

基于地理加权随机森林模型的流动人口驱动因子研究*——以湖南省县域为例

Decoupling the Driving Factors of Floating Population Using a Geographically Weighted Random Forest Model: A County-Level Case Study in Hu'nan Province

朱 珠 郑卫民 李 晟 吴 博 周宸旭 ZHU Zhu, ZHENG Weimin, LI Sheng, WU Bo, ZHOU Chenxu

摘 要 目前,县域人口流动的驱动机制研究多基于传统线性模型,难以捕捉驱动因子的非线性阈值效应和空间异质性。引入机器学习中的RF和GWRF模型,通过构建“系统阈值—地理权重”双维框架,利用2000—2020年人口普查数据融合多源地理大数据,以湖南省县域为研究对象,探究经济、社会、环境和区位4类因子对人口县内流动、跨县流入和跨县流出的作用规律。研究发现:(1) 湖南省人口流动呈现梯度分异和多中心演化趋势;(2) 湖南省人口流动主要受经济要素限制和生态要素约束;(3) 流动人口的驱动机制存在明显的阈值效应和边际递减规律;(4) 流动人口的驱动机制在空间上存在明显的“核心极化—边缘收缩”特征。研究结果可为中部地区的人口流动管理提供“动态阈值预警—空间靶向施策”的方法论参考。

Abstract Prevailing studies on the investigation of the driving mechanisms of county-level population movements relying on traditional linear models inadequately capture the nonlinear threshold effects and spatial heterogeneity. This study innovatively integrates machine learning techniques (Random Forest and Geographically Weighted Random Forest) to establish a dual-dimensional "systematic threshold-geographical weighting" analytical framework. Leveraging integrated datasets from population censuses (2000-2020) and multi-source geospatial big data across Hu'nan Province's counties, we systematically investigate how economic, social, environmental, and locational factors govern intra-county mobility, inter-county immigration, and outmigration. Key findings reveal: (1) a gradient differentiation pattern with polycentric evolutionary trends in population redistribution, (2) dual constraints from economic thresholds and ecological carrying capacities, (3) pronounced threshold effects with diminishing marginal impacts across driving factors, and (4) distinct core polarization-edge contraction characteristics in spatial mechanisms. The proposed methodology advances dynamic threshold monitoring and spatially adaptive governance strategies, offering policy insights for managing population mobility in central China's transitional regions.

关键词 流动人口;驱动因子;非线性阈值;空间异质性;地理加权随机森林模型;湖南省县域

Key words floating population; driving factors; nonlinear threshold; spatial heterogeneity; geographically weighted random forest (GWRF) model; county-level units in Hu'nan Province

文章编号 1673-8985 (2025) 06-0099-09 中图分类号 TU984 文献标志码 A

DOI 10.11982/j. supr. 20250613

作者简介

朱 珠

南华大学建筑与设计艺术学院

硕士研究生

郑卫民 (通信作者)

南华大学建筑与设计艺术学院

教授,博士,2021000028@usc.edu.cn

李 晟

南华大学建筑与设计艺术学院

副教授,博士

吴 博

南华大学建筑与设计艺术学院

讲师,硕士

周宸旭

南华大学经济与管理法学院

硕士研究生

0 引言

中国人口高速增长的时代已经过去,现正面临人口零增长乃至负增长的局势^[1]。随着人口自然增长率降低,其对区域人口变动的影 响逐渐弱化,流动人口逐渐成为区域人口变动的关键力量。

已有研究表明,人口流动特征与区域发展密切相关^[2-4],流动人口主要集聚于高度发

*基金项目:湖南省社会科学基金重点委托项目“长株潭城市群生态安全保障机制研究”(编号24WTB33);湖南省教育厅资助重点科研项目“多尺度耦合的绿地生态网络构建及优化研究”(编号22A0309)资助。

展的长三角、珠三角、京津冀等城市群,且集中于这些区域的超大特大城市^[5-10]。然而,随着21世纪的经济内陆化及国家区域协调发展战略的持续推进,人口流动趋势发生结构性转变,中西部地区以强劲的经济增长重塑了区域人口流动格局^[11-13]。在全国层面,呈现人口由东部聚集向中西部扩散回流的趋势^[14],中西部省会和区域中心城市成为新兴的人口流入增长极^[15-16]。区域层面上,东北、西北和西南地区内部联系增强,形成成渝城市群、长江中游城市群、中原城市群、西安城市群及甘肃—宁夏城市群^[17-19]。省域层面上,中西部省内流动增速普遍高于省际流动,其中,川渝地区的成渝“双核心”空间格局明显,短距离内部流动成为主导^[20]。

已有研究对人口流动驱动机制有3大理论范式^[21-22]。一是基于推拉理论的经济趋利模型,强调区域收入差距^[23]、就业机会^[24]等要素的线性作用;二是新经济地理学的“核心—边缘”理论,聚焦报酬递增与运输成本的空间交互效应^[25];三是环境迁移理论,关注生态约束^[26]和气候胁迫^[27]对人口流动的阈值调控。在研究方法上,多采用主成分分析^[28-29]、最小二乘法、多元回归^[30]等传统线性回归模型,从宏观视角揭示驱动因子的全局线性作用。近年来,机器学习模型因其非线性拟合优势被引入人口研究^[31]。机器学习中的随机森林RF模型可以较好预测自变量与因变量之间的非线性关系,有效识别自变量的影响范围。尽管这些理论和方法为解释人口流动提供了重要视角,但仍存在局限性。一方面,现有研究多从宏观视角解析驱动因子贡献度,缺乏对驱动因子非线性阈值的精细化刻画;另一方面,传统线性回归模型过度依赖全局均质性假设,难以刻画关键变量的空间分异规律。

基于此,本文引入机器学习中的随机森林RF模型和地理加权随机森林GWR模型^[32],通过构建“系统阈值—地理权重”双维框架,以湖南省122个县域为研究单元,基于2000—2020年人口普查数据及多源地理环境数据,系统探究经济、社会、环境、区位

4类因子对流动人口的非线性作用与空间异质性规律,对人口流动驱动因子进行阈值效应量化和空间显性探究,以期对湖南省人口空间布局优化提供“动态阈值预警—空间靶向施策”的治理策略,推动中部地区新型城镇化战略精准施策。

1 数据来源与研究方法

1.1 研究区域概况

湖南省位于我国中部,处于连接东部沿海和中西部地区的关键地理位置。其由全省经济增长极的长株潭城市群板块、比邻粤港澳大湾区的湘南板块、对接长江经济带的环洞庭湖生态经济板块和偏远的大湘西板块构成,内部社会经济梯度差异显著^[33]。第七次全国人口普查数据显示,长株潭三市作为极核虹吸了省内大部分流动人口,而湘西、湘南地区人口持续外流,人口空间失衡问题日益凸显,并且这种人口流动的极化效应正在拉大区域发展差异,成为湖南省实施新型城镇化建设的阻碍。

1.2 数据来源和处理

流动人口数据来源于人口普查分县数据。本文的流动人口指广义人户分离人口,即居住地与户口登记地所在的乡镇街道不一致的人口^[34]。根据人口跨越的行政边界不同,本文将县域流动人口划分为3类:县内流动人口、跨县流入人口和跨县流出人口^[35]。

1.3 研究方法

1.3.1 随机森林RF模型

随机森林RF模型是由Breiman提出的一种基于分类回归树的非参数机器学习模型,运用了集成学习的思想,利用多个弱分类器进行分类及预测^[36],能够解决单棵决策树带来的过拟合问题,同时对数据本身的噪声和异常值有较强的容错能力。在构建模型时,通常采用2/3的样本作为训练数据集,将剩下的1/3样本作为袋外数据用于交叉验证,以测量模型的性能。同时,袋外数据也可以用来评估变量的重要性得分,值越高表明变量对因变量的解释力

越强^[37]。

1.3.2 地理加权随机森林GWR模型

地理加权随机森林GWR模型是一种将地理加权回归GWR和RF优势结合在一起的机器学习算法。其工作原理是,先将研究区域划分为更小的区域,对每个小区域进行RF建模,然后估计各区域响应变量与解释变量之间的空间变化关系,综合考虑变量间的非线性影响、交互作用,以及空间异质性和空间自相关^[32]。

1.3.3 模型性能评价

本文采用判定系数 R^2 、平均绝对误差 E_{MAE} 及均方根误差 E_{RMSE} 来评价RF模型和GWR模型的拟合优度及整体模拟性能。

$$E_{MAE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (1)$$

$$E_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

式中: n 为样本数量; y_i 为实测值; \hat{y}_i 为模型的预测值; \bar{y} 为样本的平均值。

1.4 研究框架

本文基于2000年、2010年和2020年的混合截面数据库,综合了全局和局部视角,构建了人口流动驱动因子“系统阈值—地理权重”双维探究框架。通过RF模型识别全局视角下县内流动率、跨县流入率和跨县流出率的关键驱动因子,并将其与GWR模型在局部视角下识别的关键驱动因子进行综合比较,提取综合影响前8的关键驱动因子,进行非线性与空间异质性分析,捕捉各驱动因子对人口流动的非线性阈值效应,并识别不同地理单元中驱动因子的差异化作用强度。“系统阈值—地理权重”双维框架整合了空间计量方法与机器学习,突破了传统线性分析的局限,既保留了全局规律性,又兼顾了局部空间异质性,为流动人口管理提供了“动态阈值预警—空间靶

向施策”的方法论,构建了“识别—预警—干预”的闭环管理体系。

本文主要聚焦于识别驱动因子长期稳定的驱动机制,同时考虑到样本量过小而无法兼顾模型可靠性的问题,故以2000年、2010年和2020年湖南省县内流动率、跨县流入率和跨县流出率为因变量,以2000年、2010年和2020年经济、社会、环境、区位4个纬度18个因子为自变量,建立混合截面数据库(见图1)。

2 结果与分析

2.1 流动人口时空格局

近20年来,湖南省流动人口规模不断扩大,总流动人口从852万人增加到3 442万人。其中,省会中心城区流动人口规模扩张最显著,岳麓区总流动人口从13.1万人增长至109.1万人,虹吸效应显著。

从流动人口结构演变来看(见图2),全省县内流动率呈下降趋势,从29.8%下降至23.4%,以雨花区为代表,县内流动率从41.4%降至13.5%,在一定程度上反映了县内人口再分布功能的弱化;人口不断流向核心城区,芙蓉区和岳麓区流入率持续攀升,人口集聚效应持续增强;而偏远县域的流出压力也未得到缓解,城步苗族自治县的流出率始终维持在70%以上,新田县流出率虽有下降,但人口外流仍处于主导趋势。

从空间分布特征来看,2020年流动人口格局呈现出显著的梯度分异特征。长沙市内的5个区形成核心区域,流入率均高于60%,但县内流动率普遍低于25%,外来人口主导特征突出;以长沙县和望城区为代表的近郊县域呈现出双向平衡流动,流入率和流出率之比趋近于1,表明城郊过渡带的人口交互性增强;而邵阳县、南县、新邵县等偏远县域流出率仍高于80%,县内流动率多低于15%,边缘区域的人口流失困境仍旧险峻。从区域对比角度进行分析,包括长沙、株洲和湘潭在内的湘中北地区流入率均值为58.7%,净流入率均值为63.2%,形成了集聚核心;包括郴州、永州

在内的湘南地区流出率均值为74.3%,以跨县流出为主;包括张家界、湘西州在内的湘西地区县内流动率均值为28.6%,其中永定区高达54.0%,本地化潜力不断凸显,可能与生态旅

游导向的发展策略有关。

综上所述,湖南省人口流动格局呈现出“单极强化—多中心萌芽”的空间特征,其中,作为极核的长沙市中心城区的极化效应持续

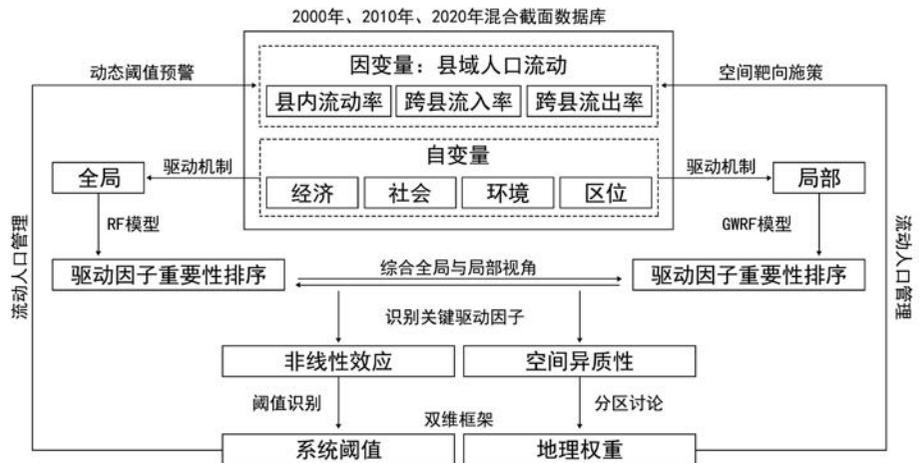


图1 “系统阈值—地理权重”双维框架图
Fig.1 Bivariate framework diagram of "systemic thresholds and geographic weights"

资料来源:笔者自绘。

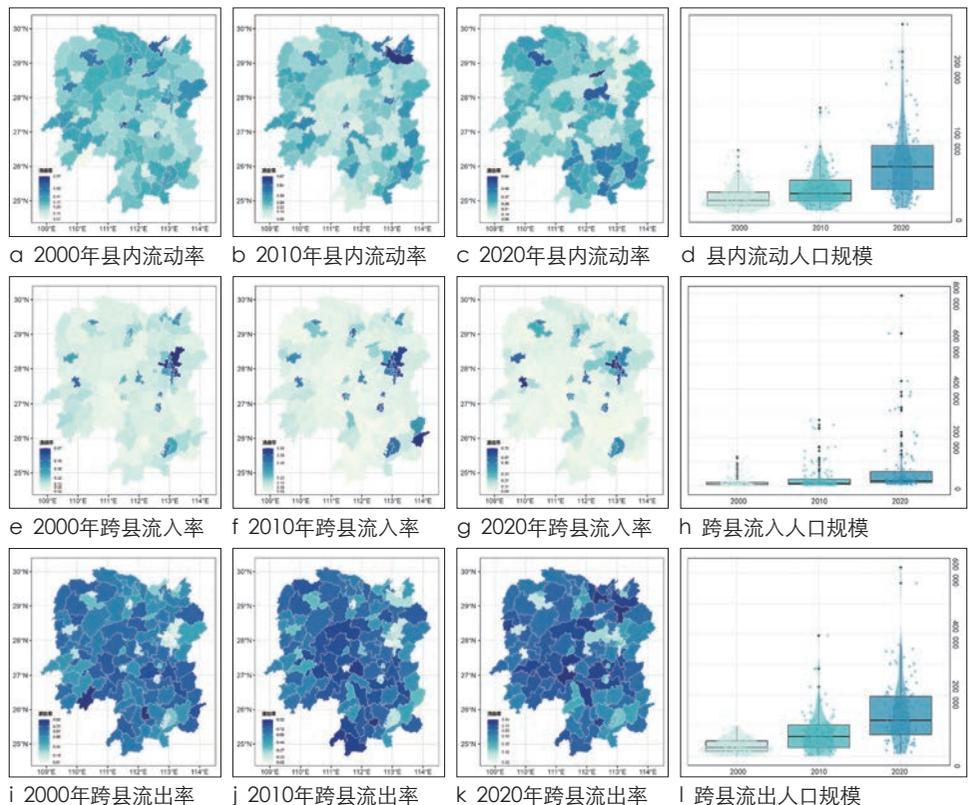


图2 湖南省县域流动人口结构时空分异图谱(2000—2020年)

Fig.2 Spatiotemporal differentiation patterns of floating population structure in Hu'nan Province's counties (2000-2020)

资料来源:笔者自绘。

强化,岳阳市与衡阳市两个次级中心对人口的强吸纳能力逐渐凸显,人口流动多中心网络初见雏形。人口持续外流的湘西湘南等偏远县域也出现一批以永定县为代表的转型县域,为边缘县域的发展提供了新思路。

2.2 流动人口驱动机制研究

2.2.1 驱动因子选取

据国内外学者对流动人口分布影响因素做出的实证研究,发现影响流动人口分布的因素十分复杂,需从自然环境、社会经济发展、历史条件等多维度进行系统分析。为探究近20年湖南省县域人口流动的驱动机制,参考前人研究结果并结合湖南省的实际情况,选取了经济、社会、环境、区位4个维度18个驱动因子展开研究(见表1)。

新古典经济学强调不均衡的区域经济格局是影响人口再分布的决定性因素,人口主要向经济发达的地区集聚,城镇化发展水平高的区域可能人口增长更快。因此,本文将城镇化率作为城镇化水平的表征指标,以就业人口数量表征城市就业水平,同时以夜间灯光指数表征城市经济发展水平,以居民年均工资表征居民收入水平,以农作物播种面积表征农业发展水平。

随着生活条件改善,区域公共服务设施配置成为影响人口流动的重要因素。本文从交通、医疗、教育、购物、娱乐、金融设施等方面对城市的公共服务水平进行衡量。

城市的居住舒适度也是影响人口流动的重要因素。绿化率对人口的吸引有正向驱动,温差、降水等气候条件对人口流动也存在一定程度的影响。本文选取了植被归一化指数、PM_{2.5}浓度、地形起伏度、年均降水量和温差作为衡量城市环境因素对人口流动影响的指标。

从“强省会”战略实施的效果来看,以长沙市为极核带动周边城市发展效果良好,可见区域中心城市在各方面都具有一定优势。区位因素上,地级市、省会城市对周边地区的辐射作用更容易吸引人口聚集。本文采用与所在地

级市市辖区最近距离及与省会城市最近距离来表征区位条件。

2.2.2 模型评价

随机选择2000年、2010年和2020年3年数据集中的70%作为训练样本,分别构建RF模型和GWR模型,并用数据集中剩余的30%来校验。县内流动率、跨县流入率和跨县流出率中RF模型的R²分别为0.837、0.967、0.942, GWR模型的R²分别为0.987、0.998、0.993,可见GWR模型的拟合优势更为突出,且其MSE和RMSE都远小于RF模型,说明GWR模型的模拟误差远小于RF模型。综合运用上述两个模型,对湖南省流动人口的驱动机制进行全局及局部的分析,探究这些驱动因子

对人口流动的非线性作用机制及空间异质性特征。

2.2.3 驱动因子的全局及局部重要性测度

RF模型和GWR模型的双模型检验显示(见图3),全局和局部的核心驱动因子具有一致性,但GWR模型引入的空间异质性参数,综合考虑了传统全局模型中忽视的地理分异规律。

从全局视角看,城镇化水平X1、农业发展水平X5在县内流动、跨县流入和跨县流出中均发挥了关键作用,研究结果稳定地验证了城镇化进程对人口空间重构的基础性作用。同时,就业水平X2、居民收入水平X4在县内流动中的协同效应显著,印证了人口迁移理

表1 变量描述

Tab.1 Variable description

维度	解释变量	指标说明	数据来源
经济因素	城镇化水平X1	各县级单位城镇人口占总人口的比重/%	湖南省人口普查分县数据
	就业水平X2	各县级单位就业人口数量/人	湖南省人口普查分县数据
	经济发展水平X3	各县级单位平均夜间灯光指数	国家地球系统科学数据中心
	居民收入水平X4	年均工资/元	湖南省统计年鉴
	农业发展水平X5	农作物播种面积/千公顷	湖南省统计年鉴
社会因素	交通设施X6	各县级单位公交、地铁站点数量/个	高德POI数据
	医疗设施X7	各县级单位卫生机构数量/个	高德POI数据
	教育设施X8	各县级单位中等学校、普通小学数量/个	高德POI数据
	购物设施X9	各县级单位商场、大型超市、便利店数量/个	高德POI数据
	娱乐设施X10	各县级单位影剧院、KTV、公园、游乐园数量/个	高德POI数据
	金融设施X11	各县级单位银行数量/个	高德POI数据
环境因素	植被覆盖度X12	各县级单位植被归一化指数	中国科学院资源环境科学与数据中心
	空气质量X13	各县级单位PM _{2.5} 浓度/ug·m ³	Zenodo欧洲科学数据开放平台
	地形地貌X14	各县级单位平均地形起伏度/m	地理空间数据云平台
	降水X15	各县级单位年平均降水量/mm	国家观测站降水数据
	温差X16	各县级单位1月与7月温差/°C	国家地球系统科学数据中心
区位因素	市辖区位X17	至地级市市辖区的空间单位距离/m	百度地图
	省城区位X18	至省会(长沙市)的空间单位距离/m	百度地图

资料来源:笔者自制。

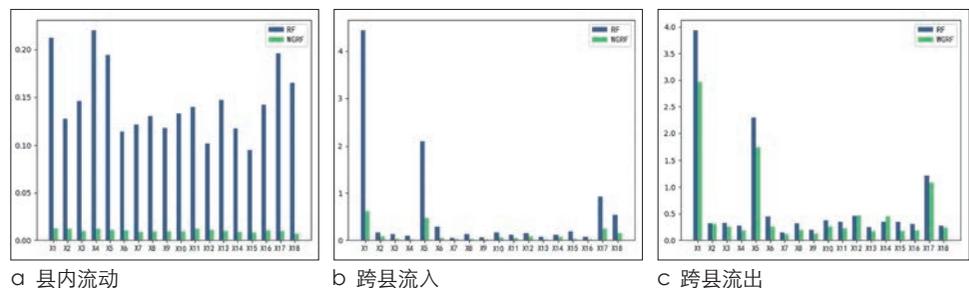


图3 驱动因子重要性得分图

Fig.3 Impact factor importance score chart

资料来源:笔者自绘。

论中的经济趋利机制。在跨县流动中,市城区位X17和省城区位X18在跨县流动中的重要性排名靠前,表明区位是影响人口跨县流动的重要因素。在跨县流出中,植被覆盖度X12、降水X15重要性突出,揭示了推拉理论中环境约束力机制。

GWRF模型的局部解析进一步弥补了传统方法的局限性。从局部视角来看,县内流动中的就业水平X2的重要性排名较全局排名上

升了10位,表明就业机会的不均衡显著改变了人口的县内流动。在跨县流动场景中,降水X15重要性排名显著下降,流入排名从6降为12,流出排名从9降为15,表明降水这类气候条件对人口的流动存在显著的区域空间效应。

综上所述,城镇化水平和农业发展水平是构成人口流动的基础驱动力,区位条件在跨县流动中驱动明显,其作用强度具有空间普适性;气候因素显著的空间异质性往往被全局模

型低估, GWRF模型为识别空间上的变化提供了有效工具。

2.2.4 驱动因子的非线性作用机制

选取3类流动人口中综合重要性排序前8位的驱动因子,分别探究其对人口流动率的作用规律。研究发现,各类驱动因子对人口结构的调控呈现复杂的非线性特征,其作用方向和强度受阈值效应、边际递减、阶级跃迁等多重机制主导(见图4)。

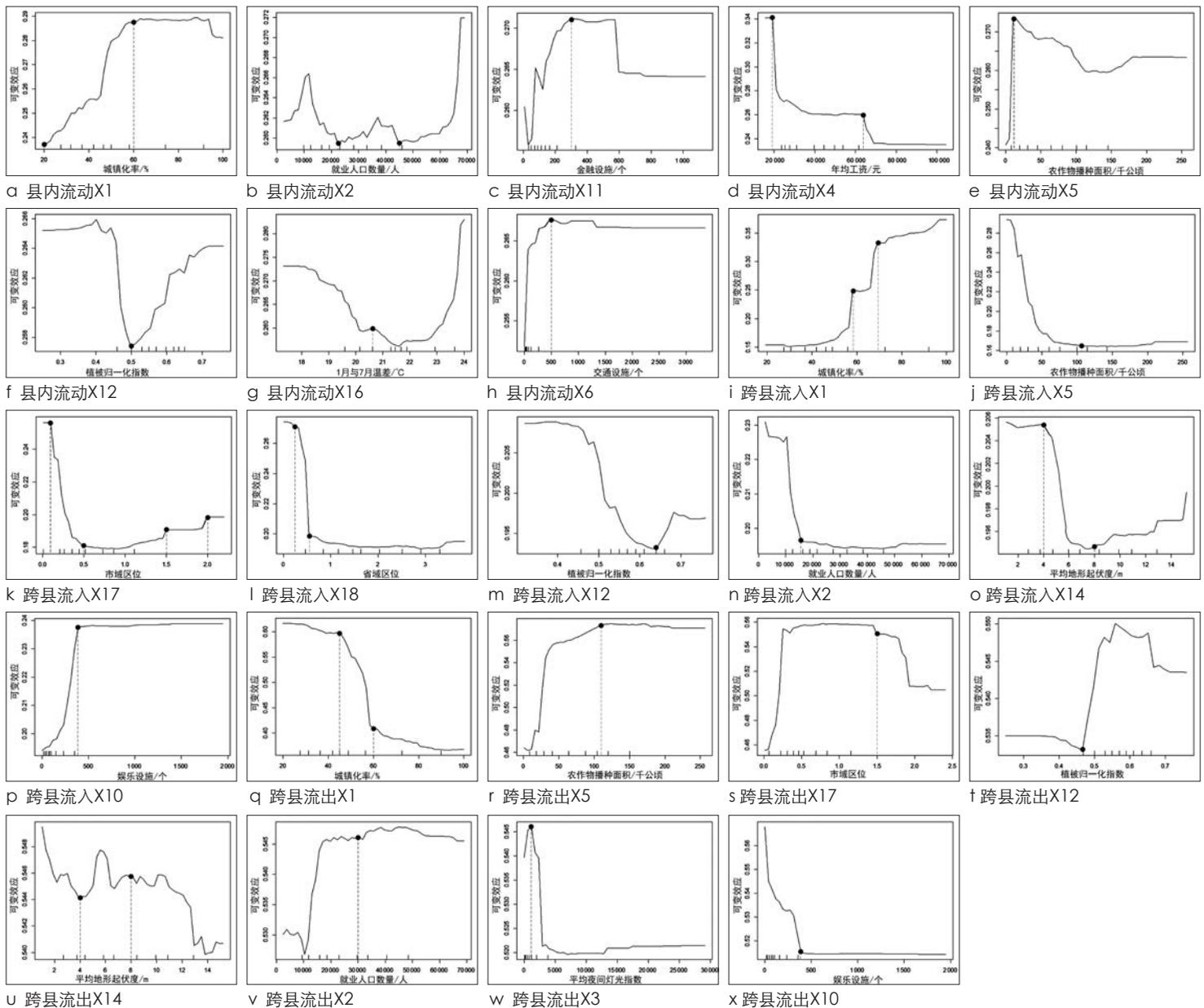


图4 驱动因子非线性作用机制图
Fig.4 Diagram of the non-linear mechanism of action of the influencing factors

资料来源:笔者自绘。

在经济因素中,城镇化率X1对县内流动率的驱动呈现边际递减效应,当城镇化率低于20%时,流动增速仅为0.3%,当城镇化率突破20%后,增速跃至2.1%,但在达到60%阈值后增速回落至0.5% (见图4a)。这种由加速集聚到稳定饱和的转变,在跨县流入和流出中都得到了体现,同时也印证了刘易斯拐点理论 (见图4i、图4q)。就业水平X2对县内流动率的作用呈现M型曲线,存在2.3万和4.5万的双重阈值,在一定程度上反映了就业规模初期扩展激活效应、中期饱和抑制、后期超载重构的递进机制 (见图4b)。其对跨县流入率呈现倒L型曲线,县域扩张存在1.5万岗位的抑制阈值,反映了本地劳动力市场对跨县人口吸纳力的刚性边界 (见图4n)。其对跨县流出率作用呈现边际递增饱和效应,存在县域就业岗位规模3.0万的虹吸阈值 (见图4v)。经济发展水平X3与跨县流出率呈非线性正相关,当夜间灯光达到200后其影响趋于稳定 (见图4w)。居民收入水平X4对县内流动率的作用呈现三阶段阈值效应,存在2万和6.2万的双重阈值,反映了迁移能力与机会成本的动态均衡 (见图4d); 农业发展水平X5的作用特征与农业规模经济的倒U型曲线一致,表明适度的农业发展有利于激活本地劳动力市场,但大规模农业生产会导致产业的结构收缩 (见图4e、图4j、图4r)。

在社会环境因素中,交通设施X6对县内流动存在规模经济500的阈值,当超过阈值后处于下降效应,表明交通网络覆盖达到饱和后,新增设施对人口吸引力贡献率显著衰减 (见图4h)。娱乐设施X10对流入流出的作用存在400的规模阈值 (见图4p、图4x)。金融设施X11对县内流动率存在单峰饱和效应,存在300个设施的最优规模阈值,超规模配置服务后会导致人口流动效率衰减 (见图4c)。

在环境因素中,植被覆盖度X12对流动率的调控呈双峰震荡,县内流动在0.50节点抑制性最低 (见图4f),流入在0.65节点抑制性最低 (见图4m),流出在0.47节点抑制性最低 (见图4t),揭示了生态宜居与资源开发约束的权衡阈

值。地形地貌X14对跨县流入率作用呈倒N型 (见图4o),其对跨县流出率作用呈现波浪形震荡 (见图4u),两者都存在4、8双重临界阈值,反映了生态约束和资源补偿机制的博弈。温差X16对于县内流动率的作用呈现U型阈值效应,存在20.5℃的临界阈值,低温阶段抑制人口流动,而高温促使人口流动 (见图4g)。

在区位因素中,市域区位X17和省域区位X18对跨县流入率呈现差异化空间衰减规律,市域区位存在0.15、0.50、1.50、2.00单位距离节点,揭示了近郊通勤成本主导、远郊资源补偿效应,而省域区位作用随距离持续衰减后趋于稳定,反映了区域辐射力的边际递减 (见图4k、图4l)。市域区位X17对于跨县流出率的作用呈现单峰饱和效应,存在1.50单位距离的市域辐射力临界值,反映近郊区人口受中心城区虹吸抑制,而远郊县域因机会缩减与流动成本降低促使人口持续外流 (见图4s)。

综上所述,城镇化水平对人口流动存在边际递减效应,农业发展对人口的推拉是技术和制度的双重作用,生态驱动因子对人口流动呈现双峰作用,体现了土地人口承载的弹性阈值。这些非线性作用规律打破了传统线性模型的静态假说,表明各驱动因子对人口流动存在动态影响,在制定相关政策时,应该建立动态模型,避免一刀切的政策。

2.2.5 驱动因子的空间异质性特征

分别选取3类流动人口中综合重要性排序前8位的驱动因子。研究发现,各类驱动因子对人口结构的调控在空间上存在地理非平稳性、地理集聚与功能分区特征,人口流动是空间异质性和复杂系统动力学协同作用的结果 (见图5)。

就人口县内流动而言,城镇化水平X1、就业水平X2、农业发展水平X5等3个驱动因子在邵阳市市区及益阳—岳阳市都市圈形成双核心集聚模式,揭示了工业化与农业化对人口流动存在协同作用的空间特征 (见图5a、图5b、图5e)。居民收入水平X4和温差X16两个驱动因子在岳阳市呈现独特的单极热点,空间异质性强度显著高于其他区域,表明洞

庭湖平原的气候与经济耦合系统对县内人口流动具有地理敏感性 (见图5d、图5g)。教育设施X8和金融设施X11分别在长沙市和邵阳市形成单极核心,印证了基础设施的枢纽辐射效应 (见图5h、图5c)。植被覆盖度X12在岳阳—新田县形成了双核异质区,反映了武陵—罗霄山脉生态屏障对人口流动产生的地理阻隔效应 (见图5f)。

就人口跨县流入而言,城镇化水平X1、农业发展水平X5、植被覆盖度X12等3个驱动因子的重要性得分在衡阳市形成高值单核,并沿湘江流域呈梯度递减,印证了区域增长极理论中的“中心化—外围扩散”效应 (见图5i、图5j、图5m),其中植被覆盖度X12的生态抑制系数揭示了南岳山脉生态屏障对人口流入的过滤作用。就业水平X2在邵阳市区形成单核心 (见图5n),娱乐设施X10在岳阳形成文化引力核心 (见图5p),符合新经济地理学中“本地市场效应”的空间规律。地形地貌X14在邵阳市、株洲市形成双核心 (见图5o)。省域区位X18以长沙市为引力核心 (见图5l),其辐射范围覆盖长株潭,而市域区位X17在湘西—湘南形成双核 (见图5k),反映了边缘区位的补偿性流动机制。

就人口跨县流出而言,城镇化水平X1、就业水平X2、植被覆盖度X12及地形地貌X14在衡邵娄交界处形成协同驱动核心 (见图5q、图5v、图5t、图5u),各因子之间存在交互影响,表明“自然约束—经济压力”的共同作用是衡邵娄人口外溢的主导原因。经济发展水平X3呈现西南向西北递减趋势 (见图5w),农业发展X5则呈现自东向西递减趋势 (见图5r),揭示了湘西欠发达地区正面临“低发展陷阱”,呈现人口持续外流现状。娱乐设施X10在邵阳市区呈现空间抑制效应 (见图5x),其重要性得分均为全域最低。市域区位X17自西北向东南呈现递减趋势 (见图5s),揭示了地理位置较偏远的西北地区对人口流出的推动作用较强,而东南部的交通和经济条件相对更具吸引力,流出程度较轻,印证了中心地理论中“距离衰减—边际替代”的复合机制。

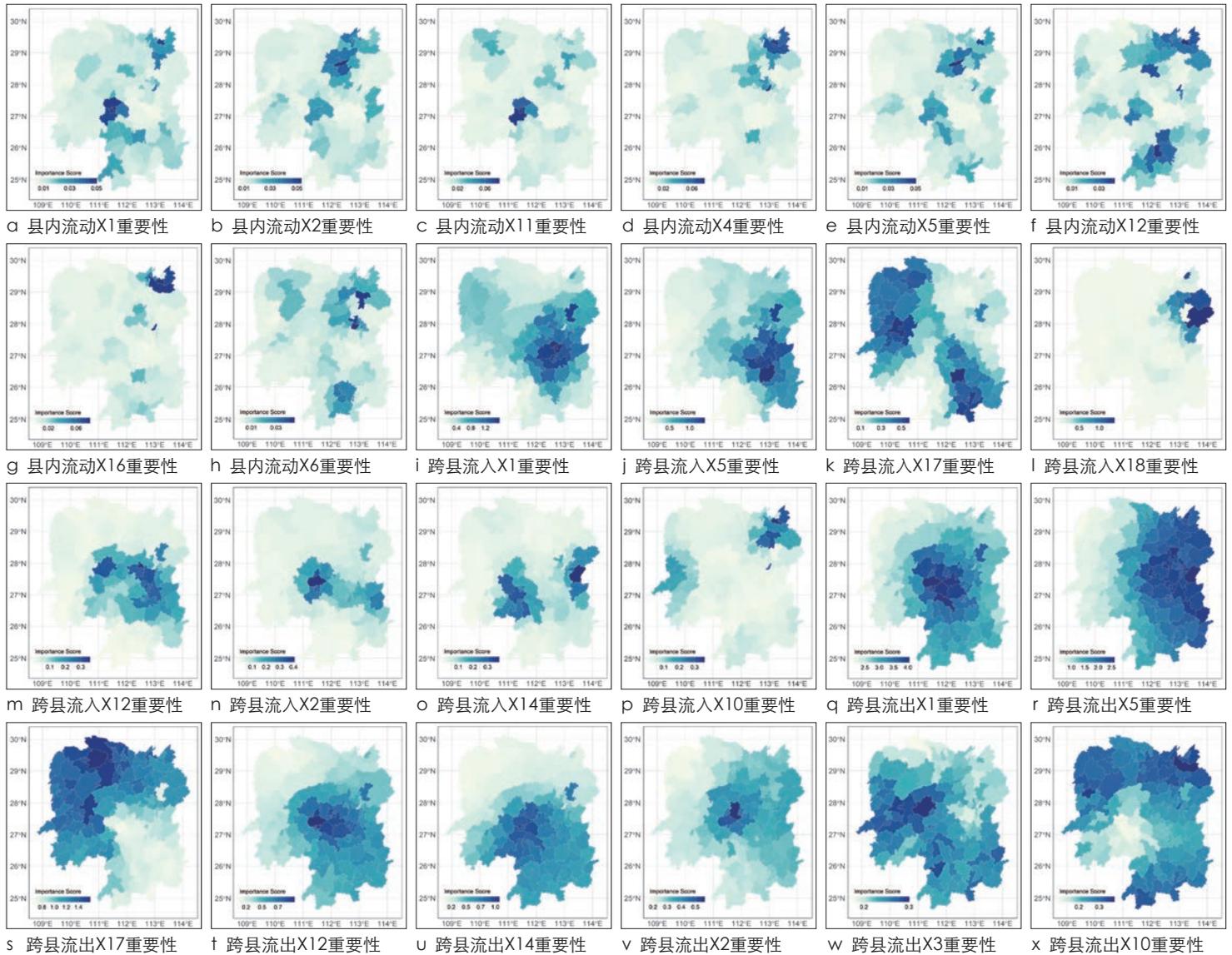


图5 驱动因子空间异质地图
Fig.5 Map of spatial heterogeneity of impact factors

资料来源: 笔者自绘。

综上所述,人口流动的驱动机制由地理要素多尺度耦合构成。其中,经济要素主要通过“核心极化—路径依赖”机制对人口流动进行约束,社会要素主要通过“枢纽辐射—过密陷阱”双重效应对人口流动进行空间重构,生态要素以“地理阻隔—阈值相应”重塑人口流动网络,区位要素对人口驱动规律遵循“边际替代—衰减梯度”效应。对这些要素的驱动规律进行空间异质性研究,突破了传统人口迁移理论的均质化假设,为人口流动的区域化管理提供了差异化视角。

3 结论与讨论

3.1 结论

本文基于RF模型和GWRF模型,以湖南省县域为研究单元,融合2000年、2010年、2020年3年人口普查数据和多元地理大数据,从全局和局部结合视角探究了人口流动驱动因子的非线性与空间异质性,突破了传统全局模型的均质化假设,为县域尺度人口流动的“驱动因子—系统阈值—地理空间”协同解析提供了新的方法论。

主要结论如下:

(1) 湖南省人口流动呈现梯度分异和多中心演化趋势。长沙市中心城区岳麓区、芙蓉区流入率超60%,通过虹吸周边人口形成核心极化区,同时衡阳、岳阳次级中心呈现萌芽趋势;郴州、永州湘南地区流出率超70%,因低发展陷阱人口持续外流;以永定县为代表的县内流动率高达54%的湘西地区,依托生态旅游发展本地化,为边缘县域的人口结构优化提供了转型范本。

(2) 湖南省人口的流动主要受经济要素限制和生态要素约束。城镇发展水平、农业发

展水平与区位条件的全局驱动作用显著;而局部模型显示,湘西地区植被覆盖度和地形地貌等驱动因子的重要性较全局提升40%,洞庭湖平原的气候—经济耦合效应突出。

(3) 流动人口的驱动机制存在明显的阈值效应和边际递减规律。城镇发展水平的作用规律呈现20%、60%双拐点的边际递减效应,就业水平的驱动机制以3万、6万的M形曲线反映了劳动力市场的三阶段演化,交通设施与金融设施分布作用规律分别在500、400节点后大幅度衰减。

(4) 流动人口的驱动在空间上存在明显的“核心极化—边缘收缩”的特征。其中,邵阳—益阳双核心区通过工业—农业协同促使县内流动人口重组,长株潭都市圈依托交通枢纽辐射形成核心极核并向外梯度递减,湘西—湘南边缘县域因交通区位衰减而触发人口补偿性外流。

3.2 讨论

本文引入RF模型和GWR模型结合全局和局部视角,对湖南省县域人口流动的驱动机制进行了非线性阈值与空间异质性特征探究,揭示了“城镇化拐点”“枢纽过密”“生态约束”“边缘收缩”等作用规律。同时,通过构建“系统阈值—地理权重”双维度框架,为以湖南省为代表的中部地区人口流动管理提供了“动态阈值预警—空间靶向施策”的方法论参考。

基于上述研究,提出以下几点政策建议:

(1) 对驱动人口流动的各项因子进行动态阈值调控,将城镇化率60%、就业岗位3万、交通设施500等阈值纳入区域规划的考虑范畴,规避过量资源配置;(2) 针对各项驱动因子进行空间靶向治理,长沙近郊县域需抑制“过密化陷阱”,湘南地区应对产业进行升级突破低发展瓶颈,湘西地区继续发展生态旅游以激活人口本地化潜力;(3) 针对人口流动的整体格局进行分区施策,在邵阳—益阳双核心区推动产城融合,在衡阳—岳阳双次级中心强化文化—生态导向,在湘西—湘南边缘区

县实施流动补偿机制。

本文仍有些问题需进一步深化。首先,存在微观缺失的问题,在驱动因子选取时,忽视了人口的性别、年龄、学历、家庭等方面对人口流动的作用,在未来的进一步研究中可以从微观视角对流动人口进行差异化分析。此外,省内流动与跨省流动的比较也能为研究提供新视角。可以通过对比湖南省与其他省份流动人口空间分布特征及其驱动机制,分析区域之间的异同,为区域协调发展提供参考。

参考文献 References

- [1] 翟振武,金光照. 中国人口负增长:特征、挑战与应对[J]. 人口研究, 2023, 47 (2): 11-20.
ZHAI Zhenwu, JIN Guangzhao. Negative population growth in China: characteristics, challenges, and responses[J]. Population Research, 2023, 47(2): 11-20.
- [2] 刘涛,卓云霞. 中国县级人口变动的空间格局及影响因素——基于第七次全国人口普查数据的新探索[J]. 人口研究, 2022, 46 (6): 72-87.
LIU Tao, ZHUO Yunxia. A county-level analysis of China's population change: insights from the 7th population census[J]. Population Research, 2022, 46(6): 72-87.
- [3] ZELINSKY W. The hypothesis of the mobility transition[J]. Geographical Review, 1971, 61(2): 219-249.
- [4] SKELDON R. Migration and development: a global perspective[M]. London: Routledge, 2014.
- [5] FAN C C. China on the move: migration, the state, and the household[J]. London: Routledge, 2007.
- [6] 刘涛,彭荣熙,卓云霞,等. 2000—2020年中国人口分布格局演变及影响因素[J]. 地理学报, 2022, 77 (2): 381-394.
LIU Tao, PENG Rongxi, ZHUO Yunxia, et al. China's changing population distribution and influencing factors: insights from the 2020 census data[J]. Acta Geographica Sinica, 2022, 77(2): 381-394.
- [7] 侯华丽,冯健,张瑞华. 基于日常人口流动的城市群等级与网络结构——对中国五大城市群的考察[J]. 城市发展研究, 2024, 31 (9): 91-100.
HOU Huali, FENG Jian, ZHANG Ruihua. Hierarchy and network structure of urban agglomerations

- based on daily population mobility: a study of five major urban agglomerations in China[J]. Urban Development Studies, 2024, 31(9): 91-100.
- [8] 邵云通,吴晓. 中国城际人口流动网络的多尺度空间差异——基于腾讯迁徙大数据的分析[J]. 地理科学, 2025, 45 (1): 189-201.
SHAO Yuntong, WU Xiao. Multi-scale spatial differences of intercity population flow network in China: based on Tencent migration big data[J]. Geographical Science, 2025, 45(1): 189-201.
 - [9] 金丹,周婧楠,许尊. 京津冀城市群流动人口空间格局及影响因素研究[J]. 城市发展研究, 2024, 31 (12): 34-44.
JIN Dan, ZHOU Jingnan, XU Zun. Study on the spatial pattern and influencing factors of floating population in the Beijing-Tianjin-Hebei urban agglomeration: analysis based on the sixth national census and the seventh national census[J]. Urban Development Studies, 2024, 31(12): 34-44.
 - [10] 王桂新. 中国省际人口迁移变化特征——基于第七次全国人口普查数据的分析[J]. 中国人口科学, 2022 (3): 2-16.
WANG Guixin. Research on characteristics of China's inter-provincial migration: based on the data of China's seventh population census[J]. Chinese Journal of Population Science, 2022(3): 2-16.
 - [11] 曹广忠,刘涛. 中国城镇化地区贡献的内陆化演变与解释——基于1982—2008年省区数据的分析[J]. 地理学报, 2011, 66 (12): 1631-1643.
CAO Guangzhong, LIU Tao. Rising role of inland regions in China's urbanization in the 21st century: the new trend and its explanation[J]. Acta Geographica Sinica, 2011, 66(12): 1631-1643.
 - [12] 林李月,朱宇,柯文前. 城镇化中后期中国人口迁移流动形式的转变及政策应对[J]. 地理科学进展, 2020, 39 (12): 2054-2067.
LIN Liyue, ZHU Yu, KE Wenqian. Changes in the form of population migration and mobility in China and corresponding policy responses at the late-intermediate stage of urbanization[J]. Progress in Geography, 2020, 39(12): 2054-2067.
 - [13] 吴瑞君,薛琪新,罗志华. 我国人口迁移和城镇化格局的转折性变化:2000—2020年[J]. 上海行政学院学报, 2022, 23 (1): 74-86.
WU Ruijun, XUE Qixin, LUO Zhihua. The turning change of the spatial patterns of population migration and urbanization in China from 2000 to 2020[J]. The Journal of Shanghai Administration Institute, 2022, 23(1): 74-86.
 - [14] 刘艳军,汤爽爽,吴康,等. 经济地理学视角下中国人口研究热点与展望[J]. 经济地理, 2021, 41 (10): 97-105.
LIU Yanjun, TANG Shuangshuang, WU Kang, et al. Hotspots and prospects of population research in China from the perspective of economic geography[J]. Economic Geography, 2021, 41(10): 97-105.
 - [15] 罗小龙,曹姝君,顾宗侃. 回流城镇化:中部地区

- 城镇化开启新路径[J]. 地理科学, 2020, 40 (5): 685-690.
- LUO Xiaolong, CAO Shujun, GU Zongni. The returning-migrant-induced urbanization: the new path of urbanization in middle China[J]. *Geographical Science*, 2020, 40(5): 685-690.
- [16] 王新贤, 高向东. 中国流动人口分布演变及其对城镇化的影响——基于省际、省内流动的对比分析[J]. 地理科学, 2019, 39 (12): 1866-1874.
- WANG Xinxian, GAO Xiangdong. The evolution of China's floating population and its impact on urbanization: a comparative analysis based on inter- and intra-provincial perspectives[J]. *Geographical Science*, 2019, 39(12): 1866-1874.
- [17] 柯文前, 肖宝玉, 林李月, 等. 中国省际城乡流动人口空间格局演变及与区域经济发展的关系[J]. 地理学报, 2023, 78 (8): 2041-2057.
- KE Wenqian, XIAO Baoyu, LIN Liyue, et al. Interprovincial urban and rural floating population evolution of China and its relationship with regional economic development[J]. *Acta Geographica Sinica*, 2023, 78(8): 2041-2057.
- [18] 赖建波, 朱军, 郭煜坤, 等. 中原城市群人口流动空间格局与网络结构韧性分析[J]. 地理与地理信息科学, 2023, 39 (2): 55-63.
- LAI Jianbo, ZHU Jun, GUO Yukun, et al. Spatial pattern of population flow and the resilience of network structure of central plains urban agglomeration[J]. *Geography and Geo-Information Science*, 2023, 39(2): 55-63.
- [19] 曹广忠, 陈思创, 刘涛. 中国五大城市群人口流入的空间模式及变动趋势[J]. 地理学报, 2021, 76 (6): 1334-1349.
- CAO Guangzhong, CHEN Sichuang, LIU Tao. Changing spatial patterns of internal migration to five major urban agglomerations in China[J]. *Acta Geographica Sinica*, 2021, 76(6): 1334-1349.
- [20] 李聪, 宗会明, 肖磊. 中国典型人口流出地区人口流动格局——以川渝地区为例[J]. 热带地理, 2021, 41 (3): 516-527.
- LI Cong, ZONG Huiming, XIAO Lei. Spatial pattern of population flow in China's typical outflow areas: a case study of the Sichuan-Chongqing area[J]. *Tropical Geography*, 2021, 41(3): 516-527.
- [21] SHEN J. Modelling interregional migration in China in 2005–2010: the roles of regional attributes and spatial interaction effects in modelling error[J]. *Population, Space and Place*, 2017, 23(3): e2014.
- [22] GU H. Understanding the migration of highly and less-educated labourers in post-reform China[J]. *Applied Geography*, 2021, 137: 102605.
- [23] STARK O. The migration of labor[M]. Hoboken: Blackwell Publishing, 1991.
- [24] HARRIS J R, TODARO M P. Migration, unemployment and development: a two-sector analysis[J]. *The American Economic Review*, 1970, 60(1): 126-142.
- [25] KRUGMAN P. Increasing returns and economic geography[J]. *Journal of Political Economy*, 1991, 99(3): 483-499.
- [26] BLACK R, ADGER W N, ARNELL N W, et al. The effect of environmental change on human migration[J]. *Global Environmental Change*, 2011, 21: S3-S11.
- [27] MCLEMAN R. International migration and climate adaptation in an era of hardening borders[J]. *Nature Climate Change*, 2019, 9(12): 911-918.
- [28] 王婧, 刘奔腾, 李裕瑞. 京津冀人口时空变化特征及其影响因素[J]. 地理研究, 2018, 37 (9): 1802-1817.
- WANG Jing, LIU Benteng, LI Yurui. Spatial-temporal characteristics and influencing factors of population distribution and floating changes in Beijing-Tianjin-Hebei region[J]. *Geographical Research*, 2018, 37(9): 1802-1817.
- [29] 王婧, 李裕瑞. 中国县城城镇化发展格局及其影响因素——基于2000和2010年全国人口普查分县数据[J]. 地理学报, 2016, 71 (4): 621-636.
- WANG Jing, LI Yurui. Spatial pattern and influencing factors of urbanization development in China at county level: a quantitative analysis based on 2000 and 2010 census data[J]. *Acta Geographica Sinica*, 2016, 71(4): 621-636.
- [30] 田盼盼, 朱宇, 林李月, 等. 省际与省内流动人口空间分布及其影响因素的差异——以福建省为例[J]. 人口学刊, 2015, 37 (6): 56-67.
- TIAN Panpan, ZHU Yu, LIN Liyue, et al. Differences in the spatial distribution and its determinants between inter- and intra-provincial floating population: the case of Fujian Province[J]. *Population Journal*, 2015, 37(6): 56-67.
- [31] 王超, 阚瑗珂, 曾业隆, 等. 基于随机森林模型的西藏人口分布格局及影响因素[J]. 地理学报, 2019, 74 (4): 664-680.
- WANG Chao, KAN Aike, ZENG Yelong, et al. Population distribution pattern and influencing factors in Tibet based on random forest model[J]. *Acta Geographica Sinica*, 2019, 74(4): 664-680.
- [32] GEORGANOS S, GRIPPA T, NIANG GADIAGA A, et al. Geographical random forests: a spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling[J]. *Geocarto International*, 2021, 36(2): 121-136.
- [33] 邓楚雄, 李民, 宾津佑. 湖南省人口分布格局时空变化特征及主要影响因素分析[J]. 经济地理, 2017, 37 (12): 41-48.
- DENG Chuxiong, LI Min, BIN Jinyou. Spatial-temporal variation characteristics and main influence factors analysis of population distribution pattern in Hu'nan Province[J]. *Economic Geography*, 2017, 37(12): 41-48.
- [34] 段成荣, 盛丹阳, 吕利丹, 等. 迁徙中国视野下我国县域人口流动状况与发展挑战[J]. 西北人口, 2022, 43 (6): 1-13.
- DUAN Chengrong, SHENG Danyang, LYU Lidan, et al. Study on migration of China's county and its challenges from the perspective of migrating China[J]. *Northwest Population Journal*, 2022, 43(6): 1-13.
- [35] 曾永明, 钟子康, 刘厚莲. 中国县域流动人口分布样态及其演变——基于柯西分布的模拟及预测[J]. 中国人口科学, 2024, 38 (1): 113-128.
- ZENG Yongming, ZHONG Zikang, LIU Houlian. The pattern and dynamics of county-level population flow in China: simulation and projection using the Cauchy Distribution[J]. *Chinese Journal of Population Science*, 2024, 38(1): 113-128.
- [36] BREIMAN L. Random forests[J]. *Machine Learning*, 2001, 45: 5-32.
- [37] GRÖMPING U. Variable importance assessment in regression: linear regression versus random forest[J]. *The American Statistician*, 2009, 63(4): 308-319.