

土壤养分空间变异及合理取样数研究

薛正平, 杨星卫, 段项锁, 陆 贤
(上海市气象科学研究所)

摘 要: 土壤养分存在明显的空间差异, 田间土壤养分合理取样数取决于养分要素自身空间变异程度和人们对数据精度的要求。若空间变异大、精度要求高则需要较多采样数, 反之需较少样点数。计算结果表明, 极大多数养分要素的合理取样数较实际取样数有较大幅度减少, 意味着可大幅降低采样及分析成本。据统计学空间自相关和半方差函数方法研究结果, 养分要素存在一定的空间自相关性, 但不同方向上的自相关性有明显差异, 独立间距也因养分要素而异。研究结果可为精准农业土壤养分样点密度设定、养分要素等值线绘制及精准施肥提供支持。
关键词: 土壤养分; 空间变异; 自相关; 半方差; 合理取样数
中图分类号: O 212 2; S151. 9 文献标识码: A 文章编号: 100226819(2002) 0420006204

土壤特性的空间差异是指土壤质地相同的分区内, 土壤特性如物理参数、状态变量等, 在同一时刻各空间点的值不相等。这种差异性直接关系到土壤养分测点的密度。确定合理取样点数, 对减少采样分析费用的作用是十分显然的; 对养分要素进行最优插值是制作等值线图及分析其空间分布的基础, 插值的合理正确与否要依据养分要素空间自相关性和独立间距的研究。这些都是精准农业的基础工作^[1, 2]。本文将前人成果^[3, 4]运用在上海精准农业示范基地中, 以期得到适用于上海实际的合理取样方法。

1 试验材料和方法

土壤养分取样地点为中科院上海精准农业示范基地。基地位于上海奉贤区境内的上海市农工商现代农业园区内。该区域是 1958 年围垦的滨海盐土。样点按 57 m × 57 m 的网格设置, 去除随机分布在

田埂、渠道、沟系中的样点, 实际取样点共 280 个。取样时间为 2000 年 5 月 10 日。前茬作物小麦。取样及观测项目有 3 大类, 即土壤养分及酸碱度、土壤可溶性离子浓度、土壤有效微量元素含量。本文着重分析耕作层的全氮、速效 N、有机质、全 P、速效 P、全盐、NaCl、速效 K 及 pH 值的空间变异特性。

2 分析与讨论

2.1 经典统计分析与合理取样数分析

2.1.1 土壤养分的空间差异性

运用数理统计中变异系数公式计算, 结果表明土壤养分要素的空间变异系数 (C_v) 存在明显的差异。其中 pH 值为弱变异 ($C_v < 10\%$); 全 N、全 P、速效 N、有机质、速效 K 为中等变异 ($10\% < C_v < 30\%$); 全盐、NaCl 和速效 P 为强变异 ($C_v > 30\%$), 见表 1。其差异性在养分要素空间分布图上得到直观反映。空间变异强弱将决定合理取样数的多寡。

表 1 土壤养分空间变异系数(上海, 2000 年)

Table 1 Soil nutrient spatial variation coefficients (Shanghai, 2000)

要 素	pH 值	全 N	全 P	速效 N	有机质	速效 K	全盐	NaCl	速效 P
变异系数 $\delta\%$	1. 8	11. 2	11. 7	12. 7	14. 1	21. 4	34. 2	40. 2	43. 1

2.1.2 经典方法确定合理取样数原理

把各测定点土壤养分看作是相互独立的随机变量是经典统计分析确定合理取样数的基本出发点。其原理概述如下^[5, 6]。

假定从均值(数学期望)为 L 、方差为 R^2 的总体中取出 n 个样本(即实际工作中的采样数), 则其平均值为 X_n , 事件 $|X_n - L| \leq \delta$ 的概率 P 所要达到的置信水平 P_1 的表达式为

$$P\{|X_n - L| \leq \delta\} = P_1 \tag{1}$$

式中 δ ——某特定精度。若是独立取样且取样数量足够多, 则大数定理成立。随机变量 $(X_n - L)/\sqrt{R^2/n}$ 服从 $R = 1, L = 0$ 的标准正态分布。根据标准正态分布的特性, 当 P_1 给定时, 例如 $P_1 =$

收稿日期: 2001208209 修订日期: 2002204216
基金项目: 中国科学院知识创新工程项目- 上海精准种植技术集成与示范(K2CX22412)
作者简介: 薛正平(1962-), 男, 高级工程师, 主要从事作物模型和农业生产模型研制, 上海蒲西路 166 号 上海市气象科学研究所, 200030

95% 或 90% 等, 由分布表可得到

$$P\{\beta(X_n - L) \leq (R^2 \hat{\sigma}_n)^{1/2} \beta^{-1} 1.960\} = 95\% \quad (2)$$
$$\text{或 } P\{\beta(X_n - L) \leq (R^2 \hat{\sigma}_n)^{1/2} \beta^{-1} 1.645\} = 90\% \quad (3)$$

由(2)、(3)可推得置信度分别为 95%、90% 时的合理取样数 (M)

$$M = 1.960^2 (R\hat{\sigma})^2 \quad \text{当 } P_1 = 95\% \quad (4)$$
$$\text{或 } M = 1.645^2 (R\hat{\sigma})^2 \quad \text{当 } P_1 = 90\% \quad (5)$$

以上是假定总体 R^2 为已知情况下, 合理取样数 M 的表达式。事实上, 随机变量总体的 R^2 是未知的, 只能用样本方差 S^2 来代替。这时随机变量 $(X_n - L) \leq (S^2 \hat{\sigma}_n)^{1/2}$ 服从 t 分布, 即

$$P\{\beta(X_n - L) \leq (S^2 \hat{\sigma}_n)^{1/2} \beta^{-1} K_\alpha\} = P_1 \quad (6)$$
式中 K_α —— t 分布的特征值, 可由显著性水平 $A = 1 - P_1$ 和自由度 $f = n - 1$ 查 t 分布表得到。此时合理取样数 (M)

$$M = K_\alpha^2 (S \hat{\sigma})^2 \quad (7)$$
(7) 式即实际计算用的公式。如 M 大于 n , 说明

合理采样数较实际数增加, 如小于 n , 则合理采样数较实际数减少。

2.1.3 土壤养分合理取样数分析

根据土壤养分取样资料, 置信度分别取 $P_1 = 95\%$ 和 90% , 对数学期望 (L) 的偏差 (Q) 分别为 5%、10% 和 15%。计算结果: 当 Q 固定时, M 随置信度提高而增加; 当置信度相同时, 合理取样数随 Q 减小 (即精度要求提高) 而增加; 当 P_1 和 Q 值同时固定时, M 随土壤养分要素空间变异系数增大而增加。反映变异性较强的养分要素的空间变化, 需要较多的采样数。当 P_1 、 Q 取相同值时, 比较不同养分要素 M 的值, 可以看到, 表 2 中 M 的排序与表 1 按 C_v 的排序 (按其数值自小至大排列) 完全一致; 除 $P_1 = 95\%$ 、 $Q = 5\%$ 时速效 P 的 M 外, 其它要素 M 值均较表 2 中 2000 年有效采样数大幅度减少。即通过合理取样分析, 可大幅度减少取样数, 从而节约取样分析成本。详见表 2。

表 2 土壤养分合理取样数

Fig 2 Reasonable sampling number for soil nutrient surveying										
要素	pH 值	全 N	全 P	速效 N	有机质	速效 K	全盐	NaCl	速效 P	
$P_1 = 95\%$	$Q = 5\%$	1	19	21	25	31	70	180	249	286
	$Q = 10\%$	1	5	6	7	8	18	45	63	72
	$Q = 15\%$	1	2	3	3	4	8	20	28	32
$P_1 = 90\%$	$Q = 5\%$	1	14	15	18	22	50	127	176	202
	$Q = 10\%$	1	4	4	5	6	13	32	44	51
	$Q = 15\%$	1	2	2	2	3	6	15	20	23
2000 年有效采样数		280	280	280	140	280	278	279	267	197

2.2 土壤空间结构自相关与半方差分析

2.2.1 空间结构的自相关分析

实际上, 在某养分要素的高值区附近, 该要素的值一般也较高; 在低值区附近则一般较低。其空间变化是不完全独立的, 只有当取采样点的间距超过一定值时, 才能认为是独立的。自相关系数是分析各取样点之间相关性的较好方法。根据有关计算式^[7]得到各土壤养分要素随样点间距离及在不同方向上的自相关系数。结果表明: 1) 不同养分要素的自相关系数差异较大。例如有机质、全 N 和速效 N, 当测点间距较大时仍具有较大的自相关系数, 表现出较高的空间相关性, 有较强的连续性, 而速效 P、pH 值和盐分的空间自相关性就较差, 独立性较强。不同要素的采样密度应有区别。2) 同一养分要素, 自相关系数随样点间距增大而减小。有机质自相关系数随间距的变化就是一个较典型的例子 (图 1)。如自相关系数随距离衰减较慢, 则连续性较好, 反之独立性强。3) 自相关的阶数也能反映养分要素的空间差异性。阶数高则空间差异性小, 连续性好。上海基地养分要素

自相关阶数总体上以东西向高于南北向, 即说东西养分要素的连续性较南北方向强。采样点密度设置也需考虑方向的差异。详见表 3。

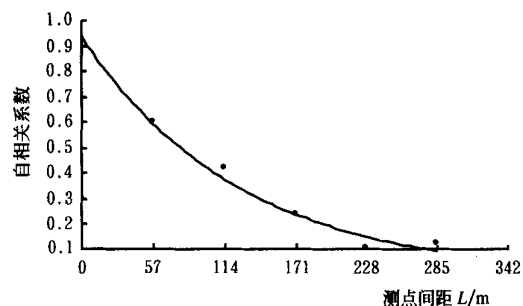


图 1 有机质自相关系数随测点间距变化 (东西向)
Fig 1 Organic matter self-correlation coefficient change with sampling points interval (E-W)

2.2.2 土壤特性空间结构的半方差分析

半方差分析是基于以下 2 个基本假设^[5, 6, 8]:

假设之一: 均值稳定。土壤特性参数的数学期望值存在且为常数, 与空间点的位置无关, 而与测点间距大小有关。

表 3 土壤养分不同方向自相关阶数

Table 3 Self-correlation coefficient grade of soil nutrient in different directions

	有机质	全 N	速效 N	速效 P	速效 K	全盐	NaCl	pH 值
自相关系数阶数(东西向)	2~ 3	2	2	1	3	0	0	1
自相关系数阶数(南北向)	2	1	1	1~ 2	0	0	0	0

假设之二: 当测点间距确定后, 要素的方差存在且为有限值, 其值仅与测点间距 L 有关。

半方差可表达为

$$r(L) = D [Y(i) - Y(i + L)]^2 \quad (8)$$

实际中通常用下式计算

$$r(L) = \frac{1}{K} \sum_{i=1}^K [Y(i) - Y(i + L)]^2 \quad (9)$$

式中 $Y(i)$ 、 $Y(i + L)$ —— 为第 i 测点和距 i 测点距离 L 的测点的养分测定值; K —— 样本数。

(9) 式的计算结果用图形更能直观地表达半方差值随测点间距 L 的变化。图 2 为速效 N 半方差值随测点距离 L 的变化。当间距小于某一值 (a) 时, $r(L)$ 随测点间距 L 增加而快速增加; 一旦间距大于 a 后, $r(L)$ 就在某个数值上下摆动, 这一数值 (图中 C_1) 即为样本方差 S^2 。 a 所代表的是该要素的相关距离, 即当 $L = a$ 时, 认为要素是与空间相关的, 用 i 测点值可推知此范围分布, 可不另设采样点; 而当 $L > a$ 时, 认为要素是空间独立的, 除 i 测点外, 须另设采样点。据计算上海精准农业示范基地速效 N 的相关距离为 285 m, 即在此距离内, 要素值分布可通过插值来反映, 若距离超过 285 m, 则须另设采样点。

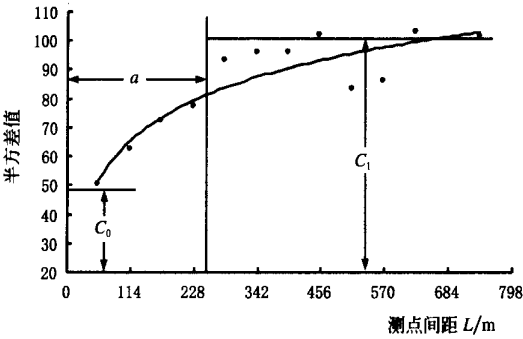


图 2 速效 N 半方差值随测点间距的变化

Fig 2 Semi-variance of NH₄N changing with sampling points interval

2.2.3 半方差函数的经验表达及物理意义

对不同形式的半方差图, 可用相应的经验函数式来表达^[5, 6, 8]。常用的有线性模型、球函数模型等。

1) 线性模型。

$$\begin{cases} r(L) = (C_1 - C_0) \times L & \text{当 } L \leq a \\ r(L) = C_1 & \text{当 } L > a \\ r(L) = 0 & \text{当 } L = 0 \end{cases} \quad (10)$$

即当 $L = a$ 时, 半方差值 $r(L)$ 随测点间距增加而线性增加; 当 $L > a$ 时, $r(L)$ 在常数 C_1 (样本方差 S^2) 上下摆动; 当 $L = 0$ 时, $r(L)$ 为 0, 即模型 (10) 为过原点的分段函数。一般情况下, $L = 0$ 时 $r(L)$ 为不等于零的值 (C_0), 图 2 中的 $r(L)$ 即不为零。不经过原点的线性函数模型变为

$$\begin{cases} r(L) = C_0 + (C_1 - C_0) \times L & \text{当 } L \leq a \\ r(L) = C_0 + C_1 & \text{当 } L > a \end{cases} \quad (11)$$

2) 球模型。球模型将 $L = a$ 时的 $r(L)$ 随测点间距变化用抛物线形式来表达; $L > a$ 的 $r(L)$ 则用线性函数表示。即

$$\begin{cases} r(L) = C_0 + C_1 [1.5L/a - 0.5(L/a)^2] & \text{当 } L \leq a \\ r(L) = C_0 + C_1 & \text{当 } L > a \end{cases} \quad (12)$$

式 (11)、(12) 中, a 为土壤要素观测值自相关范围, 当测点间距在 a 以内, 该要素间存在一定的空间相关性, 在 a 以外认为该要素值是空间独立, a 称为独立间距。

C_0 表示最小取样距离内由于土壤变异和测量误差引起的方差, 称为“金块方差”。

速效 N 的相关距离 $a = 285$ m, 金块方差 $C_0 = 50$, 半方差值最终趋于 C_1 ($C_1 = 94$, 实际样本方差为 84.2), 见图 2。

3 结 论

1) 上海精准农业示范基地 (为例) 土壤养分存在明显的空间差异。以速效 P、全盐和 NaCl 的差异最明显。

2) 经典统计方法得到的土壤养分合理取样数较 2000 年实际取样数有大幅度减少。减幅与要素的空间变异程度成反比, 即要素空间变异小则减幅大, 需要较少采样点即能反映空间分布, 反之变异大的要素减幅小, 需要较多采样点。

3) 养分要素存在一定的空间自相关性, 且自相关性强弱因方向而异。如上海基地以东西向自相关性较南北向强。在设立测点时应考虑不同方向测点密度的差异。

4) 半方差函数模型可进一步研究养分要素的空间差异, 半方差图能直观表示差异的空间意义。独立间距 a 内可依据自相关性作插值处理, 而在 a 以外, 则必须有测点资料支持才能插值绘制等值线

(面), 这是空间插值的重要依据之一。

根据上述结果, 2002 年 pH 取样数定为 10 个, 比 2000 年减少 270 个; 全 N、全 P、速效 N、有机质、速效 K, 东西方向取样间隔由原来 57 m, 扩大至 114 m, 南北方向不变, 计划取样数为 2000 年的 50%; 全盐、NaCl 和速效 P 的取样数与 2000 年相同。预计 2002 年取样总数为 2000 年的 58%。

[参 考 文 献]

[1] 汪懋华 “精细农业”发展与工程技术创新[J] 农业工程学报, 1999, 15(1): 1~ 8
[2] Yang Xinwei The primary design for demonstration base of precision faming in Shanghai[A] Beijing International Engineering Conference, 2000, one of

topic reports in A griculture Session

[3] 王 珂, 沈掌泉, John S Bailey 精准农业田间土壤空间变异与采样方式研究[J] 农业工程学报, 2001, 17(2): 33~ 36
[4] 胡克林, 李保国 农田土壤养分的空间变异特征[J] 农业工程学报, 1999, 15, 3: 33~ 38
[5] 雷志栋, 杨诗秀 土壤特性空间变异性初步研究[J] 水利学报 1985(9): 10~ 21
[6] 周慧珍, 龚子同 土壤空间变异性研究[J] 土壤学报, 1996, 33, 3: 232~ 241
[7] 安鸿志 时间序列分析[M] 华东示范大学出版社, 1992
[8] 史海滨, 陈亚新 土壤水分空间变异的套合结构模型及区域信息估计[J] 水利学报, 1994, 4: 70~ 77.

Spatial Variability of Soil Nutrient and Reasonable Sampling Number

Xue Zhengping, Yang Xingwei, Duan Xiangsuo, Lu Xian

(Shanghai Meteorology Science Research Institute, Shanghai 200030, China)

Abstract: Soil nutrient distribution has obvious spatial differences. The reasonable sampling number in field for soil nutrient surveying is based on the spatial variability of nutrient factor and the required data precision. The larger the spatial variability is, the more precise the data need to be, then the more sampling number needed. The calculated result indicates that most of the reasonable number of the soil factor can be reduced sharply, which means that a lot of cost of sampling can be saved. According to the result from self-correlation and semi-variance function of geo-statistics method, the self-correlation coefficient values have large difference in various directions, independence interval also changes with soil nutrient factors. As well, the results provide support to decide sampling point density, protracted isoline and fertilize accurately in precision faming.

Key words: soil nutrient; spatial variability; self-correlation; semi-variance; reasonable number of sampling