

文章编号:0559-9350(2010)09-1069-05

丹江口水库中长期径流量的多模型预报结果分析及综合研究

冉笃奎^{1,2}, 李 敏³, 武 晟¹, 解建仓¹

(1. 西安理工大学 西北水资源与环境生态教育部重点实验室, 陕西 西安 710048;

2. 汉江水电开发有限责任公司, 湖北 丹江口 442700; 3. 西北勘测设计研究院, 陕西 西安 710065)

摘要: 水库中长期入库径流量的预报方法有多种模型, 各有优缺点, 预报结果也存在一定差异, 应用过程中给实际决策造成了一定困难。本文以丹江口水库月径流量为研究对象, 通过分析多种预报模型计算结果的变化规律, 将精度和稳定性等作为性能指标, 构建以加权平均和最小二乘法原理为基础的综合预报模型, 并对丹江口水库月径流预报进行了建模计算。应用结果表明, 该综合模型提高了预报精度, 增强了预报稳定性, 使其具有更高可信度, 为入库径流量预报提供了一个有效的分析方法, 为水库调度决策提供了便利。

关键词: 丹江口水库; 入库径流量; 中长期径流量预报; 综合预报模型

中图分类号: P338

文献标识码: A

1 研究背景

丹江口水利枢纽是南水北调中线水源工程, 具有防洪、供水、发电、灌溉、航运和水产养殖等综合功能。水库入库径流量的准确预报是指导水资源综合利用、科学管理和优化调度最重要的依据。

近年来中长期径流预报模型的研究发展较快, 已建立了多种预报方法, 例如张晓伟^[1]等针对灰色自记忆建模中存在的缺陷, 以组合预报思想为基础, 采用BP神经网络方法修正灰色自记忆模型存在的误差, 从而建立了基于BP神经网络的灰色自记忆径流预报模型; 林剑艺等^[2]分析了支持向量机在中长期径流预报中应用的可行性; 杨道辉等^[3]从减少传统径流预报中容易产生累计误差的角度出发, 提出了一种基于粒子群小波人工神经网络组合模型的月径流过程预报算法; 于国荣等^[4]构建了混沌时间序列的支持向量机径流预报模型; 汪丽娜等^[5]探讨了支持向量机方法在降雨-径流预报中的应用, 弥补了人工神经网络在预报应用中的部分缺陷; 陈南祥^[6]等针对径流预报影响因素众多, 信息相互重叠的问题, 采用偏最小二乘回归与神经网络耦合的方法, 分析了该方法在径流预报中的实际应用效果; 曹辉等^[7]对径流预报方法进行了简单比较。从总体上看, 现有的水文预报方法主要分为两大类^[8], 一类是以水文学概念为基础, 对径流的产流过程与河道演进过程进行模拟; 另一类是以建立输入输出数据之间的最优数学关系为目标的建模方法, 近年来得到了快速发展, 主要是因为摆脱了一些无法用概念模型精确计算的过程的困扰, 同时水文数据获取能力的提高, 计算能力的飞速发展, 使得这种模型在水文预报领域受到了广泛的关注。

预报模型有各自的优势, 但预报结果的不一致性给一些决策工作带来了难度。查阅大量资料发现, 在其他研究领域也存在多模型计算结果不一致问题, 一些学者给出了较好的解决方法。例如王鹏武^[9]提出了对投入、产出指标采用不同评价方法的综合法, 秦寿康等^[10-11]提出采用兼容度、差异度极大极小化模型对多套评估结果进行优化分析。本文研究的水库径流量预报服务于水库调度计划制

收稿日期: 2009-02-20

基金项目: 国家自然科学基金项目(50979088); 国家“863”计划研究资助项目(2006AA01A126)

作者简介: 冉笃奎(1973-), 陕西宝鸡人, 博士生, 主要从事水资源利用研究。E-mail: wshmonkey@yahoo.com.cn

定，以年为预报期，以月为时段。从分析多种评估结果的特点入手，结合本文实际问题，建立综合研究模型，以期得到预报精度更高、稳定性更强的预报结果。

2 丹江口水库中长期径流量多模型预报分析

回归模型(简称PLS回归)在实际水文中长期预报应用最为普遍，该方法可减弱自变量间多重相关性在系统建模中的不利影响，具有可操作性强、便于理解的优点；人工神经网络模型(简称ANN)是近年来倍受关注的一种建模方法，该方法摆脱了一些无法用概念模型精确计算的困扰，可以直接使用历史数据作为样本实现建模；支持向量机(简称SVM)则是最近十几年开始研究的一项新技术，以统计学理论为基础，在解决小样本、非线性及高维模式识别问题中表现出许多特有的优势。为分析各种模型优缺点，本文采用这3种较为典型的方法构建丹江口水库中长期径流量的预报模型。

以丹江口水库1961—1980年的月平均径流量为样本，再以1981年的月平均径流量为预报检验样本，预报结果见表1。从表1可以看出，SVM的预报能力最强，月径流量的平均误差为4.21%，尽管PLS回归模型的预报精度是最低的，但平均相对误差也仅达到了6.92%，可见3种预报模型的预报精度差别不大。

表1 模型预报结果对比 (单位：m³/s)

时间	标准值	预报值		
		PLS回归	神经网络	SVM
1981年1月	325	351	342	311
1981年2月	319	315	299	302
1981年3月	524	562	562	551
1981年4月	1300	1380	1370	1328
1981年5月	436	410	448	419
1981年6月	928	1005	875	965
1981年7月	3020	3299	2828	2866
1981年8月	4330	4012	4143	4546
1981年9月	4340	3965	4584	4146
1981年10月	1600	1742	1653	1701
1981年11月	754	696	700	765
1981年12月	487	508	470	504
平均相对误差/%		6.92	5.24	4.21

再对单项预测的精度做进一步分析，结果见表2。从表2可以看出，尽管SVM模型有着最高的预报精度，即平均误差最小，但是在部分月份的预报精度并没有另外两种模型高，例如2、5、8和10月。由此可见，除了要关注模型的预报精确性外，还应考虑预报误差的稳定性。

表2 模型预报精度比较

时间	标准值	相对误差/%			误差最小模型	时间	标准值	相对误差/%			误差最小模型
		PLS回归	神经网络	SVM				PLS回归	神经网络	SVM	
1981年1月	325	8.15	5.19	4.21	SVM	1981年7月	3020	9.23	6.36	5.11	SVM
1981年2月	319	1.21	6.27	5.43	PLS回归	1981年8月	4330	7.35	4.33	4.99	神经网络
1981年3月	524	7.32	7.23	5.11	SVM	1981年9月	4340	8.64	5.62	4.46	SVM
1981年4月	1300	6.14	5.42	2.14	SVM	1981年10月	1600	8.87	3.32	6.32	神经网络
1981年5月	436	6.02	2.71	3.84	神经网络	1981年11月	754	7.63	7.18	1.51	SVM
1981年6月	928	8.26	5.72	4.02	SVM	1981年12月	487	4.27	3.57	3.41	SVM

3 综合研究模型的构建方法

3.1 预报结果变化规律分析 对多个中长期径流量预报模型结果进行独立分析发现，在模拟精度及预报精度都相似的情况下，预报误差的稳定性仍有着较大的区别，以两种较为典型的情况为例，加以说明，见图1。图1显示了两种模型A、B的预报结果，平均误差均为20，但稳定性显然是模型A比模型B高。然而，在单个月份的预报中，模型A的预报精度只有4个月比模型B高。这说明，两种模型的预报能力存在一定的互补性。根据这一特点，本文采用加权平均法对具有互补性的多个预报模型进行综合分析，在注重预报精度的同时关注预报能力的稳定性。

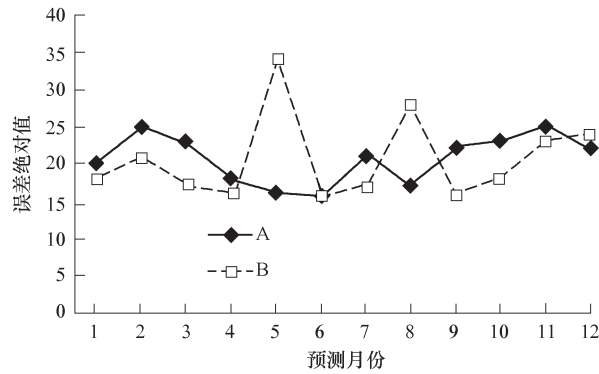


图1 预报误差变化规律分布

3.2 中长期径流量预报综合研究模型的建立 权重的分配应当考虑预报精度和稳定性两方面的问题。基于此，建立综合研究模型的过程如下。

(1)权重的确定。模型主要从提高稳定性和精度两方面考虑，稳定性主要表现在波动的幅度和整体的波动性两个方面。预报模型的波动幅度由最大与最小误差的差值表征，采用如下计算方法：

$$E_b = D_{\max} - D_{\min} \quad (1)$$

式中： E_b 为波动幅度； $D_{\max}$ 、 $D_{\min}$ 分别为最大、最小相对误差。

整体波动性主要考察多数预报结果的相对变化是否较小，在对众多建模方法进行分析比较后，采用如下计算方法对每个预报模型的整体波动性进行衡量：

$$E_z = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n \sqrt{|D_i - D_j|}}{\sum_{i=1}^n i} \times 100 \quad (2)$$

式中： E_z 为整体波动幅度； $D_{i,j}$ 为预测的第*i*、*j*个预测月份的相对误差，*n*为预报样本数量。

精度采用相对误差的均方根来表征，即：

$$E_j = \sqrt{\frac{1}{n} \sum_{i=1}^n D_i^2} \quad (3)$$

模型首先要保证预报结果的精度不降低；其次是预报的稳定性。在稳定性中首先考虑整体的波动性，其次是波动的幅度。基于以上分析，设共有*m*个模型，第*I*个预报模型的权重分配采用如下方法计算：

$$W_I = \frac{1 - 0.5 \times \frac{E_{jI}}{\sum_{k=1}^m E_{jk}} + 0.3 \times \frac{E_{zI}}{\sum_{k=1}^m E_{zk}} - 0.2 \times \frac{E_{bI}}{\sum_{k=1}^m E_{bk}}}{m - 1} \quad (4)$$

(2)根据稳定性和精度确定的权重，使加权平均值与真值更为接近的同时，误差也较为稳定，所

以可采用最小二乘法原理构建综合研究模型的构架。设综合分析每个月份的预报结果为 p ，则有

$$p = a' + b' \sum_{k=1}^m W_k a_k \tag{5}$$

式中， a' 、 b' 为任意实数， a_k 为各模型预报结果。

3.3 综合研究模型的建模步骤 3.2 节给出了基于各种预测结果的综合研究方法，具体计算分为 3 步。首先确定预报精度能够达到设计要求的基本预报模型，计算各自的预测结果；其次对比预测误差的特点，考察结果之间是否存在互补性，如果具有相互补充的可能，则可采用精度、稳定度等指标计算权重，并用最小二乘法确定参数 a' 、 b' 值，最后由式(5)计算得到综合分析结果。综合研究方法的详细步骤见图 2。

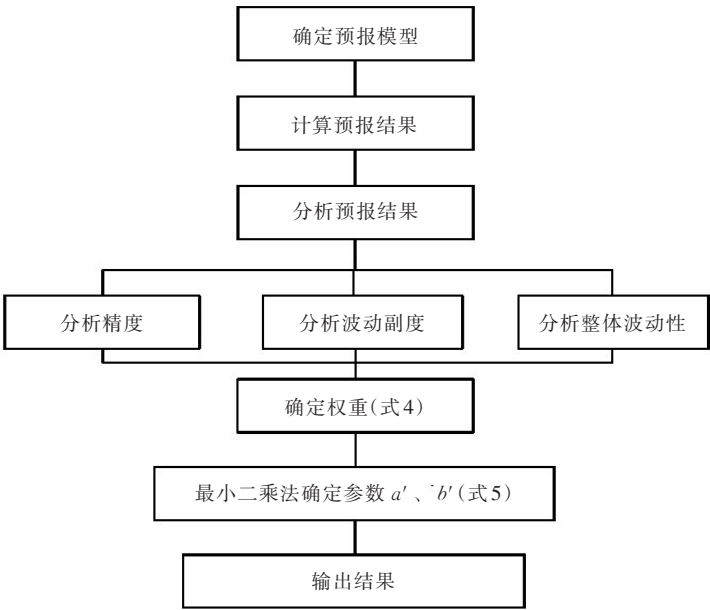


图2 综合研究方法步骤

4 综合研究模型在丹江口水库月径流量预报中的应用

从表2可以看出，PLS回归、神经网络和SVM模型的预报结果精度都较高，并且从预报误差的波动性可以看出三者具有一定的互补性，因此可以进行进一步的综合分析。

根据综合研究模型的建模方法，确定PLS回归、SVM和神经网络模型的权重分别为0.29、0.36和0.35，计算出 a' 、 b' 参数分别为1.100和1.012。

将参数带入式(5)，得到综合计算结果，并与预报模型结果从精度、波动幅度及整体波动性3个方面进行比较，如表3所示。从表3的数据比较可以看出，综合研究模型不论是在精度还是稳定性方面都有较好的表现，达到了建模的预期目标。

5 结语

水库来水量的预报，受气象、水文地质等多种因素的影响，现有的技术还不能建立一个完全准确的预报模型，对于已有的方法和模型，也有许多不完善的地方，同时预报结果的不完全一致给一些决策工作带来了困难。通过对多种模型预报结果的分析，结合其各自的优缺点，采用加权平均法，在确定权重时以预报结果的精度、整体稳定性和波动幅度3个影响因素为分配条件，并用最小二乘法构建综合研究模型，可以对多个模型取长补短，获得更优的预报结果。从丹江口水库来水量预

表3 模型性能比较

模型	神经网络	PLS回归	SVM	综合研究模型
精度/%	5.43	7.26	4.41	4.11
波动幅度/%	14.41	17.87	11.75	10.97
整体波动性/%	2.33	2.62	2.11	1.99

报的实际应用来看,该方法在精度和稳定性方面都有提高,为决策性工作提供了一个科学的分析方法,但严密的论证工作还有待进一步分析。另外,本文研究的水库来水流量预报是对月流量进行预报,适用于一般性的作业预报,从综合分析方法可以看出,参数的确定主要依赖于对历史数据进行多次的训练,所以该方法还不能应用于对洪峰流量的分析,也是需要做深入分析的一个方面。

参 考 文 献:

- [1] 张晓伟,沈冰,黄领梅,等.基于BP神经网络的灰色自记忆径流预测模型[J].水力发电学报,2009,28(1):68-71.
- [2] 林剑艺,程春田.支持向量机在中长期径流预报中的应用[J].水利学报,2006,37(6):681-686.
- [3] 杨道辉,马光文,杨梅,等.粒子群小波人工神经网络组合模型的径流预测[J].水力发电,2009,35(1):4-6.
- [4] 于国荣,夏自强,等.混沌时间序列支持向量机模型及其在径流预测中应用[J].水科学进展,2008,19(1):116-122.
- [5] 汪丽娜,李粤安,陈晓宏,等.基于支持向量机的降雨—径流预测研究[J].水文,2009,19(1):13-16.
- [6] 陈南祥,曹连海,徐建新.偏最小二乘回归神经网络在径流预测中的应用[J].灌溉排水学报,2005,24(1):54-56.
- [7] 曹辉,等.径流预测方法对比分析[J].人民黄河,2009,31(9):36-37.
- [8] 王文,马骏.若干水文预报方法综述[J].水利水电科技进展,2005,25(1):56-60.
- [9] 王鹏武.系统综合评价的一个方法[J].系统工程理论与实践,1998,7:140-144.
- [10] 秦寿康.评价方案优化系统[J].系统工程学报,2002,17(2):143-149.
- [11] 傅荣林,秦寿康,陈湛本.兼容多个综合评估方案及其分类的数学模型[J].系统工程学报,1999,14(2):179-183.

Research on multi-model forecasts in mid-long term runoff in Danjiangkou Reservoir

RAN Du-kui^{1,2}, LI Min³, WU Sheng¹, XIE Jian-cang¹

(1. Xi'an University of Technology, Xi'an 710048, China;

2. Hanjiang Hydropower Development CO., LTD, Hubei 442700, China;

3. Northwest Investigation And Design Institute Xi'an 710065, China)

Abstract: There are many models for forecasting mid-long term runoff and give different forecast results, which result in difficulties in decision making. In this paper the monthly runoff of the Danjiangkou Reservoir is adopted as the research objective, to analyze the variation of forecast results. Regarding the precision and stability as the performance index, an comprehensive forecast model is established based on the theory of weighted average and least-square method. The monthly runoffs of different models show that the comprehensive model could significantly improve the accuracy of prediction and the stability of forecast. The reliability of results can be raised.

Key words: Danjiangkou reservoir; inflow runoff; Mid-long Term Forecasting; Comprehensive forecast model

(责任编辑:王成丽)