International Journal of Natural and Engineering Sciences 7 (1): 01-05, 2013
ISSN: 1307-1149, E-ISSN: 2146-0086, www.nobel.gen.tr

Short-Term and Mid -Term Load Forecasting Using Multi Layer Perceptron's and Radial basic Function

Mohammad KARIMI[1,*] Mahdi NOOSHYAR[2] Saadat AMIRI[3] Masood NAJAFZADEH[3] Payam FARHADI[1]

[1]Young Researchers Club, Parsabad Moghan Branch, Islamic Azad University, Parabad Moghan, Iran
[2]Technical Engineering Department, University of Mohaghegh Ardabili, Ardabil, Iran
[3] Department of Electrical Engineering, Science and Research Branch Islamic Azad University, Ardabil, Iran

***Corresponding author:**
E-mail: m.karimi@iaupmogan.ac.ir

Received: June 01, 2012
Accepted: July 09, 2012

Abstract

Load forecasting plays an important role in power systems supply-demand action. For power companies, load forecasting is vital since this forecasting is a base for planning of future development, the economic dispatch, determining the security and control systems and effective operation (investment and decisions for electric generating company) in power systems. In this paper, two Neural Networks (NNs); i.e. Multi Layer Perceptron's (MLP) and Radial Basic Function (RBF) are proposed for short and mid terms load forecasting. The data of Azarbayjan Electrical Network in West-North of Iran has employed for training of these NNs. Four statistical indices used to analyze obtained results and compare abilities of MLP and RBF, these indices are: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Symmetric Mean Absolute Percentage Error (SMAPE) and Relative Error Percentage (REP).

Keywords: Mid Term Load Forecasting, Neural Networks, Short Term Load Forecasting, Realistic Power Network.

INTRODUCTION

Appropriate and exact load forecasting can properly prepare the reserve capacity; in fact, unnecessary reservations lead to reduction of power loss. Load forecasting is however an elaborated task for the following major reasons. First, because the load-series is complex and exhibits several levels of seasonality. Second, the period at a given hour or time period is dependent not only on the load at the previous period, but also is temperature-dependent [1].

Based on length of time, load forecasting can be divided into four major categories: Long-Term Load Forecasting (LTLF), Midterm Load Forecasting (MTLF), Short Term Load Forecasting (STLF) and Very Short Term Load Forecasting (VSTLF). Long-term forecasts during the period of a year or more, it is necessary for generation or power plant planning in the future, while MTLF for a few days or few weeks or months; it is important for fuel reserve planning or unit commitment. STLF having period time in a minute to three days; it is vital for unit commitment and operation, ultimately VSTLF having period time in a minute and is of great importance for real time operation.

The past works in literature have been classified based on problem solution techniques. For this, three categories are presented; i.e. neural networks (NNs) and non-NNs as well as hybrid approaches.

Different branches of NNs have been suggested for load forecasting [2-6]. Chen *et al.* have proposed a daily-based wavelet NN method for holiday STLF [2]. In [3], authors have proposed a novel method to forecast the use of regressive and artificial NN (ANN) models with the case study carried out on a given Turkish network. In this research, two methods are separately performed and then compared. It reveals that both methods result in high accurate results.

Authors in [4] have employed time lagged feed forward network made load forecasting in short time which combines conventional network topology good handling of time dependencies by means of gamma memory. The proposed technique in [5] has been organized in two parts; first, section which is based on NN and self-organizing feature maps has been used to identify those days with similar hourly load patterns, in second section feed-forward multilayer NN is designed to predict daily peak load and light load. Authors in [6] have suggested a novel hierarchical hybrid neural model for long term load forecasting.

Researchers have developed several techniques for load forecasting [7-10]. The proposed approach in [7] for STLF consists of two stages based on time-series methods. Authors have claimed that proposed technique can be used to support vector machine (SVM) according to the characteristic of historical data. Elattar *et al* in [8] have modified support vector regression (SVR) by combining the SVR and locally weighted regression (LWR), also they employed the weighted distance algorithm. The used approaches for STLF by Taylor and McSharry are: ARIMA modeling, periodic AR modeling, an extension for double seasonality of Holt-Winters exponential smoothing alternative exponential smoothing formulation and a method based on the principal component analysis (PCA)[9]. In [10], Chaotic Particle Swarm Optimization (CPSO) algorithm has been employed to choose the suitable parameter combination for a SVR model as model of electric load forecasting.

Authors have combined different techniques and presented hybrid approaches to obtain better solution for load forecasting [11-15]. The proposed hybrid approach in [11] consists of two sections: wavelet fuzzy neural network (WFNN) and fuzzy neural network Inference (FNCI) which was used the fuzzified wavelet features as the inputs to fuzzy neural networks (FNN) and the Choquet integral as the outputs of FNN. In [12], fuzzy Inductive Reasoning (FIR) has been applied to the STLF problem and simulated rebounding algorithm has been used to choose the inputs of the FIR model. Pandey *et al.*, suggested a wavelet decomposition approach with various conventional statistical and NNs as well as fuzzy inference based hybrid approaches has for STLF [13].

The proposed hybrid model in [14] for STLF has two stages: first, NNs are trained by the load patterns, then the fuzzy expert systems modify the provisional forecasted load considering the possibility of load variation due to changes in temperature and the load behavior of holiday.

Finally, in [15], a novel hybrid method based on a Self-Organizing Fuzzy NN (SOFNN) learning method with a bi-level optimization method has been proposed for STLF. In forecasting process, SOFNN and bi-level optimization methods are used to determine model structure and parameters and select the best pre-training parameters to ensure that the best fuzzy neural networks are identified, respectively.

This paper develops a comparison between two powerful branches of NNs for short and mid terms load forecasting. For this, Multi layer Perceptron's (MLP) and Radial Basic Function (RBF) has been used to forecast. This paper's contributions are as follows:

In contrary to the many conventional works which only review the past completed work, in this paper, authors have classified past works in several groups.

Two reliable NNs were suggested to solve load forecasting problem. Also, two terms have been considered for load forecasting; i.e. short and mid terms.

Four valid statistical indices have been introduced to judge between capability between MLP and RBF.

Load Forecasting using RBF

Neural Networks (NNs) have been tested on various power system problems such as planning, control, analysis, protection, design, load forecasting, and fault diagnosis. The last three ones are the most popular [16].

The RBF neural network is a forward network pattern with appropriate performance, global approximation, and also is free from trapping in the local minima. It is a multi-input, single-output system consisting of an input layer, a hidden layer, and an output layer. In data processing, the hidden layer performs nonlinear transforms for the feature extraction and the output layer gives a linear combination of output weights.

The input is n-dimensions, learning samples are (X,Y) where $X=(X_1, X_2, \ldots, X_N)$ is input variable, $X_i= (x_{i1}, x_{i2}, \ldots, x_{in})^T (1\le i\le N)$, and the expected output is $Y_i= (y_{i1}, y_{i2}, \ldots, y_{iN})$, N is the training number. If the input is X_i, the output of the ith node in the hidden layer can be expressed as,

$$G(X_i, C_j, \sigma_j) = \exp\left(\|X_i - C_j\| / 2\sigma^2\right) \quad (1)$$

where, $C_j=(c_{j1}, c_{j2},\ldots,c_{jn})^T$ and σ_j are the center and width of the Gaussian function of the jth node in the hidden layer.

For an input X_i the expected output of network is

$$y_i = \sum_{j=1}^{M} G(X_i, C_j, \sigma_j)\omega_j + e_i \quad (2)$$

which ω_j is the weight between the jth neuron of the hidden layer and output neuron. M stands for neuron number in the hidden layer, and y_i and e_i are the expected output of X_i and the error of fitting, respectively. e_i can be obtained by transposition of Eq. (2), that is

$$e_i = y_i - \sum_{j=1}^{M} G(X_i, C_j, \sigma_j)\omega_j \quad (3)$$

As for the selection of the center parameter of the Gaussian function, this paper uses the orthogonal least square algorithm. The least square algorithm is also applied to train the output weight in order to minimize total error, that is

$$\min E = \frac{1}{2}\sum_{i=1}^{N} e_i^2 \quad (4)$$

Eq. (2) can be written in a matrix form as follows:

$$Y = PW + E \quad (5)$$

where, $Y_i= [y_{i1}, y_{i2}, \ldots, y_{iN}]^T$, $Y_i= [P_1, P_2, \ldots, P_N]$is output matrix of the hidden layer, and $Y_i= [p_{i1}, p_{i2}, \ldots, p_{iN}]$ $(1\le i\le N)$. W and E are output weights and error vectors.

Carry orthogonal decomposition on P,

$$P = HA \quad (6)$$

where, A is upper triangular matrix with diagonal value 1, and matrix H contains the orthogonal vectors H_i. Eq. (7) can be changed into

$$Y = (PA^{-1})(AW) + E \quad (7)$$

suppose

$$S = AW \quad (8)$$

Then, Y is transformed as

$$Y = (PA^{-1})S + E = HS + E \quad (9)$$

Therefore, using the least square algorithm S is obtained from Eq. (9) as

$$S = (H^T H)^{-1} H^T Y \quad (10)$$

$$e_i = (s_i \omega_i^T \omega_i) / Y^T Y, \quad (1 \le i \le N) \quad (11)$$

Since the number of hidden layers, M, that satisfies the training accuracy and generalization ability are far less than training number, N, the accumulating variance can be used for sample assessment and center selection. When ρ_i meets

$$\rho_i = \sum_{j}^{i} e_j, \quad \rho_i > \rho \quad (12)$$

The ith sample with maximal ρ_i is chosen as network center. $0<\rho<1$ is preset tolerant variance limit. Weight can be determined by [17],

$$\hat{W} = A^{-1}S \quad (13)$$

Load Forecasting Using MLP

In this paper, the error back propagation algorithm is used to train the MLP network. Presenting an input pattern to the network produces an output vector. According to the difference between the produced and target outputs, the network's weights (Wij) are adjusted to reduce the output error. The error at the output layer propagates backward to the hidden layer, until it reaches the input layer.

The output of neuron i, Oi, is connected to the input of neuron j through the interconnection weight Wij. Unless neuron k is one of the input neurons, the state of neuron k is given as

$$O_k = f\left(\sum W_{ik} Q_i\right) \tag{14}$$

where, $f(x)=1/(1+e^{-x})$, and the sum is over all neurons in the adjacent layer. Let the target state of the output neuron be t. Thus, the error at the output neuron can be defined as

$$E = \frac{1}{2}\left(t_k - O_k\right)^2 \tag{15}$$

where, neuron k is the output neuron. The gradient descent algorithm adapts the weights according to the gradient error,

$$\Delta W_{ij} \alpha - \frac{\partial E}{\partial W_{ij}} = -\frac{\partial E}{\partial O_j}\frac{\partial O_j}{\partial W_{ij}} \tag{16}$$

Specifically, we define the error signal as

$$\delta_j = -\frac{\partial E}{\partial O_j} \tag{17}$$

With some manipulation, we can get the following GDR as,

$$\Delta W_{ij} = \varepsilon \delta_j O_i \tag{18}$$

where, ε is an adaptation gain. δ_j is computed based on whether or not neuron j is in the output layer. If neuron j is one of the output neurons, then

$$\delta_j = \left(t - O_j\right) O_j \left(1 - O_j\right) \tag{19}$$

If neuron j is not in the output layer, then

$$\delta_j = O_j\left(1 - O_j\right)\sum_k \delta_k W_{jk} \tag{20}$$

In order to improve the convergence characteristics, we can introduce a momentum term with momentum gain to Eq. (20), thus

$$\Delta W_{ij}(n+1) = \varepsilon \delta_j O_j + \alpha \Delta W_{ij}(n) \tag{21}$$

where, n represents the iteration index.

Once the neural network is trained, it produces a very fast output for given input data. It only requires a few multiplications, additions, and calculations of sigmoid function [18].

SIMULATION RESULTS

In this section, two networks data has been employed to train MLP and RBF for STLF and MTLF. For each, four statistical indices are used to analyze obtained results. These indices are Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Symmetric Mean Absolute Percentage Error (SMAPE) and Relative Error Percentage (REP).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|A_i - F_i\right| \tag{22}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|A_i - E\right|}{A_i} \tag{23}$$

$$SMAPE = \frac{1}{N}\sum_{i=1}^{n}\frac{\left|A_i - F_i\right|}{A_i + F_i} \times 100 \tag{24}$$

$$REP = \sqrt{\frac{\sum_{i=1}^{N}(A_i - F_i)^2}{\sum_{i=1}^{N} A_i^2}} \times 100 \tag{25}$$

where, A_i and F_i are the actual and the forecasted values, respectively, N is the testing dataset size, and i denotes the test instance index.

Data of Taghi-dizaj substation in Ardabil in North West of Iran and temperature in 1st June 2011 has been used for training of MLP and RBF in STLF. MLP has been designed in five 3, 5, 7, 9 and 11 layers. Figure 1 shows real and forecasted value in 1st June.

By considering values of Figure 1, between 1-9 hours demand of network is less than 30 kW while after hour 9 this demand increases significantly. Maximum demand occurs between 21-23 hours. Data of a part of Tabriz city electric network in 12 months of 2011 has been used to train MLP and RBF. The real and forecasted values have been illustrated in Figure2.

In all cases, RBF results are the worst solution and are very far from the real value. The values of statistical indices for STLF and MTLF have been listed in Table 1.

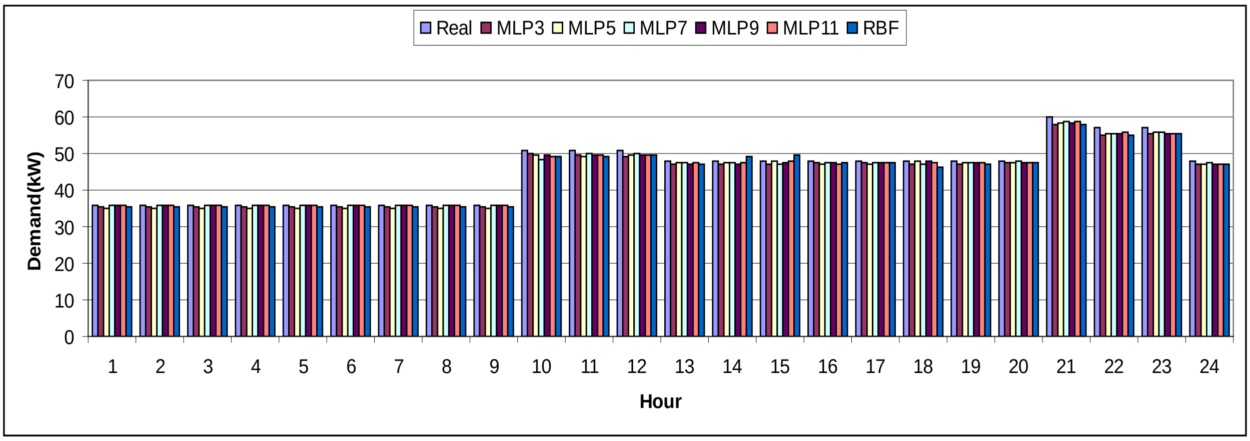

Figure 1. Short load forecasting by MLP and RBF in 1st June 2011

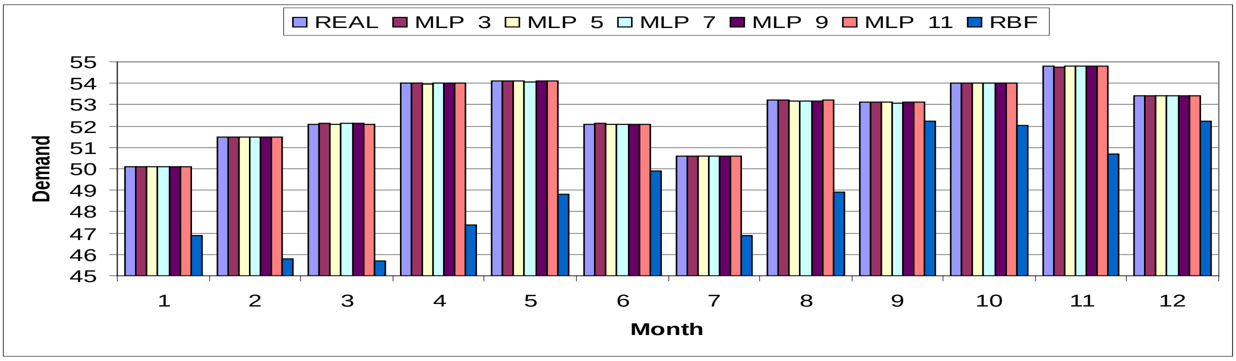

Figure 2. Mid load forecasting by MLP and RBF at 2011

Table 1. values of statistical indices for STLF and MTLF

Load Forecating	Index	MLP 3	MLP 5	MLP 7	MLP 9	MLP 11	RBF
Short Term	MAE	0.89833	0.99583	0.637917	0.71125	0.64583	1.05752
	MAPE	1.91163	2.23606	1.298228	1.47331	1.31478	2.26702
	SMAPE	0.9663	1.1325	0.6564	0.7436	0.6636	1.1446
	REP	9.613	10.656	6.826	7.611	6.911	8.891
Mid Term	MAE	0.01428	0.00514	0.00749	0.00643	0.00235	3.79773
	MAPE	0.026825	0.009592	0.01406	0.01211	0.00442	7.2047
	SMAPE	0.013415	0.00479	0.00703	0.00605	0.00221	3.77250
	REP	0.02374	0.02773	0.018545	0.01652	0.00153	24.931

In STLF case, by attention to results of Table 1, MLP7 and RBF present the best and worst MAE, respectively. MAE of MLP7 is 0.2604, 0.3579, 0.0733, 0.0079 and 0.4196 less than related parameter of MLP3, MLP5, MLP9, MLP11 and RBF, respectively. Maximum and minimum MAPE is similar to MAE. MLP7 has the lowest MAPE respect to other approaches. After MLP7, MLP9 and MLP11 present lower MAPE. MAPE of MLP7 is 0.6134, 0.9378, 0.1751, 0.0166 and 0.9688 less than MLP3, MLP5, MLP9, MLP11 and RBF, respectively. Maximum SMAPE is devoted to RBF and is 0.1783, 0.0121, 0.4882, 0.401 and 0.481more than corresponding parameter of MLP3, MLP5, MLP7 and MLP11, respectively. In this case, MLP7, MLP9, MLP11, MLP3 and MLP5 techniques are better solution, respectively. Finally, REP of MLP7 and MLP5 are minimum and maximum among six techniques, respectively. REP of MLP7 is 2.787, 3.83, 0.785, 0.085 and 2.065 less than related parameter of MLP3, MLP5, MLP9, MLP11 and RBF, respectively.

In MTLF, MLP11 is best option among MLPs. In MAE, MLP3 presents the worst values which its results is 0.00914, 0.00679, 0.00785 and 0.01193 more than related parameter of MLP5, MLP7, MLP9 and MLP11, respectively. Diagram of MAPE show MLP3 and MLP11 has maximum and minimum values, respectively. MLP5 and MLP9 as well as MLP7 are between MLP3 and MLP11, respectively. MAPE value of MLP3 is 0.0172, 0.0128, 0.0147 and 0.0224 more than MLP5, MLP7, MLP9 and MLP11, respectively. This increment in SMAPE is 0.0086, 0.0064, 0.0074 and 0.0112, respectively. REP value of MLP5 is the worst case among MLPs and its value is 0.004, 0.0092, 0.0112 and 0.0262, respectively.

DISCUSSION AND CONCLUSION

This context suggests the use of two branches of NNs for short and mid terms load forecasting. Then, concept of forecasting and adjustments of MLP and RBF for them have been formulated. Simulations have been done in two practical cases. From simulation results the followings are highlighted;

Remark i) In general, MLP presents better solution respect to RBF in solvin load forecasting problem in short and mid term cases. This fact in MTLF is more clear.

Remark ii) Forecasting in mid term is more accurate than short term.

Remark iii) Among MLPs, from the viewpoint of better solution for STLF these five kinds of MLP can be classified as follows: MLP7, MLP11, MLP9, MLP3 and MLP5.

Remark ii) Among MLPs, from the viewpoint of better solution for MTLF these five kinds of MLP can be classified as follows: MLP11, MLP5, MLP7, MLP9 and MLP3.

REFERENCES

[1] E. Souza, R. C. 2001. Neural networks for short-term load forecasting: A review and evaluation. IEEE Transaction of Power System;.16(1): 44-55.

[2] Chen Y., Luh P. B., Guan C., Zhao Y., Michel L.D., Coolbeth M. A., Friedland P. B., Rourke S. J. 2010. Short-term load forecasting: similar day-based wavelet neural networks. IEEE Transactions on Power Systems; 25(1): 332-330.

[3] basaran filik U., kurban M. 2007. A new approach for the short term load forecasting with autoregressive and artificial neural network models. International journal of computational intelligence research; 3(1):66-71.

[4] Hippert, H. S., Pedreira, C. E., Souza, R. C. 2001. Neural networks for short-term load forecasting: A review and evaluation. IEEE Transaction of Power System; 16(1):44-55.

[5] Hsu Y.Y., Yang Ch.Ch. 1991. Design of artificial neural networks for short-term load forecasting. Part I: Self-organising feature maps for day type identification. IEE Proceedings-C; 138 (5):407- 413.

[6] Carpinteiro O. A.S., Leme R. C., Zambroni de Souza A.C., Pinheiro C. A.M., oreira E. M. 2007. Long-term load forecasting via a hierarchical neural model with time integrators. Electric Power Systems Research; 77(3-4): 371–378.

[7] Wang Y., Xia Q., Kang Ch. 2011. Secondary Forecasting Based on Deviation Analysis for Short-Term Load Forecasting. IEEE Transactions on Power Systems; 26(2):500-507.

[8] Elattar E. E., Goulermas J., Wu Q. H. 2010. Electric Load Forecasting Based on Locally Weighted Support Vector Regression. IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews; 40(4): 438-447.

[9] Taylor J. W., McSharry P. E. Short-Term Load Forecasting Methods: An Evaluation Based on European Data. IEEE Transactions on Power Systems; 22(4): 2213-2219.

[10] Hong W.Ch.,2009. Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model. Energy Conversion and Management; 50(1):105–117.

[11] Hanmandlu M., Chauhan B. K. 2011. Load Forecasting Using Hybrid Models. IEEE Transactions on Power Systems; 26(1): 20-29.

[12] Hinojosa V. H., Hoese A. 2010. Short-Term Load Forecasting Using Fuzzy Inductive Reasoning and Evolutionary Algorithms. IEEE Transactions on Power Systems; 25(1): 565-574.

[13] Pandey A. Sh., Singh D., Kumar Sinha S. 2010. Intelligent Hybrid Wavelet Models for Short-Term Load Forecasting. IEEE Transactions on Power Systems; 25(3): 1266-1272.

[14] Kim K.H., Park J.K., Hwang K.J., Kim S.H. 1995. Implementation of Hybrid Short-term Load Forecasting System Using Artificial Neural Networks and Fuzzy Expert Systems. IEEE Transactions on Power Systems;10(3): 1534-1539.

[15] Mao H., Zeng X.J., Leng G., Zhai Y.J., Keane J. A. 2009. Short-Term and Midterm Load Forecasting Using a Bilevel Optimization Model. IEEE Transactions on Power Systems; 24 (2):1080-1090.

[16] Mori H. 1996. State-of-the-Art Overview on Artificial Neural Networks in Power Systems. IEEE Power Engineering Society, Edited by M. El-Sharkawi and D. Niebur: 51-70.

[17] Yun Zh., Quan Zh., Caixin S., Shaolan L., Yuming L., Yang S. 2008. RBF Neural Network and ANFIS-Based Short-Term Load Forecasting Approach in Real-Time Price Environment. IEEE Transactions on Power Systems; 23(3): 853-858.

[18] Khan M. R., Ondrusek C. 2002. Short Term Load Forecasting with Multilayer Perceptron and Recurrent Neural Networks. Journal of Electrical Engineering; 53(1-2):17-23.