

甘肃河西区出山径流统计特性分析与预测

陈元芳¹,梁忠民¹,董增川¹,沙志贵²

(1.河海大学水文水资源及环境学院,江苏 南京 210098;2.长江水利委员会水文局,湖北 武汉 450010)

摘要 根据甘肃河西区出山口水文站最新径流资料,对该区年、月径流统计特性(如自相关性、长持续性、径流年代变化、趋势性、周期性及年内变化等)利用 Hurst 系数、简单分波法等进行了分析,建立了该区年径流统计预测模型.结果表明该区年径流存在着明显的趋势性及长持续性,50 年代径流偏丰,90 年代有偏枯趋势.建立的时间序列预测模型,做出两步预测,相对误差可控制在 20% 以内.

关键词 年、月径流;出山口;趋势性;周期性;持续性;预测;甘肃省

中图分类号:TV121+.2 文献标识码:A 文章编号:1000-1980(2001)06-0083-04

1 问题的提出

甘肃省尤其是河西地区的社会经济发展明显落后于东部沿海地区.其原因是多方面的,重要原因之一是该地区水资源严重缺乏且地区分布的不均匀性.在当前西部大开发的大好形势下,为促进该地区社会经济发展,需要深入进行该地区水资源可持续利用研究.其中重要内容之一是对该区水资源变化规律作科学分析.考虑到该地区径流形成的特点,即各河出山口年、月径流量大小基本上能够反映该区水资源量的多寡.因此,本研究仅对甘肃河西地区主要内陆河流 11 个出山口水文站的径流系列作分析及预测计算.

2 资料收集及年、月径流基本特性分析

2.1 资料收集

表 1 出山口水文站基本情况

Table 1 Basic information of mountain-pass hydrological stations

站 名	所在河流	流域面积/km ²	资料起始年份
扎马什克	黑 河	4 589	1957 ~ 1995
犁 园 河	犁园河	2 240	1952 ~ 1995
沙 沟 寺	东大河	1 614	1955 ~ 1996
杂 木 寺	杂木河	851	1952 ~ 1996
冰 沟	讨赖河	6 883	1948 ~ 1996
昌 马 堡	昌马河	16 961	1953 ~ 1996
莺 落 峡	黑 河	10 009	1945 ~ 1996
红 崖 山	石羊河	14 242	1962 ~ 1995
干 沟 门	犁园河	839	1962 ~ 1983
石 家 庄	石羊河	6 805	1967 ~ 1995
党 城 湾	党 河	14 325	1966 ~ 1996

出山口水文站概况见表 1.

从表 1 可知,所研究的序列长度,多数在 40 a 以上,一般也在 30 a 左右,说明该资料序列是比较长的.由于是出山口径流站,其径流形成的自然地理条件在不同年代较为一致.

2.2 年、月径流基本特性分析

2.2.1 年径流基本统计参数的计算与分析

a. 计算方法

采用概率权重矩法估计参数 $\bar{X}$, C_V , C_S ,具体计算方法见文献[1],自相关系数 r 采用常规计算方法.在年径流序列中,有时会出现连续多个丰水年组和枯水年组交替出现的现象,这种现象表明年径流序列可能具有长持续性的特征.到底用什么指标来反映这一特性,长期以来未能得到完全解决,本文尝试采用 Hurst 系数. Hurst 系数(h_k)计算方法见文献[2].Hurst 本人曾对 800 个系列的 h_k 值作了计算,结果发现 h_k 远大于纯随机系列或 ARMA(p, q)序列($n \rightarrow$

∞)所对应的 h_k 值,即 $h_k = 0.5$. 其平均值 $\overline{h_k} = 0.73$. 由于实测系列 h_k 远大于短持续序列 $h_k = 0.5$,说明所计算序列可能存在长持续性,即所谓的 Hurst 现象.

b. 结果分析

从表 2 可以发现该区年径流量不但很小,最大年径流深仅 282 mm,而且空间分布极不均匀. 最大与最小年径流深之比为 15; C_V 平均值约 0.18,是较小的; C_S/C_V 比值均值为 4.5,最大为 8.4,最小为 2,变化较大.

对于自相关性检验,本文检验方法是,计算前 3 阶自相关系数,取显著性水平 α 为 0.05,当其中有一阶自相关系数检验未通过,则认为存在自相关. 检验结果是,总共 11 个站中有 7 个站存在着自相关,其中红崖山、石家庄两站前三阶自相关系数均高达 0.8 以上.

表 2 各站年径流基本特性计算结果

Table 2 Computational results of basic characteristics of annual runoff for all hydrological stations

站名	年径流平均值/($\text{m}^3 \cdot \text{s}^{-1}$)						$\bar{X}$		C_V	C_S	r_1	r_2	h_K	h_K (剔除趋势后)
	40年代	50年代	60年代	70年代	80年代	90年代	流量 ($\text{m}^3 \cdot \text{s}^{-1}$)	径流深 /mm						
扎马什克		27.0	22.1	21.2	24.0	16.9	22.5	154	0.15	0.87	0.11		0.73	
犁园堡		7.1	6.5	6.2	7.3	5.4	6.6	93	0.24	0.50	0.11		0.74	
沙沟寺		11.5	9.2	9.3	9.5		9.7	190	0.16	1.35	0.06		0.70	
杂木寺		10.5	7.0	6.9	7.1	6.7	7.6	282	0.29	1.38	0.33	0.05	0.83	
冰沟	16.7	22.7	19.3	21.4	19.9	17.5	20.3	93	0.19	1.70	0.32	0.05	0.76	
昌马堡		24.8	25.4	28.3	27.2	25.8	26.5	49	0.21	0.69	-0.05		0.65	
莺落峡	46.3	51.9	46.6	40.2	39.0	46.9	44.9	142	0.23	0.46	0.45	0.21	1.03	1.01
红崖山			11.3	9.0	5.9	4.6	7.9	18	0.37	0.74	0.87	0.84	1.02	0.80
干沟门			5.11	5.0	5.7		5.2	196	0.22	0.58	-0.12		0.69	
石家庄			12.1	7.8	5.4	4.4	6.7	31	0.41	1.75	0.88	0.83	0.93	0.92
党城湾			9.4	10.1	11.6	12.7	11.0	24	0.15	1.06	0.58	0.52	1.06	0.81

注 $\bar{X}$ 表示由所有年份年径流资料所计算的年径流平均值.

关于 Hurst 系数,甘肃河西区 $\overline{h_k} = 0.84$,比 Hurst 本人所计算 $\overline{h_k}$ 值 0.73 还要大. 据分析,Hurst 系数大于 0.5 原因不外乎有三种:样本容量不够大;样本系列本身不平稳;存在长持续性.

经分析,Hurst 系数在 0.9 以上的 4 个站均存在着趋势性,为了消除不平稳对 Hurst 系数值计算的影响,对这 4 个站年径流系列作剔除趋势项的计算. 结果发现,红崖山和党城湾两站的 Hurst 系数由剔除趋势之前的 1.00 以上到剔除趋势项之后的 0.8 左右,而另外两站,剔除趋势项与否,几乎不对 Hurst 系数产生影响. 说明这 4 个站可能存在长持续性.

2.2.2 不同年代平均年径流量对比分析

对各站分别计算不同年代年径流算术平均值,对于不足 10 年的年代仍然计算均值,11 个站中,有两个可计算 6 个年代,其余多数为 4~5 个年代,为清楚起见,还计算不同年代年径流均值相对距平值(因篇幅所限未列出). 按相对距平绝对值超过 10% 时为偏丰、偏枯的标准,从结果可知,50 年代偏丰偏枯站比例较高,4 个偏丰、偏枯站中全部属于偏丰,说明 50 年代偏丰. 60、70 年代偏丰、偏枯站比例较少,11 个站中仅 3 个. 80 年代比例有所增加,为 5/11. 90 年代,偏丰、偏枯比例高,以偏枯为主,且幅度大,说明 90 年代有偏枯趋势.

2.2.3 年内分配均匀性分析与计算

本次以连续最丰、最枯 4 个月径流量占全年百分比作分析计算,结果是,连续最丰 4 个月除了红崖山发生在 1~4 月,石家庄发生在 12~3 月外,其余都发生在 6~9 月,比例由 47.3%~83% 不等. 而连续最枯 4 个月除了红崖山发生在 6~9 月,石家庄发生在 4~7 月以外,其余均发生在 12~3 月,占全年总径流量比例在 21%~25% 不等. 平均而言,连续最丰 4 个月径流量是连续最枯 4 个月的 6 倍. 这说明该区各站径流年内变化很大,显然对于合理有效地开发当地水资源是十分不利的.

3 年径流组成的趋势性及周期性的分析与处理

一般水文序列可认为由趋势、跳跃、周期及随机成分的线性叠加而成,趋势、跳跃及周期成分都是确定性的,换句话说,水文序列是由确定性与随机性成分线性叠加而成. 要预测年、月径流未来的变化,必须弄清年径流组成中各项的变化规律.

3.1 趋势性的检验与剔除

3.1.1 趋势性检验

检验途径有两条：一是统计检验，二是观察年径流随时间变化过程线，前者在检验方法以及显著水平 α 确定的情况下，结果不会因人而异，而后者虽可能因人而异，但更直观，故建议同时使用。统计检验时采用 Kendall 秩次相关检验及游程检验，具体方法见文献 [4]， $\alpha = 0.05$ ，利用以上两种方法作检验时，判断有否趋势标准是：只要两种统计检验中有一种检验有趋势，则认为该序列有趋势。检验结果表明，11 个站中有 5 个站存在着趋势，且两种统计检验法其结论是一致的。从年径流过程线看统计检验结果是可接受的，即与直观的判断是相吻合的。总之，本区有趋势站的比例是相当高的。

3.1.2 趋势项的提取

对于 5 个有趋势的出山口年径流序列，逐个观察其随时间变化过程线，即 $x(t) \sim t$ 的关系 ($t = 1, 2, \dots, n$)，结果发现都具有线性趋势，故采用线性方法对各站趋势项进行提取，即分析确定方程：

$$X_t = a + bt + Z_t$$

式中： a, b ——线性回归方程系数； X_t ——第 t 年年径流； Z_t ——零均值随机成分，一般情况可当做平稳随机过程。 a, b 计算结果见表 3。对所建立回归方程是否显著作了检验，结果表明，所建立的 5 个方程均是显著的。

表 3 年径流系列趋势性检验与线性剔除成果
Table 3 Test and treatment of trend for annual runoff series

站 名	计算 u 值		线性趋势提取	
	Kendall 检验	游程检验	a	b
杂木寺	-2.81	-0.60	9.2	-0.072
红崖山	-6.18	-4.86	12.3	-0.25
莺落峡	-1.78	-2.67	49.2	-0.165
石家庄	-5.63	-5.09	10.6	-0.26
党城湾	4.44	-3.31	8.90	0.14

注：在 $\alpha = 0.05$ 时，趋势性检验临界值 $u_\alpha = 1.96$

3.2 周期成分的识别与拟合

采用简单分波法^[3]对周期成分进行识别和拟合。假设序列中有 K 年周期，将序列按 K 年一组共分成 m 组，定义检验统计量

$$F = \frac{Q_3^2(K-1)}{Q_2^2(n-K)} \sim F(K-1, n-K)$$

式中： Q_2^2 ——组内离差平方和； Q_3^2 ——组间离差平方和； n ——样本容量。

对给定的 α ，临界值记为 F_α ，当 $F > F_\alpha$ ，则认为有 K 年周期，否则不存在 K 年周期。检验时 $K = 2, 3, \dots, n/2$ 或 $(n-1)/2$ ，当检验时发现显著周期 K 不止一个时，取计算的 F 值最大的那一个作为周期值。

若检验的序列有 K 年周期，则以各组组内均值作为 K 年的周期值，在原序列中顺序减去该周期值，得到相应的余波过程 $A(t)$ 。对 $A(t)$ 重复上述计算过程，则又得到另一周期过程及相应余波，对这个过程重复 L 次，则得到 L 个周期项及其余波。序列周期成分即由这 L 个周期叠加而成，本次计算时，约定 $L = 3$ ，即只考虑到其前三个周期，认为后面的周期成分影响不大。年径流简单分波法计算结果见表 4。

表 4 周期成分分析结果

Table 4 Analytical results of periodic component of annual runoff series

站 名	周 期 组 成		
	第一周期	第二周期	第三周期
扎马什克	6	0	0
犁园堡	8	0	0
沙沟寺	11	3	0

注：其它站均不存在周期成分

4 年径流序列的拟合

从原序列中依次剔除趋势及周期成分，将剩余的序列作为平稳时间序列，用平稳 $ARMA(p, q)$ 模型拟合。该模型形式见文献 [4]。用 AIC 准则定阶，模型参数估计采用最小二乘法。为简化，计算时限定 $0 \leq p, q \leq 3$ 。模拟结果是：有 6 个站随机成分独立，5 个站可建立 ARMA 模型，即 p, q 取值不为零。拟合模型结果是：杂木寺、冰沟、莺落峡、红崖山、石家庄可分别用 $ARMA(3, 3)$ 、 $AR(1)$ 、 $AR(1)$ 、 $ARMA(3, 1)$ 及 $AR(1)$ 拟合。

5 年径流预测

按第 3 节中介绍的方法识别年径流的各组成项 ,所采用的资料是除了最新两年外的所有收集到的实测资料。对年径流的预测分两步进行 ,首先对确定性成分作预测 ,直接按时间外推计算 ,然后再对随机项采用逆函数作计算 ,该法的具体计算公式见文献 [4] ,把确定性与随机性成分预测值相加后所得的计算结果即为年径流的预测值。对于有趋势或有周期或有相关结构的 9 个站 ,利用所建立预估模型进行两步预测(即对实测最新两年年径流作预测)其结果见表 5。从表中可看出 ,一、二步预测相对误差基本上可控制在 20% 以内 ,因此可以说 ,本文所建预测模型具有较高精度。

表 5 9 个站年径流两步预测的相对误差
Table 5 Two-step prediction of annual runoff for 9 stations %

站 名	扎马什克	犁园堡	沙沟寺	杂木寺	冰沟	莺落峡	红崖山	石家庄	党城湾
一步	23.1	55.5	11.3	4.3	11.8	6.7	18.4	- 18.1	- 12.7
二步	22.2	14.4	20.8	9.3	3.6	- 6.5	- 11.2	- 26.3	22.1

6 结 语

- a. 本文利用甘肃河西地区近期最新年月径流资料 ,采用比较成熟且比较新颖的分析方法对年月径流基本特性及年径流组成成分进行了全面分析。因此 ,所得结论是比较可靠的。分析计算方法对其他地区年月径流统计特性及预测分析有一定参考价值。
- b. 甘肃河西区 11 个出山口站中有些站可能有长持续性 ,仅有三个站年径流存在周期 ,有近一半的年径流存在严重趋势性 ,且多数有下降的趋势。按年代对出山径流分析表明 ,50 年代径流偏丰 ,而 90 年代径流在多数站出现了偏枯情况 ,这些都对于当地社会经济可持续发展将是不利的 ,应引起足够重视。
- c. 在分析年径流组成特性及组成成分后 ,本文还建立了时间序列预测模型 ,并做一、二步预测 ,预估相对误差大都在 20% 以下 ,从而可认为预估具有较高精度。

参考文献 :

[1] 陈元芳 ,侯玉 .P-Ⅲ 分布参数估计方法的研究 [J].河海大学学报 ,1992 (3) 24 ~ 31.
[2] 丁晶 ,邓育仁 .随机水文学 [M].成都 :成都科技大学出版社 ,1988.77 ~ 80.
[3] 金光炎 .水文水资源随机分析 [M].北京 :中国科学技术出版社 ,1993.411 ~ 418.
[4] 杜金观 ,项静恬 ,戴俭华 .时间序列分析——建模与预报 [M].合肥 :安徽教育出版社 ,1991.174 ~ 186.

Statistic Characteristics Analysis and Prediction for Runoff from Mountain-Pass Stations in Hexi Area of Gansu Province

CHEN Yuan-fang¹ , LIANG Zhong-min¹ , DONG Zhen-chuan¹ , SHA Zhi-gui²
(1. College of Water Resources and Environment , Hohai Univ. , Nanjing 210098 , China ;
2. Hydrological Bureau of Yangtze Water Resources Commission , Wuhan 450010 , China)

Abstract : On the basis of the runoff data recently collected from mountain-pass stations in the Hexi area of Gansu Province , the statistic characteristics of runoff , including auto-correlation , long-term continuity , runoff variation in different decades , and statistic components (trend , period) are analyzed , and a predictive statistic model is established for the prediction of annual runoff. The results show that the runoff in the studied area has an obvious trend and a long-term continuity , and that the runoff is abundant in 1950s , and short in 1990s. The error of prediction by the model may be controlled within 20% .

Key words : annual and monthly runoff ; mountain-pass ; trend ; periodicity ; continuity ; prediction ; Gansu Province