

# 基于深度学习的河流径流值预测

张竞予

成都信息工程大学, 四川 成都

收稿日期: 2022年10月23日; 录用日期: 2022年11月18日; 发布日期: 2022年11月28日

## 摘要

在传统预测径流值方法中, 由于径流序列具有非常复杂的非线性特征, 所以无法充分对径流值进行信息提取并进行预测。本文提出一种基于LSTM (长短期记忆网络)和注意力机制的模型来对径流值进行预测。对于气候原因引发的自然灾害——洪水的预测来说, 该模型考虑到降雨量与上游水文站径流值对目标地区的径流值的影响, 并通过加入注意力机制对各个影响因素添加距离影响。以四川省乐山市五通桥区径流值数据为例, 通过对比2015年径流值的预测值和真实值进行对比。验证结果表明: 添加了注意力机制的LSTM模型具有误差小, 准确度高的特性, 可以较为显著提高径流值预测的能力。

## 关键词

径流值预测, LSTM, 注意力机制, 非线性特征

# Deep Learning Based River Runoff Value Prediction

Jingyu Zhang

Chengdu University of Information Engineering, Chengdu Sichuan

Received: Oct. 23<sup>rd</sup>, 2022; accepted: Nov. 18<sup>th</sup>, 2022; published: Nov. 28<sup>th</sup>, 2022

## Abstract

In the traditional method of predicting runoff values, because the runoff sequence has very complex nonlinear characteristics, it is impossible to fully extract and predict the runoff values. In this paper, a model based on LSTM (Long Short-Term Memory Network) and attention mechanism is proposed to predict runoff values. The model takes into account the effects of rainfall and upstream hydrological station runoff values on the runoff values in the target area, and adds distance effects to each influencing factor by adding attention mechanisms. Taking the runoff value data of Wutongqiao District in Leshan City, Sichuan Province as an example, the predicted and true values

of the runoff value in 2015 are compared. The verification results show that the LSTM model with the addition of attention mechanism has the characteristics of small error and high accuracy, which can significantly improve the ability to predict runoff values.

## Keywords

Runoff Value Prediction, LSTM, Attention Mechanisms, Nonlinear Features

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

## 1. 引言

对于内陆地区来说,暴雨是引起的江、河、湖水量迅速增加或水位迅猛上涨的水流现象。当流域内发生暴雨产生径流时,都依其远近先后汇集于河道的出口断面处;或者当近处的径流到达时,河水流量开始增加,水位相应上涨,以及至大部分高强度的地表径流汇集到出口断面时造成的洪水。而洪水泛滥是目前最频繁暴发的自然灾害之一。由于在暴雨过后随时都有可能发生。目前河流洪水是一个全球性的问题,给生活在河边周围的居民带来了严重的问题。四川省乐山市五通桥区位于岷江和大渡河交汇处,发生洪水灾害事件的概率更高。2020年8月18日因发生这种灾害,导致流离失所的家庭数量急剧增加。洪水风险正在由于人口增长,洪泛区的洪水风险正在增加,这个问题是由于周边环境治理差又加上气候变化的影响而更加严重,洪灾不仅仅会破坏当地基础设施建设还会带来严重传染病[1],同时也给居民的生活带来极大负担[2],所以洪水治理显得刻不容缓。而目前能够使用中径流值较为准确预测洪水泛滥是首要选择,它描绘了单位时间内通过水断面的水量。根据径流值能刻画出洪水淹没区域,能够更好的提前起到洪水灾害的预警作用,减少人员伤亡以及财产损失。

为了更好的治理和预防这种严重的自然灾害世界国内外研究人员采用不同水文模型来预测径流值(如基于物理的[3][4]、经验的[5]、概念的[6]和概率模型[7])被世界各地研究人员用于径流值预测。其中黑箱模型[5]是纯经验性的,如果数据较为缺乏,他的预测结果与真实值结果相差较远,且该模型的水文径流值的预测过程完全是由输入以及输出产生的参数之间的度量所决定的。于物理流量模型是预测径流值的方法之一,径流的本质就是河流动力学,物理模型考虑到了河流动力学中物理过程但物理模型[4]要考虑到复杂的物理特性以及流域的动态性质。概念模型[8]的校准过程涉及到集水区的物理解释,需要对集水区的物理知识有深刻的了解,需要对所考虑的集水区的物理特征有深刻的了解。水文过程的复合性质和输入参数的非线性特征使水文过程的复合性质和输入参数的非线性特征使得选择适当的洪水预报模型变得困难。适当的模型进行洪水预报。流域的性质、建模的目的、模型的适当性和输入参数的质量,如降雨量、温度、湿度、土地利用、土地覆盖和时空变化等,都会影响洪水预报的可靠性[6]。

如今,快速发展的机器为我们学习提供了大量技术改进机会。深度学习,是侧重于信息处理的算法,以解决径流模型拟合中非线性部分的性质[9]。人工神经网络(ANN)是一种广泛使用的机器学习技术,已被开发为一种替代水流的水文模型的方法[10]。在[11]中 chenchen 等采用 CNN 网络构造,在各个降雨站点之间采用反距离差分法然后规划  $12.5\text{ m} \times 12.5\text{ m}$  网格送入 CNN 网络,用于预测在目标区域的径流值,但其需要数据量较大,需要大量遥感图片来学习降雨量对目标区域径流值的影响,往往在实际应用上不是一种合适模型。在[12]中作者选择多层感知网络来对目标区域径流值进行预测,作者选取了湿度指数、

温度、降雨量为预测目标区域径流值的指标，将指标分配给网络后进行学习，但是模型没有考虑上游径流值对目标区域的径流影响较其他模型而言准确度较低[3]。作者选取了埃塞俄比亚地区巴罗 - 阿科博基地作为研究区域，该区域常年降雨导致洪水频发，作者使用 ANN 模型来对该区域河流径流值进行预测，其中采用土壤湿度指数(TWI)、降雨量、气温以及观测的径流值输入网络，降雨量的输入采用与[1]相同的放距离加权法，最终输出的测试集径流值预测结果 NSE 值为 0.88 具有较高预测准确度。针对以上问题本问题出一种添加注意力机制的 LSTM 网络对径流值进行预测，在精度上超过了以往的网络，不但添加了上游径流对下流的影响，还考虑了距离对径流值的影响。

在本文中采用 LSTM 网络对目标流域五通桥区径流值进行预测；在这实验过程中发现预测结果在极端值处的效果较差，为了解决这一问题将添加注意力机制来改善。在添加实验结果后得出的准确率较之前在极端值上有较大的提升。相较以往模型的准确率我们的模型在准确率上有提升，对于河道径流值预测有着重大意义。

## 2. 理论基础

### 2.1. LSTM (长短期记忆网络)算法

标准的 RNN 神经网络由于其网络结构的影响，在处理较长序列数据时，由于梯度消失和梯度爆炸等问题，会导致较长时间步的历史信息对输出结果出现迅速衰减或者指数增长等问题。长短期记忆网络的出现是针对这一现象的改进。记忆细胞包含了三个门结构，分别是输入门，遗忘门及输出门，细胞之间传递信息包含当前时间步信息也包含了过去时间步信息。

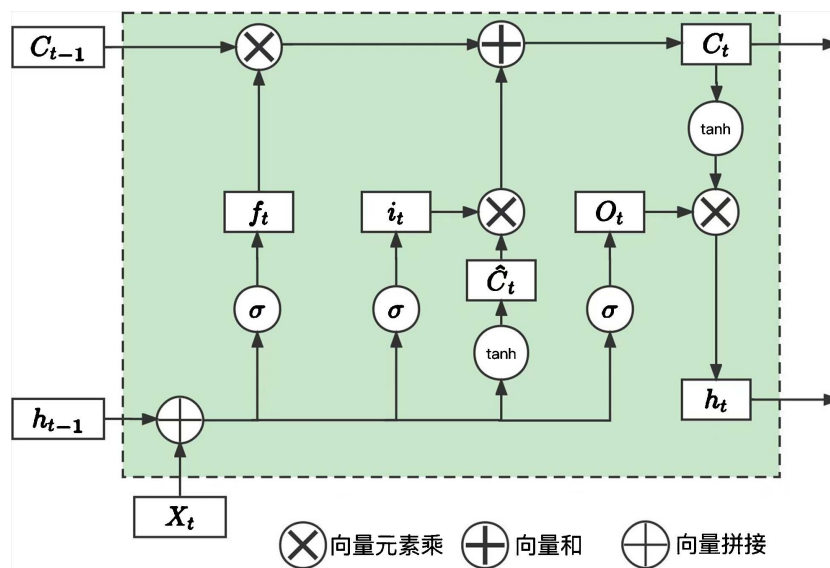


Figure 1. Lstm cell diagram  
图 1. Lstm 细胞图

LSTM 记忆单元状态更新公式如下：

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

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

$$C_t = f_t C_{t-1} + i_t \tanh(W_{xc} X_t + W_{hc} h_{t-1} + b_c) \tag{3}$$

$$O_t = \sigma(W_{x0}X_t + W_{ho}h_{t-1} + wW_{co}C_t + b_o) \quad (4)$$

$$h_t = O_t \tanh(C_t) \quad (5)$$

LSTM 结构(见图 1)中,  $i_t$ ,  $f_t$ ,  $C_t$ ,  $O_t$  分别为  $t$  时刻记忆单元的输入门、遗忘门、细胞状态和输出门;  $X_t$ ,  $h_t$  为  $t$  时刻记忆单元的输入和输出;  $W$  和  $b$  为各个门之间对应的权重矩阵和偏置向量;  $\sigma$  为 sigmoid 激活函数。

LSTM [13]网络组成主要有三个部分遗忘、输入和输出, 遗忘部分主要功能是对上个细胞传入信息进行非重要遗忘, 就是重要信息记住, 非重要的信息选择遗忘; 输入部分主要功能是对当期输入细胞进行选择性记忆, 非重要的信息加以选择性记忆; 输出部分这时候细胞会对信息做出输出处理, 此刻的细胞将会将会决定哪些隐藏状态信息将会传入下个细胞, 然后重复此过程。

## 2.2. 注意力机制

注意力机制[13]网络是对人的注意力的模拟, 通过训练权重可以帮助网络在众多信息中计算选择对当前目标更关键的信息[14]。本文将不同的时刻的 LSTM 隐藏状态注意力值量化为 0~1 之间的权重值, 并且所有隐藏值的权重和为 1。其注意力的输出计算公式如下:

$$F_t = \tanh(f_{t-1} * f_t^T + b_t) \quad (6)$$

$$q_t = \frac{\exp(F_t)}{\sum_{t=1}^T \exp(F_t)} \quad (7)$$

$$F_A = \sum_{t=1}^T q_t F_t \quad (8)$$

$F_A$  为注意力层的输出;  $f_t$  为最后一个时间步输出的隐藏状态  $f_{t-1}$  为其他细胞输出隐藏状态,  $b_t$  为注意力层偏移值。

## 3. 模型结构

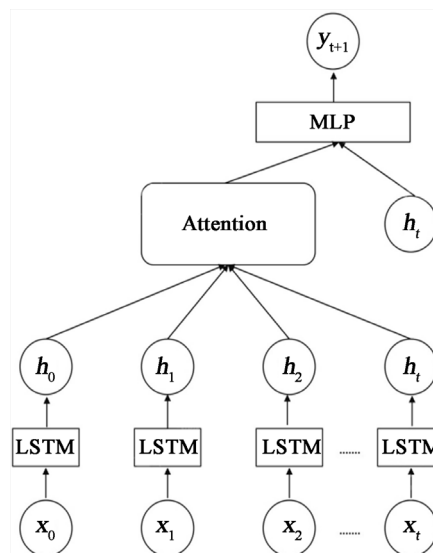


Figure 2. Process diagram

图 2. 流程示意图

模型流程图(见图 2)第一层是输入层包含了输入水文站以及降雨站。输入后到达 LSTM 层在这层中输

出每个 LSTM 细胞的隐藏状态，之后就是注意力机制层，该层的主要作用就是五通桥站点与其他站点做相关性得出注意力机制得分计算结果，最后经过全连接层得出最终预测结果。

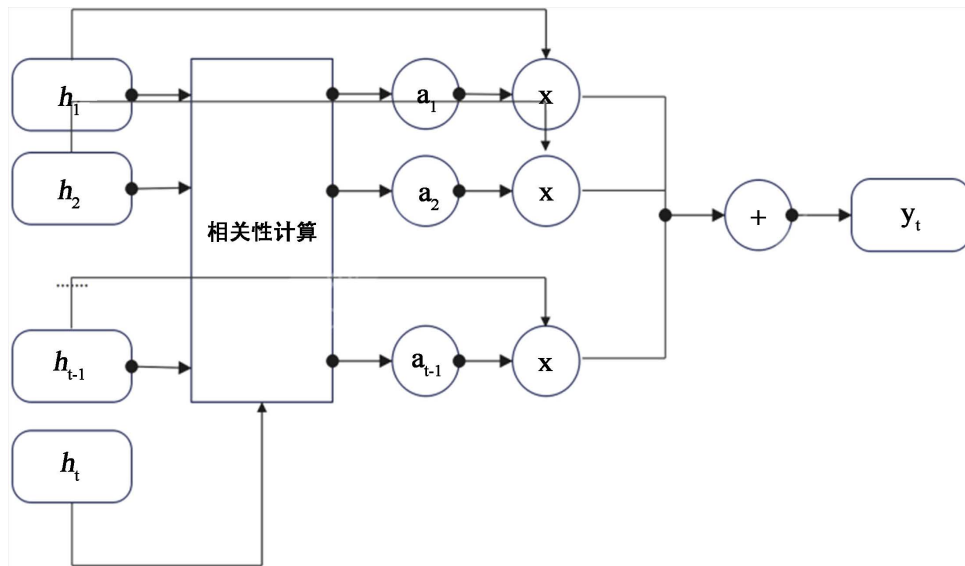


Figure 3. Attention mechanism module  
图 3. 注意力机制模块

在注意力机制模块(见图 3)中  $h_t$  分别与前面的  $h_1$  到  $h_{t-1}$  做相关性计算(如式(9)所示)得出相关性得分系数  $a_1$  到  $a_{t-1}$  在有对应对  $h$  相乘(如式(10))最后再相加得出注意力机制模块  $y_t$  (如式(11)所示)，最终与  $h_t$  一起送入全连接层进行计算，模型的损失函数如图 3 所示。

$$F_t = \tanh(h_{t-1} * h_t^T + b_t) \quad (9)$$

$$a_t = \frac{\exp(F_t)}{\sum_{i=1}^T \exp(F_i)} \quad (10)$$

$$y_t = \sum_{m=1}^{t-1} a_m F_m \quad (11)$$

## 4. 实验及结果分析

### 4.1. 数据及数据预处理

五通桥区位于四川省乐山市，在地理上它位于东经 103° 北纬 29°，全区总面积达 474 平方公里。而五通桥区在河流地理上位于大渡河与岷江交汇处。

本文采用官方获取 2012 年到 2015 年间四川省乐山市的水文站径流值(见表 1)以及降雨站点(见表 2)的降雨量。站点名以及单位如下。

数据归一化，在网络学习中，归一化能极大加快网络学习能够快速收敛同时也防止梯度爆炸和消失。归一化的具体作用是归纳统一样本的统计分布，把数据分布归一在(0, 1)之间概率分布；或者归一在(-1, 1)之间的坐标分布。无论是为了建模还是为了计算，首先基本度量单位要同一，[15]神经网络是以样本在事件中的统计分别几率来进行训练(概率计算)和预测的，且 sigmoid 函数的取值是 0 到 1 之间的，网络最后一个节点的输出也是如此，所以经常要对样本的输出归一化处理。归一化是统一在 0~1 之间的统计概率分布，当所有样本的输入信号都为正值时，与第一隐含层神经元相连的权值只能同时增加或减小，从而

导致学习速度很慢。另外在数据中常存在奇异样本数据, 奇异样本数据存在所引起的网络训练时间增加, 并可能引起网络无法收敛。为了避免出现这种情况及后面数据处理的方便, 加快网络学习速度, 可以对输入信号进行归一化, 使得所有样本的输入信号其均值接近于 0 或与其均方差相比很小。计算公式如下。

**Table 1.** Rainfall station

**表 1.** 降雨站

站名	站点编号	单位
夹江	56382	毫米
峨眉	56384	毫米
峨眉山	56385	毫米
乐山	56386	毫米
沙湾	56388	毫米
井研	56390	毫米
眉山	S6891	毫米

**Table 2.** Hydrometric station

**表 2.** 水文站

站名	站点编号	单位
彭山	60602600	立方米每秒
沙湾	60608800	立方米每秒
夹江	60612000	立方米每秒
五通桥	60600300	立方米每秒

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (12)$$

在公式中  $x_{\min}$  表示要归一化的特征的最小值,  $x_{\max}$  则表示其中的最大值,  $x$  是要归一化的值。

## 4.2. 实验

基于以往模型本文提出一种基于注意力机制的 LSTM 网络来对目标区域的径流值进行预测, 添加注意力机制能将距离对目标区域的影响添加进入网络。且该模型在精度上超越了[4]中的结果, 使得预测值与真实值之间的差距更小。

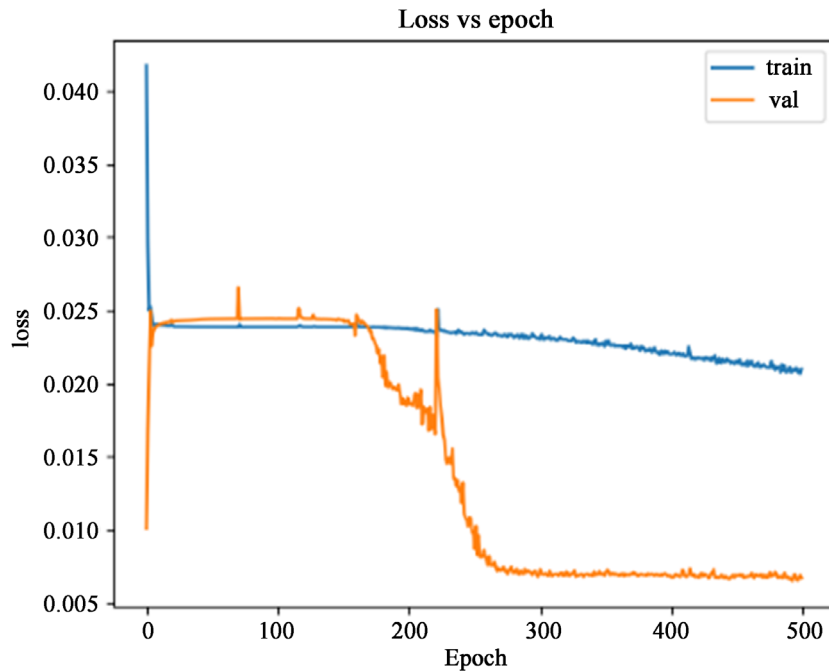
本文中所选取的模型在 RTX3060 型号显卡计算机上进行训练, 训练过程中采用训练的轮数是 500 轮, 在模型中保留训练损失(见图 4)最小的时刻得出的参数作为模型的使用参数。模型的损失函数选取的是 MSE 函数。将五通桥区径流量数据第 25 个时间步作为标签来制作数据集, 制作后的数据没 64 个为一组送入网络进行学习; 在全连接层选用 sigmoid 作为激活函数, 而输出层则是 Relu 为激活函数。

为了验证模型的准确性文中采用了 NSE(纳什效率系数)、MSE 和 RMSE 等评价指标。

## 4.3. 模型评价

### 4.3.1. Nash-Sutcliffe 效率

主要作用就是量化模型的精准度(常用于验证水文模型模拟结构的好坏)。



**Figure 4.** Loss curve, in which the network gets the optimal parameters 300 rounds ago  
**图 4.** Loss 曲线图, 图中网络在 300 轮前得到最优参数

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_o - Q_s)^2}{\sum_{i=1}^n (Q_o - Q_{mean})^2} \quad (13)$$

评价表如下。

**Table 3.** NSE evaluation form  
**表 3.** NSE 评价表

拟合度	NSE
优	$0.75 < NSE < 1$
良	$0.65 < NSE < 0.75$
中	$0.5 < NSE < 0.65$
劣	$NSE \leq 0.5$

NSE 评价表(见表 3), NSE 是预测模型对变量的预测值,  $Q_o$  是观察到的排放( $m^3/s$ ),  $Q_s$  是模拟排放( $m^3/s$ ),  $Q_o$  是平均排量( $m^3/s$ )。NSE 取值为负无穷至 1, NSE 接近 1, 表示模式质量好, 模型可信度高; NSE 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; NSE 远远小于 0, 则模型是不可信的。

#### 4.3.2. MSE 和 RMSE

MSE(均方误差)是参数的估计值与参数的真实值之差平方的期望; MSE 的值越小说明实验预测的结构约精确

$$MSE = \frac{1}{N} \sum_{i=1}^N (Q - q)^2 \quad (14)$$



RMSE(均方根误差)是参数的估计值与参数的真实值之差平方的根号的期望

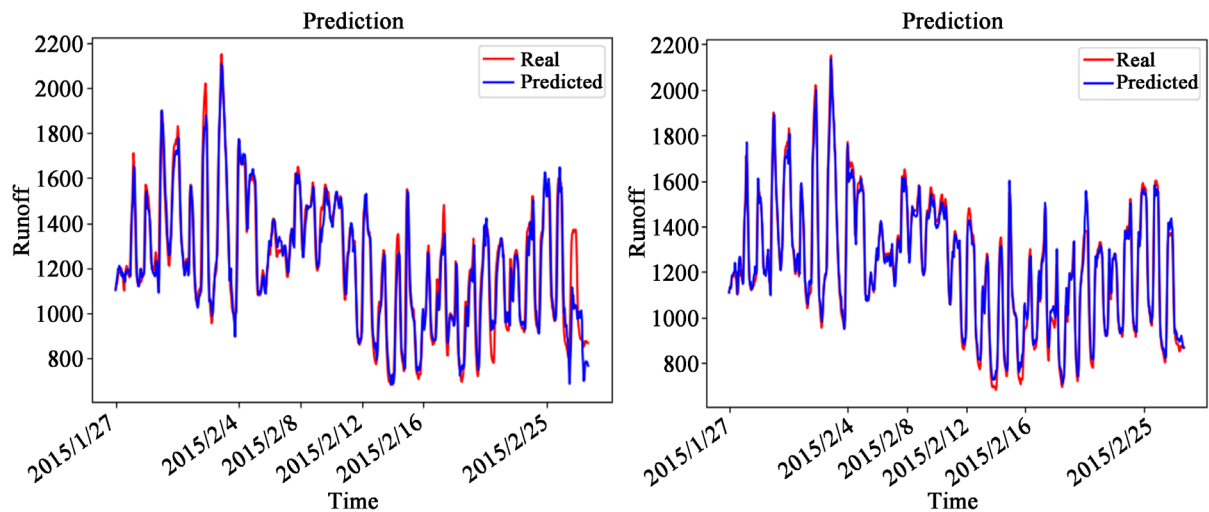
$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (Q - q)^2} \quad (15)$$

#### 4.4. 模型评价

实验数据采用了四川省五通桥区 2015 年 1 月 27 日到 2015 年 2 月 28 日小时数据用于验证模型的准确性, 模型真实值与预测值对比图如下。

**Table 4.** NSE value comparison before and after adding attention mechanism  
**表 4.** 添加注意力机制前后 NSE 值对比

模型	LSTM	LSTM + ATTENTION
NSE	0.76	<b>0.91</b>



**Figure 5.** Comparison chart of predicted value and real value  
**图 5.** 预测值与真实值对比图

对比图中(见图 5)左图为未添加注意力机制前预测值与真实值的对比图, 右图为添加了注意力机制后的预测值与真实值的对比图, 结果在测试集上添加注意力机制前的 NSE 值为 0.76 添加注意力机制后的 NSE 值为 0.91 (见表 4), 结果在精度上有较大提升。

**Table 5.** NSE value comparison between the model and other models  
**表 5.** 模型与其他模型 NSE 值对比

模型	LSTM + ATTENTION	MLP	SWAT-ANN	SWAT-MLP/WOA
NSE	<b>0.91</b>	0.88	0.79	0.81

在以往径流值预测中[4] MLP、[16] SWAT-ANN 和 SWAT-MLP/WOA 在测试集上 NSE 值为 0.88 和 0.79 以及 0.81, 准确度有明显的提升(见表 5)。其中 MSE 值为 8793, RMSE 值为 93.77。在未添加注意力机制前预测结果在较大极端值上预测的结果较差, 由于模型在没有考虑各个站点由于距离因素各个站点目标河流区站点径流值得影响; 而在添加注意力机制之后考虑到了水文站点和雨量站点距离对目标区域的影响, 改善了极端值的影响, 充分考虑了距离对目标河流区域站点的影响, 使得结果更加具有有



效性, 真实值与预测值对比图如图 5 所示。

## 5. 结束语

本文利用了四川省乐山市 12 年到 15 年降雨数据以及水文历史数据, 提出了一个基于 LSTM 网络结构的河流径流值预测模型, 该模型通过考虑了上游水文站点以及降雨站点对目标五通桥站点径流值的影响, 通过加入注意力机制增加了距离对目标区域径流值的影响, 该模型充分考虑了降雨的时空特征。输入为每小时站点数据进入一个 LSTM 细胞, 然后进行注意力机制层计算最终通过全连接层输出运算结果。实验中 2012 年到 2014 年数据参与训练得出最优参数解, 而 2015 年数据用作验证, 通过对比真实数据以及预测数据得出本文构建河流径流值模型符合预测要求。利用该模型能够有效的预测目标河流的径流值, 在准确度上较以往模型有着明显的提升, 这样对于日后洪水灾害的预防以及治理有着至关重要的意义。

在未来工作中将引用更多特征例如蒸发、植被、土壤和其他因素, 以进一步提高径流值预测能力, 同时可以进一步加入洪水淹没范围预测模型, 能够将径流值数据转化为更直接的洪水淹没范围图, 能对洪水灾害的预防提供更好的建议。

## 基金项目

四川国家应用数学中心开放基金重点项目 2022-KFJJ-02-001。

## 参考文献

- [1] Abaya, S.W. (2008) Floods and Health in Gambella Region, Ethiopia: An Assessment of the Strength and Weakness of the Coping Mechanism. Lund University, Lund, 1-44.
- [2] Cirella, G. and Iyalomhe, F. (2018) Flooding Conceptual Review: Sustainability-Focalized Best Practices in Nigeria. *Applied Sciences*, **8**, Article No. 1558. <https://doi.org/10.3390/app8091558>
- [3] Tamiru, H. and Dinka, M.O. (2021) Application of ANN and HEC-RAS Model for Flood Inundation Mapping in Lower Baro Akobo River Basin, Ethiopia. *Journal of Hydrology: Regional Studies*, **36**, 100855. <https://doi.org/10.1016/j.ejrh.2021.100855>
- [4] Graves, A. and Schmidhuber, J. (2005) Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures. *Neural Networks*, **18**, 602-610. <https://doi.org/10.1016/j.neunet.2005.06.042>
- [5] Mengistu, D.T., Moges, S.A., Sorteberg, A., et al. (2016) Revisiting Systems Type Black-Box Rainfall-Runoff Models for Flow Forecasting Application. *Journal of Water Resource and Protection*, **8**, 65-83. <https://doi.org/10.4236/jwarp.2016.81006>
- [6] Toth, E., Brath, A. and Montanari, A. (2000) Comparison of Short-Term Rainfall Prediction Models for Real-Time Flood Forecasting. *Journal of Hydrology*, **239**, 132-147. [https://doi.org/10.1016/S0022-1694\(00\)00344-9](https://doi.org/10.1016/S0022-1694(00)00344-9)
- [7] 门宝辉, 常溶溶. 基于局部均值分解和最大熵谱估计的径流分析[J]. 水力发电学报, 2021, 40(9): 27-34.
- [8] Tegegne, G., Park, D.K. and Kim, Y. (2017) Comparison of Hydrological Models for the Assessment of Water Resources in a Data-Scarce Region, the Upper Blue Nile River Basin. *Journal of Hydrology: Regional Studies*, **14**, 49-66. <https://doi.org/10.1016/j.ejrh.2017.10.002>
- [9] Tamiru, H. and Wagari, M. (2022) Machine-Learning and HEC-RAS Integrated Models for Flood Inundation Mapping in Baro River Basin, Ethiopia. *Modeling Earth Systems and Environment*, **8**, 2291-2303. <https://link.springer.com/article/10.1007/s40808-021-01175-8> <https://doi.org/10.1007/s40808-021-01175-8>
- [10] Chen, C., Hui, Q., Xie, W., Wan, S., Zhou, Y. and Pei, Q. (2021) Convolutional Neural Networks for Forecasting Flood Process in Internet-of-Things Enabled Smart City. *Computer Networks*, **186**, Article ID: 107744 <https://doi.org/10.1016/j.comnet.2020.107744>
- [11] Grimes, D.I.F., Coppola, E., Verdecchia, M. and Visconti, G. (2003) A Neural Network Approach to Real-Time Rainfall Estimation for Africa Using Satellite Data. *Journal of Hydrometeorology*, **4**, 1119-1133. [https://doi.org/10.1175/1525-7541\(2003\)004<1119:ANNATR>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1119:ANNATR>2.0.CO;2)
- [12] Dolling, O.R. and Varas, E.A. (2002) Artificial Neural Networks for Streamflow Prediction. *Journal of Hydraulic Re-*

- 
- search*, **40**, 547-554. <https://doi.org/10.1080/00221680209499899>
- [13] 石磊, 王毅, 成颖, 魏瑞斌. 自然语言处理中的注意力机制研究综述[J]. 数据分析与知识发现, 2020, 4(5): 1-14.
- [14] Wang, Y., Huang, M., Zhu, X. and Zhao, L. (2016) Attention-Based LSTM for Aspect-Level Sentiment Classification. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, 1-5 November 2016, 606-615. <https://doi.org/10.18653/v1/D16-1058>
- [15] Bahremand, A. and De Smedt, F. (2008) Distributed Hydrological Modeling and Sensitivity Analysis in Torysa Watershed, Slovakia. *Water Resources Management*, **22**, 393-408. <https://doi.org/10.1007/s11269-007-9168-x>
- [16] Lv, Z., Zuo, J. and Rodriguez, D. (2020) Predicting of Runoff Using an Optimized SWAT-ANN: A Case Study. *Journal of Hydrology: Regional Studies*, **29**, Article ID: 100688. <https://doi.org/10.1016/j.ejrh.2020.100688>