Research on Logistics Demand Forecast based on the Combination of Grey GM (1, 1) and BP Neural Network

Baochai Du and Ailing Chen

School of Management Science and Engineering, Shandong University of Finance and Economics, Erhuan East Road 7366, Jinan, China
*Corresponding Author
Email: chengchengcal@163.com

Abstract. The steady development of China's economy has led to the rapid development of the logistics industry. Nowadays, the logistics efficiency in the world has been at a high position, but compared with advanced developed countries, logistics costs are still higher. Establishing an effective logistics demand forecasting model is of great significance to reduce logistics costs and optimize the layout. This paper establishes a combined model based on the research of grey GM (1, 1) and BP neural network considering the scope of application and error. Using the combined model to fit the freight volume data of China in the past 20 years, the results show that the combined model has less error and higher precision in the forecasting of logistics demand.

1. Introduction

The development of logistics affects various industries in society, and its level of development directly affects the efficiency and quality of the overall operation of the economy. Logistics demand forecasting is the premise and foundation of all optimization schemes and decision models, and is also the support data, so logistics demand forecasting technology has received more and more attention. The current logistics demand forecasting models are mainly divided into traditional linear models and modern nonlinear models [1]. Specific classification methods can be divided into qualitative prediction and quantitative prediction, the relevant information of the former is generally non-quantitative, vague, usually with a strong subjectivity, often leading to its doubts about accuracy [2], so quantitative prediction is the mainstream method of logistics forecasting.

Grey Prediction model is one of the most important contents of grey system theory, which mainly aims at a large number of grey uncertainty prediction problems in the real world, and uses a small amount of effective data to reveal the future development trend. Professor Dong Julong first proposed the grey system theory in 1982[3][4]. After more than 30 years of development and the active expansion of many scholars, grey prediction theory has been applied in many fields such as industry, agriculture, society, economy, energy, transportation, oil and so on [5]. Neural network technology developed rapidly in the late 20th century. It has good nonlinear mapping ability, self-learning adaptability and parallel information processing ability, provides a new idea for modelling and control of uncertain nonlinear systems [6]. And its outstanding advantage is that it has strong nonlinear mapping ability and flexible network structure. Besides, the number of middle layer of the network and the number of neurons in each layer can be set arbitrarily according to the specific situation, enhancing the adaptability of the model [7].

This paper attempts to combine the grey prediction model with higher accuracy and wider application range in time series prediction with the neural network model that can independently
calculate and process the influence of each input signal on the output signal in order to integrate the advantages of time series prediction and causality prediction, making the model more accurate and adaptable.

2. Model Construction

2.1. Grey GM (1, 1) Model

The Grey GM (1, 1) model belongs to the time series prediction model, and the cumulative and decrement processing of the data is its original creation, finding out the regularity of data evolution through grey sequence generation. The model is based on the exponential fitting curve of the least square method. It can achieve better prediction results and higher accuracy without a large amount of time series data. The following is the modelling process of the grey GM (1, 1) [8].

Assume that the original time series observation data is \( X^{(0)} \)

\[
X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\}
\]  
\( (1) \)

The first-order accumulated generating operation sequence is \( X^{(1)} \)

\[
X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)\}
\]  
\( (2) \)

Where \( x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i) , \quad k = 1,2,3,\ldots,n \). Then the mean sequence is \( Z^{(1)} \)

\[
Z^{(1)} = \{z^{(1)}(1), z^{(1)}(2), \ldots, z^{(1)}(n)\}
\]  
\( (3) \)

Where \( z^{(1)}(k) = 0.5(x^{(1)}(k) + x^{(1)}(k - 1)) \). Then the basic form is as follows

\[
x^{(0)}(k) + az^{(1)}(k) = b
\]  
\( (4) \)

Where, the parameters \(-a\) and \(b\) in the GM (1, 1) model are called the development coefficient and grey action quantity respectively. Calculate \(a\) and \(b\) according to the least square method, and then construct data series \(B\) and data vector \(Y\).

\[
B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y
\]  
\( (5) \)

The time response sequence of GM (1, 1) model is

\[
\hat{x}^{(1)}(k + 1) = (x^{(0)}(1) - \frac{b}{a})e^{-ak} + \frac{b}{a} , \quad k = 1,2,\ldots,n
\]  
\( (6) \)

The restored value

\[
\hat{x}^{(0)}(k + 1) = \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k) = (1 - e^a) \left(x^{(0)}(1) - \frac{b}{a}\right) e^{-ak} , k = 1,2,\ldots,n
\]  
\( (7) \)

Where, \(\hat{x}^{(1)}(k + 1)\) is the simulation value of \(x^{(1)}(k + 1)\), \(\hat{x}^{(0)}\) is the simulation sequence of \(X^{(0)}\).

2.2. BP Neural Network Model

BP neural network system is a multi-layer network trained according to error back propagation; it can learn and store a large number of input/output relationships [9]. It uses the reverse-propagation artificial neural network to construct nonlinear functions. The simple multi-layer perceptron is shown in figure 1.
Figure 1. BP network with multi-layer perceptions. Figure 2. Reverse feedback error display.

Forward transmission as shown in figure 1, $x_1$ and $x_2$ are input layer neurons, $f(z_1), f(z_2), f(z_3), f(z_4)$ are hidden layer neurons, and $f(z_i)$ is output layer neuron, $z_i$ represents the input from the upper layer of the neuron, $w_{1i}$ represents the weight from $x_1$ to $f(z_1)$, and $w_{2i}$ represents the weight from $x_2$ to $f(z_1)$. $y_i$ represents $f(z_i)$, then the output neuron of each node can be calculated:

$$f(z_1) = w_{11}x_1 + w_{21}x_2$$
$$f(z_2) = w_{12}x_1 + w_{22}x_2$$
$$f(z_3) = w_{13}y_1 + w_{23}y_2$$
$$f(z_4) = w_{14}y_1 + w_{24}y_2$$

(8)

Reverse feedback as shown in figure 2, reverse feedback is a process in which the output layer of neural network adjusts the weights in reverse. There is an error between the output value of each neuron and the expected output value. The error is represented by $\delta_i$. When calculating it, the weight reuse of the forward propagation weight [10], so:

$$\delta_3 = w_{35}\delta_5, \delta_4 = w_{45}\delta_4, \delta_1 = w_{13}\delta_3 + w_{14}\delta_4, \delta_2 = w_{23}\delta_3 + w_{24}\delta_4$$

(9)

The purpose of reverse feedback is to calculate the weight changes in neurons. Error is an important parameter in the adjustment of weight. The change of weight adjustment is defined as follows:

$$\Delta w_i = \eta \delta_i \frac{df(z_i)}{dz_i} x_i, w'_i = w_i + \Delta w_i$$

(10)

2.3. Construction of Composite Model

Establish a linear combination model based on minimum MAPE. Assuming that the predicted sequence of grey GM (1, 1) model is $X(x_1, x_2, \cdots, x_n)$, the predicted sequence of BP neural network model is $Y(y_1, y_2, \cdots, y_n)$, the actual value sequence is $T(t_1, t_2, \cdots, t_n)$, and the error calculation adopts the mean absolute percentage error (MAPE), the goal is to make the MAPE minimum, set the combination weight are $\omega_1$ and $\omega_2$, and calculate the corresponding weights by the exhaustive method to get the minimum MAPE.

$$\min F = \frac{1}{n} \sum_{i=1}^{n} \left| \omega_1 x_i + \omega_2 y_i - t_i \right|$$

s.t. $\omega_1 + \omega_2 = 1, \omega_1, \omega_2 \in \{0,1\}$
$$x_i, y_i, t_i \geq 0 (i = 1, 2, \cdots, n)$$

(11)

Where, MAPE is:

$$\varepsilon = \frac{1}{n} \sum_{i=1}^{n} \left| \omega_1 x_i + \omega_2 y_i - t_i \right|$$

(12)

Then determine the combined weight. Determination of the combined weight uses the exhaustive method, because the constraint function is $\omega_1 + \omega_2 = 1$, that is, $\omega_2 = 1 - \omega_1$. Setting the gradient value to 0.0001, the values of $\omega_1$ are 10001 values from 0 to 1 with a tolerance of 0.0001, then $\omega_2$ is also determined. The 10001 target function values are calculated using MAT LAB calculation tools, and the values of $\omega_1$ and $\omega_2$ corresponding to the minimum objective function values are obtained.
3. Example Applications

3.1. Data Selection

According to the development of logistics industry in recent years and the unity of data standards of national statistical indicators, selecting the national freight volume of recent 20 years from the website of the National Bureau of Statistics as the forecast index of goods flow. The four indicators are selected for factor analysis—the total gross domestic product (GDP), post and telecommunications business, total retail sales of social consumer goods, residents’ consumption level. But the real influence factors may be more, the actual measurement can consider more indicators.

Table 1. Raw data

| Year | Freight volume(10 kilo-tons) | GDP(hundred million RMB) | Business total of posts and telecommunications | Total retail sales of consumer goods | Household consumption level |
|------|-----------------------------|--------------------------|-----------------------------------------------|-----------------------------------|----------------------------|
| 1997 | 1278218                     | 79715                    | 1773.29                                       | 31252.9                           | 2978                       |
| 1998 | 1267427                     | 81955.5                  | 2431.21                                       | 33378.1                           | 3126                       |
| 1999 | 1293008                     | 90564.4                  | 3330.82                                       | 35647.9                           | 3346                       |
| 2000 | 1358682                     | 100280.1                 | 4792.7                                        | 39105.7                           | 3721                       |
| 2001 | 1401786                     | 110863.1                 | 5695.8                                        | 43055.4                           | 3987                       |
| 2002 | 1483447                     | 121717.4                 | 7019.79                                       | 52516.3                           | 4606                       |
| 2003 | 1564492                     | 137422                   | 9712.3                                        | 59501                             | 5138                       |
| 2004 | 1706412                     | 161840.2                 | 9712.3                                        | 59501                             | 5138                       |
| 2005 | 1862066                     | 187318.9                 | 12028.54                                      | 68352.6                           | 5771                       |
| 2006 | 2037060                     | 219438.5                 | 15325.9                                       | 79145.2                           | 6416                       |
| 2007 | 2275822                     | 270232.3                 | 19805.03                                      | 93571.6                           | 7572                       |
| 2008 | 2585937                     | 319515.5                 | 23649.52                                      | 114830.1                          | 8707                       |
| 2009 | 2825222                     | 349081.4                 | 27193.46                                      | 133048.2                          | 9514                       |
| 2010 | 3241807                     | 413030.3                 | 31978.48                                      | 158008                            | 10919                      |
| 2011 | 3696961                     | 489300.6                 | 13333.49                                      | 187205.8                          | 13134                      |
| 2012 | 4100436                     | 540367.4                 | 15019.3                                       | 214432.7                          | 14699                      |
| 2013 | 4098900                     | 595244.4                 | 18432.24                                      | 242842.8                          | 16190                      |
| 2014 | 4167296                     | 643974                   | 21834.41                                      | 271896.1                          | 17778                      |
| 2015 | 4175886                     | 689052.1                 | 28425                                         | 300930.8                          | 19397                      |
| 2016 | 4386763                     | 743585.5                 | 43345.54                                      | 33216.3                           | 21285                      |

3.2. Prediction of Grey GM (1, 1) Model

Reference formula (5), calculating the development coefficient $a$ and grey action quantity $b$, use 1997-2016 goods traffic as an original time series observations $X^{(0)}(x^{(0)}(1), \ldots, x^{(0)}(20))$.

$$\begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y = \begin{bmatrix} -0.07865 \\ \frac{1037375.2}{1037375.2} \end{bmatrix}$$

Then the reduction sequence of the time response sequence is as follows:

$$\hat{x}(k + 1) = (1 - e^a) \left( x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} = (1 - e^{-0.07865}) \times 14467985.32e^{0.07865k}$$

Put K into the reduction sequence to get the fitting value, and calculate MAPE. Get table 2.
Table 2. Fitting values of GM (1, 1) model

| number | 1       | 2       | 3       | 4       | 5       | 6       | 7       |
|--------|---------|---------|---------|---------|---------|---------|---------|
| actual value | 1278218 | 1267427 | 1293008 | 1358682 | 1401786 | 1483447 | 1564492 |
| fitting values | 1183851.82 | 1280721.23 | 1385517.03 | 1498887.82 | 1621535.24 | 1754218.36 |
| relative error | 0 | 0.06594 | 0.00950 | 0.01975 | 0.06927 | 0.09309 | 0.12127 |
| number | 8       | 9       | 10      | 11      | 12      | 13      | 14      |
| actual value | 1706412 | 1862066 | 2037060 | 2275822 | 2585937 | 2825222 | 3241807 |
| fitting values | 1897758.36 | 2053043.61 | 2221035.17 | 2402772.74 | 2599381.10 | 2812077.06 | 3042177.01 |
| relative error | 0.11213 | 0.10256 | 0.09031 | 0.05578 | 0.00520 | 0.00465 | 0.00009 |

The MAPE of single item fitting sequence of grey GM (1, 1) model can be calculated as 6.52% by using computer calculation tools. According to Lewis (1982), MAPE is less than 10%, which is already a high-precision prediction. However, there is still a lot of room for improvement.

3.3. Prediction of BP Neural Network Model

The principle of BP neural network model belongs to the category of causal relationship prediction. The following figure 4 is hidden layer in the interval [1,100] part of the fitting lines of the neural network training, where the diamonds are the scatter diagram of the actual output value, and the lines are the neural network fitting lines. Through a wide range of trial and error experiments, the error is relatively small when the number of the hidden layer is determined between the intervals [3, 9], as shown in figure 5 below. Within this interval, it is optimal when the number of the hidden layer is 5.

Figure 4. Fitting lines of hidden layers in [1,100]. Figure 5. Fitting lines of hidden layers in [3, 9].

After 100 times of BP neural network simulation training, figure 6 shows the fluctuation of the annual simulation value. The mean value of the output of 100 times’ simulation training is taken to gain the fitting value of BP neural network model. The fitting value data are shown in Table 3.

Figure 6. Fluctuation of simulated values in 100 simulation training.
Table 3. Fitting values of BP neural network model

| number | 1      | 2      | 3      | 4      | 5      | 6      | 7      |
|--------|--------|--------|--------|--------|--------|--------|--------|
| actual value | 1278218 | 1267427 | 1293008 | 1358682 | 1401786 | 1483447 | 1564492 |
| fitting values | 1355747 | 1373237 | 1394492 | 1435233 | 1481378 | 1537169 | 1600330 |
| relative error | 0.06065 | 0.08348 | 0.07849 | 0.05634 | 0.05678 | 0.03621 | 0.02291 |

| number | 8      | 9      | 10     | 11     | 12     | 13     | 14     |
|--------|--------|--------|--------|--------|--------|--------|--------|
| actual value | 1706412 | 1862066 | 2037060 | 2275822 | 2585937 | 2825222 | 3241807 |
| fitting values | 1711099 | 1844515 | 2002015 | 2241574 | 2546280 | 2784193 | 3167702 |
| relative error | 0.00275 | 0.00943 | 0.01720 | 0.01505 | 0.01534 | 0.01452 | 0.02286 |

From the results, the MAPE of BP neural network model is 3.16%, with larger promotion in the prediction accuracy. Besides, this paper trained 100 times and took the average value to eliminate some uncertainty fluctuations, then used the combined model to improve the prediction accuracy.

### 3.4 Prediction of Combined Model

The simulated sequences fitted by grey model and BP neural network model are brought into the mathematical model (11) by using the Mat Lab tool, and the gradient is 0.0001. Then calculate the value of $\omega_1, \omega_2$ when the MAPE value is minimized: $\omega_1 = 0.1395, \omega_2 = 0.8605$. Figure 7 shows the changes of MAPE of the combined model when the $\omega_1$ takes different values. The fitting sequence of the combined model and the MAPE value are shown in table 4:

![Figure 7. Different values of $\omega_1$ corresponding to MAPE.](image)

Table 4. Fitting values of combined model

| number | 1      | 2      | 3      | 4      | 5      | 6      | 7      |
|--------|--------|--------|--------|--------|--------|--------|--------|
| actual value | 1278218 | 1267427 | 1293008 | 1358682 | 1401786 | 1483447 | 1564492 |
| fitting values | 1344932 | 1346818 | 1378621 | 1428297 | 1483821 | 1548938 | 1621798 |
| relative error | 0.052193 | 0.06264 | 0.066212 | 0.051237 | 0.058522 | 0.044148 | 0.03158 |

| number | 8      | 9      | 10     | 11     | 12     | 13     | 14     |
|--------|--------|--------|--------|--------|--------|--------|--------|
| actual value | 1706412 | 1862066 | 2037060 | 2275822 | 2585937 | 2825222 | 3241807 |
| fitting values | 1737138 | 1873605 | 2032568 | 2264061 | 2553687 | 2788083 | 3150192 |
| relative error | 0.018006 | 0.006197 | 0.002205 | 0.005168 | 0.012471 | 0.013146 | 0.028261 |

The results show that the combined prediction model has higher accuracy, higher stability and stronger adaptability compared with the prediction of single grey GM (1, 1) and BP neural network.

### 3.5 Mean Square Error Ratio Test

The average absolute percentage error is calculated in the predicted process of combined model, and the test accuracy belongs to first level. Next, the mean square error ratio test is carried out on the results of the combined prediction, and the test process is as follows:
\[ \varepsilon(k) = X^{(0)}(k) - \hat{X}^{(0)}(k) \]

\[ S_1 = \left( \frac{1}{n} \sum_{k=1}^{n} (X^{(0)}(k) - \bar{X})^2 \right)^{1/2} = 1151458.2 \]

\[ S_2 = \left( \frac{1}{n} \sum_{k=1}^{n} (\varepsilon(k) - \bar{\varepsilon})^2 \right)^{1/2} = 81380.2 \]

Where \( X^{(0)}(k) \) is the original actual sequence and \( \hat{X}^{(0)}(k) \) is the final fitting sequence of combined prediction. And \( \bar{X} = \frac{1}{n} \sum_{k=1}^{n} X^{(0)}(k) \), \( \bar{\varepsilon} = \frac{1}{n} \sum_{k=1}^{n} \varepsilon(k) \), \( k = 1,2,\ldots,20 \).

Then calculate the mean square error ratio \( C \) and the minimum error probability \( P \).

\[ C = \frac{S_2}{S_1} = 0.07068 \]

\[ P = P(|\varepsilon(k) - \bar{\varepsilon}| < 0.6745 \times S_1) = P((759237.7, 88600.9, 948232.2, 78825.1, 91244.9, 74700.6, 66515.4, 9935.8, 20748.8, 4718.3, 25510.0, 23039.9, 27929.5, 82405.4, 83264.7, 234854.8, 69420.9, 80134.2, 9206.3, 41642.7) < 776658.6) = 1 \]

It is known that the international Universal Precision Test Grade reference table, the first level of \( C \), \( P \) value requirements are \( C \leq 0.35 \), \( P \geq 0.95 \), and the \( C \) value of the combination model is insufficient 0.1, and \( P=1 \), so the prediction accuracy is primary and very high.

4. Conclusion

This paper combines the grey GM (1,1) model with the BP neural network model based on the previous research, then tries to predict the logistics demand by using the combinatorial model, and compares the single prediction results with the combined prediction results, finding that the combined prediction model can effectively predict the logistics demand. Besides, the combination model combines the advantages of time series prediction and causality prediction of two predictive types, making the combination model more efficient. It not only can combine a variety of calculation methods, but also can adapt to a variety of predictive examples. In a word, the model is applied to logistics demand forecasting, which can provide some help for the logistics industry to formulate planning plans, policies and development strategies.

5. References

[1] Zhang Lu and Muren 2015 Logistics Demand Forecasting Methods J. LogisticsSci-Tech vol 7 pp 104-106.
[2] Yangyang and Wang Xiaokun 2013 Logistics System Planning and Design (Beijing: House of Electronics Industry) pp 38-40.
[3] Deng J L 1982 Control problems of grey systems J. Systems & Control Letters 1(5) pp 288-294.
[4] Dong J L 1982 Grey Control System Journal of Huazhong Institute of Tech, 10(3) pp 9-18.
[5] Yao Tianxiang and Gong Zaiwu 2014 Grey prediction theory and its application (Beijing: Science Press) pp 1-4.
[6] Yu F and Xu X 2014 A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network J. Applied Energy 134 pp 102–113.
[7] Bizzarri F, Bongiorno M, Brambilla A, Gruouso G and Gajani G S 2013 “Model of photovoltaic power plants for performance analysis and production forecast,” IEEE Trans. Sustain. Energy, vol. 4, no. 2, pp. 278–285.
[8] Wang N, Zhao S Y and Shan X H 2018 Method of Network Public Opinion Prediction and Grading Based on Grey System Theory J. Information Studies: Theory & Application 09.
[9] Zhang Chaoli 2018 Prediction of PM_{2.5} Data in Zhengzhou Based on BP Neural Network J. Journal of Henan Institute of Education(Natural Science Edition),vol 3 pp14-17.
[10] Wang Xiaomei 2017 Introduction to Neural Networks (Beijing: Science Press) pp 82-83.