

DOI:10.3880/j.issn.1006-7647.2019.02.011

基于 LSTM 神经网络的水电站短期水位预测方法

刘亚新, 樊启祥, 尚毅梓, 樊启萌, 刘志武

(中国长江三峡集团有限公司, 北京 100038)

摘要: 针对常规水位预测方法信息挖掘能力不足和启发式算法机理不明确等缺点, 提出了一种基于长短时记忆(long short-term memory, LSTM)网络的水位预测方法。该方法采用水位和出力等直接监测数据, 避免了出入库流量等间接计算值带来的二次误差, 进而提升水位预测的准确率; 采用梯度下降法与 Broyden-Fletcher-Goldfarb-Shanno (BFGS) 算法相结合训练模型, Wolfe-Powell 线搜索方法选取步长, 提高模型收敛速率。将该方法用于葛洲坝水电站的上下游水位预测, 结果表明, 该方法能够实现下游水位连续 6 h 和上游水位连续 3 h 的准确预测, 具有较高的预测精度和实用性, 为葛洲坝水库的实时调度提供了技术支撑。

关键词: 长短时记忆网络; 短期水位预测; 梯度下降法; BFGS 算法; Wolfe-Powell 线搜索方法

中图分类号: TV131

文献标志码: A

文章编号: 1006-7647(2019)02-0056-05

Short-term water level prediction method for hydropower station based on LSTM neural network//LIU Yaxin, FAN Qixiang, SHANG Yizi, FAN Qimeng, LIU Zhiwu (China Three Gorges Corporation, Beijing 100038, China)

Abstract: In order to overcome the shortcomings such as insufficient information mining capability of conventional water level prediction methods and unclear mechanism of heuristic algorithms, an water level prediction method based on long short-term memory (LSTM) neural network is proposed. Direct monitoring data such as water level and unit output are used in this method and the middle errors caused by the indirect calculation of the outflow and inflow can be avoided, which improves the accuracy of water level prediction. A hybrid method based on the gradient descent algorithm and Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm is used to train the model, and the step length is determined by the Wolfe-Powell line search method to accelerate convergence. The proposed method is used to predict the water level at the upstream and downstream of the Gezhouba Hydropower Station. The results show that this method can continuously predict the downstream water level for 6 hours and the upstream water level for 3 hours with high accuracy, providing technical support for the real-time scheduling of the Gezhouba Reservoir.

Key words: LSTM; short-term water level prediction; gradient descent algorithm; BFGS algorithm; Wolfe-Powell line search method

准确的水位预测是水电站安全运行和增发效益的关键。然而, 受诸多因素的影响, 水位预测常常较实际情况偏差较大, 导致发电计划频繁修改, 不利于电站安全、高效、稳定地运行。特别是具有紧密水力联系的梯级电站, 如三峡水电站和葛洲坝水电站构成的梯级电站, 上游电站的出库流量预报误差会叠加到下游电站, 且因两坝间往复流、出入库不平衡等问题, 使得下游电站水位的准确预测更难实现。

大量学者对水位预测问题进行了研究, 提出了多种预测模型。王蒙蒙等^[1] 基于支持向量回归模型对洞庭湖水位进行了预测, 其输入变量为三峡水库逐日出库流量、清江逐日流量及洞庭湖“四水”逐

日入湖流量。马辉等^[2] 提出了一种将灰色模型与神经网络相结合的灰色-BP 神经网络预测方法, 得到了较高精度的预测结果。李欣等^[3] 综合考虑河流水位时空信息, 建立了基于时空序列的径向基函数(RBF)神经网络预测模型, 预测了金沙江下游向家坝水文站的水位情况。余开华^[4] 提出了小波神经网络河道流量水位预测模型, 并应用于预测盘龙河天保站的水位。Ashaary 等^[5] 采用 BP 神经网络预测了 Timah Tasoh 水库的水位变化, 采用窗处理的方式将前 w 个时刻的水位数据作为输入变量, 建立了 6 个不同的神经网络模型进行比较, 结果表明当 $w=2$ 时, 所建立的神经网络模型最优。Chang 等^[6]

作者简介: 刘亚新(1992—), 女, 助理工程师, 硕士, 主要从事水文水资源数据挖掘研究。E-mail: liu_yaxin1@ctg.com.cn
通信作者: 刘志武(1977—), 男, 教授级高级工程师, 博士, 主要从事水文水资源研究。E-mail: liu_zhiwu@ctg.com.cn

采用自适应神经模糊推理系统建立了水库水位的预测模型,并将其应用于 Shihmen 水库的水位预测,结果表明该模型可以准确预测连续 3 h 的水位。

但在梯级电站的水位预测方面,尤其是下游电站的水位预测,国内外进行的研究较少。由于梯级电站之间的紧密联系,上游电站的运行情况会影响到下游电站的水位,并且因距离的原因使得影响具有时间滞后性,因此在预测时要综合考虑这些影响因素。长短时记忆(long short-term memory, LSTM)网络是循环神经网络中一种带有门控的网络,能够考虑到数据的时序性特点,学习到数据之间的长期依赖关系,在处理时间序列数据的预测方面具有独特的优势,已被广泛应用于自然语言处理、语音识别、图片描述等领域。王鑫等^[7]将 LSTM 用于航空领域的月度故障时间序列的预测,并提出了基于多层网格搜索的 LSTM 预测模型参数优选方法。保罗等^[8]考虑磨机特征样本之间的时间尺度信息,采用编码神经网络的方法提取样本特征,使用 LSTM 建立了负荷参数的预测模型。Wielgosz 等^[9]将深度学习算法应用于大型强子对撞机超导磁体的监测和故障保护,分析了不同的 LSTM 网络结构和超参数对模型预测效果的影响,结果表明记忆单元为 128,历史步长为 16 时效果最佳。杨训政等^[10]将 LSTM 用于发电机组的污染物排放研究,并与最小二乘法、支持向量回归机进行对比,结果表明 LSTM 明显优于其他两种方法。

本文提出一种基于 LSTM 网络的水电站水位预测方法,并将其应用于预测葛洲坝水电站上下游水位,以为葛洲坝水库的实时调度提供技术支撑。

1 基于 LSTM 神经网络的水位预测模型

1.1 循环神经网络与 LSTM

循环神经网络在时间序列预测方面具有独特的优势,其通过特定的网络结构,将过去时刻的影响反映到当前的预测之中,同时由于共享不同时刻的权值矩阵,减少了参数数目,使得训练效率大大提高,并且可以处理任意长度的时间序列数据。尽管理论如此,但实际上,原始的循环神经网络由于参数共享,经过多阶段传播后的梯度倾向于消失或爆炸^[11-12],使得模型丧失了学习到更早信息的能力,这个问题也称为长期依赖问题。1997 年 Hochreiter 等^[13]提出了 LSTM 网络,很好地解决了这一问题。LSTM 通过引入门的设置来决定控制信息的流动,最初 LSTM 只有两个门,即输入门和输出门,2000 年 Gers 等^[14]对其进行了改进,加入了遗忘门使 LSTM 通过自主学习遗忘一些无关紧要的历史信息。之后

Gers 等^[15]提出了加入“窥视孔连接”的 LSTM 变体,将单元状态也作为门的输入。Cho 等^[16]将遗忘门和输入门合并为更新门,提出了 GRU (gated recurrent unit)模型。

LSTM 网络模型包含 3 个门:输入门、遗忘门和输出门。设输入序列为 $\{x_1, x_2, \dots, x_T\}$, 输出序列为 $\{y_1, y_2, \dots, y_T\}$, 其中, T 为序列长度。在 t 时刻, LSTM 的隐含层结构如图 1 所示,其中, h_t 为隐含层的输出; c_t 为隐含层的状态单元,表示当前时刻的单元状态,根据历史信息 and 当前信息来更新; f_t 为遗忘门; i_t 为输入门; \tilde{c}_t 为当前输入的单元状态; o_t 为输出门; σ 和 \tanh 为激励函数。

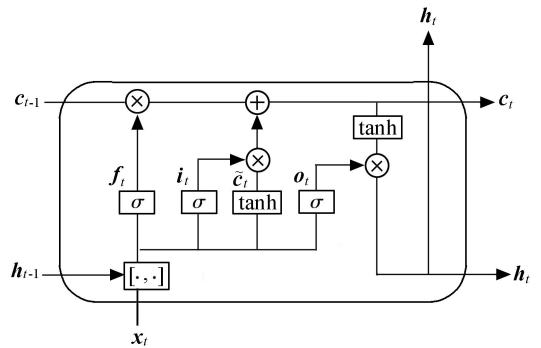


图 1 LSTM 的隐含层结构

1.2 LSTM 的前向传播过程

正如图 1 所示, LSTM 中的每个门实际上就是一个全连接层,输出是 $[0, 1]$ 之间的向量,用以控制信息的传递。门的激励函数 σ 为 sigmoid 函数,可表示为

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

根据应用的不同, LSTM 可以有不同的变体形式。本文构建的 LSTM 结构如图 2 所示,每一个时间步都有输出,隐含层结构如图 1 所示,并且隐含层单元之间有循环连接。

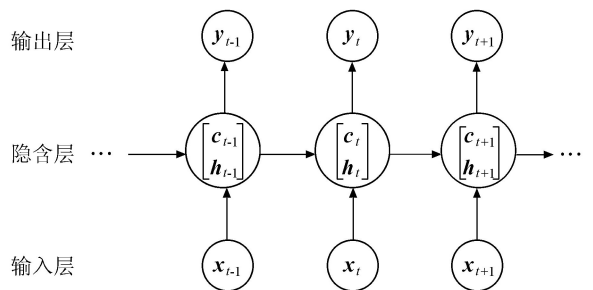


图 2 LSTM 的基本结构

图 1 和图 2 对应的 LSTM 的前向传播的具体公式为

$$\text{遗忘门} \quad f_t = \sigma(n_{f,t}) \quad (2)$$

$$\text{输入门} \quad i_t = \sigma(n_{i,t}) \quad (3)$$

$$\text{输出门} \quad \mathbf{o}_t = \sigma(\mathbf{n}_{o,t}) \quad (4)$$

$$\text{当前输入的单元状态} \quad \tilde{\mathbf{c}}_t = \tanh(\mathbf{n}_{\tilde{c},t}) \quad (5)$$

$$\text{当前的单元状态} \quad \mathbf{c}_t = \mathbf{f}_t \circ \mathbf{c}_{t-1} + \mathbf{i}_t \circ \tilde{\mathbf{c}}_t \quad (6)$$

$$\text{隐含层输出} \quad \mathbf{h}_t = \mathbf{o}_t \circ \tanh(\mathbf{c}_t) \quad (7)$$

$$\text{输出层} \quad \mathbf{y}_t = \tanh(\mathbf{n}_{y,t}) \quad (8)$$

$$\text{其中} \quad \mathbf{n}_{f,t} = \mathbf{W}_{fh} \mathbf{h}_{t-1} + \mathbf{W}_{fx} \mathbf{x}_t + \mathbf{b}_f$$

$$\mathbf{n}_{i,t} = \mathbf{W}_{ih} \mathbf{h}_{t-1} + \mathbf{W}_{ix} \mathbf{x}_t + \mathbf{b}_i$$

$$\mathbf{n}_{o,t} = \mathbf{W}_{oh} \mathbf{h}_{t-1} + \mathbf{W}_{ox} \mathbf{x}_t + \mathbf{b}_o$$

$$\mathbf{n}_{\tilde{c},t} = \mathbf{W}_{ch} \mathbf{h}_{t-1} + \mathbf{W}_{cx} \mathbf{x}_t + \mathbf{b}_c$$

$$\mathbf{n}_{y,t} = \mathbf{W}_y \mathbf{h}_t + \mathbf{b}_y$$

式中: \mathbf{W}_{fh} 和 \mathbf{W}_{fx} 分别为遗忘门中与 \mathbf{h}_{t-1} 和 \mathbf{x}_t 对应的权重矩阵; \mathbf{W}_{ih} 和 \mathbf{W}_{ix} 分别为输入门中与 \mathbf{h}_{t-1} 和 \mathbf{x}_t 对应的权重矩阵; \mathbf{W}_{oh} 和 \mathbf{W}_{ox} 分别为输出门中与 \mathbf{h}_{t-1} 和 \mathbf{x}_t 对应的权重矩阵; \mathbf{W}_{ch} 和 \mathbf{W}_{cx} 分别为当前输入的单元状态中与 \mathbf{h}_{t-1} 和 \mathbf{x}_t 对应的权重矩阵; \mathbf{W}_y 为输出层中与 \mathbf{h}_t 对应的权重矩阵; $\mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_o, \mathbf{b}_c, \mathbf{b}_y$ 分别为遗忘门、输入门、输出门、当前输入的单元状态、输出层中的偏置向量; \circ 表示两个向量对应元素相乘。

1.3 LSTM 的误差反向传播过程

LSTM 的训练算法仍然是误差反向传播算法。由于 LSTM 的前向传播过程是从左到右的有序传播,因此误差反向传播过程也是沿时间反向传递,称为 BPTT(back-propagation through time)算法。

在本文构造的 LSTM 中,要训练的参数包括权重矩阵 $\mathbf{W}_{fh}, \mathbf{W}_{fx}, \mathbf{W}_{ih}, \mathbf{W}_{ix}, \mathbf{W}_{oh}, \mathbf{W}_{ox}, \mathbf{W}_{ch}, \mathbf{W}_{cx}, \mathbf{W}_y$ 和偏置向量 $\mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_o, \mathbf{b}_c, \mathbf{b}_y$ 。将误差沿时间反向传播,需要计算每个时刻的误差项。设 t 时刻真实值 \mathbf{d}_t 与估计值 \mathbf{y}_t 的误差为 $\mathbf{y}_t - \mathbf{d}_t$, 1 到 T 时刻的总误差 E 可表示为

$$E = \sum_{t=1}^T e_t \quad (9)$$

$$\text{其中} \quad e_t = \frac{1}{2} \|\mathbf{y}_t - \mathbf{d}_t\|^2$$

定义 $\nabla_{\mathbf{n}_{y,t}} E$ 表示 E 对 $\mathbf{n}_{y,t}$ 的梯度,公式为

$$\nabla_{\mathbf{n}_{y,t}} E = \left(\frac{\partial \mathbf{y}_t}{\partial \mathbf{n}_{y,t}} \right)^T \nabla_{\mathbf{y}_t} E = (\mathbf{y}_t - \mathbf{d}_t) \circ (1 - \mathbf{y}_t^2) \quad (10)$$

进一步可得

$$\begin{aligned} \nabla_{\mathbf{h}_t} E &= \left(\frac{\partial \mathbf{n}_{y,t}}{\partial \mathbf{h}_t} \right)^T \nabla_{\mathbf{n}_{y,t}} E + \left(\frac{\partial \mathbf{n}_{i,t+1}}{\partial \mathbf{h}_t} \right)^T \nabla_{\mathbf{n}_{i,t+1}} E + \\ &\quad \left(\frac{\partial \mathbf{n}_{f,t+1}}{\partial \mathbf{h}_t} \right)^T \nabla_{\mathbf{n}_{f,t+1}} E + \left(\frac{\partial \mathbf{n}_{o,t+1}}{\partial \mathbf{h}_t} \right)^T \nabla_{\mathbf{n}_{o,t+1}} E + \\ &\quad \left(\frac{\partial \mathbf{n}_{\tilde{c},t+1}}{\partial \mathbf{h}_t} \right)^T \nabla_{\mathbf{n}_{\tilde{c},t+1}} E = \mathbf{W}_y^T \nabla_{\mathbf{n}_{y,t}} E + \mathbf{W}_{ih}^T \nabla_{\mathbf{n}_{i,t+1}} E + \\ &\quad \mathbf{W}_{fh}^T \nabla_{\mathbf{n}_{f,t+1}} E + \mathbf{W}_{oh}^T \nabla_{\mathbf{n}_{o,t+1}} E + \mathbf{W}_{ch}^T \nabla_{\mathbf{n}_{\tilde{c},t+1}} E \quad (11) \\ \nabla_{\mathbf{c}_t} E &= \left(\frac{\partial \mathbf{h}_t}{\partial \mathbf{c}_t} \right)^T \nabla_{\mathbf{h}_t} E + \left(\frac{\partial \mathbf{c}_{t+1}}{\partial \mathbf{c}_t} \right)^T \nabla_{\mathbf{c}_{t+1}} E = \end{aligned}$$

$$\nabla_{\mathbf{h}_t} E \circ \mathbf{o}_t \circ [1 - \tanh(\mathbf{c}_t)]^2 + \nabla_{\mathbf{c}_{t+1}} E \circ \mathbf{f}_{t+1} \quad (12)$$

由 $\nabla_{\mathbf{h}_t} E$ 和 $\nabla_{\mathbf{c}_t} E$ 可得:

$$\begin{cases} \nabla_{\mathbf{n}_{i,t}} E = \nabla_{\mathbf{c}_t} E \circ \tilde{\mathbf{c}}_t \circ \mathbf{i}_t \circ (1 - \mathbf{i}_t) \\ \nabla_{\mathbf{n}_{f,t}} E = \nabla_{\mathbf{c}_t} E \circ \mathbf{c}_{t-1} \circ \mathbf{f}_t \circ (1 - \mathbf{f}_t) \\ \nabla_{\mathbf{n}_{o,t}} E = \nabla_{\mathbf{h}_t} E \circ \tanh(\mathbf{c}_t) \circ \mathbf{o}_t \circ (1 - \mathbf{o}_t) \\ \nabla_{\mathbf{n}_{\tilde{c},t}} E = \nabla_{\mathbf{c}_t} E \circ \mathbf{i}_t \circ (1 - \tilde{\mathbf{c}}_t^2) \end{cases} \quad (13)$$

进而得到权重矩阵和偏置项的梯度分别为

$$\begin{cases} \nabla_{\mathbf{W}_y} E = \sum_{t=1}^T (\nabla_{\mathbf{n}_{y,t}} E) \mathbf{h}_t^T \\ \nabla_{\mathbf{W}_{ih}} E = \sum_{t=1}^T (\nabla_{\mathbf{n}_{i,t}} E) \mathbf{h}_{t-1}^T \\ \nabla_{\mathbf{W}_{fh}} E = \sum_{t=1}^T (\nabla_{\mathbf{n}_{f,t}} E) \mathbf{h}_{t-1}^T \\ \nabla_{\mathbf{W}_{oh}} E = \sum_{t=1}^T (\nabla_{\mathbf{n}_{o,t}} E) \mathbf{h}_{t-1}^T \\ \nabla_{\mathbf{W}_{ch}} E = \sum_{t=1}^T (\nabla_{\mathbf{n}_{\tilde{c},t}} E) \mathbf{h}_{t-1}^T \\ \nabla_{\mathbf{W}_{ix}} E = \sum_{t=1}^T (\nabla_{\mathbf{n}_{i,t}} E) \mathbf{x}_t^T \\ \nabla_{\mathbf{W}_{fx}} E = \sum_{t=1}^T (\nabla_{\mathbf{n}_{f,t}} E) \mathbf{x}_t^T \\ \nabla_{\mathbf{W}_{ox}} E = \sum_{t=1}^T (\nabla_{\mathbf{n}_{o,t}} E) \mathbf{x}_t^T \\ \nabla_{\mathbf{W}_{cx}} E = \sum_{t=1}^T (\nabla_{\mathbf{n}_{\tilde{c},t}} E) \mathbf{x}_t^T \end{cases} \quad (14)$$

$$\begin{cases} \nabla_{\mathbf{b}_y} E = \sum_{t=1}^T \nabla_{\mathbf{n}_{y,t}} E \\ \nabla_{\mathbf{b}_i} E = \sum_{t=1}^T \nabla_{\mathbf{n}_{i,t}} E \\ \nabla_{\mathbf{b}_f} E = \sum_{t=1}^T \nabla_{\mathbf{n}_{f,t}} E \\ \nabla_{\mathbf{b}_o} E = \sum_{t=1}^T \nabla_{\mathbf{n}_{o,t}} E \\ \nabla_{\mathbf{b}_c} E = \sum_{t=1}^T \nabla_{\mathbf{n}_{\tilde{c},t}} E \end{cases} \quad (15)$$

由以上梯度公式就可以采用梯度下降法来更新权重和偏置项。本文将梯度下降法和拟牛顿法中的 BFGS 算法相结合,在初始阶段误差比较大时,采用梯度下降法,当总误差减小到一定程度后,采用 BFGS 法来训练,并通过 Wolfe-Powell 线搜索来搜索合适的步长,提高迭代效率。设模型参数 $\omega = \{ \mathbf{W}_{fh}, \mathbf{W}_{fx}, \mathbf{W}_{ih}, \mathbf{W}_{ix}, \mathbf{W}_{oh}, \mathbf{W}_{ox}, \mathbf{W}_{ch}, \mathbf{W}_{cx}, \mathbf{W}_y, \mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_o, \mathbf{b}_c, \mathbf{b}_y \}$, 给定它的初始值 ω_0 ; 精度阈值为 ε ; 迭代次数

为 j ; 最大迭代次数为 j_{\max} ; 梯度下降法转为 BFGS 算法的阈值为 a ; D 表示逆海森矩阵的近似, 其初始矩阵设为 $D_0 = I$, I 表示单位矩阵。参数更新算法的具体步骤为: ① 计算模型训练的总误差 $E(\omega)$, 以及总误差对参数的梯度 $\nabla E(\omega)$; ② 若 $E(\omega) > a$, 则下降方向 $\delta_\omega = -\nabla E(\omega)$, 根据线搜索得到步长 α , 更新参数 $\omega_{\text{new}} = \omega + \alpha \delta_\omega$, 转至步骤⑤, 否则转至步骤③; ③ 下降方向 $\delta_\omega = D[-\nabla E(\omega)]$, 根据线搜索得到步长 α , 更新参数 $\omega_{\text{new}} = \omega + \alpha \delta_\omega$; ④ 若 $\delta_\omega^T \nabla E(\omega_{\text{new}}) > \delta_\omega^T \nabla E(\omega)$, 则 $s = \omega_{\text{new}} - \omega$, $m = \nabla E(\omega_{\text{new}}) - \nabla E(\omega)$, $h = Dm$, $k_1 = \frac{1}{s^T m}$, $k_2 = k_1 [1 + k_1 (m^T h)]$, 更新 D , 即 $D_{\text{new}} = D + k_2 s s^T - k_1 (s h^T + h s^T)$; ⑤ 若 $\|\nabla E(\omega)\| > \varepsilon$ 且 $j < j_{\max}$, 转至步骤①; 否则, 迭代结束。

2 葛洲坝水电站上下游水位预测

2.1 数据采集及分析

选取 2013 年 8 月 18 日至 2014 年 7 月 2 日三峡水电站和葛洲坝水电站非弃水期间的小时运行数据, 输入变量为三峡上游凤凰山水位、三峡左岸电站总有功、右岸电站总有功、地下电站总有功、电源电站总有功、葛洲坝大江电厂总有功、二江电厂总有功、自备电站总有功, 输出变量为葛洲坝上游 5 号站水位、下游 7 号和 8 号站水位, 训练前先将数据进行标准化处理, 消除量纲的影响。图 3(a) 显示了 7 号站和 8 号站的水位走势, 图 3(b) 显示了 5 号站的水位走势, 图 4 为葛洲坝入库流量与 5 号站水位散点图(7、8 号站散点图类似, 不再画出), 图 5 显示了三峡水库出库流量与葛洲坝入库流量之间关系。从图 4 可以看出, 葛洲坝入库流量是葛洲坝上游 5 号站水位预测的重要影响因素, 而葛洲坝入库流量又与三峡水库出库流量有着紧密的联系。从图 5 可以看出, 三峡水库出库流量和葛洲坝入库流量呈宽带状分布, 在相同的三峡水库出库流量水平下, 葛洲坝入库流量有比较大的取值区间, 仅仅采用线性回归来估计葛洲坝入库流量会产生较大的误差。另一方面, 三峡和葛洲坝梯级电站的流量数据均是计算得到, 本身含有一定的偏差, 因此本文没有选取流量数据, 而是通过水位和有功数据等实时监测数据来间接反映流量的隐含变化。

2.2 水位预测

由于三峡水库的出库流量到达葛洲坝坝前需要一定的时间, 所以三峡水电站的水位和出力变化对葛洲坝水电站水位的影响存在一定的滞后性, 当预测葛洲坝水电站上下游水位时, 应综合之前多个时刻三峡水电站和葛洲坝水电站的出力情况以及三峡

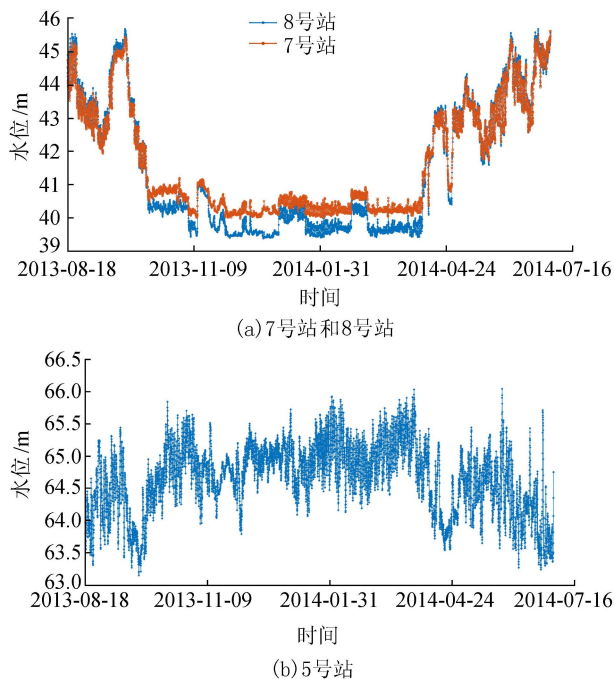


图 3 不同站位水位走势

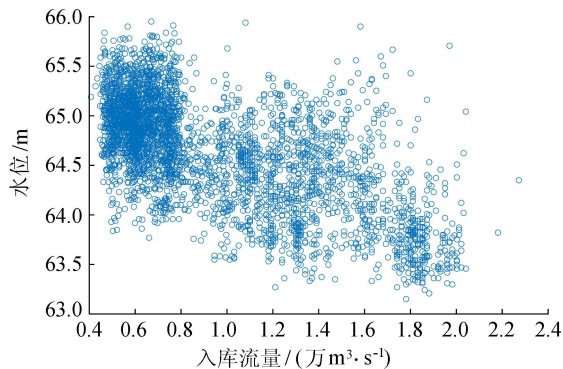


图 4 葛洲坝入库流量与 5 号站水位散点图

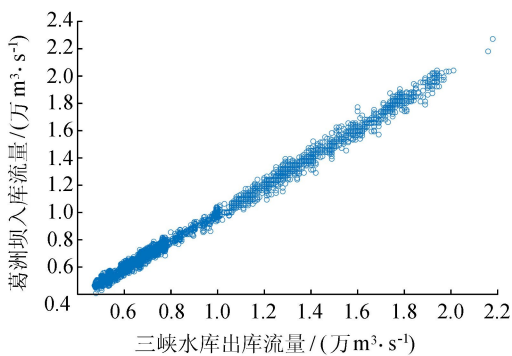


图 5 三峡水库出库流量与葛洲坝入库流量关系

上游凤凰山的水位数据。本文采用 LSTM 神经网络, 将历史每 60 个时刻的数据通过隐含状态引入到下 6 个时刻的预测之中, 每个时刻的输入为 8 维, 隐含层为 20 维, 输出为 3 维, 采用 BPTT 算法进行训练。为了评估模型的预测效果, 采用预测误差的均值、标准差、平均绝对误差 (Y_{MAE}) 和均方根误差 (Y_{RMSE}) 作为评价指标, 计算公式分别为

$$Y_{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i^{\text{test}} - \hat{y}_i^{\text{test}}| \quad (16)$$

$$Y_{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i^{\text{test}} - \hat{y}_i^{\text{test}})^2} \quad (17)$$

式中： y_i^{test} 、 \hat{y}_i^{test} 分别为测试样本中水位的真实值和预测值； N 为测试样本数。

表 1 至表 3 分别为 8 号站、7 号站和 5 号站连续 6 h 的预测误差 ($y_i^{\text{test}} - \hat{y}_i^{\text{test}}$) 的统计结果, 其中 $\sigma_{0.05}$ 和 $\sigma_{0.95}$ 分别为 0.05 分位数和 0.95 分位数。

表 1 8 号站连续 6 h 预测误差的统计结果

预见期/h	均值/m	标准差/m	$[\sigma_{0.05}, \sigma_{0.95}]$ /m	Y_{MAE} /m	Y_{RMSE} /m
1	0.00103	0.03479	[-0.04453, 0.05051]	0.02235	0.03479
2	-0.00195	0.05892	[-0.08036, 0.08103]	0.03647	0.05893
3	-0.00264	0.07873	[-0.11816, 0.11182]	0.04941	0.07874
4	-0.00281	0.09949	[-0.15017, 0.13852]	0.06098	0.09950
5	-0.00344	0.12134	[-0.18053, 0.17554]	0.07341	0.12134
6	-0.00180	0.14216	[-0.21852, 0.20157]	0.08558	0.14211

表 2 7 号站连续 6 h 预测误差的统计结果

预见期/h	均值/m	标准差/m	$[\sigma_{0.05}, \sigma_{0.95}]$ /m	Y_{MAE} /m	Y_{RMSE} /m
1	0.00019	0.02935	[-0.03788, 0.04122]	0.01835	0.02934
2	-0.00181	0.04767	[-0.06963, 0.07494]	0.02924	0.04769
3	-0.00270	0.06438	[-0.09736, 0.09207]	0.03899	0.06441
4	-0.00319	0.08385	[-0.11939, 0.11650]	0.04895	0.08388
5	-0.00395	0.10217	[-0.15044, 0.13644]	0.05936	0.10220
6	-0.00364	0.11633	[-0.18360, 0.16356]	0.06903	0.11634

表 3 5 号站连续 6 h 预测误差的统计结果

预见期/h	均值/m	标准差/m	$[\sigma_{0.05}, \sigma_{0.95}]$ /m	Y_{MAE} /m	Y_{RMSE} /m
1	0.00244	0.04558	[-0.06257, 0.07198]	0.03019	0.04562
2	0.00276	0.07719	[-0.11416, 0.11695]	0.05080	0.07720
3	0.00585	0.11831	[-0.16911, 0.19073]	0.07808	0.11841
4	0.00305	0.16189	[-0.24648, 0.26181]	0.10571	0.16185
5	0.00151	0.20470	[-0.33201, 0.33171]	0.13461	0.20462
6	0.00129	0.24605	[-0.39902, 0.41612]	0.16349	0.24595

2.3 结果分析

从表 1~表 3 可以看出, 无论是上游 5 号站水位, 还是下游 8 号站和 7 号站水位, 误差结果分析的各项指标均随着预见期的延长而增大, 说明预见期越长, 水位越难以准确预测; 对于 8 号站和 7 号站, 其连续 6 h 的预测误差的 90% 分位区间基本在 ± 0.2 m 以内, 而且 Y_{MAE} 和 Y_{RMSE} 值均较小, 总体达到水位预测的精度要求; 对于 5 号站, 连续前 3 h 的预测误差的 90% 分位区间控制在 ± 0.2 m 以内, 但是第 4 到 6 h 的 Y_{MAE} 值和 Y_{RMSE} 值均比较大, 误差波动性较大, 这主要是因为 5 号站处于三峡水电站和葛洲坝水电站之间, 水动力学特性远远比下游河道复

杂, 所以对于 5 号站水位的预测, 建议采取连续 3 h 预测, 及时更新的方法来确保预测的有效性。

3 结 语

本文建立了基于 LSTM 神经网络的梯级电站短期水位预测模型。考虑上游电站的运行对下游电站水位的影响以及时间的滞后性, 将上游电站的水位以及出力等可观测数据放入模型之中, 并采用梯度下降法与拟牛顿算法中的 BFGS 方法相结合训练模型, Wolfe-Powell 线搜索方法选取步长。将模型应用于葛洲坝水电站上下游水位预测, 结果表明模型实现了葛洲坝水电站连续 6 h 的下游水位准确预测, 以及连续 3 h 的上游水位准确预测, 为科学的调度决策提供了更加可靠的理论支撑。

参考文献:

- [1] 王蒙蒙, 戴凌全, 戴会超, 等. 基于支持向量回归的洞庭湖水位快速预测[J]. 排灌机械工程学报, 2017, 35(11): 954-961. (WANG Mengmeng, DAI Lingquan, DAI Huichao, et al. Support vector regression based model for predicting water level of Dongting Lake[J]. Journal of Drainage and Irrigation Machinery Engineering, 2017, 35(11): 954-961. (in Chinese))
- [2] 马辉, 孙颖桃, 肖艳, 等. 基于灰色-BP 神经网络组合模型的水位预测案例[J]. 人民黄河, 2016, 38(12): 89-92. (MA Hui, SUN Yingtao, XIAO Yan, et al. Case analysis of water level prediction based on grey-BP neural network combinational model[J]. Yellow River, 2016, 38(12): 89-92. (in Chinese))
- [3] 李欣, 王超, 赵虎川. 基于时空序列模型的 RBF 神经网络在河流水位预测中的应用[J]. 城市勘测, 2016(5): 34-39. (LI Xin, WANG Chao, ZHAO Huchuan. River water level forecast based on spatio-temporal series model and RBF neural network[J]. Urban Geotechnical Investigation and Surveying, 2016(5): 34-39. (in Chinese))
- [4] 余开华. 小波神经网络模型在河道流量水位预测中的应用[J]. 水资源与水工程学报, 2013, 24(2): 204-208. (YU Kaihua. Application of RBF and GRNN neural network model in forecast of water runoff and head[J]. Journal of Water Resources & Water Engineering, 2013, 24(2): 204-208. (in Chinese))
- [5] ASHAARY N A, ISHAK W H W, KU-MAHAMUD K R. Neural network application in the change of reservoir water level stage forecasting[J]. Indian Journal of Science and Technology, 2015, 8(13), 70634.

(下转第 78 页)