Flow Forecasting of the Shijiangzhuang-Jinan High Speed Railway

Yang Dongxue
Shanghai Maritime University Transportation College, Shanghai Pudong New Area, 1550 Harbour Avenue, 201306
Email:863992381@qq.com

Keywords: passenger traffic forecasting; grey correlation; gray prediction model; O-D passenger flow

Abstract: As an important part of connecting the Qingtai Passenger Transport Channel, the construction and opening of the Shiji Passenger Line will ease the state of tension in the existing lines, and play a huge role in promoting the economic development and cultural exchanges between Shijiazhuang and Jinan.

This paper focuses on the passenger traffic forecast for the Shiji passenger line, and analyzes the influence of the socio-economic situation on the passenger volume. The gray forecasting model is used to forecast the annual passenger volume. The OD passenger flow of each section is obtained. Based on this, the paper discusses the result and comes with some directions to be considered in the future.

1. Introduction

The Shiji High Speed Railway studied in this paper starts from Shijiazhuang of Hebei Province, and passes through Yucheng, Xinji, Hengshui, Jingzhou, Dezhou, Pingxian, Yucheng, Qihe and ends in Jinan. With a total length of 323km, there are 11 stations across the line, including 2 passenger stations and 9 intermediate stations. This paper will use the gray prediction model to forecast the passenger traffic of the Shiji high-speed railway in the next two years.

2. Process of passenger flow forecasting

In order to accurately predict the passenger flow of the Shiji Passenger Dedicated Line, it is necessary to determine the socio-economic conditions along the railway, including population, GDP, industrial structure and private car ownership. The determination and quantification of these indicators are crucial for forecasting.

2.1 Determination of the attraction range of railway

The railway line studied in this paper is the Shijiazhuang to Jinan section of the Shiji passenger transport route[1]. The attraction area of the route is mainly distributed in Hebei and Shandong provinces, including Shijiazhuang, Yucheng, Xinji, Hengshui and Jingzhou of Hebei Province; Dezhou, Pingyuan, Yucheng, Qihe and Jinan of Shandong Province.

2.2 Main economic indicators affecting traffic volume along the route

The factors affecting passenger traffic mainly include economic development level and structure, public transportation facilities within the city, private car ownership and population [3].

The main economic indicators of the areas along the route studied in this paper from 2006 to 2018 are shown in Table 1.
<table>
<thead>
<tr>
<th>Years</th>
<th>Railway passenger traffic / 10,000 people</th>
<th>GDP / 100 million yuan</th>
<th>Public transportation vehicle/car ownership / 10,000 cars</th>
<th>Private car ownership / 10,000 cars</th>
<th>Population (10,000 people)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total value</td>
<td>Primary industry</td>
<td>Secondary industry</td>
<td>Tertiary Industry</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>3289.0</td>
<td>7258.96</td>
<td>762.87</td>
<td>4096.5</td>
<td>3025.4</td>
</tr>
<tr>
<td>2007</td>
<td>3316.0</td>
<td>7542.01</td>
<td>785.47</td>
<td>4258.1</td>
<td>3258.9</td>
</tr>
<tr>
<td>2008</td>
<td>3426.2</td>
<td>7889.24</td>
<td>798.24</td>
<td>4478.1</td>
<td>3497.8</td>
</tr>
<tr>
<td>2009</td>
<td>3502.1</td>
<td>8063.85</td>
<td>800.95</td>
<td>4602.7</td>
<td>3582.4</td>
</tr>
<tr>
<td>2010</td>
<td>3698.0</td>
<td>8356.48</td>
<td>837.72</td>
<td>4751.4</td>
<td>3698.4</td>
</tr>
<tr>
<td>2011</td>
<td>3805.0</td>
<td>8745.23</td>
<td>956.33</td>
<td>4985.1</td>
<td>3813.5</td>
</tr>
<tr>
<td>2012</td>
<td>3900.1</td>
<td>9010.00</td>
<td>852.26</td>
<td>4189.7</td>
<td>3967.5</td>
</tr>
<tr>
<td>2013</td>
<td>4050.0</td>
<td>9350.00</td>
<td>899.42</td>
<td>4302.8</td>
<td>4036.7</td>
</tr>
<tr>
<td>2014</td>
<td>4210.0</td>
<td>9751.02</td>
<td>949.21</td>
<td>4587.9</td>
<td>4215.3</td>
</tr>
<tr>
<td>2015</td>
<td>4380.1</td>
<td>10300.5</td>
<td>1002.4</td>
<td>4710.5</td>
<td>4710.5</td>
</tr>
<tr>
<td>2016</td>
<td>4462.0</td>
<td>12103.4</td>
<td>1211.8</td>
<td>4899.3</td>
<td>4899.3</td>
</tr>
<tr>
<td>2017</td>
<td>4613.0</td>
<td>14986.7</td>
<td>1427.2</td>
<td>5123.0</td>
<td>5123.0</td>
</tr>
<tr>
<td>2018</td>
<td>4789.0</td>
<td>15782.6</td>
<td>1542.6</td>
<td>5347.9</td>
<td>5347.9</td>
</tr>
</tbody>
</table>

2.3 Calculate the grey correlation degree of each influencing factor

As can be seen from Table 1, it is necessary to find the most important influencing factors in order to predict passenger traffic based on key factors[2]. In this paper, the gray correlation degree theory is used to quantitatively analyze the above-mentioned many factors to find out the key factors affecting the passenger volume of railway. After five steps of data non-dimensionalization, difference sequence, maximum difference and minimum difference, calculation of correlation coefficient, and correlation degree, after which the correlation of GDP, primary industry, secondary industry, tertiary industry, the number of public vehicles, private car ownership, and total population can be obtained. They are defined as $\gamma_i, i=1,2,3,4,5,6,7$

$$\gamma_1=0.8968, \gamma_2=0.7342, \gamma_3=0.6245, \gamma_4=0.6875, \gamma_5=0.6439, \gamma_6=0.7086, \gamma_7=0.7512 \quad (1)$$

The numerical value of the correlation degree is between 0 and 1. The larger the value is, the closer the relationship between the factor and the predicted value is, and the greater the influence on the predicted value is, that is, the key factor; the smaller the value, the relationship between the factor and the predicted value. The more alienated, the smaller the impact on the predicted value, which is the secondary factor.

In this paper, the factors with correlation degree above 0.7500 are selected as the key influencing factors for forecasting the traffic volume from Shijiazhuang to Jinan. They are GDP and total population [4].

2.4 Grey model predicts passenger traffic

2.4.1 Method and steps for establishing GM(1,1) model

Assumption: the original time series is

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

(2)
The sequence generated by its accumulation is

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

Create a differential equation according to the cumulative generation sequence:

\[ \frac{dX^{(i)}}{dt} + aX^{(i)} = u \]  \hspace{1cm} (4)

The solution is described as discrete

\[ X^{(i)}(t+1) = \left[ X^{(0)}(1) - \frac{u}{a}\right] e^{-at} + \frac{u}{a} \]  \hspace{1cm} (5)

After the parameters \( u \) and \( a \) are determined, according to the recursion of the model, the predicted accumulated sequence can be obtained. After the test, the predicted value can be obtained by performing the subtraction.

The steps are as follows:

1. There is an original sequence. The cumulative generation sequence is calculated according to the following equation \( X^{(1)}(t) \)

\[ X^{(1)}(i) = \sum_{m=1}^{d} X^{(0)}(m) \]  \hspace{1cm} (6)

- According to \( X(1) \), the least squares method is used to determine the model parameters according to the following formula.

\[ \hat{a} = \left( \frac{a}{\mu} \right) = \left( B^T B \right)^{-1} B^T Y_N \]  \hspace{1cm} (7)

In the formula,

\[ B = \begin{bmatrix}
-\frac{1}{2} \left[ X^{(i)}(1) + X^{(i)}(2) \right] & 1 \\
-\frac{1}{2} \left[ X^{(i)}(2) + X^{(i)}(3) \right] & 1 \\
\vdots & \vdots \\
-\frac{1}{2} \left[ X^{(i)}(n-1) + X^{(i)}(n) \right] & 1
\end{bmatrix} \]

\[ Y_N = \begin{bmatrix}
X^{(0)}(2) \\
X^{(0)}(3) \\
\vdots \\
X^{(0)}(n)
\end{bmatrix} \]

- Establish a prediction model to find the cumulative sequence

\[ X^{(i)}(t+1) = \left[ X^{(0)}(1) - \frac{u}{a}\right] e^{-at} + \frac{u}{a} \]  \hspace{1cm} (8)

2. Model checking using residual analysis

- According to the future changes of the system, find the upper and lower bounds of the prediction results, that is, determine the gray plane according to the following formula.

Upper Bound

\[ X^{(i)}_{\text{max}}(n+t) = X^{(i)}(n) + t\delta_{\text{max}} \]  \hspace{1cm} (9)

Lower bound

\[ X^{(i)}_{\text{min}}(n+t) = X^{(i)}(n) + t\delta_{\text{min}} \]  \hspace{1cm} (10)

- Forecasting with models

The above model prediction is based on the prediction result of the "accumulation" generation sequence, and it is restored by the "reduction" generation method, and the predicted value of the original sequence can be obtained, and if the condition of the gray factor is satisfied, the prediction is
2.4.2 Passenger traffic forecast

This prediction uses the method of series prediction. Because of the large amount of calculation, I have selected the data of 7 years from 2012 to 2018 to make GM (1,1) prediction. The data is shown in Table 2.

<table>
<thead>
<tr>
<th>Years</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger traffic (10,000 persons)</td>
<td>3900</td>
<td>4050</td>
<td>4210</td>
<td>4380</td>
<td>4462</td>
<td>4631</td>
<td>4789</td>
</tr>
</tbody>
</table>

Establish GM (1,1) prediction model to predict passenger traffic.

Initial time series:

\[ X^{(0)} = \{3900, 4050, 4210, 4380, 4462, 4631, 4789\} \]

The first step: to accumulate the generated sequence

\[ X^{(1)} = \{3900, 7950, 12160, 16540, 21044, 25675, 30464\} \]

Step 2: Find the parameters by least squares \( \hat{a} = (a, u)^T \)

Substitute \( \hat{a} = (B^T B)^{-1} B^T Y_N \), got:

\[
\hat{a} = \begin{pmatrix} -0.0897 \\ 5174.5631 \end{pmatrix}
\]

Because \( X^{(1)}(1) = 3900 \), got:

\[
X^{(1)}(t+1) = \left[ X^{(1)}(1) - \frac{u}{a} \right] e^{-at} + \frac{u}{a}
\]

\[
= 11314.437 e^{0.0897t} - 5768.4370
\]

Step 3: Inspection

The test results are shown in Table 3. The accuracy of the table is high and the model is available.

<table>
<thead>
<tr>
<th>Years</th>
<th>Calculate data by model ( X^{(1)} )</th>
<th>Raw data ( \hat{X}^{(0)} )</th>
<th>Raw data ( X^{(0)} )</th>
<th>Absolute error</th>
<th>Relative error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>3900</td>
<td>3900</td>
<td>3900</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2013</td>
<td>11395.3</td>
<td>4071.3</td>
<td>4050</td>
<td>21.3</td>
<td>0.0036</td>
</tr>
<tr>
<td>2014</td>
<td>17660.5</td>
<td>4188.8</td>
<td>4210</td>
<td>-21.2</td>
<td>0.0034</td>
</tr>
<tr>
<td>2015</td>
<td>24187.0</td>
<td>4425.5</td>
<td>4380</td>
<td>45.5</td>
<td>0.0069</td>
</tr>
<tr>
<td>2016</td>
<td>31041.6</td>
<td>4414.4</td>
<td>4462</td>
<td>-47.6</td>
<td>0.0070</td>
</tr>
<tr>
<td>2017</td>
<td>38211.3</td>
<td>4591.3</td>
<td>4631</td>
<td>-39.7</td>
<td>0.0056</td>
</tr>
<tr>
<td>2018</td>
<td>45852.1</td>
<td>4499.7</td>
<td>4789</td>
<td>-45.8</td>
<td>0.0060</td>
</tr>
</tbody>
</table>

Step 4: Forecast passenger traffic in 2019 and 2020

2019: \( \hat{X}^{(1)}(8+1) = 11314.437 e^{0.0897*7} - 5768.4370 = 53841.4 \)

\( \hat{X}^{(0)}(9) = 5160.25 \)

2020: \( \hat{X}^{(1)}(9+1) = 11314.437 e^{0.0897*20} - 5768.4370 = 62353.8 \)
\[ X^{(0)}(10) = 6279.15 \]

### 2.5 Distribution of passenger traffic by region

According to the latest year of the material, the data of each region's GDP in 2018, Table 4, as a distribution standard, the passenger volume is finally distributed.

#### Table 4 GDP data of 2018 by region (100 million yuan)

<table>
<thead>
<tr>
<th>Number</th>
<th>Shijiazhuang</th>
<th>Gaocheng</th>
<th>Xinji</th>
<th>Hengshui</th>
<th>Jingzhou</th>
<th>Dezhou</th>
<th>Pingyuan</th>
<th>Yucheng</th>
<th>Qihe</th>
<th>Jinan</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>1010</td>
<td>490</td>
<td>346</td>
<td>331</td>
<td>74</td>
<td>750</td>
<td>189</td>
<td>174</td>
<td>208</td>
<td>1320</td>
</tr>
</tbody>
</table>

The passenger traffic allocation formula is:
Where \( i = 1, 2, 3, ..., 10 \) is the regional GDP, which is the total passenger traffic of the channel.

The calculation results are shown in Table 5.

#### Table 5 Final Annual Forecasted Transportation Volume by Region (10,000 persons)

<table>
<thead>
<tr>
<th></th>
<th>Shijiazhuang</th>
<th>Gaocheng</th>
<th>Xinji</th>
<th>Hengshui</th>
<th>Jingzhou</th>
<th>Dezhou</th>
<th>Pingyuan</th>
<th>Yucheng</th>
<th>Qihe</th>
<th>Jinan</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>388</td>
<td>188</td>
<td>132</td>
<td>127</td>
<td>28</td>
<td>287</td>
<td>72</td>
<td>66</td>
<td>79</td>
<td>506</td>
</tr>
<tr>
<td>2020</td>
<td>468</td>
<td>227</td>
<td>160</td>
<td>153</td>
<td>34</td>
<td>347</td>
<td>87</td>
<td>80</td>
<td>96</td>
<td>611</td>
</tr>
</tbody>
</table>

### 3. Conclusion

Based on the forecast of passenger traffic, this paper has mastered the prediction methods of grey correlation degree theory, regression analysis method and exponential smoothing method, and predicted the passenger traffic along the Shiji passenger railway in 2017 and 2018. This has great guiding significance for the operational decision of the Shiji Passenger Dedicated Line, such as the layout of the train operation map and the preparation of the operation plan[5].

### References


