引文格式：田诗佳,包丹文,尹俐平,等. 新冠疫情与航空运输规模的空间非平稳关系研究[J]. 地理空间信息, 2024, 22(9): 19-21. doi:10. 3969/j. issn. 1672-4623.2024.09.004

# 新冠疫情与航空运输规模的空间非平稳关系研究

田诗佳[1]，包丹文[1*]，尹俐平[1]，程　昊[1]，姚馨宇[1]
（1. 南京航空航天大学 民航学院，江苏 南京 211106）

**摘　要：**以我国通航城市为研究区，选取2020年初的疫情数据和航空运输网络数据作为实验数据，并从人口流动、区域人口特征和医疗水平3个维度选取了多个解释变量以提高模型拟合精度，构建了基于属性空间距离的情景化多尺度地理加权回归模型（CMGWR），分析了各因素对疫情的空间分异影响。结果表明，相较于传统GWR和MGWR模型，CMGWR模型对于疫情影响因素作用程度的空间分异特征具有更强的解释能力；疫情与航空运输规模的关联具有较强的空间异质性，且关联度较高的地区均为区域航空枢纽，其他因素对疫情影响力较大的范围也以区域核心城市为主。

**关键词：**航空运输；新冠疫情；影响因素；空间异质性；CMGWR模型
**中图分类号：**P237　　**文献标志码：**B　　**文章编号：**1672-4623（2024）09-0019-03

## Research on Spatial Heterogeneity in the Relation Between Air Transport and COVID-19

TIAN Shijia[1], BAO Danwen[1], YIN Liping[1], CHENG Hao[1], YAO Xinyu[1]
(1. School of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

**Abstract:** Taking the cities with airports as the research area, we used the epidemic data and air transport network data in early 2020, and five explanatory variables from population flow, regional demographic characteristics and medical level improving the model fitting accuracy to build a contextualized multi-scale geographically weighted regression model (CMGWR) based on the distance in attribute space. And then, we analyzed the spatial heterogenous impact of factors on the epidemic. The results show that the CMGWR model can better explain the spatial heterogeneity in the influence factors of the epidemic than the traditional GWR and MGWR models. The association between the epidemic and the scale of air transport has strong spatial heterogeneity, and the areas with higher associations are all regional air hubs. Other factors that have a strong influence on the epidemic are also core cities in the regions.

**Key words:** air transport, COVID-19, influence factor, spatial heterogeneity, CMGWR model

新冠病毒的传播与人口流动密切相关[1]。航空运输作为人口在区域间流动的重要方式，对疫情发展有着不可忽视的影响。国内外学者多采用多元线性回归[2]，负二项回归[3]、泊松回归[4]、空间杜宾模型[5]等方法研究不同国家和地区航空运输规模与疫情发展的相关性，但忽略了区域之间的差异以及变量回归过程中可能存在的空间非平稳性。地理加权回归（GWR）模型和多尺度地理加权回归模型（MGWR）能弥补这种缺陷，其对于空间权重矩阵的计算都是基于地理空间中位置越邻近的事物相似度越高的假设[6]。然而，我国城市体系网络具有明显的结构异质性特征[7]，在研究城市间某些特定的经济社会关系时，权重随地理距离增加而衰减的假设不再适合描述城市间的相关性。基于此，Harris R[8]等提出了属性空间的概念，并建立了基于属性空间距离的情景化地理加权回归模型（CGWR），并阐明了其在空间计量分析中的应用前景。该模型在建立空间权重矩阵时对地理学第一定律中传统地理空间邻近关系进行了情景化拓展，与GWR模型相比，其在特定情景下可有效提高模型的拟合优度。因此，本文结合CGWR模型与MGWR模型在非平稳关系建模中的优势，构建了情景化多尺度地理加权回归模型（CMGWR），并在本文设置的研究情景下与GWR、MGWR模型的拟合效果进行了对比，最终采用拟合效果最好的模型对航空运输规模与疫情的空间非平稳关联进行了研究。

## 1　研究区域与数据来源

### 1.1　研究区域和时间

本文以通航地级市为基本研究单元，由于武汉市

收稿日期：2022-08-20。
项目来源：国家自然科学基金民航联合重点资助项目（U2033203）；南京航空航天大学研究生科研与实践创新计划资助项目（xcxjh20210705）。
第一作者简介：田诗佳（1997—），硕士研究生，研究方向为交通运输规划与管理，E-mail：lexi_tian@126.com。
通信作者：包丹文（1982—），副教授，硕士生导师，研究方向为交通运输规划与管理，E-mail：baodanwen@nuaa.edu.cn。

内病例增长主要受市内人口流动影响，湖北省内其他城市的外来病例主要通过来自武汉市的高铁和公路传播[9]，因此将湖北省内通航城市排除在外。同时，去除包含缺失值、离群和噪声值的数据，本文最终选取我国128个通航市级行政区作为基本研究单元。研究时间为2020年1月23日—2月20日，该时段内各城市的确诊病例数不包含境外输入病例。

### 1.2 变量选取与数据来源

本文选取的因变量为各城市在研究时段内的确诊病例数，核心解释变量为航空运输规模。由于疫情发展还与人口流动、社会特征以及医疗水平有关，因此需进一步引入多维控制变量以提高模型的拟合精度，避免因解释变量过少而造成遗漏变量偏差。结合国内外相关研究成果，本文将武汉市迁入人口比例、城市内出行强度、人口密度、老年人口比例和医院床位数作为模型的控制变量。解释变量详见表1。本文采用方差膨胀因子（VIF）和容差对解释变量进行共线性诊断，结果见表2。所有解释变量容差均大于0.2，且VIF均小于5，说明解释变量不具备显著多重共线性，符合回归拟合的条件。未包含在模型中的由空间关系唯一决定的因素对因变量产生的影响可在一定程度上由模型中的常数项获取。

表1 解释变量描述

| 类型 | 指标名称 | 指标描述 |
|---|---|---|
| 核心解释变量 | 航空运输规模/架次 | 某机场到达航班数量 |
| 人口流动 | 武汉市迁入人口比例/% | 武汉市封城前春运期间（1月10—22日）由武汉市迁入的人口占迁出总人口的比例 |
| | 城市内出行强度 | 城市内平均出行强度 |
| 人口特征 | 人口密度/(人/km²) | 总人口数与城市建设用地面积的比值 |
| | 老年人口比例/% | 武汉市老年人口（65岁以上）占总人口的比例 |
| 医疗水平 | 医院床位数量/个 | 全市医院床位数量 |

表2 共线性诊断

| 解释变量 | 容差 | VIF |
|---|---|---|
| 航空运输规模 | 0.485 | 2.063 |
| 武汉市迁入人口比例 | 0.754 | 1.327 |
| 城市内出行强度 | 0.696 | 1.436 |
| 人口密度 | 0.731 | 1.369 |
| 老年人口比例 | 0.888 | 1.126 |
| 医院床位数量 | 0.557 | 1.796 |

本文采用的确诊病例数量来源于各省市卫健委官方公布的数据；航空运输相关数据从OAG数据库获取；人口流动数据通过Python软件从百度迁徙地图中爬取；人口密度和医院床位数量从2020年中国城市统计年鉴获取；老年人口比例通过第七次人口普查数据计算获取。

## 2 局域加权回归模型建模方法

GWR模型和MGWR模型是较常用的回归方法，建模过程中均假设地理位置较近的事物共享同样或相似的背景，这与现实生活中两个样本点地理距离较近、但社会某些方面相似度较低的现象并不一致。因此，本文采用特定属性空间邻近关系对传统地理空间邻近关系进行拓展，构建了CMGWR模型，计算公式为：

$$y_i=\beta_{bw_0}(u_i,v_i)+\sum_{k=1}^{K}\beta_{bw_k}(u_i,v_i)x_{ik}+\varepsilon_i \quad i=1,2,\cdots,n \quad (1)$$

式中，$\beta_{bw_0}(u_i,v_i)$ 为特定带宽下模型在城市 $i$ 处的截距；$bw_k$ 为校准的第 $k$ 个解释变量使用的最佳带宽，带宽选择遵循校正Akaike信息准则（AICc）；$\beta_{bw_k}(u_i,v_i)$ 为第 $k$ 个解释变量最佳带宽下的 $x_{ik}$ 系数；$\varepsilon_i$ 为相互独立的随机误差项，服从 $\varepsilon_i\sim N(0,\sigma^2 I)$。

与MGWR模型相同，CMGWR模型采用后退拟合算法进行参数估计，选取残差平方和（RSS）的比例变化作为收敛准则[10]。其不同点在于空间权重矩阵中对于距离 $d_{ij}$ 的计算，MGWR模型采用地理空间距离，而CMGWR模型采用特定属性空间距离。传统城市属性的研究依赖于城市的固有特征，如地理位置、人口和经济规模等，但该类属性并不能完全反映城市间的关联特征。根据Castells M[11]的“流动空间”理论，城市的地位特征更多依赖于其在网络中与其他城市的联系情况。本文旨在探究疫情与航空运输规模关联的空间非平稳关系，因此从航空联系的视角对通航城市属性空间特征进行探索，并构建空间权函数。采用映射法构建航空运输网络的无向拓扑图 $G=(N,E)$ 来研究通航城市在航空运输网络空间中的关系，其中 $N=\{1,2,\cdots,n\}$ 为网络中所有节点（通航城市）的集合；$E=\{(i,j)|i,j\in N,i\neq j\}$ 为网络中边（航线）的集合，采用样本点 $i$ 和 $j$ 之间的航班频率 $f_{ij}$ 对边 $(i,j)$ 进行加权处理。在航空网络中，节点度大小可反映该城市的通达性和规模[12]，计算公式为：

$$u_i=\sum_j a_{ij} \quad i,j\in N,i\neq j \quad (2)$$

式中，$a_{ij}$ 为节点 $i$ 与 $j$ 之间连接边的数量。

在节点度的基础上，考虑节点之间的联系强度，可得到节点的局部中心度。局部中心度可反映疾病在该节点的对外传播能力[12]，计算公式为：

$$v_i = \frac{u_i}{n-1}\sqrt{\frac{\sum_j f_{ij}}{M_n}} \tag{3}$$

式中，$n$ 为网络中城市的总数量；$M_n$ 为网络中所有城市对外航空联系的平均值。

因此，通航城市 $i$ 和 $j$ 在基于航空联系时间下的属性空间距离的计算公式为：

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \tag{4}$$

## 3　模型拟合效果对比

本文分别采用OLS、GWR、MGWR和CMGWR模型对疫情及其影响因素进行建模。上述模型均在Windows 10系统上进行拟合，OLS模型采用SPSS 19.0进行解算，GWR、MGWR和CMGWR模型采用Oshan TM[13]等开发的MGWR2.2软件进行解算。4种模型拟合效果对比见表3，可以看出，全局模型OLS模型的拟合优度最低，局部模型MGWR和CMGWR模型的拟合优度明显优于GWR，表明变量作用尺度效应在模型拟合中是需要重点考虑的一项内容。当采用变带宽对参数进行回归时，基于航空联系属性空间的CMGWR模型比基于地理空间的MGWR模型的修正 $R^2$ 约提高了10%，AICc值约降低了24%，RSS值约降低了25%，CMGWR模型的拟合效果最好。

表3　4种模型拟合结果对比

| 评价指标 | OLS模型 | GWR模型 | MGWR模型 | CMGWR模型 |
|---|---|---|---|---|
| 修正 $R^2$ | 0.541 | 0.657 | 0.720 | 0.789 |
| AICc | 372.058 | 288.961 | 256.613 | 194.427 |
| AICc减少 | | 83.097 | 32.348 | 62.186 |
| RSS | 56.841 | 41.405 | 35.795 | 26.973 |
| RSS减少 | | 15.436 | 5.61 | 8.822 |

## 4　回归系数空间格局分析

### 4.1　航空运输规模对疫情的空间分异影响

航空运输规模的估计系数分布在0.002~0.435之间，其中影响较大的城市在空间地理上呈分散式分布，且在秦岭淮河以南的东部沿海地区分布更密集。航空运输与疫情关联度较高的通航城市均为省会城市、直辖市或省内经济较发达的核心城市，其主要原因是核心城市的机场往往是区域的航空枢纽，其通航能力和对外联系度远远大于非核心城市的机场。

本文采用自然间断点分级法将航空运输系数分为以0.056、0.123、0.212和0.319为间隔点的5组数据，5组数据的组内方差最小、组间方差最大。系数值在由低到高各区间通航城市样本分别占样本总数的26.56%、20.31%、19.53%、16.41%和17.19%，即超过半数的通航城市疫情与航空运输量的关联程度处于较低水平。其中，排名前4的城市分别为北京（0.436）、上海（0.434）、广州（0.429）和成都（0.426），这4个城市为我国四大城市群的核心城市，疫情通过这些城市的航空网络向全国传播的风险较高，航空运输规模对疫情发展的影响较大。

### 4.2　其他控制变量对疫情的空间分异影响

本文采用自然间断点分级法对各控制变量的回归系数进行计算。在人口流动因素中，武汉市迁入人口比例在全局范围内对疫情爆发规模起正向作用，这与疫情爆发初期各通航城市的首发病例几乎都有武汉市旅居史的现象相符；而城市内出行强度与疫情爆发规模呈负相关关系，城市确诊数量越多，居民对于出行的恐惧程度就越高，出行强度也就越低。在人口特征因素中，通航城市当地的人口密度和老年人口比例对疫情爆发规模的影响在全局范围内呈正相关关系，人口密度越大，与病毒接触的风险就越高，病毒通过人群扩散的风险也就相应增加；老年人口的抵抗力较差，更容易感染病毒。在医疗水平因素中，医院床位数量与疫情爆发规模呈正相关关系。早期各地公布的确诊病例数量除与病毒的本土传播有关外，还与当地的检测水平有关，因此当地医疗水平越高，病毒携带者被检测出的几率就越大。

## 5　结　语

本文以全国128个通航地级市为研究对象，对2020年初各地疫情爆发规模与航空运输规模的空间非平稳关联开展了研究。为弥补传统地理空间邻近关系的局限性，构建了基于属性空间距离的CMGWR模型，并与全局模型OLS模型、基于地理空间的GWR、MGWR模型的建模性能进行了对比。结果表明，CMGWR模型能显著提高模型的拟合优度，航空运输规模对疫情爆发规模的影响存在显著的空间非平稳性，且影响力较大的城市均为区域的核心城市和航空枢纽。其他因素与疫情关联程度的空间分布与航空运输规模较为相似，影响力较大的地区以区域核心城市为主。本文研究成果有助于分析不同区域疫情发展对航空运输规模限制的敏感度，从而为科学防疫提供重要理论依据。未来研究可扩大数据收（下转第47页）