Research on Improved Grey Prediction Method of Airport Passenger Throughput

Jia-juan CHEN1,2, Da-chuan DING1,2,* and Chuan-tao WANG3
1School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
2Research Institute of Highway Ministry of Transport, Beijing 100044, China
3School of Mechanical-Electronic and Vehicle Engineering, Beijing University of Civil Engineering
and Architecture, Beijing 100044, China
*Corresponding author

Keywords: Improved grey prediction, Passenger throughput, Airport.

Abstract. Airport passenger throughput prediction is of great significance to the operation and development of the airport. Based on the grey prediction method, this paper studies the future passenger throughput of the Capital International Airport and the Beijing New Airport. In this paper, the traditional grey prediction model is improved and the passenger throughput of the two airports from 2019 to 2025 is predicted and analyzed. The prediction results show that the improved grey model can improve the accuracy and reduce the prediction error.

Introduction

The prediction of airport passenger throughput is the basic premise of airport and airline operation, and also an important basis for airport resource allocation. The accuracy of prediction results affects the scale of airport construction and expansion directly. This paper improves the grey model and predicts the passenger throughput of the Capital International Airport and the Beijing New Airport from 2019 to 2025, provides a reference for the future operation and management of the two airports.

Prediction methods can be roughly divided into: qualitative analysis [1], trend extrapolation [2], econometric [3], combination prediction method [4-5], etc. Different prediction methods can be used depending on the predicted scene and the predicted data. Grey prediction model (GM) is a time series model based on grey theory, which is used to deal with uncertain and rough data sets. “Grey” reveals an unclear system [6]. Grey theory can deal with incomplete and discrete data. [7], GM model is more robust to noise data and missing data [8]. This model has been proved [9] to be superior to other prediction methods in the processing of short-term prediction. For a long-term prediction, the original GM model needs to be improved.

From this point of view, many predictions are for short-term predictions, and there are fewer improvements to the model. Due to the few data of the Capital International Airport and the Beijing New Airport, the grey prediction model can better fit its data development trend, so this paper uses the grey model for prediction and improves the accuracy of the model.

Improvement of GM (1, 1) Model

The GM (1, 1) Model can be expressed as

\[
\hat{X}^{(1)(k+1)} = ce^{(ak)} + \frac{u}{a}, \quad c = X^{(0)}(1) - \frac{u}{a}, k = 0,1,...n-1.
\]

For equation (1), when \( k=0, \ldots, n-1 \), the data obtained are fitted values. When \( k \geq n \), the data obtained are the predicted values. Performing a subtraction on equation (1):

\[
X^{(0)}(k+1) = X^{(1)}(k+1) - X^{(1)}(k), \quad k = 0,1,...n-1.
\]
In the above solution, it is assumed that the background value is generated by the adjacent equal weight of a cumulative sequence, i.e. $u=0.5$, but $u=0.5$ does not mean that the prediction accuracy of the model is the best. Therefore, the $u$ value is improved by using the automatic optimization method. Let $C = c(1 - e^a)$, and put $C$ in equation (2)

$$X^{(1)}(k+1) = X^{(1)}(1 - e^a)^{-1} e^{(ak)} + \frac{u}{a}, \ k = 0, 1, \ldots, n - 1. \quad (3)$$

That is

$$X^{(0)}(k+1) = X^{(1)}(k+1) - X^{(1)}(k) = Ce^{-ak}, \ k = 1, 2, \ldots, n - 1. \quad (4)$$

The improved $C$ value is

$$C = \left[ \frac{X^{(0)}(1) - \frac{u}{a} (1 - e^a)^{-1} + \sum_{k=2}^{n} X^{(0)}(k) e^{-a(k-1)}}{(1 - e^a)^{-2} + \sum_{k=2}^{n} e^{-2a(k-1)}} \right]. \quad (5)$$

Then use equation (5) to calculate the sum of squares when $u=0$. The value of $S$ was repeatedly calculated until $u=1$, and the actual value and the predicted value under different $u$ values were compared, so as to select the corresponding $u$ value when $S$ is the smallest and use this $u$ value for predictive analysis.

**Case Study**

**Data Preparation**

Based on the radiation range of the two airports, this paper selects the air passenger throughput data for the 10 years from 2008 to 2017 as the original data.

<table>
<thead>
<tr>
<th>Years</th>
<th>Air Passenger Throughput</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>6193.25</td>
</tr>
<tr>
<td>2009</td>
<td>7276.30</td>
</tr>
<tr>
<td>2010</td>
<td>8336.57</td>
</tr>
<tr>
<td>2011</td>
<td>8887.33</td>
</tr>
<tr>
<td>2012</td>
<td>9352.92</td>
</tr>
<tr>
<td>2013</td>
<td>9820.35</td>
</tr>
<tr>
<td>2014</td>
<td>10313.06</td>
</tr>
<tr>
<td>2015</td>
<td>10951.86</td>
</tr>
<tr>
<td>2016</td>
<td>11685.17</td>
</tr>
<tr>
<td>2017</td>
<td>12274.52</td>
</tr>
</tbody>
</table>

According to the data in Table 1, the original sequence $X^{(0)}$ of the prediction data is constructed. $X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), \ldots, X^{(0)}(n)\} = \{6193.25, 7276.30, 8336.57, 8887.33, 9352.92, 9820.35, 10313.06, 10951.86, 11685.17, 12274.52\}$.

Accumulate the sequence $X^{(0)}$.

$X^{(1)} = \{X^{(1)}(1), X^{(1)}(2), \ldots, X^{(1)}(n)\} = \{6193.25, 13469.55, 21806.11, 30693.44, 40046.36, 49866.71, 60179.77, 71131.63, 82816.80, 95091.32\}$.
Model Solution and Test

The original data meet the requirements of grey prediction model, and the model is solved by MATLAB as follows:

\[
X^{(0)}(k+1) = X^{(1)}(k+1) - X^{(1)}(k) = Ce^{-ak} = 7.2702 \times 10^7 e^{-0.059k} \quad (k = 1, 2, \ldots, n-1).
\]  

(6)

The above model was used to predict the passenger throughput within the radiation range of the two airports from 2008 to 2017, and the results were compared with historical data, as shown in table 2 below.

Table 2. Improved grey model prediction results comparison [Ten thousand people]

<table>
<thead>
<tr>
<th>Years</th>
<th>Historical Curve</th>
<th>Predictive Value</th>
<th>Relative Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>6193.25</td>
<td>6194.56</td>
<td>0.0002</td>
</tr>
<tr>
<td>2009</td>
<td>7276.30</td>
<td>7711.82</td>
<td>0.0599</td>
</tr>
<tr>
<td>2010</td>
<td>8336.57</td>
<td>8180.26</td>
<td>0.0188</td>
</tr>
<tr>
<td>2011</td>
<td>8887.33</td>
<td>8677.15</td>
<td>0.0236</td>
</tr>
<tr>
<td>2012</td>
<td>9352.92</td>
<td>9204.22</td>
<td>0.0159</td>
</tr>
<tr>
<td>2013</td>
<td>9820.35</td>
<td>9763.31</td>
<td>0.0058</td>
</tr>
<tr>
<td>2014</td>
<td>10313.06</td>
<td>10356.36</td>
<td>0.0042</td>
</tr>
<tr>
<td>2015</td>
<td>10951.86</td>
<td>10985.43</td>
<td>0.0031</td>
</tr>
<tr>
<td>2016</td>
<td>11685.17</td>
<td>11652.72</td>
<td>0.0028</td>
</tr>
<tr>
<td>2017</td>
<td>12274.52</td>
<td>12360.54</td>
<td>0.0070</td>
</tr>
</tbody>
</table>

(1) Residual Test

The absolute error test and the relative error test of the sequence are carried out by using the Equations 7 and 8 respectively.

\[
\Delta^{(0)}(i) = \left| X^{(0)}(i) - \hat{X}^{(0)}(i) \right|, i = 1, 2, \ldots, n .
\]  

(7)

\[
\phi(i) = \frac{\Delta^{(0)}(i)}{X^{(0)}(i)} \times 100\%, i = 1, 2, \ldots, n .
\]  

(8)

The relative error sequence of the prediction is (0.0002, 0.0599, 0.0188, 0.0236, 0.0159, 0.0058, 0.0042, 0.0031, 0.0028, 0.007), all of them are less than 0.1, and the accuracy of the model is high, which meets the accuracy requirement of residual test.

(2) Correlation Degree Test

Equation 9 is used for the correlation degree test.

\[
r = \frac{1}{n} \sum_{k=1}^{n} \min \left\{ \frac{\hat{X}^{(0)}(k) - X^{(0)}(k)}{X^{(0)}(k)} \right\} + 0.5 \max \left\{ \frac{\hat{X}^{(0)}(k) - X^{(0)}(k)}{X^{(0)}(k)} \right\}.
\]  

(9)

Calculate the correlation degree \(r=0.9990>0.6\), the results meet the accuracy requirements.

(3) Posterior Error Test

Equation 10-11 is used for the posterior error test.
\[ C = \frac{\sum [\Delta^{(0)}(i) - \Delta^{(0)}]^2}{n-1} \]  
\[ \sqrt{\frac{\sum [X^{(0)}(i) - \bar{X}^{(0)}]^2}{n-1}} \]  

(10)

\[ P = p \left[ \Delta^{(0)}(i) - \Delta^{(0)} \right] < 0.6745 \sqrt{\frac{\sum [X^{(0)}(i) - \bar{X}^{(0)}]^2}{n-1}} \]  

(11)

Calculate \( C = 0.0946 < 0.35 \), \( P = 1 \), the results meet the accuracy requirements.

**Model Prediction Analysis**

After the model passes the validity test, the improved grey model meeting the prediction accuracy requirements is used to predict the passenger throughput of the two airports. The results are shown in table 3 below.

<table>
<thead>
<tr>
<th>Years</th>
<th>Air Passenger Throughput</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>13111.35</td>
</tr>
<tr>
<td>2019</td>
<td>13907.77</td>
</tr>
<tr>
<td>2020</td>
<td>14752.56</td>
</tr>
<tr>
<td>2021</td>
<td>15648.67</td>
</tr>
<tr>
<td>2022</td>
<td>16599.22</td>
</tr>
<tr>
<td>2023</td>
<td>17607.50</td>
</tr>
<tr>
<td>2024</td>
<td>18677.03</td>
</tr>
<tr>
<td>2025</td>
<td>19811.52</td>
</tr>
</tbody>
</table>

The predicted results after model improvement are shown in the figure 1.

The improved gray model has better prediction accuracy and more accurate fitting of the original data. However, the traditional gray model has larger deviation between the predicted value and the actual value, and the fitting degree is lower.
The improved grey model completes the prediction with less data. Through the prediction results, we can find that the passenger throughput of the two airports will show a rapid growth in the future, and the Beijing New Airport will have a significant driving role in passenger throughput.

Summary
This paper mainly predicts and analyzes the overall passenger throughput of the Capital International Airport and the Beijing New Airport. According to the characteristics of less passenger throughput data of the airport, this paper uses the grey theory and improves the traditional $GM(1,1)$ model to improve the prediction accuracy. The results show that the future passenger throughput of the two airports presents a rapid growth trend, which will reach 19.81152 million passengers in 2025, with the prediction accuracy of the model reaching 95%.

Acknowledgement
This research was financially supported by the National Key R&D Program of China (No. 2018YFB1601603).

References