

· 大气污染防治 ·

空气自动监测网络数据有效性的自动化判别

Automatic Discriminant of Data Validity of Air Auto-Monitoring Network

曲 凯, 李 彦, 崔志伟

(山东省环境信息与监控中心 济南 250012)

摘要 以山东省空气自动监测网监测数据有效性的自动化判别为例, 阐述了自动化判别的主要思路、判别程序和判别规则, 探讨了省级或区域级空气自动监测网络的数据自动化判别方法。

关键词 空气自动监测 监测数据 自动化判别

Abstract Based on automatic discriminant of data validity of air auto-monitoring network of Shandong Province, this paper exposits the main idea of automatic discriminant, the discriminated process and the discriminated rule, and it discusses the methods of Provincial or Regional Network on how to discriminate the data automatically.

Key words Air Auto-Monitoring Monitoring Data Automatic Discriminant

2000年以来, 利用空气自动监测系统可获得连续监测数据的特点, 空气自动监测已成为空气质量监测的主要手段^[1], “十一五”期间, 各省初步建立了以城市监测站为基础的空气自动监测网络, 并实现省级和国家级网络的联网, 为省级和国家级实时评估区域性的环境空气质量, 制定大气污染控制策略提供有力的技术支持^[2]。

由于省级空气自动监测网络是以城市为单位, 面向众多城市的空气自动站的统一管理, 站点众多, 而且自动站是无人值守、自动运行, 维护管理周期较长, 运行过程中出现了问题有时不能及时发现, 就会影响监测数据的准确性, 另外省级空气监测网每天的数据采集量很大, 仅依靠管理人员每日对数据进行肉眼观察, 难免出现遗漏或差错, 不可能做到全面、完整和公正^[3], 因此结合计算机的强大数据处理能力, 对监测网获取数据进行科学高效的有效性判别, 确保数据的代表性、可比性和准确性

就显得十分必要。

山东省结合自身的空气污染特征, 于2008年初建立了我国首个覆盖面最广、监测点位最多的省级空气质量监测网络, 共设监测点位144个, 覆盖了全部17个地级城市, 监测PM₁₀、SO₂、NO₂、CO、O₃ 5个污染项目和气压、温度、风向、风速、湿度5个气象参数。自运行以来, 对监测数据有效性的自动化判别(以下简称“自动化判别”)做了一些有益的探索和实践^[4], 本文就此内容作以阐述。

1 自动化判别的思路

自动化判别主要是根据反映空气自动监测数据获取情况和准确性特征指标的判别方法, 将一系列复杂的人工判别规则归纳提炼形成数学模型, 用软件模拟人工操作, 由计算机对监测网获取的大量数据进行自动判别, 通过判别的数据再由管理人员进行人工确认, 确认通过的作为有效数据参与全省城市空气质量统一评

收稿日期: 2011-04-19

基金项目: 国家自然科学基金项目(40801028)

作者简介: 曲 凯(1982-), 男, 硕士、工程师。研究方向: 环境自动监控。E-mail: qukai250014@163.com

价, 确认不通过的作为无效数据处理。

2 计算机自动判别

计算机自动判别主要是通过定义的一些规则进行自动判定, 不满足规则的数据不参加汇总统计, 并认定为无效数据, 规则来自两个方面^[5, 6]:

(1) 仪器上传的状态值, 包括报警和各种仪器操作等, 被标识的数据被认定为无效, 各种标识如下:

M: 分析仪器维护状态

P: 电源故障

B: 分析仪器运行故障

C: 分析仪器校准状态

O: 超出量程上限

U: 低于量程下限

+: 超过最大读数值限制

-: 低于最小读数值限制

R: 超过读数变化率

E: 对于R232采集数据的通道, 接收数据超时出现此标记, 对于计算通道, 公式错误或依赖数据无效均会出现此标记

>: 有效数据个数所占比例不足100%, 但大于设定比例

<: 有效数据个数所占比例小于设定比例, 数据无效

(2) 人工定义的一些判别规则, 数据出现以下几种情况系统自动判定该天、该子站、该监测仪器数据无效:

① 数据采集时间不满足要求。依据《环境空气质量标准》(GB3095-1996) 各项污染物数据统计有效性的规定, 自动统计每天全省所有子站监测仪器的数据采集频率, 对小时采集时间不足45 min、日采集时间不足18 h (PM_{10} 为12 h) 的上传数据, 不参与空气质量评价。

② 24 h内小时浓度均值变化幅度过小。主要是数据在24 h内变化幅度太小不符合污染物变化规律, 归纳的数学模型为24 h内污染物浓度的最大值(max)减最小值(min)小于一个常规合理差值。由于地域的关系, 沿海城市和内地城市以

及不同的季节这个差值是不同的, 为使管理公正, 系统设计了统计参数表, 管理人员根据季节和地域设置不同的、相对合理的差值参与统计。

③ 浓度日均值小于该污染物最低检出限。

④ 24 h内小时浓度均值出现5个以上负值。

⑤ 24 h内小时浓度均值出现连续5个以上小时无变化。

⑥ 日均值浓度小于该城市全部点位浓度日均值(该点参加均值计算)的一个设定倍数(如一半)。这个倍数由管理人员根据城市的区域范围及点位的具体情况预设系统中。

3 人工确认

人工确认主要是管理人员根据各监测点位、各城市的空气质量变化规律, 对计算机初步判别的数据进行再次确认, 确认出的异常数据由管理人员通过软件做单项或批量的数据无效标记或修正, 最终形成该天、该子站、该监测仪器的有效统计数据, 并参与全省城市空气质量评价。

为提高人工确认的效率, 系统主要采用曲线分析为主, 监测数据为辅的审核办法, 管理人员可以将任意站点、任意项目、任意时间段的数据纳入同一界面下进行比较分析, 当曲线变化不符合客观变化规律, 管理人员调出监测数据找出可疑数据, 根据审核经验或安排现场监测人员去现场调查核实, 对可疑数据进行无效认定或修正, 一些经验规则如下。

(1) 出现连续小时之间的异常跳跃点或多个恒定不变的值, 判定数据无效或修正, 如异常高值、异常低值和连续不变的值。

(2) 根据监测数据与其正常日间走势和正常季节性走势之间的偏差, 判定数据是否有效, 如臭氧浓度在大约正午时候达到其高峰, 不符合此变化趋势的认定为可疑数据; SO_2 浓度在采暖期和非采暖期变化一般比较明显, 数据变化不明显时认定为可疑数据。

(3) 同一城市子站之间、或相邻城市之间在相同时间或相似条件下数据的曲线具有相关性, 相关性差的曲线表示的子站数据可能存在问

题,见图1。

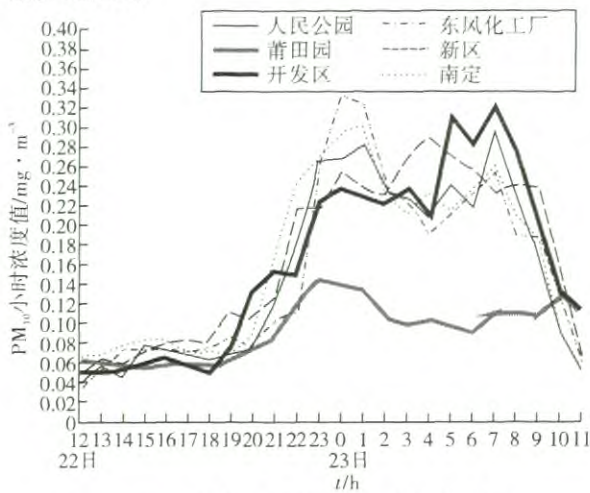


图1 同一城市某一子站曲线相关性差的情况
(以PM₁₀为例)

(4) 参考气象数据分析。比如风速的高低影响污染物的输送和混合,沙尘对大区域影响显著,降雨对大气污染物有净化作用,逆温天气中污染物不容易扩散等,都将影响监测数据的变化。图2为根据逆温天气下数据的变化规律进行判定的案例,图中为冬天某一站点SO₂、NO₂和PM₁₀一天的数据曲线,根据冬天逆温发生规律,一天内多在5:00-10:00、15:00-22:00之间有逆温现象,影响范围广,反映在数据上具有明显的时效性,

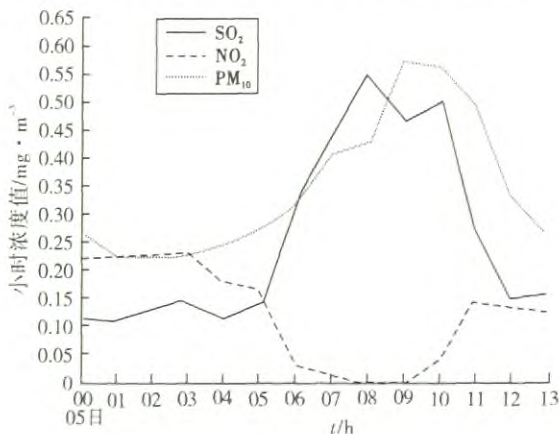


图2 根据逆温现象判定PM₁₀数据存在问题

而且各项污染物的曲线变化基本一致,SO₂和PM₁₀数据变化符合此规律,NO₂数据变化不具备上述特征,认定为可疑数据。

为保证原始监测数据库的安全性,系统设计了修正数据库,将上传数据同时保留在原始数据库和修正数据库中,管理人员所有的修正只在修正数据库中进行,一旦修正数据有误又可方便地用原始数据进行复原。

4 结语

省级空气自动监测网络一般覆盖多个城市,站点众多、数据量大,山东省环保厅结合本省的空气监测网,建立了统一的数据处理平台,开发了数据自动化判别软件,总结了一套符合本省实际的数据判别方法,对监测网上传的数据进行计算机自动判别和人工确认,确认后数据的代表性、可比性和准确性大大提高,目前已用于环境统计、定期通报、以奖代补、环境执法等环境监管的多个方面,监测数据的采用率达95%以上,因此,本文认为,山东省环保厅对数据有效性的自动化判别方法做的探索具有实践意义,适用于省级或区域级的空气自动监测网络的数据有效性判别。

参考文献

- [1] 刘方,王瑞斌,李钢.中国环境空气质量监测现状与发展[J].中国环境监测,2004,20(6):8-10.
- [2] 钟流举,郑君瑜,雷国强,等.空气质量监测网络发展现状与趋势分析[J].中国环境监测,2007,23(2):113-117.
- [3] 刘俊,袁鸾,师建中.区域空气自动监测数据审核方法与制度的探索[J].中国环境监测,2009,25(4):60-63.
- [4] 万黎,杜晓兰,李红莉.空气自动监测网络数据审核自动化实现的创新与实践[J].中国环境监测,2010,26(3):57-59.
- [5] 国家环境保护总局. HJ/T193-2005环境空气质量自动监测技术规范[S].北京:中国标准出版社,2005.
- [6] 国家环境保护总局空气和废气监测分析方法编委会.空气和废气监测分析方法(第四版)[M].北京:中国环境科学出版社,2003:252-271.