

月径流影响因子的识别

杨会龙, 王双银, 王建莹

(西北农林科技大学 水利与建筑工程学院, 陕西 杨凌 712100)

摘要:径流影响因子的合理识别是径流预报的关键和基础,能够为深入探索径流演变规律提供科学依据。以 1954—2007 年陕西省石头河流域逐月降水资料和石头河水库坝址逐月径流资料为基础,通过成因分析对月径流的可能影响因子进行了初步筛选,其次分别采用相关概率法、单相关系数法和 Spearman 相关系数法对可能影响因子与月径流之间的关系进行了分析,选取通过显著性检验的因子交集作为各月径流的主要影响因子。结果表明:石头河流域 1 月和 11 月的月径流主要影响因子是前一月径流,3 月和 12 月的月径流主要影响因子是前两个月的月径流,2 月、4 月、7 月、8 月和 10 月的月径流主要影响因子是当月降水和前一月径流,5 月、6 月和 9 月的月径流主要影响因子是当月降水。

关键词:成因分析; 相关分析; 径流影响因子; 石头河流域

中图分类号: P332.4

文献标识码: A

文章编号: 1005-3409(2011)05-0108-04

Identification of Influencing Factors on Monthly Runoff

YANG Hui-long, WANG Shuang-yin, WANG Jian-ying

(College of Water Resources and Architectural Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China)

Abstract: Identification of influencing factors on runoff is a key to runoff forecasting, which could offer a scientific support for further study the evolution of runoff. This paper takes monthly runoff as an example and uses the data on monthly runoff and the average precipitation of the Shitouhe watershed from 1954 to 2007 to select the possible factors according to cause analysis, then uses the relevant probability, single correlation coefficient and Spearman coefficient to analyze correlativity, gets the main factors from the consolidated results of above analysis methods. The results show that runoff occurring in January and November is mainly affected by runoff of previous month, monthly runoff of March and December is mainly affected by the runoff occurring in advance of that two months. Runoff of February, April, July, August and October are affected by precipitation with previous monthly runoff. Runoff of May, June and September is mainly affected by the monthly precipitation.

Key words: cause analysis; correlation analysis; runoff influencing factor; Shitouhe watershed

径流影响因子的合理识别是径流预报的关键和基础。径流影响因子识别的研究方法可以归纳为成因分析法和相关分析法,这两种方法通常结合使用。成因分析法主要是根据径流形成原理从气候因素和下垫面条件筛选径流影响因子^[1-3],由于受到现有认识程度和科技水平的限制,某些因子与径流之间的内在联系尚未被完全研究清楚,而且尺度大小也难以确定,为了更好地分析这些要素之间的相互关系,需要进一步从统计学角度对其进行相关性度量。已被广泛应用于径流影响因子识别的相关分析法有相关概

率法、单相关系数法、Spearman 相关系数法、主成分分析法及灰色关联度分析法等^[4-7]。其中,相关概率、单相关系数和 Spearman 相关系数的大小可以直接度量两组变量间的相关程度;主成分分析是对于原先提出的多组变量,通过线性变换建立尽可能少且尽可能保持原有信息的新变量,是一种数学降维方法;灰色关联度分析是根据因素之间发展趋势的相似或相异程度衡量因素间关联程度的一种方法,主要用以描述多组影响因子与被影响因子间关联度的优劣次序。这些方法反映因素相关的角度不同,实际应用中需要

综合分析使用。水文系统中,径流与其影响因子之间的关系错综复杂,目前对年径流(大尺度)和场次洪水(小尺度)的影响因素研究已经较为成熟,对于月径流(中等尺度)影响因子识别尚处于起步阶段。本文拟以陕西省石头河流域为例,采用成因分析法并结合相关概率法、单相关系数法和 Spearman 相关系数法对该流域各月径流的可能影响因子进行分析,并确定出主要影响因子,为该流域的月径流预报提供科学依据。

1 研究流域概况

石头河是渭河南岸较大一级的支流之一,发源于秦岭北麓太白山区,自南向北流经太白县,于眉县斜峪关出峪口后向北流 15 km 汇入渭河,干流全长 51.5 km,河道平均比降为 25.4‰,流域面积 778.7 km²,多年平均径流量 4.81 亿 m³[8]。流域内斜峪关以上为山区,林木茂盛,植被良好,水流清澈,水质优良,较少受人类活动影响。在斜峪关上游 1.5 km 处建有石头河水库,控制流域面积 673 km²。石头河水库是一座兼具城市供水、农田灌溉、水力发电和防洪等多功能综合利用的大型水库,总库容 1.47 亿 m³,有效库容 1.2 亿 m³,设计年调节水量 2.7 亿 m³,坝址多年平均流量 14.1 m³/s。

2 资料处理

石头河流域现设有 5 个雨量测站,本次采用加权平均法(泰森多边形法)计算流域面降水量,但是各测站的设立年份不一样,因此 1954—1970 年的逐月降水量由桃川站和斜峪关站的观测值计算,其加权系数分别为 0.842 和 0.158;而 1971—2007 年的逐月降水量由杜家庄站、桃川站、鸚鹄站、高码头站和斜峪关站的观测值计算,其加权系数分别为 0.227,0.336,0.271,0.136 和 0.030。

石头河水库坝址处原设有斜峪关水文站,1954—1983 年有连续径流流量观测值,1984 年该站撤消。1973 年在斜峪关上游设立鸚鹄水文站,控制流域面积 507 km²,距水库坝址 13 km,鸚鹄水文站于 1974 年开始测流至今。本次采用的 1954—2007 年石头河水库坝址逐月径流流量资料系列由 1954—1973 年斜峪关水文站实测月平均流量和 1974—2007 年鸚鹄水文站实测流量按面积比推算到坝址处的资料系列组成。经审查,资料系列可靠,成因一致,代表性良好。

3 影响因子初步筛选

径流的形成主要受流域气候特征和下垫面条件

两方面的综合影响。石头河流域属亚热带气候,湿润多雨,水汽主要来源于孟加拉湾和南海两处海洋,在太平洋副热带高压的影响下,随西南季风和东南季风带来大量孟加拉湾和南海的水汽,当冷暖峰相遇时,峰面抬升,大气中的水汽迅速凝结,形成较强降水[9]。径流的形成是从降水开始到水流汇集至流域出口断面的物理过程[10]。石头河流域的产流方式以蓄满产流为主,考虑到产汇流的时滞性,当包气带土壤含水量达到田间持水量时才开始产流,河川径流量可能受本时段降水、前期降水、前期蓄水量和本时段内流域蒸散发量的影响。由于石头河流域气候湿润,流域面积较小,当以月时段为时间尺度时,蒸散发量较小,可以忽略其对径流量的影响,流域前期蓄水量可以通过前期径流反映,流域主河道流程较短,月内降水主要形成本月径流,月末降水的产流最晚亦可以在下月月末流出流域(即前期降水只考虑前一月的降水量),而前期径流主要影响本时段的基流,本阶段初选前三个月的月径流作为前期径流进行分析。综上分析,石头河流域月径流与其可能的影响因子可以采用如下形式描述:

$$R_t = f(P_t, P_{t-1}, R_{t-1}, R_{t-2}, R_{t-3}) \quad (1)$$

式中: t ——当前时段(月); $t-1, t-2, t-3$ ——表示前移 1, 2, 3 个时段(月); R_t ——当月月径流流量(m³/s); P_t ——流域当月降水量(mm)。

即通过物理成因分析初步筛选的石头河流域月径流可能影响因子为流域当月降水量、前月降水量和前三个月的月径流流量。

4 相关分析

相关分析就是研究两个或多个随机变量之间的联系,给出两者或多者间相依程度的定量描述,建立一些科学合理的指标以衡量变量之间关系的密切程度。

4.1 相关概率法

相关概率法是水文气象工作者人工挑选预报因子的一种典型方法[11]。具体步骤是首先计算出径流 y 和影响因子 x 的距平,然后分析比较它们的距平符号(“+”或“-”),统计距平符号相同的百分率,在一定置信度 α 下达到标准的表明径流 y 和影响因子 x 显著相关,则可以选择该因子,达不到标准的则舍弃。常用 χ^2 检验考察其显著程度。 χ^2 值的计算公式为

$$\chi^2 = \frac{(2m-n)^2}{n} \quad (2)$$

式中: n ——样本容量, $n=52$; m —— y 与 x 的距平符号相同或相反的出现次数,统计出现次数多的。

根据样本计算出 χ^2 值以后,选择置信度 α ,在 χ^2 分布分位数表^[12]中查出自由度为 1 的 χ_α^2 数值,如果 $\chi_\alpha^2 > \chi^2$,则认为在这一信度下,两个随机变量相关性显著,这一因子可选,否则应予以舍弃。置信度 α 也称为置信水平或可靠度,选取的 α 越小,临界值越大,

通过相关性检验的因子越可靠,由于本论文研究识别月径流的影响因子,对因子数量不做要求,信度 α 取 0.05 即可。经分析计算,石头河流域月径流与其可能的影响因子的“ χ^2 ”值计算成果见表 1,在 χ^2 分布表中查得:自由度为 1 时, $\chi_\alpha^2 = 3.841$ 。

表 1 石头河流域月径流与其可能的影响因子的 χ^2 值计算结果

因子	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
$R_t - P_{t-1}$	0.961	0.176	0.176	3.314	5.667*	4.412*	0.961	7.078*	0.490	0.490	0.020	0.490
$R_t - P_t$	0.077	4.923*	1.923	22.23*	19.69*	22.23*	7.692*	17.31*	22.23*	19.69*	3.769	0.308
$R_t - R_{t-1}$	16.49*	37.23*	4.923*	7.692*	3.769	3.769	4.923*	7.692*	1.231	7.692*	6.231*	13.00*
$R_t - R_{t-2}$	4.412*	16.49*	4.923*	1.231	0.077	1.231	3.769	2.769	0.308	4.923*	0.077	11.08*
$R_t - R_{t-3}$	4.412*	4.412*	3.314	1.231	0.692	0.308	2.769	1.923	0.308	4.923*	1.923	11.08*

注:标“*”的表示通过显著性检验,则选取该因子,下表同。

由表 1 可以看出:石头河流域 1 月和 12 月的月径流仅与前三个月的月径流显著相关,2 月和 10 月的月径流与当月降水和前三个月的月径流均显著相关,3 月的月径流仅与前两个月的月径流显著相关,11 月的月径流仅与前一月的月径流显著相关,4 月和 7 月的月径流与当月降水和前一月径流显著相关,5 月和 6 月的月径流与当月降水和前一月降水显著相关,8 月的月径流与当月降水、前一月降水及前一月径流显著相关,9 月的月径流仅与当月降水显著相关。

4.2 单相关系数法

在径流影响因子的识别中经常采用单相关系数来考察径流与影响因子是否线性相关,并以此作为因子挑选的依据。单相关系数的计算公式为

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3)$$

式中: x_i ——径流的系列值; y_i ——影响因子的系列值; \bar{x} ——径流的均值; \bar{y} ——影响因子的均值; r ——线性相关系数, $0 \leq |r| \leq 1$,当 $|r|$ 越接近于 1 时,说明其线性相关越显著。

当计算出 $|r|$ 后,同样可以根据样本的容量 n 和置信度 α 在相关系数临界值 r_α 表中查出 r_α ,如果 $|r| > r_\alpha$,则表明两者线性相关程度显著,否则,线性相关程度不显著。石头河流域月径流与其可能的影响因子的 $|r|$ 值计算成果见表 2,当置信度 $\alpha = 0.05$,样本容量 n 为 52 时, $r_\alpha = 0.273$ 。

表 2 石头河流域月径流与其可能的影响因子的 $|r|$ 值计算结果

因子	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
$R_t - P_{t-1}$	0.142	0.295*	0.133	0.184	0.131	0.196	0.209	0.215	0.012	0.083	0.093	0.046
$R_t - P_t$	0.039	0.278*	0.376*	0.767*	0.80*	0.872*	0.678*	0.836*	0.782*	0.770*	0.272	0.102
$R_t - R_{t-1}$	0.625*	0.936*	0.554*	0.327*	0.468*	0.294*	0.350*	0.364*	0.403*	0.573*	0.479*	0.516*
$R_t - R_{t-2}$	0.440*	0.645*	0.458*	0.133	0.061	0.185	0.281*	0.079	0.383*	0.164	0.276*	0.493*
$R_t - R_{t-3}$	0.265	0.409*	0.308*	0.179	0.173	0.114	0.123	0.066	0.027	0.240	0.151	0.435*

由表 2 可以看出,单相关系数法的分析结果为石头河流域 1 月和 11 月的月径流仅与前两个月的月径流显著相关,12 月的月径流与前三个月的月径流均显著相关,2 月的月径流与当月降水、前月降水、前三个月的径流均显著相关,3 月的月径流与当月降水和前三个月的径流显著相关,4 月、5 月、6 月、8 月和 10 月的月径流与当月降水和前一月的月径流显著相关,7 月和 9 月的月径流与当月降水和前两月的月径流显著相关。

4.3 Spearman 相关系数法

应用单相关系数法挑选因子时,要求随机变量服从正态分布,对于不服从正态分布的随机变量则不适用。Spearman 相关系数^[13]又称秩相关系数,是利用

两变量的秩次大小作线性相关分析,对原始变量服从何种分布不作要求,属于非参数统计方法。Spearman 相关系数的计算公式为

$$R = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n (y_i - x_i)^2 \quad (4)$$

式中: R ——Spearman 相关系数; n ——样本容量; x_i ——径流的系列值; y_i ——影响因子的系列值。

首先计算出 Spearman 相关系数 R ,然后根据样本容量 n 和置信度 α 在 Spearman 相关系数临界值表^[14]中查得到 R_α ,若 $R_\alpha > R$,则选取该因子,否则舍弃。石头河流域月径流与其可能的影响因子的 R 值计算成果见表 3,当置信度 α 为 0.05 时,样本容量 $n = 52, R_\alpha = 0.297$ 。

表3 石头河流域月径流与其可能的影响因子的 |R| 值计算结果

因子	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
$R_t - P_{t-1}$	0.530*	0.065	0.447*	0.390*	0.343*	0.367*	0.290	0.113	0.275	0.520*	0.435*	0.070
$R_t - P_t$	0.032	0.322*	0.387*	0.707*	0.786*	0.812*	0.650*	0.671*	0.847*	0.582*	0.217	0.052
$R_t - R_{t-1}$	0.769*	0.936*	0.903*	0.375*	0.468*	0.459*	0.354*	0.319*	0.265	0.584*	0.416*	0.660*
$R_t - R_{t-2}$	0.048	0.101	0.502*	0.181	0.075	0.252	0.344*	0.216	0.247	0.292	0.171	0.452*
$R_t - R_{t-3}$	0.005	0.012	0.123	0.096	0.125	0.081	0.227	0.032	0.011	0.290	0.163	0.291

由表3可以看出,石头河流域1月份和11月份的月径流与前一月降水和前一月径流显著相关,2月份的月径流与当月降水和前一月径流显著相关,3月的径流与当月降水、前一月降水、前三个月的月径流均显著相关,4—6月的月径流与当月降水、前一月降水及前一月径流显著相关,7月的月径流与当月降水和前两个月的月径流显著相关,8月的月径流与当月降水和前一个月的月径流显著相关,9月的月径流仅与当月降水显著相关,10月的月径流与当月降水、前一月降水及前一月径流显著相关,12月的月径流与前两个月的月径流显著相关。

5 主要影响因子的综合识别

相关概率法和 Spearman 相关系数法其实质都是秩相关,适合等级资料的相关性度量,对于变量的总体分布形态和样本容量大小均没有要求;单相关系数反映影响因子与径流之间的线性相关程度,但是要求数据资料服从正态分布,因此这三种方法反映相关程度的角度不同,其分析结果也不同。为了能够准确有效识别各月月径流的主要影响因子,综合以上相关分析结果,取与各月径流显著相关的影响因子交集作为各月月径流的主要影响因子,识别成果认为石头河流域1月和11月的月径流主要影响因子是前一月径流,3月和12月的月径流主要影响因子是前两个月的月径流,该时段内流域降水量很小或无降水;2月、4月、7月、8月和10月的月径流主要影响因子是当月降水和前一月径流,该时段月径流受当月降水和前一月径流的共同影响,5月、6月和9月的月径流主要影响因子是当月降水,该时段降水量较大,易直接形成地表径流,致使河川径流增加。

6 结论

(1)石头河流域1月和11月的月径流主要影响因子是前一月径流,3月和12月的月径流主要影响因子是前两个月的月径流,2月、4月、7月、8月和10月的月径流主要影响因子是当月降水和前一月径流,5月、6月和9月的月径流主要影响因子是当月降水。

(2)在成因分析基础上采用相关分析法识别月径流影响因子,符合流域径流形成原理,而且所需资料

容易获取,计算方法简单。

(3)研究方法和成果均可以有效应用于秦岭北麓与石头河流域气候水文特征和下垫面条件相似的其他流域,可以将综合识别得到的各月月径流主要影响因子作为预报因子进行月径流预报。

参考文献:

- [1] Wang G L, Eltahir E A B. Use of ENSO information in medium and long range forecasting of the Nile floods[J]. *Journal of Climate*, 1999, 12(6): 1726-1737.
- [2] Whitaker D W, Wasimi S A, Islam S. The El Nino-Southern Oscillation and long-range forecasting of flows in the Ganges[J]. *International Journal of Climatology*, 2001, 21(1): 77-87.
- [3] Piechota T C, Chiew F H S, Dracup J A, et al. Seasonal streamflow forecasting in eastern Australia and the El Nino Southern Oscillation [J]. *Water Resources Research*, 1998, 34(11): 3035-3044.
- [4] 张丽霞,梁新平. 基于单相关系数法的中长期水文预报研究[J]. *水资源与工程学报*, 2008, 19(3): 49-51.
- [5] 郭霞,王正中,王双银. 基于相关分析的流域汛期划分初探[J]. *西北农林科技大学学报:自然科学版*, 2010, 28(3): 206-210.
- [6] 李红霞,许士国,范垂仁. 基于贝叶斯正则化神经网络的径流长期预报[J]. *大连理工大学学报*, 2006, 46(增刊): 174-177.
- [7] 金艳. 人类活动对黄土高原三川河流域产汇流变化影响的研究[D]. 郑州:华北水利水电学院, 2005.
- [8] 王双银,冯国章,宋松柏,等. 石头河流域产流预报模型研究[J]. *西北农林科技大学学报:自然科学版*, 2001, 29(6): 115-118.
- [9] 余汉章. 陕西水文[M]. 西安:陕西科学技术出版社, 1987.
- [10] 詹道江,叶守泽. 工程水文学[M]. 北京:中国水利水电出版社, 2000.
- [11] 汤成友,官学文,张世明. 现代中长期水文预报方法及应用[M]. 北京:中国水利水电出版社, 2008.
- [12] 刘光祖. 概率论与应用数理统计[M]. 北京:高等教育出版社, 2000: 379-394.
- [13] 王晓燕,李美洲. 浅谈等级相关系数与斯皮尔曼等级相关系数[J]. *广东轻工职业技术学院学报*, 2006, 5(4): 26-27.
- [14] 刘桂芬. 医学统计学[M]. 北京:中国协和医科大学出版社, 2007: 365-366.