

# Simulation of Ankang Reservoir Inflow Based on LSTM Model

Wei He<sup>1</sup>, Jun Wan<sup>2\*</sup>, Yuanzhong Zuo<sup>1</sup>, Jingtao Shi<sup>1</sup>, Jinjin Wu<sup>2</sup>, Jiong Guo<sup>2</sup>

<sup>1</sup>Ankang Hydropower Station, Ankang Shaanxi

<sup>2</sup>State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan Hubei  
Email: \*52jkg1.163.com

Received: Dec. 5<sup>th</sup>, 2019; accepted: Jan. 16<sup>th</sup>, 2020; published: Feb. 10<sup>th</sup>, 2020

## Abstract

The underlying surface conditions have changed greatly in the Ankang river basin due to human activities, and the error of the flood forecasting system is relatively large. Considering that hydrological data is a complex time series, it is difficult for ordinary hydrological models to capture its changing laws. LSTM (Long-short term memory), as a learning network with memory ability, can learn the complex and changeable time of hydrological data well by continuously inputting new data and learning the main features and changes of time series sequence. The LSTM network model is used to simulate the flood process of Ankang reservoir and compare with the Xinanjiang model. The adaptability of LSTM model in reservoir flood prediction is also discussed.

## Keywords

Flood Simulation, Long Short-Term Memory (LSTM), Ankang Reservoir

# 基于LSTM模型模拟安康水库洪水过程

何伟<sup>1</sup>, 万俊<sup>2\*</sup>, 左园忠<sup>1</sup>, 石静涛<sup>1</sup>, 吴金津<sup>2</sup>, 郭灵<sup>2</sup>

<sup>1</sup>安康水力发电厂, 陕西 安康

<sup>2</sup>武汉大学水资源与水电工程科学国家重点实验室, 湖北 武汉

Email: \*52jkg1.163.com

收稿日期: 2019年12月5日; 录用日期: 2020年1月16日; 发布日期: 2020年2月10日

## 摘要

安康流域受人类活动等因素影响, 下垫面条件发生较大变化, 洪水预报系统的误差较大。考虑到水文数据是一

作者简介: 何伟(1977-), 男, 陕西西乡人, 高级工程师, 从事水电运行及检修管理。

\*通讯作者。

种复杂的时间序列, 普通的水文模型难以捕捉其变化规律, LSTM (长短期记忆)作为一种有记忆能力的学习网络模型, 通过不断输入的新数据, 学习时间序列的主要特征和变化趋势, 能很好地学习水文数据这种复杂的多变的时间序列。本文利用LSTM网络模型对安康水库洪水过程进行模拟, 并与新安江模型进行比较分析, 探讨LSTM网络模型在水库洪水预报的适应性。

## 关键词

洪水模拟, 长短期记忆, 安康水库

Copyright © 2020 by author(s) and Wuhan University.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 概述

### 1.1. 研究背景

安康水库洪水预报系统于 2008 年建成, 洪水预报模型采用的是三水源新安江模型。近 10 年来流域内人类活动频繁, 自然条件变化大, 已有的新安江洪水预报模型的预报精度越来越低, 难以满足现在安康水库洪水预报精度的要求。考虑到水文数据具有高度的时间不确定性, 属于复杂的非线性时间序列[1], 传统的线性模型一般用于平稳序列的预报或模拟, 因而在水文预报中的表现往往不尽人意, 即使是像新安江三水源模型这种非线性结构的蓄满产流模型也只是对应部分固定地区有较好的模拟精度。

随着信息技术的高速发展, 流域水文气象资料的观测取得了长足进步。与此同时, 流域产汇流理论却尚未取得重大突破。在这种背景下, 一些学者试图用数据驱动的方式来解决流域产汇流问题。人工神经网络(Artificial Neural Networks, ANN)作为一种适用于拟合高度非线性系统的算法, 早在 20 世纪 90 年代初期就被应用于降雨径流模拟当中。递归神经网络 Recurrent Neural Networks, RNN)作为人工神经网络的一种, 由于其结构特别适用于时间序列数据的关系模拟, 也被应用在流域降雨产流模拟的研究中。

本文以安康水库洪水过程模拟作为研究对象, 尝试构建有时间序列记忆能力的 LSTM 神经网络模型来模拟安康水库洪水过程, 以便研究 LSTM 神经网络模型在安康水库的适用性。

### 1.2. 研究对象

安康水库位于汉江上游陕西省境内, 是一座以发电为主, 兼顾防洪、航运等综合利用的大型水利枢纽。坝址位于安康市城西 18 km 处, 控制流域集水面积 35,700 km<sup>2</sup>, 且坝址下游与安康市区间有月河加入[2]。地理上属南北过度的亚热带地区, 夏季受太平洋副高控制, 多暴雨。由于流域内水系条件较为特殊, 常遇降雨往往形成稀遇洪水。

安康水库为季调节水库, 总库容为 32 亿 m<sup>3</sup>, 其中死库容为 11.13 亿 m<sup>3</sup>, 兴利库容为 14.72 亿 m<sup>3</sup>, 最大调洪库容为 9.8 亿 m<sup>3</sup>; 水库死水位为 305 m, 正常蓄水位为 330 m, 防洪限制水位 325 m, 设计洪水位 333.10 m, 校核洪水位 337.05 m, 水库预留有 3.6 亿 m<sup>3</sup> 的防洪库容。安康水库上游修有石泉季调节水库(1975 建成), 占流域面积 60%; 安康 - 石泉区间的两个大一点支流上, 其中: 岚河干流上修建有蔺河口季水库(2003 建成), 任河干流上修建有巴山年调节水库(2008 建成), 而这两个支流正是暴雨中心, 降雨产流复杂。

本文需要的原始资料有: 遥测站或水文站日雨量资料、安康日入库流量资料, 具体如下: 安康现有遥测站 33 个, 日雨量资料统计不全, 这 33 个雨量站都分布在石泉 - 安康区间, 其日雨量资料安康仅提供了 2001~2017

年的, 暂缺石泉日雨量资料, 最后收集国家气象局的安康上游以上流域 8 个雨量站(57,028 太白、57,124 留坝、57,127 汉中、57,134 佛坪、57,211 宁强、57,232 石泉、57,238 镇巴、57,245 安康)的日雨量数据作为本项目研究的雨量数据, 系列长度为 1961~2017 年(来自国家气象信息中心的中国气象数据网), 8 个气象站点布置如图 1 所示。安康水库的入库资料由安康水电站提供, 系列长度为 1991~2017 年。

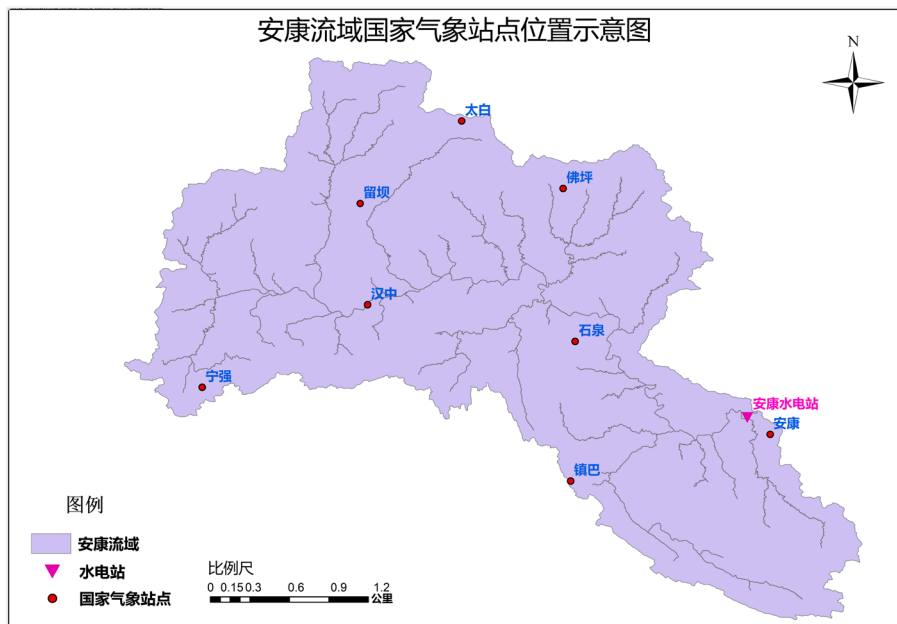


Figure 1. Location map of the meteorological and hydrological stations in the Ankang River basin  
图 1. 安康流域气象水文站点位置示意图

## 2. 模型介绍

### 2.1. LSTM 模型

Hochreiter 和 Schmidhuber 引入长短期记忆(LSTM)神经网络, 是一种特殊的循环神经网络(Recurrent Neural Network, RNN) [3], 其结构如图 2 所示, 许多人进行了改进和普及, 解决了 RNN 的长期记忆能力上的缺陷。LSTM 神经网络通过在神经网络的隐藏层(Hidden layer)中引入存储单元(Memory cell), 即输入门(Input gate)、忘记门(Forget gate)、内部回馈连结(Self-recurrent connection)、和输出门(Output gate)来选择记忆当前信息或遗忘过去记忆信息, 以增强神经网络的长期记忆能力[4]。

原始的 RNN 神经网络包含输入层、输出层和隐含层, 是一种模拟时间序列的递归神经网络。具体表现为计算网络会对前期输出的信息进行记忆并在当前输出中应用[5]。但在实际应用中, 为了降低网络的复杂度, RNN 常常只假设当前节点状态与之前邻近的几个状态有关。

在网络的前向传播过程中, 上一时刻网络的输出及权值对下一层的影响可表示为:

$$h_t = \delta(Ux_t + Wh_{t-1} + b) \quad (1)$$

式中:  $h_{t-1}$  表示  $t-1$  时刻隐含层的输入,  $h_t$  表示  $t$  时刻隐含层的输出,  $x_t$  表示  $t$  时刻输入层的输入,  $b$  表示偏置项,  $\delta$  为激励函数。

因此 RNN 是一个不停向前传递的神经网络, 沿着时间轴方向利用当前时刻输入数据和模型当前状态, 模拟下一时刻目标值。虽然 RNN 的计算过程考虑到前一个计算时刻的状态, 但如时间间距较长, 后面的节点对前面节点的感知将逐渐下降, 将会产生“梯度消失”问题[6]。

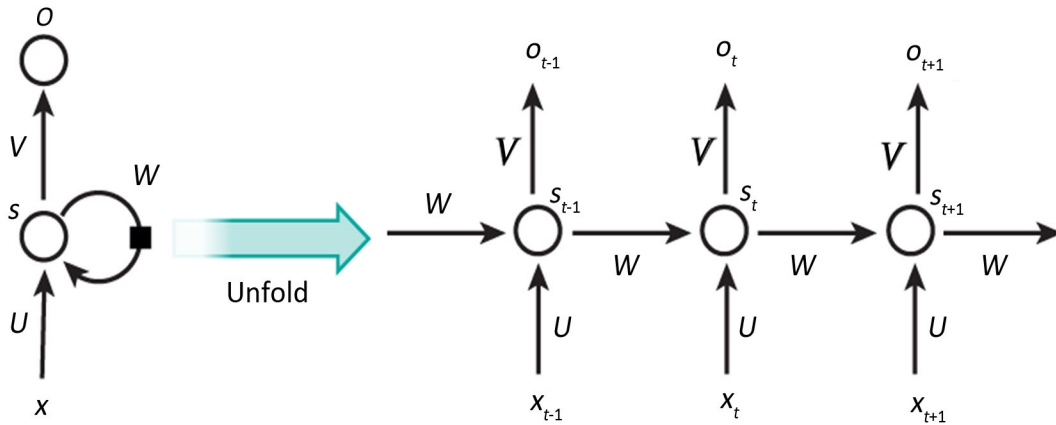


Figure 2. Recurrent neural network structure  
图 2. 循环神经网络结构

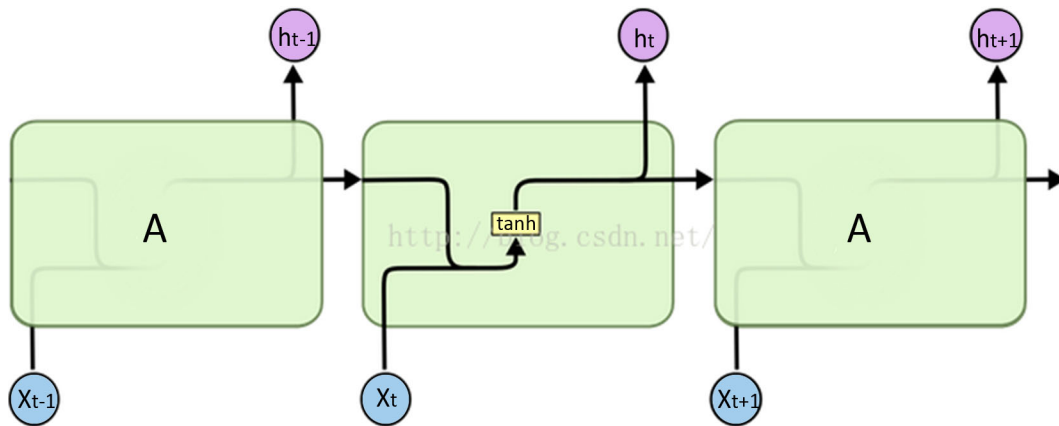


Figure 3. Complete RNN unit diagram  
图 3. 完整的 RNN 单元图

LSTM 神经网络是一种在 RNN (如图 3)基础上改进的递归神经网络,用来解决 RNN 使用过程中所产生的梯度消失、梯度爆炸以及缺乏长期记忆能力等问题。Bengio 等[7]已经表明,传统的 RNN 几乎不能记住长度超过 10 的序列,对于日流量模拟,这就意味着我们只能使用最近 10 天的气象数据作为输入来预测第二天的流量。考虑到地下水,积雪甚至冰川储存等集水区的记忆,这段时间太短,降水和排放之间的滞后时间达到几年。为解决 RNN 在计算过程中出现梯度消失的问题,LSTM 神经网络在原有的 RNN 神经网络结构的基础上,将隐含层的 RNN 计算单元换成结构更加复杂的 LSTM 计算单元,更新了计算网络的节点。LSTM 神经网络通过设置多个不同门的开关实现时间轴上的记忆和遗忘功能,解决梯度消失问题,在每个 LSTM 神经网络计算单元设置 3 个门控制器,分别为输入门、输出门和遗忘门,形成一个新的计算单元。输入门控制特征值的信息输入,遗忘门控制计算单元状态信息的保留,输出门控制信息输出。

为了解释 RNN 和 LSTM 如何的工作,我们将递归神经网络展开为有向无环图(见图 4)有最后 n 个独立变量的连续时间步长(在我们的特征值中,每日降水,最低/最高温度,太阳辐射和蒸气压)构成的输入  $x = [x_1, \dots, x_n]$  预测特定时间步长的输出并按顺序处理。在每一个时间步长  $t(1 \leq t \leq n)$ ,在网络中每层的循环单元中,目前的输入  $x_t$  被处理。

LSTM 神经网络的构建思想是在 RNN 网络的基础上,将 RNN 中每个隐藏计算单元换成具有记忆功能的计算单元。LSTM 层的计算可以表示如下(若干个计算单元组成一个 LSTM 层):

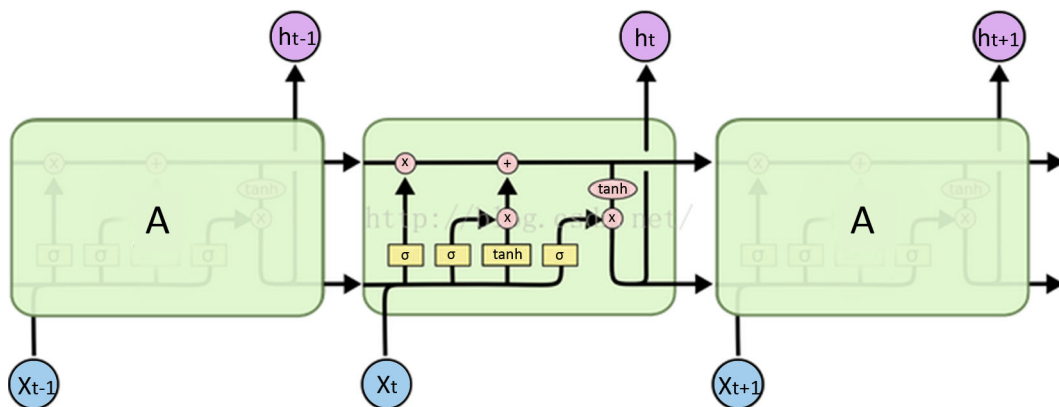


Figure 4. Complete LSTM unit diagram  
图 4. 完整的 LSTM 单元图

$$g^{(t)} = \phi(W_{gx}x^{(t)} + W_{gh}h^{(t-1)} + b_g) \tag{2}$$

$$i^{(t)} = \sigma(W_{ix}x^{(t)} + W_{ih}h^{(t-1)} + b_i) \tag{3}$$

$$f^{(t)} = \sigma(W_{fx}x^{(t)} + W_{fh}h^{(t-1)} + b_f) \tag{4}$$

$$o^{(t)} = \sigma(W_{ox}x^{(t)} + W_{oh}h^{(t-1)} + b_o) \tag{5}$$

$$h^{(t)} = s^{(t)} \tag{6}$$

由以上公式可以看出：输入门、遗忘门和输出门各自对应着一个线性单元，控制计算单元的信息输入和输出以及更新计算单元的状态。LSTM 神经网络采用基于时间的反向传播(BPTT)算法进行训练，该算法与经典反向传播算法类似[6]。采用 Adam 算法进行梯度优化算法，与其他优化算法相比，Adam 算法更适合在实际生产中应用。

神经网络是一种由处理单元相互关联，有自适应能力的，非线性系统。它可以通过神经元的有向连接，建立输入参数到输出参数的映射，这种学习能力可以省略系统建模的步骤，特别适用于规律未知的情况。在 LSTM 神经网络的计算中往往还需要通过堆积 LSTM 隐藏层来模拟较为复杂的水文时间序列[8]。如本文图 5 使用的双层 LSTM 计算流程图。输入前  $n$  个时段的特征值向量，计算第  $n$  个时段的目标值。

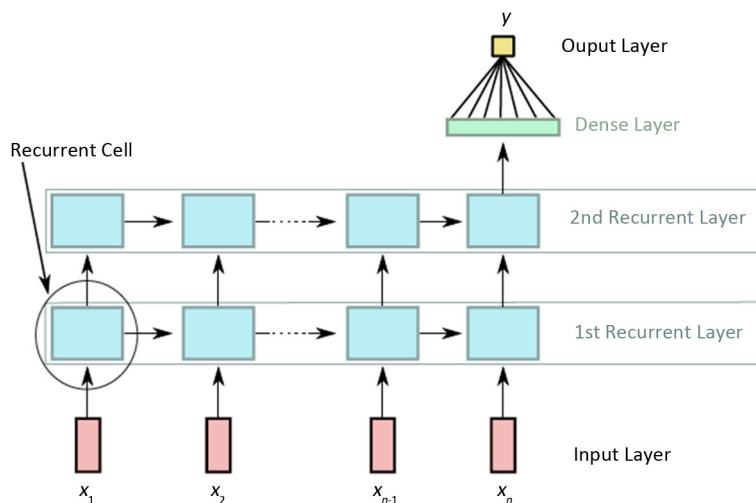


Figure 5. LSTM calculation flow chart  
图 5. LSTM 计算流程图

## 2.2. 新安江模型

新安江三水源模型是一种模拟流域上降雨径流形成的数学函数以及逻辑结构,模型输入的是降雨量和蒸散发能力(常用水面蒸发值),输出的是流域出口断面的洪水过程,广泛在我国南方湿润地区应用,并取得很好的效果[9] [10]。新安江模型是一种适用于我国湿润半湿润地区的概念性流域水文模型[11],其中,蒸散发采用三层模型计算;产流计算选择蓄满产流模型;总径流采用自由水库结构被划分为地表径流、壤中流和地下径流;汇流计算采用线性水库;河道汇流采用马斯京根演算。模型参数主要包括以下 4 类:1) 蒸散发参数;2) 产流模型参数;3) 水源划分参数;4) 汇流模型参数。对于新安江模型而言,其模型的输入变量包括流域面均降水量和蒸发。

## 3. 实例计算与分析

### 3.1. 模型评价指标

基于水文预报的规范以及从安康水库的实际情况出发,选择洪峰流量相对误差和洪量相对误差这两个指标,对两个模型进行对比分析,同时采用纳什模型效率系数评价 LSTM 模型的精度,如公式(7)所示。

$$NS = 1 - \frac{\sum_{t=1}^n (Q_{ot} - Q_{pt})^2}{\sum_{t=1}^n (Q_{ot} - \bar{Q}_o)^2} \quad (7)$$

式中:  $n$  是径流序列的长度,  $Q_{pt}$  和  $Q_{ot}$  分别表示  $t$  时刻径流的预报值和实测值,  $m^3/s$ ;  $\bar{Q}_o$  是实测径流均值,  $m^3/s$ 。该系数越接近 1, 表示模型效果越好。

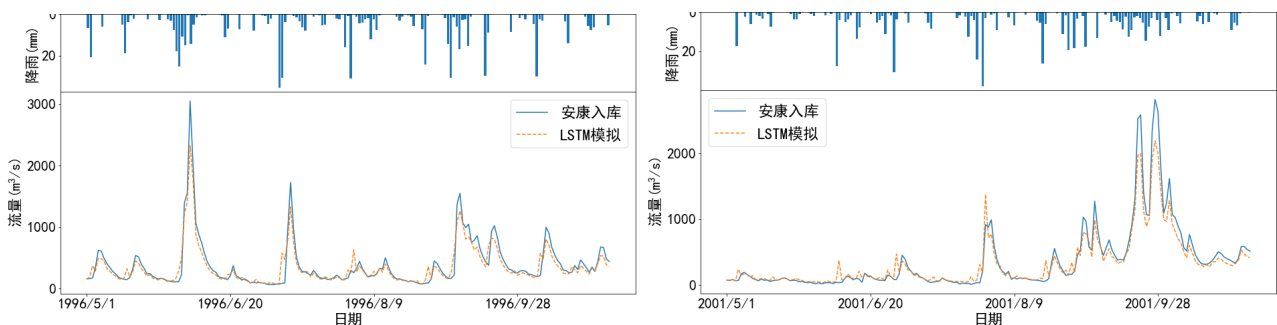
### 3.2. LSTM 模拟计算结果

LSTM 神经网络需要构建多层复杂网络结构,并且选取不同特征值来适配最优模拟目标结果,本文在尝试多种网络结构以及不同的特征值后,最终选取双层 LSTM 网络结构,选取时间、降雨、前一天入库流量作为特征值,模拟安康水库的入库洪水过程。

本文 LSTM 模型选择以日为计算时段,以 1991~2017 年序列资料为计算期,选取 1991~1995 作为率定期,1996~2017 作为检验期。检验期 22 年(1996~2017)实测值与预报值对比, LSTM 模型的 Nash 效率系数为 0.926, LSTM 的模拟效果明显优于新安江模型。图 6 绘制几个代表年汛期降雨径流过程及模拟结果, LSTM 模型模拟的入库径流与实测入库径流总体比较吻合。

### 3.3. 新安江模拟计算结果

同样,新安江模型选择以日为计算时段,以 1991~2017 年序列资料为计算期,选取 1991~1995 作为率定期,1996~2017 作为检验期,新安江模型参数率定结果如表 1 所示,检验期 22 年(1996~2017)实测值与预报值对比,新安江模型的 Nash 效率系数为 0.540。图 7 给出几个代表年汛期新安江模型模拟的汛期过程,由图可见新安江





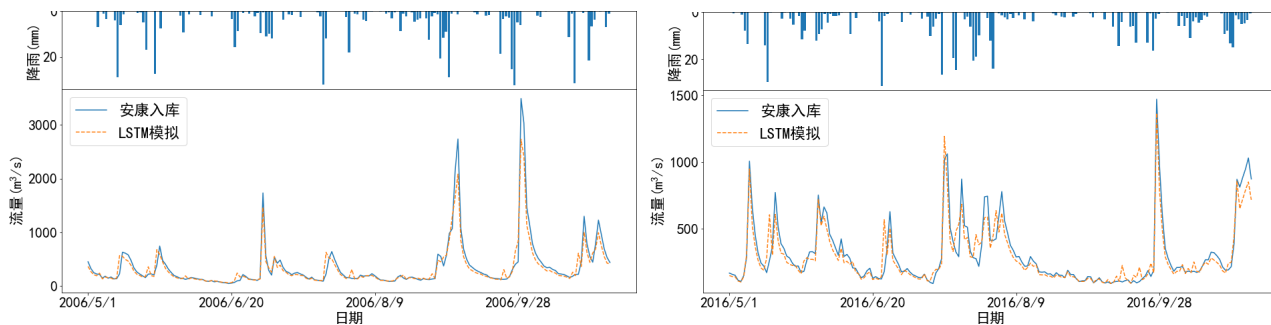


Figure 6. Flood seasonal runoff hydrograph simulated by LSTM model  
图 6. LSTM 模型模拟的汛期径流过程线

Table 1. Calibrated parameters of Xinanjiang model  
表 1. 新安江模型参数率定结果

蒸散发参数				产流参数			分水源参数			
K	WUM	WLM	C	WM	B	Imp	SM	EX	KSS	KG
1.06	28.26	38.40	0.172	72.34	0.74	0.063	35.12	1.96	0.088	0.162

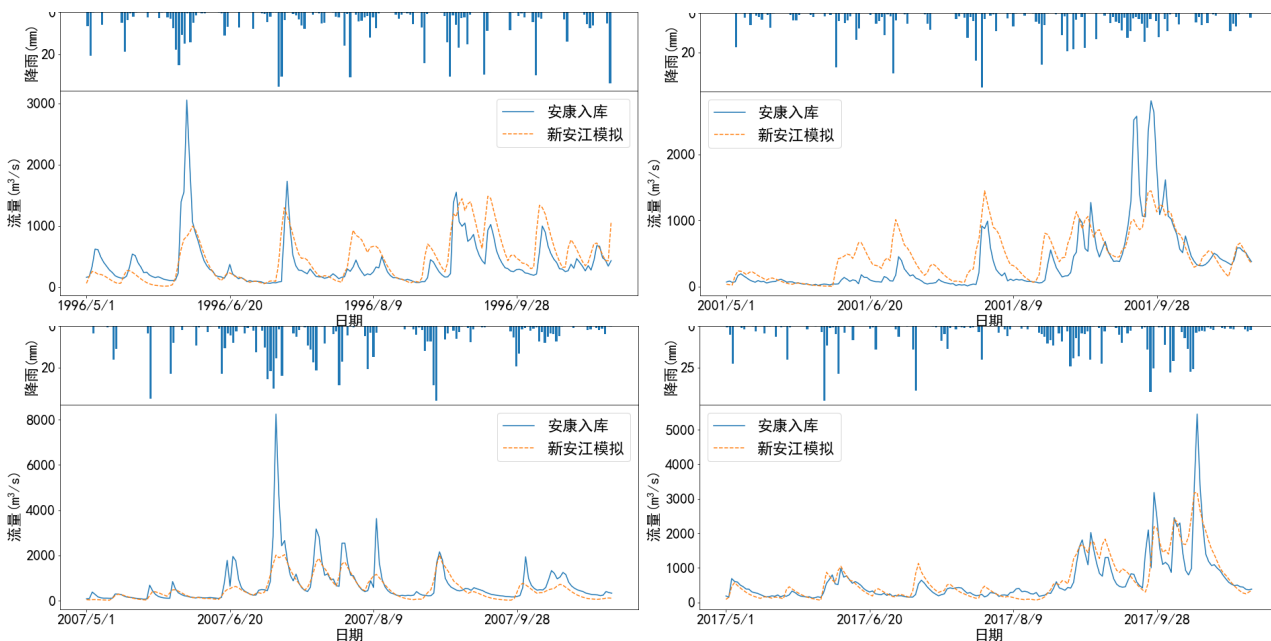


Figure 7. Flood seasonal runoff hydrograph simulated by Xinanjiang model  
图 7. 新安江模型模拟的汛期径流过程

模型模拟的效果不理想，峰值偏低，峰现时间也不太吻合。

### 3.4. 结果对比分析

为了进一步分析 LSTM 模型的实际效果，表 2 给出了检验期 22 年(1996~2017)年模拟和汛期模拟的 Nash 效率系数。LSTM 模型的 Nash 效率系数的大部分模拟结果都在 0.9 以上，而对应的新安江模型的 Nash 效率系数为 0.5 左右。从检验期 22 年(1996~2017)的均值，LSTM 模型和新安江模型的 Nash 效率系数分别为 0.926 和 0.540，可以看出 LSTM 模型的模拟效果明显优于新安江模型。

**Table 2.** Comparison of annual and flood seasonal Nash efficiency coefficients simulated by LSTM model and Xinanjiang model  
**表 2.** LSTM 和新安江模型分别模拟的年和汛期 Nash 效率系数比较

年份	LSTM 模型		新安江模型	
	全年 NS	汛期 NS	全年 NS	汛期 NS
1996	0.908	0.895	0.492	0.283
1997	0.851	0.845	0.255	0.279
1998	0.934	0.926	0.382	0.327
1999	0.948	0.95	0.527	0.555
2000	0.936	0.921	0.659	0.634
2001	0.931	0.924	0.565	0.52
2002	0.915	0.914	0.219	0.18
2003	0.943	0.939	0.612	0.582
2004	0.909	0.909	0.546	0.542
2005	0.953	0.937	0.656	0.632
2006	0.92	0.912	0.568	0.561
2007	0.946	0.938	0.559	0.507
2008	0.907	0.886	0.426	0.337
2009	0.933	0.918	0.666	0.622
2010	0.948	0.938	0.631	0.581
2011	0.93	0.919	0.659	0.626
2012	0.924	0.913	0.568	0.553
2013	0.927	0.911	0.638	0.621
2014	0.953	0.954	0.742	0.748
2015	0.918	0.918	0.418	0.383
2016	0.901	0.869	0.355	0.619
2017	0.932	0.901	0.747	0.732
平均	0.926	0.915	0.540	0.519

由图 6、图 7 可见，就洪峰流量看，两个模型模拟的结果均小于实测流量，但是 LSTM 网络模型的模拟结果更接近实际值；就峰现时间看，新安江模型模拟的洪水过程，峰现时间比实际情况晚，甚至会出现实际上不存在的洪水过程，而 LSTM 模型模拟的洪水峰现时间基本与实际洪水过程吻合，模拟精度较高。显然，从图形总体来看，LSTM 模型的模拟效果远比新安江模型的模拟效果好。

需要说明的是，本文的研究所用到的安康流域降雨站点(图 1 所示的 8 个雨量站点)偏少，导致新安江模型的模拟精度较低。同时也说明，LSTM 模型对流域降雨站点的密集程度要求不那么高，而对历史洪水过程资料整编的精度要求较高。

#### 4. 结论

依据以上的研究计算结果，可以得到以下 2 点结论：

1) LSTM 模型对安康水库洪水过程模拟时，仅使用降雨与径流的历史过程，相对于新安江模型在一定程度



上更简便, 模拟效果也明显优于新安江模型。

2) 在日洪水过程模拟的基础上, 进一步研究 LSTM 模型在时段洪水过程的模拟, 以便将 LSTM 模型应用于安康水库实时洪水预报。

## 参考文献

- [1] 武连洲, 白涛, 哈燕萍, 等. 水文序列变异对水库调度运行的影响研究[J]. 水资源与水工程学报, 2016, 27(4): 88-92.  
WU Lianzhou, BAI Tao, HA Yanping, et al. Effect of hydrological sequence variation on operation reservoir. Journal of Water Resources & Water Engineering, 2016, 27(4): 88-92. (in Chinese)
- [2] 杨光, 郭生练, 陈柯兵, 等. 基于决策因子选择的梯级水库多目标优化调度规则研究[J]. 水利学报, 2017, 48(8): 914-923.  
YANG Guang, GUO Shenglian, CHEN Kebing, et al. Multi-objective cascade reservoir optimal operation rules based on decision factor selection. Journal of Hydraulic Engineering, 2017, 48(8): 914-923. (in Chinese)
- [3] 冯钧, 潘飞. 一种 LSTM-BP 多模型组合水文预报方法[J]. 计算机与现代化, 2018(7): 82-85.  
FENG Jun, PAN Fei. A hydrologic forecast method based on LSTM-BP. Computer and Modernization, 2018(7): 82-85. (in Chinese)
- [4] 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[J]. 水资源研究, 2019(1): 1-12.  
ZHOU Yanlai, GUO Shenglian, CHANG Fi-John, et al. Hydrological forecasting using artificial intelligence techniques. Journal of Water Resources Research, 2019(1): 1-12. (in Chinese)
- [5] 陈卓, 孙龙祥. 基于深度学习 LSTM 网络的短期电力负荷预测方法[J]. 电子技术(上海), 2018, 47(1): 39-41.  
CHEN Zhuo, SUN Longxiang. Short-term electrical load forecasting based on deep learning LSTM networks. Electronics Design & Application, 2018, 47(1): 39-41. (in Chinese)
- [6] BENGIO, Y., SIMARD, P. and FRASCONI, P. Learning long-term dependencies with gradient descent is difficult. IEEE transactions on neural networks, 1994, 5(2): 157-166. <https://doi.org/10.1109/72.279181>
- [7] HOCHREITER, S., SCHMIDHUBER, J. Long short-term memory. Neural Computation, 1997, 9(8): 1735-1780.  
<https://doi.org/10.1162/neco.1997.9.8.1735>
- [8] 江衍铭, 张建全, 明焱. 集合神经网络的洪水预报[J]. 浙江大学学报(工学版), 2016, 50(8): 1471-1478.  
CHIANG Yenming, ZHANG Jianquan, MING Yan. Flood forecasting by ensemble neural networks. Journal of Zhejiang University (Engineering Science), 2016, 50(8): 1471-1478. (in Chinese)
- [9] 钱承萍, 黄川友. 新安江三水源模型与水箱模型在清江流域上的应用与比较[J]. 西北水电, 2013(2): 4-7.  
QIAN Chengping, HUANG Chuanyou. Application and comparison of Xinanjiang model and Tank model in Qingjiang catchment Northwest Hydropower, 2013(2): 4-7. (in Chinese)
- [10] 赵人俊, 王佩兰. 新安江模型参数的分析[J]. 水文, 1988(6): 4-11.  
ZHAO Renjun, WANG Peilan. Analysis of Xinanjiang model parameters. Journal of China Hydrology, 1988(6): 4-11. (in Chinese)
- [11] 刘佩瑶, 郝振纯, 王国庆, 等. 新安江模型和改进 BP 神经网络模型在闽江水文预报中的应用[J]. 水资源与水工程学报, 2017, 28(1): 40-44.  
LIU Peiyao, HAO Zhenchun, WANG Guoqing, et al. Application of Xinanjiang model and the improved BP neural network model in hydrological forecasting of the Min River. Journal of Water Resources and Water Engineering, 2017, 28(1): 40-44. (in Chinese)