

# 张家港土壤表层铜含量空间预测的不确定性评价研究\*

赵永存<sup>1</sup> 黄 标<sup>1†</sup> 孙维侠<sup>1</sup> 邵学新<sup>1,2</sup> 王志刚<sup>1,3</sup>

(1 土壤与农业可持续发展国家重点实验室(中国科学院南京土壤研究所), 南京 210008)

(2 中国林业科学院亚热带林业研究所, 浙江富阳 311400)

(3 中国科学院研究生院, 北京 100039)

**摘 要** 定量土壤重金属含量空间分布预测结果的不确定性对于控制土壤空间数据的应用风险具有重要意义。采用序贯指示模拟方法对张家港土壤表层(0~15 cm) Cu 含量的空间不确定性进行了定量评价。结果表明, 张家港东部的集约农业区及中北部的冶金和电子类企业聚集区土壤表层 Cu 含量较高, 其空间预测的波动性也较大, 而该市南部土壤 Cu 含量较低, 空间表达的波动性也较小; 序贯指示模拟的模拟实现图像没有平滑效应, 当给定一个 Cu 含量阈值时, 序贯指示模拟可以定量任一空间位置 Cu 含量大于该阈值的概率, 同时序贯指示模拟也可以评价描绘 Cu 含量大于该阈值的区域的空间不确定性。

**关键词** 土壤制图; 空间不确定性; 随机模拟

中图分类号 S159 文献标识码 A

土壤属性制图的一个重要目的就是提供有关土壤性质空间变异的信息。在应用数字土壤性质空间分布图进行分析或决策时, 人们经常假定所获取的数字化信息是完善的, 但实际上由于空间预测模型的限制等使由采样点位置的观测值来预测土壤性质空间分布的过程中通常包含一定的不确定性, 这种不确定性将通过随后的建模、分析或决策过程加以传递并对最终的结果产生深刻影响<sup>[1]</sup>。

土壤性质制图过程中最重要的一步就是采样的“点”数据到空间分布图“面”数据的空间建模预测过程, 为了实现土壤属性空间分布的定量表达和制图, 地统计学的克里格(kriging)插值与 3S 技术相结合的方法得到了广泛应用<sup>[2-5]</sup>。目前土壤性质空间预测的精度一般采用平均误差、均方根误差或方差来评价<sup>[6]</sup>。平均误差和均方根误差是从确定的位置来计算的, 是应用于整个地图的全局值, 不能指示不确定性的空间变异<sup>[7]</sup>, 而当采用地统计学方法制图时可以用方差来估计不确定性, 但克里格估计是平滑插值, 具有局部准确性和整体统计特征的不准确性的特点<sup>[8]</sup>。克里格估计的一个缺点就是平滑效应依赖于采样点的分布, 距离采样点越近平滑越小, 距离

采样点越远平滑效应越强烈, 由于这些原因使克里格估计只能提供简化变异的空间分布模式, 而在出现极端值和空间连续性模式敏感的应用中则存在很大的局限性, 比如典型的土壤污染数据<sup>[8]</sup>。克里格估计也附带了一个误差方差估计, 但克里格方差却经常被误用为衡量克里格估计可靠性的尺度, 比如估计值大于某一给定阈值, 但实际值可能低于该阈值, 因此克里格方差在度量空间不确定性方面是不完善的<sup>[9,10]</sup>。为了克服线性地统计学在评价不确定性方面的缺点, 1995 年以前实际上都是采用非线性克里格方法, 如指示克里格<sup>[11]</sup>, 但指示克里格同样存在平滑效应, 并且只能获得单独一个空间位置上土壤性质的条件累积分布函数, 对单一空间位置的局部不确定性进行定量评价, 而不能实现同时考虑多个空间位置的“空间”不确定性评价<sup>[12]</sup>。近年来人们提出采用随机模拟方法——序贯高斯模拟(sequential gaussian simulation, SGS)和序贯指示模拟(sequential indicator simulation, SIS)来进行空间不确定性建模以克服指示克里格的内在局限性<sup>[13]</sup>。SGS 和 SIS 均基于概率模型, 与克里格插值法相比, 随机模拟的优势表现在它综合考虑结果的整体统计性质和

\* 江苏省自然科学基金项目(BK2006581)、中国科学院知识创新工程领域前沿项目(ISSASIP0604)和国家重点基础研究发展规划项目(2002CB410810)资助

† 通讯作者, E-mail: bhuan@issas.ac.cn

作者简介: 赵永存(1973~), 男, 博士, 从事空间数据质量、尺度效应和不确定性建模方面的研究

收稿日期: 2006-08-23; 收到修改稿日期: 2006-11-06

模拟值的空间相关性(重现输入数据的基本统计特征和半方差模型), SGS 和 SIS 在克里格插值模型中系统地加入随机噪声, 模拟结果较插值要更真实, 此外, 随机模拟能够产生多个等概率实现(realization), 各个模拟图之间的差别正是空间不确定性的反映。Mowrer<sup>[14]</sup>采用 SGS 方法评价多年生亚高山带森林空间分布预测的不确定性; Zhao 等<sup>[1]</sup>采用 SIS 方法评价了河北省土壤碳密度空间分布表达的不确定性; 李保国等<sup>[15]</sup>采用序贯高斯模拟对农田土壤表层饱和导水率进行了条件模拟; 王学军<sup>[16]</sup>应用转向带法对北京东郊污灌区土壤铜和铅含量进行了条件模拟; 而徐英和陈亚新<sup>[17]</sup>则运用条件模拟对土壤水盐特性空间变异的各向同性进行了探讨。SGS 用于不确定性建模的前提是数据符合多高斯分布, 但在实践中很难确定多高斯分布是否存在, 而 SIS 则不考虑分布类型, 在不能确定多高斯分布是否存在时序贯指示模拟是一个明智的选择<sup>[12]</sup>。

江苏省张家港市地处长江三角洲, 改革开放以来经济增长迅速, 该地工业企业类型较多, 农业养殖业也较发达, 土壤重金属污染的潜在风险高, 探索该区土壤重金属含量空间预测的不确定性评价方法对于提供完善的土壤空间信息、控制该区土壤重金属含量空间分布信息的应用风险具有重要的理论和现实意义。

## 1 材料与方法

### 1.1 研究区概况

研究区域为江苏省张家港市(图 1), 张家港地处长江三角洲冲积平原, 地势平坦。该区属北亚热带季风气候, 年平均气温 15.2℃, 年均降雨量 1 039 mm。全市总面积 999 km<sup>2</sup>, 其中陆地面积 799 km<sup>2</sup>, 人口 85 万。该市土壤类型主要有潮土、水稻土两个土类, 潮土主要分布于北部的沿江圩田地区, 为长江冲积物母质发育而成, 其中绝大部分属灰潮土亚类。水稻土则主要分布于南部的平田地区, 由泻湖相沉积母质发育而成, 绝大部分属潴育水稻土亚类。此外黄棕壤、堆叠土和园田土在该区亦有零星分布, 但数量极少。

### 1.2 样品采集与分析

土壤采样时间为 2004 年, 共采集了 547 个土壤表层样品(0~15 cm), 其中 386 个土壤样点的布置主要根据土壤类型、土地利用状况及采样点在整个研究区分布的均匀性等因素, 另外 161 个采样点则

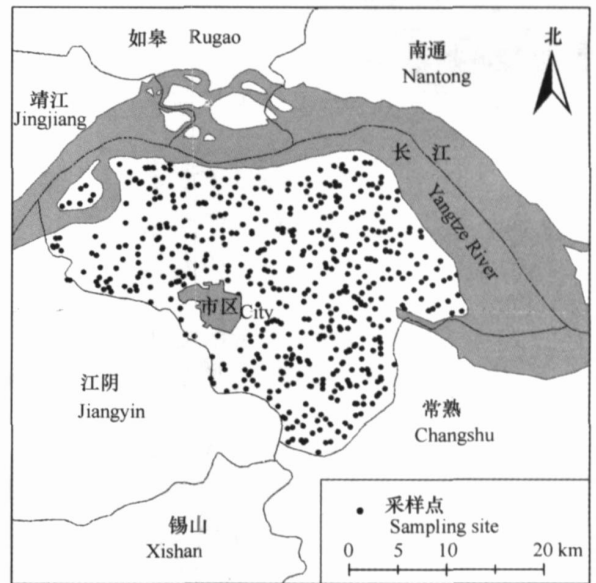


图 1 研究区域及采样点分布图

Fig. 1 Geographic location of Zhangjiagang City and the sampling site distribution map

主要分布在各类企业周围, 主要考虑了各类企业的排污方式及潜在的污染来源。每个采样点的位置均采用 GPS 记录了采样点的经纬度, 并详细记载了采样点周围的环境信息, 采样点的分布如图 1 所示, 土壤全 Cu 含量的测定见文献[18]。

### 1.3 序贯指示模拟

首先通过指示函数  $I(x, z_c)$ , 将土壤表层 Cu 含量观测值  $z(x)$  转换为指示编码(1 或 0), 以便进行 Cu 含量先验概率分布函数建模。本研究中 Cu 含量指示变换的阈值  $z_c$ (cutoff value) 设置分别为 24.12 mg kg<sup>-1</sup>(20% 分位数)、28.00 mg kg<sup>-1</sup>(40% 分位数)、31.72 mg kg<sup>-1</sup>(60% 分位数)、37.00 mg kg<sup>-1</sup>(80% 分位数)和 41.59 mg kg<sup>-1</sup>(90% 分位数)。进行指示编码变换的指示函数  $I(x, z_c)$  为:

$$I(x, z_c) = \begin{cases} 1, & z(x) \leq z_c \\ 0, & z(x) > z_c \end{cases} \quad (1)$$

通过 5 个土壤 Cu 含量阈值的指示变换可以获得 5 个 Cu 含量指示编码数据集, 然后通过半方差建模分别获得与 Cu 含量阈值相对应的 5 个指示编码数据集的半方差图和理论模型拟合参数。

定义一条遍历未采样位置的随机路径, 每次模拟过程中随机路径上的每个结点仅模拟一次, 通过以下步骤实现序贯指示模拟<sup>[8]</sup>:

(i) 对于随机路径上的一点  $i$ , 使用已经建立的指示半方差图和指示克里格来估计该点 Cu 含量小

于等于给定 Cu 含量门槛值的概率,可以获得该点 Cu 含量分别小于等于 24.12 mg kg<sup>-1</sup>、28.00 mg kg<sup>-1</sup>、31.72 mg kg<sup>-1</sup>、37.00 mg kg<sup>-1</sup>和 41.59 mg kg<sup>-1</sup>的概率值;

(ii) 对于  $i$  点使用获得的 5 个概率值来外推创建其 Cu 含量的先验条件累积分布函数;

(iii) 从  $i$  点的先验条件累积分布函数随机获取一个 Cu 含量的模拟值;

(iv) 将  $i$  点获得的 Cu 含量模拟值的指示编码加入到下一个位置的 Cu 含量先验条件累积分布函数建模中;

(v) 沿着该随机路径重复第 (i) 到 (iv) 步便获得未采样点位置 Cu 含量的一个随机模拟实现 (realization)。重复  $L$  次序贯模拟,每次模拟均使用不同的随机路径则将产生  $L$  个等概率的模拟实现。模拟的方法在重现整体统计特征的同时也建模评价了空间不确定性,这一系列的模拟实现提供了空间不确定性的可视化和定量评价的模型。

本研究中共进行了 1 000 次序贯指示模拟(栅格尺寸为 50 m × 50 m),所进行的随机模拟为条件模拟,模拟过程中采样点位置的 Cu 含量值保持不变,模拟所采用的软件为 GSLIB<sup>[13]</sup>。

## 1.4 不确定性评价

**1.4.1 局部不确定性评价** 单一空间位置  $x'$  上 (single location) Cu 含量空间表达的不确定性,即局部不确定性 (local uncertainty),可以用该点 Cu 含量  $z(x')$  大于某一给定 Cu 含量阈值  $z_t$  的概率  $Prob[z(x') > z_t]$  来进行定量评价。这个概率值可通过以下公式计算:

$$prob[z(x') > z_t] = \frac{n(x')}{1\ 000} \quad (2)$$

其中 1 000 为随机模拟的次数;  $n(x')$  为 1 000 次模拟中在  $x'$  位置上 Cu 含量模拟值大于阈值  $z_t$  的次数。

**1.4.2 空间不确定性评价** 多个位置 (multi location) 的不确定性,即空间不确定性 (spatial uncertainty),可以用于评价基于概率  $Prob[z(x') > z_t]$  来描绘 Cu 含量空间分布的可靠性。例如给定一个临界概率值  $p_c$  和 Cu 含量阈值  $z_t$ ,则基于规则  $Prob[z(x') > z_t] \geq p_c$  可确定一个 Cu 含量大于  $z_t$  的区域 A,那么由于  $p_c$  的选择差异,用区域 A 来表达 Cu 含量大于  $z_t$  的区域的不确定性有多大? 假设 A 中共包括  $n$  个

位置(栅格)  $m_1, m_2, \dots, m_n$ , 则区域 A 的 Cu 含量大于  $z_t$  的空间不确定性可由以下公式计算的联合概率来评价<sup>[12]</sup>:

$$prob[z(m_1) > z_t, z(m_2) > z_t, \dots, z(m_n) > z_t] = \frac{n(m_1, m_2, \dots, m_n)}{1\ 000} \quad (3)$$

其中 1 000 为模拟次数;  $n(m_1, m_2, \dots, m_n)$  为 1 000 次随机模拟中,区域 A 中的  $n$  个位置 Cu 含量模拟值均大于  $z_t$  的次数。

论文中图件的投影为 Albers 等积圆锥投影,将 GSLIB 模拟结果的 ASCII 码文件转换为 ASCII GRID 格式,在 ArcGIS 9.0 软件中进行图件绘制。

## 2 结果与讨论

### 2.1 描述统计分析及序贯指示模拟的实现

张家港土壤表层平均 Cu 含量为 32.01 mg kg<sup>-1</sup> (图 2),描述统计分析结果表明,土壤 Cu 含量的极差为 218.5 mg kg<sup>-1</sup>,极差范围较宽,Cu 含量的变异性较大,变异系数为 47%。偏度为 7.62,有部分 Cu 含量较高的点存在。土壤 Cu 含量的频率分布直方图表现为右侧长尾分布,说明该区土壤表层 Cu 含量的空间分布是非常不均匀的。

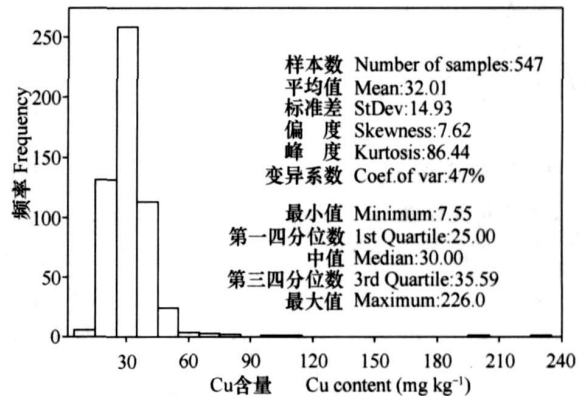


图 2 土壤表层 Cu 含量的描述统计分析与频率分布直方图

Fig. 2 Descriptive statistics and histogram of Cu contents in topsoil of Zhangjiagang City

Cu 含量的指示编码变换共使用了 5 个门槛值,相应指示编码的半方差图(图 3)和半方差模型的拟合参数(表 1)均由统计计算软件 R<sup>[19]</sup>中的 geoR<sup>[20]</sup>模块计算获得,半方差函数拟合采用的是权重最小二次方法(WLS)<sup>[13]</sup>。

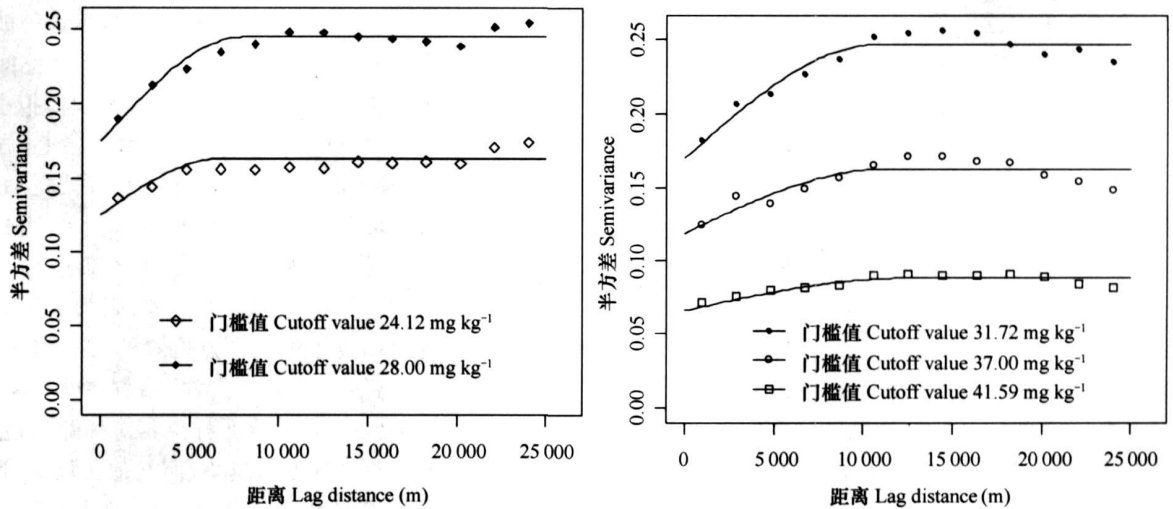


图3 指示编码的半方差图

Fig. 3 Semivariograms of the indicator codes obtained by different cutoff values

本研究中仅使用了5个阈值来建立土壤Cu含量的先验条件累计分布函数, Cattle等<sup>[21]</sup>研究认为使用较多的阈值可以获得更可靠的条件累计分布函数, 但实践中由于阈值划分过多使极端阈值条件下数据点对数过少而影响指示半方差函数建模, 此外过多的阈值也容易造成违背次序关系的出现, 这均会

对建立先验条件累计分布函数产生影响<sup>[22]</sup>。从图3和表1可以看出, 本研究中的5个指示半方差图都具有明显的空间结构性, 并且拟合的参数差异也较大, 这表明使用5个阈值能够较好地体现研究区土壤Cu含量的整体变异性, 同时这也将大大降低随机模拟过程中的计算量, 提高模拟速度。

表1 土壤Cu含量指示编码半方差函数的拟合参数

Table 1 Parameters of the theoretic model for indicator codes of total Cu contents obtained at different cutoff values

阈值 Cutoff value	理论模型 Theoretic model	块金常数 $C_0$ Nugget	基台值 Sill	变程 Range(m)	均方根误差 Root mean square error
24.12	球状 Spherical	0.125	0.163	6.666	0.002
28.00	球状 Spherical	0.175	0.245	8.000	0.001
31.72	球状 Spherical	0.170	0.247	10.800	0.002
37.00	球状 Spherical	0.119	0.163	11.350	0.002
41.59	球状 Spherical	0.066	0.088	12.500	0.002

根据以上获取的参数, 采用 GSLIB 中的 sisim 程序就可以对土壤Cu含量的空间分布进行序贯指示随机模拟, 模拟过程中的搜索半径设置为指示半方差图的最大变程。图4是从1000次模拟结果中随机抽取的

两次模拟实现图像, 每个模拟实现图像均是等可能的。模拟实现图像也反映了研究区土壤Cu含量的空间分布模式, 总体趋势表现为东部及中北部有几个高值区域出现, 南部土壤Cu含量相对较低且变异性较小。

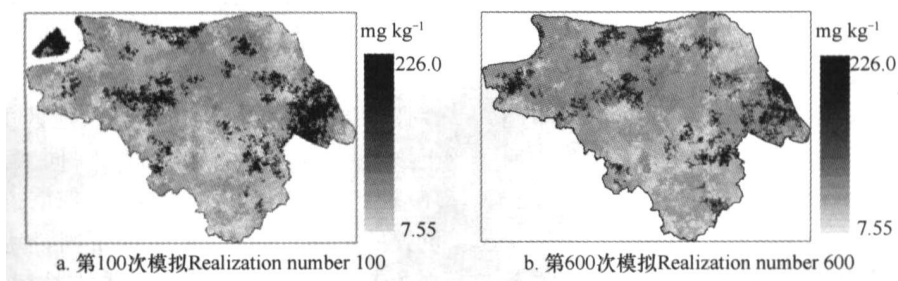


图4 从1000次模拟中随机抽取的两次模拟实现

Fig. 4 Two randomly selected realizations from 1000 realizations generated by SIS

## 2.2 局部不确定性评价

E型均值(亦称为条件均值)可以通过取1000个模拟实现的平均值而获得(图5a),它也反映了研究区土壤表层Cu含量空间分布的宏观趋势,即东部、中北部土壤Cu含量较高,而南部水稻土区的土壤Cu含量则较低。图5b是采用geoR软件计算的土壤Cu含量的普通克立格插值结果(对数变换后的半方差模型为球状模型,块金常数0.07,基台值0.093,变程6 km),1000次SIS模拟的平均值(E型均值)与普通克立格方法估计的土壤Cu含量空间分布模式类似,但由于克立格估计的平滑效应,SIS生成的模拟图像(图4)显然要比普通克立格插值结果(图5b)更现实。

由于每个随机模拟实现均是等概率、等可能的,因此条件方差(conditional variance)反映了1000次模

拟实现结果中土壤Cu含量的波动性,当然这种波动性是基于“点”的,即对于研究区域中的任一空间位置,该位置上土壤Cu含量模拟值在1000次模拟中的方差。从图5c中可以看出张家港土壤表层Cu含量空间表达波动性大的区域主要集中在具有较高Cu含量区域,这些区域具有较高的条件方差(图5c中黑色区域),土壤Cu含量空间分布表达的不确定性也较大,在采样点位置上的条件方差为零,因为采用的是条件模拟,即模拟过程中采样点位置的Cu含量值保持不变。克立格估计也提供了克立格方差(图5d),但克立格方差大的区域集中在采样点相对稀疏的地区,离采样点越近,克立格方差越小,而其他区域不管Cu含量高低,克立格方差的差异均不大,这主要是因为克立格方差仅依赖于采样点的分布情况,因此克立格方差在评价局部不确定性方面是不完善的。

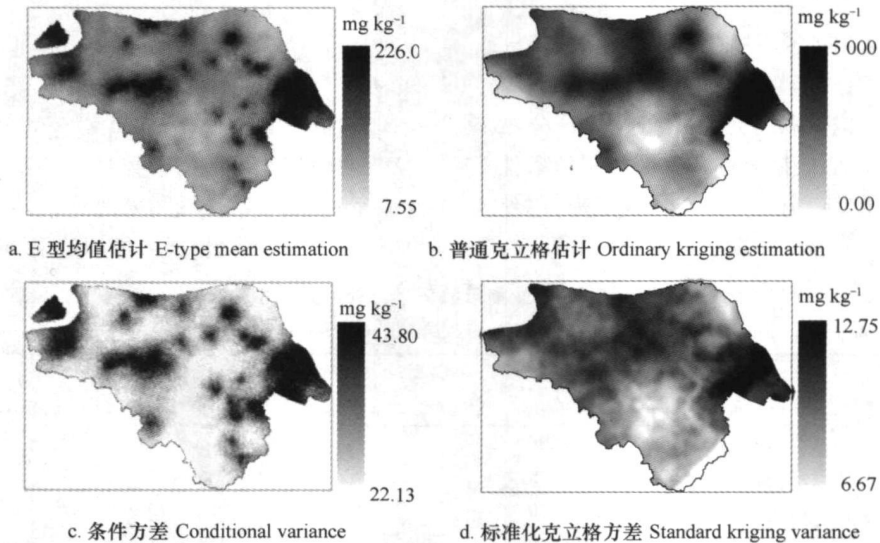


图5 土壤Cu含量的E型均值估计、条件方差及普通克立格估计和标准化克立格方差

Fig 5 E type mean estimation, conditional variance of the 1000 realizations and the ordinary kriging estimation and standard kriging variance of Cu content

土壤元素含量的不确定性可以采用置信区间表示,当随机变量的条件累积分布函数具有高度的对称性并且正态分布能够满足时,可以采用条件均值加减标准差(标准差可由条件方差计算获得)的方法来获得高斯类型的置信区间。而从E型均值估计(图5a)与中值估计(图6a)的比较来看,总体分布趋势基本一致,但局部的差异也是比较明显的,这表明土壤Cu含量的分布并非对称分布,在非对称分布条件下中值估计要较均值估计更稳健<sup>[23]</sup>。基于条件累积分布函数的分位数和1000次条件序贯指示随机模拟计算的研究区土壤表层Cu含量空间分布

的置信区间下限(2.5%分位数)和上限(97.5%分位数)如图6b、图6c所示。

由于研究区土壤表层Cu含量较高的区域波动性较大,我们选择了一个相对较高的Cu含量阈值 $35 \text{ mg kg}^{-1}$ 来阐明土壤Cu空间分布表达的不确定性评价方法。这个阈值接近该地区1985年获取的土壤表层Cu背景值的上限<sup>[24]</sup>,同时也是国家土壤环境质量标准中Cu含量的背景值,如果某一空间位置的土壤Cu含量大于该值则说明由于近20年来的工业化进程,该点可能具有被污染的“潜在”风险。基于1000次随机模拟实现计算的任一空间位

置土壤 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的概率如图 6d 所示, 可以看出张家港土壤表层 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的区域主要分布在东部及中北部地区, 这些地区 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的概率较高, 研究区大多数区域土壤表层 Cu 含量均低于  $35 \text{ mg kg}^{-1}$

(低概率区域)。土壤 Cu 含量的这种空间分布模式与研究区的工农业布局是密切相关的, 该市东部的现代集约农业集中区及中北部的冶金和电子类企业的聚集可能是造成这些区域土壤表层 Cu 含量较高的原因。

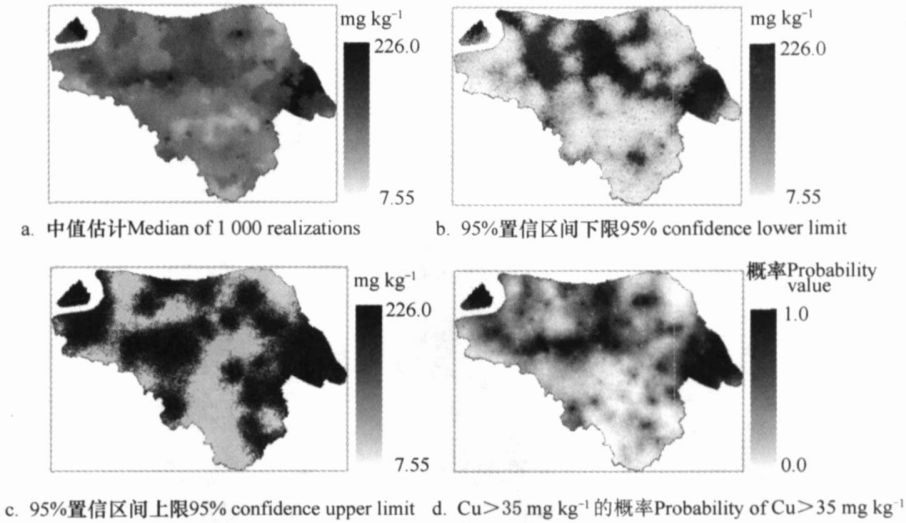


图 6 土壤 Cu 含量的中值、95% 置信区间估计及 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的概率

Fig. 6 Estimates of median and 95% confidence interval for Cu content in the study area, and the probability of Cu content being greater than  $35 \text{ mg kg}^{-1}$

### 2.3 空间不确定性评价

在给定不同临界概率  $p_c$  的前提下, 基于规则  $\text{Prob}[z(x') > 35] \geq p_c$  可以从图 6d 来确定 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的区域范围(图 7), 但以图 6d 来评价 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的区域的可靠性(即空间不确定性)是不可行的。假设以 0.95 作为临界概率  $p_c$ , 则基于规则  $\text{Prob}[z(x') > 35] \geq 0.95$  可以从图 6d 中确定

一个概率值大于等于 0.95 的区域(图 7a 中的黑色区域), 该区域中的任意一点 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的概率均大于等于 0.95, 但该黑色区域中所有点同时大于  $35 \text{ mg kg}^{-1}$  的概率(即该区域的空间不确定性)却不能通过其包含的所有点的概率值的简单相乘而获得, 因为一系列单个点的条件累积分布函数不能提供任何衡量多点联合条件累积分布函数的信息<sup>[25]</sup>。

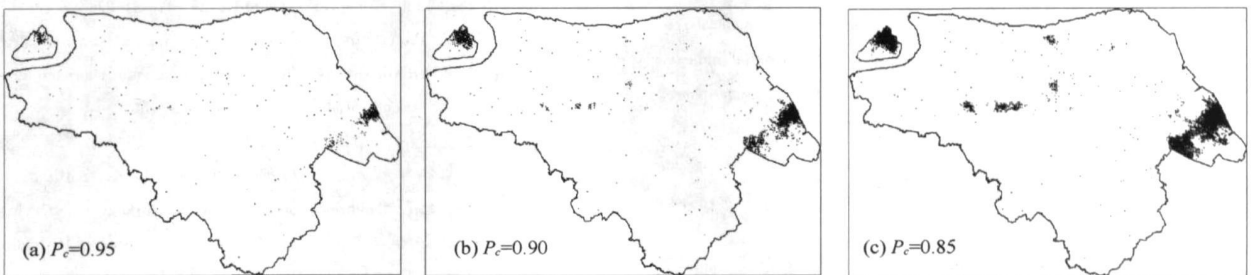


图 7 不同临界概率  $p_c$  条件下获取的土壤 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的区域分布

Fig. 7 Estimates of areas where Cu content being greater than  $35 \text{ mg kg}^{-1}$  at different critical probabilities

从图 7 可以看出随着给定的 Cu 含量临界概率值  $p_c$  的减小, 土壤 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的黑色区域(Cu 污染的潜在风险区域)个数在增加, 同时区域面积也在增大。单从概率论的角度考虑, 大于等于

0.90 的概率也是比较可靠的, 但当临界概率值  $p_c$  为 0.90 时(图 7b), 黑色区域中 5 432 个栅格点 Cu 含量同时大于  $35 \text{ mg kg}^{-1}$  的概率仅为 0.19(表 2), 即 1 000 次模拟中仅出现 190 次, 具有很大的空间不确

定性。以 0.85 为临界概率值,用图 7c 中黑色区域表达 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的区域的空間不确定性也很大,该区域出现的概率为 0.02(表 2)。但如果选择的临界概率值为 0.95,则用图 7a 中的黑色区

域来表达 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的区域则比较可靠,该区域出现的联合概率为 0.85,即 1 000 次模拟中该区域中 2 040 个栅格点同时大于  $35 \text{ mg kg}^{-1}$  的次数为 850 次,空間不确定性较小。

表 2 土壤 Cu 含量大于  $35 \text{ mg kg}^{-1}$  的区域的空間不确定性评价

Table 2 Assessment of spatial uncertainty of areas where Cu content is greater than  $35 \text{ mg kg}^{-1}$  based on joint probability

	临界概率 $p_c$ Critical probability		
	0.95	0.90	0.85
50 m × 50 m 栅格的个数 Number of cells	2 040	5 432	11 244
联合概率 Joint probability	0.85	0.19	0.02

### 3 结 论

张家港土壤表层 Cu 含量的变异性较大,变异系数为 47%,1 000 次序贯指示模拟的结果表明,张家港土壤表层 Cu 含量空间表达波动性大的区域主要集中在该市东部及中北部地区,该区的 Cu 含量也相对较高,而该市南部的土壤 Cu 含量较低,Cu 含量空间表达的波动性也小。序贯指示模拟的模拟实现图像可以反映土壤 Cu 含量的空間分布模式,并且模拟实现没有平滑效应,所有的 Cu 含量模拟实现图像一起可用于探索土壤表层 Cu 含量空間分布表达的可能模式。当给定一个 Cu 含量阈值时,序贯指示模拟可以评价任一空間位置 Cu 含量大于该阈值的概率,同时序贯指示模拟也可以评价描绘 Cu 含量大于该阈值的区域的空間不确定性。

### 参 考 文 献

- [ 1 ] Zhao Y C, Shi X Z, Yu D S, *et al.* Uncertainty assessment of spatial patterns of soil organic carbon density using sequential indicator simulation, a case study of Hebei Province, China. *Chemosphere*, 2005, 59(11): 1 527~ 1 535
- [ 2 ] Gotway C A, Hengert G W. Incorporating spatial trends and anisotropy in geostatistical mapping of soil properties. *Soil Sci. Soc. Am. J.*, 1997, 61: 298~ 309
- [ 3 ] McBratney A B, Mendonca Santos M L, Minasny B. On digital soil mapping. *Geoderma*, 2003, 117: 3~ 52
- [ 4 ] Zhao Y C, Shi X Z, Yu D S, *et al.* Soil organic carbon density in Hebei Province, China: Estimates and uncertainty. *Pedosphere*, 2005, 15(3): 293~ 300
- [ 5 ] 曹尧东,孙波,宗良纲,等. 丘陵红壤重金属复合污染的空间变异分析. *土壤*, 2005, 37(2): 140~ 146. Cao Y D, Sun B, Zong L G, *et al.* Spatial variation of combined heavy metal pollution of red soils in a hill region (In Chinese). *Soils*, 2005, 37(2): 140~ 146
- [ 6 ] Bishop T F A, McBratney A B. A comparison of prediction methods for the creation of field extent soil property maps. *Geoderma*, 2001, 103: 149~ 160
- [ 7 ] Wood J D, Fisher P F. Assessing interpolation accuracy in elevation models. *IEEE Computer Graphics and Applications*, 1993, 13: 48~ 56
- [ 8 ] Goovaerts P. Geostatistics in soil science: State of the art and perspectives. *Geoderma*, 1999, 89: 1~ 45
- [ 9 ] Journel A G. Geostatistics: Roadblocks and challenges. In: Soares A. ed. *Geostatistics Troia 92*. Dordrecht: Kluwer Academic Publishers, 1993. 213~ 224
- [ 10 ] Armstrong M. Is research in mining geostats as dead as a dodo? In: Dimitrakopoulos R. ed. *Geostatistics for the Next Century*. Dordrecht: Kluwer Academic Publishers, 1994. 303~ 312
- [ 11 ] Goovaerts P, Journel A G. Integrating soil map information in modelling the spatial variation of continuous soil properties. *Eur. J. Soil Sci.*, 1995, 46: 397~ 414
- [ 12 ] Juang K W, Chen Y S, Lee D Y. Using sequential indicator simulation to assess the uncertainty of delineating heavy metal contaminated soils. *Environmental Pollution*, 2004, 127: 229~ 238
- [ 13 ] Deutsch C V, Journel A G. *GSLIB, Geostatistical Software Library and User's Guide*. New York: Oxford Univ. Press, 1998
- [ 14 ] Mowrer H T. Propagating uncertainty through spatial estimation processes for old growth subalpine forests using sequential Gaussian simulation in GIS. *Ecol. Modell.*, 1997, 98: 73~ 86
- [ 15 ] 李保国,胡克林,陈德立,等. 农田土壤表层饱和导水率的条件模拟. *水利学报*, 2002(2): 36~ 40, 46. Li B G, Hu K L, Chen D L, *et al.* Conditional simulation of soil surface saturated hydraulic conductivity at field scale (In Chinese). *Journal of Hydraulic Engineering*, 2002(2): 36~ 40, 46
- [ 16 ] 王学军. 应用转向带法进行土壤铜和铅含量的条件模拟. *应用生态学报*, 2002, 13(12): 1 667~ 1 670. Wang X J. Application of turning band method in conditional simulation of soil heavy metals (In Chinese). *Chinese Journal of Applied Ecology*, 2002, 13(12): 1 667~ 1 670
- [ 17 ] 徐英,陈亚新. 土壤水盐特性空间变异的各向同性近似探讨. *灌溉排水学报*, 2003, 22(4): 14~ 17, 24. Xu Y, Chen Y X. On the approximation of isotropy for spatial variability of soil water and salt (In Chinese). *Journal of Irrigation and Drainage*, 2003, 22

- (4): 14~ 17, 24
- [18] 邵学新, 黄标, 孙维侠, 等. 长江三角洲典型地区工业企业的分布对土壤重金属污染的影响. 土壤学报, 2006, 43(3): 397~ 404. Shao X X, Huang B, Sun W X, *et al.* Effect of industrial distribution on soil heavy metal accumulation in a typical area of the Yangtze River Delta (In Chinese). *Acta Pedologica Sinica*, 2006, 43(3): 397~ 404
- [19] R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3 900051-07-0, URL <http://www.R-project.org>, 2005
- [20] Ribeiro Jr P J, Diggle P J. GeoR: A package for geostatistical analysis. R-NEWS (ISSN 1609 3631), 2001, 1(2): 15~ 18
- [21] Cattle J A, McBratney A B, Minasny B. Kriging method evaluation for assessing the spatial distribution of urban soil lead contamination. *J. Environ. Qual.*, 2002, 31: 1 576~ 1 588
- [22] Goovaerts P. Geostatistics for Natural Resources Evaluation. New York: Oxford Univ. Press, 1997
- [23] Isaaks E H, Srivastava R M. An Introduction to Applied Geostatistics. New York: Oxford Univ. Press, 1989
- [24] 农业环境背景值研究编写组. 农业环境背景值研究. 上海: 上海科学技术出版社, 1997. Agricultural Environment Background Value Research Group. Studies on Agricultural Environment Background Values of China (In Chinese). Shanghai: Shanghai Science and Technology Press, 1997
- [25] Goovaerts P. Geostatistical modeling of uncertainty in soil science. *Geoderma*, 2001, 103: 3~ 26

## UNCERTAINTY ASSESSMENT OF THE SPATIAL PATTERNS OF COPPER CONTENTS IN TOPSOIL IN ZHANGJIAGANG CITY, CHINA

Zhao Yongcun<sup>1</sup> Huang Biao<sup>†</sup> Sun Weixia<sup>1</sup> Shao Xuexin<sup>1,2</sup> Wang Zhigang<sup>1,3</sup>

(1 State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China)

(2 The Research Institute of Subtropical Forestry, CAF, Fuyang, Zhejiang 311400, China)

(3 Graduate School of the Chinese Academy of Sciences, Beijing 100039, China)

**Abstract** The mapping of heavy metal distribution is usually associated with certain uncertainty that needs to be considered in decision making for heavy metal pollution management. The uncertainty in prediction of heavy metal contents at unsampled locations is particularly important because such uncertainty can be propagated into subsequent modelling or decision making process and has fundamental impacts on ultimate results. Based on the Cu content data of a total of 547 topsoil (0~ 15 cm) samples in Zhangjiagang City, the sequential indicator simulation (SIS) were applied to perform a conditional stochastic simulation in this study to quantitatively assess the uncertainty of mapping total Cu content. Results show that prediction of the spatial pattern of soil Cu fluctuated the most in the eastern and central northern parts of the study area, where intensive agriculture and metallurgic and electronic enterprises dominate and the total Cu contents were high, and the least in the southern part, where the total Cu contents were low. The pattern simulated with SIS did not show any smoothing effect. With a given threshold of Cu content, SIS can quantitatively assess both local uncertainty and spatial uncertainty of Cu content that is greater the threshold.

**Key words** Soil mapping; Spatial uncertainty; Stochastic simulation