

COVID-19时空分布特征及影响因素研究

——以石家庄市为例

邓淑鑫¹, 裴鑫¹, 李明涛¹, 张娟², 柴玉珍¹

¹太原理工大学数学学院, 山西 太原

²山西大学复杂系统研究所, 山西 太原

收稿日期: 2022年4月23日; 录用日期: 2022年5月17日; 发布日期: 2022年5月24日

摘要

新型冠状病毒肺炎(COVID-19)疫情持续爆发, 对人民生活和社会经济发展产生重要影响。了解新冠肺炎疫情的时空格局和影响因素, 有助于控制疫情。本研究旨在分析石家庄市新冠肺炎疫情的时空分布特征及其与影响因素的关系。石家庄市疫情持续时间相对较短, 高发时间在一月中下旬。利用ArcGIS空间统计分析工具发现高发区域位于藁城区及其周边地区, 方向分布呈现沿“西北-东南”方向扩散的格局。采用多元回归模型发现生活和气象因素都对新冠肺炎的传播有影响, 但生活因素的影响更大。对模型参数进行提取发现生活因素对新冠肺炎的影响是正相关, 气象因素呈负相关。基于SPSS统计软件中Spearman相关系数分析发现距新发地的距离与确诊病例呈显著负相关。政策因素对新冠肺炎的发生有一定的影响, 管控在其中发挥较大作用。石家庄市新冠肺炎高发区域比较集中, 说明疫情防控措施及时有效; 生活环境政策因素等对新冠肺炎疫情的传播有一定影响, 且影响较明显; 研究也为元旦等人员流动较多情况下的流行病预防和控制提供一些建议。

关键词

新冠肺炎, 时空分布特征, 影响因素, 空间分析, 多元回归模型

Characteristics of Spatial and Temporal Distribution and Influencing Factors of COVID-19

—A Case Study of Shijiazhuang City

Shuxin Deng¹, Xin Pei¹, Mingtao Li¹, Juan Zhang², Yuzhen Chai¹

¹College of Mathematics, Taiyuan University of Technology, Taiyuan Shanxi

²Complex Systems Research Center, Shanxi University, Taiyuan Shanxi

Abstract

The outbreak of novel coronavirus (COVID-19) continues, which has an important impact on people's life and social and economic development. Understanding the temporal and spatial pattern and influencing factors of the epidemic situation in COVID-19 is helpful to control the epidemic situation. The purpose of this study is to analyze the temporal and spatial distribution characteristics of COVID-19 epidemic in Shijiazhuang and its relationship with influencing factors. The epidemic duration in Shijiazhuang is relatively short, with high incidence in the middle and late January. Using ArcGIS spatial statistical analysis tools, it is found that the high incidence area is located in Gaocheng District and its surrounding areas, and its directional distribution shows a pattern of spreading along the northwest-southeast direction. Multiple regression model shows that both life and meteorological factors have an impact on the spread of COVID-19, but life factors have a greater impact. By extracting the model parameters, it is found that the influence of life factors on COVID-19 is positively correlated, while that of meteorological factors is negatively correlated. Based on the analysis of Spearman correlation coefficient in SPSS statistical software, it is found that the distance from the new place of origin is negatively correlated with the confirmed cases. Policy factors have certain influence on the occurrence of COVID-19, in which control plays a greater role. The high incidence areas in Shijiazhuang COVID-19 are concentrated, which indicates that the epidemic prevention and control measures are timely and effective; life policy factors have certain influence on the spread of COVID-19 epidemic, and the influence is obvious. The research also provides some suggestions for epidemic prevention and control in the case of high turnover of people such as New Year's Day.

Keywords

COVID-19, Characteristics of Spatial and Temporal Distribution, Influencing Factors, Spatial Analysis, Multiple Regression Model

Copyright © 2022 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

新型冠状病毒肺炎(Corona Virus Disease 2019, COVID-19), 简称“新冠肺炎”, 2019年12月以来, 疫情已席卷世界200多个国家。世界卫生组织(WHO)于2020年1月30日宣布新冠肺炎疫情为“国际关注的突发公共卫生事件”[1]。同日, 国际病毒分类委员会(ICTV)提议将新型冠状病毒命名为 Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) [2]。随后于2020年3月11日世界卫生组织(WHO)宣布 COVID-19 成为大流行。

截至2022年3月16日, 中国累计报告了123,773例新冠肺炎确诊病例, 包括4636例死亡[3]。同一天, 全球报告超过4亿的累计确诊新冠肺炎病例, 包括6,058,322例死亡, 其在约翰霍普金斯大学新冠肺炎地图仪表盘上被在线跟踪和绘制[4]。目前, 中国的新冠肺炎疫情已经得到控制, 情况有了明显改善。然而, 由于全球确诊新冠肺炎病例的快速增长, 全球新冠肺炎大流行急剧恶化。石家庄市疫情的暴发正处于中国春节期间, 人口大量流动, 导致疫情在极短时间内扩散。新冠肺炎疫情对国内甚至全世界都产

生了极其重要的影响, 从不同专业视角对疫情进行多方面探讨, 可以为国家或地区的疫情防控提供有效支撑。

石家庄市 COVID-19 疫情的传播大致可以归纳为“病例聚集 + 流动扩散”的模式传播, 最初暴发地的感染者通过社区传染、聚集活动(如: 参加婚宴)等形式形成疫情扩散的次级暴发点, 再通过人口流动形成下一个次级暴发点, 导致疫情在人口密集的地区大幅暴发。石家庄市此次疫情与之前北京、大连等地区爆发的疫情不同点在于石家庄市疫情的暴发点为农村, 居民防控意识薄弱, 文化水平较低, 政府监控力度不够。因此, 研究病例的流行趋势、方向分布、聚集性等时空分布特征对于农村疫情防控具有重要意义。

自新冠疫情发生以来, 通过统计方法分析疫情时空分布特征的工作已有很多。例如: Kang 等[5]利用空间自相关分析, 发现新冠疫情在中国省级水平上存在显著的空间相关性; Mo 等[6]利用时空立方体的方法对中国新冠疫情病例数据进行分析, 给出各城市的疾病热点发病区域; Xiong 等[7]利用 Anselin Local Moran's I 发现湖北省病例在武汉、孝感、鄂州等中心区域聚集; Desjardins 等[8]基于 Sat Scan 软件进行时空统计扫描, 检测出美国各市县的疫情活跃集群。孙建平等[9]利用标准差椭圆发现安徽省新冠肺炎疫情在县域水平上存在较强的方向性。然而, 对新冠肺炎疫情的空间统计特征和影响因素的研究仍然缺乏, 尤其是从地市级和县级尺度上的研究, 而这对于理解新冠肺炎疫情的空间演化机制和制定防控政策具有重要作用。

疫情聚集性影响因素研究的主要思路是将病例信息和地理要素相关联, 研究者多使用多元回归分析、地理加权回归[10]等方法, 或采用仿真模型估计某种具体措施的实施对疫情的影响[11]等。新冠肺炎暴发适逢春运, 人口流动、交通运输、城市经济、政府政策等因素对疫情暴发起到了推动作用。Jia 等[12]发现疫情早期中国各省的累积确诊病例与从武汉输入当地的人数相关系数为 0.919, 而与当地人口数的相关性不显著; 王姣娥等[13]研究中国各省的疫情空间扩散情况, 发现地理邻近性、人口流动、人口规模、交通网络、疫情防控管理等因素对疫情的空间扩散具有显著影响; 汪冉等[14]证实河南省疫情与交通通达性呈现显著相关关系; Xiong 等[7]发现湖北各县市的疫情与人口、经济等因素相关; Mollalo 等[15]发现美国各县市家庭收入中位数、收入不平等性、护士从业者百分比、黑人女性百分比的分布对新冠肺炎发病率具有较高解释力。总之, 一些自然、社会和经济变量, 如相对湿度、人口和 GDP 等被发现与新冠肺炎疫情显著相关, 大多是在相对较大的规模(大城市或大省)。因此, 在较小的范围内研究影响新冠肺炎传播的因素也很重要。

此外, 目前关于气象因素对新冠肺炎传播影响的研究结果仍存在争议。Ujiie M 等[16]发现温度可在一定程度上影响新冠肺炎的传播; 而 Tosepu R 等[17]的研究发现温度与新冠肺炎之间并无显著关联。不仅以上在某一局限区域内开展的研究结果互相矛盾, 在较大范围内开展的研究结果也不一致。Xie J 等[18]对中国 122 个城市进行的研究表明, 当平均温度低于 3℃时, 每日确诊病例与温度呈正相关, 但当平均温度高于 3℃时, 两者相关性不显著; 而 Wu Y 等[19]对覆盖世界范围内 166 个国家的研究发现温度和相对湿度与每天的新增病例数和死亡人数呈负相关。

基于以上情况, 本文针对 2021 年 1 月 2 日至 2 月 14 日石家庄市发生的新冠肺炎疫情, 运用 ArcGIS 软件分析其病例的时空分布特征。主要内容包括: ① 利用描述性统计方法分析其时间、空间分布和病例属性结构特征; ② 利用方向分布分析其空间分布的方向性特征; ③ 利用局部自相关分析, 挖掘其热点发病区域以及聚类和异常值。用 R 软件分析各种影响因素对新冠肺炎疫情的影响。主要内容包括: ① 用 KMO 球和巴特利特球检验生活因素和气象因素, 看它们是否适合进行因子分析; ② 进行主成分分析和因子分析, 选择合适的分析方法对影响因素进行整理和重组; ③ 对分析结果和确诊病例进行五种建模分析, 并评价其适用性, 以选择最适合解决此类问题的模型, 并得出影响因素对新冠肺炎确诊病例的作用; ④ 采用多元线性回归模型分析政策因素对新冠肺炎疫情的影响。本研究的目的是: ① 丰富对疫情时空

演变的研究；② 了解影响新冠肺炎传播的因素；③ 为科学防控新冠肺炎提供有用信息。

2. 研究材料与方法

2.1. 研究区域

河北省石家庄市是此次河北省新冠肺炎疫情的高发区域。本文以石家庄市的 8 个区、12 个县、4 个县级市和 1 个副地级市的县级单位作为研究区域,利用新冠肺炎发生的时间和地理位置等数据,分析 2021 年 1 月 2 日至 2 月 14 日疾病在整个石家庄市的时空传播特征。同时,以新冠肺炎聚集区域藁城区为研究对象,以镇(街道)级区域为研究单位,分析该区域的疾病时空传播特征。

2.2. 数据来源

从河北省石家庄市卫生健康委员会网站(<http://wsjkw.hebei.gov.cn/>)收集 2021 年 1 月 2 日至 2 月 14 日新冠疫情监测数据,包括确诊病例的确诊日期、地理位置、行动轨迹等信息。以各市地理编码为关联字段,将疫情信息表与矢量地图相连接,建立石家庄市的新冠疫情地理信息系统。

为研究生活环境、气象、距离、政策因素对新冠肺炎的影响,从石家庄市统计年鉴中获取石家庄市各区县的 GDP 水平、人口密度、医疗机构个数、城镇化率及财产情况;从石家庄市气象局获取各区县的降水量、温度、湿度、风速和日照时数;距离的研究以这次疫情中第一例病例所在村的中心为新发地,通过计算确诊病例所在地与新发地之间的距离,再与确诊人数做相关性分析来研究距离对新冠肺炎的影响。政策则从政策公布的措施中提取重要政策来研究。

2.3. 研究方法

2.3.1. 方向分布

方向分布可用于研究数据空间分布的方向性特征,通过标准差椭圆可以确定研究对象的分布中心及其扩散程度,识别研究对象的数据分布方向和其扩散趋势[20]。椭圆面要素的生成以要素分布的平均中心为椭圆中心,其在 X 和 Y 两个方向上的标准差距离用来作为椭圆的长、短轴。椭圆长轴与短轴分别表示研究对象在空间中分布得最多与最少的方向。椭圆的面积表征离散程度,面积越大则离散程度越大,反之越聚集。椭圆的方向表示扩散的方向[21]。

2.3.2. 热点分析

热点分析是研究局部空间自相关的一种方法,可以识别高值或低值要素在空间上发生聚类的位置,即热点和冷点的分布。当一个要素具有高值,并且其周围要素也是高值,就是统计学上的热点;反之为冷点[22]。通常通过 G_i^* 统计量来衡量,在统计显著性水平下, $G_i^* > 0$ 意味着该区域为热点区域,即高值聚类; $G_i^* < 0$ 意味着该区域为冷点区域,即低值聚类; $G_i^* = 0$ 意味着呈随机分布,不存在空间聚类[23]。

2.3.3. 聚类与异常值分析

聚类和异常值分析可以探测研究区域内各县(区)与邻近县(区)的新冠肺炎发病聚集区域和异常值,一般用局部 Moran's I 指标来度量,空间 i 的局部 Moran's I 统计量 $I_i > 0$,表示相邻空间位置都具有高值或都具有低值,即或为高值的聚集或为低值的聚集; $I_i < 0$,表示为空间异常值。该统计量和 Z 得分、P 值可将局部空间分布划分为四种类型($P < 0.05$):高值聚类(HH)、低值聚类(LL)、高值由低值围绕的异常值(HL)、低值由高值围绕的异常值(LH) [24]。

2.3.4. 主成分分析

主成分分析的目的是将原来的变量通过线性组合压缩成少数的几个综合变量,但往往不具有明显的

现实意义[25]; 因子分析则是期望将一些相关联的变量压缩, 通过线性组合得到一些公共因子, 这些公共因子对变量起解释作用且往往有明显的现实意义。气象和生活因素用 KMO 和巴特利特球检验它们是否适合于因子分析。

2.3.5. 模型选择

模型建立采用泊松回归(PR)、负二项回归(NBR)、障碍回归(HR)、零膨胀泊松回归(ZIPR)和零膨胀负二项回归(ZINBR)五种计数数据回归模型, 了解生活环境与气象因素对石家庄市新冠肺炎疫情的影响。

基于石家庄市疫情发生的时间线来看, 考虑石家庄市的几项重点政策: 1月5号的闭环管控、1月6号的全市全员检测、1月8号的居家隔离、1月12日的封闭管理, 这几项政策记为管控(C_t)、检测(T_t)、隔离(I_t)、封锁(L_t)。假设这些政策与新冠每日确诊病例之间存在线性关系, 使用多元线性回归建模如下[26]:

$$y_t = \alpha_0 + \alpha_c \cdot C_t + \alpha_T \cdot T_t + \alpha_I \cdot I_t + \alpha_L \cdot L_t + \varepsilon_t$$

其中 y_t 表示每日确诊病例, α_0 、 α_c 、 α_T 、 α_I 、 α_L 分别表示无策略、管控、检测、隔离、封锁的回归系数。 ε_t 为残差项, 假设其期望为零, 方差不变且独立同分布。其中封锁为一个分类变量(根据其是否封锁, $L_t = 0/1$), 其余三个策略为数值。

3. 结果

3.1. 流行病学特征

3.1.1. 石家庄市新冠肺炎疫情的发展趋势

图1呈现了2021年1月2日至2月14日石家庄市新冠肺炎病例新增和累计数据, 结果显示: 石家庄市新增确诊病例总量在2021年1月8日出现拐点; 在2021年1月14日上升至峰值, 随后呈逐步下降的趋势; 截至2月14日, 新增确诊病例数已降至1例。疫情的发展趋势与石家庄市采取的封城、居家隔离等防控措施密切相关。

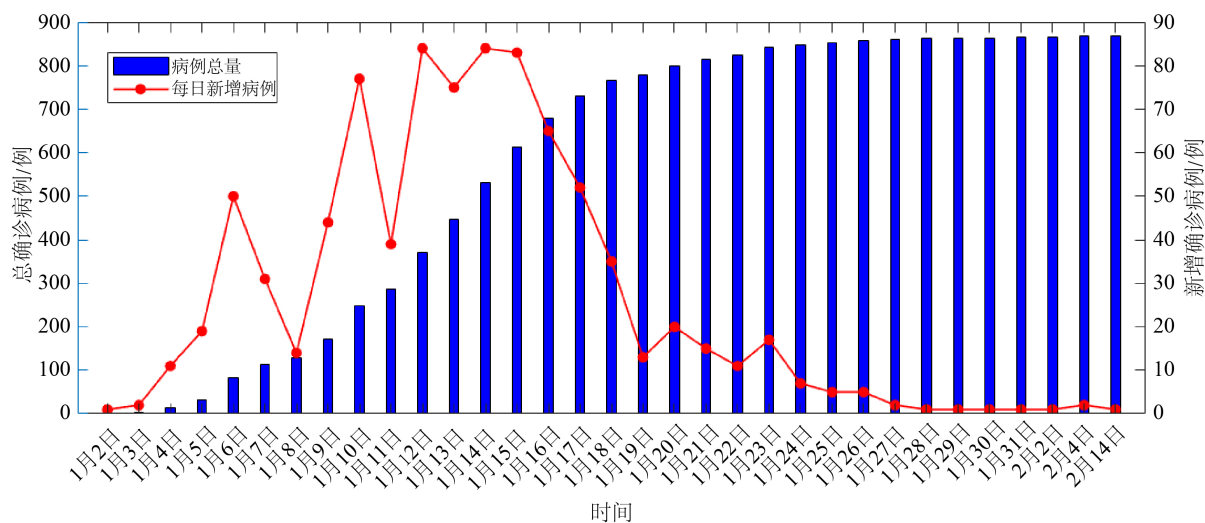


Figure 1. Development trend of total number of COVID-19 cases and new daily increment in Shijiazhuang

图1. 石家庄市新冠肺炎总确诊病例与每日新增确诊病例的发展趋势

3.1.2. 石家庄市 COVID-19 确诊病例的地理分布

石家庄市于1月8日实施居家隔离措施, 1月16日确诊病例开始逐渐下降, 1月23日各区县实行分

区分级管控，2月1日无新增确诊病例，所以我们基于 ArcGIS 软件绘制以下五个时间段下各区县新增确诊病例时空分布图，呈现石家庄市新冠肺炎确诊病例空间分布随时间的变化趋势(如图 2 所示)。

结果显示：疫情主要集中在藁城区及周边区县，前三个时间段病例较多，第四个时间段各区县确诊病例数量降低，最后新增确诊病例仅在藁城区出现，且数量很少，说明疫情防控工作显现初步成效，疫情传播范围得到控制。另外，我们发现五个时间段中藁城区的确诊病例都处于高发病人数水平，其周围的新乐市、正定县等区县次之，剩余大多数区县处于低发病水平。

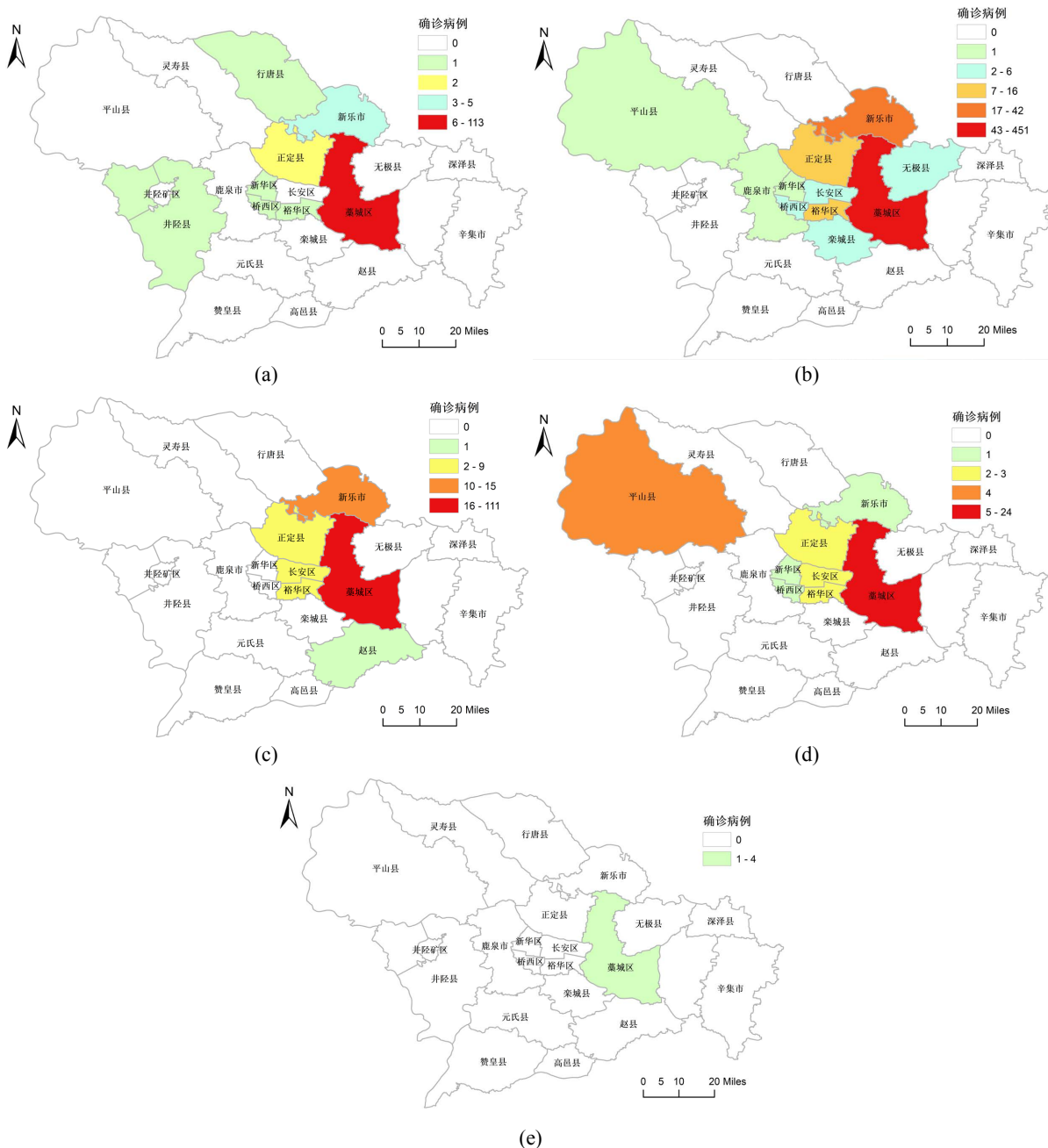


Figure 2. Spatial distribution of COVID-19 at district or county level in Shijiazhuang (a) 1.2~1.8; (b) 1.9~1.16; (c) 1.17~1.22; (d) 1.23~1.31; (e) 2.1~2.14

图 2. 新冠肺炎在石家庄市(区)县级水平空间分布图(a) 1.2~1.8; (b) 1.9~1.16; (c) 1.17~1.22; (d) 1.23~1.31; (e) 2.1~2.14

3.1.3. 病例的年龄、性别、城乡结构特征

根据所收集的病例监测数据,对此次新冠疫情确诊病例的属性结构进行分析。从年龄分布来看,按照我国现行的年龄划分标准,将年龄划分为童年(0~6岁)、少年(7~17岁)、青年(18~40岁)、中年(41~65岁)、老年(66岁以上)五个年龄段。在石家庄市新冠肺炎确诊病例中,中年(39%)和青年(29%)较多,老年人占比16%,少年占比10%,童年病例较少(见图3(b)),中年和青年占比较多是由于他们社交活跃、与确诊病例和病毒携带者接触较多、参加社区活动等;从性别分布来看,女性占比高于男性(见图3(a)),可能与身体素质、行为特征、活跃群体有关;从城乡占比来看,石家庄市疫情主要集中在农村(见图3(c)),这与此次疫情农村为爆发点有着高度的一致性。

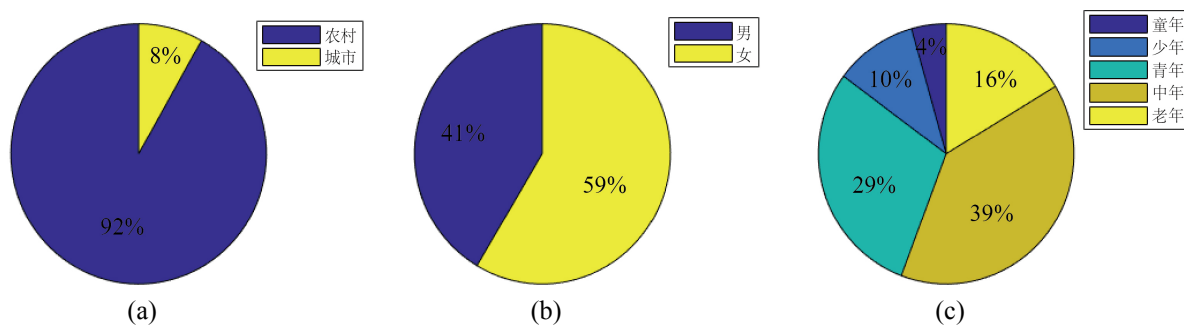


Figure 3. Age, gender and urban-rural structure of confirmed COVID-19 cases in Shijiazhuang
图3. 石家庄市新冠肺炎确诊病例年龄、性别、城乡结构

3.2. 时空分布特征

2021年1月2日石家庄市新增1例新冠肺炎确诊病例,疫情防控指挥部办公室于1月8日发布通告:为了防止疫情扩散,保障人民群众健康安全,广大群众居家7天,有效防控疫情,人员不流动、不聚集、不出社区和村庄[1],2月14日,此波疫情得到控制;且此次疫情,藁城区为石家庄市高发区域,其病例数为703例,占石家庄市总病例数的81%。为此,本节将时间分为两个时间段:1月2日至1月8日(采取居家令措施前)和1月9日至2月14日(采取居家令措施后),分析石家庄市和藁城区镇(街道)级水平的病例空间分布特征。

3.2.1. 可视化

石家庄市及藁城区按镇(街道)级水平的病例空间分布图见图4,并在藁城区可视化地图上对藁城区的村(乡)级水平的确诊病例数以点状图形式进行呈现。可以看出:两个时间段中藁城区的增村镇发病人数最多,疫情情况最为严重;1月2日至1月8日,病例集中在石家庄市的5个镇,其中增村镇的确诊病例最多,其余地区确诊病例均为1例;1月8日之后病例开始向外传播,涉及到的村镇也明显增多,但高发区域还是在增村镇。从藁城区来看,两个时间段中,小果庄村和刘家佐村疫情情况最为严重,南桥寨村和东桥寨村疫情情况较严重,其他村处于低发病水平。可以说明居家令颁布时,疫情已经传播开来,但因新冠肺炎有潜伏期的存在,采取居家令措施导致第二个时间段疫情比较严重,传播范围更广。

3.2.2. 方向分布

本节对石家庄市和藁城区新冠肺炎确诊病例进行分析,结果(见图5和表1):2个时间段中病例总体具有明显的方向性和向心性,前一时间段椭圆面积较小,病例较集中,之后椭圆开始扩张,病例开始逐渐扩散,在第2个时间段椭圆面积较大。椭圆的方向趋势也越来越显著,呈现沿“西北-东南”方向扩散的分布格局。

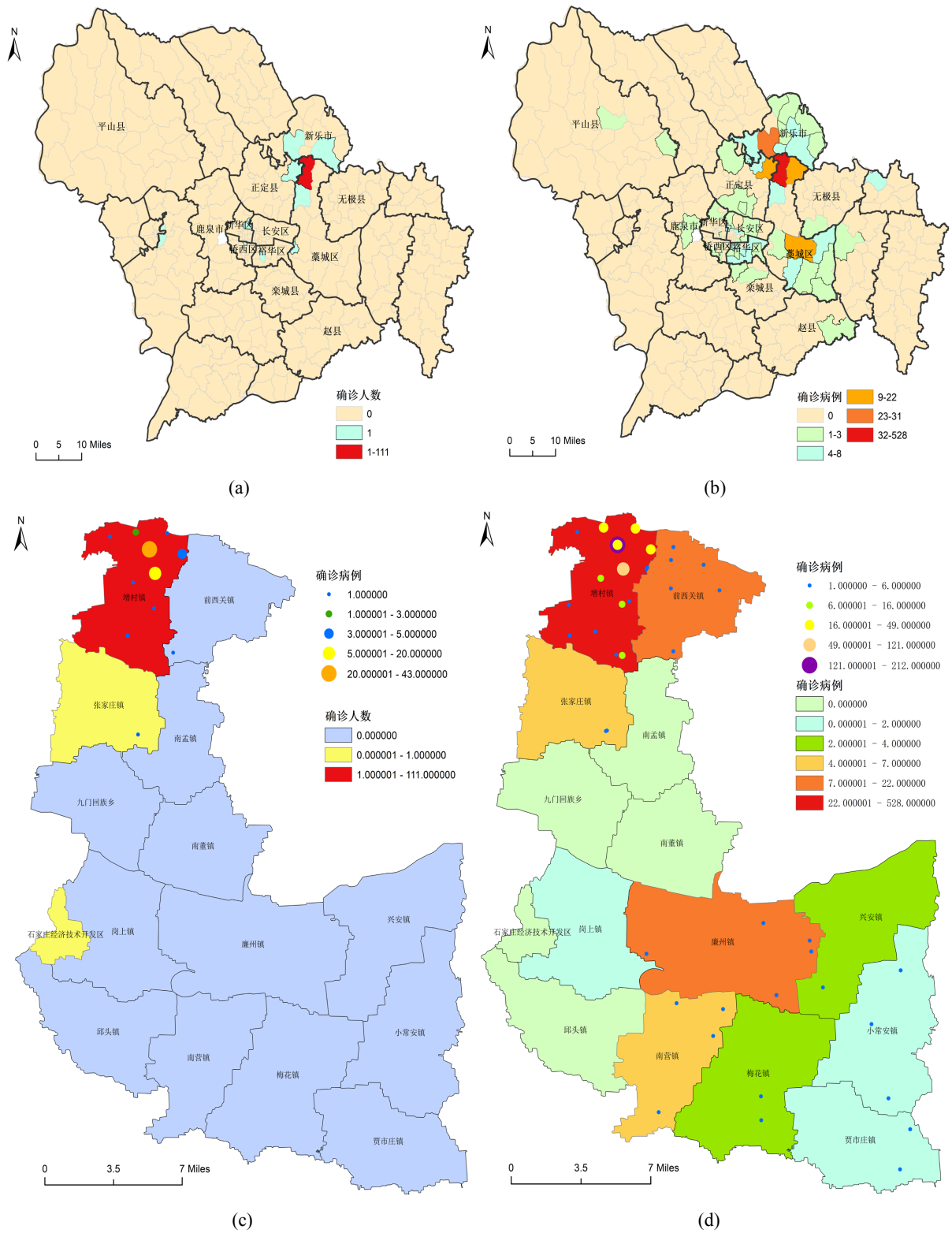


Figure 4. Spatial distribution of COVID-19 town or subdistrict level (a) Shijiazhuang 1.2~1.8; (b) Shijiazhuang 1.9~2.14; (c) Gaocheng 1.2~1.8; (d) Gaocheng 1.9~2.14

图 4. 新冠肺炎镇(街道级)水平空间分布图(a) 石家庄市 1.2~1.8; (b) 石家庄市 1.9~2.14; (c) 藁城区 1.2~1.8; (d) 藁城区 1.9~2.14

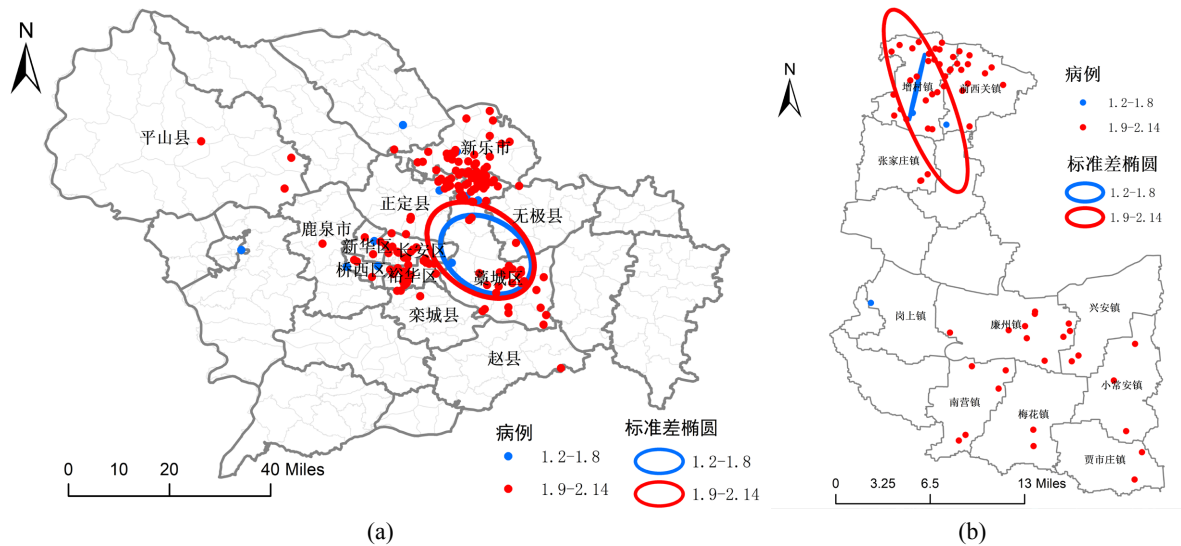


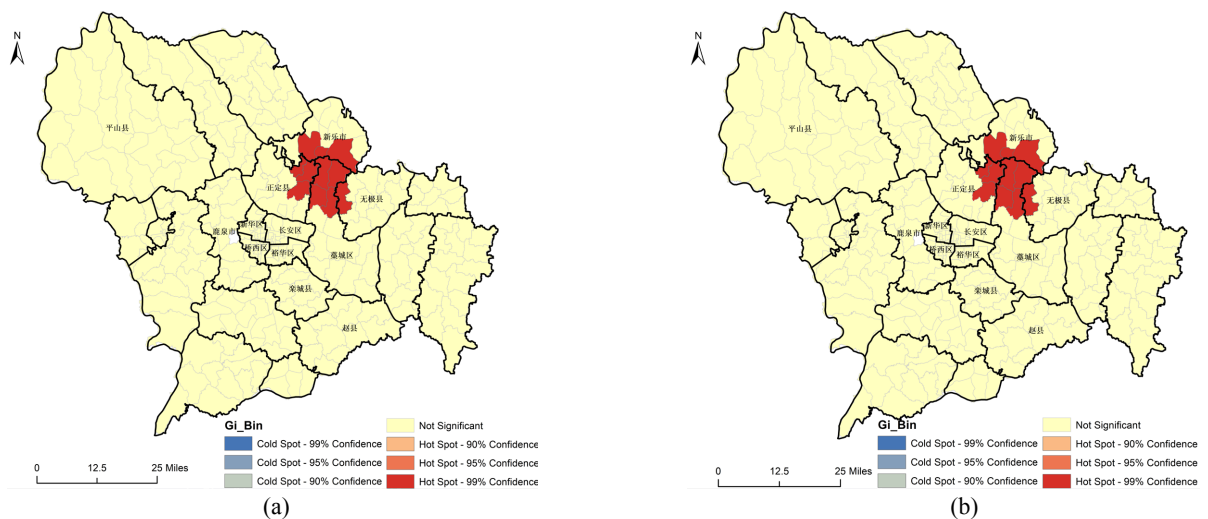
Figure 5. Directional distribution of COVID-19 (a) Shijiazhuang; (b) Gaocheng
图 5. 新冠肺炎方向分布(a) 石家庄市; (b) 藁城区

Table 1. Results of elliptical parameters in confirmed cases of COVID-19
表 1. 新冠确诊病例椭圆参数结果

地区	时间	短半轴	长半轴	面积	方向角度
石家庄市	1.2~1.8	0.100752	0.139106	0.044017	121.014752
	1.9~2.14	0.117181	0.169009	0.062205	125.329038
藁城区	1.2~1.8	0.000496	0.034917	0.000017	9.941994
	1.9~2.14	0.020422	0.100159	0.006423	164.732803

3.2.3. 热点分析

本节运用 ArcGIS 地理统计分析功能对石家庄市和藁城区的确诊病例分别进行热点分析[27] (见图 6)。结果显示：以石家庄市为研究对象，两个时间段中热点区域都集中在藁城区的 4 个镇，以及藁城区周围的正定县、新乐市和无极县的部分区(县)；以藁城区为研究对象，热点区域都集中在增村镇和前西关镇；说明疫情防控措施的实施，使得疫情传播得到有效控制。



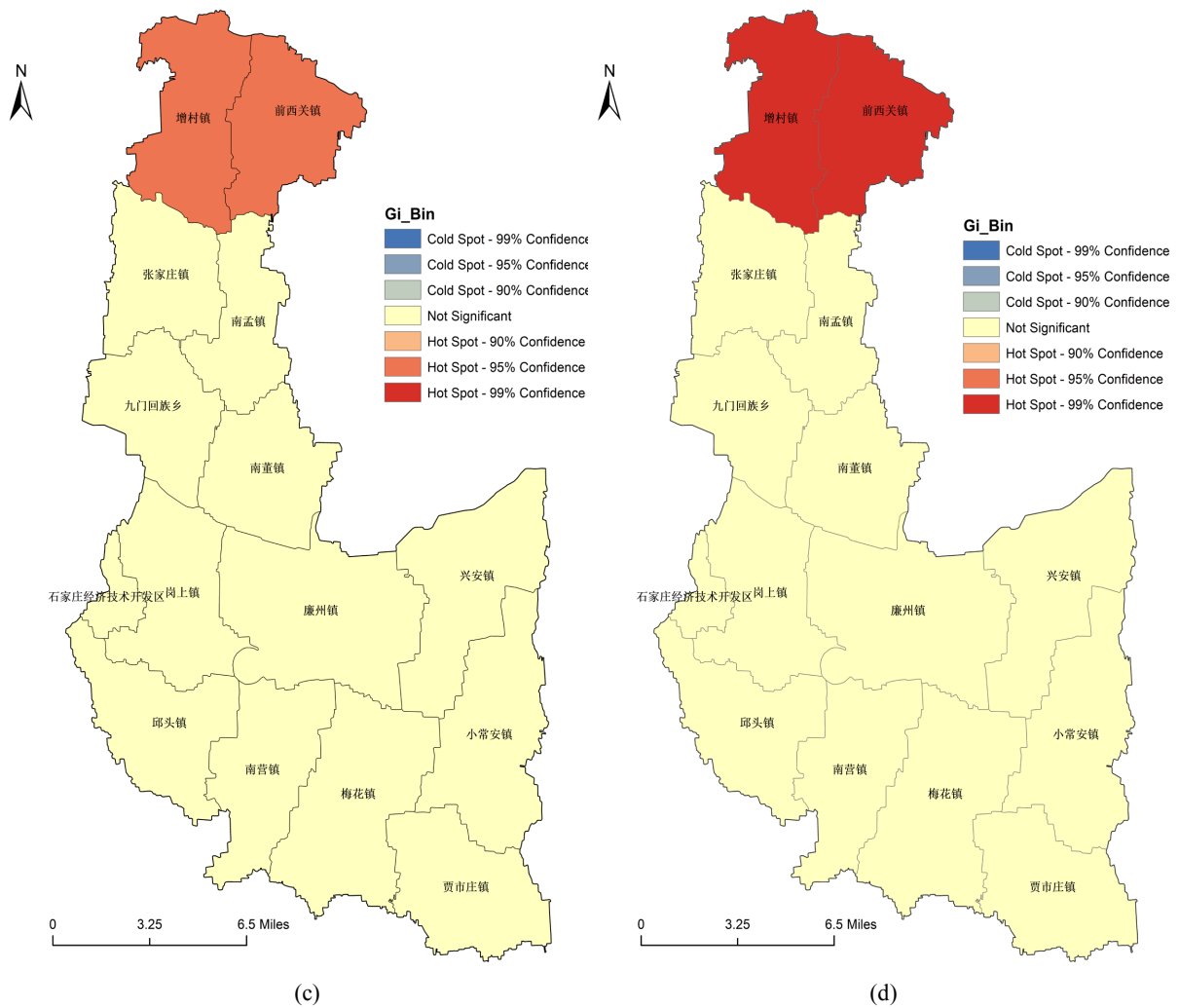


Figure 6. Hot spot analysis (a) Shijiazhuang 1.2~1.8; (b) Shijiazhuang 1.9~2.14; (c) Gaocheng 1.2~1.8; (d) Gaocheng 1.9~2.14

图 6. 热点分析(a) 石家庄市 1.2~1.8; (b) 石家庄市 1.9~2.14; (c) 藁城区 1.2~1.8; (d) 藁城区 1.9~2.14

3.2.4. 聚类与异常值分析

本节利用局部 Moran's I 指数对石家庄市和藁城区的确诊病例分别进行聚类 and 异常值分析[28] (见图 7)。结果显示: 1月2日到1月8日期间新冠肺炎确诊人数在石家庄市镇(街道)有 HH 和 LH 两种聚集模式, 高值聚类位于藁城区张家庄镇、新乐市邯邰镇、正定县新城铺镇, 意味着该区域与其周围区域都有高病例数, 藁城区前西关镇为 LH, 意味着该区域病例数较少而周围病例数较多; 1月9日至2月14日时间段内仅有 HH (高值聚类)一种聚集模式, 位于藁城区增村增、张家庄镇、前西关镇和新乐市邯邰镇、正定县新城铺镇, 意味着该区域与其周围区域都有高的病例数; 其他镇(街道)的病例数为零或者无统计学意义, 是随机分布的。

以藁城区为研究对象, 结果显示: 1月2日至1月8日期间无聚类和异常值; 1月9日至2月14日内增村镇为 HL, 前西关镇为 LH, 说明此时间段增村镇的发病人数较高, 前西关镇的发病人数较低, 其他各镇(街道)的确诊人数为零或者无统计学意义, 是随机分布的。同时说明其他各镇的疫情主要由前期藁城区增村镇引入, 各镇之间的疫情相关性不大, 说明居家令措施起到作用。

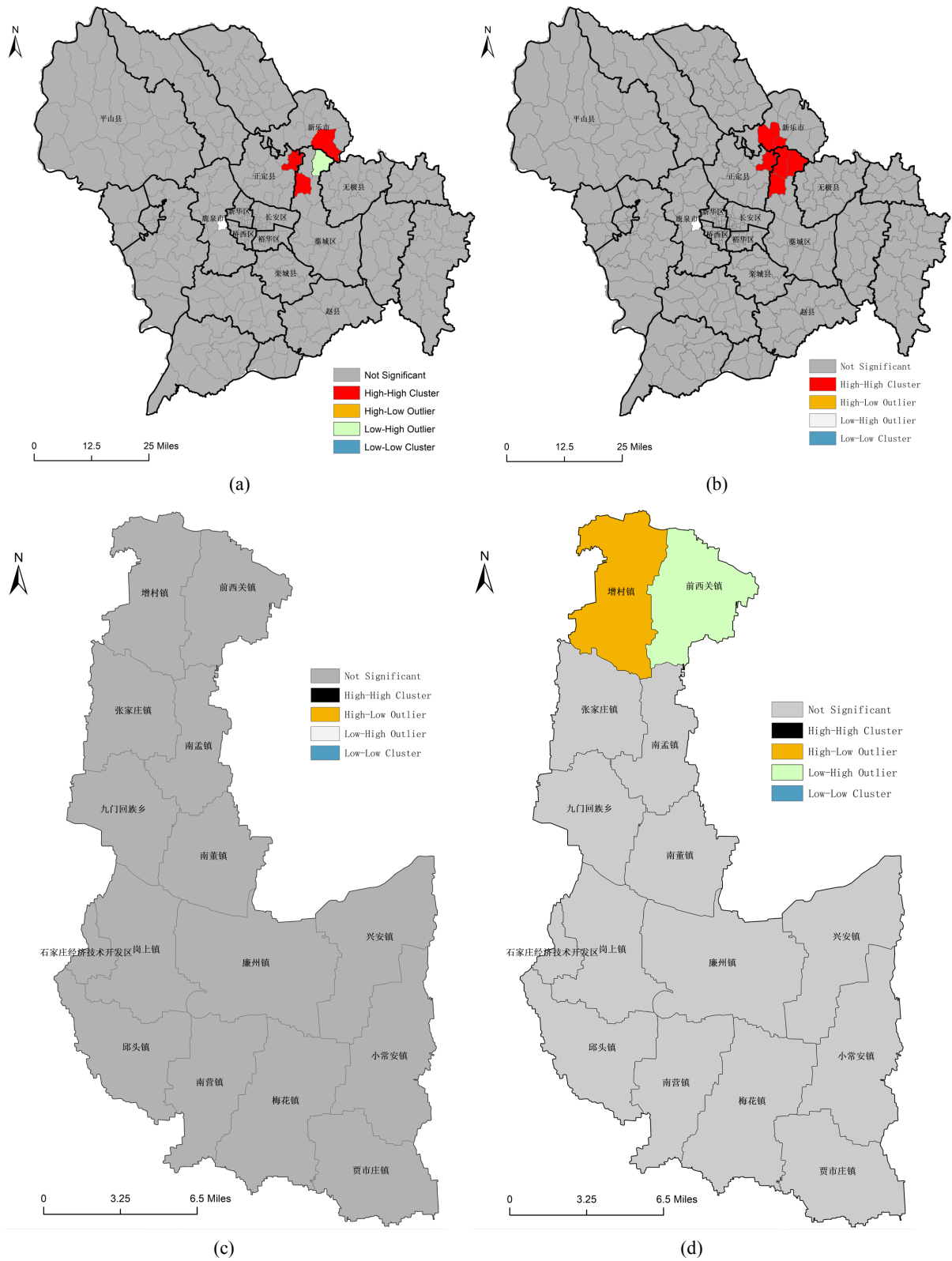


Figure 7. Clustering and outlier analysis (a) Shijiazhuang 1.2~1.8; (b) Shijiazhuang 1.9~2.14; (c) Gaocheng 1.2~1.8; (d) Gaocheng 1.9~2.14

图 7. 聚类与异常值分析(a)石家庄市 1.2~1.8; (b) 石家庄市 1.9~2.14; (c) 藁城区 1.2~1.8; (d) 藁城区 1.9~2.14

3.3. 影响因素研究

对新冠肺炎疫情影响因素的研究有助于对其传播机制和防控的最新认识。如上所述，石家庄市新冠肺炎的流行呈现明显的集中爆发和流动扩散。然而，具有这些特征的新冠肺炎疫情的影响因素仍然没有得到解决。为了给新冠肺炎疫情危险因素的识别和预防提供帮助信息，本研究利用 R 统计软件进行相关性分析，从经济、自然、政策等方面探讨了石家庄市县级新冠肺炎传播的影响因素。我们特意选取了 15 个指标，分别是 5 个生活参数(工资、人口密度、城市化率、GDP、医疗机构数量)、5 个气象指标(降雨量、风速、日照时数、湿度、温度)、1 个距离指标(距新发地距离)和 4 个政策指标进行相关性分析。特别地，本研究使用了 2021 年 1 月 2 日至 2 月 14 日全市 13 个空间单元进行相关性分析。

3.3.1. 生活因素与气象因素

将生活环境因素和气象因素一起研究。我们假设影响因素与 COVID-19 的确诊病例之间存在着线性关系。首先，对收集到的因素进行检验与因子分析。然后，将因子分析的结果作为自变量，与 COVID-19 的确诊病例进行建模和分析。最后，对模型的结果进行分析和总结，结论如下。利用 R 语言的主成分分析和因子分析，影响结果的因素见以下表 2 和图 8。

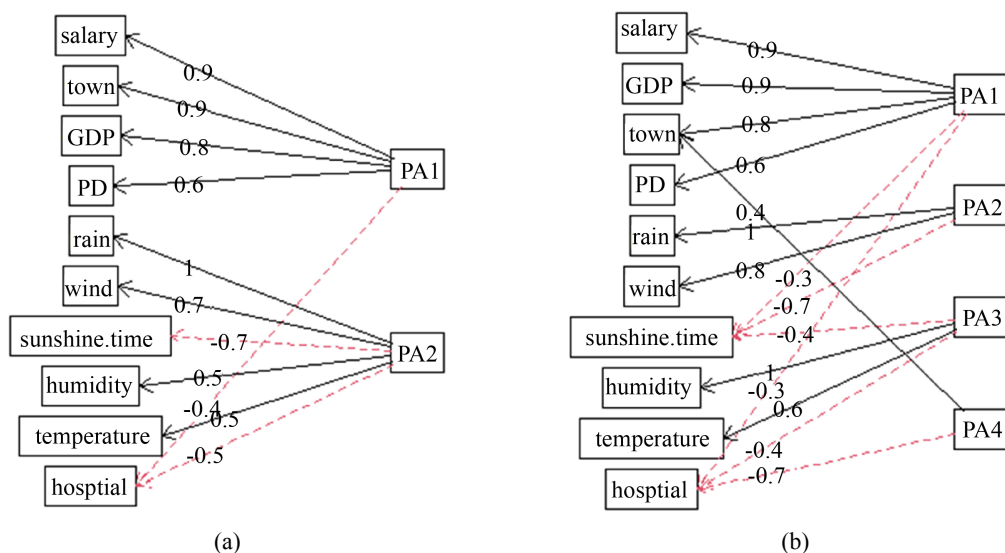


Figure 8. Factor analysis (a) Two factors; (b) Four factors
图 8. 因子分析(a) 两个因子; (b) 四个因子

Table 2. The results of factor analysis
表 2. 因子分析结果

分析方法	提取因子数	累积方差贡献
主成分分析	4	85%
因子分析	2	60%
因子分析	4	76%

主成分分析的累计方差贡献率最高(见表 2)。但在因子分析条件下提取的两个因素对影响因素的提取是明显的，体现在生活因素和气象因素两个方面(见图 8)。生活因素中的工资、城市化率等占 GDP 的比

重较大,对 PA1 的贡献也较大。气象因素中降雨量、风速、湿度的比重较大,其对 PA2 的贡献也较大。综合上述因素,考虑因子分析提取两个因子。模型结果见表 3,各模型的 AIC 值见表 4。

Table 3. Parametric results of five models

表 3. 五种模型的参数结果

参数	PA1	PA2
模型	报道病例数	
GAM	47.67653	0.06199
OLS	47.67653	0.06199
GLM	1.67491	-0.50276
ZEROINFL	1.3038	-0.49792
NBR	2.4729	-2.5294

Table 4. The AIC values of five models

表 4. 五种模型的 AIC 值

模型	GAM	GLM	NBR	OLS	ZEROINFL
AIC	287.1987	2693.1	125.91	287.1987	2560.17

提取模型参数(表 5),可以发现两个因素中生活因子所占比例更大,说明影响更强,气象因子的影响相对较弱,部分呈现负相关。此外,GLM 模式具有最高的 AIC 值(表 4)并且这两个因素对模型结果都有显著影响(表 3)。生活因素参数为正且值越大,表明生活因素对新冠肺炎的影响越大,具体表现为更高的生活质量中,新冠肺炎的确诊病例数也会更多。这是因为生活质量较高地区的人口会相对较多,导致疾病传播。气象因素参数和负值都比较小,说明新冠肺炎的气象因素也可能受到影响,但与生活因素的影响相比较小。具体来说,气象综合因素大的地区新冠肺炎确诊病例数少。

Table 5. GLM model parameter values

表 5. GLM 模型参数值

参数	Estimate	Std. Error	Z 值	Pr(> z)
PA1	1.67491	0.06535	25.629	<2e-16
PA2	-0.50276	0.05863	-8.575	<2e-16

3.3.2. 距离因素

本节以疫情爆发的中心村为新发地市场,通过计算确诊的 COVID-19 病例与新发地市场之间的距离,与确诊数量做相关性分析,研究距离对 COVID-19 的影响。以该村为对象,在 MATLAB 基础上建立函数,解决两个已知经纬度计算距离问题。

将距离与确诊人数通过 SPSS 软件做 Spearman 相关分析,结果见表 6,得出结论:距离新发地市场与问诊病例之间有明显的负相关关系。相关系数为 0.327,其相关性在 95%置信区间下是显著的。也就是说,离新发地越近,确诊病例的数量越多。这与政府在发现首例病例后及时采取封城措施有关。

Table 6. The correlation between distance and confirmed COVID-19 cases
表 6. 距离与 COVID-19 确诊病例的相关性

ρ	距新发地距离(KM)
确诊病例	-0.327*

*在 0.05 级别，相关性显著。

3.3.3. 政策因素

在本节中，我们使用多元线性回归模型来研究政府在全国范围内实施的政策对控制新冠肺炎病毒传播的影响。特别是，我们估计了四项政府政策(预测变量)，即管控、检测、隔离、封锁，对流行病指标(反应变量)报告病例数的潜在影响。对于整个国家以及每个省和每个流行病指标，我们考虑了所有可能的模型，一次采用一个、两个、三个和四个预测变量。然后，在每种情况下，我们使用 AIC 比较模型，以确定描述数据的最佳模型；AIC 值越低，描述数据的模型越好。

AIC 值见表 7。表中显示了每种策略中 AIC 值最低的模型。通过数据分析发现，这些模型的 AIC 差异并不大，这可能是由于量化政策的方法存疑。但数据显示了控制这一政策的重要性。我们观察了一次采用一项、两项或三项政策时产生的结果，发现：当采用一项措施时，控制政策下的 AIC 最小。当采取两种措施时，控制和隔离措施的 AIC 值最小。当采取三种政策时，AIC 在控制、隔离和检测措施方面是最低的。这也表明，未来的疫情应该追踪源头并加以隔离。

Table 7. The AIC value of 15 models
表 7. 15 种模型的 AIC 值

城市	报道病例数	
	政策	AIC
石家庄	C (管控)	303.6276
	T (检测)	320.5525
	I (隔离)	305.5768
	L (封锁)	315.6312
	C, T	305.121
	C, I	296.0438
	C, L	305.6003
	T, I	307.5765
	T, L	316.4279
	I, L	298.3572
	C, T, I	297.3648
	C, T, L	307.0595
	C, I, L	304.5285
	T, I, L	299.3412
	C, T, I, L	298.1512

4. 结论

本文基于 2021 年 1 月 2 日至 2 月 14 日河北省石家庄市新冠肺炎疫情监测数据和地理信息数据，运

用时空统计学方法,研究其时空传播特征,得到以下结论:1) 石家庄市新冠肺炎确诊病例主要集中在1月中下旬,到一月下旬疫情得到明显控制,高发区位于藁城区及周边的个别区(县),确诊病例中女性多于男性、中年和青年占比最多且大多集中在农村;2) 石家庄市本次疫情存在扩散现象,标准差椭圆呈现沿“西北-东南”方向扩散的分布格局;3) 疫情的热点区域位于藁城区的增村镇及其周边区域且发病热点区域没有随时间发生变化;4) 石家庄市确诊病例的聚集模式有HH和LH两种模式,高值聚类位于藁城区的增村镇、张家庄镇以及新乐市的长寿街道、邯鄲镇等区域,低值由高值围绕的异常值位于藁城区前西关镇。

运用多元回归建模等方法研究了各种影响因素对新冠肺炎的影响。得出以下结论:用KMO和巴特利特球度对生活因子和气象因子进行检验,发现它们适用于因子分析。工资水平、城市化率、GDP对生活因素影响较大,降雨量、风速、湿度对气象因素影响较大。生活因素与新冠肺炎呈正相关,而气象因素与新冠肺炎呈负相关。对因子分析结果和确诊病例进行了五次建模分析,并用AIC值评价其适用性。研究发现,GLM模型的AIC值最高,生活和气象因素都对新冠肺炎有影响,但生活因素的影响大于气象因素。距新发地的距离与确诊病例之间存在显著的负相关。政策因素对新冠肺炎的发生有一定的影响,控制作用较大。

5. 讨论

防止COVID-19的传播是目前所有国家的共同目标。石家庄市及时有效的防控措施使COVID-19得到了控制,在2021年3月初的大规模集体返岗返校活动中,疫情没有再次出现,说明石家庄市政府的防控措施和策略相当有效。

根据分析结果,本文提出以下建议:疫情的传播具有聚集性,可以根据不同的传播方式采取针对性的措施。例如,在区域层面,石家庄东北部是输入性传播的防控重点,西北部是本地传播的防控重点(图2)。在内容层面,城市间传播、街道间互动是当地疫情防控的重点(图6)。应减少亲友间的探访,注意在街上人与人交往密集的地区进行疫情检测,限制人与人之间的接触频率,避免病毒因人口流动而迅速传播。外来输入的防控措施要以外来输入为主,适当关注市内传播,筛查社区外人员,做好消毒工作,提醒公众注意自身防护。提出几点防控措施:从性别和年龄角度,重点关注高流动性返乡青年学生和外出打工的中青年劳动力的活动轨迹,重点识别此类群体(图3)。应关注儿童、老人等特殊人群,在老年人中普及新冠肺炎知识,提高他们的防控意识。从时间上来说,在疫情的不同阶段要采取不同的措施。同时,在国内疫情得到控制的同时(图4),要严防输入性病例。在空间上,要因地制宜制定政策。对于石家庄市疫情高发的藁城区,需要加强区域经济发展,改善医疗物资配置,以应对未来类似不可控的挑战。同时,无论是城市还是农村,都需要注意环境卫生,定期做好消毒工作,创造一个健康舒适的生活、工作和学习环境。

我们承认我们的研究有几个局限性。本研究中使用的数据是有限的,并带有一些不确定性。特别是,COVID-19病例只包括那些被检测并确认为阳性的病例。然而,一些研究记录了无症状或未诊断的COVID-19感染者,他们可以传播病毒,尽管他们的感染性仍不确定。因此,我们的估计可能低估了指标的价值。在我们的建模中,政策的影响被假定为是线性的。然而,传染病的传播机制,特别是在有多种政策的情况下,可能是相当非线性的。

基金项目

国家自然科学基金青年项目(11801398, 12101443); 山西省应用基础研究面上青年项目(201801D221024), 2021年度留学人员科技活动择优资助项目(20210009)。

参考文献

- [1] WHO Director-General's Statement on IHR Emergency Committee on Novel Coronavirus (2019-nCoV). [https://www.who.int/director-general/speeches/detail/who-director-general-s-statement-on-ihf-emergency-committee-on-novel-coronavirus-\(2019-ncov\)](https://www.who.int/director-general/speeches/detail/who-director-general-s-statement-on-ihf-emergency-committee-on-novel-coronavirus-(2019-ncov))
- [2] Gorbalenya, A.E., Baker, S.C., Baric, R.S., *et al.* (2020) The Species Severe Acute Respiratory Syndrome-Related Coronavirus: Classifying 2019-nCoV and Naming It SARS-CoV-2. *Nature Microbiology*, **5**, 536-544. <https://doi.org/10.1038/s41564-020-0695-z>
- [3] National Health Commission of the People's Republic of China (2022) Update on the Novel Coronavirus Pneumonia Situation in China as of 24:00 16 March. <http://www.nhc.gov.cn/xcs/yqtb/202203/c6fd4e5d3e084d7592d57d68ffecf9a.shtml>
- [4] Johns Hopkins University. The Johns Hopkins COVID-19 Map Dashboard. <https://coronavirus.jhu.edu/map.html>
- [5] Kang, D., Choi, H., Kim, J.-H., *et al.* (2020) Spatial Epidemic Dynamics of the COVID-19 Outbreak in China. *International Journal of Infectious Diseases*, **94**, 96-102. <https://doi.org/10.1016/j.ijid.2020.03.076>
- [6] Mo, C., Tan, D., Mai, T., *et al.* (2020) An Analysis of Spatiotemporal Pattern for COVID-19 in China Based on Space-Time Cube. *Journal of Medical Virology*, **92**, 1587-1595. <https://doi.org/10.1002/jmv.25834>
- [7] Xiong, Y., Wang, Y., Chen, F., *et al.* (2020) Spatial Statistics and Influencing Factors of the COVID-19 Epidemic at both Prefecture and County Levels in Hubei Province, China. *International Journal of Environmental Research and Public Health*, **17**, 3903. <https://doi.org/10.3390/ijerph17113903>
- [8] Desjardins, M.R., Hohl, A. and Delmelle, E.M. (2020) Rapid Surveillance of COVID-19 in the United States Using a Prospective Space-Time Scan Statistic: Detecting and Evaluating Emerging Clusters. *Applied Geography*, **118**, Article ID: 102202. <https://doi.org/10.1016/j.apgeog.2020.102202>
- [9] 孙建平, 胡文海, 陈若尘, 等. 安徽省新型冠状病毒肺炎疫情时空格局及集聚模式分析[J]. 安徽师范大学学报(自然科学版), 2020, 43(6): 573-579.
- [10] Xin, L., Lin, Z., Tao, J., *et al.* (2020) Influence of Urban Factors on the COVID-19 Epidemic: A Case Study of Wuhan City. *Geomatics and Information Science of Wuhan University*, **368**, 395-400.
- [11] Kraemer, M.U.G., Yang, C.-H., Gutierrez, B., *et al.* (2020) The Effect of Human Mobility and Control Measures on the COVID-19 Epidemic in China. *Science*, **368**, 493-497. <https://doi.org/10.1126/science.abb4218>
- [12] Jia, J.S., Lu, X., Yuan, Y., *et al.* (2020) Population Flow Drives Spatio-Temporal Distribution of COVID-19 in China. *Nature*, **582**, 389-394. <https://doi.org/10.1038/s41586-020-2284-y>
- [13] 王姣娥, 杜德林, 魏冶, 等. 新冠肺炎疫情的空间扩散过程与模式研究[J]. 地理研究, 2020, 39(7): 1450-1462.
- [14] 汪冉, 张明鑫, 李浩. 河南省交通通达水平对新型冠状病毒传播的影响[J]. 河南理工大学学报(自然科学版), 2020, 39(6): 68-77.
- [15] Mollalo, A., Vahedi, B. and Rivera, K.M. (2020) GIS-Based Spatial Modeling of COVID-19 Incidence Rate in the Continental United States. *Science of the Total Environment*, **728**, Article ID: 138884. <https://doi.org/10.1016/j.scitotenv.2020.138884>
- [16] Ujiie, M., Tsuzuki, S. and Ohmagari, N. (2020) Effect of Temperature on the Infectivity of COVID-19. *International Journal of Infectious Diseases*, **95**, 301-303. <https://doi.org/10.1016/j.ijid.2020.04.068>
- [17] Tosepu, R., Gunawan, J., Effendy, D.S., *et al.* (2020) Correlation between Weather and Covid-19 Pandemic in Jakarta, Indonesia. *Science of the Total Environment*, **725**, Article ID: 138436. <https://doi.org/10.1016/j.scitotenv.2020.138436>
- [18] Xie, J. and Zhu, Y. (2020) Association between Ambient Temperature and COVID-19 Infection in 122 Cities from China. *Science of the Total Environment*, **724**, Article ID: 138201. <https://doi.org/10.1016/j.scitotenv.2020.138201>
- [19] Wu, Y., Jing, W., Liu, J., *et al.* (2020) Effects of Temperature and Humidity on the Daily New Cases and New Deaths of COVID-19 in 166 Countries. *Science of the Total Environment*, **729**, Article ID: 139051. <https://doi.org/10.1016/j.scitotenv.2020.139051>
- [20] 王青芳, 靳祯, 张文义. 人感染 H7N9 型禽流感空间分布研究[J]. 公共卫生与预防医学, 2017, 28(6): 16-19.
- [21] Fischer, M.M. (2010) Handbook of Applied Spatial Analysis. Springer, Berlin, 280-286. <https://doi.org/10.1007/978-3-642-03647-7>
- [22] Getis, A. and Ord, J.K. (1992) The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, **24**, 189-206. <https://doi.org/10.1111/j.1538-4632.1992.tb00261.x>
- [23] Saini, A. (2013) Spatial Modelling of Cardiovascular Mortality Patterns: A GIS Based Study of Rural Bist Doab Region (Punjab, India). *IOSR Journal of Humanities and Social Science*, **14**, 25-30. <https://doi.org/10.9790/0837-1412530>
- [24] 王青芳, 靳祯, 王仲兵, 等. 山西省 2005-2012 年人间布鲁菌病空间分布特征研究[J]. 中国预防医学杂志, 2017,

18(12): 910-915.

- [25] Arijit, D., Sasanka, G., Kalikinkar, D., *et al.* (2021) Living Environment Matters: Unravelling the Spatial Clustering of COVID-19 Hotspots in Kolkata Megacity, India. *Sustainable Cities and Society*, **65**, Article ID: 102577. <https://doi.org/10.1016/j.scs.2020.102577>
- [26] Hsiang, S., Allen, D., Annan-Phan, S., *et al.* (2020) The Effect of Large-Scale Anti-Contagion Policies on the Covid-19 Pandemic. *Nature*, **178**, 317-330. <https://doi.org/10.1101/2020.03.22.20040642>
- [27] Ord, K. and Getis, A. (2010) Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*, **27**, 286-306. <https://doi.org/10.1111/j.1538-4632.1995.tb00912.x>
- [28] Anselin, L. (1995) Local Indicators of Spatial Association-LISA. *Geographical Analysis*, **27**, 93-115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>