

## 基于多元地统计的土壤有机质含量空间格局反演

陈锋锐<sup>1,2</sup>, 秦奋<sup>1,2</sup>, 李熙<sup>3</sup>, 彭光雄<sup>4</sup>

(1. 河南大学黄河中下游数字地理技术教育部重点实验室, 开封 475004; 2. 河南大学环境与规划学院, 开封 475004; 3. 武汉大学测绘遥感国家重点实验室, 武汉 430079; 4. 中南大学地球科学与信息物理学院, 长沙 410083)

**摘要:** 为了提高土壤有机质含量的空间预测精度, 该文采用了一种多元地统计方法来构建遥感定量反演模型。考虑到回归误差在空间上具有一定程度的聚类, 该文提出了基于局部变化均值的普通克里金方法, 然后用其构建土壤有机质含量遥感定量反演模型。对四川省西南部土壤有机质含量进行空间预测试验, 并与普通克里金、普通遥感定量反演、基于回归克里金的遥感定量反演等方法相比较。结果表明: 该文提出方法的空间预测结果最优, 其原因为该方法通过空间统计来建立采样数据与地表反射率间的联系, 充分考虑了数据间的空间相关性, 因此可以更精确地获得土壤有机质含量的遥感反演模型; 相比基于回归克里金的遥感定量反演方法, 基于局部变化均值的普通克里金假设回归误差在局部邻域内的均值也不一定为零, 更符合实际情况。该方法为农田养分管理及区域农业的可持续发展提供科学依据。

**关键词:** 土壤, 遥感, 回归分析, 有机质, 反演, 地统计

doi: 10.3969/j.issn.1002-6819.2012.20.025

中图分类号: S29; S127

文献标志码: A

文章编号: 1002-6819(2012)-20-0188-07

陈锋锐, 秦奋, 李熙, 等. 基于多元地统计的土壤有机质含量空间格局反演[J]. 农业工程学报, 2012, 28(20): 188—194.

Chen Fengrui, Qin Fen, Li Xi, et al. Inversion for spatial distribution of soil organic matter content based on multivariate geostatistics[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2012, 28(20): 188—194. (in Chinese with English abstract)

### 0 引言

土壤有机质 (soil organic matter, SOM) 是土壤固相物质的组成部分, 是植物营养的仓库, 被称为植物的“养分银行”。土壤有机质含量的多少是衡量土壤肥力的一个重要指标, 它在空间上非均匀分布, 呈现一定的异质性。通常土壤有机质含量的获得来自于土壤采样, 很难满足对面含量的实际需求, 因此, 精确地获得土壤有机质含量的空间分布, 有利于准确掌握土壤肥力状况, 为农田养分管理以及区域农业的可持续发展提供科学依据。

传统上往往采用空间插值方法, 如反距离加权、普通克里金 (ordinary Kriging, OK)<sup>[1]</sup>等, 利用采样点数据对未采样位置的含量进行估计, 获得土壤有机质含量在整个区域的空间分布。然而, 由于采样点的数量有限, 且分布不均匀, 因此, 难以获得精确的空间预测结果。自 20 世纪 70 年代开始, 许多学者开始

利用遥感影像反演土壤有机质含量的空间格局, 大量的研究表明: 土壤有机质在可见光-短波红外波段范围内具有明显的光谱特征, 随土壤有机质含量的增加, 土壤的光谱反射率减小<sup>[2-4]</sup>。Gomez 等<sup>[5]</sup>利用 Hyperion 高光谱影像中的可见光和近红外波段对新南威尔士西南部区域的土壤有机碳含量进行了定量反演。Hill 等<sup>[6]</sup>、张法升等<sup>[7]</sup>利用 TM 影像来反演土壤有机质含量的空间分布。此外, 其他学者<sup>[8-13]</sup>也从遥感影像中较好地反演出土壤有机质含量的空间格局。利用遥感影像反演土壤有机质含量具有简单, 准确等特点, 关键是如何建立并优化遥感反演模型。目前, 模型构建大都是基于经典统计方法, 试图通过回归分析找到土壤有机质的敏感波段并建立反演模型; 然而经典统计方法假定样本是独立的, 而土壤有机质含量数据之间存在空间相关性, 这与经典统计的假设相矛盾, 其将导致回归系数及其置信区间发生偏移, 难以获得土壤有机质含量精确的反演模型<sup>[14]</sup>。

空间统计考虑到数据间的空间相关性, 可以有效避免这些问题。本文旨在利用多元地统计方法精确获得遥感反演模型, 提高土壤有机质含量的空间预测精度。顾及到回归误差的空间分布特性, 对回归克里金 (regression kriging, RK)<sup>[15-16]</sup>进行改进, 提出了基于局部变化均值的普通克里金 (ordinary Kriging with

收稿日期: 2012-04-16 修订日期: 2012-08-22

基金项目: 中国博士后科学基金 (2012M511571); 国家自然科学基金 (41101413)

作者简介: 陈锋锐 (1982—), 男, 河南禹州人, 博士, 讲师。主要从事多元地统计, 生态遥感遥感信息研究。开封 河南大学黄河中下游数字地理技术教育部重点实验室, 475004。Email: fruch@gmail.com

local varying means, OKLM)，并利用该方法构建土壤有机质含量的遥感定量反演模型，以期消除基于经典统计方法构建模型的不足，获得较精确的空间分布。

## 1 数据与方法

### 1.1 研究区域及数据来源

研究区位于四川省西南部，凉山彝族自治州正南端，地理位置为 E 102°13'~102°24', N 26°33'~26°48', 见图 1。该区主要为农业用地，土壤类型以红壤，黄壤，紫色土为主，土层较深厚，质地适中，富含有机质、速效钾和多种微量元素。

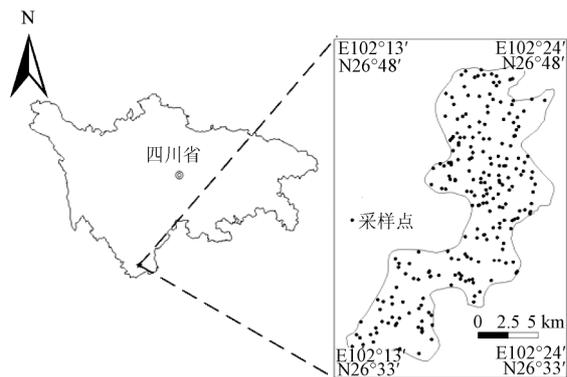


图 1 研究区地理位置和采样点空间分布  
Fig.1 Geographical location of study area and spatial distribution of sampling points

土壤有机质采样时间为 2009 年 4 月，在前茬作物收割后，烟苗栽种之前，此时土壤裸露。采样深度为 0~20 cm，在 1 m<sup>2</sup> 的范围内，均匀采集 6 次，混合形成一个样品，共采集 262 个土壤表层样品，用重铬酸钾稀释热法测定土壤有机质含量。每个采样点的位置均采用 GPS 记录其经纬度，采样点分布见图 1。土壤有机质含量概况如表 1 所示，该区域有机质含量适宜，平均值为 24.7 g/kg。但有机质含量变化幅度大：最高为 62.2 g/kg，而最低仅 6 g/kg，变异系数较大。此外，偏度为 0.84，小于 1，无需进行数据变换，可直接计算变异<sup>[17-18]</sup>。遥感影像数据为采样同时期该区域的 TM 多光谱影像，该影像共有 7 个波段，其中第 6 波段为热红外波段，分辨率为 120 m，其他均为 30 m。

表 1 土壤采样有机质质量分数描述性统计  
Table 1 Descriptive statistics of soil organic matter content

有机质质量分数			标准差	偏度	变异系数
最大值 /(g·kg <sup>-1</sup> )	最小值 /(g·kg <sup>-1</sup> )	平均值 /(g·kg <sup>-1</sup> )			
62.2	6.0	24.7	10.5	0.84	0.43

### 1.2 研究方法

区域变量的变异由 3 部分组成：确定性部分、相关部分和随机部分<sup>[19]</sup>。简单克里金和普通克里金假定区域变量满足二阶平稳或者本征假设，确定性

部分（均值）在全局或局部领域内是常量，主要对随机部分进行估计<sup>[20]</sup>。然而在实际中往往存在更一般的情况，均值在空间上或局部邻域内不再是一个常数而是一个空间变量，此时区域变量是非平稳的。区域变量  $z(u)$  可以用一个确定性漂移  $m(u)$  和一个残差部分  $r(u)$  来表示

$$z(u) = m(u) + r(u) \quad (1)$$

式中， $E[z(u)]=m(u)$ ，且  $E[r(u)]=0$ 。

如果漂移函数  $m(u)$  已知，可以用地统计对剩下的残差进行估值，然后将漂移加到相应位置的残差估值上，其和就是  $z(u)$  的估计值。由于漂移是一个空间变量，且已知。因此， $z(u)$  的空间变异函数不能用常规方法来计算，而是由残差部分  $r(u)$  来确定，于是空间预测式 (1) 改写如下

$$z(u) = m(u) + \sum_{i=1}^{n(x)} \lambda_i(u)r(u_i) \quad (2)$$

式中， $\lambda_i(u)$  为参与估计残差  $r(u_i)$  的权重。

回归克里金，也称为基于局部变化均值的简单克里金（simple kriging with local varying means, SKLM），已经广泛应用于农业<sup>[12,21-22]</sup>、气象<sup>[23-25]</sup>等领域。它利用空间数据间存在着一定程度的相关性，通过回归分析来构建漂移函数，前提是辅助数据与待预测变量之间存在显著的相关关系<sup>[23]</sup>。漂移函数构建如下

$$z(u) = f(y_1(u), y_2(u), \dots, y_n(u)) + \varepsilon(u) \quad (3)$$

式中， $y_1(u)$ ,  $y_n(u)$  为辅助变量； $f(\cdot)$  为回归函数，其数学期望满足  $E[f(\cdot)]=m(u)$ ； $\varepsilon(u)$  为误差，且  $E[\varepsilon(u)]=0$ 。此时，回归函数  $f(\cdot)$  为漂移函数  $m(u)$ ，回归误差  $\varepsilon(u)$  为漂移残差  $r(u)$ 。经典统计处理一维数据时，回归误差满足  $\varepsilon(u) \sim N(0, \sigma^2)$ ，且相互独立<sup>[26-27]</sup>。但是在处理空间数据时，由于没有考虑到数据间的空间相关性，导致误差  $\varepsilon(u)$  在空间上的分布呈现出一定的空间结构，存在空间的变异。回归克里金认为回归误差满足本征假设，即回归误差的均值处处为 0，于是空间预测式 (2) 改写如下

$$z(u) = f(y_1(u), y_2(u), \dots, y_n(u)) + \sum_{i=1}^{n(x)} \lambda_i^{SKLM}(u)\varepsilon(u_i) \quad (4)$$

$\lambda_i^{SKLM}(u)$  为参与估计误差  $\varepsilon(u_i)$  的权重， $\lambda_i^{SKLM}(u)$  满足以下约束

$$\sum_{\beta=1}^{n(x)} \lambda_{\beta}^{SKLM}(u)C_R(u_a - u_{\beta}) = C_R(u_a - u) \quad (5)$$

$$a = 1, \dots, n(x)$$

式中， $C_R(h)$  是误差  $\varepsilon(u)$  的协方差函数， $u_a$  和  $u_{\beta}$  为参与估计误差的坐标， $u$  为待估计误差的坐标。

由于回归误差在空间上呈现一定程度的聚类，

虽然它的均值在总体上为 0，但在局部邻域内不一定为 0。考虑到回归误差在空间分布上的分布特性，本文对回归克里金法进行如下改进：假设回归误差的均值在空间上是变化的，但在同一局部邻域内相等。这样就放宽了假设条件，更加符合回归误差在空间上的分布状况，改进后的多元地统计模型称为基于局部变化均值的普通克里金。式 (4) 改

$$\begin{cases} \sum_{\beta=1}^{n(x)} \lambda_{\beta}^{OKLM}(u) C_R(u_a - u_b) + \mu_0^{OKLM}(u) = C_R(u_a - u) & a = 1, \dots, n(x) \\ \sum_{\beta=1}^{n(x)} \lambda_{\beta}^{OKLM}(u) = 1 \end{cases} \quad (7)$$

式中， $C_R(h)$  是误差  $\varepsilon(u)$  协方差函数， $\mu_0^{OKLM}(u)$  是拉格朗日乘数。

利用该模型进行土壤有机质含量的遥感定量反演主要分为两部分。首先利用遥感影像敏感波段构建漂移函数，由式 (3) 可知，传统上采用经典统计构建的遥感反演模型即为漂移函数；接下来通过式 (6) 和 (7) 来反演土壤有机质含量的空间格局。

### 1.3 数据处理方法

首先对获得的 TM 影像进行几何精校正，采用横轴墨卡托投影 (UTM/WGS84)，并使几何配准精度优于 1 个像元，保证地面测量点位与图像像元位置对应。遥感影像所记录的数据是 DN 值 (digital number value)，是传感器对接收辐射能量转换后的结果。由于传感器接收的地面辐射不可避免地受到太阳高度角、大气条件等的影响，造成地物电磁波波谱的畸变，所以遥感影像需经过辐射校正以消除这些因素的影响，获得地表真实的反射率<sup>[28]</sup>。大气校正采用 ENVI 4.7 中 FLAASH (fast line-of-sight atmospheric analysis of spectral hypercubes) 大气校正模块来完成。

### 1.4 精度检验

土壤有机质含量预测结果评价包括两个部分，一是预测结果的精度评价，通过 47 个验证样本来检验 4 种方法的好坏，采用的检验指标有：均值 (mean)、平均绝对误差 (mean absolute error, MAE) 和均方根误差。平均绝对误差反映样本数据估值的总体误差或精度水平，均方根误差反映样本数据的估值灵敏度和极值。平均绝对误差与均方根误差的值越小，表明预测结果的精度越高。

另一部分是衡量预测误差的空间相关性，好的预测方法其预测误差在空间上独立分布。通过 Morans'I 来衡量预测误差在空间上的分布情况，取值范围在 -1~1 之间，正值表示该属性在空间上具有正相关性，负值表示属性在空间上分布具有负相

关性，零值表示该属性在空间上不存在相关性，即随机分布。

写如下

$$z(u) = f(y_1(u), y_2(u), \dots, y_n(u)) + \sum_{i=1}^{n(x)} \lambda_i^{OKLM}(u) \varepsilon(u_i) \quad (6)$$

$\lambda_i^{OKLM}(u)$  为参与估计误差  $\varepsilon(u_i)$  的权重， $\lambda_i^{OKLM}(u)$  满足以下约束

相关性，零值表示该属性在空间上不存在相关性，即随机分布。

## 2 结果与分析

### 2.1 遥感数据预处理

辐射校正后各波段地表反射率统计分析见表 2，可见光 (波段 1~3) 范围内的地表反射率小于近红外和短波红外 (波段 4~5)。进一步对各波段地表反射率进行相关分析 (表 3) 可知，波段 1~3 的地表反射率相关系数较高，均大于 0.8，说明这三个波段间存在较大的信息冗余，需要通过筛选手段以确定最佳的波段或波段组合；其余波段具有较大的独立性。

表 2 研究区 TM 影像各波段地表反射率统计分析  
Table 2 Statistical analysis for reflectivity of TM bands in study area

波段	波段地表反射率			
	最小值	最大值	均值	标准差
Band1	0.0623	0.147	0.0623	0.0077
Band 2	0.0227	0.0735	0.0373	0.0054
Band 3	0.0163	0.1008	0.0332	0.0076
Band 4	0.0238	0.1948	0.1136	0.0189
Band 5	0.0104	0.1722	0.0954	0.0194
Band7	0.0044	0.09177	0.0342	0.0086

表 3 各波段地表反射率相关系数  
Table 3 Correlation coefficient of reflectivity of TM bands

	Band1	Band2	Band3	Band4	Band5	Band7
Band 1	1	0.8821**	0.8432**	0.2016**	0.0151	0.0564
Band 2	0.8821**	1	0.9221**	0.1372*	0.1790	0.0483
Band 3	0.8432**	0.9221**	1	0.1421*	0.1721	0.0087
Band 4	0.2016**	0.1372*	0.1421*	1	0.0715	-0.0557
Band 5	0.0151	0.1790	0.1721	0.0715	1	0.0293
Band 7	0.0564	0.0483	0.0087	-0.0557	0.0293	1

注：\*表明 0.05 水平上的显著性相关，\*\*表明 0.01 水平上的显著性相关。

### 2.2 回归分析

构建漂移函数时需选择那些与土壤有机质含量显著相关，且相互之间具有较小信息冗余的波

段, 这样更能反映土壤有机质含量的漂移信息。由表 4 可知, 土壤有机质含量与波段 1~3 和波段 5 的相关系数均通过  $p < 0.01$  检验, 它们之间相关性显著, 因此辅助变量可以从这些波段中选择; 而其余波段 (波段 4 和波段 7) 与土壤有机质含量相关性不显著, 不满足作为辅助变量的条件。虽波段 1~3 的地表反射率均与土壤有机质含量显著相关, 但它们之间的相关系数均大于 0.8 (表 3), 存在较大的信息冗余, 需从中选择一个波段作为辅助变量。考虑到波段 2 的相关系数最大, 因此选择波段 2 和波段 5 作为辅助变量来构造土壤有机质含量的漂移函数。

表 4 土壤有机质与各波段地表反射率的相关系数分析  
Table 4 Correlation coefficients analysis between soil organic matter content and reflectivity of TM bands

	Band1	Band2	Band3	Band4	Band5	Band7
相关系数 $r$	0.52	0.61	0.56	0.03	0.46	-0.01
$p$	0	0	0	0.64	0	0.84

由于土壤有机质含量与地表反射光谱关系的复杂性, 构造土壤有机质含量漂移函数时, 同时检验它们之间的线性和非线性关系, 并通过拟合度  $R^2$  和均方根误差 (root mean squared error, RMSE) 来衡量回归函数的优劣。如果  $R^2$  越接近 1, 且均方根误差越小, 则表明该回归函数越优, 反之则越差。回归分析通过 Matlab 2008a 实现, 结果见表 5, 它们均通过  $p < 0.01$  显著性检验。其中 NL-B5 的  $R^2$  最小, 仅为 0.33, 同时它的均方根误差最大 (0.31), 因此 4 种方法中该函数拟合最差。L-B2-B5 的 VIF 值为 1.003, 表明波段 2 和波段 5 地表反射率之间不存在共线性。NL-B2-B5 的拟合度最好, 且均方误差最小, 因此选择它作为土壤有机质含量的漂移函数。其函数定义如下

$$\ln \text{SOM} = 77.2 + 320.4 \times B2 - 892.6 \times B5 - 3438 \times (B2)^2 + 2443.8 \times (B5)^2 \quad (8)$$

式中,  $\ln \text{SOM}$  为土壤有机质含量取对数;  $B2$  和  $B5$  分别为 TM 影像波段 2 和波段 5 的地表反射率。

表 5 回归函数的统计检验

Table 5 Statistical test for regression equation

函数类型	NL-B2	NL-B5	L-B2-B5	NL-B2-B5
$R^2$	0.52	0.33	0.51	0.70
均方根误差	0.27	0.31	0.27	0.24
VIF	--	--	1.003	--

注: NL-B2 和 NL-B5 分别为土壤有机质含量与波段 2 以及波段 5 地表反射率的二次回归函数; L-B2-B5 为土壤有机质含量与波段 2 和波段 5 地表反射率的多元线性回归函数; NL-B2-B5 为土壤有机质含量与波段 2 和波段 5 的多元非线性回归函数。VIF (variance inflation factor) 为方差膨胀因子。

## 2.3 空间变异分析

变异模型的计算采用 GS+ 软件, 参照所有采样点最小平均邻居距离, lag 设置为 300 m。土壤有机质含量以及回归误差的变异模型见表 6。土壤有机质含量的变异模型为高斯模型 (Gaussian), 而回归误差为指数模型 (Exponential)。变异模型反映变量在空间上呈现出的结构性和随机性等特征, 变异模型的变程越大, 说明变量的空间连续性越好, 反之则随机性增强。土壤有机质含量的变程为 1 720 m, 小于回归误差的变程 (2 280 m), 表明后者的空间连续性要大于前者。变量的空间相关性由结构方差/基台值之比来衡量, 值越大说明相关性越强。2 个变量结构方差/基台值的比值均大于 0.5, 具有中等程度的空间相关; 此外, 它们的  $R^2$  都大于 0.9, 说明变异函数具有很好的拟合度。

表 6 各变量半变异模型参数

Table 6 Parameters of semivariogram models of different variables

变量	变异模型	块金	基台	变程	结构方差/基台	$R^2$
有机质	Gaussian	29.1	58.67	1720	0.5	0.94
回归误差	Exponential	13.5	28.3	2280	0.52	0.91

注: lag 设置为 300 m。

## 2.4 空间预测结果及精度评价

为了验证本文提出方法的性能, 其他 3 种方法: 普通克里金、普通遥感定量反演、基于回归克里金的定量反演等方法被用于对比。其中普通克里金、基于回归克里金的定量反演方法采用 GSLIB<sup>[29]</sup> 实现, 普通遥感定量反演采用 Envi 4.7, 本文提出的反演方法在扩展开发 GSLIB 的基础上完成, 所有结果均在 ArcGIS 9.3 中制图 (图 2)。这些方法均反映出研究区土壤有机质含量南部和中部偏高的特征: 南部和中部大部地区土壤有机质质量分数大于 20 g/kg; 而北部地区海拔相对较高, 有机质质量分数小于 15 g/kg 的土壤大部分分布在该区。但是在细节信息呈现上存在较大差异, 普通克里金的空间预测 (图 2a) 在空间上则更趋于平滑, 而后 3 种方法 (图 2b, 2c, 2d) 由于整合了遥感信息, 土壤有机质含量在空间上呈现出不同程度的细节。

从预测结果的值域看, 普通克里金空间预测结果的范围 (10.8~53.6 g/kg) 小于采样数据的值域范围 (6.0~62.2 g/kg), 这是由于普通克里金在一定程度上对土壤有机质含量进行了平滑, 导致较高含量值被过低估计, 较低含量值被过高估计。其他方法由于遥感数据的引入则不存在此类问题, 其反演结果的值域均大于采样数据的取值范围。其中普通遥感定量反演的值域最大, 最大和最小值分别为 71.4 和 3.2 g/kg; 而基于回归克里金和 OKLM 的遥感定量反演的值域则大体相当, 它们与采样数据的值域最接近。

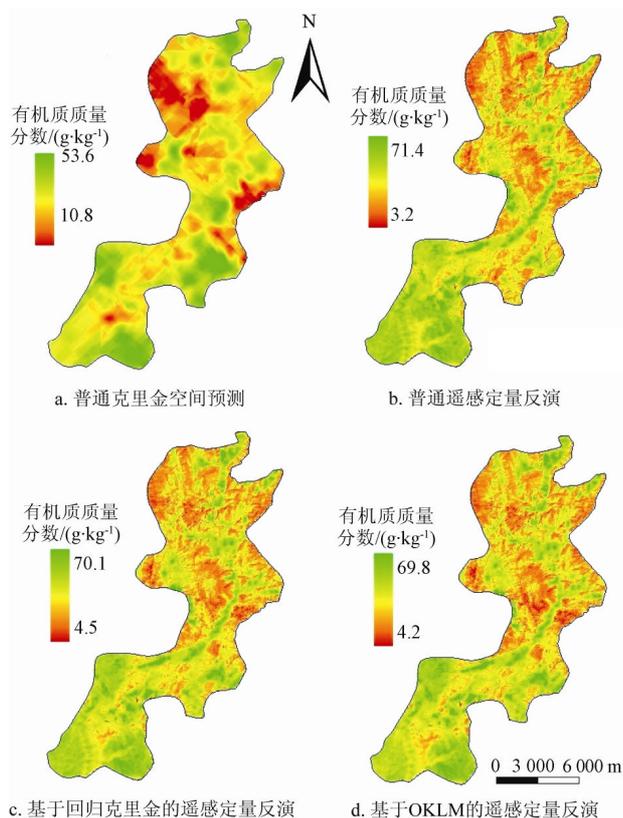


图 2 多种方法土壤有机质含量空间分布预测结果

Fig.2 Predicted mapping of soil organic matter content with different methods

4 种方法的精度评价见表 7。仅采用地面采样数据的普通克里金法精度最差，其均方根误差和平均绝对误差分别为 7.05 和 5.52，均大于其他方法。普通遥感定量反演方法由于遥感数据的引入，提供了未采样位置的先验信息，其预测精度要远优于仅采用地面采样数据的空间预测；然而该模型的构造是基于经典统计的，所以其误差大于其他 2 种基于空间统计构建模型的方法。本文提出的基于 OKLM 的遥感定量反演方法，其 3 项精度评价指标均要小于其他方法，因此，该方法的预测精度最高。

表 7 土壤有机质含量空间预测精度检验

Table 7 Prediction accuracy of soil organic matter content with different methods

	普通克里金空间预测	普通遥感定量反演	基于回归克里金的遥感定量反演	基于 OKLM 的遥感定量反演
均值	-0.27	-0.56	0.03	0
平均绝对误差	5.52	3.93	3.72	3.61
均方根误差	7.05	5.32	4.93	4.77

注：OKLM 为基于局部变化均值的普通克里金。

普通遥感定量反演的误差分析结果见图 3a，Moran's I 指数为 0.17，且通过置信度为 0.01 显著性检验，具有较强的聚集。这与回归分析要求的，误差应

独立分布的假设相矛盾，因此基于经典统计不能很好地构建遥感反演模型。其他 3 种方法(图 3b, 3c, 3d)由于采用空间统计方法，预测误差在空间上呈现显著的随机分布。本文提出基于 OKLM 的遥感反演方法，其误差的 Moran's I 为-0.06；此外，该方法误差的半变异模型是一个与 x 轴平行的直线(图 4)，这些均表明其误差在空间上不具有连续性和相关性。

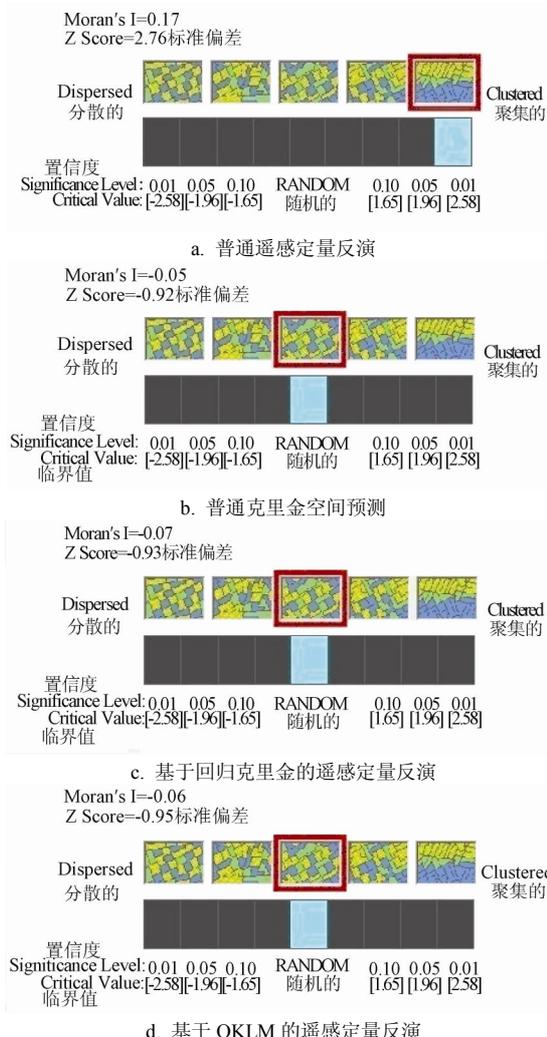


图 3 土壤有机质含量预测误差的 Moran's I

Fig.3 Moran's I of prediction error of soil organic matter content

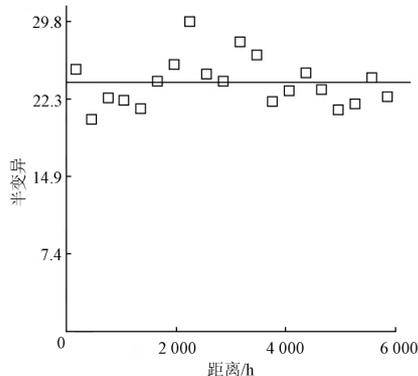


图 4 基于 OKLM 的遥感定量反演误差的半变异分析  
Fig.4 Semi variance analysis of prediction error for remote sensing retrieval based on OKLM

### 3 讨论

造成普通克里金法空间预测精度最低的原因主要有 2 点: 一是该方法仅采用采样数据进行空间预测, 而采样数据的数据量有限, 导致其空间预测精度存在一定的极限。其次, 普通克里金属于平稳地统计方法, 它的变异是整个研究区变异的数学期望, 一些局部细节信息被过滤, 整体趋势得以保留, 因此, 其空间预测结果对较大值过低估计, 而对较小值过高估计, 导致空间预测值域小于采样数据的值域。

普通遥感定量反演方法根据土壤的光谱特性与有机质含量之间的关系, 综合利用地面采样和遥感数据对土壤有机质含量进行预测, 因此其空间预测精度要远优于仅采用地面采样数据的普通克里金法。然而该方法通过经典回归分析来构建定量反演模型, 经典回归分析假定样本独立, 这样方可使预测误差相互独立。这与土壤有机质含量的空间分布特性相矛盾, 致使普通遥感定量反演方法获得的结果仍存在一定程度的偏移, 其预测精度小于采用回归克里金和 OKLM 的遥感定量反演方法。

基于多元地统计(回归克里金和 OKLM)的遥感定量反演方法在构建反演模型过程中, 不仅考虑到数据间的普通相关关系, 而且还顾及数据间的空间相关关系, 因此, 它们的预测误差在空间上呈现独立分布, 较为精确地获得研究区土壤有机质含量的空间分布。本文提出的基于 OKLM 的遥感定量反演方法, 充分考虑到回归误差在空间上的分布特性, 即回归误差在空间上呈现一定程度的聚类, 导致回归误差的均值在整个区域上是变化的, 因此, 该方法要优于基于回归克里金的遥感定量反演方法。

### 4 结论

遥感定量反演的关键是准确构建反演模型。针对传统遥感定量反演采用经典统计构建反演模型的不足, 本文提出了基于多元地统计的遥感反演方法。考虑回归误差在空间上的分布特征, 提出了 ordinary Kriging with local varying means (OKLM), 然后运用该方法来构建土壤有机质含量反演模型, 并与普通克里金、普通遥感定量反演以及基于回归克里金的遥感定量反演等方法进行对比。结果表明: 本文提出的基于 OKLM 的遥感定量反演方法的均值、平均绝对误差、均方根误差均优于其他 3 种方法, 具有最高的反演精度。同时还发现: 遥感定量反演方法由于综合了采样点以及遥感影像信息, 因此其空间预测精度要优于仅采用采样数据的空间预测方法; 相比传统的遥感定量反演方法, 采用空间统计方法构建遥感反演模型, 不仅有效提高空间预测的精度, 而且使误差在空间上独立分布。

### [参 考 文 献]

- [1] Goovaerts P. Geostatistics for Natural Resources Evaluation[M]. New York: Oxford University Press, 1997: 132—138.
- [2] 何挺, 王静, 林宗坚, 等. 土壤有机质光谱特征研究[J]. 武汉大学学报: 信息科学版, 2006, 31(2): 975—979. He Ting, Wang Jing, Lin Zongjian, et al. Spectral features of soil organic matter[J]. Geomatics and Information Science of Wuhan University, 2006, 31(2): 975—979. (in Chinese with English abstract)
- [3] 张娟娟, 田永超, 朱艳, 等. 不同类型土壤的光谱特征及其有机质含量预测[J]. 中国农业科学, 2009, 42(9): 3154—3163. Zhang Juanjuan, Tian Yongchao, Zhu Yan, et al. Spectral characteristics and estimation of organic matter contents of different soil types[J]. Scientia Agricultura Sinica, 2009, 42(9): 3154—3163. (in Chinese with English abstract)
- [4] Ben-Dor E, Inbar Y, Chen Ying. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400-2500 nm) during a controlled decomposition process[J]. Remote Sensing of Environment, 1997, 61(4): 1—15.
- [5] Gomez C, Viscarra Rossel R A, McBratney A B. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study[J]. Geoderma, 2008, 146(4): 403—411.
- [6] Hill J, Schütt B. Mapping complex patterns of erosion and stability in dry Mediterranean ecosystems[J]. Remote Sensing of Environment, 2000, 74(5): 557—569.
- [7] 张法升, 曲威, 尹光华, 等. 基于多光谱遥感影像的表层土壤有机质空间格局反演[J]. 应用生态学报, 2010, 21(4): 883—888. Zhang Fasheng, Qu Wei, Yin Guanghua, et al. Spatial pattern of surface soil organic matter based on remotely sensed multispectral image[J]. Chinese Journal of Applied Ecology, 2010, 21(4): 883—888. (in Chinese with English abstract)
- [8] 李欣宇, 宇万太, 李秀珍. 基于 TM 影像的表层土壤有机碳空间格局[J]. 生态学杂志, 2008, 27(3): 333—338. Li Xinyu, Yu Wantai, Li Xiuzhen. Spatial distribution pattern of surface organic matter based on TM image[J]. Chinese Journal of Ecology, 2008, 27(3): 333—338. (in Chinese with English abstract)
- [9] 程朋根, 吴剑, 李大军, 等. 土壤有机质高光谱遥感和地统计定量预测[J]. 农业工程学报, 2009, 25(3): 142—148. Cheng Penggen, Wu Jian, Li Dajun, et al. Quantitative prediction of soil organic matter content using hyperspectral remote sensing and geo-statistics[J]. Transaction of the Chinese Society of Agriculture Engineering (Transactions of the CSAE), 2009, 25(3): 142—148. (in Chinese with English abstract)
- [10] 邢宇, 姜琦刚, 李远华, 等. 基于 RS 和 GIS 的松辽平原黑土信息提取[J]. 农业工程学报, 2010, 26(10): 212—217. Xing Yu, Jiang Qigang, Li Yuanhua, et al. Extracting phaeozem information in Songliao Plain based on remote sensing and GIS[J]. Transaction of the Chinese Society of Agriculture Engineering (Transactions of the CSAE), 2010, 26(10): 212—217. (in Chinese with English abstract)
- [11] 徐明星, 周生路, 丁卫, 等. 苏北沿海滩涂地区土壤有机质含量的高光谱预测[J]. 农业工程学报, 2011, 27(2): 219—223. Xu Mingxing, Zhou Shenglu, Ding Wei, et al. Hyperspectral reflectance models for predicting soil organic matter content in coastal tidal land area, northern Jiangsu[J]. Transaction of

- the Chinese Society of Agriculture Engineering (Transactions of the CSAE), 2011, 27(2): 219–223. (in Chinese with English abstract)
- [12] 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. Zhang Sumei, Wang Zongming, Zhang Bo, et al. Prediction of spatial distribution of soil nutrients using terrain attributes and remote sensing data[J]. Transaction of the Chinese Society of Agriculture Engineering (Transactions of the CSAE), 2010, 26(5): 188–194. (in Chinese with English abstract)
- [13] 刘焕军, 赵春江, 王纪华, 等. 黑土典型区土壤有机质遥感反演[J]. 农业工程学报, 2011, 27(6): 211–21. Liu Huanjun, Zhao Chunjiang, Wang Jihua, et al. Soil organic matter predicting with remote sensing image in typical blacksoil area of Northeast China[J]. Transaction of the Chinese Society of Agriculture Engineering (Transactions of the CSAE), 2011, 27(6): 211–21. (in Chinese with English abstract)
- [14] 王劲峰. 空间分析[M]. 北京: 科学出版社, 2006: 2–3.
- [15] Hengl T, Heuvelink G, Stein A. A generic framework for spatial prediction of soil variables based on regression-kriging[J]. Geoderma, 2004, 20(3): 75–93.
- [16] Hengl T, Heuvelink G, Rossiter D G. About regression-kriging: From equations to case studies[J]. Computers & geosciences, 2007, 33(5): 1301–1315.
- [17] Kerry R, Oliver M. Determining the effect of asymmetric data on the variogram I: Underlying asymmetry[J]. Computers & Geosciences, 2007, 33(4): 1212–1232.
- [18] Kerry R, Oliver M. Determining the effect of asymmetric data on the variogram II: Outliers[J]. Computers & Geosciences, 2007, 33(3): 1233–1260.
- [19] 张仁铎. 空间变异理论及应用[M]. 北京: 科学出版社, 2005: 40–41.
- [20] Webster R, Oliver M. Geostatistics for Environmental Scientists[M]. New York: John Wiley and Sons Inc, 2007: 145–147.
- [21] Zhu Qu, Lin Haisheng. Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes[J]. Pedosphere, 2010, 20(5): 594–606.
- [22] Li Yong. Can the spatial prediction of soil organic matter contents at various sampling scales be improved by using regression kriging with auxiliary information?[J]. Geoderma, 2010, 159(4): 63–75.
- [23] 陈锋锐, 崔伟宏, 彭光雄, 等. 融合多源信息的地表气温插值研究[J]. 中国图象图形学报, 2011, 16(9): 1708–1715. Chen Fengrui, Cui Weihong, Peng Guangxiong, et al. Surface air temperature interpolation based on multiple sources information fusion[J]. Journal of Image and Graphics, 2011, 16(9): 1708–1715. (in Chinese with English abstract)
- [24] Moral F J. Comparison of different geostatistical approaches to map climate variables: Application to precipitation[J]. International Journal of Climatology, 2010, 30(4): 620–631.
- [25] Watt M S, Palmer D J. Use of regression kriging to develop a Carbon: Nitrogen ratio surface for New Zealand[J]. Geoderma, 2012, 183(2): 49–57.
- [26] 古扎拉蒂. 计量经济学基础[M]. 北京: 中国人民大学出版社, 2005: 27–28.
- [27] Haining R P. Spatial Data Analysis: Theory and Practice[M]. Cambridge: Cambridge Univ Press, 2003: 3–4.
- [28] 梁顺林, 范闯捷. 定量遥感[M]. 北京: 科学出版社, 2009: 121–123.
- [29] Deutsch C, Journel A. GSLIB: Geostatistical Software Library and User's Guide[M]. New York: Oxford University Press, 1998: 272–295.

## Inversion for spatial distribution of soil organic matter content based on multivariate geostatistics

Chen Fengrui<sup>1,2</sup>, Qin Fen<sup>1,2</sup>, Li Xi<sup>3</sup>, Peng Guangxiong<sup>4</sup>

(1. Key Laboratory of Geospatial Technology for Middle and Lower Yellow River Regions, Ministry of Education, Henan University, Kaifeng 475004, China; 2. College of Environment and Planning, Henan University, Kaifeng 475004, China; 3. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; 4. School of Geosciences and Info-Physics, Central South University, Changsha 410083, China)

**Abstract:** The classical statistical method is always used to construct quantitative remote sensing retrieval model. However, the method doesn't take into account the spatial relations between data, which will severely affect the retrieval accuracy. In order to improve the spatial predictive accuracy of soil organic matter, a multivariate geospatial method for making retrieval model was presented in this paper. Considering the spatial distribution characteristic of regression error, a multivariate geostatistical method called ordinary Kriging with varying local means (OKLM) was presented, which was used to construct remote sensing retrieval model. The method was illustrated using soil organic matter (SOM) content in Southwest Sichuan province, and was compared with other method, such as ordinary Kriging, ordinary remote sensing retrieval method, and remote sensing retrieval model based on regression Kriging. The results showed the proposed method improved the predictive accuracy effectively among these methods, because the proposed method was based on relations between SOM sampling data and TM images using spatial statistics, taking fully into account the spatial relations among the data, and obtained more accurate retrieval model. Compared with regression Kriging, OKLM assumed that the means of regression errors cannot always be zero in local neighborhood, which was more in line with the actual situation. The proposed method provides a scientific basis for the farmland nutrient management and sustainable development of the regional agricultural.

**Key words:** soils, remote sensing, regression analysis, organic matter, retrieval, geostatistics