J. Lake Sci.(湖泊科学), 2026, 38(2):000-000
DOI 10.18307/2026.0216

基于机器学习构建微生物指数的水生态状况评价：以鄱阳湖为例*

彭彩霞[1,2]，王秀君[2]，李梦源[2]，吴 兰[1**]，曾 巾[2,3**]
（1：南昌大学生命科学学院，鄱阳湖环境与资源利用教育部重点实验室，南昌 330031）
（2：中国科学院南京地理与湖泊研究所湖泊与流域水安全全国重点实验室，南京 211135）
（3：中国科学院南京地理与湖泊研究所鄱阳湖湖泊湿地综合研究站，九江，332899）

摘要：微生物能够对环境变化做出快速响应，具有应用于生态状况评价的潜力，尚缺乏其在水生态评价中的应用。本研究以鄱阳湖不同湖区（主湖区、碟形湖和养殖湖）为研究区域，采用高通量测序技术分析水体和沉积物中的细菌群落结构，并基于机器学习的无分类策略构建了微生物指数（Microbial Biotic Index, MBI）对不同湖泊进行评价。结果表明，不同湖泊水体和沉积物营养状态存在显著差异，驱动着细菌群落结构及关键菌属的空间分布差异；尽管水体和沉积物细菌群落融合程度较低，但随着营养状态指数的升高，水体和沉积物之间细菌群落相互作用逐渐增强，即水体中沉积物源细菌比例显著增加。通过分位数回归和机器学习算法，将水体和沉积物中的细菌 ASVs 划分为五个生态组，基于此构建的 MBI 指数与水和沉积物的营养状态指数之间呈现出高度一致的相关性。评价结果显示，主湖区和养殖湖的水体生态状况相似，均优于碟形湖；而沉积物生态状况由优至中表现为养殖湖＞碟形湖＞主湖区。鄱阳湖水体的生态状况优于沉积物，表明了沉积物可能成为水体富营养化的潜在来源，在未来湖泊健康管理中需要加强关注沉积物的生态状况。本研究通过机器学习方法构建了基于微生物群落的综合评价指数（MBI），能够为不同营养水平水体的生态状况评价提供参考。
关键词：鄱阳湖；富营养化；微生物指数；机器学习；生态评价

Assessment of Aquatic Ecological Status Based on a Microbial Biotic Index Developed by Machine Learning: A Case Study of Lake Poyang *

Peng Caixia[1,2]，Wang Xiujun[2]，Li Mengyuan[2]，Wu Lan[1**]& Zeng Jin[2,3**]
(1: School of Life Sciences, Key Laboratory of Poyang Lake Environment Resource Utilization, Ministry of Education, Nanchang University, Nanchang 330031, P.R.China)
(2: State Key Laboratory of Lake and Watershed Science for Water Security, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 211135, P.R.China)
(3: Poyang Lake Wetland Research Station, Nanjing Institute of Geography and Limnology, Chinese Academy of' Sciences, Jiujiang 332899, P.R.China)

Abstract: Microorganisms respond rapidly to environmental changes, rendering them promising candidates as bioindicators for ecological assessments. Nevertheless, their application in evaluating aquatic ecosystems remains relatively limited. This study employed high-throughput sequencing to investigate bacterial community structures in both water and sediments across three

* 2025-05-22 收稿；2025-06-25 收修改稿。
中国科学院战略性先导科技专项（XDB0810000）和江西省千人计划项目（jxsq2020101042）联合资助。
** 通信作者；E-mail：jzeng@niglas.ac.cn，ncusk724@hotmail.com。

distinct lake zones within Lake Poyang: the main lake area, dish-shaped lakes, and aquaculture lakes. A machine learning framework was utilized to develop a Microbial Biotic Index (MBI) for assessing the ecological status of these habitats. Results revealed significant disparities in nutrient levels between water and sediments across lake types, which in turn drove spatial variations in bacterial community composition and dominant genera. Although community coalescence between water and sediment bacterial assemblages was limited, elevated trophic state index (TSI) levels were associated with enhanced cross-habitat connectivity, indicated by a significantly greater proportion of sediment-derived bacteria within the water column. Using quantile regression and machine learning, Amplicon Sequence Variants (ASVs) from both water and sediment were classified into five ecological groups. The resulting MBI demonstrated strong concordance with trophic state indices for both water and sediment compartments. Ecological assessments based on the MBI indicated superior water quality in the main lake and aquaculture lakes compared to dish-shaped lakes. In contrast, sediment ecological conditions were highest in aquaculture lakes, intermediate in dish-shaped lakes, and poorest in the main lake area. Overall, water quality exceeded sediment conditions, suggesting that sediments may act as internal nutrient sources contributing to eutrophication. These findings highlight the critical need to incorporate sediment quality monitoring into future lake management strategies. This study proposes a novel microbiome-based assessment index, constructed via machine learning, that provides a robust tool for evaluating aquatic ecological conditions across varying trophic states.

Keywords: Lake Poyang; eutrophication; Microbial Biotic Index; machine learning; ecological assessment

作为我国最大的淡水湖和长江流域重要调蓄水体，鄱阳湖具有多重生态价值。2000 年以来鄱阳湖水质恶化趋势显著，总氮、总磷等指标持续超标，局部水域已出现富营养化[1,2]。水体富营养化的发生受到营养物质输入、水文条件及湖泊形态等因素影响，而各因素在不同湖区中作用强度存在明显空间异质性，导致传统单一化评估方法在应用中存在局限性[3]。传统的湖泊环境状况评估方法主要依赖于物理、化学指标，通常包括水体透明度（Secchi Depth, SD）、pH 值（pH）、溶解氧（Dissolved Oxygen, DO）、总氮（Total Nitrogen, TN）和总磷（Total Phosphorus, TP）等参数[4]。这些方法虽然提供了关于水质和水环境的基本信息，但忽略了生物群落和生态过程的复杂性，难以全面揭示湖泊生态系统的健康状态[5]。随着生物监测技术的发展，越来越多研究将鱼类、大型无脊椎动物和藻类作为指示生物[6,7]，通过这些生物对环境变化的敏感性来评价湖泊生态状况[8,9]。与浮游植物和鱼类等生物指标相比，微生物具有数量多、生长快、多样性高、代谢能力强、分布广泛等特征，使得微生物对环境变化响应敏感，对湖泊生态环境具有指示和反馈作用[10,11]。

高通量测序技术在微生物生态监测中取得了较大进展，有助于获得比传统的分类学方法更全面、准确的生态信息[12]。然而，大量微生物因不可培养或难以在实验室实现纯培养，导致许多测序序列无法与已知菌株建立对应，也制约了对其生态功能的深度解析。近年来，为解决参考数据库不完备导致的分类和注释问题，一些研究提出了基于机器学习（Machine Learning, ML）的无分类策略，通过训练模型直接从环境数据中推断和赋予微生物操作分类单元（Operational Taxonomic Unit, OTUs）或扩增子序列变体（Amplicon Sequence Variant, ASVs）生态学意义，从而大幅提高了高通量测序数据的利用率[13,14]。基于机器学习分析微生物数据，可以提高其生态系统健康评估的准确性，从而创新水生态状况评估方法。Wright 和 Ziegler 通过开发 ranger 随机森林算法，提供了一种高效处理高维数据的技术工具[15]，该算法在处理大规模生态数据时，通过提升运行效率和内存优化，能够有效处理更多分类学信息不完备的样本，这一工具的广泛应用能够为复杂的生态状况评估提供一种新的技术手段。

本研究基于高通量测序数据，构建并验证了一种新的水生态评价指数：微生物指数（Microbial Biotic Index, MBI）。研究以鄱阳湖主湖区、碟形湖与养殖湖三个典型功能区为研究区域，分析水体与沉积物的理化特性和微生物群落结构特征，并通过分位数回归（Quantile Spline Regression, QSR）和基于随机森林的机器学习算法，构建了 MBI 指数用于鄱阳湖水体和沉积物的生态状况评估，并与传统营养指数方法对比验证 MBI 指数的准确性和有效性。

1.材料与方法

1.1 研究区域和样品采集

鄱阳湖（28°25′~29°45′N，115°50′~116°44′E）是长江中下游地区重要的大型通江湖泊之一，在维系长江水量平衡和水域生态平衡方面发挥着关键作用[16]。鄱阳湖位于江西省北部，提供了生物多样性保护、水资源供给、调洪蓄水等广泛的生态系统服务功能[17,18]。鄱阳湖主湖区水深、流通性强，其水文动态是调节水量、水质的关键[19]；碟形湖为季节性湖泊，具有“高水位相通、低水位分离”的水文特性，是鄱阳湖生态完整性与生物多样性的核心区域[20,21]；养殖湖为人工控制的围隔养殖区，坚持生态化管理，在评估人类活动与生态平衡方面具有重大意义[22]。本研究于 2024 年 1 月在鄱阳湖的三种典型区域（主湖区、碟形湖、养殖湖）共设置了 9 个位点（图 1），每个位点设置 3 个取样点作为重复，收集水体和沉积物样本共 54 份。使用取水器采集水面以下 50 cm 的水样，装入无菌瓶中。采用抓斗式采泥器采集表层沉积物，样品立刻转移至 4 ℃的采样箱保存。所有样品在 24 h 内用 4 ℃的冰盒送回实验室处理。

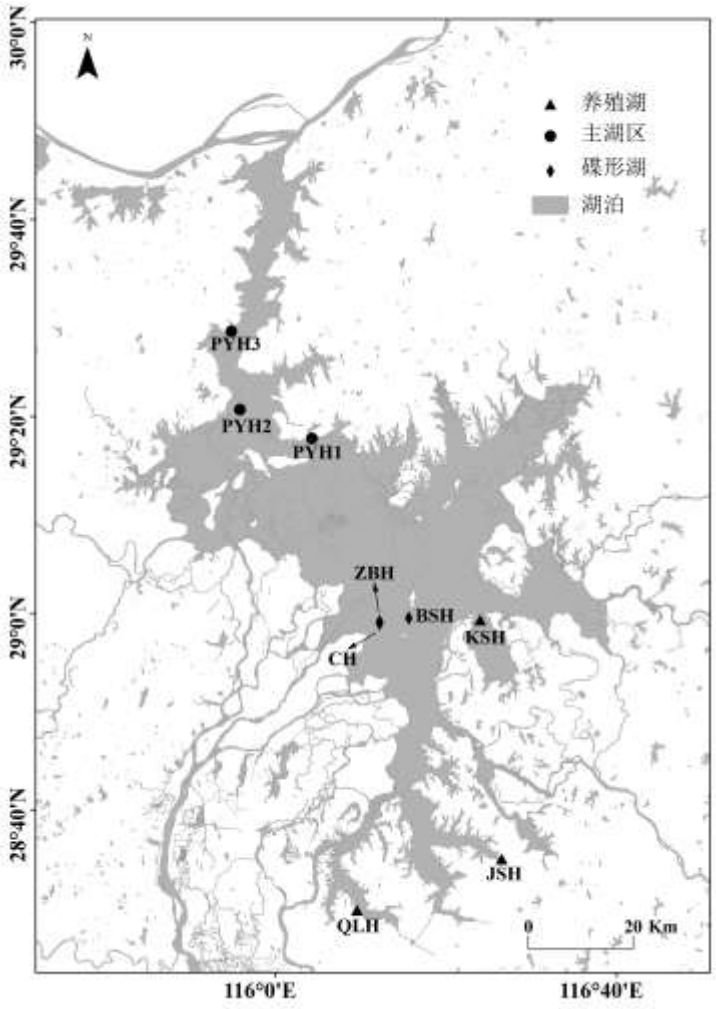

图 1 鄱阳湖采样点分布情况

主湖区：鄱阳湖 1（PYH1）、鄱阳湖 2（PYH2）和鄱阳湖 3（PYH3）；碟形湖：白沙湖（BSH）、战备湖（ZBH）和常湖（CH）；养殖湖：军山湖（JSH）、青岚湖（QLH）和康山湖（KSH）

Fig.1 Distribution of sampling sites in Lake Poyang

The main lake area: PYH1, PYH2, and PYH3; the dish-shaped lakes: Lake Baisha (BSH), Lake Zhanbei e (ZBH), and Lake Chang (CH); the aquaculture lakese: Lake Junshan (JSH), Lake Qinglan (QLH), and Lake Kangshan (KSH)

水样经 0.22 μm 的聚碳酸酯滤膜（Millipore, Billerica, MA, USA）过滤，收集浮游细菌，并将滤膜保存在-80 ℃超低温冰箱备用。部分水样通过 0.45 μm 滤膜过滤，滤膜保存至-20 ℃冰箱用于后续叶绿素 a（Chlorophyll-α, Chl α）浓度测定。沉积物样品使用真空冷冻干燥机进行冷冻干燥后研磨并均质，用于后续的理化参数测定。

1.2 理化指标测定

使用水质监测仪（YSI 6600, Yellow Springs Instrument, OH, USA）原位测量水温、pH 值、氧化还原电位和电导率，并采用透明度盘测定水体透明度（SD）。水体总氮（TN）通过碱性过硫酸钾消解分光光度法测定，总磷（TP）含量通过钼酸铵分光光度法测定。氨氮（NH_4^+-N）、硝酸盐氮（NO_3^--N）、亚硝酸盐氮（NO_2^--N）通过连续比色流动分析（Skalar SAN PLUS 系统，Skalar Analytical BV，Breda，荷兰）测定。Chl α 通过乙醇萃取分光光度法测定[23]。沉积物的理化参数，包括总磷（TP）、总氮（TN）、氨氮（NH_4^+-N）、硝酸盐氮（NO_3^--N）、亚硝酸盐氮（NO_2^--N）和 pH，均按照 Zeng 等人[24]的方法进行测定。

1.3 DNA 提取和高通量测序

水体和沉积物样本的 DNA 分别使用 E.Z.N.A.® Water DNA 试剂盒（Omega Biotek, Doraville, GA, USA）和 PowerSoil DNA Isolation 试剂盒（Mo Bio Laboratories, Inc., Carlsbad, CA, USA）提取，并使用 NanoDrop ND-1000（NanoDrop Technologies, Wilmington, DE, USA）检验 DNA 浓度和纯度[24]。提取的 DNA 样本作为模板，对 16S rRNA 基因 V4 可变区进行 PCR 扩增，所用引物为：515F（5′-GTGCCAGCMGCCGCGGTAA-3′）和 806R（5′-GGACTACHVGGGTWTCTAAT-3′）。PCR 反应条件如下：初始变性 95°C，5 min；随后进行 28 个循环（95°C 变性 30 s，55°C 退火 30 s，72°C 延伸 45 s）；最终延伸 72°C，10 min。扩增产物由凌恩生物科技有限公司（中国上海）采用 Illumina MiSeq 平台进行 PE250 双端测序。

原始序列数据采用 QIIME 2（Version 2023.9）处理[25]。首先使用 cutadapt 识别和移除引物序列，并删除短于 200 bp 的序列。之后使用 DADA2 对序列进行降噪处理[26]。经过质量过滤的序列被组装成 ASVs（基于 100%序列同一性）。基于 SILVA 138 数据库预先训练好的 Naive Bayes 分类器对细菌 ASVs 进行分类，通过 feature-classifier 进行注释，使用特征分类器的置信度为 80%[27,28]。

1.4 营养指数计算

营养状态指数（Trophic State Index, *TSI*）作为传统的水生态状况评估指标，通过整合分析水体的 Chl α、透明度、磷浓度等参数，综合反映水体的富营养化程度及其潜在生态风险。本研究基于 TN、TP、Chl α、SD 四个参数计算了 *TSI*，具体公式如下[29]：

$$TSI(\text{Chl}\,\alpha) = 10(2.5 + 1.0861\text{nChl}\,\alpha) \quad (1)$$

$$TSI(\text{TP}) = 10(9.436 + 1.6241\text{nTP}) \quad (2)$$

$$TSI(\text{TN}) = 10(5.453 + 1.6941\text{nTN}) \quad (3)$$

$$TSI(\text{SD}) = 10(5.118 - 1.941\text{nSD}) \quad (4)$$

$$TSI = 0.326TSI(\text{Chl}\,\alpha) + 0.219TSI(\text{TN}) + 0.230TSI(\text{TP}) + 0.225TSI(\text{SD}) \quad (5)$$

TSI 在 0-100 范围内量化水体的营养状态，*TSI* 值越高表示富营养程度越高。营养水平级标准为：*TSI* ＜30 表示寡营养，30≤*TSI*≤50 表示中营养，50 ＜ *TSI*≤60 表示轻度富营养化，60 ＜ *TSI*≤70 表示中度富营养化，*TSI* ＞ 70 表示重度富营养化。

参考土壤肥力指数（Soil Fertility Indices，SFI）[30]计算沉积物营养指数（Sediment-Trophic State Index，*S-TSI*）用于评估沉积物的营养状况。首先使用非线性评分模型将 6 个指数的实测值计算为分数 *S*:

$$s = \frac{1}{1 + \left(\frac{x_i}{x_m}\right)^b} \quad (6)$$

其中，x_i为实测值，x_m为平均值，对于沉积物营养水平具有较大影响的参数（如 TN、TP 等）在评分模型中 b＝-2.5；当 pH 大于 7 时设定 b＝2.5，pH 小于 7 时设定 b＝-2.5。然后通过主成分分析选择特征值大于 1 的主成分的公因子方差计算沉积物理化参数的权重，最后进行加权求和计算 *S-TSI*。

$$w_i = \frac{C_i}{\sum_{i=1}^{n} C_i} \tag{7}$$

$$S\text{-}TSI = \sum_{i=1}^{n} w_i s_i \tag{8}$$

式中，w_i是各理化参数的权重，C_i为公因子方差。*S-TSI* 值越高代表沉积物的营养水平越高。

1.5 水体与沉积物细菌群落融合分析

群落融合是指两个或多个微生物群落及其生境在特定空间或环境条件下发生的交换事件，最终形成一个新的群落及生境的过程[31,32]。探究水体和沉积物细菌群落融合（community coalescence）特征，有助于深入理解不同生态位微生物群落间的交互作用及潜在功能[33]。本研究采用基于群落级别的微生物源追踪方法（Microbial Source Tracking）[31]分析水体与沉积物中细菌群落的相互联系，量化不同水域生态环境中浮游细菌与沉积物细菌群落的潜在来源及其相对贡献。

溯源分析在 R 中使用 FEAST 包（fast expectation-maximization microbial source tracking）完成。FEAST 方法假设浮游细菌群落（受体群落）是由多个潜在源群落（如不同地点的沉积物细菌群落）以及未知来源群落按不同比例加权混合而成，能够定量估算已知来源群落及未知来源对浮游细菌群落的贡献比例[33]。

1.6 ASVs 生态组的分类

将 *TSI* 和 *S-TSI* 标准化至 0-100 范围，绘制 ASVs 相对丰度与 *TSI* 和 *S-TSI* 的散点图。本研究采用 95% 分位数回归（QSR）方法，进一步探索 ASVs 丰度与 *TSI* 及 *S-TSI* 之间的非线性依赖关系。QSR 方法能够在指定分位点（本研究中为 95%分位点）捕捉环境因子对 ASVs 丰度的影响，首先使用 splines 包中的 bs（）函数和 quantreg 包中的 rq（）函数构建模型，然后通过 pracma 包计算拟合函数的最高点（峰值），确定每个 ASVs 在预测出最高丰度时对应的营养水平范围（*TSI* 或 *S-TSI* 值），并基于此将所有 ASVs 划分到五个不同的生态组[34,35]。随后，基于随机森林的机器学习算法对 ASVs 生态组的划分结果进行验证。通过 randomForest 包，随机选 30%的 ASVs 生态组作为训练数据集，剩余 70%作为测试数据集。计算并构建 ASVs 的丰度特征与生态组划分之间的机器学习模型[36]。使用测试数据集评估该模型的性能，并通过计算 Cohen's Kappa 值以及比较预测结果与实际结果，以评估模型的准确率。

1.7 浮游和沉积物细菌指数

基于 ASVs 的生态权重及其相对序列丰度，计算微生物指数（Microbial Biotic Index, MBI），公式如下[36-38]：

$$\mathrm{MBI} = (0 \times \%\mathrm{level\ 1}) + (1.5 \times \%\mathrm{level\ 2}) + (3 \times \%\mathrm{level\ 3}) + (5 \times \%\mathrm{level\ 4}) + (12 \times \%\mathrm{level\ 5}) \tag{9}$$

式中，%level 1-5 表示归类为生态组 1-5 的 ASVs 的相对丰度之和。参考 Borja 等在 AMBI 指数中提出的生态分组方法，结合微生物对污染梯度的响应特征，通过回归优化，获得权重系数，反映各生态组在不同污染水平下的生态重要性[36]。MBI 的值越高代表生态状况越好，其分级标准为：优 $[0.8, 1.0]$；良 $[0.6, 0.8)$；中 $[0.4, 0.6)$；一般 $[0.2, 0.4)$；差 $[0, 0.2)$。

1.8 统计分析

所有统计分析在 R（Version 4.0.1）中进行，使用 vegan 包完成主成分（PCA）分析以及基于 Bray-Curtis 距离的置换多元方差分析（Permutational multivariate analysis of variance, PERMANOVA）和主坐标分析（Principal co-ordinates analysis, PCoA）[39]。分别基于 randomForest 包和 rfPermute 包进行随机森林分类、回归分析以及预测因子的统计显著性检验[40,41]。单因素方差分析通过 stats 和 multcomp 包完成，所有统计检验的显著性水平设定为 $P = 0.05$ 水平上确定，不同小写字母表示组间差异的显著性具有统计学意义（$P < 0.05$）。可视化绘图均由 ggplot2 包完成[42,43]。

2. 结果

2.1 水体与沉积物理化性质特征

水体理化性质及营养状态指数评价结果显示，三个区域的水体均达到了中度富营养水平，营养程度由

低到高依次为养殖湖、主湖区和碟形湖，其中碟形湖为重度富营养化（*TSI* = 71.60 ± 3.39），其 *TSI* 显著高于主湖区和养殖湖（$P < 0.05$）：其浊度（442.22 ± 96.71 NTU）、Chl α（23.41 ± 11.63 μg/L）及总磷（0.57 ± 0.17 mg/L）均显著高于主湖区与养殖湖（$P < 0.05$），而透明度（0.14 ± 0.02 m）最低。养殖湖 *TSI* 最低（*TSI* = 62.87 ± 2.40），但其总有机碳含量（3.02 ± 0.85 mg/L）与 Chl α 含量（16.15 ± 4.57 μg/L）显著高于主湖区（$P < 0.05$）。而主湖区的硝态氮含量（1.02 ± 0.36 mg/L）则显著高于其它区域。

对于沉积物，参照土壤肥力指数计算了沉积物营养指数 *S-TSI*（范围 0-1），结果显示，虽然鄱阳湖不同区域之间的沉积物营养状况差异不显著，但仍存在一定的梯度，由低到高分别为养殖湖（0.41 ± 0.12）、碟形湖（0.46 ± 0.07）和主湖区（0.5 ± 0.07）。养殖湖总磷含量（0.59 ± 0.23 mg/kg）低于主湖区（0.76 ± 0.39 mg/kg）和碟形湖（0.75 ± 0.22 mg/kg），氨氮（2.19 ± 1.72 mg/kg）含量高于主湖区（0.97 ± 0.85 mg/kg）和碟形湖（2.00 ± 0.68 mg/kg）。主湖区的 pH 和 NO_3^--N 含量显著高于养殖湖和碟形湖。总氮含量从高到低依次为碟形湖（1.04 ± 0.56 mg/kg）、养殖湖（1.02 ± 0.83 mg/kg）、主湖区（0.84 ± 0.26mg/kg）。

通过 PCA 分析不同富营养化程度下水体和沉积物样品的理化性质差异（图 2）。水体和沉积物环境变量的 PCA 前两轴的累积解释率分别达到了 63.14%和 68.01%，能较好的反映不同湖泊区域间环境变量的差异。与水体和沉积物营养水平的评价结果类似，主湖区、碟形湖和养殖湖两两之间的水体性质均表现出显著差异（$R^2 = 0.79, P = 0.001$）（图 2 a），而不同湖区沉积物的理化性质差异稍小（$R^2 = 0.22, P = 0.027$），主湖区和碟形湖沉积物之间差异较大（$R^2 = 0.37, P = 0.009$）（图 2 b）。此外，通过随机森林分类分析不同湖泊区域间理化性质差异的关键因素，发现水体中关键的预测因子为透明度、水温和浊度（图 2 c），而沉积物中则为 pH、NH_4^+-N 和 NO_2^--N（图 2 d）。

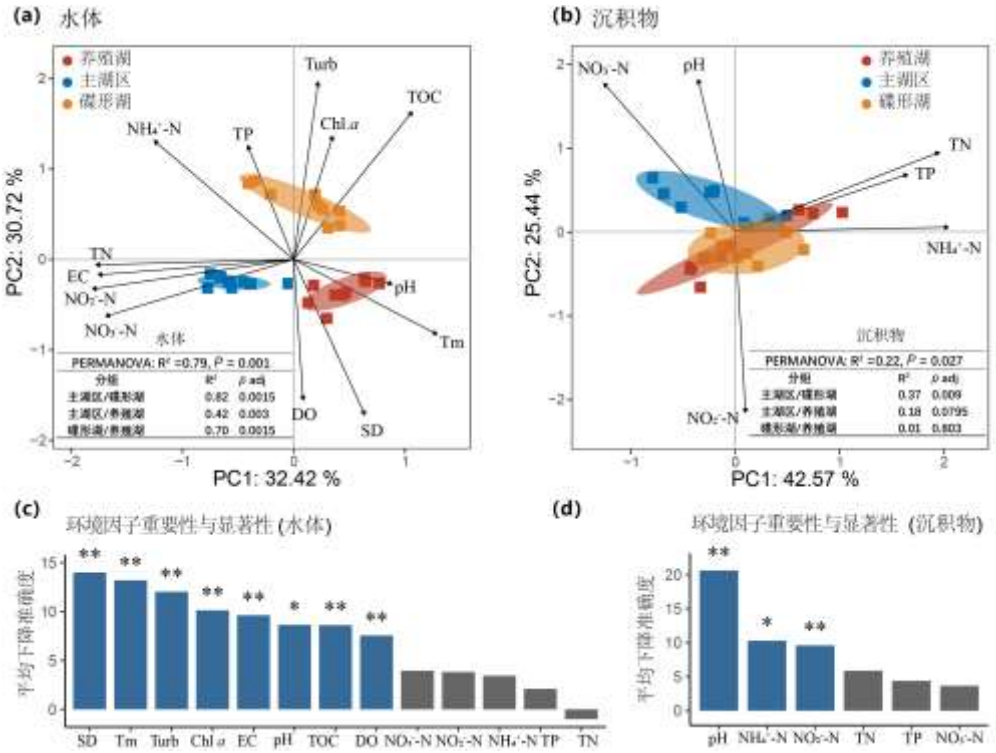

图 2 鄱阳湖不同区域水体（a）和沉积物（b）理化参数 PCA 分析及置换多元方差分析；基于随机森林分类模型评估水体（c）和沉积物（d）环境因子的相对重要性

Fig. 2. Principal component analysis (PCA) combined with PERMANOVA for physicochemical parameters in water (a) and sediment (b) across different regions of Lake Poyang; Relative importance of environmental factors in water (c) and sediment (d) assessed by a random forest classification model

2.2 水体与沉积物细菌群落对营养水平的响应

54 个水体和沉积物 DNA 样品测序获得了 1075886 和 723332 条高质量序列，分别鉴定出 2948 个浮游细菌 ASVs 和 11193 个沉积物细菌 ASVs。水体微生物群落的优势类群（相对丰度前 10）包括 γ-变形菌纲（Gammaproteobacteria）、放线菌门（Actinobacteria）、α-变形菌纲（Alphaproteobacteria）和拟杆菌门（Bacteroidota）（图 3a）；沉积物微生物群落（相对丰度前 10）则以 γ-变形菌纲（Gammaproteobacteria）、绿弯菌门（Chloroflexi）、酸杆菌门（Acidobacteria）和放线菌门（Actinobacteria）为主（图 3b）。门水平物种组成分析结果显示，水体和沉积物中主要优势菌群的相对丰度未随营养状态指数（*TSI* 和 *S-TSI*）的升高表现出明显变化趋势，PCoA 分析结果表明，无论是在水体（$R^2=0.29$，$P=0.001$）（图 3c）还是沉积物（$R^2=0.23$，$P=0.001$）（图 3d）中，细菌群落结构均呈现显著的空间分异。此外，低 *TSI*（*S-TSI*）值样品和高 *TSI*（*S-TSI*）值样品分别在 PCoA 空间中表现为紧密的聚集状态（图 3c、d）。进一步结果表明，水体和沉积物中，*TSI*（*S-TSI*）差值的绝对值与细菌群落的 Bray-Curtis 距离呈显著正相关（图 3e、f），表明营养水平是驱动细菌群落 β 多样性分异的关键因子。

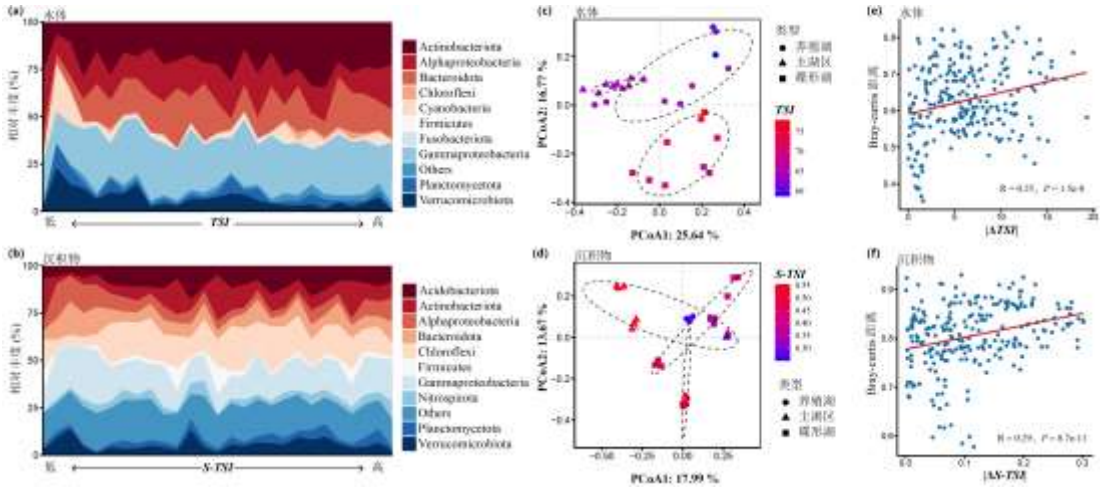

图 3 鄱阳湖水体（a）和沉积物（b）中细菌群落组成（主要门水平，Top10）沿营养状态指数（*TSI*/*S-TSI*）梯度的变化特征；基于主坐标分析（PCoA）的水体（c）和沉积物（d）细菌群落结构差异；水体（e）和沉积物（f）细菌群落距离 Bray-Curtis 距离与 *TSI*（*S-TSI*）绝对差值的相关性分析

Fig. 3. Variations in bacterial community composition (Top 10 phyla) along trophic state index (TSI for water; S-TSI for sediment) gradients in water (a) and sediment (b) of Lake Poyang; Principal coordinate analysis (PCoA) of bacterial community structure in water (c) and sediment (d); Correlation between Bray-Curtis distance and absolute differences of *TSI*/*S-TSI* values in water (e) and sediment (f) compartments

为了深入探讨水体和沉积物细菌群落对营养水平的响应模式，通过随机森林回归分析，筛选出水体（图 4a）和沉积物（图 5a）中与营养状态指数（*TSI* 和 *S-TSI*）显著相关的前 15 个属水平关键物种。在水体中，*Polymorphobacter*（Proteobacteria）、*Methylorosula*（Proteobacteria）和 *Rhodobacter*（Proteobacteria）的丰度与 *TSI* 呈显著负相关（图 4b~d），其丰度随 *TSI* 的升高显著降低。相反，*Rhodoluna*（Actinobacteriota）、*Polynucleobacter*（Proteobacteria）、*Nordella*（Proteobacteria）、*Aurantisolimonas*（Bacteroidota）、*Ferruginibacter*（Bacteroidota）、*Oligoflexus*（Bdellovibrionota）的丰度与 *TSI* 呈显著正相关（图 4e~j），这些属的丰度随 *TSI* 的升高逐渐增加。

沉积物细菌群落表现出类似的生态分化，部分属在低营养水平条件下占主导地位，而在高营养水平环境中则被其他微生物类群所取代。其中，*Arenimonas*（Proteobacteria）、*Gemmatimonas*（Gemmatimonadota）、*Porphyrobacter*（Proteobacteria）、*Rhodoferax*（Proteobacteria）、*Paraburkholderia*（Proteobacteria）和 *Sphingomonas*（Proteobacteria）的丰度与 *S-TSI* 呈负相关（图 5b~g），随着 *S-TSI* 的升高，这些属的丰度显

著降低。相反，*Meiothermus*（Deinococcota）、*Latescibacterota*（Latescibacterota）、*hgcI clade*（Actinobacteriota）和 *Crenothrix*（Proteobacteria）的丰度与 *S-TSI* 呈正相关（图 5h~k），随着 *S-TSI* 升高，这些属的丰度逐渐增加。

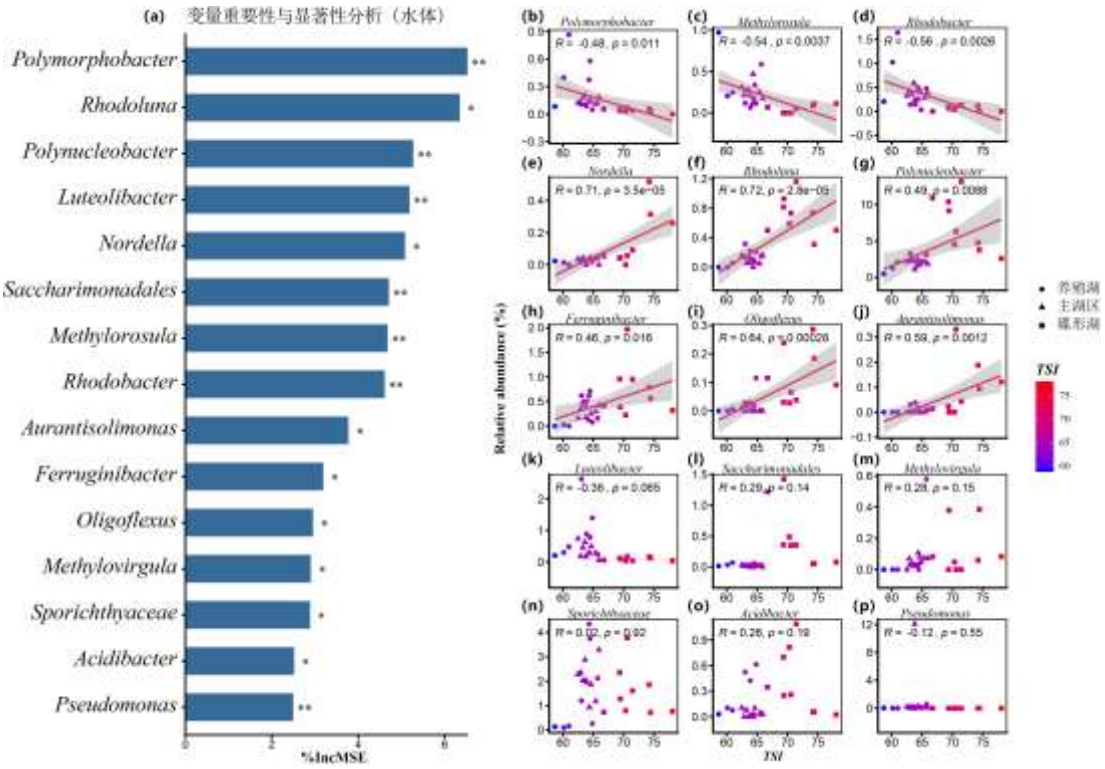

图 4 基于随机森林回归模型识别水体中与营养状态指数（*TSI*）关联的关键类群（属水平，Top15）（a）及其相对丰度沿 *TSI* 梯度的响应特征（b~p）

Fig. 4. Identification of *TSI*-associated indicator genera (Top 15) in water using random forest regression (a), with corresponding relative abundance response patterns across *TSI* (b~p)

2.3 水体与沉积物细菌群落融合特征及其与营养水平的关系

通过群落融合分析，深入探讨不同湖泊类型中水体与沉积物细菌群落的相互作用及物种交换特征。结果显示，所有位点的群落融合程度普遍较低，介于 1.31%到 7.37%之间，且大多数位点的群落融合程度低于 5.0%（图 6a、b）。表明水体与沉积物细菌群落在同一生态系统内仍保持较高的特有性，反映出二者在生态位特征、栖息环境和生境适应性方面存在明显差异。

尽管不同湖泊类型间水体与沉积物的群落融合程度差异不显著，但值得注意的是，水体群落的融合度（即水体中来源于沉积物的细菌比例）与水体的营养水平呈显著正相关：随着水体营养水平的提高，水体中来源于沉积物的细菌比例也显著升高（图 6c）。这一结果揭示，在高营养水平条件下，水体和沉积物细菌群落之间的交互作用增强，表明沉积物中的细菌可能在水体富营养化过程中扮演重要角色。

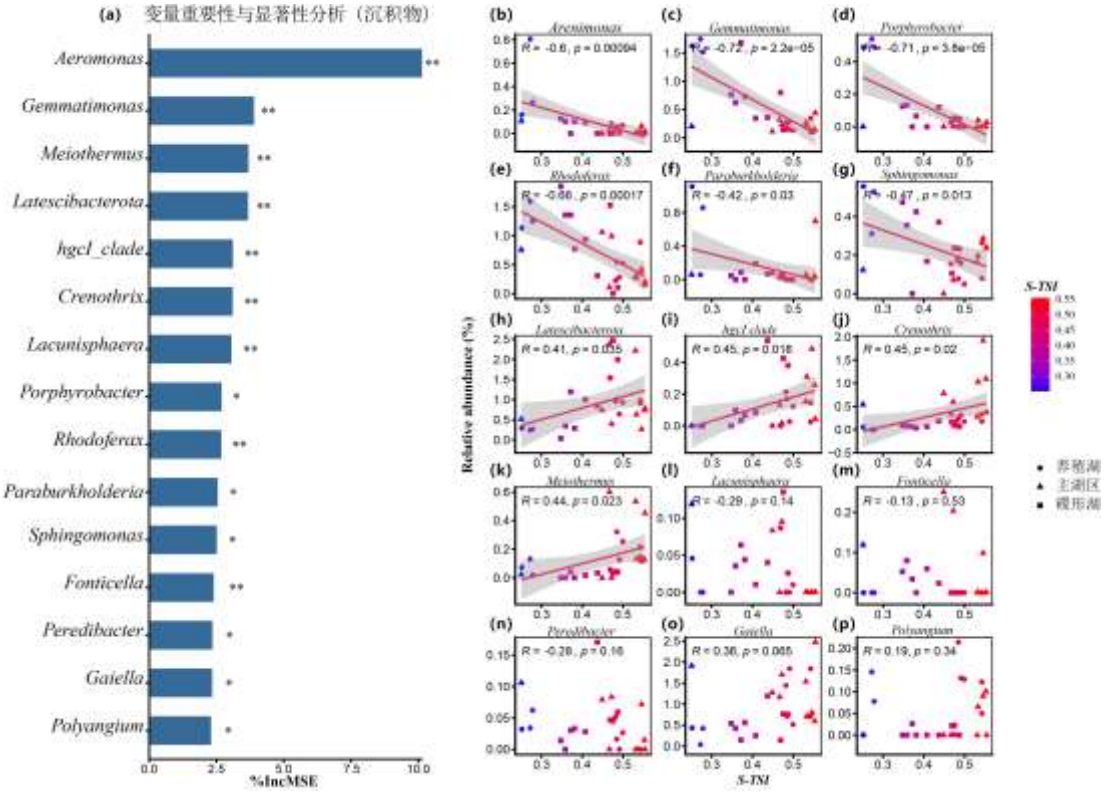

图 5 基于随机森林回归模型识别沉积物中与营养状态指数（*S-TSI*）关联的关键类群（属水平，Top15）（a）及其相对丰度沿 *S-TSI* 梯度的响应特征（b~p）

Fig. 5. Identification of *S-TSI*-associated indicator genera (Top 15) in sediment using random forest regression (a), with corresponding relative abundance response patterns across *S-TSI* (b~p)

2.4 基于机器学习划分的微生物群落 ASVs 生态组

基于 95%分位数回归模型，预测每个 ASVs 达到峰值丰度时对应的 *TSI* 或 *S-TSI* 值，并据此将所有 ASVs 划分到五个不同的生态组（图 7a、b）。结果显示，分别有 51.6%和 15.2%的 ASVs（在 *TSI* 条件下为 1520 个 ASV，*S-TSI* 条件下为 448 个 ASV）被归类到生态组 5，表明这些 ASVs 能够适应高营养水平环境（图 7c、d）。为验证基于 QSR 算法的生态组划分结果的可靠性，采用基于随机森林的机器学习模型进行验证。结果显示，水体和沉积物的生态组划分结果准确率分别达到了 99%和 98%，表明 QSR 算法在区分不同生态组方面具有较高的可靠性和有效性。

2.5 基于生态组的 MBI 指数与传统指数的比较

为评估鄱阳湖的生态状况，采用线性回归评估 MBI 指数与营养状态指数（*TSI* 和 *S-TSI*）之间的一致性，R 值越高表示一致性越强。结果显示，MBI 指数与传统营养指数间均存在显著的相关性，并且一致的趋势。具体而言，浮游细菌的 MBI 指数与 *TSI* 呈显著负相关（$R = -0.86$, $P < 0.001$）（图 8a），沉积物细菌的 MBI 指数与 *S-TSI* 呈显著负相关（$R = -0.9$, $P < 0.001$）（图 8b）。这些结果表明，基于机器学习构建的微生物指数 MBI 在评估水体和沉积物生态状况方面具有较高的敏感性和准确性，展示微生物应用于水生态监测的潜力。

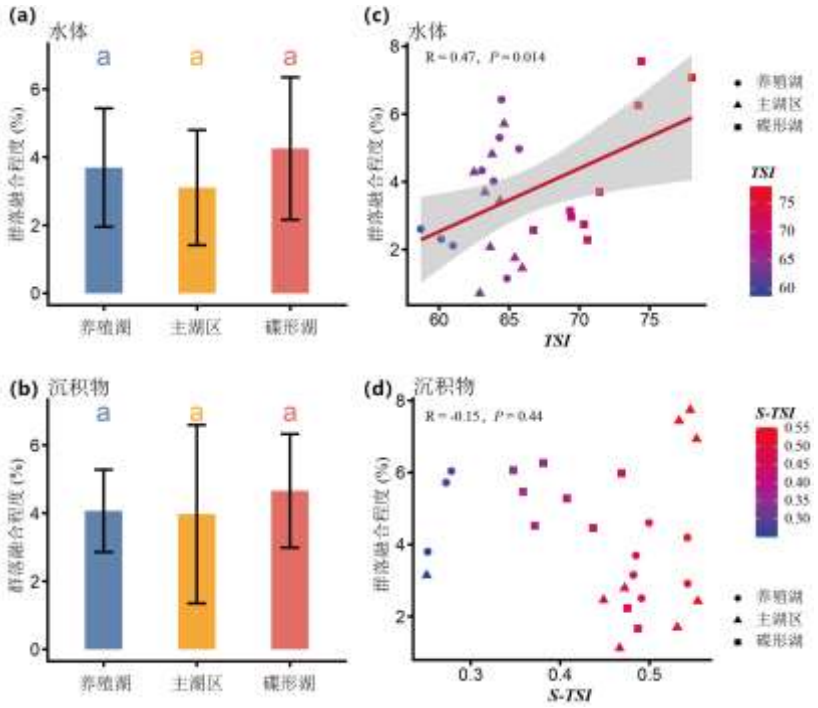

图 6 鄱阳湖不同区域水体（a）和沉积物（b）细菌群落的融合程度；水体（c）和沉积物（d）细菌群落融合程度与营养状态指数（*TSI*/*S-TSI*）的相关性分析。相同字母表示差异无统计学意义（即差异不显著，$P > 0.05$）

Fig. 6. The degree of bacterial community coalescence in water (a) and sediment (b) across different regions of Lake Poyang; Correlation between community coalescence degree and trophic state indices (*TSI* for water; *S-TSI* for sediment) in water (c) and sediment (d). The same letter indicates that the difference is not statistically significant (i.e. the difference is not significant, $P > 0.05$)

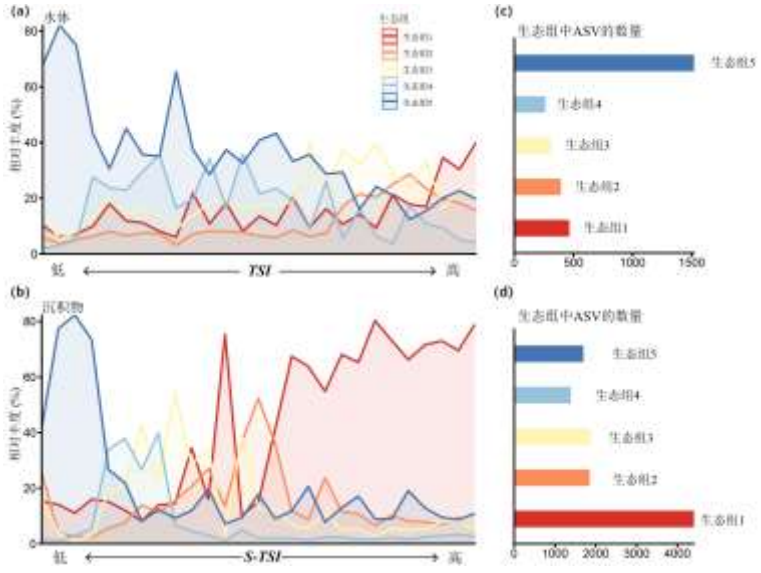

图 7 水体（a）与沉积物（b）细菌群落的生态组组成沿营养状态指数（*TSI*/*S-TSI*）梯度的变化特征；水体（c）和沉积物（d）细菌 ASV 的生态组分布

Fig. 7. Variations in bacterial ecological group composition along trophic state index (*TSI* for water; *S-TSI* for sediment) gradients in water (a) and sediment (b); Distribution of bacterial ASVs across ecological groups in water (c) and sediment (d)

进一步的生态评价结果表明，鄱阳湖不同湖泊区域的生态状况存在显著差异。主湖区水体自上游至下游由“良”（MBI = 0.61）下降至“一般”（MBI = 0.32），而沉积物仅中部样点达“一般”（MBI = 0.36）；碟形湖水质均未达中级（MBI = 0.07~0.38），沉积物评价结果呈 BSH（MBI = 0.43）> ZBH（MBI = 0.25）> CH（MBI = 0.11）的空间差异；在养殖湖中，水体评价结果在“优”至“中”之间波动（MBI = 0.48~0.90），沉积物评价结果则自“差”至“优”不等（MBI = 0.08~0.95）。

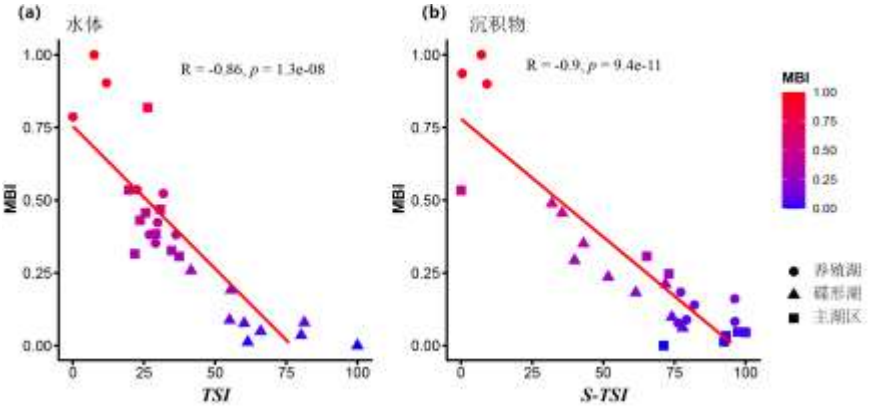

图 8 水体（a）和沉积物（b）微生物指数（MBI）与传统营养状态指数（水体：*TSI*；沉积物：*S-TSI*）相关性分析

Fig. 8. Correlation analysis between the Microbial Biological Index (MBI) for water (a) and sediment (b) and the traditional trophic state indices (*TSI* for water; *S-TSI* for sediment).

3. 讨论

3.1 鄱阳湖水体与沉积物理化特性对营养状态的指示作用

本研究系统分析了鄱阳湖不同湖区水体与沉积物的理化特征，揭示了营养水平在空间上的分异。结果显示，三类湖区水体均已达到中度及以上富营养化水平，*TSI* 从低到高依次为养殖湖、主湖区和碟形湖，其中碟形湖已达重度富营养。碟形湖为典型浅水湖泊，水体容积小、水动力弱，且存在季节性干涸现象，易导致氮磷等营养盐富集和初级生产力升高，从而加剧水体富营养化[44]。相比之下，主湖区与养殖湖水体交换频繁，*TSI* 相对较低。近年来，部分养殖湖实行生态养殖模式，如控制投喂、水质调节、种植水生植物等，有助于降低氮磷积累[22,45,46]。同时，周期性清淤可有效清除沉积有机质，减少内源营养物释放。此外，养殖湖区内部存在明显的空间异质性。不同养殖单元在投喂强度、管理措施及使用年限等方面差异显著，可能对监测结果产生影响[47,48]。尤其在水动力条件较好、物质交换较强的区域，氮磷不易富集，营养水平更易维持在相对较低范围内。值得注意的是，尽管养殖湖整体 *TSI* 低于主湖区，但本研究中水体 Chl α 浓度仍高于主湖区，表明养殖活动对初级生产力仍具有一定促进作用。因此，不同湖区营养状态的评估应结合长时间尺度的动态监测数据，以提高判断的准确性与科学性。

沉积物方面，尽管各湖区 *S-TSI* 差异不显著，但主湖区沉积物的 pH 值和 NO_3^--N 含量显著高于其他区域，反映出营养物质在沉积物中的持续积累。这可能源于主湖区接收大量外源输入，部分物质随水流外排，而有机质和营养物质仍持续沉降于底部，形成潜在的内源负荷。已有研究指出，在浅水湖泊中，沉积物在扰动条件下可释放大量氮磷，加剧水体富营养化进程[49]。碟形湖面积小、水浅、干涸频繁，同样面临营养物质在沉积物中积累与内源释放的风险。Wang 等人[50]在巢湖研究中发现，此类浅水湖泊可保留高达 58% 的氮和 48%的磷，底泥磷释放量可占外源输入的近 10%，反映出内源营养负荷在浅水湖中的重要作用。综合而言，不同湖区水体与沉积物的营养状况受湖泊形态、水动力、外源输入和人类干预等因素共同影响。为有效降低富营养风险，未来湖泊治理应加强对沉积物的长期监测与干预，并结合湖区特征实施差异化的水质管理策略。

3.2 不同类型湖泊水体与沉积环境中微生物群落对营养水平变化的响应

高通量测序结果表明，不同营养水平湖区的水体和沉积物微生物群落结构存在明显差异。群落融合分

析结果显示，总体上水体与沉积物群落融合度较低（大多数低于 5%），说明在同一湖区内部水-沉积物界面存在强烈的环境异质性，而非简单的微生物群落隔离。然而，随着营养水平的升高，水体中源自沉积物的微生物比例显著增加，可能源于富营养条件下，藻类水华发生和风浪扰动增强等因素加剧了沉积物的扰动与再悬浮，从而促使更多沉积源微生物释放并扩散至水体中。

不同湖区微生物群落对营养水平变化的响应也表现出一定的规律性。碟形湖因水浅、易受风浪扰动，增强了沉积物中营养盐和微生物的再悬浮与释放，促使高营养水平敏感菌属（如水体中的 *Polynucleobacter* 和沉积物中 *Crenothrix*）向邻近生态位迁移[44,51]；而对低营养环境具有适应能力的功能菌属（如分解难降解有机物的 *Methylorosula* 和固磷的 *Gemmatimonas*）丰度下降，进一步促进营养物质的单向迁移[52]。养殖湖中，底栖生物（如鱼、螃蟹）掘穴、排泄等活动形成了独特的生物扰动模式，可能通过改变沉积物孔隙结构，促进兼性菌（如水体中 *Rhodobacter* 和沉积物中 *Rhodoferax*）在水-沉积物界面间的扩散与迁移[53]。微生物群落作为生态状况的敏感指示物，其丰度变化可以揭示不同营养水平下的生态过程[54]。例如，低营养水平环境中占优的 *Polymorphobacter* 和 *Arenimonas* 可能参与碳固存和氮循环，而它们在富营养条件下被其它菌群取代，表明在生态系统功能发生了潜在转变。

3.3 基于机器学习的微生物指数评价鄱阳湖生态状况

本研究通过结合机器学习方法，构建了新的微生物指数（MBI），为鄱阳湖的生态状况评价提供了新的视角。研究结果表明，使用机器学习将 ASVs 分配到不同的生态组以界定其生态特征，是对微生物进行无分类生物监测以评估湖泊生态状况的有效策略。本研究中，近 99%的 ASVs 被成功划分至生态组，这些分组结果可以直接用于计算相应的 MBI 指数，以揭示湖泊生态状况变化。同时，研究证实了 MBI 指数与传统营养状态指数（*TSI* 和 *S-TSI*）存在显著的相关性，并且一致的趋势。这种无分类策略近年来已被用于生态特征未知的 OTUs 或 ASVs，用以确定其生态功能或权重，在纤毛虫、有孔虫或线虫等生物类群已有成功案例，这些研究证实了这种策略的准确性和适用性[13,55]。

基于 MBI 的综合评价揭示了鄱阳湖不同湖区的生态状况差异：主湖区（如 PYH3，MBI＝0.6075）和养殖湖（KSH，MBI＝0.9）的水体评价等级相对较高，但其沉积物的生态状况则较差（PYH3，MBI＝0.0423；KSH，MBI＝0.08），说明尽管部分区域水体的富营养化水平处于相对可控的状态，沉积物仍可能作为富营养化“储存库”存在，并在特定条件下释放营养盐，进而加剧水体的富营养化进程[56]。这种水-沉积物间生态状况的非对称性提示沉积物可能成为水体富营养化的潜在源头，这与 Lin[57]等人的研究结果一致，研究指出沉积物在水体富营养化过程中起着重要的反馈作用。在微生物群落响应层面，研究揭示了不同生态位的微生物对富营养化压力的差异化适应。例如，在鄱阳湖水体中，超过 50%的 ASVs 被归类为耐受性物种，这些物种通常出现在富营养化水体中并归类为生态组中的 5 级。相比之下，沉积物中只有 15%的 ASVs 被归类为对富营养化和水污染敏感的物种，显示出微生物群落在富营养化压力下的适应性及多样性。通过这一角度，能够在富营养化早期阶段预警潜在的生态问题，为湖泊管理和保护措施的制定提供更具前瞻性的数据支持。

4. 结论

本研究以鄱阳湖不同湖区（主湖区、碟形湖和养殖湖）为研究区域，采用高通量测序技术分析水体和沉积物中的细菌群落结构，并基于机器学习的无分类策略构建了微生物指数（MBI），并结合传统营养状态指数对不同湖区进行评价，结论如下：

1）基于营养盐参数，碟形湖表现出典型的富营养化特征，而主湖区和养殖湖的营养状态相对较好，显示出湖区间水体和沉积物营养水平的显著空间分异。

2）水体中源自沉积物的细菌比例随着富营养程度显著上升，揭示了水体与沉积物间的群落融合趋势，沉积物可能为水体营养负荷的重要潜在来源。

3）基于机器学习构建的 MBI 指数能够定量反映湖泊生态状况，生态评价结果显示，水体中主湖区与养殖湖生态状况优于碟形湖，而沉积物中生态状况依次为养殖湖 > 碟形湖 > 主湖区。沉积物生态质

量普遍低于水体，可能是内源污染和营养物释放的重要介质。

5. 参考文献

[1] Du BX, Xu LG, Zhang J *et al*. The spatial-temporal characteristics of eutrophication in Poyang Lake and its relationship with the water level. *Research of Environmental Sciences*, 2019, **32**(5): 795-801. DOI: 10.13198/j.issn.1001-6929.2018.10.07. [杜冰雪，徐力刚，张杰等. 鄱阳湖富营养化时空变化特征及其与水位的关系. 环境科学研究，2019，**32**(5)：795-801.]

[2] Xie HY, Hu M, Ji XY *et al*. Water quality evolution characteristics and pollution factor analysis in Poyang Lake from 2011 to 2019. *Environmental Science*, 2022, **43**(12): 5585-5597. DOI: 10.13227/j.hjkx.202112220. [谢慧钰，胡梅，嵇晓燕等. 2011～2019 年鄱阳湖水质演化特征及主要污染因子解析. 环境科学，2022，**43**(12)：5585-5597.]

[3] Dolman AM, Mischke U, Wiedner C. Lake-type-specific seasonal patterns of nutrient limitation in German lakes, with target nitrogen and phosphorus concentrations for good ecological status. *Freshwater Biology*, 2016, **61**: 444-456. DOI: 10.1111/fwb.12718.

[4] Bhateria R, Jain D. Water quality assessment of lake water: a review. *Sustainable Water Resources Management*., 2016, **2**(2): 161-173. DOI: 10.1007/s40899-015-0014-7.

[5] Hering D, Borja A, Carstensen J *et al*. The European water framework directive at the age of 10: a critical review of the achievements with recommendations for the future. *Science of the Total Environment*. 2010, **408**(19): 4007-4019. DOI: 10.1016/j.scitotenv.2010.05.031.

[6] Birk S, Bonne W, Borja A *et al*. Three hundred ways to assess Europe's surface waters: an almost complete overview of biological methods to implement the water framework directive. *Ecological Indicators*, 2012, **18**: 31-41. DOI: 10.1016/j.ecolind.2011.10.009.

[7] Pinto R, Patrício J, Baeta A *et al*. Review and evaluation of estuarine biotic indices to assess benthic condition. *Ecological Indicators*, 2009, **9**(1): 1-25. DOI: 10.1016/j.ecolind.2008.01.005.

[8] Sagova-Mareckova M, Boenigk J, Bouchez A *et al*. Expanding ecological assessment by integrating microorganisms into routine freshwater biomonitoring. *Water Research*, 2021, **191**: 116767. DOI: 10.1016/j.watres.2020.116767.

[9] Liu B, Hou W, Kattel G *et al*. Advancing potential of diatom indicators for stream and lake water quality assessment against nutrient-induced pollution. *Journal of Applied Phycology*, 2023, **35**(2): 721-737. DOI: 10.1007/s10811-023-02911-6.

[10] Jousset A, Bienhold C, Chatzinotas A *et al*. Where less may be more: how the rare biosphere pulls ecosystems strings. *Isme Journal*, 2017, **11**(4): 853-862. DOI: 10.1038/ismej.2016.174.

[11] Falkowski PG, Fenchel T, Delong EF. The microbial engines that drive earth's biogeochemical cycles. *Science*, 2008, **320**(5879): 1034-1039. DOI: 10.1126/science.1153213.

[12] Hering D, Borja A, Jones JI *et al*. Implementation options for DNA-based identification into ecological status assessment under the European water framework directive. *Water Research*, 2018, **138**: 192-205. DOI: 10.1016/j.watres.2018.03.003.

[13] Cordier T, Esling P, Lejzerowicz F *et al*. Predicting the ecological quality status of marine environments from eDNA metabarcoding data using supervised machine learning. *Environmental Science & Technology*, 2017, **51**(16): 9118-9126. DOI: 10.1021/acs.est.7b01518.

[14] Frühe L, Cordier T, Dully V *et al*. Supervised machine learning is superior to indicator value inference in monitoring the environmental impacts of salmon aquaculture using eDNA metabarcodes. *Molecular Ecology*, 2021, **30**(13): 2988-3006. DOI: 10.1111/mec.15434.

[15] Wright MN, Ziegler A. Ranger: a fast implementation of random forests for high dimensional data in c plus and r. *Journal of Statistical Software*, 2017, **77**(1): 1-17. DOI: 10.18637/jss.v077.i01.

[16] Li XY, Wang H, Wu XM *et al*. Characterization and prediction of dissolved oxygen fluctuation in Poyang Lake based on machine learning. *J Lake Sci,* 2025, **37**(3): 915-927. DOI: 10.18307/2025.0328. [李晓瑛，王华，吴小毛等. 基于机器学习的鄱阳湖溶解氧波动特征及预测. 湖泊科学，2025, **37**(3): 915-927.]

[17] Huang AP, Liu XB, Peng WQ *et al*. Spatiotemporal characteristics, influencing factors and evolution laws of water exchange capacity of Poyang Lake. *Journal of Hydrology*, 2022, **609**: 127717. DOI: 10.1016/j.jhydrol.2022.127717.

[18] Huang SX, Zhang K, Lin Q *et al*. Abrupt ecological shifts of lakes during the anthropocene. *Earth-Science Reviews*, 2022, **227**:

103981. DOI: 10.1016/j.earscirev.2022.103981.

[19] Li HH. Changes of water level in main lake area of Poyang Lake and in dish-shaped sub-lake and their impacts on water quality. *Resources And Environment In The Yangtze Basin*, 2018, **27**(0): 1298-1306. DOI: 10.11870/cjlyzyyhj201806013. [李海辉. 鄱阳湖主湖区与碟形湖水位变化及其对水质的影响. 长江流域资源与环境，2018，**27**(0)：1298-1306.]

[20] Niu HL, Wang MX, Shi, YB *et al*. Numbers, spatial distribution and characteristics of saucer-shaped lakes in Lake Poyang wetland and influencing factors. *J Lake Sci*, 2025, **37**(5). DOI 10.18307/2025.0552. [牛汇林，王明星，施以兵等. 鄱阳湖洪泛区碟形湖数量、空间分布及其影响因素探析. 湖泊科学，2025，**37**(5)]

[21] 胡振鹏,张祖芳,刘以珍等. 碟形湖在鄱阳湖湿地生态系统的作用和意义. .江西水利科技，2015，**41**(5)：317-323.

[22] Kamalam BS, Pandey PK. Nutrition and environment interactions in aquaculture. 2022: 407-422. DOI: 10.1007/978-981-19-5500-6_27.

[23] Zhao DY, Cao XY, Huang R *et al*. The heterogeneity of composition and assembly processes of the microbial community between different nutrient loading lake zones in Taihu Lake. *Applied Microbiology and Biotechnology*, 2017, **101**(14): 5913-23. DOI: 10.1007/s00253-017-8327-0.

[24] Zeng J, Jiao CC, Zhao DY *et al*. Patterns and assembly processes of planktonic and sedimentary bacterial community differ along a trophic gradient in freshwater lakes. *Ecological Indicators*, 2019, **106**: 105491. DOI: 10.1016/j.ecolind.2019.105491.

[25] Bolyen E, Rideout JR Dillon MR *et al*. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, 2019, **37:** 852-857. DOI: 10.1038/s41587-019-0209-9.

[26] Callahan B, McMurdie P, Rosen M *et al*. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 2016, **13**: 581-583. DOI: 10.1038/nmeth.3869.

[27] uast C, Pruesse E, Yilmaz P *et al*. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 2013, 41: 590-596. DOI: 10.1093/nar/gks1219.

[28] Bokulich NA, Kaehler BD, Rideout JR *et al*. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*, 2018, **6**(1):90. DOI: 10.1186/s40168-018-0470-z.

[29] Zhang YL, Zhou YQ, Shi K *et al*. Optical properties and composition changes in chromophoric dissolved organic matter along trophic gradients: implications for monitoring and assessing lake eutrophication. *Water Research*, 2018, **131**: 255-263. DOI: 10.1016/j.watres.2017.12.051.

[30] Yu PJ, Liu SW, Zhang L *et al*. Selecting the minimum data set and quantitative soil quality indexing of alkaline soils under different land uses in northeastern China. *Science of the Total Environment*, 2018, **616-617**: 564-571. DOI: 10.1016/j.scitotenv.2017.10.301.

[31] Shenhav L, Thompson M, Joseph TA *et al*. Feast: fast expectation-maximization for microbial source tracking. *Nature Methods*, 2019, **16**(7): 627-32. DOI: 10.1038/s41592-019-0431-x.

[32] Rillig MC, Antonovics J, Caruso T *et al*. Interchange of entire communities: microbial community coalescence. *Trends in Ecology & Evolution*, 2015, **30**(8): 470-6. DOI: 10.1016/j.tree.2015.06.004.

[33] Liu T, Zhang AN, Wang JW *et al*. Integrated biogeography of planktonic and sedimentary bacterial communities in the Yangtze river. *Microbiome*, 2018, **6**(1): 16. DOI: 10.1186/s40168-017-0388-x.

[34] Rao SJ. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. *Journal of the American Statistical Association*, 2003, **98**(461): 257-258. DOI: 10.1198/jasa.2003.s263.

[35] Wang WJ, Yan J. Shape-restricted regression splines with r package splines2. *Journal of Data Science*, 2021, **19**(3): 498-517. DOI: 10.6339/21-jds1020.

[36] Borja A, Franco J, Pérez V. A Marine Biotic Index to Establish the Ecological Quality of Soft-Bottom Benthos Within European Estuarine and Coastal Environments. Marine Pollution Bulletin, 2000, 40(12): 1100-1114. DOI: 10.1016/S0025-326X(00)00061-8.

[37] Li XH, Li FL, Min XY *et al.* Embracing eDNA and machine learning for taxonomy-free microorganisms biomonitoring to assess the river ecological status. *Ecological Indicators*, 2023, **155**: 110948. DOI: 10.1016/j.ecolind.2023.110948.

[38] Consuegra S, Aylagas E, Borja Á *et al.* Environmental status assessment using DNA metabarcoding: towards a genetics based marine biotic index (gAMBI). *PLoS ONE*, 2014, **9**(3): e90529. DOI: 10.1371/journal.pone.0090529.

[39] Dixon P. VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, 2003, **14**(6): 927-930. DOI: 10.1111/j.1654-1103.2003.tb02228.x.

[40] Zhang L, Huettmann F, Liu SR *et al.* Classification and regression with random forests as a standard method for presence-only data SDMs: A future conservation example using China tree species. *Ecological Informatics*, 2019, **52**: 46-56. DOI: 10.1016/j.ecoinf.2019.05.003.

[41] Delgado-Baquerizo M, Maestre FT, Reich PB *et al.* Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nature Communications*, 2016, **7**: 10541. DOI: 10.1038/ncomms10541.

[42] Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. *Biometrical Journal*, 2008, **50**(3): 346-363. DOI: 10.1002/bimj.200810425.

[43] Ginestet, C. ggplot2: Elegant Graphics for Data Analysis. 2011, **174**: 245. DOI: 10.1111/j.1467-985X.2010.00676_9.x.

[44] Zhou J, Leavitt P R, Zhang YB *et al*. Anthropogenic eutrophication of shallow lakes: Is it occasional? *Water Research*, 2022, **221**: 118728. DOI: 10.1016/j.watres.2022.118728.

[45] Paolacci S, Stejskal V, Toner D *et al.* Integrated multitrophic aquaculture; Analysing contributions of different biological compartments to nutrient removal in a duckweed-based water remediation system. *Plants Basel*, 2022, **11**(22): 3103. DOI: 10.3390/plants11223103.

[46] Wang HH, Fu YL, Xu XD, *et al*. Assessment on the efficiency of *Micropterus salmoides* eco-cage culture pattern in Poyang Lake District. *Chinese Journal of Eco-Agriculture*, 2013, **21**(8): 1009-1015. DOI: 10.3724/SP.J.1011.2013.01009. [王海华，傅义龙，徐先栋等. 鄱阳湖区加州鲈网箱生态养殖模式的效果评估. 中国生态农业学报，2013，**21**(8)：1009-1015.]

[47] Liu W, Chen YF, Tan QY *et al*. Distributional characteristics and evaluation of nitrogen, phosphorus and organic matter inthe surface sediments of rural ponds around the western part of Lake Poyang in summer. *J Lake Sci*, 2023, **35**(1): 192-202. DOI: 10.18307/2023.0113. [刘伟，褚一凡，谭启洋等. 鄱阳湖西侧周边农村水塘夏季表层沉积物氮、磷、有机质分布特征及评价. 湖泊科学，2023，**35**(1)：192-202.]

[48] Zhang Y, Li TJ, Li GZ *et al*. Profiling sediment bacterial communities and the response to pattern-driven variations of total nitrogen and phosphorus in long-term polyculture ponds. *Frontiers in Marine Science*, 2024, **11**: 1403909. DOI: 10.3389/fmars.2024.1403909.

[49] Yao X, Ding R, Zhou Y *et al*. How internal nutrient loading forms in shallow lakes: Insights from benthic organic matter mineralization. *Water Research*, 2023, **245**: 120544. DOI: 10.1016/j.watres.2023.120544.

[50] Wang Y, Kong X, Peng Z *et al*. Retention of nitrogen and phosphorus in Lake Chaohu, China: implications for eutrophication management. *Environmental Science and Pollution Research*, 2020, **27**: 41488-41502. DOI: 10.1007/s11356-020-10024-7.

[51] Qin BQ, Yang GJ, Ma JR *et al*. Spatiotemporal changes of cyanobacterial bloom in large shallow eutrophic lake Taihu, China. *Frontiers in Microbiology*, 2018, **9:** 451. DOI: 10.3389/fmicb.2018.00451.

[52] Qin BQ, Zhou J, Elser JJ *et al*. Water depth underpins the relative roles and fates of nitrogen and phosphorus in lakes. *Environmental Science & Technology*, 2020, **54**(6): 3191-3198. DOI: 10.1021/acs.est.9b05858.

[53] You BS, Wang TC, Fan CX *et al*. Effects of sediment resuspension on aqueous nutrient loading in grass type zone of lake Taihu. *Environmental Sciences*, 2008, **29**(01): 26-31. DOI:10.13227/j.hjkx.2008.01.029. [尤本胜，王同成，范成新等. 太湖草型湖区沉积物再悬浮对水体营养盐的影响. 环境科学，2008，**29**(1)：26-31.]

[54] Zhang L, Yuan LJ, Xiang JJ *et al.* Response of the microbial community structure to the environmental factors during the extreme flood season in Poyang Lake, the largest freshwater lake in China. *Frontiers in Microbiology*, 2024, **15**: 1362968. DOI: 10.3389/fmicb.2024.1362968.

[55] Apothéloz-Perret-Gentil L A-O, Cordonier A, Straub F *et al*. Taxonomy-free molecular diatom index for high-throughput eDNA biomonitoring. *Molecular Ecology Resources*, 2017, **17**(6): 1231-1242. DOI: 10.1111/1755-0998.12668.

[56] Le C, Zha Y, Li Y *et al*. Eutrophication of lake waters in China: cost, causes, and control. *Environmental management*, 2010, **45**(4): 662-668. DOI: 10.1007/s00267-010-9440-3.

[57] Lin SS, Shen SL, Zhou A *et al*. Assessment and management of lake eutrophication: a case study in Lake Erhai, China. *Science of the Total Environment*, 2021, **751**: 141618. DOI: 10.1016/j.scitotenv.2020.141618.