

大型矿集区识别与预测

——基于地理信息元组的分类预测理论与实践

陈 川 张晓帆 孙宝生 著



高等教育出版社
HIGHER EDUCATION PRESS

大型矿集区识别与预测

——基于地理信息元组的分类预测理论与实践

陈 川 张晓帆 孙宝生 著

封面设计：陈川 责任编辑：陈川 封面设计：陈川 责任编辑：陈川

ISBN 7-309-03383-4	定价：18.00元	印次：2008年2月第1次印刷	印数：5000册	印张：12	开本：787×960 1/16
北京鑫丰源印刷有限公司	北京鑫丰源印刷有限公司	北京鑫丰源印刷有限公司	北京鑫丰源印刷有限公司	北京鑫丰源印刷有限公司	北京鑫丰源印刷有限公司
北京市西城区西便门大街4号	北京市西城区西便门大街4号	北京市西城区西便门大街4号	北京市西城区西便门大街4号	北京市西城区西便门大街4号	北京市西城区西便门大街4号
100011	100011	100011	100011	100011	100011
010-38281000	010-38281000	010-38281000	010-38281000	010-38281000	010-38281000
010-38281118	010-38281118	010-38281118	010-38281118	010-38281118	010-38281118
http://www.jbpub.com.cn	http://www.jbpub.com.cn	http://www.jbpub.com.cn	http://www.jbpub.com.cn	http://www.jbpub.com.cn	http://www.jbpub.com.cn
010-38281118	010-38281118	010-38281118	010-38281118	010-38281118	010-38281118

高等教育出版社

材料号 33383-00

图书在版编目(CIP)数据

大型矿集区识别与预测——基于地理信息元组的分类预测理论与实践/陈川,张晓帆,孙宝生著. —北京:高等教育出版社,2008. 2

ISBN 978 - 7 - 04 - 023383 - 4

I. 大… II. ①陈…②张…③孙… III. 成矿预测 - 研究 - 中国 IV. P617. 2

中国版本图书馆 CIP 数据核字(2007)第203359 号

策划编辑 熊 威 责任编辑 徐丽萍 封面设计 王 晔 责任绘图 杜晓丹
版式设计 马敬茹 责任校对 张 颖 责任印制 韩 刚

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100011
总 机 010 - 58581000

经 销 蓝色畅想图书发行有限公司
印 刷 北京鑫丰华彩印有限公司

开 本 787 × 960 1/16
印 张 12
字 数 250 000
插 页 16

购书热线 010 - 58581118
免费咨询 800 - 810 - 0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landaco.com>
<http://www.landaco.com.cn>
畅想教育 <http://www.widedu.com>

版 次 2008 年 2 月第 1 版
印 次 2008 年 2 月第 1 次印刷
定 价 30.70 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 23383 - 00

本书由国家重要基础研究发展规划项目“中亚造山带大陆动力学过程与成矿作用”01 课题“古陆块解体与成矿物质堆积”(编号:2007CB411301)、08 课题“成矿预测体系研究与资源量潜力评估”(编号:2007CB411308)、《中国西部中亚型造山与成矿》09 课题“大型矿集区识别与预测”(编号:2001CB409809)联合资助。

序

本书是以矿集区为研究对象和目标的中尺度成矿预测研究。该研究是 20 世纪 90 年代后才进入实质阶段的。该书作者对中尺度矿集区预测进行了有益的积极探索,找到了宏观预测与微观找矿联系的结合点。作者在中亚型造山作用与成矿规律研究基础上,以中亚造山-成矿理论为指导,构建大型矿集区识别与预测体系,为中国西部大型矿床的寻找提供科学依据和技术支撑,作者在多方面进行了开拓性工作,主要有以下特点:

1. 作者通过对国内外各种成矿预测理论、技术与方法进行比较,试验研究,加之大量数据处理、运算,提出了以语义相似性度量为核心的具有创新意义的“基于地理信息元组的分类预测法”,取得了实质性和实用性进展,为现代成矿理论和成矿预测提供了新的思路与方法。

2. “基于地理信息元组的分类预测法”突出地理信息科学视角,以中亚造山-成矿理论为基础,集成地质、地球物理、地球化学和遥感多源信息,构建大型矿集区预测系统,建立了简单、可行、快速的成矿预测技术与方法,实现了预测方法的创新。

3. “基于地理信息元组的分类预测法”的相关理论、技术和方法体系建立对中亚大型矿集区的识别与预测、矿集区成矿背景,矿集区形成条件与矿床组合对比研究,提供了可量化的技术与方法。

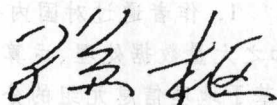
4. 作者通过建立大型矿集区概念空间,将所确定的大型矿集区概念空间形式化,以便于在计算环境下实现在该概念空间中各类别相似性的定量化度量,确定各性质维的度量值,并转化为大型矿集区语义参照系中的各数值维,组成语义参照系,对研究区进行分类或相似研究区之间进行对比研究,最后利用概率值对研究区进行形成大型矿集区有利度分级。作者机敏地将认知科学和信息科学引入并与计算机科学和地质学巧妙地结合,为现代成矿预测理论提供了新思路和新视角。

5. “基于地理信息元组的分类预测法”简单、快速、高效,可直接服务于中比例尺的找矿工作。该方法的提出,为困扰地学界的多源数据融合难题提供一种解决方案,究其原因在于作者研究的直接对象不是地理空间,而是体现地理信息空间微观本质的地理信息元组。这也是该预测方法的特色之一。

6. 作者首先从研究程度较高和矿产资源丰富的中亚地区研究入手,采用由

外向内推的方法,提出基于地理信息元组的分类预测模型,用数据驱动方式实现了中亚大型矿集区识别标志的建立,为研究程度较低的区域(新疆)进行矿集区识别与预测奠定了基础。

新疆幅员辽阔,矿产资源潜力巨大,是我国 21 世纪重要的资源接替区,将成为我国经济增长的重要增长点。为加速新疆矿产资源开发,必须突破常规找矿思路、理论和方法手段,大型矿集区的识别与预测研究就是为我国西部资源勘察寻找新思路、新理论、新方法和新手段,科学评估我国中亚成矿域的矿产资源潜力。“基于地理信息元组的分类预测法”,首次提出从地理信息科学的视角进行矿集区预测。我相信,本书的出版将对进一步发展矿产资源预测和评价的理论与方法,促进新疆地质找矿工作,尽快将西部建设成为我国可持续发展的重要矿产基地发挥重要作用。



2007 年 11 月

目 录

第 1 章 绪言	1
1.1 地理信息科学历史及现状:从 GISystems 到 GIScience	2
1.2 矿产资源评价与预测理论及方法的历史与现状	8
1.3 存在问题分析	11
1.4 分类预测理论的提出	12
1.5 研究目标及研究内容	14
第 2 章 基于地理信息元组的分类预测理论	18
2.1 基于地理信息元组的分类预测理论	18
2.2 分类预测理论框架的内涵	49
2.3 关于分类预测理论的几点说明	50
第 3 章 基于地理信息元组的分类预测模型	53
3.1 地理信息空间的作用及其构建	54
3.2 简单域的构建	55
3.3 合成域有关空间分析	63
3.4 探索式空间分析(ESDA)与矿集区预测	69
第 4 章 分类预测理论应用:以中亚 Au、Cu 矿集区预测为例	74
4.1 概述	74
4.2 研究区位置及基本地质特征	74
4.3 中亚地区含矿有利区域分类预测	80
4.4 中亚地区大型矿集区识别与预测	115
第 5 章 中亚地区矿集区识别	128
5.1 中亚地区与构造有关识别标志分析	128
5.2 中亚地区与沉积岩有关识别标志分析	144
5.3 中亚地区与侵入岩有关识别标志分析	148
5.4 中亚地区与火山岩有关识别标志分析	154
5.5 中亚地区与航磁、重力有关识别标志分析	162
第 6 章 结论	173
6.1 研究中的主要进展	173
6.2 研究中存在的主要问题	174
参考文献	175
致谢	184

第 1 章 绪 言

随着经济的发展,矿产资源越来越体现出其在全球经济与国家安全方面的重要作用。矿产资源提供了我国 90% 以上的能源和 80% 左右的工业原料。矿产资源开发利用支撑了 300 多个矿业城市的生存发展,形成了 10 300 个大型国有矿山企业,解决了 2 100 万人口就业。因此,矿产资源的持续稳定供应和大型资源基地的快速高效发现,是保证经济持续快速增长、保障国家经济安全的重要环节。在今后相当长的时期内,矿产资源仍将是我国国民经济和社会发展所依赖的重要物质基础,并将继续在国民经济发展中发挥举足轻重的作用。

然而,我国矿产资源和矿业发展形势十分严峻,主要表现为:① 矿产资源零储备;② 资源保证程度偏低;③ 矿产资源品质较差;④ 矿山资源日渐枯竭,直接威胁着国家的经济安全和社会稳定。解决上述问题的根本,必须依赖于科技进步,特别是重大成矿理论与关键技术的突破与应用,整体提升我国矿产资源勘查评价水平和科学预测能力。通过理论创新与高新技术的开发应用,快速客观地评价我国矿产资源的潜力,并据此建立矿产资源安全保障与预警系统;快速高效地发现新的大型后备矿产资源基地,增加资源补给;有效地找寻枯竭矿山的接替资源,确保矿产安全供给。

新疆是我国战略资源的重要储备区,包括新疆和邻区在内的中亚成矿域是全球矿产资源潜力最大的区域,横贯新疆的第二亚欧大陆桥是一条安全的陆路资源通道。这一区域以资源供给的巨大潜力与资源通道的特殊功能,构成我国资源安全体系的举足轻重的一翼,关系到我国的经济、国家安全和在国际格局中的核心利益。

同时,中亚丰富的能源资源和固体矿产资源已成为许多国家政府和世界各国地质学家关注的热点,与我国新疆接壤的中亚地区是世界上少数几个矿产资源密集区,拥有巨量的各种矿产资源。在世界上其他地区(北美洲和南美洲地区)矿产资源大规模开发之后,中亚地区已经成为国际矿业集团(包括各国政府)争夺的最后一块宝地。中国西部及其中亚邻国(蒙古、俄罗斯、哈萨克斯坦、吉尔吉斯斯坦、乌兹别克斯坦等)同属中亚型造山域^[1],不仅地质构造和演化历史相似,其成矿时代、类型乃至具体的成矿环境(相)都可以对比。

然而,由于社会经济发展的原因,总体上,邻国的地质矿产勘探和开发程度较高,已经查明了若干世界知名的大型矿集区。按储量和资源量看,境外矿床是

我国境内同类矿床的十倍至几十倍。开展境内矿产资源相关对比研究势在必行。

20世纪80年代后期至今,随着国际地质对比计划和各类多国编图计划的实施,境内外基础地质的对比研究有所开展;与此同时,国家305项目在“七五”期间曾开展“中国新疆周边国家矿产地质特征及成矿规律”的专题调研;地质矿产部在“七五”科技发展规划中立项,开展中国周边国家毗邻地区成矿区(带)地质矿产情报调研^[2]。此后,国土资源部信息中心和地质调查局于2001年完成了“中国西部和毗邻国家铜金找矿潜力的对比研究”项目^[3]。这些大规模的对比研究以及之后零散的一些相关研究成果^[1,4],为中国西部及其邻区的地质矿产对比研究积累了资料,为进一步开展以在我国找矿为主要目标的工作奠定了良好的资料基础^[5]。然而,已有的工作都属于调研性质,其系统性和深入程度都远不能满足当前对急于了解境内外能源、资源情况的需要^[6]。

国外在中亚区域的研究也积累了大量的数据和资料,这些不同时间、不同来源、不同内涵的数据和资料在以多元信息综合分析为主要手段的矿集区预测研究中的整合应用就形成了矿集区预测的主要内容,也决定了多元信息矿集区预测研究中地理信息科学与矿产资源评价与预测理论结合的必然性。两者的发展现状也体现出这种结合的趋势。

1.1 地理信息科学历史及现状: 从 GISystems 到 GIScience

对地理信息本质的研究,推动了地理信息系统(GISystems)到地理信息科学(GIScience)的发展^[7],将地理信息科学归入信息科学中,从而导致了从地理信息科学角度对与地理信息应用相关的地学领域进行重新认识成为一种必然趋势。在从系统到科学的转变过程中所围绕的核心问题包含以下几个方面:

- ① 什么是地理信息?
- ② 与地理信息相关的基本原理是什么?
- ③ 地理信息如何表达?
- ④ 地理信息怎样表达客观世界?
- ⑤ 地理信息信息量如何进行量化衡量?
- ⑥ 地理信息如何传递?
- ⑦ 与地理信息相关的人机交互的根本问题是什么?

从以上问题可以看出,对于地理信息本质的认识是当前地理信息科学的根本和关键,随着资源环境科学、信息科学、地球科学等学科的交叉融合,高

本问题,Goodchild 提出了地理信息的元组本质^[8],从而奠定了地理信息科学的研究基础。

1992 年 Goodchild 提出了 Geographic Information Science 作为一门新的科学而不仅仅是一种工具的观点^[9];1996 年著名国际性学术期刊 Geographic Information Systems 更名为 Geographic Information Science;在随后的近 10 年来,GIScience(即 Geographic Information Science)作为一门新的学科已经成为研究领域关注的焦点,并以各种方式被理解着,但是对于其含义、与地理信息技术的关系以及如何推动它前进等方面依然保持着不确定性。减少这些不确定性的决定性因素在于对地理信息(GI)的本质认识,即从 GISystems(即 geographic information systems)中对系统、句法、数据结构(主要包括程序语言、算法)的工具性研究转移到以认知科学和信息科学为基础的对地理信息的语义本质进行重新认识及研究上来。从 UCGIS、NCGIA 等权威研究机构所提出的优先研究领域中都反映出这一趋势,也正是这一趋势推动了从强调图形系统(systems)与算法为主要研究目标的工具时代转向以重新认识地理信息(GI)本质为主要研究内容的科学理论研究时代^[10-11],逐步实现从技术到科学的过渡。

关于地理信息科学的概念,Goodchild 在提出时并没有给出这一概念的定义,但给出了地理信息科学的研究领域及研究议程。1995 年 UCGIS(University Consortium for Geographic Information Science)正式命名成立,给地理信息科学一个科学完整的定义的使命就或多或少地交给了 UCGIS 这一研究组织。但是同 Goodchild 一样,UCGIS 依然没有给出清晰完整的定义,取而代之的是在 2002 年对该组织的研究内容给出了确切地描述:“UCGIS 致力于通过发展和应用理论、方法、技术、数据去理解地理过程、关系以及模式。地理信息科学的核心就是实现地理数据向有用信息的转化”^[10],从这一迂回的定义中我们能够发现地理信息系统与地理信息科学的区别:地理信息科学强调“理解”和“新信息(有用信息)转化”。从这一角度去理解,认知科学和信息科学无疑成为地理信息科学的理论基础,从而使地理信息系统的研究摆脱了没有理论基础而只能沦为工具的境地。在 1999 年美国国家自然科学基金会中所给出的地理信息科学的定义中也包含着类似的理解:“地理信息科学是一个基础研究领域,它探索地理概念的重新定义和这些重新定义的地理概念在地理信息系统背景下的使用;地理信息科学同时研究地理信息系统对个人及社会的影响以及社会对地理信息系统的反作用;地理信息科学在与认知科学和信息科学的整合过程中需要对一些传统的面向空间领域的学科中一些最基本的主题进行重新研究,诸如地理、制图学、地质等;地理信息科学同时也包含和汲取了一部分特殊研究领域的内容并对这些领域的发展作出贡献,诸如计算机科学、统计学、数学以及心理学等。地理信息科学支持政治学和人类学研究并将这些领域纳入地理信息与社会这一研究领域

之中。”^[12]。对于传统的面向空间领域的学科,地理信息科学这一定义无疑给这些学科中的一些最基本的议题的重新认识提供了新视角,为新的理论体系的研究提供了理论依据,本书正是从这一新视角对分类预测模型中的语义相似性问题进行了研究和探讨。

对于分类预测模型研究现状的理解,需要追溯地理信息科学的发展历史,这里所说的地理信息科学发展史本身不包含地理信息系统发展史,地理信息系统发展史本身就是一个重要的极具研究价值的研究内容,对此, Coppock 和 Rhind^[13] 以及 Foresman^[14] 都作了精辟的论述。在本书中对地理信息科学的发展史的讨论从地理信息科学出现开始。虽然在 20 世纪 60 年代甚至更早些时候地理信息系统和计算机制图学就遇上了大量概念与算法问题方面的挑战,但是关于地理信息系统软件背后可能存在着理论研究领域的思想直到 80 年代才被提出。地理信息科学发展的关键事件是在 Alber 倡议下“NCGIA”研究组织的建立,并提出了著名的“五点研究内容”^[15]:

- ① 空间分析与空间统计;
- ② 空间关系与数据库结构;
- ③ 人工智能和专家系统;
- ④ 可视化;
- ⑤ 社会、经济以及公共机构问题。

从历史的观点来看,这五点研究内容就是我们现在称为地理信息科学在作为一个新的研究领域浮现出来时的其研究范围的最初定义。

Goodchild 在上述基础上 1990 年提出了 Spatial Information Science 的概念^[16],1992 年将这一概念修改为 Geographic Information Science^[9],并在 International Journal of Geographical Information Science (IJGIS)上发表,在该文中提出了八点地理信息科学的研究内容:

- ① 数据收集和度量;
- ② 数据捕获;
- ③ 空间分析;
- ④ 数据模型和空间数据理论;
- ⑤ 数据结构、算法和处理;
- ⑥ 数据显示;
- ⑦ 分析工具;
- ⑧ 公共社团、管理以及种族问题。

上述概念和研究内容的提出是地理信息科学发展史上一个重要的里程碑。

20 世纪 80 年代期间,很多 GIS 团体的分支机构也加入到这一潮流中,1992—1993 年间一系列讨论导致 1994 年 UCGIS 成立会议的召开。1996 年

UCGIS的代表会议确定了该组织的优先研究议程,这一议程包括 10 个方面内容:

- ① 空间数据获取和整合;
- ② 分布式计算;
- ③ 地理表达扩展;
- ④ 地理信息认知;
- ⑤ 地理信息互操作;
- ⑥ 比例尺;
- ⑦ GIS 环境下的空间分析;
- ⑧ 空间信息基础平台的未来;
- ⑨ 空间数据与基于 GIS 分析的不确定性;
- ⑩ GIS 与社会。

几乎同时,NCGIA 的研究人员向美国国家自然科学基金提交了题为“Advancing Geographic Information Science”的研究申请,并获资助,即 Varenius 项目^[8]。在该项目中第一次清晰地提出了地理信息的元组本质 $\langle x, y, z, t, U \rangle$,成为地理信息科学发展史中又一重要里程碑。研究申请从新的视角提出,GIScience作为一个研究领域,应当以以下三个方面的研究为基础:

- ① 地理空间的认知模型(分类理论);
- ② 用于表达地理概念的计算方法(包括互操作、领域本体表达、地理知识发现和数据挖掘);
- ③ 信息社会中的地理。

分析上述不同研究机构、个人以及不同时期所提出的研究内容,地理信息科学的发展,从 NCGIA 的 5 点内容到 Goodchild 8 个方面的研究内容,再到 UCGIS 的 10 项优先研究内容,最后到 Varenius 项目^[8],地理信息科学最终将认知科学和信息科学确定为其基础理论,试图使地理信息在应用中摆脱缺乏理论基础的尴尬境地。也正因为如此,在地理信息应用中应当从地理信息科学的基础理论出发,对地理信息所应用领域进行以认知科学和信息科学为基础的重审视,而不局限于地理信息系统所限定的辅助工具性应用,应当着眼于地理信息本质而不再是系统与几何图形算法的研究。

Varenius 项目三方面的内容在应用过程中往往是互相联系相辅相成的,本书主要涉及地理空间的认知模型以及地理概念表达的计算方法两个方面的内容。

自 2000 年 Varenius 项目结束一直到目前,正如 Varenius 项目的标题所述,对上述内容的研究获得了长足进展,众多研究者都在以其不懈的努力对地理信息科学各个方面的发展作出卓越的贡献。如 D. Mark 和 B. Smith 致力于从哲

学角度进行地理空间认知研究^[17],并在1999年提出题为“Cognitive Models of Geographical Space”的UCGIS的优先研究内容^[18],其后续研究内容在2002年UCGIS优先研究内容规划中依然作为优先重点研究内容。此外,他们共同主持着UCGIS的另外一项重要内容——地理空间本体研究^[17-19],D. Mark研究内容反映出将认知科学和信息科学作为地理信息科学理论基础的思路。与此同时Goodchild继主持Varenius项目之后致力于UCGIS另一项优先研究内容——地理信息科学与社会的相互作用。Egenhofer在20世纪80至90年代早期主要从事地理信息系统算法与语言方面的研究,在1995年发表了开创性论文Naive Geography^[20],标志着地理信息系统向地理信息科学的转变,其文中提出的地理空间概念也表明了认知科学作为地理信息科学的理论基础必然与地理信息科学交叉整合的趋势,也因此将其研究重心转移到了地理空间中地理概念表达上。而T. Bittner和C. Freksa则对空间概念和地理对象之间的关系进行了系统地研究,并对语义粒度的概念给出了清晰地描述^[21-23]。Montelo在20世纪80—90年代中后期一直从事与地图有关的认知研究,这些研究着重于对符号的视觉变量的研究,90年代后期转向地理信息的认知研究,他的研究成果在Cognition of Geographic Information一文中具有较为系统的表述。A. Frank强调地理信息的可用性,而地理信息的可用性又由用户认知与地理信息的链接所决定^[24-25],阐明了地理信息的认知本质,他的另一个研究重点与D. Mark的地理空间本体相对应,为地理信息本体研究,并提出了地理信息形成过程中的五层体系,以此为基础,他在专著Ontology for GIS手稿中系统地阐述了地理信息科学的理论基础:从认知到本体到地理信息形成的原理和过程,第一次系统地将地理信息科学、认知科学与信息科学有机地整合起来。在众多研究者中,P. Gardenfors是比较特殊的一个,他的主要研究领域是认知科学中的语义表达模型^[26-27],但他提出的概念空间的概念在地理信息系统呼唤语义内涵时对地理信息科学具有深远的影响,概念空间弥补了两种语义表达模型的缺陷,为知识表达提供了合理的框架。受P. Gardenfors的启发,W. Kuhn提出了“semantic reference systems”的概念^[28],从应用角度提供了一条理论思路;在同一思路引导下,M. Raubal提出了概念空间形式化的设想^[29]。J. Mennis对地理信息系统应用中将认知作为数据库表达的基础,为地学信息库(知识库)中的知识发现和数据挖掘提供了思路^[30]。F. Fonseca在其博士论文中^[31]的本体驱动地理信息系统的概念,阐明了地理信息系统应用的核心应该是本体而不是本体的视觉表现(图形),提出了以认知科学和信息科学为基础的地理信息系统应用框架^[32]。Moran、Geary、Cliff、Ord、Anseline、Agterberg、Bonham-Carter等人的研究成果则在空间统计学为传统空间分析及探索式空间分析奠定了方法论基础。

从上述理论研究背景与现状分析以及其他相关文献中,不难理出地理信息

科学发展的理论思路 and 方向:从以几何表达为核心的图形系统发展到将认知科学的内容包含进该系统中,被认知的地理信息必须经过形式化后才能在计算环境中表达出来,并且所表达的包含认知语义的地理信息必须能够有效地传输^[33],这就进一步将信息科学整合进地理信息体系中。认知科学和信息科学以及地理信息科学中原有的数理理论共同形成了地理信息科学的理论基础,使地理信息系统完成了从辅助性应用工具到科学的转变。转变后的地理信息科学将地理信息系统作为它的应用工具环节。但不同于以往的是,当应用地理信息系统时应当基于地理信息科学的理论基础,即认知科学和信息科学中的有关理论,从地理信息的元组本质出发,以其认知本质而不是视觉变量为核心,这要求将地理信息分类过程中隐式地在人脑中起作用的概念(语义)或概念域在计算环境中显式地表达出来,从概念空间到概念空间形式化再到语义参照系正好为建立地理信息科学基础理论的应用提供了思路。沿着这一思路势必建立起基于地理信息科学基础理论的应用理论体系,原因如下:①当前的地理信息软件已经将认知科学和信息科学作为其基础内容包含于地学信息库构建模块中^[34-35];②以认知科学和信息科学为基础构建的地理信息空间是以概念层而不是符号层为核心的知识库(NADM1.0)^[36],这就为通过知识发现和数据挖掘方式确定概念空间的性质维提供了基础;③通过数理统计方式能够对所确定的性质维形式化^[29,37];④形式化后的概念空间能够成为语义参照系^[28],在特定的语义参照系下就能够对所研究区域进行特定的分类,而分类恰恰是大部分地学信息研究的最终目的。

综上所述,地理信息科学经过近30年的发展,已经奠定了一定的理论基础,其中最主要的进展就是以认知科学和信息科学作为其基础理论的核心内容。因此,将认知科学研究领域中的成果与地理信息科学整合是地理信息科学基础理论,同时也是应用理论体系的未来发展趋势。

在理论上,地理空间和地理信息空间的本质认识以及地理信息元组概念的提出^[7],对地理信息应用特别是在地质领域的应用理论体系的建立提供了一条理论依据和入口,从而为基于语义的地理信息应用确立了方向;认知科学领域的重要研究成果“概念空间”的提出以及哲学和认知科学领域分类理论的发展和完善,为基于语义的地理信息在以地理信息元组的方式得以应用提供了应用背景的层次框架;基于语义的驱动地理信息系统以及各种基于本体数据库表达模型的研究和建立为地理信息应用奠定了语义信息应用的数据表达基础;语义参照系的提出是认知科学与地理信息科学整合的典范,为认知科学理论成果在地理信息系计算环境下的实现提供了方法论依据。

在技术上,以 ArcGIS 为代表的新一代地理信息系统的日益完善;在地理信息表达上,以本体为核心的地理信息表达方式为基于语义的地理信息的表达及

应用提供了强有力的工具,使得原有地理信息所不能完成的知识发现、复杂环境建模等复杂应用在新地理信息系统下成为现实;在地理信息分析技术上,ArcGIS从地理信息库(知识库)、基于知识库的智能可视化,以及地理信息处理三个角度为地理信息的各种应用提供了强有力的工具支持^[35],特别是9.0版本对探索式空间数据分析方法整合使从海量日益复杂的地理信息中进行数据挖掘和知识发现可以在空间、时间、语义一体化方式下进行。

1.2 矿产资源评价与预测理论及方法的历史与现状

关于矿产资源评价与预测的历史发展有过这样的概括:“矿产勘查过程从19世纪到现在,经历了‘找矿人’勘探、传统找矿、理论勘查及目前以高新技术为代表的科学找矿阶段,而矿产资源评价与预测实质上贯穿了矿产勘查的全过程”^[38],对应于矿产勘查的发展,矿产资源评价与预测经历了从以地表露头为简单标志的单个矿床的预测方法发展到以成矿规律研究为代表对成因上有联系的系列矿床的理论找矿阶段,随着各种基础地质工作的深入以及数据获取手段的迅猛发展,导致了地质领域海量数据的迅速堆积,从而形成了与理论找矿并行的以数据驱动为特点,以较大区域矿产资源潜力(矿集区)预测为目的矿产资源评价与预测的理论与方法。从本质上看,理论找矿的直接研究对象是客观世界,而数据驱动理论与方法面对的则是信息空间,因而是两条并行的截然不同的研究方向。

理论找矿的发展得益于法国学者郎内提出的“成矿规律”的概念,并在20世纪五、六十年代得到空前发展,以毕利宾的“构造-建造”成矿预测分析方法、斯米尔诺夫的“矿床分带性”理论为代表。70年代以来,随着航天、航空、遥感技术以及各种地球物理、地球化学等找矿技术的日益广泛、成熟地应用,矿产资源评价与预测得到了空前的发展,人们开始面对大区域海量的数据,从而激发了国家从宏观决策战略对于大区域矿产资源潜力评价与预测的需求,此时的矿产资源评价与预测在区域范围上已经有了急速扩展的势头。如美国自20世纪80年代以来就开展了全国近 $5 \times 10^6 \text{ km}^2$ 的金属矿产评价工作;我国系统地开展了两轮区划工作。由于数据的迅速积累,将计算机手段引入矿产资源评价与预测也成为一种必然,从而产生了各种定量评价的方法体系。“矿产资源定量评价是以计算机信息处理技术为工具,研究各种勘查信息中所暗含的成(含)矿信息,特别是通过定量方法,研究各种多源信息与矿床资源潜力的关系模型,达到对未知区的定位、定量评价”^[3]。自五六十年代以来,阿莱斯、哈里斯、格里菲斯、康斯坦丁诺夫等人先后进行了单变量资源评价方法的探索工作,至70年代末,国

际地质科学联合会第 98 项计划推出了 6 种标准的矿产资源定量评价方法,这标志着矿产资源定量评价进入实用阶段。在这方面,一些代表性的成果有: F. P. Agtergerg、G. F. Bonham-carter、C. f. Chung 以及 D. F. Wright 于 1983 年^[39]和 1986 年^[40]提出成矿预测中的地质数据集的整合方法,并于 1989 年提出对现代数据驱动成矿预测模型起着奠基作用的证据权模型^[41-42],在此基础上又发展了加权 Logistic 回归模型^[43]; A. N. Campbell 以及 V. F. Holloster 于 1982 年提出的人工智能矿床识别程序^[44]; P. A. Burrough 于 1986 年定义的资源评价的 GIS 原则^[45]; 赵鹏大院士提出的以“求异理论”^[47]及“三联式”预测方法^[46]为代表的科学找矿评价理论方法体系^[47]; 王世称教授创立的综合信息矿产预测方法^[48]等。

随着 GIS 技术的发展,各种空间分析方法被应用于区域性的矿产资源评价与预测中,即对与成矿有关的各种地学空间信息进行综合分析,将由多个图层表示的成矿信息综合成一个预测图层,达到区域矿产资源评价与预测的目的。矿产资源评价与预测的重要目标之一就是发现新的矿床或者含矿潜力区,矿产资源评价与预测即包含描述模型,同时必须包含预测模型。在地学中,依据模型所表达的关系可将模型分为:知识驱动(理论)模型、数据驱动(经验)模型、混合模型^[41]。无论是何种模型,在矿产资源评价与预测中其主要目的都是量化地表示相关的专题空间关系,最终对若干个专题空间关系进行综合分析生成预测图。多数以 GIS 为基础的矿产预测重点研究对多个专题关系进行综合分析的方法,这些方法主要包括布尔逻辑、索引叠加、模糊逻辑、贝叶斯模型(证据权模型)、加权 Logistic 回归和神经网络等^[49-50],其中前三者属于知识驱动模型,后三者属于数据驱动模型,各种方法特点见表 1-1。

表 1-1 以空间分析为基础的多源数据整合方法概览

方法	主要特点
布尔逻辑(二值)	将各来源的要素类或专题图表示成为只包含有利(1)或不利(0)的二值图像;各二值图像不加区分地同等对待;通过简单的与操作,或者直接判断区域有利(1)或不利(0) 缺点:各要素类同等对待,判断过于简单化 适用性:只适用于极简单的判断
索引叠加 (二值要素类)	各要素类二值化后赋予不同的权值,在此基础上进行线性组合 缺点:经验权值主观随意性较大,输出结果常因人而异。并且规范所规定权值往往跟不上区域及应用背景的变化 适用性:经验和理论较丰富的专家,并且对所研究区域情况很熟悉

续表

方法	主要特点
索引叠加 (多分类要素类)	赋予各要素类不同权值的同时也赋予各要素类中各分类不同的权值,在此基础上线性组合 缺点:多权值的变化使得权值随意性成几何级数增长,不同区域及应用背景权值往往与规范有一定出入 适用性:权值的赋予判断往往在一定程度上超出研究者的理论和经验范畴,因此适用性受到一定局限
模糊逻辑	更灵活的要素类及要素类中分类的权值赋予方法,在此基础上进行非线性组合 缺点:主观随意性较大,对理论知识体系及应用背景的要求较高 适用性:往往和数据驱动方法结合,以弥补对理论知识体系和应用背景要求较高的缺陷,如扩展证据权模型
证据权法 (贝叶斯模型)	二值要素类或专题图通过对数现行组合;通过证据层权值计算和贝叶斯法则进行含矿有利判断并据此对研究区域进行分类;模型参数容易理解 缺点:贝叶斯概率要求的各证据层条件独立与地学第一原理相违背和证据层数据缺失造成后概率不稳定,造成不可信输出结果;对数据的完整性和可靠性要求较高 适用性:具备完整、可靠的多源数据,条件独立检查,对输出结果的精度要求不高
加权 logistic 回归	二值要素类或专题图通过 log 回归结合,通过对训练集判断区域的含矿有利程度并进行分类;具有较复杂的模型参数 缺点:对数据的可靠性完整性要求较高 适用性:适用于具有可靠完整数据的应用
神经网络	人工神经网络(ANN)并不用于将上述定义的空间关系综合成一幅预测图,而是进行分类的工具。建立一个包括矿床大小、距最近断层的距离、走向、局部接触密度、最近的接触类型、基岩等内容的数据库。这些信息与对已知是贫瘠地区的相似度量信息一起由 ANN 进行处理,建立输入(度量信息)与输出(矿床大小)的关联。一旦经过训练,ANN 就可以用于确定研究区任何一点的有利程度。现在的问题是开发用于处理从图件的输入到合理输出的全过程的 ANN 缺点:对输入数据的完整性和可靠性有较高要求 适用性:适用于具备完整、可靠的多源信息环境