

城市污水排放的灰色马尔柯夫预测模型

马占青¹, 崔广柏¹, 杨宏杰², 王晓强²

(1. 河海大学水文水资源及环境学院, 江苏 南京 210098; 2. 河海大学港口航道及海岸工程学院, 江苏 南京 210098)

摘要 :应用数据加载法提出了 GM(1,1) 的修正模型,通过灰色预测法和马尔柯夫链预测法的耦合,建立了城市污水排放量的灰色马尔柯夫预测模型.灰色马尔柯夫预测模型具有灰色系统应用少量数据即可建模,以及马尔柯夫链预测可以预测数据值波动较大的序列的特点.计算结果表明,城市污水排放量的预测值很好地吻合了实际值.

关键词 城市污水;灰色模型;马尔柯夫模型;预测

中图分类号 :X522 **文献标识码** :A **文章编号** :1000-198X(2000)05-0049-05

我国学者邓聚龙^[1]提出灰色理论以来,灰色预测已广泛地应用于许多科学领域,如农业、经济、运输、气象等.灰色系统是指信息不完全和不确定的系统,它介于信息完全确定的白色系统和信息完全未知的黑色系统之间.从系统论的角度出发,一个区域(如某一城市)的环境系统可以看作为一个灰色系统,其环境因子可应用灰色理论进行预测.灰色系统理论在环境科学领域的应用^[2]亦比较广泛,如大气污染、城市固体废弃物、水环境质量等的预测^[3~5].城市污水排放量的预测同其它环境因子的预测一样,由于历史数据时间序列短且波动较大,因此预测精度相对较差.本文通过灰色预测法和马尔柯夫预测法的耦合,建立了灰色马尔柯夫预测模型,并对城市污水排放量的预测问题进行了探讨.

1 一般的 GM(1,1)模型

现用 GM 表示灰色模型(Grey Model),则关于一阶模型、一个变量的情况可记为 GM(1,1).一般的 GM(1,1)模型的基本思想是,对原始无明显规律的时间序列 $\{X^{(0)}(t)\} \ t=1,2,\dots,k$ 作一次累加生成,得到序列

$$X(t)=\sum_{i=1}^tX^{(0)}(i)$$

弱化原序列中的随机性,然后建立时间连续微分方程

$$\frac{dx}{dt}+ax=\mu \tag{1}$$

式中 a,μ 为参数.解上述微分方程得时间响应模型

$$\hat{x}(t+1)=(x^{(0)}(1)-\frac{\mu}{a})e^{-at}+\frac{\mu}{a} \tag{2}$$

设方程的参数向量 $\hat{a}=[a,\mu]^T$,可由最小二乘法求解模型的参数 a,μ ,即

$$\hat{a}=[a,\mu]^T=(B^T\cdot B)^{-1}\cdot B^T\cdot Y_k$$

式中 B 为累加生成矩阵, Y_k 为向量,二者的构造分别为

$$B=\begin{bmatrix} -\frac{1}{2}(x(2)+x(1)) & 1 \\ -\frac{1}{2}(x(3)+x(2)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(x(k)+x(k-1)) & 1 \end{bmatrix} \tag{3}$$

$$Y_k = [x^{(0)}(2) \quad x^{(0)}(3) \quad \dots \quad x^{(0)}(k)]^T \tag{4}$$

式中： $x^{(0)}(i)$ ——第 i 年的原始数据； $x(i)$ ——第 i 年的一次累加生成值。

令 $t = 1, 2, \dots, k-1$ ，由 (2) 式可求得 $\hat{x}^{(0)}(t+1)$ ，即 $\hat{x}(2), \hat{x}(3), \dots, \hat{x}(k)$ ，从而有数列 $\{\hat{x}(t+1)\}$ 。
 $\hat{x}(t+1)$ 是一次累加量，还需求出还原值 $\hat{x}^{(0)}(t+1)$ ，即

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

由模型求出还原值后，还必须对灰色模型的精度进行检验。通常用小误差概率 P 和后验差比值 C 进行模型精度检验。将合格模型用于预测，并用实际值进行预测精度检验。

2 数据加载的 GM(1,1)模型

对同一灰色过程而参数不同的时间序列来说，其信息系统是相互影响的，由于受各种因素的影响，一些参数值表现出较大的波动性，另一些参数值则波动性较小。将波动性较大的数据序列加载于波动性较小的数据序列，得到新的数据序列。此数据序列波动性相对较小。在此基础上建立 GM(1,1)模型并对模型进行检验，然后进行预测。

2.1 建模方法

a. 设某灰色过程为 ω ，其原始序列为 $\{x_{j,i}^{(0)}\} (j = 1, 2, \dots, m; i = 1, 2, \dots, k)$ （其中 m 为灰色过程 ω 的原始序列总数， k 为每一原始序列 $\{x_{j,i}^{(0)}\}$ 的时间序列长度），则 ω 可用矩阵表示如下：

$$\omega = \begin{bmatrix} x_{1,1}^{(0)} & x_{1,2}^{(0)} & \dots & x_{1,k}^{(0)} \\ x_{2,1}^{(0)} & x_{2,2}^{(0)} & \dots & x_{2,k}^{(0)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1}^{(0)} & x_{m,2}^{(0)} & \dots & x_{m,k}^{(0)} \end{bmatrix}$$

首先对原始序列 $\{x_{j,i}^{(0)}\}$ 进行初始化处理得到新的序列

$$\{x'_{j,i}^{(0)}\} = \{x_{j,i}^{(0)}\} / x_{j,1}^{(0)} \quad (i = 1, 2, \dots, k)$$

然后对初始化后的序列 $\{x'_{j,i}^{(0)}\}$ 按照一般的 GM(1,1)模型进行处理，并用小误差概率 P_j 和后验差比值 C_j 对每一模型进行精度判别（表 1）。精度达到 1 级的模型可直接用于预测，对于精度未达到 1 级或不合格的模型则进行数据加载处理。

表 1 模型精度等级判别
Table 1 Criteria for model precision

检验参数	1 级 (良好)	2 级 (合格)	3 级 (勉强)	4 级 (不合格)
小误差概率 P	$0.95 \leq P$	$0.80 \leq P < 0.95$	$0.70 \leq P < 0.80$	$P < 0.70$
后验差比值 C	$0.35 \geq C$	$0.35 < C \leq 0.50$	$0.50 < C \leq 0.65$	$C > 0.65$

b. 假定某一初始化原始序列 $\{x'_{a,i}^{(0)}\} (1 \leq a \leq m)$ 所建模型达到了最佳精度，而另一初始化原始序列 $\{x'_{b,i}^{(0)}\} (1 \leq b \leq m, a \neq b)$ 所建模型的精度较差，此时可将 $\{x'_{b,i}^{(0)}\}$ 序列加载于 $\{x'_{a,i}^{(0)}\}$ 序列以构成新的序列

$$\{x'_{c,i}^{(0)}\} = \{x'_{a,i}^{(0)}\} + \{x'_{b,i}^{(0)}\}$$

然后以新构成的时间序列 $\{x'_{c,i}^{(0)}\}$ 按照一般的 GM(1,1)模型进行建模，并用小误差概率 P_c 和后验差比值 C_c 对模型进行精度检验。由 $\{x'_{c,i}^{(0)}\}$ 序列所建模型的精度好于 $\{x'_{b,i}^{(0)}\}$ 序列所建模型的精度。

由新构成的时间序列 $\{x'_{c,i}^{(0)}\}$ 和初始化原始时间序列 $\{x'_{a,i}^{(0)}\}$ 所建模型分别进行对应于初始化原始时间序列 $\{x'_{b,i}^{(0)}\}$ 的未来时段的预测，即有

$$\hat{x}'_{b,\lambda}^{(0)} = (\hat{x}'_{c,\lambda}^{(0)} - \hat{x}'_{a,\lambda}^{(0)}) \times x'_{b,1}^{(0)}$$

式中： $\lambda = k+1, k+2, \dots, p$ ； $\hat{x}'_{b,\lambda}^{(0)}$ —— $\{x'_{b,i}^{(0)}\}$ 时间序列未来时段的预测值； $\hat{x}'_{c,\lambda}^{(0)}$ —— $\{x'_{c,i}^{(0)}\}$ 时间序列未来时段的预测值； $\hat{x}'_{a,\lambda}^{(0)}$ —— $\{x'_{a,i}^{(0)}\}$ 时间序列未来时段的预测值。

最后用实际值对预测值进行检验。

2.2 模型实例

城市给排水系统中的用水总量与污水排放量密切相关 ,可视为同一灰色过程 :在同一时间序列内用水过程的同时即排放污水.现以包头市城市用水总量和污水(工业废水 + 生活污水)排放量为例 ,以 1986 ~ 1995 年数据构成时间序列(表 2) ,按上述方法进行建模 ,结果及比较见表 3.由表 3 可看出 ,用水总量的预测模型精度较高 ,小误差概率 P 和后验差比 C 达到 2 级 ,可用于未来时段预测.而污水排放量的预测模型精度较低 ,小误差概率 P 和后验差比 C 为 3 级.经数据加载后 ,模型精度得到改善 ,小误差概率 P 和后验差比 C 达到 2 级 ,以 1996 年的实际数据进行检验 ,预测结果为 8 556.45 万 t/a ,相对误差为 0.40% ,而未经数据加载时的相对误差为 2.35% .由此可见 ,污水排放量数据加载于用水总量模型后 ,所建模型的精度及预测精度较好.

表 2 污水排放量及用水总量

Table 2 Sewage drainage and total water consumption 万 t/a		
年份	污水排放量	用水总量
1986	8 020. 05	163 808. 7
1987	8 023. 05	161 998. 4
1988	8 332. 74	161 042. 2
1989	8 560. 12	162 118. 0
1990	8 575. 35	163 666. 1
1991	8 310. 29	163 690. 5
1992	8 147. 58	165 691. 9
1993	8 187. 97	167 255. 6
1994	8 371. 85	169 794. 4
1995	8 555. 66	178 498. 6
1996	8 690. 56	188 468. 5

表 3 结果比较

Table 3 Comparison of results			
项目	模型	小误差概率 P	后验差比 C
用水总量	$\hat{x}'(t+1)=90.925\,09e^{0.010\,616\,14t}-89.925\,09$	0.90	0.45
污水排放量	$\hat{x}(t+1)=1\,911.44e^{0.000\,545\,627t}-1\,910.44$	0.71	0.51
数据加载	$\hat{x}(t+1)=367.566\,6e^{0.005\,464\,304t}-365.566\,6$	0.83	0.46

3 灰色马尔柯夫预测模型

3.1 马尔柯夫预测模型

马尔柯夫链预测的理论基础是马尔柯夫过程 ,它描述的是一个随机时间序列的动态变化过程^[3].根据马尔柯夫链 ,将数据序列分成若干状态 ,以 $S_1, S_2, \dots, S_n$ 来表示 ,按时序将转移时间取为 $t_1, t_2, \dots, t_n$. $P_{ij}^{(k)}$ 表示数据序列由状态 S_i 经 k 步变为 S_j 的概率 ,即

$$P_{ij}^{(k)} = \frac{n_{ij}^{(k)}}{N_i^{(k)}} \tag{6}$$

式中 : $n_{ij}^{(k)}$ ——状态 S_i 经 k 步变为 S_j 的次数 ; $N_i^{(k)}$ ——状态 S_i 出现的总次数.则 k 步状态转移概率矩阵为

$$R^{(k)} = \begin{bmatrix} P_{11}^{(k)} & P_{12}^{(k)} & \cdots & P_{1j}^{(k)} \\ P_{21}^{(k)} & P_{22}^{(k)} & \cdots & P_{2j}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{j1}^{(k)} & P_{j2}^{(k)} & \cdots & P_{jj}^{(k)} \end{bmatrix} \tag{7}$$

若初始状态 S_i 的初始向量为 $A^{(0)}$,则经 k 步转移后 ,向量

$$A^{(k)} = A^{(0)} \cdot R^{(k)} \tag{8}$$

3.2 灰色马尔柯夫预测模型

运用上述马尔柯夫预测模型 ,以 1986 ~ 1995 年包头市污水排放量为时间序列 ,建立灰色马尔柯夫预测模型 ,并用实际值(1996 年)进行检验.

a. 对原始序列 $X_i^{(0)}(i=1, 2, \dots, n)$ 建立 GM(1, 1) 模型 ,得到时间响应函数式

$$\hat{x}(t+1)=1\,911.44e^{0.000\,545\,627t}-1\,910.44 \tag{9}$$

令
$$\hat{x}^{(0)}(t+1)=\hat{x}(t+1)-\hat{x}(t) \tag{10}$$

由式(9)(10)可求得趋势拟合值 $\hat{x}_1^{(0)}(i)(i=1, 2, \dots, n)$ 一次残差

$$\Delta x_1(i)=\hat{x}_1^{(0)}(i)-x(i) \tag{11}$$

及相对差值

$$\Delta p_1(i) = \frac{\Delta x_1(i)}{\hat{x}_1^{(0)}(i)}$$

(12)

将式(12)所得结果填入表6.

b. 对相对差值序列建立马尔柯夫模型,求出拟合值 Δp_1 ,确定状态概率转移矩阵.由于对城市用水总量和污水排放量进行预测时,状态界限是不确定的,求式(7)时,采用适算法.即:选取一组临界值,将已知相对差值序列代入式(6),然后求出概率转移矩阵,并用已知数据进行检验,符合率高的为选用转移概率矩阵.本例如选取 -1% 、 -0.1% 、 0.1% 、 1% 为界限,将相对差值序列划分为5个区间(表4),则可得到本例的相对差值序列所处的状态(表5).

表 4 概率状态划分

Table 4 Division of the probability states

概率区间	$-\infty \sim -1\%$	$-1\% \sim -0.1\%$	$-0.1\% \sim 0.1\%$	$0.1\% \sim 1\%$	$1\% \sim +\infty$
状态	I	II	III	IV	V

表 5 相对误差概率状态划分

Table 5 Probability states of ralative error

年份	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
状态	II	I	IV	V	V	II	I	I	II	V

由表5及式(6)(7)可得一步、二步、三步、四步、五步状态概率转移矩阵,由式(8)可得

$$A^{(i)} = \sum_{j=1}^5 A^{(i-j)} \cdot R^{(j)}$$

(13)

c. 用GM(1,1)模型求得预测趋势后,对相对差值进行马尔柯夫修正,可得到二次拟合值,进而可求得二次残差(表6).

表 6 灰色马尔柯夫模型预测计算

Table 6 Calculation of the Grey Markovian predicting model

时序	实际值	GM(1,1)模型			马尔柯夫模型		
		一次拟合值	相对差值	拟合相对差值	二次拟合值	二次拟合残差	二次相对差值/%
1986	8 020.05	8 286.45	- 0.03	$-1\% \sim -0.1\%$			
1987	8 023.05	8 306.59	- 0.53	$-\infty \sim -1\%$			
1988	8 332.74	8 326.36	0.08	$0.1\% \sim 1\%$			
1989	8 560.12	8 287.39	3.19	$1\% \sim +\infty$			
1990	8 575.35	8 366.06	2.44	$1\% \sim +\infty$			
1991	8 310.29	8 385.98	- 0.91	$-1\% \sim -0.1\%$	8 302.12	8.17	0.01
1992	8 147.58	8 405.95	- 3.17	$-\infty \sim -1\%$	8 321.89	- 174.31	- 2.14
1993	8 187.97	8 425.97	- 2.91	$-\infty \sim -1\%$	8 341.71	- 153.74	- 1.88
1994	8 371.85	8 446.03	- 0.89	$-1\% \sim -0.1\%$	8 450.49	- 78.64	- 0.94
1995	8 555.66	8 476.13	0.95	$1\% \sim +\infty$	8 560.90	- 5.24	0.06
1996	8 690.56	8 486.30	2.35	$1\% \sim +\infty$	8 582.16	108.40	1.24

d. 对所建立的灰色马尔柯夫模型进行小误差概率 P 和后验差比 C 的检验,检验合格后即可用于预测.本例中的 $P=0.90$ 、 $C=0.38$,因此模型精度合格,可用于预测.如1996年的预测值为8 582.16,用实际值进行检验,则相对误差为1.24%,其精度有很大改善,见图1.

4 结 语

本文运用灰色理论和马尔柯夫理论建立了城市污水排放量灰色马尔柯夫预测模型. 本方法所建模型的特点是,可运用少量的数据建模,可预测数据波动较大的序列. 灰色马尔柯夫预测模型在实际应用时,可作为动态模型,即每获得一个新的信息(实际值,非模型预测值),可按照本文所述方法建模,并进行未来时段的预测. 本文还通过数据加载法建立了数据加载 GM(1,1)修正模型,预测结果较好.

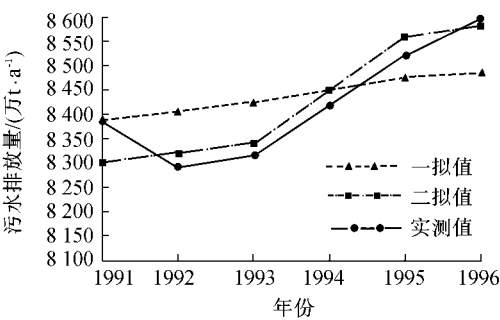


图 1 实际值与拟合值比较
Fig.1 Comparison between the fit value and real value

参考文献：

[1] 邓聚龙. 灰色系统基本方法[M]. 郑州: 河南大学出版社, 1987. 5 ~ 25.
[2] 曾光明, 杨春平. 城市开发区污染物排放量的预测计算方法[J]. 中国环境科学, 1995, 15(6) 5 ~ 9.
[3] 任成忠. 大气污染灰色预测法的研究[J]. 环境工程, 1996, 14(4) 35 ~ 38.
[4] 陈伟国, 陈翠芝. 城市固体废物的灰色预测与评价[J]. 上海环境科学, 1996, 15(1) 8 ~ 10.
[5] 冯玉国. 水环境质量灰色评价模型及应用[J]. 水资源保护, 1995, 16(1) 29 ~ 32.

Grey Markonian Model for Prediction of Urban Sewage Drainage

MA Zhan-qing¹, CUI Guang-bo¹, YANG Hong-jie², WANG Xiao-qiang²

(1. College of Water Resources and Environment, Hohai Univ., Nanjing 210098, China ;

2. College of Harbour, Waterway and Coastal Engineering, Hohai Univ., Nanjing 210098, China)

Abstract :This paper presents a high precision Grey Markovian Model on the basis of the grey theory and Markov theory. This model is used to forecast urban sewage drainage. The Grey Markovian Model is suitable for a few and wavy data. An example shows that the results of calculation are in good agreement with the measured data. Besides, a revised GM(1,1) model based on data loading is also discussed in this paper. The precision is also improved.

Key words :urban sewage ; grey model ; Markov model ; forecast