

谢云峰, 陈同斌, 雷梅, 等. 2010 空间插值模型对土壤 Cd 污染评价结果的影响 [J]. 环境科学学报, 30(4): 847–854
Xie Y F, Chen T B, Lei M, *et al*. 2010 Impact of spatial interpolation methods on the estimation of regional soil Cd [J]. Acta Scientiae Circumstantiae 30(4): 847–854

空间插值模型对土壤 Cd 污染评价结果的影响

谢云峰^{1, 2}, 陈同斌^{1*}, 雷梅¹, 郑国砥¹, 宋波¹, 李晓燕¹

1 中国科学院地理科学与资源研究所环境修复中心, 北京 100101
2 中国科学院研究生院, 北京 100049
收稿日期: 2009-07-28 修回日期: 2009-09-27 录用日期: 2010-01-28

摘要: 分析和比较了不同插值模型 (反距离加权、局部多项式、普通克里格和径向基函数) 对土壤 Cd 空间插值、污染面积估算和污染区空间分布的影响, 结果表明, 插值方法对极大值都存在不同程度的平滑, 交叉验证的 Cd 极大值插值的相对误差高于 45 %。不同插值模型估算的污染区域面积在 0~2.12% 之间。根据样点超标率估算的污染面积比插值模型估算的面积高 1.53%~3.65%。污染评价结果不确定性较大的区域主要是在局部极大值区域和高值向低值的过渡区域。因此, 土壤重金属的空间插值精度对污染评价结果有很大影响, 开展土壤重金属污染调查时应注意在浓度过渡区域加大采样密度。

关键词: Cd 空间插值; 污染评价; 极大值; 面积估算; 区域土壤环境质量

文章编号: 0253-2468(2010)04-847-08 中图分类号: X144 文献标识码: A

Impact of spatial interpolation methods on the estimation of regional soil Cd

XIE Yunfeng^{1, 2}, CHEN Tongbin^{1*}, LEI Mei¹, ZHENG Guodi¹, SONG Bo¹, LIXiaoyan¹

1 Center for Environmental Remediation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101
2 Graduate University of Chinese Academy of Sciences, Beijing 100049

Received 28 July 2009; received in revised form 27 September 2009; accepted 28 January 2010

Abstract The distribution of Cd in surface soil of the Tongzhou district (Beijing (605 km²)) was interpolated with 4 spatial interpolation methods: Inverse Distance Weighted, Local Polynomial, Ordinary Kriging and Radial Basis Function. The accuracy of every method was compared according to the extreme value prediction, the impact on area estimation and the spatial distribution of high Cd area. The cross-validation results showed that all of the methods underestimated the extreme value of Cd, and the mean relative error (MRE) was higher than 45.1%. The cadmium pollution areas calculated by different methods are from 0 to 2.12%, which were from 1.53% to 3.65% less than those estimated by sample statistics. The differences (or uncertainties) of the interpolation results among the models were mainly located in the region of extreme high concentration and the transitional region from high Cd concentration to low Cd concentration. Therefore, a proper interpolation model has a significant impact on the results of the spatial distribution estimation of soil heavy metals and that of soil pollution evaluation. While surveying the soil heavy pollution, density sampling in the soil heavy metal concentration transitional region is advised.

Keywords Cd spatial interpolation; pollution evaluation; extreme value; pollution area calculation; precision assessment; regional soil environmental quality

1 引言 (Introduction)

土壤重金属污染评价过程中受采样方式和成本的限制, 很难做到高密度的采样, 因此, 在操作中

必须根据有限的、离散点绘制区域土壤重金属空间分布图 (Bishop *et al*, 2001)。在土壤重金属污染评价中应用较多的空间插值方法有反距离加权法、克里格插值法、样条函数法、多元回归法、径向基函数

基金项目: 中国科学院知识创新工程重大项目 (No. KZCX1-YW-06-03); 资源与环境信息系统国家重点实验室开放基金; “十一五”国家科技支撑计划课题 (No. 2007BAC28B04)
Supported by the Knowledge Innovation Program of the Chinese Academy of Sciences (No. KZCX1-YW-06-03), the Foundation of LREIS and the Project in the National Science & Technology Pillar Program during the “Eleventh Five-year Plan” Period (No. 2007BAC28B04)
作者简介: 谢云峰 (1981—), 男, E-mail: xieyf07b@igsnrr.ac.cn * 通讯作者 (责任作者), E-mail: chentb@igsnrr.ac.cn
Biography: XIE Yunfeng (1981—), male, E-mail: xieyf07b@igsnrr.ac.cn * Corresponding author E-mail: chentb@igsnrr.ac.cn
© 1994-2011 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

法等 (Cheng *et al.*, 2007; Imperato *et al.*, 2003; McGrath *et al.*, 2004; 郑袁明等, 2003; McBratney *et al.*, 2003). 插值模型的选择是土壤重金属污染评价的基础. 但是, 从现有不同人的研究报道发现, 关于不同插值方法的预测精度存在着较大的差异. Gotway等 (1994)在进行土壤有机质和硝酸根离子含量制图时, 发现反距离加权法 (Inverse Distance Weight, IDW) 的插值精度要高于克里格方法. Wollenhaupt等 (1994)利用反距离加权和克里格法进行土壤 P 和 K 含量制图时, 也发现 IDW 的精度更高. Yasreb 等 (2009)比较了普通克里格法 (OK) 和反距离加权法 (IDW) 对土壤化学性质的插值精度, 认为 OK 模型要优于 IDW. Panagopoulos 等 (2006)比较了 OK、IDW 和泰森多边形对土壤总矿物氮、磷、钾、pH、电导率和土壤饱和水等的插值精度, 交叉验证的结果表明, OK 对每种土壤属性都是最好的方法.

土壤重金属污染评价中, 重点关注的是高污染风险区域的识别. 高污染风险区域的样点在空间上常表现为局部极大值. 插值模型精度的评价方法通常为交叉验证的均方根误差 (RMSE), RMSE 是对总体期望值插值精度的评价, 不能反应模型对极大值 (污染区) 的插值精度. 同时, 土壤重金属元素通常呈对数正态分布, 在空间上变异性较大. 因此, 关于土壤养分含量 N、P、K 等的插值精度评价结果不能直接推广到土壤重金属插值. 现有的土壤重金属插值方法评价研究侧重于对总体均值的预测精度评价, 关于插值方法对土壤重金属污染细节信息的平滑效应和污染评价结果的不确定性还缺乏系统研究.

本文的研究目的是: 以北京市通州区土壤 Cd 为例, 比较不同插值方法对土壤 Cd 含量极大值的预测精度, 不同插值方法对污染区面积估算、污染区空间分布和评价结果的不确定性.

2 材料与方法 (Materials and methods)

2.1 采样与分析

研究区位于北京市通州区中部, 面积约 605 km². 主要土壤类型为褐潮土、砂姜潮土. 农业种植以玉米和小麦为主. 从 20 世纪 60 年代开始, 大部分农田利用北京市的生活污水进行灌溉.

根据研究区域的特点, 采用均匀布点采样的原则在整个研究区域内共采集土壤样品 137 个, 取样

深度 0~20 cm. 土样经室内风干, 过 100 尼龙网筛. 土壤样品采用美国国家环保局 (USEPA) 方法消煮, 石墨炉-原子吸收光谱仪 (AAS Vario 6) 测定 Cd 分析过程所用试剂均为优级纯, 所用的水均为亚沸水, 分析过程均加入国家标准土壤参比物质 (GSS-1) 进行分析质量控制, 分析样品的重复数为 10%~15%. 根据实验结果, 分析中 Cd 的回收率为 87%~108% (国标要求为 86%~114%), 优于国家标准参比物质 GSS-1 给定的数值, 结果符合质控要求 (郑袁明等, 2003; 陈同斌等, 2004).

2.2 空间插值模型 (Spatial interpolation model)

2.2.1 反距离加权法 (Inverse Distance Weight, IDW) 反距离加权法是常用的空间内插方法之一. 它是以插值点与样点间的距离为权重的一种加权平均法, 离插值点越近的样本点赋予的权重越大. 距离加权系数的选择会显著影响插值结果.

2.2.2 局部多项式 (Local polynomial, LP) 插值 多项式回归时通过使用全局多项式逼近源数据点来提供内插. 局部多项式是将全局多项式方法和移动平均过程结合起来的一种内插方法. 局部多项式需要应用最小二乘多项式拟合数据, 通常选择一次、二次或者三次多项式. 局部多项式只用于拟合一个窗口定义的局部样点区域, 通过多个多项式作为局部方程式来拟合研究区域, 每个多项式都处在特定重叠的邻近区域内. 根据已知样点拟合的多项式表达式和插值点的坐标, 就可以计算插值点的变量值.

2.2.3 克里格 (Kriging) 插值 克里格插值法是以变异函数理论和结构分析为基础, 在有限区域内对区域化变量进行无偏最优估计的一种方法. 在克里格插值法中, 一个待估点的变量值就是其周围影响范围内的 n 个已知变量值的线性组合 (Isaaks *et al.*, 1990; Webster *et al.*, 2007).

针对不同的情况和目的, 克里格插值法有多种变化. 当变量满足二阶平稳假设时, 适用普通克里格法; 对数正态克里格法适用于对数正态分布的数据; 协同克里格法适用于刻画多个变量的系统区域变化; 指示克里格法适用于有特异值的数据.

2.2.4 径向基函数法 (Radial Basis Functions, RBF) 径向基函数法是用于一系列精确内插算法的统称. 径向基函数插值法如同将一个软膜插入并经过各个已知样点, 同时又使表面的总曲率最小. 径向基函数法与地统计插值方法具有相似性, 它的

优点是不需要有关样本数据的任何假设, 但是不具有地统计插值的变差图分析功能 (Smith *et al*, 2007). 常用的径向基函数有: 规则样条函数 (Completely regularized spline)、高次曲面函数 (Multiquadric functions)、薄板样条函数 (Thin plate spline)和张力度函数 (Spline with tension)等.

2.2.5 空间插值方法及参数设置 选取常用的反距离加权法 (IDW)、局部多项式 (LP)、普通克里格法 (OK)和径向基函数法 (RBF)进行土壤 Cd空间插值. 为了研究模型参数对污染评价结果的影响, IDW方法的距离加权系数选择 1、2、3、4; LP的回归次数为一次、二次和三次; OK模型的半变异函数选择 Circular(圆形)、Exponential(指数)、Gaussian(高斯)和 Spherical(球状) 4种; RBF方法选择 CRS、MQ、MQ-ST和TPS五种核函数. 普通克里格法要求插值要素符合正态分布, 本研究中的土壤 Cd含量符合对数正态分布, 所以插值过程中, 首先对数据进行对数变换. 由于土壤 Cd在空间上存在一定的分布趋势, 因此, 进行普通克里格插值时, 利用全局二次方程式拟合趋势. 详细的参数设置见表 1.

表 1 空间插值方法及参数设置
Table 1 Parameters of the interpolation methods for prediction map of Cd

插值方法 简称	插值 方法	参数设置		
		加权系数	搜索半径	
IDW 1	IDW	Power = 1	临近 15个点, 最少 10个	
IDW 2	IDW	Power = 2	临近 15个点, 最少 10个	
IDW 3	IDW	Power = 3	临近 15个点, 最少 10个	
IDW 4	IDW	Power = 4	临近 15个点, 最少 10个	
		回归系数	搜索半径	
LP1	LP	Power = 1	最少临近 10个点	
LP2	LP	Power = 2	最少临近 10个点	
LP3	LP	Power = 3	最少临近 10个点	
		半变异函数 模型	正态 转换	空间 趋势 搜索半径
OK-CIR	OK	Circular	对数	二次 临近 10个点
OK-EXP	OK	Exponential	对数	二次 临近 10个点
OK-GS	OK	Gaussian	对数	二次 临近 10个点
OK-SPH	OK	Spherical	对数	二次 临近 10个点
		核函数	搜索半径	
RBF-CRS	RBF	规则样条 (CRS)	临近 15个点, 最少 10个点	
RBF-MQ	RBF	反高次曲面 (MQ)	临近 15个点, 最少 10个点	
RBFMQ	RBF	高次曲面 (MQ)	临近 15个点, 最少 10个点	
RBF-ST	RBF	张力样条 (ST)	临近 15个点, 最少 10个点	
RBF-TPS	RBF	薄板样条 (TPS)	临近 15个点, 最少 10个点	

2.3 精度评价方法

本文主要从插值方法对 Cd含量总体和极大值的预测精度, 污染区面积估算结果和污染区域的空间分布等 3个方面来评价不同方法的插值效果.

2.3.1 空间插值精度评价 插值方法的总体精度和极大值预测精度都是利用交叉检验法 (Cross-validation)进行评价. 交叉检验法先假定每一个采样点的含量值未知, 利用周围样点的值来估算, 然后计算估计值与实际测定值的误差, 根据误差统计结果评估插值方法的优劣 (Robinson *et al*, 2006). 常用的误差统计指标有平均误差 (ME), 平均绝对误差 (MAE), 平均相对误差 (MRE)和均方根误差 (RMSE). ME越接近于 0插值误差越小; MAE和RMSE的值越小, 精度越高; MRE可以克服量纲的影响, MRE越小, 插值精度越高.

$$ME = \frac{1}{n} \sum_{i=1}^n (z(x_i) - z^*(x_i)) \tag{1}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |z(x_i) - z^*(x_i)| \tag{2}$$

$$MRE = \frac{1}{n} \sum_{i=1}^n |z(x_i) - z^*(x_i) / z(x_i)| \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (z(x_i) - z^*(x_i))^2} \tag{4}$$

式中, $z(x_i)$ 为预测值, $z^*(x_i)$ 为原始采样值.

2.3.2 污染区域面积和空间分布比较 交叉验证法是用来评价样点的插值误差, 不能反映插值误差在空间上的分布特征. 本文利用 ArcGIS Spatial Analysis的栅格分析方法比较不同模型估算的 Cd污染区的面积和空间分布的差异.

2.4 数据处理方法和工具

土壤重金属含量数据处理和统计分析软件为 MS Excel 2003和 Minitab 14, 空间结构分析和半变异函数采用 VARIOUN 2.21和 GS+ 7.0, 空间插值、污染制图和空间统计采用 ESRI ArcGIS 9.3

3 结果 (Result)

3.1 土壤 Cd含量和空间结构特征

3.1.1 Cd的统计特征 研究区土壤 Cd含量介于 0.03~0.39mg·kg⁻¹之间, 平均值为 0.109mg·kg⁻¹. 土壤 Cd总体上符合对数正态分布 (图 1). Cd的变异系数为 56.27%, 表明研究区 Cd的空间变异性较大.

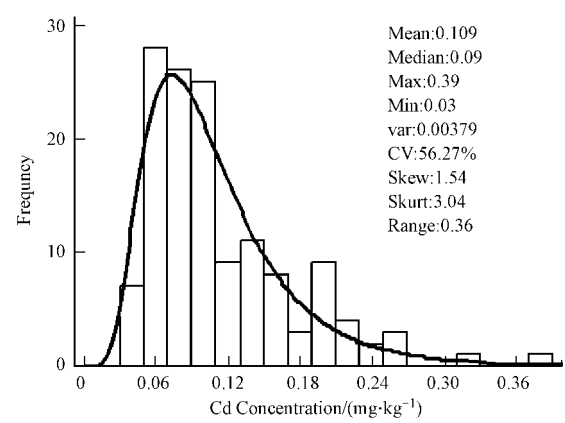


图 1 土壤 Cd 含量描述统计和直方图

Fig 1 The discriptive statistics and histogram of Cd

3 1 2 Cd的空间结构特征 从 Cd的半方差图可知(图 2), Cd的空间结构性较差, 空间上存在着漂移现象. 最佳半变异拟合函数为球状模型, 拟合的块金值(C_0)为 0 00142, 基台值(C_0+C)为 0 0033, 变程为 3300 m, 决定系数 R^2 为 0 034, 块金值/基台值为 46 06%, 表明 Cd具有中等的空间相关性.

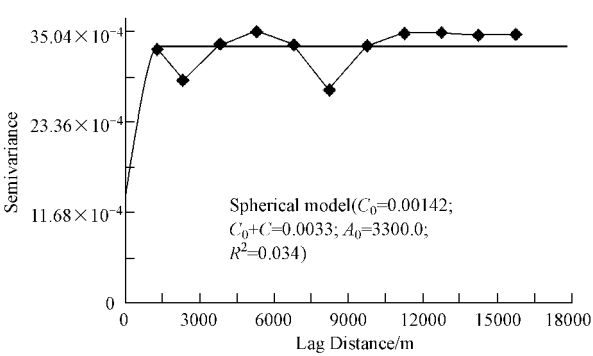


图 2 土壤 Cd 的半方差图和半变异函数拟合模型

Fig 2 The variogram and fitted model of Cd

3 2 不同插值方法的插值精度

从 RMSE 的统计结果(表 2)来看, RBF-MQ 方法的 RMSE 最小(0 059 mg kg^{-1}), LP3 的 RMSE 最大(0 116 mg kg^{-1}). DW 方法的距离加权系数越大, RMSE 越大; LP 的回归系数越大, RMSE 越大; OK 插值, 不同半变异拟合模型的 RMSE 均在 0 06 左右; RBF 方法中, RBF-TPS 的 RMSE 最大, RBF-MQ 的最小. 从 MRE 来看, RBF-MQ 和 DW1 方法

表 2 Cd 插值交叉验证误差统计结果

Table 2 The cross validation error of the interpolation methods

模型	MAE / (mg kg^{-1})	MRE	RMSE / (mg kg^{-1})	极大值 (Large 15%)		样本数 / 个	
				ME	MRE	RE> 20%	RE> 50%
DW 1	0 042	45 2%	0 060	- 0 111	46 4%	92	40
DW 2	0 043	45 9%	0 062	- 0 113	47 3%	97	40
DW 3	0 045	47 1%	0 064	- 0 115	48 4%	89	41
DW 4	0 047	48 2%	0 067	- 0 118	49 3%	95	44
LP1	0 043	47 2%	0 062	- 0 109	45 1%	84	42
LP2	0 048	49 6%	0 067	- 0 111	47 2%	89	47
LP3	0 062	61 2%	0 116	- 0 058	74 2%	103	60
OK-CR	0 043	46 6%	0 060	- 0 112	47 3%	98	40
OK-EXP	0 043	46 2%	0 060	- 0 112	47 1%	98	39
OK-GS	0 043	47 0%	0 060	- 0 112	47 2%	101	39
OK-SHH	0 043	46 7%	0 060	- 0 112	47 3%	98	40
RBF-CRS	0 044	46 6%	0 062	- 0 112	46 8%	90	41
RBF-MQ	0 042	45 1%	0 059	- 0 109	45 7%	86	38
RBF-MQ	0 046	48 0%	0 066	- 0 113	47 2%	92	47
RBF-ST	0 043	46 4%	0 061	- 0 111	46 6%	91	41
RBF-TPS	0 050	52 3%	0 069	- 0 108	46 3%	94	49

的 MRE 最小, 分别为 45 1% 和 45 2%. LP3 的 MRE 最大为 61 2%. 不同模型预测的最小 MRE 来看, DW 为 DW 1(45 2%), LP 为 LP1(47 2%), OK 模型为 OK-EXP (46 2%), RBF 方法为 RBF-MQ

(45 1%), 不同模型之间的精度差异在 2 1% 以内. 所有的插值方法中, 至少有 61 3% (84 个) 的样本的预测相对误差 (RE) 大于 20%, 超过 27 7% (38 个) 的样本的预测相对误差大于 50%.

3 3 插值方法对极大值的预测精度

为分析插值模型对极大值的预测精度, 选取样本中 Cd含量最高的 15% 样本进行评价. 从极大值的插值误差来看 (表 2), 极大值 (large15%) 插值的平均误差 (ME) 均小于 0 表明插值后极大值被低估, LP3的平均误差最小, IDW 4对极大值的预测平均误差 (ME)最大. 从 MRE 来看, LP3对极大值的预测误差最大 (74.2%), LP1、RBF-MQ 和 IDW 1对极大值的预测误差较小. IDW 方法距离加权系数越大, 对极大值插值的误差越大; LP方法的回归次数越高, 极大值插值的误差越大; OK 方法, 不同半变异拟合函数插值对极大值预测的精度相近; RBF方法中, RBF-MQ 对极大值预测误差最大, RBF-MQ对极大值的预测精度最高. 从不同模型预测的最小平均相对误差 (MRE) 来看, IDW 方法为 IDW 1 (46.4%), LP方法为 LP2(45.1%), OK 方法为 OK-EXP(47.1%), RBF 方法为 RBF-MQ (45.7%), 不同插值方法的差异在 2.0% 之间.

3 4 不同插值方法的 Cd污染评价结果比较

土壤重金属污染区域面积主要有两种估算方式: 一是根据样点超标率进行推算. 该方法利用超过某一标准的样点占总样本量的比例表示污染区面积占总面积的比例; 另一种计算方式是利用插值模型对研究区土壤重金属进行空间插值, 获取土壤重金属的空间分布. 根据土壤重金属污染评价标准对插值结果进行污染评价并统计污染区面积. 本研究利用两种方式分别计算土壤重金属污染分级面积, 比较不同计算方式面积估算结果的差异. 利用背景值-基线值的污染分级原则对研究区土壤 Cd进行污染分级评价. 根据研究区土壤 Cd的含量特征, 将研究区土壤 Cd污染分为 3个级别, 分级标准见表 3.

表 3 土壤 Cd污染评价分级标准

Table 3 Classification standards of Cd pollution assessment

Cd	污染评价分级标准 /(mg kg ⁻¹)		
	Level 1	Level 2	Level 3
污染程度	清洁区	污染警戒	轻度污染
分级标准	< 0.15	0.15~0.25	> 0.25

污染面积统计结果表明 (表 4), 不同插值模型估算的轻度污染区 (Level 3)的面积差异较大. RBF-TPS估算的轻度污染区面积比例最高, 为 2.12%, 而 OK 方法、LP1、LP2和 RBF-MQ 插值估算的结果为无轻度污染区域. IDW 方法的距离加权系数越

大, 估算的轻度污染区域面积比例越高. RBF 方法中, RBF-TPS 估算的轻度污染区面积比例最大, RBFMQ 次之, RBF-CRS和 RBF-ST 较小, RBF-MQ 方法污染面积比例为 0 与根据超标率估算的轻度污染区面积比例 (3.63%) 相比, 插值方法的估计结果要低 1.53% ~ 3.63%.

表 4 不同插值方法估算的土壤 Cd分级所占的面积比例
Table 4 Soil Cd contaminated area estimated by interpolation methods

模型	土壤 Cd分级的面积比例		
	Level 1	Level 2	Level 3
IDW 1	92.02%	7.91%	0.07%
IDW 2	85.87%	13.57%	0.56%
IDW 3	82.56%	16.45%	0.99%
IDW 4	81.08%	17.64%	1.28%
LP1	95.67%	4.33%	0
LP2	91.50%	8.50%	0
LP3	90.77%	7.52%	1.71%
OK-CR	98.49%	1.51%	0
OK-EXP	98.46%	1.54%	0
OK-GS	98.61%	1.39%	0
OK-SPH	98.49%	1.51%	0
RBF-CRS	85.54%	14.05%	0.41%
RBF-ST	87.10%	12.63%	0.27%
RBF-TPS	76.79%	21.09%	2.12%
RBFMQ	80.09%	18.69%	1.22%
RBF-MQ	97.51%	2.49%	0
样点超标率	77.37%	18.98%	3.63%

3 5 Cd污染评价结果空间分布的差异

不同模型获取的 Cd含量空间分布总体趋势是一致的, 研究区西部和中部 Cd含量较高, 东南部较低. 在 ArcGIS 软件中, 利用 Spatial Analysis的栅格运算功能对不同插值模型的污染评价结果进行差值运算. 由于插值方法较多且参数设置变化较多, 比较插值结果污染区的空间分布差异时, 选择污染面积估算较高和较低的插值方法进行比较. 插值方法中 RBF-TPS、RBFMQ、IDW 4的估算结果较高, RBF-ST、RBF-MQ、OK 和 LP方法估算结果较低, 由于估算结果较低的插值方法大部分估算结果为 0 为突出插值结果空间分布的差异, 选择 OK 方法为代表. 为分析权重对 IDW 方法插值结果的影响, 选择 IDW 4和 IDW 2进行空间差异比较. 从图 3 的污染评价结果比较可知, 不同插值模型评价的污染程度在局部存在较大差异. 在局部极大值区域, 中间样点的 Cd含量远高于临近样点的 Cd含量, RBF-TPS、RBFMQ 和 IDW 模型的污染评价结果比 OK

模型高 2 个污染级别. 在样点污染级别为 Level 2 的区域, RBF-TPS、RBF-MQ 和 IDW 模型的污染评价结果比 OK 模型高 1 个污染级别. 在 Cd 污染级别由 level 2 向 Level 1 过渡的部分区域, OK-EXP 方法比 RBF-TPS、RBFMQ 和 IDW 评价结果高 1 个级别. RBF-TPS 与 IDW 2 的差异主要是在高浓度向低浓度的过渡区域; 两种方法确定的污染区边界和形状不一致, RBF-TPS 方法确定的高污染级别范围更大.

IDW 4 和 IDW 2 的比较结果可以看出, IDW 方法距离加权系数的增大, 局部高值点的影响增加, 污染区的面积增加, 污染范围呈同心圆状增大. 总的来看, 不同方法污染评价结果的差异主要位于: ①局部极大值区域, 少数的孤立高污染样本周边的样点含量相对较低, 污染评价结果存在较大不确定性; ②在高浓度到低浓度的过渡区域, 不同插值方法确定的污染区边界范围存在很大的不确定性.

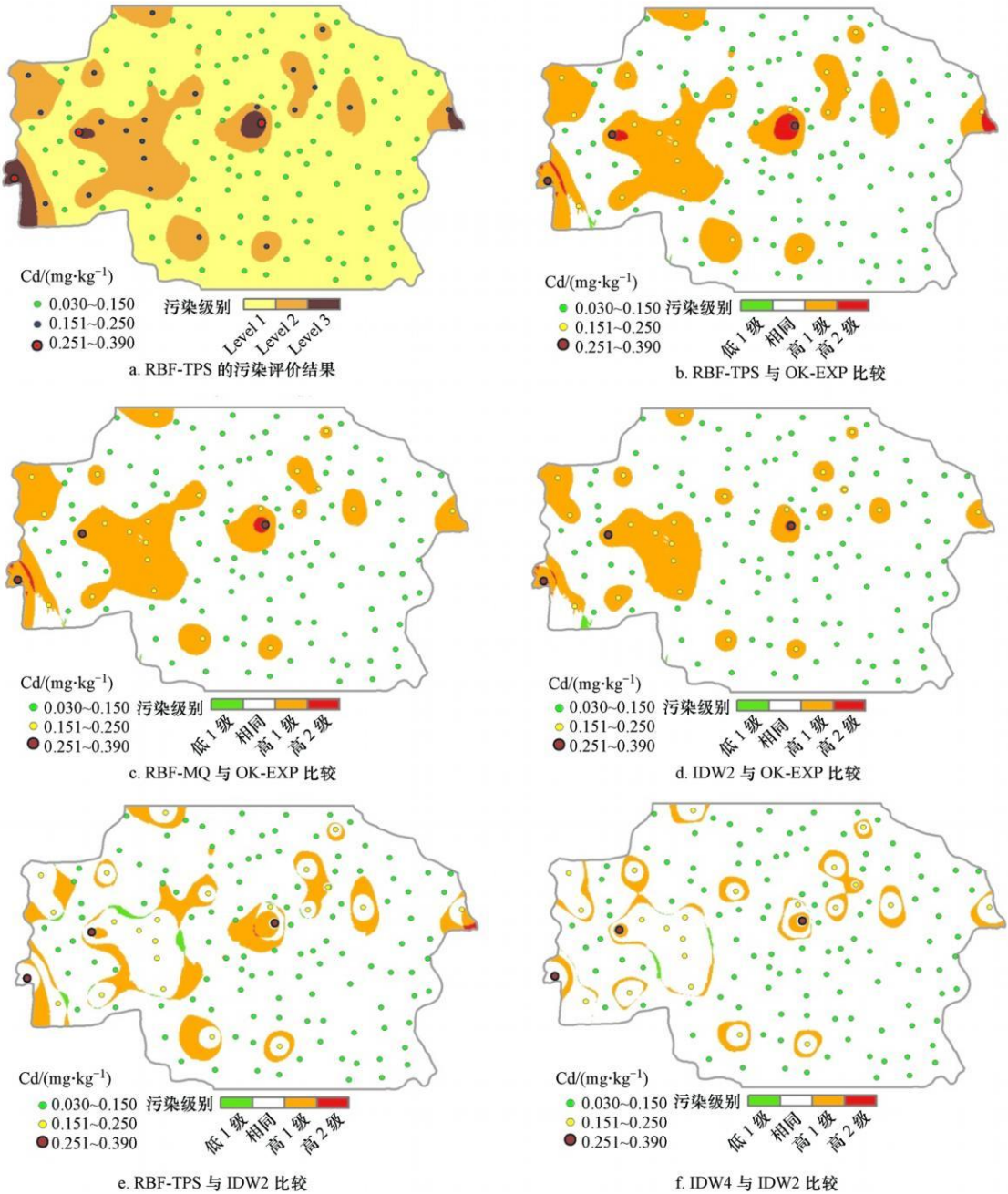


图 3 不同插值方法的污染评价分级结果比较

Fig. 3 Comparison of Cd evaluation results by different interpolation methods

4 讨论 (Discussion)

从不同插值方法的污染区面积计算结果(表 4)和污染区的空间分布对比(图 3)研究可知,插值模型选择对污染评价结果的影响较大,就同一种插值方法而言,不同的模型参数设置,污染评价结果也存在着较大的差异。而根据总体插值和极大值插值的验证结果(表 2)都表明不同插值方法的差异相对较小。两种比较结果存在差异的主要原因是交叉验证方法和直接根据插值结果计算污染面积差异的评价思路和目标存在着差异。交叉验证是根据所有采样点插值的平均误差表示插值方法的误差,极大值插值误差是 15% 极大值样本插值误差的平均值;污染区面积估算结果的差异是不同插值方法插值结果在空间上的差异。利用插值方法由土壤采样点到污染空间分布面的转换时,插值误差必然会放大。根据空间分布面积统计差异是一个累加的过程,而交叉验证法仅对采样点进行插值误差估计,且是所有样点插值误差的平均值。另外,交叉验证方法本身也存在一定的局限性。利用交叉验证法进行插值验证时,是先去掉一个点,然后通过周围的采样点来预测,根据预测结果与实测结果的差异确定预测误差,但在实际的插值过程中,所有的采样点都要参与插值运算的。因此,交叉验证的评价结果与不同插值图比较的结果存在一定差异。从总体插值的均方根误差(表 2)和污染区域面积估算结果(表 4)比较来看,对总体插值误差较小的 RBF-MQ、IDW 1 和 OK 方法污染区面积估算结果较小(或无污染区),误差较大;RBF-TPS 和 RBF-MQ 和 IDW 4 的总体插值误差较大,但对污染区面积估算结果较大,误差较小。总体插值精度是对样本总体预测精度的评价,不能反映对污染区面积估算的精度。

由于重金属元素在土壤中的迁移相对困难,容易在局部地域累积(王学军等, 1997),高污染风险区域的样点往往表现为局部极大值,而且高污染风险区的样本占总体样本量的比例较小。插值方法是以对总体期望值(平均含量)的预测为目标,必然会对局部细节的污染信息产生平滑作用,平滑的结果就是极大值降低、极小值升高。因此,根据插值模型估算的污染区面积要低于根据样点超标率估算的结果。

空间插值精度取决于模型对要素空间变异性

和相关性的反映(朱会义等, 2004)。就污染评价而言,对土壤重金属元素的空间变异性的反应更为重要。OK 方法以 Cd 的空间结构特征为基础,确定采样点对估值点的影响权重,给出对样本的总体最优无偏(最小方差)估计,对污染物的空间分布趋势有良好的预测效果,在理论上是对总体的最佳估计方法。但是,半变异函数拟合的精度对 OK 方法的精度影响较高,在样本量较小的情况下,很难获得理想的半变异拟合函数。同时,半变异函数的拟合存在较大的主观性。OK 方法在空间上存在低通滤波效应,插值过程中会丢失局部极大值和极小值信息,导致对局部极大值的低估和局部极小值的高估;在空间变异大、自相关性较差的区域,平滑效应越强(李庆谋, 2005; Goovaerts, 1999)。体现在污染评价结果上,高污染风险区域面积降低。LP 方法属于非精确性插值方法,非精确性插值法预测值在样点处一般不等于实测值。LP 主要是利用最小二乘多项式拟合土壤重金属的局部空间分布趋势,插值结果较光滑,但是对局部细节信息存在较大的平滑作用。

IDW 和 RBF 都属于精确性插值算法,即样点处的预测值与实测值相等(Smith et al., 2007)。所以 IDW 和 RBF 的插值结果都保留了土壤重金属空间分布的局部波峰(极大值)或波谷(极小值)信息。两种方法的主要差异在于 IDW 的权重是根据距离的影响确定,而 RBF 方法是根据局部的平滑趋势确定的权重。IDW 方法的距离加权系数越大,局部极大值点影响的范围就越大,污染区的估计范围就越大,所以 IDW 4 的污染区面积高于 IDW 2 的结果。RBF 方法不同核函数拟合的局部平滑趋势不一致,RBF-TPS 方法对局部极大值的预测效果较好。总的来看,在土壤重金属污染评价方面,RBF 和 IDW 方法较适宜。从不同方法污染评价结果的空间差异来看(图 3),不确定性比较大的区域主要是局部极大值区域和高含量到低含量的过渡区域。如果在这些不确定性的区域增加抽样的样本量,可以进一步提高污染评价结果的精度。

5 结论 (Conclusions)

1) 插值模型的总体插值误差较大,不同方法预测的均方根误差(RMSE)均大于 $0.059\text{mg} \cdot \text{kg}^{-1}$ 。插值的相对误差(MRE)均大于 45%,不同插值方法的 MRE 差异在 2% 以内。

2) 插值模型预测的极大值均小于原始值,超过

27.7%的Cd样本插值相对误差大于50%。插值方法对极大值预测的MRE高于45%。RBF-TPS方法对极大值预测的平均误差(ME)最小,OK方法的平均误差较大。

3)基于样点数据估算的轻度污染区面积比例比插值模型估算的面积高1.53%~3.65%。不同插值模型估算的轻度污染面积比例在0~2.12%之间。RBF、IDW模型估算的轻度污染区域面积相对较大。LP和OK方法估计的轻度污染面积区为0。不同插值模型的污染评价结果差异主要是在局部极大值区域和高值向低值的过渡区域。

4)从土壤重金属污染区面积估算和空间分布预测精度来看,RBF方法和IDW方法要优于OK方法和LP方法。

责任作者简介:陈同斌(1963—),男,研究员,博士生导师,中国科学院地理科学与资源研究所环境修复中心主任。主要从事植物修复、固体废弃物堆肥处理、区域土壤环境质量与风险评估研究。主持“863”课题、“973”课题、国家杰出青年科学基金项目等40多项,发表论文200余篇,申请或获得发明专利18项。

参考文献 (References):

Bishop T F A, McBratney A B. 2001. A comparison of prediction methods for the creation of field extent soil property maps [J]. *Geoderma* 103(1/2): 149—160

陈同斌, 郑表明, 陈煌, 等. 2004. 北京市土壤重金属含量背景值的系统研究 [J]. *环境科学*, 25(1): 117—122

Chen T B, Zheng Y M, Chen H, *et al*. 2004. Background concentrations of soil heavy metals in Beijing [J]. *Environmental Science* 25(1): 117—122 (in Chinese)

Cheng J L, Shi Z, Zhu Y W. 2007. Assessment and mapping of environmental quality in agricultural soils of Zhejiang Province, China [J]. *Journal of Environmental Sciences* 19(1): 50—54

Goovaerts P. 1999. Geostatistics in soil science: state-of-the-art and perspectives [J]. *Geoderma* 89(1/2): 1—45

Goway C A, Ferguson R B, Hergert G W, *et al*. 1996. Comparison of Kriging and inverse distance methods for mapping soil parameters [J]. *Soil Science Society of America Journal* 60(4): 1237—1247

Imperato M, Adamo P, Naimo D, *et al*. 2003. Spatial distribution of heavy metals in urban soils of Naples city (Italy) [J]. *Environmental Pollution* 124(2): 247—256

Isaaks E H, Srivastava R M. 1990. An introduction to applied

geostatistics [M]. Oxford, USA: Oxford University Press 278—301

李庆谋. 2005. 多维分形克里格方法 [J]. *地球科学进展*, 20(2): 248—256

Li Q M. 2005. Multifactorial Kriging interpolation method [J]. *Advance in Earth Sciences* 20(2): 248—256 (in Chinese)

McBratney A B, Mendonça Santos M L, Minasny B. 2003. On digital soil mapping [J]. *Geoderma* 117(1/2): 3—52

McGrath D, Zhang C, Carton O T. 2004. Geostatistical analyses and hazard assessment on soil lead in Silvermines area, Ireland [J]. *Environmental Pollution* 127(2): 239—248

Panagopoulos T, Jesus J, Antunes M D C, *et al*. 2006. Analysis of spatial interpolation for optimizing management of a salinized field cultivated with lettuce [J]. *European Journal of Agronomy* 24(1): 1—10

Robinson T P, Metternicht G. 2006. Testing the performance of spatial interpolation techniques for mapping soil properties [J]. *Computers and Electronics in Agriculture* 50(2): 97—108

Smith M J, Goodchild M J, Longley P A. 2007. *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and software tools* (2nd Edition) [M]. Troubador Publishing Ltd 235—244

王学军, 邓宝山, 张泽浦. 1997. 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 [J]. *环境科学学报*, 17(4): 414—416

Wang X J, Deng B S, Zhang Z P. 1997. Spatial structures of trace element contents in sewage irrigated soil at the eastern suburb of Beijing [J]. *Acta Scientiae Circumstantiae* 17(4): 414—416 (in Chinese)

Webster R, Oliver M. 2007. *Geostatistics for Environmental Scientists* (2nd Edition) [M]. New York: John Wiley & Sons 153—174

Wollenhaupt N C, Wolkowski R P, Clayton M K. 1994. Mapping soil test phosphorus and potassium for variable-rate fertilizer application [J]. *Journal of Production Agriculture* 7(4): 441—447

Yasrebi J, Saffari M, Fathi H, *et al*. 2009. Evaluation and comparison of Ordinary Kriging and Inverse Distance Weighting methods for prediction of spatial variability of some soil chemical parameters [J]. *Research Journal of Biological Sciences* 4(1): 93—102

郑表明, 陈同斌, 陈煌, 等. 2003. 北京市近郊区土壤镍的空间结构及分布特征 [J]. *地理学报*, 58(3): 470—476

Zheng Y M, Chen T B, Chen H, *et al*. 2003. The spatial structure and distribution of Ni contents in soils of suburbs of Beijing [J]. *Acta Geographica Sinica* 58(3): 470—476 (in Chinese)

朱会义, 刘述林, 贾绍凤. 2004. 自然地理要素空间插值的几个问题 [J]. *地理研究*, 23(4): 425—432

Zhu H Y, Liu S L, Jia S F. 2004. Problems of the spatial interpolation of physical geographical elements [J]. *Geographical Research* 23(4): 425—432 (in Chinese)