

基于 IGJO-TCN-BiGRU-SA 的水质预测模型^{*}

陈 静¹ 李政委¹ 杨 凯¹ 凌冰心¹ 李政权² 祝家福³

(1. 安徽理工大学电气与信息工程学院 淮南 232001; 2. 空军工程大学航空机务士官学校 信阳 464000;

3. 安徽理工大学材料科学与工程学院 淮南 232001)

摘 要: 针对水质数据的时序性、强非线性及维度高导致模型预测精度低的问题, 提出一种基于 IGJO-TCN-BiGRU-SA 的水质预测模型。首先, 通过三次样条插值 (cubic spline interpolation, CSI) 算法修复缺失数据, 采用箱线图 (Boxplot) 和快速傅里叶变换 (fast Fourier transformation, FFT) 算法剔除异常数据, 提升数据的完备性, 利用灰色关联分析 (grey relational analysis, GRA) 算法分析水质指标相关性, 得到强相关性的水质指标; 其次, 以双向门控循环单元 (bidirectional gated recurrent unit, BiGRU) 为基准模型, 提取水质数据中的潜在变化特征, 采用时域卷积网络 (temporal convolutional network, TCN) 对数据进行局部特征提取, 增强了模型对不同时间尺度特征的提取能力, 并通过自注意力机制 (self-attention, SA) 自适应更新 BiGRU 模型特征权重, 有效提取全局上下文信息的关键特征; 最后采用改进的金豺优化算法 (improved golden jackal optimization, IGJO) 寻优 TCN 的 L2 参数、BiGRU 的学习率和神经元个数、SA 机制的键值, 从而提升水质预测精度。以淮河流域水质历史数据为样本进行实验, 相对于 LSTM、GRU、BiGRU 和 TCN-BiGRU-SA 等经典算法, 所提出的 IGJO-TCN-BiGRU-SA 预测算法在 MAE 上降低 19.64%~68.29%, RMSE 上降低 25.47%~75.19%, MAPE 上降低 28.71%~73.36%, R^2 上提升 0.19%~3.61%。实验结果证明, 基于 IGJO-TCN-BiGRU-SA 的预测算法明显提高了水质预测精度。

关键词: 水质预测; 双向门控循环单元; 时域卷积网络; 自注意力机制; 改进的金豺优化算法; 三次样条插值

中图分类号: TP18; TN919; X832 **文献标识码:** A **国家标准学科分类代码:** 413.1040

Water quality prediction model based on IGJO-TCN-BiGRU-SA

Chen Jing¹ Li Zhengwei¹ Yang Kai¹ Ling Bingxin¹ Li Zhengquan² Zhu Jiafu³

(1. School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, China;

2. Aviation Maintenance NCO School, Air Force Engineering University, Xinyang 464000, China;

3. School of Materials Science and Engineering, Anhui University of Science and Technology, Huainan 232001, China)

Abstract: Aiming at the problems of low model prediction accuracy caused by the temporality, strong nonlinearity, and high dimensions of water quality data, a water quality prediction model based on the IGJO-TCN-BiGRU-SA has been proposed. Firstly, missing data are repaired through the CSI algorithm, while abnormal data are removed through Boxplot and FFT algorithms to enhance data completeness. The GRA algorithm is used to analyze the correlation of water quality indicators and obtain strongly correlated water quality indicators. Secondly, the BiGRU model is employed as the baseline model to extract latent change features from the data. The TCN algorithm is employed for local feature extraction from the data, enhancing the model's capability to extract features at different time scales. The SA mechanism is used to adaptively update the feature weights of the BiGRU model, effectively extracting key features from global contextual information. Finally, the IGJO algorithm is employed to optimize the L2 of TCN, the learning rate and neuron number of BiGRU, and the key of the SA mechanism, thereby improving the accuracy of water quality prediction. The historical water quality data from the Huai River Basin were used as samples for experiments. Compared to classic algorithms such as LSTM, GRU, BiGRU and TCN-BiGRU-SA,

收稿日期: 2025-12-03 Received Date: 2025-12-03

^{*} 基金项目: 国家自然科学基金项目 (51874010)、安徽省教育厅高校自然科学研究项目 (KJ2018A0087)、安徽理工大学 2025 年大学生创业基金扶持项目、安徽理工大学创新基金项目 (2025cx2083) 资助

the IGJO-TCN-BiGRU-SA prediction algorithm proposed in this research achieved reductions of 19.64% to 68.29% in MAE, 25.47% to 75.19% in RMSE, 28.71% to 73.36% in MAPE, and 0.19% to 3.61% in R^2 . Experimental results indicate that the water quality prediction algorithm based on IGJO-TCN-BiGRU-SA significantly enhances the model's predictive accuracy.

Keywords: water quality prediction; bidirectional gated recurrent unit; temporal convolutional network; self-attention mechanism; improved golden jackal optimization; cubic spline interpolation

0 引 言

水是中国经济社会发展的重要基础,水资源的保护已被纳入国家战略规划体系^[1]。随着全球经济的快速发展和人民对美好生活的需求日益增长,水质要求越来越高^[2]。良好的水质是工业用水的重要支撑,对于居民健康和工业发展尤其重要^[3-4]。2015 年国务院实施《水污染防治行动计划》以来^[5],我国的水质管理标准越来越严格,水环境保护力度持续加大。然而,由于人类活动和城市化推进,生活污水和工业废料等大量污染物被排入河流^[6],严重影响了水质安全。水质预警是水资源保护的重要措施,包括水质监测和水质预测两个环节^[7]。水质监测是水质预测的前提,能为水质预测提供必要的支撑和帮助^[8]。水质预测通过分析水质环境的变化,得出科学合理的预测结果,有利于保护水质安全、优化水资源调配和减少污染对水环境的危害^[9]。

近年来,国内外很多学者对水质预测算法展开了广泛研究,包括回归分析算法^[10]、时间序列算法^[11]、机器学习算法^[12]和人工神经网络(artificial neural network, ANN)^[13],其中 ANN 凭借其独特的拓扑结构和自适应学习能力,在水质预测领域被广泛应用^[14]。Zhang 等^[15]基于循环神经网络(recurrent neural network, RNN)预测水质,但存在梯度消失或梯度爆炸的问题。肖扬岚等^[16]采用长短期记忆(long short-term memory, LSTM)神经网络进行水质预测,解决了梯度消失或爆炸问题,但模型结构复杂,计算成本高。黎煜昭等^[17]利用门控循环单元(gated recurrent unit, GRU)进行水质预测,通过将输入门和遗忘门合并为更新门,同时引入重置门,减少了内部隐藏状态和相关门结构。然而,GRU 依靠前一时刻数据预测下一时刻数据,忽略了数据中潜在的未来变化信息,无法应对水环境发生突变的情况。

工程上,水质预测模型需要处理高维复杂的水质数据,人工调参效率低且误差较大,大多数学者常采用元启发式算法寻找最优解。Yang 等^[18]采用遗传算法(genetic algorithm, GA)模拟生物进化过程,以寻找支持向量机(support vector machine, SVM)参数最优解,但算法参数设置复杂,寻优精度低。杨潞霞等^[19]使用粒子群优化(particle swarm optimization, PSO)算法优化 LSTM-SA 模型,减少了算法参数设置,但寻优精度仍较

低。牛景辉^[20]利用灰狼优化(grey wolf optimizer, GWO)算法寻优极端决策提升树(extreme gradient boosting, XGBoost)算法的超参数,一定程度上提高了算法寻优精度,但还有待于进一步提升。

在工程实践中,水质数据采集的过程存在水质监测设备故障或网络异常等因素干扰,导致采集到的水质数据会存在缺失、异常和高频噪声等缺点^[21]。张梦迪等^[22]采用平均值修复缺失数据,并使用拉依达准则剔除数据异常值,但未考虑原始数据的非线性特征。秦文虎等^[23]采用数据加权算法和三次样条插值修复缺失数据,但未考虑不同水质指标之间的关联性。

针对上述问题,本文提出一种基于 IGJO-TCN-BiGRU-SA 算法的水质预测模型。首先,采用三次样条插值(cubic spline interpolation, CSI)算法修复缺失数据,利用箱线图(boxplot)和快速傅里叶变化(fast Fourier transformation, FFT)算法剔除异常数据,并使用灰色关联分析(grey relational analysis, GRA)算法分析水质指标间的相关性。其次,采用双向门控循环单元(bidirectional gated recurrent unit, BiGRU)作为基准模型,以挖掘数据中潜在的未来变化信息;使用时域卷积网络(temporal convolutional network, TCN)提取数据的局部特征,以避免 BiGRU 模型过拟合;考虑到 BiGRU 在训练过程中由于梯度更新衰减较大导致的梯度消失问题,引入自注意力机制(self-attention, SA)更新模型权重。最后,通过改进的金豺优化算法(improved golden jackal optimization, IGJO)寻优 TCN 的 L2, BiGRU 的学习率和神经元个数,SA 机制的键值,以避免人工调参的误差,进一步提高预测精度。

本文提出基于 TCN-BiGRU-SA 的水质预测模型。其中, BiGRU 模型能同时融合前向和后向 GRU 模块功能,捕获水质数据中潜在的未来变化特征; TCN 算法可增强 BiGRU 模型对数据不同时间尺度特征的提取能力,解决 BiGRU 模型可能出现过拟合问题; SA 机制能根据 BiGRU 模型输出特征的重要性自适应更新权重,降低 BiGRU 模型出现梯度消失的风险,提升水质预测精度。采用 IGJO 算法寻优 TCN-BiGRU-SA 模型参数。模型中, TCN 的 L2, BiGRU 的学习率和神经元个数, SA 的键值等是影响模型预测精度的关键超参数。金豺优化算法(golden jackal optimization, GJO)能一定程度上提升模型参数的寻优能力,但存在迭代后期易陷入局部最优解和收敛速

度慢的问题,IGJO 算法通过在 GJO 算法中引入佳点集初始化、动态权重更新策略和柯西变异策略,进一步提升了预测模型参数的全局寻优能力和收敛速度。

1 研究方法

1.1 数据预处理算法

1) 基于 CSI 的缺失数据处理算法

水质监测设备故障或网络异常会导致数据缺失,本文采用 CSI 算法修复缺失数据。相对于均值插值、线性插值等经典插值算法,CSI 算法能够拟合水质数据的复杂非线性变化趋势,其工作原理是在数据缺失区域构造一系列相邻的 3 次多项式函数以生成平滑曲线^[24]。三次样条插值表达式如式(1)所示。

$$y_i = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3 \quad (1)$$

$i = 0, 1, \dots, n$

式中: y_i 表示水质数据; a_i 、 b_i 、 c_i 和 d_i 分别表示 3 次多项式方程的常数项、1 次项、2 次项和 3 次项的系数; n 为三次多项式个数; i 为三次多项式顺序。

2) 基于 Boxplot 与 FFT 的异常数据处理算法

受汛期及数据传输过程存在信号干扰等原因,采集到的水质数据存在异常、高频噪声等问题。相对于散点图、直方图等异常值分析算法,Boxplot 算法可以指示数据分布状态。该方法通过计算数据的四分位数和四分位距,将异常值定义为超出 $[Q_1 - 1.5IQR, Q_3 + 1.5IQR]$ 范围的数据点,其下界和上界的表达式如式(2)~(4)所示。

$$IQR = Q_3 - Q_1 \quad (2)$$

$$L = Q_1 - 1.5IQR \quad (3)$$

$$U = Q_3 + 1.5IQR \quad (4)$$

式中: Q_1 和 Q_3 分别表示第 1 个四分位数和第 3 个四分位数; IQR 表示四分位距; L 表示下界; U 表示上界。

数据中的高频噪声仍会影响模型对水质变化规律的学习。相对于中值滤波等经典算法,FFT 算法能够精准的去掉数据中高频噪声,同时保留数据的非噪声频段特征。FFT 算法的原理是利用离散傅里叶变换 (discrete Fourier transform, DFT) 的周期性与对称性,将原始时间序列数据分解为多个短序列傅里叶变换之和,以实现更高效的计算^[25]。DFT 表达式如式(5)所示。

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{n \cdot k} \quad n = 0, 1, \dots, N-1 \quad (5)$$

式中: $X(k)$ 表示 $x(n)$ 在第 k 个离散频率点的幅值; k 为频率索引; $W_N = e^{-j\frac{2\pi}{N}}$ 是旋转因子。

为减少 DFT 计算量,FFT 将序列 $x(n)$ 依据奇偶划分为 $u(m)$ (奇序列) 和 $p(m)$ (偶序列),其表达式如式(6)所示。

$$\begin{cases} u(m) = x(2m + 1) \\ p(m) = x(2m) \end{cases} \quad (6)$$

式中: $m = 0, 1, 2, \dots, \frac{N}{2} - 1$ 。

$u(m)$ 和 $p(m)$ 的 DFT 表达式如式(7)所示。

$$\begin{cases} U(k) = \sum_{m=0}^{\frac{N}{2}-1} u(m) (W_{N/2})^{m \cdot k} \\ P(k) = \sum_{m=0}^{\frac{N}{2}-1} p(m) (W_{N/2})^{m \cdot k} \end{cases} \quad (7)$$

式中: $U(k)$ 为奇序列的 DFT; $P(k)$ 为偶序列的 DFT。

最终,FFT 的表达式如式(8)所示。

$$X(k) = U(k) W_N^k + P(k) \quad (8)$$

3) 基于 GRA 的相关性分析

水质数据中存在的冗余特征会导致其维度过高,降低模型计算效率,本文引入 GRA 算法计算水质指标间的相关性。相对于皮尔逊 (Person) 等传统相关性分析算法,GRA 算法能挖掘不同水质指标间的非线性关系。GRA 算法的工作流程如下:

(1) 定义参考序列和比较序列。在本研究中,可选择水质指标任意一项作为参考序列,其余水质指标则作为比较序列,参考序列和比较序列的表达式如式(9)和(10)所示。

$$Y = \{Y(k) \mid k = 1, 2, \dots, n\} \quad (9)$$

$$X_i = \{X_i(k) \mid k = 1, 2, \dots, n\} \quad i = 1, 2, \dots, m \quad (10)$$

式中: Y 为参考序列; X_i 为比较序列; $Y(k)$ 表示参考序列中第 k 个参数; $X_i(k)$ 表示比较序列中第 i 组序列的第 k 个参数。

(2) 归一化处理。水质指标间的属性存在差异,需要归一化处理,处理后的参考序列和比较序列如式(11)和(12)所示。

$$y_k = \left\{ \frac{Y(k)}{\bar{Y}}, k = 1, 2, \dots, n \right\} \quad (11)$$

$$x_i(k) = \left\{ \frac{X_i(k)}{\bar{X}_i}, k = 1, 2, \dots, n \right\} \quad i = 1, 2, \dots, m \quad (12)$$

式中: y_k 是参考序列的归一化值; $x_i(k)$ 是比较序列的归一化值; $\bar{Y}$ 和 $\bar{X}_i$ 分别表示 Y 和 X_i 的平均值。

(3) 灰色关联度计算。水质指标间灰色关联度 r_i 的计算公式如式(13)所示。

$$r_i = \frac{1}{n} \sum_{k=1}^n \xi_i(k) \quad (13)$$

式中: $\xi_i(k)$ 为灰色关联系数,其计算公式如式(14)所示。

$$\xi_i(k) = \frac{\min_k \Delta_i(k) + \max_k \Delta_i(k)}{\Delta_i(k) + \rho \max_k \Delta_i(k)} \quad (14)$$

式中: ρ 为分辨系数, $\rho \in (0, 1)$; $\max_i \max_k \Delta_i(k)$ 和 $\min_i \min_k \Delta_i(k)$ 分别表示绝对差值序列中的最大值和最小值; $\Delta_i(k)$ 为第 i 组序列的第 k 个绝对差值。

$\Delta_i(k)$ 计算公式如式 (15) 所示。

$$\Delta_i(k) = |y_k - x_i(k)| \quad (15)$$

根据需求, 保留相关性较高的水质指标, 以降低模型计算复杂度。

1.2 基于 BiGRU 的水质预测算法

鉴于水质环境发生突变时会影响水质预测精度的问题, 本文采用 BiGRU 算法预测水质数据。GRU 作为 LSTM 的改进模型, 将 LSTM 的输入门和遗忘门合并为更新门, 同时结合重置门和候选隐藏向量以生成下一时刻的状态向量, 降低了 LSTM 模型复杂度。相较于传统的 GRU 算法, BiGRU 通过同时整合前向和后向 GRU 模块, 能够捕获数据中潜在的未来变化特征。BiGRU 的表达式如式 (16) ~ (18) 所示, 其基本单元结构如图 1 所示。

$$\vec{h}_t = \text{GRU}(x_t, \vec{h}_{t-1}) \quad (16)$$

$$\overleftarrow{h}_t = \text{GRU}(x_t, \overleftarrow{h}_{t+1}) \quad (17)$$

$$h_t = \text{Concat}(\vec{h}_t, \overleftarrow{h}_t) \quad (18)$$

式中: $\vec{h}_t$ 和 $\overleftarrow{h}_t$ 分别代表 BiGRU 前向和后向的输出; x_t 表示当前时刻的输入; h_t 为整个 BiGRU 模型的输出; Concat 表示拼接操作。

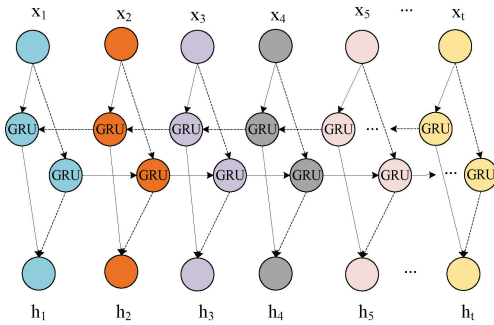


图 1 BiGRU 的基本单元结构

Fig. 1 Basic unit structure of BiGRU

1.3 基于 TCN 的特征提取算法

在大规模数据训练过程中, BiGRU 模型的递归结构导致模型复杂度增加, 容易出现过拟合问题, 本文引入 TCN 算法提取水质数据的局部特征。相对于卷积神经网络 (convolutional neural network, CNN), TCN 解决了时间建模长度受卷积核尺寸限制的难题。TCN 算法通过扩张卷积结构能够以指数增长的感受野处理水质数据, 并且其通过残差连接在训练过程中能够有效的解决 BiGRU 的递归结构导致的过拟合问题。TCN 算法的表达式如式 (19) 所示, 其模型结构如图 2 所示。

$$y_i = \sum_{j=1}^k w_j \times x_{i+j-1} \quad (19)$$

式中: y_i 表示扩张卷积运算的输出; k 为卷积核的大小; w_j 为卷积核的权重; x_{i+j-1} 为时间步 $x+j-1$ 的输入信号。

1.4 基于 SA 的权重自适应更新机制

BiGRU 模型在训练过程中的梯度更新衰减较大, 导致先输入特征被后续输入特征稀释, 序列前端的特征无法充分利用, 引发梯度消失问题, 本文引入 SA 机制更新模型权重。与传统注意力机制相比, SA 机制专注于自身进行相关信息的提取, 抑制无用特征的影响, 使模型有选择性地专注于重要信息而忽略无关的信息。SA 机制的表达式如式 (20) ~ (23) 所示, 其模型结构如图 3 所示。

$$Q = W^Q A \quad (20)$$

$$K = W^K A \quad (21)$$

$$V = W^V A \quad (22)$$

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (23)$$

式中: Q 、 K 和 V 分别表示查询、键和数值矩阵; A 表示输入矩阵; W^Q 、 W^K 和 W^V 分别表示 Q 、 K 和 V 对应的权重矩阵; $\sqrt{d_k}$ 表示 Q 、 K 和 V 的维度。

1.5 基于 IGJO 算法的模型超参数优化

工程上, TCN 的 L2, BiGRU 的学习率和神经元个数, SA 的键值需人工整定, 但存在误差大和效率低的问题, 本文提出 IGJO 算法用于寻优上述参数。

1) GJO 优化算法

GJO 算法是由 Chopra 等^[26]于 2022 年提出的一种基于种群的元启发式优化算法, 通过金豺合作狩猎行为模拟了上述待优化参数的选取最优解过程。GJO 算法的实现包括初始化种群、搜索猎物 and 包围攻击猎物 3 个阶段。

(1) 初始化种群

GJO 算法在预先设置的参数取值范围内随机生成多组初始解 (种群大小决定初始解的组数, 参数个数决定搜索空间维度), 由此形成初始猎物矩阵。在优化过程中, 本研究选择 RMSE 作为适应度函数, 并计算每个猎物的适应度, 选择最优适应度的个体为雄性金豺, 次优适应度的个体为雌性金豺。

(2) 搜索猎物 (探索阶段)

在参数空间中, 雄、雌金豺对猎物 (参数) 进行全局探索, 由雄性金豺主导, 雌性金豺跟随进行狩猎。

(3) 包围并攻击猎物 (开发阶段)

当猎物 (参数) 遭受金豺骚扰而疲惫时, 其逃脱能量会减少, 雄、雌金豺会包围猎物并发起进攻。此阶段用于对参数进行局部开发。随着算法迭代的进行, 金豺会不断靠近全局最优解, 最终输出上述参数的最优组合。

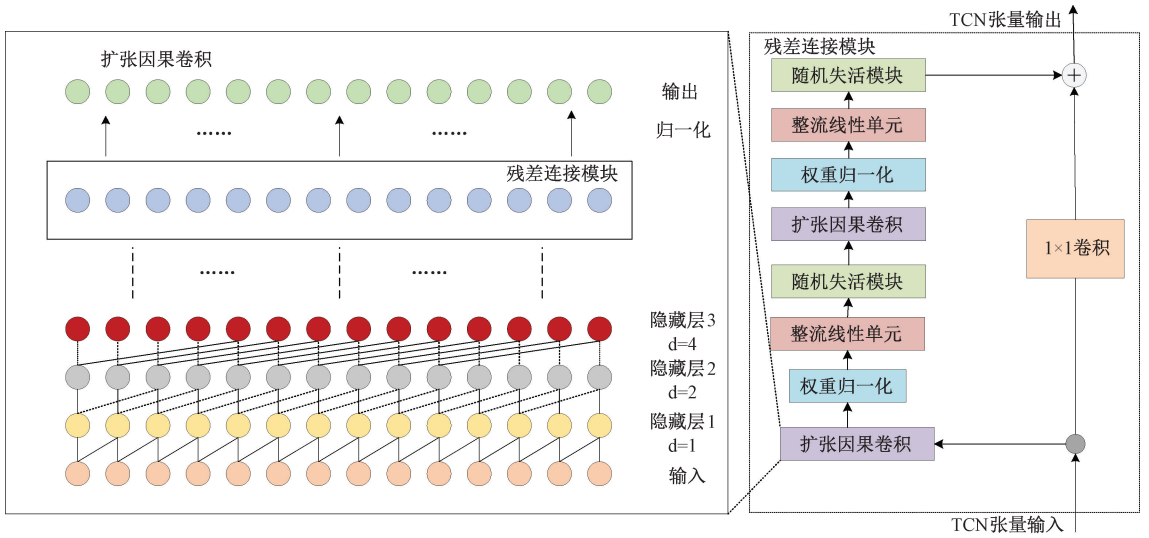


图2 TCN 模型结构

Fig. 2 Structural diagram of the TCN model

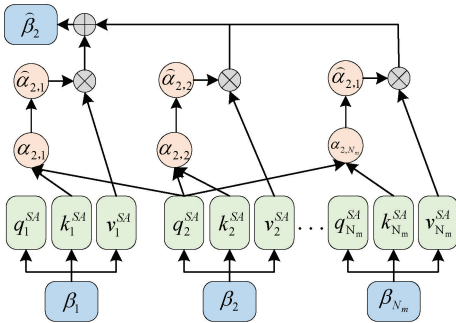


图3 SA 机制结构

Fig. 3 Structural diagram of the self-attention mechanism

2) 改进的 GJO 优化算法

GJO 算法容易陷入局部最优,且收敛速度较慢,本文引入佳点集初始化、动态权重更新策略和柯西变异策略以改进 GJO 算法。

(1) 基于佳点集初始化的种群位置

随机初始化生成的种群分布不均匀,且随机性较强,易导致 GJO 算法陷入局部最优。目前大多数研究采用混沌映射进行种群初始化,但种群分布易出现聚集现象,降低了算法的整体探索能力^[27]。佳点集初始化生成的种群在搜索空间分布更加均匀,样本点偏差较小,且偏差只受种群大小影响,从而降低算法陷入局部最优解的风险,其表达式如式(24)所示。

$$x_i(j) = lb_j + (ub_j - lb_j) \times [rj^{(i)} \cdot k] \quad (24)$$

式中: ub_j 和 lb_j 分别表示第 j 维的上界和下界; $[rj^{(i)} \cdot k]$ 表示佳点集取小数部分。

(2) 基于动态权重更新策略的探索阶段

在 GJO 算法的探索阶段,金豺的搜索范围会随着迭

代次数的增加而逐渐缩小,导致算法收敛速度降低。本文引入动态权重更新策略使金豺可自适应调整其搜索范围,增强金豺在搜索后期的寻优能力,提高算法的收敛速度,其计算公式如式(25)所示。

$$\omega = \begin{cases} \omega_{\max}, & f_i \geq f_a \\ \omega_{\min} + \frac{(\omega_{\max} - \omega_{\min}) \times (f_i - f_{\min})}{f_{\min}}, & f_i < f_a \end{cases} \quad (25)$$

式中: ω 为动态权重; f_i 为第 i 个金豺的适应度; f_a 为所有金豺的平均适应度; $f_{\min}$ 为最小适应度; $f_{\max}$ 为最大适应度; $\omega_{\max}$ 为最大权重; $\omega_{\min}$ 为最小权重。

雄、雌金豺相对猎物的位置向量更新表达式如式(26)和(27)所示。

$$Y_1(t) = Y_M(t) - E \times \omega \mid Y_M(t) - rl \cdot Prey(t) \mid \quad (26)$$

$$Y_2(t) = Y_{FM}(t) - E \times \omega \mid Y_{FM}(t) - rl \cdot Prey(t) \mid \quad (27)$$

式中: $Y_1(t)$ 和 $Y_2(t)$ 分别表示雄、雌豺狼搜索猎物更新后的位置向量; t 为当前迭代次数; $Y_M(t)$ 和 $Y_{FM}(t)$ 分别表示当前雄、雌性豺狼的位置向量; E 为猎物逃跑能量; rl 为基于 Levy 分布的向量; $Prey(t)$ 表示猎物位置向量。

(3) 基于柯西变异策略的开发阶段

在 GJO 算法的开发阶段,金豺种群会快速聚集在参数最优解附近,导致种群多样性显著降低,增大算法陷入局部最优值的风险^[28]。本文引入了柯西变异策略,对雄金豺(最优解)进行柯西扰动,扩大算法的搜索空间,以提高 GJO 算法跳出局部最优的能力。改进后的雄金豺位置向量更新表达式如式(28)所示。

$$Y_{new} = \begin{cases} Y_{best}(1 + Cauchy(0,1)), & f(Y_{best}) > f(Y_{new}) \\ Y_{best}, & f(Y_{best}) \leq f(Y_{new}) \end{cases} \quad (28)$$

式中: Y_{best} 为柯西变异前的雄金豺位置向量; Y_{new} 为柯西变异后的雄金豺位置向量; $Cauchy(0,1)$ 为标准柯西分布。

IGJO 的算法流程如图 4 所示。

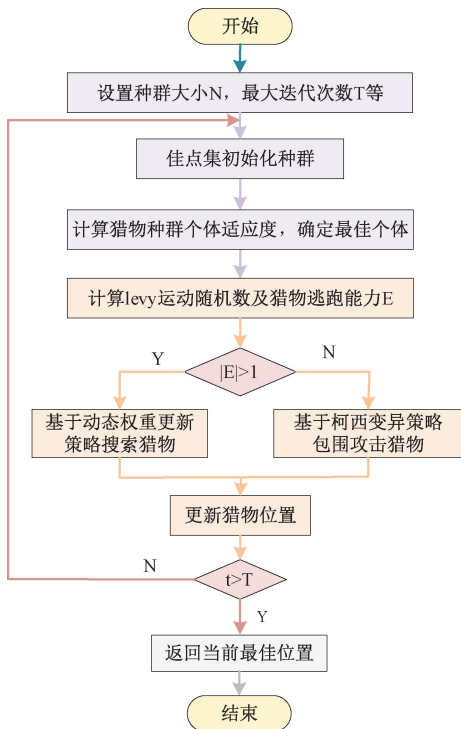


图 4 IGJO 算法流程

Fig. 4 IGJO algorithm flowchart

1.6 基于 IGJO-TCN-BiGRU-SA 的水质预测模型

IGJO-TCN-BiGRU-SA 水质预测模型由 TCN-BiGRU-SA 水质预测模块和 IGJO 算法模块两部分组成, 如图 5 所示。预测过程如下:

步骤 1) 采用 IGJO 算法优化 TCN 的 L2、BiGRU 的学习率和神经元个数及 SA 机制的键值, 得到模型超参数的最佳组合;

步骤 2) 输入层的水质数据 x_n 进入 TCN, TCN 通过扩张卷积提取数据的局部特征, 并将特征处理后的数据输入到 BiGRU 层;

步骤 3) BiGRU 层通过正向 GRU 和反向 GRU 分析序列中的上下文信息, 挖掘数据中潜在的未来变化特征;

步骤 4) SA 机制根据 BiGRU 输出数据的重要性进行加权运算, 并将结果输入至全连接层;

步骤 5) 全连接层对输入结果进行非线性变换, 生成最终的预测结果 y_n 。

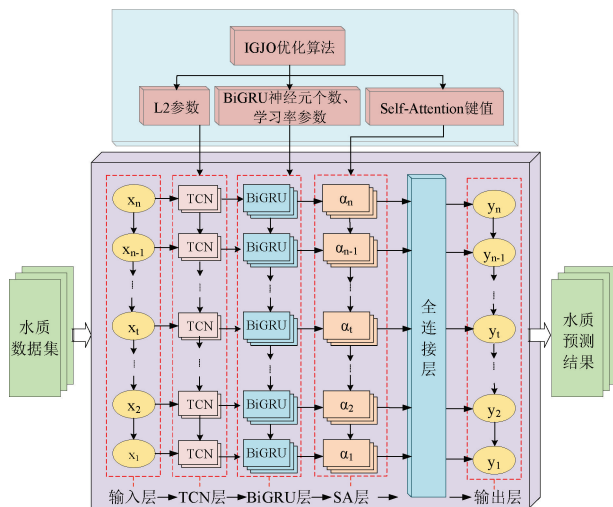


图 5 基于 IGJO-TCN-BiGRU-SA 的水质预测模型

Fig. 5 Water quality prediction model based on IGJO-TCN-BiGRU-SA

2 实验及分析

2.1 实验环境

本研究实验基于 MATLAB R2023b 仿真平台开展。实验平台硬件配置包括 AMD Ryzen 9 7945HX 中央处理器、NVIDIA GTX 4060 独立显卡和 16 GB DDR5 内存。操作系统为 Windows 11 64 位专业版。

2.2 数据来源与处理

1) 数据来源

淮河流域作为中国南北气候的分界, 它在促进南北文化交流、支撑农业生产、推动经济发展和水资源管理等方面发挥着至关重要的作用, 其地理位置如图 6 所示。本文所用水质数据集来自于淮河流域安徽省阜阳市王家坝站, 时间为 2020 年 1 月 1 日 ~ 2022 年 6 月 30 日 (每 4 h 采集一次), 参数包括的温度 (temp)、pH 值、溶解氧 (dissolved oxygen, DO)、化学需氧量 (chemical oxygen demand, COD) 质量浓度、氨氮 (ammonia nitrogen, $\text{NH}_3\text{-N}$) 质量浓度、总磷 (total phosphorus, TP) 质量浓度、电导率 (electrical conductivity, EC)、浊度 (turbidity, TU) 和总氮 (total nitrogen, TN), 共计 3 534 组。数据集以 7 : 3 的比例划分为训练集和测试集。水质数据中各参数范围及中国水质分类标准如表 1 所示。

2) 数据处理

(1) 基于 CSI 算法的缺失数据处理

对原始数据集的缺失情况进行了统计分析, 其结果如表 2 所示。

由表 2 可知, Temp、pH、DO、EC 和 TN 没有缺失数据, COD、 $\text{NH}_3\text{-N}$ 、TP 和 TU 存在不同程度的缺失, 缺失率

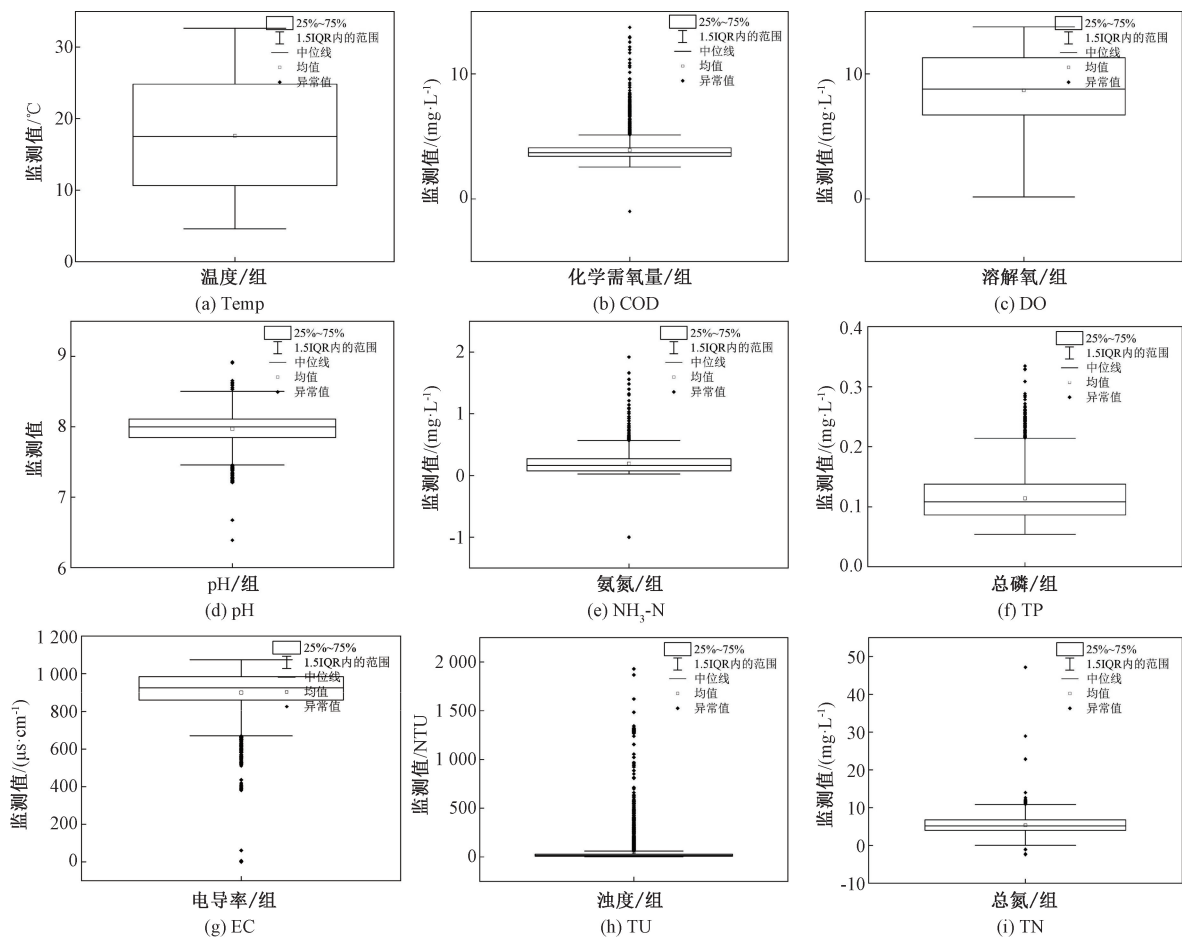


图 7 基于箱线图的异常数据分析结果

Fig. 7 Results of anomalous data analysis based on boxplots

表 3 水质异常数据分析结果

Table 3 Abnormal data analysis results of water quality indexes

水质指标	25%	50%	75%	上界	下界	异常数据数	异常率/%
Temp/℃	10. 65	17. 50	24. 84	32. 64	4. 60	0	0
pH	7. 85	8. 00	8. 11	8. 51	7. 46	105	2. 97
DO/(mg·L ⁻¹)	6. 72	8. 79	11. 31	13. 76	0. 16	0	0
COD/(mg·L ⁻¹)	3. 40	3. 99	4. 07	5. 06	2. 55	225	6. 37
ρ(NH ₃ -N)/(mg·L ⁻¹)	0. 07	0. 16	0. 27	0. 57	0. 03	127	3. 59
ρ(TP)/(mg·L ⁻¹)	0. 09	0. 11	0. 14	0. 21	0. 05	102	2. 89
EC/(μs·cm ⁻¹)	858. 80	925. 56	984. 91	1 075. 14	669. 96	158	4. 47
TU/(NTU)	8. 3	17. 18	28. 69	59. 23	0. 98	389	11
ρ(TN)/(mg·L ⁻¹)	4. 03	5. 13	6. 75	10. 81	0. 05	23	0. 65

(4) 数据归一化处理

为加快模型的收敛,本研究通过数据归一化将不同量纲的数据统一映射至区间[0,1],使尺度不同的数据在模型训练中具有相同的影响权重,数据归一化通过优化模型学习率的设置加快模型收敛,其基本公式如式(29)所示。

$$N_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \tag{29}$$

式中: N_i 表示归一化处理的数据; X_i 表示原始数据; $X_{\max}$ 表示数据的最大值; $X_{\min}$ 表示数据的最小值。

2.3 基于 GRA 算法的相关性分析

设置水质指标中的任意一项为参考序列,其余水质指标则为比较序列,使用 GRA 算法计算水质指标间的灰

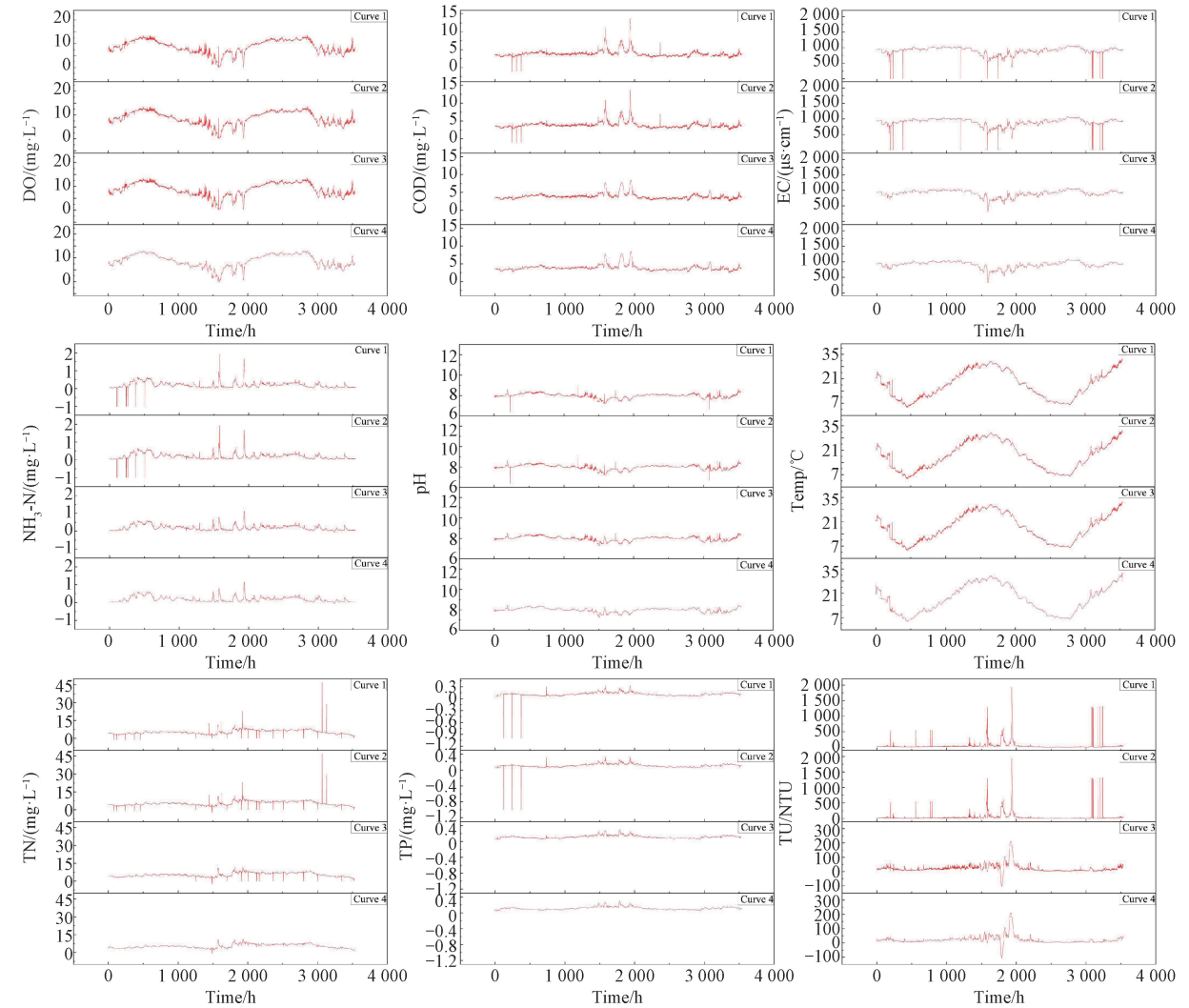


图8 缺失数据和异常数据处理

Fig. 8 Processing of missing data and abnormal data

色关联度,结果如图 9 所示。根据 GRA 算法原理,任意水质指标与自身的灰色关联度恒为 1,与其他水质指标的灰色关联度介于[0,1]内。为区分水质指标间的相关程度,根据灰色关联度的区间对相关程度进行划分,如表 5 所示。

表 4 水质指标数据
Table 4 Experimental data

水质指标	最小值	最大值	平均值	标准偏差
温度/℃	4.6	32.64	17.60	7.78
pH	7.58	8.44	8.01	0.17
DO/(mg·L ⁻¹)	0.43	13.02	8.69	2.75
COD/(mg·L ⁻¹)	2.67	4.70	3.87	0.39
ρ(NH ₃ -N)/(mg·L ⁻¹)	0.01	0.54	0.19	0.12
ρ(TP)/(mg/L)	0.03	0.21	0.11	0.03
EC/(μs·cm ⁻¹)	680.27	1 132.57	899.20	83.82
TU/(NTU)	0.06	60.00	50.33	11.56
ρ(TN)/(mg·L ⁻¹)	1.23	10.55	5.33	1.57

表 5 相关程度的划分
Table 5 Classification of correlation levels

相关程度	区间
弱相关	[0, 0.5]
一般相关	[0.5, 0.7]
强相关	[0.7, 1]

由图 9 可知,水质指标任意一项与其余指标之间的相关系数均介于 0~1,选择强相关性参数作水质指标的输入数据。其中,TP 与其余水质指标的相关性参数皆小于 0.7,故根据专家经验选择相关性参数最大的 3 个指标作为输入数据,水质预警模型的输入和输出如表 6 所示。

2.4 模型评价指标

为评价模型的预测性能,选择均方根误差 (root mean square error, RMSE)、平均绝对误差 (mean absolute error, MAE)、平均绝对百分比误差 (mean absolute

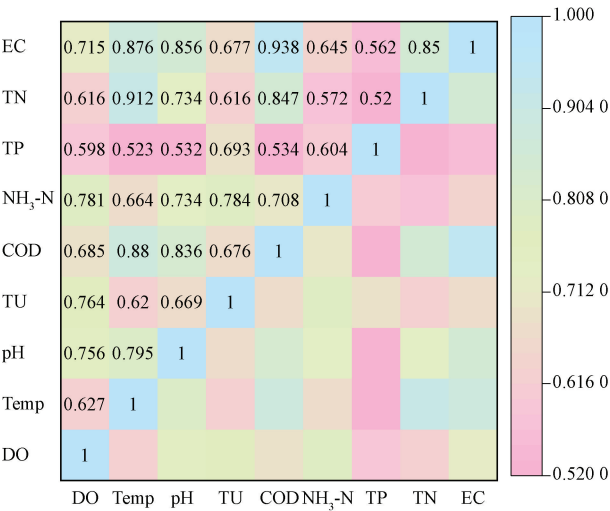


图 9 基于 GRA 的相关性分析结果

Fig. 9 Results of correlation analysis based on GRA

percentage error, MAPE) 和决定系数 (coefficient of determination, R^2) 作为评价指标。其中, RMSE 的范围为

$[0, +\infty]$, 表示预测误差的波动强度; MAE 的范围为 $[0, +\infty]$, 表示预测误差的平均水平; MAPE 的取值范围为 $[0, +\infty]$, 表示误差相对于真实值的比例; R^2 的范围为 $[0, 1]$, 表示模型对数据的拟合程度。RMSE、MAPE、MAE 的值越接近于 0 表示模型的预测误差越小, R^2 值越接近 1 表示模型的拟合效果越好, 计算公式如式 (30) ~ (33) 所示。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y(i) - f(X_i))^2}$$
 (30)

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y(i) - f(X_i)|$$
 (31)

$$MAPE = \frac{1}{N} \times \sum_{i=1}^N \frac{|Y(i) - f(X_i)|}{Y(i)} \times 100\%$$
 (32)

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y(i) - f(X_i))^2}{\sum_{i=1}^N (Y(i) - \bar{Y})^2}$$
 (33)

式中: N 表示样本数量; $Y(i)$ 表示实际值; $f(x_i)$ 表示预测值; $\bar{Y}$ 表示平均值。

表 6 水质预测模型的输入和输出

Table 6 Inputs and outputs of the water quality prediction model

输出	pH	COD	EC	DO	TN	NH ₃ -N	TP	TU
输入	pH	COD	EC	DO	TN	NH ₃ -N	TP	TU
	DO	Temp	DO	pH	Temp	DO	TU	DO
	Temp	pH	Temp	TU	pH	pH	DO	NH ₃ -N
	COD	NH ₃ -N	pH	NH ₃ -N	COD	TU	NH ₃ -N	
	NH ₃ -N	TN	COD	EC	EC	COD		
	TN	EC	TN					
	EC							

2.5 IGJO 优化算法性能验证

为验证 IGJO 算法的优化性能, 设置 GA、PSO、GWO 和 GJO 算法作为对照组与 IGJO 算法进行比较。实验选用 9 组基准测试函数对各算法进行性能评估, 涵盖单峰函数 ($F_1 \sim F_3$)、多峰函数 ($F_4 \sim F_6$) 及固定维多峰函数 ($F_7 \sim F_9$), 每类各包含 3 个函数, 如表 7 所示。其中, 单峰函数用于测试算法的收敛速度; 多峰函数用于评估算法跳出局部最优能力; 固定维多峰函数用于评估算法的全局寻优性能。

1) IGJO 算法的全局收敛性及收敛速度分析

为确保对比实验的公平性, 所有算法均按照统一的参数设置, 种群规模和最大迭代次数分别设为 100 和 500。为分析各算法的全局收敛性及收敛速度, 根据每轮迭代中算法适应度值的变化, 绘制出相应的适应度曲线, 如图 10 所示。

由图 10 可知, 与其他优化算法相比, IGJO 算法能够在最少的迭代次数找到最优的适应度值, 表明其收敛精

度更高, 且收敛速度更快。

2) IGJO 算法的寻优精度及稳定性分析

为降低随机误差对实验结果的干扰, 将上述测试实验独立重复 30 次, 得到 30 个最优适应度值, 并计算其平均值 (Mean) 和标准差 (Std)。这些指标被作为优化算法寻优精度及稳定性的评价指标。不同算法在测试函数中的寻优结果如表 8 所示。

由表 8 可知, 相对于上述对比算法, IGJO 在所有测试函数上最优适应度的平均值均最接近理论最优值, 且在 F_1 和 F_6 测试函数中达到了理论最优值, 表明 IGJO 算法的寻优精度更高; IGJO 算法在所有测试函数上最优适应度的标准差均最低, 且在 F_1 、 F_2 、 F_5 和 F_6 测试函数上标准差等于 0, 表明 IGJO 算法多次寻优的输出结果一致, 其稳定性更强。

2.6 基于训练集的水质预测模型参数和结构学习

本研究基于训练集进行水质预测模型参数和结构学

习,初始水质预测模型设置为:学习次数为 100 次、随机失波器个数、大小和区块数分别为 128、5 和 8,BiGRU 神经元活率为 0.2、L2 参数为 0.000 1,学习率为 0.001、TCN 的滤个数为 20、注意力机制键值为 50、优化器为 Adam。

表 7 测数函数

Table 7 Measurement functions				
种类	函数	维度	取值范围	理论最优值
单峰	$F_1(x) = \sum_{i=1}^n (\sum_{j=1}^i x_j)^2$	30	$[-100,100]$	0
	$F_2(x) = \max_i \{ x_i , 1 \leq i \leq n \}$	30	$[-100,100]$	0
	$F_3(x) = \sum_{i=1}^n ([x_i + 0.5])^2$	30	$[-100,100]$	0
多峰	$F_4(x) = \sum_{i=1}^n -x_i \sin(\sqrt{ x_i })$	30	$[-500,500]$	-12 569.487
	$F_{5(x)} = -20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^n x_i^2}\right) - \exp\left[\frac{1}{n}\sum_{i=1}^n \cos(2\pi x_i)\right] + 20 + e$	30	$[-32,32]$	0
	$F_6(x) = \frac{1}{4\,000}\sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos(\frac{x_i}{\sqrt{i}}) + 1$	30	$[-600,600]$	0
固定维多峰	$F_7(x) = -\sum_{i=1}^4 [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	$[0,10]$	-10.153 2
	$F_8(x) = -\sum_{i=1}^7 [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	$[0,10]$	-10.402 9
	$F_9(x) = -\sum_{i=1}^{10} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	$[0,10]$	-10.536 4

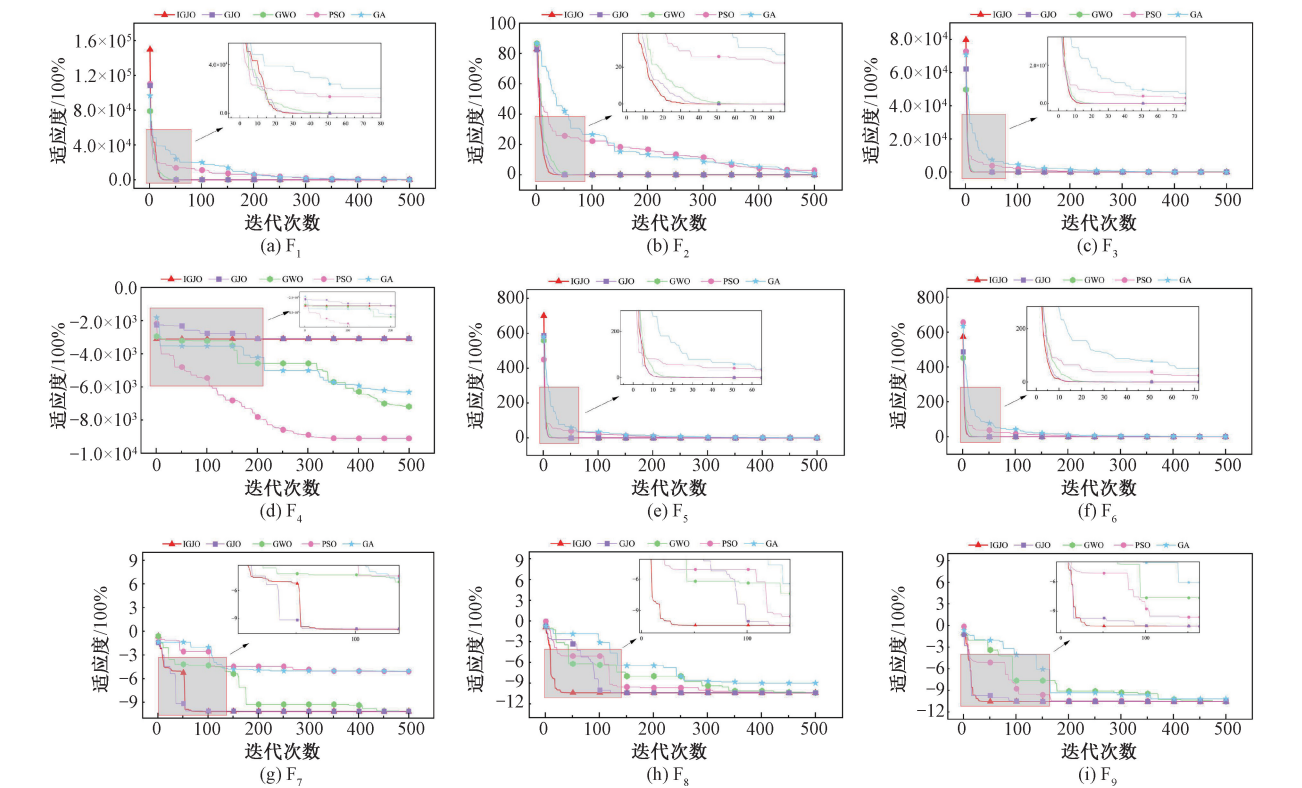


图 10 各测试函数收敛曲线

Fig. 10 Convergence curves of different test functions

1) 模型学习次数的确定

学习次数不足会导致模型欠拟合,无法充分捕获数

据特征,而学习次数过多又会导致模型过拟合。本实验使用不同的学习次数来测试模型,得到模型评价指标,如

表 8 不同算法测试函数寻优结果对比

Table 8 Comparison of test function optimization results of different algorithms

函数名	指标	IGJO	GJO	GWO	PSO	GA
F ₁	Mean	0	2.006 8×10 ⁻²⁵	1.544 6×10 ⁻¹¹	8.312 5×10 ²	2.855 7×10 ¹
	Std	0	8.700 7×10 ⁻²⁵	4.884 8×10 ⁻¹¹	1.264 1×10 ³	1.127 8×10 ¹
F ₂	Mean	5.099 3×10 ⁻¹⁸⁷	1.881 5×10 ⁻²²	2.276 9×10 ⁻¹⁰	3.841 0×10 ⁰	7.370 8×10 ⁻¹
	Std	0	3.518 3×10 ⁻²²	3.571 7×10 ⁻¹⁰	1.026 8×10 ⁰	2.172 0×10 ⁻¹
F ₃	Mean	2.255 3×10 ⁻⁵	1.589 5×10 ⁻⁴	5.780 6×10 ⁻⁴	2.500 5×10 ⁻²	1.232 8×10 ⁻²
	Std	2.430 7×10 ⁻⁵	1.203 9×10 ⁻⁴	2.626 8×10 ⁻⁴	7.654 1×10 ⁻³	3.394 6×10 ⁻³
F ₄	Mean	-2.875 7×10 ³	-4.425 6×10 ³	-6.372 2×10 ³	-9.259 7×10 ³	-7.997 4×10 ³
	Std	2.758 4×10 ²	1.287 7×10 ³	6.675 4×10 ²	4.665 6×10 ²	7.470 8×10 ²
F ₅	Mean	4.440 9×10 ⁻¹⁶	5.062 6×10 ⁻¹⁵	2.590 5×10 ⁻¹⁴	1.910 9×10 ⁻³	7.954 3×10 ⁻¹
	Std	0	1.655 9×10 ⁻¹⁵	3.623 5×10 ⁻¹⁵	4.261 8×10 ⁻³	6.522 4×10 ⁻¹
F ₆	Mean	0	0	2.898 1×10 ⁻³	1.821 6×10 ⁻²	6.162 9×10 ⁻¹
	Std	0	0	6.348 3×10 ⁻³	2.218 5×10 ⁻²	8.180 5×10 ⁻²
F ₇	Mean	-1.015 3×10 ¹	-9.642 7×10 ⁰	-9.133 6×10 ⁰	-8.210 6×10 ⁰	-7.557 8×10 ⁰
	Std	6.206 3×10 ⁻¹⁵	1.710 7×10 ⁻¹¹	2.074 1×10 ⁰	2.632 1×10 ⁰	3.324 8×10 ⁰
F ₈	Mean	-1.040 3×10 ¹	-9.671 5×10 ⁰	-9.012 9×10 ⁰	-5.993 7×10 ⁰	-4.643 3×10 ⁰
	Std	4.664 9×10 ⁻¹⁶	2.180 3×10 ⁻¹¹	4.776 4×10 ⁻⁴	9.704 3×10 ⁻¹	1.348 3×10 ⁰
F ₉	Mean	-1.053 6×10 ¹	-1.006 5×10 ¹	-1.053 5×10 ¹	-9.219 0×10 ⁰	-8.942 5×10 ⁰
	Std	2.618 2×10 ⁻¹⁵	1.633 2×10 ⁻¹¹	5.311 5×10 ⁻⁴	2.738 1×10 ⁰	2.740 2×10 ⁰

表 9 所示。

表 9 不同学习次数的评价指标

Table 9 Evaluation indicators for different learning times

学习次数	MAE	MAPE	RMSE	R ²
100	0.069 456	0.019 782	0.095 803	0.963 59
150	0.057 025	0.016 6	0.0766 21	0.976 71
200	0.056 437	0.015 946	0.075 16	0.977 59
250	0.040 089	0.011 332	0.054 48	0.988 23
300	0.041 224	0.011 695	0.055 426	0.987 81
350	0.041 381	0.011 739	0.055 689	0.987 70
400	0.043 068	0.012 171	0.057 97	0.986 67
450	0.045 997	0.012 774	0.060 044	0.985 70
500	0.050 858	0.014 696	0.067 327	0.982 02

由表 9 可知,模型学习次数为 250 的 MAE、MAPE、RMSE 和 R² 分别为 0.040 089、0.011 332、0.054 48 和 0.988 23,相对于其余学习次数的测试,此时的模型预测性能最佳。当学习次数超过 250 后,MAE、MAPE 和 RMSE 逐渐增加,R² 逐渐下降,表明模型出现过拟合现象。故综合考虑,将模型的学习次数设置为 250。

2) TCN 网络结构的确定

TCN 的滤波器个数通常设置为 64、128 或 256,卷积核大小通常设置为 3、5 或 7,膨胀系数一般设置为 2。根据专家知识,将 TCN 网络结构的滤波器个数设置为 128、卷积核大小设置为 5 和膨胀系数设置为 2。

3) TCN 中 L2 的确定

TCN 的 L2 通过添加惩罚项进行工作,用于避免过拟合问题。采用 IGJO 算法寻优 TCN 的 L2,为了确定合理

的参数搜索空间,通过多组预实验测试确定边界条件为 [0.000 1,0.001],所得最佳 L2 参数为 0.000 106。

4) BiGRU 学习率的确定

BiGRU 学习率用于控制参数更新的步长。学习率过大会导致局部最优;而学习率过小又会参数更新效率低。为平衡训练速度与稳定性,采用 IGJO 算法寻优 BiGRU 学习率,通过多组预实验测试确定边界条件为 [0.001,0.01],所得最佳学习率参数为 0.008 6。

5) BiGRU 神经元个数的确定

BiGRU 神经元个数是指隐层层的神经元数量,增加神经元数量能够提升模型的学习能力,但会降低运算效率。采用 IGJO 算法寻优 BiGRU 神经元个数,基于文献[14]中的 GRU 神经元个数设置确定边界条件为 [10,50],所得最佳 BiGRU 神经元个数为 32。

6) SA 机制键值的确定

SA 机制的键值影响模型的权重分配,采用 IGJO 算法寻优 SA 机制的键值,为覆盖参数敏感区域并减少无效搜索空间,通过多组预实验测试确定边界为 [2,50],所得最佳键值参数为 43。

7) 模型预测持续时间的确定

工程实践中,需要针对具体时间序列预测问题选择合适的预测策略^[29]。预测策略主要分为单步预测和多步预测,不同的预测策略对应的预测持续时间不同,导致模型的预测性能也会有差异。本实验设置模型预测持续时间为 1~6 d,分别测试模型,进一步探索了不同预测时间下 IGJO-TCN-BiGRU-SA 模型的预测性能,得到的模型评价指标如表 10 所示。

表 10 不同预测持续时间的模型预测评价指标
Table 10 Evaluation metrics for model predictions with different prediction durations

预测持续 时间	MAE	MAPE	RMSE	R^2
1	0.037 321	0.010 627	0.051 389	0.989 52
2	0.090 369	0.026 524	0.111 49	0.950 65
3	0.110 43	0.031 891	0.143 6	0.918 14
4	0.151 58	0.043 101	0.198 91	0.842 97
5	0.199 38	0.056 419	0.260 97	0.729 72
6	0.202 12	0.056 232	0.273 57	0.703 05

由表 10 可知,模型预测持续时间为 1 d 的 MAE、MAPE、RMSE 和 R^2 分别为 0.037 321、0.010 627、0.051 389 和 0.989 52,评价指标的表现均最优,故将模型预测持续时间设置为 1 d。

2.7 基于测试集的模型预测结果分析

为了验证 IGJO-TCN-BiGRU-SA 模型的有效性,基于测试集对模型进行性能测试。同时为了证明 IGJO-TCN-BiGRU-SA 模型的优越性,设置 LSTM、GRU、BiGRU、和 TCN-BiGRU-SA 为对照组与 IGJO-TCN-BiGRU-SA 进行对比。各模型的预测评价指标如表 11 所示。由于篇幅限制,仅阐述对比实验中 DO 的预测结果图,如图 11 所示。

表 11 基于测试集的各模型预测评价指标
Table 11 Evaluation metrics for each model prediction based on the test set

水质指标	预测模型	MAE	MAPE	RMSE	R^2
pH	LSTM	0.020 489	0.002 578 6	0.032 445	0.969 05
	GRU	0.0196 04	0.002 459 3	0.029 583	0.974 27
	BiGRU	0.016 508	0.002 071 0	0.025 891	0.980 29
	TCN-BiGRU-SA	0.013 454	0.001 689 8	0.021 612	0.986 26
	IGJO-TCN-BiGRU-SA	0.009 1	0.001 142 2	0.015 013	0.993 37
DO	LSTM	0.323 44	0.042 637	0.500 25	0.961 62
	GRU	0.321 79	0.040 52	0.454 47	0.968 32
	BiGRU	0.275 21	0.028 607	0.317 08	0.984 61
	TCN-BiGRU-SA	0.127 64	0.015 574	0.166 48	0.995 76
	IGJO-TCN-BiGRU-SA	0.102 57	0.011 102	0.124 07	0.997 64
COD	LSTM	0.072 401	0.020 46	0.098 127	0.961 77
	GRU	0.064 586	0.018 151	0.081 954	0.973 34
	BiGRU	0.055 904	0.015 271	0.070 791	0.980 10
	TCN-BiGRU-SA	0.048 224	0.013 278	0.067 109	0.982 12
	IGJO-TCN-BiGRU-SA	0.034 748	0.010 072	0.048 134	0.990 80
$\rho(\text{NH}_3\text{-N})$	LSTM	0.021 482	0.264 99	0.025 832	0.944 32
	GRU	0.014 784	0.228 61	0.020 979	0.963 25
	BiGRU	0.010 645	0.225 40	0.015 158	0.980 81
	TCN-BiGRU-SA	0.009 686	0.197 19	0.013 249	0.985 33
	IGJO-TCN-BiGRU-SA	0.006 913 2	0.190 17	0.011 564	0.988 82
$\rho(\text{TP})$	LSTM	0.004 147 3	0.040 349	0.005 388 3	0.959 35
	GRU	0.002 488 9	0.024 733	0.003 471 6	0.983 12
	BiGRU	0.002 167 7	0.021 495	0.003 008 0	0.987 33
	TCN-BiGRU-SA	0.001 609 4	0.016 651	0.002 102 0	0.993 81
	IGJO-TCN-BiGRU-SA	0.001 35	0.013 557	0.001 835 3	0.995 29
EC	LSTM	11.008	0.011 456	14.030	0.957 09
	GRU	8.683 6	0.008 875 3	11.882	0.969 24
	BiGRU	7.519 8	0.007 959 1	10.268	0.977 03
	TCN-BiGRU-SA	5.236 5	0.005 5	7.440 2	0.987 93
	IGJO-TCN-BiGRU-SA	4.747 7	0.004 908 2	7.385 2	0.988 11
TU	LSTM	1.627 40	0.178 57	2.502 0	0.918 98
	GRU	1.457 50	0.156 86	2.180 4	0.938 47
	BiGRU	1.323 00	0.136 08	2.035 1	0.946 37
	TCN-BiGRU-SA	1.080 90	0.105 84	1.642 0	0.965 18
	IGJO-TCN-BiGRU-SA	0.969 43	0.102 99	1.442 7	0.973 04
$\rho(\text{TN})$	LSTM	0.291 84	0.056 541	0.404 59	0.944 53
	GRU	0.219 05	0.045 231	0.283 64	0.972 72
	BiGRU	0.131 30	0.025 208	0.175 20	0.989 59
	TCN-BiGRU-SA	0.108 57	0.022 914	0.124 88	0.994 71
	IGJO-TCN-BiGRU-SA	0.082 41	0.017 265	0.115 23	0.995 5

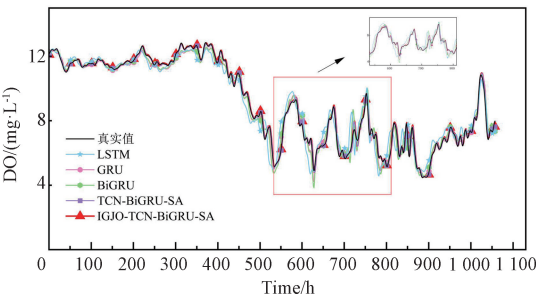


图 11 模型预测结果对比

Fig. 11 Comparison of model prediction results

由表 11 和图 11 可知,相对于 LSTM、GRU、BiGRU、BiGRU-SA、TCN-BiGRU 和 TCN-BiGRU-SA, IGJO-TCN-BiGRU-SA 模型的 MAE 分别降低了 68.29%、68.12%、

29.07%和 19.64%,表明其预测误差的平均值更低;RMSE 分别降低了 75.19%、72.7%、37.85%和 25.47%,表明其预测误差波动更小;MAPE 分别降低了 73.36%、72.6%、38.07%和 28.71%,表明其误差相对于真实值的比例更低; R^2 分别提高了 3.61%、2.94%、0.38%和 0.19%,表明其对数据的拟合效果更好。

综上所述,相对于传统的预测算法,IGJO-TCN-BiGRU-SA 模型预测精度明显提升。

2.8 模型消融实验

为了明确 TCN-BiGRU-SA 模型各组件对预测精度的影响,进行 BiGRU、BiGRU-SA、TCN-BiGRU 模型与 TCN-BiGRU-SA 的消融实验,得到消融实验评价指标,如表 12 所示。基于表 12 绘制了评价指标的雷达图,如图 12 所示。

表 12 消融实验的评价指标结果

Table 12 Evaluation index results of ablation experiment

水质指标	预测模型	MAE	MAPE	RMSE	R^2
pH	BiGRU	0.016 508	0.002 071 0	0.025 891	0.980 29
	BiGRU-SA	0.016 134	0.001 944 7	0.025 037	0.981 56
	TCN-BiGRU	0.015 215	0.001 808 8	0.023 664	0.983 53
	TCN-BiGRU-SA	0.013 454	0.001 689 8	0.021 612	0.986 26
DO	BiGRU	0.275 21	0.028 607	0.317 08	0.984 61
	BiGRU-SA	0.163 86	0.019 736	0.212 11	0.993 08
	TCN-BiGRU	0.144 60	0.017 928	0.199 64	0.993 87
	TCN-BiGRU-SA	0.127 64	0.015 574	0.166 48	0.995 76
COD	BiGRU	0.055 904	0.015 271	0.070 791	0.980 10
	BiGRU-SA	0.053 339	0.014 850	0.070 410	0.980 34
	TCN-BiGRU	0.052 797	0.013 676	0.067 490	0.981 94
	TCN-BiGRU-SA	0.048 224	0.013 278	0.067 109	0.982 12
$\rho(\text{NH}_3\text{-N})$	BiGRU	0.010 645	0.225 40	0.015 158	0.980 81
	BiGRU-SA	0.010 570	0.204 08	0.015 056	0.981 20
	TCN-BiGRU	0.010 133	0.200 34	0.014 151	0.983 39
	TCN-BiGRU-SA	0.009 686	0.197 19	0.013 249	0.985 33
$\rho(\text{TP})$	BiGRU	0.002 167 7	0.021 495	0.003 008 0	0.987 33
	BiGRU-SA	0.002 131 5	0.021 066	0.002 853 6	0.988 61
	TCN-BiGRU	0.001 760 3	0.018 072	0.002 366 8	0.992 16
	TCN-BiGRU-SA	0.001 609 4	0.016 651	0.002 102 0	0.993 81
EC	BiGRU	7.519 8	0.007 959 1	10.268	0.977 03
	BiGRU-SA	6.938 5	0.007 356 5	8.656 7	0.983 73
	TCN-BiGRU	6.815 8	0.007 232 3	8.486 5	0.984 36
	TCN-BiGRU-SA	5.236 5	0.005 5	7.440 2	0.987 93
TU	BiGRU	1.323 00	0.136 08	2.035 1	0.946 37
	BiGRU-SA	1.292 20	0.133 27	1.806 6	0.949 22
	TCN-BiGRU	1.162 20	0.113 74	1.720 8	0.961 82
	TCN-BiGRU-SA	1.080 90	0.105 84	1.642 0	0.965 18
$\rho(\text{TN})$	BiGRU	0.131 30	0.025 208	0.175 20	0.989 59
	BiGRU-SA	0.130 25	0.025 002	0.168 27	0.990 42
	TCN-BiGRU	0.124 51	0.024 502	0.151 51	0.992 23
	TCN-BiGRU-SA	0.108 57	0.022 914	0.124 88	0.994 71

由表 12 和图 12 可知,相较于 BiGRU、BiGRU-SA 和 TCN-BiGRU,TCN-BiGRU-SA 模型的 MAE 分别降低了 53.62%、22.10%和 11.73%;RMSE 分别降低了 47.50%、21.51%和 16.61%;MAPE 分别降低了 45.56%、21.09%

和 13.13%; R^2 分别提高了 1.13%、0.27%和 0.19%。结果表明,相较于 BiGRU、BiGRU-SA 和 TCN-BiGRU 模型,TCN-BiGRU-SA 模型的预测精度显著提升,证明了 TCN-BiGRU-SA 模型的有效性。

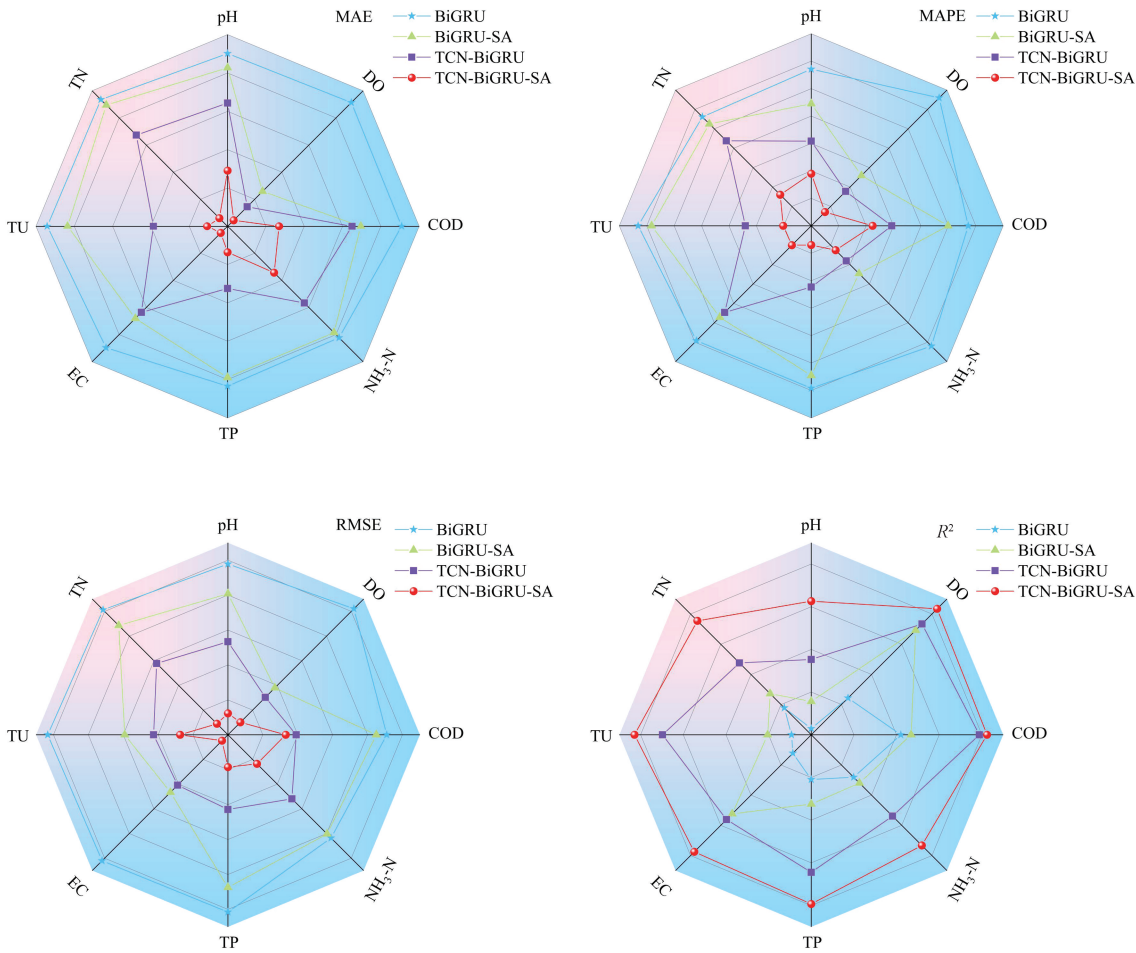


图 12 消融实验的评价指标雷达图
Fig. 12 Radar chart of predictive evaluation indicators based on ablation experiments

3 结 论

本文提出一种 IGJO-TCN-BiGRU-SA 预测算法,用于实现淮河流域的水质预测,并基于历史数据进行了实验验证。首先,采用 CSI 算法修复缺失数据、利用 Boxplot 和 FFT 算法剔除异常数据,以提高数据的完备性;采用 GRA 算法分析水质指标间的相关性,以提升模型的运算效率。其次,利用 BiGRU 挖掘数据中潜在的未来变化特征;采用 TCN 算法提取水质数据的局部特征,解决了 BiGRU 模型的过拟合问题;引入 SA 机制,根据 BiGRU 输出特征的重要性更新权重,解决了梯度消失问题。最后,针对传统优化算法寻优精度低的问题,采用 GJO 算法优化 TCN 的 L2、BiGRU 的学习率和神经元个数及 SA 机制

的键值,但存在迭代后期易陷入局部最优解和收敛速度慢的问题,通过引入佳点集初始化种群、动态权重更新策略和柯西变异策略改进 GJO 算法,提升了全局寻优能力和收敛速度。

然而,本文仅用一个监测站的水质数据,未考虑多水质监测站的位置拓扑关系,及预测模型的多步预测预测性能下降问题。未来的研究方向是在水质时序属性数据基础上,纳入水质数据的空间属性,考虑多步预测性能下降问题,构建并改进基于 GNN 的水质时空关联预测模型。

参考文献

[1] 张京新,蔡庆旺,康子怡,等. 渭河流域水质分析评价及预测模型 [J]. 环境科学, 2025, 46 (9): 5619-5640.

- ZHANG J X, CAI Q W, KANG Z Y, et al. Water quality analysis, evaluation and prediction model for the Wei river basin [J]. *Environmental Science*, 2025, 46(9): 5619-5640.
- [2] 陶志勇, 曹琦, 徐光宪. 基于 FOA-GRNN 模型的太湖水质预测研究[J]. *电子测量与仪器学报*, 2021, 35(11): 83-90.
- TAO ZH Y, CAO Q, XU G X. Research on taihu lake water quality prediction based on the FOA-GRNN Model [J]. *Journal of Electronic Measurement and Instrumentation*, 2021, 35(11): 83-90.
- [3] 杨昆, 喻臻钰, 罗毅, 等. 湖泊表面水温预测与可视化方法研究[J]. *仪器仪表学报*, 2017, 38(12): 3090-3099.
- YANG K, YU ZH Y, LUO Y, et al. Research on prediction and visualization methods of lake surface water temperature [J]. *Chinese Journal of Scientific Instrument*, 2017, 38(12): 3090-3099.
- [4] TIAN Q, LUO W, GUO L. Water quality prediction in the Yellow River source area based on the DeepTCN-GRU model[J]. *Journal of Water Process Engineering*, 2024, 59: 105052.
- [5] 徐若诗, 逢勇, 罗缙, 等. 基于 WQI 的南水北调东线江苏段水质评价及时空分布特征[J]. *环境科学*, 2024, 45(9): 5227-5234.
- XU R SH, PANG Y, LUO J, et al. Water quality evaluation and spatiotemporal distribution characteristics of the Jiangsu section of the East Route of the South to North Water Diversion Project based on WQI [J]. *Environmental Science*, 2024, 45(9): 5227-5234.
- [6] ALIZAMIR M, MORADVEISI K, AHMED K O, et al. An efficient data fusion model based on Bayesian model averaging for robust water quality prediction using deep learning strategies [J]. *Expert Systems with Applications*, 2025, 261: 125499.
- [7] 秦梓萱, 郭健, 许模. 基于 ARIMA-BP 模型的北京市平谷区地下水水质双尺度预测[J]. *兰州大学学报 (自然科学版)*, 2023, 59(1): 121-128.
- QIN Z X, GUO J, XU M. Dual scale prediction of groundwater quality in Pinggu District, Beijing based on ARIMA-BP model [J]. *Journal of Lanzhou University (Natural Science Edition)*, 2023, 59(1): 121-128.
- [8] 任安虎, 白静静, 林英新. 远程水质自动监测系统上位机设计与实现[J]. *国外电子测量技术*, 2019, 38(10): 115-118.
- REN AN H, BAI J J, LIN Y X. Design and implementation of remote water quality automatic monitoring system host computer[J]. *Foreign Electronic Measurement Technology*, 2019, 38(10): 115-118.
- [9] YU J W, KIM J S, LI X, et al. Water quality forecasting based on data decomposition, fuzzy clustering and deep learning neural network [J]. *Environmental Pollution*, 2022, 303: 119136.
- [10] 陈昭明, 王伟, 赵迎, 等. 改进主成分分析与多元回归融合的汉丰湖水水质评估及预测[J]. *环境监测管理技术*, 2020, 32(4): 15-19.
- CHEN ZH M, WANG W, ZHAO Y, et al. Improved principal component analysis and multiple regression fusion for water quality assessment and prediction of Hanfeng Lake [J]. *Environmental Monitoring Management and Technology*, 2020, 32(4): 15-19.
- [11] DASTORANI M, MIRZAVAND M, DASTORANI M T, et al. Simulation and prediction of surface water quality using stochastic models[J]. *Sustainable Water Resources Management*, 2020, 6(4): 74.
- [12] 陈亚松, 刘家雯, 赵云鹏, 等. 基于机器学习的人工湿地出水水质预测与影响因素[J]. *中国环境科学*, 2025, 45(6): 3161-3170.
- CHEN Y S, LIU J W, ZHAO Y P, et al. Prediction of effluent water quality from constructed wetlands using machine learning and its influencing factors[J]. *Chinese Environmental Science*, 2025, 45(6): 3161-3170.
- [13] GAO Z, WANG G, ZHU Y, et al. Prediction of water quality parameters and pollution exceedance analysis in typical rivers of semi-arid regions based on interpretable deep learning models [J]. *Environmental Pollution*, 2025: 126801. <https://doi.org/10.1016/j.envpol.2025.126801>.
- [14] 陈静, 李海洋. 基于 GRA-GRU 的淮河流域水质预测研究[J]. *安全与环境学报*, 2024, 24(1): 376-387.
- CHEN J, LI H Y. Research on water quality prediction in the Huaihe River basin based on GRA-GRU [J]. *Journal of Safety and Environment*, 2024, 24(1): 376-387.
- [15] ZHANG Y F, FITCH P, THORBURN P J. Predicting the trend of dissolved oxygen based on the kPCA-RNN model[J]. *Water*, 2020, 12(2): 585-585.
- [16] 肖扬岚, 沈惠柔, 许一涵, 等. 基于 GBDT-LSTM 的闽江流域水质预测[J]. *生态环境学报*, 2024, 33(4): 597-606.
- XIAO Y L, SHEN H R, XU Y H, et al. Water quality prediction of minjiang river basin based on GBDT-LSTM[J]. *Journal of Ecological Environment*, 2024, 33(4): 597-606.
- [17] 黎煜昭, 刘启亮, 邓敏, 等. 基于物理约束 GRU 神经网络的河流水质预测模型[J]. *地球信息科学学报*,

- 2023, 25(1): 102-114.
- LI Y ZH, LIU Q L, DENG M, et al. River water quality prediction model based on physically constrained GRU neural network[J]. Journal of Earth Information Science, 2023, 25(1): 102-114.
- [18] YANG Z, ZOU L, XIA J, et al. Inner dynamic detection and prediction of water quality based on CEEMDAN and GA-SVM models [J]. Remote Sensing, 2022, 14(7): 1714.
- [19] 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 ZH Y, SHEN SH J, et al. Research on wastewater quality prediction based on PSO-LSTM-SATN Model [J]. Industrial Water Treatment 2025, 45(6): 159-166.
- [20] 牛景辉. 基于 GWO-XGBoost 的工业污水水质关键数据预测算法[J]. 工业水处理, 2024, 44(1): 184-190.
- NIU J H. A key data prediction algorithm for industrial wastewater quality based on GWO-XGBoost [J]. Industrial Water Treatment, 2024, 44(1): 184-190.
- [21] AHMED A N, OTHMAN F B, AFAN H A, et al. Machine learning methods for better water quality prediction [J]. Journal of Hydrology, 2019, 578: 124084.
- [22] 张梦迪, 徐庆, 刘振鸿, 等. 基于动态滑动窗口 BP 神经网络的水质时间序列预测[J]. 环境工程技术学报, 2022, 12(3): 809-815.
- ZHANG M D, XU Q, LIU ZH H, et al. Water quality time series prediction based on dynamic sliding window BP neural network [J]. Journal of Environmental Engineering Technology, 2022, 12(3): 809-815.
- [23] 秦文虎, 陈溪莹. 基于长短时记忆网络的水质预测模型研究[J]. 安全与环境学报, 2020, 20(1): 328-334.
- QIN W H, CHEN X Y. Research on water quality prediction model based on long short term memory network[J]. Journal of Safety and Environment, 2020, 20(1): 328-334.
- [24] 欧阳鹏, 邝先验, 叶景贞. 基于改进天牛群算法和三次样条插值的路径规划[J]. 电子测量与仪器学报, 2025, 39(7): 259-268.
- OU Y P, KUANG X Y, YE J ZH. Path planning based on improved beetle swarm algorithm and cubic spline interpolation[J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(7): 259-268.
- [25] 熊川羽, 廖晓红, 何诗英, 等. 使用快速傅里叶变换优化周期参数的 EMD-FFT-SARIMA 光伏发电预测模型[J]. 强激光与粒子束, 2024, 36(8): 121-127.
- XIONG CH Y, LIAO X H, HE SH Y, et al. An EMD-FFT-SARIMA photovoltaic power generation forecasting model optimized using periodic parameters via the fast fourier transform[J]. Intense Laser and Particle Beam, 2024, 36(8): 121-127.
- [26] CHOPRA N, ANSARI M M. Golden jackal optimization: A novel nature-inspired optimizer for engineering applications [J]. Expert Systems with Applications, 2022, 198: 116924.
- [27] 谢国民, 王润良. 基于改进金豺算法的短期负荷预测[J]. 电力系统及其自动化学报, 2024, 36(3): 65-74.
- XIE G M, WANG R L. Short-term load forecasting based on an improved golden jackal algorithm[J]. Journal of Power Systems and Automation, 2024, 36(3): 65-74.
- [28] 张彼德, 陈广, 廖其龙, 等. 基于 IGJO-DHKELM 的光伏阵列故障诊断[J]. 电子测量与仪器学报, 2024, 38(11): 79-89.
- ZHANG B D, CHEN G, LIAO Q L, et al. Fault diagnosis of photovoltaic array based on IGJO-DHKELM[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(11): 79-89.
- [29] SHI J, NI L, LIU J, et al. Water quality prediction at large-Scale: Systematic assessment of decomposition-deep learning models [J]. Journal of Cleaner Production, 2025, 519: 145995.

作者简介



陈静, 2014 年于上海大学获得博士学位, 现为安徽理工大学副教授, 主要研究方向为数据挖掘与数据融合、深度学习与信息处理。

E-mail: jchen@aust.edu.cn

Chen Jing received his Ph. D. degree from Shanghai University in 2014. Now he is an associated professor in Anhui University of Science and Technology. His main research interests include data mining and information fusion, deep learning and information processing.



李政委 (通信作者), 2024 年于安徽理工大学获得学士学位, 现为安徽理工大学硕士研究生, 主要研究方向为深度学习与信息处理。

E-mail: yue09999duan@163.com

Li Zhengwei (Corresponding author) received his B. Sc. degree from Anhui University of Science and Technology in 2024. Now he is an M. Sc. candidate at Anhui University of Science and Technology. His main research interests include deep learning and information processing.