J. Lake Sci.(湖泊科学),2025,**37**(1):279-292
DOI 10.18307/2025.0143
© 2025 by *Journal of Lake Sciences*

基于概念性水文模型与长短时记忆模型耦合的汉江上游流域径流模拟*

邓　超[1,2**],孙培源[1,2],尹　鑫[3],邹佳成[4],王卫光[1,2]
(1:河海大学水文水资源学院,南京 210098)
(2:河海大学,水灾害防御全国重点实验室,南京 210098)
(3:南京水利科学研究院,水灾害防御全国重点实验室,南京 210029)
(4:赣江下游水文水资源监测中心,宜春 336000)

摘　要:为了探究概念性水文模型(GR4J)与长短时记忆模型(LSTM)耦合过程中物理模型参数反馈机制以及考虑土壤含水量作为中间变量对物理引导机器学习(PIML)模型径流模拟的影响,本研究构建了 PIML 模型并设置考虑参数反馈、考虑中间变量和两者同时考虑的 3 种方案(依次简称为 H1、H2、H3),以安康站为控制站的汉江上游流域开展实例研究。结果表明:(1)3 种 PIML 模型径流模拟效果均优于 LSTM 模型,其中验证期纳什系数(*NSE*)平均提升 10.6%,而 PIML-H1 与 PIML-H3 模型径流模拟效果优于 GR4J 模型,验证期 *NSE* 平均提升 4.2%,其中 PIML-H3 模型在 3 种 PIML 模型中表现最佳,表明基于 LSTM 和 GR4J 模型构建的 PIML 模型对径流模拟效果有所改善,且同时考虑中间变量和参数反馈因素时对径流模拟效果改善最佳;(2)3 种 PIML 模型对低水流量的模拟效果均优于 GR4J 和 LSTM 模型,且 PIML-H3 模型效果最佳,而对于高水流量,3 种 PIML 模型均表现不佳,说明 PIML 模型往往在模拟低流量事件中更占优势;(3)3 种 PIML 模型的径流模拟效果均表现出较强的季节性变化,PIML-H2 和 PIML-H3 模型相较于 PIML-H1 模型受到的季节性变化影响更为明显,说明 PIML 模型模拟径流结果的季节性变化受到中间变量的影响。研究可为汉江上游流域径流模拟预报提供技术支撑。
关键词:物理引导机器学习;径流模拟;中间变量;GR4J;LSTM;汉江

Runoff simulation in the upper Han River Basin using physics-informed machine learning-model*

Deng Chao[1,2**], Sun Peiyuan[1,2], Yin Xin[3], Zou Jiacheng[4] & Wang Weiguang[1,2]
(1: *College of Hydrology and Water Resources*, *Hohai University*, *Nanjing* 210098, *P.R.China*)
(2: *The National Key Laboratory of Water Disaster Prevention*, *Hohai University*, *Nanjing* 210098, *P.R.China*)
(3: *The National Key Laboratory of Water Disaster Prevention*, *Nanjing Hydraulic Research Institute*, *Nanjing* 210029, *P.R.China*)
(4: *Hydrology and Water Resources Monitoring Center of Lower Ganjiang River*, *Yichun* 336000, *P.R.China*)

Abstract: This study investigated the impact of coupling the conceptual hydrological model (GR4J) with the long short-term memory model (LSTM) in a physics-informed machine learning (PIML) framework for runoff simulation. Three scenarios (H1, H2 and H3) were designed to examine the effects of the physical model parameter feedback mechanism, the consideration of soil moisture as an intermediate variable, and the former both on the PIML models, respectively. The case study was conducted in the upper Han River Basin, with the Ankang hydrological station as the control station. The main findings were as follows: (1) Compared with the LSTM model, all three PIML models had improved performance on runoff simulation, with a 10.6% increase in average Nash-Sutcliffe efficiency (*NSE*) during the validation period. Additionally, both the PIML-H1 and PIML-H3 models exhibited better performance than the GR4J model, with a 4.2% increase in average *NSE* during the validation period. Notably, the PIML-H3

* 2024-03-14 收稿;2024-05-09 收修改稿。
国家重点研发计划项目(2022YFC3202802)和国家自然科学基金项目(51979071)联合资助。
** 通信作者;E-mail:dengchao@hhu.edu.cn。

model outperformed other PIML models, indicating that coupling GR4J and LSTM models simultaneously considering intermediate variables and parameter feedback yielded the most significant improvement in the model performance of runoff simulation. (2) For low flows, all three PIML models outperformed the GR4J and LSTM models, and the PIML-H3 model achieved the best performance. For high flows, the performance of all three PIML models was not high, implying that PIML models were suitable in simulating low flows events. (3) The runoff simulations from the three PIML models exhibited significantly seasonal variations during both the training and validation periods. The seasonal variations in the PIML-H2 and PIML-H3 models were more pronounced compared to that in the PIML-H1 model, indicating that the seasonal variations in simulated runoff results of the PIML model were influenced by intermediate variables. This study contributed to a better understanding of the performance differences among various PIML model schemes in runoff simulation, providing technical support for runoff simulation and forecasting in the study area.

Keywords: Physics-informed machine learning; runoff simulation; intermediate variable; GR4J; LSTM; Han River

近年来,受气候变化和高强度人类活动共同影响,极端水文事件发生频次及量级远超历史,流域水安全和水资源管理问题更加突出[1]。我国政府明确提出要以全面提升水安全保障能力为目标,完善流域防洪减灾体系为重点。径流模拟与流域防洪减灾、水资源管理等密切相关,一直以来都是流域水文学的关键任务[2-6]。作为流域径流模拟的重要工具,流域水文模型通常可分为两类:基于物理概念的模型和基于数据驱动的模型[7]。相较于物理概念模型,数据驱动模型由于计算复杂性低且模型精度高,已广泛用于径流模拟预报[8],尤其是机器学习(machine learning,ML)模型,如极端梯度提升(extreme gradient boosting,XGBoost)、循环神经网络(recurrent neural network,RNN)、长短时记忆模型(long short-term memory model,LSTM)[8-9]。近年来,利用 LSTM 开展流域径流模拟得到了广泛的应用与验证[10-12]。然而,由于其“黑箱”性质,模型仅考虑输入、输出变量之间的统计学关系而忽略了流域水文过程的物理机制[13],使得模型结果存在过度拟合以及过程缺乏解释性等问题[14]。

针对 ML 模型存在的上述问题,许多学者将物理模型和机器学习耦合开展水文模拟预报研究[15-16]。物理引导机器学习(physics-informed machine learning,PIML)作为一种耦合方法,能够结合物理模型对流域水文过程机理的概化描述和 ML 模型的数据挖掘能力,通过将已知的物理定律或规律整合到 ML 中,并对 ML 模型的输出进行约束或限制从而提高预测的稳定性,该类方法已经在降雨径流预测中得到有效应用[17-18]。目前,常用的 PIML 方法是将基于物理模型得到的输出结果(如模拟流量,水文中间变量)作为 ML 模型的输入[19-21],或构建独立的参数通道以耦合物理模型与 ML 模型[17]。例如,Yu 等[22]将 HBV 水文模型与 LSTM 模型相结合,考虑了 LSTM 模型和 HBV 模型之间的参数反馈,但其训练机器学习模型的输入信息中未考虑水文中间变量(如蒸散发、土壤含水量)对径流过程的影响。水文中间变量作为降雨和径流之间的桥梁,充分考虑其对径流的影响有助于增强有效信息的利用,进而提高 ML 模型模拟的稳定性和精准度[23-24]。Bhasme 等[25]提出的 PIML 方案将概念性水文模型与机器学习算法相结合,通过概念性水文模型提取的变量函数关系确定 ML 模型输入来先后预测中间变量(即实际蒸散发)和目标变量(即径流),该方法考虑了中间变量对径流预测的影响。但在构建 PIML 模型时,概念水文模型与 ML 模型之间的直接反馈联系在该框架中尚未得到充分探讨。此外,考虑中间变量影响的 GR4J 模型和 ML 模型耦合的问题研究尚缺乏。

因此,本研究基于概念性水文模型(Génie Ruralà 4 paramètres Journalier,GR4J)与 LSTM 模型构建 PIML 模型,考虑了 GR4J 与 LSTM 模型的参数反馈以及中间状态变量对于 PIML 模型构建的影响,通过设置 3 种不同 PIML 方案,以汉江上游流域开展流域日径流模拟研究。

1 数据资料

1.1 流域概况

汉江作为长江最大的支流,全长 1577 km。其发源地位于陕西省宁强县冢山,流域横跨湖北、陕西、甘肃、四川、重庆 5 个省(市),总面积达 15.9 万 km^2[26-27]。本文以安康水文站为控制站的汉江上游流域为研究对象,该站点位于陕西省安康市汉滨区,地理坐标为 33.67°N,109°E(图 1),以安康站为控制站的集水面积达 3.98 万 km^2,占据整个汉江流域总面积的 25%。该地区北靠秦岭,南与大巴山相接,地势两边高、中间低,属于亚热带湿润性季风气候。受秦岭、巴山的影响,该地区同时兼有暖湿带和中温带山地气候特征,四季分

明,雨量充沛,无霜期长,多年平均年降水量为 976 mm,年平均潜在蒸发量为 897 mm,年平均径流深为 446 mm[28]。此外,该地区历史上植被覆盖面积广,以常绿落叶阔叶混交林为主。由于该地区气温较高,降水较多,岩石的物理、化学风化和溶蚀作用也比较强烈,在缓坡地带形成了厚厚的松散坡积物和棕黄色的亚黏土堆,形成重力地貌。

1.2 数据来源与处理

研究使用的数据主要包括地面观测数据和遥感数据。

地面观测数据包括流量和气象数据。其中流量资料来源于水文年鉴安康水文站的实测逐日平均流量。气象数据包括降水(P)、温度、相对湿度、日照时速、气压、风速,该数据来源于国家气象科学数据中心(https://data.cma.cn)。因考虑到数据质量以及数据序列的连续性,选取了汉江上流流域内及相邻 11 个气象站点(图 1)的逐日观测数据。

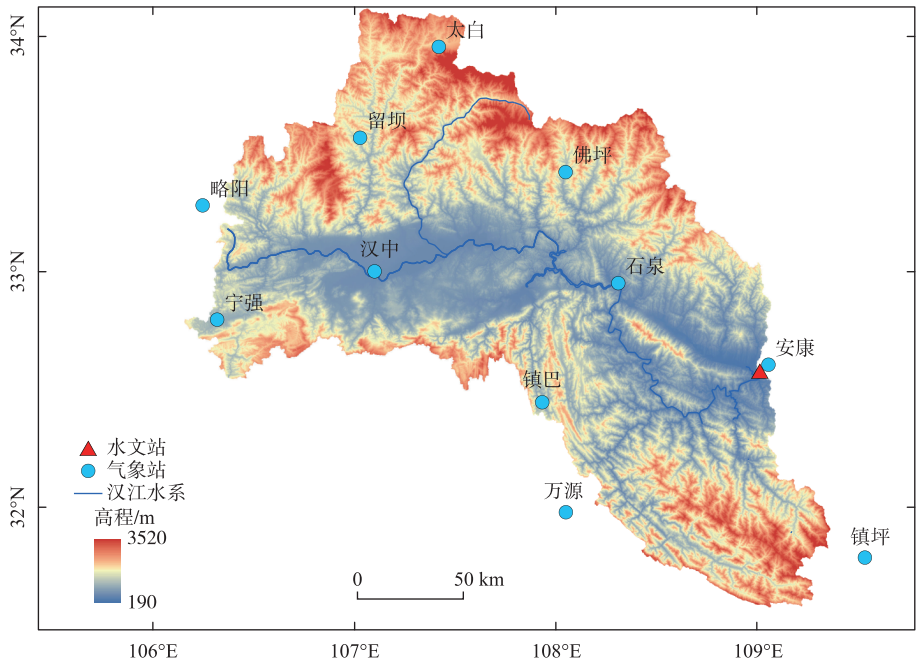


图 1 汉江上游流域地理位置及气象站、水文站分布

Fig.1 Location of the upper Han River Basin and the distribution of meteorological and hydrological stations

遥感数据为逐日土壤含水量(S)数据,该数据来源于美国国家航空航天局全球陆地数据同化系统(global land data assimilation system,GLDAS)的卫星数据产品。本文采用的数据为 GLDAS-2.1 版本,序列起止时间为 2000 年至今,时间分辨率为 3 h,空间分辨率为 0.25°,来源于 https://doi.org/10.5067/G90R32A924YM。研究区的面平均逐日土壤含水量基于流域边界并采用 Google Earth Engine(https://earthengine.google.com/)提取得到,其中面平均方法采用空间数据的区域统计方法,使用 GEE 中自带的 reduceRegion 函数,且函数 reducer 属性设置为 Reducer.mean。

此外,根据水文模型输入数据要求,潜在蒸散发(PET)数据根据已有气象资料采用彭曼-蒙特斯(Peman-Monteith)[29]公式计算得出。考虑到降水、潜在蒸散发以及土壤含水量与径流数据在时间序列上的匹配性及数据质量,本文选取 2000 年 1 月 1 日和 2011 年 12 月 31 日作为本次研究的数据序列起止时间。

2 研究方法

2.1 GR4J 模型

GR4J 模型[30]是一个集总式概念性水文模型,其采用两个非线性水库来模拟产流和汇流过程,模型输入

为日降水 P(mm)和日潜在蒸散发 PET(mm)两部分。该模型已广泛应用于国内外湿润、半湿润地区的径流模拟与预报中[31-32]。该模型仅有 4 个参数,分别为产流水库最大容量 x_1(mm)、地下水交换系数 x_2(mm)、汇流水库最大容量 x_3(mm)、单位线汇流时间 x_4(d)。

2.2 LSTM 模型

LSTM 模型是由 Hochreiter 和 Schmidhuber[33] 在 1997 年首次提出的一种特殊类型的循环神经网络(RNN),该模型广泛应用于基于时间序列数据的分类、处理和预测[34]。与 RNN 相比,LSTM 在隐藏层中添加了一个记忆单元,取代了 RNN 中的原始单元,可以对带有时间段的序列数据进行特征提取,克服了 RNN 的梯度爆炸和消失弱点[35]。本文中使用的单一 LSTM 模型结构采用两层 LSTM 层,隐藏层为一层,Dropout 层数为两层。LSTM 模型结构可由以下方程描述:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{4}$$

$$O_t = \sigma(W_O \cdot [h_{t-1}, x_t] + b_O) \tag{5}$$

$$h_t = O_t \times \tanh(C_t) \tag{6}$$

式中,t 表示时间步长;x_t为输入到 LSTM 单元的输入向量;f_t为遗忘门的激活向量;i_t为输入门的激活向量;O_t为输出门的激活向量;h_t为隐藏状态向量,也称为 LSTM 单元的输出向量;$\tilde{C}_t$为细胞输入激活向量;C_t为细胞状态向量;W 和 b 为需要在训练过程中更新的权重矩阵和偏差向量参数。

2.3 PIML 模型

本文采用的 PIML 模型框架由 Bhasme 等[20] 提出,其实质是从概念性物理水文模型中提取水文变量之间的关系,随后用于确定基于 ML 模型预测目标变量的输入。本研究基于 GR4J 与 LSTM 模型构建 PIML 模型,并分别设置了考虑参数反馈、考虑中间变量和两者同时考虑共 3 种方案(依次简称为 H1、H2、H3),各方案流程如图 2 所示。

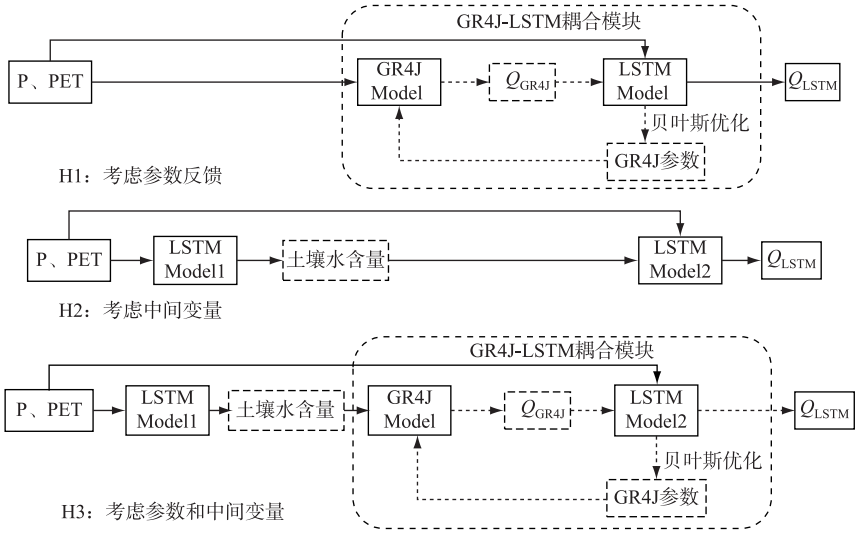


图 2 3 种 PIML 模型方案流程

Fig.2 The architectural framework of three scenarios of PIML model

H1 方案考虑了 GR4J 模型与 LSTM 模型之间的输入输出反馈;H2 方案中,根据产汇流原理,通过采用两层 ML 模型来考虑中间变量的约束作用;H3 方案则是一种综合考虑参数变量和中间变量的方案。3 种 PIML

模型方案的原始输入包括降水和潜在蒸散发。

在 H1 方案中,建立 GR4J 模型和 LSTM 模型的耦合模块。该模块的输入包括 P、PET,输出为 LSTM 模型的模拟流量 Q_{LSTM},其中 GR4J 模型的模拟流量 Q_{GR4J} 被视为中间变量。在此过程中,采用贝叶斯优化算法对 LSTM 模型的超参数和 GR4J 模型的参数进行调优。调优过程中设置初始的 GR4J 模型参数集(x_1、x_2、x_3、x_4)以及 LSTM 的超参数,并在取值范围内迭代搜索使得模型性能最优的参数组合。当前迭代过程得到新的最优 GR4J 模型参数集(x_1^1、x_2^1、x_3^1、x_4^1)将会替换原参数以生成 Q_{GR4J1}。更新后的 GR4J 模型将再次与 LSTM 模型耦合,LSTM 模型以 P、PET 和新的 Q_{GR4J1} 作为输入,生成 Q_{LSTM1}。通过对 GR4J 模型参数不断更新实现参数反馈,直至搜寻到满足 Q_{LSTM} 模拟最优条件下的参数组合,并输出最优径流模拟结果。

H2 方案是通过从 GR4J 模型中提取的变量函数关系来确定 LSTM 模型的输入来先后预测中间变量(土壤含水量)和目标变量(即径流),该方法考虑了中间变量对径流预测的影响。

在 H2 和 H3 方案中,降水和潜在蒸散发数据首先输入到第一层 LSTM 模型模拟中间变量,随后,中间变量、降水和潜在蒸散发一并输入到下一层 LSTM 模型中;在两个方案中,GR4J 模型的中间变量土壤水含量计算模块用 ML 模型替换。H2 和 H3 两个方案的不同之处在于,H2 的中间变量和原始输入被代入第二层 LSTM 模型,而 H3 则将其代入 GR4J 模型中,其余的过程与 H1 相似。

本研究基于 GR4J、LSTM、PIML-H1、PIML-H2、PIML-H3 5 种水文模型进行汉江上游流域的日径流模拟,各模型的输入输出变量如表 1 所示。GR4J 模型参数率定基于 Python 3.10 版本下 Scipy 库中的 L-BFGS-B 优化函数实现。针对涉及到的机器学习模型,均采用贝叶斯优化[36]进行超参数优化。它是一种高效、鲁棒且具有理论支持的方法,能够更快地找到模型的最佳超参数组合,提高模型的性能和泛化能力[37]。

表 1 水文模型输入、输出变量概况

Tab.1 Description of input and output variables of the hydrologic models

模型	输入变量	中间变量	输出变量	目标变量
GR4J	P、PET	—	Q_{GR4J}	Q_{obs}
LSTM	P、PET	—	Q_{LSTM}	Q_{obs}
PIML-H1	P、PET、Q_{GR4J}	—	GR4J 参数、$Q_{Hybrid1}$	Q_{obs}
PIML-H2	P、PET、Q_{GR4J}	土壤水含量	中间变量、$Q_{Hybrid2}$	Q_{obs}
PIML-H3	P、PET、Q_{GR4J}	土壤水含量	中间变量、GR4J 参数、$Q_{Hybrid3}$	Q_{obs}

2.4 评价指标

选取纳什系数(*NSE*)和均方根误差(*RMSE*)两种指标来评价水文模型在不同方案下的径流模拟效果,各指标的计算公式为:

$$NSE = 1 - \frac{\sum_{t=1}^{n}(Q_{sim,t} - Q_{obs,t})^2}{\sum_{t=1}^{n}(Q_{obs,t} - \overline{Q}_{obs,t})^2} \tag{7}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(Q_{sim,t} - Q_{obs,t})^2}{n}} \tag{8}$$

式中,$Q_{sim,t}$ 和 $Q_{obs,t}$ 分别表示 t 时刻模拟流量和实测流量;$\overline{Q}_{obs}$ 表示实测流量的平均值;n 表示时间序列长度;其中,*NSE* 越接近 1,*RMSE* 越小,说明模型径流模拟效果越好,反之越差。

3 结果分析

3.1 参数率定

GR4J 参数取值范围基于其物理概念并根据前期文献进行设定[38-39],模型参数率定结果如表 2 所示。对于 4 种机器学习模型,采用贝叶斯优化方法对模型的待估计超参数进行调参,超参数取值范围参考前期

表 2 GR4J 模型参数率定结果

Tab.2 Optimal values of the GR4J model's parameters

模型参数	最优值	取值范围
x_1	100	40~1000
x_2	−1.11	−10~0
x_3	120	50~150
x_4	1.34	1~3

文献设定,训练期 4 种模型的超参数取值范围相同[40-41],验证期均采用训练期的最优超参数值,参数估计结果如表 3 所示。上述水文模型率定得到的模型参数最优值均在取值范围内,说明模型各参数取值均在合理范围内分布,可进一步用于流域径流模拟。

3.2 模型径流模拟评价

3.2.1 模型整体评价 本研究将数据集划分为预热期、训练期、验证期 3 部分,其中数据集时间跨度为 2000 年 1 月 1 日−2011 年 12 月 31 日。数据集前 1 个月的数据(2000 年 1 月 1 日−1 月 31 日)用作模型预热,剩余数据的前 80%用于模型训练,后 20%用于模型验证。

表 3 机器学习模型超参数优化结果

Tab.3 Optimal values of hyperparameters of the machine learning models

模型	丢弃率	神经元数	学习率	迭代次数	批量大小
LSTM	0.45	78	0.0095	67	39
PIML-H1	0.1	128	0.001	55	96
PIML-H2	0.54	45	0.0016	115	90
PIML-H3	0.13	78	0.0012	86	112
取值范围	0~0.9	30~512	0.001~0.01	40~120	32~1024

表 4 和图 3 分别展示了 5 种模型在训练期和验证期(2000 年 2 月 1 日−2011 年 12 月 31 日)的模拟径流评价结果及径流模拟过程。总体而言,PIML-H3 模型在所有模型中径流模拟结果最好,PIML-H1 模型次之,而 PIML-H2 与 LSTM 模型的模拟结果相对较差。具体而言,3 种 PIML 模型径流模拟效果均优于 LSTM 模型,其中验证期 *NSE* 平均提升 10.6%;仅 PIML-H1 与 PIML-H3 模型径流模拟效果优于 GR4J 模型,验证期 *NSE* 平均提升 4.2%,PIML-H2 模型的改进效果并不显著,这说明基于物理概念公式确定目标变量的输入无法完全确保 ML 对其模拟精度的提高[42-43],仅考虑土壤含水量作为补充输入对 ML 的改善效果有限。个别模型方案验证期精度高于训练期,主要由于采用贝叶斯优化机器学习超参数可能使模型正则化过多,从而导致模型在训练期和验证期模拟结果相差较大。此外,模型在训练期与验证期均方根误差均较大,主要是由于模型在高流量时的模拟值偏低。验证期模拟流量与实测流量绝对误差较大,主要由于验证期内 2011 年夏季出现的洪水洪量偏大,但训练期出现过与验证期内相同量级的洪水,因此,在进行参数率定时已将洪水造成的模拟影响考虑在内[3]。从总体上看,5 种模型在汉江上游的径流模拟效果较为合理。

表 4 训练期与验证期水文模型径流模拟评价

Tab.4 Evaluation of runoff simulations obtained from the five hydrologic models during the calibration and validation periods

模型评价指标	训练期		验证期	
	NSE	*RMSE*/(m^3/s)	*NSE*	*RMSE*/(m^3/s)
GR4J	0.55	792.92	0.71	507.10
LSTM	0.64	507.16	0.68	709.94
PIML-H1	0.81	373.41	0.74	650.01
PIML-H2	0.65	488.66	0.69	760.65
PIML-H3	0.84	447.17	0.76	401.07

从 5 种模型 2000 年 2 月 1 日−2011 年 12 月 31 日的流量模拟值与实测值的散点图(图 4)可以看出,GR4J 模型和 PIML-H1、PIML-H3 模型的散点图呈现出比其他模型更为紧密的分布,而 LSTM 和 PIML-H2 模型的散点图明显较为松散。在 3 个 PIML 模型中,PIML-H1(R^2 = 0.83)表现最好,其次是 PIML-H3(R^2 =

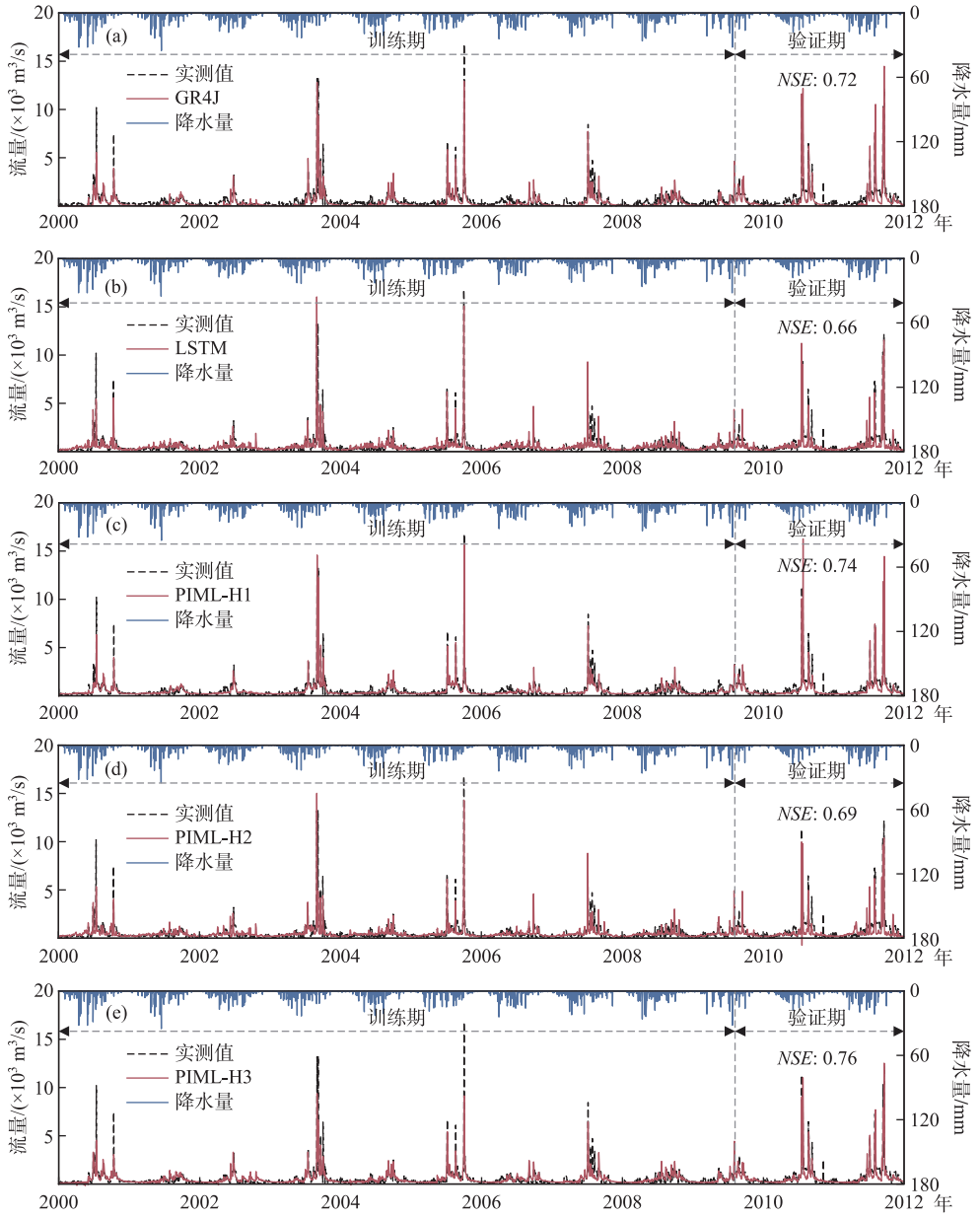


图 3 训练期与验证期 5 种水文模型实测和模拟流量

Fig.3 Observed and simulated flow by the five models during the calibration and validation periods

0.78),最后是 PIML-H2(R^2=0.74)。尽管 PIML-H3 模型在主要评价指标上相较 GR4J 模型有所提升,但是 R^2却有所下降,这是由于模型 PIML-H3 模型比 GR4J 模型更复杂,模拟过程中更容易出现过拟合现象,导致数据泛化能力较差[44]。

3.2.2 高水低水评价　图 5 展示了 5 种模型实测流量和模拟流量的流量历时曲线(FDCs)。总体而言,除 PIML-H1 模型外,其他 4 种模型均在一定程度上低估了流量,尤其在高水部分更为显著。具体而言,PIML-H1、PIML-H2、PIML-H3 模型以及 LSTM 模型在低水流量方面倾向于高估(图 5b),而 GR4J 模型则倾向于低估。在模拟低水流量时,3 种 PIML 模型效果均优于 GR4J 和 LSTM 模型,且 PIML-H3 模型效果最佳;而在模

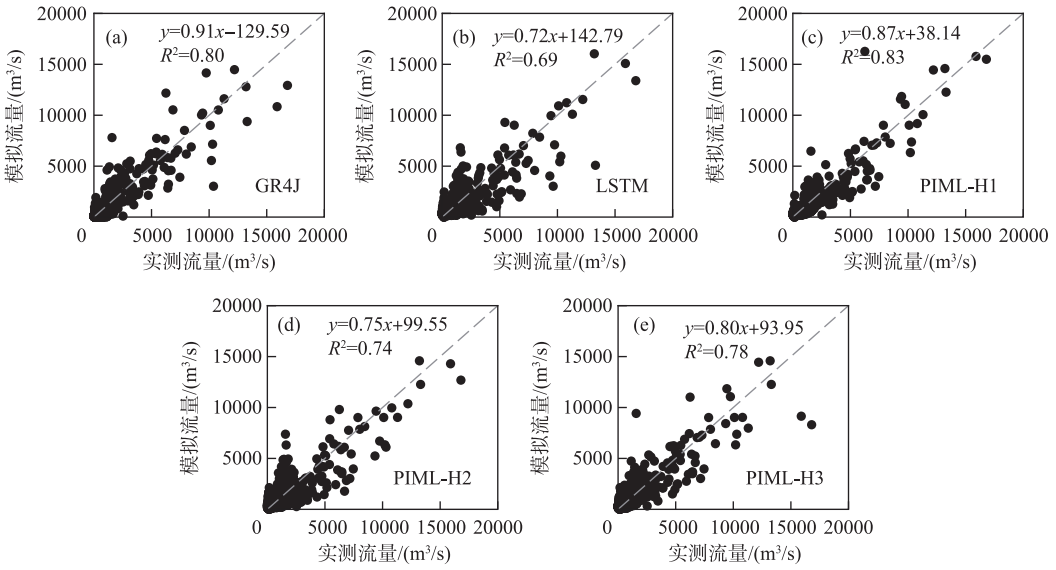


图 4 训练期与验证期 5 种模型模拟流量和实测流量散点图

Fig.4 Scatterplot of simulated and observed flow of five models during the calibration and validation periods

拟高水流量时,3 种 PIML 模型均表现不佳(图 5a),说明 PIML 模型在模拟低水流量事件中更占优势。

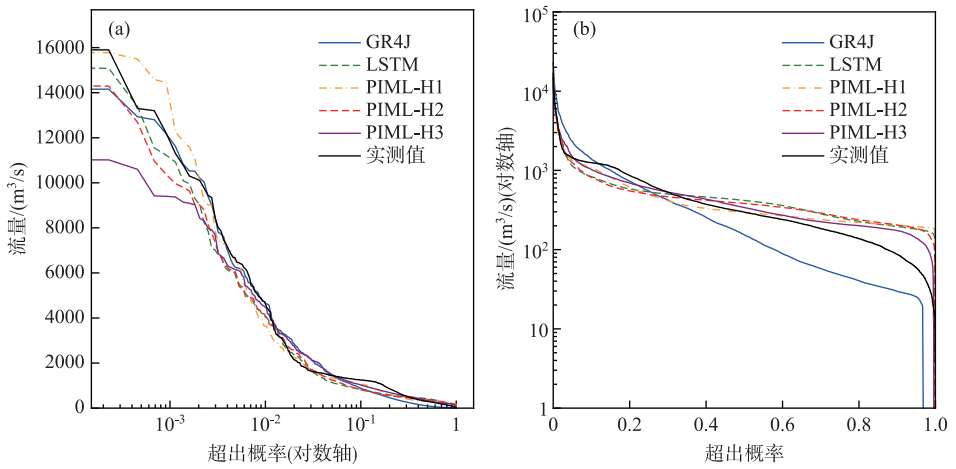


图 5 训练期与验证期 5 种模型实测流量和模拟流量的流量历时曲线

Fig.5 Flow duration curves (FDCs) of observed and simulated flow of five models during the calibration and validation periods

图 6 呈现了 2010 年和 2011 年各模型的降水、流量变化情况。总体而言,除 LSTM 模型外,其他模型在 2011 年的预测中普遍表现为优于 2010 年。5 种模型均未能准确捕捉 2010 年 7 月初—8 月末的流量动态,其中 GR4J 模型、PIML-H1 模型、PIML-H3 模型在该阶段的模拟效果优于 LSTM、PIML-H2,除 GR4J 和 PIML-H3 模型外,其他 3 种模型在 2010 年 8 月的高流量情况下均存在低估。GR4J 模型在 2011 年 8 月高估了峰值流量,而 LSTM 和 PIML-H2 模型则存在低估。其中,PIML-H1 和 PIML-H3 模型在 2011 年 8 月的模拟峰值流量与实测流量非常接近。2011 年 10 月和 11 月期间,所有模型均呈现出流量低估的趋势,其中 GR4J 和 PIML-H3 模型相较于其他 3 个模型表现出的低估程度较小。相反,PIML-H2 模型在该阶段表现出了异常的高估情况,这与该时

期土壤含水量和短暂的小洪峰事件有关,较小的洪峰事件对初始土壤含水量更为敏感[45]。尽管 PIML-H3 同样涉及土壤含水量变量,但因其耦合过程受到多种因素制约,导致影响不明显。

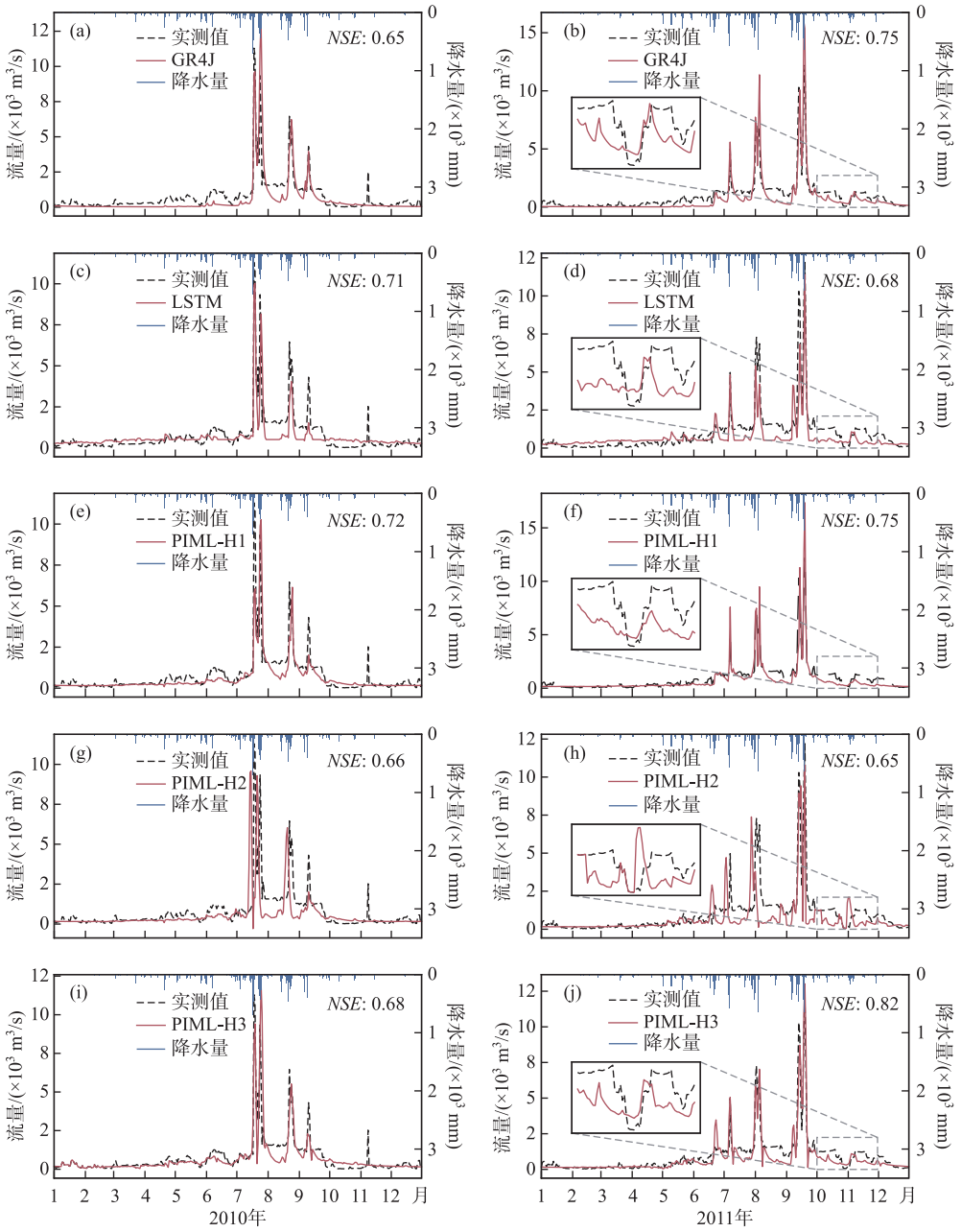


图 6 验证期模型模拟流量和实测流量的对比(左列为 2010 年,右列为 2011 年)

Fig.6 Comparison between simulated and observed flow during the validation period years (left and right columns represent the year of 2010 and 2011, respectively)

3.2.3 年际间变化趋势　图 7 给出了 3 种 PIML 模型在月尺度和多年平均月尺度上的径流模拟效果。从左列观察,3 种 PIML 模型均在 6－10 月间有较大的径流深模拟量,在年份维度上,3 种 PIML 模型均在 2003－

2005 年以及 2010－2011 年有较大的径流深模拟量。从右列观察,PIML-H1 和 PIML-H3 模型在整个数据集上的模拟效果明显优于 PIML-H2 模型。在月份维度上,表现较好的模拟结果集中在 6－10 月之间,说明在该时期 PIML 模型的模拟能力相对较强,且 PIML-H3 模型相较于 PIML-H1 模型受到的季节性变化影响更为

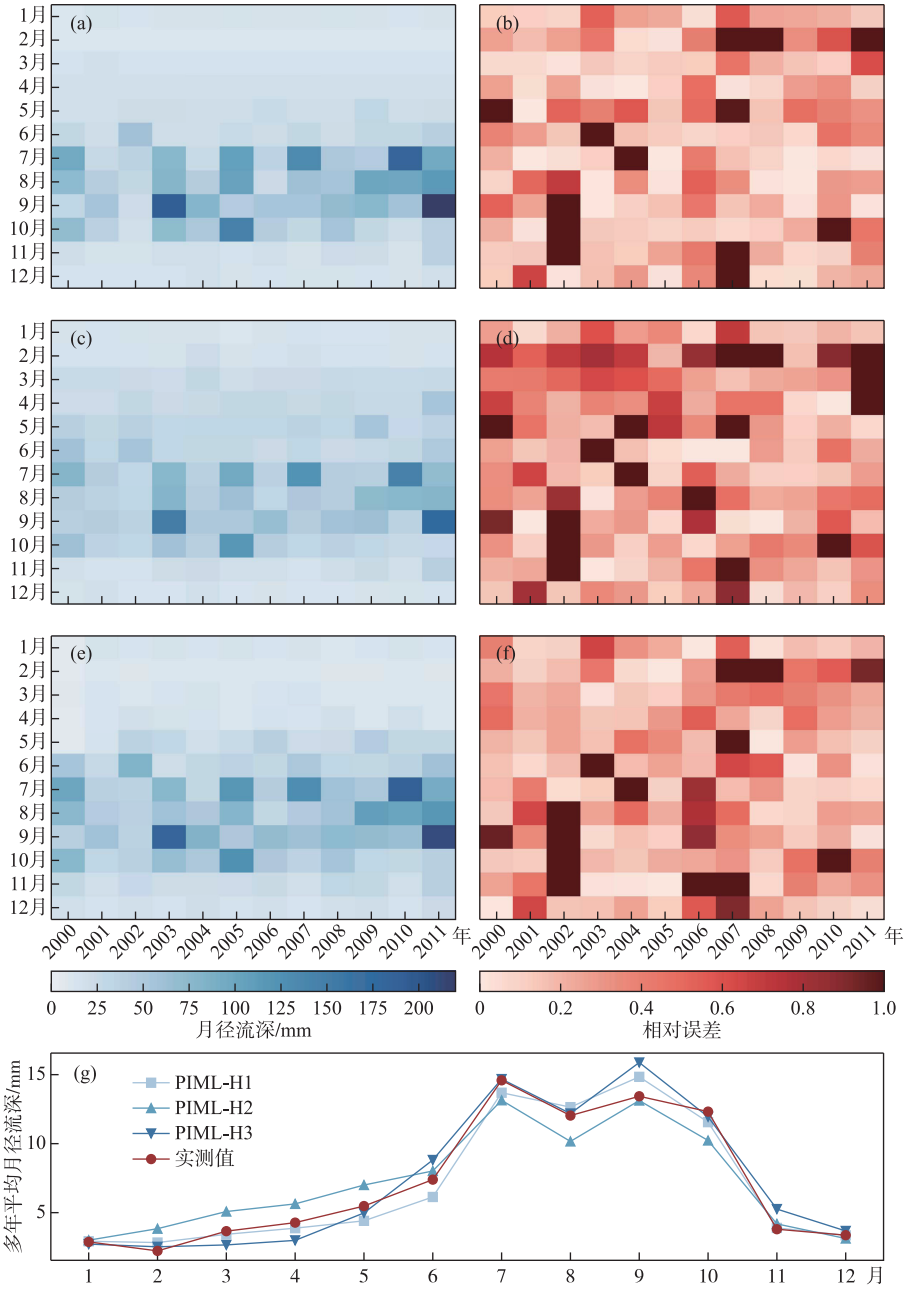


图 7 训练期和验证期 3 种 PIML 模型月径流深热图及多年平均月径流深折线图
(图 a~b、c~d、e~f 分别代表 PIML-H1、PIML-H2、PIML-H3 3 种模型的月径流深热图)

Fig.7 Thermal plots illustrating monthly runoff depth during the calibration and validation periods for three PIML models (PIML-H1(a-b), PIML-H2(c-d) and PIML-H3(e-f)) along with line graphs depicting the multi-year average monthly runoff depth)

明显,这与 PIML-H3 耦合过程中受到土壤含水量变化的影响有关[46]。在年份维度上,PIML-H1 模型模拟较好的年份主要分布在 2003－2011 年之间,而 PIML-H3 模型模拟较好的年份则集中在 2003－2005 年以及 2009－2011 年,但 3 种模型模拟效果并未呈现出明显的差异。图 7g 展示了 3 种 PIML 模型模拟与实测的多年平均月径流深对比,总体上讲,PIML-H1 模型模拟与实测的多年平均月径流深偏离程度最小,其次为 PIML-H3,PIML-H2 最差。具体而言,1－6 月期间,PIML-H1 与 PIML-H3 均在不同程度上低估了径流深,而 PIML-H2 则倾向于高估,6－8 月以及 10－11 月之间,PIML-H3 模拟值与实测值的平均差值最小,PIML-H2 模型平均差值最大,说明同时考虑中间变量和参数反馈因素在短时间内径流急剧变化下的径流模拟效果较好。

图 8 为 5 种模型在训练期和验证期各季节的径流模拟效果泰勒图,图中径向线反映线性相关系数,水平轴和垂直轴表示标准差,虚线则表示均方根标准差(*RMSD*)。本研究按照 3－5 月为春季、6－8 月为夏季、9—11 月为秋季、12 月—次年 2 月为冬季来划分四季[47]。总体而言,在春、夏季阶段,相较于 PIML-H2 和 LSTM 模型,GR4J、PIML-H1、PIML-H3 模型的相关性较高,但标准差和 *RMSD* 较低;3 种模型中,GR4J 模型呈现出最低的标准差和均方根标准差,而 PIML-H1 模型则展现出最大的线性相关系数。5 种模型在秋季的径流模拟效果较好,而在春季和冬季则有所下降。具体而言,除 LSTM、PIML-H2 模型外,其余 3 种模型在夏季的模拟效果均较春季有所提升,从夏季到秋季,5 种模型的线性相关系数均有所升高,且标准差下降,整体模拟效果提升。从秋季到冬季,5 种模型的标准差均有较大幅度增加,模型的模拟效果整体下降。

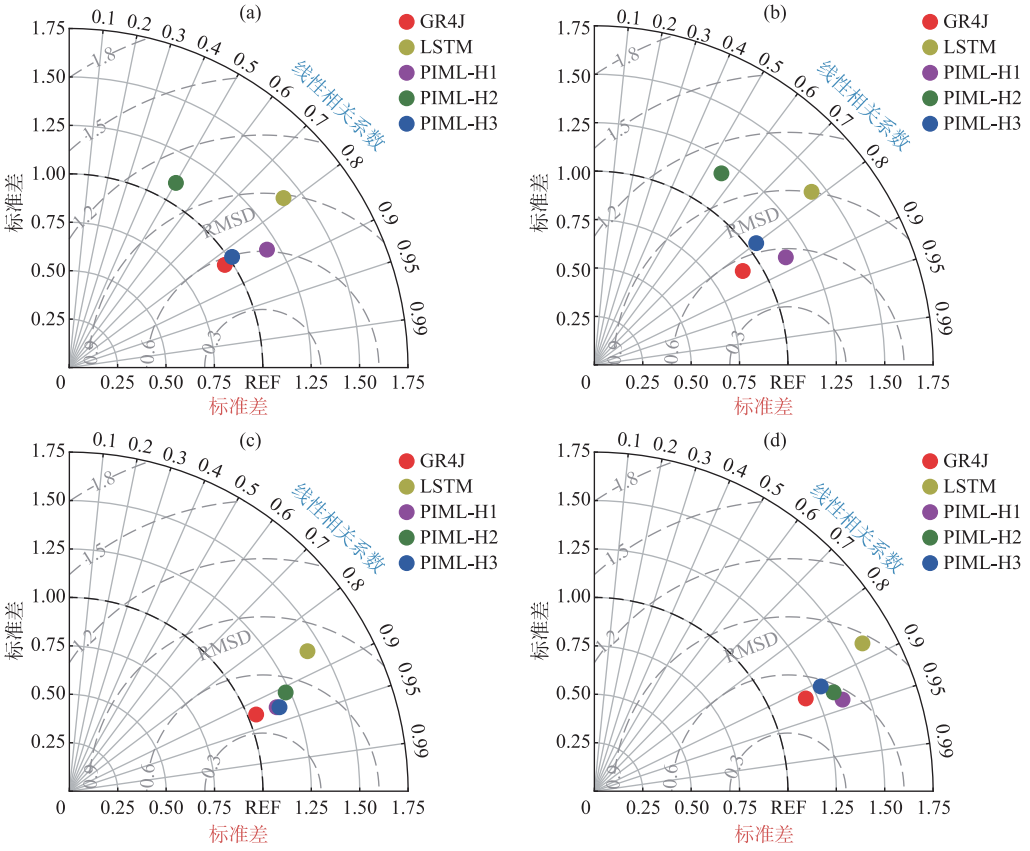


图 8 春(a)、夏(b)、秋(c)、冬(d)季训练期与验证期 5 种模型的径流模拟精度泰勒图

Fig.8 Performance of Taylor diagrams for five models during calibration and validation periods over spring (a), summer (b), autumn (c) and winter (d)

4 结论

本研究基于 GR4J 和 LSTM 模型构建了 PIML 模型,并分别设置了考虑参数反馈、考虑中间变量和两者同时考虑的 3 种方案(依次简称为 H1、H2、H3),探讨了 GR4J 与 LSTM 耦合过程中物理模型参数反馈机制以及考虑土壤含水量作为中间变量对 PIML 模型构建的影响,其主要结论如下:

1)3 种 PIML 模型径流模拟效果均优于 LSTM 模型,其中验证期平均 *NSE* 平均提升 10.6%,PIML-H1 与 PIML-H3 模型径流模拟效果优于 GR4J 模型,验证期 *NSE* 平均提升 4.2%,PIML-H3 模型在 3 种 PIML 模型中表现最佳,说明基于 LSTM 和 GR4J 模型构建的 PIML 模型对径流模拟效果有所改善,同时考虑中间变量和参数反馈因素对径流模拟效果改善最佳。

2)除 PIML-H1 模型外,其他 4 种模型均在一定程度上低估了流量,尤其在高水部分更为显著。在模拟低水流量时,3 种 PIML 模型的模拟效果均优于 GR4J 和 LSTM 模型,PIML-H3 模型效果最佳。而对于高水流量,3 种 PIML 模型均表现不佳,可见 PIML 模型往往在模拟低水流量事件中更占优势。

3)3 种 PIML 模型在训练和验证期间模拟结果均出现较强的季节性变化特征,并在 6－10 月均表现出较好的模拟效果。PIML-H2 和 PIML-H3 模型相较于 PIML-H1 模型受到的季节性变化影响更为明显,可见 PIML 模型模拟结果的季节性变化与耦合过程中土壤含水量变化有关。

总之,同时考虑参数反馈以及土壤含水量作为中间变量对 PIML 模型耦合的影响,在一定程度上能提升径流模拟效果。为了充分体现中间变量和参数反馈的必要性,在未来研究中可考虑更合适的中间变量以及参数反馈架构来提高径流模拟精度。另外,如何尽可能减小机器学习模型预测水文模型参数过程中的不确定性、探究不同中间变量对 PIML 耦合结果的影响,还有待深入研究。

5 参考文献

[1] Qiao PW, Wang S, Li JB *et al*. Process, influencing factors, and simulation of the lateral transport of heavy metals in surface runoff in a mining area driven by rainfall: A review. *Science of the Total Environment*, 2023, **857**(Pt1): 159119. DOI: 10.1016/j.scitotenv.2022.159119.

[2] Luo S, Xu YP, Wang Q *et al*. The evolution of annual maximum daily runoff and its influencing factors under the back-ground of urbanization: An example of Qinhuai River Basin in the lower reaches of the Yangtze River. *J Lake Sci*, 2023, **35**(6): 2123-2132. DOI: 10.18307/2023.0642.[罗爽, 许有鹏, 王强等. 城市化背景下年最大日径流演变及影响因素研究. 湖泊科学, 2023, **35**(6): 2123-2132.]

[3] Wang WG, Zou JC, Deng C. Comparison of data assimilation based approach for daily streamflow simulation under multiple scenarios in Ganjiang River Basin. *J Lake Sci*, 2023, **35**(3): 1047-1056. DOI: 10.18307/2023.0323.[王卫光, 邹佳成, 邓超. 赣江流域多种数据同化方案的径流模拟比较. 湖泊科学, 2023, **35**(3): 1047-1056.]

[4] Fan HX, He HD, Xu LG *et al*. Simulation and attribution analysis based on the long-short-term-memory network for detecting the dominant cause of runoff variation in the Lake Poyang Basin. *J Lake Sci*, 2021, **33**(3): 866-878. DOI: 10.18307/2021.0319.[范宏翔, 何菡丹, 徐力刚等. 基于长短记忆模型的鄱阳湖流域径流模拟及其演变的归因分析. 湖泊科学, 2021, **33**(3): 866-878.]

[5] Zhao LX, Guo ZL, Nie XF *et al*. Effects of extreme rainfall events on runoff and sediment in the southern red soil area: A long series analysis based on the Upper Lianjiang River of Ganjiang River (1984-2020). *J Lake Sci*, 2023, **35**(6): 2133-2143. DOI: 10.18307/2023.0643.[赵利祥, 郭忠录, 聂小飞等. 极端降雨事件对南方红壤区径流和输沙的影响——基于赣江支流濂江上游流域的长序列分析(1984－2020 年). 湖泊科学, 2023, **35**(6): 2133-2143.]

[6] Guan TS, Bao ZX, He RM *et al*. Uncertainties of model parameters regionalization in ungauged basins. *Advances in Water Science*, 2023, **34**(5): 660-672.[关铁生, 鲍振鑫, 贺瑞敏等. 无资料地区水文模型参数移植不确定性分析. 水科学进展, 2023, **34**(5): 660-672.]

[7] Devia GK, Ganasri BP, Dwarakish GS. A review on hydrological models. *Aquatic Procedia*, 2015, **4**: 1001-1007. DOI: 10.1016/j.aqpro.2015.02.126.

[8] Yaseen ZM, El-shafie A, Jaafar O *et al*. Artificial intelligence based models for stream-flow forecasting: 2000-2015. *Journal of Hydrology*, 2015, **530**: 829-844. DOI: 10.1016/j.jhydrol.2015.10.038.

[9] Lin GF, Chou YC, Wu MC. Typhoon flood forecasting using integrated two-stage Support Vector Machine approach. *Journal of Hydrology*, 2013, **486**: 334-342. DOI: 10.1016/j.jhydrol.2013.02.012.

[10] Xiang ZR, Yan J, Demir I. A rainfall-runoff model with LSTM-based sequence-to-sequence learning. *Water Resources Research*, 2020, **56**(1): e2019WR025326. DOI: 10.1029/2019WR025326.

[11] Kratzert F, Klotz D, Brenner C *et al*. Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth*

System Sciences, 2018, **22**(11): 6005-6022. DOI: 10.5194/hess-22-6005-2018.

[12] Gao S, Huang YF, Zhang S *et al*. Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation. *Journal of Hydrology*, 2020, **589**: 125188. DOI: 10.1016/j.jhydrol.2020.125188.

[13] Beven K. Deep learning, hydrological processes and the uniqueness of place. *Hydrological Processes*, 2020, **34**(16): 3608-3613. DOI: 10.1002/hyp.13805.

[14] Gilpin LH, Bau D, Yuan BZ *et al*. Explaining explanations: An overview of interpretability of machine learning. 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). Turin, Italy. IEEE, 2018: 80-89. DOI: 10.1109/DSAA.2018.00018.

[15] Yang SY, Yang DW, Chen JS *et al*. A physical process and machine learning combined hydrological model for daily streamflow simulations of large watersheds with limited observation data. *Journal of Hydrology*, 2020, **590**: 125206. DOI: 10.1016/j.jhydrol.2020.125206.

[16] Nourani V. An Emotional ANN (EANN) approach to modeling rainfall-runoff process. *Journal of Hydrology*, 2017, **544**: 267-277. DOI: 10.1016/j.jhydrol.2016.11.033.

[17] Jiang SJ, Zheng Y, Solomatine D. Improving AI system awareness of geoscience knowledge: Symbiotic integration of physical approaches and deep learning. *Geophysical Research Letters*, 2020, **47**(13): e2020GL088229. DOI: 10.1029/2020GL088229.

[18] Kashinath K, Mustafa M, Albert A *et al*. Physics-informed machine learning: Case studies for weather and climate modelling. *Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences*, 2021, **379**(2194): 20200093. DOI: 10.1098/rsta.2020.0093.

[19] Humphrey GB, Gibbs MS, Dandy GC *et al*. A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network. *Journal of Hydrology*, 2016, **540**: 623-640. DOI: 10.1016/j.jhydrol.2016.06.026.

[20] Okkan U, Ersoy ZB, Ali Kumanlioglu A *et al*. Embedding machine learning techniques into a conceptual model to improve monthly runoff simulation: A nested hybrid rainfall-runoff modeling. *Journal of Hydrology*, 2021, **598**: 126433. DOI: 10.1016/j.jhydrol.2021.126433.

[21] Badrzadeh H, Sarukkalige R, Jayawardena AW. Hourly runoff forecasting for flood risk management: Application of various computational intelligence models. *Journal of Hydrology*, 2015, **529**: 1633-1643. DOI: 10.1016/j.jhydrol.2015.07.057.

[22] Yu Q, Jiang LG, Wang YJ *et al*. Enhancing streamflow simulation using hybridized machine learning models in a semi-arid basin of the Chinese loess Plateau. *Journal of Hydrology*, 2023, **617**: 129115. DOI: 10.1016/j.jhydrol.2023.129115.

[23] Bian LK, Qin XE, Zhang CL *et al*. Application, interpretability and prediction of machine learning method combined with LSTM and Light-GBM-A case study for runoff simulation in an arid area. *Journal of Hydrology*, 2023, **625**: 130091. DOI: 10.1016/j.jhydrol.2023.130091.

[24] Deng C, Chen CY, Yin X *et al*. Catchment runoff simulation by coupling data assimilation and machine learning methods. *Advances in Water Science*, 2023, **34**(6): 839-849. [邓超, 陈春宇, 尹鑫等. 融合数据同化与机器学习的流域径流模拟方法. 水科学进展, 2023, **34**(6): 839-849.]

[25] Bhasme P, Vagadiya J, Bhatia U. Enhancing predictive skills in physically-consistent way: Physics Informed Machine Learning for hydrological processes. *Journal of Hydrology*, 2022, **615**: 128618. DOI: 10.1016/j.jhydrol.2022.128618.

[26] Pan X, Qin CC. Study on temporal and spatial evolution law of precipitation in Hanjiang River Basin. *Express Water Resources & Hydropower Information*, 2023, (6): 333-343. [潘鑫, 秦创创. 汉江流域降水时空演变特征研究. 水利水电快报, 2023, (6): 333-343.]

[27] Wang D, Wu DD, Xie XB *et al*. Study on spatio-temporal variation of runoff in flood season in Hanjiang River Basin. *Pearl River*, 2020, **41**(4): 30-39. [王栋, 吴栋栋, 解效白等. 汉江流域汛期径流时空变化特征研究. 人民珠江, 2020, **41**(4): 30-39.]

[28] Liu YW, Li JK, Ding Q *et al*. Characteristic analysis and prediction of rainfall and runoff in Ankang section of Hanjiang River Basin. *Pearl River*, 2021, **42**(6): 59-69. [刘易文, 李家科, 丁强等. 汉江流域安康段降雨径流特征分析及预测. 人民珠江, 2021, **42**(6): 59-69.]

[29] Allen RG, Pereira LS, Raes D *et al*. Crop evapotranspiration: Guidelines for computing crop requirements. FAO Irrigation and Drainage Paper, 1998: 56.

[30] Perrin C, Michel C, Andréassian V. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 2003, **279**(1/2/3/4): 275-289. DOI: 10.1016/s0022-1694(03)00225-7.

[31] Kunnath-Poovakka A, Eldho TI. A comparative study of conceptual rainfall-runoff models GR4J, AWBM and Sacramento at catchments in the upper Godavari River Basin, India. *Journal of Earth System Science*, 2019, **128**(2): 33. DOI: 10.1007/s12040-018-1055-8.

[32] Wang HY, Guo SL, Wang J *et al*. Comparative study of GR4J and Xinanjiang models in humid area. *Journal of Water Resources Research*, 2021, (1): 11-20. [王何予, 郭生练, 王俊等. GR4J 与新安江模型在湿润地区比较研究. 水资源研究, 2021, (1): 11-20.]

[33] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, **9**(8): 1735-1780. DOI: 10.1162/neco.1997.9.8.1735.

[34] Sherstinsky A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 2020, **404**: 132306. DOI: 10.1016/j.physd.2019.132306.

[35] Alhirmizy S, Qader B. Multivariate Time Series Forecasting with LSTM for Madrid, Spain pollution. 2019 International Conference on Computing and Information Science and Technology and Their Applications (ICCISTA). Kirkuk, Iraq. IEEE, 2019: 1-5. DOI: 10.1109/ICCISTA.2019.8830667.

[36] Berkenkamp F, Krause A, Schoellig AP. Bayesian optimization with safety constraints: Safe and automatic parameter tuning in robotics. *Machine Learning*, 2023, **112**(10): 3713-3747. DOI: 10.1007/s10994-021-06019-1.

[37] Joy TT, Rana ST, Gupta S *et al*. Fast hyperparameter tuning using Bayesian optimization with directional derivatives. *Knowledge-Based Systems*, 2020, **205**: 106247. DOI: 10.1016/j.knosys.2020.106247.

[38] Deng PX, Wang YT, Hu QF *et al*. Application of GR4J in daily runoff simulation for Ganjiang River Basin. *Journal of China Hydrology*, 2014, **34**(2): 60-65. [邓鹏鑫, 王银堂, 胡庆芳等. GR4J 模型在赣江流域日径流模拟中的应用. 水文, 2014, **34**(2): 60-65.]

[39] Wei XJ, Guo SL, Xiong LH. Improving efficiency of hydrological prediction based on meteorological classification: A case study of GR4J model. *Water*, 2021, **13**(18): 2546. DOI: 10.3390/w13182546.

[40] Habtemariam ET, Kekeba K, Martínez-Ballesteros M *et al*. A Bayesian optimization-based LSTM model for wind power forecasting in the Adama district, Ethiopia. *Energies*, 2023, **16**(5): 2317. DOI: 10.3390/en16052317.

[41] Skomski E, Lee JY, Kim W *et al*. Sequence-to-sequence neural networks for short-term electrical load forecasting in commercial office buildings. *Energy and Buildings*, 2020, **226**: 110350. DOI: 10.1016/j.enbuild.2020.110350.

[42] Wang S, Peng H, Hu Q *et al*. Analysis of runoff generation driving factors based on hydrological model and interpretable machine learning method. *Journal of Hydrology: Regional Studies*, 2022, **42**: 101139. DOI: 10.1016/j.ejrh.2022.101139.

[43] Prendin F, Pavan J, Cappon G *et al*. The importance of interpreting machine learning models for blood glucose prediction in diabetes: An analysis using SHAP. *Scientific Reports*, 2023, **13**: 16865. DOI: 10.1038/s41598-023-44155-x.

[44] Chollet Ramampiandra E, Scheidegger A, Wydler J *et al*. A comparison of machine learning and statistical species distribution models: Quantifying overfitting supports model interpretation. *Ecological Modelling*, 2023, **481**: 110353. DOI: 10.1016/j.ecolmodel.2023.110353.

[45] Grillakis MG, Koutroulis AG, Komma J *et al*. Initial soil moisture effects on flash flood generation—A comparison between basins of contrasting hydro-climatic conditions. *Journal of Hydrology*, 2016, **541**: 206-217. DOI: 10.1016/j.jhydrol.2016.03.007.

[46] Wang S, Fu BJ, Gao GY *et al*. Soil moisture and evapotranspiration of different land cover types in the Loess Plateau, China. *Hydrology and Earth System Sciences*, 2012, **16**(8): 2883-2892. DOI: 10.5194/hess-16-2883-2012.

[47] Xiong Y, Zhou JZ, Sun N *et al*. Monthly runoff prediction based on self-adaptive variational mode decomposition and long short-term memory network. *Journal of Hydraulic Engineering*, 2023, **54**(2): 172-183, 198. [熊怡, 周建中, 孙娜等. 基于自适应变分模态分解和长短期记忆网络的月径流预报. 水利学报, 2023, **54**(2): 172-183, 198.]