Improved BP neural network for forecasting industrial electricity consumption in China

Xingping Zhang, Rui Gu

***Abstract*—**An improved BP Neural Network with additional momentum and adaptive learning is proposed in the paper to predict the growth rate of industrial electricity consumption. Firstly the industry is classified into Ⅰ, Ⅱ and Ⅲ sectors based on the developmental economics. In the BP neural model, the growth rate of current year industrial value-added output, and the growing ratios of Ⅰ, Ⅱ and Ⅲ sector electricity consumption are taken as input variables, and the output is the growth rate of next year industry electricity consumption. Matlab7 is used as modeling tool to design the model. The simulation result is compared with that of traditional BP Neural Network model, which show the feasibility of the model proposed in the paper. Furthermore, superiority of improved BP Neural Network is validated.

***Keywords*—**Electricity, Forecasting Model, Industry Structure, Neural Network

I. INTRODUCTION

THE industry electricity consumption is a complex phenomenon with multiple influencing factors such as macro economic environment, subsector structure, state of industry development and electricity using fashion [1]. Because of the difficulty to establish definitive function between industry electricity consumption and these factors, the regular prediction methods usually do not fare well.

Artificial neural network (ANN) is a kind of nonlinear system that simulates the structure, character and function of human brain [2]-[3]. It is of the character of self-organizing, selfadaptive, selflearning and defaulting, as well as of the merit of strongly nonlinear output-input mapping and prone to learn and train, so it behaves with particular function when dealing with intercrossing variables. It is why ANN is often used in prediction of complex system.

In developmental economics, the neoclassification of industry sector can explain the character of industry growth in a more scientific fashion and provide a new insight into understanding industry structure classification. Therefore in the paper an improved BP ANN model for predicting industry electricity consumption based on neoclassification of industry sector is proposed.

Manuscript received February 4, 2007. Revised Manuscript received April 29, 2007.

X. Z. is with the Business and Administration School, North China Electric Power University, Beijing City, 102206, China. (phone: 86-10-80795308; e-mail: zhangxingping302@ 163.com).

R. G. is with the Business and Administration School, North China Electric Power University, Beijing City, 102206, China. (e-mail: gurui@ncepu.edu.cn).

II. NEO-CLASSIFICATION OF INDUSTRY SECTOR IN DEVELOPMENTAL ECONOMICS

According to Chenery's classification [4] to industry sector into nonage, metaphase and anaphase ones, in the paper the industry sector is divided into three kinds of catalogue, among which, the Ⅰcatalogue mainly includes the sectors highly correlated with daily lives, such as food, beverage, tobacco, textile, wood and furniture, clothes and fibers. Catalogue Ⅱ is mainly heavy industry including petroleum refining and coking, chemical, chemical fiber, rubber, black metal, color metal and metalwork. Catalogue Ⅲ mainly includes medicine, ordinary machinery, specific equipment, transportation equipment, electric machinery, electronic instrument and appearance, office facility and other equipment. According to the corresponding relation provided in [1] for economic statistics and electricity statistics, the electricity consumption for the three catalogues is collected.

The widely accepted view is that China's spontaneous industrialization process began from the urban economic reform in 90th. Therefore it is of importance to analyze the industry electricity consumption since then for judging the trend in the future.

III. BP ANN AND ITS IMPROVEMENT

A. Principle of BP ANN

BP ANN is a kind of multilayer forward propagation NN based on the inverse error propagation algorithm. The BP model is consisted of one input layer, some hidden layers and one output layer, with every layer consisting of some neurons, connection of the different layer neurons by connective weight and threshold and nonexistence of connection among same layer neurons. The sketch of standard BP model is shown in figure 1.

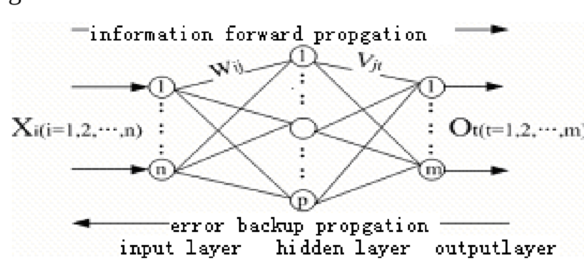

Fig.1 Structure of three-layer BP Neural Network

B. Standard BP ANN Learning Algorithm

The standard BP learning algorithm [4]-[5] is to modify the network connective weights and thresholds to make the error function descending along negative grad direction.

The k^{th} input vector in the network is $X_k = (x_1^k, x_2^k, \cdots, x_n^k)$, and the k^{th} output vector is $Y_k = (y_1^k, y_2^k, \cdots, y_m^k)$. The k^{th} hidden layer neuron input is $S_k = (s_1^k, s_2^k, \cdots, s_p^k)$, and the output $B_k = (b_1^k, b_2^k, \cdots, b_p^k)$. The input vector of k^{th} neuron in the output layer is $L_k = (l_1^k, l_2^k, \cdots, l_m^k)$, and the output vector is $O_k = (o_1^k, o_2^k, \cdots, o_m^k)$. The connective weight in the hidden layer is $w_{ij}(i = 1,2,\cdots,n; j = 1,2,\cdots,p)$, and that in output layer is $v_{jt}(j = 1,2,\cdots,p; t = 1,2,\cdots,m)$. The threshold value of the hidden layer neurons is $\theta_j(j = 1,2,\cdots,p)$, and that in output layer is $\lambda_t(t = 1,2,\cdots,m)$. Among which, $k = 1,2,\cdots,h$ is parameter, and h is the number of training sample.

Step 1 We give initial value for every connective weight in the model, and set up the maximum permitted error ε and learning rate α (0< α <1).

Step 2 We randomly choose a set of input, objective output sample, and $Y_k = (y_1^k, y_2^k, \cdots, y_m^k)$ for the NN.

Step 3 We calculate the output of hidden layer neuron b_j^k according to hidden layer stimulating function, which generally is sigmoid function as:

$$s_j^k = \sum_{i=1}^{n} w_{ij} x_i^k - \theta_j \tag{1}$$

$$b_j^k = f(s_j^k) \tag{2}$$

where $j = 1,2,\cdots,p$.

Step 4 We calculate the output of hidden layer neuron o_t^k according to hidden layer output function as:

$$l_t^k = \sum_{j=1}^{m} v_{jt} b_j^k - \gamma_t \tag{3}$$

$$o_t^k = f(l_t^k) \tag{4}$$

where $t = 1,2,\cdots,m$.

Step 5 We randomly provide next sample with the NN and return to step3 until all the h samples are finished.

Step 6 We calculate the overall error E of NN by using (5). If $E \le \varepsilon$, the learning is stopped, else return to next step.

$$E = \frac{1}{2}\sum_{k=1}^{h}\sum_{t=1}^{m}\left(y_t^k - o_t^k\right)^2 \tag{5}$$

Step 7 We calculate the generalized error of every neuron in the output layer as:

$$d_t^k = (y_t^k - o_t^k) \times o_t^k \times (1 - o_t^k) \tag{6}$$

where $t = 1,2,\cdots,m$.

Step 8 We calculate the generalized error of every neuron in the hidden layer as:

$$e_j^k = \left[\sum_{t=1}^{m} d_t^k \times v_{jt}\right] \times b_j^k \times (1 - b_j^k) \tag{7}$$

where $j = 1,2,\cdots,p$.

Step 9 According to (2) and (6), we modify the connective weights and threshold values for the output layer as follows:

$$_{\Delta} v_{jt}(N) = \alpha \times \sum_{k=1}^{h}\left(d_t^k \times b_j^k\right) \tag{8}$$

$$_{\Delta} \gamma_t(N) = \alpha \times \sum_{k=1}^{h} d_t^k \tag{9}$$

$$v_{jt}(N+1) = v_{jt}(N) + {}_{\Delta} v_{jt}(N) \tag{10}$$

$$\gamma_t(N+1) = \gamma_t(N) + {}_{\Delta} \gamma_t(N) \tag{11}$$

where $t = 1,2,\cdots,m$; $j = 1,2,\cdots,p$.

Step 10 According to (7), we modify the connective weights and threshold values of hidden layer as follows:

$$_{\Delta} w_{ij}(N) = \alpha \times \sum_{k=1}^{h}\left(e_j^k \times x_i^k\right) \tag{12}$$

$$_{\Delta} \theta_j(N) = \alpha \times \sum_{k=1}^{h} e_j^k \tag{13}$$

$$w_{ij}(N+1) = w_{ij}(N) + {}_{\Delta} w_{ij}(N) \tag{14}$$

$$\theta_j(N+1) = \theta_j(N) + {}_{\Delta} \theta_j(N) \tag{15}$$

where $i = 1,2,\cdots,n$; $j = 1,2,\cdots,p$.

Step 11 We resample a set of input and objective output vector from h samples and return to step3 for new round of learning.

C. Improvement to Standard BP Algorithm

(1) Deficiency of standard BP algorithm

Firstly, because the curve of error function is multidimensional complex surface, there exists local minimum point and the movement to multidirection will induce increase of error. So the BP algorithm is always unable to reach global minimum for absorption of local minimum. Secondly, the learning rate is presumed to be fixed. However, too big the learning rate will induce the instability of the learning process and too little will induce longevity of the training time. The usual improvement is to introduce additional momentum and adaptive learning [6]-[7].

(2) Additive momentum

Additive momentum is to consider not only the grad direction but also the moving trend in modification of NN weight. The details of the method is consider the last time change of weight and threshold value in modifying the current weight and threshold value, and recalculate the weight and threshold value according to error BP algorithm. Taken (10) and (11) for example, the calculation [8]-[9] is:

$$v_{jt}(N+1)=v_{jt}(N)+(1-mc)\times{}_{\Delta}v_{jt}(N)+mc\times{}_{\Delta}v_{jt}(N-1)\ ;$$
$$\gamma_t(N+1)=\gamma_t(N)+(1-mc)\times{}_{\Delta}\gamma_t(N)+mc\times{}_{\Delta}\gamma_t(N-1)\ ; \qquad (16)$$

where mc is momentum factor, and $0\leq mc<1$; while ${}_{\Delta}v_{jt}(N-1)$ and ${}_{\Delta}\gamma_t(N-1)$ is the last time change in weight and threshold value.

When mc is equal to one, the modification of weight and threshold value is equal to the last time change and when it is zero, the change is only determined by current change. So the essence of additive momentum is to lower the sensitivity of the NN and decrease the vibration of learning process by the addition of damp term.

(3) Adaptive learning

The idea of adaptive learning is to accelerate convergence by modifying learning rate according to the change in grad of error surface. The calculation process is: to check whether the modifications of connective weight and threshold value really lower the error, if so add to the learning rate, else, decrease it. The modification equation [4] is as:

$$\begin{cases}\alpha(N+1)=1.05\alpha(N)\ ,\ E(N+1)<E(N)\\ \alpha(N+1)=0.7\alpha(N)\ ,\ E(N+1)>1.04E(N)\\ \alpha(N+1)=\alpha(N)\ ,\ else\end{cases} \qquad (17)$$

IV. Improved BP ANN Prediction Model based on Neo-Classification of Industry Structure

Electricity is tightly correlated with the development of economy. On the other hand, power consumption haves inertia of the correlation in next two years because of the existence of inertia in economy. So in the paper to describing the correlation of current year industry electricity consumption and current industry output, the total industry power consumption is classified into Ⅰ, Ⅱ and Ⅲ sector consumption according the developmental economics classification catalogue. The input variables of the model are current year industry output growth rate, growing ratio of Ⅰ, Ⅱ and Ⅲ sector power consumption and the output variable is next year total industry electricity consumption growing rate.

Existing theory has proved that a three layer BP ANN can approach any kinds of nonlinear relationship in any precision, so in the paper a three layer improved 4×p×1 BP model is used for prediction. In the model, the number of neurons in input layer is four, p in the hidden layer and is one in output layer. For determining the number of neurons in hidden layer there is no proved theory but an empirical method as $2N+1$ by Komogorol theory, where N is the number of input layer neurons. In the paper the number of neurons in hidden layer is preliminarily determined as five. The stimulating function in both the hidden and output layer of the NN is Sigmoid function as $f(x)=1/(1+e^{-x})$.

V. Case Study

Matlab7.0 is used in the paper for model implementation.

A. Sampling Data *and Training of ANN*

The , and sector electricity consumption growing ratio for 1991-2005 are calculated (Table Ⅰ), together with industry output in value-added are taken as input variables in the prediction model.

TABLE I
GROWTHS OF INDUSTRIAL PRODUCTION VALUE AND POWER CONSUMPTION OF THREE KINDS INDUSTRY IN 1991 - 2005

Year	Elec. con. Growth rate of Ⅰ	Elec. con. Growth rate of Ⅱ	Elec. con. Growth rate of Ⅲ	Growth rate of industry output	Growth rate of industry Elec. Con.
1991	0.0787	0.0829	0.0750	0.1792	0.0808
1992	0.0836	0.1057	0.1110	0.2717	0.1034
1993	0.0554	0.0974	0.1161	0.3796	0.0941
1994	0.0510	0.0881	0.1015	0.3730	0.0850
1995	0.0693	0.0831	0.1076	0.2689	0.0843
1996	0.0345	0.0638	0.0486	0.1766	0.0585
1997	0.0199	0.0318	0.0170	0.1145	0.0287
1998	-0.0051	0.0144	0.0048	0.0301	0.0111
1999	0.0715	0.0621	0.0896	0.0509	0.0663
2000	0.1291	0.1010	0.1642	0.1129	0.1116
2001	0 0834	0 0843	0 1037	0 0852	0 0866

The sample data is input in the model to training it. The number of hidden layer neurons is set as 11, maximum total error 0.00001 and maximum training times 10000. By repeated experiment, finally the number of hidden layer neurons is determined as 9. The errors curve is depicted as in figure 2 and the total error at the time is 0.000009536.

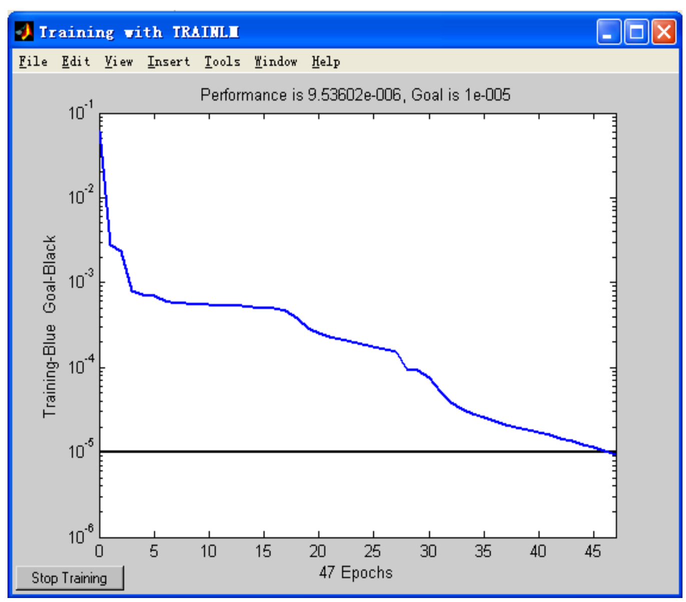

Fig. 2 The training error curve of improved BP Neural Network

Figure 3 is the training errors curve of standard BP model with same other condition. By comparison it is obvious that improved BP model is better.

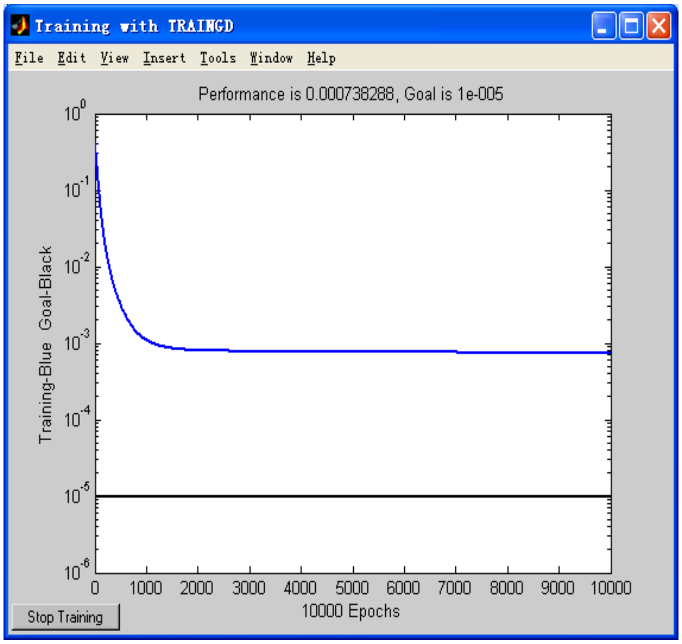

Fig. 3 Training error curve of standard BP Neural Network

B. *Test of Training Results*

All the sampling data are tested respectively by standard BP and improved BP training results and the test results are listed in table Ⅱ, which is shown that, as compared with standard BP model, the prediction of improved BP model is better in less error.

TABLE Ⅱ
COMPARISON OF FORECASTING RESULTS OF TWO MODELS

Year	Actual	Improved BP ANN		Standard BP ANN	
		Forecast	Error	Forecast	Error
1992	0.1034	0.1012	-0.0022	0.0832	-0.0202
1993	0.0941	0.0992	0.0051	0.0987	0.0046
1994	0.0850	0.0877	0.0027	0.1040	0.0190
1995	0.0843	0.0842	-0.0001	0.0985	0.0142
1996	0.0585	0.0545	-0.004	0.0970	0.0385
1997	0.0287	0.0296	0.0009	0.0794	0.0507
1998	0.0111	0.0171	0.0060	0.0650	0.0539
1999	0.0663	0.0667	0.0004	0.0594	-0.0069
2000	0.1116	0.1103	-0.0013	0.0831	-0.0285
2001	0.0866	0.0894	0.0028	0.1063	0.0197
2002	0.1243	0.1276	0.0033	0.0910	-0.0333
2003	0.1667	0.1689	0.0022	0.1175	-0.0492
2004	0.1687	0.1701	0.0014	0.1368	-0.0319
2005	0.1492	0.1555	0.0063	0.1358	-0.0134

C. *Prediction of Industry Electricity Consumption Growth*

When the ANN model is trained, the data of 2005 industry output growth rate, growing ratio of Ⅰ, Ⅱ and Ⅲ sector power consumption are taken as input of the improved BP ANN the 2006 industry electricity consumption growth is predicted as 14.1%. According to the relation between industry electricity consumption and total society electricity consumption, the growth rate of total consumption should be among 13.5-14%.

VI. CONCLUSION

The industry electricity consumption is complex with multiple influencing factors so regular prediction model usually don't fare well for its prediction.

Theory in developmental economy on neo-classification of industry catalogue provides new insight into the prediction and explanation of industry electricity consumption growth.

Improved BP ANN model fares better than traditional methods and is validated by simulation study.

The method proposed in the paper is of reference for future study and in the future more statistical data sample will be collected for practical application.

REFERENCES

[1] Yuan Jia-hai, "Research on power economic relationship: theoretical methods, empirical analysis and forecast application," Dr. thesis, North China Electric Power University, China, 2006.
[2] Ye Zhong-wei, Shi Zhong-zhi, Principle of Artificial Neuron Network (Book style). Beijing: Machinery Publication House, China, 2004, pp. 98-102.
[3] Dai Kui, *Design of Principle of Artificial Nneuron Network* (Book style). Beijing: Machinery Publication House, China, 2002, pp.58-61.
[4] Chenery, *The Comparative Study of Industrialization and Development* (Book style). People Publication House of Shanghai, China, 1996, pp. 125-132.
[5] Wang Zheng-wu, Zhang Liu Rui-ping, "One improved BP Neural Network learning algorithm," *Mathematical Theory And Applications*, Vol. 25, no.1, pp. 31-34, 2005.
[6] Parlos A G, "An accelerated learning algorithm for multiplayer perception networks," *IEEE Trans on Neural Networks*, Vol., no. 3, pp. 86-8, 1994.
[7] XU Zhong-hua, Yang Wei-feng, Jiang Wei-jin, "Improvement and application of the learning algorithm of BP Neural Networks," *Computer Engineering &Science*, Vol. 26, no. 12, pp. 62-65, 2004.
[8] Sheng Wanxing, "A recognition frame of intelligent control and realization, " *Proceedings of the CSEE*, Vol.21, no.1, pp. 48-51, 2001.
[9] Hu Zhao-guang, Fang Yan-pin, "Intelligent engineering and load forecast, " *Power System Technology*, Vol. 23, no. 5, pp. 15-18, 1999.

Xingping Zhang was born in ShanXi province, China. B.Sc. in Management, Shanxi Finance and Economics University, Taiyuan City, 1993. M.Sc. in Management, China University of Mining and Technology, Beijing City, 1997. Dr. in Management, China University of Mining and Technology, Beijing City, 2001. Postdoctoral Fellow, Liaoning Engineering and Technology University, Fuxin City, 2004.

He is Full Professor at School of Business and Administration, North China Electric Power University, Beijing City. Research interests include Energy Economics and Electric Power Market.

Dr. Zhang is member of the Energy Economic Institute of Beijing.

Rui Gu was born in Beijing, China. B.Sc. in Mathematics, North China Electric Power University, Beijing City, 2004. M.Sc. in Management, North China Electric Power University, Beijing City.

She is a graduate student in Management tutored by Dr. Zhang.