

DOI:10.19853/j.zgjsps.1000-4602.2022.14.005

采用公众信息大数据分析污水厂运行状态

姜 阳, 杨成建, 李伟志, 胥 昂, 卢 猛, 李志华

(西安建筑科技大学环境与市政工程学院 西北水资源与环境生态教育部重点实验室, 陕西 西安 710055)

摘 要: 为给政府、企业提供管理污水处理厂的有效建议举措,对微博第三方监管平台和西安市生态环境局的公众信息大数据进行分析,研究了国内污水处理厂的运行现状。结果表明,污水处理厂投诉量与GDP、规模以上工业企业数呈正相关关系,污水超标以化学需氧量、氨氮超标为主,其中79%为短期超标,21%为长期超标,冬季超标严重,企业对超标的反馈以设备故障为主。政府需畅通公众参与渠道、保证排放标准和考核指标的科学性,企业应保证反馈内容的真实性,及时检修设备,需重点关注冬季污水处理厂的超标问题。

关键词: 污水处理厂; 微博; 生态环境局; 公众信息大数据

中图分类号: TU992 **文献标识码:** A **文章编号:** 1000-4602(2022)14-0033-05

Using Public Information Big Data to Analyze the Operation Status of Sewage Treatment Plants in China

JIANG Yang, YANG Cheng-jian, LI Wei-zhi, XU Ang, LU Meng, LI Zhi-hua

(Northwest China Key Laboratory of Water Resources and Environment Ecology, School of Environmental and Municipal Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China)

Abstract: In order to provide government and enterprises with effective suggestions for the management of sewage treatment plants, we analyzed the public information big data from the Micro-blog third-party supervision platform and the Xi'an Ecological Environmental Protection Bureau, and studied the current status of domestic sewage treatment plants. The results show that the number of complaints from sewage treatment plants is positively correlated with GDP and the number of industrial enterprises over designated size. Higher chemical oxygen demand and ammonia nitrogen concentration usually are the two main causes exceeding the wastewater treatment limits. Seventy-nine percent cases exceeding the limits are short-term. Twenty-one percent cases are long-term. The severe exceeding limits phenomenon occurs in winter. The company's feedback on exceeding limits is mainly based on the equipment failure. The government needs to unblock public participation channels and ensure the scientificity of emission standards as well as the assessment indicators. Sewage treatment plants should ensure the authenticity of the feedback content, repair equipment in time, and pay attention on the problem of exceeding limits in

基金项目: 国家自然科学基金资助项目(52070149、51878539); 陕西省重点研发计划项目(2018ZDXM-SF-025); 上海市国资委技术创新和能级提升项目(2018001)

通信作者: 李志华 E-mail: lizhigua@xauat.edu.cn

winter.

Key words: sewage treatment plant; Micro-blog; ecological environment bureau; public information big data

环境问题不仅单方面依靠政府、企业整治,还需要公众参与下的社会监督,我国“十三五”规划纲要指出,要形成“公众、企业、政府”共治的环境治理体系。网络监管是环境共治的一大手段,挖掘网络公开的公众信息大数据的价值是一项挑战。王辰成^[1]利用微博大数据研究地方政府回应网络投诉的影响因素,分析了政府管理在环境污染治理中的作用。李宗敏等^[2]通过收集新冠疫情期间的微博辟谣帖子,运用自然语言处理方法,通过文本特征为辟谣提供新的思路;Li 等^[3]基于网络的搜索引擎和社交媒体文本内容可靠地预测疫情暴发的发展,对于采取行动防止其蔓延至关重要。但是,利用公众信息大数据对国内污水处理厂的研究报道较少。因此,通过研究“公众投诉、企业反馈、行政处罚”的公众信息大数据,以期由政府区域性整治污水处理厂、制定环保政策提供新的分析方法和有效的建议举措。

1 数据来源及研究方法

1.1 数据来源

公众投诉类公告来源于新浪微博“绿色江南公众环境关注中心”下属的 11 个第三方环保监管账号,有效数据约 9 300 条,数据样本时间段为 2018 年 1 月—2020 年 10 月;企业反馈类公告来源于新浪微博“IPE 数据库”,有效数据约 400 条,数据样本时间段为 2018 年 1 月—2020 年 10 月;行政处罚类公告来源于西安市生态环境局,有效数据约 3 400 条,数据样本时间段为 2010 年 1 月—2020 年 10 月。

对采集到的数据进行清洗,删除无关公众投诉、企业反馈、行政处罚公告和缺失的行列,对数据进行整合及标准化。

1.2 研究方法

采用相关性分析、自然语言处理、时间序列分析等研究方法,其中自然语言处理采用短语提取和词频统计,提取公众投诉、企业反馈、行政处罚类公告的中文短语,统计词频及排序,获得部分关键词。本研究中,短语提取和词频统计方法采用 GitHub 开源项目 HanLP。

2 结果与讨论

2.1 省份投诉量与中国年鉴数据的相关性分析

图 1 为 2019 年全国各省份污水超标公众投诉量与中国年鉴部分数据相关性的热力图,GDP、规模以上工业企业数量、城市污水日处理能力有着较强的互相关性,公众投诉量与规模以上工业企业数量、GDP 有着较强的相关性。本次研究中全国公众投诉量约 9 300 条,以投诉量高的山东省(1 239 条投诉)、江苏省(1 473 条投诉)、浙江省(829 条投诉)举例,这些省份经济发达、水资源丰富,政府对工业废水处理设施投资高^[4],有较为公开的信息、畅通的投诉渠道,以及高 GDP、高信息公开度是投诉量高的主要原因。畅通的公众投诉渠道可以找出需重点关注的企业及污水处理厂,为政府进行专项整治提供依据。

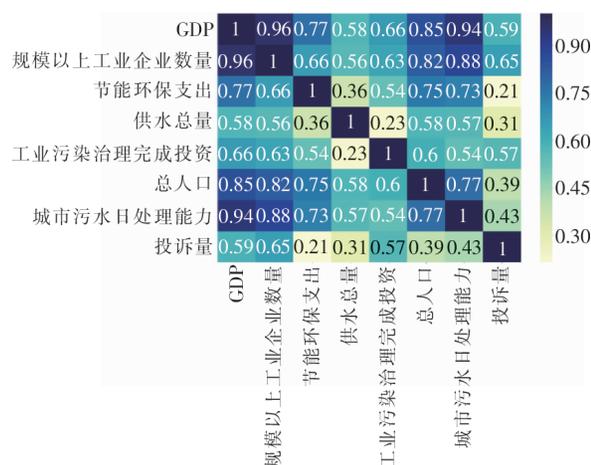


图 1 2019 年全国各省份投诉量与中国年鉴数据相关性的热力图

Fig.1 Correlation diagram between China Yearbook data and the number of public complaints in 2019

针对公众投诉量与 GDP 和规模以上工业企业数量的关系,拟合了如图 2 所示的 2019 年全国各省份投诉量与 GDP、规模以上工业企业数量拟合的气泡图。通过观察拟合的点位变化规律和气泡的大小,发现投诉量与 GDP 和规模以上工业企业数量呈正相关。值得注意的是广东省较江苏省、山东省、浙江省不符合总体变化规律,广东省为国内 GDP 最

高的省份,但其投诉量较少,这可能是由于广东省污水治理成效较好或者排污监测信息的公开程度低。

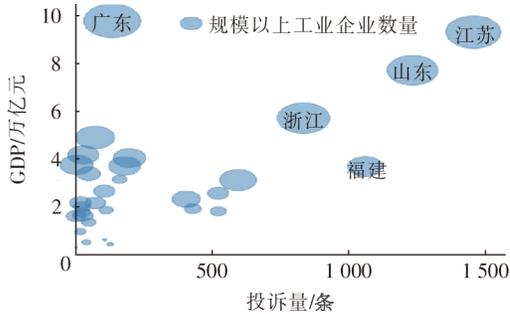


图2 2019年全国各省份投诉量与GDP、规模以上工业企业数量拟合的气泡图

Fig.2 Bubble chart fitting the number of public complaints across the country with GDP and the number of industrial enterprises above designated size

2.2 污水处理厂超标问题的成因分析

在清洗后的有效公众投诉公告中,检索到关于废气(氮氧化物、二氧化硫、烟尘、颗粒物、一氧化碳)超标公告2 486条,污水水质指标(化学需氧量、氨氮、总氮、总磷、pH、重金属)超标投诉公告6 659条,从问题占比上看,27%是关于废气超标问题投诉,73%是关于污水超标问题投诉,污水问题远大于废气问题。图3为6 659条关于污水水质指标超标的公众投诉公告中超标的水质指标占比,污水超标主要由化学需氧量导致(41%),其次是氨氮(17%)、pH(14%)、总氮(13%)、总磷(12%)、重金属(3%)。这可能是由于公众投诉多为产生工业废水的企业,工业废水进水污染物浓度波动幅度大导致碳氮比不稳定,且工业废水污染物毒性大,造成化学需氧量、氨氮成为最容易超标的指标,同时,制药、化工废水的氯离子浓度导致化学需氧量在线监测仪测出的数据不准确^[5]。企业在平时的运行管理中应更注重应对突发水污染事故,重点关注化学需氧量和氨氮超标问题。

企业反馈内容包含污水处理厂超标的原因与解决措施,由此可以分析出造成污水处理厂出水水质超标的原因。图4为企业反馈内容的词云图,其关键词的大小表示该关键词在文本内容中出现的频率。原版词云图内容价值小,包含生态环境局、在线监测、设备超标等无用信息。经过清洗后的词云图能为研究带来更多信息,由图4(b)所示的词云

图可发现企业反馈内容主要是解释超标原因,企业大多指出超标是由设备故障所致,另外零点漂移、采到杂质也是重要原因。同时,通过观察公司类型、水质指标等关键词的大小,发现科技有限公司超标次数较多,化学需氧量超标较总氮、总磷等超标次数更多。

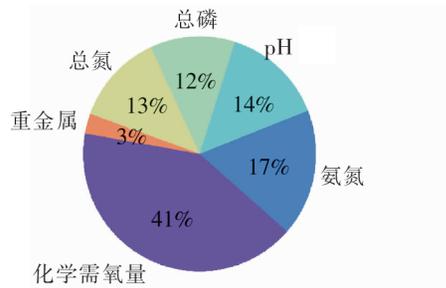


图3 全国2019年污染物超标种类比例

Fig.3 Proportion of pollutants exceeding the standard in China in 2019



a. 企业反馈内容原版词云图 b. 企业反馈内容人工清洗后的词云图

图4 企业反馈内容的词云图

Fig.4 Word cloud diagram of corporate reply content

企业反馈的公告中,大部分公告表达的内容都为企业无重要责任,但是会重点关注。如果企业反馈的内容真实,环保部门需要思考排放标准和考核指标是否科学,如果企业反馈内容多为虚假的推托责任,环保部门需要重点关注企业反馈内容的真实性,并加大监管力度。

因词云图反映的信息有局限性,手动整理了企业反馈公告的内容,并进行了分类分析。图5为企业反馈内容类型的饼状图,其内容包括五类。企业否认超标的原因是在线监测数据与真实数据不一致或已向政府部门备案;设备故障是污水处理厂超标最多的原因,在举报后大多数企业都进行了设备维修,且恢复正常;短期调试也会导致超标,如零点

漂移实验、短期清扫、标准液更换等；工艺改造、管理不当、冬季低温来临没做好相应准备等导致污水处理厂超标归于污水处理厂自身原因；部分污水处理厂也解释是因停电、暴雨、偷排等外在原因导致的超标。

其中，否认超标、设备故障、短期调试可以归为短期超标，总占比 79%，说明大多数投诉问题都能在短期内得到有效解决；自身原因、外界原因超标可以归为长期超标，占比 21%，因为工艺调试、事故恢复需要一定的时间。值得注意的是，在 2020 年冬季的企业反馈公告中，有的企业指出因为新冠疫情导致的消毒剂用量过高，进水冲击过大导致了短期的水质超标^[6]，也有的企业指出因为疫情等限制^[7]，污水处理厂需使用的药剂没有及时供应，这一点发现也为新冠疫情的其他研究提供了信息。

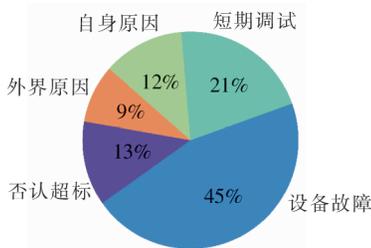


图 5 企业反馈的污水处理厂出水超标原因比例

Fig.5 Proportion diagram of the reasons for the effluent exceeding standard from the sewage treatment plant

2.3 政府关注度的时间序列分析

行政处罚的内容主要是针对违规污水处理厂的行政处罚公示，政府对污水处理厂的处罚次数能表示政府对污水处理行业违规的整治力度，各城市的生态环境局网站是独立的，信息公开程度层次不齐，很多城市生态环境局的数据暂未公开或储存的数据量较少没有代表性。以西安市为例(见图 6)，从行政处罚类公告时间序列上的变化来看，从 2015 年开始，西安市政府加大了对污水处理行业违规企业的整治力度，处罚数大幅提升，其原因可能是 2015 年国家落实《水污染防治行动计划》大力推动了我国水环境产业发展，随着监管体系的完善，政府加强了对排污单位的污染物排放的监管，要求企业必须高度重视^[5]，从处罚量的变化也可以看出政府的确有严格执行。

行政处罚污水处理厂的公告发布时间在四季中的比例可以说明污水处理厂超标的季节性变化，

西安市 2010 年 1 月—2020 年 10 月政府对污水处理厂进行处罚公告发布时间的季节占比见图 7。从季节变化上来看，夏季是行政处罚量最少的季节，冬季为行政处罚量最多的季节，再次说明污水水质超标一定程度上与温度有关，冬季微生物活性低，是污水处理厂运营的一大难题^[8]，企业应重点关注污水处理厂安全过冬的问题。

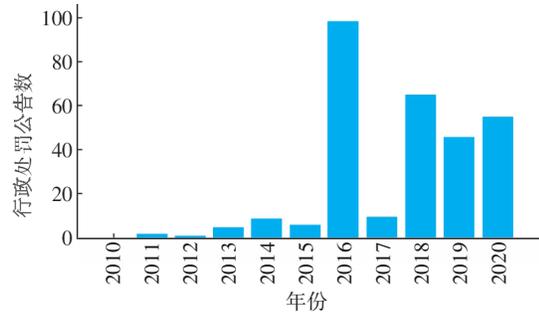


图 6 西安市 2010 年—2020 年行政处罚污水处理厂公告数

Fig.6 Histogram of the number of posts in Xi'an administrative penalized sewage treatment plants from 2010 to 2020

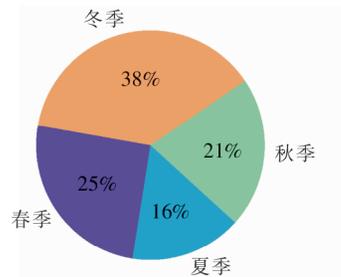


图 7 行政处罚公告的季节比例

Fig.7 Seasonal proportion of administrative penalty posts

3 结论

① 从 2015 年开始，政府加大了对污水处理行业违规企业的整治力度；污水处理厂的公众投诉量与规模以上工业企业数量和氨氮、GDP 呈正相关性，污水超标最多的指标为化学需氧量和氨氮、超标最多的季节为冬季；环境问题 27% 为废气超标，73% 为污水超标，污水超标问题中 79% 为短期影响，21% 为长期影响，企业大多将超标责任归于设备故障，少数企业承认自身原因。

② 公众信息大数据具有重要价值，能为政府整治污水处理厂、制定环保政策提供新的分析方法和有效的建议举措。政府在发展城市工业的同时应加强整治环境，畅通公众投诉的渠道，思考排放标准 and 考核指标是否科学；企业应增加反馈内容的

可信度,注意设备的检修,并重点关注污水处理厂安全过冬的问题。

③ 本研究不足之处在于数据来源有一定局限性,新浪微博为第三方平台,非官方信息发布平台,分析结果可能与实际存在部分偏差;各省市生态环境局信息公开情况参差不齐,建议政府增加环境监管信息的公开程度,以便于后续研究。

参考文献:

- [1] 王辰成. 环境投诉与地方政府回应[D]. 南京:南京大学, 2018:69-74.
WANG Chencheng. Environmental Complaints and Local Government Responses [D]. Nanjing: Nanjing University, 2018:69-74(in Chinese).
- [2] 李宗敏, 张琪, 杜鑫雨. 基于辟谣微博的互动及热门评论情感倾向的辟谣效果研究——以新冠疫情相关辟谣微博为例[J]. 情报杂志, 2020(11):90-95,110.
LI Zongmin, ZHANG Qi, DU Xinyu. Research on rumor-refutation effectiveness based on the interactions and popular comments' emotional tendencies of the rumor-refuting microblogs: taking rumor-refuting microblogs related with COVID-2019 as an example [J]. Journal of Intelligence, 2020(11):90-95,110(in Chinese).
- [3] LI C L, CHEN L J, CHEN X, *et al.* Retrospective analysis of the possibility of predicting the COVID-19 outbreak from internet searches and social media data, China, 2020 [J]. Eurosurveillance, 2020. DOI: 10.2807/1560-7917.ES.2020.25.10.2000199.
- [4] 张统, 李志颖, 董春宏, 等. 我国工业废水处理现状及污染防治对策[J]. 给水排水, 2020, 46(10): 1-3,18.
ZHANG Tong, LI Zhiying, DONG Chunhong, *et al.* Current situation of industrial wastewater treatment and countermeasures of pollution control in China [J]. Water & Wastewater Engineering, 2020, 46(10): 1-3, 18(in Chinese).
- [5] GUO Y, ZENG W, LI N, *et al.* Effect of electron acceptor on community structures of denitrifying polyphosphate accumulating organisms in anaerobic-anoxic-oxic (A²O) process using DNA based stable-isotope probing (DNA-SIP) [J]. Chemical Engineering Journal, 2018, 334: 2039-2049.
- [6] DICKENS B L, KOO J R, WILDER-SMITH A, *et al.* Institutional, not home-based, isolation could contain the COVID-19 outbreak [J]. The Lancet, 2020, 395(10236): 1541-1542.
- [7] 杜昱, 丁西明, 孙月驰, 等. 《水污染防治行动计划》对水处理领域的影响[J]. 环境卫生工程, 2018, 26(2): 76-79.
DU Yu, DING Ximing, SUN Yuechi, *et al.* Impact of "Water Pollution Prevention Action Plan" on water treatment field [J]. Environmental Sanitation Engineering, 2018, 26(2): 76-79(in Chinese).
- [8] 杨小丽, 叶峰, 宋海亮, 等. 基于污水厂运行数据的低温生物脱氮强化研究[J]. 中国给水排水, 2009, 25(1): 82-85,88.
YANG Xiaoli, YE Feng, SONG Hailiang, *et al.* Enhanced biological nitrogen removal at low-temperature based on operation data of wastewater treatment plants [J]. China Water & Wastewater, 2009, 25(1): 82-85, 88(in Chinese).

作者简介:姜阳(1995-),男,四川成都人,硕士研究生,研究方向为污水处理过程优化与控制、物联网与大数据分析。

E-mail:lizhihuagroup@xauat.edu.cn

收稿日期:2020-12-24

修回日期:2021-01-13

(编辑:丁彩娟)