

# Container Throughput Forecast and Spillover Effect Analysis of Guangzhou Port

Zhaowu Xu

School of Shanghai, Maritime University, Shanghai 201306, China

## Abstract

Taking Guangzhou port as the research object, this paper establishes a time series model to predict the container throughput of Guangzhou port. This paper selects 123 groups of data from 2009.01 to 2019.03, and tests the stability of container throughput of Guangzhou port, and constructs ARIMA model according to the test results. The forecast results show that the container throughput of Guangzhou port will maintain a steady upward trend in the coming year, which provides reference value for port operation. In addition, this paper studies the information transmission between the container throughput of Guangzhou port and the stock price of Guangzhou port, which provides a reference for port investors.

## Keywords

Guangzhou port; ARIMA model; Spillover effect; Pulse analysis.

## 1. Introduction

Container transportation has become the main mode of port transportation because of its economic and social benefits. At present, as an important factor to evaluate the comprehensive capacity of ports, port container throughput has important reference value in port development and construction. Therefore, it is necessary to study the prediction of port container throughput. Guangzhou port is a container hub port in South China. It is of great significance to accurately predict its container throughput. By improving the accuracy of port container throughput prediction, it will play a positive role in the conventional operation, resource allocation, terminal scheduling, planning and construction of the port.

The prediction methods of port container throughput are mainly divided into multiple regression method and time series analysis method. The research on multiple regression method mainly includes: Liu Yiqun and others used the GDP of Dalian City, the total freight volume and the GDP of the three eastern provinces and Inner Mongolia as the influencing factors to predict the container throughput of Dalian port; Liu Bin and others used the gross national product, foreign trade volume, port fixed assets investment and interest rate as the influencing factors to predict the total port throughput of China.

Time series analysis methods mainly include index smoothing method, trend extrapolation method and ARIMA model. The main research contents are as follows: Liu Yulu et al. Constructed Arima prediction model by analyzing the change trend and cyclical characteristics of Wuhan port cargo throughput data; Zhao Shangwei and others used SARIMA and VaR combination forecasting method to forecast the throughput of seven major ports in China, reflecting the time series method Prediction has advantages.

Bakshi G (2011) found that the growth rate of BDI index can predict the return rate of some stock markets; in the case of the growth of global economic activities, BDI index can promote the entity and financial sectors. Oral Erdogan, Kenan Tata, can karahasan, etc. (2012) studied the Dow Jones average index and the Baltic dry bulk index, and found that there is linkage between the shipping market and the stock market, and the degree of information spillover

between the two will change constantly; the Dow Jones average can explain the change of BDI index in a short period of time. Amir H. Alizadeh and gulnur muradoglu (2014) analyzed the relevant data of S & P 500, S & P 400, S & P 600 and BDI index from February 1989 to October 2010, and found that shipping freight rate has certain predictability to the stock markets of the United States and the world, and its predictability to the stock price of the large market index, small and medium-sized enterprises and companies in different industries is significant. On the whole, there are relatively few articles about the linkage between shipping market and stock market in foreign literature, and the existing linkage analysis is slightly simple, which makes the research of this paper more practical.

This paper will take Guangzhou port as the research object, through the analysis of the data growth trend of port container throughput over the years, and use ARIMA time series method to establish a prediction model to predict the short-term container throughput of Guangzhou port. It also analyzes the information linkage between container throughput and stock price of Guangzhou port.

## 2. Model Establishment

### 2.1. ARIMA Model

Time series forecasting method is a regression prediction method, ARIMA model is widely used in time series analysis. It is an autoregressive integrated moving average model for estimating non seasonal and seasonal stationarity. Different from the general regression model, K exogenous variables,  $x_1, x_2$  In XK ARIMA model, the random error term and the lag term of YT are used to explain the variable. Arima method can find a suitable model when the data pattern is unknown, so it is widely used in the economic field. Its concrete form can be expressed as ARIMA (P, D, q). The order of autoregression is represented by P, the number of difference is expressed by D, and the order of moving average process is expressed by Q. If the data is not stable, it is necessary to process the data to make it stable, which is called d-order difference. Stationarity means that there is no growth or decline in the data. In other words, demand data fluctuates around a constant mean, independent of time. If the data has obvious growth trend or large fluctuation, it is necessary to eliminate the unstable trend. The best method is difference method. The difference method is a method that uses the difference or change of each observation value in the original time series to convert the data into a stable time series like white noise. The existence of white noise means that the data is no longer in the data There are trend characteristics.

The first order difference is as follows:

$$\Delta D_t = D_t - D_{t-1} \quad (1)$$

The second order difference is as follows:

$$\Delta^2 D_t = \Delta D_t - \Delta D_{t-1} \quad (2)$$

The d-order difference is as follows:

$$\Delta^d D_t = \Delta^{d-1} D_t - \Delta^{d-1} D_{t-1} \quad (3)$$

Once the data is transformed from unstable state to stationary state, the new sequence generated by difference can be replaced.

$$w_t = \Delta^d D_t \tag{4}$$

Where  $W_t$  is a stationary sequence. Yes,  $W_t$  establishes ARMA (P, q) model, determines the order of P and Q, and obtains  $W_t$  expression

$$w_t = \phi_1 w_{t-1} + \phi_2 w_{t-2} + \dots + \phi_p w_{t-p} + \mu_1 w_{t-1} + \mu_2 w_{t-2} + \dots + \mu_p w_{t-p} + c \tag{5}$$

Then, the  $w_t$  prediction value is obtained by static prediction method or dynamic simulation method, and the  $D_t$  prediction value is obtained by introducing it into equations (1), (2) and (3).

## 2.2. Spillover Effect Model

In this paper, a vector autoregressive model (VAR) is established to study the spillover effect between the container throughput and the stock price of Guangzhou port, and further explore the transmission relationship between the stock price and the port throughput. Spillover effect refers to that when an organization carries out an activity, it will not only produce the expected effect of the activity, but also have an impact on people or society outside the organization. Guangzhou port (601228) was issued in Shanghai and Shenzhen on March 29, 2017, so the data selected in this section are from April 2017 to March 2019, a total of 24 groups of data.

$$\Delta D_{1,t} = a_{11} D_{1,t-1} + a_{12} D_{2,t-1} + b_{11} \Delta D_{1,t-1} + b_{12} \Delta D_{2,t-1} + c_1 \tag{6}$$

$$\Delta I_{2,t} = a_{21} D_{1,t-1} + a_{22} D_{2,t-1} + b_{21} \Delta D_{1,t-1} + b_{22} \Delta D_{2,t-1} + c_2(\theta) \tag{7}$$

$$\Delta D_{i,t} = \frac{D_{i,t} - D_{i,t-1}}{D_{i,t-1}} \tag{8}$$

Where  $D_{1,t}$  is the container throughput of Guangzhou port in t month,  $D_{2,t}$  is the average stock price of Guangzhou port in t month.  $\Delta D_{i,t}$  is the rate of return for the month. Formula (6) and formula (7) can be vectorized as follows:

$$\Delta_t = A D_{t-1} + B R_{t-1} + C \tag{9}$$

Where,  $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ ,  $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ ,  $C = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

This paper further establishes the pseudo impulse model (QIRF) to measure the spillover effect between the container throughput and the stock price of Guangzhou port. The model is based on the assumption that at time t the original series (...  $D_{i,t-2}$ ,  $D_{i,t-1}$ ,  $D_{i,t}$ ,  $D_{i,t+1}$ ,  $D_{i,t+2}$  ...) After stimulation with a unit standard deviation  $\delta$ , the sequence after the change was calculated (...  $D_{i,t-2}$ ,  $D_{i,t-1}$ ,  $\tilde{D}_{i,t}$ ,  $D_{i,t+1}$ ,  $D_{i,t+2}$  ...) Impact on the target market.

$$\tilde{D}_{i,t} = D_{i,t} + \delta \tag{10}$$

The pseudo pulse model of sequence  $D_{i,t}$  can be expressed as  $\Delta_{i,s}(\tilde{D}_{i,t})$ .

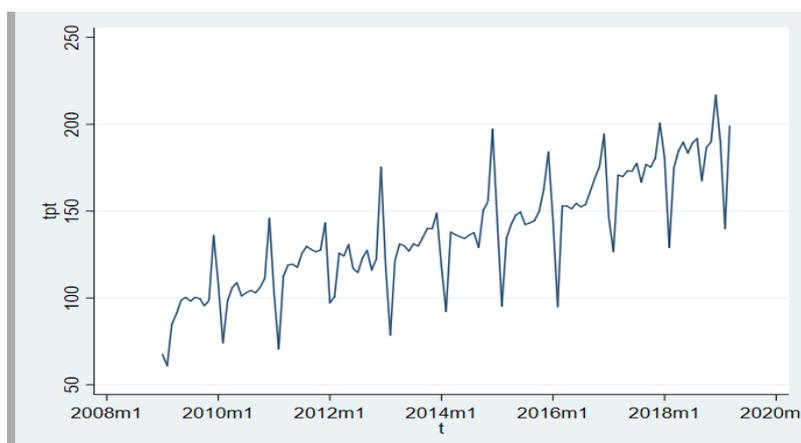
$$\Delta_{j,s}(\tilde{r}_{i,t}) = \Delta \tilde{D}_{i,t+s} - \Delta D_{i,t+s}, \tag{11}$$

### 3. Results

#### 3.1. Forecast Results of Container Throughput of Guangzhou Port

As an important transportation hub in China's port system, Guangzhou port is an important port for foreign trade in South China. It is composed of Nansha port, Huangpu port, Huadu port, Xintang port, etc. At present, the container transportation of Guangzhou port covers more than 80 countries and regions, involving more than 300 world ports. With the deepening of one belt, one road policy, Guangzhou port has ranked the top 10 in the world port ranking and ranks among the world's first-class ports.

This paper selects 123 sets of container throughput data of Guangzhou port from January 2009 to March 2019 after the financial crisis in 2008. It is found that the container throughput of Guangzhou port has strong periodicity and shows an increasing trend. It can be seen that the ordinary linear regression is not accurate in describing the periodic series, while ARIMA model in the time series model can eliminate the periodic characteristics through the difference method, so as to improve the accuracy of the prediction model.



**Figure 1.** Container throughput of Guangzhou Port

Null Hypothesis: X has a unit root  
 Exogenous: Constant  
 Lag Length: 11 (Automatic - based on SIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	0.928446	0.9956
Test critical values: 1% level	-3.490210	
5% level	-2.887665	
10% level	-2.580778	

\*MacKinnon (1996) one-sided p-values.

(a)

Null Hypothesis: D(X) has a unit root  
 Exogenous: Constant  
 Lag Length: 10 (Automatic - based on SIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-12.06956	0.0000
Test critical values: 1% level	-3.490210	
5% level	-2.887665	
10% level	-2.580778	

\*MacKinnon (1996) one-sided p-values.

(b)

**Figure 2.** ADF inspection results of Guangzhou Port Container Throughput

First of all, ADF test is carried out for the stationarity of the sequence. It is found that the original sequence is non-stationary, and the ADF test is passed after the first-order difference processing.

Figure 2 shows the test results of container throughput stability of Guangzhou port. (a) figure shows the results of the original series of tests, and the results show that the T values are greater than the inspection values under different confidence levels, and the Prob value is far greater than 0.05, which proves that the data are non-stationary series, that is, the mean value or covariance function of the time series changes with time; (b) the figure shows the stability test after the first-order difference of the original series The test results show that t value is less than the test value under different confidence levels, and the Prob value is less than 0.05, which proves that the data results become stable after the first-order difference, then the value of D is determined to be 1, that is  $\Delta D_t = D_t - D_{t-1}$ ,  $D_t$  is the throughput of period t.

Then, the values of P and Q are estimated by means of the autocorrelation function graph and partial autocorrelation function graph generated by Eviews (Fig. 3).

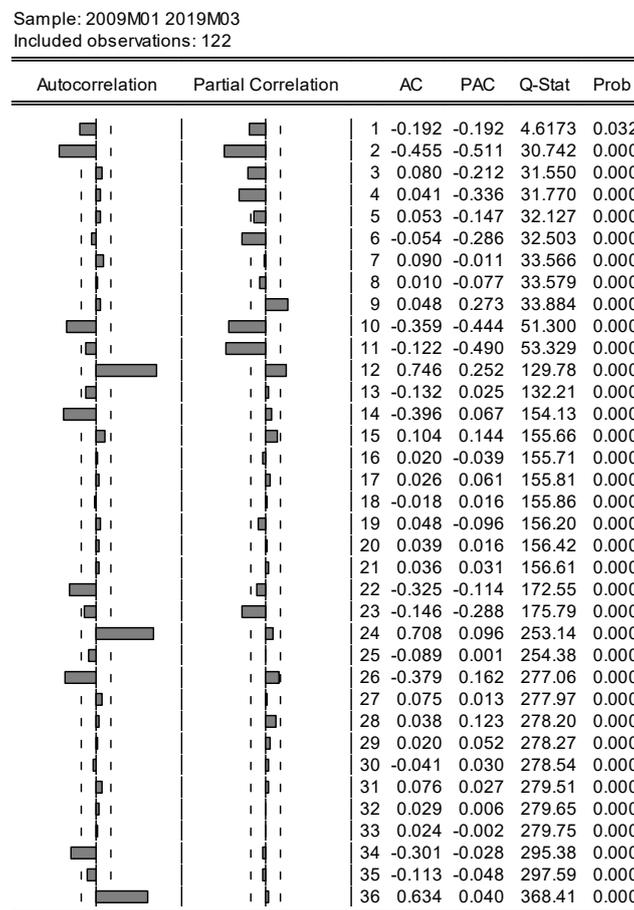


Figure 3. Autocorrelation function and partial correlation function of Guangzhou Port Throughput

Through observation, we can find that the autocorrelation function graph shows "tailing" phenomenon, and the partial autocorrelation function graph is initially judged as "truncated" after 12 periods, then the values of P, D and Q are 12, 1, 0 respectively, and ARIMA (12, 1, 0) model is constructed. The ARIMA model was tested, and the results (as shown in Figure 4) showed that the test value of phase 12 was not significant. If you go back to the figure above, it is found that the partial autocorrelation function of phase 12 is in the range of two standard deviations, so try to change the p value to 11 and construct the ARIMA (11,1,0) model.

Dependent Variable: Y  
 Method: ARMA Maximum Likelihood (OPG - BHHH)  
 Date: 05/07/19 Time: 15:41  
 Sample: 2009M02 2019M03  
 Included observations: 122  
 Convergence achieved after 43 iterations  
 Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.845159	0.160929	5.251762	0.0000
AR(1)	-0.637900	0.103425	-6.167746	0.0000
AR(2)	-0.849287	0.113608	-7.475583	0.0000
AR(3)	-0.672986	0.146795	-4.584529	0.0000
AR(4)	-0.753426	0.168537	-4.470383	0.0000
AR(5)	-0.616887	0.159391	-3.870266	0.0002
AR(6)	-0.686815	0.167013	-4.112349	0.0001
AR(7)	-0.496224	0.163563	-3.033832	0.0030
AR(8)	-0.607544	0.159934	-3.798728	0.0002
AR(9)	-0.485446	0.135887	-3.572431	0.0005
AR(10)	-0.589639	0.148986	-3.957680	0.0001
AR(11)	-0.505357	0.107989	-4.679711	0.0000
AR(12)	0.180852	0.118554	1.525485	0.1301
SIGMASQ	103.9014	12.46727	8.333930	0.0000
R-squared	0.803545	Mean dependent var	1.077869	
Adjusted R-squared	0.779898	S.D. dependent var	23.09227	
S.E. of regression	10.83375	Akaike info criterion	7.810032	
Sum squared resid	12675.97	Schwarz criterion	8.131805	
Log likelihood	-462.4119	Hannan-Quinn criter.	7.940726	
F-statistic	33.98037	Durbin-Watson stat	1.911051	
Prob(F-statistic)	0.000000			
Inverted AR Roots	.83+.48i	.83-.48i	.49+.85i	.49-.85i
	.26	-.00+.99i	-.00-.99i	-.48-.86i
	-.48+.86i	-.85-.49i	-.85+.49i	-.87

Figure 4. Evaluation of ARIMA (12,1,0) model

The results show that T statistics pass the significance test, R2 is close to 1, DW test is close to 2, indicating that the model fitting effect is very good.

Dependent Variable: Y  
 Method: ARMA Maximum Likelihood (OPG - BHHH)  
 Date: 05/07/19 Time: 15:49  
 Sample: 2009M02 2019M03  
 Included observations: 122  
 Convergence achieved after 38 iterations  
 Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.834768	0.129299	6.456089	0.0000
AR(1)	-0.758444	0.070759	-10.71874	0.0000
AR(2)	-0.996590	0.069479	-14.34373	0.0000
AR(3)	-0.796546	0.118715	-6.709714	0.0000
AR(4)	-0.904324	0.135965	-6.651165	0.0000
AR(5)	-0.744668	0.141352	-5.268186	0.0000
AR(6)	-0.847262	0.134470	-6.300764	0.0000
AR(7)	-0.641109	0.133668	-4.796274	0.0000
AR(8)	-0.774161	0.109338	-7.080470	0.0000
AR(9)	-0.633728	0.098158	-6.456227	0.0000
AR(10)	-0.769444	0.065369	-11.77085	0.0000
AR(11)	-0.643809	0.063492	-10.14006	0.0000
SIGMASQ	107.2819	13.13959	8.164781	0.0000
R-squared	0.797153	Mean dependent var	1.077869	
Adjusted R-squared	0.774822	S.D. dependent var	23.09227	
S.E. of regression	10.95796	Akaike info criterion	7.822692	
Sum squared resid	13088.39	Schwarz criterion	8.121481	
Log likelihood	-464.1842	Hannan-Quinn criter.	7.944051	
F-statistic	35.69601	Durbin-Watson stat	1.711126	
Prob(F-statistic)	0.000000			
Inverted AR Roots	.84-.48i	.84+.48i	.50-.85i	.50+.85i
	-.00+.99i	-.00-.99i	-.47+.86i	-.47-.86i
	-.81	-.84+.49i	-.84-.49i	

Figure 5. ARIMA (11,1,0) model evaluation table

Then ARIMA (11,1,0) model was constructed,

$$\begin{aligned} \Delta D_t = & -0.7584\Delta D_{t-1} - 0.9966\Delta D_{t-2} - 0.7965\Delta D_{t-3} - 0.9043\Delta D_{t-4} - 0.7447\Delta D_{t-5} \\ & - 0.8473\Delta D_{t-6} - 0.6411\Delta D_{t-7} - 0.7742\Delta D_{t-8} - 0.6337\Delta D_{t-9} - 0.7694\Delta D_{t-10} \\ & - 0.6438\Delta D_{t-11} + 0.8348 \end{aligned}$$

According to the fitted model and the existing actual data, the container throughput of Guangzhou port in 12 months from April 2019 to March 2020 is predicted. The forecast results are shown in Table 1:

**Table 1.** Forecast of container throughput of Guangzhou Port

date	throughput	First order differential prediction of throughput
2019M04	201.9402656	2.940265599
2019M05	194.5518119	-7.388453719
2019M06	193.2768061	-1.275005748
2019M07	198.9248782	5.648072054
2019M08	200.6754625	1.750584279
2019M09	178.1082348	-22.56722768
2019M10	196.1449326	18.03669779
2019M11	201.0388458	4.893913252
2019M12	221.3674407	20.32859483
2020M01	200.4300398	-20.93740082
2020M02	160.903718	-39.52686802
2020M03	203.8626645	42.9549264

The results show that Guangzhou port will continue to maintain a steady growth trend. According to the prediction results, it is suggested that in the peak and trough period of container cargo transportation, active resource allocation, wharf scheduling, planning and construction should be carried out to meet the demand of port container transportation.

### 3.2. Analysis of Container Throughput Spillover Effect in Guangzhou Port

This paper establishes vector autoregressive model and pseudo pulse model to analyze the spillover effect between container throughput and stock price of Guangzhou port. Container throughput is one of the profit indicators of the port, and the port stock price represents the investors' expectation of the port's future operation, so the port container throughput has a positive spillover effect on the port stock price.

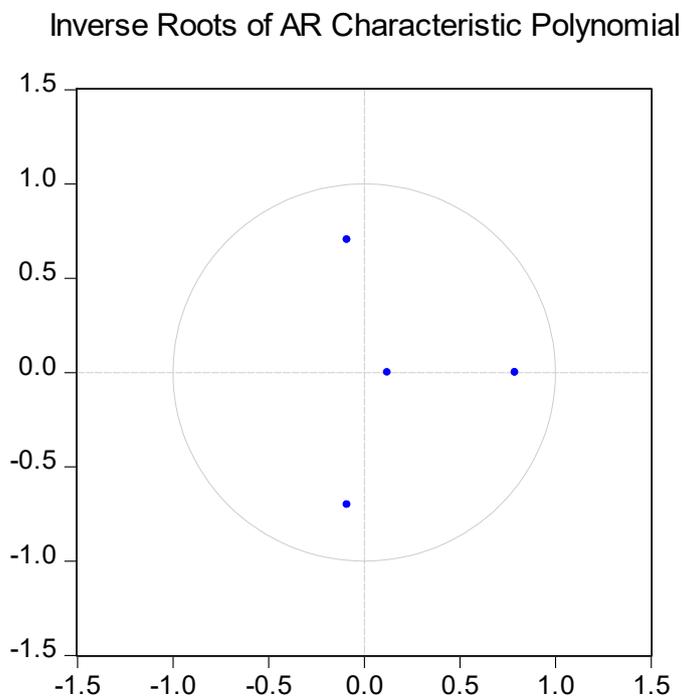
**Table 2.** Regression results of VAR (1,1) model

a11	a12	b11	b12	c1
-0.0114*** (0.0018)	-0.0424*** (0.0146)	-1.3268*** (0.2013)	-0.1290 (0.2628)	2.3312*** (0.3719)
a21	a22	b21	b22	c2
-0.0009 (0.0012)	-0.0256*** (0.0097)	-0.1896 (0.1332)	-0.0787 (0.1739)	0.2850 (0.2462)

Note:\*\*\* It is significant at 99% confidence level

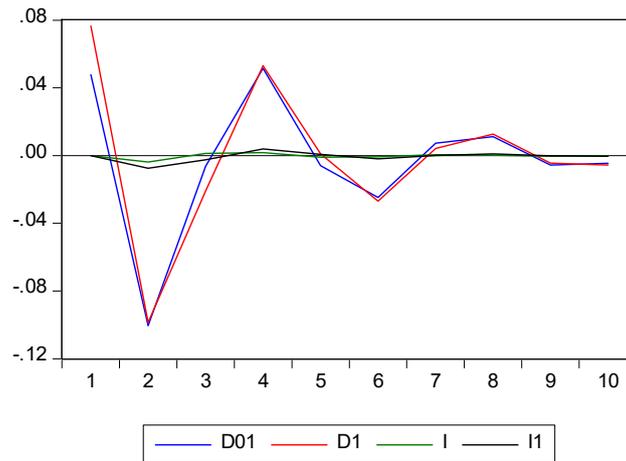
As can be seen from table 1, a11 is significant, which indicates that the container throughput of Guangzhou port has an impact on its throughput, A12 and A22 are significant, indicating that the stock price of Guangzhou port has a negative impact on its container throughput and stock

price; B11 shows that the return rate of Guangzhou port's container throughput has an extension effect on itself. In order to verify the rationality of the model, AR root test is carried out on the structure. As shown in Fig. (6), the AR root test results are all in the unit Park, proving that the model is stable. Next, impulse analysis of VAR model is carried out.

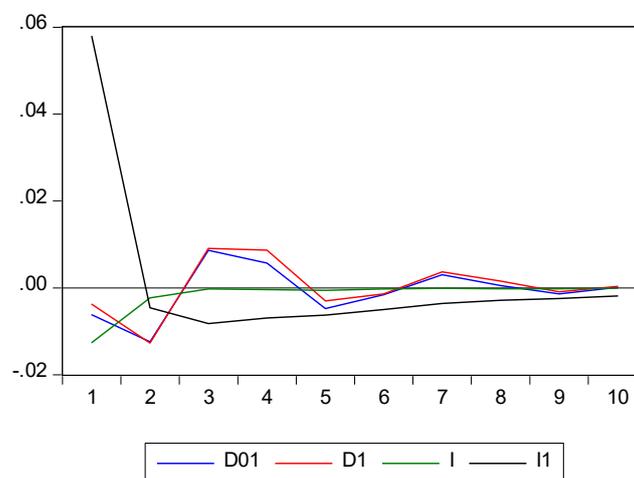


**Figure 6.** AR root test results of VAR (1,1) model

Figure 7 shows the pulse diagram of container throughput and stock price of Guangzhou port. D01 shows the impact of the previous period's port container throughput on the current port container throughput and stock price. It can be seen that the port container throughput has a positive impact on itself in the next month, but has a negative impact on the next month, which reaches the maximum in the next month, and finally disappears after 7 months. For the stock price, the port container throughput has a negative impact on the stock price return, which becomes positive after two months, and gradually disappears after five months. D1 represents the impact of port container throughput rate of return on port container throughput and return on share price. It can be seen that the effect of port container return on stock price return is similar to that of port container throughput. I and I1 represent the impact of the previous period's stock price and the previous period's return on the port container throughput and stock price return respectively. The results show that the stock price and the stock price return will not affect the port container throughput. Thus, the new information is first reflected in the port's container throughput, and then reflected in the port's share price. This is because the container throughput of a port is one of its operation indicators. The higher the port container throughput is, the higher the port's revenue is, and the higher the investors' expectation of it is.



(a) Pulse chart of container throughput



(b) Port stock return pulse chart

**Figure 7.** Pulse chart of container throughput and stock return of Guangzhou Port

## 4. Conclusion

This paper takes Guangzhou port as the research object, uses 123 sets of data from 2009.01 to 2019.03 to analyze the changing trend characteristics of container throughput data, and constructs Arima prediction model. The prediction model based on the time series of each container is very universal. Finally, the ARIMA prediction model is used to predict the container throughput of Guangzhou port in the next year. It is believed that it can provide some reference for the development and construction, daily operation and planning layout of Guangzhou port. The spillover effect between the container throughput of Guangzhou port and the stock price of Guangzhou port is studied. The results show that the new information first shows in port container throughput and then in port stock price, which provides information for investors.

## References

- [1] Liu Yu Lu, Chen Dong lin. Prediction of cargo throughput of Wuhan port based on ARIMA model [J]. Water transportation in China, 2016(10):45-47.
- [2] Liu Ting, Lin . Study on prediction method of port container throughput [J]. Journal of Suzhou University of science and Technology (Engineering and Technology Edition), 2011, 24(4):44-46.

- [3] Liu Bin, Qiu Guodong, Liu Chao. Linear regression model of container port throughput [J]. Journal of Dalian Maritime University, 2003, 29(2): 27-30.
- [4] Zhao Shangwei, Zhou Jianhong. Prediction of China's port container throughput: Based on combined time series [J]. Systems science and mathematics, 2018, 38 (2): 210-219.
- [5] Ma Huihui, Guo Qingran, Ding Cuicui. *Views statistical analysis and application (3rd Edition)* [M]. Electronic Industry Press, 2016.
- [6] Zhu Xiaomeng, Luan Weixin, Zhu Yisheng. Medium and long term prediction of container throughput of China's seaports based on time series causality combination method [J]. Journal of Dalian Maritime University (SOCIAL SCIENCE EDITION). 2014 (05)
- [7] Gao Fengjiao, Shi Xiaofa. Analysis of influencing factors on container throughput of Shanghai port [J]. Logistics technology. 2009 (09)
- [8] Liu Changjian, Zhang Qingnian. Dynamic prediction of container throughput based on time series BP neural network [J]. Water transportation engineering. 2007 (01)
- [9] Research on the correlation between BDI index and Shanghai Composite Index under the financial crisis [D]. Wan Zhong. Dalian Maritime University, 2009
- [10] Kavussanos M G , Visvikis I D , Dimitrakopoulos D N . Economic spillovers between related derivatives markets: The case of commodity and freight markets[J]. Transportation Research Part E: Logistics and Transportation Review, 2014, 68:79-102.
- [11] Alizadeh A H , Muradoglu G . Stock market efficiency and international shipping-market information[J]. Journal of International Financial Markets, Institutions and Money, 2014, 33:445-461.
- [12] Kavussanos M G, Visvikis I D, Dimitrakopoulos D N. Economic spillovers between related derivatives markets: The case of commodity and freight markets[J]. Transportation Research Part E, 2014, 68(4):79-102.
- [13] Kavussanos M G, Visvikis I D, Batchelor R A. Over-the-counter forward contracts and spot price volatility in shipping[J]. Transportation Research Part E Logistics & Transportation Review, 2004, 40(4):273-296.
- [14] Alizadeh A H. Trading volume and volatility in the shipping forward freight market[J]. Transportation Research Part E Logistics & Transportation Review, 2013, 49(1):250-265.
- [15] Adland R , Cariou P , Wolff F C . The influence of charterers and owners on bulk shipping freight rates[J]. Transportation Research Part E: Logistics and Transportation Review, 2016, 86:69-82.