International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-1 Number-2 Issue-2 December 2011

Comparisons of Short Term Load Forecasting using Artificial Neural Network and Regression Method

Rajesh Deshmukh [1], Amita Mahor [2]
Department of Power Electronics, NRI Institute of Science and Technology Bhopal [1,2]

Abstract

In power systems the next day's power generation must be scheduled every day, day ahead short-term load forecasting (STLF) is a necessary daily task for power dispatch. Its accuracy affects the economic operation and reliability of the system greatly. Under prediction of STLF leads to insufficient reserve capacity preparation and in turn, increases the operating cost by using expensive peaking units. On the other hand, over prediction of STLF leads to the unnecessarily large reserve capacity, which is also related to high operating cost. the research work in this area is still a challenge to the electrical engineering scholars because of its high complexity. How to estimate the future load with the historical data has remained a difficulty up to now, especially for the load forecasting of holidays, days with extreme weather and other anomalous days. With the recent development of new mathematical, data mining and artificial intelligence tools, it is potentially possible to improve the forecasting result. This paper presents a new neural network based approach for short-term load forecasting that uses the most correlated weather data for training, validating and testing the neural network. Correlation analysis of weather data determines the input parameters of the neural networks. And its results compare to regression method.

Index terms

Load Forecasting, artificial neural network, short term and linear regression.

1. 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.

Currently, power utilities are using various load forecasting techniques worldwide. Most of the developed methods can be broadly categorized into three groups, namely parametric, nonparametric, and artificial intelligence based methods .In the parametric methods, a mathematical or statistical relationship is developed between the load and the factors affecting it. Some examples of these models are time functions, polynomial functions, linear regressions, Fourier series In time-series methods, the load is treated as a time series signal with known periodicity such as seasonal, weekly, or daily. Such repetitive cycle gives a rough prediction of the load at the given season, day of the week, and time of the day. The difference between the estimated and actual load can be considered as a stochastic process, which can be then analyzed using Kalman filter methods. Nonparametric methods forecast the load directly from historical data. For example, using nonparametric regression, the load can be forecasted by calculating an average of historical loads and then assign weights to different loads using a multivariate product kernel. Recently, significant interests and efforts have been directed towards the application of artificial intelligence techniques to load forecasting. This includes the application of expert systems to load forecasting, and comparing its performance to traditional methods. It also includes the use of fuzzy inference and fuzzy-neural models However, the models that have received a high share of efforts and focus are the artificial neural networks (ANNs). The main advantage of ANNs is their outstanding performance in data classifications and function approximation. ANN is also capable of detecting dependencies from historical data without the need to develop a specific regression model. First publications on ANN application to the load forecasting problem were made in late 1980's and early 1990's. Since then ANN have been well accepted in practice, and are used by many utilities. Most of the conventional ANN-based load forecasting methods deals with 24-hour-ahead load forecasting or next day peak load forecasting by using forecasted temperature. The drawback of this method is that when rapid changes in temperature of the forecasted day occur, load power changes significantly, which leads to high forecast error. In addition, conventional neural networks use all similar day's data throughout the training process. However, training of the neural networks using all similar day's data is a complex task and it does not suit learning of neural network.

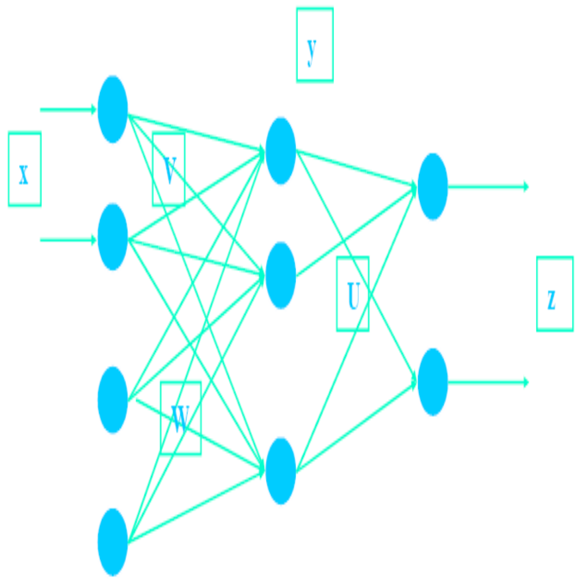


Fig.1 Artificial neural network structure

In the above fig. input signal are represented by X, Y is hidden layer V,W is the interconnection between input layer and hidden layer and Z is the output layer. In case of STLF, the performance of the forecast depends largely on the proper selection of the load affecting variables. In STLF, the key variables are time, forecasted weather variables, and historical load . Therefore, it is vital to identify the input variables, which have significant impacts on the system load. This is particularly important since inclusion of irrelevant inputs or inputs with no significant impact on the target outputs can distort the forecast performance, increase the training time, increase network complexity and reduce the network execution time. One approach to identify the most affecting input variables are by evaluating the statistical correlation between such input variables and the target output. In this paper, the linear correlation coefficient index is calculated between each input variable and the output target. Correlation coefficient with absolute values near 1 implies a high influence of a particular input variable to the target output.

Input variables

Input data of may 2011 collected from state load dispatch center Jabalpur, In case of STLF, the performance of the forecast depends largely on the proper selection of the load affecting variables. In STLF, the key variables are time, forecasted weather variables, and historical load . Therefore, it is vital to identify the input variables, which have significant impacts on the system load. This is particularly important since inclusion of irrelevant inputs or inputs with no significant impact on the target outputs can distort the forecast performance, increase the training time, increase network complexity and reduce the work execution time. In this approach, the forecast time interval is taken to be one hour i.e. forecast is done for each hour. Therefore, it is assumed that the hourly values of the weather parameters can capture the most conservative conditions that may happen during this hour.

Training data

Training is the process by which ANN determines the different network parameters such as weights and biases. In general, the training data set should cover a wide range of input patterns sufficient enough to train the network to recognize and predict the relationship between input variables and target output. Typically, ANNs are trained following a supervised pattern, i.e. the desired output is given for each input and the training process then adjusts the weights and biases to match the desired output . A new method for selecting the training vector is presented in this paper. In this method, the minimum distance between the forecasted input variable and its desired outcome is calculated for the entire historical database.

Network topology

In this paper, a three-layer feed forward neural network is used to construct the STLF model. In theory, a two-layer feed forward network can be used for predictions . For the number of neurons in each layer are 10 and 20, the number of units in the input and output layers are fixed by the number of inputs and the number of outputs, respectively. Since the target output is the forecasted hourly load, the model has one output representing the forecasted load of the target hour. Due to seasonal load variations, four case studies, related to the different four seasons are performed. The number of inputs for each case is dependent on the number of the effective parameters as determined by the correlation analysis as previously discussed in the Load Analysis Section. One of the challenges in the design of ANN is the proper selection of the number of neurons in the hidden layer, which affects the learning capability and varies by the complexity of the problem. In general, a tradeoff between accuracy and generalization ability can be achieved by selecting the proper number of hidden units. While, there is no rigorous set of rules to determine the optimal number of hidden units, the fundamental rule is to select the minimum number of hidden neurons just enough to ensure the complexity of the problem, but not too many to cause over fitting of the training set and losing generalization ability. The approach presented in is also adopted here. The method starts by setting the estimated optimal number of hidden neurons as the square root of the product of the number of inputs times the number of the outputs. Then the number of hidden neurons is gradually incremented by one.

The parameters of three layer artificial neural network for hour ahead load forecasting are as given below:

- No. of layers: 3 (Input layer, Hidden layer, Output layer)
- No of neurons in hidden layer: 10 to 20
- No of neurons in output layer: 1
- Activation function of hidden layer: logsig
- Activation function of output layer: Linear
- Training algorithm: Back-Propagation
- Learning rate (α): 0.1
- No of data sets in each epoch: 63
- No. of epochs for training: 100

2. Regression method

The results obtained from ANN method for day to day forecasting ,week to week forecasting and 24 hrs of the day is divided into three intervals 1:00 am to 08:00 am 09:00 am to 16:00 pm and 17:00pm to 24:00 am. They are comparing to linear regression method, using curve fitting (CF) toolbox.

$$F(X) = P1x + P2$$

Where P1 and P2 is known as regression coefficients.

3. Applications

1. Adaptive learning**:** An ability to learn how to do tasks based on the data given for training or initial experience.
2. Self-Organization**:** An ANN can create its own organization or representation of the information it receives during learning time.
3. Real Time Operation**:** ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.

4. Research work

The main purpose of the research work is to from which method obtained best results for short load forecasting using back propagation training algorithm in artificial neural network and curve fitting toolbox in regression method.

5. Results

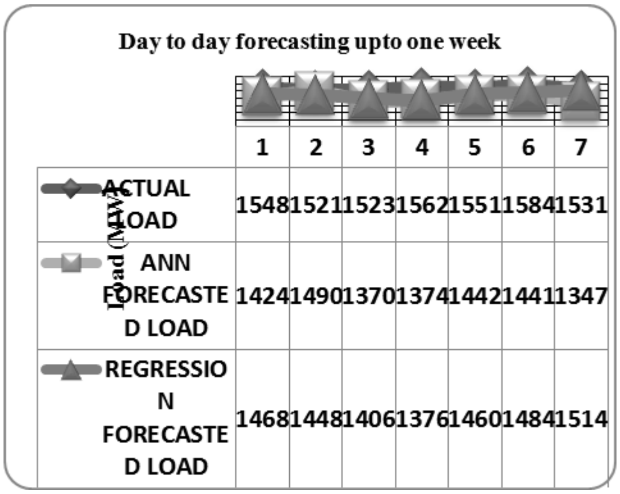
The performance of the developed method for short-term load forecast is tested using the actual hourly load data for month May 2011 which is collected from state load dispatch center Jabalpur, the results obtained by ANN method are compare to the other linear regression method for short-term load forecasting. In this method, the main driving factors, which are impacting the load consumption, are identified. Functional relationships are then established between load and its driving factors by carrying out multiple regression analysis. to verify the performance of the proposed method, four cases are carried out for one hour load forecasting for the four seasons of the year. The neural networks developed have one input layer with the number of neurons equal the number of the input variables, one hidden layer with enough number of neurons for generalization, and one output layer neuron.

The input variables of the ANN are selected based on the correlation analysis. The mean of accuracy obtained between the predicted and actual loads for day to next day , week to next week and hole day divided in to three intervals 1:00 am to 08:00 am, 09:00 am to 16:00pm and 17:00pm to 24:00am has been calculated and presented in the table. This represents a high degree of accuracy in the ability of neural networks to forecast electric load.

6. Comparisons of results

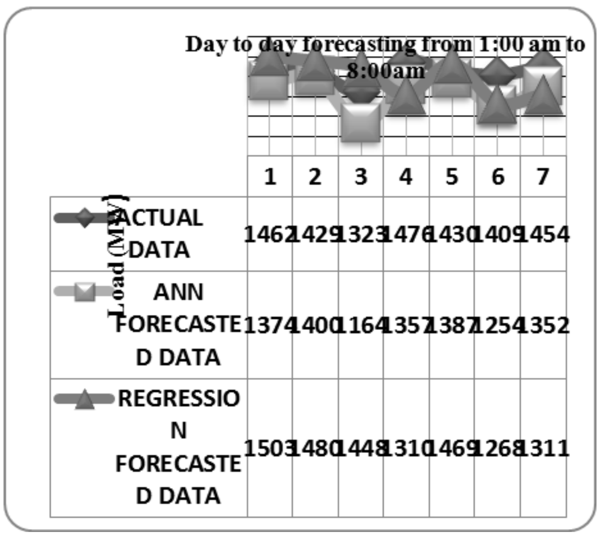
Table 1: Result Comparison

Mode of Forecasting	ANN Method	Regression Method
Day to Day forecasting	92.32%	93.71%
Week to Week forecasting	93.41%	94.82%
1:00 am to 08:00 am forecasting	93.05%	91.72%
9:00 am to 16:00 pm forecasting	93.81%	92.42%
17:00 am to 24:00 am forecasting	92.12%	91.57%



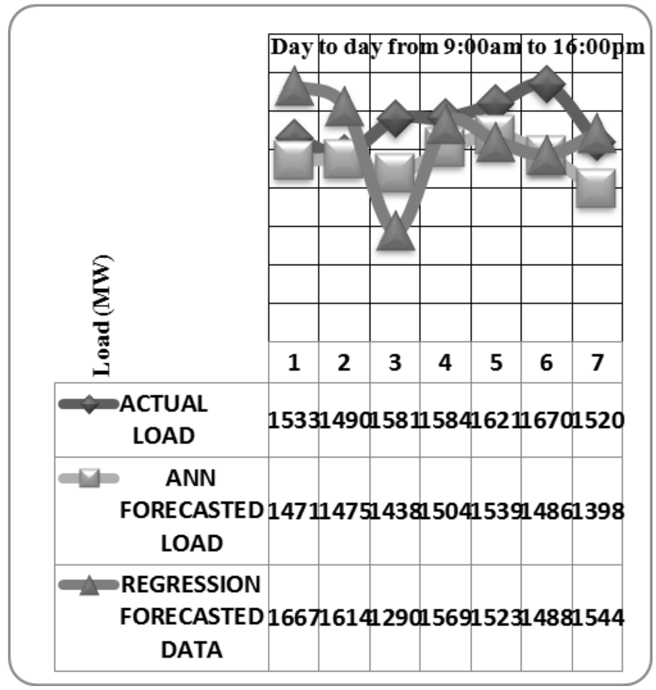
	1	2	3	4	5	6	7
ACTUAL LOAD	1548	1521	1523	1562	1551	1584	1531
ANN FORECASTED LOAD	1424	1490	1370	1374	1442	1441	1347
REGRESSION FORECASTED LOAD	1468	1448	1406	1376	1460	1484	1514

Fig.2 Day to Day Forecasting of May (02.05.2011 to 08.05.2011)



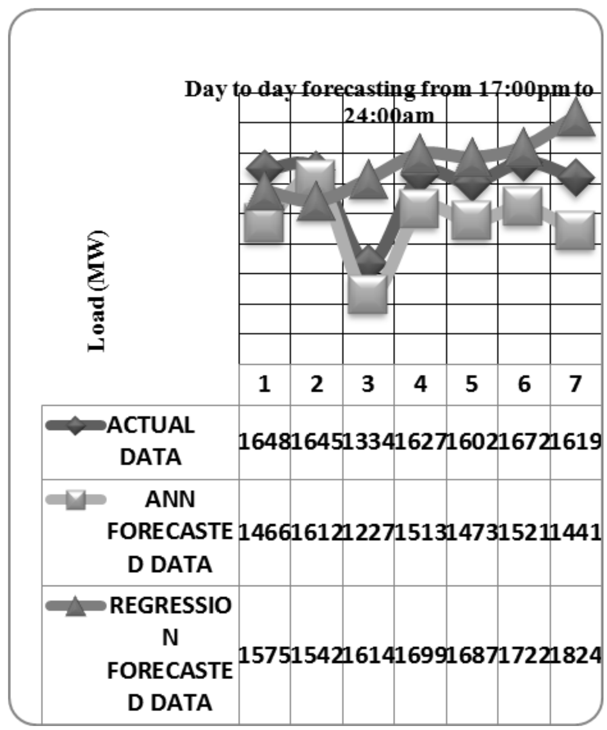
	1	2	3	4	5	6	7
ACTUAL DATA	1462	1429	1323	1476	1430	1409	1454
ANN FORECASTED DATA	1374	1400	1164	1357	1387	1254	1352
REGRESSION FORECASTED DATA	1503	1480	1448	1310	1469	1268	1311

Fig.3 Day to day forecasting of may 02.05.2011 from 01:00 am to 08:00 am.



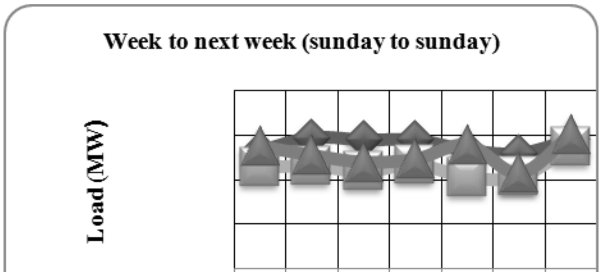
	1	2	3	4	5	6	7
ACTUAL LOAD	1533	1490	1581	1584	1621	1670	1520
ANN FORECASTED LOAD	1471	1475	1438	1504	1539	1486	1398
REGRESSION FORECASTED DATA	1667	1614	1290	1569	1523	1488	1544

Fig.4 Day to day forecasting of may 02.05.2011 from 09:00 am to 16:00 pm.



	1	2	3	4	5	6	7
ACTUAL DATA	1648	1645	1334	1627	1602	1672	1619
ANN FORECASTED DATA	1466	1612	1227	1513	1473	1521	1441
REGRESSION FORECASTED DATA	1575	1542	1614	1699	1687	1722	1824

Fig.5 Day to day forecasting of may 02.05.2011 from 17:00 pm to 24:00 am.



	1	2	3	4	5	6	7
ACTUAL DATA	1531	1591	1580	1580	1534	1519	1580
ANN FORECASTED DATA	1454	1463	1437	1469	1411	1412	1548
REGRESSION FORECASTED DATA	1552	1511	1480	1493	1558	1430	1604

Fig.6 Week to week forecasting from May (01.05.2011 to 14.05.2011)

Conclusion

The result of MLP network model obtained from both ANN and regression method for days to day , week to week and 24 hrs of the day is divided in to three intervals for short term load forecast for the west zone region, ANN method shows that MLP network has a good performance and reasonable prediction accuracy was achieved for this model as quit compare to regression method. Its forecasting reliabilities were evaluated by computing the mean accuracy between the exact and predicted values. We were able to obtain a high degree of accuracy.

References

[1] T.W.S. Chow, C.T. Leung, 'Nonlinear autoregressive integrated neural network model for short-term load forecasting', IEE Proceedings, Generation, Transmission and Distribution, 1996, 143, 500 – 506.

[2] H. Mori, and N. Kosemura, 'Optimal regression tree based rule discovery for short-term load forecasting', Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference, 2001, 1, 421 – 426.

[3] Wenjin Dai , ping Wang 'Application of Pattern Recognition and Artificial Neural Network to Load Forecasting in Electric Power System.

[4] A.A. Desouky, M.M. Elkateb, 'Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA', IEE Proceedings of Generation, Transmission and Distribution, 2000,147(4), 213 – 217.

[5] J.Y. Fan, J.D. McDonald, 'A real-time implementation of short – term load forecasting for distribution power systems', IEEE Transactions on Power Systems, 1994, 9, 988 – 994.

[6] S.E. Skarman, M. Georgiopoulous, 'Short-term electrical load forecasting using a fuzzy ARTMAP neural network', Proceedings of the SPIE, (1998), 181 - 191.

[7] M. Peng, N.F. Hubele, G.G. Karady, 'Advancement in the application of neural networks for short-term load forecasting', IEEE Transactions on Power Systems, 1992, 7, 250 – 257.

[8] A.G.Bakirtzis, et al., 'A neural network short-term load forecasting model for the Greek power system', IEEE Transactions on Power Systems, 1996, 11, 858 – 863.