

土壤养分变异与合理取样数量

姜 城, 杨俐苹, 金继远, 张维理

(中国农业科学院土壤肥料研究所, 北京 100081)

摘要:利用地统计学方法、地理信息系统技术,结合土壤养分状况系统研究法对一定条件下的土壤合理取样数量作了细致的研究。结果表明,大部分土壤养分都具有较为良好的半方差结构,空间自相关距都比较大。在平衡取样成本和精确度的前提下必须考虑土壤养分的空间变异程度。利用地理信息系统等手段可以充分表现土壤养分变异的分布情况,从而为设置取样点提供依据。在本研究条件下,利用分层取样的最适分配法获得 34.5hm^2 耕地上的最佳取样数量,针对土壤速效钾的取样以 95% 的置信水平 10% 的相对误差为宜,取样数量为 24 个;针对土壤速效磷的取样以 95% 的置信水平 20% 的相对误差为宜,取样数量 10 个。

关键词:地统计学; 地理信息系统; 取样数量; 最适分配法

中图分类号:S159.2 **文献标识码:**A **文章编号:**2001-505X(2001)03-0262-09

对土壤养分时空变异的了解是进行土壤养分科学管理和合理施肥的基础^[1]。要正确评价整个区域的土壤肥力状况,只能通过科学的抽样技术,用统计学方法研究土壤样本各种养分测试值的统计量,从而估计总体(整个区域)的土壤肥力状况。一般来说,虽然适合土壤肥力评价的取样方法有很多,如名单取样(包括纯随机取样和分层随机取样)、整体取样和网格取样等,但无论哪种取样方法都应以土壤养分变异状况为依据^[1]。一些人对土壤肥力特性变异与取样的关系进行了研究。如 Jacob 和 Klute^[2]研究了在一定条件下的取样小区、取样数量以及测定方法之间的最佳组合; Cipra^[3]等研究发现地块与地块之间的差异最大,其次才是地块内取样区域之间的变异,同一取样区域内样本间的变异最小。近年来,在许多大农场里,以随机混合样本决定肥料用量的方法正在逐渐为更为精确的以网格取样为基础的地理信息系统管理变量施肥技术所代替,但我国在这方面的研究还很少。在施肥等农艺措施越来越趋于精细准确的今天,如何准确了解土壤肥力状况,从而为精确施肥打下牢固的基础,仍是一个崭新的课题。本研究旨在结合地理信息系统技术、网格取样技术及土壤养分状况系统研究法,研究一定条件下的土壤合理取样数量,为精准农业技术的推广和应用提供理论依据和基本技术。

1 材料与方法

1.1 样品的采集和分析

试验在河北省邯郸市陈刘营村 34.5hm^2 玉米田上进行。土壤为褐土,质地沙壤,前作为小麦,取土时间 1999 年 5 月 17 日。

收稿日期:2000-01-28

作者简介:姜城(1969—),男,吉林省吉林市人,博士,主要从事土壤养分管理及精准农业研究。
本研究在农业部植物营养学重点开放实验室完成。

利用网格取样法采集土壤样本 47 个,取样间隔 80m。方法为在以网格点为圆心 10m 为半径的范围内采集 10 钻 0~20cm 的耕层土壤组成代表该点的混合样本。取样点分布如图 1 所示。

土壤中的速效磷(P)、钾(K)、铜(Cu)、铁(Fe)、锰(Mn)和锌(Zn)采用国际农化服务中心的 ASI 联合浸提剂^[4]同时浸提(0.25mol/L NaHCO₃-0.01 mol/L EDTA-0.01mol/L NH₄F);速效硼(B)和硫(S)用 0.08mol/L 的磷酸钙溶液浸提;铵态氮(N)和速效钙(Ca)、镁(Mg)用 1mol/L 的 KCl 浸提。土壤样品用容量法取样。

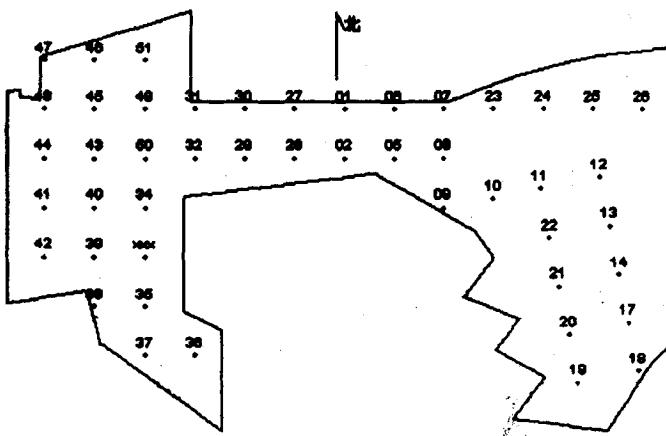


图 1 玉米田土壤取样点分布状况

Fig. 1 The distribution of sampling sites in corn field

1.2 数据处理

数据分析采用地统计学方法,主要包括:

1.2.1 半方差函数 半方差函数是描述土壤性质空间变异的一个函数,即:

$$r(h) = \frac{1}{2} v_{ar}[Z(x+h) - Z(x)]$$

式中:r(h)是间距为 h 的半方差,该值随 h 的增加而增加。

1.2.2 Kriging 插值 采用 Kriging 插值法,对点数据进行空间内插,然后利用 GIS 进行有关图幅的处理和相关的资料统计。

Kriging 插值是目前地统计学中应用最为广泛的内插方法,它是利用已知点的数据去估计任意点(x_0)处的数值,其算法可表示为:

$$Z(x_0) = \sum_{i=1}^n \lambda_i Z(x_i)$$

式中: λ_i 是考虑了半方差图中表示空间的权重,所以,Z 值的估计应该是无偏的,因为:

$$\sum_{i=1}^n \lambda_i = 1$$

估计偏差可以由下列方程求出:

$$\sigma_d^2 = b' \begin{bmatrix} \lambda \\ \mu \end{bmatrix}$$

式中 b 是被估计点与其它点之间的半方差矩阵, μ 为拉格朗日参数。

1.2.3 GIS 平台及统计软件 主要利用 MapGIS 地理信息系统软件及 GS+(Gamma Design Software)地统计学分析软件进行。

1.3 取样数目的估计

经常用到的是 Cochran^[5]针对于区域纯随机取样而构造的最佳取样数量计算公式:

$$n = (t * Std)^2 / d^2$$

式中 n 为需要的取样数量, t 为与显著性水平 α 相对应的标准正态偏差, Std 为样本标准差, d = 样本平均值 * 相对误差 (%)。

为使土壤取样更加合理, 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。计算公式分别为:

比重分配法:

$$\text{总取样数: } n = \sum W_h s_h^2 / V$$

$$\text{各类型区取样数: } n_h = n * W_h$$

最适分配法:

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

$$\text{各类型区取样数: } n_h = n * (N_h s_h / \sum N_h s_h)$$

式中: $V = (d^2/t^2)$, $W_h = N_h / N$, V -样本估计的必需方差, d -允许误差范围, t -t 检验值, 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 -区域总单元数目, W_h 各区层比例。

2 结果与分析

2.1 土壤养分的空间变异特征

对 34.5hm^2 玉米田用网格取样法采集的 47 个土壤样本进行分析, 结果见表 1。

表 1 玉米田土壤养分含量范围及临界值指标¹⁾

Table 1 Concentration and criteria for soil available nutrients in corn field

项目 Items	pH	OM	Ca	Mg	K	N	P	S	B	Cu	Fe	Mn	Zn
最大值	8.1	1.0	4509.0	505.4	348.0	17.0	150.1	101.5	4.1	22.8	28.9	5.6	2.7
最小值	7.7	0.4	1623.2	209.0	35.2	0.2	23.1	17.3	0.4	1.0	7.1	1.9	0.8
平均值	8.0	0.7	2358.3	316.4	94.3	5.5	51.7	55.6	1.4	2.6	12.9	3.0	1.2
中值	8.0	0.7	2264.5	314.7	74.3	4.9	47.8	59.0	1.3	1.6	11.5	2.9	1.1
变异系数 (%)	1.1	15.7	21.9	20.1	67.7	64.6	41.0	34.8	55.4	149.0	34.9	25.3	37.4
临界值	-	-	400.8	121.5	78.2	50.0	12.0	12.0	0.2	1.0	10.0	5.0	2.0

1) OM 为有机质, 单位 %, 其它养分单位为 $\mu\text{g}/\text{mL}$, 下同。OM : organic matter, unit is %; others unit are $\mu\text{g}/\text{mL}$, same as follows.

利用半方差模型对网格取样的土壤养分进行拟合, 各养分的半方差特征参数见表 2。本研究所采用的半方差模型在各养分之间有一定的差异, 主要的依据是采用 R^2 最大的模型。结果看出, 除 Mn 以外, 大部分养分都具有较好的半方差结构, 拟合程度最好的 Mg 的 R^2 高达 0.976; P、pH、Fe 的半方差模型拟合效果与其它元素相比稍差一些。从各养分的空间自相关距结果看出, 该研究区域内各养分的空间自相关距都比较大, 空间依赖性大小可以排列为 OM > pH, P > Fe > Zn > Mg > S > Cu > K > Ca > Mn。其中, OM 的空间自相关距最大, 达到了 1805m, 远高于其它元素, 而最小的为 Mn, 也达到了 90m。结果表明, 在本地块所采用的取样间隔(80m)基本上能满足土壤养分空间变异评价需要。

表 2 玉米田土壤速效养分空间变异半方差特征参数
Table 2 Semi-variance parameters of soil nutrients in corn field

项目 Item	c_0	$c_0 + c$	有效范围 Effective range	R^2	模型 Model
pH	0.0074	0.0074	733	0.242	Linear
OM	0.0075	0.0253	1805	0.978	Gaussian
Ca	19000	350800	364	0.889	Linear to sill
Mg	740	4987	465	0.976	Spherical
K	250	5040	396	0.871	Linear to sill
P	377.2	431.6	733	0.150	Linear
S	142.7	447.1	457	0.925	Linear to sill
Cu	0.01	19.61	399	0.841	Linear to sill
Fe	8.1	20.97	546	0.253	Exponential
Mn	0.081	0.565	90	0.012	Exponential
Zn	0.0424	0.2248	483	0.856	Exponential

2.2 土壤合理取样数量

2.2.1 把整个取样区域作为一个均质的整体取样 根据 Cochran^[5]的公式,可以求出在一定的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,这是在作推荐施肥中所经常采用的办法。毫无疑问,土壤养分变异越大,为达到一定精确度所要求的取样数量也会相应加大。由土壤养分的统计分析结果可以看出,各种元素的变异系数大小顺序为 Cu>K>N>B>P>Zn>Fe>S>Ca>Mg>OM>pH。为达到一定精确度的各元素取样数量也呈大致一致的顺序(表 3)。例如,Cu 变异系数最大,其所要求的取样数量也最高,95% 置信水平 10% 的相对误差条件下为 905 个土壤样本,次之为 K,达到 191 个;而在同等条件下,P 只需

表 3 玉米田土壤合理取样数量分析
Table 3 Adequate sampling amount for corn field (Cochran)

置信水平 Given level	相对误差 Relative error	95 %			90 %		
		5 %	10 %	20 %	5 %	10 %	20 %
pH		1	1	1	1	1	1
OM		43	11	3	30	8	2
Ca		82	21	6	58	15	4
Mg		69	18	5	48	12	3
K		762	191	48	533	134	34
N		656	164	41	458	115	29
P		274	69	18	192	48	12
S		186	47	12	130	33	9
B		471	118	30	329	83	21
Cu		3620	905	227	2529	633	159
Fe		192	48	12	134	34	9
Mn		104	26	7	73	19	5
Zn		224	56	14	157	40	10

采 69 个土壤样本。pH 要求的取样数量最少,因为整个区域的 pH 值水平相对比较均匀,所以在整个 34.5hm^2 的田块上只要取 1 个土壤样本就能基本上反映该区域的土壤酸碱水平。

表 3 还看出,随着对置信水平和相对误差要求的降低,各种元素所需要的取样数量都有大幅度的下降。如速效 P,5% 相对误差 95% 置信水平条件下需要 274 个土壤样本,而 10% 的相对误差只需要 69 个;速效 K,在 5% 的相对误差 95% 置信水平的要求下高达 762 个,10% 的相对误差下只要 191 个,而在 10% 相对误差 90% 置信水平条件下只需要 134 个土壤样本就能反映整个区域养分的真实状况。

Cameron^[7]研究指出,合理取样相对误差的确定应根据土壤养分的实际平均浓度和临界水平来确定。本研究中,土壤速效 P 的平均水平为 $50.8\mu\text{g}/\text{mL}$,远高于临界值 $12\mu\text{g}/\text{mL}$,刚好处于系统研究法速效 P 分级中的 $40\sim60\mu\text{g}/\text{mL}$ 级别的中间部位,而且,如果取 95% 置信水平 10% 的相对误差,则对应的波动范围为 $45.72\sim55.88\mu\text{g}/\text{mL}$,仍处于 $40\sim60\mu\text{g}/\text{mL}$ 的级别内。如果取 20% 的相对误差,则变动的范围会增加到 $40.64\sim60.96\mu\text{g}/\text{mL}$,与系统研究法速效 P 分级中的 $40\sim60\mu\text{g}/\text{mL}$ 级别相近。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,所以对于土壤速效 P 来说选取 20% 的相对误差,即 18 个土壤样本就已经足够了。

土壤速效 K 的平均水平为 $91.5\mu\text{g}/\text{mL}$,10% 的相对误差为 $9.15\mu\text{g}/\text{mL}$,在这个误差范围内,对平均值的估测值会出现在 $82.35\sim100.65\mu\text{g}/\text{mL}$ 之间,而土壤速效钾的临界值为 $78.2\mu\text{g}/\text{mL}$,其上一个级别为 $117.3\mu\text{g}/\text{mL}$,因而,选择 10% 的相对误差较为合理,即应采取 191 个土壤样本。如果选取 20% 的允许误差,则变动的范围为 $73.2\sim109.8\mu\text{g}/\text{mL}$,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临界值以下或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。

2.2.2 分层取样

按分层取样理论,如果土壤养分在不同类型区间的标准差相差很大,则需分别估计其在各类型区所要求的未来取样数目,即把研究区域分成性质较为均匀的区层,然后独立地从每一类型区中随机抽取样本,再根据各类型区估计值采用加权法估计整个区域的平均值和变异程度^[6]。

首先,对所获得的养分数据赋予相对的地理坐标,使之具有地理属性,然后利用地理信息系统 MapGIS 稀疏网格加密和 Kriging 空间内插的功能作出各种养分的分布图,再与地块图进行图层叠加后就可以明确各种养分在地块上的分布状况。图 2 为地块土壤速效 P 和速效 K 的分布图。从图中可看出,两种养分的变异程度在不同的地块之间都存在着一定的差异,其中以土壤速效 K 更为明显。在整个地块的右下侧速效 K 的变异较大,几乎包括了所有速效 K 的级别;而在地块的左上侧速效 K 的变异则较小,只有两个级别。根据这种情况,我们可以考虑引入分层取样的方法,但是必须解决两个问题,即如何划分分层取样的取样单元和如何确定各单元的权重。在这里,我们依据的是在 GIS 上所展示的土壤速效养分变异状况及田间可操作性,尽可能把变异程度不同的区域划分开来。本研究按土壤速效 K 的变异情况把整个区域划分成 ABCD 4 个亚区(图 3)。在 MapGIS 的区属性浏览功能可以统计出各个亚区的面积(表 4),而把各亚区的面积比作为各亚区分配取样数量的权重。

对各亚区的土壤样本数据重新进行统计,结果如表 5。可以看出,各个亚区的变异系数有很大差别,土壤速效钾在 B 区和 C 区的变异系数都在 40% 左右,而 A 区为 29.02%,D 区的变异系数最小,只有 2.42%。B 区的土壤速效磷与其它三个区也有很大的差别,这就从

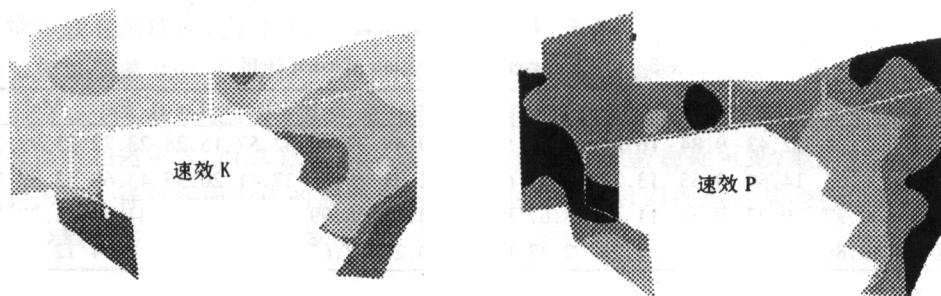


图 2 土壤速效养分分布图

Fig. 2 Distribution of some available nutrients in soil on GIS

(颜色深浅代表养分浓度的高低。Dark and light colour means high and low nutrient concentration, respectively)

表 4 各亚区的面积及比例

Table 4 Area and proportion for each sub-region

分区 Region	A	B	C	D	Total
面积(m^2)	159932	62029	131036	15364	368360
占总面积比例(%)	43.42	16.84	35.57	4.17	100

另一方面证明了各地块变异程度存在着一定的差异,因而利用 Cochran 的纯随机取样方法把整个区域作为一个总体来计算取样数量并不十分合适,其局限性在于把土壤看成了均质的统一体。而事实上,正如 Ball^[8]指出的那样,空间变异是土壤本身存在的一种自然属性。

利用比重分配法计算各亚区土壤速效磷钾的合理取样数量,结果见表 6。与 Cochran 的方法相似,由于比重法是利用样本的标准差和相对误差计算出总的取样数量,然后根据各个亚区的面积来分配,这样就会忽视变异程度在各地块之间差异的存在的客观事实,其结果表现为地块面积越大,所需要的取样数量就会越多。如土壤速效 K 在 95% 置信水平 10% 相对误差的条件下,A 区最大则需要 25 个土样,C 区次之,需要 19 个土样,这显然并不十分合理。

利用最适分配法计算的整个区域的取样数量与 Cochran 方法相比,在确定置信水平和允许相对误差条件所要求的取样数量有较大幅度的下降。如土壤速效 K,10% 相对误差 95% 置信水平条件下需 24 个土壤样本,而 Cochran 方法则需 191 个;相同条件下土壤速效 P 需 40 个土样,而 Cochran 方法则需 9 个土样(表 7)。根据上述分析,在 20% 相对误差 95% 置信水平下的 10 个土壤样本对于土壤速效 P 的估计就已足够,因而利用最适分配法所取得的取样数量在很大程度上降低了取样成本。

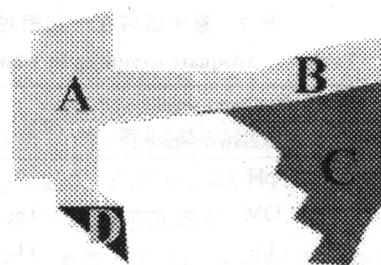


图 3 土壤速效钾亚区分布示意图

Fig. 3 Distribution of sub-region for available K sampling

表 5 各亚区的变异系数

Table 5 CV% for each sub-region of corn field

项目 Item	pH	OM	Ca	Mg	K	N	P	S	B	Cu	Fe	Mn	Zn
A 亚区	1.31	12.43	9.84	16.94	29.02	50.25	40.49	22.46	62.57	15.28	23.72	23.99	36.48
B 亚区	0.53	14.60	8.63	13.26	41.84	66.83	22.04	26.20	27.41	20.53	43.69	33.37	17.89
C 亚区	0.97	9.37	9.41	11.43	38.67	86.28	43.85	38.09	32.30	26.67	34.40	21.08	33.59
D 亚区	0.88	11.38	1.59	4.74	2.42	86.15	43.27	2.63	91.44	17.42	7.74	12.41	8.66

表 6 在一定相对误差和置信水平条件下土壤速效磷钾的合理取样数量(比重法)

Table 6 Adequate Sampling under given level of relative error and significance

置信水平 Given level	分区 Region	95 %			90 %		
		5 %	10 %	20 %	5 %	10 %	20 %
速效 K	A	99	25	7	69	18	5
	B	36	9	3	25	7	2
	C	75	19	5	52	13	4
	D	9	3	1	7	2	1
速效 P	A	108	27	7	76	19	5
	B	42	11	3	30	8	2
	C	89	23	6	62	16	4
	D	11	3	1	8	2	1

表 7 整个区域在一定相对误差和置信水平条件下合理取样数量(最适法)

Table 7 Adequate sampling in whole area under given level of relative error and significance

置信水平 Given level	95 %			90 %		
	5 %	10 %	20 %	5 %	10 %	20 %
pH	1	1	1	1	1	1
OM	18	5	2	13	4	1
Ca	11	3	1	8	2	1
Mg	23	6	2	16	4	1
K	96	24	6	67	17	5
N	421	106	27	294	74	19
P	159	40	10	111	28	7
S	62	16	4	44	11	3
B	288	72	18	201	51	13
Cu	47	12	3	33	9	3
Fe	90	23	6	63	16	4
Mn	66	17	5	46	12	3
Zn	96	24	6	67	17	5

最适分配法与比重分配法的计算结果有很大的差异,如 A 区土壤速效钾,在 10 % 相对误差 95 % 置信水平条件下比重法需取 25 个土壤样本,D 区需取 3 个样本(表 6)。土壤样本需要数量的多少与各区面积一致,为 A>C>B>D。最适法则不同,在相同条件下各亚区要求的取样数量大小顺序为 C>A>B>D,其中最多的 C 区需取 13 个样本,而最少的 D 区只

取一个样本就有足够的代表性(表 8)。

20% 相对误差 95% 置信水平条件下,对土壤速效磷的取样各亚区 A、B、C 和 D 的取样数量分别为 6、1、4 和 1,之所以 A 区取样数量大于 C 区,可能是由于 A 区和 C 区的土壤速效磷变异系数相差不大,而 A 区面积又远大于 C 区的缘故。由此可见,最适分配法一方面可以减少由于取样过多而造成的浪费,另一方面也可以有针对性地对某些特定区域适当地加大或减少取样密度,提高取样效率。

表 8 一定相对误差和置信水平条件下各亚区的合理取样数量(最适法)

Table 8 Adequate sampling in each sub-region under given level of relative error and significance

置信水平 相对误差	Given level Relative error	分区 Region	95 %			90 %		
			5 %	10 %	20 %	5 %	10 %	20 %
速效钾		A	28	7	2	20	5	2
		B	18	5	2	13	3	1
		C	50	13	4	35	9	3
		D	2	1	1	1	1	1
速效磷		A	85	22	6	59	15	4
		B	15	4	1	11	3	1
		C	54	14	4	38	10	3
		D	6	2	1	4	1	1

3 结论

3.1 在本研究中,大部分土壤养分都具有较良好的半方差结构,有较高的拟合程度。该研究区域内各养分的空间自相关距都比较大,最小为 Mn,也达到了 90m。这种结果证明,在本地块所采用的取样间隔(80m)基本上能满足土壤养分空间变异评价需要。

3.2 利用地理信息系统所作的土壤养分分布图对于表现土壤养分变异,确定合理的取样数量具有一定的参考意义。从本文的分析结果可以看出,土壤养分在地块的不同区域上的变异程度并不一致,因而利用 Cochran 的公式把整个区域作为一个均质的整体来确定合理取样数目有一定的不足。对于某些特殊的区域尤其是在土壤养分变异在地块之间存在着较大差异的情况下,可以结合土壤养分分布图采用分层取样。最适分配法具有较强的针对性,可以获得较为理想的取样数量,降低取样成本,提高取样效率。

3.3 在本研究条件下,利用分层取样的最适分配法获得 34.5hm² 耕地上的最佳取样数量,土壤速效 K 的取样以 95% 的置信水平 10% 的相对误差为宜,取样数量为 24 个;土壤速效 P 的取样以 95% 的置信水平 20% 的相对误差为宜,取样数量 10 个。

参 考 文 献:

- [1] Diaz O A et al . Soil nutrient variability and soil sampling in the Everglades agricultural area[J]. Commu. of Soil Sci. and Plant Analy., 1992,23 (17-20):2313-2337.
- [2] Jacob W C and Klute A. Sampling soils for physical and chemical properties[J]. Soil. Sci. Soc. Am. Proc., 1956, 20: 170-172.

- [3] Cipra J E, Bidwell O W, Whitney K A and Feyerherm A M. Variation with distance in selected fertility measurements of pedons of western Kansas Ustoll[J]. Soil Sci. Soc. Am. Proc., 1972, 36:111-115.
- [4] 金继运,等. 土壤养分状况系统研究法在土壤肥力研究及测土施肥中的应用[J]. 植物营养与肥料学报,1996,2(1):8-14.
- [5] Cochran W G. Sampling Techniques(3rd ed.)[M]. John Wiley and Sons, Inc., New York, 1977.
- [6] 马育华. 试验统计[M]. 北京:农业出版社,1985.
- [7] Cameron D R, Nyborg M, Toogood J A and Laverty D H. Accuracy of field sampling for soil tests[J]. Can. J. Soil Sci., 1971, 51:165-175.
- [8] Ball D F and Williams W M. Variability of soil chemical properties in two uncultivated brown earth[J]. J. Soil Sci., 1968, 19:379-391.

Soil nutrients variability and rational sampling quantity

JIANG Cheng, YANG Li-ping, JIN Ji-yun, ZHANG Wei-li

(Soil and Fertilizer Institute of CAAS, Beijing 100081, China)

Abstract: Rational sampling of soil under certain condition was studied by using Geo-statistics, GIS and Systematic Approach for Soil Nutrients Status Evaluation. The results showed that most of soil nutrients possess semi-variogram structure with longer range. It is essential to consider the spatial variability of soil nutrients in order to balance sampling cost and accuracy. GIS could give us better information about soil nutrient variability and supply the theoretical basis for arranging sampling site. In this study, the rational sampling quantity could be determined by using Stratified and Rational Distributed Sampling. The rational sampling quantity for available K in 34.5hm^2 would be 24 under 95% significant level and 10% relative error and P would be 19 under 95% significant level and 20% relative error.

Key words: geo-statistics; GIS; sampling quantity; rational distributed sampling