

回增帅, 赵立杰, 范文玉. 基于长短时记忆改进的 Informer 模型的污水 COD 与总氮预测[J]. 水利水电技术(中英文), 2026, 57 (8): 93-107. DOI: 10.13928/j.cnki.wrahe.2026.08.007

HUI Zengshuai, ZHAO Lijie, FAN Wenyu. Prediction of wastewater COD and total nitrogen based on Informer model improved by Long Short-Term Memory [J]. Water Resources and Hydropower Engineering, 2026, 57 (8): 93-107. DOI: 10.13928/j.cnki.wrahe.2026.08.007

基于长短时记忆改进的 Informer 模型的 污水 COD 与总氮预测

回增帅¹, 赵立杰², 范文玉¹

(1. 沈阳化工大学 环境与安全工程学院, 辽宁 沈阳 110142; 2. 沈阳化工大学
信息工程学院, 辽宁 沈阳 110142)

摘 要:【目的】监测和管理污水处理厂的出水水质指标对于环境保护至关重要。然而, 污水处理具有高度可变性以及难以预测的复杂生化过程。【方法】通过建立一个基于长短时记忆网络(Long-Short Term Memory, LSTM)-Informer 的污水处理厂出水 COD 和总氮预测模型, 首先采用长短时记忆模块提取进水水质指标中的时间依赖关系, 然后将提取的时间特征融入 Informer 网络模块中解决长期依赖的问题。并通过对几种典型模型(Convolutional Neural Network, CNN、LSTM、CNN-LSTM、Informer)性能评估指标的分析, 比较了各个模型的拟合效果和预测效果。【结果】试验结果表明: LSTM-Informer 模型能够准确预测出水 COD、TN 浓度, 其在试验数据集的预测值和实际值的均方误差分别为 $0.630 \text{ mg} \cdot \text{L}^{-1}$ 和 $0.105 \text{ mg} \cdot \text{L}^{-1}$, 均方根误差分别为 $0.79 \text{ mg} \cdot \text{L}^{-1}$ 和 $0.32 \text{ mg} \cdot \text{L}^{-1}$, 平均绝对误差分别为 $0.60 \text{ mg} \cdot \text{L}^{-1}$ 和 $0.24 \text{ mg} \cdot \text{L}^{-1}$, 平均绝对百分比误差分别为 2.16% 和 2.26%, 决定系数 R^2 分别为 0.97 和 0.90, 各项指标均优于几种典型神经网络模型; 在统计学结果显著性检验中, 验证在显著性水平为 0.05 时, p 值均大于其余对比模型。【结论】结果表明: 基于长短时记忆网络捕获短期时间依赖于 Informer 的多头稀疏自注意力获取长期时序特征所建立的 LSTM-Informer 模型为污水出水水质提供了长期监测和预测保证, 其准确性在评价指标与统计学检验中都取得了一定优势, 综合来讲可以为污水处理厂提前预警水质变化, 对实现智能化控制和优化处理模式具有重要意义。

关键词: 污水处理; 神经网络; LSTM-Informer 模型; COD 预测; 总氮预测; 注意力机制; 长时序预测; 水质预警

DOI: 10.13928/j.cnki.wrahe.2026.08.007

开放科学(资源服务)标志码(OSID):

中图分类号: X703

文献标志码: A

文章编号: 1000-0860(2026)08-0093-15



Prediction of wastewater COD and total nitrogen based on Informer model improved by Long Short-Term Memory

HUI Zengshuai¹, ZHAO Lijie², FAN Wenyu¹

(1. College of Environmental and Safety Engineering, Shenyang University of Chemical Technology, Shenyang 110142,

收稿日期: 2025-05-29; 修回日期: 2025-08-01; 录用日期: 2025-08-14; 网络出版日期: 2025-10-28

基金项目: 国家重点研发计划(2018YB1700200); 2023 年度辽宁省研究生教育教学改革研究项目(2023-132)

作者简介: 回增帅(1999—), 男, 硕士研究生, 主要从事水污染控制技术研究。E-mail: 1194831943@qq.com

通信作者: 范文玉(1972—), 男, 副教授, 硕士, 主要从事水污染控制与防护研究。E-mail: 676748879@qq.com

©Editorial Department of Water Resources and Hydropower Engineering. This is an open access article under the CC BY-NC-ND license.

Liaoning, China; 2. College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, Liaoning, China)

Abstract: [Objective] Monitoring and managing effluent quality indicators of wastewater treatment plants is critical for environmental protection. However, wastewater treatment involves highly variable and unpredictable complex biochemical processes. [Methods] A prediction model for effluent COD and total nitrogen in wastewater treatment plants was established based on a Long Short-Term Memory (LSTM)-Informer. The LSTM module was first used to extract temporal dependencies from the influent quality indicators. Subsequently, the extracted temporal features were integrated into the Informer module to address the issue of long-term dependency. Additionally, through the analysis of performance evaluation indicators of several typical models (Convolutional Neural Network, CNN, LSTM, CNN-LSTM, and Informer), the fitting performance and prediction performance of each model were compared. [Results] The experimental result showed that the LSTM-Informer model could accurately predict the effluent COD and TN concentrations. On the experimental dataset, the mean squared errors between the predicted and actual values were $0.630 \text{ mg} \cdot \text{L}^{-1}$ and $0.105 \text{ mg} \cdot \text{L}^{-1}$, respectively. The root mean square errors were $0.79 \text{ mg} \cdot \text{L}^{-1}$ and $0.32 \text{ mg} \cdot \text{L}^{-1}$, the mean absolute errors were $0.60 \text{ mg} \cdot \text{L}^{-1}$ and $0.24 \text{ mg} \cdot \text{L}^{-1}$, the mean absolute percentage errors were 2.16% and 2.26%, and the coefficients of determination (R^2) were 0.97 and 0.90, respectively. All these indicators outperformed those of several typical neural network models. In the statistical significance test, it was verified that the p -values of the LSTM-Informer model were all greater than those of the other comparative models at a significance level of 0.05. [Conclusion] The result demonstrate that the LSTM-Informer model, which captures short-term temporal dependencies via the LSTM network and extracts long-term temporal features through the multi-head sparse self-attention mechanism of the Informer, provides a reliable basis for the long-term monitoring and prediction of wastewater effluent quality. Its accuracy demonstrates advantages in both evaluation indicators and statistical tests. Overall, this model can provide early warning of water quality changes for wastewater treatment plants, which is crucial for achieving intelligent control and optimizing treatment processes.

Keywords: wastewater treatment; neural network; LSTM-Informer model; COD prediction; total nitrogen prediction; attention mechanism; long-term time series prediction; water quality early warning

0 引言

随着污水排放标准的逐步提高,我国对污水处理的运行管理和自动化水平提出了更严格的要求。活性污泥法作为一种常用的污水处理工艺,依托活性污泥中的微生物群落,通过吸附、分解和氧化污水中的有机物,并通过硝化、反硝化、磷的释放与吸收等反应,有效去除水中的氮磷污染物,从而确保出水水质符合排放标准^[1-2]。

活性污泥法工艺涉及生物反应、化学反应和物理反应,反应机制复杂且相互影响,受到多种因素制约。这使得污水处理厂在脱氮除磷过程中表现出不确定性、耦合性强和线性回归难以拟合等特征^[3]。传统的预测模型通常基于反应机理构建,但由于对复杂的生物、化学和物理反应的模拟精度有限,这些机理模型在出水水质预测中难以发挥有效作用。

随着人工智能领域的飞速发展,产生了机器学习、深度学习等诸多方法,其中深度学习具有极强的适应学习能力和动态映射能力,可以广泛应用到污水处理中。污水水质数据因其复杂非线性的特点,很难通过

机理性模型预测出水水质指标,而通过对污水时序性排列,寻找污水时间相关性的预测模型例如 ARIMA^[4]、决策树^[5]、随机森林^[6]、LSTM^[7]、GRU^[8]、Transformer^[9]等被开发出来。CAO 等^[10]开发了一种循环神经网络评估纳米 FeCu 在污水处理中的性能,将 pH、硝酸盐、亚硝酸盐和氨用作输入,建立了一对一与三对三模型,其中一对一模型的精度达到 R^2 为 0.87 到 0.98,三对三的平均 R^2 为 0.91。作为 RNN 的改进网络结构,LSTM 的出现与发展也为本研究方向提供了很大帮助。ZHANG 等^[11]开发了基于 LSTM 神经网络的深度学习模型,用于检测 WWTPs 摄入量中的延时 WQI,在案例中所提模型取得了良好的准确性, R^2 达到 0.914 1 和 0.923 9。马笑文等^[12]提出对污水处理过程中的进水负荷和出水水质采用长短时记忆网络以及卷积神经网络进行预测,取得了预期成果,说明长短时记忆与卷积网络的结合可以满足污水处理水质预测的要求。AN 等^[13]基于双阶段注意力改进的 LSTM 网络(DA-LSTM),与单个 LSTM 模型相比,DA-LSTM 模型对 COD、TP 和 TN 的预测精度分别提高了 5.10%、2.11%、14.47%。这个结果表明,输

入注意力(IA)和时间注意力(TA)显著提高了 LSTM 的预测性能,说明注意力机制的加入可以有效提升特征自适应加权和长期记忆聚焦的能力。陈湛峰等^[14]提出了基于时间和特征双注意力机制的 BiLSTM 模型对珠江口水质进行预测,相比于单一模型取得了更好的效果, MSE 、 MAE 、 $MAPE$ 分别降低了 17.4%、19.4%、26.4%,说明通过时间和局部特征两个方向提取信息对模型效果具有正向作用。张铭玮等^[15]提出了基于 CNN-LSTM 的水质预测模型,该模型结合了 CNN 与 LSTM 两者的优点,首先将水质数据输入到卷积神经网络中进行局部特征提取,降低数据维度;然后将水质数据特征输入到 LSTM 网络中进行学习训练,并输出预测结果,预测速度与精度均得到了提高。以上说明使用复合模型可以改进模型的数据处理能力,更好地满足出水水质预测的精度和效率要求。

近年来新提出的 Informer 模型,作为 Transformer 的改进形式,在预测长序列数据时具有良好的精度^[16]。2024 年何志铎团队基于 Informer 模型建立了工作面矿压预测,取得了很好的预测效果^[17]。CAO 等^[18]提出了一种基于改进的 Stacking 集成算法的 LSTM-Informer 模型(ISt-LSTM-Informer),通过整合两个基础模型的优点,实现了准确的中短期光伏发电功率预测。WANG 等^[19]提出了一种基于集成学习的 LSTM-Informer 模型来解决长期电力负荷预测问题,该模型底层采用长短期记忆网络(LSTM)作为学习者捕捉电力负荷的短期时间相关性,顶层采用 Informer 模型解决电力负荷的长期依赖预测,结果表明,该模型的长期预测比 LSTM 模型的 MSE 和 MAE 分别提高了 0.58 和 0.38。LSTM-Informer 在电力负荷时间序列预测中的应用为本文污水处理出水的时序预测提供了实验方法。安昱宁等^[20]基于 Informer 模型建立了污水处理厂出水总磷预测模型,结果表明 Informer 模型相比于 LSTM 的 MSE 、 $RMSE$ 、 MAE 分别降低了 21.95%、11.60%、28.65%, R^2 提高了 19.94%,具有明显的预测优势。但是传统的机器学习模型存在很多缺陷^[21-23],包括在复杂、非线性、多扰动的数据处理中效率不高,难以解决时序类与非框架化的数据自回归的问题^[24-25],这些缺陷的凸显,需要更高效,准确率更高的方法来解决。目前,深度学习在污水处理领域的应用日益广泛,其中包括水质预测、指标监测、系统优化以及参数测量等^[26-28]。这些技术不仅能够分析和掌握污水厂进水水质、各处理单元的运行状况以及出水水质的变化规律,对于提升污水厂脱氮除磷的效果具有重要的实际意义,还能帮助降低能耗

和药剂使用量,助力实现碳中和目标,推动我国污水管理的可持续发展^[29]。

本文以 Informer 模型为基准模型,结合 LSTM 长短时记忆方法进行改进,提出了一种 LSTM-Informer 模型的污水处理出水水质预测方法,并与 CNN、LSTM、CNN-LSTM、Informer 模型进行了对比测试。

1 试验方法与数据

1.1 数据采集过程

本试验所采用的试验数据均来自辽宁省某污水处理厂,该厂采用改进式 A^2/O 工艺。为控制每次试验数据变量,每次取样位置均在进水口、曝气池水面下 50 cm 处以及出水口处,切为保证数据精确性,每次取样均在一组东西两侧曝气池和二组东西两侧曝气池各取一次样本作为试验对象,对采集回来的样本数据进行加权平均值处理作为本试验的样本数据。

试验共持续 36 个月,每天取一次样本,单位为 d,共计采集样本 1 095 个。每个样本包括进水流量(Q)、进水 COD_{Cr} 浓度($iCOD$)、进水 BOD_5 浓度($iBOD$)、进水氨氮浓度(iNH_3-N)、进水 SS 浓度(iSS)、进水总氮浓度(iTN)、进水总磷浓度(iTP)、进水 pH(ipH)、生化池中溶解氧浓度(DO)、硝化液回流比(nei)、污泥回流比(wai)作为模型的输入变量;出水化学需氧量 COD_{Cr} ($oCOD$)、出水生化需氧量 BOD_5 ($oBOD$)、出水总氮(oTN)作为输出变量。

数据按 7 : 3 的比例进行数据分割,使用前 767 个数据集作为训练集,后 328 个数据集作为验证集,以此来检验模型的预测准确性。

1.2 数据预处理

本试验使用了 Hampel 滤波算法进行数据异常值与离群值处理,Hampel 滤波法是一种强大的工具,用于识别和去除时间序列或信号中的异常值,特别适用于处理含噪声的数据^[30]。它通过比较每个点与其周围的数据来判断它是否是异常点。如果该点与其周围数据的差异超出了某一门限值,则认为它是一个异常点,并用中位数进行替代。

其步骤为:首先定义一个窗口(通常为奇数),用于确定当前点的邻居数据。这个窗口定义了检测异常的范围。对于每个数据点,计算其邻域中的中位数(median)和中位绝对偏差(Median Absolute Deviation, MAD)。中位绝对偏差的计算公式为

$$MAD = \text{median}(|x_i - \text{median}(x)|) \quad (1)$$

计算一个阈值来判定异常值,其公式为

$|Threshold| = n_{\sigma} \times k \times MAD \tag{2}$

式中, k 为一个常数 1.482 6, n_{σ} 是可以调整的倍数, 通常为 3。

2021—2023 年各个水质特征值箱线如图 1 所示,

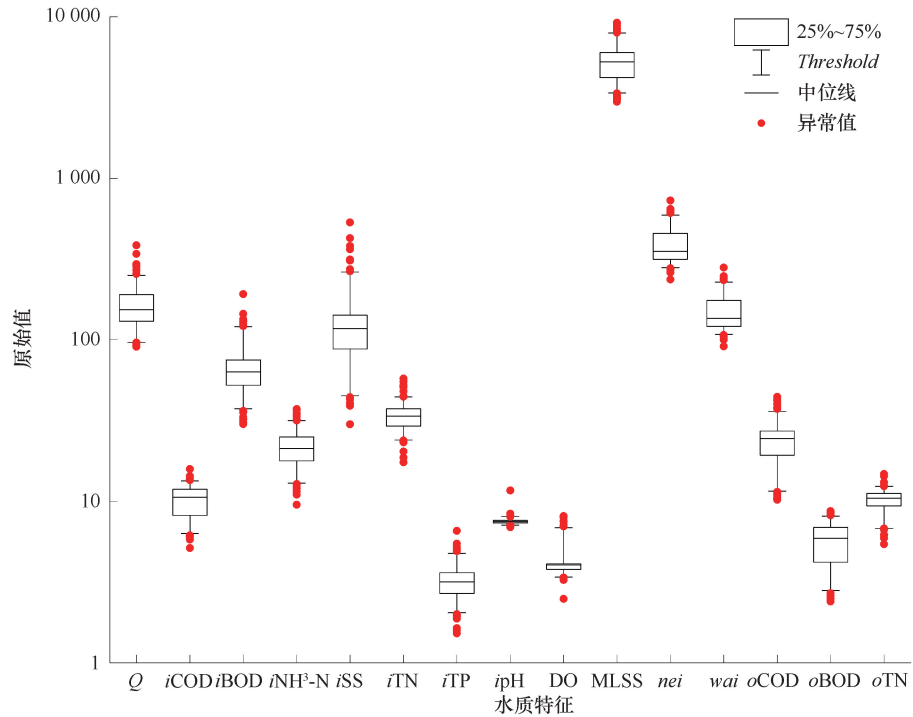


图 1 2021—2023 年各个水质特征值箱线图

Fig. 1 Boxplot of water quality characteristics (2021—2023)

2 基于 LSTM-Informer 建模

2.1 LSTM 模型

LSTM 模型结构如图 2 所示。

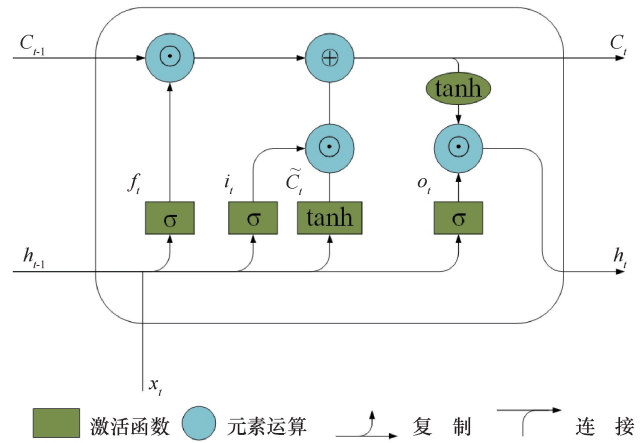


图 2 LSTM 模型结构

Fig. 2 LSTM model structure

循环神经网络(RNN)模型是一种处理数据节点之间依赖关系的人工神经网络(ANN)。RNN 模型中引入了隐藏状态的概念, 隐藏状态可以提取数据的特征进

将每个点与窗口的滚动中位数比较, 计算与中位数的绝对差值。如果该差值超过上述计算的阈值, 则该点被视为异常值。如果检测到异常值, 用其所在窗口内的中位数替代。

行变换然后输出。LSTM 模型作为 RNN 的改进变体, 引入了门控机制, 输入门(input gate)控制新信息进入细胞状态的程度; 遗忘门(forget gate)调节历史信息的保留与丢弃比例; 输出门(output gate)则筛选细胞状态中应当传递至下一时间步的信息^[31]。其公式为

$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$

$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{4}$

$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{5}$

$\tilde{C}_t = \tanh(x_t W_c \cdot [h_{t-1}, x_t] + b_c)$

$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$

$h_t = o_t \cdot \tanh(C_t) \tag{7}$

式中, i_t 、 f_t 、 o_t 分别为激活输入、遗忘、输出门; C_t 为更新记忆元; $\tilde{C}_t$ 为候选记忆元; h 为关于时间的隐含层; σ 为 sigmoid 激活函数; $\tanh$ 为双曲正切函数; W_i 、 W_f 、 W_c 、 W_o 分别为权重矩阵; b_i 、 b_f 、 b_c 、 b_o 分别为偏置项。

2.2 Informer 模型

Informer 模型结构如图 3 所示。由编码器(encoder, EN)和解码器(decoder, DE)组成。Informer

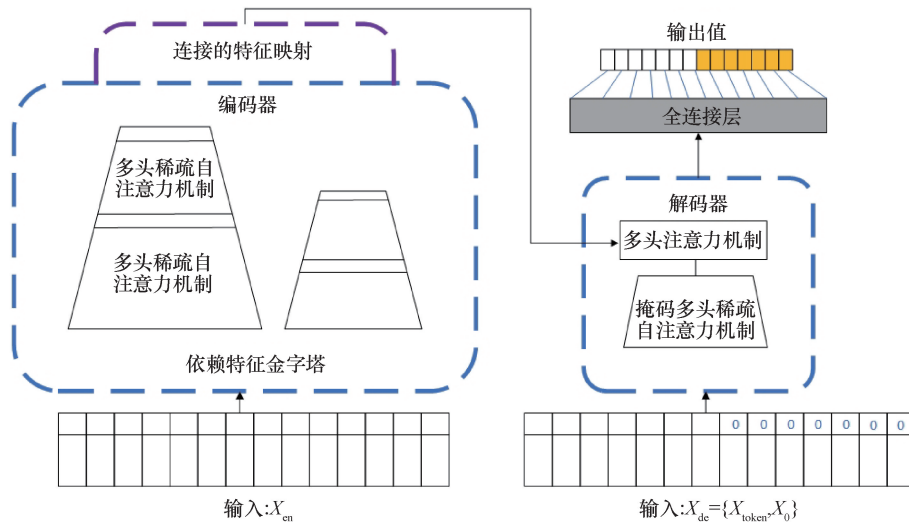


图3 Informer 模型结构

Fig. 3 Informer model structure

模型在输入向量时, 会添加位置编码(时间戳)信息来标记不同时间的位置关系, 以处理时间的相关性。与 Transformer 模型相比, Informer 模型的 ProbSparse 机制有选择地保留重要的注意力, 忽略不重要的信息, 从而解决长期依赖问题。DE 的输入一方面来自 EN 输出的隐含特征映射, 另一方面来自原始的输入向量, 将待预测的值赋为 0, 进行掩码操作以防止注意未来时间信息。最后将 DE 连接全连接层, 并输出预测值^[16]。

2.2.1 编码器-解码器结构

Informer 模型可通过编码器结构提取输入数据的特征并生成有关全局信息的向量, 然后将向量输入到解码器中, 通过解码可以得到预测数据^[32]。在实际预测过程中, 若输入输出数据为多指标序列, 会由于长度不同而导致部分指标预测精度不佳。编码器-解码器结构可以将输入数据中影响输出预测指标的因素通过编码器编码, 建立向量矩阵在输入到解码器中, 使得输入与输出长度保持一致从而达到合适的预测效果。在长时间序列预测中, 诸如 RNN 以及 LSTM 都存在随着时间步的增加, 会在其隐状态中压缩过多的信息, 导致信息损失, 使预测精度下降, 而编码器-解码器结构可以将长时间序列建立为向量矩阵解决这一问题。

2.2.2 多头概率稀疏自注意力机制

在长序列中, 部分元素对全局依赖关系的贡献非常小, 因此可以忽略这些不重要的注意力得分, 只保留重要的贡献部分。在计算自注意力时, 不再考虑每一个 Query 与所有 Key 的注意力得分, 而是基于稀疏

策略, 仅关注那些重要的键值对。通过稀疏选择, 降低计算复杂度, 同时保留模型对序列关键位置的关注能力。其公式为

$$A(Q, K, V) = \sum_j \frac{k(q_i, k_j)}{\sum_j k(q_i, k_j)} v_j = E_p \left(\frac{k_j}{q_i} \right) [v_j] \quad (8)$$

式中, A 为自注意力机制的计算过程; q_i 为 Q 中的第 i 行; k_i 、 v_i 分别为 K 、 V 中的第 j 行, Q 、 K 、 V 分别为 3 个输入(query, key, value); E 为概率期望。

稀疏性评价公式为

$$M(q_i, K) = \ln \sum_{j=1}^{L_K} e^{\frac{q_i k_j^T}{\sqrt{d}}} - \frac{1}{L_K} \sum_{j=1}^{L_K} \frac{q_i k_j^T}{\sqrt{d}} \quad (9)$$

式中, M 为评定方法; $\ln \sum_{j=1}^{L_K} e^{\frac{q_i k_j^T}{\sqrt{d}}}$ 为最大平滑值函数,

$\frac{1}{L_K} \sum_{j=1}^{L_K} \frac{q_i k_j^T}{\sqrt{d}}$ 为算术平均值。

最终得到注意力公式为

$$A(Q, K, V) = \text{Softmax} \left(\frac{\bar{Q} K^T}{\sqrt{d}} \right) V \quad (10)$$

Informer 还结合了多头注意力机制, 使得模型能够在不同的子空间中关注序列的多种模式, 其中每个头独立执行概率稀疏自注意力, 各头的输出经过线性映射后合并, 捕获不同维度和时间尺度, 提高模型能力。公式为

$$A_M(Q, K, V) = \text{Concat} (head_1, head_2, \dots, head_n) \quad (11)$$

式中, $A_M(Q, K, V)$ 为多头概率稀疏自注意力; $head_n$ 为第 n 个“头”, $head_n = A(Q_n, K_n, V_n)$, Q_n, K_n, V_n 分别为第 n 个“头”的 Q, K, V 矩阵。

2.3 LSTM-Informer 模型

基于上述单一 LSTM 与 Informer 模型的原理介绍, Informer 关键依赖自注意力机制捕获全局信息, 缺点是对局部细节敏感性不足, 易忽略短期变化或局部特征弱化, 而引入 LSTM 后, 其时间依赖性能增强模型对时间连续性的理解, 同时 LSTM 自带记忆机制, 能提供稳定的浅层信息提取, 辅助自注意力机制平衡全局依赖于局部细节。

结合污水处理厂水质数据建立的 LSTM-Informer 模型流程如图 4 所示。

本研究以 LSTM 作为建立模型的基学习器, 以 Informer 作为模型的元学习器。该流程为先使用基学习器 LSTM 对进水水质数据进行前期处理, 解决短期依赖的问题, 提取时间相关特征, 然后添加元学习器 Informer 模型来解决时间序列预测问题。

具体实施机制为串联策略, 在 Informer 模型编码

和解码过程中, 将 LSTM 嵌入自注意力层中, 将依照时间排列的进水水质数据特征, 经过 Hampel 滤波后得到如 $iCOD(t)$ 、 $iBOD(t)$ 、 $iNH_3-N(t)$ 等, 输入 LSTM 层中捕获短期时间相关性。公式为

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t \tag{12}$$

在当前时间点, 前一状态的记忆 C_{t-1} 和当前新输入的候选状态 $\tilde{C}_t$ 结合以形成新的状态, 即形成新的特征值。将原特征值格式更改为 epoch, range, characteristic 对应 Informer 中的 query, key, value 矩阵输入到 Informer 模型中再经过训练, 利用多头稀疏自注意力机制找出特征变化。公式为

$$F_m = A_M[y'(t)W_E, y'(t)W_R, y'(t)W_C] \tag{13}$$

式中, F_m 为特征变化; $y'(t)$ 为 LSTM 提取的时间特征; W_E, W_R, W_C 分别为 epoch、range、characteristic 矩阵。

再对 F_m 进行蒸馏操作, 提取出主要成分特征。蒸馏操作的步骤为: 对输入特征序列进行一次池化, 对时间维度进行采样, 得到长度降至原来的 1/2 至 1/4。

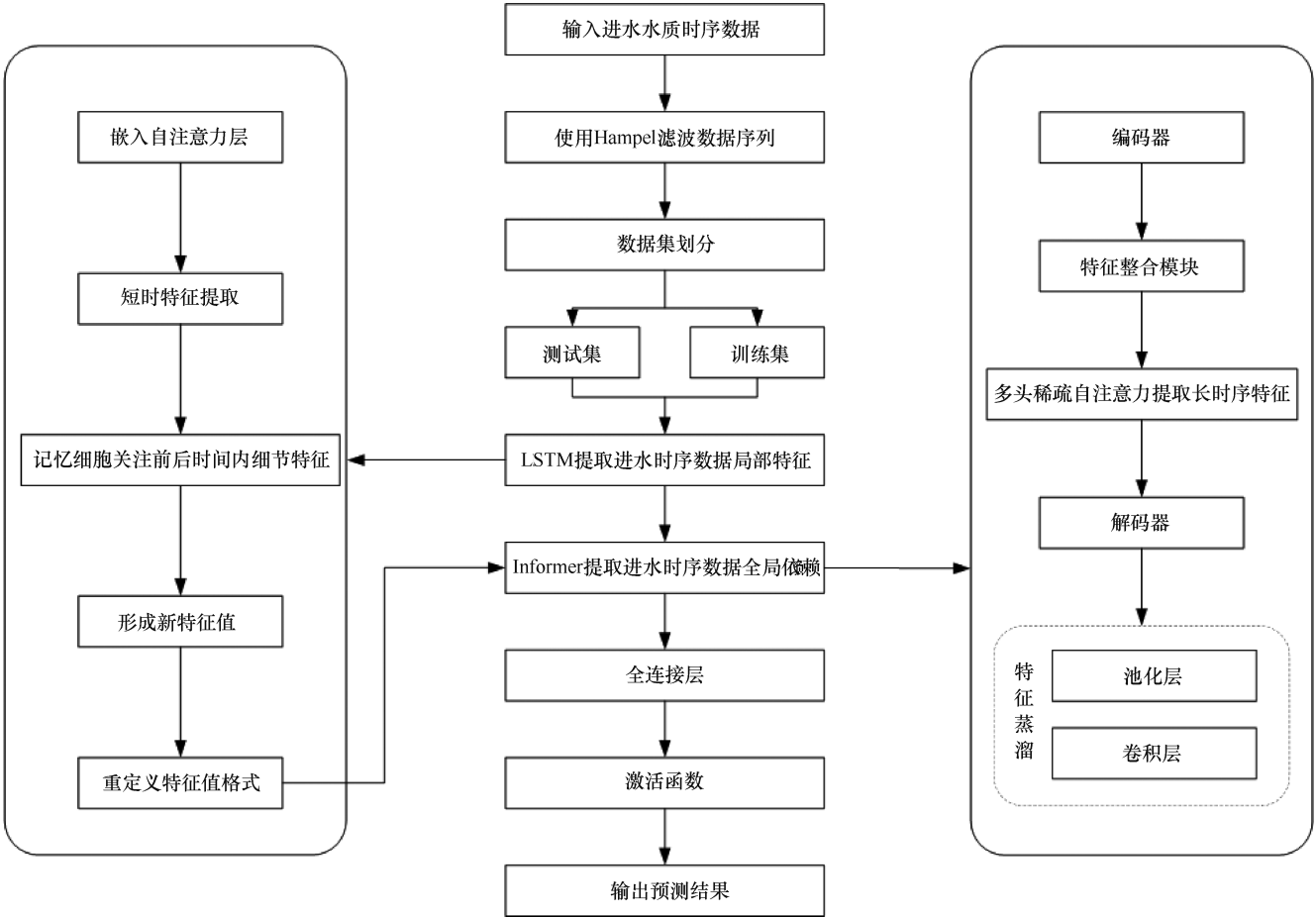


图 4 LSTM-Informer 模型流程结构
Fig.4 LSTM-Informer model structure

$$F_m^{pool}[i, j] = \max_{t \in [i \cdot s, i \cdot s + k - 1]} F_m[t, j] \quad (14)$$

式中, $pool$ 为最大池化; s 为步长; k 为窗口大小; t 为时间参数。

紧接着, 用一维卷积层作用于池化后的序列, 用于学习局部特征融合将得到的浓缩序列作为下一层的输入。公式为

$$F'_m = ReLU[Conv(F_m^{pool})] \quad (15)$$

式中, $ReLU$ 为激活函数; $Conv$ 为一维卷积。

通过多头稀疏自注意力进一步聚焦进水水质数据特征, 最后将特征向量经过全连接层处理输出多步预测值, 可以得到预测结果。

3 方法性能评估

3.1 试验环境

试验环境为 python3.10 torch2.5 框架,

windows10 专业版, 32 GB, Intel (R) Core (TM) i5-12600CPU@2.50 GHz, NVIDIA GeForce 3090Ti 显卡。

3.2 数据相关性分析

皮尔逊相关系数(PCC)衡量两个变量间线性关系的程度, 取值范围为 $[-1, 1]$, 绝对值越接近 1 表示相关性越强, 符号则代表相关方向。图 5 为各进水水质与出水水质之间相关性热力图。PCC 通常用 r 表示, 定义公式为

$$r = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{Y})^2}} \quad (16)$$

式中, x_i 为特征向量的值; y_i 为输出向量的值; $\bar{X}$ 和 $\bar{Y}$ 分别为两者的样本均值。

输出向量($oCOD$ 、 $oBOD$ 、 oTN)的 PCC 值特征

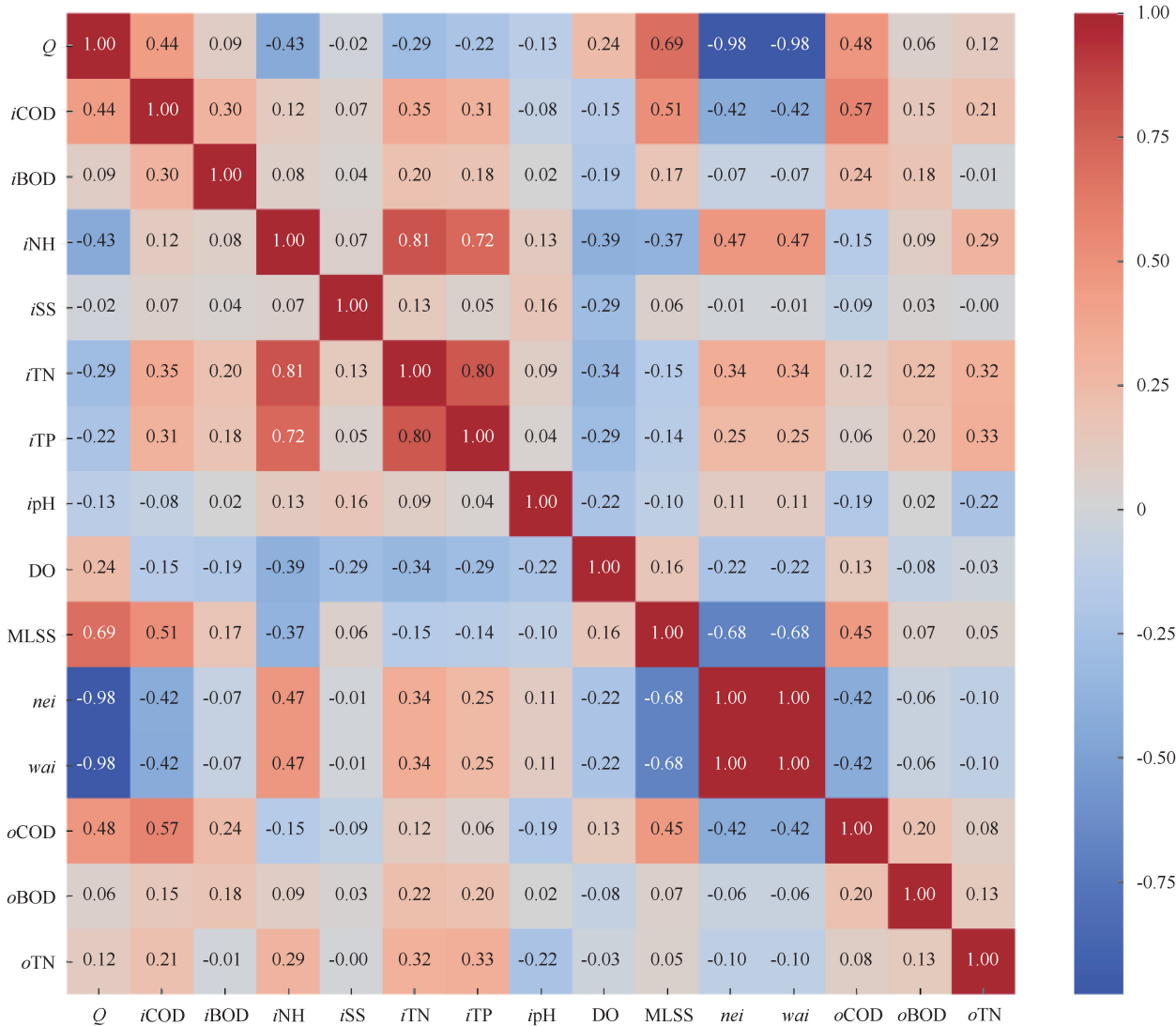


图 5 特征向量与输出向量相关性的热力图

Fig. 5 Heatmap of correlations between eigenvectors and output vectors

向量值如表 1 所列。

表 1 特征向量与输出向量的 PCC 值

Table 1 PCC values between eigenvectors and output vectors			
特征向量	出水 COD 浓度	出水 BOD 浓度	出水 TN 浓度
进水量	0.48	0.06	0.12
进水 COD _{Cr} 浓度	0.57	0.15	0.21
进水 BOD ₅ 浓度	0.24	0.18	-0.01
进水 NH ₃ -N 浓度	-0.15	0.09	0.29
进水悬浮固体浓度	-0.09	0.03	0.00
进水 TN 浓度	0.12	0.22	0.32
进水 TP 浓度	0.06	0.20	0.33
进水 pH 值	-0.19	0.02	-0.22
生化反应池 DO 浓度	0.13	-0.08	-0.03
生化反应池 MLSS	0.45	0.07	0.05
内回流比	-0.42	-0.06	-0.10
外回流比	-0.42	-0.06	-0.10

由表 1 可知，为避免冗余变量对模型训练造成干扰，选取 PCC 值大于 0.2 的特征向量作为模型输入数据。对于输出向量 $oCOD$ 、 Q 、 $iCOD$ 、 $iBOD$ 、 $MLSS$ 、内外回流比可作为特征向量；对于输出向量 oTN 、 $iCOD$ 、 iNH_3-N 、 iTN 、 iTP 、 $i pH$ 可作为特征向量；而对于输出向量 $oBOD$ ，其与特征向量的 r 值很低，故无法作为预测的输出向量考虑。综合来看，考虑到污水处理工艺是一个物理化学生物耦合的复杂动态过程，其具有非线性特点，可说明各特征向量与输出向量的相关性不高的现象是普遍的。

3.3 数据评估方法

通过计算均方误差 (MSE)、均方根误差 ($RMSE$)、平均绝对误差 (MAE)、平均绝对百分比误差 ($MAPE$) 以及决定系数 (R^2) 对模型精度进行评价，各评价指标的计算公式为

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{17}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{18}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{19}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{20}$$

$$R^2 = \frac{\left(n \sum_{i=1}^n y_i \cdot \hat{y}_i - \sum_{i=1}^n y_i \sum_{i=1}^n \hat{y}_i \right)^2}{\left[n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right] \left[n \sum_{i=1}^n \hat{y}_i^2 - \left(\sum_{i=1}^n \hat{y}_i \right)^2 \right]} \tag{21}$$

式中， n 为样本数量； y_i 为实际值； $\hat{y}_i$ 为 y_i 对应的预测值。

3.4 模型性能评估

3.4.1 模型网络参数

本文为验证 LSTM-Informer 模型的准确性，分别建立了 4 种神经网络预测模型进行对比测试，包括 CNN、LSTM、CNN-LSTM、Informer 模型。通过对本次试验所使用数据规模与复杂度，前人所做研究的经验，以及所使用设备、训练速度等多角度分析，各模型参数如表 2—表 6 所列。其中，部分超参数例如初始学习率设定为 0.01~0.001；batch-size 受设备限制设定为 64；优化器设定为 Adam；损失函数据本次试验使用数据为复杂型非线性数据，设定为 MSE 。

表 2 卷积神经网络参数

Table 2 Convolutional neural network parameters	
参数类型	参数值
卷积层	3
池化层	3
卷积核	3
全连接层神经元	100
激活函数	Relu
优化器	Adam 梯度下降算法
最大训练次数	500
初始学习率	0.001
下降因子	0.1

表 3 LSTM 网络参数

Table 3 LSTM network parameters	
参数类型	参数值
隐含层神经元	150
堆叠层	3
遗忘门激活函数	sigmoid
输入门激活函数	tanh
输出门激活函数	tanh
优化器	Adam 梯度下降算法
最大训练次数	500
初始学习率	0.01
下降因子	0.1
损失函数	MSEloss
Dropout	0.05

3.4.2 模型评价指标与分析

测试集中选择 2023 年 1—4 月共 110 d 的出水 COD、TN 实际值与预测值进行对比。如表 7 所列，为出水 COD 的 5 种预测模型的各项性能评估指标；实际值与预测值在测试集上的拟合结果如图 6 所示。

表 4 CNN-LSTM 网络参数

Table 4 CNN-LSTM network parameters

参数类型	参数值
卷积层	3
池化层	3
卷积核	3
全连接层神经元	100
激活函数	Relu
LSTM 隐含层神经元	125
遗忘门激活函数	sigmoid
输入门与输出门激活函数	tanh
优化器	Adam 梯度下降算法
最大迭代次数	500
初始学习率	0.01
下降因子	0.5
损失函数	MSEloss

表 5 Informer 网络参数

Table 5 Informer network parameters

参数类型	参数值
编码器	2
解码器	1
模型维度	32
多头注意力个数	8
激活函数	Gelu
Batch-size	64
优化器	Adam 梯度下降算法
最大训练次数	100
初始学习率	0.001
早停比例	2
损失函数	MSEloss

表 6 LSTM-Informer 网络参数

Table 6 LSTM-Informer network parameters

参数类型	参数值
隐含层神经元	150
堆叠层	3
遗忘门激活函数	sigmoid
输入门激活函数	tanh
输出门激活函数	tanh
编码器	3
解码器	3
模型维度	32
多头注意力个数	8
激活函数	Gelu
Batch-size	64
优化器	Adam 梯度下降算法
最大训练次数	100
初始学习率	0.001
早停比例	2
损失函数	MSEloss

表 7 出水 COD 预测模型的各项性能评估指标对比

Table 7 Comparison of different performance evaluation indicators for effluent COD prediction models

模型名称	MSE / $mg \cdot L^{-1}$	$RMSE$ / $mg \cdot L^{-1}$	MAE / $mg \cdot L^{-1}$	$MAPE$ /%	R^2
CNN	8.905	2.984	2.376	8.3	0.551
LSTM	3.368	1.835	1.334	4.7	0.830
CNN-LSTM	1.318	1.148	0.667	2.4	0.933
Informer	1.920	1.386	0.961	3.4	0.903
LSTM-Informer	0.631	0.795	0.602	2.2	0.968

出水 TN 的 5 种预测模型的各项性能评估指标如表 8 所列；实际值与预测值在测试集上的拟合结果如图 7 所示。

表 8 出水 TN 的预测模型的各项性能评估指标对比

Table 8 Comparison of different performance evaluation indicators for effluent TN prediction models

模型名称	MSE / $mg \cdot L^{-1}$	$RMSE$ / $mg \cdot L^{-1}$	MAE / $mg \cdot L^{-1}$	$MAPE$ /%	R^2
CNN	0.633	0.795	0.562	5.6	0.424
LSTM	0.348	0.590	0.453	4.4	0.683
CNN-LSTM	0.234	0.484	0.358	3.5	0.787
Informer	0.269	0.518	0.337	3.3	0.755
LSTM-Informer	0.105	0.324	0.235	2.2	0.904

由表 7 和表 8 可知，CNN 模型的预测精度较低且拟合效果不佳，这可能与其本身更适合处理二维图像有关。当将其调整为一维卷积以处理时间序列时，可能会导致训练时间变长、收敛速度变慢，以及最终收敛的损失值较高，从而降低预测精度。此外，CNN 的特点是侧重于局部相关特征的提取，而忽视了对全局依赖特征的捕捉。这一点也可以说明，在预测出水浓度时，长期依赖特征起到了至关重要的作用。

LSTM 模型预测一个时刻的出水水质浓度时能同时考虑该时刻前段时间内的输入特征向量、过程向量、标签向量和当前时刻的输入特征向量，符合实际处理污水过程中的时间贯通性，故在时间序列预测领域具有明显优势。图 6 和图 7 反映了出水 COD 和 TN 真实值和预测值的变化趋势和拟合程度，LSTM 模型的变化趋势在真实数据和预测数据上基本一致，且拟合程度较 CNN 模型高、预测稳定性稍强，但仍然存在梯度消失或梯度爆炸的问题导致总体预测效果不佳。

CNN-LSTM 模型，结合了 CNN 与 LSTM 模型二者的特点，首先将数据输入到 CNN 中进行卷积局部特征提取，对数据进行了降维，然后将降维后数据放入 LSTM 网络模型中进行训练学习，对经过卷积处

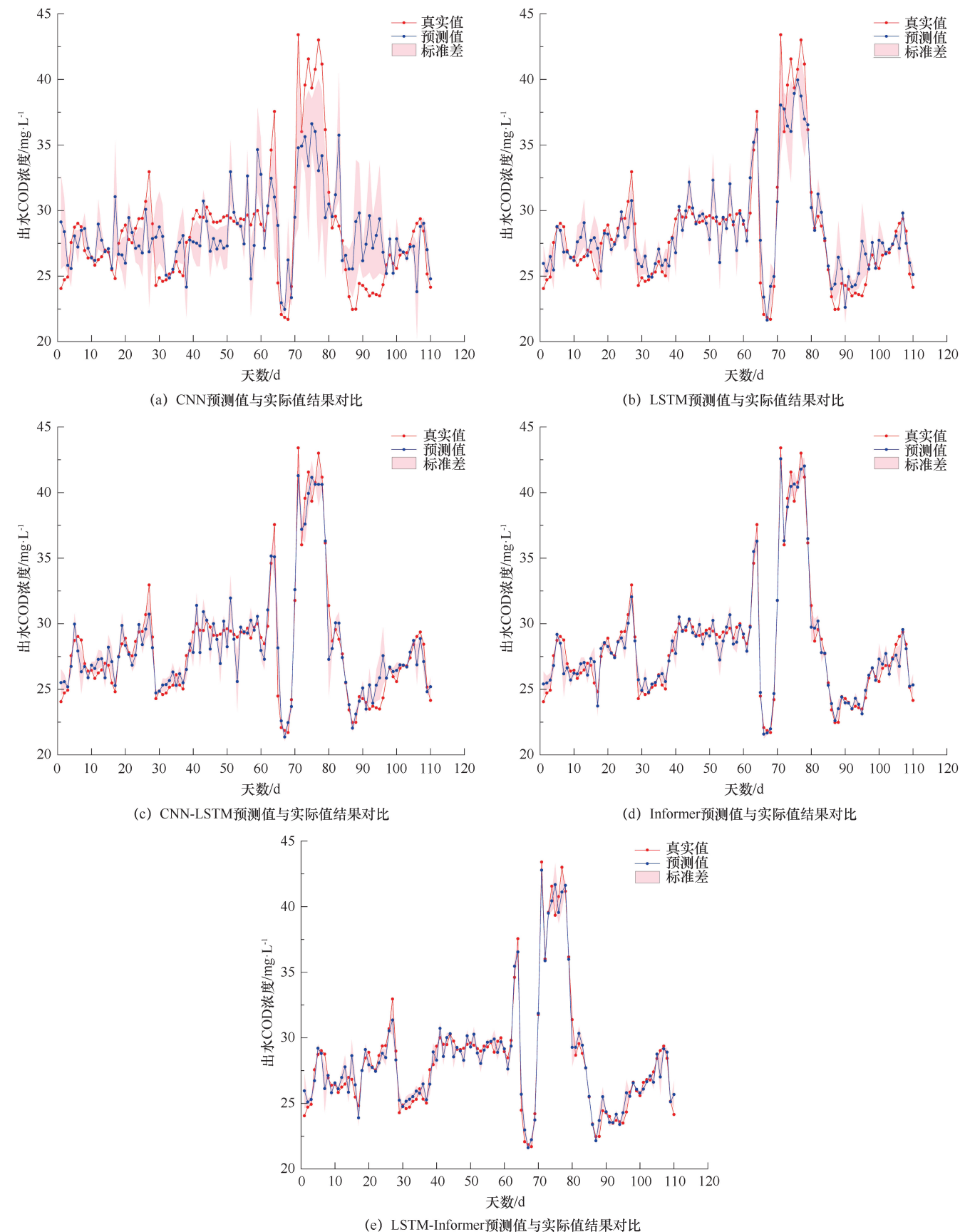
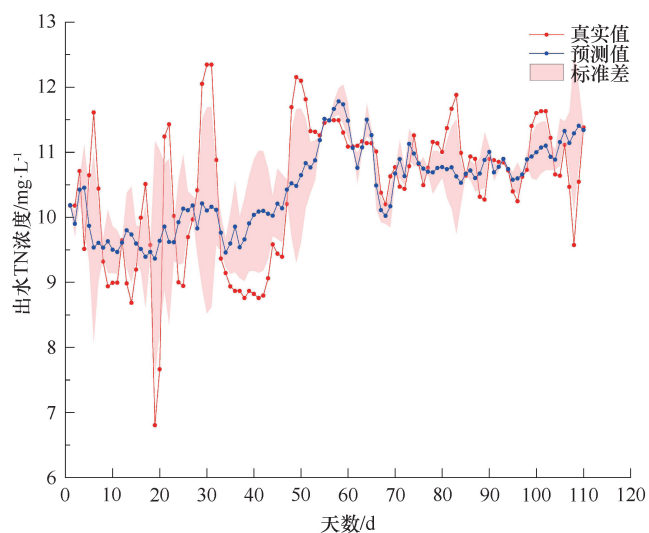
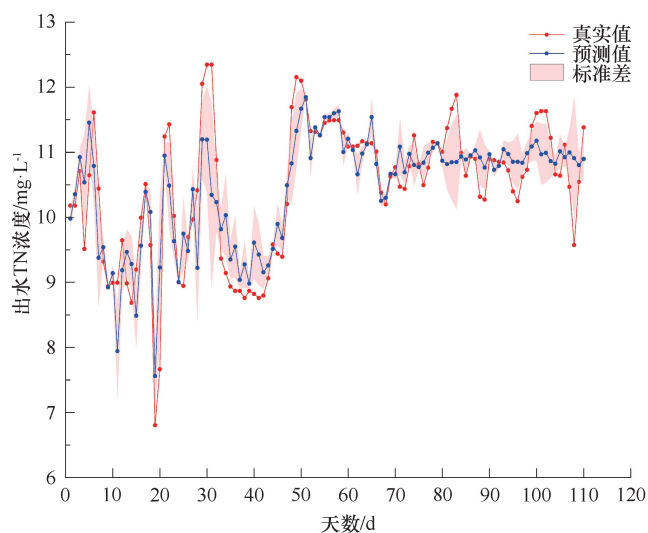


图6 各预测模型在2023年1—4月的出水COD预测结果

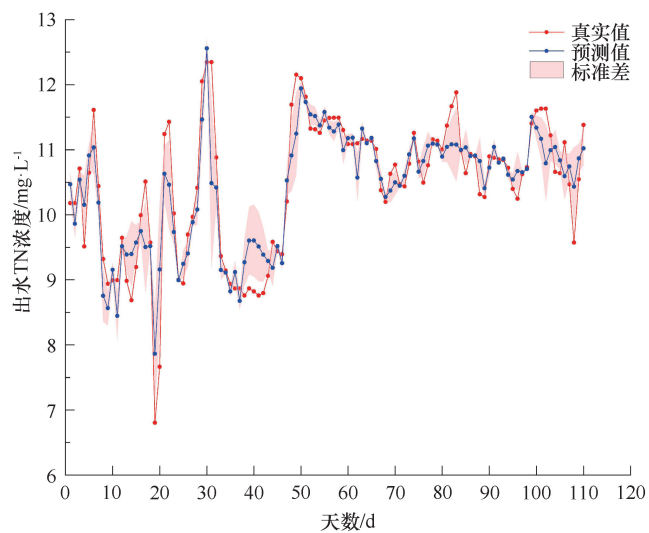
Fig. 6 Effluent COD prediction results of different models from January to April 2023



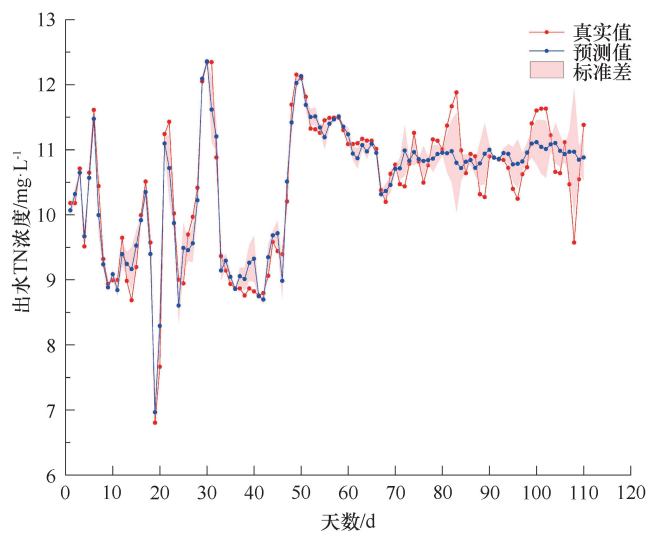
(a) CNN预测值与实际值结果对比



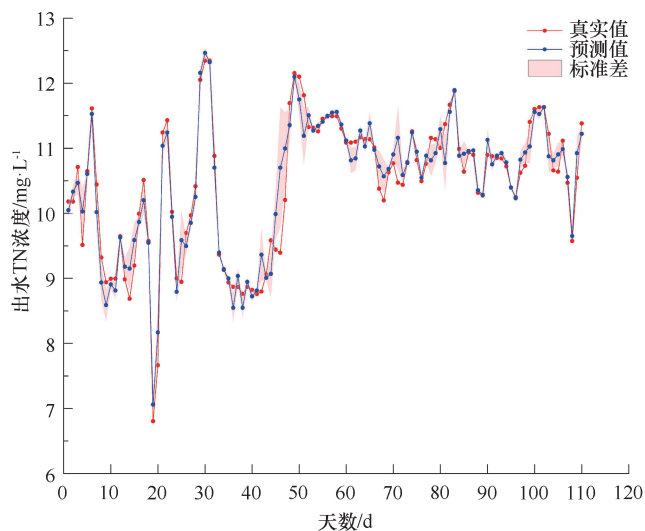
(b) LSTM预测值与实际值结果对比



(c) CNN-LSTM预测值与实际值结果对比



(d) Informer预测值与实际值结果对比



(e) LSTM-Informer预测值与实际值结果对比

图7 各预测模型在2023年1—4月的出水TN预测结果

Fig. 7 Effluent TN prediction results of different prediction models from January to April 2023

理后的序列特征进行时间维度上的依赖建模, 并输出预测结果。该模型评估结果显示, 较 CNN 与 LSTM 单个模型比较, 其各指标 (MSE 、 $RMSE$ 、 MAE 、 $MAPE$) 等均有所降低。通过结合 CNN 和 LSTM, 模型实现了空间与时间特征的高效融合, 这说明有效的针对单一模型的优点进行耦合, 建立的复合模型在一定程度上会提高模型的准确率。

Informer 模型凭借其自注意力机制, 将长序列数据中任意两点间的距离压缩至 1, 能够高效学习长时序依赖关系。同时, 多头概率稀疏自注意力机制可以从多个层面提取重要的特征向量, 如进水量、进水 COD 和进水 NH_3-N 、TN 等, 赋予这些关键特征更高的权重, 从而使模型更加关注核心信息, 减少次要信息的干扰, 避免关键信息的遗漏。在保障较高预测精度的同时, 该机制还显著降低了运算的时间和空间复杂度。此外, 自注意力蒸馏机制通过卷积和池化操作对每层编码器输入的时间序列进行降维, 进一步减少了维度和网络参数数量, 从而提升了预测效率。这些机制使得模型能够更有效地处理稀疏的时间序列数据, 从表来看, 预测结果具有良好的平稳性和较强的泛化能力, 能够更精准地挖掘出水水质变化趋势的特征。

LSTM-Informer 复合模型结合了 LSTM 和 Informer 的优势, 首先将水质时间序列数据输入 LSTM 模型以提取短期时间依赖于局部特征; 接着, 这些时序特征被输入到 Informer 模型通过 ProbSparse 机制捕捉长期依赖特征; 随后, 采用蒸馏操作来去除冗余的特征组合, 并通过稀疏自注意力机制进一步聚焦长时间序列特征; 最终, 经过全连接层输出预测结果。其三个评估指标 MSE 、 $RMSE$ 、 MAE 相比于 LSTM 和 Informer 单模型在出水 COD 浓度预测中分别降低了 81.3%、57.4%、54.9% 和 67.2%、23.2%、37.5%, 且在总氮的预测中也表现良好。由此可知该复合模型极大改善了预测精度。

3.4.3 模型结果显著性检验

为验证本试验所建立的 LSTM-Informer 模型的预测结果在统计学上具有优势, 采用结果差异显著性检验的方法。由于本试验采用的水质数据集不满足正态分布, 故采用 Wilcoxon 符号秩检验, 原假设为真实值与预测值之间的差值的中位数等于 0, 结果如表 9 和表 10 所列。

观察表中数据可知, 表 9 中各模型在显著性水平 $\alpha=0.05$ 下, p 值均大于 0.05, 说明各模型的预测结果与真实值的差异很小, 均具有一定的预测能力。而

表 9 各模型在出水 COD 中的结果差异显著性检验

Table 9 Significance test of differences in effluent COD results among models

模型名称	样本数 (N)	检验统计	标准化检验 统计(Z 值)	渐进显著 性(p 值)
CNN	110	3 454	1.197	0.231
LSTM	110	3 208	0.464	0.643
CNN-LSTM	110	2 989	-0.189	0.850
Informer	110	3 101	0.145	0.885
LSTM-Informer	110	3 096	0.130	0.897

表 10 各模型在出水 TN 中的结果差异显著性检验

Table 10 Significance test of differences in effluent TN results among models

模型名称	样本数 (N)	检验统计	标准化检验 统计(Z 值)	渐进显著 性(p 值)
CNN	110	2 793	-0.774	0.439
LSTM	110	3 233	0.538	0.590
CNN-LSTM	110	3 170	0.350	0.726
Informer	110	3 140	0.261	0.794
LSTM-Informer	110	3 047	-0.160	0.987

本试验建立的 LSTM-Informer 模型的 p 值为 0.897, 大于其余对比模型, 说明预测结果与真实值存在差异的可能性很小, 进而证明了本模型在统计学上具有优势。表 10 数据所体现的结果与表 9 的分析一致。综上, 有理由说明 LSTM-Informer 相较于其他模型的性能优势具有统计意义。

3.5 模型泛用性

通过对四种典型模型在测试集上的评估指标和结果差异显著性检验的对比, 该复合模型在出水水质预测性能方面确实优于其他模型。为进一步验证本试验建立的 LSTM-Informer 模型的泛用性, 使用该模型对 2022 年 1—4 月的水质数据集进行出水水质预测, 预测结果如图 8 所示。可以看到无论是在出水 COD 还是出水 TN 的预测效果上, 预测曲线与实际曲线都较为贴合, LSTM-Informer 模型可以有效地预测出水水质的变化规律, 善于捕捉长时序数据中的时间依赖关系, 对于水质观测任务具有良好的预报效果。

3.6 结果讨论

LSTM-Informer 模型的优势在于 LSTM 的短期时间记忆能力以及 Informer 引入的注意力机制, 提升了模型的可解释性。LSTM 层主要针对短期时间依赖和局部特征进行捕捉, 在本次试验中重点针对进水数据中如进水量波动较大时刻, 溶解氧缺乏时刻, 以及投加药剂以补充碳源、氮源时刻等, 通过记忆细胞获取这些时刻的水质变化信息, 以保证短期时间信息的

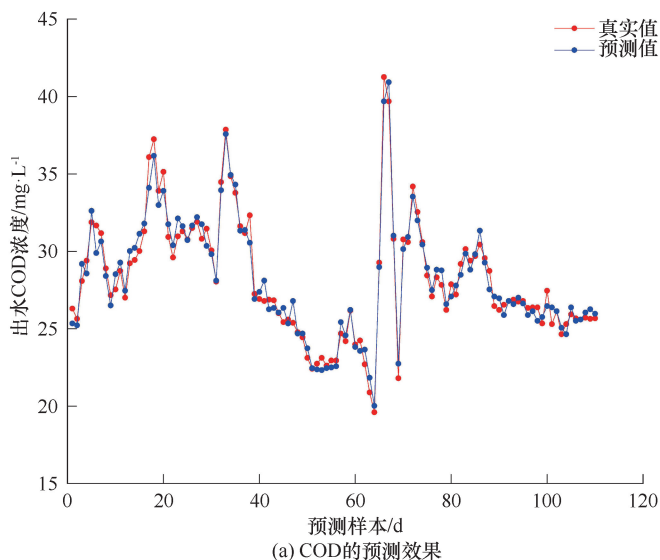
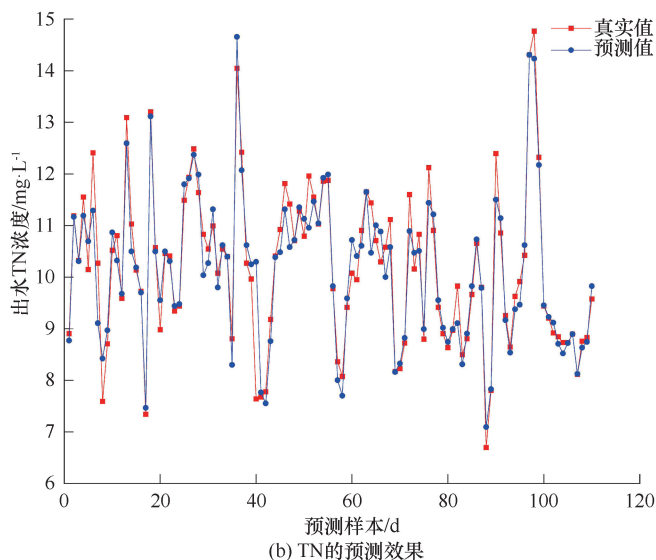


图8 2022年1—4月LSTM-Informer模型预测效果

Fig. 8 Prediction performance of LSTM-Informer model from January to April 2022

及时更新。Informer 模型中的注意力机制则需要关注长期水质的变化信息,重点关注如季节性变化,雨季前后,节假日周边,污染负荷剧烈变化等。通过对2022年和2023年全年预测结果的分析,注意力机制关注到11月以后至第二年3月为气温渐暖,冬去春来之际,微生物代谢活动增强,酶促反应迅速,碳源趋于充分分解,造成出水COD显著降低。而部分时间出水COD的异常升高,LSTM-Informer记录到了工业生产的高峰期,对应了进水COD及进水量的异常变化,说明了所建模型的适用性。与同类研究案例结果对比,如使用LSTM模型预测出水总氮^[33]、Informer模型预测出水总磷^[20]、BiLSTM预测出水COD和TN^[34],LSTM-Informer模型的预测准确性均有所提高。在出水COD的预测中,比BiLSTM模型的 R^2 值提高了1.2%;在出水TN预测中对比LSTM模型, $RMSE$ 降低了69%。通过上述分析,肯定了本试验所建模型的优势,但也存在一些问题亟待解决如下。

(1)本试验针对出水COD和TN进行了全年的数据训练与部分月份的数据预测,但所选水质特征指标数量不足,可能导致模型泛化性出现漏洞。

(2)本试验所采用数据序列长度为以天为单位的数据集,而所建模型为长时序预测,在数据序列长度上会由于时间较短使模型捕获特征不明显,数据预测说服力不足。

(3)本试验在统计学上使用了Wilcoxon符号秩检验的方法,一般来说,对于时间序列预测模型采用

Deibold-Mariano(DM)检验更准确。

(4)未来可建立一种利用废水表面图像和水质数据预测废水质量的受脑启发的多模式方法^[35],将图像分析与工艺参数结合的水质预测方法,从更多维度进行水质指标分析,从而构建更智能的自动化水质监测系统。

4 结 论

(1)在本试验中,建立了一种基于长短时记忆神经网络(LSTM)改进的Informer网络预测模型,其中利用LSTM的记忆细胞提取水质数据短期时间依赖和局部特征,并实现记忆信息的实时更新,再将水质时间序列数据转化为epoch, range, characteristic 矩阵输入到Informer网络中,多头稀疏自注意力关注长期的水质变化进行长时序预测,重点关注季节变化、污染负荷剧烈变化、工业生产大幅增减时的水质特征,结合水质短期剧烈变化和长期周期性变化,得到准确的出水指标预测。

(2)LSTM-Informer模型在出水COD浓度的预测评估指标 MSE 、 $RMSE$ 、 MAE 、 $MAPE$ 、 R^2 分别为 $0.630 \text{ mg} \cdot \text{L}^{-1}$ 、 $0.795 \text{ mg} \cdot \text{L}^{-1}$ 、 $0.602 \text{ mg} \cdot \text{L}^{-1}$ 、2.2%、0.968;在出水TN浓度的预测评估指标 MSE 、 $RMSE$ 、 MAE 、 $MAPE$ 、 R^2 分别为 $0.105 \text{ mg} \cdot \text{L}^{-1}$ 、 $0.324 \text{ mg} \cdot \text{L}^{-1}$ 、 $0.235 \text{ mg} \cdot \text{L}^{-1}$ 、2.3%、0.904。结果表明本文所建立的LSTM-Informer模型效果优于同类型的网络模型CNN、LSTM、CNN-LSTM、Informer,能更加准确地预测未来长期污水出水水质的变化趋势,为污水处理厂的自动化控制和管理提供了更有效

的技术支持。

(3)LSTM-Informer 模型在结果差异显著性上也具有优越性,在显著性水平为 0.05 的检验结果中,出水 COD 浓度预测显著性检验 p 值为 0.897,大于其他对比模型的 p 值;出水 TN 浓度预测显著性检验 p 值为 0.987,大于其他对比模型的 p 值,说明 LSTM-Informer 的预测值与真实值存在差异的可能性更小,预测结果更接近真实值。

(4)本研究中构建的预测模型采用多指标输入预测单一指标的模式,并基于 12 个月的样本预测 4 个月的出水 COD 和 TN 数据,尽可能涵盖所有影响污水处理厂工艺的关键特征指标,使模型能够适应各种复杂环境。同时,本试验将污水处理厂的水质数据与各构筑物的运行状态相结合,通过模型预测结果定性分析出水 COD 和 TN 浓度变化的原因,并进一步定位需调控的进水、加药指标。这为污水处理厂的出水指标控制、超标预警、提升质量与效率以及节能降耗提供了理论依据。

参考文献 (References) :

[1] 张宝军,王国平,袁永军,等. 水处理工程技术[M]. 重庆:重庆大学出版社,2015.
ZHANG B J, WANG G P, YUAN Y J, et al. Water Treatment Engineering Technology [M]. Chongqing: Chongqing University Press, 2015.

[2] 张志文. 污水处理过程中软测量技术的研究及应用[J]. 城市建设理论研究(电子版), 2018(16): 70.
ZHANG Z W. Research and application of soft measurement technology in sewage treatment[J]. Theoretical Research in Urban Construction, 2018(16): 70.

[3] 黄道平,刘乙奇,李艳. 软测量在污水处理过程中的研究与应用[J]. 化工学报, 2011, 62(1): 1-9.
HUANG D P, LIU Y Q, LI Y. Soft sensor research and its application in wastewater treatment[J]. CIESC Journal, 2011, 62(1): 1-9.

[4] ÖMER FARUK D. A hybrid neural network and ARIMA model for water quality time series prediction[J]. Engineering Applications of Artificial Intelligence, 2010, 23(4): 586-594.

[5] GORGAN-MOHAMMADI F, RAJAEI T, ZOUNEMAT-KERMANI M. Decision tree models in predicting water quality parameters of dissolved oxygen and phosphorus in lake water[J]. Sustainable Water Resources Management, 2022, 9(1): 1.

[6] SUN J J, GUAN X, SUN X J, et al. Water quality prediction and carbon reduction mechanisms in wastewater treatment in Northwest cities using Random Forest Regression model[J]. Scientific Reports, 2024, 14: 31525.

[7] 杨璐霞,王智瑜,沈帅杰,等. 基于 PSO-LSTM-SATN 模型的污水水质预测研究[J]. 工业水处理, 2025, 45(6): 159-166.

YANG L X, WANG Z Y, SHEN S J, et al. Research on sewage water quality prediction based on the PSO-LSTM-SATN model[J]. Industrial Water Treatment, 2025, 45(6): 159-166.

[8] 许睿亭,姜翠玲,孙磊,等. 基于深度学习的环太湖河流气象水质协同预测模型研究[J]. 水利水电技术(中英文), 2025, 56(7): 228-238.
XU R T, JIANG C L, SUN L, et al. Deep learning-based water quality prediction model combined with meteorological data for rivers around Taihu Lake [J]. Water Resources and Hydropower Engineering, 2025, 56(7): 228-238.

[9] ZHANG Y T, WANG J H, LI C L, et al. Attention-based deep learning models for predicting anomalous shock of wastewater treatment plants[J]. Water Research, 2025, 275: 123192.

[10] CAO D D, CHAN M, NG S. Modeling and forecasting of nanoFeCu treated sewage quality using recurrent neural network (RNN) [J]. Computation, 2023, 11(2): 39.

[11] ZHANG Y T, LI C L, DUAN H P, et al. Deep learning based data-driven model for detecting time-delay water quality indicators of wastewater treatment plant influent [J]. Chemical Engineering Journal, 2023, 467: 143483.

[12] 马笑文. 基于神经网络的污水处理厂进水负荷与出水水质预测建模研究[D]. 南京:南京大学, 2020.
MA X W. Predictive Modeling of Influent Load and Effluent Water Quality of Wastewater Treatment Plants Based on Neural Network [D]. Nanjing: Nanjing University, 2020.

[13] AN T, FENG K L, CHENG P J, et al. Adaptive prediction for effluent quality of wastewater treatment plant: Improvement with a dual-stage attention-based LSTM network [J]. Journal of Environmental Management, 2024, 359: 120887.

[14] 陈湛峰,李晓芳. 基于注意力机制优化的 BiLSTM 珠江口水质预测模型[J]. 环境科学, 2024, 45(6): 3205-3213.
CHEN Z F, LI X F. Water quality prediction model for the Pearl River Estuary based on BiLSTM improved with attention mechanism [J]. Environmental Science, 2024, 45(6): 3205-3213.

[15] 张铭玮,李正权,方志豪. 基于量子粒子群优化的 CNN-LSTM 水质预测模型[J]. 中国计量大学学报, 2022(3): 303-309.
ZHANG M W, LI Z Q, FANG Z H. A CNN-LSTM water quality prediction model based on quantum particle swarm optimization[J]. Journal of China University of Metrology, 2022(3): 303-309.

[16] ZHOU H Y, ZHANG S H, PENG J Q, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(12): 11106-11115.

[17] 何志铎,熊祖强. 基于 Informer 神经网络的工作面矿压预测研究[J]. 矿业研究与开发, 2024, 44(7): 142-148.
HE Z H, XIONG Z Q. Research on Ming pressure prediction of working face based on informer neural network[J]. Mining Research and Development, 2024, 44(7): 142-148.

[18] CAO Y S, LIU G, LUO D H, et al. Multi-timescale photovoltaic power forecasting using an improved Stacking ensemble algorithm based LSTM-Informer model[J]. Energy, 2023, 283: 128669.

[19] WANG K, ZHANG J L, LI X W, et al. Long-term power load

- forecasting using LSTM-informer with ensemble learning [J]. Electronics, 2023, 12(10): 2175.
- [20] 安显宁,朱四富,刘静,等. 基于深度学习的污水处理厂出水总磷预测方法[J]. 工业水处理, 2024, 44(10): 143-150.
- AN Y N, ZHU S F, LIU J, et al. Prediction of total phosphorus in wastewater treatment plant effluent based on deep learning [J]. Industrial Water Treatment, 2024, 44(10): 143-150.
- [21] 王朱贺,李楠,张希瑞,等. 基于机器学习的深基坑三维土层重建[J]. 重庆大学学报, 2021, 44(5): 135-145.
- WANG Z H, LI N, ZHANG X R, et al. 3D soil layer reconstruction of deep foundation pit based on machine learning [J]. Journal of Chongqing University, 2021, 44(5): 135-145.
- [22] LIU H B, JIAN H R. ANN-based prediction of ammonia nitrogen for wastewater discharge indicators under carbon neutral trend [J]. Frontiers in Ecology and Evolution, 2023, 11: 1199870.
- [23] BI J, LIN Y Z, DONG Q X, et al. Large-scale water quality prediction with integrated deep neural network [J]. Information Sciences, 2021, 571: 191-205.
- [24] BERTONE E, STEWART R A, ZHANG H, et al. Hybrid water treatment cost prediction model for raw water intake optimization[J]. Environmental Modelling & Software, 2016, 75: 230-242.
- [25] HAO X D, LIU R B, HUANG X. Evaluation of the potential for operating carbon neutral WWTPs in China [J]. Water Research, 2015, 87: 424-431.
- [26] 姚怡帆,荆玉姝,王丽艳,等. 基于集成模型的污水处理厂出水总氮预测方法[J]. 工业水处理, 2023, 43(9): 187-194.
- YAO Y F, JING Y S, WANG L Y, et al. Prediction method of total nitrogen in wastewater treatment plant based on integrated model[J]. Industrial Water Treatment, 2023, 43(9): 187-194.
- [27] VERMA A, WEI X P, KUSIAK A. Predicting the total suspended solids in wastewater: A data-mining approach [J]. Engineering Applications of Artificial Intelligence, 2013, 26(4): 1366-1372.
- [28] WEI X P, KUSIAK A, SADAT H R. Prediction of influent flow rate: Data-mining approach[J]. Journal of Energy Engineering, 2013, 139(2): 118-123.
- [29] 张涛,陈猷鹏. 人工智能在工业园区废水处理中的应用研究现状及展望[J]. 土木与环境工程学报(中英文), 2024, 46(6): 192-202.
- ZHANG T, CHEN Y P. Research status and prospect of the application of artificial intelligence in industrial park wastewater treatment [J]. Journal of Civil and Environmental Engineering, 2024, 46(6): 192-202.
- [30] PARK C H, CHANG J H. Robust shrinkage range estimation algorithms based on hampel and skipped filters [J]. Wireless Communications and Mobile Computing, 2019, 2019(1): 6592406.
- [31] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.
- [32] 史红伟,陈祺,王云龙,等. 基于 encoder-decoder 框架的城镇污水厂出水水质预测[J]. 中国农村水利水电, 2023(11): 93-99.
- SHI H W, CHEN Q, WANG Y L, et al. Water quality prediction of town sewage plant effluent based on encoder-decoder framework [J]. China Rural Water and Hydropower, 2023(11): 93-99.
- [33] 余铭铨,师浩铭. 基于 LSTM 模型的污水处理厂出水总氮预测研究[J]. 山东科学, 2024, 37(6): 116-124.
- YU M Q, SHI H M. Prediction of effluent total nitrogen in wastewater treatment using LSTM neural network [J]. Shandong Science, 2024, 37(6): 116-124.
- [34] 陈树龙,黎志伟,黄祖安,等. 基于 BiLSTM 深度学习模型的污水处理厂水质预测[J]. 当代化工研究, 2024(5): 96-98.
- CHEN S L, LI Z W, HUANG Z A, et al. Water quality prediction of sewage treatment plant based on BiLSTM deep learning model [J]. Modern Chemical Research, 2024(5): 96-98.
- [35] LI J C, LIN S J, ZHANG L, et al. Brain-inspired multimodal approach for effluent quality prediction using wastewater surface images and water quality data [J]. Frontiers of Environmental Science & Engineering, 2023, 18(3): 31.

(责任编辑 王璐)