

人工智能在乡镇废水处理中的创新应用

李巧艳

山东省青岛市莱西市店埠镇人民政府, 山东 青岛

收稿日期: 2025年11月16日; 录用日期: 2025年12月22日; 发布日期: 2025年12月31日

摘要

面向乡镇污水治理的复杂性与不确定性, 本文系统梳理了人工智能在污染物精准去除、水质监测预警与工艺优化控制中的创新应用与可行路径。针对乡镇污水水量水质波动大、氮磷负荷高、季节性差异显著等特征, 构建了以监督式与无监督式学习为主的技术路线: 通过回归与分类模型实现COD、氨氮、总氮、总磷等关键指标的预测与分类决策, 借助聚类、异常检测与降维提升工况识别、冲击负荷预警与可视化分析能力, 并以特征工程与正则化、集成学习增强小样本与不平衡样本条件下的泛化性能。在工程实践层面, 强化数据治理与可解释性成为落地关键: 以标准化采集、清洗、转换、特征选择与降维构建高质量数据底座, 采用K均值、DBSCAN、PCA等方法支撑实时监测与运维闭环; 依托SHAP、部分依赖图与局部解释提升黑箱模型透明度, 结合白箱模型与规则库满足监管可追溯要求。与此同时, 本文提出伦理与前瞻视角下的治理框架: 以模型透明度、问责机制与全生命周期环境足迹评估保障可持续应用, 通过数据公平与隐私安全治理弥合城乡差异; 面向中长期, 倡导数据驱动与机理模型深度融合的混合建模, 提升模型在工况突变下的鲁棒性与可迁移性, 并以多源真实运行数据持续校准, 推动“预测-优化-控制”闭环在更多乡镇场景中规模化复用, 最终实现环境效益、经济效率与社会接受度的协同提升。

关键词

乡镇污水治理, 人工智能模型, 数据预处理, 可解释性

Innovative Applications of Artificial Intelligence in Rural Wastewater Treatment

Qiaoyan Li

People's Government of Dianbu Town, Laixi City, Qingdao City, Shandong Province, Qingdao Shandong

Received: November 16, 2025; accepted: December 22, 2025; published: December 31, 2025

Abstract

Addressing the complexity and uncertainty of township wastewater treatment, this paper

文章引用: 李巧艳. 人工智能在乡镇废水处理中的创新应用[J]. 人工智能与机器人研究, 2026, 15(1): 58-64.
DOI: 10.12677/airr.2026.151007

systematically reviews innovative applications of artificial intelligence in pollutant removal, real-time monitoring, and process optimization. Considering the challenges of large fluctuations in flow and water quality, high nitrogen and phosphorus loads, and pronounced seasonal variability in township wastewater, we propose a technical route centered on supervised and unsupervised learning: regression and classification models for predicting and classifying key indicators such as COD, ammonia nitrogen, total nitrogen, and total phosphorus; clustering, anomaly detection, and dimensionality reduction for condition identification, shock load early warning, and visualization; and feature engineering enhanced by regularization and ensemble methods to improve generalization under small-sample and imbalanced settings. For field deployment, data governance and interpretability are critical enablers: a standardized pipeline of acquisition, cleaning, transformation, feature selection, and dimensionality reduction builds a high-quality data foundation; K-means, DBSCAN, PCA, and related methods underpin real-time monitoring, maintenance closed loops, and operator interpretability; explainable AI with SHAP and partial dependence plots complements white-box models and rule bases to meet regulatory traceability. Meanwhile, this paper proposes a governance framework from an ethical and forward-looking perspective: ensuring sustainable applications through model transparency, accountability mechanisms, and full life-cycle environmental footprint assessment; and bridging urban-rural disparities through data fairness and privacy-preserving governance. Looking to the medium and long term, we advocate for hybrid modeling that deeply integrates data-driven and mechanistic models to enhance the robustness and transferability of models under sudden changes in operating conditions. This approach, continuously calibrated with multi-source real-world operational data, will promote the large-scale reuse of the “prediction-optimization-control” closed loop in more rural scenarios, ultimately achieving synergistic improvements in environmental benefits, economic efficiency, and social acceptance.

Keywords

Rural Sewage Treatment, Artificial Intelligence Model, Data Preprocessing, Interpretability

Copyright © 2026 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

乡镇污水治理作为保护农村居民健康与生态环境的关键环节，通过有效削减排入周边水体的污染物负荷，对维护乡村水生态安全具有重要作用[1]。随着我国乡村振兴战略推进和乡镇产业快速发展，农村污水产生特征与处理需求正发生显著变化——人口集聚带来污水量持续增长，特色农业与小型加工业发展导致污染物类型日趋多元，传统治理模式在应对这些新挑战时逐渐显现出适应性不足的问题。当前主流处理技术多沿用城市污水治理思路，对乡镇污水特有的低浓度波动大、氮磷等营养物占比高、季节性排放差异显著等特点考虑不足，存在监测指标覆盖不全、系统调控依据有限等问题，由此加剧了处理过程中营养物质向周边水体的异常排放风险。乡镇污水治理的复杂性既来源于降雨径流、畜禽养殖等自然与人为活动的动态干扰，也与处理设施工艺适配性、运维管理水平等直接相关；这些因素相互交织，导致处理系统运行效果存在显著不确定性，具体表现为进水水量水质的随机波动(如雨季污水稀释、旱季浓度骤升)、处理效率的不稳定(如生物反应受温度影响显著)以及资源能源回收的低效性。面对乡镇地区污染物类型多样、处理成本敏感且需兼顾生态与经济效益的现实需求，亟需探索更贴合乡村实际的治理技术体系与管理模式[2]，通过精准化参数监测、适应性工艺优化及资源高效循环利用，切实提升乡镇污水治理的综合效能。

近年来,人工智能技术以其强大的数据处理与模式识别能力,逐步渗透至环境工程领域,尤其在废水处理过程中展现出前所未有的潜力[3]。人工智能技术通过建立数学预测模型,可利用历史数据估算这些参数。运用统计数学方法预测废水特性及污水处理厂运行效果的主要优势,在于能全面解析其背后的物理化学与生物过程。此外,将这些模型与化学反应相结合,还能深化对结果解读的理解。尽管已有大量研究聚焦于 AI 在污染物去除、水质监测与故障诊断等方面的应用,然而,随着技术实践的深入,一系列新型问题与系统局限性也逐渐浮出水面[4]。本文旨在重新审视人工智能在废水处理中的角色定位,从模型构建、数据治理、系统整合到伦理与社会影响等多个维度展开探讨,以期对未来研究提供批判性视角与创新路径。

2. 面向乡镇低质量数据的 AI 治理与可解释技术路径

在乡镇污水处理这一特殊而复杂的应用场景中,人工智能模型从理论构想走向工程实践的征途,始终伴随着数据稀疏、标注成本高昂以及物理化学机理认知模糊等核心挑战的掣肘。这些挑战并非孤立存在,而是相互交织,共同构成了阻碍 AI 技术深度赋能乡镇水务的“最后一公里”难题[4]。具体而言,乡镇污水处理厂普遍存在监测设备陈旧、连续运行稳定性差的问题,导致关键水质参数(如 COD、氨氮)数据存在大量随机缺失与异常片段;其进水来源分散且受季节性农耕、节假日人口流动影响巨大,导致水质、水量波动剧烈[5][6],数据分布极不均衡,难以形成具有统计意义的高质量样本集;更重要的是,专业运维人员的长期匮乏,使得为历史数据打上精准、可靠的工况标签(如“污泥膨胀”、“硝化抑制”等)几乎成为一种奢望,严重制约了监督学习模型的训练与应用。面对这一现实困境,本文毅然将技术探讨的焦点从宽泛的应用枚举,收敛至“低质量数据治理”与“可解释人工智能”这两大最具基础性与紧迫性的技术路径上,旨在构建一个从数据源头治理到决策终端透明的全链路、可信赖的 AI 方法体系[5]。

乡镇污水处理数据的“低质量”特征主要体现在“不全”、“不匀”和“不准”三个方面。为解决这些问题,我们提出了一套分层递进的数据治理框架,其核心是前沿的数据增强技术与面向特定场景的迁移学习策略的有机结合。在数据增强层面,传统的插值、平滑等方法对于处理大规模连续缺失或生成符合复杂物理规律的新样本往往力不从心[7]。为此,我们引入了基于时序生成对抗网络(Time-series Generative Adversarial Network, T-GAN)的增强方法。T-GAN 由一个生成器(Generator)和一个判别器(Discriminator)构成。生成器通过学习有限历史水质时序数据(如进水 COD、流量、pH 的逐时序列)的潜在分布,生成与真实数据在统计特性、动态模式和噪声分布上高度一致的合成数据;判别器则全力甄别输入数据来自真实样本还是生成样本[6]。二者在对抗博弈中不断优化,最终使生成器能够产出“以假乱真”的高质量时序数据。这一技术的直接价值在于,它能有效弥补因设备故障或采样频率低造成的样本数量不足,为后续模型训练提供丰沛的“数据燃料”。例如,在某镇级污水厂的实际应用中,我们利用其三个月不完整的监测数据,通过 T-GAN 生成了长达一年的、涵盖不同季节特征的合成数据集,使可供模型训练的有效样本量扩大了四倍。然而,仅仅增加数据量仍不足以保证模型在特定乡镇厂站的优异表现。迁移学习策略的引入,旨在解决目标站点(目标域)标签数据极度稀缺的根本性难题。我们的策略是“先广博后精深”:首先,在数据采集条件优越、监测体系完善的城市大型污水处理厂(源域),利用其海量的、标注齐全的历史运行数据[8],对一个通用的水质预测模型进行预训练。此阶段的目标是让模型掌握污水处理过程中普遍的、共性的动态规律,如底物降解与微生物增长的基本动力学、溶解氧对硝化反应的关键影响等。然后,将这个在“大城市”练就的“通才”模型,迁移到目标乡镇污水厂(目标域),仅利用该厂站有限的本地数据(可能仅占源域数据的百分之几)进行微调。微调过程本质上是让模型“入乡随俗”,通过少量几次迭代训练,使其适应目标厂站独特的进水特性、工艺参数和设备条件。实践表明,该策略在乡镇

进水 COD 和氨氮的预测任务中表现卓越, 仅需使用 30% 的本地标注数据, 即可使模型达到甚至超过使用 100% 本地数据从头训练的模型的预测精度(判定系数 R^2 稳定在 0.85 以上), 极大地降低了对昂贵标注数据的依赖[9]。

3. AI 驱动的水污染物精准去除与优化控制

3.1. 有机污染物的智能预测与工艺优化

面向成分复杂、降解路径多样的有机污染物(如抗生素、工业染料、新兴微污染物), 人工神经网络与集成学习已成为降解/吸附过程预测与优化的核心工具[10]。以水凝胶吸附孔雀石绿、亚甲蓝等典型染料为例, 径向基函数网络(RBF)凭借局部逼近特性, 拟合效果与实验高度一致; 而在涉及多污染物共存、竞争吸附或高级氧化等强非线性场景中, 多层感知器(MLP)凭借更深层次结构展现更强的泛化能力。大量研究显示, 复杂任务下模型预测与实测的决定系数普遍达到 $R^2 > 0.98$, 显著优于传统响应面法与线性模型。针对乡镇站点常见的冲击负荷(如染料/表面活性剂波动), AI 可基于历史运行数据进行快速聚类, 精准匹配相似“历史工况”, 并调用经验证的最佳控制参数模板, 从而缩短恢复稳定时间、增强抗冲击能力[11][12]。

3.2. 无机污染物与营养盐的协同优化控制

在无机污染物控制方面, RBF 网络与随机森林对重金属离子去除效率的预测表现稳健, 模型预测与实测的相关系数多次验证超过 0.99。更重要的是, 这类“黑箱”模型能够揭示 pH、反应温度、吸附剂投加量等关键参数间的交互效应[13][14], 为工艺优化提供可操作的机理洞察。面向营养盐(氮、磷)去除这一难点, AI 实现了从“单点优化”到“协同调控”的跨越: 以前馈神经网络精准模拟生化过程, 结合自组织映射(SOM)对运行状态进行可视化聚类, 系统可快速锁定同时满足高效脱氮与深度除磷的最佳运行窗口[15]。由此, 在确保出水总氮、总磷稳定达标前提下, 智能平衡硝化—反硝化与生物除磷对溶解氧、碳源等资源的需求, 实现达标与降耗的最优折中。

3.3. 从预测到控制: 构建闭环智能运行体系

面向污水处理的复杂工况, 将高精度预测模型与专家规则库、智能优化器深度耦合, 构建“预测—优化—控制”闭环, 使工艺在秒级内完成从状态识别到执行指令的链路。在线数据经清洗对齐后, 模型滚动预测出水水质与关键状态, 优化器以能耗、药耗、碳排为目标函数, 结合曝气、回流、污泥浓度、药剂投加等约束求解最优动作; 执行结果由在线仪表与软测量回传, 触发偏差校正与再优化, 形成自适应的反馈校正机制。该体系兼顾多变量耦合与大时滞, 在乡镇污水场景中显著提升对来水波动的适应性, 保障稳定达标与能耗最小化并行。为保障安全与可运维, 闭环系统采用分层自治: 现场边缘控制负责快速执行与异常限幅, 云端大脑进行全局优化与策略下发, 并通过数字孪生与仿真预演对策略进行离线验证与在线校核。结合异常预警、远程操控与移动巡检, 实现“监测—诊断—处置”的一体化联动; 同时以知识图谱沉淀专家经验, 固化为可解释的规则与约束, 在极端天气与冲击负荷下维持韧性与合规性[16](见图 1)。

图 1 展示了一个智能环境监测与决策系统的整体架构。系统始于“多源数据采集”与“感知输入层”, 对有机污染物、重金属离子、营养盐等多种污染指标进行数据预处理。随后利用包括 ANN/MLP、RBF 在内的“AI 模型集群”进行预测与分析, 并结合“专家规则库”与“多目标优化器”实现诊断与优化。最终通过“SOM 可视化聚类”模块呈现结果, 形成一个从数据感知到智能决策的完整“智能闭环”, 驱动环保决策的“智能决策大脑”



Figure 1. Architecture diagram of intelligent monitoring and decision-making closed-loop system for sewage treatment
图 1. 污水处理智能监测与决策闭环系统架构图

4. 人工智能在乡镇污水处理的伦理与前瞻：通向可持续水管理的智能之路

4.1. 技术赋能与工程价值

AI 正成为提升污染物去除效率与工艺韧性的核心引擎。面向成分复杂、路径多样的有机污染物(如抗生素、工业染料、新兴微污染物)，深度神经网络与集成学习显著提升了降解/吸附等过程的预测精度与泛化能力；在存在多污染物共存、竞争吸附或高级氧化的强非线性场景下，模型可稳定输出可操作的调控策略。针对冲击负荷(染料/表面活性剂波动)带来的扰动，基于工况聚类与相似历史模板复用的快速匹配机制，可在分钟级给出最优参数组合，显著缩短恢复时间并增强系统抗扰动能力。在无机污染物与营养盐协同优化方面，模型不仅提供高精度的去除效率预测，还能量化 pH、温度、吸附剂/碳源投加等关键因子的交互效应，支撑更精细的工艺调控。结合前馈神经网络对生化过程与自组织映射对运行状态的可视化聚类，系统可快速锁定兼顾高效脱氮与深度除磷的运行窗口，在确保总氮、总磷稳定达标的同时，优化曝气、回流、污泥浓度与药剂投加，实现能耗与药耗的协同降低，形成“预测 - 优化 - 控制”闭环，为厂站长周期稳定、经济高效运行提供可复制的工程价值[17]。

4.2. 伦理挑战与治理要点

推进 AI 应用需同步构建伦理与治理框架：一是透明度与问责，AI 决策(如优化曝气以平衡能耗与处理效率)需具备可解释、可追溯的决策逻辑，确保符合法规并与社区预期一致；二是环境足迹，在推广智能曝气控制等节能应用的同时，应评估数据中心、传感器网络等基础设施的能耗与电子垃圾影响[18]，开展全生命周期评估(LCA)；三是数据偏见与公平性，避免仅以基础设施完善的富裕地区“优质数据”训练模型，需整合多元、具代表性的数据集以确保公平；四是隐私与安全，海量数据的采集与使用需建立清

晰的数据治理规范，界定所有权、使用权与访问权限，切实保护个人隐私与社区权益[19]-[21]。

4.3. 前瞻方向与落地路径

为克服当前局限并充分释放 AI 潜力，可从四方面发力：发展机理 - 数据融合的混合建模，以更低数据需求融合“黑箱”预测与“白箱”机理解释；提升可解释性，增强模型在突变工况下的适应性与对输入波动的稳定性，推进可解释 AI (XAI)与标准化架构选择；以真实污水厂运行数据为输入，提升预测精度与控制的实时性；面向全链条智能决策，贯通废水排放 - 收集 - 处理 - 污泥管理 - 环境影响 - 经济效益 - 政策制定，并在报告与使用数据时完整标注来源、地理背景、工艺环境与本体特征，确保可追溯性与结果可靠性。

5. 结论

面向乡镇污水治理的现实需求，需以低质量数据治理与可解释人工智能为主线，构建适配乡村场景的 AI 方法体系。同时，引入可解释性机制与专家规则库，将黑箱预测转化为可审计、可追责的运维决策，增强系统韧性与合规性。最终，以数据 - 模型 - 控制 - 治理的一体化路径，推动乡镇污水从经验驱动走向智能驱动、从被动应对转向主动优化，为乡村振兴与水生态安全提供可持续的技术支撑。

参考文献

- [1] Aghdam, E., Mohandes, S.R., Manu, P., Cheung, C., Yunusa-Kaltungo, A. and Zayed, T. (2023) Predicting Quality Parameters of Wastewater Treatment Plants Using Artificial Intelligence Techniques. *Journal of Cleaner Production*, **405**, Article 137019. <https://doi.org/10.1016/j.jclepro.2023.137019>
- [2] Alam, G., Ihsanullah, I., Naushad, M. and Sillanpää, M. (2022) Applications of Artificial Intelligence in Water Treatment for Optimization and Automation of Adsorption Processes: Recent Advances and Prospects. *Chemical Engineering Journal*, **427**, Article 130011. <https://doi.org/10.1016/j.cej.2021.130011>
- [3] Altowayti, W.A.H., Allozy, H.G.A., Shahir, S., Goh, P.S. and Yunus, M.A.M. (2019) A Novel Nanocomposite of Aminated Silica Nanotube (MWCNT/Si/NH₂) and Its Potential on Adsorption of Nitrite. *Environmental Science and Pollution Research*, **26**, 28737-28748. <https://doi.org/10.1007/s11356-019-06059-0>
- [4] Altowayti, W.A.H., Shahir, S., Othman, N., Eisa, T.A.E., Yafouz, W.M.S., Al-Dhaqm, A., et al. (2022) The Role of Conventional Methods and Artificial Intelligence in the Wastewater Treatment: A Comprehensive Review. *Processes*, **10**, Article 1832. <https://doi.org/10.3390/pr10091832>
- [5] Alwis, L.S.M., Sun, T. and Grattan, K.T.V. (2016) Fibre Grating-Based Sensor Design for Humidity Measurement in Chemically Harsh Environment. *Procedia Engineering*, **168**, 1317-1320. <https://doi.org/10.1016/j.proeng.2016.11.359>
- [6] Asad, S., Amoozegar, M.A., Pourbabaee, A.A., Sarbolouki, M.N. and Dastgheib, S.M.M. (2007) Decolorization of Textile Azo Dyes by Newly Isolated Halophilic and Halotolerant Bacteria. *Bioresource Technology*, **98**, 2082-2088. <https://doi.org/10.1016/j.biortech.2006.08.020>
- [7] Asadi, A., Verma, A., Yang, K. and Mejabi, B. (2017) Wastewater Treatment Aeration Process Optimization: A Data Mining Approach. *Journal of Environmental Management*, **203**, 630-639. <https://doi.org/10.1016/j.jenvman.2016.07.047>
- [8] Asadi, M., Guo, H. and McPhedran, K. (2020) Biogas Production Estimation Using Data-Driven Approaches for Cold Region Municipal Wastewater Anaerobic Digestion. *Journal of Environmental Management*, **253**, Article 109708. <https://doi.org/10.1016/j.jenvman.2019.109708>
- [9] Ba-Alawi, A.H., Loy-Benitez, J., Kim, S. and Yoo, C. (2022) Missing Data Imputation and Sensor Self-Validation Towards a Sustainable Operation of Wastewater Treatment Plants via Deep Variational Residual Autoencoders. *Chemosphere*, **288**, Article 132647. <https://doi.org/10.1016/j.chemosphere.2021.132647>
- [10] Ba-Alawi, A.H., Al-masni, M.A. and Yoo, C. (2023) Simultaneous Sensor Fault Diagnosis and Reconstruction for Intelligent Monitoring in Wastewater Treatment Plants: An Explainable Deep Multi-Task Learning Model. *Journal of Water Process Engineering*, **55**, Article 104119. <https://doi.org/10.1016/j.jwpe.2023.104119>
- [11] Bagheri, M., Mirbagheri, S.A., Ehteshami, M. and Bagheri, Z. (2015) Modeling of a Sequencing Batch Reactor Treating Municipal Wastewater Using Multi-Layer Perceptron and Radial Basis Function Artificial Neural Networks. *Process Safety and Environmental Protection*, **93**, 111-123. <https://doi.org/10.1016/j.psep.2014.04.006>

-
- [12] Bahramian, M., Dereli, R.K., Zhao, W., Giberti, M. and Casey, E. (2023) Data to Intelligence: The Role of Data-Driven Models in Wastewater Treatment. *Expert Systems with Applications*, **217**, Article 119453. <https://doi.org/10.1016/j.eswa.2022.119453>
- [13] Beyaztaş, U. (2021) Prediction of Copper Ions Adsorption by Attapulgite Adsorbent Using Tuned-Artificial Intelligence Model. *Chemosphere*, **276**, Article 130162.
- [14] Bolón-Canedo, V., Morán-Fernández, L., Cancela, B. and Alonso-Betanzos, A. (2024) A Review of Green Artificial Intelligence: Towards a More Sustainable Future. *Neurocomputing*, **599**, Article 128096. <https://doi.org/10.1016/j.neucom.2024.128096>
- [15] Bozkurt, H., van Loosdrecht, M.C.M., Gernaey, K.V. and Sin, G. (2016) Optimal WWTP Process Selection for Treatment of Domestic Wastewater—A Realistic Full-Scale Retrofitting Study. *Chemical Engineering Journal*, **286**, 447-458. <https://doi.org/10.1016/j.cej.2015.10.088>
- [16] Bui Hamanh, B.H., Perng Yuanshing, P.Y. and Duong Huonggiangthi, D.H. (2016) The Use of Artificial Neural Network for Modeling Coagulation of Reactive Dye Wastewater Using Cassia Fistula Linn. Gum.
- [17] Bustillo-Lecompte, C.F. and Mehrvar, M. (2017) Treatment of Actual Slaughterhouse Wastewater by Combined Anaerobic-Aerobic Processes for Biogas Generation and Removal of Organics and Nutrients: An Optimization Study towards a Cleaner Production in the Meat Processing Industry. *Journal of Cleaner Production*, **141**, 278-289. <https://doi.org/10.1016/j.jclepro.2016.09.060>
- [18] Cai, Y., Zaidi, A.A., Shi, Y., Zhang, K., Li, X., Xiao, S., *et al.* (2019) Influence of Salinity on the Biological Treatment of Domestic Ship Sewage Using an Air-Lift Multilevel Circulation Membrane Reactor. *Environmental Science and Pollution Research*, **26**, 37026-37036. <https://doi.org/10.1007/s11356-019-06813-4>
- [19] Chen, J.C., Chang, N.B. and Shieh, W.K. (2003) Assessing Wastewater Reclamation Potential by Neural Network Model. *Engineering Applications of Artificial Intelligence*, **16**, 149-157. [https://doi.org/10.1016/s0952-1976\(03\)00056-3](https://doi.org/10.1016/s0952-1976(03)00056-3)
- [20] Chen, J., Song, L., Wainwright, M., *et al.* (2018) Learning to Explain: An Information-Theoretic Perspective on Model Interpretation. *Proceedings of the International Conference on Machine Learning*.
- [21] Cheng, H., Liu, Y., Huang, D. and Liu, B. (2019) Optimized Forecast Components-SVM-Based Fault Diagnosis with Applications for Wastewater Treatment. *IEEE Access*, **7**, 128534-128543. <https://doi.org/10.1109/access.2019.2939289>