

陈剑飞,李勇,刘俊江,等.耦合陆面水文模型和机器学习方法的岩滩水库径流量预报及应用[J].气象研究与应用,2022,43(1):01-07.
Chen Jianfei, Li Yong, Liu Junjiang, et al. Yantan Reservoir runoff forecast and application coupled with land surface hydrological model and machine learning method[J]. Journal of Meteorological Research and Application, 2022, 43(1): 01-07.

耦合陆面水文模型和机器学习方法的 水库径流量预报及应用

陈剑飞¹, 李 勇¹, 刘俊江², 钟利华^{1*}, 史彩霞¹, 袁 星^{2,3}, 钟华昌⁴

(1.广西壮族自治区气象灾害防御技术中心, 南宁 530022; 2.南京信息工程大学水文与水资源工程学院, 南京 210044;
3.中国科学院东亚区域气候-环境重点实验室, 北京 100029; 4.广西桂冠电力股份有限公司, 南宁 530029)

摘要:采用智能网格预报等多种气象预报数据与陆面水文模型及机器学习方法进行耦合,以岩滩水库流域区间为例,将区间预报径流量与历史径流量数据输入长短期记忆网络模型(LSTM)进行水库入库径流量预报,通过对日径流量模拟试验和业务试用分析,探明该方法在短期水文预报中的适用性。结果表明,耦合气象-水文-机器学习的径流量预报模型在率定期和验证期纳什效率系数(NSE)在0.65左右,在强降水过程训练试验中,日径流量和洪峰预报合格率 $\geq 87.5\%$,达到甲级预报精度标准;在业务试用中,24h、48h、72h日径流量预报合格率分别为87.3%、70.4%、75.5%,达到甲级或乙级预报精度标准,满足发布正式预报的精度要求;3次较大降水过程径流量和峰现时间预报合格率均为100%,达到甲级预报精度标准;峰值预报为66.7%,达到丙级预报标准,可用于参考性预报。将物理模型与人工智能方法进行有机耦合,可提高水文预报产品精度和适用性。

关键词:耦合;陆面水文模型;机器学习;径流量预报;物理模型;人工智能

中图分类号: P338 **文献标识码:** A **doi:** 10.19849/j.cnki.CN45-1356/P.2022.1.01

引言

水电作为清洁能源,其可靠供应是社会经济发展的重要支撑。广西是我国水利发电的大省(区)之一,利用好水资源、增加水力发电电量是水电经济调度的重要工作。岩滩水电站作为红水河流域广西电力部门直接调度的第一个季调节水电站,其库区精准的径流量监测预报信息对流域水库群联合优化调度起到重要的作用,有利于提高水库群发电效益和供水安全。以往岩滩水库径流量预报方法主要是传统的统计预报方法^[1-2],对流域范围内下垫面山坡河网等微地形影响研究不多,难以描述细致的地表、地下产汇流过程和水文要素的时空异质性,而单纯的物理模型或经验模型也难以准确地预报水库径流量,需要人工智能模型加以融合分析。在人工智能

模型方面,当前比较流行的模型是长短期记忆网络模型(LSTM),它不仅能够发现输入变量和输出变量之间的隐藏关系,也能够更好地识别并记忆具有长期相关性的信息。

对于耦合气象预报、物理水文模型或利用机器学习方法提高水库径流量预报准确率问题的探讨,许多学者从多个方面开展了研究。包红军等^[3]开展了洪水预报中定量降水预报应用进展研究,认为融合预报员预报的格点化定量降水预报技术,是提高面向洪水预报中流域降水预报精度的重要方法;崔春光等^[4]开展了定量降水预报与水文预报耦合的预报试验,发现考虑预见期内的降雨相对于未考虑预见期降雨对洪水预报准确性提高具有明显的优势;曾鹏等^[5]、彭筱等^[6]、卢小凤等^[7]、史彩霞等^[8]通过开展智能网格、中国气象局广东快速更新同化数值预

收稿日期: 2022-01-28

基金项目: 国家自然科学基金项目(41861124005)、国家重点研发计划“全球变化及应对”专项课题(2018YFA0606002)

作者简介: 陈剑飞(1975—),男,高工,主要从事专业气象技术服务及管理工作。E-mail: 497933494@qq.com

* 通讯作者: 钟利华(1962—),女,正研级高工,主要从事专业气象预报服务工作。E-mail: 1427962612@qq.com

报系统(CMA-GD)预报产品的降尺度研究和应用检验分析,以期提升模式预报产品的实际预报能力;郭玉雪等^[9]以舟山岛水库群为例,基于三种神经网络模型建立了不同预报因子组合和预见期的径流量预报模型,发现 LSTM 通过耦合降雨预报信息可提高径流量预报精度;殷兆凯等^[10]基于 LSTM 建立不同预见期流域径流量预报模型的结果显示,预见期在 0~3d 的径流量预报效果 LSTM 模型优于新安江模型;梁振清等^[11]基于卷积长短期记忆神经网络,构建深度学习的人工智能降水短临预报系统,显示出在 1h 内预测准确率保持在 80%以上。这些研究对耦合后的径流量预报精度进行了不同方面的论述,为后续改进提供了有益的参考作用。

以红水河流域岩滩水库范围作为研究区域,Liu 等^[12]在传统水文气象预报模型中引入机器学习方法,使用全球交互式大集合中心的欧洲中心回报数据(TIGGE-ECMWF),驱动 CSSPv2 (Conjunctive Surface-Subsurface Process model version 2)陆面水文模型进行了气象-水文回报试验后,将结果输入 LSTM 模型中进行校正,结果表明,耦合气象-水文-机器学习方法,与气象-水文方法建模比较,径流量预报与实况相比能够在前 72h 内提升 6%的精度。本文将采用多套短期定量降水预报模式优选结果,驱动陆面水文模型计算区间径流量,并将实况径流量输入 LSTM 模型中,完成对岩滩水库入库径流量的预报。通过对历史洪水过程及业务运行期间的径流量预报进行评估,系统性分析耦合气象-水文-机器学习方法的岩滩水库径流量预报能力,为水库运行优化调度决策提供参考。

1 资料和研究区域

1.1 资料来源

气象资料采用岩滩水电站流域内 97 个自动气象站 2013 年 1 月 1 日至 2020 年 12 月 31 日逐时降水、气温、相对湿度、风速、气压等实况资料,全国智能网格、欧洲中期天气预报中心数值模式(ECMWF)和 CMA-GD 模式预报产品,以及中国气象局陆面数据同化系统同期逐时气压、比湿、风速、降水、短波太阳辐射等数据。水文资料采用岩滩、甲篆、罗富、龙滩 4 个水库同期逐日平均径流量数据,岩滩、龙滩 2 个水库逐时径流量数据。

1.2 研究区域

岩滩水电站位于广西大化县,流域内含岩滩、甲

篆、罗富 3 水文观测站。利用 GIS 技术,通过提取岩滩水库流域范围水系、沿水系分水岭、主河道、流向、汇流量数据,划分为甲篆库区、罗富库区、岩滩上库区、岩滩中库区、岩滩下库区 5 个流域区间,图 1 所示为岩滩水库 5 个流域区间及气象站点和水文站点分布图。

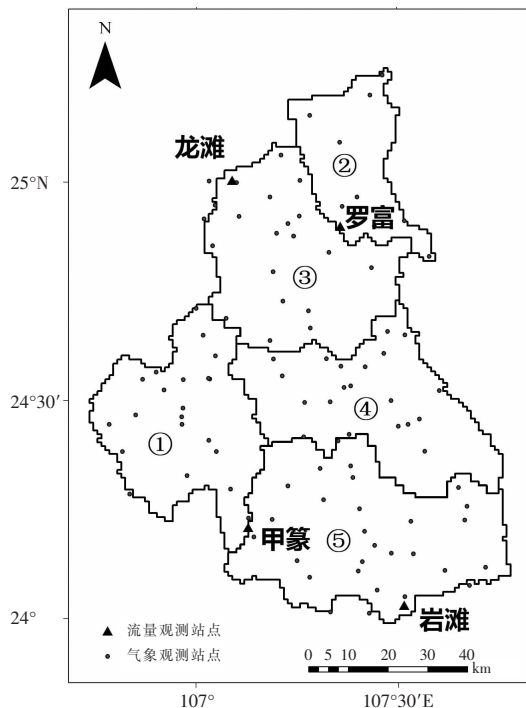


图 1 岩滩水库 5 个流域区间及气象站点和水文站点分布图

①甲篆库区;②罗富库区;③岩滩上库区;
④岩滩中库区;⑤岩滩下库区

2 研究方法

2.1 CSSPv2 模型和 LSTM 机器学习模型

CSSPv2 模型是在 Common Land Model 模型^[13-14]基础上发展起来的一种分布式、基于网格的陆面水文模型,具有完善的陆面水文过程描述,适合高分辨率模拟^[15-16]。CSSPv2 模型分为模型前处理、模拟模块和后处理三个部分。首先通过生成气象强迫数据的插值权重以及根据权重进行气象强迫数据插值,并由陆面模式输入模型进行模拟,最后对陆面模式输出的网络通用数据格式(NetCDF)进行后期处理,根据经纬度将格点和站点进行匹配,找到站点对应的格点位置,选取需要的输出要素进行提取,完成径流量预报。

长短期记忆网络模型(LSTM)记忆单元输入包括 $t-1$ 时刻的隐藏层状态变量 c_{t-1} 、记忆单元状态变量 h_{t-1} 以及当前输入变量 x_t ,输出当前时刻隐藏层状

态变量 c_t 与输出变量 h_t , 其运算主要通过遗忘门、输入门和输出门三个步骤完成。

文中所使用的 LSTM 是由 Keras 深度学习框架所构建, 通过输入历史 24h 实况入库径流量、上游出库径流量及 CSSPv2 预报的 72h 区间径流量进行训练, 得到 72h 长度的入库径流量预报。

2.2 CSSPv2 模型参数率定

通过对产流参数和汇流参数进行率定实现物理模型的本地化应用, 其中模型产流模块采用可变入渗容量产流 (VIC) 方案, 是当前水文模型中运用较广的基于土壤水含量的土壤产流参数化方案。其采用地表入渗曲线来表示网格内部地表入渗能力的分布, 对于地下基流则借助基流曲线计算。产流参数率定使用 SCE-UA 进化算法^[17], 以 NSE 作为目标函数对地表入渗曲线和基流曲线形状参数进行最大 2000 代率定。模型汇流模块则使用隐式求解差分一维运动波方程^[18-19], 适合目标区域的多山地形。河道汇流受格点河道密度、河道坡度、河道宽度与河道深度以及河道糙率参数影响, 其初值可使用精细 DEM 进行提取; 参数率定使用人工试错法, 以多个代表性强降水过程为目标进行率定。

经产汇流参数率定后, 日尺度的 CSSPv2 模式模拟流量与实况流量差值较小, 而小时流量模拟对于上游的瞬时出库响应较差, 波形坦化较重, 需使用人工智能方法加以校正。

2.3 模式降水预报产品优选方案

选取智能网格、ECMWF 模式和 CMA-GD 模式降水预报产品, 经过双线性插值得到岩滩流域降水量预报产品, 同时使用自动站实况降水进行距离反比权重插值制作岩滩流域实况降水量产品, 采用算术平均法计算岩滩水库甲篆库区、罗富库区、岩滩上库区、岩滩中库区、岩滩下库区 5 个流域区间面雨量。使用绝对误差检验方法, 即: 绝对误差 = 预报面雨量 - 实况面雨量, 对 3 种不同降水预报产品进行优选。采用动态评定方法, 按照智能网格、ECMWF 模式和 CMA-GD 模式降水预报前 15 日的平均绝对误差评分的大小进行排名, 采用排名第 1 的降水预报作为最优预报, 代入预报模型运行, 形成径流量预报产品; 若评分最优预报缺测, 将排名第 2 的代入, 进行依次递补。

2.4 预报模型建立及模型效果

2.4.1 预报模型建立

预报模型建立包括 3 个步骤, 即: 驱动数据制

作、物理模型模拟和机器学习预报。其中驱动数据制作过程采用距离反比权重法对流域自动站插值制作气象实况数据, 采用双线性插值法对智能网格、ECMWF 模式、CMA-GD 模式预报产品插值制作气象预报数据, 经预报降水优选后, 为物理模型模拟提供驱动数据。

物理模型模拟采用 Restart 方式, 将前一天的流域状态文件输入气象实况数据进行状态更新, 再输入气象预报数据进行 72h 流域区间径流量的预报。其中初始状态数据通过 2013—2020 年连续多次循环模拟后, 提取相同日期的气候态流域状态制作, 经过两个月的业务运行达到稳定后得到。

机器学习预报方面, 模型的训练使用 2013—2020 年岩滩水库入库实况径流量和模拟区间径流量数据。其中, 取 2013—2017 年为训练期, 训练期预报样本共计 1826 个, 2018—2020 年为验证期, 验证期预报样本共计 1096 个。输入的 49 个数据分别是 0~23h 前岩滩水库入库径流量、0~23h 前上游龙滩水库出库径流量以及距离起报时刻 t 小时物理模型模拟的区间径流量, 预报数据为距起报时刻 t 小时的岩滩水库入库流量; 总构建 72 个模型, 完成 $t+1$ 至 $t+72h$ 的岩滩水库入库径流量预报。

2.4.2 模型检验方法

选取纳什效率系数 (NSE)、相对误差合格率方法、洪峰出现时间时距差作为径流量预报性能评价指标, 计算公式如下:

$$NSE = 1 - \frac{\sum_{t=1}^T (Q_t^f - Q_m^t)^2}{\sum_{t=1}^T (Q_t^f - \bar{Q}_0)^2} \quad (1)$$

$$\text{相对误差} = (\text{预报} - \text{实况}) / \text{实况} \times 100\% \quad (2)$$

其中, Q_t^f 指 t 时刻实况径流量, Q_m^t 指 t 时刻预报径流量, \bar{Q}_0 表示实况径流量的总平均, NSE 介于 $(-\infty, 1]$ 之间, 取值为 1 时表示预报结果完美拟合实况值; 当相对误差绝对值 $\leq 15\%$ 时视为预报合格。根据《水文情报预报规范》(GB/T 224822-2008)^[20], 使用合格率来判断预报精度, 合格预报次数与预报总次数之比的百分数为合格率, 当合格率 $\geq 85\%$ 预报精度为甲级, $70\% \leq \text{合格率} < 85\%$ 预报精度为乙级, $60\% \leq \text{合格率} < 70\%$ 预报精度为丙级; 合格率 $\geq 70\%$ 的预报, 可用于发布正式预报, $60\% \leq \text{合格率} < 70\%$ 的预报, 可用于参考性预报; 对于洪峰出现时间误差以预报时间至实测出现时间之间时距的 30% 作为许可误差, 选取峰现时间误差 $\leq 3h$ 判定为合格。

3 预报模型模拟效果分析

从 2013—2020 年岩滩水库径流量数据中,选取 48h 内滑动平均径流量 $>3000\text{m}^3\cdot\text{s}^{-1}$ 或 24h 滑动流域平均面雨量和 $>12\text{mm}$ 的强降水过程作为预报模型强降雨过程检验样本样本,其中 2013—2017 年 15 次强降雨过程作为训练期样本,2018—2020 年 8 次强降雨过程作为验证期样本。图 2 实例中的强降雨过程径流量大多在 $3000\sim 3500\text{m}^3\cdot\text{s}^{-1}$, 其中图 2a、2b 显示了土壤湿度未饱和时产生蓄满产流的情况,图

2c 显示了土壤湿度未饱和且降水未达到土壤蓄满,流量以超渗产流为主的情况,图 d 显示了土壤湿度饱和产生蓄满产流的情况。试验表明,在土壤湿度未饱和时,蓄满产流峰现时间相对土壤湿度饱和情景下能够产生延迟,且降水未使土壤饱和时,产流洪峰并未出现,表明了预报模型能根据土壤湿度条件对洪峰是否出现、峰现时间延迟长度进行调节。采用 NSE 评估,训练期强降雨过程 NSE 为 0.66,验证期强降雨过程 NSE 为 0.65,较不使用 LSTM 模型,仅使用物理模型汇流的模拟 NSE 提升了 0.05。

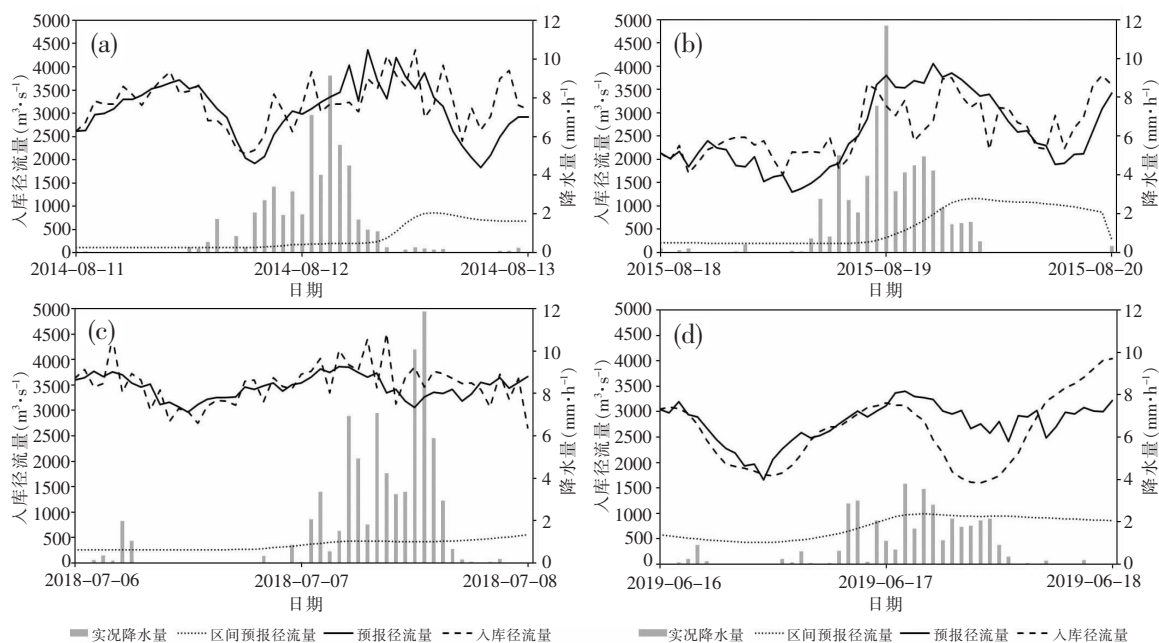


图 2 岩滩水库径流量预报模型模拟效果图

(a)2014 年 8 月 11 日—13 日;(b)2015 年 8 月 18 日—20 日;(c)2018 年 7 月 6 日—8 日;(d)2019 年 6 月 16 日—18 日

采用 48h 预报与实况径流量的相对误差为合格率标准进行评估,预报效果见表 1。从表 1 可见,日径流量预报相对误差绝对值训练期 15 次过程均小于 15%,预报合格率为 100%,达到甲级预报精度,验证期 8 次过程有 7 次过程小于 15%,预报合格率为 87.5%,达到甲级预报精度。对于洪峰预报相对误差绝对值,训练期 15 次过程有 14 次小于 15%,预报合格率为 93.3%,达到甲级预报精度,验证期 8 次过程均小于 15%,预报合格率为 100%,达到甲级预报精度。对于峰现时间预报误差,训练期 15 次过程有 13 次 $\leq 3\text{h}$ (达到合格要求),合格率为 86.7%,验证期 8 次过程有 5 次 $\leq 3\text{h}$ (达到合格要求),合格率为 62.5%。表明引入 LSTM 模型考虑上游放水影响

后,岩滩水库径流量预报有较好的预报精度;日径流量预报和洪峰预报,达到甲级预报精度,可用于发布正式预报;峰现时间预报,训练期的强降雨过程预报精度合格率达到甲级,可用于发布正式预报,而在验证期为丙级,可用于参考性预报。

4 业务应用效果

2021 年 12 月,通过建立岩滩水库径流量预报系统,实现岩滩水库径流量预报产品的业务试运行。图 3 给出 2021 年 12 月 18 日—2022 年 2 月 10 日 24~72h 日径流量预报效果。从图 3 可见,24h、48h、72h 日径流量预报与实况的相对误差绝对值分别为 48d、38d、40d 小于等于 15%,预报合格率分别为

表 1 岩滩水库训练期和验证期强降水过程日径流量预报、洪峰相对误差及峰现时间误差

强降水过程		峰值出现时间 (年/月/日/时)	峰值 ($\text{m}^3 \cdot \text{s}^{-1}$)	日径流量预报 相对误差(%)	峰值 相对误差(%)	峰现时间 误差(h)
训练期	1	2014/6/5/17	3290	12	5	-5
	2	2014/7/24/19	4090	-4	17	2
	3	2014/8/12/15	4360	4	0.1	3
	4	2015/4/29/21	3140	13	12	0
	5	2015/6/15/8	4060	-15	1.3	-6
	6	2015/7/1/16	3740	-5	6	2
	7	2015/8/19/13	4050	-3	-7	2
	8	2015/9/12/17	5150	-14	-7	1
	9	2016/4/7/20	3383	8	13	-1
	10	2016/6/11/20	3925	0.7	5	-2
	11	2016/7/6/13	3586	-2	11	2
	12	2017/6/5/15	3411	-5	-3	2
	13	2017/7/12/11	4644	-7	-8	-3
	14	2017/8/15/3	4631	-1	-3	-2
	15	2017/9/10/13	6007	-5	9	3
验证期	1	2018/5/28/17	3361	20	15	3
	2	2018/6/12/17	4698	3	-4	-3
	3	2018/6/25/17	4078	3	-15	5
	4	2018/7/7/1	4511	-1.6	2	7
	5	2019/6/17/2	3131	8	-15	2
	6	2019/7/24 11	4038	5	0.1	8
	7	2019/7/19 15	4319	14	-5	1
	8	2020/5/25 8	1792	4	3	1

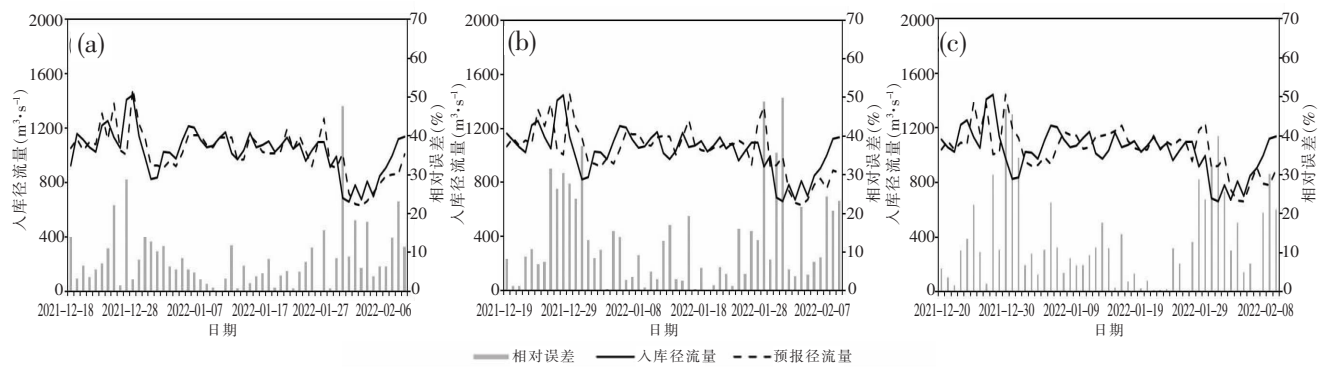


图 3 2021 年 12 月 18 日至 2022 年 2 月 10 日岩滩水库日径流量预报效果
(a)24h 预报效果;(b)48h 预报效果;(c)72h 预报效果

87.3%、70.4%、75.5%，表明 24h 预报达到甲级预报精度，48h、72h 预报达到乙级预报精度，均可用于发布正式预报。

表 2 给出岩滩水库径流量 $>1200\text{m}^3\cdot\text{s}^{-1}$ 的 3 次较大降水过程(2021 年 12 月 22—23 日,2021 年 12

月 25—26 日,2022 年 1 月 6—7 日)。从表 2 可见，其中 3 次过程日径流量相对误差均 $\leq 5\%$ ，峰现时间误差均 $\leq 1\text{h}$ ，预报合格率均为 100%，峰值相对误差绝对值有 2 次 $\leq 14\%$ ，有 1 次为 17%，预报合格率为 66.7%。表明 3 次较大降水过程径流量和峰现时间预

表 2 岩滩水库 2021 年 1 月—2022 年 2 月 3 次降水过程日径流量预报、洪峰相对误差及峰现时间误差

峰值出现时间 (年/月/日/时)	峰值 ($\text{m}^3 \cdot \text{s}^{-1}$)	起报时间 (年/月/日/时)	日径流量预报 相对误差(%)	峰值 相对误差(%)	峰现时间 误差(h)
2021/12/23/1	2919	2021/12/22/8	5	-17	0
2021/12/26/1	2501	2021/12/25/8	1	14	0
2022/1/7/2	2471	2022/1/6/8	4	6	1

报效果好,到达甲级预报精度,可用于发布正式预报,而峰值预报为丙级预报标准,可用于参考性预报。

5 结论

基于水文物理模型耦合机器学习方法,通过对全国智能网格、ECMWF 和 CMA-GD 模式预报产品的优选,驱动 CSSPv2 陆面水文模型计算区间径流量,在实况径流量中输入 LSTM 模型对岩滩入库径流量进行预报,得到以下结论:

(1)将陆面水文模型模拟的区间径流量结果输入 LSTM 模型进行预报后,岩滩水库日水文过程具有较好的模拟效果,率定期和验证期 NSE 效率系数在 0.65 左右,与仅使用陆面水文模型模拟相比,NSE 效率系数提升了 0.05。

(2)在强降水评估实验中,日径流量预报和洪峰预报验证期预报合格率 $\geq 87.5\%$,为甲级预报精度标准,对该流域具有较好的应用价值;峰现时间预报合格率为 62.5%,为丙级预报标准。

(3)在业务试用中,24h、48h、72h 日径流量预报合格率分别为 87.3%、70.4%、75.5%,其中 24h 预报合格率达到甲级预报精度,48h、72h 预报合格率达到乙级预报精度,满足发布正式预报的精度要求;3 次较大降水过程径流量和峰现时间预报合格率均为 100%,达到甲级预报精度,满足发布正式预报的精度要求;峰值预报合格率为 66.7%,达丙级预报标准,可用于参考性预报。

在岩滩流域水库径流量预报业务中,将物理模型与人工智能进行有机地耦合,充分利用机器学习的优势,弥补模型的不足之处,对提高水文预报产品的精度和适用性起到很好的借鉴作用。但对于降水未达土壤湿度饱和状态时,径流量主要由上游放水控制,LSTM 模型对于径流量的削减过多,降低了峰值径流量,在今后研究中需要进一步改进。

参考文献:

[1] 韦润新,林芳.岩滩水电站水文预报方案研究[J].企业科

技与发展,2011,312(18):60-61,73.

[2] 黄彩珍.岩滩水库水文预报技术分析[J].企业科技与发展,2014,385(21):34-36

[3] 包红军,王莉莉,沈学顺,等.气象水文耦合的洪水预报研究进展[J].气象,2016,42(9):1045-1057.

[4] 崔春光,彭涛,沈铁元,等.定量降水预报与水文模型耦合的中小流域汛期洪水预报试验[J].气象,2010,36(12):56-61.

[5] 曾鹏,钟利华,郑凤琴,等.基于智能网格降水产品的降尺度释用技术研究[J].气象研究与应用,2021,42(1):74-79.

[6] 彭筱,陈晓燕,黄武斌.2016 年夏季不同分辨率 GRAPES_MESO 模式的西北地区预报效果检验[J].气象研究与应用,2019,40(4):6-11.

[7] 卢小凤,李仲怡,陈剑飞,等.智能网格产品在西江流域面雨量预报中的应用检验[J].气象研究与应用,2020,41(2):45-49.

[8] 史彩霞,郑凤琴,曾鹏,等.智能网格预报在重大活动气象服务保障中的应用——以环广西公路自行车世界巡回赛为例[J].气象研究与应用,2020,41(1):61-64.

[9] 郭玉雪,许月萍,陈浩,等.基于多种递归神经网络的海岛水库径流预报[J].水力发电学报,2021,40(9):14-26.

[10] 殷兆凯,廖卫红,王若佳,等.基于长短时记忆神经网络(LSTM)的降雨径流模拟及预报[J].南水北调与水利科技,2019,17(6):1-9,27.

[11] 梁振清,陈生.基于深度学习和雷达观测的华南短临预报精度评估[J].气象研究与应用,2020,41(1):41-47.

[12] Liu J, Yuan X, Zeng J H, et al. Ensemble Streamflow Forecasting over a Cascade Reservoir Catchment with Integrated Hydrometeorological Modeling and Machine Learning [J]. Hydrology and Earth System Science, 2022 (26): 265-278.

[13] Dai Y J, Dickinson R, Wang Y P. A Two-Big-Leaf Model for Canopy Temperature, Photosynthesis, and Stomatal Conductance [J]. Journal and Climate, 2004, 17 (12): 2281-2299.

[14] Dai Y, Zeng X, Dickinson R. The Common Land Model [J]. Bull. Am. Meteorol. Soc., 2003, 84(8): 1013-1024.

[15] Yuan X, Ma F, Wang L, et al. An Experimental Seasonal Hydrological Forecasting System over the Yellow River

- Basin-Part 1: Understanding the Role of Initial Hydrological Conditions[J]. *Hydrology and Earth System Science*, 2016, 20(6): 2437–2451.
- [16] Yuan X, Ji P, Wang L, et al. High - Resolution Land Surface Modeling of Hydrological Changes over the Sanjiangyuan Region in the Eastern Tibetan Plateau: 1. Model Development and Evaluation[J]. *Journal of Advances in Modeling Earth Systems*, 2018(10): 2806–2828.
- [17] Duan Q Y, Sorooshian S, Gupta V K. Optimal Use of SCEUA Global Optimization Method for Calibrating Watershed Models[J]. *Journal of Hydrology*, 1994(158): 265–284.
- [18] Getirana A C V, Boone A, Yamazaki D, et al. The Hydrological Modeling and Analysis Platform (HyMAP): Evaluation in the Amazon Basin[J]. *Journal of Hydrometeorology*, 2012(13): 1641–1665.
- [19] Luo X, Li H Y, Leung L R, et al. Modeling Surface Water Dynamics in the Amazon Basin Using Mosart-Inundation V1.0: Impacts of Geomorphological Parameters and River Flow Representation [J]. *Geoscientific Model Development Discussions*, 2017, 10(3): 1–42.
- [20] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. 水文情报预报规范: GB/T22482–2008[S]. 北京: 中国标准出版社, 2008.

Yantan Reservoir runoff forecast and application coupled with land surface hydrological model and machine learning method

Chen Jianfei¹, Li Yong¹, Liu Junjiang², Zhong Lihua^{1*}, Shi Caixia¹, Yuan Xing^{2,3}, Zhong Huachang⁴

(1. Guangxi Meteorological Disaster Prevention Center, Nanning 530022, China; 2. School of Hydrology and Water Resources, Nanjing University of Information Science and Technology, Nanjing 210044, China; 3. Key Laboratory of Regional Climate-Environment for Temperate East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029; 4. Guangxi Guiguan Electric Power Co., Ltd., Nanning 530029, China)

Abstract: Using a variety of meteorological forecast data such as intelligent grid forecasting to couple with land surface hydrological models and machine learning methods, taking the Yantan Reservoir watershed as an example, the interval forecast runoff and historical runoff data were input into the long-term and short-term memory network model (LSTM) to forecast the inflow runoff of the reservoir. The applicability of this method in short-term hydrological forecasting was proved through the simulation test of daily runoff and the analysis of operational trial. The results show that the Nash efficiency coefficient (NSE) of the runoff prediction model coupled with meteorology-hydrology-machine learning is about 0.65 in the calibration period and verification period. In the training test of heavy rainfall process, the qualified rate of daily runoff and flood peak prediction is greater than or equal to 87.5%, reaching the standard of A-level forecast accuracy. In the operational trial, the qualified rates of 24h, 48h, and 72h daily runoff forecasts are 87.3%, 70.4%, and 75.5%, respectively, reaching the A or B level forecast accuracy standards and meeting the accuracy requirements for issuing official forecasts. The qualified rate of runoff and peak present time forecasts for the three major precipitation processes are all 100%, reaching the A-level forecast accuracy standard. The peak forecast is 66.7%, which meets the C-level forecast standard and can be used for reference forecasting. The organic coupling of physical models and artificial intelligence methods can improve the accuracy and applicability of hydrological forecast products.

Key words: coupling; land surface hydrological model; machine learning; runoff forecast; physical model; artificial intelligence