第24卷第3期 水资源与水工程学报 Vol.24 No.3
2013年 6月 Journal of Water Resources & Water Engineering Jun.,2013

灰色理论在龙江突发镉污染水质预测中的应用

魏智宽[1], 蒋世云[1], 李少旦[2]
(1.广西工学院 生物与化学工程学院, 广西 柳州 545006; 2.广西柳州疾病预防控制中心, 广西 柳州 545007)

摘 要: 利用灰色预测理论,建立了龙江突发水污染水质变化趋势的GM(1,1)预测模型,而后对模型进行残差修正。结果表明:经残差修正后的GM(1,1)模型更为合理,预测精度明显提高。本研究中灰色理论能够有效地对突发性水污染事故中的水质变化趋势进行短期预测,为相关部门及时采取相应的应急措施提供参考,减小事故风险。

关键词: GM(1,1)模型; 残差修正; 水质预测; 龙江
中图分类号:X522 **文献标识码**: A **文章编号**: 1672-643X(2013)03-0135-03

Application of grey theory in water quality prediction of sudden cadmium pollution in Longjiang River

WEI Zhikuan[1], **JIANG Shiyun**[1], **LI Shaodan**[2]
(1. *Department of Biological and Chemical Engineering, Guangxi University of Technology, Liuzhou* 545006, *China*;
2. *Guangxi Liuzhou Center for Disease Control and Prevention, Liuzhou* 545007, *China*)

Abstract: Through application of grey forecast theory, the paper built GM(1,1) forecast model of pollution water quality change trend of the Longjiang River burst and modified the model residual. The results show that the GM(1,1) model after residual modification is more reasonable , and prediction accuracy is obviously improved. The residual GM(1,1) model can effectively forecast the trend of water quality during a short period in sudden water pollution accidents, provide reference for the relevant department to adopt the corresponding emergency measures and reduce the risk of accidents.

Key words: GM(1,1)model; residual modification; water quality forecast; Longjiang river

水质预测是通过对河流水质现状、污染物迁移的特点和污染物排放情况来预测水质未来的变化趋势。在突发性水污染事故中,水质预测显得尤为重要,它可以为人民群众提供饮用水的安全信息。目前,用于突发性污染水质预测的方法主要有水质一维模型、水质二维模型、水质三维模型[1-2]等。由于在突发性污染中,污染源的位置、总量,以及污染发生的时间,发生污染时的天气、地理和水文条件等都不确定,加之水质的变化有很大的随机性,这给传统的确定性水质数学模型的使用带来了很大困难。灰色系统分析方法对于信息不完整或不完全的情况具有良好的适用性。灰色系统是黑箱概念的一种推广,灰色系统理论是由华中科技大学邓聚龙教授于1982年提出的,30多年来,引起了很多国内外学者的关注,并得到了长足的发展。其中GM(1,1)模型在水质预测中得到广泛应用[3],并取得了很好的效果。当GM(1,1)模型精度不够准确时,可以用残差建立GM(1,1)模型,来对原模型进行修正。本文应用灰色系统理论建立了残差GM(1,1)模型,尝试对龙江突发镉污染水质中镉浓度的变化进行预测,取得了良好的预测效果。

1 模型的建立

1.1 GM(1,1)模型

GM(1,1)模型是一种一阶微分方程,其建模的主要步骤如下[4]:

(1) 由原始序列 $X^{(0)}$ 计算一次累加序列 $X^{(1)}$;

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

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

将无规律的 $X^{(0)}$ 经过一次累加后生成的 $X^{(1)}$ 变得较有规律。其中, $x^{(1)}(k) = \sum_{i=1}^{k} x(i), (k = 1,2,\cdots,n)$

(2) 由 $X^{(1)}$ 构造紧邻均值生成序列 $Z^{(1)}$,其中,

收稿日期:2013-01-14; **修回日期**:2013-03-25
基金项目:广西卫生厅科研项目(Z2012549); 广西壮族自治区科学技术厅科研项目(桂科攻1355007-2)
作者简介:魏智宽(1986-),男,河北沧州人,在读研究生,主要研究方向为污水处理技术。
通讯作者:李少旦(1975-),男,湖南邵阳人,副主任技师,主要从事水污染预警与应急处理。

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

$$z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1))$$

(3) 建立白化形式微分方程:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \tag{1}$$

其中:a 为发展系数; u 为灰色作用量。

(4) 建立矩阵 B,y

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

$$y = (x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(n))^{T}$$

(5) 作最小二乘估计,求参数 a 和 u

$$\hat{U} = \begin{bmatrix} a \\ u \end{bmatrix} = (B^{T}B)^{-1}B^{T}y$$

(6) 建立 GM(1,1) 模型时间响应方程

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

(7) 对 $\hat{x}(k+1)$ 作累减还原,得到 GM(1,1) 预测模型:

$$\begin{aligned} \hat{x}^{(0)}(k+1) &= \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \\ &= (1-e^{a})\left(x^{(0)}(1) - \frac{u}{a}\right)e^{-ak}, \\ & k = 1,2,\cdots,n \end{aligned} \tag{3}$$

1.2 模型精度的检验

为确定模型是否合理,采用后验差检验法[5]对模型的精确度进行检验:

(1) 分别计算原始序列 $X^{(0)}$ 及残差序列 $\varepsilon^{(0)}(k)$ 的方差 S_1 和 S_2:

$$S_1^2 = \frac{1}{n}\sum_{k=1}^{n}[x^{(0)}(k) - \bar{x}]^2$$

$$S_2^2 = \frac{1}{n}\sum_{k=1}^{n}[e(k) - \bar{e}]^2$$

其中, $\varepsilon^{(0)}(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$,$(k = 1,2,\cdots,n)$, $\bar{x} = \frac{1}{n}\sum_{k=1}^{n}x^{(0)}(k)$,$\bar{e} = \frac{1}{n}\sum_{k=1}^{n}e(k)$

(2) 分别计算后验差比 C 及小误差概率 p:

$$C = S_2/S_1,\ p = P\{|e(k) - \bar{e}| < 0.6745S_1\}$$

C 和 p 是后验差检验的两个重要指标,指标 C 越小,就表示尽管原始数据很离散,但模型所得的计算值与实际值之差并不是很离散。很明显 C 越小越好。一般将模型的精度等级分为 4 级,具体见表 1。

表 1 灰色模型等级对照表

预测精度等级	C	p
好	< 0.35	> 0.95
合格	< 0.45	> 0.80
勉强	< 0.50	> 0.70
不合格	≥ 0.65	≤ 0.70

1.3 残差修正 GM(1,1) 模型

当 GM(1,1) 模型的精度达不到要求时,可以用残差序列 $\varepsilon^{(0)}(k)$ 建立 GM(1,1) 模型,来对原模型进行修正,以提高精度[3]。

计算原始值与模型计算值之间的残差 $\varepsilon^{(0)}$,取 $\varepsilon^{(0)}(k)$,$(k = k_0, k_0+1, \cdots, n)$ 为可建模的残差尾段,对其作一阶累加生成,建立 GM(1,1) 模型, 求出残差修正值方程:

$$\begin{aligned} \hat{\varepsilon}^{(0)}(k+1) &= (1 - e^{a_\varepsilon}) \cdot \\ & \left(\varepsilon^{(0)}(k_0) - \frac{u_\varepsilon}{a_\varepsilon}\right)e^{-a_\varepsilon(k-k_0)},\ k \geq k_0 \end{aligned} \tag{4}$$

若残差 $\varepsilon^{(0)}(k)$ 序列中的最小值 $\varepsilon^{(0)}(k)_{\min} < 0$,则将横坐标下移 $2|\varepsilon^{(0)}(k)_{\min}|$ 单位,得到非负残差序列 $\dot{\varepsilon}^{(0)}(k) = \varepsilon^{(0)}(k) + 2|\varepsilon^{(0)}(k)_{\min}|$[6],将 $\dot{\varepsilon}(k)$ 作一阶累加生成,建立 GM(1,1) 模型,求出此时的残差 GM(1,1) 模型的时间响应方程并作累减还原得到残差修正值方程, 再将横坐标上移 $2|\varepsilon^{(0)}(k)_{\min}|$,得:

$$\begin{aligned} & \hat{\varepsilon}^{(0)}(k+1) \\ & = (1 - e^{a_\varepsilon})\left(\varepsilon^{(0)}(k_0) - \frac{u_\varepsilon}{a_\varepsilon}\right)e^{-a_\varepsilon(k-k_0)} - \\ & \quad 2|\varepsilon^{(0)}(k)_{\min}| \qquad k \geq k_0 \end{aligned} \tag{5}$$

用 $\hat{\varepsilon}^{(0)}(k+1)$ 修正同时刻的 $x^{(0)}(k+1)$,即得残差修正 GM(1,1) 模型:

$$\hat{x}^{(0)}(k+1) = \begin{cases} (1-e^{a})(x^{(0)}(1) - \frac{u}{a})e^{-ak}, k < k_0 \\ (1-e^{a})(x^{(0)}(1) - \frac{u}{a})e^{-ak} + \\ (1-e^{a_\varepsilon})\left(\varepsilon^{(0)}(k_0) - \frac{u_\varepsilon}{a_\varepsilon}\right)e^{-a_\varepsilon(k-k_0)} - \\ 2|\varepsilon^{(0)}(k)_{\min}|, k \geq k_0 \end{cases} \tag{6}$$

2 应用实例

2012 年 1 月 15 日,广西龙江河拉浪水电站网箱养鱼出现少量死鱼现象,龙江河宜州拉浪码头前

200 m 水质重金属超标 80 倍。时间正值农历龙年春节，龙江河段检测出重金属镉含量超标，其下游的柳江是广西的工业重镇柳州市工农业用水及生活用水的重要水源，沿岸居民饮水安全遭到严重威胁。至事故发生中期，经过政府的治污工作，沿岸污染源基本被关闭，污染团已被控制水情基本稳定，此时污染团已经到达龙江西门崖处，龙江西门涯处距事故发生地约 35 km，距柳州的河西水厂约 45 km，疾控中心工作人员于 1 月 29 日至 2 月 2 日，每日 6 时和 18 时，连续 10 个时间段，在龙江西门崖处附近水域对水中镉浓度进行监测，检测仪器使用 PinAAcle900T 原子吸收分光光度计（美国 PerkinElmer 公司），监测数据详见表 2。

表 2　镉浓度时段监测值　μg/L

时段	1	2	3	4	5	6	7	8	9	10
镉浓度	17.58	16.27	15.93	16.48	13.66	16.75	17.56	19.61	20.76	23.22

下面通过前 8 个时段的监测数据建立 GM(1,1) 模型来预测后 2 个时段的镉浓度变化。

原始序列为：

$$X^{(0)} = \{17.58, 16.27, 15.39, 16.48, 13.66, 16.75, 17.56, 19.61\}$$

将该序列作一次累加生成，得到 $X^{(0)}$ 的 1 − AGO 序列 $X^{(1)}$：

$$X^{(1)} = \{17.58, 33.85, 49.78, 66.26, 79.92, 96.67, 114.23, 133.84\}$$

$X^{(1)}$ 的紧邻均值生成序列 $Z^{(1)}(k)$ 为：

$$Z^{(1)}(k) = \{25.7150, 41.8150, 58.0200, 73.0900, 88.2950, 105.4500, 124.0350\}$$

作最小二乘估计，求参数 a 和 u：

$$\hat{U} = \begin{bmatrix} a \\ u \end{bmatrix} = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}y = \begin{bmatrix} -0.0307 \\ 14.3405 \end{bmatrix}$$

得时间响应方程为：

$$\hat{x}^{(1)}(k+1) = 484.6973\mathrm{e}^{0.0307k} - 467.1172 \quad (7)$$

作累减生成得预测模型为：

$$\hat{x}^{(0)}(k+1) = 14.2205\mathrm{e}^{0.0307k} \quad (8)$$

由此得出实测值、模型预测值以及他们之间的相对误差，见表 3。

表 3　实测值、模型预测值及相对误差计算结果　μg/L，%

时间	时段	实测值	预测值	残差	相对误差
01-29 6:00	1	17.58	17.5800	0.0000	0.0000
01-29 18:00	2	16.27	15.1121	1.1579	7.1168
01-30 6:00	3	15.93	15.5839	0.3461	2.1726
01-30 18:00	4	16.48	16.0704	0.4096	2.4854
01-31 6:00	5	13.66	16.5722	−2.9122	−21.3192
01-31 18:00	6	16.75	17.0896	−0.3396	−2.0275
02-01 6:00	7	17.56	17.6231	−0.0631	−3.5933
02-01 18:00	8	19.61	18.1733	1.4367	7.3264

经后验差法检验，$C = 0.8369 > 0.65$，模型不合要求，需采用残差模型进行修正，取 $k_0 = 4$（不需取全部数据进行残差修正），由于残差尾段序列中的最小值 $\varepsilon^{(0)}(5) = -2.122 < 0$，需作非负处理，将坐标轴下移 $2|\varepsilon^{(0)}(5)| = 5.8244$，建立残差 GM(1,1) 模型，求出残差修正值方程，再将坐标值上移得：

$$\hat{\varepsilon}^{(0)}(k+1) = 2.5893\mathrm{e}^{0.237(k-4)} - 5.8244 \quad (9)$$

用 $\hat{\varepsilon}(k+1)$ 同时刻的预测数据，得残差修正 GM(1,1) 模型：

$$\hat{x}^{(0)}(k+1) = \begin{cases} 14.2205\mathrm{e}^{0.0307k}, & k < 4 \\ 14.2205\mathrm{e}^{0.0307k} + 2.5893\mathrm{e}^{0.237(k-4)} - 5.8244, & k > 4 \end{cases} \quad (10)$$

根据修正后的模型进行计算，得出实测值、模型预测值以及他们之间的误差，见表 4。

表 4　残差修正后的实测值、模型预测值及相对误差的计算结果　μg/L，%

时间	时段	实测值	预测值	残差	相对误差
01-30 18:00	4	16.48	16.4800	0.0000	0.0000
01-31 6:00	5	13.66	14.3718	−0.6580	4.8169
01-31 18:00	6	16.75	15.3409	1.4094	8.4125
02-01 6:00	7	17.56	17.0866	0.4734	2.6959
02-01 18:00	8	19.61	19.7270	−0.117	0.5966

经检验，$C = 0.399 < 0.45$，$p = 1$，根据表 4，可判断模型合格。将时段 $k = 9$ 代入预测模型，可得 2 日上午 6 时的镉浓度预测值为 22.23 μg/L，与实测值相比较，相对误差为 6.613%。将时段 $k = 10$ 带入模型，可得 2 日 18 时的镉浓度预测值为 25.13 μg/L，与实测值相比较误差率为 7.6%。说明此方法可以对龙江镉污染水质趋势进行预测。

（下转第 141 页）

表 2　隐层节点数对精度影响　s

序号	隐层节点数	训练时间	训练次数	均方差 MSE
1	15	321.70	2876	5.19759
2	16	12.36	332	4.21332×10^{-7}
3	17	1.58	562	0.00019952
4	18	45.06	587	0.000199191
5	19	8.98	325	1.44438×10^{-6}
6	20	4.58	33	1.28396×10^{-5}
7	21	2.36	46	1.61245×10^{-5}
8	22	4.987	399	9.84118×10^{-7}
9	23	3.928	522	9.82206×10^{-6}
10	24	6.069781	925	1.30926×10^{-7}
11	25	41.6670	4152	3.83574×10^{-5}
12	27	27.59	3927	0.00019961
13	30	32.345	236	9.31619×10^{-6}
14	35	6.998	56	1.87623×10^{-6}
15	40	15.236	127	1.65773×10^{-6}
16	45	25.693	366	0.00017357

新新邱露天煤矿地下水环境质量分为 4 类,以此基础上采用 BP 神经网络方法,以各个指标做为输入节点,以标准化矩阵 R_1 每行的加权之和为输出接点,对数据进行训练的和预测。得出主要结论如下:

(1)用模糊聚类法能解决矿区地下水污染程度等级分类问题,需要针对不用地区的特点,建立适合本地区的模型和和聚类界限及 λ 的选取问题,只有这样才能更加符合实际。

(2)模糊聚类分析法理论简单明,具体计算可在计算机上进行,能迅速而准确地做出稳定性判断, MATLAB 的 BP 神经网络工具箱在针对数据的训练和预测运行速度快且精确,因此对选取的样本数和反映各个样本的因素是决定分类和预测准确的关键因素。

参考文献:

[1] 姜建军. 矿山环境五大问题亟待解决[J]. 地质与勘探, 2005(3):70.

[2] Maree J P, Gerber A, Mclaren A R, et al. A biological treatment of mining effluents[J]. Environ Tech Letters, 1987, 8:53 - 64.

[3] 韦冠俊. 矿山环境工程[M]. 北京:冶金工业出版社, 2001:92 - 96.

[4] Maree J P, Hill E. Biological removal of sulphate from industrial effluents and concomitant production of slphur[J]. Sci Tech, 1989, 21:265 - 276.

[5] 胡文容, 高延耀. 酸性矿井水的处理方法和利用途径[J]. 煤矿环境保护, 1994, 8(1):17 - 21.

[6] 肖利萍, 张春婵, 潘纯林, 等. 生活污水为碳源处理硫酸盐矿山废水可行性实验[J]. 水资源与水工程学报, 2008, 19(6):76 - 78.

[7] 刘志斌, 范军富, 丛 鑫. 煤矸石山对地下水环境质量影响的分析研究[J]. 露天采煤技术, 2002(2):6 - 8 + 12.

[8] 韦冠俊. 矿山环境保护[M]. 北京:冶金工业出版社, 1990.

[9] 刘志斌. 露天煤矿排土场地下水环境质量影响的模糊综合评价[J]. 露天采煤技术, 2003(2):16 - 18.

[10] 高新波. 模糊聚类分析及其应用[M]. 西安:电子科技大学出版社, 2004.

[11] 袁曾任. 人工神经元网络及其应用[M]. 北京:清华大学出版社, 1999.

[12] 丛 爽. 面向 MATLAB 工具箱的神经网络理论与应用[M]. 北京:中国科学技术大学出版社, 1998.

(上接第 137 页)

3　结　语

(1)根据灰色预测理论,建立了龙江镉污染水质变化预测模型,由预测结果可知,灰色预测理论在本研究中有着很好的适用性。

(2)将实测数据与预测数据对比来看,经残差修正后的模型更为合理,预测结果更为可信。可以为相关部门提前采取应急措施提供参考,减小事故风险。

(3)由于影响水质变化的因素众多以及人为采取的措施,水中镉浓度不会无限制增加,仅用以往的数据建立的模型可能会失效,但这并不否定灰色预测方法在本研究中的有效性。因此灰色预测模型可用于对水质进行短期预测,而在长期预测中还应不断的向模型补充新信息,删除旧信息,建立新的预测模型。

参考文献:

[1] 李 丹, 张 涛, 王祥三. 河流随机水质预测模型及其应用[J]. 中国农村水利水电, 2006(11):32 - 35.

[2] 孙佳颖, 徐 卫. 东河流水质预测模型研究进展[J]. 山西建筑, 2010, 36(36):360 - 361.

[3] 王 伟, 程永清, 石砚秀. 灰色预测模型在渭河水环境信息系统中的应用[J]. 环境保护科学, 2007, 33(6), 100 - 102.

[4] 刘思峰, 党耀国, 等. 灰色系统理论及其应用[M]. 北京:科学出版社, 2010.

[5] 李如忠, 汪家权, 钱家忠. 基于灰色动态模型群法的河流水质预测研究[J]. 水土保持通报, 2002, 22(4):10 - 12.

[6] 杨建飞, 刘俊民, 陈 琳. 基于灰色残差模型的灌区地下水最小埋深预测[J]. 人民黄河, 2011, 33(4):101 - 105.