

DOI :10.3876/j.issn.1000-1980.2009.03.002

多因子逐步回归周期分析在中长期水文预报中的应用

葛朝霞^{1 2}, 薛 梅^{1 2}, 宋颖玲^{1 2}

(1. 河海大学水文水资源与水利工程科学国家重点实验室, 江苏 南京 210098 ;
2. 河海大学水文水资源学院, 江苏 南京 210098)

摘要 : 利用多因子逐步回归周期分析法建立长江宜昌站年平均流量的中长期水文预报模型. 结果表明 : 所选优势因子主要是一组反映冷暖空气活动的指数, 冷暖空气的活动和强度对宜昌的年平均流量影响很大, 各优势因子的优势隐含周期 4 ~ 9 个不等, 且普遍都具有 25 a 和 22 a 的变化周期 ; 多因子逐步回归周期分析法对长江宜昌站年平均流量的拟合和试预报合格率都达到了 100% , 具有较高的可信度.

关键词 : 中长期水文预报 ; 逐步回归 ; 优势因子 ; 优势隐含周期 ; 宜昌站

中图分类号 : P338 **文献标识码 :** A **文章编号 :** 1000-1980(2009)03-0255-03

中长期水文预报是根据前期水文气象要素, 用成因分析与数理统计的方法, 对未来较长时间的水文要素进行科学的预测. 该学科是介于水文学、气象学与其他学科之间的一门边缘学科. 随着社会的不断发展, 国民经济各个部门对水文预报提出的要求越来越高, 如防汛抗旱的指挥, 大中型水利、水电、水运工程的兴建及管理运行等, 都要求水文部门能提供预见期长、准确性高的中期与长期预报. 由于影响因素的复杂与目前科学水平的限制, 中长期水文预报还处于探索、发展阶段^[1-4].

目前, 用统计学方法作中长期水文预报大体可以分为 2 类, 即单要素预报和多要素综合预报. 前者主要有历史演变法、时间序列法和周期分析法等 ; 后者主要包括多元回归分析、逐步回归分析等方法^[3-11]. 其中, 用逐步回归的方法可以挑选出物理意义清晰的优势因子, 建立径流和优势因子的预报模型, 但这种方法受资料的限制, 预报期不可能太长. 用逐步回归周期分析的方法来进行研究, 根据径流自身的变化规律建立预报模型, 可以使预报期增长, 但物理因子的重要信息不能得到充分利用. 为此, 本文采取了综合上述 2 种方法优点的多因子逐步回归周期分析法, 建立长江宜昌站年平均流量的中长期水文预报模型, 可得到较高的预报精度.

1 多因子逐步回归周期分析的原理

1.1 基本思想^[12-13]

首先利用逐步回归方法, 以各气象因子对预报方程的方差贡献大小作为判断标准, 逐个引入对径流影响显著的物理因子, 同时逐个剔除对径流影响不显著的物理因子, 最终挑选出优势因子. 这些优势因子一般都存在多个明显的准周期, 即隐含周期. 在这些隐含周期中, 有些隐含周期是起主导作用的, 它们能够显著地影响径流的变化趋势, 因而可以称为优势隐含周期. 这样可以用逐步回归周期分析方法来分析这些优势隐含周期, 并外推它们未来的变化情况. 其预报模型为

$$y = b_0 + \sum_{j=1}^m b_j \left[\sum_{i=1}^k a_{ij} f_{ij}(t) + \varepsilon_j(t) \right] \quad (t = n + 1, n + 2, \dots) \tag{1}$$

式中 : b_0, b_j ——回归系数 ; $f_{ij}(t)$ ——隐含周期长度为 l_i 的周期波序列 ; a_{ij} —— $f_{ij}(t)$ 的系数 ; $\varepsilon_j(t)$ ——白噪声 ; m ——优势因子的个数 ; k ——优势隐含周期的个数 ; n ——资料系列长度.

1.2 计算步骤

a. 首先从物理成因上考察因子, 并通过相关性计算粗选出对径流变化影响显著的物理因子, 然后用逐

收稿日期 : 2008-05-16
基金项目 : 国家自然科学基金(40233037, 90211011)
作者简介 : 葛朝霞(1954 —) , 女, 江苏南京人, 教授, 主要从事水文气象和中长期水文预报研究.

步回归方法挑选优势因子,建立预报方程:

$$y = b_0 + \sum_{j=1}^m b_j x_j$$

(2)

b. 对挑选出来的优势因子 x_j 根据周期分析法推算出各自的周期波序列,并用逐步回归方法挑选出优势隐含周期 $f_{ij}(t)$,同时确定出 a_{ij} 值,方程为

$$x_j = \sum_{i=1}^k a_{ij} f_{ij}(t) + \varepsilon_j(t)$$

(3)

c. 外推优势因子的优势隐含周期 $\{f_{ij}(n+1), f_{ij}(n+2), \dots\}$,并根据式(3)确定其外推值 $\{x_j(n+1), x_j(n+2), \dots\}$.

d. 根据优势因子的外推值和式(2)预报宜昌站年平均流量未来的变化情况.

2 实例分析

本文所用的资料为 1951~2005 年长江宜昌站的月径流资料和 1951~2005 年 74 项大气环流因子的月资料(由国家气候中心提供).

2.1 预报模型的建立

a. 先从物理成因上对气象因子进行考察,确定对宜昌年径流量影响显著的物理因子,然后计算各月的气象因子与长江宜昌站年平均流量的相关系数,取置信度 $\alpha = 0.01$,初步筛选出 19 个大气环流因子.

b. 对初选出的 19 个大气环流因子进行逐步回归分析,当给定置信度 $\alpha = 0.05$ 时,挑选出 5 个影响显著的优势因子,建立宜昌站年平均流量的预报方程:

$$y = 19713.05 + 591.41x_1 - 34.01x_2 + 73.84x_3 - 64.27x_4 - 25.57x_5$$

(4)

式中: x_1 ——4 月冷空气强度,冷空气越强(弱),则年平均流量越大(小); x_2 ——亚洲区 8 月极涡面积指数,极涡面积指数越大(小),则年平均流量越小(大); x_3 ——北半球 8 月极涡中心强度,极涡中心强度越大(小),则年平均流量越大(小); x_4 ——12 月西太平洋副高面积指数,副高面积指数越大(小),则年平均流量越小(大); x_5 ——太平洋区 9 月极涡面积指数,极涡面积指数越大(小),则年平均流量越小(大).综合这些分析可以得出,所选因子均反映了冷暖空气的强弱,可见,长江宜昌站的年平均流量与冷暖空气的强度关系十分密切.

c. 利用逐步回归周期分析法分析优势因子的优势隐含周期.表 1 列出了优势因子的优势隐含周期和复相关系数.其优势隐含周期 4~9 个不等,周期长度短的 5 a,长的 25 a.复相关系数也都通过了置信度为 0.05 时的显著性检验.

2.2 历史资料的拟合

根据确定的预报模型(方程(4)),对 1951~2000 年长江宜昌站的年平均流量进行拟合.以相对误差小于 20% 作为合格标准,统计结果表明历史拟合合格率为 100%,拟合效果令人满意.图 1 给出了宜昌站年平均流量的实际曲线和拟合曲线,可以看出 2 条曲线走势基本一致,特别是 20 世纪 90 年代以来 2 条曲线基本重合,除有 2 年的相对误差大于 3% 外,其余的相对误差都小于 3%.

2.3 宜昌站年平均流量的试预报

模型的实际意义在于对未来几年的流量进行预测,利用方程(4)试预报 2001~2005 年宜昌站的年平均流量,结果见表 2.由表 2 可知,预报误差都在允许误差之内,相对误差均小于 15%,合格率达到 100%,预报精度较高.

表 1 优势因子的优势隐含周期统计
Table 1 Statistical table of dominant hidden cycles of dominant factors

优势因子	优势隐含周期/a	复相关系数
x_1	17 22 21 25 15 13 18	0.9562
x_2	25 16 22 19 9 15	0.9012
x_3	25 18 24 16 22 21 15 23	0.9564
x_4	25 20 23 22	0.8404
x_5	25 17 23 16 6 7 19 11 5	0.9634

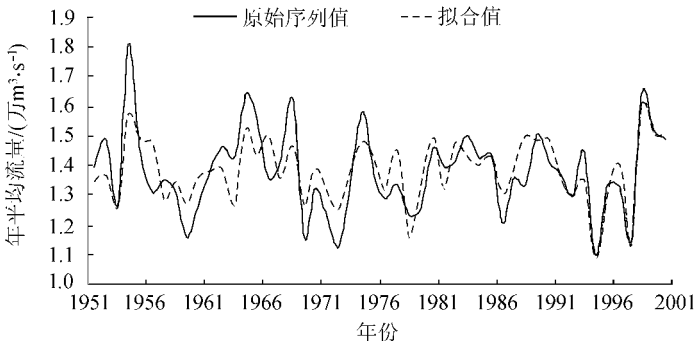


图 1 宜昌站年平均流量的实际曲线和拟合曲线
Fig. 1 Actual and fitting curves of mean annual discharge at Yichang station

3 结 论

- a. 多因子逐步回归周期分析法结合了逐步回归和逐步回归周期分析 2 种方法的优点,既可以增长预见期,又能充分利用物理因子的重要信息,同时在计算过程中还能挑选出起显著作用的优势因子和优势隐含周期.
- b. 通过分析,优势因子为 4 月冷空气、亚洲区 8 月极涡面积指数、北半球 8 月极涡中心强度、12 月西太平洋副高面积指数及太平洋区 9 月极涡面积指数,表明宜昌的年平均流量与冷暖空气的活动和强度关系极为密切.
- c. 经分析,各优势因子的优势隐含周期 4~9 个不等,周期长度短的 5 a,长的 25 a,且各优势因子普遍都具有 25 a 和 22 a 的周期.
- d. 多因子逐步回归周期分析法对长江宜昌站年平均流量的拟合和试预报合格率都达到了 100%,具有较高的可信度,这种方法具有很高的实用价值,不失为中长期水文预报的一种较好方法.

参考文献:

[1] 鹿坤.浅谈中长期水文预报方法[J].贵州水力发电,2004,18(2):17-19.

[2] 王钦钊,向奇志.中长期水文预报方法的探讨[J].江西水利科技,2006,32(2)87-90.

[3] 范钟秀.中长期水文预报[M].南京:河海大学出版社,1999.

[4] 汤成友,宫学文,张世明.现代中长期水文预报方法及其应用[M].北京:中国水利水电出版社,2008.

[5] 江秋阳,宋飙,胡文全.线性回归最小一乘法在水文流量预测中的应用[J].水利水电科技进展,2004,24(增刊):72-74.

[6] 曾小凡,梁忠民,朱成涛.长期水文统计预报方法在长江大通站径流预报中的应用[J].河海大学学报:自然科学版,2003,31(增刊):155-157.

[7] 李茂贵,朱承山.酒埠江水库入库流量预报的逐步回归方法[J].水电能源科学,2001,19(1):49-51.

[8] 李杰友,熊学农,罗清标.新丰江水库月径流预报模型[J].河海大学学报,1998,26(5):104-106.

[9] 葛朝霞,王会容,曹丽青,等.柘溪水库月入库流量预报方案的建立与对比分析[J].河海大学学报:自然科学版,2004,32(5)508-511.

[10] 葛朝霞,顾月红,曹丽青,等.海温和气象因子对洪家度水电站月径流量的影响[J].河海大学学报:自然科学版,2006,34(6)606-609.

[11] 何丽娟,白玉良,赫明天.双重逐步回归分析在中长期水文预报中的应用[J].东北水利水电,2006,24(2)34-35.

[12] 张利平,王德智,夏军等.基于气象因子的中长期水文预报方法研究[J].水电能源科学,2003,21(3)4-6.

[13] 冯利华.基于优势隐含周期分析的气象灾害预报[J].灾害学,2007,22(2):15-17.

Application of multi-factor stepwise regression cycle analysis in medium and long-term hydrological forecast

GE Zhao-xia^{1, 2}, XUE Mei^{1, 2}, SONG Ying-ling^{1, 2}

- (1. State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China ;
2. College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China)

Abstract : A medium and long-term hydrological forecast model of the mean annual discharge at Yichang station on the Yangtze River was established using the multi-factor stepwise regression cycle analysis. The result shows that the dominant factors are mainly a group of indices that can reflect the activity of the cold and warm air, and the activity and the intensity of the cold and warm air have a great impact on the mean annual discharge at Yichang station. The various dominant factors have 4-9 dominant hidden cycles and generally have cycles of 25 years and 22 years. The simulation and forecast accuracy of the mean annual discharge at Yichang station on the Yangtze River with the multi-factor stepwise regression cycle analysis reaches 100%. The present method is highly reliable.

Key words : medium and long-term hydrological forecast ; stepwise regression ; dominant factor ; dominant hidden cycle ; Yichang station

表 2 宜昌站年平均流量试预报成果

Table 2 Predicted results of mean annual

预测年份	年平均流量/(m ³ ·s ⁻¹)		相对误差/%	合格与否
	预测值	实测值		
2001	12 944	13 135	1.46	合格
2002	14 242	12 385	- 14.99	合格
2003	12 867	12 933	0.50	合格
2004	14 512	13 081	- 10.94	合格
2005	13 893	14 397	3.50	合格