

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^{*}

范宏翔^{1,2}, 何菡丹³, 徐力刚^{1,4**}, 张明睿^{1,5}, 姜加虎¹

(1:中国科学院南京地理与湖泊研究所,中国科学院流域地理学重点实验室,南京 210008)

(2:中国科学院大学,北京 100049)

(3:江苏省水资源服务中心,南京 210029)

(4:中国长江三峡集团有限公司长江生态环境工程研究中心,北京 100038)

(5:安徽工业大学,马鞍山 243002)

摘要: 气候变化和人类活动直接或间接的影响着全球和区域水文循环过程,是导致水文水资源时空分布的主要因素,同时也是流域—湖泊水文情势变化的根本原因。本文基于长短记忆模型构建了鄱阳湖气象—径流模型,同时引入了基准期的概念,定量区分了导致鄱阳湖流域径流变化的主要影响因素。研究结果表明:在同时考虑计算效率和模拟效果的前提下,采用 10 d 预测窗口大小来构建鄱阳湖气象—径流模型能够很好地捕捉径流的极值,并且对径流的短期波动也能有很好的体现。训练期模型在各个子流域的纳什效率系数均高于 0.94,而在验证期,模型在各个子流域的纳什效率系数均高于 0.90。基于径流模拟结果,定量区分了人类活动和气候变化对鄱阳湖径流的影响,研究结果显示:人类活动对径流的影响主要发生在春、秋季,其中,人类活动在春季主要会造成径流的增加,平均增加幅度约为 $139.47 \text{ m}^3/\text{s}$,而在秋、冬季,人类活动则会导致径流平均减少约 $34.37 \text{ m}^3/\text{s}$ 。对比二者的相对贡献率,可以发现,春季人类活动对径流造成的影响较大,平均相对贡献率为 77.26%。而在其余季节,鄱阳湖流域径流过程的改变主要是由于气候变化,平均相对贡献率约 75.84%。研究结果能够为鄱阳湖流域水资源管理提供科学依据和理论指导。

关键词: 长短记忆模型;气候变化;人类活动;鄱阳湖;径流变化

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^{*}

Fan Hongxiang^{1,2}, He Handan³, Xu Ligang^{1,4**}, Zhang Mingrui^{1,5} & Jiang Jiahu¹

(1: Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, P.R.China)

(2: University of Chinese Academy of Sciences, Beijing 100049, P.R.China)

(3: Water Resources Service Center of Jiangsu Province, Nanjing 210029, P.R.China)

(4: Eco-Environmental Engineering Research Center, China Three Gorges Corporation, Beijing 100038, P.R.China)

(5: Anhui University of Technology, Maanshan 243002, P.R.China)

Abstract: Climate change and human activities directly or indirectly affect the global and regional hydrologic cycle, which is the main factor leading to the temporal and spatial distribution of water resources. Based on the long-short-term-memory (LSTM) model, this paper attributes the contribution of climate change and human activities on runoff variation in the Lake Poyang Basin. Results show that the runoff process was simulated well by LSTM framework with a window size of 10 days for both accuracy and computational efficiency. The overall performance of the model was good with Nash-Sutcliffe Efficiency varied from 0.94 to 0.95 in the

* 2020-07-01 收稿;2020-09-07 收修改稿。

国家重点研发计划项目(2018YFE0206400,2018YFC0407606)、国家自然科学基金项目(41971137,41771235)、青海省科技支撑项目(2019-HZ-818)、中国科学院STS项目(KFJ-STS-QYZD-098)、中国长江三峡集团有限公司项目(201903145)、中国科学院南京地理与湖泊研究所引进人才启动项目(NIGLAS2019QD005)和中国科学院流域地理学重点实验室开放基金项目(WSGS2020003)联合资助。

** 通信作者;E-mail: lgxu@niglas.ac.cn.

training period and 0.90 to 0.98 in the test period, respectively. Based on the simulation results, the contribution of human activities and climate change on runoff change in Lake Poyang Basin was then quantified. Results show that the runoff increased by 139.47 m³/s due to human activities in spring, accounting for 77.26% of runoff change. Whereas climate change was the dominant factor in changing runoff in the other seasons, which led the runoff to decrease by 34.37 m³/s, accounting for about 75.84% of runoff changes. The results can provide scientific basis and theoretical guidance for water resources management in Lake Poyang Basin.

Keywords: Long-short-term-memory network; climate change; human activities; Lake Poyang; runoff

气候变化和人类活动直接或间接地影响着全球和区域水文循环过程,是导致水文水资源时空分布的主要因素,同时也是湖泊一流域水文情势变化的主要原因。近年来,全球气候发生了巨大的变化,其中很多变化是数十上百年来前所未见的^[1],因此,研究气候变化作用下区域水循环的变化对人类社会、生态系统和区域水资源管理具有十分重要的意义^[2]。上个世纪,特别是1980s以后,全球变暖显著改变了全球和区域的水循环过程。在全球尺度上有研究表明,北半球的平均地表温度在1901—2010年间每年增加了0.010°C,而在1979—2010年间其增长速率达到0.026°C/a^[3]。气温变化和降雨变化是目前研究气候变化对水文情势影响的研究重点,如Labat等^[4]就发现陆地表面温度每上升1°C就会导致地表径流量增加5%。而在区域尺度上,Zhang等^[5]在长江流域的研究表明气候变化是1931—1960年长江流域径流变化的主要影响因素,其导致的径流增长占总体径流变化的67.10%~78.70%。人类活动则通过毁林、造林、开垦和城市化等土地利用方式的转变,影响水文循环过程的各个环节。目前对人类活动的研究主要可以分为两类:(1)土地利用变化,如Guo等^[6]利用SWAT模型研究了土地利用变化对鄱阳湖流域径流变化的影响,结果表明土地利用变化会显著影响季节性的径流变化并改变年内的径流过程。(2)水利工程,如Nilsson等^[7]指出全球大部分的大型河流水文情势受到大坝建设的影响;而Agostinho等^[8]也指出,水利工程建设是巴西Paraná流域水文改变变化的主要驱动因素;同样地,Arias等^[9]研究也发现湄公河支流上的大坝是柬埔寨TonleSap洪泛平原水文改变的主要原因,类似的研究还有Arias等^[9]、Ty等^[10]和Cohen等^[11]。

关于气候变化和人类活动对水文情势演变归因分析的研究方法目前主要有两类^[12]:(1)通过改变水文模型的输入,以观察由此产生的径流变化;(2)水热耦合平衡方法。Githui等^[13]、Zhang等^[14]和Narsimlu等^[15]利用SWAT模型分别对不同流域分区的径流对气候变化或者人类活动的响应关系进行了深入研究,取得了丰硕的研究成果。采用模型的手段虽然具有一定的物理意义,但是也有诸多问题,比如数据时空转换的问题,模型参数率定的不确定性等。同时构建模型时,还需要大量的观测数据资料,这些都制约着水文模型的大规模使用。相较而言,水热耦合平衡方法在减少模型的不确定性方面具有一定的优势,可以有效区分气候变化和人类活动对流域径流过程的影响。Ye等^[16]利用水热耦合公式定量区分了气候变化和人类活动在不同年代对鄱阳湖径流变化的影响,其结果表明气候变化是鄱阳湖流域径流变化的主要因素,其对径流增加的贡献率为105%~212%,而人类活动主要导致径流量的减少,其贡献率为-5%~112%。但是,水热耦合平衡方法的稳态假设限制了其应用的时间尺度^[17],以往研究大多在长时间尺度上(10~30 a)应用水热耦合平衡方法^[18-20],因此其无法对精细尺度上的径流变化影响因素进行区分。为了解决上述研究难点,亟需一种简单高效的手段来对流域径流过程进行精细尺度的模拟,从而定量区分导致其变化的主要影响因素。近年来,有许多研究都证明了数据驱动型方法(如多元线性回归、支持向量机以及人工神经网络)在水文领域的适用性^[21],并且也讨论了其与传统水文模型、方法之间的对比。研究表明,以上机器学习方法在水位预测^[21]、地下水模拟^[22]、土壤水分^[23]反演以及径流模拟^[24]等方面都取得了较传统方法一致甚至更好的结果。然而,传统的机器学习方法,如人工神经网络,并不是专门为处理时间序列数据而设计的,不能直接处理时间序列间依赖性。而长短记忆模型的提出,解决了时间序列的长短期依赖问题,大大提高了时间序列预测的精度,其在水文领域的应用也越来越广泛^[25-28]。利用深度学习网络构建流域径流预测模型,并通过与历史时期的进行对比,可以在更加精细的时间尺度上区分影响流域径流变化的主要因素。值得注意的是,长短记忆模型的一个重要参数是模拟窗口的大小,在以往的研究中,往往会给定一个经验值^[24],忽视窗口大小对模拟结果的影响,因此,针对流域实际情况,确定合适的模拟窗口大小,也是提高模拟精度的重要手段之一。

综上所述,本文旨在(1)基于利用长短记忆网络,构建鄱阳湖流域气象—径流模型,通过设置不同的模拟窗口大小来确定适合鄱阳湖流域的最佳模拟窗口长度;(2)通过引入基准期的概念,利用构建的鄱阳湖流

域气象—径流模型模拟自然条件下流域径流的变化过程;(3)通过与实际观测数据对比,定量区分鄱阳湖流域径流改变的主要影响因素.研究结果能够为鄱阳湖流域水资源管理提供技术支持,同时为定量区分人类活动和气候变化对流域径流变化的影响提供一种新思路.

1 材料与方法

1.1 数据

1.1.1 气象数据 根据资料搜集情况,选取鄱阳湖流域时间序列最长的14个国家级气象站(图1)的日气象数据;其中降水数据时段为1955—2015年,其他气象数据(包括气温、气压、湿度、日照时数、风速、饱和水汽压等)时段为1960—2015年.

1.1.2 水文数据 鄱阳湖流域由赣江流域、抚河流域、信江流域、饶河流域以及修水流域5个子流域组成,鄱阳湖入湖水量主要来自五大河,其中大部分来自赣江(55.0%),其次是信江(14.4%)和抚河(12.1%),三者之和超过总入湖径流量的80%^[20].选取5个子流域下游出口断面流量监测站(图1)月平均径流数据来分析鄱阳湖流域水文时间序列所蕴含的趋势与规律.

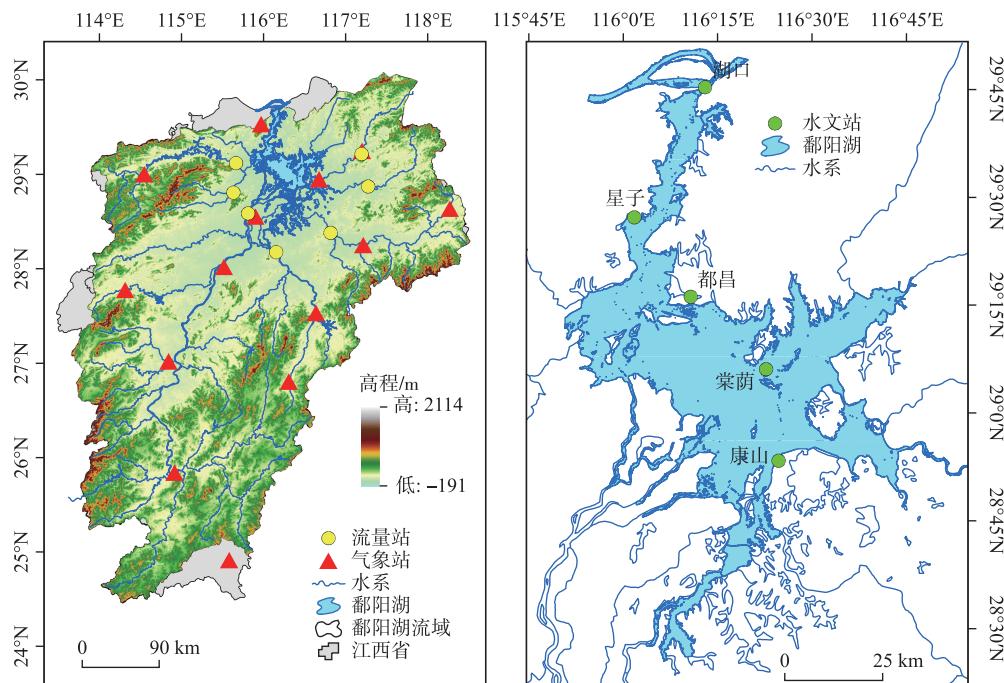


图1 鄱阳湖流域地理位置及气象站、水文站分布

Fig.1 Location of Lake Poyang Basin and the meteorological and hydrological stations inside the basin

本文所用到的水文站基本情况如表1所示.

1.2 研究方法

1.2.1 长短记忆模型原理 长短记忆模型(long-short-term-memory)是一种特殊的循环神经网络(recurrent neural network)结构,旨在解决传统循环神经网络在学习长期依赖问题上的不足. LSTM最早由Hochreiter等^[29]提出,并在最近被Graves等^[30]推广和改进. LSTM在时间维度上具有深层结构,通过巧妙设计的门和记忆单元(图2),LSTM能够决定在何时遗忘状态信息和将状态信息保存多久.在很多问题上,LSTM都取得了很大的成功,并得到了广泛的应用^[23-24,31-34].

LSTM的关键就是细胞状态,细胞状态类似于传送带.直接在整个链上运行,只有一些少量的线性交互,使得信息能够保持不变.细胞状态由3个巧妙设计的门进行控制.在t时刻,记忆块(当前神经元)的输入包

表 1 鄱阳湖流域主要水文站基本信息

Tab.1 List of main hydrological stations in Lake Poyang Basin

水文站	所在水系	集水面积/($\times 10^4 \text{ km}^2$)	资料序列
外洲	赣江	8.09	1960—2015 年
李家渡	抚河	1.58	1960—2015 年
梅港	信江	1.55	1960—2015 年
虎山	饶河—昌江	0.64	1960—2015 年
渡峰坑	饶河—昌江	0.50	1978—2015 年
万家埠	修水—潦河	0.35	1960—2015 年
虬津	修水	0.99	1978—2015 年

括:当前时刻输入变量 X_t 、前一时刻隐藏层状态变量 h_{t-1} 和前一时刻记忆单元状态变量 C_{t-1} ,之后模型依次通过遗忘门 f_t 、输入门 i_t 、输出门 o_t ,记忆块的输出包括:当前时刻输出变量 h_t 和当前时刻记忆单元状态变量 C_t .

1.2.2 模型基本设定 在本文中,基于长短记忆模型框架构建了一个神经网络模型来模拟鄱阳湖流域的径流过程. 模型框架如图 3 所示. 构建的模型有一个输入层,共包含 128 个 LSTM 神经元,用来输入流域的气象参数. 在此模型中,所用的气象数据包括降雨、气温、太阳辐射、相对湿度和近地面风速. 另外,为了防止模型出现过拟合的现象,在输入层之后设置了一个丢弃层 (dropout layer),并设置其丢弃率 (dropout rate) 为 0.4^[35],即神经元有 60% 的概率被保存. 之后,建立了一个输出层,用来输出 5 个独立的径流时间序列(分别代表赣江、抚河、信江、饶河和修水的径流过程). 前人的研究大多采用降雨作为径流模型的驱动变量,但是其余气象要素如气温、相对湿度、太阳辐射等也会对径流过程产生影响,因此,通过输入多个气象要素(降雨、相对湿度、气温、太阳辐射、日照时间和风速)训练了鄱阳湖流域气象—径流模型. 在前人的研究中,多将 1960—1969 年作为基准期,即该时期的人类活动较少,可以反映自然条件下鄱阳湖流域的径流特征^[16,19,36]. 同时,考虑到 LSTM 模型的特殊性,基准期选取时也考虑了不同的流域来水情景. 根据水文情报预报规范 (GB/T 22482—2008),按照鄱阳湖流域年最大流量进行水文年的划分. 其中,1962 年为典型丰水年,1963 年为典型枯水年,1966 年和 1968 年为典型平水年. 因此,为了进一步区分人类活动与气候变化对鄱阳湖径流的影响,在模型训练和验证时仅使用基准期(即 1960—1969 年)的气象水文数据. 基于此,将基准期划分为训练期和验证期两个阶段:模型的训练周期为 1960 年 1 月 1 日—1968 年 12 月 31 日,共计 3287 d. 模型的验证周期为 1969 年 1 月 1 日—1969 年 12 月 31 日,共计 365 d.

1.2.3 模型预测窗口设定 由于长短记忆模型在本质上是一种时间序列预测模型,其本身需要先前的时间序列作为输入变量. 先前时间序列的长度称为预测窗口的长度或大小,其取值对模型最终的模拟效果至关重要. 因此,在本章中,分析了不同窗口大小对模拟效果的影响,并尝试找出模拟的最优窗口大小. 考虑到计算效率和模拟时间,设置了 9 组窗口大小,分别是 1、5、10、15、25、30、60、90 和 180 d. 模拟效果最好的窗口被作为最终的模拟窗口大小.

1.2.4 超参数率定 神经网络模型往往包含大量的超参数(hyper-parameters). 超参数是在开始学习过程之前设置值的参数,而不是通过训练得到的参数数据. 通常情况下,需要对超参数进行优化,给学习机选择一组最优超参数,以提高学习的性能和效果. 超参数优化或模型选择是为学习算法选择一组最优超参数时的问题,通常目的是优化算法在独立数据集上的性能的度量. 通常使用交叉验证来估计这种泛化性能^[37]. 常用

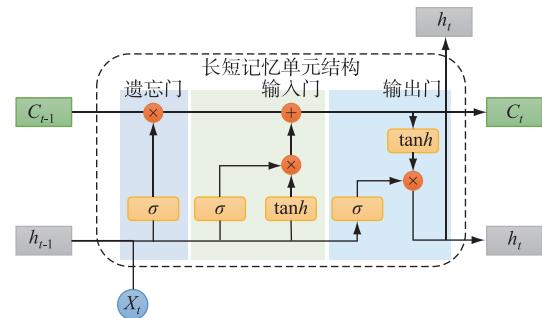


图 2 长短记忆模型记忆单元结构示意图

Fig.2 The architecture of LSTM memory cell

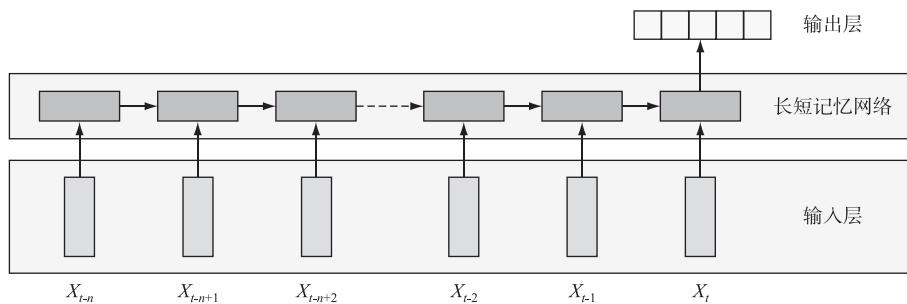


图3 长短记忆模型结构示意图

Fig.3 Structure of the LSTM model

的率定方法有网格搜索^[38]、贝叶斯优化^[39]、随机搜索^[40]和基于梯度的优化^[41]等。在本文中,使用均方根误差函数作为损失函数来进行超参数的优化。

常见的超参数包括学习率、训练次数、输出维度等。其中,学习率表示梯度下降算法中的步长(本章使用随机梯度下降算法),在鄱阳湖径流模型中,将初始学习率设置为0.2并设置了相应的衰减速率来更新每次训练过程中的学习率。另外,训练次数是指神经网络学习一次全部数据的过程,其作用是将训练过程分解为独立的不同阶段。训练时间过长会导致模型出现过拟合的现象,即会让模型学习到只有训练数据中才含有的特征。反之,训练时间过短则会出现欠拟合的现象,让模型无法充分学习数据中的特征。在本章中,首先设置了一个较长的训练次数:200,随后,发现最佳的拟合结果出现在训练50次后。因此将后续的训练次数都设置为50。其他的一些超参数如输出维度等则通过网格搜索方式进行优化。模型主要参数如表2所示。

表2 模型主要参数

Tab.2 Model parameters

参数类型	参数名称	调参方式	本文取值
全局超参数	隐藏层层数(layers)	经验调参	2
	单个隐藏层中神经元个数(hidden size)		128
	训练数据包大小(batch size)		500
	预测窗口大小(window size)		10
	丢弃率(dropout rate)		0.4
训练时超参数	学习率(learning rate)		0.2
	迭代次数(epoch)		50
普通参数	输入全连接层参数(\$w_{in}, b_{in}\$)	自动调参	模型自动优化
	循环体隐藏层参数(\$w_{mn}, b_{mn}\$)		
	输出全连接层参数(\$w_{out}, b_{out}\$)		

1.2.5 模型评价指标 模型的模拟效果采用常规的统计指标来量化评价,在本章中主要采用的指标为纳什效率系数(NSE)和均方根误差(RMSE),其计算公式分别为:

$$NSE = 1 - \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - y'_i)^2}{n}} \quad (2)$$

式中,\$y'\$为模拟序列,\$y\$为观测序列,\$\bar{y}\$和\$\bar{y}'\$分别为观测序列和模拟序列的均值,\$n\$为序列长度。

1.2.6 情景设置 本文通过引入基准期的概念,应用 LSTM 模型定量区分气候变化和人类活动对鄱阳湖径流变化的贡献率。为了达成此目标,本文拟设置 3 个情景,具体如表 3 所示。通过对 S1 和 S0,可以定量区分气候变化对鄱阳湖流域径流过程的影响分量;而通过对 S2 和 S1,则可以定量区分人类活动对鄱阳湖流域径流变化的影响贡献。由于不同情景方案模拟结果误差高度相似,因此基于情景方案的数值模拟结果相比可消除模型系统误差的影响^[42]。

表 3 情景设置
Tab.3 Model scenarios

情景	流量过程
S0	1960—1969 年鄱阳湖流域五河多年平均日流量过程
S1	气象—径流模型模拟的 1970—2015 年鄱阳湖流域五河多年平均日流量过程
S2	实测的 1970—2015 年鄱阳湖流域五河多年平均日流量过程

1.2.7 影响程度计算方法 参照 Ye 等^[16]关于定量区分气候变化和人类活动对鄱阳湖流域径流量影响程度的分析方法,由气候变化和人类活动引起的鄱阳湖流域径流变化量可以分别用下式来解释:

$$\begin{cases} \Delta Q_c = Q_1 - Q_0 \\ \Delta Q_h = Q_2 - Q_1 \end{cases} \quad (3)$$

式中, Q_0 、 Q_1 和 Q_2 分别代表 S0、S1 以及 S2 情景下鄱阳湖流域五河多年平均日流量过程, ΔQ_c 和 ΔQ_h 分别代表气候变化和人类活动引起的鄱阳湖流域径流变化的绝对分量。

气候变化以及人类活动对鄱阳湖流域径流变化的影响程度计算方法为:

$$\begin{cases} \eta_c = \frac{|\Delta Q_c|}{|\Delta Q_c| + |\Delta Q_h|} \times 100\% \\ \eta_h = \frac{|\Delta Q_h|}{|\Delta Q_c| + |\Delta Q_h|} \times 100\% \end{cases} \quad (4)$$

式中, η_c 和 η_h 分别代表气候变化和人类活动对鄱阳湖径流变化的相对影响程度。

2 研究结果

2.1 不同预测窗口下模型的预测精度

为了研究不同预测窗口大小对模型预测结果的影响,用 9 组窗口大小训练了 9 个独立的 LSTM 模型。模型的表现用 RMSE 和 NSE 来衡量,RMSE 越低则说明模拟值与真实值之间的差距越小,而 NSE 越高则说明模拟值与真实情况越接近。不同窗口大小下的模拟结果如图 4 所示。从图 4 中可以发现,在预测窗口增大的初期,RMSE 迅速下降,并在预测窗口为 10 d 时达到最小值。而当预测窗口继续增大至 60 d 时,RMSE 出现了明显的增大。这有可能意味着模型出现了过拟合现象。而当预测窗口超过 60 d 后,窗口大小的继续增加对模型效果影响并不大。类似的变化特征在纳什效率系数上也得到了相应的体现。当预测窗口小于 15 d 时,窗口越大 NSE 越高,模拟效果越好;当窗口大小超过 15 d 时,窗口大小增加会导致 NSE 迅速下降。而当窗口超过 60 d 时,窗口大小的增加对模拟效果的影响会减弱。RMSE 和 NSE 的变化趋势都说明了预测窗口大小并不是越大越好,在同时考虑计算效率和模拟效果的前提下,在鄱阳湖径流模型中使用 10 d 的预测窗口大小较为合理。

2.2 模型精度评价

鄱阳湖气象—径流模型的总体表现如表 4 所示。训练期,模型在赣江、抚河、信江、饶河和修水 5 个子流域内的 RMSE 分别为 630.80、143.64、194.74、129.46 和 144.01 m³/s;而在验证期,模型在赣江、抚河、信江、饶河和修水 5 个子流域内的 RMSE 分别为 671.03、158.90、212.04、139.52 和 162.79 m³/s。训练期模型在赣江、抚河、信江、饶河和修水 5 个子流域的 NSE 分别为 0.94、0.95、0.95、0.95 和 0.94;而在验证期,模型在赣江、抚河、信江、饶河和修水 5 个子流域的 NSE 分别为 0.90、0.95、0.95、0.98 和 0.96。

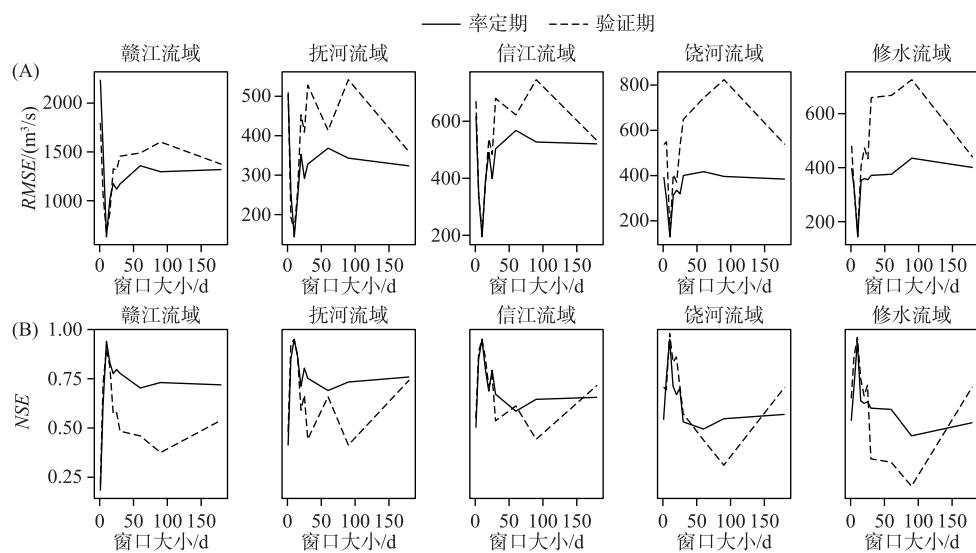


图 4 不同窗口大小下的模型表现

Fig.4 Model performance with different window sizes

表 4 气象—径流模型总体表现

Tab.4 Simulation performance of LSTM

流域	训练期		验证期	
	RMSE/(m ³ /s)	NSE	RMSE/(m ³ /s)	NSE
赣江	630.80	0.94	671.03	0.90
抚河	143.64	0.95	158.90	0.95
信江	194.74	0.95	212.04	0.95
饶河	129.46	0.95	139.52	0.98
修水	144.01	0.94	162.79	0.96

为了更加直观地反映气象—径流模型在径流过程模拟上的表现,将气象—径流模型模拟的时间序列与观测数据做了对比,结果如图 5A 所示。从气象—径流模型在训练期的模拟结果可以发现,LSTM 能够很好地模拟流域的径流过程,在训练期模型能够充分学习到鄱阳湖气象要素与径流过程间的相互关系,能够很好地捕捉径流的变化过程和极值(图 5B)。尽管气象—降雨模型在训练期取得了很好的拟合效果,但是其在验证期的结果仍然有待检验。由模型在验证期的结果可以发现,气象—径流模型较降雨—径流模型能够更好地捕捉径流的剧烈变化过程,对径流极值的模拟效果也比较好(图 5C),因此可以认为构建的气象—径流模型能够很好地模拟鄱阳湖流域的径流过程。

由鄱阳湖气象—径流模型的季节性模拟效果可见,在率定期模型在各个季节的模拟效果均较好,其中,春、夏、秋、冬季的平均 NSE 值分别为 0.90、0.96、0.73、0.74。在验证期的模拟效果略差于率定期,春、夏、秋、冬季的平均 NSE 值分别为 0.89、0.96、0.64、0.74(图 6)。Moriasi 等^[43]的研究结果表明,模型模拟的 NSE 在 0.75 左右即表示模型模拟效果良好。因此,本文构建模型的模拟精度较高,能够反映鄱阳湖流域径流的季节变化特征。

2.3 气候变化与人类活动对鄱阳湖流域径流过程的影响区分

利用训练的鄱阳湖流域气候—径流模型模拟了鄱阳湖 1970—2015 年的径流过程。基于模拟结果,进一步定量区分了人类活动与气候变化对鄱阳湖流域径流过程的贡献率,结果如图 7 所示。从图 7A 可以发现,人类活动对径流的影响主要发生在春、秋季,其中,人类活动在春季主要会造成径流的增加,平均增加幅度

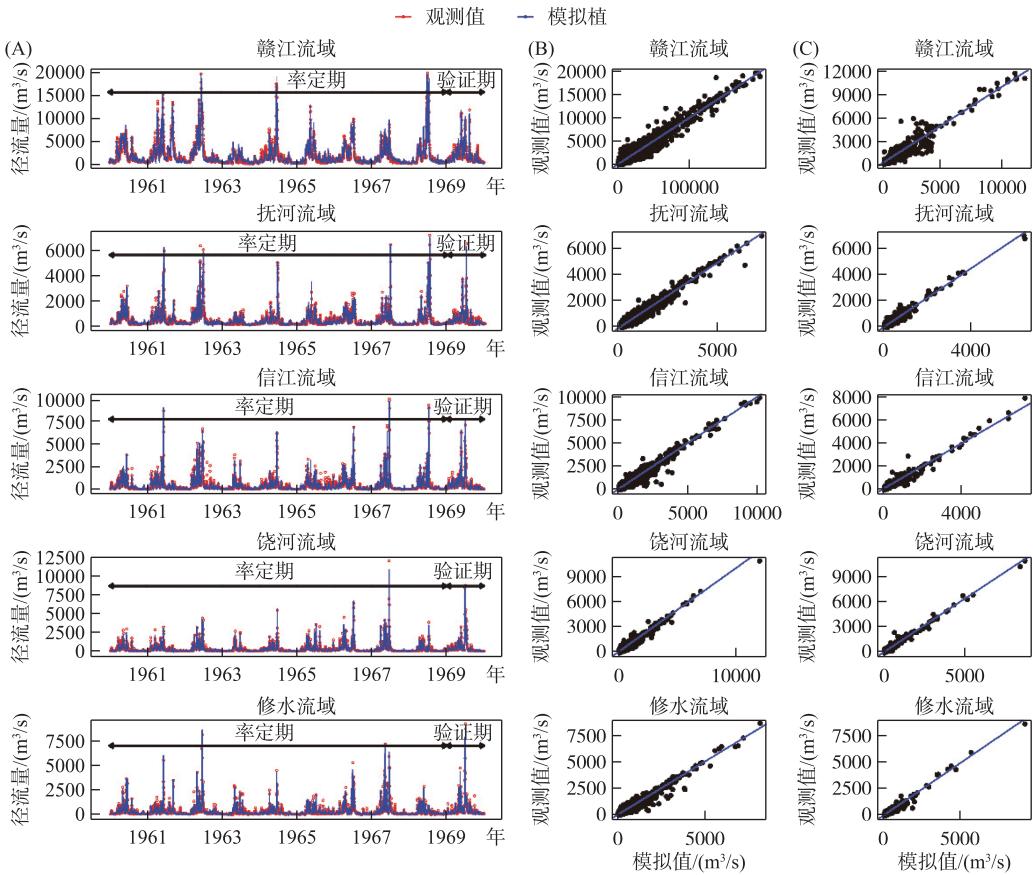


图 5 鄱阳湖流域气象—径流模型模拟结果(A: 总体模拟效果; B: 训练期模拟效果; C: 验证期模拟效果)

Fig.5 The simulation performance of the LSTM model in Lake Poyang Basin (A: the overall performance; B: model performance during the training period; C: model performance during the validation period)

约为 $139.47 \text{ m}^3/\text{s}$,而在秋、冬季,人类活动则会导致径流量平均减少约 $34.37 \text{ m}^3/\text{s}$. 图 7B 为人类活动和气候变化对径流过程影响相对贡献率的年内变化特征. 从相对贡献率来看,春季人类活动对径流造成的影响较大,平均相对贡献率为 77.26%. 其中,人类活动在春季对赣江流域径流变化的影响贡献最大,达到了 96.97%,而抚河、信江、饶河以及修水流域人类活动对径流变化的影响贡献分别为 62.77%、82.71%、93.31% 和 50.54%. 在其余季节,鄱阳湖流域径流过程的改变主要是由于气候变化,平均相对贡献率约为 75.84%. 其中,气候变化对秋、冬季径流变化的影响较大,平均贡献率分别达到了 82.58% 和 77.46%;而在夏季,气候变化对径流变化的贡献率要略低一些,平均为 67.49%.

3 讨论

近年来,气候发生了相当大的变化^[44],有学者认为气候变化很可能是对水循环产生影响的主导因素,继而影响到水资源,造成洪旱灾害^[2,45-49]. 鄱阳湖流域的气候在近年来也发生了剧烈的变化. 降雨和蒸发作为影响流域水量平衡的两大因素,在流域径流过程的变化上起着决定性作用. 有研究显示,近年来鄱阳湖流域的潜在蒸发量呈现长期下降趋势^[19],而实际蒸发量也呈现出波动下降的趋势^[50]. 蒸发量的下降会在一定程度上导致鄱阳湖流域径流量的增加. 而降水作为影响径流量的最主要因素,其年内变化会对流域径流量产生决定性影响. Zhang 等^[51]通过对比不同时期水汽通量变化,认为冬季水汽从自太平洋及南海北上增加导

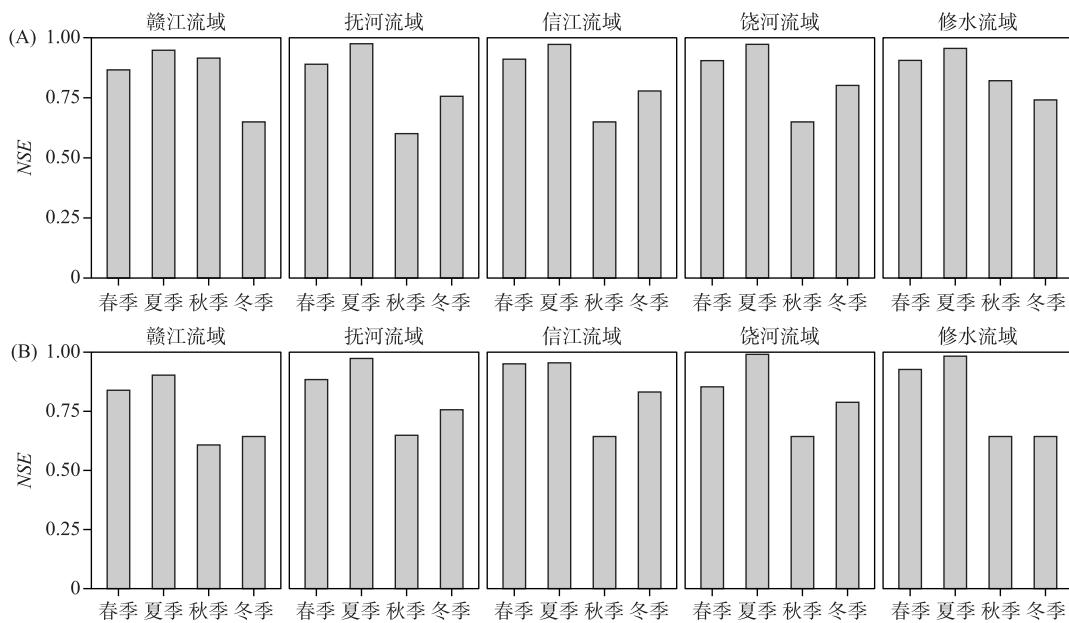


图 6 鄱阳湖流域气象—径流模型季节性模拟效果 (A: 训练期模拟效果; B: 验证期模拟效果)

Fig.6 The seasonal simulation performance of the LSTM model in Lake Poyang Basin

(A: model performance during the training period; B: model performance during the validation period)

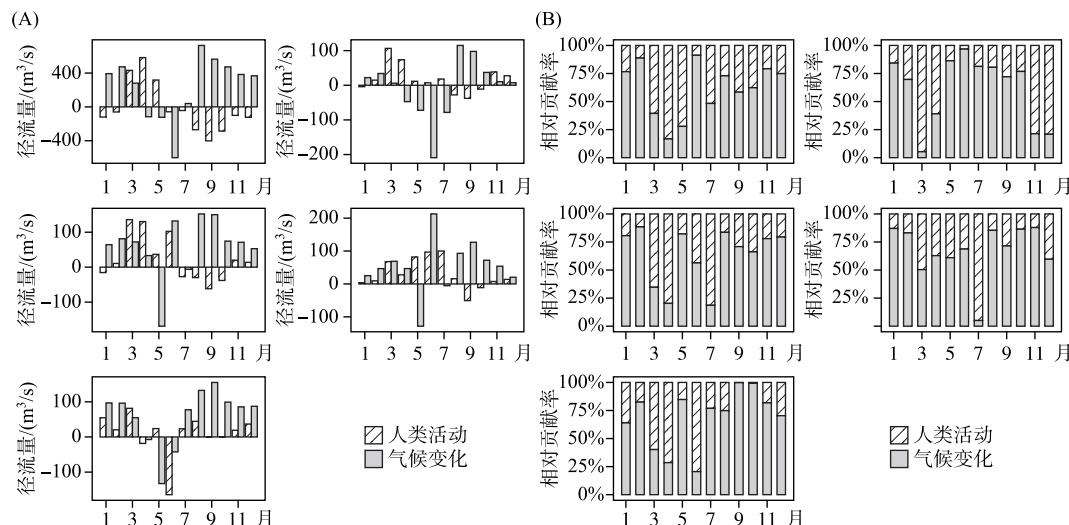


图 7 人类活动和气候变化对径流过程影响的年内变化 (A: 绝对贡献率; B: 相对贡献率)

Fig.7 Contributions for intra-annual runoff variations (A: absolute contribution rate; B: relative contribution rate)

致鄱阳湖流域秋、冬季降水量增加,与本文气候变化增加秋、冬径流量的结果基本一致。同时,Zhu 等^[52]研究指出,夏季水汽向北输送能力减弱,水汽输送量减少,使长江流域水汽含量下降,水汽上升运动减弱,导致长江流域夏季降水量显著减少,这也很好地解释了气候变化导致鄱阳湖流域夏季径流量的减小(图 6)。

以往的研究往往基于 Budyko 假设^[53]来定量区分气候变化和人类活动对径流变化的影响分量,并且都

指出气候变化对流域径流变化起着决定性的作用^[5,16,54]. 其中, Ye 等^[16]指出, 整个鄱阳湖流域 1970—2007 年年均径流变化相对于 1960s 主要受气候变化的影响, 而人类活动起着补充作用. 本文结果显示除了春季, 鄱阳湖流域径流过程的改变主要是由于气候变化, 其相对贡献率约为 75%, 和以往的研究结果相符. 但是, 由于 Budyko 假设是基于长时间的水量平衡方程, 往往需要长时间尺度的数据^[17,45,49], 并且其结果仅可达到年尺度^[5,54]. 对于月尺度或者更加精细尺度的区分, Budyko 假设则无法胜任. 而本文通过引入基准期的概念, 利用 LSTM 模拟了日尺度的流域径流过程, 可以在月尺度或者更加精细的时间尺度上定量区分气候变化以及人类活动对流域径流过程的影响分量. 本文在月尺度上发现人类活动对流域径流的改变存在不可忽视的贡献作用. 人类活动对水文过程的影响主要体现在两个方面, 一是改变流域下垫面特征, 从而改变降雨—径流关系; 有研究表明^[55], 从 1980s 开始, 鄱阳湖流域的标准化植被指数(*NDVI*)呈现显著的上升趋势, 有学者认为部分是由于鄱阳湖流域内的退耕还林工程造成的^[56]. 同时, 随着的江西山江湖开发治理工程的实施^[57], 1985—2001 年, 全省水土流失面积从 330 万 hm^2 下降到 130 万 hm^2 ; 全省植树造林 230 万 hm^2 , 基本消灭了宜林荒山, 森林覆盖率由 31.5% 上升到 59.7%. 植被覆盖的大量增加也导致了流域径流的迅速减少. 此外, 人类活动兴建的大型水利工程改变了河道径流在短时间尺度上的自然分配特性, 并且, 一般水利工程, 特别是大型水库的建设还可能增加流域中自由水体的面积, 从而增加蒸发量. 要威等^[58]的研究指出, 三峡水库调度对中下游水情影响主要表现为 11 月至次年 4 月径流量增大, 5—10 月径流量减小. 3 月径流量增加最多, 10 月减少最多, 9 月也减少较为明显. 这与本研究结果基本一致, 即人类活动在春、冬季会增加鄱阳湖流域径流量, 而在夏、秋季则会减少流域的径流量(图 6).

尽管神经网络技术改变了学术领域的方方面面, 但是其在水文领域的应用一直相对比较缺乏^[20]. 同时, 由于神经网络作为一个黑箱模型, 模拟结果无法从水文的角度进行解释, 也在一定程度上阻碍其在水文领域的应用前景. 最近 Kratzert 等^[24]通过神经元状态与气象水文观测数据的对比, 指出神经网络可以在某种程度上解释一定的水文现象, 但是其合理性仍有待于验证. 尽管如此, LSTM 出色的模拟效果证明了它可以用来进行高效的径流模拟. 近年来, 随着很多高级的神经网络框架如 Keras 和 Pytorch 被开发出来, 构建一个复杂的神经网络正在变得越来越方便, 使得决策者和科学家能够快速进行应用. 同时, 和传统水文模型相比, LSTM 需要的计算量更小, 模型更加轻量化, 因此其率定和验证的效率会大大提高, 使得其能够运用在水文水资源的实时预报方面, 为洪涝灾害的预警提供有力的工具.

4 结论

1) 长短记忆模型的预测窗口对模型精度影响至关重要. 预测窗口并不是越大越好, 过大的预测窗口会使模型发生过拟合的现象, 造成模拟精度的下降. 在同时考虑计算效率和模拟效果的前提下, 在鄱阳湖径流模型中使用 10 d 的预测窗口大小较为合理.

2) 基于长短记忆模型构建的鄱阳湖流域气象—径流模型在训练期和验证期在各个子流域的 NSE 均在 0.90~0.98 之间, 可以很好地模拟鄱阳湖流域的径流过程, 能够捕捉径流的极值, 并且对径流的短期波动也能有很好的体现.

3) 人类活动对径流的影响主要发生在春、秋季, 其中, 人类活动在春季主要会造成径流的增加, 平均增加幅度约为 139.47 m^3/s , 而在秋、冬季人类活动则会导致径流平均减少约 34.37 m^3/s . 对比二者的相对贡献率, 可以发现, 春季人类活动对径流造成的影响较大, 最大相对贡献率超过了 70%. 而在其余季节, 鄱阳湖流域径流过程的改变主要是由于气候变化, 平均贡献率约为 75.84%.

5 参考文献

- [1] Solomon S, Qin D, Manning M et al. Climate change 2007—the physical science basis: Working group I contribution to the fourth assessment report of the IPCC. Cambridge University Press, 2007.
- [2] Choudhury B. Evaluation of an empirical equation for annual evaporation using field observations and results from a bio-physical model. *Journal of Hydrology*, 1999, **216**(1): 99-110.
- [3] Jones P, Lister D, Osborn T et al. Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010. *Journal of Geophysical Research: Atmospheres*, 2012, **117**(D5).

- [4] Labat D, Goddérus Y, Probst JL et al. Evidence for global runoff increase related to climate warming. *Advances in Water Resources*, 2004, **27**(6) : 631-642.
- [5] Zhang D, Hong HY, Zhang Q et al. Attribution of the changes in annual streamflow in the Yangtze River Basin over the past 146 years. *Theoretical and Applied Climatology*, 2015, **119**(1/2) : 323-332. DOI: 10.1007/s00704-014-1121-3.
- [6] Guo H, Hu Q, Jiang T. Annual and seasonal streamflow responses to climate and land-cover changes in the Poyang Lake basin, China. *Journal of Hydrology*, 2008, **355**(1/2/3/4) : 106-122. DOI: 10.1016/j.jhydrol.2008.03.020.
- [7] Nilsson C, Malm-Renfält B. Linking flow Regime and water quality in rivers: a challenge to adaptive catchment management. *Ecology & society*, 2008, **13**(2) : 18.
- [8] Agostinho AA, Bonecker CC, Gomes LC. Effects of water quantity on connectivity: the case of the upper Paraná River floodplain. *Ecohydrology & Hydrobiology*, 2009, **9**(1) : 99-113.
- [9] Arias M, Piman T, Lauri H et al. Dams on Mekong tributaries as significant contributors of hydrological alterations to the Tonle Sap Floodplain in Cambodia. *Hydrology and Earth System Sciences*, 2014, **18**(12) : 5303-5315.
- [10] Ty TV, Sunada K, Ichikawa Y. A spatial impact assessment of human-induced intervention on hydrological regimes: a case study in the upper Srepok River basin, Central Highlands of Vietnam. *International Journal of River Basin Management*, 2011, **9**(2) : 103-116.
- [11] Cohen Liechti T, Matos JP, Boillat JL et al. Influence of hydropower development on flow regime in the Zambezi River Basin for different scenarios of environmental flows. *Water Resources Management*, 2015, **29**(3) : 731-747. DOI: 10.1007/s11269-014-0838-1.
- [12] Sankarasubramanian A, Vogel RM, Limbrunner JF. Climate elasticity of streamflow in the United States. *Water Resources Research*, 2001, **37**(6) : 1771-1781.
- [13] Githui F, Gitau W, Mutua F et al. Climate change impact on SWAT simulated streamflow in western Kenya. *International Journal of Climatology*, 2009, **29**(12) : 1823-1834.
- [14] Zhang X, Srinivasan R, Hao F. Predicting hydrologic response to climate change in the Luohe River basin using the SWAT model. *Transactions of the ASABE*, 2007, **50**(3) : 901-910.
- [15] Narsimlu B, Gosain AK, Chahar BR. Assessment of future climate change impacts on water resources of Upper Sind River Basin, India using SWAT model. *Water Resources Management*, 2013, **27**(10) : 3647-3662.
- [16] Ye XC, Zhang Q, Liu J et al. Distinguishing the relative impacts of climate change and human activities on variation of streamflow in the Poyang Lake catchment, China. *Journal of Hydrology*, 2013, **494**(12) : 83-95. DOI: 10.1016/j.jhydrol.2013.04.036.
- [17] Donohue RJ, Roderick ML, McVicar TR. Assessing the differences in sensitivities of runoff to changes in climatic conditions across a large basin. *Journal of Hydrology*, 2011, **406**(3/4) : 234-244. DOI: 10.1016/j.jhydrol.2011.07.003.
- [18] Roderick ML, Farquhar GD. A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties. *Water Resources Research*, 2011, **47**(12). DOI: 10.1029/2010wr009826.
- [19] Fan HX, Xu LG, Tao H et al. Accessing the difference in the climate elasticity of runoff across the Poyang Lake Basin, China. *Water*, 2017, **9**(2) : 135. DOI: 10.3390/w9020135.
- [20] Shen C. A trans-disciplinary review of deep learning research and its relevance for water resources scientists. *Water Resources Research*, 2018, **54**(11) : 8558-8593. DOI: 10.1029/2018WR022643.
- [21] Li YL, Zhang Q, Werner AD et al. Investigating a complex lake-catchment-river system using artificial neural networks: Poyang Lake (China). *Hydrology Research*, 2015, **46**(6) : 912-928. DOI: 10.2166/nh.2015.150.
- [22] Zhang J, Zhu Y, Zhang X et al. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *Journal of Hydrology*, 2018, **561** : 918-929. DOI: 10.1016/j.jhydrol.2018.04.065.
- [23] Fang K, Shen C, Kifer D et al. Prolongation of SMAP to spatiotemporally seamless coverage of continental U.S. using a deep learning neural network. *Geophysical Research Letters*, 2017, **44**(21) : 11030-11039. DOI: 10.1002/2017GL075619.
- [24] Kratzert F, Klotz D, Brenner C et al. Rainfall-Runoff modelling using Long-Short-Term-Memory (LSTM) networks. *Hydro Earth Syst Sci Discuss*, 2018, **22**(11) : 6005-6022.
- [25] Guo J, Zhang Y, Wang J et al. Application of the long short-term memory networks for flood forecast. *Journal of Water Resources Research*, 2019, **8**(1) : 24-32. [郭炅, 张艳军, 王俊勃等. 长短期记忆模型在小流域洪水预报上的应用研究. 水资源研究, 2019, 8(1) : 24-32.]

- [26] Guo Y, Lai X. Water level prediction of Lake Poyang based on long short-term memory neural network. *Journal of Lake Sciences*, 2020, **32**(3) : 865-876. DOI: 10.18307/2020.0325. [郭燕, 赖锡军. 基于长短时记忆神经网络的鄱阳湖水位预测. 湖泊科学, 2020, **32**(3) : 865-876.]
- [27] Hu Q, Cao S, Yang H et al. Daily runoff predication using LSTM at the Ankang Station, Hanjing River. *Progress in Geography*, 2020, **39**(4) : 636-642. [胡庆芳, 曹士圯, 杨辉斌等. 汉江流域安康站日径流预测的LSTM模型初步研究. 地理科学进展, 2020, **39**(4) : 636-642.]
- [28] Zhou Y, Guo S, Chang F et al. Hydrological forecasting using artificial intelligence techniques. *Journal of Water Resources Research*, 2019, **8**(1) : 1-12. [周研来, 郭生练, 张斐章等. 人工智能在水文预报中的应用研究. 水资源研究, 2019, **8**(1) : 1-12.]
- [29] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural computation*, 1997, **9**(8) : 1735-1780.
- [30] Graves A, Schmidhuber J. Framework phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, 2005, **18**(5/6) : 602-610.
- [31] Li Y, Cao H. Prediction for tourism flow based on LSTM neural network. *Procedia Computer Science*, 2018, **129** : 277-283. DOI: 10.1016/j.procs.2018.03.076.
- [32] Qi Y, Li Q, Karimian H et al. A hybrid model for spatiotemporal forecasting of PM_{2.5} based on graph convolutional neural network and long short-term memory. *Science of the Total Environment*, 2019, **664** : 1-10. DOI: 10.1016/j.scitotenv.2019.01.333.
- [33] Qin J, Liang J, Chen T et al. Simulating and predicting of hydrological time series based on TensorFlow deep learning. *Polish Journal of Environmental Studies*, 2018. DOI: 10.15244/pjoes/81557.
- [34] Yu Z, Yang K, Luo Y et al. Spatial-temporal process simulation and prediction of chlorophyll-a concentration in Dianchi Lake based on wavelet analysis and long-short term memory network. *Journal of Hydrology*, 2020, **582** : 124488. DOI: 10.1016/j.jhydrol.2019.124488.
- [35] Fan H, Jiang M, Xu L et al. Comparison of long short term memory networks and the hydrological model in runoff simulation. *Water*, 2020, **12**(1) : 175.
- [36] Wang XL, Han JY, Xu LG et al. Soil characteristics in relation to vegetation communities in the wetlands of Poyang Lake, China. *Wetlands*, 2014, **34**(4) : 829-839. DOI: 10.1007/s13157-014-0546-x.
- [37] Bergstra J, Bengio Y. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 2012, **13** (Feb) : 281-305.
- [38] Hsu CW, Chang CC, Lin CJ. A practical guide to support vector classification. 2003.
- [39] Hutter F, Hoos HH, Leyton-Brown K. Sequential model-based optimization for general algorithm configuration. International Conference on Learning and Intelligent Optimization, 2011: 507-523.
- [40] Bergstra JS, Bardenet R, Bengio Y et al. Algorithms for hyper-parameter optimization. Advances in neural information processing systems, 2011: 2546-2554.
- [41] Chapelle O, Vapnik V, Bousquet O et al. Choosing multiple parameters for support vector machines. *Machine learning*, 2002, **46**(1-3) : 131-159.
- [42] Azevedo IC, Bordalo AA, Duarte PM. Influence of river discharge patterns on the hydrodynamics and potential contaminant dispersion in the Douro estuary (Portugal). *Water Research*, 2010, **44**(10) : 3133-3146.
- [43] Moriasi DN, Arnold JG, Van Liew MW et al. Model evaluation guidelines for systematic quantification of accuracy in Watershed Simulations. *Transactions of the ASABE*, 2007, **50**(3) : 885-900.
- [44] O'Neill BC, Oppenheimer M, Warren R et al. IPCC reasons for concern regarding climate change risks. *Nature Climate Change*, 2017, **7**(1) : 28-37.
- [45] Donohue RJ, Roderick ML, McVicar TR. Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model. *Journal of Hydrology*, 2012, **s436/437** : 35-50.
- [46] Donohue RJ, Roderick ML, McVicar TR et al. Impact of CO₂ fertilization on maximum foliage cover across the globe's warm, arid environments. *Geophysical Research Letters*, 2013, **40**(12) : 3031-3035.
- [47] Yang H, Yang D, Hu Q. An error analysis of the Budyko hypothesis for assessing the contribution of climate change to runoff. *Water Resources Research*, 2014, **50**(12) : 9620-9629.
- [48] Yang HB, Yang DW. Derivation of climate elasticity of runoff to assess the effects of climate change on annual runoff. *Wa-*

- ter Resources Research, 2011, 47. DOI: 10.1029/2010wr009287.
- [49] Yang HB, Yang DW, Lei ZD et al. New analytical derivation of the mean annual water-energy balance equation. *Water Resources Research*, 2008, 44(3). DOI: 10.1029/2007wr006135.
- [50] Guiwu Q, Li Y, Aimin C. Prediction of future water resources quantity in Poyang Lake Basin. *Journal of Anhui Agricultural Sciences*, 2017, 45(8): 53-54, 73. [钱贵伍, 游黎, 陈爱民. 鄱阳湖流域水资源量预测. 安徽农业科学, 2017, 45(8): 53-54, 73.]
- [51] Zhang Q, Liu Y, Yang G et al. Precipitation and hydrological variations and related associations with large-scale circulation in the Poyang Lake basin, China. *Hydrological Processes*, 2011, 25(5): 740-751.
- [52] Zhu Y, Wang H, Zhou W et al. Recent changes in the summer precipitation pattern in East China and the background circulation. *Climate Dynamics*, 2011, 36(7/8): 1463-1473.
- [53] Budyko MI ed. Climate and life. New York: Academic Press, 1974.
- [54] Xu X, Yang D, Yang H et al. Attribution analysis based on the Budyko hypothesis for detecting the dominant cause of runoff decline in Haihe basin. *Journal of Hydrology*, 2014, 510(6): 530-540.
- [55] Piao S, Ciais P, Huang Y et al. The impacts of climate change on water resources and agriculture in China. *Nature*, 2010, 467(7311): 43-51. DOI: 10.1038/nature09364.
- [56] Liu Q, McVicar TR. Assessing climate change induced modification of Penman potential evaporation and runoff sensitivity in a large water-limited basin. *Journal of Hydrology*, 2012, 464: 352-362. DOI: 10.1016/j.jhydrol.2012.07.032.
- [57] Dai X, Luo B. From “harnessing” to “enrichment”—research of strategic upgrading of mountain-river-lake engineering: providing strong support for development of Changjiang Economic Zone. *Yangtze River*, 2016, 47(24): 1-5, 11. [戴星照, 罗斌华. 从“治”到“富”:“山江湖工程”战略升级研究. 人民长江, 2016, 47(24): 1-5, 11.]
- [58] Yao W, Yu Q, You Z. Study on reservoirs operation at upper Yangtze River for alleviating low water situation of Poyang Lake. *Yangtze River*, 2015, 46(24): 1-5, 20. [要威, 余启辉, 游中琼. 长江上游水库调度缓解鄱阳湖枯水情势作用研究. 人民长江, 2015, 46(24): 1-5, 20.]