Application of grey system theory in prediction of accident prone sea area

Wang xi-kun¹, Wang yi-xin¹ and Xu you-hao¹
¹ Shandong Jiaotong University, Weihai, Shandong, 264200, China
*Corresponding author’s e-mail: 2564546101@QQ.com

Abstract. The GM (1,1) model is established by using the grey system theory, Realizes the goal that the accidents in the multi accident sea area can be accurately predicted by the prediction model. The key point is to establish the model correctly and apply it to the prediction of accident prone sea area. Looking up the sea area with the characteristic of frequent accidents, building the model, and testing the result of substituting the data into the model prediction, the result shows that the viewpoint of this paper is correct and can predict the maritime traffic accidents with high accuracy.

1. Introduction
GM (1,1) model of grey system is based on the comprehensive data of many known factors in the system, and the time series of this data is fitted by differential equation to force the dynamic process described by "near time series", and then extrapolate to achieve the purpose of prediction.

2. Grey system theory and GM (1,1) model

2.1. On grey system theory and GM (1,1) model
The original sequence is recorded as:

\[ x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(t)\} \quad (2.1) \]

Due to the randomness of the occurrence of ship accidents, there is no specific law. To weaken the randomness of ship accidents, we need to adopt the method of accumulation to deal with the time data and get the accumulated generated number as follows:

\[ x^{(1)}(t) = \sum_{i=1}^{n} x^{(0)}(i), \ldots, \quad n \quad (2.2) \]

Thus, we can get the generating terms without less randomness and volatility. The original sequence of the cumulative generating number is recorded as follows:

\[ x^{(1)}(t) = \{x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(k)\}, \quad k = 1, 2, \ldots, \quad n \quad (2.3) \]

The number of progressive generation is:

\[ x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1), \quad x^{(1)}(0) = 0. \quad (2.4) \]

Then, according to a series of time data obtained above, \( x^{(1)} \) change shows exponential change, and the new data sequence is assumed to satisfy the first-order univariate ordinary differential equation:

\[ \frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b \quad ( a \text{ and } b \text{ are constants not equal to 0}) \quad (2.5) \]
The prediction model is established by data accumulation:

\[ x^{(0)}(k+1) = [x^{(0)}(1) - \frac{b}{a}] e^{-ak} + \frac{b}{a} \quad (k=0, 1, 2, \ldots, n) \]  

The final grey prediction model is as follows:

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

2.2. Test the model

First, the model is tested by residual and relative residual methods:

\[ E(t) = x^{(0)}(t) - \tilde{x}^{(0)}(t) \]  

\[ e(t) = \frac{E(t)}{x^{(0)}(t)} = \frac{x^{(0)}(t) - \tilde{x}^{(0)}(t)}{x^{(0)}(t)} \]

In two formulas, \( x^{(0)}(t) \) is the number of ship traffic accidents in the original year \( t \), \( \tilde{x}^{(0)}(t) \) is predict the number of ship traffic accidents that will occur in the \( T \) year, \( e(t) \) the smaller the model is, the higher the prediction accuracy is; otherwise, the opposite is true.

In addition to the above methods, the posteriori difference test is also a conventional test method.

Posterior error test: set \( S_1 \) as the mean square deviation of the original ship traffic accident sequence and \( S_2 \) as the mean square deviation of the above residual sequence, and calculate them separately.

\[ S_1^2 = \frac{1}{n} \sum_{k=1}^{n} [x^{(0)}(k) - \tilde{x}^{(0)}] \]  

\[ S_2^2 = \frac{1}{n} \sum_{k=1}^{n} [E(k) - E]^2 \]  

\[ C = S_2/S_1 \] and \( P = P[E(t) - E|<0.6745S_1] \) the ratio of posterior error and the frequency of small error, respectively, is an index to measure the accuracy of the prediction model. The smaller the \( C \) is, the higher the accuracy is, the larger the \( P \) is, the higher the accuracy is, and vice versa. The combination of these two indicators can comprehensively determine the accuracy of the prediction model. Generally speaking, the accuracy of the model can be divided into four grades (see Table 1).

| Accuracy grade | \( e \) | \( k_0 \) | \( C_0 \) | \( p_0 \) |
|----------------|-------|-------|-------|-------|
| LEVEL 1        | 0.01  | 0.90  | 0.35  | 0.95  |
| LEVEL 2        | 0.05  | 0.80  | 0.50  | 0.80  |
| LEVEL 3        | 0.10  | 0.70  | 0.65  | 0.70  |
| LEVEL 4        | 0.20  | 0.60  | 0.80  | 0.60  |
3. The practice of ship traffic accidents in multi accident sea area in grey system theory

3.1. Recent accidents in Liaoning and Qingdao sea areas

Table 2 and table 3 respectively show the number of maritime vessel traffic accidents in Liaoning sea area and Qingdao sea area, which are frequent accidents in China from 2012 to 2018.

Based on the investigation of ship traffic accidents in recent years, the data in the following table are obtained:

| Table 2 Traffic accident volume of ships in Liaoning sea area in 2012-2018 |
|---------------------|---|---|---|---|---|---|---|
| Particular year     | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
| Traffic accident volume | 165 | 159 | 154 | 149 | 148 | 150 | 148 |

| Table 3 Traffic accident volume of ships in Qingdao sea area of China in 2012-2018 |
|---------------------|---|---|---|---|---|---|---|
| Particular year     | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
| Traffic accident volume | 154 | 149 | 142 | 137 | 135 | 138 | 136 |

The data of ship traffic accidents in these two accident prone areas are superposed and integrated into a new table 4.

| Table 4 Total traffic accidents in Liaoning sea area and Qingdao sea area |
|---------------------|---|---|---|---|---|---|---|
| Particular year     | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
| Total traffic accidents | 319 | 308 | 296 | 286 | 283 | 288 | 284 |

Through table 4, we can directly read the total number of ship traffic accidents in the multi accident sea area from 2012 to 2018.

3.2. Modeling

First, we directly use the original data of ship traffic accidents to build the model, accumulate the original data of ship traffic accidents from 2012 to 2018, and get the original sequence:

$[x(n)/432,397,358,326,298,312,281]$  \( (3.1) \)

Finally, the GM (1,1) prediction model is obtained as follows:

$x^{(1)}(t+1) = -4127.215139e^{-0.109362k} + 5174.216951$  \( (3.2) \)

Test the model: $\bar{e} = 0.12, e = 0.11$, test this model: according to the comparison of accuracy level table, the accuracy level of this prediction model is level 3, and the level is low. As a prediction model, it is obviously unqualified.

The second-order weakening of the original data makes the data present a certain regularity to establish a prediction model:

$XD^2 = XDD = [x(1)d^2, x(2)d^2, \ldots, x(n)d^2]$  \( (3.3) \)

$x(k)d^2 = \frac{1}{n-k+1}[x(k)d + x(k+1)d + \ldots + x(n)d], k = 1, 2, \ldots, n$  \( (3.4) \)

Substitute the original traffic data of the ship into:

$x^{(0)} = [x^{(0)}(t)/432,397,358,326,298,312,281]$  \( (3.5) \)

After the second order weakening, we get a set of different sequences:
\[ X^{(0)}D^2, X^{(1)}D^2 = (295, 268, 254, 228, 216, 231) \quad (3.6) \]

New sequence \( X^{(i)} \) by accumulation
\[ X^{(i)} = [X^{(i)}(1), X^{(i)}(2), X^{(i)}(3), X^{(i)}(4), X^{(i)}(5), X^{(i)}(6)] = (295, 268, 254, 228, 216, 231) \quad (3.7) \]

Finally, the whitening equation is obtained:
\[ \frac{dX^{(i)}}{dt} + aX^{(i)} = u \quad , \quad (3.8) \]

Two parameters \( a \) and \( u \) are obtained:
\[
Y_n = \begin{bmatrix}
268 \\
254 \\
228 \\
216 \\
231 \\
\end{bmatrix}, \quad B = \begin{bmatrix}
-477 & 1 \\
-784 & 1 \\
-1287 & 1 \\
-1465 & 1 \\
-1794 & 1 \\
\end{bmatrix} \quad (3.9)
\]

\( a = -0.013376, \quad u = 316.243201 \)

Get the prediction model:
\[ x^{(i)}(t+1) = -21324.567884e^{-0.032543k} + 21365.567884 \quad (3.10) \]
\[ X^{(0)} = [X^{(0)}(1), X^{(0)}(2), X^{(0)}(3), \ldots, X^{(0)}(6)] = (295, 268, 254, 228, 216, 231) \quad (3.11) \]

3.3. Establish absolute correlation
Calculation \( X^{(0)} \) and \( X^{(i)} \) absolute correlation, That is to establish the absolute correlation between the original data and the data processed by the second-order weakening operator.

Set the original data as \( X^{(0)} \), The data processed by the second order weakening operator is \( X^{(i)} \), Establish the absolute correlation between the two data, \( X^{(0)} \) and \( X^{(i)} \) the absolute correlation coefficient between them is \( K \), if \( k_0 > 0 \), then \( k > k_0 \) the model is very reasonable and effective.

If \( k = 0.914 \geq 0.9 \) is calculated by the calculation method of absolute correlation degree, then the model can be checked to be qualified. After the model is determined to be qualified, the formula can be obtained
\[ x^{(i)}(t+1) = -23303.763654e^{-0.012654k} + 23773.763654 \quad (3.24) \]

This formula can be used as the prediction model of grey system. Through this model, we can calculate the prediction of ship traffic accidents in Qingdao sea area and Liaoning sea area in 2018, 2019 and 2012 respectively (1), and they are 265, 245 and 231 respectively.

3.4. Model prediction test
The second-order weakening operator is used to process the original data of ship traffic accidents in 2012-2018, and then the prediction model is used to obtain the accident prediction data of 2019, 2020 and 2021 respectively, with the prediction value data. See table 5 for the prediction value data.
Table 5. Prediction results of GM (1,1) model.

| Particular year | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 |
|-----------------|------|------|------|------|------|------|------|------|------|------|
| K               | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    |      |
| $X^{(1)}$       | 245  | 561  | 1042 | 1212 | 1463 | 1644 | 1832 |      |      |      |
| $X^{(2)}$       | 241  | 233  | 225  | 217  | 220  | 218  |      |      |      |      |
| Original value  | 319  | 308  | 296  | 286  | 283  | 288  | 284  |      |      |      |
| Predicted value (1) | 296 | 291  | 286  |      |      |      |      |      |      |      |
| Predicted value (2) | 278 | 265  | 254  |      |      |      |      |      |      |      |

4. Conclusion

By combining the data of Liaoning sea area and Qingdao sea area with the prediction analysis calculated by the grey system model, it is a very simple and effective method to build GM (1,1) model to predict maritime traffic accidents. However, only two sea area traffic accident data are selected for superposition, the number of samples is small, and this model generally has high prediction accuracy under short-term and constant conditions. Therefore, in order to predict the sea area with frequent accidents accurately and stably in a long term, this model needs to be further optimized and improved.

References

[1] Deng J L. Control problems of grey systems[J]. Systems & Control Letters, 1982, 1(5): 288-294.
[2] Tao Jun, Li Yaodong. Research on ship traffic supervision under poor visibility in Zhanjiang port [J]. Zhujian water transport, 2017, 27 (15): 72-73.
[3] Tu Haiyang, Yuan Zhangxin. Research on simulation system of ship traffic accident recurrence [J]. Shanghai Institute of ship transportation science, 2001.
[4] DEHDARA F, YAPB S F, NAGHAVIC M S, et al. Charting the future global status of oil and natural gas using grey forecasting[J]. Institutions and Economies, 2016, 8(3): 105-125.
[5] Kong Xue, Wang Li, Feng Yihua. Application status and Prospect of Grey Prediction GM (1,1) model [J]. Journal of Qilu University of technology, 2018, 32 (6): 49-53.
[6] XU Q Y, YANG M L, LIU W D, et al. Product form feature evolution forecasting based on IGMBPM model [J]. Computer-Aided Design and Applications, 2016, 13(4): 431-439.
[7] Deng Julong. Basic method of grey system [M]. Wuhan: Huazhong University of Technology Press, 1987.
[8] Liu Sifeng, Xie Naiming. Grey system theory and its application [M]. Beijing: Science Press, 2018.