

基于样点的数字土壤属性制图方法及样点设计综述^①

张淑杰^{1,2}, 朱阿兴^{1,3}, 刘京³, 杨琳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 中国科学院大学, 北京 100049;

3 Department of Geography, University of Wisconsin-Madison, Madison WI 53706)

摘要: 土壤剖面数据与土壤类型图按照某种原则进行连接是目前获取土壤属性空间分布信息的主要方法, 这种传统的土壤属性制图方法以土壤专家的“经验”和手工描绘为基础, 耗费资本高、生产周期长。数字土壤制图通过借鉴先进的空间信息处理技术和高分辨率地形数据的优势, 能够快速准确地获取高精度、高分辨率的土壤属性空间变化信息, 是一种精细、高效、经济的土壤属性制图技术。本文详细介绍了基于样点进行数字土壤属性制图的3种方法: ①基于空间自相关的方法; ②基于空间自相关和土壤-环境关系混合相关的方法; ③基于土壤-环境关系的方法。同时, 为保证样点能够全面地捕捉到研究区内土壤属性空间变异特征, 以上3种方法都对样点的数量、分布或典型性提出了较为严格的要求, 即样点应具有全局代表性。因此, 如何设计样点成为数字土壤属性制图中的一个重要问题。依据样点设计过程中是否能够整合已有样点, 本文将样点设计方案分为采样设计方案和补样设计方案两种, 并对其分别进行了详细的综述。

关键词: 数字土壤属性制图; 全局代表性; 采样方案设计; 补样方案设计

中图分类号: P934

土壤属性空间分布信息是进行土壤侵蚀模拟、水资源管理等方面研究的重要基础信息。目前, 我国1:100万土壤属性图是通过将土壤剖面数据与土壤类型图按照某种原则连接而获取的^[1-2]。这种基于多边形的传统土壤类型图所获取的土壤属性图无论在空间分辨率还是在属性精度方面都无法满足上述研究对土壤信息详细度的要求。数字土壤制图技术通过借鉴先进的空间信息处理技术和高分辨率地形数据的优势, 为快速地获取高精度、高分辨率的土壤分布信息提供了有效的途径。相比于以土壤专家的“经验”和手工操作为基础并以多边形为表达方式的传统土壤图, 以栅格形式表达土壤空间变化特征的数字土壤图, 在节约制图成本、缩短制图时间的同时提高了制图精度和表达的详细程度^[3], 数字土壤制图已经成为当前最重要的土壤制图方法。

数字土壤属性制图方法充分利用了土壤属性空间分布的自相关关系和土壤-景观关系, 在大多数情况下野外采集的样点是挖掘土壤属性空间自相关关系和建立土壤-环境间关系知识的主要来源。这是因为, 挖掘土壤空间自相关关系和建立土壤-环境间关系知识的来源包括3个方面: 有经验的土壤普查专家、已有的

土壤图和野外采集的土壤样点^[4]。基于当地土壤普查专家建立空间自相关和土壤-环境间关系知识的过程是利用人工智能手段定量地表达专家脑海中描述性的知识, 但所获知识的正确性在很大程度上取决于专家经验的丰富程度, 同时由于我国大多数地区缺乏对当地土壤空间分布规律有丰富经验的普查专家, 这种方法在实际应用中受到了很大的限制^[5-6]; 基于已有土壤图建立土壤-环境间关系知识的过程是利用数据挖掘的手段, 还原土壤专家当年制图过程中隐含在土壤图中的知识, 但所获知识的正确性很大程度上取决于原有土壤图的精度, 同时由于我国大部分地区不存在高精度的土壤图, 已有精度较低的传统土壤图一般只能作为数字土壤制图的辅助信息, 使得这种方法在我国大面积的推广应用同样受到很大限制^[7]。因此, 野外采集的土壤样点在很多情况下成为进行数字土壤属性制图唯一可用的数据源。

目前, 基于样点进行数字土壤属性制图的方法可以归纳为3大类: ①基于空间自相关的数字土壤属性制图方法; ②基于空间自相关和土壤-环境关系混合相关的数字土壤属性制图方法; ③基于土壤-环境关系的数字土壤属性制图方法。其中, 前两种方法都是建立

①基金项目: 国家自然科学基金项目(40971236、41023010、41001298)和科技部国际科技合作项目(2010DFB24140)资助。

作者简介: 张淑杰(1983—), 女, 山东德州人, 博士研究生, 主要从事数字土壤制图及样点设计研究。E-mail: zhangsj@reis.ac.cn

在大量具有全局代表性样点基础上的制图方法,样点的数量和布局往往对推测结果有重大影响^[8-10]。第三种基于土壤-环境关系知识模型的方法,即可以建立在大量具有全局代表性样点的基础上也可以建立在少量能够代表土壤属性空间变化特征的典型样点基础上^[11-12]。由此可见,不论对于哪种方法,都对样点的数量、分布或典型性提出了严格的要求,以保证样点集的全局代表性。

因此,基于样点进行数字土壤属性制图的方法,以及设计能够满足制图方法要求的具有全局代表性的样点是数字土壤属性制图中两个重要的研究内容。本文在第一部分首先总结了基于样点的数字土壤属性制图方法,在第二部分和第三部分根据样点设计过程中是否兼顾已有样点资源将样点设计方案分为采样设计方案和补样设计方案两大类,并对其分别进行了详细的综述,最后是展望部分。

1 基于样点的数字土壤属性制图方法

1.1 基于空间自相关的数字土壤属性制图方法

基于空间自相关推测土壤属性空间分布的方法是根据给定的一组或多组离散样点的土壤属性值与其空间位置或距离拟合函数方程,使该方程能够充分反映样点属性值与空间位置或距离间的数学关系,从而推断一定区域范围内其他未采样点土壤属性的方法^[13]。代表性的方法有趋势面分析、样条函数、反比距离加权、普通克里格等方法。

趋势面分析是以样点的地理坐标为自变量,以样点位置土壤属性值为因变量的低阶多项式,趋势面分析的目的是对样点集属性特征进行全局和局部规律的研究。Edmonds 和 Campbell^[14]根据美国弗吉尼亚州和邻近各州的观测站数据,利用趋势面分析的方法来确定土壤的温度值。王会肖和张超^[15]在陕西省绥德县辛店沟一典型区,基于规则采集的样点利用趋势面分析的方法推测了土壤水分的空间变化趋势。

样条函数法首先对研究区进行分区,在每一分区以多项式的形式拟合表面,之后连接各分区表面建立能够满足最优平滑原则的曲面。该方法在样点数据充足且冗余度低的时候,计算结果快速且光滑性较好^[16-17]。赫晓慧和温仲明^[18]在黄土高原丘陵区一小流域,利用薄板光滑样条函数插值法,对各地形因子对土壤水分空间变异性的影响与趋势进行了研究。

反比距离加权法是一种局部插值方法,认为空间上越临近的点越具有相似的特征,根据已知样点与待推测点的空间距离来确定加权系数。该方法简便易行,

其结果受所采用的局域空间范围大小和参与的样本数量影响很大,在样点密度高且样点在空间上规则分布的情况下能获得较高的精度^[19]。庞国伟等^[20]利用该方法对陕西省水蚀土壤因子的土壤空间变异特征进行研究。

普通克里格方法同样认为空间上距离相近的点之间比距离较远的点之间具有更大的相似性,与反比距离方法不同的是,克里格方法充分考虑了土壤属性在空间上的自相关特征。该方法首先需要通过大量的野外样点来估计方差,并建立半方差函数,之后确定土壤插值模型(如球状模型、指数模型、高斯模型等)^[21-22],最后根据邻近样点的土壤属性值对待推测点进行空间插值。Burgess 和 Webster^[23]最早开始在土壤制图领域使用地统计学的方法,目前已经得到了广泛的应用。

综上所述,基于空间自相关推测土壤属性空间分布的方法要求样点密度高,在样点能够很好地捕捉土壤属性空间自相关特征的情况下能够得到理想的结果。

1.2 基于空间自相关和土壤-环境关系混合相关的数字土壤属性制图方法

这种基于空间属性混合相关的数字土壤属性制图方法,是在上述空间自相关模型的基础上引入了协变量要素,即在充分考虑土壤属性空间分布具有自相关特征的同时,以土壤发生学原理为基础,引入气候、生物、岩石、地形等成土因素作为协同变量共同推测土壤属性的空间变异情况。该方法同时考虑空间自相关性和环境要素相关性。在土壤属性与协同变量要素相关性较好时,可以显著提高模型的插值精度^[24]。代表方法有协同克里格插值、回归克里格插值等。

协同克里格插值方法:在普通克里格插值考虑到样本本身空间自相关特征的基础上,协同克里格插值方法在推测土壤属性的空间分布时,同时兼顾了在空间上连续分布的其他变量(即协同变量)来提高土壤属性的制图精度^[25-26]。这些协同变量可以是与待推测点土壤属性具有协同变化关系的其他土壤属性,也可以是根据数字高程模型和遥感数据获得的协同环境变量(如:坡度、植被指数等)^[27]。

回归克里格插值方法:将基于参数关系的统计方法与基于空间自相关的统计方法进行结合,从而产生了回归克里格插值方法。该方法的思路是首先将土壤属性值与相关环境要素进行回归分析,然后将回归模型的残差项作为区域化变量进行普通克里格内插,最后的推测结果值为普通克里格的内插值与回归模型预

测值的加和^[28-29]。

综上所述，基于空间自相关和土壤-环境关系混合相关的数字土壤制图方法同样需要符合某种分布规则的大量样点来捕获土壤属性、相关环境要素的空间分布特征以及二者之间的关系。

1.3 基于土壤-环境相关的数字土壤属性制图方法

基于土壤-环境关系知识模型推测土壤属性空间分布的方法，是现有数字土壤制图方法中最为常用的方法之一。俄国著名土壤学家道库恰耶夫的土壤发生学理论和美国土壤学家 Jenny^[30]的土壤-景观模型“CLORPT”($S = f(Cl, O, R, P, T, \dots)$)， S 为土壤， Cl 为气候， O 为生物， R 为地形， P 为母质， T 为时间，省略号为其他尚未确定或认知的因素)中指出，土壤是成土母质在一定地热水热条件和生物作用下，经过一系列的物理、化学和生物化学过程形成的。“CLORPT”函数关系式明确指出，土壤是气候、生物、地形、母质、时间共同作用形成的，同时还提出不同地区、不同类型的土壤，往往有不同的成土因素占优势。基于土壤-环境相关的数字土壤属性制图方法依据土壤发生学理论和土壤-景观理论，利用土壤属性和环境要素之间的协同关系，由环境条件的空间变化来推测土壤的空间变化，其中获取土壤-环境关系知识是该方法的关键。基于样点获取土壤-环境关系知识的方法可依据样本量分为：基于大量样本的知识获取方法和基于少量样本的知识获取方法。

基于大量样本的知识获取方法：基于大量土壤样本的方法通过对土壤属性与环境数据进行分析，如回归分析^[31-32]、神经网络^[33-34]、回归树/决策树^[35-36]等，提取隐含在已有数据中的知识，并用于推测土壤属性的空间分布。因此，这些方法需要大量的野外样点来保证样点集的全局代表性进而确保提取到全局的土壤-环境关系知识。

基于少量样本的知识获取方法：这类方法降低了对样本数量的要求。当样点为能够反映研究区土壤变化特征的典型样点时，可以提取蕴含于典型样本的全局土壤-环境关系知识推测整个研究区的土壤空间变化^[37]，这种方法的关键是如何确定样点是能够全面反映当地土壤-环境关系的典型样点^[38-39]；当样点的全局代表性不好时，则可以基于样点的“个体代表性”推测样点集能够代表范围内的土壤属性值^[40]。这种方法产生的土壤图并不能覆盖整个研究区，需要进一步设计补充样点来完备样点集的全局代表性。

综上所述，目前基于样点获取土壤-环境关系知识的数字土壤制图方法或者对样点提出了严格的要求或

推测结果只能部分覆盖研究区。

2 采样方案设计方法

野外采集的样点是进行土壤属性制图的重要数据来源，所获取的样点应该在数量、分布或典型性方面符合现有数字土壤制图方法的要求，才能够保证推理结果的准确性和完整性。按照采样方案设计方式的不同，本文将其分为3类：①空间均匀布样方案；②土壤及其协同环境要素属性域内的均匀布样方案；③兼顾地理空间域和属性域的均匀布样方案。

2.1 空间均匀布样方案

为使样点在研究区空间范围内尽量均匀分布，一般采取规则格网的设计方案，该方案包括格网形状的设计(如正方形格网、三角格网、六边形格网)和格网间距的设计。尽管 Yfantis 等^[41]和王珂等^[42]经研究得出三角格网的采样方式要优于正方形格网和六边形格网的结论，但正方形格网仍然是实际采样过程中应用最多的采样方式。一般情况下格网间距越小，得到的土壤属性估计值越精确，但同时较大的采样数量会产生更多的采样费用。格网间距一般都是基于研究区土壤属性空间变异规律的特征来确定，当土壤属性的空间自相关性较大时，可以适当加大格网间的距离，反之应设计较小的格网间距^[43-44]。王秀等^[45]指出在田间尺度上，对于大多数作物采样间隔为32m时得到的结果比较理想，采样间隔最大也不应该超过90m。然而，不同的土壤属性具有不同的空间变异规律，针对特定的土壤属性，应设置不同的采样间距^[46-47]。规则格网的布样方案，在实际应用中往往会遇到很多问题，有些网格点可能坐落在建筑物覆盖地、水域或者无法到达的高山峡谷地区，致使样点无法获取。同时，当基于获取的样点进行属性推理时，边界地区尤其是不规则的边界地区误差往往较大^[48]。

2.2 土壤及其协同环境要素属性域内的均匀布样方案

这种采样方案设计的样点，适合用土壤-环境间关系知识推测土壤属性空间分布的数字土壤制图方法。采样人员希望采集那些能够更好地代表土壤属性与协同环境要素之间关系的样点，要求样点位置处环境变量的值能够涵盖环境变量的整个值域范围，从而能够有效地提高采样效率和(或)降低采样数量^[49]。Gessler 等^[50]提出了基于复合地形指数(compound topographic index)随机选择采样单元的方法。Brus 等^[51]首先对地理空间进行聚类分析，然后将聚类中心作为采样点。Hengl 等^[52]提出了基于环境协同变量的主成分分析结

果进行采样设计, 建议采样数量与各环境变量的主成分分数成比例。Zhu 等^[39]和赵量等^[53]首先根据研究区内与土壤属性协同变化的环境要素进行模糊 c 均值聚类, 然后在每一类别中选择隶属度高的典型点作为样点, 这种目的性采样方案既能保证所采集的样点在属性空间中具有很好的代表性, 又能够有效地降低采样数量提高采样效率, 成为目前一种经济、实用的新型采样方案。

2.3 兼顾地理空间域和属性域的均匀布样方案

如若样点能够同时兼顾覆盖地理空间域和属性域, 则它们能够更全面地捕捉土壤属性的空间变化特征。最具代表性的采样方案有: 分层采样、方差树采样以及拉丁超立方采样。

分层采样首先是对研究区按照一定的指标特征进行划分, 之后在每一分区内利用简单随机或其他方法布设样点。McKenzie 和 Ryan^[54]首先基于地形属性、气候特征和地质数据将研究区划分为不同的景观单元, 之后在每一类同质的景观中随机选取采样单元; 姚荣江等^[55]基于土壤盐分空间变异情况及田间的可操作性, 综合考虑土壤含盐量的理论分级以及实际田间变异程度对研究区进行划分, 然后采用最适分配法确定每个区的样点数量。

方差树采样首先在研究区内进行粗尺度的采样, 然后在属性变化比较大的区域再进行加密采样, 从而确保采样区域内属性的变化能够被有效地表达, 最终确定较为有效的采样方案^[56-57]。

拉丁超立方体采样方案的基本思想为: 影响土壤形成的多个协同变量在属性空间中形成一个超立方体的分布形式, 其中空间位置也可以作为一个属性变量加入到这个超立方体空间中, 样点设计时应该尽可能地覆盖这个超立方体。拉丁超立方体采样完全依赖于原始数据的分布, 是一种无参数的采样方法^[58-59]。

3 补样方案设计方法

由上可知, 采样方案在样点设计的过程中并不能兼顾已有的土壤样点, 造成了资源的浪费。如何在已有样点的基础上, 设计补充样点并确定每一个样点的最优位置成为样点设计人员另一个关注的焦点。本文将兼顾已有样点的设计方案称为补样方案, 并从样点位置优化准则和所使用的优化算法两个方面对设计补样方案进行综述。

3.1 确定样点采样位置的优化准则

常用的优化准则包括: 最小化平均估计方差准则、最小化最大估计方差准则、WM (warrick-myers) 准则、

平均最短距离最小化准则 (MMSD) 等。除此之外, 还有极大熵准则、分形维度、均方距离准则等^[60]。

在对研究区土壤空间结构具有一定先验知识的情况下, 可首先基于已有样点利用克里格插值的方法计算每一栅格位置的估计方差, 然后采用最小化平均估计方差或最小化最大估计方差的优化准则对补充样点的位置进行优化。最小化平均估计方差的准则能够保证所得推测结果在总体上符合制图者的精度要求, 但不能保证推测结果的任意位置方差都很小, 基于这种优化准则设计的样点推测的土壤属性空间分布图往往会在研究区边界部分具有较大估计方差; 最小化最大估计方差的准则能够保证推测结果任意位置的方差都在制图者要求的范围内, 其所设计的样点向边界位置靠近从而避免这些地区方差相对较大的问题^[51,61]。

在对研究区土壤空间分布认识不足的情况下, 可采用 WM 准则、MMSD 准则对样点进行空间位置优化。WM 准则尽量使得设计样本点对的空间分布能够更好地拟合预先定义好的理想分布 (比如均匀分布) 的实验变差函数的点对空间分布状况。由于该准则只考虑了样本点之间的距离分布, 为了避免样本点的聚集, 通常的做法是: 首先将部分样本点按系统抽样的方式大致布满整个研究区, 然后再使样本点对的分布与设计分布最大可能地保持一致^[62-63]。平均最短距离最小化准则通过使研究区内任意一个位置的点到与它最邻近的样点之间的距离的期望值最小, 进而达到样本点在地理空间中尽可能均匀覆盖的目的。

3.2 确定样点采样位置的优化算法

样点设计的优化准则明确以后, 便可以通过最小化 (或最大化) 优化准则来确定样点的最优位置。枚举法用于解决这类问题最简单也最直接的方法, 该方法将所有可能选择的样本位置均看作预选样点, 然后分别计算出每种组合情况下的目标值, 最后选择最优的样本组合位置。由于该方法对所有可能布设样点的位置组合全部进行了对比, 因此枚举法确定的样本一定是最优的样本。当可供选取的样本数量较少时, 利用枚举的方式寻求最佳的样本位置简单方便。然而, 在数字土壤制图中可供选择的备选样本量一般都很大, 这时样本位置的不同组合方式会异常复杂, 其计算量相当大从而转变为“组合爆炸”问题。为解决大规模组合优化问题, 许多学者提出了一些亚启发式、启发式、仿生等优化算法, 如贪婪法、爬山算法、禁忌算法、模拟退火算法、遗传算法、蚁群算法、粒子群算法等。这些算法大大提高了求解问题的效率, 其中模拟退火算法成为土壤样点设计方案中应用最广泛

的优化算法^[64-66]。

4 研究展望

(1) 数字土壤属性制图方面: 基于空间自相关的数字土壤属性制图方法、基于空间自相关和土壤-环境关系混合相关的数字土壤属性制图方法以及基于土壤-环境相关的数字土壤属性制图方法是现有数字土壤属性制图的3种基本方法, 以上3种方法对样点的数量、分布或典型性都提出了一定的要求。只有当这些样点具备一定的数量并且在分布上符合某种特定的规则(如规则分布或随机分布等)或是能够反映土壤空间变化规律的典型样点时, 才能保证样点集对研究区内土壤空间分布规律的全局代表性, 进而保证数字土壤图产品的精度^[67]。然而, 我国大多数地区经过历次土壤普查或服务于特定研究目的的野外考察所累积的野外样点很难满足上述3种制图方法的要求。因此, 研究如何充分利用这些不符合某种特定规则或典型性不好的样点, 合理地推测土壤属性的空间变化是数字土壤属性制图中一个重要的研究方向。

另外, 现有数字土壤属性制图方法大都只是基于某一个数据源, 即或者是通过挖掘土壤普查专家脑海中的知识并对其量化表达后进行推理制图, 或者是基于大量样点直接进行空间插值, 或者是利用数据挖掘的方法(如神经网络、决策树、判别分析等)从大量样点或典型样点或传统土壤图中挖掘土壤-环境关系知识后进行推理制图。但是, 对于某一个特定地区来讲, 这几种数据源往往是共同存在的, 因为大多数地区经过历次的土壤普查或针对特定目的的野外调查已经积累了一定数量的样点, 同时1:100万的土壤类型图已经覆盖我国全部范围。另外, 调查报告或出版的文献(如全国土种志、地方土种志)中也有对某些样点的位置及其周边环境信息的描述。因此, 如何综合地、有效地挖掘隐含于土壤普查样点、传统土壤图、历史文献资料等宝贵遗留数据中的土壤-环境关系知识, 提高数字土壤图的精细程度, 必是一个重要的研究方向。

(2) 样点合理布设方面: 基于经典统计理论的样点设计方案(如: 规则布样、分层布样等)和基于地统计理论的样点设计方案(如: 克里格布样)所设计的样点都很难兼容已有的土壤样点(如土种志中记录的土壤普查样点)。然而, 这些已有的土壤样点都包含了采样者对当地土壤-环境间关系知识的认识, 样点设计过程中应当充分整合这些已有样点。因此, 在充分利用已有样点资源的基础上, 设计高效、逐次地补充

样点, 也是数字土壤属性制图中的一项重要研究内容。

参考文献:

- [1] 张定祥, 史学正, 于东升, 潘贤章, 孙维侠. 中国 1:100 万数据库建立的基础. 地理学报, 2002, 57(B12): 82-86
- [2] 张定祥, 潘贤章, 史学正, 杨金玲, 林杰. 中国 1:100 万土壤数据库建设中的几个问题. 土壤通报, 2003, 34(2): 81-84
- [3] 朱阿兴, 李宝林, 杨琳, 裴韬, 秦承志, 张甘霖, 蔡强国, 周成虎. 基于 GIS、模糊逻辑和专家知识的土壤制图及其在中国的应用前景. 土壤学报, 2005, 42(5): 844-851
- [4] 朱阿兴, 李宝林, 裴韬, 秦承志, 杨琳, 王平, 刘峰, 李润奎, 刘京. 精细数字土壤普查模型与方法. 北京: 科学出版社, 2008: 22-23
- [5] Zhu AX, Lawrence EB, Barry D, Thomas JN. Automated soil inference under fuzzy logic. Ecological Modelling, 1996: 123-145
- [6] Zhu AX. A personal construct-based knowledge acquisition process for natural resource mapping. International Journal of Geographical Information Science, 1999, 13(2): 119-141
- [7] Qi F, Zhu AX. Knowledge discovery from soil maps using inductive learning. International Journal of Geographical Information Science, 2003, 17(8): 771-795
- [8] Mueller TG, Pierce FJ. Soil carbon maps: Enhancing spatial estimates with simple terrain attributes at multiple scales. Soil Science Society of America Journal, 2003, 67(1): 258-267
- [9] 朱会义, 贾绍凤. 降雨信息空间插值的不确定性分析. 地理科学进展, 2004, 23(2): 34-42
- [10] 姚丽贤, 周修冲, 蔡永发, 陈婉珍. 不同采样密度下土壤特性的空间变异特征及其推估精度研究. 土壤, 2004, 36(5): 538-542
- [11] 赵量, 赵玉国, 李德成, 孙孝林, 张甘霖. 基于模糊集理论提取土壤-地形定量关系及制图应用. 土壤学报, 2007, 44(6): 961-967
- [12] 杨琳, 朱阿兴, 秦承志. 运用模糊隶属度进行土壤属性制图的研究——以黑龙江省鹤山农场研究区为例. 土壤学报, 2009, 46(1): 9-15
- [13] Goovaerts P. Geostatistics in soil science: State-of-the-art and perspectives. Geoderma, 1999, 89: 1-45
- [14] Edmonds WJ, Campbell JB. Spatial estimates of soil temperature. Soil Science, 1984, 138: 203-208
- [15] 王会肖, 张超. 利用 MATLAB 研究土壤水分空间变异初探. 中国生态农业学报, 2007, 15(1): 127-130
- [16] 时培忠. 空间插值和缺值方法的研究与应用(硕士学位论文). 北京: 中国科学院地理科学与资源研究所, 2000

- [17] 张俊, 徐绍辉, 刘建立, 张佳宝. 应用三次样条函数计算土壤水分扩散率. 土壤, 2006, 38(3): 338-341
- [18] 赫晓慧, 温仲明. 小流域地形因子影响下的土壤水分空间变异性研究. 水土保持研究, 2008, 15(2): 80-83
- [19] Chang KT. Introduction to Geographic Information Systems. New York: McGraw Hill, 2004: 400
- [20] 庞国伟, 杨勤科, 张爱国, 李锐. 陕西省水蚀土壤因子指标插值方法比较研究. 水土保持通报, 2009, 29(2): 176-182
- [21] 李连发. 地学空间数据分析研究及应用 (博士学位论文). 北京: 中国科学院地理科学与资源研究所, 2005
- [22] 王劲峰. 空间分析. 北京: 科学出版社, 2006: 62-68
- [23] Burgess TM, Webster R. Optimal interpolation and isarithmic mapping of soil properties: I. The semivariogram and punctual kriging. Journal of Soil Science, 1980, 31: 315-331
- [24] 朱会义, 刘述林, 贾绍凤. 自然地理要素空间插值的几个问题. 地理研究, 2004, 23(4): 425-432
- [25] McBratney AB, Webster R. Optimal interpolation and isarithmic mapping of soil properties: V. Co-regionalization and multiple sampling strategy. Journal of Soil Science, 1983, 34: 137-162
- [26] 孙波, 宋歌, 曹尧东. 丘陵区水稻土 Cu 污染空间变异的协同克里格分析. 农业环境科学学报, 2009, 28(5): 865-870
- [27] Odeh IOA, McBratney AB, Chittleborough DJ. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma, 1994, 63: 197-214
- [28] 赵永存, 史学正, 于东升, 赵彦锋, 孙维侠, 王洪杰. 不同方法预测河北省土壤有机碳密度空间分布特征的研究. 土壤学报, 2005, 42(3): 379-385
- [29] 张慧智, 史学正, 于东升, 王洪杰, 赵永存, 孙维侠, 黄宝荣. 中国土壤温度的空间预测研究. 土壤学报, 2009, 46(1): 1-8
- [30] Jenny H. Factors of Soil Formation: A System of Quantitative Pedology. New York: Dover Publ., 1941
- [31] Odeh IOA, McBratney AB, Chittleborough DJ. Further results on prediction of soil properties from terrain attributes: Heterotopic cokriging and regression-kriging. Geoderma, 1995, 67: 215-225
- [32] 孙孝林, 赵玉国, 秦承志, 李德成, 赵量, 张甘霖. DEM 栅格分辨率对多元线性土壤—景观模型及其制图应用的影响. 土壤学报, 2008, 45(5): 971-977
- [33] Zhu AX. Mapping soil landscape as spatial continua: The neural network approach. Water Resources Research, 2000, 36(3): 663-677
- [34] 沈润平, 丁国香, 魏国栓, 孙波. 基于人工神经网络的土壤有机质含量高光谱反演. 土壤学报, 2009, 46(3): 391-397
- [35] Behrens T, Förster H, Scholten T, Steinrücken U, Ernst-Dieter S, Goldschmitt M. Digital soil mapping using artificial neural networks. Journal of Plant Nutrition and Soil Science, 2005, 168: 21-33
- [36] 张秀英, 孙棋, 王珂, 蒋玉根, 林芬芳, 韩凝. 基于决策树的土壤 Zn 含量预测. 环境科学, 2008, 29(12): 3508-3512
- [37] Shi X, Zhu AX, Burt JE, Qi F, Simonson D. A case-based reasoning approach to fuzzy soil mapping. Soil Science Society of America Journal, 2004, 68: 885-894
- [38] Webster R, Oliver MA. Statistical Methods in Soil and Land Resource Survey. Oxford: Oxford University Press, 1990: 316
- [39] Zhu AX, Yang L, Li BL, Qin CZ, English E, Burt JE, Zhou CH. Purposive sampling for digital soil mapping for areas with limited data // Hartemink AE, McBratney AB, Mendonca Santos ML. Digital Soil Mapping with Limited Data. New York: Springer-Verlag, 2008: 233-245
- [40] 刘京. 基于样点个体代表性的大范围土壤属性制图方法研究 (硕士学位论文). 北京: 北京师范大学, 2010
- [41] Yfantis EA, Flatman GT, Behar JV. Efficiency of kriging estimation for square, triangular and hexagonal grids. Mathematical Geology, 1987, 19(3): 193-205
- [42] 王珂, 沈掌泉, John SB, Jordan CF. 精确农业田间土壤空间变异与采样方式研究. 农业工程学报, 2001, 17(2): 33-36
- [43] 许红卫, 王珂. 田间土壤采样数据的统计特征与空间变异性研究. 浙江大学学报(农业与生命科学版), 2000, 26(6): 665-669
- [44] 任振辉, 吴宝忠. 精细农业中最佳土壤采样间距确定方法的研究. 农机化研究, 2006(6): 82-85
- [45] 王秀, 赵春江, 孟志军, 叶涛, 马智宏, 薛绪掌. 精准农业土壤采样栅格划分方法的研究. 土壤学报, 2005, 42(2): 199-205
- [46] 齐文虎, 谢高地, 丁贤忠. 精准农业土壤采样密度研究—以上海精准农业试验示范基地为例. 中国生态农业学报, 2003, 11(1): 199-205
- [47] 张淑娟, 方慧, 何勇. 精准农业田间信息采样策略. 农业机械学报, 2004, 35(4): 88-92
- [48] Heuvelink GBM, Brus DJ, de Gruijter JJ. Optimization of sample configurations for digital soil mapping of soil properties with universal kriging // Lagacherie P, McBratney AB, Voltz M. Digital Soil Mapping –An Introductory Perspective. New York: Elsevier, 2007: 137-151
- [49] Minasny B, McBratney AB. Latin hypercube sampling as a tool for digital soil mapping // Lagacherie P, McBratney AB, Voltz M. Digital Soil Mapping –An Introductory Perspective. New York: Elsevier, 2007: 153-165
- [50] Gessler PE, Moore ID, McKenzie NJ, Ryan PJ. Soil-landscape modelling and spatial prediction of soil attributes. Int. J.

- Geographic Information Systems, 1995, 9(4): 421-432
- [51] Brus DJ, de Gruijter JJ, van Groenigen JW. Designing spatial coverage samples using the K-means clustering algorithm // Lagacherie P, McBratney AB, Voltz M. Digital Soil Mapping –An Introductory Perspective. New York: Elsevier, 2007: 183-192
- [52] Hengl T, Rossiter DG, Stein A. Soil sampling strategies for spatial prediction by correlation with auxiliary maps. Australian Journal of Soil Research, 2003, 41: 1 403-1 422
- [53] 赵量, 赵玉国, 李德成, 孙孝林, 张甘霖. 基于模糊集理论提取土壤-地形定量关系及制图应用. 土壤学报, 2007, 44(6): 961-967
- [54] McKenzie NJ, Ryan PJ. Spatial prediction of soil properties using environment correlation. Geoderma, 1999, 89: 67-94
- [55] 姚荣江, 杨劲松, 姜龙. 黄河三角洲土壤盐分空间变异性与合理采样数研究. 水土保持学报, 2006, 20(6): 89-94
- [56] Minasny B, McBratney AB. The variance quadtree algorithm: Use for spatial sampling design. Computers & Geoscience, 2007, 33: 383-392
- [57] 姚荣江, 杨劲松, 赵秀芳, 李晓明, 刘梅先. 滩涂土壤电磁感应仪与方差四叉树法采样布局研究. 农业机械学报, 2010, 41(7): 174-180
- [58] Minasny B, McBratney AB. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computers & Geoscience, 2006, 32: 1 378-1 388
- [59] Carre F, MacBratney AB, Minasny B. Estimation and potential improvement of the quality of legacy soil samples for digital soil mapping. Geoderma, 2007, 141: 1-14
- [60] 姜成晟. 空间样点代表性评价与计算方法 (博士学位论文). 北京: 中国科学院地理科学与资源研究所, 2009
- [61] van Groenigen JW. The influence of variogram parameters on optimal sampling schemes for mapping by kriging. Geoderma, 2000, 97: 223-236
- [62] Dobermann A, Simbahan GC. Methodology for using secondary information in sampling optimisation for making fine-resolution maps of soil organic carbon // Lagacherie P, McBratney AB, Voltz M. Digital Soil Mapping –An Introductory Perspective. New York: Elsevier, 2007: 167-181
- [63] 姜成晟, 王劲峰, 曹志冬. 地理空间抽样理论研究综述. 地理学报, 2009, 64(3): 368-380
- [64] van Groenigen JW, Pieters G, Stein A. Optimizing spatial sampling for multivariate contamination in urban areas. Environmetrics, 2000, 11: 227-244
- [65] Debba P, van Ruitenbeek FJA, van der Meer FD, Carranza EJM, Stein A. Optimal field sampling for targeting minerals using hyperspectral data. Remote Sensing of Environment, 2005: 373-386
- [66] Brus DJ, Heuvelink JBM. Optimization of sample patterns for universal kriging of environmental variables. Geoderma, 2007, 138: 86-95
- [67] Goovaerts P. Geostatistics in soil science: State-of-the-art and perspectives. Geoderma, 1999, 89: 1-45

Sample-based Digital Soil Mapping Methods and Related Sampling Schemes

ZHANG Shu-jie^{1,2}, ZHU A-xing^{1,3}, LIU Jing³, YANG Lin¹

(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 University of Chinese Academy of Sciences, Beijing 100049, China;

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

Abstract: Traditional soil mapping methods, relating soil survey profiles with soil type maps under some criterions, based on soil surveyors' experience and manual delineation process are time-consuming and cost-prohibitive. The digital soil mapping (DSM), which takes advantage of the advancement in spatial information processing techniques and the increasing availability of high resolution spatial data, is feasible in providing soil distribution information at the required level of accuracy and spatial details. Field samples are the most important data sources for DSM. There are three DSM methods based on samples which are the ones based on spatial autocorrelation, the ones based on relationship between soil properties and environment co-varieties as well as the hybrid ones based on both spatial auto-correlation and the relationship between soil properties and environment co-variates. In order to capture the global representativeness of samples for characterizing the spatial variation of soil property, these kinds of methods all have serious requirement for the size, distribution and typicality of field samples. So, sampling design is an important process for DSM. Schemes for sampling design can be grouped into two categories based on whether the existing samples can be integrated in the sampling process: simple sampling scheme and integrative sampling scheme. This paper provides an overall review over these aspects of digital soil mapping.

Key words: Digital soil property mapping, Global representativeness, Sampling scheme, Integrative sampling scheme