

A PREDICTION MODEL OF ROAD TRAFFIC FATALITIES IN HANOI USING IMPROVED GREY MODEL GM (1,1)

VUONG XUAN CAN^{1,2}, MOU RUI-FANG¹, VU TRONG THUAT²

¹School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, China

²University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.

Corresponding author's email: 1323972686@qq.com

Abstract: This study improves the first-order and one-variable grey differential equation model (abbreviated as GM(1,1) model), which requires only a limited amount of data to estimate the behavior of a system with unknown and known information, to predict road traffic fatalities (RTFs) in Hanoi in short-term with experimental data of RTFs from 2010 to 2018. Firstly, we introduce basic concept of original GM(1,1), and then an improved GM(1,1) model is established based on the modification of the errors. Finally, we compare the performance of forecasting models to predict RTFs in Hanoi. The results of prediction models show that the forecasting accuracy of the improved GM(1,1) model is higher than the original GM(1,1) model. Comparison with original GM(1,1) model, mean absolute percentage error (MAPE) of the improved GM(1,1) model is considerably decreased from 8.30% to 1.70%. This proves the applicability of the improved GM(1,1) model to make a short-term prediction for RTFs in Hanoi.

Keywords: prediction model, grey model GM (1,1), road traffic fatalities

Received: 06/04/2020

Accepted: 13/05/2020

Published online: 14/06/2020

I. INTRODUCTION

With the rapid development of urbanization and motorization across Hanoi, as well as other cities in Vietnam, road transportation system is playing an increasingly role in urban transportation system. However, the occurrence of road traffic accidents (RTAs) in Hanoi still remain high as a result of the speedy growing of motorized vehicles, and it has become a more and more important factor that restrict the development of economy and threaten the safety of human and public health. According to statistic data, about 7.4 percent of RTAs and 7.1 percent of total fatalities by RTAs in Vietnam have occurred in Hanoi [1], and the rate of the road traffic fatalities (RTFs) is approximately 8 deaths per 100,000 inhabitants [2]. Accurate forecasting of RTAs would greatly contribute to reasonable road networks planning and the improvement management of road traffic safety [3], especially in Vietnam where RTAs remain high. So far, there are many prediction models of RTAs that have been announced, such as regression models, grey models [4-6], artificial neural network [7], hybrid models [8], and so forth. Among them, grey prediction models, which are one of most important parts of the grey system theory, propounded by Deng [9], have been commonly used in the time series prediction due to their

simplicity and ability to characterize an unknown system with high accuracy and limited data [10,11].

In Hanoi, as well as other cities in Vietnam, forecasting of RTAs has not respected in both short and long term. Besides, the statistics of RTAs are still difficult and exist many inadequacies, such as lack of systematization, under-reporting, incompleteness, incorrectness and inaccessibility [1,8,12]. There are many data sources but the data is not sufficient, it is difficult to apply traditional statistical analysis methods that require data and information large enough with many factors considered. Considering the complexity and uncertainty of the influencing factors on RTAs, forecasting of RTAs can be regarded as a grey system with unknown and known information [6], so forecasting of RTAs in the context of Vietnamese transport using by grey prediction models is very appropriate.

The first-order one-variable grey differential equation model, abbreviate as GM(1, 1) is used most widely among various grey models, and has been successfully applied to various fields, e.g., forecasting of RTAs [4,5,6]. The GM(1,1) model has many advantages [13], such as small amount of data needed, simple modeling process, and easy to learn and use. However, the GM (1,1) model is monotonic due to the exponential function of the prediction model, the results of the GM (1,1) model sometime may not be satisfactory for the random fluctuation series. Recently the GM (1,1) model has been expanded to improve the accuracy and prediction ability.

In this paper, we first introduce basic concept of GM(1,1), and then present an improved GM (1,1) model using Fourier series [14,15,16,17]. Finally, we compare the performance of forecasting models to predict in Hanoi in short-term.

II. MAIN CONTENTS

2.1. GM (1,1) model

GM (1, 1) grey prediction model is that through the monotone sequence has been given, by means of once additive generates a group of new albinism differential equation. The GM(1, 1) model has the three main operations: (1) Accumulating Generation Operator (AGO), (2) grey modelling, (3) Inverse Accumulating Generation Operator (IAGO), and it is summarized as follows [18].

Assume that a non-negative original time-series of data $X^{(0)}$ with n samples is expressed as:

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)\} \quad (1)$$

where, the superscription (0) of $X^{(0)}$ represents the original series; $x^{(0)}(k)$, ($k=1, 2, \dots, n$) is the data point at time k , and $n \geq 4$. Let a new cumulative $X^{(1)}$ be the first-order AGO of $X^{(0)}$, whose elements are generated from $X^{(0)}$:

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), \dots, x^{(1)}(n)\} \quad (2)$$

where $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$, for $k = 1, 2, \dots, n$. The generated mean sequence $Z^{(1)}$ of $X^{(1)}$ is defined as:

$$Z^{(1)} = \{z^{(1)}(1), z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)\} \quad (3)$$

where $z^{(1)}(k)$ is the mean value of adjacent data, and is calculated by:

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1), k \geq 2 \quad (4)$$

The GM (1, 1) model can be formed by establishing a first order differential equation for $X^{(1)}$ as follows:

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = b \quad (5)$$

where a and b are the developing coefficient and the endogenous control coefficient, respectively. They can be calculated by traditional least square method as follows:

$$[a, b]^T = (B^T \cdot B)^{-1} \cdot B^T \cdot Y \quad (6)$$

where Y is called data series, B is called data matrix, as follows:

$$Y = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]^T \quad (7)$$

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \dots & \dots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad (8)$$

According to Eq. (5), the solution of $X^{(1)}$ as follows:

$$\begin{cases} \hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak} + \frac{b}{a}, k \geq 1 \\ \hat{x}^{(1)}(1) = x^{(0)}(1) \end{cases} \quad (9)$$

Applying the IAGO, the forecasted value can be obtained as follows:

$$\begin{cases} \hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), k \geq 1 \\ \hat{x}^{(0)}(1) = x^{(0)}(1) \end{cases} \quad (10)$$

or

$$\begin{cases} \hat{x}^{(0)}(k+1) = (1 - e^a) \left(x^{(0)}(1) - \frac{b}{a}\right) e^{-ak}, k \geq 1 \\ \hat{x}^{(0)}(1) = x^{(0)}(1) \end{cases} \quad (11)$$

2.2. Improved GM (1,1) model

In order to improve the accuracy of forecasting models, this study uses Fourier series in modifying the errors in the GM(1,1) model [14,15,16,17]. The improved model is implemented as follows.

Suppose that $\varepsilon^{(0)} = \{\varepsilon^{(0)}(2), \varepsilon^{(0)}(3), \dots, \varepsilon^{(0)}(n)\}$ is the error sequence of $X^{(0)}$ where $\varepsilon^{(0)}(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$. If all errors are positive, then a remnant GM(1,1) model can be built [19]. When the errors can be positive or negative, the errors can be expressed using by Fourier series, and are rewritten as follows [14].

$$\hat{\varepsilon}^{(0)}(k) \cong \frac{1}{2}a_0 + \sum_1^z \left[a_i \cos\left(\frac{2\pi i}{n-1}k\right) + b_i \sin\left(\frac{2\pi i}{n-1}k\right) \right], k = 2, 3, \dots, n \quad (12)$$

where $z = [(n-1)/2] - 1$ is called the minimum deployment frequency of Fourier series [15], and it only take integer number. Therefore, Equation (12) can be estimated as:

$$\varepsilon^{(0)} \cong PC \quad (13)$$

where C is a vector of coefficients; P is a matrix. They can be expressed as

$$P = \begin{bmatrix} \frac{1}{2} & \cos(2\frac{2\pi \cdot 1}{n-1}) & \sin(2\frac{2\pi \cdot 1}{n-1}) & \dots & \cos(2\frac{2\pi \cdot z}{n-1}) & \sin(2\frac{2\pi \cdot z}{n-1}) \\ \frac{1}{2} & \cos(3\frac{2\pi \cdot 1}{n-1}) & \sin(3\frac{2\pi \cdot 1}{n-1}) & \dots & \cos(3\frac{2\pi \cdot z}{n-1}) & \sin(3\frac{2\pi \cdot z}{n-1}) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{1}{2} & \cos(n\frac{2\pi \cdot 1}{n-1}) & \sin(n\frac{2\pi \cdot 1}{n-1}) & \dots & \cos(n\frac{2\pi \cdot z}{n-1}) & \sin(n\frac{2\pi \cdot z}{n-1}) \end{bmatrix} \quad (14)$$

$$C = [a_0, a_1, b_1, a_2, b_2, \dots, a_z, b_z]^T \quad (15)$$

The parameters $a_0, a_1, b_1, a_2, b_2, \dots, a_z, b_z$ are obtained by using the ordinary least squares method whose results are in the following equation:

$$C = (P^T P)^{-1} P^T \varepsilon^{(0)T} \quad (16)$$

From the original prediction series ($\hat{x}^{(0)}$) and the error ($\hat{\varepsilon}^{(0)}$), prediction series of the improved grey model using Fourier series ($\hat{x}_F^{(0)}$) is determined by

$$\begin{cases} \hat{x}_F^{(0)}(k) = \hat{x}^{(0)}(k) + \hat{\varepsilon}^{(0)}(k), k \geq 2 \\ \hat{x}_F^{(0)}(1) = \hat{x}^{(0)}(1) \end{cases} \quad (17)$$

In order to evaluate the performance of the proposed model, we use mean absolute percentage error (MAPE), which are defined as follows:

$$MAPE = \frac{1}{n} \sum_{k=1}^n \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \quad (18)$$

where, $x^{(0)}(k)$ and $\hat{x}^{(0)}(k)$ are the actual value and forecasted value, respectively. The MAPE value is more than 10%, the result is an inaccurate forecast, 5%–10% is a reasonable forecast, 1%–5% is a good forecast, and less than 1% is an excellent forecast [8].

2.3. APPLICATION OF THE MODELS

RTFs data of Hanoi from 2010 to 2018 collected from Hanoi Department of Transport (HDOT) [20] is selected to establish prediction models. Number of RTFs from 2010 to 2015 is used as the in-sample, and number of RTFs from 2016 to 2018 is used as out-of-sample. As shown in Table 1, although the trend of RTFs in Hanoi from 2010 to 2018 has decreased slightly.

Table 1. RTFs data of Hanoi from 2010 to 2018.

Year	RTFs	Year	RTFs	Year	RTFs
2010	807	2013	626	2016	594
2011	749	2014	609	2017	583
2012	619	2015	602	2018	543

The actual values and forecasted values implemented by PYHTON 3.8 are shown in Table 2 to compare relative errors of the models. MAEPs for the different stage are shown in Table 3.

Table 2. Models values and prediction errors of RTFs in Hanoi.

Stage	Year	Actual value	Original GM(1,1)		Improved GM(1,1)	
			Model value	Error (%)	Model value	Error (%)
2010-2015	2010	807	807	0	807	0
	2011	749	705	-5.85	739	-1.30
	2012	619	671	8.47	629	1.57
	2013	626	639	2.12	616	-1.55
	2014	609	609	-0.06	619	1.59
2016-2018	2015	602	579	-3.74	592	-1.61
	2016	594	552	-7.12	604	1.63
	2017	583	525	-9.90	573	-1.67
	2018	543	500	-7.89	553	1.79

Table 3. MAPEs for the different prediction models.

Stage	MAPE (%)		Performance	
	Original GM(1,1)	Improved GM(1,1)	Original GM(1,1)	Improved GM(1,1)
2010-2015	3.37	1.27	Good	Good
2016-2018	8.30	1.70	Reasonable	Good

Table 2 shows that the improved GM(1,1) model is smaller than the original GM(1,1) model by comparing the relative error. As shown in Table 3, it can be seen that MAPE of the improved GM(1,1) model and original GM(1,1) from 2016 to 2018 are 8.3% and 1.7%, respectively. In comparison with original GM(1,1), MAPE of the improved GM(1,1) is considerably decreased. The effectiveness and accuracy of improved GM(1,1) model is higher than the original GM(1,1). The trend of RTFs in Hanoi from 2010-2018 by actual values and models values is shown in Figure 1. This figure also shows that the improved GM(1,1) model has the better forecasting precision in 2016-2018, it is more suitable to make a short-term prediction.

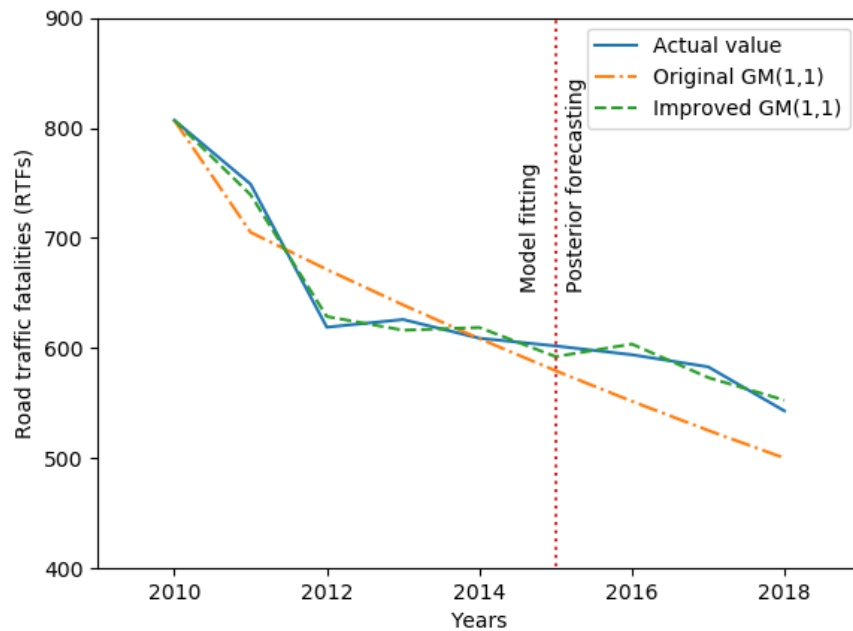


Figure 1. The trend of RTFs in Hanoi from 2010-2018.

III. CONCLUSIONS

The grey prediction models could deal with the problems with incomplete and also the small sample, so this paper uses it to predict RTFs in Hanoi. An improved GM(1,1) model also is established based on original GM(1,1) model and the modification of the errors. The results show that the accuracy of improved GM(1,1) model in forecasting value for 2010 to 2018 is higher than the original GM(1,1) model. The study results of the paper not only supplement the gap in prediction of RTAs in Vietnam, but also provide useful information for people who work in traffic management. However, the paper is limited to short-term prediction with limited data, the next step will need to be further developed for long-term studies considering many influencing factors.

ACKNOWLEDGMENT

This research is funded by the Ministry of Education and Training of Vietnam under grant No. CT.2019.05.03.

References

- [1] R.-F. Mou, X.-C. Vuong, and T.-T. Vu, Analysis of Road Accident in Hanoi, Vietnam, in Proceedings of the Eastern Asia Society for Transportation Studies, Vol.12, 2019.
- [2] X.-C. Vuong, R.-F. Mou, H.-S. Nguyen, and T.-T. Vu, Signal Timing Optimization of Isolated Intersection for Mixed Traffic Flow in Hanoi City of Vietnam Using VISSIM, in International Conference on Smart Vehicular Technology, Transportation, Communication and Applications, 2018, pp. 133–139.
- [3] J. Zheng and X. Wu, Prediction of road traffic accidents using a combined model based on IOWGA operator, Periodica Polytechnica Transportation Engineering, 43(3) (2015) 146–153.
- [4] H. Tongyuan and W. Yue, Forecasting Model of Urban Traffic Accidents Based on Gray Model GM (1,1), in Second Workshop on Digital Media and its Application in Museum & Heritages (DMAMH

2007), 2007, pp. 438–441.

- [5] U. Becker and H. Manz, Grey systems theory time series prediction applied to road traffic safety in Germany, *IFAC-PapersOnLine*, 49 (3) (2016) 231–236.
- [6] A. Mohammadi, L. Moradi, A. Talebnejad, and A. Nadaf, The use of grey system theory in predicting the road traffic accident in Fars province in Iran, *Aust. J. Bus. Manag. Res.*, 1(9)(2011) 18-23.
- [7] A. P. Akgüngör and E. Doğan, An artificial intelligent approach to traffic accident estimation: Model development and application, *Transport*, 24(2)(2009)135–142.
- [8] X.-C. Vuong, R.-F. Mou, and T.-T. Vu, Analysis and Forecast of Road Traffic Accidents in Vietnam Based on Grey BP Neural Network, in *The 11th National Conference on Fundamental and Applied IT Research*, 2018, pp. 35–41.
- [9] D. Julong, Introduction to grey system theory, *J. grey Syst.*, 1(1)(1989) 1–24.
- [10] L. Qu, D. He, and R. Jia, Optimized grey model based on cuckoo search algorithm and its prediction application, *J. Inf. & COMPUTATIONAL Sci.*, 11(5)(2014) 1419–1426.
- [11] S. Liu, J. Forrest, and Y. Yang, A brief introduction to grey systems theory, in *Proceedings of 2011 IEEE International Conference on Grey Systems and Intelligent Services*, 2011, pp. 1–9.
- [12] N. H. Duc, D. T. M. Hoa, N. T. Huong, and N. N. Bao, Study on quality of existing traffic accident data in Vietnam, in *Proceedings of the Eastern Asia Society for Transportation Studies*, 2011, vol. 8.
- [13] B. Zeng, H. Duan, Y. Bai, and W. Meng, Forecasting the output of shale gas in China using an unbiased grey model and weakening buffer operator, *Energy*, 151(2018) 238–249.
- [14] C. L. Tan and S. P. Chang, Residual correction method of Fourier series to GM (1, 1) model, in *Proceedings of the first national conference on grey theory and applications*, Kauhsiung, Taiwan, 1996, pp. 93–101.
- [15] Y.-L. Huang and Y.-H. Lee, Accurately forecasting model for the stochastic volatility data in tourism demand, *Mod. Econ.*, 2(05)(2011) 823-829.
- [16] Z. Guo, X. Song, and J. Ye, A Verhulst model on time series error corrected for port throughput forecasting, *J. East. Asia Soc. Transp. Stud.*, 6(2005) 881–891.
- [17] W. Chia-Nan and P. Van-Thanh, An improved nonlinear grey Bernoulli model combined with Fourier series,” *Math. Probl. Eng.*, (2015) No.740272.
- [18] J. L. Deng, *Grey forecasting and decision*, Huazhong Univ. Sci. Technol. Press. Wuhan, pp. 97–134, 1986.
- [19] S. Liu and J. Y. L. Forrest, *Grey systems: theory and applications*. Springer Science & Business Media, 2010.
- [20] HDOT, *Traffic Safety Annual Reports (2010-2018)*, Hanoi, Vietnam, 2018.