

# 基于时变权重组合与贝叶斯修正的中长期径流预报

李伶俐<sup>1</sup>, 王银堂<sup>1</sup>, 胡庆芳<sup>1</sup>, 刘勇<sup>1</sup>, 刘定忠<sup>2</sup>, 崔婷婷<sup>1</sup>

(1. 南京水利科学研究院水文水资源与水利工程科学国家重点实验室, 南京 210029;

2. 云南龙江水利枢纽开发有限公司, 云南 德宏 678400)

**摘要:**高精度的中长期径流预报信息是水资源规划管理与水利工程经济运行的重要基础支撑。论文在组合预报与误差修正2类径流预报后处理方法串联应用的技术框架下,考虑径流的高度非平稳与非线性等特征,提出了基于时变权重组合和贝叶斯修正的中长期径流预报方法。应用该方法开展了云南龙江水库年、月入库径流预报的实例研究,结果表明时变权重组合平衡了已建立的随机森林与支持向量机模型在建模期与检验期预报性能的差异,经贝叶斯修正后的预报精度接近或优于两阶段各自的最优单一模型。根据年径流预报结果判断水文年型的正确率达到77.2%,月预报径流的不确定性系数超过0.90。因此,该方法在提升中长期径流预报精度方面具有积极效果。

**关键词:**中长期径流预报;时变权重组合;贝叶斯修正;云南龙江水库

中长期径流预报是理解和认识未来水文情势演变的重要技术手段,对于防汛抗旱、水资源规划管理及水利工程经济运行等具有重要意义<sup>[1]</sup>。然而,大气环流、降水、下垫面变化及水利工程调控等多重自然与人为因素的交织影响增大了中长期径流预报的不确定性及难度。因此,如何提高中长期径流预报精度一直是水文水资源领域的研究重点和难点。

近年来,学界围绕预报因子识别、预报模型构建等环节开展了大量研究,最优子集回归<sup>[2]</sup>、Copula熵<sup>[3]</sup>、决策树<sup>[4]</sup>等丰富了最佳预报因子的筛选方法,同时随机森林<sup>[4-5]</sup>、神经网络<sup>[6-7]</sup>等数据驱动模型的引入大大丰富了预报模型,这些研究不同程度地推动了中长期径流预报的发展。但中长期径流预报的精度与实践需求之间仍存在明显差距,因此国内外学者从不同途径开展提升预报精度的探索,而对预报结果后处理则是预报技术链条上一个重要的努力方向。后处理方法主要包括组合预报和预报误

差修正2大类。前者主要通过融合多种模型提升预报精度,如Chowdhury等<sup>[8]</sup>、Zhang等<sup>[9]</sup>与Men等<sup>[10]</sup>采用时变权重组合、贝叶斯加权集合与交叉熵等方法进行组合预报,降低了中长期径流预报的不确定性。但组合预报在局部时段也有引入较大误差的风险,导致计算结果可能并未优于最佳单一预报模型。如孙惠子<sup>[11]</sup>对比分析了多种线性和非线性组合方法的适用条件及效果,发现组合预报的平均精度常介于最优与最劣模型之间。而在预报误差修正方面,一系列适应径流非线性特点的方法得到深入探索,如Wu等<sup>[12]</sup>基于建模期实测值与预报值的频率曲线建立非线性误差订正函数,显著降低了海流兔河流域年径流预报误差;Liang等<sup>[13]</sup>、鲁帆等<sup>[14]</sup>和冯小冲<sup>[15]</sup>根据贝叶斯原理建立了确定性预报结果的后验概率分布,有效提高了预报精度。综观已有文献,组合预报和预报误差修正2类方法独立应用于中长期径流预报的研究较为多见,但两者的联合应用却鲜有报道。因此,将2种方法结合,分析其对

收稿日期:2019-11-03;修订日期:2019-12-16。

基金项目:国家重点研发计划项目(2016YFC0400902, 2016YFC04009010);国家自然科学基金项目(51609140, 51809252)。

[Foundation: National Key Research and Development Program of China, No. 2016YFC0400902 and 2016YFC04009010; National Natural Science Foundation of China, No. 51609140 and 51809252.]

第一作者简介:李伶俐(1992—),男,山西吕梁人,硕士,工程师,主要从事水文水资源研究。E-mail: ljli@nhri.cn

引用格式:李伶俐,王银堂,胡庆芳,等. 基于时变权重组合与贝叶斯修正的中长期径流预报[J]. 地理科学进展, 2020, 39(4): 643-650. [Li Lingjie, Wang Yintang, Hu Qingfang, et al. Mid- and long-term runoff prediction based on time-varying weight combination and Bayesian correction. Progress in Geography, 2020, 39(4): 643-650.] DOI: 10.18306/dlkxjz.2020.04.011

改善中长期径流预报效果的作用,值得深入探讨。

鉴于此,本文基于组合预报与误差修正相结合的中长期径流预报技术框架,考虑适应径流过程高度非平稳和非线性等特点,提出基于时变权重组合和贝叶斯修正的中长期径流预报方法。针对云南龙江水库年、月入库径流预报,以建立的随机森林与支持向量机模型为基础,应用所提方法开展预报试验,验证其对改善预报精度的有效性,以期为中长期径流预报提供一种新方法。

### 1 研究方法

组合预报与误差修正相结合的中长期径流预报技术框架以多种模型的预报结果为基础,通过某种集成方法得到组合预报结果,然后将此结果输入误差校正模型,进而得到中长期径流预报的最终结果。对于组合预报方法,其关键环节是合理确定各模型成员的权重。目前已有的误差倒数平方、多元线性回归、信息熵等恒定权重方法,对于单一模型预报误差的非平稳性考虑不足,而采用时变权重方法可以根据预报性能的变化,对权重适时作出调整,从而优化组合预报结果。在误差修正方面,时间序列分析、卡尔曼滤波等方法已经日趋发展成熟;基于贝叶斯原理的修正方法从径流概率分布出发,能够较好地反映其高度非线性特征。因此,本文选择时变权重组合与贝叶斯修正加以结合,提出一种中长期径流预报的新方法,方法流程如图1所示。

#### 1.1 时变权重组合

根据组合预报与单一模型预报结果之间的函数关系,组合预报可分为线性和非线性2类<sup>[6]</sup>。线性组合预报表达形式简单,应用最为广泛:

$$f_c = \alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_g f_g \tag{1}$$

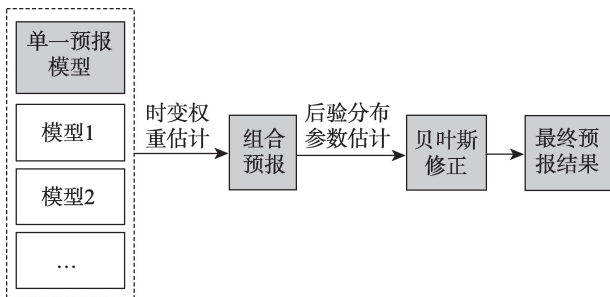


图1 基于时变权重组合和贝叶斯修正的中长期径流预报方法流程

Fig.1 Flowchart of mid- and long-term runoff forecast based on time-varying weight combination and Bayesian correction

式中: $f_c$ 为组合预报结果; $f_1, f_2, \dots, f_g$ 为 $g$ 个单一模型的预报值; $\alpha_1, \alpha_2, \dots, \alpha_g$ 为各单一模型权重,  $\alpha_1 + \alpha_2 + \dots + \alpha_g = 1, \alpha_i \geq 0, i = 1, 2, \dots, g$ 。

对于模型成员权重,考虑时变与否,可划分为恒定权重和时变权重2种方法。恒定权重方法是指在确定各单一模型权重后,保持权重不变进行组合预报。这种方法缺乏对单一模型预报精度时变特征的考虑,即根据平均效果赋予较优单一模型更高的权重,但在某些时段可能会出现较优模型预报误差明显增大的情形,此时仍采用恒定权重,则无法充分吸收其他模型在该时段的优势,导致具有较小误差预报信息的损失。与这种方法不同,时变权重方法可以根据各单一模型预报性能的变化而适时调整,从而实现预报信息的优化组合。常用的恒定权重方法通过一定的数学处理,均可实现动态变化。本研究以误差倒数平方方法为例,该方法认为各单一模型预报误差相互独立,权重与误差倒数平方成正比[式(2)]。考虑到与预报时段邻近时期的预报误差对于模型在局部时段预报能力的反映更加贴切,因此考虑各前序时段与预报时段的时间距离,采用非等权方法计算方差[式(3)],即可实现权重动态变化,从而对多种中长期径流预报模型的结果进行时变权重组合预报。

$$\alpha_{i,t} = (1/\sigma_{i,t-1}^2) / \sum_{j=1}^g (1/\sigma_{j,t-1}^2) \tag{2}$$

$$\sigma_{i,t-1}^2 = \sum_{s=1}^{t-1} s(f_{i,s} - o_s)^2 / \sum_{s=1}^{t-1} s \tag{3}$$

式中: $\alpha_{i,t}$ 为第 $i$ 个模型在 $t$ 时段的权重; $\sigma_{i,t-1}^2$ 为第 $i$ 个模型在 $1 \sim t-1$ 时段的预报方差; $s$ 为时段序号, $s = 1, 2, \dots, t-1$ ;  $f_{i,s}$ 为第 $i$ 个模型在 $s$ 时段的径流预报值; $o_s$ 为 $s$ 时段的径流实测值。

#### 1.2 贝叶斯修正

Krzysztofowicz<sup>[7]</sup>提出的贝叶斯预报处理器(Bayesian Process of Forecasts, BPF)是水文预报不确定性分析的有力工具。BPF突破了常规确定性水文模型在信息利用方面的局限性,有助于改善确定性预报结果的精度。令确定性径流预报结果为 $F$ ,实测流量为 $O$ ,2个变量的实现值分别用 $f, o$ 表示。在预报时刻 $k$ ,实测流量 $(o_1, o_2, \dots, o_{k-1})$ 与确定性预报结果 $(f_1, f_2, \dots, f_{k-1}, f_k)$ 已知, $o_k$ 未知,BPF利用已知信息对 $o_k$ 进行概率预报。实测流量的不确定性考虑为贝叶斯公式的先验分布,确定性预报模型结构与参数的不确定性转化为模型输出结果的不

确定性,并用似然函数体现。利用贝叶斯公式将先验分布与似然函数耦合得到  $o_k$  的后验概率密度函数:

$$\phi_k(o_k|o_1, o_2, \dots, o_{k-1}, f_1, \dots, f_k) = \frac{h_k(f_1, \dots, f_k|o_1, \dots, o_k)g_k(o_1, \dots, o_k)}{\int h_k(f_1, \dots, f_k|o_1, \dots, o_k)g_k(o_1, \dots, o_k)do_k} \quad (4)$$

式中:  $\phi_k$  为  $o_k$  的后验概率密度函数;  $g_k$  为先验概率密度函数; 对于确定性预报结果  $(f_1, f_2, \dots, f_k)$ ,  $h_k(f_1, \dots, f_k|o_1, \dots, o_k)$  为  $(o_1, o_2, \dots, o_k)$  的似然函数。

$g_k, h_k$  的计算难度较大, Krzysztofowicz 假设  $(o_1, o_2, \dots, o_k)$  为一阶马尔柯夫过程、 $(f_1, f_2, \dots, f_k)$  相互独立、 $f_i$  与  $o_j (i \neq j)$  相互独立,  $g_k, h_k$  均采用线性-正态模型简化计算, 经过严密的数学推导得到后验概率分布为正态分布, 推导过程见相关文献<sup>[14,18]</sup>。在实际预报中, 首先根据实测径流与同期确定性预报结果估计先验概率分布与似然函数的参数, 进而计算后验概率分布的均值和标准差, 最后取后验分布  $P=50\%$  的分位数作为 BPF 修正预报结果。

## 2 研究实例

### 2.1 研究区域与数据

云南龙江水库位于龙江流域干流下游河段(图2), 兼有防洪、发电、灌溉、旅游等综合效益, 为大(I)型水库, 具有年调节性能。坝址以上干流长度约 300 km, 河床平均坡降 5‰ 左右, 控制流域面积 5758 km<sup>2</sup>。降水时空分布不均, 由上游到下游逐步

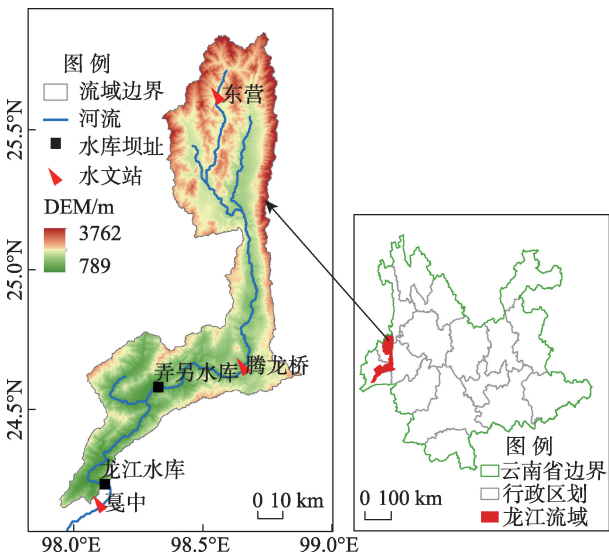


图2 云南龙江水库位置示意图

Fig.2 Location of the Longjiang Reservoir in Yunnan Province

减少, 时程上集中于 5—10 月。龙江水库上游建有弄另水库, 属大(II)型季节性调节水库, 其调度运行对龙江水库入库流量具有一定影响。

收集了上游腾龙桥水文站 1960—2010 年逐月流量, 采用水文比拟法并考虑弄另水库调节作用(逐月出库与入库流量的平均比例系数), 推算得到龙江水库相应时段入库径流, 将其与由水量平衡法反推的 2011—2018 年入库径流连接, 经一致性审查与还原计算, 得到 1960—2018 年径流系列, 以年(水文年: 1960—2017 年, 6 月至翌年 5 月)、月径流作为预报对象。收集了 1960—2018 年国家气候中心的 130 项逐月大气环流指数、NOAA 全球月平均海温格点数据(2°×2°)和 NCEP/NCAR Reanalysis 1 的 500 hPa 和 100 hPa 月平均气压位势格点数据(2.5°×2.5°), 作为中长期径流预报因子的备选集合。

### 2.2 单一模型预报效果

对于龙江水库年、月入库径流预报, 以环流指数、海温、气压和前期月径流为备选预报因子, 以 1961—2002 年为建模期(由于需要利用前期月径流, 起始年份推迟 1 a), 以 2003—2017 年为检验期, 预见期为 1 月, 选择随机森林(Random Forest, RF)和支持向量机(Support Vector Machine, SVM)构建 2 套“非滚动”径流预报模型, 即径流预报仅以前 1 年各预报因子作为输入。针对备选预报因子, 经相关系数计算、物理背景分析、显著相关因子挑选<sup>[19]</sup>, 选定基础预报因子集, 进而采用决策树对各因子重要性由高到低排序。针对不同预报因子规模(按照重要性排序逐步扩充预报因子组合)建立预报模型, 选用结合交叉验证的粒子群方法<sup>[20]</sup>, 优化确定最佳预报因子组合与模型参数。模型精度采用合格率(Qualification Rate, QR)和平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)评估<sup>[13]</sup>。

表 1 为 RF 与 SVM 模型对龙江水库年、月入库径流的预报精度。对于年入库径流, RF 与 SVM 模型的合格率均超过 90%, MAPE 均小于 12%; 建模期 RF 模型的定量误差低于 SVM, 而检验期则 SVM 更小; 由建模期到检验期, SVM 模型的 MAPE 增幅明显小于 RF, 即 SVM 模型的泛化能力较强。对于月入库径流, 综合两阶段精度指标, 8 月、9 月和翌年 2 月预报效果优于其他月份, 翌年 5 月预报精度明显较差。对比 2 种模型在不同阶段的表现, 在建模期, RF 模型的 QR 介于 71.4%~100% 之间, MAPE 介于 5.56%~19.24% 之间, 而各月 SVM 模型的 QR 不及 RF, 同时 MAPE 明显偏大, 最大 MAPE 接近

表1 基于RF与SVM模型的龙江水库年、月入库径流预报精度

Tab.1 Evaluation metrics of random forest (RF) and support vector machine (SVM) models for annual and monthly inflow forecast of the Longjiang Reservoir (%)

时段	建模期				检验期			
	RF模型		SVM模型		RF模型		SVM模型	
	QR	MAPE	QR	MAPE	QR	MAPE	QR	MAPE
年	100	3.83	97.6	6.77	93.3	11.45	93.3	9.39
6月	71.4	15.67	69.0	19.88	66.7	21.29	53.3	20.32
7月	95.2	8.25	83.3	10.81	60.0	22.22	60.0	18.78
8月	100.0	5.56	88.1	10.44	73.3	16.62	93.3	12.51
9月	95.2	7.58	83.3	10.23	80.0	13.02	73.3	16.63
10月	90.5	10.56	81.0	16.87	53.3	24.21	53.3	23.70
11月	90.5	8.75	69.0	11.59	53.3	32.99	40.0	28.66
12月	90.5	8.80	85.7	11.49	60.0	25.13	53.3	22.36
翌年1月	97.6	7.84	85.7	10.80	66.7	20.97	73.3	19.70
翌年2月	95.2	7.14	78.6	9.28	73.3	16.13	80.0	12.11
翌年3月	88.1	9.09	71.4	13.95	60.0	21.51	53.3	25.55
翌年4月	81.0	12.14	57.1	16.93	66.7	18.36	46.7	22.08
翌年5月	71.4	19.24	64.3	24.69	40.0	37.31	66.7	37.99

25%；在检验期，RF模型的预报QR总体上仍高于SVM，但6—8月、10月至翌年2月的定量误差大于SVM，其他月份则较小。由建模期到检验期，各月RF模型的MAPE增大幅度总体大于SVM，6—7月、10月至翌年1月、翌年3月和翌年5月RF模型检验期的MAPE均超过了20%。因此，年径流和部分月份径流的2种预报模型的泛化能力存在差异，建模期与检验期预报性能的优劣关系发生反转；同时部分月份定量误差相对突出，尚有较大的改善空间。

2.3 本文方法预报效果

在获取RF与SVM模型预报结果的基础上，应用本文方法预报龙江水库年、月入库径流。以年与6—8月入库径流为例，图3给出了2种模型的时变权重。由图可知，RF模型权重总体上大于SVM，但局部时段SVM却超过RF，并且SVM权重呈现先减小后增大的趋势，表明SVM模型在建模期初期预报误差趋于增大，到检验阶段又趋于减小，预报性能逐步接近RF。这充分说明了2种模型的预报性

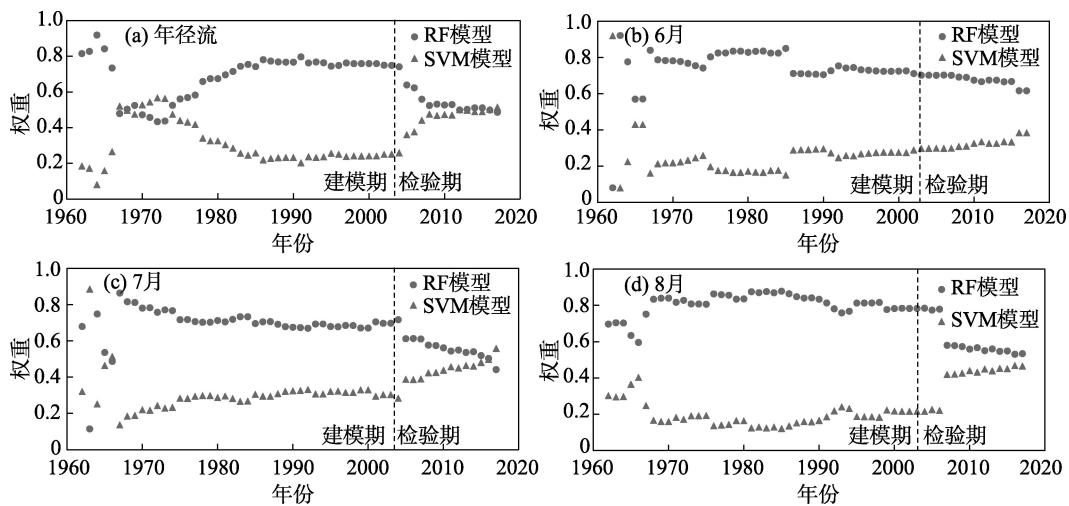


图3 龙江水库年入库径流与6—8月入库径流的RF与SVM模型的时变权重

Fig.3 Time-varying weight of random forest (RF) and support vector machine (SVM) models for annual runoff and monthly (June, July, and August) inflow forecast of the Longjiang Reservoir

能具有显著的非平稳特征,采用时变权重能够适时地根据邻近时段预报误差水平的变化,优化利用单一模型预报信息,从而得到组合预报结果。

基于时变权重组合预报结果,采用BPF修正方法建立后验分布,以50%分位数作为组合预报结果的修正值。图4对比了RF、SVM与本文方法在建模期与检验期的预报精度。对于年入库径流,时变权重组合与BPF修正后的QR与RF模型一致;从定量误差来看,组合后建模期与检验期的MAPE介于RF和SVM之间,平衡了二者在两阶段预报性能的差异,降低了不确定性;经BPF修正后,组合预报的定量误差进一步压缩,建模期与检验期的MAPE分别与各阶段最优单一模型持平,从而达到了两阶段较小误差预报信息的集成利用。对于月入库径流,从合格率来看,时变权重组合的QR总体上分布于RF和SVM之间;BPF修正后,建模期6月、11月、翌年2—5月的QR明显超过了单一模型和组合预报结果,其他月份与最优单一模型接近或一致,在检验期7月、10月的QR得到明显改善。定量误差方面,组合优化后多数月份建模期的MAPE接近最优单一模型,BPF修正则进一步使6—9月和翌年1—5月径流的MAPE低于最优单一模型;到检验期,各月组合预报结果的MAPE均低于最差单一模型,BPF修正使7月、10月至翌年1月、翌年3月和翌年5月的定量误差小于RF与SVM中的最优者,其他月份则与最优单一模型接近。另外,也可以看到由

组合优化到BPF修正,少数月份的预报精度并未得到改善,如6月检验期MAPE增大、翌年1月建模期QR降低等,这是由于实际径流与组合预报结果的统计关系与BPF模型假设存在一定出入而引起的。同样以年和6—8月入库径流为例,图5给出了不同方法的预报结果。可以看到,本文方法预报结果与实际径流具有较高的时程同步性,时变权重组合后多数时段预报值介于RF与SVM之间,BPF修正使预报结果更贴近实际值,局部极值处的预报效果得到明显改善。

综上,采用本文方法的时变权重组合能有效吸收RF和SVM模型在不同时段的较高精度预报信息,平衡了二者对于年和部分月份径流在建模期与检验期预报效果的差异,BPF修正改善了多数月份预报精度,部分峰谷径流的预报误差也得到了明显压缩。

以BPF修正值作为本文方法的最终预报结果,由图4知年径流预报的QR超过90%,MAPE低于10%;各月建模期MAPE全部小于15%,而到检验期,7—9月、翌年1—2月和翌年4月MAPE低于20%,其他月份的定量误差仍相对突出。基于径流预报结果,研判水文年型对于水资源规划管理和水利工程优化调度等具有重要指导意义。对龙江水库实际年入库径流序列进行水文频率分析(采用适线法估计P-III型分布参数),得到年入库径流的概率分布,计算采用本文方法预报年径流的理论概率

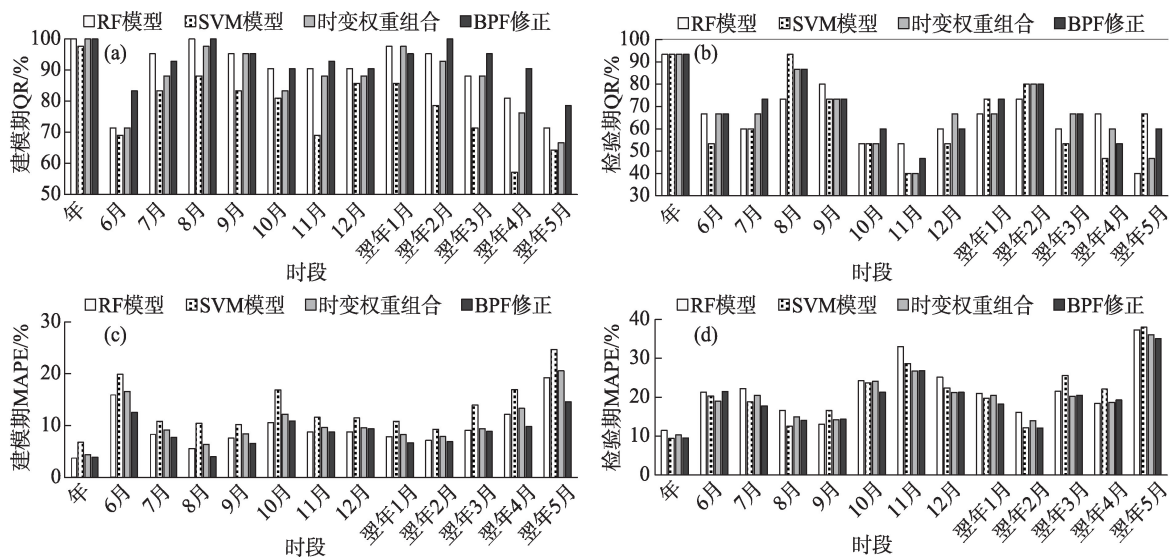


图4 本文方法与RF、SVM模型对龙江水库年、月入库径流预报精度的比较

Fig.4 Comparison of prediction performance between the proposed method and random forest (RF) and support vector machine (SVM) models for annual and monthly inflow of the Longjiang Reservoir

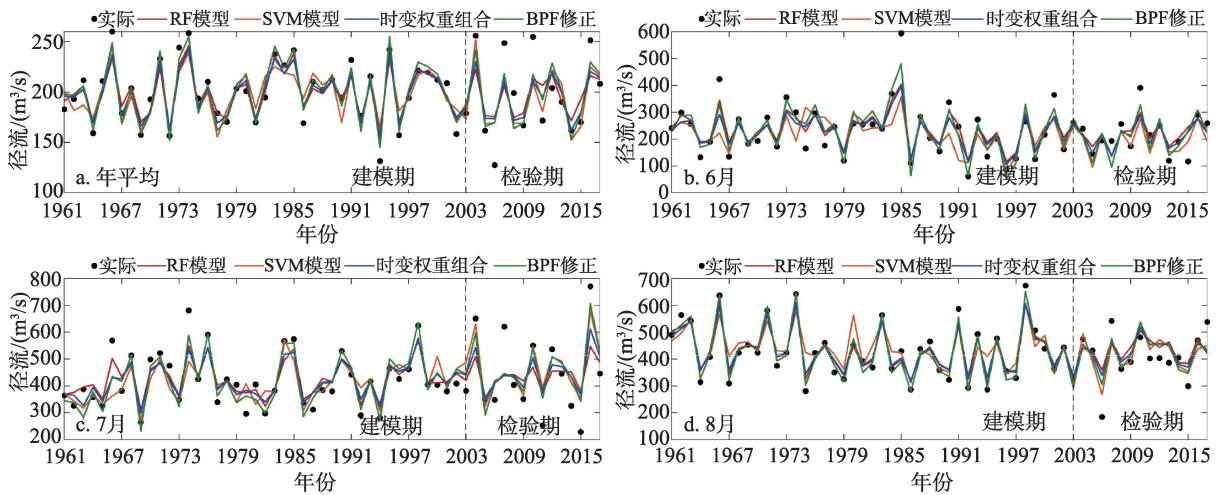


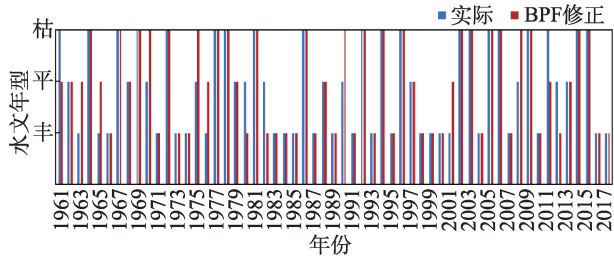
图5 本文方法与RF、SVM模型对龙江水库年径流及6—8月径流的预报结果

Fig.5 Annual and monthly (June, July, and August) inflow predictions of the Longjiang Reservoir based on the proposed approach and random forest (RF) and support vector machine (SVM) models

并确定水文年型,如图6所示。由该图知,1961—2017年大多数年份由预报年径流判断的水文年型与实际情况相同,正确率为77.2%。因此,根据本文方法的预报结果能够较准确研判下一年来水的总体丰枯情势。对于本文方法输出的“非滚动”月预报径流按时间排序,获得1961—2017年月预报径流序列,如图7所示。预报径流与实际径流的峰谷对应关系较好,建模期1961—2002年预报月径流的确性系数达到0.97,检验期2003—2017年为0.90。总体上,本文方法对于龙江水库年、月入库径流的预报效果较为理想。

### 3 结语

本文结合组合预报与预报误差修正2类中长期径流预报后处理技术的优势,考虑不同具体方法的



注:丰水年指超过累积频率(P)<37.5%,平水年指37.5%≤P<62.5%,枯水年指P≥62.5%。

图6 基于本文方法预报年径流的水文年型判断结果  
Fig.6 Hydrological year type classification based on predicted annual inflow by the proposed method

特性,提出了基于时变权重组合和BPF修正的中长期径流预报方法。龙江水库年、月入库径流预报的研究实例表明,时变权重组合可以有效平衡RF与SVM模型在建模期与检验期预报性能的差异,BPF修正后两阶段预报精度接近或优于最优单一模型。年径流的MAPE压缩到10%以下,各月径流建模期的MAPE全部低于15%,部分月份检验期定量误差的削减效益相对有限。根据1961—2017年年预报径流正确判断水文年型的比例为77.2%,月预报径流的确性系数超过0.90,总体预报精度较高。因此,本文方法融合了时变权重组合与BPF修正的优势,在吸收各模型不同时段较高精度信息、降低不确定性和提升预报精度等方面展现了积极效果,丰富了中长期径流预报方法。

但鉴于中长期径流预报问题的复杂性,本文方法的有效性及其可靠性还需更多的实例验证;另外,在BPF修正时主要应用了非滚动月预报径流的概

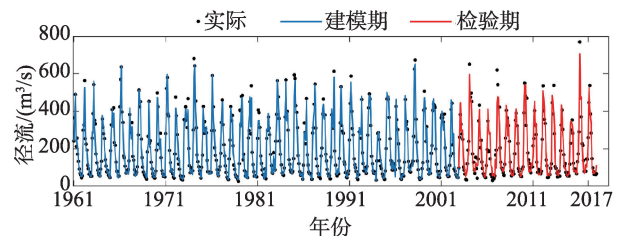


图7 基于本文方法预报的龙江水库1961—2017年逐月径流

Fig.7 Monthly inflow predictions of the Longjiang Reservoir using the approach proposed in this study, 1961—2017

率分布信息,在此基础上还可考虑分析月径流预报误差的时间序列特征,从而实现滚动修正,进一步改善预报效果。

### 参考文献(References)

- [1] 林沛榕, 张艳军, 洗翠玲, 等. 不同时间尺度的中长期水文预报研究 [J]. 水文, 2017, 37(6): 2-8. [Lin Peirong, Zhang Yanjun, Xian Cuiling, et al. Time-scale features of medium- and long-range hydrological forecasting. Journal of China Hydrology, 2017, 37(6): 2-8. ]
- [2] 刘勇. 基于物理成因的中长期水文预报方法与应用研究 [D]. 南京: 河海大学, 2011. [Liu Yong. Methods and application for mid-long term hydrological forecast based on physical cause. Nanjing, China: Hohai University, 2011. ]
- [3] 陈璐, 叶磊, 卢韦伟, 等. 基于 Copula 熵的神经网络径流预报模型预报因子选择 [J]. 水力发电学报, 2014, 33(6): 25-29. [Chen Lu, Ye Lei, Lu Weiwei, et al. Determination of input variables for artificial neural networks for flood forecasting using Copula entropy method. Journal of Hydroelectric Engineering, 2014, 33(6): 25-29. ]
- [4] 赵铜铁钢, 杨大文, 蔡喜明, 等. 基于随机森林模型的长江上游枯水期径流预报研究 [J]. 水力发电学报, 2012, 31(3): 18-24, 38. [Zhao Tongtiegang, Yang Dawen, Cai Ximing, et al. Predict seasonal low flows in the upper Yangtze River using random forests model. Journal of Hydroelectric Engineering, 2012, 31(3): 18-24, 38. ]
- [5] Yang T, Asanjan A A, Welles E, et al. Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information [J]. Water Resources Research, 2017, 53(4): 2786-2812.
- [6] Tan Q F, Lei X H, Wang X, et al. An adaptive middle and long-term runoff forecast model using EEMD-ANN hybrid approach [J]. Journal of Hydrology, 2018, 567: 767-780.
- [7] Le X H, Ho H V, Lee G, et al. Application of Long Short-Term Memory (LSTM) neural network for flood forecasting [J]. Water, 2019, 11(7): 1387. doi: 10.3390/w11071387.
- [8] Chowdhury S, Sharma A. Multisite seasonal forecast of arid river flows using a dynamic model combination approach [J]. Water Resources Research, 2009, 45(10): 2381-2386.
- [9] Zhang L, Yang X L. Applying a multi-model ensemble method for long-term runoff prediction under climate change scenarios for the Yellow River Basin, China [J]. Water, 2018, 10(3): 301. doi: 10.3390/w10030301.
- [10] Men B H, Long R S, Zhang J H. Combined forecasting of streamflow based on cross entropy [J]. Entropy, 2016, 18(9): 336. doi: 10.3390/e18090336.
- [11] 孙惠子. 中长期径流的多种组合预测方法及其比较 [D]. 杨凌: 西北农林科技大学, 2012. [Sun Huizi. A variety of combination forecasting method and comparing of medium and long-term stream-flow. Yangling, China: Northwest A & F University, 2012. ]
- [12] Wu J H, Qian H, Li P, et al. A system-theory-based model for monthly river runoff forecasting: Model calibration and optimization [J]. Journal of Hydrology and Hydromechanics, 2014, 62(1): 82-88.
- [13] Liang Z M, Li Y J, Hu Y M, et al. A data-driven SVR model for long-term runoff prediction and uncertainty analysis based on the Bayesian framework [J]. Theoretical and Applied Climatology, 2018, 133(1-2): 137-149.
- [14] 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. [Lu Fan, Jiang Yunzhong, Yin Junxian. Bayesian probabilistic low flow forecasting research based on BP ANN. Water Resources and Power, 2008, 26(3): 7-10. ]
- [15] 冯小冲. 水库中长期水文预报模型研究 [D]. 南京: 南京水利科学研究所, 2010. [Feng Xiaochong. Study on mid-long term hydrological forecasting model for reservoir. Nanjing, China: Nanjing Hydraulic Research Institute, 2010. ]
- [16] 陈华友. 组合预测方法有效性理论及其应用 [M]. 北京: 科学出版社, 2008: 40-41. [Chen Huayou. The validity theory of combined forecasting method and its application. Beijing, China: Science Press, 2008: 40-41. ]
- [17] Krzysztofowicz R. Bayesian models of forecasted time series 1 [J]. Journal of the American Water Resources Association, 1985, 21(5): 805-814.
- [18] 张洪刚. 贝叶斯概率水文预报系统及其应用研究 [D]. 武汉: 武汉大学, 2005. [Zhang Honggang. Probabilistic hydrologic forecasting system based on the Bayesian method. Wuhan, China: Wuhan University, 2005. ]
- [19] 李运刚, 何娇楠, 李雪. 基于 SPEI 和 SDI 指数的云南红河流域气象水文干旱演变分析 [J]. 地理科学进展, 2016, 35(6): 758-767. [Li Yungang, He Jiaonan, Li Xue. Hydrological and meteorological droughts in the Red River Basin of Yunnan Province based on SPEI and SDI indices. Progress in Geography, 2016, 35(6): 758-767. ]
- [20] 周婷, 金菊良, 李荣波, 等. 基于小波支持向量机的径流预测性能优化分析 [J]. 水力发电学报, 2017, 36(10): 45-55. [Zhou Ting, Jin Juliang, Li Rongbo, et al. Performance optimization analysis for inflow prediction using wavelet Support Vector Machine. Journal of Hydroelectric Engineering, 2017, 36(10): 45-55. ]

## Mid- and long-term runoff prediction based on time-varying weight combination and Bayesian correction

LI Lingjie<sup>1</sup>, WANG Yintang<sup>1</sup>, HU Qingfang<sup>1</sup>, LIU Yong<sup>1</sup>, LIU Dingzhong<sup>2</sup>, CUI Tingting<sup>1</sup>

(1. State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering,

Nanjing Hydraulic Research Institute, Nanjing 210029, China;

2. Yunnan Longjiang Water Conservancy Project Development Co., Ltd., Dehong 678400, Yunnan, China)

**Abstracts:** The mid- and long-term runoff prediction with satisfactory accuracy plays an important role as basic information in water resources planning & management and optimal operation of water conservancy projects. Combination and bias reduction are two common post-processing approaches in runoff forecast. Applying them in turn, considering the complicated non-stationary and nonlinear characteristics of runoff, a new mid- and long-term runoff prediction method by connecting time-varying weight combination and Bayesian correction is proposed. This method was used to study the annual and monthly inflow prediction of the Longjiang Reservoir in Yunnan Province. The results show that time-varying weight combination balances the performance difference of the established random forest (RF) and support vector machine (SVM) models in the modeling period and the test period. As a consequence of Bayesian correction, the prediction metrics are close to or better than the best of the predictions of the two individual stages. The proportion of correctly classified hydrological year type reaches 77.2% by employing the forecasted annual runoff, and the Nash-Sutcliffe efficiency coefficient of predicted monthly runoff series is over 0.90. Overall, the method put forward in this study has achieved positive effects in improving the forecast performance.

**Keywords:** mid- and long- term runoff prediction; time- varying weight combination; Bayesian correction; Longjiang Reservoir in Yunnan Province