

11-17-2020

Modeling and Relevance Analysis of Urban Epidemic Transmission and Work Resumption Intensity

Gan Mi

*1.School of Transportation and Logistics,Southwest Jiaotong University,Chengdu 610031,China;
;2.National Engineering Laboratory of Integrated Transportation Big Data Application
Technology,Southwest Jiaotong University,Chengdu 610031,China; ;3.National United laboratory of
Comprehensive Transportation,Southwest Jiaotong University,Chengdu 610031,China;*

Yunyi Tian

*1.School of Transportation and Logistics,Southwest Jiaotong University,Chengdu 610031,China;
;2.National Engineering Laboratory of Integrated Transportation Big Data Application
Technology,Southwest Jiaotong University,Chengdu 610031,China; ;*

Wenchang Zhang

*1.School of Transportation and Logistics,Southwest Jiaotong University,Chengdu 610031,China;
;2.National Engineering Laboratory of Integrated Transportation Big Data Application
Technology,Southwest Jiaotong University,Chengdu 610031,China; ;*

Xihan Zhao

*1.School of Transportation and Logistics,Southwest Jiaotong University,Chengdu 610031,China;
;2.National Engineering Laboratory of Integrated Transportation Big Data Application
Technology,Southwest Jiaotong University,Chengdu 610031,China; ;*

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Modeling and Relevance Analysis of Urban Epidemic Transmission and Work Resumption Intensity

Abstract

Abstract: On the basis of multi-source big data, the model for analyzing the population migration changes and manpower gaps caused by the COVID-19 epidemic in 34 typical cities across the country is constructed, and the work resumption intensity of other cities is predicted by using the migration base constructed. The SEIR model is used to estimate the basic reproduction number in each city since the simulate results show that it can emulate the transmission trend of this epidemic accurately, and the retrospective matrix analysis of the work resumption intensity is combined with the manpower gap to summarize the anti-epidemic experience of China. In the stage of relevance analysis, regression analysis is performed on the city clusters under K-means unsupervised clustering. The results show that for most cities, the work resumption intensity significantly correlated with the manpower gap, basic reproduction number, and GDP per capita.

Keywords

COVID-19, epidemic transmission, work resumption intensity, basic reproduction number, modeling, relevance analysis

Recommended Citation

Gan Mi, Tian Yunyi, Zhang Wenchang, Zhao Xihan. Modeling and Relevance Analysis of Urban Epidemic Transmission and Work Resumption Intensity[J]. Journal of System Simulation, 2020, 32(11): 2258-2271.

城市疫情传播和复工强度建模及相关性分析

甘蜜^{1,2,3}, 田昀翊^{1,2}, 张文畅^{1,2}, 赵夕涵^{1,2}

(1. 西南交通大学交通运输与物流学院, 四川 成都 610031; 2. 西南交通大学综合交通大数据应用技术国家工程实验室, 四川 成都 610031; 3. 西南交通大学综合交通国家地方联合实验室, 四川 成都 610031)

摘要: 以多源大数据为基础构建模型, 分析全国 34 个典型城市因疫情导致的人口迁徙变化和人力缺口, 并通过迁徙基数推算其他各城市的复工强度。仿真结果证明 SEIR 仓室模型能够较好地模拟此次疫情发展趋势, 利用其估计各城市内部新型冠状病毒感染肺炎的基本再生数, 结合人力缺口对复工强度进行回顾性的矩阵分析, 以总结我国此次抗疫经验。相关性分析阶段对 K-means 无监督聚类后的城市集群进行回归分析, 结果表明对于大部分城市而言, 复工强度的大小与其人力缺口、基本再生数以及人均 GDP 水平有明显的相关关系。

关键词: 新冠肺炎; 疫情传播; 复工强度; 基本再生数; 建模; 相关性分析

中图分类号: TP391.9 文献标识码: A 文章编号: 1004-731X (2020) 11-2258-13

DOI: 10.16182/j.issn1004731x.joss.20-0305

Modeling and Relevance Analysis of Urban Epidemic Transmission and Work Resumption Intensity

Gan Mi^{1,2,3}, Tian Yunyi^{1,2}, Zhang Wenchang^{1,2}, Zhao Xihan^{1,2}

(1.School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China;

2.National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Southwest Jiaotong University, Chengdu 610031, China; 3.National United laboratory of Comprehensive Transportation, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: On the basis of multi-source big data, the model for analyzing the population migration changes and manpower gaps caused by the COVID-19 epidemic in 34 typical cities across the country is constructed, and the work resumption intensity of other cities is predicted by using the migration base constructed. The SEIR model is used to estimate the basic reproduction number in each city since the simulate results show that it can emulate the transmission trend of this epidemic accurately, and the retrospective matrix analysis of the work resumption intensity is combined with the manpower gap to summarize the anti-epidemic experience of China. In the stage of relevance analysis, regression analysis is performed on the city clusters under K-means unsupervised clustering. The results show that for most cities, the work resumption intensity significantly correlated with the manpower gap, basic reproduction number, and GDP per capita.

Keywords: COVID-19; epidemic transmission; work resumption intensity; basic reproduction number; modeling; relevance analysis

引言

2019 年 12 月底爆发的新型冠状病毒肺炎



收稿日期: 2020-06-04 修回日期: 2020-07-14;
基金项目: 科技部国家重点研发计划(2018YFB1601402);
作者简介: 甘蜜(1984-), 女, 湖南, 博士, 副教授,
研究方向为智慧物流、数据驱动的复杂系统; 田昀翊
(1997-), 女, 四川, 硕士生, 研究方向为智慧物流大
数据分析; 张文畅(1994-), 女, 四川, 硕士生, 研究
方向为智慧物流大数据分析。

(COVID-19, 简称“新冠肺炎”), 由于其在全人群易感, 具有很高的传染性和致病率, 在 2020-03-11 被 WHO 列为全球性大流行病, 其严重程度超过了 2003 年的重症急性呼吸系统综合征 (SARS)^[1-3]。面对此次重大突发公共卫生事件, 在政府的有力调控和全社会各界的积极协作下, 我国疫情防控工作已经取得重大战略成果, 自

2020-02-18 起, 全国新增治愈病例超过了新增确诊病例, 累计治愈出院病例超过了现有确诊病例, 且绝大部分地区已经完成复工复产任务, 投入了正常的经济生产活动。但境外多个国家现仍处于疫情较高峰, 截止 2020-06-04, 全球新冠肺炎确诊病例累计已超 628 万例, 累计死亡病例超 37 万例, 给全球经济带来巨大冲击, 待疫情逐步缓和, 复工复产将成为这些国家的主要任务。

国内有关新冠肺炎疫情传播的时空特征及社会影响的研究主要从省级和市级角度展开。省级层面, 大部分研究集中于刻画某个省份新冠肺炎疫情的时空分布格局与演变过程^[4]、多个维度的时空扩散过程^[5-6]、聚集性疫情的流行特征^[7-8]等。市级层面, 也有学者结合多种方法探究城市新型冠状病毒肺炎疫情的演变趋势、基本特征和空间分布^[9-10], 或者分析城市特大疫情防控机制^[11]等。在分析、预测疫情走向方面, 目前最为常见的是从流行病学角度出发, 通过建立 SEIR(Susceptible Exposed Infected Recovered)仓室模型^[12]计算基本再生数, 模拟此类流行性病毒的传播趋势, 进而为疫情的防控和政策制定提供理论支持, 其中大部分研究得出该值在 1~3 之间^[13-16]。

这些研究都基于病例数据时间序列展开, 部分学者通过搜集疫情期间全国范围内多个城市的人口迁徙和变化大数据, 分别从时、空维度构建模型, 比较分析各个城市的疫情传播状况, 表明通过百度迁徙出行大数据和疫情实时监控数据构建模型, 结合时空分析、回归分析等, 是客观反映疫情防控效果和获得地区活动指标的重要手段^[17]。然而, 已有研究主要是对病毒传播与人口移动的关系进行建模和分析, 尚没有对我国不同城市复工情况进行建模, 并分析病毒传播与复工情况之间相关关系的研究。复工方面, 仅有少量文献资料通过对电力数据建立指数模型, 评估某个地区的复产情况^[18], 更多的研究则局限于讨论如何通过政府宏观调控策略^[19]、地方局部刺激政策^[20]等促进企业复工复产。尚缺乏利用城市人员真实出行活动数据来刻画

和分析城市复工强度, 并对疫情期间影响不同地区复工强度的关键因素进行分析讨论的相关研究。

事实上, 通过系统的建模和刻画我国在疫情期间的复工情况, 能有效的为其它国家和地区提供相关经验。鉴于此, 本文搜集 2020 年 2 月 1 日-2 月 18 日春运期间, 也是国内新冠疫情防控形势最为严峻期间, 全国 34 个典型城市的春运数据、百度迁徙出行数据, 常住人口数据等, 尝试通过构建一系列数学模型, 定量描述该段时间内城市的疫情传播状况和复工强度, 并且在检验了 SEIR 仓室模型的仿真效果的基础上, 利用其计算各城市的基本再生数, 随后进行矩阵分析、K-means 无监督聚类下的回归分析, 为目前尚处于疫情重灾区的国家和地区提供抗疫经验和复工建议。

1 全国春运期间人口流动情况

据交通部发布的春运迁徙数据^[21], 2020 年全国春运(从 1 月 10 日-2 月 18 日, 共计 40 天), 铁路、道路、水路、民航共累计发送旅客 14.76 亿人次, 比去年同期下降 50.3%。其中, 铁路发送旅客 2.10 亿人次, 下降 47.3%; 道路发送旅客 12.11 亿人次, 下降 50.8%; 水路发送旅客 1689.1 万人次, 下降 58.6%; 民航发送旅客 3839.0 万人次, 下降 47.5%。以铁路运输为例, 节前(1 月 10 日-1 月 24 日)共发送旅客量 1.68 亿人次, 节后(1 月 25 日-2 月 18 日)仅发送旅客量 4248 万人次, 与 2019 年春运同期数据相比, 减少了 2.2 亿人次。

可见, 疫情防控措施的升级使得各城市大批的航班、铁路、水路班次被取消, 对大部分城市而言, 今年的春运展现了 2 个明显特征:(1) 受疫情影响, 旅客发送量下降, 返城高峰客流也大幅缩减;(2) 返城的时间线延长数倍, 前后巨大的流量差也揭露了疫情初期, 全国绝大部分企业的复工复产都存在一定程度的人力空缺。

1.1 国内典型城市迁入迁出变化趋势

1.1.1 数据来源

百度自 2014 年春运期间推出的百度迁徙服务

系统，可以利用百度地图 LBS(Location-Based Service)平台跟踪、记录和整合海量第三方 APP 手机终端用户位置的时空变化轨迹，最终实现以可视化的手段、实时地呈现国内各个城市的逐日人口流动方向和强度。本研究通过百度迁徙平台^[22]挖掘了在 2019 年和 2020 年春节前后，以北京、郑州、昆明、武汉等全国 34 个典型城市之间 840 914 939 人次的出行 OD(Origin and Destination nodes)；迁出、迁入百分比数据和城市内部出行活跃程度数据。为比较城市现有人口和常住人口之间的比值，以确定城市劳动力缺口情况，本文的数据集还包括中华人民共和国国家统计局^[23]发布的 2018 年末 34 个典型城市的常住人口、国内生产总值(GDP)等统计数据。本文使用上述数据推算各个城市该时间段内的迁出人口，及其现在人数与常住人口的差距。

1.1.2 2019 年春运期间各典型城市人口迁移情况

由图 1 可知，在常态情况(2019 年度)下，深圳、上海、温州市属于典型的迁出人口远多于迁入人口的城市，即以迁出人口为主型城市；乌鲁木齐、呼

和浩特、西宁、福州、成都、拉萨市人口流动较为均衡，迁出与迁入人口数量差较小；迁入人口比高的城市表明是春运期间人们旅游或返乡的主要目的地，抑或是人们旅游、返乡、出国的主要中转地，其中郑州市人口迁入量居全国第一，验证了河南劳动力输出大省的地位。

1.1.3 2020 年春运期间各典型城市人口迁移情况

如图 2 所示，在新冠肺炎疫情影响下，2020 年春节期间，各典型城市的人口迁移情况产生了巨大的变化。除重庆市仍属于以迁入人口为主的的城市外，图 1 中其它城市几乎都转化为以迁出人口为主的的城市；而少数迁入-迁出比低于中位数的城市，如哈尔滨、沈阳、南宁、兰州、香港等城市人口变化较去年更为平缓，成为了相对平衡型城市。一线城市、区域中心城市及劳动力密集城市等展示出其强大的人口吸纳能力。前后结果对比表明，2020 年春运期间全国范围内的迁出人口量剧增，大量异地务工人员、学生群体未能在春节后返回到务工或学习的城市。

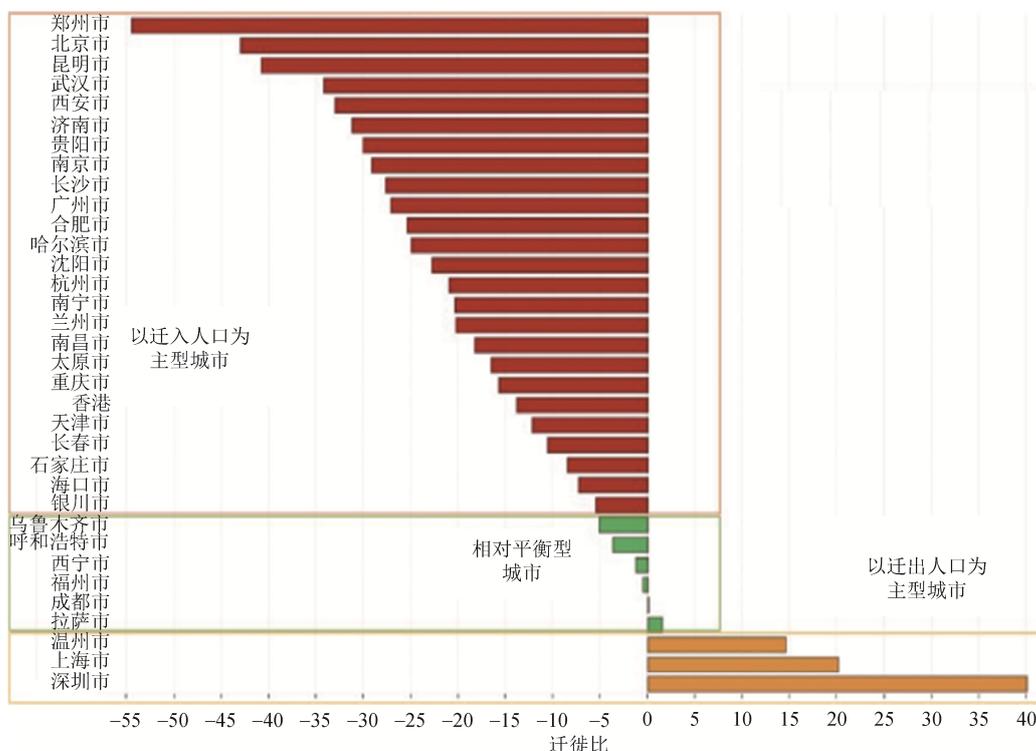


图 1 2019 年春运期间国内典型城市人口迁移结构
Fig. 1 Migration structure of typical cities in China during Spring Festival in 2019

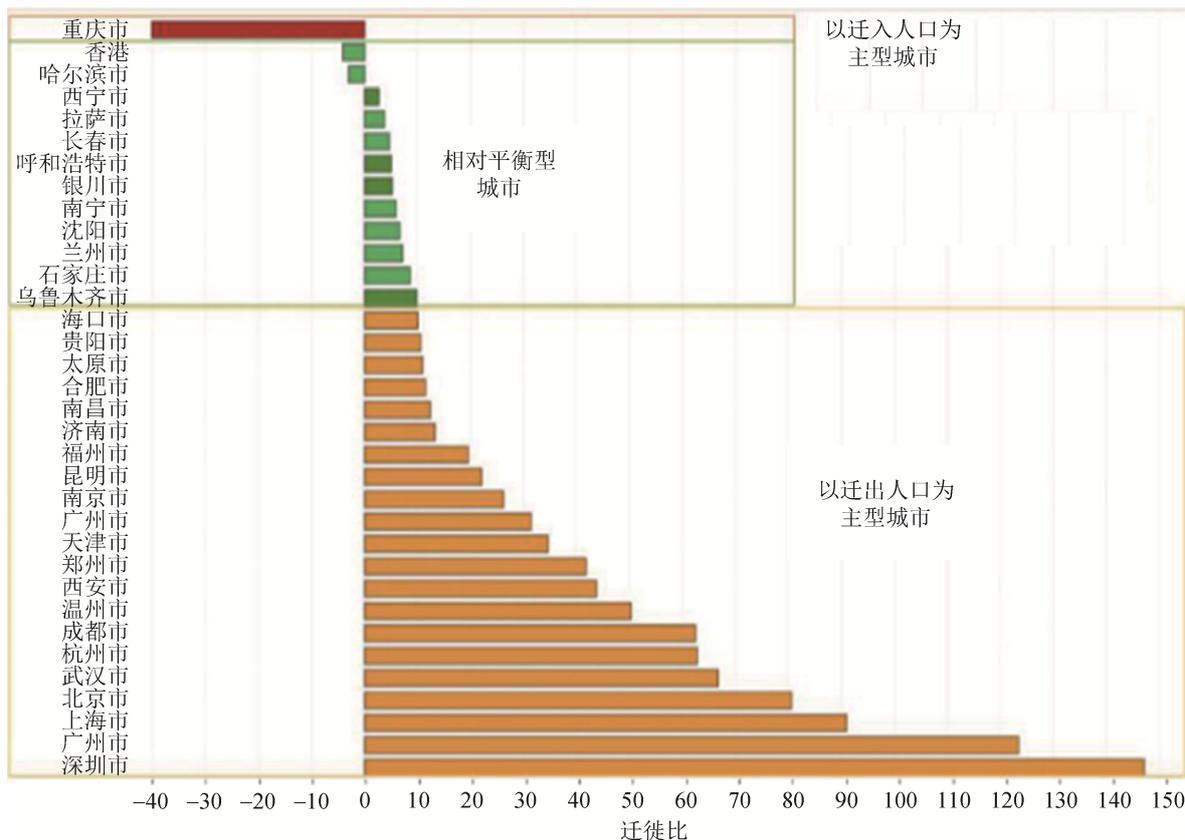


图 2 2020 年春运期间国内典型城市人口迁移结构

Fig. 2 Migration structure of typical cities in China during Spring Festival in 2020

1.2 2020 年国内典型城市回流规模情况

复工潮来临后, 城市人口开始回流, 将逐步恢复到 2019 年常态下的情况。按照图 3 所示城市回流规模(人口迁出比与迁入比的差值)大小排名, 广州、北京、深圳、武汉、郑州、杭州、西安、上海、昆明、成都、长沙、南京等城市的返工客流较大。

由图 4 可知, 春运后国内典型城市逐步进入了回流高峰(图中每个方块代表某城市某天的客流回流规模, 方块颜色越深代表对应典型城市当日的回流客流越大, 反之则越小), 尤其是以广州市, 深圳市, 上海市, 北京市、成都市、长沙市和昆明市为甚, 2 月 1 日-2 月 3 日、2 月 9 日-2 月 10 日、2 月 16 日-2 月 17 日出现返程高峰期是因为这 3 个时间段为企业正常复工时间和 2 次延期复工时间。

以迁出基数式(1)所示为桥梁, 再结合最新的

常住人口普查数据, 可以推测出其他各典型城市在相应时间段内的现有人数比例式(2)~(3), 以及相应的迁出人数, 即人力缺口。

$$P = B \times M \quad (1)$$

$$L = N - P \quad (2)$$

$$R = (N - P) / N \quad (3)$$

式中: B 为百度迁徙统计的总的迁徙人数; M 为各城市迁出比; P 为各城市具体的迁出人口数; N 为各城市的常住人口数据; L 为各市的人力缺口。由式(3)给出了各城市春运期间的现有人数与常住人口之比 R 的计算方法, 其分布如图 5 所示, R 越高代表城市人力缺口少, 否则缺口大。其中, 广州、深圳、杭州、温州等城市的春节空城指数较高, 与其人口迁徙情况和回流规模相对应。

城市	客流回流规模	城市	客流回流规模
广州市	-150.0	沈阳市	-29.4
北京市	-123.5	太原市	-27.4
深圳市	-106.9	兰州市	-27.3
武汉市	-100.9	南宁市	-26.2
郑州市	-96.1	哈尔滨市	-21.8
杭州市	-83.6	福州市	-20.0
西安市	-76.7	海口市	-17.3
上海市	-70.8	石家庄市	-17.0
昆明市	-62.9	长春市	-15.2
成都市	-62.4	乌鲁木齐市	-14.8
长沙市	-58.9	银川市	-10.5
南京市	-55.2	银川市	-10.5
天津市	-46.7	香港	-9.5
济南市	-44.4	呼和浩特市	-8.7
贵阳市	-41.4	西宁市	-3.9
合肥市	-36.9	拉萨市	-2.2
温州市	-35.5	重庆市	24.5
南昌市	-30.6		

图 3 2020 年春运后国内典型城市回流规模

Fig. 3 Backflow scale of typical cities in China after Spring Festival in 2020

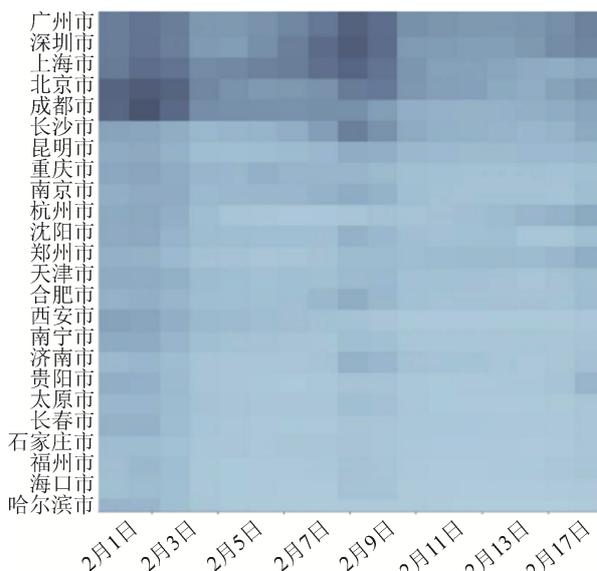


图 4 2020 年 2 月 1 日-2 月 18 日各典型城市回流客流监控数据

Fig. 4 Monitoring data of passenger backflow in typical cities from February 1st to February 18th, 2020

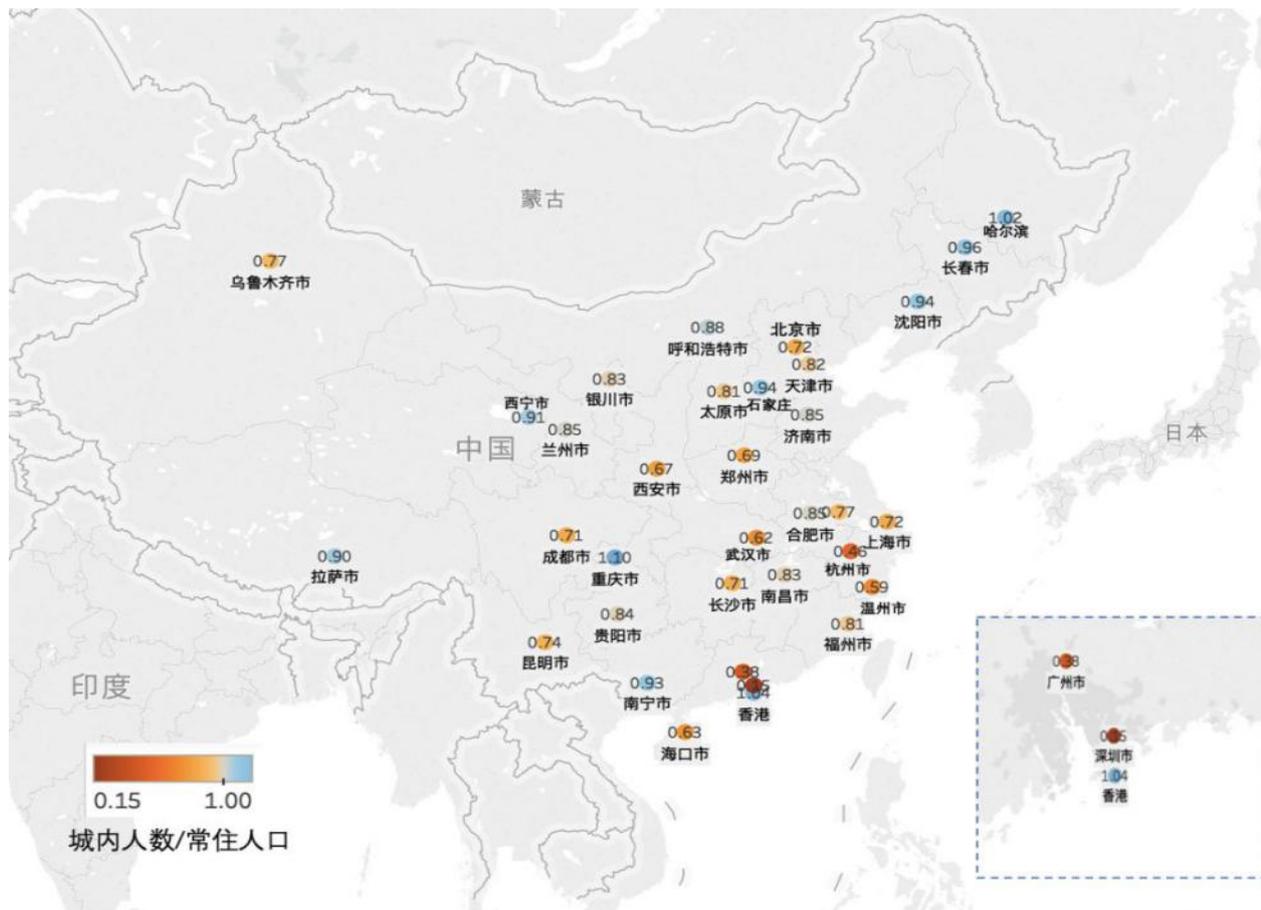


图 5 2020 年春运期间典型城市现有人数比例分布

Fig. 5 Proportional distribution of remaining population in typical cities during Spring Festival in 2020

2 基于出行大数据的城市复工强度估计

以不同于城市用电电力数据为视角, 本文通过城市内有出行的人数与该城市居住人口比值的指数化结果, 即出行强度作为判断依据, 构建复工强度估计的方法。如前所述, 通过在百度迁徙上挖掘 34 个城市自 2019-01-01 至 2019-22-18, 与 2020-01-01 至 2020-02-18 共计 3 332 条城市内部出行强度指数, 数据字段包括: 时间、城市、出行强度。基于出行强度的计算逻辑, 进一步构建复工强度计算模型, 评估 34 个典型城市在春运期间的复工状况。

2.1 疫情期间假期出行强度变化

根据 2020 年和 2019 年无复工行为的假期阶段出行强度的变化来评价疫情状态非复工情况下全城人员的出行频率的变化, 该变化可以利用 2020 年春节假期期间各城市的出行频次与 2019 年春节假期期间各城市的出行频次的比值式来求得, 如

(4)~(5)所示。该比值记作 K , 体现了与 2019 年春节假期相比, 不同城市的个体在 2020 年春节假期出行频次的降低比例。

$$F = I \times N \quad (4)$$

$$K = F_t / F_r \quad (5)$$

式中: I 为各城内出行强度; F 为各城内出行频次; N 为各城市常住人口数据; F_t 和 F_r 分别为 2020 年与 2019 年春运期间各城市的出行频次。

由图 6 可知, 受当地疫情和防控措施的双重影响, 春节期间各个城市的个人出行频次下降比例呈现出了不同程度的变化。以武汉、深圳、广州、杭州、温州、西安、海口、长沙和成都为代表的典型城市春节空城率高, 个体出行频次降低了 50%~88%; 南宁、天津、上海、乌鲁木齐、南昌等城市的个人出行频次降低了 40%~50%; 拉萨、长春、沈阳、贵阳、呼和浩特、银川、西宁、香港等地区的个体出行频次降低率不足 40%; 重庆和哈尔滨作为年前年后单向迁入城市, 人口流动量大, 个体出行频次变化较小。

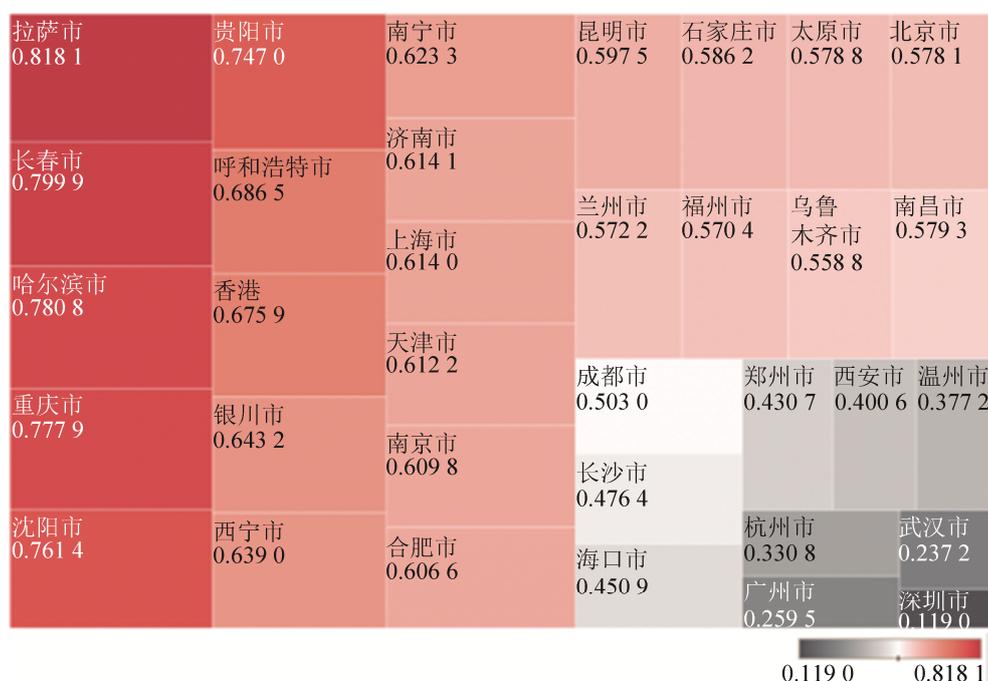


图 6 2020 年各典型城市 k 值分布图

Fig. 6 Distribution graph of K value of each typical city in 2020

2.2 全国城市复工强度分布

通过 2.1 节的分析, 已经获得了在疫情期间无复工行为下, 2020 年相比 2019 年出行频次的下降程度 K , K 也体现了在疫情状态下, 城市居民非复工出行行为的降低情况, 因此, 要准确刻画城市的复工强度, 一方面需要比较复工后与复工前城市出行强度的变化; 另一方面, 也需要考虑疫情对各城市非复工的普通出行频次降低的影响, 因此需要利用 K 值来修正城市正常状态下的出行强度。将春运前(2020-01-10 前)城市内部的个体出行频次, 记作 $F-normal$, $F-normal$ 一定程度上反应了一个城市在非疫情正常工作情况下城市出行的频次, 通过利用 K 值相乘修正后, 可仿真在有疫情情况下城市复工后正常的出行频次; 令 2020-02-10 日至 2020-02-18 期间的城内个体出行频次, 记作

$F-covid$, 后者可以反映疫情情况下逐步复工状态下城内个体出行的真实情况。通过对二者进行比较, 可实现在一定程度上判断和反映城市在今年春运期间的整体复工强度 Q , 如式(6)所示:

$$Q = F - covid / (K \times F - normal) \quad (6)$$

通过计算获得全国典型城市复工强度排名情况和复工强度分布图, 如图 7 和图 8 所示。

图 7 显示, 2 月 18 日时, 全国典型城市复工强度的均值已经恢复到 45.6%。其中, 西南和西北地区的复工强度普遍较高, 西宁、兰州、重庆、昆明、成都、银川等城市复工率均超过 50%; 东北地区除受疫情影响较大的哈尔滨外, 沈阳、长春等城市复工率较高, 超过全国复工均值; 华北地区的太原、天津、合肥、呼和浩特等城市复工强度也超过了 40%; 位于长三角地带的南京、杭州、合肥等城市的复工强度则在 35%~40% 之间。

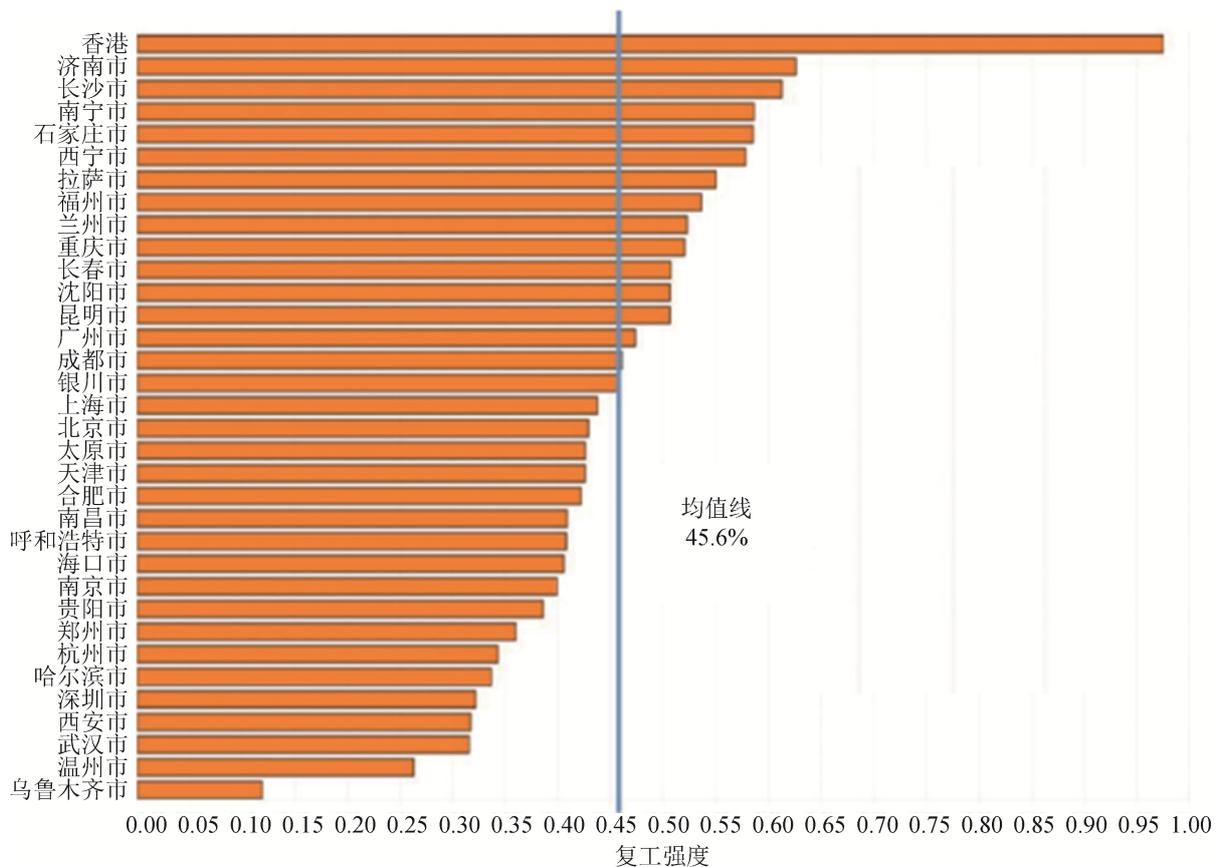


图 7 全国典型城市复工强度排名情况

Fig. 7 Ranking of resumption intensity in typical cities



图 8 各典型城市复工强度分布图

Fig. 8 Distribution graph of resumption intensity in each typical city

图 8 所示为全国典型城市复工强度分布, 图中每个城市对应的标记颜色深浅表示其复工强度的高低。可以发现, 图 8 中标记颜色越深的城市, 在图 5 中对应的标记颜色越浅, 即该城市的现有人数比例越高, 说明所得的复工强度指数与根据实际人口统计数据推算出的城内人数之间具有明显的正相关关系, 导致企业复工难的主要原因就是人力资源不足, 表明构建的复工强度计算模型能够较好地刻画城市的复工现状, 从而为地方狙击疫情提供科学的研判依据。

3 基于矩阵分析的我国复工情况回顾性分析

全国的疫情防控取得显著成效, 统计数据显示, 目前全国近 20 个省份的现有确诊人数已清零, 超过 90% 的地区已连续多日无新增确诊病例, 西藏、新疆和青海等省份的零新增确诊病例记录甚至已经过百日, 绝大部分城市的复工进程也已基本进入尾声, 各行业的贸易往来活动恢复正常, 表明我国对复工复产的整体布控是有效且成功的。因此, 利用 SEIR 病毒传播仿真模型估算 2020 年春运期间各典型城市内新型冠状病毒感染肺炎的基本再生数, 结合前文所得的人力缺口和复工强度, 通过矩阵分析的方法, 探讨疫情传播风险下的复工策略, 并提出相应的防控建议, 为将来各地管理部门处理类似情况提供有益借鉴。

3.1 城市疫情传播仿真模型

基本再生数(basic reproduction number, R_0)是指在一个全是易感染态个体构成的群体中, 一个感染态的个体在恢复之前平均能感染的人数, 若 $R_0 > 1$, 表明疾病将流行, 若 $R_0 < 1$, 则表明疾病将逐渐被消灭, 故 R_0 是反映传染病传播特征的重要指标之一。通过计算基本再生数, 可以真实地仿真和刻画对病毒在某区域内的传播能力和趋势。

3.1.1 SEIR 仓室仿真模型的建立

本文将新冠肺炎早期的自由传播用“易感染-潜伏态-感染态-移除态”的 SEIR 仓室模型^[21]描述, 则“被感染后处于潜伏期的人群”在单位时间内转变为“潜伏期之后已具有感染能力的人群”的概率为 γ_1 , 随后在单位时间内转变为“各类不对流行病传播动力学产生影响的人群”的概率为 γ_2 , 基本再生数可表示为:

$$R_0 = 1 + \lambda T_g + \rho(1 - \rho)(\lambda T_g)^2 \quad (7)$$

式中, $\lambda = \ln Y(t)t$ 为早期确诊患者在指数增长时的增长率; $Y(t)$ 为截至到 t 时刻有症状的感染人数; $T_E = 1/\gamma_1$ 为潜伏期, 感染期时长可表示为和 $T_I = 1/\gamma_2$ 为感染期, 则生成时间 $T_g = T_E + T_I$; $\rho = T_E / T_g$ 为潜伏期占生成时间的比例。

3.1.2 模型仿真结果

为验证模型对病毒传播仿真的适用性, 以武汉市为例, 通过 Python 建模仿真 SEIR 模型, 如图 9 所示, 可得到易感人数、暴露人数、感染人数和康复人数随病毒第一天爆发后的分布图, 通过 SEIR 模型模拟, 可知累计确诊病例数在 84 天(即 2020-03-01)左右达到峰值, 约为 48 778 例。在疫情爆发约 125 天后(即 2020 年 4 月初), 日新增病例接近 0, 逐步接近尾声。由上述仿真结论可知, 该模型仿真结果较为接近真实情况, 能较好的模拟病毒传播情况。

依据各城市卫生健康委员会公布的全国疫情初期的新增和累计的疑似、确诊、死亡和出院病例等数据, 并借鉴 SARS 的流行病学关键参数^[24],

分别计算各市内部的新冠病毒潜伏期 T_E 和感染期 T_I , 进而计算出 T_g 的值, 并代入公式得到相应的 ρ 值。利用上述数据和 SEIR 仿真模型, 计算得到各城市 2 月 18 日的新冠病毒基本再生数, 如图 10 所示仿真值, 从而实现基于 SEIR 模型对当时各城市疫情传播情况进行回溯。

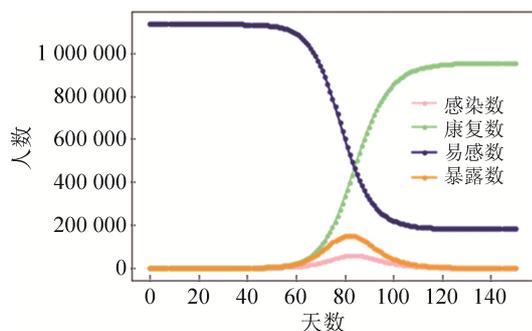


图 9 基于 SEIR 模型的武汉市新型冠状病毒肺炎传播趋势仿真结果

Fig. 9 Simulation results of COVID-19 epidemic transmission trend in Wuhan based on SEIR model

城市	基本再生数	城市	基本再生数
武汉市	2.477	郑州市	1.542
重庆市	1.749	杭州市	1.537
温州市	1.735	南京市	1.508
北京市	1.710	福州市	1.464
深圳市	1.707	昆明市	1.454
广州市	1.679	南宁市	1.444
上海市	1.646	济南市	1.433
哈尔滨市	1.637	青岛市	1.416
长沙市	1.632	贵阳市	1.363
南昌市	1.623	海口市	1.350
合肥市	1.604	石家庄市	1.342
成都市	1.554	太原市	1.308
西安市	1.551	天津市	1.231

图 10 基于 SEIR 模型的 2 月 18 日各城市新冠病毒基本再生数仿真结果

Fig. 10 Simulation results of basic reproduction number of COVID-19 in each city on February 18 based on SEIR model

3.2 基本再生数与复工强度

从仿真结果可以看出, 从 2 月 2 日-2 月 18 日, 全国各地的新冠病毒基本再生数有很大的降低, 且最终集中于 1.36~2.5 之间, 除湖北省内城市外, 其它城市的基本再生数都降低到了 1.3~1.8 之间, 与季节性流感以及新型流感病毒相当, 表明该段时间内全国各地的抗疫攻坚战收效显著。

(1) 选取复工强度和基本再生数分别作为纵横坐标, 通过图 11 所示的矩阵分析可知, 几乎无

基本再生数小而复工强度低的城市, 这表明复工初期的疫情防控效果理想, 企业的复工愿望强烈, 同时也证明我国复工政策科学合理。

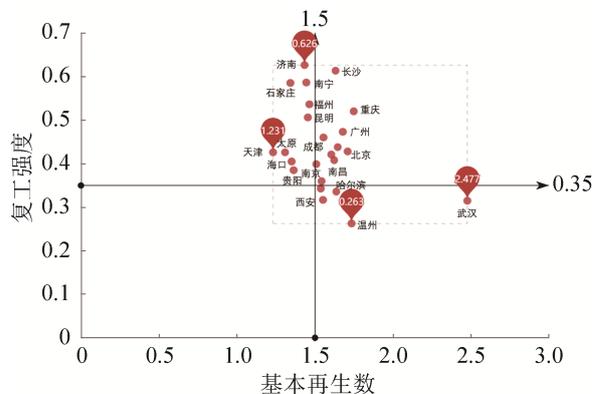


图 11 复工强度与基本再生数矩阵分析

Fig. 11 Matrix analysis of resumption intensity and basic reproduction number

(2) 复工强度较高且基本再生数也小的城市包括: 济南、南宁、福州、太原、海口、贵阳、天津、昆明、石家庄, 这类城市遭遇疫情时的传播风险较小, 因此可在合理布施防疫措施后适当降低企业复工审批要求, 满足各行业生产需要, 加快恢复地区经济秩序。

(3) 复工强度较低, 而基本再生数较大的城市包括: 武汉、温州、西安、深圳、杭州、哈尔滨, 由于这些城市面对此类大面积传播性疾病时的病毒传播风险相对较高, 尤其是深圳、西安、杭州等地还表现出了人力缺口大的问题, 为降低此类情况下的复工风险, 应以降低风险为首要目标, 严格遵守和坚持复工审批手续, 人员排查和回程复工人员 14 天隔离制度等, 实现安全有序复工。

(4) 复工强度较高, 而基本再生数也较高的城市包括: 长沙、重庆、广州、上海、北京、成都、合肥、南京、南昌等, 数据显示这类城市在复工初期就有大量人员回城复工, 同时疫情传播风险也较高, 同样情况下复工后, 必须从源头控制病毒扩散渠道, 建议可全方位应用 AI 和大数据手段, 在密集人群交通工具(公交、地铁、车站等)、用餐场所、工作场所等地, 通过判断人群众体温, 人群热力等实现聚集传播预警。

3.3 人力缺口和复工强度

根据各典型城市的春节前后人口流动情况和常住人口数, 推算出当时各城市的人力缺口状况, 结合表现出的复工强度进行矩阵分析, 探究面临不同程度人力资源问题时的复工策略。

(1) 图 12 的矩阵分析显示, 此次新冠疫情爆发后的复工初期, 国内复工强度高而人力缺口大的城市几乎没有, 表明我国各地区复工复产策略科学合理, 没有出现劳动力严重缺乏导致的复工困难。

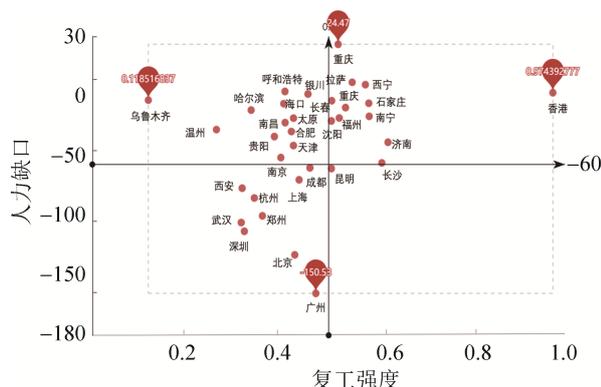


图 12 复工强度与人力缺口矩阵分析

Fig. 12 Matrix analysis of resumption intensity and manpower gap

(2) 重庆、香港、西宁、石家庄、济南、拉萨等城市表现出的复工强度较高且人力缺口小, 表明这些城市抗疫初期的复工状态良好, 类似情况下只需要提供足够的个人防护用品, 保持供应链上下端流通, 地方企业产能便快速恢复。

(3) 乌鲁木齐、呼和浩特、贵阳、温州等城市总体复工强度较低, 人力缺口较小, 说明在人员充足的情况下, 这些城市仍未展现出快速复工的趋势, 导致经济复苏缓慢, 为避免出现相同情况, 建议在以后面对类似传染性疾疾病时, 对于健康员工, 企业应考虑采用班车点到点接送, 午餐分时分餐, 或提供宿舍等形式, 引导员工积极复工。

(4) 北京、深圳、广州、上海、杭州、郑州等国内 GDP 强市呈现出了较低的复工强度和较大的人力缺口, 即在面对高传染性疾疾病时, 这些城市的复工需求依然强烈, 此时除通过加强企业间的合

作, 采取共享员工等形式充分利用好本地稀缺人力外, 还应针对拟回流人员以往的驻地, 回流风险, 回流后的隔离管理等进行精细化的预先评价和分析, 做好分批次回流的规划工作和风险防控, 降低人力回流过程中传播病毒的风险。

4 基于 K-means 无监督聚类分类回归的复工强度影响因素分析

在上述回顾性分析的基础上, 本节对各城市复工强度数据进行回归分析, 以探索影响城市复工强度的关键因素。由于各个城市经济活动特性有所不同, 因此有必要先分别依据经济发展水平(即人均 GDP)、基本再生数和人力缺口对城市进行聚类后, 再利用回归分析判断各因素与复工强度之间的相关性。采用 K-means 无监督聚类算法对选取的典型城市进行聚类分析, 具体过程如下:

输入无标记数据集: $\{x^{(1)}, x^{(2)}, \dots, x^{(n)}\}$, 其中 $x^{(i)} \in R^n$, 随后初始化聚类中心: $\mu_1, \mu_2, \dots, \mu_k \in R^n$, 建立度量有关类分配和类重心失真的目标函数, 定义为所有点到其类重心距离的平方和:

$$J(C, \mu) = \sum_{i=1}^m \|x^{(i)} - \mu_{C^{(i)}}\|^2$$

重复以下 2 个步骤, 实现在交替固定 C 和 μ 的过程中, 相对于另一个变量令 J 最优化:

1) 将所有的点分配给离它最近的类重心:

$$C^{(i)} = \arg \min_j \|x^{(i)} - \mu_j\|^2$$

2) 将类重心更新为每一类中所有点的均值:

$$\mu_j = \frac{\sum_{i=1}^m r_{ij} \{C^{(i)} = j\} x^{(i)}}{\sum_{i=1}^m r_{ij} \{C^{(i)} = j\}}$$

式中: r_{ij} 为指示变量, 当 $i=j$ 时, $r_{ij}=1$, 否则 $r_{ij}=0$; 随着函数 J 的值不断减小, 直到收敛, 最终得到聚类集合 $\{C_1, C_2, \dots, C_k\}$, 其中新的对应类别的数据集合为: $\{x^{(C_1)}, x^{(C_2)}, \dots, x^{(C_k)}\} \subseteq \{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$ 。

通过 python 分别对 3 个影响因素下的复工强度对应的数据集: $\{x_1^{(1)}, x_1^{(2)}, \dots, x_1^{(n)}\}$, $\{x_2^{(1)}, x_2^{(2)}, \dots, x_2^{(n)}\}$, $\{x_3^{(1)}, x_3^{(2)}, \dots, x_3^{(n)}\}$ 进行 K-means 算法的迭代计算, 得到聚类结果如图 13 所示。

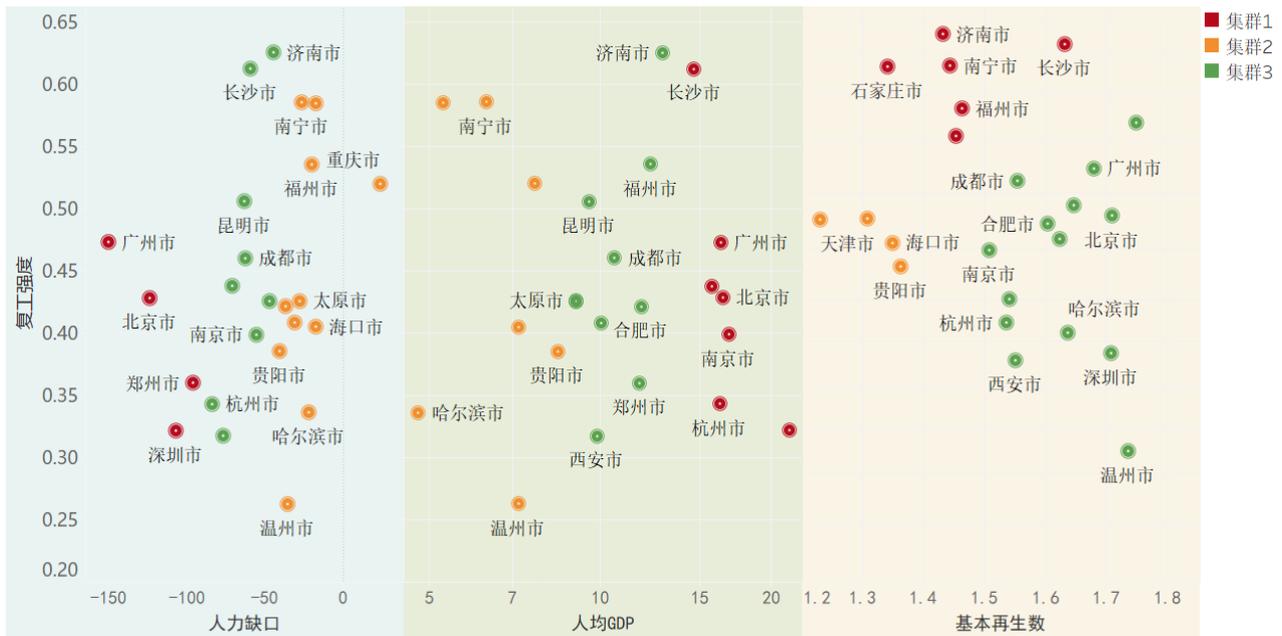


图 13 K-means 无监督聚类结果分布图

Fig. 13 Distribution graph of K-means unsupervised clustering results

根据聚类结果, 综合分析后获得 3 类集群, 对每一群进行变量拟合, 分别讨论每类城市在突发疫情后的复工强度与 3 个影响因素的相关性关系:

(1) 集群 1 包括北京、深圳、上海、广州、南京、杭州等城市。进行回归分析后集群的总体 p 值提示不显著, 但人均 GDP 对复工强度的影响水平较高, 且 R 方达到了 0.720, 说明这些城市的复工强度与这 3 个因素, 尤其是人均 GDP 之间存在一定的相关关系。

(2) 集群 2 包括长沙、济南、福州、昆明、南宁、石家庄等城市。该集群的 R 方较弱, 即这些城市的复工强度与人均 GDP、基本再生数以及人力缺口等无明显相关性, 其原因可能与空间距离、交通量的差异有关。

(3) 集群 3 包括太原、天津、贵阳、海口等城市。值得注意的是, 回归分析结果显示这一集群的 R 方=1.000, 且 p 值 \ll 0.000 1, 证明这些城市的复工强度几乎完全由上述因素决定, 相关性十分显著。

如今, 新型冠状病毒感染肺炎疫情仍在持续蔓延, 给世界经济和国际秩序带来了长期、严重的负面影响, 为把疫情造成的经济、生命财产等损失降

到最低, 各个国家和地区在疫情得到有效控制后, 首要任务就是在常态化的疫情防控中稳步恢复各行业生产生活秩序, 抓紧推进企业复工复产进程。从回归分析结果可以看出, 除政府宏观调控政策等因素外, 各个地区的复工复产活动主要受当地社会经济发展的驱动, 个人和企业复工需求较高的城市可以首先因时、因地制宜地制定计划开展复工复产活动, 主要管理结论包括:

(1) 当病毒传播的基本再生数处于 1.3~1.8 之间时, 其传染性与季节性流感基本相当, 此时在有效的防控下, 可制定相应措施, 尽快复工复产, 恢复地方经济秩序。

(2) 对于人均 GDP 高的城市, 对于复工复产需求也高, 而在疫情期间要有序地恢复复工复产, 人力缺口大是其关键的制约因素, 各地的管理部门应当首先正确判断各行业、企业存在的实际人力资源缺口后, 然后根据自身地理区位、交通设施、流量等因素, 并且综合考虑当地的卫生医疗条件、取得的疫情防控成效以及潜在的病毒传播风险等情况, 制定一系列防控方案措施, 有序调动各类社会资源安排人员安全返城、复工。

以上有关分析对于局部地区在有效避免聚集性病毒感染事件发生的同时, 加快促进局部生产力和经济秩序恢复的工作具有一定指导意义。

5 结论

基于百度迁徙数据、春运数据、常住人口普查数据以及新冠肺炎病例数, 本文通过构建模型量化分析 2020-02-01 至 2020-02-18 期间各典型城市的疫情传播情况和复工强度, 且计算所得的各城市的现有人数比例与复工强度之间具有明显的正相关关系, 证明所构建的模型能够合理地判断城市的复工状况。通过所得的人力缺口, 和利用 SEIR 仓室仿真模型估计的疫情初期, 各城市新冠肺炎病毒感染肺炎的基本再生数, 分别与复工强度进行矩阵分析, 从多个角度总结“中国经验”, 从而给各国抗击新冠肺炎疫情提供有效支持。结合 K-means 无监督聚类 and 回归分析方法, 发现人力缺口、基本再生数、城市人均 GDP 是影响城市复工强度的关键因素。

参考文献:

- [1] WHO. Corona Virus Disease 2019(COVID-19) Situation Report-51[EB/OL]. (2020-03-11)[2020-03-15]. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_10.
- [2] 叶莹, 范威, 王文华, 等. 新型冠状病毒肺炎聚集性疫情中无症状感染者与确诊患者的流行差异[J/OL]. 中国感染控制杂志: 1-6[2020-05-30]. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=GRKZ202006002&v=TC1jXruv1g6NutnRDQXjNGkOuzHtLXrN8CVv14IXv8EtVPGAbFXaxoEtQ0DIXIT6>
Ye Ying, Fan Wei, Wang Wenhua, et al. Difference in Epidemic Characteristics between Asymptomatic Infected Persons and Confirmed Cases in COVID-19 Clustered Epidemics[J/OL]. Chinese Journal of Infected Control: 1-6[2020-05-30]. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=GRKZ202006002&v=TC1jXruv1g6NutnRDQXjNGkOuzHtLXrN8CVv14IXv8EtVPGAbFXaxoEtQ0DIXIT6>
- [3] 陈苏虹, 戴俊明, 胡俏, 等. 2019 冠状病毒病毒 (COVID-19) 暴发疫情下公众焦虑状况及其影响因素[J/OL]. 复旦学报(医学版): 1-7[2020-05-30]. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=SHYK202003015&v=VCoD1s%25mmd2FvESjECWTG6xB9KBTNh5dbZH84Wjhk6nVIVwIm%25mmd2BKbXZvg1%25mmd2Bo26ElvxEpZe>
Chen Suhong, Dai Junming, Hu Qiao, et al. Public Anxiety and Its Influencing Factors in the Initial Outbreak of COVID-19[J/OL]. Fudan University Journal of Medical Sciences: 1-7[2020-05-30]. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=SHYK202003015&v=VCoD1s%25mmd2FvESjECWTG6xB9KBTNh5dbZH84Wjhk6nVIVwIm%25mmd2BKbXZvg1%25mmd2Bo26ElvxEpZe>
- [4] 王皎贝, 李钢, 王建坡, 等. 陕西省 COVID-19 疫情时空演化与风险画像[J/OL]. 热带地理: 1-18[2020-05-30]. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=RDDD202003007&v=%25mmd2FfApa7sGF5sofZbET9yBGhzjzIxjEukzuoRA7NLhlzcRA11DtF3gdYrtPpuFmh1>
Wang Jiaobei, Li Gang, Wang Jianpo, et al. Spatiotemporal Evolution and Risk portrait of COVID-19 Epidemic in Shaanxi Province[J/OL]. Tropical Geography: 1-18[2020-05-30]. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=RDDD202003007&v=%25mmd2FfApa7sGF5sofZbET9yBGhzjzIxjEukzuoRA7NLhlzcRA11DtF3gdYrtPpuFmh1>
- [5] 刘勇, 杨东阳, 董冠鹏, 等. 河南省新冠肺炎疫情时空扩散特征与人口流动风险评估—基于 1243 例病例报告的分析[J]. 经济地理, 2020, 40(3): 24-32.
Liu Yong, Yang Dongyang, Dong Guanpeng, et al. The Spatio-Temporal Spread Characteristics of 2019 Novel Coronavirus Pneumonia and Risk Assessment Based on Population Movement in Henan Province: Analysis of 1243 Individual Case Reports[J]. Economic Geography, 2020, 40(3): 24-32.
- [6] 刘逸, 李源, 黎卓灵, 等. 新冠肺炎疫情在广东省的扩散特征[J/OL]. 热带地理: 1-9[2020-05-30]. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=RDDD202003001&v=%25mmd2FfApa7sGF5uFxmMcumJ83829rNctOEB7siyTTIF51kUhaBcaOP4jCalCo0pbJd7g>
Liu Yi, Li Yuan, Li Zhuoling, et al. The Spreading Characteristics of New Coronary Pneumonia in Guangdong province[J/OL]. Tropical Geography: 1-9[2020-05-30]. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=RDDD202003001&v=%25mmd2FfApa7sGF5uFxmMcumJ83829rNctOEB7siyTTIF51kUhaBcaOP4jCalCo0pbJd7g>

- <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=RDDD202003001&v=%25mmd2FfApa7sGF5uFxmCmumJ83829rNctOEB7siyTTIF51kUhaBcaOP4jCalCo0pbJd7g>
- [7] 陈芳, 吴家兵, 姜静静, 等. 安徽省新型冠状病毒肺炎聚集性疫情流行特征与防控措施分析[J]. 中国公共卫生, 2020, 36(4): 469-472.
Chen Fang, Wu Jiabing, Jiang Jingjing, et al. Epidemiological Characteristics of Clusters of COVID-19 in Anhui province[J]. Chin J Public Health, 2020, 36(4): 469-472.
- [8] 王刚, 李玥琪, 蒋俊俊, 等. 广西不同地区新型冠状病毒肺炎的流行特征[J]. 热带医学杂志, 2020, 20(4): 431-433, 572.
Wang Gang, Li Yueqi, Jiang Junjun, et al. The Epidemic Characteristics of COVID-19 in Different Regions of Guangxi[J]. Journal of Tropical Medicine, 2020, 20(4): 431-433, 572.
- [9] 赵宏波, 魏甲晨, 王爽, 等. 大城市新冠肺炎疫情风险评估与精准防控对策—以郑州市为例[J]. 经济地理, 2020, 40(4): 103-109, 124.
Zhao Hongbo, Wei Jiachen, Wang Shuang, et al. The Risk Assessment of Covid-2019 Epidemic in Metropolis and Precise Prevention and Control Measures: A Case Study of Zhengzhou City[J]. Economic Geography, 2020, 40(4): 103-109, 124.
- [10] 卿菁. 特大城市疫情防控机制: 经验、困境与重构—以武汉市新冠肺炎疫情防控为例[J]. 湖北大学学报(哲学社会科学版), 2020, 47(3): 21-32.
Qing Jing. Mechanisms for Epidemic Prevention and Control in Megacities: Experience, Dilemma, and Reconstruction-Taking Wuhan New Coronary Pneumonia Epidemic Preventi and Control as an Example[J]. Journal of Hubei University (Philosophy and Social Science), 2020, 47(3): 21-32.
- [11] 王英鉴, 张娜, 吕涵路, 等. 上海市新型冠状病毒肺炎发病趋势初步分析[J]. 上海预防医学, 2020, 32(2): 142-146.
Wang Yingjian, Zhang Na, Lv Hanlu, et al. Preliminary Analysis on the Incidencetrend of Novel Coronavirus Pneumonia in Shanghai[J]. Shanghai Journal of Prevention Medicine, 2020, 32(2): 142-146.
- [12] 耿辉, 徐安定, 王晓艳, 等. 基于 SEIR 模型分析相关干预措施在新型冠状病毒肺炎疫情中的作用[J]. 暨南大学学报(自然科学与医学版), 2020, 41(2): 175-180.
Geng Hui, Xu Anding, Wang Xiaoyan, et al. Analysis of the Role of Current Prevention and Control Measures in the Epidemic of Corona Virus Disease 2019 Based on SEIR Model[J]. Journal of Jinan University (Natural Science & Medicine Edition), 2020, 41(2): 175-180.
- [13] Wu J T, Leung K, Leung G M. Nowcasting and forecasting the potential domestic and international Fpread of the 2019-nCoV Outbreak Originating in Wuhan, China: a Modelling Study[J/OL]. The Lancet, [2020-02-20]. [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)30260-9/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30260-9/fulltext). DOI:10.1016/S0140-6736(20)30260-9.
- [14] Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-infected Pneumonia[J]. New England Journal of Medicine, 2020, 382(13): 1199-1207. DOI: 10.1056/NEJMoa2001316.
- [15] Tang B, Wang X, Li Q, et al. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions[J]. Journal of Clinical Medicine, 2020, 9(2): 462. DOI: <https://doi.org/10.3390/jcm9020462>.
- [16] Zhao S, Musa S S, Lin Q, et al. Estimating the Unreported Number of Novel Coronavirus (2019-nCoV) Cases in China in the First Half of January 2020: A Data-Driven Modelling Analysis of the Early Outbreak[J]. Journal of Clinical Medicine, 2020, DOI: <https://doi.org/10.3390/jcm9020388>.
- [17] Jia J S, Lu X, Yuan Y, et al. Population Flow Drives Spatio-temporal Distribution of COVID-19 in China[J]. Nature, 2020, 582: 389-394. DOI: <https://doi.org/10.1038/s41586-020-2284-y>.
- [18] 杨桥桥, 孙益楠, 龚强. 基于电力数据的企业复产指数研究[J]. 财富时代, 2020(4): 41.
Yang Jiaoqiao, Sun Yinan, Gong Qiang. Research on Enterprise Resumption Index Based on Electricity Data[J]. Times of Fortune, 2020(4): 41.
- [19] 刘光浩. 国内外大数据助力疫情防控经验的启示[J]. 中国电信业, 2020(4): 68-71.
Liu Guanghao. Inspiration from the Experience of Big Data at Home and Abroad in Assisting Epidemic Prevention and Control[J]. China Telecommunications Trade, 2020 (4): 68-71.
- [20] 吴兆春, 林柳琳. 广州企业在复工复产中遇到的新困难及对策建议[J]. 广东经济, 2020(5): 68-71.
Wu Zhaochun, Lin Liulin. New Difficulties Encountered by Guangzhou Enterprises in Resuming Production and Production and Countermeasures[J]. Guangdong Economy, 2020(5): 68-71.
- [21] 中华人民共和国交通部. 2020 年春运全国共发送旅客

- 14.8 亿人次 [BM19000004][2020-02-20]. <http://www.mot.gov.cn/>.
- [22] 百度地图慧眼. 百度迁徙[京 ICP 证 030173 号]. [2020-02-20]. <https://qianxi.baidu.com/>.
- [23] 中华人民共和国国家统计局. 主要城市年度数据(2018) [EB/OL][2020-02-03]. <http://data.stats.gov.cn/easyquery.htm?cn=E0105>.
National Bureau of Statistics of the People's Republic of China. Annual Data of Major Cities(2018)[EB/OL] [2020-02-03]. <http://data.stats.gov.cn/easyquery.htm?cn=E0105>.
- [24] 周涛, 刘权辉, 杨紫陌, 等. 新型冠状病毒感染肺炎基本再生数的初步预测[J]. 中国循证医学杂志, 2020, 20(3): 359-364.
Zhou Tao, Liu Huiquan, Yang Zimo, et al. Preliminary Prediction of the Basic Reproduction Number of the Novel Coronavirus 2019-nCoV[J]. Chinese Journal of Evidence-Based Medicine, 2020, 20(3): 59-364.
- [25] 盛华雄, 吴琳, 肖长亮. 新冠肺炎疫情传播建模分析与预测[J]. 系统仿真学报, 2020, 32(5):759-766.
Sheng Huaxiong, Wu Lin, Xiao Changliang, Modeling Analysis and Prediction on NCP Epidemic Transmission [J]. Journal of System Simulation, 2020, 32(5): 759-766.
- [26] 焦松, 楚威, 毛少杰. 基于聚类和相关性分析的作战效能仿真评估方法[J]. 系统仿真学报, 2015, 27(9): 2083-2088.
Jiao Song, Chu Wei, Mao Shaojie. Simulation Evaluation Method of Operational Effectiveness Based on Clustering Analysis and Correlation Analysis[J]. Journal of System Simulation, 2015, 27(9): 2083-2088.