Computer Life (CPL) ISSN: 1819-4818
DELIVERING QUALITY SCIENCE TO THE WORLD

Predicting Tianjin Port Container Throughput Based on Grey Relational Analysis

Li Yang

College of Transport and Communications, Shanghai Maritime University, Shanghai, China

891857322@qq.com

Abstract: The port throughput forecast is an essential reference for the development of port development plans and infrastructure investments. This article analyzes the impact indicators such as GDP, import and export trade volume, and industrial production value of Tianjin Port. Based on the statistics of container throughput of Tianjin Port in recent years, this paper uses the grey relational analysis to select the key indicators affecting the container throughput forecast and establishes a multiple linear regression model to predict the port throughput of Tianjin Port. The validity and practicability of the model in port container throughput forecasting are demonstrated by an example.
Keywords: Port throughput forecast, Grey correlation method, Linear multiple regression.

---

1. INTRODUCTION

Port throughput forecasting is an important part of port planning and its feasibility study. Its correct prediction is related to port development strategy, management tactics, the layout of port, berth locations, and so on. Tianjin Port is located in Tianjin, radiating inland hinterlands such as Northeast, North China and Northwest China. It connects Northeast Asia and Central and Western Asia. It is the gateway to the Beijing-Tianjin-Hebei region. It is the starting point of the eastern boundary of the Sino-Mongolian-Russian Economic Corridor, the important node of the New Eurasian Continental Bridge and the strategic fulcrum of the Century Maritime Silk Road in the 21st century. In 2017, the container throughput of Tianjin Port exceeded 15.07 million TEUs, ranking 11th in the world. With the increase in container throughput, it is necessary to improve the port infrastructure, so it is necessary to scientifically and accurately predict port throughput.

There are many methods for predicting port throughput. Wu (2019) predicts the container throughput of Guangzhou Port based on the time series model. The results show that the relative error between the predicted value and the actual value is less than 10%, which reflects the reliability of the model [3]. Liu and Tian (2012) used the Grey-Markov Combined Model to predict cargo throughput at Qinhuangdao Port [1]. In addition, An and Xu (2009) combined the grey forecasting model GM (1,1) with the Markov prediction method to predict the port throughput of Tianjin Port [2]. Zheng (2018) used the GM (1,1) model to predict the cargo throughput of Ningbo Port and verified the validity of the model [4]. However, these documents did not screen the factors affecting port throughput.

In order to deeply analyze the role and potential of Tianjin Port's national strategy of resource service, this paper takes Tianjin Port's container throughput over the years as the research object, uses the grey relational analysis to select the two factors with the highest degree of correlation, and establishes a multi-linear model to predict the container throughput of Tianjin Port in the next five years.

2. METHODOLOGY

2.1 Grey relational analysis

Grey Relational Analysis is a grey system method to measure the degree of correlation between factors. The specific steps are as follows:

Step1: Target behavior sequence

$$X_1 = (x_1(1),\ x_1(2),\ \ldots,\ x_1(m)) \tag{2-1}$$

Reference behavior data sequence:

$$X_i = (x_i(1),\ x_i(2),\ \ldots,\ x_i(m)),\quad i = 2,3,\ldots,n \tag{2-2}$$

Step2: Initialization is used to dimensionless the index sequence with the correlation operator to form the following matrix.

$$(X_1', X_2', \ldots X_m')^T = \begin{bmatrix} x_1'(1) & \ldots & x_1'(m) \\ M & M & M \\ x_n'(1) & \ldots & x_n'(m) \end{bmatrix} \tag{2-3}$$

$$x_i' = i\frac{x_i(k)}{\frac{1}{m}\sum_{k=1}^{m} x_i(k)} \qquad i = 1,2,\ldots,n; k = 1,2,\ldots,m \tag{2-4}$$

Step3: Calculate the absolute value of the difference between the target sequence and the reference sequence.

$$\Delta_i(k) = \left| x_1'(k) - x_i'(k) \right| \qquad k = 1,2,\ldots,m \qquad i = 1,2,\ldots,n \tag{2-5}$$

Make:

$$M = \max_{i=1}^{n} \max_{k=1}^{m} \Delta_i(k) = \left| x_1'(k) - x_i'(k) \right| \tag{2-6}$$

$$m = \min_{i=1}^{n} \min_{k=1}^{m} \Delta_i(k) = \left| x_1'(k) - x_i'(k) \right| \tag{2-7}$$

Step4: Computation of correlation coefficient and correlation degree

The correlation coefficients (Eq. 2-8) and correlation degrees (Eq. 2-9) of the corresponding elements of each reference sequence and the target sequence are calculated, and then the influencing factors are sorted according to the magnitude of correlation degree.

$$r(x_1'(k), x_i'(k)) = \frac{m + p \times M}{\Delta_i(k) + p \times M} \qquad k = 1,2,\ldots,m \qquad 0 \le p \le 1 \tag{2-8}$$

$$r(X_1, X_i) = \frac{1}{m}\sum_{k=1}^{m} r_{1i}(k) \tag{2-9}$$

2.2 Regression model

Here we defined the port container throughput as the dependent variable Y, and defined the two influencing factors with the largest correlation degree as the independent variables x1, x2, then established a regression model (Eq. 2-10).

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \tag{2-10}$$

3. PREDICTING CONTAINER THROUGHPUT OF TIANJIN PORT

3.1 Analysis of influencing factors of container throughput in Tianjin Port

3.1.1 Port's own influencing factors

The influencing factors of the port itself mainly include its natural conditions such as hydrology and meteorology, the length of the inbound channel, the area of the warehouse, the number of berths at the dock, and the capacity of the container yard. Because the natural conditions of the port are difficult to quantify, the length of the inbound channel does not change significantly, and the warehouse area varies with the berth number, this paper chooses the berth number (number) and container yard capacity (10,000 TEU) as the influencing factors of the model.

3.1.2 Objective regional influence factor

The objective influencing factors of port throughput are as follows: In terms of the economic environment, its main influencing factor is GDP value. In terms of port operations, it mainly refers to import and export business, and the total export volume of foreign trade (in millions of US dollars) reflects the number of port throughput. The operation of the port is becoming more and more autonomous, and the direct intervention of the state will be less and less. The state's macroeconomic regulation and control are inseparable from taxation, so the total tax revenue (100 million yuan) is also used as an influencing factor.

To sum up, this paper finally chooses the berth number (number), container yard capacity (10,000 TEU), social fixed assets investment (100 million yuan), Tianjin GDP (100 million yuan), total foreign trade exports (million dollars), total tax revenue (100 million yuan) as the impact factors of the model.

3.2 Predicting container throughput

Port throughput forecasting is the basic traffic forecasting of the port, which will play an important role in the future port development and management tactics.

3.2.1 Original data of factors affecting container throughput in Tianjin Port

There are several factors affecting port throughput. According to the daily experience, the data of conventional factors are selected to analyze and predict, and then the key factors can be selected according to the degree of relevance. Figure 1 shows the container throughput of the Tianjin area from 2007 to 2017, and Table 1 shows the data of the port throughput impact factors from 2007 to 2017.

3.2.2 Grey relational analysis of influencing factors of port container throughput

This article uses MATLAB to calculate the grey correlation degree, and the steps are as follows:

Firstly, the reference column is selected, and then dimensionless quantization is processed to get Table 2.

Secondly, the absolute difference is calculated, as shown in Table 3.

Thirdly, select p=0.5, calculating the grey correlation coefficient of the port throughput influencing factors, as shown in Table 4.

Finally, using the grey correlation formula, the correlation degree is calculated according to the correlation coefficient table, as shown in Table 5.

Sorting by the size of the correlation degree in Table 5, two factors that have a great impact on the throughput of Tianjin Port are berth number and industrial added value.

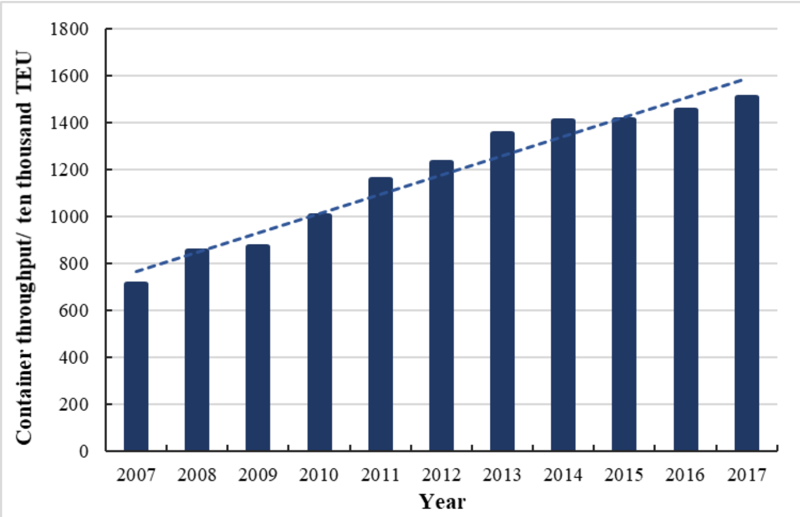

Fig.1 Container throughput of Tianjin Port from 2007 to 2017

Table 1 Influencing factors of container throughput forecast in Tianjin Port

Particular year	Berth number	Container yard capacity (10,000 TEU)	Social fixed assets investment (RMB 100 million)	Tianjin GDP (RMB 100 million)	Total foreign trade exports ($1 million)	Total tax revenue (RMB 100 million)	Industrial added value （RMB 100 million）
2007	142	41.75	2353	5252	71549	438	2669
2008	139	42.01	3389	6719	80538	546	3534
2009	134	49.05	4738	7521	63944	614	3750
2010	151	60.23	6278	9224	82200	776	4411
2011	154	59.52	7067	11307	103391	1004	5381
2012	158	62.41	7934	12893	115622	1105	6123
2013	159	62.87	8752	14370	128528	1261	6679
2014	162	65.57*	11654	15722	133912	1486	7083
2015	173	74.77*	13066	16538	114347	1578	6981
2016	176	77.28*	14629	17885	102651	1624	7239
2017	178*	78.95*	11274	18595	115790	1612	6864

Data source: Tianjin Statistical Yearbook of National Economic and Social Development, Port Yearbook, “*” is an estimate.

Table 2 Dimensionless quantitative processing results

Particular year	Berth Number	Container yard capacity	Social fixed assets investment	Tianjin GDP	Total foreign trade exports	Total tax revenue	Industria l added value
2007	0.797753	0.5294864	0.160845	0.282442	0.534299	0.269704	0.368697
2008	0.780899	0.5327838	0.231663	0.361334	0.601425	0.336207	0.488189
2009	0.752809	0.6220672	0.323877	0.404464	0.477508	0.378079	0.518027
2010	0.848315	0.7638554	0.429148	0.496047	0.613836	0.477833	0.609338
2011	0.865169	0.754851	0.483082	0.608067	0.772082	0.618227	0.743335
2012	0.88764	0.7915029	0.542347	0.693358	0.863418	0.680419	0.845835
2013	0.893258	0.7973367	0.598264	0.772788	0.959794	0.776478	0.922641
2014	0.910112	0.8315789	0.796637	0.845496	1	0.915025	0.97845
2015	0.97191	0.9482562	0.893157	0.889379	0.853897	0.971675	0.96436
2016	0.988764	0.9800888	1	0.961818	0.766556	1	1
2017	1	1	0.770661	1	0.864672	0.992611	0.948197

Table 3 Absolute difference

Particular year	Berth Number	Container yard capacity	Social fixed assets investment	Tianjin GDP	Total foreign trade exports	Total tax revenue	Industria l added value
2007	0.202247	0.4705136	0.839155	0.717558	0.465701	0.730296	0.631303
2008	0.219101	0.4672162	0.768337	0.638666	0.398575	0.663793	0.511811
2009	0.247191	0.3779328	0.676123	0.595536	0.522492	0.621921	0.481973
2010	0.151685	0.2361446	0.570852	0.503953	0.386164	0.522167	0.390662
2011	0.134831	0.245149	0.516918	0.391933	0.227918	0.381773	0.256665
2012	0.11236	0.2084971	0.457653	0.306642	0.136582	0.319581	0.154165
2013	0.106742	0.2026633	0.401736	0.227212	0.040206	0.223522	0.077359
2014	0.089888	0.1684211	0.203363	0.154504	0	0.084975	0.02155
2015	0.02809	0.0517438	0.106843	0.110621	0.146103	0.028325	0.03564
2016	0.011236	0.0199112	0	0.038182	0.233444	0	0
2017	0	0	0.229339	0	0.135328	0.007389	0.051803

Table 4 Coefficient table of correlation degree

Particular year	Berth Number	Container yard capacity	Social fixed assets investment	Tianjin GDP	Total foreign trade exports	Total tax revenue	Industrial added value
2007	0.674752	0.4713872	0.333333	0.368977	0.47395	0.36489	0.399263

2008	0.656946	0.4731399	0.353205	0.396485	0.512835	0.387289	0.450486
2009	0.62927	0.5261092	0.382931	0.41333	0.445378	0.402859	0.465396
2010	0.734474	0.639871	0.423632	0.454319	0.520735	0.445532	0.517844
2011	0.756801	0.6312032	0.448029	0.517033	0.648	0.523588	0.620454
2012	0.788773	0.6680377	0.478298	0.577756	0.754419	0.567642	0.731299
2013	0.797192	0.6743009	0.510861	0.648708	0.912555	0.65243	0.844329
2014	0.823565	0.713569	0.673543	0.730868	1	0.831583	0.951148
2015	0.937253	0.8902154	0.797039	0.791359	0.741721	0.936761	0.921707
2016	0.973919	0.9546946	1	0.916589	0.642517	1	1
2017	1	1	0.646582	1	0.756125	0.982694	0.890104

Table 5 Correlation degree table

	Berth Number	Container yard capacity	Social fixed assets investment	Tianjin GDP	Total foreign trade exports	Total tax revenue	Industrial added value
Correlation degree	0.79754	0.6947753	0.549769	0.619584	0.673476	0.645024	0.708366

3.2.3 Prediction of container throughput in Tianjin Port

As can be seen from Table 5, the correlation degree between berth number and industrial added value is relatively high. Therefore, the two factors are selected as relevant variables, and a regression model is established to predict the container throughput of Tianjin Port.

Taking the container throughput of Tianjin Port as dependent variable Y, the berth number and industrial added value as independent variables x1 and x2 respectively, and the regression model is established and solved by Excel.

According to the regression analysis table, the forecast model of Tianjin Port container throughput is Eq.3-1.

$$Y = 4.0120709x_1 + 0.133957x_2 - 192.17264 \qquad (3\text{-}1)$$

The results obtained by using EXCEL solution are shown in Table 6.

It can be seen from Table 6 that although the maximum relative error reaches 4.35%, the average error is 1.87% (average of relative error), less than 3%, and achieves high precision, which is an effective prediction. Fig.2 shows that the fitting results of the two variables are good. Therefore, the model can predict the container throughput of Tianjin Port. The forecast value of container throughput of Tianjin Port in the next five years is shown in Table 7.

Table 6 Tianjin Port container throughput forecast

Particular year	Observation value	Predictive value Y	Residual	Relative error (%)
2007	710	735.0726308	-25.07263075	0.035313564
2008	850	838.9092151	11.09078489	0.013047982
2009	870	847.7835708	22.21642918	0.025536125
2010	1000	1004.534346	-4.534346141	0.004534346
2011	1158	1146.50884	11.49116037	0.009923282
2012	1230	1261.95321	-31.95321011	0.02597822
2013	1350	1340.445368	9.554632238	0.007077505
2014	1406	1406.600205	-0.600204543	0.000426888
2015	1411	1437.069371	-26.06937088	0.018475812
2016	1452	1483.666487	-31.66648702	0.021808875
2017	1507	1441.456757	65.54324276	0.04349253

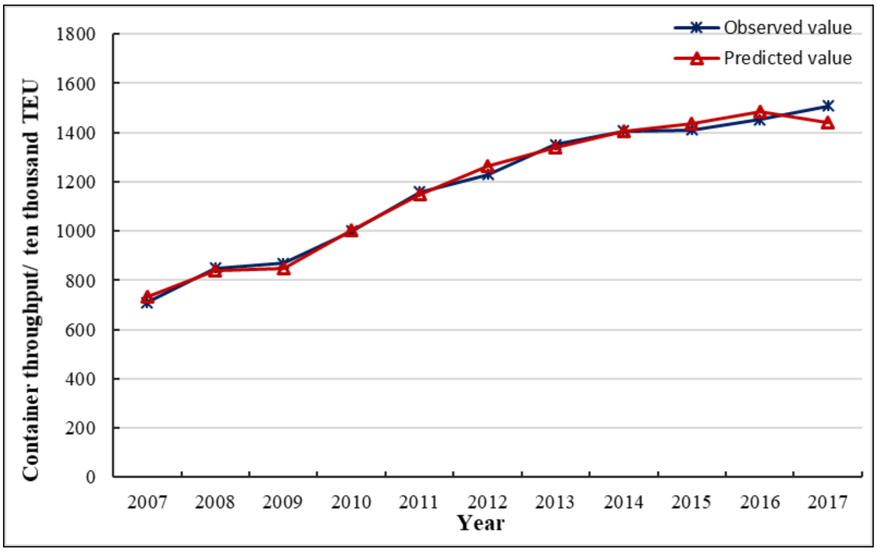

Fig.2 Comparison of observed and predicted values

Table 7 Container throughput forecast of Tianjin Port in 2018-2022

Particular year	2018	2019	2020	2021	2022
Container throughput (10,000 TEU)	1662.5	1742.04	1821.57	1901.1	1984.64

4. CONCLUSION

In this paper, the factors affecting container throughput are determined by grey relational analysis, and the regression model is used to predict the container throughput of Tianjin Port from 2018 to 2022. The forecast results show that the container throughput of Tianjin Port will increase steadily. According to this growing trend, the existing port conditions in Tianjin Port cannot meet the needs of its ports. Therefore, the port should strengthen the construction of the transportation network system, build a combined port group, and coordinate development.

REFERENCES

[1] Liu S., Tian L., The application of Grey-Markov combined model for port cargo throughput forecasting[C]. International Conference on Fuzzy Systems & Knowledge Discovery. 2012.

[2] An L., Xu H., Prediction Based on Grey Model of Cargo Throughput of Tianjin Port[C]. International Conference on Information Engineering & Computer Science. IEEE, 2009.

[3] Wu Chen, Prediction of Port Container Throughput Based on Time Series Model[J]. Pearl River Water Transport, 2019, 477 (05): p.75-76. (In Chinese)

[4] Zheng H., Research on Forecast of Cargo Throughput of Ningbo Zhoushan Port Based on Grey Model[J]. Market Weekly, 2018(10): p.5-6+13. (In Chinese)