# Influencing Factors Analysis and Forecast of Regional Logistics Demand: a case of Wuhan

## Rongyan Zhu[1,2,a*] Lizhe Shi[3,b] and Shufan Zhu [1,c]

[1]College of Logistics Wuhan Technology and Business University, No. 3 Huangjiahu West Road, Baishazhou Avenue, Wuhan, China 430065

[2]Hubei Business Service Research Development Center, No. 3 Huangjiahu West Road, Baishazhou Avenue, Wuhan, China 430065

[3]College of Science & Arts of Jianghan University, No.1 Shenghai Avenue, Huangpi District, Wuhan, China 430062

[a]13007132492@163.com, [b] 38257022@qq.com, [c]307947682@qq.com
* The Corresponding Author

**Keywords:** Regional logistics; Logistics demand; Influencing factors; Demand forecast

**Abstract.** Regional logistics demand has an important guiding significance for regional logistics development planning. However, due to the wide range of areas and departments involved in logistics activities, and the corresponding statistics is unavailable. Therefore, the demand for regional logistics is much more difficult to determine. Through the influencing factors analysis of regional logistics, this paper chooses several indicators and constructs a multiple linear regression model. Then an empirical analysis of Wuhan in China is carried out and some indicators are eliminated. Finally, a regression model is established with gross domestic product, gross agricultural output and total retail sales of consumer goods. The model has statistical significance and it is used to predict the logistics demand of Wuhan in 2020, 2025 and 2030. The results show that the region's logistics demand will continue to grow.

## Introduction

In recent years, China's logistics demand has increased rapidly. Various local governments have introduced various incentive policies. The logistics industry has ushered in a good opportunity for development and has been developing well and rapidly. All kinds of logistics infrastructure are built up in a short time. But due to the lack of adequate analysis and research, the logistics infrastructure is repeated constructed and does not meet the actual logistics demand. This will certainly not conducive to the sustainable development of the logistics industry. Therefore, analyzing the influencing factors of regional logistics demand, and scientifically and rationally forecasting the logistics demand is the basis for formulating the regional logistics development policy and the logistics system construction planning [1], an important guarantee for the rational allocation of the logistics resources and the sustainable and healthy development of the logistics industry.

## Influencing Factors Analysis of Regional Logistics Demand

Because logistics activities involve a wide range of fields and departments, and the connotation and denotation of statistical indicators in those fields and departments are different. There is no statistical data on logistics volume at present, and there are no other statistical indicators and statistical data that can directly reflect the scale of logistics. At meanwhile, the larger circulation range of commodities, the growing market share of small enterprises, the complexity of sales market, the rapid response of enterprises, and the requirements of logistics strategy such as JIT, that leads the regional logistics demand to become more and more scattered [2]. So, it is difficult for the logistics demand statistics. Most studies use freight volume and freight turnover to reflect the logistics demand, and a few studies directly use the specific volume of goods as the logistics demand. For example, literature [3] uses the consumption of meat, aquatic products, fruits, vegetables and other refrigerated commodities as logistics needs, which is of great significance for the study of logistics demand in a specific field. But regional logistics demand generally covers a wider range of demand fields. Therefore, this paper will use more commonly used indicator that is

freight volume (FV, $y$ , 10000 tons) to measure the local logistics demand.

Considering that the demand for logistics is derived from other industries, most studies find the law of common change by establishing the link between logistics demand and its influencing factors. According to this law, logistics demand is predicted through a certain forecasting method. Wang[4] selected seven factors, including GDP growth, total import and export and total retail sales of consumer goods and so on, and established a multi-factor grey forecasting model, which is applied to the empirical research of logistics demand prediction. Qiu etal. [5] established the system dynamic model between urban logistics subsystem and production subsystem, civil life subsystem, natureal environment subsystem, and chose GDP, freight volume, technology, production, finished goods inventory, order production, population, urban pollution, congestion and other indicators to carry out the experiment. The results showed that average transport capacity and average fuel use are the indirect causes of urban logistics demand dynamics in Beijing, but all have a direct impact on the city environment. Yang etal.[6] subjectively chose several influence factors of logistics demand, including the output of the primary, secondary and tertiary industries, regional retail sales, regional total import and export, residents' consumption level and so on, then the chaos optimization method was applied to select indicators for support vector machine to predict the logistics demand in Shanghai. Chen etal. [7] divided logistics demand into three parts: agricultural logistics demand, manufacturing logistics demand and commercial logistics demand. Then the influencing factors of logistics demand in each industry were analyzed separately to predict logistics demand. Miao etal [8] believed that the conditions of weather, humidity, and temperature will affect logistics demand and predicted the dynamic logistics demand based on these indicators.

Generally speaking, regional logistics demand is closely related to the development of regional economy and industry. First of all, regional logistics demand is inevitably affected by the level of regional economic development, and this effect is stable in the short term and complies with certain laws in the long term. Therefore, there must be a certain functional relationship between the demand of regional logistics and the level of economic development. Secondly, the demand for logistics comes mainly from the displacement caused by the consumption of industrial and agricultural products, and the distribution of social consumer goods. Therefore, the scale of logistics demand should also be directly related to industrial, agricultural and commercial activities. Combined with the reference to the literature and taking the availability of data into consideration, regional economic indicators, industrial structure indicators, regional consumption level indicators and other indicators are chosen to study their impact on regional logistics demand. These indicators include gross domestic product (GDP, $x_1$ ,100 million yuan), residents' consumption level (RCL, $x_2$ , yuan), population (POP, $x_3$ ,10 thousand people), total above scale industrial output (TSIO, $x_4$ ,100 million yuan), gross agricultural output (GAO, $x_5$ ,100 million yuan), gross output of construction industry (GOCI, $x_6$ ,100 million yuan), total retail sales of consumer goods (TRSCG, $x_7$ ,100 million yuan), total import and export (TIE, $x_8$ ,100 million yuan).

## Model Construction and Calculation Steps

The forecasting methods of regional logistics demand are very many. Basically they can be divided into time series prediction method and causality prediction method [9]. Time series prediction method is a kind historical data extension prediction method through the collection and analysis of time series. According to the development process and trend reflected in the time series, this method predicts the future potential level by analogy. The commonly used methods are GM (1, 1) and so on. Causality prediction method mainly seeks for the causal relationship and interaction relationship among things' development by system analysis. On this basis, function relation is established to predict the development and change of things. Specific functional relationships include one element regression, multiple regression or linear and nonlinear forms. In order to better understand the relationship between local logistics demand and its various influencing factors, the multiple linear regression method is selected to analyze in this paper. The multi linear regression equation can be

established between the regional logistics demand and the influencing factors, as shown in Eq. 1.

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_8 x_8 + \varepsilon \qquad (1)$$

In which, $b_0$ is a constant term, $b_1, \cdots, b_8$ are partial regression coefficients, which mean that the average change of logistics demand when each independent variable changes by one unit with other factors unchanged. $\varepsilon$ is the residual, which is the difference between the measured value($y$) and the estimated value($\hat{y}$) of the freight volume.

In order to avoid the excessive correlation of some variables in the independent variable, those less important independent variables are gradually removed. The specific calculation steps are as follows:

Collect data and construct multiple linear regression models.

SPSS is used to operate models, and the way of independent variables entering into the model is set to backward method in SPSS. All variables that meet the rejection criteria are excluded from the equation until no variables could be eliminated, so as to filter the independent variables of the final multiple linear regression equation.

The final linear regression model is determined by SPSS, and the model is analyzed according to the test criteria.

The future value of the respective independent variables in the linear regression equation is determined by the average growth rate. Then the values are used to predict the regional logistics demand.

## Empirical Analysis

Based on the foregoing analysis, the paper chooses Wuhan in China to carry out an empirical analysis. Wuhan is a big city in Central China and one of the core cities in the Yangtze River economic belt. Its excellent geographical location provides an advantage for its development of an comprehensive transportation hub and modern logistics. The city is the only city in Central China that can directly sail the four continents in the world at present. In the city's "13th Five-Year Plan" of logistics industry, its regional logistics development was positioned to become the national logistics center in 2020. The logistics industry added value of the city exceeded 100 billion yuan in 2015. Looking up the Wuhan statistical yearbook from 2003 to 2017, the original data from 2002 to 2016 obtained are shown in Table 1. The last row of Table 1 shows the calculated average growth rate(AGR) of each variable.

Table 1　Raw data and average growth rate of each variable

| | $y$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ |
|---|---|---|---|---|---|---|---|---|---|
| 2002 | 15942.6 | 1467.8 | 7820.28 | 823.7 | 1020.84 | 141.25 | 251.52 | 798.64 | 182.3 |
| 2003 | 16609.8 | 1622.18 | 8524.56 | 836.8 | 1138.98 | 151.79 | 382.29 | 885.69 | 259.8 |
| 2004 | 17044.8 | 1882.24 | 9564.05 | 845.43 | 1334.49 | 165.65 | 580.77 | 996.22 | 355 |
| 2005 | 19611.7 | 2261.17 | 10849.72 | 858 | 1678.34 | 180.6 | 728.98 | 1128.64 | 506.7 |
| 2006 | 20817.75 | 2679.33 | 12359.98 | 875 | 2286.69 | 191.21 | 909.15 | 1293.33 | 641.2 |
| 2007 | 22554.93 | 3209.47 | 14357.64 | 891 | 2818.59 | 215.92 | 1246.02 | 1518.3 | 761.8 |
| 2008 | 29142.98 | 4115.51 | 16712.44 | 897 | 3525.63 | 244.64 | 1611.53 | 1850.05 | 984.3 |
| 2009 | 34409.19 | 4620.86 | 18385.02 | 910 | 4338.28 | 251.79 | 1967.61 | 2164.09 | 780.5 |
| 2010 | 40287.93 | 5565.93 | 20806.32 | 978.54 | 5032.18 | 281.09 | 2424.52 | 2570.4 | 1227.3 |
| 2011 | 41804.45 | 6762.2 | 23738.09 | 1002 | 6424.59 | 329.49 | 3180.5 | 3031.79 | 1485.1 |
| 2012 | 43892.49 | 8003.82 | 27061 | 1012 | 9018.88 | 476.04 | 4040.7 | 3467.37 | 1285.1 |
| 2013 | 44528.75 | 9051.27 | 29821.22 | 1022 | 10394.07 | 530.27 | 4879.24 | 3916.6 | 1349.6 |
| 2014 | 48529.99 | 10069.48 | 33270.39 | 1033.8 | 11764.59 | 559.44 | 5840.5 | 4573.54 | 1624.2 |
| 2015 | 48185.19 | 10905.6 | 36436 | 1060.77 | 12374.92 | 620.28 | 6016.26 | 5102.24 | 1749.3 |
| 2016 | 49981.81 | 11912.61 | 39737 | 1076.62 | 13159.09 | 669.65 | 6980.37 | 5610.59 | 1570.1 |
| AGR | 0.088 | 0.163 | 0.123 | 0.019 | 0.204 | 0.121 | 0.274 | 0.150 | 0.188 |

**Model establishment.** The correlation coefficients between the freight volume and other variables are shown in Table 2. It can be seen from table 2 that there is a significant correlation between the freight volume and the chosen indicators, which provides a certain basis for the multiple linear regression model.

Table 2    Correlation coefficients between indicators

|  | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ |
|---|---|---|---|---|---|---|---|---|
| $y$ | 0.962** | 0.961** | 0.982** | 0.945** | 0.917** | 0.941** | 0.950** | 0.967** |

**. Correlation is significant at the 0.01 level (2-tailed).

In accordance with the preceding calculation steps, SPSS is used to remove the indicators according to the standards that P value is over 0.1. The system carried out 6 times to fit, successively excluding residents' consumption level, population, total above scale industrial output, total import and export and gross output of construction industry. In the end, there are three indicators that are gross domestic product, gross agricultural output and total retail sales of consumer goods. Using these three indicators to carry on multiple linear regression, the multiple linear regression model is obtained as shown in Eq. 2.

$$y = 19410.913 + 14.253x_1 - 99.542x_5 - 12.807x_7 + \varepsilon$$

（2）

**Model Analysis.** The fitting condition of the model is shown in Table 3. Table 3 shows that the determination coefficient (R Square) is 0.989, and the adjusted determination coefficient (Adjusted R Square) is 0.986, which indicates that the linear regression model fits well and has statistical significance. The test statistic of Durbin-Watson is 2.522. The number of variables k in regression model (including constant) equals 4, the number of samples n equals 15. Based on these two value, $d_L = 0.685, d_U = 1.977$ can be obtained by checking DW statistics table of critical values. Combined with the criteria of DW values (see Table 4), the test gets no conclusion, namely cannot determine whether there is autocorrelation between random errors or not. The result may be due to the sample size being too small. F value is 323.056 and P value is 0.000. In summary, at a significant level of 0.05, there is a linear relationship between regional logistics demand and GDP, GAO and TRSCG.

Table 3    Model Summary

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | | Durbin-Watson |
|---|---|---|---|---|---|---|---|---|---|---|
|  |  |  |  |  | R Square Change | F Change | df1 | df2 | Sig. F Change |  |
| 1 | 0.994 | 0.989 | 0.986 | 1,568.0231 | 0.989 | 323.056 | 3 | 11 | 0.000 | 2.522 |

Table 4    Test criteria for DW values

| DW value | Test Result |
|---|---|
| $4 - d_L < DW < 4$ | Negative hypothesis, negative correlation exists |
| $0 < DW < d_L$ | Negative hypothesis, positive correlation exists |
| $d_U < DW < 4 - d_U$ | Accept the hypothesis, no autocorrelation exists |
| $d_L < DW < d_U$ or $4 - d_U < DW < 4 - d_L$ | No conclusion, undetermined |

The test of partial regression coefficients in the linear regression model is shown in Table 5. Table 5 shows that the partial regression coefficient P values of GDP, GAO and TRSCG are 0.000, 0.000 and 0.007 respectively by T-test. Those values indicate that at a given significant level of 0.05, the linear effect of each variable on the dependent variable has statistical significance. However, The partial regression coefficient of GAO and TRSCG is negative, which indicates that the growth of these two indicators will lead to a decline in the freight volume. This result does not seem to be consistent with our common sense, and there may be two possible reasons for this result. On the one hand, the growth of GAO and TRSCG reflects the growth of local demand for agricultural products and consumer goods. However, this demand may be mainly accomplished through distribution by some private individuals or self-employed persons, and freight volume is usually the statistics for the regional freight traffic of those regular enterprises entering and leaving the region. Therefore, the completion of logistics demand for agricultural products and consumer goods is not completely reflected in statistics. On the other hand, there are still some problems to use freight volume instead of regional logistics demand. Freight transport is only a part of the logistics activities. Although it takes a large share and the indicator can roughly reflect the change of logistics demand. It still can not completely replace the regional logistics demand. Therefore, find a more suitable indicator to reflect logistics demand is still very necessary.

Table 5    Coefficients

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 19,410.913 | 1,101.500 | | 17.622 | 0.000 |
| | $x_1$ | 14.253 | 1.765 | 3.920 | 8.073 | 0.000 |
| | $x_5$ | -99.542 | 15.402 | -1.402 | -6.463 | 0.000 |
| | $x_7$ | -12.807 | 3.843 | -1.575 | -3.333 | 0.007 |

**Regional Logistics Demand Forecast.** According to the multiple linear regression model shown in Eq. 2, the predicted value of regional logistics demand can be obtained by substituting GDP, GAO and TRSCG in each year. Taking into account the average growth rate of GDP, GAO and TRSCG, their value of 2020, 2025 and 2030 can be estimated. By substituting them into Eq. 2, the corresponding regional logistics demand of each year can be gotten too. Specific prediction results and deviations are shown in Table 6. From table 6, it can be seen that the logistics demand in Wuhan will continue to increase in the future, and the demand and requirements for logistics infrastructure in the region will also be raised accordingly.

Table 6    Prediction results and deviations of logistics demand in Wuhan

| Year | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|
| y | 15942.6 | 16609.8 | 17044.8 | 19611.7 | 20817.75 | 22554.93 |
| $\hat{y}$ | 16042.78 | 16079.11 | 16990.50 | 19207.28 | 22001.93 | 24217.07 |
| ε | -100.18 | 530.69 | 54.30 | 404.42 | -1184.18 | -1662.14 |
| Year | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 |
| y | 29142.98 | 34409.19 | 40287.93 | 41804.45 | 43892.49 | 44528.75 |
| $\hat{y}$ | 30023.15 | 32492.22 | 37841.99 | 44165.39 | 41695.63 | 45473.34 |
| ε | -880.17 | 1916.97 | 2445.94 | -2360.94 | 2196.86 | -944.59 |
| Year | 2014 | 2015 | 2016 | 2020 | 2025 | 2030 |
| y | 48529.99 | 48185.19 | 49981.81 | - | - | - |
| $\hat{y}$ | 48668.72 | 47758.64 | 50686.60 | 85724.63 | 129522.16 | 173319.69 |
| ε | -138.73 | 426.55 | -704.79 | - | - | - |

## Conclusion

In this paper, Wuhan is taken as an example to analyze the regional logistics demand. The results show that GDP, GAO and GRSCG in the chosen eight indicators were selected by a certain standard. The multi linear regression model was ultimately determined by the three indicators as independent variables and the freight volume as the dependent variable. The model is significant, and the linear effect of each independent variable on the dependent variable also has statistical significance. Through the model analysis, it is found that the sample capacity needs to be increased, and there is a limitation to replace the regional logistics demand by freight volume. According to the established regression model, the logistics demand of Wuhan in 2020, 2025 and 2030 is predicted. The forecast results indicate that the logistics demand in this region will continue to grow in the future.

## Acknowledgment

## References

[1] Xiaoyan Xie, Xueting Wei, Lin Wang, Prediction of logistics demand based on exponential smoothing model for Hohhot, Baotou and Ordos, Journal of Arid Land Resources and Environment. 27(2013) 58-62.

[2] Guilbault M, Gouvernal E, Transport and logistics demand: new input from large surveys of shippers in France, Transportation Research Record: Journal of the Transportation Research Board. 2168 (2010) 71-77.

[3] Zhang Y, Xu H, Zheng Y, Chinese residents' cold chain logistics demand forecasting based on GM (1, 1) model, African Journal of Business Management. 14 (2012) 5136-5141.

[4] Xiaoli Wang, Logistics demand forecast based on multi-factor grey model, Statistics & Decision. 14 (2013) 86-87.

[5] Qiu Y, Shi X, Shi C, A system dynamics model for simulating the logistics demand dynamics of metropolitans: A case study of Beijing, China, Journal of Industrial Engineering and Management. 3(2015) 783-803.

[6] Yang H, Zhou Y, Liu H, Chaos optimization SVR algorithm with application in prediction of regional logistics demand//International Conference in Swarm Intelligence. Springer, Berlin, Heidelberg, (2010) 58-64.

[7] Chen S, Gan M, Tang Y, Analysis of predicting the diversity regional logistics demand based on SVR: the case of Sichuan in China, Applied Mathematics & Information Sciences. 2(2013) 645-651.

[8] Miao X, Xi B, Agile Forecasting of Dynamic Logistics Demand, Transport: Research Journal of Vilnius Gediminas Technical University and Lithuanian Academy of Sciences, 1(2008) 26-30.

[9] Peng L. Individual Choice and Reputation Distribution of Cooperative Behaviors among Heterogeneous Groups. Chaos, Solitons & Fractals, 2015, 77: 39-46.

[10] Hu Z, Zhang Y, Yao L, Radial basis function neural network with particle swarm optimization algorithms for regional logistics demand prediction, Discrete Dynamics in Nature and Society, Article ID 414058, 13 pages, (2014). doi:10.1155/2014/414058.