

高等学校教材



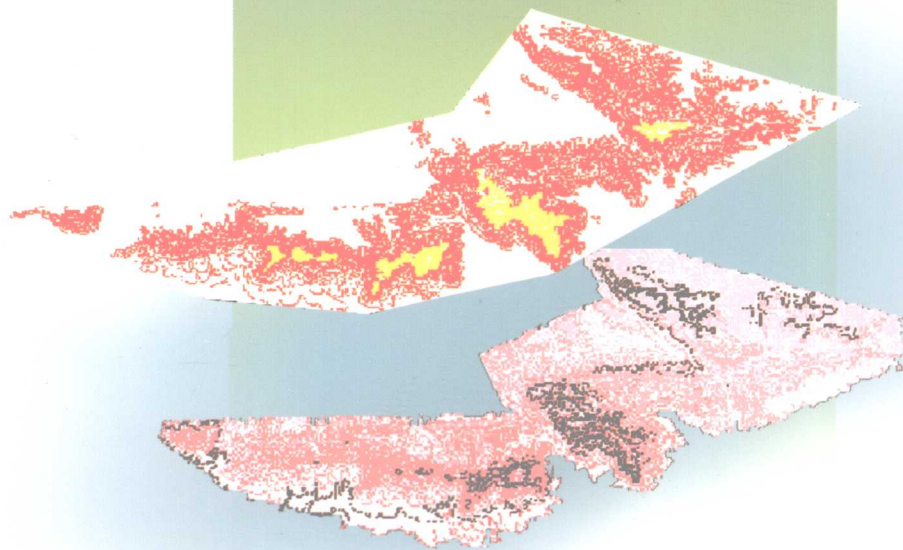
测绘科技专著出版基金资助

CEHUI KEJI ZHUANZHU CHUBAN JIJIN ZIZHU

地理计算原理

李霖 应申 朱海红 编著

与方法



PRINCIPLES AND METHODS OF GEOCOMPUTATION

测绘出版社

高等学校教材
测绘科技专著出版基金资助

地理计算原理与方法

Principles and Methods of GeoComputation

李霖 应申 朱海红 编著

测绘出版社

· 北京 ·

内 容 简 介

地理计算属当前 GIS 和地理学研究的前沿问题。该书主要内容:一是紧扣地理计算的地理性和计算性,比较详细地介绍了地理计算的内涵、特征和应用前景;二是介绍了高性能计算环境下各种计算技术在地理空间分析和地理计算中的应用;三是以某些新的技术与方法为例(如元胞自动机,分形分维)介绍了其在地理计算中的具体应用;四是分析和讨论了地图可视化、虚拟地理环境等可视化技术给地理计算注入的活力及带来的问题。

本书适合于地理学、GIS 及相关专业的本科生和研究生做参考教材,也可从事相关工作的科技人员和管理人员提供业务参考。

图书在版编目(CIP)数据

地理计算原理与方法/李霖,应申,朱海红编著. —北京:测绘出版社,2008.8

ISBN 978-7-5030-1869-5

I. 地… II. ①李…②应…③朱… III. 计算方法—应用—地理学 IV. K90

中国版本图书馆 CIP 数据核字(2008)第 100197 号

责任编辑 杨蓬莲

封面设计 杨晓明

出版发行 测绘出版社

社 址 北京西城区复外三里河路 50 号

邮政编码 100045

电 话 010-68512386 68531558

网 址 www.sinomaps.com

印 刷 北京通州区次渠印刷厂

经 销 新华书店

成品规格 184mm×260mm

印 张 16.75

字 数 418 千字

版 次 2008 年 8 月第 1 版

印 次 2008 年 8 月第 1 次印刷

印 数 0001—3000

定 价 33.00 元

书 号 ISBN 978-7-5030-1869-5/P·488

如有印装质量问题,请与我社发行部联系

前言

地理数据的采集、定位和绘制是地理分析的第一步,接下来的分析、建模是地理分析最重要的部分,也是原始地理数据能够产生效能的关键所在。地理计算是个令人耳目一新的话题,但它到底是什么?本书试图让读者对地理计算的体系和应用有一个比较直观的认识和了解。

随着计算机的运算速度越来越快,计算科学已经成为解决许多科学问题的有效方法,计算科学对地理学像对社会学一样带来了巨大的冲击,地理计算(GeoComputation,有文献译为“地学计算”,本书采用“地理计算”这个术语)就是计算科学在地理或地学环境中的应用(这里的地学环境包括人文系统和自然系统)。地理计算从字面意义上来看和计算科学是密不可分的,计算科学是20世纪末发展起来的交叉学科,但目前还没有统一的概念。计算科学依赖于计算技术的发展和硬件性能的提高,使用计算机来研究各种科学问题,实现科学调查中的各种理论和方法。一旦计算机硬件达到要求,可以通过科学计算可视化来观察各种实验的中间过程。可以说,科学计算能让研究者通过模拟来测试现有的理论,并有可能创建新的理论。

地理计算的出现与发展,得益于计算机技术和计算科学理论与方法的飞速发展。20世纪90年代并行超级计算机硬件的成功实现,GPS,RS,GIS技术在获取大容量、整体性地理数据信息中的成功应用,以超级计算机为基础的一系列高性能计算新方法的实现,使计算、实验与理论共同构成了人类认知客观世界的有效科学工具。地理计算已不是传统意义上的利用计算机求解地理问题的计算,它是以向量或并行处理器为基础的超级计算机为工具,对整体的、大容量的资料所表现的地理问题实施高性能计算,探索构筑新的地理学理论和应用模型。

当前地理计算的研究包括很多部分,其中比较重要的包括:高性能计算、人工智能以及其更抽象的表达——计算智能、处理大规模空间数据库的全球GIS。当然,这些不是唯一,还包括许多传统的内容,如统计技术、数学建模、计算机地理模拟等。地理计算的研究内容归纳为5个方面:①理论发展;②经验分析;③建模和模拟;④社会方面;⑤自动分析和建模功能。

地理计算是交叉学科,其相关学科包括统计学、模式识别、人工智能、神经计算以及GIS等,本书将简述有关学科的内容及其与地理计算的结合。

地理学的许多研究方法,无论是定性的或是定量的,在高性能计算的时代,都将成为可能。地理计算的概念、方法、模型对一部分地理学家来说,可能都还比较陌生,但正如GIS在经过相当长一段时间的地学领域研究和实践后才为大地学工作者所接受的那样,地理计算也将为地理学家们所接受而成为地理科学殿堂中的重要一员。可以肯定地说,地理计算的原理、方法、模型、算法的不断完善,高性能计算设备的发展和普及,地理学家们的广泛参与,终将把地理科学推向一个全新的水平。

本书的主要内容如下:

- 介绍了地理计算的历史渊源和发展过程,地理计算的概念、特点以及当前的主要研究和发展,也介绍了地理计算中的主要方法和分析手段。

- 针对地理计算中的计算特点,介绍了高性能计算技术的发展,以及并行计算、网络计算和网格计算相关概念和技术体系。

• 地理空间数据是地理计算的对象,作为地理计算知识完整性的组成部分,本书简述了地理空间数据的基本特征、表达形式以及组织管理方法。

• 地理计算利用各种传统和新的方法来建模,分析各种地理现象,本书以元胞自动机、专家系统、分形方法、遗传算法等技术和方法为例,介绍了相关原理和基本技术,相关案例都可以作为地理计算的具体应用。

• 地理可视化是地理计算融合计算技术和地理空间数据的重要体现,本书介绍了地图可视化、GIS 可视化等基本原理和方法,这些可视化技术,是地理计算技术的重要组成部分。

在本书的编撰过程中,一批富有朝气的青年学者为此作出了积极的贡献,无论在资料收集、整理中,还是在具体案例的研究和实施中,都有他们的身影。他们是:张晓通、桂胜、苗蕾、赵虎、胡洁、刘小飞、李茂勋、王琤、汪飞等。在此,对他们表示衷心的感谢。

本书适合于地理学、GIS 及相关专业的本科生和研究生作参考教材,也可供从事 GIS、地理学、资源与环境信息系统、土地、规划、旅游、建筑等相应工作的科技人员和管理人员参考。

当然,由于作者水平和眼界有限,书中难免出现一些不妥之处,有些观点可能以偏盖全,甚至出现错误,恳请读者批评指正。

作者

2008 年 4 月

目 录

第 1 章 地理计算概述	1
§ 1.1 地理学中的定量分析	1
1.1.1 计量地理学的发展	1
1.1.2 地理学的定量分析	2
1.1.3 计量地理和地理计算的关系	3
§ 1.2 地理计算的基本概念	4
1.2.1 地理计算的起源	4
1.2.2 地理计算的特点	5
1.2.3 地理计算的概念	7
1.2.4 地理计算的研究和发展	9
§ 1.3 地理计算中的主要方法	11
§ 1.4 地理计算与空间分析	14
第 2 章 基本计算技术	16
§ 2.1 高性能计算	16
§ 2.2 并行计算	18
2.2.1 什么是并行计算	19
2.2.2 并行计算机的类型	19
2.2.3 并行的方法和分解策略	21
2.2.4 何时使用并行计算	22
2.2.5 怎样使用并行编程	22
2.2.6 并行计算的体系结构	22
2.2.7 网络并行计算系统	27
2.2.8 并行算法和并行效率测评	31
2.2.9 地理学中的超级计算	33
2.2.10 并行可视计算	35
§ 2.3 网络计算	40
2.3.1 网络计算的含义	40
2.3.2 网络计算机	41
2.3.3 网络计算的结构模式	41
2.3.4 网络计算的特点和优点	42
2.3.5 网络计算	43
2.3.6 网络计算与 GIS	44
§ 2.4 网格计算	45
2.4.1 元计算与网格计算	45
2.4.2 网格的概念与意义	45

2.4.3	网格体系结构	47
2.4.4	网格计算的特点及其关键技术	50
2.4.5	著名的网格工程	52
2.4.6	网格 GIS	53
第3章	地理空间数据	56
§ 3.1	引言	56
§ 3.2	地理空间数据的特征	58
3.2.1	概述	58
3.2.2	科学数据的特征	58
3.2.3	地理空间数据的特征	59
§ 3.3	地理空间数据的表达、组织与管理	63
3.3.1	地理空间数据的表达	63
3.3.2	地理空间数据的组织与管理	75
第4章	元胞自动机原理及其应用	84
§ 4.1	引言	84
§ 4.2	元胞自动机理论	84
4.2.1	元胞自动机概念	84
4.2.2	元胞自动机的提出与发展	85
4.2.3	元胞自动机的构成	86
4.2.4	几种典型的元胞自动机模型	87
4.2.5	元胞自动机的一般特征	88
4.2.6	元胞自动机的分类	89
4.2.7	元胞自动机的应用	89
§ 4.3	元胞自动机与地理系统研究	90
4.3.1	元胞自动机用于地理系统研究的必然性	90
4.3.2	元胞自动机用于地理系统研究的可行性	90
§ 4.4	元胞自动机的扩展	91
§ 4.5	元胞自动机在地理学中的应用	92
4.5.1	元胞自动机在地理学中的应用概述	92
4.5.2	元胞自动机模型在地理学中的几种主要应用模型	92
§ 4.6	基于元胞自动机的城市扩张模拟	97
4.6.1	基本原理	98
4.6.2	基于元胞自动机的实现	99
4.6.3	实验	100
§ 4.7	本章总结	103
第5章	专家系统及其与 GIS 的结合	105
§ 5.1	专家系统理论	105
5.1.1	人工智能和专家系统	105
5.1.2	专家系统的发展	106
5.1.3	专家系统概述	107

5.1.4 专家系统的基本组成和结构	107
5.1.5 专家系统的特征	109
5.1.6 专家系统的分类	110
5.1.7 专家系统中的知识表示和知识库的组织	111
5.1.8 推理方式和推理机工作机制	112
5.1.9 产生式规则专家系统	113
5.1.10 专家系统的开发	116
5.1.11 专家系统的应用	119
§ 5.2 GIS 与专家系统的结合	119
5.2.1 GIS 与专家系统(ES)的关系	119
5.2.2 GIS 与专家系统的结合	120
5.2.3 地理专家系统概述	121
5.2.4 地理专家系统国内外研究现状及进展	121
§ 5.3 GIS 与专家系统一体化的发展趋势	123
第 6 章 遗传算法及其应用	125
§ 6.1 遗传算法的基本理论	125
6.1.1 引言	125
6.1.2 遗传算法基本操作	127
6.1.3 遗传算法表示模式	132
6.1.4 遗传算法求解问题的过程	138
§ 6.2 遗传算法在 GIS 领域中的应用	142
6.2.1 遗传程序设计	143
6.2.2 遗传算法在 GIS 中的应用	147
6.2.3 存在的问题	161
§ 6.3 本章小结	165
第 7 章 分形方法及其应用	166
§ 7.1 引言	166
§ 7.2 分形理论的历史与现状	166
§ 7.3 分形理论的基本应用	167
7.3.1 影像处理	167
7.3.2 地形模拟	167
7.3.3 地图制图	168
7.3.4 水系分析	168
7.3.5 地震预报和其他应用	169
§ 7.4 分形及分形维数的定义	169
7.4.1 分形	169
7.4.2 分形维数	170
§ 7.5 分形维数的基本计算方法	171
7.5.1 改变观察尺度法	171
7.5.2 测度关系法	171

7.5.3	分布函数法	172
7.5.4	计盒法	172
7.5.5	频谱法	172
7.5.6	变异元法	172
7.5.7	几种方法的比较	173
§ 7.6	典型地貌的分形分析实验	174
7.6.1	背景知识	174
7.6.2	地貌的分形维数计算方法	176
7.6.3	研究过程	182
7.6.4	研究流程	182
7.6.5	结果分析	186
§ 7.7	本章小结	189
第 8 章	动态地图可视化与 3 维可视化	190
§ 8.1	引言	190
8.1.1	科学可视化的兴起与地学领域可视化发展	190
8.1.2	可视化的相关概念	192
§ 8.2	动态地图可视化	193
8.2.1	地图可视化	194
8.2.2	动态地图可视化的出现与发展	196
8.2.3	动态地图可视化的研究现状	198
8.2.4	地图可视化的动态使用过程	201
8.2.5	动态地图可视化的实现	210
8.2.6	人口变化动态地图	222
8.2.7	地理可视化	227
8.2.8	GIS 可视化	228
8.2.9	可视化相关技术	233
§ 8.3	3 维 GIS 可视化	234
8.3.1	从 2 维 GIS 到 3 维 GIS	234
8.3.2	3 维可视化相关技术	235
8.3.3	3 维 GIS 空间数据模型	236
8.3.4	3 维可视化实例分析	238
§ 8.4	地学可视化的特征与研究框架	247
§ 8.5	作为地理计算工具的可视化	248
8.5.1	地理数据的特点	248
8.5.2	视觉属性	248
8.5.3	可视化的设计	248
8.5.4	可视化在地球科学中的应用	249
8.5.5	视觉真实性	249
8.5.6	动画和互感器的使用	250
参考文献		251

第1章 地理计算概述

§ 1.1 地理学中的定量分析

1.1.1 计量地理学的发展

世界上的任何事物都可以用数值来度量,对地理和地理信息也是如此。地理学的发展总是与数学有着紧密的联系。在现代地理学中,传统方法是数学方法的基础,数学方法是传统方法的重要补充,新型计算是地理计算的研究热点。

地理学的发展史可划分为3个基本阶段:①古代地理学,以记载地理知识为主体;②近代地理学,对各种地理现象进行条理化归纳,并对它们之间的关系进行解释性描述;③现代地理学,采用定性与定量相结合的方法,规范研究与实证研究并举,解释各种地理现象的内在机制并预测其未来演变(徐建华,2006)。古代地理学和近代地理学中的数学方法限于定量地描述、记载和解释地理知识。现代地理学中运用数学方法,是为了深入地进行量化研究,揭示地理现象发生、发展的内在机制及运动规律,从而为地理系统的预测及优化调控提供科学依据。

Rees等(1998)认为,地理计算的历史与计算机在科学研究中的引进和应用的历史一样悠久,至少可以追溯到20世纪50年代末至60年代初,即地理学中的计量革命。由于受当时计算机的内存的大小和中央处理器的运算速度的限制,面对现实地理世界的特殊性和复杂性,当时的地理计算仅仅应用了统计学等数学的方法,而这些方法早在计算机应用之前就已经存在。60年代以后,由于计算机性能的不断提高,并行计算技术的应用,极大地提高了计算机求解地理问题的能力,推动了计量地理学向地理计算学的发展。总的来说,计量地理学的发展经历了4个阶段(徐建华,2006)。

第一阶段(20世纪50年代末到70年代)为计量革命(统计模型):把统计学方法引入地理学研究领域,构造了一系列统计量来定量地描述地理要素的分布特征,应用了各种概率分布函数、平均值、方差、标准参数以及简单的两要素间的一元线性差、变异系数等统计特征回归分析方法。70年代末期,多元统计分析方法和电子计算机技术在地理学的研究中得到了广泛应用。

第二阶段(20世纪70年代末到80年代)为数学模型革命(数理模型和规划模型):数学方法有概率论与数理统计方法、规划方法、决策方法、网络分析方法、数学物理方法、模糊数学方法、分形几何学方法、非线性分析方法、投入产出分析方法等。

第三阶段(20世纪80年代)为地理信息系统(Geographic Information System, GIS)革命,标志为1983年Dobson提出自动地理学(automated geography)。

第四阶段(20世纪90年代初至今),由于后现代数学方法和GIS的引入,地理学的定量化和理论化进入了全新的阶段,新的定量地理学与计算机科学汇流发展成为地理计算学。

随着计算机及相关的信息获取技术的不断进步,各种类型的数据库逐步建立并以越来越

低的成本提供大批量数据,科学研究的重点自然地转向现有数据库中的数据挖掘或者称为隐含信息提取。由于空间数据量大、多维和存在自相关等原因,其数据挖掘较其他数据类型更为复杂,地理计算成为定理地理学的发展趋势。1994年在英国里兹大学正式建立全球第一个计算地理中心,20世纪90年代中期国外学者正式创立“地理计算学”(geocomputation)一词。1996年起,已多次举行全球地理计算学术年会,出版论文专集。作为计量地理学的深层次发展,地理计算学的出现与发展,对整个地理学科,尤其是对人文地理学、经济地理学的理论模型和应用研究,已经产生了并将继续产生深远的影响。

但是,有关地理学的定量化、数学化等还有一定的纷争。定量化不等于数学化,有些地理问题虽然不能定量化,但却可以建立数学模型,只要数学化和定量化双管齐下,地理学就可望从经验科学上升成为理论科学(陈彦光,2005)。现在对计量地理学存在一定的争论,包括反定量化、定量化和非定量化等(徐建华,2007)。

- 反定量化:反对地理学定量化研究,认为地理现象十分复杂,不能用简单的数学方法来解释,对数学方法采取拒绝和否定态度。代表人物有 David Smith 和 Philip Ogden 等。

- 定量化:推崇地理学定量化研究,认为数学方法不仅是一种分析技术,而且能够导出普遍性的规律,能够解决地理学传统研究方法所不能解决的理论问题。代表人物如 Christaller, Bunge, Chorley 以及 Haggett 等。

- 非定量化:认为数学方法只是地理学研究方法之一,只能用来研究地理要素之间的数量关系和地理事物的空间格局,不能用来描述和解释地理规律,不能导出地理学理论,但其观点摇摆不定。

计量地理学有广义和狭义之分,广义的计量地理学内容广博,包括所有借助数学和统计工具的地理研究分支;而狭义的计量地理学,主要是指以统计分析为核心的地理学量化处理和计算方法。早先的定量地理学可以概略地分为数学地理学和统计地理学,前者对应于理论地理学,后者对应于狭义的计量地理学;前者着重于地理学的建模与推理,目标是建设地理学的基础理论;后者着重于观测数据的整理与分析,不以建立基础理论为目标,主要用于解决应用研究的系统和预测分析问题。数学地理学是理论地理学的核心,但数学地理学并非都是定量分析,因为并不是所有的数学模型都要与数据进行拟合,而统计地理学则一定是计量分析。但是,如果需要对数学模型进行拟合检验和参数考察,则数学地理研究有时需要用到统计地理学的有关方法或者分析技术,如回归分析等。

近年来,国际上从事地理计量分析最为活跃的人物之一,Openshaw 建议将统计地理学与数学地理学区分开来,并将20世纪50年代至60年代的计量地理学称为古典的计量地理学,而将近年的发展结果称为计算的人文地理学,可以肯定,对定量地理学的外延适当加以限制有利于这门学科朝着内涵的方向深入发展。英国著名定量地理学家 Fotheringham(2000)认为,定量地理学主要由空间数据的数值分析,空间理论的开发,空间过程数学模型的建立与检验几方面构成。所有这些研究的目的是为了加强我们对空间过程的理解,理解空间过程可以是直接的,也可以是间接的。在间接的情况下,空间过程需要借助一定的逻辑推断才能得以认识。

1.1.2 地理学的定量分析

总结起来,地理学的定量分析和应用有如下几点(徐建华,2002;2006)。

- 分布型分析:对地理要素的空间分布特征及规律进行定量分析。
- 相互关系分析:对地理要素、地理事物之间的相互关系进行定量分析。
- 类型研究:对地理事物的类型和各种地理区域进行定量划分。
- 网络分析:对水系、交通网络、行政区划、经济区域等地理要素的空间结构进行定量分析。
- 趋势面分析:通过地理要素的趋势等值线图,展示地理要素的空间分布规律。
- 空间相互作用分析:定量分析各种“地理流”在不同区域之间流动的方向、强度以及相互作用。
- 系统仿真研究:步骤为:①对复杂地理系统的各种系统要素之间的相互关系与反馈机制进行分析,构造系统结构;②建立描述系统的数学模型;③以适当的计算方法与算法语言将数学模型转化为计算机可以识别运行的工作模型;④运行工作模型,对真实系统进行模拟仿真,从而揭示其运行机制与规律。
- 过程模拟与预测研究:通过对地理过程的模拟,定量地揭示地理事物及地理现象随时间、空间变化的规律,预测其未来发展趋势。
- 空间扩散研究:定量地揭示各种地理现象,包括自然现象、经济现象、社会现象、文化现象、技术现象在地理空间中的扩散规律。
- 地理系统的优化调控研究:运用系统控制论的有关原理与方法,研究人地相互作用的地理系统的优化调控问题,寻找人口、资源、环境与社会经济协调发展的方法、途径与措施。
- 地理系统的复杂性研究,目前还未形成一种非常有效的方法体系。

1.1.3 计量地理和地理计算的关系

数学方法是人们进行数学运算和求解的工具,能以严密的逻辑和简洁的形式描述复杂的问题、表述丰富的实质性思想。但是在地理学研究中,数学方法有其局限性,现代地理学中数学方法的形成和发展与计算机应用技术密切相关,地理学的定量分析与采用的数据和模型紧密相关。在计量地理的数据处理中要注意地理数据的筛选与质量检验,以及模型的选择和建造中的变量等问题。

研究一些复杂的地理问题,需要综合应用多种数学方法,建立一系列具有分析、模拟、仿真、预测、规划、决策、调控等多种功能的众多模型组成的模型系统。同时还需要海量的地理数据,这些都离不开GIS的支持。GIS的基本技术以及建造空间分析模型都需要借助有关的数学方法来实现。近几年来出现的基于知识的空间决策支持系统就是数学方法、人工智能与GIS技术在地理学应用研究领域相互结合的成功典范,而正是它们的结合促使了新学科——地理计算的发展。

地理计算和计量地理的关系如何呢?其实关系很简单,地理计算的范围要大一些,它包含计量地理的所有方法和工具,但它会逐渐地削减计量地理的弱点,重新表达计量地理来使之适应21世纪的需求(Openshaw,1998)。地理计算为地理环境下的应用科学提供了新的前景。另有一种观点认为地理计算和计量地理没有太大的差别,仅是将自然和人文地理问题中的计算方法应用到地理科学中去。

地理计算关心新的计算技术、算法与范式,它们之间相互独立,可以充分利用高性能计算(HPC)技术。地理计算主要包括4个边缘技术。

- (1) 地理信息系统(GIS)技术:用来创建数据;
- (2) 人工智能(AI)和计算智能(CI)技术:用于提供尖端技术和工具;
- (3) 高性能计算(HPC)技术:用于提供动力;
- (4) 地理计算科学理论:用于提供理论基础。

因此,地理计算不仅仅是将计算机应用到地理中,也不是企图将几十年前衰落的计量地理再复活,更不是用计算机来实现数学,大规模的数学计算仅仅是它的一个特点。Openshaw(2000)指出,由于现在的计算机运算速度还不够快,也没有足够大的内存来使地理计算应用转换为实际可行的一般应用,所以说地理计算的高性能计算是全新的。这种高性能不再是1994年的Cray T3D(512个处理器、32 GB内存、40万亿次)和1998年的Cray T3E(576个处理器、148 GB内存、122万亿次),可以预见,在不久的将来高性能计算可能超过1亿万亿次。

我们是否需要这种超性能的计算能力呢?许多地理学者认为,现在的PC机已远比20世纪七八十年代的计算机性能好多了,而那时就能做到分析、建模和统计,现在的条件应该更没问题。某种程度上讲他们的说法是正确的,但是如果在高性能计算时代还局限于这个条件,就不能推进地理学的新发展。诚然,在低性能计算条件下的各种数学模型和统计分析方法现在还都在使用,但事实上这些模型和方法的计算结果与现在使用的技术所产生的结果差得很远。

Macmillan(1998)发现地理计算和传统的计量地理之间具有较强的关联性,因为它们都源自相同的科学理论。这种现象可以简述为地理计算囊括了计量地理,但是地理计算比计量地理涵盖的范围要广。现代的计量地理可以认为是从低速计算时代中遗留下来的各种统计和数学方法的堆积,仅仅是计算技术中很小的一部分,它们不是以数据为中心的,这种技术和相应观点在现代计量地理中仍存在。为了弥补计算技术的不足,解析逼近和更高效的数学处理方法后来在计量地理中出现。事实上,近年来计算技术得到了飞速发展,Cray T3E运行12小时的程序让一般PC机运行则需要4~8年。地理计算的挑战就是发展和创造思想、方法、模型和范式,利用不断增加的计算速度来做地理环境中的“有用的、值得做的、革新性的”科学研究。

§ 1.2 地理计算的基本概念

地理科学不仅仅是测绘方面的技术,地理数据的采集、定位和绘制是地理分析的第一步;接下来的分析、建模是地理分析最重要的部分,也是原始地理数据能够产生效能的关键所在。地理计算是个令人耳目一新的话题,但它到底是什么?

地理计算就是使用各种地理及相关数据,拓展出各种科学的方法和工具,解决地理学中的学术、理论和应用问题。它将计算机科学和人工智能中的有关技术和概念转变成地理信息分析中的操作工具。从计算机发展并应用到地理学中的那一刻开始,许多学者都在不断探讨这个问题。

1.2.1 地理计算的起源

地理计算从字面意义上来看和计算科学是密不可分的,计算科学是20世纪末发展起来的交叉学科,但目前还没有统一的概念。计算科学依赖于计算技术的发展和硬件性能提高,使用计算机来研究各种科学问题,实现科学调查中的各种理论和方法。一旦计算机硬件达到要求,

那么计算问题就会变成物理实验的替代品,并且还可以通过科学计算可视化来观察各种实验的中间过程。可以说,科学计算能让研究者通过模拟来测试现有的理论,并有可能创建新的理论。

随着计算机的运算速度越来越快,计算科学已经成为解决许多科学问题的有效方法,同时它作为第三代方法,填补了物理实验和解析方法之间的空白。许多复杂的现象(如核爆)采用解析的方法无法处理,用物理实验的方法太昂贵或太危险,而计算机模拟技术却可以为其提供定性和定量的分析。

计算科学对地理学和社会学一样带来了巨大的冲击,地理计算(GeoComputation)就是计算科学在地理或地学环境中(Geo)的应用,这里的地学环境包括人文系统和自然系统。英文字母“C”大写是 Bob Abrahams 倡导的,用来强调计算的成分及其重要性。

地理计算的出现与发展,得益于计算机技术和计算科学理论与方法的飞速发展。20世纪90年代并行超级计算机硬件的成功实现,GPS,RS,GIS技术在获取大容量、整体性地理数据信息中的成功应用,以超级计算机为基础的一系列高性能计算新方法的实现,使计算、实验与理论共同构成了人类认知客观世界的有效科学工具。地理计算已不是传统意义上的利用计算机求解地理问题的计算,它以向量或并行处理器为基础的超级计算机为工具,对整体的、大容量的资料所表现的地理问题实施高性能计算,探索构筑新的地理学理论和应用模型(徐建华,2006)。

1994年英国里兹大学成立了地理计算中心,其最开始的目标是为人文地理发展一些新的计算范式,但后来扩展到自然地理的许多领域中。Openshaw(1994)称之为计算人文地理研究时代,后来发现其研究的内容却是自然地理比人文地理多,而 GeoComputation(地理计算)一词更好地解决了这个问题。Computational/Computing Geography(计算地理)的概念太狭小了,有许多地理环境中的问题本应该在其范围内而事实上却被排除在外。

1.2.2 地理计算的特点

地理计算不仅仅是计算机在地理学中的应用,也不仅仅是关于计算,它是以大规模的科学计算为工具进行的所有地理领域的研究,包括有关的数学建模、模拟、统计等。以地理计算的英文构词来看,GeoComputation由前缀 Geo和主词 Computation组合而成,前者指地理计算要做什么,后者则指如何做。Openshaw(1998)指出了地理计算的两个特点:Geo主题和高性能计算。

1. Geo 主题

地理学,自其产生之日起,就与数学有着不解之缘。在古代,地理学与数学的源泉科学——几何学,几乎都是研究地表现象的。《辞海》关于几何学的解释是:“古代埃及为兴建尼罗河水利工程,曾经进行过测地工作,它逐渐发展为几何学。”因此,在来自希腊文的西方文字中,几何学有“测地”之意,如几何学的英文 Geometry,与地理学(Geography)、地貌学(Geomorphology)、地植物学(Geobotany)、地生态学(Geoecology)等术语有着一个共同的前缀“Geo”。

在古代地理学时期,人们为了测算河流长度,山体高度,土地面积,不得不运用几何学原理和方法。在近代地理学时期,经济学中的区位论被移植到地理学中,开创了地理学运用分析数学的先河。但“Geo”是指“地理”、“地质”还是泛指“地学”,很难作出定论,但可从数量地理学

20 世纪 80 年代以来发展的几个重要事件中得到一些启示。1983 年,随着 GIS 技术在地理学中的应用不断扩展,Dobson(1983)首次提出了“自动地理学”(Automated Geography)的概念,并将其定义为“描述空间性质,解释地理现象,求解地理问题的计算方法和计算技术”。90 年代初,在著名理论与数量地理学家 Wilson 教授领衔的英国里兹大学地理学院成立了全球第一个地理计算中心,正式启用了计算地理学的词语。1994 年 Openshaw 在应用人文地理学(Applied Human Geography)的基础上,综合提出了计算人文地理学(Computational Human Geography)的概念和研究议程。1996 年,里兹大学召开了第一次国际地理计算大会,并正式定名为 GeoComputation,或许这是“地理计算”合成新词语的第一次出现,会议组织者将会议的宗旨和新词语的应用定义为“利用不断发展中的高性能计算机和计算方法,对各种地理求解问题的研究进行聚合”(Rees, et al, 1998)。1998 年,在 GeoComputation 的会议公告中,则作了进一步定义:“地理计算学代表了计算机科学、地理学、地理信息科学、数学和统计学的聚合和趋同”(Unwin, 1998)。看来,GeoComputation 的前缀“Geo”更多的是地理的意义,而不一定是泛指意义上的“地学”概念。历次会议的组织者和与会者,如美国的 Goodchild, Kelly, 英国的 Batty, Openshaw, Rees, 奥地利的 Fischer 等都是当今最负盛名的理论和计量地理学家,他们把握了 GeoComputation 的发展主流,也使地理计算的地理学色彩更加明显。

地理计算学科上的关注点,不仅仅是地理学,还处理所有类型的地理或空间信息。计量地理学里就没有多少强调地理方面的东西,现在还很惊奇为什么它里面应用的各种方法竟然和地理相关性不大(Openshaw, 2000)。可能是他们认为地理数据和其他数据一样,就将自然科学中的方法原封不动地引入地理学中。应该来说,这种引进对初期计量地理学的发展起着推动作用,但是却减缓了人们认识地理数据的特殊性。Geo 特征在地理计算中是十分重要的,也是其主要特色,以前没有得到足够的认识是没有充分意识到其复杂性。

2. 高性能计算

计算是地理计算的方法和手段,地理计算是通过计算来寻找已有问题的解决方案或者提高其处理能力,以及处理新的或以前未解决的问题。

在科学意义上讲,计算(computation)这个词可有两层意义:广义上泛指利用计算工具(计算机)进行各种类型的分析,无论是其是量化分析或其他形式的分析,如果以计算机为工具,都可以被描述为“计算的”(computational)。地理计算即是这种凭借计算机工具对地理学问题进行定量或非定量分析的抽象概括和综合研究。狭义上讲,计算(computation)指的是一组具体有序的“计数”(counting)、“计算”(calculating)、“估算”(reckoning)或“估计”(estimating)等所有产生量化分析结果的操作运算,是一种自建模至获取所求结果这一过程中模型运行的手段。因此,地理计算的内涵可定义为对地理学时间与空间问题所进行的基于计算机的量化分析。

计算所描述的任何模型,无论是其是广义或是狭义上的,都包含有特殊的运算对象,即运算域(operand)和运算算子(operator)集合而成的某种数学表达式。对于以时空演变为特征的地理现象,其运算所涉及的运算域和算子都可以是空间的或是非空间的,空间与非空间的运算域,与运算算子所包容的空间态与非空间态相互配合与依存,构成整个地理计算的对象与方法体系。如空间运算域可以是几何的影像地图、体积或形状,也可能是拓扑的图、单纯图形,而空间算子可以是经典的几何变换(平移、旋转、反射),也可以是拓扑变换(伸缩、挠曲、卷迭等)。它们的组合体现在地理模型的类型与结构上,地理计算的模型将涉及元胞自动机、分形、形状

语法(shape grammara)、渗流(percolation),以及其他一些包含有递归模式生成的模型,而在技术方法上将体现在地理影像的几何变换、拓扑变换、地理主题的虚拟现实表现,以及包含了基本空间算子的所有基于GIS的方法。

计算意味着特殊的范式,它是基于数值逼近,而不是解析(Openshaw,2000)。地理计算是基于数据驱动的高性能计算的归纳工具,并试图解决以前不能解决的问题,而不是推理演绎的方法。从数据挖掘的角度来说,它是数据驱动的,或者说大规模的实验仅仅是个理论上的工具,用来理解系统在建模和模拟时的复杂度。

1.2.3 地理计算的概念

地理计算被认为是继GIS之后地理学的一次继续革命,预计在21世纪将会取得快速的发展和重要的成就(Openshaw,1998)。地理计算是地理学、数学、计算科学、GIS等诸多学科与技术交融的产物,其发展源流之一就是计量地理学,这一点从Openshaw对定量地理学的时段性分类可以看出。活动在地理计算领域的人员构成具有以下几个特点:

- (1)参与地理计算发展运动的学者,包括许多在计量地理学领域十分活跃的人物;
- (2)极力倡导地理计算理论的地理学家,如Openshaw在计量地理学方面有独到、精辟的见解;
- (3)从事地理分形等理论变革的少壮派骨干,如Longley等在地理计算领域非常活跃。

地理计算的概念最先是由Openshaw于1996年9月在第一次地理计算国际会议上正式提出的。Rees和Turton(1998)提议将其定义为:“应用计算技术求解地理问题的理论、方法和过程”。Gahegan(1999)在给地理计算下定义时,认为地理计算的核心是“用一系列的工具和方法来建模,分析一些复杂的、非确定性的问题来丰富地理学……从地理学和计算机科学来寻求更扎实的基础,是一个真正能动的定量地理分析计算,为计算机科学家提供关于计算和表达方法更丰富的资源。”这是从较实用的角度来分析的。Longley(1998)从其研究背景出发,认为对地理计算“研究者和参与者做的不多也不少”,这种说法有点中文的中庸之道。Conclelis(1998)给出了一个相对简单的定义,认为地理计算是“计算方法和技术的应用,用来描述空间属性,解释地理现象和解决地理问题”,并称地理计算是“计算技术应用到空间问题的总称”,其发展目的是广泛利用各种机会“采用计算方法为复杂地理问题求解”(Longley, et al, 1998)。地理计算不是简单的计算机在地理中应用,包含它但不限于它。真正地理计算所蕴涵的是利用前沿的计算技术解决地理相关问题。Macmillan(1998)也支持地理计算的创新性体现在现代科学与技术上。Longley(1998)指出地理计算的背景是GIS,是应用GIS的创新和探索,更强调“形式之上的过程、静态之上的动态、被动反应之上的交互”,是描述空间过程的计算工具和方法。Longley的观点的重要之处指出了地理计算在GIS之上,认为地理计算的精髓是“匹配了技术与环境、过程与数据模型、几何与应用、分析与背景、理论和实践”。丰富的GIS数据环境为现实世界的数字化提供了无限可能,地理计算的技术则提供了更好的结构化、分析和理解的方法,如果说GIS主要是关于数字地图信息的话,那么地理计算的重点是在不同的领域中应用它们。更明确地说,地理计算不是GIS的附加物,也不是GIS的一部分。

从地理计算的特点中可以看出,地理计算包含3个组成部分(Openshaw,1998):

- (1)地理或环境数据:对这些数据没有具体限制其数据形式和数据格式,也不一定要从GIS中来;

(2)现代计算技术:运用这些技术可以为涉及地理或环境数据的问题寻求答案,解决方法包括统计、数学、数值、智能等有关技术和它们的组合,这是解决地理问题的实际操作方案;

(3)高性能计算硬件:包括从小型机、工作站、多处理器工作站、矢量计算机(vector machine)到大规模并行处理计算机群。

本质上讲,地理计算的范围十分广泛,它关心计算科学在地理现象(人文和自然)中的应用,包括具有革新性的和前瞻性的问题,它十分广泛地借鉴了其他方法和技术,来处理几乎不能解决的、纷乱复杂的问题。目前理解的地理计算囊括各种以计算机为基础的模型和技术,其中大多模型和技术来自人工智能领域以及近年才出现的计算智能领域。这些模型和技术包括专家系统、仿生数学体系、后现代数学建模、模糊集合、虚拟现实与多媒体、探索性数据分析与数据挖掘,等等。因此,地理计算并不神秘,它是基于范式的计算依赖的科学,来广泛研究自然地理和人文地理中的各种问题,其最为突出的一点是强调基于科学方法来解决各种问题的计算。为更清楚地表达地理计算,Openshaw(2000)指出:“地理计算不是 GIS 的代名词,不是计量地理,不是极端的归纳,不缺乏理论,不缺少哲学思想,不是工具集。”

Ehlen 等(2002)对比了 5 次国际会议中对地理计算的描述,为了避免主观意识影响,他们用文字处理软件来处理。首先把每次会议论文的所有摘要做成一个文件,删除里面的括号、数字、公式、特殊字符以及引用等,使用完整的拼写,只有很常用的(如 GIS, www)才用缩写。统一采用英式英文拼写,然后交由美国马萨诸塞大学的计算机科学系的智能信息处理中心进行文字和短语解析。分析结果生成两个文件,一个是词语,另一个是段落,每个文件里都按照其(词或短语)出现的降序排列。对产生的结果还要进一步处理,包括删除仅出现在一个摘要中的词或短语、无意义的介词(*versa priori* 等)和短语(*paper describes*)以及短语处理等(Feng, et al, 2000; Ehlen, et al, 2000)。从众多学者的研究来看,对地理计算概念的探讨主要集中在 4 个问题上:一是是否集中于高性能计算,二是地理计算是不是工具集合(*grab bag*),三是 GIS 和地理计算的关系如何,四是地理计算的关键概念是什么。最终的结果并不像我们预期的那样,但是该分析结果给出了那些自认为是地理计算学者的人关于地理计算的研究方向或内容。

汇总 1996 年至 2000 年在地理计算中使用频率最高的 25 个英文单词(见表 1-1),从中可以看出,数据(*date*)、空间(*spatial*)、地理信息系统(*GIS*)出现频率最高,更出乎我们意料的是,结果中并不包含 Openshaw 倡导的“高性能计算”或“探索分析”等词。从分析的结果来看:①作为一个学科,地理计算还不完善,缺乏能被人们广泛接受的定义和内容框架;②地理计算借鉴了一些学科的子领域中的内容,因此并不是一个新的主题,而是一个交叉领域,是地学空间分析的扩展;③地理计算以地学问题或地理信息为研究背景,其科学理论的基础还不够完善。

地理计算使用一系列计算方法和工具,建立地理模型,并从中分析其复杂的、具有不确定性的地理问题,从而丰富了地理学的研究。地理计算回归计量革命时代的地理模型分析,吸收了新的计算机科学成果,如高性能计算机、模式识别、分类、预测与模型技术、知识挖掘、可视化等,其目标是:①将计算机工具引入地理学领域中;②设计合适的地理数据挖掘和知识发现操作;③研发时空尺度上集群算法;④获得超越目前软件、硬件能力的地理分析方法;⑤使用可视化和虚拟现实的手段实现人们对地理问题的理解与交流。