ISSN 2616-5775 Vol. 6, Issue 1: 26-31, DOI: 10.25236/AJCIS.2023.060105

Analysis and Forecast of Cross-border E-commerce Scale in Guangdong Province Based on Grey Model

Min Liang*, Minjing Peng

School of Economics & Management, Wuyi University, Jiangmen, China
**Corresponding author*

Abstract: *With the advent of the era of big data, data analysis and cloud computing have become important tools for analysis and prediction. With the rapid development of the Internet, Guangdong Province, as a big province opening to the outside world, the analysis and prediction of cross-border e-commerce based on the grey model is an important way and basis to improve the business in Guangdong Province. Firstly, the important factor is calculated from affecting scale factors cross-border e-commerce by the grey correlation degree. Further, based on the factor, a grey prediction model is established to predict the cross-border e-commerce development in Guangdong Province. The results show that the local financial allocation for science and technology is an important factor affecting cross-border e-commerce development in Guangdong Province. Moreover, the grey prediction model with a relative error of 8.2% proves that the impact will continue to grow in the future.*

Keywords: *cross-border e-commerce, grey correlation degree, grey prediction, data analysis*

1. Introduction

It has become indispensable to analyze its past development and forecast its future growth trend with scientific methods. As the largest foreign trade province in China, Guangdong plays an important role in the cross-border e-commerce circle. At present, cross-border e-commerce in Guangdong Province has entered a stage of rising and rapid development. Even in recent years, affected by epidemic situation, changes in policy environment and fierce market competition, many cross-border e-commerce enterprises have maintained growth, which fully verifies the vitality and growth toughness of the cross-border e-commerce export market [1]. Therefore, in order to further promote its development, we analyze the scale of cross-border e-commerce in Guangdong Province through grey model. Firstly, the factors affecting cross-border e-commerce are found according to the grey correlation model, and then the scale of cross-border e-commerce is predicted by the grey prediction model.

2. Research Methods and Variable Setting

2.1. Research methods of data processing

Grey correlation analysis refers to the quantitative description and comparison of the development and change of a system. The basic idea is to determine the geometric similarity between the reference data column and several comparison data columns to judge whether the connection is close. It reflects the degree of correlation between curves [2]. In a grey system, it is necessary to know the relative strength of projects affected by other factors.

2.1.1. Determining characteristic sequence and parent sequence

The parent sequence is determined to reflect the index characteristics of the research object and the subsequence that affects the transformation of the parent sequence, represented by α. α is the sequence matrix of each index.

$$A = \begin{bmatrix} \gamma_1(1) & \cdots & \gamma_1(k) \\ \vdots & \ddots & \vdots \\ \gamma_1(k) & \cdots & \gamma_i(k) \end{bmatrix} \quad (1)$$

$where \gamma_i(k)$ is the kth sub-sequence in the ith parent sequence.

ISSN 2616-5775 Vol. 6, Issue 1: 26-31, DOI: 10.25236/AJCIS.2023.060105

2.1.2. Dimensionless treatment

Because of the different physical meanings of the selected influencing factors, it is inconvenient or difficult to compare the grey correlation analysis. Thus, it is necessary to carry out dimensionless treatment first. In order to reduce the error residual term as much as possible, the original data of indicators need to be dimensionless to improve the correlation between different factors. There are two mainstream ways to standardize the index data: initial value (Eq. (2)) and average value (Eq. (3)).

$$X'_{oi} = \frac{X_{oi}}{x(1)} \quad (2)$$

$$X'_{oi} = \frac{X_{oi}}{\bar{x}} \quad (3)$$

2.1.3. Correlation coefficient

The absolute value difference between the data in the independent variable sequence and the corresponding data in the dependent variable sequence is (K).

$$\Delta_{oi}(K) = |Y'(K) - X'_{oi}(K)| \quad (4)$$

The correlation coefficient is calculated as

$$\xi_{oi}(K) = \frac{min(min\Delta_{oi}(K)) + \rho max max(\Delta_{oi}(K))}{\rho max(max\Delta_{oi}(K)) + \Delta_{oi}} \quad (5)$$

where the resolution coefficient is $0<\rho<1$, generally $\rho=0.5$.

$$R^{oi} = \frac{1}{n}\sum_{oi=1}^{n} \xi_{oi}(K)$$

The grey correlation degree is normalized to obtain the index weight.

2.2. Data selection and variable setting

2.2.1. Tested variable

The total amount of cross-border e-commerce transactions in Guangdong Province includes the import and the export volume. In order to make the analysis results more intuitive, the total amount of cross-border e-commerce transactions in Guangdong Province is taken as the explained variable and expressed by Y.

2.2.2. Explanatory variable

There are many factors affecting the volume of cross-border e-commerce transactions, such as economy, culture, logistics and trade scale [3]. Based on the research of domestic and foreign scholars, we select the provincial GDP, local financial science and technology grants, freight volume, patent authorization, the number of college graduates, wholesale and retail commodity volume, and RMB exchange rate as the index system to study the influencing factors of cross-border e-commerce. In addition, as a large coastal province, sea and land transportation is an important way of commodity transportation [4]. The trading of e-commerce commodities cannot be separated from transportation. Therefore, this study takes the freight volume into account.

Table 1: Data on cross-border e-commerce import and export volume and influencing factors of cross-border e-commerce in Guangdong Province from 2014 to 2020

Year	Y	x1	x2	x3	x4	x5	x6	x7
2014	46.47	67792.2	274.33	179953	353732	44.09	60351.79	614.28
2015	167.30	72812.5	569.55	241176	376434	47.69	53553.5	622.84
2016	227.96	79512	742.97	259032	377645	48.94	60245.99	664.23
2017	441.90	89879	823.89	332648	400601	51.12	71916	675.18
2018	759.88	97277.77	1034.71	478082	424996	52.39	73208.99	661.74
2019	1107.90	107670	1168.79	527389	374823	52.21	81832.55	689.85
2020	1726.46	110800	951.36	709725	356221	55.01	91537.89	689.76

Therefore, we select the per capita GDP of Guangdong Province (x1, unit: 100 million yuan), Guangdong Provincial Local Financial Science and technology Appropriation (x2, unit: 100 million yuan), the number of patents authorized in Guangdong Province (x3, unit: PCS), freight volume (x4, unit: 10000 tons), the number of college graduates (x5, unit: 10000), and wholesale and retail commodity

ISSN 2616-5775 Vol. 6, Issue 1: 26-31, DOI: 10.25236/AJCIS.2023.060105

wholesale volume (x6, unit: 100 million yuan). Table 1 shows the data selection of the central parity rate (x7, unit: yuan) between RMB and US dollars.

3. Empirical Analysis Process

3.1. Grey correlation analysis

3.1.1. Correlation coefficient

Figure 1 shows the import and export volume of cross-border e-commerce in Guangdong Province from 2014 to 2020. According to Eq. (1), the original data of each variable in Table 1 is defined as parent sequence and sub sequence. The import and export volume of cross-border e-commerce in Guangdong Province has a stable increasing trend from 2014 to 2020. Therefore, the initial dimensionless processing is adopted (Eq. (2)). Secondly, since the data in Table 1are dimensionless with initial values, the absolute value data of the scale and the influencing factors of cross-border e-commerce in Guangdong Province can be obtained by substituting into Eq. (4). By substituting the data into Eq. (5), the correlation coefficient data of the transaction scale and influencing factors can be calculated. Table 2 shows the correlation coefficient of each variable.

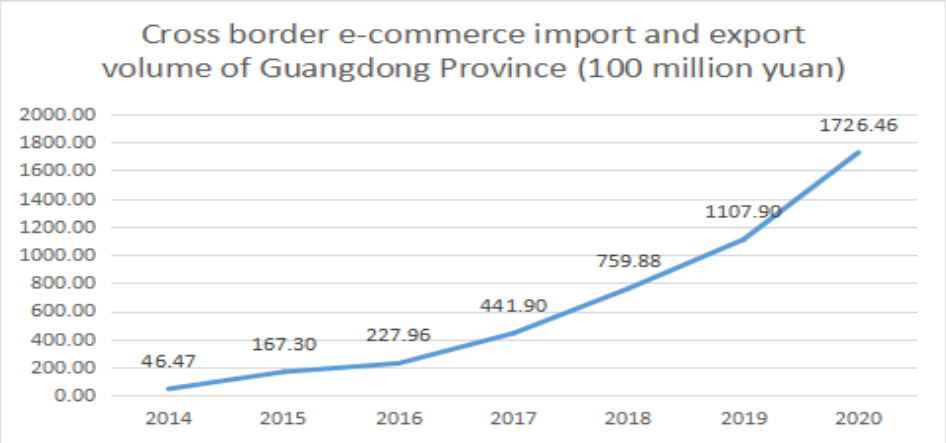

Figure 1: Data of import and export volume of cross-border e-commerce in Guangdong Province from 2014 to 2020

Table 2: Correlation coefficient results

Year	*Correlation coefficient results*						
	x1	**x2**	**x3**	**x4**	**x5**	**x6**	**x7**
2014	0.83937	0.99936	0.66304	0.50023	1	0.85445	0.99841
2015	0.82973	0.99887	0.59495	0.48476	0.99966	0.86896	0.99872
2016	0.81702	0.99855	0.57767	0.48400	0.99950	0.85505	0.99877
2017	0.79831	0.99893	0.51588	0.46940	0.99890	0.83202	0.99934
2018	0.78576	0.99923	0.42583	0.45488	0.99801	0.83012	0.99972
2019	0.76863	0.99983	0.40215	0.48646	0.99703	0.81431	0.99883
2020	0.76446	0.99782	0.33334	0.49965	0.99531	0.79764	0.99709

3.2. Grey correlation degree

Table 3: Grey correlation degree and ranking of influencing factors

Influence factor	Grey correlation degree	Ranking
x1	0.8	5
x2	0.999	1
x3	0.502	6
x4	0.483	7
x5	0.998	3
x6	0.836	4
x7	0.999	2

ISSN 2616-5775 Vol. 6, Issue 1: 26-31, DOI: 10.25236/AJCIS.2023.060105

The correlation coefficient represents the correlation degree value on the corresponding dimension of the sub sequence and the parent sequence. That is, the larger the number, the stronger the correlation. Based on the correlation degree to provide analysis reference, the grey correlation degree analysis is used, and the resolution coefficient is taken as 0.5. Combined with the correlation coefficient to calculate the correlation value, the correlation degree value is calculated for evaluation and judgment. The correlation coefficient data of the transaction scale and influencing factors shown in Table 2 is substituted into Eq. (6) to obtain the grey correlation degree and ranking of each factor, as shown in Table 3.

Combining the results of the above correlation coefficient for weight, the final correlation degree value is obtained, and the correlation degree value is used to rank the 7 evaluation objects. The correlation degree value is between 0 and 1. The larger the value, the stronger the correlation between the parent sequence and the parent sequence, which means the higher the evaluation. Table 3 shows that for the 7 evaluation items, the provincial local financial science and technology Appropriation (x2) has the highest evaluation (correlation degree: 0.999), followed by the RMB dollar foreign currency central parity rate (x7) (correlation degree: 0.999). According to the correlation degree of the three factors in Table 3, the order is

R1 > R7 > R5 > R6 > R1 > R3 >R4

3.3. Analysis summary

According to the results of grey correlation analysis, the correlation coefficient of other variables is greater than 0.5 except that the freight volume less than 0.5. This indicates that the indicators have a significant impact on the development of China's cross-border e-commerce. The influence degree of the six factors is different. Among them, the most relevant is the provincial local financial science and technology appropriation of Guangdong Province, which proves that this variable is the main factor for the trade volume of cross-border e-commerce in Guangdong Province.

4. Gray Prediction

The grey prediction model is to establish mathematical model and make prediction through a small amount of incomplete information [5,6].

4.1. Original data series

According to the correlation analysis, the scale of cross-border e-commerce in Guangdong Province in the next few years is predicted based on the GM (1,1) model. Equation (7) is used to establish the original data series $X^{(0)}$:

$$X^{(0)} = [x^{(0)}(1), x^{(0)}(2), x^{(0)} \quad x^{(0)} \quad x^{(0)} \quad x^{(0)}$$

4.2. Grade ratio test judgment

The grey prediction model must be tested to judge whether it is reasonable. Only the tested model can be used for prediction. Therefore, before establishing the grey prediction model GM (1,1), the order ratio test is carried out on the time series. Table 4 shows the results of level ratio test on the original data series.

Table 4: Grade ratio test results

	Year	Original value	Grade ratio
$x^{(0)}(1)$	2014	274.33	-
$x^{(0)}(2)$	2015	569.55	0.482
$x^{(0)}(3)$	2016	742.97	0.767
$x^{(0)}(4)$	2017	823.89	0.902
$x^{(0)}(5)$	2018	1034.71	0.796
$x^{(0)}(6)$	2019	1168.79	0.885
$x^{(0)}(7)$	2020	951.36	1.229

The stage ratio interval of $x^{(0)}(4), x^{(0)}(5), x^{(0)}(6)$ and $x^{(0)}(7)$ is in the interval $(e^{-\frac{2}{n+1}}, e^{\frac{2}{n+1}})$. The calculated interval is (0.778, 1.284), but other data are not included in this interval. Therefore, the

ISSN 2616-5775 Vol. 6, Issue 1: 26-31, DOI: 10.25236/AJCIS.2023.060105

sequence is translated so that the translated sequence meets the order ratio test. Table 4 shows the stage ratio after translation conversion. The sequence value after translation conversion meets the stage ratio test. It shows that the shift sequence is suitable for constructing the grey prediction model.

Table 5: Stage ratio test after translation conversion

Year	2014	2015	2016	2017	2018	2019	2020
Original value	1443.33	1738.55	1911.97	1992.89	2203.71	2337.79	2120.36
Grade ratio	-	0.83	0.909	0.959	0.904	0.943	1.103

4.3. First order accumulation generation

Since the shifted sequence is suitable for constructing the grey prediction model, the first-order accumulation is generated from the data.

$$X^{(1)}$$

4.4. Determine parameters

After the estimated values of A and B are obtained by regression analysis, the posterior error ratio C is calculated, as shown in Table 5. The development coefficient represents the development law and trend of the series, and the grey action quantity reflects the change relationship of the series. The posterior error ratio verifies the accuracy of grey prediction. The smaller the posterior error ratio is, the higher the accuracy of grey prediction is. Generally, if the posterior error ratio C value is less than 0.35, the model accuracy is high. If the C value is less than 0.5, the model accuracy is qualified. If the C value is less than 0.65, the model accuracy is basically qualified. If the C value is greater than 0.65, the model accuracy is unqualified. According to the analysis in the above Table 5, the posterior error ratio is 0.13 and the model precision is high.

Table 6: Parameter value and posterior error ratio C value

Development coefficient A	**Grey action B**	**Posterior error ratio C value**
-0.046	1716.289	0.13

4.5. Model fitting establishment and results

According to the calculated development coefficient a and grey action quantity B, the grey prediction model GM (1,1) can be obtained.

$$x^{(1)}(k+1) = 38753.964e^{0.046k} -$$

The results are restored and calculated to obtain the grey prediction fitting results. The residual and relative error values are used to test the model fitting, and the test results of the simulation values are shown in Table 7. The smaller the relative error, the better. Generally, the fitting is good with the means of less than 20%. The average relative error of the model is 8.222%, which means that the fitting effect of the model is good.

Table 7: Prediction fitting results

Year	**Estimate**	**Residual**	**Relative error (%)**
2014	274.33	0	0
2015	654.73	-85.18	14.956
2016	740.11	2.86	0.385
2017	829.488	-5.598	0.679
2018	923.05	111.66	10.791
2019	1020.992	147.798	12.645
2020	1123.52	-172.16	18.096

The order prediction is carried out according to the prediction model, and the prediction results are shown in Table 8.

Table 8: Forecast results

Prediction order	1	2	3	4	5	6	7
Estimate	1230.847	1343.199	1460.811	1583.93	1712.812	1847.728	1988.960

ISSN 2616-5775 Vol. 6, Issue 1: 26-31, DOI: 10.25236/AJCIS.2023.060105

5. Conclusions

The analysis and prediction of the scale of cross-border e-commerce in Guangdong Province shows that the development of cross-border e-commerce in Guangdong Province is closely related to the provincial local financial science and technology grants. Science and technology is the core of progress which promotes local development. For the coastal Guangdong Province, the progress of science and technology is particularly important. Therefore, the local government needs to encourage enterprises to develop science and technology. Especially for small and medium-sized enterprises, the government must give them more guidance and support for innovative technology to strengthen the innovation and upgrading of cross-border e-commerce in Guangdong Province.

Acknowledgment

This research was supported by The Natural Science Foundation of Guangdong Province (Grant No. 2015A030313642), and Science and Technology Planning Project of Guangdong Province (Grant No. 2014B040404072).

References

[1] Wang Jianfeng, and Jin Congying, "Literature review of cross-border e-commerce development." Business Economics, 11(2021): 100-104. (In Chinese)
[2] Wang Dongfeng, "Analysis on Influencing Factors of e-commerce transaction volume based on grey correlation degree." Journal of Zhengzhou Institute of aeronautical industry management, 32 (2014): 53-56.(In Chinese)
[3] Zheng Chunfang, Zhang Yanqiu, "Research on influencing factors and potential of China's cross-border e-commerce export." Journal of Graduate School of Chinese Academy of Social Sciences, 04(2021): 63-72.(In Chinese)
[4] Qi Yaman, "Analysis on the correlation between China's cross-border e-commerce and socio-economic development." Industrial innovation research, 04(2018): 35-37.(In Chinese)
[5] Zhang Luping, "E-commerce index prediction based on grey system prediction model," Modern marketing (next ten days), 03(2022): 125-127. (In Chinese)
[6] Jin Mingyue, and Cao Hongping, "Forecasting the scale of China UK e-commerce transactions based on GM model." Logistics technology, 44.02 (2021): 66-68. (In Chinese)