Analysis of Port Handing Capacity Based on X11-ARIMA Model

Lixin Zhang, Cuifang Yang and Zhenqi Fan[*]
College of Information Engineering, Tarim University, Alar Xinjiang843300, China
[*]Corresponding author

Abstract—The analysis of port handing capacity is of great significance to the construction and planning of the port. Using the data of the last 12 years, X11 method is used to separate the long-term trend. The results show that the port handing capacity tends to increase linearly. We use difference method to eliminate the effect of the long-term trend and the seasonal effect, choose ARIMA model to fit the development of the sequence after repeated attempts. With these data that from May to August in 2017 as the comparison sample，the model predicted the data very precisely.

Keywords—port handing capacity; long-term trend; sequence

I. INTRODUCTION

Port handing capacity refers to the total quantity of goods exported by water, imported into port area and handled in a year. It is an important quantitative index reflecting the achievement of port's production and operation activities. Correct prediction of port handing capacity to provide data to determine the reasonable port layout, comprehensive transportation infrastructure investment scale, operation strategy, development strategy and related transportation planning is also very important, so the prediction of port handing capacity is of great significance to the port.At present, there are studies based on the neural network method for the prediction of port handing capacity.The difficulty in variable selection affects the practicability of the model to some extent, as in [1-2]. Ning Chen and other authors make smooth predictions based on logarithmic quadratic exponential functions, but the choice of smoothing coefficient affects the accuracy and sensitivity of the prediction to a certain extent[3]. Xiaodan Shi adopted improved numerical attribute reduction algorithm based on neighborhood rough set to obtain key indicators, and then used genetic programming methods to train key index samples to build prediction model. However, this algorithm has a certain dependence on the selection of the initial group, and it is implemented by many parameters. The choice of these parameters seriously affects the quality of the solution[4]. Changxin Xu et al applied the system dynamics method to the forecasting model, But the selection of different independent variables affects the results of the prediction[5]. Chaofeng Deng using linear regression prediction model, because there are many factors affecting the port handing capacity, and random factors on the system is often difficult to use a certain function to describe, linear regression is used for simple prediction. In fact, the relationship between the factors is often nonlinear[6].Some authors use different methods to predict the different goods. But in this way, we must first classify the goods according to certain standards, and we need to predict different goods by different ways, which increases

the complexity of prediction to a certain extent[7-8].Several methods are also combined to reduce the systematic error of prediction and improve the effect of prediction, but the selection of combined weights has no unified standard[9-11].This paper tries to predict port handing capacity by using the X11-ARIMA model, and uses the data from May to August in 2017 as an assessment sample to verify the accuracy of the model.

II. MATERIALS AND METHOD

A. Data Material

From the website of the National Bureau of Statistics, we collected the monthly data on port handing capacity from January in 2005 to August in 2017.

B. Research Methods

1) X11 model: The X11 model assumes that the time series consists of four elements: long-term trends, seasonal variations, cyclical fluctuations, and random fluctuations. The following three moving average methods are used in the X11 program to achieve accurate decomposition of various sequences: simple center moving average, Henderson weighted moving average, Musgrave asymmetric moving average.

2) ARIMA model: Assuming that x_t is the t-th observation of a time series, the model $ARIMA(p,d,q)$ is:

$$\nabla^d x_t = \frac{\Theta(B)}{\Phi(B)} \varepsilon_t \qquad (1)$$

The t here represents time, B stands for delay operator, then $\nabla = 1 - B$,

$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ represents the regression polynomial,

$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$ represents the moving average polynomial.

The multiple seasonal model of order $(p,d,q) \times (P,D,Q)_S$ can be expressed as

$$\nabla^d \nabla_S^D x_t = \frac{\Theta(B)\Theta_S(B)}{\Phi(B)\Phi_S(B)} \varepsilon_t \qquad (2)$$

The S here represents the seasonal cycle, then $\nabla_S = 1 - B^s$,

$\Phi_s(B) = 1 - \varphi_1 B^S - \varphi_2 B^{2S} - \cdots - \varphi_P B^{PS}$ represents the seasonal autoregressive polynomial.

$\Theta_s(B) = 1 - \theta_1 B^S - \theta_2 B^{2S} - \cdots - \theta_Q B^{QS}$ represents the seasonal moving average polynomial.

III. EMPIRICAL ANALAYSIS

A. Long-Term Trend Analysis

We select the data of port handing capacity from January in 2005 to April in 2017, marked as sequence x_t.In order to show the trend of changes in the sequence, draw the timing diagram, as shown in figure 1.
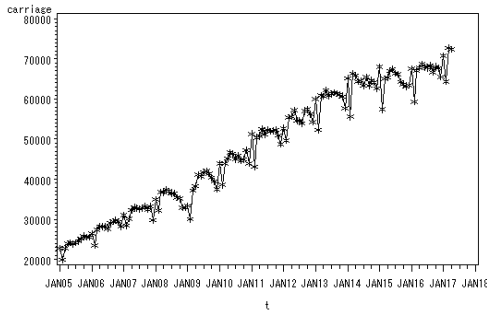

FIGURE I. ORIGINAL SEQUENCE DIAGRAM

From the sequence diagram as a whole we can see the linear increasing trend, seasonal effect for one year period and the influence of random fluctuations, using X11 method to decompose the long-term trend, and draw the long-term trend chart, as shown in figure 2.
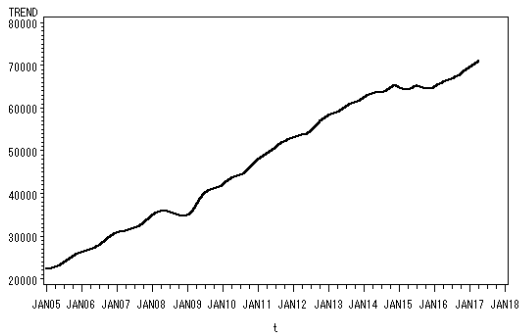

FIGURE II. LONG-TERM TREND CHART

From the long-term trend chart, it can be clearly seen that the port handing capacity has shown a linear growth trend in the past 12 years. The reason is that the main factor affecting port handing capacity has been growing rapidly in recent years, such as GDP, total export-import volume. Under this influence, port handing capacity has been growing rapidly.

B. ARIMA Model Analysis

1) Stability test: Since the sequence has a very clear linear increasing trend and the seasonal effect of one year cycle, we eliminate the linear trend by first order difference, then make a 12-step difference to eliminate the seasonal effect. The sequence after differences is recorded as y_t, and the stability test is carried out. The test results are shown in table 1, where P-value is far less than the critical value of 0.05. Therefore, the differential sequence is stable.

TABLE I. STABILITY TEST

Type	Delay Order	τ -Statistic	P-Value
Type1	0	-17.87	<.0001
	1	-9.54	<.0001
Type2	0	-17.80	<.0001
	1	-9.50	<.0001
Type3	0	-17.73	<.0001
	1	-9.47	<.0001

2) White noise test: The white noise test of the sequence after the differences is shown in table 2.

TABLE II. WHITE NOISE TEST

Delay Order	χ^2 -Statistic	DF	P-Value
6	37.15	6	0.0001
12	69.08	12	<.0001
18	87.00	18	<.0001
24	123.50	24	<.0001

P-Value is far less than the critical value of 0.05, so the sequence is stable and is not a white noise sequence, there is a certain degree of relationship between the sequence, which need to establish a model to fit the changing trend of this sequence.

3) Model establishment: The order of the model is determined according to the characteristics of the autocorrelation coefficient and the partial autocorrelation coefficient of the post-differential sequence y_t. Autocorrelation coefficient and partial autocorrelation coefficient are shown in table 3:

TABLE III. AUTOCORRELATION COEFFICIENT AND PARTIAL AUTOCORRELATION COEFFICIENT

Delay Order	Autocorrelation Coefficient	Partial Autocorrelation Coefficient	Standard Deviation
0	1.00000	--	0
1	-0.41203	-0.41203	0.086066
2	0.17704	0.00876	0.099612
3	-0.17651	-0.12118	0.101916
4	-0.07063	-0.22648	0.104156
5	0.06847	-0.03426	0.104510
6	-0.15889	-0.18319	0.104842
7	0.20661	0.03303	0.106611
8	-0.09865	0.00084	0.109537
9	0.14072	0.06891	0.110193
10	-0.06781	0.02879	0.111516
11	0.07548	0.11557	0.111821
12	-0.36389	-0.38631	0.112198
13	0.15348	-0.11323	0.120624
14	-0.09243	-0.08883	0.122062
15	0.11833	-0.04158	0.122579
16	0.06640	0.00532	0.123422
17	-0.06101	-0.01118	0.123687
18	0.24662	0.18708	0.123909
19	-0.33345	-0.06251	0.127493
20	0.15830	-0.05556	0.133798
21	-0.16155	-0.00044	0.135178
22	0.07123	-0.00696	0.136600
23	0.15713	0.20286	0.136876
24	-0.18333	-0.24608	0.138206

Both the autocorrelation coefficient and the partial autocorrelation coefficient are not truncated, so we try to use the $ARMA(p,q)$ model to extract short-term related information. We select $ARMA(1,12)$ model after repeated attempts, combined with the previous differential operation, the final model is $ARMA(1,1,0) \times (0,1,1)_{12}$, that is

$$(1-B^{12})(1-B)x_t = \frac{1-\theta_{12}B^{12}}{1-\varphi_1 B} \varepsilon_t \qquad (3)$$

The conditional least square method was used to estimate the parameters. The fitted model is:

$$(1-B^{12})(1-B)x_t = \frac{1-0.62449B^{12}}{1+0.46178B} \varepsilon_t \qquad (4)$$

The test results show that the model successfully passes the residual's white noise test and the significance test of the parameters, which show that the model fully extracts the information of the data. The test results are shown in table 4 and table 5.

TABLE IV. RESIDUAL'S WHITE NOISE TEST

Delay Order	χ^2-Statistic	DF	P-Value
6	7.43	4	0.1149
12	10.05	10	0.4360
18	12.08	16	0.7382

TABLE V. PARAMETER TEST

Parameter	Delay Order	Estimated Value	t-Value	P-Value
MA1，1	12	0.62449	9.06	<.0001
AR1，1	1	-0.46178	-5.98	<.0001

In order to display the effect of the model fitting more intuitively, the fitting value and the observation value are drawn in figure 3, where the curve represents the fitting value and the asterisk represents the observation. It can be intuitively seen from the figure that the model fits well with the data.
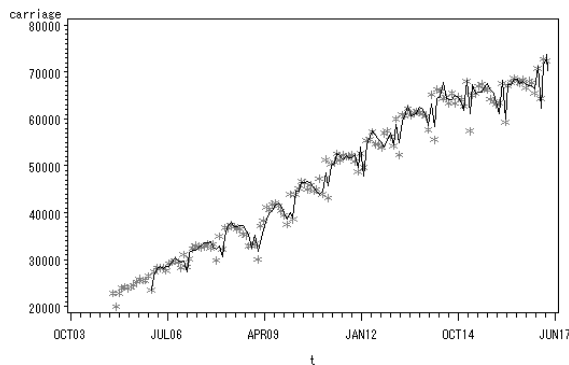

FIGURE III. FITTING DIAGRAM

4) Model prediction: According to (4), port handing capacity from May to December in 2017 are predicted. The data from May to August are taken as the test samples. The test results are shown in table 6.It is obvious from the table that the relative error between the predicted value and the actual value is no more than 1.5%, which indict that the model predicts the data very accurately.

TABLE VI. PREDICTED RESULTS

Month	Actual Value (10 thousand ton)	Predictive Value (10 thousand ton)	Relative Error(%)
May	74415	73631.14957	1.05335
June	72906	72980.5555	-0.10226
July	73679	72614.73535	1.444461
August	72991	73066.42609	-0.10334
September		71813.25918	
October		72382.20412	
November		71874.48523	
December		70346.74655	

IV. CONCLUSION

According to the changing tendency of port handing capacity, we separate the long-term trend by X11 method. The result shows that port handing capacity tends to increase linearly. The reason is that the main factor affecting the port handing capacity has been growing rapidly in recent years, such as GDP. Using the difference method to eliminate the effects of long-term trends and seasonal effects, after repeated attempts, we use $ARMA(1,1,0) \times (0,1,1)_{12}$ model to fit the development of the sequence and forecast the port handing capacity from May to December in 2017. The predicted value from May to August show that the model predicts the data very accurately, which can provide reference for the planning and construction of the port.

REFERENCES

[1] Meilian Liu, Meihua Zhu. The Port Throughput Forecast Model Based on BP Neural Network[J]. Chinese Journal of Systems Science,2012, 20(04):88-91.

[2] Ling Su.Research on Port Throughput Prediction Based on radial basis function neural network[D].Shanghai Maritime Univeristy,2006.

[3] Ning Chen,Meiqi Zhu,Zhenwen Yu.Port Throughput Prediction Based on logarithmic two exponential smoothing[J].Journal of Wuhan University of Technology,2005,(09):77-79.

[4] Xiaodan Shi.Prediction Research on Port Throughput on Rough Set and Genetic Programming[D]. Beijing Jiaotong University,2014.

[5] Changxin Xu,Yixin Yan,Ping Zhang. Prediction Model of Port Throughput Based on System Dynamics[J].Port & Warterway Engineering, 2006,(05):26-28+40.

[6] Chaofeng Deng. Study on Throughput Prediction Model of Main Port in the Bohai Bay Port Cluster[D]. Wuhan University of Technology,2006.

[7] Haibo Kuang. Research on the prediction model of Chinese coastal Port Throughput [J]. Science Research Management, 2009,30(03):187-192.

[8] Hongfu Chi,Jiahong Wang,Yu Chen.Preliminary Study on the Methods of Forecast of Harbor Handling Capacity[J]. Journal of Zhejing Ocean University(Natural Science), 2001,(02):101-106.

[9] Wenya Zang.The combination forecast model of port throughput about Chongqing[D]. Chengdu University of Technology,2012.

[10] Fengping Wu,Yanguo Cao.Combined prediction method for port throughput based on induced ordered weighted averaging operator[J].Journal of Economics of Water Resources, 2013,31(05):11-14+75.

[11] Xiao Wu,Beihua Zong.Nonlinear combined prediction of port throughput based on support vector machine regression[J].Journal of Hohai University(Natural Sciences) ,2005,(04):482-484

[12] Meilian Liu,Meihua Zhu,Jian Huang.Screening of Influential Indices for Port Throughput Forecasting[J].Port & Waterway Engineering, 2011,(03):76-80.

[13] Jing Li,Jing Lv.Influence of Hiner's Economic Development on Port Throughput. Port & Waterway Engineering, 2007,(11):49-51

[14] Jinhe Xu.Study on Intrinsic Factors of Port Throughput Based on Principal Component Analysis[J]. Port & Waterway Engineering, 2010,(01):1-4.