

# 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 Jiao tong 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, the effect is not ideal. The combination model has been applied in recent years, for example, combination prediction based on Grey Markova, combination prediction method 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<sup>[1]</sup> is not so much and it ' s 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:

$$\hat{y}_{t+T} = a_t + b_t T + c_t T^2. \quad (1)$$

Suppose time series is  $y_1, y_2, \dots, y_t$ , the I times of phase t exponential smoothing value is  $S_t$ , then the exponential smoothing calculation formula is:

$$S_t^{(1)} = \alpha y_t + (1 - \alpha) S_{t-1}^{(1)}. \quad (2)$$

$$S_t^{(2)} = \alpha S_t^{(1)} + (1 - \alpha) S_{t-1}^{(2)}. \quad (3)$$

$$S_t^{(3)} = \alpha S_t^{(2)} + (1 - \alpha) S_{t-1}^{(3)}. \quad (4)$$

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:

$$a_t = 3S_t^{(1)} - 3S_t^{(2)} + S_t^{(3)}. \quad (5)$$

$$b_i = \frac{\alpha}{2(1-\alpha)^2} [(6-5\alpha)S_i^{(1)} - 2(5-4\alpha)S_i^{(2)} + (4-3\alpha)S_i^{(3)}] \quad (6)$$

$$c_i = \frac{\alpha^2}{2(1-\alpha)^2} (S_i^{(1)} - 2S_i^{(2)} + S_i^{(3)}) \quad (7)$$

### 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  $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ , we make an accumulation on  $X^{(0)}$ , and then get the generate series is:

$X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$ , it can be get from the formula as follows:

$$X^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i) \quad k=1, 2, \dots, n. \quad (8)$$

2) GM model's differential equation in the form of an albino is:

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

Parameter  $\hat{a} = [a \ u]^T = (B^T B)^{-1} B^T Y_N$ .

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

$$Z^{(1)}(k) = 0.5X^{(1)}(k) + 0.5X^{(1)}(k-1), k=2, 3, \dots, n,$$

get data matrix B and  $Y_N$  are as follows:

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \dots & \dots \\ -z^{(1)}(n) & 1 \end{bmatrix},$$

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

4) Determine the discrete solution:

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

5) Restore to the original data:

$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$ , And the result is:

$$\hat{x}^{(0)}(k+1) = -a(x^{(0)}(1) - u/a)e^{-ak}. \quad (12)$$

#### 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:

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

$$e(k) = x^{(0)}(k) - \hat{x}^{(0)}(k). \quad (14)$$

The formula of relative error is:

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

### 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  $y_i (i=1, 2, \dots, k)$ , and to forecast the n periods. Record  $y_{it}$  is the first I kind of forecasting model's exact number in t time,  $\hat{y}_{it}$  is the first I kind of forecasting model's predication,  $e_{it}$  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:

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

Assume that a single forecasting model fitting error is:  $e_{it} = y_{it} - \hat{y}_{it} (i = 1, 2, \dots, k; t = 1, 2, \dots, n)$  then the prediction model can form fitting error matrix:

$$E = \begin{bmatrix} \sum_{t=1}^n e_{1t}^2 & \sum_{t=1}^n e_{1t}e_{2t} & \cdots & \sum_{t=1}^n e_{1t}e_{kt} \\ \sum_{t=1}^n e_{2t}e_{1t} & \sum_{t=1}^n e_{2t}^2 & \cdots & \sum_{t=1}^n e_{2t}e_{kt} \\ \vdots & \vdots & & \vdots \\ \sum_{t=1}^n e_{kt}e_{1t} & \sum_{t=1}^n e_{kt}e_{2t} & \cdots & \sum_{t=1}^n e_{kt}^2 \end{bmatrix}. \quad (17)$$

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:

$$\min Q = \sum_{t=1}^n e_t^2, \sum_{i=1}^k p_i = 1. \quad (18)$$

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

$$\min Q = \sum_{t=1}^n e_t^2 = P^T E P, \sum_{i=1}^k R^T P = 1. \quad (19)$$

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

$$P = \frac{E^{-1}R}{R^T E^{-1}R}. \quad (20)$$

The minimum value of the objective function is:

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

## 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<sup>[10]</sup> for modeling calculation respectively.

TABLE I 2013 NATIONAL HIGHWAY PASSENGER TRAFFIC (TEN THOUSAND).

month	Jan.	Feb.	Mar.	Apr.	May.	Jun.
traffic	309398	321130	302412	301245	305909	301530
month	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
traffic	304765	311055	315035	330585	314156	329636

### A. Exponential Smoothing Model Prediction

TABLE II PREDICTED RESULTS COMPARED WITH THREE KINDS OF SMOOTHING COEFFICIENT.

Month	Traffic	Prediction			Absolute error			Relative error (%)		
		0.3	0.4	0.5	0.3	0.4	0.5	0.3	0.4	0.5
1	309398	323154	306538	321024	13756	-2860	11626	4.45	0.92	3.76
2	321130	339617	324054	333025	18487	2924	11895	5.76	0.91	3.70
3	302412	322669	309718	313443	20257	7306	11031	6.70	2.42	3.65
4	301245	323148	312066	312056	21903	10821	10811	7.27	3.59	3.59
5	305909	328772	319489	316697	22863	13580	10788	7.47	4.44	3.53
6	301530	323564	316478	311969	22034	14948	10439	7.31	4.96	3.46
7	304765	325403	320449	315110	20638	15683	10345	6.77	5.15	3.39
8	311055	329304	326626	321394	18249	15570	10339	5.87	5.01	3.32
9	315035	329503	329325	325276	14468	14290	10241	4.59	4.54	3.25
10	330585	340334	342940	341080	9749	12355	10495	2.95	3.74	3.17
11	314156	317097	322354	323882	2941	8198	9726	0.94	2.61	3.10
12	329636	324869	333435	339572	-4767	3799	9936	1.45	1.15	3.01
Mean	312238	327286	321956	322877	15048	9718	10639	4.89	3.13	3.41

The key in calculation of exponential smoothing is the selection of  $\alpha$ . In general, if the data is volatile, value should be larger, by which can increase the influence of the prediction results of recent data. If the data fluctuate smoothly, value of  $\alpha$  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:  $\hat{x}^{(0)}(k+1) = 302192.3 \times e^{0.0061k}$ , predicted value can be obtained by using Grey model in Table 3.

TABLE III THREE KINDS OF PREDICTION MODEL PREDICTION VALUE AND ERROR.

Month	Traffic	GM value	GM RE	Exponential prediction	Exponential Relative error	Portfolio prediction	Portfolio Absolute error	Portfolio Relative error
1	309398	309398.00	0.00%	306538.42	0.92%	308571.46	-826.54	0.27%
2	321130	303100.55	5.61%	324053.66	0.91%	309156.10	-11973.90	3.73%
3	302412	304945.03	0.84%	309718.00	2.42%	306324.42	3912.42	1.29%
4	301245	306800.72	1.84%	312066.37	3.59%	308322.39	7077.39	2.35%
5	305909	308667.71	0.90%	319489.37	4.44%	311795.06	5886.06	1.92%
6	301530	310546.06	2.99%	316477.77	4.96%	312260.39	10730.39	3.56%
7	304765	312435.85	2.52%	320448.48	5.15%	314751.65	9986.65	3.28%
8	311055	314337.13	1.06%	326625.48	5.01%	317888.61	6833.61	2.20%
9	315035	316249.98	0.39%	329325.41	4.54%	320028.66	4993.66	1.59%
10	330585	318174.47	3.75%	342940.26	3.74%	325331.71	-5253.29	1.59%
11	314156	320110.67	1.90%	322353.84	2.61%	320758.99	6602.99	2.10%
12	329636	322058.66	2.30%	333434.69	1.15%	325346.42	-4289.58	1.30%

### 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  $E = \begin{pmatrix} 769228268 & 253612711 \\ 253612711 & 1521931276 \end{pmatrix}$ , by the above

equation; the optimal weight vector is calculated

$$P = [0.711 \ 0.289]^T$$

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.

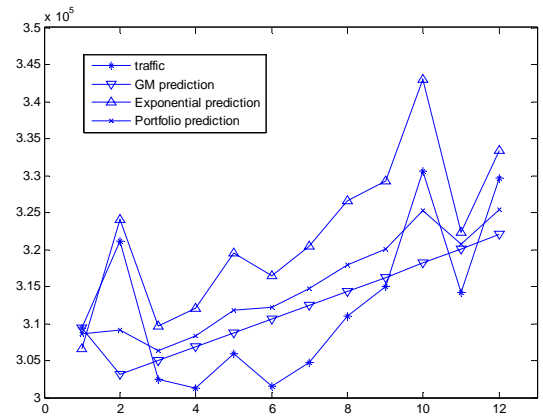


Fig. 1 Comparing the results of several kinds of prediction.

By analyzing the combination model and figure 1, we can draw 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

highway passenger traffic volume of January-May, 2014, the results are as follows:

TABLE IV THE PASSENGER TRAFFIC VOLUME OF COMBINATION FORECAST MODEL IN 2014.

Month	Jan.	Feb.	Mar.	Apr.	May.
Prediction (ten thousand)	320041.95	319534.97	318739.49	317655.55	316283.21

## 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

- [1] XIE Anoxia-Ru, JIANG Hui-yuan, SHEN Yao-we. Passenger Traffic Volume Forecast based on Gray-linear Regression Combination Model, [J]. RAILWAY TRANSPORT AND ECONOMY, 2008, 30(8).
- [2] LIU Si-eng, DENG J-long. The Range Suitable for GM (1, 1) [J]. SYSTEMS ENGINEERING—THEORY&PRACTICE, 2000, 20(5).
- [3] BAO Beng, TANG Chen-min. Modified Grey Model of Passenger Volume of City Public Traffic [J]. JOURNAL OF BEIJING JIAOTONG UNIVERSITY, 2005, 29(3).
- [4] MA Anoxia-ake, WANG Ci-Guangzhou. The Application of the Cubic Exponential Smoothing Method on Volume Forecasting of Da-Ain Railway, [J]. JOURNAL OF EAST CHINA JIAOTONG UNIVERSITY, 2005, 22(3).
- [5] ZHANG Ling-gang, NIU De-ning, MENG Zhao-min, FAN Jia-dong. Application of Combination Forecasting Model in Traffic Volume Forecasting, [J]. TECHNOLOGY & ECONOMY IN AREAS OF COMMUNICATIONS, 2010, 12(5).
- [6] HOU Li-min, MA Guo-feng. Forecast of Railway Passenger Traffic Based on a Grey Linear Regression Combined Model, [J]. COMPUTER SIMULATION, 2011, 28(7).
- [7] ZHAO Ling, XU Hongke, CHENG Hongliang. Road traffic accidents prediction based on optimal weighted combined model, [J]. Computer Engineering and Applications, 2013, (24).
- [8] Xinkai Wua, Henry X. A shockwave profile model for traffic flow on congested urban arterials, [J]. Transportation Research Part B: Methodological. Volume 45, Issue 10, December 2011, Pages 1768–1786.
- [9] Kranti Kumara, M. Paridab, V.K. Katiyar . Short Term Traffic Flow Prediction for a Non Urban Highway Using Artificial Neural Network Procedia, [J]. Social and Behavioral Sciences. Volume 104, 2 December 2013, Pages 755–764.
- [10] Xiaojian HU, Wei WANG, Hui SHENG. Urban Traffic Flow Prediction with Variable Cell Transmission Model, [J]. Journal of Transportation Systems Engineering and Information Technology. Volume 10, Issue 4, August 2010, Pages 73–78.