

不同插值方法下的土壤重金属污染三维空间分析

陈 诚

上海富嗣工程咨询有限公司，上海

收稿日期：2022年3月11日；录用日期：2022年4月14日；发布日期：2022年4月21日

摘 要

为确定某地块区域内土壤污染物的空间变异与三维分布状况，选择退役某合金制造工厂内土壤重金属Cu为研究对象，对比Kriging、IDW和Nearest Neighbor 3种三维空间插值方法对污染预测的不确定性影响。结果表明，不同插值模型计算结果差异较大。交叉验证结果表明：Kriging模型插值精度最高，其预测结果能较真实地反映实际污染情况；三维可视化方法相较于传统的二维空间分析可以解决二维平面不能直观展示污染物垂直空间变化的问题，可以多方式多角度展示土壤信息，反映土壤内部污染物形态。

关键词

土壤重金属，污染，三维空间插值

Three-Dimensional Spatial Analysis of Soil Heavy Metal Pollution under Different Interpolation Methods

Cheng Chen

Shanghai Fusi Engineering Consulting Co., Ltd., Shanghai

Received: Mar. 11th, 2022; accepted: Apr. 14th, 2022; published: Apr. 21st, 2022

Abstract

In order to determine the spatial variability and three-dimensional distribution of soil pollutants in contaminated sites, the soil heavy metal Cu in a certain alloy manufacturing plant was selected as the research object, and the uncertainty effects of three 3D spatial interpolation methods on

pollution prediction, Kriging, IDW and Nearest Neighbor were compared. The results show that the results of different difference models are large. The cross-validation results show that the Kriging model has the highest interpolation accuracy, and the prediction results can reflect the actual pollution situation. Compared with the traditional two-dimensional spatial analysis, the three-dimensional visualization method can solve the problem that the two-dimensional plane cannot visually display the vertical spatial changes of pollutants, and can display soil information in multiple ways and at multiple angles, reflecting the form of pollutants in the soil.

Keywords

Heavy Metals in Soil, Pollution, Three-Dimensional Spatial Interpolation

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在我国经济飞速发展的状态下,各种途径而来的重金属污染物质相继侵入到土壤中,从而导致大区域土壤重金属元素的聚集,严重影响着人们的健康生存与未来发展。土壤中积累的重金属亦会通过食物链富集最终对人类的健康产生危害,严重地威胁居民的身体健康[1][2]。当退役工厂,用地性质发生改变的时候,为了防止带来新的环境问题,非常有必要开展退役工业地块调查评估工作,并对调查评估发现的高风险污染地块进一步开展修复治理工作[3]。

目前国内外环境工作者、学者针对土壤污染物空间分布模拟方面开展了大量研究[4][5],现有研究都是将整个土壤按照一定的深度划分为若干个水平土层进行比较分析,忽略了不同深度土壤层之间的相互影响,没有模拟土壤整体的变异特征。因此,必须通过一种能同时在水平、垂直方向上进行插值的方法以解决以上问题。而三维的统计学方法无疑是最具潜力的方法之一,目前已有学者将 Kriging、IDW 和 Nearest Neighbor 等确定型三维插值方法应用到土壤盐分、土壤重金属污染等相关领域的三维数字土壤制图研究中。但对土壤重金属进行三维可视化研究并不多,相比较于传统的水平面的空间变异理论研究,三维污染物模型可以解决二维平面不能直观展示污染物垂直空间变化的问题,可以多方式多角度展示土壤信息,反映土壤内部污染物形态。对研究土壤污染物三维分布特征,确认污染范围与污染源,为日后污染治理与修复技术提供理论支撑,具有一定的理论意义与实际意义。本文拟以某退役合金制造工厂土壤 Cu 污染为研究案例,在分析污染物空间结构特征基础上,对比研究常用的三维克里格、反距离加权和邻近点模型在不同特征参数设置下的预测精度,并对土壤 Cu 含量进行三维可视化,探究其空间分布特征,为指导该地块的修复治理提供依据,以期对日后该类型地块风险评估与土壤重金属修复治理等工作提供科学依据。

2. 材料与方法

2.1. 样品采集与分析

本文样品采集使用系统布点与经验布点相结合的原则,结合研究区水文地质资料以及现场实际状况,将研究区域由上至下分为三层(0~1 m、1~3 m、3~6 m)进行土壤样品的采集,采样深度为 0~6 m,共布设采样点 44 个,钻孔样点空间分布图如图 1 所示,每个钻孔按照分层深度采样,每层样不同位置取样品 3

个, 现场利用光离子化检测器(PID)进行样品筛选, 选取本层浓度最高的土样装入 250 mL 直口玻璃瓶中密封, 填写标签, 放入装有低温蓝冰的保存箱中送实验室进行检测。土壤中 Cu 含量采用 USEPA 6020A-2007 方法进行测定[6]。

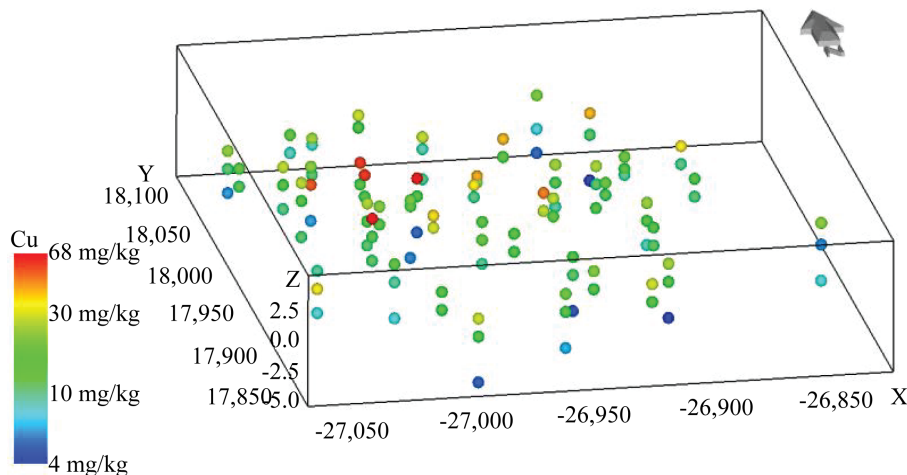


Figure 1. Sampling point spatial distribution
图 1. 采样点空间分布

2.2. 三维插值模型

目前常见的三维空间插值方法有三维克里格插值法(Kriging)、反距离加权插值法(IDW)、最邻近点法(Nearest Neighbor) [7], 这些插值模型已经在土壤养分、土壤盐分、土壤电导率、土壤污染物等领域有应用。

三维克里格插值法是二维克里格插值在三维空间中的扩展, 三维克里格必须在三维空间中采样, 并在水平(X-Y 方向)和垂直方向(Z 方向)上的半方差函数进行结构分析, 选取模型拟合最优的半方差函数, 组合水平和垂直方向上的半方差函数构建三维半方差函数[8], 要求数据满足二阶平稳假设, 其变量的平均值为已知的常数, 计算公式为

$$r(h_3) = \frac{1}{2N(h_3)} \sum_{i=1}^{N(h_3)} [Z(x_i) - Z(x_i + h_3)]^2 \quad (1)$$

式中 h_3 为两样点之间的空间距离, $N(h_3)$ 是以 h_3 为间距的所有样本对数, $Z(x_i)$ 和 $Z(x_i + h_3)$ 分别是点 x_i 和 $x_i + h_3$ 处的值。

通过半方差函数分析后, 如果样点数据存在空间自相关性, 可采用三维克里格进行空间插值, 公式为

$$Z(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (2)$$

式中 $Z(x_0)$ 为 x_0 处预测值, n 为预测点周围已知样点数, λ_i 为各样点的权重, $Z(x_i)$ 为 x_i 处实测值。

反距离加权插值(IDW)最早是由 Shepard 提出, 是一种多元插值的方法, 与克里格插值有些类似, 可统一用公式(2)表示, 但是反距离加权插值在赋权重时只考虑未知点与已知样点间的距离, 而克里格插值不仅要考虑距离, 还需要通过半方差函数和结构分析, 考虑到已知点的空间分布和与未知点的空间位置关系, 反距离加权插值要求样点分布要尽可能均匀, 且布满整个插值区域。

最邻近点法(Nearest Neighbor)又称泰森多边形方法,在地理分析中常使用该方法对区域进行快速赋值[9],一般当空间数据较多且均匀间隔分布或只有少数点没有取值时,用最邻近点法可以空间布点。在搜索半径内,选择距离估计点最近的已知点作为估计值。其权重系数为常数

$$h_j = \begin{cases} 1, & h = \min(h_1, h_2, \dots, h_3) \\ 0, & \text{else} \end{cases} \quad (3)$$

式中 h_j 在二维插值中为 $h_j = \sqrt{(x-x_j)^2 + (y-y_j)^2}$, 在三维空间插值中为

$$h_j = \sqrt{(x-x_j)^2 + (y-y_j)^2 + (z-z_j)^2}。$$

2.3. 插值精度评价

本次研究采用常用的交叉验证来评各种插值模型的计算精度,采用均方根误差(RMSE)与平均误差(ME)作为误差的统计指标, RMSE 越小; ME 越接近 0 插值精度越高。

$$ME = \frac{1}{n} \sum_{i=1}^n [u(x_i) - u^*(x_i)] \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [u(x_i) - u^*(x_i)]^2} \quad (5)$$

式中: $u(x_i)$ 为预测值, $u^*(x_i)$ 为原始样点值。

3. 结果与讨论

3.1. Cu 含量数据统计特征分析

对土壤中 Cu 含量进行描述分析,得出各项统计指标,各层土壤中 Cu 含量描述性统计结果如表 1 所示。从表 1 中可以看出,各层土壤中 Cu 含量的平均值变化范围为 7.26 mg/kg~21.66 mg/kg,每层样点含量的最小值与最大值差异较大,从表中 Cu 的最小值、最大值以及平均值可以看出,第一层 Cu 含量最高,随着深度的增加,污染物浓度降低,表明最上层土壤污染最为严重,三层土壤中 Cu 的变异系数分别为 0.39、0.42、0.30,均大于 0.1,反映土壤中 Cu 的空间分布呈中等变异性,空间变异特征较强,而且空间分异在垂直方向上差异不大。

Table 1. Statistical analysis of Cu content in each layer of soil

表 1. 各层土壤中 Cu 含量统计特征分析

分类统计	最小值 mg/kg	最大值 mg/kg	平均值 mg/kg	峰度	偏度	变异系数	标准差 mg/kg
第一层	8.50	46.30	21.66	1.94	1.45	0.39	8.49
第二层	5.52	36.70	12.85	8.34	2.45	0.42	5.41
第三层	3.88	11.90	7.26	-0.98	1.30	0.30	2.23

3.2. 土壤 Cu 元素的空间结构

二阶平稳假设意味着两个样点之间的差异依赖于两者的方向和距离,与具体位置无关[10],所以二维水平面的空间变异分析同样适用于三维空间,本文利用地统计分析软件 GS+对研究区内土壤重金属 Cu 进行半方差函数分析,研究土壤中 Cu 含量在空间上的变异特征,拟合最优半方差函数,首先需要对数据进行正态检验,要求数据因符合正态分布,否则会存在比例效应[11]。在 SPSS 数据统计软件中进行正态

分布检验, 样本总数为 112, 选择 K-S 检验, 结果显示, 第一层第三层土壤铜含量数据符合正态分布要求($P > 0.05$)。第二层土壤铜含量数据不符合正态分布, 经过对数函数变换后 P 值为 0.82 大于 0.05。选择决定系数最大和残差最小的为最优模型, 经计算得到模型的参数、决定系数和残差如表 2 所示。

Table 2. Theoretical model of soil Cu semivariance function and its related parameters

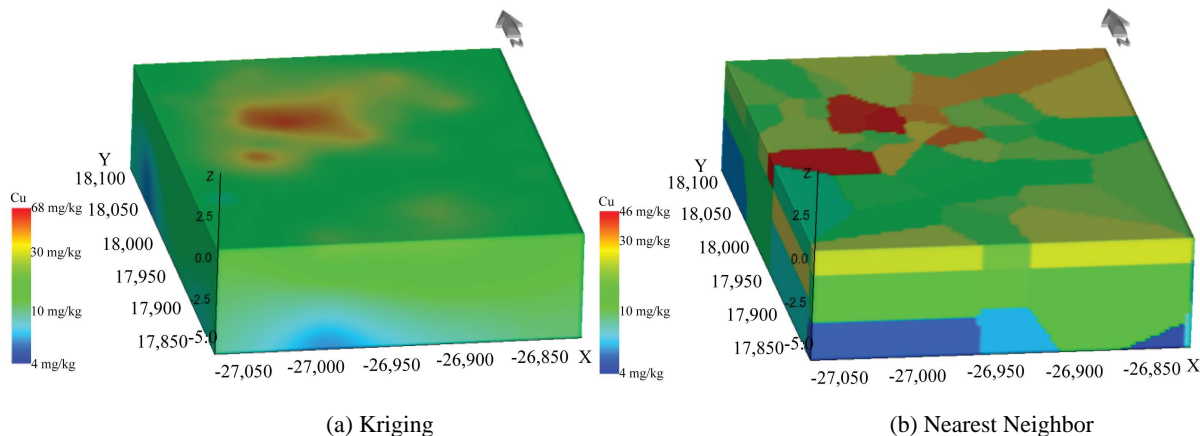
表 2. 土壤 Cu 半方差函数理论模型及其相关参数

层位	模型	C_0	$C_0 + C$	A (m)	$C_0/C_0 + C$	RSS	R^2
第一层	球状	68.67	90.71	133.17	0.76	422	0.47
第二层	球状	9.49	18.99	204.8	0.5	115	0.42
第三层	线性	3.36	6.721	55.7	0.08	6.78	0.41

通过对研究区内不同深度的 Cu 含量进行半方差分析, 得到土壤 Cu 含量在空间中的变异特征。表 2 中空间变异分析表明研究区内不同深度的土壤 Cu 含量空间相关性差异明显, 块金值与基台值的比值表明了变量的空间相关程度, 也就是由随机性因素引起的空间变异性占系统总变异的比, 又称块金系数。块金系数小于 25%, 说明系统具有强烈的空间相关性; 块金系数在 25%~75%之间, 说明它们的空间变异主要是由随机因素和结构因素共同作用引起的, 具有中等强度的空间相关性; 大于 75%说明系统空间相关性较弱[12]。由表 2 的比值可知, 研究区内第一层重金属 Cu 块金系数 76%, 其值大于 75%说明研究区表层土壤 Cu 空间变异主要是由随机因素引起的, 空间相关性较弱; 第二层块金系数为 50%, 其值在 25%~75%之间, 说明在研究区内 3 m 深处土壤 Cu 空间变异是由结构因素和随机因素共同作用引起的, 具有中等强度空间相关性; 第三层块金系数为 8%, 其值小于 25%, 表明在研究区内 6 m 深处土壤 Cu 空间变异主要是由结构因素作用, 具有较强的空间相关性。随机性因素主要包括施肥、灌溉、交通、工业生产等人为活动因素, 而结构性因素则包括气候、母质、土壤类型、地形等自然因素。

3.3. 不同插值模型 Cu 的三维分布

通过不同插值模型, 得到土壤重金属 Cu 的三维空间分布(见图 2), 图中可见, 三种插值模型下土壤 Cu 污染总体分布范围一致, 污染较重的区域主要集中在研究区中部偏西北位置, 该区域原为企业车屑锭堆场与堆焊车间, 主要用作合金材料的堆放与焊接加工, 其他区域多为空地或企业办公室与成品车间, 污染程度相对较轻。



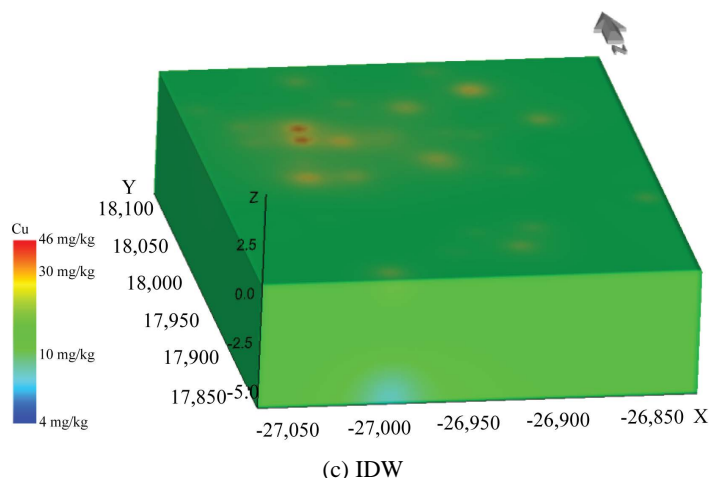


Figure 2. Three-dimensional spatial distribution of heavy metal Cu pollution
图 2. 重金属 Cu 污染三维空间分布图

从三种不同插值结果图可以看出，三种插值结果差异较大，krig 模型在预测过程中考虑了污染物含量空间结构信息，能够较好的反映出污染物三维空间分布特征，IDW 模型在三维空间的插值结果为不连续的独立的“孤岛”状。对比 krig 与 NN 模型在相同参数值下的插值结果，IDW 模型的结果相对“保守”一些，重度污染区相对略小。而 NN 模型具有明显的锅齿，棱角较为明显。在 krig 模型的基础上对重污染区域垂直方向进行切片(见图 3)，截面纵贯研究区南 - 北、东 - 西两个方向，根据截面的 Cu 分布信息，污染分布主要在上层土壤，随着深度的增加，土壤 Cu 含量逐渐减小，污染物从地表核心污染区向地下四周迁移扩散。研究区东部和南部土壤 Cu 含量变化不明显，且含量不高。

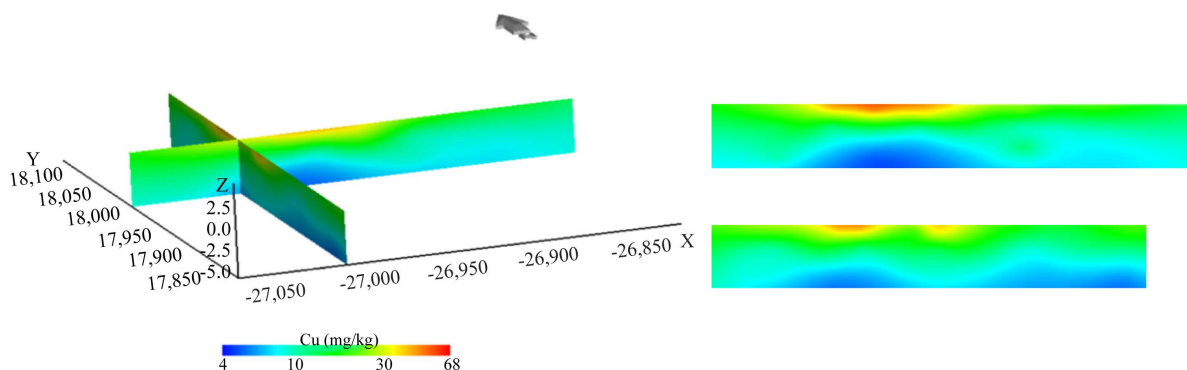


Figure 3. Longitudinal spatial distribution of heavy metal Cu
图 3. 重金属 Cu 纵断空间分布

3.4. 预测精度检验

不同插值模型可以得到不同的三维污染分布表征，但是不同插值模型精度和插值效果都不尽相同，本文选用交叉验证中常见的平均误差(ME)与均方根误差(RMSE)来分析不同插值模型的预测精度，交叉验证结果如表 3 所示，IDW 模型和 NN 模型在插值过程中只依据数据的几何结构特征进行计算，不能很好的反映出变量在空间上的变异特征。因此，克里格插值模型更适合模拟土壤重金属 Cu 的三维空间分布特征，这与刘庚等[13]和韩霁昌[14]等的研究结果相似。

Table 3. Prediction errors of different interpolation models
表 3. 不同插值模型的预测误差

	插值模型	第一层	第二层	第三层
均方根误差(RMSE)	Kriging	5.25	7.89	3.33
	IDW	7.72	5.55	4.92
	NN	8.31	5.44	4.85
平均误差(ME)	Kriging	0.17	0.33	0.18
	IDW	0.21	0.42	0.19
	NN	-1.12	-0.22	-0.23

4. 结论

插值结果表明,该地块内土壤重金属 Cu 含量主要呈“岛”状分布,高值区集中在中部偏北的生产车间与废金属堆场,从垂直方向上看,随着深度增加,污染减轻。从不同模型的三维空间分布预测精度与土壤重金属 Cu 污染评价结果来看, Kriging 模型优于 IDW 模型和 NN 模型,前者对污染物范围的预测与污染实际状况差异最小,适合该地块土壤 Cu 的三维污染评价。

参考文献

- [1] 郑黎明,袁静. 重金属污染土壤植物修复技术及其强化措施[J]. 环境科技, 2017, 30(1): 75-78.
- [2] 祁迎春,应婷,田浩,等. 陕北公路旁土壤、农作物重金属污染特征与评价[J]. 环境科技, 2016, 29(3): 52-57.
- [3] 周以富,董亚英. 几种重金属土壤污染及其防治的研究进展[J]. 环境科学动态, 2003(1): 15-16.
- [4] 唐小亮,吴以中,张瑜. 氯丹和灭蚁灵在典型污染地块的空间分布研究[J]. 土壤通报, 2012, 43(4): 942-948.
- [5] Liu, G., Niu, J., Guo, W., *et al.* (2017) Assessment of Terrain Factors on the Pattern and Extent of Soil Contamination Surrounding a Chemical Industry in Chongqing, Southwest China. *Catena*, **156**, 237-243. <https://doi.org/10.1016/j.catena.2017.04.005>
- [6] USEPA (United States Environmental Protection Agency) (1996) Method 3050B: Acid Digestion of Sediments Sludges and Soils.
- [7] 巍文侠,宋博宇,李培中,等. 三维可视化建模方法在污染地块中的应用[J]. 环境工程技术学报, 2016, 6(4): 384-390.
- [8] Pan, G., Moss, K., Heiner, T., *et al.* (1992) A Fortran Program for Three Dimensional Cokriging with Case Demonstration. *Computers & Geosciences*, **18**, 557-578. [https://doi.org/10.1016/0098-3004\(92\)90094-8](https://doi.org/10.1016/0098-3004(92)90094-8)
- [9] Triantafyllis, J., Odeh, I.O.A. and McBratney, A.B. (2001) Five Geostatistical Models to Predict Soil Salinity from Electromagnetic Induction Data across Irrigated Cotton. *Soil Science Society of America Journal*, **65**, 869-878. <https://doi.org/10.2136/sssaj2001.653869x>
- [10] 李洪义. 滨海盐土三维土体电导率空间变异及可视化研究[D]: [博士学位论文]. 杭州: 浙江大学, 2008.
- [11] 王政权. 地统计学及在生态学中的应用[M]. 北京: 科学出版社, 1999.
- [12] Webster, R. and Norteliff, S. (1984) Improved Estimation of Micronutrients in Hectare Plots of the Sonning Series. *European Journal of Soil Science*, **35**, 667-672. <https://doi.org/10.1111/j.1365-2389.1984.tb00623.x>
- [13] 刘庚,毕如田,权腾,等. 某焦化地块土壤中多环芳烃分布的三维空间插值研究[J]. 生态学报, 2014, 34(11): 2876-2883.
- [14] 韩霖昌,李晓明,孙剑虹,等. 卤泊滩典型田块土壤盐分三维空间分布研究[J]. 自然资源学报, 2014, 29(5): 847-854.