Highway Passenger Traffic Volume Prediction of Cubic Exponential Smoothing Model Based on Grey System Theory

Wenwen Liu, Yong Qin, Honghui Dong, Yanfang Yang, Zhao Tian
School of Traffic and Transportation, Beijing Jiaotong University
Beijing Research Center of Urban Traffic Information Intelligent Sensing and Service Technologies
Beijing, China
Email: 13120869@bjtu.edu.cn

Abstract—In consideration of highway passenger transportation system is a Grey system of incomplete information; this paper constructs a Grey model and a cubic-exponential-smoothing model at first. In order to avoid the shortcomings of a single model, it constructs a cubic-exponential-smoothing combination model based on Grey system theory by using the optimal weighted principle. Compared with the single model, the prediction accuracy of the cubic-exponential-smoothing combination model is higher, and the predicted results are much more reliable. So it is better suitable for highway passenger traffic volume prediction. At last, this paper gives the prediction of the National highway passenger traffic volume in the first five months of 2014.

Keywords- highway passenger traffic volume prediction; Grey model; cubic exponential smoothing model; combination model

I. INTRODUCTION

To predict the traffic volume and its development trends, characteristics and regularity in perspective and scientifically, is an important theory basis of establishing the highway passenger transport development plan as well as for road bus field. Many scholars have carried on various researches of traffic volume prediction and proposed a variety of prediction methods. Such as time series method, regression analysis, Gray system method, neural network, adaptive filtering method, the method of linear trend model, particle swarm optimization algorithm and support vector machine (SVM), etc. Although these methods are able to forecast the highway passenger traffic volume prediction, the effect is not ideal. The combination model has been applied in recent years, for example, combination prediction based on Grey Markova, combination prediction model based on multivariate statistical and the improved BP network, combination prediction method based on exponential smoothing and Markova model, etc.

This paper is on the basis of fully studies on the working principle of each single model and applicability established in the Grey model and cubic exponential smoothing model, and then established the combination model. In order to contrast with the cubic exponential smoothing model, the article adopted the relative error inspection in Gray model error inspection. Moreover, this paper calculated and compared the prediction results in cubic exponential smoothing model by using different smoothing coefficient, rather than the subjective choice, which makes calculation much more accurate. The prediction results show that the combination prediction can improve the prediction accuracy effectively.

II. THREE EXPONENTIAL SMOOTHING MODEL PREDICT PROCESS

Materials used in the exponential smoothing model is not so much and its operation is simple, and the exponential smoothing is the correction process, by which can automatically correct prediction error. Therefore, it can be used for highway passenger traffic volume forecast. Its computation formula is:

\[ \hat{y}_{t+T} = a_t + b_t T + c_t T^2 \]  \hspace{1cm} (1)

Suppose time series is \( y_1, y_2, \ldots, y_t \), the I times of phase exponential smoothing value is \( S_i \), then the exponential smoothing calculation formula is:

\[ S_i^{(1)} = \alpha y_i + (1 - \alpha) S_{i-1}^{(1)} \]  \hspace{1cm} (2)

\[ S_i^{(2)} = \alpha S_i^{(1)} + (1 - \alpha) S_{i-1}^{(2)} \]  \hspace{1cm} (3)

\[ S_i^{(3)} = \alpha S_i^{(2)} + (1 - \alpha) S_{i-1}^{(3)} \]  \hspace{1cm} (4)

At the above formula, \( \alpha \) is the smooth coefficient. Equation (2) is the first exponential smoothing value, Equation (3) is the secondary exponential smoothing value and Equation (4) is the third exponential smoothing value.

Model coefficient's calculating formula is:

\[ a_t = 3S_t^{(1)} - 3S_t^{(2)} + S_t^{(3)} \]  \hspace{1cm} (5)
\[ b_i = \frac{\alpha}{2(1-\alpha)^2}[(6-5\alpha)S_{i}^{(1)} - 2(5-4\alpha)S_{i}^{(2)}]. \quad (6) \]

\[ c_i = \frac{\alpha^2}{2(1-\alpha)^2}(S_{i}^{(1)} - 2S_{i}^{(2)} + S_{i}^{(3)}). \quad (7) \]

### III. Gray Model Forecasting Process

Grey forecasting method \[2-3\] does not consider the influence factors of the research object, but rather to find useful information in its own time series. Grey forecasting model is set up under the condition of incomplete information or message that not know for sure, therefore it can be used for traffic volume forecasting, which is feasible and has certain practical significance.

#### A. The Process of the Construction of the Gray Model

1) Suppose \(X^{(0)}\) is a negative original sequence, it has the form of \(X^{(0)}(0), X^{(0)}(1), \ldots, X^{(0)}(n)\). We make an accumulation on \(X^{(0)}\), and then get the generate series is:

\[X^{(1)}(0), X^{(1)}(1), \ldots, X^{(1)}(n)\], it can be get from the formula as follows:

\[X^{(1)}(k) = \sum_{i=0}^{k} X^{(0)}(i) \quad k = 1, 2, \ldots, n. \quad (8)\]

2) GM model’s differential equation in the form of an albinos is:

\[\frac{dx^{(1)}}{dt} + ax^{(1)} = u. \quad (9)\]

Parameter \(\hat{a} = [a \quad u]^T = (B^TB)^{-1}B^TY_N\).

3) Construct the data matrix B and data vector \(Y_N\), make

\[Z^{(1)}(k) = 0.5X^{(1)}(k) + 0.5X^{(1)}(k-1), k = 2, 3, \ldots, n\],

get data matrix B and \(Y_N\) are as follows:

\[
B = \begin{bmatrix}
-x^{(1)}(2)
-x^{(1)}(3)
\vdots
-x^{(1)}(n)
\end{bmatrix}
\]

\[Y_N = [x^{(0)}(2), x^{(0)}(3), x^{(0)}(4), \ldots, x^{(0)}(n)]^T. \quad (10)\]

4) Determine the discrete solution:

\[x^{(1)}(k + 1) = (x^{(0)}(1) - \frac{u}{a})e^{-ak} + \frac{u}{a}. \quad (11)\]

5) Restore to the original data:

\[x^{(0)}(k + 1) = (x^{(1)}(k + 1) - x^{(1)}(k)) \text{. And the result is:} \]

\[x^{(0)}(k + 1) = -a(x^{(0)}(1) - u/a)e^{-ar}. \quad (12)\]

#### B. Grey Model Error Inspection

The precision of the Grey model test \[4\] generally has three methods: relative error size test, correlation test and posterior error method. In order to compare the prediction results with the cubic exponential smoothing model, this paper uses relative error size test.

Calculate residual error is:

\[E = (e(1), e(2), \ldots, e(n)) = X^{(0)} - \hat{X}^{(0)}. \quad (13)\]

The formula of relative error is:

\[rel(k) = \frac{e(k)}{x^{(0)}(k)} \times 100\%, k = 1, 2, \ldots, n. \quad (15)\]

### IV. The Construction of Combination Model

Combination model \[5-6\] means to forecast the same object at first by using different forecast methods, and then in order to improve the prediction precision, constructs a new model through weighted combination. The optimal weighting method \[7\] is based on certain standards to construct the objective function \(Q\), calculate the weighted coefficients of combination model under the constraints (such as the weight sum to 1) as well as minimize \(Q\).

Suppose traffic volume forecast \[8-9\] problem have \(k\) single forecasting models \(y_i (i = 1, 2, \ldots, k)\), and to forecast the \(n\) periods. Record \(y_{n}^{(i)}\) is the first \(i\) kind of forecasting model’s exact number in \(t\) time, \(y_{n}^{(i)}\) is the first \(i\) kind of forecasting model’s prediction, \(e_{n}^{(i)}\) is the first \(i\) kind of forecasting model’s forecasting error, \(p_{i}\) is the first \(i\) kind of forecasting model’s weighting coefficient and determine...
\[ \sum_{i=1}^{k} p_i = 1 , \] through which we can get the combined forecasting model form:

\[ Y = p_1 \hat{y}_1 + p_2 \hat{y}_2 + p_k \hat{y}_k = \sum_{i=1}^{k} p_i \hat{y}_i . \] (16)

Assume that a single forecasting model fitting error is:

\[ e_{it} = y_{it} - \hat{y}_{it} (i = 1, 2, \ldots, k; t = 1, 2, \ldots, n) \] then the prediction model can form fitting error matrix:

\[
E = \begin{bmatrix}
\sum_{t=1}^{n} e_{i1}^2 & \sum_{t=1}^{n} e_{i1} e_{i2} & \cdots & \sum_{t=1}^{n} e_{i1} e_{ik} \\
\sum_{t=1}^{n} e_{i2} e_{i1} & \sum_{t=1}^{n} e_{i2}^2 & \cdots & \sum_{t=1}^{n} e_{i2} e_{ik} \\
\vdots & \vdots & \ddots & \vdots \\
\sum_{t=1}^{n} e_{ik} e_{i1} & \sum_{t=1}^{n} e_{ik} e_{i2} & \cdots & \sum_{t=1}^{n} e_{ik}^2
\end{bmatrix} . \] (17)

The process of solving the combination model is to get the extreme value in the objective function under the constraint condition by using least square method. The constraint conditions is:

\[ \min Q = \sum_{i=1}^{n} e_{i}^2 + \sum_{i=1}^{k} p_i = 1 . \] (18)

Define R is column vector with component of the whole of 1, namely \( R = [1, \ldots, 1]^T \), record \( P = [p_1, p_2, \ldots, p_m] \), so by the constraint conditions can obtain:

\[ \min Q = \sum_{i=1}^{n} e_{i}^2 = P^T E P, \sum_{i=1}^{k} R^T P = 1 . \] (19)

Calculate the above formulas by using the Lagrange multiplier method, the optimal weight vector is obtained as follows:

\[ P = E^{-1} R \] (20)

The minimum value of the objective function is:

\[ \min Q = \frac{1}{R^T E^{-1} R} . \] (21)

V. EXPERIMENTAL ANALYSIS

The data is from the national statistics bureau's monthly statistics, this paper selected the transport of passengers inside. Make use of highway passenger traffic of January - December, 2013 at first, utilized two kinds of prediction model\[10\] for modeling calculation respectively.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic</td>
<td>309398</td>
<td>321130</td>
<td>302412</td>
<td>301245</td>
<td>305909</td>
<td>301530</td>
</tr>
<tr>
<td>July</td>
<td>304765</td>
<td>311055</td>
<td>315035</td>
<td>330585</td>
<td>314156</td>
<td>329636</td>
</tr>
</tbody>
</table>

A. Exponential Smoothing Model Prediction

<table>
<thead>
<tr>
<th>Month</th>
<th>Prediction</th>
<th>Absolute error</th>
<th>Relative error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan.</td>
<td>309398</td>
<td>306538</td>
<td>315024</td>
</tr>
<tr>
<td>Feb.</td>
<td>321130</td>
<td>324054</td>
<td>333025</td>
</tr>
<tr>
<td>Mar.</td>
<td>302412</td>
<td>312066</td>
<td>32056</td>
</tr>
<tr>
<td>Apr.</td>
<td>301245</td>
<td>320449</td>
<td>316697</td>
</tr>
<tr>
<td>May.</td>
<td>305909</td>
<td>319489</td>
<td>31697</td>
</tr>
<tr>
<td>Jun.</td>
<td>301530</td>
<td>322669</td>
<td>313443</td>
</tr>
<tr>
<td>July</td>
<td>304765</td>
<td>312066</td>
<td>32056</td>
</tr>
<tr>
<td>Aug.</td>
<td>311055</td>
<td>320449</td>
<td>316697</td>
</tr>
<tr>
<td>Sep.</td>
<td>315035</td>
<td>322669</td>
<td>313443</td>
</tr>
<tr>
<td>Oct.</td>
<td>330585</td>
<td>322669</td>
<td>313443</td>
</tr>
<tr>
<td>Nov.</td>
<td>314156</td>
<td>322669</td>
<td>313443</td>
</tr>
<tr>
<td>Dec.</td>
<td>329636</td>
<td>322669</td>
<td>313443</td>
</tr>
</tbody>
</table>

| Mean   | 302238     | 312238         | 327286             | 321956              | 322877              | 15048               | 9718                | 10639              | 4.89                | 3.13               | 3.41               |

188
The key in calculation of exponential smoothing is the selection of $\alpha$. In general, if the data is volatile, $\alpha$ should be larger, by which can increase the influence of the prediction results of recent data. If the data fluctuate smoothly, $\alpha$ should be smaller. Theorists think that when time series have fluctuations, but the long-term trend change is not big, we can pick and choose the larger value, in which usually ranges from 0.2 to 0.5. This paper selected three values (0.3, 0.4, and 0.5) and calculated the prediction, then analyzed the error.

By analyze the results in Table 2, we can get that when $\alpha$ equal to 0.4, the average absolute error and relative error are the minimum, which has a better prediction results. So, the smoothing coefficient is 0.4. When $\alpha$ take 0.4, we can calculated $a_t = 315836.2$, $b_t = -5063.49$, $c_t = -513.904$, and then obtained the prediction model is:

$$\hat{y}_{12+T} = 315836.2 - 5063.49T - 513.904T^2$$

### B. Grey Model Prediction

By using Grey theory to deal with the original data, the GM (1, 1) model is established. Among them $a = -0.0061$, $u = 300304.9517$. Then the prediction model is:

$$x^{(0)}(k + 1) = 302192.3 \times e^{0.0061k}$$

predicted value can be obtained by using Grey model in Table 3.

### Table III Three Kinds of Prediction Model Prediction Value and Error.

<table>
<thead>
<tr>
<th>Month</th>
<th>Traffic</th>
<th>GM value</th>
<th>GM RE</th>
<th>Exponential prediction</th>
<th>Exponential Relative error</th>
<th>Portfolio prediction</th>
<th>Portfolio Absolute error</th>
<th>Portfolio Relative error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>309398</td>
<td>309398.00</td>
<td>0.00%</td>
<td>306538.42</td>
<td>0.92%</td>
<td>308571.46</td>
<td>-826.54</td>
<td>0.27%</td>
</tr>
<tr>
<td>2</td>
<td>321130</td>
<td>303110.55</td>
<td>5.61%</td>
<td>324053.66</td>
<td>0.91%</td>
<td>309156.10</td>
<td>-11973.90</td>
<td>3.73%</td>
</tr>
<tr>
<td>3</td>
<td>302412</td>
<td>304945.03</td>
<td>0.84%</td>
<td>309718.00</td>
<td>2.42%</td>
<td>306324.42</td>
<td>3912.42</td>
<td>1.29%</td>
</tr>
<tr>
<td>4</td>
<td>301245</td>
<td>306800.72</td>
<td>1.84%</td>
<td>312066.37</td>
<td>3.59%</td>
<td>308322.39</td>
<td>7077.39</td>
<td>2.35%</td>
</tr>
<tr>
<td>5</td>
<td>305909</td>
<td>308667.71</td>
<td>0.90%</td>
<td>319489.37</td>
<td>4.44%</td>
<td>311795.06</td>
<td>5886.06</td>
<td>1.92%</td>
</tr>
<tr>
<td>6</td>
<td>301530</td>
<td>310546.06</td>
<td>2.99%</td>
<td>316477.77</td>
<td>4.96%</td>
<td>312260.39</td>
<td>10730.39</td>
<td>3.56%</td>
</tr>
<tr>
<td>7</td>
<td>304765</td>
<td>312435.85</td>
<td>2.52%</td>
<td>320448.48</td>
<td>5.15%</td>
<td>314751.65</td>
<td>9986.65</td>
<td>3.28%</td>
</tr>
<tr>
<td>8</td>
<td>311055</td>
<td>314373.13</td>
<td>1.06%</td>
<td>326625.48</td>
<td>5.01%</td>
<td>317888.61</td>
<td>6833.61</td>
<td>2.20%</td>
</tr>
<tr>
<td>9</td>
<td>315035</td>
<td>316249.98</td>
<td>0.39%</td>
<td>329325.41</td>
<td>4.54%</td>
<td>320028.66</td>
<td>4993.66</td>
<td>1.59%</td>
</tr>
<tr>
<td>10</td>
<td>330585</td>
<td>318174.47</td>
<td>3.75%</td>
<td>342940.26</td>
<td>3.74%</td>
<td>325331.71</td>
<td>-5253.29</td>
<td>1.59%</td>
</tr>
<tr>
<td>11</td>
<td>314156</td>
<td>320110.67</td>
<td>1.90%</td>
<td>322353.84</td>
<td>2.61%</td>
<td>320758.99</td>
<td>6602.99</td>
<td>2.10%</td>
</tr>
<tr>
<td>12</td>
<td>329636</td>
<td>322058.66</td>
<td>2.30%</td>
<td>333434.69</td>
<td>1.15%</td>
<td>325346.42</td>
<td>-4295.58</td>
<td>1.30%</td>
</tr>
</tbody>
</table>

### C. Combination Model Prediction

Suppose $y_1$ is the prediction of GM (1, 1) model, $y_2$ is the prediction of cubic exponential smoothing model. Using the two models of the known data, you can get $E = \begin{pmatrix} 769228268 \\ 253612711 \end{pmatrix}$, by the above equation; the optimal weight vector is calculated $P = \begin{bmatrix} 0.711 \\ 0.289 \end{bmatrix}$

Then get the optimal combination model:

$$y = 0.711y_1 + 0.289y_2$$

By analyze of Table 3 we can concluded that the combination model has an average relative error absolute value 0.95%, which is far less than the gray model MAPE 2.01% and the cubic exponential smoothing model MAPE 3.13%. Therefore, the combination forecast model of prediction effect is better than a single model, the prediction error is smaller.

By analyzing the combination model and figure 1, we can drawn that it has the highest precision, as well as has obvious advantages in traffic volume forecast, so the use of combination model to predict can get much more accurate results. Using the combination model to forecast national traffic volume.
highway passenger traffic volume of January-May, 2014, the results are as follows:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>320041.95</td>
<td>319534.97</td>
<td>318739.49</td>
<td>317655.55</td>
<td>316283.21</td>
</tr>
</tbody>
</table>

**VI. CONCLUSION**

1) The optimal weighted combination model comprehensively utilized the effective information of the GM (1, 1) model, the cubic exponential smoothing model, by which avoided the limitation of single theory, improved the forecasting precision, reduced the prediction error, and made the prediction effect better.

2) In this paper the choice of smoothing coefficient in cubic exponential smoothing model is based on general experience, only selected three coefficients to analyze, so there still have shortcomings need to improve.

3) The rationality and feasibility of the model is verified, and the combination model can improve the prediction accuracy according to the national highway passenger traffic data collected from the national bureau of statistics.

4) Predicted the national highway passenger traffic volume of 2014 by using the combination model, which has an important reference value to transportation management department in arranging the passenger transport management reasonable and strengthening the management of bus passenger terminals.

**ACKNOWLEDGEMENT**

Manuscript received April 22, 2014. This work is supported by the National 863 program (2012AA112401) and Science and Technology Supporting Program (2014BAG01B02), and is also supported by the National Nature Science Foundation under Grant61104164.

**REFERENCES**


