

大气环境监测中大数据技术的应用实践探究

刘堂银

达州市渠县生态环境监测站 四川 达州 635200

摘要：大数据技术在大气环境监测中应用广泛且关键。大气环境监测数据复杂、实时性及精准性要求高，大数据技术可有效应对。其在数据采集与整合上实现“空天地”一体化；通过多种方法评估与分析大气环境质量；利用受体模型等解析污染源并溯源；基于机器学习算法预测与预警大气环境；还为环境管理与决策提供支持。实践表明，大数据技术提升了大气环境监测的效率与精准度，为大气环境治理提供了有力支撑，推动环境管理迈向数据驱动新时代。

关键词：大气环境监测；大数据技术；应用实践

引言

大气环境质量与人类健康及生态平衡紧密相连，大气环境监测意义重大。然而，传统监测手段在数据采集、处理和分析上存在短板，难以应对日益复杂的监测任务。大数据技术的出现为大气环境监测带来了新机遇。它凭借强大的数据处理、分析及可视化能力，能整合多源异构数据，挖掘数据深层价值，达成实时监测、精准评估与科学预测。深入探讨其应用实践，可提升监测水平，为大气环境治理提供科学支撑，推动人与自然和谐共生。

1 大数据技术概述

大数据技术源于数据量的急剧增长。随着互联网、物联网、社交媒体等的飞速发展，数据产生的速度和规模呈指数级上升。从电商平台的交易记录、社交网络的用户动态，到工业生产中的传感器数据，每时每刻都在生成海量信息^[1]。这些数据来源广泛、类型多样，涵盖了结构化、半结构化和非结构化数据，为大数据技术提供了丰富的素材。大数据技术的核心在于其强大的数据处理能力。传统的数据处理方法在面对海量、复杂的数据时往往力不从心，而大数据技术通过分布式计算、并行处理等技术手段，能够高效地对数据进行存储、管理和分析。利用分布式文件系统可以将数据分散存储在多个节点上，提高数据的可靠性和读写性能；通过并行计算框架，如MapReduce，能够同时对大量数据进行处理，大大缩短了数据处理时间。在数据分析方面，大数据技术融合了多种先进的方法和技术。机器学习算法可以从数据中自动学习模式和规律，实现对数据的分类、预测和聚类；深度学习则通过构建多层神经网络，对数据进行更深入的挖掘和分析，在图像识别、语音识别等领域取得了显著成果。此外，数据挖掘技术能够发现数据中隐

藏的关联关系和潜在价值，为企业决策提供有力支持。大数据技术的最终目标是将数据转化为有价值的信息和知识，并实现可视化展示。通过数据可视化工具，将复杂的数据以直观的图表、图形等形式呈现出来，使用户能够快速理解数据背后的含义。这不仅有助于企业进行市场分析、风险评估和战略规划，也为政府部门的决策提供了科学依据。总之，大数据技术正以其独特的优势，引领着各行业迈向智能化、数字化的新时代。

2 大气环境监测中大数据技术的应用需求

2.1 大气环境监测数据的复杂性

大气环境监测数据的多元性与复杂性，远超传统数据处理范畴。数据采集渠道丰富，从地面固定监测站点的高频采样，到卫星遥感的广域扫描，再到无人机、移动监测车的灵活探测，形成“点-面-动态”的立体化数据网络。数据类型更是涵盖气象参数、污染物浓度、地理信息等多维度信息，例如气象数据包含温度、湿度、风速风向等变量，污染物浓度数据涉及PM_{2.5}、PM₁₀、SO₂等十余种指标，地理信息数据则包含地形地貌、城市功能区分布等内容。这些数据不仅种类繁多，还具有时空分辨率高、体量大的特点。地面监测站点每小时甚至每分钟生成数据，卫星遥感影像分辨率可达30米，海量数据以TB级规模持续增长。数据格式上，有结构化的数据库表、半结构化的JSON/XML文件，以及非结构化的遥感影像、文本报告，传统处理方法在数据清洗、格式统一和关联分析上效率低下，难以挖掘数据深层价值，亟需大数据技术实现高效处理与整合。

2.2 大气环境监测的实时性要求

大气环境瞬息万变，污染物扩散、气象条件突变等过程以分钟甚至秒级速度演变，传统监测手段因数据传输延迟、处理缓慢，无法满足实时预警需求。例如突

发工业泄漏、交通拥堵导致的局部污染,若不能及时发现并处置,将快速扩散造成更大危害。大数据技术凭借高速数据采集与实时处理能力,成为解决这一问题的关键。物联网传感器实时采集数据,5G、NB-IoT等通信技术实现低延迟传输,分布式计算框架如SparkStreaming可秒级处理海量数据流。一旦监测数据触发预设阈值,系统能立即发出预警,为应急响应争取时间。如深圳市大气监测系统通过大数据实时分析,将污染事件响应时间从小时级缩短至分钟级,显著降低污染影响范围与危害程度。

2.3 大气环境监测的精准性需求

大气环境监测的核心目标是为环境管理提供科学依据,传统单一数据源、简单统计分析难以精准刻画环境状况。例如仅依靠固定自动监测站的监测数据,无法判断污染物来源及传输路径,导致治理措施缺乏针对性。大数据技术通过多源数据融合与深度分析,大幅提升监测精准度。将气象数据、污染物浓度数据、地理信息数据交叉分析,利用机器学习算法挖掘数据间非线性关系,可精准识别污染源并量化其贡献比例。基于历史数据构建的预测模型,如LSTM神经网络,能考虑多种影响因素预测未来污染趋势,误差率较传统模型降低20%-30%。

3 大数据技术在大气环境监测中的应用实践

3.1 数据采集与整合

在大气环境监测体系构建中,数据采集与整合是开展后续分析与决策的基石。传统监测方式依赖固定站点,数据采集存在空间覆盖不足、时间滞后等问题,而大数据技术推动下的“空天地”一体化采集体系,实现了监测维度的全面拓展。地面监测站点作为核心数据采集单元,部署了高精度的污染物浓度监测设备,能够实时获取PM_{2.5}、PM₁₀、臭氧、一氧化碳、二氧化硫、氮氧化物等关键指标数据。我国目前已建成的国控空气质量监测站点达1436个,覆盖全国主要城市和区域,为大气环境质量监测提供了基础数据支撑。与此卫星遥感凭借其广域覆盖的优势,从高空捕捉气溶胶光学厚度、臭氧柱浓度等宏观数据,如我国高分五号卫星,其搭载的光谱成像仪可实现对大气污染物的高光谱监测,空间分辨率达30米,为大范围污染态势分析提供了可能。无人机监测则深入复杂地形或工业园区,利用搭载的微型传感器实现局部区域的精细化监测,尤其适用于固定站点难以覆盖的偏远地区或污染热点区域^[2]。物联网技术成为连接这些分散监测设备的关键纽带,通过5G、NB-IoT等通信技术,将采集的数据实时传输至数据中心。在数据传输过程中,

采用SSL/TLS加密协议、数字签名等技术,保障数据在传输过程中的安全性和完整性,防止数据被窃取、篡改。当数据抵达数据中心后,需经过清洗、转换和集成等预处理操作。借助Python编写的数据清洗脚本,或ETL工具,可自动识别并剔除因设备故障、信号干扰产生的异常值、重复记录;利用数据转换工具,将JSON、XML、CSV等不同格式的数据统一为标准格式;最后,通过数据集成技术,将多源异构数据融合到数据仓库中。

3.2 大气环境质量评估与分析

大气环境质量评估是对大气环境状况的量化判断,需综合考虑气象条件、污染物浓度等多种因素。大数据分析技术的引入,使得评估过程更加科学、精准和高效。模糊综合评价法是常用的评估方法之一,该方法通过构建隶属函数,将各指标的监测值映射到不同的质量等级隶属度,再结合权重计算综合评价结果,能够直观呈现区域大气环境质量等级。在对某城市空气质量进行评估时,可将PM_{2.5}、PM₁₀、二氧化硫等污染物浓度作为评价指标,通过模糊综合评价法得出该城市空气质量属于优良、轻度污染等具体等级。神经网络模型则凭借强大的自学习能力,通过训练大量历史数据,自动学习数据间的复杂非线性关系,实现对大气环境质量的定量评估。数据挖掘技术在分析历史数据方面发挥着重要作用。通过关联规则挖掘,能够发现不同因素之间的潜在关联,如研究发现湿度与臭氧浓度之间存在一定的负向相关关系,低湿度条件可能促进臭氧的生成;聚类分析则能将具有相似污染特征的区域归为一类,为管理者制定差异化的管控策略提供依据。生态环境部利用大数据分析模型,对全国城市空气质量进行实时评估并发布日报,通过可视化图表展示各城市空气质量指数、排名等信息,不仅为公众提供了清晰的环境质量信息,也为管理者决策提供了重要参考。

3.3 大气污染源解析与溯源

大气污染源解析旨在确定大气污染物的来源及其贡献率,是制定精准污染治理措施的关键环节。大数据技术通过融合大气环境监测数据、气象数据、地理信息数据等多源数据,运用受体模型、源模型等方法,实现对大气污染源的精准解析和溯源。化学质量平衡模型(CMB)是常用的源解析方法之一,该模型基于受体样品中化学组分的测量值,结合源成分谱数据,通过数学计算确定各污染源对污染物的贡献率。在北京的PM_{2.5}源解析中,利用CMB模型明确了机动车尾气贡献率约为30%、工业排放约为25%、扬尘约为15%等,为后续针对性治理提供了数据支撑。多源数据融合分析进一步提升

了污染源溯源的准确性。将气象数据中的风向、风速与污染物浓度数据结合,可追踪污染物的传输路径;地理信息系统(GIS)则以地图形式直观展示污染源分布,标记出高污染区域和关键排放源^[3]。上海市通过GIS可视化平台,将化工园区、交通主干道等污染源清晰呈现,环保执法人员可通过该平台快速定位污染源头,开展精准执法行动,极大提高了执法效率和效果。

3.4 大气环境预测与预警

基于大数据分析和机器学习算法建立的大气环境预测模型,能够有效预判未来大气环境质量变化趋势,为污染防控争取宝贵时间。时间序列分析利用历史数据的趋势性和周期性,对未来污染物浓度进行初步预测,例如通过分析过去一年的 $PM_{2.5}$ 浓度数据,预测未来一周的浓度变化趋势。支持向量机(SVM)通过寻找最优分类超平面,实现对污染等级的分类预测,判断未来大气环境质量属于优良、轻度污染等不同等级。深度学习中的长短期记忆网络(LSTM),擅长处理时间序列数据,能够精准捕捉污染变化规律,在复杂的污染预测场景中表现出色。为提高预测模型的准确性,需不断优化模型参数。通过交叉验证调整模型超参数,引入新的影响因素数据扩充训练集,如将交通流量、工业生产负荷等数据纳入模型训练。依据大气环境质量标准设定预警阈值,如 $PM_{2.5}$ 日均浓度超过 $75\mu g/m^3$ 时触发橙色预警,臭氧8小时平均浓度超过 $160\mu g/m^3$ 时发布臭氧污染预警。

3.5 环境管理与决策支持

大数据技术为环境管理部门提供了全面的数据支持和科学的决策依据,推动环境管理从经验驱动向数据驱动转变。大气环境管理信息平台整合了大气环境监测数据、污染源数据、环境执法数据等多源数据,形成动态监管中枢。环境管理人员通过平台,可实时查看各区域

空气质量指数、污染源排放数据,利用数据可视化图表快速定位异常区域。当发现某工业园区污染物排放超标时,管理人员可立即调取周边监测数据和执法记录,查看该园区近期排放变化趋势、历史违规记录等信息,迅速安排现场检查,实现对环境问题的快速响应和处理^[4]。大数据分析结果为环境政策制定、环境规划编制、污染治理项目实施等提供了科学依据。在制定环保政策时,通过分析不同行业减排对空气质量的影响,确定重点管控行业和减排目标;编制环境规划时,参考历史污染数据和预测结果,合理布局生态保护区和工业区域,避免污染集中。

结语

大数据技术在大气环境监测中的应用实践成效显著。从数据采集整合到质量评估分析,从污染源解析溯源到环境预测预警,再到环境管理与决策支持,大数据技术贯穿始终,提升了监测的全面性、实时性和精准性。它为环境管理部门提供了科学依据,助力精准治理。未来,随着大数据技术的不断发展,其在大气环境监测中的应用将更加深入和广泛,有望实现更高效的环境监测与治理,推动大气环境质量持续改善,为建设贡献力量。

参考文献

- [1]杨禄川.大气环境监测中大数据技术的应用实践探究[J].皮革制作与环保科技,2022,3(8):68-70.
- [2]周光.环境监测在大气污染治理中的实践探究[C]//2024智慧施工与规划设计学术交流会议论文集.2024:1-3.
- [3]武云燕.环境监测治理技术在大气污染中的实践当议[J].资源节约与环保,2023(1):51-54.
- [4]钱锴.大气污染中环境监测治理应用探究[J].区域治理,2021(13):145-146.