

控制与决策

Control and Decision

基于面板数据灰色面角关联模型的区域PM_{2.5}健康风险相关因素识别

安艺萌, 党耀国, 杨德岭, 王俊杰, 冯宇

引用本文:

安艺萌, 党耀国, 杨德岭, 等. 基于面板数据灰色面角关联模型的区域PM_{2.5}健康风险相关因素识别[J]. 控制与决策, 2025, 40(3): 899–908.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.1748>

您可能感兴趣的其他文章

Articles you may be interested in

[基于分类的多策略预测方法求解动态多目标优化问题](#)

[Classification-based multi-strategy prediction method for dynamic multiobjective optimization problems](#)

控制与决策. 2021, 36(7): 1569–1580 <https://doi.org/10.13195/j.kzyjc.2019.1320>

[嵌入重采样技术的C4.5决策树集成分类算法的临床医学预测](#)

Clinical prediction of C4.5 decision tree classification algorithm with embedded resampling technique

控制与决策. 2021, 36(6): 1342–1350 <https://doi.org/10.13195/j.kzyjc.2019.1247>

[基于前景优化和概率估计的显著性检测](#)

Saliency detection based on foreground optimization and probability estimation

控制与决策. 2021, 36(4): 917–924 <https://doi.org/10.13195/j.kzyjc.2019.1101>

[基于近端强化学习的股价预测方法](#)

Method of stock prices forecast based on proximal reinforcement learning

控制与决策. 2021, 36(4): 967–973 <https://doi.org/10.13195/j.kzyjc.2019.1245>

[基于多维泰勒网的超前d步预测模型](#)

d-step-ahead predictive model based on multi-dimensional Taylor network

控制与决策. 2021, 36(2): 345–354 <https://doi.org/10.13195/j.kzyjc.2019.0722>

基于面板数据灰色面角关联模型的区域PM_{2.5} 健康风险相关因素识别

安艺萌^{1,2}, 党耀国¹, 杨德岭³, 王俊杰^{1†}, 冯宇^{1,4}

- 南京航空航天大学经济与管理学院, 南京 211106;
- 贝尔法斯特女王大学 IEEE 计算机科学学院, 贝尔法斯特 BT9 5AH;
- 许昌电气职业学院经济管理系, 许昌 461002;
- 德蒙福特大学 计算机科学学院, 莱斯特 LE1 9BH)

摘要: 准确识别大气污染下健康风险的主要经济-社会要素, 对于环境质量管理与健康风险防控具有重要意义。通过对长江三角洲 (YRD) 城市群 PM_{2.5} 污染进行空间相关性分析, 发现城市间 PM_{2.5} 污染存在显著的正向空间依赖性。鉴于区域大气污染具有空间临近效应和时间趋势特征, 针对现有灰色关联模型未能将面板数据整体特征与不同维度趋势特征相整合的问题, 提出面板数据灰色面角关联模型 (GDAIM), 以识别空间临近区域 PM_{2.5} 健康风险的主要影响因素。通过构建三维数据网络曲面并研究曲面夹角关系, 构建面角差异性测度公式, 给出 GDAIM 建模过程。以 YRD 地区 PM_{2.5} 污染“高-高”聚集区为研究对象, 运用 GDAIM 梳理该区域 PM_{2.5} 健康风险与经济-社会因素间的关联关系。结果表明, 颗粒物和二氧化硫排放是影响 YRD 地区 PM_{2.5} 健康风险的主要因素, 其次是乡村人口数量和氮氧化物排放。研究结果可为大气污染联防联控战略制定提供理论支持。

关键词: 面板数据; 灰色关联分析; PM_{2.5} 污染; 健康风险; 空间相关性; 长江三角洲地区

中图分类号: N941.5 文献标志码: A

DOI: 10.13195/j.kzyjc.2023.1748

引用格式: 安艺萌, 党耀国, 杨德岭, 等. 基于面板数据灰色面角关联模型的区域 PM_{2.5} 健康风险相关因素识别 [J]. 控制与决策, 2025, 40(3): 899-908.

Identification of regional PM_{2.5} health risk factors based on panel data grey dihedral angle incidence model

AN Yi-meng^{1,2}, DANG Yao-guo¹, YANG De-ling³, WANG Jun-jie^{1†}, FENG Yu^{1,4}

- College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;
- School of Electronics, Electrical Engineering & Computer Science, Queen's University Belfast, Belfast BT9 5AH, UK;
- Department of Economic and Management, Xuchang Electrical Vocational College, Xuchang 461002, China;
- School of Computer Science and Informatics, De Montfort University, Leicester LE1 9BH, UK)

Abstract: Identifying the main economic-social influencing factors of health risks of PM_{2.5} pollution is significant for the development of environmental quality management and health risk prevention. Through spatial correlation analysis of PM_{2.5} pollution in the Yangtze River Delta (YRD) region, significant positive spatial dependence of PM_{2.5} pollution is observed. Given the spatial proximity and temporal trend effects of regional air pollution, and the limitations of the existing grey incidence models in integrating different dimensions trend characteristics of panel data, this paper proposes a grey dihedral angle incidence model (GDAIM) to identify the main influencing factors of regional PM_{2.5} pollution health risk. By constructing the three-dimensional data matrix network structures and studying the angle relationship between two network planes, a difference-measure method is developed, and the modeling steps of the GDAIM are given. Focusing on the 'high-high' clustering area of PM_{2.5} pollution in the YRD, the GDAIM is used to

收稿日期: 2023-12-15; 录用日期: 2024-05-09.

基金项目: 国家自然科学基金项目 (72271120, 72001107, 71771119, 51979106); 中国博士后科学基金项目 (2020 T130297, 2019M660119); 教育部人文社会科学研究青年基金项目 (19YJC630167); 中央高校基本科研业务费专项资金项目 (NP2022104); 江苏省研究生科研与实践创新计划项目 (KYCX23_0404); 南京航空航天大学博士生跨学科创新基金项目 (KXKCXJJ202303); 中国国家留学基金资助项目。

†通信作者. E-mail: wangjj@nuaa.edu.cn.

本文附带电子附录文件, 可登录本刊官网该文“资源附件”区自行下载阅览.

examine the association between $PM_{2.5}$ health risk and the economic-social factors. Results show that particulate matter and sulfur dioxide emissions are the main factors, followed by rural population and nitrogen oxide emissions. The findings can provide theoretical support for joint prevention and control of air pollution.

Keywords: panel data; grey incidence analysis; $PM_{2.5}$ pollution; health risks; spatial correlation; Yangtze River Delta

0 引言

大气污染是人类健康的主要威胁之一,长期暴露于 $PM_{2.5}$ 大气污染会对人体健康产生有害影响^[1]. 已有大量学者利用流行病学和统计学等方法,对 $PM_{2.5}$ 污染带来的经济损失与健康风险进行了量化和评估. 例如,文献 [2] 利用离线分析方法,调查了省级城市的颗粒物来源. 文献 [3] 通过离子色谱法和电感耦合法,研究了 $PM_{2.5}$ 中水溶性离子和金属元素对血常规参数的影响. 文献 [4] 运用一般均衡性模型对 $PM_{2.5}$ 污染的健康经济效益进行了评估. 现有 $PM_{2.5}$ 健康风险的研究大多基于病理学与统计分析视角,探究污染物对人体的致病机理,从管理学视角识别区域 $PM_{2.5}$ 污染下人群健康风险的经济和社会影响因素的研究较少. 同时,考虑到城市尺度的卫生健康登记系统近年来才逐步完善,并且经济与社会数据往往为年度和季度数据,样本量较小,因此在研究大气污染与健康风险关系时存在信息结构不完整、数据信息不充分等问题. 灰色系统理论以有限信息源的不确定性问题为研究对象,通过挖掘已有信息,实现对系统的运行规律进行正确地描述和控制.

灰色关联模型 (GIMs) 是灰色系统理论的重要分支,是运用关联系数测度因素变量与系统特征变量之间关联程度的一种因素分析方法^[5]. GIMs 已广泛应用于驱动要素识别^[6-7]、排序研究^[8]、聚类分析^[9] 和多目标优化^[10] 等领域中. 从建模数据类型来看, GIMs 的研究对象从时间序列扩展到面板数据. 就建模机制而言, GIMs 构建的关键之一是如何测度两个研究对象之间的差异性特征.

在测度两面板数据差异性特征时,一种思路是将面板数据从时间维和截面维进行拆分,分别测度两面板数据两个维度的差异性,再将二者进行组合. 刘震等^[11] 将面板数据网格拆分为时间维和截面维的线段,利用线段在空间中的斜率差值,构建了灰色网格关联系数. 崔立志等^[12] 从截面纬衡量指标发展速度差异性,从时间维衡量了增速差异性,构建了面板数据矩阵相似关联模型. 罗党等^[13] 分别利用斜率差和均值之比构建了时间维横向关联度和截面维纵向关联度,通过权重将二者组合构建了灰色 B 型关联模型 (BGIM). Sun 等^[14] 利用向量夹角测度时间维趋

势相似度,利用数据增量和空间距离测度截面维相似程度,通过加权得到综合关联度. 党耀国等^[15] 在时间维量化指标的横向发展特征,在对象维量化指标的纵向振荡特征,构建了灰色向量模关联模型 (GVMIM). 上述 GIMs 均存在“维度剥离”问题,即先将面板数据按照不同维度进行拆分,分别计算时间维和截面维两类关联度,然后通过设置偏好权重将二者加权求和. 这种方法在识别两个面板数据两个维度的相似性特征时是依次进行的,即先测度一个维度的相似性特征,再测度另一个维度的相似性特征,导致在测度时间维相似性时忽略了截面维相似性,反之亦然,面板数据整体特征被忽略.

另一种思路是直接比较两个面板数据的形状特征,从整体视角分析两面板数据的相关关系. 张可等^[16] 将两个面板数据的相似度转化为两个曲面形状的接近程度,利用两面板数据曲顶柱体体积的差值,构建了扩展灰色绝对关联模型 (EGAIM). 吴鸿华等^[17] 利用两面板数据的四面体有向体积的差值,构建了灰色四面体网格关联模型 (TGIM). 从面板数据的整体特征出发,可以较为清晰地描述单个面板数据的几何特征,但上述 GIMs 在识别两个面板数据结构特征时是非同步的,忽略了两个面板数据在某一段时间内的发展趋势差异性. 综合上述分析,由于面板数据同时蕴含截面纬度和时间维度的特征,在构建面板数据关联分析模型时,如何充分利用面板数据两个维度的信息,在不丢失时间趋势特征的前提下,将面板数据两个维度的特征相整合,而非简单地线性加权处理,还有待进一步研究.

Tobler 第一地理定律^[18] 表明,临近区域的变量值之间往往存在一定的相关性,距离越近空间相关性越强. 受气流传播^[19] 和空间溢出效应^[20] 影响,城市群大气污染存在着显著的空间相关性. 城市群由许多在经济和社会上密切相关的行政单位组成,某城市由大气污染引发的健康风险不仅受到当地生产活动和环境法规等的影响,还受到来自临近城市的影响. 因此,在识别区域 $PM_{2.5}$ 健康风险 (ΔE) 影响因素的研究中,在考虑相关因素与 ΔE 时间趋势相似性特征的同时,还要考虑临近城市间空间相关性的时间趋势特征. 鉴于此,本文为研究长江三角洲

(YRD)地区城市群PM_{2.5}污染健康风险的影响因素, 构建基于三维数据矩阵的灰色面角关联模型(GDAIM). 该模型同步整合了面板数据在截面维度的相关性特征和时间维度的趋势特征, 通过构建三维数据矩阵并利用其在几何坐标系中的平面性质以及位置关系, 构建灰色面角差异性测度, 为多变量复杂系统的关联分析问题提供一种新的思路.

不同于现有模型, GDAIM 利用同一个差异性测度整合了两面板数据在两个维度的差异性特征, 即对截面维度和时间维度的差异性测度是同步进行的, 而非单独测度再进行加和, 解决了现有面板数据综合GIMs中“维度剥离”的问题. 此外, GDAIM 在比较两面板数据形状特征的同时, 不仅考察了时间维的变化特征, 还考察了截面维相关性的发展特征, 解决了现有面板数据GIMs中因注重考察面板数据形状相似性而忽略面板数据两维度趋势特征的问题. 本文在PM_{2.5}健康风险的流行病学研究基础上, 利用GDAIM模型对YRD地区的PM_{2.5}健康风险与经济社会因素进行相关性分析, 以期从管理学视角识别影响YRD地区PM_{2.5}污染下健康风险的相关因素.

1 三维数据网络曲面构建

定义1 设系统中有 w 个观测指标、 m 个观测对象和 n 个离散的连续观测时刻点, 记指标 $s(s = 1, 2, \dots, w)$ 的第 $l(l = 1, 2, \dots, m)$ 个观测对象在时刻 $t(t = 1, 2, \dots, n)$ 处的观测值为 $x_s(l, t)$, 则称

$$X_s(l, t) = \begin{bmatrix} x_s(1, 1) & x_s(1, 2) & \dots & x_s(1, n) \\ x_s(2, 1) & x_s(2, 2) & \dots & x_s(2, n) \\ \vdots & \vdots & \ddots & \vdots \\ x_s(m, 1) & x_s(m, 2) & \dots & x_s(m, n) \end{bmatrix}$$

为指标 s 的面板数据; 称

$$X(l, t, x_s) = \begin{bmatrix} (1, 1, x_s(1, 1)) & \dots & (1, n, x_s(1, n)) \\ (2, 1, x_s(2, 1)) & \dots & (2, n, x_s(2, n)) \\ \vdots & \vdots & \vdots \\ (m, 1, x_s(m, 1)) & \dots & (m, n, x_s(m, n)) \end{bmatrix}$$

为 $X_s(l, t)$ 的三维数据矩阵.

三维数据矩阵 $X(l, t, x_s)$ 将面板数据 $X_s(l, t)$ 中 $m \times n$ 个观测值 $x_s(l, t)$ 投影为三维空间中的 $m \times n$ 个散点 $(l, t, x_s(l, t))$. 三维数据矩阵可为面板数据关联模型的构建提供便利, 基于解析几何理论明晰建模机理. 如图1所示, 在 $X(l, t, x_s)$ (蓝色)中选取临近的4个点 $S_1(l, t, x_s(l, t)), S_2(l, t + 1, x_s(l, t + 1)), S_3(l + 1, t, x_s(l + 1, t)), S_4(l + 1, t + 1, x_s(l + 1, t + 1))$

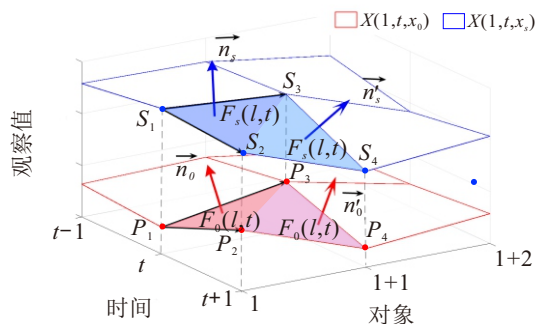


图1 指标 I_0 和 I_s 在区域 E_{lt} 内的平面法向量

1)), $t = 1, 2, \dots, n - 1, l = 1, 2, \dots, m - 1$. 其中: S_1, S_2, S_3 可确定唯一平面 $F_s(l, t)$; S_2, S_3, S_4 确定唯一平面 $F'_s(l, t)$.

定义2 在三维数据矩阵 $X(l, t, x_s)$ 中, 平面 $F_s(l, t)$ 由 $\{S_1(l, t, x_s(l, t)), S_2(l, t + 1, x_s(l, t + 1)), S_3(l + 1, t, x_s(l + 1, t))\}$ 三点确定, 平面 $F'_s(l, t)$ 由 $\{S_2(l, t + 1, x_s(l, t + 1)), S_3(l + 1, t, x_s(l + 1, t)), S_4(l + 1, t + 1, x_s(l + 1, t + 1))\}$ 三点确定, 称

$$G_s = \{G_s(l, t) = \angle(F_s(l, t), F'_s(l, t)) \mid l = 1, 2, \dots, m - 1, t = 1, 2, \dots, n - 1\} \quad (1)$$

为 $X(l, t, x_s)$ 的三维数据网络曲面. 其中: $G_s(l, t)$ 为区域 $E_{lt} = \{(x, y, z) \mid x \in [t, t + 1], y \in [l, l + 1]\}$ 内的网络单元, 可视为由点 P_1, P_4 和直线 P_2P_3 组成的二面角; $X(l, t, x_s)$ 可视为由 $(m - 1) \times (n - 1)$ 个网络单元 $G_s(l, t)$ 组成的网络曲面. 二面角 $G_s(l, t)$ 由 $F_s(l, t)$ 过 P_1 的半平面和 $F'_s(l, t)$ 过 P_4 的半平面组成. 由于平面与其半平面几何方程相同, 不特别对二者区分. 式(1)给出了三维数据网络结构的几何表示形式.

2 基于三维数据矩阵的灰色面角关联模型

2.1 模型构建

设 $\{O_l \mid l = 1, 2, \dots, m\}$ 为对象集, $\{I_0, I_1, \dots, I_w\}$ 为指标集, 其中 I_0 为特征指标, $I_s(s = 1, 2, \dots, w)$ 为因素指标, 其三维数据矩阵分别为 $X(l, t, x_0)$ 和 $X(l, t, x_s)$. 如定义2所述, 在网络曲面 G_s 中(蓝色), 存在 S_1, S_2, S_3, S_4 组成曲面单元 $G_s(l, t) = \angle(F_s(l, t), F'_s(l, t))$. 同理, 在网络曲面 G_0 (红色)中, 存在临近4点 $P_1(l, t, x_0(l, t)), P_2(l, t + 1, x_0(l, t + 1)), P_3(l + 1, t, x_0(l + 1, t)), P_4(l + 1, t + 1, x_0(l + 1, t + 1))$ 组成的曲面单元 $G_0(l, t) = \angle(F_0(l, t), F'_0(l, t))$, 如图1所示.

对于指标 $I_0, \overrightarrow{P_1P_2}$ 为对象 O_l 在 $[t, t + 1]$ 内的时间发展趋势; $\overrightarrow{P_1P_3}$ 蕴含了对象 O_l 与 O_{l+1} 在 t 时刻的

相关关系. 结合指标 I_s 进行分析, 若 $\overrightarrow{P_1P_2} \parallel \overrightarrow{S_1S_2}$, 则认为 $[t, t+1]$ 时段内, O_l 在指标 I_0 和 I_s 上表现出相同的时间变化趋势; 若 $\overrightarrow{P_1P_3} \parallel \overrightarrow{S_1S_3}$, 则认为 t 时刻, 在指标 I_0 和 I_s 两个方面, O_l 与 O_{l+1} 具有相同的相关关系. 特别地, 若 $\overrightarrow{P_1P_3} \parallel \overrightarrow{S_1S_3}$ 且 $\overrightarrow{P_2P_4} \parallel \overrightarrow{S_2S_4}$, 则认为对象 O_l 与 O_{l+1} 在指标 I_0 和 I_s 上表现出相同的相关关系, 且这种相关关系不随时间的推移发生变化. 这一特征对于临近区域大气污染健康风险相关因素的识别是必要的. 若 O_1 与 O_2 为两临近城市, 则考虑到污染物的扩散和溢出效应, 区域 O_1 污染物浓度增加会影响临近区域 O_2 的污染物浓度, 将此现象视为 O_1 与 O_2 在大气污染 I_s 上呈现的相关关系; 同时, O_1 区域健康风险增加, 临近城市 O_2 的健康风险也发生变化, 将此现象视为 O_1 、 O_2 在健康风险指标 I_0 上呈现的相关关系; 临近城市 O_1 与 O_2 在 I_s 和 I_0 上表现出的相关关系均与空间临近效应有关, 且在稳定系统中持续存在. 这种情况下可以得到 $F_0(l, t) \parallel F_s(l, t)$ 且 $F'_0(l, t) \parallel F'_s(l, t)$, 此时认为指标 I_0 和 I_s 在时间和对象两个维度上均表现出较强的相关性. 当 $\{P_1, P_2, P_3, P_4, S_1, S_2, S_3, S_4\}$ 中任意一点发生变化, 这种强相关性会被破坏, $F_0(l, t)$ 与 $F_s(l, t)$ 或 $F'_0(l, t)$ 与 $F'_s(l, t)$ 的空间位置关系由平行变为相交.

因此, 可利用 $F_0(l, t), F_s(l, t), F'_0(l, t), F'_s(l, t)$ 之间的位置关系, 测度 $X(l, t, x_0)$ 和 $X(l, t, x_s)$ 在区域 E_{lt} 内的差异性. 这样不仅能够通过向量关系来反映面板数据在各个维度上的趋势相似性特征, 而且可以通过网络曲面的几何形态, 直观地表现面板数据的整体相似性特征. 如何利用网络曲面的空间几何特征, 从整体视角构建差异性测度是本研究的关键.

定理 1 设 $X(l, t, x_0)$ 和 $X(l, t, x_s)$ 分别为特征指标 I_0 和因素指标 I_s 的三维数据矩阵, $G_0(l, t) = \angle(F_0(l, t), F'_0(l, t))$ 和 $G_s(l, t) = \angle(F_s(l, t), F'_s(l, t))$ 分别为 $X(l, t, x_0)$ 和 $X(l, t, x_s)$ 在区域 E_{lt} 内的行为曲面单元. 取平面 $F_0(l, t)$ 、 $F_s(l, t)$ 、 $F'_0(l, t)$ 和 $F'_s(l, t)$ 的法向量分别记为 $\vec{n}_0, \vec{n}_s, \vec{n}'_0$ 和 \vec{n}'_s , 则 \vec{n}_0 与 \vec{n}_s 向量夹角 θ , 以及 \vec{n}'_0 与 \vec{n}'_s 向量夹角 θ' 满足如下等式:

$$\cos \theta = \cos \langle \vec{n}_0, \vec{n}_s \rangle = \frac{A_0 \cdot A_s + B_0 \cdot B_s + C_0 \cdot C_s}{\sqrt{A_0^2 + B_0^2 + C_0^2} \sqrt{A_s^2 + B_s^2 + C_s^2}}; \quad (2)$$

$$\cos \theta' = \cos \langle \vec{n}'_0, \vec{n}'_s \rangle = \frac{A'_0 \cdot A'_s + B'_0 \cdot B'_s + C'_0 \cdot C'_s}{\sqrt{A'^2_0 + B'^2_0 + C'^2_0} \sqrt{A'^2_s + B'^2_s + C'^2_s}}. \quad (3)$$

其中

$$\begin{cases} A_0 = x_0(l+1, t) - x_0(l, t), \\ B_0 = x_0(l, t+1) - x_0(l, t), \\ A_s = x_s(l+1, t) - x_s(l, t), \\ B_s = x_s(l, t+1) - x_s(l, t), \\ C_0 = C_s = -1; \end{cases} \quad (4)$$

$$\begin{cases} A'_0 = x_0(l, t+1) - x_0(l+1, t+1), \\ B'_0 = x_0(l+1, t) - x_0(l+1, t+1), \\ A'_s = x_s(l, t+1) - x_s(l+1, t+1), \\ B'_s = x_s(l+1, t) - x_s(l+1, t+1), \\ C'_0 = C'_s = -1. \end{cases} \quad (5)$$

定义 3 设 $G_0(l, t)$ 和 $G_s(l, t)$ 分别为 $X(l, t, x_0)$ 和 $X(l, t, x_s)$ 在区域 E_{lt} 内的行为曲面单元, $\cos \langle \vec{n}_0, \vec{n}_s \rangle$ 和 $\cos \langle \vec{n}'_0, \vec{n}'_s \rangle$ 如定理 1 所述, 记

$$d_{0s}(l, t) = \text{sgn}_{0s}(l, t) \cdot \cos \langle \vec{n}_0, \vec{n}_s \rangle. \quad (6)$$

$$\text{sgn}_{0s}(l, t) = \begin{cases} 1, & (\overrightarrow{P_1H} \cdot \vec{n}_0)(\overrightarrow{P_1H} \cdot \vec{n}_s) \geq 0; \\ -1, & (\overrightarrow{P_1H} \cdot \vec{n}_0)(\overrightarrow{P_1H} \cdot \vec{n}_s) < 0. \end{cases} \quad (7)$$

$$d'_{0s}(l, t) = \text{sgn}'_{0s}(l, t) \cdot \cos \langle \vec{n}'_0, \vec{n}'_s \rangle. \quad (8)$$

$$\text{sgn}'_{0s}(l, t) = \begin{cases} 1, & (\overrightarrow{P_4H} \cdot \vec{n}'_0)(\overrightarrow{P_4H} \cdot \vec{n}'_s) \geq 0; \\ -1, & (\overrightarrow{P_4H} \cdot \vec{n}'_0)(\overrightarrow{P_4H} \cdot \vec{n}'_s) < 0. \end{cases} \quad (9)$$

其中: $H\left(l + \frac{1}{2}, t + \frac{1}{2}, \frac{1}{2}(x_s(l, t+1) + x_s(l+1, t))\right)$ 为 $S_2(l, t+1, x_s(l, t+1))$ 与 $S_3(l+1, t, x_s(l+1, t))$ 的中点; 称 $d_{0s}(l, t)$ 为平面 $F_0(l, t)$ 与 $F_s(l, t)$ 的面角差异性测度; $d'_{0s}(l, t)$ 为 $F'_0(l, t)$ 与 $F'_s(l, t)$ 的面角差异性测度; $d_{0s}(l, t)$ 和 $d'_{0s}(l, t)$ 统称为面角差异性测度.

将由过点 P_1 的半平面 $F_0(l, t)$ 与过点 S_1 的半平面 $F_s(l, t)$ 组成的二面角记为 α , 则定义 3 中面角差异性测度 $d_{0s}(l, t)$ 为 α 的余弦值, 即 $d_{0s}(l, t) = \cos \alpha$; 同理可记 $d'_{0s}(l, t) = \cos \alpha'$, α' 为由过点 P_4 的半平面 $F'_0(l, t)$ 与过点 S_4 的半平面 $F'_s(l, t)$ 组成的二面角. 在区域 E_{lt} 内, α 与 α' 越小, 行为曲面单元 $G_0(l, t)$ 与 $G_s(l, t)$ 越相似, $d_{0s}(l, t)$ 与 $d'_{0s}(l, t)$ 越大; 特别地, 当 $\alpha = \alpha' = 0$ 时, $F_0(l, t) \parallel F_s(l, t)$, 且 $F'_0(l, t) \parallel F'_s(l, t)$, 此时行为曲面单元 $G_0(l, t)$ 与 $G_s(l, t)$ 的几何特征完全相同, $d_{0s}(l, t) = d'_{0s}(l, t) = 1$.

推论 1 记 $d_{0s}(l, t) = \cos \alpha$, θ 为向量 \vec{n}_0 与 \vec{n}_s 的夹角; $d'_{0s}(l, t) = \cos \alpha'$, θ' 为向量 \vec{n}'_0 与 \vec{n}'_s 的夹角, $\alpha, \alpha' \in [0, \pi)$, 则有:

- 1) 当 $\overrightarrow{P_1H} \cdot \vec{n}_0$ 与 $\overrightarrow{P_1H} \cdot \vec{n}_s$ 同号时, $\alpha = \theta$;
- 2) 当 $\overrightarrow{P_1H} \cdot \vec{n}_0$ 与 $\overrightarrow{P_1H} \cdot \vec{n}_s$ 异号时, $\alpha = \pi - \theta$;
- 3) 当 $\overrightarrow{P_4H} \cdot \vec{n}'_0$ 与 $\overrightarrow{P_4H} \cdot \vec{n}'_s$ 同号时, $\alpha' = \theta'$;
- 4) 当 $\overrightarrow{P_4H} \cdot \vec{n}'_0$ 与 $\overrightarrow{P_4H} \cdot \vec{n}'_s$ 异号时, $\alpha' = \pi - \theta'$.

在式(6)中, 差异性测度 $d_{0s}(l, t)$ 由两部分组成. 一部分是 $\cos \langle \vec{n}_0, \vec{n}_s \rangle = \cos \theta$, 为两平面法向量夹角 θ 的余弦值. 由于两平面的二面角与法向量夹角存在相等或互补两种情形(如推论1所述), 为了得到 α 的余弦值, 需要对二面角 α 与法向量夹角 θ 的关系进行判定. 另一部分通过引入符号函数 $\text{sgn}_{0s}(l, t)$, 给出判定二面角 α 与向量夹角 θ 相等或互补的条件:

当 $(\vec{P}_1 \vec{H} \cdot \vec{n}_0)(\vec{P}_1 \vec{H} \cdot \vec{n}_s) \geq 0$ 时, α 与 θ 相等, 有 $\cos \alpha = \cos \theta$;

当 $(\vec{P}_1 \vec{H} \cdot \vec{n}_0)(\vec{P}_1 \vec{H} \cdot \vec{n}_s) < 0$ 时, α 与 θ 互补, 有 $\cos \alpha = -\cos \theta$.

定义4 设 $d_{0s}(l, t)$ 和 $d'_{0s}(l, t)$ 为定义3所述面角差异性测度, 称 $\varepsilon_{0s}(l, t)$ 为 $X(l, t, x_0)$ 和 $X(l, t, x_s)$ 在区域 E_{lt} 内的灰色面角关联系数, 其中

$$\varepsilon_{0s}(l, t) = \frac{1}{4}[1 + d_{0s}(l, t)] + \frac{1}{4}[1 + d'_{0s}(l, t)]. \quad (10)$$

由于 $\alpha, \alpha' \in [0, \pi)$, 易知 $\varepsilon_{0s}(l, t) \in (0, 1]$, 且当 α 和 α' 越大时, 关联系数 $\varepsilon_{0s}(l, t)$ 越小.

定义5 设 $X(l, t, x_0)$ 和 $X(l, t, x_s)$ 分别为特征指标 I_0 和因素指标 I_s 的三维数据矩阵, $\varepsilon_{0s}(l, t)$ 为定义4所述灰色面角关联系数, 记

$$\gamma_{0s} = \frac{1}{(m-1)(n-1)} \sum_{l=1}^{m-1} \sum_{t=1}^{n-1} \varepsilon_{0s}(l, t), \quad (11)$$

称 γ_{0s} 为 $X(l, t, x_0)$ 和 $X(l, t, x_s)$ 的灰色面角关联度.

2.2 模型性质

性质1 基于三维数据矩阵的灰色面角关联度满足: 1) 规范性, $0 < \gamma_{0s} \leq 1$, 当 $X(l, t, x_0) = X(l, t, x_s)$ 时, 有 $\gamma_{0s} = 1$; 2) 对称性, $\gamma_{0s} = \gamma_{s0}$.

性质2 设面板数据 $X_i(l, t), i \in \{0, s, f\}$ 及其三维数据矩阵 $X(l, t, x_i)$ 如定义1所示. $X_i(l, t)$ 经平移变换得到面板数据 $Y_i(l, t), Y_i(l, t) = X_i(l, t) + \mu, \mu = \text{const}$, 其三维数据矩阵为 $Y(l, t, y_i)$. 设 $X(l, t, x_s), X(l, t, x_f)$ 与 $X(l, t, x_0)$ 的灰色面角关联度分别为 γ_{0s} 和 γ_{0f} , $Y(l, t, y_s), Y(l, t, y_f)$ 与 $Y(l, t, y_0)$ 的灰色面角关联度分别为 γ_{0^*s} 和 γ_{0^*f} , 则对 $\forall \mu$ 有

$$\gamma_{0s} > \gamma_{0f} \implies \gamma_{0^*s} > \gamma_{0^*f}. \quad (12)$$

2.3 建模步骤

GDAIM模型建模步骤总结如下:

step 1: 确定对象集合、指标集合、时间集合. 选定复杂系统中 N 个观测指标、 m 个观测对象和 n 个观测时刻点, 构建对象集 $\{O_1, O_2, \dots, O_m\}$, 时间集 $\{T_1, T_2, \dots, T_n\}$, 指标集 $\{I_0, I_1, I_2, \dots, I_w\}$. 选取 I_0 为特征指标, $I_s(s = 1, 2, \dots, w)$ 为其影响因素指标.

step 2: 形成三维数据矩阵. 根据定义1构建观

测指标的三维数据矩阵集合 $X = \{X(l, t, x_s) \mid s = 1, 2, \dots, w, l = 1, 2, \dots, m, t = 1, 2, \dots, n\}$.

step 3: 计算差异性测度. 根据定义3计算 $X(l, t, x_0)$ 和 $X(l, t, x_s)$ 在区域 E_{lt} 内的面角差异性测度 $d_{0s}(l, t)$ 和 $d'_{0s}(l, t)$.

step 4: 计算关联系数和关联度. 根据定义4计算 $X(l, t, x_0)$ 和 $X(l, t, x_s)$ 在区域 E_{lt} 内的灰色面角关联系数 $\varepsilon_{0s}(l, t)$. 根据定义5计算 $X(l, t, x_0)$ 和 $X(l, t, x_s)$ 的灰色面角关联度 γ_{0s} .

step 5: 得到指标面板数据的关联度序列. 重复step 3和step 4, 得到各影响因素指标 I_s 与特征指标 I_0 的灰色面角关联度集合 $\Gamma = \{\gamma_{0s} \mid s = 1, 2, \dots, w\}$.

step 6: 确定关联序. 按照关联度的大小确定各因素指标的关联序, 进而确定特征指标的各影响因素的主次关系. 若 $\gamma_{01} > \gamma_{02}$, 则称与指标 I_2 相比, 指标 I_1 与特征指标 I_0 具有较强相关性, 记为 $I_1 \succ I_2$.

3 算例分析

如引言所述, EGAIM^[16]和TGIM^[17]从面板数据的整体视角出发, 考察面板数据形状相似性特征; BGIM^[13]和GVMIM^[15]分别从对象维和时间维对面板数据关联性进行测度, 再加权求和. 本文模型将面板数据的时间维和对象维视为统一的整体进行相关性测度. 因此, 本文选取EGAIM、TGIM、BGIM和GVMIM作为对比模型, 以研究本文模型的优势. 设有一组面板数据为 $\{X_i \mid i = 0, 1, \dots, 9\}$, 其在三维空间中的网络曲面图如图2所示, 其中 X_0 为特征指标, $\{X_s \mid s = 1, 2, \dots, 9\}$ 为因素指标. 分别运用本文GDAIM模型与4个对比模型EGAIM、TGIM、BGIM、GVMIM求得特征指标与各因素指标的关联度如表1所示.

由图2可以看出: $\{X_1, X_2, X_6, X_7\}$ 与 X_0 的面板结构最为相似; 面板结构波动性较大的 $\{X_3, X_4, X_5\}$ 与 X_0 的结构相似性较小; 呈现反向趋势特征的 $\{X_8, X_9\}$ 与 X_0 面板结构特征相差较大. 表1中关联度显示, 5个模型均具有较大的 $\{\gamma_{01}, \gamma_{02}, \gamma_{06}, \gamma_{07}\}$ 和较小的 $\{\gamma_{08}, \gamma_{09}\}$, 且得到一致结果 $\{X_1, X_2, X_6, X_7\} \succ \{X_3, X_4, X_5\} \succ \{X_8, X_9\}$. 5个模型均能有效识别与 X_0 网络结构较为相似的因素指标.

对于指标组 $\{X_3, X_4, X_5\}$, 5个模型得到的组内关联序结果各不相同. 从图2可以看出: $\{X_3, X_4, X_5\}$ 在对象和时间两个维度的趋势特征均与 I_0 具有较大差异, 但三者之间在两个维度上具有较高的趋势相似性. 从关联度分辨率来看, 由于BGIM和GVMIM均是从两个维度分别对面板数据的趋势相似性进行

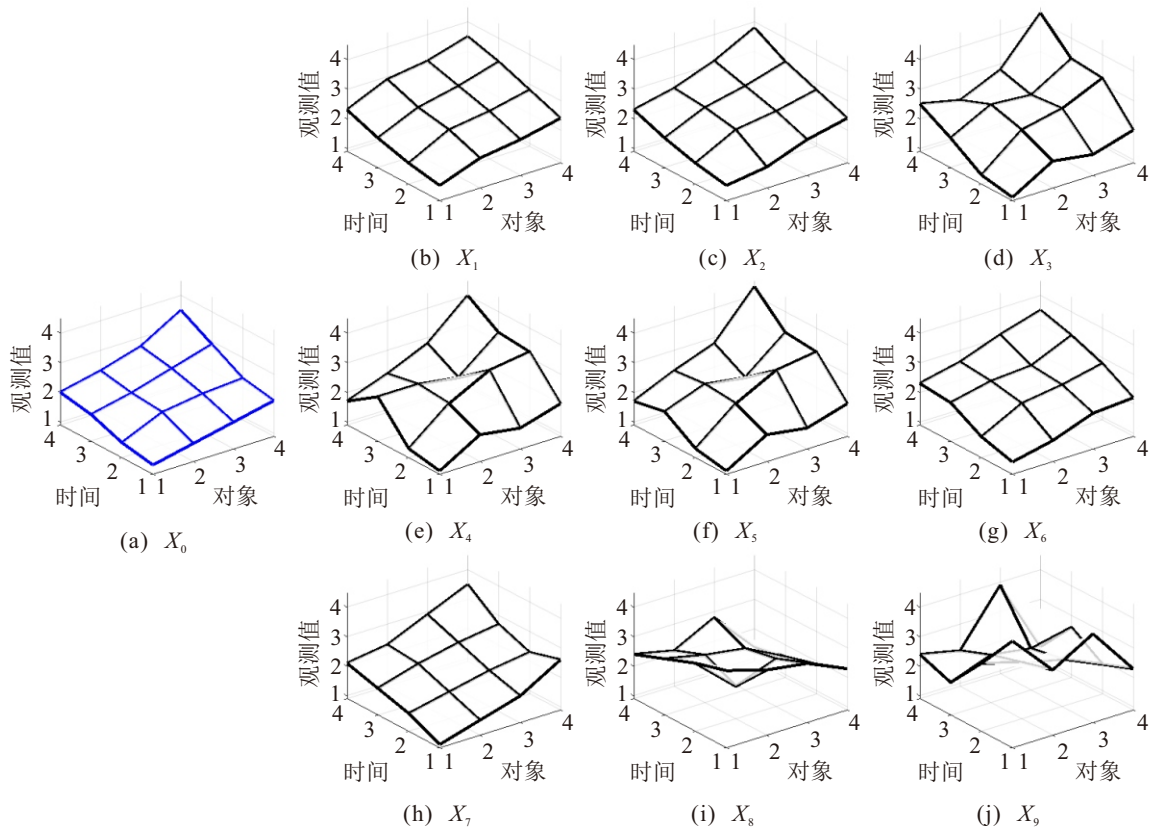


图2 特征指标及其因素指标三维数据网络曲面

表1 对比模型关联度计算结果

关联度	GDAIM	EGAIM	TGIM	BGIM	GVMIM
γ_{01}	0.9915	0.9040	0.8723	0.8939	0.8627
γ_{02}	0.9944	0.9197	0.8430	0.9110	0.8958
γ_{03}	0.6555	0.8682	0.7378	0.8010	0.8120
γ_{04}	0.5474	0.8932	0.6831	0.8021	0.8382
γ_{05}	0.6039	0.8286	0.7631	0.8066	0.8435
γ_{06}	0.9884	0.9909	0.8305	0.8975	0.8588
γ_{07}	0.8783	0.9830	0.8164	0.9092	0.8658
γ_{08}	0.5224	0.5241	0.6181	0.6236	0.5750
γ_{09}	0.5222	0.5340	0.5176	0.5584	0.6262

测度, 没有考虑面板数据的整体特征, 与另外 3 个模型相比, γ_{03} 、 γ_{04} 、 γ_{05} 的区分度较低, 无法有效识别 3 个因素指标与 I_0 在面板数据整体形状特征上的差异性. 从关联度数值大小来看, BGIM 和 GVMIM 的 $\{\gamma_{03}, \gamma_{04}, \gamma_{05}\}$ 数值相对较大, 接近于指标组 $\{X_1, X_2, X_6, X_7\}$ 的关联度, 不符合经验判定结果, 可见从整体出发分析面板数据的形状相似性特征是必要的. 从面板数据整体特征来看, $\{X_4, X_5\}$ 与 X_0 具有明显的形状差异, X_3 与 X_0 的结构特征更具有相似性, 因此判定 $\{X_3\} \succ \{X_4, X_5\}$. 本文模型是 5 个对比模型中唯一得到此结论的模型. 总体来看, 本文模型得到的关联序结果与经验判定结果更为符合, 可有效识别面板数据在三维空间中的结构相似性特征, 且在因素指标间面板结构特征较为相似的情况下,

本文模型对因素指标的排序结果更具可靠性.

4 区域 PM_{2.5} 污染健康风险相关因素识别

4.1 研究区域及其 PM_{2.5} 污染空间效应分析

YRD 地区是位于长江中下游平原的一个集聚区, 包括上海市及江苏省、浙江省、安徽省全域, 共 41 座城市. YRD 地区作为中国特大城市群之一, 具有经济发展活跃、开放程度高、创新能力强等特征. 伴随着经济的快速增长, YRD 地区面临着大气污染问题^[21]. 同时 YRD 地区也是中国人口最稠密的地区之一, 人口密度为 656 人/km², 由大气污染引发的人群健康风险问题备受各界关注^[22]. 利用 Moran 指数^[23]对 YRD 地区 PM_{2.5} 污染水平进行空间相关性分析, 可检验各城市 PM_{2.5} 污染是否具有空间依赖性.

图 3 展示了 YRD 地区 41 市月度 Moran 指数 (选用地理空间权重矩阵). 结果显示, 2016 ~ 2022 年各月份 Moran 指数均大于 0 ($p > 0.001$), 表明 YRD 地区城市间具有显著的正向空间依赖性. 特别地, 每年 1 ~ 2 月、11 ~ 12 月 Moran 指数值较高, 城市间具有较大的空间相关性. An 等^[24]指出, YRD 地区秋冬季节较高的 PM_{2.5} 污染浓度是由部分城市的重污染引起的. 进一步利用各城市 PM_{2.5} 污染浓度最大值的均值绘制 Moran 散点图以分析局域集聚特征, 如图 4 所示. 可以看到呈现“高-高”集聚类型的城市

(第1象限)均分布于安徽和江苏两省,包括安徽省的淮北、宿州、亳州、蚌埠、阜阳、淮南、六安、合肥、滁州、铜陵、马鞍山、池州、安庆、芜湖和宣城市,江苏省的徐州、宿迁、连云港、淮安、扬州和镇江市,表明两省大部分城市具有较高的PM_{2.5}污染浓度,且其周边邻近城市的污染也较为严重.而位于浙江省的城市(湖州、嘉兴、杭州、绍兴、宁波、金华、衢州、台州、丽水、温州、舟山)均呈现“低-低”集聚类型(第3象限),说明浙江省各城市PM_{2.5}污染较轻,且周边邻近城市污染较小.

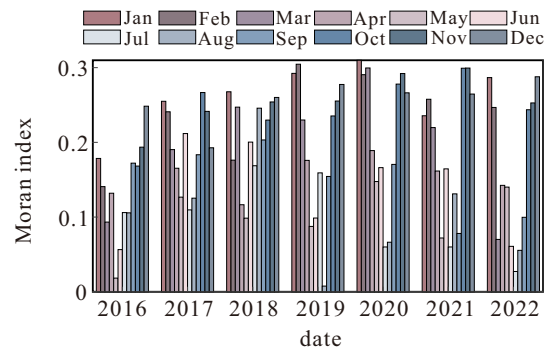


图3 YRD地区41市PM_{2.5}污染月度Moran指数

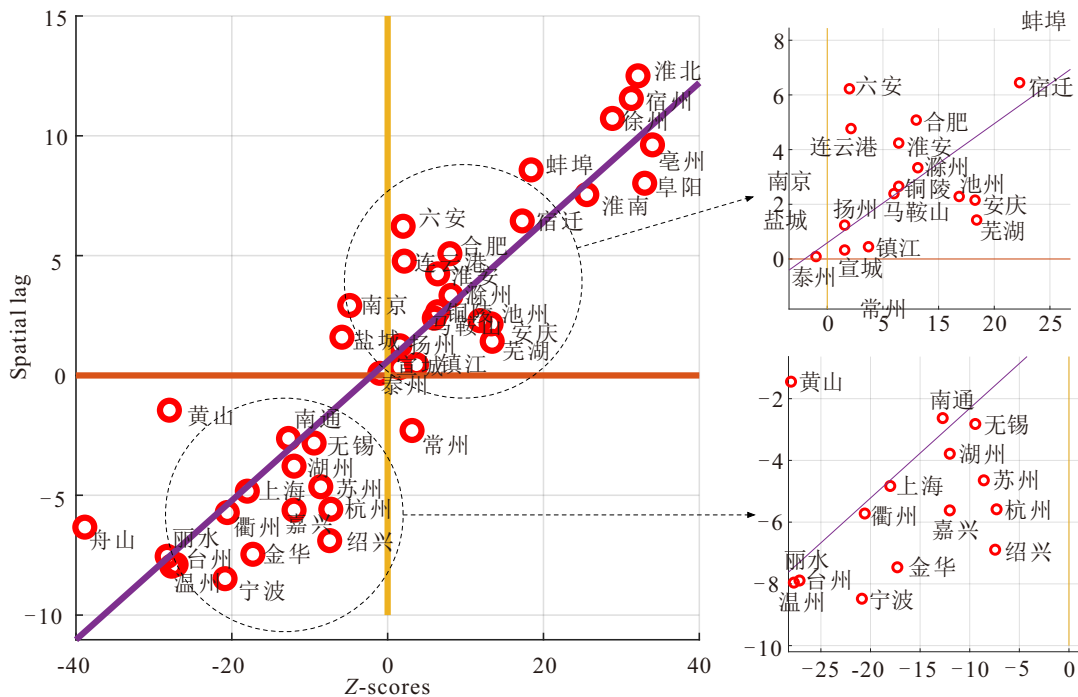


图4 YRD地区41市PM_{2.5}污染峰值Moran散点图

基于上述分析,在YRD地区PM_{2.5}污染的健康风险相关研究中,考虑到江苏省和安徽省秋冬季PM_{2.5}污染较为严重,是危害YRD地区人群健康的潜在风险,且两省PM_{2.5}污染呈现出显著的空间相关性和“高-高”集聚类型,因此可以利用本文构建的GDAIM模型,该模型可整合对象纬度相关性与时空趋势相似性,以识别与YRD地区PM_{2.5}污染健康风险相关的经济和社会因素.

4.2 人群健康风险评估方法

根据流行病学中污染物浓度与健康效应之间的暴露-反应关系以及泊松回归相对危险度模型^[25],可估算归因于PM_{2.5}污染的人群健康风险

$$E = E_0 \exp[\beta(c - c_0)],$$

$$\Delta E = P(E - E_0). \quad (13)$$

其中: E 为PM_{2.5}实际浓度下的健康风险, E_0 为PM_{2.5}参考基准浓度下的健康风险, β 为暴露-反应关

系数, c 为PM_{2.5}实际浓度($\mu\text{g} \cdot \text{m}^{-3}$), c_0 为PM_{2.5}参考基准浓度($\mu\text{g} \cdot \text{m}^{-3}$), P 为暴露人口数量, ΔE 为归因于PM_{2.5}浓度变化带来的健康风险效应变化量.

本文选取门诊内科就诊人次为健康风险终点,根据优先选择本地化系数的原则,选取文献[25]中暴露-反应关系系数和基准发生率分别为0.000 49和0.411 05%;选取常住人口为暴露人群;选取PM_{2.5}国家二级浓度限值($35 \mu\text{g} \cdot \text{m}^{-3}$)为参考基准浓度 c_0 ;选取每年中PM_{2.5}月均浓度最大值为实际浓度.依据式(13)可计算出江苏省和安徽省2016~2021年PM_{2.5}污染的人群健康风险.

4.3 基于GDAIM模型的相关因素识别

考虑经济活动、人口分布以及污染物排放3个方面对健康风险的影响,本文选取地区生产总值(GDP)、地区人均可支配收入(PCDI)、发电量(PG)、第二产业增加值占比(PASI)、城镇人口(UP)、乡村

人口 (RP)、人口密度 (PD)、颗粒物排放量 (PM)、二氧化硫排放量 (SO₂) 和氮氧化物排放量 (NO_x) 作为 ΔE 的影响因素指标. 其中: GDP 是衡量地区经济活动总量的重要指标; PCDI 反映了地区居民的生活水平和消费能力; YRD 地区以江苏省、安徽省等为代表的省份通常是电力输入省份, 工业用电量污染排放高; UP、RP 和 PD 体现了暴露人口的分布特点; PM、SO₂ 和 NO_x 是主要的污染排放指标.

根据 GDAIM 建模步骤可以得到 YRD 地区各因素指标与 ΔE 之间的关联度, 如表 2 所示. 关联序为 PM > SO₂ > RP > NO_x > PASI > PD > UP > PG > GDP > PCDI. 建模结果表明, 3 种排放要素 PM、SO₂、NO_x 均与 ΔE 有较大相关性. 已有大量文献研究表明, 大气污染会对人群健康产生影响, 并与过早死亡有关^[26]. 本文选取的特征指标为归因于 PM_{2.5} 污染的人群健康风险, 并且颗粒物是大气 PM_{2.5} 污染形成的主要原因^[27], 因此颗粒物排放与 ΔE 的关联度较大. 同时也表明本文模型在本案例中是有效的. 由于 SO₂ 和 NO_x 是 PM_{2.5} 污染的主要前体物质^[28], SO₂ 和 NO_x 与 ΔE 具有一定相关性. YRD 地区可加强污染物排放末端管控政策的制定, 在重点管控颗粒物排放的同时, 也需要关注 SO₂ 和 NO_x 的排放, 以防范大气污染物排放带来的健康风险.

表2 PM_{2.5} 污染健康风险与其影响因素的关联度

指标	GDAIM		EGAIM		TGIM		BGIM		GVMIM	
	γ_{0s}	序	γ_{0s}	序	γ_{0s}	序	γ_{0s}	序	γ_{0s}	序
GDP	0.418	9	0.500	10	0.858	3	0.309	10	0.662	10
PCDI	0.393	10	0.500	10	0.868	1	0.397	5	0.669	8
PG	0.419	8	0.500	8	0.849	6	0.362	9	0.667	9
PASI	0.559	5	0.504	5	0.759	10	0.468	1	0.695	5
UP	0.435	7	0.500	7	0.853	4	0.375	7	0.673	7
RP	0.625	3	0.531	4	0.867	2	0.468	1	0.750	2
PD	0.443	6	0.502	6	0.829	8	0.371	8	0.673	6
PM	0.845	1	0.645	3	0.849	5	0.451	3	0.734	4
SO ₂	0.654	2	0.818	1	0.848	7	0.396	6	0.735	3
NO _x	0.585	4	0.806	2	0.820	9	0.428	4	0.762	1

在经济要素中, GDP、PCDI 与 ΔE 相关性较小, PASI 与 ΔE 相关性较大. 经济增长可能会增加大气污染的排放量, 对人群健康产生潜在风险. 但由于经济要素本身受到多因素的影响, 例如不同产业部门对 GDP 贡献率的变化会引起经济结构的变化^[28], 经济要素指标本身具有较大的不确定性, 所以与排放要素相比, GDP 对 ΔE 的影响相对较低. 第二产业增加值占比对 PM_{2.5} 污染健康风险具有一定的影响, 工业部门的增加往往会增加污染排放量, 因此 YRD 地区可通过调整经济结构, 调配各行业的污染排放份

额来管控工业排放对暴露人群产生的健康风险.

在人口要素中, 乡村人口 RP 与 ΔE 的关联度 (0.6245) 明显大于城镇人口 UP (0.4349). Liu 等^[29]指出, 由于在烹饪和取暖等活动中使用非清洁能源, 农村地区人口的与 PM_{2.5} 相关心血管疾病风险会更高.

基于上述分析, YRD 地区 PM_{2.5} 污染引发的人群健康风险受到大气污染物排放和经济结构等外部因素的影响, 需要通过政府“有形之手”进行干预. YRD 地区在制定 PM_{2.5} 污染下健康风险防控政策时, 一方面要加强颗粒物等大气污染物排放的管控和治理; 另一方面可通过调整产业结构以降低污染物排放对暴露人群产生的健康风险. 此外, 乡村地区人口所面临的 PM_{2.5} 污染健康风险问题还需得到更多的关注, 可进一步考虑城乡差异性, 针对性地制定乡村地区健康风险防控政策, 降低乡村人口的暴露水平.

利用不同时间尺度实验数据以及滚动时间窗方法, 可对 GDAIM 模型稳定性进行测试, 建模结果如表 3 所示. 在对不同时间尺度数据进行建模时, 除了少部分指标与其排序临近指标前后位置调换之外 (例如 NO_x 与 PASI, GDP 与 PG), 指标排序结果几乎没有改变. 在 3 个滚动时间阶段的建模结果中, PM 始终是排位第 1 的指标, 并且除了 NO_x 排序位置发生了较大变化之外, 少部分指标仅是与临近指标调换位置, 从总体上看, 3 个时间阶段的指标排序结果是相似的. 综上分析, GDAIM 模型是稳定的, 可用于区域大气污染健康风险影响因素的识别研究.

表3 GDAIM 模型稳定性测试

时间/年	指标排序结果
2016~2021	PM > SO ₂ > RP > NO _x > PASI > PD > UP > PG > GDP > PCDI
2017~2021	PM > SO ₂ > RP > PASI > NO _x > PD > UP > GDP > PG > PCDI
2018~2021	PM > SO ₂ > RP > NO _x > PASI > PD > UP > PG > GDP > PCDI
2019~2021	PM > SO ₂ > RP > NO _x > PASI > PD > UP > GDP > PG > PCDI
2016~2019	PM > NO _x > RP > PASI > SO ₂ > PD > UP > PG > GDP > PCDI
2017~2020	PM > RP > SO ₂ > PASI > NO _x > PD > PG > UP > GDP > PCDI
2018~2021	PM > SO ₂ > RP > NO _x > PASI > PD > UP > PG > GDP > PCDI

4.4 与已有模型对比分析

本节仍选取 EGAIM、TGIM、BGIM 和 GVMIM 作为对比模型, 分析不同功能面板数据关联模型在实例中的应用效果, 各模型关联度计算结果及指标排序结果分别如表 2 和表 4 所示.

从关联度数值来看, TGIM、BGIM、GVMIM 模型指标关联度区分度较差, 关联度数值区间分别为

表4 5个面板数据灰色关联模型的指标排序结果

模型	指标排序结果
GDAIM	PM _x >SO ₂ >RP>NO _x >PASI>PD>UP>PG>GDP>PCDI
EGAIM	SO ₂ >NO _x >PM>RP>PASI>PD>UP>PG>GDP = PCDI
TGIM	PCDI>RP>GDP>UP>PM>PG>SO ₂ >PD>NO _x >PASI
BGIM	PASI=RP>PM>NO _x >PCDI>SO ₂ >UP>PD>PG>GDP
GVMIM	NO _x >RP>SO ₂ >PM>PASI>PD>UP>PCDI>PG>GDP

[0.759 3, 0.867 7]、 [0.309 2, 0.467 9]、 [0.661 6, 0.761 8]; 在 TGIM 中, 所有影响因素指标均与 ΔE 具有较大的相关性. 相反, 在 BGIM 模型中, 所有影响因素指标均与 ΔE 有较小的相关性. 由此可见, 仅考虑面板数据的形状特征 (TGIM) 或仅考虑面板数据维度特征 (BGIM、GVMIM) 不能满足邻近区域大气污染健康风险影响因素识别的需求, TGIM、BGIM、GVMIM 三种模型对 YRD 地区 PM_{2.5} 污染健康风险的关键影响因素识别效果不佳. 从排序结果来看, GDAIM 模型和 EGAIM 模型对影响因素指标的排序较为一致, 3 个排放因素指标 PM、SO₂、NO_x 以及人口指标 RP 与 ΔE 相关性较高. 但对 3 个排放因素指标的排序不同, 在 GDAIM 建模结果中 PM_{2.5} 是影响 ΔE 的主要因素, 而在 EGAIM 模型中, SO₂ 和 NO_x 排放量对 ΔE 的影响要高于 PM_{2.5} 排放量, 结合 4.2 节分析, 由于特征指标为归因于 PM_{2.5} 污染的人群健康风险, 因此本文模型建模结果更为可靠.

5 结论

本文探讨了面板数据在空间坐标系中的几何表示方法, 构建了三维数据矩阵的空间网络结构, 并针对其几何特征, 利用平面法向量和夹角相关理论, 构建了面角差异性测度, 进而构建了基于三维数据的灰色面角关联模型 GDAIM. 该模型将面板数据的时间维和截面维视为一个整体, 从整体性视角研究两个面板数据的发展趋势特征. 算例分析验证了本文模型的有效性及其优势, 并且本文模型对于面板数据波动性特征的识别效果更佳. 通过对 YRD 地区 PM_{2.5} 污染进行 Moran 相关性分析, 发现 YRD 地区江苏和安徽两省的 PM_{2.5} 污染具有显著的空间相关性. 为同时兼顾面板数据截面维度对象间的空间相关性以及时间维的发展趋势特征, 将 GDAIM 模型应用于识别 YRD 地区 PM_{2.5} 污染下健康风险的主要影响因素. 结果表明颗粒物、二氧化硫和氮氧化物的排放是 YRD 地区 PM_{2.5} 污染健康风险的主要相关因素. 在经济因素中, 与 GDP 相比, 第二产业增加值占比对 YRD 地区 PM_{2.5} 污染健康风险的影响较

大. 在人口要素中, 与城镇人口相比, 乡村人口与 YRD 地区 PM_{2.5} 污染健康风险的相关性较大. 在未来的研究中, 当研究对象数量 ≥ 3 时, 其排列方式对建模结果的影响尚需进一步研究. 考虑到城市间具有地理方位等固有信息 (例如经纬度坐标), 基于研究对象在系统中蕴含的固有信息, 面板数据中截面维度研究对象的排序规则, 构建规则约束下的三维数据矩阵, 以期完善面板数据建模理论.

参考文献 (References)

- [1] Rajagopalan S, Al-Kindi S G, Brook R D. Air pollution and cardiovascular disease JACC state-of-the-art review[J]. *Journal of the American College of Cardiology*, 2018, 72(17): 2054-2070.
- [2] Huang R J, Zhang Y L, Bozzetti C, et al. High secondary aerosol contribution to particulate pollution during haze events in China[J]. *Nature*, 2014, 514(7521): 218-222.
- [3] Wang S, Kaur M, Li T F, et al. Effect of different pollution parameters and chemical components of PM_{2.5} on health of residents of Xinxiang City, China[J]. *International Journal of Environmental Research and Public Health*, 2021, 18(13): 6821.
- [4] 黄德生, 张世秋. 京津冀地区控制 PM_{2.5} 污染的健康效益评估[J]. *中国环境科学*, 2013, 33(1): 166-174. (Huang D S, Zhang S Q. Health benefit evaluation for PM_{2.5} pollution control in Beijing-Tianjin-Hebei region of China[J]. *China Environmental Science*, 2013, 33(1): 166-174.)
- [5] 韦保磊, 谢乃明. 广义灰色关联分析模型的统一表述及性质[J]. *系统工程理论与实践*, 2019, 39(1): 226-235. (Wei B L, Xie N M. Unified representation and properties of generalized grey relational analysis models[J]. *Systems Engineering - Theory & Practice*, 2019, 39(1): 226-235.)
- [6] Wang J J, Cai Y, Feng Y, et al. Novel grey dynamic trend relational analysis models with different types of accumulation delay effects for time-delay systems[J]. *Expert Systems with Applications*, 2024, 238: 121661.
- [7] Feng Y, Dang Y G, Wang J J, et al. A novel grey projection incidence model for assessing the relationships between cardiovascular diseases and air pollutants[J]. *ISA Transactions*, 2023, 135: 398-409.
- [8] Huang M D, Xia T B, Chen Z, et al. A DEA integrated grey factor analysis approach for efficiency evaluation and ranking in uncertain systems[J]. *Computers & Industrial Engineering*, 2021, 162: 107681.
- [9] Liu Y, Zhang R S. A three-way grey incidence clustering approach with changing decision objects[J]. *Computers & Industrial Engineering*, 2019, 137: 106087.
- [10] Javed S A, Mahmoudi A, Liu S F. Grey absolute decision analysis(GADA) method for multiple criteria group decision-making under uncertainty[J]. *International Journal of Fuzzy Systems*, 2020, 22(4): 1073-1090.
- [11] 刘震, 党耀国, 钱吴永, 等. 基于面板数据的灰色网格

- 关联度模型[J]. 系统工程理论与实践, 2014, 34(4): 991-996.
- (Liu Z, Dang Y G, Qian W Y, et al. Grey grid incidence model based on panel data[J]. *Systems Engineering-Theory & Practice*, 2014, 34(4): 991-996.)
- [12] 崔立志, 刘思峰. 面板数据的灰色矩阵相似关联模型及其应用[J]. 中国管理科学, 2015, 23(11): 171-176.
- (Cui L Z, Liu S F. Grey matrix similar incidence model for panel data and its application[J]. *Chinese Journal of Management Science*, 2015, 23(11): 171-176.)
- [13] 罗党, 张曼曼. 基于面板数据的灰色 B 型关联模型及其应用[J]. 控制与决策, 2020, 35(6): 1476-1482.
- (Luo D, Zhang M M. Grey incidence model of B-mode based on panel data and its application[J]. *Control and Decision*, 2020, 35(6): 1476-1482.)
- [14] Sun J, Dang Y G, Zhu X Y, et al. A grey spatiotemporal incidence model with application to factors causing air pollution[J]. *Science of the Total Environment*, 2021, 759: 143576.
- [15] 党耀国, 尚中举, 王俊杰, 等. 基于面板数据的灰色指标关联模型构建及其应用[J]. 控制与决策, 2019, 34(5): 1077-1084.
- (Dang Y G, Shang Z J, Wang J J, et al. Grey incidence model for relationship between indicators of panel data and its application[J]. *Control and Decision*, 2019, 34(5): 1077-1084.)
- [16] 张可, 刘思峰. 灰色关联聚类在面板数据中的扩展及应用[J]. 系统工程理论与实践, 2010, 30(7): 1253-1259.
- (Zhang K, Liu S F. Extended clusters of grey incidences for panel data and its application[J]. *Systems Engineering - Theory & Practice*, 2010, 30(7): 1253-1259.)
- [17] 吴鸿华, 刘思峰, 方志耕. 基于面板数据的灰色四面体网格关联分析模型及应用[J]. 控制与决策, 2022, 37(11): 3033-3041.
- (Wu H H, Liu S F, Fang Z G. Grey tetrahedral grid incidence analysis model based on panel data and its application[J]. *Control and Decision*, 2022, 37(11): 3033-3041.)
- [18] Brunsdon C, Fotheringham A S, Charlton M E. Geographically weighted regression: A method for exploring spatial nonstationarity[J]. *Geographical Analysis*, 1996, 28(4): 281-298.
- [19] Miao L Z, Liu C L, Yang X, et al. Spatiotemporal heterogeneity analysis of air quality in the Yangtze River Delta, China[J]. *Sustainable Cities and Society*, 2022, 78: 103603.
- [20] Feng T, Du H B, Lin Z G, et al. Spatial spillover effects of environmental regulations on air pollution: Evidence from urban agglomerations in China[J]. *Journal of Environmental Management*, 2020, 272: 110998.
- [21] Feng R D, Wang K Y, Wang F Y. Quantifying influences of administrative division adjustment on PM_{2.5} pollution in China's mega-urban agglomerations[J]. *Journal of Environmental Management*, 2022, 302: 113993.
- [22] Li N, Zhang H R, Zhu S H, et al. Secondary PM_{2.5} dominates aerosol pollution in the Yangtze River Delta region: Environmental and health effects of the Clean air Plan[J]. *Environment International*, 2023, 171: 107725.
- [23] Moran P A P. The interpretation of statistical maps[J]. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 1948, 10(2): 243-251.
- [24] An Y M, Dang Y G, Wang J J, et al. Identification of heavily polluted areas based on a novel grey integrated incidence model: A case study of the Yangtze River Delta, China[J]. *Sustainable Cities and Society*, 2023, 92: 104466.
- [25] 李惠娟, 周德群, 魏永杰. 我国城市 PM_{2.5} 污染的健康风险及经济损失评价[J]. 环境科学, 2018, 39(8): 3467-3475.
- (Li H J, Zhou D Q, Wei Y J. An assessment of PM_{2.5}-related health risks and associated economic losses in Chinese Cities[J]. *Environmental Science*, 2018, 39(8): 3467-3475.)
- [26] Burnett R, Chen H, Szyszkowicz M, et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter[J]. *Proceedings of the National Academy of Science*, 2018, 115(38): 9592-9597.
- [27] Wu D, Zheng H T, Li Q, et al. Toxic potency-adjusted control of air pollution for solid fuel combustion[J]. *Nature Energy*, 2022, 7: 194-202.
- [28] Geng G N, Zheng Y X, Zhang Q, et al. Drivers of PM_{2.5} air pollution deaths in China 2002-2017[J]. *Nature Geoscience*, 2021, 14(9): 645-650.
- [29] Liu C, Chan K H, Lv J, et al. Long-term exposure to ambient fine particulate matter and incidence of major cardiovascular diseases: A prospective study of 0.5 million adults in China[J]. *Environmental Science & Technology*, 2022, 56(18): 13200-13211.

作者简介

安艺萌 (1995-), 女, 博士生, 主要研究方向为灰色系统理论、时间序列预测建模, E-mail: anyimeng699@nuaa.edu.cn;

党耀国 (1964-), 男, 教授, 博士生导师, 主要研究方向为灰色系统理论、数量经济, E-mail: iamdangyg@163.com;

杨德岭 (1963-), 男, 教授, 博士, 主要研究方向为高等教育管理、风险评价, E-mail: zzb8913@163.com;

王俊杰 (1988-), 男, 副教授, 博士生导师, 主要研究方向为灰色系统理论、数量经济, E-mail: wangjj@nuaa.edu.cn;

冯宇 (1993-), 男, 博士生, 主要研究方向为灰色决策与不确定性分析, E-mail: fengyu3091@nuaa.edu.cn.