Freight Demand Analysis

Shuang LIU
School of Economics and Management, Chongqing Jiaotong University, Chongqing, CHINA

Abstract: In the current, the level of China's economic development is still lagging behind and urban development is not balanced in all aspects of the situation, so Urban logistics demand analysis has an important social and economic significance. In this paper, we use multiple linear regression forecasting method, exponential smoothing forecasting method and grey forecasting method to forecast the changing trend of freight volume from 2000 to 2014. By comparing three kinds of forecasting methods, the forecasting accuracy is sorted as follow: multiple linear regression forecasting method > grey prediction method > exponential smoothing method.

Keywords: Freight demand analysis; Multiple regression; Exponential smoothing; Grey prediction

1. Introduction

After the implementation of "10th Five-Year", "11th Five-Year", "12th Five-Year" the three Five-year plans, China's economy grows rapidly. With the economy growing, the freight volume has also increased rapidly. From 2000 to 2014, the total volume of cargo from 13billion 586million and 820 thousand tons to 43 billion 868 million tons, increasing by 3.2 times, at the annual compound growth rate of 8.73%. Behind the rapid growth in the volume of freight, there are also a wide range of problems in the regional logistics: the need to improve the logistics service quality, logistics demand structure needs continuous development, logistics demand and supply balance and flexibility still need to be improved.

Therefore, in order to carry out the effective investment, avoid excessive investment on logistics facilities and service behavior, reduce waste, it is necessary to forecast the accurate of logistics demand .So, Many scholars have made a research on freight demand forecasting.

Li Shi, Cai Guoqiang, Li Mingzhu (2008) [1]take the freight industry in Sichuan as an example, using the method of multiple regression time series to establish the prediction model, the prediction and evaluation of the model are carried out by SPSS software. Song Caiping, Han Fei (2009) [2] Select annual volume of freight, gross national products, total retail sales of social consumer goods, fixed assets investment amount, total fixed assets investment in department of transportation and telecommunications the five index factors as influence factor to conduct Multivariate linear analysis for the cargo volume of Harbin city. Zhang Shu, Yu Xia (2014) [3] use the three times exponential smoothing method to forecast the freight volume of enterprises, which provides theoretical support for the enterprise freight planning and logistics planning. Jiang Liming, Wu Ruilin (2003) [4] use the grey model for analyzing the freight demand of new towns which are lacking of basic information and get a satisfactory results. Tian Gang, Li Nan (2010) [5] think that the freight logistics system is a typical grey system, by constantly adding new information, getting rid of the old information which are of little significance, the dynamic grey prediction model which can better reflect the characteristics of the system can be established

Single prediction method can achieve the forecast of freight volume, but the fitting accuracy needs to be improved. Many scholars try to use the Markov chain to correct the prediction method to improve the prediction accuracy.

Zhang Cheng, Zhou Xiangfeng (2007) [6] combine the grey forecasting method, Markov chain and qualitative analysis, which improve the predicted results greatly. Wu Qunqi, Rui Haitian (2013) [7] by contrasting one time exponential smoothing method, two times exponential smoothing method and three times exponential smoothing method prediction results, find out that the two times exponential smoothing method should be used when the time series variation is linear, while the prediction accuracy of the three exponential smoothing method is higher when the time series variation is in the curve trend, therefore, it is better to use the three times exponential smoothing method to predict the trend of the road passenger traffic, then use Markov model to modify and improve the accuracy.

Through the research of the above scholars, it is known that all kinds of forecasting methods have different characteristics and the prediction accuracy is also different. In order to further compare the prediction accuracy of various forecasting methods, this paper has chosen the multiple linear regression prediction method, Exponential smoothing method and gray prediction method for study, to select the best fitting forecasting method of logistics demand datas.

2. Model Introduction

2.1. Multiple linear regression forecasting method

Regression analysis forecasting method mainly through the analysis of the relationship between traffic volume and their impact factors, establish the mathematical model which is based on this relationship to predict the freight volume.

The mathematical model of multiple linear regression is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon$$

In this model, y is the variable which is explained by other variables, x is Explanatory variable, $\beta_0, \beta_1 \cdots \beta_n$ is the unknown parameters of the model, Respectively represent for regression constant and Partial regression coefficients, is a random error.

The general steps of regression analysis are:

1.Comfirm the y and x in the regression equation
2.Establish the regression equation
3.Test the regression equation
4.Use the equation to predict

2.2. Exponential smoothing forecasting method

The exponential smoothing method is improved from the moving average method, and it is considered that the nearer in the time axis, the greater influence on the predicted value. The exponential smoothing method can be divided into four types, including one time exponential smoothing method, two times exponential smoothing method, three times exponential smoothing method and high order smoothing method. The most commonly used is the two smoothing method.

The formula of the two exponential smoothing method is:

$$\begin{cases} st^{(1)} = \alpha y_t + (1-\alpha) S_{t-1}^{(1)} \\ st^{(2)} = \alpha s_t^{(1)} + (1-\alpha) S_{t-1}^{(2)} \end{cases}$$

In this formula, the value of define the proportion of the new data and the original forecast value in the new predictive value, the greater the value, the greater the proportion of new data. is the forecast value for the period t. $st^{(1)}$ is one time exponential smoothing forecasting value for the period t. $st^{(2)}$ is two times exponential smoothing forecasting value for the period t.

2.3. Grey forecasting method

The main feature of gray prediction is that the model uses not the original data sequence, but the generated data sequence. This prediction method does not need a lot of datas, and can solve the problems of less historical data, low integrity of the sequence and low reliability.

GM (1,1) represent for gray model which use the first order differential equation and only include one variable.

The steps of GM (1,1) model prediction are:

1.Data checking and processing
2.Build model
3.Test predictive value
4.Forecast

3. Example forecast analysis

3.1. Multiple linear regression forecasting method

1) Determine variable

In order to predict the freight demand, select the total freight volume of the whole country as the variable y. The explanatory variables are selected by taking the following aspects into consider:

1.With the rapid economic growth, transportation infrastructure investment is also increasing, the ability to transport and the scale and scope of transport continue to upgrade. At the same time, the development of transportation promotes the economic growth. Thus, there is a very important relationship between economy and freight volume, so choose GDP as the explanatory variable.

2.With the improvement of people's living standard, the demand for material increases, and it will promote the freight transportation, in other words, it has a certain influence on freight volume, so choose the per capita income of urban residents as the explanatory variable.

3. People's production and life are inseparable from energy, such as coal and oil and other energy. Between the place where they are exploited and where they are used, there is a certain distance and need transportation to connect supply and demand, so choose the total national energy consumption as explanatory variables.

4.Internationaltrade and international investment help China achieve an unprecedented economic growth, and transport is a key factor to the realization of foreign trade. There is an indivisible relationship between foreign trade and transportation, so choose the total amount of import and export as the explanatory variable.

5.With the increase of transport demand, transport mode will continue to develop. Modern transport is divided into 5 types according to different modes of transport: road transport, rail transport, water transport, air transport and pipeline transport. The application of different modes of transport and the freight volume are not the same. The proportion of average value of each mode of transport volume in the year 2000-2014 is shown in Figure 1.

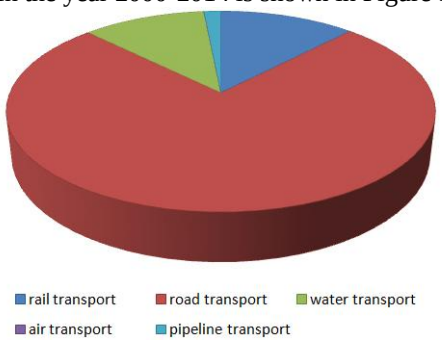

Figure 1. The proportion of5 types of transport volume

As can be seen from the figure, road, rail, water transport volume accounted for the most proportion. There is a direct relationship between the volume of freight and the number of vehicles. For road transport and water transport, the number of truck and barge can be the indicators to predict the volume of freight traffic. While when it comes to rail transport, air transport and pipeline transport, they don't have a particularly large change in the number of transport vehicles. So choose the number of truck and barge as the explanatory variable. The 2000-2014 year data of each variable are shown in Table 1.

Table 1(a). Data table of explained variables and explaining variables

year	Freight volume (10000 tons)	GDP (billion yuan)	total imports and exports (yuan)	Per capita income of urban residents (yuan)
	Y	X1	X2	X3
2000	1358682	99214.6	39273.2	6280
2001	1401786	109655.2	42183.6	6859.6
2002	1483447	120332.7	51378.2	7702.8
2003	1564492	135822.8	70483.5	8472.2
2004	1706412	159878.3	95539.1	9421.6
2005	1862066	184937.4	116921.8	10493
2006	2037060	216314.4	140974	11759.5
2007	2275822	265810.3	166863.7	13785.8
2008	2585937	314045.4	179921.5	15780.8
2009	2825222	340902.8	150648.1	17174.7
2010	3241807	401512.8	201722.1	19109.4
2011	3696961	473104	236402	21809.8
2012	4100436	519470.1	244160.2	24564.7
2013	4098900	568845.2	258168.9	26955.1
2014	4386800	636138.7	264241.8	29381

Table 1(b). Data table of explained variables and explaining variables

year	Per capita income of rural residents (yuan)	The total national energy consumption (Ten thousand tons of standard coal)	The number of truck (10,000 car)	The number of truck (one ship)
	X4	X5	X6	X7
2000	2253.4	146964	716.32	44658
2001	2366.4	155547	765.24	41457
2002	2475.6	169577	812.22	37041
2003	2622.2	197083	853.51	40457
2004	2936.4	230281	893	43846
2005	3254.9	261369	955.55	41394
2006	3587	286467	986.3	36555
2007	4140.4	311442	1054.1	34227
2008	4760.6	320611	1126.1	31943
2009	5153.2	336126	1368.6	27565
2010	5919	360648	1597.6	22783
2011	6977.3	387043	1788	21292
2012	7916.6	402138	1894.8	20282
2013	8895.9	416913	2010.6	17214
2014	9892	426000	2125.5	17003

2) Use SPSS to solve

Import data into SPSS and apply the SPSS to regression analysis

With the goodness of fit test being carried out, the adjusted R square is 0.996, which is closed to 1, so we can think the fitting degree was high.

Then conduct the significance test of regression equation. Set the significance level α as 0.05, due to the probability p values close to 0, less than α , we should refuse to the null hypothesis of the significance test of regression equation and think the regression coefficients are not 0 at the same time, the linear relationship between the explained variable and the explaining variables is significant to establish a linear model.

In the end, conduct the significance test of regression coefficients, the probability p values which are respectively resulted from the regression coefficient significance t-test of explaining variables x1, X6, X4 are higher than the significantlevel of α , we should refuse to the null hypothesis and think these partial regression coefficients have significant difference with 0, so we can get regression model:

Y=491311.414+6.485X1+912.329X6-212.027X4

3) Fitting effect

The model calculation results are shown in Table 2.The growth trend of the original value and the predicted value is shown in Figure 2.

3.2. Exponential smoothing forecasting method

Select α =0.2.Calculate the results of the model by using MATLAB, the results are shown in Table 3.The growth trend of the original value and the predicted value is shown in Figure 3.

3.3. Grey forecasting method

The calculation results of the model are obtained by using MATLAB programming, as shown in Table 4.The growth trend of the original value and the predicted value is shown in Figure 4.

Table 2. Multiple linear regression forecasting effect

Serial number	Year	Original value	Predicted value	Residual error	Relative error
1	2000	1358682	1310455.962	48226.0	0.035495
2	2001	1401786	1398835.337	2950.7	0.002105
3	2002	1483447	1487786.793	4339.8	-0.002925
4	2003	1564492	1594826.997	30335.0	-0.019390
5	2004	1706412	1720235.904	13823.9	-0.008101
6	2005	1862066	1872279.747	10213.7	-0.005485
7	2006	2037060	2033399.542	3660.5	0.001797
8	2007	2275822	2298864.124	23042.1	-0.010125
9	2008	2585937	2545866.414	40070.6	0.015496
10	2009	2825222	2858062.005	32840.0	-0.011624
11	2010	3241807	3297625.303	55818.3	-0.017218
12	2011	3696961	3711249.996	14289.0	-0.003865
13	2012	4100436	3910177.437	190258.6	0.046400
14	2013	4098900	4128448.481	29548.5	-0.007209
15	2014	4386800	4458418.596	71618.6	-0.016326

Table 3. Exponential smoothing forecasting effect

Serial number	year	Original value	st1	st2	predicted value	residual error	relative error
1	2000	1358682	1358682	1358682			
2	2001	1401786	1367303	1360406	1358682	43104	0.030749
3	2002	1483447	1390532	1366431	1375924	107523.4	0.072482
4	2003	1564492	1425324	1378210	1420657	143834.9	0.091937
5	2004	1706412	1481541	1398876	1484216	222195.8	0.130212
6	2005	1862066	1557646	1430630	1584873	277193	0.148863
7	2006	2037060	1653529	1475210	1716417	320643.5	0.157405
8	2007	2275822	1777988	1535765	1876428	399394	0.175494
9	2008	2585937	1939578	1616528	2080765	505171.6	0.195353
10	2009	2825222	2116706	1716564	2343390	481832.4	0.170547
11	2010	3241807	2341727	1841596	2616885	624922.1	0.19277
12	2011	3696961	2612773	1995832	2966889	730071.5	0.197479
13	2012	4100436	2910306	2178726	3383951	716485.3	0.174734
14	2013	4098900	3148025	2372586	3824780	274119.7	0.066876
15	2014	4386800	3395780	2577225	4117323	269477	0.061429

Table 4. Grey model forecasting effect

Serial number	Year	Original value	Predicted value	Residual error	Relative error
1	2000	1358682	1358682	0	0
2	2001	1401786	1324743	77043	0.05496
3	2002	1483447	1458378	25069	0.01689
4	2003	1564492	1605493	-41001	0.02621
5	2004	1706412	1767448	-61036	0.03577
6	2005	1862066	1945741	-83675	0.04494
7	2006	2037060	2142019	-104959	0.05152
8	2007	2275822	2358098	-82276	0.03615
9	2008	2585937	2595973	-10036	0.00388
10	2009	2825222	2857844	-32622	0.01155
11	2010	3241807	3146131	95676	0.02951
12	2011	3696961	3463500	233461	0.06314
13	2012	4100436	3812884	287552	0.07012
14	2013	4098900	4197512	-98612	0.02406
15	2014	4386800	4620940	-234140	0.05337

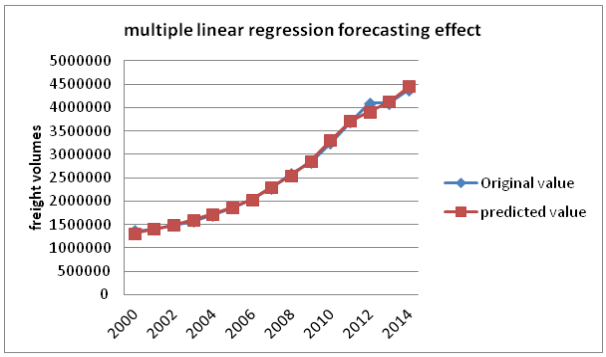

Figure 2. Comparison of original and predicted values of multiple linear regression method

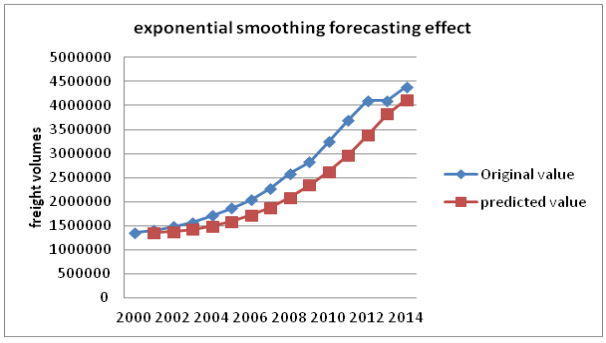

Figure 3. Comparison of original and predicted values of exponential smoothing forecasting method

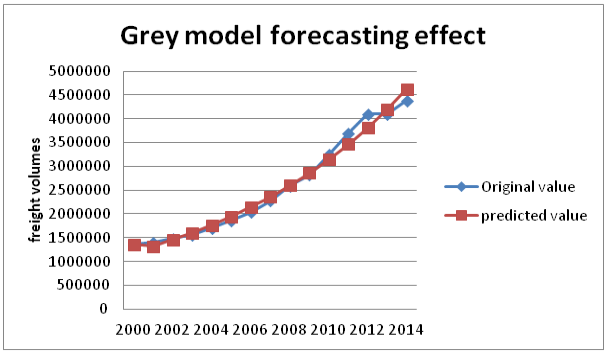

Figure 4. Comparison of original and predicted values of grey forecasting method

4. Conclusions

Through the forecast effect charts which are resulted from the three prediction method, the forecasting accuracy is sorted as follow: multiple linear regression forecasting method > grey prediction method > exponential smoothing method.

In order to further confirm the conclusion, we compare the results by using the quantitative comparison. The average relative errors of the three prediction methods respectively are: 1.36%, 13.33%, 3.48%. The Ranking of prediction accuracy result obtained is consistent with the previous result.

References

[1] Li Shi, CaiGuoqiang, Li Mingzhu. Application of multivariate regression time series method in freight forwarding market [J].logistics technology, 2008 (193): 160-162.

[2] Song Caiping, Han Fei. Analysis of the influence factors of freight volume [J]. modern commerce and trade industry, 2009 (6): 29-30.

[3] Zhang Shu, Yu Xia. [J]. forestry machinery and woodworking equipment based on exponential smoothing method in forecasting the freight volume of enterprises, 2014 (1): 55-57.

[4] Jiang Li Ming, Wu Ruilin. Freight demand forecasting for new towns in the gray model industry [J]. Journal of Huazhong University of Science and Technology (Urban Science Edition), 2003 (1): 65-67.

[5] Zhang Cheng, Zhou Xiangfeng. Prediction of railway freight volume based on grey prediction Markov chain - qualitative analysis [J]. Journal of China Railway, 2007 (5): 15-21.

[6] Wu Qunqi, Yuan Huazhi, FengZhongxiang, Zhu Wenying,. Based on exponential smoothing and Markoff model Prediction method of highway passenger traffic volume [J]. Journal of traffic and transportation engineering, 2013 (4): 87-93.

[7] Si Shoukui, Mathematical modeling algorithms and applications. Beijing: National Defence Industry Press, 2013.