

基于 GIS、模糊逻辑和专家知识的土壤制图 及其在中国应用前景*

朱阿兴^{1, 2} 李宝林¹ 杨琳¹ 裴韬¹ 秦承志¹ 张甘霖³ 蔡强国¹ 周成虎¹

(1 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, 北京 100101)

(2 Department of Geography, University of Wisconsin-Madison, Madison, WI 53706, USA)

(3 中国科学院南京土壤研究所, 南京 210008)

摘要 详细的土壤空间与属性的信息已成为环境模型和土地管理的基本参数, 传统的以类别多边形和手工编制为基础的传统土壤制图效率低精度也较差。本文基于 GIS、模糊逻辑和专家知识, 建立了土壤—环境推理模型(SoLIM), 通过基于土壤—环境关系模型的土壤相似度模型与对该模型进行赋值的推理技术来编制土壤图, 从而克服了传统土壤制图中的简化。通过两个小区的研究表明, 与传统土壤制图相比, 通过 SoLIM 得出的土壤信息在空间详细度和属性精确度都有较大的提高, 也能够大量减少调查的时间和经费, 从而大大提高土壤调查的效率。SoLIM 方法在我国推广十分必要且具有一定的条件, 但仍需要进一步完善

关键词 土壤—环境推理模型; GIS; 模糊逻辑; 专家知识; 土壤制图

中图分类号 P934 **文献标识码** A

土壤信息是土地管理和环境模型的基本参数, 随着现代精准农业的发展和环境模型模拟精度的提高, 对土壤的空间和属性信息的详细程度都提出了更高的要求^[1, 2]。传统的土壤调查首先通过野外调查建立区域土壤与环境关系模型, 然后通过手工根据航片或地形图等将不同的土壤或土壤组合绘制在空间范围上, 如美国的传统土壤普查技术^[3, 4], 我国两次土壤普查也是采用这种方法^[5-7]。传统土壤调查的准确性和效率主要受两个因素限制: 基于多边形的图形表达和手工完成的制图过程。基于多边形的土壤图常使小的土壤斑块在制图综合时常被综合掉, 从而产生土壤图在空间上的简化; 同时由于不能表达土壤性状的渐变和连续, 从而产生属性上的简化。另外, 以大量野外调查和手工制图为基础的传统土壤普查周期很长, 例如, 美国的 1:24 000 土壤图的更新周期大约为 100 a, 精度却只有 50% ~ 60%。由此可见, 常规土壤调查已不能满足信息时代环境模型和土地管理模型所需的详细土壤信息。近些年信息技术的发展为土壤制图提供了新的技术与数据支持, 利用 GIS 来生产土壤类型图和土壤属

性图的数字土壤制图已成为土壤调查研究的热点问题^[8]。早期主要是通过 GIS 空间分析来确定景观类型, 土壤类型与景观类型之间的关系仍以定性的专家知识为主^[9, 10]。后来, 数字土壤制图侧重建立土壤类型或土壤属性与环境要素之间的定量关系, 到目前为止, 获取它们之间关系出现了很多的方法, 如线性回归^[11]、分类法^[12, 13]、回归决策树法^[14-16]、神经网络法^[17-19]和地质统计学^[20-22]等方法。土壤的形成是多要素制约的非线性过程, 而以回归为基础的线性回归、回归决策树方法和神经网络等方法以土壤和环境关系在空间上是固定的这一假设为基础, 在实际中的应用受到较大的限制。另外, 土壤本身非常大的变异性也不能满足地质统计学对数据稳定性的要求, 也使这类方法的应用受到制约。目前这些方法大部分作为研究探讨, 都没有作为土壤普查实用技术为政府部门所采纳。本文通过 GIS 和遥感技术获得土壤形成的环境因子, 根据土壤专家或者野外调查来提取土壤和环境之间的关系, 构造基于模糊逻辑的推理模型——土壤—环境推理模型(Soil-Land Inference Model, SoLIM), 然后推导出土

* 美国农业部和威斯康星大学、中国科学院“百人计划”项目、国家自然科学基金项目(40101028)和中国科学院地理科学与资源研究所知识创新项目(CX10G-D02-02)支持

作者简介: 朱阿兴(1962~), 男, 博士, 研究员, 主要从事地理信息系统、人工智能、模糊数学理论集成及其应用研究, 发表学术论文 60 余篇, 其中第一作者 SCI 论文 9 篇。E-mail: Axing@lreis.ac.cn

收稿日期: 2004-12-30; 收到修改稿日期: 2005-04-18

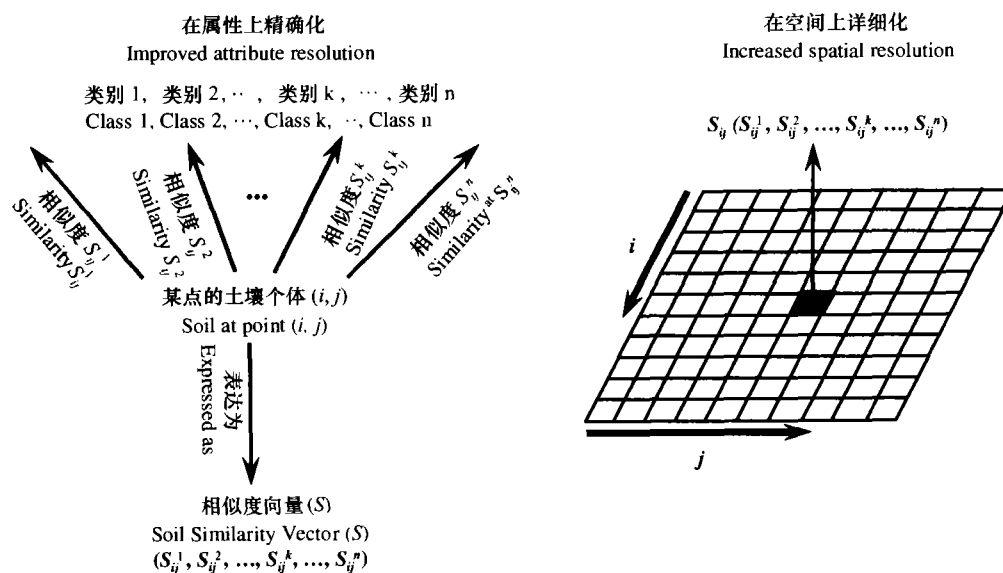
壤的空间分布,从而克服传统土壤调查的局限性,使土壤普查信息在空间详细度和属性精确度等都有较大提高,目前美国农业部已将这种方法作为土壤调查的标准技术进行推广。

1 土壤——环境推理模型

1.1 土壤类型的连续表达——土壤相似度模型

土壤相似度模型在空间上采用栅格表示,在属性上则采用相似度来表达。空间上的栅格表示可使土壤在地理空间上的简化大大降低,其详细度可通过空间分辨率变化来实现。属性上的相似度表达则

是基于模糊逻辑^[23],每个像元的土壤根据其与各典型土壤类型的相似程度分配给多个土壤类型,这些分配的比例称为隶属度,这样所有的隶属度就形成一个 n 维向量(土壤相似度向量,或称为隶属度向量) $S_{ij}(S_{ij}^1, S_{ij}^2, \dots, S_{ij}^k, \dots, S_{ij}^n)$, n 是给定的土壤类型的个数,第 k 个元素 S_{ij}^k 代表在像元 (i, j) 的土壤和土壤类型 k 的相似度。这样某一地区的土壤可以表示成为一组像元,每一像元的土壤属性被表达成相似度向量(图 1),由此土壤空间就可以被表达成空间上和属性上都连续的面,在很大程度上克服了土壤在地理空间和属性空间上的简化问题。



(空间上表达为像元,属性上表达为相似度向量)

(Soil bodies are presented as pixels in spatial domain and as similarity vectors in parameter domain)

图 1 土壤相似度模型

Fig. 1 Soil similarity model

1.2 土壤相似度模型的赋值——模糊逻辑下的土壤自动推理

土壤相似度向量中各隶属度值主要是根据土壤因子方程^[24, 25]和土壤环境模型来确定^[4],采用某一土壤类型典型的成土环境与某一点当地特定的成土环境之间的相似度 S' 来近似地代替该地土壤与这种典型土壤类型的相似程度:

$$S' = \int f_1(E) dt \quad (1)$$

式(1)中, t 代表时间; f_1 代表土壤与形成环境之间的关系; E 为描述气候、地形、母质和植被等环境条件的变量。在实践中具体描述 t 因子是非常困难

的,并且 t 中的信息通常在其他成土环境因子(如地形位置)或当地土壤专家的知识中有所表达,因此本模型中式(1)可简化为:

$$S' = f(E) \quad (2)$$

在实际工作中,土壤相似度模型赋值,首先要建立土壤环境关系 f ,土壤和环境关系可以通过专家知识、神经网络计算、案例推理和数据挖掘等方法来获得^[26~29],本文主要采用专家知识的方法;然后将获取的土壤环境关系与描述土壤成土环境因子相结合,环境因子主要通过 GIS 技术来获得,由于数据的可获取性和各地区成土过程的不同,不同地区包括的环境变量并不确定;最后通过模糊推理得到

S' ^[30,31]。现在我们将获取的土壤环境关系储存在知识库中,土壤成土环境数据存储在GIS数据库中,通过一系列的模糊逻辑推理技术(模糊推理机)连接知识库和GIS数据库,从而自动推导出土壤相似向量(图2)。一般而言,对于像元(i, j),推理机从GIS数据库提取土壤成土环境因子的数据,然后将GIS数

据和从知识库中获取土壤类型 k 的土壤环境关系结合,计算当地环境和土壤类型 k 的典型环境之间的相似度 S'_{ij}^k 来代替 S_{ij}^k 。当该点环境条件与所有典型土壤类型的环境条件相似程度计算完后,这一像元的土壤相似向量 S_{ij} 就得到了。当所有的像元都计算完毕,就完成了整个区域土壤相似度的推导工作。

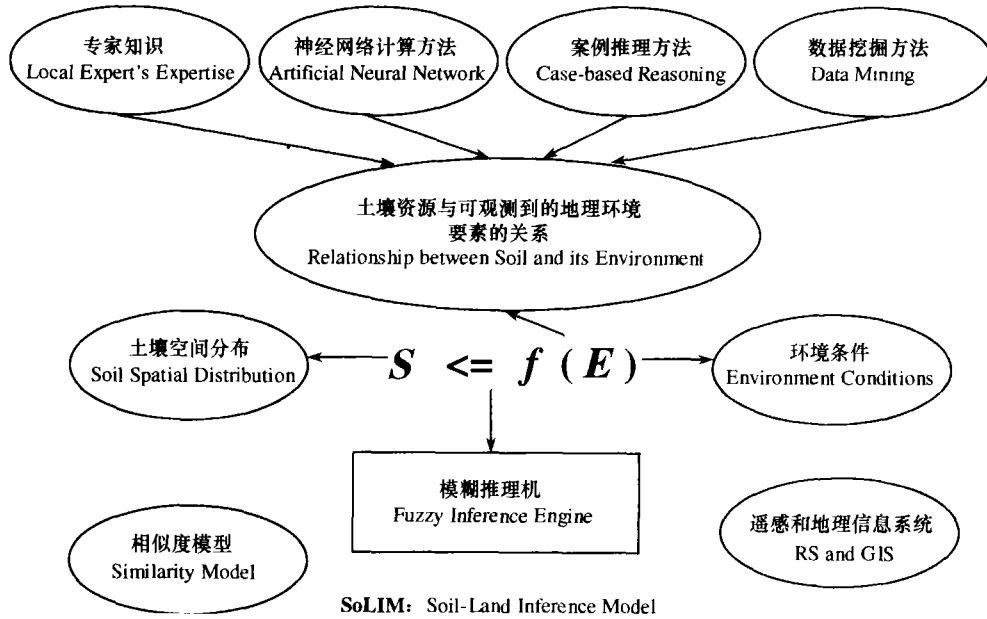


图2 模糊逻辑下的自动土壤推理

Fig.2 The automated soil inference under fuzzy logic

1.3 土壤相似度模型的应用

根据土壤相似向量除了可以获得每种土壤类型的隶属度分布图外,还可以通过硬化(Hardening)土壤相似向量中隶属度来获得土壤类型图,硬化是指将相似向量中最大的隶属度值所代表的土壤类型作为该点的土壤类型。例如,某一点的相似向量是(0.2,0.4,0.1,0.3),向量中的隶属度分别代表土壤与土壤类型 A、B、C、D 的相似程度,这里土壤类型 B 的隶属度值最大,这点的土壤类型确定为 B。在通过硬化编制的土壤类型图基础上,可根据一定域值去掉部分特别小的斑块,并将栅格形式的土壤图转变为以多变形为基础的常规土壤图(图3),使之能够和传统的土壤图进行对比分析。在相似度模型的基础上,也可对分类误差进行全面定义,从而可

以精确地、详细地描写分类结果中不确定性在空间上的变化^[32](图4)。

土壤属性图的编制是通过点(i, j)的土壤环境与各土壤类型 k 典型环境条件的相似度 S_{ij}^k 的隶属度线性加权的方法来进行,这主要是基于当地土壤的成土环境与给定土壤类型的成土环境相似,则当地土壤的性状就与给定的土壤类型性状相似的假设:

$$V_{ij} = \frac{\sum_{k=1}^n S_{ij}^k V^k}{\sum_{k=1}^n S_{ij}^k} \tag{3}$$

这里 V_{ij} 表示某点(i, j)的土壤属性值; V^k 代表典型土壤类型 k 的土壤属性值, n 代表这一地区给定的土壤类型的总数。

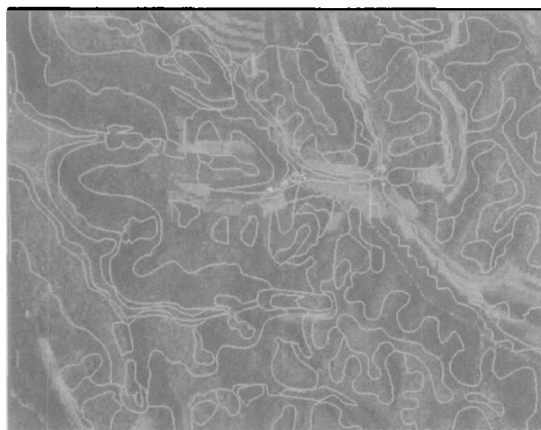


图 3 在硬化土壤图基础上生成的传统土壤图

Fig. 3 Conventional soil map derived through hardening similarity vectors

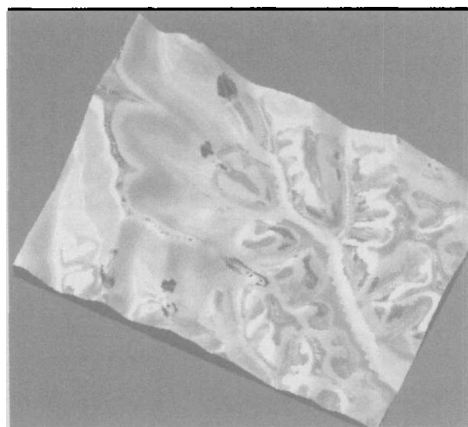


图 4 根据土壤相似度向量编制的土壤不确定性图

Fig.4 Soil uncertainty map based on soil similarity vector

2 土壤—环境推理模型的评价

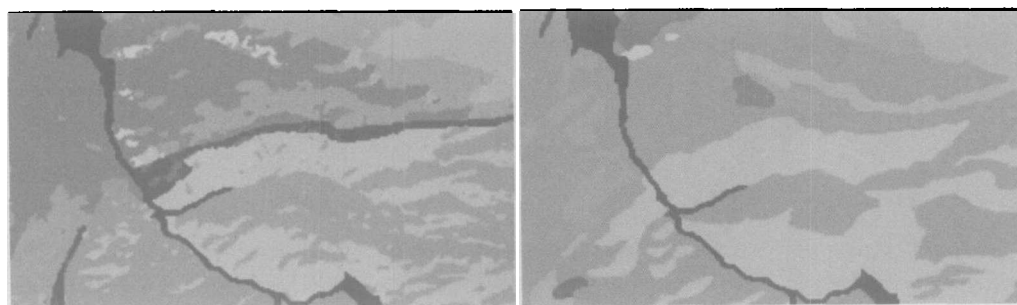
2.1 土壤图精度评价

本文在美国蒙大纳州西部的 Lubrecht 和威斯康星州 La Crosse 县东部的 Raffelson 流域检验 SoLIM 模型,采样方法为横剖面采样结合分层采样。图 5 是 Lubrecht 研究区 SoLIM 编制的土系图和传统土系图,SoLIM 推导出的土系图要比传统土系图详细得多。在 64 个点中(表 1),SoLIM 的结果有 52 个点正确(准确率 81%),而传统土系图只有 39 个点正确(准确率 61%)。SoLIM 绘制的土系图与传统土系图不一致的错配点中,其中 71%是 SoLIM 编制的土系图结果正确。Raffelson 研究区在 99 个样点中,SoLIM 推断的有 83 个正确(准确率 83.8%),而传统土系图中只有 66 个样点正确(正确率 66.7%)。

从 Lubrecht 研究区 A 层厚度空间变化图来看(图 6),SoLIM 推出的厚度图更具连续的空间变化,而传统土壤图只在土壤多边形边界处显示出这种变化,这种情况在这个研究区是不现实的。对于检验 A 层厚度的 33 个观测点,SoLIM 推出的深度(与实测值相关系数为 0.602)比传统土系图得出的深度(与实测值相关系数为 0.436)更接近现实情况。

2.2 土壤调查过程评估

(1) 采用 SoLIM 自动制图可以在整个区域使用同一个土壤—环境模型,因而不同土壤类型的制图标准是一致的,而传统土壤普查容易产生不同土壤



SoLIM 编制的土系图
SoLIM-derived soil map

传统的土系图
Conventional soil map

图 5 Lubrecht 研究区土系分布图

Fig.5 Soil series distribution map in Lubrecht

表 1 Lubrecht 研究区 SoLIM 编制的土系图和传统土系图的精度对比

Table 1 Accuracy of soil series map based on SoLIM and that of conventional soil series map for the Lubrecht study

	全区总计 Overall total			错配总计 Mismatches total		
	正确数 Correct	样点数 Samples	百分比 Percentage	正确数 Correct	样点数 Samples	百分比 Percentage
SoLIM	52	64	81	17	24	71
土壤图 Soil map	39	64	61	4	24	17

类型制图标准不一致的现象,图 7 中红色线条为传统土壤图的土壤类型边界,在坡肩部位夸大十分明

显,而黑色线条的 SoLIM 编制的土壤图边界则不存在这个问题。

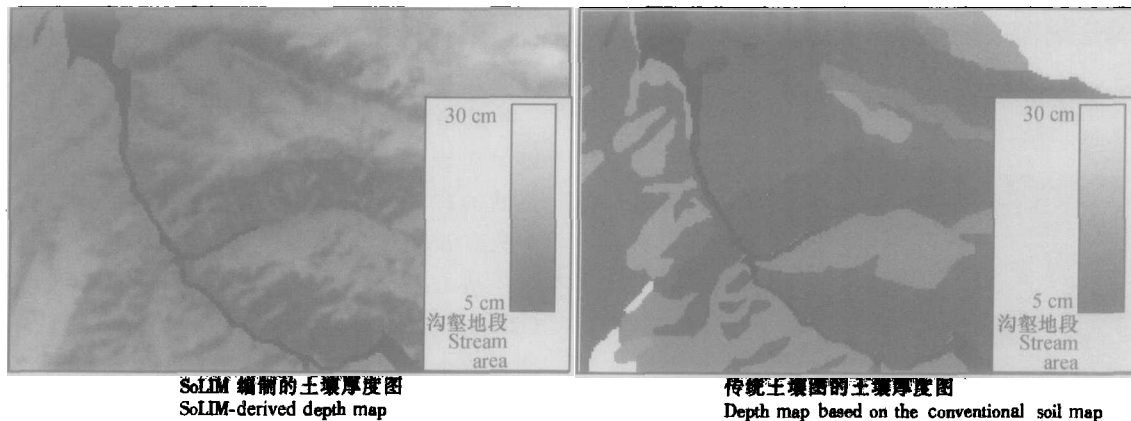


图 6 Lubrecht 研究区 A 层土壤厚度图

Fig.6 Maps of soil A-horizon depth in Lubrecht

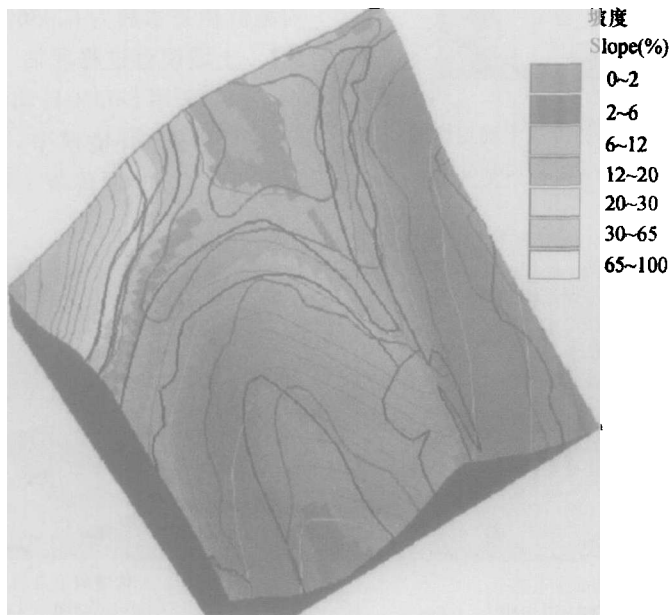


图 7 传统土系图与 SoLIM 土系图的比较

Fig.7 Comparison of soil series map from convention approach with that based on SoLIM

(2) 每个地区的 GIS 数据库和知识库都是数字化的,因此 SoLIM 可很快地生成每个地区新的土壤数据库,土壤调查更新大大快速,根据在美国威斯康星州的实验表明,普查速度比常规手段要快 10 倍左右。

(3) GIS 数据库、知识库和模糊推理机都是可以重新利用的,所以大多数最初调查和更新都是可利用的,费用大为减少,根据在美国威斯康星州的实验表明,费用比常规手段要减少 2/3 左右。

(4) SoLIM 标准设计把整个土壤调查过程分成几个部分,而每一部分都可由最适合的专家操作。这样土壤专家可从费时的制图中解放出来,更集中于发现土壤—环境关系。

(5) SoLIM 模型中,土壤环境关系的知识被清楚地表达,不会因为当地有经验的土壤专家退休而流失,可保持不同时代土壤专家之间对土壤—环境关系模型理解的一致性。

(6) 从模糊逻辑推理得出的结果都是数字化的产品,土壤数据可以直接在 GIS 或地图中应用,而不需经过复杂的数字化处理过程。数字化工作的减少也减少了数字化过程中和属性表输入过程中的错误。

3 土壤—环境推理模型在中国的应用前景

3.1 采用 SoLIM 进行土壤调查的必要性及可行性

3.1.1 必要性 (1) 虽然建国以来我国在 1958 年至 1960 年和 1979 年至 1984 年进行了两次全国土壤普查,近 20 a 随着我国经济的迅速发展,我国土地资源利用状况发生了巨大的变化,土壤也随之发生了很大变化,进行新一轮的土壤资源调查也是十分必要的。两次土壤普查,尤其是第二次土壤普查动员了全国的力量、历时 5 a,动用了大量的物力财力^[33]。采用新的土壤普查技术,使土壤普查更为快速、精度更高显得十分必要。(2) 最近一次的第二次土壤普查是从乡一级开始,按比例尺 1:1 万逐级汇总到县级 1:5 万、地区级 1:10 万、省级 1:50 万、直至全国 1:100 万、1:250 万和 1:400 万^[5]。从目前情况来看,一般能拿到最大比例尺的土壤图为县级 1:5 万土壤图,乡级和更大比例尺的土壤图则更难获得,随着时间的推移,原始的大比例尺图件被保存的可能性更小,目前这种以中小比例尺为主的土壤图已经不能满足目前农业生产的需要,需要更为精细的土壤图。(3) 我国的两次土壤普查基本上是以地形图为基础,结合航空相片和遥感数据来确定土壤形成的景观单元来编制土壤图^[5~7],也是基于多边形的图形表达和手工完成的制图过程,这种土壤图对表达土壤类型的空间差异,尤其是土壤属性的空间差异,即使是 1:1 万大比例尺的土壤图,也已不能满足精细农业的需求。

3.1.2 可能性 近 20 年来随着信息技术的发展,获得土壤环境条件能力迅速增强,国家对于基础地理信息数字化进行了大量的投入,已完成七大江河区域范围的 1:1 万的 DEM,完成了全国大部分地区 1:5 万地形

图扫描校正后的栅格数据库^[34],1:1 万的地形图也覆盖了全国的大部分地区,这些都可作为 SoLIM 模型的运行提供基本的环境参数。另外,SoLIM 需要现代信息技术尤其是 GIS 技术的支持,目前 GPS、GIS 和 RS 技术已经广泛应用于我国资源与环境的各个领域,全国各高校和科研单位都有与 GIS 和 RS 有关的研究单位,这可以为 SoLIM 应用提供有力的人员支撑。最后,目前我国土壤分类学在中国科学院和高等院校仍然有比较雄厚的力量,各地农业技术推广站等有关单位,也有对土壤分类比较熟悉的工作人员,对土壤分类方面也有比较充分的技术保证。

3.2 SoLIM 方法局限性及其改进方向

SoLIM 土壤图的质量取决于两个方面:GIS 环境下的各环境要素和从当地土壤专家获取的土壤与环境关系模型。SoLIM 方法在美国主要应用在环境梯度比较大的地区,在地形相对平缓地区的效果相对较差。如境梯度较大的 Lubrecht 研究区,用 USGS30 米 DEM 和 1:24 000 地质图就能取得很好的效果,而在环境变梯度化中等的 Raffelson 研究区,则需要 10 m 的 DEM 和 1:24 000 或者更大比例尺的地质图才能取得比较好的效果。尽管我国是以山地为主的地区,但环境梯度较小平原区使用 SoLIM 制图,可能需要更为详细的 DEM 等环境因子参数。

目前 SoLIM 需要当地土壤专家提供土壤与环境关系模型。根据我们的经验,提供这个模型需要一个有经验的当地土壤专家。尽管我国有熟悉土壤分类的专家,但是与美国不同的是,我国缺少美国那种主要工作就是研究当地土壤与环境关系的土壤普查专家,尤其是比较偏远的地区。与此同时,通过密集的采样来获得土壤与环境关系,又需要大量的人力、物力和财力。在我国两次土壤普查中也只把重点放在农地上,土壤剖面大多数在农地,山区较少,因此需要在土壤—环境关系模型方面需要更为深入的工作,尤其是研究从相对少量的野外工作来获取土壤与环境关系的技术,如改变原来的概率采样为目的的采样或者是挖掘现有的土壤图知识来获取土壤与环境关系的知识,但这两都需要更为深入的工作。

4 结论与讨论

通过 GIS 和遥感技术获得土壤形成的环境因子,根据土壤专家或者野外调查来提取土壤和环境之间的关系,构造基于模糊逻辑的推理模型——土壤—环境推理模型(SoLIM),在此基础上推导出的土

壤类型图和属性图精度与传统土壤图相比精度有较大的提高,详细程度也要详细得多,研究区较小制图单元可以在美国二级和三级详细程度的土壤普查中都能够表达,而这在传统的土壤图中是做不到的。土壤类型也表现为过渡性而不只是一种土壤类型的近似表达,使得土壤类型的描述更为准确。

SoLIM方法取得较好的效果在于在模糊逻辑下土壤—环境关系知识与原理和GIS技术的有效集成。空间上的栅格表示和属性上的相似度表达,使得土壤作为一种地理空间和属性空间的连续面,从而使土壤在地理空间和属性空间的简化大大降低。GIS对数据数字化的处理能力使得环境参数的变异可以在很详细程度上做到量化,也大大减少了土壤的概括和错误的解释。

SoLIM土壤调查法不仅能提高土壤详查的质量,也使土壤调查更新更为高效,费用更低,因此对我国土壤普查具有很重要的借鉴意义。我国近20a信息技术的发展也为SoLIM方法在我国的推广提供了有利条件。采用SoLIM进行土壤制图对环境数据和基于专家知识的土壤与关系模型要求较高,在这方面还要根据我国的实际情况进行深入的工作。

参考文献

- [1] Band L E, Moore I D. Scale: Landscape attributes and geographical information systems. *Hydrological Processes*, 1995, 9: 401 ~ 422
- [2] Peterson C. Precision GPS navigation for improving agricultural productivity. *GPS World*, 1991, 2: 38 ~ 44
- [3] United States Department of Agriculture (USDA). Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Survey. <http://soil.usda.gov/technical/classification/taxonomy/>. 2004-11-20
- [4] Hudson B D. The soil survey as paradigm-based science. *Soil Science Society of America Journal*, 1992, 56: 836 ~ 841
- [5] 全国土壤普查办公室. 中国土壤普查技术. 北京: 中国农业出版社, 1992. China Soil Survey Department. Techniques of China Soil Survey (In Chinese). Beijing: China Agricultural Press, 1992
- [6] 庄卫民著. 土壤调查与制图技术: 理论 方法 应用. 北京: 中国农业科技出版社, 1995. Zhuang W M. Techniques of Soil Survey and Mapping: Theories, Methodology, Application (In Chinese). Beijing: China Agricultural Science and Technology Press, 1995
- [7] 宋达泉编. 土壤调查手册. 北京: 科学出版社, 1955. Song D Q. ed. Soil Survey Manual (In Chinese). Beijing: Science Press, 1955
- [8] McBratney A B, Mendonca Santos M L, Minasny B. On digital soil mapping. *Geoderma*, 2003, 117: 3 ~ 52
- [9] Hu Y M, Dai J, Wang R C. GIS-based red soil resources classification and evaluation. *Pedosphere*, 1999, 9(2): 131 ~ 138
- [10] 杨胜天, 朱启疆, 李天杰. RS和GIS支持下的土壤系统分类制图方法研究——以贵州省贵阳市为例. *土壤学报*, 2001, 38(1): 41 ~ 48. Yang S T, Zhu Q J, Li T J. The research on the cartography of soil taxonomy on the basis of RS and GIS-Taking Guiyang City, Guizhou Province as sample (In Chinese). *Acta Pedologica Sinica*, 2001, 38(1): 41 ~ 48
- [11] Lane P W. Generalized linear models in soil science. *European Journal of Soil Science*, 2002, 53: 241 ~ 251
- [12] De Bruin S, Stein A. Soil-landscape modeling using fuzzy c-means clustering of attribute data derived from a Digital Elevation Model (DEM). *Geoderma*, 1998, 83: 17 ~ 33
- [13] Triantafyllis J, Ward W T, Odeh I O A, et al. Creation and interpolation of continuous soil layer classes in the lower Namoi valley. *Soil Science Society of America Journal*, 2001, 65: 403 ~ 413
- [14] Bui E N, Moran C J. A strategy to fill gaps in soil survey over large spatial extents: An example from the Murray-Darling basin of Australia. *Geoderma*, 2003, 111: 21 ~ 44
- [15] Lagacherie P, Holmes S. Addressing geographical data errors in a classification tree soil unit prediction. *International Journal of Geographical Information Science*, 1997, 11: 183 ~ 198
- [16] 沙晋明, 李小梅. 基于遥感信息的哈夫曼优化树在山地土壤资源调查中的应用——以浙江省龙游县为例. *山地学报*, 2002, 22(2): 223 ~ 227. Sha J M, Li X M. Application of Huffman Tree based on the information of remote sensing for investigating soil resource—Taking Longyou County, Zhejiang Province as sample (In Chinese). *Journal of Mountain Science*, 2002, 22(2): 223 ~ 227
- [17] Minasny B, McBratney A B. The neurom method for fitting neural network parametric pedotransfer functions. *Soil Science Society of America Journal*, 2002, 66: 352 ~ 361
- [18] Fidêncio P H, Ruisanchez I, Poppi R J. Application of artificial neural networks to the classification of soils from São Paulo state using near-infrared spectroscopy. *Analyst*, 2001, 126: 2 194 ~ 2 200
- [19] 付强, 王志良, 梁川. 自组织竞争人工神经网络在土壤分类中的应用. *水土保持通报*, 2002, 22(1): 39 ~ 43. Fu Q, Wang Z L, Liang C. Application in soil classification using Self-Organizing Mapping (SOM) network (In Chinese). *Bulletin of Soil and Water Conservation*, 2002, 22(1): 39 ~ 43
- [20] Goovaerts P. Geostatistics in soil science: State-of-the-art and perspectives. *Geoderma*, 1999, 89: 1 ~ 45
- [21] Heuvelink G B M, Webster R. Modelling soil variation: Past, present, and future. *Geoderma*, 2001, 100: 269 ~ 301
- [22] Odeh I O A, McBratney A B, Chittleborough D J. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. *Geoderma*, 1994, 63: 197 ~ 214
- [23] Zhu A X. A similarity model for representing soil spatial information. *Geoderma*, 1997, 77: 217 ~ 242
- [24] Glinka K D. *The Great Soil Groups of the World and Their Development*. Ann Arbor, MI: Edwards Bros, 1927
- [25] Jenny H E W. *Hilgard and the birth of modern soil science*. Berkeley, CA: Farallo Publication, 1961
- [26] Zhu A X. A personal construct-based knowledge acquisition process for natural resource mapping using GIS. *International Journal of Geographic Information Science*, 1999, 13(2): 119 ~ 141

- [27] Zhu A X. Mapping soil landscape as spatial continua: The neural network approach. *Water Resource Research*, 2000, 36: 663 ~ 677
- [28] Shi X, Zhu A X, Burt J E, *et al.* A case-based reasoning approach to fuzzy soil mapping. *Soil Science Society of America Journal*, 2004, 68: 885 ~ 894
- [29] Qi F, Zhu A X. Knowledge discovery from soil maps using inductive learning. *International Journal of Geographical Information Science*, 2003, 17(8): 771 ~ 795
- [30] Zhu A X, Hudson B, Burt J E, *et al.* Soil mapping using GIS, expert knowledge and fuzzy logic. *Soil Science Society of America Journal*, 2001, 65: 1 463 ~ 1 472
- [31] Zhu A X, Band L E. A knowledge-based approach to data integration for soil mapping. *Canadian Journal of Remote Sensing*, 1994, 20: 408 ~ 418
- [32] Zhu A X. Measuring uncertainty in class assignment for natural resource maps using a similarity model. *Photogrammetric Engineering & Remote Sensing*, 1997, 63: 1 195 ~ 1 202
- [33] 龚子同. 面临新世纪挑战的土壤地理学. *土壤*, 1999, (5): 236 ~ 243. Gong Z T. The prospect of soil geography under new challenges in the 21st century. *Soils*, 1999, (5): 236 ~ 243
- [34] 国家基础地理信息中心. 国家基础地理信息系统. <http://nfgis.nsd.gov.cn/>. 2004-12-01. National Fundamental Geographic Information Center. National Fundamental Geographic Information System. <http://nfgis.nsd.gov.cn/>. 2004-12-01

PREDICTIVE SOIL MAPPING BASED ON A GIS, EXPERT KNOWLEDGE, AND FUZZY LOGIC FRAMEWORK AND ITS APPLICATION PROSPECTS IN CHINA

Zhu Axing^{1, 2} Li Baolin¹ Yang Lin¹ Pei Tao¹ Qin Chengzhi¹ Zhang Ganlin³ Cai Qiangguo¹
Zhou Chenghu¹

(1 *State Key Laboratory of Environment and Resources Information System, Institute of Geographical Sciences and Resources Research, Chinese Academy of Sciences, Beijing 100101, China*)

(2 *Department of Geography, University of Wisconsin-Madison, Madison, WI 53706, USA*)

(3 *Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China*)

Abstract Detailed soil spatial and attribute information are now basic parameters for environmental modeling and land management applications. The accuracy and efficiency of conventional soil surveys, based on the polygon model and the manual mapping practice, are quite low. A geographical information system (GIS) and expert knowledge based-fuzzy soil inference scheme (soil-land inference model, SoLIM) was developed to overcome the problems faced by the conventional soil survey. The scheme consists of three major components: (i) a model employing a similarity representation of soils, (ii) a set of inference techniques for deriving similarity representation, and (iii) application of the similarity representation. According to case studies conducted in Wisconsin, U. S. A., SoLIM improves the accuracy and efficiency of soil survey. Soil type and properties maps based on SoLIM are better than these based on conventional techniques in term of both spatial detail and attribute accuracy. The accuracy of the soil series maps based on SoLIM is about 80%. Moreover, soil mapping by means of SoLIM is about ten times faster than by conventional ones and saves about 2/3 in cost. However, how the SoLIM works highly depends on the availability and quality of environmental data and the quality of soil, environmental relationship model for the study area.

The Second Soil Survey in China was conducted 20 years ago. Recent intensive land use activities may have greatly impacted the soil conditions. It's quite necessary to update the soil data in China for agricultural purposes. The SoLIM framework can provide great assistance in updating soil resource inventory. At the same time, the availability of spatial data, spatial information processing technology and human resources related to GPS, GIS and RS also make it possible to apply the SoLIM approach in the Chinese soil survey. However, the degree of success of the SoLIM highly depends on the quality of knowledge on soil environmental relationships in the study area. The knowledge about soil-environmental relationships used for the SoLIM in USA comes mainly from local soil survey experts. Although experts in soil taxonomy can provide much help for soil survey in China, the number of experts who focus on soil mapping in local areas is small, which will pose a major problem in using the SoLIM approach in China due to its reliance on human expertise for soil-environmental relationships. If we try to discern the soil-environmental relationships by field sampling, we will need a large amount of samples, which is too expensive and too slow. Thus, we must develop new approaches for extraction of knowledge on soil-environmental relationships. These new approaches could include new sampling methods, such as purposive sampling to reduce fieldwork and data mining for extracting knowledge from existing soil maps.

Key words Soil-Land Inference Model (SoLIM); GIS; Fuzzy Logic; Expert Knowledge; Soil mapping