



# 空间自相关支撑下的地类分布模式一致性评价

罗芳<sup>1</sup> 艾廷华<sup>2</sup> 贾小斌<sup>3</sup>

1 深圳市规划和自然资源局调查测绘中心,广东 深圳,518034

2 武汉大学资源与环境科学学院,湖北 武汉,430079

3 华图国际数字智能科技深圳有限公司,广东 深圳,518000

**摘要:**空间分布模式是否保持一致是土地利用数据综合质量评价的一项重要内容。针对当前的研究缺少量化分析和位置表达的现状,提出了一种新的空间数据特有的自相关性评价方法。首先利用语义距离建立空间权重矩阵,随后通过莫兰指数(Moran's  $I$ )计算数据处理前后全局和局部自相关度,最后利用莫兰(Moran)散点图和空间关联的局部指标(local indicators of spatial association, LISA)集聚图相结合的方法对综合前后的土地利用分布模式进行可视化对比。相较传统评价方法,所提方法顾及数据语义关系,计算可量化聚集程度,以直观可视化方法对比展示,更好地对土地利用数据在综合前后的全局空间分布模式一致性进行了评价。认知实验结果符合人类认知,表明所提方法切实有效。

**关键词:**空间自相关;一致性评价;分布模式;可视化

**中图分类号:**P208

**文献标志码:**A

土地利用数据通过土地利用分类(简称地类)图斑来表达各种分类的空间位置、分布模式、利用类型、数量质量以及权属界线等信息,具有空间全覆盖、无重叠、无缝隙以及语义上多层次划分的特点<sup>[1]</sup>。针对土地利用数据在综合前后位置精度、属性精度、完整性保持、面积一致性保持和时间精度等5个方面<sup>[2]</sup>,目前已有不少研究。位置精度方面,文献[3]基于极化变换对点群综合进行质量评价,文献[4]利用傅里叶变换对数据进行形状识别,文献[5]对空间点群目标进行相似度计算。完整性方面,文献[6]对土地利用数据综合前后各地类面积变化情况进行评价;文献[7]量化区分微观、中观和宏观3个尺度,从图形位置精度、拓扑一致性、面积平衡性、语义精度等方面对土地利用数据综合进行评价。面积一致性方面;文献[8]利用地类面积一致性对土地利用数据进行评价,提出一种尺度变换过程中数据质量变化的宏观控制方法。

但是高质量的土地利用数据综合结果除了图斑数量减少、几何形状简化以及空间上全覆盖、无重叠、无缝隙等特征保留之外,尤其需要保

持各种地类在空间分布上的特征<sup>[9]</sup>。文献[10]从土地利用景观格局层面设计土地利用数据综合的指标描述模型;文献[11]利用结构相似度、布局相似度、图斑适宜度指数来衡量土地利用现状图缩编的质量。但是这些研究基本上是利用综合前后各种地类的面积值进行对比计算,缺乏基于空间位置的描述。

空间自相关统计作为一种经典的地理学理论方法,广泛应用于空间统计、时空格局分析、空间分布特征探测等方面<sup>[12-13]</sup>,包括不同时期空间自相关分析结果比较、宏观对象的地理分布状态研究<sup>[14]</sup>以及要素空间关联模式随时间变化规律的探索<sup>[15]</sup>等。地类在空间上一般呈现聚集、离散和随机3种分布模式,具有显著的空间自相关性,综合前后3种分布模式的一致性土地利用数据综合质量评价比较直观的依据。

本文基于空间自相关理论,提出了一种顾及语义特征的可量化计算聚集程度的评价方法,并利用莫兰(Moran)散点图和空间关联的局部指标(local indicators of spatial association, LISA)集聚图可视化表达,在整个分类和整幅地图两个层次

收稿日期:2021-04-07

项目资助:自然资源部城市国土资源监测与仿真重点实验室开放基金(KF-2020-05-0076)。

第一作者:罗芳,博士,高级工程师,主要从事测绘数据质量的研究、检测和评价工作。whulfgis@163.com

通讯作者:贾小斌,博士。Ji Xiaobin\_123@126.com

上对地类的空间分布模式一致性进行评价,并选取土地利用综合数据进行实验。与传统的基于地类面积值的评价方法相比,本文的可视化表达方法可以从空间位置上直观体现地类分布模式的保持情况,同时,本文顾及地类的语义特征改良了空间权重矩阵,使得评价方法更适用于土地利用数据。

## 1 基于空间自相关的分布模式评价

### 1.1 空间自相关

Tobler<sup>[16]</sup>在地理学第一定律概念中对空间自相关进行了说明,即地理实体及其语义在空间上具有相关性,距离较近比较远关联性更强。空间自相关性研究属于探索空间数据分布的一种方法<sup>[17]</sup>,可用来探测聚类、离散和随机3种空间数据分布模式。空间自相关在层次上可分为全局和局部2个指标<sup>[18]</sup>:全局自相关指标研究区域内某一属性相关性,局部自相关指标研究区域内某一单元与周围邻居单元就某一属性的空间相关性。空间自相关从性质上可进一步分为正相关和负相关2种情形:正相关表明变化趋势相同,如某一

属性空间分布上呈现低的地方周围也低,高的地方周围也高;负相关则说明变化趋势相反。

### 1.2 基于语义距离的空间权重矩阵

空间权重矩阵是空间自相关分析中量化观测值之间空间依赖关系的重要工具,定义了空间单元之间相互作用的结构和强度<sup>[19]</sup>。对数据空间的相关性进行计算应考虑空间权重矩阵与属性数据真实的空间结构是否吻合,不应使用单一的空间权重矩阵<sup>[20]</sup>。土地利用数据具有很强的语义性,综合中应尽量保持地类分布稳定,即综合前后各地类总面积应该差异不大。

为更真实准确地量化土地利用数据空间自相关性,本文基于各地类之间的语义距离计算空间权重矩阵。针对土地利用中的8个地类,分别建立对应的语义距离序列。如耕地根据其语义邻近程度,表达为序列{耕地,园地、林地、草地、城镇村及工矿用地、交通运输用地、其他土地、水域及水利设施用地}。对任一地类 $C_i$ ,根据语义序列中与其他地类的邻近关系得到语义距离集合 $\{F_1, F_2, F_3, \dots, F_n\}$ ,随后依此建立 $n$ 维向量空间,维数 $X_i$ 代表语义距离集合元素 $F_i$ 。由此,建立基于语义距离的空间权重矩阵,如表1所示。

表1 土地利用数据空间权重矩阵

Tab.1 Spatial Weight Matrix of Land Use Data

地类	耕地	园地	林地	草地	城镇村及工矿用地	交通运输用地	水域及水利设施用地	其他用地
耕地	1	0.88	0.75	0.63	0.50	0.38	0.13	0.25
园地	0.88	1	0.88	0.75	0.63	0.50	0.25	0.38
林地	0.75	0.88	1	0.88	0.75	0.63	0.38	0.50
草地	0.63	0.75	0.88	1	0.88	0.75	0.50	0.63
城镇村及工矿用地	0.50	0.63	0.75	0.88	1	0.88	0.63	0.75
交通运输用地	0.38	0.50	0.63	0.75	0.88	1	0.75	0.88
水域及水利设施用地	0.13	0.25	0.38	0.50	0.63	0.75	1	0.88
其他用地	0.25	0.38	0.50	0.63	0.75	0.88	0.88	1

### 1.3 空间自相关指数

选取莫兰指数(Moran's  $I$ )计算数据的空间自相关程度,量化评价空间分布模式。

1)全局 Moran's  $I$ 。将某区域 $A$ 的全局空间自相关 Moran's  $I$ 指数简写为 $I_A^G$ ,计算方法为:

$$I_A^G = \frac{n}{S_0} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

式中, $x_i$ 为区域 $A$ 中空间单元 $i$ 的某一属性观测值; $\bar{x}$ 为所有空间单元观测值的期望值; $w_{ij}$ 表示空间单元 $i$ 和 $j$ 自相关程度的空间权重; $S_0$ 为区域 $A$

中所有空间单元与其他单元的空间权重之和,即

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}; n \text{ 为区域 } A \text{ 中空间单元的总个数。}$$

2)局部 Moran's  $I$ 。局部空间自相关用来度量属性值在特定局部区域的自相关程度。将某区域 $A$ 局部单元 $i$ 的空间自相关指数 Moran's  $I_i$ 简写为 $I_i^L$ ,计算公式为:

$$I_i^L = \sum_{j=1}^n \sum_{j \neq i}^n w_{ij} z_i z_j \quad (2)$$

式中, $z_i$ 和 $z_j$ 分别代表单元 $i$ 和 $j$ 观测值的均值标准化;权重矩阵元素 $w_{ij}$ 标准化后为 $\sum_{i=1}^n \sum_{j \neq i}^n w_{ij} = n$ ;

$I_i^L$  值表示空间单元  $i$  与邻居单元观测值的加权平均乘积。根据  $I_A^G$  和  $I_i^L$  的表达式及定义,两者的关系表示为:

$$I_A^G = \frac{\sum_{i=1}^n \sum_{j \neq i}^n w_{ij} z_i z_j}{S^2 \sum_{i=1}^n \sum_{j \neq i}^n w_{ij}} = \frac{1}{n} \sum_{i=1}^n (z_i \sum_{j \neq i}^n w_{ij} z_j) = \frac{1}{n} \sum_{i=1}^n I_i^L \quad (3)$$

1.4 基于 Moran's I 的分布模式评价

1) 局部空间分布模式评价。定义空间单元  $i$  的观测值评价函数为  $f_E(i)$ , 公式为:

$$f_E(i) = \begin{cases} H, x_i \geq \bar{x} \\ L, x_i < \bar{x} \\ 0, x_i = \bar{x} \end{cases} \quad (4)$$

取值 H 时, 观测值较大; 取值 L 时, 观测值较小。

同理, 定义空间单元  $i$  的邻居单元集合  $N$  观测值为  $f_N(i)$ , 评价函数为  $f_E(iN)$ , 公式为:

$$f_E(iN) = \begin{cases} H, f_N(i) \geq \bar{x} \\ L, f_N(i) < \bar{x} \\ 0, f_N(i) = \bar{x} \end{cases} \quad (5)$$

取值 H 时, 观测值较大; 取值 L 时, 观测值较小。

定义某区域  $A$  的局部单元  $i$  的空间自相关评价 (spatial autocorrelation evaluation, SAE) 函数为  $f_{SAE}(i)$ , 公式为:

$$f_{SAE}(i) = \begin{cases} HH, f_E(i) = H \text{ 且 } f_E(iN) = H \\ LL, f_E(i) = L \text{ 且 } f_E(iN) = L \\ HL, f_E(i) = H \text{ 且 } f_E(iN) = L \\ LH, f_E(i) = L \text{ 且 } f_E(iN) = H \\ R, f_E(i) f_E(iN) = 0 \end{cases} \quad (6)$$

若  $f_{SAE}(i) = HH$ , 则空间单元  $i$  与邻居单元呈高高聚集分布模式; 若  $f_{SAE}(i) = LL$ , 则呈低低聚集分布模式; 若  $f_{SAE}(i) = HL$ , 则呈高低聚集分布模式; 若  $f_{SAE}(i) = LH$ , 则呈低高聚集分布模式; 若  $f_{SAE}(i) = R$ , 则呈随机分布模式。

2) 全局空间分布模式评价。区域  $A$  全局空间自相关 Moran's I 指数  $I_A^G$ , 根据定义及构成, 其取值范围为  $[-1, 1]$ ; 进一步依据显著性水平,  $f_{SAE}(A)$  的计算公式为:

$$f_{SAE}(A) = \begin{cases} HH // LL, 0 < I_A^G \leq 1 \\ HL // LH, -1 \leq I_A^G < 0 \\ R, I_A^G = 0 \end{cases} \quad (7)$$

当  $0 < I_A^G \leq 1$  时, 区域显著空间正相关, 若  $f_{SAE}(A) = HH$ , 则为高高聚集模式; 若  $f_{SAE}(A) = LL$ , 则为低低聚集模式, 最后形成一种空间聚集模式,  $I_A^G$  值越大, 该属性空间对象正相关性越大。当  $-1 \leq I_A^G < 0$  时, 区域显著空间负相关, 若  $f_{SAE}(A) = HL$ , 则为高低聚集模式; 若  $f_{SAE}(A) = LH$ , 则为低高聚集空模式, 最后形成一种空间分散模式,  $I_A^G$  值越小, 该属性空间对象负相关性越大。当  $I_A^G = 0$  时, 则区域不存在空间自相关, 形成一种随机分布模式。

2 分布模式评价的可视化方法

本文使用 Moran 散点图和 LISA 聚集图对数据空间分布模式进行可视化展示。Moran 散点图用来表达全局自相关, 展示全局范围内的空间单元就某一属性在不同时点上的聚集结构。LISA 聚集图用来表达局部自相关, 反映局部具体空间位置的聚集显著度。空间聚集模式通常分为 4 种, 可用 4 象限直观展示<sup>[21]</sup>, 如图 1 所示。坐标原点  $\bar{x}$  代表区域内所有空间单元某一属性的观测值期望, 空间单元  $i$  的观测值  $x_i$  表示横坐标, 空间单元  $i$  的邻居单元坐标集合  $f_N(i)$  表示纵坐标。结合式 (7) 不难看出, 当  $f_{SAE}(i) = HH$  时, 空间单元  $i$  落入第一象限, 与邻居单元呈现一种高高聚集分布模式; 当  $f_{SAE}(i) = LH$  时, 空间单元  $i$  落入第二象限, 呈现一种低高聚集分布模式; 当  $f_{SAE}(i) = LL$  时, 空间单元  $i$  落入第三象限, 呈现一种低低聚集分布模式; 当  $f_{SAE}(i) = HL$  时, 空间单元  $i$  落入第四象限, 呈现一种高低聚集分布模式; 当  $f_{SAE}(i) = R$  时, 空间单元  $i$  与原点重合, 呈现一种随机分布模式。通过 4 个象限中的散点密度可判断空间主要聚集形态。

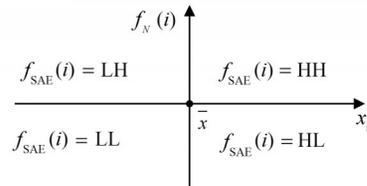


图 1 Moran 散点图的 4 种空间聚集类型  
Fig.1 4 Types of Spatial Aggregation in Moran Scatter Plot

3 认知实验与讨论

地类图斑一般呈现聚集、分散和随机 3 种分布模式, 对区域依赖性很强, 地类在不同区域表

现出不同聚集分布模式。如建设用地在城市聚集分布,而在农村表现出离散特征;耕地在城市离散分布,而在农村表现出聚集性。数据综合应保持同一地类在处理前后聚集模式的一致。

本文基于空间自相关方法进行综合前后地类分布模式的一致性评价。全局方面,比较区域内综合前后各地类分布模式的变化情况,并定量研究不同区域的地类分布整体结构特点;局部方面,分析对比区域内典型地类分布模式的变化,并可视化展示。选择位于城市(区域1)和农村(区域2)两个典型片区的数据进行实验,原始数

据比例尺为1:1万,综合后为1:10万和1:25万,变换操作主要是小图斑综合。

认知实验步骤如下:首先,采用规则空间网格取样,将各地类离散化到大小为500 m×500 m网格中;然后,计算综合前后各网格内各地类面积占比;再利用各网格内面积占比计算全局与局部Moran's *I*;最后制作Moran散点图,进行全局与局部空间自相关评价及可视化分析。区域1中8种地类的Moran's *I*值如表2所示,绘制区域1的相应空间自相关图(见图2~4)全面评价综合前后各地类空间分布模式的保持情况。

表2 综合前后各地类Moran's *I*值及排序

Tab.2 Moran's *I* and Ranking Before and After Land Types Scale Transformation

空间权重矩阵	用地类型Moran's <i>I</i> 值及排序	耕地	园地	林地	草地	交通运输用地	水域及水利设施用地	其他土地	城镇村及工矿用地
Queen式邻接矩阵	1:1万	Moran's <i>I</i> 值 0.527 4	0.482 5	0.676 9	0.445 6	0.437 6	0.533 4	0.414 2	0.635 2
		排序 4	5	1	6	7	3	8	2
	1:10万	Moran's <i>I</i> 值 0.659 8	0.585 3	0.696 7	0.535 1	0.449 6	0.534 9	0.435 1	0.655 7
		排序 2	4	1	5	7	6	8	3
语义距离权重矩阵	1:25万	Moran's <i>I</i> 值 0.668 1	0.673 2	0.699 3	0.556 3	0.510 2	0.569 2	0.493 6	0.658 5
		排序 3	2	1	6	7	5	8	4
	1:1万	Moran's <i>I</i> 值 0.563 9	0.385 6	0.736 5	0.383 4	0.357 1	0.582 7	0.218 7	0.612 2
		排序 4	5	1	6	7	3	8	2
Queen式邻接矩阵	1:10万	Moran's <i>I</i> 值 0.579 8	0.420 6	0.755 8	0.353 4	0.359 5	0.579 9	0.242 3	0.634 3
		排序 4	5	1	7	6	3	8	2
	1:25万	Moran's <i>I</i> 值 0.675 0	0.500 2	0.759 7	0.426 1	0.360 2	0.597 8	0.287 7	0.628 8
		排序 2	5	1	6	7	4	8	3

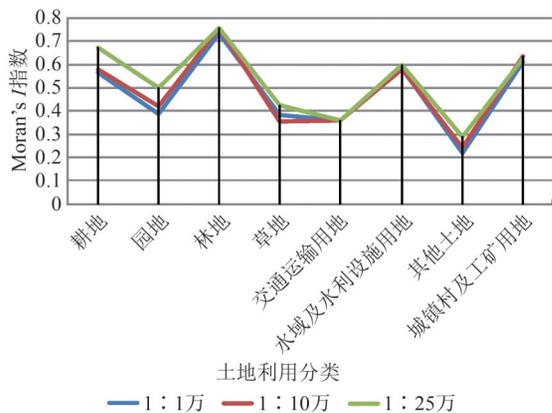


图2 综合前后各地类空间自相关指数变化趋势对比  
Fig.2 Comparison of Spatial Autocorrelation Index Before and After Land Use Data Scale Transformation

3.1 全局空间自相关评价及可视化

采用两种空间权重矩阵进行计算,一种是本文采用的语义距离空间权重矩阵,另一种是Queen式邻接矩阵。表2展示了使用两种空间权重矩阵计算出的各地类综合前后的Moran's *I*值

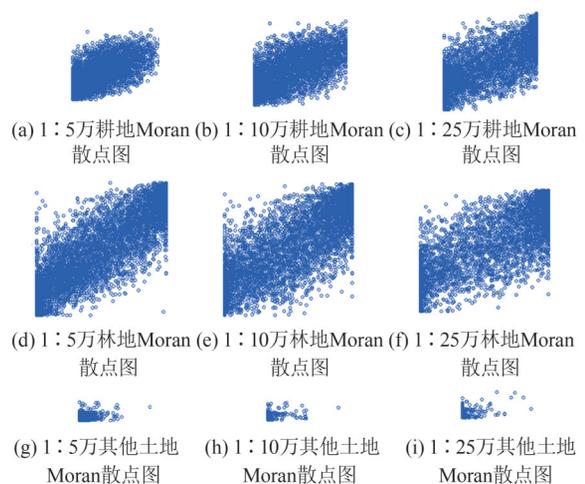


图3 1:5万、1:10万、1:25万典型地类Moran散点图可视化

Fig.3 Scale 1:50 000 to 1:250 000 of Moran Scatter Plot Visualization for Land Use Types Distribution Characteristics

及排序。对比两种空间权重矩阵的计算结果分析发现:(1)Queen式邻接矩阵计算出来的各地类

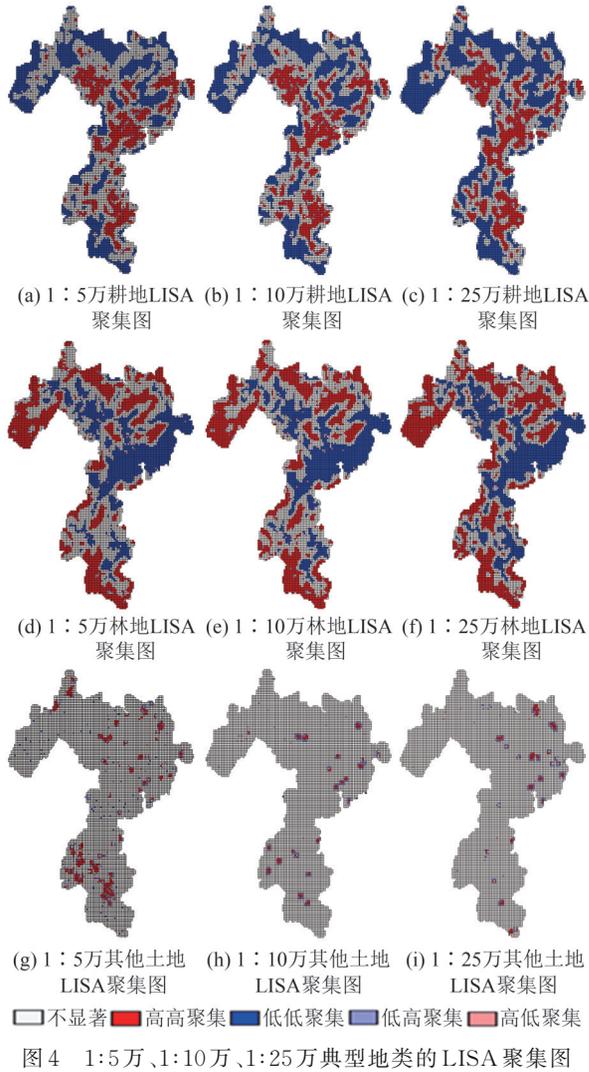


图 4 1:5 万、1:10 万、1:25 万典型地类的 LISA 聚集图可视化  
 Fig.4 Scale 1:50 000 to 1:250 000 of LISA Aggregation Map Visualization for Land Use Types Distribution Characteristics

Moran's  $I$  值比较相近,综合前后都在 0.4~0.7 之间,且综合程度越高,各地类的 Moran's  $I$  值差异越小,较难体现出各地类在综合前后聚集程度的变化;(2)在综合过程中,除了交通运输用地和其他土地之外的 6 种地类的 Queen 式邻接矩阵的 Moran's  $I$  值过于接近,以至于排序发生了紊乱,明显与综合规律的认知不符。可见该方法不太适用于土地利用数据。

分析表 2 中语义距离空间权重矩阵的计算结果发现:(1)8 种地类的 Moran's  $I$  值均为正,呈全局空间正相关,表明 8 种地类在空间上均呈聚集分布特征,但是聚集程度差异较大;(2)由 1:1 万变换到 1:10 万时,各地类的 Moran's  $I$  值都增加,说明本文方法可以正确地反映综合变换的规律,但是排序基本保持不变,说明该阶段的综合质量

较好;(3)变换到 1:25 万时有 4 种地类排序发生变化,说明尺度变化越大,地类分布在综合前后的一致性越差,这与综合的一般认知相符;(4)3 种尺度下林地、城镇村及工矿用地、水域及水利设施用地、耕地 4 类用地的 Moran's  $I$  值始终位居前列,验证了综合中聚集性强的图斑更容易“吞食”聚集性弱的小图斑,呈现“强者恒强”的分布格局。图 2 中反映出各地类 Moran's  $I$  值随着比例尺变小呈增加趋势,说明综合导致聚集效应增强,这与实际认知相符。计算两种变换操作下 Moran's  $I$  值差值,可看出变换到 1:10 万过程中,草地和水域及水利设施用地聚集程度不升反降,出现异常,说明其变换处理存在问题。分析综合操作算法,发现是由于草地及水利设施用地的不规则小图斑过多,综合程度过大。

以上分析说明,基于语义距离空间权重矩阵计算全局空间自相关值可以量化体现地类在综合前后的聚集程度,辅助判断地类在综合前后空间分布模式是否保持一致,进而对变换操作的质量进行评价。

### 3.2 局部空间自相关评价及可视化

选取 Moran's  $I$  值最大的林地、居中的耕地、最小的其他土地作为代表性地类,制作其 Moran 散点图和 LISA 聚集图分别如图 3、图 4 所示。可以看出 3 种地类分布模式在综合前后的一致性都较好,具体表现在:(1)纵向上比较 3 种地类发现其空间分布特征差异较大,Moran 散点图显示林地较多地分布于第一和第三象限,LISA 聚集图也显示林地主要呈高高聚集和低低聚集分布,两个图呈现的结果一致;而其他土地的 Moran 散点图主要分布于第一象限,但 LISA 聚集图显示其主要呈不显著分布,而不是高高聚集,两种图呈现结果不一致,说明两种图结合使用的必要性;(2)横向上比较各地类在 3 个比例尺下的 Moran 散点图,可以看出,随着比例尺变小,位于第一象限的点都增多了,体现了比例尺变小导致聚集效应增强;(3)横向比较各地类在 3 个比例尺下的 LISA 聚集图,发现林地 HH 类型大幅增加,而其他土地 HH 类型大幅减少,体现了由于综合过程大面积图斑不断“吞食”小面积图斑,导致大面积分布的地类聚集效应会不断增强。上述分析说明 Moran 散点图和 LISA 聚集图能直观展示综合过程中各地类的分布特征,从而评价分布模式有无发生不正常变异。

为验证不同类型区域的适用性,本文采用同

样方法分析区域2(地类以农用地为主),综合比例尺为从1:5万到1:10万,耕地和城镇村及工矿

用地综合前后空间分布对应的Moran散点图及LISA聚集图见图5。

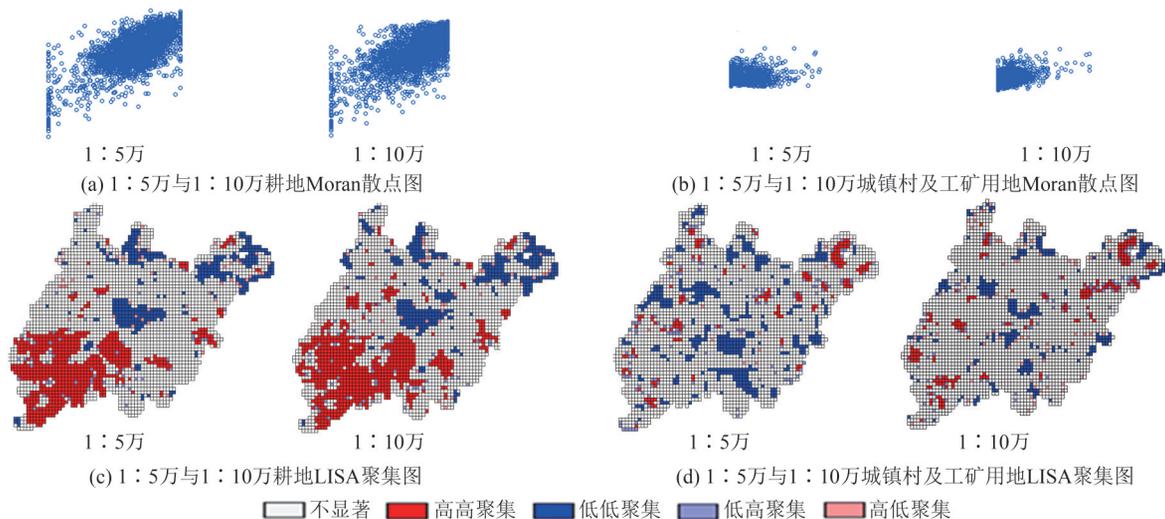


图5 1:5万和1:10万区域2耕地和城镇村及工矿用地Moran散点图及LISA聚集图可视化

Fig.5 Scale 1:50 000 and 1:100 000 of Moran Scatter and LISA Aggregation Map Visualization for Cultivated Land, Urban Village and Industrial-Mining Land in Area 2

从图5中可以看出:(1)综合前后,区域2中耕地的Moran散点图大多分布于第一象限,呈现局部空间正自相关,LISA聚集图中主要为高高聚集类型,两种图的可视化表达结果一致,说明综合前后空间分布模式一致性较好。(2)城镇村及工矿用地大多分布于第一象限,LISA聚集图中主要为不显著类型,两种图的可视化表达结果不一致,这说明了两种方法结合使用的必要性,也说明聚集特征越明显的地类,可视化表达效果越好。(3)综合前,耕地的全局Moran's  $I$ 指数为0.487 3,城镇村及工矿用地的全局Moran's  $I$ 指数为0.166 4,说明耕地空间分布比城镇村及工矿用地表现出更强的聚集性。同时,综合后两者的全局Moran's  $I$ 指数均变大,分别为0.514 2和0.205 2,验证了空间聚集特征会随综合而增强的结论,说明本文方法可量化表达区域的地类分布总体特征。

## 4 结 语

土地利用数据是一种专题地图数据,以数字形式来体现土地利用现状的调查成果,图斑的地类、面积、分布是其核心表达内容,高质量的土地利用数据综合结果需要保持各种地类在空间分布上的特征。综合处理前后地类空间分布模式一致性的判断,对于综合数据质量评价非常重要,然而该方面量化及可视化的研究相对欠缺。针对土地利用数据在语义上具有较强的空间自

相关特征,本文提出了一种地类分布模式一致性保持评价方法:首先利用各地类的语义距离建立空间权重矩阵,随后计算综合前后的全局和局部自相关程度,最后采用Moran散点图和LISA聚集图相结合的方法对地类空间分布模式一致性保持进行可视化表达。认知实验中选取2个不同性质区域的土地利用数据进行验证,并对比分析了两种空间权重矩阵,实验结果表明,本文方法以量化的结果体现区域内不同地类的分布特征和不同区域的地类总体特征,进而通过可视化的方式对综合前后的地类分布模式进行表达,对比分析各地类的分布特征是否得到了保持。实验结果表明本文方法实用有效,结果符合人类认知;但本文中数据质量评价的分布模式类型仅为聚集、离散、分布模式,稍显单一,有待进一步深入研究。

## 参 考 文 献

- [1] Bader M, Weibel R. Detecting and Resolving Size and Proximity Conflicts in the Generalization of Polygonal Maps [C]//The 18th International Cartographic Conference, Stockholm, Sweden, 1997
- [2] GoodChild M. Attribute Accuracy [M]//Guptill S C, Morrison J L. Elements of Spatial Data Quality. Oxford: Progamon, 1995
- [3] Qian Haizhong, Wu Fang, Zhang Linlin, et al. Quality Assessment of Point Group Geometry Generalization with Polarization Transformation[J]. *Acta Geo-*

- daetica et Cartographica Sinica*, 2005, 34 (4) : 361-369 (钱海忠,武芳,张琳琳,等. 基于极化变换的点群综合几何质量评估[J]. 测绘学报,2005,34(4): 361-369)
- [4] Liu Pengcheng. Applications of Shape Recognition in Map Generalization[D]. Wuhan: Wuhan University, 2009(刘鹏程. 形状识别在地图综合中的应用研究[D]. 武汉:武汉大学, 2009)
- [5] Liu Tao, Du Qingyun, Yan Haowen. Spatial Similarity Assessment of Point Clusters[J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(10): 1150-1153 (刘涛, 杜清运, 闫浩文. 空间点群目标相似度计算[J]. 武汉大学学报·信息科学版, 2011, 36(10): 1150-1153)
- [6] Ai Tinghua, Yang Fan, Li Jingzhong. Land-Use Data Generalization for the Database Construction of the Second Land Resource Survey [J]. *Geomatics and Information Science of Wuhan University*, 2010, 35(8): 9-13 (艾廷华, 杨帆, 李精忠. 第二次土地资源调查数据建库中的土地利用图综合缩编[J]. 武汉大学学报·信息科学版, 2010, 35(8): 9-13)
- [7] Luo Fang. Quality Evaluation for Land-Use Data Generalization[D]. Wuhan: Wuhan University, 2013 (罗芳. 土地利用数据综合结果的质量评价[D]. 武汉:武汉大学, 2013)
- [8] Jia Xiaobing. Agent Technology Applied in Map Generalization: Case Study for Land-Use Patch[D]. Wuhan: Wuhan University, 2013 (贾小斌. 地图综合的智能体技术: 以土地利用图斑综合为例[D]. 武汉:武汉大学, 2013)
- [9] Huang Wanli, Li Hu, Lin Guangfa, et al. A Comprehensive Study of Land Use Type Data of Scale Change [J]. *Journal of Geo-informatin Science*, 2010, 12(3): 329-335 (黄万里, 李虎, 林广发, 等. 尺度变化的土地利用类型数据的综合研究[J]. 地球信息科学学报, 2010, 12(3): 329-335)
- [10] Liu Yaolin, Jiao Limin. Land Use Data Generalization Indices Based on Scale and Landscape Pattern [J]. *Acta Geodaetica et Cartographica Sinica*, 2009, 38(6): 549-555 (刘耀林, 焦利民. 顾及尺度效应和景观格局的土地利用数据综合指标研究[J]. 测绘学报, 2009, 38(6): 549-555)
- [11] Shi Zhikuan, Tang Guoan. Quality Assessment of Land Use Data Generalization[J]. *Geographical Research*, 2011, 30(2): 233-242 (石志宽, 汤国安. 土地利用现状数据缩编的质量评价方法[J]. 地理研究, 2011, 30(2): 233-242)
- [12] Anselin L. Spatial Econometrics: Methods and Models [M]. Netherlands: Kluwer Academic Publishers, 1988
- [13] Zhao Yong. The Origin and Application History of Statistical Analysis of Spatial Data[J]. *Geographical Research*, 2018, 37(10): 2059-2074 (赵永. 空间数据统计分析的思想起源与应用演化[J]. 地理研究, 2018, 37(10): 2059-2074)
- [14] Han Chunmeng, Liu Huiping, Zhang Yanghua. et al. Multi-scale Spatial Distribution Analysis of Leisure Agriculture in Beijing Based on Kernel Density Estimation [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2019, 35(6): 271-278 (韩春萌, 刘慧平, 张洋华, 等. 基于核密度函数的多尺度北京市休闲农业空间分布分析[J]. 农业工程学报, 2019, 35(6): 271-278)
- [15] Jiang Botao, Wang Yandong, Ye Xinyue. Detecting Development Pattern of Urban Business Facilities Using Reviews Data[J]. *Acta Geodaetica et Cartographica Sinica*, 2015, 44(9): 1022-1028 (蒋波涛, 王艳东, 叶信岳. 使用点评数据探测城市商业服务设施的发展规律[J]. 测绘学报, 2015, 44(9): 1022-1028)
- [16] Tobler W R. A Computer Movie Simulating Urban Growth in the Detroit Region[J], *Economic Geography*, 1970, 46: 234-240
- [17] Hubert L J, Golledge R G, Costanzo C M. Generalized Procedures for Evaluating Spatial Autocorrelation [J]. *Geographical Analysis*, 1981, 13(3): 224-233
- [18] Ma Ronghua, Pu Yingxia, Ma Xiaodong. GIS Spatial Association Pattern Discovery [M]. Beijing: Science Press, 2007 (马荣华, 蒲英霞, 马晓冬. GIS空间关联模式发现[M]. 北京:科学出版社, 2007)
- [19] Zhang Keyun, Wang Yujin, Wang Jing. Research on the Specification Methods of Spatial Weight Matrix [J]. *Regional Economic Review*, 2017 (1) : 19-25 (张可云, 王裕瑾, 王婧. 空间权重矩阵的设定方法研究[J]. 区域经济评论, 2017(1): 19-25)
- [20] Zhang Yongshu, Wang Pei. Regional Economic Spatial Autocorrelation Analysis Based on Multiple Spatial Weight Matrices: A Case Study of Henan Province [J]. *Geomatics and Spatial Information Technology*, 2020, 43(1): 45-49 (张永树, 王培. 基于多种空间权重矩阵的区域经济空间自相关分析: 以河南省为例[J]. 测绘与空间地理信息, 2020, 43(1): 45-49)
- [21] Cliff A, Ord J. Spatial Processes: Models and Applications [M]. London: Pion, 1981

## Consistency Evaluation of Land Use Distribution Pattern Supported by Spatial Autocorrelation

LUO Fang<sup>1</sup> AI Tinghua<sup>2</sup> JIA Xiaobin<sup>3</sup>

1 Planning and Natural Resources Survey Center of Shenzhen Municipality, Shenzhen 518034, China

2 School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China

3 Huatu International Digital Intelligent Technology Shenzhen Co. Ltd, Shenzhen 518000, China

**Abstract: Objectives:** The consistency of spatial distribution pattern before and after generalization is an important factor to evaluate the quality of land use data. Considering the lack of quantitative analysis and visualization in spatial distribution model evaluation of data, we put forward a new evaluation method based on the unique autocorrelation of spatial data. **Methods:** Firstly, we establish the spatial weight matrix using the semantic distance of data and calculate global and local autocorrelation of data before and after processing by Moran's  $I$  index. Secondly, we use Moran scatter plot and LISA (local indicators of spatial association) aggregation map to visualize the quality of spatial distribution patterns. **Results:** Compared with the traditional data quality evaluation method, the proposed method can better evaluate the consistency of spatial distribution pattern of global map by taking into account the semantic relation of the data, obtaining the quantifiable aggregation degree and contrasting by the visual method. **Conclusions:** In cognitive experiments, we take the land type data been scale transformed which cause data quality problems mostly as an example. The experimental results accord with human cognitions and the experimental method is practical.

**Key words:** spatial autocorrelation; consistency evaluation; distribution pattern; visualization

**First author:** LUO Fang, PhD, senior engineer, specializes in surveying and mapping data quality. E-mail: whulfgis@163.com

**Corresponding author:** JIA Xiaobin, PhD. E-mail:jjaxiaobin\_123@126.com

**Foundation support:** The Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources (KF-2020-05-0076).

**引文格式:** LUO Fang, AI Tinghua, JIA Xiaobin. Consistency Evaluation of Land Use Distribution Pattern Supported by Spatial Autocorrelation[J]. Geomatics and Information Science of Wuhan University, 2022, 47(7):1017-1024. DOI:10.13203/j.whugis20200179 (罗芳, 艾廷华, 贾小斌. 空间自相关支撑下的地类分布模式一致性评价[J]. 武汉大学学报·信息科学版, 2022, 47(7):1017-1024. DOI:10.13203/j.whugis20200179)