
基于多源遥感数据的 山地城市人口空间化及演变特征研究 ——以重庆市为例¹

祝汉收^{1, 2, 3}, 李月臣^{1, 2, 3}

(1. 重庆师范大学 地理与旅游学院, 重庆 401331;

2. 重庆市高校 GIS 应用研究重点实验室, 重庆 401331;

3. 三峡生态环境遥感研究所, 重庆 401331)

【摘要】: 随着数字化、信息化、大数据等时代的到来, 人口统计数据空间化对生态环境评价、灾害预估、区域社会和经济等的细致化和空间化研究具有重要意义。通过夜间灯光和 NDVI 二种遥感数据构建人居指数模型, 并以多元回归方法模拟人口分布并预测误差, 然后根据人口预测误差通过 DEM 数据对模型进行修正和重构, 最终模拟出山地城市人口空间分布特征。研究结果表明, 人居指数与人口之间存在极强的相关性, 由于山地城市地形的复杂性, 模拟精度较低, 通过 DEM 对人居指数的修正, 模拟精度得到有效提高, 2000 年模拟的平均相对误差从 72.53% 降低至 45.29%, 2005 年从 62.23% 降低至 36.08%, 2010 年从 44.69% 降低至 21.17%。重庆市人口空间分布模式主要呈现出人口逐渐从渝西主城区向周围扩张; 渝东北人口密度相对较小, 但在逐渐增大; 而渝东南地区人口分布密度最小, 增长最缓慢。

【关键词】: 人口空间化; 夜间灯光; NDVI; 人居指数

【中图分类号】: C922

【文献标识码】: A

中共十九大以来, 大数据再一次成为热点, 运用多种遥感数据将人口统计数据进行空间可视化模拟, 是对大数据时代的呼应。在数字化、信息化的当今时代, 许多领域内均会用到人口数据, 人口空间化为生态资源环境评价、社会经济与人口关系研究、地质灾害评估等众多领域的空间研究提供了一种有效的人口空间数据支撑。目前, 国内外对人口空间化的研究模型众多, 如范一大等、Merniis 将人口数据运用面插值方法实现了人口数据空间化, 这种方法简单易实现, 但往往不能有效反映真实的人口分布状况; 杨小唤等、田永中等、廖顺宝等^[6]根据土地利用数据与人口空间分布的关系对人口数据进行空间化模拟, 这种模型

¹[收稿日期]: 2017-12-08

[基金项目]: 国家自然科学基金项目: 山地城市产业地理集中与分散(项目编号: 41201133)。

[作者简介]: 祝汉收(1990-), 男, 山东成武人, 硕士研究生, 研究方向: 资源环境遥感与 GIS 应用。

[通讯作者]: 李月臣(1974-), 男, 教授, 重庆市人, 研究方向: 资源环境遥感、GIS 应用与开发以及区域生态环境等。

更符合实际，但模型默认各土地类型内的人口均匀分布，不能有效反映同一土地类型内人口分布的差异性。近年来，因夜间灯光数据的灯光强度和人口数量之间存在极高的相关系数，越来越多的被运用于人口空间化模型中，但夜间灯光数据往往存在空间分辨率较低、灯光过饱和和像元溢出等问题。NDVI 数据作为衡量植被茂密程度的指标之一，往往能有效反映人类活动的规律，且分辨率相对较高，与夜间灯光数据存在互补的作用。Lu 等利用夜间灯光数据和 NDVI 数据提出一种能反映人类居住适宜度的人居指数，模拟出的人口空间分布情况更加符合实际。本文首先结合夜间灯光数据与 NDVI 数据构建人居指数模型，然后将人居指数与人口统计数据做回归分析，根据回归方程模拟出各区县的人口。由于山地城市地形复杂，本研究将 DEM 数据对人居指数进行了优化，形成新的人居指数模型，最终模拟出重庆市 2000、2005 和 2010 年三期人口空间分布图。

重庆市是中国中西部唯一的直辖市、长江上游的经济中心，地理位置和经济位置均极其重要，是各种生态、资源和环境、地质灾害等与人口关系的研究热点区域。本研究将重庆市作为研究区，将为该地区各课题提供更有效的人口数据。重庆市是典型的山地城市，地形地貌极大的限制了人口的空间分布，人口空间化模拟对了解该地区的人口分布规律，解析山地城市人口分布影响因子有重大意义。

1、数据处理与研究方法

1.1 数据获取与处理

本次使用的数据主要包括：①重庆市地理信息基础数据；②来《重庆市统计年鉴》的各区县各时期的人口统计数据；③DEM 数据来源于地理空间数据云，重采样为 25m；④NDVI 逐月数据来源于美国 LAADSAC (<https://ladsweb.nascom.nasa.gov/>)，空间分辨率重采样为 250m；⑤DMSP-OLS 夜间灯光数据来源于美国国家地理数据中心 (<https://www.ngdc.noaa.gov/eog/>)，空间分辨率为 1km，因 2013 年以后传感器更换，夜间灯光数据不同，本研究未模拟 2015 年人口空间数据。将所有数据投影为 Albers 等面积投影，用重庆市矢量边界对数据进行裁剪，得到重庆市范围内的各数据集。将 NDVI 数据和夜间灯光数据重采样成空间分辨率为 25m 的数据，以利于结合 DEM 数据，同时，剔除 NDVI 小于 0.1 和大于 0.9 不适宜人口居住的区域，并用水系数据剔除水体。

1.2 研究方法

人居指数能有效反映区域人口居住环境的适宜程度，根据 NDVI 和 DMSP-OLS 灯光数据构建人居指数，因 NDVI 数据的值在 -1-1 之间，应对 DMSP-OLS 数据做标准化处理。本文参考 LU 等构建的人居指数对重庆市地区构建人居指数 (HSI) 模型，具体公式如下：

$$OLS_{nor} = \frac{OLS}{OLS_{max}} \quad (1)$$

$$HSI = \frac{(1-NDVI_{max})+OLS_{nor}}{(1-OLS_{nor})+NDVI_{max}+OLS_{nor} \times NDVI_{max}} \quad (2)$$

式中：OLS 为夜间灯光数据值；OLS_{max}为夜间灯光数据的最大值；HIS 为人居指数；NDVI_{max}为 NDVI 年最大值；OLS_{nor}为经过标准化的夜间灯光数据。人居指数越高表示越适宜人居住。

人口模拟误差检验是对模拟结果精度的计算，具体计算公式如下：

$$MPE = \frac{\sum_{i=1}^n |(RE)_i|}{n} \quad (3)$$

$$RE = \frac{POP_t - POP_c}{POP_c} \times 100\% \quad (4)$$

式中：MPE 为平均相对误差；RE 为相对误差；POP_s 为区县模拟人口；POP_c 为区县统计人口；n 为重庆市区县个数。

2、结果分析

2.1 重庆市各区县人口空间化的初步模拟

运用 SPSS 软件对 2000、2005 和 2010 年重庆市各区县的人口密度值和人居指数密度值做回归分析，发现 R² 大于 0.9，两者具有极强的相关性，对模拟的结果进行误差检验（见表 1），发现平均相对误差极大，不能用于相关研究中。

表 1 人居指数与人口密度相关系数及模拟结果误差检验表

年份	2000 年	2005 年	2010 年
相关系数 (R2)	0.95	0.97	0.91
平均相对误差 (MPE)	72.53%	62.23%	44.69%

在未考虑高程的情况下，人口模拟误差极大。通过分析发现相对误差小于 20% 的区县平均高程基本都小于 600m。高程大于 600m 的区县相对误差基本大于 50%，因此，本文将高程小于 600m 的区县运用人居指数模型单独计算，高程大于 600m 的区县引入 DEM 高程数据对人居指数模型进行修订。将重庆市高程大于 600m 区县的平均 DEM 与人口密度做相关性分析（见表 2）。

表 2 平均高程与人口密度相关系数表

年份	2000 年	2005 年	2010 年
相关系数 (R2)	-0.72	-0.73	-0.78

通过表 2 可知，平均高程与人口密度的相关系数 R² 均大于 0.7，证明高程与人口分布存在着较大的相关性。因此，这里利用 DEM 数据对高程大于 600m 区县的人居指数进行修订，参考杨续超等的研究，公式如下：

$$HSI_{new} = \frac{(1 - NDVI_{max}) + OL_{new}}{(1 - OL_{new}) + NDVI_{max} + OL_{new} \times NDVI_{max}} \times e^{-aDEM} \quad (5)$$

式中：HSI_{new} 修订后的人居指数，a 为 e 指数函数里 DEM 的系数。

根据修正后的人居指数模型对人居指数和人口数据做相关性分析和回归分析，根据回归方程模拟重庆市 2000、2005 和 2010 年三期人口空间分布图，并对结果进行误差检验（表 3）。

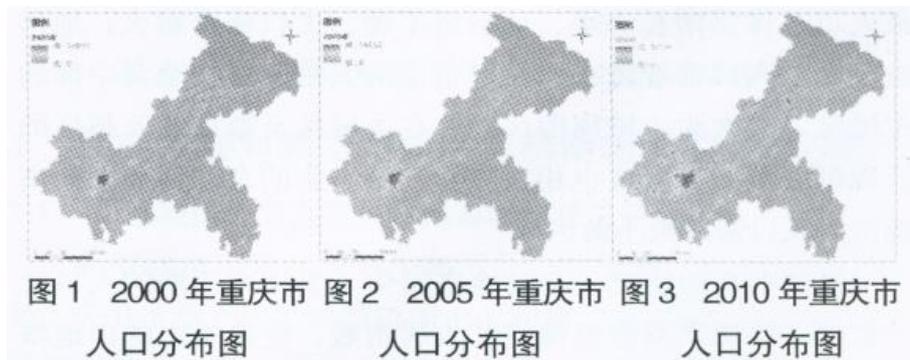
表 3 修正后人居指数与人口密度相关系数及模拟结果误差体验表

年份	2000 年	2005 年	2010 年
相关系数 (R2)	0.98	0.97	0.97
平均相对误差 (MPE)	45.29%	36.08%	21.17%

根据表 3 和表 1 可知, 修订后的人居指数与人口的相关系数更高, 模拟出的人口精度远高于修订前, 说明增加了高程因子使人口模拟效果更好。2000 年模拟的人口平均相对误差从修订前的 72.53% 降低至 45.29%, 2005 年人口平均相对误差从 62.23% 降至 36.08%, 2010 年人口平均相对误差从 44.69% 降至 21.17%, 且时间越近模拟的精度越高。通过分析发现, 人口误差较大的区县多是海拔相对较高且发展相对落后的区县, 分析原因可能是本研究用夜间灯光数据是根据灯光强度判断人口分布的情况, 发展相对落后的地区夜间用灯较少或灯光微弱, 导致夜间灯光捕捉不完全, 而时间越近, 经济越发达, 灯光捕捉率则越高, 模拟出的人口精度也就越高。同时, 重庆市复杂的地形地貌条件, 使人口分布极度不均匀, 以区县为边界对人口数据取平均值进行回归分析, 本身误差较大, 但经过 DEM 的修订后, 模拟精度得到极大的提高, 说明本研究的方法是可行的, 在适当提高人口统计边界精度的前提下, 运用本研究的方法可以较好的模拟出反映重庆市人口的实际空间分布情况。

2.2 重庆市人口空间模拟结果及分布特征

结合修订前和修订后的人居指数对重庆市人口进行模拟, 得到 2000、2005 和 2010 年三期 25m×25m 分辨率的人口空间分布图 (见图 1~3)。



根据模拟出的三期人口空间分布图可以发现, 重庆市人口在历史发展中均集中分布在渝西中部的主城区, 主城区外围则有多个聚居地环绕, 渝东北地区中部万州区附近分布多个人口聚居地, 渝东南人口密度总体最低。随着时间的推移, 重庆市人口呈逐渐增加的趋势, 人口从主城区逐渐向外扩展, 特别是 2005 至 2010 年间人口急剧增加, 在渝西西部地区的多个聚居地人口规模普遍增大; 同时, 人口呈现从主城区向渝东北地区蔓延的趋势, 在主城区至渝东北中部的万州区一带, 人口有多个聚居地, 且人口规模逐渐增大, 至 2010 年, 这一长廊形区域人口已普遍分布, 并形成多个较大的人口聚居地。渝东南地区人口密度在三个时期均是最低的区域, 且在发展过程中只有渝东南最南部人口有明显增加趋势, 其他区域人口密度则普遍较低, 人口规模增长不明显。

根据 DEM 矫正后的人居指数模型和重庆市的地形地貌分布规律分析, 渝西地区为浅丘地貌, 地势较平, 特别是渝西中部的主城区, 地势平缓, 人居指数极高, 是重庆市人口分布的主要地区。渝西至渝东北的长廊形地区, 地势相对较平, 人居指数相对较高, 并受主城区和万州区的带动, 经济发展水平得到有效提高, 人口密度较大; 渝东东北部位于大巴山山脉, 地形起伏度较大, 人居环境较差, 人口数量分布也较少。渝东南地区地处武陵山和大娄山山脉, 地势较复杂, 交通不便, 经济欠发达, 导致人居环境较差, 人口增长幅度极小, 人口分布是整个重庆市最少的地区。

通过分析重庆市人口在 2000、2005 和 2010 年的分布情况, 反映了人口的分布受地形地貌的限制极大, 地形起伏度较小, 高程较低的地区, 人居环境较好, 较适宜人类的居住, 是人口数量和密度较大的地区; 地势复杂、高程越高和交通不便的地区, 人居环境较差, 人口密度则相对较低。

3、结论与讨论

人口统计数据受行政区划的限制,存在时间和空间分辨率低的缺点,在当今数字化和信息化的时代,人口统计数据已远远不能满足研究的需要,人口数据空间化为解析人口具体空间分布规律、灾害风险精细化评估及人口与资源环境的协调性空间评价等领域提供了可能。对于山地城市而言,高程因子是人口分布的必要性因子,通过 DEM 数据对人居指数的修订,能更有效精确地模拟山地城市的人口空间分布。此外,分析发现人口模拟精度随时间的推移精度越来越高,从侧面反映人口模拟精度一定程度上受地区经济发展的影响。重庆市人口总体呈增长态势,地形越平缓,人口密度越大;地形越复杂,人口密度越小。人口密度最大的区域在渝西中部的主城区,其次为主城周围区县,在主城区至渝东北万州区的长廊形区域人口密度也相对较大,渝东北的东北部和渝东南地区为人口密度最小的区域。

通过本文的研究,有效地空间化了重庆市三期的人口统计数据,并用高程数据修订了人居指数,使模拟出的山地城市人口分布状况更符合实际,但本研究也存在许多不足之处,虽然人居指数经过高程数据的修订,模拟精度有较大的提高,但模拟结果仍存在较大误差,在后续研究中,应增加其他影响人口空间分布的因子,对人居指数模型进一步修改完善,并提高人口统计数据的边界分辨率,以提高模拟精度。

[参考文献]:

- [1]司春霞,胡瑞芝.我国农村居民点布局存在的问题及对策研究[J].农村经济与科技,2006(12):64-65.
- [2]范一大,史培军,辜智慧,等.行政单元数据向网格单元转化的技术方法[J].地理科学,2004,24(1):105-108.
- [3]Mennis J. Generating surface models of population using dasymetric mapping[J].The Professional Geographer,2003,55(1):31-42.
- [4]杨小唤,江东,王乃斌,等.人口数据空间化的处理方法[J].地理学报,2002,57(增刊):70-75.
- [5]田永中,陈述彭,岳天祥,等.基于土地利用的中国人口密度模拟[J].地理学报,2004,59(2):283-292.
- [6]廖顺宝,李泽辉.基于人口分布与土地利用关系的人口数据空间化研究——以西藏自治区为例[J].自然资源学报,2003(06):659-665.
- [7]Prosperie L, Eyton R. The relationship between brightness values from a nighttime satellite image and Texas County population[J].Southwestern Geographers,2002,92(2):224-240.
- [8]Amaral S, Monteiro A, Camara G, et al. DMSP-OLS night-time light imagery for urban population estimates in the Brazilian Amazon[J].International Journal of Remote Sensing,2006,27(5):855-870.
- [9]Sutton P. Modeling population density with night-time satellite imagery and GIS[J]. Computer, Environment and Urban System,1997,21(3):227-244.
- [10]Chowdhury P K R, Maithani S, Dadhwal V K. Estimation of urban population in Indo-Gangetic Plains using night-time OLS data[J].International Journal of Remote Sensing,2012,33(8):2498-2515.
- [11]卓莉,陈晋,史培军,等.基于夜间灯光数据的中国人口密度模拟[J].地理学报,2005(02):266-276.

[12]Roy Chowdhury P, Maithan I S. Monitoring growth of built-up areas in indo-gangetic plain using multi-sensor remote sensing data[J].Journal of the Indian Society of Remote Sensing, 2010, 38 (2) : 291-300.

[13]Lu D, Tian H, Zhou G, et al. Regional mapping of human settlements in southeastern China with multisensory remotely sensed data[J].Remote Sensing of Environment, 2008, 112 (9) : 3668-3679.

[14]Lu D, Weng Q, Li G. Residential population estimation using a remote sensing derived impervious surface approach[J].International Journal of Remote Sensing, 2006, 27 (16) : 3553-3570.

[15]杨续超, 高火伟, 丁明军, 等. 基于多源遥感数据及 DEM 的人口统计数据空间化——以浙江省为例[J]. 长江流域资源与环境, 2013 (06) : 729-734.