

# 空间统计分析方法与 GIS 在区域经济分析中的应用

陈 斐<sup>1</sup> 杜道生<sup>1</sup>

(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)

**摘 要:** 首先概述了度量空间自相关、空间关联的一些空间统计分析方法以及识别区域空间关联的标准, 然后探讨了将空间统计分析嵌入到一个 GIS 系统中的可行性, 开发了一个分析空间关联的模块, 并举例说明其在区域经济分析中的应用。

**关键词:** 空间统计分析; GIS; 空间自相关; 空间关联; 区域经济分析

**中图法分类号:** P208

目前, 空间统计学已广泛应用于农业、地质、土壤、水文、环境、经济以及地理等领域<sup>[1, 2]</sup>。不少学者<sup>[3~11]</sup>先后对空间统计的一些基本理论进行了广泛的研究, 形成了一些新的统计方法。

GIS 技术作为一种交互式的、可视化的决策支持工具, 和传统分析方法、分析系统一样, 在区域经济分析中发挥着重要的作用<sup>[12~14]</sup>, 特别是应用于地方、区域和全国各级水平的经济发展决策过程<sup>[12, 15, 16]</sup>。

近年来, 不少学者致力于空间统计学与 GIS 的结合在社会经济发展领域的应用研究<sup>[17~21]</sup>。这种结合集中体现于空间统计分析在区域发展分析中的应用<sup>[22]</sup>, 但目前 GIS 在空间统计分析方面的能力有限<sup>[23, 24]</sup>, 限制了 GIS 应用到区域经济分析领域。对于究竟如何将空间统计分析方法与 GIS 结合在一起, 不同的研究者提出了不同的观点<sup>[17~21, 23~26]</sup>, 大体上认为空间统计分析方法与 GIS 的集成应包括两个完全不同但同样有效的解决办法: ①将空间统计分析功能嵌入到一个 GIS 中; ②将选定的 GIS 功能嵌入到一个空间统计分析环境中。

## 1 空间统计分析方法

在地理情形分析中, 空间统计的基本出发点是空间抽样。通常认为一个区域单元上的某种地理现象或某一属性值是与其邻近区域单元上同一现

象或属性值相关的<sup>[3, 4, 7]</sup>。几乎所有空间数据都具有空间依赖或空间自相关特征<sup>[26]</sup>。空间依赖的存在打破了大多数古典统计分析中相互独立的基本假设, 将古典统计方法应用于与地理位置相关的数据时, 通常不能获取这些数据的空间依赖性, 会引起各种问题, 而大多数的城市和区域分析的完成依赖于不同地理区域或地带的离散化数据集<sup>[25]</sup>。因此, 在处理地理区域或地带的离散化数据时, 需要引入一些合适的空间统计分析方法。

在本文中, 空间统计指的是一个狭义上的定义, 指适合处理地理区域或地带的离散化数据的统计方法, 而不是所有分析空间数据的统计方法。在这种情况下, 空间统计分析的核心是认识与地理位置相关的数据间的空间依赖、空间关联或空间自相关<sup>[3, 4, 7~10]</sup>, 涉及空间权重矩阵的构建、空间自相关、空间关联的度量与检验、空间关联的识别等。空间统计并不是抛弃所有的古典统计技术, 而是对这些技术加以修改以使它们能够适用于空间数据分析<sup>[20, 21]</sup>。

### 1.1 空间权重矩阵

由 GIS 生成的拓扑信息, 提供了空间链接或空间邻近的基本度量。空间链接或空间邻近广泛应用于空间数据分析中<sup>[18, 19]</sup>。通常定义一个二元对称空间权重矩阵  $W_{n \times n}$  来表达  $n$  个位置的空间邻近关系, 可以根据邻接标准或距离标准来度量, 还可以根据属性值  $x_i$  和二元空间权重矩阵来定义一个加权空间邻近度量方法。根据邻接标

准,当  $i$  和  $j$  邻接时,空间权重矩阵的元素  $w_{ij}=1$ , 否则  $w_{ij}=0$ 。根据距离标准,当位置  $i$  和  $j$  之间的距离在一给定的距离 ( $d$ ) 范围内时,空间权重矩阵  $W$  的元素  $w_{ij}=1$ , 否则  $w_{ij}=0$ 。习惯上将  $W$  的所有对角线元素  $w_{ij}$  设为 0。

### 1.2 空间自相关、空间关联的度量

空间自相关、空间关联反映的是一个区域单元上的某种地理现象或某一属性值与邻近区域单元上同一现象或属性值的相关程度<sup>[3,4,7]</sup>。可以使用全局指标和局部指标两种不同等级的指标加以度量。

Moran 系数(MC)和 Geary 比率(GR)是两个用来度量空间自相关的全局指标<sup>[3,6,7]</sup>。Moran 系数反映的是空间邻接或空间邻近的区域单元属性值的相似程度。空间权重矩阵  $W$  确定了位置相似性,  $(x_i - \bar{x})(x_j - \bar{x})$  反映属性相似性, 确定了位置邻近关系  $w_{ij}$  和属性相似性  $c_{ij}$ , 就可以计算出全局 Moran 系数。

全局指标仅仅使用一个单一的值来反映一定范围内的自相关, 很难发现存在于不同位置区域的空间关联模式<sup>[10,27]</sup>。  $G_i(d)$  统计、局部 Moran 和局部 Geary 统计是可选择的局部指标<sup>[9,10,27]</sup>。

$G_i(d)$  统计可以定义如下:

$$G_i(d) = \left[ \sum_{j, j \neq i}^n w_{ij} x_j \right] \left/ \sum_{j, j \neq i}^n x_j \right. \quad (1)$$

式中,  $x_j$  为位置  $j$  的某一属性值, 二元对称空间权重矩阵  $W$  基于距离标准构建。为了便于解释,

定义  $Z(G_i) = \frac{G_i - E_i(G_i)}{\sqrt{\text{var}(G_i)}}$ ,  $Z(G_i)$  为  $G_i(d)$  的一种标准化的形式, 具体参见文献 9, 10, 27]。

根据 Getis 和 Ord<sup>[8]</sup>, 对于每个位置  $i$ , 局部 Moran 统计、局部 Geary 统计分别可以定义为:

$$I_i = (Z_i / S^2) \sum_{j \neq i}^n w_{ij} Z_j \quad (2)$$

$$C_i = \sum_{j \neq i}^n w_{ij} (Z_i - Z_j)^2 \quad (3)$$

式中,  $Z_i$  和  $Z_j$  为位置  $i$  和位置  $j$  的属性值与均值的偏差, 即  $Z_i = x_i - \bar{x}$ ;  $w_{ij}$  是二元对称空间权重矩阵的元素;  $S^2 = \sum_j (x_j - \bar{x})^2 / (n - 1)$  且  $j \neq i$ ;  $w_{ij} Z_j$  为周围位置属性值偏差的加权平均。与局部 Moran 统计  $I_i$  不同的是, 局部 Geary 统计  $C_i$  是对属性值偏差之差的平方和的加权度量。

### 1.3 空间关联的识别

MC 和 GR 是负相关关系, 根据一个指标可以表达另一个指标<sup>[3,4,7,20]</sup>。从统计上讲, MC 也是很有效的, 因此, 本文使用 MC 来度量全局空

间自相关。在一个正态性分布的假定下, MC 的期望值  $E[I]$ 、方差  $\text{var}(I)$  的数字表达式参见文献 [4, 7], 并可以得到:

$$Z(I) = (I - E[I]) / \sqrt{\text{var}(MC)} \quad (4)$$

根据式(4)计算出检验统计量, 可以对零假设  $H_0$  ( $n$  个区域单元的属性值之间不存在空间自相关) 进行显著性检验, 即检验所有区域单元的属性值之间是否存在空间自相关。  $MC = -1 / (n - 1)$  ( $n$  是总的区域单元数) 或  $GR = 1$  时, 表示一种随机的地理分布模式; 当  $MC > -1 / (n - 1)$  或  $0 < GR < 1$ , 且  $MC$  或  $GR$  显著时, 表示在地理分布中相似的属性值倾向于聚集在一起 (正的空间自相关); 当  $MC < -1 / (n - 1)$  或  $GR > 1$ , 且  $MC$  或  $GR$  显著时, 表示不同的属性值倾向于聚集在一起 (负的空间自相关)。当  $n$  是一个比较大的数值时, MC 的期望值收敛于 0; 而且一个正值和正的空间自相关联系在一起, 负值和负的空间自相关联系在一起。

标准化的  $G$  统计  $Z(G_i)$  可以应用于属性值为正或为负的情形。对零假设  $H_0: G_i = 0$ , 可以执行一个  $t$  检验<sup>[9,10,27]</sup>。  $Z(G_i)$  将位置  $i$  排除在外。  $G$  统计可以用来确定空间凝聚的模式, 判定是高值簇还是低值簇。然而,  $G$  统计不能发现相似性 (正关联) 或非相似性 (负关联) 的空间模式。

与  $G_i(d)$  统计相比, 局部 Moran、局部 Geary 统计具有一些优势。对于一个随机分布假设,  $I_i$  的期望值  $E[I_i]$ 、方差  $\text{var}(I_i)$  的数字表达式参见文献 [10]。

同时, 可得:

$$Z(I_i) = (I_i - E[I_i]) / \sqrt{\text{var}[I_i]} \quad (5)$$

根据式(5)计算出检验统计量, 可以对有意义的局部空间关联进行显著性检验。根据一个“条件”随机方法或排列方法, 可以获得  $I_i$  的一个伪显著性水平。  $p$  值同样为零假设  $H_0$  检验提供了基础, 即检验所有的属性值在空间上是否随机分布。

对局部 Moran 的解释与  $G$  统计相似。一个小的  $p$  值 (如  $p < 0.05$ ) 表明与位置  $i$  相关联的周围位置的属性值较大, 而一个大的  $p$  值 (如  $p > 0.95$ ) 表明与位置  $i$  相关联的周围位置的属性值较小。

局部 Geary 的伪显著性水平  $p$  值的计算与局部 Moran 的  $p$  值的计算相似<sup>[19,27]</sup>。一个大的  $p$  值 (如  $p > 0.95$ ) 表明有一个小的极值  $C_i$ , 这间接表明观测点  $i$  与它周围的观测点之间存在一

个正的空间关联(++或--), 而一个小的  $p$  值(如  $p < 0.05$ )表明有一个大的极值  $C_i$ , 这间接表明观测点  $i$  与它周围的观测点之间存在一个负的空间关联(+--或-+-)。

## 2 空间统计分析与 GIS 的集成

GIS 最关键的特性是能够将各种与空间有关的数据与地理位置链接在一起, 从而可以从空间角度出发表达、分析各种数据。近年来, 不少学者致力于空间统计学与 GIS 的结合在社会经济发展领域的应用研究<sup>[17~21]</sup>。这种结合集中体现于空间统计分析在区域经济分析决策中的应用。

对于究竟如何将空间统计分析与 GIS 结合在一起, 不同的研究者提出了不同的观点<sup>[17~21, 23~26]</sup>。Goodchild (1992)、Ding & Fortheingham (1992)、Zhang & Griffith (1997, 2000)等认为空间统计分析与 GIS 的集成应包括两个完全不同的但同样有效的解决办法: ①将空间统计分析功能嵌入到一个 GIS 中; ②将选定的 GIS 功能嵌入到一个空间统计分析环境中。到目前为止, 注意力主要放在前者<sup>[20, 21]</sup>, 第二种集成方法基本上被忽略了。

结合区域经济分析的实际是, 将前述的各种空间统计分析方法嵌入到一个 GIS 中就可以满足分析存在于各区域单元之间的空间关联的要求, 同时可使用户在同样的环境中执行空间统计分析和可视化分析。因此, 在 ArcView 中使用 Avenue 开发了一个用户友好的统计模块, 使空间统计技术和一个区域描绘过程结合在一个 GIS 环境中, 为区域经济决策提供了一种灵活方便的、交互式的可视化支持工具。下面举例说明其在区域经济分析中的应用。

## 3 实例分析

笔者根据 1978~1999 年新疆维吾尔自治区各县市 GDP 数据, 计算出各县市年平均 GDP 增长速度, 并利用这个指标来进行分析。在此, 仅计算全局 Moran 系数、各县市的局部 MC 系数, 并借助局部 Moran 系数散点图来确定空间显著特征点。根据计算出的全局 MC、各县市的局部 MC 可大致说明上述空间统计分析和 GIS 技术在确定、量化经济区域内存在的空间关联的有效性。

空间统计分析的第一步是要生成一个空间权重矩阵。在笔者设计的模块中, 采用邻近多边形

列表来表示区域单元空间邻近关系, 具体参见文献[20]与图 1。图 2 是实际分析中生成的邻近多边形列表的一部分。

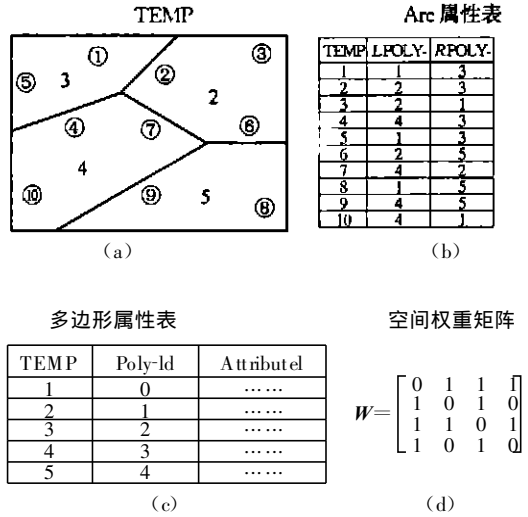


图 1 利用 Arc/Info 拓扑信息建立空间权重矩阵

Fig. 1 Constructing Spatial Weights Matrix Using Arc/Info Topology



图 2 区域单元之间空间关系的邻近多边形列表表示

Fig. 2 Spatial Neighbor List of the Spatial Relations Among Regional Units

生成邻近多边形列表后, 可以计算出隐含在数据集中的空间自相关的性质和强度。根据图 3 中的计算结果, 同时进行显著性检验(取  $\alpha = 0.05$ , 查表得  $Z_{0.05} = 1.96$ , 由于  $Z > 1.96$ , 故拒绝



图 3 全局空间自相关系数及显著性检验统计量

Fig. 3 Global Spatial Autocorrelation Coefficient and Its Significant Testing Statistics

$H_0$ ), 可以表明 1978 ~ 1999 年新疆各县市 GDP 年平均增速之间存在显著的的正的空间自相关。

根据前述的方法, 计算出县级水平的局部 Moran 系数  $I_i$ , 可以进一步考察可能存在于新疆各县市之间的局部空间经济关联模式。借助笔者开发的模块及前述的空间关联识别标准, 可以很好地完成以上计算和分析, 如表 1 (仅列出了显著性的统计值) 及图 4 所示。根据表 1 中的局部 Moran 系数  $I_i$  及检验统计量  $Z$  值, 可以发现各县市局部区域之间既存在显著的的正的空间关联, 又存在显著的负的空间关联。如库尔勒市的年均 GDP 增速与其周围各县市的 GDP 增速之间存在显著的正相关, 鄯善县与其周围各县市之间存在显著的负相关。进一步考察各县市初始数据<sup>[29]</sup>, 可以从另一方面说明存在的局部空间关

左下角的点, 来确定可能存在局部空间关联的类型<sup>[20]</sup>。

## 4 应用前景

对于大多数区域研究而言, 相对独立的经济区构成了分析的重要基础。笔者将计划采用文献 [28] 中的分区方案, 对 10 个子区域分别展开分析, 并将空间关联类型与区域经济概念结合在一起, 对结果进行进一步分析, 来重点认识核心区与邻近区域之间可能存在的以下几种空间经济关联<sup>[19]</sup>: ①扩散效应(++)——周围地区的增长与经济核心区的快速增长相关联; ②离心效应(-)——周围地区的增长与经济核心区的缓慢增长相关联; ③极化效应(+-)——经济核心区的增长与周围地区的缓慢增长或下降相关联; ④无关性——周围地区的增长与经济核心区的经济活动变化没有紧密的关联。

对于考察区从小等级的区域单元到更大区域单元的变化, GIS 很适合, 而且有利于理解一个更大的区域对其组成部分的影响。基于 GIS 的技术, 使得这些数据分析问题变得更加便于管理, 同时为评价空间因素在经济和环境问题上所起的作用提供了更高级的方法。

## 参 考 文 献

- Stein A, Meer V D, Gorte B. Spatial Statistics for Remote Sensing. Dordrecht; Kluwer Academic Publishers 1999
- 马洪超, 李德仁. 基于空间统计学的空间数据窗口大小的确定. 武汉大学学报·信息科学版, 2001, 26(1): 18 ~ 23
- Cliff A D, Ord J K. Spatial Autocorrelation. London; Pion., 1973
- Cliff A D, Ord J K. Spatial Processes; Models and Applications. London; Pion., 1981
- Griffith D A. Theory of Spatial Statistics. In: Spatial Statistics and Models. Boston; D. Reidel Publishing Company, 1984. 1 ~ 15
- Miron J. Spatial Autocorrelation in Regression Analysis; a Beginner's Guide. In: Spatial Statistics and Models. Boston; D. Reidel Publishing Company, 1984. 201 ~ 222
- Goodchild M F. Spatial Autocorrelation. Norwich; Geo-Books 1986
- Getis A, Ord J K. The Analysis of Spatial Association by the Use of Distance Statistics. Geographical Analysis 1992, 24(3): 189 ~ 206
- Ord J K, Getis A. Local Autocorrelation Statistics; Distributional Issues and an Application. Geographical Analy-

表 1 各县市局部 Moran 系数及检验统计量 (部分)

Tab. 1 Local Moran Coefficient and Its Testing Statistics for Different Counties(One Section)

县、市	代码	$I_i$	$Z$ -值
和静县	41	17.161 14	8.032 15
鄯善县	50	-7.032 72	-3.253 38
库尔勒市	56	14.837 59	6.948 30
尉犁县	66	8.176 91	3.841 34
和田市	88	6.712 25	3.503 60
和硕县	81	13.879 01	6.501 16
博湖县	82	15.735 55	7.367 16
墨玉县	83	10.966 43	5.142 55
洛浦县	84	7.612 84	3.578 23
皮山县	86	6.781 95	3.539 73

注: 基于 Bonferroni 标准, 当总体显著性水平  $\alpha = 0.05$ , 样本个体的显著性水平为  $\alpha/n$ , 但必须注意当  $n$  相当大时, 采用此标准可能过于保守。

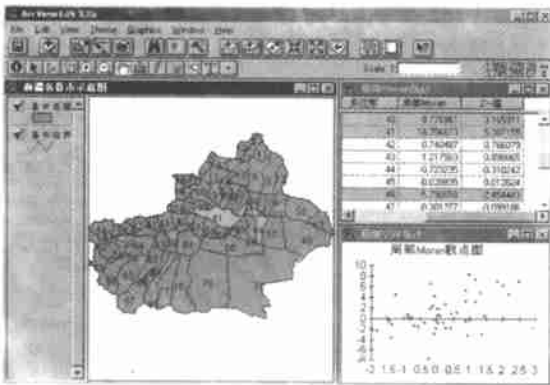


图 4 空间统计分析结果的多窗口链接示意图

Fig. 4 A Sketch Map of Multi-window Linkages for the Results of Spatial Statistical Analysis

联。根据图 4 所示的多窗口链接功能, 也可以通过考察局部 Moran 散点图中的右上角、右下角及

- sis, 1995, 27(4): 286~306
- 10 Anselin L. Local Indicators of Spatial Association—LISA. *Geographical Analysis* 1995, 27(2): 93~115
  - 11 Moore M. *Spatial Statistics: Methodological Aspects and Applications*. New York: Springer, 2001
  - 12 Drummond W J. GIS as a Visualization Tool for Economic Development. *Comput., Environ. and Urban Systems*, 1993, 17(5): 469~479
  - 13 Goodchild M F. A Spatial Analytical Perspective on Geographical Information Systems. *Int. J. Geographical Information Systems*, 1987, 1(4): 327~334
  - 14 Morrill R. Geographic Variation in Change in Income Inequality Among US States 1970~1990. *The Annals of Regional Science* 2000, 34(2): 109~130
  - 15 Shaw S L. GIS for Urban Travel Demand Analysis: Requirements and Alternatives. *Comput., Environ. and Urban Systems* 1993, 17(1): 15~29
  - 16 Clarke G. Applied Spatial Modeling For Business and Service Planning. *Comput., Environ. and Urban Systems*, 1997, 21(3): 373~376
  - 17 Ding Y M, Fotheringham A S. The Integration of Spatial Analysis and GIS. *Comput., Environ. and Urban Systems* 1992, 16(1): 3~19
  - 18 Shen Q, Berkeley C. An Application of GIS to the Measurement of Spatial Autocorrelation. *Comput., Environ. and Urban Systems* 1994, 18(3): 167~191
  - 19 Bao S, Henry M S, Barkley D. RAS: A Regional Analysis System Integrated With Arc/Info. *Comput., Environ. and Urban Systems* 1995, 19(1): 37~56
  - 20 Zhang Z Q, Griffith D A. Integrating GIS Components and Spatial Statistical Analysis in DBMSs. *Int. J. Geographical Information Science*, 2000, 14(6): 543~566
  - 21 Zhang Z Q, Griffith D A. Developing User-friendly Spatial Statistical Analysis Modules for GIS: an Example Using ArcView. *Comput., Environ. and Urban Systems* 1997, 21(1): 5~29
  - 22 Fischer M, Scholten H, Unwin D. *Spatial Analytic Perspectives in GIS*. London: Taylor & Francis 1996
  - 23 Anselin L, Getis A. Spatial Statistical Analysis and Geographic Information Systems. In: Fischer M M, Nijkamp P, eds. *Geographic Information Systems Spatial Modelling and Policy Evaluation*. Berlin: Springer-Verlag, 1993. 103~114
  - 24 Aeselin L, Dodson R, Hudak S. Linking GIS and Spatial Data Analysis in Practice. *Geographical Information Systems* 1993, 1(1): 3~23
  - 25 Anselin L, Getis A. Spatial Statistical Analysis and Geographic Information Systems. *Annals of Regional Science* 1992, 26(1): 19~33
  - 26 Goodchild M, Haining R. Integrating GIS and Spatial Data Analysis: Problems and Possibilities. *Int. J. Geographical Information Systems* 1992, 6(5): 407~423
  - 27 Getis A, Ord J K. Local Spatial Statistics: An Overview. In: Long P eds. *Spatial Analysis, Modelling in a GIS Environment*. Cambridge: GeoInformation International, 1996, 261~177
  - 28 韩德林, 冯亚斌, 高志刚, 等. 新疆开放与外向型经济发展——参与中亚次区域经济集团化新思维. 乌鲁木齐: 新疆科技卫生出版社, 2000
  - 29 高志刚. 新疆区域经济差异及其预警研究. [博士论文]. 北京: 中国科学院地理科学与资源研究所, 2001

作者简介: 陈斐, 讲师, 博士生. 现主要从事区域经济、地理信息系统应用研究. 代表成果: 中国陆疆开放系统与重点区产业建设研究(西北区段)、西部地区开发决策信息支持系统、新疆参与周边国家次区域经济集团化研究等.

E-mail: chenfei1208@sina.com

## Application of the Integration of Spatial Statistical Analysis with GIS to the Analysis of Regional Economy

CHEN Fei<sup>1</sup> DU Daosheng<sup>1</sup>

(1 National Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing,  
Wuhan University, 129 Luoyu Road, Wuhan, China 430079)

**Abstract:** Spatial autocorrelation means the self-correlation or spatial dependence among observations of a geo-referenced attribute. There are two different scales for spatial dependence: global indicators and local indicators. In this paper, the authors summarize a few spatial statistical analysis methods concerning about how to measure spatial autocorrelation and spatial association firstly, then discuss the criteria for the identification of spatial association by the use of global Moran Coefficient, Local Moran and Local Geary.

Secondly, the authors make a brief review about the integration of the spatial statistical analysis with GIS which is believed to occur in two different ways: embedding spatial statistical analysis into a GIS environment and embedding selected GIS functions into a spatial statistical analysis environment. Based on what has been done in this area, the authors point out that it is necessary and worthwhile to develop a user-friendly statistical module combining spatial statistical analysis methods with GIS visual techniques in GIS directly, and provide an example to illustrate how this can be implemented in Arcview using Avenue.

Constructing an adjacency spatial weight matrix is the first step to deal with further analysis. A two-dimensional matrix can be expressed as a one-dimensional array by using the "List" class. In this paper, we use a spatial neighbor list table to represent spatially adjacent relations among different regional units. With the use of Avenue, users can view, choose, input and report important information and results, and report error messages. All of these window-based operations make it possible for users to execute abstract statistical analysis simply by pointing and clicking a friendly GUI (graphical user interface). We take Xinjiang Uyger Autonomous Region as a research area and utilize mean annual GDP increasing velocity from 1978 to 1999 in different counties, then calculate global Moran and local Moran based on those data, and illustrate the usefulness of that module in identifying the characteristic and significance of spatial association among observed locations over space.

**Key words:** spatial statistical analysis; GIS; spatial autocorrelation; spatial association; regional economic analysis

**About the author:** CHEN Fei, lecturer, Ph. D candidate. His major research orientation include regional economy development, spatial statistics, GIS and its application to regional economic analysis. His typical achievements are Opening System of Inland Border Area in China and Industrial Construction for Key Section, Decision-making and Information Support System for the Exploitation in Western China, Participating in Sub-regional Economic Collectivization for Xinjiang Along Neighbor Countries etc.

E-mail: chenfei1208@sina.com

## 《武汉大学学报·信息科学版》编辑委员会

名誉主任: 宁津生

主任: 李德仁

委员: 毋河海 王新洲 刘 甬 刘经南 刘耀林 朱元泓 朱灼文

仲思东 张正禄 张祖勋 苏光奎 杜清运 杜道生 李建成

李清泉 郑肇葆 柳建乔 晁定波 龚健雅 舒 宁 詹庆明

主 编: 李德仁(兼)

副 主 编: 柳建乔(常务)