

DOI:10.14188/j.1671-8844.2015-06-001

文章编号:1671-8844(2015)06-0739-05

中长期径流预报的时间尺度

冼翠玲,张艳军,邹霞,张素琼,邹静

(武汉大学水资源与水电工程科学国家重点实验室,湖北 武汉 430072)

摘要:为研究时间尺度在中长期水文预报中的影响,选取在实际运用中较成熟的预报模型——最近邻抽样回归模型对不同时间尺度的径流序列进行建模预报.通过对三峡水库 1980—2012 年的逐日平均入库流量,高场、汉口、大通 3 个水文站 1980—2012 年的逐日平均流量,进行时间聚集,分别得到对应的 3 d、周、旬、月、季、年的平均径流量.用所建模型对以上 4 个站点的 6 个不同时间尺度的径流量进行模拟预测,结果显示随着时间尺度的增加,模拟预测结果的相对误差并非单调递增或者递减,而是先变小,后变大.研究中长期水文预报中时间尺度给预报结果带来的影响,可为水文预报精度的提高开辟新思路.

关键词:中长期径流预报;最近邻抽样回归模型;时间尺度

中图分类号:TV 11 文献标志码:A

Analysis of hydrological time-scale of mid-long term runoff forecasting

XIAN Cuiling, ZHANG Yanjun, ZOU Xia, ZHANG Suqiong, ZOU Jing

(State Key Laboratory of Water Resources and Hydropower Engineering
Science, Wuhan University, Wuhan 430072, China)

Abstract: In order to study the influence of hydrological time-scale on mid-long term runoff forecasting, the Nearest Neighbor Bootstrapping Regressive Model which well applied in practice is selected and established on different time scales for runoff prediction. Taking daily inflow series of Three Gorges Reservoir and daily runoff series of Gaochang, Hankou, and Datong hydrological stations from 1980 to 2012 as the research subjects, the corresponding average runoff series of six time scales (three days, week, ten days, month, season, year) are obtained by time aggregation. The runoff series at those six different time scales are simulated. The results show that the prediction accuracy is relatively low at smaller or larger scales. This study can provide new ideas for improving the accuracy of hydrological forecasting.

Key words: mid-long term runoff forecasting; nearest neighbor bootstrapping regressive model; hydrological time-scale

中长期径流预报,通常泛指预见期超过流域最大汇流时间,且在 3 d 以上、1 a 以内的径流预报^[1,2],其中预见期在 3~10 d 的称为中期预报,预见期在 10 d~1 a 的称为长期预报,若其预见期超过 1 a,则称为超长期预报.中长期径流预报所具有的较长预见期,能够在一定程度上对防洪抗旱、兴利调度等各用水部门起指导作用,使用水获取最大的效

益.而另一方面,正是由于其较长的预见期,使得直接或间接影响因素变化的不确定性增大,造成预报难度进一步提高,但要使预报结果具有一定的意义及价值,则预报必须具有准确性、可靠性和及时性.用于中长期径流预报的方法主要有传统方法及新方法两大类,传统方法包括成因分析法、水文统计法;新方法包括模糊分析、人工神经网络、支持向量机等

收稿日期:2015-02-23

作者简介:冼翠玲(1991-),女,硕士研究生,研究方向为现代水文学, E-mail: xiancuiling0426@126.com.

通讯作者:张艳军(1982-),男,博士,副教授,研究方向为水文水环境, E-mail: zhangyj1015@gmail.com.

基金项目:国家自然科学基金项目(编号:51209162,51379149,51279140,51279143).

方法.

多年以来,对中长期径流预报的研究重点都集中在预报方法的算法上,试图通过算法上的改进来提高中长期径流预报的精确度,而对从其他角度来提高预报精度的研究则较少.水文尺度的提出^[3]给中长期水文预报带来了新的研究思路,水文尺度是指水文过程、水文观测或水文模型的特征时间或长度,研究较多的分别有年、季、月、旬、周这几个尺度,在 Dooge^[4]的概念里这些尺度均属于中观时间尺度,也是流域水文学研究的主要时间尺度^[5].但是,现阶段的研究成果均集中在单个时间尺度的预报上,包括年径流量预报^[6,7]、季径流量预报^[8,9]、月径流量预报^[10,11]、旬径流量预报^[12,13],都取得了不错的成果.王文川等^[14]建立基于 PSO 的 SVM 年径流预报模型并应用于伊犁河雅马渡水电站的年径流预报,结果表明该模型值得在年径流预报中推广应用;鲁帆等^[15]建立基于 BP 神经网络的贝叶斯概率枯季径流预报模型,在枯季径流预报中取得良好效果;刘晔^[16]采用不同预报模型对三峡 2 个控制站进行不同尺度的预报研究,发现在旬径流预报和月径流预报中最近邻抽样回归模型预报效果较理想.然而很少人对时间尺度进行比较研究,高燕^[17]在研究流域洪水预报模型时指出水文模型模拟尺度应该有个上下阈值,并在阈值范围内存在一个最佳模拟尺度;明波等^[18]采用 4 种不同的模型对宜昌水文站进行不同尺度的径流预报,对比分析结果表明各模型在日尺度模拟效果均较好,月尺度预报效果较差.

因此,时间尺度问题值得深入研究,本文尝试将时间尺度引入中长期水文预报中,通过时间聚集、时间解集的方法进行不同时间尺度的转换,选择对最近邻抽样回归模型在不同预报时间尺度作径流预报,分析不同预见期对径流预报精度的影响.

1 研究方法

1.1 最近邻抽样回归模型

国内外常用的水文预报方法有 3 类:确定性预报方法、随机水文模型预报方法和确定与随机相耦合的预报方法.这些传统预报方法一般假定水文水资源系统是线性的,而真实水文水资源系统是高维、非线性、非正态的,故上述方法预测效果不尽如人意.究其原因,除了水文水资源系统具有不确定性和随机性外,还存在着灰色性、模糊性、非线性和混沌性,因而一些新预报方法,如灰色系统预测法、模糊系统预测模型、人工神经网络模型、水文混沌预测模型等应运而生.这些新理论、新方法预测精度有一定

提高,但它们仍然没有摆脱“假定-拟合-预测”这种模式.近几十年来发展了一种探索性的数据分析方法——投影寻踪技术,能较好地解决非线性、非正态问题,且预测效果较好,但该技术数学理论较复杂,编程困难.为避免这些问题,本文使用基于数据驱动、不需识别参数的非参数模型——最近邻抽样回归模型^[19](nearest neighbor bootstrapping regressive model,简称 NNBR).它的基本原理是:客观世界的发生与发展存在着一定联系,未来的运动轨迹与历史具有相似性,即未来发展模式可以从已知众多模式中去寻求^[20].此模型结构简单,使用方便,不需对数据作预处理,避免了模型选择和参数不确定性的问题,对预测而言较有优势.

已知水文时间序列 $\{X_t\}_n$, X_t 依赖于前 P 个相邻历史值 $X_{t-1}, X_{t-2}, \dots, X_{t-P}$. 定义特征矢量 $D_t = (X_{t-1}, X_{t-2}, \dots, X_{t-P})$, X_t 为 D_t 的后续值($t = P+1, P+2, \dots, n$).最近邻抽样回归模型认为:在已有的特征矢量 D_t ($t = P+1, P+2, \dots, n$) 中,总有 K 个与当前特征矢量 D_i 最近邻(最相似),从而在 D_i 中可得到 K 个最近邻特征矢量 $D_{1(i)}, D_{2(i)}, \dots, D_{K(i)}$,其对应的后续值分别为 $X_{1(i)}, X_{2(i)}, \dots, X_{K(i)}$ ^[17]. D_i 与 D_t 间的欧氏距离越小,则视为越近邻.欧氏距离计算:

$$r_{t(i)} = \left[\sum_{j=1}^P (d_{ij} - d_{ij})^2 \right]^{1/2} \quad (1)$$

式中: $r_{t(i)}$ 表示 D_i 与 D_t 间的欧氏距离; d_{ij}, d_{ij} 分别为 D_i, D_t 的第 j 个元素; P 为特征矢量维数.

K 个最近邻特征矢量 $D_{1(i)}, D_{2(i)}, \dots, D_{K(i)}$ 与 D_i 间的欧氏距离记为 $r'_{1(i)}, r'_{2(i)}, \dots, r'_{K(i)}$. $r'_{j(i)}$ 越小,说明 $D_{j(i)}$ 与 D_i 越近邻,则 $X_i = X_{j(i)}$ 的可能性 $W_{j(i)}$, $j = 1, 2, \dots, K$ 越大^[19].把 $W_{j(i)}$ 记为 $X_{j(i)}$ 的抽样权重,可见 $W_{j(i)}$ 与欧氏距离 $r'_{j(i)}$ 成反比.

根据上述思想,单因子 NNBR 模型的基本形式为

$$X_i = \sum_{j=1}^K W_{j(i)} X_{j(i)} \quad (2)$$

式中: K 称为最近邻数; $\sum_{j=1}^K W_{j(i)} = 1.0$; 其余符号同前.

最近邻数 K 一般取 $K = \text{int}(\sqrt{m})$, int 表示取整, m 为特征矢量个数;特征矢量维数 P 为与 X_t 相关的前相邻历史值个数,由时间序列自相关图和偏相关图确定,同时要求满足 $P \geq 2$.抽样权重 $W_{j(i)}$ 的选择具有多样性,一般要求 $\sum_{j=1}^K W_{j(i)} = 1.0$,选择原则是:1) 距离与抽样权重成反比,距离越小,权重越

大;2) 权重函数简便实用. 目前被广泛使用的计算公式如下:

$$W_{j(i)} = \frac{1}{\sum_{j=1}^K \left(\frac{1}{L}\right)}, j = 1, 2, \dots, K \quad (3)$$

式中: $W_{j(i)}$, K 同前. 当 K 确定后, $W_{j(i)}$ 只需计算 1 次.

1.2 不同时间尺度转换

文中的时间尺度是针对预报时间而言的,即将预报时间分为年、季、月、旬、周及 3 d 这 6 个时间尺度,在研究不同时间的过程中,需要对已有资料进行尺度转换,时间尺度转换就是使转换对象在时域上压缩或扩展而幅度不发生变化,对应的方法分别为时间聚集和时间解集^[17].

时间聚集方法比较简单,就是将小尺度的资料在时间上进行累积求和,根据需要的的时间尺度,累积成所需要的等时间间隔的数据资料. 许多水文要素都是通过累加而成的,例如年水量由月水量累加而成,月水量由日水量累加而得等.

时间解集方法是根据已有的高聚水平的时间序列来模拟低聚水平时间序列的过程,即采用一定的规则,把大时间尺度的资料分解到较小时间尺度上. 解集是与聚集刚好相反的一个过程,例如某年的年水量可以分解成年内各月水量,月水量可以分解成该月各日水量. 其显著特点在于能保持水量平衡和连续分解.

本次研究将各站点的逐日径流资料在时间上进行聚集,分别累积成 3 d、周、旬、月、季、年尺度上的径流量,最后平均得到以上 6 个时间尺度的平均径流量.

2 研究过程及结果

2.1 研究区域

长江流域介于东经 $90^{\circ}33' \sim 122^{\circ}25'$, 北纬 $24^{\circ}30' \sim 35^{\circ}45'$ 之间, 横跨中国东部、中部和西部三大经济区共计 19 个省、市、自治区. 流域总面积 180 万 km^2 , 占中国国土面积的 18.8%, 长江干流全长 6 300 余 km; 支流众多, 其中支流长度 500 km 以上的 18 条, 流域面积超过 1 000 km^2 支流达 437 条; 长江流域年均水资源总量 9 960 亿 m^3 , 全流域水能理论蕴藏量约 2.8 亿 kW, 可开发量约 2.6 亿 kW, 森林资源蕴藏量大, 林木蓄积量、矿产资源丰富.

研究的对象包括三峡水库、长江上游断面的高场水文站、中游断面的汉口水文站、下游断面的大通水文站, 是长江流域从上游至下游具有代表性的站

点, 分别收集了 1980—2012 年的逐日平均径流量资料. 其中需要说明的是三峡水库的逐日平均流量资料是经还原后的入库流量资料. 研究站点地理位置如图 1 所示.

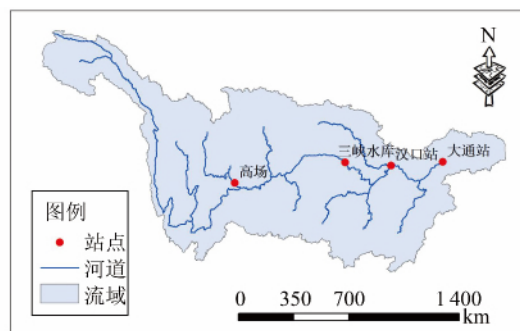


图 1 研究站点分布图

Fig. 1 Distribution of the study station

2.2 计算结果

将 1980—2012 年逐日平均流量用时间聚集成各尺度的径流量, 其中 1980—2000 年作为率定期来构造特征向量建立模型, 2001—2012 年作为检验期来预测径流检验模型的预测效果, 各站点的最近邻数 K 及特征矢量维数 P 的取值表见表 1, 模拟预测结果如表 2 及图 2 所示.

表 1 各站点不同时间尺度参数取值

Table 1 Parameter selection in different time scales of each station

	参数	3 d	周	旬	月	季	年
三峡	K	3	6	6	5	3	3
	P	6	5	3	3	3	2
高场	K	4	3	5	6	4	4
	P	2	6	6	3	5	3
汉口	K	7	6	4	5	3	5
	P	6	4	4	4	2	3
大通	K	6	7	6	3	4	4
	P	3	4	6	2	5	2

表 2 各站点不同时间尺度径流预报相对误差

Table 2 Relative error of runoff forecasting in different time scales of each station

		3 d	周	旬	月	季	年	%
三峡	率定	8.49	10.75	11.01	11.62	12.28	5.2	
	检验	11.05	15.93	16.68	18.52	24.2	11.86	
高场	率定	10.06	11.93	10.14	10.55	9.64	7.85	
	检验	15.66	17.59	17.82	18.08	17.38	15.45	
汉口	率定	3.52	8.02	9.46	12.06	12.47	5.9	
	检验	4.86	12.03	12.26	15.78	23.25	8.68	
大通	率定	3.22	5.84	8.1	12.76	13.81	5.55	
	检验	3.9	10.26	10.97	17.29	16.21	13.02	

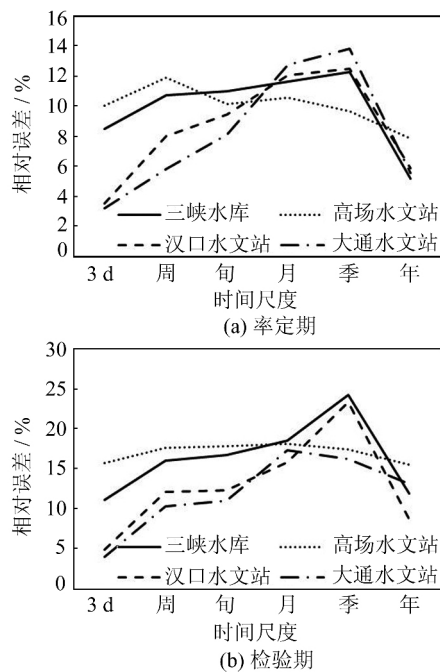


图 2 各站点径流不同预报时间尺度相对误差
Fig. 2 Relative error of runoff forecasting in different time scales of each station

从表 2 及图 2 可知:1) 对三峡水库入库径流量从 3 d~1 a 的 6 个时间尺度进行模拟与预测时,相对误差的趋势是先增加后减少的,其中在季时间尺度上达到最大(率定期 12.28%,检验期 24.20%),季时间尺度在本次计算中的效果最差,而在两端的 2 个时间尺度(3 d 时间尺度、年时间尺度)的模拟效果均较好,在检验期的相对误差控制在 12% 以内;2) 对高场水文站的径流量从 3 d~1 a 的 6 个时间尺度进行模拟与预测时,相对误差的趋势是先增加后减少的,在率定阶段,周时间尺度上的相对误差达最大(11.93%),在预测阶段,月时间尺度上的相对误差达最大(18.08%),相对误差在 6 个时间尺度上的变幅不大,率定期从 7.85%~11.93%,检验期从 15.45%~18.08%;3) 对汉口水文站的径流量从 3 d~1 a 的 6 个时间尺度进行模拟与预测时,相对误差的趋势是先增加后减少的,在较小的时间尺度(3 d)及较大的时间尺度(a)的模拟预测时,相对误差小,效果非常好;从 3 d 到季时间尺度,相对误差在逐步增加,并在季时间尺度上达最大(率定期 12.47%,检验期 23.25%);4) 对大通水文站的径流量从 3 d~1 a 的 6 个时间尺度进行模拟与预测时,相对误差的趋势是先增加后减少的,在率定阶段,季时间尺度上的相对误差最大(13.81%),在检验阶段,月时间尺度上的相对误差最大(17.29%),其中,3 d 时间尺度的效果明显优于其他 5 个不同的时间尺度,相对误差控制在 4% 以内。

2.3 结果分析与讨论

在率定期和检验期,4 个不同站点在 6 个不同预报时间尺度的径流量模拟时,相对误差的变化趋势是一致的,均是先增加后减少,曲线呈现出两端较低,中间较高的形态.这个变化趋势与张素琼等^[21]采用人工神经网络方法对水文站径流进行模拟和预测所得出的相对误差变化趋势是一致的。

在率定期,除上游的高场水文站在周时间尺度的相对误差最大外,其他各站点的相对误差最大值均出现在季时间尺度上;而张素琼等^[21]的计算结果显示,率定期上游水文站相对误差曲线最大值出现在旬尺度,中游和下游水文站在月尺度相对误差最大.在检验期,三峡水库及汉口水文站的相对误差在季时间尺度达最大,上游的高场及下游的大通水文站的相对误差在月尺度达最大;而张素琼等^[21]计算得到的检验期上游水文站和中游螺山站相对误差曲线最大值出现在季尺度,上游的汉口站和下游的大通站在月尺度上相对误差最大.总体上看,从上游至下游,小时间尺度(3 d、周、旬)径流预报相对误差逐渐变小,径流量的预报模拟效果逐渐变好;对于大时间尺度(月、季、年)趋势并不明显。

相对误差曲线呈现上述变化规律的原因是径流序列值之间同时存在相关性和线性的影响.序列值的相关性随着时间尺度的增大而减弱,而径流序列时间尺度越大,非线性特征越不明显.由于日流量资料样本非常丰富,时间序列样本的长度最接近于无穷大,而且小时间尺度序列值的相关性很强,所以 3 d 时间尺度上的模拟预测效果非常好;随着预报时间尺度增加,序列值的相关性变弱,而整体的线性特征仍不够明显,因此预报的相对误差增大,预测效果变差;但当预报时间尺度继续增加到年尺度时,由于年径流时间序列变幅在某一均值附近波动,大时间尺度综合了一些扰动因素,使其变化基本是线性的,因而预报相对误差又逐渐变小,预报模拟结果较理想。

从长江流域上游至下游计算过程中,较小时间尺度预报相对误差呈现出逐渐变小的趋势,但较大时间尺度径流预报中此趋势不明显.这一结果表明随着流域面积的增大,较小时间尺度径流序列的非线性结构也逐步减弱,计算精度得到提高,此结论与 Wang 等^[22]在 1981 年研究中发现的规律一致^[21];较大时间尺度中,预报结果可能受到人类活动以及各种水库调蓄的影响比较突出,所以上下游的相对误差规律性并不明显。

3 结语

中长期径流预报所具有的较长预见期,能够在一定程度上对防洪抗旱、兴利调度等各用水部门起指导作用,使用水获取最大效益.而另一方面,正是由于其较长的预见期,使得直接或间接影响因素变化的不确定性增大,造成预报难度进一步提高,但要使预报结果具有一定的意义及价值,则预报必须具有准确性、可靠性和及时性.目前用于计算中长期水文预报的方法已非常多,如何在现有方法的基础上,从其他角度来考虑其预报精度,即将重点转移到方法之外而非方法本身,是一个值得关注的问题.本文以时间尺度为切入点,采用中长期径流预报中较成熟的最近邻抽样回归模型对长江流域上具有代表性的三峡水库入库径流(高场、汉口、大通水文站径流)进行模拟预测,分别选取了从3 d~1 a之间的6个常用的水文预报尺度,探讨了不同预报时间尺度对于模型精度的影响,为中长期水文预报研究提供了新思路.今后可深入研究时间序列的机理,以进一步揭示本次研究的本质.

参考文献:

- [1] 陈金荣, 范钟秀. 中国中长期水文预报的现状与发展趋势[C]//长江流域规划办公室. 1981年全国水文预报学术讨论会水文预报论文集. 北京: 水利电力出版社, 1985: 174-179.
- [2] 水利电力部水文水利调度中心. SD138-85水文情报预报规范[S]. 北京: 水利电力出版社, 1985.
- [3] Blöschl G, Sivapalan M. Special issue on the scale issues in hydrological modeling[J]. Hydrology Processes, 1995, (9): 251-290.
- [4] Dooge J C I. Looking for hydrologic laws[J]. Water Resources Research, 1986, 22(9): 46s-58s.
- [5] 夏军. 水文尺度问题[J]. 水利学报, 1993, (5): 32-37.
- [6] Zhang Q, Wang B D, He B, et al. Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting[J]. Water Resources Management, 2011, 25(11): 2683-2703.
- [7] Zhou Y, Zhou P, Jin J, et al. Probabilistic forecasting model for annual runoff based on bootstrap[J]. IAHS-AISH Publication, 2011: 627-632.
- [8] Robertson D E, Pokhrel P, Wang Q J. Improving statistical forecasts of seasonal stream flows using hydrological model output[J]. Hydrology and Earth System Sciences, 2013, 17(2): 579-593.
- [9] Laugesen R, Tuteja N K, Shin D, et al. Seasonal Streamflow forecasting with a workflow-based dynamic hydrologic modelling approach[C]//19th International Congress on Modelling and Simulation. Perth, Australia, 2011: 2352-2358.
- [10] Okkan U, Dalkilic H Y. Monthly runoff model for Kemer Dam with radial based artificial neural networks[J]. Teknik Dergi, 2012, 23(2): 5957-5966.
- [11] 刘冀, 徐刚, 彭涛, 等. 基于参数优化的集对分析月径流预测[J]. 水文, 2013, 33(1): 8-11.
- [12] 梁国华, 董霞, 王国利, 等. 旬降雨径流非线性模型及其应用研究[J]. 南水北调与水利科技, 2009, 7(1): 53-55.
- [13] Liu C H, Chen C S, Su H C, et al. Forecasting models for the ten-day streamflow of Kao-Ping River[C]//Fuzzy Systems, 2006 IEEE International Conference on Vancouver, BC, IEEE, 2006: 1527-1534.
- [14] 王文川, 和吉, 邱林. 基于PSO的SVM年径流预报模型研究[J]. 人民黄河, 2012, 34(4): 17-19.
- [15] 鲁帆, 蒋云钟, 殷峻暹. 基于BP神经网络的贝叶斯概率枯季径流预报研究[J]. 水电能源科学, 2008, 26(3): 7-10.
- [16] 刘晔. 长江三峡枯季入库径流中长期水文预报研究[D]. 重庆: 重庆交通大学, 2012.
- [17] 高燕. 时间尺度对半干旱半湿润流域洪水预报模型的影响研究[D]. 太原: 太原理工大学, 2008.
- [18] 明波, 刘冀, 吕翠美, 等. 径流多时间尺度预报方法对比分析[J]. 中国农村水利水电, 2013, (3): 30-33.
- [19] 王文圣, 向红莲, 丁晶. 最近邻抽样回归模型在水文水资源预报中的应用[J]. 水电能源科学, 2001, 19(2): 8-10.
- [20] 朱聪. 径流中长期预测模型研究[D]. 成都: 四川大学, 2005.
- [21] 张素琼, 张艳军, 宋星原, 等. 基于神经网络的中长期径流预报时间尺度研究[J]. 中国农村水利水电, 2014, (8): 110-114.
- [22] Wang C T, Gupta V K. A geomorphologic synthesis of nonlinearity in surface runoff[J]. Water Resources Research, 1981, 17(3): 545-554.