

基于大数据的区域水质检测结果趋势研判应用方法研究

颜燕玉

无锡市江阴生态环境监测站

DOI:10.32629/eep.v8i10.2910

[摘要] 本文聚焦于基于大数据的区域水质检测结果趋势研判应用方法。首先阐述了大数据在水质检测领域应用的重要性,接着详细介绍了区域水质检测大数据的获取途径与数据特点。然后深入探讨了趋势研判的具体应用方法,包括时间序列分析、空间关联分析以及综合模型构建等方面。通过实际案例分析展示了该方法在预测水质变化、发现潜在污染区域等方面的有效性,最后对未来的发展方向进行了展望,旨在为区域水质管理和保护提供科学有效的决策支持。

[关键词] 大数据; 区域水质检测; 趋势研判; 应用方法

中图分类号: X832 文献标识码: A

Research on the Application Method of Trend Judgment for Regional Water Quality Detection Results Based on Big Data

Yanyu Yan

Jiangyin Ecological Environment Monitoring Station

[Abstract] This paper focuses on the application method of trend judgment for regional water quality detection results based on big data. Firstly, it elaborates on the importance of big data application in the field of water quality detection. Then, it details the acquisition channels and data characteristics of regional water quality detection big data. Subsequently, it delves into the specific application methods of trend judgment, including time series analysis, spatial correlation analysis, and comprehensive model construction. Through the analysis of actual cases, it demonstrates the effectiveness of this method in predicting water quality changes and identifying potential pollution areas. Finally, it looks forward to the future development direction, aiming to provide scientific and effective decision support for regional water quality management and protection.

[Key words] Big Data; Regional Water Quality Detection; Trend Judgment; Application Method

引言

水是生命之源,区域水质状况直接关系到当地居民的生活质量、生态环境的稳定以及经济的可持续发展。传统的水质检测方法往往依赖于有限的采样点和定期的检测,难以全面、及时地掌握水质动态变化。随着信息技术的发展,大数据为区域水质检测带来了新的机遇。通过对大量水质检测数据的收集、整合与分析,能够更准确地研判水质变化趋势,提前发现潜在的水质问题,为水质管理和保护提供科学依据。因此,研究基于大数据的区域水质检测结果趋势研判应用方法具有重要的现实意义。

1 区域水质检测大数据的获取与特点

1.1 数据获取途径

区域水质检测大数据的获取依赖于多源、多模态的监测体系。首先,传统检测站点由环保部门在河流、湖泊、水库等关键水体布设,按照固定时间间隔开展人工采样与实验室分析,可稳

定获取水温、酸碱度(pH)、溶解氧(DO)、化学需氧量(COD)等常规水质指标数据,具有较高的准确性和权威性。其次,在线监测设备被广泛部署于水体敏感或重点管控区域,能够对水质参数进行高频次甚至连续实时监测,并通过无线网络将数据自动上传至数据中心,显著提升了数据的时效性与响应能力。此外,移动监测平台(如搭载传感器的船只、无人机等)凭借其灵活性和机动性,可在大范围水域内动态巡航,有效覆盖固定站点和在线设备难以触及的盲区,实现空间维度上的补充监测。最后,为构建更全面的水质评估模型,还需整合气象数据(如降雨量、气温)、土地利用信息、工业废水排放记录等外部辅助数据源。这些数据虽非直接水质指标,却与水体污染成因及演变趋势密切相关,为深入解析水质变化机制提供重要支撑。

1.2 数据特点

区域水质检测大数据呈现出典型的“四高”特征:①海量性;②多样性;③时效性;④复杂性。首先,由于监测网络覆盖

广、指标种类多、时间跨度长,系统持续积累的数据体量极为庞大,涵盖从分钟级到年际尺度的多维时序信息。其次,数据形态高度多样,既包括结构化的数值型数据(如各类污染物浓度、理化参数),也包含非结构化的文本描述(如现场检测日志、异常事件报告)以及带有地理坐标的矢量空间数据,对数据融合与标准化处理提出更高要求。再者,水质状况易受突发污染事件或极端天气影响,部分应用场景(如应急预警、污染溯源)亟需近实时的数据采集、传输与分析能力,对系统的响应速度和计算效率构成挑战。最后,水质变化本质上是自然过程与人类活动交织作用的结果——既受降水、蒸发、径流等水文气象因素驱动,又受工业排放、农业面源污染、污水处理效能等人为干预影响,导致各变量间存在非线性、滞后性及交互耦合关系,使得数据内在规律挖掘与趋势预测面临高度复杂性。^[1]

2 趋势研判应用方法

2.1 时间序列分析

时间序列分析是研究区域水质动态演变规律的核心方法之一,尤其适用于对长期、连续监测数据的趋势识别与未来预测。该方法通过系统梳理历史水质指标(如化学需氧量COD、氨氮、总磷等)的时间序列,挖掘其内在的周期性、趋势性和随机波动特征。

以某河流COD月均值为例,首先需对原始序列进行预处理,包括缺失值插补、异常值剔除及平稳性检验(如ADF检验)。若序列呈现非平稳性,则通过一阶或二阶差分使其转化为平稳序列,为后续建模奠定基础。随后,依据自相关函数(ACF)和偏自相关函数(PACF)图确定自回归(AR)与移动平均(MA)项的阶数,构建合适的ARIMA(p, d, q)模型。模型拟合后需进行残差白噪声检验和参数显著性评估,确保其统计合理性。最终,利用训练好的模型对未来若干时段的水质指标进行滚动预测,不仅可预判水质恶化风险,还能为环保部门制定前置性管控策略(如加强排污监管、调度生态补水)提供量化支撑。此外,结合季节性分解(如STL)或引入外部协变量(如降雨量),可进一步提升模型在复杂水文条件下的适应性与预测精度。

2.2 空间关联分析

区域水质不仅随时间变化,在不同空间位置也存在差异。空间关联分析旨在揭示水质数据在空间上的分布特征和相关性。借助地理信息系统(GIS)平台,可将各监测站点的水质实测值与其精确地理坐标进行空间匹配,生成直观的水质热力图或等值线图,清晰展现污染“热点”与“冷点”区域。^[2]

在此基础上,采用空间统计方法(如全局Moran's I指数)定量评估水质数据的空间自相关性:若Moran's I显著大于0,表明高值或低值在空间上呈集聚分布,暗示可能存在区域性污染源(如工业园区废水排放、农业面源径流);若接近0,则说明水质状况呈随机分布。为进一步解析成因,可叠加土地利用类型、人口密度、河网结构等空间图层,开展地理加权回归(GWR)或多尺度空间聚类分析,识别不同下垫面条件下水质响应的敏感区域。例如,工业用地密集区往往伴随COD、重金属浓度升高,

而农业区则可能与氮磷营养盐超标密切相关。此类空间关联洞察有助于精准划定优先管控单元,优化监测网络布局,并为流域尺度的分区治理提供科学依据。^[3]

2.3 综合模型构建

面对水质变化受多重时空因素耦合驱动的复杂现实,单一分析方法难以全面刻画其演化机制,因此亟需构建融合多维信息的综合研判模型。当前,基于机器学习的集成建模方法展现出显著优势。

以随机森林(Random Forest)为例,该算法能够高效整合来自不同维度的数据:时间维度上纳入历史水质序列及其滞后项,空间维度上引入监测点坐标、缓冲区内的土地利用比例、邻近污染源距离等地理特征,同时耦合气象因子(如降水量、蒸发量)、社会经济数据(如工业产值、污水处理率)等外部协变量,共同作为输入特征集,目标输出则为未来特定时段的水质指标预测值。随机森林通过构建大量决策树并集成其结果,不仅能自动处理高维、非线性、交互性强的数据关系,还具备良好的抗过拟合能力和特征重要性评估功能,有助于识别关键驱动因子。此外,也可结合深度学习模型(如LSTM用于时序建模、图神经网络用于空间关系建模)或贝叶斯融合框架,进一步提升模型的表达能力与鲁棒性。此类综合模型不仅能实现“时空一体”的水质趋势预测,还可支持情景模拟(如减排政策效果评估、极端气候影响推演),为智慧水务和精准治污提供强有力的技术支撑。^[4]

3 典型城市区域水质趋势智能研判与治理应用实例

3.1 数据收集与整理

本案例以某南方中等规模城市为研究对象,该区域水系发达,涵盖3条主要河流及2个大型湖泊,周边土地利用类型复杂,包括约15 km²的工业园区、80 km²的农田以及密集的居民聚居区。研究团队系统收集了2019年1月至2023年12月共60个月的水质监测数据,覆盖区域内28个固定监测站点,每月定期检测pH值(均值7.2±0.4)、溶解氧(DO,平均6.8 mg/L)、氨氮(NH₃-N,平均0.85 mg/L)、总磷(TP,平均0.22 mg/L)等关键指标。同时,整合了气象局提供的逐月气温(年均温17.5℃)与降水量(年均降水1200 mm)、自然资源部门的土地利用矢量图,以及生态环境局监管的工业废水排放台账(年均排放量约1200万吨,含氮污染物占比18%)。所有数据经统一时空基准对齐、异常值清洗和标准化处理后,构建了结构化多源数据库,为后续趋势研判奠定坚实基础。^[5]

3.2 趋势研判过程

在趋势研判阶段,首先开展时间序列分析:对氨氮浓度建模发现,其夏季(6-8月)均值达1.32 mg/L,显著高于年均值(p<0.01),ARIMA(1, 1, 2)模型拟合优度R²=0.87,揭示出明显的季节性上升趋势,推测与高温促进有机物分解、农业施肥高峰期及降雨冲刷面源污染有关。其次,空间关联分析借助ArcGIS平台生成水质热力图,全局Moran's I指数为0.43(p=0.002),表明存在显著空间集聚;局部LISA聚类进一步显示,位于工业园区下游5 km范围内的3个监测点持续呈现“高一高”集聚(氨氮>1.5 mg/L,总磷>0.3

mg/L), 印证工业排放的累积影响。最后, 构建随机森林综合模型, 输入包含12维特征(如距工业区距离、前3月平均降水、季节虚拟变量等), 经5折交叉验证, 模型对氨氮未来1个月预测的RMSE为0.11mg/L, R^2 达0.91, 显著优于单一时间或空间模型。^[6]

3.3 应用效果

基于上述研判结果, 当地环保部门于2023年5月提前启动预警响应机制: 一方面对工业园区内12家重点排污企业实施限产与在线监控强化, 另一方面联合农业农村局推广生态沟渠与缓冲带建设, 减少农田氮磷流失。同时, 在预测高风险河段新增2个临时监测点, 将监测频率由月度提升至每周一次。实施干预后, 2023年7-8月该河段实测氨氮浓度稳定在0.95-1.05 mg/L, 较2022年同期(1.42mg/L)下降约30%, 未出现超标(标准限值1.5mg/L)事件。^[7]

4 未来发展方向

面向未来, 区域水质趋势研判的发展将呈现多维度融合与智能化升级的特征。首先, 在数据层面, 亟需打破环保、水利、气象、住建、农业等部门间的数据壁垒, 推动跨行业、跨层级的多源异构数据深度融合, 构建统一、开放、标准化的水质大数据共享平台, 整合地面监测、卫星遥感、企业排污、土地利用及社会经济等多元信息, 全面提升数据质量与时效性。其次, 在模型方法上, 应持续优化现有研判体系, 积极引入深度学习(如LSTM、图神经网络)、强化学习等先进算法, 增强模型对非线性、高维、时空耦合过程的刻画能力, 并针对不同区域水文特征和水体类型(如河流、湖泊、城市内河)开发可迁移、可解释的个性化预测模型。此外, 在应用响应方面, 需加快部署基于物联网与5G的实时水质传感网络, 结合边缘计算与智能分析技术, 构建“感知—传输—预警—响应”一体化的智能预警系统, 实现对水质异常的秒级识别与分级告警, 为突发污染事件提供快速决策支持。最后, 水质趋势研判应超越单一环境管理范畴, 深度融入水资源优化配置、国土空间规划、生态保护修复及“双碳”战略等更广泛的政策领域——例如, 在流域调水中同步考虑水质约束条件, 在城市更新中依据水质风险布局绿色基础设施, 从而为区域高质量发展和人水和谐共生提供系统性、前瞻性、协同化的综合决策支撑。^[8]

5 结束语

综上所述, 基于大数据的区域水质变化趋势智能研判方法, 为新时代水环境精细化管理和科学决策提供了有力支撑。通过系统整合多源异构的水质监测数据, 融合时间序列分析、空间关联建模与机器学习等先进技术, 不仅能够精准识别水质演变规律, 还能有效预判潜在污染风险, 显著提升水环境治理的前瞻性与主动性。本文所开展的实际案例分析充分验证了该方法在复杂城乡交错区域中的适用性与实践价值。展望未来, 随着数据基础设施的完善、人工智能算法的演进以及跨领域协同机制的深化, 水质趋势研判将在实时感知、动态预警、区域适配和多目标协同等方面持续优化, 进一步融入水资源管理、生态保护与国土空间规划等宏观治理体系。该方法的发展与应用, 将为筑牢水生态安全屏障、推动绿色低碳转型和实现区域高质量可持续发展提供坚实的技术支撑与决策依据。

[参考文献]

- [1]王理想.基于多指标时序数据的城市河道水质异常检测方法研究[D].浙江大学,2021.
- [2]梁庆.基于数据驱动的水质预测方法研究及监测系统开发[D].大连理工大学,2022.
- [3]韩晓霞,陈媛,胡冠宇等.一种融合专家知识和监测数据的水质预测模型[J].浙江工业大学学报,2021,49(05):569-576.
- [4]宋岳.大数据技术在水质监测中的应用[J].信息记录材料,2021,22(01):125-126.
- [5]李雪清,郑航.基于多源数据机器学习的区域水质预测方法研究[J].水利水电技术(中英文),2021,52(11):152-163.
- [6]张泽鹏,张云,张志.基于大数据的水质集散化监测系统的研究[J].科技资讯,2019,17(11):11-12.
- [7]吉绪新.数据驱动模型与机理模型耦合的水质预警预测研究[D].华南理工大学,2021.
- [8]胡前明,金松.基于环境保护大数据的监测与智能诊断研究[J].化工设计通讯,2022,48(09):167-169.

作者简介:

颜燕玉(1986--),女,汉族,环境工程专业,本科,工程师,主要从事环境监测工作。