

Research on Prediction of Electric Quantity based on Artificial Neural Network

Guorong Zhu¹, Yihong Zhang¹, Fei Cheng², Wei Wang², Yingfeng Zhou^{3,*}

¹State grid Zhejiang economic research institute, China

²Zhejiang Huayun Information Technology Co., Ltd, China

³School of Informatics and Electronics, Zhejiang Sci-Tech University, China

*Corresponding author e-mail: 2231884973@qq.com

Abstract. In this paper, BP neural network is used to predict electric quantity of electric power system. BP neural network model is established, input data is processed, hidden layer number is determined, while initial weight value is determined by selection of learning parameters and training samples for electricity prediction and so on. The electric quantity prediction of power system is very important for power network planning and improving the economic benefit of power department.

1. Introduction

Short-term load forecasting of power system is an essential task in power system operation and dispatching. It is related to the safe and economic operation of power system and the scientific rationality of power network management and dispatch. In this paper, several factors affecting the change of power load are studied, and the method of BP neural network is used to forecast the load. Short-term electricity forecasting based on BP neural network has three main parts: generating data sample set, determining network type and structure, and training and testing network.

2. Selection of Historical Data Samples

The first step of electricity forecasting is to select the historical data, as not all the historical data are appropriate. To get the ideal prediction effect, we must deal with the historical data properly, otherwise it will lead to the failure of the prediction.

The training samples of neural networks should not be too many or too few. Too many choices will increase the network training time greatly and even lead to network deterioration, while too little will make the network cannot get enough information and will not get the correct results. According to experience, the training sample is 5 to 10 times of the connection weight. When selecting the factors that influence the load change, we should first select the factors which have a great influence on the electricity quantity prediction, such as weather factors, and ignore the smaller ones. Finally, the sample data is divided into two categories: training data and test data.

3. Preprocessing of Input Data

Usually, the historical data we select will be defective because of various reasons, and there will often be bad data deviating from the true value in the data, which will have a negative impact on the accuracy



of the electricity quantity prediction. If all the data fluctuates are within $\pm 10\%$, the sample is considered eligible and the data beyond $\pm 10\%$ is bad data, which can be processed by the following methods:

(a) Horizontal treatment method

That is to select the data to be processed before and after two times as a benchmark, set a load fluctuation interval, if the data exceeds this interval, then the average method is used to process the data.

If

$$|Y(d, t) - Y(d, t - 1)| > \alpha(t) \quad (1)$$

$$|Y(d, t) - Y(d, t + 1)| > \beta(t) \quad (2)$$

Then

$$Y(d, t) = \frac{Y(d, t + 1) + Y(d, t - 1)}{2} \quad (3)$$

In the formula, $Y(d, t)$ is the load value at the time of day d t , $\alpha(t)$, $\beta(t)$ is the threshold value.

(b) Vertical Processing Method

Because of the periodicity of electric load, we think that the load value at the same time of different date is similar, and the difference between the two values has a certain range. If the difference value is outside this range, it should be corrected.

If

$$|Y(d, t) - m(t)| > \gamma(t) \quad (4)$$

Then

$$Y(d, t) = m(t) + \gamma(t) \quad Y(d, t) > m(t) \quad (5)$$

$$Y(d, t) = m(t) - \gamma(t) \quad Y(d, t) < m(t) \quad (6)$$

In the formula, $m(t)$ is the average load values in recent days in history, $\gamma(t)$ is the threshold value.

4. Quantification of Influencing Factors

When we use BP neural network to predict electric power, we usually choose sigmoid function as the excitation function of the neuron node. According to the characteristic of the curve, the input amount is too large or too small will enter the saturation region of the function. The output value does not change significantly. So before BP network training, we use the following formula A to calculate the data needed for prediction (0.1, 0.9), after completion of the prediction, we use the formula B to convert the load value.

$$a. y_i = 0.1 + (0.9 - 0.1) \frac{x_i - \min}{\max - \min} \quad i = 1, 2, \dots, P \quad (7)$$

$$b. x_i = \min + \frac{(y_i - 0.1)(\max - \min)}{0.9 - 0.1} \quad i = 1, 2, \dots, P \quad (8)$$

In the formula, x_i represents the original historical data value, y_i as values after normalization. “Min” is the minimum values before normalization for similar sample sets, max is the maximum value before normalization for a similar sample set.

The meteorological factors such as temperature, wind, humidity and sunshine make up the atmospheric environment on which living things depend. The reason why human body temperature is constant is, on the one hand, because of human regulation. On the other hand, it is the result of heat exchange between human beings and the atmospheric environment. Therefore, the human body is very sensitive to the change of external temperature, and the power load is most significantly affected by the temperature.

When considering the influence of temperature change on load, the temperature parameters can be chosen to measure. In a certain range, the influence of temperature on load is similar. In a certain temperature range, the change of temperature has little effect on coincidence. But when the temperature increases or decreases to a certain extent, it will have a great influence on load change. The piecewise quantization of temperature can be represented by figure 1.

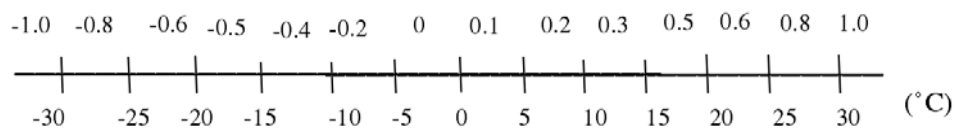


Figure 1. Temperature range and its corresponding values

5. Determination of the Node in the Input Layer and the Output Layer

The determination of neural network structure is the determination of input layer, hidden layer and output layer. The reasonable construction of the network can correctly simulate the law of load change, improve the efficiency of training and learning, and then improve the accuracy of prediction. Using artificial neural network to predict electric quantity is to use the nonlinear processing ability of data sample to determine the output in the case of historical load data input, so to determine the network structure must first determine the input and output nodes.

(a) Determination of the input layer nodes

The number of nodes in the input layer depends on the dimension of the data source and the dimension of the input feature vector. The number of choices has the most appropriate value, not as much as possible. The following characteristics should be taken into account in the selection process: independence, each feature used should be independent of each other, and reliability, for things of the same nature. There should be little difference in the size of their eigenvalues; distinctiveness; distinct differences in the eigenvalues of different classes of objects. The complexity of the pattern recognition system and the number of training network samples increase with the increase of the number of features. In order to improve the accuracy of electric quantity prediction, the load variation must be taken into account on the surface. All kinds of influencing factors, however, if all factors are taken into account together, the network structure is too complicated, the training time is greatly increased, and the prediction accuracy will be reduced because of the indeterminacy of the feature quantity. Therefore, the forecast work must influence the factor reasonable choice.

This paper selects the first year, the first two years, the first three years as the corresponding time load and the highest and lowest values of the corresponding temperature. The date type is the input of the forecast network, and the temperature and date type are quantified according to the method described in the 3.1 section, and the prediction model is a set of data samples.

(b) Determination of output layer nodes

Because we want to predict the electric quantity in the next 12 days, we can choose the multimodel single variable prediction method. We use the output nodes of 12 neural network models to correspond to each hour of the day. This method can avoid the phenomenon of over fitting when a single network is small. Because the single variable method is adopted in this paper, there is only one output node in

each output layer of the self-model. The output layer uses the linear activation function `purelin()`, which can produce output of any size.

6. Determination of Hidden Layer

For a multi-layer neural network, the number of hidden layers should be determined first. The feature of the hidden layer is the information that can be extracted from the input layer. From a theoretical point of view, the processing ability of the number of neural networks with hidden layers is proportional to the relationship. But in practice, too much hidden layer will greatly increase the training complexity of the network and the training needs. The number of samples, at the same time, will increase the training time and reduce the efficiency. According to the Kolmogorov theorem, a three layer neural network can achieve any mapping from m dimension to n -dimension. If a neural network contains a reasonable hidden layer of a neuron, it can solve the related problems well. Therefore, if a three-layer BP network has enough neurons, it can solve the general data processing problem. Therefore, this paper chooses to build a BP neural network model with an implicit layer, that is, a three-layer network structure.

The hidden layer has the function of extracting and storing the inherent rules from the sample, too many hidden layer nodes will appear excessive coincidence, reduce the generalization ability and increase the training time, the network will not be able to fully extract information if there are too few hidden layer nodes. It is difficult to generalize the inherent law of the sample. The number of hidden layer nodes is affected by the object of study, input layer nodes and output layer nodes. The exact number of hidden layer nodes is determined by the "trial and error method", which depends on the researchers' previous experience and repeated experiments. The samples are trained by the network with different number of hidden layer nodes. When the error is minimum or the number of learning times is minimum, the network is stable. That's the best number. The principle of minimum test error should be followed in determining the number of hidden layer nodes, and some empirical formulas for determining hidden layer nodes can be applied.

As empirical formula $n_1 = \sqrt{n + m} + a$ (m is the number of nodes in the output layer, n is the number of nodes in the input layer, n_1 is the number of hidden layers. A is the constant between 0 and 9) Given the minimum initial value of 3 and the maximum value of 12, the number of nodes is increased in turn, and the best performance value is selected as the hidden layer value of the electric quantity prediction model. The relation between the hidden layer node value and the mean error is shown in Table 1. It can be seen that the minimum error value is the node value of 9.

Table 1. Relationship between the average error and the number of hidden layer nodes in the network

Hidden layer node number	3	4	5	6	7	8	9	10	11	12
Average error	0.028	0.019	0.022	0.023	0.004	0.015	0.002	0.003	0.011	0.027

The hidden layer activation function uses the hyperbolic tangent function Sigmoid (), as shown in figure 2.

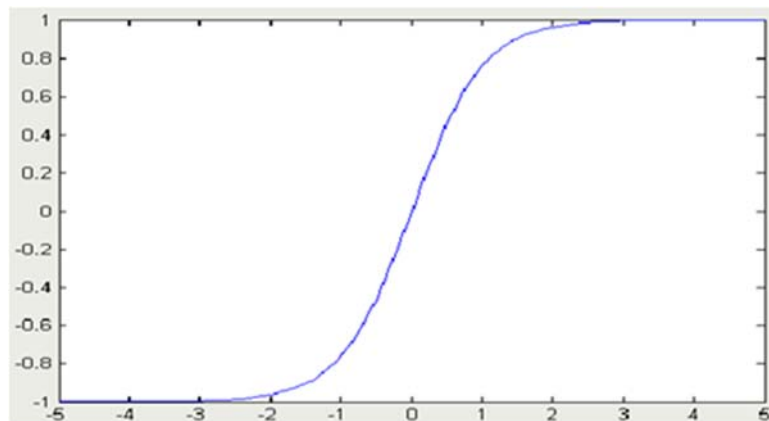


Figure 2. Hyperbolic tangent function

The parameters of the above network model are shown in Table 2.

Table 2. Network model parameters

Neural network unit layer		Node description
Input layer	1~3	Load values for the year preceding the forecast, the first two years, and the corresponding time for the first three years
	4~9	The highest and lowest temperatures in the year before the forecast time, the first two years, and the corresponding time period of the first three years
Hidden layer node	9	
output layer	Daily load value to be foretasted	

7. Determination of initial weights

Improper selection of initial weights can lead to a number of problems, as illustrated by the E-W curve in figure 3.

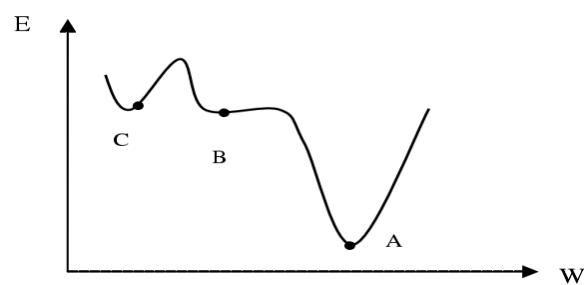


Figure 3. E-W diagram

(a) If the initial weight value is chosen near C point and there is not enough energy to jump out of C region, the local minimum solution will be obtained, but the global minimum solution will not be obtained.

(b) After entering the flat area, the training error remains unchanged for a long time without falling into a local minimum. But after a period of time, the error will continue to decrease from point B in the graph to point A, and in this process the error will continue to decrease. When $\frac{\partial E}{\partial \omega_i}$ reaches to 0, ω 's correction value is small. Then the network may not converge for a long time.

(c) The selection range of initial weights is related to the activation function. When the activation function is an exponential function, the value is generally in the interval (0 ~ 1), and when it is a tangent function, the weight should be in (-1, 1).

8. Determination of Learning Parameters

The learning parameters of BP neural network include momentum factor α and learning rate η .

(a) Determination of momentum factor

α Coefficient plays a regulating role in the training of neural network under the action of additional momentum factor. It can reduce the possibility of turbulence in network learning, avoid the network falling into local minima, and reduce its sensitivity to the details of error surface. Then the convergence is improved and the optimal solution is obtained. However, this method also has its obvious defects. It requires that the direction of error descent of the position of its own value on the error curve must be the same as the direction of motion of the minimum error value. The results show that the training effect of the network is better when the momentum factor is larger than the learning rate.

(b) Determination of learning rate

The learning rate determines the weight change of each cycle in network training. The learning rate should not be too large and the system will be unstable. Otherwise, the training time will increase and the convergence will slow down. Increasing the learning rate can reduce the number of training times, but to a certain extent, increasing the learning rate will not have any obvious effect. Decreasing the learning rate can avoid the network error falling into the local minimum error because it cannot jump out of the error surface trough. The learning rate should be chosen in the range of 0.011 / 0.8.

9. Summary

This paper introduces the work of using the BP network to predict the quantity of electricity. It lays the foundation for the modeling of BP network electricity forecasting from several aspects such as how to select the historical data, determine the topology of the network, and determine the network parameters and the processing data.

References

- [1] DU Song-Huai, HOU Zhi-Jian, Jiang Chuan-Wen. Grey forecasting price, station and its simulation [J]. Journal of Grey System, 2013, (1): 43 ~ 48.
- [2] Granger C W J. Combining forecasts-twenty years later [J]. Journal of Forecasting, 2009, 8 (3): 167 ~ 173.
- [3] Park D C, El-SharKawi M.A, Marks J, et al. Electric load forecasting using a neural net work [J]. IEEE Trans on Power Systems, 2011, 6(2): 442 ~ 449.
- [4] Chow Mo-Yuen, Zhu Jin-Xiang, Hahn Tram. Applications of fuzzy multi-objective decision making in spatial load forecasting [J]. IEEE Trans on Power Systems, 2008, 13 (3): 1185 ~ 1190.