-linear weights have been extensively discussed in the literature.

Various techniques are available; Intelligence techniques include Expert Systems, Fuzzy inference and Fuzzy – neural models. For decades the problem of improving the accuracy of load forecasts has been an important topic of research. Different types of load forecasting methodologies have their own advantages. Load forecasting techniques such as regression analysis, statistical methods, artificial neural networks, genetic algorithm, fuzzy logic etc., Jenkins transfer functions [11], ARMAX models [12], optimization techniques [13], non-parametric regression [14], structural models and curve – fitting procedures [15]. The most popular ones are linear regression ones models [6-8], threshold auto-regressive models[9], methods based on Kalman-filtering[10], Box –. Artificial Time series models (load is modelled as a function of its past observed values) [5], multiplicative auto-regressive models [6], dynamic linear or non-linear weights have been extensively discussed in the literature.
2. Procedure

![Figure 1. Block Diagram](image)

In this paper four Models have been designed and discussed for forecasting with each of its advantages and disadvantages, along with the combined model (averaging model) that gave best results. Those models are:

A) Neural Network model (NN): using NN which is a part of artificial intelligence (AI) to forecast is very good due to the flexibility provided by NN and their ability to do nonlinear relationships. The whole model is a matrix of weights that will be trained by the historical loads and then using these Weights we can predict any output for any given input.

Total weighted input, using the formula in Eqn. (1) below

\[ X_j = \sum X_i W_{ij} \]  \hspace{1cm} (1)

where \( X_i \) is the activity level of the \( W_{ij} \)th unit in the previous layer and \( W_{ij} \) is the weight of the connection between the \( i \)th and the \( j \)th unit. Next, the unit calculates the activity \( Y_j \) using some function of the total weighted input. Typically the sigmoid function in as given Eqn. (2) is given by.

\[ Y_j = \frac{1}{1 + e^{-x_j}} \] \hspace{1cm} (2)

Once the activities of all output units are determined, the network computes the error \( E \), which is defined by the expression as in Eqn. (3).

\[ E = \frac{1}{2} \sum (Y_j - d_j)^2 \] \hspace{1cm} (3)

Where \( Y_j \) the activity level of the \( j \)th unit in the top is layer and \( d_j \) is the desired output of the \( j \)th unit [40].
B) Regression Trees Model: Regression is one of the most widely used statistical techniques. For electric load forecasting regression methods are usually used to model the relationship of load consumption and other factors such as weather, day type, and customer class [36]. Engle et al., [37] presented several regression models for the next day peak forecasting. Their models incorporate deterministic influences such as holidays, stochastic influences such as average loads, and exogenous influences such as weather. References [19, 31, 16, 3] describe other applications of regression models to load forecasting. Regression trees specify the form of the relationship between predictors and response; we first build the tree and then fit the leaves values to the input predictors like in Neural Networks.

C) Multiple Linear Regression Model: This model is good to estimate the general trends of the data however it cannot predict nonlinearities, it basically produces coefficient vector that can be multiplied with any input to provide an estimated output.

A linear regression model that contains more than one predictor variable is called a multiple linear regression model. The following model is a multiple linear regression model with two predictor variables, \( x_1 \) and \( x_2 \) as shown in eqn.4

\[
Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon
\]  

(4)

The model is linear because, it is linear in the parameters \( \beta_0 \), \( \beta_1 \), and \( \beta_2 \) is linear. The model describes a plane in the three-dimensional space of \( Y \), \( x_1 \), and \( x_2 \). The parameter \( \beta_0 \) is the intercept of this plane. Parameters \( \beta_1 \) and \( \beta_2 \) are referred to as partial regression coefficients. Parameter \( \beta_1 \) represents the change in the mean response corresponding to a unit change when it is held constant. Parameter \( \beta_2 \) represents the change in the mean response corresponding to a unit change in \( x_2 \) when \( x_1 \) is held constant. Consider the following example of a multiple linear regression model with two predictor variables, \( x_1 \) and \( x_2 \) as shown in Eqn.(5)

\[
Y = 30 + 5 x_1 + 7 x_2 + \epsilon
\]  

(5)

This regression model is a first order multiple linear regression Fmodels. This is because the maximum power of the variables in the model is 1. The true regression model is usually never known (and therefore the values of the random error terms corresponding to observed data points remain unknown). However, the regression model can be estimated by calculating the parameters of the model for an observed data set.

Consider a multiple linear regression model with \( k \) predictor variables as shown in Eqn.(6)

\[
Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \epsilon
\]  

(6)

Let each of the \( k \) predictor variables, \( x_1, x_2, \ldots, x_k \), have \( n \) levels. Then \( x_{ij} \) represents the \( i^{th} \) level of the \( j^{th} \) predictor variable \( x_j \). For example, \( x_{51} \) represents the fifth level of the first predictor variable \( x_1 \), while \( x_{10} \) represents the first level of the ninth predictor variable, \( x_9 \). Observations, \( y_1, y_2, \ldots, y_n \), recorded for each of these \( n \) levels can be expressed in the following way as shown in Eqn.(7),(8),(9),(10).

\[
y_{ij} = \beta_0 + \beta_{i,1} x_{i1} + \beta_{i,2} x_{i2} + \ldots + \beta_{i,k} x_{ik} + \epsilon_i
\]  

(7)

\[
y_{i2} = \beta_0 + \beta_{i,1} x_{i1} + \beta_{i,2} x_{i2} + \ldots + \beta_{i,k} x_{ik} + \epsilon_i
\]  

(8)

\[
y_{i3} = \beta_0 + \beta_{i,1} x_{i1} + \beta_{i,2} x_{i2} + \ldots + \beta_{i,k} x_{ik} + \epsilon_i
\]  

(9)
\[ y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \ldots + \beta_s x_{ns} + \epsilon_n \]  

(10)

The system of \( n \) equations shown previously can be represented in matrix notation as follows in Eqn.(11),(12),(13),(14)(15)

\[ y = X\beta + \epsilon \]  

(11)

Where

\[ y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \]

\[ X = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1u} \\ 1 & x_{21} & x_{22} & \ldots & x_{2u} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{nu} \end{bmatrix} \]  

(13)

\[ \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_s \end{bmatrix} \]  

(14)

\[ \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_s \end{bmatrix} \]  

(15)

The matrix \( X \) is referred to as the design matrix. It contains information about the levels of the predictor variables at which the observations are obtained. The vector \( \beta \) contains all the regression coefficients. To obtain the regression model, \( \beta \) should be known. \( \beta \) is estimated using least square estimates. The following equation is used in Eqn.(16)

\[ \hat{\beta} = (X'X)^{-1}X'y \]  

(16)
Where ‘\( ^\top \)’ represents the transpose of the matrix while ‘\(-1\)’ represents the matrix inverse. Knowing the estimates, \( \hat{\beta} \), the multiple linear regression model can now be estimated as in Eqn.17

\[
y = X \hat{\beta}
\]  

(17)

The estimated regression model is also referred to as the fitted model. The observations, \( y_i \), may be different from the fitted values \( \hat{y} \) obtained from this model. The difference between these two values is the residual, \( e_i \). The vector of residuals, \( e \), is obtained as given in Eqn.(18)

\[
e = y - \hat{y}
\]  

(18)

The fitted model can also be written as follows, given in Eqn.(19),(20),(21) & (22)

\[
\hat{\beta} = (X^\top X)^{-1} X^\top y 
\]  

(19)

\[
y = X \hat{\beta}
\]  

(20)

\[
y = X (X^\top X)^{-1} X^\top y = Hy 
\]  

(21)

Where \( H = X (X^\top X)^{-1} X^\top \). The matrix, \( H \), is referred to as the hat matrix. It transforms the vector of the observed response values, \( y \), to the vector of fitted values, \( \hat{y} \).

**D) Curve Fitting:** Curve Fitting Prediction software is a collection of graphical user interfaces (GUIs), prediction functions for curve and surface fitting that operate in the MATLAB technical computing environment [1]. A typical flowchart for curve fitting prediction is given in Figure 2.

![Figure 2. A Typical Flowchart for Curve Fitting Prediction Methods](image-url)
Basic principle of curve fitting is using the historical data to find coefficients of an equation; the equations could be multiple sin waves, Fourier or polynomial equation, etc., Matlab Curve fitting GUI is good way to find the potential of this Model. In this case the equations are Fourier of the form as given in Eqn.(13)

\[ F(x) = a_0 + a_1 \cos(xw) + b_1 \sin(xw) + \ldots + a_n \cos(nxw) + b_n \sin(nxw) \]  

(13)

Where \( a_n, b_n \) are the coefficients which are used to find “fit” function in MATLAB.

E) Averaging Model:

In this model the advantage of ANN in nonlinearities and curve fitting in general trends of data and regression trees in good responses and average those models to get a model that has results better than the other three models individually.

3. Discussion

Designing Load forecasting model is best done by visualizing historical data coupled with understanding the trends of the curves and what parameters can affect the Load at certain hour. From the MATLAB programming following plots of one year data have been observed, the observations are:

A variation of Load based on the month is given in Figure 3.

![Figure 3. Variation of Load w.r.t Month](image)

The variation of load w.r.t to the day of the month \( i.e. \) Monday-Sunday is given in Figure 4.
The variation of load based on the day of the week is given in Figure 5.

A general trend throughout the hours of the day where the peak Load happens in the afternoon is given by Figure 6.
Variations of load through all days of the year are given by Figure 7.

4. Results

A. Neural Network Model:

This model has provided the best forecasting and best fitting of historical data as it can be seen from Absolute Mean Percent Error (MAPE) of the modelled data to the historical input data, and also MAPE of 24 hour loads of one day ahead which gave 2.9%, the corresponding outputs are shown in Figures. 8.1, 8.2 & 8.3.
Figure 8.1. MATLAB Result

Figure 8.2. Actual Vs. Predicted Load

B. Regression Trees:

This model is also found good and the results are shown in Figures 9.1, 9.2 & 9.3

Figure 9.1. MATLAB Result
C. Multiple linear regressions: This model has more MAPE, since it cannot work with nonlinear behaviours. Corresponding plots are shown in Figures 10.1, 10.2 & 10.3.

D. Curve Fitting Model (Fourier): This model is quite good especially when it works with multiple linear regression as shown in Figures 11.1, 11.2 & 11.3.
Figure 11.1 MATLAB Result

Figure 11.2 Actual Vs. Predicted Load

E. Curve Fitting + Regression Trees + Neural Network (averaging Model):
This is the best model of all as it uses the advantages of all previous models and it reached MAPE of 2.66% of one day ahead loads as shown in Figures. 12.1, 12.2 & 12.3.

Figure 12.1. MATLAB Result
F. Finally the following plot shows, the prediction of all previous models when compared to the actual loads:

5. Comparison

<table>
<thead>
<tr>
<th>S.No</th>
<th>Technique Used</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ANN</td>
<td>2.9</td>
</tr>
<tr>
<td>2</td>
<td>Regression Trees</td>
<td>4.067</td>
</tr>
<tr>
<td>3</td>
<td>Multiple Linear Regression</td>
<td>4.665</td>
</tr>
<tr>
<td>4</td>
<td>Curve Fitting</td>
<td>4.92</td>
</tr>
<tr>
<td>5</td>
<td>Average Model(Curve Fitting + Regression Trees + Neural Network)</td>
<td>2.66</td>
</tr>
</tbody>
</table>

6. Conclusion

It can be concluded that using forecasting techniques, it is possible to design good models that can do short term Load forecasting that can help electricity generating plants to work efficiently and meet the load demands with less discrepancy between demand and supply.
Acknowledgments

The authors are highly acknowledged S.V.E.C, Tirupathi for the encouragement and support during the research of this paper.

References

Authors

Kumar Reddy Cheepati, was born in Kadapa, Andhra Pradesh, India on July 1st, 1988. He received the B.E., in Electrical & Electronics Engineering from Anna University, Chennai, in 2009, M. Tech from V.T.U. Belgama in 2011. Currently he is working as an Asst. Professor at S.V.E.C, Tirupathi which is affiliated to J.N.T.U, Ananthapaur and doing Ph.D in Electrical & Electronics Engineering from J.N.T.U, Kakinada. His research interests include Artificial Neural Networks, Genetic Algorithms, Particle Swarm Optimization, Power Quality Improvement, Smart Grid and Multi Level Inverters.

T. Nageswara Prasad working as Professor & Head of Electrical and Electronics department at S.V.E.C. Tirupathi. He received Ph. D in Electrical Engineering from S.V University, Tirupathi. His research interests include Artificial Neural Networks, Genetic Algorithms, Particle Swarm Optimization, Power Quality Improvement, Smart Grid and Multi Level Inverters.