

基于半格网错位统计的平面点位密度图制图研究

——以中国县城点密度制图为例

朱瑞 汤国安 吴文博 王丹妮 房彩申

(南京师范大学地理科学学院, 江苏 南京 210046)

摘要: 本文提出了“半格网错位统计法”绘制密度图的原理和方法。以 1:100 万比例尺的全国各县城矢量图为实验数据, 进行了多尺度对比实验和分析验证, 结果证实了该方法绘制出的密度图较为有效。

关键词: 密度制图, 格网统计, 多分辨率

1 序言

密度图作为专题地图的一个类别, 可以反映出多种地理对象在地理空间的分布密度。多种专题密度图在国家的灾害评估、疾病防控、统计决策等各个方面都发挥着巨大的作用。多年来, 基于格网统计的密度图制图方法得到广泛的应用, 例如, 王瑞林(1987)制作的江苏省人口密度图, 王英安(2003)的山东省人口空间分布模拟等, 都较好地展示了研究对象的空间分布特征^[1-2]。但是, 大量的实践也发现, 当采样点集中于统计格网边界时、经过统计生成的二次采样点位于每个统计格网的中心时, 空间内插后必然导致热点地区密度值的降低和密度值的水平偏移。可见, 统计格网的元参数(形状、大小、方位、原点), 都在一定程度上影响制图的精度, 值得进行深入的研究。本文认为, 可以首先从格网的层面数与原点布局入手, 提高制图精度。即: 变原有单层格网统计模式为双层格网, 且通过起始原点的对分错位, 在不参加格网分辨率的前提下, 实现统计采样密度的增加, 提高密度制图的精度。作者将这种制图方法称为“半格网错位统计密度制图法”。

2 原理

在原来的规则格网统计单元基础上, 使用同样分辨率的格网进行半格网错位嵌套。若原来的格网为 M 行 \times N 列, 则错位嵌套格网的行列数同样为 M 行 \times N 列。错位嵌套格网的每个中心点与原统计格网的顶点重合。这样在进行错位嵌套格网的统计时, 就可以对落于原统计

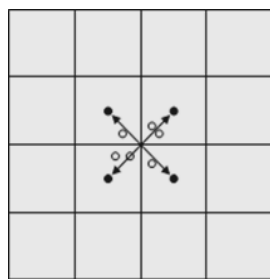


图1 误差原因分析

Fig.1 The reason of the error

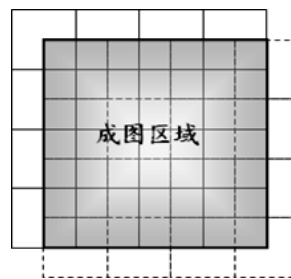


图2 成图区域

Fig.2 The area of the map

1 收稿日期:2008-10; 修定日期:***. 2 基金项目:南京师范大学本科生创新训练项目(1812200013).

3 作者简介:朱瑞, 男, 汉族, 在读本科生, E-mail:911zhurui@163.com.

4 汤国安(1961-), 通讯作者.

格网顶点四周的数据点，进行有效地空间位置约束，高密度地区得以较好的拟合。具体统计时，将每个格网内的采样点总数赋值给该格网中心点，使之成为该中心点的一个属性值。然后将该属性值除以该格网代表的实际面积，得到该格网的平均密度。其次，将错位格网与原统计格网内的二次采样点叠置为一个图层。最后，进行空间内插生成密度图。

3 实验过程

3.1 实验数据

在 ArcGIS9.2 环境下，使用全国 1：100 万比例尺的各县域矢量数据为实验数据。该数据共有数据点 2447 个。总体来看，中国县域在山东、河北、河南、山西和陕西等地及东南沿海分布较密集；而在东、西部大部分地区分布较稀疏并同时伴有“线状”排列现象；全国其他各地的县域分布相对比较平均，局部地区有“聚集”分布现象。由此，该数据点可以代表点状地物应该具有的不同空间分布形态，可以满足实验需要。

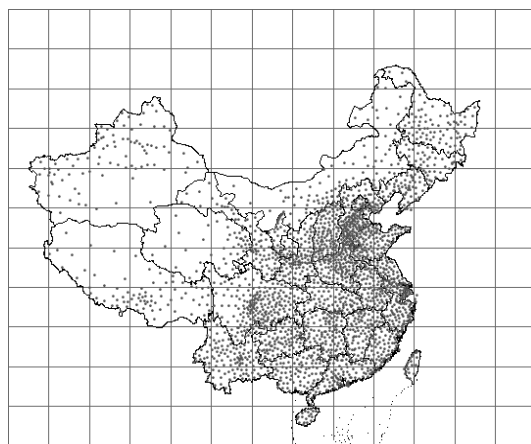


图3 中国县城点状分布

Fig.1 The distribution of Chinese county seats

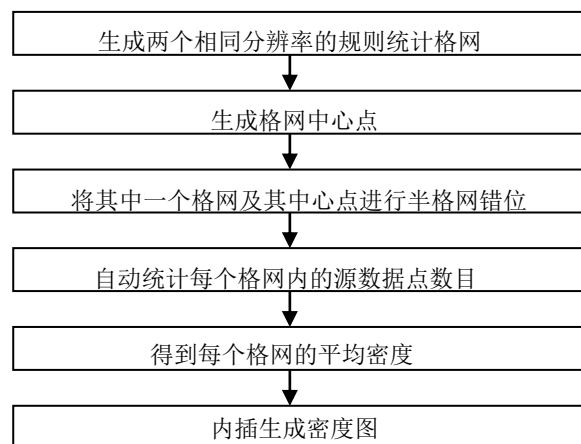


图4 技术流程图

Fig.4 The flow chart of the technical

3.2 实验原则、步骤

实验遵循的原则：①坐标原点统一；②考虑尺度效应对密度图精确性的影响；③生成的目标图层图幅大小保持一致；④统计格网尽量能以最小范围覆盖整个研究区域。

实验主要以网络密度分析法为指导。具体实验时主要有以下几个步骤：①建立相同分辨率的规则统计格网和错位统计格网；②分别自动统计出每个格网内的数据点数目，并且赋值给格网中心点；③用以下公式计算每个格网的平均密度：

$$P_i = \frac{N_i}{S_i} \quad (1)$$

式中 P_i 表示每个格网的平均密度； N_i 表示每个统计格网内的目标点数目； S_i 表示每个格网代表的实际区域面积。④将两个格网的中心点叠置为一个图层；⑤对该图层的中心点进行全局内插生成空间连续分布的密度图。⑤将半格网错位统计法与原格网统计法进行分辨率为 100km 到 800km、100km 等间隔的多尺度对比分析。

4 实验分析

4.1 数据对比分析

为了从理论上证明用该方法生成的密度图密度拟合性更好,首先必须对实验数据作一个评定。标准差 σ 是对一组测量数据可靠性的估计,标准误差值小,表示平均数的变动范围小,如果重复实验,则每次实验的平均数之间的差异小,实验数据的精确度高^[6-7]。

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad (2)$$

式中 n 表示内插点的数目; x_i 表示每个内插点的密度属性值; \bar{x} 表示算术平均值。表3的统计结果显示,在相同统计尺度下,半格网错位统计法生成的二次采样数据,标准差小于传统统计法的标准差。并且当格网分辨率提高到400km时,半格网统计法的数据稳定性已经好于规则格网的最优数据稳定性。

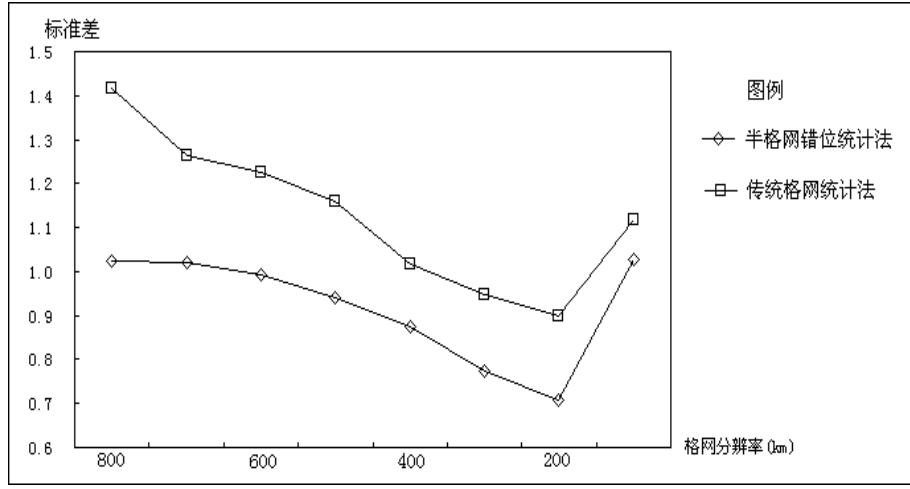


图5 多尺度标准差对比

Fig.5 The comparison of standard deviations based on multiple grids

4.2 密度值对比分析

用检查点法对生成的密度值精度进行评定^[8-11]。根据图5所示,检查点A即为规则格网的内插中心点。这些点均匀分布于整个研究区域,有很好的代表性。将这些点的密度值和内插生成的密度值B、C逐一比较,最后生成中误差(RMSE)折线图。

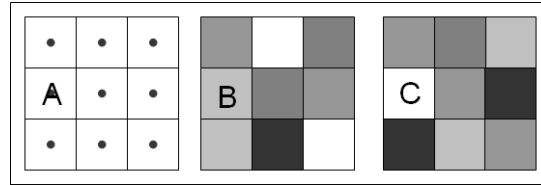


图6 密度值对比分析

Fig.5 The comparison of different densities

设检查点的密度值为 Z_k ($k=1, 2, 3, \dots, n$), 在建立密度图之后, 密度图上与之对应的密度值为 R_k , 则生成的密度图精度为:

$$RMSE = \sqrt{\frac{1}{n-1} \sum_k^n (R_k - Z_k)^2} \quad (3)$$

根据以上公式，在统计格网分辨率为 100-800km、100km 等间隔时，计算使用两种方法获得的密度图精度值。经计算、统计后如图 7 所示，可以清晰地看出在不同尺度下的两条精度折线。随着统计格网分辨率的增大，内插点数目增加，中误差增大。当分辨率小于 500km 时，半格网错位统计法得到的中误差，均略高于单统计格网得到的中误差值。当分辨率为 500km 时，中误差分别为 0.0243 和 0.0287。当分辨率大于 500km 后，半格网错位统计法得到的中误差，均低于单统计格网得到的中误差，并且这一趋势随着分辨率的提高而急剧加大。

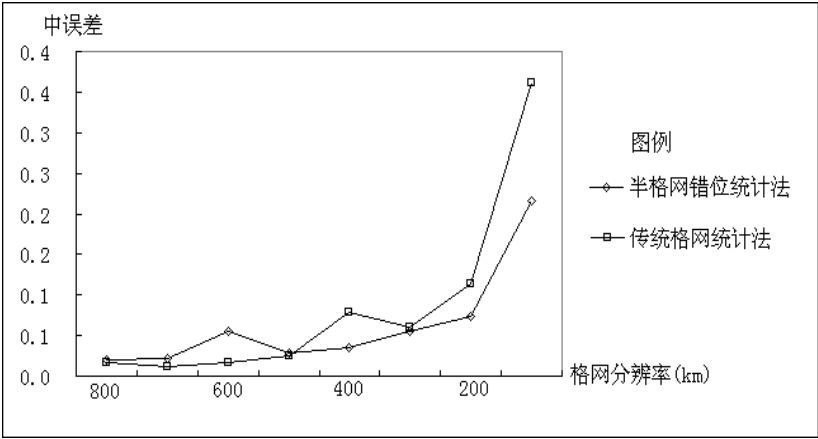


图7 多尺度的中误差对比

Fig.7 The comparison of mean square errors based on multiple grids

由此，经过对比分析，综合空间内插数据的精度、密度图表面的拟合情况以及生成的密度图精确度三者考虑，可以得到当统计格网分辨率为 200km 时，拟合出的中国县域密度图(图 8) 效果最佳。

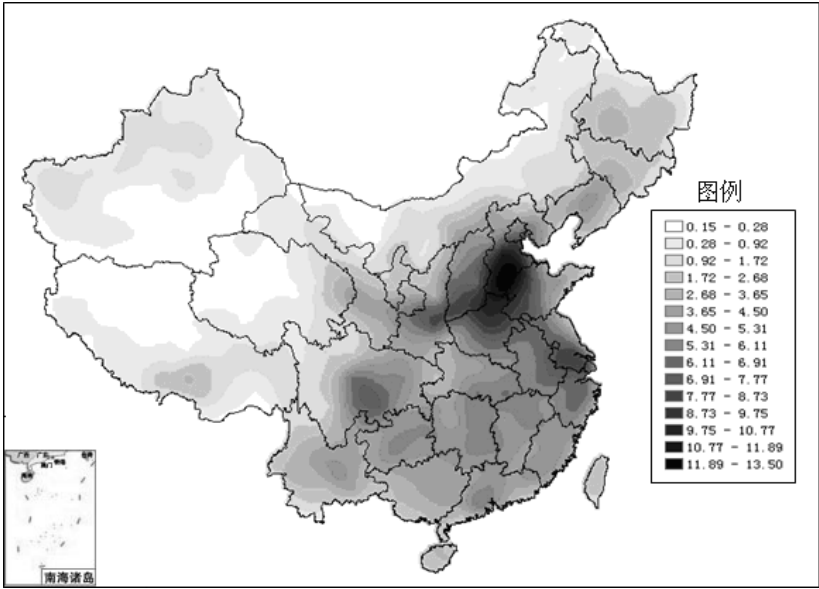


图8 中国县域密度图

Fig.8 The density map of Chinese county seats

5 结束语

以半格网错位法进行密度图制图，在一定程度上克服了以单网格为统计的弊端，提高了

制图精度与图面效果。今后的研究中,应强化对统计栅格的形状、大小、方位、原点等栅格元参数对密度制图的影响研究,完善该方法。

参考文献

- [1] 王瑞林. 格网人口图的计算机制图方法. 地理学报, 1987, 42(4): 357-365.
- [2] 王英安, 岳天祥. 基于格网生成方法的山东省人口密度空间分布模拟. 鲁行经院学报, 2003, 3: 14-16.
- [3] 金君, 印洁, 李成名, 林宗坚. 人口密度推求的技术方法研究. 测绘通报, 2002, 5: 58-60.
- [4] 吕安民, 李成名, 林宗坚, 史文中. 人口密度的空间连续分布模型. 测绘学报, 2003, 32(4): 344-348.
- [5] 李明杰, 钱乐祥, 陈健飞, 肖燕, 彭军超. 聊城市人口密度空间化方法探讨与运用. 广州大学学报(自然科学版), 2008, 7(2): 71-74.
- [6] 刘湘南, 黄方, 王平, 佟志军. GIS空间分析原理与方法. 北京: 科学出版社, 2005. 213.
- [7] 史文中. 空间数据与空间分析不确定性原理. 北京: 科学出版社, 2005. 180-192.
- [8] 汤国安, 刘学军, 闫国年. 数字高程模型及地学分析的原理和方法. 北京: 科学出版社, 2005. 192.
- [9] 王远飞, 何洪林. 空间数据分析方法. 北京: 科学出版社, 2007. 66.
- [10] 舒宁, 马洪超, 孙和利. 模式识别的理论与方法. 武汉: 武汉大学出版社, 2004. 6-8.
- [11] 黎夏, 刘凯. GIS与空间分析——原理与方法. 北京: 科学出版社, 2006. 92-94.

How to Create the Density Map by Means of Staggering Half Length of Statistical Grids

—— Set Chinese County Seats Density Map for Example

ZHU Rui, TANG-Guoan, WU Wen-bo, WANG Dan-ni, FANG Cai-shen

(School of Geographical Science, Nanjing Normal University, Nanjing 210046, China)

Abstract: To begin with, this essay points out the influences when dealing with density maps by means of regulated grids. Secondly, the author creates the principle of *the statistic based on staggering half length grids* and gives the method to solve the problem on mapping density maps. In addition, with the 1 / 1,000,000 Chinese country seats vector map, we compared the new method with the original method by means of different scales and proved that it is indeed more accurate than before.

Key words: Grids, Different Scales, Density Mapping