

· 监管新论 ·

# 环境空气自动监测数据审核中异常数据判断和处理

管擎宇

(上海市嘉定区环境监测站, 上海 201822)

**摘要:**以上海嘉定区环境空气自动监测数据为依据,利用中美合作-上海市环境空气质量发布系统示范项目(AIRNOW-International)提供的DMS数据自动审核模块,结合市区两级人工审核系统,对数据有效性审核中的异常数据情形进行归类和总结。以实例的方式介绍了环境空气常规六参数常见的异常数据情形,并提出应对异常数据的审核程序和办法。

**关键词:**环境空气;自动监测数据;数据审核;异常数据

中图分类号:X831

文献标志码:B

文章编号:1674-6732(2016)05-0059-05

## Judgment and Handling of Abnormal Data during Ambient Air Automatic Monitoring Data Audit

GUAN Qing-yu

(Jiading Environmental Monitoring Station, Shanghai 201822, China)

**Abstract:** With the wide application of ambient air automatic monitoring data, the data audit is one of the key links to ensure data validity. Based on the ambient air automatic monitoring data of Jiading district, using the automatic data audit modular of DMS provide by USEPA and Shanghai EPB cooperation project (AIRNOW-International), combined with urban two-stage artificial audit system, data validity of abnormal data was classified and summarized. This article introduced six conventional parameters of abnormal data of the ambient air using practical examples and proposed audit procedures and methods of dealing with the abnormal data.

**Key words:** Ambient air; Automatic monitoring data; Data audit; Abnormal data

随着环境空气自动监测技术的不断发展,自动监测技术已经逐步趋于成熟和稳定<sup>[1-2]</sup>,环境空气自动监测站点逐步建成了国家、省级和市级3个层面的环境空气质量监测网络,实现了对全国主要城市环境空气质量数据的实时监测和发布。而环境空气自动站因仪器故障、运维不当、恶劣天气等原因造成在线数据偏离正常测量范围的情况时有发生,过去仅依靠人工审核数据的方式已愈来愈不能满足实际的工作要求,因此急需建立一套较为完整的审核体系来应对工作量巨大的自动监测数据审核工作,确保环境空气自动监测数据的有效性、连续性和准确性。

### 1 数据审核方法和流程

嘉定区空气自动监测数据审核体系主要依靠上海市环境监测中心的AIRNOW系统的DMS软件自动审核和市区两级人工审核两部分组成。AIRNOW-国际版是中美环保部门合作开发的环

境空气质量数据管理和发布平台,通过DMS数据审核模块,设置判断规则能够实现自动站监测数据的自动判断和处理<sup>[3]</sup>。自动审核规则的设定主要根据《环境空气质量自动监测技术规范》(HJ/T 193-2005)中第5条中异常值的取舍以及《环境空气质量标准》(GB 3095-2012)中第6条数据统计有效性的相关规定<sup>[4-5]</sup>。

人工审核采用市区两级专业技术人员审核体系架构(图1)。一审由区县站负责,主要职责包括:对DMS模块自动审核数据的有效性进行复核,即对软件自动审核为无效的数据(带有标记“D”“B”的数据)进行再次确认,确保软件自动审核数据的准确性;通过站点间趋势变化、关联参数同步分析以及仪器参数分析等方法综合判断异常数据;对审核怀疑的异常数据,要求运维人员现场检查,

收稿日期:2016-05-19;修订日期:2016-05-27

作者简介:管擎宇(1985—),女,工程师,硕士,从事环境工程工作。

根据实际情况对仪器进行校正和维护后判断数据是否异常并在审核平台提交异常数据和原因。二审由市监测中心负责,主要职责为:对区县提交的异常数据进行再次确认。

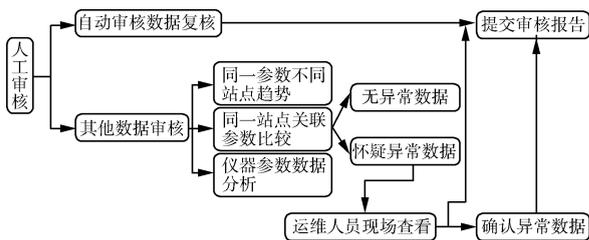


图1 人工审核流程

## 2 空气自动监测仪器及分析方法

以上海市嘉定区3个环境空气自动站的监测数据为依据,嘉定区环境空气自动站在线监测仪器全部引进美国赛默飞世尔科技(中国)有限公司,空气自动站在线监测参数分别为:SO<sub>2</sub>、NO<sub>2</sub>、CO、O<sub>3</sub>、PM<sub>10</sub>、PM<sub>2.5</sub>。环境空气自动监测站由六项环境空气质量评价参数和气象五参数构成,仪器的型号和分析方法见表1。

表1 环境空气自动监测仪器分析方法及仪器型号

监测项目	仪器型号	在线分析方法
SO <sub>2</sub>	43i	紫外脉冲荧光法
NO <sub>2</sub>	42i	化学发光法
CO	48i	非分散红外吸收法
O <sub>3</sub>	49i	紫外吸收法
PM <sub>10</sub>	TEOM1405	微量振荡天平法
PM <sub>2.5</sub>	TEOM1405 - FDMS	微量振荡天平法 - 膜动态测量系统

## 3 异常数据的判断和处理

陈伟<sup>[6]</sup>、刘三长<sup>[7]</sup>等对美国API公司生产的SO<sub>2</sub>、NO<sub>x</sub>和O<sub>3</sub>分析仪在运行中常见的故障进行了整理和归类。姚文辉<sup>[8]</sup>对长光程大气自动监测中异常数据进行了分析和处理。毛茂南<sup>[9]</sup>在环境空气质量自动监测数据审核中提到了三级审核的概念。这些仪器故障判断和异常数据的分析对自动监测数据有效性审核提供了良好的经验借鉴,但是随着自动监测技术的不断发展,环境空气质量评价参数的增加以及自动监测仪器的不断更新,在数据审核过程中除了要借鉴传统的异常数据判断经验外,更要结合实际情况多方面分析和判断,在审核过程中不断积累经验 and 提高异常数据判断能力,才

能对自动监测数据进行合理和准确的有效性审核。

对嘉定区3个环境空气自动站2014—2016年自动监测数据审核中发现的异常情况进行分类和整理,包括以下4类:(1)仪器故障导致在线监测数据的测量离群(偏高或偏低);(2)监测点周围有异常污染源导致监测数据明显高于其他监测点位;(3)运维人员疏忽导致异常数据的产生以及其他一些不可控因素(停电、断网)等造成的在线数据异常;(4)恶劣天气导致测量值的连续异常。

### 3.1 仪器故障导致的异常数据

#### 3.1.1 CO分析仪常见的仪器故障及典型实例

CO分析仪采用非分散红外吸收法测量,测量原理为CO对4.67 μm的红外光辐射有吸收作用。因此红外光源强度及其稳定性是影响CO分析仪测量精度的关键因素之一。在数据审核中曾发现因CO红外光源损坏导致测量值出现异常高正值和负值。图2为某日A站点CO分析仪红外光源损坏的情形。A站点21:00—22:00 ρ(CO)从0.2 mg/m<sup>3</sup>突增至2 188 mg/m<sup>3</sup>,然后23:00突降至-402 mg/m<sup>3</sup>。运维人员到现场检查后发现CO分析仪红外光源被打穿后导致测量值出现负数,更换新的红外光源后恢复正常。因此,从该日22:00—次日8:00 A站点CO的测量数据均为无效并剔除。

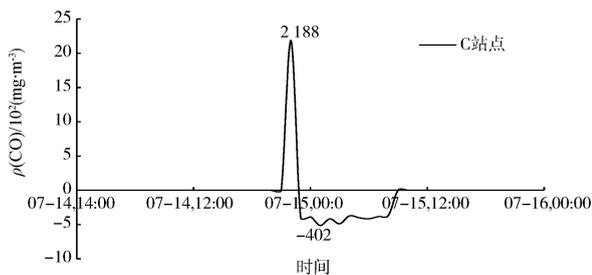


图2 2015年CO分析仪红外光源故障异常情况

除了红外光源外,有时会发生CO分析仪泵轴卡死的情况,在线测量数据会突然降至0点附近,图3为某日B站点的CO分析仪泵轴卡死的典型情况,B站点ρ(CO)从某日01:00开始从0.4 mg/m<sup>3</sup>突降至0.06 mg/m<sup>3</sup>,随后7 h较其原先小时浓度值一直处于较低的水平。运维人员在现场对仪器进行检查,发现泵轴卡死无法抽到气体,更换泵轴后09:00使用CO标准气体与动态稀释仪对CO分析仪工作点(3.5 mg/m<sup>3</sup>)进行查核后于

10:00 恢复正常。因此,从该日 01:00—09:00 B 站点 CO 的测量数据均为无效并剔除。

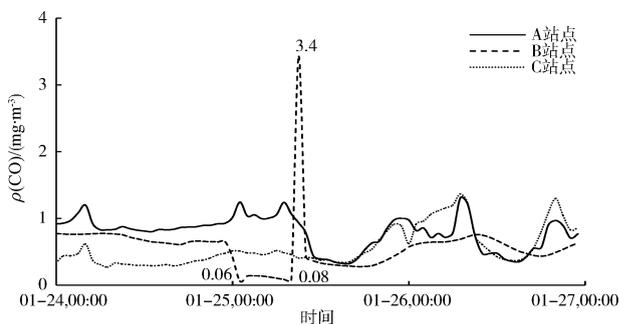


图3 2016年CO分析仪泵轴卡死异常情况

### 3.1.2 $\rho(\text{PM}_{2.5})$ 异常原因分析

$\text{PM}_{2.5}$  在线分析仪采用微量振荡天平法进行测量,流量是影响颗粒物在线测量准确性的关键指标之一。在日常数据审核中, $\text{PM}_{2.5}$  数据有效性审核除通过多个站点的趋势变化分析外,还要结合  $\text{PM}_{10}$  的相关指标浓度值进行比较,通过两者之间的占比变化和变化趋势判断颗粒物数据是否异常。

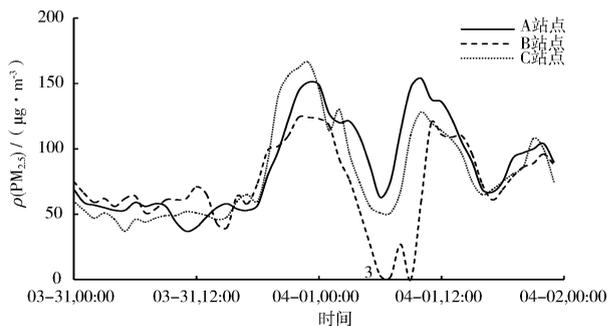


图4 2016年 $\text{PM}_{2.5}$ 分析仪异常数据

图4为 $\text{PM}_{2.5}$ 在线分析仪流量异常情况典型案例。从某日05:00开始B站点 $\rho(\text{PM}_{2.5})$ 突然呈直线下降趋势,到06:00已下降至 $3 \mu\text{g}/\text{m}^3$ ,而A、C 2个点位数据从05:00—06:00虽然也呈下降趋势,但是其总体水平维持在 $50 \sim 60 \mu\text{g}/\text{m}^3$ 范围内。随后2h内B站点 $\rho(\text{PM}_{2.5})$ 始终保持在较低水平。

观察相同时段内 $\rho(\text{PM}_{10})$ 变化趋势,见图5。根据 $\text{PM}_{10}$ 监测数据显示,A、B、C 3个站点在相同时段内变化趋势与 $\rho(\text{PM}_{2.5})$ 总体一致,且B站点 $\rho(\text{PM}_{10})$ 为 $56 \sim 144 \mu\text{g}/\text{m}^3$ 。与相同时段内的 $\rho(\text{PM}_{2.5})$ 相比未出现明显下降的趋势。

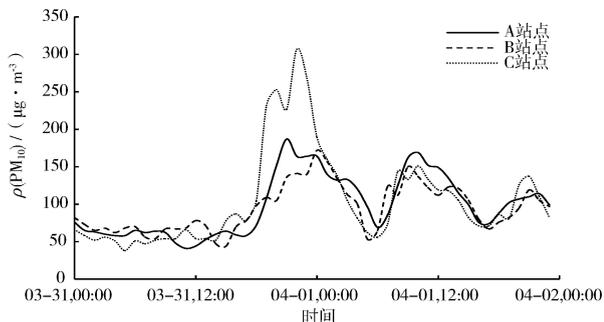


图5 2016年 $\text{PM}_{10}$ 分析仪在线监测数据

根据上述数据趋势变化以及相关指标之间的浓度值分析,判断该时段内 $\text{PM}_{2.5}$ 监测数据异常,通过远程查看 $\text{PM}_{2.5}$ 该时段内的仪器状态参数,滤膜负载率在 $55\% \sim 56\%$ ,环境空气湿度在 $94\% \sim 97\%$ ,机箱温度为 $35 \pm 0.2 \text{ }^\circ\text{C}$ ,颗粒物分析仪主流量在 $3.0 \pm 0.03 \text{ L}/\text{min}$ ,辅流量变化幅度较大。运维人员对主、辅流量分别进行测量后发现辅流量从 $13.67 \text{ L}/\text{min}$ 降低至 $8.42 \text{ L}/\text{min}$ ,因当日下雨空气潮湿,辅流量滤芯堵塞导致流量变化异常,更换滤膜后测量值恢复正常。因此B站点05:00—08:00  $\text{PM}_{2.5}$ 的测量值均审核为无效数据并剔除。

### 3.2 站点周围短时污染源导致异常数据

在日常数据审核中,除了因仪器零配件故障或者管路堵塞造成的异常数据外,有时因站点周围污染源排放导致在线数据测量值异常偏高。该类数据在有效性审核过程中不能剔除,应予以保留。

#### 3.2.1 $\rho(\text{SO}_2)$ 异常偏高

图6为 $\rho(\text{SO}_2)$ 因周围污染源短时排放造成突增的典型示例。某日07:00发现C站点 $\rho(\text{SO}_2)$ 在线数据较其他2个站点明显偏高,A站点值为 $39 \mu\text{g}/\text{m}^3$ ,B站点值为 $17 \mu\text{g}/\text{m}^3$ ,而C站点值为 $156 \mu\text{g}/\text{m}^3$ 。通过前推24h查看3个站点的 $\rho(\text{SO}_2)$ ,3个站点的值在 $(20 \pm 5) \mu\text{g}/\text{m}^3$ 范围内,且站点间数据一致性较好。通过后推8h观察站点的的变化,除07:00的C站点值有明显偏高外,后8h的3个站点的的变化也逐渐趋于平稳和接近,3个站点值在 $(40 \pm 10) \mu\text{g}/\text{m}^3$ 范围内。由此可见,07:00 C站点 $\rho(\text{SO}_2)$ 偏高仅是单个小时的数据异常。

通过Teamviewer软件远程查看C站点05:00—09:00  $\text{SO}_2$ 在线分析仪的电压、压力、流量、光电倍增管、灯压和灯信号(表2)。根据在线仪器

的各参数显示, C 站点在异常时间点 07:00 前后 2 h 内仪器的各项参数均正常, 由此看出在线仪器运行正常。

通过仪器状态参数查看和数据分析显示, C 站点 SO<sub>2</sub> 在 05:00—09:00 之间仪器各项参数均在正常范围内, 且灯强、光电倍增管、反应室温度、采样流量等参数稳定性较好, 由此判断该时间段内 SO<sub>2</sub> 在线监测分析仪的测量正常。

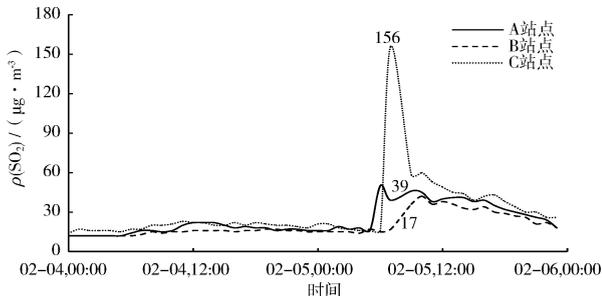


图 6 2016 年 SO<sub>2</sub> 分析仪在线监测数据

表 2 2016 年 SO<sub>2</sub> 在线分析仪仪器参数情况统计

时间	02-05,05:00	02-05,06:00	02-05,07:00	02-05,08:00	02-05,09:00	范围
内部温度/℃	22.4	22.5	22.4	22.6	22.7	15~45
压力/10 <sup>4</sup> Pa	9.902	9.907	9.915	9.902	9.919	5.333~13.33
PMT voltage/V	-736.3	-736.7	-736.3	-736.3	-736.3	
SO <sub>2</sub> coef	0.927	0.927	0.927	0.927	0.927	0.85~1.25
反应室温度/℃	44.8	44.7	45.1	45.0	44.8	43~47
采样流量/(L·min <sup>-1</sup> )	0.597	0.597	0.598	0.597	0.597	0.35~0.75
φ(SO <sub>2</sub> bkg)/10 <sup>-9</sup>	35.1	35.1	35.1	35.1	35.1	
灯电压/V	859.081	857.868	858.677	854.634	859.485	750~1200
Lamp INT/Hz	89	89	89	89	88	40~100

排除仪器故障后, 运维人员赶到现场对仪器进行进一步的复核确认, 此外对站点附近情况进行现场调查发现: C 站点附近 20 m 处的内河 07:00 因维护柴油巡逻艇导致短时 ρ(SO<sub>2</sub>) 迅速上升, 待维护结束后, 随着大气扩散和污染源停止排放后, 08:00 ρ(SO<sub>2</sub>) 迅速降低至正常浓度水平。由此判断得出该异常高值并非仪器原因, 而是因站点周围的短时污染排放导致测量值短时上升, 因此该小时浓度值数据有效性审核为有效, 予以保留并填写异常偏高原因。

### 3.3 运维不当造成数据异常

有时因运维人员遗忘关闭通道或者运维结束后忘记及时开启通道, 数据标记位显示为运维状态, 系统自动审核为无效。表 3 为某日 C 站点 O<sub>3</sub> 数据标记位状态表, 根据空气自动站的 OPcode 和 QCcode 显示运维人员于 11:00 对 C 站点 O<sub>3</sub> 进行维护和校准, 运维人员将 C 站点的数据上传通道切换为维护状态, 数据 OPcode 显示为“B”, QCcode 显示为“无效”, 软件将该时段数据自动审核为无效数据。

12:00—16:00 的 C 站点 O<sub>3</sub> 的 OPcode 始终显示为“B”维护状态, QCcode 根据维护状态“B”的标记自动将该时段的数据审核为“无效”。根据标记位显示, 该站点从 11:00—16:00 O<sub>3</sub> 在线分析仪一直处于运维状态, 但根据审核经验判断,

表 3 2014 年 10 月 9 日 C 站点 O<sub>3</sub> 数据标记位 (OPcode)

时间	监测值/(μg·m <sup>-3</sup> )	OPcode <sup>①</sup>	QCcode <sup>②</sup>
09:00	107		
10:00	184		
11:00	314	B	无效
12:00	123	B	无效
13:00	113	B	无效
14:00	114	B	无效
15:00	115	B	无效
16:00	90	B	无效
17:00	70		

①为数据标记位代码; ②为数据有效性判断。

O<sub>3</sub> 在线分析仪的常规运维时间约为 20~30 min。因此对该时间段的 O<sub>3</sub> 运维状态“B”的在线数据提出异常数据怀疑。比较异常时段 C 站点的 O<sub>3</sub> 监测值以及前推 24 h 相同时段的 O<sub>3</sub> 监测值见图 7。

根据图 7 比较显示, 除 11:00 的数据有明显差距外, 10 月 8 日和 10 月 9 日的 ρ(O<sub>3</sub>) 变化趋势相同, 浓度值也较为接近。其中 C 站点 10 月 8 日的 ρ(O<sub>3</sub>) 总体低于 10 月 9 日。且随着时间的推移呈逐渐降低趋势, 这与 O<sub>3</sub> 在一天中的变化规律也相似[在晴天条件下, ρ(O<sub>3</sub>) 呈单峰变化趋势, 13:00 左右最高, 随后紫外线逐渐降低, 其值也随之下降]。由此判断 12:00—16:00 的 O<sub>3</sub> 的数据标记位“B”数据有异常。

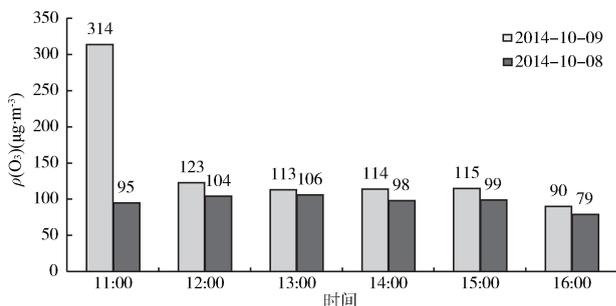


图7 异常时段内C站点 $\rho_{\text{小时}}(\text{O}_3)$ 变化比较

与运维人员电话沟通后得知其在11:00仪器校正结束后,忘记把运维通道切换成正常测量通道,导致监测数据虽然已经正常测量,但 $\text{O}_3$ 的数据标记仍然显示为维护“B”状态,因此12:00—16:00测量数据为正常测量,应将其审核为有效。

### 3.4 其他异常情况

在日常数据审核中,除了因仪器故障、站点周边短时污染源、运维人员主观因素等情况导致在线测量数据的异常外,有时因恶劣天气(如暴雨、雷电、大风)也会造成数据异常。例如:暴雨会导致 $\rho(\text{PM}_{10})$ 出现负值,见图8。

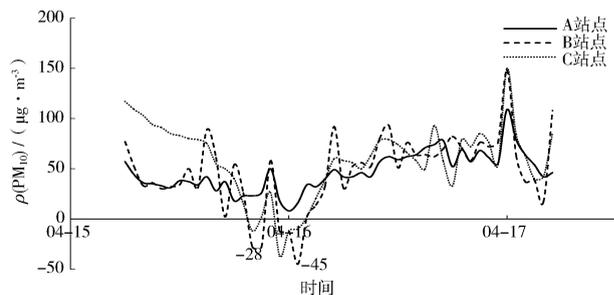


图8  $\rho(\text{PM}_{10})$ 出现负数异常数据

$\text{PM}_{10}$ 分析仪采用微量振荡天平法测量,随着空气中湿度增加,从某日零点开始,随着降雨量逐渐增加,空气中颗粒物 $\rho(\text{PM}_{10})$ 先逐渐降低,随着雨量进一步增加,环境空气中的颗粒物已经趋于较低水平,此时滤膜上水分的挥发量远大于颗粒物浓度的增加量,导致测量值出现负数。这与包权<sup>[3]</sup>在颗粒物数据审核要点中提出连续多天下雨时往往会产生大量瞬时负值的结论是相同的。

### 3.5 异常数据审核统计

据统计,2014—2016年嘉定区3个环境空气

自动监测子站数据有效性审核中共发现异常情况40次。其中,系统自动审核发现的异常数据3次,占总数的7.5%;两级人工审核发现的异常数据32次,占总数的80.0%;通过经验判断对审核无效的数据进行保留的数据5次,占总数的12.5%。由此看出,两级人工审核对自动监测数据有效性审核中异常数据的判断起到了关键的作用。

## 4 结语

环境空气自动监测数据有效性审核过程主要包括3个方面:

(1) 自动化审核规则设定,通过AIRNOW系统的DMS软件设定自动审核规则可以大大提高自动监测数据审核效率;

(2) 两级人工审核,区县站负责数据一审,发现并查找异常数据及原因,市级站负责数据二审,核查异常数据原因并确认数据有效性;

(3) 审核人员的数据审核经验积累,异常数据的发现和判断需要综合分析,不仅需要对仪器原理和常见故障等知识不断积累,更要对子站周边情况进行熟悉和了解,综合各方面后才能对自动监测数据的有效性进行及时和准确的判断。

### [参考文献]

- [1] IVAN K, MICHAEL W F, AMIRHOSSEIN H, et al. Air quality and environmental impacts of alter-native vehicle technologies in Ontario [J]. Canada International Journal of Hydrogen Energy, 2010, 35(10):5145-5153.
- [2] 廖乾邑, 陈建文, 李亮, 等. 浅谈空气自动监测数据质量现状与处理方法[J]. 四川环境, 2011, 30(2):30-32.
- [3] 包权. 市和区县两级环境空气自动监测数据审核系统在上海的应用[J]. 环境监控与预警, 2015, 7(2):26-28.
- [4] 环境保护部. 环境空气质量自动监测技术规范: HJ/T 193-2005[S]. 北京: 中国环境科学出版社, 2005.
- [5] 国家环境保护总局. 环境空气质量标准: GB 3095-2012[S]. 北京: 中国环境科学出版社, 2012.
- [6] 陈伟, 吴楠. 环境空气自动监测系统的常见故障及排除[J]. 环境监测管理与技术, 2002, 14(1):36-37.
- [7] 刘三长, 黎泳珊, 杜娟. 环境空气自动监测异常数据的判断与处理[J]. 环境监测管理与技术, 2004, 16(2):33-34.
- [8] 姚文辉. 长程大气自动监测中异常数据的分析及处置[J]. 中国环境监测, 2004, 20(4):29-31.
- [9] 毛茂南, 陈建文, 刘艳蓉, 等. 环境空气质量自动监测系统数据的三级审核[J]. 四川环境, 2007, 26(5):31-32.