

采样点数量对长三角典型地区土壤肥力指标空间变异解析的影响^①

王志刚^{1,2}, 赵永存^{1*}, 黄标¹, Darilek Jeremy Landon^{1,2}, 孙维侠¹

(1 中国科学院土壤环境与污染修复重点实验室 (南京土壤研究所), 南京 210008; 2 中国科学院研究生院, 北京 100049)

摘要: 通过对江苏省如皋市 888 个采样点的不重复随机抽样, 探讨了采样点数量对土壤肥力指标空间变异解析的影响。从半方差函数估计的可靠性角度考虑, 在长江冲积物形成的土壤上, 针对县级农业管理和生态环境规划的土壤肥力指标调查采样, 采集 250 个样点较为合适。另外, 仅使用交互验证和独立验证评价半方差函数估计的可靠性及确定采样点数量是不完善的, 而平均克里格方差理论上随采样点数量的增加而单调递减, 可以作为不同采样点数量条件下, 综合评价半方差函数估计可靠性及确定采样点数量的补充指标。

关键词: 采样密度; 半方差函数; 交互验证; 独立验证; 克里格方差

中图分类号: S159.3

土壤资源利用中的许多问题都与定量评估土壤性质的空间分布密切相关, 比如土壤养分分区管理与合理施肥、土壤环境质量综合评价等。为了回答这些问题就必须对研究区的土壤进行调查采样, 通过采样点数据的统计分析和建模实现土壤性质空间变异结构的解析及空间分布的定量表征。虽然土壤是具有属性空间 (feature space) 和地理空间 (geographic space) 连续性的历史自然体, 但是任何区域调查采样都只是一个有限的采样过程, 同时, 采样预算和成本在很大程度上也限制了采样点的数量, 因此研究采样点数量对土壤性质空间变异结构的解析及空间分布定量表征的影响对于建立合理的土壤采样布点方案、提高空间预测精度、实现土壤资源高效可持续利用具有重要的理论和现实意义。

土壤性质的空间表征在很大程度上取决于对其空间变异结构的认知, 而采样点数量对半方差函数的准确推断具有重要影响。Webster和 Oliver^[1]通过对一个模拟的 256 × 256 (共 65536) 个栅格的自相关随机域 (autocorrelated random fields) 进行网格或断面采样, 认为对一个具有正态分布的各向同性变量, 采用矩法

估计时, 400 个样点可以稳健地估计半方差函数, 225 个样点可以较可靠地估计半方差函数, 150 个样点可以满足要求, 100 个样点是最低的要求, 而 50 个样点则不能满足要求。Gascuel-Oudou和Boivin^[2]通过从分布在 288 hm²上的 561 个栅格样点中抽取不同样点数量下的各 20 次重复, 探讨了不同样点数量下半方差函数估计的统计学特征, 并推荐 200 个样点比较经济合理。但是, 其在文中也指出, 由于使用的是重复抽样, 在抽取样点数量较大的情况下样点重复使用过多, 导致结论存在可靠性的问题, 例如从 561 个样点中抽取 200 个样点并重复 20 次。

虽然这些研究较为系统地探讨了采样点数量对于半方差函数估计影响的统计学特征, 但是在实际应用中还仍然存在着许多适用性的问题。首先, 土壤是一个多尺度的历史自然体, 通过模拟数据得到的推荐采样点数量, 只是在数学上提供了一个准确估计半方差函数时采样点数量需要满足的最低要求, 很难与实际的土壤尺度相匹配。Webster等^[3]指出对于空间预测制图, 采样点数量本身并不重要, 重要的是采样密度要与空间变异尺度相匹配。Kerry和Oliver^[4]认为样点间

^①基金项目: 国家自然科学基金项目 (40601039)、江苏省自然科学基金项目 (BK2009340) 和中国科学院知识创新工程重要方向项目 (KZCX2-YW-409) 资助。

* 通讯作者 (yczhaoy@issas.ac.cn)

作者简介: 王志刚 (1981—), 男, 湖北枣阳人, 博士研究生, 主要从事土壤空间预测和不确定性评价、地理信息系统应用和土壤地球化学等方面的研究。E-mail: wzg592@tom.com

距要小于变程的一半,而小于变程的 1/10 则是一种浪费。Skøien等^[5]则探讨了空间采样的尺度三要素,即幅度(extent)、样点间距(spacing)和支撑(support)对半方差函数估计的影响。其次,一些具体的研究一般只适合于他们的特定研究,不一定适合其他的区域或调查目的。McBratney^[6]指出对于县域范围的土壤进行调查采样通常是针对区域农业管理和生态环境规划的,而适用于生态环境规划的支撑在 0.02 ~ 2 km。长江三角洲地区一个县域的幅度大约在 30 km (对应面积约 1 000 km²),适合的支撑一般在 0.02 ~ 2 km,在幅度和支撑确定的情况下,通过样点间距即可确定采样点数量。

另外,在评价采样点数量对半方差函数估计可靠性的影响方面,尚缺乏完善的指标体系。虽然在实际应用中,人们大量采用交互验证(cross-validation)评价半方差函数估计可靠性以及采样点数量是否足够^[7-9],但许多学者通过实际的研究指出交互验证存在明显的不足。Zhang^[10]研究发现交互验证不能识别非高斯分布数据广义线性混合模型空间变异函数估计的明显失误;Mueller等^[11]认为不能单独使用交互验证作为评价不同插值方法优劣的标准,同时还指出独立验证(validation)的结果也受基底效应和半方差函数拟合 R^2 等的影响;Robinson和Metternicht^[12]研究表明,对不同的插值方法,交互验证得出的均方根误差非常接近,而Davis^[13]则早在 1987 年就提出,交互验证只是一种数据探索的技术,而不是假设检验的标准,它可以用作校正和重新构建模型,使模型能更好地符合已有的数据,而不能作为检验一个模型是否最优的标准。交互验证是一种具有严格限制的验证方法,它所确定的“最佳模型”是针对所选定的评价指标、样点分割方法、模型种类和参数而言的,改变它们中的任何一个就会产生新的“最佳模型”。同时,一个特定样本下的“最佳模型”并不是所有样本的“最佳模型”,交互验证主要误用之一,就是将在一个特定样本下得出的结

论推广到所有的样本。由此可见,交互验证作为不同采样点数量下半方差函数估计可靠性的评价标准是不完善的。

针对目前采样点数量对土壤性质空间变异结构解析研究中存在的这些问题,本文以长江三角洲典型地区如皋市为例,探讨适合县域土壤肥力指标空间预测的合理采样点数量;完善不同采样点数量下,具有尺度特性的半方差函数估计可靠性的评价指标。

1 材料与方法

1.1 研究区域概况

江苏省如皋市地处长江三角洲北部,位于 32°00'~32°30'N, 120°20'~120°50'E 之间,属北亚热带湿润季风气候,年平均气温 14.6℃,年平均降水量为 1 059.7 mm。该市地形平坦,陆地面积 1 402.84 km²,成土母质主要由江淮冲积物、江淮冲积浅湖相沉积物、古河汉浅湖相沉积物和长江新冲积物组成。根据中国土壤系统分类检索^[14],如皋市的土壤分属雏形土和人为土二个土纲。如皋市是我国著名的六大长寿市(县)之一,主要以农业生产为主,工业不发达,作物以小麦、水稻、玉米和油菜为主。该市的成土母质、农业种植历史及耕作方式在长三角地区具有较强的典型性和代表性。

1.2 肥力指标的选取、分析及随机样本抽取

研究区主要土壤肥力指标中,有机质含量受土壤稳定的性质影响较大,其空间自相关变程大、变异性较弱,而速效P含量受人为因素,如施肥、耕作、灌溉等影响较大,变程较短、变异性较强^[15-18]。这两个土壤肥力指标的空间变异特点,对于如皋市主要土壤肥力指标的变程阈值具有较强的代表性。因此,本文测定了如皋市 888 个采样点(图 1)的土壤有机质和速效P含量,其中有机质采用高温外加热重铬酸钾氧化-容量法测定,速效P采用Olsen法浸提-钼锑抗比色法测定^[19]。

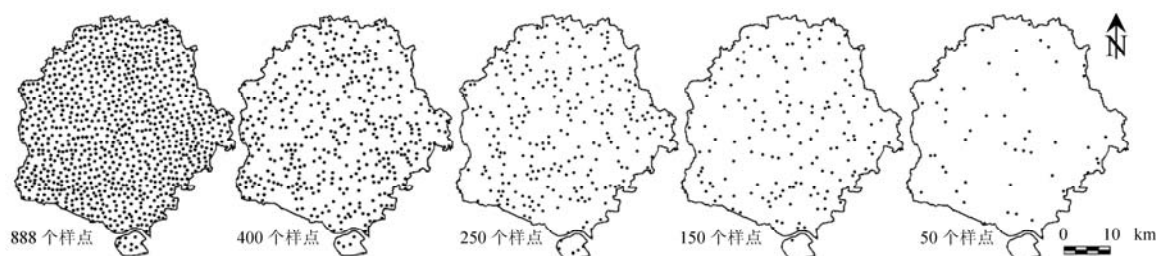


图 1 如皋市采样点及随机抽取的采样点分布图

Fig. 1 Soil sampling schemes based on different sampling densities in Rugao County

为了便于与基于模拟数据获得的半方差函数估计所需的采样点数量进行对比分析, 本文选取的不同采样点数量尽可能与 Webster 和 Oliver^[1] 提出的样点数量序列相对应, 同时, 为避免样点重复使用导致的结论不可靠, 本文采用了不重复随机抽样的方法, 即从 888 个样点中不重复地随机抽取 400、250、150 和 50 个样点 (图 1), 组成 4 个子样本来探讨不同采样点数量对有机质及速效 P 空间变异结构解析及空间分布表征的影响。采样点的平均间距和幅度采用 Blöschl^[20] 的公式计算, 即:

$$\text{样点间距} = \sqrt{\text{区域面积} / \text{样本量}} \quad (1)$$

$$\text{幅度} = \sqrt{\text{区域面积}} \quad (2)$$

888 个采样点和随机抽取的 400、250、150 和 50 个采样点的对应的平均样点间距分别为 1.26、1.87、2.37、3.06 和 5.30 km, 研究区的幅度为 37.45 km。

1.3 半方差函数估计与克里格插值

本文使用 GS+ 3.0 软件, 采用 Matheron^[21] 的矩法估计半方差函数, 以权重最小二次方方法拟合理论半方差函数模型^[22]。土壤肥力指标使用 ArcGIS 8.3 软件, 采用点普通克里格方法进行空间插值。

1.4 半方差函数估计可靠性的评价

目前广泛使用的评价半方差函数估计可靠性的方法为交互验证和独立验证。这两种方法中最常用和具有代表性的评价指标是平均误差 (ME)、均方根误差 (RMSE) 和平均绝对误差 (MAE)^[23], 其中平均误差趋近于 0、均方根误差或平均绝对误差尽量小, 是评价半方差函数估计可靠性的首要标准^[13]。作为比较, 本文计算了交互验证和独立验证的平均误差和平均绝对误差, 公式如下:

$$ME = \frac{1}{n} \sum_{i=1}^n (\hat{Z}_i - Z_i) \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{Z}_i - Z_i| \quad (4)$$

式中, Z_i 和 \hat{Z}_i 分别为样点实测值和预测值, 交互验证中 n 为样本量, 独立验证中 n 为独立验证集的样本量。本文中的独立验证集为从 888 个样点中抽出该样本后剩余的所有样点, 如 250 个样点的样本对应的验证集为 888 个样点中的另外 638 个样点。

另外, 选择平均克里格方差 (average kriging variance, AKV) 作为半方差函数估计可靠性的评价指标。使用如下公式计算研究区域内所有预测点的克里格方差^[21]和平均克里格方差:

$$\sigma_k^2 = 2 \sum_{i=1}^n \lambda_i \gamma(x_0, x_i) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(x_i, x_j) \quad (5)$$

$$AKV = \frac{1}{N} \sum_{i=1}^N \sigma_{k,i}^2 \quad (6)$$

式中, σ_k^2 为克里格方差, x_0 和 x_i 、 x_j 为待估样点和实测样点, λ_i 和 λ_j 为权重系数, N 为预测点数量。由于块段克里格会影响克里格方差的计算, 所以本文采用点克里格方差。

2 结果与讨论

2.1 描述统计分析

如阜市 888 个土壤采样点的有机质平均含量为 14.15 g/kg, 与中值 14.10 g/kg 相差不大, 变异系数仅为 10.1%, 变异性较弱 (表 1)。400、250、150 和 50 个样点的有机质含量平均值、中值、标准差及变异系数均与 888 个样点的统计结果相差不大。同时, Kolmogorov-Smirnov 检验表明, 不同采样点数量的有机质含量均符合正态分布。888 个样点速效 P 平均含量为 7.02 mg/kg, 中值 6.76 mg/kg, 平均值略高于中值,

表 1 不同抽样量样本土壤有机质和速效 P 含量的描述统计结果

Table 1 Descriptive statistics of soil organic matter and available P contents in Rugao County at different sample sizes

项目	样本量	平均值	中值	标准差	变异系数 (%)	偏度	峰度
有机质 (g/kg)	888	14.15	14.10	1.42	10.1	-0.08	0.77
	400	14.12	14.10	1.46	10.4	-0.15	1.20
	250	14.14	14.20	1.42	10.0	-0.11	0.28
	150	14.16	14.00	1.34	9.5	0.21	0.83
	50	13.99	13.95	1.50	10.9	0.04	-0.09
速效 P (mg/kg)	888	7.02	6.76	2.29	32.6	0.92	1.59
	400	6.96	6.72	2.25	32.4	1.08	2.58
	250	7.06	6.80	2.21	31.4	0.89	1.44
	150	7.06	6.77	2.49	35.3	0.86	1.00
	50	7.24	6.61	2.40	33.3	0.45	-0.77

变异系数为 32.6%，变异性较强（表 1）。400、250、150 和 50 个样点下速效 P 含量的平均值、中值、标准差、变异系数与 888 个样点相比亦相差不大。而不同采样点数量下速效 P 含量的 Kolmogorov-Smirnov 检验则表明，如皋土壤速效 P 含量均为对数正态分布。不同采样点数量下土壤有机质和速效 P 含量平均值和标准差等统计量的差异较小，说明虽然抽取的采样点数量不同但 4 个子样本对总体仍有较强的代表性。

2.2 采样点数量对有机质和速效磷空间变异结构解析的影响

由于真实的半方差函数是无法知道的，这里我们以全部 888 个样点推断的半方差函数作为比较的基

准。从有机质的理论半方差图和拟合参数来看（图 2），400 和 250 个采样点拟合的有机质理论半方差函数与 888 个样点的相似。150 个和 50 个样点与 888 个样点的差异较大，主要体现在块金常数和偏基台值的估计上存在较大的误差，特别是块金常数明显低估，两者的基底效应（块金/总基台值）分别为 14.0% 和 8.8%，而 888、400 和 250 个样点的基底效应分别为 44.0%、38.3% 和 50.0%。另外，150 个样点下有机质的变程也明显地小于 888 个样点的拟合值。因此，从本研究看，250 个样点可较可靠地推断有机质的半方差函数模型及参数，而 150 个样点，即平均样点间距大于 3.06 km 时不能保证有机质半方差函数估计的可靠性。

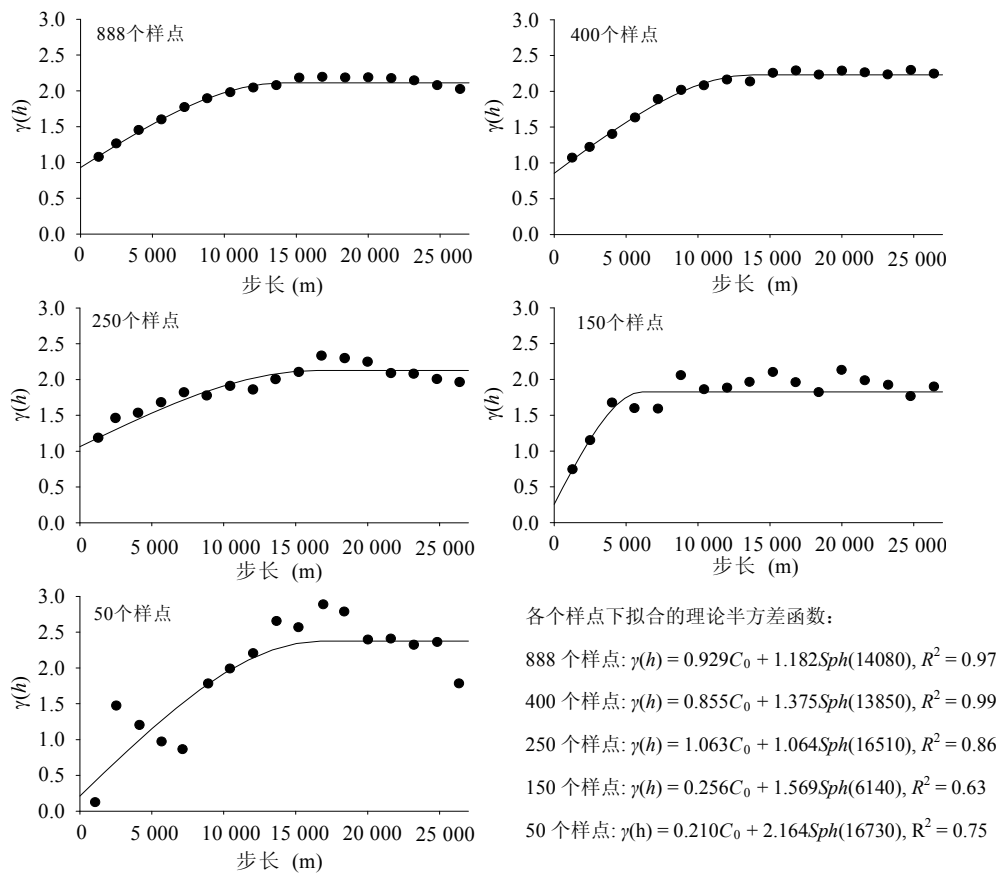


图 2 各个采样点数量下土壤有机质的半方差函数

Fig. 2 Semivariograms of soil organic matter content in Rugao County at different sample sizes

对于研究区土壤速效 P 而言，由于速效 P 含量为对数正态分布，为稳定样点数据方差，对速效 P 数据作自然对数转换后再进行半方差函数推断。不同采样点数量下速效 P 的变异结构分析表明（图 3），50 个

采样点的理论半方差函数表现为纯块金效应，与 888 个样点的推断结果差异较大，而其他几个采样点数量下速效 P 的理论半方差函数图及拟合参数均差异不大。

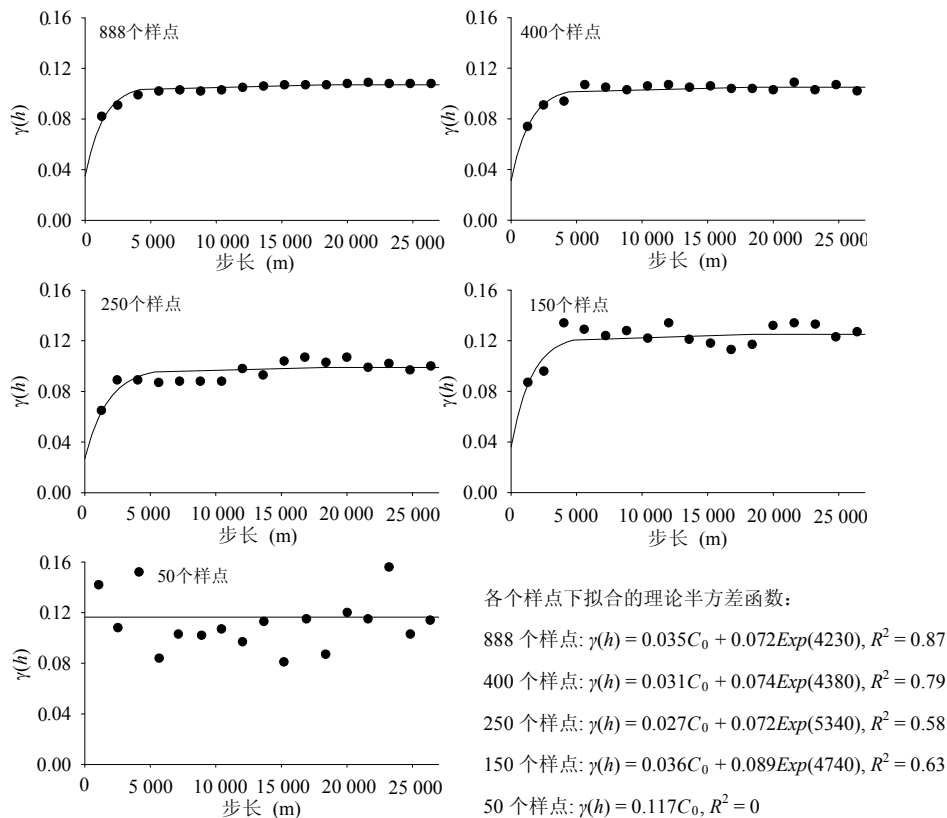


图 3 各个采样点数量下土壤速效 P 的半方差函数

Fig. 3 Semivariograms of soil available P content in Rugao County at different sample sizes

2.3 不同采样点数量下有机质和速效磷半方差函数估计可靠性的评价

2.3.1 交互验证和独立验证 从表 2 可以看出，不同采样点数量下有机质的平均误差均接近 0，速效 P 的平均误差略小于 0，其中最大的平均误差，即 50 个样点下有机质的独立验证和速效 P 的交互验证，也分别仅占其平均数的 1.1% 和 4.7%，这表明不同的采样点数量均可以保证克里格插值在总体上的无偏性，也说明交互验证和独立验证的平均误差不能明确反映采样点数量对半方差函数估计的准确性。

除有机质交互验证的平均绝对误差外，其他的平均绝对误差均随样本量的减少而增大（表 2），说明克里格插值的准确度随样本量的减少而降低。但是，交互验证和独立验证的平均绝对误差在不同采样点数量上的变化幅度均很小，不能明确指示出 150 和 50 个样点下有机质和 50 个样点下速效 P 半方差函数估计的不可靠，甚至交互验证还给出了错误的判断，即 150 和 50 个样点有机质的平均绝对误差低于 250 个样点的。因此，仅使用交互验证和独立验证评价采样点数量对

半方差函数估计可靠性的影响是不完善的。

2.3.2 平均克里格方差评价不同采样点数量下半方差函数估计的可靠性 从表 3 可见，888、400 和 250 个样点有机质的平均克里格方差随样本量的减少依次增大，而 150 和 50 个样点的平均克里格方差却比 888、400 和 250 个样点的小。理论上，对于一个研究区域，相同采样点数量的随机样本反映了相同的采样尺度，并且采样点数量越大反映的空间详尽程度越高。克里格方差只与半方差函数和样点的空间格局相关，与采样点位置土壤的实测值无关^[24]。当空间变异为各向同性时，空间格局是指预测点到采样点的距离，而本研究中，有机质和速效 P 的空间分布不存在明显的各向异性。所以，采样点数量越少则预测点到采样点的平均距离越大，平均克里格方差就越大。理论上讲，随机样本的平均克里格方差随采样点数量增加而减小，是采样点数量的单调递减函数，因此，平均克里格方差波动也可以反映采样尺度变化。然而，研究区 150 和 50 个样点估计的有机质含量半方差函数则出现了小样本量下的平均克里格方差小于大样本量下的情况

表 2 不同采样点数量下有机质和速效 P 半方差函数估计的交互验证和独立验证评价结果

Table 2 Evaluation results of semivariogram estimation of soil organic matter and available P contents in Rugao County at different sample sizes using cross validation and validation

样本量	交互验证				独立验证			
	有机质		速效 P		有机质		速效 P	
	ME	MAE	ME	MAE	ME	MAE	ME	MAE
888	0.00	0.81	-0.26	1.61				
400	-0.00	0.83	-0.26	1.62	-0.00	0.85	-0.31	1.65
250	-0.00	0.92	-0.26	1.65	0.01	0.88	-0.31	1.68
150	0.02	0.82	-0.33	1.82	-0.04	0.91	-0.31	1.70
50	-0.01	0.88	-0.34	1.88	-0.16	0.97	-0.12	1.73

表 3 不同采样点数量下有机质和速效 P 的克里格方差及克里格方差超过基台值区域的面积百分比 (速效 P 为对数尺度)

Table 3 Kriging variances and area percentages with kriging variances exceeding sill of soil organic matter and available P contents in Rugao County at different sample sizes (data of available P is in logarithmic scale)

样本量	有机质					速效 P				
	AKV	最小值	最大值	基台值	%	AKV	最小值	最大值	基台值	%
888	1.14	1.09	1.62	2.111	0	0.07	0.05	0.13	0.107	0.1
400	1.18	1.04	1.89	2.230	0	0.08	0.05	0.13	0.105	2.2
300	1.18	1.05	1.80	2.102	0	0.08	0.06	0.14	0.107	4.8
250	1.37	1.25	1.97	2.127	0	0.08	0.04	0.14	0.099	5.0
200	1.22	0.99	2.41	2.086	0.2	0.10	0.04	0.14	0.104	32.3
150	1.11	0.43	2.54	1.825	3.9	0.11	0.06	0.17	0.125	23.7
50	0.97	0.36	2.62	2.374	0.4	0.12	0.12	0.14	0.117	100

(表 3), 这表明 150 和 50 个样点数量条件下估计的有机质含量半方差函数是不可靠的, 并且平均克里格方差能够识别这种不可靠性。888、400 和 250 个样点有机质含量的最大克里格方差均小于总基台值, 不存在预测点距最近采样点的距离超过变程, 而 150 和 50 个样点有机质的最大克里格方差均大于总基台值, 有少量区域 (分别占总面积 3.9% 和 0.4%) 内的预测点距最近采样点的距离超过变程。

400 和 250 个样点速效 P 的平均克里格方差略大于 888 个样点的平均克里格方差, 150 和 50 个样点速效 P 的平均克里格方差分别比 888 个样点的大 57.1% 和 71.4% (表 3)。另外, 当预测点到最近采样点的距离超过变程时, 克里格插值就不能得到这个预测点的最优估计, 在克里格方差上的表现就是, 预测点的克里格方差超过总基台值^[24]。当出现大量克里格方差超过总基台值时, 也说明采样点数量明显不够。不同采样点数量下速效 P 含量的最大克里格方差均大于总基台值。888、400 和 250 个样点的克里格方差大于基台值的面积百分比均略大于零, 而 150 和 50 个样点条件下则分别高出 23.7% 和 100%。150、50 个样点与 888、

400 和 250 个样点间, 平均克里格方差和克里格方差超过总基台值的面积存在明显差异, 表明 150 及 50 个样点不能保证速效 P 半方差函数估计和空间插值的可靠性。

为了进一步验证 250 个样点的合理性, 同样采用不重复抽样的方法抽取了另外一个采样点数量序列, 即 300、250 和 200 个样点。从有机质的平均克里格方差看, 200 个样点的小于 250 个样点的, 同样违背了平均克里格方差单调递增的规律, 而对于速效 P, 200 个样点下克里格方差超过基台值区域的面积达到 32.3% (表 3)。因此, 200 个样点不能够保证有机质和速效 P 半方差函数推断的可靠性。

2.4 采样点数量对有机质和速效磷空间分布表征的影响

不同采样点数量下有机质和速效 P 的空间预测结果表明 (图 4), 400 和 250 个样点下有机质和速效 P 的空间分布格局基本上都能反映出 888 个样点预测的空间分布格局, 而 150 个和 50 个样点的分布格局与 888 个样点的存在明显的差异。主要表现在有机质含量的局部变异不能充分准确地表达, 888、400 和 250

个样点下获得的有机质空间分布模式表明, 如皋市东北部边缘有机质含量大部分处于 14.15 ~ 14.76 g/kg 之间, 而 150 个样点下则不能反映这一格局。一般认为有机质具有较高的空间自相关性, 其空间变异也较平滑^[15]。从 888、400 和 250 个样点下有机质的空间分布格局来看, 如皋土壤有机质的空间变异模式比较平滑, 但是 150 个样点下有机质的空间分布显得比较零乱, 这可能是由于 150 个样点显著地低估了有机质的变程

所致。速效P则由于变程较短, 888、400 及 250 个样点条件下其空间分布均呈现出复杂的岛状格局, 而 150 个样点的空间分布图则过滤了大量的岛状分布结构, 50 个样点下平滑效应更加强烈、局部岛状分布不能体现, 同时研究区西南部速效P含量也明显低估。因此, 从采样点数量对空间分布表征的影响来看, 250 个样点可以准确地表征土壤有机质和速效P的空间分布格局, 而且比较经济。

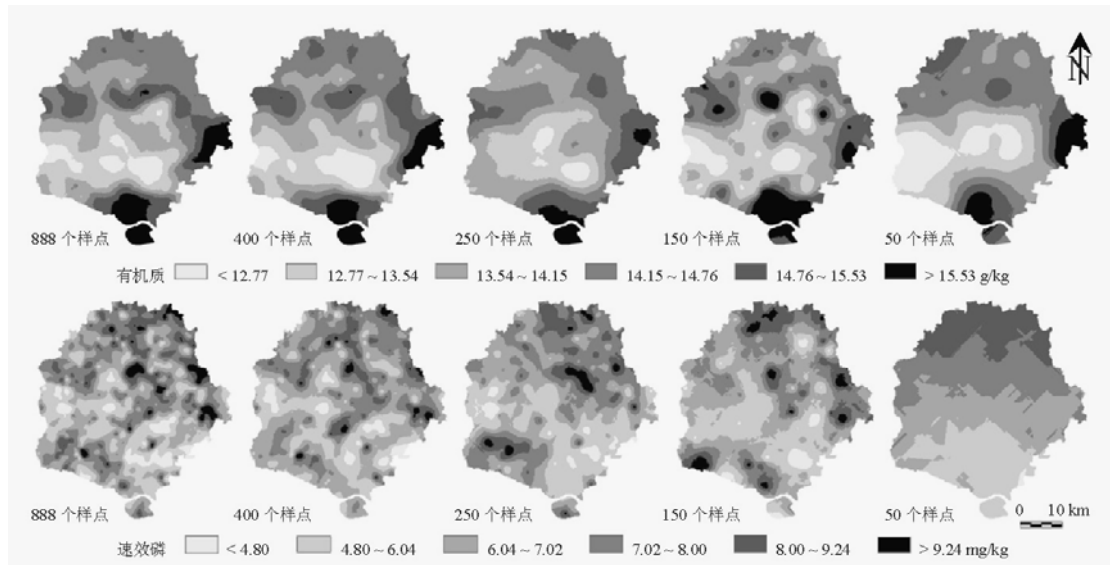


图4 不同采样点数量下克里格预测的有机质和速效P空间分布图

Fig. 4 Kriging of soil organic matter and available P contents in Rugao County at different sample sizes

3 结论

不同采样点数量条件下, 如皋市土壤有机质和速效P含量的平均值、标准差等描述统计量差异不大, 不同抽样量的样本对总体仍有较强的代表性, 但 150 和 50 个样点不能保证土壤肥力指标空间变异结构的解析的可靠性。在长江冲积物形成的土壤上, 针对县级区域农业管理和生态环境规划的土壤肥力指标调查采样, 250 个样点比较合适。此外, 单独使用交互验证难以准确评价不同采样点数量下半方差函数估计的可靠性, 独立验证的平均误差和平均绝对误差在不同采样点数量上的变化幅度较小, 也不能为采样点数量是否足够提供一个明确的判断, 而综合运用交互验证、独立验证及平均克里格方差判断 3 种方法则能够对不同采样点数量条件下半方差函数估计的可靠性做出完善评价。

参考文献:

[1] Webster R, Oliver MA. Sample adequately to estimate variograms

of soil properties. *Journal of Soil Science*, 1992, 43: 177-192

[2] Gascuel-Oudoux C, Boivin P. Variability of variograms and spatial estimates due to soil sampling: A case study. *Geoderma*, 1994, 62: 165-182

[3] Webster R, Welham SJ, Potts JM, Oliver MA. Estimating the spatial scales of regionalized variables by nested sampling, hierarchical analysis of variance and residual maximum likelihood. *Computers & Geosciences*, 2006, 32: 1320-1333

[4] Kerry R, Oliver MA. Average variograms to guide soil sampling. *International Journal of Applied Earth Observation and Geoinformation*, 2004, 5: 307-325

[5] Skøien JO, Blöschl G. Scale effects in estimating the variogram and implications for soil hydrology. *Vadose Zone Journal*, 2006, 5: 153-167

[6] McBratney AB, Odeh IOA, Bishop TFA, Dunbar MS, Shatar TM. An overview of pedometric techniques for use in soil survey. *Geoderma*, 2000, 97: 293-327

[7] Tabios GQ, Salas JD. A comparative analysis of techniques for

- spatial interpolation of precipitation. *Water Resources Bulletin*, 1985, 21: 365–380
- [8] Kerry R, Oliver MA. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. *Geoderma*, 2007, 140: 383–396
- [9] Bargaoui ZK, Chebbi A. Comparison of two kriging interpolation methods applied to spatiotemporal rainfall. *Journal of Hydrology*, 2009, 365: 56–73
- [10] Zhang H. Optimal interpolation and the appropriateness of cross-validating variogram in spatial generalized linear mixed models. *Journal of Computational and Graphical Statistics*, 2003, 12: 1–16
- [11] Mueller TG, Pusuluri NB, Mathisa KK, Cornelius PL, Barnhisel RI, Shearer SA. Map quality for ordinary and inverse distance weighted interpolation. *Soil Science Society of America Journal*, 2004, 68: 2042–2047
- [12] Robinson TP, Metternicht G. Testing the performance of spatial interpolation techniques for mapping soil properties. *Computers and Electronics in Agriculture*, 2006, 50: 97–108
- [13] Davis BM. Uses and abuses of cross-validation in geostatistics. *Mathematical Geology*, 1987, 19: 241–248
- [14] 中国科学院南京土壤研究所土壤系统分类课题组, 中国土壤系统分类课题研究协作组. 中国土壤系统分类检索. 合肥: 中国科学技术大学出版社, 2001
- [15] Cambardella CA, Moorman TB, Novak JM, Parkin TB, Karlen DL, Turco RF, Konopka AE. Field-scale variability of soil properties in central Iowa soils. *Soil Science Society of American Journal*, 1994, 58: 1501–1511
- [16] Berndtsson R, Bahri A, Jinnó K. Spatial dependence of geochemical elements in a semiarid agricultural field: II. Geostatistical properties. *Soil Science Society of America Journal*, 1993, 57: 1323–1329
- [17] Zhao YC, Xu XH, Sun WX, Huang B, Darilek JL, Shi XZ. Uncertainty assessment of mapping mercury contaminated soils of a rapidly industrializing city in the Yangtze River Delta of China using sequential indicator co-simulation. *Environmental Monitoring and Assessment*, 2008, 138: 343–355
- [18] 廖菁菁, 黄标, 孙维侠, 邹忠, 苏建平, 丁峰, 黄耀. 农田土壤有效磷的时空变异及其影响因素分析——以江苏省如皋市为例. *土壤学报*, 2007, 44(4): 620–628
- [19] 鲁如坤主编. 土壤农业化学分析方法. 北京: 中国农业科技出版社, 1999
- [20] Blöschl G. Scaling issues in snow hydrology. *Hydrological Processes*, 1999, 13: 2149–2175
- [21] Burgess TM, Webster R. Optimal interpolation and isarithmic mapping of soil properties. I the semi-variogram and punctual kriging. *Journal of Soil Science*, 1980, 31: 315–331
- [22] Cressie N. Fitting variogram models by weighted least squares. *Mathematical Geology*, 1985, 17: 563–586
- [23] 孙孝林, 赵玉国, 张甘霖, 李德成. 预测性土壤有机质制图中模糊聚类参数的优选. *农业工程学报*, 2008, 24(9): 31–37
- [24] Burgess TM, Webster R, McBratney AB. Optimal interpolation and isarithmic mapping of soil properties. IV sampling strategy. *Journal of Soil Science*, 1981, 32: 643–659

Effects of Sample Size on Spatial Characterization of Soil Fertility Properties in an Agricultural Area of the Yangtze River Delta Region, China

WANG Zhi-gang^{1,2}, ZHAO Yong-cun¹, HUANG Biao¹, DARILEK Jeremy-landon^{1,2}, SUN Wei-xia¹

(1 Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China;

2 Graduate University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Reasonable sample size or sampling density is an assurance of the accuracy of soil spatial variability analysis which is important for sustainable use of soil resources. This study discussed the influence of sample size on variogram estimation and the criterion used to judge the adequacy of sample size by comparing non-replicated random samples of 400, 250, 150, and 50 from a total of 888 samples. For reliability of variogram estimation, we suggested that a minimum sample size of 250 should be used for soils developed from Yangtze River alluvium in the Yangtze River Delta Region. This sample size recommendation is suitable for agricultural and environmental management and municipal zoning of a county. The results also showed that it is inaccurate to evaluate the reliability of variogram estimation and the adequacy of a sample size by cross-validation and validation alone. However, theoretically, average kriging variance monotonical decrease with the increase of sample size is a suitable complementary criterion to assess the reliability of variogram estimation and the adequacy of a sample size.

Key words: Sampling density, Variogram, Cross-validation, Validation, Kriging variance