Available online www.jsaer.com

Journal of Scientific and Engineering Research, 2023, 10(2):63-71

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Data Detection based on Fractional GM (1,1) and BP Neural Network

Teng Sun*, Bingxu FU, Ruiyuan Liu

School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo 255000, China

Abstract The traditional grey GM (1,1) and BP neural network models are highly dependent on the original data sequence and slow in convergence. The idea of fractional accumulation is introduced into the GM (1,1) model, and then the traditional BP neural network is improved by the layer-by-layer training algorithm. At the same time, based on China's electricity data from 2010 to 2014, a fractional GM (1,1) and BP neural network combination model is built to predict the total electricity generation in 2015 and 2016. The experimental results show that the combined model based on fractional GM (1,1) and BP neural network has better data fitting effect and higher prediction accuracy than GM (1,1) model, fractional GM (1,1) model and the combined model of GM (1,1) and BP neural network.

Keywords Fractional GM (1,1); BP neural network; Layer by layer training; Power forecast

1. Introduction

The grey system theory is a new theory established by Chinese scholar Professor Deng Julong in 1982 to solve the problem of "small sample, poor information" uncertain system [1]. In order to reduce the disturbance of the model, Wu Lifeng constructed the Caputo fractional grey prediction model and designed the rule of information priority [2]. Meng Wei established a fractional grey prediction model and designed an algorithm for order optimization [3]. Hui Zhihao and others optimized the grey prediction model and predicted the school-age population of basic education in Pingdingshan City [4]. In the 20th century, foreign scholars U. Kumar and V.K. Jain used the grey Markov model to predict the energy consumption of India by combining the grey model, rolling mechanism and singular spectrum analysis (SSA) [5]. Chen Zhiqiang et al. adopted the grey GM (1,1) prediction model in the energy management project to provide better prediction advantages for long-term prediction problems. The prediction performance of the improved GM (1,1) model was verified by using the China Energy Database, and compared with the results of artificial neural network and time series. The experimental results showed that the precision model of grey prediction was significantly improved [6]. Wang Jianzhou et al. used the multi-period Hubbert model to predict the annual production of natural gas in China. In addition, he also proposed a small sample effective rolling GM (1,1). The experimental results show that the gap between supply and demand in China will become larger and larger [7]. The fractional cumulative GMC (1,1) model is optimized for the effectiveness of the original data. The model can simulate well and has significant prediction performance compared with the traditional grey prediction model [8]. Yang Yang and Xue Dingyu use the fractional order accumulation method to establish the grey prediction model of interval grey number. Compared with the integer order model, this method has good modeling and prediction performance and can be applied to the prediction of small batch interval historical industrial series samples [9]. Wei Baolei et al. proposed a single variable discrete grey polynomial model and designed an adaptive selection model structure. The results show that the method is effective [10]. Wu Lifeng applied the perturbation theory of the least squares method to explain that the cumulative generation operator violated the system theory of the new grey information priority principle, and proposed a new fractional cumulative grey system model. The example calculation results have very significant prediction performance compared with the traditional prediction methods [11]. Mao Shuhua et al. believed that the classic GM (1,1) model has the following shortcomings: it cannot reflect the new information priority principle. If it is to achieve ideal modeling effect, the raw data must meet the ratio requirement [12]. Wang An and others successively designed the grey combination optimization model [13-14] and a new grey coupling model [15], and proved its effectiveness through an example.

The prediction model in this paper is mainly divided into two parts: fractional grey GM (1,1) model and BP neural network model based on layer-by-layer training algorithm. The grey system seeks its change rule by mining and sorting out the original data. The way to find the data reality rule is called grey sequence generation [16]. The superiority of grey prediction to the prediction of small amount of data can accurately predict the data of small samples, that is, select the factors that affect the total power generation: the power component of the total power generation and the power output to establish the fractional grey GM (1,1) prediction model. For example, the power component of the total national power generation from 2010 to 2014: hydropower generation, thermal power generation, nuclear power generation, wind power generation, import power and power output are predicted as factors affecting the total power generation, and the power component and power output from 2015 to 2016 are predicted; BP neural network has good generalization ability and certain generalization ability and association ability for unknown situations. The structure of BP neural network model based on layer-by-layer training algorithm is the same as that of traditional BP neural network. The algorithm is layer-by-layer training algorithm, which can effectively improve the slow convergence speed of BP neural network. By taking the power component and power output from 2010 to 2016 as the input data of the BP neural network prediction model based on the layer-by-layer training algorithm, the predicted value of total power generation is output. All power data in this paper are from China Statistical Yearbook.

2.Establishment of fractional grey GM (1,1) model

The grey GM (1,1) model is an exponential growth model, which is mainly aimed at the single variable system, seeks the system change rule through the randomness of the weakening sequence generated by accumulation, and establishes a prediction model about time based on this. The fractional-order grey prediction model is a dynamic model that uses data to establish fractional-order prediction model, and then optimizes the design order. First, the original sequence is accumulated fractional-order, and then the fractional-order accumulation is accumulated one by one. The model is built according to the fractional-order accumulation sequence, and finally the order is optimized, so that the precision of the fractional-order grey prediction model is improved.

The fractional GM (1,1) model and GM (1,1) model have the same structure. The essential difference is that the GM (1,1) model uses the first-order accumulation of the original sequence $X^{(0)}$ to generate sequence $X^{(1)}$; The fractional order GM (1,1) model uses the r-order accumulation of the original sequence $X^{(0)}$ to generate the sequence $X^{(\mathrm{r})}$ as the modeling sequence, which has a certain rule and is fitted with a curve. The specific fractional GM (1,1) prediction model establishment process is as follows:

2.1 Definitions

Let the nonnegative original sequence $X^{(0)}=\left(x^{(0)}(1),\ x^{(0)}(2),\cdots,\ x^{(0)}(\mathrm{n})\right)$ be added to the original sequence $X^{(0)}$ to obtain the sequence $X^{(\mathrm{r})}=\left(x^{(\mathrm{r})}(1),\ x^{(r)}(2),\cdots,\ x^{(r)}(\mathrm{n})\right)$, where,

$$x^{(\mathrm{r})}(k)=\sum_{i=1}^{k}\frac{\Gamma(r+k-i)}{\Gamma(k-i+1)\Gamma(r)}x^{(0)}(i)$$

$$=\sum_{i=1}^{k}\frac{\int_0^\infty e^{-x}x^{(r+k-i)}dx}{\int_0^\infty e^{-x}x^{(k-i)}dx\int_0^\infty e^{-x}x^{(r-1)}dx}x^{(0)}(i)$$

$$k=1,2,\cdots,n$$

$x^{(r-1)}(k)+az^{(r)}(k)=b$ is called fractional gray prediction model, where:

$$Z^{(r)}(k)=\frac{x^{(r)}(k)+x^{(r)}(k-1)}{2},$$
$$k=2,3,\cdots,n$$

2.2.1 Theorem 1

The parameter estimation of the fractional grey prediction model is $\hat{u}=[a,b]^T=\left(B^TB\right)^{-1}B^TY$, where Y and B are respectively:

$$Y=\begin{bmatrix} x^{(r-1)}(2) \\ x^{(r-1)}(3) \\ \vdots \\ x^{(r-1)}(n) \end{bmatrix}, \quad B=\begin{bmatrix} x^{(r-1)}(2) & 1 \\ x^{(r-1)}(3) & 1 \\ \vdots & \vdots \\ x^{(r-1)}(n) & 1 \end{bmatrix}$$

2.2.2 Theorem 2

The whitening differential equation of the fractional grey prediction model is $\frac{dx^{(r)}}{dt}+ax^{(r)}=b$, and its time response function is:

$$\hat{x}^{(r)}(t)=\left[x^{(r)}(1)-\frac{b}{a}\right]e^{-at}+\frac{b}{a} \tag{1}$$

The time response sequence of fractional grey prediction model $x^{(r-1)}(k)+az^{(r)}(k)=b$ is:

$$\hat{x}^{(r)}(k)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-a(k-1)}+\frac{b}{a} \tag{2}$$
$$k=2,3,\cdots,n$$

Restore value is:

$$\hat{x}^{(0)}(k)=\left(\hat{x}^{(r)}\right)^{(-r)}(k)$$
$$=\sum_{i=1}^{k-1}(-1)^i\frac{\Gamma(r+1)}{\Gamma(i+1)\Gamma(r-i+1)}\hat{x}^{(r)}(k-i) \tag{3}$$
$$k=2,3,\cdots,n$$

Among them, $\hat{x}^{(0)}(1)=x^{(0)}(1)$

2.3 Algorithm steps

According to the principle of fractional grey prediction model, the algorithm steps can be summarized as follows:

Step 1 Generate r-order accumulation sequence and r-order accumulation sequence

Let the non-negative original sequence $X^{(0)}=\left(x^{(0)}(1),\ x^{(0)}(2),\cdots,\ x^{(0)}(\mathrm{n})\right)$, and the original sequence $X^{(0)}$ is accumulated to obtain the sequence: $X^{(\mathrm{r})}=\left(x^{(\mathrm{r})}(1),\ x^{(r)}(2),\cdots,\ x^{(r)}(\mathrm{n})\right)$, where:

$$x^{(\mathrm{r})}(k)=\sum_{i=1}^{k}\frac{\Gamma(r+k-i)}{\Gamma(k-i+1)\Gamma(r)}x^{(0)}(i)$$
$$=\sum_{i=1}^{k}\frac{\int_0^\infty e^{-x}x^{(r+k-i)}dx}{\int_0^\infty e^{-x}x^{(k-i)}dx\int_0^\infty e^{-x}x^{(r-1)}dx}x^{(0)}(i)$$
$$k=1,2,\cdots,n$$

Including:

$$Z^{(\mathrm{r})}(k)=\frac{x^{(r)}(k)+x^{(r)}(k-1)}{2},$$
$$k=2,3,\cdots,n$$

$X^{(\text{-r})}=\left(x^{(\text{-r})}(1),\ x^{(-r)}(2),\cdots,\ x^{(-r)}(\mathrm{n})\right)$ is the r-order cumulative generation operator of $X^{(0)}$, where:

$$x^{(\text{-r})}(k)=\sum_{i=0}^{k-1}\frac{\Gamma(r+1)}{\Gamma(i+1)\Gamma(r-i+1)}x^{(0)}(k-i)$$
$$=\sum_{i=0}^{k-1}(-1)^i\frac{\int_0^\infty e^{-x}x^{(r)}dx}{\int_0^\infty e^{-x}x^{i}dx\int_0^\infty e^{-x}x^{r-i}dx}x^{(0)}(k-i)$$
$$k=1,2,\cdots,n$$

Step 2 Establish the first order differential equation

$$\frac{dx^{(r)}}{dt}+ax^{(r)}=b$$

It is the whitening differential equation of fractional order operator GM (1,1) model.

Step 3 Calculate the parameter values of a and b, and the parameter vector $\hat{u}=[a,b]^T$ in the GM (1,1) model $x^{(\mathrm{r}-1)}(\mathrm{k})+\mathrm{az}^{(\mathrm{r})}(\mathrm{k})=\mathrm{b}$ can be estimated by the least square method

$$\hat{u}=\left(B^TB\right)^{-1}B^TY$$

Step 4 Find the time response of fractional GM (1,1)

$$\hat{x}^{(r)}(k)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-a(k-1)}+\frac{b}{a},\quad k=2,3,\cdots,n$$

Step 5 Reduce the value in r order to get the predicted value

$$\hat{x}^{(0)}(k) = \left(\hat{x}^{(r)}\right)^{(-r)}(k)$$
$$= \sum_{i=1}^{k-1} (-1)^i \frac{\Gamma(r+1)}{\Gamma(i+1)\Gamma(r-i+1)} \hat{x}^{(r)}(k-i)$$
$$k = 2, 3, \cdots, n$$

Step 6 Model verification

The residual test formula is:

$$\varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) \tag{4}$$

The relative error test formula is:

$$\Delta_k = \frac{|\varepsilon(k)|}{x^{(0)}(k)} \tag{5}$$

The average absolute percentage error test formula is:

$$\Delta = \frac{100\%}{n} \sum_{k=1}^{n} \Delta_k \tag{6}$$

The algorithm flow chart of order optimization of fractional grey prediction model is shown in Figure 1

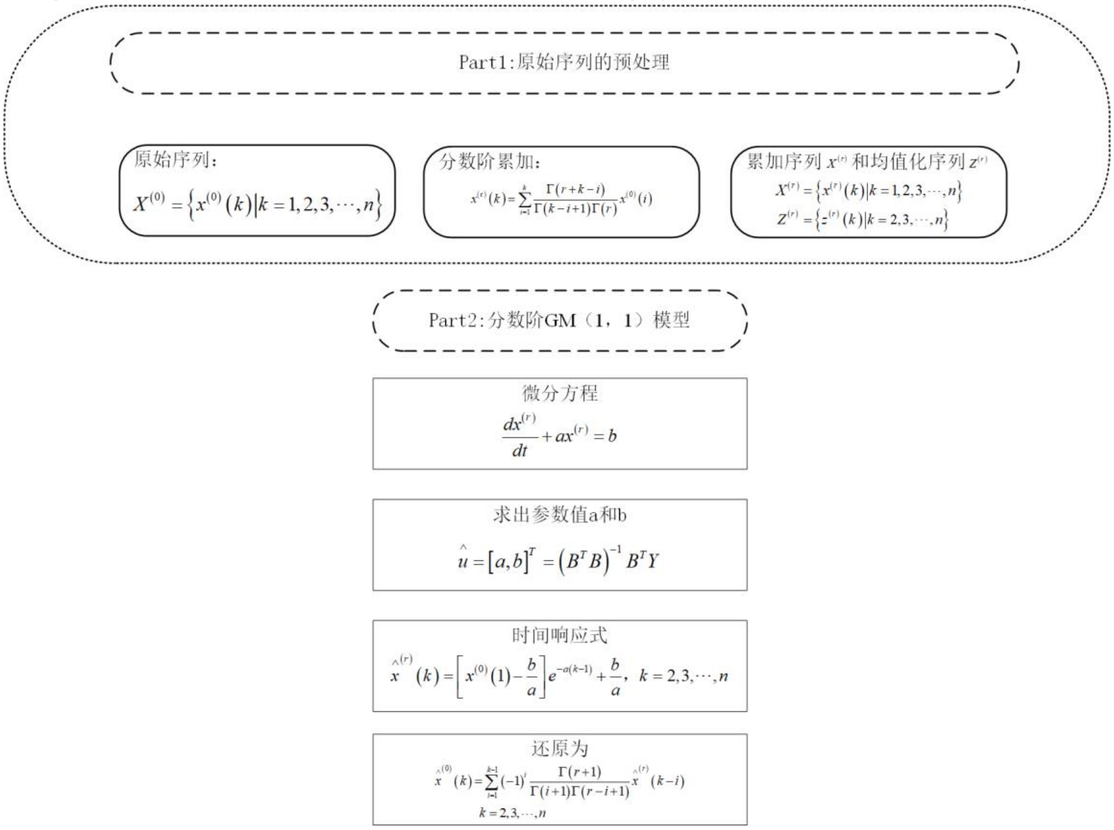

Figure 1: Algorithm flow of fractional grey prediction model order optimization

3. BP neural network prediction model based on layer-by-layer training algorithm

BP neural network is a multilayer feedforward neural network. The main characteristics of BP neural network are signal forward transmission and error back propagation [17]. In forward propagation, the input signal is processed by the hidden layer and then transmitted to the output layer. If the output layer node fails to achieve the expected output, it will enter the reverse propagation phase of the error, return the output error to the input layer through the hidden layer in a certain seed form, and allocate it to the hidden layer node and the input layer node, so as to obtain the error signal of each layer unit as the basis for modifying the weight value of each unit.

BP algorithm only uses the information of the first derivative (gradient) of the mean square error function to the weight and threshold value, so the convergence speed of the algorithm is slow, and it is easy to fall into local minima and other defects. In order to solve this problem, Hinton et al. [18] proposed the unsupervised greedy layer-by-layer training algorithm in 2006, that is, a machine learning method of deep neural network based on human brain learning thought, which brings hope to solve the optimization problems related to the deep structure. The main idea of the layer-by-layer training algorithm is to train only one layer of the network at a time, and each layer is trained separately. First train a network with only one hidden layer, and then train a network with two hidden layers after the training of this layer is completed, and so on (Figure 2). In each step, we fix the trained first k-1 layer, and then add the k layer (that is, take the output of the trained first k-1 as the input). The weights obtained from the individual training of each layer are used to initialize the weights of all BP neural networks, and all layers are put together to optimize the training error on the labeled training set.

This paper establishes a combination model based on fractional grey GM (1,1) model and BP neural network.

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\left(Y_i - \hat{Y}_i\right)}{Y_i}\right| \tag{7}$$

The above formula is the objective function of BP neural network, namely the mean relative error (MRE).

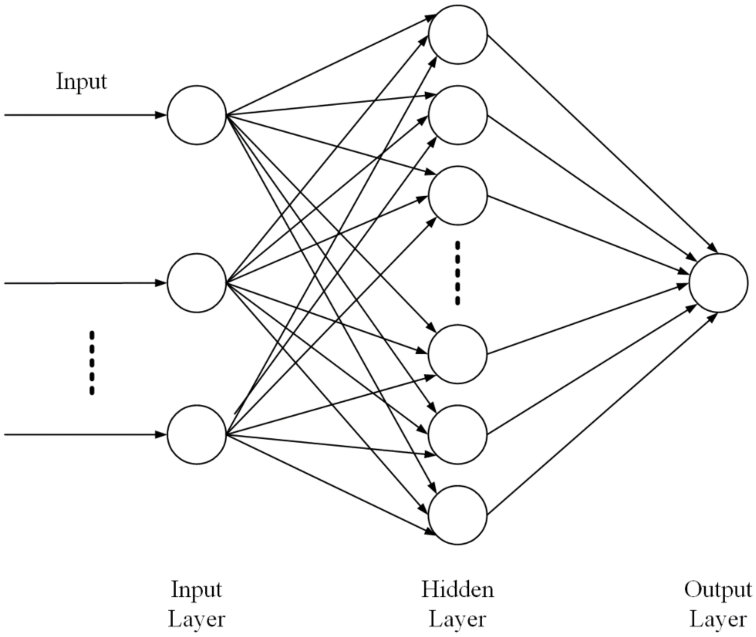

Figure 2: BP neural network structure

4. Application

The five components of total power generation and power output in 2010-2014 are substituted into the fractional GM (1,1) prediction model as modeling data to obtain the simulation data in 2010-2014 and the prediction data in 2015-2016 (see Table 1). Compared with the original data, the average relative error accuracy of the simulated values of five power components and power output predicted by the fractional GM (1,1) model is low, so it can be used as the input signal of BP neural network.

The input layer of BP neural network prediction model based on layer-by-layer training algorithm is 5 power components and power output of total power generation, and the hidden layer is set with 12 nodes to output prediction data. The five power components and power output predicted by the fractional GM (1,1) model (as shown in Table 1) are substituted into the BP neural network prediction model based on the layer-by-layer training algorithm as the six nodes of the input layer to obtain the power data of 2015 and 2016 (as shown in Table 2).

Table 1: Forecast data of total power generation component and power output (unit: 100 million kWh)

Year	hydropower	thermal power	nuclear power	wind power	import	power output	total power generation
2010	7221.7	33319.3	738.8	446.2	55.5	51936.5	42071.6
2011	6989.5	38337	863.5	703.3	65.6	47002.7	47130.2
2012	8721.1	38928.1	973.9	959.8	68.7	49767.7	49875.5
2013	9202.9	42470.1	1116.1	1412	74.4	54204.1	54316.4
2014	10643.4	42686.5	1325.4	1560.8	67.5	56371.8	56495.8
2015	11823.1	45554.4	1565.1	1835.7	66.8	60258.1	
2016	13108.9	48042.2	1868.7	2099.5	64.4	63960.5	
Average relative error (%)	2.682	1.449	0.474	5.083	2.283	0.735	

Table 2: Forecast data of total power generation (unit: 100 million kWh)

Year	hydropower	thermal power	nuclear power	wind power	import	power output	total power generation	Total power generation forecast
2010	7221.7	33319.3	738.8	446.2	55.5	51936.5	42071.6	42071.59765
2011	6989.5	38337	863.5	703.3	65.6	47002.7	47130.2	47130.19921
2012	8721.1	38928.1	973.9	959.8	68.7	49767.7	49875.5	49875.21484
2013	9202.9	42470.1	1116.1	1412	74.4	54204.1	54316.4	54316.39843
2014	10643.4	42686.5	1325.4	1560.8	67.5	56371.8	56495.8	56495.79687
2015	11823.1	45554.4	1565.1	1835.7	66.8	60258.1		60630.71875
2016	13108.9	48042.2	1868.7	2099.5	64.4	63960.5		64921.55468

It can be seen from the total power generation prediction results of the four models in Table 3 that the simulated value of this model in 2010-2014 is very close to the original data, and the average relative error between the simulated value and the original data is also very low. The combined model in this paper is more than the single GM (1,1), The accuracy of fractional GM (1,1) and GM (1,1) combined with BP neural network forecasting model has been significantly improved. It is very suitable for power load forecasting and can accurately predict the total power generation in China in the future.

Table 3: Error comparison of total power generation forecast data of several models (unit: 100 million kWh)

Year	total power generation	GM（1，1）Analog value	FGM（1，1）Analog value	GM (1,1) and BP neural network Analog value	Fractional GM (1,1) And BP neural network Analog value
2010	42071.6			42066.02	42071.597
2011	47130.2	47178.2	47130.2	47127.8	47130.199
2012	49875.5	50217.7	50271.7	49878.88	49875.214
2013	54316.4	53453.1	53493.9	54318.66	54316.398
2014	56495.8	56896.9	56857.4	56494.44	56495.796
2015		60562.6	60391.7	61172.72	60630.718
2016		64464.4	64117.4	65964.01	64921.554
Average relative error (%)		0.7719	0.7372	0.00634	0.0001175

Figure 3 is the comparison curve between the combined forecasting model of this paper and the single GM (1,1) forecasting model, the fractional GM (1,1) forecasting model, and the GM (1,1) and BP neural network combined forecasting model. The curve fitting of the simulated value curve of the total power generation of this model from 2010 to 2014 and the original data is very high, and the curve fitting error between the simulated curve of other forecasting models and the original data is relatively low.

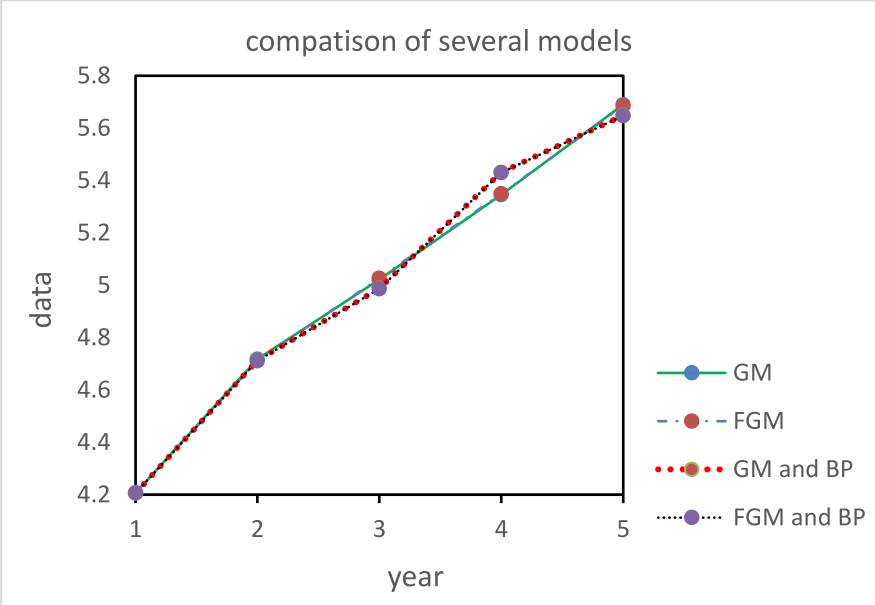

Figure 3: Comparison of several prediction curves

5. Conclusion

Both the traditional fractional GM (1,1) model and the BP neural network model are suitable for dealing with nonlinear data. The GM (1,1) model can also deal with small sample data, but it has the disadvantages of high dependence on the original sequence and slow convergence. This paper uses the layer-by-layer training algorithm to improve the BP neural network, and combines the fractional GM (1,1) model to establish a combined model for power load forecasting. The empirical test shows that the data fitting effect is good and the accuracy is high. The results meet the demand of power load forecasting and provide an important scientific basis for the development of power and economy in China.

References

[1]. Deng Julong. Grey control system [J]. Journal of Huazhong Institute of Technology, 1982, 10 (3): 9-18
[2]. Wu Lifeng. Research on fractional grey prediction model and its application [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2015
[3]. Meng Wei. Grey prediction model based on fractional expansion operator [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2015
[4]. Hui Zhihao, Ma Fuqiang, Yang Jinwei. Prediction and analysis of the evolution of the number and structure of school-age population in Pingdingshan [J]. Journal of Qufu Normal University (Natural Science Edition), 2015, 41 (4): 51-54, 68
[5]. Kumar U，Jain V K． Time series models (Grey — Mark- ov，Grey Model with rolling mechanism and singular spec- trum analysis) to forecast energy consumption in India [J]．Energy，2010，35(4): 1709 — 1716．
[6]. Chen Z Q，Wang X J．Applying the Grey Forecasting Model to the Energy Supply Management Engineering [J]． Systems Engineering Procedia，2012，5: 179—184．
[7]. Wang J Z，Jiang H Y，Zhou Q P，et al． China’s natural gas production and consumption analysis based on the mul ticycle Hubbert model and rolling Grey model [J]. Renew- able and Sustainable Energy Reviews，2016，53: 1149 — 1167．
[8]. WU l F，GAO X H，XIAO Y L，et al． Using a novel multi — variable grey model to forecast the electricity consumption of Shandong Province in China [J]．Energy，2018，157: 327 — 335．

[9]. Yang Y，Ue D Y．An actual load forecasting methodology by interval grey modeling based on the fractional calculus [J]．ISA Transactions，2018，82: 200 － 209．
[10]. Wei B L，Xie N M，Yang Y J．Data － based structure selection for unified discrete grey prediction model [J]. Expert Systems with Applications，2019，136: 264—275．
[11]. Wu L F，Liu S F，Yan S L，et al. Grey system model with the fractional order accumulation [J]．Commun Non- linear Sci Numer Simulat，2013，18(7): 1775 － 1785．
[12]. Mao S H，Gao M Y，Xiao X P，et al．A novel fractional grey system model and its application [J]. Applied Mathematical Modelling，2016，40(7 /8): 5063 － 5076．
[13]. Wang An, Yang Yu. Grey combination forecasting model based on least square method and its application [J]. Practice and understanding of mathematics, 2018, 48 (22): 291-299.
[14]. Wang An, Wang Wenhu, Hu Jiao, et al. Application of the optimal combination forecasting model in China's natural gas demand forecast [J]. Journal of Pingdingshan University, 2018, 33 (5): 32-39
[15]. Wang An, Yang Yu, Xie Lihua. A new grey coupling model and its application [J]. Journal of Pingdingshan University, 2020, 35 (5): 11-15, 21.
[16]. Liu Sifeng, Yang Yingjie, Wu Lifeng, et al. Grey system theory and its application [M]. Beijing: Science Press, 2014
[17]. Wang Xiaochuan, Shi Feng, Yu Lei, Li Yang. Analysis of 43 cases of MATLAB neural network [M]. Beijing: Beijing University of Aeronautics and Astronautics Press, 2013.
[18]. Hinton G E，Salakhutdinov R. Reducing the dimensionality of data with neural networks [J]. Science, 2006，313(5786): 504—507.