

文章编号:1671-8879(2009)04-0087-04

# 基于可变参数的道路交通事故动态灰色预测模型

马壮林, 邵春福, 李霞, 郝合瑞

(北京交通大学 交通运输学院, 北京 100044)

**摘要:**以中国 1990~2006 年万车死亡率为研究对象,在固定参数动态灰色预测模型基础上,根据人口数量、机动车数量和公路里程优化参数  $P$  值,建立了可变参数动态灰色预测模型;分别运用固定参数动态灰色预测模型、可变参数动态灰色预测模型和基于三点平均技术的可变参数动态灰色预测模型进行预测,得到的平均相对误差分别为 6.97%、5.67%和 4.61%。研究表明:可变参数动态灰色预测模型优于固定参数动态灰色预测模型;基于三点平均技术的可变参数动态灰色预测模型可以进一步平滑预测结果,增加预测结果的可靠性。

**关键词:**交通工程;交通安全;死亡率;可变参数;动态灰色预测模型

中图分类号:U491.3

文献标志码:A

## Forecasting grey model of traffic accidents based on variable parameter value rolling

MA Zhuang-lin, SHAO Chun-fu, LI Xia, HAO He-rui

(School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China)

**Abstract:** Taking the death rates from 1990 to 2006 in Chinese accidents as basic datum, a variable parameter value rolling grey forecasting model (VRGM) is proposed based on the fixed parameter value rolling grey forecasting model (FRGM), which has a constant parameter and equals 0.5. It is apparent that if the parameter value is decided in consideration of the factors, such as the number of population, the number of vehicles and the length of highway, which influence the forecasting objective, better result will be achieved. According to death rates per 10 thousand vehicles in China, a VRGM (1,1) model is constructed. The average residual error of FRGM (1,1) model, VRGM (1,1) model and VRGM (1,1) model plus 3-points average are 6.97%, 5.67% and 4.61% respectively. The research suggests that VRGM (1,1) model is better than FRGM (1,1) model, and VRGM (1,1) model plus 3-points average can further smooth the prediction results and increase the reliability. 3 tabs, 15 refs.

**Key words:** traffic engineering; traffic safety; death rate; variable parameter value; rolling grey forecasting model

## 0 引言

道路交通事故预测是指以某一地区或一条道路

为研究范围,通过调查等手段获得研究范围内与道路交通事故有关的信息(如历史事故、道路设施、人口、车辆、道路线形和天气等信息),依据这些信息,

收稿日期:2008-08-20

基金项目:国家科技支撑计划项目(2007BAK35B06)

作者简介:马壮林(1980-),男,辽宁鞍山人,工学博士研究生,E-mail:06114185@bjtu.edu.cn.

利用现代数学,采用定量与定性相结合的方法,推断未来一段时间研究范围内的道路交通事故发生状况,为掌握和控制事故发展趋势提供参考<sup>[1]</sup>。国内外对于道路交通事故预测方法进行了多方面的研究,提出了一些较为实用的道路交通事故预测方法,如基于泊松分布模型的预测方法、基于负二项分布的预测方法、多因素时间序列预测方法和基于神经网络的预测方法<sup>[2-5]</sup>等。虽然这些方法都能从不同角度对道路交通事故进行预测,但对于具有非线性、随机性、动态性及不确定性的道路交通系统,这些方法表现出明显不足。且传统的事统计方法普遍存在着“小样本、长周期、大区域、低信度”的缺陷<sup>[6-9]</sup>,因此直接影响到交通事故分析的准确性和分析精度。而灰色系统理论是一种研究数据较少、信息不足问题的新方法,并依据信息覆盖,通过序列算子的作用,探索事物运动的现实规律<sup>[10-12]</sup>。

道路交通事故分析指标主要有绝对指标和相对指标。目前,较为常用的绝对指标有事故起数、死亡人数、受伤人数和经济损失;相对指标有10万人口死亡率、万车死亡率和亿车公里事故率等。这些指标分别从不同的角度表征一个国家或地区道路交通安全的整体状况。在绝对指标中,交通事故死亡人数是最精确的指标,但是不同时期的人口数量、机动车数量和道路里程与交通事故死亡人数密切相关,因而采用死亡率这类相对指标更合适。与10万人口死亡率相比,万车死亡率能更真实和客观地反映一个区域的交通安全水平<sup>[13]</sup>。为此,本文运用可变参数动态灰色预测模型,对中国道路交通事故死亡率进行预测。

## 1 可变参数动态灰色预测模型

动态GM(1,1)模型是应用动态数据处理技术进行预测的一种方法。在预测模型中,动态数据处理技术将每个新得到的信息送入数据列中,同时去除一个最陈旧的数据,使模型具有一种新陈代谢的功能,可以使模型更加符合实际情况<sup>[14-15]</sup>。

实际应用中,该方法有时可得到较好的结果,有时的结果却不理想。这是由于一般的GM(1,1)模型,在构造Z序列时往往采用固定参数值的方法,即参数P一般取值为0.5(Z为GM(1,1)模型中的序列;P为参数)。但是,如果P值在0.1~0.9区间取值,可得到的误差比P取0.5时的误差更小。因此,可变参数动态灰色预测模型可以使预测结果更为精确。

由于Z序列的构造过程就是对原始数据隐含规律的灰处理过程,并利用它形成灰色微分方程并计算系数,因此,可以通过分析外界因素的影响来更准确地构造Z序列。本文认为,可通过建立若干个影响因素与历史最优P值之间的函数关系,获得预测的P值,然后再利用此P值进行动态灰色预测。

## 2 可变参数动态灰色预测模型的建立

设 $X^{(0)}$ 为原始数据序列,则

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

式中: $x^{(0)}(k) \geq 0 (k = 1, 2, \dots, n)$ ,为数列中的值。

每次截取r个数据进行预测,那么用于预测的原始序列为

$$X^{(0)}(i; K) = \{x^{(0)}(i), x^{(0)}(i+1), \dots, x^{(0)}(K)\}$$

其中, $r = K - i + 1$ ,且 $K \leq n$ 。

对 $X^{(0)}(i; K)$ 作一次累加,生成数据序列 $X^{(1)}(i; K)$ 为

$$X^{(1)}(i; K) = \{x^{(1)}(i), x^{(1)}(i+1), \dots, x^{(1)}(K)\} \quad (1)$$

根据 $X^{(1)}(i; K)$ 生成数据序列 $Z^{(1)}(i; K)$ ,则

$$Z^{(1)}(i; K) = \{z^{(1)}(i), z^{(1)}(i+1), \dots, z^{(1)}(K)\}$$

$$z^{(1)}(j+1) = Px^{(1)}(j) + (1-P)x^{(1)}(j+1) \quad (2)$$

其中, $i \leq j \leq K-1$ ,且式(2)中参数P根据分析影响因素的影响程度而获得。

灰色微分方程 $X^{(0)}(i; K) + aZ^{(1)}(i; K) = b$ 的最小二乘法估计参数a、b满足

$$\begin{bmatrix} a \\ b \end{bmatrix} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y}_N \quad (3)$$

$$\text{式中: } \mathbf{B} = \begin{bmatrix} z^{(1)}(i+1) & 1 \\ -z^{(1)}(i+2) & 1 \\ \vdots & \vdots \\ -z^{(1)}(K) & 1 \end{bmatrix};$$

$$\mathbf{Y}_N = \begin{bmatrix} x^{(0)}(i+1) \\ x^{(0)}(i+2) \\ \vdots \\ x^{(0)}(K) \end{bmatrix}。$$

$X^{(1)}(i; K)$ 数据序列的灰色预测模型为

$$\hat{x}^{(1)}(i+t) = (x^{(0)}(i) - \frac{b}{a})e^{-at} + \frac{b}{a} \quad (4)$$

式中: $0 \leq t \leq r-1$ ;  $\hat{x}^{(1)}(i+t)$ 为预测值。

运用累减还原式(5),进行累减还原

$$\hat{x}^{(0)}(i+t+1) = \hat{x}^{(1)}(i+t+1) - \hat{x}^{(1)}(i+t) \quad (5)$$

得到原始数据序列 $X^{(0)}(i; K)$ 的预测值序列 $\hat{X}^{(0)}(i; K)$ 为

$$\hat{X}^{(0)}(i;K) = \{\hat{x}^{(0)}(i), \hat{x}^{(0)}(i+1), \dots, \hat{x}^{(0)}(K)\}$$

其中,  $\hat{x}^{(0)}(i) = x^{(0)}(i)$ 。

根据式(5),可以计算得到  $\hat{x}^{(0)}(K+1)$ ,而  $\hat{x}^{(0)}(K+1)$  就是可变动态灰色预测模型对预测对象的预测值。

从  $i=1$  开始至  $i=n-4$ ,重复操作上述步骤,可以得到  $(n-r)$  个  $\hat{x}^{(0)}(K+1)$ ,形成序列  $\hat{x}^{(0)}(r+1, n) = \{\hat{x}^{(0)}(r+1), \hat{x}^{(0)}(r+2), \dots, \hat{x}^{(0)}(n)\}$ ,该序列为可变动态灰色预测模型所得预测值。

### 3 模型精度的检验

本文采用平均相对误差来检验模型的优劣。

设残差序列为

$$\varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) \quad (6)$$

则相对误差为

$$e(k) = \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\% \quad (7)$$

其中,  $k \leq n$ 。

因此,平均相对误差  $\Delta$  为

$$\Delta = \frac{1}{n-4} \sum_{k=5}^n e(k) \quad (8)$$

### 4 实证研究

本文以中国1990~2006年的万车死亡率为研究对象,运用可变参数动态灰色预测模型(VRGM)进行预测,然后比较固定参数动态灰色预测模型(FRGM)、可变参数动态灰色预测模型(VRGM)和基于三点平均技术的可变参数动态灰色预测模型(VRGM)的预测精度。

由于道路交通系统由人、车、路组成,交通事故的发生是3个影响因素单独作用、两两作用或三者相互作用的结果。因此,本文认为,中国道路交通事故死亡率受到人口数量、汽车数量和公路里程的影响。根据1990~2006年中国道路交通事故万车死亡率,在采用4期数据滚动的情况下,得到1994~2006年最小误差的最优参数值,如表1所示。

根据这些数据,在最优参数  $P$  值与人口增长率  $x$ 、汽车增长率  $y$  和公路里程增长率  $z$  之间进行回归,得到回归模型( $R^2$  为相关系数),即

$$P = -1.7289 + 92.4694x + 11.3812y + 0.0966z \quad R^2 = 0.642 \quad (9)$$

将1994~2006年中国的人口增长率  $x$ 、汽车增长率  $y$  和道路里程增长率  $z$  重新代入式(9),得到

表1 最小误差的最优参数值

年份	人口增长率/%	汽车增长率/%	公路里程增长率/%	P 值
1994	1.12	15.21	3.17	0.9
1995	1.06	10.41	3.51	0.1
1996	1.05	5.78	2.49	0.1
1997	1.01	10.82	3.43	0.9
1998	0.92	8.22	4.25	0.1
1999	0.82	10.13	5.73	0.1
2000	0.76	10.73	3.77	0.1
2001	0.70	12.00	21.05	0.1
2002	0.65	13.94	3.96	0.1
2003	0.60	16.06	2.53	0.9
2004	0.59	13.04	3.36	0.2
2005	0.59	17.30	78.82	0.9
2006	0.53	17.02	3.34	0.9

$P$  值,分别为1.03、0.43、-0.11、0.43、0.05、0.18、0.19、0.29、0.45、0.64、0.29、0.84和0.68。因此,FRGM(1,1)、VRGM(1,1)和基于三点平均技术的VRGM(1,1)的预测结果及误差比较如表2所示。

表2 预测模型的预测结果及误差比较

年份	实际值	FRGM(1,1)		VRGM(1,1)		基于三点平均技术的VRGM(1,1)	
		预测值	相对误差/%	预测值	相对误差/%	预测值	相对误差/%
1994	70.45	73.83	4.80	71.98	2.17	71.98	2.17
1995	68.74	64.28	6.50	64.56	6.09	67.40	1.95
1996	66.95	63.67	4.91	65.67	1.92	65.20	2.61
1997	60.59	65.28	7.75	65.39	7.93	63.37	4.60
1998	59.17	57.76	2.39	59.07	0.18	60.03	1.45
1999	57.49	54.74	4.78	55.64	3.22	57.06	0.75
2000	58.33	56.05	3.91	56.47	3.20	56.56	3.04
2001	58.78	57.49	2.20	57.58	2.05	57.84	1.61
2002	53.27	59.50	11.69	59.47	11.62	56.25	5.59
2003	43.8	51.99	18.71	51.71	18.06	50.19	14.59
2004	39.75	38.79	2.42	39.39	0.90	41.23	3.72
2005	31.25	33.41	6.92	32.59	4.29	33.03	5.71
2006	24.19	27.49	13.60	27.12	12.09	27.12	12.09
平均值			6.97		5.67		4.61

注:实际值、预测值的单位为“死亡人数/万车”。

从表2可以看出,FRGM(1,1)模型的平均误差、最大误差和最小误差分别为6.97%、18.71%和2.2%;VRGM(1,1)模型的平均误差、最大误差和最小误差分别为5.67%、18.06%和0.18%;基于三点平均技术的VRGM(1,1)模型的平均误差、最大误差和最小误差分别为4.61%、14.59%和0.75%。

预测模型的误差对比分析如表3所示。

表3 预测模型的误差对比分析

年份	FRGM (1,1)	VRGM(1,1)		基于三点平均技术 的 VRGM(1,1)	
	相对误差/ %	相对误差/ %	误差减少 程度/%	相对误差/ %	误差减少 程度/%
1994	4.80	2.17	54.74	2.17	54.74
1995	6.50	6.09	6.35	1.95	69.96
1996	4.91	1.92	60.79	2.61	46.77
1997	7.75	7.93	-2.26	4.60	40.67
1998	2.39	0.18	92.42	1.45	39.24
1999	4.78	3.22	32.53	0.75	84.22
2000	3.91	3.20	18.25	3.04	22.35
2001	2.20	2.05	6.68	1.61	26.87
2002	11.69	11.62	0.55	5.59	52.18
2003	18.71	18.06	3.45	14.59	22.02
2004	2.42	0.90	62.67	3.72	-53.96
2005	6.92	4.29	38.00	5.71	17.50
2006	13.60	12.09	11.12	12.09	11.12

从表3可以看出,只有在1997年,VRGM(1,1)模型的误差大于FRGM(1,1)模型,相对于FRGM(1,1)模型,VRGM(1,1)模型的误差最大减少程度达到92.42%,误差年平均减少程度达到29.65%;同理,只有在2004年,基于三点平均技术的VRGM(1,1)模型的误差大于FRGM(1,1)模型,相对于FRGM(1,1)模型,基于三点平均技术的VRGM(1,1)模型的误差最大减少程度达到84.22%,误差年平均减少程度达到33.36%。由此可见,可变参数动态灰色预测模型优于固定参数动态灰色预测模型,基于三点平均技术的可变参数动态灰色预测模型可以进一步平滑预测值,具有更好的可靠性。

## 5 结 语

(1)VRGM(1,1)模型不仅考虑到数据的时间性,还考虑到其他影响因素的作用;通过对中国道路交通事故死亡率的研究,说明该模型具有较好的预测精度。

(2)基于三点平均技术的VRGM(1,1)模型,考虑了当前状态的前一个状态和后一个状态的预测值,使其预测的结果具有更好的平滑性、可靠性。

(3)采用多元线性回归模型建立的函数关系,应注意影响因素与最优参数值之间函数关系的选择;实际上不同的预测对象或不同的影响因素,都可能产生其他函数关系,特别是随着时间的推移,在每一期新数据的加入后,都需要建立新的函数关系。

## 参考文献:

### References:

- [1] 裴玉龙,王 炜. 道路交通事故成因及预测对策[M]. 北京:科学出版社,2004.
- [2] 罗石贵,周 伟. 路段交通冲突的调查技术[J]. 长安大学学报:自然科学版,2003,23(1):71-75.  
LUO Shi-gui, ZHOU Wei. Survey way of road-traffic-conflict technique[J]. Journal of Chang'an University: Natural Science Edition, 2003, 23(1): 71-75.
- [3] Poch M, Mannering F. Negative binomial analysis of intersection accident frequencies[J]. Journal of Transportation Engineering, 1996, 122(2): 105-113.
- [4] 季彦婕,王 炜,邓 卫. 道路交通事故多因素时间序列宏观预测模型[J]. 武汉理工大学学报:交通科学与工程版,2006,30(3):433-436.  
JI Yan-jie, WANG Wei, DENG Wei. Multi-factor time series model for macro-level road accident prediction [J]. Journal of Wuhan University of Technology: Transportation Science & Engineering, 2006, 30(3): 433-436.
- [5] 李 娟,邵春福. 基于BP神经网络的交通事故预测模型[J]. 交通与计算机,2006,24(2):34-37.  
LI Juan, SHAO Chun-fu. Traffic accident prediction based on BPNN[J]. Computer and Communications, 2006, 24(2): 34-37.
- [6] 周维新. 交通事故灰色预测模型的研究[J]. 西安公路交通大学学报,2000,20(2):73-75.  
ZHOU Wei-xin. Study on the forecast model of traffic accidents with grey theory[J]. Journal of Xi'an Highway University, 2000, 20(2): 73-75.
- [7] 李相勇,张 南,蒋葛夫. 道路交通事故灰色马尔可夫预测模型[J]. 公路交通科技,2003,20(4):98-100,104.  
LI Xiang-yong, ZHANG Nan, JIANG Ge-fu. Grey-markov for model forecasting road accidents[J]. Journal of Highway and Transportation Research and Development, 2003, 20(4): 98-100, 104.
- [8] 王福建,李铁强,俞传正. 道路交通事故灰色 Verhulst 预测模型[J]. 交通运输工程学报,2006,6(1):122-126.  
WANG Fu-jian, LI Tie-qiang, YU Chuan-zheng. Grey Verhulst predictive model of road traffic accidents [J]. Journal of Traffic and Transportation Engineering, 2006, 6(1): 122-126.
- [9] 周 伟,罗石贵. 路段交通事故的多发点的冲突判定方法[J]. 中国公路学报,2000,13(1):81-86.