J. Lake Sci.(湖泊科学), 2020, **32**(3): 865-876
DOI 10. 18307/2020. 0325

基于长短时记忆神经网络的鄱阳湖水位预测*

郭　燕[1,2], 赖锡军[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:中国科学院大学资源与环境学院,北京 100049)

摘　要: 湖泊水位是维持其生态系统结构、功能和完整性的基础. 鄱阳湖受流域"五河"和长江来水双重影响,水位变化复杂. 为了准确预测鄱阳湖水位变化,采用长短时记忆神经网络方法(LSTM)构建了鄱阳湖水位预测模型. 该模型以赣江、抚河、信江、饶河和修水"五河"入湖流量和长江干流流量作为输入条件,预测鄱阳湖湖区不同代表站(湖口、星子、都昌、吴城和康山)的水位过程. 研究以 1956－1980 年的水文时间序列数据作为训练集,1981－2000 年作为验证集,探讨了 LSTM 模型输入时间窗、隐藏神经元数目、初始学习率等模型参数对预测精度的影响,并确定了鄱阳湖水位预测模型的最优参数. 结果表明,采用 LSTM 神经网络方法可基于流域"五河"和长江来水量历时数据合理预测鄱阳湖不同湖区的水位过程,五站水位预测的均方根误差为 0.41~0.50 m,纳什效率系数和决定系数达 0.96~0.98. 为考察模型训练数据集对鄱阳湖水位预测结果的影响,进一步选取了随机 5 年(1956－1960 年)的资料和 5 个典型水文年(1954 年、1973 年、1974 年、1977 年和 1978 年)的日均流量资料来训练模型. 结果显示随机 5 年资料作为训练数据的预测精度要差于典型年水文资料训练得到的模型,尤其是洪、枯水位的预测;由于典型水文年数据量仍远低于 20 年的资料,故其总体预测精度要略低于采用 20 年资料训练的模型. 建议应用这类基于数据驱动的模型时,应该尽可能多选取具有代表性的资料来训练.
关键词: 湖泊水位;LSTM 循环神经网络;模型参数;训练集;鄱阳湖

Water level prediction of Lake Poyang based on long short-term memory neural network*

GUO Yan[1,2] & LAI Xijun[1**]
(1: *Key Laboratory of Watershed Geographic Sciences*, *Nanjing Institute of Geography and Limnology*, *Chinese Academy of Sciences*, *Nanjing* 210008, *P.R.China*)
(2: *College of Resources and Environment*, *University of Chinese Academy of Sciences*, *Beijing* 100049, *P.R.China*)

Abstract: Lake water level is the basis for maintaining the structure, function, and integrity of its ecosystem. The water level change of Lake Poyang is complicated as it was affected by five rivers within the basin and the Yangtze River. To accurately predict the water level change of Lake Poyang, the long short-term memory (LSTM) is used to construct the water level prediction model of Lake Poyang. The model uses the flows of the Ganjiang River, Fuhe River, Xinjiang River, Raohe River, Xiushui River and the mainstream of the Yangtze River as input conditions to predict the water level process of different representative stations in the Lake Poyang area (Hukou, Xingzi, Duchang, Wucheng and Kangshan). The hydrological time series data from 1956 to 1980 is used as the training set, and data from 1981 to 2000 was used as the verification set. The influence of model parameters such as input time window, hidden neuron nodes and initial learning rate on prediction accuracy is discussed. The optimal parameters of the Lake Poyang water level prediction model are determined. The results show that the LSTM can accurately predict the water level at different parts of Lake Poyang based on the water flow from the five rivers and the Yangtze River. The *RMSE* value of the five stations is 0.41-0.50 m, and the *NSE* and R^2 are 0.96-0.98. In order to investigate the impact of the model training set on the water level prediction results of Lake Poyang, the study further selects data from 5 random years (1956-1960) and 5 typical hydrological years (1954, 1973, 1974, 1977 and 1978) daily average flow data to train the model. The results show that the prediction accura-

* 2019-06-10 收稿;2019-11-06 收修改稿.
中国科学院战略性先导科技专项(A 类)(XDA230402)资助.
** 通信作者;E-mail:xjlai@ niglas.ac.cn.

cy of random 5 years data as training set is worse than that of typical annual hydrological data training, especially the prediction of flood and dry water level; since the typical hydrological data volume is still much lower than 20 years of data, the overall prediction accuracy is slightly lower than the model with 20 years of data training. Therefore, representative data should be selected as much as possible for training, when applying such a data-driven LSTM neural network model.

Keywords: Lake water level; LSTM neural network; model parameters; training set; Lake Poyang

鄱阳湖是中国第一大淡水湖,水量主要来自流域"五河". 汛期,长江水可能倒灌鄱阳湖. 鄱阳湖是典型的季节性吞吐型湖泊,水位年内年际变化剧烈,具有"洪水一片,枯水一线"的特征. 每年4—6月为鄱阳湖主汛期,湖水位随"五河"来水的增加而升高;7—9月是长江主汛期,受到长江高水位的顶托作用,湖水位进一步升高;每年10月—次年3月鄱阳湖属于枯水期,水位较低. 水位变化会导致湿地植被分布以及候鸟栖息等也相应发生改变[1-3]. 近十几年来,在气候变化和高强度的人类活动影响下,鄱阳湖水情发生了剧烈的变化,特别是枯水问题突出,出现了枯水发生时间提前、持续时间延长的现象. 湖区生态环境因水量平衡的变化面临威胁[4-5].

鄱阳湖水量平衡近年成为了学界关注热点. 为阐明鄱阳湖的水量变化,开展了大量卓有成效的研究,建立了诸多水文水动力模型和数理统计模型,分析了气候变化和不同类型的人类活动对鄱阳湖水量的影响过程以及江湖交互作用下湖泊水位的变化特征[6-9]. 水动力模型可实现过程的精细模拟,但是水动力系统庞大、需要非常完善的基础资料、计算复杂、耗时较长,对水位高效预测仍存挑战[10-11]. 数理统计方法(如:多元线性回归(MLR)、累积距平曲线(CPC)、概率分布函数(PDF)、泰尔森方法(TSA))简单易行,主要用于长时间尺度的变化检验分析[12-16]. 鄱阳湖水位受流域"五河"和长江来水多重控制,其变化与"五河"来水和长江来水有着非常复杂的非线性关系[17],难以采用简单的统计模型模拟预测水位和来水的响应关系. 人工神经网络方法作为一种数据驱动的自适应性方法,没有先验假设,可较好地模拟预测复杂的非线性作用关系,也在鄱阳湖得到了应用[18]. 近年来,基于人工神经网络的深度学习算法得到了飞速发展. 特别是长短时记忆神经网络方法(long short-term memory,LSTM),它通过循环反馈结构存储历史信息,具有较强的时间序列问题求解能力,在模拟预测水文时间序列问题中受到关注. Zhang 等[19]基于雨量器和水位传感器的在线数据,比较了不同神经网络结构在模拟和预测挪威 Drammen 联合下水道结构水位方面的性能,证实了 LSTM 比没有显式细胞记忆的传统架构更适合于多步预测. Zhang 等[20]使用 LSTM 预测农业地区的地下水位,并将基于 LSTM 方法的预测结果与传统神经网络的预测结果进行了比较,发现前者的性能优于后者. Lee 等[21]利用基于物理的水文模型 SWAT 和数据驱动的深度学习算法 LSTM,在湄公河下游 Kratie 站进行了径流模拟. Liu 等[22]提出了改进的 CA-LSTM 上下文感知神经网络,基于收集到的洪水因子序列数据对洪水进行了预测.

鄱阳湖水位受江湖共同制约,其变化与流域"五河"来水量及长江来水量具有非常复杂的非线性关系. 本文选取鄱阳湖为研究对象,采用 LSTM 神经网络方法,建立鄱阳湖水位日尺度的预测模型,探讨其用于预测江湖交互作用下鄱阳湖水位变化过程的潜力,为鄱阳湖水量平衡研究提供一个快速有效的预测方法.

1 数据和方法

1.1 研究区概况

鄱阳湖(28°24′~29°46′N,115°49′~116°46′E)位于长江中下游南岸,北部较平坦,东西南三面环山,全流域总体地势是南高北低,南部地形起伏变化较大,形成了一个以鄱阳湖为底,向北开口连接长江的巨大盆地. 鄱阳湖流域由赣江、抚河、信江、饶河和修水5个相互独立的子流域组成,与湖区共同形成了一个完整的流域系统,流域总面积为 16.2×10^4 km^2[16, 23]. 多年平均径流量为1491亿 m^3,水位年内变幅为9.59~15.36 m.

1.2 研究数据

选用"五河"流域外洲站、李家渡站、梅港站、渡峰坑站、虎山站、万家埠站和长江中游汉口站7站逐日平均流量数据以及湖口站、星子站、都昌站、吴城站和康山站5站逐日平均水位数据建模,时间序列长度均为1956—2000年,监测站点分布见图1.

1.3 LSTM 基本原理

1.3.1 LSTM 前向计算 LSTM 是1997年 Hochreiter 和 Schmidhuber[24]为了解决 RNN 模型的梯度消失或爆炸

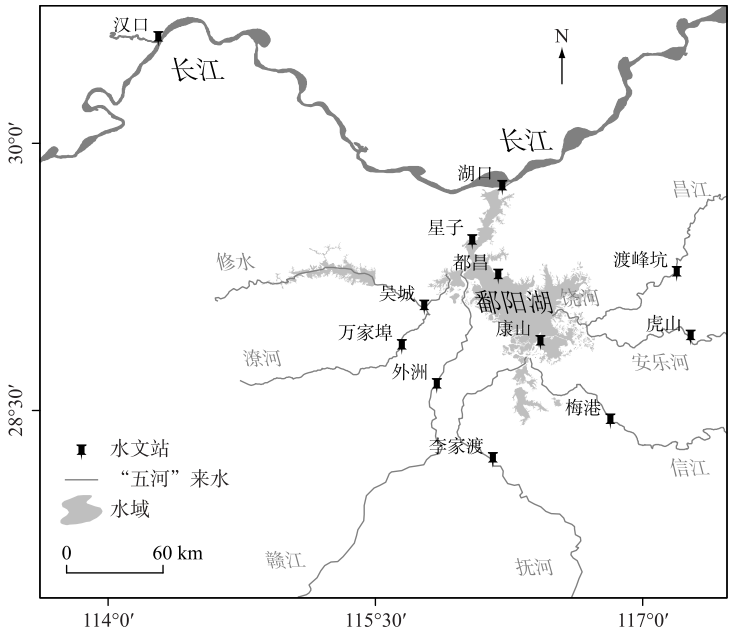

图 1 鄱阳湖监测站点的空间分布

Fig.1 Location of sampling sites in the Lake Poyang

缺陷而开发的一种复杂的递归模型. Gers 等[25]正式提出 LSTM 网络层是由遗忘门 f、输入门 i、记忆单元 C 和输出门 o 组成(图 2). LSTM 的关键是记忆单元,它像传送带一样,将信息从上一个单元传递到下一个单元.

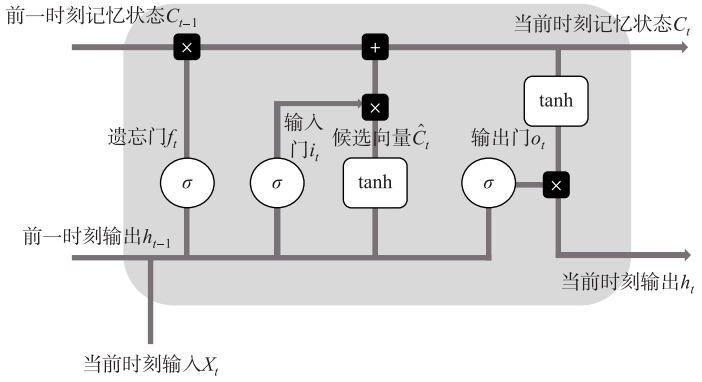

图 2 LSTM 内部结构示意图

Fig.2 Schematic diagram of the internal structure of LSTM

首先,决定从记忆单元中增减多少信息. 遗忘门本质上是一个 σ 神经网络层,根据前一时刻的输出 h_{t-1} 和当前的输入 X_t ,产生一个介于 0 和 1 之间的 f_t 值:

$$f_t = \sigma(W_{xf} \cdot X_t + W_{hf} \cdot h_{t-1} + b_f) \tag{1}$$

接下来,确定将在记忆单元中存储哪些新信息. 这一部分由一个 σ 神经网络层和一个 tanh 神经网络层两部分组成. 输入门根据前一时刻的输出 h_{t-1} 和当前的输入 X_t ,产生一个介于 0 和 1 之间的 i_t 值:

$$i_t = \sigma(W_{xi} \cdot X_t + W_{hi} \cdot h_{t-1} + b_i) \tag{2}$$

tanh 神经网络层产生一个新的候选向量 $\hat{C}$:

$$\hat{C} = \tanh(W_{xc} \cdot X_t + W_{hc} \cdot h_{t-1} + b_c) \tag{3}$$

从而更新记忆单元:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \hat{C} \tag{4}$$

最后,我们决定输出什么信息. 首先,运行一个 σ 神经网络层来决定输出记忆单元的哪些部分. 然后,通过 tanh 神经网络层将 C_t 值调整为-1~1,并将其乘以 σ 神经网络层的输出,最终得到目标信息.

$$o_t = \sigma(W_{xo} \cdot X_t + W_{ho} \cdot h_{t-1} + b_o) \tag{5}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{6}$$

式(1~6)中,f、i、o、C 和 h 分别是遗忘门、输入门、输出门、记忆单元和输出信息,W 是相应的权重矩阵,b 是偏差矩阵,σ 和 tanh 是激活函数.

1.3.2 LSTM 反向训练 神经网络训练的过程是去寻找最优参数,使得模型收敛,即损失函数达到极小甚至最小的过程. 网络通过反向传播损失函数,利用梯度下降法迭代更新网络的权重[26-27]. 我们选取最常用的均方误差计算损失函数:

$$\text{loss} = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 \tag{7}$$

式中,loss 为损失函数,y_i 和 $\hat{y}_i$ 分别为 i 时刻的观测值和预测值.

1.4 模型评估标准

LSTM 是通过已有的训练样本(即已知数据及对应的输出)去学习得到一个最优模型,再利用该模型将所有的输入映射为相应的输出. 我们采用均方根误差(*RMSE*)、纳什效率系数(*NSE*)和决定系数(R^2)3 个指标对模型模拟的效果进行评价[28].

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i^{\text{obs}} - Y_i^{\text{pre}})^2}{n}} \tag{8}$$

RMSE 值为 0 表示观测值与预测值完全吻合.

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n}(Y_i^{\text{obs}} - Y_i^{\text{pre}})^2}{\sum_{i=1}^{n}(Y_i^{\text{obs}} - Y^{\text{mean}})^2}\right] \tag{9}$$

NSE 是一个参数,它决定了剩余方差(噪声)相对于测量数据(信息)中的方差的相对重要性,取值范围为负无穷至 1. 接近 1,表示模拟效果好,模型可信度高;接近 0,表示模拟结果接近观测值的平均值水平,即总体结果可信,但过程模拟误差大.

$$R^2 = \left[\frac{(\sum_{i=1}^{n}(Y_i^{\text{pre}} - Y_{\text{pre}}^{\text{mean}})(Y_i^{\text{obs}} - Y^{\text{mean}}))^2}{\sum_{i=1}^{n}(Y_i^{\text{pre}} - Y_{\text{pre}}^{\text{mean}})^2 \sum_{i=1}^{n}(Y_i^{\text{obs}} - Y^{\text{mean}})^2}\right] \tag{10}$$

式(8~10)中,Y_i^{obs}、Y_i^{pre}、Y^{mean} 和 $Y_{\text{pre}}^{\text{mean}}$ 分别表示观测值、预测值、观测值的平均值和预测值的平均值,n 为数据长度. R^2 取值为 0~1,越大表示模型拟合效果越好.

2 鄱阳湖水位预报模型构建

2.1 模型结构

采用单层 LSTM 循环神经网络建模,输入长江和"五河"来水量,其中长江水量以汉口站流量为代表,"五河"水量以赣江外洲站、抚河李家渡站、信江梅港站、饶河渡峰坑站和虎山站、修水万家埠站流量为代表,输出湖口站、星子站、都昌站、吴城站和康山站 5 个代表站的水位,实现根据"五河"来水量和长江来水量组合预测鄱阳湖湖区未来 1 d 水位的空间分布,神经网络结构见图 3,表达式为:

$$h_i^{t+1} = f(Q_j^t, Q_j^{t-1}, \cdots, Q_j^{t-n+1}) \tag{11}$$

式中,h_i^{t+1} 为未来 1 d 的水位,Q_j^t 为当前时刻流量,Q_j^{t-n+1} 为输入时间窗为前 n 天时刻的流量.

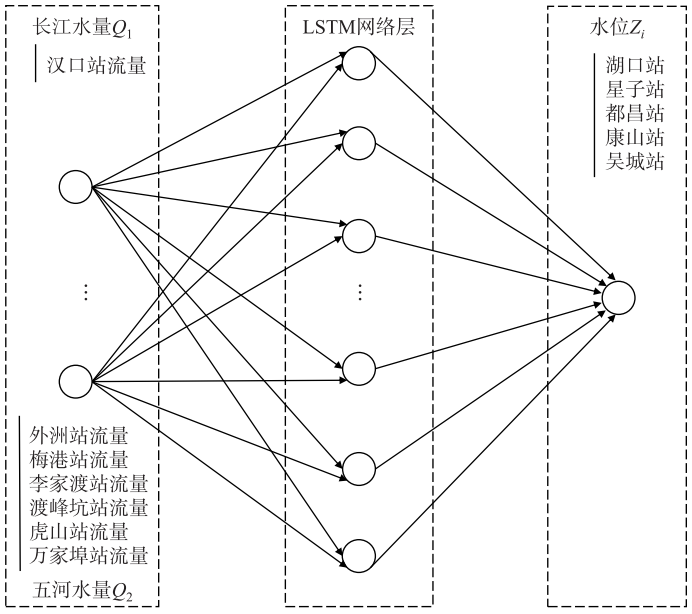

图 3 单层长短时记忆网络 LSTM 模型结构

Fig.3 Structure of a single layer LSTM model

数据之间的差异性会对模型的学习能力产生负面影响. 因此,为了保证构建的模型中参数能够稳定收敛,在神经网络训练之前需对其进行预处理,以确保所有变量保持在相同的量表上[29-30]. 我们对所有数据进行[0,1]归一化处理:

$$X = \frac{X_{\text{ori}} - X_{\min}}{X_{\max} - X_{\min}} \tag{12}$$

式中,X 是归一化后的数据,X_{ori}是原始数据,$X_{\max}$和 $X_{\min}$分别是原始数据的最大值和最小值.

模型损失函数选用均方根误差,优化选取基于梯度下降的 ADAM 算法,采用 Dropout 正则化方法防止模型过拟合[31-33]. 训练集为 1956－1980 年的水文时间序列数据,验证集数据长度为 1981－2000 年. 在 LSTM 神经网络中,输入时间窗、隐藏神经元数目、初始学习率大小等重要参数会直接影响到模型预测效果. 所以,接下来我们将对模型中的这些参数进行优选.

2.2 模型参数优选

2.2.1 输入时间窗长短对模型模拟效果的影响　鄱阳湖作为大型通江洪泛湖泊,其入湖流量与湖泊水位关系呈现明显的非线性特征,加之长江与湖泊之间相互作用的复杂性,湖泊流量与水位之间呈现多种对应关系[34]. “五河”水量经一段时间传输入湖,鄱阳湖接受“五河”及长江来水进行调蓄作用,因而水位和流量并不总是同步变化,存在明显的相位滞后效应,输入时间窗的优选计算就是确定流量—水位之间滞后时长的过程. 我们计算并比较了不同长短的输入时间窗(1~10 d,即分别将当前时刻的流量、前 2 d 内的流量、……、前 10 d 内的流量作为输入变量)下,根据“五河”来和长江来水量,组合预测鄱阳湖湖区未来 1 d 的水位空间分布. 同时,为了确保模型较高的准确率,隐藏神经元数目应尽量多取,设置了 100 个;初始学习率大小应定义为较小的值,设置为 0.001. 不同输入时间窗模式下模型拟合评价指标计算结果如图 4 所示.

虽然神经网络模型的外延效果一般不是很理想,但可通过反复对模型调试计算、选取最适宜的输入时间窗来提高模型预测效果. 图 4 结果表明:不同输入时间窗下,湖口站水位预测的效果均最好,其次分别是星子站、都昌站和吴城站,康山站水位预测效果为五站中最差. 利用当前时刻的七站流量来预测未来 1 d 的鄱阳湖五站水位时,湖口站训练和验证过程的 *RMSE* 均较大,分别为 0.91 m 和 0.79 m. *NSE* 和 R^2相对较小,训练阶段 *NSE* 为 0.94,R^2为 0.95;验证阶段 *NSE* 和 R^2均为 0.96. 随着输入时间窗逐步延长,模型训练和验证

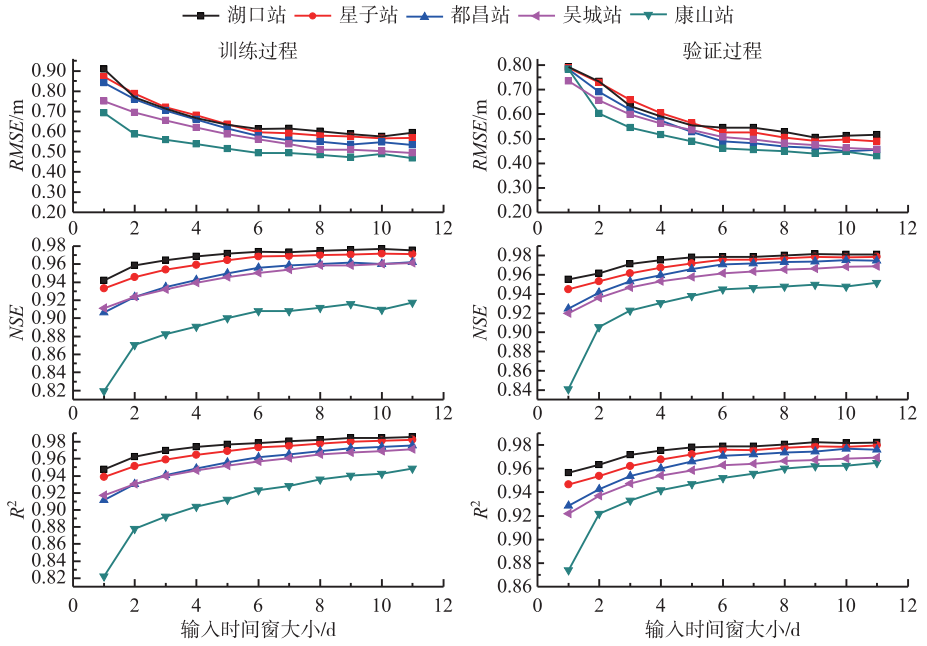

图 4 不同输入时间窗下模型训练和验证过程的模拟结果

Fig.4 Simulation results of model training and verification under different input time windows

阶段的模拟效果都稳步提高. 当输入时间窗为 7 d 时,模型各项性能指标已达到很好的效果,继续增大输入时间窗,模型预测效果没有明显提高. 因此,我们选择利用"五河"六站及长江干流汉口站这七站前 7 d 内的流量来预测鄱阳湖湖口、星子、都昌、吴城和康山站未来 1 d 的水位. 其中,湖口站训练和验证阶段的 *RMSE* 分别为 0.62 和 0.55 m,*NSE* 分别为 0.97 和 0.98,R^2 均为 0.98.

2.2.2 隐藏神经元数目对模型模拟效果的影响 LSTM 网络层中隐藏神经元数目是影响模型预测结果准确率的重要参数之一. 若数量太少,网络不能很好地学习,需要训练的次数也多,训练精度也不高;若数量太多,训练时间又较长,甚至可能导致模型不收敛[35-36]. 因此,我们进一步计算了隐藏神经元数分别为 1、5、10、20、30、40、50、100、200、500 共 10 种模式下,根据前 7 d"五河"六站及汉口站流量组合预测鄱阳湖 5 站未来 1 d 的水位空间分布,相应的训练和验证阶段模型拟合评价指标的计算结果如图 5 所示.

计算结果表明,一定范围内随着隐藏神经元数目的增加,模型在训练和验证阶段的水位预测效果均稳步提高,但当隐藏神经元数目增加到一定值之后,继续增加其数量模型预测效果变化不大. 隐藏神经元数为 1 个时,各站水位预测效果最差. 湖口站水位训练和验证阶段的 *RMSE* 均高达 1.0 m 以上,分别为 1.02 和 1.13 m,*NSE* 和 R^2 均相对较低. 隐藏神经元数目为 50 个时,各站水位预测效果的评估指标均达到平稳. 其中,作为 5 个水位站中预测效果最差的康山站,*RMSE* 仅为 0.45 m,*NSE* 为 0.95,R^2 为 0.96. 考虑到模型预测精度以及计算时间,我们设置了 50 个隐藏神经元节点. 模型训练阶段的 *RMSE* 为 0.61 m,*NSE* 为 0.97,R^2 为 0.98;验证阶段的 *RMSE* 为 0.53 m,*NSE* 为 0.98,R^2 也为 0.98.

2.2.3 初始学习率大小对模型模拟效果的影响 学习率表示每次迭代后权重的更新量,学习率太小,模型更新速度慢;学习率过大,模型可能错过最优解. 为了找到模型的全局最优解,我们借助基于梯度下降的具有相当鲁棒性的 ADAM 自适应学习率优化算法作为优化器[37],算法描述为:

$$m_t = \beta_1 \cdot m_{t-1} + (1-\beta_1)\, g_t \tag{13}$$

$$v_t = \beta_2 \cdot v_{t-1} + (1-\beta_2)\, g_t^2 \tag{14}$$

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \tag{15}$$

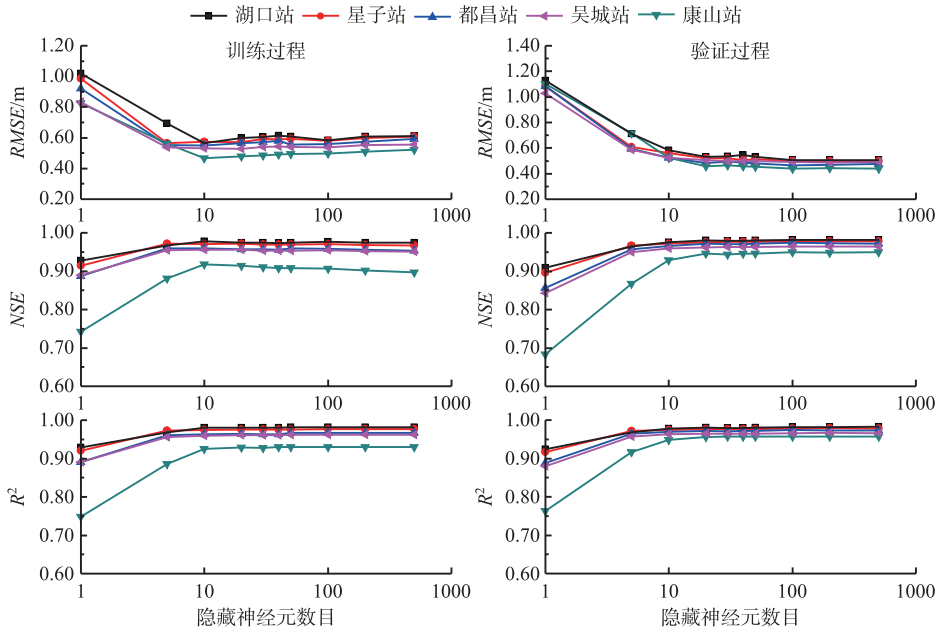

图 5 隐藏神经元数目对模型训练和验证过程的模拟结果的影响

Fig.5 The effect of the number of hidden nodes on the simulation results of model training and verification

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{16}$$

$$W_{t+1} = W_t - \frac{\eta}{\sqrt{\hat{v}_t}+\varepsilon}\hat{m}_t \tag{17}$$

式中，m_t 和 v_t 分别为一阶动量项和二阶动量项；β_1 和 β_2 为动力值大小，通常分别取 0.9 和 0.999；$\hat{m}_t$ 和 $\hat{v}_t$ 为各自的修正值；W_t 表示第 t 次迭代时模型的参数；g_t 为梯度；ε 是一个取值很小的数(一般为 10^{-8})，为了避免分母为 0.

计算了 4 种初始学习率(分别为 0.1、0.01、0.001 和 0.0001)下模型训练和验证阶段损失函数的大小(图 6). 我们发现，初始学习率为 0.1 时，损失函数曲线随着迭代次数的增加发生不同幅度的震荡，此时学习率选择过大；当初始学习率为 0.01 时，训练阶段的损失函数具有较好的学习过程，但验证阶段的损失函数曲线迅速下降，模型学得太快，基于对模型精度和速度的综合考虑，认为初始学习率过大；当初始学习率为 0.0001 时，则经过较长时间模型才得以收敛. 因此，我们选取训练和验证阶段均有很好的学习过程曲线的 0.001 作为最适初始学习率，且模型快速地收敛于 0.03. 此时，五站中水位预测效果最好的湖口站，训练和验证阶段的 *RMSE* 分别为 0.58 和 0.50 m，*NSE* 均为 0.98，R^2 也均为 0.98. 预测效果相对最差的康山站，验证阶段的 *RMSE* 为0.42 m，*NSE* 和 R^2 分别可达 0.95 和 0.96. 康山站的 *RMSE* 值反而比湖口站的低，是因为各站年内水位变幅存在差异，康山站多年年内变幅均值为 6.03 m，而湖口站可达 12.00 m.

3 水位预测

3.1 预测结果分析

综上所述，最终我们建立了含 50 个隐藏神经元的单层 LSTM 神经网络，采用基于梯度下降的自适应学习率的 ADAM 优化算法，初始学习率为 0.001，利用“五河”六站和长江干流共 7 个流量站前 7 d 内的日均流量，组合预测鄱阳湖从北到南 5 个水位站未来 1 d 的日均水位. 将 1956－1980 年水文序列数据作为训练集，1981－2000 年数据对模型进行验证. 构建的 LSTM 模型对湖口、星子、都昌、康山和吴城站的水位预测精度

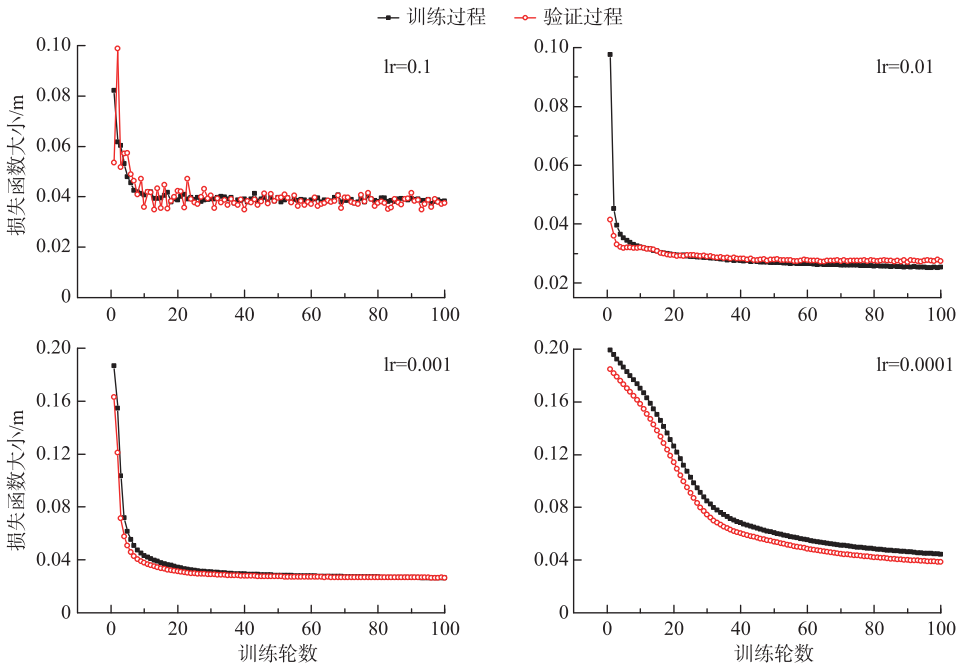

图 6 初始学习率大小对模型训练和验证过程的模拟结果的影响

Fig.6 The effect of the initial learning rate on the simulation results of model training and verification

较高,模型拟合的各项评价指标如表 1 所示,训练阶段和验证阶段 *RMSE* 都很小,范围分别为 0.48~0.60 m 和 0.41~0.50 m. 同时,训练和验证阶段的 *NSE* 和 R^2 都很大,几乎接近于 1. 其中,训练阶段和验证阶段的 *NSE* 范围分别为 0.92~0.98 和 0.96~0.98;两个阶段的 R^2 范围则分别为 0.93~0.98 和 0.96~0.98. 湖口站和星子站 1981—2000 年水位预测的验证结果如图 7 所示,以两站 2000 年的水位过程线为例,湖口站年均水位实测值和预测值分别为 12.98 和 13.26 m. 最高日均水位实测值为 18.10 m,预测值较之大 0.62 m,但与年均水位相比,相差较小,具有较强的可靠性. 日均最低水位的实测值与预测值分别为 7.73 和 7.61 m,较年均和年最高水位值的预测结果更精确. 构建的 LSTM 模型对湖泊最下游湖口站的水位预测精度最高,从最上游康山站至湖口站,模型对五站水位预测精度有逐步增强的趋势,这可能和湖区各站受长江的顶托关系逐渐趋弱有关.

表 1 最佳模型评价指标计算结果

Tab.1 Calculation result of best model evaluation index

数据类型	评价指标	湖口站	星子站	都昌站	康山站	吴城站
训练数据	*RMSE*/m	0.58	0.60	0.58	0.48	0.54
	NSE	0.98	0.97	0.96	0.92	0.95
	R^2	0.98	0.97	0.96	0.93	0.96
验证数据	*RMSE*/m	0.50	0.49	0.46	0.41	0.50
	NSE	0.98	0.98	0.97	0.96	0.96
	R^2	0.98	0.98	0.97	0.96	0.97

3.2 模型能力比较

为了充分验证本文构建的 LSTM 模型的水位模拟精度,突出模型的优势以及后续模型应用,我们也用一般的循环神经网络 BP 神经网络对鄱阳湖水位进行模拟,并与 LSTM 模型的模拟精度进行对比. BP 神经网

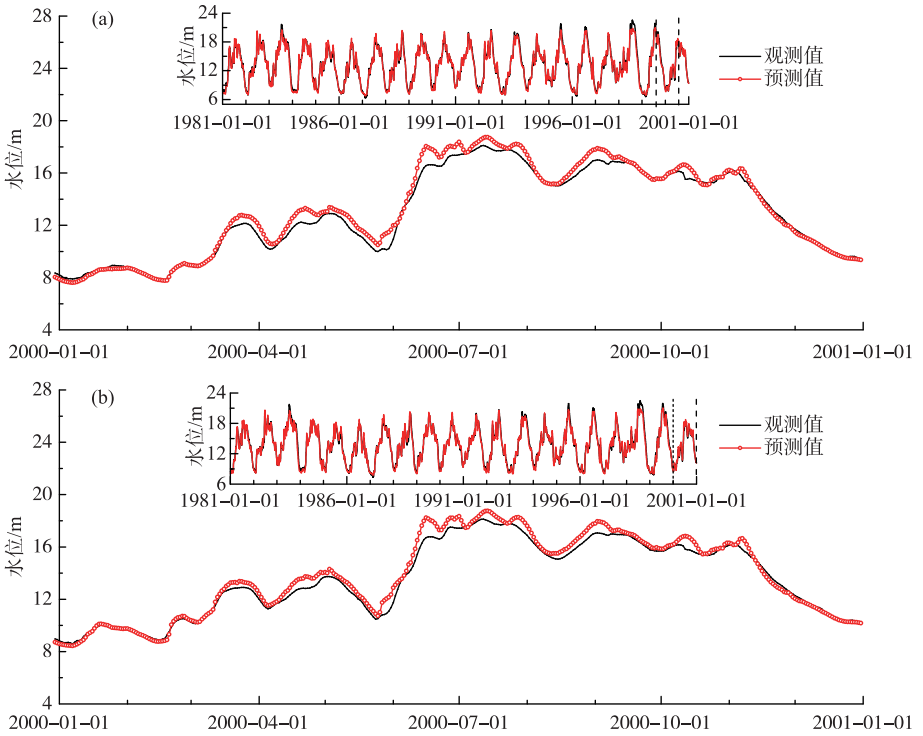

图 7 1981－2000 年湖口站(a)和星子站(b)水位预测结果

Fig.7 Water level prediction results of Hukou station (a) and Xingzi station (b) from 1981 to 2000

络同样是将 1956－1980 年水文序列数据作为训练集,1981－2000 年数据对模型进行验证,利用“五河”六站和长江干流共 7 个流量站前 7 d 内的日均流量,组合预测鄱阳湖从上游到下游 5 个水位站未来 1 d 的日均水位. 比较了两种方法的各项评价指标(表 2).

表 2 两种神经网络方法验证阶段评价指标计算结果对比

Tab.2 Comparison of evaluation index calculation results of two methods in the testing stage

模型	评价指标	湖口站	星子站	都昌站	康山站	吴城站
LSTM	*RMSE*/m	0.50	0.49	0.46	0.41	0.50
	NSE	0.98	0.98	0.97	0.96	0.96
	R^2	0.98	0.98	0.97	0.96	0.97
	计算速度/s			21.66		
BPNN	*RMSE*/m	0.85	0.86	0.87	0.77	0.82
	NSE	0.95	0.94	0.91	0.84	0.90
	R^2	0.95	0.94	0.92	0.87	0.91
	计算速度/s			138.82		

对五站水位的预测,LSTM 模型得到的最佳 *RMSE* 值均低于 0.50 m,而 BPNN 模型得到的最佳 *RMSE* 值除最上游康山站为 0.77 m,其他四站的 *RMSE* 值均高于 0.80 m;LSTM 模型得到的 *NSE*,预测精度最低的康山站水位也达到 0.96,而 BPNN 得到的 *NSE*,预测精度最高的湖口站水位也仅有 0.95. 不难看出,我们构建的 LSTM 模型模拟精度明显强于 BPNN 模型. 除了模拟精度外,计算速度也是衡量模型性能的一个重要指标. 计算了两种模型的计算速度,结果表明,LSTM 模型的计算速度(21.66 s)明显快于 BPNN 的计算速度(138.82 s). 综上,构建的 LSTM 模型在精度和速度上较 BPNN 模型均具有明显的优势.

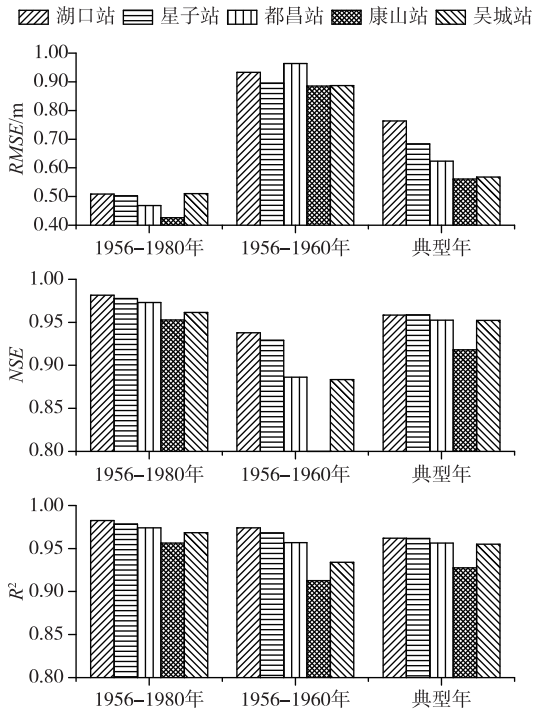

图 8 三种训练数据集下模型验证结果比较

Fig.8 Comparison of model verification results under three training data sets

3.3 训练数据集的影响

实际情况中,应尽可能选取足够多的数据来建模,但由于各种不可抗因素,往往无法获取完整的水文时间序列数据. 为此,我们需要考虑选用不同的训练数据集,分析模型的效果. 我们基于水文时间序列数据自身的特点,利用上述参数设计好的LSTM 模型,设计 3 种方案分别对五站水位进行预测,分别为:①训练集数据为长时间序列(1956－1980 年);②训练集数据较短时间序列(1956－1960 年);③训练集数据时间序列为 5 个典型年:1954 年、1973 年、1974 年、1977 年和 1978 年. 根据《水文情报预报规范》(GB/T 22482－2008)中的距平百分率划分径流丰平枯的标准,1954 年和 1973 年是典型的丰水年,1977 年是平水年,1974 年和 1978 年为典型的枯水年. 利用 1981－2000 年的数据进行验证,比较 3 种训练集数据模式下模型拟合的效果.

五站水位预测效果的各项评价指标计算结果如图 8 所示,同种预测方案中,湖口站水位预测效果均最好,康山站均最差. 方案①中,湖口站验证阶段 *RMSE*、*NSE* 和 R^2分别为 0.51 m、0.98 和 0.98;康山站相应的指标计算结果分别为 0.43 m、0.95 和 0.95. 方案②中,五站水位预测的效果明显降低. 湖口站 *RMSE*、*NSE* 和 R^2分别为 0.93 m、0.94 和 0.97;康山站相应的指标计算结果分别为 0.88 m、0.80 和 0.91. 方案①和方案②中训练集的样本量分别为 9126 和 1827. 而将样本量为 1820 个的 5 个典型年的数据作为训

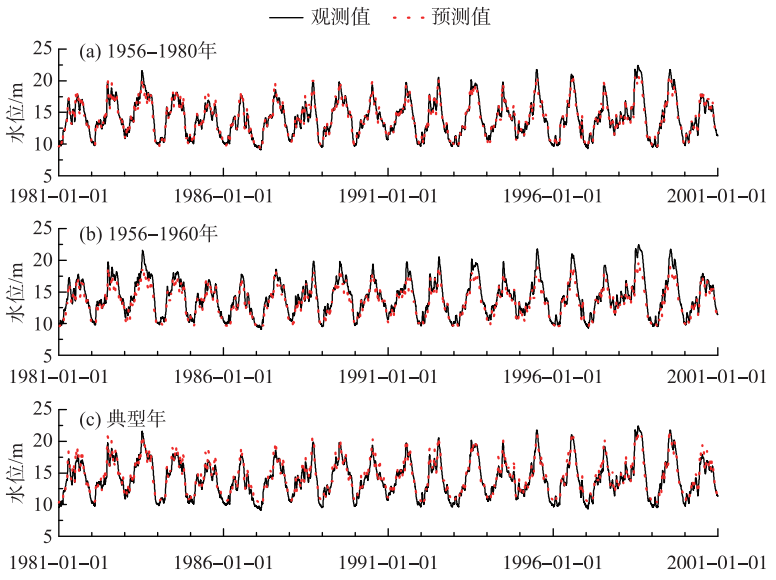

图 9 三种训练集训练模式下都昌站 1981－2000 年水位预测结果

Fig.9 Water level prediction results of Duchang station from 1981 to 2000 under three training sets

练数据进行模型训练时,五站水位的预测效果与方案①相比较差,但明显优于方案②的预测效果. 湖口站 *RMSE*、*NSE* 和 R^2 分别为 0.76 m、0.96 和 0.96;康山站相应的指标计算结果分别为 0.56 m、0.92 和 0.93. 所以,我们在利用 LSTM 神经网络预测湖泊水位时,应尽可能选取足够长时间序列的数据. 若因为不可抗因素无法获取完整的数据序列,选取涵盖各种代表性数据的典型年数据进行训练,也可以获得较好的模型预测效果.

都昌站代表鄱阳湖主湖区水位,3 种方案下其水位预测的结果如图 9 所示(其他水位站预测结果与其相似,故图省略),当 1956－1980 年数据作为训练时间序列时,五站多年水位的预测值与真实值之间具有非常高的对应关系. 1956－1960 年是鄱阳湖典型的枯水期,因而模型通过训练能够较好地反映低水位特征. 5 个典型年涵盖了几个重要的丰、枯水年信息,总体预测效果介于前两者之间.

4 参考文献

[1] Hui FM, Xu B, Huang HB *et al*. Modelling spatial-temporal change of Lake Poyang using multitemporal Landsat imagery. *International Journal of Remote Sensing*, 2008, **29**(20): 5767-5784. DOI: 10.1080/01431160802060912.

[2] Kanai Y, Ueta M, Germogenov N *et al*. Migration routes and important resting areas of Siberian cranes (*Grus leucogeranus*) between northeastern Siberia and China as revealed by satellite tracking. *Biological Conservation*, 2002, **106**(3): 339-346. DOI: 10.1016/s0006-3207(01)00259-2.

[3] Liu XY, Guan YN, Guo S *et al*. Vegetation distribution and water level change response of Poyang Lake Wetland based on time series harmonic analysis. *J Lake Sci*, 2016, **28**(1): 195-206. DOI: 10.18307/2016.0123. [刘旭颖, 关燕宁, 郭杉等. 基于时间序列谐波分析的鄱阳湖湿地植被分布与水位变化响应. 湖泊科学, 2016, **28**(1): 195-206.]

[4] Ouyang QL, Liu WL. Study on the variation characteristics of Poyang Lake water level in the past 50 years. *Resources and Environment of the Yangtze River Basin*, 2014, (11): 1545-1550. DOI: 10.11870/cjlyzyyhj201411009. [欧阳千林, 刘卫林. 近 50 年鄱阳湖水位变化特征研究. 长江流域资源与环境, 2014, (11): 1545-1550.]

[5] Min Q. Changes in the shape and water regime of Poyang Lake in the past 50 years and its relationship with reclamation. *Advances in Water Science*, 2000, **11**(1): 76-81. [闵骞. 近 50 年鄱阳湖形态和水情的变化及其与围垦的关系. 水科学进展, 2000, **11**(1): 76-81.]

[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] Ye XC, Zhang Q, Bai L *et al*. A modeling study of catchment discharge to Poyang Lake under future climate in China. *Quaternary International*, 2011, **244**(2): 221-229. DOI: 10.1016/j.quaint.2010.07.004.

[8] Lai XJ, Huang Q, Zhang YH *et al*. Impact of lake inflow and the Yangtze River flow alterations on water levels in Poyang Lake, China. *Lake and Reservoir Management*, 2014, **30**(4): 321-330. DOI: 10.1080/10402381.2014.928390.

[9] Zhang Q, Liu JY, Singh VP *et al*. Evaluation of impacts of climate change and human activities on streamflow in the Poyang Lake basin, China. *Hydrological Processes*, 2016, **30**(14): 2562-2576. DOI: 10.1002/hyp.10814.

[10] Lai XJ, Jiang JH, Liang QH *et al*. Large-scale hydrodynamic modeling of the middle Yangtze River Basin with complex river-lake interactions. *Journal of Hydrology*, 2013, **492**: 228-243. DOI: 10.1016/j.jhydrol.2013.03.049.

[11] Nourani V, Baghanam AH, Adamowski J *et al*. Applications of hybrid wavelet-Artificial Intelligence models in hydrology: A review. *Journal of Hydrology*, 2014, **514**: 358-377. DOI: 10.1016/j.jhydrol.2014.03.057.

[12] Hu Q, Feng S, Guo H *et al*. Interactions of the Yangtze river flow and hydrologic processes of the Poyang Lake, China. *Journal of Hydrology*, 2007, **347**(1/2): 90-100. DOI: 10.1016/j.jhydrol.2007.09.005.

[13] Zhao GJ, Hörmann G, Fohrer N *et al*. Streamflow trends and climate variability impacts in Poyang lake basin, China. *Water Resources Management*, 2010, **24**(4): 689-706. DOI: 10.1007/s11269-009-9465-7.

[14] Guo H, Hu Q, Zhang Q *et al*. Effects of the three gorges dam on Yangtze river flow and river interaction with Poyang Lake, China: 2003-2008. *Journal of Hydrology*, 2012, **416/417**: 19-27. DOI: 10.1016/j.jhydrol.2011.11.027.

[15] 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**: 83-95. DOI: 10.1016/j.jhydrol.2013.04.036.

[16] Zhang Q, Ye XC, Werner AD *et al*. An investigation of enhanced recessions in Poyang Lake: Comparison of Yangtze River and local catchment impacts. *Journal of Hydrology*, 2014, **517**: 425-434. DOI: 10.1016/j.jhydrol.2014.05.051.

[17] Li XH, Zhang Q, Xu CY. Suitability of the TRMM satellite rainfalls in driving a distributed hydrological model for water balance computations in Xinjiang catchment, Poyang lake basin. *Journal of Hydrology*, 2012, **426/427**: 28-38. DOI: 10.1016/j.jhydrol.2012.01.013.

[18] 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.

[19] Zhang D, Lindholm G, Ratnaweera H. Use long short-term memory to enhance internet of things for combined sewer overflow monitoring. *Journal of Hydrology*, 2018, **556**: 409-418. DOI: 10.1016/j.jhydrol.2017.11.018.

[20] Zhang JF, Zhu Y, Zhang XP *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.

[21] Lee GH, Jung SH, Lee DE. Comparison of physics-based and data-driven models for streamflow simulation of the Mekong river. *Journal of Korea Water Resources Association*, 2018, **51**(6): 503-514.

[22] Wu YR, Liu ZY, Xu WG *et al*. Context-Aware Attention LSTM Network for Flood Prediction. 24th International Conference on Pattern Recognition (ICPR), 2018. Beijing. New York, USA: IEEE, 2018. DOI: 10.1109/icpr.2018.8545385.

[23] Dai X, Wan R, Yang G. Non-stationary water-level fluctuation in China's Poyang Lake and its interactions with Yangtze River. *Journal of Geographical Sciences*, 2015, **25**(3): 274-288. DOI: 10.1007/s11442-015-1167-x.

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

[25] Gers FA, Schmidhuber J, Cummins F. Learning to forget: Continual prediction with LSTM. *Neural Computation*, 2000, **12**(10): 2451-2471. DOI: 10.1162/089976600300015015.

[26] Wythoff BJ. Backpropagation neural networks. *Chemometrics and Intelligent Laboratory Systems*, 1993, **18**(2): 115-155. DOI: 10.1016/0169-7439(93)80052-j.

[27] Wen Y, Zhang K, Li Z *et al*. A Discriminative feature learning approach for deep face recognition. *Lecture Notes in Computer Science*, 2016: 499-515. DOI: 10.1007/978-3-319-46478-7_31.

[28] Nash JE, Sutcliffe JV. River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 1970, **10**(3): 282-290. DOI: 10.1016/0022-1694(70)90255-6.

[29] Sønderby SK, Sønderby CK, Maaløe L *et al*. Recurrent Spatial Transformer Networks. arXiv preprint, 2015. arXiv: 1509.05329.

[30] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint, 2015. arXiv: 1502.03167.

[31] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, **60**(6): 84-90. DOI: 10.1145/3065386.

[32] Srivastava N, Hinton G, Krizhevsky A *et al*. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 2014, **15**(1): 1929-1958.

[33] Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. *Computer Science*, 2014: 1-15.

[34] Zhang XL, Zhang Q, Wang XL. Analysis of nonlinear characteristics of water level-flow relationship in floodplain. *Resources and Environment of the Yangtze River Basin*, 2017, **26**(5): 78-84. [张小琳, 张奇, 王晓龙. 洪泛湖泊水位-流量关系的非线性特征分析. 长江流域资源与环境, 2017, **26**(5):78-84.]

[35] Zhang D, Lin JQ, Peng QD *et al*. Modeling and simulating of reservoir operation using the artificial neural network, support vector regression, deep learning algorithm. *Journal of Hydrology*, 2018, **565**: 720-736. DOI: 10.1016/j.jhydrol.2018.08.050.

[36] Xue LQ, Cui GB, Chen KQ. Dynamic water level neural network prediction model for non-stationary time series. *J Lake Sci*, 2002, **14**(1): 19-24. DOI: 10.18307/2002.0103. [薛联青, 崔广柏, 陈凯麒. 非平稳时间序列的动态水位神经网络预报模型. 湖泊科学, 2002, **14**(1): 19-24.]

[37] Smith LN. Cyclical Learning rates for training neural networks. *Computer Science*, 2015: 464-472. DOI: 10.1109/WACV.2017.58.