The use of Grey System Theory in predicting the road traffic accident in Fars province in Iran

Ali Mohammadi: Corresponding author
Management Department, Shiraz University, Iran
amohamadi11@gmail.com

leyla Moradi
Management Department, Shiraz University, Iran
leilamoradi19@yahoo.com

Ahmad Talebnejad
Management Department, Shiraz University, Iran
taleb149@yshoo.com

Abdolreza Nadaf
Management Department, Shiraz University, Iran
dr.arnadaf@gmail.com

ABSTRACT

Traffic accidents have become a more and more important factor that restrict the development of economy and threaten the safety of human beings. Considering the complexity and uncertainty of the influencing factors on traffic accidents, traffic accident forecasting can be regarded as a grey system with unknown and known information, so be analyzed by grey system theory. Grey models require only a limited amount of data to estimate the behavior of unknown systems. In this paper, first, the original predicted values of road traffic accidents are separately obtained by the GM (1,1) model, the Verhulst model and the DGM(2,1) model. The results of these models on predicting road traffic accident show that the forecasting accuracy of the GM(1,1) is higher than the Verhulst model and the DGM(2,1) model. Then, the GM(1,1) model is applied to predict road traffic accident in Fars province.

Keywords: *Road traffic accident, Grey system theory, The GM(1.1) model, Forecasting.*

1. INTRODUCTION

With the rapid development of economy, road transportation system is playing an increasingly role in comprehensive transportation system. However, the occurrence of traffic accident shows a rising tendency as a result of the quick growing of motor vehicles. Though the occurrence of traffic accidents is occasional, it can be predicted scientifically based on the related statistical indexes. Accurate forecasting of the transport accident is important not only for governments, but also for insurance companies making decisions for the future.

Grey system theory, propounded by Deng [2], is a truly multidisciplinary theory dealing with grey systems that are characterized by both partially known and partially unknown information. It has been widely used in several fields such as agriculture, industry and environmental systems studies. As an essential part of grey system theory, grey forecasting models have gained in popularity in time-series forecasting due to their simplicity and ability and high precision to characterize an unknown system by using as few as four data points [7, 10].

In recent years, the grey system theory has been widely used to forecast in various fields and demonstrated satisfactory results. For instance, Tongyuan and Yue (2007) , had applied grey model to predict the urban traffic accident in China [6] , Lin et al (2009) had presented the adaptive and high precision grey forecasting model to predict the stock index in Taiwan , Zhu (2010) had used the composite grey BP neural network forecasting model to motor vehicle fatality risk in China [11], Zhou et al (2006) had used a trigonometric grey prediction approach to forecasting electricity demand [10], Lu (2007) had used the grey system theory to analyze and forecast the road traffic safety improvement in Netherlands [5], Zhan- li and Jin- hua (2011) had applied the Grey- Markov model in forecasting fire accidents in China [9], Wu and Chen (2005) had used a prediction model using the grey model GMC(1, n) combined with the grey relational analysis to predict the internet access population in Taiwan [8].

Considering the complexity and uncertainty of the influencing factors on traffic accidents, traffic accident forecasting can be regarded as a grey system with unknown and known information, so be analyzed by grey system theory. [11]

In this paper, we first introduce basic concept of GM(1,1), Verhulst model and DGM(2,1) mode, and then compare the performance of three models to predict traffic accident. Finally, the GM(1,1) model is applied to predict road traffic accident in Fars province.

2- GREY MODELS

2-1- The GM (1, 1) model

The most commonly used grey forecasting model is GM(1, 1), which indicates one variable is employed in the model and the first order differential equation is adopted to match the data generated by the accumulation generating operation (AGO). The AGO reveal the hidden regular pattern in the system development. [4]
Before the algorithm of GM(1, 1) [1,4, 11] is described, the raw data series is assumed to be

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

where n is the total number of modeling data.
The AGO formation of $X^{(1)}$ is defined as:

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

Where

$$x^{(1)}(k) = \sum_{j=1}^{k} x^{(0)}(j) \ (k = 1,2,\ldots,n) . \quad (3)$$

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

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = u \quad (4)$$

where parameters a and u are called the developing coefficient and grey input, respectively.

In practice, parameters a and u are not calculated directly from Eq. (4). Therefore, the solution of (4) can be obtained by using the least square method. That is,

$$\hat{x}^{(1)}(k+1) = [x^{(0)}(1) - \frac{u}{a}]e^{-ak} + \frac{u}{a} \quad (5)$$

where

$$\hat{a} = [a,u]^T = (B^T B)^{-1} B^T Y$$

and

$$B = \begin{bmatrix} -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{bmatrix} \quad (6)$$

$$Y_N = (x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n))^T . \quad (7)$$

Applying the inverse accumulated generation operation (IAGO). Namely

$$\hat{x}^{(0)}(k) = [x^{(0)}(1) - \frac{u}{a}](1 - e^{a})e^{-a(k-1)} \qquad (8)$$

Where

$$\hat{x}^{(1)}(1) = x^{(0)}(1) \quad (k = 2,3,\ldots,n).$$

2-2- The grey Verhulst model

The Verhulst model [3, 11] was first introduced by a German biologist Pierre Franois Verhulst. The main purpose of Velhulst model is to limit the whole development for a real system.

For an initial time sequence,

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

The initial sequence $X^{(0)}$ is used to directly construct the Verhulst model.

$$\frac{dX^{(0)}}{dt} + aX^{(0)} = u(X^{(0)})^2 \qquad (9)$$

Where a is development coefficient, u is grey action quantity. The solution of the parameter vector

$$\hat{a} = [a,u]^T$$

can be obtained by utilizing the least square method. Where

$$\hat{a} = [(A \vdots B)^T (A \vdots B)]^{-1} (A \vdots B)^T Y$$

and

$$A = \begin{bmatrix} -\frac{1}{2}(x^{(0)}(1) + x^{(0)}(2)) \\ -\frac{1}{2}(x^{(0)}(2) + x^{(0)}(3)) \\ \vdots \\ -\frac{1}{2}(x^{(0)}(n-1) + x^{(0)}(n)) \end{bmatrix} \quad (10) \quad B = \begin{bmatrix} [\frac{1}{2}(x^{(0)}(1) + x^{(0)}(2))]^2 \\ [\frac{1}{2}(x^{(0)}(2) + x^{(0)}(3))]^2 \\ \vdots \\ [\frac{1}{2}(x^{(0)}(n-1) + x^{(0)}(n))]^2 \end{bmatrix} \quad (11)$$

$$Y = [x^{(0)}(2) - x^{(0)}(1), x^{(0)}(3) - x^{(0)}(2), \ldots, x^{(0)}(n) - x^{(0)}(n-1)]^T \qquad (12)$$

The resolution of (9) is:

$$\hat{x}^{(0)}(k+1) = \frac{ax^{(0)}(1)}{ux^{(0)}(1) + (a - ux^{(0)}(1))\exp(ak)} \qquad (13)$$
$$(k = 0,1,2,\ldots,n)$$

2-3- The DGM (2, 1) model

The DGM (2, 1) model [11] is a single sequence second-order linear dynamic model and is fitted by differential equations.

Assume an original series to be $X^{(0)}$,

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

A new sequence $X^{(1)}$ is generated by the accumulated generating operation (AGO).

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

Where

$$x^{(1)}(k) = \sum_{j=1}^{k} x^{(0)}(j) \ (k = 1,2,\ldots,n)$$

setting up a second-order differential equation.

$$\frac{d^2 X^{(1)}}{dt^2} + a\frac{dX^{(1)}}{dt} = u \qquad (14)$$

Where

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

$$Y = \begin{bmatrix} (x^{(0)}(2) - x^{(0)}(1)) \\ (x^{(0)}(3) - x^{(0)}(2)) \\ \vdots \\ x^{(0)}(n) - x^{(0)}(n-1)) \end{bmatrix} \quad (16) \qquad B = \begin{bmatrix} -x^{(0)}(2) & 1 \\ -x^{(0)}(3) & 1 \\ \vdots & \vdots \\ -x^{(0)}(n) & 1 \end{bmatrix} \quad (15)$$

According to (14), we have

$$\hat{x}^{(1)}(k+1) = (\frac{u}{a^2} - \frac{x^{(0)}(1)}{a})e^{-ak} + \frac{u}{a}(k+1) + (x^{(0)}(1) - \frac{u}{a})(\frac{1+a}{a}) \qquad (17)$$

The prediction values of original sequence can be obtained by applying inverse AGO to $\hat{x}^{(1)}$. Namely

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \ (k = 0,1,2,\ldots,n)$$

3. CASE STUDY

In this section, the GM(1, 1),the Verhulst model and the DGM(2,1) are used for comparison. The road traffic accident data in Fars province from 2002 to 2010 is adopted to demonstrate the effectiveness and practicability of these models. The road traffic accident data in 2002-2007 is employed to set up the three grey prediction models and the traffic accident data from 2008 to 2010 is used as data set to compare the three models accuracy.

The evaluation criterion [11, 10] is the mean relative percentage error (MRPE), which measures the percent of prediction accuracy.

$$MRPE=\frac{1}{n}\sum_{k=1}^{n}[|x^{(0)}(K)-x^{(0)}(K)|/x^{(0)}(K)]$$

The real and forecasted values are shown in Table 1 to compare the three model accuracy and relative error. The corresponding calculated results (the mean error in the different stage) are shown in Table 2.

Table 1 demonstrates that the GM(1,1) prediction model is smaller than the others by comparing the relative error. From Table 2, it can be seen that the MRPE of the GM(1,1) model, the Verhulst model and the DGM(2,1) from 2008 to 2010 are 6.6%, 12.8% and 8.3% ,respectively. The effectiveness and accuracy of GM(1,1) model is higher than the Verhulst model and the DGM(2,1) model.

	Year	Real Value	**GM(1,1)**		**Verhulst**		**DGM(2,1)**	
			Model value	Error R (%)	Model value	Error R (%)	Model value	Error R (%)
Model Building stage	2002	4472	4472	0	5711	-27.7	4472	0
	2003	5561	6097	9.6	6928	-24.5	5137	7.6
	2004	7177	6846	4.6	8007	-11.5	6298	12.2
	2005	8364	7686	8.1	8883	-6.2	7389	11.6
	2006	8382	8630	-2.9	9545	-13.8	8412	-0.3
	2007	9632	9689	-0.5	10017	-3.9	9374	2.6
Ex-post building stage	2008	11368	10879	4.3	10341	9.03	10276	9.7
	2009	12695	12214	3.8	10557	16.8	11123	12.4
	2010	12263	13715	-11.8	10698	10.2	11919	2.8

Table1, Model Values and Prediction Error of the Traffic Accident in Fars province

Stage	GM(1,1)	Verhulst	DGM(2,1)
	MRPE (%)	MRPE (%)	MRPE (%)
2002-2007	4.28	14.6	5.7
2008-2010	6.6	12.8	8.3

Table 2.Error Analytical Results for the Different Prediction Model

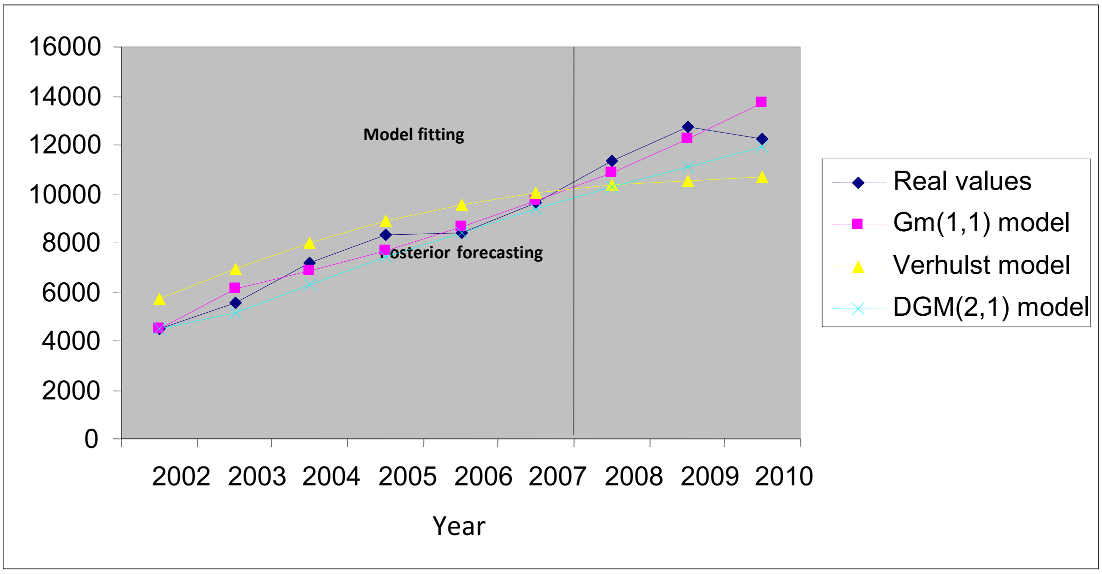

Figure 1, Real values and models values for road traffic accident in Fars province from 2002-2010.

Fig.1 shows that the GM(1,1) model and the DGM(2,1) model have the better forecasting precision in 2008-2010, but the GM(1,1) prediction model seems to obtain the lowest post- forecasting errors and it is more suitable to make a short-term prediction, so the GM(1,1) model uses to predict road traffic accident for 2011 and 2012 in Fars province. The result of forecasting:

Year	2007	2008	2009	2010	2011	2010
Real values	9632	11368	126995	12263		
Model values	9632	11637	12063	12505	12963	13438

Table3. Forecasting values for road traffic accident in Fars province for 2011& 2012

4. CONCLUSIONS

In this paper, we compare the accuracy of the grey forecasting models to predict road traffic accident in Fars province. The grey system theory could deal with the problems with incomplete or unknown information and also the small sample, so this paper uses it.

The result show that the accuracy of GM(1,1)model in forecast value for 2008 to 2010 is higher than the Verhulst grey model and the DGM(2,1) model. Based on the above analysis, the GM(1,1) model appeals to be intrinsically better because it has merits of both simplicity of application and high forecasting precision, thus we use the GM(1,1) model to predict road traffic accident in Fars province for 2011 and 2012.

ACKNOWLEDGEMENTS

The authors would like to thank the Traffic office and Transport organization, Fars province, for providing the traffic accident data.

REFERENCES

1. Hsu. Ch. Ch and Ch. Y. Chen (2003), "Application of improved grey prediction model for power demand forecasting", Energy Conversion & Management 44, 2241-2249.
2. Julong. D (1989), "Introduction to Grey System Theory", Journal of Grey System, 1-24.
3. Kayacan. E., Kaynak. O. and B. Ulutas (2010). "Grey System Theory- Based models in time Series prediction", Expert Systems with Application37, 1784-1789.
4. Lin. Y. H., Lee. P. Ch and T. P. Chang (2009), "Adoptive and high- precision grey forecasting model", ELSEVIER, 9658-9662.
5. Lu, M (2007). "Analysis and Forecast of Road Traffic Safety Improvement in Netherlands", ARS Traffic & Transport Technology, 1-14.
6. Tongyuan. H. and W. Yue (2007)," Forecasting Model Of Urban Traffic Accidents Based on Grey Model-GM (1, 1)", Institute of Computer, Chongqing Institute of Technology, 438-441.
7. Tseng. P., Lin. F. and Ch. Feng (2003),"The analysis of human and vehicle factors for Taiwan freeway traffic accidents", Journal of the Eastern Asia Society for transportation Studies, Vol.5.
8. Wu. W. Y and Sh. P. Chen (2005), " A Prediction method Using the Grey Model GMc(1,n) Combined With the Grey Relational Analysis: a case study on internet access population forecast", *ELSEVIER,* 198-217.
9. Zhan Li. M and S. Jin-Hua (2011), "Application of Grey- Markov Model in Forecasting Fire Accidents", *Precedia Engineering11*, 314-318.
10. Zhou. P., Ang. B.W and K. L. Poh (2006), "A trigonometric Grey Prediction Approach to Forecasting Electricity Demand", Energy Conversion & Management31, 2839-2847.
11. Zhu. X (2010), "Application of Composite Grey BP Neural Network Forecasting Model to Motor Vehicle Fatality Risk", Second International Conference on Computer Modeling and Simulation, 236-240.