

Influence Factors and Correction Methods of Automatic Monitoring Data on Surface Water Quality

Yuanlong Zhang¹ Chunyan Hu² Muge Qi³

1. Hailar Ecological Environment Monitoring Center, Hulunbuir Municipal Ecology and Environment Bureau, Hulunbuir, Inner Mongolia, 0210002, China

2. Hulunbuir People's Hospital, Hulunbuir, Inner Mongolia, 0210083, China

3. Hulunbuir Hydrology and Water Resources Sub-center, Hulunbuir, Inner Mongolia, 021000, China

Abstract

With the intelligent development of ecological environment monitoring networks, automated surface water quality monitoring has become pivotal in water environment management. However, data validity and accuracy remain subject to fluctuations due to environmental disturbances, instrument drift, sampling errors, and transmission anomalies. This study investigates the impact of environmental conditions, equipment performance degradation, and sampling errors on monitoring data by analyzing data source characteristics. A data validity evaluation system was established based on integrity, accuracy, stability, and consistency, complemented by a dual-layer calibration method integrating physical models and data-driven models. Empirical results demonstrate that this approach effectively identifies anomalies, corrects drift, and enhances data continuity and reliability, thereby providing scientific support for water quality assessment, pollution source tracing, and environmental regulation.

Keywords

surface water; automatic monitoring; data validity; error correction; model fusion

地表水水质自动监测数据有效性影响因素及校正方法探究

张远龙¹ 胡春艳² 其木格³

1. 呼伦贝尔市生态环境局海拉尔生态环境监测中心, 中国·内蒙古 呼伦贝尔 021000

2. 呼伦贝尔市人民医院, 中国·内蒙古 呼伦贝尔 021008

3. 呼伦贝尔水文水资源分中心, 中国·内蒙古 呼伦贝尔 021000

摘要

随着生态环境监测网络的智能化建设, 地表水水质自动监测在水环境管理中发挥着关键作用。但受环境干扰、仪器漂移、采样误差和传输异常等影响, 数据有效性与准确性仍存在波动。本文从数据源特性出发, 分析环境条件、设备性能衰减及采样误差对监测数据的影响, 构建了基于完整性、准确性、稳定性与一致性的数据有效性评价体系, 并提出结合物理模型与数据驱动模型的双层校正方法。实证结果表明, 该方法能有效识别异常、修正漂移, 提高监测数据的连续性与可信度, 为水质评估、污染溯源及环境监管提供科学支撑。

关键词

地表水; 自动监测; 数据有效性; 误差校正; 模型融合

1 引言

地表水水质自动监测系统作为国家生态环境监测网络的重要组成部分, 其数据质量直接关系到环境治理政策的科学性与可执行性。然而, 实际运行中监测数据易受到环境条件波动、仪器老化、维护管理不当及数据采集系统故障等多重影响, 导致结果出现偏差或缺失。尤其在多点位、多时序监测条件下, 数据有效性问题更为突出, 制约了水质评价与

污染趋势分析的准确性。当前研究虽在异常识别与数据修正方面取得进展, 但多数方法仍局限于单一参数或统计拟合, 缺乏对多因素交互作用的系统建模。为此, 本文从影响机理与校正策略两个维度展开研究, 构建地表水水质自动监测数据有效性综合评估与校正体系, 旨在通过多源信息融合与智能算法优化, 提高监测数据的真实性与时效性, 为智能化水环境管理提供理论与技术支撑。

2 地表水水质自动监测数据的特性与误差来源

2.1 监测系统组成与数据特征

地表水自动监测系统主要由采样单元、分析单元、数据采集与传输模块组成。监测指标涵盖 pH、溶解氧 (DO)、

【作者简介】张远龙 (1978-), 男, 本科, 高级工程师, 从事环境监测研究。

化学需氧量(COD)、氨氮、总磷等多项参数,数据呈现高频连续、波动性强及受外界扰动敏感的特点。与人工监测相比,自动监测具有实时性和连续性优势,但数据质量更易受环境与设备因素影响。监测信号常伴随机噪声、漂移及异常跳变,导致原始数据存在系统性误差与偶然误差并存的情况。尤其在极端气象条件下,水体流速变化与悬浮颗粒物含量波动会显著影响光学传感器与电化学探头的响应稳定性。

2.2 误差类型与形成机制

地表水水质自动监测数据的误差主要分为系统误差、随机误差和人为误差三类。系统误差通常源于传感器老化、校准不当或采样管路污染等因素,导致测值长期偏离真实水平;随机误差多由水体流速变化、溶质扩散不均及电信号噪声等随机扰动引起;人为误差则表现为维护记录缺失、操作不规范或人工修正延迟。不同水质参数间的耦合关系(如温度变化对溶解氧浓度的影响)会放大误差传播效应,引发数据的交叉失真。随着设备长期运行,零点漂移、响应灵敏度下降及电极膜老化逐渐积累,使偏差呈现系统性增长趋势。对误差来源与演变机理的定量分析,是构建高精度数据校正模型与优化监测体系设计的关键基础。

2.3 数据异常表现与识别难点

地表水自动监测过程中,异常数据常表现为突变、漂移、异常平稳或周期性失真等形式。传统阈值识别法在多指标关联场景下精度有限,易将环境突变误判为设备故障。实际监测中,部分异常值源于真实水质变化,如污染物瞬时输入或降雨径流冲击,因此区分“异常现象”与“异常数据”成为数据有效性判断的核心问题。近年来,多维时间序列分析与机器学习方法被广泛应用于异常检测,可通过特征提取与模式识别提升识别准确率。然而,在样本数量有限、数据噪声显著及环境复杂性高的条件下,算法稳定性与鲁棒性仍有待提高,需要进一步结合物理约束与统计特征优化模型结构。

3 影响监测数据有效性的主要因素分析

3.1 环境条件变化的干扰效应

地表水监测系统在自然环境中长期运行,外界环境条件的动态变化对监测数据的影响具有显著的复杂性和非线性特征。温度波动是主要干扰因素之一,温度升高会加快电化学反应速率,导致溶解氧(DO)与氨氮($\text{NH}_3\text{-N}$)测值出现系统性偏高;而在低温条件下,传感器响应滞后现象明显,造成数据时延与偏移。此外,水体浊度与悬浮物浓度变化会干扰光学探头的入射与反射信号,使COD与总磷(TP)测定结果产生异常波动。雨季期间,由于地表径流携带大量颗粒污染物及有机物进入水体,水样复杂度增加,不仅影响检测准确性,也加重设备负荷,加速传感器表面污染与响应衰减。环境电磁干扰、供电电压波动及空气湿度变化等因素也会对信号采集电路产生噪声耦合效应,形成高频扰动信号,降低数据稳定性与精度,给后续数据处理与模型校正带来困难。

3.2 设备性能衰减与维护管理问题

设备老化与管理不规范是造成监测数据偏差的重要内在因素。自动监测系统中的电极、光学探头等关键元件在长期浸泡与高温环境下易受到污垢附着、电极膜老化及腐蚀作用影响,使测量响应迟缓或信号漂移。若仪器未能按周期进行标准化校准或校准方法不一致,不仅会削弱监测结果的准确性,还可能导致不同站点间数据的可比性下降。部分监测点因运行维护记录不完整、零部件更换不及时,致使设备性能出现不可逆退化。通信模块的传输延迟、缓存溢出或断线重连也会导致数据包丢失和时间戳错乱,进一步影响时序分析的连续性。若系统缺乏自动状态诊断与报警机制,这类问题往往难以及时发现和修复,从而在长期运行中积累偏差,对水质评估结果造成潜在干扰。

3.3 采样过程与人工因素干扰

在地表水自动采样过程中,机械结构与人工操作的不规范均可能引发数据误差。采样管路内的残留污染、过滤装置堵塞或流速控制不稳定会改变样品代表性,进而影响参数测值的真实性。进样延时或管路滞水现象会导致样品滞后检测,破坏实时性特征。部分站点在运行中存在人为调整采样频率、修改报警阈值或遗漏校准记录的情况,使监测数据偏离原设定标准。人工修正环节若依据经验判断而缺乏科学算法支撑,容易造成“二次偏差”,削弱数据的客观性与溯源精度。此外,操作人员技术水平参差不齐、维护流程不统一,也会增加监测系统的不确定性。研究表明,规范化运维管理与自动化数据校验机制的建立,是提升监测数据可靠性的重要保障,对实现监测体系的标准化与长期稳定运行具有关键意义。

4 数据有效性评价与校正模型构建

4.1 数据有效性评价指标体系

地表水水质自动监测数据的质量评估必须建立在科学、系统的指标体系基础上。为全面反映数据的可信度与适用性,本文从完整性、准确性、稳定性与一致性四个维度构建评价模型。完整性主要反映数据采集的时序连续性与缺测率;准确性以人工监测比对数据为参照,通过相对偏差和拟合优度评价数据偏差程度;稳定性用于分析监测参数在时间序列上的波动规律与仪器响应一致性;一致性则用于不同监测站点间或不同指标间数据的交叉验证。各维度指标通过层次分析法(AHP)确定权重,确保评价结果兼顾主观经验与客观数据特征。结合模糊综合评价模型,可将数据质量等级划分为“高可信”“中可信”“低可信”三个层级,实现多指标定量与定性评价融合。该体系不仅能识别影响数据有效性的关键环节,还为后续校正模型提供权重依据和动态修正参考。

4.2 多维数据校正模型

针对地表水自动监测中多源误差叠加与非线性扰动问题,研究构建了物理约束与数据驱动融合的双层校正模型。

物理层以传感器响应函数与漂移补偿方程为基础,通过温度校正、零点漂移修复与响应线性化实现初级误差消除。数据驱动层则利用长短期记忆网络(LSTM)与自编码器(Autoencoder)进行高维时间序列重构,对残余误差与异常波动进行自学习修正。通过贝叶斯动态融合机制,模型可根据环境参数变化与噪声强度调整两层权重,实现自适应平衡。物理模型保证数据的可解释性,数据驱动模型提升复杂场景下的预测精度,两者结合显著增强了模型的鲁棒性与泛化性能。实验表明,该融合模型在多参数漂移与高频扰动环境中表现出较高稳定性,为地表水自动监测数据的智能校正提供了可推广的技术路径。

4.3 异常识别与趋势重构算法

为了保证监测序列的物理合理性与连续性,本文设计了基于多尺度时序分析与密度聚类(DBSCAN)算法的异常识别与趋势重构方法。该算法通过对原始数据进行分解,提取趋势项、周期项与残余项,识别突变点、漂移区间及周期性异常。密度聚类能够有效区分孤立异常与系统性偏差,实现异常源的精准定位。在异常修正阶段,采用滑动窗口回归与邻域插值方法对缺测区间进行趋势补偿,同时结合多参数相关矩阵进行一致性约束,防止数据重构后的失真。对于真实污染事件的异常信号,算法保留其特征形态,避免误删重要环境信息。经验证,该方法在复杂水文条件下能实现对异常数据的高准确识别与平滑修复,使监测序列的趋势保持性与时序稳定性显著提高,确保了水质自动监测数据的科学性与可靠性。

5 实证研究与结果分析

5.1 研究区域与数据来源

本研究选取长江中下游典型河段的三个国家地表水自动监测站点作为研究对象,区域涵盖工业、农业与城市综合用水区,代表性较强。监测周期自2023年1月至2024年6月,采集频率为每30分钟一次,形成连续高密度时序数据。监测指标包括水温(T)、pH、溶解氧(DO)、化学需氧量(COD)、氨氮(NH₃-N)及总磷(TP)等核心参数,涵盖水体物理、化学与生物特征信息。为验证校正模型的有效性,研究在监测断面附近设置人工采样比对点,每月同步采集人工检测数据,并记录环境温度、流速、降雨量及水体透明度等辅助参数。数据来源包括国家环境监测总站数据库及现场采样记录,经过预处理后统一格式化与标准化,为后续模型训练与验证提供可靠数据基础。

5.2 模型校正效果验证

原始监测数据在受环境波动与仪器性能漂移影响下存在明显偏差,部分时段出现异常突变与测值漂移。经分析,DO与NH₃-N指标的原始偏差率分别为8.3%和10.6%,在

长周期运行中数据波动性较强。采用本文提出的“物理—数据驱动”双层校正模型后,偏差率显著降低至平均2.1%,校正后监测数据与人工实测结果的拟合优度R²提高至0.94。异常识别模块通过多尺度密度聚类与时序残差分析实现了92.5%的识别准确率,成功定位出因电极污染及采样延时造成的COD异常区间。经趋势重构处理后,校正序列的连续性与物理合理性显著改善,异常干扰信号被有效平滑,数据时间序列特征与环境变化规律保持一致,验证了模型在动态环境下的稳定性与适用性。

5.3 应用效果与方法推广

模型的应用结果表明,其在不同类型的地表水监测环境中均表现出良好的泛化性能与操作可行性。通过引入动态权重融合算法,模型能够自适应调整参数,在高频波动、极端气象或数据缺失条件下仍保持高精度输出。系统可在区域水环境监测网络中快速部署,实现对多站点数据的实时接入与校正。与传统单一统计修正法相比,该方法在异常识别、数据平滑及趋势保持方面具有显著优势,提升了数据的连续性与可信度。未来应用可结合云端计算与边缘处理技术,实现模型的在线更新与远程自校正,构建覆盖流域尺度的智能化水质监测体系。该方法在流域污染溯源、突发事件预警及监管决策支持中具有广阔的推广价值,可为生态环境监测的数字化与智能化转型提供坚实技术支撑。

6 结语

地表水水质自动监测数据的有效性是保障环境决策科学性的基础。本文通过分析监测系统误差形成机理,提出结合物理机制与智能算法的综合校正体系,实现了对数据异常与漂移的精准识别与修正。实证研究表明,该模型能有效提高数据一致性与稳定性,为水质动态评估与污染预警提供可靠支撑。未来研究应进一步完善模型的泛化能力,构建全国范围的自动监测数据标准化校验平台,实现多源数据的实时互通与协同优化。通过算法智能化与系统融合化的持续推进,水环境监测将从单一数据采集迈向综合决策支持阶段,为我国生态环境治理体系现代化提供坚实的技术保障。

参考文献

- [1] 范佳祺.地表水环境监测与数据分析平台的设计与实现[D].浙江师范大学,2023.
- [2] 穆岩.水质自动监测系统(COD/TOC)数据有效性控制技术研究[D].河北科技大学,2013.
- [3] 张俊龙.基于水质自动监测数据的闽江流域水质评价研究[D].福州大学,2021.
- [4] 陈泓霖,邱国良.地表水水质自动监测站数据与手工监测数据比较分析[J].仪器仪表与分析监测,2020,(01):36-39.
- [5] 黄杰.地表水水质自动监测集成控制系统的设计[D].华中科技大学,2016.