



基于建筑物信息的城市街道尺度人口估算

孟瑶^{1,2} 李锐¹ 蒋捷³ 王顺利¹ 吴华意¹

1 武汉大学测绘遥感信息工程国家重点实验室,湖北 武汉,430079

2 航天工程大学士官学校,北京,102249

3 北京建筑大学测绘与城市空间信息学院,北京,102616

摘要:随着城镇化进程的加快与城市人口的迅速膨胀,街道尺度的人口数据在城市经济、社会、资源与环境发展等方面都发挥着愈发重要的作用。研究如何利用高分辨率遥感影像进行城市街道尺度上的人口估算,对促进城市可持续发展具有十分重要的理论意义与实际价值。利用遥感影像建筑物信息与人口普查数据,分析街道建筑物信息与人口数量之间的关联性,提出基于建筑物信息的城市街道尺度人口估算模型,通过多元逐步回归法与赤池信息准则确定建筑物显著特征变量,建立了建筑物数量、几何特征与人口数量的估算模型。实验证明,提出的人口估算模型能够以较高的精度估算街道尺度的人口数量。

关键词:人口估算;遥感影像;建筑物信息;逐步回归

中图分类号:P208

文献标志码:A

城市人口是城市国土空间规划基本且重要的内容之一,人口数据是城市应急响应、城市承载功能设计、城市资源开发与分配、城市产业结构调整、城市交通规划设计等方面着重考虑的基础信息要素,对促进城市经济社会的全面协调可持续发展具有十分重要的意义^[1]。人口普查是现今世界各国广泛采用的获取人口资料的方法,其数据真实、准确、完整,但人口普查成本高、工作量大,且普查周期为 10 年,无法提供普查间隔内的人口资料^[2]。因此,人口普查数据在时间分辨率上的天然缺陷使其不能满足许多与人口相关研究与应用的需要,如在研究人口目标区域上更短时间内的人口变化、解决人口目标区域与人口普查区边界不一致的问题及对更小空间尺度上人口数据的需求等^[3]。

随着摄影测量与遥感技术的迅速发展,利用影像进行地表信息的快速采集与更新逐渐成为重要方法之一。摄影测量数据与遥感影像在数据获取方式及时空分辨率上的优势使其成为新型人口调查的重要辅助数据^[4-5]。半个多世纪以来,国内外学者提出了许多基于遥感影像和地理信息系统(geographic information system, GIS)辅

助的人口估算方法。利用遥感影像进行人口估算研究的基本思想如下:首先,从遥感影像中获取各类与人口数量可能相关的信息;然后,利用先验知识与数学方法建立人口数量与所获取信息的数学模型,从而实现人口数量估算,这是交叉学科的应用之一。从影像中获取的信息通常包括城市面积^[6-7]、土地利用类型^[8-9]、居住单元^[10-11]、影像的光谱反射值或纹理^[12-13]。多数人口估算研究集中在城市面积估算法与图像像元特征估算法上,方法较为成熟,但中等分辨率遥感影像的低空间分辨率缺陷使其在城市街道尺度的人口估算研究上具有不适用性。

建筑物相关信息与城市人口数据联系十分紧密,亦有许多学者从高分辨率遥感影像中提取建筑物轮廓、高度等信息。Wu 等^[14]将人口普查区域依照建筑物体积进行分块,使用总建筑体积与住房建筑物面积对人口数量进行回归分析,结果表明,子块区域越小,人口估算的精度越高。鹿琳琳等^[15]根据深圳市福田区的建设情况把住宅类型依据楼层分为 3 类,基于遥感与 GIS 辅助数据对研究区域内各居委会的人口数量进行估算。Lwin 等^[16]讨论了基于人口普查区和建筑物

收稿日期:2020-05-27

项目资助:国家重点研发计划(2016YFB0502301);国家自然科学基金(41771426)。

第一作者:孟瑶,硕士,主要研究方向为时空计算与数据挖掘。yao.meng@whu.edu.cn

通讯作者:蒋捷,博士,教授。jiangjie@bucea.edu.cn

足迹数据集估算建筑物人口的 GIS 方法,并通过目视判读、空间统计分析等手段评估估算结果。冯甜甜等^[17]针对美国居住人口的分布特点将建筑物分为单居民住宅区、多居民住宅区与非居民区建筑,建立居民建筑与人口数量的线性关系,实现美国城市人口估算。Ural 等^[18]利用航空影像和其他地理数据提出一种改进的建筑物人口估算模型,考虑了不同类型住宅建筑对人口数量的影响。封静^[19]通过实地调查获取楼层与家庭人口信息,从遥感影像中提取建筑物面积与高度信息,并将建筑物依据高度进行分类,在估算建筑物面积后除以人均居住面积,实现对上海松江新城 17 个居委会的人口估算。利用高分辨率遥感影像的建筑物信息可以有效地估算街道尺度人口数量。在数学模型的应用上,各类回归模型被用于建立估算模型,但对较为经典的逐步回归分析法的使用较少,且变量选择主要依靠经验值,极少进行统计检验。

本文通过从高分辨率遥感影像中获得的的城市街道建筑物的数量与几何特征,分析人口数量和街道建筑物特征变量之间的数学关系,提出基于街道建筑物信息的城市街道尺度人口估算模型,从而实现对城市街道尺度人口数量的估算。

1 建筑物信息与人口数量关联性

1.1 样本数据及预处理

本文采用 2010 年全国第六次人口普查数据、2017 年全国乡镇/街道行政边界与 2018 年第 1 季度遥感影像数据为本文研究的基础数据。利用软件 GlobalMapper 提取全国范围内 24 个主要城市,共 287 条街道的建筑物数量与几何特征,其中,几何特征包括建筑物的占地面积、周长与楼层信息。

数据预处理。首先,筛选街道建筑物,剔除占地面积超过 1 500 m²的非居民地建筑物;然后,筛选街道,剔除人均建筑面积大于 120 m²的非居民地建筑物占比显著的街道;最后,对保留的 232 条样本街道,共计 348 403 条建筑物数据进行分类,划分为层数 ≤ 2 的低层建筑物、2 < 层数 ≤ 10 的中层建筑物、层数 > 10 的高层建筑物 3 类。以 2 层作为低层与中层分类的依据是:不高于两层的建筑物,包括一层与两层,通常为城市边缘地区如村镇等人口的主要居住建筑,多为独户家庭居住环境,具有较为特殊的人口密度;以 10 层作为中层与高层分类的依据为采取城市中心地区

建筑物楼层分类的经验值。

1.2 关联性分析

本文选取街道建筑物信息中与人口数量分布存在相关性的建筑物特征变量,对各建筑物特征变量间进行关联性分析。根据居住单元估算法,人口数量与建筑物数量之间存在密切相关关系,由此选取街道总建筑物数量 N_T ,由于不同高度的建筑物人均占有面积不同,因此,选取将街道建筑物依据楼层分类后的低层建筑物数量 N_L 、中层建筑物数量 N_M 与高层建筑物数量 N_H 。仅利用建筑物数量信息无法刻画建筑物的几何特征,因此,增加选取建筑物面积、周长与楼层,同理,增加选取特征变量 $A_T, A_L, A_M, A_H, C_T, C_L, C_M, C_H, F_T, F_L, F_M, F_H$ 。本文共选取 16 个建筑物特征变量,分别是 $N_T, N_L, N_M, N_H, A_T, A_L, A_M, A_H, C_T, C_L, C_M, C_H, F_T, F_L, F_M, F_H$,即街道建筑物数量、面积、周长、楼层和经建筑物分类后的各对应变量。

由于选取的变量之间可能存在一定的相关性,完全无关的变量没有共同建模的必要,同时,要求各变量间存在较小的共线性,从而提高建模效果,因此,需要对各变量进行关联性分析。对街道建筑物总数量、总面积、总周长、总楼层 4 个特征变量进行的关联性分析如表 1 所示。

表 1 建筑物特征总量间相关系数

Tab.1 Correlation Coefficients Between Total Number of Building Features

变量	相关系数			
	N_T	A_T	C_T	F_T
N_T	1.00	0.98	0.99	0.86
A_T	0.98	1.00	1.00	0.86
C_T	0.99	1.00	1.00	0.86
F_T	0.86	0.86	0.86	1.00

表 1 显示 N_T, A_T 和 C_T 三者之间具有非常强的两两相关性,由于大样本建筑物占地面积平均值趋于定值,且建筑物的形状大多为规则几何形状,因此,在大样本数据统计结果中,这三者呈现非常强的正线性相关关系。 N_T, A_T, C_T 和 F_T 之间的相关性相对弱于前三者之间的相关性,主要是现代城市建设更多地向三维空间扩展,与建筑物平均占地面积相比,建筑物的平均高度更具差异性,导致 F_T 与前三者之间的相关性相对较弱,但仍处于 0.86 的较高水平。预期仅使用街道建筑物特征变量总数会较大地削弱建筑物信息对人口数量的估算能力,同时,也掩盖了不同街道建

筑物的特性,从第1个方面支持了对街道建筑物分类的必要性。

街道建筑物数量、面积、周长、楼层经过建筑

物分类后得到的12个变量的关联性统计结果如表2所示。

表2 建筑物分类特征变量间相关系数

Tab.2 Correlation Coefficients Between Classified Building Features

变量	相关系数											
	N_L	A_L	C_L	F_L	N_M	A_M	C_M	F_M	N_H	A_H	C_H	F_H
N_L	1.00	0.97	0.99	0.98	0.30	0.39	0.35	0.18	0.29	0.28	0.30	0.24
A_L	0.97	1.00	0.99	0.94	0.26	0.37	0.32	0.14	0.26	0.25	0.26	0.20
C_L	0.99	0.99	1.00	0.96	0.28	0.37	0.33	0.15	0.27	0.26	0.27	0.21
F_L	0.98	0.94	0.96	1.00	0.36	0.44	0.40	0.23	0.36	0.35	0.37	0.30
N_M	0.30	0.26	0.28	0.36	1.00	0.97	0.98	0.96	0.55	0.52	0.54	0.47
A_M	0.39	0.37	0.37	0.44	0.97	1.00	0.99	0.92	0.54	0.51	0.53	0.45
C_M	0.35	0.32	0.33	0.40	0.98	0.99	1.00	0.94	0.54	0.51	0.53	0.45
F_M	0.18	0.14	0.15	0.23	0.96	0.92	0.94	1.00	0.52	0.49	0.51	0.43
N_H	0.29	0.26	0.27	0.36	0.55	0.54	0.54	0.52	1.00	0.99	1.00	0.98
A_H	0.28	0.25	0.26	0.35	0.52	0.51	0.51	0.49	0.99	1.00	1.00	0.97
C_H	0.30	0.26	0.27	0.37	0.54	0.53	0.53	0.51	1.00	1.00	1.00	0.97
F_H	0.24	0.20	0.21	0.30	0.47	0.45	0.45	0.43	0.98	0.97	0.97	1.00

表2中,在低层、中层与高层建筑物的跨建筑物类型中,各特征变量之间的相关性大幅度减弱,显著增强了街道建筑物信息对人口数量的可描述性,从第2个方面支持了对街道建筑物分类的必要性。

本文分析了街道建筑物特征变量与城市街道人口数量之间的相关性,表3为16个变量与人口数量的相关系数。

表3 全变量与人口数量的相关系数

Tab.3 Correlation Coefficients Between Each Building Feature and Population

变量	相关系数	变量	相关系数	变量	相关系数	变量	相关系数
N_L	0.12	A_L	0.09	C_L	0.10	F_L	0.18
N_M	0.76	A_M	0.74	C_M	0.76	F_M	0.82
N_H	0.52	A_H	0.50	C_H	0.52	F_H	0.48
N_T	0.57	A_T	0.60	C_T	0.60	F_T	0.79

表3显示 N_T 、 A_T 、 C_T 、 F_T 和人口数量之间的相关性均弱于对应中层建筑物特征变量与人口数量之间的相关性,预期利用4个特征总量建立模型的估算效果弱于分类后的模型。另外,中层建筑物特征变量和人口数量之间的相关性明显突出,高层建筑物特征变量与人口数量之间的相关性明显强于低层建筑物特征变量与人口数量之间的相关性。这是由于在当前城市街道建筑物体系中,中层建筑物占有相当大的比例,而且

在相同数量的低层与高层建筑物中,高层建筑物的人口容纳量远高于低层建筑物的人口容纳量。预期对于城市街道尺度人口数量的描述,主要依赖于中层与高层建筑物的数量与几何特征,低层建筑物的数量与几何特征起辅助补充作用,从第3个方面支持了对街道建筑物分类的必要性。

2 城市街道尺度人口估算模型

2.1 估算模型设计

依据对城市街道建筑物信息与人口数量的关联性分析,利用街道建筑物总数量 N_T 建立线性回归模型1a。由于不同高度街道建筑物人均占有面积不同,在回归模型中表现为系数的不同,因此,在模型1a的基础上,利用分类后的建筑物数量 N_L 、 N_M 与 N_H 建立模型1b。同理,利用建筑物面积、周长与楼层总量建立模型2a、3a、4a,利用各分类后的特征变量建立模型2b、3b、4b。

模型1~模型4均利用街道建筑物单一几何特性进行人口估算,变量数量较少,方法简单,计算简便。以此为基础,本文综合构建上述模型的所有变量,利用逐步线性回归法挑选街道建筑物显著特征变量,建立线性回归综合模型5a与模型5b。利用变量 N_T 、 A_T 、 C_T 、 F_T 建立模型5a,利用变量 N_L 、 N_M 、 N_H 、 A_L 、 A_M 、 A_H 、 C_L 、 C_M 、 C_H 、 F_L 、 F_M 、 F_H 建立模型5b。模型表达式如表4所示。

表 4 模型表达式
Tab.4 Model Expressions

模型	模型表达式
1a	$\hat{P} = \text{floor}(46\,332.554 + 19.365N_T)$
1b	$\hat{P} = \text{floor}(36\,110.249 - 8.712N_L + 40.910N_M - 58.469N_H)$
2a	$\hat{P} = \text{floor}(43\,580 + 0.052A_T)$
2b	$\hat{P} = \text{floor}(36\,430 - 0.041A_L + 0.099A_M + 0.108A_H)$
3a	$\hat{P} = \text{floor}(43\,260 + 0.233C_T)$
3b	$\hat{P} = \text{floor}(34\,980 - 0.149C_L + 0.453C_M + 0.561C_H)$
4a	$\hat{P} = \text{floor}(33\,140 + 5.587F_T)$
4b	$\hat{P} = \text{floor}(33\,710 - 1.414F_L + 7.660F_M + 3.250F_H)$
5a	$\hat{P} = \text{floor}(31\,280 - 74.36N_T - 0.19A_T + 1.525C_T + 7.996F_T)$
5b	$\hat{P} = \text{floor}(30\,070 - 96.94N_M - 0.175A_M + 1.602C_M + 11.71F_M - 219.4N_H + 16.10F_H)$

2.2 估算模型建立

模型 1~模型 4 均采用多元线性回归法构建,在回归人口数与实际人口数残差平方和最小的前提下,利用最小二乘法求解模型的回归系数^[20]。

模型 5a 与模型 5b 利用赤池信息准则 (Akaike information criterion, AIC) 进行向后逐步回归挑选最佳模型^[21]。AIC 由日本统计学家赤池弘次^[22-23]提出,用于衡量给定数据集中不同统计模型的相对质量。向后逐步回归的主要思想是:依次排除模型变量后,对模型进行检验,从而确定剔除变量,直到没有满足继续剔除条件的变量为止,确定最优模型^[24]。具体步骤为:(1)构建全变量回归方程,计算方程 AIC 值。(2)逐次建立某一变量系数为 0 的回归方程,计算该方程 AIC 值并检验变量显著性。(3)根据新的 AIC 值与统计检验值判断是否剔除该变量,若无需要剔除的变量,则向后逐步回归过程结束;若需要剔除某一变量,则将该变量剔除,进入下一步。(4)建立剩余变量的回归方程,计算 AIC 值,返回步骤 (2)。选定用于建立模型的街道建筑物显著特征变量后,利用多元线性回归构建回归方程,此即候选模型集合中的最优模型。模型 5b 最终保留的街道建筑物显著特征变量为中层建筑物的总数量、总面积、总周长、总楼层与高层建筑物的总数量、总楼层。低层建筑物数量与几何特征全部被剔除,中层建筑物数量与几何特征全部被保留,高层建筑物保留数量与楼层特征与关联性分析预期的结果一致。

3 实验与分析

3.1 模型比较

基于街道建筑物数据与人口普查数据建立模型,统计各模型决定系数与残差标准差如表 5 和图 1 所示。

表 5 模型决定系数与残差标准差

Tab.5 Model Determination Coefficients and Residual Standard Deviations

模型	决定系数	残差标准差/人	模型	决定系数	残差标准差/人
1a	0.321	25 860	3b	0.631	19 170
1b	0.606	19 810	4a	0.626	19 180
2a	0.355	25 210	4b	0.699	17 300
2b	0.613	19 650	5a	0.711	17 020
3a	0.354	25 210	5b	0.757	15 680

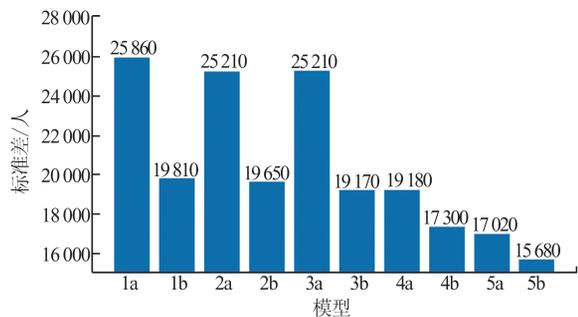


图 1 模型残差标准差

Fig.1 Model Residual Standard Deviations

表 5 与图 1 表明,在对街道建筑物进行低层、中层与高层建筑物分类后,模型的估算效果具有显著提升。模型 1b、2b、3b 的决定系数与残差标准差非常接近,这与关联性分析中建筑物总数量、总面积与总周长的极强相关性相符合。而模型 1a、2a、3a 的决定系数与残差标准差均远差于模型 1b、2b、3b 的,证明了考虑楼层信息对街道建筑物进行分类的有效性。模型 4a 的各项统计量与模型 1b~模型 3b 相当,这是由于在统计总楼层时同样考虑了高度信息,而模型 4b 将高度信息进一步细化分类,模型估算效果明显优于模型 4a,再次证明了对街道建筑物分类的必要性。在利用单一街道建筑物特征变量建立的模型 1~模型 4 中,利用楼层数建立的模型 4 估算效果显著优于利用数量、面积与周长特征变量建立的模型 1、模型 2、模型 3,这是因为在现代城市建筑中,高度空间的利用越发重要,事实上楼层信息才更加细致地刻画了人群居住空间的特点。

采用向后逐步回归法建立的模型 5a 与模

型4b的估算效果相当,与模型4a相比,在性能上存在显著的提升,说明街道建筑物数量、面积与周长3类特征信息对于人口数量估算具有明显的辅助作用,这一点从模型5b与模型4b的比较中亦能得到印证。模型5b在统计上具有最优精度,决定系数为0.757,残差标准差为15 680人。这是由于该模型对街道建筑物特征信息进行了最充分的利用,在模型4b的基础上,加入了其他街道建筑物特征辅助信息;在模型5a的基础上,增加了依据楼层对街道建筑物的分类。

各模型的估算精度表明,在实际进行城市街道尺度人口估算时,街道建筑物楼层或高度信息是必要的,一方面可以直接将楼层数据应用于估算模型,另一方面楼层数据亦是对街道建筑物进行分类的依据;将街道建筑物按照层数分为低层、中层与高层建筑物,可以显著提高模型的人口估算效果;模型5b在人口估算整体上具有最佳性能,但当街道建筑物数据种类有限时,采用模型4b与模型5a同样可以得到较好的估算效果。

3.2 模型验证

由于模型1a~模型4a与模型1b~模型3b的精度均较低,估算偏差较大,因此,在后续研究中,仅对模型4b、模型5a与模型5b进行讨论。为验证模型的有效性与适用性,以长沙市为例,对长沙市进行街道尺度的人口数量估算。首先,基于深度学习中Robosat工具从高分二号遥感影像中提取街道建筑物轮廓信息;然后,从网上公开的长沙市建筑物数据提取街道建筑物高度信息,匹配给从影像中提取的相应街道建筑物。经数据预处理后,保留8条测试街道,共计19 126条街道建筑物数据,代入模型4b、模型5a与模型5b,各模型人口估算相对误差如表6所示。

表6 长沙市人口估算相对误差/%
Tab.6 Relative Errors of Population Estimation in Changsha City/%

街道	模型4b	模型5a	模型5b
岳麓	7.630	20.500	0.734
左家塘	19.526	15.766	16.065
新开铺	2.794	17.918	0.039
洪山	33.202	61.525	48.245
湘龙	24.686	7.233	15.354
观沙岭	40.268	65.476	43.693
雨花亭	1.157	15.550	13.307
青园	61.982	89.814	29.040

从表6可知,本文提出的模型4b、模型5a与模型5b均取得良好的估算效果,其中,模型5b的

综合估算精度最佳,模型4b次之,模型5a估算效果略差。模型4b与模型5b对岳麓街道与新开铺街道的估算精度最佳,这是由于两条街道低中高层建筑物数量比例约为3:4:1,是较为标准的街道规模,同样也是样本街道的主要规模,因此,估算效果非常理想。当街道人口数在12万人左右时,模型5b能够对模型4b的估算效果进行有效的修正,这是由于模型5b在模型4b仅利用建筑物楼层信息估算人口的基础上考虑到建筑物数量、面积和周长特征对人口数量的影响,如左家塘街道与湘龙街道,前者的相对误差由20%降低至16%,后者由25%降低至15%,此时的街道规模依然在模型5b能够精细刻画的范围。但对街道人口数为17万人的雨花亭街道,模型5b的估算效果远低于模型4b,这是由于此时的街道规模已经超出模型5b能够修正的范围,对人口数量更为精细的描述反而会造成更大的精度损失。

4 结 语

本文依据城市街道尺度人口数量与街道建筑物分布呈现密切相关的特点,对基于高分辨率遥感影像建筑物信息估算城市街道尺度人口数量进行了研究。通过对街道建筑物数量、面积、周长与楼层各变量之间,以及各变量与人口数量之间的关联性分析,提出了基于街道建筑物信息的城市街道尺度人口估算模型,其中,模型5b具有综合最佳性能,并证明了本文提出的街道建筑物分类方法能够显著提高模型估算精度。以长沙市为例对模型进行了验证,实验结果表明,本文提出的模型能够有效地估算城市街道尺度人口数量。

受研究主题及篇幅限制,本文采用的数据为行政区划中的街道尺度数据,但实际上提出模型的估算范围并不局限于此,理论上在跨街道区域的人口估算应用中同样有效,这一点有待在后续研究中进行验证。

参 考 文 献

- [1] Ding Chengri, Shi Xiaodong, Niu Yi, et al. Urban Population Prediction and Its Significance to Urban Planning: A Case Study of Beijing[J]. *City Planning Review*, 2018(9): 21-27 (丁成日, 石晓冬, 牛毅, 等. 城市人口预测及其城市规划意义——以北京为例[J]. 城市规划, 2018(9): 21-27)
- [2] Feng Tiantian. Urban Small Area Population Estimation Based on High-Resolution Remote Sensing

- Data[D]. Wuhan: Wuhan University, 2010 (冯甜甜. 基于高分辨率遥感数据的城市精细尺度人口估算研究[D]. 武汉: 武汉大学, 2010)
- [3] Yan Qingwu, Bian Zhengfu, Wang Hong. Census Spatialization Based on Thiessen Polygons and Grids Smoothing: A Case Study in Xuzhou[J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(8): 987-990 (闫庆武, 卞正富, 王红. 利用泰森多边形和格网平滑的人口密度空间化研究: 以徐州市为例[J]. 武汉大学学报·信息科学版, 2011, 36(8): 987-990)
- [4] Azar D, Graesser J, Engstrom R, et al. Spatial Refinement of Census Population Distribution Using Remotely Sensed Estimates of Impervious Surfaces in Haiti[J]. *International Journal of Remote Sensing*, 2010, 31(21): 5 635-5 655
- [5] Chen K. An Approach to Linking Remotely Sensed Data and Areal Census Data[J]. *International Journal of Remote Sensing*, 2002, 23(1): 37-48
- [6] Kuang Wenhui, Zhang Shuwen, Zhang Yangzhen. Research on Estimating Urban Land Use Floor Area in Changchun Based on High-Resolution Satellite Images[J]. *Journal of Chongqing Jianzhu University*, 2007, 29(1): 18-21 (匡文慧, 张树文, 张养贞. 基于遥感影像的长春城市用地建筑面积估算[J]. 重庆建筑大学学报, 2007, 29(1): 18-21)
- [7] Yang Cunjian, Bai Zhong, Jia Yuejiang, et al. Study on the Relationship Between Residential Area from Multi-source Remote Sensing Images and Multi-level Population Data[J]. *Geographical Research*, 2009, 28(1): 19-26 (杨存建, 白忠, 贾月江, 等. 基于多源遥感的聚落与多级人口统计数据的关系分析[J]. 地理研究, 2009, 28(1): 19-26)
- [8] Hardin P J, Jackson M W, Jensen R R. Modelling Housing Unit Density from Land Cover Metrics: A Midwestern US Example[J]. *Geocarto International*, 2008, 23(5): 393-411
- [9] Mao Yingying. Study on Urban Population Data Spatialization in City: A Case Study in Downtown Fuzhou[D]. Fuzhou: Fujian Normal University, 2016 (毛莹莹. 城市人口数据空间化研究——以福州市中心城区为例[D]. 福州: 福建师范大学, 2016)
- [10] Zhang Baoguang. Application of Remote Sensing Technology to Population Estimation[J]. *Chinese Geographical Science*, 2003, 13(3): 267-271
- [11] Guo Hongxu, Huang Ying, Zhao Daiqing. Modeling of Spatial Distribution of Urban Population Density: A Case Study of Tianhe District, Guangzhou[J]. *Tropical Geography*, 2013, 33(1): 81-87 (郭洪旭, 黄莹, 赵黛青. 城市居住人口空间分布的模拟研究——以广州市天河区为例[J]. 热带地理, 2013, 33(1): 81-87)
- [12] Harvey J. Small Area Population Estimation Using Satellite Imagery[J]. *Statistics in Transition*, 2000, 4(4): 611-633
- [13] Liu X, Clarke K, Herold M. Population Density and Image Texture: A Comparison Study[J]. *Photogrammetric Engineering and Remote Sensing*, 2006, 72(2): 187-196
- [14] Wu Shuosheng, Wang Le, Qiu Xiaomin. Incorporating GIS Building Data and Census Housing Statistics for Sub-Block Population Estimation[J]. *Professional Geographer*, 2008, 60: 121-135
- [15] Lu Linlin, Guo Huadong. Automatic Estimation of Urban Population with Remote Sensing Imagery Based on ArcGIS[J]. *Remote Sensing Information*, 2008(2): 64-68 (鹿琳琳, 郭华东. 利用遥感影像自动估算深圳福田城市人口[J]. 遥感信息, 2008(2): 64-68)
- [16] Lwin K K, Murayama Y. A GIS Approach to Estimation of Building Population for Micro-spatial Analysis[J]. *Transactions in GIS*, 2009, 13(4): 401-414
- [17] Feng Tiantian, Gong Jianya. Investigation on Small-Area Population Estimation Based on Building Extraction[J]. *Remote Sensing Technology and Application*, 2010(3): 323-327 (冯甜甜, 龚健雅. 基于建筑物提取的精细尺度人口估算研究[J]. 遥感技术与应用, 2010(3): 323-327)
- [18] Ural S, Hussain E, Shan J. Building Population Mapping with Aerial Imagery and GIS Data[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2011, 13(6): 841-852
- [19] Feng Jing. Urban Small Area Population Estimation Based on High-Resolution Remote Sensing Data[D]. Shanghai: East China Normal University, 2012 (封静. 基于高分辨率遥感影像的城市精细尺度人口估算[D]. 上海: 华东师范大学, 2012)
- [20] Leng Jianfei, Gao Xu, Zhu Jiaping. Application of Multiple Linear Regression Statistical Prediction Model[J]. *Statistics and Decision*, 2016(7): 82-85 (冷建飞, 高旭, 朱嘉平. 多元线性回归统计预测模型的应用[J]. 统计与决策, 2016(7): 82-85)
- [21] Yang Ren, Wang Donggen. Studies on the Prediction Model of People Trip Flow in Urban Area[J]. *Journal of Wuhan Technical University of Surveying and Mapping*, 1990, 15(3): 77-83 (杨仁, 王冬根. 近代回归分析在交通调查分析建模中的应用[J]. 武汉测绘科技大学学报, 1990, 15(3): 77-83)

- [22] Hong Yu. Research of Structure Learning Algorithm for Bayesian Network Based on Information Theory [D]. Shanghai: Donghua University, 2016 (洪宇. 基于信息论的贝叶斯网络结构学习算法研究[D]. 上海: 东华大学, 2016)
- [23] Wang Yarong. Tuning Parameter Selection Using ERIC Criterion in the Generalized Linear Model [D]. Lanzhou: Lanzhou University of Finance and Economics, 2018 (王亚荣. 广义线性模型下 ERIC 方法的调节参数选择[D]. 兰州: 兰州财经大学, 2018)
- [24] Sun Liqin, Zhang Fuhao, Yang Shuwen, et al. The Method of Selecting Geographically and Temporally Weight Regression Variable Based on Stepwise Regression [J]. *Science of Surveying and Mapping*, 2019, 44(1): 73-78 (孙立琴, 张福浩, 杨树文, 等. 逐步回归的时空地理加权变量选取方法[J]. 测绘科学, 2019, 44(1): 73-78)

Urban Street Scale Population Estimation Based on Building Information

MENG Yao^{1,2} LI Rui¹ JIANG Jie³ WANG Shunli¹ WU Huayi¹

1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

2 School of Non-commissioned Officer, Space Engineering University, Beijing 102249, China

3 School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, Beijing 102616, China

Abstract: Objectives: With the acceleration of urbanization and the rapid expansion of urban population, the population data on the street scale plays an important role in urban economic, social, resource and environmental development. Therefore, estimating population on urban street scales with high-resolution remote sensing images is of great theoretical and practical value for promoting sustainable urban development. **Methods:** This paper uses the building information in remote sensing images and census data to analyze the correlation between street building information and population data, and the city street scale population estimation models based on building information are proposed. The best estimation model for the number and geometric characteristics of buildings and population is established based on the multivariate stepwise regression method and the Akaike information criterion that determine the significant feature variables of the building. **Results:** The Experiments show that the population estimation model proposed in this paper can estimate the population of street scale with high accuracy. On the one hand, the floor data can be directly applied to the estimation model. On the other hand, the floor data is also the basis for classifying street buildings. Dividing street buildings into low-rise, middle-rise and high-rise buildings according to the number of floors can significantly improve the population estimation effect of the model. **Conclusions:** The population estimation model proposed in this paper can estimate the population of street scale with high accuracy. This paper uses street-scale data in administrative divisions, but the estimation range of the proposed model is theoretically also valid in population estimation across street areas, which needs to be verified in subsequent studies.

Key words: population estimation; remote sensing images; building information; stepwise regression

First author: MENG Yao, master, specializes in space-time computing and data mining. E-mail: yao.meng@whu.edu.cn

Corresponding author: JIANG Jie, PhD, professor. E-mail: jiangjie@bucea.edu.cn

Foundation support: The National Key Research and Development Program of China (2016YFB0502301); the National Natural Science Foundation of China (41771426).

引文格式: MENG Yao, LI Rui, JIANG Jie, et al. Urban Street Scale Population Estimation Based on Building Information[J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(8):1194-1200. DOI:10.13203/j.whugis20190343 (孟瑶, 李锐, 蒋捷, 等. 基于建筑物信息的城市街道尺度人口估算[J]. *武汉大学学报·信息科学版*, 2021, 46(8):1194-1200. DOI:10.13203/j.whugis20190343)