

Research on Feight Volume Prediction in Sichuan and Chongqing based on Multiple Regression and Exponential Smoothing Method

Lingyuan Tian¹, Zhihao Zhang², Xinyue Wang³

¹School of Modern Postal, Chongqing University of Posts and Telecommunications; China

²School of Modern Postal, Chongqing University of Posts and Telecommunications; China

³Chongqing No.2 Foreign Language School, China

Abstract: As the starting point of the national "One Belt One Road" development and the important economic belt of the Yangtze River in the central and western regions of the country, the Sichuan-Chongqing region is of great significance to study the influencing factors of its logistics. Freight volume is an important indicator for measuring the level of logistics development in a country or region. This article will use SPSS software to analyze the freight volume data of Sichuan-Chongqing region from 2005 to 2019, and use the quadratic exponential smoothing method to model. It is predicted that the freight volume of Sichuan-Chongqing region will continue to grow in the next five years, with a growth rate of approximately 5%. In 2024, the freight volume in Sichuan and Chongqing will reach 3.6 billion tons.

Keywords: Freight volume; Regression analysis; Time series analysis; Forecast.

1. Introduction

Since the reform and opening up, especially with the continuous advancement of the western development strategy, the Sichuan-Chongqing region has developed rapidly in economy and society, and its comprehensive strength has been continuously improved. It is in the forefront of my country's western region. With the support of the national economic development strategy, the Sichuan-Chongqing region has new development advantages, but also faces many unknown challenges. In the process of studying logistics demand, one of the important indicators is freight volume. The accurate forecast of freight volume can not only provide necessary reference value for the development of local logistics economy, but also an important basis for analyzing the level of logistics development and future trends of a country or region. Therefore, the study of freight volume in the Sichuan-Chongqing region is particularly important for the development of the region.

The logistics development in Sichuan and Chongqing has unique geographical advantages. As one of the most important provincial capitals in the core area of China's western region, Sichuan Province is a bridgehead that connects the east and the west. Chongqing is located in the upper reaches of the Yangtze River and the central and western regions, and it is one of the few large cities in the southwest that can use waterways for transportation. In addition, the Sichuan-Chongqing region has initially formed a multi-faceted large logistics pattern with roads, railways, aviation, water transportation, and pipeline transportation interconnected with each other^[1]. However, the outdated, low-modernization and low-specialization of logistics facilities in Sichuan-Chongqing area restrict the development of Sichuan-Chongqing logistics industry.

Therefore, this article uses SPSS data processing software to use multiple linear regression analysis, time series analysis and other analytical methods to explore the total freight volume and the main influencing factors of the freight volume

in the Sichuan-Chongqing region in the past 15 years. A reasonable forecast of the future freight volume is made by establishing a time series model. The significance of the research is as follows:

1. Effectively promote regional economic integration.

The integration of regional economy can greatly improve the utilization rate of production resources and technology between regions and even regions. Logistics as the basis and conditions for economic development, the main factors in Sichuan and Chongqing freight analysis of the regional economy and the whole western Sichuan and Chongqing have guiding significance.

2. Provide a feasible basis for logistics planning and layout

Establishing a reasonable forecasting model and quantitatively predicting the development trend of freight volume is conducive to the rational planning and layout of logistics resources and technology, and can also provide reference value for relevant policy makers, which is conducive to promoting the future development of my country's logistics industry.

2. Regression Model Test

Regression analysis method refers to the use of data statistics principles to mathematically process a large number of statistical data, and determine the correlation between dependent variables and certain independent variables, and establish a regression equation with good correlation to predict future changes.

In the selection of the important factors affecting the freight volume in Sichuan and Chongqing, this article combines the research results of several scholars on the most important factors affecting the freight volume, and selects seven indicators after multiple considerations^[2]. The seven indicators are the output value of the primary, secondary and tertiary industries, the total retail sales of consumer goods, the length of transportation routes, the total import and export of goods and GDP. In order to further explore the degree of each impact on the freight volume, the SPSS25.0 data statistical

analysis software is used to analyze the correlation of the above-mentioned variable data after the pretreatment. The above variables are defined as x_1, \dots, x_7 , and the freight volume is

defined as y . Obtain the correlation coefficient table between the variables, as shown in **Table 1**.

Table 1. Correlation coefficient table

	y	x_1	x_2	x_3	x_4	x_5	x_6	x_7
y	1	0.923	0.947	0.864	0.874	0.941	0.507	0.914
x_1	0.923	1	0.994	0.984	0.991	0.928	0.699	0.994
x_2	0.947	0.994	1	0.965	0.978	0.939	0.671	0.992
x_3	0.864	0.984	0.965	1	0.996	0.875	0.745	0.982
x_4	0.874	0.991	0.978	0.996	1	0.895	0.749	0.991
x_5	0.941	0.928	0.939	0.875	0.895	1	0.568	0.925
x_6	0.507	0.699	0.671	0.745	0.749	0.568	1	0.676
x_7	0.914	0.994	0.992	0.982	0.991	0.925	0.676	1

It can be found from **Table 1** that, except for the correlation coefficient of total import and export of goods, the correlations of other variables are close to 1. This is related to the geographical characteristics of the Sichuan-Chongqing region. Because it is located inland, there is no large seaport, and the cost of import and export of goods is high, the correlation with freight volume is not as obvious as other variables, but there is still a moderate correlation. On the other hand, it is easy to find that the correlation coefficient between

the independent variables is relatively high overall through the correlation coefficient table. It shows that there is multicollinearity among various variables.

The significance test of the regression equation is performed to determine whether it is appropriate to describe the relationship between the respective variables and the freight volume with a linear model^[3]. Perform F tests on the data:

Table 2. Model summary

R	R squared	Adjusted R squared	Standard estimation error
.993a	.986	.972	11081.403

From **Table 2**, it can be seen that $R = 0.993$, $R^2 = 0.986$, indicating that the goodness of fit of the model is high, and there is a strong linear correlation between the explanatory variables and the explained variables.

Checking the goodness of fit of the regression equation shows that when 1, the degree of freedom of the molecule is 7. Query the F distribution table to get $F_{\alpha=0.05}(7,7)=3.787 < 70.887$, so the H_0 hypothesis is rejected, indicating that the model is very significant. Through t tests to analyze the regression coefficient of the

model, the regression equation is obtained:

$$y = -0.253x_1 + 2.138x_2 + 1.992x_3 - 3.315x_4 + 0.319x_5 + 0.024x_6 + 0.063x_7 \quad (1)$$

Carry out t test on the coefficient value in this model, make its significance level $\alpha = 0.05$, and according to the degree of freedom of 7, check the t distribution table to get that $t_{\alpha/2} = t_{0.025} = 2.365$. The t -test values of the regression coefficients of all variables are less than 2.365, indicating that all variables have not passed the test.

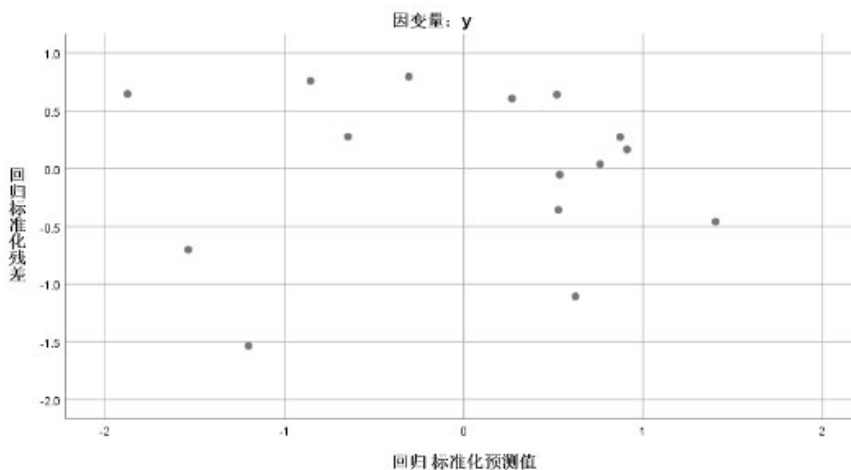


Figure 1. Residual error graph

From **Figure 1**, it can be clearly seen that the variance of the residuals decreases with the increase of the explanatory variables, and the overall appearance is a clear bell mouth. So we can determine that the regression model has

heteroscedasticity.

By judging the value of the variance expansion factor to determine whether the model has collinearity. Get **Table 3**:

Table 3. Collinearity diagnosis table

	Collinearity statistics	
	Tolerance	VIF
x_1	.002	514.417
x_2	.003	338.588
x_3	.005	217.466
x_4	.001	1064.714
x_5	.076	13.192
x_6	.092	10.844
x_7	.002	636.210

According to **Table 3**, it is obvious that the tolerance of the seven explanatory variables is very small, the largest is 0.092, and the smallest is only 0.002. Moreover, the value of the variance expansion factor exceeds 10, which indicates that there is serious multicollinearity among the seven explanatory variables in the model.

for forecasting the short-term economic development trend. It is an analysis and forecasting method developed on the basis of the moving average method^[4]. It can be seen from the time series chart of the freight volume in Sichuan and Chongqing from 2005 to 2019 that there is a clear upward trend, so it is more appropriate to use the second smoothing method in the exponential smoothing method for extrapolation.

3. Exponential Smoothing Analysis

The exponential smoothing method was a common method

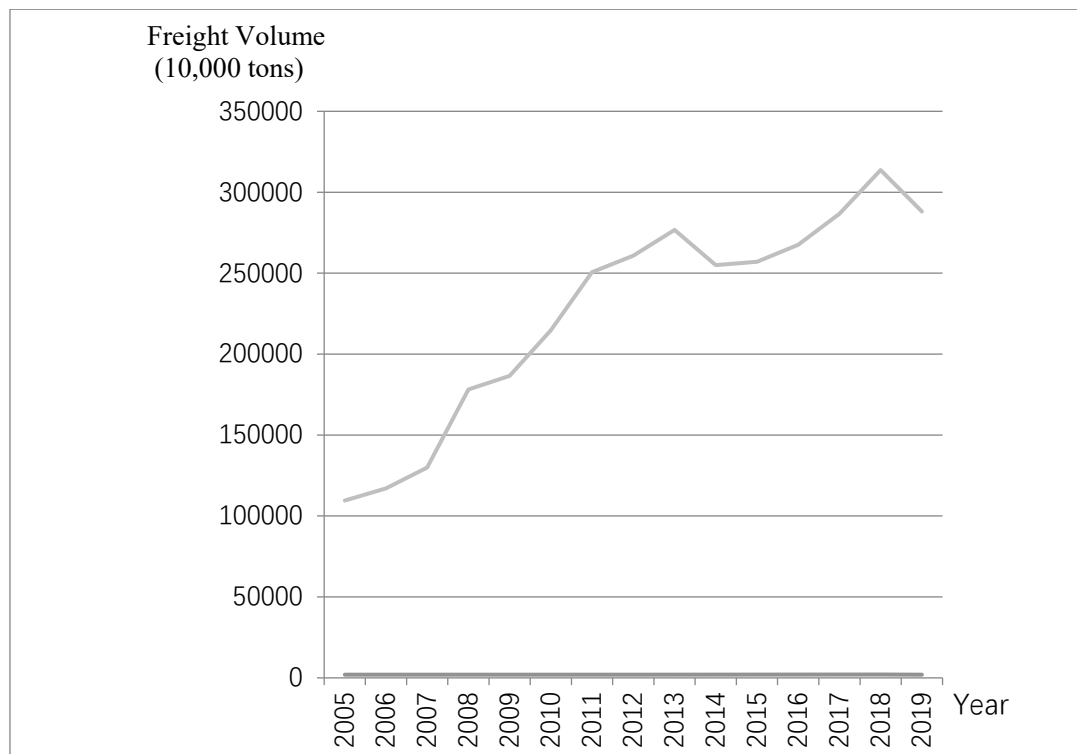


Figure 2. Time sequence diagram of freight volume from 2005 to 2019

Through multiple experiments, select the α value with the smallest average absolute error.

$$MAE = \frac{1}{n} \sum_{i=1}^m |(y_i - y'_i)| = 3735.11 \quad S_t^{(1)} = \alpha y_t + (1 - \alpha) S_{t-1}^{(1)} \quad (2)$$

Use the quadratic exponential smoothing formula:

$$S_t^{(1)} = \alpha y_t + (1 - \alpha) S_{t-1}^{(1)} \quad (1)$$

$$S_t^{(2)} = \alpha y_t + (1 - \alpha) S_{t-1}^{(2)} \quad (2)$$

Can be calculated:

$$S_{2019}^{(1)} = 290346.57, S_{2019}^{(2)} = 292112.70 \quad (3)$$

So the prediction equation is

$$a_{2019} = 2f_{2019}^1 - f_{2019}^2 = 288580.44 \quad (4)$$

$$b_{2019} = \frac{0.9}{0.9-1}(f_{2019}^1 - f_{2019}^2) = 15895.15 \quad (5)$$

When $m = 1, 2, 3, 4, 5$, the predicted value of freight volume in Sichuan and Chongqing from 2020 to 2024 is shown in Table 4.

Table 4. Forecast value of freight volume from 2020 to 2024 (ten thousand tons)

Years	m	f_{t+m}
2020	1	304475.59
2021	2	320370.73
2022	3	336265.88
2023	4	352161.02
2024	5	368056.17

4. Analysis of Prediction Results

According to the quadratic exponential smoothing method, when the smoothing coefficient is 0.9, it is predicted that the freight volume in Sichuan-Chongqing region will reach 304,755,900 tons in 2020, 3203,707,300 tons in 2021, 336,658,800 tons in 2022, and 352,161,200 tons in 2023. In 2024, it will reach 36805617 tons. According to the

influencing factors of the freight volume in Sichuan and Chongqing in the past 15 years, the steady growth of the freight volume in the next five years may be due to the growth of the output value of the secondary industry in Sichuan and Chongqing. At the same time, it brings positive effect to the development of other industries. In addition, in terms of transportation modes, with the national focus on the western region and policy support in recent years, the transportation network in Sichuan-Chongqing region has been gradually improved, the frequency of cargo transportation has increased year by year, and the choice of combined transportation has become more flexible, promoting the rapid development of freight volume[5].

References

- [1] Liu Z. Research on the Economic Linkage and Spatial Structure of the Chengdu-Chongqing Double-city Economic Circle—Based on the Perspective of the Strength of Economic Linkage and Urban Flow[J]. Journal of Sichuan Administration Institute, 2021(02): 77-87.
- [2] Zhang Y. The Three Gorges Freight Volume Demand Forecast Method Based on Variable Scale Regression[J]. Journal of Wuhan University of Technology (Transportation Science and Engineering Edition), 2020, 44(05): 938-942.
- [3] Zeng L. Forecast of Heilongjiang Grain Freight Volume Demand Based on Exponential Smoothing Method[J]. Research on Land and Natural Resources, 2015(01):17-22.
- [4] Zhang H. Analysis and forecast of national freight volume based on exponential smoothing method[J]. Bohai Economic Outlook, 2020(01):162.
- [5] Cai W, Peng Yi, Chen Q. Forecast of Air Transport Passenger Volume Based on Multiple Regression Model[J]. Aviation Computing Technology, 2019, 49(04): 50-53+58.