第13卷第6期
1993年12月

水土保持通报
Bulletin of Soil and Water Conservation

Vol.13 No.6
Dec.,1993

用自回归模型预报枯季径流初探

鲁小清　冯　平

(宁夏回族自治区固原水文水资源勘测大队·固原县·756000)

提　要

该文根据颉河流域三关口水文站25年(1966年至1991年)的观测资料,应用建立的自回归模型,对枯季径流总量及各月平均流量进行预报,明显提高了预报精度。

关键词: 自回归模型　枯季径流　预报

Applying the Autoregressive Model to Predict Runoff in Withered Season

Lu Xiaoqing　*Feng Ping*

(*Guyuan Survey Large Body of Hydrography and Water Resources, Ningxia Autonomous Region of the Hui Nationality, Guyuan, Ningxia* 756000)

Abstract

According to the hydrographical data investigated from Sanguankou hydrologic station in Xiehe basin during 26 years (1966——1991), the mean discharge for every month and the total runoff amount for withered season in this area were predicted with autoregressive model. The predictive precision is improved remarkably.

Key words autoregressive model　runoff in withered season　prediction

自回归模型常简写为AR模型,它是描述时间序列相依特性的数学模型。该模型具有时间相依的非常直观的形式,同时建立模型和具体应用比较简单。因此,在水文学中得到重视并在水资源规划和设计中得到广泛的应用。

在宁夏南部山区,降水与径流分配很不均匀,汛期6~9月,集中了70%左右的降水和50%~90%以上的径流量,汛期径流多为山洪,泥沙甚大,利用受限制。为了防洪和减少淤积,水库多系空库度汛,汛末开始蓄水至次年春夏灌溉。而本地区春旱夏旱十分严重,在农作物生长期有"十年九旱"、"三年两头旱"的农谚。因此,枯季径流对于水库蓄水和利用显然十分重要。

该区的河流,枯季历时长而水量稳定。降水对枯季径流有一定影响,但影响不大。枯季径流主要由地下水补给,而地下水补给与雨季产生的流域蓄水直接关联。每年汛末流量逐渐变小,冬春则主要受气温和土壤冻结等因素影响。考虑枯季径流退水过程流量渐变的相依性,本文试用AR(2)模型作出枯季径流预报方案。

一、AR(2)模型

(一)模型的表达　设时间序列$\{X_t\}$为平稳序列,服从均值为u、方差为σ^2的正态分布,具有自

收稿日期:1993—10—25

回归相关（或时间相依结构），即$\{X_t\}$为自回归序列，则阶数为 2 的自回归模型 AR(2)可表达为：

$$X_t = u + \varphi_1(X_{t-1} - u) + \varphi_2(X_{t-2} - u) + \varepsilon_t \tag{1}$$

式中：ε_t与X_t无关，并且本身是独立随机变量（均值为零、方差为σ_ε^2）；φ_1、φ_2为自回归权重系数。由于独立随机序列ε_t的方差σ_ε^2和序列X_t的方差σ^2存在着一定的关系，所以在 AR(2)模型中的参数有u、σ和φ_1、φ_2。u表示序列的平均水平，σ表示序列X_t围绕均值变化的程度，φ_1、φ_2表示序列在时间上的相依程度。

（二）参数估计　AR(2) 模型中包含有四个基本参数：u、σ、φ_1和φ_2。

其中

$$u = \frac{1}{n}\sum_{i=1}^{n} X_t \tag{2}$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(X_t - u)^2 \tag{3}$$

φ_1和φ_2可用尤尔—沃尔克(Yule－Walker)方程进行矩法估计，$p = 2$时，有

$$\left.\begin{aligned}\rho_1 &= \varphi_1 + \varphi_2\rho_1\\ \rho_2 &= \varphi_1\rho_1 + \varphi_2\end{aligned}\right\} \tag{4}$$

联解可得

$$\left.\begin{aligned}\varphi_1 &= \rho_1(1-\rho_2)/(1-\rho_1^2)\\ \varphi_2 &= (\rho_2-\rho_1^2)/(1-\rho_1^2)\end{aligned}\right\} \tag{5}$$

以样本序列的γ_1和γ_2代替式(5)中的ρ_1和ρ_2，便可获得估计值φ_1和φ_2。除此以外，φ_1和φ_2也可由二元线性回归直接求出。

（三）主要统计特性　一个序列的相依特性用该序列的自相关函数来表示。对于 AR(2)序列，由 Yule－Walker 方程得

$$\rho_k = \varphi_1\rho_{k-1} + \varphi_2\rho_{k-2} \tag{6}$$

其中$\rho_0 = 1$，$\rho_{-k} = \rho_k$，利用式(6)便可递推出 k=1,2,…时的自相关系数ρ_1，ρ_2，…最后获得AR(2)序列的自相关函数。

序列在频率域上的特性以频谱表示。对AR(2)模型，方差谱密度（函数）为

$$S(f) = \frac{(1+\varphi_2)(1-\varphi_1^2+\varphi_2^2-2\varphi_2)}{(1-\varphi_2)\{1+\varphi_1^2+\varphi_2^2-2\varphi_1(1-\varphi_2)\cos 2\pi f - 2\varphi_2\cos 4\pi f\}} \tag{7}$$

二、利用 AR(2)模型进行预报

正态平稳序列X_t的m步预报，就是根据现在时刻t和以前各时刻$t-1$，$t-2$，…的序列值X_t，X_{t-1}，X_{t-2}，…，对未来时刻$t+1$的随机变量X_{t+1}作出预报，预报值记为$\hat{X}_t(m)$，它等于X_t，X_{t-1}，X_{t-2}，…的线性组合。一般地用平稳线性最小方差预报。所谓平稳线性最小方差预报，是指预报值与真值之差的平方和最小，即

$$E[X_{t+m} - \hat{X}_t(m)] = min \tag{8}$$

在正态条件下，$\hat{X}_t(m)$是X_{t+1}对于序列值X_t，X_{t-1}，…的条件数学期望，即

$$\hat{X}_t(m) = E(X_{t+m} \mid X_t, X_{t-1}, \cdots) \tag{9}$$

根据这一定义和预报法则，可得出 AR(2)序列的连续预报递推公式：

$$\left.\begin{aligned}\hat{X}_t(1) &= \varphi_1 X_t + \varphi_2 X_{t-1}\\ \hat{X}_t(2) &= \varphi_1\hat{X}_t(1) + \varphi_2 X_t\\ &\cdots\quad\cdots\quad\cdots\\ \hat{X}_t(m) &= \varphi_1\hat{X}_t(m-1) + \varphi_2\hat{X}_t(m-2)\end{aligned}\right\} \tag{10}$$

由 AR(2)模型中的 φ_1,φ_2 和实测值 X_t, X_{t-1} 代入上式，作出“一步预报”$\hat{X}_t(1)$；又由 $\hat{X}_t(1)$ 与 X_t 作出“二步预报”$\hat{X}_t(2)$，如此递推，可连续作出“m 步预报”$\hat{X}_t(m)$。

在 $t+1$ 时刻，已取得新的实测值 X_{t+1}，与预报值 $\hat{X}_t(1)$ 的差$[X_{t+1}-\hat{X}_t(1)]$表示取得新的实测值后带来的“新息”，此时可进行现时改正预报。公式如下：

$$\hat{X}_{t+1}(m)=\hat{X}_t(m+1)+Gm[X_{t+1}-\hat{X}_t(1)] \tag{11}$$

上式表示 t+1 时刻的 m 步预报等于 t 时刻 m+1 步预报值与“新息”的加权和。“权”为 G_m。

三、预报方案实例

(一)建立模型，进行预报 颉河位于宁夏回族自治区南部山区，是泾河的一条支流。流域内多为土石山地，部分属陇东黄土高原。植被主要为山地草原森林小区和落叶小区，覆盖率约 60%。流域内多年平均降水量 620mm，在该区境内属降水量较丰的地区。因此，地下水量较为丰富。三关口水文站位于颉河靠甘宁交界处，流域面积 218km^2，枯季由 10 月至翌年 4 月，枯季径流占年径流量的 49.5%。现直接用实测的 20 年资料建立 AR(2)模型，预留 5 年资料作检验。

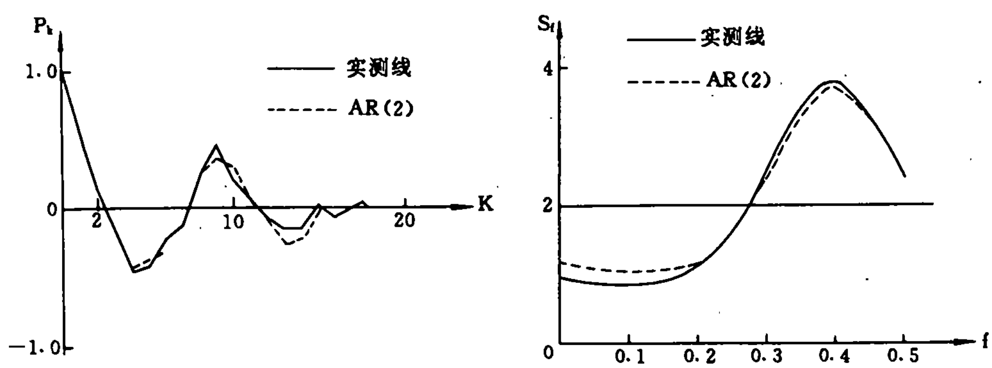
AR(2)模型：

$$Q_{i,j}=\overline{Q}_j+\varphi_{1j}(Q_{i,j-1}-\overline{Q}_{j-1})+\varphi_{2j}(Q_{i,j-2}-\overline{Q}_{j-2})+\varepsilon_{i,j} \tag{12}$$

式中 $Q_{i,j}$表示第 i 年第 j 月的平均流量； $\overline{Q}_j$ 表示第 j 截口的均值。参数 φ_{1j}和 φ_{2j}用二元回归方程逐个算出。于是可得 10 月至翌年 4 月各月平均流量的 AR(2)模型。

由 AR(2)模型对各月平均流量进行预报，得到各月平均流量的预报值，并由此推算出枯季径流总量的预报值(各月径流之和)。

(二)模型识别 模型识别就是识别模型的主要统计特性与样本是否吻合。由式(6)和式(7)分别计算出序列的自相关函数和谱密度函数，成果见图 1。可见 AR(2)模型与样本吻合较好，由此可以判断对样本选用 AR(2)模型是比较合适的。

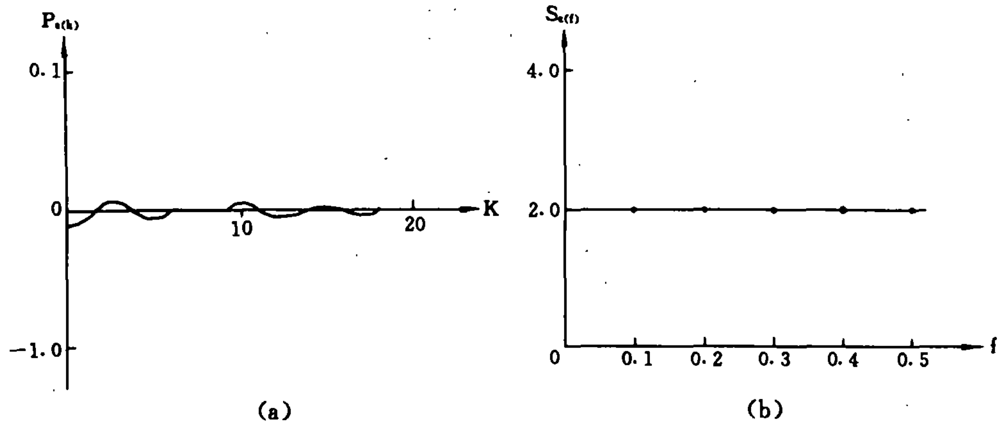


(a)相关图 (b)谱密度图

图 1 模型识别图

(三)模型检验 预报值与真值之差为误差项。模型检验就是要检验误差项 $\varepsilon_{i,j}$是否是白噪声(独立随机序列)。独立随机序列的自相关函数和方差谱密度函数分别围绕自相关系数为零和方差

谱密度为 2 的水平轴上、下波动。本模型计算成果见图 2。由此可以判断模型的误差项是独立的。



(a)　　(b)

图 2　AR(2)模型随机项的相关图(a)及方差谱密度图(b)

表 1　AR(2)模型预报方案评定表

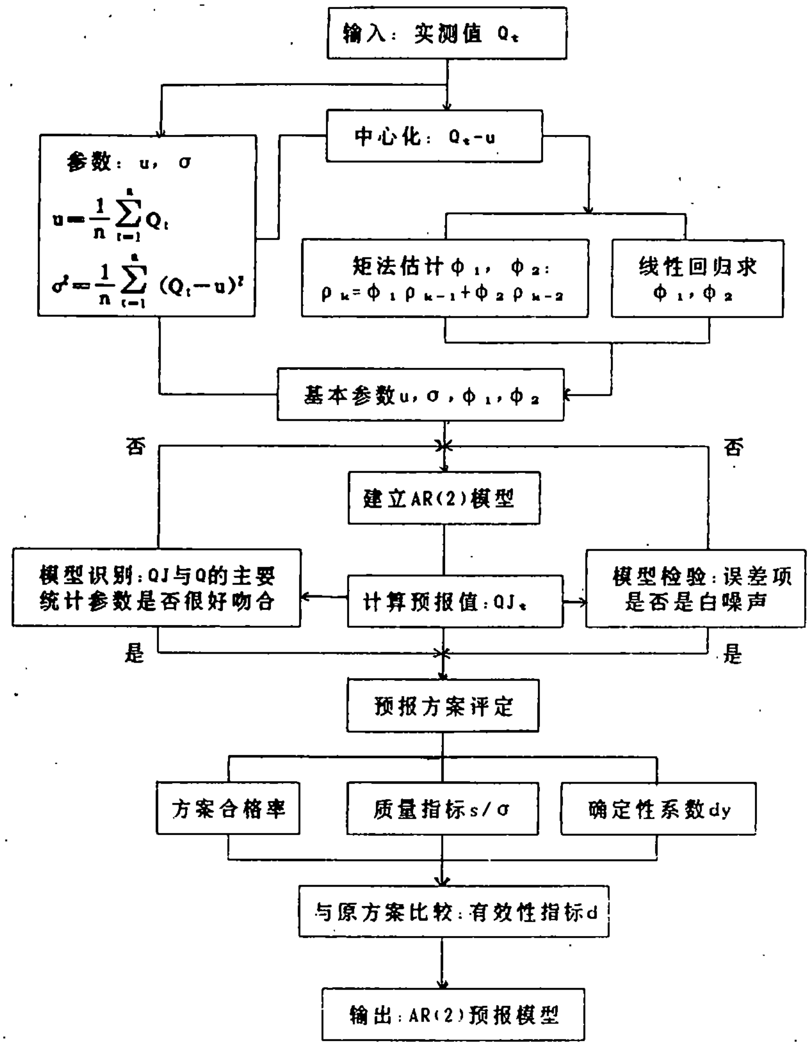
项目	月份	方案合格率(%)				方案质量指标			
		30%标准	方案等级	40%标准	方案等级	S/σ	方案精度	确定性系数 d	有效性指标 d
各月平均流量	10	93	甲	100	甲	0.40	好	0.84	0.16
	11	95	甲	100	甲	0.26	好	0.94	0.89
	12	85	甲	95	甲	0.42	好	0.82	0.70
	01	85	甲	100	甲	0.47	好	0.78	0.68
	02	95	甲	100	甲	0.37	好	0.87	0.94
	03	90	甲	90	甲	0.43	好	0.81	0.92
	04	80	乙	90	甲	0.89	不好	0.21	0.90
枯季径流总量		95			甲	0.20	好	0.96	0.95

注：枯季径流总量的预报误差以实测值的 20%为标准；1973 年 4 月 27 日发生第 3 位大洪水($634m^3/s$)对序列方差影响很大，若扣除其影响则 4 月方案精度为好。

(四)预报方案的评定　根据水利部水文水利调度中心《水文情报预报规范》，“对于小流域的年、月平均流量预报的允许误差，以实测值的 30%～40%为标准”及其对质量指标 s/σ 和确定性系数 dy 的有关规定，对方案的合格率和精度统计如表 1。

按规范的标准，预留的 5 年资料作检验，结果有 25%的预报值属优，27. 5%属优良，35%合格，总合格率为 87. 5%。因此，本方案属甲（优）级，精度高且预见期长，是完全可以发布正式预报的。

(五)方案比较　固原水文水资源勘测大队自 1977 年开始发布该站的枯季径流预报，其方案一直采用单纯的消退系数法，至今已有 14 年预报值。按上述方法统计，其各月平均流量的总合格率为 50%，而枯季径流总量的合格率仅为 21. 4%，均远远低于 AR（2）模型。按规范规定，有效性指标“$d>0$，表示新模型优于旧模型”。d 值见表 1。因此，AR（2）模型（见框图）的优越性是显而易见的。



AR(2)预报模型框图

四、结　语

1. 应用自回归模型预报枯季径流是一条很好的途径，预报总合格率为 87.5%。

2. 自回归模型较单纯利用消退系数的预报方案，精度高 66.1%，且适用性强。

参　考　文　献

〔1〕. 丁晶等.《随机水文学》. 成都市：成都科技大学出版社，1988 年

〔2〕. 华东水利学院主编.《水文学的概率统计基础》. 北京：水利电力出版社，1981 年

〔3〕. 高荣松等.《用自回归模型模拟沱江登瀛岩的枯水过程》. 成都科技大学学报，1986 年，第 4 期