Container throughput prediction of some ports in the Yangtze River Delta Based on GM (1, 1)

Xinna Liu*
College of Transport and Communications, Shanghai Maritime University, Shanghai, 201306, China

*Corresponding author e-mail: 202030610163@stu.shmtu.edu.cn

Abstract. The MATLAB implementation of grey neural network prediction modeling has been widely used.[1] Firstly, this paper briefly introduces the prediction process of grey GM (1, 1) model, constructs the data of container throughput of some main ports in the Yangtze River Delta from 2016 to 2020, uses the powerful scientific calculation function of MATLAB, first accumulates the correlation coefficients a and B, and obtains the corresponding prediction model; Secondly, verify the original data with residuals and relative errors, and then predict the container throughput from 2021 to 2025. According to the prediction results, it provides suggestions for the logistics development of Shanghai port, which can provide a certain theoretical basis for the government departments to formulate the future development plan of the port to a certain extent.

Key words: container throughput; GM (1, 1) model; Matlab calculation.

1. Introduction
At present, the world economic growth is sluggish, international economic and trade frictions are intensified, and the downward pressure on the domestic economy is increasing. Port throughput prediction plays an important role in the formulation of port development planning, channel engineering feasibility study, resource allocation and operation and management strategy.

As a smart and efficient container hub port and an international shipping center with global influence, Shanghai's container throughput is an important indicator of port construction capacity and international competitiveness. The accurate prediction of container throughput is to timely resolve bottlenecks and risks, an important strategic goal of port development, as well as Shanghai's port city linkage, promoting the city with the port, port scientific planning an important basis and strategic demand for sustainable development.

There are many models and methods for predicting port container throughput, such as trend curve model, grey prediction method, BP neural network method, exponential smoothing model method, regression model method, etc. [2].

Zhu Jian [3] found that the development of the world economy has the greatest impact on the cargo throughput of China's coastal ports. Huang Yuchua and Chen Haishan [4] put forward the optimized GM (1, 1) power model, which is better for port throughput prediction.
In this paper, GM (1, 1), a widely applicable grey system theory and prediction model with low requirements for original data but high prediction accuracy, is selected. Its main principle is to use accumulation operation to weaken the dynamic randomness of actual data, generate a strong regular data sequence, and then establish differential equations to predict the future development status and trend of things.

2. Prediction principle of GM (1, 1) model

2.1. Model Steps

Firstly: Set sequence \( X^{(0)} = X^{(0)}(1), X^{(0)}(2), X^{(0)}(3), \ldots, X^{(0)}(n) \), among them, \( X^{(0)}(k) \geq 0, k = 1, 2, \ldots, n \), \( X^{(1)} \) is \( X^{(0)} \) ' sequence of 1–AGO: \( X^{(1)} = X^{(1)}(1), X^{(1)}(2), X^{(1)}(3), \ldots, X^{(1)}(n) \), then \( X^{(1)}(k) = \sum_{i=1}^{k} X^{(0)}(i), k = 1, 2, \ldots, n \);

Secondly: Generation of nearest neighbor mean for \( X^{(1)} \), make \( Z^{(1)}(k) = \frac{1}{2}X^{(1)}(k) + X^{(1)}(k - 1), k = 2, 3, \ldots, n \), we can get the generation sequence of \( X^{(1)} \) adjacent to the mean: \( Z^{(1)} = Z^{(1)}(2), Z^{(1)}(3), Z^{(1)}(4), \ldots, Z^{(1)}(n) \);

Thirdly: Establish GM (1, 1) model,

\[
X^{(0)}(k) + az^{(1)}k = b;
\]

Fourthly: Solving the differential equation: let \( X^{(1)} \) satisfy the first-order differential equation,

\[
\frac{dx^{(1)}}{dt} + ax^{(1)} = b.
\]

Among them, the parameter \( m \) is the development coefficient of the model, which is mainly used to express the development trend of \( X^{(0)} \) original data sequence and \( X^{(1)} \) newly generated data sequence; Parameter \( n \) is the coordination coefficient of the model, which is mainly used to express the transformation relationship between data sequences. The development coefficient \( m \) and coordination coefficient \( n \) are usually calculated by the least square method, as follows:

\[
\left( a \right) = \left( B^T N \right)^{-1} B^T Y,
\]

Among them,

\[
Y = \begin{bmatrix} X^{(0)}(2) \\ X^{(0)}(3) \\ \vdots \\ X^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -Z^{(2)} \\ -Z^{(3)} \\ \vdots \\ -Z^{(n)} \end{bmatrix}, 1
\]

Fifthly: Solving time response series

GM (1, 1) model,

\[
\hat{x}^{(1)}(k + 1) = \left[ x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a}
\]

Its corresponding time series is:

\[
\hat{x}^{(1)}(k + 1) = \left[ x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a}
\]

After progressive reduction, the prediction model of \( x^{(0)} \) original data series is obtained:

\[
x^{(0)}(k + 1) = \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k)
\]

2.2. Error Inspection

Residual: \( \varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) \)

Relative error: \( \varepsilon(k) = \frac{\varepsilon(k)}{x^{(0)}(k)} \times 100\% \)

When using GM (1, 1) model to predict, sample data must be tested first. The development coefficient \( m \) in the model affects the prediction accuracy. The greater \( -m \) is, the greater the prediction error is.

Therefore, GM (1, 1) model is suitable for prediction only when -\( m \) in a specific range.

(1) If \( -m \leq 0.3 \), GM (1, 1) model is suitable for medium and long-term prediction;
(2) If 0.3 < -m ≤ 0.5, GM (1, 1) model is suitable for short-term prediction and should be used with caution for medium and long-term prediction;
(3) If 0.5 < -m ≤ 0.8, GM (1, 1) model shall be used with caution for short-term prediction;
(4) If 0.8 < -m < 1, GM (1, 1) model needs to be corrected with residual;
(5) If -m > 1, GM (1, 1) model is not suitable for prediction.

GM (1, 1) model is suitable for the prediction of systems with uncertain factors. When predicting systems with less historical data and less fluctuation, it can get more ideal prediction values.

3. Container throughput forecast of some major ports in the Yangtze River Delta

3.1. Data Establishment
This paper selects the container throughput of each port from 2016 to 2020 as the original data, as shown in Table 1. The initial sequence is accumulated to generate the sequence next to the mean value, and then the prediction model is obtained by using the least square method. Finally, the prediction result is obtained by restoring the calculation result of the model.

| particular year | 2016  | 2017  | 2018  | 2019  | 2020  |
|-----------------|-------|-------|-------|-------|-------|
| Shanghai port   | 3713.3| 4023.3| 4201  | 4330.3| 4350  |
| Ningbo Zhoushan Port | 2156  | 2461  | 2635.1| 2559.7| 2872.2|
| Lianyungang Port | 500.92| 471.07| 474.56| 478.11| 480   |
| Suzhou port     | 469   | 590   | 636   | 627   | 629   |
| Taicang port    | 409   | 450   | 507   | 515.2 | 521   |

3.2. Data Operation
1. According to table 1, the port throughput sequence is established. The original sequence is accumulated by 1-mgo once, and the container throughput of each port is accumulated.

\[
X^0(k) = \begin{bmatrix} 3713.3 & 4023.3 & 4201 & 4330.3 & 4350 \\ 2156 & 2461 & 2635.1 & 2559.7 & 2872.2 \\ 500.92 & 471.07 & 474.56 & 478.11 & 480 \\ 469 & 590 & 636 & 627 & 629 \\ 409 & 450 & 507 & 515.2 & 521 \end{bmatrix}
\]

That is, \( X^{(0)} \) is generated by accumulation:

\[
X^{(1)} = X^{(1)}(1), X^{(1)}(2), X^{(1)}(3), \ldots, X^{(1)}(n)
\]

\[
X^{(1)} = \begin{bmatrix} 3713.3 & 7736.6 & 11937.6 & 16267.9 & 20617.9 \\ 2156 & 4617 & 7252.1 & 9811.8 & 12684 \\ 500.92 & 971.99 & 1446.55 & 1924.66 & 2404.66 \\ 469 & 1059 & 1695 & 2322 & 2951 \\ 409 & 859 & 1366 & 1881.2 & 2402.2 \end{bmatrix}
\]

2. Generate the nearest neighbor mean for \( X^{(1)}, Z^{(1)}(k) = \frac{1}{2} X^{(1)}(k) + \frac{1}{2} X^{(1)}(k-1), k = 2, 3, \ldots, n \), get:

\[
Z^{(1)} = Z^{(1)}(2) + Z^{(1)}(3) + Z^{(1)}(4) + Z^{(1)}(5)
\]

Namely:

\[
Z^{(1)} = \begin{bmatrix} 5724.85 & 9837.1 & 14102.75 & 18442.9 \\ 3386.5 & 5934.55 & 8531.95 & 11274.9 \\ 736.455 & 1209.27 & 1685.605 & 2164.66 \\ 764 & 1377 & 2008.5 & 2636.5 \\ 634 & 1112.5 & 1623.6 & 2141.7 \end{bmatrix}
\]
3.3. Calculation parameters $a, b$

\[
B = \begin{bmatrix}
-Z(2) & 1 \\
-Z(3) & 1 \\
-Z(4) & 1 \\
-Z(5) & 1 \\
\end{bmatrix}, \quad Y = \begin{bmatrix}
x(0)(2) \\
x(0)(3) \\
x(0)(4) \\
x(0)(5) \\
\end{bmatrix},
\]

and \( (a, b) = (B^T B)^{-1} B^T Y \)

Therefore, the coefficients $a$ and $B$ of each port are as follows:

1. Shanghai port: $(a, b) = \left[ \begin{array}{c} -0.0260518 \\ 3912.826949 \end{array} \right]$
2. Ningbo Zhoushan port: $(a, b) = \left[ \begin{array}{c} -0.044462 \\ 2308.523106 \end{array} \right]$
3. Lianyungang port: $(a, b) = \left[ \begin{array}{c} -0.00637048 \\ 466.704189 \end{array} \right]$
4. Suzhou port: $(a, b) = \left[ \begin{array}{c} -0.0171604 \\ 591.3874356 \end{array} \right]$
5. Taicang port: $(a, b) = \left[ \begin{array}{c} -0.043478 \\ 438.3894967 \end{array} \right]$

3.4. Get the prediction model

According to the above budget, the prediction equations of GM (1, 1) can be established as follows:

1. Shanghai port:
\[
\frac{dx(1)}{dt} - 0.0260518x(1) = 3912.826949
\]

2. Ningbo Zhoushan port:
\[
\frac{dx(1)}{dt} - 0.044462x(1) = 3912.826949
\]

3. Lianyungang port:
\[
\frac{dx(1)}{dt} - 0.00637048x(1) = 466.704189
\]

4. Suzhou port:
\[
\frac{dx(1)}{dt} - 0.0171604x(1) = 591.3874356
\]

5. Taicang port:
\[
\frac{dx(1)}{dt} - 0.043478x(1) = 438.3894967
\]

The time response formula is:
\[
x(1)(k + 1) = x(0)(1) - \frac{b}{a} e^{-ak} + \frac{b}{a}
\]

1. Shanghai port:
\[
x(1)(k + 1) = 153907 - 150194
\]

2. Ningbo Zhoushan port:
\[
x(1)(k + 1) = 54078e^{0.0444616k} - 51922
\]

3. Lianyungang port:
\[
x(1)(k + 1) = 73761.3e^{0.00637048k} - 73260.4
\]
(4) Suzhou port:

\[ x^{(1)}(k+1) = 3493.14e^{0.0171604k} - 3446.24 \]

(5) Taicang port:

\[ x^{(1)}(k+1) = 10492e^{0.043473k} - 10083 \]

3.5. Calculate the simulated value \( x^{(1)} \) and its reduced value \( \hat{x}^{(0)} \) of each port \( \chi^{(1)} \):

Table 2. Slove the simulated \( x^{(1)} \) and restored values \( \hat{x}^{(0)} \) for each port \( \chi^{(1)} \):

| port            | K=0 | K=1 | K=2 | K=3 | K=4 |
|-----------------|-----|-----|-----|-----|-----|
| Analog value    | 3713| 7775| 11945| 16224| 20617|
| Restore value   | 3713| 4062| 4170| 4279| 4393|
| Shanghai port   |     |     |     |     |     |
| Analog value    | 2156| 4615| 7185| 9872| 12682|
| Restore value   | 2156| 2459| 2570| 2687| 2810|
| Ningbo Zhoushan |     |     |     |     |     |
| Analog value    | 500.9| 972.3| 1446.7| 1924.1| 2404.6|
| Restore value   | 500.9| 471.4| 474.4| 477.4| 480.5|
| Lianyungang port|     |     |     |     |     |
| Analog value    | 469| 1073.6| 1688.7| 2314.4| 2951.0|
| Restore value   | 469| 583.6| 615.1| 625.7| 636.6|
| Suzhou port     |     |     |     |     |     |
| Analog value    | 409| 875.2| 1362.2| 1870.8| 2402.0|
| Restore value   | 409| 466.2| 487| 508.6| 531.2|
| Taicang port    |     |     |     |     |     |

3.6. Error inspection

According to the above calculation, based on GM (1, 1) model, taking the actual data of port cargo throughput of each port from 2016 to 2020 as the calculation data sample, the predicted simulation data of each port from 2021 to 2025 can be obtained. By comparing the actual data with the simulation data, Calculated as residual \( \epsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) \) and relative error: \( \epsilon(k) = \frac{|\epsilon(k)|}{\hat{x}^{(0)}(k)} \times 100\% \).

Table 3. GM (1, 1) model test error

| port            | Serial number | Actual data | Analog data | residual | relative error (%) |
|-----------------|---------------|-------------|-------------|----------|--------------------|
| Shanghai port   | 1             | 4023.3      | 4062.25     | -38.95999| 0.968359           |
|                 | 2             | 4201        | 4169.45     | 31.55    | 0.751012           |
|                 | 3             | 4330.3      | 4279.5      | 51.8     | 1.196222           |
|                 | 4             | 4350        | 4392.5      | -42.5    | 0.977011           |
| Ningbo Zhoushan | 1             | 2461        | 2458.64     | 2.36     | 0.068188           |
|                 | 2             | 2635.1      | 2570.43     | 64.67    | 0.245418           |
|                 | 3             | 2559.7      | 2687.29     | -127.59  | 4.984565           |
|                 | 4             | 2872.7      | 2809.44     | 62.76    | 2.184704           |
| Lianyungang port| 1             | 471.07      | 471.395     | -0.325   | 0.069001           |
|                 | 2             | 474.56      | 474.405     | 0.1550   | 0.032668           |
|                 | 3             | 478.11      | 477.44      | 0.67     | 0.140135           |
| Suzhou port     | 1             | 590         | 604.61      | -14.61   | 2.476271           |
|                 | 2             | 636         | 615.07      | 20.93    | 3.290881           |
|                 | 3             | 627         | 625.72      | 1.28     | 0.204147           |
|                 | 4             | 629         | 636.55      | -7.55    | 1.200318           |
| Taicang port    | 1             | 450         | 466.234     | -16.2    | 3.6                |
|                 | 2             | 507         | 486.956     | 20       | 3.944773           |
|                 | 3             | 515.2       | 508.59      | 6.6      | 1.281056           |
|                 | 4             | 521         | 531.19      | -10.2    | 1.957774           |
It can be seen from the accuracy test results that the relative errors between the actual value and the simulated value of GM (1, 1) model are basically between 0 ~ 3.6%, the development coefficient M = -0.0256 ≤ 0.3, and the accuracy test results are relatively ideal. Therefore, the sample data can use GM (1, 1) model to predict the container throughput of other ports in the Yangtze River Delta in the medium and long term.

3.7. Prediction Results
According to the above analysis, the port container throughput of each port from 2021 to 2025 can be predicted based on GM (1, 1) prediction model: the prediction results are shown in Table 3.

Table 4. Forecast of port container throughput of each port from 2021 to 2025

| Port       | K=6  | K=7  | K=8  | K=9  | K=10 |
|------------|------|------|------|------|------|
| Shanghai port | 29752 | 34502 | 39377 | 44380 | 49516 |
| Estimate   | 4627 | 4750 | 4875 | 5003 | 5136 |
| Ningbo Zhoushan port | 18690 | 21900 | 25256 | 28765 | 32434 |
| Estimate   | 3071 | 3210 | 3356 | 3509 | 3669 |
| Lianyungang port | 3374.8 | 3864.6 | 4357.5 | 4853.5 | 5352.8 |
| Estimate   | 486.6 | 489.8 | 492.9 | 496 | 499.3 |
| Suzhou port | 4257.3 | 4927.5 | 5609.3 | 6302.8 | 7008.4 |
| Estimate   | 658.8 | 670.2 | 681.8 | 693.5 | 705.6 |
| Taicang port | 3536.2 | 4141.4 | 4773.5 | 5433.7 | 6123.2 |
| Estimate   | 579.4 | 605.2 | 632.1 | 660.2 | 689.5 |

According to the data in Table 3, from 2021 to 2025, the container throughput of Shanghai port will increase from 46.21 million TCU in 2021 to 51.19 million TCU in 2025, and will increase by 4.98 million TCU in five years. On the whole, the increasing trend of container throughput of Shanghai port is relatively mild, but at present, the use of site and related equipment for Shanghai port terminals has tended to be saturated, Compared with the slow growth of container throughput, scientific planning and accelerating the development of port logistics are also essential and important measures.

4. Conclusion
The integrated development of the Yangtze River Delta has become a national strategy, which has brought new development opportunities to the ports in the Yangtze River Delta. Social and economic development, transportation first, steady progress in port integration in the Yangtze River Delta, high-grade inland waterways connected to high standards, and interconnection, integration and coordination of ports, waterways and other infrastructure.

Container throughput is also increasing day by day. Therefore, we will continue to realize the optimal allocation of resources, promote the integrated development of iron water intermodal transport, river sea intermodal transport and public water intermodal transport, and improve logistics efficiency. The planning and arrangement of throughput of regional ports need to be further transformed and upgraded to improve service support capacity, quality and efficiency.

References
[1] Zhao Yahong, Niu Qintao, Zhou Wenguo. Matlab implementation of grey neural network prediction modeling [J]. Shanxi architecture, 2017, 43 (21): 255-257.
[2] Du baisong. Container throughput prediction of Shenzhen Port Based on Grey Markov model [J]. Journal of Zhejiang Ocean University (NATURAL SCIENCE EDITION), 2019 (3): 180-186.
[3] Zhu Jian. The study found that the development of the world economy has the greatest impact on the cargo throughput of China's coastal ports.
[4] Huang Yuehua, Chen Haishan. Application of optimized GM (1, 1) power model in Port
Throughput Prediction [J]. China navigation, 2016, 39 (2): 131-134.

[5] Zhang Shukui, Lu ziai. Research on grey neural network prediction model of port container throughput [J]. Journal of Jiangsu University of science and Technology (NATURAL SCIENCE EDITION), 2014, 28 (3): 216-219224. Doi: 10.3969/j.issn.1673-4807.2014.03.003.

[6] Li guangru, Zhang Xin, Zhu Qinghui. Improved Elman neural network port throughput prediction method based on AdaBoost [J]. Journal of Chongqing Jiaotong University (NATURAL SCIENCE EDITION), 2021, 40 (05): 1-5.