

文章编号:1001-4810(2011)03-0260-10

GM(1,1)分解模型与 ARIMA 模型 在岩溶地下水模拟中的对比研究 ——以柳林泉流量模拟为例

李华敏¹, 吴敬², 赵娇娟¹, 郝永红³, 王亚捷³, 曹碧波¹

(1. 天津师范大学城市与环境科学学院, 天津 300387; 2. 天津财经大学理工学院统计系, 天津 300222;
3. 天津师范大学天津市水环境与水资源重点实验室, 天津 300387)

摘要:本文运用 GM(1,1)分解模型和 ARIMA 模型分别模拟柳林泉流量。根据影响特征将泉水流量变化分为两个时段研究:1957—1973 年泉水流量处于自然状态;1974—2009 年泉水流量受气候变化和人类活动双重影响。运用第 1 时段的数据建模获得自然状态下泉水流量的模型,将模型外推,获得第 2 阶段自然状态下泉水流量,然后根据水量平衡原理,减去同期实测流量,获得人类活动对泉水流量衰减的贡献。GM(1,1)分解模型的结果显示,从 20 世纪 70 年代到 21 世纪初柳林泉衰减量为 $2.26 \text{ m}^3/\text{s}$;ARIMA 模型的结果为 $2.32 \text{ m}^3/\text{s}$;与同期实际衰减量 $2.27 \text{ m}^3/\text{s}$ 比,相对误差分别为 0.44% 和 2.20%,表明两种模型都适用于泉水流量的模拟。对比人类活动和气候变化对柳林泉流量衰减的贡献,两个模型得到同样结果,即人类活动的贡献是气候变化的 8~9 倍。实证研究显示,GM(1,1)模型适用于指数序列的模拟,对具有周期波动的泉水流量,可通过周期修正提高精度;而 ARIMA 模型能够较好地反映泉水流量相对于降水量的时间滞后效应,能比较准确地模拟泉水流量与降水量的量化关系。

关键词:GM(1,1)分解模型;ARIMA 模型;分时段分析;岩溶泉流量

中图分类号:P641.2 **文献标识码:**A

0 引言

GM(1,1)分解模型和 ARIMA 模型均被普遍应用于水文水资源的模拟中,并取得了良好的效果。郭纯青基于岩溶地下水系统的不确定性,将灰色系统理论和方法应用于岩溶地下水资源的评价中,为我国岩溶地下水评价提供了新方法^[1];夏军将水资源系统看作一个部分信息已知,部分信息未知的灰色系统,提出了灰色系统水文学^[2]。近年来,灰色系统理论的不

断发展^[3-7],为其在岩溶地下水系统研究中的中应用创新提供了条件。为了研究中国柳林泉域的降水量和泉水流量的关系,Hao 等提出了灰色系统 GM(1,2)时滞模型和 GM(1,1)分解模型^[8-9],并且应用灰色系统模型分别研究了气候变化与人类活动对中国岩溶泉流量衰减的影响^[10-11]。与此同时,ARIMA 模型在水文水资源中的应用也日臻成熟,Durdu 和 Mishra 等分别用 ARIMA 模型预测了河流的干旱情况^[12-13];Carlos 运用 ARIMA 模型预测地下水中硝

基金项目:国家自然科学基金(40972165, 40572150)、天津市科技发展战略研究计划项目 09JCYBJC27500、天津市水资源与水环境重点实验室开放基金 52XS1015

第一作者简介:李华敏(1986—),女,硕士研究生,主要从事 GIS 在水文水资源中的应用研究。

通讯作者:郝永红(1964—)男,教授,博士生导师,主要研究方向为水资源系统分析。E-mail: haoyhong@yahoo.com;haoyh@sxu.edu.cn。

收稿日期:2011-03-07

酸盐污染浓度,为地下水污染控制提出了一种新方法^[14];Mohan 等利用 ARIMA 模型对季风气候下,流入水库的月流量进行了建模预测^[15];吴志峰等采用 ARIMA 对水生态足迹计算结果进行了模型拟合,并预测了随后 5 年的水生态足迹^[16]。

那么,GM(1,1)分解模型和 ARIMA 模型在水文水资源中应用效果有何差异?它们各自的优势在哪里?适用条件是什么?本文分别运用 GM(1,1)分解模型和 ARIMA 模型研究了气候变化和人类活动对柳林泉流量衰减的影响,实证分析了这两种模型的应用效果,为其日后在水文水资源中的应用提供参考。

1 灰色系统 GM(1,1)分解模型与 ARIMA 模型简介

1.1 灰色系统 GM(1,1)分解模型

灰色系统 GM(1,1)分解模型是将时间序列(如泉水流量序列)分解为趋势量、周期波动量和随机变化量的 GM(1,1)扩展模型。通过 GM(1,1)模型模拟时间序列的趋势量;通过 GM(1,1)模型残差的周期分析获得周期波动量;合并趋势量与周期波动量得到时间序列的周期修正量,然后,通过对周期修正量的残差分析获得随机变化量。最后通过合并趋势量、周期修正值和随机变化量得到预测模型^[9-10]。

1.1.1 时间序列的趋势量

利用 GM(1,1)模型分析时间序列的趋势量,首先对原始数据 $X^{(0)} = (X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(n))$ 进行一次累加生成,获得新的数据列: $X^{(1)} = (X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(n))$,它弱化了数据序列的随机性,增加了规律性,然后建立如下所示的微分方程:

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

式中: a 为系统的发展系数; u 为灰色作用量。

求解该微分方程得到时间响应函数:

$$\hat{x}^{(1)}(t+1) = (x^{(1)}(0) - \frac{u}{a})e^{-at} + \frac{u}{a}; t=1, 2, \dots, n \quad (2)$$

再进行累减还原得到时间序列的趋势量:

$$\hat{x}^{(0)}(t) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t), t=2, \dots, n \quad (3)$$

1.1.2 时间序列的周期波动量

时间序列的周期波动量是对 GM(1,1)模型的残

差进行周期分析得到的,依次分段选择不同长度的周期和变幅,用正弦(或余弦)曲线拟合残差序列,计算公式为:

$$S(t) = A_i \sin \frac{2\pi [t - \sum_{j=1}^i (j-1) T_j]}{T_i} \quad (4)$$

式中: t 满足 $[\sum_{j=1}^i (j-1) T_j] \leq t \leq [\sum_{j=1}^{i+1} (j-1) T_j]$; $S(t)$ 为残差第 i 周期 t 时刻的残差模拟值; A_i 为第 i 周期的最大变幅; T_i 为第 i 周期的长度。对残差序列拟合最好的正弦(或余弦)的周期和振幅就是时间序列的波动周期和最大变幅。

然后,将残差拟合值分别叠加到同一时刻的 GM(1,1)模型的趋势量 $\hat{x}^{(0)}(t)$ 上,即获得时间序列的周期修正值 $\hat{x}^{(0)}(t)$:

$$\hat{x}^{(0)}(t) = \hat{x}^{(0)}(t) + S(t) \quad (5)$$

1.1.3 时间序列的随机波动性

时间序列的随机波动性可以由以下公式得到:

$$R(t) = x^{(0)}(t) - \hat{x}^{(0)}(t) \quad (6)$$

最后合并趋势量、周期波动量和随机变化量获得时间序列的变化过程:

$$Q(t) = \hat{x}^{(0)}(t) + S(t) \pm |R(t)| \quad (7)$$

式中: $Q(t)$ 为模拟的时间序列范围。

由于随机波动量由其绝对平均值确定,所以,公式(7)得到的模拟值是一个范围,用其上限和下限来表达。

1.2 差分自回归移动平均模型(ARIMA)

ARIMA 模型适用于模拟线性动态结构,描述时间序列与滞后序列之间的线性关系,用于线性预测^[17],利用过去自身的时间序列值以及当前和过去的相关时间序列值进行预测。正如现在的岩溶泉流量受过去的降水量和泉水流量的共同影响,我们可以使用 ARIMA 模型去拟合它。ARIMA 模型包括一类由自回归模型(AR)与移动平均模型(MA)组成的平稳自回归移动(ARMA)模型^[17]。

P 阶自回归模型(AR(p))被定义为:

$$Y_t = b_1 Y_{t-1} + b_2 Y_{t-2} \dots + b_p Y_{t-p} + \epsilon_t \quad (8)$$

式中: $\{Y_t\}$ 为 t 时刻的时间序列(比如泉水流量); $\{b_i, i=1, 2, \dots, p\}$ 为需要估计的参数; $\{\epsilon_t\}$ 为随机干扰误差也称为白噪声序列,并且 $\{\epsilon_t\} \sim WN(0, \sigma^2)$, 也就是 ϵ_t 的方差为 σ^2 , 期望为 σ^2 , 并且 $\{\epsilon_t\}$ 不相关。

AR 模型也可以表示为:

$$\varphi(B)Y_t = \epsilon_t \quad (9)$$

式中： $\varphi(B)$ 为自回归算子，由滞后算子多项式表示： $\varphi(B)=1-b_1B\cdots-b_pB^p$ ，其中 B 称为滞后算子。公式(8)和(9)用线性回归的方式将现在的状态 Y_t 用过去 p 期的状态 $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ 表示。

q 阶移动平均模型(MA(q))被定义为：

$$Y_t = \varepsilon_t + a_1\varepsilon_{t-1} + a_2\varepsilon_{t-2} \cdots + a_q\varepsilon_{t-q} \quad (10)$$

式中： $\{a_i, i=1, 2, \dots, p\}$ 为需要估计的参数。

MA(p)也可以用下列形式表示：

$$Y_t = \theta(B)\varepsilon_t \quad (11)$$

式中： $\theta(B)$ 为移动平均算子，由滞后算子多项式表示： $\theta(B)=1-a_1B\cdots-a_pB^p$ 。

MA模型描述的是一个白噪声过程的移动平均时间序列。因为没有时间序列的将来值，我们不能得到将来的残差值，所以，只能用MA(q)模型预测未来 q 期的值。

合并AR和MA模型就得到了常用的平稳自回归模型(ARMA)，它可以表示为：

$$Y_t = b_1Y_{t-1} + b_2Y_{t-2} \cdots + b_pY_{t-p} + \varepsilon_t + a_1\varepsilon_{t-1} + a_2\varepsilon_{t-2} \cdots + a_q\varepsilon_{t-q} \quad (12)$$

ARMA(p, q)也可以用下列形式表示：

$$\varphi(B)Y_t = \theta(B)\varepsilon_t \quad (13)$$

式中： $p, q \in N$ ，(p, q)称为模型的阶数，可以表示为 $\{Y_t\} \sim \text{ARMA}(p, q)$ 。

ARMA模型只针对平稳的时间序列，平稳的时间序列反映了一些随时间不变的特性，是做统计推断的必备条件。但是，现实中的时间序列并不都是平稳的，例如具有时间趋势(比如缓慢的增长)或者具有周期性的时间序列。差分变换是一种去除趋势和季节性周期的简单而有效的方法。对原始数据进行差分变换得到的模型称为单整自回归移动模型(ARIMA)。时间序列 $\{Y_t\}$ 称为 p, d, q 下的单整自回归移动模型，记为 $\{Y_t\} \sim \text{ARIMA}(p, d, q)$ ，其中差分序列 $Z_t = (1-B)^d Y_t$ ，满足平稳性， $d \in N$ 。

除了过去的时间序列值和过去的误差，我们还可以在模型中使用输入序列(其他序列过去和现在的值)。这种带有输入序列的模型被称为ARIMAX模型，它通常表示为：

$$\varphi(B)(Y_t - \mu - \sum_i \psi_i(B)X_{i,t}) = \theta(B)\varepsilon_t \quad (14)$$

式中： Y_t 为响应序列或者差分变换后的响应序列， μ 为中项， $\psi_i(B)$ 为转换函数，可以表示为： $\psi_i(B)=1-c_{i1}B\cdots-c_{i_p}B^p$ ， $\{X_{i,t}\}$ 代表第 i 个输入序列。包含一个输入序列的ARIMAX模型为：

$$(1-b_1B\cdots-b_pB^p)(Y_t - \mu - c_1X_{r-1} - c_2X_{r-2} \cdots - c_sX_{r-s}) = \varepsilon_t + a_1\varepsilon_{t-1} + a_2\varepsilon_{t-2} \cdots + a_q\varepsilon_{t-q} \quad (15)$$

式中 $\{c_r, r=1, 2, \dots, s\}$ 是需要估计的参数。

2 柳林泉流量的GM(1,1)分解模型与ARIMA模型

2.1 柳林泉概况

柳林泉是我国北方的岩溶大泉，泉水出露于山西省柳林县城以东约3 km处的三川河两岸及河底。泉源出露区东西长2.4 km，出露高程790~801 m，泉域总面积约5 100 km²，见图1。1957—2009年平均降雨量为515 mm，60~70%的降水发生在7—9月，同期，柳林泉实测平均流量为2.96 m³/s。

柳林泉域的地质单元属鄂尔多斯台向斜东翼吕梁背斜西翼单斜构造。区内东部为中高山区，一般高程1 200~1 500 m，大部分地区基岩出露，由北至南均有灌木植被。西部为中低山丘陵区，高程800~1 200 m，相对高程200 m左右。区内黄土广布，植被稀疏，地面分割破碎，水土流失严重，沟底大部分有基岩出露。中间夹离石—中阳山间盆地，盆地面积443 km²。区内总的地势东高西低，由北、东、南向中部倾斜。

柳林泉属寒武、奥陶系碳酸盐岩溶裂隙水，主要补给来源为大气降水，通过泉域裸露的碳酸盐岩裂隙入渗。在三川河三条支流(北川河、南川河、东川河)河谷有六个碳酸盐岩渗漏段，长33 km，河水通过渗漏补给柳林泉。岩溶地下水由北、东、南三个方向向柳林县的三川河谷集中，受石炭、二叠纪隔水层阻挡，溢流成泉，为侵蚀溢流泉。

从20世纪50年代开始，柳林泉流量总体上呈现出减少的趋势，特别是到了70年代以后，工业和农业的需水量大大增加，水库的修建，泉域周边地下水的超采对泉域地下水产生袭夺，矿坑排水，生态环境的破坏等等，在这些因素的综合作用下，柳林泉域的地下水位逐渐下降，泉流量变化呈现出以衰减为主的特征。针对以上问题，学者们从不角度对柳林泉域地下水进行了研究，取得了可喜的成果。吴吉春等先后建立了柳林泉局部溶质运移二维数值模型、裂隙发育区运移三维数值模型和二维区域地下水渗透模型，并将这些模型用于模拟柳林电厂水源地对柳林泉造成的影响^[18-20]。Wang等利用地质统计和地球化学分析

方法对柳林泉域地表水渗漏过程进行了分析,为区域地下水保护提供了基础资料^[21]。马腾采用灰色系统中的 GM(1,1)等维递补预测模型对柳林泉的渗透量进行了建模预测,并分析了该模型在预测中的准确性^[22]。

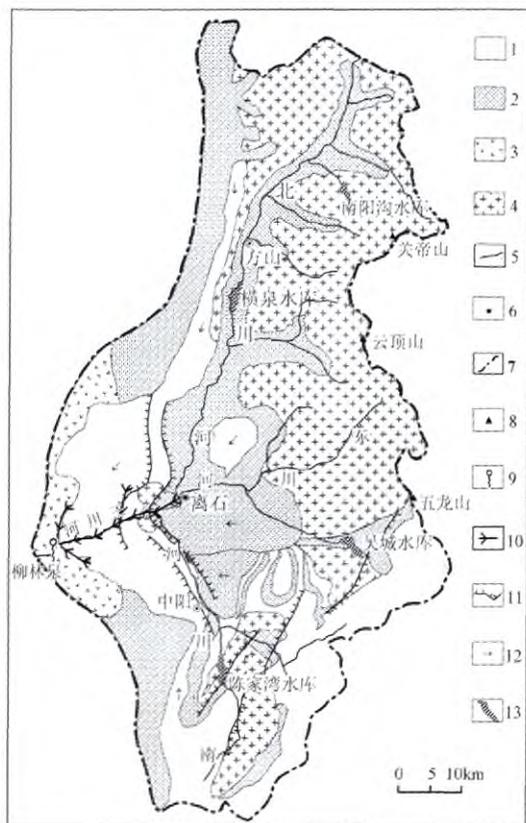


图 1 柳林泉域水文地质图

Fig. 1 Hydrogeologic map of the Liulin Spring Basin

- 1. 奥陶系石灰岩; 2. 第四系黄土沉积层; 3. 石炭二叠系砂岩; 4. 太古代变质岩; 5. 河流; 6. 城镇; 7. 泉域范围边界; 8. 山峰; 9. 泉; 10. 地下水强径流带; 11. 断层; 12. 地下水径流方向; 13. 水库

2.2 柳林泉流量序列的分时段特征

从 20 世纪 50 年代开始,柳林泉流量呈现逐年减少的趋势,见图 2。仔细分析柳林泉流量变化和岩溶地下水开采过程,可以将泉水流量划分为两个阶段,主要依据如下:

从经济技术条件上讲,20 世纪 70 年代之前,柳林泉域是一个极其不发达的地区,经济落后,人口稀少。根据柳林泉域地质条件和开采技术条件,柳林泉域岩溶地下水的开采需要开凿 150~450 m 左右的深井,成井后地下水位平均埋深在 100 m 以下^[23]。20 世纪 70 年代之前,当地没有经济实力引进开凿深井的现代化的钻井和抽水设备,在这一阶段人们主要利用地表水和松散层地下水。1973 年之后随着地方经济的发展,逐渐开发利用深层岩溶地下水(图 2),尤其是随着改革开放,柳林泉域的经济和人口蓬勃发展(图 3),对深层岩溶地下水的开采不断增强。因此,柳林泉水文过程可分为两个时期。第一阶段(1957—1973 年),柳林泉处于自然状态,气候变化是影响岩溶地下水系统的主要因素,人类活动的影响可以忽略。在第二阶段(1974—2009 年)岩溶地下水系统受气候变化和人类活动的双重影响。于是,我们采用 1957—1973 年柳林泉流量序列,运用灰色系统 GM(1,1)分解模型和 ARIMAX 模型分别建立柳林泉自然状态下的模拟模型,继而,将模型外推得到 1974—2009 年柳林泉自然条件下的流量。然后,根据水量平衡原理,减去同期实测流量,获得人类活动对泉水流量衰减的贡献。由此可见,采用分段分析法可以将气候变化和人类活动对柳林泉流量衰减的贡献区别开,并且量化了 1974—2009 年气候变化和人类活动对柳林泉流量衰减的贡献。我们通过实证比较灰色系统 GM(1,1)分解模型和 ARIMA 模型的结果,分析了两种模型的准确性和适用条件。

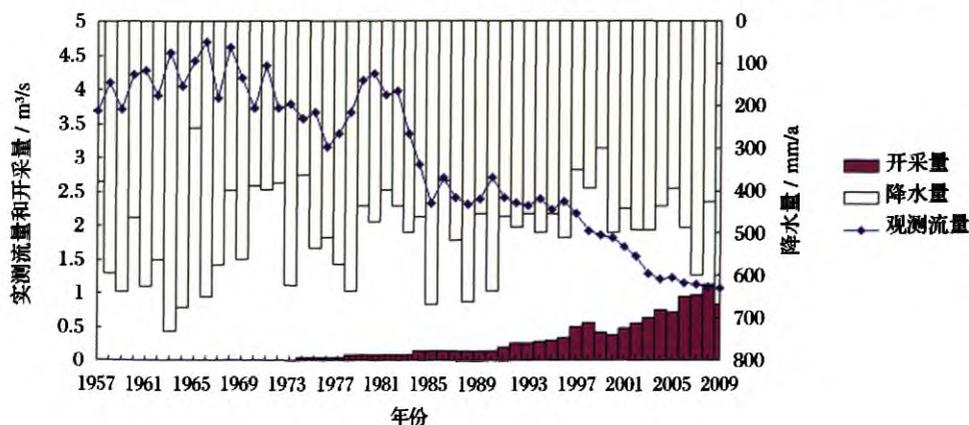


图 2 柳林泉年降水量、实测流量和开采量

Fig. 2 Annual precipitation, measured discharge and pumping yield in Liulin Spring Basin

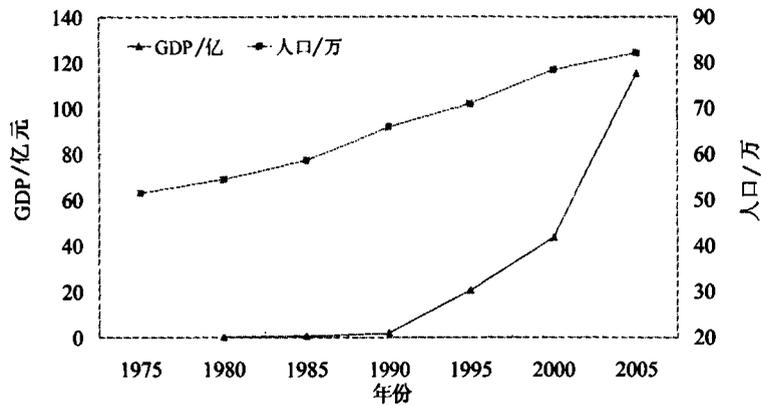


图3 柳林泉域人口和GDP发展状况

Fig. 3 Development of population and GDP in Liulin Spring Basin

2.3 GM(1,1)分解模型对柳林泉流量的模拟

灰色系统理论认为影响岩溶地下水系统的因素众多,包括气象条件、地表植被、地形地貌、地质条件、社会经济发展水平、科技水平等。因此,地下水系统是一个部分信息已知,部分信息未知的灰色系统,可以运用灰色系统理论与方法对它进行研究。运用灰色系统GM(1,1)分解模型对柳林泉流量进行模拟过程如下^[24]。

2.3.1 柳林泉流量的趋势量

将1957-1973年的柳林泉流量作为原始数据,建立GM(1,1)模型,求得预测模型为:

$$x^{(1)}(t+1) = -2422.794e^{-0.001723751t} + 2426.576 \quad (16)$$

由公式(3)就可以获得趋势量 $\hat{x}^{(0)}(t)$ 。

2.3.2 柳林泉流量的周期量

GM(1,1)残差曲线见图4。从图4中可以看出,由于边界效应的影响1957到1961年不是一个完整的周期。1961年以后表现为一个完整周期,振幅为0.14 m³/s,周期长度为16年。将以上参数代入公式(4)中,得到柳林泉流量的周期波动量为:

$$S(t) = 0.14 \sin \frac{2\pi [t - 1961]}{16} \quad (17)$$

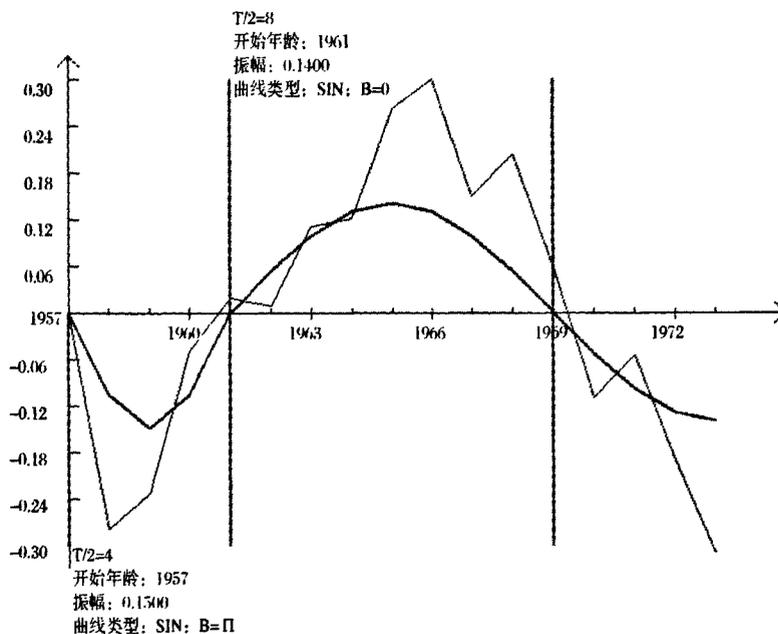


图4 柳林泉流量的GM(1,1)残差周期拟合曲线

Fig. 4 Fitted curve of GM(1,1) residual of the Liulin Spring flow

2.3.3 柳林泉流量的随机波动量

由公式(6)得到随机波动变化量的平均值 $|R(t)|$ 为 $0.08 \text{ m}^3/\text{s}$ 。

根据公式(7),我们将以上获得的趋势量、周期波动量和随机波动变化量叠加,就可以得到1957—1973年柳林泉的GM(1,1)分解模型,它模拟了柳林泉在自然状态下的泉水流量变化过程。使用该模型继续推算,就得到1974—2009年柳林泉在自然状态下的流量。根据水量平衡原理,从柳林泉1974—2009年自然状态下的流量中减去实际观测流量,得到人类活动造成的泉水流量的衰减,结果见图5。

2.4 ARIMA 模型对柳林泉流量的模拟

大气降水的一部分入渗地下,进入包气带,继续下渗进入地下水饱和带形成压力波,地下水压力波在含水层中传播,当压力波传播到适当的地点,地下水从含水层流出,形成泉水。因此,岩溶泉流量是地下水系统对区域降水量的响应。随机理论认为,泉水流量和降水量是两个随机变量,它们之间的关系可以通过ARIMA模型进行研究。本文利用Statistics Analysis System(SAS)对数据进行分析建立ARIMA模型,将降雨量作为输入时间序列,泉水流量作为响应时间序列。其分析过程经过平稳性检验、参数估计和诊断分析、建立预测模型三个阶段。

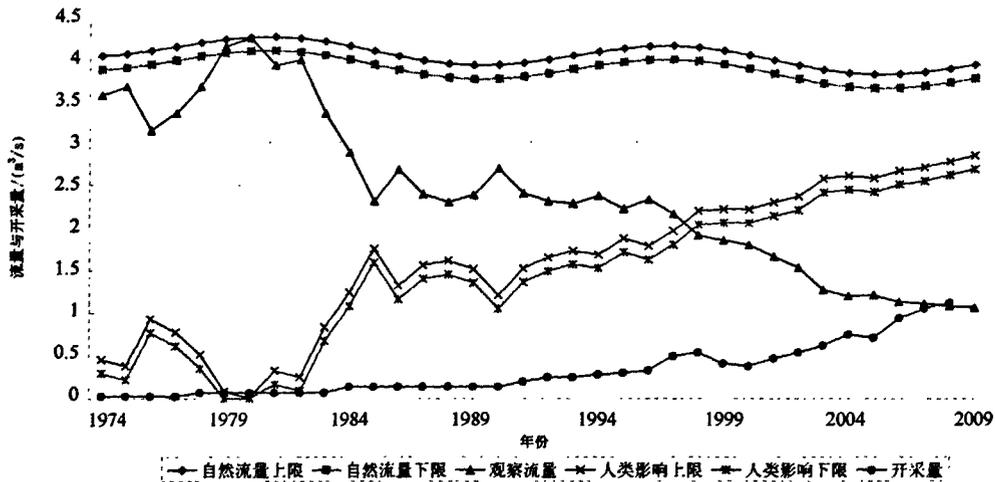


图5 GM(1,1)分解模型获得的人类活动对柳林泉流量衰减的贡献

Fig.5 The contribution of human activities on Liulin Spring discharge decline by GM(1,1) decomposition model

2.4.1 平稳性检验

我国北方岩溶泉流量稳定,属于稳定性的

泉^[23,25]。对柳林泉时间序列平稳性检验,结果(表1)说明柳林泉的流量是平稳的。

表1 1957—1973年柳林泉流量的平稳性检验

Tab.1 Stationarity test to the Liulin Spring flow during the first period (1957—1973)

类型	滞后阶数	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
无截距项和无趋势项	0	-0.097 1	0.643	-0.19	0.599 3		
	1	-0.061 8	0.649 7	-0.28	0.567 3		
	2	0.006 5	0.665 4	0.07	0.689		
只有截距项	0	-18.992 9	0.001 1	-4.62	0.002 6	10.69	0.001
	1	-12.965 9	0.021 2	-2.04	0.269 8	2.1	0.565 2
	2	-3.386 7	0.567 2	-1.17	0.653 5	0.7	0.889 4
有截距项和趋势项	0	-18.758	0.011 5	-4.4	0.015 9	10.17	0.006 4
	1	-10.849 4	0.231 5	-1.75	0.677 7	2.37	0.718 5
	2	-1.051 6	0.976 8	-0.76	0.944 7	12.41	0.001

2.4.2 参数估计和诊断分析

参数估计与诊断分析的目的是找到1957—1973年泉水流量的最佳模型参数以及相应的置信区间。用ARIMA模型的不同形式去拟合泉水流量,并根据不同模型的准确性进行选择,从而得到最理想的预测模型。

因为MA(q)模型只能预测未来q期值,而AR(p)可以实现滚动预测,所以最好的模型应该包含AR(q)模型。因此,用于拟合的模型应该是AR,ARMA和ARMAX模型。AR(3)和AR(4)系数的估计值在模型一(AR)中是显著的,但是它们在模型二、

三、四(ARMA)中是非显著的(见表2)。所以模型二、三、四不适合用于预测。但可以通过在ARMA模型中使用降水量建立ARMAX模型,寻找最合适的预测模型。在模型五(ARMAX)中AR(1), X_{t-2} 和 X_{t-4} 是显著的,它的信息准则(AIC)和贝叶斯准则(SBC)是五个模型中最小的,而且其相对误差仅为2.4%(用AIC和SBC来判断模型拟合好坏,它们越小,拟合精度越高)(表2)。结果显示模型五能够很好的拟合原始序列,且残差的自相关检验说明残差序列是白噪声(表3)。

表2 柳林泉流量的ARIMA模型回归结果
Tab.2 Results of ARMA regressions for the Liulin Spring discharge

参数	模型一 (AR)	模型二 (ARMA)	模型三 (ARMA)	模型四 (ARMA)	模型五 (ARMAX)
b_1					-0.660 47 * (0.314 18)
b_2					
b_3	0.385 74 * (0.217 33)	0.391 54 (0.248 07)	0.315 90 (0.263 89)		
b_4	-0.636 32 ** (0.231 90)	0.593 53 (0.438 24)		-0.253 60 (0.454 69)	
a_1					
a_2		-0.619 12 * (0.314 43)			
a_3					
a_4		-0.047 13 (0.493 12)	0.774 48 *** (0.219 55)	0.728 42 * (0.399 13)	
c_1					
c_2					0.002 130 9 *** (0.000 184 1)
c_3					
c_4					0.000 642 6 *** (0.000 181 5)
AIC	4.930 994	5.601 435	4.401 645	5.263 81	-13.147 6
SBC	7.430 634	8.101 075	6.901 285	7.763 45	-10.887 8

*** 显著性水平为0.01; ** 显著性水平为0.05; * 显著性水平为0.1。

表3 模型五的残差白噪声检验
Tab.3 Test to residuals for model 5

To Lag	Chi-Square	DF	Pr > ChiSq	自相关函数					
6	7.89	5	0.162 2	0.096	-0.234	0.034	0.378	0.002	-0.362
12	10.64	11	0.474	-0.185	0.041	-0.108	-0.087	-0.072	-0.003

综上所述,模型五适合用于对柳林泉流量做预测。该模型可以用下面公式表示:

$$(1 + 0.066047B)(Y_t - 2.60123 - 0.0021309X_{t-2} - 0.0006426X_{t-4}) = \epsilon_t \quad (18)$$

它也可以用以下形式表示:

$$Y_t = 4.319264 - 0.66047Y_{t-1} + 0.0021309X_{t-2} +$$

$$0.001407X_{t-3} + 0.0006426X_{t-4} + 0.000424X_{t-5} + \epsilon_t \quad (19)$$

公式(18)中的AR(1)系数的估计值在10%的显著性水平下显著, X_{t-2} 和 X_{t-4} 系数的估计值在1%的显著性水平下显著。公式(19)是根据柳林泉流量和降水量建立的ARIMAX模型,它说明当前年份的泉水

流量与过去一年的泉水流量以及过去 2~5 年的降水量有关。

2.4.3 预测

利用公式(18)我们预测到了第二阶段在自然条件下的柳林泉流量(1974-2009),从中减去柳林泉流量的实际观测值,就得到泉域内的人类活动造成的泉

水流量的衰减量,即

$$Y_{human} = Y_t - Y_{observation} \quad (20)$$

式中: Y_{human} 表示人类活动造成的泉水流量衰减量; $Y_{observation}$ 表示泉水流量观测值。

计算结果见图 6,图中给出了置信概率为 95%时的泉水流量的置信区间。

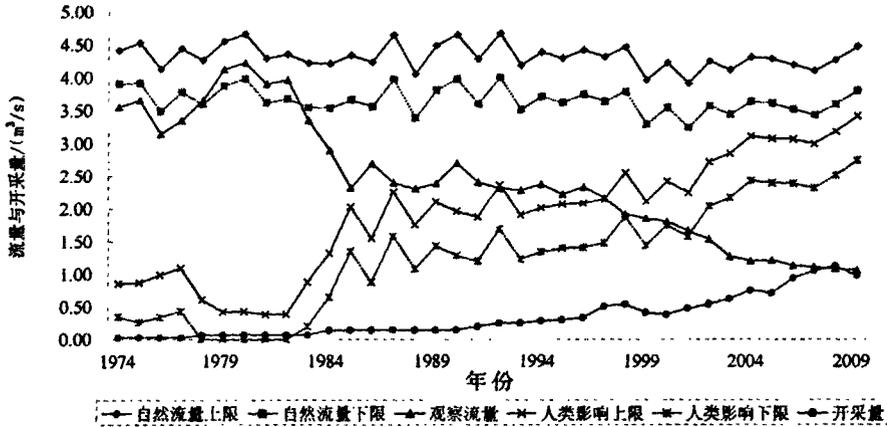


图 6 ARIMA 模型获得的人类活动对柳林泉流量衰减的贡献

Fig. 6 The contribution of human activities on Liulin Spring discharge decline by ARIMA model

3 结果与讨论

GM(1,1)分解模型,从系统分析的角度研究地下水的水文循环过程。在这里泉水流量是系统特征变量,通过建立系统特征变量的动态变化模型来完成对系统的描述。由图 5 可见泉水流量具有总体下降的趋势,同时由图 4 显示泉水流量具有 16 年的周期变化。ARIMAX 模型把降水作为输入,泉水流量作为输出,它反映降水转化为泉水流量的机理,即岩溶化程度高的区域,含水层的岩溶管道畅通,是地下水运移的强径流带,在这些区域地下水传播速度快;在岩溶化程度低的区域,岩溶管道不畅,地下水运移速度慢。公式(19)显示泉水流量与过去第 2~5 年的降水量有关系,反映出柳林泉域岩溶化程度具有较大差异。根据降水量系数可以推断出过去第 2 年的降水量对柳林泉流量的贡献最大,过去第 3 年和第 5 年的降水的贡献其次,过去第 4 年的降水的贡献最小。

图 5 列出了 GM(1,1)分解模型的模拟结果,在自然条件下,泉水流量呈现出在波动中平稳下降的趋势,泉水流量从 20 世纪 70 年代到 21 世纪初下降了 0.23 m³/s,同期由于人类的活动导致的泉水流量下降了 2.03 m³/s,两者之和 2.26 m³/s 即为泉水流量的总衰减量,其中气候变化对泉水流量衰减的贡献占总衰减量的 10%,人类活动对泉水流量衰减的贡献占总衰减量的 90%。同样地,图 6 列出了 ARIMAX 模型的模拟结果,在自然条件下,泉水流量呈现出平稳下降的态势,在自然状态下泉水流量从 20 世纪 70

年代到 21 世纪初下降 0.20 m³/s,同期人类活动造成的泉水流量下降量为 2.12 m³/s,两者相加为 2.32 m³/s,它代表泉水流量的总衰减量,其中气候变化对泉水流量衰减的贡献占总衰减量的 8.6%,人类活动对泉水流量衰减的贡献占总衰减量的 91.4%。从泉水流量的观测数据得知泉水流量的实际总衰减量为 2.27 m³/s,灰色系统 GM(1,1)分解模型的误差为 0.01 m³/s;ARIMAX 模型的误差为 -0.05 m³/s。通过实证比较可见,灰色系统 GM(1,1)分解模型和 ARIMA 模型都具有比较高的精度,在柳林泉流量模拟中所得结果非常接近,互相验证了模拟结果的准确性和可靠性。对比人类活动和气候变化对柳林泉流量衰减的贡献发现,人类活动是气候变化贡献的 8~9 倍,由此可见,人类活动是影响柳林泉流量减少的主要因素。

在柳林泉域,岩溶地下水开采、煤矿采煤排水和河流上建坝等人类活动,直接减少了地下水的补给量,造成泉水流量减少。表 4 为不同年代地下水开采量占人类活动造成的泉水衰减量的比例。由表 4 可见,随着时间的推移地下水开采量所占人类活动对流量衰减的百分比在不断的增加,两种模型推算得到的开采量占人类活动对泉水流量衰减贡献的百分比相近,分别从最初的 9.60%和 8.78%上升到了 29.46%和 29.84%,这些再次验证了两种模型在水文水资源分析中的正确性。同时,结果显示随着经济的发展人们对地下水的开采强度在日益增强。

表4 地下水开采量占人类活动造成泉水衰减量的百分比
Tab. 4 The occupancy of groundwater exploitation in the spring discharge depletion by human activities

时间	GM(1,1)分解模型	ARIMA模型
20世纪70年代	9.60%	8.78%
20世纪80年代	11.53%	11.57%
20世纪90年代	18.79%	18.11%
21世纪初	29.46%	29.84%

由灰色系统GM(1,1)模型可以看出,它适用于累加生成后服从指数规律的时间序列,见公式(1)和(2)。我国北方岩溶泉流量相对稳定,由于降水的丰枯变化,致使泉水流量呈现出一定的周期波动,因此,在运用灰色系统GM(1,1)模型时,需要采用正弦(或余弦)函数去拟合周期波动性,这样可以提高模型精度。ARIMA模型在泉水流量拟合中,首先对柳林泉流量进行了平稳性检验,之后它考虑了降水补给地下水在时间上的滞后效应,实际是把泉水流量看作若干年前降水的线性迭加,它同样比较准确地模拟出了降水与泉水流量的定量关系。

在对长时间序列建模过程中,通过分析不同阶段的特性,分时段进行模拟可以得到更多信息。本文根据流量序列的特征,将它分解为2个阶段,继而分别计算出气候变化和人类活动对泉水流量衰减的贡献量,比较客观地揭示出了泉流量的变化特征,对柳林泉的开发保护有积极的参考意义。

4 结论

(1) GM(1,1)分解模型和ARIMA模型的理论基础不同,前者从系统论的角度,将岩溶地下水看作一个不确定系统,泉水流量是系统的特征变量,通过对系统特征变量的发展态势的动态模拟获得泉水流量的变化规律;后者是将泉水流量和降水量当作随机变量,通过研究它们之间的统计规律,使泉水流量得到定量的描述。虽然两个理论的出发点不同,但都得到了相近的结果,都对柳林泉流量进行了准确描述。由此可见,GM(1,1)分解模型和ARIMA模型都是研究岩溶地下水的有效手段。

(2) GM(1,1)分解模型和ARIMAX模型对柳林泉20世纪70年代以来衰减量的模拟结果分别为 $2.26\text{m}^3/\text{s}$ 和 $2.32\text{m}^3/\text{s}$,与同期实际衰减量 $2.27\text{m}^3/\text{s}$ 的相对误差百分比分别为0.44%和2.20%,具有较高的精度,反映出这两种模型都适合泉水流量的模拟。同时,两模型分别计算了20世纪70年代、80年代、90年代和21世纪初地下水开采量占人类活动影响的百分比,结果相当吻合,再次证明这两种模型对岩溶泉流量的模拟具有很好的效果。

(3) GM(1,1)分解模型对于指数序列的模拟精确度较高,对于具有周期波动的泉水流量,可以通过周期修正达到精度要求;ARIMA模型能够较好地反映泉水流量相对于降水量的时间滞后效应,因此,ARIMA模型能够比较准确地模拟出泉水流量与降水的量化关系。

(4) 将分时段建模方法应用于柳林泉流量模拟过程中,将柳林泉流量分为2个阶段,根据不同时段的特征分析研究,成功地将气候变化和人类活动对柳林泉流量衰减的贡献分别模拟出来。因此,分时段分析法可以增加时间序列分析的信息量,是信息挖掘的一个有效手段。

(5) GM(1,1)分解模型和ARIMA模型对柳林泉的模拟结果显示,人类活动对泉水流量衰减的贡献是气候变化贡献的8~9倍,这使我们清楚地认识到人类活动是造成我国北方岩溶泉流量衰减的主要原因,不能把所有的责任都归咎于气候变化,用气候变化作为挡箭牌来开脱我们的责任。从全球意义上讲,人类活动是造成气候变化的主要因素之一,但是就柳林泉域而言,这里的人类活动是造成泉水流量衰减的主要原因。

(6) 柳林泉作为我国北方典型的岩溶泉,代表了我国北方岩溶地下水的开发利用过程,20世纪50年代到60年代末,我国北方岩溶地下水基本处于自然状态,维持着良好的水文循环。从20世纪70年代以来,开采量逐年增大造成地下水水位下降,泉水流量减少,甚至一些泉已断流(例如晋祠泉于1994年断流)。因此,加强我国北方岩溶地下水系统的管理与保护迫在眉睫。

参考文献

- [1] 郭纯青. 岩溶地下水评价的灰色系统理论与方法研究[M]. 北京:地质出版社,1993,4-8.
- [2] 夏军. 灰色系统水文学[M]. 武汉:华中理工大学出版社,2000,24-53.
- [3] 刘思峰,党耀国,方志耕,等. 灰色系统理论及其应用[M]. 北京:科学出版社,2004.
- [4] 肖新平,宋中民,李峰. 灰技术基础及其应用[M]. 北京:科学出版社,2005.
- [5] 王学萌,聂宏声. 灰色系统模型在农村经济中的应用[M]. 武汉:中理工大学出版社,1989.
- [6] 王学萌. 等维灰数递补动态预测[J]. 华中工学院学报,1989,4:9-16.
- [7] 张岐山. 灰朦胧集的差异信息理论[M]. 北京:石油工业出版社,2002.
- [8] Hao Y H, Yeh T-C J, Gao Z Q, et al. A Gray System Model for Studying the Response to Climatic Change: the Liulin Karst Spring, China[J]. Journal of Hydrology, 2006,328(3-4): 668-676.
- [9] Hao Y H, Yeh T-C J, Wang Y R, et al. Analysis of Karst Aq-

- uifer Spring flows with a Gray System Decomposition Model [J]. *Ground Water*, 2007, 45(1): 46–52.
- [10] 郝永红, 王玮, 王国卿, 等. 气候变化及人类活动对中国北方岩溶泉的影响[J]. *地质学报*, 2009, 83(1): 139–144.
- [11] Hao Y H, Wang Y J, Zhu Y E, et al. Response of Karst Springs to Climate Change and Anthropogenic Activities; the Niangziguan Springs, China[J]. *Progress in Physical Geography*, 2009, 33(5): 634–649.
- [12] Omer Faruk Durdu. Application of Linear Stochastic Models for Drought Forecasting in the Buyuk Menderes River Basin, Western Turkey[J]. *Stoch Environ Res Risk Assess*, 2010, DOI 10.1007/s00477-010-0366-3.
- [13] Mishra A K, Desai V R. Drought Forecasting using Stochastic Models[J]. *Stoch Environ Res Res Risk Assess*, 2005, 19, 326–339.
- [14] J. Carlos García Díaz. Monitoring and Forecasting Nitrate Concentration in the Groundwater using Statistical Process Control and Time Series Analysis; a Case Study[J]. *Stoch Environ Res Res Risk Assess*, 2010, DOI 10.1007/s00477-010-0371-6.
- [15] S. Mohan and S. Vedula. Multiplicative Seasonal ARIMA Model for Longterm Forecasting of Inflows[J]. *Water Resources Management*, 1995, 9: 115–126.
- [16] 吴志峰, 胡永红, 李定强, 等. 城市水生态足迹变化分析与模拟[J]. *资源科学*, 2006, 5(28): 152–156.
- [17] Fan J, and Yao Q. *Nonlinear time series: Nonparametric and parametric 517 methods* [M]. New York, Springer-Verlag, 2003.
- [18] 吴吉春, 薛禹群, 黄海, 等. 山西柳林泉局部区域溶质运移二维数值模拟[J]. *水利学报*, 2001, (8): 38–43.
- [19] 吴吉春, 薛禹群, 黄海, 等. 山西柳林泉裂隙发育区域溶质运移三维数值模拟[J]. *南京大学学报*, 2000, 36(6): 728–734.
- [20] 吴吉春, 薛禹群, 黄海, 等. 山西柳林泉域地下水水流数值模拟[J]. *水文地质工程地质*, 2001, (2): 18–20.
- [21] Wang Y, Ma T, Luo Z. Geostatistical and geochemical analysis of surface water leakage into groundwater on a regional scale: a case study in the Liulin karst system, northwestern China[J]. *Journal of hydrology*. 2001, 246: 223–234.
- [22] 马腾, 王国卿, 张庆保. 山西省柳林泉域地表径流渗透量的灰色预测[J]. *地球科学——中国地质大学学报*, 1997, 22(1): 90–93.
- [23] 韩行瑞, 鲁荣安, 李庆松, 等. 山西岩溶大泉研究[M]. 北京: 地质出版社, 1993, 294–305.
- [24] 王亚捷, 郝永红, 王学萌, 等. 岩溶地下水灰色系统分析软件开发[J]. *中国岩溶*, 2010, 29(4): 389–395.
- [25] 何宇彬, 邹成杰. 中国南北方喀斯特水特征对比[J]. *中国岩溶*, 1991, 15(3): 259–268.

Comparative study on karst ground water simulation between GM(1,1) decomposition model and ARIMA model: A case study on discharge simulation of the Liulin Spring

LI Hua-min¹, WU Jing², ZHAO Jiao-juan¹, HAO Yong-hong³, WANG Ya-jie³, CAO Bi-bo¹

(1. College of Urban and Environmental Sciences, Tianjin Normal University, Tianjin 300387, China;

2. Department of Statistics, Tianjin University of Finance and Economics, Tianjin 300222, China;

3. Tianjin Key Lab of Water Environment and Water Resources, Tianjin Normal University, Tianjin 300387, China)

Abstract: The discharge of the Liulin spring is simulated with GM(1,1) decomposition model and ARIMA model respectively. According to the hydrological characteristics, the Liulin spring flow series could be divided into two periods. First, from 1957 to 1973 the spring flow was under natural state; second, from 1974 to 2009 the spring flow was impacted by both climate change and human activities. Using the data of first period, the spring flow under the natural state is fitted with GM(1,1) decomposition model and ARIMA model, and then the models are extrapolated to obtain the second periods' spring flow under the natural state. According the water balance principle, the spring flow decrement contributed by human activities is acquired by subtracting the observed discharge from simulated spring flow of the second period under the natural state. Thus, it is differentiated the effects of human activities from climate change. The simulated Liulin Springs' attenuation from 1970s to early 21st century is 2.26 m³/s by GM (1,1) decomposition model and 2.36 m³/s by ARIMA model with the relative error being 0.44% and 2.20% respectively, showing both GM (1,1) decomposition model and ARIMA model are suitable for spring flow simulation. Comparing the effects of human activities and climate change to the depletion of the Liulin Spring's discharge, the authors find that the contribution of human activities is 8 to 9 times higher than that of the climate change. The empirical studies have shown that the GM (1,1) model is of high precision in simulating the exponential series. It can also improve accuracy by periodic amendment, when simulate the spring flow with periodic fluctuations. ARIMA model could reflect time-lag between precipitation and spring discharge and accurately simulate their quantitative relation.

Key words: GM (1,1) decomposition model; ARIMA model; piecewise analysis; karst spring discharge