

基于多源数据的大城市人口数据空间化研究进展

唐子茹¹, 罗瑞麒¹, 蒋文芮¹, 谢高源¹, 庞婧颖², 魏传华¹

¹中央民族大学理学院, 北京

²中央民族大学经济学院, 北京

收稿日期: 2025年3月18日; 录用日期: 2025年4月9日; 发布日期: 2025年4月22日

摘要

人口数据空间化是利用多源数据与建模技术, 将传统的以行政区划为统计单元的人口数据转化为高分辨率格网化分布数据, 以揭示人口的真实空间异质性的技术方法。本文系统性文献回顾了基于多源数据的大城市人口数据空间化研究成果, 重点讨论了过往研究中变量的选择与建模方法的演变, 多维度探析其数据处理、人口空间化模型选择面临的挑战与领域突破的路径构想。通过人口空间化算法的适应性改进与多源数据融合, 可构建更高精度的人口空间化框架, 并为经济数据空间化、灾害风险评估及公共卫生资源配置等提供科学支撑。

关键词

多源数据, 人口数据空间化, 多元线性回归, 机器学习

Advances in Multi-Source Data-Driven Spatialization of Urban Population Distributions

Ziru Tang¹, Ruiqi Luo¹, Wenrui Jiang¹, Gaoyuan Xie¹, Jingying Pang², Chuanhua Wei¹

¹College of Science, Minzu University of China, Beijing

²School of Economics, Minzu University of China, Beijing

Received: Mar. 18th, 2025; accepted: Apr. 9th, 2025; published: Apr. 22nd, 2025

Abstract

Population spatialization is a technical methodology that transforms traditional administrative

文章引用: 唐子茹, 罗瑞麒, 蒋文芮, 谢高源, 庞婧颖, 魏传华. 基于多源数据的大城市人口数据空间化研究进展[J]. 统计学与应用, 2025, 14(4): 169-177. DOI: 10.12677/sa.2025.144098

unit-based demographic statistics into high-resolution gridded distribution data through multi-source data integration and modeling techniques, aiming to reveal the authentic spatial heterogeneity of human populations. This study systematically reviews research advances in megacity population spatialization driven by multi-source datasets, focusing on the evolution of variable selection criteria and modeling methodologies. Through multidimensional analysis, it critically examines persistent challenges in data processing and model optimization, while proposing innovative pathways for disciplinary breakthroughs. The findings suggest that adaptive refinement of spatialization algorithms combined with cross-domain data fusion could establish a higher-accuracy population mapping framework, which may provide scientific support for economic data spatialization, disaster risk quantification, and public health resource allocation.

Keywords

Multi-Source Data, Population Data Spatialization, Multiple Linear Regression, Machine Learning

Copyright © 2025 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

随着全球城镇化进程加速,人口空间分布的动态监测与精准预测成为优化城市治理的核心议题。2025年3月5日,国务院总理李强在《政府工作报告》中提出:“2025 年要推进新型城镇化和区域协调发展,持续推进城市更新和城镇老旧小区改造,统筹城市低效用地再开发,加快健全城市防洪排涝体系,加强燃气、给排水、热力、地下管廊等建设和协同管理。”中国作为人口流动规模最大的国家之一,了解人口分布的发展变化对城镇规划、灾害风险评估与公共卫生资源配置等具有重要作用。常见的人口分布数据——传统人口普查数据通常以乡镇为最小统计单元,难以捕捉城市内部街区尺度的人口密度差异,且存在更新滞后等问题,严重制约精细化的城市管理。因此,高分辨率人口栅格数据的建模研究已成为人口空间异质性分析的关键需求,驱动了多源数据融合框架下的深度探索。

如今,学界围绕人口数据空间化的研究现状如下:第一,自变量选择方面,多源数据融合[1]已成为显著趋势,土地利用类型数据、夜间灯光数据、DEM 数据和 POI 数据等数据的融合有效提升模型精度;第二,因变量优化方面,手机信令数据[2]、微信定位数据[3]等新型数据源的引入使得研究的精度大幅提升;第三,研究主题集中于技术创新,建模方法从早期的人口密度模型和空间插值模型演变为使用统计回归模型,包括多元线性回归模型、地理加权回归模型、空间回归模型、多因素融合模型等,并逐步发展到机器学习模型和深度学习模型的应用。

本文对人口数据空间化的变量选择、建模方法进行全面梳理,总结现今存在的研究困难与解决方法展望,旨在进一步优化人口数据空间化方法,以支持城市规划、公共卫生应急及灾害风险评估,从而为人口地理学提供方法论创新,亦为智慧城市治理与可持续发展目标的实现提供科学工具。

2. 人口空间化研究中的变量选择

人口空间化研究通过建模变量将行政单元人口统计数据转化为地理空间分布数据,其核心在于准确反映人口分布的真实格局。变量选择直接影响模型精度与适用性,是研究的关键环节。传统方法多依赖单一变量,因覆盖影响因素不足导致预测准确率较低,难以刻画复杂异质性。近年来,手机信令、POI 等多源数据的广泛应用显著提升研究数据的丰富度与精度,不同类型数据可互补分辨率与准确性

缺陷。部分学者[4]-[7]基于多源数据开展了省市级人口空间化研究或分析人口时空分布特征[8]，但变量选择仍缺乏统一标准，导致模型结果差异显著。本文通过系统梳理现有研究，归纳变量选择中的核心问题，探讨不同数据类型的适用场景及其对模型精度的影响机制，为优化人口空间化模型构建提供理论参考。

2.1. 自变量的选择

影响人口分布的因素可大致分为自然因素和社会经济因素两类。现有研究常通过土地利用类型数据、夜间灯光数据、DEM 数据和 POI 数据等多源数据综合反映这些因素。

土地利用类型数据通过多级分类反映地表自然状况。现有研究中，研究者常依据研究地区的土地利用特点来决定建模因子，选取方法主要分为两种：其一以用地类型的人口密集度为标准，合并较小，保留较大，并按二级地类分类[5]。其二将数据按照一级地类汇总作为建模因子[5][9]。通过总结现有研究，在建设用地类型对人口分布影响显著的城市地区可考虑使用前者，当研究区域为自然条件主导的省级或更大区域面积时，更适宜选择后者进行研究。通过针对性选择作为建模因子的地类层级，可在一定程度上提升建模精度。影响人口分布的自然因素还包括坡度，通过 DEM 高程数据予以表征，以其高精度、广覆盖和多维度分析的优势，精准反映地表高程变化。

夜间灯光数据可直观反映夜间人类活动情况。现有常用的数据源有 DMSP-OLS、NPP/VIIRS [5]、珞珈一号数据[10]三种，其中珞珈一号数据精度最高，所需预处理步骤最少。通过对夜间灯光数据进行重采样、噪声去除等预处理，达到反映区域人口分布的目的。冯磊等[11]研究发现，将夜间灯光数据与卫星遥感数据、土地利用数据等综合使用，能显著提高人口空间分布研究的精度，更准确体现影响人口分布的社会经济因素。

POI 数据是新兴的地理信息数据类型，包含商业设施、公共服务、交通枢纽等人类活动场所的位置信息，可直观反映人类活动的空间分布特征，为人口分布与流动研究提供关键支持。通过与其他空间数据[4][5]结合，能更精准揭示社会经济因素对人口分布的影响。

综上所述，本研究总结出不同自变量的数据源、获取方式以及在人口分布研究的作用，见表 1。

Table 1. The selection of independent variables

表 1. 自变量的选择

自变量数据源	获取方式	作用
土地利用类型数据	OSM (OpenStreetMap)官网	反映地表自然状况
夜间灯光数据	珞珈一号数据	反映夜间人类活动情况
POI 数据	高德地图 POI 数据	反映人类活动的空间分布特征

2.2. 因变量的选择

本文基于现有文献整理发现，人口空间化研究中的因变量为人口相关数据，主要有三种类型：各省、市统计年鉴中所给出的数据，可供开源下载的人口格网数据集，以及含有位置信息的人口数据。

统计年鉴中的人口数据易于获得，且较为精准。在现有研究中，学者通常从统计年鉴中取得数据[11]-[13]，后结合行政区划数据、道路数据[10]，对研究区域制作渔网化格网，求得单个格网中的人口密度[4]，作为因变量参与后续建模。然而，这类人口数据精度较低，仅能细化到市、县级，同时更新频率低，最新数据尚未发布只能根据前一年数据及人口增速估算[10]，可能导致结果偏差。

人口格网数据集数据来源多、更新率快、分辨率较高，支持多重尺度分析。其中，WorldPop 数据集

最为常用[4] [9]，能够提供多尺度数据且无需预处理人口密度。但该数据仍存在误差较大的问题，需要采用多数据集交叉验证[4]。

部分学者[6]-[8]采用手机信令数据、腾讯位置大数据等含有位置信息的人口数据，其优势在于高取样率、广覆盖范围和全天活动记录能力，有效反映人口时空分布规律。但这类数据存在获取成本高，需从不同渠道购买，以及信号漂移、数据重复等异常问题，可能影响分析结果的可靠性。

以上研究发现，人口空间化研究中的因变量选择需根据研究需求权衡。统计年鉴数据易获取且权威，但分辨率低、更新慢，适用于宏观分析；人口格网数据集(如 WorldPop)提供多尺度、高分辨率数据，更新较快，但存在误差，需多源数据校正。对于高精度或时空动态分析，含有位置信息的数据更具优势，但成本高、存在信号漂移、数据重复等问题。未来，多源数据融合与技术进步有望提升研究精度与应用范围。

综上所述，本研究总结出因变量的数据源、获取方式以及在人口分布研究的优势和劣势，见表 2。

Table 2. The selection of dependent variable data
表 2. 因变量数据的选择

因变量数据源	获取方式	优势	劣势
统计年鉴人口数据	各级行政单位统计年鉴	易于获取，较为精准	精度较低，更新频率低
人口格网数据集	WorldPop 数据集	来源多、更新率快、分辨率高	误差较大
手机信令数据、腾讯位置大数据	通过专业渠道购买	取样率高，含位置信息	成本高、存在信号漂移、数据重复等问题

2.3. 数据预处理

现有研究中数据预处理主要包含统一坐标系、重采样、分辨率标准化、格网化等常规流程。针对不同数据类型需进一步实施个性化处理：夜间灯光数据需根据像元特征采用差异化降噪方法提升精度[14]；POI 数据则通过投影变换、去重清洗、分类存储[4]三步优化数据结构以适配格网模型。为增强模型精度，常基于夜间灯光[14]、GDP [13]、土地利用类型[6]等特征进行分区处理，将自然与经济条件相似的区域划分为同一单元，通过分区回归减少拟合误差，提升预测结果的准确性与可靠性。

3. 基于多源数据的人口数据空间化主要方法

随着遥感和地理信息系统(GIS)技术的发展，人口数据空间化的建模方法和数据也在不断变化和丰富。建模方法从早期基于单一数据的人口密度模型和空间插值模型逐渐演变为使用基于多源大数据的统计回归模型，包括多元线性回归、地理加权回归、空间回归模型(空间滞后模型和空间误差模型)，并逐步发展到基于多源数据的机器学习模型和深度学习模型的应用。

3.1. 多因素融合思想

人口数据空间化中的多因素融合思想是将各种数据来源和技术结合起来，精确绘制人口分布图。这种思想的空间化策略是将道路、坡度、土地利用、夜间灯光以及城市建筑物等信息加以融合。针对所有影响指标开展相关性分析，选择恰当因子用以构建影响人口空间分布的多源融合模型。同时确定各因子对人口空间分布影响的权重数值，最终借助加权求和达成人口在格网单元的空间分布呈现。图 1 展示了多因素融合思想的基本技术路线。

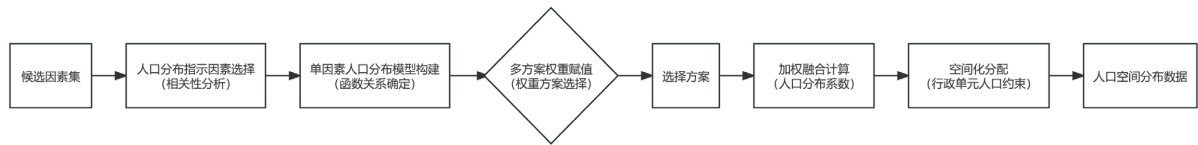


Figure 1. The fundamental concept of multi-factor integration

图 1. 多因素融合思想的基本思路

该方法最具代表性的研究为 Bhaduri [15] 等人开发的人口分布数据库 LandScan USA 模型, 这是一种多维的密度制图建模方法, 能创建高分辨率的人口时空分布数据。该模型以美国人口普查数据为核心, 结合多源辅助数据, 通过分配权重的方式将人口数据细化到更精细的网格单元, 分别构建夜间和白天的的人口分布模型, 如王珂靖、蔡红艳[16]等结合 DMSP/OLS 夜间灯光数据对城镇用地进行再分类, 在人口特征分区的基础上构建人口数据空间化模型。

多因素融合模型综合考虑多源数据对人口分布的影响, 因其对多源数据的适配性和整合能力, 在现阶段被广泛应用于人口数据空间化建模。

3.2. 主要模型分析

3.2.1. 多元线性回归模型

多元线性回归模型在人口数据空间化研究中主要用于构建人口与土地利用/覆盖(夜间灯光)等蕴含人口分布信息的空间数据的关系模型。Lo [17] 在研究亚特兰大市人口估计时, 构建基于不同土地利用类型的多元线性回归模型, 将高密度城市用地、低密度城市用地、农田和森林作为自变量, 人口数量作为因变量, 建立四分类 OLS 回归模型。

在此基础上, 杨晓荣和陈楠[18]将夜间灯光、DEM、路网等多源数据与人口统计数据建立多元回归关系。多元线性回归模型在人口数据空间化的应用假定同一分区下同一土地利用类型人口密度相同, 遵从“无土地无人口”原则[19]。该方法的优点是原理相对直观, 可利用多种数据进行预测和模拟, 但由于线性假设的局限性, 无法准确描述人口分布与各影响因素之间复杂的非线性关系。同时人口数据通常具有空间自相关性, 即相邻空间位置的人口数据存在相互影响, 多元线性回归方法通常难以直接处理。

3.2.2. 地理加权回归模型

地理加权回归(GWR)的最初概念出现于 20 世纪 90 年代中期, Brunsdon 和 Fotheringham 等[20] [21] 分别介绍了地理加权回归作为一种解决回归分析中空间非平稳性的方法, 是对传统回归方法在处理空间数据方面的局限性的补充。地理加权回归通过为数据集中的每个位置构建单独的回归模型实现预测, 每个模型都包含目标位置周围定义的邻域或带宽内的特征的依赖变量和解释变量。观测值的影响随着与目标位置的距离而减小, 使用高斯函数或 bis-squared 等核函数来确定权重, 使得地理加权回归模型可以捕获在全局模型被掩盖的关系中的局部变化。

在此基础上, Yu [22] 将地理加权回归与面板数据分析相结合, 开发了地理加权面板回归(GPWR)。该方法利用时间序列数据增强时间维度, 具有更高的估计效率。Fotheringham 等[23]在 2017 年首次提出多尺度地理加权回归(MGWR)的概念, 多尺度地理加权回归是地理加权回归框架的扩展, 允许每个解释变量在不同的空间尺度上运行, 能够更准确地估计局部系数和减少多重共线性问题。魏传华等[22]学者构建了空间变系数地理加权自回归模型, 综合考虑空间自相关性和空间异质性, 改进了传统模型仅单独考虑这两种特性的局限。

3.2.3. 空间回归模型

空间回归模型包括空间滞后模型(SLM)和空间误差模型(SEM), 用于分析和理解数据中的空间模式,

已被广泛应用于各个领域。空间计量经济学的早期研究为空间回归模型奠定理论基础,突出了空间依赖性在数据分析中的重要性[24]。与传统回归模型的独立性假设相悖,空间滞后模型和空间误差模型都假设接近的观测值往往具有相近的值,即空间自相关。

该方法的代表性研究有 Chi 和 Zhu [25]利用空间回归模型,以威斯康星州 1990~2000 年人口变化数据为例进行分析。Sannigrahi 和 Pilla [26]利用空间回归模型,空间滞后模型(SLM)、空间误差模型(SEM)等方法,对 31 个欧洲国家的 COVID-19 病例和死亡数据以及社会人口数据进行分析。Teng 和 Ding [27]利用 MCMCINLA 估计方法,结合空间滞后模型和变系数空间滞后模型,对 2015~2020 年中国大陆 31 个地区的社会经济数据进行分析。空间回归模型考虑空间相关性,在提高估计精度、揭示空间异质性、提供空间预测并结合多源数据方面具有独特优势。

3.2.4. 机器学习模型

近年来,机器学习(ML)和深度学习(DL)在人口数据空间化的应用取得显著进展,以提高人口制图的准确性和效率。Miao He、Yongming Xu [28]等人利用多源遥感数据和镇级人口普查数据,确定随机森林算法在人口估计中精度最高,并运用该算法进行北京市人口数据空间化研究。Nikparvar 和 Thill [29]强调理解空间属性在机器学习应用中的重要性,讨论空间依赖性、异质性和尺度对模型性能的影响。Zhang 和 Rodavia (2023) [30]使用卷积神经网络(CNN)来改进人口空间分布模型,强调了卷积神经网络在从复杂数据集提取相关空间特征方面的作用。机器学习模型在人口数据空间化中的算法灵活多样,具有强大的非线性建模和数据适应性,对特征提取与选择的可控性使其实现高精度的预测和模拟。

3.3. 主要模型对比分析

多元线性回归原理直观,可利用多源数据预测模拟,具有显著的可解释性,能够通过检验回归系数来了解每个变量对人口分布的影响,这些系数进行归一化以确定哪些自变量对人口密度的影响最大,从而为人口模式的驱动因素提供参考。但其存在线性假设局限性,无法处理人口分布与影响因素间的非线性关系,该模型假设整个研究区域的变量之间的关系恒定,无法解释空间一致性,也难以处理人口数据的空间自相关性。

地理加权回归及其衍生方法(如 GPWR、MGWR)能解决回归分析中的空间非平稳性问题,通过为每个位置构建单独模型捕获局部变化,捕捉人口分布格局的空间异质性。GPWR 增强时间维度, MGWR 可让解释变量在不同空间尺度运行,减少多重共线性。但这类方法相对复杂,计算量较大,且对空间权值的规定很敏感,而带宽的选择(决定空间影响的程度)会显著影响结果。不适当的带宽选择可能导致平滑过度(缺少局部变化)或平滑不足(引入过多噪声)。此外,当模型包含许多具有空间自相关的预测因子时,地理加权回归模型可能受到局部变量之间多重共线性的影响。

空间回归模型(SLM 和 SEM)能分析数据空间模式,考虑空间自相关,有效捕获人口数据的空间自相关性,且考虑了传统回归模型忽略的邻近区域的影响,提供对人口分布的空间扩散过程和空间外部性的思路。然而空间回归模型引入较传统回归模型更大的复杂性,需要仔细指定定义邻域关系的空间权重矩阵,不恰当的空间权重矩阵可能导致空间关系的错误表示及结果偏差。

机器学习这类模型可以捕捉到空间变量和人口密度间复杂的非线性关系,提供内置的特征重要性评估,能够确定空间变量对人口分布模式的影响程度。它能处理复杂数据,在人口估计和改进人口空间分布模型上表现出色,提高人口制图的准确性和效率。但这类模型计算量大,需要更多处理资源,模型可解释性相对较差,可能无法解释驱动人口分布的潜在因素,对空间自相关的描述也不够完善。

综上所述,本研究总结出各个模型的基本原理以及其应用在人口分布研究上的优点和局限性,见表 3。

Table 3. Comprehensive comparison of major models in population spatialization
表 3. 人口数据空间化主要模型综合比较

模型名称	原理	优点	局限性
多元线性回归	构建人口与土地利用/覆盖(夜间灯光)等蕴含人口分布信息的空间数据的关系模型，遵从“无土地则无人口”原则	原理直观，可解释性强，根据回归模型可判断自变量对人口数据的影响程度	无法解释空间异质性和空间自相关问题，不能有效捕捉控件变量于人口密度间非线性关系
地理加权回归	考虑变量之间的局部关系，定义系数在空间上变化来扩展传统回归，捕捉不同地区人口密度及其决定因素之间关系的空间非平稳性	可解决回归分析中的空间非平稳性问题，捕捉人口分布格局的空间异质性	计算量较大，对带宽的选择和空间权值规定敏感，对于包含许多具有空间自相关的预测因子时，可能受局部变量间多重共线性的影响
空间滞后模型	结合空间滞后的因变量作为额外的预测因子，考虑邻近区域对所研究区域的影响，使用空间权重矩阵来定义不同区域之间的关系	有效地考虑人口数据的空间自相关性及传统回归模型忽略的邻近区域的影响	不恰当的空间权重矩阵可能导致空间关系的错误表示并预测错误
随机森林模型	使用 bagging (bootstrap aggregation)和特征随机性来创建不同的决策树，结合多个决策树对其预测进行聚合。	有效刻画空间变量和人口数据之间复杂的非线性关系，根据其内置的特征重要性评估可判断空间变量对人口分布模式的影响成高度	可解释性差、不能直接解释变量间具体关系，对调整参数要求高，无法完全描述空间自相关

4. 基于多源数据的人口数据空间化研究现状

党的二十大报告明确提出优化国土空间布局、推进新型城镇化和区域协调发展的战略目标。人口数据空间化研究通过精准刻画人口分布动态，为资源规划和政策制定提供科学依据。本文系统梳理了当前基于多源数据的人口空间化研究进展，从变量选择、空间化方法两个方面得出以下结论。

人口数据空间化的数据和支撑方法不断变化和丰富，在变量选择方面，多源数据融合已成为显著趋势。人口数据空间化的数据来源从最早的基于单一变量进行建模发展到基于土地利用/覆盖数据、人口密度数据、夜间灯光数据、POI 数据等多源数据进行建模。在空间化方法方面，多因素融合的空间化方法成为显著趋势。人口数据空间化方法由最早的人口密度模型和空间插值方法演变为基于多源数据的多元线性回归模型、地理加权模型、空间滞后模型和机器学习模型及其衍生模型的应用。

在当前多源数据融合背景下，人口数据空间化研究面临的挑战一是多源数据存在噪声、格式异构等问题，需通过标准化处理提升数据可靠性；二是大尺度研究的数据获取与计算成本较高，现有模型在省级以上范围应用时效率有所下降；三是数据异质性导致模型泛化能力受限，不同区域需针对性调整参数组合。

因此，关于人口数据空间化的研究，未来学者可以考虑从以下三个方面进行。其一，引入新型数据源(如移动基站数据、社交媒体数据等)以拓展人口刻画维度，进一步研究人口的时空分布；其二，进行多因素多模型的综合应用，本文提到的常用建模方法各有其优缺点，单一模型难以精确刻画人口空间分布。如何综合利用多种模型，综合考虑各方面优势是当前研究的难点；其三，高分辨率人口数据空间化研究在实际应用场景广泛，如何将人口数据空间化经度提高到单个建筑物尺度，提高人口数据空间化精度是未来的研究方向。

基金项目

大学生创新创业训练计划项目(编号: URTP2024110097)。

参考文献

- [1] 王熙. 基于多源数据并顾及空间差异的北京市人口空间化方法研究[D]: [硕士学位论文]. 北京: 中国测绘科学研究院, 2021.
- [2] 陆俊贤, 付茜子, 覃立伟, 李俺义, 陈静仪, 牛强. 疫情前后武汉市职住人口分布对比分析及片区韧性模式探索[J]. 山西建筑, 2025, 51(4): 1-5, 20.
- [3] 王熙, 宁晓刚, 张翰超, 王浩, 郝铭辉. 融合 LJ1-01 夜间灯光和微信定位数据的人口空间化——以北京市为例[J]. 测绘科学, 2022, 47(2): 173-183.
- [4] 邹雅婧. 多源数据支持下的人口空间化研究[D]: [硕士学位论文]. 北京: 中国矿业大学, 2020.
- [5] 刘美君. 基于多源数据的辽宁省人口空间化研究[D]: [硕士学位论文]. 阜新: 辽宁工程技术大学, 2023.
- [6] 葛熙豪. 基于多源数据的青岛市人口数据空间化方法及其应用[D]: [硕士学位论文]. 北京: 中国矿业大学, 2023.
- [7] 孙铭悦. 基于多源数据的人口空间化方法研究[D]: [硕士学位论文]. 北京: 北京建筑大学, 2021.
- [8] 王梅红, 贺风, 司连法, 等. 精细地图融合手机信令的城市人口时空分布特征研究[J]. 测绘地理信息, 2023, 48(1): 49-52.
- [9] 郭子翰, 陈斐, 刘晓丽, 等. 基于土地利用类型提高人口格网数据空间精度的方法[J]. 测绘通报, 2021(12): 66-70.
- [10] 檀成龙. 基于夜光遥感的人口空间化估算[J]. 测绘与空间地理信息, 2022, 45(0): 103-105, 109.
- [11] 冯磊, 刘根源, 孙康. 基于夜间灯光遥感的人口空间化研究现状及热点[J]. 科技创新与应用, 2024, 14(31): 84-87.
- [12] 龚亚西, 储云, 程珊珊. 基于夜间灯光数据和空间回归模型的人口格网化方法研究[J]. 金陵科技学院学报, 2024, 40(1): 1-12.
- [13] 潘颖, 黄万里, 叶士琳, 等. 基于土地利用类型数据的统计人口空间化模型研究: 以福建省为例[J]. 云南大学学报(自然科学版), 2020, 42(4): 701-711.
- [14] 肖东升, 杨松. 基于夜间灯光数据的人口空间分布研究综述[J]. 国土资源遥感, 2019, 31(3): 10-19.
- [15] Bhaduri, B., Bright, E., Coleman, P. and Urban, M.L. (2007) Landsat USA: A High-Resolution Geospatial and Temporal Modeling Approach for Population Distribution and Dynamics. *GeoJournal*, 69, 103-117. <https://doi.org/10.1007/s10708-007-9105-9>
- [16] 王珂靖, 蔡红艳, 杨小唤. 多元统计回归及地理加权回归方法在多尺度人口空间化研究中的应用[J]. 地理科学进展, 2016, 35(12): 1494-1505.
- [17] Lo, C.P. (2008) Population Estimation Using Geographically Weighted Regression. *GIScience & Remote Sensing*, 45, 131-148. <https://doi.org/10.2747/1548-1603.45.2.131>
- [18] 杨晓荣, 陈楠. 基于多源数据的福建省人口数据空间化研究[J]. 贵州大学学报: 自然科学版, 2019, 36(2): 79-84, 95.
- [19] 董南, 杨小唤, 蔡红艳. 人口数据空间化研究进展[J]. 地球信息科学学报, 2016, 18(10): 1295-1304.
- [20] Brunson, C., Stewart Fotheringham, A. and Charlton, M.E. (1996) Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28, 281-298.
- [21] Fotheringham, A.S. (1997) Trends in Quantitative Methods I: Stressing the Local. *Progress in Human Geography*, 21, 88-96. <https://doi.org/10.1191/030913297676693207>
- [22] Yu, D.L. (2010) Exploring Spatiotemporally Varying Regressed Relationships: The Geographically Weighted Panel Regression Analysis. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38, 134-139.
- [23] Fotheringham, A.S., Yang, W. and Kang, W. (2017) Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, 107, 1247-1265. <https://doi.org/10.1080/24694452.2017.1352480>
- [24] LeSage, J.P. (2014) What Regional Scientists Need to Know about Spatial Econometrics. *SSRN Electronic Journal*, 44, 13-32. <https://doi.org/10.2139/ssrn.2420725>
- [25] Chi, G. and Zhu, J. (2007) Spatial Regression Models for Demographic Analysis. *Population Research and Policy Review*, 27, 17-42. <https://doi.org/10.1007/s11113-007-9051-8>

-
- [26] Sannigrahi, S., Pilla, F., Basu, B., Basu, A.S. and Molter, A. (2020) Examining the Association between Socio-Demographic Composition and COVID-19 Fatalities in the European Region Using Spatial Regression Approach. *Sustainable Cities and Society*, 62, Article ID: 102418. <https://doi.org/10.1016/j.scs.2020.102418>
- [27] Teng, J., Ding, S., Zhang, H. and Hu, X. (2023) MCMCINLA Estimation of Varying Coefficient Spatial Lag Model—A Study of China's Economic Development in the Context of Population Aging. *PLOS ONE*, 18, e0279504. <https://doi.org/10.1371/journal.pone.0279504>
- [28] He, M., Xu, Y. and Li, N. (2020) Population Spatialization in Beijing City Based on Machine Learning and Multisource Remote Sensing Data. *Remote Sensing*, 12, Article 1910. <https://doi.org/10.3390/rs12121910>
- [29] Nikparvar, B. and Thill, J. (2021) Machine Learning of Spatial Data. *ISPRS International Journal of Geo-Information*, 10, Article 600. <https://doi.org/10.3390/ijgi10090600>
- [30] Zhang, X. and Rodavia, M.R. (2023) Population Spatialization Based on Convolutional Neural Network and Multi-Source Geospatial Big Data. *Academic Journal of Science and Technology*, 7, 213-216. <https://doi.org/10.54097/ajst.v7i2.12276>