

王 丹, 宋 轩, 陈 杰. 基于环境因子的土壤养分预测制图——以河南省封丘县为例[J]. 江苏农业科学, 2014, 42(4): 277–280.

# 基于环境因子的土壤养分预测制图 ——以河南省封丘县为例

王 丹, 宋 轩, 陈 杰

(郑州大学水利与环境学院, 河南郑州 450001)

**摘要:**土壤养分空间变异由气候、母质、地形、植被等结构性因素和施肥、耕作措施等随机因素共同作用。以河南省封丘县为研究对象, 选择高程、坡度、平面曲率、剖面曲率、地形湿度指数、归一化植被指数 (NDVI) 等环境因子为辅助变量模拟土壤结构性分量, 并分别采用普通克里格插值法和反距离权重法处理空间相关的随机分量, 提出了结合环境因子的空间插值方法, 并对土壤有机质和全氮进行预测制图。交叉验证结果表明, 结合环境因子的空间插值方法能够提高土壤养分空间分布的预测精度, 是一种有效的土壤养分预测制图方法。

**关键词:**土壤养分; 环境因子; 克里格插值法; 反距离权重法; 河南封丘

**中图分类号:** S158.9    **文献标志码:** A    **文章编号:** 1002-1302(2014)04-0277-04

土壤并非匀质体, 其空间变异由气候、母质、地形、植被、土壤类型等结构性因素和施肥、耕作措施、种植制度等随机因素共同作用<sup>[1]</sup>。土壤养分的空间变异对环境预测、精准农业和自然资源管理具有重要意义, 已成为土壤科学研究的热点之一。由于受多方面因素限制, 采样点数量是有限的, 为了解整个研究区土壤养分的空间分布状况, 就需要采用空间插值技术。20 世纪 80 年代初, Mathern 的区域化变量理论和地统计方法被引入到土壤特性的研究中<sup>[2]</sup>, 土壤养分空间分布预测逐渐由定性转为定量。此后地统计方法广泛应用于土壤制图领域, 被证明是分析土壤要素空间分布特征最为有效的方法之一<sup>[3]</sup>。20 世纪 90 年代以来, 随着地理信息系统 (GIS) 和遥感技术 (RS) 的发展, 使用环境变量辅助预测土壤养分的方法开始受到土壤调查人员的青睐。研究表明, 量化的环境因子与土壤属性之间存在显著相关性, 可以辅助预测土壤属性<sup>[4]</sup>。这种基于环境因子进行土壤预测性制图的方法, 如空间回归模型, 弥补了较少考虑影响土壤养分空间分布结构性因素的地统计学方法的不足。然而, 在环境变化较为均一的地区, 环境因子与土壤属性的相关性较低, 仅利用环境因子进行预测制图难以得到理想的效果。刘静等结合海拔高度改进空间插值方法, 该方法对土壤有机质预测效果明显优于反距离权重法和普通克里格方法<sup>[5]</sup>。Sumfleth 等认为, 将环境因子与插值方法相结合是土壤制图的有效方法<sup>[6]</sup>。因此, 将环境因子与插值方法相结合是土壤养分精确制图的关键。本研究以河南省封丘县为对象, 以高程、坡度、平面曲率、剖面曲率、地形湿度指数、归一化植被指数等环境因子为辅助变量模拟土壤结构化分量, 分别采用普通克里格插值法和反距离权

重法处理空间相关的随机分量, 对研究区有机质和全氮含量进行了数字化制图研究, 以期为准农业和生态环境建设提供精确的数字土壤地图。

## 1 材料与方法

### 1.1 研究区概况

封丘县位于河南省东北部, 界于 144°14′ ~ 114°46′ E、34°53′ ~ 35°14′ N, 总面积 1 225 km<sup>2</sup>; 地处黄河故道, 地貌复杂, 沙岗、平原、洼地兼有; 黄河大堤以南地势较高, 其余地势低洼, 整体由西南向东北倾斜, 属暖温带大陆性季风气候, 年均气温 13.9 ℃, 年均降水量 615.1 mm, 无霜期 214 d。主要土壤类型是新成土和雏形土, 占全县土壤总面积的 98% 以上。

### 1.2 研究方法

**1.2.1 土壤样品采集与分析** 利用 SPOT 遥感图像, 兼顾代表性和均匀性, 共布设 355 个土壤样点, 样点分布如图 1 所示。取 0~20 cm 的表层土壤, 采用四分法取样, 并通过 GPS 获取采样点坐标。所有土壤样品经自然风干后, 采用重铬酸钾容量法测定有机质含量, 采用半微量凯氏定氮法测定全氮含量。

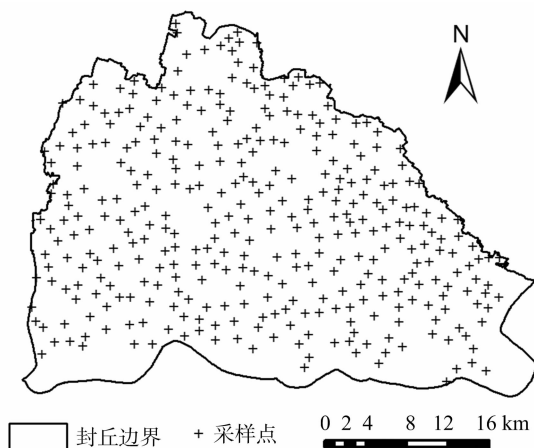


图1 采样点分布

收稿日期: 2013-08-29

基金项目: 国家自然科学基金 (编号: 40971128); 郑州大学研究生创新基金。

作者简介: 王 丹 (1989—), 女, 河南许昌人, 硕士研究生, 从事土地资源遥感与信息技术研究。E-mail: wdzu2012@126.com。

通信作者: 宋 轩。E-mail: songxuan@zzu.edu.cn。

1.2.2 环境因子的提取与处理 研究选择的环境因子包括地形因子和植被因子,其中,地形因子包括高程( $H$ )、坡度( $\beta$ )、平面曲率( $C_h$ )、剖面曲率( $C_v$ )和地形湿度指数( $\Phi$ ,也称复合地形指数)<sup>[6]</sup>,是在 ArcGIS9.3 支持下,利用研究 30 m 分辨率的数字高程模型计算得到;植被因子选取反映农作物长势和营养状况的归一化植被指数(NDVI)<sup>[3]</sup>,该指数采用同期 Landsat ETM+ 遥感数据计算得到。不同环境因子的量纲是不同的,需要进行标准化处理<sup>[7]</sup>。采用主成分分析可以消除环境因子间的多重共线性,用主成分替代原始环境因子可以提高土壤养分建模的预测精度<sup>[8]</sup>。

1.2.3 土壤养分数据处理 土壤养分数据有明显的阈值限制,并且其与环境因子之间的函数关系是不确定且复杂的,简单的线性函数并不能准确描述这种关系。研究采用 Logit 变换对土壤养分进行处理<sup>[9]</sup>,如公式(1):

$$Z^{++} = \ln\left(\frac{Z^+}{1 - Z^+}\right), 0 < Z^+ < 1 \tag{1}$$

式中: $Z^{++}$ 是目标变量  $Z$  经过 Logit 变换后的值; $Z^+$ 是目标变量  $Z$  标准化至 0 到 1 范围内的值,可由公式(2)求得:

$$Z^+ = \frac{Z - Z_{\min}}{Z_{\max} - Z_{\min}}, Z_{\min} < Z < Z_{\max} \tag{2}$$

式中: $Z_{\max}$ 与 $Z_{\min}$ 分别是目标变量  $Z$  的物理最大值和最小值。经过变换后,所有预测值都被限定在极值之间。为了避免  $\ln(0)$  情况的出现, $Z$  值不能与  $Z_{\max}$ 或 $Z_{\min}$ 相等。

土壤养分预测值反变换至原始取值范围如公式(3):

$$\hat{Z}(x_0) = \frac{e^{\hat{Z}^{++}(x_0)}}{1 + e^{\hat{Z}^{++}(x_0)}}(Z_{\max} - Z_{\min}) + Z_{\min} \tag{3}$$

式中: $\hat{Z}^{++}(x_0)$ 为采样点  $x_0$  处的土壤养分回归预测值; $\hat{Z}(x_0)$ 为经过反变换后的土壤养分预测值。

1.2.4 基于环境因子的土壤养分插值方法 采用多元线性回归分析研究土壤养分和环境因子之间的变化关系,回归系数表示土壤养分随环境因子的变化率。在此基础上,分 3 个步骤对土壤养分进行插值制图。

第一步,消除土壤养分中受环境因子影响的部分,采用公式(4)将其修正到相同环境条件下:

$$Z'(X_0) = Z(X_0) - \sum_{j=1}^m a_j \cdot [EI_j(x_0) - EI_{j\min}] \tag{4}$$

式中: $Z'(X_0)$ 为采样点  $x_0$  处土壤养分修正值; $Z(X_0)$ 为采样点  $x_0$  处土壤养分实测值; $m$ 为环境因子个数; $a_j$ 为土壤养分随第  $j$  个环境因子的变化率; $EI_j(x_0)$ 为采样点  $x_0$  处的第  $j$  个环境因子值; $EI_{j\min}$ 为采样点第  $j$  个环境因子最小值。修正后的土壤养分可以被认为是空间相关的随机变量,不包含任何与环境因子相关的结构化信息<sup>[10]</sup>。

第二步,分别选用反距离权重法和普通克里格插值法等常用插值方法,对修正后的土壤养分进行插值。

第三步,插值得到的栅格图像与土壤养分中被消除的部分叠加,生成土壤养分插值图,如公式(5):

$$Z^+(X_0) = Z^*(X_0) + \sum_{j=1}^m a_j \cdot [EI_j(x_0) - EI_{j\min}] \tag{5}$$

式中: $Z^*(X_0)$ 为第二步的插值结果; $Z^+(X_0)$ 为土壤养分实际预测值。

1.3 精度验证方法

采用交叉验证法评估不同插值方法,选取平均绝对误差

(MAE)和均方根误差(RMSE)作为预测精度的评估指标。MAE 和 RMSE 分别根据公式(6)和公式(7)来求算。

$$MAE = \frac{1}{n} \cdot \sum_{i=1}^n |Z^+(X_i) - Z(X_i)| \tag{6}$$

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n [Z^+(X_i) - Z(X_i)]^2} \tag{7}$$

式中: $Z^+(X_i)$ 与 $Z(X_i)$ 分别是土壤养分预测值与实际测定值。MAE、RMSE 越小说明预测精度越高。

2 结果与分析

2.1 土壤养分统计特征

根据阈值识别法剔除采样点异常值,其中,有机质异常值 6 个,全氮异常值 7 个。对剔除异常值后的土壤养分进行统计分析,结果见表 1。从变异系数(CV)来看,CV<10%为弱变异性,10%≤CV≤100%为中等变异性,CV>100%为高度变异性<sup>[11]</sup>,故有机质和全氮均为弱变异性。采用柯尔莫哥洛夫-斯米尔诺夫(Kolmogorow-Smirnov, K-S)方法对有机质和全氮进行正态分布检验<sup>[12]</sup>,结果表明有机质和全氮含量均服从正态分布。

表 1 土壤养分统计特征

土壤养分	样本数 (个)	最小值 (g/kg)	最大值 (g/kg)	平均值 (g/kg)	标准差 (g/kg)	变异系数 (%)
有机质	349	6.10	21.60	14.06	2.40	5.78
全氮	348	0.50	1.28	0.87	0.13	0.02

2.2 土壤养分与环境因子的关系

利用 ArcGIS 所提取的环境因子与所测土壤有机质、全氮进行相关性分析。由表 2 可见,土壤有机质和全氮含量与高程相关系数分别为-0.213 和-0.294,均呈极显著负相关关系。一般而言,随着海拔的升高,风力加大,太阳辐射增强,水热条件加快了有机质的分解速度,并且在重力和侵蚀作用下,海拔高处土壤中的物质和能量被运移到海拔较低处,海拔高处的有机质和全氮含量比海拔低处低。有机质和全氮含量与植被指数相关系数分别为 0.315 和 0.271,均呈极显著正相关关系,表明土壤有机质和全氮对植物生长影响较大。通常情况下,良好的植被覆盖可以有效减少侵蚀作用,利于土壤保水、保肥,植被长势良好,覆盖度大,凋落物及田间残留物相对较多,腐殖化作用显著,表层土壤有机质含量较高,而有机质对土壤水、气、热等各种肥力因素起重要调节作用,影响养分的保持和供给,使氮素等其他养分含量增高<sup>[3]</sup>。除此之外,全氮与坡度呈显著负相关关系( $r = -0.118$ ),表明坡度越陡,全氮含量越低,这是由于坡度越陡,土壤养分越容易流失造成的。全氮与地形湿度指数呈正相关关系( $r = 0.110$ ),地形湿度指数在一定程度上与土壤含水量和沉积物运移相关,土壤水分含量较高的地区全氮含量较大。有机质与地形湿度指数呈正相关关系,但不显著。地形湿度指数与土壤养分的相关性较低,远没有达到预期效果,这与连纲等的研究结论<sup>[4,13]</sup>相似。

表 2 土壤养分和环境因子间的相关性

土壤养分	$H$	$\beta$	$C_h$	$C_v$	$\Phi$	NDVI
有机质	-0.213**	-0.055	0.065	0.015	0.062	0.315**
全氮	-0.294**	-0.118*	-0.037	0.015	0.110*	0.271**

注:“\*”表示  $P < 0.05$ ;“\*\*”表示  $P < 0.01$ 。下同。

剔除相关性较低的环境因子,对剩余的环境因子(有机质:高程、归一化植被指数;全氮:高程、归一化植被指数、坡度、地形湿度因子)进行主成分分析,以消除环境因子间的多重共线性。选择可解释原变量超过 85% 信息的成分作为主成分,在此基础上进行回归分析。由表 3 可见,2 个主成分 SPC1 和 SPC2 可以代表有机质与环境变量之间的关系,全氮与环境之间的关系则需要 3 个主成分 SPC1、SPC2 和 SPC3;有机质随 SPC1 和 SPC2 的变化率分别为 0.79 和 -0.205,全氮随 SPC1、SPC2 和 SPC3 的变化率分别为 -0.030、0.032 和 -0.004; $R^2$  值表明,利用主成分分别能解释有机质含量及全氮含量在空间上的变异为 11.5% 和 12.3%, $R^2$  值均不高; $F$  检验结果表明,回归模型能较好地表达有机质和全氮含量的变异。本研究尺度较大,土地利用类型变化较为复杂,这可能影响土壤养分的拟合效果。另外,如果使用高分辨率、大比例尺的数字高程模型及高精度 GPS、更为详尽的环境变量(如土

表 3 土壤养分与环境因子主成分回归结果						
土壤养分	SPC1	SPC2	SPC3	常数项	$R_2$	$F$
有机质	0.790 **	-0.205 *		14.059 **	0.115	21.880 **
全氮	-0.030 **	0.032 **	-0.004	0.867 **	0.123	15.384 **

壤类型、气温等),可能会提高模型解释变异的能力<sup>[14]</sup>。

2.3 土壤有机质和全氮的空间分布

本研究首先采用传统反距离权重法(IDW)和普通克里格方法(OK),然后采用结合环境因子的反距离权重法和普通克里格插值法对土壤有机质和全氮进行空间预测制图,结果如图 2、图 3 所示。总体上看,4 种插值方法得到的土壤养分空间格局分布结果相似:有机质空间分布呈现北部偏高、南部偏低的发展趋势,全氮空间分布呈现东北部偏高、东南部偏低的趋势;有机质、全氮空间分布的整体趋势与研究区南高北低的地势相反,较高值出现于地势相对低洼的地区。

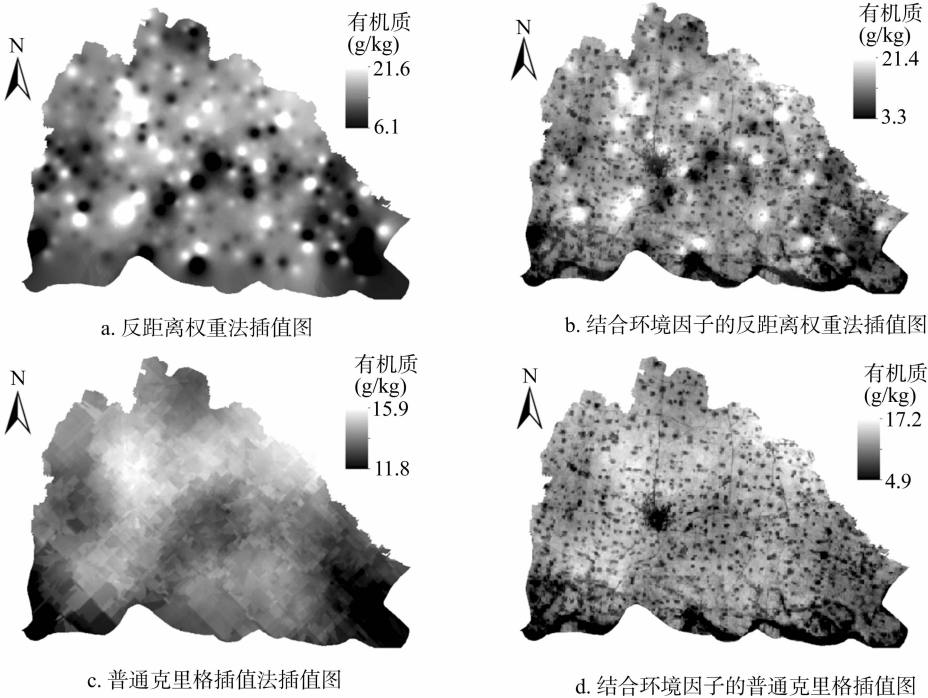


图2 封丘县土壤有机质空间分布图

通过对比原方法(IDW、OK)与改进方法(改进 IDW、改进 OK)得到的养分空间分布图发现,改进后的方法表现出更加明显的梯度变化,能更准确地反映出土壤养分随环境的变化趋势和由于环境变化所带来的局部差异。

通过结合环境因子的插值方法得到的 4 副预测图(图 2、图 3)可以看出,研究区离散分布着许多小块状的低值区,以及中部明显的块状低值区和南部出现的条带状低值区。这是由于归一化植被指数反映植被覆盖状况,能有效地识别植被覆盖地区和非植被覆盖地区,这些低值区与非植被覆盖地区相吻合,离散的小块状以及中部明显的块状低值区均为建设用地,南部条带状低值区则为黄河滩地。因此,应用结合环境因子的插值方法生成的土壤养分图,在一定程度上也反映了土地利用情况,2 种改进方法相比,克里格插值法有效地消除了反距离权重法的“牛眼”现象,在目视效果上优于改进的反

距离权重法。

2.4 不同插值方法精度比较

克里格插值法主要根据周围样点与采样点的相关程度来确定权重系数,而反距离权重法仅以距离反比作为周围样点的权重系数。由表 4 可见,普通克里格法插值精度略优于反距离权重法,但优势并不明显,对于有机质,MAE 和 RMSE 分别提高了 0.49% 和 0.34%,对于全氮,MAE 和 RMSE 分别提高了 0.11% 和 0.41%;结合环境因子的 2 种插值方法精度均比原方法有所提高,对于有机质,结合环境因子的反距离权重法,MAE 和 RMSE 精度分别提高了 2.04% 和 5.55%,结合环境因子的普通克里格插值法,MAE 和 RMSE 精度分别提高了 2.08% 和 5.66%;对于全氮,结合环境因子的反距离权重法,MAE 和 RMSE 精度分别提高了 2.68% 和 4.75%,结合环境因子的普通克里格插值法,MAE 和 RMSE 精度分别提高了 2.56% 和

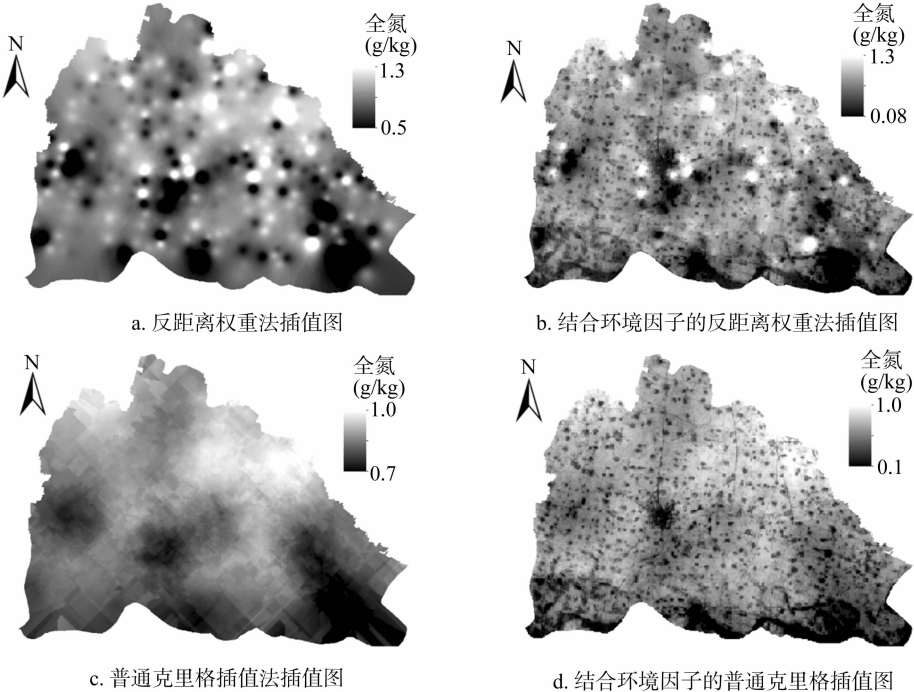


图3 封丘县土壤全氮空间分布图

表 4 不同插值方法精度比较

土壤养分	插值方法	IDW		OK	
		MAE	RMSE	MAE	RMSE
有机质	原方法	1.671 7	2.406 4	1.663 5	2.398 1
	改进方法	1.637 5	2.272 8	1.628 9	2.262 4
全氮	原方法	0.088 8	0.124 8	0.088 7	0.124 3
	改进方法	0.086 3	0.118 8	0.086 3	0.117 9

注:MAE 为平均绝对误差;RMSE 为均方根误差;IDW 为反距离权重插值法;OK 普通克里格插值法。

5.07%。改进后的插值方法精度较原方法虽然仍未得到大幅提高,但是,精度提高程度已经远大于普通克里格方法相对于反距离权重法的提高程度。综合比较 4 种方法,结合环境因子的普通克里格插值法精度最高。

3 小结与讨论

封丘县土壤有机质呈现北部偏高、南部偏低的空间分布趋势,全氮呈现东北部偏高、东南部偏低的空间分布趋势。研究区内土壤养分变异较低,属弱变异性。环境因子对土壤养分空间分布具有一定影响,但这种影响在不同环境因子间表现出差异性。高程和归一化植被指数是影响土壤养分分布的主要因素,其对土壤有机质、全氮具有极显著影响。坡度和地形湿度指数显著影响全氮的空间分布,平面曲率和剖面曲率对土壤养分没有显著影响。与传统普通克里格插值法和反距离权重法相比,结合环境因子的插值方法考虑了影响土壤养分空间分布的环境因素,可以有效地提高土壤养分空间分布预测精度。从制图目视效果和预测精度来看,结合环境因子的普通克里格插值法整体最优。

影响土壤养分空间分布的因素复杂,除了考虑环境因素外,还可以考虑如人为因素、土地利用状况等其他影响因素,这样可以更加精确地了解研究区的土壤养分空间分布状况。

参考文献:

[1] 贾树海,张琦,孟维军,等. 基于 GIS 与地统计学土壤养分空间变异特性研究[J]. 水土保持通报,2009,29(3):197-202.

[2] 王珂,许红卫,史舟,等. 土壤钾素空间变异性和空间插值方法的比较研究[J]. 植物营养与肥料学报,2000,6(3):318-322,344.

[3] 张素梅,王宗明,张柏,等. 利用地形和遥感数据预测土壤养分空间分布[J]. 农业工程学报,2010,26(5):188-194.

[4] 连纲,郭旭东,傅伯杰,等. 基于环境相关法和地统计学的土壤属性空间分布预测[J]. 农业工程学报,2009,25(7):237-242.

[5] 刘静,蔡国学,刘洪斌. 西南丘陵区土壤有机质含量的空间插值法研究[J]. 西南大学学报:自然科学版,2008,30(3):107-112.

[6] Sumfleth K, Duttman R. Prediction of soil property distribution in paddy soil landscapes using terrain data and satellite information as indicators[J]. Ecological Indicators,2008,8(5):485-501.

[7] AnselinL, Syabri I, Kho Y. GeoDa: an introduction to spatial data analysis[J]. Geographical Analysis,2006,38(1):5-22.

[8] Hengl T, Heuvelink G B, Stein A. A generic framework for spatial prediction of soil variables based on regression-kriging[J]. Geoderma,2004,120(1/2):75-93.

[9] Gessler P E, Moore I D, McKenzie N J, et al. Soil-landscape modeling and spatial prediction of soil attribute[J]. International Journal of Geographical Information Systems,1995,9(4):421-432.

[10] 李新,程国栋,卢玲. 青藏高原气温分布的空间插值方法比较[J]. 高原气象,2003,22(6):565-573.

[11] 连纲,郭旭东,傅伯杰,等. 黄土高原小流域土壤容重及水分空间变异特征[J]. 生态学报,2006,26(3):647-654.

[12] 徐丽华,谢德体,李兵. 距离倒数和克里克方法在土壤养分制图中的比较研究[J]. 农机化研究,2011,33(11):56-61,70.

[13] 宋轩,李立东,寇长林,等. 黄水河小流域土壤养分分布及其与地形的关系[J]. 应用生态学报,2011,22(12):3163-3168.

[14] 孙孝林,赵玉国,赵量,等. 应用土壤-景观定量模型预测土壤属性空间分布及制图[J]. 土壤,2008,40(5):837-842.