

精准农业中土壤养分分析的适宜取样数量的确定^{*}

赵 伟¹ 谢德体¹ 刘洪斌^{1**} 王晓东² 丁声源²

(1. 西南大学资源环境学院 重庆 400716; 2. 西南大学经济管理学院 重庆 400716)

摘 要 为研究适合丘陵地区的土壤养分取样数量,减少采样分析成本,本文采用 Cochran(1977)公式对丘陵地区土壤特性的取样点数量进行了探讨。结果发现:在 1 km² 研究区域内,利用 Cochran 公式对土壤养分进行预估,其取样数目比网格取样法明显减少,土壤中的 N 以 15% 的相对误差为宜,在置信度分别为 90% 和 95% 条件下,分别只采集 2 个和 3 个样本就能反映其含量水平;P 以 15% 的均值误差为宜,在置信度分别为 90% 和 95% 条件下,分别采集 17 个和 25 个样本就足够反映其基本情况;Mg 以 5% 的均值误差为宜,在置信度分别为 90% 和 95% 下,取样数量分别为 11 个和 15 个。

关键词 精准农业 置信度 取样数量 土壤养分

中图分类号: S131+.3 **文献标识码**: A **文章编号**: 1671-3990(2008)02-0318-05

Proper soil sample numbers for soil nutrient analysis in precision agriculture

ZHAO Wei¹, XIE De-Ti¹, LIU Hong-Bin¹, WANG Xiao-Dong², DING Sheng-Yuan²

(1. College of Resources and Environment, Southwest University, Chongqing 400716, China;

2. College of Economics and Management, Southwest University, Chongqing 400716, China)

Abstract The Cochran formula was used to determine suitable sampling plot numbers of soil characteristics in order to determine proper soil sampling numbers in hill slope areas and to reduce costs of sampling and analysis. Research result shows that pre-estimation of soil nutrients using Cochran formation can greatly reduce sampling plot number. 2 to 3 samples can reliably reflect soil nitrogen content in 1 km² within the accuracy range of 90% ~ 95%, and only 15% proper relative error. 17 to 25 samples can sufficiently reflect soil phosphorus content within the accuracy range of 90% ~ 95%, and 15% proper relative error. Accordingly, 11 to 15 samples can sufficiently reflect soil magnesium content within the accuracy range of 90% ~ 95%, with 5% proper relative error.

Key words Precision agriculture, Accuracy range, Sample number, Soil nutrient

(Received Dec. 5, 2006; accepted Aug. 10, 2007)

精准农业是一种基于信息和知识管理的现代农业生产系统^[1],是由美国明尼苏达大学的土壤学者于 20 世纪 90 年代倡导环境保全型农业的通称。具体而言,所谓精准农业,是指在微观层次上按照田间每一操作单元的具体条件,精细准确地调整土壤和作物的各项管理措施,最大限度地优化各项农业投入,在保护农业生态环境、保护土地等农业自然资源的前提下,获取最大经济、环境和生态效益^[2-5]。

精准农业的实现要求精确的土壤养分状况分布图,这样的分布图可以利用 Kriging 技术对所获得的取样点属性数据进行内插而得到^[6,7]。在实际应

用中,利用一个位点上的样本属性数据来代表它周围的大面积土壤会存在一定的误差。在对土壤养分进行空间变异分析时,要求取样数据真实而有代表性,一般多采用规则栅格取样技术,以便使取样数据符合内蕴假设。除取样方法外,另一个重要的问题是如何合理确定取样的样本容量。若取样数过少,即使取样方法相当完善,也会使取样结果缺乏代表性而失去可信的价值;若取样数过多,则又耗费过多的人力、物力、财力。因此,合理确定取样数量在精准农业研究中十分重要。

本文以重庆江津市吴滩镇为例,运用 Cochran

^{*} 重庆市科委重点攻关项目(2006AB1015)资助

^{**} 通讯作者, E-mail: lwzb2000@163.com

赵伟(1982~),男,江苏连云港人,博士研究生,研究方向为 3S 技术应用及开发,土地资源可持续利用。E-mail: zhaowei811330@163.com

收稿日期:2006-12-05 接受日期:2007-08-10

(1977)公式对土壤特性的取样点数量进行了探讨,指出在一定置信度条件下合理样本容量的确定方法。

1 研究区概况

江津市位于重庆市西南郊百余里的长江中游两岸,东经 $105^{\circ}49'24'' \sim 106^{\circ}4'00''$,北纬 $28^{\circ}30'05'' \sim 29^{\circ}26'42''$ 之间,地势南高北低,相对高差 1 530.2 m,全县属丘陵低山地貌类型。研究区位于重庆市区西南部的江津市吴滩镇,面积 1 km^2 。监测村地貌以丘陵为主,海拔最低点 300 m,最高处 349.9 m,相对高差 49.9 m。土壤主要为侏罗系紫色砂泥岩发育的中性紫色土。

2 研究方法

2.1 采样方法

采用网格取样法采集土壤样本 121 个,取样间隔 100 m。方法为在以网格点为圆心、10 m 为半径的范围内采集 10 钻 0~20 cm 的耕层土壤组成代表该点的混合样本(目的是得到平均的或消除小距离范围内的土壤变异),并记录采样点的坐标数据(大地坐标和相对公里网坐标),生成研究区的样点分布图。土壤 pH、有机质(OM)、速效磷(P)、速效钾(K)、速效铜(Cu)、速效铁(Fe)、速效锰(Mn)、速效钙(Ca)、速效镁(Mg)、速效硫(S)、速效硼(B)和速效锌(Zn)采用国际植物营养研究所(IPNI)的 ASI 联合浸提剂同时浸提($0.25 \text{ mol} \cdot \text{L}^{-1} \text{NaHCO}_3 + 0.01 \text{ mol} \cdot \text{L}^{-1} \text{EDTA} + 0.01 \text{ mol} \cdot \text{L}^{-1} \text{NH}_4\text{F}$),铵态氮(N)用 $1 \text{ mol} \cdot \text{L}^{-1}$ 的 KCl 浸提。室内 pH 测定采用 pH 计法(水土比为 1:25),有机质测定采用重铬酸钾容量法,速效养分测定参照文献[8]进行。

2.2 分析方法

对于区域土壤肥力的合理评价,确定合理的样本量是非常重要的。通常有几种方法用来估计样本数量,但多数都依赖于几个关键的因素:第一为取样研究的性状、特征的变异性;第二是估测过程所要求的精度;第三是可以接受的显著水平。经常用到的是 Cochran(1977)针对于区域纯随机取样而造成的最佳取样数量计算公式^[9]:

$$n = (t \cdot Std)^2 / d^2 \quad (1)$$

式中, n 为需要的取样数量, t 为与显著性水平相对应的标准正态偏差, Std 为样本标准差, d = 样本平均值 \times 相对误差(%)。如果计算所得样本数 n 大于总体样本容量 N 的 10%,则采用不重复抽样公式^[10]:

$$n' = n / (1 + n/N) \quad (2)$$

在合理取样分析的基础上,可以用比重分配法或最适分配法来进行分层取样可能性的分析。所谓分层取样即把研究区域分成性质较为均匀的区层(类型区),然后独立地从每一类型区中随机抽取 n 个样本,计算各类型区每种养分的平均数及方差等,再根据各类型区估计值采用加权法以估计整个区域(总体)平均值和变异程度。分层取样确定抽样数目的方法包括比重分配法和最适分配法,前者是按类型区单元数目的比重分配,后者是求出一种最小取样误差来确定区层取样数目,计算公式分别为:

比重分配法:

$$\text{总取样数: } n = \sum W_h S_h^2 / V \quad (3)$$

$$\text{各类型区取样数: } n_h = n \cdot W_h \quad (4)$$

最适分配法:

$$\text{总取样数: } n = (\sum W_h S_h)^2 / V \quad (5)$$

$$\text{各类型区取样数: } n_h = n \cdot (N_h S_h / \sum N_h S_h) \quad (6)$$

式中, $V = d^2 / t^2$, $W_h = N_h / N$, V 为样本估计的必需方差, d 为允许误差范围, N_h 和 S_h 分别为各类型区单元数目和各类型区的标准差, N 为区域总单元数目, W_h 为各区层比例。

3 结果与分析

3.1 取样点数量预估

把整个取样区域作为一个均质的整体取样,根据 Cochran(1977)的公式,可以求出在一定置信水平和相对误差要求下的整个区域所需要的取样数量,这是推荐施肥中所常采用的办法。毫无疑问,土壤养分变异越大,为达到一定精度所要求的取样数量相应也会增加。本研究分别取置信度 $P_1 = 90\%$ 和 $P_2 = 95\%$,在不同均值允许误差(5%、10%、15%)条件下样本容量计算见表 1。

在均值允许误差为 5%,置信度为 90%、95% 的条件下,土壤特性所需样本量除 pH、Ca、K(Mg 在置信度为 90% 所需样本容量为 11)外,其余土壤特性所需样本容量都超过了总体样本容量的 10%(即 12 个)。B 的变异系数最大,在置信度 $P_1 = 90\%$ 和 $P_2 = 95\%$,均值允许误差 15% 条件下样本容量分别为 53 个和 76 个,其余条件下,合理取样数量都超过本研究实际的取样数量(121 个);Cu 和 P 在均值允许误差 5%,置信度分别为 90% 和 95% 的条件下,所需样本容量大于实际取样数量(121 个),其余情况下合理取样数量均小于本研究实际取样量,其余土壤特性在所研究的条件下合理取样数量均小于

表 1 研究区域土壤养分合理取样数量

Tab. 1 Reasonable sampling number for soil nutrient survey

土壤特性 Soil characteristic	分布类型 Distribution type	平均值 Mean	标准差 SD	变异系数 CV(%)	取样数量 Sample number					
					$P_1 = 90\%$			$P_2 = 95\%$		
					$Q = 5\%$	$Q = 10\%$	$Q = 15\%$	$Q = 5\%$	$Q = 10\%$	$Q = 15\%$
pH	N	4.76	0.29	6.09	4	1	1	6	1	1
Ca	LgN	3.03	0.21	6.93	5	1	1	7	2	1
K	LgN	1.79	0.16	8.94	9	2	1	12	3	1
Mg	LgN	1.79	0.16	9.87	11	3	1	15	4	2
S	LgN	1.66	0.23	13.86	21	5	2	29	7	3
$\text{NH}_4^+\text{-N}$	LgN	1.21	0.17	14.05	21	5	2	30	8	3
Mn	LgN	1.42	0.24	16.90	31	8	3	44	11	5
OM	N	0.68	0.15	22.06	53	13	6	75	19	8
Zn	LgN	0.50	0.13	26.00	73	18	8	104	26	12
Fe	N	197.18	54.31	27.54	82	21	9	117	29	13
Cu	N	2.23	0.76	34.08	126	31	14	178	45	20
P	LgN	1.08	0.41	37.96	156	39	17	221	55	25
B	N	0.63	0.42	66.67	481	120	53	683	171	76

实际取样量。说明本研究中的样本容量不能满足该区域中土壤 B 在置信度 $P_1 = 90\%$ 和 $P_2 = 95\%$, 均值允许误差分别为 5% 和 10% 条件下的空间变异性的研究; 本研究中实际采集的样本量(121 个)也掩盖了 Cu 和 P 在均值允许误差 5%, 置信度分别为 90% 和 95% 下的空间变异情况。

表 1 的计算结果表明, 在不同置信度、不同均值允许误差、不同变异系数条件下, 土壤特性所需采样点不同, 主要表现在以下几方面: (1) 当均值允许误差固定时, 合理取样数量随置信度提高而增加; (2) 当置信度相同时, 合理取样数随均值允许误差减小(即精度要求提高)而增加; (3) 当置信度和均值允许误差同时固定时, 合理取样数量随土壤特性空间变异系数增加而增加。

3.2 取样点数量确定

表 2 统计了本研究区域土壤养分含量分别在均值允许误差范围内的波动范围值, 然后根据(表 3)土壤养分系统研究法养分指标范围, 确定该区域土壤养分在置信度分别为 90% 和 95% 条件下合理取样的相对误差, 从而确定合理的样本容量。合理取样相对误差的确定遵循以下原则:

(1) 如果土壤养分含量分别在 5%、10%、15% 误差范围内, 对平均值的估测值都介于土壤养分系统研究法养分指标的同一范围之间, 则应选取误差

范围最大情况下土壤养分的合理取样数量。本研究中, 土壤 P 的平均水平为 $17.73 \mu\text{g} \cdot \text{mL}^{-1}$, 高于临界值 $12 \mu\text{g} \cdot \text{mL}^{-1}$, 刚好处于土壤养分系统研究法养分指标范围 P 分级的中等肥力水平 $13 \sim 25 \mu\text{g} \cdot \text{mL}^{-1}$ (见表 3), 而且, 如果分别取均值误差为 5%、10%、15% 时, 则对应的养分含量波动范围分别为 $16.84 \sim 18.61 \mu\text{g} \cdot \text{mL}^{-1}$ 、 $15.96 \sim 19.5 \mu\text{g} \cdot \text{mL}^{-1}$ 、 $15.07 \sim 20.39 \mu\text{g} \cdot \text{mL}^{-1}$, 均处于土壤养分系统研究法养分指标范围的中等肥力水平。由于应用此方法进行施肥推荐时处于该级别的样品将推荐同一施肥量, 所以对于土壤中 P 来说, 选取 15% 的均值误差为宜, 在置信度分别为 90% 和 95% 的条件下, 研究区域分别采集 17 个和 25 个样本就足够了(见表 1)

土壤 N 含量的平均值为 $17.33 \mu\text{g} \cdot \text{mL}^{-1}$, 远低于土壤养分系统研究法的临界值 $50 \mu\text{g} \cdot \text{mL}^{-1}$, 5%、10%、15% 的相对误差分别为 $0.87 \mu\text{g} \cdot \text{mL}^{-1}$ 、 $1.73 \mu\text{g} \cdot \text{mL}^{-1}$ 、 $2.6 \mu\text{g} \cdot \text{mL}^{-1}$, 在这些误差范围内, 对平均值的估测值分别在 $16.46 \sim 18.19 \mu\text{g} \cdot \text{mL}^{-1}$ 、 $15.59 \sim 19.06 \mu\text{g} \cdot \text{mL}^{-1}$ 、 $14.73 \sim 19.92 \mu\text{g} \cdot \text{mL}^{-1}$ 之间, 均处于土壤 N 含量低肥力水平 $0 \sim 75 \mu\text{g} \cdot \text{mL}^{-1}$ 范围。从推荐施肥和节约成本的角度考虑, 应选取 15% 的相对误差, 在置信度分别为 90% 和 95% 条件下, 研究区域分别只采集 2 个和 3 个样本, 从而减少

表 2 土壤养分推荐取样数量

Tab.2 Recommended sampling number for soil nutrient survey

项目 Item	平均值 Mean ($\mu\text{g}\cdot\text{mL}^{-1}$)	含量范围 Content range($\mu\text{g}\cdot\text{mL}^{-1}$)			取样相对误差 Sampling relative error	推荐取样量 Recommended sampling number	
		5%	10%	15%			
						$P_1 = 90\%$	$P_2 = 95\%$
Ca	1 196.53	1 136.70 ~ 1 256.36	1 076.88 ~ 1 316.18	1 017.06 ~ 1 376.01	15%	1	1
Mg	245.16	232.90 ~ 257.42	220.64 ~ 269.68	208.39 ~ 281.93	5%	11	15
K	65.03	61.78 ~ 68.28	58.53 ~ 71.53	55.27 ~ 74.78	15%	1	1
N	17.33	16.46 ~ 18.19	15.59 ~ 19.06	14.73 ~ 19.92	15%	2	3
P	17.73	16.84 ~ 18.61	15.96 ~ 19.50	15.07 ~ 20.39	15%	17	25
S	51.22	48.66 ~ 53.78	46.10 ~ 56.34	43.54 ~ 58.90	15%	2	3
B	0.54	0.51 ~ 0.56	0.48 ~ 0.59	0.45 ~ 0.62	15%	53	76
Cu	2.22	2.11 ~ 2.33	2.00 ~ 2.44	1.89 ~ 2.55	15%	14	20
Fe	197.18	187.32 ~ 207.04	177.47 ~ 216.90	167.61 ~ 226.76	15%	9	13
Mn	30.59	29.06 ~ 32.12	27.53 ~ 33.65	26.00 ~ 35.18	15%	3	5
Zn	3.29	3.13 ~ 3.46	2.96 ~ 3.62	2.80 ~ 3.79	15%	8	12

表 3 土壤养分分级标准

Tab.3 Rang and critical value of soil nutrients classification

 $\mu\text{g}\cdot\text{mL}^{-1}$

项目 Item	Ca	Mg	K	N	P	S	B	Cu	Fe	Mn	Zn
养分水平 低 Low	0~400.8	0~97.2	0~78.2	0~75	0~12	0~12	0~0.2	0~1	0~20	0~5	0~2
Nutrient 中 Middle	400.8~801.6	97.2~243	78.2~156.4	75~150	13~25	13~25	0.21~0.5	1.1~3	21~40	5.0~10	2.1~4
level 高 High	801.6~3 607.2	243~1 093.5	156.4~586.5	151~300	26~75	26~75	0.51~4.0	3.1~10	41~200	11.0~50	4.1~18
临界值 Criteria value	400.8	121.5	78.2	50	12	12	0.2	1	10	5	2

了取样数量,节约取样成本。

(2)土壤养分含量在不同均值误差范围内,对于养分含量的波动范围介于土壤养分系统研究法养分指标范围的,则应选取该均值误差条件下的土壤合理样本容量,但对于微量元素,如果不同的均值误差条件下取样数量差距很大,则应考虑取样成本,从而选取精度较低情况下的土壤合理取样数量。如土壤中 B 的平均值为 $0.54 \mu\text{g}\cdot\text{mL}^{-1}$,5% 的相对误差为 $0.027 \mu\text{g}\cdot\text{mL}^{-1}$,养分含量变动范围为 $0.51 \sim 0.56 \mu\text{g}\cdot\text{mL}^{-1}$,刚好介于土壤养分系统研究法高肥力水平 $0.51 \sim 4.0 \mu\text{g}\cdot\text{mL}^{-1}$ 之间,而中等肥力水平的范围为 $0.48 \sim 0.59 \mu\text{g}\cdot\text{mL}^{-1}$,理论上应选择 5% 的均值误差较为合理,即在置信度为 90% 和 95% 下,研究区域分别应采取 481 个和 683 个土壤样本(表 1)。如果选取 10%、15% 的允许误差,则变动范围分别为 $0.48 \sim 0.59 \mu\text{g}\cdot\text{mL}^{-1}$ 、 $0.45 \sim 0.62 \mu\text{g}\cdot\text{mL}^{-1}$,根据正态分布的原则,就意味着有相当部分样本的估测数值会由于误差的存在而出现在另一个推荐级别里,因而有可能造成较大的施肥误

差。但在实际施肥管理中,由于 B 属于微量元素,施用量很少,在没有特别的精度要求情况下,没有必要花费过多的成本。因此,对于微量元素,可适当降低取样精度,即加大取样误差,减少合理取样数量。本研究中, B 在 15% 误差范围,置信度为 90%、95% 的条件下,研究区域土壤样本合理取样数量分别为 53 个、76 个(见表 1),取样数量较均值误差为 5% 的情况下明显减少,因此确定 15% 的均值误差为合理的取样误差范围。

(3)如果土壤养分含量在不同的均值误差范围内,养分含量都跨越土壤养分系统研究法养分指标范围的两个肥力水平,则应选择精度比较高的,即均值误差尽可能小的条件下,来确定土壤合理取样数量。如本研究中土壤 Mg 含量的平均值为 $245.16 \mu\text{g}\cdot\text{mL}^{-1}$,在 5%、10%、15% 误差范围内,变动范围分别为 $232.90 \sim 257.42 \mu\text{g}\cdot\text{mL}^{-1}$ 、 $220.64 \sim 269.68 \mu\text{g}\cdot\text{mL}^{-1}$ 、 $208.39 \sim 281.93 \mu\text{g}\cdot\text{mL}^{-1}$,均跨越土壤养分系统研究法养分指标范围的中、高肥力水平,根据此原则,应选取 5% 的均值误差,在置信度分别

为 90% 和 95% 下,研究区域采集土壤的样本数量分别为 11 个和 15 个。但在推荐施肥中,应根据土壤养分含量的平均值来确定,本研究中 Mg 含量的平均值介于土壤养分系统研究法养分指标范围的高等肥力水平 $243 \sim 1\,093.5 \mu\text{g} \cdot \text{mL}^{-1}$ 之间,因此应推荐此肥力水平的施肥措施。在此情况下,选取精度比较高的条件下,增加土壤合理取样数量,其目的在于能更准确地确定土壤养分含量的平均值水平,为推荐施肥提供更可靠的施肥标准。

表 2 主要是根据以上原则确定土壤养分合理的样本容量,其取样数量明显减少,避免了盲目取样而造成的不必要的浪费,以这种原则确定的取样数量对于推荐施肥研究中样本容量的确定有一定参考意义。

4 小结

利用 Cochran 公式确定的土壤养分分析的取样数目明显减少。研究表明:土壤中的各速效养分在置信度分别为 90% 和 95% 条件下,只需少量的样本便可反映其基本情况,这样便避免了盲目取样而造成的不必要的浪费。利用 Cochran 的纯随机取样方法是把整个研究区域作为一个总体来确定合理取样数目,但实际上土壤养分变异在地块之间存在着较大的差异,不同类型区间的标准差相差很大。利用地理信息系统技术,与地块图进行图层叠加后,明确土壤各种养分在地块上的分布状况,进行分层取样,可进一步减少取样数量,节约成本,同时也可有针对性的对某些特定区域适当地增加或减少取

样密度,提高取样效率。但如何划分分层取样的取样单元和如何确定各单元的权重仍然是今后进一步研究的方向。

参考文献

- [1] 安凯,谢高地,冷允法,等. 精准农业农田地理信息系统设计[J]. 中国生态农业学报,2003,11(1):66-69
- [2] Blakemore B. S. Precision farming [J]. Outlook on Agriculture,1994,23(4):275-280
- [3] Blakemore B. S. An Information System for Precision Farming [C]//Brighton Conference: Pests and Diseases. Brighton: British Crop Protection Council,1996:18-21
- [4] 金继运.“精准农业”及其在我国的应用前景[J]. 植物营养与肥料学报,1998,4(1):1-7
- [5] 陈建能. 21 世纪精确农业的新内涵及我国的发展对策[J]. 福建农林大学学报(哲学社会科学版),2003,6(1):8-11
- [6] Franzen D. W., Peck T. R. Field soil sampling density for variable rate fertilization[J]. J. Prod. Agric.,1995,8:568-574
- [7] Wollenhaupt N. C., Wolkowski R. P., Clayton M. K. Mapping soil test phosphorus and potassium for variable-rate fertilizer application [J]. J. Prod. Agric.,1994,7:441-448
- [8] 中国科学院南京土壤研究所. 土壤理化分析[M]. 上海: 上海科学技术出版社,1978
- [9] Cochran W. G. Sampling Techniques [M]. New York: John Wiley & Sons,1977
- [10] Sachs L. Applied Statistics [M]. New York: Springer-Verlag,1982