Forecasting of Maritime Traffic Accident based on Residual Error Unbiased Grey Forecast Model

Sun Shuang1,*, Chen Yan1, Zhang Jinsong1 and Li Taoying1

1School of Maritime Economics and Management, Dalian Maritime University, Dalian, China

*470582807@qq.com

Abstract. In order to enhance the safety of maritime transportation and improve the accuracy of maritime traffic accident prediction, an unbiased grey forecast model based on residual error is applied to predict maritime traffic accident. Based on the historical data of maritime traffic accidents from 2008 to 2017, the traditional unbiased grey model prediction and residual error unbiased grey model prediction are carried out, and the fitting curves of actual value and predicted value of the two models are drawn. The results show that the prediction accuracy and fitting curve of residual error unbiased grey model are better than those of traditional unbiased grey model, which can truly reflect the development trend of comprehensive safety of maritime traffic, and the prediction results have certain reliability and practicability.

1. Introduction

Grey theory was first put forward by Deng Julong in 1982.[1] It is a new method to study the problems of minority data, poor information and uncertainty. This theory takes the uncertain system of "small sample and poor information" with "part information known and part information unknown" as the research object. Through the generation and development of some known information, it extracts valuable information, and realizes the correct description and effective monitoring of the system's operation behaviour and evolution law. In grey forecasting model, GM (1,1) model of time series forecasting is the most commonly used model. It has been widely used in industry and data mining forecasting. [2] However, when using GM (1,1) model to model and predict, it sometimes has big deviation, and has some defects, such as low learning ability to data and low prediction accuracy.[3]

The basic idea of maritime traffic accident prediction is to reasonably infer and analyse the possible maritime traffic accidents based on the statistics, analysis and processing of the accident data and the causes and rules of the accidents. [4] Common forecasting methods include regression analysis method, time series forecasting method, grey forecasting method, BP neural network forecasting method, etc. [5] Practice shows that each prediction method has its own characteristics and can basically meet the expected requirements, but at the same time there are corresponding defects. [6] For example, the grey model uses the accumulated new data to model, ignoring the randomness of the predicted data. [7] Although BP neural network has high prediction accuracy, it needs very complete data, otherwise the network extension trained will be very poor.[8]
2. Model establishment

2.1. grey GM(1,1) model

1) Selecting n-year maritime traffic accident volume as the initial time series $X^{(0)}$ of grey model, a set of original time series samples of water traffic accident are obtained,

$$ X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)) $$

In formula (1), $x^{(0)}(k) \geq 0, k = 1, 2, 3, \cdots, n$, the sequence reflects the changing process of maritime traffic accidents over time.

2) In order to weaken the randomness and volatility of water traffic accidents, a new time series is generated by 1-AGO accumulation of the original series $X^{(0)}$,

$$ X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)) $$

In formula (2),

$$ z^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 1, 2, 3, \cdots, n. $$

3) The smoothness and quasi-exponential test of the generated new time series $X^{(1)}$ are carried out. When $k > 3$, if $\rho(k) < 0.5$ and $1 < \sigma(k) < 1.5$, the data satisfy the rules of smoothness and quasi-exponential, the GM (1,1) model can be established. The formula for testing smoothness and quasi-exponential law is as follows:

$$ \rho(k) = x^{(0)}(k)/x^{(1)}(k) $$

$$ \sigma^{(1)}(k) = x^{(1)}(k)/x^{(1)}(k-1) $$

4) Generate a series of adjacent mean values. Adjacent mean values of Sequence $X^{(1)}$:

$$ Z^{(1)} = (z^{(1)}(1), z^{(1)}(2), \cdots, z^{(1)}(n)) $$

The formula is

$$ z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1)), k = 2, 3, \cdots, n. $$

5) Determine parameters of weighted GM(1,1) model by least square method for $X^{(1)}$:

$$ \hat{a} = [a, u]^T = (B^T B)^{-1} B^T Y $$

In formula (6),

$$ B^T = \begin{bmatrix} -z^{(1)}(2) & -z^{(1)}(3) & \cdots & -z^{(1)}(n) \\ 1 & 1 & \cdots & 1 \end{bmatrix}; Y^T = \begin{bmatrix} x^{(0)}(2) & x^{(0)}(3) & \cdots & x^{(0)}(n) \end{bmatrix}. $$

6) By substituting the parameters $a$ and $u$ into the prediction model derived from albino differential equation, the accumulative series of predictions is obtained.

$$ \hat{x}^{(1)}(k) = [x^{(0)}(1) - \frac{u}{a}]e^{-a(k-1)} + \frac{u}{a} $$

The initial value of cumulative reduction is: $x^{(0)}(k+1) = x^{(1)}(k+1) - x^{(1)}(k)$.  

2.2. Residual error unbiased grey GM(1,1) model

1) According to the grey GM(1,1) model, the parameters $a$ and $u$ are obtained, from which the parameters $b$ and $A$ of unbiased grey GM (1,1) model can be obtained: [9]

$$ b = \ln\left(\frac{(2-a)/(2+a)}\right) $$

$$ A = 2u/(2+a) $$

2) By calculating the values of $b$ and $A$, the fitting model of unbiased grey GM (1,1) model is:

$$ \bar{x}^{(0)}(1) = x^{(0)}(1), \bar{x}^{(0)}(k+1) = Ae^{bk}, k = 1, 2, 3, \cdots, n $$

According to the model, a group of predictions can be listed as follows:
The difference between the original series and the predicted series is 
\[ e^{(0)}(t) = x^{(0)}(t) - \bar{x}^{(0)}(t) \],
then the series of residuals is listed as:
\[ e^{(0)}(t) = \{e^{(0)}(1), e^{(0)}(2), e^{(0)}(3), \ldots, e^{(0)}(n)\} \] (12)

If \( e^{(0)}(t) \) series has the negative number, it should be positively processed, and the corresponding absolute value should be taken, after which a new series can be obtained:
\[ e^{(1)}(t) = \{e^{(1)}(1), e^{(1)}(2), e^{(1)}(3), \ldots, e^{(1)}(n)\} \] (13)

The GM (1,1) model of \( e^{(1)}(t) \) is built and the results are as follows:
\[
\begin{aligned}
&\bar{e}^{(0)}(t+1) = e^{(0)}(1) - \frac{\mu_e}{a_e} \left( e^{-a_e^{(t+1)}} - e^{-a_e^{(t-1)}} \right), (t = 1, 2, \ldots, n-1) \\
&k \geq n. The state transition matrix \( P \) is introduced.
\end{aligned}
\] (15)

The mean and variance of the original series \( (0)X \) are also calculated:
\[
\begin{aligned}
\bar{x}^{(0)}(k+1) = \bar{x}^{(0)}(t+1) + \eta \bar{e}^{(0)}(t+1), (t = 1, 2, \ldots, n-1)
\end{aligned}
\] (16)

Among it, the correction factor is:
\[
\eta = \begin{cases} 
1 & x^{(0)}(k) - \bar{x}^{(0)}(k) \geq 0 \\
-1 & x^{(0)}(k) - \bar{x}^{(0)}(k) < 0 
\end{cases}
\] (17)

Markov state transition matrix is used to judge the sign of residual prediction value when \( k \geq n \).

2.3. Test of Forecast Model

Posterior error test and small error probability test are commonly used in grey forecast model test.

2.3.1. Posterior error test. The mean \( \bar{X} \) and variance \( S_1 \) of the original series \( X^{(0)} \) are calculated:
\[
\begin{aligned}
\bar{X} &= \frac{1}{n} \sum_{k=1}^{n} X^{(0)}(k) \ ; \ S_1 = \frac{1}{n} \sum_{k=1}^{n} (X^{(0)}(k) - \bar{X})^2 
\end{aligned}
\] (18)

The mean \( \bar{X} \) and variance \( S_1 \) of the original series \( e^{(0)}(t) \) are also calculated:
\[
\begin{aligned}
\bar{e} &= \frac{1}{n} \sum_{k=1}^{n} e^{(0)}(k) \ ; \ S_2 = \frac{1}{n} \sum_{k=1}^{n} (e^{(0)}(k) - \bar{e})^2
\end{aligned}
\] (19)

Calculating mean square deviation ratio: \( C = S_2 / S_1 \) (20)

2.3.2. Small error probability test.
\[
p = p \left\{ |e(k) - \bar{e}| < 0.6745S_1 \right\} 
\] (21)
Generally speaking, the accuracy level of the model is divided into the following table 1: [11]

| Precision grades | C  | P  | Precision grades | C  | P  |
|------------------|----|----|------------------|----|----|
| Excellent(Level 1)| 0.35| 0.95| Qualified(Level 3)| 0.65| 0.70|
| Good(Level 2)    | 0.50| 0.80| Poor(Level 4)     | 0.80| 0.60|

### 3. Model validation and analysis

Taking the total number of maritime accidents in China from 2008 to 2017 published by the Ministry of Communications as the research object (see Table 2). From the historical statistics of maritime traffic accidents in recent 10 years, it can be seen that the number of maritime traffic accidents in China is on the decline as a whole.

**Table 2. 2008-2017 Distribution table of maritime traffic accidents.**

| year | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 |
|------|------|------|------|------|------|------|------|------|------|------|
| accidents | 342 | 358 | 331 | 298 | 270 | 262 | 260 | 231 | 196 | 196 |

#### 3.1. Establishment of unbiased grey forecast model

Taking the number of maritime traffic accidents from 2008 to 2017 as the original sequence, \( X^{(0)} = (342, 358, 331, 298, 270, 262, 260, 231, 196, 196) \). The original sequence is accumulated once and according to formula (6), it can be obtained that:

\[
[a, u] = (B^T B)^{-1} \begin{bmatrix} 0.078402161 \\ 395.8989012 \end{bmatrix}, \text{So we can get: } b = -0.07844236, A = 3.809646742.
\]

The unbiased grey GM (1,1) forecast model is obtained by introducing \( b \) and \( A \) into formula (10):

\[
\begin{align*}
\hat{x}^{(0)}(1) &= x^{(0)}(1) \\
\hat{x}^{(0)}(k + 1) &= 3.809646742 e^{-0.07844236 k}, (k = 1, 2, 3, \ldots, n)
\end{align*}
\]

The predicted values and residuals are obtained by formula (22) as shown in table 3.

**Table 3. Results of unbiased grey forecast model for maritime traffic accidents.**

| Year | Actual Value | Prediction Value | Residual Error | Year | Actual Value | Prediction Value | Residual Error |
|------|--------------|------------------|----------------|------|--------------|------------------|----------------|
| 2008 | 342          | 342              | 0              | 2014 | 260          | 238              | 22             |
| 2009 | 358          | 352              | 6              | 2015 | 213          | 220              | -7             |
| 2010 | 331          | 326              | 5              | 2016 | 196          | 203              | -7             |
| 2011 | 298          | 301              | -3             | 2017 | 196          | 188              | 8              |
| 2012 | 270          | 278              | -8             | 2018 | -            | 174              | -              |
| 2013 | 262          | 257              | 5              | 2019 | -            | 161              | -              |

According to table 3, although unbiased grey prediction fits the overall distribution characteristics of maritime traffic accidents and shows an exponential downward trend, it fails to fully reflect the random and fluctuating characteristics of accident volume. According to formula (18) - (21), \( C = 0.1564, p = 1 \) is obtained.
3.2. Establishment of unbiased grey forecast model based on residual error
Taking the residual error of maritime traffic accidents prediction results as the original sequence, $e^{(0)}(t) = (0, 6, 5, -3, -8, 5, 22, -7, -7, 8)$. Based on the modelling theory described above, the unbiased grey model of residual error can be obtained that: $[a, u]^T = (B^T B)^{-1} = \begin{pmatrix} -0.063662856 \\ 5.883508929 \end{pmatrix}$. The predicted values and residuals are obtained and shown in Table 4.

**Table 4.** Results of residual error unbiased grey forecast model for maritime traffic accidents.

| Year  | Actual Value | Prediction Value | Residual Error | Year  | Actual Value | Prediction Value | Residual Error |
|-------|--------------|------------------|----------------|-------|--------------|------------------|----------------|
| 2008  | 342          | 342              | 0              | 2014  | 260          | 246              | 14             |
| 2009  | 358          | 358              | 0              | 2015  | 213          | 211              | 2              |
| 2010  | 331          | 333              | -2             | 2016  | 196          | 194              | 2              |
| 2011  | 298          | 294              | 4              | 2017  | 196          | 198              | -2             |
| 2012  | 270          | 271              | -1             | 2018  | -            | 185              | -              |
| 2013  | 262          | 265              | -3             | 2019  | -            | 173              | -              |

According to formula (18) - (21), $C = 0.1549$, $p = 1$ is obtained. The comparison results of the two models’ accuracy are shown in Table 5.

**Table 5.** The comparison results of the two model’s accuracy.

| Model                              | Relative Error E | Mean square deviation ratio C | Small error probability P |
|------------------------------------|------------------|-------------------------------|---------------------------|
| unbiased grey forecast model       | 0.0285           | 0.1564                        | 1                         |
| residual error unbiased grey model | 0.0118           | 0.1549                        | 1                         |

**Figure 1.** Comparison of two forecast models and real values.
The prediction value of unbiased grey forecast model and residual error unbiased grey forecast model are compared with the actual value. The results are shown in figure 1. From figure 1, it can be seen that the residual error unbiased grey model improves the prediction accuracy of unbiased grey model, and the prediction results are close to the actual values. From Table 5, it can be seen that the prediction accuracy of the unbiased grey prediction model with residual error is much higher than that of the unbiased grey prediction model, and the average relative error is reduced by 58.60%. For the posterior difference ratio, the unbiased grey prediction model and the unbiased grey prediction model with residual correction both reach the first level, but the value with residual error is smaller. Therefore, it is better to establish the residual error grey forecast model to predict maritime traffic accidents, and the method is feasible.

4. Summary
The residual error unbiased grey forecast model is used to predict maritime traffic accidents, and the results are satisfactory. The prediction accuracy of residual error unbiased grey forecast model is obviously higher than that of the unbiased grey forecast model. It shows that the residual grey prediction model has certain practicability and effectiveness in the forecasting of maritime traffic accidents. In addition, in the process of correcting unbiased grey prediction model, the residual error model can be superimposed many times in order to obtain higher accuracy.

Acknowledgments
The authors are very grateful to the editors and reviewers for their valuable comments and suggestions. This work was supported by the National Natural Science Foundation of China (Grant No. 71271034), National Social Science Foundation of China (Grant No.L15CGL031).

References
[1] Ju-Long D 2002 THE BASIS OF GREY THEORY.
[2] Xiong-Bin W 2009 Application of residual error gray forecast Model in traffic accident forecast Technology & Economy in Areas of Communications 11(1):33-4.
[3] Li K, Lu L, Zhai J, Taghi M K, Li T 2016 The improved grey model based on particle swarm optimization algorithm for time series prediction Engineering Applications of Artificial Intelligence 55(C):285-91.
[4] Lang-Xiong G, Bao-Gang Z, Yuan-Zhou Z and Jia-Min P 2008 Application of the grey model theory to forecast maritime traffic accident Ship & Ocean Engineering 37(6):99-102.
[5] Yang Z Q 2011 Marine traffic accident prediction based on particle swarm optimization-based RBF neural network International Conference on Computer Research & Development.
[6] Zhao L 2013 Traffic accident prediction based on equal dimension and new information unbiased grey Markov model Computer Engineering & Applications 49(7):35-383.
[7] Qi W and Zhiming W 2013 Forecasting of maritime traffic accidents based on the improved SCGM(1,1) c—Markov Model Navigation of China 36(4):119-24.
[8] Gang L, Huang T, Yan H, Huang X 2009 Grey residual error model of highway traffic accident forecast Journal of Traffic & Transportation Engineering 9(5):88-93.
[9] Ji P R, Huang W S, Hu X Y 2000 An Unbiased Grey Forecasting Model Systems Engineering & Electronics 22(6):6-8.
[10] Lin Y, Chen S, Chen Y and Zhai J 2013 Grey-Markov model and algorithm for traffic accidents forecasting Journal of Wuhan University of Technology 37(5):924-8.
[11] Changyuan C, Ran D, Jijun F, Xingwang Y and Jie Z 2017 Improved grey model applied in maritime traffic safety index prediction Navigation of China 40(1):58-63.