姜明栋,王馨阳,颜 蓉,等.考虑经济活动强度的区域臭氧浓度多尺度预测 [J]. 中国环境科学, 2025,45(9):4737-4748.
Jiang M D, Wang X Y, Yan R, et al. Multi scale prediction of regional ozone concentration considering economic activity intensity [J]. China Environmental science, 2025,45(9):4737-4748.

考虑经济活动强度的区域臭氧浓度多尺度预测

姜明栋[1],王馨阳[2],颜 蓉[3],于欣鑫[4],毋泽鹏[5*] (1.东南大学公共管理系,东南大学长三角碳中和战略发展研究院,江苏 南京 211189；2.College of Business and Economics,Australian National University, Canberra, Australia 2601；3.盐城工学院绿色低碳发展研究院,江苏 盐城 224051；4.中国人民大学生态环境学院,北京 100872；5.北京大学环境科学与工程学院,北京 100871)

摘要：以夜间灯光作为经济活动强度的代理变量,借助随机森林技术构建耦合人类活动和气象因素的 O_3 浓度预测模型,并基于京津冀及周边地区("2+26"城市)2014~2022 年数据展开实证研究,在此基础上进一步比较研究经济活动强度对不同时间尺度,不同地区和不同污染物浓度预测的贡献差异.结果发现:①在不同时间尺度,经济活动强度均对 O_3 浓度预测均有显著贡献,在年尺度,经济活动强度对 O_3 浓度预测的贡献度达到 0.7053,超过各类气象因子成为首要贡献因素,但在日尺度,气温等气象因素的贡献度相对更高,经济活动强度贡献度仅为 0.0869.②不同因素对 O_3 浓度高值区与低值区的预测贡献存在显著差异,气温和经济活动对高值区 O_3 浓度影响显著,而低值区城市表现出不同因子预测贡献的区域性和不连续性特点.③在长时间尺度,人类经济活动对城市 $PM_{2.5}$ 和 O_3 浓度均有决定性贡献;在短时间尺度,经济活动对 $PM_{2.5}$ 浓度的影响仅次于气温,但对 O_3 浓度的驱动贡献较小.

关键词：随机森林模型；臭氧(O_3)；细颗粒物($PM_{2.5}$)；气象因素；夜间灯光；污染预测；"2+26"地区
中图分类号：X511 **文献标识码**：A **文章编号**：1000-6923(2025)09-4737-12

Multi scale prediction of regional ozone concentration considering economic activity intensity. JIANG Ming-dong[1], WANG Xin-yang[2], YAN Rong[3], YU Xin-xin[4], WU Ze-peng[5*] (1.Department of Public Administration, Yangtze River Delta Carbon Neutrality Strategy Development Institute, Southeast University, Nanjing 211189, China；2.College of Business and Economics, Australian National University, Canberra, Australia 2601；3.Green and Low-carbon Development Institute, Yancheng Institute of Technology, Yancheng 224051, China；4.School of Ecology and Environment, Renmin University of China, Beijing100872, China；5.College of Environmental Sciences and Technology, Peking University, Beijing100871, China). *China Environmental science*, 2025,45(9)：4737~4748

Abstract：In this study, nighttime light intensity was adopted as a proxy for economic activity, and a random forest model was constructed to integrate both anthropogenic and meteorological factors for O_3 concentration prediction. Panel data from the Beijing-Tianjin-Hebei region and surrounding areas (the "2+26" cities) during 2014~2022 were employed to empirically examine the differential contributions of economic activity intensity across various temporal scales, spatial regions, and pollutant types. The findings showed that:①economic activity intensity consistently exhibited significant predictive power across different temporal scales. On the annual scale, its importance reached 0.7053, surpassing all meteorological variables and becoming the dominant predictor. In contrast, on the daily scale, temperature and other meteorological factors were found to be more influential, with the contribution of economic activity reduced to 0.0869. ②Marked spatial heterogeneity was observed between high- and low- O_3 concentration areas. Temperature and economic activity had stronger impacts in high- O_3 regions, while low-concentration regions displayed irregular and spatially discontinuous patterns of predictor relevance. ③ Over long time horizons, economic activity was identified as a decisive factor in explaining both $PM_{2.5}$ and O_3 concentrations. However, on shorter time scales, although economic activity was the second most important factor for $PM_{2.5}$ prediction after temperature, its contribution to O_3 prediction remained relatively limited.

Key words：random forest model；ozone (O_3)；fine particulate matter ($PM_{2.5}$)；meteorological factors；nighttime light；pollution prediction；"2+26"urban agglomeration

收稿日期：2025-02-18
基金项目：国家自然科学基金资助国际合作项目(W2412157);国家社会科学基金资助重点项目(21AZD060);国家社会科学基金资助项目(23BGL317);江苏省社会科学基金资助重点项目(22EYA001)

* 责任作者，讲师, zpwu@pku.edu.cn

随着经济快速发展和城市化不断推进,人类经济活动强度显著提高,以京津冀及周边城市为代表的大型城市群大气污染问题愈发严峻[1].在“2+26”城市联防联控机制等政策推动下,该区域$PM_{2.5}$等传统大气污染得到有效控制,但以 O_3 为代表的其他污染物管控形势依然严峻,成为当前大气环境治理工作的重点任务[2].已有不少研究探讨了前体物排放和气象因素对 O_3 污染的影响[3].相关研究多以北京、广州等重点城市或典型城市群为研究对象,证实了挥发性有机化合物(VOCs)和氮氧化物(NO_x)排放在区域 O_3 污染加剧中发挥的关键作用,并提出削减 VOCS 和 NO_x 排放是当前 O_3 污染防控的重点方向[4-6].在此基础上,相关研究进一步探讨了气温、风速等不同气象因素或汽车尾气、工业排放、化石燃料燃烧、城市化进程等人类活动因素对区域 O_3 浓度的贡献[7-13].

伴随机器学习技术的发展,学术界基于该方法提出了多种模型剖析 O_3 浓度的驱动因素并进一步预测其变化趋势.相关研究以全国或典型城市为研究对象,比较分析比湿、气压等不同气象条件与 O_3 浓度间的非线性关系[14-15],同时探讨了 O_3 前体物信息在城市 O_3 建模中的关键性作用[16].还有研究针对中国不同省份,借助机器学习方法研究 O_3 浓度的时空变化特征及其成因,从而为区域性 O_3 污染治理提供依据[17-19].总体而言,这些模型方法在研究精度和适用性方面各有优势,已在多个城市和地区得到了应用.其中,在探究近地层 O_3 浓度与气象条件的关系时,随机森林模型因具有对气象因子非线性影响的拟合能力,已成为 O_3 污染影响因素分析和预测中的一种有效工具[20].相关研究证实其预测精度优于传统统计估算方法,且能有效捕捉 O_3 与气象因子的非线性关系[21].

尽管现有文献已从不同角度对 O_3 污染的成因和预测模型进行了广泛探讨,但仍存在进一步拓展的空间.以往研究大多聚焦于气象因素对 O_3 浓度的影响,较少学者考虑人类活动因素,更鲜有文献将二者同时纳入 O_3 浓度预测模型.与现有研究成果相比,本文的边际贡献主要体现在三方面:一是以夜间灯光作为综合反映经济活动强度的代理变量,将其纳入 O_3 污染的影响因素及预测分析框架中,并进一步探究经济活动因素和自然气象因素对区域 O_3 浓度的交互作用;二是从日度和年度不同时间尺度出发,识别 O_3 浓度短期波动和长期演变的不同贡献因子,并进一步比较研究经济活动强度对不同地区 O_3 浓度预测的贡献差异;三是对比分析经济活动强度和气象因子对 O_3 及 $PM_{2.5}$ 浓度在不同时间尺度上的驱动贡献差异,以期为制定具有针对性的大气污染协同防控策略提供参考.

1 材料与方法

1.1 研究区域

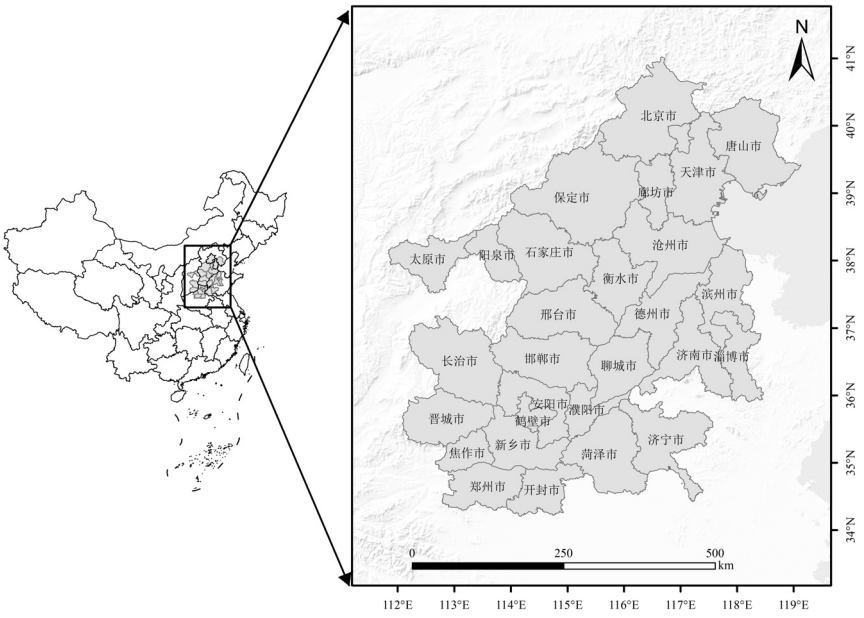


图 1 “2+26”城市区域示意

Fig.1 Schematic diagram of the "2+26" urban area

审图号:GS(2024)0650 号

本文以京津冀及周边地区(“2+26”城市)为研究对象,其范围如图 1 所示,包括北京市、天津市以及河北省 8 个城市、山东省 7 个城市、河南省 7 个城市、山西省 4 个城市.该地区地处华北平原,地理范围介于东经 111°30'~119°50'、北纬 34°16'~41°36'之间,地势呈现出西北高、东南低的特征[22].京津冀及周边地区重工业集聚,高耗能、高污染企业分布广泛.2022 年,“2+26”城市群的平均优良天数比例为 66.7%,较上一年同比下降 0.5 个百分点;$PM_{2.5}$ 平均浓度为 44μg/m^3,同比上升 2.3%;O_3 平均浓度为 179μg/m^3,同比上升 4.7%[23],大气污染防控形势依旧严峻.

1.2　基于随机森林的 O_3 浓度预测方法

随机森林是集成学习中的经典方法之一,属于 Bagging(Bootstrap Aggregating)类技术.该方法通过从原始数据集中进行有放回的随机抽样,构建多个数据子集,并对每个子集训练一个独立模型,最后融合所有模型的预测结果,从而有效提升了模型的稳定性和准确率[24].凭借强大的处理能力和鲁棒性,随机森林在环境监测领域,尤其是区域 O_3 预测中得到了广泛应用.在构建每棵决策树时,随机森林采用了两个关键的随机化步骤:首先,从原始数据集中有放回地随机抽取 M 个样本作为训练集;其次,在决策树构建过程中,每次分裂节点时,从 N 个特征中随机抽取 n 个特征(n 远小于 N),并基于这些特征选择最佳分裂特征,以此构建每棵决策树.在建模过程中,使用 O_3 日最大 8h 滑动平均值作为建模指标.其中,2014~2021 年的数据用于构建模型,2022 年数据用于模型验证.根据研究目标,分别从逐日尺度和逐年尺度开展建模与评估,以探索模型在不同时间尺度下的表现差异.具体步骤如下:首先对数据进行预处理,剔除缺失值和异常值,生成训练数据集 D;接着,通过 Bootstrap 重采样生成多个样本子集 D*,并用这些子集构建子回归树模型,重复此过程 n 次,得到 n 个样本子集和对应的袋外数据集 O*,通过袋外误差确定决策树数量(ntree).在确定 ntree 的最优值后,再通过交叉验证等方法遍历随机特征数(mtry)取值范围,计算每个候选 mtry 下的均方误差(MSE),从而选择最优 mtry 值.最终,依据选定的 ntree 和 mtry 参数构建随机森林模型,并平均 n 棵回归树的预测结果,得到最终预测值.为验证模型在不同时间尺度下稳健性,本文进一步引入 R^2,MSE 等拟合指标进行分析.

1.2.1　特征因子定义在近地层 O_3 污染生成的气象条件研究领域,尽管不同地区存在一定差异,但国内外关于气温,相对湿度等因素对 O_3 浓度影响的研究成果在总体上具有一致性.鉴于不同类型的高低空环流形势以及多尺度气象条件的组合会致使大气边界层呈现出不同的动力和热力条件,进而对 O_3 的生成与传输产生影响[25],因此本文选择气压(PRES)、比湿(QLML)、风速(SPDL)、气温(TLML)和降水量(PREC)等相关变量作为研究的特征因子.

除了上述气象因素外,本文在模型中还引入人类经济活动变量以探究其在 O_3 浓度预测中的作用.具体而言,本文以夜间灯光强度(LIGH)作为经济活动强度的代理变量,用于刻画人类活动对污染物生成的影响.夜间灯光与社会经济活动密切相关[26],引入夜间灯光数据作为经济活动的代理变量,能够综合捕捉人类经济活动对 O_3 浓度的潜在影响[27-28].尽管夜间灯光强度直接反映城市在夜间的活跃程度,但不少研究表明,其与地区生产总值(GDP)、人口密度及城市化水平具有显著相关关系[29-30],并被广泛运用于间接衡量城市综合经济活跃度.相比于 GDP 等传统的统计指标,夜间灯光数据具备更高的时空分辨率和时序一致性,尤其适用于缺乏细粒度经济数据的多时间尺度污染建模研究[31].因此,综合考虑指标可信度和数据可得性,本文采用夜间灯光作为经济活动强度的代理变量.

1.2.2　模型参数调优本文基于 Anaconda 平台进行模型参数调优.考虑到算法的计算效率,仅针对模型性能影响最为显著的 4 个参数进行调优.决策树数量对模型性能的影响最为突出,且呈现单调增加的趋势[32-33].参考相关文献中的做法,在研究过程中,以 1000 作为基准进行调整,当决策树数量为 800 时,误差逐渐趋于稳定.此外,对决策树的最大深度进行调优,采用较小的深度能够有效避免模型出现过拟合现象;对叶节点最小样本数进行调优,设置较大的值可防止模型过度复杂;最后,对节点分裂最小样本数进行调优.为确定各参数的最优取值,本文采用了十折交叉验证与网格搜索相结合的调优方法.

1.2.3　变量重要性评估 IncMSE(Increase in Mean Squared Error)是随机森林算法中常用的变量重要性度量方法.当对某一变量进行随机扰动后,如果模

型预测误差显著增加,则说明该变量对模型的预测效果贡献较大.具体计算方式如下:

$$Inc=MSE_{pem}-MSE_{orig} \quad (1)$$

式中:MSE_{orig} 为原始模型在验证集上的误差,MSE_{perm} 为打乱变量后重新预测的误差.该过程通常在随机森林的每棵决策树上分别计算,最终求平均值并估计标准误差,以提供稳定性评估.IncMSE 值越高,说明该变量在模型中的预测作用越大;反之,若接近于 0 或为负值,则说明其贡献较低甚至可能冗余[34-36].

1.3 空间插值方法

为了分析 O_3 浓度的空间分布特征,本文采用反距离加权插值法(IDW)对“2+26”城市群进行空间插值.IDW 不依赖于空间变异函数的建模,适用于采样点有限或分布不均的情形.尽管其不提供置信区间估计,但在污染物浓度、气象变量的可视化估算中仍具良好适用性[37].具体而言,IDW 假设空间位置越接近,其属性值越相似,因而通过距离的倒数对观测值加权.对某一待估位置 S_0,其插值结果为:

$$\hat{y}(S_0)=\sum_{i=1}^{n}\lambda_i y(S_i) \quad (2)$$

式中:λ_i 是位置 S_i 的权重,计算方式如下:

$$\lambda_i=\frac{d_{0i}^{-\alpha}}{\sum_{j=1}^{n}d_{0j}^{-\alpha}} \quad (3)$$

$$\sum_{i=1}^{n}\lambda_i=1 \quad (4)$$

式中:d_{0i} 表示 S_0 与 S_i 之间的欧氏距离;α 为距离衰减系数.当 α 较小时,权重趋于平均值;当 α 增大时,权重更倾向于距离较近的样本点,远点的权重迅速减小.

1.4 数据来源

本文采用的 O_3 浓度数据来源于 NCDC 的公开 FTP 服务器 (ftp://ftp.ncdc.noaa.gov/pub/data/noaa/isd-lite/),选取了“2+26”城市区域 2014~2022 年的观测站点数据;逐小时的站点气象数据,包括压强、气温、比湿、24h 降水量和水平风速,均取自中国气象数据中心的历史记录 (https://q-weather.info/weather/).这些数据涵盖了共 275 个气象监测站点的观测结果,为研究提供了详实的气象背景资料.受数据可得性约束,本文仅考虑水平风速对 O_3 浓度的影响.夜间灯光数据来自 NASA 数据库,该数据库将全球根据经纬度进行划分,以 10°×10°的网格进行分割,每个分块包含 2400×2400 个格点,分辨率约为 500m×500m,数据均为 HDF5 格式.通过将栅格与行政区进行匹配,最终得到市域数据.

为确保研究结果的准确性,本文严格遵循《环境空气质量标准》[38]对站点原始数据进行了严格的质量控制.为避免模型因输入变量不全而失真,排除长期缺测或数据波动异常的站点,同时剔除了原始数据缺失或异常的观测记录.其中,异常站点或观测记录的识别遵循 1.5 倍四分位距法则.最终筛选出“2+26”区域城市的 208,096 条有效观测数据.

2 结果与分析

2.1 O_3 浓度时空变化特征

本文采用反距离加权插值法对“2+26”城市群 O_3 浓度进行空间插值,得到 2014~2022 年间 O_3 浓度的逐年分布变化特征,如图 2 所示.从年度变化趋势看,2014~2018 年间,随着经济社会发展以及工业活动日益频繁,O_3 浓度呈现逐年上升态势.然而,2018 年后,随着一系列大气污染防治措施的有效实施,重点城市 O_3 浓度增长速度开始放缓,部分地区的 O_3 污染状况得到一定程度的改善.但 2022 年 O_3 浓度有所反弹,这主要与当年高温日数偏多,静稳天气频发以及后疫情时期经济活动快速恢复所致的前体物排放增加密切相关,凸显出区域 O_3 污染防控形势的复杂性与反复性.

从空间分布看,O_3 污染浓度表现出南高北低的地域特征.河北省的石家庄、唐山等地区是 O_3 污染的高值区,其日最大 8h 平均浓度超过 73.2μg/m^3,且随时间推移,高值区逐渐向南部扩展.这种分布与地理条件和排放特征密切相关.河北地处华北平原,北部被山脉环绕,南部为开阔的平原地形.这种地形使得污染物在平原地区容易积聚,尤其是在静风天气条件下,污染物难以扩散.此外,河北省作为我国重要的钢铁生产基地,石家庄、唐山等城市的钢铁冶炼产业规模庞大,生产过程中排放出的大量的氮氧化物,极大地促进了 O_3 的生成.同时,机动车保有量的增加和频繁行驶所排放的尾气,也进一步加重了该地区的 O_3 污染负担.相比之下,山西长治、晋城及河南鹤壁等中小城市,其 O_3 浓度均不足 62.8μg/m^3.

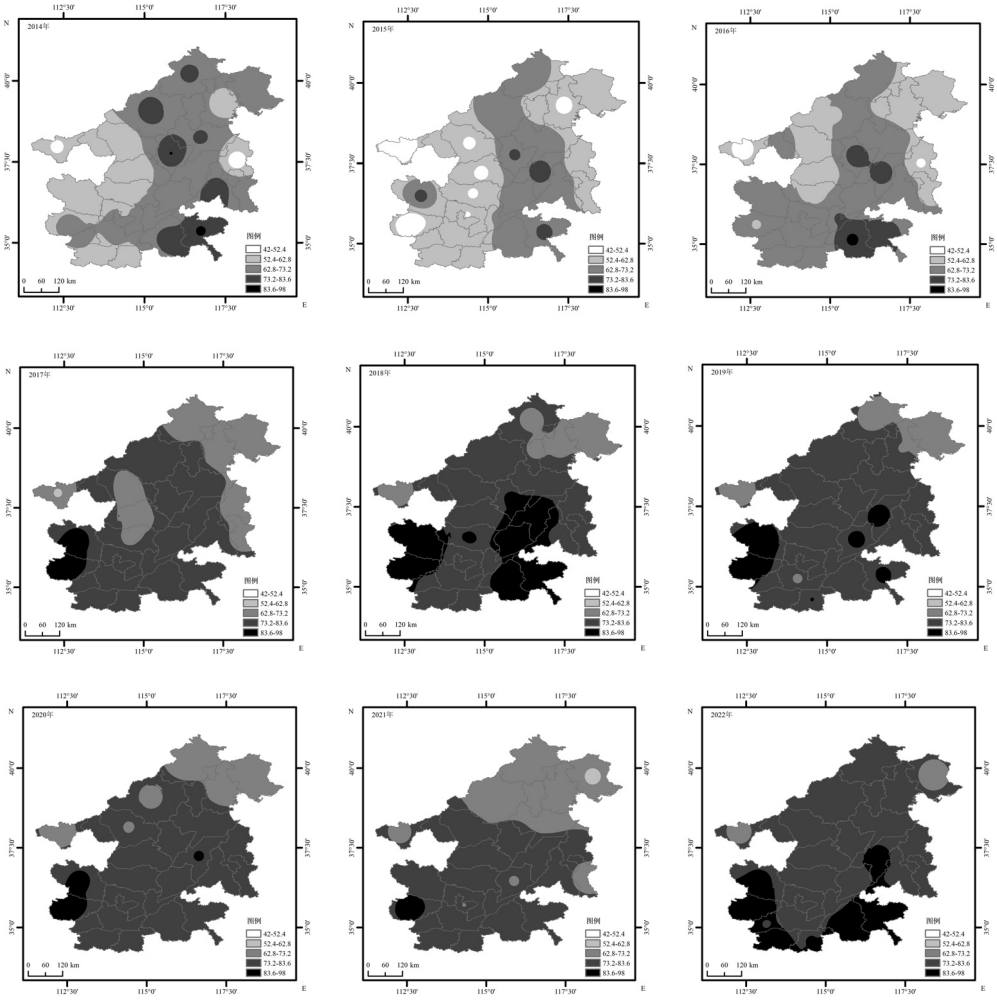


图 2　2014~2022 年 O_3 时空分布特征

Fig.2　Temporal and spatial distribution characteristics of O_3 from 2014 to 2022

审图号:GS(2024)0650 号

2.2　模型验证

2.2.1　拟合度分析拟合度分析可以评估现有模型与实际观测值之间的匹配程度.本文通过线性回归检验,比较模型拟合结果与实际数据的吻合性来评估模型的表现.从图 3(a)的残差结果显示,残差呈现单峰和钟形曲线,符合正态分布,表明模型的残差主要集中在较小范围内,模型拟合效果较好;从图 3(b)可以看出,预测值与实际值之间存在紧密的线性关系,大部分数据点集中在回归线附近,且回归线的斜率接近于 1,说明预测值与实际值之间的比例关系较为准确.

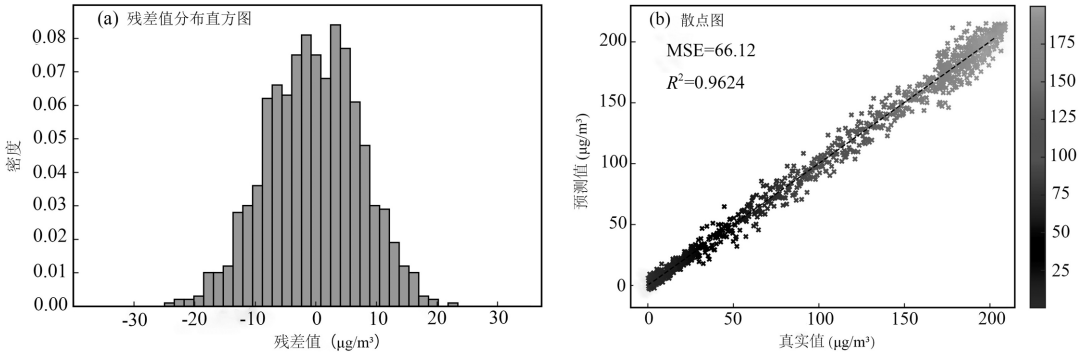


图 3　随机森林模型拟合度分析

Fig.3　Fitting analysis of the random forest model

2.2.2 模型预测能力评估为评估随机森林模型对O_3预测能力,以2014~2021年O_3逐日和逐年数据作为训练集,通过6个影响因子对2022年O_3浓度进行预测,得到预测值和观测值对比,结果如图4所示.其中数值点为样本索引,基于数据质量和可视化需求,分别从日度与年度样本中随机抽取有效样本进行展示.从图 4(a)的实际值与预测值时间序列图来看,日时间尺度下,随机森林对 O_3 浓度的预测较为准确,预测值与实际值高度重合,这与较高的 R^2 值一致,表明模型能够很好地解释 O_3 浓度的变化.同时,MSE值较低,说明模型的预测误差较小.在图4(b)的年度模型下,虽然 R^2值略低于日度,但仍然表明模型有很强的解释能力.预测值与实际值的波动趋势基本一致,但有部分预测值偏离实际值,特别是在峰值处,MSE 值更低,表明该模型的整体预测误差更小.

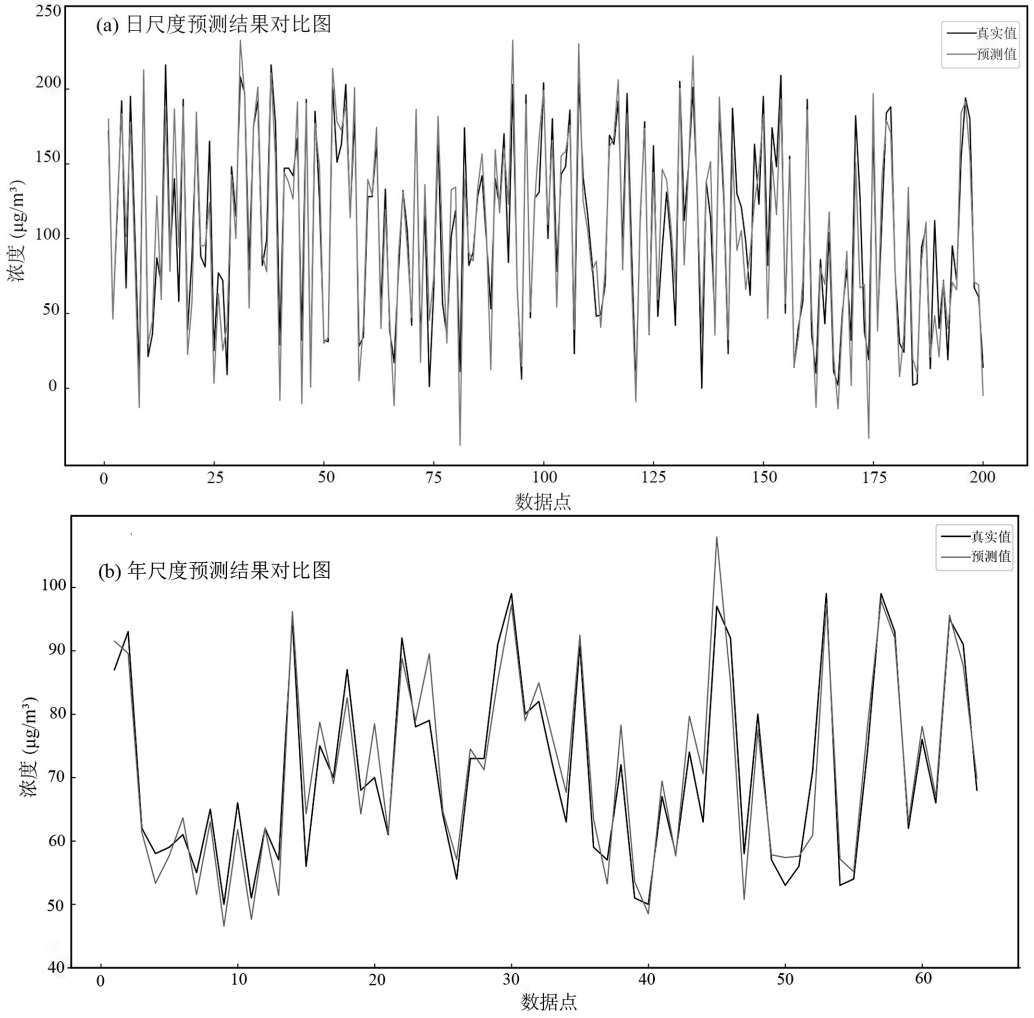


图4 O_3浓度的随机森林预测结果年度分析

Fig.4 Annual analysis of prediction results for O_3 concentration using the random forest model

2.3 不同时间尺度模型的因子重要性评估

在应对日益严重的污染问题时,不同时间尺度的研究具有重要意义.日度尺度研究能够揭示短期内污染物的波动以及气象条件的瞬时影响,而年度尺度则能够反映出污染物的长期趋势和累积效应,有助于评估长期政策的实施效果.例如,在短期内,临时管控措施可以通过限制交通和工业活动有效降低污染物浓度;而从长期角度来看,减排政策和产业结构调整则能逐步减少污染源的总量.因此,综合分析不同时间尺度下的污染物变化,对于制定更加科学有效的污染控制措施至关重要.

2.3.1 整体结果分析重要性评估是随机森林中识别和理解关键变量的重要依据.本文以 O_3 浓度为数据基础,采用随机森林模型评估气象因素与经济因

素在日度和年度不同时间尺度下的影响力.具体来看,根据前期对照实验结果,加入经济因素后,日度模型的均方误差从 138.9959 降至 135.9101,尽管 R^2 从 0.9379 上升至 0.9387,提升幅度较小,但 MSE 的下降表明模型预测误差有所改善;年度模型中,均方误差从 13.2841 显著降低至 3.7939,R^2 从 0.7347 提升至 0.9407.年度模型预测精度的提升主要得益于数据汇总效应所带来的短期波动平滑,显著降低了噪声干扰,提高了模型拟合效果.然而,进一步的变量重要性评估显示,在年度尺度下,夜间灯光的重要性评分提升至 0.7053,表明夜间灯光作为经济活动代理变量,在年尺度下的贡献显著提升,表明在长期尺度上,经济因素的作用更为稳定和显著,具备结构性解释能力.

表 1　不同时间尺度下 O_3 预测模型的拟合优度指标

Table 1　Goodness-of-fit metrics (MSE and R^2) for O_3 prediction models at different temporal scales

指数	O_3(日度)	O_3(年度)
MSE	135.9101	3.7939
R^2	0.9387	0.9407

如图 5 所示,在日时间尺度下,气温是最重要的因素,其重要性数值约为 0.5759,对 O_3 浓度的预测贡献最大.较高的气温会加速光化学反应的速率,使得空气中的氧气分子与 NO_x 和 VOCS 更活跃地反应,从而生成更多的 O_3[39-41].气压对 O_3 浓度日度变化的驱动贡献同样较为显著.高气压系统往往对应静稳天气,边界层高度低,垂直混合弱,污染物易在地表区域积聚,进而抬升 O_3 浓度;而低气压系统则常伴随对流增强,有利于污染物的稀释与输送[42-43].相比之下,湿度,风速对 O_3 浓度的影响在日度情况下相对较小.湿度通过影响光化学反应过程间接影响 O_3 生成,较高的湿度增加水蒸气含量,部分吸收紫外线,降低光化学反应效率,但其对 O_3 浓度的影响主要体现在更长时间尺度上,通过作用于气溶胶和大气化学过程,日时间尺度的湿度变化难以显著改变 O_3 浓度.水平风速通过影响 O_3 扩散和稀释从而改变其局地浓度.在日度尺度上,风速变化时间较短,其对 O_3 浓度的影响主要局限于局地扩散层面,对其生成速率影响有限.城市环境中,局地排放源主导 O_3 的生成和消耗过程[44],因而日度变化中风速的作用常被局地排放的变化所掩盖.

日度模型中,夜间灯光(LIGH)的重要性为 0.0869,位列第五,表明其对 O_3 预测有一定影响,但贡献相对较小.然而,在年度模型中,夜间灯光的重要性跃升至第一位,表明相较于其他因素,经济活动在长期尺度下对 O_3 浓度的影响更显著.频繁的现代人类活动在提高地区夜间灯光的同时带来了更大规模的工业生产,交通运输和终端能源消耗,从而导致更多的 NO_x 和 VOCs 的排放.作为 O_3 的前体物质,NO_x 和 VOCs 的增加直接影响地面 O_3 的生成.人类活动通过增加 O_3 前体物质的排放量,间接促进 O_3 浓度的提升[45].这种效应在年度时间尺度上逐渐累积,导致夜间灯光在长期模型中对 O_3 浓度产生显著影响.年度模型中,气温同样是较为重要的因素之一.在较长的时间尺度上,气温变化会持续增加 O_3 浓度,因此,气温的年度变化对 O_3 的整体浓度水平有重要贡献.此外,风速与降水亦在模型中展现出较强的解释力.适宜的风速有助于 O_3 及其前体物的扩散,而持续低风速则可能导致污染物滞留,提升局地 O_3 浓度.降水则主要通过清除气溶胶和水溶性前体物间接影响光化学反应,从而抑制 O_3 浓度提升.

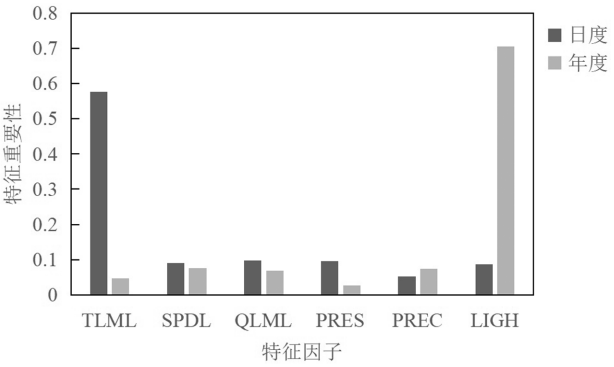


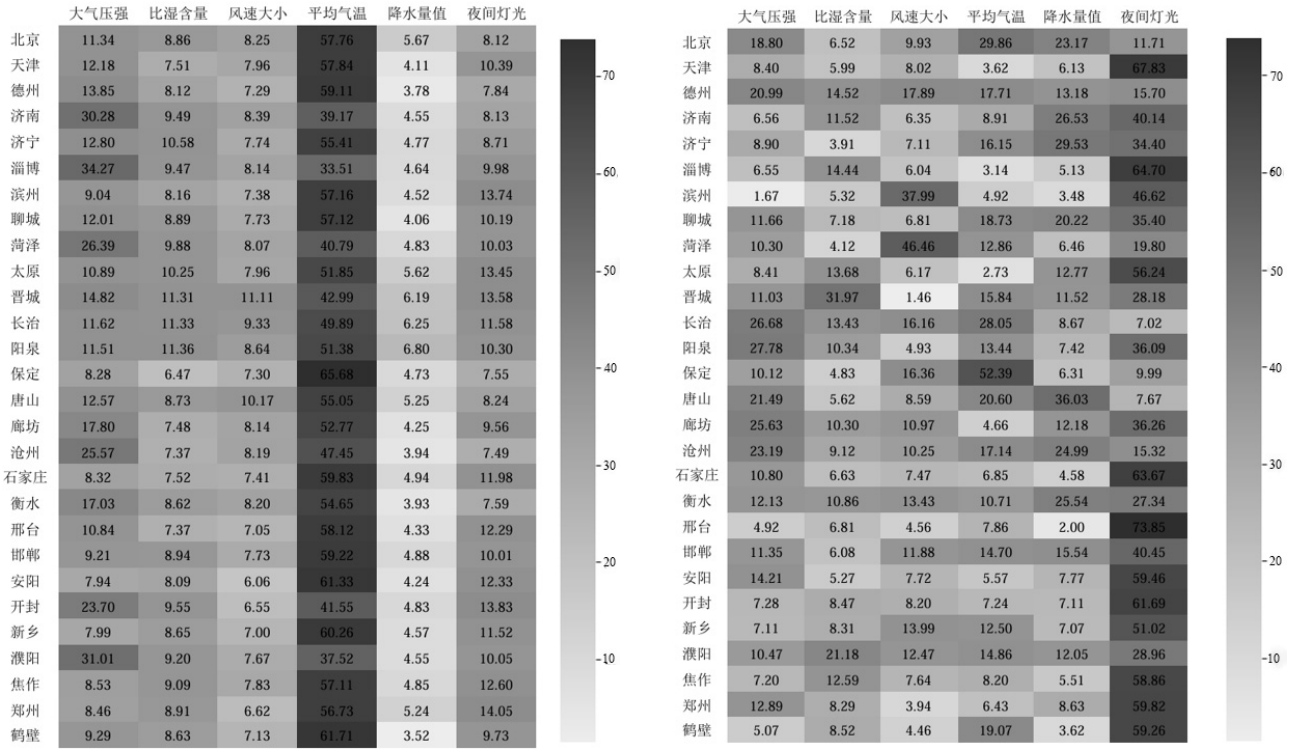
图 5　不同时间尺度下预测 O_3 浓度的特征重要性

Fig.5　Feature importance in predicting O_3 concentration at different time scales

2.3.2　城市尺度结果分析由于各个城市的气象条件、地理位置和环流背景存在差异,因此影响 O_3 浓度的主导因子也有所不同.如图 6 所示,从整体来看,气温、气压和夜间灯光是影响 O_3 浓度变化的三个主要因素.在日度数据中,有 13 个城市的气温重要性评分超过 50,如北京、天津和德州等城市.高温有利

于光化学反应的进行,从而增加 O_3 浓度;气压对部分城市有显著影响,如济南、德州和沧州,尤其在日度数据中,这些城市的气压评分超过 20,稳定的高气压系统会抑制污染物扩散,导致 O_3 积累.夜间灯光作为经济活动的代理变量,在年度数据中对多个城市的 O_3 浓度具有显著影响,特别是在邢台、淄博和石家庄等城市,夜间灯光的评分超过 50,反映出经济活动对 O_3 长期浓度的显著影响.

结合 O_3 浓度在 2014~2022 年间的情况,将“2+26”城市群的年均 O_3 日最大 8h 滑动平均浓度进行排序后,选取排名前 6 的城市定义为“高值区”,即 O_3 污染水平相对较高区域;选取排名后 6 的城市定义为“低值区”,代表 O_3 水平相对较低的区域,并分析浓度高值区(北京、天津、济南、德州)与低值区(石家庄、太原、长治)之间的变化差异.在高值区中,气温和夜间灯光对 O_3 浓度的影响最显著,这些因素在多个城市的重要性评分接近或超过 50.高值区城市普遍表现出较强的高温条件,有利于 O_3 的生成和累积,进一步加剧 O_3 浓度的升高.低值区中,主导驱动因子较为多样化,具有明显的区域性和不连续性特点.其中,石家庄主要受到气压和风速的影响,太原主要受到气温和湿度的影响,而长治则主要受到人类活动强度的影响.



(a)日度 (b)年度

图 6 “2+26”城市群不同因子的重要性贡献度

Fig.6 Importance contribution of the "2+26" urban agglomeration at daily and annual scale

3 讨论

$PM_{2.5}$ 作为另一关键大气污染物,其浓度同样受到多种因素的综合影响[46-50].但由于 $PM_{2.5}$ 形成的物理化学过程与 O_3 不同,两类污染物受人类经济活动及自然气象因子的影响呈现出不同的特征.综合分析不同因子对两类污染物的驱动贡献异同有助于制定具有针对性的大气污染协同防控策略.因此,本节在前文分析的基础上,进一步采用随机森林模型对 $PM_{2.5}$ 的驱动因子进行分析,并同样以夜间灯光强度作为经济活动强度的代理指标,捕捉人类经济活动对区域 $PM_{2.5}$ 和 O_3 浓度的影响,结果如表 2 和图 7 所示.

由模型结果可知,对于日度模型,经济活动强度对 $PM_{2.5}$ 的影响较 O_3 更为突出,贡献因子仅次于气温,反映出城市经济活动对 $PM_{2.5}$ 浓度的短期提升具有直接贡献.这是由于 O_3 作为一种典型的光化学二次污染物,其生成依赖于 NO_x 和 VOCs 在太阳辐射

条件下发生的链式反应[51-52],这一过程相较于 $PM_{2.5}$ 更加依赖气温条件.相比之下,$PM_{2.5}$ 的生成既包括一次排放,也包含前体物的物理化学转化,受人类活动强度的影响更直接.在气象因子中,气温同时是日时间尺度下 $PM_{2.5}$ 和 O_3 浓度的首要贡献因子,但其对 O_3 浓度的贡献因子(0.5759)超过 $PM_{2.5}$(0.2807)的 2 倍;相比之下,气压、比湿、风速、降水等其他气象因子对 $PM_{2.5}$ 浓度的影响均高于 O_3.对于年度模型,夜间灯光同时是 $PM_{2.5}$ 和 O_3 浓度的首要贡献因子,这意味着在更长时间尺度下,人类经济活动同时成为两类主要大气污染物浓度的决定性因素.与日度模型类似,经济活动强度对 $PM_{2.5}$ 浓度的影响同样大于 O_3,但二者差距明显缩小.尽管 O_3 的生成主要依赖光化学反应而非人类经济活动的直接排放,但从长时间尺度看,前体物排放长期累积引发的 O_3 浓度间接上升问题依旧不可小觑.在年时间尺度上,气象因子对两类大气污染物的相对贡献均有所下降.具体而言,风速和气温对城市 $PM_{2.5}$ 浓度的年度变化影响更大,而气压、比湿和降水对 O_3 浓度的贡献更为明显.

表 2　不同时间尺度下 $PM_{2.5}$ 预测模型的拟合优度指标

Table 2　Goodness-of-fit metrics (MSE and R^2) for $PM_{2.5}$ prediction models at different temporal scales

指标	$PM_{2.5}$(日度)	$PM_{2.5}$(年度)
MSE	249.6123	12.7139
R^2	0.8957	0.9419

上述比较分析的结果表明,$PM_{2.5}$ 和 O_3 浓度在长时间尺度均与人类经济活动高度相关,对二者的管控均依赖于长期的结构性调整和技术性变革,应从源头着手,从产业布局、交通布局、国土空间规划布局层面加强对 $PM_{2.5}$ 和 O_3 浓度的协同管控.而限产、停产等对经济活动的临时管控措施虽然可以有效控制 $PM_{2.5}$ 浓度,但对 O_3 的短期贡献较小.这意味着过去在一些重大活动前通过限制部分经济活动有效控制局地 $PM_{2.5}$ 浓度的经验性举措可能对 O_3 而言难以发挥同样的效果,因此对于 O_3 的管控尤为需要从顶层设计的战略性视角出发,制定长远规划,施行计划性,长期性调控策略.

需要指出的是,受篇幅和数据可得性限制,本文预测模型未能纳入 O_3 生成的光化学反应路径中较为关键的 VOCs 组分因素,限制了对人类经济活动的污染生成机制以及对 $PM_{2.5}$ 与 O_3 浓度驱动机制差异的进一步讨论.未来研究可结合 VOCs 高时空分辨率观测数据,进一步增强对模型反应机制和作用机理的解释力.此外,囿于高时空分辨率经济社会统计数据的缺失,本文采取多数文献的做法,仅采用夜间灯光数据反映人类经济活动强度,未来若能突破数据限制,可进一步细分经济活动类型,并将政策变量纳入预测模型,从而进一步提高模型的适应性和解释力.

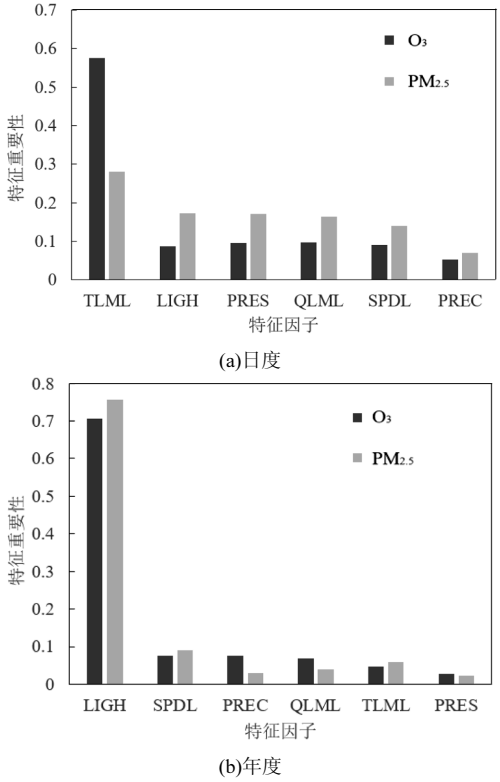


图 7　不同时间尺度下 $PM_{2.5}$ 因子重要性结果

Fig.7　Factor importance results of $PM_{2.5}$ on different scales

4　结论

4.1　以夜间灯光作为代理变量的经济活动强度在不同时间尺度都对 O_3 浓度预测均有显著贡献.在年度模型中,经济活动强度对 O_3 浓度预测的贡献度达到 0.7053,超过各类气象因子成为首要贡献因素;但在日度模型中,以气温为代表的气象因子对 O_3 浓度的影响更大.经济活动是影响 O_3 浓度年度变化的决定性因素,分析和预测其长时间尺度变化时需考虑经济变量的贡献.

4.2 不同因素对 O_3 浓度高值区与低值区的预测贡献存在差异.在北京、天津、济南、德州等 O_3 浓度高值区,气温和经济活动的影响较大;而在低值区,不同因子的预测贡献具有明显的区域性和不连续性特点.其中,石家庄的 O_3 浓度主要受气压和风速影响,气温和湿度对太原 O_3 浓度的影响较大.因此,对于不同类型城市的 O_3 浓度预测和管控,应因地制宜做好关键气象因子的监测预警工作,同时重点加强对 O_3 浓度高值区城市排放源头的监管.

4.3 通过比较 O_3 与 $PM_{2.5}$ 的多尺度随机森林模型预测结果发现,在长时间尺度,人类经济活动对城市 $PM_{2.5}$ 和 O_3 浓度均有决定性贡献;在短时间尺度,经济活动对 $PM_{2.5}$ 浓度的影响仅次于气温,但对 O_3 浓度的驱动贡献较小.从长期来看,通过结构性调整和技术性变革等经济活动加强对 $PM_{2.5}$ 和 O_3 浓度的协同管控;而通过限制部分经济活动有效控制局地 $PM_{2.5}$ 浓度的经验性举措可能对 O_3 而言难以发挥同样的效果.

参考文献:

[1] 鲍捷萌,曹 娟,高 锐,等.欧洲环境空气臭氧污染防治历程,经验及对我国的启示 [J]. 环境科学研究, 2021,34(4):890–901.
Bao J M, Cao J, Gao Ri, et al. Process and experience of ozone pollution prevention and control in Europeand enlightenment to China [J]. Environmental Science Research, 34(4):890–901.

[2] 王 情,朱欢欢,杜 鹏,等.京津冀及周边地区“十四五”及中长期 $PM_{2.5}$ 污染控制目标的健康效益预估研究 [J]. 环境科学研究, 2021,34(1):220–228.
Wang Q, Zhu H H, Du P, et al. Health benefit of “14th Five-Year” and medium & long-term $PM_{2.5}$ control targets in Beijing-Tianjin-Hebei and its surrounding areas [J]. Environmental Science Research, 2021,34(1):220–228.

[3] 姜 华,高 健,李 红,等.我国大气污染协同防控理论框架初探 [J]. 环境科学研究, 2022,35(3):601–610.
Jiang H, Gao J, Li H, et al. Preliminary research on theoretical framework of cooperative control of airPollution in China [J]. Environmental Science Research, 2022,35(3):601–610.

[4] Li C L, Liu Y F, Cheng B F, et al. A comprehensive investigation on volatile organic compounds (VOCs) in 2018 in Beijing, China: characteristics, sources and behaviors in response to O_3 formation [J]. Science of the Total Environment, 2022,806:150247.

[5] Tao T H, Shi Y S, Gilbert K M, et al. Spatiotemporal variations of air pollutants based on ground observation and emission sources over 19 Chinese urban agglomerations during 2015~2019 [J]. Scientific Reports, 2022,12(1):4293.

[6] 黄继章,高 健,余美芳,等.广州市 2022 年臭氧污染特征与成因分析 [J]. 中国环境科学, 44(6):3100–3110.
Huang J Z, Gao J, Yu M F, et al. Analysis of characteristics and cause of ozone pollution in Guangzhou in 2022 [J]. China Environmental Science, 2024,44(6):3100–3110.

[7] Gao J, Zhang J, Li H, et al. Comparative study of volatile organic compounds in ambient air using observed mixing ratios and initial mixing ratios taking chemical loss into account – A case study in a typical urban area in Beijing [J]. Science of the Total Environment, 2018,628–629:791–804.

[8] Zhang Y J, Zhao Y C, Li J, et al. Modeling ozone source apportionment and performing sensitivity analysis in summer on the North China plain [J]. Atmosphere, 2020,11(9):992.

[9] 张 新.中国三大城市群对流层臭氧时空变化特征及影响因素研究 [D]. 北京:中国气象科学研究院, 2023.
Zhang X. Research on the spatiotemporal variation characteristics and influencing factors of tropospheric ozone in the three major urban agglomerations of China [D] Beijing: Chinese Academy of Meteorological Sciences, 2023.

[10] 米艺华,杜惠云,李 杰,等.我国重点地区臭氧变化特征及与气象要素的关系 [J]. 中国环境科学, 2024,44(4):1894–1904.
Mi Y H, Du H Y, Li J, et al. Ozone pollution characteristics and its relationship with meteorological factors of key regions in China [J]. China Environmental Science,2024,44(4):1894–1904.

[11] Dang R J, Liao H, Fu Y. Quantifying the anthropogenic and meteorological influences on summertime surface ozone in China over 2012~2017 [J]. Science of the Total Environment, 2021,754:142394.

[12] 薛 鑫,陈 镇,邬双双,等.气象条件对珠三角秋季地表臭氧年变化及趋势的影响 [J]. 环境科学, 2024,45(7):3808–3814.

[13] Xue X, Chen Z, Wu S S, et al. Meteorological influences on annual variation and trend of autumn surface ozone inthe Pearl River Delta [J]. Environmental Science, 2024,45(7):3808–3814.

[14] 洪声杨,任志彬,王诚聪,等.2019 年中国城市大气臭氧污染时空格局及其影响因素研究 [J]. 地理科学, 2024,44(10):1706–1717.
Hong S Y, Ren Z B, WangC C et al. Spatial and temporal patterns of ozone pollution and its influencing factors in China in 2019. Scientia Geographica Sinica, 2024,44(10):1706–1717.

[15] Weng X, Forster G L, Nowack P. A machine learning approach to quantify meteorological drivers of ozone pollution in China from 2015 to 2019 [J]. Atmospheric Chemistry and Physics, 2022,22(12):8385–8402.

[16] 张容硕,谢沛远,陈宏飞,等.基于机器学习的郑州市大气 $PM_{2.5}$ 与 O_3 浓度预测方法及气象因子的影响分析 [J]. 环境科学研究, 2024, 37(3):469–478.
Zhang R S Xie P Y, Chen H F, et al. Prediction method and meteorological factors impact analysis of atmospheric $PM_{2.5}$ and O_3 concentration in Zhengzhou City based on machine learning [J]. Environmental Science Research, 2024,37(3):469–478.

[17] Balamurugan V, Balamurugan V, Chen J. Importance of ozone precursors information in modelling urban surface ozone variability using machine learning algorithm [J]. Scientific Reports, 2022,12(1).

[18] 齐艳杰,于世杰,杨 健,等.河南省臭氧污染特征与气象因子影响分析 [J]. 环境科学, 2020,41(2):587–599.
Qi Y J, Yu S J, Yang J, et al. Analysis of characteristics and meteorological influence factors of ozone pollution in Henan Province

[J]. Environmental Science, 2020,41(2):587−599.

[19] 钱　悦,许　彬,夏玲君,等.2016~2019 年江西省臭氧污染特征与气象因子影响分析 [J]. 环境科学, 2021,42(5):2190–2201.
Qian Y, Xu B, Xia L J, et al. Characteristics of ozone pollution and relationships with meteorological factorsin Jiangxi Province [J]. Environmental Science, 2021,42(5):2190−2201.

[20] 符传博,徐文帅,丹　利,等.2015~2020 年海南省臭氧时空变化及其成因分析 [J]. 环境科学, 2022,43(2):675–685.
Fu C B, Xu W S, Dan L, et al. Temporal and spatial variations in ozone and its causes over Hainan Province from 2015 to 2020 [J]. Environmental Science, 2022,43(2):675−685.

[21] 马润美,张亚一,班　婕,等.基于随机森林模型的京津冀地区近地面臭氧站点浓度预测 [J]. 环境与健康杂志, 2019,36(11):954–957.
Ma R M, Zhang Y Y, Ban J, et al. Prediction of ground ambient ozone site concentration in Beijing−Tianjin−Hebei region based on random forest model [J]. Journal of Environment and Health, 2019,36(11): 954−957.

[22] 宋龙娟,胡睿琪,刘　兵,等.基于随机森林的川东北城市群臭氧污染预报研究 [J]. 中国测试, 2024,50(11):150−159.
Song L J, Hu R Q, Liu B, et al. Ozone pollution forecasting in northeast Sichuan City Cluster based on random forest [J]. Chinese Journal of Testing, 2024,50(11):150−159.

[23] 李　慧,王淑兰,张文杰,等.京津冀及周边地区“2+26”城市空气质量特征及其影响因素 [J]. 环境科学研究, 2021,34(1):172−184.
Li H, Wang S L, Zhang W J, et al. Air quality characteristics and influencing factors of the "2+26" cities in the Beijing−Tianjin−Hebei Region and its surrounding areas [J]. Environmental Science Research, 2021,34(1):172−184.

[24] 生态环境部.2022 年中国生态环境状况公报 [R]. 北京:生态环境部, 2023.
Ministry of Ecology and Environment. 2022 China ecological and environmental status bulletin [R]. Beijing: Ministry of Ecology and Environment, 2023.

[25] Breiman L. Bagging predictors [J]. Machine Learning, 1996,24(2): 123–140.

[26] 康　平,张梓桓,王天琦,等.南亚高压对四川盆地夏季区域持续性 O_3 污染的影响 [J]. 中国环境科学, 2022,42(12):5484–5496.
Kang P, Zhang Z H, Wang T Q, et al. Effects of South Asian high pressure on regional persistent ozone pollution in Sichuan Basin during summer [J]. China Environmental Science, 2022,42(12):5484−5496.

[27] 刘贤赵,张国桥,杨文涛,等.长株潭城市群 $PM_{2.5}$ 和 O_3 浓度时空分布特征及影响因素分析 [J]. 环境科学, 2022,43(12):5354−5366.
Liu X Z, Zhang G Q, Yang W T, et al. Analysis of spatiotemporal distribution characteristics and influencing factors of $PM_{2.5}$ and O_3 in Changsha−Zhuzhou−Xiangtan urban agglomeration, China [J]. Environmental Science, 2022,43(12):5354−5366.

[28] 王桂林,张　炜.中国城市扩张及空间特征变化对 $PM_{2.5}$ 污染的影响 [J]. 环境科学, 2019,40(8):3447−3456.
Wang G L, Zhang W. Effects of urban expansion and changes in urban characteristics on $PM_{2.5}$ pollution in China [J]. Environmental Science, 2019,40(8):3447−3456.

[29] Liu H, He X, Bai Y, et al. Nightlight as a proxy of economic indicators: Fine−grained GDP inference around Chinese mainland via attention−augmented CNN from daytime satellite imagery [J]. Remote Sensing, 2021,13:2067.

[30] Henderson J V, Storeygard A, Weil D N. Measuring economic growth from outer space [J]. American Economic Review, 2012,102(2):994–1028.

[31] Bickenbach F, Bode E, Nunnenkamp P, et al. Night lights and regional GDP [J]. Review of World Economics, 2016,152:425–447.

[32] Zhai G, Zhang M, Hu Y, et al. Nighttime light metrics for analysing urban−rural economic disparities: A case study in 36 Chinese metropolitan areas [J]. Cities, 2025,162:105930.

[33] Oshiro T M, Perez P S, Baranauskas J A. How many trees in a random forest? [C]. Lecture Notes in Computer Science, Springer, 2012.

[34] Probst P, Boulesteix A L. Hyperparameters and tuning strategies for random forest [R]. Journal of Machine Learning Research, 2017.

[35] Lu G Y, Wong D W. An adaptive inverse−distance weighting spatial interpolation technique [J]. Computers & Geosciences, 2008,34(9): 1044–1055.

[36] Zhang X, Li L, Wang Y. A machine learning approach to ozone concentration prediction and feature importance analysis in urban areas of China [J]. Environmental Pollution, 2023,320:121045.

[37] Li J, Wang M, Liu Z, et al. Understanding the spatiotemporal variability of surface ozone in the North China Plain using Random Forests [J]. Science of The Total Environment, 2024,902:166150.

[38] Kumar R, Ghosh S, Singh A. Comparative evaluation of feature importance in air quality prediction using ensemble machine learning models [J]. Atmospheric Environment, 2023,298:119701.

[39] GB 3095−2012　环境空气质量标准 [S].
GB 3095-2012　Ambient air quality standard [S].

[40] 王　玫,郑有飞,柳艳菊,等.京津冀臭氧变化特征及与气象要素的关系 [J]. 中国环境科学, 2019,39(7):2689–2698.
Wang M, Zheng Y F, Liu Y Y, et al. Characteristics of ozone and its relationship with meteorological factors in Beijing−Tianjin−Hebei Region [J]. China Environmental Science, 2019,39(7):2689−2698.

[41] Coates J, Marl A, Ojha N, et al. The influence of temperature on ozone production under varying NO_x conditions – a modeling study [J]. Atmospheric Chemistry and Physics, 2016,16(16).

[42] 董　瑾,崔荣国,程立海,等.基于随机森林方法的京津冀地区臭氧精细时空预测 [J]. 时空信息学报, 2024,31(3):348−358.
Dong J, Cui R G, Cheng L H, et al. Fine spatiotemporal prediction of ozone in the Beijing−Tianjin−Hebei based on random forest. Journal of Spatio−temporal Information, 2024,31(3):348−358.

[43] Liao Z, Gao M, Zhang J, et al. Mixing−layer−height−referenced ozone vertical distribution in the lower troposphere of Chinese megacities: stratification, classification, and meteorological and photochemical mechanisms [J]. Atmospheric Chemistry and Physics, 2024,24:3541–3557.

[44] Liu L, Zhang X, Wang J, et al. Changes in the height of the pollution boundary layer and their meteorological effects on the distribution of surface ozone concentrations [J]. Frontiers in Environmental Science, 2022,10:1094404.

[45] Zheng X, Yang X, Gu H, et al. Characterization and sources of volatile organic compounds (VOCs) during 2022 summer ozone pollution

control in Shanghai, China [J]. Atmospheric Environment, 2024,327: 120464.

[46] 张运江,雷若媛,崔世杰,等.2015~2020 年我国主要城市 $PM_{2.5}$ 和 O_3 污染时空变化趋势和影响因素 [J]. 科学通报, 2022,67(18):2029–2042.
Zhang Y J, Lei R Y, Cui S J, et al.Spatiotemporal trends and impact factors of $PM_{2.5}$ and O_3 pollution in major cities in China during 2015~2020 [J]. Chinese Science Bulletin, 2022,67(18):2029−2042.

[47] 宋小涵,燕 丽,刘 伟,等.2015~2021 年京津冀及周边地区 $PM_{2.5}$ 和臭氧复合污染时空特征分析 [J]. 环境科学, 2023,44(4):1841–1851.
Song X H, Yan L, Liu W, et al. Spatiotemporal distribution characteristics of co−pollution of $PM_{2.5}$ and ozone over BTH with surrounding area from 2015 to 2021 [J]. Environmental Science, 2023,44(4):1841−1851.

[48] 肖巧俐,王 跃.基于夜间灯光数据的区域经济收敛性研究 [J]. 统计与决策, 2024,40(12):108–112.
Xiao Q L, Wang Y. Research on regional economic convergence based on nighttime light data [J]. Statistics and Decision, 2024,40(12): 108−112.

[49] 吴舒祺,金团囡,郑冬阳,等. “2+26”城市 $PM_{2.5}$ 与气象因子的尺度依存关系及影响因素分析 [J]. 环境科学, 2023,44(12):6441–6451.
Wu S Q, Jin J N, Zheng D Y, et al. Scale dependence between $PM_{2.5}$ and Meteorological Factors and its Influencing factors in "2+26" Cities [J]. Environmental Science, 2023,44(12):6441−6451.

[50] Zhou L H, Zhang J, Lu T W, et al. Pollution patterns and their meteorological analysis all over China [J]. Atmospheric Environment, 2021,246:118108.

[51] 花 丛,江 琪,迟茜元,等.我国中东部地区 2015~2020 年夏半年 $PM_{2.5}$ 和臭氧复合污染气象特征分析 [J]. 环境科学研究, 2022,35(3):650–658.
Hua Dan, Jiang Qi, Chi Xiyuan, et al. Meteorological characteristics of $PM_{2.5}$–O_3 air combined pollution in Central and Eastern China in the summer half years of 2015~2020 [J]. Environmental Science Research, 2022,35(3):650−658.

[52] Hao Y, Zhou J, Zhou J P, et al. Measuring and modeling investigation of the net photochemical ozone production rate via an improved dual−channel reaction chamber technique [J]. Atmospheric Chemistry and Physics, 2023,23:9891–9910.

[53] 邢怡然,牛月圆,闫雨龙,等.典型工业城市长治市夏季大气挥发性有机物来源及对臭氧生成贡献 [J]. 中国环境科学, 2024,44(8):4202–4210.
Xing Y R, Niu Y Y, Yan Y L, et al. Research on the sources of volatile organic compounds in the summer atmosphere and the contribution to ozone generationin Changzhi, a typical industrial city [J]. China Environmental Science, 2024,44(8):4202−4210.

作者简介：姜明栋(1997−),男,江苏盐城人,讲师,博士,主要从事环境与可持续发展管理研究.发表论文 40 余篇.jmd@seu.edu.cn.