Prediction of Ship Arrival Quantity Based on Optimized GM (1, 1) Power Model

Mingxia Zheng¹, mengying Fan ² and hu Liu ³
¹ Maritime School of Zhejiang Ocean University, Zhoushan, Zhejiang, 316022, China
*Corresponding author’s e-mail: liu-hu-175@163. com

Abstract. Since the ship arrivals is vulnerable to climate, market season, economic development and other factors and is characterized for volatility, the prediction for ship arrivals by means of the conventional GM (1, 1) power model has such problems as low precision, non-relevance between the modeling method and the model testing. In this paper, the optimal matching model parameters for the ship arrival data are found and the GM (1, 1) power model with optimized model parameters is established using the nonlinear programming method, based on the conventional GM (1, 1) power model. Comparisons in term of accuracy in the application of the predictions are made between the ship arrivals data in Cengang port area of Ningbo - Zhoushan port in May for the last 10 years and the predictions. The predicted ship arrivals for May, 2018 is 860, with an average relative error of 1.67%, and the average relative error of the optimized power model is greatly reduced, which shows that the average relative error of the optimized GM (1, 1) power model has been significantly reduced, with a relative residual not exceeding 7%. The results show that the optimized GM (1, 1) power model is able to further improve the prediction accuracy and meet the actual requirements, providing theoretical support for planning of the port anchorage.

1. Introduction
As an important node of waterway transportation, the rational planning of ports and the full utilization of resources are crucial to the development of regional economy. According to An overall plan for Ningbo -- Zhoushan port for the period 2014-2030, in 2016 Ningbo - Zhoushan port cargo throughput of the first breakthrough 900 million tons port, cargo throughput of prediction in 2020 and 2030 will reach 1.17 billion tons and 1.44 billion tons, what’s more, its annual average growth rate of 5.0% and 2.1% [1]. It can be seen that the port throughput presents an obvious growth trend, the ship to port will grow, however, port to ship to the anchorage of the contradiction between the demand and the supply of the port of anchorage capacity gradually revealed. Lack of anchorage resources, low utilization rate and lagging development planning not only affect the security of ports, but also restrict the development of ports. Therefore, the prediction of ship arrival based on the prediction theory is of great significance to the rational planning of port anchorage and the safety guarantee of waterway traffic.

2. Selection of prediction model
Ship arrival at port is a typical random arrival event. The prediction of random arrival events has been studied by a large number of scholars. Among them, exponential smoothing method and elastic coefficient method are widely used in port throughput prediction and port traffic flow. Liu yiqun and Zeng Ming studied these common prediction methods in the study of port container handling and
handling prediction methods and drew the following conclusions [7]: Compared with the grey prediction method and the multiple linear regression analysis and prediction method, the exponential smoothing method and the elastic coefficient method are mostly based on experience and have certain errors in the determination of parameter values or variable selection, which also leads to the large discrepancy between the predicted results and the actual values.

The following grey models are commonly used at present: GM (1, 1) model, Grey Verhulst model, GM (1, N) model. Both GM (1, 1) model and GM (1, N) model are linear models in form, while port to port quantity is a complex process affected by multiple factors with nonlinear characteristics. GM (1, 1) power model is a kind of nonlinear gray model, whose background value and gray derivative components can satisfy the translational relation. Depending on different values of $\alpha$, this model is suitable for modeling of different original sequence. It is obvious that such parameter selection has some drawbacks. Literature [6] determines parameter according to the information coverage principle of the gray system. The example shows that the traditional GM (1, 1) power model obtained by this method is superior to the gray Verhulst model in terms of application scope and prediction accuracy. By the known, in view of the random events port vessel arrival has the characteristics of increasing trend with abnormal fluctuations data, based on traditional GM (1, 1) power model, using the nonlinear programming method to find the best fit to the port of measured data of model parameters, to optimize the model parameters of GM (1, 1) power model, forecast ship arrival quantity, the more close to the actual forecast results are achieved.

3. Establishment and solution of GM (1, 1) power model

3.1. Establishment of model

The ship's original arrival data is denoted as $X^{(0)}$, and $X^{(1)}$ as the original ship to the port of first-order accumulated data. The immediately adjacent mean value series which is generated based on the sequence $X^{(1)}$ is $Z^{(1)}$.

The calculation formula of $x^{(1)}(k)$ and $z^{(1)}(k)$ is:
\[
X^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(k), \quad k = 1, 2, ..., n
\]
\[
Z^{(1)}(k) = \frac{x^{(1)}(k) + x^{(1)}(k-1)}{2}
\]

The grey differential equation is:
\[
x^{(0)}(k) + \alpha * z^{(1)}(k) = b * Z^{(1)}(k)^{\alpha}, \quad \alpha \neq 1
\]

Thus, the whitening differential equation of the GM (1, 1) power model can be obtained as:
\[
\frac{dx^{(1)}}{dt} + ax^{(1)} = bx^{(1)}^{\alpha}
\]

3.2. Solution of model parameters

3.2.1. A general estimation method for parameter $\alpha$. In equation (3), $a$, $b$ and $\alpha$ are unknown, and $t$ is time. When $t = k$, $t = k + 1$, the expression of parameter $\alpha$ is obtained according to the idea of grey system information coverage:
\[
\alpha = Q \times W^{-1}
\]
\[
Q = \begin{bmatrix}
\left[x^{(0)}(k + 1) - x^{(0)}(k)\right] * z^{(1)}(k + 1) * z^{(1)}(k) * x^{(0)}(k) - \\
\left[x^{(0)}(k) - x^{(0)}(k - 1) * z^{(1)}(k + 1) * z^{(1)}(k) * x^{(0)}(k + 1)\right]
\end{bmatrix}
\]
\[
W = \begin{bmatrix}
\left[x^{(0)}(k + 1)\right]^2 * z^{(1)}(k) * x^{(0)}(k) - \left[x^{(0)}(k)\right]^2 * z^{(1)}(k + 1) * x^{(0)}(k + 1)
\end{bmatrix}
\]

According to equation (4), there are $n - 2$ values of $\alpha$ that can be obtained for $k = 2, 3, ..., n - 1$, which denoted as $\alpha_k$.

At the same time set $g(\alpha) = \sum_{k=2}^{n-1} (\alpha - \alpha_k)^2$
The value of $\alpha$ that minimizes $g(\alpha)$ is the undetermined constant where
\[\alpha = \frac{1}{n-2} \sum_{k=2}^{n-1} a_k\] (6)

3.2.2. The estimation method of parameters $a$ and $b$. After $\alpha$ is calculated, the following equation can be obtained by performing the least squares estimation for the parameters sequence $(a, b)^T$ according to the whitening differential equation:
\[(a, b)^T = (B^TB)^{-1}B^TY\] (7)

Among them:
\[B = \begin{bmatrix} -z^{(1)}(2) & ( - z^{(1)}(2) )^a \\ \vdots & \vdots \\ -z^{(1)}(n) & ( - z^{(1)}(n) )^a \end{bmatrix}, \quad Y = \begin{bmatrix} x^{(0)}(2) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}\]

The time response equation can be obtained by solving the above albino differential equation:
\[\hat{x}^{(1)}(k + 1) = \left[ \frac{b}{a} + \frac{[x^{(0)}(1)]^{1-a} - \frac{b}{a} e^{-ak(1-a)}}{1-a} \right]^{1-a}\] (8)

In the above equation, $k = 2, 3, \ldots, n - 1$; By performing a first-order reduction calculation on equation (8), the predicted value of the gray power model of the original ship arrival sequence $\hat{x}^{(0)}$ is obtained.
\[\hat{x}^{(0)}(k + 1) = \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k)\] (9)

3.3. Calculation of optimization parameters

According to formula (6), if $\alpha$ is determined, then parameters $a$ and $b$ in the formula model will be determined accordingly. The objective is to minimize the average relative error of the predicted arrival quantity, and the optimization model is established with the relation between the power index and parameters $a$ and $b$ as the constraint condition, and the value of the optimal power index $\alpha$ is obtained by the exhaustive method, so that the absolute value of the average relative error of the prediction model of arrival quantity can be minimized in theory.
\[\min_{\alpha} \text{ARPE} = \frac{1}{n-1} \sum_{i=2}^{n} \left| \frac{\hat{x}^{(0)}(k) - \hat{x}^{(0)}(k)}{\hat{x}^{(0)}(k)} \right|\] (10)

3.4. Model accuracy evaluation

In order to ensure the accuracy of the predicted values, the prediction results of the grey model must be tested for accuracy. There are three common testing methods: the conformity model, small error probability conformity model for the residual conformity inspection and the mean square deviation ratio. And the commonly used accuracy levels are shown in Table 1.

| Level | TRPE value | C value | P value |
|-------|------------|---------|---------|
| 1 Level | <0.01 | <0.35 | >0.95 |
| 2 Level | <0.05 | <0.5 | >0.8 |
| 3 Level | <0.1 | <0.65 | >0.75 |
| 4 Level | >=0.2 | >-0.8 | <=0.6 |

In this paper, the residual inspection test is employed, and the error ratio of 5% is set as the defining standard, that is, when the relative error of the residual is less than 5%, it is considered as a model with satisfactory residual, and when the relative error ratio is more than 10%, it is considered that the residual is not satisfactory. Of which, the residual is the difference between the actual ship arrivals and the predicted ship arrivals; the relative error of the ship arrivals at time $k$ is denoted as $1$, and $\hat{x}$ is the relative accuracy of the simulated ship arrivals at time $k$; the average relative error of the ship arrivals at all times is denoted as $\text{TRPE}$, and the expression is as follows:
\[\text{TRPE}(k) = \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{\hat{x}^{(0)}(k)} \right| \times 100\%\] (11)
\[ \text{TARPE} = \frac{1}{n-1} \sum_{k=2}^{n} TRPE(k) \] (12)

4. Application examples and comparative analysis

4.1. Original data analysis

In this paper, the simulation and prediction are made taking the data of ship arrivals in Canggang port area, Zhoushan port under the jurisdiction of Zhejiang province as the instance. Because there exists the low season and high season for ship arrivals, one particular month is selected for the monthly prediction, and the sum of the monthly prediction results is the final result of the annual prediction. In this paper a separate prediction is made for the data of May.

| Year | Arrivals of ships |
|------|-------------------|
| 2009 | 325               |
| 2010 | 380               |
| 2011 | 398               |
| 2012 | 432               |
| 2013 | 454               |
| 2014 | 459               |
| 2015 | 542               |
| 2016 | 548               |
| 2017 | 921               |

4.2. Calculation of traditional model

The model is established with the data of May between the years 2009-2014, and a comparison in term of accuracy is made against the data of May between the years 2015 - 2017. An initial predicted value \( \hat{x}^{(1)}(1) = x^{(1)}(1) = 2 \) is taken. According to equations (1) and (2), the data in Table 3 are obtained through process the raw data.

| Year | Original data | First-order summation | The immediately adjacent mean value series |
|------|---------------|-----------------------|------------------------------------------|
| 2009 | 325           | 325                   |                                          |
| 2010 | 380           | 705                   | 515                                      |
| 2011 | 398           | 1103                  | 904                                      |
| 2012 | 432           | 1535                  | 1319                                     |
| 2013 | 454           | 1989                  | 1762                                     |
| 2014 | 459           | 2448                  | 2218.5                                   |

According to the idea of gray system information coverage, it is concluded for formula (4) that, when \( k = 2, 3, \ldots, n - 1 \), the value of \( n - 2 \)'s \( \alpha \), which is denoted as \( \alpha_k \), and the specific values are as shown in Table 4.

| \( K \) | \( \alpha_k \) |
|-------|--------------|
| 1     | -0.29693273  |
| 2     | 0.334378974  |
| 3     | -0.36649823  |
| 4     | 0.432950803  |
| 5     | 0.739973571  |

From formula (5) and (6), it is determined that a, which has minimized \( g(\alpha) \), is the undetermined constant value, and \( \alpha = \frac{1}{n-2} \sum_{k=2}^{n-1} \alpha_k \) at this time. Finally, the parameter \( \alpha = 0.30259 \) in the conventional power model is obtained. According to equation (7), the model parameters \( a = 0.03905196, b = 55.999892 \) are obtained, which are introduced into the equations 9 and 10 to obtain the time response equation:

\[ \hat{x}^{(1)}(k + 1) = (1433.984321 - 1428.228983e^{-0.211029292k})^{1.433876773} \]

Then a further first-order reduction calculation is performed and, the simulated value of the grey GM (1, 1) power model of the original ship arrivals sequence \( \hat{x}^{(0)} \) as well as the relative error of the residual are computed.
4.3. Calculation of optimization model

The calculation of the optimization model has been performed using the exhaustive method, on the basis of calculations with the conventional model. According to equations (4) (5) (6) (7) (10), the model parameters $a = -0.082373027, b = 521.1449327, \alpha = -0.07$ are obtained, which are introduced into the equations 9 and 10 to obtain the optimization time response equation:

$$x^{(1)}(k + 1) = \{-6326.64537 + 6327.312435e^{0.0749k}\}_{0.934579439}$$

Performing a first-order reduction calculation on equation (8), the predicted value of the gray power model of the original ship arrival sequence $\hat{x}^{(0)}$ is obtained, and the specific values are computed.

4.4. Comparison of model accuracy

Through the above calculations, the simulation and prediction results with the two models are obtained. Comparisons in terms of accuracy are made between the simulation and prediction results of the conventional model and the optimized model, and the data in Table 5 are obtained.

| Year | The actual value | The traditional model | The optimization model |
|------|------------------|-----------------------|------------------------|
|      |                  | Simulation value       | TRPE%                  | Simulation value | TRPE%                  |
| 2009 | 325              | 325.00                 | 0.00                   | 325.00           | 0.00                   |
| 2010 | 380              | 344.53                 | 9.34                   | 379.59           | 0.11                   |
| 2011 | 398              | 398.83                 | 0.21                   | 398.15           | 0.04                   |
| 2012 | 432              | 437.83                 | 1.35                   | 423.47           | 1.97                   |
| 2013 | 454              | 467.06                 | 2.88                   | 453.22           | 0.17                   |
| 2014 | 459              | 489.36                 | 6.61                   | 486.79           | 6.06                   |
|      | TARPE%           | 4.08                   | 1.67                   |
| 2015 | 542              | 506.42                 | 6.56                   | 524.06           | 3.31                   |
| 2016 | 548              | 519.38                 | 5.22                   | 565.09           | 3.12                   |
| 2017 | 921              | 529.01                 | 42.56                  | 610.06           | 33.76                  |

As can be seen from Table, for the conventional model, TRAPE = 4.08%, for the optimized model, TRAPE = 1.67%, both are less than 5%, meeting the accuracy requirements for the mode and are satisfactory models, with the optimized model having a relative residual of not higher than 7% and consequently having a higher accuracy. In addition, the precision of the predictions for the years 2015-2017 is higher than that of the conventional model. Therefore, the optimization model is superior to the traditional model in model accuracy.

5. Conclusion

The gray GM (1, 1) power model is a nonlinear gray model, which is suitable for the prediction of a sequence with small sample oscillations. In this paper, the optimal matching model parameters for the ship arrival data are found and the GM (1, 1) power model with optimized model parameters is established using the nonlinear programming method. Comparisons in term of accuracy in the application of the predictions are made between the ship arrivals data in Cengang port area of Ningbo - Zhoushan port in May for the last 10 years and the predictions. The predicted ship arrivals for May, 2018 is 860, with an average relative error of 1.67%, and the average relative error of the optimized power model is greatly reduced, which shows that the average relative error of the optimized GM (1, 1) power model has been significantly reduced, with a relative residual not exceeding 7%. The results show that the optimized GM (1, 1) power model is able to further improve the prediction accuracy and meet the actual requirements, providing theoretical support for planning of the port anchorage.

Project support:
This paper was supported by Zhejiang University Student Science and Technology Innovation Program and New Miao Talent Program, No. 2018R411015.
References:

[1] Shi, Q. F. (2017) Marine legislation and planning: master plan of Zhoushan port of Ningbo (2014-2030) was approved. China ocean yearbook. Ocean press, 167-168.

[2] Huang, Y. H., Chen H. S. (2016) Optimized GM (1, 1) Power Model for Predicting Port throughput [J]. China maritime., 39(02): 131-134.

[3] Lu, H. X., Zhang, D. S. (2017) Modelling of GM (1, 1) Power Model Based on Oscillating Sequence [J]. Henan science, 35(03): 360-364.

[4] Hu, P. (2017) To Optimize the Background Value of GM (1, 1) Power Model and Its Application [J]. Mathematical practice and cognition, 47(19):99-104.

[5] Liu, Y. L., Li, F., Cao W. J. (2017) Application of Modified GM (1, 1) Power Model in Prediction of Discharge Population [J]. China health statistics, 34(03): 483-484+487.

[6] Wang, Z. X., Dang, Y. G., Liu, S. F., Lian, Z. W. (2009) Solution of GM (1, 1) Power Model and Its Properties [J]. Systems engineering and electronic technology, 31(10): 2380-2383.

[7] Liu, Y. Q., Zeng, M. (2008) The Research on Forecasting Method of Port Container Throughput [J]. Logistics engineering and management, 40(08): 85-87.

[8] Lu, P. W., Zhao, C. L., Li, F. Y. (2017) Research on Automobile Traffic Accident Prediction Based on Residual Modified GM (1, 1) Model [J]. Internal combustion engine and accessories, (15): 89-91.