Forecast Model of Coal Demand Based on Improved Tandem Gray BP Neural Network

Guohua Gou

***Abstract*—In this study, we build a new coal demand prediction model of tandem gray BP neural network. Firstly we use 2000-2015 years coal demand data to establish three gray prediction models: GM(1,1), WPGM(1,1) and pGM(1,1); Secondly, by comparison, we select the best prediction model pGM(1,1) and at the same time take coal demand factors as the BP neural network input, 200-2015year of coal demand date for training and testing. Lastly we proceed to predict coal demand in China in 2016 and 2020. Prediction result is: mean relative error of the improved tandem gray BP neural network prediction results is 1.92%, which is lower 0.158% than pGM(1, 1) model and 0.28% than BP neural network model respectively.**

***Index Terms*—BP neural network, gray forecast, coal demand forecast, tandem gray BP neural network.**

I. Introduction

Coal accounts for 70% of the primary energy consumption structure in China. The coal industry supports the rapid development of the national economy. Coal plays an important role in China's economic and social development. For a long time, the coal industry due to the lack of scientific planning, resulting in supply and demand imbalances affect the healthy development of the national economy. At present, it is necessary to formulate and implement corresponding development strategies and plans to ensure the overall balance and long-term balance of coal production and demand, and scientific and accurate coal demand forecast is the prerequisite for formulating coal development strategies and plans.

There are many methods for coal demand forecasting, which can be divided into single forecasting model and combination forecasting model. Single-item prediction model mainly includes traditional time series method, elasticity coefficient method, co-integration and error correction model, input-output method, system dynamic model; artificial intelligence model (ANN, SVM).Its main drawback is difficult to fit the trend of changes in coal demand, forecasting results are to be improved [1]. The combination forecasting model improves and optimizes the nonlinear flaw and model combination of the single forecasting model, but its selection is biased towards the method orientation, and does not closely follow the characteristics of the coal demand change, and the prediction accuracy needs to be improved.

In view of limitations of the above prediction method, this study proposes a tandem gray BP neural network model to predict coal demand. First of all, substitute the original data of into the three gray prediction model GM (1,1), WPGM (1,1) and pGM (1,1), then compare the prediction accuracy of the three models and last select the predictive results of the most accurate gray model and the main factors of impacting coal demand as the input of BP neural network. This method combines the strengths and weaknesses of the gray theory and artificial neural network algorithm, the results show that the relative error of predicted results of the improved tandem gray BP neural network model is smaller than the actual results, which has a good feasibility on application of coal demand prediction.

Manuscript received November 2, 2016; revised January 12, 2017.

Guohua Guo is with the Department of Modern Management, Zhengzhou Technical College, P. C. 450121 China (e-mail: gouguohua0930@126.com).

II. The Gray BP Neural Network Model

A. Gray Theory Prediction Model

The basic idea of the gray system theory is to take the system of mall sample size, poor information and the uncertainty based on partial information known and some of the information unknown as the research object, extract valuable information mainly through the generation and development of the part known information, make sure the correct description and effective monitoring to the system running behavior and the evolution[2] In this study we select three gray forecasting model GM (1,1), WPGM (1,1) and PGM (1,1) to predict the Coal Demand.

GM(1,1)model: Assume that time series have n observations, $x^{(0)} = \left\{x^{(0)}(1), x^{(0)}(2), \cdots x^{(0)}(n)\right\}$ generate new series by accumulating, $x^{(1)} = \left\{x^{(1)}(1), x^{(1)}(2), \cdots x^{(1)}(n)\right\}$,then the corresponding differential equations of GM (1,1) model is: $\frac{dx^{(1)}}{dt} + ax^{(1)} = \mu$, in which a called the developing gray number, μ known as endogenous control gray number. Set $\hat{a}$ as a be estimated parameters vector $\hat{a} = (\frac{a}{\mu})$, which can be solved by the least squares method, the solution is:

$$\boldsymbol{B} = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \vdots & \vdots \\ -\frac{1}{2}[x^{(1)}(n-1) + x^{(1)}(n)] & 1 \end{bmatrix} \quad Y_n = \begin{bmatrix} x^{(1)}(2) \\ x^{(1)}(3) \\ \vdots \\ x^{(1)}(n) \end{bmatrix} \tag{1}$$

doi: 10.18178/ijiet.2017.7.11.987

Solve the differential equations, then we can obtain Discrete-time response function of gray prediction:

$$\hat{x}(k+1) = [x^{(0)}(1) - \frac{\hat{\hat{\mu}}}{\hat{\hat{a}}}]e^{-\hat{a}k} + \frac{\mu}{a} \quad k = 0,1,2\cdots n \quad (2)$$

Is the accumulating prediction value, restore the predictive value; we can see the gray prediction model:

$$\begin{cases} \hat{x}^{(0)}(1) = x^{(0)}(1) \\ \hat{x}^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1) = (1-e^{\hat{\hat{a}}})[x^{(0)}(1) - \frac{\hat{\mu}}{\hat{a}}]e^{-a(k-1)} \end{cases} \quad (3)$$

WPGM (1, 1) model: If the original data is an index sequence, that is:

$$x^{(0)} = Ae^{a(k-1)} \qquad k = 1,2\cdots N \quad (4)$$

Its one time accumulated generating sequence:

$$x^{(1)} = A(1-e^{a(k-1)})/(1-e^{a}) \qquad k = 1,2\cdots N \quad (5)$$

Using the GM (1, 1) to build a model, we obtain:

$$x^{(0)}(k) + az^{(1)}(k) = u \quad (6)$$

$$B = \begin{bmatrix} -\frac{1}{2}A\frac{2-e^{a}-e^{2a}}{1-e^{a}} & 1 \\ -\frac{1}{2}A\frac{2-e^{ka}-e^{(k+1)a}}{1-e^{a}} & 1 \\ \vdots & \vdots \\ -\frac{1}{2}A\frac{2-e^{(N-1)a}-e^{Na}}{1-e^{a}} & 1 \end{bmatrix} \quad Y_n = \begin{bmatrix} Ae^{a} \\ Ae^{2a} \\ \vdots \\ Ae^{(N-1)a} \end{bmatrix} \quad (7)$$

After derivation available:

$$(\hat{a},\hat{u}) = (B^T B)^{-1} B^T Y_n = [\frac{2(1-e^{a})}{1+e^{a}}, \frac{2A}{1+e^{a}}]^T \quad (8)$$

The ultimate simulation result is:

$$\hat{x}^{(0)}(1) = A; \quad \hat{x}^{(0)}(k) = \frac{Ae^{a}(1-e^{a})}{1-e^{a}} e^{-\hat{a}(k-1)} \quad (9)$$

By (9), we can get:

$$a = \ln\frac{2-\hat{a}}{2+\hat{a}}, A = \frac{2\hat{u}}{2+\hat{a}} \quad (10)$$

Use the GM (1, 1) model parameters $\hat{a}$, $\hat{\mu}$ to express the parameters of the original data sequence. Assume the established model on index sequence is:

$$\hat{x}^{(0)}(k) = \hat{A}e^{\hat{a}(k-1)}; k = 1,2,3\cdots N \quad (11)$$

If, then $\hat{a} = a$, at this point (11) is the no deviation model of (4)

pGM (1, 1) model: Set the original data sequence is, $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots x^{(0)}(n)\}$. Its one time accumulated generating sequence, $x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \cdots x^{(1)}(n)\}$ sequence's albino equation of PGM (1,1) model is:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \quad (12)$$

In which the albino value of gray parameters. $\beta = [a, u]^T$ Determine the optimal weights and generate background value sequence:

$$z' = \{-z'(2), -z'(3), \cdots - z'(n)\} \quad (13)$$

In which $z'(t+1) = px(t+1) + (1-p)x(t)$

Using the least squares method to calculate β, then:

$$\beta = [a,u]^T = (A^T A)^{-1} A^T B \quad (14)$$

In which $A = \begin{bmatrix} -z(2) & -z(3) & \cdots & -z(n) \\ 1 & 1 & \cdots & 1 \end{bmatrix}^T$

$$B = (x^{(0)}_{(2)}, x^{(0)}_{(3)}, \cdots, x^{(0)}_{(n)})^T$$

Substitute the obtained gray parameters into (14) and then find the solution of differential equations:

$$\hat{x}^{(1)}_{(t+1)} = (x^{(0)}_{(1)} + u/a)e^{-at} + u/a \quad (15)$$

If $\hat{x}^{(1)}_{(t+1)}$ is the model calculated value, tired less generate it, you can get the analog value of the model, $\hat{x}^{(0)}_{(t+1)}$ that is:

$$\hat{x}^{(0)}_{(1)} = x^{(1)}_{(0)}$$

$$\hat{x}^{(0)}_{(t+1)} = (1-e^{a})(y^{(0)}_{(1)} - u/a)e^{-at} = x^{(1)}_{(t+1)} - x^{(1)}_{(t)} \quad (16)$$

(15) and (16) is a specific formula for calculating PGM (1, 1) model.

B. BP Neural Network Prediction Model

BP network has a three-tier structure, namely the input layer, the hidden layer and the output layer, which are fully connected. Set the input layer is *i*, the hidden layer is *h* and the output layer is *j*, the number of nodes of three layers respectively are *ni* , *nh* , *nj* , the threshold value of the hidden layer nodes and output layer nodes respectively are θh and θj , the wiring weight between the input layer nodes and the hidden layer node is, the wiring weight of hidden layer nodes and output layer nodes is, each node input is *x* [3].

- Normalize the input and output sample
- The initialization. Assume the input and output samples after normalized are:

$$\{x_{k,j}, d_{k,j} | k = 1,2,\cdots nk; i = 1,2,\cdots ni; j = 1,2,\cdots nj\}$$

- *nk* = The sample capacity, each connection weights{ w_{ih} }{ w_{hj} }and threshold {θh }{θj } =Take n to a random value in the interval (-0.1， 0.1)
- Set *k*=1, provide the input and output samples to the network
- Calculate the input and output of each node of the hidden layer (*h* = 1, 2, ...,*nh*)

$$x_h = \sum_{i=1}^{ni} w_{ih} * x_{ki} + \theta_h; y_h = 1/(1+e^{-xh}) \qquad (17)$$

- Calculate the input and output of each node of the hidden layer (*j*=1, 2, ...,*nj*)

$$x_j = \sum_{h=1}^{nh} w_{hj} * y_h + \theta_j; y_j = 1/(1+e^{-xj}) \qquad (18)$$

- The calculation of the change rate of the total input changes the output layer node receives a single sample error.

$$\frac{\partial E_k}{\partial x_j} = y_j(1-y_j)(y_j - d_{k,j})(j = 1,2,\cdots nj) \qquad (19)$$

- The calculation of the change rate of the total input changes the Hidden layer node receives a single sample error.

$$\frac{\partial E_k}{\partial x_h} = y_k(1-y_k)\sum_{j=1}^{nj}(\frac{\partial E_k}{\partial x_j} * w_{ij})(h = 1,2,\cdots nh) \qquad (20)$$

- The correction of the connection weights and thresholds.

$$w_{hj}^{t+1} = w_{hj}^t - \eta \frac{\partial E_k}{\partial x_j} y_j + a(w_{hj}^t - w_{hj}^{t-1}) \qquad (21)$$

$$\theta_j^{t+1} = w_j^t - \eta \frac{\partial E_k}{\partial x_j} + a(\theta_j^t - \theta_j^{t-1}) \qquad (22)$$

$$w_{ih}^{t+1} = w_{ih}^t - \eta \frac{\partial E_k}{\partial x_h} x_{k,i} + a(w_{ih}^t - w_{ih}^{t-1}) \qquad (23)$$

$$\theta_h^{t+1} = \theta_h^t - \eta \frac{\partial E_k}{\partial x_h} + a(\theta_h^t - \theta_h^{t-1}) \qquad (24)$$

In which correction number is, Learning rate, momentum factor, algorithm converges is slowly while is smaller, algorithm converges is faster while is larger, but it is in-stable, may shock and the function of is opposite.

Set, $k = k+1$ provide ($(x_{k,i}, d_{k,j})$ to the network, then go to step 4), until all the samples are completely trained.

Repeated steps 3 to 9 until the network global error function.

$$E = \sum_{k=1}^{nk} E_k = \sum_{k=1}^{nk}\sum_{j=1}^{nj} \left| y_j - d_{k,j} \right|^2 \qquad (25)$$

Learning frequency is bigger than the preset value or less than a smaller value of the pre-set.

C. Improved Tandem Gray BP Neural Network Model

Tandem gray BP neural network is to take the results of gray prediction model as input of neural network, utilizing the non-linear fitting ability to obtain the final predicted value. But this approach ignores the impact of the other main factors to results prediction, based on which, we put forward a Improved tandem gray BP neural network model to predict thermal coal demand In this study, that is: select a best forecast model among GM (1, 1), WPGM (1, 1) and PGM (1, 1), at the same time take the main factors of effecting coal demand as the input of the neural network to achieve the best fit. As shown in Fig. 1:

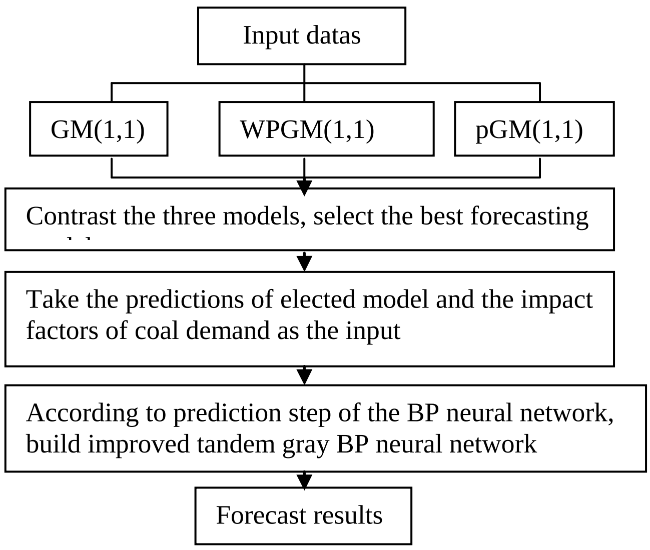

Fig. 1. Overall framework of coal demand prediction model.

III. COAL DEMAND BASED ON THE GRAY BP NEURAL NETWORK MODEL

The selection of the basic data: Neural network and gray neural network model need some data including coal

Consumption, the coal production, and fixed assets investment of coal industry, Wages of coal workers as well as GDP. The specific data is shown in Table 1[4].

TABLE I: THE ORIGINAL SAMPLE DATA

years	coal Consumption (10,000tons)	the coal production (10,000tons)	Coal industry Fixed assets(100million yuan)	Wages of coal workers (100 million yuan)	GDP (100 million yuan)
2000	127040	101018	337.6	312.2	98562
2001	132710	107031	366.1	370.3	108683
2002	165870	114238	408.0	443.3	119765
2003	188290	134972	436.4	662.9	135719
2004	191120	158085	690.4	831.8	160290
2005	226310	177275	1162.9	1031.2	184576
2006	207402	189691	1459.0	1259.6	217247
2007	225795	205527	1804.6	1500.5	268631
2008	229237	213058	2399.2	1847.3	318737
2009	240666	219719	3056.9	2089.1	345046
2010	249568	239712	3784.7	2458.8	407138
2011	271704	264658	4907.3	3174.2	479576
2012	275465	267493	5370.2	3600.7	532872
2013	280999	270523	5212.6	3833.2	583197
2014	281160	263520	4684.5	3728.2	644043
2015	283871	254486	4784.5	3989.2	676708

Source: China Statistical Yearbook 2015

TABLE II: FORECAST RESULTS OF THREE GRAY MODELS

Years	Actual value	Predictive value of GM	Predictive value of WPGM	Predictive value of PGM
2000	127040	134852.98	134022.23	132706.00
2001	132710	141030.91	140255.71	138469.61
2002	165870	176518.87	174875.94	172488.23
2003	188290	200058.15	198745.97	197553.87
2004	191120	202764.94	201536.22	198764.80
2005	226310	240337.36	238837.01	234892.89
2006	207402	220877.92	218667.78	216512.05
2007	225795	239130.70	238507.58	236177.53
2008	229237	244135.29	242349.06	239543.30
2009	240666	255116.29	253745.86	249813.34
2010	249568	264534.09	263630.07	259061.31
2011	271704	288837.96	286727.06	283114.71
2012	275465	292533.77	292261.06	286481.56
2013	280999	298991.94	297342.05	292509.22
2014	281160	299210.48	297354.80	292322.06
2015	283871	301811.61	299711.07	295254.22

D. The Gray Model Selection of Coal Demand

According to the original sample data in Table I, we establish gray prediction model GM(1,1), WPGM(1,1), pGM(1,1), program Three models using MATLAB language and forecast coal demand from1994 to 2010, the predicted results are shown in Table II.

According to Table II, relative error is using forecasting model GM (1,1), WPGM (1,1), pGM (1,1), which is shown in Table III.

TABLE III: ANALOGY PERFORMANCE OF THE THREE MODELS

Model	GM (1,1)	WPGM (1,1)	pGM(1,1)
Average relative Error (%)	6.2457	5.6115	4.1793

As is shown in Table II, the analog performance of pGM (1, 1) is higher than GM(1, 1) and WPGM(1, 1). According to Table IV, the relative error for GM(1, 1), WPGM (1, 1) and pGM(1, 1) belong to level 2, the simulation accuracy is acceptable.

TABLE IV: ACCURACY CLASS REFERENCE TABLE

Accuracy class	level 1	level 2	level 3	Level4
Relative error (%)	1	5	10	20

In summary, pGM (1, 1) is compared to GM (1, 1) and WPGM (1, 1) whose predictive effect is better, so pGM (1, 1) is put into the combined model.

E. The Accuracy Validation of Coal Demand Model Based on Improved Series Gray BP Neural Network

Coal demand factor is the input of coal demand projections network, in theory, the number of impact factors is the number of input layer neurons, the coal production, Wages of coal workers, GDP, and fixed assets investment of coal industry as the input, therefore, input nodes of coal demand artificial neural network model are four. Based on trial and error method, a hidden layer has 13 neurons according to empirical formula of node algorithm; the network training adopts the Levenberg-Marquardt algorithm [5].

After training the data from 2000 to 2010 using above neural network model, predicted results of coal demand are shown in Table V from 2011 to 2015.

TABLE V: COAL DEMAND PREDICTIVE RESULTS WITH ARTIFICIAL NEURAL NETWORK MODEL FROM 2006 TO 2010

Years	Predictive value	Actual value	Relative error (%)
2011	282843.9	271704	0.041
2012	285932.7	275465	0.038
2013	294206.0	280999	0.047
2014	293531.0	281160	0.044
2015	295140.7	283871	0.0397
Average relative error			0.04194

According to above analysis, pGM (1,1) is compared to GM (1,1) and WPGM (1,1) whose predictive effect is better, pGM (1,1) model prediction results is taken as the input of improved series gray neural network model. Combined with the above trained neural network structure, we can see the structure of the improved series gray neural network model: The input layer 5 neurons: pGM (1,1) model prediction results, coal consumption, fixed assets investment of coal industry, Wages of coal workers and GDP, Output layer has one neuron, hidden layer has 13 neurons, network training adopt the Levenberg-Marquardt algorithm[6]. the network structure is shown in Fig. 2.

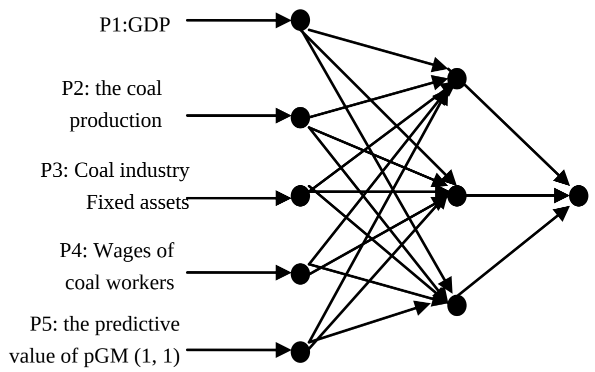

Fig. 2. Network diagram.

Forecasting coal demand from 2011 to 2015 adopts the improved series gray neural network, prediction results are shown in Table VI.

TABLE VI: COAL DEMAND PROJECTIONS WITH THE IMPROVED SERIES GRAY BP NEURAL NETWORK

Years	Predictive value	Actual value	Relative error (%)
2011	276377.3	271704	0.0172
2012	279707.2	275465	0.0154
2013	285579.3	280999	0.0163
2014	285405.5	281160	0.0151
2015	288214.2	283871	0.0153
Average relative error			0.01586

Mean relative error of this prediction results is 1.586%, which is smaller for 4.179% than pGM（1,1）model and for 2.608% than neural network model. Therefore improved series gray neural network can be used to predict coal demand of China from 2016 to 2020.

F. Coal Demand Forecast of China

We can find GDP, coal production, fixed assets investment of coal industry, Wages of coal workers data from 2009 to 2011 from China Statistical Yearbook, China Electric Power Yearbook and other relevant information. As is known from the content, prediction effect of pGM (1, 1) is best, so through the pGM (1, 1) model, forecast the data of all the factors in 2012 and 2020 on the basis of historical data. Specific results are shown in Table VII.

TABLE VII: PREDICTIVE VALUE OF 2016-2020 YEARS CHINA'S COAL DEMAND

years	2016	2017	2018	2019	2020
Predictive value (million tons)	2984	3106	3228	3350	3472

IV. CONCLUSION

Gray forecasting model has the characteristics of less sample data, no need to consider its distribution and trend, simple modeling and convenient operation. But it lacks the ability of self-learning, self-organization and self-adaptability, and the processing ability of nonlinear information is weak . The gray forecasting model is used to predict the nonlinear system. The error of prediction and actual value will be relatively large, and the accuracy of prediction can not meet the requirement. Artificial neural network is an effective non-linear modeling method, in which error back propagation (BP) algorithm is more mature and widely used algorithms, BP neural network has a high degree of mapping ability, can be arbitrarily approximated to any nonlinear Function, more suitable for some complex problem modeling. GM (1, 1) model can not describe the fluctuation trend of data, so the residual GM (1, 1). the study sets up a newly improved series gray neural network model for china's coal demand, with the problems that there are not enough long-term forecasting data for coal demand, we established three gray prediction models which include pGM(1,1), GM(1,1), WPGM(1,1) using time series of coal demand from 2000 to 2015. Combining the advantages and disadvantages of gray theory and neural network algorithm, a new model that the best accuracy pGM(1,1) model, coal production, and fixed assets investment of coal industry, Wages of coal workers as well as GDP are all taken as the input of neural network is constructed. Then the new model is put to test the coal demand from 2011 to 2015, the average relative error of the predicted results is 1.586% which is smaller than pGM (1, 1) model for 4.179% and is also smaller than neural network model for 2.608%. Finally, we make a forecasting for coal demand of china in 2016 and in 2020. Therefore, the newly improved series gray neural network model has a higher prediction accuracy that can be regarded as an effective way for china's coal demand**.**

REFERENCES

[1] G. H. Zhao and S. F. Guo, "An Investigation of coal demand in China based on the variable weight combination forecasting model," *Journal of Resources and Ecology*, vol. 2, 2011, pp. 126-131.

[2] J. L. Deng and Z. Wang, *Gray Forecast and Decision Model*, Science Press, Beijing, 2009, pp. 1-8.
[3] X. M. Hu and G. H. Zhao, "Forecasting model of coal demand based on Matlab BP neural network," *Chinese Journal of Management Science*, vol. 10, 2008, pp. 522-526.
[4] National Bureau of Statistics, "*2015* China statistical yearbook," China Statistics Press, Beijing, 2015.
[5] T. Hagan, H. Demuth, and M. Beale *Neural Network Design*, Boston, MA: PWS Publishing, 1996, pp. 35-40.
[6] J. G. Zhou, M. Lin, and X. D. Pan, "Thermal power industry NOX emissions forecast based on improved tandem gray BP neural network," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 5, 2013, pp. 4716-4721.

Guohua Gou is associate professor. He was born in 1981. He graduated from Northeast Petroleum University in 2004 with a bachelor's degree in management. He received a master's degree in management from Harbin Institute of Technology in 2006. Nowadays, he is in the Department of Management of Zhengzhou technical College, who is engaged in teaching and scientific research. His research direction is regional economics.