

# 人口迁徙格局在重大突发公共卫生事件的时空关联性研究

张 岩<sup>1</sup> 李英冰<sup>1</sup> 郑 翔<sup>2</sup>

<sup>1</sup> 武汉大学测绘学院,湖北 武汉,430079

<sup>2</sup> 武汉大学信息管理学院,湖北 武汉,430072

## Temporal and Spatial Correlation of Population Migration Pattern in Magnitude Outburst Public Health Incident

ZHANG Yan<sup>1</sup> LI Yingbing<sup>1</sup> ZHENG Xiang<sup>2</sup>

<sup>1</sup> School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China

<sup>2</sup> School of Information Management, Wuhan University, Wuhan 430072, China

**摘 要:**重大突发公共卫生事件是指某种传染病在短时间内发生,波及范围广泛,出现大量的病人或死亡病例,且发病率远超过常年发病水平。突发的新型冠状病毒(COVID-19)疫情对交通、旅游、文化、教育、娱乐、零售等产业产生了巨大破坏和深远影响,本文利用百度迁徙数据,分析人口在湖北省内城市之间迁徙的起止(origin-destination, OD)关系,提取了2020-01春运期间的省内净流量比指标矩阵,并进行主成分降维和地理加权回归计算。计算结果表明,武汉确诊人数显著高于模型预测结果,宜昌、荆门等地确诊人数显著低于模型预测结果,“封城”阻断了人口流动,使武汉与周边城市产生显著性差异,其他城市对来人口流动严格管控,防止了疫情继续蔓延。

**关键词:**人口迁徙格局;突发公共卫生事件;时空关联性;地理加权回归;疫情评估

**中图分类号:**P208;P954

**文献标志码:**A

**Abstract:** Major public health emergencies are the occurrence of a certain infectious disease in a short time, which spreads widely and causes a large number of patients or deaths. The incidence rate is far beyond the level of perennial disease. The outbreak of novel coronavirus (COVID-19) has caused great damage and far-reaching impact on transportation, tourism, culture, education, entertainment and retail industries. This paper uses Baidu migration data to analyze the relationship between population migration and migration in Hubei city, and extracts the index of net flow ratio in January of spring January. The main components are reduced and the geographical weighted regression is calculated. The re-

sults show that the number of confirmed cases in Wuhan is significantly higher than that predicted by the model, and the number of confirmed cases in Yichang, Jingmen and other places is significantly lower than that predicted by the model. The closure of the city blocks the flow of population, which makes a significant difference between Wuhan and the surrounding cities. Other cities control the flow of population strictly, thus prevent the spread of the epidemic.

**Key words:** population migration pattern; public health emergencies; temporal spatial correlation; GWR(geographically weighted regression); epidemic situation assessment

2020年春季新型冠状病毒肺炎暴发,为了阻止病毒在全国范围内蔓延,武汉于2020-01-23“封城”,全国都采取了积极控制措施。疫情爆发后,31个省份1386个县受到影响,其中大部分感染病例出现在湖北省<sup>[1]</sup>,Liu等<sup>[2]</sup>认为本次新冠疫情传染的基本再生数 $R_0$ 大于2003年非典疫情,均值约为3.28;Dong等<sup>[3]</sup>利用WebGIS技术,对全球疫情进行了可视化系统开发;Chinazzi等<sup>[4]</sup>认为武汉“封城”措施将整体疫情进展推迟了3~5d,模拟结果显示,社区流动限制辅以交通管制能够有效控制疫情蔓延。中国疾病预防控制中心编写了疫情进展和风险评估报告,服务于本次防疫控制。许多研究团队开展了疫情发展趋势预测研究,预测结果不尽相同,与疫情的实际发展也有差异,特别是在政策干预、人员流动阻隔等措施干预下,传染病的发展走向可能会被完全改变。闫磊等<sup>[5]</sup>对比了美国、欧洲与中国的流感病

**引用格式:**张岩,李英冰,郑翔.人口迁徙格局在重大突发公共卫生事件的时空关联性研究[J].测绘地理信息,2020,45(5):66-71(Zhang Yan, Li Yingbing, Zheng Xiang. Temporal and Spatial Correlation of Population Migration Pattern in Magnitude Outburst Public Health Incident[J]. Journal of Geomatics, 2020, 45(5): 66-71)

**基金项目:**国家重点研发计划(2018YFC0807000)。

毒检测与预报体系,并提出相关建议。陈江平等<sup>[6]</sup>利用空间自相关等方法,对甲型 H1N1 流感疫情的空间聚集性进行了讨论。刘嵩等<sup>[7]</sup>利用传统的 SIR (susceptible infected recovered) 模型对甲流感染进行模拟。徐敏等<sup>[8]</sup>利用时空扫描统计方法探索时空异常聚集范围,将地理信息系统应用于流感病毒的研究。

通过监测及时发现传染病爆发趋势、切断传染路径是传染病防控的重要举措。但传统的传染病监测系统存在监测范围局限、预警时间滞后等不足。随着大数据时代的到来,如互联网,移动通信等产生大量的关联数据,将之与传统监测数据相结合,提高监测的及时性和敏感性<sup>[9]</sup>。在此背景下,如何对重大传染病进行准确高效的监测,特别是人口迁徙对疫情发展的影响,受到广泛关注。

本文技术路线如图 1 所示,利用百度迁徙数据构建湖北省内城市 OD(origin-destination) 关系网络,构建了城市之间的时空净流量关系矩阵,利用主成分-地理加权回归模型,进行迁徙结果-确诊人数的时空关系检验。

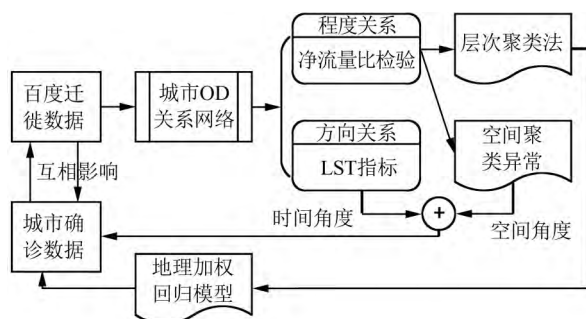


图 1 技术路线示意图

Fig. 1 Technical Route

## 1 数据获取与人口迁徙格局

### 1.1 迁徙数据获取

百度迁徙指数是对用户的签到记录、购票信息以及手机导航等数据的综合计算结果,实验获取 2020-01-01~01-30(农历腊月初七到元月初六),湖北省地级行政单位之间的 OD 流向与权重数据。

### 1.2 OD 计算

建立了湖北省 17 个地级市构成的 OD 流向网络,其中网络中节点的大小代表重要程度,连接边宽度代表强度<sup>[10]</sup>,坐标系采用 WGS-84 坐标系。

武汉市是我国疫情最严重的城市,也是湖北省劳务输入的主要城市,春节期间存在大量的人口流动。图 2 是武汉市与黄冈市人口迁徙对比图,1 月

23 日后武汉的迁入曲线和迁出曲线都下降严重,净增加曲线趋近于 0,呈现净减少状态。由于武汉市是大学生、外来务工人员较多的城市,由于“封城”影响与春运影响叠加,迁出高峰发生在 23 日以前。而黄冈市在该时段内一直呈现净流入状态。

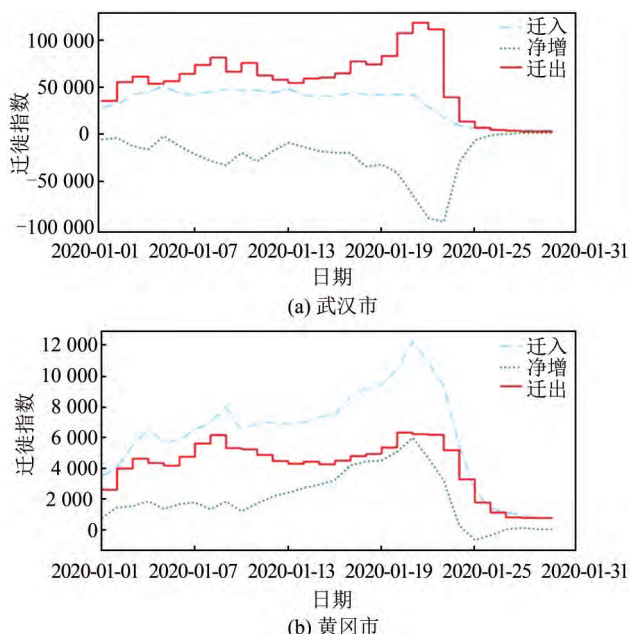


图 2 武汉市与黄冈市迁徙行为对比

Fig. 2 Comparison of Migration Behavior Between Wuhan and Huanggang

### 1.3 城市交通网络建立

依据人口在城市之间的迁徙关系和百度迁徙数据,对“封城”前的网络共绘制了 291 条具有方向性的有向连接,如图 3(a)所示。图 3 中线的粗细代表了人员流动的多少,如武汉与孝感两条连接边,较粗的一条代表武汉流向孝感迁徙数量,较细的一条代表孝感流向武汉数量。该时段内迁徙强度最高的线路是武汉-孝感,武汉-荆州和武汉-黄冈。

“封城”以后人口流动急剧衰减,图 3(b)是“封城”后的迁徙数据图形,以同等比例绘制了 284 条有向连接,武汉-孝感,武汉-黄冈依然是人数较多的迁徙路线。孝感、黄冈的防疫压力不容忽视,事实上,这两个城市疫情也很严重。

由于疫情爆发初期恰好在春运期间,而春运人员人口流动最多的时段,大量的人口流动使得疫情快速扩散。图 4 是湖北省内各城市在研究时间内累计迁徙的统计结果。除武汉市外,其余城市都出现人口净增长,且湖北省他城市的净增长高于武汉市的净减少。

图 5 是截止到 02-20 湖北省各地市累计确诊人数。通过分析迁徙数据,挖掘人口迁徙与感染人数之间的关联。





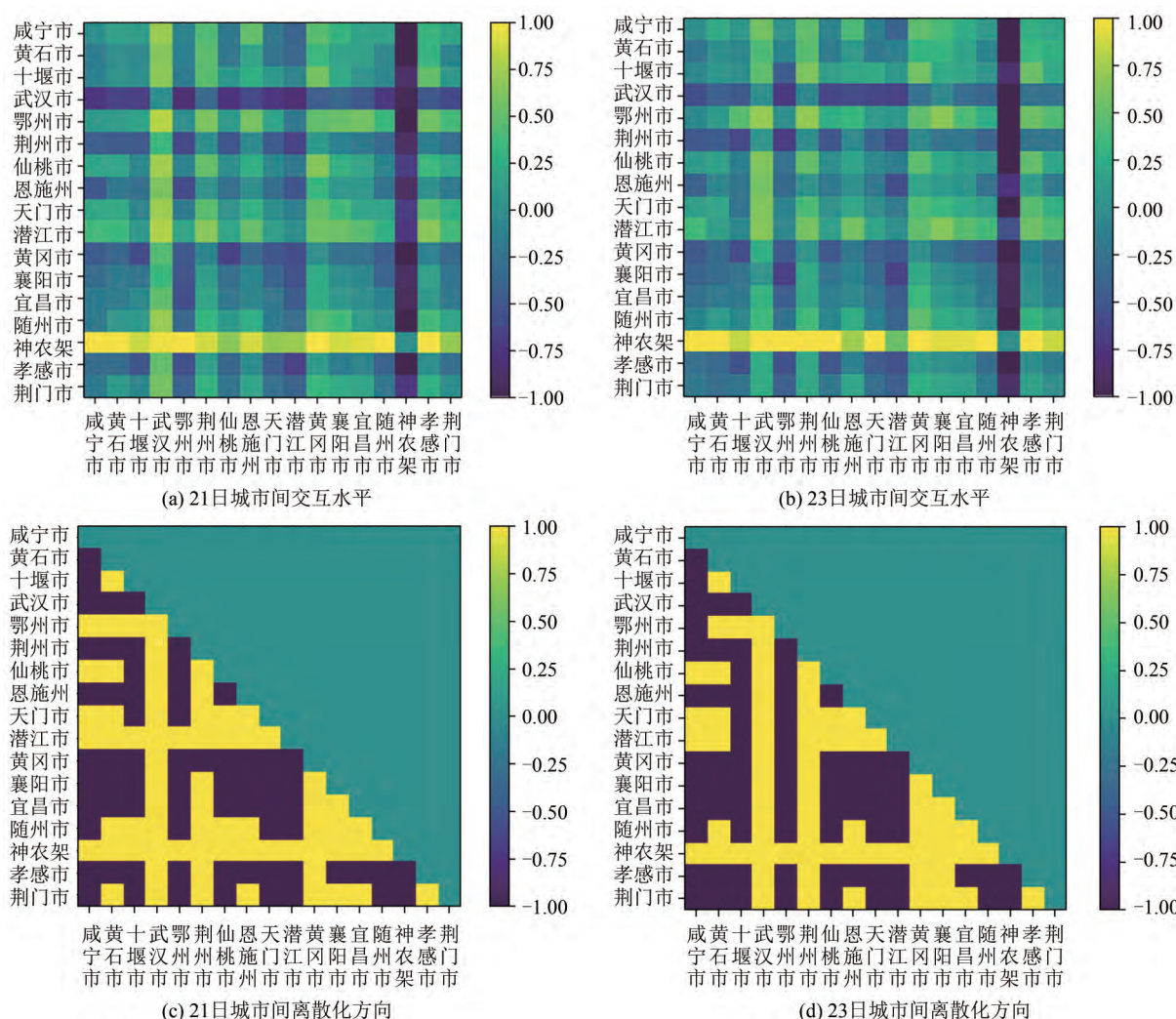


图 7 带有方向性的城市交互水平

Fig. 7 City Interaction Level with Directionality

州-咸宁由鄂州的净流入(为正)转移为鄂州流出至咸宁的趋势,潜江-十堰趋势由潜江的净流入转移为十堰流至至潜江的趋势,襄阳-荆州,宜昌-荆州,随州-仙桃,随州-恩施,随州-十堰,孝感-襄阳等城市对的流向都发生了逆转。

### 2.3 层次聚类分析

层次聚类法(hierarchical cluster methods)是无监督聚类方法,通过树状结构表示不同的类别划分,优点在于事先不需要对分类进行训练,并且可以形象地展示要素之间的层次结构。

在 NFR 矩阵基础上进行层次聚类分析,计算结果如图 8 所示。依据出行特征将城市划分为 6 类:黄石市、宜昌市、十堰市出行类型划为类 1,鄂州市、神农架出行类型划为类 2,黄冈市、恩施州、天门市出行类型划为类 3,襄阳市、随州市、荆州市、咸宁市出行类型划为类 4,潜江市、仙桃市、荆门市、孝感市出行类型划为类 5,武汉市的出行类别划为单独

一类。城市聚类结果显示,武汉市与其他城市区别很大,体现了武汉在湖北省内独特的地位。时间聚类结果显示,25 日至 30 日划分为一类,其余时间划分为一类。

### 2.4 空间自相关分析

空间自相关分析是认识空间分布特征的最常用方法,其计算公式如下:

$$I = \frac{N}{W_{ij}} \times \frac{\sum \sum W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{x_i - \bar{x}} \quad (3)$$

式中,  $N$  表示空间实体数目;  $x_i$  表示空间实体的属性值;  $W_{ij}=1$  表示空间实体  $i$  和  $j$  相邻,  $W_{ij}=0$  表示空间实体  $i$  和  $j$  不相邻。

莫兰指数是空间自相关的常用指标。指数为正表示空间要素“高-高”聚集或者“低-低”聚集,即具有同样高或同样低的属性值的要素邻近,且该要素是聚类的一部分,表示空间正相关。负值表示空间要素“高-低”聚集或“低-高”聚集,空间要素的邻近

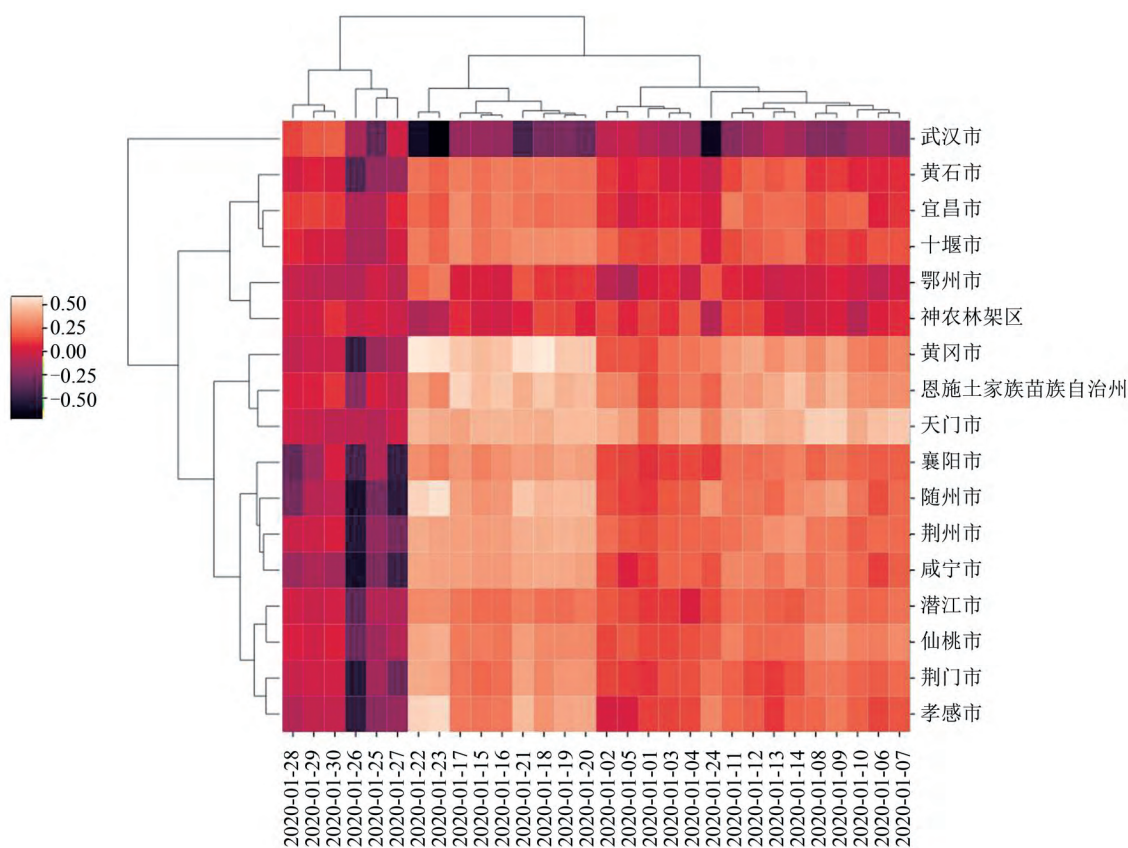


图 8 城市/时间层次聚类

Fig. 8 City/Time Level Clustering

要素为不同值,即该要素为异常值,其中又分为由低值围绕的高值“高-低”聚集和由高值围绕的低值“低-高”聚集,空间实体呈离散分布,指数为零表示空间实体是随机分布的。

抽取 21 日的 NFR 指数,对各个城市迁徙指数进行空间自相关分析,如图 9 所示,指数呈现负相关性,表示研究区域内城市吸引力的空间差异性较大。武汉市与神农架处于低-高区间,即区域内人口吸引力下降而周边区域吸引力上升,随州市处于高-高区间,即随州与周边省市都处于人口吸引力上升状态,黄石与鄂州处于低-低区间,表明这两个城市与周边城市都处于人口吸引力下降阶段,黄冈市处于高-低区间,表明此区域吸引力上升而周边区域吸引力下降。

## 2.5 主成分-地理加权回归

地理加权回归 (geographically weighted regression, GWR) 是是基于加权回归模型的空间扩展,将空间位置信息嵌入回归模型,考虑了区域之间存在的局部性差异,并且考虑到了变量的空间自相关<sup>[11]</sup>,计算公式为:

$$Y = \beta_0(\mu_i, v_i) \sum_{k=1}^n X_{ik} \beta_k(\mu_i, v_i) + \epsilon_i \quad (4)$$

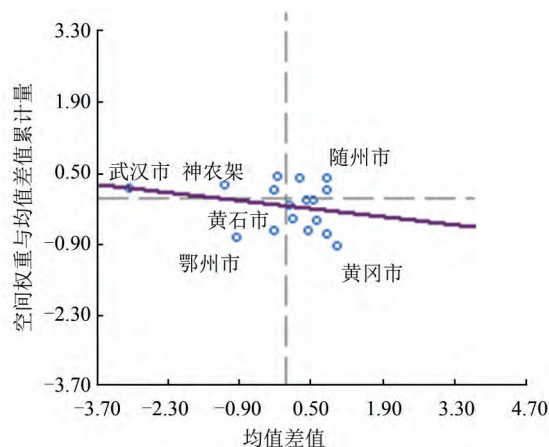


图 9 1 月 21 日的莫兰指数

Fig. 9 Moran's Index on January 21st

式中,  $(\mu_i, v_i)$  表示空间要素的位置;  $\beta_k(\mu_i, v_i)$  表示基于位置的特定系数估计。独立变量  $X_{ik}$  包括统计要素不同维度的值。利用 GWR 模型对湖北省每个城市计算  $t$  统计量,检验其在 0.05 水平即 95% 概率下的统计显著性和相关性。

考虑到图 8 生成的矩阵,每一个城市都有约 30 个不同的属性,因此采用主成分分析对矩阵进行降维,减少属性个数,再进行地理加权计算。通过对 NFR 矩阵降维,降维后前 3 个主成分分别占原信息 77.53%, 10.87%, 10.87%, 总计 99%。



建立主成分-地理加权回归模型,参数采用 NFR 矩阵降维后的结果,观测值为各城市的确诊人数。计算结果如图 10 所示,拟合  $R^2$  为 0.713 435,调整的  $R^2$  为 0.582 593,省内每一个城市的拟合  $R^2$  都在 0.7 以上,说明主成分降维的人口迁徙净流量比矩阵对感染情况有着一定的解释作用。

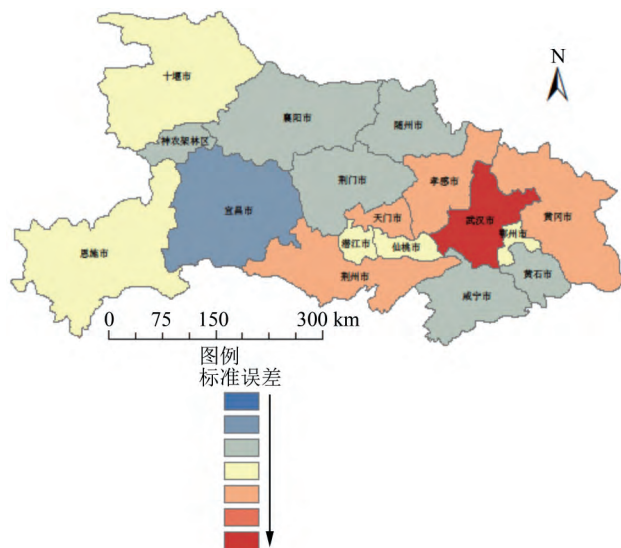


图 10 地理加权回归拟合结果

Fig. 10 Geographical Weighted Regression Fitting Results

在回归模型中,武汉市的差异较大,其感染结果显著高于回归值,原因在于武汉市是重灾区,“封城”阻断了人口流动,使武汉与周边城市产生显著性差异。黄冈、孝感、天门、荆州等地确诊人数高于预测模型,应对输入病例的压力较大,而宜昌市、随州、荆门、咸宁、黄石等地拟合结果低于实际确诊人数,如果仅考虑人口流动角度带来的防疫压力因素,这些城市对外来人口流动的有效管理防止了疫情的进一步扩散。

### 3 结束语

本文利用百度迁徙数据,结合主成分分析与地理加权回归的优点,建立确诊人数与降维后的前 3 个主成分之间的回归模型,分析了人口迁徙与感染人数之间的关联关系。

1)湖北省交通 OD 网络和迁徙过程 NFR 矩阵表明:武汉作为湖北省城市交通网络的中心城市,其人口迁徙特征与湖北省其他任何城市的人口迁徙特征差别都很大,呈现核心地位。武汉与其他城市处于武汉净流出的趋势,1月21日至23日出现净迁出高峰,孝感和黄冈两地是武汉流出人口最多的城市,相较于周边区域,城市对人口的吸引力偏高。疫情发展与人口迁徙存在关联关系。

2)建立新冠肺炎确诊人数与降维后的主成分之间的地理加权回归模型,计算的武汉市的实际感染结果显著高于回归值,说明武汉“封城”阻断了人口流动,使武汉与周边城市产生显著性差异,有效控制了新冠疫情的传播。宜昌、荆门、咸宁、黄石等地确诊人数显著低于模型预测结果,说明其他城市对外来人口流动的管理防止了疫情继续蔓延。

### 参考文献

- [1] 中国疾病预防控制中心新型冠状病毒肺炎应急响应机制流行病学组. 新型冠状病毒肺炎流行病学特征分析[J]. 中华流行病学杂志, 2020,41(2): 145-151
- [2] Liu Y, Gayle A A, Wilder-Smith A, et al. The Reproductive Number of COVID-19 Is Higher Compared to SARS Coronavirus[J]. Journal of Travel Medicine, 2020,27(2),doi:10.1093/jtm/taaa021
- [3] Dong E, Du H, Gardner L. An Interactive Web-Based Dashboard to Track COVID-19 in Real Time[J]. The Lancet Infectious Diseases, 2020,20(5):533-534
- [4] Chinazzi M, Davis J T, Ajelli M, et al. The Effect of Travel Restrictions on the Spread of the 2019 Novel Coronavirus (COVID-19) Outbreak [J]. Science, 2020,368(6 489):395-400
- [5] 闫磊,任玉环,黄晓霞,等. 国内外流感监测系统的地理流行病学比较研究[J]. 地球信息科学学报, 2010,12(3): 410-418
- [6] 陈江平,张兰兰,余远剑,等. 利用空间自相关的中国内陆甲型 H1N1 流感疫情分析[J]. 武汉大学学报·信息科学版, 2011,36(11): 1 363-1 366
- [7] 刘嵩,石润,张兰兰. 遗传算法在甲流 SIR 模型参数求解中的应用[J]. 测绘地理信息, 2012,37(6): 6-9
- [8] 徐敏,曹春香,程锦泉,等. 甲流感疫情时空聚集性的 GIS 分析[J]. 地球信息科学学报, 2010,12(5): 707-712
- [9] 李杰,赵阳. 基于 WebGIS 的突发事件网络舆情可视化设计与实现[J]. 测绘地理信息, 2014,39(4): 38-41
- [10] 张岩,李英冰,郑翔. 基于微博数据的台风“山竹”舆情演化时空分析[J/OL]. 山东大学学报(工学版),2020: 1-9[2020-04-21]. <http://kns.cnki.net/kcms/detail/37.1391.T.20200221.1529.004.html>
- [11] 周璐瑶,刘艳芳,刘耀林,等. 多中心视角下的武汉市路网与经济活动空间关系分析[J]. 测绘地理信息, 2020,45(2): 8-14

收稿日期:2020-05-14

第一作者:张岩,硕士生,主要研究领域为时空大数据分析等。  
E-mail:sggzhang@whu.edu.cn

通讯作者:李英冰,博士,副教授,主要研究方向为空间大数据分析。  
E-mail:ybli@whu.edu.cn