

文章编号: 1674-5566(2026)01-0105-14

DOI: 10.12024/jso.20241204730

基于FC-TCN-GRU模型的凡纳滨对虾养殖水中氨氮和化学需氧量的预测

王智华¹, 吴昊², 周英娴¹, 李桂娟¹, 江敏^{1,3}

(1. 上海海洋大学 海洋科学与生态环境学院, 上海 201306; 2. 上海海洋大学 水产与生命学院, 上海 201306; 3. 上海海洋大学 水域环境生态上海高校工程研究中心, 上海 201306)

摘要: 基于2014—2018年及2021—2024年某水产养殖合作社凡纳滨对虾(*Litopenaeus vannamei*)养殖池塘的水质检测数据,选取总氮(TN)、总磷(TP)、活性磷(AP)、硝态氮(NO_3^- -N)、亚硝态氮(NO_2^- -N)、氨氮(TAN)、化学需氧量(COD)、温度(T)和pH等在内的关键水质参数,构建了基于时域卷积网络(Temporal convolutional network, TCN)和门控循环单元(Gate recurrent unit, GRU)的TAN和COD水质预测模型。首先搭建FC-TCN-GRU的混合模型,即采用TCN对数据特征进行提取和降维处理,再将处理后的数据输入GRU模型,最后通过全连接层(Fully connected layers, FC)映射为预测结果。对搭建好的FC-TCN-GRU模型进行评估,其绝对误差(MAE)、均方误差(MSE)和决定系数(R^2)在对TAN预测中分别为0.255、0.089和0.861;在对COD的预测中分别为1.750、4.840和0.332。将模型与PCA-LSTM、基本LSTM和基本GRU模型对TAN和COD的预测结果进行比较,结果显示:FC-TCN-GRU模型对TAN和COD指标的预测精度优于其他3种模型,在TAN预测中表现出色,但对COD的预测效果尚有提升空间。

关键词: 凡纳滨对虾; 水质预测; 全连接层; 门控循环单元; 时域卷积网络

中图分类号: S 912 **文献标志码:** A

凡纳滨对虾(*Litopenaeus vannamei*)是全球最广泛养殖的对虾品种之一,因其温度和盐度适应范围广、生长速度快、肉质鲜美而备受青睐^[1]。氨氮(TAN)和化学需氧量(COD)等水质指标是影响对虾生长和诱发对虾暴发疾病的关键因素。凡纳滨对虾对TAN的耐受性极低,即使TAN浓度未达致死水平(0.3~0.5 mg/L),也会影响其生长,导致养殖周期延长和经济损失;COD的升高会导致异养菌大量繁殖,在一定程度上抑制硝化菌的活性,进一步加剧TAN的积累。对TAN和COD指标进行快速准确的预测,可对水质恶化提前做出预警,降低养殖风险^[2]。

当前,关于水质预测模型的研究主要聚焦于两类神经网络:前馈神经网络和循环神经网络。CAO等^[3]和REN等^[4]利用前馈神经网络分别对一般养殖水体和鱼菜共生系统中的溶解氧指标

进行了预测,前者运用了正则化极限学习机与经验模态分解进行组合预测,后者则运用遗传算法优化后的模糊神经网络进行预测,提高了模型的泛化能力和预测精度。

养殖环境除了受到自然环境、气候的影响外,还受到养殖品种、养殖管理等的人为干扰,水质具有复杂性和不确定性,要求预测模型具有较高的准确性和适应性,能够处理非线性、时变性和多变量的数据,并能够适应环境变化。利用前馈神经网络构建的模型不能捕捉数据之间的时间相关性,而以长短期记忆网络(Long short-term memory, LSTM)和门控循环单元(Gate recurrent unit, GRU)为代表的循环神经网络模型在时间序列预测任务中则表现出良好的性能,在一定程度上弥补了前馈神经网络模型的缺点^[5],因而被广泛应用于水质预测。王昱文等^[6]通过结合GRU

收稿日期: 2025-01-28 修回日期: 2025-04-30

基金项目: 上海市现代农业产业技术体系建设项目(沪农科产字[2022]第5号)

作者简介: 王智华(2000—),男,硕士研究生,研究方向为渔业水域环境保护研究。E-mail: 1171944789@qq.com

通信作者: 江敏, E-mail: mjiang@shou.edu.cn

版权所有 ©《上海海洋大学学报》编辑部(CC BY-NC-ND 4.0)

Copyright © Editorial Office of Journal of Shanghai Ocean University (CC BY-NC-ND 4.0)

<http://www.shhydxxb.com>

和全连接层(Fully connected layers, FC),提出了1个多元水质参数预测(MWQPP)模型,用于预测自然水体中的4个水质参数(pH、DO、COD和TAN),结果表明该模型相比于传统的单模型,具有更好的鲁棒性。LI等^[7]通过结合RNN和改进后的Dempster-Shafer(D-S),提出了RNNs-DS模型,用于预测自然水体中的4个水质参数(COD、pH、TP和DO),并对比了SVR和BPNN等多个模型。研究表明,RNNs-DS模型在水质预测上具有更高的准确性和稳定性。

冗余的时间序列变量会引入噪声,从而降低多变量时间序列预测模型的精度。YU等^[8]为减少噪声对自然水体的DO和TAN预测结果的干扰,将鄱阳湖水质数据切割为若干个子序列,获得子序列后利用c均值聚类法进行重组,再针对每个聚类序列建立BiGRU预测模型。陈英义等^[9]为解决对杂交鮰鱼养殖池塘中溶解氧预测时出现的数据冗余问题,对数据采用小波变换和卷积网络联用的方法进行降噪。CAO等^[10]在螃蟹养殖池塘的溶解氧预测研究中,在GRU模型中引入了注意力机制,使模型更好地处理长序列场景。

本研究提出构建FC-TCN-GRU水质预测混合模型,将时域卷积模型(Temporal convolutional network, TCN)的数据特征提取能力和GRU的非线性拟合能力相结合,再通过FC将高维特征映射为最终预测结果,以期对养殖水体的TAN和COD变化进行预测。

1 材料与方法

1.1 数据来源

实验数据源自课题组对上海市奉贤区思阳水产养殖合作社2014—2018年和2021—2024年的常规水质监测,共计832组。在养殖期间,每5天对3个凡纳滨对虾露天池塘直接测量水温T,现场用pH计和溶氧仪测定pH、DO;采集水样带回实验室后分别用碱性高锰酸钾法(HY003.4—91)、碱性过硫酸钾消解紫外分光光度法(GB 11894—89)、纳氏试剂分光光度法(GB7479—87)、分子吸收分光光度法(GB7493—87)、紫外分光光度法(HJ/T 346—2007)和钼酸铵分光光度法(GB 11893—89)测定COD、TN、TAN、NO₂⁻-N、NO₃⁻-N、TP和AP等指标。

1.2 数据预处理

1.2.1 数据正态检验

Shapiro-Wilk检验基于数据的顺序统计量(排序后的样本值)与理论正态分布期望值的相关性。其原理是:若数据服从正态分布,则样本顺序统计量与理论期望值应高度相关。检验统计量W的计算公式为

$$W = \frac{\left[\sum_{i=1}^n a_i x_i \right]^2}{\sum_{i=1}^n (x_i - \bar{x})} \quad (1)$$

式中: x_i 为排序后的样本数据; \bar{x} 为样本数据的平均值; a_i 为预先计算的系数; n 为样本量。

1.2.2 异常数据和缺失值的处理

由1.2.1节可知,本研究使用的数据集属于非正态分布数据。因此,采用马氏距离(Mahalanobis distance)对水质数据进行识别与筛选,以剔除异常数据:

$$D = \sqrt{(X - \mu)^T S^{-1} (X - \mu)} \quad (2)$$

式中: X 为单个数据点的特征向量; μ 为均值向量; T 为矩阵的转置; S^{-1} 为协方差矩阵的逆矩阵。

对于时间间隔较小的缺失数据,通过计算缺失数据点前后5日同一时刻的数据均值来完成数据填补;对于时间跨度较大的缺失数据,采用1种基于气象条件相似性的插补方法,即选取与缺失数据点在天气类型上相似的时间段,利用这些时段的数据进行缺失值的补全^[11]。

1.2.2 归一化处理

归一化能够提高梯度下降求最优解的速度,也能在一定程度上提高模型的精度和泛化能力^[12]。本研究采用的归一化方法是离差标准化(Min-Max normalization),该方法会将特征值映射到[0,1]之间,适用于样本数据比较集中的情况。

$$X_{\text{norm}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \quad (3)$$

式中: X_{norm} 为归一化后的值; X 表示原始数据; X_{max} 和 X_{min} 分别为数据集中的最大值和最小值。

1.2.3 数据集划分

数据清洗后,将数据集的前80%划分为训练集^[13],即2014—2018年采集的数据;后20%为测试集,即2021—2024年采集的数据。

1.3 基本模型

1.3.1 门控循环单元GRU

GRU的核心结构包括2个主要的门:更新门(Update gate)和重置门(Reset gate),见图1。更新门的值(z_t)在0到1之间,值越接近1,表示对先前记忆的保留程度越高,计算公式为

$$z_t = \sigma(W_z^* [h_{t-1}, x_t]) \quad (4)$$

式中: σ 为 sigmoid 函数; W_z 为权重矩阵; h_{t-1} 为前一时刻的隐藏状态; x_t 为当前时刻的输入; $*$ 为矩阵乘法。

阵乘法。

重置门的值(r_t)也在0到1之间,值越接近0,表示对之前信息的遗忘程度越高,计算公式为

$$r_t = \sigma(W_r * [h_{t-1}, x_t]) \quad (5)$$

最后,GRU根据更新门和重置门的值更新当前时刻的隐藏状态:

$$\tilde{h}_t = \tanh(W_r * [r_t \cdot h_{t-1}, x_t]) \quad (6)$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (7)$$

式中: \cdot 为内积; \tilde{h}_i 为候选隐藏状态。

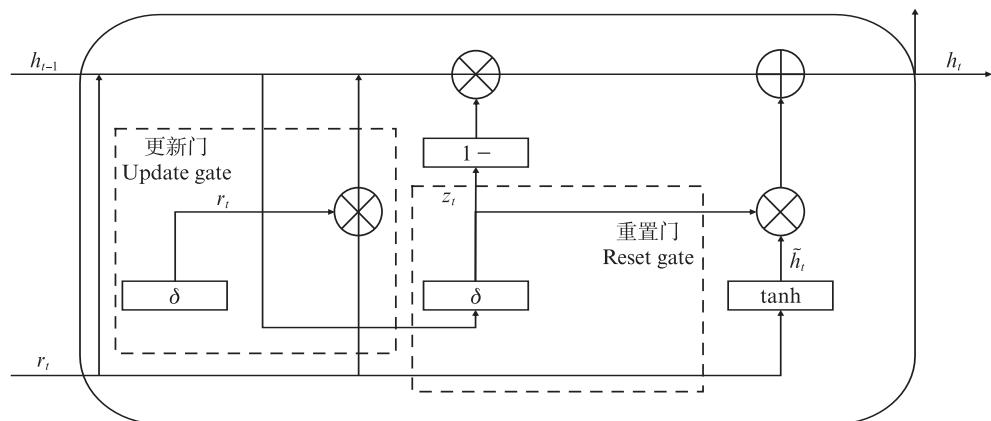


图1 GRU核心结构

1.3.2 时域卷积模型 TCN

TCN 的结构如图 2 所示,包含扩张因果卷积的残差块是 TCN 的主要结构。残差块的设计使得 TCN 能够有效地处理时间序列数据,同时保持了模型的深度和复杂性。TCN 能够增加感受野

的宽度：

$$r = 1 + \sum_{i=1}^{n-1} 2 \cdot (k-1) \cdot b \quad (8)$$

式中: r 为网络的感受野宽度; b 为扩张基; k 为内核大小; n 为残差块数量。

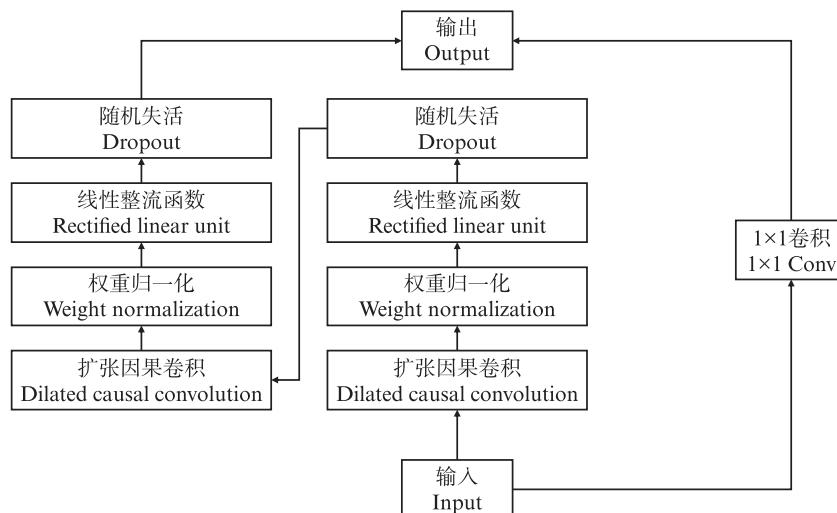


图2 TCN残差块

1.3.3 全连接层 FC

全连接层 FC 的结构如图 3 所示, 输入向量 x 通过 1 个权重矩阵 W 完成线性变换, 再依次进行添加偏置项 b 并通过激活函数 σ 完成非线性变换, 得到输出向量 y ^[14]:

$$y = \sigma(Wx + b) \quad (9)$$

1.4 特征构造

通过滞后效应建模将投喂量构建为时序滞后特征, 进而将其融入 FC-TCN-GRU 混合模型, 该方法能够有效量化人为管理的直接与间接影响。滞后效应建模如下:

$$\text{Lag}_k(X_t) = X_{t-k} \quad (k = 1, 2, 3 \dots T) \quad (10)$$

式中: $\text{Lag}_k(X_t)$ 为时间序列 X 在时间步 t 的 k 阶滞后

后值; t 为当前时间步; X 为投喂量; k 为滞后步长; T 为最大滞后时间。

1.5 FC-TCN-GRU 模型

本研究将 TCN 与 GRU 的优势相结合, 构建 TAN 序列预测模型, 其主要结构如图 4 所示, 模型建立流程如图 5 所示。首先, 将预处理后的数据输入 TCN 层, 通过内核大小为 3, 扩张因子为 1 的残差块实现对输入序列更广泛的感受野覆盖, 并对所得特征进行了深入挖掘和维度缩减; 将 TCN 处理过后的数据输入 GRU, GRU 按 1 个时间步向后处理输入序列, 这种方式使得 GRU 不会随时间步的增加而清除以前信息, 而是会保留相关信息并传递到下一单元; 最后再经由全连接层将高维特征映射为最终输出序列。

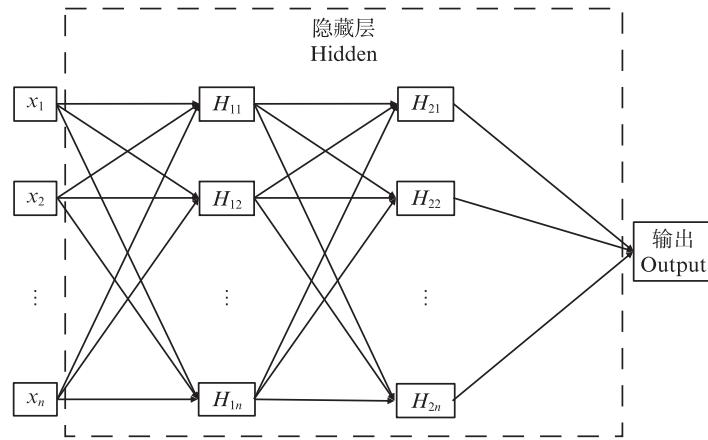


图 3 全连接层
Fig. 3 Fully Connected Layer

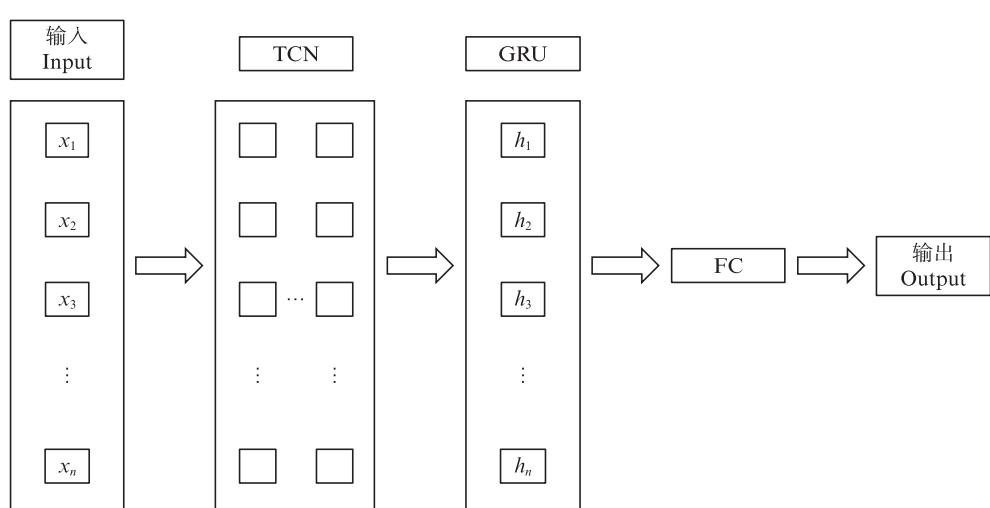


图 4 FC-TCN-GRU 结构
Fig. 4 Structure of FC-TCN-GRU

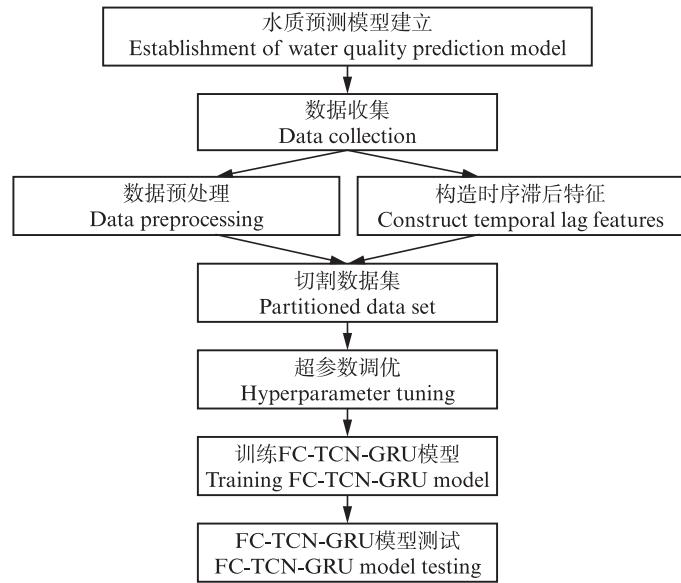


图5 模型建立流程

Fig. 5 Model establishment process

1.6 模型超参数调优

模型超参数(Hyperparameters)用于控制学习过程,帮助确定模型的学习能力和泛化能力;本研究主要调试的超参数是学习率(Learning rate, lr)和隐含层大小(Size)。

使用能够多轮次并行运行的Hyperband算法在一定范围内自动搜索最优超参数组合,算法流程如图6所示。

图6中: R 为单个超参数组合最多分配的资源; μ 为每轮中保留的超参数组合比例; s_{max} 为最大评估轮数; B 为总预算; s 为当前评估轮数; p 为超参数组合数; r 为每个超参数组合实际分配的资源量; i 为迭代次数,用于控制循环; p_i 为当前迭代中的超参数组合数量; r_i 为当前迭代中的资源量。

1.7 PCA-LSTM、基本LSTM、基本GRU模型的建立

参照习文双^[15]等和LEE^[16]等的研究,与FC-TCN-GRU模型采用同一数据集,分别建立PCA-LSTM、基本LSTM和基本GRU模型,旨在与本研究提出的FC-TCN-GRU模型进行性能对比分析。通过这一比较,旨在深入探讨不同模型在处理相同数据集时的效能差异,从而验证FC-TCN-GRU模型的优势和适用性。

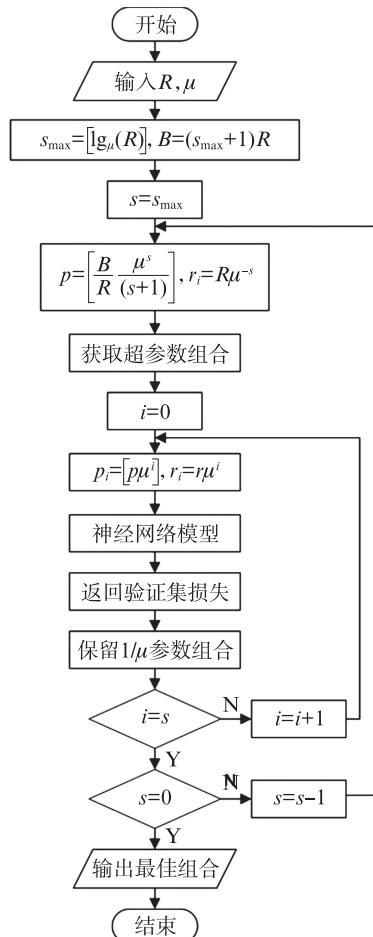


图6 Hyperband算法流程

Fig. 6 Hyperband algorithm process

数据输入PCA-LSTM模型时,首先对数据进行PCA降维,再将降维后特征值大于1的5个主成分(表1)输入LSTM进行训练。因此FC-TCN-GRU、LSTM和GRU模型的输入节点为5,PCA-LSTM为10。

1.8 解释模型SHAP

SHAP通过量化每个特征对预测结果的贡献,对机器学习算法的预测结果进行解释。

$$\varphi_i = \sum_{S \times (F \setminus \{i\})} \frac{|S|! (|F| - |S| - 1)}{|F|!} \times \left[f_{SU\{i\}}(x_{SU\{i\}}) - f_s(x_s) \right] \quad (11)$$

式中: φ_i 为特征*i*的Shapley值;|*F*|为特征总数;*F*\{*i*\}为除特征*i*外所有可能的特征组合的集合;*S*为*F*\{*i*\}的特征集合;*f_s(x_s)*为*S*中特征的模型预测;*f_{SU\{i\}}(x_{SU\{i\}})*−*f_s(x_s)*为特征*i*在已有特征*S*的情况下,对预测的边际贡献。

基于此,解释模型为

$$G(z') = \varphi_0 + \sum_{i=1}^M \varphi_i z' \quad (12)$$

式中: $z' \in \{0, 1\}^M$ 为相应特征是否对模型预测有贡献;*M*为输入特征数; φ_0 为模型常数; φ_i 为特征*i*的Shapley值。

1.9 模型评估指标

本研究采用平均绝对误差(MAE)^[17]、均方误差(MSE)^[18]、决定系数(R^2)^[19]和相对误差(E_r)^[20]对模型的效能进行评估。MAE和MSE常用于评估预测值和真实值之间的误差大小; R^2 能较好地体现模型对数据的拟合优劣程度; E_r 能够帮助判断预测值是否处于合理范围内:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - y'_i| \quad (13)$$

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_i - y'_i)^2 \quad (14)$$

$$R^2 = \frac{\sum_{i=1}^n (y'_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (15)$$

$$E_r = \frac{y_i - y'_i}{y_i} \quad (16)$$

式中:*m*、*n*为样本数;*y_i*为第*i*时刻的真实值;*y'_i*为第*i*时刻的预测值; \bar{y} 为样本平均值。

2 结果

2.1 正态性分布检验

使用SPSSPRO对数据集进行正态检验(表2和图7),结果显示TAN和COD呈现显著性(*P*<0.05),因此可认为TAN与COD均不符合正态分布。

2.2 目标因子与各特征因子的相关性分析

对水质因子进行相关性分析,可排除相关性低的变量,有助于降低网络规模和提高网络性能,从而改善由于冗余变量导致的模型精度下降问题。由图8可知,与TAN呈正相关的指标有:TN、AP、NO₂⁻-N、DO、TP和NO₃⁻-N,呈负相关的指标有:T、pH和COD。相关程度排序:NO₃⁻-N、TP、TN、AP、COD、NO₂⁻-N、pH、T和DO。与TAN相关性最大的5个因子为NO₃⁻-N、TP、TN、AP和COD;同理可得,与COD相关性最大的指标为TN、TP、T、DO和pH。

表1 主成分分析
Tab. 1 Principal component analysis

成分 Component	特征值 Eigen value	初始贡献率 Percentage of variance/%	累积贡献率 Cumulative/%
1	2.098 85	20.988 48	20.988 48
2	1.914 38	19.143 81	40.132 29
3	1.358 56	13.585 56	53.717 86
4	1.118 24	11.182 39	64.900 24
5	1.021 52	10.215 22	75.115 46
6	0.770 23	7.702 33	82.817 80
7	0.655 65	6.556 54	89.374 34
8	0.497 31	4.973 08	94.347 42
9	0.393 26	3.932 63	98.280 05
10	0.171 99	1.719 95	100.000 00

表2 正态性检验结果
Tab. 2 Normality test results

指标 Factor	中位数 Median	平均值 Average value	标准差 Standard deviation	偏度 Skewness	峰度 Kurtosis	Shapiro-Wilk	P
氨氮 COD/ (mg/L)	10.220	10.439	5.429	1.847	10.735	0.888	0
化学需氧量 TAN/ (mg/L)	0.735	1.291	1.565	3.081	11.238	0.635	0

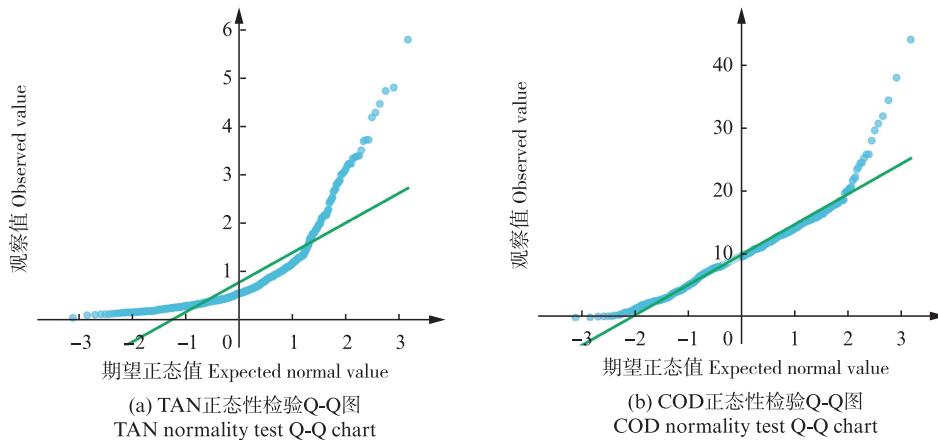


图7 正态分布Q-Q图
Fig. 7 Normal distribution Q-Q diagram

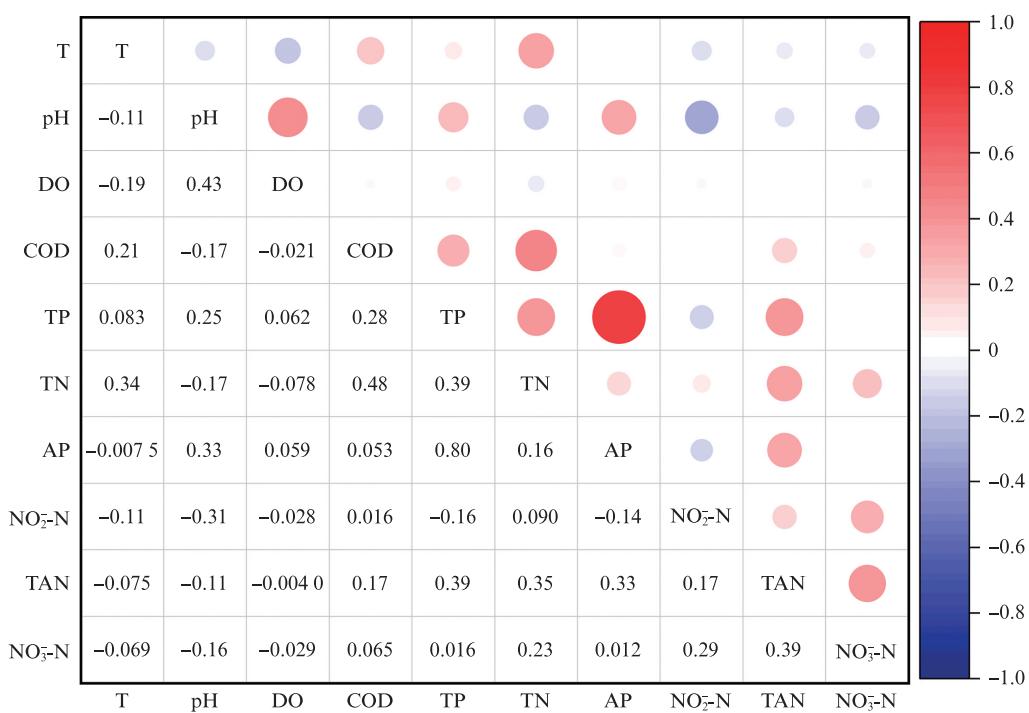


图8 水质因子的相关性分析
Fig. 8 Correlation analysis of water quality factors

2.3 SHAP模型解释结果

基于SHAP模型,计算各指标因子的SHAP值并输出蜂群图(图9),图9a和图9b分别为对TAN和COD的预测任务特征蜂群图。横坐标

SHAP值表示各指标因子对模型输出的贡献,纵坐标均表示特征值。

由图9a可知,对预测TAN影响程度最大的5个因子依次为TN、COD、T、DO和TP,且大部分

指标因子的SHAP值呈现“红左蓝右”的分布态势,表明大多指标因子与TAN呈现负相关。

由图9b可知,对预测COD影响程度最大的5个因子依次为TAN、TN、NO₃-N、DO和AP,部分指标因子的SHAP值呈现“红左蓝右”的分布态势,表明大多指标因子与COD呈现负相关。

SHAP模型解释结果与2.2节的相关性分析结果并不完全相同。这是因为,相关性分析不考虑特征之间的交互作用和多变量影响;而SHAP模型能够反映特征对模型预测的全局或局部贡献,即多变量分析,并且能够综合考虑特征间的交互作用和模型的非线性因素。

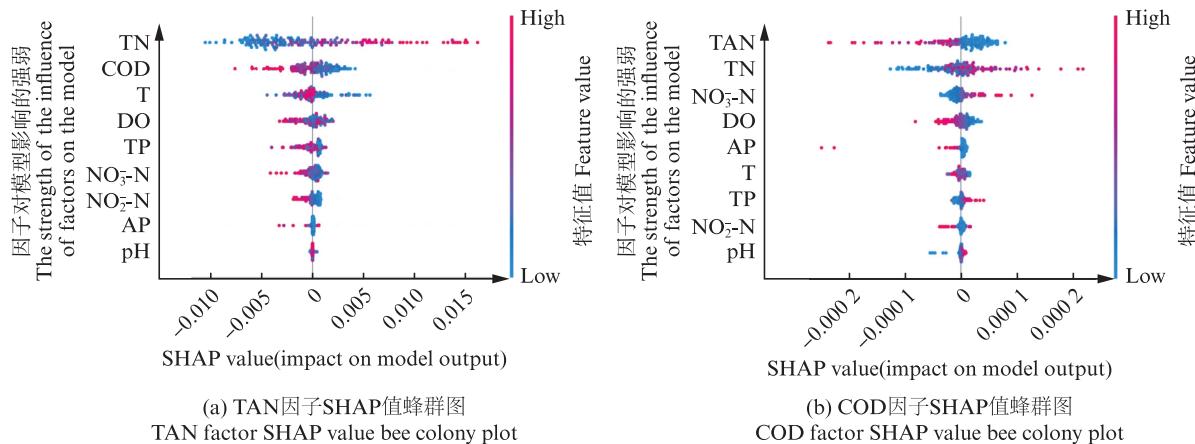


图9 各指标因子SHAP值蜂群图
Fig. 9 Colony plot of SHAP values for each indicator factor

2.3 模型超参数优化

通过Hyperband算法在一定范围内自动搜索

的最优超参数配置列于表3,图10a和图10b分别为TAN和COD预测模型的损失函数图。

表3 最优超参数配置
Tab. 3 Optimal hyperparameter configuration

预测模型 Prediction model	学习率 Learning rate	隐含层大小 Hidden size	损失函数 Loss
TAN	0.000 213 191 026 046 070 40	128	0.007 372 960 913 926 363
COD	0.000 106 102 874 305 149 34	256	0.020 959 311 863 407 493

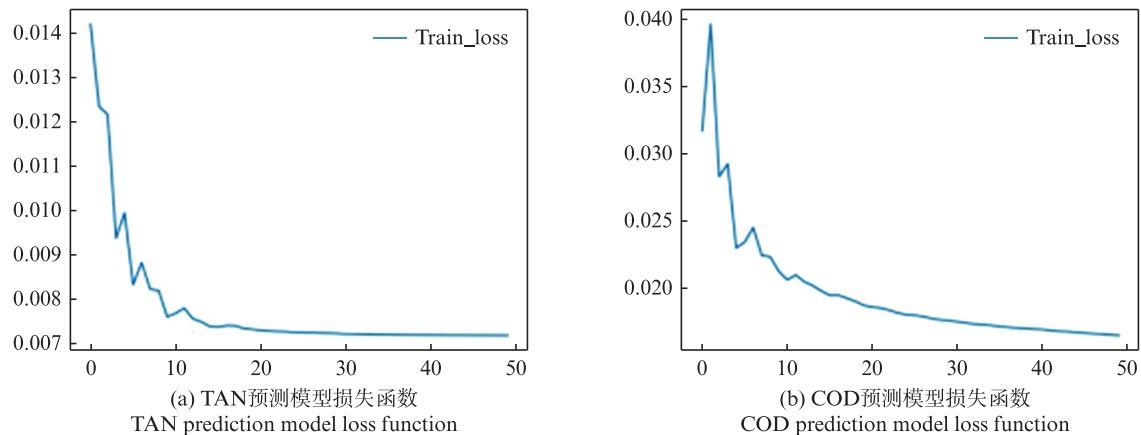


图10 TAN和COD预测模型的LOSS函数图
Fig. 10 LOSS function diagram of TAN and COD prediction models

2.4 模型评估

在统计学和机器学习领域,有多个指标可用

于模型性能的评估。MAE是衡量预测准确性的1个重要指标,当MAE值小于1时,普遍认为模型

表现出色,即预测结果与实际值之间的平均绝对差异较小,模型具有较高的预测精度。MSE是另一个衡量预测准确性的指标,其值越接近0,意味着模型预测的误差平方和越小,即模型预测的准确性越高。 R^2 则用于衡量模型对数据变异性的解释程度,其值越接近1,表明模型对数据的拟合效果越好,解释能力越强;通常认为, R^2 值超过0.7时,表明模型性能优秀,具有较高的预测准确性。当然,这些指标的解释力和预测准确性的判断还需结合具体的应用场景和数据范围来综合评估,以确保模型评估的准确性和适用性。

表4数据显示,TAN预测模型的MAE和MSE均显著低于COD预测模型,而 R^2 则与之相反,表明本研究构建的FC-TCN-GRU模型在TAN的预测精度上显著优于COD。同时,TAN预测模型的MAE(0.255)远小于1,MSE仅为0.089, R^2 (0.861)则大于0.7,符合模型性能优秀的标准。因此,可以认为FC-TCN-GRU模型在TAN指标的预测中表现出色,具有良好的预测准确性和解释能力。

而COD预测模型的MAE值为1.75,超过了1,MSE为4.84, R^2 为0.332,低于0.7。表明所构建的FC-TCN-GRU模型在COD指标的预测任务中表现不佳,预期的预测准确性和解释能力还有待提

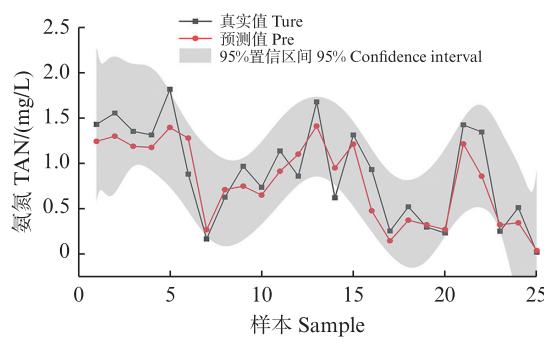
高,需要进一步进行优化以提高其预测性能。

表4 TAN和COD预测模型的各评价指标
Tab. 4 The evaluation indices of the predictive model for TAN and COD

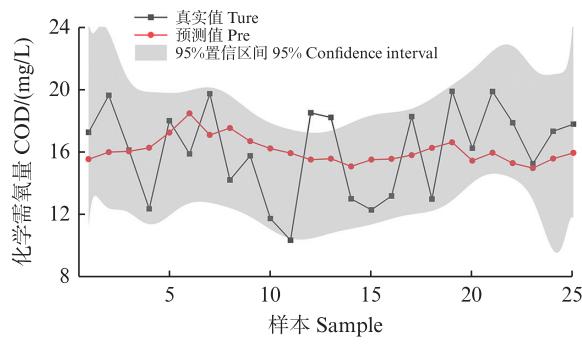
预测指标 Predictive factor	平均绝对 误差 MAE	均方误差 MSE	决定系数 R^2
氨氮 TAN	0.255	0.089	0.861
化学需氧量 COD	1.750	4.840	0.332

2.5 模型预测结果

图11为2021—2024年TAN和COD实测值与预测值的变化趋势。由图11a可知,TAN预测结果中95%的数据点都位于95%置信区间内,可认为预测值和实测值吻合度很高。虽然以MAE、MSE和 R^2 值评价COD的预测模型结果不理想,但图11b中COD预测数据基本都位于95%置信区间内。MAE、MSE和 R^2 值聚焦于评估模型预测值与实际值之间的偏差程度,主要作用是衡量模型预测精度。而置信区间则侧重于关注预测结果的不确定性和可信度。评价指标表现不佳,表示模型的预测误差较大;而预测值均落于95%置信区间之中,则表示模型在整体趋势和预测范围仍是可信的。



(a) TAN测试集样本 Test set samples of TAN



(b) COD测试集样本 Test set samples of COD

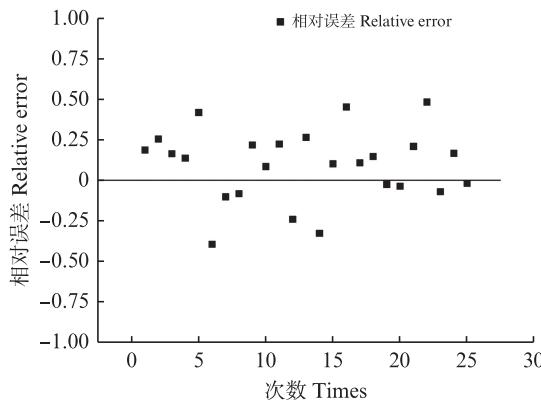
图11 预测值95%置信区间
Fig. 11 95% confidence interval for forecasts

通过公式(12)计算并将误差可视化,结果如图12a所示。TAN的整体相对误差都较小,大部分数据都在0.25之间,虽然个别点位仍有一定误差,但总体上趋于一致。仅在第6次与第22次的相对误差超出0.5,其中第6次的实测值 $y_6=0.88$ mg/L,预测值为 $y_6'=1.27$ mg/L;第22次的实测值 $y_{22}=1.34$ mg/L,预测值为 $y_{22}'=0.86$ mg/L。回顾这2次数据采集时的情况可知:第6次数据采样日

期为2022年11月30日,上海受冷空气影响出现剧烈降温;第22次数据的采样日期为2024年9月16日,在采样前一天,台风“贝碧嘉”登陆上海。两次极端天气对养殖水质产生了影响,导致预测结果与真实值存在偏差。

由于COD数据数量级大于TAN,并且COD数据在相对误差图中的离散程度较大,如图12b所示。从第6次数据开始,相对误差超出0.5,后

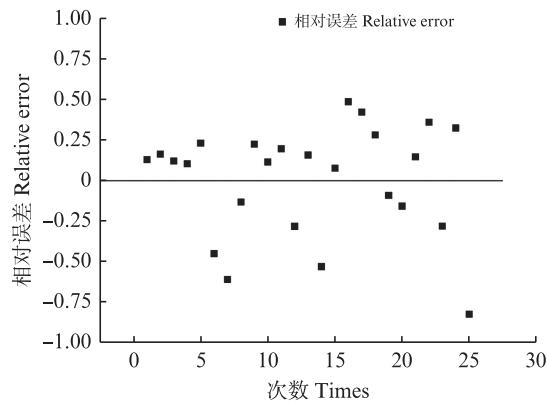
续数据相对误差逐步变大,出现最大的相对误差超过0.75。因此本研究认为,FC-TCN-GRU模型在对COD指标的预测性能方面,仍存在进一步优化和提升的潜力。未来研究可以探索更复杂的网络架构、更精细的特征工程或更先进的优化算法,以期实现对COD指标更为精确的预测。



(a) TAN预测相对误差 TAN prediction relative error

2.6 多种模型对比

将同一数据集分别输入FC-TCN-GRU与PCA-LSTM、LSTM、GRU模型,并对训练结果进行比较。4种模型的超参数及输入节点如表5所示。



(b) COD预测相对误差 COD prediction relative error

图12 预测值的相对误差

Fig. 12 The relative error of the predicted values

表5 4种模型的输入节点及超参数优化结果

Tab. 5 The evaluation indices of the predictive models for TAN and COD

参数 Parameter	LSTM		GRU		PCA-LSTM		FC-TCN-GRU	
	TAN	COD	TAN	COD	TAN	COD	TAN	COD
输入 Input	5	5	5	5	10	10	5	5
学习率 lr	0.001	0.000 1	0.001	0.002	0.000 9	0.000 5	0.000 2	0.000 1
隐含层 size	64	256	128	64	64	64	128	256
损失值 LOSS	0.007	0.021	0.006	0.006	0.007	0.02	0.007	0.02

模型建立完成后,对2021—2024年某养殖场的TAN进行预测,采用MAE、MSE和 R^2 的值对各模型预测能力进行对比。

图13显示,FC-TCN-GRU模型在TAN和COD预测精度上优于其余3种模型,其中GRU模型表现优于LSTM,原因是GRU更适合处理小数据集。此外,FC-TCN-GRU模型显示出更强的鲁棒性、泛化能力和更高的预测精度,更适合应用于凡纳滨对虾养殖水体的预测。

3 讨论

3.1 FC-TCN-GRU模型的优势

目前在水产养殖领域,LSTM是预测水质变

化的重要模型,因为其能够很好地解决循环神经网络因长期依赖而带来的梯度消失与梯度爆炸问题,但由于LSTM有3个不同的门,其所需要的参数较多,导致了实现的难度较大^[21]。GRU模型是在LSTM基础上进行优化的模型,在超参数全部调优的情况下,二者的性能相当,但GRU的结构更为简单,需求的训练样本比LSTM更少,在极大程度上降低了实现的难度与成本。本研究提出的FC-TCN-GRU模型,首先使用TCN对数据特征进行降维处理的同时能够更好地捕捉局部的依赖关系,进而提高GRU的预测精度。

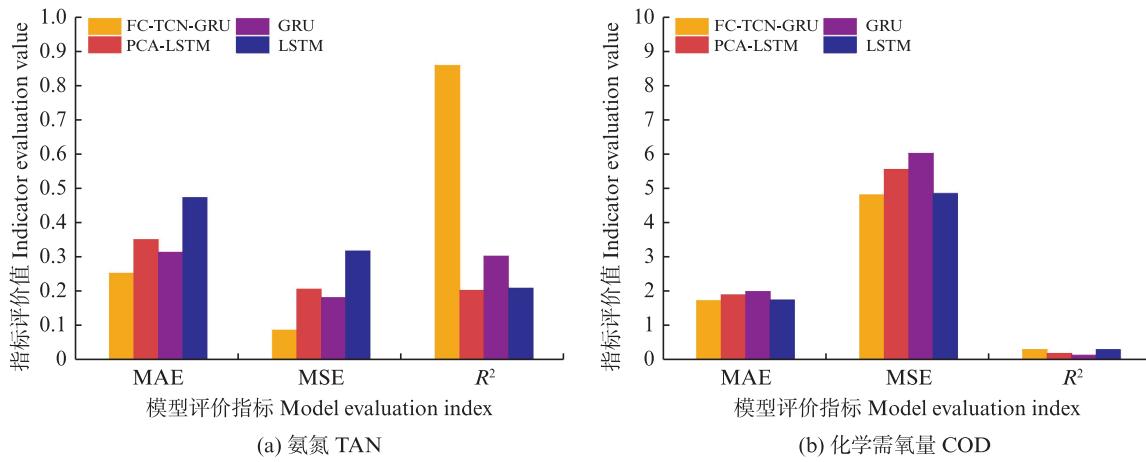


图13 4种模型预测结果精度分析

Fig. 13 Accuracy analysis of prediction results of four models

3.2 模型预测效果的制约因素分析

水质指标预测模型是池塘养殖中不可或缺的工具,通过精准预测水质变化,为科学管理和健康养殖提供了强有力的数据支持。本研究提出的FC-TCN-GRU模型虽在2021—2024年某养殖场的TAN指标预测中的效果良好,而COD指标的预测并不理想,仍有个别时刻预测结果偏差较大,原因可能有以下两点:其一,台风、暴雨等极端天气事件会对养殖池塘的水质产生显著影响。这些突发性的自然现象可能会导致水质参数的剧烈波动,而这种波动在历史数据中可能不常见,因此模型难以准确预测。其二,在数据预处理阶段,本研究未能充分考虑采样当天的天气情况,仅依据数据是否出现剧烈波动来识别异常值。这种方法可能导致模型无法捕捉到特定天气条件下的数据特征,进而影响模型在类似情况下的预测准确性。

值得注意的是,养殖水体相较于天然水体,一个显著的区别在于人为干预的存在与否。本研究构建的模型在纳入大量数据进行训练的同时,特别考虑了人为干预这一外部因素,从而使得模型在预测过程中能够综合评估人为干预的影响。这种综合考量显著提升了模型对TAN指标的预测准确性,但对COD指标的预测精度提升并不明显,而导致COD预测模型效果不理想的原因主要是:第一,COD指标的水质特征具有高度复杂性,这要求模型必须具备足够的能力来捕捉和模拟这些复杂的变化,而现有数据样本可能不足以支持模型学习到COD变化趋势的全貌,导致模型在泛化能力上存在局限。第二,

模型复杂度的不足也是一个关键因素,可能限制了模型捕捉COD数据特征的完整性和深度。这可能是由于模型结构设计未能充分考虑COD数据的内在特性,或者是模型参数设置未能优化至最佳状态。第三,未将足够的人为干预因素进行量化,且构造的投喂量-时序滞后特征并不能充分地将投喂量指标量化,导致模型预测精度下降。第四,TCN固定感受野对长周期波动的适应性不足,本模型中TCN的感受野设计更适用于捕捉局部时序特征,但对跨度数周至数月的长周期波动可能缺乏拟合能力。第五,GRU对非线性交互与多尺度特征的整合能力有限,其隐藏状态更新过程是线性变换与非线性激活的叠加,对高度非线性的水质参数交互作用可能拟合不足。

3.3 模型在养殖水质预测中面临的挑战与局限性

目前,FC-TCN-GRU预测模型仍然在多个方面存在局限性,其中最亟需解决的是黑箱模型可解释性较差的问题。虽然本研究使用了SHAP解释模型量化了各个特征的贡献度,但是模型中更深层的非线性交互机制仍然缺乏较为直观的解释。尤其在误差较大的COD预测任务中,模型的决策逻辑的透明度不足可能影响应用推广。同时,与天然水体相比,养殖水体会受到人为干扰,其DO、pH、营养元素和有机物等水质指标之间的相关性更加复杂。例如养殖池塘中常开启增氧机以提高DO水平,同时还带来了对水的扰动,进而导致与氧化还原反应有关的各水质参数之间的关系变得更加复杂,增加了模型准确捕捉关系的难度,使模型预测精度有所下降^[22-23]。同时,由

于养殖池塘中的生物量远超自然水体,对虾等养殖生物的活动会高频率地引发生物扰动现象,而生物扰动会加速底泥悬浮,释放更多的耗氧物质,导致捕捉DO的变化规律需考虑高频动态变化^[23-24],增加预测难度。

3.4 模型优化建议

在模型优化方面,已有研究指出,基于多光谱信息融合的水质COD预测模型能够提高预测精度和泛化能力,这表明结合多种信息源对于提高模型性能具有积极作用^[25]。此外,卷积神经网络(Convolutional neural network, CNN)与紫外-可见光谱分析法相结合的模型在水质COD预测方面也显示出了较高的预测精度^[26]。这些研究结果为提高COD预测模型的性能提供了有益的参考。因此,为了提升COD预测模型的效果,需要从数据收集、模型设计、训练策略等多个角度进行综合考虑和优化。

综上所述,对虾养殖水质预测过程中可对模型进行以下改进:第一,建议纳入极端气象事件的数据并整合多种人工干预因素,以增强模型对异常气候条件下水质变化的预测能力及提高模型在实际应用中的准确性和可靠性。第二,将人工干预作为一个特征因子与时间关联起来,可以使模型能够在预测时根据不同的时间考虑不同的人工干预因素,综合考量后提升模型预测的精度。第三,针对数据预处理阶段中的异常值剔除,建议采取更为审慎的态度。对于可能蕴含关键信息的异常值,应考虑予以保留,或探索其他先进的数据处理技术,以充分挖掘其潜在价值,避免因简单剔除而损失重要数据特征。第四,在处理数据中的缺失值时,建议采用多元化的插补策略,并对比评估不同方法对模型性能的具体影响,这样可以更好地保证数据完整性,同时减少因插补引入的偏差,进一步提升模型的泛化能力和预测精度。第五,引入自适应感受野机制(如注意力机制等),通过在TCN中嵌入注意力模块,动态调整不同时间步的权重,以增强对长周期波动的敏感性。第六,在TCN与GRU间加入交叉特征交互层(如自注意力或外部记忆单元等),捕捉水质参数间的非线性关系。第七,通过引入更丰富的数据样本、设计更复杂的模型结构、采用有效的正则化和数据增强技术,以及借鉴已有的研究成果,可有望提高模型在COD预测任务上的

性能。第八,优化养殖管理,通过精准投喂和控制养殖密度,减少对虾的过度活动,降低底泥扰动的风险。

作者声明本文无利益冲突。

参考文献:

- [1] HUANG H H, LIAO H M, LEI Y J, et al. Effects of different carbon sources on growth performance of *Litopenaeus vannamei* and water quality in the biofloc system in low salinity [J]. Aquaculture, 2022, 546: 737239.
- [2] AKRATOS C S, PAPASPYROS J N E, TSIHRINTZIS V A. Total nitrogen and ammonia removal prediction in horizontal subsurface flow constructed wetlands: use of artificial neural networks and development of a design equation [J]. Bioresource Technology, 2009, 100 (2) : 586-596.
- [3] CAO W J, HUAN J, LIU C, et al. A combined model of dissolved oxygen prediction in the pond based on multiple-factor analysis and multi-scale feature extraction [J]. Aquacultural Engineering, 2019, 84: 50-59.
- [4] REN Q, ZHANG L, WEI Y G, et al. A method for predicting dissolved oxygen in aquaculture water in an aquaponics system [J]. Computers and Electronics in Agriculture, 2018, 151: 384-391.
- [5] SHU T X, CHEN J H, BHARGAVA V K, et al. An energy-efficient dual prediction scheme using LMS filter and LSTM in wireless sensor networks for environment monitoring[J]. IEEE Internet of Things Journal, 2019, 6 (4): 6736-6747.
- [6] 王昱文, 杜震洪, 戴震, 等. 基于复合神经网络的多元水质指标预测模型[J]. 浙江大学学报(理学版), 2022, 49(3): 354-362, 375.
- [7] WANG Y W, DU Z H, DAI Z, et al. Multivariate water quality parameter prediction model based on hybrid neural network [J]. Journal of Zhejiang University (Science Edition), 2022, 49(3): 354-362, 375.
- [8] LI L, JIANG P, XU H, et al. Water quality prediction based on recurrent neural network and improved evidence theory: a case study of Qiantang River, China [J]. Environmental Science and Pollution Research, 2019, 26 (19): 19879-19896.
- [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] 陈义英, 方晓敏, 梅思远, 等. 基于WT-CNN-LSTM的溶解氧含量预测模型[J]. 农业机械学报, 2020, 51 (10): 284-291.

- CHEN Y Y, FANG X M, MEI S Y, et al. Dissolved oxygen prediction model based on WT-CNN-LSTM [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(10): 284-291.
- [10] CAO X K, REN N, TIAN G L, et al. A three-dimensional prediction method of dissolved oxygen in pond culture based on attention-GRU-GBRT [J]. Computers and Electronics in Agriculture, 2021, 181: 105955.
- [11] 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 Z H, et al. Prediction of water quality time series based on the dynamic sliding window BP neural network model [J]. Journal of Environmental Engineering Technology, 2022, 12 (3) : 809-815.
- [12] 秦子轩,张晓东,白广芝,等.基于多尺度特征融合的航空发动机剩余寿命预测[J].航空发动机,2024,50(4):114-120.
- QIN Z X, ZHANG X D, BAI G Z, et al. Aeroengine remaining useful life prediction based on multi-scale feature fusion[J]. Aeroengine, 2024, 50(4): 114-120.
- [13] 付博,杨永锋,刘向真,等.数据集划分及预处理方法对烟叶化学成分近红外定量模型的影响[J/OL].河南农业大学学报. <https://doi.org/10.16445/j.cnki.1000-2340.20241009.001>.
- FU B, YANG Y F, LIU X Z, et al. Influence of dataset partitioning and spectral pre-processing methods on the near infrared quantitative model of chemical ingredients in tobacco leaves [J/OL]. Journal of Henan Agricultural University. <https://doi.org/10.16445/j.cnki.1000-2340.20241009.001>.
- [14] 李健. SDN环境下的DDoS攻击检测溯源技术研究[D].成都:电子科技大学,2024.
- LI J. Research on DDoS attack detection and traceability technology in SDN environment [D]. Chengdu: University of Electronic Science and technology of China, 2024.
- [15] 习文双,江敏,吴昊,等.基于PCA-LSTM神经网络的凡纳滨对虾养殖水质预测[J].上海海洋大学学报,2023,32(1):108-117.
- XI W S, JIANG M, WU H, et al. Prediction of water quality in *Litopenaeus vannamei* aquaculture ponds based on the PCA-LSTM neural network model [J]. Journal of Shanghai Ocean University, 2023, 32(1): 108-117.
- [16] LEE J H, LEE J Y, LEE M H, et al. Development of a short-term water quality prediction model for urban rivers using real-time water quality data [J]. Water Supply, 2022, 22(4): 4082-4097.
- [17] WILLMOTT C J, MATSUURA K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance [J]. Climate Research, 2005, 30: 79-82.
- [18] WANG Z, BOVIK A C. Mean squared error: love it or leave it? A new look at signal fidelity measures [J]. IEEE Signal Processing Magazine, 2009, 26 (1) : 98-117.
- [19] CAMERON A C, WINDMEIJER F A G. An *R*-squared measure of goodness of fit for some common nonlinear regression models[J]. Journal of Econometrics, 1997, 77 (2): 329-342.
- [20] SURESH A S, PARTHASARATHY S S. Probabilistic principal component analysis and long short-term memory classifier for automatic detection of Alzheimer's disease using MRI brain images[J]. Journal of the Institution of Engineers (India): Series B, 2021, 102(4): 807-818.
- [21] 李道亮,刘畅.人工智能在水产养殖中研究应用分析与未来展望[J].智慧农业,2020,2(3):1-20.
- LI D L, LIU C. Recent advances and future outlook for artificial intelligence in aquaculture [J]. Smart Agriculture, 2020, 2(3): 1-20.
- [22] 于辉辉.基于机器学习的池塘养殖水质关键因子预测方法研究[D].北京:中国农业大学,2018.
- YU H H. Prediction research of water quality in aquaculture based on machine learning method [D]. Beijing: China Agricultural University, 2018.
- [23] 周劲风,温琰茂,李耀初.珠三角养殖池塘水质模型建立及其应用研究[J].中山大学学报(自然科学版),2010,49(1): 119-124.
- ZHOU J F, WEN Y M, LI Y C. Establishment of water quality model of farming ponds in the pearl river delta and its application [J]. Acta Scientiarum Naturalium Universitatis Sunyatseni, 2010, 49(1): 119-124.
- [24] 张敬旺,谢骏,龚望宝,等.池塘底泥耗氧与环境因子的关系[C]/渔业科技创新与发展方式转变——2011年中国水产学会学术年会论文摘要集.厦门:中国水产学会,2011: 313.
- ZHANG J W, XIE J, GONG W B, et al. The pond's sediment oxygen demand and the relationship with environmental factors [C]/Fishery Science and Technology Innovation and the Transformation of Development Patterns: Proceedings of the 2011 Annual Conference of the Chinese Society of Fisheries. Xiamen: China Society of Fisheries, 2011: 313.
- [25] 邱远见.基于多光谱信息融合的水质COD预测模型研究[D].秦皇岛:燕山大学,2019.
- QIU YUANJIAN. Study on COD predicted model of water quality based on multi-spectral information fusion [D]. Qinhuangdao: Yanshan University, 2019.
- [26] 贾文坤,张恒之,马洁,等.基于紫外-可见光谱与深度学习CNN算法的水质COD预测模型研究(英文)[J].

光谱学与光谱分析, 2020, 40(9): 2981-2988.

JIA W S, ZHANG H Z, MA J, et al. Study on the predication modeling of COD for water based on UV-VIS

spectroscopy and CNN algorithm of deep learning [J].

Spectroscopy and Spectral Analysis, 2020, 40(9): 2981-2988.

Prediction of ammonia nitrogen and chemical oxygen demand in *Litopenaeus vannamei* aquaculture ponds based on the FC-TCN-GRU model

WANG Zhihua¹, WU Hao², ZHOU Yingxian¹, LI Guijuan¹, JIANG Min^{1,3}

(1. College of Oceanography and Ecological Science, Shanghai Ocean University, Shanghai 201306, China; 2. College of Fisheries and Life Sciences, Shanghai Ocean University, Shanghai 201306, China; 3. Shanghai Higher Education Engineering Research Center for Water Environment Ecology, Shanghai Ocean University, Shanghai 201306, China)

Abstract: Based on water quality data from *Litopenaeus vannamei* aquaculture ponds in the same aquaculture farm during 2014–2018 and 2021–2024, this study selected key water quality parameters including total nitrogen (TN), total phosphorus (TP), active phosphorus (AP), nitrate nitrogen (NO₃⁻-N), nitrite nitrogen (NO₂⁻-N), total ammonia nitrogen (TAN), chemical oxygen demand (COD), temperature (T), and pH values to develop water quality prediction models for TAN and COD using temporal convolutional network (TCN) and gated recurrent unit (GRU). A hybrid FC-TCN-GRU model architecture was constructed, which employed TCN for feature extraction and dimensionality reduction of data features, fed the processed data into GRU, and finally mapped the results through fully connected layers (FC) to generate predictions. Mean absolute error (MAE), mean squared error (MSE), and coefficient of determination (R^2) values of the FC-TCN-GRU model for TAN prediction were 0.255, 0.089 and 0.861, respectively, while achieved 1.750, 4.840 and 0.332 for COD prediction. Compared with PCA-LSTM, basic LSTM and basic GRU models, the FC-TCN-GRU model showed better predictive accuracy for both TAN and COD prediction. The model performs superior in TAN prediction, but it still needs improvement in COD prediction.

Key words: *Litopenaeus vannamei*; water quality prediction; fully connected layers; gate recurrent unit; temporal convolutional network