

Road Traffic Accident Forecast Based on Optimized Grey Verhulst Model

Subing Liu, Congwei Wu

College of Science, Rocket Force University of Engineering, China

*liusubing2002@163.com

Keywords: road traffic accident; grey verhulst model; prediction; accuracy test

Abstract. Road traffic safety system is a grey system, which can be forecasted by the optimized Grey Verhulst model developed in this paper. The original Grey Verhulst model is optimized and improved from two aspects: the grey whitening derivative problem and the initial values. The new optimized model is established to forecast road traffic accidents, meanwhile the detail modeling steps is also given. Taking Chinese's road traffic accident statistics as an example, the optimized model is tested. The result shows that the improved Grey Verhulst model is with high prediction accuracy and low error, which provides a reference for the accurate prediction of road traffic accidents.

1 Introduction

Traffic accident forecasting is the prerequisite and basis for road traffic security administration, plans and decisions. The purpose of forecasting is to know well the future growth of traffic accidents and to estimate the feasibility and effects of traffic safety measures reasonably. It can also be used to control every influencing factor effectively, and to achieve the goal of decreasing traffic accidents.

The purpose of traffic accident prediction is to find the law in the past related data and to predict the future situation according to the found law. Many methods had been used to forecast traffic accident, such as multiple regressions, artificial neural networks and so on. In those methods the connection between exogenous factors and accidents had been used to find the related law. Yuan wei [1] did the regression analysis by considering many key influencing factors, such as motor vehicle population, population quantity, road mileage, turnover amount of passenger and freight transport, national control power. Shi zhian [2] analyzed the main influencing factors of the traffic accident by the grey relational analysis and used the BP neural networks model to forecast the traffic accident. Sliupas [3] did the same analysis by introducing a variety of road influencing factors. Qin Liyan [4] developed a forecasting model based on FM network predicting model and genetic algorithms. Liu Xiuqing [5] established a forecasting model basing on radial basis neural network, and verified the validity of the model.

Traffic accidents acted randomly, accidentally and fuzzy, so there are many accidents caused by uncertainties, such as traffic flow, motor speed, motor status and human factor. The collection and determination of those factors was difficult, and it caused the feasible prediction in many cases. So exploring the law from the development of the accident itself has become another way to predict the traffic accident. Theory of road traffic safety system has grey property, as part of the information such as road conditions, traffic facilities are known, but other factors are unknown such as driver's mental state or weather situation. Grey forecasting model can be used for small samples and poor information, and it can also be used to predicting system's operation trend. Therefore, in recent years the amount of road traffic accidents were often considered as the behavioral characteristics of road traffic safety system processing and the grey model has been used to forecast the traffic accidents [6~10].

GM(1,1) and Grey Verhulst model were used in the literature [11]. In the literature [12] the combined model of the grey model and Markov chain was developed and had fine forecasting effect. But the accuracy of its predictions was influenced by the state of division obviously, and the repeated comparisons of different division increase the complexity of the model. Due to the limitation of the original model, the original model has been optimized and improved from two aspects: the grey whitening derivative problem and the initial values. The outline of the paper is presented as follows:

Section 2 introduces the traditional Grey Verhulst model. Section 3 develops the new optimized model. Section 4 performs application example. Finally, Section 5 summarizes the paper.

2 Traditional Grey Verhulst Model

Definition 1. Consider the following data as non-negative data sequence

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

With the first-order data accumulation, the 1-AGO sequence can be obtained:

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

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

Definition 2. Define the follow equation as Grey Verhulst model

$$x^{(0)}(k) + az^{(1)}(k) = b(z^{(1)}(k))^2 \quad (1)$$

where, $\mathbf{Z}^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n))$ is the neighbors mean sequences of $\mathbf{X}^{(1)}$, it can be written as:

$$z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k-1) + x^{(1)}(k)) \quad (2)$$

Definition 3. Define the follow equation as Grey Verhulst model's winterization equation

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b(x^{(1)})^2 \quad (3)$$

Set the parameters column as $\hat{a} = (a, b)^T$. Based on Eq.1 and the least square method, the column $\hat{a} = (a, b)^T$ can be obtained as $\hat{a} = (a, b)^T = (B^T B)^{-1} B^T Y$,

$$\text{Where: } B = \begin{pmatrix} -z^{(1)}(2) & (z^{(1)}(2))^2 \\ -z^{(1)}(3) & (z^{(1)}(3))^2 \\ \vdots & \vdots \\ -z^{(1)}(n) & (z^{(1)}(n))^2 \end{pmatrix}, \quad Y = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{pmatrix}.$$

Grey Verhulst model's time-dependent equation is listed as follows:

$$\hat{x}^{(1)}(k+1) = \frac{1}{\frac{a}{b} + \left(\frac{1}{x^{(1)}(1)} - \frac{a}{b}\right)e^{ak}} \quad (4)$$

The reducing sequence is defined as:

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1), \quad k = 2, \dots, n \quad (5)$$

3 Establishment of Improved Grey Verhulst Model

3.1 Analysis of Grey Verhulst Model.

According to the modeling process of traditional Grey Verhulst model, the forecasting accuracy of the model depended on the precision of Eq.1 and the selected initial value. On the one hand, according to the original model, parameters which is estimated from Eq.1 is applied to Eq.3 which can derived the time depended response function and carried out forecasting. So the accuracy of the model is depended on the accuracy of Eq.1. In fact, the above modeling method is the direct use of $x^{(0)}(t)$ as a

grey derivative term $dx^{(1)}/dt$, and used the trapezoidal formula to estimated the area between the fitting curve and the time axis. By using the above mentioned method $z^{(1)}(t)$ can be obtained. Substituting $z^{(1)}(t)$ into Eq.3, we can get Eq.1. This modeling method cut off the necessary connection between the difference equations and differential equations [14].

On the other hand, the default initial value $(1, x^{(0)}(1))$ was selected by the traditional Grey Verhulst model, and this is not tally with the actual situation. Because the optimized fitted curve is not always going through a certain point in historical data. So, in this paper, both of two above mentioned questions are considered, and then the new optimized Grey Verhulst model is developed. First, according to Eq.14 a new Verhulst model is definite, and then based on the optimized initial value, the new model is established.

3.2 The Improved Grey Verhulst Model.

(1) Using the trapezoid formulation to whiten derivative, then a new definition of Verhulst model can be obtained.

On account of the modeling idea of grey model, the smooth curve fitting $x^{(1)}(t)$ of the first-order data accumulation can be obtained. According to the trapezoid formula, $x^{(1)}(t+1)$ can be obtained as:

$$x^{(1)}(t+1) = x^{(1)}(t) + \frac{1}{2} \left[\frac{dx^{(1)}(t+1)}{dt} + \frac{dx^{(1)}(t)}{dt} \right] \quad (6)$$

Where:

$$x^{(1)}(t+1) = x^{(1)}(t) + x^{(0)}(t+1) \quad (7)$$

According to Eq.6 and Eq.7, we can get

$$\frac{dx^{(1)}(t+1)}{dt} + \frac{dx^{(1)}(t)}{dt} = 2x^{(0)}(t+1) \quad (8)$$

The following equation can be obtained from the Grey Verhulst model's winterization equation

$$\begin{cases} \frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b(x^{(1)}(t))^2 \\ \frac{dx^{(1)}(t+1)}{dt} + ax^{(1)}(t+1) = b(x^{(1)}(t+1))^2 \end{cases} \quad (9)$$

Adding the two equations in upper Eq.9, we have the following expression:

$$\frac{dx^{(1)}(t)}{dt} + \frac{dx^{(1)}(t+1)}{dt} + a(x^{(1)}(t) + x^{(1)}(t+1)) = b[(x^{(1)}(t))^2 + (x^{(1)}(t+1))^2]$$

Substituting Eq.9 into Eq.8, we can get

$$x^{(0)}(t+1) + \frac{a}{2}[x^{(1)}(t) + x^{(1)}(t+1)] = \frac{b}{2}[(x^{(1)}(t))^2 + (x^{(1)}(t+1))^2] \quad (10)$$

By the discretion of Eq.10, a new definition can be derived as:

$$x^{(0)}(k+1) + \frac{a}{2}[x^{(1)}(k) + x^{(1)}(k+1)] = \frac{b}{2}[(x^{(1)}(k))^2 + (x^{(1)}(k+1))^2] \quad (11)$$

According to the least square method, $\hat{a} = (a, b)^T = (B^T B)^{-1} B^T Y$,

$$\text{Where: } B = \begin{pmatrix} -z^{(1)}(2) & (x^{(1)}(1))^2 + (x^{(1)}(2))^2 \\ -z^{(1)}(3) & (x^{(1)}(2))^2 + (x^{(1)}(3))^2 \\ \vdots & \vdots \\ -z^{(1)}(n) & (x^{(1)}(n-1))^2 + (x^{(1)}(n))^2 \end{pmatrix}, Y = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{pmatrix}.$$

(2) Establishment of initial value of the optimized model:

$$\text{Let } \hat{x}^{(1)}(k) = \frac{1}{\frac{b}{a} + ce^{a(k-1)}} \quad (12)$$

By using the least square method and under the law of minimum square sum of error, unconstrained optimized model can be established:

$$\min_c f(c) = \min_c \sum_{k=2}^n (\hat{x}^{(1)}(k))^{-1} - (x^{(1)}(k))^{-1})^2 = \min_c \sum_{k=2}^{n-1} (\frac{b}{a} + ce^{a(k-1)} - (x^{(1)}(k))^{-1})^2 \quad (13)$$

Let $df(c)/dc = 0$, we can get the only stationary point, that is:

$$c = \frac{\sum_{k=2}^{n-1} ((x^{(1)}(k))^{-1} - \frac{b}{a})e^{a(k-1)}}{\sum_{k=2}^{n-1} e^{2a(k-1)}} \quad (14)$$

Substituting Eq.14 into Eq.12, we can get the sequence $\hat{x}^{(1)}(k)$.

3.3 Modeling Steps of Improved Grey Verhulst Model

- (1) Based on sequence $X^{(0)}$, get the 1-AGO sequence $X^{(1)}$.
- (2) Using the trapezoid formulation to whiten derivative and then the new definition of Verhulst model can be obtained.
- (3) By using the least square method, the parameter column a, b can be obtained.
- (4) Establishing the unconstrained optimized model to optimized the initial value and to get the parameter of c .
- (5) Substituting c into Eq.12, the new time depended sequence $\hat{x}^{(1)}(k)$ can be got.
- (6) Proceed Error Analysis.

4 Application Examples

Basing on 2003-2011 year's historical data of Chinese road traffic accident (shown in Table 1), the new model is tested for the forecasting. As the original data has the property of S-curve, and the gray Verhulst model is mostly used to describe the saturated process which has S-curve property, so it can be used for predictions. For further analysis the efficiency of optimized Grey Verhulst model, the numerical test is carried out, and the comparison between two models is also shown.

In this paper the data from 2003-2009 year are selected as simulated data, and data from 2010 - 2011 year are selected as forecasted data. The traditional and the optimized Grey Verhulst model are carried out for the simulation and prediction. The simulated values and the relative errors are shown in Table 2, while, the predicted values and the relative errors are shown in Table 3.

Table 1. The Historical Data of Chinese Road Traffic Accident [Ten Thousand]

Year	2003	2004	2005	2006	2007	2008	2009	2010	2011
Original data	10.4	9.4	9.8738	8.8455	8.1649	7.3484	6.7759	6.5225	6.238
1-AGO data	10.4	19.8	29.6738	38.5193	46.6842	54.0326	60.8085	67.3310	73.5690

Table 2. Comparison of Simulated Values and Relative Errors by Using Two Different Models

Year	Actual value	Traditional model		Optimized model	
	$x^{(1)}(k)$	Emulated value	Relative error	Emulated value	Relative error
2003	10.4	10.4	0.0	10.4	0.0
2004	19.8	16.7158	0.1558	19.7849	0.0008
2005	29.6738	25.2034	0.1507	29.1475	0.0177
2006	38.5193	35.0812	0.0893	39.3576	0.0218
2007	46.6842	44.8227	0.0399	48.7118	0.0434
2008	54.0326	52.9811	0.0195	56.0094	0.0366
2009	60.8085	58.9290	0.0309	61.0209	0.0035
Average relative error		0.0810		0.0206	

Table 3. Comparison of Predicted Values and Relative Errors

Year	Actual value	Traditional model		Optimized model	
	$x^{(1)}(k)$	Predicted value	Relative error	Predicted value	Relative error
2010	67.3310	62.8411	0.0677	64.1684	0.0470
2011	73.5690	65.2430	0.1132	66.0357	0.1024

From Table 2 and Table 3, we can see that the optimized model has fine accuracy than the traditional model. Meanwhile, the predicted error of the first and second step which calculated by the optimized model is also lower than that of the traditional model. As the traditional model ignore the influence of the error caused by the initial value and the background value. The results show that the optimized model which developed in this article has higher accuracy than the traditional model, so the new model can mine the information and development law of the original data effectively, and it can raise the forecasting ability of the model.

5 Conclusions

Aiming at improving the prediction accuracy of traffic accident and basing on the statistics of traffic accidents, a new forecasting model is developed in this article and applied to the forecasting of Chinese traffic accident. The results show that the new optimization model has higher forecast accuracy than the traditional Grey Verhulst model. By using the trapezoid formula to whiten derivative and optimizing the initial value, the new model can improve forecasting performance, so it will be a practical, simple and high accuracy method for forecasting of the traffic accident.

References

- [1] Yuan Wei, Fu Rui, Guo Ying-shi etc. Prediction model of death toll resulted from road traffic accidents[J], Journal of Traffic and Transportation Engineering, 2007, 7(4): 112-116.
- [2] SHI Zhi. An Road Traffic Accident Prediction Based on MATLAB Neural Network Model[D]. Chang An University, 2010.
- [3] SLIUPAS T. The Impact of road parameters and the surrounding area on traffic accidents[J]. Transport, 2009, 24(1): 42-47.
- [4] Qin Li-yan, Shao Chun-fu, Zhao Liang. Macro Prediction Model of Road Traffic Accident Based on Neural Network [J]. Journal of Wuhan University of Technology(Transportation Science & Engineering), 2010, 34(1): 154-157.
- [5] LIU Xiu-qing, WANG Xiao-yuan, YU Ren-de. Study on traffic accidents prediction model based on RBF neural network[J]. Computer Engineering and Applications, 2009, 45(17): 188-190.
- [6] QIAN Wei-dong, LIU Zhi-qiang. Road Traffic Accident Forecast Based on Grey-Markov Model [J]. China Safety Science Journal, 2008, 18(3): 33-36.

- [7] Garcia-Ferrer A, De Juan A, Poncela P. Forecasting traffic accidents using disaggregated data[J]. *International Journal of Forecasting*, 2006, 22(2): 203-222.
- [8] Mao M, Chirwa E C. Application of grey model GM (1, 1) to vehicle fatality risk estimation [J]. *Technological Forecasting and Social Change*, 2006, 73(5): 588-605.
- [9] LIU Zhao-hui, WANG Chao, XU Hong-guo. Grey Prediction Model for Road Traffic Accident and Its Application Research[J]. *Journal of Changchun University of Science and Technology (Natural Science Edition)*. 2006, 29(4): 42-44.
- [10] ZHAO Jiang-ping, DING Jia-li. Road Accidents Prediction based on Wavelet Analysis of the Grey GM(1,1) Model[J]. *Mathematics in Practice and Theory*. 2015, 45(12): 119-124.
- [11] WANG Yan-jun, SUN You-xin. Forecast of road accidents based on GM(1,1) model[J]. *Technology & Economy in Areas of Communications*, 2007, 9(5): 97-98.
- [12] LIN Yan, CHEN Shuai, CHEN Yan, etc. Grey-Markov Model and Algorithm for Traffic Accidents Forecasting[J]. *Journal of Wuhan University of Technology (Transportation Science & Engineering)*, 2013, 37(5): 924-929.
- [13] LIU Si-feng etc. Grey systems theory and application[M]. Beijing: Science Press, 2014: 198-199.
- [14] WANG Zheng-xin, DANG Yao-guo, SHEN Chun-guang. Improvement on Derivative Item in Grey Verhulst model[J]. *Statistics & Information Forum*, 2010, 25(6): 19-22.