

城市空间信息规则网格与不规则网格的数据转换

李德仁¹ 彭明军^{2,3}

(1 武汉大学测绘遥感信息工程国家重点实验室,武汉市珞喻路 129 号,430079)
(2 武汉大学遥感信息工程学院,武汉市珞喻路 129 号,430079)
(3 武汉市规划土地管理信息中心,武汉市三阳路 13 号,430014)

摘 要:对从不规则网格向规则网格进行社会经济信息转换的方法进行了比较和研究,并采用蒙特卡罗和 GIS 相结合的方法,以人口数据为例,利用土地利用分类信息将以行政区为基础采集的人口数据转换到规则网格中。实验证明,转换后的人口分布数据较好地体现了人口在空间的分布情况,并可根据不同的需要,聚合或不同大小的网格,能较好地满足城市微观建模和宏观社会经济信息统计分析的需要。

关键词:规则网格;不规则网格;转换

中图法分类号:P208

我国对社会经济信息的统计传统上都按行政区划进行,采用地图的形式进行表达,如按照市-省-国家的方式进行人口数据的汇总,并编制人口分布地图。而相同等级的行政区面积相差很大,按统一的分类和分级标准生成的统计调查数据,在传统地图上的渲染效果可能给人以失真的印象,无法反映社会经济信息与空间自然现象(山川、河流等)之间的关系。以行政区作为统计单元得到的人口分布,是一种按照行政区域统计的人口分布平均值或密度,很难反映自然村、镇的发展状况及其与自然、生态、环境、城市、交通的相互关联,几乎无地理意义。

采用网格的方法则可以避免以上问题。李德仁等指出了从数字地图到空间信息多级网格的发展趋势,提出在我国西部、农村和城市采用 10 km×10 km、1 km×1 km、100 m×100 m 网格进行社会经济数据统计^[1,2],并以人口调查为例,设计了一个带空间信息编码的人口普查多级格网 SIMGPC(spatial information multi grid for population census)应用方案^[3],对全球表面网格划分、网格编码策略、存储方式、网格编码与地理坐标转换方法进行了分析^[4]。以适当大小的网格为单元来表现人口的分布情况,能比较贴切地反映

高原、高山、荒漠、河谷、绿洲、耕地与人口分布的相互关系,可以提高人口统计分析的地理精度,避免在规划、管理工作中的盲目性。行政单元与流域单元的网格化处理,通过网格的四叉树多层结构,可以把自然与人类活动的空间关系完整地表达出来^[5]。

为充分发挥利用规则网格进行社会经济数据宏观统计和分析的优势,需要将目前以不规则网格为基础统计的社会经济数据转换到规则网格中。

1 转换方法

社会经济数据从不规则网格向规则网格转换时,可采用内插方法。内插法是指在一个研究区域内,已知在一种分区(网格)系统中各统计单元的值,求同一研究区内另一种分区系统下各统计单元的值。在两种分区(网格)系统中,各统计单元的边界一般是不兼容的,这两种分区系统分别称为源区(source zones)和目标区(target zones)。下面以人口数据为例,研究人口数据从不规则网格向规则网格的转换方法。如图 1 所示,虚线分割所形成的部分为不规则网格(源区),实线分割所形成的部分为规则网格(目标区)。

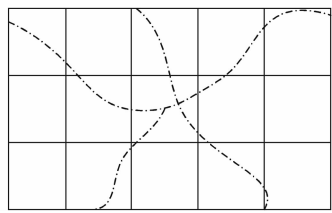


图 1 将不规则网格中的数据转换到规则网格中
Fig. 1 Transferring Social Economic Data from Irregular Grid to Regular Grid

1.1 基于点的内插方法

基于点的内插方法是指统计变量值保持一致的面插值法,它是基于点对多边形的插值^[6,7]。其基本思路是将面域的信息用点来表示,通过这些点进行内插,形成人口密度表面,再将规则网格与人口表面密度数据叠加,得到规则网格内的人口数。

该算法的特点是:① 选择质点位置具有很大的不确定性;② 无法保证源区与目标区的人口数相等;③ 算法相对简单,结果的准确性与源区、目标区的大小和点的分布有关。

1.2 简单面域加权法

简单面域加权的转换方法认为,面域区域内的社会经济数据是均匀分布的,数据的多少只与面积有关,其计算方法也比较简单^[8,9]。其计算方法为:

$$P_t = \sum_{i=1}^n \frac{A'_i}{A_i} P_i = A'_i \cdot D_i$$

其中, P_t 为规则网格中的人口数; A_i 为不规则网格的面积; A'_i 为不规则网格与规则网格相交部分的面积; P_i 为不规则网格中的人口数; D_i 为不规则网格 i 的人口密度; n 为与规则网格相交的不规则网格的数目。

该算法的特点是:① 假定人口在源区内的分布是均匀的;② 能够保证源区与目标区的人口数相等。

该方法的假设不符合实际的人口分布状况,其使用受到很大限制。

1.3 修正面域加权法

修正面域加权法与简单面域加权法的最大差别在于采用了面积以外的辅助信息。该方法认为,研究范围内的人口分布是不均匀的,这种不均匀的分布和城市的土地利用强度密切相关。为此引入密度区的概念,将研究范围划分为若干个密度区,在每个密度区内部,其密度是相同的。设研究范围内有 C 个用地类型, N 个密度区,第 i 个密度区中的第 k 个土地利用类型的权重为 W ,土地

利用图斑的面积为 A_i ,研究区域内的人口总数为 P_s ,则图斑 A_i 中的人口数 P_t 为:

$$P_t = \frac{A_i \times W_i^k}{\sum_{i=1}^m A_i \times W_i^k} P_s$$

将人口分配到每个土地利用图斑后,与规则网格进行叠加,实现人口数据向规则网格的转换。

该方法的特点是:采用控制区的概念将区域划分为不同的匀质区,或引入与社会经济信息分布相关的辅助信息。与简单面域加权相比,提高了精度,并能够保证源区与目标区的人口数相等。

1.4 最大化保留法

最大化保留法(pycnophylatic)是简单面域加权法的扩展,其基本原理就是对面域加权方法计算出的目标区域属性值采用邻近区域的计算结果进行修正^[10]。该方法步骤如下:① 对研究区域生成标准大小的网格;② 对每个网格利用面源比重法进行赋值;③ 对每个网格利用其邻域单元计算结果进行平滑;④ 汇总整个源区域网格,得到源区域属性值;⑤ 对在同一源区域的目标区的单元网格赋相同权重,保证源区域属性值的稳定;⑥ 重复步骤③~⑤,直至达到预定目标,输出的结果可以是栅格、等值线或连续分布的面状地图。

该算法的特点是:① 保证平滑前后统计单元的人口数一致;② 不要求统计单元中变量分布的均质性,但单元边界处变量的较大变化会影响插值的效果,因此边界处的变量值差别不应太大。

1.5 随机方法

蒙特卡罗(Monte Carlo)方法是一种随机方法,将其与 GIS 进行结合,可将人口数据离散化到栅格上。由于栅格的大小是相同的,根据概率统计的基本原理,将数据分配到某一个栅格的概率为:

$$p_i = \frac{W_i^k}{\sum_{i=1}^m W_i^k}$$

分配到某一个栅格的人口数为:

$$P_t = \frac{W_i^k}{\sum_{i=1}^m W_i^k} P_s$$

其中, W_i^k 为第 i 个密度区中的第 k 个土地利用类型的权重; P_s 为总人口数; P_t 为栅格中的人口数。文献[11-13]对该方法进行研究,采用该方法将人口离散到栅格的过程如下。

- 1) 根据城市不同的密度和建设强度,将研究范围划分为若干密度不同的区域。
- 2) 将城市土地利用数据和密度区数据进行

叠加,得到每个用地类型所属的密度区,根据土地利用类型(居住用地、商住用地、商业用地、工业用地、交通用地、绿地、水域等)和密度区,赋予不同的权重。

3) 将叠加后的城市土地利用数据,根据土地利用性质,转换成栅格数据,同时将土地利用类型的权重值也赋给对应的栅格,将所有栅格的权重进行累加,得到权重之和 W 。

4) 对每个栅格的权重值进行顺序累加,设第 1~2 个栅格的权重为 10,第 3~6 个栅格的权重为 20,则第 1 个栅格的取值范围为 1~10,第 2 个栅格的取值范围为 11~20,第 3 个栅格的取值范围为 21~40,依次类推。

5) 从 1~ W 中产生一系列随机数,当随机数落在某一个栅格的取值范围内时,该栅格的值就加 1。重复该过程,直到产生的随机数的个数和研究范围内的人口总数相等为止。

6) 得到离散到栅格的人口数。

7) 将规则网格和栅格进行叠加,将落在规则网格内的栅格中的人口数进行累加,得到规则网格中的人口数,进而实现不规则网格统计的人口数据向规则网格的转换。

使用该方法进行离散化后的人口总数与离散化前的人口总数相同。将人口数据离散到面积较小的栅格后,可根据人口阈值对网格进行聚合,形成不同大小的网格,以表示人口的疏密程度。

采用随机方法实现不规则网格中的数据向规则网格的转换过程中,各种土地利用类型中人口分配的权重是非常重要的,可采用以下方法来确定各类土地利用的权重。

1) 通过外业调查数据或从其他合适的数据源中确定各种土地利用类型的人口比重。

2) 通过统计建模的方法确定各类土地利用类型中的人口分布密度。采用全局回归或区域回归方法得到回归模型,并将回归系数作为最初的密度分配比例,再进行适当的调整。

3) 通过选取样本的方法确定各类土地利用类型中的人口分布密度和比例。

2 实验研究

基于以上分析,本文采用武汉市 2004 年以行政区(街)为基础的人口统计数据 and 武汉市 2004 年城市土地利用数据,利用随机方法,实现人口数据从不规则网格向规则网格的转换。为提高数据转换精度,将房屋数据和道路数据进行叠加,得到

房屋的建筑密度,采用详细的城市土地利用分类数据,提高转换精度。在转换时,采用 $50\text{ m} \times 50\text{ m}$ 的栅格进行人口数据的离散化,再将栅格中的人口数据聚合到规则网格中,实现人口数据从不规则网格向规则网格的转换。

对武汉市 2004 年城市土地利用现状数据进行详细分类,并对各类用地设置不同的权重。设各类土地利用人口分布权重总和为 100,各类用地权重分配为:居住用地为 70,商业用地为 15,工业用地为 5,仓储用地为 5,农村居民点为 5,道路、河流、湖泊、绿地、山体、耕地为 0。封二彩色插图 1 是以行政区(街)显示的 2004 年武汉市人口密度分布图。可以看出,这种方式显示的人口分布是一种平均化的效果,不能真实地反映人口在城市中的空间分布情况。封二彩色插图 2 是 2004 年武汉市城市土地利用现状图。在 ArcGIS (含 Spatial Analyst 模块)中根据土地利用类型,将土地利用数据转换成 $50\text{ m} \times 50\text{ m}$ 大小的栅格,采用随机方法将人口数据离散到 $50\text{ m} \times 50\text{ m}$ 栅格上,如封三彩色插图 3 所示。可以明显看出,中心城区人口多,城郊结合部人口少,城市中河流、湖泊、绿地、道路的人口数为 0,城市中公共活动场所的人口数较少,符合城市人口空间分布的实际情况,离散化之后的人口较好地反映了人口与城市道路、水系以及其他各类城市用地之间的关系。根据应用需要,可采用不同分辨率的规则网格与离散化之后的人口数据进行叠加,并转换到规则网格中。封三彩色插图 4、图 5 分别是将离散到栅格的人口数据采用 $1\text{ km} \times 1\text{ km}$ 、 $100\text{ m} \times 100\text{ m}$ 网络进行转换后显示的结果。

与基于点的内插法、简单面域加权法、修正面域加权法、最大化保留法等方法相比,随机方法具有以下优点:引入了与城市人口分布高度相关的城市土地利用为权重进行人口数据的离散化,通过调整土地利用类型的权重,可以计算白天和夜间的人口分布;可根据需要将人口离散到不同分辨率的栅格上;离散前后的人口总数保持不变等。同时,采用这种方法得到的人口分布具有很强的可视化效果,可以作为地图产品的补充形式;可满足城市微观建模的需要^[14],进行人口在空间上变化规律的分析^[15]。如将两个不同时间的人口采用以上方法离散到 $1\text{ km} \times 1\text{ km}$ 的规则网格后,可采用探测性空间数据分析(exploratory spatial data analysis,ESDA)方法对人口变化规律进行分析。同时,我国目前正处在经济高速发展阶段,由于城市扩张、旧城改造等原因,各种行政边界经

常变化,这种方法相对于采用行政区为统计单元的人口时序分析具有明显的优势。采用这种转换方法,可以充分利用已有的以行政单元为基础的社会经济统计数据,将其转换到规则网格中去。

对于其他社会经济数据,可采用同样的方法实现不规则网格中的社会经济数据向规则网格的转换,如在分析 GDP 的分布情况时,可采用随机方法,将城市中的工业用地、商业用地、仓储用地等设置较大的权重,而居住用地、道路用地等的权重设置为 0,就可以方便地将 GDP 数据离散到栅格上,再根据需要,采用不同分辨率的规则网格进行数据统计。

3 结 语

采用规则网格的方法,可以较好地反映人口在城市中真实的空间分布状况。本文对规则网格与不规则网格之间的社会经济数据的转换方法进行了研究,并采用随机方法,以城市土地利用数据为权重,将武汉市 2004 年按照行政单元(街)统计的人口数据离散到栅格,再采用规则网格进行叠加统计,实现了从不规则网格统计的社会经济数据向规则网格的转换。实验结果表明,离散化后的人口数据体现了人口在城市中的真实分布情况,得到较好的可视化效果,可满足城市微观建模、人口变化规律分析以及宏观社会经济信息统计分析等方面的需要。

参 考 文 献

- [1] 李德仁,朱欣焰,龚健雅. 从数字地图到空间信息网格——空间信息多级网格理论思考[J]. 武汉大学学报·信息科学版,2003,28(12): 642-650
- [2] Li Deren, Zhu Xinyan, Gong Jianya. From Digital Map to Spatial Information Multi-grid[C]. IEEE International Geoscience and Remote Sensing Symposium, Anchorage, USA, 2004
- [3] 李德仁,邵振峰,朱欣焰. 论空间信息多级网格及其典型应用[J]. 武汉大学学报·信息科学版,2004,29(11): 945-950
- [4] 李德仁,肖志峰,朱欣焰,等. 空间信息多级网格的划分方法及编码研究[J]. 测绘学报,2006,35(1): 52-56
- [5] 陈述彭. 人口统计的时空分析[J]. 中国人口·资源与环境,2002,12(4): 3-7
- [6] Okabe A, Sadahiro Y. Variation in Count Data Transferred from a Set of Regular Zones Through the Point-in-polygon Method [J]. International Journal of Geographical Information Science, 1997, 11(1): 93-106
- [7] Lam N S. Spatial Interpolation Methods; a Review [J]. The American Cartographer, 1983(10): 129-150
- [8] Goodchild M F, Lam N S. Areal Interpolation; a Variant of the Traditional Spatial Problem[J]. Geoprocessing, 1980(1): 297-312
- [9] Flowerdew R, Green M. Areal Interpolation and Types of Data[M]//Fotheringham A S, Rogerson P A. Spatial Analysis and GIS. London: Taylor and Francis, 1994
- [10] Tobler W R. Smooth Pycnophylactic Interpolation for Geographical Regions[J]. Journal of the American Statistical Association, 1979(74): 519-530
- [11] Huang Zhengdong. Data Integration for Urban Transport Planning[D]. Utrecht: Utrecht University, 2003
- [12] Spiekermann K, Wegener M. Freedom from the Tyranny of Zones: Towards New GIS-based Spatial Models[M]//Wegener M. Spatial Models and GIS: New Potential and New Models. London: Taylor and Francis, 2000
- [13] Wegener M. New Spatial Planning Models[J]. International Journal of Applied Earth Observation and Geoinformation, 2001, 3(3): 224-237
- [14] Moeckel R, Schurmann C, Wegener M. Microsimulation of Urban Land Use[C]. 42nd European Congress of the Regional Science Association, Dortmund, 2002
- [15] Gregory I N. The Accuracy of Areal Interpolation Techniques: Standardising 19th and 20th Century Census Data to Allow Long-term Comparisons[J]. Computers, Environment and Urban Systems, 2002 (26): 293-314

第一作者简介:李德仁,教授,博士生导师,中国科学院院士,中国工程院院士,欧亚科学院院士。现主要从事以遥感、全球定位系统和地理信息系统为代表的空间信息科学技术的科研和教学工作。代表成果:高精度摄影测量定位理论与方法;GPS辅助空中三角测量;SPOT 卫星像片解析处理;数学形态学及其在测量数据库中的应用;面向对象的 GIS 理论与技术;影像理解及像片自动解译以及多媒体通信等。已发表论文 400 余篇。
E-mail: dli@wtusm.edu.cn

Transformation Between Urban Spatial Information Irregular Grid and Regular Grid

LI Deren¹ PENG Mingjun^{2,3}

- (1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)
- (2 School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)
- (3 Wuhan Urban Planning and Land Administration Information Center, 13 Sanyang Road, Wuhan 430014, China)

Abstract: The multiple areal interpolation methods to transfer social economic data from irregular grid to regular grid are discussed, the advantages and disadvantages of each method are outlined. The integration of Monte Carlo simulation and GIS is employed to realize the transformation of population from irregular grid to regular grid. The population statisticed in irregular grids (administration boundaries) are disaggregated into pixels to obtain the population surface by assigning different weight to different urban landuse types, then the population attached to each pixel is aggregated to regular grid. The experiment result indicates that the regular grids can represent the real situation of population distribution in the urban areas. And different resolution grids can be selected to aggregate the population to meet different requirement, and the population represented in the regular grids can meet the requirement of urban micro modeling and macro social economic data statistics.

Key words: regular grid; irregular grid; transformation

About the first author: LI Deren, professor, Ph. D supervisor, academican of the Chinese Academy of Sciences, academican of the Chinese Academy of Engineering, academican of the Euro-Asia International Academy of Sciences. He is concentrated on the re search and education in spatial information science and technology represented by remote sensing (RS), global positioning system (GPS) and geographic information system (GIS). He has made unique and original contribution in the area s of theories and methods for high precision photogrammetric positioning, GPS aerotriangulation, analysis and processing of SPOT imagery, mathematical morphology and its application in spatial databases, theories of object-oriented GIS, image understanding and automatic photointerpretation, multi-media communication and mobile mapping systems, etc. These research findings have promoted the progress of the technology directly and are being turned into products. His published papers are more than 400.

E-mail: dli@wtusm.edu.cn

我国陆地最低点艾丁湖开始重测

艾丁湖位于新疆吐鲁番市南约 40 km 处,是我国陆地最低点。1979 年测得艾丁湖洼地最低点的海拔高程为-154 m。27 a 后,艾丁湖区域的环境发生了很大的变化,测绘技术也得到了很大的提高,为掌握更加准确的地理信息,新疆测绘部门决定重测艾丁湖的最低海拔高程,这次测量数据将至少精确到 cm。

据悉,这次重测艾丁湖洼地最低点海拔高程为期 6 个月。测绘人员将通过全球定位系统、重力测量、精化大地水准面等目前世界上最专业的技术手段进行重测工作。