

# 基于PM2.5浓度研究京津冀环境一体化

祝本航, 熊海怡, 陈南, 陈铭城, 栾函渝

中国矿业大学(北京)理学院, 北京

收稿日期: 2025年3月18日; 录用日期: 2025年4月9日; 发布日期: 2025年4月22日

---

## 摘要

本文旨在探究京津冀地区的PM2.5浓度分布及其相互之间的影响, 通过分析京津冀地区供暖季与非供暖季的差异, 探索京津冀地区PM2.5浓度的空间相关性。首先, 使用十三个市的勘测点数据计算供暖季PM2.5浓度的全局莫兰指数和局部莫兰指数, 以分析各地区间的相关性, 结论为京津冀地区PM2.5浓度具有较高的相关性。接着, 在京津冀地区对供暖季和非供暖季的PM2.5浓度进行克里金插值, 识别浓度变化显著的区域, 并进行PM2.5浓度差异原因分析以及高污染区域成因分析, 探究京津冀PM2.5浓度相互之间的影响, 有针对性地提出了相应环境保护建议。最后, 由于勘测点的不均衡性可能导致结果偏差, 本文依据克里金插值结果在京津冀地区扩大勘测点位重新计算各市供暖季的PM2.5平均值, 并对比前后的莫兰指数, 得出的结论为京津冀13个市PM2.5浓度相关性较强, 污染呈现明显聚类, 北部地区呈现“低-低”相聚, 南部地区呈现“高-高”相聚, 并且供暖季污染面以城市为中心规律扩散。

---

## 关键词

全局莫兰指数, 局部莫兰指数, 克里金插值, 空间相关性

---

# Investigating Environmental Integration in the Beijing-Tianjin-Hebei Region Based on PM2.5 Concentration

Benhang Zhu, Haiyi Xiong, Nan Chen, Mingcheng Chen, Hanyu Luan

School of Science, China University of Mining and Technology, Beijing

Received: Mar. 18<sup>th</sup>, 2025; accepted: Apr. 9<sup>th</sup>, 2025; published: Apr. 22<sup>nd</sup>, 2025

---

## Abstract

This study aims to investigate the spatial distribution of PM2.5 concentrations and their inter-regional influences in the Beijing-Tianjin-Hebei (BTH) region, with a focus on analyzing seasonal

文章引用: 祝本航, 熊海怡, 陈南, 陈铭城, 栾函渝. 基于PM2.5浓度研究京津冀环境一体化[J]. 统计学与应用, 2025, 14(4): 178-185. DOI: [10.12677/sa.2025.144099](https://doi.org/10.12677/sa.2025.144099)

variations between heating and non-heating periods. The spatial correlation characteristics of PM2.5 concentrations were systematically examined through spatial statistical approaches. Initially, Global Moran's I and Local Moran's I indices were calculated using monitoring data from 13 municipal-level cities during heating seasons, revealing significant spatial autocorrelation of PM2.5 concentrations across the BTH region. Subsequently, Kriging interpolation was implemented to map PM2.5 concentration patterns for both seasons, identifying regions with substantial variation and conducting causal analyses for concentration discrepancies and pollution hotspots. Based on these findings, targeted environmental protection recommendations were proposed. Finally, due to the uneven distribution of monitoring points potentially leading to biased results, this study recalculated the average PM2.5 levels during the heating season for each city in the Beijing-Tianjin-Hebei region by expanding the monitoring points based on Kriging interpolation results. A comparison of Moran's Index before and after the adjustment revealed that the PM2.5 concentrations in the 13 cities of the region exhibited strong spatial correlations, with pollution showing significant clustering. The northern areas displayed a "low-low" aggregation pattern, while the southern areas exhibited a "high-high" aggregation. Additionally, during the heating season, pollution spread outward from urban centers in a regular pattern.

## Keywords

Global Moran's I, Local Moran's I, Kriging Interpolation, Spatial Correlation

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着京津冀协同发展战略的深入推进,区域发展已从单一的经济要素合作逐步演进为涵盖生态环境治理、社会治理与制度创新的系统性协同工程。作为国家重大区域发展战略的实践载体,该区域在产业结构优化、交通网络互通和公共服务均等化等方面取得显著进展,然而面对环境治理这一复杂系统工程,当前协同机制仍呈现多维张力。本文将从PM2.5的角度探究京津冀环境的相关性从而佐证京津冀环境协调治理的必要性。

本文使用了2023年3月15日~2024年3月15日京津冀的环境监测数据2,提取了其中每日二十四小时的PM2.5浓度数据。本文通过京津冀地区的供暖时间将时间划分为两个季度:供暖季(2023年11月15日~2024年3月15日)和非供暖季(2023年3月16日~2023年11月14日)。依据供暖季的PM2.5浓度对京津冀地区做分析,探究京津冀环境的相关性,并且通过克里金插值<sup>[1]</sup>将供暖季与非供暖季做对比分析。对数据预处理剔除不合理数据后,在京津冀供暖季提取了93个有效勘测点的数据,非供暖季提取了116个有效勘测点的数据。其中,天津市有效勘测点最多(15个左右),廊坊、沧州、衡水有效勘测点最少(3,4个)。本文对数据的处理方式是采用平均值,先对勘测点每天二十四小时的PM2.5浓度取平均值得到每天每个勘测点的平均PM2.5浓度;然后再以供暖季、非供暖季为界限对数据取平均得到每个勘测点供暖季、非供暖季的平均PM2.5浓度;最后对供暖季每个市所有勘测点取平均得到供暖季各个市的平均PM2.5浓度(其中邯郸市最高,PM2.5浓度为74.09 μg/m<sup>3</sup>;张家口最低,PM2.5浓度为24.22 μg/m<sup>3</sup>)。

## 2. 莫兰指数

在空间相互作用和空间扩散的双重影响下,各地区PM2.5浓度值不再相互独立,而是具有特定的相关性。本文使用莫兰指数从全局和局部两个角度对京津冀地区PM2.5浓度进行空间相关分析<sup>1</sup>[2]。

<sup>1</sup> 数据来源于网站: <https://quotsoft.net/air/#archive>。

## 2.1. 全局莫兰指数

计算公式为：

$$I = \frac{N}{S_0} \cdot \frac{\sum_{i=1}^N \omega_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

其中， $N$  是空间单元的总数； $\omega_{ij}$  是空间权重矩阵元素； $S_0 = \sum_{i=1}^N \sum_{j=1}^N \omega_{ij}$  是权重矩阵的总和； $x_i$  是第  $i$  个空间单元的属性值； $\bar{x} = \sum_{i=1}^N x_i$  是所有单元属性值的平均值。

鉴于数据的自然地理属性，本文采用了 Queen 邻接方式来定义空间权重(即相邻单元共享边界或定点时，权重  $\omega_{ij}$  赋值为 1) [3]。基于各个市供暖季平均 PM2.5 浓度数据，可得结果如表 1 所示。

**Table 1.** Global Moran's index  
**表 1.** 全局莫兰指数

全局莫兰指数汇总/Summary of Global Moran's Index	
全局莫兰指 (Global Moran's I)	0.542839
预期指数(Expected Index)	-0.083333
方差(Variance)	0.035541
z 得分(z-score)	3.321478
p 值(p-value)	0.000895

由表 1 可知，z 得分为 3.32，远高于显著性水平 0.01 对应的临界值 2.58。原假设为属性值在空间上随机分布，不存在自相关性；而 p 值为 0.000895 ( $<0.01$ )，说明随机产生此聚类模式的概率小于 1%，结果在 99% 置信水平下显著，表明结果是可信的、显著拒绝原假设。全局莫兰指数为 0.543 (显著不为 0)，这一正值说明各市 PM2.5 浓度值存在显著的正空间自相关，即高浓度区域倾向于邻近高浓度区域，低浓度区域邻近低浓度区域，整体呈现空间聚集特征。预期指数为 -0.083，实际值远高于预期，进一步支持空间聚集的结论。

PM2.5 浓度在空间上表现出显著聚集，这种现象可能受多种因素的综合影响，如局部污染源、气象条件以及地形特征等。为进一步识别具体的聚集区域(如热点或冷点)，需结合局部莫兰指数进行深入分析。

## 2.2. 局部莫兰指数

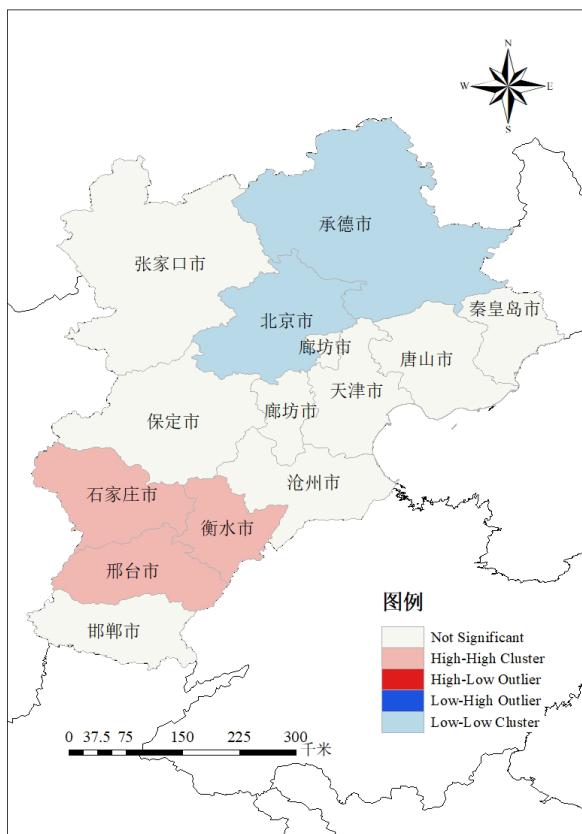
计算公式为：

$$I_i = \frac{(x_i - \bar{x})}{s^2} \sum_{j=1}^n \omega_{ij} (x_j - \bar{x})$$

其中  $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$  是样本方差，其余变量与全局莫兰指数相同。

对京津冀地区 PM2.5 浓度值的局部莫兰指数进行可视化分析，如图 1 所示。

由图 1 可知，京津冀地区存在显著的空间聚类特征。承德和北京为低低相聚(冷点)；石家庄、衡水和邢台为高高相聚(热点)。具体而言，低值聚集区域在环境质量或经济发展等方面存在一定的优势，而高值聚集区域在相应方面面临较大的挑战。这些空间聚类特征对于制定针对性的区域发展策略具有重要的参考价值。

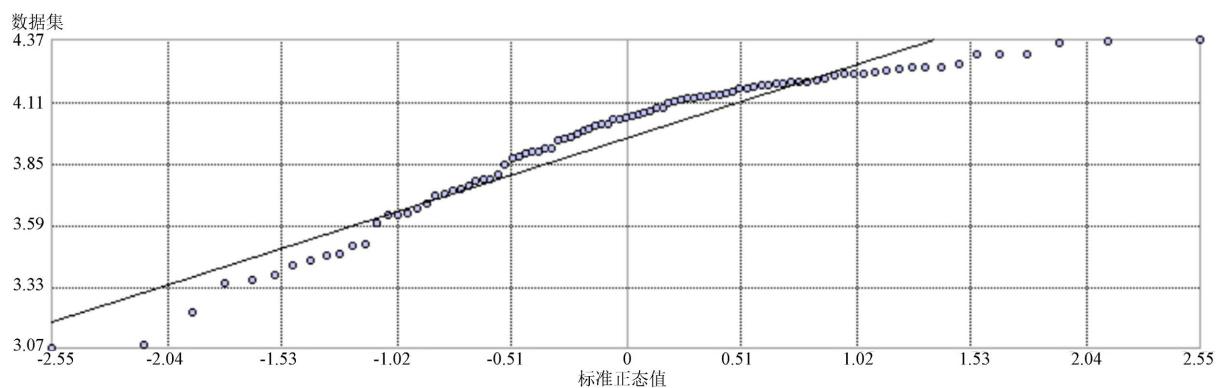


**Figure 1.** Visualization of local Moran's index  
**图 1.** 局部莫兰指数可视化

### 3. 克里金插值

#### 3.1. 数据预处理

进行克里金插值需要满足平稳性假设[4]，数据需要满足正态分布，如果不满足，则需要对数据进行适当变换以满足正态分布[5]。本文对数据进行对数变换，使之满足正态分布，并用正态 QQ 图检验数据的正态性，相应图示如图 2 所示：



**Figure 2.** Normal QQ plot after logarithmic transformation of the data  
**图 2.** 对数据做对数变换后的正态 QQ 图

正态 QQ 图中数据近似在直线左右, 基本满足正态分布。

考虑到京津冀地势呈现出显著的高低差异, 不同地区 PM2.5 的期望不同, 本文选用泛克里金插值法 [6]。

### 3.2. 数值结果(供暖季)

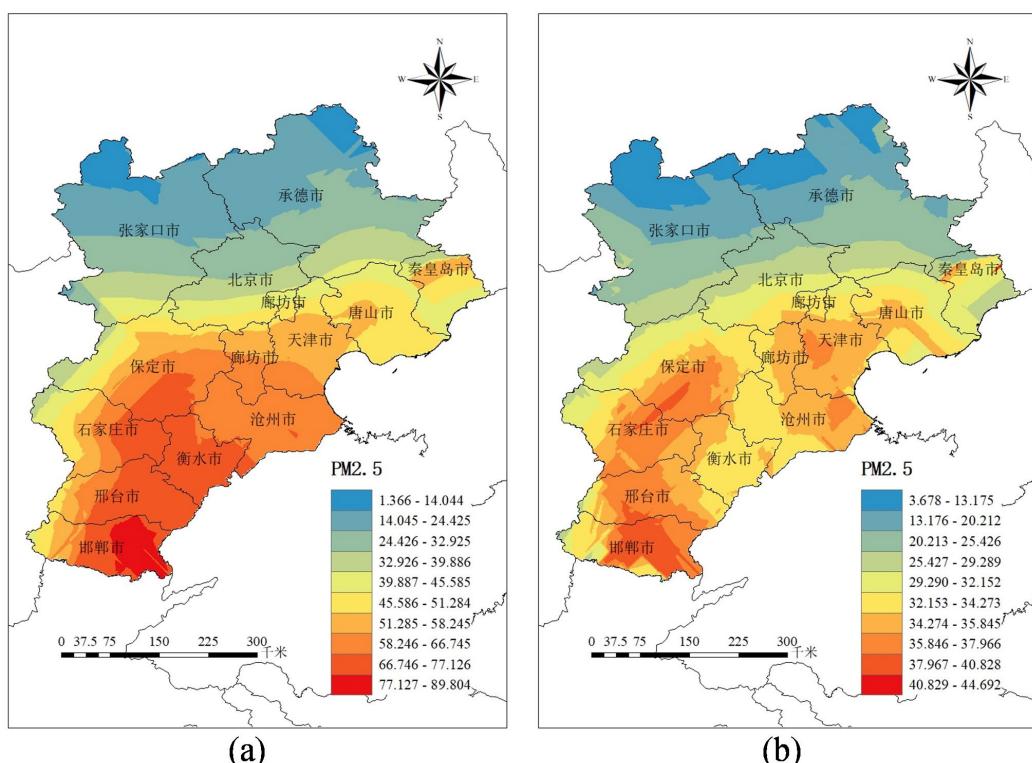
京津冀供暖季的 PM2.5 均方根误差为 4.4146, 即  $RMSE = 4.4146$ , 表明预测误差的典型幅度为 4.41  $\mu\text{g}/\text{m}^3$ , 需结合 PM2.5 浓度范围(如京津冀冬季 PM2.5 常处于 50~150  $\mu\text{g}/\text{m}^3$ )评估相对误差。若浓度中位数为 100  $\mu\text{g}/\text{m}^3$ , 则相对误差约为 4.4%, 属于合理范围。

标准化平均值为-0.0680, 接近于 0 (理想值为 0) 以及标准化均方根为 0.9385, 接近 1 (理想值为 1), 表明模型在标准化尺度下无显著偏差, 且误差分布接近正态, 模型整体校准较好。

标准化误差指标表现良好, 说明泛克里金插值模型在统计假设(如正态性、无偏性)下较为稳健。空间自相关性(块金效应占比 15%)利用充分, 核心变异结构建模合理。

### 3.3. 泛克里金插值结果

在泛克里金插值模型的基础上, 对供暖季和非供暖季京津冀各地区的 PM2.5 浓度值进行泛克里金插值, 如图 3 所示。



**Figure 3.** Comparison of universal kriging interpolation between heating and non-heating seasons; (a) Universal kriging interpolation in heating season; (b) Universal kriging interpolation in non-heating season

**图 3.** 供暖季与非供暖季泛克里金插值对比; (a) 供暖季泛克里金插值; (b) 非供暖季泛克里金插值

图 3(a) 供暖季 PM2.5 浓度的泛克里金插值结果表明, 供暖季泛克里金插值与局部莫兰指数的结果相吻合, 污染主要集中在石家庄、保定、衡水、邢台和邯郸; 张家口、北京和承德地区的空气质量在供暖季

没有明显下降。图 3(b)非供暖季 PM2.5 浓度的泛克里金插值结果表明污染出现明显的聚类，例如保定、石家庄、沧州、邯郸、廊坊、天津等地。

下面依据插值结果分析成因，并给出一些建议。

地理原因：北京、张家口、承德：北京地处华北平原北部，张家口和承德位于燕山 - 太行山生态屏障区，地形开阔且风力较强，有利于污染物扩散。此外，张家口作为冬奥会举办地，生态保护力度大，植被覆盖率较高，进一步抑制了扬尘污染。石家庄、保定、衡水等南部城市：位于太行山东麓，冬季易形成静稳天气，逆温层频繁，导致污染物难以扩散。加之城市群密集，工业和交通排放叠加，从而加剧污染积累。

供暖方式：北京及周边地区已完成大规模“煤改气”“煤改电”，清洁能源占比高。例如，北京市 PM2.5 年均浓度连续三年达标(32 微克/立方米)，燃气供暖比例超 99%。而石家庄、邯郸等城市仍依赖燃煤供暖，尤其是农村地区散煤燃烧问题突出。邯郸市 2023 年 1 月 PM2.5 浓度位列全省倒数第一，与燃煤污染密切相关。

工业结构：张家口、承德以旅游业和轻工业为主，重污染企业较少。石家庄、邢台以及邯郸等地钢铁、建材、化工等高耗能产业集中，冬季生产叠加供暖排放，导致 PM2.5 浓度显著升高。

非供暖季图对比：在非供暖季污染明显集中在各市的城市地区，且呈现出区域化污染。其中天津作为工业重镇，钢铁、石化等重工业集中，非供暖季虽无燃煤供暖压力，但工业生产排放的 PM2.5、VOCs (挥发性有机物) 和 NOx (氮氧化物) 仍居高不下。沧州市化工、建材等高污染行业密集，其在经济开发区应急预案中提到需对 39 个重点行业企业实施差异化管控，但在非供暖季时部分企业仍存在排放超标问题。廊坊地区虽推进绩效分级减排，但工业聚集区(如燕郊)的 VOCs 排放使得其夏季臭氧层被严重污染。

治理方案：加快清洁能源替代，推进“煤改气”“煤改电”全覆盖，尤其是农村地区散煤治理；参照京津冀秋冬季攻坚方案，建立三地统一的 O<sub>3</sub> 和 PM2.5 协同防控机制，共享污染源清单和预警信息；淘汰落后产能，严控钢铁、化工等高耗能行业新增产能，推动沧州、天津等地的传统产业向绿色制造转型。

#### 4. 利用克里金插值结果重新计算全局和局部莫兰指数分析相关性

本文初步计算得到的莫兰指数来源于京津冀地区勘测点的数据，然而勘测点在地图上表现出明显聚集现象，分布不均衡，且有些地区勘测点数量极其稀少。为得到更准确的莫兰指数，本文利用克里金插值选取预测点(各个市的预测点数取各个市勘测点的五倍)，将勘测点的数量扩大，使勘测点的分布变得均匀，以得到更为准确的莫兰指数(图 4 展示了北京市的勘测点以及预测点)。

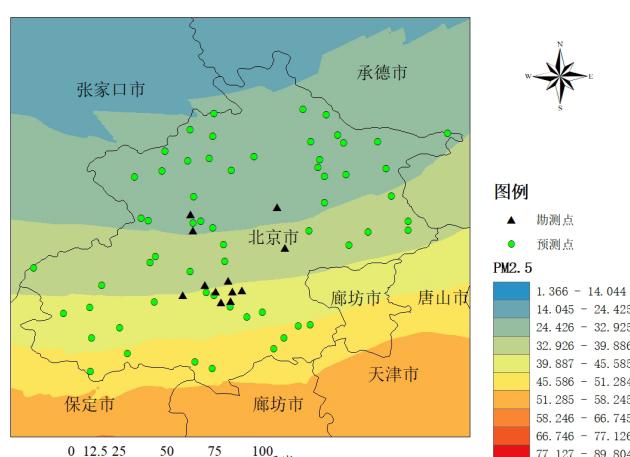


Figure 4. Beijing sampling points and random points

图 4. 北京勘测点位与随机点位

#### 4.1. 全局莫兰指数

Table 2. Global Moran's index  
表 2. 全局莫兰指数

全局莫兰指数汇总/Summary of Global Moran's Index	
全局莫兰指数(Global Moran's I)	0.572537
预期指数(Expected Index)	-0.083333
方差(Variance)	0.034579
z 得分(z-score)	3.527068
p 值(p-value)	0.000420

由表 2 可知 z 得分为 3.53, p 值为 0.000420, 结果在 99% 置信水平下显著; z 得分均高于 0.01 显著性水平的临界值(2.58), p 值均小于 0.01; 说明随机产生此聚类模式的概率小于 1%, 结果高度显著。表 2 中呈现 z 得分更高、p 值更小, 说明其空间聚类模式的随机性概率更低, 统计可靠性更稳健。

由表 1、表 2 可知, 初步全局莫兰指数为 0.543, 呈正值, 表现出空间聚集性, 完善后全局莫兰指数为 0.573, 呈正值且大于初步全局莫兰指数, 空间聚集性更强。两次全局莫兰指数均远高于预期指数(-0.083), 表明 PM2.5 浓度存在显著的空间聚类。

#### 4.2. 局部莫兰指数

对局部莫兰指数进行空间聚类可视化分析, 如图 5 所示。

由图 5 可知, 京津冀地区在 PM2.5 浓度上存在明显且具有区域特征的空间聚类现象。张家口、北京和承德显示出明显清洁聚类(冷点); 衡水和邢台呈现出显著污染聚类(热点)。相比原始 PM2.5 浓度数据, 采用克里金插值法对进行处理后的结果更稳健, 在实际应用方面更有指导意义。

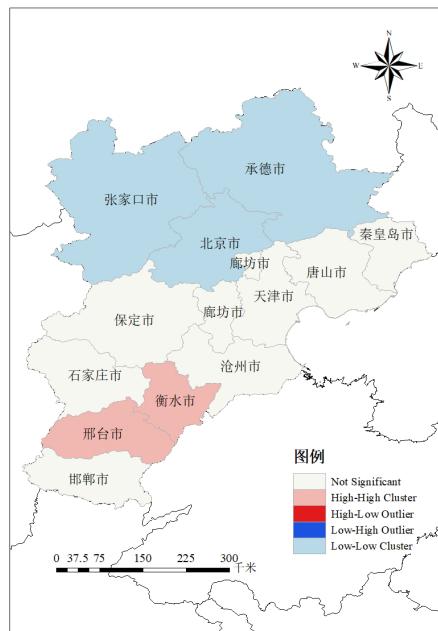


Figure 5. Visualization of local Moran's index  
图 5. 局部莫兰指数可视化

## 5. 总结

从整体上来看,京津冀 13 个市 PM2.5 浓度的空间相关性很强。从地理、生活方式、工业生产等方面都有密切联系。从局部来看,高污染区域集中在京津冀中南部(如石家庄、保定等工业城市),低污染区位于北部山区(如承德、张家口等环保城市),并且非供暖季污染主要聚集在京津冀中南部城市地区,且呈现明显聚集现象(以天津为中心;以石家庄、邢台为中心)。供暖季污染扩散面很大,对京津冀中南地区污染很严重,但对京津冀北部影响较弱。京津冀地区污染格局受地理环境、能源结构与产业布局等多重因素制约。在地理维度,北部张家口、承德依托燕山-太行山生态屏障及风力资源优势形成天然扩散通道,而南部石家庄、保定等地因太行山东麓地形阻滞与静稳天气频发,叠加密集城市群形成的“污染辐合效应”,导致冬季 PM2.5 浓度比北部高。能源结构差异尤为突出,北京通过“煤改气”工程使清洁供暖率达 99%,PM2.5 年均浓度稳定达标,而石家庄、邯郸等地农村散煤燃烧仍贡献冬季 PM2.5 排放量的 32%~40%。产业空间分异加剧污染梯度,北部以旅游业为主的产业结构 PM2.5 排放强度远低于南部钢铁、化工产业集群区的 PM2.5 排放强度。非供暖季污染呈现新型复合特征,天津、沧州等工业重镇因石化、钢铁生产排放的 PM2.5-VOCs-NOx 复合污染突出。

综上,京津冀地区不仅仅需要经济一体化,环境也需要一体化联合治理。

## 基金项目

大学生创新训练项目(202407017)。

## 参考文献

- [1] 周体鹏. 基于克里金插值法的昆明市 PM2.5 预测[D]: [硕士学位论文]. 昆明: 云南大学, 2016.
- [2] 盛晴, 洪志敏, 陈女珍. 京津冀地区 PM2.5 空间分布特征及其影响因素分析[J]. 环境保护科学, 2023, 49(5): 68-75.
- [3] 王翠婷, 童童, 汤萌萌, 等. 基于莫兰指数的丘陵地区高标准农田建设时序分区——以安徽省滁州市凤阳县为例[J]. 江苏农业学报, 2024, 40(1): 83-92.
- [4] 刘爱利, 王培法, 丁园圆. 地统计学概论[M]. 北京: 科学出版社, 2012.
- [5] 肖格新. 空间统计实战[M]. 北京: 科学出版社, 2018.
- [6] 刘婕. 泛克里金插值法在北京市 PM2.5 空间模型建立中的应用[D]: [硕士学位论文]. 北京: 中国疾病预防控制中心, 2020.