

DOI:10.3876/j.issn.1000-1980.2020.06.006

基于多因子最近邻抽样回归模型的径流相似性预报

谭乔凤¹, 陈 然¹, 朱 阳², 胡立春², 闻 昕¹

(1. 河海大学水利水电学院, 江苏 南京 210098; 2. 国电大渡河流域水电开发有限公司, 四川 成都 610041)

摘要: 针对传统径流预报精度不高、预见期不足的问题, 提出基于降雨、径流相似性的径流预报方法, 采用大数据挖掘在历史降雨产流过程中搜索相似过程, 预测后期径流最可能的过程线。为了延长径流预报预见期, 实时接入降雨预报信息, 提出 3 种径流滚动预报方式, 实现了 7 d 预见期的径流逐日滚动预报; 针对流域在涨退水等不同阶段的产汇流特性, 建立可根据实时水雨情自适应切换的降雨、径流输入模式, 进一步提高径流预报精度。该研究成果在大渡河的应用表明预报效果达到预期: 3 d 预见期的纳什系数大于 0.9, 平均相对误差小于 10%; 7 d 预见期的纳什系数大于 0.8, 平均相对误差小于 15%。

关键词: 径流预报; 径流相似性; 多因子最近邻抽样回归模型; 预见期; 滚动预报

中图分类号: P338 **文献标志码:** A **文章编号:** 1000-1980(2020)06-0521-07

Runoff similarity forecast based on multi-factor nearest neighbor bootstrapping regressive model

TAN Qiaofeng¹, CHEN Ran¹, ZHU Yang², HU Lichun², WEN Xin¹

(1. College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China;
2. Dadu River Hydropower Development Co., Ltd., Chengdu 610041, China)

Abstract: Focusing on the low accuracy and insufficient foreseen period of traditional runoff forecast, this study proposed a runoff forecast method based on the similarity of rainfall and runoff. Data mining was used to search for the similar historical rainfall and runoff process, and the most likely runoff hydrograph in the later period was predicted. To prolong the runoff foreseen period to seven days, the real-time rainfall forecast information was inputted into the model and three rolling forecast schemes were proposed. The forecast models could be adaptively switched according to real-time rainfall conditions to further improve the forecast accuracy. The application in Dadu River showed that the Nash coefficients of forecasting the third day and seventh day were greater than 0.9 and 0.8, and the average relative errors were less than 10% and 15%, respectively. The research is of great significance to improve the forecast accuracy, extend the foreseen period, and promote the management and operation level of the reservoir group.

Key words: runoff forecast; runoff similarity; multi-factor nearest neighbor bootstrapping regressive model; foreseen period; rolling forecast

基金项目: 国家重点研发计划(2018YFC0407902)
作者简介: 谭乔凤(1991—), 女, 副教授, 博士, 主要从事水文预报和水库调度研究。E-mail: qiaofengtan@126.com
通信作者: 闻昕, 教授。E-mail: njwenxin@163.com
引用本文: 谭乔凤, 陈然, 朱阳, 等. 基于多因子最近邻抽样回归模型的径流相似性预报[J]. 河海大学学报(自然科学版), 2020, 48(6): 521-527.
TAN Qiaofeng, CHEN Ran, ZHU Yang, et al. Runoff similarity forecast based on multi-factor nearest neighbor bootstrapping regressive model[J]. Journal of Hohai University(Natural Sciences), 2020, 48(6): 521-527.

径流预报是防洪减灾、水资源保障、电力生产等流域管理和调度决策的关键依据,也是全球变化下水文水资源领域研究的前沿热点^[1-2]。目前广泛运用的径流预报方法主要分为过程驱动和数据驱动两类^[3]。前者从径流的物理成因出发,通过一系列含参数的数学物理方程描述产汇流过程,且每个参数都有明确的物理意义,但是当流域自然地理环境或产汇流条件复杂时,模型参数率定困难^[4-5];后者是以建立输入输出数据最优数学关系为目标的黑箱子方法^[6-7],无需考虑径流形成的物理机制,预报精度依赖于可靠且海量的样本数据^[8]。

近年来,由于人工智能和大数据技术的快速发展,流域内现代化的水文测站网络的建立与完善,以及历史水文气象数据的积累,为突破径流预报瓶颈开辟了新思路和新途径。借助大数据技术,可对成千上万条历史数据实现全面、多层次的分析,隐藏于数字背后的水文规律可以“挖掘”出来。实际上,对于某一特定流域,在一定的地理环境条件下,制约降雨的主导天气系统和降雨类型会反复出现,在相似天气系统条件下所产生的降雨(或暴雨)过程及其径流(或洪水)过程也将是相似的。当具有较长时间、较多场次的历史降雨、径流过程时,采用数据挖掘技术,找出相似降雨-径流过程,“参考过去预测未来”,即根据历史相似降雨、径流提供的信息预测未来径流,这对延长径流预报预见期,提高流域水库群精细化管理及精益化调度水平具有重要意义。

1 径流相似性预报

1.1 多因子最近邻抽样回归模型

研究将基于实时降雨、历史降雨和径流的相似性特征,利用多因子最近邻抽样回归模型(NNBR)在历史降雨产流过程中搜索相似过程,预测后期径流过程。NNBR无需对研究对象的相依形式和概率分布作假定,能避免确定函数关系所带来的不确定性问题,是一类数据驱动的非参数模型^[9]。其基本原理是:客观世界的发生发展存在一定联系,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。对于确定控制断面的径流预报,首先识别出与其相关性较高的前滞降雨、径流因子,建立降雨、径流特征指标矢量:

$$\mathbf{X}_j = (\mathbf{X}_{1,j}, \mathbf{X}_{2,j}, \dots, \mathbf{X}_{p,j}, \mathbf{X}_{p+1,j}, \mathbf{X}_{p+2,j}, \dots, \mathbf{X}_{p+q,j}) \tag{1}$$

式中: $\mathbf{X}_j$ ——预报因子; p ——降雨特征指标个数; q ——径流特征指标个数,共有 m 个特征指标, $m=p+q$ 。

已知当前特征矢量 $\mathbf{X}_0$,要预报其相应的值 y_0 ,NNBR 的基本思想是:在历史已有的特征矢量 $\mathbf{X}_j(j=1,2,\dots,N)$ 中,总有 K 个与当前特征矢量 $\mathbf{X}_0$ 最近邻(最相似),从而在 $\mathbf{X}_j$ 中可得到 K 个最近邻特征矢量,记为 $\mathbf{X}_j^{(1)}, \mathbf{X}_j^{(2)}, \dots, \mathbf{X}_j^{(K)}$,其对应的预报对象值分别为 $y_j^{(1)}, y_j^{(2)}, \dots, y_j^{(K)}$ 。最近邻可根据 $\mathbf{X}_j$ 与 $\mathbf{X}_0$ 间的欧氏距离、相关系数、灰色关联度等判断。以欧式距离 $r_j^{(i)}$ 判断为例,距离越小, $\mathbf{X}_j^{(i)}$ 与 $\mathbf{X}_0$ 越近邻。

考虑到降雨和径流在数值量级上存在差异,引入降雨影响权重 α ,提出降雨、径流综合欧氏距离的计算公式:

$$r_j^{(i)} = \alpha \left[\sum_{p_s=1}^p (\mathbf{X}_{p_s,j} - \mathbf{X}_{p_s,i})^2 \right]^{\frac{1}{2}} + \left[\sum_{q_s=1}^q (\mathbf{X}_{q_s,j} - \mathbf{X}_{q_s,i})^2 \right]^{\frac{1}{2}} \tag{2}$$

式中: $r_j^{(i)}$ ——实时样本与历史样本 j 的降雨、径流综合欧氏距离; α ——降雨影响权重; p_s ——降雨特征指标索引; q_s ——径流特征指标索引。

K 个最近邻特征矢量 $\mathbf{X}_j^{(1)}, \mathbf{X}_j^{(2)}, \dots, \mathbf{X}_j^{(K)}$ 与 $\mathbf{X}_0$ 间的欧氏距离记为 $r_{0j}^{(1)}, r_{0j}^{(2)}, \dots, r_{0j}^{(K)}$ 。 $r_{0j}^{(i)}$ 越小,说明 $\mathbf{X}_j^{(i)}$ 与 $\mathbf{X}_0$ 距离越近,则 $y_j^{(i)}=y_0$ 的可能性越大。用 $w_j^{(i)}$ 表示样本相似程度,称为抽样权重。距离越小表示越近邻, $w_j^{(i)}$ 与欧氏距离 $r_j^{(i)}$ 成反比。

根据上述思想,用式(3)表示多因子 NNBR 的基本形式:

$$y_0 = \sum_{i=1}^K w_j^{(i)} y_j^{(i)} \tag{3}$$

其中 $w_j^{(i)}$ 取值满足下述约束条件:

$$\sum_{i=1}^K w_j^{(i)} = 1 \tag{4}$$

模型涉及的参数有:最近邻数、特征矢量维数、抽样权重、降雨影响权重。

a. 最近邻数 K :模型根据已有的特征矢量,寻找历史中最近邻特征矢量的个数,也是模型的抽样个数。 K 的选取主要考虑两个因素:(a)降雨、径流的历史数据长度;(b)预报的径流量级在历史中出现的频率。首先根据历史数据长度预选一个最近邻数,当模型建立之后,测试不同 K 值下的预报效果,选取预报效果最好的值作为最终的 K 值。

b. 特征矢量维数:径流特征矢量维数由径流时间序列在不同滞时下的自相关性确定;降雨特征矢量维数需要结合降雨、径流物理成因,分析降雨与预报断面径流响应时间来确定,这个时间不仅与降雨量级有关,还与降雨中心的地理位置到预报断面的距离有关。

c. 抽样权重 $w_j^{(i)}$:只要满足式(4)且符合距离越近,抽样权重越大的原则即可。本研究采用待识别样本与最近邻样本的欧氏距离倒数的归一化数据作为抽样权重:

$$w_j^{(i)} = \frac{\frac{1}{r_j^{(i)}}}{\sum_{i=1}^K \frac{1}{r_j^{(i)}}} \tag{5}$$

d. 降雨影响权重:研究结合遗传算法等现代启发式算法,以预报效果最优为目标,优化降雨影响权重 α 。其中,预报效果评价指标主要包括纳什系数(NS)、均方根误差(RMSE)、平均绝对误差(MAE)和平均相对误差(MARE)^[10]。

1.2 滚动预报方式

为了延长预见期,实时接入滚动降雨预报信息,滚动预报未来径流过程,提出 3 种滚动预报方式。

a. 方式 1:不考虑降雨预报和径流预报信息,直接根据已经发生的实测降雨和径流信息寻找降雨、径流相似过程,以预报 $t \sim t+f$ 时段的径流过程为例。

$$\left. \begin{matrix} P_{t-p_T}, \cdots, P_{t-2}, P_{t-1} \\ Q_{t-q_T}, \cdots, Q_{t-2}, Q_{t-1} \end{matrix} \right\} \Rightarrow Q'_t, Q'_{t+1}, \cdots, Q'_{t+f} \tag{6}$$

式中: $P_{t-p_T}, \cdots, P_{t-2}, P_{t-1}$ ——已经发生的、实测的前期降雨过程; $Q_{t-q_T}, \cdots, Q_{t-2}, Q_{t-1}$ ——已经发生的、实测的前期径流过程; $Q'_t, Q'_{t+1}, \cdots, Q'_{t+f}$ ——预见期内逐时段的径流预报结果; p_T ——前期降雨滞时长度; q_T ——前期径流滞时长度; f ——预见期长度。不管是预报 t 时段的径流 Q'_t ,还是 $t+1$ 时段的径流 Q'_{t+1} ,都是依据相同的前期降雨 $P_{t-p_T}, \cdots, P_{t-2}, P_{t-1}$ 和前期径流 $Q_{t-q_T}, \cdots, Q_{t-2}, Q_{t-1}$ 在历史降雨、径流样本中寻找相似过程。

b. 方式 2:考虑降雨预报信息滚动更新降雨相似性,同时不断加入预报的径流作为输入滚动更新径流相似性以预报下一时段的径流。以预报 $t \sim t+f$ 时段的径流过程为例:

$$\left\{ \begin{matrix} P_{t-p_T}, \cdots, P_{t-2}, P_{t-1} \\ Q_{t-q_T}, \cdots, Q_{t-2}, Q_{t-1} \end{matrix} \right\} \Rightarrow Q'_t$$
$$\left\{ \begin{matrix} P_{t-p_T+1}, \cdots, P_{t-1}, P'_t \\ Q_{t-q_T+1}, \cdots, Q_{t-1}, Q'_t \end{matrix} \right\} \Rightarrow Q'_{t+1}$$
$$\vdots$$
$$\left\{ \begin{matrix} P_{t-p_T+f}, \cdots, P'_{t-2+f}, P'_{t-1+f} \\ Q_{t-q_T+f}, \cdots, Q'_{t-2+f}, Q'_{t-1+f} \end{matrix} \right\} \Rightarrow Q'_{t+f} \tag{7}$$

预报 t 时段的径流 Q'_t 与滚动方式 1 思路一致。预报 $t+1$ 时段的径流 Q'_{t+1} 时,需要将 t 时段的径流预报结果 Q'_t 以及 t 时段的预报降雨 P'_t 加进模型的输入中。

c. 方式 3:考虑降雨预报信息滚动更新降雨相似性,但是不将上一时段预报的径流作为输入,预测下一时段径流。

$$\left\{\begin{array}{l}P_{t-p_T}, \cdots, P_{t-2}, P_{t-1} \\ Q_{t-q_T}, \cdots, Q_{t-2}, Q_{t-1}\end{array}\right\} \Rightarrow Q'_t$$
$$\left\{\begin{array}{l}P_{t-p_T+1}, \cdots, P_{t-1}, P'_t \\ Q_{t-q_T}, \cdots, Q_{t-2}, Q_{t-1}\end{array}\right\} \Rightarrow Q'_{t+1}$$
$$\vdots$$
$$\left\{\begin{array}{l}P_{t-p_T+f}, \cdots, P'_{t-2+f}, P'_{t-1+f} \\ Q_{t-q_T}, \cdots, Q_{t-2}, Q_{t-1}\end{array}\right\} \Rightarrow Q'_{t+f}$$

(8)

滚动方式3与方式2的区别是:在预报 $t+1$ 时段的径流 Q'_{t+1} 时,将 t 时段的预报降雨 P'_t 输入模型,而径流输入不变。

1.3 多模式自适应切换

针对流域在涨退水等不同阶段的产汇流特性,建立可根据实时水雨情自适应切换的降雨、径流输入模式,以进一步提高径流预报精度。例如,根据考虑前期降雨滞时长度的不同,分为短降雨滞时模式和长降雨滞时模式。短降雨滞时模式只考虑短期内的实测降雨,例如只考虑前1d的降雨;长降雨滞时模式考虑更长时期内的实测降雨,例如考虑前3d之内的降雨。短降雨滞时模式以退水期预报为主,长降雨滞时模式以涨水期预报为主,模型将根据前期径流、前期降雨、预报降雨等因素自适应切换预报模式。另外,历史库中不同量级径流出现频率不同,对于大量级洪水过程,历史上出现的相似降雨和径流过程较少,模型将对径流进行实时追踪,当径流数值到达某一个量级时,自动减小最近邻数以保证抽样样本具有相似性。

2 应用实例

2.1 研究区域与数据

大渡河流域是国家规划的十三大水电基地之一,提高径流预报精度,延长预见期,对流域水库群精细化管理及精益化调度具有重要意义。本文以大渡河干流上游丹巴断面为研究对象。断面以上共有21个雨量遥测站点,3个流量站(绰斯甲、足木足、马尔康),站点分布情况见图1。为降低模型的输入维度,将预报断面以上的流域降雨站点分为5个子区域,并建立各分区点雨量到面雨量的空间映射关系,在体现流域降雨空间分布格局的前提下降低降雨数据输入维度。区域1包含色达、壤塘、一林场3个雨量站;区域2包含班玛、灯塔2个雨量站;区域3包含俄热、太阳河2个雨量站;区域4包含日部、草登、足木足、马尔康、刷经寺5个雨量站;区域5包含丹东、边尔、布科、东谷、沙冲、牦牛、抚边河、小金、达维9个雨量站。研究收集了各站点2008—2019年的逐日历史实测降雨、径流资料,并以2008—2016年的数据构建历史降雨、径流样本,对2017—2019年的径流进行相似性预测,以评价模型未来1d和未来7d逐日的预测效果。

2.2 未来1d预报

2.2.1 预报方案设置

预见期为1d的预报因子分为3类:上游雨量(点或面雨量)、上游径流、丹巴本站径流。根据相关性分析,上游前1d雨量与丹巴径流相关性最高,因此可以初步取降雨滞时为1d。丹巴径流自相关系数在前5d分别为:0.96、0.87、0.79、0.69、0.61。前4d滞时下的相关性系数明显下降,因此丹巴本站前期径流滞时取3d。另外,研究测试了预报结果为绝对值和相对值两种输出模式。绝对值模式根据相似性样本下一时段径流的绝对

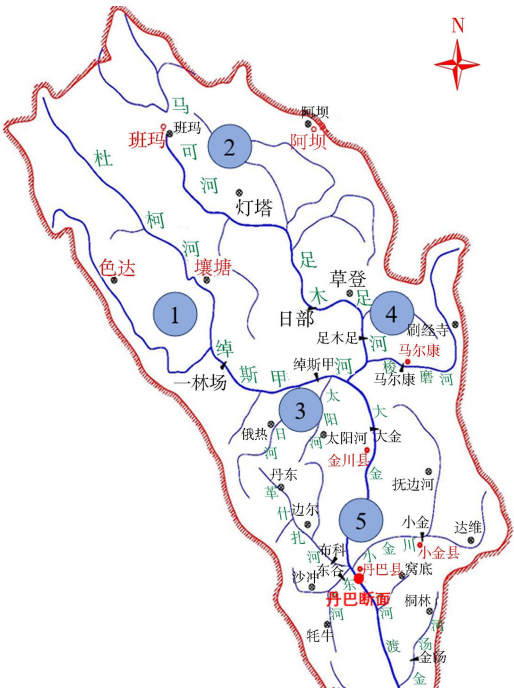


图1 大渡河流域丹巴以上水文站点分布
Fig.1 Distribution of hydrological stations above Danba in Dadu River Basin

值预报实时径流;相对值模式根据相似性样本下一时段径流相对于前一时段径流的变化量预报实时径流。

根据不同的输入因子和输出模式,研究设置了多种预报方案,如表 1 所示。其中,方案 1~4 采用绝对值模式,方案 1 与方案 2 是为了对比点雨量和面雨量输入对预报效果的影响。方案 3 与方案 2 是为了对比上游支流站点径流对预报效果的影响。方案 4 的预报因子只考虑丹巴本站前期径流,可测试只考虑径流自身时间序列的预报效果。方案 5~10 采用相对值模式,6 种不同滞时组合下预报方案形成对比。值得注意的是,方案 5 与方案 2 的预报因子输入完全相同,但是方案 2 采用绝对值模式,而方案 5 采用相对值模式,因此这两种方案可直接对比 2 种输出模式的预报效果。

表 1 未来 1 d 的预报方案

Table 1 Forecast schemes with foreseen period of 1 day

方案	预报因子输入	输出模式
1	上游 5 个站点雨量($t-1$ d) + 上游 3 个站点径流($t-1$ d) + 丹巴本站前期径流($t-1$ d、 $t-2$ d、 $t-3$ d)	绝对值
2	上游 5 个区域面雨量($t-1$ d) + 上游 3 个站点径流($t-1$ d) + 丹巴本站前期径流($t-1$ d、 $t-2$ d、 $t-3$ d)	绝对值
3	上游 5 个区域面雨量($t-1$ d) + 丹巴本站前期径流($t-1$ d、 $t-2$ d、 $t-3$ d)	绝对值
4	丹巴本站前期径流($t-1$ d、 $t-2$ d、 $t-3$ d)	绝对值
5	上游 5 个区域面雨量($t-3$ d) + 上游 3 个站点径流($t-1$ d) + 丹巴本站前期径流($t-1$ d、 $t-2$ d、 $t-3$ d)	相对值
6	上游 5 个区域面雨量($t-2$ d) + 上游 3 个站点径流($t-1$ d) + 丹巴本站前期径流($t-1$ d、 $t-2$ d、 $t-3$ d)	相对值
7	上游 5 个区域面雨量($t-1$ d) + 上游 3 个站点径流($t-1$ d) + 丹巴本站前期径流($t-1$ d、 $t-2$ d、 $t-3$ d)	相对值
8	上游 5 个区域面雨量($t-3$ d、 $t-2$ d、 $t-1$ d) + 上游 3 个站点径流($t-1$ d) + 丹巴本站前期径流($t-1$ d、 $t-2$ d、 $t-3$ d)	相对值
9	上游 5 个区域面雨量($t-2$ d、 $t-1$ d) + 上游 3 个站点径流($t-1$ d) + 丹巴本站前期径流($t-1$ d、 $t-2$ d、 $t-3$ d)	相对值
10	上游 5 个区域面雨量($t-3$ d、 $t-2$ d) + 上游 3 个站点径流($t-1$ d) + 丹巴本站前期径流($t-1$ d、 $t-2$ d、 $t-3$ d)	相对值

2.2.2 预报结果分析

方案 1~10 的预报精度评价结果见表 2。从表 2 可知,采用面雨量输入的方案 2 各项指标均优于采用点雨量输入的方案 1。采用相对值输出的方案 5~10 的预报效果整体优于采用绝对值输出的方案 1~4,因此预报模型输入取面雨量、输出取相对值更适合本案例。方案 2 与方案 3 的精度对比可知,是否考虑前 1 d 上游 3 站的径流对丹巴预报效果影响不大。与方案 3 相比,方案 4 的预报精度明显降低,说明前期降雨对丹巴径流影响显著。方案 5~10 共 6 种预报方案的预报效果整体接近,在 NS 等于 0.96 的方案 7~9 中,方案 7 的 RMSE 虽然比方案 8、方案 9 略高,但是 MARE 最低,因此确定方案 7 是预见期 1 d 的最优方案。

从 2017—2019 年 3 a 的预报结果来看,当预见期为 1 d 时,模型的整体预报精度较高,但是在个别洪峰处的预报效果较差。以 2018 汛期(5—10 月)为例(图 2),在年最大洪峰处的预报值明显低于实际值,可能的原因有两个:一是 NNBR 的预报结果是历史相似样本的加权平均值,容易出现极大值预报偏小和极小值预报偏大的情况;二是历史相似的大量级洪水场次较少,导致模型抽样得到的样本相似性较弱。

表 2 预见期为 1 d 的预报精度评价

Table 2 Forecast accuracy with foreseen period of 1 day

方案	NS	RMSE/ ($\text{m}^3 \cdot \text{s}^{-1}$)	MAE/ ($\text{m}^3 \cdot \text{s}^{-1}$)	MARE/%
1	0.90	193.49	106.11	6.31
2	0.91	190.23	105.68	6.29
3	0.91	192.27	110.42	6.64
4	0.87	204.56	135.74	7.13
5	0.95	158.87	97.32	4.28
6	0.95	154.67	92.70	5.75
7	0.96	130.39	69.22	4.27
8	0.96	129.60	71.60	5.39
9	0.96	127.56	69.42	5.28
10	0.95	155.40	92.80	5.73

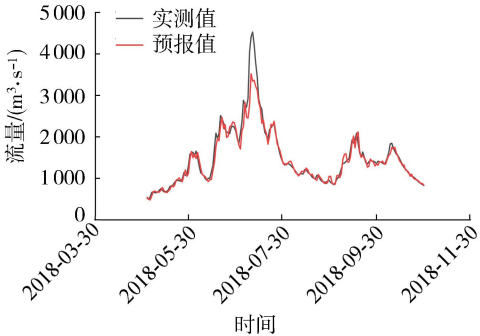


图 2 预见期为 1 d 的预报结果

Fig. 2 Forecast results with 1-day forecast period

2.3 未来 1 周预报

2.3.1 预报方案设置

为了延长径流预报预见期,实时接入 7 d 降雨预报信息,预报流域未来 1 周的径流过程。由于缺失历史降雨预报资料,用实测降雨代替降雨预报信息。在预见期为 1 d 的预报方案基础上,提出 3 个滚动预报方案。3 种方案的预报因子均是上游 5 个区域前 1 d 面降雨和丹巴本站前 3 d、前 2 d、前 1 d 径流,并且均采用相对值

输出模式。3种方案分别对应章节1.2所介绍的3种滚动预报方式,具体方案如表3所示。

2.3.2 预报结果分析

滚动预报方案1~3的预报精度见表4。滚动方案2在7d内预报效果最好,滚动方案1与滚动方案3的预报效果很接近。且随着预见期的增加,预报精度显著下降。图3显示了滚动方案2对2019年汛期典型洪水的预报过程图,黑色线为实测径流过程,每一个日期图例表示以该日期为预报起点的未来7d的径流预报过程,预报结果逐日滚动更新。典型径流过程的预报结果与实测值较为接近,整体预报精度较高。

表4 预见期为7d的预报精度评价

Table 4 Forecast accuracy with foreseen period of 7 days

预见期/d	方案1				方案2				方案3			
	NS	RMSE/ ($\text{m}^3 \cdot \text{s}^{-1}$)	MAE/ ($\text{m}^3 \cdot \text{s}^{-1}$)	MARE/%	NS	RMSE/ ($\text{m}^3 \cdot \text{s}^{-1}$)	MAE/ ($\text{m}^3 \cdot \text{s}^{-1}$)	MARE/%	NS	RMSE/ ($\text{m}^3 \cdot \text{s}^{-1}$)	MAE/ ($\text{m}^3 \cdot \text{s}^{-1}$)	MARE/%
1	0.96	130.46	69.34	4.29	0.96	130.46	69.34	4.29	0.96	130.46	69.34	4.29
2	0.86	258.80	153.01	9.08	0.90	217.55	117.31	7.00	0.87	250.15	147.15	8.83
3	0.74	354.99	226.33	13.53	0.84	279.05	154.64	9.10	0.74	352.90	223.80	13.36
4	0.62	429.33	285.13	16.96	0.78	324.38	183.55	10.68	0.60	436.68	289.19	17.19
5	0.49	496.65	337.29	19.96	0.73	358.21	208.95	11.98	0.46	507.00	343.27	20.39
6	0.35	554.46	376.90	22.38	0.69	384.69	231.76	13.19	0.33	566.80	384.55	22.88
7	0.22	609.47	412.68	24.59	0.66	404.28	248.88	14.21	0.21	615.08	416.50	24.83

2.4 多模式自适应预报

在进行未来1周径流预测时发现当前1d的降雨 P_{t-1} 延长至前3d降雨 $P_{t-3}, P_{t-2}, P_{t-1}$ 时,能够更好地预测大渡河丹巴断面的汛期涨水过程。通过分析 P_{t-1} 和 $P_{t-3}, P_{t-2}, P_{t-1}$ 两种输入的滚动预测效果,研究确定了模型自适应判断条件,当满足判断条件时,模型将从1d模式切换至3d模式;否则,默认采用1d模式进行滚动预报。同时,研究发现,当前期径流量级超过 $3500\text{ m}^3/\text{s}$ 时,应适当减少最近邻数,推荐取4。

采用多模式自适应切换对2017—2019年的滚动预报效果见表5。总体精度相比于原滚动方案2有显著提升。预见期为3d的NS大于0.9, MARE小于10%;预见期为7d的NS大于0.8, 差MARE小于15%。与原滚动方案2相比,在径流发展的不同阶段自适应预报效果均有提升,其中涨水阶段的预报效果提升最为明显(图4)。自适应预报带来的精度提升主

表5 自适应模式的预报精度评价

Table 5 Forecast accuracy of adaptive mode

预见期/d	NS	RMSE/ ($\text{m}^3 \cdot \text{s}^{-1}$)	MAE/ ($\text{m}^3 \cdot \text{s}^{-1}$)	MARE/%
1	0.98	102.57	58.23	3.73
2	0.95	159.02	92.35	5.94
3	0.92	197.78	118.60	7.58
4	0.89	228.45	142.19	8.95
5	0.87	252.77	162.15	10.04
6	0.84	278.23	178.92	10.94
7	0.81	298.31	191.63	11.60

表3 预见期为7d的滚动预报方案

Table 3 Forecast schemes with foreseen period of 7 days

滚动方案	预报因子	滚动预报方式	输出模式
1	上游5个区域前1d面雨量 丹巴本站前3d、前2d、前1d径流	方式1	相对值
2	上游5个区域前1d面雨量 丹巴本站前3d、前2d、前1d径流	方式2	相对值
3	上游5个区域前1d面雨量 丹巴本站前3d、前2d、前1d径流	方式3	相对值

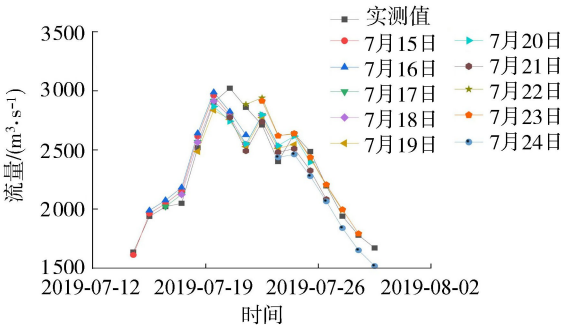


图3 典型洪水过程的滚动预报结果

Fig. 3 Rolling forecast results of typical flood

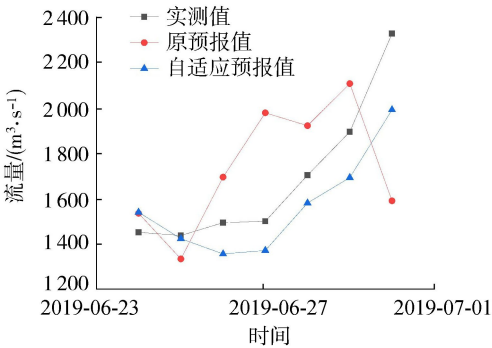


图4 典型径流过程的自适应预报结果

Fig. 4 Adaptive forecast results of typical runoff

要源于最近邻样本相似性增强和累计误差的减小。通过判断当前降雨、径流状态,选择不同的预报模式寻找相似性样本,可以有效提高预报精度,减小累计误差。

3 结 语

基于多因子最近邻抽样回归模型开展径流相似性预报,预报效果整体较好。同时,针对流域在涨退水等不同阶段的产汇流特性,建立可根据实时水雨情自适应切换的降雨、径流输入模式,进一步提高了径流预报精度。在大渡河丹巴断面的应用效果表明预报效果达到预期,3 d 预见期的 NS 大于 0.9,MARE 小于 10%;7 d 预见期的 NS 大于 0.8,MARE 小于 15%。未来可以对洪峰处的径流过程做进一步深入研究,优化不同应用环境下的模型参数,优选特定径流过程的相似性样本,全面提升径流预报精度。

参考文献:

[1] LIU Saiyan, HUANG Shengzhi, XIE Yangyang, et al. Identification of the non-stationarity of floods: changing patterns, causes, and implications[J]. Water Resources Management, 2019, 33(3):939-953.

[2] MAGGIONI V, MASSARI C. On the performance of satellite precipitation products in riverine flood modeling: a review[J]. Journal of Hydrology, 2018,558:214-224.

[3] 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. (WANG Wen, MA Jun. Overview of several hydrological forecast methods[J]. Advances in Science and Technology of Water resources,2005,25(1):56-60. (in Chinese))

[4] 刘金涛,宋慧卿,张行南,等. 新安江模型理论研究的进展与探讨[J]. 水文, 2014, 34(1): 1-6. (LIU Jintao, SONG Huiqing, ZHANG Xingnan, et al. Progress and discussion on the research of Xin'anjiang model theory[J]. Hydrology, 2014, 34(1):1-6. (in Chinese))

[5] 李晓,李致家,董佳瑞. SWAT 模型在伊河上游径流模拟中的应用[J]. 河海大学学报(自然科学版), 2009,37(1):23-26. (LI Xiao, LI Zhijia, DONG Jiarui. Application of SWAT model in runoff simulation of upper Yihe River[J]. Journal of Hohai University (Natural Science), 2009,37(1):23-26. (in Chinese))

[6] 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. (LIN Jianyi, CHENG Chuntian. Application of support vector machine in medium and long term runoff forecast[J]. Journal of Hydraulic Engineering, 2006,37(6):681-686. (in Chinese))

[7] 屈亚玲,周建中,刘芳,等. 基于改进的 Elman 神经网络的中长期径流预报[J]. 水文,2006,26(1): 45-50. (QU Yaling, ZHOU Jianzhong, LIU Fang, et al. Medium and long-term runoff forecasting based on improved Elman neural network[J]. Hydrology,2006,26(1):45-50. (in Chinese))

[8] YASEEN Z M, SULAIMAN S O, DEO R C, et al. An enhanced extreme learning machine model for river flow forecasting: state-of-the-art, practical applications in water resource engineering area and future research direction[J]. Journal of Hydrology, 2019,569:387-408.

[9] 王文圣,袁鹏,丁晶. 最近邻抽样回归模型在水环境预测中的应用[J]. 中国环境科学, 2001, 21(4):367-370. (WANG Wensheng, YUAN Peng, DING Jing. Application of nearest neighbor bootstrap regressive model in prediction of water environment[J]. China Environment Science, 2001,21(4):367-370. (in Chinese))

[10] 谭乔凤,王旭,王浩,等. ANN、ANFIS 和 AR 模型在日径流时间序列预测中的应用比较[J]. 南水北调与水利科技, 2016, 14(6):12-17. (TAN Qiaofeng, WANG Xu, WANG Hao, et al. Comparative study of ANN, ANFIS and AR model for daily runoff time series prediction[J]. South-to-North Water Transfers and Water Science & Technology, 2016,14(6):12-17. (in Chinese))

(收稿日期:2020-10-30 编辑:张志琴)