https://doi.org/10.62836/emi.v4i1.284

Global Science Publishing

EMI

Economics & Management Information

https://ojs.sgsci.org/journals/emi

Article

Qingdao Port Throughput Prediction-Based on Grey Prediction Model

Yanping Yang

College of Economics and Management, Qingdao University of Science and Technology, Qingdao266100, China

Abstract: Since Qingdao has been integrated into the "Belt and Road" initiative, the throughput of Qingdao Port has been developing rapidly, but at the same time, the rapid growth of throughput has overwhelmed port infrastructure. Therefore, it is of great significance to measure the throughput scale and study the development trend of throughput. Based on the data of cargo throughput and container throughput of Qingdao port from 2019 to 2022, the grey prediction GM(1,1) model was established and tested using EXCEL, Pycharm64, MATLAB, and mathematical modeling methods. The model is proven to be available, and the cargo and container throughputs of Qingdao Port from 2023 to 2026 are predicted. The results show that the cargo and container throughputs of Qingdao Port increase annually, and the development potential of container transportation is huge. The port department should strengthen the planning and design of the port's throughput capacity, improve it, and meet the growing actual needs of the port.

Keywords: Qingdao port; throughput; grey prediction model

1. Introduction

Ports play a key role in economic development, as an important part of international trade. Qingdao Port is one of the most important ports in China and one of the busiest ports in the world. The container throughput of Qingdao Port in 2020. In 2020, the cargo throughput of Qingdao Port ranked fifth globally, container throughput ranked sixth globally, and first in Northeast Asia. In 2021, Qingdao Port's cargo and container throughput will rank fifth in the country. In 2022, Qingdao Port's cargo and container throughput will rank fourth in the country. Forecasting the throughput of Qingdao Port is of great significance for understanding the development trend of China and global trade, formulating logistics strategies, and optimizing resource allocation. Port throughput is affected by many factors, including domestic and foreign economic situations, trade policy, route arrangement, cargo demand, and transportation technology. The change in these factors complicates and makes the prediction of port throughput challenging. The accurate prediction of port throughput is of great significance for port operations and management. This can help the port formulate reasonable plans and strategies to adapt to the changing cargo flow and to make resource allocation, equipment scheduling, and personnel arrangements in advance. In addition, throughput prediction plays a guiding role in port development planning, port expansion, and the decision-making of freight companies. To accurately understand the future development trend of cargo throughput and container throughput of Qingdao Port and develop a good port development plan, it is necessary to establish a prediction model with high accuracy to predict the cargo throughput and container throughput of

Received: 20 December 2024; Accepted: 6 January 2025.

Corresponding: Yanping Yang (779673085@qq.com)

Qingdao Port.

For the prediction of port throughput, relevant experts and scholars also use different methods, including statistical analysis methods based on historical data, such as time series analysis, regression analysis, etc., as well as prediction models based on machine learning and artificial intelligence, such as neural networks, support vector machines, random forests, etc. This paper takes Qingdao Port as the research object. Because many factors affect port cargo and container throughput, the quantification of many influencing factors is difficult. In the case where the internal and external influencing factors of port throughput cannot be known and there are not many analytical samples available, it is reasonable to use the grey prediction model to construct the throughput prediction model of Qingdao Port from the historical data of port cargo throughput and container throughput. The GM(1, 1) model has the advantages of requiring a small number of data samples and a high short-term prediction accuracy [1]. Therefore, this study uses the grey GM(1,1) model to predict and analyze the cargo throughput and container throughput of Shenzhen Port from 2023 to 2026 based on the data of cargo throughput and container throughput of Shenzhen Port from 2019 to 2022 to provide scientific and accurate methods and models. It provides a practical reference for port departments to make scientific decisions and helps port managers and decision makers make reasonable predictions, plans, and decisions to promote the sustainable development and economic prosperity of ports.

At present, in the study of port throughput, the artificial neural network model is mainly used to predict the freight volume of Florida port in the United States (Klodzinski & Al-Deek, 2015) [2], the multiple linear regression and three-layer BP neural network model are used to predict the throughput of Fangcheng port (Fan Linsheng et al., 2015) [3], and the gray Markov model and ARMA-weighted Markov model are used to predicting the throughput of the port (Shi Leilei, 2015) [4]. Using system clustering and typical index analysis, the main influencing factors were selected, and a multiple regression model was established to predict the throughput of Yingkou Port (Li Guizhi et al., 2015) [5]. Using a regression model, the GDP data of Fujian Province were used to predict the cargo throughput of coastal ports in 2020 (Lin Jian, 2016) [6]. The ant colony algorithm was used to optimize the BP neural network model, fuzzy neural network prediction model, RBF prediction model, and BP prediction model to predict the throughput of a port. The ant colony algorithm optimizes the BP neural network model with the fastest convergence speed (Li, 2020) [7].

2. The Application of GM(1,1) Model in Qingdao Port Throughput Prediction

The data for this study are from the “China Port Statistics” (Table 1). For the convenience of research, the letters g and C are used to represent the cargo throughput and container throughput respectively, where the cargo throughput is represented by g_{SZ} and the container throughput is represented by C_{SZ}.

Table 1. Data of cargo throughput and container throughput of Qingdao Port from 2019 to 2022.

Year	2019	2020	2021	2022
Cargo throughput (10 million tons)	57.74	60.46	63.03	65.75
Container throughput (10 million TEU)	2.101	2.201	2.371	2.567

2.1. Stage Ratio Test and Modeling Feasibility Judgment

2.1.1. Establish Port Cargo Throughput and Container Throughput Data Time Series

According to the gray prediction model, the established time series of port cargo throughput and container throughput data are as follows：

$$X^{(0)}g_{sz} = \left(X^{(0)}g_{SZ}(1), X^{(0)}g_{SZ}(2), X^{(0)}g_{SZ}(3), X^{(0)}g_{SZ}(4) \right) = (57.74,\ 60.46,\ 63.03,\ 65.75)$$

$$X^{(0)}C_{sz} = \left(X^{(0)}C_{SZ}(1), X^{(0)}C_{SZ}(2), X^{(0)}C_{SZ}(3), X^{(0)}C_{SZ}(4) \right) = (2.101,\ 2.201,\ 2.371,\ 2.567)$$

2.1.2. Get the Grade Ratio

Run with the matlab code according to the following formula:

$$\lambda g_{SZ}(k)=\frac{X^{(0)}g_{SZ}(k-1)}{X^{(0)}g_{SZ}(k)}$$

The code was run with the results as shown belows:

$$\lambda g_{SZ}(k)=\left(\lambda g_{SZ}(2),\lambda g_{SZ}(3),\lambda g_{SZ}\ (4)\right)=(1.0471, 1.0425, 1.0432)$$

ditto,

$$\lambda C_{SZ}(k)=\frac{X^{(0)}C_{SZ}(k-1)}{X^{(0)}C_{SZ}(k)}$$

After running with matlab code, the level is obtained as follows:

$$\lambda C_{SZ}(k)=\left(\lambda C_{SZ}(2),\lambda C_{SZ}(3),\lambda C_{SZ}\ (4)\right)=(1.0476, 1.0772, 1.0827)$$

2.1.3. Grade Ratio Judgment

Calculate the maximum and minimum values of the level ratio and determine whether it is within a predetermined range. The predetermined range is as follows:

$$\lambda g(k)\in\left(e^{-\frac{2}{n+1}},e^{\frac{2}{n+1}}\right),\lambda C(k)\in\left(e^{-\frac{2}{n+1}},e^{\frac{2}{n+1}}\right)$$

From the stage ratio obtained in Section 2.2.2, it can be seen that, Because all $\lambda C_{SZ}(k)$ and $\lambda C_{SZ}(k)$ are in the boundary region $[0.6703,\ 1.4918]$, the GM(1,1) model with $X^{(0)}g_{sz}$、 $X^{(0)}C_{sz}$ are satisfactory.

2.2. Construction of GM(1,1) Model

2.2.1. Make an Accumulation of the Original Data

Through the running results of pycharm software, it can be seen that:

$$X^{(1)}g_{GZ}(k)=\sum_{m=1}^{k}X^{(0)}(m)\,(k=1,2,3,4)$$

$$X^{(1)}g_{SZ}=(X^{(1)}g_{SZ}(1),X^{(1)}g_{SZ}(2),X^{(1)}g_{SZ}(3),X^{(1)}g_{SZ}(4))$$

Then: $X^{(1)}g_{SZ}$=(57, 74, 118.2, 181.23, 246.98)

Similarly, the running results of pycharm software show that

$$X^{(1)}C_{GZ}(k)=\sum_{m=1}^{k}X^{(0)}(m)\,(k=1,\ 2,\ 3,\ 4)$$

$$X^{(1)}C_{SZ}=(X^{(1)}C_{SZ}(1),X^{(1)}C_{SZ}(2),X^{(1)}C_{SZ}(3),X^{(1)}C_{SZ}(4))$$

Then: $X^{(1)}C_{SZ}$=(2.101, 4.302, 6.67, 9.24)

2.2.2. Construct Data Matrix and Data Vector

Data matrix *Bg* and data vector *Yg* are constructed.

Then:

$$\begin{cases} Z^{(1)}g_{GZ}(2)=\frac{1}{2}\left[X^{(1)}g_{GZ}(1)+X^{(1)}g_{GZ}(2)\right]=\frac{1}{2}[57.74+118.2]=87.97 \\ Z^{(1)}g_{GZ}(3)=\frac{1}{2}\left[X^{(1)}g_{GZ}(2)+X^{(1)}g_{GZ}(3)\right]=\frac{1}{2}[118.2+181.23]=149.71 \\ Z^{(1)}g_{GZ}(4)=\frac{1}{2}\left[X^{(1)}g_{GZ}(3)+X^{(1)}g_{GZ}(4)\right]=\frac{1}{2}[181.23+246.98]=214.11 \end{cases}$$

So we get:

$$Yg_{GZ}=\begin{bmatrix}X^{(0)}g_{GZ}(2)\\X^{(0)}g_{GZ}(3)\\X^{(0)}g_{GZ}(4)\end{bmatrix}=\begin{bmatrix}60.46\\63.03\\65.75\end{bmatrix}$$

$$Bg_{sz}=\begin{bmatrix}-Z^{(0)}g_{GZ}(2) & 1\\-Z^{(0)}g_{GZ}(3) & 1\\-Z^{(0)}g_{GZ}(4) & 1\end{bmatrix}=\begin{bmatrix}-87.97 & 1\\-149.71 & 1\\-214.11 & 1\end{bmatrix}$$

Data matrix *Bc* and data vector *Yc* are constructed.

Then:

$$\begin{cases} Z^{(1)}C(2) = \frac{1}{2}\left[X^{(1)}g_{GZ}(1) + X^{(1)}g_{GZ}(2)\right] = \frac{1}{2}[2.101 + 4.302] = 3.2015 \\ Z^{(1)}gC_{GZ}(3) = \frac{1}{2}\left[X^{(1)}g_{GZ}(2) + X^{(1)}g_{GZ}(3)\right] = \frac{1}{2}[4.302 + 6.67] = 5.4875 \\ Z^{(1)}gC_{GZ}(4) = \frac{1}{2}\left[X^{(1)}g_{GZ}(3) + X^{(1)}g_{GZ}(4)\right] = \frac{1}{2}[6.67 + 9.24] = 7.9565 \end{cases}$$

So we get:

$$YC_{GZ} = \begin{bmatrix} X^{(0)}g_{GZ}(2) \\ X^{(0)}g_{GZ}(3) \\ X^{(0)}g_{GZ}(4) \end{bmatrix} = \begin{bmatrix} 2.201 \\ 2.371 \\ 2.567 \end{bmatrix}$$

$$BC_{sz} = \begin{bmatrix} -Z^{(1)}g_{GZ}(2) & 1 \\ -Z^{(1)}g_{GZ}(3) & 1 \\ -Z^{(1)}g_{GZ}(4) & 1 \end{bmatrix} = \begin{bmatrix} -3.2015 & 1 \\ -5.4875 & 1 \\ -7.9565 & 1 \end{bmatrix}$$

2.2.3. Least Squares Estimation for Parameter Series

The least squares estimation of the parameter sequence Pg = (a, b)^T

$$P^{\wedge}g_{GZ} = (a^{\wedge}b^{\wedge})^T = (Bg_{GZ}{}^T Bg_{GZ}{}^T)^{-1} Bg_{GZ}{}^T Yg_{GZ}{}^T = \begin{pmatrix} -0.0419 \\ 56.7638 \end{pmatrix}$$

So get a = -0.0419, b = 56.7638.

Similarly, the least squares estimation of the parameter sequence Pc = (a, b)^T

$$P^{\wedge}C_{GZ} = (a^{\wedge}b^{\wedge})^T = (BC_{GZ}{}^T B_{GZ}{}^T)^{-1} BC_{GZ}{}^T YC_{GZ}{}^T = \begin{pmatrix} -0.077 \\ 1.9524 \end{pmatrix}$$

So get a = -0.077, b = 1.9524

2.2.4. Modeling

According to the parameters obtained in part 2.2.3, the gray prediction model is established as follows:

$$X^{(0)}g_{SZ}(k) - 0.0419Z^{(1)}g_{SZ}(k) = 56.7638$$

$$X^{(0)}C_{SZ}(k) - 0.077Z^{(1)}C_{SZ}(k) = 1.9524$$

The time response sequence is solved as

$$X^{\wedge(1)}g_{GZ}(k+1) = \left(X^{(0)}g_{GZ}(1) - \frac{a^{\wedge}}{b^{\wedge}}\right)e^{-a^k} + \frac{b^{\wedge}}{a^{\wedge}}$$

$$X^{\wedge(1)}C(k+1) = \left(X^{(0)}C_{GZ}(1) - \frac{a^{\wedge}}{b^{\wedge}}\right)e^{-a^k} + \frac{b^{\wedge}}{a^{\wedge}}$$

2.2.5. Find the Generated Sequence Value and Model Reduction Value

Find the generated sequence value $X^{\wedge(0)}g(k+1)$ and the model reduction value $X^{\wedge(0)}g(k+1)$

Substituting k = 1, 2, 3 into the time response function, we can get $X^{\wedge(1)}g_{GZ}(k+1)$:

$$X^{\wedge(1)}g_{SZ} = (57.74,\ 118.2,\ 181.23,\ 246.98)$$

(Where $X^{\wedge(1)}g_{GZ}(1) = X^{\wedge(0)}g_{GZ}(1) = X^{\wedge(0)}g_{GZ}(1) = 57.74$)

$X^{\wedge(0)}g_{GZ}(k) = X^{\wedge(1)}g_{GZ}(k) - X^{\wedge(1)}g_{GZ}(K-1)$ is generated by decrement.

The reduction value is obtained:

$$X^{\wedge(0)}g_{GZ} = (X^{\wedge(0)}g_{GZ}(1), X^{\wedge(0)}g_{GZ}(2), X^{\wedge(0)}g_{GZ}(3), X^{\wedge(0)}g_{GZ}(4)) = (57.74,\ 60.44,\ 63.03,\ 65.73)$$

Similarly, Find the generated sequence value $X^{\wedge(1)}C_{GZ}(k+1)$ and the model reduction value $X^{\wedge(0)}C_{GZ}(k+1)$

Substituting k = 1, 2, 3 into the time response function, we can get $X^{\wedge(1)}C_{GZ}(k)$:

$$X^{\wedge(1)}C_{GZ} = (2.101, 4.302, 6.673, 9.24)$$

(Where $X^{\wedge(1)}C_{GZ}(1) = X^{\wedge(0)}C_{GZ}(1) = X^{\wedge(0)}C_{GZ}(1) = 2.101$)

$X^{\wedge(0)}C_{GZ}(k) = X^{\wedge(1)}C_{GZ}(k) - X^{\wedge(1)}C_{GZ}(K-1)$ is generated by decrement.

The reduction value is obtained:

$$X^{\wedge(0)}g_{GZ}=(X^{\wedge(0)}g_{GZ}(1),X^{\wedge(0)}g_{GZ}(2),X^{\wedge(0)}g_{GZ}(3),X^{\wedge(0)}g_{GZ}(4))=(2.1010,2.1977,2.3737,2.5637)$$

2.3. Model Test

2.3.1. Cargo Throughput

As we know above, the original and forecast values of the cargo throughput of Qingdao Port from 2019 to 2022 are shown in the table below (As shown in Table 2):

Table 2. Inspection Table of Cargo Throughput of Qingdao Port.

Year	2019	2020	2021	2022
Raw value	57.74	60.46	63.03	65.75
Predicted value	57.74	60.44	63.03	65.73

The average residual $e=0.007387$

The variance of historical data $S_{12}=8.8456$

The residual variance $S_{22}=8.2043\times10^{-5}$

The minimum error probability $P=1>0.95$

The mean square deviation ratio $C=\dfrac{S_{22}}{S_{12}}=9.275\times10^{-6}<0.35$(as shown in Table 3)

Table 3. Model Accuracy Level.

Model accuracy level	Mean square deviation ratio
Level 1 (Excellent)	C ⩽ 0.35
Level 2 (Pass)	0.35 < C ⩽ 0.50
Level 3 (Basic Pass)	0.50 < C ⩽ 0.65
Level 4 (Fail)	C > 0.65

It was verified that the model error was small and that it could be predicted.

2.3.2. Container Throughput

Also known from the above, the original and forecast values of the container throughput of Qingdao Port from 2019 to 2022 are shown in the table below (as shown in Table 4):

Table 4. Inspection Table of Container Throughput of Qingdao Port.

Year	2019	2020	2021	2022
Raw value	2.101	2.201	2.371	2.567
Predicted value	2.101	2.1977	2.3737	2.5637

The average residual $e=0.00019$

The variance of historical data $S_{12}=0.03133$

The residual variance $S_{22}=6.24919\times10^{-6}$

The minimum error probability $P=1>0.95$

The mean square deviation ratio $C=\dfrac{S_{22}}{S_{12}}=0.000199<0.35$(as shown in Table 3)

It was verified that the model error was small and that it could be predicted.

2.4. Qingdao Port Throughput Forecast

According to the gray model, the throughput of Qingdao Port from 2022 to 2026 was predicted, and the following results were obtained (as shown in Table 5):

Table 5. Data of cargo throughput and container throughput of Qingdao Port from 2023 to 2026.

Year	2023	2024	2025	2026
Cargo throughput (10 million tons)	68.54867	71.48484	74.54677	77.73986
Container throughput (10 million TEU)	2.76886	2.99049	3.22987	3.48841

3. Conclusion and Discussion

According to the relevant data of cargo throughput and container throughput of Qingdao port, we established a GM(1, 1) grey prediction model to predict the future throughput. Through the analysis of the model, we obtained the average phase residual error and accuracy index, which verified the availability and accuracy of the model. Specifically, the forecast results of the model show that the cargo throughput and container throughput of Qingdao port will continue to grow in the next few years.

According to the forecast data, the cargo throughput of Qingdao port is expected to increase by about 20 million tons in 2026 compared with 2019, while the container throughput is expected to increase by about 13.874 million TEUs. This growth trend is similar to that of Shenzhen Port, and the cargo throughput and container throughput of the latter also show an increasing trend year by year. This shows that with the continuous development of global trade and the prosperity of the regional economy, the demand for port throughput capacity will continue to increase. However, the increasing throughput of Qingdao port will inevitably cause great pressure on the carrying capacity of the port. Although the throughput growth of Qingdao Port is gratifying, the design and planning of its cargo throughput capacity and container throughput capacity lag behind the actual development needs. This mismatch may lead to congestion in the peak period of the port, affecting the timely transportation of goods and the overall operational efficiency of the port. Therefore, the design and planning of Qingdao Port ‘s throughput capacity must be adapted to the actual needs to improve the comprehensive benefits of the port. To this end, the port management department should formulate corresponding strategies and measures based on the results of the prediction model to ensure that port facilities and services can meet future needs. This includes upgrading the existing facilities, adding new handling equipment, optimizing the operation process, etc.

Through the accurate prediction of the future development trend of Qingdao port cargo throughput and container throughput, the port department can provide scientific reference for the future development strategy of the port, the planning and construction of container terminals and other major decisions. This not only helps to improve the operational efficiency of the port, but also promotes the sustainable development of the port and ensures that Qingdao Port maintains its advantages in the fierce market competition. In summary, the establishment of GM(1,1) grey prediction model provides important theoretical support and practical guidance for the future development of Qingdao Port. Through scientific prediction and reasonable planning, Qingdao Port can better cope with future challenges and achieve efficient and sustainable development.

Funding

Not applicable.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable.

Conflicts of Interest

The author declares no conflict of interest.

References

1 Yan X. Research on Supply and Demand Forecasting of High-Skilled Logistics Talents in Guangdong Province Based on Grey Model. *Practice and Understanding of Mathematics* 2017; **47**(**17**): 313–320.

2 Al-DeekHM K. Transferability Ofaninter Modal Freight Transportation Forecasting Model to Major Florida Lseaports. *Transportation Research Record* 2003; **1820**: 36–45.

3 Fan L, Deng J, Hao M, *et al*. Fangchenggang Port Throughput Prediction Based on TEI@I Methodology. *Logistics Technology* 2015; **19**: 75–79.

4 Shi L. Application of Combination Prediction Based on Port Throughput Prediction. *Modern Economic Information* 2015; **16**: 86.

5 Li G, Bai Q. Analysis of Influencing Factors of Yingkou Port Throughput Based on System Clustering. *Logistics Technology* 2015; **1**: 189–191.

6 Lin J. Forecast and Analysis of Port Throughput in Fujian Province during the 13th Five-Year Plan Period. *Port Economy* 2016; **11**: 14–16.

7 Li C, Lu X, Wu Z, *et al*. Port Throughput Prediction Based on Ant Colony Algorithm to Optimize Backpropagation Neural Network. *Journal of Metrology* 2020; **41**(**11**): 1398–1403.

© *The Author(s) 2025. Published by Global Science Publishing* (*GSP*).

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.