

DOI: 10.11766/trxb201903180031

黄思华, 濮励杰, 解雪峰, 朱明, 阚博颖, 谭言飞. 面向数字土壤制图的土壤采样设计研究进展与展望[J]. 土壤学报, 2020, 57(2): 259–272.

HUANG Sihua, PU Lijie, XIE Xuefeng, ZHU Ming, KAN Boying, TAN Yanfei. Review and Outlook of Designing of Soil Sampling for Digital Soil Mapping[J]. Acta Pedologica Sinica, 2020, 57(2): 259–272.

面向数字土壤制图的土壤采样设计研究进展与展望*

黄思华^{1, 2}, 濮励杰^{1, 2†}, 解雪峰³, 朱明^{1, 2}, 阚博颖^{1, 2}, 谭言飞^{1, 2}

(1. 南京大学地理与海洋科学学院, 南京 210023; 2. 自然资源部海岸带开发与保护重点实验室, 南京 210023; 3. 浙江师范大学地理与环境科学学院, 浙江金华 321004)

摘要: 全球化土壤环境问题的出现对基础输入数据的精度、尺度和时序提出了更高要求, 面向数字土壤制图的土壤采样研究得到了快速发展。首先利用文献计量学的方法对国内外土壤采样研究学科分布和研究热点变化; 随后重点梳理了国内外土壤采样研究的文献, 根据不同的土壤调查目的、调查区历史采样点将土壤采样设计分为: 土壤全面采样设计、土壤补充采样设计、土壤验证采样设计和土壤监测采样设计; 最后介绍了基于样点的推理制图方法。在此基础上, 对未来在多尺度的土壤采样设计、土壤-环境因子关系的新型假设和采样设计中现实问题的量化等方面进行了展望, 旨在为数字土壤调查工作的开展提供参考依据。

关键词: 数字土壤制图; 土壤调查; 采样策略; 土壤-环境关系

中图分类号: S159.3 **文献标志码:** A

Review and Outlook of Designing of Soil Sampling for Digital Soil Mapping

HUANG Sihua^{1, 2}, PU Lijie^{1, 2†}, XIE Xuefeng³, ZHU Ming^{1, 2}, KAN Boying^{1, 2}, TAN Yanfei^{1, 2}

(1. School of Geography and Ocean Science, Nanjing University, Nanjing 210023, China; 2. Key Laboratory of Coastal Zone Exploitation and Protection, Ministry of Natural Resources, Nanjing 210023, China; 3. College of Geography and Environmental Science, Zhejiang Normal University, Jinhua, Zhejiang 321004, China)

Abstract: The appearance of soil environmental problems, such as pollution and degradation, has stimulated researches on hydro-ecological simulation, soil resource management, soil carbon and nitrogen monitoring, etc., thus putting forward higher requirements on basic input data, like types and attributes of soils, in accuracy, scale and timing sequence and causing rapid development of the research on soil sampling oriented towards digital soil mapping. In this study, the bibliometric method was applied to quantitatively analyze variation of the researches at home and abroad in distribution of disciplines hotspot during the recent four decades since 1980. Based on collation and review of the literature, summarization was performed of methods widely used nowadays in soil sampling and speculative mapping, and discussions conducted about future trends of the research on

* 国家自然科学基金项目(41871083, 41230751) 江苏省国土资源科技项目(2018018) 资助 Supported by the National Natural Science Foundation of China (Nos. 41871083, 41230751), Land and Resources Technology Research Program of Jiangsu Province (No. 2018018)

† 通讯作者 Corresponding author, E-mail: ljpu@nju.edu.cn

作者简介: 黄思华(1995—), 女, 广西桂林人, 博士研究生, 主要研究领域为土地利用与环境效应。E-mail: huangsihua@smail.nju.edu.cn

收稿日期: 2019-03-18; 收到修改稿日期: 2019-05-31; 优先数字出版日期(www.cnki.net): 2019-07-18

sampling designing for digital soil mapping, in an attempt to provide a reference for development of digital soil survey. Results show: (1) Over the last four decades, hotspots of the research on soil sampling have been focused on theories, methods, techniques, means and strategies of soil sampling, soil management and utilization, digital soil mapping, etc., involving disciplines that have developed from a single field of soil science into a transdisciplinary research covering agronomy, engineering, soil science, environmental science and geography, etc., with theories, techniques and means turning from mere application of probability theory into combination of geostationary models, deep learning and knowledge mining, and focuses laid on application in ecological monitoring and protection, precision agriculture and polluted remediation. (2) Soil sampling designing is a process of selecting an appropriate sampling method to meet the specific goal of a soil survey based on certain prior knowledge. According to the purposes of a soil survey and soil sampling history of the surveyed area, soil sampling designs can be divided into four categories, i.e. comprehensive sampling, supplemental sampling, verification sampling and monitoring sampling. In regions lacking historical soil sampling data, comprehensive sampling can be implemented by appointing sampling sites randomly based on the designed sampling method to achieve uniform coverage of their geospatial space, or by laying out sampling sites with reference to environmental factors to realize coverage and optimization of feature spaces, while in regions rich in data, comprehensive sampling may better adopt model-based sampling methods. For supplemental sampling, sampling points should be laid out in locations low in similarity of environmental factors, or high in uncertainty of predictive mapping or both. For validation sampling, independent sampling points should be arranged in line with the sampling design for better validation effect. And monitoring sampling could be designed into spatial sampling and temporal sampling or both with sampling sites laid out based on the design and the model in combination. And (3) soil mapping is a process of realizing point-plane expansion with the aid of mathematical methods or spatial models based on soil-environment relationship and spatial autocorrelation of soil attributes. In soil mapping, cartographers should pay attention to adoption of environmental factors and efficiency of the method utilizing the information of variables of the sampling sites. Hereby, cartographers should choose a corresponding inference model to implement spatial expression of soil information. So studies in future should be oriented towards application and theory, like designing of multi-scale soil sampling, new hypothesis of soil-environment relationship hypothesis and quantification of realistic problems in soil sampling designing.

Key words: Digital soil mapping; Soil survey; Sampling strategy; Soil-environment relationship

土壤调查是获取土壤属性特征和时空演变信息的有效方式^[1]。传统土壤调查服务于农业生产和管理,土壤专家凭借土壤知识及主观判断在野外采样,存在周期长、成本高、过程复杂和主观性等缺点^[2]。21世纪初,基于地理信息系统、地表数据获取技术和数据挖掘技术的数字土壤制图(Digital Soil Mapping, DSM)逐渐兴起,成为高效表达土壤空间分布的技术方法,为全球化研究、生态水文动态模拟、土壤资源管理、可持续土地利用提供大尺度、高精度土壤信息^[1-2]。

土壤采样通过选择代表性样点为数字土壤制图提供数据源。土壤采样往往结合统计推断、模型模拟和数字制图形成完整映射链,科学的采样设计能有效避免后续统计推断问题^[1]。基于设计的采样方法受传统抽样设计中概率统计理论的影响,认为土壤属性的空间变异具有随机性,样本的选择基于给

定的误差和概率,主要包括简单随机采样、系统采样和分层随机采样等^[3-4]。然而,土壤属性变化在地理空间中呈现空间自相关性,属于地统计学研究对象,由此,以地统计学理论为基础形成了基于模型的采样方法,主要工具包括协方差函数和变异函数,结合克里金插值方法,依据土壤的空间变异性和自相关特性来获取全局代表性样点^[5];近年来,土壤采样研究开始挖掘土壤本身的形成、发生以及与环境协变量之间的协同变化关系,在土壤-景观模型理论的基础上利用环境因子辅助采样,如基于专家知识采样、基于环境因子分层的拉丁超立方体采样、基于环境因子相似性的多等级代表性采样、基于环境因子变化程度的方差四叉树采样和基于土壤-环境因子关系的响应表面采样等^[2, 6-9]。土壤推理制图反映了土壤空间分布特征和规律,主要利用土壤-环境因子关系和土壤属性空间自相关性选择相应的

数学方法或空间模型实现点面拓展^[10]。土壤采样和推理制图相互联系，采样点质量是制约制图精度的关键因素^[11]，制图精度常被用于控制所需的样本量，而样点的布设规则直接影响推理模型的选择^[12]。

土壤调查的历史采样点由于数量有限、分布不符合某种规则、典型性不够或缺乏精确的地理参考，难以满足数字土壤制图的要求，往往需要重新采样获取额外样点^[13]。实际土壤调查过程中，土壤采样造成的误差远超过样品处理、实验室数据分析相关的误差，很大程度上决定了制图精度，同时采样工作受成本、时间和精度的限制。采样设计实际上是权衡精度和成本的过程，一方面可以通过增加样点数量提高精度，同时又要通过控制代表性样点的数量降低采样成本^[14]。本文据此梳理了国内外学者关于土壤采样的研究，系统总结了当前常用的几种土壤采样策略和推理制图的方法，并进一步讨论土壤采样设计的未来研究趋势。

1 国内外土壤采样研究热点变化

基于中国知网数据库和 Web of Science 核心数据库，分别以“土壤采样”、“土壤采样优化”、“土壤样点”和“soil sampling design”、“soil sampling strategy”、“soil sampling optimization”为关键词检索国内外土壤采样研究公开发表文献，去除重复和无关条目，统计文献发表数量随时间变化特征（如表 1）。结果表明 1980 年至 2018 年间共发表了 458、8 923 篇中、英文论文；1990 年初系统开展土壤采样的研究，这一时期主要得益于 GIS、遥感技术的

支持和地统计学的应用，21 世纪以来伴随全球数字土壤制图的兴起，土壤采样研究快速发展；国内在该领域的研究起步较晚，早期将地统计学应用于采样数量和采样密度的研究中，近年来发展迅速，主要集中在耕地质量监测样点、样带的布设和基于土壤-景观模型的采样方法的应用，其次作为数字土壤制图的子研究，国内学者在国际期刊发表了大量新型采样方法和策略的相关成果。

以文献为数据源，通过 CiteSpace（版本 5.3.R4.SE）软件^[15]定量分析土壤采样研究学科和热点的变化并进行可视化展示。图 1 为研究学科、主题和关键词的聚类结果，节点代表分析对象，越大表示研究热度越高，节点间的连线表示两者之间具有相关性。从图中可以看出，国外对土壤采样的研究从农学、工程学、土壤科学、环境科学和地理学等多学科交叉的角度进行，运用遥感、物理、化学分析和计算机技术；国内主要从土壤科学的角度研究，关键词包括空间变异、土壤养分、耕地质量、重金属和地统计学等。表 2 展示了不同时期土壤采样的研究热点，国外研究主要涉及土壤采样方法和技术手段、采样策略、土壤管理应用、数字土壤制图等方面，而国内则主要应用地统计学方法探讨土壤特性的空间分异，以此进行采样设计。从整体来看，土壤采样的研究从土壤科学的单一领域走向多学科交叉研究，在方法技术手段上从概率理论的应用走向对地统计学模型、深度学习和知识挖掘的算法结合，应用研究上也重点关注生态监测与保护、精准农业、污染修复等全球化问题。

表 1 土壤采样研究论文发表数量

Table 1 Number of papers published on soil sampling

时期	1980-	1985-	1990-	1995-	2000-	2005-	2010-	2015-	合计
Period	1984	1989	1994	1999	2004	2009	2014	2018	Total
英文 English	7	5	278	714	1 078	1 645	2 380	2 816	8 923
中文 Chinese	1	1	3	2	13	76	206	156	458

2 土壤采样设计方法

采样设计中合理利用先验知识可以提高样点全局代表性、降低采样时间和成本。先验知识来自辅

助地图、专家知识、历史土壤图和历史采样点数据。其中，辅助地图包括数字地形图和遥感影像，从中提取的全局环境因子包括与土壤形成相关的环境因子以及其他可能影响土壤属性空间变化的因素（如表 3），作为辅助数据可直接应用于后续建模。专家

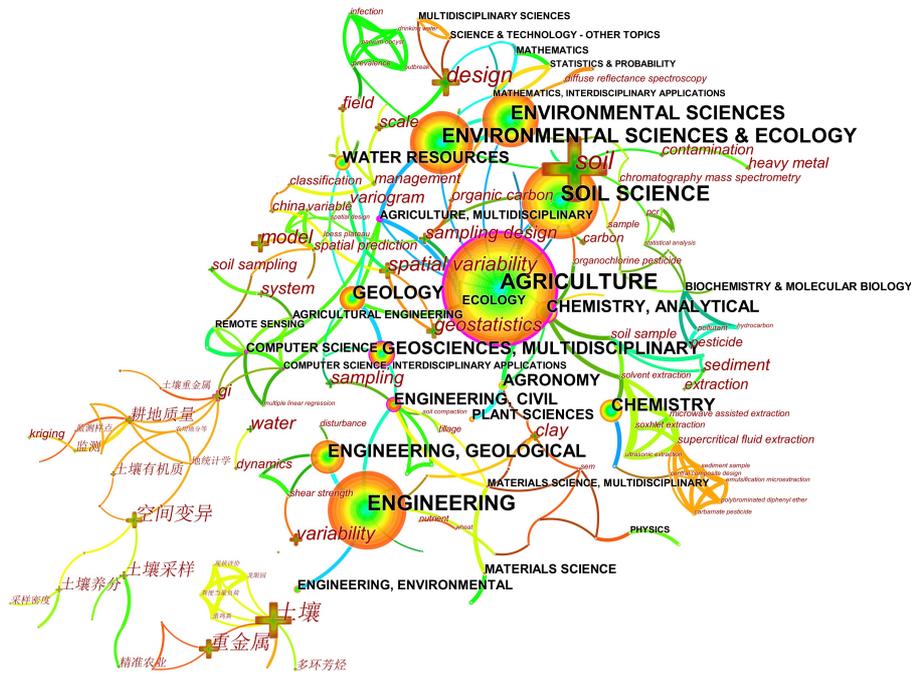


图 1 土壤采样研究学科、主题和关键词的共线分布示意图

Fig. 1 Sketch of co-occurrence of disciplines, themes and key words in research on soil sampling

表 2 不同时期土壤采样研究关键词

Table 2 Key words of the soil sampling research relative to period

研究方向 Fields	研究时期 Period	研究关键词 Key words
土壤采样理论、方法及技术手段 Theory; Method; Technical Methods	1990 至今	采样理论: 区域随机抽样理论; 普通克里金; 多元回归技术; 个体预测土壤制图; 经典抽样理论
	1990 至今	采样方法: 变差函数模型; 响应表面方法; 合理的样本量; 嵌套; 定量关系; 空间预测; 聚类算法; 灵活的拉丁超立方体; 拉丁超立方体; 空间覆盖采样; 随机抽样; 基于设计的空间采样; 迭代主成分分析; 协变量选择; 模拟退火
	2000—2016	采样尺度: 空间尺度; 采样密度; 采样网格
	1990 至今	采样技术手段: 土壤取样; 信息提取; 遥感; R 软件
土壤采样设计 Sampling design	2010 至今	补充采样设计: 不确定性导向抽样方法; 个体预测土壤制图
	1990—2017	土壤质量监测; 土壤生态功能监测网络; 精确监测; 土壤水分监测; 土壤监测网; 监测森林土壤特性; 环境质量等级划分; 耕地质量监测
管理应用 Management	1990—2014	土壤修复调查: 生态系统修复; 盐渍土; 土壤质量管理; 土壤侵蚀; 区域气候变化; 土壤形态; 土体刚度
	1995—2014	土壤养分: 有效氮含量; 空间变异特征; 土壤碳储量; 分布特征; 土地利用模式; 有机碳库动态; 土壤有机质
	1995—2016	空间变异: 表层土壤; 土壤盐分分布; 土壤水分; 地统计分析; 尺度依赖关系
制图 Mapping	1995 至今	深层土壤: 土体保存; 土体结构; 运输技术的研究; 采样深度
	2010 至今	全球数字土壤制图

表 3 土壤采样设计中常用的环境协变量指标

Table 3 Environmental covariate indices commonly used in soil sampling design

因子 Factors	具体指标 Indicators	尺度 Scale	获取方式 Method of obtaining	文献 Refs
母质 Parent material	地质图	①	历史地质图; USGS geological map	
	风化图	①	历史地貌图	[25-26]
	地貌图	②	历史地貌图	
气候 Climate	年均温度、年积温	②	遥感: AVHRR、GOES、TIROS	
	年均降水	②	AVHRR 数据; 气象监测站: 气象数据插值	[27-28]
	蒸散发量	②	能量平衡方程计算 遥感: Landsat TM 数据、AVHRR 数据	
地形 Terrain	高程、坡度、坡向、平面/剖面曲率	①	DEM 计算	[14]
	地形湿度指数	①	DEM 计算	[29-30]
	复合地形指数	①	DEM 计算	[31]
生物 Biological	模糊坡位	③	对空间渐变的坡位(坡肩、背坡)模糊表达	[32]
	土地覆被	①	全球土地覆盖产品	[30-33]
	增强植被指数	①	可见光近红外范围内的高光谱图像、微波图像; 波段计算 遥感: Landsat 数据; 微波图像	[34]
时间 Time	土壤调节植被指数	①	从可见光和红外反射中获得的, 并且最近通过微波图像得到了增强	[30]
	生物量和产量: 归一化植被指数 NDVI	①	可见光近红外范围内的高光谱图像、微波图像; $NDVI = \frac{Bands(4-3)}{(4+3)}$	[28]
人类活动 Human activity	年龄	①	第四纪地质图; 样点数据年龄; 遥感数据	[35]
	农作物轮作类型	①	遥感影像的作物解译	[36]
	净初级生产力 NPP	①	MODIS 数据	[27]
近地传感 Near-ground sensing	土地利用类型	①	遥感影像解译	[28]
	电导率/多光谱数据	①	近地传感器	[37]
其他 Other	近红外反射率 NIR	①	Chinese vis-NIR soil spectral library (CSSL)	[38]
	空间坐标	①	栅格单元中心位置坐标	[27, 39]
	物候信息: 季节长度、季节幅度	①	基于卫星光谱数据测量植被指数时间序列, 拟合局部多项式函数 指示季节性气候周期	[40]

注: ①任何尺度 All, ②全球/国家/区域 Global/National/Regional, ③平缓小流域区域 Gentle small watershed area

知识以描述性知识为主, 包括土壤专家对土壤类型和属性真假判断以及对土壤-景观环境关系的描述。这部分知识的应用往往借助布尔逻辑理论和感知计算理论将描述性单词映射至制图单元, 获取典型土

壤类型/属性对应的环境知识并集成至土壤预测模型, 目前已有的方法包括专家知识系统 (Expert Systems) [16] 和模糊隶属函数 [17]。历史土壤图是土壤调查者对区域土壤、景观、地形、自然环境等的综

合认识,利用神经网络^[18]、决策树^[19]、随机森林^[19]和贝叶斯^[20]等方法获取其中包含的土壤-环境关系知识可以指导采样。历史遗留土壤样点是不同时段、不同目的下土壤调查的成果,大多存在分布不遵循统计标准、典型性不够和缺乏精确的地理参考等问题^[13, 21]。评估历史采样点数据的可用性、时效性、代表性和信息完整性,是后续历史采样点参与土壤采样和制图的前提。对此,张忠启等^[22]提出了揭示特定时段土壤有机碳变化所需采样数量的方法,解决了不同历史时段土壤样点的利用问题。Carré等^[21]和 Stumpf 等^[23]通过拉丁超立方体确定历史采样点在环境协变量超立方体的占有率,从而评估历史采样点的质量,指导布设补充采样的位置,实现了历史土壤样本的整合。An 等^[24]利用环境协变量聚类来近似代替土壤变化类型,选择位于环境协变量聚类质心的历史采样点为代表性样品参与采样和制图。

根据土壤调查者的采样目的将土壤采样设计计划分为:基于不同先验知识对区域的全面采样,基于历史采样点的补充采样,用于评价制图质量的验证采样和反映土壤空间分布实时信息的监测采样(如图2)。

2.1 土壤全面采样设计

2.1.1 无历史采样点区域的全面采样 无历史采样点区域进行全面采样,方法的选择取决于先验知识。大尺度土壤调查对制图精度的要求较低,简单随机采样和网格采样无疑是快速而实用的方法^[41]。中小尺度土壤调查对制图精度要求较高,需要借助环境因子辅助采样提高采样点的全局代表性。特征空间是由一组环境变量范围限定形成的虚拟空间^[31]。环境因子辅助采样的主要原则是选择样点覆盖或者优化这个虚拟空间。为了实现特征空间的全局覆盖,样点常被布设于能够完整代表环境空间差异性的区域。在地理空间上,这些区域可以通过某种典型的环境梯度组合切割分层,如 Mckenzie 等^[42]结合地形地质参数和植被数据对研究区分层后随机选择样点; Hengl 等^[31]将环境协变量分布频率作为分层依据,通过等距设计样点实现特征空间的均匀扩展; Minasny 等^[8]基于方差四叉树法根据环境因子的变化程度对研究区分层采样。从优化特征空间的角度,将环境协变量进行聚类或分层,在特定环境组合区域布设采样点,形成的特征空间能够最大程度代表现实空间的特征属性,常用方法包括模糊均值聚类采样^[43]、K 均值聚类采样^[39]、多等级代表性采样^[44]、

拉丁超立方体采样^[7]、条件拉丁超立方体采样^[14]等。此外,样点在地理空间的良好分布也是采样设计的重要原则之一。平衡采样是优化地理空间分布引导的空间覆盖采样,利用目标变量与环境协变量之间的线性关系抽样,样点的布设需要满足样点环境协变量均值等于总体均值^[34]。实现平衡采样的方法包括 Horvitz-Thompson 估计器、排斥程序、枚举法、局部关键算法、空间相关泊松采样、立方体方法和广义随机-曲面分层抽样法等^[45-47]。双重平衡空间采样在平衡采样的基础上进行了优化,可以实现空间平衡良好分布的同时避免选择相邻单元^[34, 48]。响应表面采样是基于模型预测的采样,在目标变量和环境协变量之间的关系可以拟合线性或二次回归模型的假设下,通过布设样点优化模型参数,降低模型残差的空间自相关效应^[6]。响应表面采样最初利用土壤电导率数据估算土壤盐度,随后开发的 ESAP 软件允许输入环境遥感数据,为大尺度土壤采样提供可能性,但只能生成 6、12 或 20 个样本^[6]。

2.1.2 有历史采样点区域的全面采样 有历史采样点区域进行全面采样,方法的选择取决于历史采样点中隐含知识,包括土壤变异信息和土壤-环境知识。经典统计学 Cochran 公式是目前计算区域最优采样数的常用方法^[41],该公式的参数涉及置信水平、精度和先验样本的变异系数。在此基础上,最适分配法计算确定分层采样中每层最佳采样数^[49]。经典统计学计算方法可以对区域进行整体大致趋势和特征的研究,但无法决定样点的空间位置,同时也忽略了区域土壤特征的空间变异性^[50]。基于模型的采样方法能够弥补这些不足,该方法利用大量历史样点建立可靠的空间变异模型,进而拟合克里金插值模型绘制土壤图。伴随以上过程产生的空间变异模型相关参数和未知点插值预测误差可以指导现阶段土壤采样。对此,国内学者展开了一系列合理采样数的研究,如张志霞等^[50]结合半方差函数模型和克里金插值结果交叉验证,综合精度评价指标 RMSE、 R^2 和空间结构性指标最小化时确定合理采样数;赵业婷等^[51]通过对比普通克里金法和协同克里金法的合理采样数和优化采样数量的适用性,认为协同克里金方法能够更好地优化采样数量,提供更多局部变异信息。地统计学模型除了应用于确定合理采样数,还能指导采样点的空间优化布局,如空间变异模型的变程值可以反映网格采样的间隔^[50]。一些研

究中采样点的布设以减少预测误差为目标，如陈天恩等^[52]根据克里金插值绘制估计值方差等值线，在估计误差的方差大于给定阈值的区域加密采样点；Li等^[53]利用方差四叉树算法结合半方差函数对目标变量插值的方差较大的区域不断四分得到等方差的区层，对变异较大的区域增加采样密度。而历史采样点相对较少的区域，难以建立模型准确描述区域土壤属性的空间变异情况，通常将已有样点与环境因子进行相关性分析，选取典型的环境协变量组合，借助环境因子辅助采样实现样点布设^[23]。或基于对土壤景观关系的不同假设，通过建立多元回归函数、协同克里金、泛克里金、随机森林等模型作为目标函数，通过模拟退火算法优化生成最优空间布局的样点集^[5, 23]。

2.2 土壤补充采样设计

当历史采样点在数量、分布和典型性上无法达到制图要求时，需要设计典型补充样点去提高土壤信息空间表达精度，这样的采样设计称为补充采样。土壤补充采样的关键在于整合现有资源，挖掘遗留样本中包含的局部土壤-环境知识和空间属性变异信息。

2.2.1 基于环境因子相似性的补充采样设计 基于推理不确定性的思想来源于样点个体代表性预测土壤制图^[54]。在景观单元环境越相似其土壤性质越相似的假设下，通过已有样本与未采样点之间的环境相似性推测两者之间土壤性质的相似性，以此估计土壤属性预测的不确定性^[54]。刘京等^[55]提出了度量土壤属性推理不确定性的方法并成功应用于大尺度土壤制图。基于该理论，未采样点和采样点的环境越不相似，其土壤属性推测的不确定性越高，当未采样点的不确定性值高达一定程度，那么该点属性值无法通过采样点推理获得^[55]。基于环境相似性的补充采样设计以推理不确定信息分布图为指导，样点的布设考虑如何扩大可预测范围或降低推理结果的不确定性^[13]。张淑杰等^[13]提出了逐次、高效地设计补充样点方案，在不确定性较高的区域布设补充样点直至新样点集的空间代表范围覆盖整个研究区，具有确定补充样点的数量、位置和重要性次序等优点，但未考虑如何提高推理精度。以降低推理结果不确定性为目标的补充样点布设方案通过模拟退火算法实现，该方法不能提供补充样本的采样顺序，也没有明确最优补样数^[56]。启发式不确定性补

充采样考虑了以上两种原则，利用遍历算法和贪婪算法补充最少的样点扩展预测区域覆盖整个研究区，在不确定性值和面积较大的区域对环境因子聚类，在最大的环境类别上布设样点以减少土壤属性预测的不确定性^[26]。张磊等^[57]提出了一种基于分融策略的样点设计方法，能有效避免冗余样点，包括分化阶段和融合阶段，分化阶段分化推测可信度低的样点，增加样点来降低推测不确定性，融和阶段将环境条件过于相似的样点进行融合以降低冗余。

2.2.2 基于预测制图不确定性的补充采样设计

基于预测制图不确定性的补充采样目的是通过添加样点降低目标地理变量的整体空间预测不确定性。土壤空间插值模型和土壤-环境关系预测模型的预测制图误差可以指示空间预测的不确定性。在不同的克里金模型中，泛克里金模型在精度上更有优势，其方差包含趋势估计误差和空间插值误差的方差两个分量，能有效平衡特征空间和地理空间^[5]。相对于构建克里金模型对样本密度、统计假设等要求，随机森林模型预测方差是土壤制图的副产品，不需要额外的处理步骤，避免了统计复杂性，能够满足实际不确定性度量的需求^[5, 23]。补充采样的候补区是具有较大预测方差的区域，补充样点布设方法包括拉丁超立方体和模拟退火算法，在迭代过程中通过设定算法停止的阈值（足够小的预测误差）决定样本数量和位置，实现环境协变量空间与采样效果之间的平衡^[58]。

Li等^[59]提出基于属性域和空间域的不确定性补充采样方法，结合了上述两种思想。属性域是环境协变量的关系域，在属性域中补充采样借鉴了启发式不确定性补样，将不确定性由高至低排序选择补样点；空间域为空间自相关域，依据平均/最大克里金方差划分不确定性，利用模拟退火算法选择补样点，最终补样点为两种方案的并集。

2.3 土壤验证采样设计

对生产者 and 使用者而言预测制图的质量验证必不可少^[60-61]。部分基于模型（如克里金插值）或基于环境相似度的土壤制图会产生不确定性分布图等附加产品，可以指示制图结果的可靠性^[2, 55, 62]。目前常规的定量评价土壤制图精度的方法包括数据分裂、交叉验证和附加概率采样验证^[61]。前两种方法基于已有样点实现，数据分裂将校准样点依据一定比例（20%~30%）随机分为训练样点和验证样点，

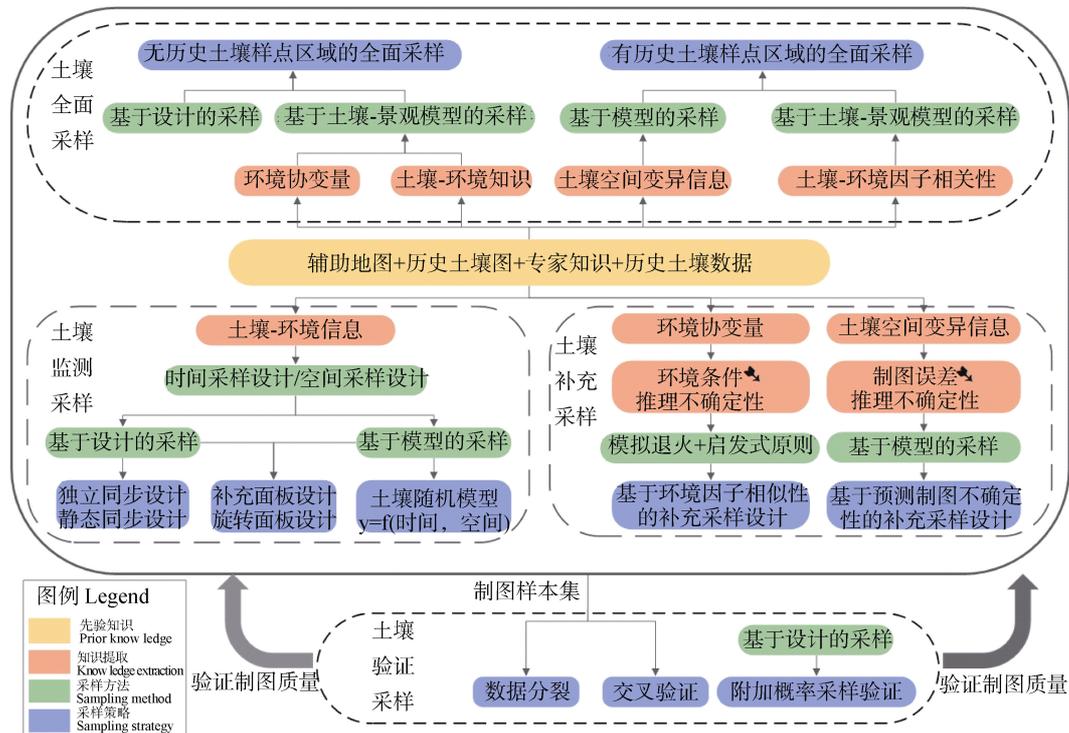


图2 基于先验知识的土壤采样方法选择决策树

Fig. 2 Soil sampling strategy selection and decision tree based on prior knowledge

训练样点推理制图后由验证样点评价制图精度^[50, 63]。交叉验证通过重复分割校准样点集去验证，本质上是数据分裂的迭代过程，相对于数据分裂更有效，包括留一交叉验证和多折交叉验证^[64]。目的采样获得的制图样点集，分割形式无法改变样点本身的偏向性，预测误差或分类错误率存在着空间自相关，验证精度高估了实际精度，因此，数据分裂和交叉验证难以实现无偏和有效的地图精度估计^[61, 65]。而附加概率采样验证通过概率抽样选择独立验证点与预测地图单元进行比较，不需要模型估算地图精度，避免对预测误差的空间自相关做出假设，能够有效地指示制图精度^[65]。Brus等^[61]提出利用基于设计的采样方法布设附加验证点，包括简单随机采样、分层简单随机采样、系统随机采样、聚类采样和两阶段随机采样。在保证验证质量的前提下，Gruijter等^[66]提供了上述不同采样方法所需的最小样本量的计算方法，通过抽样概率和参数估计的方差计算验证样点总数，并根据其特有的布设方式分配样点。一般而言，增加验证采样强度有利于提升预测质量，但并不能增加实际的制图精度，因此实际验证采样设计中需要权衡验证采样成本与验证质量的关系。

前人研究表明，综合实际操作性、误差估计及精度，可优先选择分层简单随机采样获取验证样点^[61]，而大尺度地图验证采样受成本和可达性的限制，选择聚类采样和两阶段随机采样更有效^[60]。有学者提出利用成本限制下的条件拉丁超立方体采样选取验证点^[67]，该方法虽然随机选取样点位置但并不属于概率抽样设计，难以保证对地图质量的无偏估计^[68]。对此，Yang等^[68]提出将逐点访问成本作为衡量变量加入简单随机采样、成比例概率采样和分层随机采样选择验证点。

2.4 土壤监测采样设计

目前，世界许多国家或区域已经建立了土壤监测网络，用于定期观测土壤肥力、土壤污染和土壤侵蚀等变化。监测站点的选取一方面基于历史采样点或者预设样点的空间变异性信息实现监测样点的再优化，如欧洲土壤监测系统应用 50 km×50 km 格网采样^[69]，国内耕地质量监测系统应用地统计学变异函数^[70]优化布设样点。另一方面，通过土壤或者景观环境信息实现监测样带的选择，如德国 BDF-SH 长期土壤监测计划通过景观单元、土壤类型和土地利用选择代表区^[71]，国内耕地质量监测系

统依据自然条件、利用水平和收益水平等因素组合确定耕地质量均值区域选择监测点^[72]。以上采样方法均从静态层面上选取样点，未考虑土壤属性的时空变异性，缺乏数学模型的统计推断和空间抽样理论的验证。

土壤属性特征具有时空变异性。近十年，国外土壤监测采样研究从时间、空间两个维度考虑，分为时间采样设计、空间采样设计和时空采样设计^[66]。基于概率抽样理论、地统计理论和时间序列分析，形成完全基于设计、完全基于模型和混合方法的土壤监测的样本布设方法^[66, 73]。完全基于设计的方法利用概率采样选择抽样单元和抽样时间，其统计参数通过概率采样所确定的包含概率推断，具有对时空平均值的无偏估计和量化抽样误差所导致的估计总量不确定性等优点^[74-75]。完全基于模型的方法通过历史数据建立随机模型来描述土壤属性的时空变化，利用模型推断参数和预测时空平均值^[66]。混合方法涉及了基于设计和模型的推理，对样本位置进行概率采样构建离散化空间均值的时间序列模型^[73-75]。土壤监测网中最佳采样数量需要满足两个目标：能代替土壤属性空间均值和空间变化，通过功率分析法和最小可检测差异法计算^[76]。Brus 等^[73]提出四种时空设计概念性例子：独立同步设计、静态同步设计、补充面板设计和旋转面板设计，对监测采样设计具有重要指导意义。

理论模型和算法确定的采样点在现场操作过程中受成本、时间和地形等限制，预设位置可能会落在无法访问区。针对这些现实问题，选择容易到达的替换位置采样是其中一个解决思路，如基于道路网不同距离设置采样尺度，寻找最佳的土壤采样布局^[77]；有学者尝试结合不同方法进行交替采样，对于初始采样方案中无法到达的点，通过环境因子聚类选择具有相似环境因子组合的代替点^[78]；此外，通过到达采样点所花费的时间或成本量化可达性限制，将其加入条件拉丁超立方体目标函数，这样选取的采样点既能够满足统计有效性又落入了可到达区域，但避开成本过高的采样点后其精度会有所下降^[14]；对于拉丁超立方体采样的样本量不确定性产生的计算需求量过大的问题，渐进拉丁超立方体采样可以生成一系列较小的子集进行多阶段或顺序抽样，可以有效改进相关分析的收敛性和采样结果的稳定性^[79]。灵活拉丁超立方体采样通过帕累托优化

对多目标同时进行优化，实现特征空间的覆盖以及成本的控制，适合大区域采样^[80]。

3 推理制图

土壤推理制图是以数学方法和空间分析为手段，利用土壤属性的空间自相关性和土壤-环境协变量关系，将点映射至面以体现土壤空间分布特征和规律的过程^[10]。尽管大尺度数字土壤制图更多地借助于土壤近地传感、土壤光谱和卫星遥感技术，但野外采样点仍是至关重要的数据源。土壤样点的数量和布设规则将影响土壤空间推理模型的选择^[1]。基于样点的土壤制图方法和采样设计的思路相似，大致可概括为利用土壤-环境因子关系、土壤属性空间自相关性推测区域土壤的空间分布^[81]。前者主要利用机器学习、数据挖掘和数学模型等方法挖掘土壤属性与环境协变量的关系知识作为制图依据，后者则利用给定的一组离散土壤样点建立目标属性的空间自相关模型，通过空间插值模型，如趋势面分析、克里金插值、样条函数、反距离加权法和最邻近法等制图^[2, 82]。以上两者结合的制图法同时考虑了土壤属性的空间自相关特征和土壤与环境因子的关系，主要方法包括回归克里金插值、协同克里金插值和地理加权回归模型^[82-84]。

一般而言，制图精度会随着样点的数量增加而逐渐提高，在方法上，土壤-环境模型制图法相对普通克里金法和线性回归模型更有效，回归克里金法能有效结合土壤-环境模型法和空间插值的优势，制图精度优于普通克里金法^[84-86]。然而，一些方法对比研究发现，土壤预测制图效果并不完全取决于方法的精密和复杂性，也需要关注辅助环境因子的应用以及方法对样点变量信息的利用效率^[82]。土壤采样和推理制图是两个相互联系的过程，制图精度常被用于控制所需的样本量，而样点的布设规则是选择推理模型的关键因素^[12]。基于模型的采样设计的样点以模型参数估算方差和插值预测误差最小化为目标，所建立的空间自相关模型具有较小的预测不确定性，利用空间插值模型制图，或者在空间自相关基础上引入环境因子辅助推测制图具有较好的效果^[1]。环境因子辅助采样点相对于基于模型的采样点具有数量较少、代表性高且包含土壤-环境关系的特点，适合基于土壤-环境关系制图的方法^[2]。制图

者需要根据特定制图要求选择相应的土壤-环境关系表达模型推理制图。采样设计其中一个重要依据是土壤-环境因子间线性或非线性假设,土壤属性制图也通常利用这一关系推理制图,广泛使用的方法包括线性回归模型^[87]、随机森林^[82]、决策树^[85]和人工神经网络模型^[85]等。土壤类型制图则依据特定的土壤类型-环境组合知识推理制图,代表方法包括土壤-景观推理模型(Soil-Landscape Inference Model, SoLIM)^[88]和语义模型模糊推理模型^[12]。近年来,大尺度的土壤采样设计如多等级代表性采样、基于不确定性的补充采样等,样点布设基于环境越相似土壤属性越相似的假设,点面拓展方法涉及模糊隶属加权平均法和个体预测土壤制图法^[54, 89]。随着全球数字土壤制图工作的开展,对大尺度采样设计和推理制图提出挑战,加强环境协变量的应用,提升推理方法对样点变量信息的利用效率成为重要研究方向。

4 结论与展望

本文系统梳理了国内外关于土壤采样策略的研究,根据不同的土壤调查目的、调查区历史采样点将土壤采样分为:土壤全面采样设计、土壤补充采样设计、土壤验证采样设计和土壤监测采样设计。无历史采样点区域的全面采样通过基于设计的采样方法随机选择样本位置实现地理空间均匀覆盖,或通过环境因子辅助样本布设实现特征空间的覆盖和优化,有历史采样点区域的全面采样更适合基于模型的采样方法。补充采样的样点应布设于环境因子相似性较低、预测制图不确定性较高或两者兼备的位置。验证采样选择基于设计的方法获取独立验证点更可取。监测采样分为空间采样、时间采样和时空采样,将基于设计和基于模型的方法组合布设样点。基于样点的推理制图方法主要包括基于土壤-环境因子关系制图和基于土壤属性空间自相关性制图。制图者应关注环境因子的应用及方法对样点变量信息的利用效率,根据特定制图要求选择相应的推理模型完成土壤信息的空间表达。未来土壤采样研究的发展趋势应包括:

(1) 多尺度的土壤采样设计。目前土壤采样的研究尺度以流域、农场、田块为主,选取的调查区土地利用/覆被结构相对简单,对已提出的采样方法

具有良好的应用性。在国家以及全球尺度上,由于土壤变异情况、土地利用结构和人类活动更复杂,现有提出的大尺度采样一般是基于设计的方法,这样得到的结果公正且有效,但样点缺乏代表性且需要较多成本和时间^[69]。因此,为满足全球数字土壤制图和全球化研究的需要,探索大尺度基于土壤-景观模型的采样设计和基于预测不确定性的补充采样设计等新型采样方法的适用性,实现采样方法尺度扩张是未来重要的研究方向。

(2) 土壤-环境因子关系的新型假设。土壤采样并不是一个独立的过程,往往会结合统计推断、模型模拟和制图形成完整的映射链,这种映射关系基于一定的假设。在采样方法中,部分不依赖于这种假设,如概率采样,因为它对总体参数的估计无偏。其他采样方法基于的假设大致可分为两类,一是基于克里金插值模型后具有独立残差的假设,如基于模型和模拟退火算法的采样方法;其次就是基于土壤-环境因子关系,如响应表面采样和平衡采样,认为土壤属性与环境因子之间存在线性关系^[90]。由于土壤-环境关系是一个黑箱,难以通过简单的线性关系进行表达^[2]。通过机器学习和数学挖掘等方法来获取新型土壤属性与环境因子关系,指导建立更加有效的采样方法是未来重要的研究方向。

(3) 采样设计中量化现实问题。现有土壤采样设计的目标主要针对单一目标土壤属性,并不适用于大尺度土壤调查,如何通过一次采样满足多个土壤属性变量空间分布已有初步研究,有学者通过建立线性模型同时最小化多个土壤变量的平均克里金方差得到最优样本量和样本分布模式^[91]。此外,多数采样方案确定的样本位置在野外操作过程中受成本、时间及可达性限制而无法准确获取该样点。有学者尝试在拉丁超立方体采样和概率采样中加入成本限制来解决这一问题,但导致采样精度降低^[67]。因此,如何在采样设计中满足多个目标土壤属性分布以及权衡提升野外可操作性带来了精度降低等现实需求,还需要深入研究。

参考文献 (References)

- [1] McBratney A B, Odeh I O A, Bishop T F A, et al. An overview of pedometric techniques for use in soil survey. *Geoderma*, 2000, 97 (3/4): 293—327.
- [2] Zhu A X, Yang L, Fan N Q, et al. The review and outlook of digital soil mapping. *Progress in Geography*, 2018,

- 37 (1): 66—78. [朱阿兴, 杨琳, 樊乃卿, 等. 数字土壤制图研究综述与展望. 地理科学进展, 2018, 37 (1): 66—78.]
- [3] Wallenius K, Niemi R M, Rita H. Using stratified sampling based on pre-characterisation of samples in soil microbiological studies. *Applied Soil Ecology*, 2011, 51 (1): 111—113.
- [4] Zhang J, Zhang C. Sampling and sampling strategies for environmental analysis. *International Journal of Environmental Analytical Chemistry*, 2012, 92 (4): 466—478.
- [5] Brus D J, Heuvelink G B M. Optimization of sample patterns for universal kriging of environmental variables. *Geoderma*, 2007, 138 (1/2): 86—95.
- [6] Corwin D L, Lesch S M, Segal E, et al. Comparison of sampling strategies for characterizing spatial variability with apparent soil electrical conductivity directed soil sampling. *Journal of Environmental & Engineering Geophysics*, 2010, 15 (3): 147—162.
- [7] Minasny B, McBratney A B. Latin hypercube sampling as a tool for digital soil mapping. *Developments in Soil Science*, 2007, 31: 153—165, 606.
- [8] Minasny B, McBratney A B, Walvoort D J J. The variance quadtree algorithm: Use for spatial sampling design. *Computers & Geosciences*, 2007, 33 (3): 383—392.
- [9] Yang L, Zhu A X, Qin C Z, et al. A soil sampling method based on representativeness grade of sampling points. *Acta Pedologica Sinica*, 2011, 48 (5): 938—946. [杨琳, 朱阿兴, 秦承志, 等. 一种基于样点代表性等级的土壤采样设计方法. 土壤学报, 2011, 48(5): 938—946.]
- [10] Lagacherie P, McBratney A. Spatial soil information systems and spatial soil inference systems: Perspectives for digital soil mapping. *Developments in Soil Science*, 2006, 31: 3—22.
- [11] Vaysse K, Lagacherie P. Evaluating Digital Soil Mapping approaches for mapping global soil map soil properties from legacy data in Languedoc-Roussillon (France). *Geoderma Regional*, 2015, 4: 20—30.
- [12] Yang L. Relationships between sample selection and mapping accuracy under purposive sampling: A case study in digital soil mapping. Beijing: Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, 2009. [杨琳. 目的性采样下样本设计与制图精度的关系研究: 以数字土壤制图为例. 北京: 中国科学院地理科学与资源研究所, 2009.]
- [13] Zhang S J, Zhu A X, Liu J, et al. An integrative sampling scheme for digital soil mapping. *Progress in Geography*, 2012, 31 (10): 1318—1325. [张淑杰, 朱阿兴, 刘京, 等. 整合已有土壤样点的数字土壤制图补样方案. 地理科学进展, 2012, 31 (10): 1318—1325.]
- [14] Godinho Silva S H, Owens P R, Silva B M, et al. Evaluation of conditioned Latin hypercube sampling as a support for soil mapping and spatial variability of soil properties. *Soil Science Society of America Journal*, 2015, 79 (2): 603—611.
- [15] Chen C M. Science mapping: A systematic review of the literature. *Journal of Data and Information Science*, 2017, 2 (2): 1—40.
- [16] de Menezes M D, Silva S H G, Owens P R, et al. Digital soil mapping approach based on fuzzy logic and field expert knowledge. *Ciência e Agrotecnologia*, 2013, 37 (4): 287—298.
- [17] Zhu A X, Yang L, Li B L, et al. Construction of membership functions for predictive soil mapping under fuzzy logic. *Geoderma*, 2010, 155 (3/4): 164—174.
- [18] Li Q, Cheng J, Hu Y, et al. Spatial interpolation of soil nutrients based on bp neural network. *Agricultural Science & Technology*, 2014, 15 (3): 506.
- [19] Subburayalu S K, Slater B K. Soil series mapping by knowledge discovery from an ohio county soil map. *Soil Science Society of America Journal*, 2013, 77 (4): 1254—1268.
- [20] Brus D J, Bogaert P, Heuvelink G B M. Bayesian maximum entropy prediction of soil categories using a traditional soil map as soft information. *European Journal of Soil Science*, 2008, 59 (2): 166—177.
- [21] Carré F, McBratney A B, Minasny B. Estimation and potential improvement of the quality of legacy soil samples for digital soil mapping. *Geoderma*, 2007, 141 (1/2): 1—14.
- [22] Zhang Z Q, Yu F Z, Yu D S, et al. Temporal variability of soil organic carbon and rational sample size in red soil region. *Acta Pedologica Sinica*, 2016, 53(4): 891—900. [张忠启, 于法展, 于东升, 等. 红壤区土壤有机碳时间变异及合理采样点数量研究. 土壤学报, 2016, 53(4): 891—900.]
- [23] Stumpf F, Schmidt K, Behrens T, et al. Incorporating limited field operability and legacy soil samples in a hypercube sampling design for digital soil mapping. *Journal of Plant Nutrition and Soil Science*, 2016, 179 (4): 499—509.
- [24] An Y M, Yang L, Zhu A X, et al. Identification of representative samples from existing samples for digital soil mapping. *Geoderma*, 2018, 311: 109—119.
- [25] McBratney A B, Mendonça Santos M, Minasny B. On digital soil mapping. *Geoderma*, 2003, 117(1/2): 3—52.
- [26] Zhang S J, Zhu A X, Liu J, et al. An heuristic uncertainty directed field sampling design for digital soil mapping. *Geoderma*, 2016, 267: 123—136.
- [27] de Brogniez D, Ballabio C, Stevens A, et al. A map of the topsoil organic carbon content of Europe generated by a generalized additive model. *European Journal of Soil Science*, 2015, 66 (1): 121—134.
- [28] Yang R M, Zhang G L, Liu F, et al. Comparison of boosted regression tree and random forest models for

- mapping topsoil organic carbon concentration in an alpine ecosystem. *Ecological Indicators*, 2016, 60: 870—878.
- [29] Ashtekar J M, Owens P R. Remembering knowledge: An expert knowledge based approach to digital soil mapping. *Soil Horizons*, 2013, 54 (5): 1—6.
- [30] Brungard C W, Boettinger J L. Conditioned Latin hypercube sampling: Optimal sample size for digital soil mapping of arid rangelands in Utah, USA//*Digital Soil Mapping*. Dordrecht: Springer Netherlands, 2010: 67—75.
- [31] Hengl T, Rossiter D G, Stein A. Soil sampling strategies for spatial prediction by correlation with auxiliary maps. *Australian Journal of Soil Research*, 2003, 41 (8): 1403—1422.
- [32] Qin C Z, Lu Y J, Qiu W L, et al. Application of fuzzy slope positions in predicting spatial distribution of soil property at finer scale. *Geographical Research*, 2010, 29 (9): 1706—1714. [秦承志, 卢岩君, 邱维理, 等. 模糊坡位信息在精细土壤属性空间推测中的应用. *地理研究*, 2010, 29 (9): 1706—1714.]
- [33] Hengl T, Mendes de Jesus J, Heuvelink G B M, et al. Soilgrids250m: Global gridded soil information based on machine learning. *PLoS One*, 2017, 12 (2): e0169748.
- [34] Brus D J. Balanced sampling: A versatile sampling approach for statistical soil surveys. *Geoderma*, 2015, 253/254: 111—121.
- [35] Noller J S. Applying geochronology in predictive digital mapping of soils//*Digital Soil Mapping*: Dordrecht: Springer Netherlands, 2010: 43—53.
- [36] Song M, Yang L, Zhu A X, et al. Mapping soil organic matter in farming areas with crop rotation. *Chinese Journal of Soil Science*, 2017, 48 (4): 778—785. [宋敏, 杨琳, 朱阿兴, 等. 轮作模式在农耕地土壤有机质推测制图中的应用. *土壤通报*, 2017, 48 (4): 778—785.]
- [37] Viscarra Rossel R A, McBratney A B, Minasny B. Proximal soil sensing. Dordrecht: Springer Netherlands, 2010.
- [38] Shi Z, Ji W, Viscarra Rossel R A, et al. Prediction of soil organic matter using a spatially constrained local partial least squares regression and the Chinese vis-NIR spectral library. *European Journal of Soil Science*, 2015, 66 (4): 679—687.
- [39] Brus D J, de Grujter J J, van Groenigen J W. Designing spatial coverage samples using the k-means clustering algorithm//*Developments in Soil Science*. Amsterdam: Elsevier 2006: 183—192.
- [40] Poggio L, Gimona A. 3D mapping of soil texture in Scotland. *Geoderma Regional*, 2017, 9: 5—16.
- [41] Zhang S M, Xu M X, Zhang Z X, et al. Methods of sampling soil organic carbon in farmlands with different landform types on the Loess Plateau. *Journal of Natural Resources*, 2018, 33 (4): 634—643. [张圣民, 许明祥, 张志霞, 等. 黄土高原不同地貌类型区农田土壤有机碳采样布点方法研究. *自然资源学报*, 2018, 33 (4): 634—643.]
- [42] McKenzie N J, Ryan P J. Spatial prediction of soil properties using environmental correlation. *Geoderma*, 1999, 89 (1/2): 67—94.
- [43] Lin Q H, Li H, Luo W, et al. Optimal soil-sampling design for rubber tree management based on fuzzy clustering. *Forest Ecology & Management*, 2013, 308 (4): 214—222.
- [44] Yang L, Zhu A X, Zhao Y G, et al. Regional soil mapping using multi-grade representative sampling and a fuzzy membership-based mapping approach. *Pedosphere*, 2017, 27 (2): 344—357.
- [45] Deville J C. Efficient balanced sampling: The cube method. *Biometrika*, 2004, 91 (4): 893—912.
- [46] Stevens D L Jr, Olsen A R. Spatially balanced sampling of natural resources. *Journal of the American Statistical Association*, 2004, 99 (465): 262—278.
- [47] Tillé Y. Ten years of balanced sampling with the cube method: An appraisal. *Survey Methodology*, 2011, 37 (2): 215—226.
- [48] Grafström A, Tillé Y. Doubly balanced spatial sampling with spreading and restitution of auxiliary totals. *Environmetrics*, 2013, 24 (2): 120—131.
- [49] Yao R J, Yang J S, Jiang L. Study on spatial variability and appropriate sampling quantity of soil salinity in Yellow River Delta. *Journal of Soil and Water Conservation*, 2006, 20 (6): 89—94. [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.]
- [50] Zhang Z X, Xu M X, Liu J, et al. Spatial variation and reasonable sampling number of soil organic carbon under different geomorphic types on the Loess Plateau. *Journal of Natural Resources*, 2014, 29 (12): 2103—2113. [张志霞, 许明祥, 刘京, 等. 黄土高原不同地貌区土壤有机碳空间变异与合理取样数研究. *自然资源学报*, 2014, 29 (12): 2103—2113.]
- [51] Zhao Y T, Chang Q R, Li Z P, et al. Study on spatial distribution of total nitrogen in arable soil layer and optimization of number of soil samples for the study based on cokriging. *Acta Pedologica Sinica*, 2014, 51 (2): 415—422. [赵业婷, 常庆瑞, 李志鹏, 等. 基于 cokriging 的耕层土壤全氮空间特征及采样数量优化研究. *土壤学报*, 2014, 51 (2): 415—422.]
- [52] Chen T E, Chen L P, Wang Y J, et al. Optimal arrangement of soil nutrient sampling based on geo-statistics. *Transactions of the Chinese Society of Agricultural Engineering*, 2009, 25 (S2): 49—55. [陈天恩, 陈立平, 王彦集, 等. 基于地统计的土壤养分采样布局优化. *农业工程学报*, 2009, 25 (S2): 49—55.]

- [53] Li Y, Shi Z, Wu C F, et al. Optimised spatial sampling scheme for soil electrical conductivity based on Variance Quad-Tree (VQT) method. *Agricultural Sciences in China*, 2007, 6 (12): 1463—1471.
- [54] Zhu A X, Liu J, Du F, et al. Predictive soil mapping with limited sample data. *European Journal of Soil Science*, 2015, 66 (3): 535—547.
- [55] Liu J, Zhu A X, Zhang S J, et al. Large-scaled soil attribute mapping method based on individual representativeness of sample sites. *Acta Pedologica Sinica*, 2013, 50 (1): 12—20. [刘京, 朱阿兴, 张淑杰, 等. 基于样点个体代表性的大尺度土壤属性制图方法. *土壤学报*, 2013, 50 (1): 12—20.]
- [56] Zhang S J, Zhu A X, Liu J, et al. Soil sampling scheme based on simulated annealing method. *Chinese Journal of Soil Science*, 2013, 44 (4): 820—825. [张淑杰, 朱阿兴, 刘京, 等. 基于模拟退火算法的土壤样点设计方法研究. *土壤通报*, 2013, 44 (4): 820—825.]
- [57] Zhang L, Zhu A X, Yang L, et al. A sample differentiation and fusion strategy for designing of soil sampling. *Acta Pedologica Sinica*, 2017, 54 (5): 1079—1090. [张磊, 朱阿兴, 杨琳, 等. 基于分融策略的土壤采样设计方法. *土壤学报*, 2017, 54 (5): 1079—1090.]
- [58] Zhu Z Y, Stein M L. Spatial sampling design for prediction with estimated parameters. *Journal of Agricultural Biological & Environmental Statistics*, 2006, 11 (1): 24—44.
- [59] Li Y, Zhu A X, Shi Z, et al. Supplemental sampling for digital soil mapping based on prediction uncertainty from both the feature domain and the spatial domain. *Geoderma*, 2016, 284: 73—84.
- [60] Bishop T F A, Horta A, Karunaratne S B. Validation of digital soil maps at different spatial supports. *Geoderma*, 2015, 241/242: 238—249.
- [61] Brus D J, Kempen B, Heuvelink G B M. Sampling for validation of digital soil maps. *European Journal of Soil Science*, 2011, 62 (3): 394—407.
- [62] Webster R, Oliver M A. Geostatistics for environmental scientists. *Journal of the Royal Statistical Society*, 2009, 43 (4): 499.
- [63] Ramirez-Lopez L, Schmidt K, Behrens T, et al. Sampling optimal calibration sets in soil infrared spectroscopy. *Geoderma*, 2014, 226/227 (4): 140—150.
- [64] Schmidt K, Behrens T, Daumann J, et al. A comparison of calibration sampling schemes at the field scale. *Geoderma*, 2014, 232/234 (12): 243—256.
- [65] Knotters M, Brus D J. Purposive versus random sampling for map validation: A case study on ecotope maps of floodplains in the Netherlands. *Ecohydrology*, 2013, 6 (3): 425—434.
- [66] de Gruijter J J, Bierkens M F P, Brus D J, et al. Sampling for natural resource monitoring. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006.
- [67] Godinho Silva S H, Owens P R, Duarte de Menezes M, et al. A technique for low cost soil mapping and validation using expert knowledge on a watershed in Minas Gerais, Brazil. *Soil Science Society of America Journal*, 2014, 78 (4): 1310—1319.
- [68] Yang L, Brus D J, Zhu A X, et al. Accounting for access costs in validation of soil maps: A comparison of design-based sampling strategies. *Geoderma*, 2018, 315: 160—169.
- [69] Morvan X, Saby N P A, Arrouays D, et al. Soil monitoring in Europe: A review of existing systems and requirements for harmonisation. *Science of the Total Environment*, 2008, 391 (1): 1—12.
- [70] Zhu J X, Xu B G, Zhang L Y. Optimization layout method of monitoring sample points of cultivated land quality based on semi-variance analysis and grade combination. *Transactions of the Chinese Society of Agricultural Engineering*, 2015, 31 (19): 254—261. [祝锦霞, 徐保根, 章琳云. 基于半方差函数与等别的耕地质量监测样点优化布设方法. *农业工程学报*, 2015, 31 (19): 254—261.]
- [71] Neger R, Beylich A, Fohrer N. Long-term monitoring of soil quality changes in Northern Germany. *Geoderma Regional*, 2016, 7 (2): 239—249.
- [72] Yu S Q, Zhang B B, Xiang H, et al. Layout method for monitoring sample point of arable land quality level based on combination of factors. *Transactions of the Chinese Society of Agricultural Engineering*, 2014, 30 (24): 288—297. [余述琼, 张蚌蚌, 相慧, 等. 基于因素组合的耕地质量等级监测样点布控方法. *农业工程学报*, 2014, 30 (24): 288—297.]
- [73] Brus D J. Statistical sampling approaches for soil monitoring. *European Journal of Soil Science*, 2014, 65 (6): 779—791.
- [74] Brus D, Gilkes R J, Prakongkep N. Design-based and model-based sampling strategies for soil monitoring. *World Congress of Soil Science: Soil Solutions for A Changing World*, 2010.
- [75] Brus D J, de Gruijter J J. A hybrid design-based and model-based sampling approach to estimate the temporal trend of spatial means. *Geoderma*, 2012, 173/174 (2): 241—248.
- [76] Chang X F, Bao X Y, Wang S P, et al. Exploring effective sampling design for monitoring soil organic carbon in degraded Tibetan grasslands. *Journal of Environmental Management*, 2016, 173: 121—126.
- [77] Han Z W, Huang W, Luo Y, et al. Application of simulated annealing method and neural network on optimizing soil sampling schemes based on road distribution. *Chinese Journal of Applied Ecology*, 2015, 26 (3): 891—900. [韩宗伟, 黄魏, 罗云, 等. 基于路网的土壤采样布局优化——模拟退火神经网络算法. *应用生态学报*, 2015, 26 (3): 891—900.]

- [78] Kidd D, Malone B, McBratney A, et al. Operational sampling challenges to digital soil mapping in Tasmania, Australia. *Geoderma Regional*, 2015, 4: 1—10.
- [79] Sheikholeslami R, Razavi S. Progressive Latin Hypercube sampling: An efficient approach for robust sampling-based analysis of environmental models. *Environmental Modelling & Software*, 2017, 93: 109—126.
- [80] Clifford D, Payne J E, Pringle M J, et al. Pragmatic soil survey design using flexible Latin hypercube sampling. *Computers & Geosciences*, 2014, 67 (3): 62—68.
- [81] Minasny B, McBratney A B. Digital soil mapping: A brief history and some lessons. *Geoderma*, 2016, 264: 301—311.
- [82] Zhao Y F, Li H J, Chen J, et al. Comparison between soil mapping approaches based on their ability explaining soil variability—A case of mapping soil organic carbon density of soil (0~1 m) in Nanyang district. *Acta Pedologica Sinica*, 2018, 55 (1): 43—53. [赵彦锋, 李豪杰, 陈杰, 等. 基于土壤变异解释力的几种土壤制图方法的对比研究——以南阳市 1m 土体土壤有机碳密度制图为例. *土壤学报*, 2018, 55 (1): 43—53.]
- [83] Keskin H, Grunwald S. Regression kriging as a workhorse in the digital soil mapper's toolbox. *Geoderma*, 2018, 326: 22—41.
- [84] Li X, Pan Y C, Zhao C J, et al. Estimation of soil organic matter based on four methods and effect of sampling number on estimation accuracy. *Scientia Geographica Sinica*, 2007, 27 (5): 689—694. [李翔, 潘瑜春, 赵春江, 等. 利用不同方法估测土壤有机质及其对采样数的敏感性分析. *地理科学*, 2007, 27(5): 689—694.]
- [85] Heung B, Ho H C, Zhang J, et al. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. *Geoderma*, 2016, 265: 62—77.
- [86] Guo Z X, Yuan Y Z, Guo Y, et al. Optimal estimation model of soil organic carbon based on the terrain factor. *Acta Pedologica Sinica*, 2017, 54(2): 331—343. [郭治兴, 袁宇志, 郭颖, 等. 基于地形因子的土壤有机碳最优估算模型. *土壤学报*, 2017, 54 (2): 331—343.]
- [87] 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 (3): 197—214.
- [88] Zhu A X, Hudson B, Burt J, et al. Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Science Society of America Journal*, 2001, 65 (5): 1463—1472.
- [89] Yang L, Zhu A X, Zhao Y G, et al. Regional soil mapping using multi-grade representative sampling and a fuzzy membership-based mapping approach. *Pedosphere*, 2017, 27 (2): 344—357.
- [90] Brus D J. Sampling for digital soil mapping: A tutorial supported by R scripts. *Geoderma*, 2019, 338: 464—480.
- [91] Vašát R, Heuvelink G B M, Borůvka L. Sampling design optimization for multivariate soil mapping. *Geoderma*, 2010, 155 (3/4): 147—153.

(责任编辑: 檀满枝)