DOI:10.3880/j.issn.1006-7647.2022.03.010

Seq2Seq 模型的短期水位预测

刘 艳[1],张 婷[1],康爱卿[2],李建柱[1],雷晓辉[2]

(1.天津大学水利工程仿真与安全国家重点实验室,天津 300072;
2.中国水利水电科学研究院流域水循环模拟与调控国家重点实验室,北京 100038)

摘要:为有效预测未来一定时间内的连续水位,提出了基于序列到序列(Seq2Seq)的短期水位预测模型,并使用一个长短期记忆神经网络(LSTM)作为编码层,将历史水位序列编码为一个上下文向量,使用另一个 LSTM 作为解码层,将上下文向量解码来预测目标水位序列。以流溪河为研究对象,针对不同预测长度分别建立水位预测模型,并与 LSTM 模型和人工神经网络(ANN)模型进行了对比。结果表明:Seq2Seq 模型对连续 6 h、12 h 和 24 h 水位预测的纳什效率系数最高分别为 0.93、0.90 和 0.85;当预测长度为 6 h 时,LSTM 和 Seq2Seq 模型预测结果相似,ANN 模型精度较低;当预测长度为 12 h 和 24 h 时,Seq2Seq 模型相比 LSTM 模型和 ANN 模型预测效果更好,收敛速度更快。

关键词:水位预测;序列到序列模型;长短期记忆神经网络;人工神经网络;深度学习

中图分类号:TV124 **文献标志码**:A **文章编号**:1006-7647(2022)03-0057-07

Short-term water level prediction based on Seq2Seq model//LIU Yan[1], ZHANG Ting[1], KANG Aiqing[2], LI Jianzhu[1], LEI Xiaohui[2](1. *State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin* 300072, *China*; 2. *State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing* 100038, *China*)

Abstract: In order to effectively predict the continuous water level in a certain period in advance, a short-term water level prediction model based on sequence to sequence (Seq2Seq) was proposed. The model used a long short-term memory (LSTM) network to encoder the historical water level sequence into a context vector, and another LSTM network was used to decode the context vector to predict the target water level sequence. Taking the Liuxi River as the research object, the water level prediction model was established for different prediction lengths, and compared with LSTM model and artificial neural network (ANN) model. The results show that the highest Nash-Sutcliffe efficiency coefficient predicted by the Seq2Seq model for continuous 6 h, 12 h and 24 h is 0.93, 0.90 and 0.85, respectively. When the prediction length is 6h, the performance of the LSTM and Seq2Seq model are similar, and the ANN model has lowest precision. When the prediction length is 12h and 24h, the Seq2Seq model has better performance and faster convergence speed compared with the LSTM model and ANN model.

Key words: water level prediction; sequence to sequence model; long short-term memory neural network; artificial neural network; deep learning

水位预测对水资源的合理开发、利用及优化配置具有重要意义。由于河流水位受多方面因素的影响,变化频繁,因此短期的精细化预测相对中长期的趋势预测更有实际价值。目前,水位预测的研究方法可以归纳为两类,即过程驱动模型和数据驱动模型。过程驱动模型通过有物理依据及经验性的公式,描述水流运动与影响因素的关系,由于影响水位的因素较多,考虑降雨、地形、水工建筑物等的物理模型会十分复杂,模型精度很难达到水位控制要求[1]。数据驱动模型通过挖掘输入与输出数据之间的相关信息进行预测,而不研究物理过程。常用的模型包括以数理统计为基础的时间序列分析,以及以人工神经网络(artificial neural network,ANN)、支持向量机(support vector machine,SVM)为代表的机器学习。已有的研究采用一种或集合多种数据驱动方法建立水位预测模型,并使用遗传算法等优化方法获得最优参数[2-5]。数据驱动模型在水位预测方面取得了丰硕的成果,但仍存在很大的改进空间,

基金项目:国家重点研发计划(2018YFC0407902)

作者简介:刘艳(1994—),女,博士研究生,主要从事水文预报研究。E-mail:liuy_an@163.com

通信作者:张婷(1988—),女,副教授,博士,主要从事水文水资源研究。E-mail:zhangting_hydro@tju.edu.cn

如传统 SVM 模型只能单步输出,以及浅层神经网络不足以充分挖掘数据的所有特征等。

近年来,人工智能和大数据驱动技术给水文研究带来了新的研究思路和技术方法,以深度学习为代表的新一代人工神经网络已经在水文领域进行探索性应用[6]。深度学习是一种特殊的机器学习方法,从简单的线性网络发展到经典的生成式对抗网络(generative adversarial networks,GAN),经历了深度信念网络(deep belief network,DBN)、稀疏编码(sparse coding)、卷积神经网络(convolutional neural network,CNN)、结构递归神经网络(recursive neural network,RNN)的快速迭代[7]。长短期记忆网络(long short-term memory,LSTM)是 RNN 的改进版,通过引入记忆门控单元缓解梯度消失,是进行时间序列数据仿真模拟及预测的首选。在水位预测方面,LSTM 模型受到越来越多的关注,相关研究成果已表明 LSTM 模型预测精度较高,预测效果优于 SVM、RNN 等模型[8-14]。然而,现有的大部分研究是预测未来某个时间的水位,而不是水位序列。有少数研究对多步水位进行预测,但没有考虑数据输出的先后顺序。

序列到序列(sequence-to-sequence,Seq2Seq)模型是一种由序列生成序列,且序列长度互不影响的模型,被广泛应用于机器翻译、行为预测等任务中[15-17]。水位序列与语言上下文以及其他时间序列具有相似的特征,但 Seq2Seq 模型在水位预测方面的研究甚少。因此,本文提出了一种基于 Seq2Seq 模型的短期多步水位预测模型,以克服上述数据驱动模型存在的不足。以流溪河的水位数据为例,对所提模型进行验证,并与 LSTM 和 ANN 模型进行比较,分析评价预测精度和计算效率。

1 研究区及数据处理

1.1 研究区概况

流溪河位于广东省广州市北部,地理位置介于 113°10′12″E ~ 114°2′00″E,23°12′30″N ~ 23°57′36″N 之间。流溪河是广州市辖区内唯一的一条全流域河流,也是集雨面积最大的河流[18]。河流全长约为 171 km,流域面积约为 2 300 km^2。该流域处于亚热带季风气候区,气候温湿,雨量丰沛。年平均降雨量为 1 800 mm,年径流量约为 28.4 亿 m^3。84% 的降雨集中在汛期(4—9 月),有较强的季节性。

流溪河梯级闸坝工程开发力度较大,从上至下共有 11 座拦河闸坝,依次为良口坝、青年坝、胜利坝、卫东坝、人工湖坝、水厂坝、街口坝、大坳坝、牛心岭坝、李溪坝和人和坝[19]。流溪河是广州市最早的供水水源地,目前仍然是重要的工农业和生活用水水源地,广州市 30% 的自来水来自流溪河,被誉为广州的“母亲河”。拦河闸坝的修筑对于广州市水资源开发利用发挥着重要作用。本文选取良口坝、大坳坝和李溪坝为研究对象,如图 1 所示。

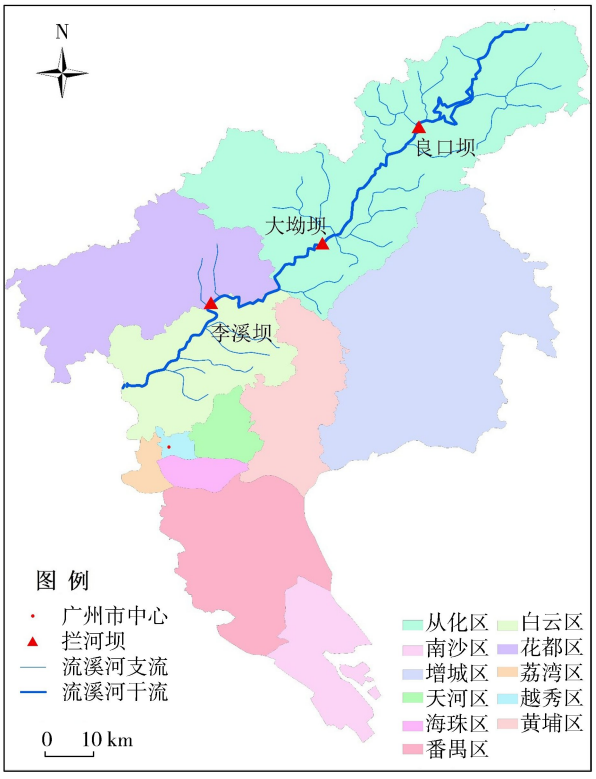

图 1 流溪河及拦河坝位置示意图

1.2 数据处理

本文采用的数据长度为 2014 年 10 月至 2020 年 5 月,监测频率为 1h,其中,2014 年 10 月至 2018 年 2 月为训练期,2018 年 3 月至 2020 年 5 月为验证期。

数据质量是影响数据驱动模型效果的关键因素。在建模之前要对原始数据进行预处理,包括数据整理和数据变换。针对水位数据,数据整理主要是噪声数据识别和空缺值填充。噪声数据是指在测量和收集数据时由机器采集或者人为造成的误差。噪声数据远远超过或不及其他数据,明显偏离一般水平,一般将其剔除然后按照空缺值的方法填充。空缺值是指较少且不连续的数据缺失,一般采用平均值、众数等进行填充。考虑到研究对象水位时间间隔较短,对于空缺值采用线性插值的方式处理。数据变换是对数据进行规范化,本文采用标准化算法对模型输入数据进行处理[20],处理后的序列均值为 0,标准差为 1。计算公式为

$$x^* = \frac{x - \bar{x}}{\sigma} \tag{1}$$

式中:x、x^*、$\bar{x}$ 为给定序列的初始值、标准值和均值;σ 为给定序列的标准差。当给定的数据大于其所在

序列均值时,会得到一个正的标准化数值;反之,会得到一个负的标准化数值。本文在划分训练集和测试集的基础上,采用训练集的均值和标准差分别对两组数据集进行标准化,并采用同样的均值和标准差对模型结果进行反标准化。

2 模型构建与评价

2.1 模型构建

本文利用历史水位作为输入,预测未来一定时间内的连续水位,因此选用处理时间序列的神经网络。LSTM 是 RNN 的一种特殊类型,可以从隐藏的长期依赖信息中学习[21]。LSTM 通过引入门控单元统筹并传递对当前时刻重要的历史信息,提升了循环神经网络的长时间记忆,并保持了训练过程中梯度下降的稳定性。但 LSTM 的每一个输入都对应一个输出,因此只能处理输入输出序列时序一致的数据。

Seq2Seq 模型最初是因机器翻译而诞生,随后被广泛用于处理输出序列长度与输入序列长度无关的相关任务。Seq2Seq 模型由编码器和解码器组成,编码器和解码器本质上是循环神经网络。编码器将可变长度序列数据编码为一个固定长度的中间向量,解码器将固定长度的中间向量进行解码作为预测输出,实现了任意长度的输入序列可以映射到任意长度的输出序列的目的。

在实际应用中,编码器和解码器有多种选择,本文选用两个 LSTM 模块组成 Seq2Seq 模型对水位进行预测。当输入历史时间序列的长度为 n,预测输出时间序列的长度为 m 时,构造的输入向量 $\boldsymbol{X}$ 及输出向量 $\boldsymbol{Y}$ 如下:

$$\boldsymbol{X}=(X_1,X_2,X_3,\cdots,X_N) \tag{2}$$

$$\boldsymbol{Y}=(Y_1,Y_2,Y_3,\cdots,Y_N) \tag{3}$$

$$\boldsymbol{X}_i=(x_{1+m(i-1)},x_{2+m(i-1)},x_{3+m(i-1)},\cdots,x_{n+m(i-1)}) \tag{4}$$

$$\boldsymbol{Y}_i=(x_{n+1+m(i-1)},x_{n+2+m(i-1)},x_{n+3+m(i-1)},\cdots,x_{n+mi}) \tag{5}$$

式中:N 为样本总数;$\boldsymbol{X}_i$、$\boldsymbol{Y}_i$ 分别为第 i 个样本的输入和输出($i=1,2,\cdots,N$)。此种输入输出向量构造方法可同时得到一个预测长度的各步长预测结果,避免了使用前一步预测结果作为真实值输入所造成的误差累积。

将样本($\boldsymbol{X}_1$,$\boldsymbol{Y}_1$)代入 Seq2Seq 模型中,其中$\boldsymbol{X}_1=(x_1,x_2,\cdots,x_n)$表示输入的历史水位,$\boldsymbol{Y}_1=(y_1,y_2,\cdots,y_m)$表示输出的预测水位。编码器由单层 LSTM 单元组成,将输入序列按照时间顺序分步读入,每一时刻的隐藏层状态 h_t 都由当前时刻的输入数据 x_t 与上一时刻的隐藏层状态 h_{t-1} 和细胞状态 c_{t-1} 共同决定,即

$$h_t=f(x_t,h_{t-1},c_{t-1}) \tag{6}$$

式中:h_t、h_{t-1} 分别为 t 时刻、$t-1$ 时刻编码层的 LSTM 神经元隐藏层状态;x_t 为 t 时刻的输入;c_{t-1} 为 $t-1$ 时刻编码层的 LSTM 神经元细胞状态。

LSTM 单元内部由输入门、遗忘门和输出门组成。首先,遗忘门决定舍弃多少细胞状态 c_{t-1};接下来,输入门决定保留多少当前外部信息(即输入 x_t),并生成候选细胞状态 c_t;然后,根据遗忘门和输入门的结果更新细胞状态 c_t;最后,输出门决定输出细胞状态 c_t 的哪些状态特征,并生成隐藏层状态 h_t。当$t=1$ 时,h_0 和 c_0 对应编码层起始状态隐藏层的输出,采用 0 进行初始化。上述过程对应公式如下:

$$f_t=\sigma(\boldsymbol{W}_{\mathrm{f}}(h_{t-1},x_t)+\boldsymbol{b}_{\mathrm{f}}) \tag{7}$$

$$i_t=\sigma(\boldsymbol{W}_{\mathrm{i}}(h_{t-1},x_t)+\boldsymbol{b}_{\mathrm{i}}) \tag{8}$$

$$\bar{c}_t=\tanh(\boldsymbol{W}_{\mathrm{c}}(h_{t-1},x_t)+\boldsymbol{b}_{\mathrm{c}}) \tag{9}$$

$$c_t=c_{t-1}f_t+\bar{c}_t i_t \tag{10}$$

$$o_t=\sigma(\boldsymbol{W}_{\mathrm{o}}(h_{t-1},x_t)+\boldsymbol{b}_{\mathrm{o}}) \tag{11}$$

$$o_t=o_t\tanh c_t \tag{12}$$

式中:$\boldsymbol{W}_{\mathrm{f}}$、$\boldsymbol{W}_{\mathrm{i}}$、$\boldsymbol{W}_{\mathrm{c}}$、$\boldsymbol{W}_{\mathrm{o}}$分别为遗忘门、输入门、输出门和门控单元的权重向量;$\boldsymbol{b}_{\mathrm{f}}$、$\boldsymbol{b}_{\mathrm{i}}$、$\boldsymbol{b}_{\mathrm{c}}$、$\boldsymbol{b}_{\mathrm{o}}$分别为遗忘门、输入门、输出门和门控单元的偏置向量;f_t、i_t、$\bar{c}_t$、o_t 均为模型中间参数;$\sigma(\cdot)$ 为 Sigmoid 激活函数;$\tanh(\cdot)$为双曲正切激活函数。

编码器经过 n 个时间步的更新,输入序列被编码为最终时间步的隐藏层状态 h_n。由于循环神经网络具有长期记忆功能,理论上 h_n 包含输入序列的完整信息。解码器由单层 LSTM 单元组成,接受编码器的最终状态 h_n 作为初始输入值。解码器每一时刻隐藏层状态 h'_t 通过当前时刻的输入 h_n 与上一时刻的隐藏层状态 h'_{t-1} 和细胞状态 c'_{t-1} 更新,其表达式如下:

$$h'_t=f(h_n,h'_{t-1},c'_{t-1}) \tag{13}$$

式中:h'_t、h'_{t-1} 分别为 t 时刻、$t-1$ 时刻解码层的 LSTM 神经元隐藏层状态;c'_{t-1} 为 $t-1$ 时刻解码层的 LSTM 神经元细胞状态。

解码层更新 h'_{t-1} 的具体操作与编码层更新 h_t 一致。当 $t=1$ 时,h'_0 和 c'_0 为解码层起始状态隐藏层的输出,采用 0 进行初始化。隐藏状态 h'_t 经全连接层的学习输出为 y_t,其表达式为

$$y_t=F(h'_t) \tag{14}$$

式中 y_t 为 t 时刻的输出。经过分步解码,形成最终的输出序列。

2.2 模型参数设置

对良口坝、大坳坝和李溪坝的逐小时水位进行连续 6 h、12 h 和 24 h 的预测,即模型输出长度 m 分别是 6 h、12 h、24 h。深度学习模型参数有两类:一类是可训练参数,通过训练集自动率定,主要是研究方法中提到的权重和偏置;另一类是超参数,根据训练过程中评价指标的结果进行经验调参。Seq2Seq 模型的计算步骤及参数设置如下:

步骤 1　对数据进行预处理。

步骤 2　设置模型预测目标及输入数据。模型预测长度 m 分别为 6 h、12 h、24 h，模型输入长度 n 分别设为 72 h、168 h、168 h。

步骤 3　建立编码层与解码层的网络结构。编码层和解码层均为单层 LSTM，编码层隐藏层单元数量分别为 32、128、128，解码层隐藏层单元数量分别为 8、64、64。

步骤 4　利用训练集的样本训练模型。训练过程以均方误差为损失函数，使用 ADAM 优化模型参数，学习率为 0.001。由于数据集较大，将批量尺寸参数 batch-size 设置为 256，最大训练次数设为 200。

步骤 5　对模型输出进行反归一化处理，得到水位预测结果。

步骤 6　结果分析。

2.3　评价指标

本文采用纳什效率系数(NSE)、均方根误差(RMSE)和平均绝对误差(MAE)定量评价模型的预测性能，NSE 的变化范围从 $-\infty$ 到 1，越接近 1，拟合越完美，模型可信度越高。同时采用 RMSE 和 MAE 衡量模拟值和实测值之间的偏差，变化范围从 0 到 $+\infty$，越接近 0，模型整体模拟效果越好。当 MAE 相等时，RMSE 越大，数据离散度越高。

3　结果分析

3.1　整体预测效果评价

选择 NSE、RMSE 和 MAE 对 Seq2Seq 模型在良口坝、大坳坝和李溪坝的短期水位预测结果进行评价，结果见表 1。由表 1 可知：在所有预测时间，MAE 均小于 0.1 m；NSE 均大于 0.7，最大值为 0.98；RMSE 最大值为 0.17 m；Seq2Seq 模型取得了较好的预测效果。对比同一站点不同预测长度的评价指标结果可以发现，当预测长度为 6 h 时，NSE 均大于 0.9，模型的预测效果最佳。随着预测长度的增加，RMSE 和 MAE 不断增大，而 NSE 不断减小。以良口坝为例，预测长度从 6 h 增加到 24 h 时，测试集的 NSE 降低了 0.08，RMSE 和 MAE 分别增加了 0.06 m、0.02 m。

表 1　各站点预测结果对比

闸坝工程	m/h	训练集			测试集		
		NSE	RMSE/m	MAE/m	NSE	RMSE/m	MAE/m
良口坝	6	0.98	0.12	0.08	0.93	0.11	0.06
	12	0.98	0.17	0.09	0.90	0.13	0.07
	24	0.97	0.16	0.09	0.85	0.17	0.08
大坳坝	6	0.95	0.06	0.04	0.93	0.07	0.04
	12	0.92	0.08	0.05	0.83	0.10	0.05
	24	0.87	0.10	0.06	0.78	0.12	0.06
李溪坝	6	0.92	0.05	0.03	0.92	0.05	0.03
	12	0.89	0.06	0.04	0.82	0.07	0.05
	24	0.82	0.07	0.05	0.73	0.08	0.06

图 2 给出了 3 种不同预测长度下各站点测试集的预测与实测水位对比，从中可以看出，Seq2Seq 模型能够较好地反映坝上水位随时间的波动趋势。各个站点连续 6 h 的预测水位与实测水位最接近，连续 24 h 的预测水位不能很好地再现实测水位涨落

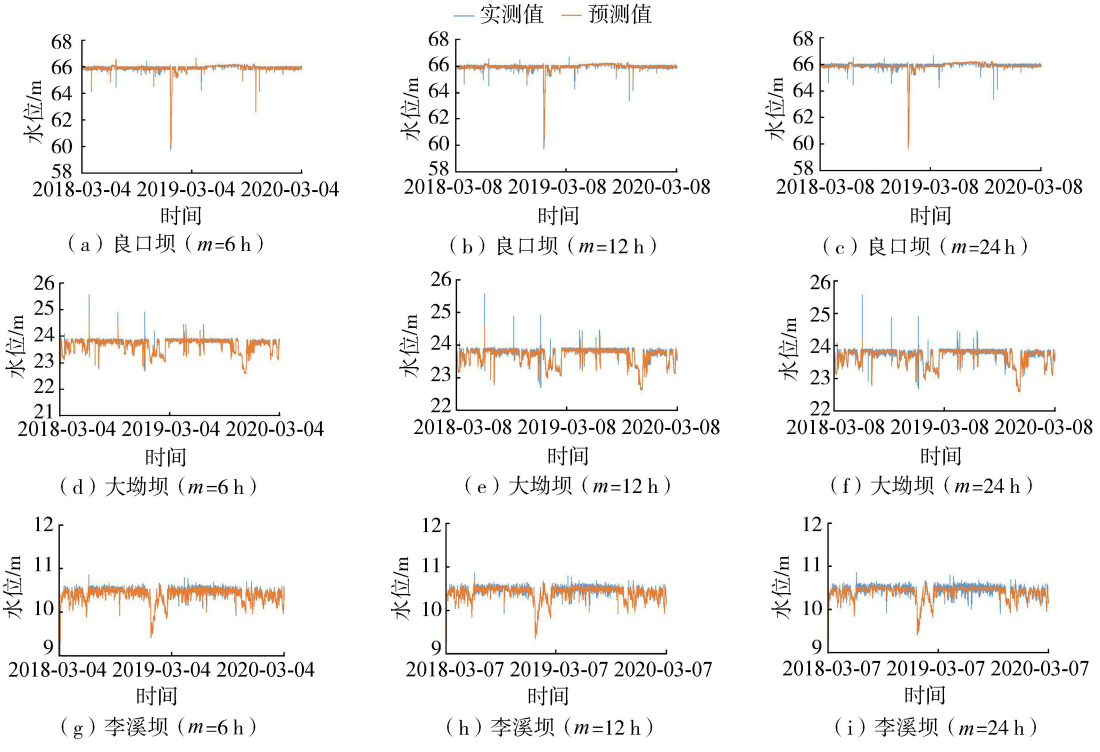

图 2　基于 Seq2Seq 的水位预测对比

过程。Seq2Seq 模型对于高水位或低水位(极值)的预测效果相对较差。为了定量分析预测误差大小,计算良口坝水位小于 64 m,大坳坝水位大于 24 m 以及李溪坝水位小于 10 m 范围内水位预测的最大误差、最小误差和 MAE,结果如表 2 所示。由表 2 可知,总体上 MAE 随预测长度的增大而增大。与表 1 相比,极值预测的 MAE 远高于全过程的 MAE,说明模型整体预测性能较好,但对极值的预测略有不足。这主要是由于影响水位的因素较多,比如闸坝调度、灌溉取水等,只考虑历史水位数据不能很好地呈现水位突变的趋势。

表 2　各站点极值的预测误差对比

闸坝工程	*m*/h	最大误差/m	最小误差/m	MAE/m
良口坝	6	2.604 2	−0.004 4	0.331 5
	12	2.560 8	0.001 8	0.594 0
	24	2.586 9	0.003 4	0.709 1
大坳坝	6	−1.399 5	0.004 2	0.427 8
	12	−1.971 9	−0.009 5	0.633 8
	24	−1.875 2	−0.047 2	0.709 5
李溪坝	6	0.586 5	0.000 1	0.063 7
	12	0.631 6	−0.000 1	0.063 2
	24	0.757 2	0.000 7	0.104 4

此外,对于测试集,在预测长度分别为 6 h、12 h、24 h 的情况下,大坳坝的 NSE 比良口坝的 NSE 依次小 0、0.07 和 0.07,李溪坝的 NSE 比良口坝的 NSE 依次小 0.01、0.08 和 0.12。结合 NSE 和水位过程对比可以看出,良口坝的水位预测效果最好,大坳坝次之,李溪坝较差。这是由于从河流上游至下游,随着预测站点到河源的距离增大,影响水位的因素增加;同时,当预测长度增大时,上下游站点之间差异有增大的趋势。

3.2　与 LSTM 和 ANN 模型的对比

为了便于模型之间的比较,除 Seq2Seq 模型外,本文还构建了 LSTM 模型和 ANN 模型。ANN 模型采用的是最基本的 3 层结构,即输入层、隐藏层、输出层,又称为全连接神经网络(full connected neural networks, FCNN)[22]。LSTM 模型和 ANN 模型隐藏层的神经元数量与 Seq2Seq 模型的编码层神经元数量一致,其他参数也与 Seq2Seq 模型相同。基于测试集,对比分析 Seq2Seq、LSTM 和 ANN 模型的预测结果。

图 3 ~ 5 为 3 种模型在不同预测长度下的评价指标。从图 3 ~ 5 可以看出,ANN 模型的 NSE 均小于 0.9,最小值为 0.67,RMSE 和 MAE 的最大值分别超过了 0.20 m 和 0.10 m。LSTM 和 Seq2Seq 模型的 NSE 最小值分别为 0.70 和 0.73,MAE 最大值分别为 0.09 m 和 0.08 m,可见这两个深度学习模型各项指标均优于简单神经网络模型。当预测长度为 6 h 时,Seq2Seq 和 LSTM 模型的各项指标基本相等;

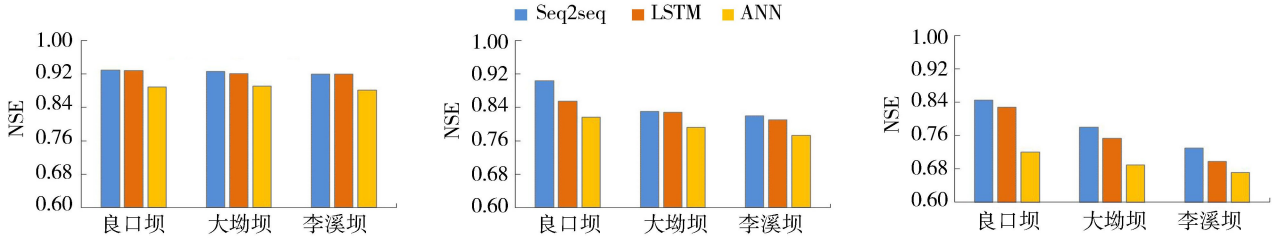

图 3　不同模型预测结果的 NSE 对比

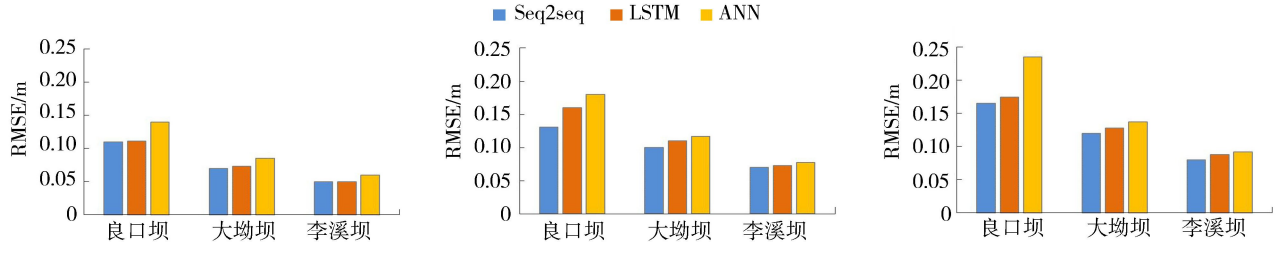

图 4　不同模型预测结果的 RMSE 对比

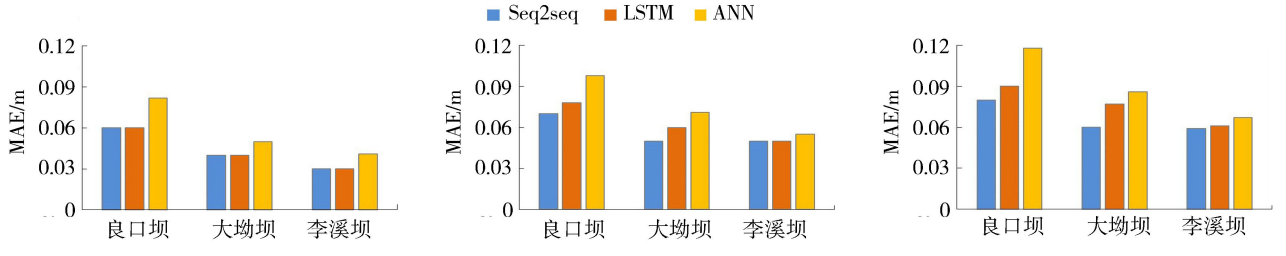

图 5　不同模型预测结果的 MAE 对比

当预测长度为 12 h 时，在良口坝的水位预测中，Seq2Seq 模型相对 LSTM 模型来说，NSE 提高了大约 0.05，RMSE 和 MAE 分别降低了大约 0.03 m 和 0.01 m，其他站点的各项指标值较为接近；当预测长度为 24 h 时，Seq2Seq 模型在 3 个站点的各项指标结果均有不同程度的改进。

LSTM 模型和 ANN 模型预测结果同 Seq2Seq 模型一样，都表现出随着预测长度的增加，预测性能逐渐下降的特性。对于评价指标值的变化幅度，当预测长度增加到 12 h 时，Seq2Seq 模型预测结果最小，LSTM 模型预测结果和 ANN 模型接近；当预测长度增加到 24 h 时，ANN 模型预测结果最大，Seq2Seq 模型在大坳坝和李溪坝预测结果最小，而 LSTM 模型在良口坝预测结果最小。此外，LSTM 和 ANN 模型受站点位置的影响规律与 Seq2Seq 模型一致，即中上游站点的预测效果优于下游站点，且下游站点与中上游站点之间的评价指标值差距随着预测长度的增大而增大。

由于 ANN 模型的预测效果较差，仅对 Seq2Seq 与 LSTM 模型的计算效率进行比较。以良口坝为例，在相同环境下两模型完成一次计算的时间如表 3 所示。由表 3 可以看出，预测长度为 12 h 的计算时间大于预测长度为 24 h 的计算时间，表明神经元数量相同时，样本数越多运行时间越长；预测长度为 6 h 的计算时间小于预测长度为 12 h 的计算时间，表明样本数虽多，但神经元数量减少，模型运行时间依然较短。由以上分析可知，样本数和神经元数量都影响计算效率，且神经元数量的影响较大。对比同一预测长度下两模型的计算时间可以发现 LSTM 模型计算时间更短。图 6 给出了不同预测长度下两模型的训练过程。由图 6 可知，Seq2Seq 模型的 RMSE 均小于 LSTM。当预测长度为 6 h 和 12 h 时，LSTM 先收敛；当预测长度为 24 h 时，Seq2Seq 模型的收敛速度较快。大坳坝和李溪坝计算效率的对比结果与良口坝相似。以上对预测精度和计算效率的对比表明，在实际应用中，需根据对精度和效率的要求选择合适的预测模型。

表 3　不同模型的计算时间对比

m/h	样本数		神经元数量		计算时间/s	
	训练集	测试集	编码层	解码层	Seq2Seq	LSTM
6	4976	3280	32	8	146.71	140.46
12	2488	1632	128	64	964.74	946.48
24	1240	816	128	64	554.22	460.46

4　结　论

a. 基于 Seq2Seq 模型建立了未来 6 h、12 h 和 24 h的逐小时水位预测模型，使输入与输出的长度设置不受限制，模型训练过程平稳，预测效果较好。

b. 当预测长度为 6 h、12 h 和 24 h 时，Seq2Seq 模型的 MAE 均小于 0.1 m，NSE 均大于 0.7，表明该模型具有较高的预测精度。同时表现出预测长度越小，评价指标越优的趋势；相同预测长度下，上游水位比中下游水位的预测效果更好。

c. 本文测试集 ANN 模型、LSTM 模型和 Seq2Seq 模型的 NSE 最小值分别为 0.67、0.70 和 0.73，MAE 最大值分别为 0.12 m、0.09 m 和 0.08 m，可见 Seq2Seq 模型整体效果优于 LSTM 模型和 ANN 模型。LSTM 模型的计算时间较 Seq2Seq 模型短，但当预测长度为 24 h 时，Seq2Seq 模型收敛速度更快。

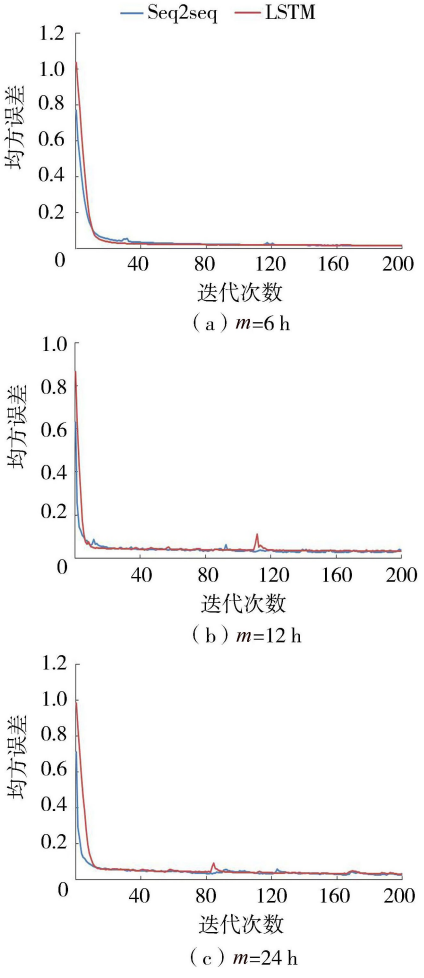

图 6　不同模型的训练过程对比

参考文献：

[1] PAN M, ZHOU H, CAO J, et al. Water level prediction model based on GRU and CNN[J]. IEEE Access, 2020, 8:60090-60100.

[2] 易云飞，盛康. 基于时间序列分析的水位短期预测模型仿真[J]. 计算机工程与设计，2016，37(5)：1331-1334，1339. (YI Yunfei, SHENG Kang. Time series analysis based short-term forecasting model simulation of water

level [J]. Computer Engineering and Design, 2016, 37(5):1331-1334,1339. (in Chinese))
[3] 王景雷,吴景社,孙景生,等. 支持向量机在地下水位预报中的应用研究[J]. 水利学报,2003,34(5):122-128. (WANG Jinglei, WU Jingshe, SUN Jingsheng, et al. The application of support vector machine in groundwater level prediction [J]. Journal of Hydraulic Engineering, 2003, 34(5):122-128. (in Chinese))
[4] WEI C C. Wavelet kernel support vector machines forecasting techniques: case study on water-level predictions during typhoons [J]. Expert Systems with Applications, 2012, 39(5):5189-5199.
[5] LEE G, LEE S, PARK H. Improving applicability of neuro-genetic algorithm to predict short-term water level: a case study [J]. Journal of Hydroinformatics, 2014, 16(1):218.
[6] 刘昌军. 基于人工智能和大数据驱动的新一代水文模型及其在洪水预报预警中的应用[J]. 中国防汛抗旱,2019,29(5):11,22. (LIU Changjun. A new generation of hydrological model based on artificial intelligence and big data and its application in flood forecasting and early warning[J]. China Flood & Drought Management, 2019, 29(5):11,22. (in Chinese))
[7] LIU Y, ZHANG T, KANG A, et al. Research on runoff simulations using deep-learning methods [J]. Sustainability, 2021, 13(3):1336.
[8] LIANG C, LI H, LEI M, et al. Dongting lake water level forecast and its relationship with the three gorges dam based on a long short-term memory network [J]. Water, 2018, 10(10):1389-1392.
[9] 郭燕,赖锡军. 基于长短时记忆神经网络的鄱阳湖水位预测[J]. 湖泊科学,2020,32(3):865-876. (GUO Yan, LAI Xijun. Water level prediction of Lake Poyang based on long short-term memory neural network [J]. Journal of Lake Sciences, 2020, 32(3):865-876. (in Chinese))
[10] ZHANGA J, ZHUB Y, ZHANGA X, et al. Developing a long short-term memory (LSTM) based model for predicting water table depth in agricultural areas [J]. Journal of Hydrology, 2018, 561:65-69.
[11] 王枭雄. 内河航道智能水位预测模型研究[D]. 大连:大连海事大学,2019.
[12] 唐鸣,雷晓辉,龙岩,等. 基于长短时记忆网络(LSTM)的南水北调中线水位预测[J]. 中国农村水利水电,2020(10):189-193. (TANG Ming, LEI Xiaohui, LONG Yan, et al. Water level forecasting in middle route of the south-to-north water diversion project (MRP) based on long short-term memory(LSTM) [J]. China Rural Water and Hydropower, 2020(10):189-193. (in Chinese))
[13] 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,2018,275(7):86-89,96. (FENG Jun, PAN Fei. A hydrologic forecast method based on LSTM-BP[J]. Computer and Modernization, 2018, 275(7):86-89,96. (in Chinese))
[14] 王亦斌,孙涛,梁雪春,等. 基于 EMD-LSTM 模型的河流水量水位预测[J]. 水利水电科技进展,2020,40(6):44-51. (WANG Yibin, SUN Tao, LIANG Xunchun, et al. Prediction of river water flow and water level based on EMD-LSTM model [J]. Advances in Science and Technology of Water Resources, 2020, 40(6):44-51. (in Chinese))
[15] DENG Y, WANG L, JIA H, et al. A sequence-to-sequence deep learning architecture based on bidirectional GRU for type recognition and time location of combined power quality disturbance [J]. IEEE Transactions on Industrial Informatics, 2019, 15(8):4481-4493.
[16] 刘旭林,赵文芳,唐伟. 应用 CNN-Seq2Seq 的 PM 2.5 未来一小时浓度预测模型[J]. 小型微型计算机系统,2020,41(5):1000-1006. (LIU Xulin, ZHAO Wenfang, TANG Wei. Forecasting model of PM 2.5 concentration one hour in advance based on CNN-Seq2Seq [J]. Journal of Chinese Computer Systems, 2020, 41(5):1000-1006. (in Chinese))
[17] 游兰,韩雪薇,何正伟,等. 基于改进 Seq2Seq 的短时 AIS 轨迹序列预测模型[J]. 计算机科学,2020,47(9):175-180. (YOU Lan, HAN Xuewei, HE Zhengwei, et al. Improved sequence-to-sequence model for short-term vessel trajectory prediction using AIS data streams [J]. Computer Science, 2020, 47(9):175-180. (in Chinese))
[18] 贾超,邹牡娇,肖春意,等. 广州流溪河流域水质污染现状调查与研究[J]. 广东化工,2018,45(15):138-139. (JIA Chao, ZOU Mujiao, XIAO Chunyi, et al. Investigation and research on the current situation of water pollution for the LiuXi River, Guangzhou [J]. Guangdong Chemical Industry, 2018, 45(15):138-139. (in Chinese))
[19] 张艳艳,熊德迟,李添,等. 广州流溪河梯级闸坝工程鱼道布局研究[J]. 人民珠江,2016(2):83-87. (ZHANG Yanyan, XIONG Dechi, LI Tian, et al. The distribution study of the fish resources in Liuxi River of Guangzhou [J]. Pearl River, 2016(2):83-87. (in Chinese))
[20] 殷兆凯,廖卫红,王若佳,等. 基于长短时记忆神经网络(LSTM)的降雨径流模拟及预报[J]. 南水北调与水利科技, 2019, 17(6): 1-9, 27. (YIN Zhaokai, LIAO Weihong, WANG Ruojia, et al. Rainfall-runoff modelling and forecasting based on long short-term memory(LSTM) [J]. South-to-North Water Transfers and Water Science & Technology, 2019, 17(6):1-9,27. (in Chinese))
[21] SHOAIB M, SHAMSELDIN A Y, MELVILLE B W, et al. A comparison between wavelet based static and dynamic neural network approaches for runoff prediction [J]. Journal of Hydrology, 2016(1):211-225.
[22] 夏瑜潞. 人工神经网络的发展综述[J]. 电脑知识与技术,2019,15(20):227-229. (XIA Yulu. A review of the development of artificial neural networks [J]. Computer Knowledge and Technology, 2019, 15(20):227-229. (in Chinese))

(收稿日期:2021-03-18 编辑:骆超)