

空间数据与空间分析 不确定性原理

Principles of Modeling
Uncertainties in Spatial Data
and Spatial Analyses

(第二版)

史文中/著

 科学出版社

空间数据与空间分析 不确定性原理

(第二版)

史文中 著

科学出版社

北京

内 容 简 介

全书共七篇计 18 章,系统地阐述空间数据与空间分析的不确定性理论体系。第一篇为概述,主要介绍不确定性的相关来源与概念,以及不确定性理论的数学理论基础。第二篇阐述空间数据不确定性理论,包括空间数据位置不确定性模型、属性不确定性建模以及位置与属性不确定性综合模型。第三篇介绍空间模型不确定性理论。第四篇讲述空间分析不确定性理论,包括叠置分析位置不确定性建模、缓冲区分析位置不确定性建模和线简化位置不确定性建模。第五、六篇分别给出空间数据质量控制理论和空间数据质量信息表达方法。第七篇总结未来的发展方向。

本书系统地反映了空间数据与空间分析不确定性处理的理论与方法,也叙述了国际上该领域的最新研究进展,可供 GIS、遥感及相关专业研究人员、高校教师、研究生阅读使用。

图书在版编目(CIP)数据

空间数据与空间分析不确定性原理/史文中著.—2版.—北京:科学出版社,2015.3

ISBN 978-7-03-043522-

I. ①空… II. ①史… III. ①空间信息系统-不确定系统-研究 IV. ①P208

中国版本图书馆 CIP 数据核字(2015)第 040348 号

责任编辑:彭胜潮 李 静/责任校对:赵桂芬

责任印制:肖 兴/封面设计:铭轩堂

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

中国科学院印刷厂印刷

科学出版社发行 各地新华书店经销

*

2005 年 6 月第 一 版 开本:787×1092 1/16

2015 年 3 月第 二 版 印张:22 1/4 插页:6

2015 年 3 月第三次印刷 字数:522 000

定价:98.00 元

(如有印装质量问题,我社负责调换)

序*

我十分高兴地看到一本关于空间数据与空间分析不确定理论著作的出版。空间数据质量与不确定性理论是地球空间信息科学的最核心支撑理论之一。因此,本书的出版对完善和发展地球空间信息科学将起到积极的作用。

作者史文中博士于1985年毕业于武汉测绘科技大学(现武汉大学)航空摄影测量与遥感系,获学士学位并被评为全校的优秀毕业生之一,于1988年在该校获硕士学位。他于1994年在德国 Osnabrueck 大学以优异的成绩 (*magna cum laude*) 获得博士学位。随后应聘在香港理工大学任教至今。

史文中教授一直致力于地球空间信息科学的理论与技术的研究。本书系统地总结了他在过去十余年中在空间数据与空间分析不确定性理论方面的主要研究成果。

全书共五篇计十五章,系统地阐述空间数据与空间分析的不确定性理论体系。第一篇为概述,主要介绍不确定性理论的形成与发展,系统地论述了不确定性误差源,介绍了空间数据不确定性理论主要依赖的若干基础理论。第二篇是关于空间数据的不确定性理论模型,包括 GIS 目标的位置不确定性模型与理论、数字地面高程模型的不确定性建模,以及属性不确定性建模。第三篇是不确定性关系模型和空间分析的不确定性理论,其中包括属性与位置不确定性的综合、不确定性拓扑关系理论、空间叠置分析的不确定性建模、缓冲区分析的不确定性分析模型。第四篇是关于空间不确定性的表达问题,包括不确定性的可视化以及空间数据的元数据描述。第五篇给出了空间不确定性理论的应用,如基于不确定性的空间数据挖掘、地籍数据的质量控制,以及运用网络服务技术分发数据质量与空间分析的不确定信息。

纵观全文,本书具有以下特点。

科学性: 本书在科学上的贡献是为地球空间信息科学从确定性发展到不确定性从一个方面奠定了理论基础。所提出的一系列新的理论模型与方法建立在严密的数学理论基础之上,如基于置信区间理论的误差带证明、基于随机化图形学理论的面状目标不确定性模型、基于模糊拓扑学的不确定关系模型、基于解析几何与积分方法的 TIN 模型平均误差公式的推导、基于多进

* 此文为本书第一版的序。

制小波理论的影像融合技术等。

系统性：系统地构建了空间数据不确定性的研究体系和理论框架，创立了空间数据与空间分析的不确定性理论体系。包括位置不确定性模型（点、直线、曲线、面以及体目标的建模）、属性不确定性模型、位置与属性不确定性的集成、DEM 的精度估计、不确定空间关系、空间分析的不确定性理论、数据质量控制理论、不确定信息的可视化等。

创新性：在多个主要方面取得了创新性的成果，如创立了空间数据与空间分析的不确定性理论体系；发现了国际上使用了近 40 年空间线状要素的 ϵ -带误差模型的两个重大理论缺陷，对此建立了一系列新的误差带理论模型；在空间位置不确定性方面，系统地建立了点、直线、曲线、面和体目标的位置不确定性数学理论模型；在属性数据不确定方面，提出了综合属性不确定性和位置不确定性“S-带”模型等；在基于矢量 GIS 的空间分析方面，开创了一个新的不确定性国际理论前沿等。

前瞻性：在多个方面提出空间数据不确定性理论发展的新趋势，并积极地通过理论研究予以实现和发展。如首次将静态 GIS 数据的不确定性模型与方法延拓至动态的空间分析不确定性理论；创新性地实现了从数据的误差描述逐步地发展到数据的质量控制从而达到减小数据不确定性的实践；基于模糊拓扑学理论将定性的不确定性空间拓扑关系描述扩展至定量的不确定性空间拓扑关系表达等。

GIS 中空间数据与空间分析不确定理论研究涉及空间不确定性、时态不确定性、属性不确定性和语义不确定性，是一个值得深入研究并使之能用于解决实际问题的挑战性命题。中国年轻的学者在空间数据不确定性理论及其他领域的成就在国际上得到了认可，我感到十分欣慰。同时，也期望在不久的将来可以看到更多的中国学者在地球空间信息科学理论发展的不同领域取得卓越的成就，努力推动 GIS 理论和实践向前发展。

李德仁

2005 年 6 月 6 日于武汉珞珈山

Foreword^{*}

Michael F. Goodchild^{**}

On February 3 1998 a military aircraft of the US Marine Corps struck and severed the cables supporting a gondola in Cavalese, Italy, leading to the deaths of 20 people. The plane had wing and tail damage but was able to return to its base. It subsequently became clear that the cable-car was not shown on the maps being used by the pilot, and in the subsequent trial the pilot claimed also that the aircraft's height-measuring system had malfunctioned. Errors, omissions, and uncertainties in geographic information do not often result in international incidents of such significance, but it is an inescapable fact that any knowledge of the Earth's surface is subject to uncertainty of some kind, whether it be in the positions of features, their existence, or their description. It is impossible to know the exact location of anything on the Earth's surface, since our methods of measurement are always subject to error; and uncertainties creep into our maps, databases, and written records through a wide range of additional mechanisms. Many of the classification schemes used to map aspects of the Earth's surface, from soils to land use, are inherently uncertain, with the result that two people mapping the same area will never produce identical maps. Despite centuries of progress in mapping technology, the creation of geographic information remains as much an art as a science.

From this perspective it is important for users of geographic information to have some awareness of the uncertainties that are likely to influence their decisions—to know what the database does not tell them about the real world, and about the reliability of what it does tell them. As geographic information technologies have developed over the past few decades it has become clear that one ignores issues of uncertainty at ones peril. For example, regulations in many countries now make it difficult to construct in areas classified as wetland. Decisions are made daily about the uses to which private land can be put. But such decisions are clearly open to legal challenge if they can be shown to have been based on maps that are inherently uncertain.

* 此文为本书英文版的序(2010)。

** National Center for Geographic Information and Analysis, and Department of Geography, University of California, Santa Barbara, CA 93106-4060, USA. Phone +1 805 893 8049, FAX+1 805 893 3146, Email good@geog.ucsb.edu.

Research on the description and modeling of uncertainty in geographic information began in earnest in the late 1980s, and has accelerated over the past two decades. Today, a large literature describes successful efforts to address the issue, and tools are increasingly available to allow the effects of uncertainty to be propagated, so that uncertainties can be associated with the results of analysis as well as with the inputs.

John Shi has been one of the leaders in this research area. He has made very significant contributions, particularly in the modeling of uncertainties in geographic features of complex geometry, and has also made a very valuable contribution as the organizer of a series of conferences on spatial data quality, the International Symposia on Spatial Data Quality, that have occurred every two years since 1999. The conferences provide a forum for a very broadly based discussion of recent research on uncertainty that is not limited to any single paradigm or theoretical framework, because experience over the past two decades has shown that this problem of uncertainty is so pervasive, and so multidimensional, that no single approach can possibly address it.

The contents of this book provide an excellent introduction to this multidimensional problem. Early research tended to focus on error and accuracy, on the grounds that the creation of geographic information was similar to any problem of scientific measurement, and could be addressed through the application of the theory of errors. While this is conceptually simple, in reality geographic information tends to have some very awkward properties that complicate the approach enormously. Maps are not collections of independent measurements, but instead represent the culmination of a long and complex process that induces very strong correlations in errors. All positions will be subject to error, but nearby positions will have more similar errors than distant positions—in other words, relative errors of position over short distances tend to be much less than absolute errors. To handle this, models of uncertainty need to incorporate strong spatial autocorrelations, and to require comparatively advanced mathematics.

By the mid 1990s, however, it was clear that some aspects of the problem of uncertainty derived in part from the inherent vagueness of definitions, and that these could be handled much more effectively using the theoretical constructs of fuzzy sets. For example, we may not know exactly what is meant by wetland, but nevertheless it may make sense to be able to say that this area is more like wetland than that area. The approach was immediately attractive to many users of geographic information, who found it more intuitive and accessible than the statistical approach.

Both statistical and fuzzy frameworks are covered in this book, which provides a comprehensive overview of the current state of the field. At the same time it strongly reflects John Shi's own approaches, and the very significant contributions he has made. It should be indispensable reading for anyone interested or actively engaged in this research area, and desirable reading for anyone using geographic information to solve real prob-

lems. We have made much progress in the past two decades, and today few users of geographic information systems are willing to assume that their outputs are exact and correct just because they came from a computer. At the same time we are still some distance from the goal of placing a plus or minus on every output, and of incorporating uncertainty into every decision made with geographic information. But this book may help us get closer to that goal.

前 言

地理信息科学(GI Science)中,主要基础理论包括地理现象的数字表达方法、地学编码方法、地理数据的本质、不确定性理论及尺度理论。本书主要从理论研究的角度深入介绍其中一个基础理论——空间数据与分析的不确定性理论。

本书所介绍的空间数据与分析不确定性理论是作者在该领域中近 20 年研究工作的系统总结,特别是最新成果的介绍。全书共分为七篇:①概述;②空间数据不确定性理论;③空间模型不确定性理论;④空间分析不确定性理论;⑤空间数据质量控制理论;⑥空间数据质量信息表达方法;⑦结束语。

在第一篇概述部分,第 1 章和第 2 章分别介绍了不确定性的有关概念和来源。在此基础上,第 3 章介绍了不确定分析的数学基础,包括概率论、统计分析、模糊数学、证据理论、模糊拓扑学和信息论等。

第二篇介绍空间数据的不确定性理论。第 4 章系统地描述 GIS 中线特征误差模型。在回顾现有误差模型的基础上,介绍了我们所提出的从一维到多维的置信域误差模型及线特征通用统计分布模型,之后又拓展到曲线目标的误差模型、面状目标的误差模型。第 5 章介绍属性数据误差理论,主要有抽样方法、内部和外部测试、误差矩阵方法以及缺陷率统计模型等。第 6 章介绍位置与属性误差的综合问题,主要有基于概率及基于证据理论两种方法。

第三篇介绍空间模型的不确定性理论,其中第 7 章介绍空间目标间不确定性的模型拓扑关系的描述方法。第 8 章介绍数字高程模型不确定性模型。

第四篇介绍空间分析不确定性建模理论。第 9 章介绍叠置分析位置不确定性建模,采用解析与模拟的两种方法对叠置分析的位置不确定性进行建模。第 10 章介绍缓冲区空间分析位置不确定性建模理论,该方法是基于严格的解析数学推证的。第 11 章介绍了线简化空间分析位置不确定性建模方法。

第五篇介绍空间数据质量控制理论,其中第 12 章应用数据误差及测量中平差统计分析的方法,提出了地籍数据的质量控制方法。第 13 章给出了基于场的空间数据的质量控制方法,首先介绍采用基于点的变换模型的几何纠正方法来控制基于场的空间数据(卫星影像)的质量;在此基础上,提出了采用基于线矢量的变换模型进行基于场的空间数据质量控制。第 14 章介绍了提高数字高程模型质量的方法,重点描述两种改进的 DEM 插值方法:①混合插值方法;②双方向性插值方法。

第六篇侧重于不确定性的可视化及元数据建模方法。在第 15 章的可视化技术中,介绍了各种方法对于不确定性的可视化方法,如误差椭圆、三维显示、等值线表达等方法。第 16 章介绍了空间数据质量元数据,引入了面向对象的元数据组织方法,在对象级的层面上提供不确定性元数据。第 17 章主要介绍两个地理信息数据质量信息系统:

①空间数据质量检查系统；②基于 Web 服务的空间数据质量信息服务系统。

最后，为了为推进空间数据与分析不确定性领域知识的进一步发展，第 18 章（第七篇）阐述了空间数据与空间分析不确定性理论研究的未来发展方向。

希望本书的出版，可以起到抛砖引玉的作用——引起更多同行对空间数据及分析不确定性理论问题的研究兴趣，从而为地理信息科学理论发展奠定更为坚实的基础。

本书可作为研究人员、大学教师及研究生理论研究的参考文献及地理信息科学课程参考书。

本书的再版得到了科技部-国家高技术研究发展计划(课题号：2012AA12A305)和国家自然科学基金重点项目(项目号：41331175)的支持。

目 录

序

Foreword

前言

第一篇 概 述

第 1 章 绪论	3
1.1 不确定性	3
1.2 不确定性的普遍性	5
1.3 空间数据不确定性	6
1.4 本章小结	10
参考文献	10
第 2 章 空间数据与空间分析不确定性来源	12
2.1 引言	12
2.2 客观世界的不确定性	12
2.3 人类认知局限带来的不确定性	14
2.4 量测不确定性	15
2.5 空间分析和数据处理的不确定性	19
2.6 本章小结	21
参考文献	21
第 3 章 数学基础	22
3.1 引言	22
3.2 概率理论	22
3.3 数理统计理论	29
3.4 D-S 证据理论	38
3.5 模糊集合理论	42
3.6 粗集理论	50
3.7 信息论和熵	55
3.8 本章小结	60
参考文献	60

第二篇 空间数据不确定性理论

第 4 章 空间数据位置不确定性模型	63
4.1 引言	63
4.2 现有误差模型概述	64
4.3 空间实体的定义	65
4.4 点元的位置误差模型	69
4.5 置信域模型	71
4.6 线元的概率分布模型	80
4.7 线元的 G-带误差模型	83
4.8 线元的标准误差带及其质量指标	90
4.9 曲线的误差模型	102
4.10 多边形误差模型	112
4.11 本章小结	115
参考文献	116
第 5 章 属性不确定性建模	117
5.1 引言	117
5.2 属性不确定性建模的方法	121
5.3 本章小结	131
参考文献	132
第 6 章 位置与属性不确定性综合	134
6.1 引言	134
6.2 问题定义	135
6.3 “S-带”模型	136
6.4 基于概率论的解决方案	136
6.5 基于确定因子的解决方案	137
6.6 一个综合不确定性建模的例子	141
6.7 本章小结	146
参考文献	146

第三篇 空间模型不确定性理论

第 7 章 不确定性拓扑关系模型	149
7.1 拓扑关系模型综述	149
7.2 目标拓扑关系模型	151
7.3 拓扑关系的不确定性建模	155
7.4 应用实例	162
7.5 本章小结	164
参考文献	164

第 8 章 数字高程模型中的不确定性理论	166
8.1 引言	166
8.2 地表建模	166
8.3 DEM 误差的来源	170
8.4 TIN 模型精度估计	171
8.5 规则格网 DEM 的精度估计	178
8.6 本章小结	184
参考文献	185

第四篇 空间分析不确定性理论

第 9 章 叠置分析位置不确定性建模	189
9.1 引言	189
9.2 现有的矢量叠置分析模型的回顾	190
9.3 叠置分析生成的多边形的不确定性建模	190
9.4 本章小结	200
参考文献	200
第 10 章 缓冲区分析位置不确定性建模	201
10.1 引言	201
10.2 现有矢量缓冲区分析误差模型	202
10.3 基于概率论的缓冲区分析不确定性模型	203
10.4 本章小结	214
参考文献	214
第 11 章 线简化位置不确定性建模	215
11.1 引言	215
11.2 线简化中的不确定性	216
11.3 基于线性属性度量的误差模型	216
11.4 本章小结	221
参考文献	221

第五篇 空间数据质量控制理论

第 12 章 基于对象的空间数据质量控制	225
12.1 引言	225
12.2 区域对象误差的平差处理方法	226
12.3 地籍数据质量控制方法	233
12.4 实例研究	235
12.5 本章小结	241
参考文献	241

第 13 章 基于场的空间数据的质量控制	242
13.1 引言	242
13.2 二维变换模型	243
13.3 像方空间到物方空间的三维变换模型	244
13.4 基于线的变换模型	246
13.5 基于点变换模型的实验研究	248
13.6 基于线变换模型的实验研究	251
13.7 本章小结	253
参考文献	254
第 14 章 数字高程模型的改进插值方法	255
14.1 引言	255
14.2 混合插值方法	255
14.3 双方向性插值方法	261
14.4 本章小结	267
参考文献	267

第六篇 空间数据质量信息表达方法

第 15 章 空间数据不确定性的可视化	271
15.1 引言	271
15.2 误差椭圆方法	272
15.3 矢量箭头的描述方法	273
15.4 灰度图方法	278
15.5 色彩图方法	280
15.6 其他基于符号的方法	283
15.7 三维方法	284
15.8 动画方法	286
15.9 本章小结	287
参考文献	287
第 16 章 空间数据质量元数据	289
16.1 简介	289
16.2 元数据质量指标	293
16.3 面向对象的元数据系统	302
16.4 本章小结	305
参考文献	305
第 17 章 GIS 数据质量信息服务系统	306
17.1 引言	306
17.2 SDQS 数据质量检查系统	307
17.3 基于 Web 服务的数据质量服务系统 DQIS	319
17.4 本章小结	328

第七篇 结束语

第 18 章 结束语	333
18.1 未来研究方向	333
18.2 有关数学理论	336
18.3 应用实践问题	337
18.4 与其他学科的相互关系	338
18.5 总结	339

彩图

第一篇 概 述

