

• 人工智能 •

DOI:10.15961/j.jsuese.202300315



本刊网刊

基于融合注意力机制LSTM网络的地下水位自适应鲁棒预测

佃松宜¹, 厉潇滢¹, 杨 丹², 芮胜阳², 郭 斌^{1*}

(1. 四川大学 电气工程学院, 四川 成都 610065; 2. 成都兴蓉市政设施管理有限公司, 四川 成都 610065)

摘 要: 地下水水位是旱天污水管网地下水入渗量的重要影响因素, 快速精准地预测地下水水位能有效提升旱天污水管网地下水入渗量估算准确度, 辅助优化管网病害治理与维护策略。针对目前城市复杂水文预测存在的准确度低、灵敏度低、泛化能力弱等问题, 本文提出了一种新的鲁棒自适应水位预测算法。首先, 对水文数据进行预处理, 解决了数据时间跨度大、噪声多、缺失及异常、非平稳等问题。其次, 针对不同输入特征对预测指标的影响, 在模型训练阶段提出一种新的空间变量注意机制, 可快速识别与水位关联的关键变量, 并对输入特征赋予不同的影响权重。然后, 针对不同序列长度对预测效果的影响, 还设计了自适应时间注意力机制, 帮助网络自适应地找出与不同时间序列长度预测指标相关的编码器隐藏状态, 以更好地捕捉时间上的依赖关系。在此基础上, 以上下文向量作为输入, 提出一种融合注意力机制的长短时记忆网络水文预测算法。最后, 通过意大利Petrignano水文数据验证了所提算法的有效性, 并与GRU、Elman、LSTM、VA-LSTM和S-LSTM等方法进行预测性能比较。结果表明, 基于融合注意力机制的LSTM网络在面临大规模、噪点多的复杂数据时有优于其它几种算法的预测效果, 表明该算法具有强自适应性和鲁棒性。本文研究结果可以为市政排水策略合理调整、及时控制提供参考。

关键词: 地下水位预测; 时间与空间注意力机制; LSTM网络; 自适应预测; 鲁棒预测

中图分类号: TV211.12

文献标志码: A

文章编号: 2096-3246(2024)01-0054-11

2019年, 住房和城乡建设部、生态环境部和发展改革委联合发布了《城镇污水处理提质增效三年行动方案》(2019—2021)^[1-2], 方案明确要求进水生化需氧量(BOD)浓度低于100 mg/L的污水处理厂应开展系统化整治工作。城市污水管网的降雨入流及地下水入渗问题是污水处理厂进水浓度偏低的重要原因^[3]。在实际生活中, 旱天外来水入流、入渗水量受管网服务区域排水规律、管道结构条件等复杂因素影响^[4]。因此, 快速精准地预测地下水水位对预测污水浓度、准确计算旱天污水管网地下水入渗量^[5]、辅助管网病害治理与维护具有重要意义。

城市地下水水位除受到降水入流、旱时入渗等因素影响, 还受到其所在区域地质层及所铺设管网分布特征制约, 导致其水位动态变化, 呈现出复杂的

非线性、趋势性、季节性、滞后性及随机性等特征^[6], 导致其数据的不规律和杂乱性远超其他普通的时间序列数据, 其预测面临着独特的挑战和严峻性。在实际水文测量结果中, 易出现数据缺失、异常值、非平稳等质量问题^[7]; 除此之外, 水位数据的时空相关性和周期性规律差, 难以捕捉, 即特征和地下水位间的关联难以被模型挖掘。这系列问题为城市区域级地下水水位预测带来了重大挑战。

目前常用的地下水水位预测方法可大致分为以下3种: 基于物理模型的方法^[8]、基于时间序列分析的方法^[9]和数据驱动的预测方法^[10]。基于物理模型的方法是依据地下水流方程和边界条件等物理规律, 采用数值方法求解模型预测地下水位。但该方法原理复杂、水文参数定标难、模型开发周期长, 还易受

收稿日期: 2023-04-24

基金项目: 国家重点研发计划项目(2020YFB1709705)

作者简介: 佃松宜(1972—), 男, 教授, 博士。研究方向: 先进控制与人工智能算法; 机器人与自动化。E-mail: scudiansy@scu.edu.cn

* 通信作者: 郭 斌, E-mail: bguodxl@scu.edu.cn

网络出版时间: 2023-08-24 17:39:19

网络出版地址: <https://link.cnki.net/urlid/51.1773.TB.20230823.0915.001>

假设局限性和不确定性的影响^[11]。基于时间序列分析的方法是根据趋势、季节性和周期性^[12]等特征,通过对历史数据进行拟合来预测未来的数据^[13],但主要适用于具有周期性和趋势性的数据。以上两种传统的预测方法往往无法充分捕捉地下水位数据中的非线性关系和时空依赖性。基于数据驱动的方法是对已有数据进行学习和分析,通过构建模型完成预测^[14],其原理是基于数据的内在规律和模式识别。例如,杨建飞等^[15]通过建立灰色残差模型对宝鸡峡灌区地下水位进行预测,结果显示该模型能够有效克服数据序列不稳定带来的误差并提高预测精度;Barzegar等^[16]结合小波变换和神经网络模型对地下水位短期动态进行预测,取得了较为理想的效果。

由于水文数据的特殊性^[17],现有的基于数据驱动的方法不能保证对不同规模数据集的自适应能力^[18]。该方法需要大量数据来训练模型,但水文数据常见缺失、异常值等问题,数据质量会影响模型预测的准确性和鲁棒性^[19]。其次,数据的时空相关性和周期性并不显著,仅仅使用简单的线性模型或传统的神经网络模型不能够充分挖掘水文特征与地下水位间的关联^[20],无法满足预测的准确性和鲁棒性要求,故选择合适的预测模型尤为重要。随着深度学习模型的迅速发展,越来越多性能良好的模型被提出并被应用于不同的场景,例如长短时记忆网络(LSTM)^[21]、门控单元神经网络(GRU)^[22]等。其中,LSTM神经网络因方便进行时间序列建模,且具有长期记忆功能,在水位预测方面得到了广泛应用^[23-25]。但LSTM网络不能专注于不同时间序列长度的不同变量,在预测过程中不能保证预测效果的鲁棒性^[26]。

针对城市地下水位预测中的水位数据随机性与非线性强、时空依赖明显、数据波动规律难捕捉等特点,结合现有预测模型单一、处理时空相关任务弱、灵敏度低、泛化能力弱、鲁棒性能力低等问题,本文提出基于时空注意力机制的LSTM(STA-LSTM)水位预测算法。主要创新点如下:1)基于空间注意力机制快速识别与水位相关联的关键变量,准确捕捉影响地下水位的多因素间的依赖模式和关联性;2)时间注意力机制捕捉时间上的依赖关系并自适应地找出不同时间序列长度下与地下水位相关的编码器隐藏状态,可更好捕捉地下水位数据的复杂性和不确定性;3)能够动态地调整权重,提高水位的预测精度,且算法对复杂水位数据有较好的自适应性和较强的鲁棒性,对于城市区域级地下水位预测具有重要的参考价值与工程意义。

1 基于STA-LSTM的水位预测方法

传统的LSTM网络是一种循环神经网络,但相比于原始的循环神经网络的隐层部分,LSTM增加了细胞状态,能保证信息不变地流过整个循环神经网络。LSTM网络还引入门机制用于控制特征的流通和损失,门机制包括输入门、遗忘门和输出门。门能够有选择性地决定让哪些信息通过,通过门机制对细胞状态进行删除或者添加信息。细胞状态和门机制使LSTM网络能够处理长期依赖关系。

本文提出的基于STA-LSTM的自适应鲁棒水位预测方法是在传统LSTM网络的基础之上加入空间注意力机制和时间注意力机制以提高该模型对关键信息的关注程度,主要可分为模型训练阶段和模型预测阶段。在模型训练阶段,基于训练数据计算局部注意力度量并构建相应的地下水位预测模型;在模型预测阶段,利用前一阶段所构建的预测模型进行地下水位预测。

1.1 模型训练阶段

本文提出的STA-LSTM是在传统LSTM网络基础上加入融合注意力机制^[27],其遵循编码器-解码器体系^[28]结构,如图1所示。

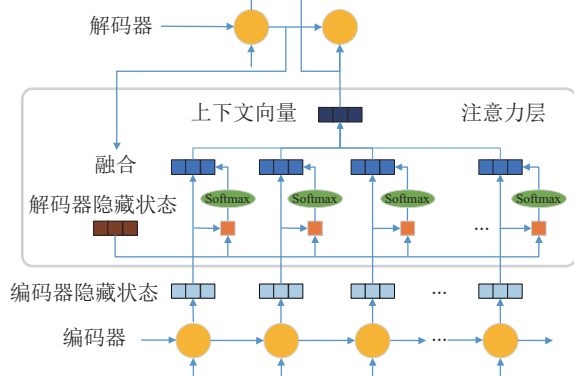


图1 基于注意力机制的编码器解码器结构

Fig. 1 Encoder-decoder structure based on attention mechanism

STA-LSTM首先在编码器中引入空间注意力机制(S-A),可快速识别与水位相关联的关键变量,并为输入变量分配不同的空间注意力值。可变的注意力加权样本数据作为编码器LSTM的新输入。之后,编码器LSTM网络学习新的加权输入的隐藏状态。在第二步中引入时间注意力机制,自适应地找出不同时间序列长度下与地下水位相关的编码器隐藏状态,捕捉时间上的依赖关系。最后,以自适应上下文向量为输入,由另一个解码器LSTM预测输出地下水位的预测值。完整的STA-LSTM算法结构如图2所示。

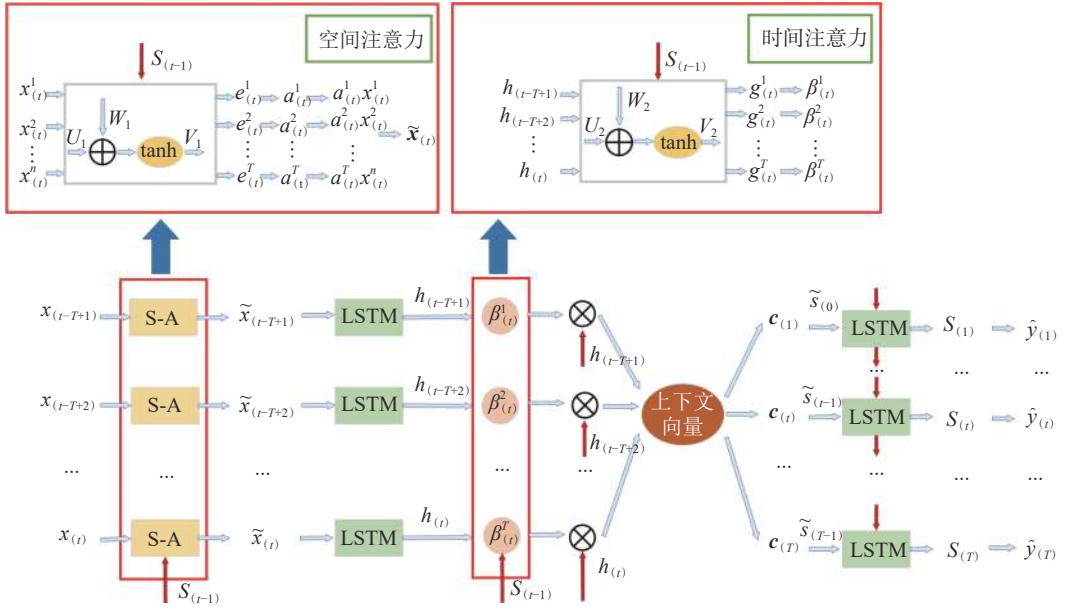


图 2 STA-LSTM算法结构

Fig. 2 Structure of STA-LSTM algorithm

1.1.1 空间注意力机制

由于输入的特征对预测指标影响程度不同,不能简单地给输入特征赋予相同的影响权重,故在算法模型训练阶段提出空间变量注意力机制,可快速识别与水位关联的关键变量。

给定输入序列 $\{x_{(1)}, x_{(2)}, \dots, x_{(T)}\}$, 假设每个样本有 n 个输入变量, 即 $x_{(t)} = \{x^1_{(t)}, x^2_{(t)}, \dots, x^n_{(t)}\}$ 。通过参考先前的解码器隐藏状态获得每个输入变量与地下水位指标之间的关系, 通常使用两层网络来获取可变的注意力值:

$$e^i_{(t)} = V_1^i \tanh(W_1^i s_{(t-1)} + U_1^i x^i_{(t)} + b_1^i) \quad (1)$$

$$\alpha^i_{(t)} = \frac{|e^i_{(t)}|}{\sum_{j=1}^n |e^j_{(t)}|}, 1 \leq i \leq n \quad (2)$$

式(1)~(2)中: V_1^i 、 W_1^i 、 U_1^i 和 b_1^i 是要学习的参数, 由于地下水位及其相关的输入特征存在负值, 故选取 $\tanh$ 为激活函数; $e^i_{(t)}$ 表示第 i 个输入变量在时间 t 对地下水位的重要性的注意力值; $s_{(t-1)}$ 包含与地下水位数据相关的信息, 一旦确定了空间注意力网络的参数, 就通过解码器隐藏状态 $s_{(t-1)}$ 和输入数据 $x_{(t)}$ 之间的相似性计算来获得输入和输出之间的相关性; $\alpha^i_{(t)}$ 表示空间注意力权重, 确保在每个采样时刻 t 所有输入变量的注意值和为 1。

当得到样本 $x_{(t)}$ 的输入变量空间注意力值时, 空间注意力加权样本表示为:

$$\tilde{x}_{(t)} = \{\alpha^1_{(t)} x^1_{(t)}, \alpha^2_{(t)} x^2_{(t)}, \dots, \alpha^n_{(t)} x^n_{(t)}\} \quad (3)$$

使用 LSTM 网络从可变注意力加权输入 $\tilde{x}_{(t)}$ 中学习编码器的隐藏状态。

$$f_{(t)} = \sigma(W_{fx} \tilde{x}_{(t)} + W_{fh} h_{(t-1)} + b_f) \quad (4)$$

$$i_{(t)} = \sigma(W_{ix} \tilde{x}_{(t)} + W_{ih} h_{(t-1)} + b_i) \quad (5)$$

$$o_{(t)} = \sigma(W_{ox} \tilde{x}_{(t)} + W_{oh} h_{(t-1)} + b_o) \quad (6)$$

$$\tilde{c}_{(t)} = \tanh(W_{cx} \tilde{x}_{(t)} + W_{ch} h_{(t-1)} + b_c) \quad (7)$$

$$m_{(t)} = f_{(t)} \odot m_{(t-1)} + i_{(t)} \odot \tilde{c}_{(t)} \quad (8)$$

$$h_{(t)} = o_{(t)} \odot \tanh(m_{(t)}) \quad (9)$$

式(4)~(9)中: $f_{(t)}$ 、 $i_{(t)}$ 、 $o_{(t)}$ 、 $\tilde{c}_{(t)}$ 表示编码器 LSTM 单元的遗忘门、输入门、输出门和中间状态; $\odot$ 表示两个向量的逐点乘法; σ 是非线性 sigmoid 激活函数; W_{f*} 、 W_{o*} 、 W_{i*} 、 W_{c*} 、 b_f 、 b_i 、 b_o 、 b_c 是需要学习的参数; $m_{(t)}$ 为存储单元。在基于空间注意的 LSTM 之后, 隐藏状态 $h_{(t)}$ 被用作基于时间注意力机制的 LSTM 的输入。

1.1.2 时间注意力机制

针对不同序列长度对预测效果的影响, 设计了自适应时间注意力机制, 帮助网络自适应地找出与不同序列长度预测指标相关的编码器隐藏状态, 以更好地捕捉时间上的依赖关系。

编码器之后, 使用另一个基于 LSTM 的解码器输出地下水位预测值 $\hat{y}_{(t)}$ 。在时间 t 的隐藏状态时间注意

值为:

$$g_{(t)}^k = V_2^k \tanh(W_2^k s_{(t-1)} + U_2^k h_{(t-T+k)} + b_2^k) \quad (10)$$

$$\beta_{(t)}^k = \frac{|g_{(t)}^k|}{\sum_{m=1}^T |g_{(t)}^m|}, 1 \leq k \leq T \quad (11)$$

式(10)~(11)中, V_2^k 、 W_2^k 、 U_2^k 和 b_2^k 是关于第 k 条数据要学习的参数, $g_{(t)}^k$ 表示时间 t 的第 k 个编码器隐藏状态的注意值。式(11)用于计算时间注意力权重 $\beta_{(t)}^k$, 确保所有隐藏状态的注意值权重和为1, T 为子窗口大小。计算所有编码器隐藏状态的时间加权之和, 以获得自适应上下文向量:

$$\mathbf{c}_{(t)} = \sum_{k=1}^T h_{(t-T+k)} \beta_{(t)}^k \quad (12)$$

$\mathbf{c}_{(t)}$ 与目标序列 $\{y_{(1)}, y_{(2)}, \dots, y_{(t-1)}\}$, 即地下水位真实值相结合, 以更新解码器隐层状态:

$$\tilde{s}_{(t-1)} = \mathbf{W}_3 y_{(t-1)} + \mathbf{V}_3 s_{(t-1)} + \mathbf{b}_3 \quad (13)$$

$$s_{(t)} = f_l(\mathbf{c}_{(t)}, \tilde{s}_{(t-1)}) \quad (14)$$

式(13)~(14)中, $\mathbf{W}_3$ 和 $\mathbf{V}_3$ 为权重矩阵, $\mathbf{b}_3$ 为偏置向量, $f_l(\cdot)$ 为LSTM的基础单元, 地下水位输出 $\hat{y}_{(t)}$ 为:

$$\hat{y}_{(t)} = F(y_{(1)}, y_{(2)}, \dots, y_{(t-1)}, x_{(1)}, x_{(2)}, \dots, x_{(t)}) = \mathbf{V} \cdot (f_l([\tilde{s}_{(t-1)}; \mathbf{c}_{(t)}])) + \mathbf{b}_v \quad (15)$$

式中, $[\tilde{s}_{(t-1)}; \mathbf{c}_{(t)}]$ 为 $(t-1)$ 时刻的解码器隐藏状态和 t 时刻的上下文向量的串联, $\mathbf{V}$ 为权重矩阵, $\mathbf{b}_v$ 为偏置向量。

模型训练选用Adam算法, 其在计算时间和内存需求方面优于动量梯度下降法和均方根反向传播算法。使用标准反向传播最小化均方误差 (MSE) 学习模型的参数:

$$\Theta_{\text{MSE}} = \frac{1}{T_{\text{training}}} \sum_{t=1}^{T_{\text{training}}} (y_{(t)} - \hat{y}_{(t)})^2 \quad (16)$$

式中, $y_{(t)}$ 和 $\hat{y}_{(t)}$ 分别是采样时间 t 时的真实值和预测值。 T_{training} 为训练数据的总数。

1.2 模型预测阶段

在模型预测阶段, 对于历史水文数据, STA-LSTM网络训练出了最佳参数, 网络结构和权重加载到内存中。当把测试集的水文数据输入到训练好的网络中时, STA-LSTM网络利用输入的序列数据, 预测出下一个时刻的地下水水位输出。

STA-LSTM预测地下水位的流程如下所示。通过时间反向传播计算参数的梯度, 使用Adam算法更新网络参数。

1) 设置输入及初始参数: 数据集 $D = (x_{(t)}, y_{(t)}), t = 1, 2, \dots, k$, $x_{(t)} = \{x_{(t)}^1, x_{(t)}^2, \dots, x_{(t)}^n\}$, 学习率 r , 隐层神经元个数 N , 单次用以训练的数据个数 M , 训练轮数 E 。

编码器阶段:

2) 利用式(1)对 a (从1到 T)计算空间变量注意值

$$e_{(t-a+1)}^i: e_{(t-a+1)} \leftarrow V_1 \tanh(W_1 s_{(t-a)} + U_1 x_{(t-a+1)})。$$

3) 利用式(2)对 a (从1到 T)计算注意力归一化 $\alpha_{(t-a+1)}^i$:

$$\alpha_{(t-a+1)}^i \leftarrow \frac{|e_{(t-a)}^i|}{\sum_{j=1}^n |e_{(t-a)}^j|}, 1 \leq i \leq n。$$

4) 计算空间注意力权重样本:

$$\tilde{x}_{(t-a+1)} \leftarrow \alpha_{(t-a+1)} \cdot x_{(t-a+1)}。$$

5) 更新编码器隐层状态:

$$h_{(t-a+1)} \leftarrow f_{\text{LSTM1}}(\tilde{x}_{(t-a+1)}, h_{(t-a)})。$$

解码器阶段:

6) 对 b (从1到 T)计算时间变量注意值:

$$g_{(t)}^b \leftarrow V_2 \tanh(W_2 s_{(t-1)} + U_2 h_{(t-b+1)})。$$

7) 对 b (从1到 T)计算注意力归一化:

$$\beta_{(t)}^b \leftarrow \frac{g_{(t)}^b}{\sum_{a=1}^T g_{(t)}^a}。$$

8) 计算上下文向量:

$$\mathbf{c}_{(t)} \leftarrow \beta_{(t)}^1 h_{(t)} + \beta_{(t)}^2 h_{(t-1)} + \dots + \beta_{(t)}^T h_{(t-T+1)}。$$

9) 更新解码器隐层状态:

$$s_{(t)} \leftarrow f_{\text{LSTM2}}(\mathbf{c}_{(t)}, s_{(t-1)}, y_{(t-1)})。$$

10) 输出: 预测的地下水位 $\hat{y}_{(t)}, t = 1, 2, \dots, k$ 。

$$\hat{y}_{(t)} \leftarrow f_{\text{linear}}(s_{(t)})。$$

2 实例分析

为了验证基于时空注意力机制的LSTM网络对地下水位预测的有效性, 将其应用于水务领域地下水位预测过程中。模拟计算机的配置为: 操作系统为Windows 10, CPU是Intel(R) Core(TM) i7-8550U CPU @ 1.80 GHz 1.99 GHz, GPU是NVIDIA GeForce MX 150, RAM为8 GB, 编码工具为Python 3.6。

图3为基于时空注意力机制的LSTM网络(STA-LSTM)预测框架。

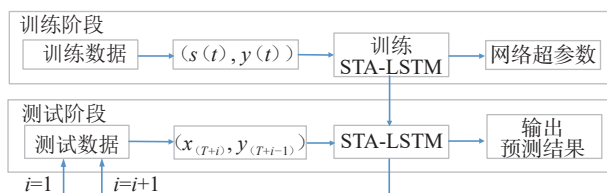


图 3 STA-LSTM模型框架

Fig. 3 STA-LSTM model framework

为比较多种模型的性能,分别建立了GRU、Elman^[29]、标准LSTM、VA-LSTM^[30]、S-LSTM^[31]和STA-LSTM。其中,GRU是类似LSTM的递归神经网络,具有更简化的结构和较少的参数,善于在建模长期依赖关系和短期相关性之间实现平衡。Elman循环神经网络借助隐藏层和反馈连接捕捉时序数据的依赖关系;LSTM通过记忆单元和门控机制来捕捉和处理序列数据中的长期依赖关系;VA-LSTM综合了变分自编码器和LSTM的思想,提高模型对不确定数据的建模能力;S-LSTM借助监督机制来捕捉数据间的时序关系和依赖关系。上述方法对水位类时序预测任务进行了研究,同样适用于地下水位的预测任务,将上述算法作为比较算法更具有合理性。选取平均绝对误差(θ_{MAE})、均方根误差(θ_{RMSE})和相关系数(θ_{R^2})作为评价指标, θ_{MAE} 度量预测值和真实值之间的平均差异,且对异常值不敏感、易于解释和计算,能体现模型的鲁棒性; θ_{RMSE} 度量预测值与真实值之间的均方根误差,对大误差的影响更加敏感; θ_{R^2} 衡量模型对观测数据的解释能力。3个评价指标能够更全面地综合评估模型的鲁棒性和自适应能力,计算方式如下:

$$\theta_{MAE} = \frac{1}{T_{\text{testing}}} \left(\sum_{i=1}^{T_{\text{testing}}} |y_{(i)} - \hat{y}_{(i)}| \right) \quad (17)$$

$$\theta_{RMSE} = \sqrt{\sum_{i=1}^{T_{\text{testing}}} (y_{(i)} - \hat{y}_{(i)})^2 / (T_{\text{testing}} - 1)} \quad (18)$$

$$\theta_{R^2} = 1 - \sum_{i=1}^{T_{\text{testing}}} (y_{(i)} - \hat{y}_{(i)})^2 / \sum_{i=1}^{T_{\text{testing}}} (y_{(i)} - \bar{y})^2 \quad (19)$$

式(17)~(19)中, T_{testing} 为测试样本数, $\bar{y}$ 是测试数据中真实值的平均值。

2.1 数据集及数据预处理

意大利小镇Petrignano的地下水位由Chiascio河以及地下含水层提供,本文将STA-LSTM网络应用于预测该镇的地下水位。地下水位受降雨量、地下水深度、温度和自来水管的取水量、Chiascio河的水位参数等的影响。数据时间跨度为2009年1月1日—2020年6月30日。数据来自Acea Smart Water Analytics^[32]。

输入特征包括降雨量、温度、河水位置、自来水取水量和地下水水位。

在自来水厂取水量和河水位置处出现异常零值时,将其替换成NaN后清除。对于清除后的数据以及缺失数据,选取插值填充的方式。具体的处理结果如图4所示。

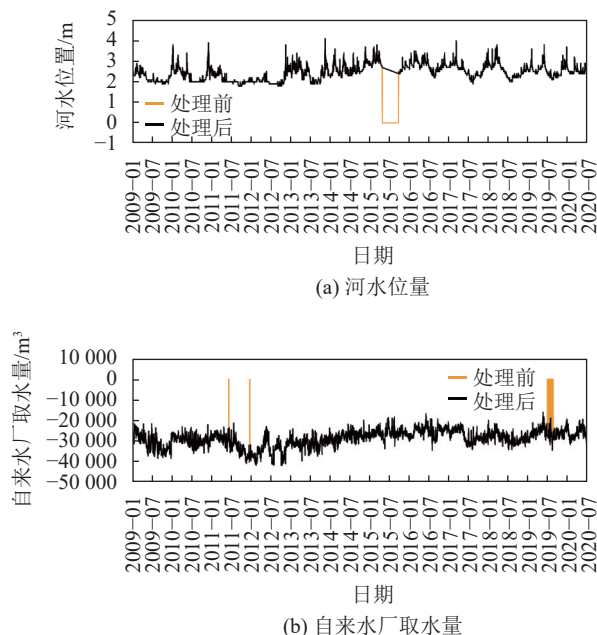


图 4 数据异常、缺失处理结果

Fig. 4 Results of processing data anomalies and missing data

数据可视化结果显示除温度外所有特征都具有非恒定的均值和方差。由于动态均值和动态方差的计算均要使用历史数据,动态均值和动态方差从2010年初开始记录,特征平稳性检验结果如图5所示。

本文采取差值处理的方法实现平稳性,以地下水水位的处理效果为例,差值处理过后数据趋于平稳,增广迪基-富勒检验(ADF)数据也证明该序列的平稳性增加,如图6所示。

2.2 结果和讨论

2.2.1 参数调整

总共收集了4 199条数据。对于模型训练和测试,用2 940个样本组成训练集,剩余1 259个样本为测试数据。对于STA-LSTM,需要优化序列长度参数,候选集合{3, 4, 5, 6}的 θ_{MAE} 、 θ_{RMSE} 和 θ_{R^2} 如图7所示。在该实验中序列长度选择为4, θ_{MAE} 、 θ_{RMSE} 都取得最小值, θ_{R^2} 取得最大值,此时训练模型达到最佳性能。

除此之外,还需要选择最适合的batchsize,不同的batchsize所对应的详细 θ_{MAE} 、 θ_{RMSE} 和 θ_{R^2} 如图8所示,在batchsize为50时, θ_{MAE} 、 θ_{RMSE} 都取得最小值, θ_{R^2} 取得最大值,此时模型达到最佳效果。另外,神经网络在空间和时间注意力机制都是两层神经网络。

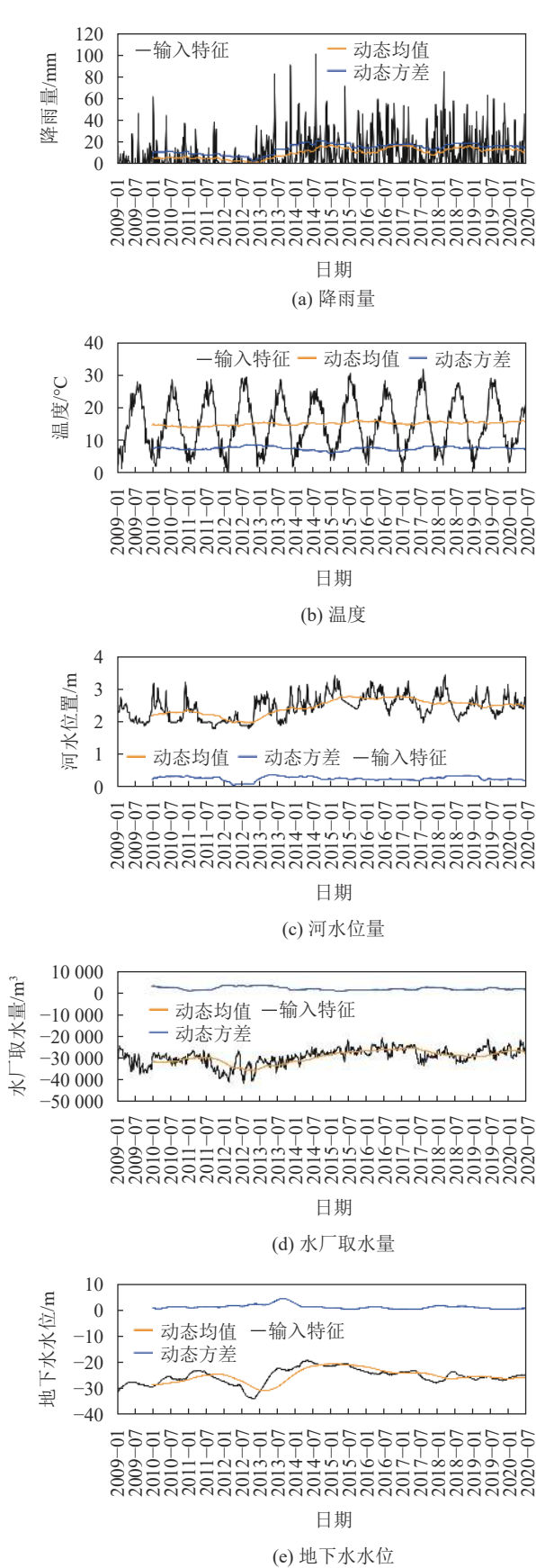


Fig. 5 Feature Stationarity test diagram

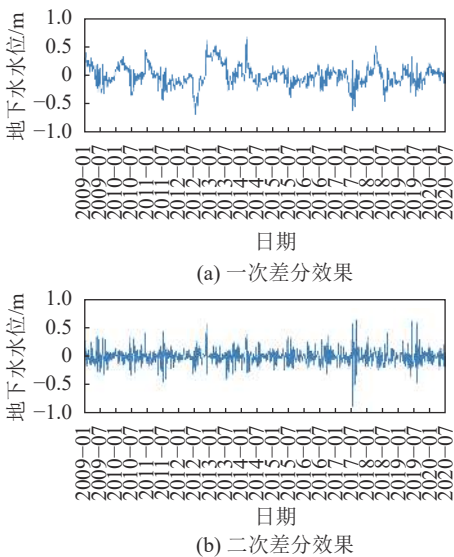


Fig. 6 Effect of difference processing to achieve stability

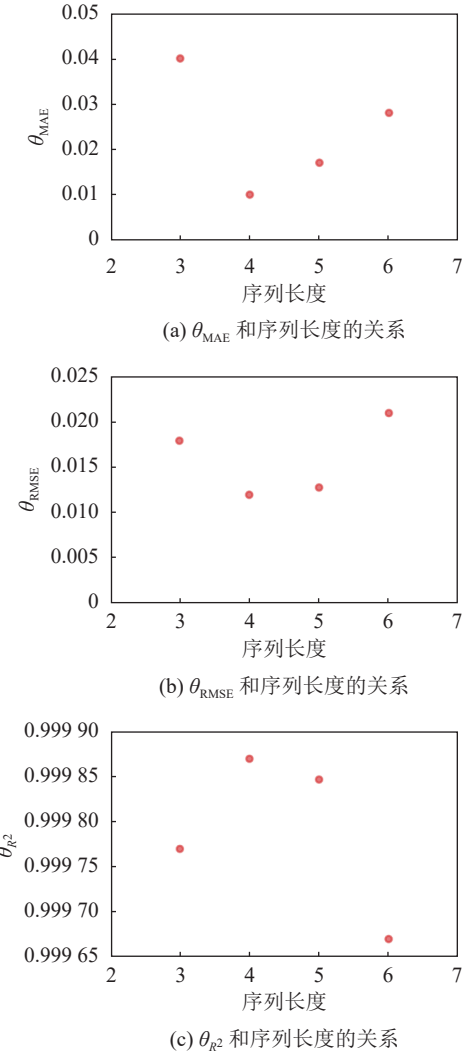


Fig. 7 Relationship between evaluation index and sequence length

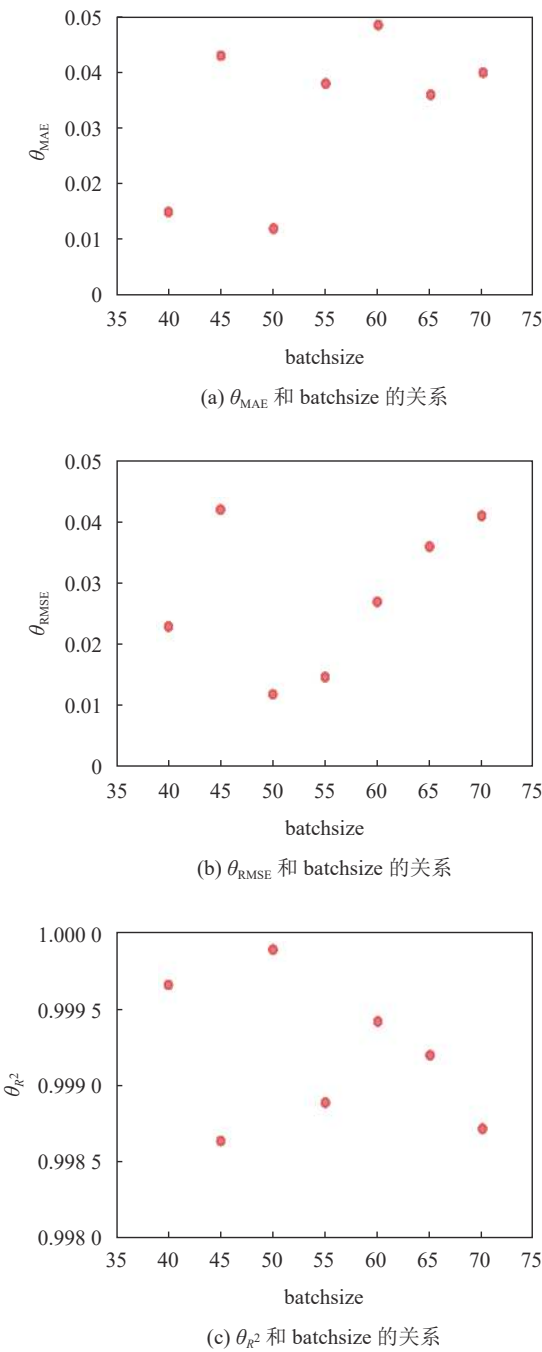


图 8 评价指标与batchsize之间的关系
Fig. 8 Relationship between evaluation index and batchsize

2.2.2 模型有效性验证

为了证明模型的有效性,有必要验证该模型的结构和参数.因此对该模型进行了一些统计测试,如残差自相关(ACF)以及残差偏自相关(PACF)。

$$P_{ACF} = \frac{Cov(X_t, X_{t-k})}{Var(X_t)} \quad (20)$$

$$P_{PACF} = Corr(X_t, X_{t-k} | X_{t-1}, X_{t-2}, \dots, X_{t-k+1}) \quad (21)$$

式(20)~(21)中,自相关函数的 $Cov(X_t, X_{t-k})$ 是时间点 t 和时间点 $(t-k)$ 的观测值的协方差, $Var(X_t)$ 是时间序列 X_t 的方差.偏自相关函数是当前时刻的观测值和滞后 k 期的相关系数,偏自相关函数通过递归的方式进行,从而得到每个滞后期的偏自相关系数.计算结果如图9所示,由图9(a)和9(b)可知,残差是相互独立的,表明该模型表现良好。

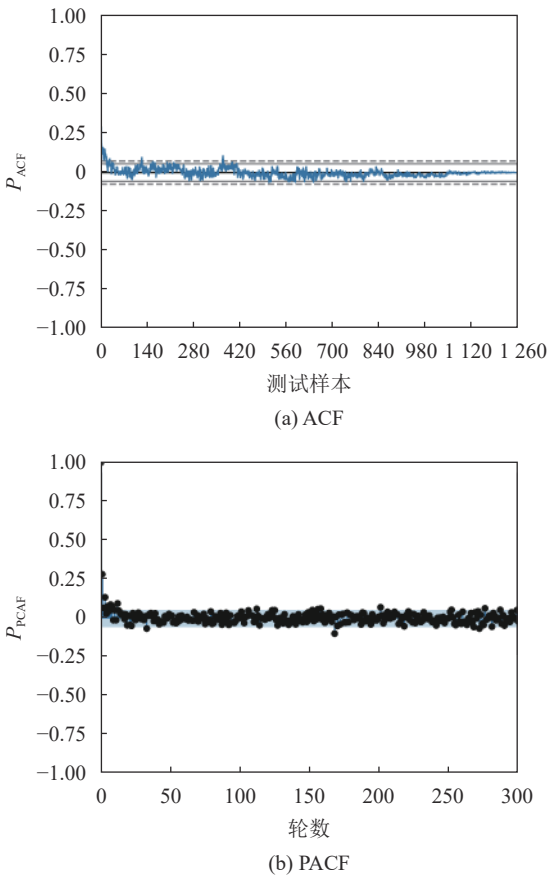


图 9 模型性能评价图
Fig. 9 Model performance diagram

2.2.3 实验结果对比

原始LSTM网络通过非线性激活函数提取数据的非线性特征,并使用存储单元存放用于预测的长期信息.但没有对不同时间序列长度的数据给予不同的关注,故可能丢失重要隐藏状态信息.基于注意力的LSTM模型利用注意力机制以解码器前一次隐藏状态作为参考,为编码器隐藏状态分配不同的关注权重,以达到高精度预测的目的.本文使用的STA-LSTM模型不仅能自适应地发现跨时间序列长度的相关样本,还能自适应地识别与预测指标相关的输入变量.图10为6种算法的预测值和真实值的对比. GRU网络参数相对较少,对于复杂地下水位数据处理难度稍大,图10(a)表明,GRU在极值处存在较为明显的预测失效;Elman网络因其只有一层全连接层

而无法充分捕捉数据中的复杂非线性关系, 图10(b)表现出预测趋势与真实值接近但预测误差较大, 其 θ_{RMSE} 高达0.799 43; 传统的LSTM由于注意力不足导致无法充分考虑数据中不同时间点之间的重要性差异, 在地下水位突变处尤容易出现预测失误; VA-LSTM重点处理数据的不确定性, 对于长期趋势和季节性预测十分受限, 面对时间跨度长达十余年的数据, 该算法表现较差, 其 θ_{MAE} 和 θ_{RMSE} 分别高达1.033 50和

1.274 78; S-LSTM未能充分考虑地下水位数据的特殊性从而影响预测性能, 预测值相对于真实值普遍偏低; 由于STA-LSTM算法中的空间-时间融合注意力机制不仅可快速识别与水位相关联的关键变量还能有效捕捉时间上的依赖关系, 实现动态调整权重, 故能有效提高水位的预测精度, 且算法对复杂水位数据有较好的自适应性和较强的鲁棒性。其 θ_{MAE} 和 θ_{RMSE} 低至0.011 66和0.014 62。

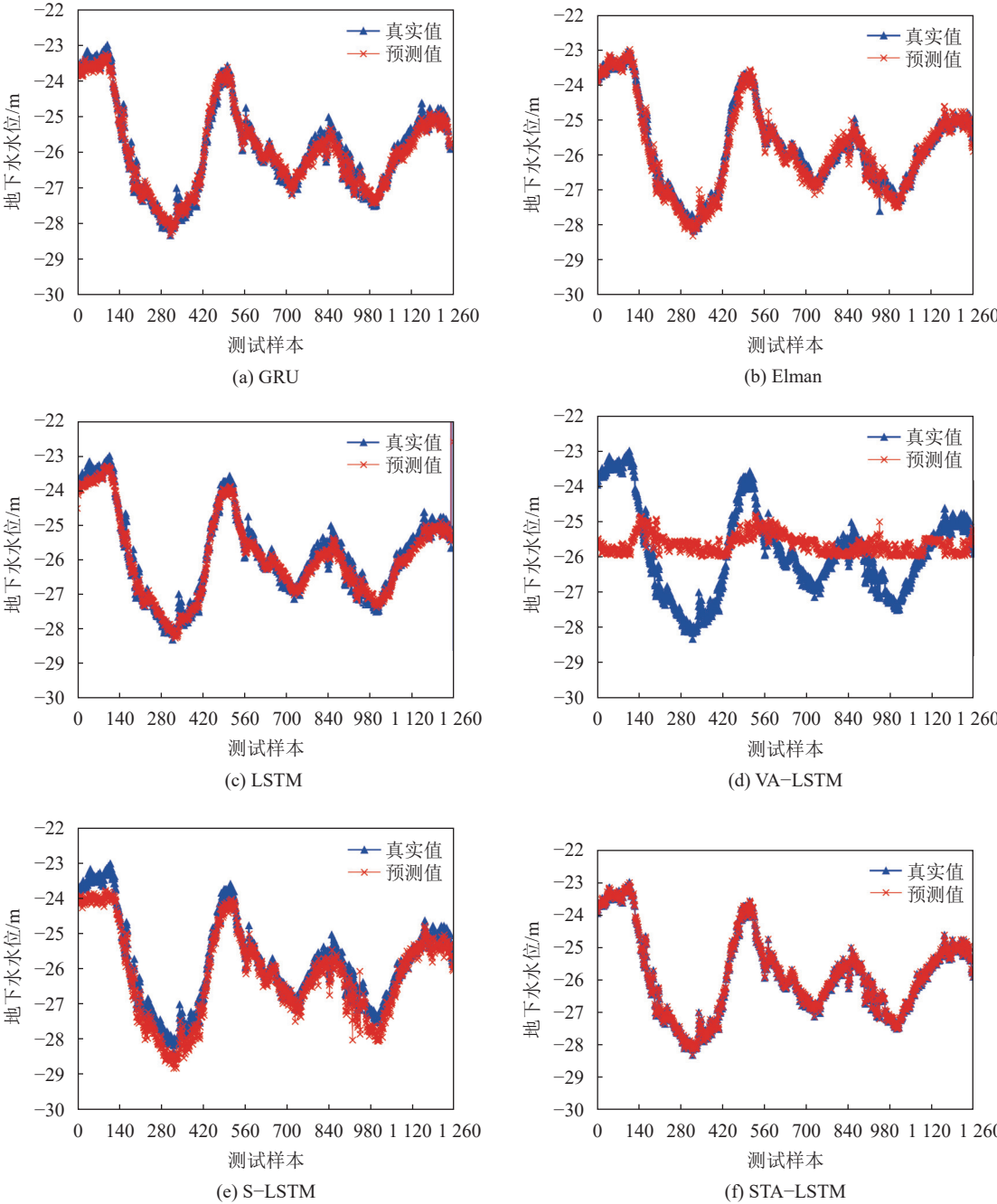


图 10 不同算法预测效果对比

Fig. 10 Comparison of prediction effects

图11展示了STA-LSTM算法的损失图和残差图, 可以看出, STA-LSTM的损失在100轮以前已经降至

较低水平, 残差较小且都稳定集中分布在0左右。表1比较了使用测试数据集的6个不同算法的预

测性能。从表1可以看出, STA-LSTM算法具有最低的 θ_{MAE} 和 θ_{RMSE} , 以及最高的相关系数, 证明其预测误差最小, 也证明了该算法在处理复杂、时间跨度大的数据集时展现出的较好的鲁棒性和自适应能力。由图10、表1可知, STA-LSTM弥补了LSTM面对长期历史数据时的注意力低的问题, 同时充分考虑地下水位数据中不同时间点之间的重要性差异以及数据的特殊性, 显示了本文所提出的算法的优越性与有效性。

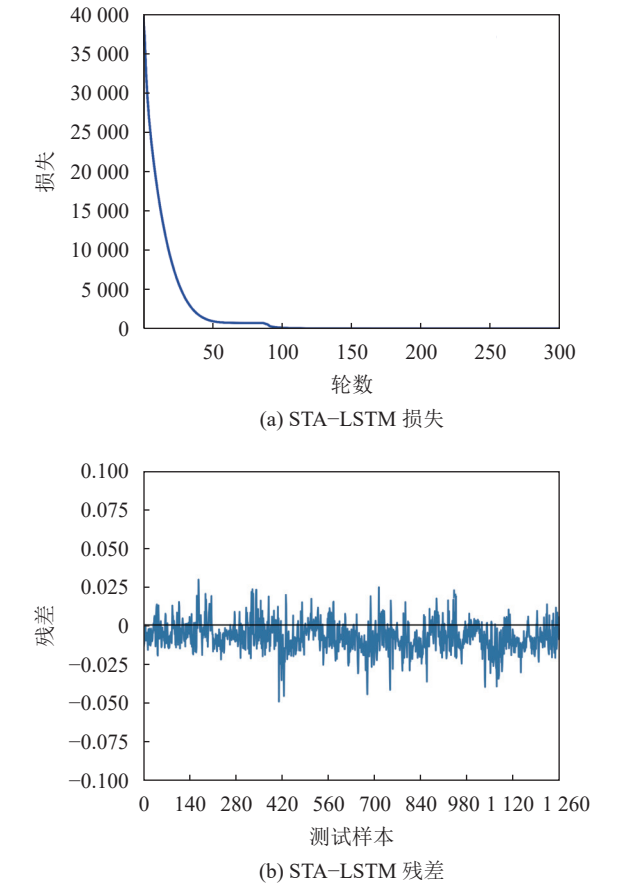


图 11 STA-LSTM损失图、残差图

Fig. 11 STA-LSTM loss diagram and residual diagram

表 1 不同预测方法的预测效果

Tab. 1 Prediction effects of different methods			
预测方法	θ_{MAE}	θ_{RMSE}	θ_{R^2}
GRU	0.205 87	0.517 94	0.957 98
Elman	0.679 74	0.799 43	0.823 74
LSTM	0.246 74	0.554 72	0.942 33
S-LSTM	0.402 07	0.575 55	0.790 07
VA-LSTM	1.033 50	1.274 78	0.054 53
STA-LSTM	0.011 66	0.014 62	0.999 87

为体现本文所使用的STA-LSTM算法面对复杂数据时的强适应性, 除了使用意大利小镇Petrignano的地下水位数据集验证算法有效性, 本文另外使用Powell湖水氧含量数据集和化工领域脱丁烷塔数据集来验证算法的泛化性能。Powell湖水氧含量数据集

是通过流入水量、总排出量、湖水pH值、湖水蒸发量、湖水温度、水压等对湖水氧含量进行预测; 脱丁烷塔数据集是通过塔顶温度和压力、回流流量、底部温度、流向下一阶段的流量等对丁烷浓度进行预测。表2和图12显示其预测效果仍然保持较好水平, 表明本文的算法在面对不同复杂数据时的强适应性和强鲁棒性。

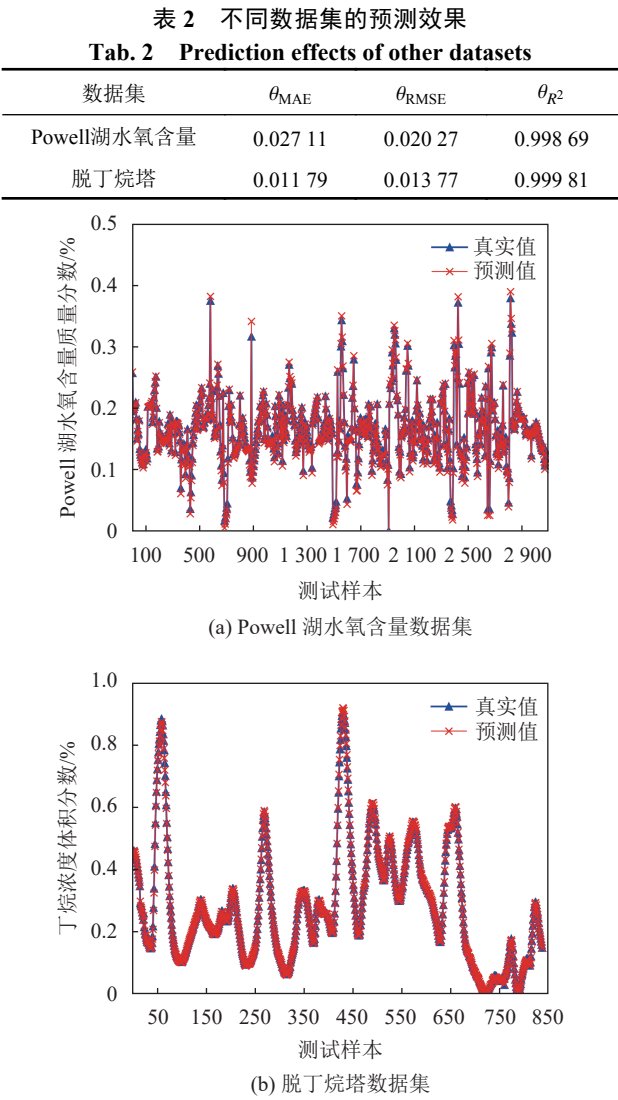


图 12 不同数据集预测效果对比

Fig. 12 Comparison of prediction effect of other datasets

3 结 论

本文针对水文预测中存在的准确度低、灵敏度低、泛化能力弱等问题, 提出了基于融合注意机制LSTM预测模型(STA-LSTM)。融合了时间注意力机制和空间注意力机制的LSTM网络可以动态地调整权重。其中, 空间注意力机制可快速识别与水位相关的关键变量, 更准确捕捉影响地下水位的多因素间的依赖模式和关联性; 时间注意机制能有效捕捉

时间上的依赖关系并自适应地找出不同时间序列长度下与地下水位相关的编码器隐藏状态。本文算法解决了传统算法中数据时空挖掘能力低、处理非线性能力弱、预测结果准确性与鲁棒性等问题,有效补偿了由于数据影响多元带来的数据杂乱和规律复杂造成的预测难度,而且能有效提高水位的预测精度,对复杂水位数据有较好的自适应性和较强的鲁棒性。与对比算法的对比结果表明,提出算法的综合指标优于对比算法,平均绝对误差、均方根误差和相关系数分别为0.011 66、0.014 62、0.999 87。

参考文献:

- [1] Sun Yongli, Wu Fansong, Li Wenqiu, et al. Reflections on the centralized collection rate of municipal sewage pollutants and the influent concentration of wastewater treatment plants [J]. *Water & Wastewater Engineering*, 2023, 49(1): 41–46. [孙永利, 吴凡松, 李文秋, 等. 城市生活污水集中收集率和污水处理厂进水浓度问题的思考[J]. *给水排水*, 2023, 49(1): 41–46.]
- [2] 中华人民共和国住房和城乡建设部. 住房和城乡建设部生态环境部 发展改革委关于印发城镇污水处理提质增效三年行动方案(2019—2021年)的通知[R/OL]. (2019-05-10)[2023-04-24]. https://www.mohurd.gov.cn/gongkai/zhengce/zhengcefilelib/201905/20190510_240490.html.
- [3] Wang Xiaoting, Yao Yue, Zhou Wei, et al. Quantification of inflow and infiltration in urban sewer systems based on triangle method[J]. *Water Pollution and Treatment*, 2019(4): 152–159. [王小婷, 姚越, 周炜, 等. 基于三角分析法的城市污水管网入渗量评估[J]. *水污染及处理*, 2019(4): 152–159.]
- [4] Franz T, Krebs P. Statistical methods towards more efficient infiltration measurements[J]. *Water Science and Technology*, 2006, 54(6/7): 153–160.
- [5] Cui Baoshan, Hua Yanyan, Wang Chongfang, et al. Estimation of ecological water requirements based on habitat response to water level in Huanghe River Delta, China[J]. *Chinese Geographical Science*, 2010, 20(4): 318–329.
- [6] Zhang Jianfeng, Zhu Yan, Zhang Xiaoping, 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: 918–929.
- [7] He Siwei, Nan Zhuotong. Application of rank set pair analysis method to predicting groundwater dynamic[J]. *Journal of Sichuan University(Engineering Science Edition)*, 2013, 45(Supp2): 55–60. [何思为, 南卓铜. 使用秩次集对方法预测地下水位动态变化[J]. *四川大学学报(工程科学版)*, 2013, 45(增刊2): 55–60.]
- [8] Afan H A, Ibrahim Ahmed Osman A, Essam Y, et al. Modeling the fluctuations of groundwater level by employing ensemble deep learning techniques[J]. *Engineering Applications of Computational Fluid Mechanics*, 2021, 15(1): 1420–1439.
- [9] Gribovszki Z, Kalicz P, Szilágyi J, et al. Riparian zone evapotranspiration estimation from diurnal groundwater level fluctuations[J]. *Journal of Hydrology*, 2008, 349(1/2): 6–17.
- [10] Zhao Zheng, Chen Weihai, Wu Xingming, et al. LSTM network: A deep learning approach for short-term traffic forecast[J]. *IET Intelligent Transport Systems*, 2017, 11(2): 68–75.
- [11] Liu Bo, Xiao Changlai, Liang Xiujuan. Application of combining SOM and RBF Neural Network Model for Groundwater Levels Prediction[J]. *Journal of Jilin University(Earth Science Edition)*, 2015, 45(1): 225–231.
- [12] Zhu Hongsheng, Wang Jihua, Chen Xin. Application of RBF neural network model in groundwater depth prediction[J]. *Yellow River*, 2020, 42(3): 50–54. [朱洪生, 王继华, 陈新. 基于改进人工蜂群算法的地下水埋深预测研究[J]. *人民黄河*, 2020, 42(3): 50–54.]
- [13] Abdel-Nasser M, Mahmoud K. Accurate photovoltaic power forecasting models using deep LSTM-RNN[J]. *Neural Computing and Applications*, 2019, 31(7): 2727–2740.
- [14] Le X, Ho H V, Lee G, et al. Application of long short-term memory(LSTM) neural network for flood forecasting[J]. *Water*, 2019, 11(7): 1387.
- [15] Yang Jianfei, Liu Junmin, Chen Lin. Forecast of the minimum depth of groundwater in irrigation district based on the residual model of GM(1,1)[J]. *Yellow River*, 2011, 33(7): 51. [杨建飞, 刘俊民, 陈琳. 基于灰色残差模型的灌区地下水最小埋深预测[J]. *人民黄河*, 2011, 33(7): 51.]
- [16] Barzegar R, Fijani E, Asghari Moghaddam A, et al. Forecasting of groundwater level fluctuations using ensemble hybrid multi-wavelet neural network-based models[J]. *Science of the Total Environment*, 2017, 599/600: 20–31.
- [17] Sun J S, Lou Y S, Chen Y J. Outlier detection of hydrological time series based on ARIMA-SVR model[J]. *Computer & Digital Engineering*, 2018, 2: 225–230.
- [18] Pan Mingyang, Zhou Hainan, Cao Jiayi, et al. Water level prediction model based on GRU and CNN[J]. *IEEE Access*, 2020, 8: 60090–60100.
- [19] Ma Zhijin, Zhang Bo. Research on the forecast dates of Ningxia-inner Mongolia reach freeze-up based on LSTM[J]. *Yellow River*, 2021, 43(2): 45–48. [马志瑾, 章博. 基于LSTM模型的宁蒙河段封河时间预测研究[J]. *人民黄河*, 2021, 43(2): 45–48.]
- [20] Tao Chao, Pan Hongbo, Li Yansheng, et al. Unsupervised spectral-spatial feature learning with stacked sparse autoencoder for hyperspectral imagery classification[J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(12): 2438–2442.
- [21] Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735–1780.
- [22] Zhang Zhanyu, Liang Zhenhua, Feng Baoping, et al. Groundwater level forecast based on principal component analysis and multivariate time series model[J]. *Advances in Water*

- Science,2017,28(3):415–420.[张展羽,梁振华,冯宝平,等.基于主成分-时间序列模型的地下水位预测[J].水科学进展,2017,28(3):415–420.]
- [23] Shen Chaopeng.A transdisciplinary review of deep learning research and its relevance for water resources scientists[J]. *Water Resources Research*,2018,54(11):8558–8593.
- [24] Ardana P H,Redana I W,Yekti M I,et al.Prediction of groundwater level in the shallow aquifer using artificial neural network approach[J]. *IOP Conference Series:Earth and Environmental Science*,2021,837(1):012010.
- [25] Gao Yifei,Wang Jianping,Li Linfeng.Outlier detection method of hydrological time series based on Cauchy distribution[J]. *Journal of Hohai University(Natural Sciences)*,2020,48(4):307–313.[高熠飞,王建平,李林峰.基于柯西分布的水文序列异常值检测方法[J].河海大学学报(自然科学版),2020,48(4):307–313.]
- [26] Kratzert F,Klotz D,Brenner C,et al.Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks[J]. *Hydrology and Earth System Sciences*,2018,22(11):6005–6022.
- [27] Vaswani A,Shazeer N,Parmer N,et al.Attention is all You need[EB/OL].2017:arXiv:1706.03762.[2023-04-24].<https://arxiv.org/abs/1706.03762.pdf>.
- [28] Cho K,van Merriënboer B,Gulcehre C,et al.Learning phrase representations using RNN encoder–decoder for statistical machine translation[C]// *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.Doha:Association for Computational Linguistics,2014.
- [29] Elman J L.Finding structure in time[J]. *Cognitive Science*,1990,14(2):179–211.
- [30] Yuan Xiaofeng,Li Lin,Wang Yalin.Nonlinear dynamic soft sensor modeling with supervised long short-term memory network[J]. *IEEE Transactions on Industrial Informatics*,2020,16(5):3168–3176.
- [31] Yuan Xiaofeng,Li Lin,Wang Yalin,et al.Deep learning for quality prediction of nonlinear dynamic processes with variable attention-based long short-term memory network[J]. *The Canadian Journal of Chemical Engineering*,2020,98(6):1377–1389.
- [32] Acea Group.Acea Smart Water Analytic[DB/OL].[2023-04-24].<https://www.gruppo.acea.it/en>.

Adaptive Robust Prediction of Groundwater Level Based on Fusion Attention Mechanism LSTM Network

DIAN Songyi¹, LI Xiaoying¹, YANG Dan², RUI Shengyang², GUO Bin^{1*}

(1.School of Electrical Eng., Sichuan Univ., Chengdu 610065, China;

2.Chengdu Xingrong Municipal Facilities Management Co. Ltd., Chengdu 610065, China)

Abstract: Groundwater level is an important factor affecting groundwater infiltration of sewage pipe network in dry weather. Accurate prediction of groundwater level can effectively improve the accuracy of groundwater infiltration estimation in dry weather, and assist in optimizing pipe network disease control and maintenance strategies. Aiming at the problems of low accuracy, low sensitivity, and weak generalization ability in the current urban complex hydrological prediction, a new robust adaptive water level prediction algorithm was proposed in this paper. First, a prior processing was carried out on the hydrological data, which solved the problems of large time span, high noise, missing and abnormal, and non-stationary data. Secondly, in view of the influence difference of input features on predictive indicators, a new spatial variable attention mechanism was proposed in model training stage, which can quickly identify key variables associated with water levels and assign different influence weights to input features. Furthermore, in view of the influence difference of various sequence lengths on the prediction effect, an adaptive temporal attention mechanism was also designed to adaptively find out the hidden state of the encoder related to the predictors of different sequence lengths, so as to capture time dependencies. On this basis, with the context vector as the input, an LSTM hydrological prediction algorithm integrating attention mechanism was proposed. Finally, the effectiveness of the proposed algorithm was verified by the hydrological data of Petriano, Italy. The prediction performance was compared with GRU, Elman, LSTM, VA–LSTM and S–LSTM methods. The results showed that the proposed STA–LSTM network based on the fusion attention mechanism has a better prediction effect than other algorithms when faced with complex, large-scale, and noisy data, indicating the strong adaptability and robustness of the algorithm. The research results of the paper provide a reference for the reasonable adjustment and timely control of municipal drainage strategies.

Key words: groundwater level prediction; spatial-temporal attention mechanism; LSTM; adaptive prediction; robust prediction

(编辑 陈 雪)

引用格式: Dian Songyi,Li Xiaoying,Yang Dan,et al.Adaptive robust prediction of groundwater level based on fusion attention mechanism LSTM network[J]. *Advanced Engineering Sciences*,2024,56(1):54–64.[佃松宜,厉潇滢,杨丹,等.基于融合注意力机制LSTM网络的地下水位自适应鲁棒预测[J].工程科学与技术,2024,56(1):54–64.]