

海南岛人口空间化数据集(2013–2018)

赖先龙, 韩念龙*, 黄鹏锦

海南大学公共管理学院土地资源管理系, 海口 570228

摘要: 本数据集基于 2013–2018 年 NPP/VIIRS 夜间灯光数据, 进行影像合成、不稳定光源去除以及逐年校正等处理, 消除短暂性光源和背景噪声的影响。将已处理的夜间灯光数据与人口统计数据相关性分析建立回归模型, 按照误差分类进行人口空间化建模, 以消除山地地形、人口数量少等因素对模拟精度的影响, 最终生成海南岛人口空间分布数据集。该数据集以.tif 格式存储, 空间分辨率为 500 m, 由 26 个数据文件组成, 11.0 MB (压缩为 1 个文件, 168 KB), 反映 2013–2018 年海南岛人口时空分布变化特征。

关键词: 海南岛; 人口; NPP/VIIRS 夜间灯光数据; 回归模型; 人口空间化

DOI: <https://doi.org/10.3974/geodp.2022.01.08>

CSTR: <https://cstr.escience.org.cn/CSTR:20146.14.2022.01.08>

数据可用性声明:

本文关联实体数据集已在《全球变化数据仓储电子杂志 (中英文)》出版, 可获取:

<https://doi.org/10.3974/geodb.2021.08.02.V1> 或 <https://cstr.escience.org.cn/CSTR:20146.11.2021.08.02.V1>.

1 前言

夜光遥感影像能够为国民生产总值、人口、电力消费、温室气体排放、贫困指数、基尼系数等社会经济参数的估算提供重要依据^[1]。目前普遍使用的夜间灯光数据为 DMSP/OLS 夜间灯光数据和 NPP/VIIRS 夜间灯光数据, 均来源于美国国家海洋与大气管理局, 两者的区别在于 DMSP/OLS 夜间灯光遥感影像的时间序列为 1992–2013 年, 空间分辨率为 1 km×1 km, 而 NPP/VIIRS 夜间灯光遥感影像的时间序列为 2012 年至今, 空间分辨率为 500 m。同时, NPP/VIIRS 夜间灯光数据由 Suomi-NPP 卫星利用 VIIRS 拍摄, 采用极地轨道, 由多幅无云影像拼接得到^[2], NPP–VIIRS 传感器有 22 个波段, 波长范围在 0.4–12 μm, 涵盖可见光和红外光谱区间, 光谱分辨率为 16 bit, 空间分辨率为 500 m, 更宽波段的辐射探测计和在轨辐射校正技术有效地提高了灯光影像的质量^[3]。NPP/VIIRS 夜间灯光数据得益于其更强的灯光捕获敏感度、更高的空间分辨率与时间分辨率, 更适合中小范围的人类活动研究^[4], 基于此提取的人口空间分布信息能够直观地反映人口的时空演变特征。因此本

收稿日期: 2021-08-26; 修订日期: 2021-11-17; 出版日期: 2022-03-25

基金项目: 海南省 (HNSK(ZD)19-119); 海南省自然科学基金 (2019RCO16); 海南大学 (kyqdsx1962)

*通讯作者: 韩念龙, 海南大学, nlhan@hainanu.edu.cn

数据引用方式: [1] 赖先龙, 韩念龙, 黄鹏锦. 海南岛人口空间化数据集(2013–2018)[J]. 全球变化数据学报, 2022, 6(1): 58–64. <https://doi.org/10.3974/geodp.2022.01.08>. <https://cstr.escience.org.cn/CSTR:20146.14.2022.01.08>.

[2] 赖先龙, 韩念龙, 黄鹏锦. 海南岛夜间灯光与统计数据融合的人口 500-m 栅格数据集 (2013–2018) [J/DB/OL]. 全球变化数据仓储电子杂志, 2021. <https://doi.org/10.3974/geodb.2021.08.02.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2021.08.02.V1>.

数据集基于 NPP/VIIRS 夜间灯光数据进行处理，并与海南岛 18 个市县的人口统计数据进行分析相关性分析回归建模得到海南岛人口空间分布信息，以期为海南省国土空间规划或科学研究提供数据基础及决策依据。

2 数据集元数据简介

《海南岛夜间灯光与统计数据融合的人口 500-m 栅格数据集 (2013–2018)》^[5]的名称、作者、地理区域、数据年代、时间分辨率、空间分辨率、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表 1。

表 1 《海南岛夜间灯光与统计数据融合的人口 500-m 栅格数据集 (2013–2018)》元数据简表

条 目	描 述
数据集名称	海南岛夜间灯光与统计数据融合的人口 500-m 栅格数据集 (2013–2018)
数据集短名	PopulationHaiNan_2013-2018
作者信息	赖先龙, 海南大学, 1309842628@qq.com 韩念龙, 海南大学, nlhan@hainanu.edu.cn 黄鹏锦, 海南大学, 1538541468@qq.com
地理区域	海南岛
数据年代	2013–2018
时间分辨率	逐年
空间分辨率	500 m
数据格式	.tif
数据量	11.0 MB (压缩后 168 KB)
数据集组成	本数据集由海南岛 2013–2018 年夜间灯光与统计数据融合的人口 500-m 栅格数据集和海南岛 2013–2018 年间增长人口空间分布 500-m 栅格数据组成
基金项目	海南省社科重大项目 (HNSK(ZD)19-119); 海南省自然科学基金 (2019RCO16); 海南大学 (kyqdsx1962)
数据计算环境	ArcGIS
出版与共享服务平台	全球变化科学研究数据出版系统 http://www.geodoi.ac.cn
地址	北京市朝阳区大屯路甲 11 号 100101, 中国科学院地理科学与资源研究所
数据共享政策	全球变化科学研究数据出版系统的“数据”包括元数据 (中英文)、通过《全球变化数据仓储电子杂志 (中英文)》发表的实体数据集和通过《全球变化数据学报 (中英文)》发表的数据论文。其共享政策如下: (1) “数据”以最便利的方式通过互联网系统免费向全社会开放, 用户免费浏览、免费下载; (2) 最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源; (3) 增值服务用户或以任何形式散发和传播 (包括通过计算机服务器) “数据”的用户需要与《全球变化数据学报 (中英文)》编辑部签署书面协议, 获得许可; (4) 摘取“数据”中的部分记录创作新数据的作者需要遵循 10% 引用原则, 即从本数据集中摘取的数据记录少于新数据集总记录量的 10%, 同时需要对摘取的数据记录标注数据来源 ^[6]
数据和论文检索系统	DOI, CSTR, Crossref, DCI, CSD, CNKI, SciEngine, WDS/ISC, GEOSS

3 数据研发方法

3.1 研究区概况及数据来源

本文研究区域为海南岛陆域范围 (不含三沙市) 的 18 个市县, 分别是海口市、三亚市

和儋州市 3 个地级市以及 15 个省直辖县。海南岛位于东经 108°37′–111°03′，北纬 18°10′–20°10′之间，全岛地势四周低平，中间高耸，呈穹隆山地形。

本数据集的原始夜间灯光数据主要是 2013–2018 年间的 NPP/VIIRS 夜间灯光遥感月度合成均值影像，来源于美国国家海洋和大气管理局（NOAA）中的 NPP/VIIRS 传感器数据集¹，分辨率为 500 m；海南岛行政区划数据来源于中国自然资源数据库²；海南岛的人口统计数据来自海南省统计年鉴^[7]。

3.2 技术路线

本研究技术路线包括 NPP/VIIRS 夜间灯光数据获取，数据处理以及人口空间化建模三部分（图 1）。首先是原始数据的获取；其次是数据处理，虽然 NPP/VIIRS 夜间灯光数据的月度数据剔除了闪电、月光照射及云覆盖的影响，但仍然存在极光、火光、船只灯光等短暂性光源和背景噪声的影响，因此需要将原始数据进行处理得到稳定的年度均值数据，因此 NPP/VIIRS 夜间灯光数据处理包括年度均值影像合成、消除负值和极高值、去除不稳定光源以及连续校正等过程^[8,9]。

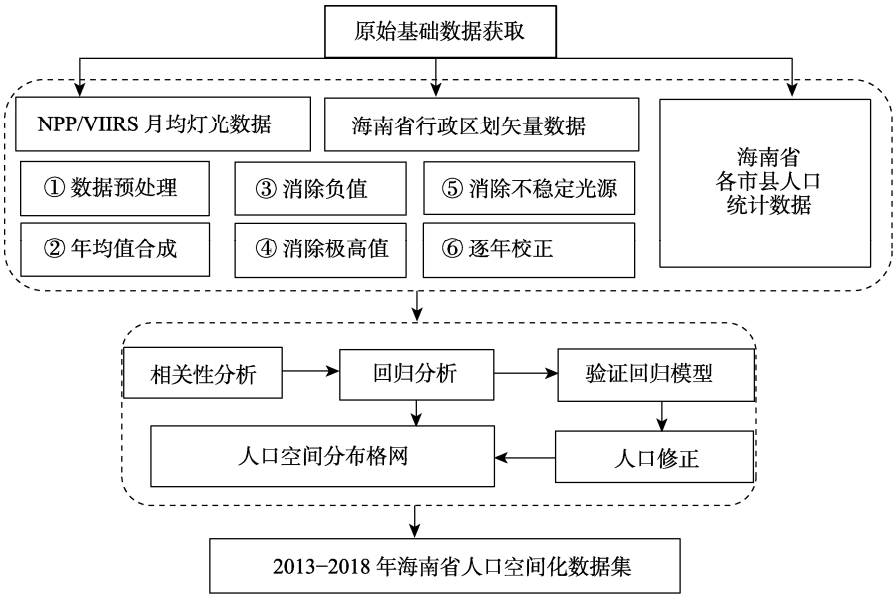


图 1 数据研发技术路线图

人口空间化建模包括相关性分析、回归建模及修正等内容。首先以市县为单位，对 2013–2017 年的各市县夜间灯光亮度值与相对应的人口统计数据进行相关性分析。进而对人口统计数据和夜间灯光亮度值进行回归建模分析。结果显示三次线性函数模型的 R^2 最高，因此采用三次线性函数模型。利用 2018 年数据验证上述模型的人口模拟精度，结果表明，海南岛整体模拟人口平均误差为 36.55%，并存在个别市县误差远大于平均误差。为进一步提高人口模拟的精度，将海南岛 18 个市县按照误差大小进行分类分别二次建模并检验

¹ 美国国家海洋和大气管理局. <https://ngdc.noaa.gov>.

² 中国自然资源数据库. <http://www.naturalresources.csdb.cn/index.asp>.

模型精度，结果表明，总体误差为 18.19%，人口模拟精度得到了显著提高。最后，由于模拟的人口数量存在一定的误差，因此引入修正系数要进一步消除误差，得到最靠近真实的人口空间分布信息。

4 数据结果与验证

4.1 数据集组成

2013–2018 年海南岛人口空间化数据集由 26 个数据文件组成，数据命名采用“主题+时间”的形式命名，如 P2013。

4.2 数据结果

将原始数据经过上述技术路线处理后得到分辨率为 500 m×500 m 的《海南岛夜间灯光与统计数据融合的人口 500-m 栅格数据集 (2013–2018)》，其中数据单位为 500m 栅格上的人口总量。由数据集得到的 2013–2018 年海南岛人口数量时空分布规律表明，海南岛人口分布整体上较为分散，城区人口密集，而乡镇人口稀疏，沿海市县人口多，而中部市县人口少，形成北海口、南三亚的两极分布特征（图 2、图 3）。在 2013–2018 年间增长的人口主要分布在各市县的城区外围，而城区中心人口增长较少，如海口、三亚两市该情况尤为明显，整体上增长人口主要分布在沿海市县（图 4）。

4.3 数据结果验证

将 2013–2018 年的 NPP/VIIRS 夜间灯光数据通过模型反演得出历年人口数量，并依据统计数据验证反演的误差。经过误差分类进行的二次建模人口模拟精度较第一次建模人口

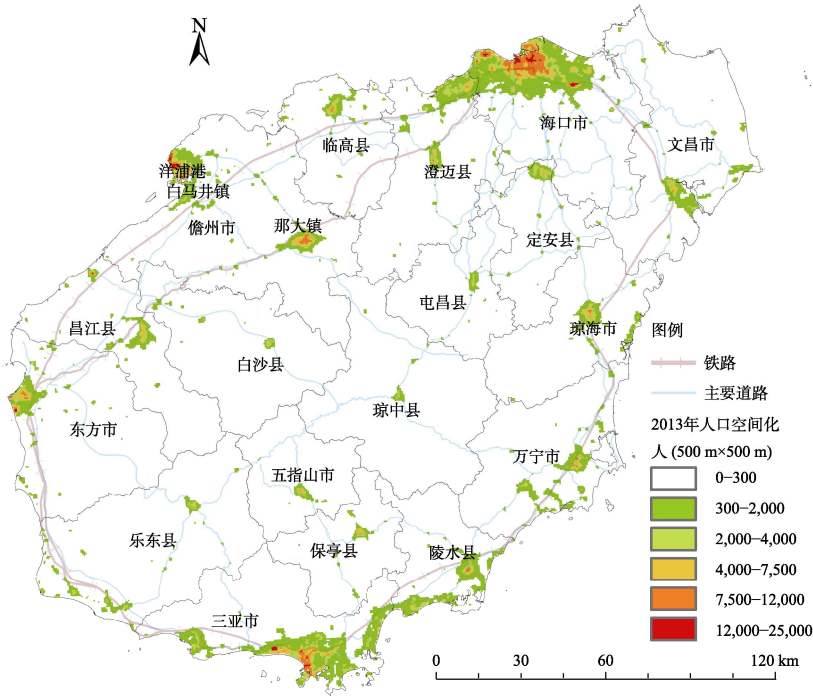


图 2 2013 年海南岛人口空间化

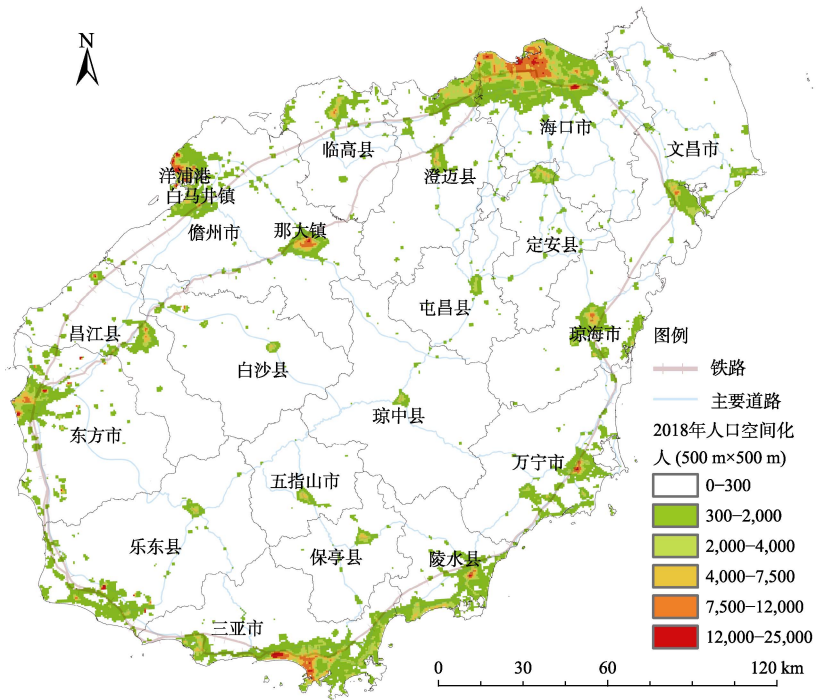


图 3 2018 年海南岛人口空间化

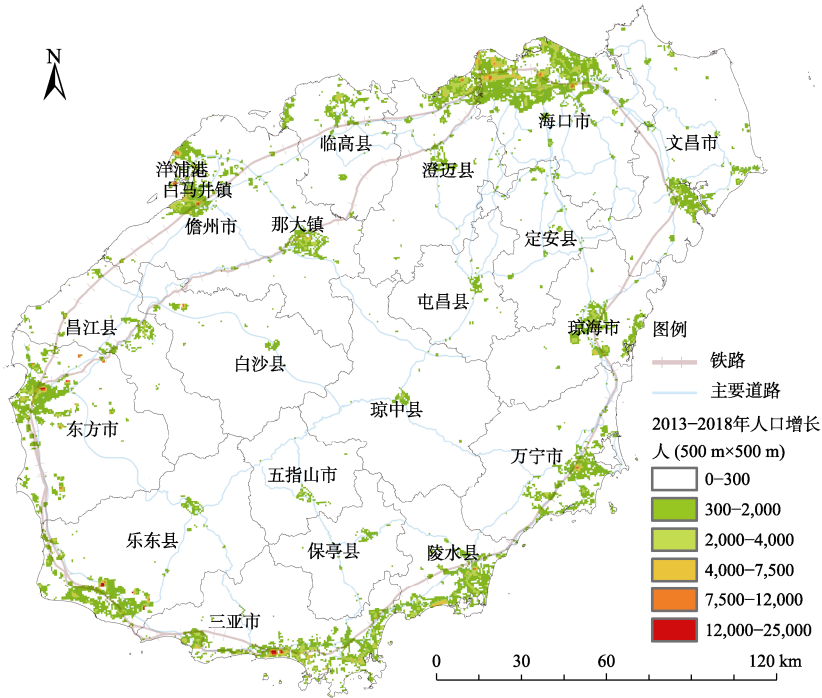


图 4 2013-2018 年人口增长

模拟精度结果有显著提高，回归系数 R^2 均为 0.9 以上，与曹丽琴等基于线性模型估算湖北地级市人口^[11] (R^2 为 0.923) 以及李翔基于空间回归模型估算上海市各乡镇街道的人口^[10]

(R^2 为 0.91) 较为接近。模型反演结果的平均相对误差为 18.19%，精度与赵宇等基于 NPP/VIIRS 夜间灯光数据构建线性回归模型估算大连市各辖区人口的平均相对误差为 18% 相近^[12]，高于王璇璇等基于 NPP/VIIRS 夜间灯光数据构建空间滞后模型估算江苏省各县市常住人口，平均相对误差为 26.09%^[13]和陈晴等基于 DMSP/OLS 夜间灯光数据构建线性回归模型估算黄河三角洲高效生态经济区 19 个县市区人口，平均相对误差为 30%^[14]。

在市县层面，各市县的反演结果中五指山市的误差较大，原因在于五指山市位于海南岛中部山区腹地，城市地形起伏较大且人口较为稀疏，单凭夜间灯光数据难以准确反演其人口空间分布数量。根据海南省行政区划，各市县按照地级市、县级市、普通县和少数民族自治县划分，可以发现二次建模的结果中地级市的人口模拟误差最小，其次是少数民族自治县，县级市误差略高于少数民族自治县，一般普通县的误差最大，不同行政规模的市县之间的人口模拟精度并没有规律性区别。

同时，付晶莹等人针对我国人口空间分布特征及区域差异进行人口区划，筛选出与人口分布相关性强的土地利用类型作为建模参数，建立多元回归模型实现人口空间化，并利用城市人口密度、交通状况、DEM 及总量控制四个因子对模型进行校正^[15]，最终得到全国人口分布数据集研究产品，其中海南岛 2010 年人口空间化数据的平均相对误差仅为 0.92%^[16]，因此，未来可参考这些方法结合夜间灯光数据进行人口空间化以提升精度。

表 2 海南岛人口数量反演结果及误差统计

类别	市县	统计人口 (万人)	一次模拟人口 (万人)	误差	二次模拟人口 (万人)	误差
地级市	海口市	230.23	258.12	12.12%	194.61	15.47%
	三亚市	77.39	121.09	56.46%	74.51	3.72%
	儋州市	99.84	90.34	9.52%	117.89	18.08%
	平均相对误差			26.03%		12.42%
县级市	文昌市	56.89	42.34	25.58%	51.57	9.35%
	琼海市	51.57	43.34	15.97%	52.66	2.11%
	万宁市	57.86	47.28	18.28%	57.13	1.27%
	五指山市	10.71	26.27	145.25%	16.36	52.77%
	东方市	42.97	62.63	45.74%	51.95	20.89%
	平均相对误差			50.16%		17.28%
普通县	定安县	29.76	30.05	0.99%	39.55	32.88%
	屯昌县	26.85	24.67	8.11%	35.09	30.68%
	澄迈县	49.44	57.02	15.33%	69.19	39.94%
	临高县	45.1	35.25	21.84%	44.32	1.73%
	平均相对误差			11.57%		26.31%
少数民族自治县	乐东县	48.27	48.94	1.39%	59.08	22.39%
	琼中县	18.02	24.86	37.95%	14.87	17.49%
	保亭县	15.28	26.68	74.59%	16.8	9.93%
	陵水县	33.39	52.65	57.67%	43.06	28.95%
	白沙县	17.34	26.71	54.05%	16.83	2.91%
	昌江县	23.35	36.68	57.07%	27.26	16.75%
	平均相对误差			47.12%		16.40%
	整体平均相对误差			36.55%		18.19%

5 讨论和总结

NPP/VIIRS 夜间灯光数据人口空间化建模能基本上反映人口的空间分布,但是五指山市的模拟精度较差。其原因可能是由于五指山市位于中部山区,地形起伏度较大,使得夜间灯光数据受较大影响,且该市的人口数量与其他市县差距较大^[14]。针对该问题,未来将考虑对研究区进行更详细的人口区划,同时结合土地利用数据,POI 等多源数据^[17]并采用多地理因子加权法、空间加权回归或神经网络等模型以提高反演精度。

本数据集是基于 NPP/VIIRS 夜间灯光数据反演的 500-m 分辨率海南岛人口空间化数据,该数据集能够反映出海南岛 2013–2018 年的人口时空动态分布情况,可为当前的海南国土空间规划提供辅助依据。

作者分工: 韩念龙对数据集的开发做了总体设计;赖先龙采集和处理了夜间灯光数据数据;赖先龙设计了模型和算法;黄鹏锦做了数据验证;赖先龙撰写了数据论文等。

利益冲突声明: 本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献

- [1] 李德仁,李熙.论夜光遥感数据挖掘[J].测绘学报,2015,44(6):591–601.
- [2] 李峰,米晓楠,刘军等.基于 NPP-VIIRS 夜间灯光数据的北京市 GDP 空间化方法[J].国土资源遥感,2016,28(3):19–24.
- [3] 李峰,孙广通,王秋玲等.NPP-VIIRS 夜间灯光数据的县级电力消费空间化研究[J].测绘与空间地理信息,2018,41(4):8–11,15,18.
- [4] 高义,王辉,王培涛.基于人口普查与多源夜间灯光数据的海岸带人口空间化分析[J].资源科学,2013,35(12):2517–2523.
- [5] 赖先龙,韩念龙,黄鹏锦.海南岛夜间灯光与统计数据融合的人口 500-m 栅格数据集(2013–2018)[J/DB/OL].全球变化数据仓储电子杂志,2021. <https://doi.org/10.3974/geodb.2021.08.02.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2021.08.02.V1>.
- [6] 全球变化科学研究数据出版系统.全球变化科学研究数据共享政策[OL]. <https://doi.org/10.3974/dp.policy.2014.05> (2017 年更新).
- [7] 海南省统计局,国家统计局海南调查总队.海南统计年鉴[M/OL].2014–2019. <http://www.hainan.gov.cn/hainan/tjnj/list3.shtml>.
- [8] 胡为安,刘传立,詹淇雯.中国区域 NPP-VIIRS 年度夜间灯光数据的合成方法与对比验证[J/OL].桂林理工大学学报,2020:1–10 [2020-12-23].
- [9] 周翼,陈英,刘洋等.NPP-VIIRS 年度夜间灯光数据的合成方法与验证[J].遥感信息,2019,34(2):62–68.
- [10] 李翔,陈振杰,吴洁璇等.基于夜间灯光数据和空间回归模型的城市常住人口格网化方法研究[J].地球信息科学学报,2017,19(10):1298–1305.
- [11] 曹丽琴,李平湘,张良培.基于 DMSP/OLS 夜间灯光数据的城市人口估算:以湖北省各县市为例[J].遥感信息,2009(1):83–87.
- [12] 赵宇,王利,李欣欣等.基于 NPP/VI RS 夜间灯光和土地利用的大连市人口空间化研究[J].资源开发与市场,2018,34(10):1372–1376.
- [13] 王璇璇,赵明松,谢毅等.基于 NPP-VIIRS 夜光遥感的江苏省人口数据空间化研究[J].黑龙江工程学院学报,2020,34(3):31–35.
- [14] 陈晴,侯西勇,吴莉.基于土地利用数据和夜间灯光数据的人口空间化模型对比分析:以黄河三角洲高效生态经济区为例[J].人文地理,2014,29(5):94–100.
- [15] 付晶莹,江东,黄耀欢.中国公里网格人口分布数据集[J].地理学报,2014,69(增刊):41–44.
- [16] 付晶莹,江东,黄耀欢.中国公里网格人口分布数据集[J/DB/OL].全球变化数据仓储,2014. <https://doi.org/10.3974/geodb.2014.01.06.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2014.01.06.V1>.
- [17] 赵鑫,宋英强,刘轶伦等.基于卫星遥感和 POI 数据的人口空间化研究:以广州市为例[J].热带地理,2020,40(1):101–109.