

# 基于多智能体模型与建筑物信息的高空间分辨率人口分布模拟

卓 莉<sup>1</sup>, 黄信锐<sup>1</sup>, 陶海燕<sup>1</sup>, 王 芳<sup>2</sup>, 谢育航<sup>1</sup>

(1. 中山大学地理科学与规划学院, 综合地理信息研究中心,  
广东省城市化与地理环境空间模拟重点实验室, 广州 510275; 2. 广州大学地理科学学院, 广州 510006)

**摘要:** 自上而下的人口分布模拟模型自动化程度较低, 难以分析人口分布成因, 且因精细尺度的人口样本较难获取而不太适用于高空间分辨率人口分布模拟。提出了一种基于多智能体模型和建筑物信息的高空间分辨率人口分布模拟模型。首先利用建筑物三维分布数据提取住宅建筑, 构建智能体人口分布模拟模型的环境; 然后基于统计、调研数据定义智能体属性, 确定智能体居住选择行为规则; 最后以泰日社区为例进行了居住人口分布仿真模拟。研究结果表明, 基于建筑物信息的人口分布多智能体模型, 可以获取每栋建筑物上的人口, 改进了当前高分辨率人口模拟主要只模拟小区或者居委会人口的不足; 多智能体模型具有较高的自动化程度, 不仅能获得较好的模拟结果, 而且可在一定程度上从微观机理解释宏观居住分布模式, 是对传统统计模型的有益补充。

**关键词:** 人口分布; 人口密度; 高空间分辨率; 多智能体; 建筑物

DOI: 10.11821/dlyj201403011

## 1 前言

20世纪以来, 世界人口的骤增加剧了全球范围内的资源短缺、环境污染及生态恶化, 阻碍了社会和经济的可持续发展, 人口问题也因此成为世界面临的核心问题之一<sup>[1]</sup>。获取并掌握人口分布信息, 尤其是精细尺度的人口空间分布信息, 逐渐成为制定区域发展规划、灾害风险防范与救助、经济建设、环境与生态保护等政策的重要科学依据<sup>[2]</sup>。

获取人口分布信息早期主要是通过区域统计或人口调查, 这种方法虽然能获得统计单元上比较准确的人口信息, 但耗时长、工作量大、成本较高<sup>[3]</sup>, 而且忽略统计单元内人口分布的差异, 有可能导致许多与人口分布相关分析的不准确。随着遥感、地理信息系统、人工智能、优化计算等新技术方法的发展, 人口分布模拟方法也得到快速发展。各国学者相继开展了全球尺度、国家尺度、省级尺度、县级尺度乃至像元尺度的人口分布模拟研究<sup>[4-33]</sup>, 并形成了一些具有较大影响力的全球人口空间化产品, 如哥伦比亚大学国际地球科学信息网络中心的GPW (Gridded Population of the World) 人口空间化产品<sup>[30]</sup>、美国国家橡树岭实验室建立的LansScan人口空间分布数据产品<sup>[31]</sup>等。这些涉及多种空间尺度的人

收稿日期: 2013-03-22; 修订日期: 2013-09-23

基金项目: 国家自然科学基金项目 (41371499); 广东省自然科学基金项目 (S2012010010517); 中山大学柳林教授千人计划科研启动项目 (2011-2014)

作者简介: 卓莉, 女, 湖南慈利县人, 副教授, 主要从事地理信息科学与资源环境遥感研究。

E-mail: zhuoli@mail.sysu.edu.cn

口分布模拟研究,为灾害、城市规划与管理等领域的相关研究提供了有效的基础数据。但目前仍然存在一些不足:①目前的人口估算/分布模拟多采用统计回归方法<sup>[37, 16, 17]</sup>,即用与人口相关的指标建立人口估算模型来获取人口的空间分布。统计回归方法虽然简单,但较依赖于样本,普适性较差。并且在估算小尺度的人口时,通常直接将较大统计单元获得的模型用于小尺度人口估算,较少考虑模型的尺度效应。由于规律往往与尺度相关,较大尺度上存在的规律并不一定适用于较小尺度。因此,这种自上而下的推演存在理论上的不完善。②数据处理较多依赖于人工操作,自动化程度相对较低。③高空间分辨率人口分布模拟研究较少,主要只估算了社区或者小区尺度上的人口,而对于更精细尺度上的人口估算则较少涉及。

多智能体模型采用自下而上的模拟思路,具有强大的复杂计算和时空动态模拟功能,在许多领域都得到了较好的应用<sup>[34-40]</sup>。智能体能够模拟人的行为,因此可从模拟人类居住行为的角度获取人口分布。原理上,基于智能体的人口分布模拟不仅能获得宏观上的人口分布模式,而且能在一定程度上从微观角度解释人口分布的成因,有可能改善自上而下统计方法的不足,且具有较高的自动化程度。目前有少数学者开展了相关研究,但主要是模拟居委会、住宅小区上的人口,而对于更精细的尺度则较少涉及。由于高空间分辨率居住人口的分布与建筑物个数、密度、面积、高度、体积、新旧、分布等密切相关<sup>[16, 17]</sup>,因此本文利用建筑物信息,基于多智能体方法建立自下而上的高空间分辨率居住人口分布模拟模型。

## 2 研究方法

人口分布具有时空动态特性,为了简化模型,本文暂时忽略其时间动态,主要探讨静态的居住人口空间分布模拟方法。根据已有研究成果,虽然影响精细尺度居住人口分布的因子很多,但建筑物体积是精细尺度人口分布模拟的最优因子<sup>[16, 17]</sup>。考虑到模型的易于使用与推广,本文着重研究基于建筑物信息的多智能体人口分布模拟模型,并分析模拟效果。建筑物信息主要利用高空间分辨率卫星遥感数据以及立体像对遥感数据,采用面向对象分析方法和立体像对分析方法来获取,这一工作已经在前期研究中完成<sup>[45]</sup>。本文将进一步利用前期已获取的建筑物空间分布数据提取出居住建筑物,并对建筑物进行高度一层数转换;然后基于多智能体原理、人口分布模拟的特点与目标,定义多智能体的属性,确定模拟环境,制定多智能体的居住选择行为规则,以此建立高空间分辨率人口空间分布模拟模型;最后对模拟效果进行分析。

### 2.1 研究区与数据

由于已经对上海金汇镇中部的泰日居委会进行过建筑物提取方法的研究<sup>[45]</sup>,因此继续选择这一区域,基于已有的建筑物三维分布信息进行高空间分辨率人口分布模拟研究。研究所用的主要数据包括:研究区GeoEye-1高空间分辨率全色与多波段的融合影像、泰日居委会建筑物三维空间分布数据、《奉贤年鉴2010》、奉贤区地图。研究区的位置概况及GeoEye-1高空间分辨率全色与多波段融合影像(空间分辨率为0.5 m,图1)。

由于研究区建筑物三维分布数据包含了所有建筑,而非住宅建筑物理论上不居住人口,因此首先将这部分非住宅建筑物从当前数据中剔除。以建筑物三维分布数据为基础,结合奉贤区地图及实地调查,运用面向对象分类方法剔除工厂厂房、行政办公楼、学校教

学楼、其他公共设施楼等非住宅建筑物，并用目视判断进行后处理，以保障住宅建筑物提取的准确率。

目前高分辨率人口估算多采用建筑物高度信息。考虑到建筑物层数比建筑物高度与人口之间的关系更为密切，而且在模型中使用层数信息可以减少建筑物高度反演误差对人口分布模拟精度的影响，同时也可以使模型运算效率得到一定程度的提高，因此，在模型中用建筑物层数代替常用的高度信息。根据中国住宅建筑层高控制规定，并结合研究区的调研数据，建立楼高与楼层的转换规则（表1），将研究区住宅建筑物的高度转换为层数。研究区反演的建筑物高度最大值为22 m，调研所获得的建筑物最高楼层为6层，由于楼顶的天台，储水池等信息也包含在建筑物高度信息中，因此将17-22 m高的建筑物转换为6层建筑物。

为了减少住宅建筑信息提取误差对人口分布模拟精度的影响，对提取的建筑物个数与层数进行精度评价。随机抽取20个样本，以调研数据、街景图作为评价标准，计算住宅建筑物提取个数及楼层信息的正确率 $n/20$ （ $n$ 为住宅建筑物或楼层信息提取正确的个数），计算结果约为93%，精度满足后续人口分布模拟。

2.2 基于多智能体的高空间分辨率人口分布模拟模型

2.2.1 环境设定 因研究区域较小，可认为自然环境大致相同，因此在建立多智能体人口分布模拟模型时，环境因素主要考虑建筑环境，包括住宅价格、建筑密度、楼层、区位、新旧等。其中，住房价格是影响居住选择行为的重要因素，但价格信息相对较难获取，而住房价格往往与建筑物楼层新旧、区位、密度等信息密切相关，因此在建模过程中选择建筑物占地面积、楼层、新旧程度、区位，以及密度这5个指标作为环境图层。环境图层的获取主要基于预处理后的住宅建筑物空间分布数据，为了去除量纲等影响，所有图层均进行归一化处理，归一化处理后的图层见图2。

2.2 基于多智能体的高空间分辨率人口分布模拟模型

2.2.1 环境设定 因研究区域较小，可认为自然环境大致相同，因此在建立多智能体人口分布模拟模型时，环境因素主要考虑建筑环境，包括住宅价格、建筑密度、楼层、区位、新旧等。其中，住房价格是影响居住选择行为的重要因素，但价格信息相对较难获取，而住房价格往往与建筑物楼层新旧、区位、密度等信息密切相关，因此在建模过程中选择建筑物占地面积、楼层、新旧程度、区位，以及密度这5个指标作为环境图层。环境图层的获取主要基于预处理后的住宅建筑物空间分布数据，为了去除量纲等影响，所有图层均进行归一化处理，归一化处理后的图层见图2。

（1）建筑物面积图层。根据公式（1）对居住建筑物面积进行归一化。

$$\alpha_{A,i} = \frac{A_i}{A_{\max}}$$

(1)

式中： $\alpha_{A,i}$ 为建筑物 $i$ 的面积归一化值； $A_i$ 为建筑物 $i$ 基脚轮廓的面积； $A_{\max}$ 为研究区居住建筑物基脚面积最大值；“1, 2, ...,  $n$ ”为不同的建筑物，以下相同。

（2）建筑物层数图层。根据公式（2）对居住建筑物楼层进行归一化。



图1 研究区位置及GeoEye-1全色与多波段融合影像（2009年2月）

Fig. 1 Location and Geoeeye-1 image of the study area

表1 研究小区住宅建筑物高度与层数转换规则

Tab. 1 Transformation rule of buildings' height to floors in the study area

高度(m)	≤4	5-7	8-10	11-13	14-16	17-22
层数	1	2	3	4	5	6

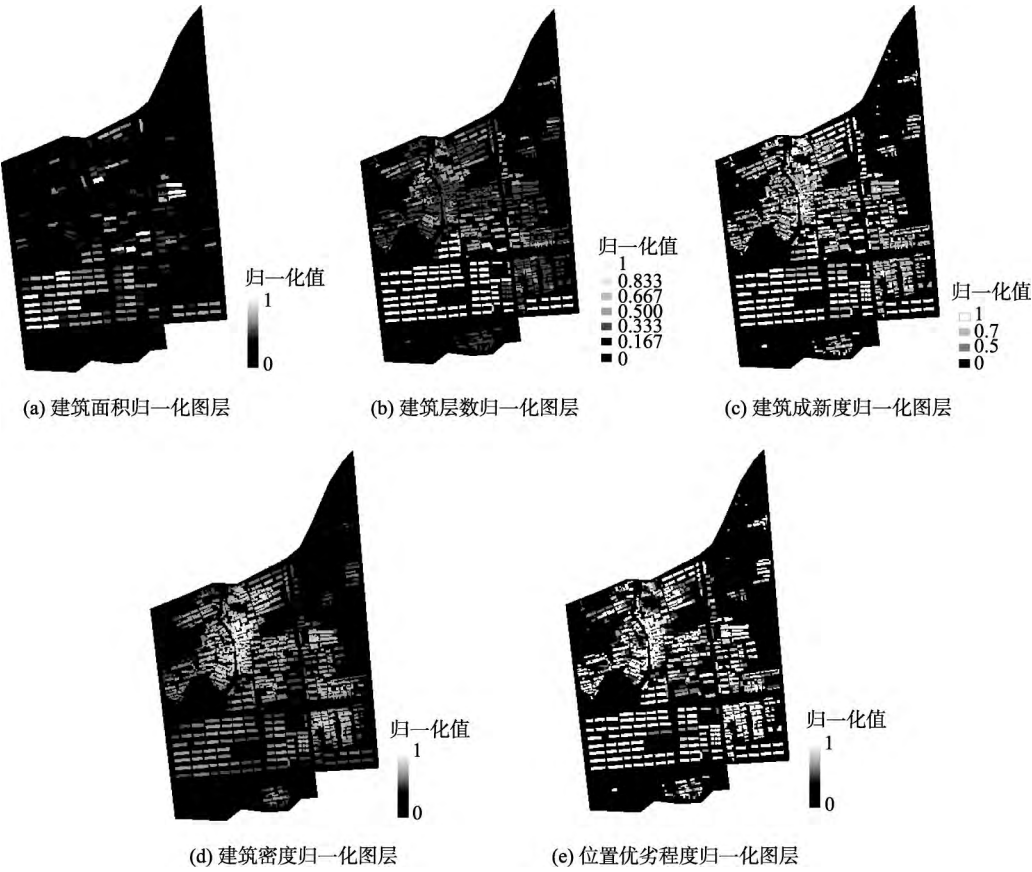


图2 环境图层

Fig. 2 Environment layers

$$\alpha_{F,i} = F_i / F_{\max} \tag{2}$$

式中： $\alpha_{F,i}$  为建筑物*i*的层数归一化值； $F_i$ 为建筑物*i*的层数； $F_{\max}$ 为研究区建筑物层数最大值，为6。

(3) 建筑物成新度图层。依据我国城乡建设环境保护部批准的《房屋完损等级评定标准》(1984年)规定，建筑物成新度的确定主要依据房屋的完损等级。房屋的完损等级可划分为5类：完好房、基本完好房、一般损坏房、严重损坏房和危险房，相对应的成新度对应如下：完好房为十、九、八成新；基本完好房为七、六成新；一般损坏房为五、四成新；严重损坏房及危险房为三成新以下。根据2010年7月对研究区的实地调查可知，研究区住宅建筑物主要为完好房、基本完好房和少数一般损坏房。这些不同等级的房屋在GeoEye-1高分辨率全色与彩色融合后的遥感影像呈现出不同的颜色(图1)。因此结合遥感图像、地域建筑特点及调研，采用监督分类方法将住宅建筑物分为：完好房、基本完好房、一般损坏房三类，并根据完损等级与成新度的对应关系，将这三类房屋的成新度 $\alpha_{N,i}$ 赋值如下：

$$\alpha_{N,i} = \begin{cases} 1 & (\text{完好房}) \\ 0.7 & (\text{基本完好房}) \\ 0.5 & (\text{一般损坏房}) \end{cases} \quad (i = 1, 2, \dots, n) \tag{3}$$



(4) 住宅建筑密度图层。住宅建筑密度是指项目用地范围内所有住宅建筑的基底总面积与规划建设用地面积之比。对于建筑物*i*所在局部区域的密度,则用建筑物*i*为中心、半径为100 m内的建筑面积与圆面积的比值表示,因此,建筑物*i*的密度归一化值 $a_{C,i}$ 可表示为式(4):

$$a_{C,i} = C_{100i} / C_{100\max} \quad (4)$$

式中: $C_{100i}$ 为建筑物*i*为中心;半径为100米内所有住宅建筑物基底面积的总和; $C_{100\max}$ 为所有 $C_{100i}$ 中的最大值。

(5) 住宅位置优劣程度图层。因研究区较小,环境差异不大,故仅从住宅建筑与工厂、医院等非住宅建筑之间的距离来度量居住位置的优劣。根据《上海市城市规划管理技术规定》(2010年),居住建筑物与非居住建筑物的离界最小距离为3~24 m。研究区住宅用地距离非住宅建筑用地的距离一般多于24 m,故设定24 m为居住建筑物与非居住建筑物之间的离界阈值,在离界阈值以内,居住地距离非住宅用地越近,噪声越大,居住优劣程度降低,据此建立建筑物位置优劣程度公式如下:

$$a_{G,i} = \begin{cases} 1 & (r \geq 24) \\ \frac{r}{24} & (r < 24) \end{cases} \quad (5)$$

式中: $r$ 为住宅建筑物*i*平面中心到最近非住宅建筑物平面中心的距离。

**2.2.2 智能体属性定义** 居民的居住选择行为受到许多属性的影响,如收入、职业、年龄、家庭成员等,其中收入对居住选择行为影响较大<sup>[33, 39]</sup>。在本文中主要考虑“收入”属性对个体居住选择行为的影响,其他属性的影响将在后续研究中探讨。收入属性取值为:低收入、中收入、高收入。

**2.2.3 智能体居住地选择行为规则** 智能体居住地选择行为主要受自身属性、以及建筑物吸引力的影响。据此,对于智能体*j*,若其当前所在的建筑物为*i*,则完成一次移动的行为规则可定义如下:

(1) 搜索智能体*j*视野*R*内的所有建筑物,得到建筑物集合 $U_{j1}$ (不包含当前所在建筑物*i*)。不同属性智能体的视野有所不同,结合研究区的实际建筑情况,采用扩展型的摩尔邻域作为人口智能体的邻域,低收入人口智能体邻域 $R_{\text{低}}=10$ ,中收入人口智能体邻域 $R_{\text{中}}=12$ ,高收入人口智能体的邻域 $R_{\text{高}}=15$ 。

(2) 剔除 $U_{j1}$ 中人口数量已达到其上限值的所有建筑物,得到建筑物集合 $U_{j2}$ 。建筑物可容纳人口数量的上限值 $E$ ,用建筑物总面积与人均居住面积之比来表示。分析该区域的统计数据可知,研究区人均居住面积约为20 m<sup>2</sup><sup>[43]</sup>,因此,若 $A_i$ 为建筑物*i*单层的建筑面积,则某建筑物*i*可容纳人口数量的上限值 $E_i$ 可用公式(6)表示为:

$$E_i = \text{Round}\left(\frac{A_i}{20}\right) \times F_i \quad (6)$$

(3) 计算 $U_{j2}$ 中所有建筑物的综合吸引力指数*P*,以及智能体*j*当前所在建筑物*i*的综合吸引力指数 $P_i$ 。建筑物综合吸引力包括两个方面,一是建筑物环境,二是建筑物内的人口状况,即其他智能体对房屋的占据状况。因此,从这两方面确定建筑物的综合吸引力。

假设 $Q_i$ 为建筑物*i*的环境状况对智能体产生的吸引力, $W_i$ 为建筑物*i*内人口状况对智能体产生的吸引力,则建筑物*i*对智能体产生的综合吸引力可定义为:

$$P_i = W_i Q_i \quad (7)$$

$$Q_i = b_1 \alpha_{A,i} \times b_2 \alpha_{F,i} \times b_3 \alpha_{N,i} \times b_4 (1 - \alpha_{C,i}) \times b_5 \alpha_{G,i} \quad (8)$$

$$W_i = 1$$

对于低收入智能体

(9)

$$W_i = \begin{cases} 1 & (D_i \leq \frac{2}{3}E_i) \\ 1 - \frac{D_i}{E_i} & (D_i > \frac{2}{3}E_i) \end{cases}$$

对于中收入智能体

(10)

$$W_i = \begin{cases} 1 & (D_i \leq \frac{1}{2}E_i) \\ 1 - \frac{D_i}{E_i} & (D_i > \frac{1}{2}E_i) \end{cases}$$

对于高收入智能体

(11)

式中： $\alpha_{A,i}$ 、 $\alpha_{F,i}$ 、 $\alpha_{N,i}$ 、 $\alpha_{C,i}$ 、 $\alpha_{G,i}$ 分别为建筑物*i*的面积、层数、新旧、密度及位置优劣程度的归一化值； $b_1$ 、 $b_2$ 、 $b_3$ 、 $b_4$ 、 $b_5$ 为权重值； $D_i$ 为建筑物*i*的当前人口数。

确定权重的常用方法主要有四种<sup>[42]</sup>：人为判别法、排序法、比率法和成对比较法。其中，排序法是确定影响因子权重比较简单而有效的方法，使用排序法，公式如下：

$$b_j = \frac{n - r_j + 1}{\sum_k (n - r_k + 1)}$$

(12)

式中： $b_j$ 为第*j*个影响因子的归一化权重值； $n$ 是影响因子的数量； $r_j$ 是第*j*个影响因子在影响因子排序中的序号。

不同属性的智能体对环境影响因子的排序不同，对不同属性的智能体分别根据公式(12)计算各影响因子的权重，计算结果见表2所示。

(4) 智能体的移动规则。目前智能体的移动一般基于理性选择机制，即智能体选择最具吸引力的建筑物居住。然而，现实情况下，智能体在进行居住选择时，由于多种制约往往难以实现最优的理性选择，即在几个满足需求的可供选择目标中随机选择其一。为了比较这两种选择机制的效果，在模型中分别考虑这两种选择机制。

理性选择机制下智能体的移动规则为：智能体比较建筑物*i*的综合吸引力指数 $P_i$ 与其他建筑物吸引力 $P$ ，若存在大于 $P_i$ 的 $P$ 值，找到 $P$ 的最大值，智能体移动到 $P$ 值最大的建筑物上（若存在多个最大值时，随机选择一个）；若不存在大于 $P_i$ 的 $P$ 值，则智能体不发生移动，但此时产生一个随机因子，使智能体有0.2%的几率随机移动到 $U_j$ 中的某个建筑物上。

有限理性选择机制下智能体的移动规则为：智能体比较 $P_i$ 与其他建筑物的吸引力指数 $P$ 值，当存在大于 $P_i$ 的 $P$ 值时，在这些 $P$ 值中随机选择一个移动到其对应的建筑物上；当不存在大于 $P_i$ 的 $P$ 值时，智能体不发生移动，但此时会产生一个随机因子，使智能体有0.2%的几率随机移动到 $U_j$ 中的某个建筑物上。

**2.2.4 基于多智能体的人口分布模拟模型** 常用的多智能体仿真平台有Netlogo、Swarm、Repast等，这些平台在使用上各有优劣。本文采用Repast多智能体仿真平台，运

表2 环境因子对不同属性智能体的吸引力权重

Tab. 2 Attraction weights of various environment factors to agents with different properties

	$b_1$	$b_2$	$b_3$	$b_4$	$b_5$
低收入	0.2	0.2	0.2	0.2	0.2
中收入	0.333	0.133	0.267	0.067	0.2
高收入	0.067	0.333	0.133	0.2	0.267

用Java语言进行编程,这种方式具有代码简洁、平台独立等特点。在确定好智能体属性、环境以及不同属性智能体的居住选择行为、以及模型的数据格式后,在Repast多智能体仿真平台上,根据图3所示流程建立了基于建筑物信息和多智能体的高分辨率人口分布模拟模型。

3 模型运行及结果分析

根据2.2节中所建模型,以2009年泰日居委会为例进行了人口分布模拟实验。金汇镇泰日居委会位于金汇镇东部地区,占地1.4 km<sup>2</sup>,总户数3875户,常住人口12213人,管辖7个小区,1个老镇区。模型运行前首先需设定一些初始值以及运行结束条件,运行结束后需对模拟结果进行分析,具体如下所述:

(1) 模型初始化。主要包括不同属性智能体数量的初始化和智能体空间分布的初始化。智能体总数量我们设定为研究年份研究区的统计人口总数,此外根据研究区统计资料<sup>[47]</sup>,确定模型中不同属性智能体的数量比为:低收入:中收入:高收入=1:3:1。人口智能体在建筑物*i*内的初始值*S<sub>i</sub>*为区域建筑面积上的平均人口与某建筑物建筑物面积的乘积。

(2) 运行结束条件。当99%以上的智能体不再发生移动时,模型停止运行。

(3) 模拟结果。分别用理性选择与有限理性选择机制进行模拟,获得了2010年上海市金汇镇泰日居委会在两种情形下的人口分布模拟结果,分别如图4和图5所示。

(4) 结果分析。为了对模拟结果进行检验,我们通过实地调查获取了20个建筑物上的人口数量作为模拟结果的参考值,然后找到对应建筑物的模拟结果,根据公式(13)分别计算理性选择与有限理性选择模拟结果的相对误差。样本的参考值、模拟值和相对误差如表3所示。

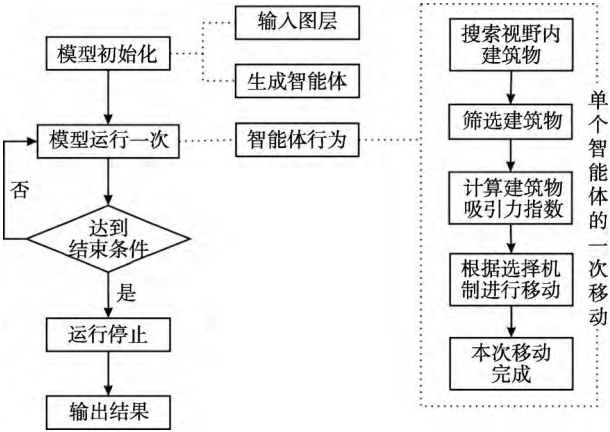


图3 MAS人口分布模拟模型运行流程图

Fig. 3 Flow chart of the multi-agent system based population distribution model

$$E_i = |P_i^m - P_i^t| / P_i^t \tag{13}$$

由表3可知,理性选择模型的相对误差绝对值的平均值为17.84%,其中,有80%的模拟值的相对误差在±30%以内;而有限理性选择模型的相对误差绝对值的平均值为12.00%,其中,有90%的模拟值的相对误差在±30%以内。为了更直观地观察模拟结果误差的分布,以每栋建筑物人口参考值为横轴,以两种选择机制下的模拟结果为纵轴做散点图(图6),从图中可以看出,理性选择在建筑物人口较多(此处为大于60人)的情况下,模拟值大于真实值的可能性增大,这与许多人去竞争“最好”的建筑物有关;而有限理性选择在同样情形下,模拟值大于与小于真实值的概率接近,总体偏离较小。此外,这两种模型都表现出在低值区模拟结果较好,而在高值区模拟结果误差较大,因此还有待于进一步完善模型。总体而言,基于有限理性选择机制的模拟方法获得的模拟结果精度更高,更接近现实分布。

4 结论与讨论

4.1 结论

基于建筑物三维空间分布数据，首先通过剔除非住宅建筑物以及楼高一楼层的转换获取建模所需的居住建筑物密度、楼层、区位等信息，然后根据多智能体基本原理建立了基于多智能体的人口分布模拟模型，最后以此模型模拟了2009年上海市金汇镇泰日居委会的人口空间分布，获得了较好的结果。相比同类型的研究，本文具有以下特色：

(1) 基于多智能体的自下而上的人口分布模拟方法，不仅改善了自上而下人口分布模拟方法存在理论不完善的缺陷，而且有利于从微观上探索人群的居住选择行为，具有一定的创新性。这种自下而上的模拟思路不是仅局限于多智能体的方法，也可以运用其他一些自下而上的模拟方法，不是仅局限于建筑物上的人口模拟，对于像元尺度的模拟也有一定的借鉴意义。

(2) 人口分布模拟的结果精细到了单个建筑物，空间分辨率较高，在当前的人口分布模拟研究中还不多见。这种高分辨率的人口分布模拟结果，更能体现人口居住的空间局部差异性，反映人口的真实分布状况，此外也为需要高空间分辨率人口分布数据的部门或相关研究提供了一个较好的数据源。

(3) 方法自动化程度较高，在完成模型的构建之后，只需要少量的输入数据，不需再交互式的调整即可按设定的优化目标获得最后的人口分布结果。这种较小的数据限制和操作上的简单，使得文中所建模型更加容易推广应用，可为各行业提供服务。

(4) 利用多智能体模型模拟人口分布考虑了居住选择行为的随机性，因而可以模拟得到相同楼层、相似面积的建筑物上具有不同数量的人口，且能通过程序自动实现，改进了采用传统宏观统计模型需多次分配才能得以实现的局限。

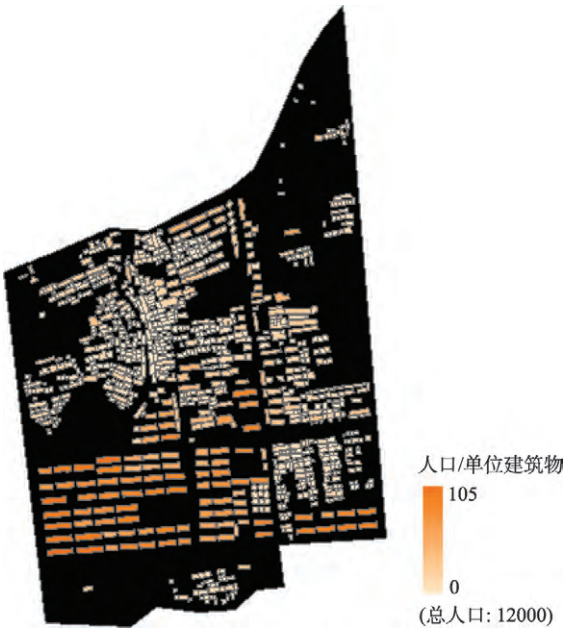


图4 理性选择模型人口分布模拟结果图  
Fig. 4 Result of rational choice model

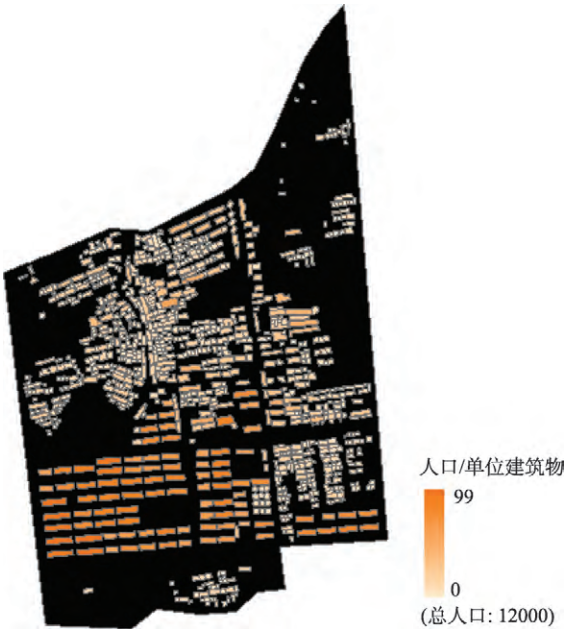


图5 有限理性选择模型人口分布模拟结果图  
Fig. 5 Result of limited rational choice model



4.2 讨论

虽然本文提出的方法对人口分布模拟研究具有一定的贡献，但目前的研究还处于前期探索，仍存在一些不足之处：

(1) 由于影像空间分辨率相对较高，数据量较大，在影像的处理、模型的运行方面都会消耗较多的计算机资源，为了提升计算效率，本文仅选择了小区域作为研究区，此外由于资源和条件的限制，也未能在其他区域验证论文的方法，因此，大范围、高效率的应用推广有待进一步的研究。

(2) 虽然人口分布模拟的精度整体较

好，但存在局部精度不够理想的情况，如在建筑物人口较多的情形下偏差较大。这可能是在模型建立过程中对某些因素考虑不周所致，比如多智能体属性的目前只设置了收入，实际上职业、年龄等属性也会影响居住选择行为，因此在后续研究中将考虑多种属性的影响；此外由于研究区域较小而将所有智能体的自然环境视为同一的假设可能存在不足，因

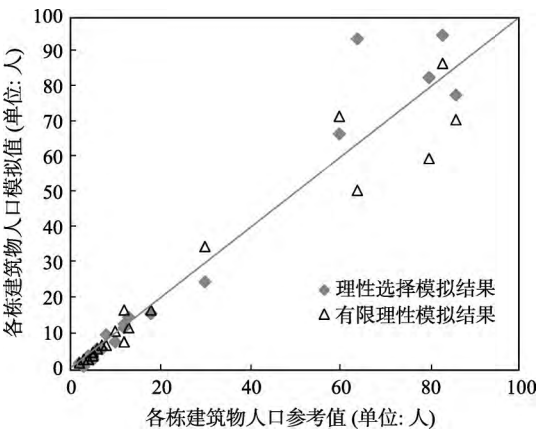


图6 人口分布模拟误差分析  
Fig. 6 Error Analysis of the Simulation

表3 采样值、模拟值与相对误差  
Tab. 3 Comparison of sampled data and simulated results

建筑物 编号	人口 参考值	理性选择 模型模拟值	有限理性选择 模型模拟值	理性选择 模型相对误差	有限理性选择 模型相对误差
1	2	2	2	0.00%	0.00%
2	3	2	3	-33.33%	0.00%
3	3	1	3	-66.67%	0.00%
4	4	4	3	0.00%	-25.00%
5	5	4	4	-20.00%	-20.00%
6	5	3	5	-40.00%	0.00%
7	6	6	6	0.00%	0.00%
8	7	6	7	-14.29%	0.00%
9	8	10	7	25.00%	-12.50%
10	10	8	11	-20.00%	10.00%
11	12	12	8	0.00%	-33.33%
12	12	13	17	8.33%	41.67%
13	13	15	12	15.38%	-7.69%
14	18	16	17	-11.11%	-5.56%
15	30	25	35	-16.67%	16.67%
16	60	67	72	11.67%	20.00%
17	64	94	51	46.88%	-20.31%
18	80	83	60	3.75%	-25.00%
19	83	95	87	14.46%	4.82%
20	86	78	71	-9.30%	-17.44%
绝对平均值				17.84%	13.00%

为即便在小区域自然环境实际上可能还是有所区别, 可以加入植被等信息做进一步区分; 对于区位优势程度, 目前考虑还较为单一, 可以加入比较详细的道路图等信息, 使区位优势程度更准确, 从而有可能获得更加合理的人口分布模拟结果。因此模型还有待随着研究的深入进一步完善, 从而提高模拟精度。

## 参考文献(References)

- [1] 杨彦彦. 人口、资源与环境经济学. 北京: 中国经济出版社, 1999. [Yang Yunyan. The Economics of Population, Resource and Environment. Beijing: China Economic Publishing House, 1999.]
- [2] 中国科学院可持续发展研究组. 中国可持续发展战略报告. 北京: 科学出版社, 2000. [Sustainable Development Strategy Study Group, Chinese Academy of Sciences. China Sustainable Development Report. Beijing: Science Press, 2000.]
- [3] 李素, 庄大方. 基于RS和GIS的人口估计方法研究综述. 地理科学进展, 2006, 25(1): 109-121. [Li Su, Zhuang Dafang. A review on RS- and GIS based population estimation methods. Progress in Geography, 2006, 25(1): 109-121.]
- [4] Lo C P, Welch R. Chinese urban population estimate 1977. Annals of the Association of American Geographers, 1977, 67(2): 246-253.
- [5] Kraus S P, Senger L W. Estimating Population from photographically determined residential land use types. Remote Sensing of Environment, 1974, 3(1): 35-42.
- [6] Lisaka J, Hegedus E. Population estimation from landsat imagery. Remote Sensing of Environment, 1982, 12: 259-272.
- [7] Weber C. Per-zone classification of urban land use cover for urban population estimation in environmental remote sensing from regional to global scales. In: Foody G M, Curran P J (eds.). Chichester, UK: New York, NY: Wiley, 1994: 142-148.
- [8] Sutton P, Roberts D. A comparaton of nighttime satellite imagery and population density for the continental United States. Photogram Metric Engineering and Remote Sensing, 1997, 63(11): 1303-1313.
- [9] Yuan Y., Smith R. M., Limp W. F. Remodeling census population with spatial information from Landsat TM imagery. Computer, Environment and Urban Systems, 1997, 21: 245-258.
- [10] Lo C P. Modeling the population of China using DMSP operational linescan system nighttime data. Photogrammetric Engineering and Remote Sensing, 2001, 67: 1037-1047.
- [11] Langford M, Harvey J T. The use of remotely sensed data for spatial disaggregation of published census population counts. IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, 2001: 260-264.
- [12] Harvey J T. Population estimation models based on individual TM pixels. Photogram Metric Engineering and Remote Sensing, 2002, 68(11): 1181-1192.
- [13] Li G., Weng Q. Using Landsat ETM imagery to measure population density in Indianapolis, Indiana, USA. Photogrammetric Engineering and Remote Sensing, 2005, 71: 947-958.
- [14] Lu D S, Weng Q, Li G Y. Residential population estimation using a remote sensing derived impervious surface approach. International Journal of Remote Sensing, 2006, 27(16): 3553-3570.
- [15] Zhuo L, Ichinose T, Zheng J, et al., Modeling population density of China at pixel level based on DMSP/OLS non-radiance calibrated nighttime light image. International Journal of Remote Sensing, 2009, 30(4): 1003-1018.
- [16] Dong P, Ramesh S, Nepali A. Evaluation of small-area population estimation using LiDAR, Landsat TM and parcel data. International Journal of Remote Sensing, 2010, 31(21): 5571-5586.
- [17] Silva'n-ca'rdenas J L, Wang L, Rogerson P, et al. Assessing fine-spatial-resolution remote sensing for small-area population estimation. International Journal of Remote Sensing, 2010, 31(21): 5605-5634.
- [18] 高志强, 刘纪远, 庄大方. 基于遥感和GIS的中国土地资源生态环境质量同人口分布的关系研究. 遥感学报, 1999, 3(1): 66-70. [Gao Zhiqiang, Liu Jiyuan, Zhuang Dafang. The relations analysis between ecological environmental quality of Chinese land resource and population. Journal of Remote Sensing, 1999, 3(1). 66-70.]
- [19] 杨小唤, 江东, 王乃斌, 等. 人口数据空间化的处理方法. 地理学报, 2002, 57(增刊): 070-075. [Yang Xiaohuan, Jiang Dong, Wang Naibin, et al. Method of pixelizing population data. Acta Geographica Sinica, 2002, 57(supp): 71-75.]
- [20] 刘建军, 李春来. 基于遥感和GIS的巢湖流域人口信息提取. 科学通报, 2002, 47(23): 1835-1837. [Liu Jianjun, Li Chunlai. Population information extraction of the Chaohu Basin based on Remote Sensing and GIS. Chinese Science

- Bulletin, 2002, 47(23): 1835-1837.]
- [21] 冯健. 杭州市人口密度空间分布及其演化的模型研究. 地理研究, 2002, 21(5): 635-646. [Feng Jian. Modeling the spatial distribution of urban population density and its evolution in Hangzhou. Geographical Research, 2002, 21(5): 635-646.]
- [22] 高向东, 吴文钰. 20世纪90年代上海市人口分布变动及模拟. 地理学报, 2005, 60(4): 637-644. [Gao Xiaodong, Wu Wenyu. Population redistribution and modelling of Shanghai City in 1990s. Acta Geographica Sinica, 2005, 60(4): 637-644.]
- [23] 刘纪远, 岳天祥, 王英安, 等. 中国人口密度数字模拟. 地理学报. 2003. 58(1). 17-24. [Liu Jiyan, Yue Tianxiang, Wang Ying'an, et al. Digital simulation of population density in China. Acta Geographica Sinica, 2003, 58(1): 17-24.]
- [24] 廖顺宝, 孙九林. 基于GIS的青藏高原人口统计数据空间化. 地理学报, 2003, 58(1): 25-33. [Liao Shunbao, Sun Jiulin. GIS based spatialization of population census data in Qinghai-Tibet Plateau. Acta Geographica Sinica, 2003, 58(1): 25-33.]
- [25] 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. 地理学报. 2004. 59(2). 283-292. [Tian Yongzhong, Chen Shupeng, Yuetianxiang, et al. Simulation of Chinese population density based on landuse. Acta Geographica Sinica, 2004, 59(2): 283-292.]
- [26] 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. 地理学报, 2005, 60(2). 266-275. [Zhuo Li, Chen Jin, Shi Peijun, et al. Modeling population density of China in 1998 based on DMSP/OLS nighttime light image. Acta Geographica Sinica, 2005, 60(2): 266-275.]
- [27] 吕安民, 李成名. 基于遥感影像的城市人口密度模型. 地理学报, 2005, 60(1): 158-164. [Lu Anmin, Li Chengming. Spatial continuous surface model of population density. Acta Geographica Sinica, 2005, 60(1): 158-164.]
- [28] 杜国明, 张树文. 基于遥感的城乡人口分布模拟. 遥感学报, 2007, 11(2). 252-256. [Du Guoming, Zhang Shuwen. Simulation urban-rural population distribution based remote sensing: Taking Songyuan City as a case study. Journal of Remote Sensing, 2007, 11(2). 252-256.]
- [29] Serkan U, Ejaz H, Jie S. Building population mapping with aerial imagery and GIS data. International Journal of Applied Earth Observation and Geoinformation, 2011, 13(6): 841-852.
- [30] Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT). 2005. Gridded Population of the World, Version 3 (GPWv3). Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University.
- [31] Bhaduri B, Bright E, Coleman P, et al. LandScan USA: A high resolution geospatial and temporal modeling approach for population distribution and dynamics. GeoJournal, 2007, 69: 103-117.
- [32] Gallego F J. A population density grid of the European Union. Population and Environment, 2010, 31: 460-473.
- [33] 冯甜甜, 龚健雅. 基于建筑物提取的精细尺度人口估算研究. 遥感技术与应用, 2010, 25(3): 323-327. [Feng Tiantian, Gong Jianya. Investigation on small-area population estimation based on building extrcation. Remote Sensing Technology and Application, 2010, 25(3): 323-327.]
- [34] 史忠植. 智能主体及其应用. 北京: 科学出版社, 2002. [Shi Zhongzhi. Intelligent Agents and Their Applications. Beijing: Science Press, 2002.]
- [35] Zielinska A L, Jankowski P. Agent-based models as laboratories for spatially explicit planning policities. Environment and Planning B: Planning and Design, 2007, 34(2): 316-335.
- [36] 吴静, 王铮. 2000年来中国人口地理演变的Agent模拟分析. 地理学报, 2008, 63(2): 186-194. [Wu Jing, Wang Zheng. Agent- based simulation on the evolution of population geography of China during the past 2000 years. Acta Geographica Sinica, 2008, 63(2): 186-194.]
- [37] 陶海燕, 黎夏, 陈晓翔, 等. 基于多智能体的地理空间分异现象模拟: 以城市居住空间演变为例. 地理学报, 2007, 62(6): 579-588. [Tao Haiyan, Li Xia, Chen Xiaoxiang, et al. Method exploration of geographical spatial differentiation based on multi-agent: A case study of urban residential simulations. Acta Geographica Sinica, 2007, 62(6): 579-588.]
- [38] 张鸿辉, 曾永年, 金晓斌, 等. 多智能体城市土地扩张模型及其应用. 地理学报, 2008, 63(8): 869-881. [Zhang Honghui, Zeng Yongnian, Jin Xiaobin, et al. Urban land expansion model based on multi-agent system and application. Acta Geographica Sinica, 2008, 63(8): 869-881.]
- [39] 陶海燕, 黎夏, 陈晓翔. 基于多智能体的居住空间格局演变的真实场景模拟. 地理学报, 2009, 64(6): 665-676. [Tao Haiyan, Li Xia, Chen Xiaoxiang. Simulation for evolvement of residential spatial patterns in real scene based on multi-agent. Acta Geographica Sinica, 2009, 64(6): 665-676.]

- [40] 刘小平, 黎夏, 陈逸敏, 等. 基于多智能体的居住区位空间选择模型. 地理学报, 2010, 65(6): 695-707. [Liu Xiaoping, Li Xia, Chen Yimin, et al. Agent-based model of residential location. *Acta Geographica Sinica*, 2010, 65(6): 695-707.]
- [41] 黄信锐. 基于建筑物信息和多智能体的高空间分辨率人口分布模拟. 广州: 中山大学, 2012. [Huang Xinrui. High spatial resolution population distribution simulation. Guangzhou: Sun Yet-sen University, 2012.]
- [42] 黎夏, 刘凯. GIS 与空间分析: 原理和方法. 北京: 科学出版社, 2006. [Li Xia, Liu Kai. GIS and spatial analysis: Theory and methods. Beijing: Science Press, 2006.]
- [43] 上海奉贤区人民政府. 奉贤年鉴. 上海: 学林出版社, 2010. [People's Government of Fengxian District. Shanghai. Yearbook of Fengxian 2020. Shanghai: Xuelin Press, 2010.]

## High spatial resolution population distribution simulation based on building information and multi-agent

ZHUO Li<sup>1</sup>, HUANG Xinrui<sup>1</sup>, TAO Haiyan<sup>1</sup>, WANG Fang<sup>2</sup>, XIE Yuhang<sup>1</sup>

(1. Center of Integrated Geographic Information Analysis, Guangdong Key Laboratory for Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China;

2. School of Geography Sciences, Guangzhou University, Guangzhou 510006, China)

**Abstract:** Knowledge of population distribution is essential for understanding and responding to many social, political, economical and environmental problems. Many studies have been done to estimate population distribution. However, most of the existing methods cannot produce high spatial resolution results or require data which are difficult to acquire. In this paper, we present a method of simulating high spatial resolution population distribution, which uses a multi-agent model based on building information. Firstly, simulation environment of the multi-agent model is constructed based on building information, such as height, area, density; Secondly, the property and behavior rules are defined, based on statistical data and survey reports; Finally, the multi-agent model is used to simulate population distribution of a community named Tairi, which is located in Fengxian District, Shanghai. Results prove that our multi-agent population model is capable of producing high spatial resolution population distribution results at acceptable accuracy level. The method is highly automated and the data it requires are easily available, thereby it has great potential of application.

**Key words:** population distribution; population density; high spatial resolution; multi-agent; building