CHEMICAL ENGINEERING TRANSACTIONS
VOL. 51, 2016

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Tichun Wang, Hongyang Zhang, Lei Tian

ISBN 978-88-95608-43-3; **ISSN** 2283-9216
DOI: 10.3303/CET1651099

Research on Power Load Forecasting Based on the Improved Elman Neural Network

Jiarui Zhang

College of Electrical Engineering & New Energy,China Three Gorges University, Yichang Hubei province China
zhangjiarui2055@126.com

The forecast of the electric power load is related to the development of economy. It is also related to the national security and the daily operation of the society. The result of the power load forecasting is not only directly related to the economic benefits of the power enterprise, but also is very important to the users. From improving the accuracy of the power load forecasting, this paper proposes an improved Elman neural network algorithm to predict the power load. Then, we analyze and compare the experimental results by the numerical experiments. We find that the experimental results are better than the traditional prediction algorithm. The experimental results show that the algorithm is reliable and effective.

1. Introduction

The power load changes with the change of the climate and the industrial structure adjustment (Wang and Chui, 2015). In addition, with the development of the economy and the improvement of people's living standard, the electric power load is also increasing. These factors make it difficult to manage the power system (Koprinska, 2015). It can improve the safety and economy operation of the power grid to study the power load. It also has the profound theoretical guidance and practical application value of studying the power system. In addition, it has the great significance for the stability and the economic operation of the power system (Yang, 2015).
Power load characteristic index is the number of load characteristics, that is, the characteristic value of the load change.
The load characteristic index in common use of china is shown as follows.

Table 1: Load characteristic index in common use of china

Description class (absolute)	Maximum load, minimum load, average load, peak valley difference, maximum load utilization hours
Comparative class (relative)	Loading rate, average daily load rate, minimum load rate, peak valley ratio, month rate of balanced production, annual rate of balanced production, at the same time rate, the rate, peak load
Curve class	load curve

Table 2: Classification of the power load forecasting (Time)

Prediction type	Time	Main use
Ultra short term power load forecasting	10min~1hour	Power grid on-line control
short term power load forecasting	1day~1year	Generation, power supply and distribution plan
Medium term power load forecasting	1year~5year	Grid operation plan and development plan
Long term power load forecasting	More than 5year	Power grid development planning

Please cite this article as: Zhang J.R., 2016, Research on power load forecasting based on the improved elman neural network, Chemical Engineering Transactions, 51, 589-594 DOI:10.3303/CET1651099

Power load forecasting is used to predict the time distribution and spatial distribution of future power grid load. The power load forecasting is classified by the time as follows.
And the power load forecasting is classified by the different industries as follows.

*Table 3: Classification of the power load forecasting (*industries*)*

Serial number	Name
1	Urban civil load
2	Commercial load
3	Rural load
4	Industrial load

Considering the influence on the power load by the economic, social, climate and other uncertain factors, if we only consider a change trend or a single factor, it is difficult to accurately describe the actual complex variation law of the power load forecasting. Now the methods of the power load forecasting are mainly the following.
(1) Grey forecasting method. The grey forecasting method has the advantages of not considering the load distribution and the changed trend of the load. However, the value of grey prediction in the later period is larger (Bahrami, 2014).
(2) Artificial neural network. This method has strong robustness and strong learning ability. The disadvantage is that the learning convergence speed is slow (Wang, 2012).
(3) Combination forecasting method. The method is to obtain the final results by the weighted average of the predicted values obtained by the multiple prediction methods. (Lee and Hong, 2015).
(4) Fuzzy theory. In recent years, the system modeling, the selection of prediction algorithm and the improved algorithm of the short term load forecasting of power system based on Fuzzy theory are in the exploration and development stage. Fuzzy theory is often combined with other methods to achieve good prediction results. The research of influencing factors is the precondition of using fuzzy theory to predict. Therefore, the application of fuzzy theory in power load forecasting has some limitations (Thair et al., 2015).
The characteristics of power load forecasting are shown in the following figure.

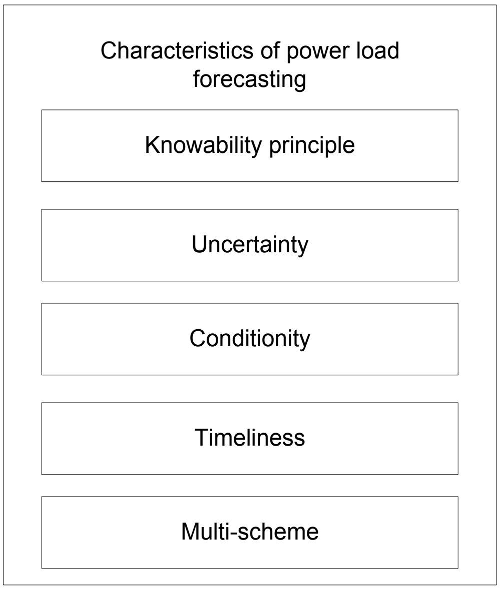

Figure 1: The characteristics of power load forecasting

2. Elman structure

2.1 Network structure

In addition to the input layer, the hidden layer and the output layer, neural network also has a special undertaking layer. Its structure is shown in figure 1. The undertaking layer is also called the context layer or the state layer. It is used to memory the output value of the previous unit time for the hidden layer. It can be considered as one step delay operator and it makes the Internet have the function of the dynamic memory(Muhammad and Saeed, 2010).

2.2 Learning algorithm

In the figure 2, we assume that the number of the neurons of the input layer, the hidden layer and the output layer are M, N, L. $W_{1ij}(t)$ is the connected weight matrix from the input layer to the hidden layer. $W_{2jq}(t)$ is the connected weight matrix from the hidden layer to the output layer. $W_{3rt}(t)$ is the connected weight matrix from the undertaking layer to the hidden layer. $f(\cdot)$ is the nonlinear vector function which is constituted by the excitation function of the hidden layer unit. $g(\cdot)$ is the nonlinear vector function which is constituted by the excitation function of the output layer unit (Sanjay and Sanjay, 2012).

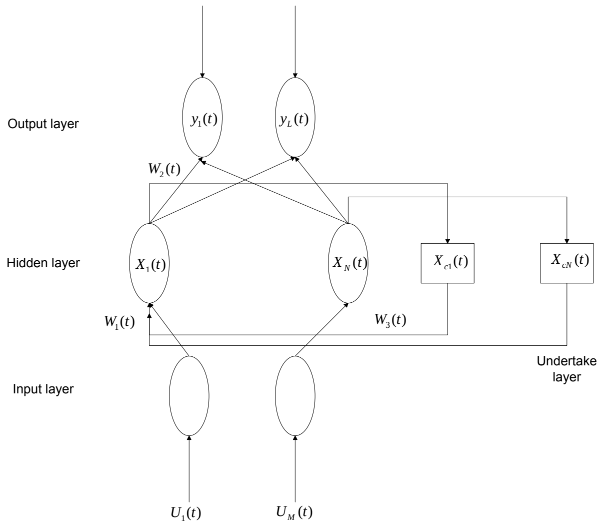

Figure 2: Structure of Elman neural network

According to the network structure, the nonlinear state space expression is as follows.
The output of the hidden layer is,

$$X_L(t) = f[\sum_{i=1}^{M} W_{1ij}(t)U_i(t) + \sum_{r=1}^{N} W_{3rj}(t)X_{Cr}(t)] \tag{1}$$

Where

$$X_{Cr}(t) = X_j(t-1) \tag{2}$$

The output of the output layer is,

$$y_q(t) = g[\sum_{j=1}^{N} W_{2jq}(t)X_j(t)] \tag{3}$$

The learning algorithm of *Elamn* neural network can adopt the LM learning algorithm. Compared with the conventional gradient descent method of BP and its improved algorithm, the LM algorithm has fast convergence and good robustness (Zhao et al., 2013).
The weight value $W_{1ij}(t)$ of *Elamn* neural network and the correction of $W_{2jq}(t)$ are similar to BP neural network. The main difference is $W_{3rj}(t)$. We can use the chain derivative rule to obtain (Liu et al., 2015).

$$W_{3rj}(t) = \sum_{t=1}^{T} [(y_0(t) - y_q(t))W_{2jq}(t)]f'(X_L(t))X_L(t-1) \tag{4}$$

3. Improved ELAMN neural network

3.1 The selection of the network parameter

Before the adjustment of the connection weight, it needs to give a random initial weight to each connection weight. Because the system is nonlinear, the initial value has the great influence on whether the learning is convergence and is be trapped in local minimum or not (Wang et al., 2014). In general, we hope that the initial weight value can make the state value of each neuron is close to zero. At the same time, they are not zero or equal to a constant. Otherwise, the system cannot be trained. The initial weight value is a random number in (-0.3,+0.3). It can guarantee the normal training of the network and it can make the network maintain in a good state. During the training process, we can consider the adjusted learning rate automatically. The formula of the adjusting the learning rate is as follows.

$$\eta(k+1)=\begin{cases}1.1\eta(k), SSE(k)>SSE(k-1)\cdot 1.05\\ 0.8\eta(k), SSE(k)<SSE(k-1)\\ \eta(k), \quad otherwise\end{cases} \tag{5}$$

Where, *SSE* is the number of the network output errors. The selection of the initial learning rate $\eta(0)$ has big randomness. In general, we take 0.1. In addition, in order to avoid that the network connection weight coefficient matrix correction is in minimum energy predicament, we add the momentum item when correcting the correction weight coefficient matrix. In the training of the additional momentum, we can use the following conditions to adjust the momentum value.

$$mc=\begin{cases}0, SSE(k)>SSE(k-1)\cdot 1.05\\ 0.9, SSE(k)<SSE(k-1)\\ mc, \quad otherwise\end{cases} \tag{6}$$

When

$$SSE(k)<SSE(k-1) \tag{7}$$

This shows that the number of the error in this moment is smaller than the number of the error in the last one. Therefore, in general, the initial value of the momentum factor *mc* is 0.95.

3.2 Improvement of the excitation function

In neural network, in a certain sense, the excitation function can determine the nature of the error function *E*. In the general *Elamn* neural network, the hidden function *f*(*x*) adopts *sigmoid* function. That is,

$$f(x)=1/(1+e^{-x}) \tag{8}$$

The output layer *f*(*x*) function adopts the linear function. However, *sigmoid* function may makes the convergence rate of the network slow and makes the network fall into the local minimum. Therefore, it needs to be improved.

Aiming at *sigmoid* function curve, the improved method is to add the independent variable factor *k*, the adjustable bias parameter *b*, the constant *a* and the extended function gain *c*. The improved *f*(*x*) function can be expressed as follows.

$$f(x,a,b,c,k)=a+c/(1+e^{-k(x+b)}) \tag{9}$$

Derivative function is,

$$f'=-k/c(f-(2a+c)/2)^2+kc/4 \tag{10}$$

In the learning algorithm of the neural network, The common δ learning rule (Error correction learning) is established based on the idea of minimizing the variance of the error output. The learning rate of the neural network is related to *f'*. The bigger *f'* is, the quicker the learning rate is. Through the adjustment of the parameters, we can realize the translation and expansion of the image of the excitation function. By the excitation function and its derivative, when the function value *f* is close to (2a+c)/2, the derivative value is bigger and the convergence speed of the function is faster. At the same time, the maximum value of *f'* is kc/4. With the increase of the *k* and *c*, the value of *f'* is bigger and the convergence speed is faster. In the above formula, the coefficients *k* and *c* are mainly used to change the amplitude of the excitation function and its derivative. The parameters *a* and *b* are mainly used to translate in the horizontal direction and the vertical direction. Through the adjustment of the parameters, it can make the convergence rate and the prediction accuracy of the function *f* achieve an optimal value.

3.3 Data pre-processing

In order to avoid the saturation of the neuron, we normalize the input data in the input layer. We transform the numerical value to the interval of [0,1]. In the output layer, we adopt the formula (10) to transform the actual value of the output value.

$$y_i = (x_i - x_{\min})/(x_{\max} - x_{\min}) \tag{11}$$

$$x_i = (x_{\max} - x_{\min}) \cdot y_i + x_{\min} \tag{12}$$

Where, x_{max} and x_{min} are the maximum value and the minimum value of the input data. y_i is the normalized value of the sample.

4. Experiment

In order to validate the reliability and validity of the improved *Elamn* neural network algorithm which is proposed by this paper, we use this algorithm to forecast the power load for one province. Among them, the monthly load historical data of the first 5 years are the training set and the monthly load data of the sixth year is the test set. The predicted results are shown as follows.

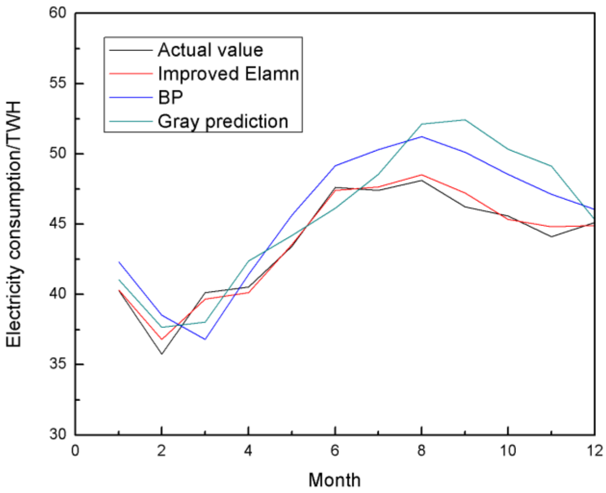

Figure 3: Comparison of prediction effects of different methods

The specific values are shown in Table 4.

Table 4: The power load value of the seventh year and the eighth year

Month	Actual value	Improved Elman	BP	Gray Prediction
1	40.28	40.31	42.31	41.07
2	35.76	36.78	38.54	37.66
3	40.13	39.66	36.79	38.03
4	40.52	40.12	41.42	42.38
5	43.33	43.56	45.64	44.21
6	47.63	47.41	49.17	46.15
7	47.41	47.65	50.30	48.54
8	48.12	48.53	51.21	52.11
9	46.25	47.21	50.11	52.45
10	45.58	45.34	48.56	50.34
11	44.11	44.80	47.12	49.13
12	45.12	44.91	46.03	45.31

From Figure 3 and table 4, we can see that the curve of the actual value of the improved neural network is the most similar to the curve of the forecasting value compared with other methods. The predicted results are closer to the actual value. It shows that the proposed method is most accurate. The experimental results show that the method is effective and reliable.

5. Conclusion

Power system load forecasting is an important and routine work in the power sector. The result of the forecast is not only directly related to the economic benefits of the power enterprise, but also related to the livelihood of the people. Therefore, it is very necessary to study the power load forecasting. In this paper, we propose an improved neural network algorithm and use this algorithm to study the power load forecasting. In this paper, we mainly do the following work. Firstly, we introduce the research background of the power load forecasting. Secondly, we introduce the neural network algorithm. Thirdly, we improve the neural network algorithm and propose the improved neural network algorithm. Then, we use this algorithm to carry on the research to the electric power load forecast. The experimental results show that the algorithm has good validity and reliability.

References

Ardalani-Farsa M., Zolfaghari S., 2010, Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks, Neurocomputing, 73: 2540-2553, DOI: 10.1016/j.neucom.2010.06.004

Bahrami S., Hooshmand R.A., Parastegari M., 2014, Short term electric load forecasting by wavelet transform and grey model improved by PSO (particle swarm optimization) algorithm, Energy 72:434-442, DOI: 10.1016/j.energy.2014.05.065

Kelo S., Dudul S., 2012, A wavelet Elman neural network for short-term electrical load prediction under the influence of temperature, International Journal of Electrical Power & Energy Systems, 43: 1063-1071, DOI: 10.1016/j.ijepes.2012.06.009

Koprinska I., Rana M., Agelidis V.G., 2015, Correlation and instance based feature selection for electricity load forecasting, Knowledge-Based Systems, 82:29-40, DOI: 10.1016/j.knosys.2015.02.017

Lee W., Hong J.Y., 2015, A hybrid dynamic and fuzzy time series model for mid-term power load forecasting, International Journal of Electrical Power & Energy Systems, 64:1057-1062, DOI: 10.1016/j.ijepes.2014.08.006

Liu H., Tian H.Q., Liang X.F., Li Y.F., 2015, Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks, Applied Energy, 157: 183-194, DOI: 10.1016/j.apenergy.2015.08.014

Li Y., Bao Y.Q., Yang B., Chen C., Ruan W.J., 2015, Modification method to deal with the accumulation effects for summer daily electric load forecasting, International Journal of Electrical Power & Energy Systems, 73: 913-918, DOI: 10.1016/j.ijepes.2015.06.027

Lou C.W., Dong M.C., 2015, A novel random fuzzy neural networks for tackling uncertainties of electric load forecasting, International Journal of Electrical Power & Energy Systems, 73:34-44, DOI: 10.1016/j.ijepes.2015.03.003

Mahmoud T.S., Habibi D., Hassan M.Y., Bass O., 2015, Modelling self-optimised short term load forecasting for medium voltage loads using tunning fuzzy systems and Artificial Neural Networks, Energy Conversion and Management, 106: 1396-1408, DOI: 10.1016/j.enconman.2015.10.066.

Wang J.J., Zhang W.Y., Li Y.N., Wang J.Z., Dang Z.L., 2014, Forecasting wind speed using empirical mode decomposition and Elman neural network, Applied Soft Computing, 23: 452-459, DOI: 10.1016/j.asoc.2014.06.027

Wang Y.L., Niu D.X., Ji L., 2012, Short-term power load forecasting based on IVL-BP neural network technology, Systems Engineering Procedia, 4:168-174, DOI: 10.1016/j.sepro.2011.11.062

Zhao J., Zhu X.L., Wang W., Liu Y., 2013, Extended Kalman filter-based Elman networks for industrial time series prediction with GPU acceleration, Neurocomputing, 118: 215-224, DOI: 10.1016/j.ijepes.2012.06.009