

# GIS 数据之属性不确定性的研究

史文中<sup>1)</sup> 王树良<sup>1),2)</sup>

<sup>1)</sup>(香港理工大学土地测量与地理资讯学系, 香港)

<sup>2)</sup>(武汉大学测绘遥感信息工程国家重点实验室, 武汉 430079)

**摘要** GIS 中的不确定性可分为位置、属性、时域、逻辑关系和完整性等几个方面。属性数据的不确定将直接影响基于 GIS 决策的准确性和可靠性, 特别是对侧重于属性分析的领域。本文首先总结了属性不确定性的含义、来源、传播、它和位置不确定性的关系, 以及属性不确定性分析和可视化的理论和方法。最后概述了属性不确定性的研究进展。

**关键词** 属性不确定性 GIS 理论和方法

中图法分类号: P208 文献标识码: A 文章编号: 1006-8961(2001)09-0918-07

## State of the Art of Research on the Attribute Uncertainty in GIS Data

SHI Wen-zhong<sup>1)</sup>, WANG Shu-liang<sup>1),2)</sup>

<sup>1)</sup>(Department of Land Surveying and Geo-informatics, The Hong Kong Polytechnic University, Kowloon, Hong Kong)

<sup>2)</sup>(National Laboratory of Information Engineering, Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079)

**Abstract** Uncertainty of spatial data in GIS can be in the aspect of position, attribute, temporary, logical relationship and completeness. Among them, attribute uncertainty can directly affect the quality of GIS-based decision making. In this paper, the concept and propagation of the attribute uncertainty are firstly described. Then the relationship between the attribute and position uncertainties is analyzed. Moreover, the theories and methods to study and visualize the attribute uncertainty are discussed. Attribute uncertainty of spatial describes the difference between the observed values of an entity attribute with their true values in spatial and temporal space. There are a number of theories and developed methods to deal with attribute uncertainties. These may include, for example, (1) object- and filed-based model; (2) probability theory, evidence theory, spatial statistics, "S" band model and probability vector which deal with random attribute uncertainties, while fuzzy set, rough set, genetic algorithms which deal with imprecise attribute uncertainties. Cloud theory can be used to deal with both randomness and fuzziness of attribute uncertainty. The relationship between  $x$  and  $\mu(x)$ , fuzzy set is one to one, rough set is one to a region, and cloud theory is on one to a piece of cloud; (4) Visualization of attribute uncertainty which is based on Bertin geographical parameters and their extensions. Three-dimensional visualization and virtual reality are also helpful. Finally, the future research issues on attribute uncertainty are given.

**Keywords** Attribute uncertainty, GIS, Theories and methods

## 0 引言

GIS 中的模型是对不确定现实世界的一种近似表达。作为一个空间信息决策支持系统, GIS 的应用日益广泛, 已经从测绘、地理等地学学科扩展到环

境、交通、规划、农林等学科领域, 并与遥感等技术集成<sup>[1,2]</sup>。可是在复杂多变的现实世界中, 空间实体多相互混杂, 界限有时不很分明。人们在利用 GIS 认识和改造自然的过程中, 常常不易明确地将空间实体分类, 所获信息也不可能穷尽现实世界的全部。

而且,获取的信息在被导入计算机系统并用于空间分析决策的过程中,又被部分舍弃或删除(如目标模型的地学点、线、面抽象,制图综合)。GIS数据的误差本质上可分为3种:系统误差、随机误差、人为粗差和三者累积误差的影响。在GIS应用过程中,空间数据质量中的不确定性问题也越来越为人重视。

经典的数据处理方法认为空间分布可以用一组离散点、线、面来表达。通常假设已经检核了属性数据,并把属性不确定性和位置不确定性隔离讨论。这造成GIS不确定性研究的重点,主要基于数值型的统计方法分析位置的不确定性,而对属性数据不确定性的研究相对较少<sup>[3]</sup>,同时,传统误差一般指测量值或计算值与真实数据或假定真实数据间的差值,而获取大量空间数据的真值并不容易。所以,有必要研究属性不确定性及其和位置不确定性的一致框架,如S带模型<sup>[4]</sup>,延拓误差为更广义的不确定性概念<sup>[5]</sup>。

实际上,在以属性数据分析为主的土地评估学、土壤化学、环境科学和农林科学等许多GIS应用领域中,属性的不确定性带来的影响较之位置不确定性更大。如在评估土地的价格属性时,整个过程基本上都是在做属性不确定性分析,即使地价的定位也要利用地价属性不确定性在空间位置的表现形式——均质地域<sup>[6]</sup>。由于正确使用不确定性在GIS决策支持中具有避免利用错误信息导致决策失误和度量信息支持度的作用<sup>[2]</sup>,故在许多侧重于属性分析的领域中,属性数据的不确定问题将直接影响GIS分析决策结果的准确性和可靠性。现在,属性不确定性的研究正在兴起,并取得了相当的成果。

## 1 属性不确定性

美国国家数字制图数据标准委员会的空间数据标准包括历程(linage)、位置精度、属性精度、逻辑一致性和完整性<sup>[2,7]</sup>。后来又提出了表达时间的时域精度。因历程是描述数据的获取和处理等过程,故GIS的数据精度应该包括空间精度、时域精度、属性精度、拓扑一致性和完整性5个方面。可见,属性精度是精度范畴中的一个重要组成部分。GIS的属性数据是对点、线、多边形、面、体等矢量数据或栅格数据属性值(或属性类别)的描述。

误差指观测值与其真值间的差异<sup>[8,9]</sup>,具有统计意义,而不确定性则为更广义的误差概念延拓,是被测量对象知识缺乏的程度,表现为空间数据所具有

的误差、不精确性、随机性和模糊性,且受尺度、分辨率、抽样等因素的影响<sup>[10,11]</sup>。混沌理论认为空间实体因各要素的相互作用而处于力的不平衡状态,综合导致了其不确定性,并表现为空间上的要素不均匀性和不稳定性,时间上的发展非周期性<sup>[12]</sup>。含有不确定性的空间实体可能包含发展阶段所存储保留的信息,也随时接收来自其他实体的信息和自身发射信息。

属性不确定性是在采集、描述和分析真实世界中客观实体的过程中,实体属性的量测、分析值围绕其属性真值,在时间和空间内的随机不确定性变化域。也可以认为属性不确定性是更广义的属性误差问题,或者是属性误差问题的纵横延伸。连续属性数据的不确定性,可以使用和位置不确定性相同的误差传播定律等方法用量测误差度量;非连续属性数据的可以通过评价一组分类结果来实现<sup>[2]</sup>。研究属性不确定性可概括为探讨不确定性的产生、传播和控制等几个方面的问题。

### 1.1 属性不确定性的来源

GIS中的误差与所使用的数据、算法和数据结构皆有关。主要有明显误差、源于自然或原始量测值的误差和源于数据处理的误差三类<sup>[5]</sup>。因此,GIS在使用属性数据描述空间实体的属性时,由于计算机在采集源数据、分析数据和传送表达结果的过程中,不可避免地要给出空间实体的属性定义,并对复杂的空间实体进行取舍、抽象和离散。同时其准确性也会受到传感器和数学模型等设备技术因素的影响,所以就产生了GIS属性的不确定性及其传播问题。由此可见,属性不确定性主要来自属性的定义、数据源、数据建模和分析过程中引入的不确定性等。其中,数据源的不确定性又源于数据采集过程中的量测、人为判断和假设等<sup>[10]</sup>。利用遥感技术获取的土地覆盖分类数据的属性不确定性,则主要来自空间、光谱、波谱和时域特性的不确定性<sup>[2]</sup>。

### 1.2 属性不确定性的传播

应用GIS时,常利用空间数据库中的基础属性数据派生部分新属性数据。例如根据某区域的土地利用类型、地价、交通便利度和环境污染度属性数据产生一幅关于在该区域投资房地产的适宜性地图。这种GIS通过对多个空间属性层进行空间操作得到新属性域的方法,常表现为逻辑模型和算术模型。传统的布尔逻辑模型,只能处理确定性的属性数据分析,文献<sup>[10]</sup>中把它扩展为基于模糊逻辑的不确

定性传播模型. 他们把任意位置的属性值满足属性值条件的程度作为隶属度, 对  $n$  层单属性专题图进行逻辑操作, 得到具有  $n$  种属性的新专题层. GIS 属性数据分析中的算术模型, 是派生新属性数据的另一类模型, 如地价评估计算模型.

### 1.3 位置与属性不确定性的关系

地理信息的基本单元是一个连接时空位置  $x$  和系列属性  $z$  的数组  $(x, z)$ <sup>[1]</sup>, GIS 主要使用位置和属性两种数据描述、分析和处理现实世界的空间实体. 在 GIS 应用中, 位置和属性的关系非常紧密. 在目标模型用点、线和面状目标的位置、属性或拓扑关系特性描述 GIS 空间实体时, 空间目标可以是界址点、道路和宗地的空间定位坐标, 其相应的属性(如界址点使用的时间, 道路的宽度, 以及宗地的权属、拓扑关系、与相邻宗地的关系等)则依附在位置信息上, 域模型则把二者联系在一起. 所以, 位置不确定性和属性不确定性的关系也非常紧密.

边界不确定性是属性与空间不确定性紧密相关的典型表现. 空间坐标定位的不确定导致属性类别的分类不确定; 另一方面, 类别属性的同物异谱或同谱异物则使得相邻类别的边界难以准确定位. 而且位置在用确切边界和均质多边形模型表示物质时, 其误差就是同时含有属性和边界位置不确定性. 遥感影像的混合象素, 则是在地面位置上相邻的不同地物类别属性的综合反映<sup>[11]</sup>.

连续属性数据是连续的随机变量, 它的性质类似于位置数据. 连续属性数据的不确定性可以用量测误差决定, 它的建模可以使用和位置不确定性相同的误差传播定律等方法. 非连续属性数据的评估可以通过评价一组分类结果来实现. 对于一个明确定义的面状目标, 位置不确定性与属性不确定性的共同作用, 意味着其边界并非总是几何意义上的线, 而有时是具有一定宽度的过渡区域. 边界内部(属性)是空间匀质分布的, 且其不确定性可以用概率来描述(例如地籍纪录的平均正确概率). 包含和自相关两个指标已被用来描述位置和属性不确定性的关系的误差分布模式<sup>[13]</sup>, 而“S-带”模型把位置和属性不确定性转化为用同一种指标描述, 在同一理论基础下用统一的方法集成分析了其本质<sup>[4]</sup>.

## 2 属性不确定性的研究内容

现实空间可被分为渐变的连续空间和跳跃的离散空间两种. 属性数据根据连续与否, 也有了离散数

据和连续数据之分. 连续空间常假定区域内部属性是均质的, 其重要特征的变化发生在边界上. 由于区域边界是经过人为判断和边界定位确定的. 因此, 区域多边形中属性数据的不确定性将主要取决于属性分类不确定性、边界定位误差和属性测量误差.

传统的属性不确定性<sup>[14]</sup>主要是针对卫星、航测和雷达等遥感影像的分类, 如土壤、森林和植被等. 同时, 属性区域边界存在的定位误差也导致了边界附近区域属性的分类不确定性. 区域差异的显著之处在中央部位, 从中心向边缘差异的显著程度渐次递减. 这样, 在两个质地不同的区域之间必然存在着一个能够容纳不确定性的、有宽度的边界过渡带, 而不是纯几何的线状目标. 故由边界定位误差引起的属性不确定性, 与边界线的定位精度、空间实体的形状和面积、相邻实体中属性特征值的变化程度都有关. 属性区域内部的属性值通过抽样测量得到, 空间属性观测量的统计不确定性由相关函数表达<sup>[10]</sup>.

属性不确定性涉及的范围很广, 基本类型包括: (1)对现实世界本身的定义不明确. 如土地价格属性的高低. (2)难以确定的实体与实体间的边界, 很多实体没有明确的边界, 有些则与其他实体相互混杂, 如商住综合用地<sup>[6]</sup>; 有的是一个渐变的过程, 如土壤的边界; 有的是一个不确定区域, 如城乡相互渗透的城乡结合部. 在遥感中也有类似的混合用地的分类问题, 一块长有疏草的土地, 难以准确划分草地和裸露土壤的界线; 一个水中长草的湖泊, 更是难以定出草地和水域明确界线. (3)定义或语义模糊的实体关系, 如在南湖的“附近”, 是距离南湖 5km, 还是 50km. (4)空间分析中的不确定性, 如某块土地对于某种作物的适宜程度, GIS 分析的结果可能介于适宜和不适宜之间, 为比较适宜或不太适宜的情况. (5)查询过程中的不确定性, 如土地交易中, 买主要寻找具有在南湖附近, 地价适中, 交通便利等条件的土地. (6)在矢量专题图的 GIS 应用分析中, 属性数据的不确定性问题主要是属性不确定性的度量和在 GIS 分析过程中的传播<sup>[10]</sup>等等.

## 3 研究属性不确定性的理论和方法

在可用于属性不确定性分析的方法中, 目标模型和域模型是经典的数据处理模型. 概率论、概率矢量、证据理论和空间统计学具有统计意义, 粗集、模糊集合、云理论、遗传算法、混沌理论、灰色理论等应用广泛.

### 3.1 目标模型和域模型

目标模型和域模型在GIS中分别对应于矢量和栅格数据。目标模型用点、线和面状目标的位置、属性或拓扑关系特性描述空间实体,较易表达目标的位置不确定性。域模型通过对每个像素赋予属性来描述空间实体,较适于模糊、含混的空间实体<sup>[2]</sup>,可以用于描述异质数据不确定性的空间分布以及渐变区域的不确定性。在表达GIS不确定性的空间分布时,目标模型和域模型是互补的。

### 3.2 概率论

概率论分析由属性随机误差造成的不确定性。它在获得大量的观测值后,把不确定性表示成给定条件下、某一假设为真的条件概率。文献<sup>[2]</sup>中提出了能够描述属性不确定性空间分布的、基于最大似然分类的概率矢量法。一个像素的概率矢量的零元素被剔除并按降序排列,概率矢量中的第一个元素表示其所属类别,其他元素描述分类的不确定性,兼有描述绝对不确定性、相对不确定性、像素混合度和证据完整性4个参数。提出了基于概率论方法和具有概率解释的确定因子模型,研究并实现了结合位置和属性不确定性的“S-带”模型。文献<sup>[5]</sup>中的影象隶属概率矢量和文献<sup>[11]</sup>中的概率场,则没有顺序和剔除零元素。此外,还有随机影象的不确定性研究<sup>[16]</sup>。

### 3.3 证据理论

证据理论是概率论的一个扩展(又称 Dempster-Shafer 理论)<sup>[17]</sup>,是由可信度函数(度量已有证据对假设支持的最低程度)和可能函数(衡量根据已有证据不能否定假设的最高程度)所确定的一个区间。当证据的未支持部分为空时,证据理论等同于传统概率论。证据理论可以利用结合规则或两两比较的方法分析<sup>[18]</sup>属性不确定性。

### 3.4 空间统计学

空间统计学利用有序模型描述无序事件<sup>[8]</sup>,根据不确定性和有限信息分析、评价和预测空间属性数据,其主要运用空间自协方差结构、变异函数或与其相关的自协变量或局部变量值的相似程度来描述空间属性的不确定性<sup>[3]</sup>。空间统计学能改善GIS对随机过程的处理,估计模拟决策分析的不确定性范围;分析空间模型的误差传播规律;并为分析连续域的空间相关性提供理论依据和量化工具等。用地理统计数据、栅格数据、点数据和物体可描述现实世界。文献<sup>[11]</sup>则应用空间统计学的克吕格方法,由已

知点位直接估计未知点位的类别验后概率,从影象上获取分类信息。

### 3.5 模糊集合

模糊集合,是用隶属函数确定的隶属度描述不精确的属性数据,重在处理不精确的概率<sup>[2,19]</sup>。模糊集合在GIS中把类型、空间实体分别视为模糊集合、集合元素,空间实体对备选类制论域连续隶属度区间为 $[0,1]$ 。元素隶属度用于表示实体属于某类型的程度。具有类型混合、居间或渐变不确定性的实体可用元素隶属度描述,如一块含有土壤和植被的土地,可以由两个元素隶属度表示。在模糊集的应用过程中,文献<sup>[3]</sup>中讨论了不确定性数据的模糊布尔逻辑模型;文献<sup>[20]</sup>中提出融模糊综合评判和模糊聚类分析为一体的模糊综合法,基于绝对均值距离的模糊聚类分析,评估土地的地价和级别属性不确定性;文献<sup>[2]</sup>中使用模糊隶属度表达遥感影象中的不确定相邻边界的像素类别。

### 3.6 云理论

云理论由云模型、不确定性推理和云变换构成<sup>[21]</sup>。云在空间由系列云滴组成,具有期望值、熵和超熵3个数字特征。期望值完全隶属于该定性概念;熵是定性概念模糊度的度量,其值越大,概念越模糊;超熵反映云滴的离散程度,其值越大,隶属的随机离散度越大。云理论构成定性和定量相互间的映射,处理GIS中容模糊性(定性概念的亦此亦彼性)和随机性(隶属度的随机性)为一体的属性不确定性。云理论已经用于空间关联规则的挖掘、空间数据库的不确定性查询<sup>[22]</sup>。

### 3.7 粗集理论

粗集理论由上近似和下近似组成,适于处理不精确、不确定和不完全的数据<sup>[23,24]</sup>。粗集从集合论的观点出发,在给定论域中以知识足够与否作为实体分类的标准,并给出划分类型的精度。上近似中的实体具有足够必要的信息和知识,确定为属于该类别;论域全集以内且下近似以外的实体没有必要的信息和知识,确定为不属于该类别;上近似和下近似的差集中的实体没有足够必要的信息和知识,无法确切地判断是否属于该类别,为类别的边界。若两个实体有完全相同的信息,则它们为等价关系,不可区分。根据利用统计信息与否,现存的粗集模型及其延伸可以分为代数型和概率型两大类<sup>[25]</sup>。粗集可用于分析GIS属性数据库属性表的一致性和属性的可靠性、简化属性依赖和属性表、生成最小决策和分类

算法,已被用于描述属性 ROSE 不确定模型,分辨不精确的空间影象<sup>[2]</sup>,发现不确定属性的知识<sup>[22]</sup>,集成多源不确定的属性数据<sup>[26]</sup>,分类遥感影象<sup>[27]</sup>,面向目标的软件评估<sup>[28]</sup>.

### 3.8 灵敏度分析

GIS 的属性不确定性的灵敏度分析,在理论上属于不确定性传播分析.它通过在地理分析输入中添加模拟理论干扰变量,研究所加输入对输出成果的作用<sup>[29]</sup>.主要用于讨论属性不确定性对 GIS 成果的影响规律,分析不能用数学模型表达的属性不确定性,检查和划定 GIS 分类产品的等级.属性不确定性有 3 个基本参数:不足率(真实存在而属性没有)、过量率(真实没有而属性出现)和紊乱率(属性拥有和真实不符的错误值)<sup>[7]</sup>.对于 GIS 属性不确定性的灵敏度分析,文献<sup>[29]</sup>中作了较为系统的回顾,可以通过分析 5 个灵敏度指标、添加理论误差<sup>[9]</sup>和噪声概率模型<sup>[7]</sup>实现.灵敏度分析在栅格数据和地图叠置中发现符号语义对结果有各种等级的影响<sup>[30,31]</sup>、属性不确定性在矢量 GIS 中可以传播<sup>[29]</sup>.

此外,研究 GIS 属性不确定性的还有基于复制、交叉和变异的遗传算法<sup>[32]</sup>;基于灰色分析的灰色系统<sup>[33]</sup>;基于信息无序互动的混沌理论<sup>[34]</sup>;基于未确知数的未确知数学<sup>[35]</sup>;基于可靠性的类别变换区域模型<sup>[6]</sup>;基于加权平均的降雨影响土壤分布随机场<sup>[16]</sup>;利用 Latin Hypercube Sampling Strategy 传播土壤重金属富集度不确定性的随机模型<sup>[5]</sup>;度量和传播分析属性数据不确定性的 GIS 逻辑操作和算术操作<sup>[10]</sup>,等等.

因此,目标模型是基于目标的空间数据模型,重在用传统的测量平差理论讨论位置不确定性,采用域讨论属性的不确定性问题优于使用目标模型. GIS 的灵敏度分析是理论上的属性不确定性传播分析.基于经典的确定集理论,概率论和空间统计学研究随机的属性不确定性.由于误差矩阵不能描述不确定性的空间分布,概率矢量被提出表示不确定性的空间分布.证据理论是概率论的一个扩展.经典集合也被扩展,用于研究不能精确描述的属性不确定性.模糊集合重在模糊属性,基础为模糊隶属度;云理论重在兼有模糊性和随机性的属性,基础为云变换;粗集重在不完备的属性信息,基础为上、下近似集.在属性自变量集  $x$  和类别因变量集  $\mu(x)$  之间,模糊数学是一对一关系,即对一个特定的  $x_k$ ,只存在唯一的隶属度  $\mu(x_k)$ ,且  $\mu(x_k) \in [0, 1]$ ;粗集是

一对区域关系(Pawlak 1981),即对一个特定的  $x_k$ ,存在隶属区域 $[\{\min_{-}\mu(x_k)\}, \{\max_{-}\mu(x_k)\}]$ ;云理论则是一对一云滴关系,云滴根据隶属度在空间随机离散分布,聚集到一定程度成为一朵云.当然,上述理论和方法不是孤立的,属性不确定性分析常常要综合应用它们,还要充分借鉴成功的位置不确定性分析方法、常规的属性数据库技术和数据挖掘技术等.

## 4 属性不确定性的可视化

GIS 属性不确定性的可视化是空间属性数据的视觉表达与分析.它借助计算机技术,把抽象的属性数据转换为人的视觉可以直接感受的具体图形图象.属性不确定性的可视化技术包括 Bertin 变量分类系统及其扩展、动画和三维虚拟现实等.

Bertin 变量分类系统<sup>[36]</sup>包括 7 个基本地理变量(大小、数值、纹理、颜色、方向、形状和位置),组合起来可以可视化平面符号的变化、不同的数据水平和变量联合效用.如也门和沙特阿拉伯的国界应用不同宽度的线分别表示确定的等值线和海图上未标明的测深;苏格兰汇水边界的超载,红色是警戒,雾翳或模糊表示不确定的边界<sup>[37]</sup>.类别边界的不确定性以混合形状或混合颜色表示,内部的不确定性则用影象灰度的褪变表达<sup>[30]</sup>.

Maceachren 在 Bertin 系统中又加入了符号焦点变量<sup>[38]</sup>.以表达轮廓是否清楚、填充是否明晰、雾翳和判断的能力.轮廓涉及地图符号边界的调整,以及实体符号的外围界线.确定的实体为“明显、狭窄的线”,而不确定的实体为“模糊、宽大的线”.对于拥有分离边界和内部地理属性的实体,实体填充能通过调整来反映关联不确定性.填充模式定义越清晰,实体越明确.它固有的作用即是在出现薄雾时,利用雾翳寻找模糊的不确定的地图实体,通过雾翳产生的可视化模糊度来描绘不确定性.简化决策通过调整判断力来表明与地图特征连接的不确定性.

Mcgranaghan 把现实和时间也视为表达数据质量的符号化地理变量<sup>[31]</sup>.在地图上,数据质量越低,实体越模糊;数据质量越高,实体越清晰<sup>[3]</sup>.图形动画则应用了时间变量,把图形化的属性不确定性按时间序列实时处理,利用闪烁、渐变和移动表征数据质量的不同方面.闪烁通过实体出现和消失的时间量,表示和实体相关的数据质量,质量高的实体出现的时间长于消失的时间.渐变借助实体的一个地

理特征的摆动,反映相关的数据质量.如一条河流摆动在绿色和蓝色之间(可靠性较高),或者绿色和红色之间(可靠性较低).动画移动则是通过将图形实体移向多个位置,来表达位置的不确定性.

三维虚拟现实技术把GIS属性不确定性的抽象信息具体化<sup>[39,40]</sup>,集成不确定信息和单层图形数据,利用动摇、摆动、不确定轮廓和厚实表面来可视化不确定性.Clarke和Teague(2000)利用虚拟现实模型语言,从已有的方法中选取亮度、闪烁、摆动和摇晃,构造了表示图形不确定性的描绘技术<sup>[41]</sup>.实体的亮度值根据实体的不确定性水平细化设置.实体越确定越亮,越不确定越暗.动画闪烁利用实体的出现和消失时间,把不确定性传送与地图用户,实体的不确定性越高,闪烁等级越大,越不稳定;反之越小越稳定.颜色摇摆使用色度之间的运动来描绘实体的位置和属性的不确定性.确定的实体不摆动,而不确定的实体在不同颜色之间摆动.大小的摆动是通过变换道路线段的宽度来描绘位置的不确定性.最后,他们根据用户眼睛在虚拟场中的焦点,设计了借助大量用户对此反映的统计结果来测试有效性的方法.

## 5 总 结

本文讨论了属性不确定性的含义、来源、传播,以及它和位置不确定性的关系.并且总结了概率矢量等多种属性不确定性分析的理论和方法.认为:

(1) GIS的属性数据是对点、线、多边形等的属性值或类别的描述.属性不确定性是在采集、描述和分析现实世界中的客观实体的过程中,实体属性的量测或分析值围绕其属性真值,在时间或空间内的随机不确定性变化域.属性不确定性主要来自属性的定义、数据源、数据建模和分析过程中引入的不确定性等.利用GIS对空间数据分析的过程中,属性不确定性也随之而传播.在许多情况下,位置和属性不确定性的关系非常紧密.

(2) 属性不确定性涉及的问题很多,包括难以确定的实体与实体间的边界,定义模糊的实体关系,空间分析中的不确定性及其传播,不确定性查询、度量不确定性等.

(3) 传统的概率论可用来分析由属性随机误差而产生的不确定性.基于最大似然分类的概率矢量,可用于指出分类之不确定性及其空间分布不确定性.证据理论是基于可信度函数和可能函数的一个区间,是概率论的一个扩展.空间统计学根据不确定

性和有限信息分析、评价和预测空间属性数据,主要运用空间自协方差结构、变异函数或与其相关的自协变量或局部变量值的相似程度来描述空间的不确定性.目标模型和域模型是互补的.

(4) 概率论和空间统计学研究随机的属性不确定性,不能精确描述的属性不确定性可以考虑模糊集、粗集和云理论.在属性自变量集 $x$ 和类别因变量集 $\mu(x)$ 之间,模糊数学是一一对应关系,粗集是一一对区域关系,云理论则是一一对云滴关系.

(5) GIS可视化是空间数据的视觉表达与分析,属性不确定性的可视化是它的重要组成部分,已经出现了包括Bertin地理变量及其延伸、动画和三维显示技术在内的很多技术.

(6) 各种研究属性不确定性的理论和方法都不是相互孤立的,应该综合使用.属性不确定性是GIS决策分析的重要基础支持信息,它的研究也应该充分借鉴其他领域发展成熟的理论和方法.随着对属性不确定性研究的不断深入,其内容和方法将会日渐丰富和改善.

## 参 考 文 献

- 1 Goodchild M F. Measurement-based GIS[A]. In: Proceedings of the International Symposium on Spatial Data Quality'99 [C], Hong Kong: The Hong Kong Polytechnic University, 1999: 1~9.
- 2 史文中. 空间误差处理的理论和方法[M]. 北京: 科学出版社, 1998.
- 3 Burrough P A. GIS and geostatistics: Essential partners for spatial analysis [A]. In: Proceedings of the International Symposium on Spatial Data Quality'99 [C], Hong Kong: The Hong Kong Polytechnic University, 1999: 10~20.
- 4 Shi W Z. Modeling positional and thematic uncertainties in integration of remote sensing and geographic information systems [M]. Enschede: ITC Publication, 1994.
- 5 Burrough P A, FRANK A U (eds). Geographic objects with indeterminate boundaries[M]. Basingstoke: Taylor and Francis, 1996.
- 6 王树良, 孙春生, 严春. 基准地价中的土地利用类型探讨[J]. 中国土地科学, 1999, 13(1): 5~9.
- 7 Bonin O. New advances in error simulation in vector geographical databases[A]. In: Accuracy 200: Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences [C]. Amsterdam: University of Amsterdam, 2000, 59~65.
- 8 Mikhail E M, Ackermann F. Observations and Least Squares [M]. New York: IEP-A Dun-Donnelley Publisher, 1976.
- 9 黄幼才等. GIS空间数据误差分析和处理[M]. 武汉: 中国地质大学出版社, 1995.

- 10 刘文宝,邓敏,夏宗国. 矢量 GIS 中属性数据的不确定性分析[J]. 测绘学报,2000,29(1):76~81.
  - 11 张景雄,杜道生. 位置不确定性与属性不确定性的场模型[J]. 测绘学报,1999,28(3):244~249.
  - 12 郭达志,胡召玲,陈云浩. GIS 中空间对象的不确定性研究[J]. 中国矿业大学学报(自然科学版),2000,29(1):20~24.
  - 13 Goodchild M F. Attribute accuracy[A]. In: Elements of Spatial Data Quality[C], New York: Elsevier, 1995:139~151.
  - 14 Congalton R G, Kass G. Assessing the accuracy of remotely sensed data: Principles and practices[M]. Boca Raton: Lewis Publications, 1999.
  - 15 Canters F. Evaluating the uncertainty of area estimates derived from fuzzy land-cover classification [J]. Photogrammetric Engineering and Remote Sensing, 1997, (63):403~414.
  - 16 Journel A G. Modelling uncertainty and spatial dependence: stochastic imaging [J]. International Journal of Geographical Information Systems, 1996, 10(5):517~522.
  - 17 Shafer G. A mathematical theory of evidence. Princeton: Princeton University Press, 1976.
  - 18 Chrisman N C. Exploring geographic information systems[M]. New York: Wiley & Sons, 1997.
  - 19 Zadeh L A. Fuzzy sets[J]. Information and Control, 1965, (8): 338~353.
  - 20 王新洲,王树良. 模糊综合法在土地定级中的应用[J]. 武汉测绘科技大学学报, 1997, 22(1):42~46.
  - 21 李德毅,史雪梅,孟海军. 隶属云和隶属云发生器[J]. 计算机研究和发展, 1995, 42(8):32~41.
  - 22 邱凯昌. 空间数据发掘和知识发现的理论和方法[博士论文][D]. 武汉:武汉测绘科技大学, 1999.
  - 23 Pawlak Z. Rough sets[J]. International Journal of Computer and Information Sciences, 1982, 11(5):341~356.
  - 24 Pawlak Z. Rough sets [M]. Norwell: Kluwer Academic Publishers, 1991.
  - 25 Yao Y Y, Wong S K M, LIN T Y. A review of rough set models [A]. In: Rough Sets and Data Mining Analysis for Imprecise Data[C], Norwell: Kluwer Academic Publishers, 1997:47~75.
  - 26 WANG S L, WANG X Z. Data mining applied in land use control in city-country combinative area[A]. Proceedings of the 19<sup>th</sup> ISPRS Congress[C], Amsterdam: the 19<sup>th</sup> ISPRS, 2000, 1677~1683.
  - 27 Ahlqvist O, Keukelaar J, Oukbir K. Rough classification and accuracy assessment[J]. International Journal of Geographical Information Science, 2000, 14(5):475~496.
  - 28 Gunther R. Rough set-based analysis in goal-oriented software measurement [A]. In: Proceedings of METRICS'96[C], 1996, 10~19.
  - 29 Lodwick W A *et al.* Attribute error and sensitivity analysis of map operations in GIS: suitability analysis. International Journal of Geographical Information Systems, 1990, 4(4):413~428.
  - 30 Fisher P F. Visualizing uncertainty in soil maps by animation [J]. Cartographica, 1994, (30):20~27.
  - 31 Mcgranaghan M. A cartographic view of spatial data quality[J]. Cartographica, 1993, (30):8~19.
  - 32 Buckless B P, Petry F E. Genetic algorithms [M]. Los Alamitos: IEEE Computer Press, 1994.
  - 33 邓聚龙. 灰色系统基本方法[M]. 武汉:华中工学院出版社, 1987.
  - 34 Awrejcewicz J. Bifurcation and chaos in simple dynamical systems[M]. Singapore: World Scientific, 1989.
  - 35 刘开第等. 未确知数学[M]. 武汉:华中理工大学出版社, 1997.
  - 36 Bertin J. Semiologie graphique: Les diagrammes, les reseaux, les cartes, Mouton and Gauthiers-Villars[M], Berlin: Mouton and Gauthiers-Villars, 1967.
  - 37 Slocum T A. Thematic cartography and visualization [M]. London: Prentice Hall, 1999.
  - 38 Maceachren A M. Visualizing uncertain information [J]. Cartographic Perspectives, 1992, (13):10~19.
  - 39 龚建华等. 地学可视化探讨[J]. 遥感学报, 1998, 3(3):236~243.
  - 40 刘大杰,史文中等. 空间误差处理的理论和方法[M]. 上海:上海科学技术文献出版社, 1998.
  - 41 Clarke K C, Teague P D. Representation of cartographic uncertainty using virtual environments [A]. Accuracy 200: Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences [C], Amsterdam: University of Amsterdam, 2000, 109~116.
- 史文中 1963年生,博士,香港理工大学土地测量与地理资讯学系副教授,地球资讯科技研究中心主任.主要从事地理信息系统和遥感的研究.出版学术专著3部,发表学术论文百余篇.
- 王树良 1975年生,博士生,香港理工大学研究助理.主要从事地理信息系统和空间数据挖掘的研究.出版5部学术专著的章节,发表学术论文30余篇.