

土壤微量元素含量空间分布的条件模拟*

张泽浦 王学军

(北京大学城市与环境学系, 北京 100871)

CONDITIONAL SIMULATION FOR SPATIAL DISTRIBUTION OF TRACE ELEMENTS IN SOILS

Zhang Ze-pu Wang Xue-jun

(Department of Urban and Environmental Science, Peking University, Beijing 100871)

关键词 土壤, 微量元素, 条件模拟

中图分类号 S153.61

土壤微量元素的空间分布特征是土壤母质成分和外源输入长期作用的结果, 了解这一分布特征是进行土壤环境评价的重要前提。

地统计学中的条件模拟方法由于同时考虑了分布在二维或三维空间的变量整体上的空间结构特点和局部的不确定性, 因而在研究变量的空间分布、变化和相互联系等方面具有独到之处, 并在有关领域得到了广泛的应用。

本文根据地统计学原理, 对北京东郊污灌土壤表层微量元素含量的小尺度分异进行了条件模拟, 为考察该方法的特点, 同时进行了克里格插值以作对比分析。

1 原理和方法

1.1 原理概述

地质统计学是数学地质的重要分支, 它最早在矿石品位和矿床储量计算中引进了数理统计方法, 应用随机函数研究变量的空间分布规律。

地统计学在其发展早期主要应用于两个方面, 即: 以离差方差、估计方差和半变异函数为基本工具, 描述区域化变量的随机性和结构性, 分析矿床的空间结构, 用于控制生产品位; 以克里格法为基本方法, 给出有限区域内区域化变量的最佳线性无偏估计量, 用于

* 国家教委跨世纪人才培养计划资助。

收稿日期: 1996-10-21; 收到修改稿日期: 1997-06-09

估计实际资源量。

但克里格法作为一种插值方法,不能很好地再现区域化变量的空间结构。1973年, Matheron 提出转向带法,开始了地统计模拟的研究。

假设 $z(x)$ 为所研究的变量 z 在点 x 的值。地质统计学把 $z(x_i)$ 解释为随机变量 $Z(x_i)$ 在点 x_i 的特定的现实。这些自相关的随机变量的集合 $\{Z(x), x \text{ 在研究区域内变化}\}$ 构成一个随机函数。区域化变量 $z(x)$ 可以解释为这一随机函数 $Z(x)$ 的一个现实。这个随机函数可以用一个分布函数和二阶矩(协方差函数或半变异函数)来表征。模拟即是从这个随机函数提取其他的现实,该现实与观测到的现实同构,即具有相同的分布和二阶矩。

假设已有 $z(x)$ 在 k 个点的实测数据 $z_0(x_a), a=1, \dots, k$, 并有 $z(x)$ 的一个模拟现实 $z_s(x)$ 。如果进一步将模拟条件化到实测数据,即要求实测点处的模拟值等于实测值:

$$z_s(x_a) = z_0(x_a), \quad \forall x_a$$

则称该现实为区域化变量 $z(x)$ 的条件模拟,记为 $z_{sc}(x)$ 。相应地称未条件化到实测数据的模拟为非条件模拟,可记为 $z_s(x)$ 。

根据具体模拟方法的不同,条件化有时是在非条件模拟的基础上独立进行,有时是与非条件模拟同时进行。

地统计模拟的突出特点表现在以下几方面:

(1)由于地质、矿产问题本身的特点,地统计模拟是较早提出的针对二维或三维空间变量的数值模拟方法,它关注的核心问题始终是变量值在空间的分布、变化、联系及差异。

(2)模拟现实具有与实测数据相同的频率直方图。更重要的是,模拟现实与真实现实具有相同的空间自相关结构(用自相关函数、自协方差函数或半变异函数描述)。这实际上是构造地统计模拟各种方法的基本要求。因为区域化变量在空间各点的值既有差异又相互影响,形成了特定的空间结构,模拟现实只有再现了真实现实的空间结构,才能在一定程度上代表真实现实,从而具有预测功能。

(3)地统计模拟一般指的是条件模拟,即模拟值以临近点的实测数据为条件:如果模拟点与某一实测点临近,则模拟值与该点的实测值相近;如果模拟点与某一实测点重合,则模拟值与该点的实测值相等。这样,模拟现实不仅空间变化规律上与真实现实相符,而且量值与实际相近,因而对特定问题更具实际意义。

地统计模拟自提出以来发展很快,产生了多种方法,如转向带法、矩阵分解方法、序贯方法等,应用领域也不断拓宽,在石油勘探、采矿、水文地质、工程地质等方面均已有了成功应用的例子^[1-5]。

1.2 模拟过程

本文应用条件模拟方法研究了北京东郊污灌土壤表层元素 Cr、Cu 和 Pb 的含量。对 10×10 米规则网格点进行了 10 次条件模拟,同时进行了克里格插值。模拟点和插值点均为 46 行 73 列共计 3358 个。

所用的模拟方法是以转向带法为基础的。转向带法的基本思想是降维,即不直接模拟二维(或三维)过程,而是先沿若干条直线独立地模拟一维过程,然后迭加得到所需的二

维(或三维)过程。

以下用“结构函数”统称直接或间接反映了变量空间结构的函数(在本文中指自协方差函数)。模拟过程的主要步骤归纳如下:

(1)由实测数据计算待模拟的随机场的结构函数。

(2)求出与多维结构函数相对应的一维结构函数。

(3)在若干条直线上分别独立地模拟符合(2)得出的结构函数的一维过程。本文采用滑动平均进行一维模拟。

(4)将多个一维现实带权叠加,得到多维随机场的模拟现实。

(5)将非条件模拟结果条件化到已知数据。从非条件模拟结果与克里格插值结果的比较中得出每一网格点的误差,加到克里格插值结果上。

(6)结果检验与分析。

由于本文研究的变量存在较明显的漂移,因此用趋势面拟合其漂移部分,然后对残差进行模拟。模拟完成后,将残差模拟值与趋势值相加,得到原始数据的条件模拟。趋势面采用多项式曲面。考虑实际需要和拟合精度两个因素并计算比较了3至7次趋势面之后确定采用五次曲面进行拟合。

本文中需要模拟的随机函数的一元分布近似对数正态分布,因此先把数据进行高斯变换,根据由变换后的数据计算的协方差函数进行模拟,并用变换后的数据对非条件模拟结果进行条件化,最后再对条件模拟结果进行逆高斯变换。

分析变量空间结构的主要工具是半变异函数。本文对三种元素半变异函数的拟合采用球状模型,并把所研究的变量近似视为各向同性。

首先采用滑动平均在 0° 、 30° 、 60° 、 90° 、 120° 、 150° 六条直线上分别独立地产生一维模拟,最后进行转向带迭加。

模拟过程中半变异函数拟合及克里格插值利用了GEO-EAS1.2.1软件,其余计算全部用Visual Basic3.0语言编制程序,在PC机上运行通过。结果图件分别用MSEXcel5.0和Surfer5.01绘制。

2 样品采集与分析

土壤样品采自北京东郊通惠河畔的污灌土壤,该区土壤为潮土,种植水稻和蔬菜,采样区域为 760×500 米的矩形。自60年代以来一直利用通惠河和南北灌溉渠进行污水灌溉。采样按均匀布点方式进行,共采集10行19列共计190个表土样品,南北向采样点间距为50米,东西向采样点间距为40米。

用原子吸收分光光度计测定土样中Co、Cr、Cu、Fe、Mn、Ni、Pb和Zn8种元素的含量。

采样区域位置、采样点分布及分析方法请参见文献^[6]。

3 结果与讨论

以Cr元素为例(Cu、Pb的结果类似)从以下几个方面考察条件模拟结果,同时与克里格插值结果进行对比分析。十次模拟结果是类似的,因此只取其中若干次给出图示。

3.1 微量元素含量空间分布基本趋势

从原始数据、两次条件模拟以及克里格插值结果的平面等值线图和三维立体图(图1)中可以看出,模拟结果的元素含量空间分布基本趋势与原始数据是相符的。进一步的分析证实,在实测点上,模拟值和实测值相等。在非实测点上,不同次的模拟给出略有差异的模拟值,而克里格方法在每一非实测点给出唯一的插值。由于受到空间总体变化趋势的控制,可以推知非实测点上的变量值将处于某一合理范围内,但在各种未知因素的影响下,又难以唯一确定其值。从这一方面来看,不同次模拟结果之间存在的局部波动比克里格方法的唯一插值结果更加合理。这是模拟与估计的一个重要区别。对研究空间的一点(不是控制点),估计值是一定的,模拟值则可以有多,事实上,模拟现实的个数没有任何限制。因此,经常把随机场的模拟现实作为蒙特卡罗模拟或其它模型的输入数据(或参量),由同一数据系列的多个条件模拟现实得到多个模型输出,可计算输出的统计规律,也可分析输入(或参量)与输出的统计关系,空间不确定性对输出的影响,以及模型的灵敏度等。

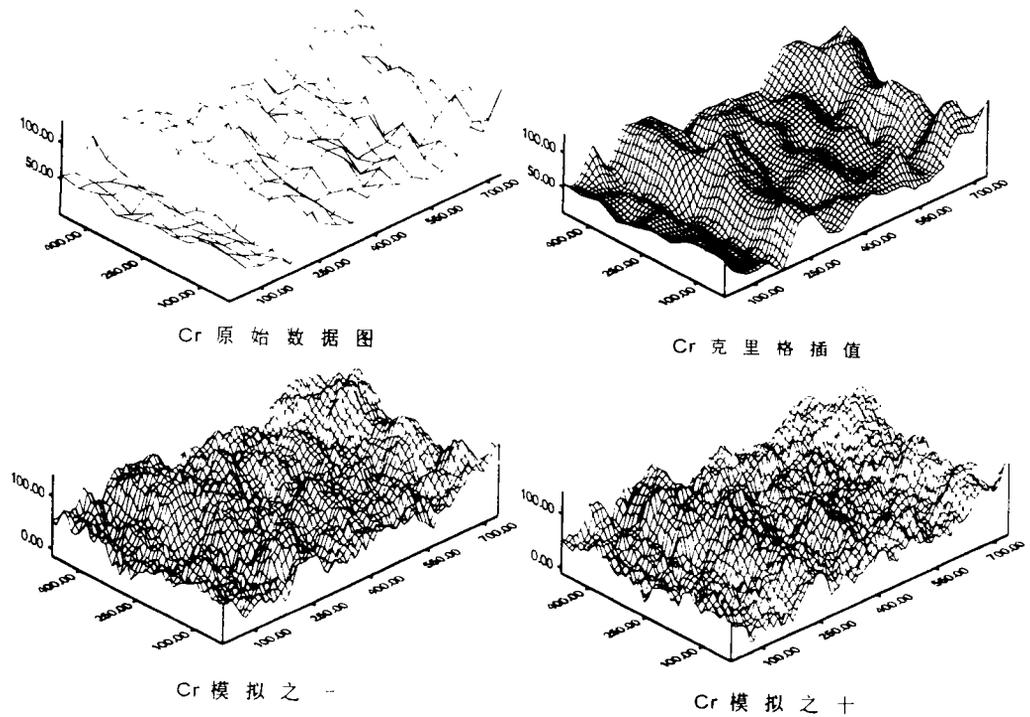


图1 Cr的原始数据、两次模拟结果和克里格插值结果

直观上看,模拟值是符合原始数据的变化趋势的,下文的分析揭示出模拟值和原始值在空间结构上存在更深刻的一致性,即半变异函数相同。从图1可以看出克里格插值结果存在明显的平滑效应,下面的分析通过半变异函数对此给出了更精确的说明。

3.2 基本统计特征

图2显示了描述模拟结果、插值结果以及原始数据离散特征的基本统计量。从图中可以看出,克里格插值的数据范围与原始数据相同;模拟结果极端值范围比原始数据和插值

结果的相应范围明显宽些,但大部分数据所处的范围三者相近;模拟值、插值和原始值三者的均值及中位数相近;另据检验,在控制点上模拟值与克里格插值均与实测值严格相等。以上特点显示出克里格插值的内插特点和条件模拟在实测数据控制下的波动性。

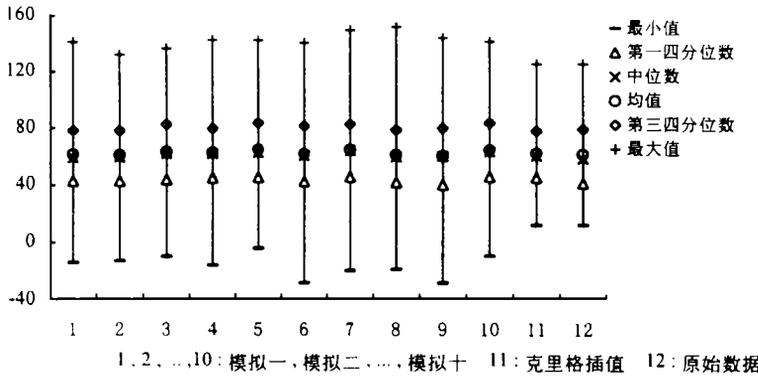


图2 Cr基本统计量

图 3 显示了模拟结果、插值结果以及原始数据的概率分布,可以认为三者是基本吻合的。

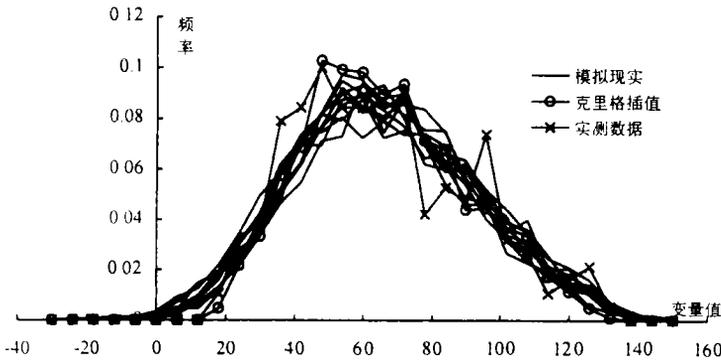


图3 Cr的频率分布

3.3 半变异函数

半变异函数是方向和两点之间距离的函数,定义为两点变量值之差的二阶矩的一半。半变异函数是分析变量空间结构的主要工具,它反映了变量的空间自相关性。对一定的距离,函数值越大,说明变量在该距离上自相关性越弱,即变化性越大。

对离散特征和频率分布的考察说明模拟和插值均有效地再现了原始数据的基本统计特征。但从图 4 和图 5 可以看出,模拟结果的半变异函数与原始数据符合得很好,插值结果的半变异函数明显小于原始数据和模拟结果的半变异函数,表现出平滑效应。

是否符合原始数据的半变异函数,正是模拟与插值的重要不同之处。模拟与估计(插值)是研究区域化变量时常用的两种方法,二者在理论意义和应用领域上均有不同。估计的主要目的是在每一点 x 提出一个尽可能接近未知的真实值 $z_0(x)$ 的估计量 $z^*(x)$, 估计的

质量用无偏性和最小均方误差或者估计方差来衡量;估计有明显的“平滑”效应,改变了变量的空间结构。模拟的目的主要是在模拟值接近真实值的前提下,再现变量的空间变化特点,而模拟值并不是最优估值。因此当空间变化对所研究的问题有重要意义时,就要用模拟值,而不是用估计值。

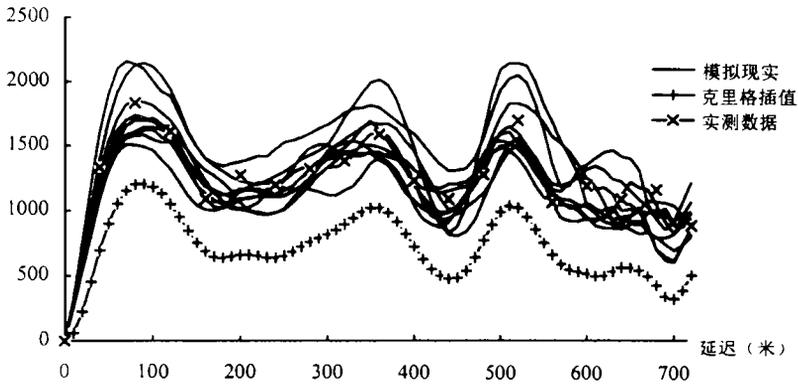


图4 Cr半变异函数(东西方向)

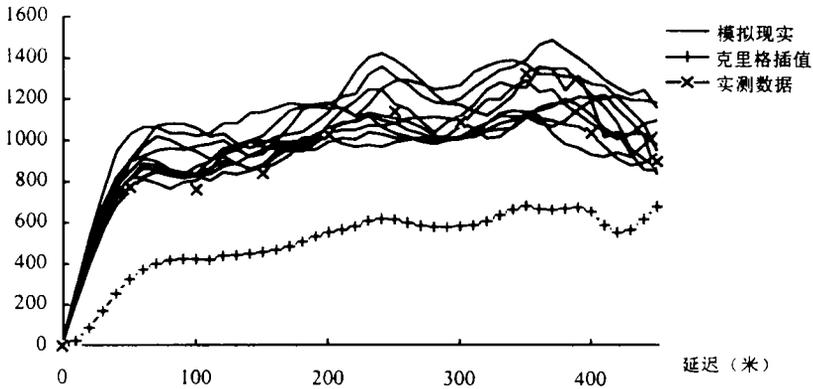


图5 Cr半变异函数(南北方向)

根据上文所述研究分析结果,与克里格等空间插值方法相比,空间模拟方法可以很好地再现土壤微量元素的空间结构特征,从而为揭示土壤微量元素含量的空间结构特征提供了一个有力的工具。

鉴于环境问题的空间属性和环境研究的发展趋势(趋势之一就是描述地表化学物质在二维、三维空间的分布及迁移转化特征),地统计模拟方法在环境研究中有广阔的应用前景。本文只对北京东郊污灌区土壤重金属含量小尺度空间结构特征进行了模拟研究,而实际上,该方法在了解区域大尺度土壤化学物质空间特征方面同样有着明显的优势。

目前,环境污染问题日益引起人们的重视。应用空间模拟手段有助于了解土壤、水体和大气环境中污染物的来源、分布,从而为环境评价、规划及解决有关问题提供技术支持。

参 考 文 献

1. Gomez-Hernandez J J, Srivastava R M. ISIM3D: an ANSI-C three-dimensional multiple indicator conditional simulation program. *Computers & Geosciences*, 1990, 16(4): 395—440
2. Journel A. Geostatistics for conditional simulation of ore bodies. *Economic Geology*, 1974, 69: 673—687
3. Journel A G, Isaaks E H. Conditional indicator simulation: application to a Saskatchewan uranium deposit. *Math. Geol.*, 1984, 16(7): 685—718
4. 周叶, 王家华. 估计渗透率的空间分布的方法. 见: 中国数学地质(3). 北京: 地质出版社, 1991
5. Journel A G, Zhu H. Indicator conditioned estimator. *Trans. Soc. Mining Engineering*, 1989, 286: 1880—1886
6. 王学军, 邓宝山, 张泽浦. 北京东郊污灌区表层土壤微量元素的小尺度空间结构特征. *环境科学学报*, 1996, 待发