The Combined Model of Gray Theory and Neural Network which is based Matlab Software for Forecasting of Oil Product Demand

Song Zhaozheng[1], Jiang Yanjun[2], Jiang Qingzhe[1]

1State Key Laboratory of Heavy Oil for Processing, University of Petroleum, Beijing102249, China

2.Turpan-Hami Oilfield,Shanshan Xinjiang838202, China

Abstract: Under the situation that single prediction can't meet the accuracy requirements, combined forecasting method is usually the first choice. Using the history data of domestic oil products consumption, by optimized combination of grey forecasting GM(1, 1) model and artificial neural network, the combined forecasting model was built. It can be getting from the combined forecasting model that on the 2020, the gasoline demand is 101.52 million ton and the diesel demand is187.57 million ton.

Keywords: oil products demand; grey prediction model; BP neural network; combined forecasting

1. Introduction

Energy problems are vital to national well-being and the people's livelihood as well as national safety. In the past decades, they have been paid more attentions in the world. Long-term planning of energy resources supply and demand is desired to meet the requirements of sustainable economic development, among which forecasting of oil products' demands is an essential part. Forecasting can help the decision makers gain insight of scale and development trend of oil products' demand, such that sound planning for development and utilization of oil resources can be generated. Accurate forecasting can also make the investments from refinery corporations more reasonable, which is useful for their engineering planning, reconstruction and expansion activities. In addition, forecasting is helpful for rational allocation of gasoline/diesel ratio to maximally utilize the resources(Jiang, 2001).

Generally, the forecasting methods can be categorized into two classes: causal forecasting and historical data-based forecasting methods. Causal forecasting assumes the cause-and-effect relationships between the oil products' demand and other independent impact factors so that the relationships model can be formulated accordingly. Typical methods of this kind include multivariate linear regression, multivariate nonlinear regression, and artificial neural network. Historical data-based forecasting relies only on historical observation and empirical data. Through time series analysis of data, the variation laws of the variables are addressed and the future trends are identified. Typical methods of this kind include time series analysis, and grey forecasting method. Among the aforementioned forecasting approaches, grey forecasting based on grey system theory is becoming more and more popular, and has been applied in a wide range of fields (Deng, 1982). Grey forecasting has strong adaptation ability because it only requires less data and the distribution information is not necessary. Only a few discrete data are sufficient to characterize an unknown system and formulate the real-time models, leading to the practicality in tackling the real-world problems associated with uncertainty. However, grey forecasting may encounter difficulties in handling information due to the lack of self-learning, self-organization and self-adaptive abilities. Moreover, its prediction precision is smaller and uncontrollable because of deficiencies of error feedback adjustment and validation. The computational processes are time-consuming and complex because they involve a number of uncertain information (Akay and Atak, 2007; Deng, 1987).

Artificial neural network is effective for dealing with the above problems due to strong self-learning, self-organization and self-adaptive abilities. It can adjust and validate the error feedbacks when learning samples. It also has strong abilities of computation and error validation (Shan et al., 2007). Thus, one potential approach for improving the grey forecasting is to integrate artificial neural network and grey system theory into a framework, leading to a grey neural network forecasting model for dealing with complex nonlinear problems. Therefore, the objective of this study is to propose a grey neural network combination (GNNC) model for predicting oil products' demand and improving the prediction precision. The Matlab program will be developed for helping the implementation of the model.

2. Methodology

2.1 Grey model GM(1, 1)

The grey model GM(1, 1) is one of the most widely used grey forecasting model, which is a time series forecasting model. This model relies only on the original data to search the intrinsic regularity of data. Its principle is to

978-1-935068-37-2

generate regular data sequence $x^{(1)}(k)=(x^{(1)}(1),x^{(1)}(2),\cdots,x^{(1)}(n))$ through the accumulated generating operation (AGO) based on the initial sequence $x^{(0)}(t)=(x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n))$ with irregularity. The accumulated generating operation is written as:

$$x^{(1)}(k)=\sum_{t=1}^{k}x^{(0)}(t) \tag{1}$$

GM(1, 1) is a single variable first-order grey model, which weakens the randomness of the original data series. The first-order differential equations of the grey model GM(1, 1) can be formulated as follows (Hsu and Chen, 2003; Zhou et al., 2006; Wang and Hsu, 2008; Wang and Sun, 2007):

$$\frac{dx^{(1)}}{dt}+ax^{(1)}=u\,, \tag{2}$$

where a is the development coefficient for reflecting the development trends of $x^{(0)}$ and $x^{(1)}$, and u is the grey input. The values of a and u can be obtained by applying the least square method to equation (2) as follows (Yao and Chi, 2004; Mu et al., 2002):

$$\begin{bmatrix} a \\ u \end{bmatrix}=(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{Y}_N \tag{3}$$

where $$\mathbf{B}=\begin{bmatrix} -\frac{1}{2}\left[x^{(1)}(2)+x^{(1)}(1)\right] & 1 \\ -\frac{1}{2}\left[x^{(1)}(3)+x^{(1)}(2)\right] & 1 \\ \cdots & \cdots \\ \cdots & \cdots \\ -\frac{1}{2}\left[x^{(1)}(n)+x^{(1)}(n-1)\right] & 1 \end{bmatrix},$$

$$\mathbf{Y}_N=\left[x^{(0)}(2),x^{(0)}(3),\cdots,x^{(0)}(n)\right]^T$$

Substituting the values of a and u into equation (2), the solutions of equation (2) can be obtained as follows:

$$x^{(1)}(k+1)=(x^{(0)}(1)-\frac{u}{a})e^{-ak}+\frac{u}{a}\,, \tag{4}$$

Equation (3) is the forecasting equation. By using the inverse accumulated generating operation, the predicted value of $x^{(0)}(k+1)$ at time (k+1) can be obtained as:

$$x^{(0)}(k+1)=x^{(1)}(k+1)-x^{(1)}(k) \tag{5}$$

The residual error is formulated as:

$$q^{(0)}(k)=\hat{x}^{(0)}(k)-x^{(0)}(k) \tag{6}$$

The relative error is estimated as:

$$e(t)=\frac{q^{(0)}(k)}{x^{(0)}(k)}\,. \tag{7}$$

Generally, the smaller the value of $|e(t)|$ is, the higher the prediction accuracy.

Generally, the model and the obtained parameters needs to be validated before usage for forecasting. If the error is large enough, residual GM(1, 1) model should be formulated to modify the residual series. The commonly used three methods for validation of grey model GM(1, 1) include residual forecasting model, correlation model, and small probability of error. The most commonly used is the relative error criterion. When a $\le$ 0.3, the grey model GM(1, 1) is suitable for medium- and long-term forecasting.

2.2 Back propagation neural network

Back propagation (BP) neural network is one of the most frequently used artificial neural networks due to its strong self-learning and self-adaptive abilities. The learning process of BP neural network consists of two stages: forward pass of signals and backward pass of errors. In the forward pass stage, sample signals from the input layer are transmitted forward and passed to the output layer via one or multiple middle layers by determining the weight value of each signal in each layer. Output of each layer is merely affected by that of the previous layer. If the error between the actual output of the output layer and the expected output does not fall into a fixed precision range, the error signals will propagate backward through the network. Back propagation of the errors is to return the error signals along the network to amend the weight values of each layer and calculate the adjusted weight values accordingly. By the two iterative processes, the weight values of each layer are continuously adjusted until the error falls into an acceptable range, or the predetermined number of learning cycle is reached (Zhao et al., 2007; Tian et al., 2007).

2.3 Grey neural network combination model

BP neural network has a number of advantages, including high short-term forecasting accuracy, ability of achieving local optimum and modeling the nonlinear relationships. However, it may have poor performance of extension because it requires a wide range of data for various scenarios. In this study, a grey artificial neural network forecasting combination

(GNNC) model is presented. Firstly, forecasting is conducted based on the original irregular data by using the grey model. Then, the combination model is formulated by combining the predicted and observed values. The proposed combination model can make full use of the new data obtained from grey residual errors model through accumulated generation, weaken the randomness of the original data, and make it easy to search the regularity of data. Thus, the combination model requires less data without considering the correlations of data, and has strong abilities of self-learning, nonlinear mapping and parallel processing (Fu et al., 2006; Yin and Ding, 2002). The training, validation, and application of GNNC model can be implemented by using the Matlab 6.5 toolboxes (Cong, 2002). The forecasting precision of GNNC model is better than that of single grey model.

3. Results analysis

3.1 Grey model

Based on the annual consumptions of gasoline and diesel in China Statistical Yearbook published by National Bureau of Statistics of China, energy consumption sequence $x^{(0)}$ for years 1996-2009 was established. The predicted values of $\hat{x}^{(0)}$, residual error $\varepsilon(k)$, and relative error $e(t)$ can be obtained by using Matlab programs, as shown in Table 1.

The gasoline and diesel demands from 2010 to 2020 were predicted based on the forecasting equation (2), as shown in Table 2. The increasing rate of diesel demand would become slow compared to that of gasoline demand. Thus, ratio of diesel/gasoline would decrease after reaching its maximum in 2008. After that, diesel demand would be predicted through modifying the model with two years as a development phase.

3.2 Grey neural network combination model

GNNC model uses the original consumptions of oil products from 1999 to 2009 and predicted consumptions of oil products from 1996 to 2009 as training data. After being trained, it is used to predict the demand of oil products from 2010 to 2020. The predicted results are shown in Tables 3 and 4. Figures 1 and 2 show the scatter plots using Matlab.

Table 1 The predicted consumptions of gasoline and diesel from 1996 to 2009 (Unit: 10^4 tons)

Year	Gasoline consumption				Diesel consumption			
	$x^{(0)}$	$\hat{x}^{(0)}$	$q^{(0)}(k)$	$e(t)$	$x^{(0)}$	$\hat{x}^{(0)}$	$q^{(0)}(k)$	$e(t)$
1996	3182	3182	0	0	4692	4692	0	0
1997	3312	2819	-493	-14.89%	5291	5199	-92	-1.74%
1998	3329	3029	-300	-9.01%	5283	5670	387	7.33%
1999	3381	3254	-127	-3.75%	6232	6184	-48	-0.77%
2000	3505	3496	-9	-0.25%	6774	6744	-30	-0.45%
2001	3598	3757	159	4.41%	7108	7355	247	3.47%
2002	3750	4036	287	7.64%	7667	8021	354	4.61%
2003	4072	4337	265	6.50%	8410	8747	337	4.01%
2004	4696	4660	-36	-0.77%	9895	9539	-356	-3.59%
2005	4853	5006	153	3.16%	10972	10403	-569	-5.19%
2006	5242	5379	137	2.62%	11836	11346	-490	-4.14%
2007	5519	5780	260	4.72%	12392	12373	-19	-0.15%
2008	6343	6210	-133	-2.09%	13886	13494	-391	-2.82%
2009	7195	6672	-523	-7.26%	13859	14716	857	6.19%

Table 2 The predicted demands of gasoline and diesel for years 2010-2020 (Unit: 104 tons)

Year	2010	2012	2014	2016	2018	2020
Gasoline	7169	8276	9554	11030	12733	14699
Diesel	16049	19088	22702	27001	32113	38194

978-1-935068-37-2 © 2010 SciRes.

Table 3 The predicted demand of gasoline from 1996 to 2006 in China (Unit: 104 tons)

Year	Consumption of gasoline			Consumption of diesel		
	Actual value	Predicted value from GNNC	Relative error	Actual value	Predicted value from GNNC	Relative error
1999	3381	3383	0.07%	6232	6185	-0.98%
2000	3505	3435	-1.99%	6774	6760	-0.27%
2001	3598	3625	0.76%	7108	7236	2.43%
2002	3750	3834	2.25%	7667	7684	0.28%
2003	4072	4134	1.52%	8410	8362	-0.70%
2004	4696	4517	-3.81%	9895	9723	-2.42%
2005	4853	4878	0.51%	10972	11251	3.64%
2006	5242	5221	-0.40%	11836	11641	-2.32%
2007	5519	5649	2.35%	12392	12468	0.77%
2008	6343	6289	-0.85%	13886	13833	-0.48%
2009	7195	7184	-0.15%	13859	13885	0.22%

Table 4 The predicted demands of gasoline and diesel from 2007 to 2020 (Unit: 104 tons)

Year	2010	2012	2014	2016	2018	2020
Gasoline	7674	8913	9599	9825	9934	10152
Diesel	14058	14860	16044	17355	18310	18757

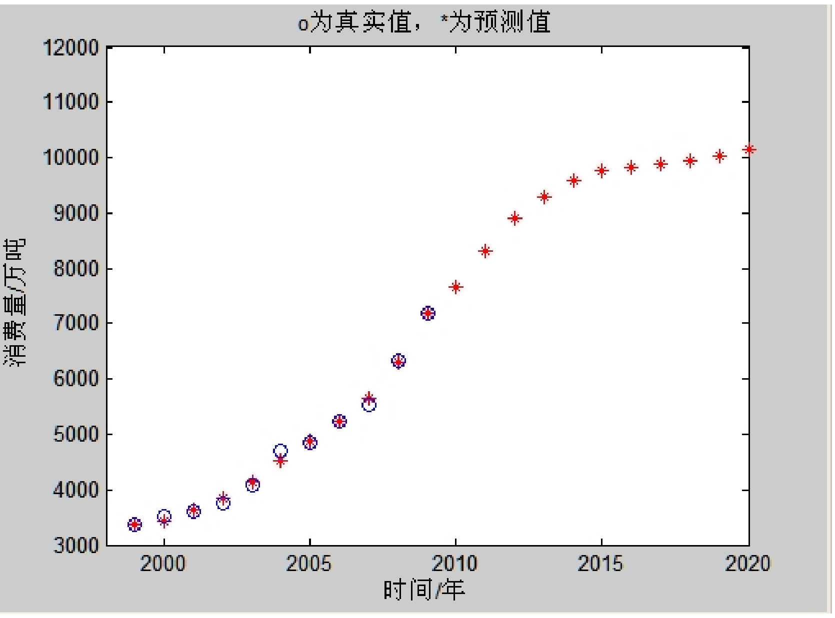

Figure 1 Scatter plot for gasoline demand from GNNC model

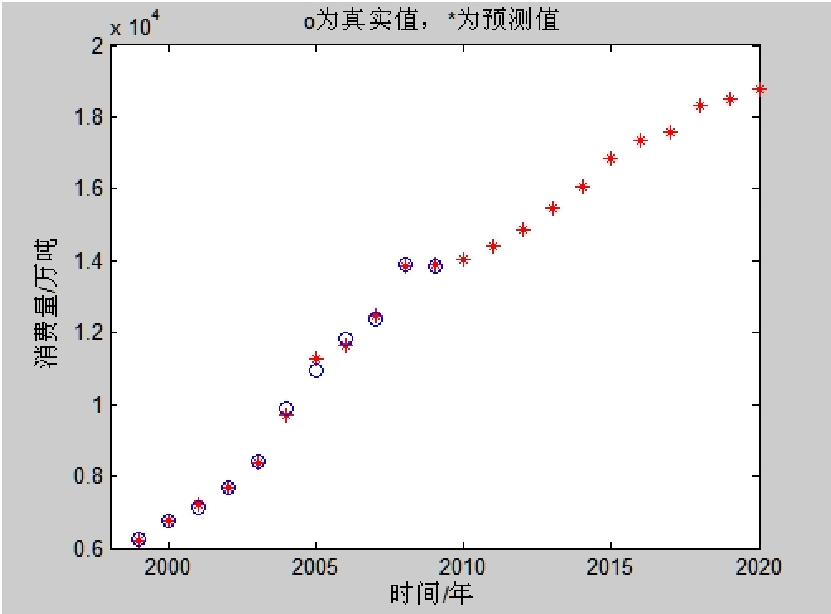

Figure 2 Scatter plot for diesel demand from GNNC model

Table 3 indicates that the prediction precision of GNNC model is better than that of single model, and the maximum relative error is no larger than 4%. Table 4 shows that the demands of gasoline and diesel would be 7674 × 10^4 and 14058 × 10^4 tons in 2010, respectively (10152 × 10^4 and 18757 × 10^4 tons in 2020, respectively). It is indicated that uncertainty associated with automotive industry development has the greatest effects on consumption of oil products, and oil products' prices significantly contribute to the demand of oil products, especially diesel.

4. Conclusions

This paper developed a grey neural network combination model by incorporating grey model and neural network into a framework. The proposed model is distinctive due to less data requirements, simplicity, and strong abilities of modeling the nonlinear relationships, error-tolerating, self-organization, and self-adaptation. The demand of oil products from 2010 to 2020 was predicted. The forecasting results indicated that the combination model had higher prediction precision, and was effective for forecasting the future demand of oil products. The demands of gasoline and diesel in 2020 would be 10152 × 10^4 and 18757 × 10^4 tons, respectively. oil apparent consumption of China reached 40837.5 × 10^4tons in year 2009,which just ranked behind the United States and was the second in the world ,the amount of oil imports reached 21888.5× 10^4tons and oil External dependence was high as 53.4%,According to the report from State Information Center it would be 66% in year 2020[16]. It is indicated that the energy shortage problems are vital to build up a resources-saving and environmental-friendly society in terms of energy security.

Due to the complexity and uncertainty of the real-world systems, it is impossible to develop a mathematical model for fully reflecting the system features without any assumptions. The prediction results are not fully accurate because they rely mainly on the historical data directly or indirectly. Many factors including variations of economic structures, national policies, and development of automotive industry can cause significant variations of oil products' demands in the future. The forecasting model is only a tool for facilitating energy resources management and providing decision support.

References

[1] Akay, D., and Atak, M. (2007). Grey Prediction with rolling mechanism for electricity demand forecasting of Turkey. Energy, 32: 1670-1672.

[2] Cong, S. (2002). Theory and application of neural network using MATLAB toolboxs. Beijing: China University of Science and Technology Press.

[3] Deng, J. L. (1987). Grey system fundamental method. Huazhong University of Science and Technology, Wuhan, China.

[4] Fu, J. F., Cai, G. T., and Zhang, L. (2006). Combination forecasting model of energy consumption based on gray model and neural network. Resource Development & Market, 2(3):216-217.

[5] Hsu, C. C., and Chen, C. Y. (2003). Applications of improved grey prediction model for power demand forecasting. Energy Conversion and Management, 44: 2242-2243.

[6] Jiang, Q. Z. (2001). Research on development planning of clean fuel oil. Beijing: China Petroleum University Press.

[7] Mu, H. L., Dong, X. H., Wang, W., Sakamoto, K., Ning, Y. D., Tonooka, Y. and Zhou, W. S. (2002). Improved gray forecast

978-1-935068-37-2 © 2010 SciRes.

models for China's energy consumption and CO2 emission. Journal of Desert Research, 22(2):142-149.
[8] Shan, M. X., Yu, F., Liu, B. L. (2007). Mineral resources forecasting model combining grey system and BP neural network. Computing Techniques for Geophysical and Geochemical Exploration, 29(5), 443-445.
[9] Tian, J. S., Yu, Q. Y., and Zhang, F. (2007). Energy consumption prediction with combined model. Resource Development & Market, 2(9):803.
[10] Wang, C. H., and Hsu, L. C. (2008). Using genetic algorithms grey theory to forecast high technology industrial output. Applied Mathematics and Computation, 195:258.
[11] Wang, X., and Sun, H. (2007). Application study of traffic flow forecasting model based on grey neural network. Technology & Economy in Areas of Communications, 2:69.
[12] Yao, A. W. L., and Chi, S. C. (2004). Analysis and design of a taguchi - grey based electricity demand predictor for energy management systems. Energy Conversion and Management, 45:1205-1217.
[13] Yin, Z. M., and Ding, C. L. (2002). Mastering MATLAB 6. Beijing: Tsinghua University Press.
[14] Zhao, Y. N., Zhang, Q., and Ma, Y. L. (2007). Application of grey and BP neural network combined forecasting model in small sample problems. Statistics and Decision, 2: 14.
[15] Zhou, P., Ang, B. W., and Poh, K. L. (2006). A trigonometric grey prediction approach to forecasting electricity demand. Energy, 31:2840-2841.