

试验研究

利用灰色马尔可夫模型预测煤矿安全事故

张林华, 刘玉洲

(河南理工大学 安全科学与工程学院, 河南 焦作 454003)

摘 要: 在煤矿事故死亡人数灰色预测的基础上引入马尔可夫链预测理论, 建立了煤矿事故灰色马尔可夫预测模型, 该模型兼有灰色预测和马尔可夫预测的优点, 不仅提高了波动性较大的随机变量的预测精度, 同时还拓宽了灰色预测的应用范围。实例证明灰色马尔可夫模型预测精度明显高于灰色 GM (1, 1) 残差修正模型预测精度, 从而说明灰色马尔可夫模型预测煤矿事故死亡人数的发展规律是可行有效的。

关键词: 灰色预测; 马尔可夫链预测; 煤矿安全

中图分类号: X913.4; X936 **文献标志码:** A **文章编号:** 0253-2336 (2006) 11-0026-05

Application of Grey Markov Mode to predict mine safety accident

ZHANG Lin-hua, LIU Yu-zhou

(School of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo 454003, China)

Abstract: Base on the grey prediction of the fatality number in a mine accident, the Markov chains prediction theory was introduced to build a grey Markov prediction model of mine accident. The model with a grey prediction function and a Markov prediction function can not only improve the high change random prediction accuracy and also develop the application range of the grey prediction. The practices showed that the prediction accuracy of the grey Markov model was higher than the prediction accuracy of the grey residue revised model. Therefore the grey Markov model to predict the development law of fatality number in a mine accident is feasible and efficient.

Key words: grey prediction; Markov chain prediction; mine safety

由于事故的发生与它的过去和现状紧密相连, 从而就有可能通过对事故的现状和历史进行综合分析, 推测它未来的发展趋势, 为事故的预防提供可靠的信息。煤矿安全事故预测对于探究煤矿安全事故的发生规律, 分析现有的煤矿生产条件下安全事故的未来发展趋势以及煤矿生产安全控制等具有重要的意义。本文基于煤矿安全事故灰色预测方法, 引入马尔可夫链预测理论, 建立了煤矿安全事故的灰色马尔可夫预测模型。

传统的灰色系统理论在建立模型时, 不需要大量的数据, 它把观测数据序列看作随时间变化的灰色过程, 通过累加生成挖掘出系统潜藏的有序的指数规律, 从而建立相应的预报模型, 它能够反映预测对象的总体发展趋势。灰色建模的实质是以指数型曲线去拟合原始数据, 其预测结果的几何图形是

一条较为平滑的曲线, 因而对于波动性较大的数据列拟合较差, 预测精度较低。虽然灰色模型本身也有一些提高预测精度的方法, 如残差辨识法、提高模型阶数等, 但对于波动性大的非平稳数列的预测, 结果不理想, 甚至可能增大误差, 而马尔可夫预测理论则为问题的解决提供了可能。根据马尔可夫链理论, 系统将来所处状态只与现在系统状态有关, 而与系统过去状态无关, 即无后效性。马尔可夫预测是根据系统状态之间的转移概率来预测系统的未来发展, 转移概率反映了各种随机因素的影响程度, 反映了各状态之间的内在规律性, 马尔可夫链适合于随机波动性较大问题的预测, 能够揭示出系统受各种复杂因素影响的随机性。由于煤矿安全事故的发生受多种因素的影响, 且具有随机性, 所以利用灰色预测和马尔可夫预测各自特点建立煤矿安全事故的灰色马尔可夫预测模型, 用灰色预测来揭示煤矿安全事故时序变化的总体趋势, 用马尔可

夫预测来确定状态的转移规律,通过两种模型的有机结合达到科学预测煤矿安全事故的目的。

1 灰色马尔可夫建模过程

1.1 建立灰色 GM (1, 1) 模型

给定原始数据列 $X^{(0)}(t) = (X^{(0)}(1), X^{(0)}(2), X^{(0)}(3), \dots, X^{(0)}(n))(t = 1, 2, \dots, n)$, 运用灰色系统理论, 建立 GM(1, 1) 模型为

$$\begin{aligned} X^{(1)}(t+1) &= [X^{(0)}(1) - b/a]e^{-at} + b/a \quad (1) \\ X^{(0)}(t+1) &= X^{(1)}(t+1) - X^{(1)}(t) = \\ &= (1 - e^a)[X^{(1)}(1) - b/a]e^{-at} \quad (2) \end{aligned}$$

式中, a 为常数, b 为对系统的定常输入。

$X^{(0)}(t+1)$ 曲线反映了实际观测数据的变化总趋势, 但是, 由于系统的各种因素会导致系统的随机变化, 而该模型未曾考虑随机性, 所以单纯利用它来预测还不完善。

1.2 状态划分

对于一个具有马尔可夫链特点的非平稳随机序列, 将数据序列划分成为若干状态, 用 $E_1, E_2, \dots, E_n$ 来表示, 状态划分的多少可以根据样本数量的多少来定, 状态的数量多, 预测精度会相应提高。任一状态可以表示为

$$\begin{aligned} E_i &= [E_{1i}, E_{2i}] \\ E_{1i} &= X^{(0)}(t+1) + a_i \\ E_{2i} &= X^{(0)}(t+1) + b_i \end{aligned}$$

其中 $i=1, 2, \dots, n$; a_i 和 b_i 的值可根据预测对象不同、资料的多少, 结合经验而赋值, 目前对于数据序列的状态如何划分, 还没有统一的标准, 本文根据灰色预测的误差情况来划分事故状态。

1.3 转移概率的计算和状态转移矩阵的建立

根据对数据序列的状态划分, 称数据序列由状态 E_i 经过 m 步转移到 E_j 的概率为 m 步转移概率, 记为 $p_{ij}^{(k)}$ 。则有

$$p_{ij}^{(k)} = \frac{m_{ij}^{(k)}}{M_i} \quad (3)$$

式中, $m_{ij}^{(k)}$ 为状态 E_i 经过 k 步的转移到状态 E_j 的次数; M_i 为状态 E_i 出现的次数。

由于数据序列的最后状态转向不明确, 故计算 M_i 的时候要去掉数据列中最末的 k 个数据。根据状态转移的概率, 可以写出状态 k 步转移概率矩阵:

$$P^{(k)} = \begin{bmatrix} p_{11}^{(k)} & p_{12}^{(k)} & \cdots & p_{1n}^{(k)} \\ p_{21}^{(k)} & p_{22}^{(k)} & \cdots & p_{2n}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}^{(k)} & p_{n2}^{(k)} & \cdots & p_{nn}^{(k)} \end{bmatrix} \quad (4)$$

该矩阵称为马尔可夫链的 k 步转移矩阵, 其中

$$\sum_{j=1}^n p_{ij}^{(k)} = 1。$$

1.4 编制预测表格

系统各种状态转移的统计规律在状态转移概率矩阵 $P^{(k)}$ 中得到了反映, 可以根据状态转移概率矩阵 $P^{(k)}$ 来预测系统未来的发展变化。编制预测表格时选取离预测时刻最近的 j 个时刻, 按距离预测时刻的远近, 转移步数分别定为 $1, 2, \dots, j$ 步, 在转移步数所对应的转移矩阵中, 取起始状态对应的行向量, 从而组成新的概率矩阵, 对新的概率矩阵将其列向量求和, 其和最大的列向量的状态即为待预测状态。

1.5 计算预测值

根据上述灰色模型和马尔可夫状态转移概率矩阵的建立, 就可以计算预测值了。确定了预测对象的状态转移矩阵, 就确定了根据 GM (1, 1) 模型拟合的灰区间 $E_i = [E_{1i}, E_{2i}]$, 利用编制出的预测表格, 确定出预测对象最有可能的状态转移。

对于最终的预测值, 取灰区间的中间值, 根据前文状态划分的描述, $E_{1i} = X^{(0)}(t+1) + a_i$, $E_{2i} = X^{(0)}(t+1) + b_i$, 即有:

$$X^{(0)}(t+1) = \frac{E_{1i} + E_{2i}}{2} = X^{(0)}(t+1) + \frac{a_i + b_i}{2} \quad (5)$$

根据式 (5) 就可以求得预测对象在 $k+1$ 时刻的值。

2 煤矿安全事故的灰色马尔可夫预测

2.1 煤矿安全事故 GM (1, 1) 模型的建立

本文中, 以历年煤矿安全事故死亡人数为准, 收集了从 1986 ~ 2004 年的全国煤矿年死亡人数 (表 1), 以表 1 中的数据为原始数据列, 建立煤矿安全事故年死亡人数 GM (1, 1) 残差模型。根据 GM (1, 1) 模型的建模方法, 得到的 GM (1, 1) 残差模型为

$$\begin{aligned} X^{(0)}(t+1) &= \\ &\begin{cases} 9\,429.260\,3\exp(-0.020\,27t) & (t < 13) \\ 9\,429.260\,3\exp(-0.020\,27t) - 2\,066.435\,7 \times \\ \quad \exp[-0.404\,454(t-13)] & (t \geq 13) \end{cases} \end{aligned} \quad (6)$$

由式 (6) 求得历年的煤矿死亡人数拟合值见表 1。

表1 煤矿安全事故年死亡人数

年份	死亡人数	GM(1,1)残差预测值	相对误差/%	状态
1986	7 937	7 937	0	E_3
1987	7 977	9 240	-15.83	E_1
1988	7 980	9 055	-13.47	E_1
1989	8 370	8 873	-6.012	E_2
1990	8 725	8 695	0.34	E_4
1991	8 127	8 520	-4.84	E_3
1992	7 920	8 349	-5.42	E_2
1993	8 620	8 182	5.08	E_5
1994	9 384	8 018	14.56	E_6
1995	9 659	7 857	18.66	E_6
1996	9 974	7 699	22.81	E_7
1997	7 083	7 545	-6.52	E_2
1998	7 400	7 393	0.09	E_4
1999	6 078	7 245	-19.2	E_1
2000	5 798	5 721	1.34	E_4
2001	5 670	6 037	-6.47	E_2
2002	6 528	6 203	4.97	E_4
2003	6 434	6 271	2.54	E_4
2004	6 027	6 273	-4.08	E_3

2.2 事故状态的划分

根据煤矿死亡人数的实际数据,以及残差 GM(1,1)模型的预测误差情况,将死亡数据序列划分为7个状态。

$$E_1: E_{11} = X^{(0)}(t+1) - 0.20y$$

$$E_{12} = X^{(0)}(t+1) - 0.1y$$

$$E_2: E_{21} = X^{(0)}(t+1) - 0.1y$$

$$E_{22} = X^{(0)}(t+1) - 0.05y$$

$$E_3: E_{31} = X^{(0)}(t+1) - 0.05y$$

$$E_{32} = X^{(0)}(t+1)$$

$$E_4: E_{41} = X^{(0)}(t+1)$$

$$E_{42} = X^{(0)}(t+1) + 0.05y$$

$$E_5: E_{51} = X^{(0)}(t+1) + 0.05y$$

$$E_{52} = X^{(0)}(t+1) + 0.1y$$

$$E_6: E_{61} = X^{(0)}(t+1) + 0.1y$$

$$E_{62} = X^{(0)}(t+1) + 0.20y$$

$$E_7: E_{71} = X^{(0)}(t+1) + 0.2y$$

$$E_{72} = X^{(0)}(t+1) + 0.25y$$

其中 y 为历年死亡人数的平均值, $y = 7\,667.95$ 。根据状态划分情况,历年的状态划分见表1。

2.3 构造状态转移矩阵

根据马尔可夫链预测理论,依据前述已划分出

的7个状态,构造出如下的状态转移概率矩阵:

$$P^{(1)} = \begin{bmatrix} 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3/5 & 1/5 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 1/5 & 1/5 & 2/5 & 1/5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$P^{(2)} = \begin{bmatrix} 0 & 2/3 & 0 & 1/3 & 0 & 0 & 0 \\ 1/4 & 0 & 1/4 & 1/4 & 0 & 1/4 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$P^{(3)} = \begin{bmatrix} 0 & 2/3 & 0 & 1/3 & 0 & 0 & 0 \\ 1/4 & 0 & 1/4 & 1/4 & 0 & 1/4 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$P^{(4)} = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/3 & 1/3 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

2.4 编制预测表格

根据马尔可夫预测理论,对2005年煤矿死亡人数进行试报,选取距离预测时刻最近的4个时段,其转移步数分别定为1,2,3,4步。在转移步数所对应的转移矩阵中,选择起始状态所对应的行向量,得到新的概率矩阵,见表2。

表2 状态预测计算

状态	转移步数	E_1	E_2	E_3	E_4	E_5	E_6	E_7
E_2	4	0	1/3	0	0	1/3	1/3	0
E_4	3	0	1/4	1/4	1/2	0	0	0
E_4	2	0	1/4	1/4	1/2	0	0	0
E_3	1	1/2	1/2	0	0	0	0	0
合计	—	1/2	4/3	1/2	1	1/3	1/3	0

根据表2中的状态转移概率列向量求和结果,可以发现2005年的状态处于 E_2 状态的可能性最大,表示2005年煤矿死亡人数拟合值与实际值的误差最可能处于 E_2 状态,对2005年的预测煤矿事故死亡人数为

$$X^{(0)}(t+1) = \frac{E_{11} + E_{21}}{2} = X^{(0)}(t+1) + \frac{a_i + b_i}{2} =$$

$$9\,429.260\,3\exp(-0.020\,27t) - 2\,066.435\,7 \times \exp[-0.404\,454(t-13)] = 5\,658\text{人}$$

而实际2005年煤矿事故死亡人数为5 986人,灰色马尔可夫预测相对误差绝对值为5.5%,这与残差GM(1,1)模型预测得到的相对误差4.12%相比较,相差不大。同理对2006年进行灰色马尔可夫预测,可以得到2006年煤矿的死亡人数为5 590人,这与国家对2006年煤矿死亡人数计划控制在5 777人相比较,相对误差绝对值为3.24%,而用GM(1,1)残差模型预测的值为6 165人,相对误差绝对值为6.71%,考虑到人们对煤矿安全生产越来越重视和国家对煤矿安全投入的加大,安全科学技术的不断进步,灰色马尔可夫模型对2006年的预测值较为符合实际情况。根据公式(5)和(6),对1986~2004年的煤矿死亡人数进行回测比较,结果见表3。

根据表3可以看出,灰色马尔可夫模型的预测精度在整体上要优于用单纯的灰色建模预测精度,特别是对于数据波动性较大的数据,灰色马尔可夫的优越性更能体现出来,表3中1994~1996年的数据用GM(1,1)残差模型预测的结果相对误差较大,相对误差绝对值最大达22.8%,而用马尔可夫模型预测的结果,相对误差明显减小,相对误差绝对值最大为6.75%,精度明显得到了提高。通过灰色马尔可夫回测得到的结果,表明灰色马尔可夫预测是可行和可靠的,所预测的结果是与实际较为符合的。

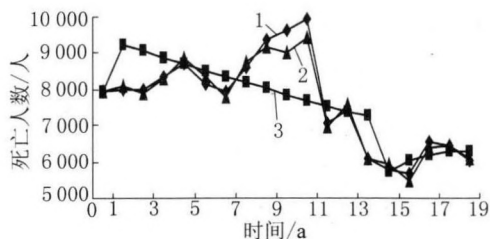
为更加直观地比较出灰色马尔可夫模型预测和GM(1,1)残差模型预测对实际数值的拟合程度差别,绘出灰色马尔可夫模型预测值,GM(1,1)残差模型预测值和实际值的曲线图来比较,如图1所示。

由图1可以看出,GM(1,1)残差模型预测拟合出的曲线反映出煤矿事故实际煤矿死亡人数总体趋势是下降的,但是对死亡人数的随机性要求不能很好地反映;而灰色马尔可夫模型预测值和实际

值的曲线拟合程度很好,它所拟合出的曲线则很好地反映了在一定趋势下的数据的随机性。故利用灰色马尔可夫模型预测煤矿事故死亡人数精度较高,适合用于预测分析。

表3 1986~2004年煤矿事故死亡人数
灰色马尔可夫回测结果

年份	死亡人数	GM(1,1)残差模型预测		灰色马尔可夫模型	
		拟合值	相对误差/%	拟合值	相对误差/%
1986	7 937	7 937	0	7 937	0
1987	7 977	9 240	15.83	8 090	1.415
1988	7 980	9 055	13.47	7 904	-0.95
1989	8 370	8 873	6.009	8 298	-0.86
1990	8 725	8 695	-0.34	8 887	1.852
1991	8 127	8 520	4.841	8 329	2.482
1992	7 920	8 349	5.423	7 774	-1.84
1993	8 620	8 182	-5.08	8 757	1.59
1994	9 384	8 018	-14.6	9 168	-2.3
1995	9 659	7 857	-18.7	9 007	-6.75
1996	9 974	7 699	-22.8	9 424	-5.51
1997	7 083	7 545	6.519	6 970	-1.6
1998	7 400	7 393	-0.09	7 585	2.5
1999	6 078	7 245	19.2	6 095	0.276
2000	5 798	5 721	-1.34	5 912	1.971
2001	5 670	6 037	6.47	5 462	-3.67
2002	6 528	6 203	-4.97	6 395	-2.04
2003	6 434	6 271	-2.54	6 463	0.444
2004	6 027	6 273	4.085	6 081	0.904



1—实际死亡人数; 2—灰色马尔可夫模型预测值;
3—GM(1,1)残差模型预测值

图1 模型预测值拟合曲线比较

3 结 论

灰色马尔可夫模型兼有灰色预测和马尔可夫链预测的优点,既考虑了从时序数列中挖掘数据的演变规律,又考虑了数据的随机响应性,具有严密的科学性。用灰色预测来揭示煤矿事故死亡人数随时

序变化的总体趋势,用马尔可夫链预测来反映随机波动性对事件预测的影响,确定状态的转移规律,通过两种模型的有机结合达到科学预测煤矿事故死亡人数的目的。

灰色马尔可夫模型预测原理简单易行,它是建立在历史数据统计的基础上的,需要对数据列进行状态的划分,但是目前对数据列状态的划分还缺乏统一的标准,有待于进一步的研究。收集到的历史数据越准确,灰色马尔可夫模型预测的精度越高,利用不断增加的监测数据来实时修正预测模型,可以实现高精度的预测。将灰色马尔可夫模型应用于煤矿事故死亡人数预测是合理可行的,通过预测值和最大状态概率可以更加准确地把握煤矿事故死亡人数的总体发展趋势,它比单纯应用灰色 GM (1, 1) 模型预测时效果要好。

参考文献:

- [1] 韩 乔. 中国力争今年煤矿事故死亡人数下降 3.5% [EB/OL]. <http://politics.people.com.cn/GB/1027/4075973.html>, 2006-02-05.
- [2] 林晓言, 陈有孝. 基于灰色-马尔可夫链改进方法的铁路货

运量预测研究 [J]. 铁道学报, 2005 (3).

- [3] 王露璐, 邓 丽, 刘先涛. 股票投资价值灰色马尔可夫预测 [J]. 商业研究, 2004 (285).
- [4] 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 (4).
- [5] 刘志彬, 施 斌. 灰色马尔可夫链在深基坑沉降预测中的应用 [J]. 煤田地质与勘探, 2002 (6).
- [6] 张景林, 崔国璋. 安全系统工程 [M]. 北京: 煤炭工业出版社, 2002.
- [7] 国家安全生产监督管理局, 国家煤矿安全监察局. 中国安全生产年鉴 [M]. 北京: 煤炭工业出版社, 2002.
- [8] 钱家忠, 朱学愚, 吴剑锋, 等. 矿井涌水量的灰色马尔可夫预报模型 [J]. 煤炭学报, 2000 (1).
- [9] 刘思峰, 郭天榜, 党耀国, 等. 灰色系统理论及其应用 [M]. 北京: 科学出版社, 2000.
- [10] 国家经济贸易委员会安全生产局. 中国安全生产年鉴 [M]. 北京: 民族出版社, 2000.

作者简介: 张林华 (1982-), 男, 安徽庐江人, 硕士研究生, 研究方向为瓦斯灾害预测与防治。Tel: 0391-3574185, E-mail: lukeliu@hpu.edu.cn

收稿日期: 2006-06-09; 责任编辑: 朱拴成

欢迎订阅 2007 年《矿业安全与环保》

《矿业安全与环保》杂志是属国家安全生产监督管理总局 (国家煤矿安全监察局) 主管, 由煤炭科学研究总院重庆分院与国家煤矿安全技术工程研究中心共同主办, 面向国内外公开发行的国家级科技期刊, 系全国中文核心期刊, 中国学术期刊综合评价数据库来源期刊, 中文科技期刊数据库原文收录期刊, 中国期刊网、中国学术期刊 (光盘版) 全文收录期刊, 万方数据数字化期刊群全文上网期刊。

报道内容: 以煤矿及非煤矿山安全技术、矿山环境保护技术为主, 包括矿井瓦斯、煤与瓦斯突出防治技术与装备; 矿井通风防灭火技术与装备; 工业粉尘及可燃性气体、粉尘爆炸防治技术与装备; 矿井水害防治技术与装备; 采掘工程安全技术; 安全与环境检测、控制技术; 物探与岩土工程技术; 安全管理与评价; 矿山污染治理与综合利用等环保技术; 兼营广告业务。

主要栏目: 试验研究, 问题探讨, 综述, 应用技术, 技术经验, 安全管理, 事故分析等。

发行范围: 全国各地煤矿和非煤矿山的企事业单位、管理部门、科研院所、大专院校、图书馆及图书信息部门, 以及美、德、俄、朝等国家及香港、台湾地区。

读者对象: 从事安全、环保技术的各级领导、干部, 科研与设计人员, 生产技术人员, 院校师生, 各级监察机关、管理部门人员, 以及相关设备、仪器等生产企业的技术管理干部和技术人员。

《矿业安全与环保》为双月刊、国际标准大 16 开本, 2007 年定价 10.00 元/册, 全年 6 期 60.00 元。本刊刊号为 ISSN 1008-4495、CN 50-1062/TD, 邮发代号 78-35, 广大读者可到当地邮局订阅, 也可随时直接汇款到杂志社订阅 (邮局和银行汇款请注明“订刊款”字样)。

欢迎订阅/投稿/刊登广告

地 址: 重庆市沙坪坝区上桥·煤炭科学研究总院重庆分院内

电 话: 023-65239221, 65235167

电子信箱: bianjibu@cqccri.com

传 真: 023-65239221

账 号: 3100024709008904327

邮政编码: 400037

收款单位: 煤炭科学研究总院重庆分院

开户银行: 重庆市工商银行上桥分理处