

基于ARIMA-BiGRU模型的水质分类预测

王佳仪, 王逸蓓, 张越溪

燕山大学理学院, 河北 秦皇岛

收稿日期: 2025年2月11日; 录用日期: 2025年3月4日; 发布日期: 2025年3月17日

摘要

本文聚焦于地表水水质分类预测, 选取2018年至2024年间秦皇岛市戴河口监测点的水质数据, 通过结合传统时间序列模型(ARIMA)和双向门控循环单元模型(BiGRU), 系统评估了未来水质变化的预测性能。研究表明, 两者的组合模型(ARIMA-BiGRU)进一步发挥了协同优势: 先利用ARIMA提取序列的线性成分和周期性特征, 再通过BiGRU对残差中的复杂非线性动态进行深度解析, 从而实现对水质参数多层次时空关联的精准建模。相较于单一模型, ARIMA-BiGRU组合模型在短期和中期预测任务中均展现出更强的鲁棒性, 对水质分类的预测误差显著降低。该模型不仅为水质预测提供了兼顾线性统计规律与深度特征挖掘的融合框架, 其多尺度预测能力还可支持管理者制定从短期污染预警到长期生态治理的差异化决策方案, 对提升水环境管理的科学性与前瞻性具有重要实践价值。

关键词

地表水, 水质预测, 传统时间序列模型, 双向门控循环单元模型

Water Quality Classification Prediction Based on ARIMA-BiGRU Model

Jiayi Wang, Yibei Wang, Yuexi Zhang

School of Science of Yanshan University, Qinhuangdao Hebei

Received: Feb. 11th, 2025; accepted: Mar. 4th, 2025; published: Mar. 17th, 2025

Abstract

This study focuses on the classification prediction of surface water quality, utilizing water quality data from the Daihekou monitoring site in Qinhuangdao City between 2018 and 2024. By integrating the traditional time series model (ARIMA) and the bidirectional gated recurrent unit model (BiGRU) into a combined model (ARIMA-BiGRU), the predictive performance for future water quality changes is systematically evaluated. The results demonstrate that the ARIMA model effectively captures linear

trends and seasonal patterns in water quality indicators through its differencing operations and autoregressive mechanisms. Meanwhile, the BiGRU model enhances nonlinear prediction accuracy by employing bidirectional temporal modeling, which integrates both historical and potential future information. The combined ARIMA-BiGRU model further leverages synergistic advantages: it first utilizes ARIMA to extract linear components and periodic features from the sequence, then employs BiGRU to deeply analyze the complex nonlinear dynamics within the residuals, achieving precise modeling of multi-level spatiotemporal correlations in water quality parameters. Compared to individual models, the ARIMA-BiGRU combined model exhibits stronger robustness in short- and medium-term prediction tasks, significantly reducing prediction errors in water quality classification. This model not only provides a hybrid framework that balances linear statistical principles and deep feature extraction for water quality forecasting but also supports differentiated decision-making plans—from short-term pollution alerts to long-term ecological management—through its multi-scale predictive capability. This approach holds significant practical value for enhancing the scientific rigor and forward-looking perspective of water environment management.

Keywords

Surface Water, Water Quality Prediction, ARIMA, BiGRU

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来,随着人口的快速增长和经济高速发展,水资源压力与水环境污染问题变得尤为严峻。水量短缺、水质恶化等挑战不断浮现,为科学应对这些问题,系统性地规划和管理成为关键,准确预测地表水体的污染程度及其趋势是有效管理的重要组成部分。为此,地表水水质预测模型应运而生,为决策者制定有效的污染防治策略提供了重要的数据支持。为了对秦皇岛市地表水水质情况进行更好地分析和预测,以秦皇岛市戴河口为例,对研究区地表水开展研究来进行水质分析和预测。传统水质数学模型依赖大量难以获取的水文和水质参数,这限制了其适用性和准确性。相比之下,机器学习方法能够直接处理实测数据,自动识别复杂模式,减少了对预先定义参数的需求。这种方法不仅提高了预测的精度,还增强了模型的适应性。随着人工智能技术的进步,包括深度学习在内的多种算法为水质预测提供了新的解决方案,使得研究人员可以更准确地评估和应对水质问题。

在水质预测中,考虑到水质数据既具有时间变化的周期性,又受到多种因素的影响而表现出复杂的变化特征,结合经典统计模型与深度学习的优势能够充分发挥 ARIMA 对线性趋势和季节性的捕捉能力,以及 BiGRU 对复杂非线性时序依赖的敏感性。文章以秦皇岛市戴河口点位自动监测站提供的数据为基础,建立了不同预测模型:传统时间序列模型 ARIMA、深度学习模型 BiGRU 以及 ARIMA-BiGRU 组合模型。其中,ARIMA 通过差分运算和自回归移动平均机制,能够有效提取序列中的线性趋势与周期性规律,尤其适用于具有稳定特征的水质指标预测;而 BiGRU 作为双向门控循环网络,通过同时捕获数据的前后向依赖关系,可挖掘水质参数中隐含的非线性动态特征,在应对突发污染事件等复杂变化时表现出更强的适应性。两者的组合模型通过融合 ARIMA 的显式趋势分解能力与 BiGRU 的深度特征学习优势,先利用 ARIMA 提取序列的线性成分,再通过 BiGRU 对残差中的非线性模式进行建模,从而实现对水质变化规律的多层次解析。相较于单一模型,这种组合策略既保留了传统统计方法在可解释性方面的优势,

又继承了深度学习对复杂关系的表征能力。通过对比不同模型(包括 ARIMA、BiGRU 以及 ARIMA-BiGRU 组合模型)对溶解氧、氨氮等关键指标的预测效果,研究发现组合模型在长短期预测任务中均展现出更高的精度和稳定性。这种融合方法不仅为水质预测提供了兼顾线性与非线性建模的新思路,其预测结果还能为管理者提供跨时间尺度的决策支持,助力制定兼顾短期应急响应与长期生态保护的水环境治理方案。

2. 研究现状分析

20 世纪 20 年代至 80 年代,水质模型的发展经历了初级阶段,这一时期的起点是 Streeter-Phelps 模型[1]对水体中有机物的氧化过程进行了描述,在此基础上 O'Connor 和 Dobbins [2]改进了复氧速度常数,并引入了时变模型的概念。20 世纪 80 年代至 90 年代中期,水质模型的研究迎来了快速发展期,这些模型通过数值方法求解方程来实现对污染物动态变化的有效分析[3],如一维稳态模型 QUAL II [4]及二维和三维模型 EFDC [5]。在实际应用这些模型时需要精确设定边界条件和模型参数,并划分细致的计算网格,这对资料的完整性和准确性提出了较高要求,想要通过数学方法完全准确地描述这些过程非常困难。

20 世纪 90 年代中期以来,机器学习技术在水质预测模型研究中的应用取得了长足的发展。支持向量机(SVM)能够通过最大化分类间隔来寻找最优决策边界,适用于小样本、非线性及高维模式识别问题,因此在水质评估和预测方面具有较高的实用价值[6]。随着算法的不断优化,神经网络及其变种模型逐渐成为水质预测领域的主流,以其强大的非线性映射能力显著提升了水质预测的准确性和可靠性。李霖等[7]提出由 CNN 卷积神经网络驱动的区域预测模型,它能够同时考虑水质的时间和空间变化规律,有效捕捉水质数据中复杂的非线性结构,提高了赣江流域水质预测精度。Zhi 等[8]采用长短期记忆网络(LSTM)构建了湖泊溶解氧的预测模型, LSTM 模型能够捕捉到长期依赖关系,适合处理具有时间序列特性的水质数据,能够更好地模拟水质指标随时间的变化规律,提高预测的准确性。王昱文等[9]展示了复合神经网络在长江流域四项水质指标预测中的优越性,该模型相较于传统神经网络和其他机器学习方法,显示出更好的适用性和准确性。Shi 等[10]证明了神经网络在高频地表水水质预测任务中的高效性。郑湔等[11]针对渭河咸阳段的水质数据,利用 BP 神经网络的自适应与自组织能力进行了模拟和验证,确定了此方法在预测水质变化情况中的可行性。周朝勉等[12]提出一种基于卷积神经网络(CNN)和长短期记忆网络(LSTM)的混合模型来预测安徽省蚌埠闸水质中的溶解氧浓度,试验结果表明 CNN-LSTM 混合模型要优于单一的 LSTM 模型和 BP 神经网络模型。

近年来,研究者们还探索了更多先进的深度学习架构,如门控循环单元(GRU)等。徐倩文[13]建立门控循环神经网络模型(GRU)对维多利亚港的水质进行预测,结果表明 GRU 神经网络模型的预测精度要高于 LSTM 模型,具有可靠性和适应性。韩旭等[14]提出了一种基于多头注意力机制的 CNN-BiGRU 混合模型用于水质预测,通过融合 CNN 的特征提取能力和 BiGRU 的时间序列预测优势,实现了高精度的水质预测,为水环境污染治理提供了技术支持。BiGRU 继承了 GRU 的门控机制,减少了参数数量,提高了计算效率,尤其适合处理具有前后向依赖关系的水质数据。在预测河流水质变化时, BiGRU 可以同时考虑上游和下游的影响,提供更加全面的预测结果。

3. 变量指标与数据获取

本文的数据来源于国家地表水水质自动监测实时数据发布系统(国家水质自动综合监管平台),特别选取了位于河北省秦皇岛市海河流域的戴河口断面监测站点。本次研究所使用的数据涵盖了从 2018 年 6 月至 2024 年 3 月的时间段,期间采取每 4 小时一次的高频次采样策略,旨在捕捉水质参数的动态变化特性。针对水质评估与预测的目的,我们选择了以下六项重要指标: pH 值、电导率(EC)、溶解氧(DO)、高

锰酸盐指数(IMn)、总磷(TP)以及氨氮($\text{NH}_4^+\text{-N}$)。这些指标被广泛认为是衡量水质状况的关键参数,能够有效反映水体的物理、化学及生物性质。通过分析这些参数的变化趋势,可以为水资源管理、污染控制措施制定提供科学依据,并对未来的水质状况做出合理的预测。此外,为了确保数据的准确性和可靠性,所有采集的数据均经过严格的质量控制程序,包括数据校验、异常值检测和处理等步骤,提升了研究结果的可信度,也为进一步深入探讨水质变化规律及其驱动因素奠定了坚实基础。

为了确保数据采集的准确性和及时性,我们根据各个数据源发布数据的频率设定了相应的定时任务。国家地表水水质自动监测实时数据发布系统每 4 小时更新一次各站点的水质监测数据,具体更新时间为每日的 0 时、4 时、8 时、12 时、16 时和 20 时左右。基于此,我们利用 Python 爬虫技术设置了每 4 小时执行一次的数据采集任务,确保获取最新发布的水质数据。通过对收集到的数据进行初步质量审查,分析数据缺失情况、数据异常情况等,剔除异常值确保了后续分析的准确性,从 2018 年 6 月 17 日至 2024 年 3 月 17 日的时间跨度内,共整理出 69 个月的数据用于后续研究。

4. 模型的建立与研究

4.1. 模型设计思路及数据预处理

针对水质时间序列数据中复杂的线性和非线性特征,本文提出了一种基于 ARIMA-BiGRU 的混合模型。该模型结合了传统时间序列分析方法(ARIMA)与深度学习技术双向门控循环单元(BiGRU),旨在通过以下方式提升分类性能:线性特征提取,利用 ARIMA 模型捕捉水质数据中的趋势性、周期性和平稳性特征;非线性特征挖掘,通过 BiGRU 网络建模数据中的复杂时序依赖关系和非线性模式;特征融合与优化,将 ARIMA 的残差即未捕捉的噪声信息与原始特征结合,作为 BiGRU 的输入,从而增强模型对数据多尺度特征的表达能力。在进行数据加载与清洗时,以“监测时间”为索引,按时间顺序对齐水质指标(水温、pH 值、溶解氧等)及类别标签(I-劣 V 类),采用线性插值法填充缺失值,确保时序连续性。然后进行标准化与编码,对连续型特征进行 Z-score 标准化,消除量纲差异,使用 LabelEncoder 将水质类别 I-劣 V 类映射为数值标签 0~5,适配分类任务需求。再进行平稳性检验与差分处理,通过 ADF 检验(Augmented Dickey-Fuller Test)验证时间序列的平稳性,对非平稳序列进行差分操作,直至通过检验($p\text{-value} < 0.05$)。

4.2. ARIMA 模型构建

ARIMA(自回归积分滑动平均)模型是一种经典的时间序列分析方法,适用于捕捉数据中的线性特征,如趋势和周期性。其核心思想是通过自回归(AR)、差分(I)和滑动平均(MA)三个部分对时间序列进行建模。ARIMA 模型表示为 ARIMA(p, d, q),其中:p 为自回归阶数,表示当前值与过去 p 个值的线性关系;d 为差分阶数,用于使非平稳时间序列平稳化;q 为滑动平均阶数,表示当前值与过去 q 个残差的线性关系。具体模型公式为:

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) (1-L)^d y_t = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t$$

其中 L 为滞后算子, ϕ_i 、 θ_i 为系数, ε_t 为白噪声。

通过 ADF 检验确定差分阶数 d,使序列平稳。利用自相关图(ACF)和偏自相关图(PACF)分别确定 q 和 p。基于 AIC 准则优化参数组合,拟合 ARIMA 模型。计算 ARIMA 预测残差 $e_t = y_t - \hat{y}_t$,反映模型未捕捉的非线性信息。将残差作为新特征输入至 BiGRU 模型,增强特征表达能力。通过 ARIMA 模型,能够有效提取时间序列中的线性特征,为后续 BiGRU 建模提供重要补充。

4.3. BiGRU 模型构建

双向门控循环单元(BiGRU)是一种改进的循环神经网络(RNN),通过引入双向结构,能够同时捕捉时间序列数据中的前向和后向依赖关系,适用于建模复杂的非线性时序特征。BiGRU 在传统 GRU 的基础上,增加了反向传播路径,从而更全面地提取时序信息。BiGRU 由两个独立的 GRU 层组成,分别从前向和后向处理输入序列:前向 GRU:从时间步 $t=1$ 到 $t=T$ 处理序列,生成隐藏状态 $\vec{h}_t$;后向 GRU:从时间步 $t=T$ 到 $t=1$ 处理序列,生成隐藏状态 $\overleftarrow{h}_t$ 。最终, BiGRU 的输出为前向和后向隐藏状态的拼接: $H_t = [\vec{h}_t \oplus \overleftarrow{h}_t]$, 其中, $\oplus$ 表示向量拼接操作。每个 GRU 单元通过以下公式更新隐藏状态:重置门 $r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$;更新门 $z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$;候选隐藏状态 $\tilde{h}_t = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t])$;最终隐藏状态 $h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$ 。其中, σ 为 Sigmoid 函数, $\odot$ 表示逐元素乘法, W_r 、 W_z 、 W_h 为可学习参数。

在网络设计中,输入层是接收维度为 (N, T, F) 的时序数据,其中 N 为样本数, T 为时间步长, F 为特征数。BiGRU 层隐藏单元数为 64,输出维度为 $(N, T, 128)$ (双向拼接后)。Dropout 层的丢弃率为 0.5,用于防止过拟合。输出层通过 Softmax 函数输出六类水质的概率分布。损失函数中的稀疏交叉熵,适用于多分类任务。Adam 优化器,初始学习率为 0.0005,配合动态学习率衰减。正则化中早停策略(patience = 10),监控验证集损失以避免过拟合。类别权重则是根据样本频率来计算,提升模型对少数类的分类性能。通过 BiGRU 模型,能够有效捕捉水质时间序列中的非线性依赖关系,结合 ARIMA 提取的线性特征,显著提升分类性能。

4.4. ARIMA-BiGRU 模型

本文通过结合 ARIMA 与 BiGRU,实现对水质时间序列数据的高效建模与分类。其核心思想是通过 ARIMA 提取线性特征,并通过 BiGRU 捕捉复杂的非线性依赖关系,从而实现特征的多尺度融合与优化。ARIMA 模块用于捕捉时间序列中的线性特征如趋势性和周期性,通过拟合 ARIMA 模型生成预测值 $\hat{y}_t$,并计算残差 $e_t = y_t - \hat{y}_t$,反映 ARIMA 未能捕捉的非线性信息。BiGRU 模块用于建模时间序列中的非线性特征和复杂依赖关系,输入为原始特征与 ARIMA 残差的拼接,输出为水质类别的概率分布,此模块输出的分类结果为 $P(y_t | X_t) = \text{Softmax}(W_0 \cdot H_t + b_0)$,其中 H_t 为 BiGRU 的隐藏状态, W_0 和 b_0 为输出层的权重和偏置。

ARIMA-BiGRU 模型通过特征级融合将 ARIMA 的线性特征与 BiGRU 的非线性特征结合,将原始特征 X_t 与 ARIMA 残差 e_t 拼接,形成增强特征集 $X'_t = [X_t \oplus e_t]$,其中 $\oplus$ 表示特征维度拼接。ARIMA 残差 e_t 包含线性模型未捕捉的噪声和非线性信息, BiGRU 通过双向结构进一步提取时序依赖关系,增强后的特征集 X'_t 输入 BiGRU 网络,同时建模线性和非线性特征。ARIMA 模块负责显性趋势如长期水质变化, BiGRU 模块挖掘隐性模式如突发污染事件的特征关联,残差传递机制使模型自适应调整特征权重,避免单一模型对特定特征的过拟合。在模型训练过程中采用适用于多分类任务的稀疏交叉熵损失函数,即

$$L = - \sum_{t=1}^T \sum_{c=1}^C y_{t,c} \cdot \log(P(y_{t,c} | X'_t))$$

其中 $y_{t,c}$ 为真实标签, $P(y_{t,c} | X'_t)$ 为预测概率。在模型优化过程中 Adam

优化器的初始学习率设为 0.0005,配合动态学习率衰减,当验证损失停滞时按因子 0.2 降低学习率。设置早停机制来监控验证损失,若连续 10 轮无改善则提前终止训练,恢复最佳权重。并进行类别权重调整,根据训练集样本分布计算权重,提升模型对少数类(劣 V 类)的分类敏感性。Dropout 层设置丢弃率 = 0.5,在 BiGRU 层后加入,随机屏蔽部分神经元,防止过拟合。

本文构建的 ARIMA-BiGRU 混合模型在水质分类任务中展现出较高的分类性能与泛化能力,验证了时序线性特征与非线性特征结合的有效性。通过实验结果分析,得出以下结论,模型在测试集上取得 88.8%

的准确率，五折交叉验证平均准确率达 88.0%，表明其具有较强的稳定性。训练过程中验证准确率从初始 35.3%逐步提升至峰值 90.3%，并在训练后期稳定于 88.0%-89.6%区间，证明模型能有效学习水质数据的时序特征与非线性关系。

Table 1. Classification report
表 1. 分类报告

类别	精确度	召回率	F1 值
I	0.62	1.00	0.77
II	0.93	0.95	0.94
III	0.91	0.85	0.88
IV	0.87	0.83	0.85
V	0.79	0.96	0.87
劣 V	0.94	0.68	0.79
宏平均	0.84	0.88	0.85
加权平均	0.89	0.89	0.89

本文通过分类模型对多个类别进行了预测，并基于预测结果生成了详细的分类报告。根据上表 1 的各项指标，模型的加权平均精确度、召回率和 F1 分数分别为 0.89、0.89 和 0.89，表明模型在整体数据集上展现出了良好的性能。这证明了所选模型在处理此类分类任务时的有效性和可靠性。II 类(F1-score 0.94)、V 类(F1-score 0.87)水质识别效果最佳，宏平均精确度、召回率和 F1 分数分别为 0.84、0.88 和 0.85，这些指标提供了对所有类别性能的总体评估，即整体性能良好。

从图 1 的混淆矩阵及评估结果来看，模型在水质分类任务中展现出较好的整体性能。混淆矩阵主对角线值显著高于非对角线元素，表明模型对多数水质类别具备较强的辨识能力。测试集准确率达到 88.8%，验证了 ARIMA-BiGRU 混合架构在时序水质数据特征提取中的有效性。

从图 2 基于多分类的 ROC 曲线也可以看出 ARIMA-BiGRU 混合模型在水质分类任务中展现出卓越的判别能力。所有水质类别的 AUC 值均高于 0.98，其中 I 类(Class 0)和IV类(Class 3)的 AUC 达到 1.00，表明模型能够近乎完美地区分正负样本，验证了其在高维时序特征提取与非线性关系建模中的有效性。I 类水质(Class 0)的 AUC 为 1.00，结合混淆矩阵中 100%的召回率，表明模型对高等级水质(如 I 类)的识别具有绝对可靠性。劣 V 类(Class 5)的 AUC 为 0.98，虽低于其他类别，但仍处于极高水平，其召回率偏低 (68%)更多源于样本复杂性而非模型判别能力不足。高 AUC 值(均 > 0.98)表明模型在实际水质监测场景中具备较强的鲁棒性，尤其在区分相邻水质类别(III类与IV类)时表现稳定，可为环境管理部门提供高可信度的分类结果。

从表 2 中可以看出，五次交叉验证的平均准确率为 88.01%，且各折叠间准确率波动较小 (86.99%~89.56%)，表明模型在不同数据子集上均表现出较强的稳定性与泛化能力。II 类水质(平均召回率 > 90%)和 V 类水质(平均召回率 > 92%)的识别效果最佳，宏平均 F1 值为 84.60%，加权平均 F1 值为 88.00%，说明模型在整体分类任务中具备较高的综合性能。

总体研究结果表明模型在整体数据集上表现出较高的预测能力，加权平均精确度、召回率和 F1 分数均达到较高水平，但也存在一些误差因素会对模型性能造成影响。首先，数据质量与特征选择对模型性能具有直接影响。数据中的噪声、异常值以及特征选择的合理性均可能限制模型的预测准确性，未来会通过特征工程筛选更具代表性的特征，以减少冗余信息对模型性能的干扰。其次，模型复杂度与类别

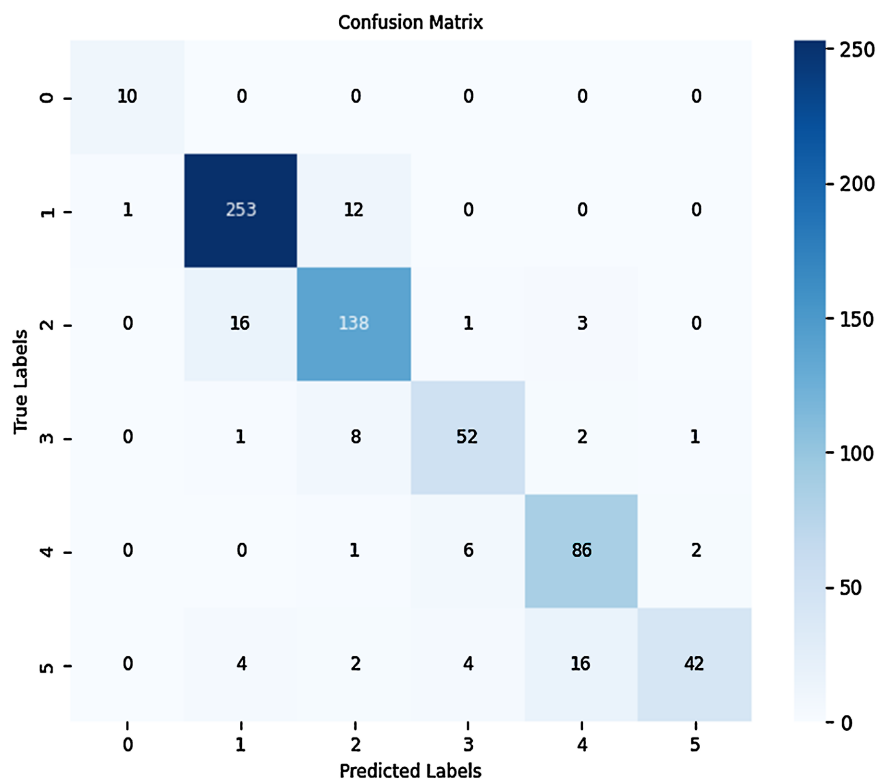


Figure 1. Confusion matrix

图 1. 混淆矩阵

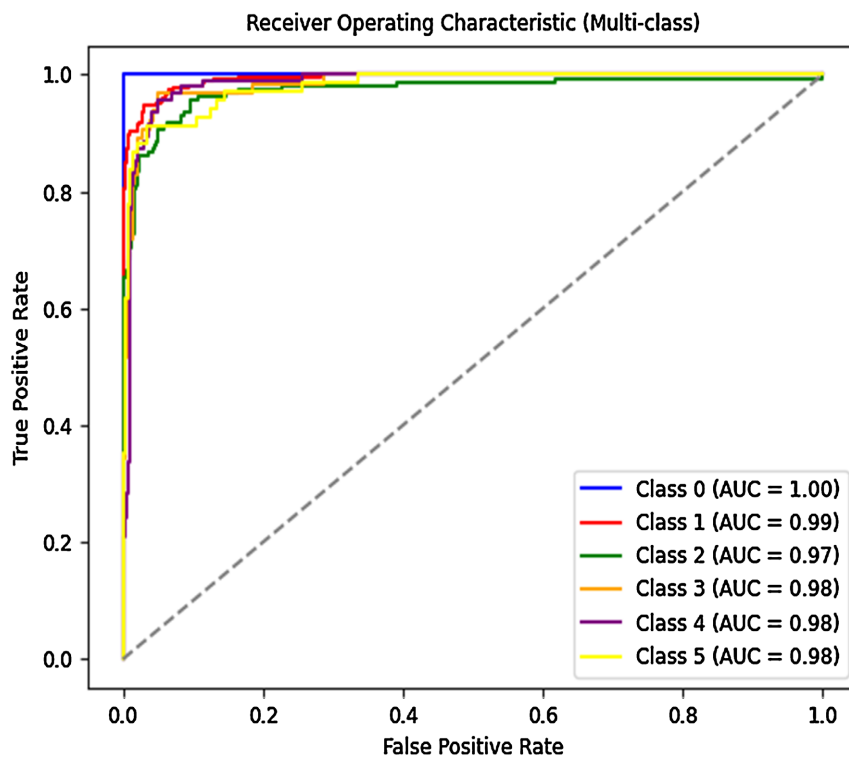


Figure 2. ROC curve and AUC value

图 2. ROC 曲线及 AUC 值

Table 2. Cross-validation results**表 2.** 交叉验证结果

组别	准确率	召回率	F1 值
Fold1	86.99%	82.00%	84.00%
Fold2	89.56%	88.00%	88.00%
Fold3	87.88%	80.00%	81.00%
Fold4	88.03%	87.00%	85.00%
Fold5	87.58%	85.00%	85.00%
平均	88.01%	84.60%	84.60%

不平衡问题也是影响预测结果的重要因素，可能使模型在少数类别上的预测表现不佳。未来会探索更先进的模型架构如深度学习模型来捕捉数据中的复杂关系，并采用重采样或代价敏感学习等方法处理类别不平衡问题。此外，水质指标的选择与测量准确性对预测结果具有重要影响，未来会充分考虑水质指标间的相互作用及其对水质类别判定的影响。时间尺度也是关键因素，水质状况会随时间不断变化，模型需具备捕捉时间动态的能力以提高预测准确性。综上所述，未来的研究会进一步优化数据预处理、特征选择、模型架构，并充分考虑水质指标和时间尺度的影响，以提升模型的预测准确性和泛化能力。

5. 结论

本文提出的基于 ARIMA-BiGRU 混合模型的时序水质分类方法，从实际监测数据中提取多维特征进行实证研究。实验结果表明，通过双向 GRU 层与 ARIMA 的结合，模型能够有效捕捉水质数据的时序依赖性与非线性关系。利用五折交叉验证对模型性能进行评估，平均准确率达到 88.01%，验证了模型的稳定性与泛化能力。通过混淆矩阵与 ROC 曲线分析，进一步证实了模型在高精度分类任务中的优越性，尤其在多数类(II 类、V 类水质)识别中表现突出。该模型可为水质监测与评估提供科学依据，为环境管理部门决策提供支持。

参考文献

- [1] Streeter, H.W. and Phelps, E.B. (1925) A Study of Pollution and Natural Purification of the Ohio River. *Public Health Bulletin*, **13**, 76-79.
- [2] O'Connor, D.J. (1967) The Temporal and Spatial Distribution of Dissolved Oxygen in Streams. *Water Resources Research*, **3**, 65-79. <https://doi.org/10.1029/wr003i001p00065>
- [3] 张晨, 王立义, 高英, 等. 引滦入津工程黎河段机理与非机理水质预测模型对比分析[J]. 安全与环境学报, 2011, 11(1): 149-152.
- [4] Câmara, A.S. and Randall, C.W. (1984) The QUAL II Model. *Journal of Environmental Engineering*, **110**, 993-996. [https://doi.org/10.1061/\(asce\)0733-9372\(1984\)110:5\(993\)](https://doi.org/10.1061/(asce)0733-9372(1984)110:5(993))
- [5] Song, W., Xu, Q., Fu, X., Wang, C., Pang, Y. and Song, D. (2019) EFDC Simulation of Fishway in the Diversion Dashaerteng River to Danghe Reservoir, China. *Ecological Indicators*, **102**, 704-715. <https://doi.org/10.1016/j.ecolind.2019.03.025>
- [6] 王成杰, 张森. 基于主成分和粒子群优化支持向量机的水质评价模型[J]. 环境工程学报, 2014, 8(10): 4545-4549.
- [7] 李霖, 王琨, 刘强, 等. 基于 CNN 的赣江水质时空规律分析与预测[C]//中国统计教育学会, 教育部高等学校统计学类专业教学指导委员会, 全国应用统计专业学位研究生教育指导委员会. 2021 年(第七届)全国大学生统计建模大赛获奖论文集(二). 北京: 中国统计出版社, 2021: 26.
- [8] Zhi, W., Feng, D., Tsai, W., Sterle, G., Harpold, A., Shen, C., et al. (2021) From Hydrometeorology to River Water Quality: Can a Deep Learning Model Predict Dissolved Oxygen at the Continental Scale? *Environmental Science & Technology*, **55**, 2357-2368. <https://doi.org/10.1021/acs.est.0c06783>

- [9] 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.
- [10] Shi, B., Wang, P., Jiang, J. and Liu, R. (2018) Applying High-Frequency Surrogate Measurements and a Wavelet-ANN Model to Provide Early Warnings of Rapid Surface Water Quality Anomalies. *Science of the Total Environment*, **610**, 1390-1399. <https://doi.org/10.1016/j.scitotenv.2017.08.232>
- [11] 郑昊, 薛惠锋, 冯海涛. 基于 BP 神经网络的地表水污染指标短期预测分析[J]. 陕西科技大学学报: 自然科学版, 2016, 34(3): 156-160.
- [12] 周朝勉, 刘明萍, 王京威. 基于 CNN-LSTM 的水质预测模型研究[J]. 水电能源科学, 2021, 39(3): 20-23.
- [13] 徐倩文. 基于 GRU 的水质数据预测[J]. 中国新通信, 2020, 22(22): 123-124.
- [14] 韩旭, 段中兴, 尚旭东. 基于多头注意力机制和 CNN-BiGRU 模型的水质预测方法[C]//中国自动化学会. 2024 中国自动化大会论文集. 2024: 6.