A New BP Neural Network Model for the Prediction Problem of Equally Spaced Time Sequences and Its Application

Mengxia Li[1,2], Ruiquan Liao[2] , Yong Dong[3*]

ABSTRACT

For the prediction of equally spaced time sequences, this paper proposes a new construction method for training datasets based on the method which is used to determine parameters of the *ARIMA* model and builds a new BP neural network predictive model. For the actual data of annual power consumption in China from 1980 to 2016, data from 1980 to 2013 are chosen for this paper to construct the training dataset, and then the model proposed in this paper and the standard BP model are used to predict the power consumption from 2014 to 2016. Finally, after comparing the results obtained from the model proposed in this paper and the standard BP model, the prediction accuracy obtained by the model used here is found to be no more than 2%.

Key Words: BP Neural Network, Time window translation, *ARIMA* model

DOI Number: 10.14704/nq.2018.16.6.1585 **NeuroQuantology 2018; 16(6):28-32**

Introduction

Equally spaced time sequences widely exist in real life; examples of this are the annual grain yield sequence, the stock daily price sequence, the stock monthly price sequence, social power consumption over the years, and so on (National Bureau of Statistics of the People's Republic of China, 2012; 2013; 2014; 2015; 2016; 2017). The power consumption data of the whole society is an important economic indicator in the power market, and it reflects the total scale and the aggregate level of the power consumption over a period of time. What's more, it can show the power demand and the change rule in the whole. It is the main part of the power market analysis and prediction to predict society's power consumption, and the prediction results are related to power source construction, power system planning, the generation of power marketing, and so on. Currently, the prediction methods for power consumption mainly include the regression method, the moving average method, exponential smoothing, and so on. The common advantages of these methods are that they are simple and easy to operate, but these methods cannot guarantee satisfactory accuracy (Du *et al*., 2018; Yuan *et al*., 2018; Jin *et al*., 2018). The power market is a complex non-linear system. It is influenced by many factors, such as the macroeconomic environment, the development situation in all industries, the power consumption mode of people, and so on. There are not only qualitative factors but also quantitative factors; it is difficult to establish the definite function relation between the objectives. Moreover, because of the interrelated interaction of the factors, the prediction problems of social power consumption belong to nonlinear

Corresponding author: Yong Dong
Address: [1]School of Computer Science, Yangtze University, Jingzhou Hubei 434023, China; [2]The Branch of Key Laboratory of CNPC for Oil and Gas Production, Yangtze University, Wuhan Hubei 430100, China; [3]School of Information and Mathematics, Yangtze University, Jingzhou Hubei 434023, China
e-mail ✉ dongyong80@126.com
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Received: 3 March 2018**; Accepted:** 28 April 2018

mapping, and the general prediction methods therefore hardly provide good predictive results.

Artificial neural network methods have strong nonlinear approximation ability. Feedforward neural networks with a certain number of neuron nodes can approximate a continuous function with arbitrary accuracy. Hence, the neural network method is one of the most useful system modelling methods (Hu *et al.*, 2016; Isah *et al.*, 2017; Li, 2016; Wang *et al.*, 2016; Zhao *et al.*, 2006). In the research and applications of the neural networks, optimal setting connecting the weight and threshold value has been a valuable research topic.

According to the power consumption data of the whole society from 1980 to 2016 in China, this paper firstly builds a training dataset based on the time window translation method, and then determines the construction of a neural network based on the idea of the *ARIMA* model and the determination method of parameters. Finally, a BP neural network is built to predict the power consumption in the coming years. It will provide support for the analysis and control of the power market.

Neutral Network

Neural Network Model

The typical three-layer neural network model is given in Fig.1. In Fig. 1, Layer L1 is the input layer. Layer L2 and L3 are the hidden layers. +1 corresponds to the bias term. The critical parameters of the neural network include the number of neurons in the input layer and hidden layer neurons, the connection weight, the neural network threshold, activation function, and so on. If the connection weight is determined by the BP algorithm, then the network is called a BP neural network. BP neural networks are relatively mature in both network theory and performance. The distinctive advantage is that it has a strong nonlinear mapping capability and flexible network structure. The number of middle layers in the network and the number of neurons in each layer can be arbitrarily set according to the specific circumstances, and the performance will differ according to structure. However, the BP neural network has some drawbacks, such as slow learning speed, being trapped into the local minimum, and a lack of corresponding theoretical guidance for choosing the neuron number in each layer. In the existing research, the most studied contents are how to accelerate the convergence speed of networks and avoid the problem of easily getting trapped in the local minimum (Zhang *et al.*, 2007).

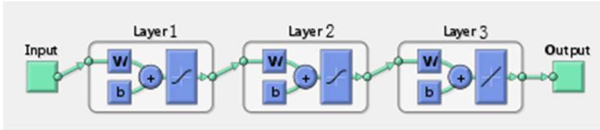


Figure1. Structure diagram of three-layer neural network model

Determination of the Number of Neurons in the Input Layer

This paper used the time window translation method to build the training dataset. The power consumption data of the whole society in China is denoted as $x=(x_1, x_2,..., x_{37})$. The training dataset is built based on 34 pieces of data. The principle of the time window translation is shown in Fig. 2.

$$x_1, x_2, \cdots, x_K \quad \rightarrow \quad x_{K+1}$$
$$x_2, x_3, \cdots, x_{K+1} \quad \rightarrow \quad x_{K+2}$$
$$\cdots$$
$$x_{32-K}, \cdots, x_{32}, x_{33} \quad \rightarrow \quad x_{34}$$

Figure 2. Principle of the time window translation

In Fig. 2, to the left of the arrow is the input data, and to the right of the arrow is the expected value of the predicted data, which is the actual value. Obviously, a key problem is to determine the appropriate K. It needs to be sure that there exists a close relationship between the power consumption in a given year and the one for several years before. The p and q values in the ARIMA model can be used to determine the value of K.

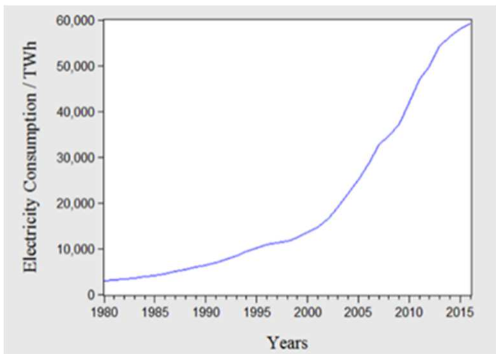


Figure 3. Time sequence diagram of power consumption

From Fig. 3, it can be seen that the power consumption data is not stable (James, 1994). Firstly, take the logarithm of the power consumption data which is denoted as logx and take the first-order difference which is denoted as *D*logx. *D*logx is shown in Fig. 4.

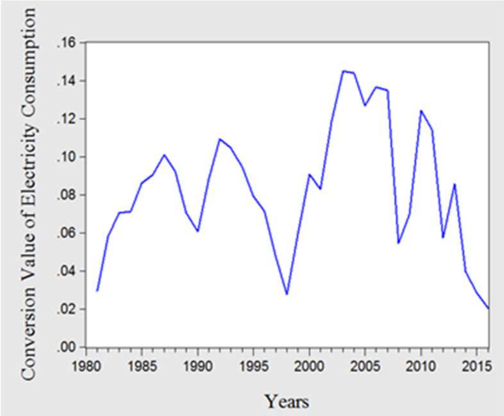
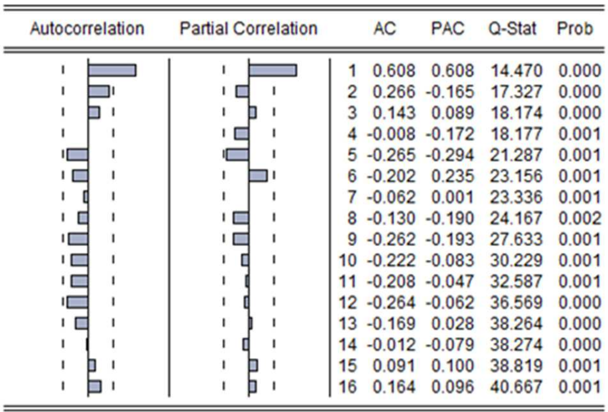


Figure 4. Time sequence diagram of *D*logx

It is easy to see that fluctuates around its mean value and doesn't represent the tendency. The autocorrelogram and the partial correlogram are shown in Fig. 5 (Ronald, 2016; Liu *et al.*, 2001).



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.608	0.608	14.470	0.000
		2	0.266	-0.165	17.327	0.000
		3	0.143	0.089	18.174	0.000
		4	-0.008	-0.172	18.177	0.001
		5	-0.265	-0.294	21.287	0.001
		6	-0.202	0.235	23.156	0.001
		7	-0.062	0.001	23.336	0.001
		8	-0.130	-0.190	24.167	0.002
		9	-0.262	-0.193	27.633	0.001
		10	-0.222	-0.083	30.229	0.001
		11	-0.208	-0.047	32.587	0.001
		12	-0.264	-0.062	36.569	0.000
		13	-0.169	0.028	38.264	0.000
		14	-0.012	-0.079	38.274	0.000
		15	0.091	0.100	38.819	0.001
		16	0.164	0.096	40.667	0.001

Figure 5. Autocorrelation and partial correlation of *D*logx

It can be seen that the autocorrelation coefficient and the partial correlation coefficient are both located along the dotted line, which means double of the standard deviation after first-order delay and shows the trailing characteristics. Hence, the *ARIMA*(1, 1) model can be built to reflect the characteristics of data. The order of the *ARIMA*(1, 1) model shows that some component has an important correlation to the previous three adjacent components. Therefore, the input layer of the neural network includes three neurons.

Determination of the number of hidden layer neurons

Given are the following assumptions: The number of input layer neurons is n_1. The number of hidden layer neurons is n_2. The output layer only includes one neuron. Then there are $(n_1+1)n_2$ connection weights and the total number of neuron bias is n_2+1. The number of parameters which needs to be determined is $(n_1+2)n_2+1$.

This paper used the time window translation method to obtain 34-3=31 training data. Then $(n_1+2)n_2+1 \leq 31$ holds. Thus $n_2 \leq 6$. $n_2=4$, 5, 6 might as well be applied to the model. The appropriate number of hidden layer neurons is determined by comparing the prediction accuracy.

BP Neural Network and Analysis of Predictive Results

Construction of BP Neural Network

The BP neural network is composed of an input layer, hidden layer, output layer, and neurons between each layer that are connected together. A three-layer BP neural network consists of an input layer, hidden layer, and output layer. The number of neurons of each layer is n_1, n_2, and n_3. It forms the so-called n_1-n_2-n_3 structure. The connection weight between the input layer and the hidden layer is ω_{1kj}, and the one between the hidden layer and the input layer is ω_{2kj}. The connection weight represents the connection strength between neurons.

30

Training Method for BP Neural Network

It is the premise for solving practical problems by using ANN to train the network. The aim is to evaluate the connection weight between networks from the known samples. For the trained BP networks, if the input data is not from the training sample, the BP neural network uses a set of connection weights obtained by the training process to calculate the corresponding response output. Thus, solutions to the practical problems are obtained. The training method of the BP neural network is the BP algorithm, which belongs to the supervised algorithm. The main idea is to use the BP neural network to calculate the actual outputs for a set of input samples. The error between the actual outputs of the BP neural network and output samples is used to modify the connection weights of the network until the error satisfies the set value. It is called the fitting error given by the sum of the square of errors which is $E=\frac{1}{2}\sum_{k=1}^{n_3}(t_k-z_k)^2$. t_k is the sample output and z_k is the actual output.

The BP algorithm which is used to train the BP neural network is given as follows:

Step 1 Initialize the connection weights. At the beginning, the connection weight is unknown. Generally, small random numbers are taken as the initial values of the connection weights of each layer.

Step 2 Calculate the outputs of each layer.

$$y_j = f_1\left(\sum_{i=0}^{n_1} \omega_{1ji} x_i + b_i\right) \qquad (j=1,2,\cdots,n_2) \tag{1}$$

$$z_k = f_2\left(\sum_{i=0}^{n_1} \omega_{2kj} y_j + b_j\right) \qquad (k=1,2,\cdots,n_3) \tag{2}$$

f_1 and f_2 are activation functions which are sigmoid functions or linear functions.

Step 3 Modify the connection weights. The gradient descent method is used to modify the connection weights. The modified values and the error functions are in proportion. The modifications of the connection weights are shown in Eq. (3) and Eq. (4).

$$\Delta\omega_{2kj} = -\eta \frac{\partial E}{\partial \omega_{2kj}} = \eta(t_k - z_k) f_{2'} y_j \qquad (k=1,2,\cdots,n_3;\ j=1,2,\cdots,n_2) \tag{3}$$

$$\Delta\omega_{1ji} = -\eta \frac{\partial L}{\partial \omega_{1ji}} \eta \sum_{i=1}^{n_3} (t_k - z_k) f_{2'} \omega_{2kj} f_{1'} x_i \qquad (j=1,2,\cdots,n_2;\ i=0,1,2,\cdots,n_1) \tag{4}$$

In Eq. (3) and Eq. (4), η is the learning rate, and f_1 and f_2 are the derivative of the activation functions.

Add the initial weights and the corresponding adjustment quantity, and then calculate the new weights. Repeat this process until the output of the sum of the square of errors of the output layer obtains the set value.

According to the conclusion of section 1.2, the training dataset is (x_i, x_{i+1}, x_{i+2}, x_{i+3}), i=1, 2,..., 31. The first three components are the input data, and the fourth one is the corresponding output data. For *NN*(3, 3, 1), *NN*(3, 4, 1), *NN*(3, 5, 1), *NN*(3, 6, 1), the weights and bias values are respectively determined by the training dataset and training neural network. The annual power consumption data from 2011 to 2013 are taken as the input data and used to predict the power consumption data in 2014. The annual power consumption data from 2012 to 2013 and the predictive power consumption data in 2014 are taken as the input data, and then the power consumption data in 2015 are evaluated through the trained neural network. Finally, the annual power consumption data in 2013 and the predictive power consumption data in 2014 and 2015 are taken as input data, and the power consumption data in 2016 are predicted by the neural network. Tab. 1 to Tab. 4 show the prediction results of neural networks with different structures and the actual power consumption data.

Table 1. Predict results of *NN*(3, 3, 1) neural network

Time	Actual power consumption (TWh)	Predictive power consumption (TWh)	Absolute error (%)	Relative error (%)
2014	56383.7	57170	786.3	1.4
2015	58020	57214	806	1.4
2016	59198	57238	1960	3.3
mean				2.0

Table 2. Predict results of *NN*(3, 4, 1) neural network

Time	Actual power consumption (TWh)	Predictive power consumption (TWh)	Absolute error (%)	Relative error (%)
2014	56383.7	55256	1127.7	2.0
2015	58020	55239	2781	4.8
2016	59198	55273	3925	6.6
mean				4.5

Table 3. Predict results of *NN*(3, 5, 1) neural network

Time	Actual power consumption (TWh)	Predictive power consumption (TWh)	Absolute error (%)	Relative error (%)
2014	56383.7	53960	2423.7	4.3
2015	58020	56236	1784	3.1
2016	59198	145680	86482	146.1
mean				51.0

Table 4. Predict results of *NN*(3, 6, 1) neural network

Time	Actual power consumption (TWh)	Predictive power consumption (TWh)	Absolute error (%)	Relative error (%)
2014	56383.7	101250	44866.3	79.6
2015	58020	629610	571590	985.2
2016	59198	691270	632072	1067.7
mean				710.8

Comparing the relative errors between the predictive results and the actual results, on the whole, the predictive effect of *NN*(3, 3, 1) is the best, followed by *NN*(3, 4, 1). The relative errors of *NN*(3, 5, 1)and *NN*(3, 6, 1) are quite large.

From the perspective of the predictive time span, the smaller the time span is, the higher the predictive accuracy is. Therefore, the *NN*(3, 3, 1) prediction model has been chosen in this paper. Fig. 6 shows that the *NN*(3, 3, 1) prediction model has a good predictive effect.

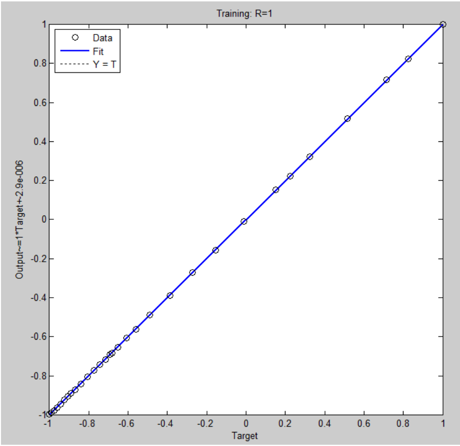


Figure 6. Predictive effect of $NN(3,3,1)$

The data shown in Figure 6 are converted from the original interval to the interval [-1, 1].

Conclusions

The model proposed in this paper is based on the time sequence itself and doesn't use other data. Therefore, a method of training data by a neural network built by time sequences is proposed in this paper. A comparison between the predictive results and actual results shows that the model proposed in this paper is effective. The training results of the neural network indicate that the BP algorithm converges slowly and costs a lot of computational time. When the training process is repeated, the training results show some differences.

Acknowledgement

The authors will thank people in the Branch of Key Laboratory of CNPC for Oil and Gas Production for their great help. This paper is supported by National Natural Science Foundation of China (61572084 and 51504038), National Key S&T Special Projects (2016ZX05056004-002, 2016ZX05046004-003 and 2017ZX05030-005), and China Petroleum Innovation Fund Project (2015D-5006-0208). This paper is also supported by Educational Commission of Hubei Province of China (Q20171302).

References

Du Y, Ma RY. Application of Different Improved ARIMA Models in the Prediction of Hydrological Time Series. Water Power 2018; (04):12-14+28.

Hu LJ, Tang L, Pan Q, Song H, Wen PG. Research and analysis of PI control strategy based on neural network in power grid. Mathematical Modelling of Engineering Problems 2016; 3(1): 25-28.

Isah OR, Usman AD, Tekanyi AMS. A hybrid model of PSO algorithm and artificial neural network for automatic follicle classification. International Journal Bioautomation 2017; 21(1): 43-58.

James D. Hamilton. Time Series Analysis. Princeton NJ: Princeton University Press. 1994.

Jin SH, Zhu X, Hao WG, Wang B, Mei F, Li YJ, Yuan XL and Liu HM. Forecasting of bus-bar net load based on PSR. ARIMA and SVR. Power Demand Side Management 2018; 20(2):20-24.

Li J. Application of BP neural network algorithm in biomedical diagnostic analysis, International Journal Bioautomation 2016; 20(3): 417-26

Liu W and Chen Z. Modeling method and operation skills of econometric software EViews6.0. Beijing: China Machine Press, 2011.

National Bureau of Statistics of the People's Republic of China. 2012 China Statistical Yearbook. Beijing: China Statistics Press, 2012.

National Bureau of Statistics of the People's Republic of China. People's Republic of China 2017 national economic and social development statistical bulletin. Beijing: China Statistics Press, 2018.

National Bureau of Statistics of the People's Republic of China.2013 China Statistical Yearbook. Beijing: China Statistics Press, 2013.

National Bureau of Statistics of the People's Republic of China.2014 China Statistical Yearbook. Beijing: China Statistics Press, 2014.

National Bureau of Statistics of the People's Republic of China.2015 China Statistical Yearbook. Beijing: China Statistics Press, 2015.

National Bureau of Statistics of the People's Republic of China.2016 China Statistical Yearbook. Beijing: China Statistics Press, 2016.

National Bureau of Statistics of the People's Republic of China.2017 China Statistical Yearbook. Beijing: China Statistics Press, 2017.

Ronald D. Yockey. SPSS Demystified: A Step by Step Approach (2nd Edition), New York: Taylor & Francis Group. 2016.

Wang TC, Xie YZ, Yan H. Research of multi sensor information fusion technology based on extension neural network. Mathematical Modelling of Engineering Problems 2016; 3(3): 129-34.

Yuan HL, Hua XH, Gong GD and Ding LH. Analysis and Prediction of build Settlement Deformation Based on Wavelet and Time Series Combination Model. Journal of Geomatics 2018; 43(2), 73-75.

Zhang HJ and Sun ZQ. Modern Detection Technology. Beijing: Chemical Industry Press, 2007.

Zhao XR, Peng XY and Jiang YG. China Northeast Power Network Short-term Load Forecasting Based on Neural Network. Journal of System Simulation 2006; 18(06): 1677-80.