

目次

加速能源转型与产业结构调整的环境健康协同效益评估:以京津冀鲁地区为例	杨玺, 孙奕生, 常世彦, 李胜悦, 郑昊天, 王书肖, 张希良 (3627)
基于不同电力需求的中国减排降碳协同增效路径	向梦宇, 王深, 吕连宏, 张楠, 白梓函 (3637)
城市碳达峰碳中和行动评估方法与应用	张保留, 白梓函, 张楠, 吕连宏, 阳平坚 (3649)
基于可解释性机器学习的城市 O ₃ 驱动因素挖掘	董佳奇, 胡冬梅, 闫雨龙, 彭林, 张辉辉, 牛月圆, 段小琳 (3660)
基于响应曲面法的臭氧生成敏感性分析	朱禹寰, 陈冰, 张雅茹, 刘晓, 李光耀, 舍静, 陈强 (3669)
基于 WRF-CMAQ/ISAM 模型的榆林市夏季 O ₃ 来源解析	王伊凡, 仝纪龙, 陈羽翔, 林鑫, 刘永乐, 敖丛杰, 刘浩天 (3676)
基于轻量级梯度提升机的南京大气臭氧浓度预测	朱珈莹, 安俊琳, 冯悦政, 贺婕, 张玉欣, 王俊秀 (3685)
2022 年广东省冬季一次臭氧污染过程的气象成因及潜在源区分析	李婷苑, 陈靖扬, 龚宇, 沈劲 (3695)
夏季对流层臭氧辐射强迫对华北地区天气和空气质量的影响	杜楠, 陈磊, 廖宏, 朱佳, 李柯 (3705)
气象、本地光化学生成和外围传输对长沙市 2018~2020 年臭氧污染趋势变化影响的识别	杨俊, 杨雷峰, 丁华, 谢丹平, 刘妍妍, 余涛, 吕明, 袁自冰 (3715)
2000~2021 年成渝城市群 PM _{2.5} 时空变化及驱动机制多维探测	徐勇, 郭振东, 郑志威, 戴强玉, 赵纯, 黄雯婷 (3724)
基于 GTWR-XGBoost 模型的四川省 PM _{2.5} 小时浓度估算	吴迪, 杜宁, 王莉, 吴宇宏, 张少磊, 周彬, 张显云 (3738)
2013~2021 年邢台 PM _{2.5} 重污染过程输送和增长特性	江琪, 盛黎, 靳雨晨, 王继康, 尤媛, 王飞 (3749)
基于高分辨率在线观测数据分析上海市城区秋冬季大气有机气溶胶化学特征及污染来源	朱书慧 (3760)
机动车尾气碳质气溶胶排放因子及其稳定碳同位素特征	于鸣媛, 王谦, 付明亮, 戈畅, 谢峰, 曹芳, 章炎麟 (3771)
沈阳工业区夏季 VOCs 组成特征及其对二次污染形成的贡献	关璐, 苏枫枫, 库盈盈, 胡建林, 于兴娜 (3779)
疫情管控期间深圳市城区 VOCs 的变化特征及减排效果评估	云龙, 林楚雄, 李成柳, 邱志诚, 古添发, 李光程, 张明棣, 郭健锋 (3788)
西安市冬季道路扬尘中有机质组成特征及其氧化潜势	王擎雯, 陈庆彩, 王超, 王瑞鹤, 刘国瑞, 李豪, 李艳广 (3797)
淮北孙疃矿区地表尘中多环芳烃类化合物的污染特征及致癌风险评价	徐振鹏, 钱雅慧, 洪秀萍, 罗钟庚, 高秀龙, 梁汉东 (3809)
雄安建设初期白洋淀水质时空差异及其对降雨和土地利用的响应	王子铭, 杨丽虎, 宋献方 (3820)
流域尺度污染溯源模拟-优化防控方法:以铜陵市顺安河流域为例	刘国王辰, 陈磊, 李佳奇, 张钰晨, 赵奕欣, 刘妍琪, 沈珍瑶 (3835)
南昌市浅层地下水水质评价及监测指标优化	郑紫吟, 储小东, 徐金英, 马志飞 (3846)
闪电河流域“三水”氢氧同位素特征及水体转换分析	杨丽娜, 贾德彬, 高瑞忠, 苏文旭, 卢方园, 郝玉胜 (3855)
金沙江丰富类和稀有类浮游真核微生物的分布特征与影响因素	燕炳成, 崔戈, 孙胜浩, 王沛芳, 王超, 吴程, 陈娟 (3864)
嘉陵江不同干扰断面河道沉积物细菌群落多样性	竹兰萍, 张拓, 李佳宁, 王佳颖, 郭伟亮, 徐飞, 张富斌 (3872)
水库水深变化对不同浮游微生物群落及网络互作关键种的影响	王洵, 廖琴, 王沛芳, 袁秋生, 胡斌, 邢小蕾, 徐浩森 (3881)
典型工业区地下水细菌群落多样性特征与环境因子响应初探	吴建强, 张书源, 王敏, 陈敏, 叶文娟, 熊丽君, 黄沈发 (3892)
太湖流域肥料施用策略调整对典型作物系统氮磷流失的影响	俞映倬, 王逸之, 杨根, 杨林章, 段婧婧, 韩雪梅, 薛利红 (3902)
基于机器学习的长江流域农田氮径流流失负荷估算	张有福, 潘哲祺, 陈丁江 (3913)
千岛湖地区上梧溪流域地表径流非点源氮污染分类识别	俞珂, 严璇, 唐张轩, 张方方, 何圣嘉, 姜培坤 (3923)
沱江和涪江水系干支流氮磷营养盐的空间分布特征	李子阳, 周明华, 徐鹏, 陈露, 刘祥龙, 林洪羽, 江南, 任兵, 张博文 (3933)
外秦淮河底泥污染及疏浚效果	张沐, 任增谊, 张曼, 赵琼, 尹洪斌 (3945)
新污染物多环芳烃衍生物的来源、分布与光化学行为	葛林科, 王子宇, 曹胜凯, 车晓佳, 朱超, 张蓬, 马宏瑞 (3957)
镉砷在针铁矿界面共吸附的相互作用机制	苏子贤, 刘赛红, 管玉峰, 陶亮 (3970)
环糊精键合 Fe-TAML 催化剂的制备及其活化 H ₂ O ₂ 氧化水中有机微污染物	刘清泉, 蔡本哲, 蔡喜运 (3978)
Co(II) 活化过一硫酸盐降解氨基三甲膦酸的性能及反应机制	朱敬林, 汪舒 (3990)
微塑料诱导下污泥造粒潜能变化及微生物富集特征	谢晴帆, 俞楠, 张妮, 谢周云, 单珂欣, 吴亦馨, 唐力, 夏静芬, 杨国靖 (3997)
中国农田土壤 Cd 累积分布特征及概率风险评价	王静, 魏恒, 潘波 (4006)
渝东南典型地质高背景区土壤重金属来源解析及污染评价	蒋玉莲, 余京, 王锐, 王佳彬, 李瑜, 余飞, 张云逸 (4017)
基于 PMF 模型的农田土壤重金属源暴露风险综合评价:以浙江省某电子垃圾拆解区为例	方嘉, 何影, 黄乃涛, 支裕优, 傅伟军 (4027)
乌鲁木齐市郊农田土壤及农作物中多环芳烃的污染特征及风险评价	范悦, 曹双瑜, 艾力江·努尔拉, 于芸云江·吗米提敏, 阿不都艾尼·阿不里, 谢宣宣, 古丽斯坦·阿不都拉, 刘河疆 (4039)
荒漠绿洲土壤抗生素抗性基因分布特征及驱动机制	黄福义, 周曙仡, 苏建强, 朱永官 (4052)
近 15 年我国土壤抗生素污染特征与生物修复研究进展	赵晓东, 乔青青, 秦宵睿, 李晓晶, 李永涛 (4059)
改性生物炭修复砷镉复合污染土壤研究进展	吕鹏, 李莲芳, 黄晓雅 (4077)
沼液还田对土壤-作物系统重金属累积的影响:Meta 分析	赵奇志, 杨志敏, 孔凡靖, 熊海灵, 朱康文, 陈玉成 (4091)
两种铁改性生物炭对微碱性砷镉污染土壤的修复效果	梁欣冉, 何丹, 郑墨华, 付庆灵, 胡红青, 朱俊 (4100)
秸秆离田对土壤 Cd 生物有效性及水稻 Cd 积累的影响	王子钰, 周航, 周坤华, 谭文韬, 蒋毅, 唐棋, 伍港繁, 辜娇峰, 曾鹏, 廖柏寒 (4109)
秸秆与脱硫石膏配施改良黄河三角洲盐碱地的理化性质	赵惠丽, 于金艺, 刘涛, 王丽, 赵英 (4119)
基于环境 DNA 的复合污染土壤生物评价和胁迫诊断	黄湘云, 钟文军, 刘训杰, 毕婉娟, 钱林皓, 张效伟 (4130)
地质高背景区土壤-玉米重金属综合质量评价	张传华, 王钟书, 刘力, 刘燕 (4142)
浙江省蔬菜生产系统重金属污染生态健康风险	张述敏, 刘翠玲, 杨桂玲, 邓美华 (4151)
生物炭配施磷肥对土壤养分、酶活性及紫花苜蓿养分吸收的影响	刘鑫裕, 王冬梅, 张泽洲, 张鹏, 樊桐桐 (4162)
低密度聚乙烯微塑料对空心菜生长和生理特征的影响	周颖, 蒋文婷, 刘训悦, 朱高荻, 唐荣贵, 章海波, 蔡延江 (4170)
沼液施用对麦茬口期土壤微生物群落结构特征及功能的影响	乔宇颖, 奚辉, 李娜, 陈喜靖, 沈阿林, 喻曼 (4179)
大同铅锌尾矿不同污染程度土壤细菌群落分析及生态功能特征	刘泽勋, 庄家尧, 刘超, 郑康, 陈玲 (4191)
马尾松采伐迹地火烧黑炭对土壤有机碳组分和碳转化酶活性的影响	姚智, 焦鹏宇, 吴晓生, 严强, 刘先, 胡亚林, 王玉哲 (4201)
《环境科学》征订启事(3648) 《环境科学》征稿简则(3659) 信息(3714, 4150, 4178)	

基于 GTWR-XGBoost 模型的四川省 PM_{2.5} 小时浓度估算

吴迪, 杜宁*, 王莉, 吴宇宏, 张少磊, 周彬, 张显云

(贵州大学矿业学院, 贵阳 550025)

摘要: 卫星气溶胶光学厚度(AOD)和气象数据已被广泛用于估算空气动力学直径 $\leq 2.5 \mu\text{m}$ 的地表颗粒物(PM_{2.5})浓度. 研究高时间分辨率、高精度的 PM_{2.5} 浓度估算方法, 对及时准确的空气质量预报和大气污染的预防及缓解具有重要意义. 使用 Himawari-8 AOD 小时产品和 ERA5 气象再分析资料作为估算变量, 提出 GTWR-XGBoost 组合模型, 对四川省 PM_{2.5} 小时浓度进行估算. 结果表明: ①提出的组合模型运用于全数据集的性能优于 KNN、RF、AdaBoost、GTWR、GTWR-KNN、GTWR-RF 和 GTWR-AdaBoost 模型, 拟合精度指标 R^2 、MAE 和 RMSE 分别为 0.96、3.43 $\mu\text{g}\cdot\text{m}^{-3}$ 和 5.52 $\mu\text{g}\cdot\text{m}^{-3}$, 验证精度指标 R^2 、MAE 和 RMSE 分别为 0.9、4.98 $\mu\text{g}\cdot\text{m}^{-3}$ 和 7.92 $\mu\text{g}\cdot\text{m}^{-3}$. ②该模型作用于 PM_{2.5} 浓度小时估算上, 具有较高的拟合优度(全数据集的 R^2 为 0.96, 不同时刻的 R^2 介于 0.91 ~ 0.98 之间), 表明该模型对于小时估算具有较好的时间稳定性, 能为区域空气质量评估提供精准的估算信息. ③在时间上年均 PM_{2.5} 小时浓度估算变化总体上呈现“倒 U”型趋势, 09:00 开始升高至 11:00 达到峰值 [$\rho(\text{PM}_{2.5})$ 为 44.56 $\mu\text{g}\cdot\text{m}^{-3}$] 后逐渐降低; 且季节性变化非常明显, 呈现冬季 > 春季 > 秋季 > 夏季的趋势. ④在空间分布上总体呈现东高西低和局部污染程度较高的特点, 高值区主要分布在城市快速发展和人口密集的东部地区, 低值区主要分布在经济发展落后和人口稀疏的西部地区.

关键词: PM_{2.5}; Himawari-8 AOD; GTWR-XGBoost 模型; 机器学习; 时空分布

中图分类号: X513 文献标识码: A 文章编号: 0250-3301(2023)07-3738-11 DOI: 10.13227/j.hjxx.202207179

Estimation of PM_{2.5} Hourly Concentration in Sichuan Province Based on GTWR-XGBoost Model

WU Di, DU Ning*, WANG Li, WU Yu-hong, ZHANG Shao-lei, ZHOU Bin, ZHANG Xian-yun

(Mining College, Guizhou University, Guiyang 550025, China)

Abstract: Aerosol optical depths of satellites and meteorological factors have been widely used to estimate concentrations of surface particulate matter with an aerodynamic diameter $\leq 2.5 \mu\text{m}$. Research on a high time resolution and high-precision PM_{2.5} concentration estimation method is of great significance for timely and accurate air quality prediction and air pollution prevention and mitigation. Himawari-8 AOD hour product and ERA5 meteorological reanalysis data were used as estimation variables, and a GTWR-XGBoost combined model was proposed to estimate hourly PM_{2.5} concentration in Sichuan Province. The results showed that: ① the performance of the proposed combination model was better than that of the KNN, RF, AdaBoost, GTWR, GTWR-KNN, GTWR-RF, and GTWR-AdaBoost models in the full dataset; the fitting accuracy indexes R^2 , MAE, and RMSE were 0.96, 3.43 $\mu\text{g}\cdot\text{m}^{-3}$, and 5.52 $\mu\text{g}\cdot\text{m}^{-3}$, respectively; and the verification accuracy indexes R^2 , MAE, and RMSE were 0.9, 4.98 $\mu\text{g}\cdot\text{m}^{-3}$, and 7.92 $\mu\text{g}\cdot\text{m}^{-3}$, respectively. ② The model had a high goodness of fit (R^2 of the whole dataset was 0.96, and R^2 of different times ranged from 0.91 to 0.98) when applied to the estimation of PM_{2.5} concentration hour. It showed that the model had good time stability for hourly estimation and could provide accurate estimation information for regional air quality assessment. ③ In terms of time, the annual average PM_{2.5} hourly concentration estimation showed an inverted U-shaped trend. It began to increase gradually at 09:00 am to a peak of 44.56 $\mu\text{g}\cdot\text{m}^{-3}$ at 11:00 and then gradually decreased. Moreover, the seasonal variation was very obvious, with winter > spring > autumn > summer. ④ In terms of spatial distribution, it showed the characteristics of high in the east and low in the west and a high degree of local pollution.

Key words: PM_{2.5}; Himawari-8 AOD; GTWR-XGBoost model; machine learning; spatial and temporal distributions

随着经济全球化的潮流,我国城市化率快速提升,而大量聚集的工业生产和人类生活排放引起 PM_{2.5} 等空气污染问题日益突出^[1]. PM_{2.5} 是指空气动力学直径不超过 $2.5 \mu\text{m}$ 的颗粒物,对全球气候、大气环境质量和人体健康等有着重要影响^[2]. 近地表空气中的 PM_{2.5} 可能被吸入人的肺部,从而导致人体呼吸系统,心血管系统,生殖系统或血液系统出现严重的健康问题^[3]. 目前对 PM_{2.5} 的监测主要是通过地面监测和遥感监测. 地面监测虽然可以获取实时的高精度数据,但受限于地面站点数量和监测覆盖范围,缺乏空间上的延续性,从而难以提供丰富的

监测结果^[4]. 相较于地面监测,遥感监测具有空间分辨率高、监测范围广、速度快、成本低和全天候实时监测的特点,可以弥补地面监测的不足^[5-7].

遥感监测 PM_{2.5} 主要是以 AOD 和气象数据为自变量,建立回归模型估算 PM_{2.5} 浓度^[8-10]. 但 PM_{2.5} 的化学成分和光学特性是不均匀的,具有显著的非线性特征,简单的线性模型难以正确估算其浓度. 机

收稿日期: 2022-07-17; 修订日期: 2022-09-12

基金项目: 贵州省省级科技计划项目(黔科合支撑[2022]一般 204)

作者简介: 吴迪(1996~),男,硕士研究生,主要研究方向为资源与环境遥感, E-mail: 1290225517@qq.com

* 通信作者, E-mail: ndul@gu.zu.edu.cn

器学习算法(K 最邻近值、随机森林和 boost 算法等)计算速度快,对异常值和噪声容忍度高,适用于解决非线性问题. 已有许多学者尝试使用机器学习模型应用于空气质量估算,克服了空气质量与气象变量之间非线性关系的困难^[11,12]. Li 等^[13]在 GWR 中融合 AOD 数据和 DMSP/OLS 夜间灯光影像数据,对美国东北部地区的 PM_{2.5} 浓度进行估算. He 等^[14]结合空间分辨率为 3 km 的 AOD 数据,运用改进的地理时间加权回归模型(GTWR)估算中国大陆地区的 PM_{2.5} 日均浓度. 侯俊雄等^[15]提出了一种融合了随机森林算法和 WRF 大气模式的 PM_{2.5} 浓度实时预报方法. Pan^[16]采用 XGBoost 模型对小时 PM_{2.5} 浓度进行估算,获得较好结果. Zhai 等^[17]提出一种 AdaBoost、XGBoost 和遗传算法(GA-MLP)集成的堆叠模型,对北京 PM_{2.5} 日均值进行了估算,结果表明集合模型的异变解释能力优于单个模型. Chen 等^[18]利用 Himawari-8 号卫星数据,结合 AdaBoost、XGBoost 和随机森林(RF)这 3 个机器学习的子模型建立一种堆叠模型,其性能优于单一模型. 以上研究存在以下问题:①虽均以连续的 AOD、PM_{2.5} 数据做估算,但时间分辨率较粗;②研究区域的选择大多是东部沿海地区,对内陆地区的研究较少;③以上这些研究使用的模型考虑到了数据源的非线性问题,但忽略了时空异质性问题. 因此对作用于内陆地区、达到小时估算分辨率且同时考虑时空异质性和非线性的组合模型研究,有助于大气污染联防联控与公众环境空间质量信息服务的改善.

本研究将 2017 年 Himawari-8 AOD 小时产品、地面监测站点 PM_{2.5} 小时质量浓度和 ERA5 气象数据结合,在分析了气象变量对 PM_{2.5} 浓度影响的基础上,提出了 GTWR 和 XGBoost 结合的 PM_{2.5} 估算模型,对四川省 PM_{2.5} 小时浓度进行估算,以期对区域空气质量评估提供精准参考.

1 材料与方法

1.1 研究区域概况

四川省位于中国西南腹地,属于中国大陆地势第一阶梯和第二阶梯,介于 97°21′~108°33′E 和 26°03′~34°19′N 之间,地势复杂,东西高低悬殊,域内面积 48.6 万 km². 以龙门山-大凉山一线为界,东部为四川盆地及盆缘山地,西部为川西高山高原及川西南山地,域内最高点位于贡嘎山,海拔 7 556 m,最低点位于广安市邻水县的御临河出省处,海拔仅 184 m,高差约 7 300 m,见图 1. 因四川盆地独特的地理位置,冷空气易沿盆周向底部运动,在盆地底部聚集,形成典型的地形逆温现象,又因人口集中于川东

盆地地区,人类生产生活密集,加之盆地风向受反气旋控制明显,全年静风较多,不利于污染物扩散,造成污染物积聚,使得四川省中东部成为空气污染较为严重的地区.

1.2 数据来源及处理

1.2.1 PM_{2.5} 站点数据

PM_{2.5} 数据为中国环境监测总站(CNEMC, <http://106.37.208.233:20035/>) 2017 年空气质量监测站逐小时的污染物监测数据,研究区范围内共包括 94 个站点. PM_{2.5} 研究区域及地面监测站分布见图 1.

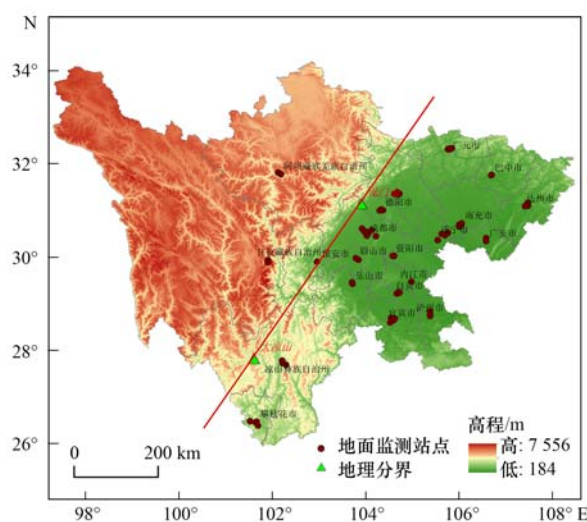


图 1 研究区域及地面监测站分布示意

Fig. 1 Research area and distribution of ground monitoring stations

1.2.2 Himawari-8 AOD 数据

Himawari-8 卫星是 2015 年 7 月 7 日正式运行的新一代静止气象卫星,该卫星具有最高 500 m 的空间分辨率及全盘扫描每次 10 min 的时间分辨率. Zang 等^[19]的研究将 Himawari-8 AOD 产品与中国及其周边多地区的 42 个 AERONET 站点的地面实际观测结果进行分析,结论得出两者之间具有较高相关性;Chen 等^[18]的研究利用 Himawari-8 卫星数据结合模型对中国中东部地区进行了逐时性的 PM_{2.5} 浓度估算,与地面站点实测小时浓度验证得到较高的拟合精度. 综合前者研究表明 Himawari-8 AOD 小时产品可用于对地面 PM_{2.5} 浓度进行小时估算. 本文使用 Himawari-8 AOD 产品中置信度为“非常好”的 Level-3 小时产品,通过日本宇航局的 P-Tree 系统(<http://www.eorc.jaxa.jp/ptree/>)下载.

1.2.3 气象数据

气象数据采用欧洲中期天气预报中心(European centre for medium-range weather forecasts, ECMWF)发布的 ERA5 再分析数据集(<https://www.ecmwf.int/en/forecasts/datasets/>), ERA5 是继前身 ERA-

Interim 之后 ECMWF 对全球气候的第五代大气再分析,相较于 ERA-Interim, ERA5 的空间分辨率达到了 0.125° ,时间分辨率达到了 1 h,同化了更多的观测变量.本研究选用四川省地区 2017 年 ERA5 气象数据,气象变量包括边界层高度(BLH)、低植被覆盖指数(CVL)、高植被覆盖指数(CVH)、气压(SP)、2 m 地面温度(T2M)、2 m 露点温度(D2M)、10 m 风速 u 分量(U_{10})、10 m 风速 v 分量(V_{10})、总降水量(RAIN)和相对湿度(RH),下文中均用缩写表示.

1.2.4 数据匹配

由于卫星过境时并非准点时刻,获得的 Himawari-8 AOD 数据与 ERA5 再分析数据存在时间与空间上的差异,所以要对两种数据进行时空匹配处理.时间上,以 Himawari-8 AOD 数据为基准,取卫星过境时间前后 1 h 的 $PM_{2.5}$ 站点数据均值与

Himawari-8 AOD 数据匹配,得到 AOD 过境时的 $PM_{2.5}$ 浓度,取与 Himawari-8 卫星过境时刻最邻近的 ERA5 数据作为 AOD 过境时的气象数据;空间上,采用改进的重采样方法对 ERA5 数据做降尺度处理,将 ERA5 的空间分辨率调整到与 Himawari-8 AOD 一致^[20].

1.3 研究方法

1.3.1 $PM_{2.5}$ 浓度与气象因素相关性分析

本文将气象因子作为 $PM_{2.5}$ 浓度预测的重要影像因素,已有研究证实了 $PM_{2.5}$ 浓度与边界层高度、气压、温度、湿度和风速等因素之间具有密切关系^[21~23].对研究区内 $PM_{2.5}$ 浓度与气象因素进行皮尔逊相关性分析,分析结果见表 1,其中所示 $PM_{2.5}$ 浓度与 SP 和 RH 呈正相关关系,与 BLH、CVH、T2M、D2M、 U_{10} 、 V_{10} 和 RAIN 呈负相关关系.

表 1 $PM_{2.5}$ 与气象因素的皮尔逊相关系数

Table 1 Pearson correlation coefficient of $PM_{2.5}$ concentration and meteorological factors

项目	BLH	CVH	SP	T2M	D2M	U_{10}	V_{10}	RH	RAIN
相关系数	-0.210	-0.094	0.077	-0.413	-0.290	-0.071	-0.142	0.090	-0.152

1.3.2 多重线性回归分析共线性

本研究收集的气象变量较多,估算变量之间的线性关系会隐蔽变量的显著性,增加参数估计的误差,影响模型预测的性能,因此需诊断各变量之间是否存在线性关系.用多重线性回归分析计算标准化系数和 VIF(方差膨胀因子)系数作为共线性诊断的重要指标.当 $VIF \geq 10$ 说明变量间存在多重共线性问题, $5 \leq VIF < 10$ 说明变量间可能存在多重共线性关系, $VIF < 5$ 说明变量间不存在多重共线性关系^[24];标准化系数越接近 1(或 -1)对于因变量的正向影响(或负向影响)越显著.多重线性回归分析结果见表 2,表中 RAIN($VIF = 48.056$)、 V_{10} ($VIF = 36.901$)、D2M($VIF = 24.385$)和 T2M($VIF = 17.149$)这 4 个变量 VIF 值超过 10 应该剔除,结合标准化系数 T2M(-0.039)、RH(0.047),经多次回归分析,当剔除 T2M、 V_{10} 和 RH 时各变量的 VIF 值均小于 2,结果最优.故将剔除 T2M、 V_{10} 和 RH 变量的数据集输入 GTWR 模型训练;因 XGBoost 模型能分布式处理高维稀疏特征,无需考虑变量之间的共线性关系,可将 T2M、 V_{10} 和 RH 变量加入 XGBoost 拟合.

1.3.3 地理时间加权回归模型(GTWR)

Brunsdon 等^[25]在总结前人提出的局部回归分析和变参数研究的基础上提出了地理加权回归方法(GWR),用于检验自变量和因变量之间的空间差异关系. Huang 等^[26]在 GWR 的基础上引入时间纬度,

表 2 多重线性回归分析结果¹⁾

Table 2 Analysis results of multiple linear regression

项目	标准化系数	VIF	剔除变量后 VIF
AOD	0.493	1.167	1.065
BLH	-0.272	3.826	1.336
CVH	-0.043	1.061	1.042
SP	0.131	1.511	1.417
T2M	-0.039	17.149	—
U_{10}	0.101	8.016	1.560
V_{10}	0.259	36.901	—
D2M	-0.417	24.385	1.374
RH	0.047	7.813	—
RAIN	-0.252	48.056	1.540

1) 因变量为 $PM_{2.5}$; “—”表示该变量剔除

拓展为地理时间加权回归模型(GTWR),用于检验自变量和因变量之间的时空异质性关系.之后,众多学者对 GTWR 的时空异质性进行研究,验证了 GTWR 能够有效地解决空间数据的时空非平稳性问题^[27,28].GTWR 的基本公式见式(1):

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_k \beta_k(u_i, v_i, t_i)x_{ik} + \varepsilon_i \quad (1)$$

式中, i 为空间点的索引; y_i 为样本点的因变量值; (u_i, v_i, t_i) 为时空点的坐标; $\beta_0(u_i, v_i, t_i)$ 为样本点的回归常数; x_{ik} 为样本点的第 k 个解释变量.此时的关键问题是估计自变量 k 在每个时空位置的回归系数 $\beta_k(u_i, v_i, t_i)$,采用矩阵的方式, $\beta_k(u_i, v_i, t_i)$ 的估计可以用式(2)表示:

$$\beta(u_i, v_i, t_i) = [X^T W(u_i, v_i, t_i) X]^{-1} X^T W(u_i, v_i, t_i) y \quad (2)$$

式中, $W(u_i, v_i, t_i) = \text{diag}(\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in})$ 为时空权重矩阵; n 为观测值总数; 对角元素 $\alpha_{ij} (1 \leq j \leq n)$ 为观测点 i 附近回归时空位置 (u, v, t) 相对于该点的时空距离函数^[29].

1.3.4 极端梯度提升模型(XGBoost)

XGBoost(extreme gradient boosting)算法由 Chen 等^[30] 提出, 是一种基于 GBDT (gradient boosting decision tree) 的改进 boost 模型, 是用多个弱分类器组合成一个强分类器(图 2), 广泛应用于股票金融、房价、顾客行为和企业经济预测等方向. 具有高精度、不易过拟合和可扩展性等特点, 能分布式处理高维稀疏特征^[31] 及非线性映射能力. XGBoost 集成了多棵分类回归树(CART)以弥补单棵 CART 无法满足预测精度的不足, 预测结果等于所有 CART 的得分总和^[32]. 模型表达公式见式(3):

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (3)$$

式中, i 为样本点数量; $\hat{y}_i$ 为样本预测值; k 为回归树数量; x_i 为数据的特征向量; F 为 CART 树的集合空间.

XGBoost 通过对代价函数进行二阶泰勒展开, 使用一阶和二阶导数, 在训练集上可以更快收敛, 有效提高训练速度, 并且将正则化项加到损失函数上, 可以降低模型的复杂度和过拟合的风险^[33].

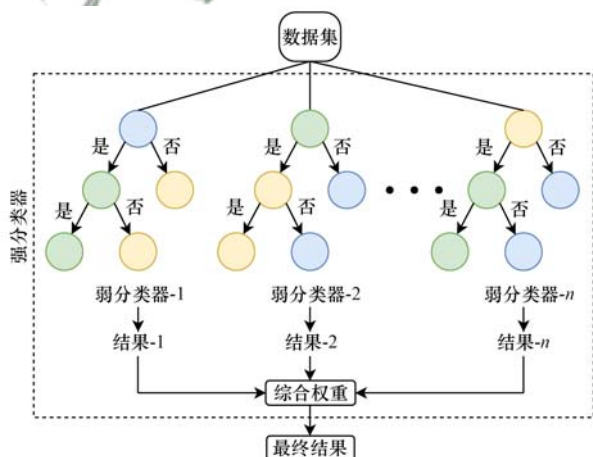


图2 XGBoost 示意

Fig. 2 Schematic of XGBoost

输入 XGBoost 模型的变量包含 GTWR 模型的估算值、T2M、RH、 V_{10} 和时空信息. 考虑到 PM_{2.5} 在大气中的传输过程及其具有时间自相关的特性, PM_{2.5} 监测站点附近 n 个站点和同站点前 m 天的 PM_{2.5} 观测值对于估算 PM_{2.5} 具有参考意义. 分别以 S-PM_{2.5} 和 T-PM_{2.5} 表示空间信息项及时间信息

项^[20], 其中时空信息项表示为:

$$S\text{-PM}_{2.5} = \frac{\sum_{p=1}^n w_{s_p} \text{PM}_{2.5,p}}{\sum_{p=1}^n w_{s_p}}, \quad w_{s_p} = \frac{1}{ds_p^2} \quad (4)$$

$$T\text{-PM}_{2.5} = \frac{\sum_{q=1}^m w_{t_q} \text{PM}_{2.5,q}}{\sum_{q=1}^m w_{t_q}}, \quad w_{t_q} = \frac{1}{dt_q^2} \quad (5)$$

式中, m 和 n 均为 3; $\text{PM}_{2.5,p}$ 为在空间上靠近该站点的第 p 个站点的 PM_{2.5} 观测值; $\text{PM}_{2.5,q}$ 为同一站点之前第 q 天的 PM_{2.5} 观测值; w_{s_p} 和 w_{t_q} 分别为空间上靠近该站点的第 p 个站点的距离加权系数和同一站点之前第 q 天的时间加权系数; ds_p 和 dt_q 分别为空间上靠近该站点的第 p 个站点与该站点的空间距离和同一站点之前第 q 天与该站点观测时刻的时间间隔.

1.3.5 模型构建

GTWR-XGBoost 组合预测模型构建流程见图 3, 包括数据预处理、变量分析和模型构建.

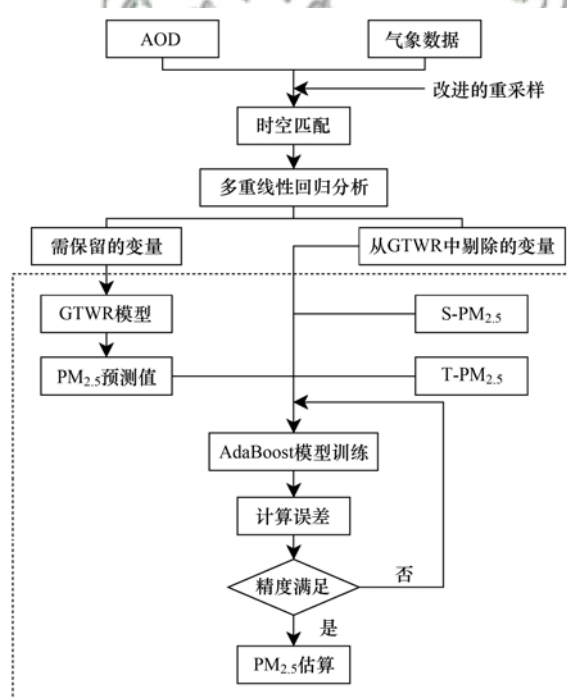


图3 GTWR-XGBoost 模型预测流程

Fig. 3 GTWR-XGBoost model prediction process

(1)数据预处理 主要包括数据清洗、时空匹配和归一化处理.

(2)变量分析 计算所有估算变量的标准化系数和 VIF 值, 经过逐步多重线性回归分析, 得到应用于模型所需保留的变量.

(3)模型构建 将保留的气象估算变量输入 GTWR 模型预测, 从 GTWR 中剔除的变量同 S-

PM_{2.5}、T-PM_{2.5} 加入 XGBoost 模型拟合,再使用 GTWR-XGBoost 组合模型估算 2017 年四川省 PM_{2.5} 浓度分布.

1.3.6 模型验证

本文使用决定系数(R^2)、平均绝对误差(MAE)和均方根误差(RMSE)作为评价模型估算精度的指标,对比 KNN、RF、AdaBoost、XGBoost、GTWR、GTWR-KNN、GTWR-RF、GTWR-AdaBoost 和 GTWR-XGBoost 模型的拟合效果,并采用十折交叉验证方法^[34]对各模型进行验证,各指标计算见式(6):

$$\left. \begin{aligned} R^2 &= 1 - \frac{\sum_{i=1}^N [PM_{2.5}^{obs}(i) - PM_{2.5}^{pre}(i)]^2}{\sum_{i=1}^N [PM_{2.5}^{obs}(i) - \overline{PM_{2.5}^{obs}}]^2}, \\ RMSE &= \sqrt{\frac{1}{N} \sum_{i=1}^N [PM_{2.5}^{obs}(i) - PM_{2.5}^{pre}(i)]^2}, \\ MAE &= \frac{1}{N} |PM_{2.5}^{obs}(i) - PM_{2.5}^{pre}(i)|. \end{aligned} \right\} \quad (6)$$

式中, N 为样本总数; $PM_{2.5}^{obs}$ 为实测 PM_{2.5} 浓度; $PM_{2.5}^{pre}$ 为估算 PM_{2.5} 浓度; $\overline{PM_{2.5}^{obs}}$ 为实测 PM_{2.5} 浓度的平均值.

2 结果与讨论

2.1 模型拟合及交叉验证

使用全数据集进行训练考察模型的拟合效果,通过十折交叉验证考察模型的估算与泛化能力.

2.1.1 模型拟合结果对比与分析

为了验证组合模型 GTWR-XGBoost 的精度,本研究选取单一 KNN、RF、AdaBoost、XGBoost 和 GTWR 以及相应组合模型 GTWR-KNN、GTWR-RF 和 GTWR-AdaBoost 作为对比实验.图 4 和表 3 为不同模型拟合值和实测值的回归分析图及评价指标统计表.

结合图 4 和表 3 可知,单一模型中,KNN(R^2 为 0.72)拟合能力较差,RF、AdaBoost、XGBoost 和 GTWR 模型 R^2 分别为 0.81、0.83、0.84 和 0.93;在将单一拟合能力较好且具有解释时空异质性能力的 GTWR 模型与其余单一模型结合后,组合模型拟合性能较原单一模型均有显著提升,GTWR-KNN 模型 R^2 、MAE 和 RMSE 分别达到了 0.91、7.44 $\mu\text{g}\cdot\text{m}^{-3}$ 和 10.21 $\mu\text{g}\cdot\text{m}^{-3}$,GTWR-RF 模型 R^2 、MAE 和 RMSE 分别为 0.94、4.35 $\mu\text{g}\cdot\text{m}^{-3}$ 和 6.75

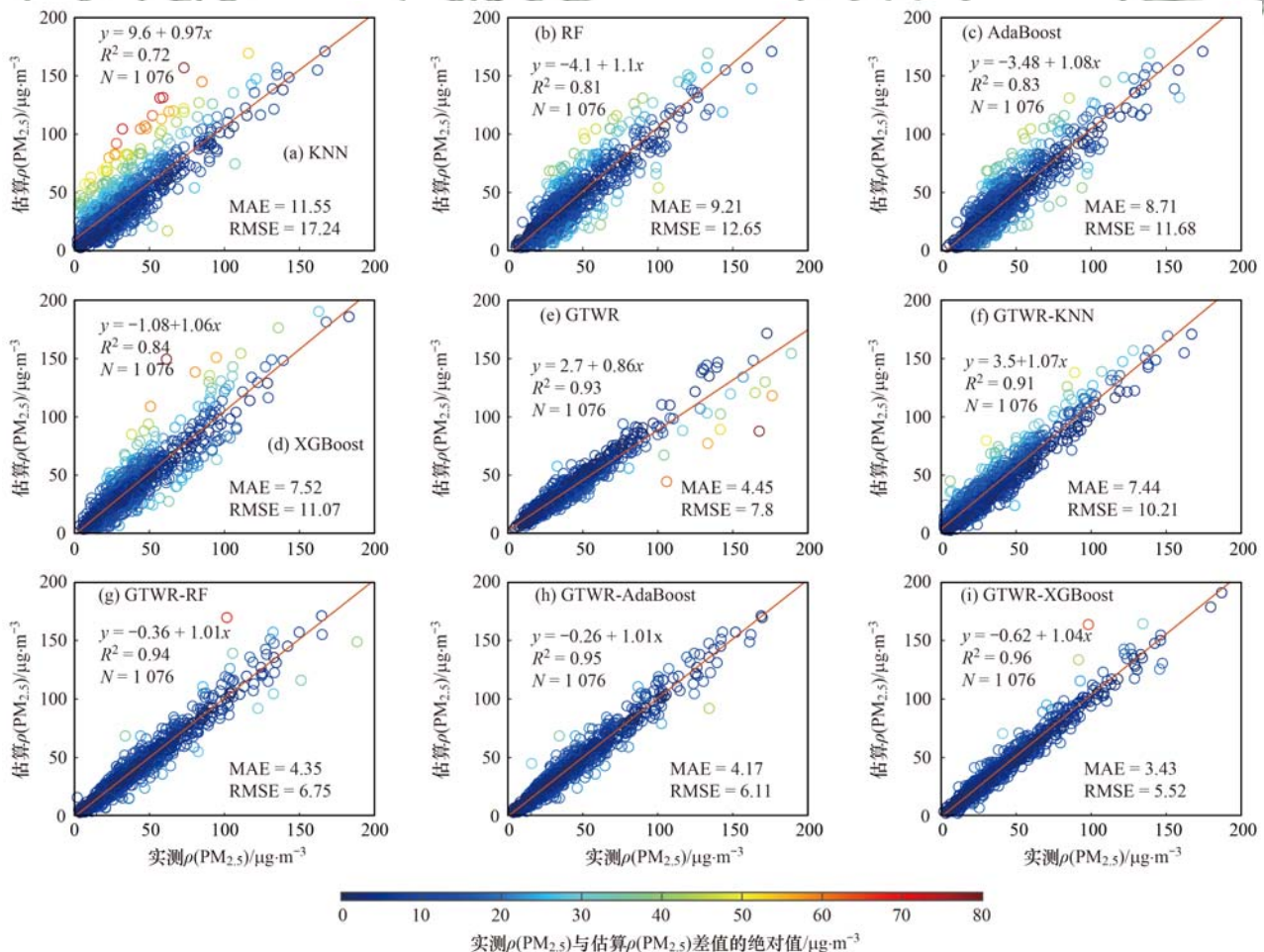


图 4 模型拟合结果对比

Fig. 4 Comparison of model fitting results

$\mu\text{g}\cdot\text{m}^{-3}$, GTWR-AdaBoost 模型 R^2 、MAE 和 RMSE 分别为 0.95、4.17 $\mu\text{g}\cdot\text{m}^{-3}$ 和 6.11 $\mu\text{g}\cdot\text{m}^{-3}$, GTWR-XGBoost 模型 R^2 、MAE 和 RMSE 分别达到 0.96、3.43 $\mu\text{g}\cdot\text{m}^{-3}$ 和 5.52 $\mu\text{g}\cdot\text{m}^{-3}$. 总体说明, GTWR-XGBoost 模型对于 PM_{2.5} 浓度的变异解释能力最强, 估算 PM_{2.5} 浓度与实测 PM_{2.5} 浓度符合程度最高, 误差最小, 具有最佳的拟合能力.

表 3 模型拟合结果评价指标统计

模型	模型拟合		
	R^2	MAE/ $\mu\text{g}\cdot\text{m}^{-3}$	RMSE/ $\mu\text{g}\cdot\text{m}^{-3}$
KNN	0.72	11.55	17.24
RF	0.81	9.21	12.65
AdaBoost	0.83	8.71	11.68
XGBoost	0.84	7.52	11.07
GTWR	0.93	4.45	7.8
GTWR-KNN	0.91	7.44	10.21
GTWR-RF	0.94	4.35	6.75
GTWR-AdaBoost	0.95	4.17	6.11
GTWR-XGBoost	0.96	3.43	5.52

2.1.2 模型验证结果对比与分析
使用十折交叉验证方法用于 9 种模型的验证.

图 5 和表 4 为不同模型交叉验证值和实测值的回归分析图及评价指标统计表. 结果表明, 这 9 种模型的交叉验证 R^2 介于 0.41 ~ 0.9 之间, MAE 范围在 4.98 ~ 12.78 $\mu\text{g}\cdot\text{m}^{-3}$ 之间, RMSE 范围在 7.92 ~ 21.44 $\mu\text{g}\cdot\text{m}^{-3}$ 之间.

结合图 5 和表 4 可知, 单一模型验证结果的精度均下降明显, 其中 KNN 模型 (R^2 为 0.41) 的验证结果较差, 下降最为显著, RF 模型 R^2 为 0.65, AdaBoost 模型 R^2 为 0.69, XGBoost 模型 R^2 为 0.67, GTWR 模型 (R^2 为 0.78) 即便具有解释时空异质性的能力, 但验证结果同样明显下降. 在相同条件下, 组合模型的验证结果普遍优于单一模型. 除 GTWR-KNN 模型 (R^2 、MAE 和 RMSE 分别为 0.78、8.92 $\mu\text{g}\cdot\text{m}^{-3}$ 和 12.61 $\mu\text{g}\cdot\text{m}^{-3}$) 验证结果稍差, 其余组合模型均保持着较好的验证精度, GTWR-RF 模型 R^2 、MAE 和 RMSE 分别为 0.88、5.34 $\mu\text{g}\cdot\text{m}^{-3}$ 和 8.39 $\mu\text{g}\cdot\text{m}^{-3}$, GTWR-AdaBoost 模型 R^2 、MAE 和 RMSE 分别为 0.89、5.22 $\mu\text{g}\cdot\text{m}^{-3}$ 和 8.11 $\mu\text{g}\cdot\text{m}^{-3}$, 拟合精度最佳的 GTWR-XGBoost 模型验证精度依旧最佳, R^2 、MAE 和 RMSE 分别为 0.9、4.98 $\mu\text{g}\cdot\text{m}^{-3}$ 和

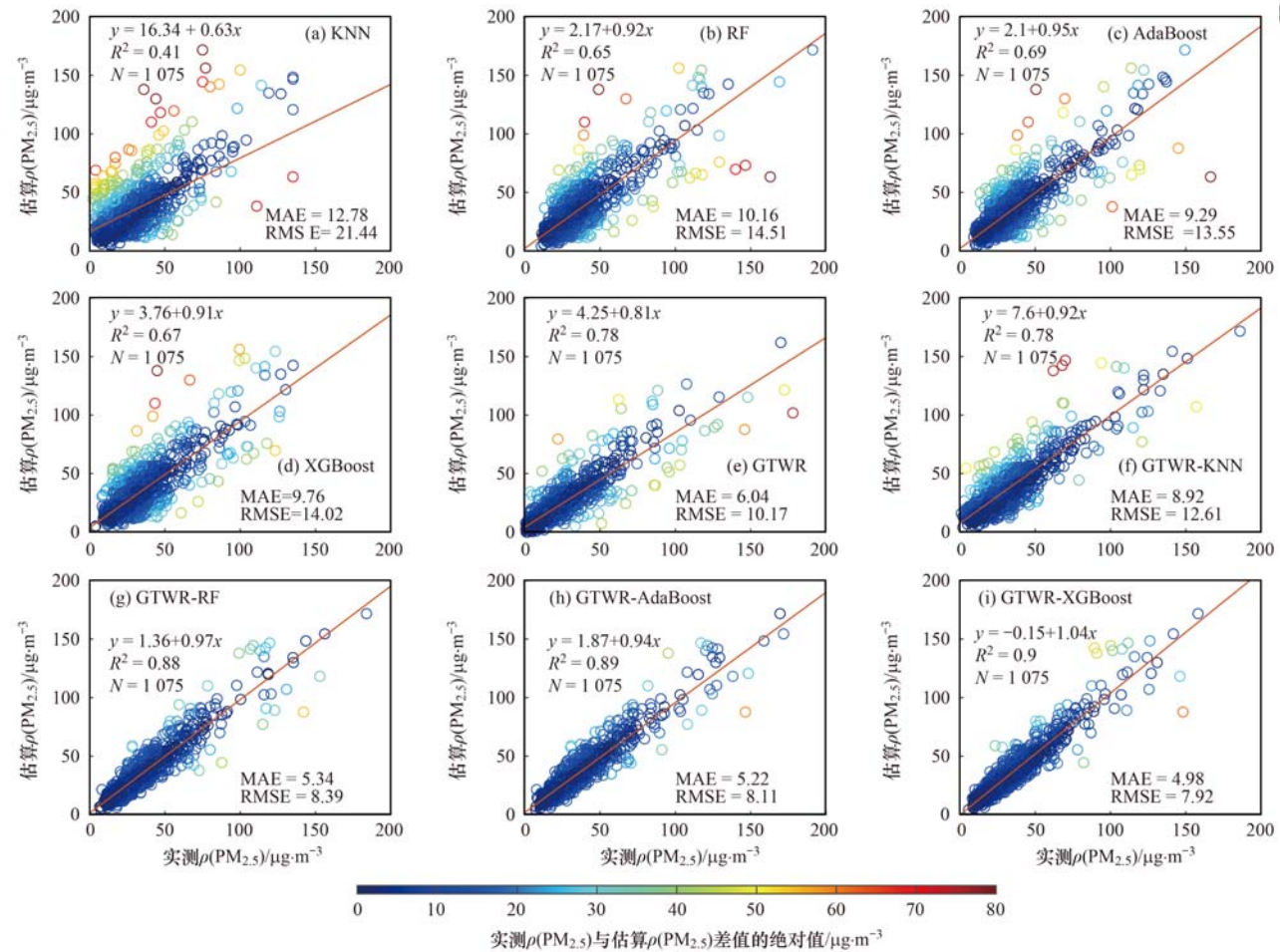


图 5 模型交叉验证结果对比

Fig. 5 Comparison of model cross-validation results

7.92 $\mu\text{g}\cdot\text{m}^{-3}$. 综上所述,对 $\text{PM}_{2.5}$ 浓度进行小时估算采用 GTWR-XGBoost 模型的精度最优.

表 4 模型交叉验证结果评价指标统计

Table 4 Statistical table of evaluation index

模型	交叉验证		
	R^2	MAE/ $\mu\text{g}\cdot\text{m}^{-3}$	RMSE/ $\mu\text{g}\cdot\text{m}^{-3}$
KNN	0.41	12.78	21.44
RF	0.65	10.16	14.51
AdaBoost	0.69	9.29	13.55
XGBoost	0.67	9.76	14.02
GTWR	0.78	6.04	10.17
GTWR-KNN	0.78	8.92	12.61
GTWR-RF	0.88	5.34	8.39
GTWR-AdaBoost	0.89	5.22	8.11
GTWR-XGBoost	0.9	4.98	7.92

2.2 $\text{PM}_{2.5}$ 时空分布分析

2.2.1 模型 $\text{PM}_{2.5}$ 浓度估算时间变化分析

为验证 GTWR-XGBoost 模型的稳定性,使用该模型对研究区内 $\text{PM}_{2.5}$ 浓度小时尺度估算值与实测值进行拟合分析,结果如图 6 所示. 从中可知, GTWR-XGBoost 模型具有较高的拟合优度(全数据集的 R^2 为 0.96,不同时刻的 R^2 介于 0.91~0.98 之

间),表明该模型对于小时估算具有较好的时间稳定性. 且该模型在中午和下午时段内性能较好,在 15:00 性能最佳, R^2 、MAE 和 RMSE 分别为 0.98、3.14 $\mu\text{g}\cdot\text{m}^{-3}$ 和 4.46 $\mu\text{g}\cdot\text{m}^{-3}$,这一发现与 Chen 等^[18] 和 Zang 等^[35] 研究的结果一致. 该现象是由于中午温度较高,行星边界层向上延伸,促进了粒子在垂直方向上的混合,从而有利于细颗粒物浓度的估算^[36].

2.2.2 模型 $\text{PM}_{2.5}$ 浓度估算时空分布

四川省近地表不同时刻(09:00~16:00)年均 $\text{PM}_{2.5}$ 浓度估算值时空分布,见图 7 所示. 从空间上可以看出,低值区主要分布在西部的阿坝、甘孜、凉山和攀枝花等地,全年 $\rho(\text{PM}_{2.5})$ 均值为 15.27 $\mu\text{g}\cdot\text{m}^{-3}$,高值区主要分布在东部的成都、德阳、绵阳、广元、巴中、达州和眉山等地,全年 $\rho(\text{PM}_{2.5})$ 均值为 45.63 $\mu\text{g}\cdot\text{m}^{-3}$. 形成这一时空分布特征,从人口经济方面分析是由于西部人烟稀少、交通不便、工业生产分散,而东部人口密集、工业生产积聚、交通发达,从而污染排放高;从地理环境方面分析是由于四川盆地独特的地形,致使大气环流受限,大

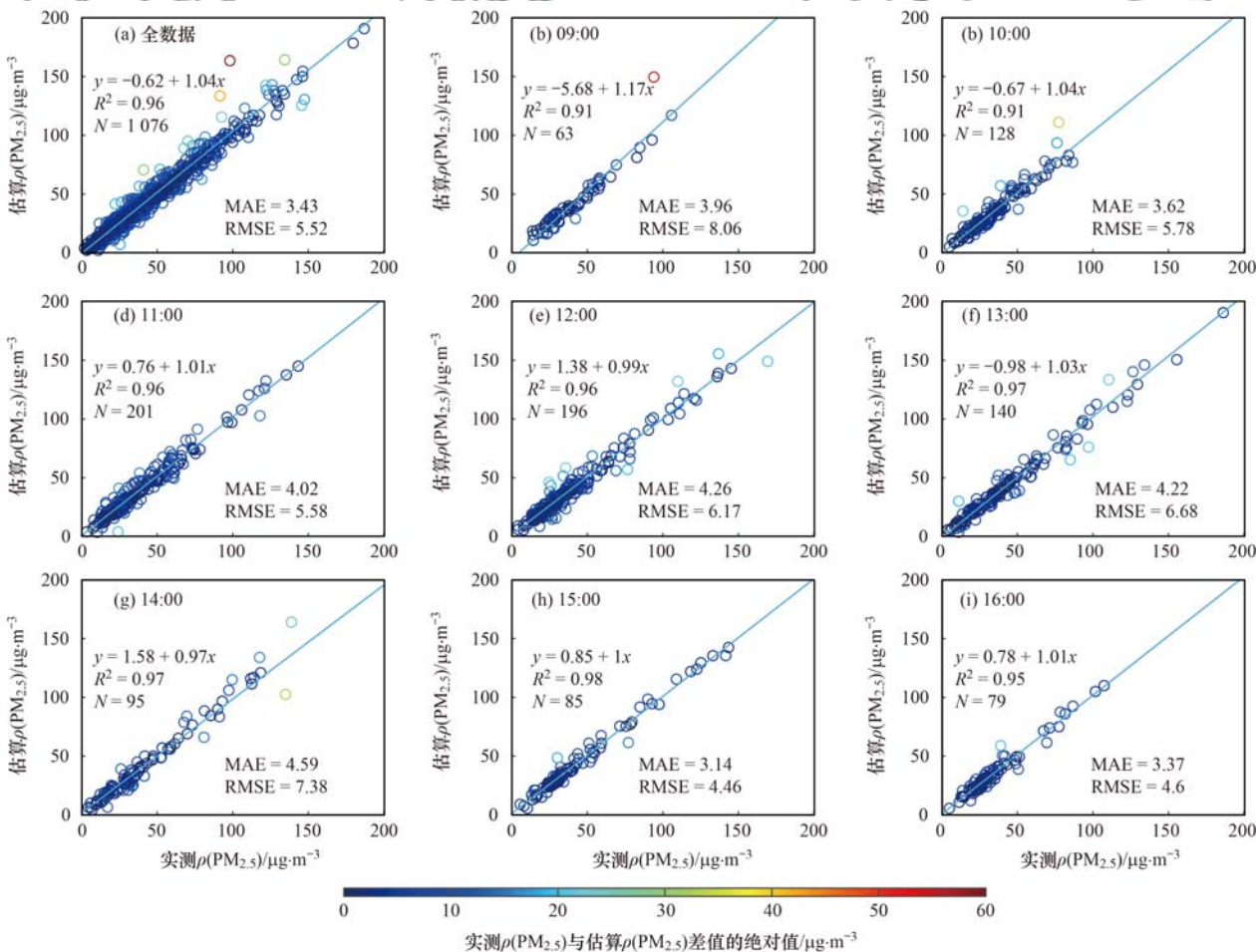


图 6 09:00~16:00 模型拟合结果

Fig. 6 The 09:00-16:00 model fitting results

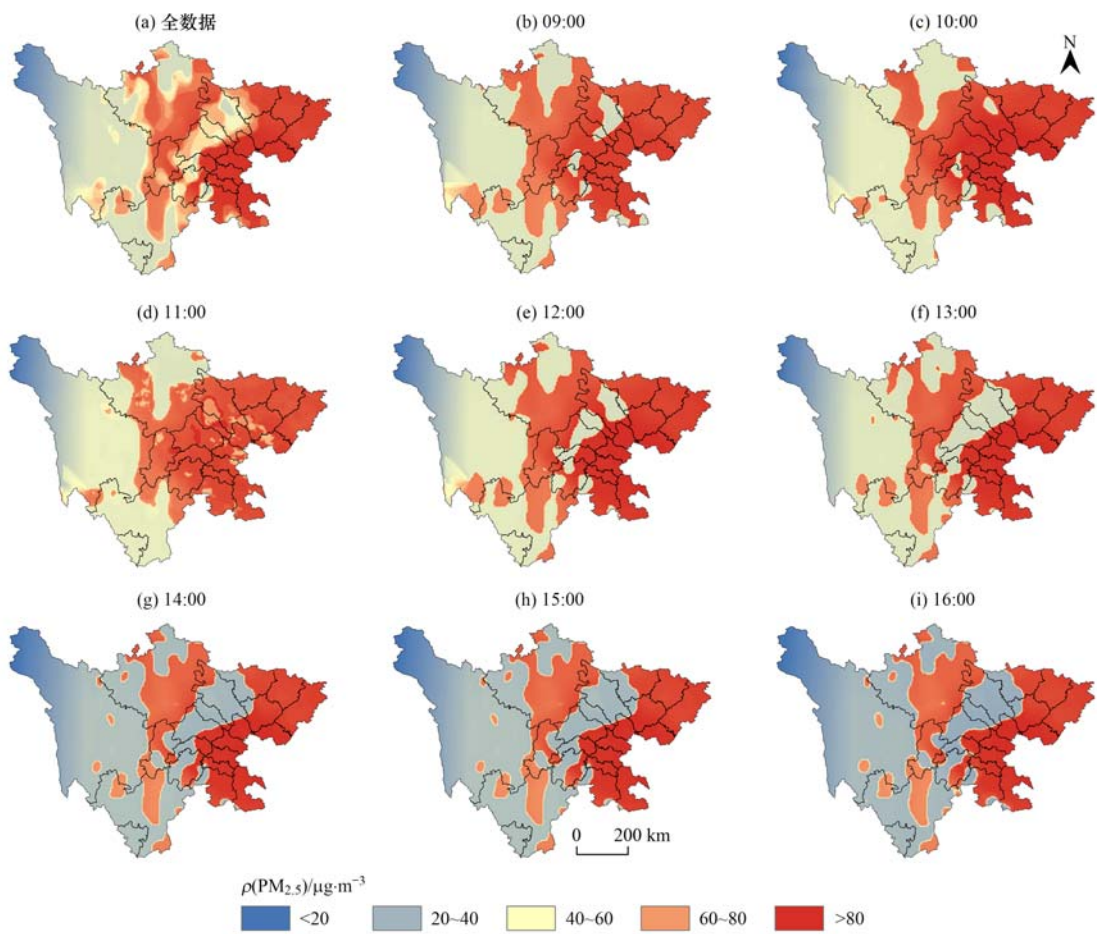


图7 09:00~16:00 PM_{2.5}时空分布
Fig. 7 The 09:00-16:00 spatial and temporal distributions of PM_{2.5}

气输送受阻,污染物聚集,从而 PM_{2.5}浓度较高^[37]。

基于四川省地表不同时刻 PM_{2.5}浓度估算值空间分布求取所有栅格的平均值,即为四川省 09:00~16:00 时段小时 PM_{2.5}浓度均值,见图 8 所示。2017 年四川省年均 PM_{2.5}小时浓度估算变化总体上呈现“倒 U”型趋势,从 09:00 开始升高至 11:00 达到峰值[$\rho(\text{PM}_{2.5})$ 为 $44.56 \mu\text{g}\cdot\text{m}^{-3}$],随后逐渐降低。09:00~11:00 时段 PM_{2.5}浓度升高可能是由于正值上班高峰期,汽车尾气排放量大,人们生活和工业生产活动增加,同时这个时段正是太阳升起的时刻,近地面大气温度比上层大气温度低,容易出现逆温现象,从而污染物积聚。11:00~16:00 时段 PM_{2.5}浓度逐渐下降是因为各类活动降低,汽车排放量减少,同时光照强,上下大气温差大,容易形成对流,有利于污染物向高空扩散。

为更精准地验证模型对 PM_{2.5}小时浓度估算的空间分布变化,将估算结果分为春季(3~5月)、夏季(6~8月)、秋季(9~11月)和冬季(12至次年2月)这4个季节,分析季均 PM_{2.5}小时浓度变化,如图 9 所示(由于时空分布图数量较多,此处仅展示季均 PM_{2.5}小时浓度变化折线图)。

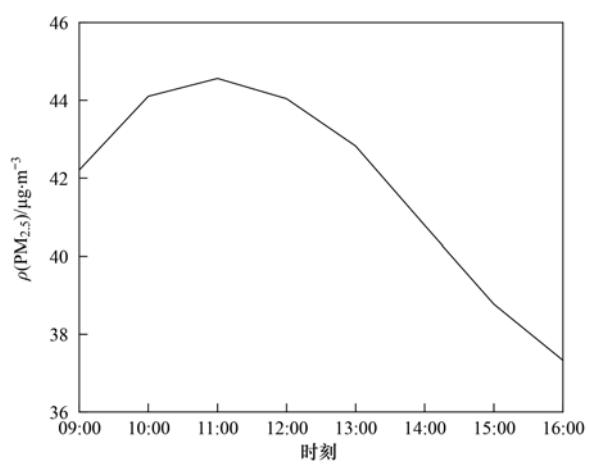


图8 年均 PM_{2.5}小时浓度变化
Fig. 8 Annual hourly concentration change in PM_{2.5}

由图 9 可知:①PM_{2.5}浓度的季节性变化非常明显。冬季污染最为严重[$\rho(\text{PM}_{2.5})$ 均值为 $58.49 \mu\text{g}\cdot\text{m}^{-3}$,变化范围 $43.38 \sim 65.96 \mu\text{g}\cdot\text{m}^{-3}$],这主要是因为冬季燃煤采暖的大范围使用,加之温度低,气压高,降雨量少和逆温现象,导致近地面大气湍流作用弱,扩散条件差,所以污染物难以扩散而堆积^[23,38];春季[$\rho(\text{PM}_{2.5})$ 均值为 $37.54 \mu\text{g}\cdot\text{m}^{-3}$,变化范围

29.31 ~ 47.84 $\mu\text{g}\cdot\text{m}^{-3}$] 与秋季 [$\rho(\text{PM}_{2.5})$ 均值为 34.04 $\mu\text{g}\cdot\text{m}^{-3}$, 变化范围 26.49 ~ 40.24 $\mu\text{g}\cdot\text{m}^{-3}$] 污染次之; 夏季污染最轻 [$\rho(\text{PM}_{2.5})$ 均值为 24.69 $\mu\text{g}\cdot\text{m}^{-3}$, 变化范围 19.39 ~ 30.99 $\mu\text{g}\cdot\text{m}^{-3}$], 这主要因为夏季降水量增多使空气湿度变大, 地面尘土和工业排放的污染物等不易悬浮扩散^[39], 又因地表温度高, 导致上升气流为主, 有利于空气中悬浮的污染物扩散, 难以在空气中堆积, 所以夏季 $\text{PM}_{2.5}$ 浓度值最低。②四季 $\text{PM}_{2.5}$ 小时浓度估算的变化虽不尽相同, 但基本符合“倒 U”趋势。四季的 $\text{PM}_{2.5}$ 小时浓度估算均从 09:00 开始上升, 春季和冬季在 11:00 达到峰值, 夏季在 10:00 达到峰值, 秋季在 13:00 达到峰值, 随后下降过程中均略有浮动。四季的不同趋势变化可能是当季的太阳照射时长、光照强度、环境温度、空气湿度以及人类活动时间等不同因素共同影响。

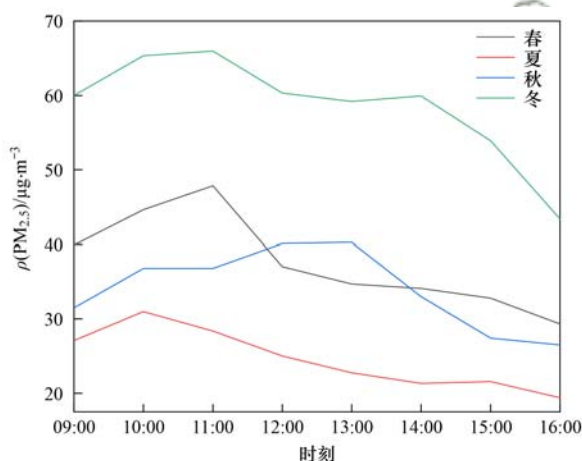


图9 季均 $\text{PM}_{2.5}$ 小时浓度变化

Fig. 9 Seasonal average hourly concentration change in $\text{PM}_{2.5}$

3 结论

(1) 提出的 GTWR-XGBoost 组合模型运用于全数据集对 $\text{PM}_{2.5}$ 浓度的小时估算性能最优, 拟合精度指标 R^2 、MAE 和 RMSE 分别为 0.96、3.43 $\mu\text{g}\cdot\text{m}^{-3}$ 和 5.52 $\mu\text{g}\cdot\text{m}^{-3}$, 验证精度指标 R^2 、MAE 和 RMSE 分别为 0.9、4.98 $\mu\text{g}\cdot\text{m}^{-3}$ 和 7.92 $\mu\text{g}\cdot\text{m}^{-3}$ 。明显优于其他 8 组对照模型。充分表明 GTWR-XGBoost 模型对四川省小时 $\text{PM}_{2.5}$ 浓度估算效果最好, 精度最高, 结果最接近地面站点测量的真实值, 能够为区域空气质量评估提供精准估算信息。

(2) 该模型作用于 $\text{PM}_{2.5}$ 浓度小时估算上, 具有较高的拟合优度 (全数据集的 R^2 为 0.96, 不同时刻的 R^2 介于 0.91 ~ 0.98 之间), 在中午和下午性能较好, 于 15:00 达到性能最佳, R^2 、MAE 和 RMSE 分别为 0.98、3.14 $\mu\text{g}\cdot\text{m}^{-3}$ 和 4.46 $\mu\text{g}\cdot\text{m}^{-3}$ 。表明该模

型在小时 $\text{PM}_{2.5}$ 浓度估算中具有优异的时间稳定性。

(3) 从时间上看, 年均 $\text{PM}_{2.5}$ 小时浓度估算变化总体上呈现“倒 U”型趋势, 上午 9:00 开始升高至 11:00 达到峰值 [$\rho(\text{PM}_{2.5})$ 为 44.56 $\mu\text{g}\cdot\text{m}^{-3}$], 随后逐渐降低, 季均 $\text{PM}_{2.5}$ 小时浓度估算变化虽不尽相同, 但基本符合“倒 U”趋势; 且季节性变化非常明显, 为: 冬季 > 春季 > 秋季 > 夏季。

(4) 从空间分布上看, 总体上呈现东高西低和局部污染程度较高的特点, 高值区主要分布在城市快速发展和人口密集的东部地区, 低值区主要分布在经济发展落后和人口稀疏的西部地区。

(5) 本文仍存在不足。由于缺乏夜间的 AOD 数据, 文中仅对白天的 $\text{PM}_{2.5}$ 浓度做出估算研究。在未来的工作中, 可将地面站点观测值、夜间灯光数据与卫星 AOD 产品整合, 以获得更高时间范围和更高精度的数据集用于估算研究。

参考文献:

- [1] Kim Y, Manley J, Radoias V. Medium- and long-term consequences of pollution on labor supply: evidence from Indonesia[J]. IZA Journal of Labor Economics, 2017, 6(1), doi: 10.1186/s40172-017-0055-2.
- [2] Li S X, Chen L L, Huang G, et al. Retrieval of surface $\text{PM}_{2.5}$ mass concentrations over North China using visibility measurements and GEOS-Chem simulations[J]. Atmospheric Environment, 2020, 222, doi: 10.1016/j.atmosenv.2019.117121.
- [3] Pope III C A, Hansen J C, Kuprov R, et al. Vascular function and short-term exposure to fine particulate air pollution[J]. Journal of the Air & Waste Management Association, 2011, 61(8): 858-863.
- [4] 于雪, 赵文吉, 孙春媛, 等. 大气 $\text{PM}_{2.5}$ 遥感反演研究进展[J]. 环境污染与防治, 2017, 39(10): 1153-1158.
Yu X, Zhao W J, Sun C Y, et al. Progress study on remote sensing retrieval of atmospheric $\text{PM}_{2.5}$ concentration[J]. Environmental Pollution & Control, 2017, 39(10): 1153-1158.
- [5] 何秀, 邓兆泽, 李成才, 等. MODIS 气溶胶光学厚度产品在地面 PM_{10} 监测方面的应用研究[J]. 北京大学学报(自然科学版), 2010, 46(2): 178-184.
He X, Deng Z Z, Li C C, et al. Application of MODIS AOD in surface PM_{10} evaluation[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2010, 46(2): 178-184.
- [6] Wang Z F, Chen L F, Tao J H, et al. Satellite-based estimation of regional particulate matter (PM) in Beijing using vertical-and-RH correcting method[J]. Remote Sensing of Environment, 2010, 114(1): 50-63.
- [7] 王静, 杨复沫, 王鼎益, 等. 北京市 MODIS 气溶胶光学厚度和 $\text{PM}_{2.5}$ 质量浓度的特征及其相关性[J]. 中国科学院研究生院学报, 2010, 27(1): 10-16.
Wang J, Yang F M, Wang D Y, et al. Characteristics and relationship of aerosol optical thickness and $\text{PM}_{2.5}$ concentration over Beijing[J]. Journal of the Graduate School of the Chinese Academy of Sciences, 2010, 27(1): 10-16.
- [8] 赵佳楠, 徐建华, 卢德彬, 等. 基于 RF-LUR 模型的 $\text{PM}_{2.5}$ 空间分布模拟——以长江三角洲地区为例[J]. 地理与地理信

- 息科学, 2018, **34**(1): 18-23.
- Zhao J N, Xu J H, Lu D B, *et al.* The spatial distribution simulation of PM_{2.5} concentration based on RF-LUR model; a case study of Yangtze River delta [J]. *Geography and Geo-information Science*, 2018, **34**(1): 18-23.
- [9] Kloog I, Nordio F, Coull B A, *et al.* Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal PM_{2.5} exposures in the mid-Atlantic states [J]. *Environmental Science & Technology*, 2012, **46**(21): 11913-11921.
- [10] 何文通. 基于 GWR 模型的 PM_{2.5} 质量浓度遥感分析方法研究[D]. 北京: 中国科学院大学, 2015.
- He W T. Study on remote sensing analysis method of PM_{2.5} mass concentration based on GWR model[D]. Beijing: University of Chinese Academy of Sciences, 2015.
- [11] Bai Y, Li Y, Wang X X, *et al.* Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions [J]. *Atmospheric Pollution Research*, 2016, **7**(3): 557-566.
- [12] Jiang Z F, Mao B, Meng X X, *et al.* An air quality forecast model based on the BP neural network of the samples self-organization clustering[A]. In: *IEEE 2010 Sixth International Conference on Natural Computation*[C]. Yantai: IEEE, 2010. 1523-1527.
- [13] Li X K, Zhang C R, Li W D, *et al.* Evaluating the use of DMSP/OLS nighttime light imagery in predicting PM_{2.5} concentrations in the northeastern United States[J]. *Remote Sensing*, 2017, **9**(6), doi: 10.3390/rs9060620.
- [14] He Q Q, Huang B. Satellite-based mapping of daily high-resolution ground PM_{2.5} in China via space-time regression modeling[J]. *Remote Sensing of Environment*, 2018, **206**: 72-83.
- [15] 侯俊雄, 李琦, 朱亚杰, 等. 融机器学习与 WRF 大气模式的 PM_{2.5} 预报方法[J]. *测绘科学*, 2018, **43**(2): 114-120, 141.
- Hou J X, Li Q, Zhu Y J, *et al.* PM_{2.5} forecasting method based on machine learning and WRF hybrid model [J]. *Science of Surveying and Mapping*, 2018, **43**(2): 114-120, 141.
- [16] Pan B Y. Application of XGBoost algorithm in hourly PM_{2.5} concentration prediction[J]. *IOP Conference Series: Earth and Environmental Science*, 2018, **113**, doi: 10.1088/1755-1315/113/1/012127.
- [17] Zhai B X, Chen J G. Development of a stacked ensemble model for forecasting and analyzing daily average PM_{2.5} concentrations in Beijing, China[J]. *Science of the Total Environment*, 2018, **635**: 644-658.
- [18] Chen J P, Yin J H, Zang L, *et al.* Stacking machine learning model for estimating hourly PM_{2.5} in China based on Himawari 8 aerosol optical depth data[J]. *Science of the Total Environment*, 2019, **697**, doi: 10.1016/j.scitotenv.2019.134021.
- [19] Zang L, Mao F Y, Guo J P, *et al.* Estimating hourly PM₁ concentrations from Himawari-8 aerosol optical depth in China[J]. *Environmental Pollution*, 2018, **241**: 654-663.
- [20] 吴宇宏, 杜宁, 王莉, 等. 基于 iLME + Geoi-RF 模型的四川省 PM_{2.5} 浓度估算[J]. *环境科学*, 2021, **42**(12): 5602-5615.
- Wu Y H, Du N, Wang L, *et al.* Estimation of PM_{2.5} concentration in Sichuan province based on improved linear mixed effect model and geo-intelligent random forest[J]. *Environmental Science*, 2021, **42**(12): 5602-5615.
- [21] 方晓婷, 段华波, 胡明伟, 等. 气象因素对大气污染物影响的季节差异分析及预测模型对比——以深圳为例[J]. *环境污染与防治*, 2019, **41**(5): 541-546.
- Fang X T, Duan H B, Hu M W, *et al.* The seasonal differential effects of meteorological parameters on atmospheric pollutants and the prediction model comparison: a case study of Shenzhen[J]. *Environmental Pollution & Control*, 2019, **41**(5): 541-546.
- [22] 胡译文, 臧增亮, 马晓燕, 等. 气象资料同化对 PM_{2.5} 预报影响的模拟分析[J]. *中国环境科学*, 2019, **39**(2): 523-532.
- Hu Y W, Zang Z L, Ma X Y, *et al.* Research on the effects of assimilation meteorological observation data on aerosol concentration[J]. *China Environmental Science*, 2019, **39**(2): 523-532.
- [23] 金团团, 杨兴川, 晏星, 等. 京津冀及周边 MAIAC AOD 和 PM_{2.5} 质量浓度特征及相关性分析[J]. *环境科学*, 2021, **42**(6): 2604-2615.
- Jin J N, Yang X C, Yan X, *et al.* MAIAC AOD and PM_{2.5} mass concentrations characteristics and correlation analysis in Beijing-Tianjin-Hebei and surrounding areas [J]. *Environmental Science*, 2021, **42**(6): 2604-2615.
- [24] Retalis A, Sifakis N. Urban aerosol mapping over Athens using the differential textural analysis (DTA) algorithm on MERIS-ENVISAT data [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2010, **65**(1): 17-25.
- [25] Brunson C, Fotheringham A S, Charlton M E. Geographically weighted regression: a method for exploring spatial nonstationarity [J]. *Geographical Analysis*, 1996, **28**(4): 281-298.
- [26] Huang B, Wu B, Barry M. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices [J]. *International Journal of Geographical Information Science*, 2010, **24**(3): 383-401.
- [27] 肖燕婷, 田铮, 郭文艳. 时空地理加权模型回归关系的非平稳性 Bootstrap 检验[J]. *统计与决策*, 2014, **30**(9): 8-12.
- Xiao Y T, Tian Z, Guo W Y. Nonstationary Bootstrap test for regression relationship of spatio-temporal geographically weighted model[J]. *Statistics & Decision*, 2014, **30**(9): 8-12.
- [28] 杨毅. 顾及时空非平稳性的地理加权回归方法研究[D]. 武汉: 武汉大学, 2016.
- Yang Y. Research on geographically and temporally weighted regression for spatial and temporal nonstationarity[D]. Wuhan: Wuhan University, 2016.
- [29] 杨东阳. 基于分层时空建模的长江三角洲地区 PM_{2.5} 浓度估算研究[D]. 上海: 华东师范大学, 2019.
- Yang D Y. Estimating PM_{2.5} concentrations using a hierarchical spatio-temporal model in the Yangtze River Delta[D]. Shanghai: East China Normal University, 2019.
- [30] Chen T Q, Guestrin C. XGBoost: a scalable tree boosting system [A]. In: *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*[C]. San Francisco, CA: Association for Computing Machinery, 2016. 785-794.
- [31] 叶倩怡, 饶泓, 姬名书. 基于 Xgboost 的商业销售预测[J]. *南昌大学学报(理科版)*, 2017, **41**(3): 275-281.
- Ye Q Y, Rao H, Ji M S. Sales prediction of stores based on Xgboost algorithm[J]. *Journal of Nanchang University (Natural Science)*, 2017, **41**(3): 275-281.
- [32] Wu Y C, Qi S F, Hu F, *et al.* Recognizing activities of the elderly using wearable sensors: a comparison of ensemble algorithms based on boosting[J]. *Sensor Review*, 2019, **39**(6): 743-751.
- [33] 康俊锋, 谭建林, 方雷, 等. XGBoost-LSTM 变权组合模型支持下短期 PM_{2.5} 浓度预测——以上海为例[J]. *中国环境科学*, 2021, **41**(9): 4016-4025.
- Kang J F, Tan J L, Fang L, *et al.* Short-term PM_{2.5}

- concentration prediction based on XGBoost and LSTM variable weight combination model: a case study of Shanghai[J]. *China Environmental Science*, 2021, **41**(9): 4016-4025.
- [34] 郝静, 孙成, 郭兴宇, 等. 京津冀内陆平原区 $PM_{2.5}$ 浓度时空变化定量模拟[J]. *环境科学*, 2018, **39**(4): 1455-1465.
Hao J, Sun C, Guo X Y, *et al.* Simulation of the Spatio-temporally resolved $PM_{2.5}$ aerosol mass concentration over the inland plain of the Beijing-Tianjin-Hebei region [J]. *Environmental Science*, 2018, **39**(4): 1455-1465.
- [35] Zang L, Mao F Y, Guo J P, *et al.* Estimation of spatiotemporal $PM_{1.0}$ distributions in China by combining $PM_{2.5}$ observations with satellite aerosol optical depth [J]. *Science of the Total Environment*, 2019, **658**: 1256-1264.
- [36] Guo J P, Miao Y C, Zhang Y, *et al.* The climatology of planetary boundary layer height in China derived from radiosonde and reanalysis data[J]. *Atmospheric Chemistry and Physics*, 2016, **16**(20): 13309-13319.
- [37] Zhang L, Guo X M, Zhao T L, *et al.* A modelling study of the terrain effects on haze pollution in the Sichuan Basin [J]. *Atmospheric Environment*, 2019, **196**: 77-85.
- [38] 袁杨森, 刘大锰, 车瑞俊, 等. 北京市秋季大气颗粒物的污染特征研究[J]. *生态环境*, 2007, **16**(1): 18-25.
Yuan Y S, Liu D M, Che R J, *et al.* Research on the pollution situation of atmospheric particulates during Autumn in Beijing city [J]. *Ecology and Environment*, 2007, **16**(1): 18-25.
- [39] 杨晓辉, 宋春杰, 范丽行, 等. 京津冀地区高分辨率 $PM_{2.5}$ 浓度时空变化模拟与分析[J]. *环境科学*, 2021, **42**(9): 4083-4094.
Yang X H, Song C J, Fan L H, *et al.* High-resolution Estimation of Spatio-temporal Variation in $PM_{2.5}$ concentrations in the Beijing-Tianjin-Hebei region[J]. *Environmental Science*, 2021, **42**(9): 4083-4094.



CONTENTS

Assessing the Environmental and Health Co-benefits of Accelerated Energy Transition and Industrial Restructuring: A Case Study of the BTHS Region	YANG Xi, SUN Yi-sheng, CHANG Shi-yan, <i>et al.</i> (3627)
Synergistic Paths of Reduced Pollution and Carbon Emissions Based on Different Power Demands in China	XIANG Meng-yu, WANG Shen, LÜ Lian-hong, <i>et al.</i> (3637)
Evaluation Method and Application for Urban Carbon Peaking & Neutrality Performance	ZHANG Bao-liu, BAI Zi-han, ZHANG Nan, <i>et al.</i> (3649)
Revealing Driving Factors of Urban O ₃ Based on Explainable Machine Learning	DONG Jia-qi, HU Dong-mei, YAN Yu-long, <i>et al.</i> (3660)
Sensitivity Analysis of Ozone Formation Using Response Surface Methodology	ZHU Yu-huan, CHEN Bing, ZHANG Ya-ru, <i>et al.</i> (3669)
Analysis of O ₃ Sources in Yulin City in Summer Based on WRF-CMAQ/ISAM Model	WANG Yi-fan, TONG Ji-long, CHEN Yu-xiang, <i>et al.</i> (3676)
Atmospheric Ozone Concentration Prediction in Nanjing Based on LightGBM	ZHU Jia-ying, AN Jun-lin, FENG Yue-zheng, <i>et al.</i> (3685)
Meteorological Formation Mechanisms and Potential Sources of an Ozone Pollution Process in Winter of 2022 in Guangdong Province	LI Ting-yuan, CHEN Jing-yang, GONG Yu, <i>et al.</i> (3695)
Impact of Summer Tropospheric Ozone Radiative Forcing on Meteorology and Air Quality in North China	DU Nan, CHEN Lei, LIAO Hong, <i>et al.</i> (3705)
Identification of Impacts from Meteorology and Local and Transported Photochemical Generation on Ozone Trends in Changsha from 2018 to 2020	YANG Jun, YANG Lei-feng, DING Hua, <i>et al.</i> (3715)
Spatio-temporal Variation and Multi-dimensional Detection of Driving Mechanism of PM _{2.5} Concentration in the Chengdu-Chongqing Urban Agglomeration from 2000 to 2021	XU Yong, GUO Zhen-dong, ZHENG Zhi-wei, <i>et al.</i> (3724)
Estimation of PM _{2.5} Hourly Concentration in Sichuan Province Based on GTWR-XGBoost Model	WU Di, DU Ning, WANG Li, <i>et al.</i> (3738)
Transmission and Growth Characteristics of Severe PM _{2.5} Pollution Events from 2013 to 2021 in Xingtai, Hebei	JIANG Qi, SHENG Li, JIN Yu-chen, <i>et al.</i> (3749)
Chemical Characteristics and Source Apportionment of Organic Aerosols in Urban Shanghai During Cold Season Based on High Time-resolution Measurements of Organic Molecular Markers	ZHU Shu-hui (3760)
Emission Factors of Carbonaceous Aerosol and Stable Carbon Isotope for In-use Vehicles	YU Ming-yuan, WANG Qian, FU Ming-liang, <i>et al.</i> (3771)
Composition Characteristics of Volatile Organic Compounds and Associated Contributions to Secondary Pollution in Shenyang Industrial Area in Summer	GUAN Lu, SU Cong-cong, KU Ying-ying, <i>et al.</i> (3779)
Characteristics of VOCs and Assessment of Emission Reduction Effect During the Epidemic Lockdown Period in Shenzhen Urban Area	YUN Long, LIN Chu-xiong, LI Cheng-liu, <i>et al.</i> (3788)
Characteristics of Organic Matter Composition and Oxidation Potential in Road Dust in Winter in Xi'an	WANG Qing-wen, CHEN Qing-cai, WANG Chao, <i>et al.</i> (3797)
Contamination Characteristics and Risk Assessment of Polycyclic Aromatic Compounds in Surface Dust of Suntuan Mining Area in Huaibei	XU Zhen-peng, QIAN Ya-hui, HONG Xiu-ping, <i>et al.</i> (3809)
Spatial-temporal Variation in Water Quality and Its Response to Precipitation and Land Use in Baiyangdian Lake in the Early Stage of the Construction of Xiong'an New Area	WANG Zi-ming, YANG Li-hu, SONG Xian-fang (3820)
Simulation of Pollution Apportionment and Optimization of Control Methods in Watershed Scale: A Case Study of the Shun'an Watershed in Tongling City	LIU Guo-wangchen, CHEN Lei, LI Jia-qi, <i>et al.</i> (3835)
Evaluation of Shallow Groundwater Quality and Optimization of Monitoring Indicators in Nanchang	ZHENG Zi-yin, CHU Xiao-dong, XU Jin-ying, <i>et al.</i> (3846)
Hydrogen and Oxygen Isotopic Characteristics and Influencing Factors of "Three Waters" in Shandian River Basin	YANG Li-na, JIA De-bin, GAO Rui-zhong, <i>et al.</i> (3855)
Distribution Characteristics and Influencing Factors of Abundant and Rare Planktonic Microeukaryotes in Jinsha River	YAN Bing-cheng, CUI Ge, SUN Sheng-hao, <i>et al.</i> (3864)
Bacterial Community Diversity in Channel Sediments of Different Disturbance Sections of the Jialing River	ZHU Lan-ping, ZHANG Tuo, LI Jia-ning, <i>et al.</i> (3872)
Effects of Reservoir Water Depth on Different Plankton Communities and Keystone Species of Network Interaction	WANG Xun, LIAO Qin, WANG Pei-fang, <i>et al.</i> (3881)
Correlation Between the Diversity Characteristics of Groundwater Bacterial Community and Environmental Factors in Typical Industrial Areas	WU Jian-qiang, ZHANG Shu-yuan, WANG Min, <i>et al.</i> (3892)
Effects of Fertilizer Application Strategy Adjustments on Nitrogen and Phosphorus Loss from Typical Crop Systems in Taihu Lake Region	YU Ying-liang, WANG Yi-zhi, YANG Bei, <i>et al.</i> (3902)
Estimation of Cropland Nitrogen Runoff Loss Loads in the Yangtze River Basin Based on the Machine Learning Approaches	ZHANG Yu-fu, PAN Zhe-qi, CHEN Ding-jiang (3913)
Classification and Identification of Non-point Source Nitrogen Pollution in Surface Flow of the Shangwu River Watershed in the Qiandao Lake Region	YU Ke, YAN Yan, TANG Zhang-xuan, <i>et al.</i> (3923)
Spatial Distribution of Nitrogen and Phosphorus Nutrients in the Main Stream and Typical Tributaries of Tuojiang River and Fujiang River	LI Zi-yang, ZHOU Ming-hua, XU Peng, <i>et al.</i> (3933)
Sediment Pollution and Dredging Effect of Waqinhui River	ZHANG Mu, REN Zeng-yi, ZHANG Man, <i>et al.</i> (3945)
Critical Review on Environmental Occurrence and Photochemical Behavior of Substituted Polycyclic Aromatic Hydrocarbons	GE Lin-ke, WANG Zi-yu, CAO Sheng-kai, <i>et al.</i> (3957)
Cadmium and Arsenic Interactions During Co-adsorption onto Goethite	SU Zi-xian, LIU Sai-hong, GUAN Yu-feng, <i>et al.</i> (3970)
Preparation of Catalyst Cyclodextrin-Fe-TAML to Activate H ₂ O ₂ and Oxidize Organic Micropollutants in Water	LIU Qing-quan, CAI Ben-zhe, CAI Xi-yun (3978)
Performance and Reaction Mechanism of Co(II) Mediated Activation of Peroxymonosulfate for Degrading Nitrilotris (Methylene Phosphonic Acid)	ZHU Jing-lin, WANG Shu (3990)
Change in Granulation Potential and Microbial Enrichment Characteristics of Sludge Induced by Microplastics	XIE Qing-fan, YU Nan, ZHANG Ni, <i>et al.</i> (3997)
Accumulation Characteristics and Probabilistic Risk Assessment of Cd in Agricultural Soils Across China	WANG Jing, WEI Heng, PAN Bo (4006)
Source Analysis and Pollution Assessment of Soil Heavy Metals in Typical Geological High Background Area in Southeastern Chongqing	JIANG Yu-lian, YU Jing, WANG Rui, <i>et al.</i> (4017)
Integrated Analysis on Source-exposure Risk of Heavy Metals in Farmland Soil Based on PMF Model: A Case Study in the E-waste Dismantling Area in Zhejiang Province	FANG Jia, HE Ying, HUANG Nai-tao, <i>et al.</i> (4027)
Pollution Characteristics and Risk Assessment of Polycyclic Aromatic Hydrocarbons in Farmland Soil and Crops in the Suburbs of Urumqi	FAN Yue, CAO Shuang-yu, Nuerla Ailijiang, <i>et al.</i> (4039)
Distribution and Driving Mechanisms of Antibiotic Resistance Genes in Desert-Oasis Continuum	HUANG Fu-yi, ZHOU Shu-yidan, SU Jian-qiang, <i>et al.</i> (4052)
Characteristics of Antibiotic Contamination of Soil in China in Past Fifteen Years and the Bioremediation Technology: A Review	ZHAO Xiao-dong, QIAO Qing-qing, QIN Xiao-rui, <i>et al.</i> (4059)
Modified Biochar for Remediation of Soil Contaminated with Arsenic and Cadmium: A Review	LÜ Peng, LI Lian-fang, HUANG Xiao-ya (4077)
Effect of Biogas Slurry Return to Field on Heavy Metal Accumulation in Soil-crop System: A Meta-analysis	ZHAO Qi-zhi, YANG Zhi-min, KONG Fan-jing, <i>et al.</i> (4091)
Remediation Effect of Two Iron-modified Biochars on Slightly Alkaline Arsenic and Cadmium Contaminated Soil	LIANG Xin-ran, HE Dan, ZHENG Zhao-hua, <i>et al.</i> (4100)
Effects of Straw Removal Measure on Soil Cd Bioavailability and Rice Cd Accumulation	WANG Zi-yu, ZHOU Hang, ZHOU Kun-hua, <i>et al.</i> (4109)
Application of Desulphurized Gypsum with Straw to Improve Physicochemical Properties of Saline-alkali Land in Yellow River Delta	ZHAO Hui-li, YU Jin-yi, LIU Tao, <i>et al.</i> (4119)
Biological Evaluation and Key Stress Factor Diagnosis of Compound Contaminated Soil Based on Environmental DNA	HUANG Xiang-yun, ZHONG Wen-jun, LIU Xun-jie, <i>et al.</i> (4130)
Comprehensive Quality Assessment of Soil-Maize Heavy Metals in High Geological Background Area	ZHANG Chuan-hua, WANG Zhong-shu, LIU Li, <i>et al.</i> (4142)
Ecological Risk and Health Risk of Heavy Metal Pollution in Vegetable Production System of Zhejiang Province	ZHANG Shu-min, LIU Cui-ling, YANG Gui-ling, <i>et al.</i> (4151)
Effect of Biochar with Phosphorus Fertilizer on Soil Nutrients, Enzyme Activity, and Nutrient Uptake of <i>Alfalfa</i>	LIU Xin-yu, WANG Dong-mei, ZHANG Ze-zhou, <i>et al.</i> (4162)
Effects of Low-density Polyethylene Microplastics on the Growth and Physiology Characteristics of <i>Ipomoea aquatica</i> Forsk	ZHOU Ying, JIANG Wen-ting, LIU Xun-jie, <i>et al.</i> (4170)
Effects of Biogas Slurry Application on Soil Microbial Communities Structure and Function During Wheat-rice Stubble Period	QIAO Yu-ying, XI Hui, LI Na, <i>et al.</i> (4179)
Analysis of Soil Bacterial Community Structure and Ecological Function Characteristics in Different Pollution Levels of Lead-zinc Tailings in Datong	LIU Ze-xun, ZHUANG Jia-yao, LIU Chao, <i>et al.</i> (4191)
Effect of Fire-deposited Charcoal on Soil Organic Carbon Pools and Associated Enzyme Activities in a Recently Harvested <i>Pinus massoniana</i> Plantation Subjected to Broadcast Burning	YAO Zhi, JIAO Peng-yu, WU Xiao-sheng, <i>et al.</i> (4201)