

引用格式:孙亚男,费锦华.基于机器学习的雾霾污染精准治理[J].资源科学,2021,43(5):872-885.[Sun Y N, Fei J H. Precise governance of haze pollution based on machine learning[J]. Resources Science, 2021, 43(5): 872-885] DOI: 10.18402/resci.2021.05.02

# 基于机器学习的雾霾污染精准治理

孙亚男,费锦华

(山东财经大学工商管理学院,济南 250014)

**摘要:**实施雾霾污染精准治理是应对局部地区雾霾污染的重要举措。本文利用2001—2016年263个中国城市PM<sub>2.5</sub>数据,基于机器学习构建了决策树递归分析方法,量化了雾霾污染分区因素和治理因素的交互影响,进而识别出雾霾污染精准治理区域及其治理因素。结果表明:①地理区位、行政级别、产业结构、区域规划、经济地带是识别中国城市雾霾污染精准治理区域的分区因素,据此本文识别出4类城市雾霾污染精准治理区域。②对于沿海城市,提高经济发展水平、降低第二产业占比有利于提高雾霾污染治理水平;对于非沿海城市,降低人口密度有助于缓解雾霾污染,其中非沿海、非省会城市提高科技进步水平也可以有效控制雾霾污染。③中国城市雾霾污染治理分区因素的长期演变趋势与中国“五年规划”期间的国家政策具有高度契合性。因此,科学施策、精准调控有利于雾霾污染精准治理。

**关键词:**机器学习;城市雾霾污染;精准治理;影响因素;递归分析

DOI:10.18402/resci.2021.05.02

## 1 引言

绿色和平组织和Air Visual联合发布的《2018年世界空气质量报告》指出,从全国看,中国雾霾污染已退出倒数10名榜单,整体空气质量得到明显改善,然而从城市看,PM<sub>2.5</sub>平均浓度前50位城市中,中国城市仍占据22个席位。当前中国城市化进程的快速推进加剧了中国局部雾霾污染<sup>[1,2]</sup>,中国城市雾霾污染已成为中国生态文明建设的“绊脚石”,加强中国城市雾霾污染治理迫在眉睫。

雾霾的产生是地区经济社会发展的结果<sup>[1,3]</sup>。受城市发展初始水平、政策法规、地理区位、资源禀赋等因素的影响<sup>[4,5]</sup>,中国城市雾霾污染正呈现出显著的空间非均衡分布特征<sup>[6]</sup>,这导致中国整体空气质量的提升未能带动全部城市雾霾污染的同步改善。因此,因地制宜地开展中国城市雾霾污染精准治理成为必然的选择。同时,中国在2018年《打赢蓝天保卫战三年行动计划》中明确提出:“强化区域

联防联控,有效应对重污染天气。”换言之,如何破解城市雾霾污染客观存在的异质性与区域联防联控治理同效性之间的矛盾,成为成功实施中国城市雾霾污染精准治理的关键。

中国雾霾污染的产生尽管受到环境、气象、生态等诸多因素的影响<sup>[7,8]</sup>,但归根结底源于粗放式经济增长方式<sup>[9,10]</sup>。廓清中国雾霾污染与经济发展之间的关系已成为学者们研究的核心内容之一。马丽梅等<sup>[11]</sup>运用空间计量方法分析中国31个省份的雾霾污染(PM<sub>2.5</sub>)与经济发展之间存在“U”型关系。邵帅等<sup>[10]</sup>则进一步从模型内生性、空间权重选择、遗漏变量等方面完善了马丽梅等学者的研究,同样得出中国省际雾霾污染(PM<sub>2.5</sub>)与经济发展之间存在“U”型关系。而刘华军等<sup>[12]</sup>采用中国160个城市雾霾污染(PM<sub>2.5</sub>、PM<sub>10</sub>)数据研究发现,中国城市雾霾污染(PM<sub>10</sub>)与经济发展之间存在倒“U”型关系,而城市雾霾污染(PM<sub>2.5</sub>)与经济发展之间

收稿日期:2020-03-11 修订日期:2020-08-11

基金项目:山东省自然科学基金项目(ZR2018MG010);教育部人文社会科学基金项目(19YJCZH107);山东省高等学校人文社会科学计划项目(J17RA080)。

作者简介:孙亚男,男,山东蓬莱人,博士,教授,主要研究方向为经济增长与绿色发展。E-mail: sun\_ynan@163.com

2021年5月

不存在倒“U”型关系,在进一步控制了产业结构、人口密度和地理区位等变量之后,揭示出城市雾霾污染(PM<sub>2.5</sub>、PM<sub>10</sub>)与经济发展之间呈现线性递减关系。而这一结论与王敏等<sup>[13]</sup>利用112个城市数据得出的雾霾污染(PM<sub>10</sub>)与经济发展之间呈“U”型曲线关系的结论不同。虽然上述研究在实证模型、估计方法和数据来源上存在差异,但随着雾霾污染研究对象由省级转向城市,数据规模的增长以及行政区划的改变,使得空间分布、社会经济等因素的异质性对科学判断雾霾污染与经济发展关系产生了愈加明显的影响。

为考察因素异质性的存在对中国雾霾污染与经济发展之间关系的影响,学者们进一步从经济发展水平异质性和污染物排放量<sup>[14]</sup>、工业及工业结构和经济地带<sup>[15]</sup>、外商直接投资<sup>[16]</sup>、城镇化推进模式<sup>[1]</sup>等多个视角实证分析了经济增长、工业份额、技术进步等经济因素对雾霾污染产生的影响,结果均表明因素异质性的存在将不同程度地影响中国雾霾污染治理成效。同时,雾霾污染具有显著的空间依赖性<sup>[17]</sup>,考虑雾霾污染空间异质性的影响,无疑可以更加精准有效地衡量雾霾污染与经济发展之间的关系<sup>[10]</sup>。然而,部分学者依据地理区位<sup>[18]</sup>、经济地带<sup>[19]</sup>、区位规划<sup>[20]</sup>对影响雾霾污染的经济因素,例如外商直接投资、经济发展水平、第二产业占比,进行实证分析后并未得到一致的结果。

已有大量研究证明,加强区域联防联控是治理雾霾的有效方式<sup>[21,22]</sup>。然而,面对雾霾污染日益突出的区域差异,科学分析和优化区域联防联控范围成为雾霾治理亟待解决的首要问题<sup>[21,23]</sup>。部分学者针对上述问题展开了探讨。刘华军等<sup>[24]</sup>利用26个城市雾霾污染数据的相关系数,选取相关系数值大于0.5为阈值,识别出东北、京津冀、华北、中三角、长三角、珠三角、成渝和大西北八大雾霾治理区。该研究采用较小的样本量和描述统计,定性地论述了联防联控区域的选择原则和初步的区域范围,但对于上述区域范围划分的成因缺乏量化解释。正如陈诗一等<sup>[23]</sup>指出的,客观存在的雾霾成因异质性已成为区域雾霾污染联防联控的现实困境,为此,本文将进一步量化分析雾霾治理联防联控区域划

分的分区因素,并揭示出分区因素的异质性对区域划分的影响,为更加精准地实施区域联防联控,提升区域雾霾污染联防联控成效提供参考。

从以上研究不难看出,当存在多个分区因素时,学者们仅主观选取单一因素开展雾霾污染治理研究,易导致结果出现偏误。事实上,当空间、经济等分区因素存在差异时,上述分区因素之间的交互作用将产生明显的叠加影响,这意味着一方面需要科学地识别出分区因素,避免叠加影响产生的内生性问题导致的实证结果有偏,另一方面也需要揭示出分区因素对雾霾污染治理因素的影响层级和影响程度,为科学、精准地治理雾霾提供决策支持。

机器学习方法与现代计量经济学研究的结合,为更加精准地识别和预测研究对象之间存在的影 响机制提供了新途径<sup>[25]</sup>。机器学习在分析变量之间影响机制的优势,主要体现在机器学习具有揭示未知复杂结构的能力<sup>[26]</sup>,这其中主要包括模型变量的选择和分类识别能力<sup>[27]</sup>。而机器学习的上述优势特别有助于解决处置效应(Treatment Effects)的异质性<sup>①</sup>估计问题。Athey等<sup>[28]</sup>提出了一个估计异质性处置效应的递归分区(Recursive Partitioning)方法,进一步将上述框架扩展为基于随机森林的非参数估计方法,作为广义随机森林算法<sup>[29]</sup>,上述方法借助机器学习可以实现传统的局部广义矩估计或者局部最大似然估计。Asher等<sup>[30]</sup>也借鉴Athey等<sup>[28]</sup>的可靠树估计方法,针对不同样本中相同模型的变量选择存在差异的问题提出了矩估计森林(Moment Forests)方法,该方法采用一种基于条件矩估计的半参数模型一致性收敛算法对协变量异质性进行分区和估计。

以上研究借助机器学习的分类算法,利用多个协变量的异质性对样本进行分区,并解释协变量的跨分区变化,为本文试图解决雾霾污染异质性的识别和影响程度估计问题提供了研究思路。然而,正如Varian<sup>[27]</sup>所说,样本数量的增加极大地减轻了样本不确定性,传统的计量经济学将研究重点放在样本不确定性,而模型不确定性在大样本条件下才是更应该关注的问题。事实上,随着可获得样本规模的增长,模型不确定性也能突显出变量异质性的存

① Athey等<sup>[32]</sup>指出异质性来自可观察的协变量(Covariate Space)。

在。Zeileis等<sup>[31]</sup>提出了一个能够较好地嵌入回归模型的递归分析方法,特别是针对应用最大似然估计的模型,以识别协变量异质性的影响。但与Athey、Asher等学者的研究不同,基于模型的递归分析方法并未给出关键和有效的统计检验,无法准确地判断研究结果的可靠性和鲁棒性<sup>[32]</sup>。而且,中国城市雾霾污染存在较为典型的空间依赖特征<sup>[21]</sup>,简单套用上述模型,在空间均质和空间独立的假设下进行分析,忽视中国城市雾霾污染的空间异质性,必然导致估计结果有偏<sup>[10,33]</sup>。为此,借助机器学习,本文采用基于空间自回归模型的决策树递归分析方法,一方面借鉴Athey等学者的思想,利用多个变量的异质性对样本进行分区,另一方面采用Zeileis等学者基于模型的递归分析思想,基于模型不确定性识别变量异质性的显著影响。通过递归分析雾霾污染分区因素与雾霾治理因素之间的交互影响,自动筛选出中国城市雾霾污染的分区因素,量化分析分区因素对中国城市雾霾污染治理因素的影响层级与影响程度,并同步实证揭示中国城市雾霾污染的分区因素的演变过程,以期对中国城市雾霾科学精准治理提供决策参考和政策建议。

## 2 研究方法和数据说明

### 2.1 研究方法

#### 2.1.1 基于分区异质性的空间自回归模型构建

考虑到中国城市雾霾污染的空间依赖性<sup>[10,34]</sup>,本文将因变量的空间滞后项纳入研究模型,假设向量 $\beta$ 在递归分区形成的各个城市分区之间可以变化。构建基于分区异质性的空间自回归方程:

$$y_i = \rho w_i^T y + x_{ij}^T \beta^{g(i)} + \varepsilon_i \quad (1)$$

式中: $i=1,2,3,\dots,263$ 表示样本城市; $y_i$ 为第 $i$ 个城市PM2.5浓度均值; $x_{ij}$ 代表第 $i$ 个城市第 $j$ 个影响因素; $\rho$ 为空间滞后项的参数; $w_i$ 为第 $i$ 个城市的空间权重向量; $T$ 表示矩阵转置; $\beta^{g(i)}$ 为包含第 $i$ 个城市的 $g(i)$ 分区的参数向量,其中 $g(i)$ 分区通过雾霾污染异质性变量进行递归分区得到; $\varepsilon$ 为随机误差项。

#### 2.1.2 未知分区与空间滞后系数的递归估计

现有文献对于给定分组或者无分组条件下空间滞后模型的研究较为成熟,但在面对既要确定未知的分区,又需要同时估计模型系数的情况时,直

接采用现有成熟的估计方法不能解决上述问题。为此,本文借鉴Sela等<sup>[35]</sup>提出的迭代思想,估计基于分区异质性的空间自回归模型的空间滞后系数。具体而言,迭代估计主要包括两个过程:①给定分区结构时估计空间滞后系数;②给定空间滞后系数时估计分区结构。当回归树不再发生变化或者样本数量小于限值时,上述迭代过程停止。详细过程如下:

①对于所有的观测样本 $i$ 给定组 $g(i) \in \{1, 2, \dots, G\}$ ,式(1)是一个关于分组和回归变量的空间滞后模型,可估计得到 $\hat{\rho}$ 以及 $\hat{\beta}^g$ 。

②依据估计得到的空间滞后系数 $\hat{\rho}$ ,构建 $\tilde{y}_i = y_i - \hat{\rho} w_i^T y$ ,对被解释变量进行调整,并将观测样本分区为 $g(i)$ 。随后采用式(2)估计参数 $\beta^{g(i)}$ ,具体模型下:

$$\tilde{y}_i = x_{ij}^T \beta^{g(i)} + \varepsilon_i \quad (2)$$

以上两步迭代过程均可作为初始过程,步骤①作为初始条件,则表明所有的观测样本在一个分区内,而步骤②作为初始条件,则意味着空间滞后系数为零时进行递归分区。经进一步验证,上述初始条件的差异不影响最终的实证结果。

#### 2.1.3 基于模型的决策树估计

估计空间滞后系数时,借助参数稳定性检验(Parameter Stability Test)评估观测样本是否需要分区。依据任何一个分区变量进行检验,如果式(2)存在显著的参数不稳定性,则根据该分区节点将观测样本分为子样本,并不断重复检验所有分区变量。若在上述递归的过程中,不能找到更显著的不稳定的分区变量,则递归分区过程停止并形成决策树。为此,根据Andrews<sup>[36]</sup>提出的Quandt-Andrews检验法,构造连续型分区变量回归系数的稳定性检验模型,具体如下:

$$\text{SupLM} = \sup_{i=1, \dots, i} \left\{ \frac{i}{n} \left( 1 - \frac{i}{n} \right) \right\}^{-1} \left\| \hat{V}^{-1/2} n^{-1/2} \sum_{l: z_l \leq i} x_l \hat{\varepsilon}_l \right\|_2^2 \quad (3)$$

式中:SupLM表示上确界似然比; $z_i$ 表示分区变量中观测值,按增序排列; $i$ 代表分区个数, $n$ 代表观测值的数量; $\hat{\varepsilon}$ 代表回归残差项,基于子样本的参数估计; $\hat{V} = n^{-1} \sum_{i=1}^n \hat{\varepsilon}_i^2 x_i x_i^T$ 是梯度外积的协方差估计,用



2021年5月

于标准化得分向量  $x_i \hat{\varepsilon}_i$ ;  $x_i$  为连续型分区变量。而对于离散型分区变量,则根据式(4)进行参数稳定性检验:

$$\chi^2 = \sum_{c=1, \dots, C} \left\| \hat{\nu}^{-1/2} n_c^{-1/2} \sum_{i: z_i=c} x_i \hat{\varepsilon}_i \right\|_2^2 \quad (4)$$

式中:  $c=1, \dots, C$  代表离散型分区变量的类别,用  $n_c$  代表  $z_i$  的数量。

两种检验的渐进  $P$  值可以分别根据平方束缚贝塞尔过程的上确界极限分布和卡方分布算出。同时,本文选择 Bonferroni 法控制多重比较时累积 I 类错误概率增大的问题。在递归分区的过程中,筛选的变量必须满足在 5% 的显著性水平下能够拒绝参数稳定的原假设,若某一回归过程中出现多个分区变量同时满足该条件,选择  $P$  值最小的变量作为分区变量。如果没有检测到更显著的参数不稳定性或子样本数量过小,则递归分区停止。

## 2.2 数据来源及指标

### 2.2.1 被解释变量

本文选取中国地级市 PM2.5 年均浓度为被解释变量,并参照邵帅等<sup>[10]</sup>、李光勤等<sup>[4]</sup>的数据处理方法,采用哥伦比亚大学社会经济数据和应用中心发布的基于卫星监测的 PM2.5 栅格数据(最新数据截至

2016 年),利用 ArcGIS 对其进行解析得到中国地级市 PM2.5 年均浓度数据。

### 2.2.2 解释变量

邵帅等<sup>[1,10]</sup>、陈诗一等<sup>[9]</sup>学者较为全面地分析了中国城市雾霾污染的经济治理因素。依据上述研究,本文选取以下城市雾霾污染的治理因素作为解释变量,包括:①经济发展水平( $\ln Pgdp$ );②产业结构( $SEC$ );③对外开放水平( $FDI$ );④科技进步水平( $SCI$ );⑤交通基础设施( $ROAD$ );⑥能源消耗( $EC$ );⑦人口密度( $POP$ )。具体解释详见表 1。各变量均为“五年规划”期间均值。除特殊声明外,上述数据均来自《中国城市统计年鉴》、EPS 全球数据统计/分析平台。

### 2.2.3 分区因素

通过对现有研究文献的梳理不难发现,当前空间分区因素主要包括城市规模(是否属于大中城市)、经济地带(东、中、西部划分)、行政级别(是否属于省会城市)、区域规划(京津冀、长三角、珠三角经济圈)、地理区位(是否属于沿海城市)等方面;而经济分区因素方面,经济发展、产业结构、技术进步等是环境污染治理的关键分区因素<sup>[13]</sup>,成为现有研究考虑雾霾污染异质性的主要因素。为此,参考马

表 1 变量说明

Table 1 Description of variables

类型	名称	描述
被解释变量	雾霾污染	PM2.5
解释变量	经济发展水平	$\ln Pgdp$
	产业结构	$SEC$
	对外开放水平	$FDI$
	科技进步水平	$SCI$
	交通基础设施	$ROAD$
	能源消耗	$EC$
	人口密度	$POP$
分区因素	经济发展水平	$Spgdp$
	产业结构	$Ssec$
	技术进步	$Sinn$
	城市规模	$Devp$
	经济地带	$East$
	行政级别	$Cap$
	区域规划	$Eci$
	地理区位	$Coa$

注:将北京、天津、上海、重庆 4 个直辖市作为省会城市进行处理。

丽梅等<sup>[15]</sup>、刘华军等<sup>[21]</sup>、陈诗一等<sup>[33]</sup>、邵帅等<sup>[1,10]</sup>的研究,选取经济发展、产业结构、技术进步、城市规模、经济地带、行政级别、区域规划、地理区位等作为分区变量(表1)。因本文试图验证国家政策制定是否对雾霾污染治理存在影响,并考虑到后续稳健性考察的需要,在剔除了考察期内数据严重缺失的城市后,为与被解释变量考察期一致,选择2001—2016年间中国263个城市雾霾污染影响因素进行相关分析,为消除量纲影响,对解释变量和被解释变量进行了标准化处理。数据主要来源于《中国城市统计年鉴》、EPS全球数据统计/分析平台。

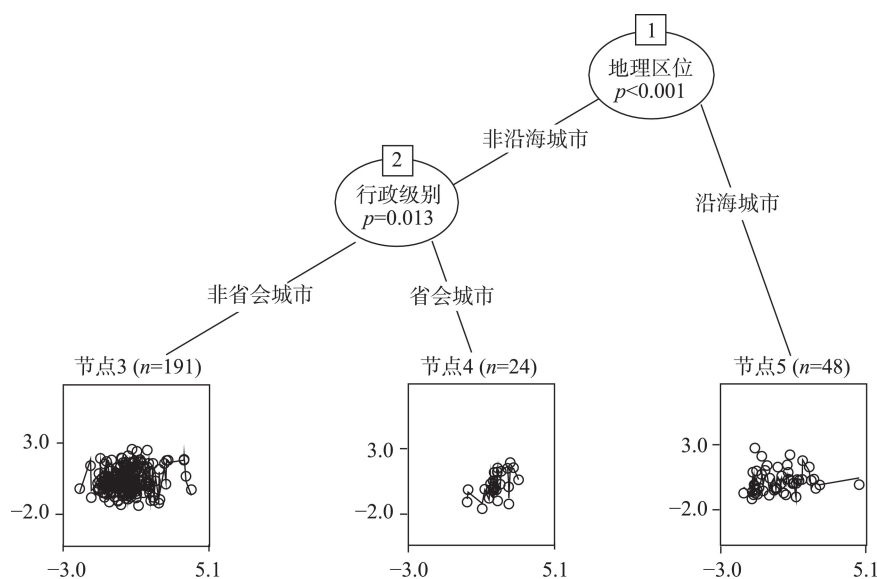
### 3 结果与分析

借助探索性空间数据分析(ESDA)对雾霾污染的空间自相关性进行检验,Moran's  $I$ 结果显示,地理距离权重下雾霾污染具有显著的空间自相关性<sup>②</sup>。因此,将地理距离空间权重纳入研究模型,采用基于空间自回归模型的决策树递归分析方法,利用对雾霾污染分区因素和治理因素的同步检验和交叉分析,对263个中国城市雾霾污染治理因素的差异化影响进行分析。

### 3.1 分区因素的影响层级与影响程度分析

“十五”计划(2001—2005年)期间共5次递归分析过程<sup>③</sup>,第一次回归检验结果表明,经济地带分区变量、区域规划分区变量、地理区位分区变量均通过了5%的显著性水平检验,但在本次回归过程中,经济地带分区变量和区域规划分区变量的 $P$ 值均大于地理区位分区变量的 $P$ 值,因此,依据前文提出的最小显著性水平的分区变量选择标准,选定地理区位分区变量作为“十五”计划期间的首个分区变量。换言之,是否属于沿海城市是“十五”计划期间城市雾霾污染治理因素存在异质性影响的首要分区因素。因未满足前述迭代终止条件,将进一步识别其他分区因素。

“十五”计划期间5次递归分析后,得出“十五”计划期间决策树,如图1所示。结果表明:①从分区因素影响层级来看,是否属于沿海城市处于第一层级,是该时期城市雾霾污染因素存在差异影响的首要分区因素。内陆城市间仍存在显著的分区因素,是否属于省会城市处于第二层级,应区分沿海城市、内陆省会城市、内陆非省会城市三部分,实行差



注:决策树内部节点显示筛选出的分区变量以及来自参数稳定性检验的 $P$ 值,终端节点显示“五年规划”期间PM2.5浓度均值与“五年规划”初始年份人均GDP之间的散点图,下同。

图1 “十五”计划时期决策树

Figure 1 Decision tree during the 10th Five-Year Plan period

② 根据城市经纬度数据构建地理距离权重矩阵。限于篇幅,未报告检验结果,有需要的读者可向作者索取。

③ 限于篇幅,未报告递归分析的参数稳定性检验结果,有需要的读者可向作者索取。

2021年5月

异化雾霾污染治理。②从分区因素影响程度上看,“十五”计划期间,地理区位分区变量的估计系数为50.08,行政级别分区变量的估计系数为24.47,差异影响程度呈递减趋势。③从子群的划分数量上看,依据地理区位和行政级别2个分区因素将中国263个城市分为3个子群,具体包括:非沿海非省会城市子群(节点3),样本数为191;非沿海省会城市子群(节点4),样本数为24;沿海城市子群(节点5),样本数为48。

“十一五”规划(2006—2010年)期间经过7次递归分析过程,筛选出地理区位、行政级别和产业结构3个分区因素。图2给出“十一五”规划期间决策树。①从分区因素影响层级来看,是否属于沿海城市仍是该期间城市雾霾污染治理因素存在差异影响的首要分区因素;是否为省会城市处于第二层级;产业结构处于第三层级。②从分区因素影响程度上看,“十一五”规划期间,地理区位分区变量的估计系数为52.41,行政级别分区变量的估计系数为31.42,产业结构分区变量的估计系数为30.78,差异影响程度呈递减趋势。③从子群的划分数量上看,依据上述3个分区因素将中国263个城市划分为4个子群,具体包括:非沿海非省会且第二产业占比 $\leq$

54.96%的城市子群(节点4),样本数为138;非沿海非省会且第二产业占比 $>$ 54.96%的城市子群(节点5),样本数为53;非沿海省会城市子群(节点6),样本数为24;沿海城市子群(节点7),样本数为48。

“十二五”规划(2011—2015年)期间经过7次递归分析过程,共筛选出地理区位、区域规划和经济地带3个分区变量,图3给出“十二五”规划期间决策树。①从分区因素影响层级来看,“十二五”规划期间首要分区因素与之前2个“五年规划”期间一致,污染治理因素存在差异影响的首要分区因素仍是地理区位;是否属于经济圈处于第二层级;是否属于东部地区城市处于第三层级,说明该时期进行雾霾污染治理时,对非沿海非经济圈城市仍不能一概而论,应进一步区分是否属于东部地区城市,以实现精细化、差异化雾霾污染治理策略。②从分区因素影响程度来看,“十二五”规划期间,地理区位分区变量估计系数为47.12,区域规划分区变量估计系数为30.60,经济地带分区变量估计系数为26.36,差异影响程度呈递减趋势。③从子群的划分数量上看,依据上述3个分区因素将中国263个城市划分为4个子群,具体包括:非沿海非经济圈非东部的“三非”城市子群(节点4),样本数为157;非沿海非

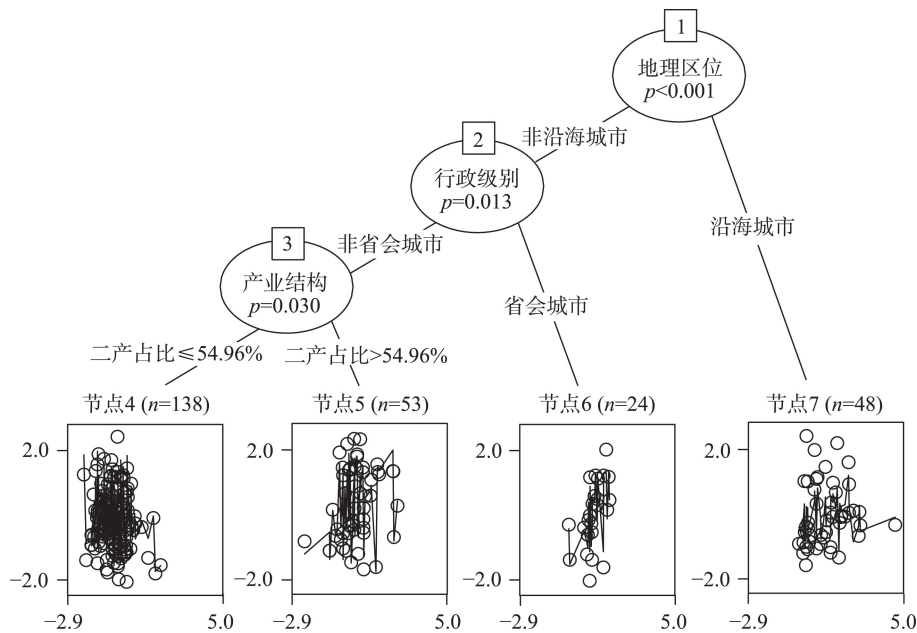


图2 “十一五”规划时期决策树

Figure 2 Decision tree during the 11th Five-Year Plan period

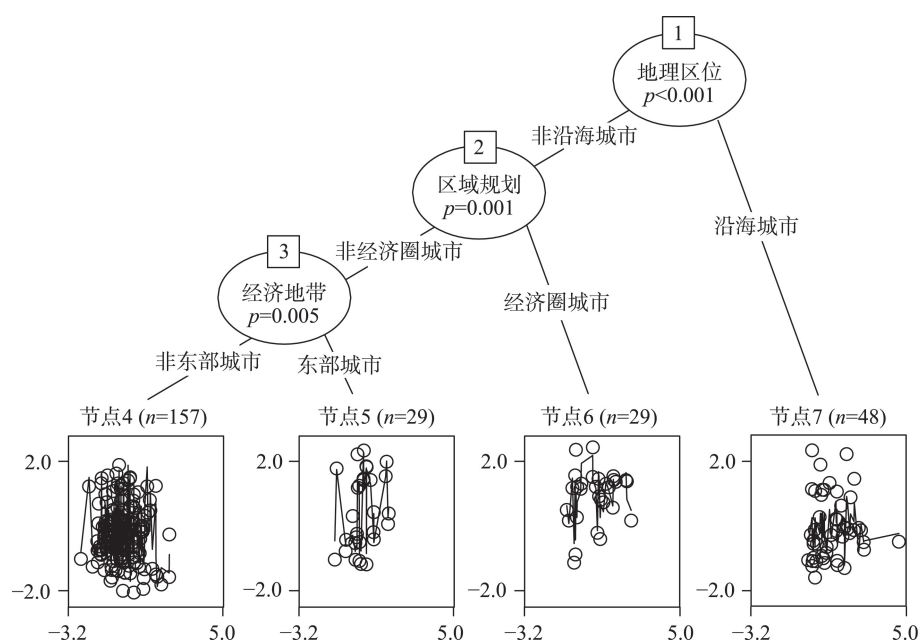


图3 “十二五”规划时期决策树

Figure 3 Decision tree during the 12th Five-Year Plan period

经济圈东部城市子群(节点5),样本数为29;非沿海经济圈城市子群(节点6),样本数为29;沿海城市子群(节点7),样本数为48。

进一步分析发现,中国城市雾霾治理分区因素的长期演变趋势与中国“五年规划”期间的国家政策具有高度契合性。“十一五”规划时期更加注重产业结构优化升级,明确提出服务业增加值占国内生产总值的比重要提高3个百分点。在此期间,产业结构分区因素对城市雾霾污染治理因素存在显著性影响。《“十二五”规划纲要》明确提出优化格局,促进区域协调发展的战略要求,“推进京津冀、长江三角洲、珠江三角洲地区区域经济一体化发展”,“推进东部地区城市在改革开放中先行先试,在转变经济发展方式、调整经济结构和自主创新中走在全国前列”。因此,区域规划和经济地带分区因素成为该时期影响城市雾霾污染治理的显著性因素。

### 3.2 雾霾污染治理因素分析

表2报告了“五年规划”期间雾霾污染治理因素递归分析结果。从中可以看出,“十五”计划期间雾霾污染精准治理的影响因素如下:①在191个内陆非省会城市中(图1),经济发展水平、人口密度和科技进步水平均通过了1%的显著性检验,其中经济发展水平和人口密度的提高促使雾霾污染加剧,而

科技进步水平提升则具有显著的减霾效果。②对于24个内陆省会城市而言,人口密度通过了1%的显著性检验,表明人口密度增大加剧了该类城市雾霾污染。经济发展水平对雾霾污染具有“促增”效应,但统计上并不显著。③从48个沿海城市的实证结果看,在10%的显著性水平上,经济发展水平抑制了沿海城市雾霾污染,表明“十五”计划时期,沿海城市已越过环境库兹涅茨曲线的拐点。科技进步水平则加剧了沿海城市雾霾污染,表明政府科技资金的投入尚未实质性助力减霾技术的提升。同时产业结构通过了1%的显著性检验,表明第二产业占比过高加剧了沿海城市的雾霾污染。

“十一五”规划期间雾霾污染精准治理的影响因素如下:①对于138个内陆非省会且第二产业占比 $\leq 54.96\%$ 的城市,经济发展水平和人口密度通过了1%的显著性检验,对外开放水平通过了5%的显著性检验,均加剧了雾霾污染。而科技进步水平通过了5%的显著性检验,具有显著的减霾效果。②53个第二产业占比 $> 54.96\%$ 且为内陆非省会城市的实证结果表明,人口密度也通过了1%的显著性检验,经济发展水平通过了5%的显著性检验,均加剧了雾霾污染。而科技进步水平通过了5%的显著性检验,具有显著的“减霾”效果。与第二产业占比



2021年5月

表2 “五年规划”期间雾霾污染影响因素递归分析结果

Table 2 Recursive partitioning analysis results of haze pollution influencing factors during the Five-Year Plan periods

	<i>n</i>	<i>Coa</i>	<i>Cap</i>	<i>Sec</i>	<i>Eci</i>	<i>East</i>	<i>Const.</i>	<i>lnPdgp</i>	<i>SEC</i>	<i>FDI</i>	<i>SCI</i>	<i>ROAD</i>	<i>EC</i>	<i>POP</i>
“十五”计划期间	3	no	no	—	—	—	-0.02 (0.046)	0.24*** (0.076)	-0.05 (0.046)	0.04 (0.062)	-0.39*** (0.132)	-0.23 (0.150)	-0.08 (0.121)	0.84*** (0.044)
	4	no	yes	—	—	—	-0.33 (0.200)	0.45 (0.285)	0.22 (0.254)	-0.03 (0.194)	-0.02 (0.099)	-0.20 (0.343)	-0.29 (0.348)	0.78*** (0.176)
	5	yes	—	—	—	—	-0.08 (0.118)	-0.26* (0.134)	0.42*** (0.094)	0.01 (0.053)	0.14* (0.073)	-0.07 (0.100)	0.07 (0.090)	0.00 (0.054)
“十一·五”规划期间	4	no	no	≤54.96%	—	—	-0.25** (0.118)	0.31*** (0.101)	-0.05 (0.056)	0.16** (0.067)	-1.45** (0.566)	-0.32 (0.202)	-0.25 (0.226)	0.86*** (0.057)
	5	no	no	>54.96%	—	—	0.36** (0.162)	0.27** (0.123)	-0.30** (0.118)	-0.03 (0.075)	-0.72** (0.331)	0.29 (0.202)	-0.30*** (0.089)	0.75*** (0.079)
	6	no	yes	—	—	—	-0.13 (0.218)	0.21 (0.308)	0.33 (0.270)	-0.17 (0.151)	0.17 (0.214)	0.08 (0.266)	-0.63** (0.293)	0.76*** (0.144)
“十二·五”规划期间	7	yes	—	—	—	—	-0.20* (0.113)	-0.10 (0.126)	0.44*** (0.089)	0.12* (0.064)	0.04 (0.056)	-0.15 (0.102)	0.17 (0.110)	0.00 (0.054)
	4	no	—	—	no	no	-0.20*** (0.071)	0.21** (0.079)	-0.04 (0.039)	0.05 (0.051)	-0.55*** (0.262)	-0.23 (0.145)	-0.08 (0.078)	0.73*** (0.051)
	5	no	—	—	no	yes	0.15 (0.136)	0.11 (0.204)	0.12 (0.181)	0.08 (0.166)	-0.77 (0.536)	0.33 (0.345)	-0.26 (0.334)	0.90*** (0.167)
	6	no	—	—	yes	—	0.19 (0.119)	-0.02 (0.198)	0.10 (0.136)	0.13 (0.105)	-0.09 (0.136)	-0.19 (0.214)	-0.32 (0.201)	0.96*** (0.174)
	7	yes	—	—	—	—	-0.16 (0.124)	-0.05 (0.134)	0.28*** (0.079)	0.19*** (0.067)	0.03 (0.071)	-0.14 (0.087)	0.15 (0.110)	0.01 (0.060)

注:\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平。



≤54.96%的城市不同,产业结构通过了5%的显著性检验,能源消耗通过了1%的显著性检验,表明扩大电能使用,替代化石能源,能够显著降低雾霾污染。③在24个内陆省会城市中,能源消耗和人口密度分别通过了5%和1%的显著性检验,其中扩大电能使用具有显著的减霾效果,而人口密度则显著地加剧了雾霾污染。④对于48个沿海城市而言,产业结构通过了1%的显著性检验,对外开放水平通过了10%的显著性检验,均加剧了雾霾污染。

“十二五”规划期间雾霾污染精准治理的影响因素如下:①157个非沿海非经济圈非东部的“三非”城市的实证结果表明,经济发展水平通过了5%的显著性检验,人口密度通过了1%的显著性检验,加剧了雾霾污染。科技进步水平通过了1%的显著性检验,在“三非”城市中发挥出积极的减霾效果。②在29个非沿海非经济圈东部城市中,人口密度通过了1%的显著性检验,是加剧雾霾污染的主要原因。③29个非沿海经济圈城市的人口密度通过了1%的显著性检验,对雾霾污染具有显著的促进作用。④48个沿海城市的产业结构、对外开放水平通过了1%的显著性检验,均加剧了沿海城市的雾霾污染。经济发展水平未通过显著性检验,但与沿海城市雾霾污染仍呈现负相关关系。

综合3个“五年规划”期间递归分析结果可以看出,除“十二五”规划时期经济圈城市外,考察期内非沿海城市经济发展水平对雾霾污染均具有促进作用,表明该类城市仍未越过环境库兹涅茨曲线的拐点。而沿海城市以及“十二五”规划时期非沿海经济圈城市,经济发展水平的提高会抑制雾霾污染加剧的状况,表明该类城市已越过库兹涅茨曲线的拐点,环境库兹涅茨曲线在本文得到了部分验证,这较为符合Wang等<sup>[37]</sup>的研究结论,即中国环境治理处于“局部改善,整体恶化”状态。同时,随着第二产业占比的增加,为产业结构优化调整提供了空间,也为能源合理使用提出了更高的要求<sup>[38]</sup>。如何通过对外开放政策带动产业结构调整,以及通过产业结构调整合理使用外资,形成产业结构与对外开放的良性互动,成为沿海地区雾霾污染治理的长期挑战。另一方面,人口密度较大的地区,经济增长诉求愈发强烈,生产、生活排放加剧了雾霾污染,提

高科技进步水平,倡导清洁能源的使用则是降低雾霾污染的关键途径。

### 3.3 中国城市雾霾污染精准治理区域识别

雾霾污染精准治理不仅要识别雾霾污染治理因素的差异影响,同时要确保区域治理成效的同质性,实现高效减霾。经过上述分析可以发现,在沿海地区、省会城市、东部城市、经济圈等分区因素的叠加影响下,中国城市雾霾治理产生了显著分化。263个中国城市中,有205个城市在连续3个“五年规划”期间可以组成4个雾霾污染精准治理区域(图4)。具体而言,48个沿海城市,如天津、唐山和沧州等一直处于同一雾霾污染精准治理区域内;18个非沿海非经济圈非东部的省会城市,如太原、呼和浩特和长春等同样构成稳定的雾霾污染精准治理区域;非沿海非经济圈非东部非省会城市中,109个第二产业占比≤54.96%的城市,如大同、忻州和晋中等被划归为同一雾霾污染精准治理区域,同时有30个第二产业占比>54.96%的城市,如阳泉、长治和晋城等也构成稳定的雾霾污染精准治理区域。

进一步探究发现,48个沿海城市中,广州、深圳和珠海等6个城市位于珠三角城市群,上海、南通和盐城等9个城市位于长三角城市群,天津、唐山、沧州和秦皇岛4个城市位于京津冀城市群。109个“四非”且第二产业占比≤54.96%的城市中,景德镇、鹰潭和宜春等22个城市属于长江中游城市群,自贡、泸州和德阳等13个城市属于成渝城市群,吉林、四平、辽源、齐齐哈尔和牡丹江5个城市属于哈长城市群。可以看出,中国城市雾霾治理必须打破各区域“各自为战”的治理模式,建立跨城市群的联防联控治理体系。具体而言,在城市群中依据分区因素与治理因素的交互影响,在4个雾霾污染治理区域内实施“求同存异”的联防联控政策,这是实现中国城市雾霾污染精准高效治理的核心路径和重要方式。

事实上,强化区域联防联控是打赢蓝天保卫战的关键。这是因为,一方面雾霾污染存在显著的空间溢出特征<sup>[22]</sup>。PM<sub>2.5</sub>浓度表现出明显的空间集聚性与扩散性<sup>[39-41]</sup>,自然及人文影响因素对雾霾污染空间溢出具有显著影响<sup>[8]</sup>;另一方面,地方政府在雾霾污染区域协同治理过程中的行为选择具有“逐底竞争”特征<sup>[42,43]</sup>。这意味着在经济利益的驱动下,各

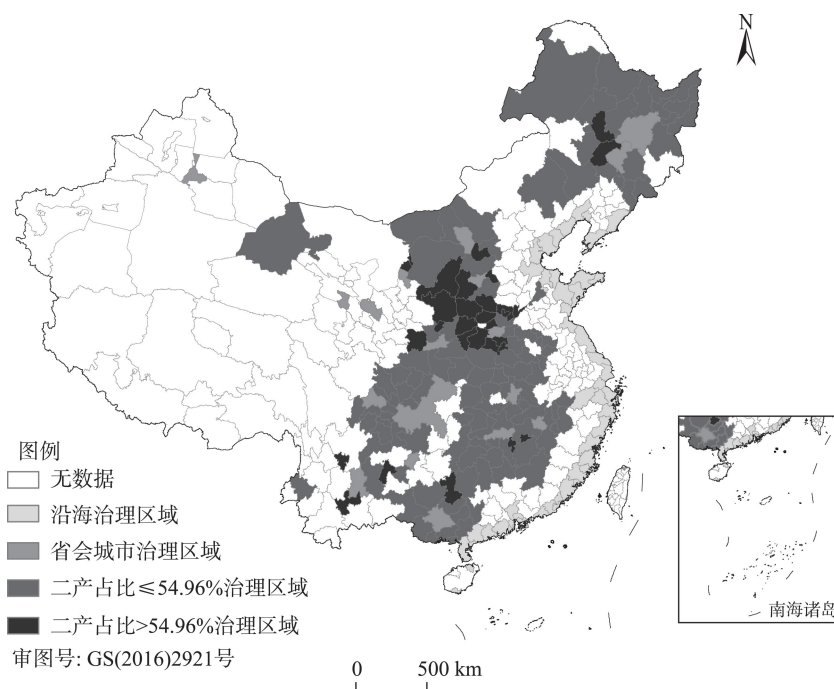


图4 中国城市雾霾污染精准治理区域

Figure 4 Precise governance regions of urban haze pollution in China

地区存在通过放松治理力度来争夺企业、人力和技术等流动性要素资源的策略性竞争行为,加剧潜在雾霾污染源的地区间转移。更为重要的是,由于雾霾污染具有较强的空间溢出效应,本地区的治霾努力也不一定能够显著提升当地环境质量。经济利益驱动和空间溢出的双重效应将进一步加剧“逐底竞争”态势。综上所述,依靠单边治霾和局部治霾难以从整体上、根本上解决区域性雾霾污染问题,应采用系统治理方式,这样不仅可以有效降低治霾成本,也能确保达到预期治理成效。

#### 4 稳健性检验

本文采用更换跨期数据及空间权重两种方式对上述结果进行稳健性检验,具体结果如下:

①将研究数据更换为2001—2010年和2001—2016年的均值。地理区位、行政级别、经济地带分区变量仍然是治理城市雾霾污染的显著性分区因素,但产业结构未被识别为显著的异质性分区变量,其原因可能在于“十一五”规划时期由于政策带

来的产业结构在更长的跨期数据下不能被识别,同时,区域规划在2001—2010年未被识别,在2001—2016年被识别为显著的分区因素,这进一步验证出上述分区因素与国家政策具有高度契合性的结论具有稳健性。②空间权重矩阵的设定存在人为因素,因此分析结果会受到不同权重矩阵设定的影响<sup>[44]</sup>,为避免由于权重矩阵引起的偏误,本文将地理距离权重矩阵( $W_1$ )更换为经济地理距离权重矩阵( $W_3$ )和经济与地理距离的嵌套权重矩阵( $W_4$ )进行递归分区检验,其中,经济地理距离权重矩阵为地理距离权重矩阵与经济距离权重矩阵<sup>④</sup>的乘积,经济与地理距离的嵌套权重矩阵( $W_4$ )参照邵帅等<sup>[10]</sup>,按照 $W_4=0.5W_1+0.5W_2$ 构建,矩阵均进行标准化处理。实证结果表明,地理区位、行政级别、经济地带分区变量仍然是治理城市雾霾污染的显著性分区因素,但产业结构未被识别为显著的分区因素。从雾霾污染治理因素上看,上述两种方式下,稳健回归检验结果的影响程度和显著性水平略有区别,但

④ 经济距离权重矩阵( $W_2$ )由人均GDP差值的倒数构建。本文在经济距离权重矩阵下进行了递归分析检验,仍然筛选出地理区位、行政级别、经济地带、区域规划4个分区变量,同样证明本文结果具有稳健。但因为经济距离权重矩阵下Moran's I值未通过显著性检验,故本文未汇报结果。限于篇幅,其他稳健性检验结果有需要的读者可向作者索取。

对不同分区下雾霾污染的影响效果基本一致,表明本文研究结论稳健。

## 5 结论和启示

### 5.1 结论

本文借助机器学习,利用2001—2016年中国263个城市雾霾污染数据,应用空间自回归模型的决策树递归分析方法,实现了雾霾污染分区因素与治理因素的同步检验与交互分析,量化了不同分区因素对中国城市雾霾污染的影响层级与影响程度,进而揭示出多个分区因素影响下雾霾污染治理因素的影响机制及联防联控治理区域,破解了城市雾霾污染客观存在的差异与分区联防联控治理同效性之间的矛盾,为精准高效制定中国城市雾霾污染治理政策提供理论借鉴和决策参考。主要结论如下:

(1)地理区位、行政级别、产业结构、区域规划、经济地带是中国城市雾霾污染精准治理的主要分区因素。在“求同存异”的治理目标下,本文进一步提出由205个城市构成的4个雾霾污染联防联控治理区域,具体为:沿海治理区域、省会城市治理区域、二产占比 $>54.96\%$ 和 $\leq 54.96\%$ 的治理区域。

(2)产业结构和人口密度分别是加剧沿海城市和非沿海城市雾霾污染的首要因素。对非沿海非省会城市及非沿海非经济圈非东部城市而言,提高科技进步水平可以有效治理雾霾污染。“十五”计划、“十一五”规划期间非沿海城市以及“十二五”规划期间非沿海非经济圈城市经济发展水平加剧了雾霾污染,未越过库兹涅茨曲线的拐点;而沿海城市越过环境库兹涅茨曲线的拐点,经济发展水平的提高抑制了雾霾污染。

(3)中国城市雾霾污染治理分区因素的长期演变趋势与“五年规划”期间国家政策具有高度契合性。“十一五”重视产业结构优化调整,产业结构新增为“十一五”规划时期的分区变量。“十二五”提出优化发展格局,推动经济圈一体化发展及东部地区先行先试,区域规划和经济地带成为“十二五”规划时期的分区因素。

### 5.2 启示

基于上述研究结论,本文得到如下启示:

(1)在“求同存异”的目标下,制定跨城市群的中国城市雾霾污染精准治理的联防联控政策。各城市自然地理条件与社会经济因素的差异使得雾霾污染治理不是简单地合并相加,为避免属地治理高成本低效益的弊端,同时破解城市雾霾污染客观存在的差异与分区联防联控治理同效性之间的矛盾,需要根据分区因素建立城市雾霾污染联防联控治理区域。

(2)根据雾霾污染治理因素差异进行全局规划,精准实施城市雾霾治理策略。明确城市雾霾污染精准治理因素中存在的差异影响,并据此区别制定雾霾污染治理政策是高效减霾的重点。注重发挥科技手段在节能减排、污染防治领域的重要功能,加大政府绿色创新投入和绿色产业扶持,以点带面发挥科技治霾的示范作用。

(3)在生态优先,绿色发展的要求下,一方面可将减排评估纳入政府官员业绩评价体系,完善减霾治理及监督评估体系,提高政府的公信力与环境污染惩处力度。另一方面利用大数据分析技术,提高信息预警能力,借助科技手段提高国家政策的预见性,保证国家政策实施的连贯性,最终取得中国城市雾霾污染精准治理攻坚战的胜利。

## 参考文献(References):

- [1] 邵帅,李欣,曹建华.中国的城市化推进与雾霾治理[J].经济研究,2019,54(2):148-165.[Shao S, Li X, Cao J H. Urbanization promotion and haze pollution governance in China[J]. Economic Research Journal, 2019, 54(2): 148-165.]
- [2] Yang X C, Jiang L, Zhao W J, et al. Comparison of ground-based PM<sub>2.5</sub> and PM<sub>10</sub> concentrations in China, India, and the U. S.[J]. International Journal of Environmental Research and Public Health, 2018, 15(7): 1382.
- [3] 陈世强,张航,齐莹,等.黄河流域雾霾污染空间溢出效应与影响因素[J].经济地理,2020,40(5):40-48.[Chen S Q, Zhang H, Qi Y, et al. Spatial spillover effect and influencing factors of haze pollution in the Yellow River Basin[J]. Economic Geography, 2020, 40(5): 40-48.]
- [4] 李光勤,秦佳虹,何仁伟.中国大气PM<sub>2.5</sub>污染演变及其影响因素[J].经济地理,2018,38(8):11-18.[Li G Q, Qin J H, He R W. Spatial-temporal evolution and influencing factors of China's PM<sub>2.5</sub> pollution[J]. Economic Geography, 2018, 38(8): 11-18.]



2021年5月

- [5] Zhang X D, Chen L L, Yuan R S. Effect of natural and anthropic factors on the spatiotemporal pattern of haze pollution control of China[J]. Journal of Cleaner Production, 2019, DOI: 10.1016/j.jclepro.2019.119531.
- [6] Li H, Zhang M, Li C, et al. Study on the spatial correlation structure and synergistic governance development of the haze emission in China[J]. Environmental Science and Pollution Research, 2019, 26(12): 12136–12149.
- [7] Chen Z Y, Cai J, Gao B B, et al. Detecting the causality influence of individual meteorological factors on local PM<sub>2.5</sub> concentration in the Jing-Jin-Ji Region[J]. Scientific Reports, 2017, 7(1): 40735.
- [8] 刘海猛, 方创琳, 黄解军, 等. 京津冀城市群大气污染的时空特征与影响因素解析[J]. 地理学报, 2018, 73(1): 177–191. [Liu H M, Fang C L, Huang J J, et al. The spatial-temporal characteristics and influencing factors of air pollution in Beijing-Tianjin-Hebei Urban Agglomeration[J]. Acta Geographica Sinica, 2018, 73(1): 177–191.]
- [9] 陈诗一, 陈登科. 雾霾污染、政府治理与经济高质量发展[J]. 经济研究, 2018, 53(2): 20–34. [Chen S Y, Chen D K. Air pollution, government regulations and high-quality economic development [J]. Economic Research Journal, 2018, 53(2): 20–34.]
- [10] 邵帅, 李欣, 曹建华, 等. 中国雾霾污染治理的经济政策选择: 基于空间溢出效应的视角[J]. 经济研究, 2016, 51(9): 73–88. [Shao S, Li X, Cao J H, et al. China's economic policy choices for governing smog pollution based on spatial spillover effects[J]. Economic Research Journal, 2016, 51(9): 73–88.]
- [11] 马丽梅, 张晓. 中国雾霾污染的空间效应及经济、能源结构影响[J]. 中国工业经济, 2014, (4): 19–31. [Ma L M, Zhang X. The spatial effect of China's haze pollution and the impact from economic change and energy structure[J]. China Industrial Economics, 2014, (4): 19–31.]
- [12] 刘华军, 裴延峰. 我国雾霾污染的环境库兹涅茨曲线检验[J]. 统计研究, 2017, 34(3): 45–54. [Liu H J, Pei Y F. An empirical test of the Environmental Kuznets Curve of China's haze pollution[J]. Statistical Research, 2017, 34(3): 45–54.]
- [13] 王敏, 黄滢. 中国的环境污染与经济增长[J]. 经济学(季刊), 2015, 14(2): 557–578. [Wang M, Huang Y. China's environmental pollution and economic growth[J]. China Economic Quarterly, 2015, 14(2): 557–578.]
- [14] 范建双, 周琳, 虞晓芬. 高速铁路和城市蔓延对雾霾污染的异质性影响[J]. 地理研究, 2021, 40(4): 1146–1164. [Fan J S, Zhou L, Yu X F. The heterogeneity effect of high-speed railway and urban sprawl on haze pollution[J]. Geographical Research, 2021, 40(4): 1146–1164.]
- [15] 马丽梅, 张晓. 区域大气污染空间效应及产业结构影响[J]. 中国人口·资源与环境, 2014, 24(7): 157–164. [Ma L M, Zhang X. Spatial effects of regional air pollution and the impact of industrial structure[J]. China Population, Resources and Environment, 2014, 24(7): 157–164.]
- [16] 祝德生, 景维民. 金融发展、空间溢出与雾霾污染: 基于广义空间两阶段最小二乘法的实证研究[J]. 经济问题探索, 2020, (8): 134–143. [Zhu D S, Jing W M. Financial development, spatial spillover and haze pollution: An empirical research based on Generalized Space Two-Stage Least Square Method[J]. Inquiry into Economic Issues, 2020, (8): 134–143.]
- [17] Wang Z B, Fang C L. Spatial-temporal characteristics and determinants of PM<sub>2.5</sub> in the Bohai Rim Urban Agglomeration[J]. Chemosphere, 2016, 148: 148–162.
- [18] 冷艳丽, 冼国明, 杜思正. 外商直接投资与雾霾污染: 基于中国省际面板数据的实证分析[J]. 国际贸易问题, 2015, (12): 74–84. [Leng Y L, Xian G M, Du S Z. Foreign direct investment and haze pollution: An empirical analysis based on provincial panel data[J]. Journal of International Trade, 2015, (12): 74–84.]
- [19] 庞雨蒙, 刘震, 潘雨晨. 财政科教支出与雾霾污染治理的空间关联效应[J]. 经济经纬, 2020, 37(6): 128–138. [Pang Y M, Liu Z, Pan Y C. Spatial relevance effect of fiscal technology and education expenditure on haze pollution[J]. Economic Survey, 2020, 37(6): 128–138.]
- [20] 戴宏伟, 回莹. 京津冀雾霾污染与产业结构、城镇化水平的空间效应研究[J]. 经济理论与经济管理, 2019, (5): 4–19. [Dai H W, Hui Y. The spatial effects of Beijing-Tianjin-Hebei Region's smog pollution, industrial structure and urbanization[J]. Economic Theory and Business Management, 2019, (5): 4–19.]
- [21] 刘华军, 孙亚男, 陈明华. 雾霾污染的城市间动态关联及其成因研究[J]. 中国人口·资源与环境, 2017, 27(3): 74–81. [Liu H J, Sun Y N, Chen M H. Dynamic correlation and causes of urban haze pollution[J]. China Population, Resources and Environment, 2017, 27(3): 74–81.]
- [22] 王少剑, 高爽, 陈静. 基于GWR模型的中国城市雾霾污染影响因素的空间异质性研究[J]. 地理研究, 2020, 39(3): 651–668. [Wang S J, Gao S, Chen J. Spatial heterogeneity of driving factors of urban haze pollution in China based on GWR model[J]. Geographical Research, 2020, 39(3): 651–668.]
- [23] 陈诗一, 张云, 武英涛. 区域雾霾联防联控治理的现实困境与政策优化: 雾霾差异化成因视角下的方案改进[J]. 中共中央党校学报, 2018, 22(6): 109–118. [Chen S Y, Zhang Y, Wu Y T. Study on practical difficulties and policy optimization of regional joint prevention and control of smog: Scheme improvement from the perspective of smog cause differentiation[J]. Journal of The Party School of the Central Committee of the C.P.C. (Chinese Academy of Governance), 2018, 22(6): 109–118.]



- [24] 刘华军, 雷名雨. 中国雾霾污染联防联控的区域范围优化分析[J]. 人文杂志, 2019, (2): 121-128. [Liu H J, Lei M Y. Optimization analysis on the regional scope of joint prevention and control of haze pollution in China[J]. The Journal of Humanities, 2019, (2): 121-128.]
- [25] Athey S, Imbens G W. Machine learning methods economists should know about[J]. Annual Review of Economics, 2019, 11(1): 685-725.
- [26] Mullainathan S, Spiess J. Machine learning: An applied econometric approach[J]. Journal of Economic Perspectives, 2017, 31(2): 87-106.
- [27] Varian H R. Big data: New tricks for econometrics[J]. Journal of Economic Perspectives, 2014, 28(2): 3-27.
- [28] Athey S, Imbens G W. Recursive partitioning for heterogeneous causal effects[J]. Proceedings of The National Academy of Sciences, 2016, 113(27): 7353-7360.
- [29] Athey S, Imbens G W. The state of applied econometrics: Causality and policy evaluation[J]. Journal of Economic Perspectives, 2017, 31(2): 3-32.
- [30] Asher S, Nekipelov D, Novosad P, et al. Moment Forests[EB/OL]. (2018-02-19) [2020-03-11]. <https://cpb-us-west-2-juc1ugur1qwqqq04.stackpathdns.com/sites.wustl.edu/dist/5/501/files/2018/02/momentTrees-2m5z1ds.pdf>.
- [31] Zeileis A, Hothorn T, Hornik K. Model-based recursive partitioning[J]. Journal of Computational and Graphical Statistics, 2008, 17(2): 492-514.
- [32] Athey S, Imbens G W. The econometrics of randomized experiments[J]. Handbook of Economic Field Experiments, 2017, 1: 73-140.
- [33] 陈诗一, 王建民. 中国城市雾霾治理评价与政策路径研究: 以长三角为例[J]. 中国人口·资源与环境, 2018, 28(10): 71-80. [Chen S Y, Wang J M. Evaluation and policy innovation of urban haze governance in China: Taking the Yangtze River Delta area as an example[J]. China Population, Resources and Environment, 2018, 28(10): 71-80.]
- [34] Wu Z, Zhang S. Study on the spatial-temporal change characteristics and influence factors of fog and haze pollution based on GAM[J]. Neural Computing and Applications, 2019, 31: 1619-1631.
- [35] Sela R J, Simonoff J S. RE-EM trees: A data mining approach for longitudinal and clustered data[J]. Machine Learning, 2012, 86(2): 169-207.
- [36] Andrews D W K. Tests for parameter instability and structural change with unknown change point[J]. Econometrica, 1993, 61(4): 821-856.
- [37] Wang Y P, Komonpipat S. Revisiting the Environmental Kuznets Curve of PM<sub>2.5</sub> concentration: Evidence from prefecture-level and above cities of China[J]. Environmental Science and Pollution Research, 2020, 27(2): 1-13.
- [38] 孙亚男, 杨名彦. 中国绿色全要素生产率的俱乐部收敛及地区差距来源研究[J]. 数量经济技术经济研究, 2020, 37(6): 47-69. [Sun Y N, Yang M Y. Research on club convergence and the sources of regional gaps of green total factor productivity in China[J]. The Journal of Quantitative and Technical Economics, 2020, 37(6): 47-69.]
- [39] 张义, 王爱君, 黄寰. 权力协同对中国雾霾防治的影响研究[J]. 经济与管理研究, 2019, 40(12): 96-113. [Zhang Y, Wang A J, Huang H. Research on the influence of power synergy on haze prevention and control in China[J]. Research on Economics and Management, 2019, 40(12): 96-113.]
- [40] 熊欢欢, 梁龙武, 曾赠, 等. 中国城市PM<sub>2.5</sub>时空分布的动态比较分析[J]. 资源科学, 2017, 39(1): 136-146. [Xiong H H, Liang L W, Zeng Z, et al. Dynamic analysis of PM<sub>2.5</sub> spatial-temporal characteristics in China[J]. Resources Science, 2017, 39(1): 136-146.]
- [41] Liao S M, Wang D, Liang Z J, et al. Spatial spillover effect and sources of city-level haze pollution in China: A case study of Guangdong Provinces[J]. Polish Journal of Environmental Studies, 2020, 29(5): 3213-3223.
- [42] 刘华军, 彭莹. 雾霾污染区域协同治理的“逐底竞争”检验[J]. 资源科学, 2019, 41(1): 185-195. [Liu H J, Peng Y. “Race to the bottom” test of collaborative management in haze pollution area[J]. Resources Science, 2019, 41(1): 185-195.]
- [43] 李建明, 罗能生. 1998-2015年长江中游城市群雾霾污染时空演变及协同治理分析[J]. 经济地理, 2020, 40(1): 76-84. [Li J M, Luo N S. Temporal and spatial evolution and coordination governance of haze pollution in urban agglomerations in the Middle Reaches of the Yangtze River from 1998 to 2015[J]. Economic Geography, 2020, 40(1): 76-84.]
- [44] Seya H, Yamagata Y, Tsutsumi M. Automatic selection of a spatial weight matrix in spatial econometrics: Application to a spatial hedonic approach[J]. Regional Science and Urban Economics, 2013, 43(3): 429-444.

## Precise governance of haze pollution based on machine learning

SUN Yanan, FEI Jinhua

(School of Business Administration, Shandong University of Finance and Economics, Jinan 250014, China)

**Abstract:** The implementation of precise governance of haze pollution is an important measure to deal with haze pollution. Using the PM<sub>2.5</sub> of 263 Chinese cities from 2001 to 2016, this study constructed a recursive partitioning analysis method of decision tree based on machine learning, quantified the interactions between haze pollution zoning factors and governance factors, and then identified the haze pollution governance regions and governance factors. The results show that: (1) Geographic location, administrative level, industrial structure, regional planning, and economic zone are the main factors to identify the precise governance regions of urban haze pollution in China. Based on these factors, four types of urban haze pollution governance regions were identified. (2) In coastal cities, increasing the level of economic development and reducing the proportion of the secondary industry are conducive to improving the level of haze pollution governance. In non-coastal cities, reducing population density is helpful for alleviating haze pollution. Furthermore, in non-coastal and non-provincial capital cities, improving the level of science and technology progress is beneficial for controlling haze pollution. (3) The long-term trend of urban haze pollution governance zoning factors in China is highly consistent with the national policies during the “Five-Year Plan” periods. Therefore, scientific policies and precise regulation are conducive to the Precise governance of haze pollution.

**Key words:** machine learning; urban haze pollution; precise governance; influencing factors; recursive partitioning