A method for forecasting alpine area load based on artificial neural network model

Guoliang Wu¹, Guangxin Zu¹* and Jun Zheng¹

¹ Electric Power Research Institute, State Grid Heilongjiang Electric Power Company Ltd, Harbin, Heilongjiang Province, 150030, China

*Corresponding author: Guangxin Zu (zgxmy2@126.com)

Abstract. In order to realize the modernization of power grid management, mid-term power load forecasting has become an important research in power dispatching. Accurate mid-term load forecasting is of great significance to future power system dispatching and safe operation, and is a prerequisite for safe and economical operation of power systems. Based on error back propagation (BP) neural network has the advantages of generalization ability, self-learning ability and self-adapting ability, and simple local calculation, this paper proposes to use the BP neural network model to forecast the power load in the alpine region. Firstly, the historical load data and meteorological data of a certain area in Harbin are studied, analyzed and screened, and they are used as the input data of the network, a mid-term load forecasting model based on BP neural network is established, and then the model is simulated and improved through simulation to predict changes in the future development of electric load demand. The prediction data results and images show that the medium-term load prediction accuracy and training speed based on the BP neural network can reach the target, and the performance is good.

1. Introduction

As an important part of the planning and operation of the power system, Power load forecasting is a behavior that takes the regional historical data as the main reference and combines the analysis of various factors such as electricity price and economy to forecast the power load of the region in a certain time span in the future[1]-[2]. The forecast time of mid-term power load forecasting is generally forecasted in months, when performing short-term power load forecasting, factors such as temperature and humidity that have a greater impact on load changes will be considered, so that improving the accuracy of load forecasting. Accurate short-term power load forecasting results can ensure that power plant dispatching and power transmission make accurate decisions, at the same time, it can provide key data support for real-time dispatch, dispatch plan optimization, safe and stable operation of the power system, and finally achieve the goal of maximizing economic and social benefits [3].

BP neural network method has been extensively researched in literature [4]-[5]. In power load forecasting, the BP neural network algorithm because of its strong ability of autonomous learning, can be fully considered in load forecasting, such as temperature, humidity, electricity and other factors, just put the parameter input into the trained model can predict the future one day or days of the load situations, so it is one of mature used neural network[6]-[8]. Because BP neural network can realize the fitting of complex nonlinear structure, it is widely used in the load research of power system.
Aiming at the forecasting of power load changes in a certain place in Harbin, this paper proposes a method of load forecasting based on the BP neural network model. By selecting input and output data which have influence on load, the selected data are normalized and reverse-normalized, a BP neural network is built, and finally the load forecast for the area is realized through simulation, and the simulation forecast results basically meet the load trend.

2. BP neural network

2.1. Principle Analysis

BP neural network is an algorithm based on error calculation that is transferred in the opposite direction of the network calculation direction. It has excellent nonlinear fitting ability, and its network structure is similar to multi-layer sensors, which can be divided into three layers: input layer, hidden layer and output layer. Fig.1 shows the BP neural network neuron model[9].

The BP neural network is a "supervised" learning rule model. A virtual benchmark is used to supervise the training of input and output values, and multiple trainings are performed to reduce the error value between the expected response and the actual response, and the training is stopped until the target expectation is met. The learning rule model can be represented as shown in the Fig.2.

![Artificial neuron model](image1)

![Supervised learning network](image2)

2.2. BP neural network training process

2.2.1. Positive process

Assuming that the number of neurons in the hidden layer is q, the analysis of the forward information transfer process can be as follows. For a certain variable $x_i$, the information transmitted from the input layer to the $j_{th}$ neuron in the hidden layer is:

$$net_j = \sum_{i=1}^{n} w_{ij} x_i (i = 1, 2, \cdots, n)$$  (1)

In the above formula, $w_{ij}$ is the link weight of $x_i$ from the input layer to the $j_{th}$ neuron in the hidden layer. The output of the $j_{th}$ neuron of the hidden layer is:

$$Q_j = f (net_j + b_j)$$

$$= \frac{1}{1 + e^{- (net_j + b_j)}} (j = 1, 2, \cdots, q)$$  (2)

In the formula, $b_j$ represents the threshold value in the $j_{th}$ neuron; $f(\cdot)$ represents the activation function of the hidden layer neuron, using the S-shaped sigmoid function, its output is a continuous value between [0,1], which enables the network to realize any non-linear mapping from input to output.

The transfer information from the hidden layer to the $k_{th}$ neuron in the output layer is:
\[ net_k = \sum_{j=1}^{q} w_{jk} Q_j \quad (j = 1, 2, \cdots, q) \tag{3} \]

The output of the \( k \)th neuron in the output layer is:

\[ Q_k = f(net_k) \tag{4} \]

The formulas (1)-(4) are the approximate analysis of the forward transfer process of the BP neural network.

2.2.2 Reverse process

Within the allowable error range, we can set a value in advance. Compared with this value, if the error between the output result of the network output layer and the expected output is larger, the error will be transferred layer by layer through the hidden layer and returned to the input layer. At the same time, this result is used as a basis to modify the size of the weight and threshold, which is the process of error back propagation. Taking the \( p \)th training sample in the training sample space as an example, the output error \( E_p \) expression is:

\[ E_p = \frac{1}{2} \sum_{k=1}^{n} (d_{pk} - Q_{pk})^2 \tag{5} \]

For the entire training sample space, the total cumulative error of the BP neural network can be expressed as:

\[ E = \sum_{p=1}^{\rho} E_p = \frac{1}{2} \sum_{p=1}^{\rho} \sum_{k=1}^{n} (d_{pk} - Q_{pk})^2 \tag{6} \]

The sample data goes through the above two processes several times, the final network structure can only be obtained when the error is reduced to a certain range or the number of network training times reaches a certain limit.

3. The establishment and parameter setting of BP neural network

The choice of input parameters has a great influence on the establishment of a forecasting model and is the most basic content for load forecasting. When predicting the load of the micro-grid, if the data of historical load, meteorological factors that affect load changes, day types and other related factors are not processed as the input of the load forecasting model, it will have a great impact on the convergence speed and output results of the model \[10\]-[11], so it is necessary to analyze and process the initially collected data.

3.1. Parameter selection

3.1.1 Day type factor

It mainly includes working days, weekends and national legal holidays. A large number of load statistics show that the overall load changes regularly every 7 days, and the electricity load from Monday to Friday is higher than the electricity load on weekends and holidays. The main reason is that during weekends and holidays, factories and enterprises that usually consume a lot of electricity stop working, or work with a lower load.

3.1.2 Meteorological factors

Meteorological factors include humidity, rainfall, temperature, wind speed and so on, among which the influence of temperature on load changes is more prominent.

With the increase in temperature in summer and the decrease in temperature in winter, enterprises, large public places, homes and other large-scale use of air-conditioning and other temperature
regulating equipment have further increased the electricity load. Therefore, the electricity load in summer and winter is generally higher than that in spring and autumn. In order to simplify the model input, temperature, average humidity, etc. are selected as the model considerations in this paper.

3.2. Normalization and inverse normalization of data
Due to the input units of various types of data are different and the orders of magnitude vary greatly, in order to improve the convergence of the training network and reduce the inaccurate prediction results caused by different factors, the input features of various data contribute to the output as consistent as possible, at the same time, the convergence speed of the model should be accelerated, and the input data can be normalized. In short, it is to scale down the data so that it can be mapped to a smaller range, and at the same time eliminate some obviously wrong unqualified data. After the result is obtained, that is, after the training of the network is completed and the result is obtained, inverse normalization is required, that is, the output data needs to be inversely normalized according to the original normalized contrary method, so as to realize the restoration of the load forecasting data that meets the requirements.

3.2.1 Normalized processing of load data
The normalization of data application processing generally has the following several ways:
a. Normalization of maximum value method
   \[ y(\hat{n}, t) = \frac{y(n, t) - y_{\text{min}}}{y_{\text{max}} - y_{\text{min}}} \]  
   Where, \( y(\hat{n}, t) \) is the normalized data; \( y(n, t) \) is the original data of the sample; \( y_{\text{min}} \) is the minimum value in the original data; \( y_{\text{max}} \) is the maximum value in the original data.
b. Peak method normalization
   \[ y(\hat{n}, t) = \frac{y(n, t)}{\max(|y_i|)}, \quad i = 1, 2, \ldots, n \]  
   Where \( \max(|y_i|) \) is the maximum value among the absolute values of all data, and \( n \) is the total number of data among them.
c. Normalization by sum method.
   \[ y(\hat{n}, t) = \frac{y(n, t)}{\sum_{i=1}^{n} |y_i|} \]  
   Where \( \sum_{i=1}^{n} |y_i| \) is the sum of the absolute values of all data.

3.2.2 Normalization of temperature data
The load changes of the power grid are relatively complicated, and the temperature changes throughout the year are also relatively large. Therefore, in this short-term load forecasting study, the influence of temperature on load is emphasized.

In order to fully present the effect of temperature on load, this paper uses three data variables of maximum temperature, minimum temperature and average temperature as the main input.

3.2.3 Normalization of humidity
The normalization formula used in this paper:
   \[ H'_i = \frac{H_i - H_{\text{min}}}{H_{\text{max}} - H_{\text{min}}}, \quad i = 1, 2, \ldots, n \]  
   Where \( H_i \) is the humidity of the sample; \( H_{\text{max}} \) and \( H_{\text{min}} \) are the maximum and minimum values of all the humidity data.
In formula (13), $H'_i$ is the normalized humidity formula, $H_i$ is the original humidity data, $H_{\text{max}}$ is the maximum value in the original data, and $H_{\text{min}}$ is the minimum value in the original data.

4. Case Research

4.1 Modeling the mid-term load forecast of the power system

This paper uses the electricity consumption data of a certain area of Harbin City from January 1, 2012 to April 9, 2012 and various meteorological data of the urban power grid as input data. By preprocessing these data, select the appropriate load forecasting method, and establish a network model for load forecasting. The partial load meteorological data is shown in Table 1.

| Date       | Maximum temperature(℃) | Lowest temperature(℃) | Average temperature(℃) | Relative humidity | Rainfall (mm) | Demand load (10000KW h) |
|------------|-------------------------|------------------------|-------------------------|------------------|---------------|------------------------|
| 20120101   | 19.5                    | 12.1                   | 15.8                    | 63               | 0             | 30.6595                |
| 20120102   | 20                      | 13                     | 16                      | 59               | 0             | 43.3730                |
| 20120103   | 18.7                    | 14.2                   | 15.8                    | 72               | 0             | 54.6110                |
| 20120104   | 14.9                    | 9.3                    | 10.9                    | 62               | 0             | 58.1355                |
| 20120105   | 9.2                     | 5.1                    | 6.9                     | 78               | 2.9           | 59.1234                |
| 20120106   | 11                      | 6.3                    | 8.2                     | 92               | 3.3           | 58.5933                |
| 20120107   | 12.4                    | 9.4                    | 10.5                    | 80               | 0             | 56.9117                |
| 20120108   | 16.6                    | 9.9                    | 12.2                    | 75               | 0             | 51.6896                |
| 20120109   | 17.3                    | 9.4                    | 12.4                    | 75               | 0             | 54.8819                |
| 20120110   | 17.1                    | 9.4                    | 13.2                    | 73               | 0             | 55.1386                |

By setting parameters such as algorithm selection, training times, learning rate, error accuracy, etc., the BP network is trained to calculate the required load forecasting data. Through continuous adjustment of prediction methods and model parameters, combined with other factors, the accuracy of the algorithm is improved. The thought flow chart of the overall power load forecasting model is shown in Fig.3, and the program flow diagram of the implementation process is shown in Figure 4.

Figure 3. BP network prediction model.
4.2 Forecast Result
The power consumption of this area was simulated by MATLAB software. In this program, the value of each row of matrix A is the daily electricity consumption data for 25 consecutive days. Normalize the input value and self-adaptive network transformation to establish a BP neural network. The simulation results are as follows:

Figure 5. BP neural network prediction value and actual value comparison.

Figure 6. BP neural network power load forecast error.
Figure 5 is for the comparison between the actual value and the predicted value. Figure 6 is the error change curve between the actual value of the daily load and the predicted value. From Figure 5 and Figure 6, it can be seen that the BP neural network predicts load error values within single digits. Through the analysis of MATLAB simulation results and various related data feature maps, it can be seen that the BP neural network has good accuracy in regional short-term power load forecasting in the form of total daily power consumption. When the number of hidden layers is 11, the overall error is controlled within 14%.

5. Conclusion
This paper proposes a medium-term load forecasting method based on BP neural network. The effectiveness of the method is verified by predicting the load change of a certain place in Harbin. From the analysis of various result images, it can be seen that the load forecasting method based on BP neural network has excellent nonlinear fitting ability. The load forecast value output by this method is smoother than the actual value, although the daily error of the forecast method fluctuates greatly. However, the ideal accuracy range of medium-term load forecasting can be achieved in general, and we believe that the proposed method can be extended to other forecasting applications.

Each figure should have a brief caption describing it and, if necessary, a key to interpreting the various lines and symbols on the figure.

Acknowledgment
This paper is supported by load forecasting and key factor analysis of Harbin area using artificial intelligence, project number: 52243719000T

Reference
[1] Høverstad, B. A. Tidemann, A. Langseth, H. Özturk, P. (2015) Short-term load forecasting with seasonal decomposition using evolution for parameter tuning. IEEE Trans. Smart Grid, 6: 1904–1913.
[2] JIA, Q. (2020) Power load time series forecasting model based on probability statistics. Modern Electronics Technique, 43: 179-186.
[3] LI, C. ZHU, A. WANG, S. LIU, Y. ZHOU, D. YUN, Q. (2019) Short-term power load forecasting under power big data. Power System and Big Data, 22: 66–70.
[4] ZHANG, D. HU, G. LU, J. YIN, X. REN, Q. (2019) Short-term load forecasting based on GAD-BP neural network. Electronic Measurement Technology, 42: 143–147.
[5] WANG, K. ZHANG, R. (2019) Research on short-term power load forecasting method based on improved BP neural network. Electrical Measurement & Instrumentation, 56: 115–121.
[6] XU, J. HUANG, N. WANG, W. (2016) Hourly Solar Radiation Forecasting Based on GA-ELM Neural Network. Power System and Clean Energy, 32: 106-109.
[7] WANG, D. WANG, B. (2013) Medium- and long-term load forecasting based on variable weights buffer grey model. Power System. Technology, 37: 167–171.
[8] Borges, C. E. Penya,Y. K. Fernandez, I. (2013) Evaluating combined load forecasting in large power systems and smart grids. IEEE Trans. Ind. Informat, 9: 1570–1577.
[9] WANG, S. (2017) Micro-grid Short-term Load Forecasting Based on Data Mining and GSA-BP Multi-model Neural Network.
[10] ZHU, H. LI, L. XIONG, H. (2016) Power network load forecasting based on improved BP neural network. Modern Electronics Technique, 39: 64-66.
[11] WANG, J. Wang, J. Li, Y. (2014) Techniques of applying wavelet de-noising into a combined model for short-term load forecasting. International Journal of Electrical Power & Energy Systems, 62: 816-824.