

环境监测数据异常值的判定及处置

尹卫萍¹, 侯鹏², 徐亮¹, 宋兴伟¹, 胡玲¹

(1. 江苏省环境监测中心, 江苏 南京 210019; 2. 江苏省苏力环境科技有限责任公司, 江苏 南京 210036)

摘要:准确、可靠的环境监测数据是环境管理的重要现实依据,因此正确鉴别和处理异常数据是环境监测工作中的关键环节。简述了环境监测异常数据的类型和产生的原因,对异常监测数据的处理措施进行了分析。提出,应建立健全环境监测质量管理体系、重视现场采样的信息化管理、正确填写监测原始记录和加强监测数据的审核。

关键词:环境监测;异常数据;判定;处置

中图分类号:X830

文献标志码:C

文章编号:1674-6732(2020)06-0063-04

Detection and Handling of Outliers in Environmental Monitoring Data

YIN Wei-ping¹, HOU Peng², XU Liang¹, SONG Xing-wei¹, HU Ling¹

(1. Jiangsu Provincial Environmental Monitoring Center, Nanjing, Jiangsu 210019, China; 2. Jiangsu SuLi Environmental Science and Technology Co. Ltd., Nanjing, Jiangsu 210036, China)

Abstract: Accurate and reliable environmental monitoring data is an important practical basis for environmental management, therefore, the correct identification and processing of abnormal data is a key link in environmental monitoring. The types and causes of abnormal environmental monitoring data are briefly described, and the processing methods of abnormal monitoring data are analyzed. It is proposed that a sound quality management system for environmental monitoring should be established, attention should be paid to the information management of on-site sampling, the original monitoring records should be filled correctly, and the monitoring data audits should be strengthened.

Key words: Environmental monitoring; Abnormal data; Detection; Handling

环境监测作为“环境保护的眼睛”,在我国环境保护中起着不可或缺的重要作用。环境监测的重要性在于提供科学准确的环境监测数据,便于环境评价、管理以及规划工作的顺利开展。提供准确的监测数据是环境监测的立身之本,但由于环境监测数据的准确性受现场采样、样品保存和运输、样品制备、实验室分析等多个环节、多种因素的影响,在实际监测中经常会出现监测结果异常,严重时甚至干扰后续管理决策。为了确保环境监测数据的准确性,需要对异常监测数据及时进行判定和处理。

1 异常监测数据的判定

1.1 异常监测数据的类型

环境监测异常数据主要指数据反映的内容与实际不符,具有较大差异,超出一般规定范围的数据。

(1) 离群数据。是指与多次监测的平均值或历史大量日常监测数据具有很大差距,无法反映实际情况的异常数据。根据统计学中规定,一组测定值中与平均值偏差超过2倍标准差的测定值就是异常数据。例如,污染源废气或废水监测时,连续监测4次,某一次测定的值明显低于或高于其他几次测定的结果均值。再如,地表水或环境空气质量

收稿日期:2019-12-16;修订日期:2020-02-16

基金项目:国家水环境监测监控及流域水环境大数据平台构建关键技术研究基金资助项目(2017zx07302);土壤监测装备标准化研究基金资助项目(1618)

作者简介:尹卫萍(1967—),女,正高级工程师,本科,主要从事环境监测工作。

监测时,存在同一断面或点位的监测值与历史同期数据差异很大而原因不明的情况。

(2) 不合理的数据。主要是指监测结果与正常逻辑不相符的数据。例如,废气监测时,相关联的监测因子的监测结果不合理,如表征恶臭的综合指标臭气浓度较低,特征因子氨气、硫化氢等恶臭类物质监测结果反而较高,或者相反;废气中,表征一类有机物的非甲烷总烃较高,而挥发性有机物反而较低甚至未检出。废气处理效率监测时,环保处理设施污染物进出口浓度出现倒置。污染源监测时,某次手工监测与正常的日常监测数据或其他来源数据如在线监测数据相差较大。地表水环境质量监测时,在同一监测时段内,某一河道水质中上游断面监测因子浓度明显高于下游断面,存在明显的不合理之处。

(3) 不符合实际的数据。主要是指监测结果与实际经验数据或经严格理论计算推算的结果数据相差较大的数据。例如,企业实际运行中,根据物料平衡,进口污染物排放浓度较高,但实际监测时,进口浓度却很低,或者明明使用原料很少,实际监测时的进口浓度却很高等情况。

1.2 产生异常数据的原因

监测数据获取的全过程包括现场采样、样品保存和运输、样品制备、实验室分析等技术过程。环境监测过程中,产生异常数据的原因很多,如何判定和处理异常值,需要结合实际,分析监测数据获取的全过程,才能从源头上降低异常数据的产生概率。

现场采样过程中,现场运行工况条件不稳定或不连续会导致离群数据的产生;污染物处理设施运行条件不正常等会导致不合理数据的产生;现场采样方式不正确、点位选择不正确等均会导致不符合实际的数据产生。样品保存及运输过程中,由于采样瓶或采样管未洗干净,会导致离群数据的产生;样品保存条件不合理会导致不符合实际的数据产生。分析过程中,分析操作不当会导致离群数据的产生;分析方法选择不正确也会导致不合理或不符合实际的数据产生。

2 异常数据的处理

异常数据不一定就是错误的。当监测结果出现异常时,首先应查找异常数据产生的原因,如果从获取数据本身的技术上发现问题,即可将异

常值作为不准确数据予以剔除。当技术上无法断定又非客观实情时就应该在出现监测结果异常时采取相应的措施。

异常数据一般在现场监测和数据审核过程中发现。现场测定的数据,在现场监测过程中就可以被发现。采样后送实验室分析出现的异常数据通过审核而发现。

2.1 现场监测结果出现异常数据的处理

(1) 仪器出现问题。首先应重新启动仪器校正测试,并以备用的校正液标准品重新确认测量仪器是否正常。或者使用备用的仪器设备,重新监测确认。

(2) 控制条件发生变化。如pH值测定时,试样与标准缓冲液温度相差太大,可以将缓冲溶液在现场平衡一段时间后再监测。

(3) 核查现场测试是否符合监测技术规范要求。如污染源监测时,监测低温、高湿的废气中SO₂和NO_x浓度,需要将采样管加热,对烟气预处理,防止水分在采集管里冷凝。在实际采样过程中,有时会直接用聚四氟乙烯管代替加热烟枪进行采样。又如在固定污染源中监测颗粒物浓度时,由于找不到合适的采样断面,就将靠近弯道处作为采样断面,导致采集到的数据异常,无法正确地反映客观事实^[1-2]。

(4) 核查所用的监测仪器测定范围、使用条件是否满足监测要求。如污染源废气监测时,高浓度一氧化碳(CO)会对定电位电解法测定SO₂产生正干扰,而对非分散红外法基本无干扰。因此,在高浓度CO工况时测试SO₂应选择非分散红外仪器。又如机场噪声监测时,飞机起飞或降落时噪声声级达到130 dB以上,但所用声级计测定上限为120 dB,这时应选择可扩展监测上限的声级计。

(5) 污染源现场监测时,核查生产设备和污染治理设施是否正常运行,生产工艺是否与设计相符,所用的原辅材料、设备运行参数及投加的药剂是否与设计参数相符。同时对现场周围环境进行观察,并找出其他可能测值的原因。当现场采样时发现工况异常偏低,废气参数出现异常等情况(如含氧量异常偏高),应立即停止采样,与企业沟通恢复正常后再继续采样。地表水监测时,某河道的闸口上下游地表水采样时,闸口开闸放水,导致沉积物冲刷,产生异常数据(如悬浮物、总磷数值异常偏高)。当现场检查过程中发现异常状况,应如实记

录,拍照存证,以便后期数据分析时查找原因。

2.2 审核数据时发现异常数据的处理

数据审核基本上通过数据合理性评估、监测分析方法适用性或干扰因素确认、历次监测数据比对等方式进行,筛选不合理数据,对异常数据进行确认,判定数据的准确性。

(1) 调阅现场采样记录,包括查看现场采样记录中关于工况、周边环境、天气状况、固定剂添加情况、点位布设、采样方法、采样时间、仪器是否过检定周期的记录,了解是否存在异常状况。例如,某化工企业进行洗涤塔氯化氢去除效率测试时,结果发现进口氯化氢浓度低于出口氯化氢浓度的不合理现象。经与现场人员核实在采样规范性、同步性均符合要求的情况下,得知洗涤塔淋洗液长期未更换,淋洗液已接近吸收饱和,洗涤塔起不到吸收效果,吸收液中氯化氢反而存在进一步释放的情况,导致进出口浓度倒置的不合理数据产生。又例如,某验收项目,连续两天的监测过程中,出现某一次废气监测值明显低于其他监测结果的离群数据情况。经与企业核实,发现由于生产装置(反应釜)非连续运行,每次投料、反应、出料整个运行周期仅1 h,连续几次采样的时段中有一次并未在反应周期内,因而导致监测结果离群。当因现场采样条件及方法等原因导致数据异常时,应排除相关原因后复测。

(2) 调阅分析报告与原始数据,进行检查。对质控数据进行审核,包括对空白值、标准曲线的相关系数、标样、平行样、加标回收率等数据审核。进行数据整体性的合理性分析,针对无法合理分析的测值,进行分析确认。分析确认具体方式见图1。

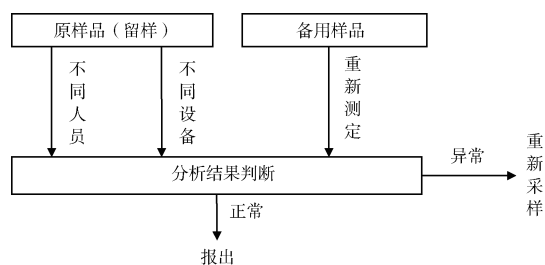


图1 分析确认方式及步骤

(3) 监测分析方法适用性或干扰因素的确认。例如,地表水总磷监测,地表水中泥沙含量、样品固定剂添加量、样品保存时间、样品前处理方式等对样品的分析结果干扰较大,实际分析时,需要对干

扰条件进行确认,以便能准确反映水质总磷浓度状况。污染源监测时,当废气中CO浓度较高,超出干扰实验允许范围时,定电位电解法不再适用于SO₂的测定,此时需要更换方法。

3 提高环境监测数据质量的措施

3.1 建立健全环境监测质量管理体系

监测能力是基础性资源,高品质且具有活力的人力资源是产生优质监测数据产品的必备条件^[3]。建立与监测业务相适应的监测质量管理体系,重点加强高水平环境监测人员队伍建设,加强环境监测技术开发和维护,提升环境监测能力,做好环境监测质量监督检查。开展质量检查、同步监测、比对监测、能力验证以及质控考核等多种形式的质量监督活动。在实施质量监督的同时开展监测方法、操作条件以及质量控制措施等多方面的深入研究,切实解决监测过程中存在的问题和难点,提升质量监督活动的实效。制定相关的工作程序、采样规范、作业指导书,对监测人员进行有效的培训和监督,确保现场采样、样品保存和运输、样品制备、实验室分析、数据传输等各个环节按规范执行。

3.2 重视现场采样的信息化管理

现场采集具有代表性的样品是确保监测数据准确的前提条件。由于现场采样情况复杂多变,环境恶劣,长期以来可采取的质控措施非常有限。近年来,随着信息技术的不断发展,应加强对现场的信息化管理。现场采样可采取全程视频监控的方式,地理信息定位系统追踪现场监测点位,现场采集数据实时传输,确保现场监测数据的溯源性。

3.3 正确填写监测原始记录

监测原始记录在第一时间及时、完整、规范地填写,是监测质量控制的关键^[4]。原始记录格式设计不能过于简单,内容应包括编制检验报告所需的全部信息及能复原监测的足够信息,包括:样品名称、编号、监测项目、地点、日期、测试依据和方法、监测结果及参加监测人员(监测人、复核人)的签名以及监测环境条件、仪器名称等。监测原始记录应由监测人员在监测过程中及时填写,不得凭回忆追记或事后抄正,不得随意更改。必要时对现场发现的异常情况备注说明。

3.4 加强监测数据的审核

通过监测数据审核,对数据进行科学合理的分析,可有效减少异常数据的产生^[5-6]。按《检验检

测机构资质认定生态环境监测机构评审补充要求》,对现场监测原始记录、实验室分析原始记录进行三级审核,选择具有丰富的现场监测经验、实验室分析经验的人员担当审核人员,明确各级审核人员的审核要求。

4 结语

随着国家和民众对环保工作的日益重视,出具准确可靠、高质量的监测数据也成为环境监测机构越来越迫切的要求。管理一方面是靠数据说话,另一方面也要用数据说话,环境监测作为环境管理的技术支撑手段,其依据就是监测数据。环境监测人员要不断提高自身素质,不断更新监测理论与技术,以认真严谨的工作态度保证环境监测数据的准确可靠,更好地服务于环保事业,促进生态文明建设的良好发展。

(上接第15页)

低值区主要分布在西南区域, $\rho(\text{PM}_{2.5})$ 在东西方向上的空间异质性程度要大于南北方向。随着站点之间空间距离的增加,各个站点局地污染分布因素的差异性逐渐增大。

(3)受市区重点污染源分布和气象条件影响,常州市区 $\rho(\text{PM}_{2.5})$ 呈现沿东北向西南区域依次递减的分区特征,高值区位于常州市区中心偏北、偏东地区,低值区位于市区西南部区域,与年均值空间分布结果具有较强的一致性。 $\rho(\text{PM}_{2.5})$ 分区特征具有明显的季节变化特征,影响因素主要为气象扩散条件。

[参考文献]

- [1] 陈渤黎,雷正翠,董芹,等. 2012—2014年常州市大气重污染日的气象条件分析[J]. 气象与环境科学, 2017, 40(1): 87-94.
- [2] 余倩楠,徐茜,周陶冶,等. 长三角地区2015年大气重污染特征及其影响因素[J]. 环境科学学报, 2018, 38(8): 3185-3196.
- [3] 温佳薇,贺军亮,赵晴,等. 京津冀地区 $\text{PM}_{2.5}$ 的时空效应研究[J]. 环境污染与防治, 2018, 40(9): 1033-1043.
- [4] 张丽华,武捷春,包玉海,等. 西安市 $\text{PM}_{2.5}$ 的时空分布特征研究[J]. 环境工程, 2016(2): 87-90.
- [5] 赵阳,胡恭任,于瑞莲,等. 2013年南昌市区 $\text{PM}_{2.5}$ 的浓度水平及时空分布特征与来源[J]. 环境科学研究, 2017, 30(6): 854-863.

[参考文献]

- [1] 胡勇,冯永超. 环境监测中异常数据研究及具体实例的探讨[J]. 资源节约与环保, 2018(4): 68-69.
- [2] 牟永铭. 环保验收监测异常数据的分析与处理[J]. 环境监测管理与技术, 2013, 25(2): 4-6.
- [3] 郁建桥,丁曦宁,刘宁锴. 持续提升环境社会检测数据质量工作探讨[J]. 环境监控与预警, 2019, 11(3): 54-57.
- [4] 蔡同锋. 现场监测数据审核方法及异常数据的判断研究[J]. 环境科学与管理, 2014, 39(10): 60-62.
- [5] 张芳,袁福的,李六一,等. 环境监测数据审核中各类数据逻辑关系的研究[J]. 环境科学与管理, 2008, 33(11): 138-141.
- [6] 李志明. 环境监测数据审核及异常数据的处理[J]. 新疆环境保护, 2013, 35(2): 41-44.

栏目编辑 谭艳 周立平

- [6] 张云芝,阿拉腾,图娅,等. 基于地面监测数据的2014年京津冀地区 $\text{PM}_{2.5}$ 时空分布研究[J]. 环境污染与防治, 2016, 38(12): 31-37.
- [7] 顾康康,祝玲玲. 合肥市主城区 $\text{PM}_{2.5}$ 时空分布特征研究[J]. 生态环境学报, 2018, 27(6): 1107-1112.
- [8] 张祥志,陈文泰,黄樱等. 江苏省2013—2016年臭氧时空分布特征[J]. 中国环境监测, 2017, 33(4): 50-59.
- [9] 周义昌,翁佳烽,周静,等. 肇庆城区空气污染 EOF 分解及与气象要素的相关性分析[J]. 广东气象, 2018, 40(5): 48-50.
- [10] 滕加泉,程钟,梁丹妮,等. 常州市大气 $\text{PM}_{2.5}$ 中水溶性离子组成及来源[J]. 长江流域资源与环境, 2016, 25(9): 1368-1374.
- [11] 王强,戴玄吏,巢文军,等. 常州市春季大气 $\text{PM}_{2.5}$ 中金属元素的分析及污染特征[J]. 环境工程学报, 2015, 9(1): 323-330.
- [12] 李清,黄雯倩,马帅帅,等. 常州春季 $\text{PM}_{2.5}$ 中 WSOC 和 WSON 的污染特征与来源解析[J]. 环境科学, 2019, 40(1): 94-103.
- [13] 滕加泉,王唯,蒋少杰,等. 常州市大气 $\text{PM}_{2.5}$ 主要排放源的成分谱研究[J]. 环境科技, 2015, 28(6): 56-64.
- [14] 葛跃,王明新,白雪,等. 苏锡常地区 $\text{PM}_{2.5}$ 污染特征及其潜在源区分析[J]. 环境科学学报, 2017, 37(3): 803-813.
- [15] 刘菲,薛银刚,屠博文. 常州市春季 $\text{PM}_{2.5}$ 中细菌群落特征及影响因素[J]. 中国环境科学, 2018, 38(9): 3254-3261.
- [16] 江晓栋,薛银刚,魏永,等. 基于成组生物毒性测试的 $\text{PM}_{2.5}$ 毒性[J]. 环境科学, 2019, 40(8): 3840-3848.
- [17] 陈志永,陈小岳,董慧斌,等. 常州市大气细颗粒物对老年人死亡影响的时间序列研究[J]. 现代预防医学, 2019, 46(7): 1181-1185.