

张春来, 陆来谋, 杨慧, 等. 岩溶区土壤有机质空间变异性分析[J]. 中国岩溶, 2022, 41(2): 228-239.
DOI: 10.11932/karst20220205

岩溶区土壤有机质空间变异性分析

张春来¹, 陆来谋^{1,2}, 杨慧¹, 黄芬¹

(1. 中国地质科学院岩溶地质研究所/自然资源部、广西岩溶动力学重点实验室, 广西 桂林 541004;
2. 广西环投金溢环境科技有限公司, 广西 南宁 530299)

摘要: 采用GIS和地统计学研究土壤有机质(SOM)的空间分布、影响因素和预测是指导农业生产、环境治理和土壤碳储计量的重要手段。基于广西马山县北部岩溶区表层土壤(0~20 cm)的441个SOM数据, 建立普通克里格(OK)、回归克里格(RK), 以及结合辅助变量的地理加权回归克里格(GWRK)、残差均值(MM_OK)和中值(MC_OK)均一化克里格的5种模型, 并比较其预测精度, 旨在探讨岩溶区SOM制图中地统计学方法的适用性。结果表明: (1) SOM的变异系数为37.30%, 属于中等空间变异; (2) 岩溶区SOM空间变异受土地利用方式、土壤类型和地形因子等因素共同影响, SOM高值区分布在西北部、西部和东部等石灰土分布的岩溶区和水田, 低值区位于北部红水河沿岸的冲积土地带; (3) RK、GWRK、MM_OK和MC_OK对SOM解释能力均较优, 可用于岩溶区SOM预测制图。结合辅助变量因子的GWRK预测模型能有效消除空间变异因素的影响, 克服岩溶区SOM含量的空间非平稳性, 从而提高SOM含量模型的稳定性和精度, 同时MC_OK模型能提高预测的准确度。

关键词: 土壤有机质; 地理加权回归; 空间分布; 岩溶

中图分类号: S153 文献标识码: A

文章编号: 1001-4810(2022)02-0228-12 开放科学(资源服务)标识码(OSID):



0 引言

土壤表层是受人类活动影响最为频繁的区域, 土壤有机质(Soil Organic Matter, SOM)含量容易受到地质背景、气候、农业活动和土地利用等因素影响^[1-2], 具有复杂的空间变异性和非稳定性, 致使对其空间分布定量预测难度增大^[3]。然而, 揭示表层SOM空间变异规律及其影响因素对指导农业生产有着极其重要的作用。研究表明: 地理信息系统(Geographic Information System, GIS)和地统计学结合应用于土壤属性的空间变异研究, 能更好地预测其空间分布特征, 在改善农业、管理耕地等方面具有重要意义, 是土壤学、地球化学、农学研究的重点

内容之一^[4-5]。Wang等^[6]使用反距离加权法(Inverse Distance Weighted, IDW)来生成土壤有机碳(Soil Organic Carbon, SOC)含量的空间分布; Zhao等^[7]使用半方差函数和普通克里格(Ordinary Kriging, OK)插值获得SOC含量的空间变异性。OK方法简单、灵活性显著, 但预测结果仅能通过采样点密度情况描述土壤属性的整体空间分布规律, 对局部信息描述不够详细, 且平滑效应严重^[8]。添加辅助变量是一种提高数字土壤测绘精度的方法, 可提高空间预测的效率^[9]。回归克里格法(Regression Kriging, RK)在空间预测上既考虑土壤属性分布过程中的结构性因素, 还有不确定性的随机因子, 以及空间上的邻近点和其他地理要素的相关性, 有助于提高预测

基金项目: 国家自然科学基金重点项目(41530316); 广西科技计划项目(桂科AD20297090); 中国地质调查评价项目(DD2016032403)

第一作者简介: 张春来(1984—), 男, 助理研究员, 主要从事岩溶地球化学和岩溶碳循环研究。E-mail: chlzhang@yeah.net。

收稿日期: 2021-02-25

精度^[10-12]。

岩溶地区具有极其复杂的二元空间异质体结构^[13],地形地貌、碳的水气固交换条件、植被条件、岩石风化等,与非岩溶区均存在较大差异^[14-15],岩溶地区的 SOM 空间分布具有独特的特征^[16-17],SOM 含量空间变异性较其他区域变化较大,且不同部位的影响因素不同^[18]。高度异质性的小生境决定了土壤有机碳的空间结构和格局,而受重度人类活动干扰的人工林和未受干扰的原生林空间异质性较次生林降低^[19]。Bai 和 Zhou^[20]对贵州普定后寨流域有机碳的研究发现,SOC 呈中等程度变异,成土母质、土壤类型、土地使用、坡位、坡向和岩石裸露率对 SOC 产生了显著影响。钟聪等^[1]对广西表层 SOM 空间变异研究发现,SOM 含量在不同土壤类型和不同土地利用方式条件下具有一定差异。吴敏等^[21]基于岩石裸露率、土深校正的空间插值方法大大降低了估算岩溶高基岩裸露率坡地表层 SOC 储量和碳密度的误差。

在以往的研究中,由于岩溶区土壤的高度异质性,使得对 SOC 的空间预测的精度不高,土壤变异性和缺乏足够的采样数据是预测精度较低的原因^[22]。本文利用 1:5 万土地质量地球化学调查的大批量数据,分别采用普通最小二乘法 (Ordinary Least Squares, OLS) 模型和地理加权回归 (Geographically Weighted Regression, GWR) 模型拟合,得到结合土地利用、土壤类型、地形因子的 SOM 趋向项和残差项,

使用文献 [4] 的研究方法分别建立 OK、RK、地理加权回归克里格 (Geographically Weighted Regression Kriging, GWRK)、中值中心化克里格 (Ordinary Kriging of Median Centralization, MC_OK) 和均值中心化克里格 (Median Modified Ordinary Kriging, MM_OK) 5 种方法模型,对预测结果进行比较,优选出研究区域最适合的预测模型,并对研究区 SOM 分布格局进行讨论,以期了解不同土地利用方式、土壤类型下 SOM 变异性,分析异质性变化趋势和影响因子,为今后岩溶区 SOM 研究的空间样品采集布设、农业生产和重金属防控等提供依据。

1 研究区概况

研究区位于广西马山县北部(图 1),地处东经 108.09°~108.2°,北纬 23.73°~23.85°,面积为 64 km²,属于亚热带季风气候区,年平均气温为 21.3 °C,年均降雨量在 1 480~1 667 mm,但四季分布不均,春—秋多干旱,夏季偏洪涝。研究区属于中国南方典型岩溶地貌,土地利用方式多样,水田、旱地、灌木林地、有林地、果园和草地的分布比例分别为 10.55%、29.39%、29.49%、12.86%、0.70% 和 0.66%,其中,水田位于谷地,土壤类型为水稻土;旱地分布在谷地两侧坡底和岩溶洼地底部,土壤类型为红壤、赤红壤和棕色石灰土、冲积土;灌木林地、有林地分布在山体,主要土壤类型为棕色石灰土、赤红壤、红壤。

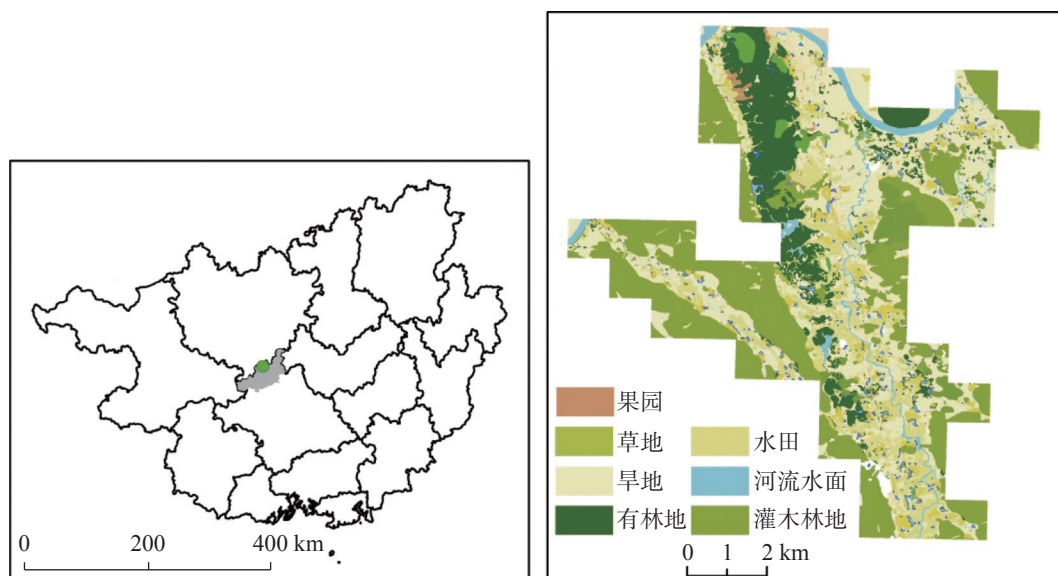


图 1 研究区位置及土地利用方式图

Fig. 1 Location of the study area and the land use

2 研究方法

2.1 样品采集与测试

在叠加了等高线和公里网的 1:5 万土地利用图上,以 500 m×500 m 为单元兼顾土地利用方式和土壤类型采集表层土壤,采用“S”形、“X”形或棋盘形以 20~50 m 的间距采集 3~5 等份的子样组合成一

个样品,4~12 个样·km⁻²,布设在土壤易于汇集的平缓坡地、山间平坝、低洼地带等土壤分布具有代表性的部位。各样品类型数量见表 1。土壤样品采集、加工和测试于 2017 年 7 月进行,使用重铬酸钾—硫酸标准溶液作为氧化剂,硫酸亚铁铵溶液为滴定剂测定土壤有机质含量^[23],测试精度满足《土地质量地球化学评价规范》(DZT0295-2016)要求。

表 1 SOM 描述性统计
Table 1 Descriptive statistics of SOM

指标	样本数	最大值/ %	最小值/ %	平均值/ %	标准差 (SD)	变异系数 (CV)/%	峰度	偏度	K-S检验	
土地利用方式	水田	83	5.03	1.66	3.56 ^a	0.69	19.90	-	-	-
	旱地	259	4.45	0.81	1.92 ^c	0.58	30.20	-	-	-
	林地	99	3.93	1.02	2.51 ^b	0.64	25.50	-	-	-
土壤类型	水稻土	98	5.03	0.91	3.19 ^a	0.93	29.10	-	-	-
	红壤	66	4.48	0.93	2.12 ^b	0.77	36.10	-	-	-
	赤红壤	122	4.45	1.02	2.18 ^b	0.66	30.20	-	-	-
	石灰土	95	4.06	1.03	2.28 ^b	0.65	28.30	-	-	-
	冲积土	60	4.30	0.81	1.78 ^c	0.72	40.40	-	-	-
总样本	441	5.03	0.81	2.36	0.88	37.30	2.83	0.76	0.00	
对数转换	-	-	-	-	-	-	5.52	0.03	0.15	

注:平均值列的不同小写字母对应显著性表现。

2.2 数据来源与方法

2.2.1 数据来源和处理

数字高程模型(Digital Elevation Model, DEM)数据来源于地理空间数据云网站,分辨率为 30 m,利用 ArcGIS 10.2 提取各采样点的坡度、坡向以及高程等地形因子(图 2)。

研究区土地利用方式和土壤类型数据来源于 2014 年度土地利用现状变更调查和全国第二次土壤普查成果。通过哑变量赋值方法^[24]将土壤类型和土地利用类型等定性变量转为定量变量,以满足地统计分析,如设旱地类的值为 1,其他值为 0,林地为参照组;水稻土值为 1,其他值为 0,冲积土为参照组等。再分别利用 SPSS 20 中的 Pearson 参考点二列相关方法与 SOM 做相关性分析^[25]。

采用平均值±3×标准差法剔除 SOM 异常值,最终得到 441 个样品数据(表 1),利用 ArcGIS 软件地统计模块将其中的 80% 作为训练集,其余 20% 作为验证集。

2.2.2 OLS、RK、GWR 和 GWRK 模型与插值

普通最小二乘法是一种全局的线性回归,利用一个或多个解释变量对因变量进行最佳预测的方法。通过利用多个观测值求得一组解释变量与因变量之间残差平方和最小的拟合。其公式如下:

$$y = \sum_{i=1}^N x_i \beta_i + \varepsilon \quad (1)$$

式中: y 为研究区 SOM 的预测值; x_i 为样品 SOM 含量; β_i 为未知的模型拟合参数; N 为建模中的样品数; ε 为模型的残差项,服从正态分布 $N(0, \rho^2)$ 。通过下列矩阵方程估计:

$$\widehat{\beta}_i = (X'X)^{-1} X'Y \quad (2)$$

式中: $\widehat{\beta}_i$ 为线性模型系数; X 为数据集; X' 为 X 的转置矩阵; Y 为目标数据集。

RK 是 OLS 模型的进一步扩展分析,是对 OLS 模型拟合的残差部分进行 OK,然后将残差部分与 OLS 模型拟合的预测值进行相加,表达式如下:

$$Y_{RK} = \sum_{i=1}^N x_i \beta_{(u_i, v_i)} + \widehat{\varepsilon}_{ok}(u_i, v_i) \quad (3)$$

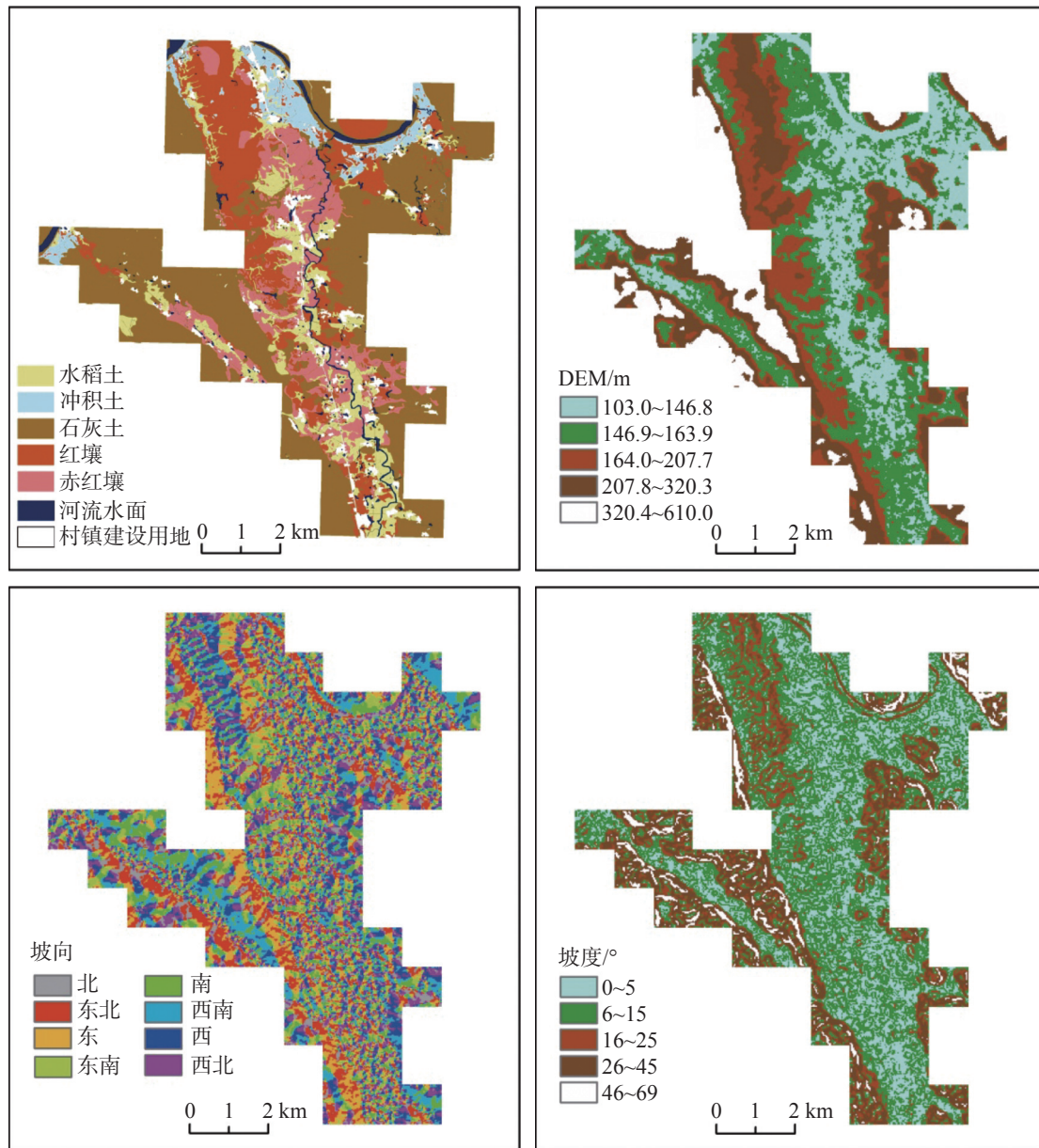


图 2 研究区土壤类型及地形因子空间特征
Fig. 2 Spatial features of ancillary variables in the study area

式中: (u_i, v_i) 为 i 点的坐标; $\beta_{(u_i, v_i)}$ 为 i 点的回归参数; $\hat{\varepsilon}_{ok}(u_i, v_i)$ 为 OK 模型在 i 点的残差值。

GWR 模型将数据集中的各要素构建成局部回归方程, 并将空间距离权重加入到拟合模型中, 构建空间权重矩阵, 从而更好地表示局部区域的空间非平稳性。通过对未知点带宽内所有的观测值进行局部线性回归以获取该点的预测值, 并消除空间异质性对模型拟合的干扰, 是一种局部参数的估计方法。其表达式如下:

$$Y_{GWR} = \beta_k(\mu_i, v_i) + \sum_{k=1}^N \beta_k(\mu_i, v_i) x_{ik} + \varepsilon_i \quad (4)$$

式中: (μ_i, v_i) 为 i 点的坐标; $\beta_k(\mu_i, v_i)$ 为 i 点的回归参数; 当 $k=0$ 时, $\beta_0(\mu_i, v_i)$ 为 i 点的截距项; ε_i 为残差项, 遵循正态分布 $N(0, \rho^2)$ 。

估计模型矩阵方程如下:

$$\hat{\beta}_i = (X' W_i X)^{-1} X' W_i Y \quad (5)$$

式中: W_i 为 i 点处周围参与建模的观测值的空间权重。

GWR 模型中通过 Gaussian 函数计算各点权重, 表达式如下:

$$W_{ij} = 1 - \left[\left(\frac{d_{ij}}{b} \right)^2 \right]^2 \quad (6)$$

式中: d_{ij} 为 i 点到周围观测样点的距离, 其值越大, 所对应的权重越小; b 为带宽, 带宽为 GWR 模型的关键, 本研究模型带宽的选择以修正赤池信息量准则 (Correction Akaike Information Criterion, AICc) 作为准则, 值最小时的 b 值作为最佳带宽。

GWRK 模型是通过 GWR 模型拟合得到趋势项和残差项, 对残差项进行普通克里格插值 (OK), 然后空间加和, 得到趋势项与残差项之和, 即为预测结果。

对创建训练集数据、OLS 残差、GWR 残差、GWR 残差均值与中值中心化, 进行 Kolmogorov-Smirnov 正态分布检验, 利用半方差函数拟合最优模型函数和模型参数^[1], 再进行普通克里格插值, 然后和经土地利用、土壤类型、地形因子等辅助变量因子回归获得的 OLS、GWR 趋势项空间加和, 最终得到 OK、RK、GWRK、OK_MC 和 MM_OK 等 5 种 SOM 分布预测图。

2.2.3 模型精度验证

模型精度验证主要通过预测模型训练集 (内部验证) 和验证集 (外部验证) 两种手段^[4]。内部验证: 标准平均值误差 (Mean Standardized Error, MSE) 与 0 最接近、标准均方根 (Root-Mean-Square Standardized, RMSS) 预测误差和 1 最接近、均方根 (Root-Mean-Square, RMS) 预测误差达到最小、平均标准误差 (Average Standard Error, ASE) 和均方根 (Root-Mean-Square, RMS) 预测误差最接近的结果为最优。外部验证: 均方根误差 (Root Mean Squares Error, RMSE) 越小模型越稳定、平均绝对值误差 (Mean Absolute

Error, MAE) 越小模型精度越高^[26], 准确度 (Accuracy, AC) 可评价预测的准确性, 取值范围为 0~1, 越接近于 1 说明预测越准确^[27]。

3 结果与分析

3.1 描述性统计及方差分析

研究区水热条件较好, 土壤较为肥沃, SOM 观测值在 0.81%~5.03%, 平均值为 2.36%±0.88%。利用 SPSS 20.0 软件对不同样品 SOM 数据进行描述性统计分析 (表 1), 并通过单因素方差分析进行显著性差异检验。研究区域不同土地利用类型的 SOM 均值均存在显著性差异 ($P < 0.01$, 下同), 其中具体表现为水田 > 林地 > 旱地; 同时水稻土、冲积土和其他土壤类型呈显著性差异, 土壤类型均值具体表现为水稻土 > 石灰土 > 赤红壤 > 红壤 > 冲积土。总体样本 CV 为 37.3%, 属于中等空间变异, 偏度为 0.76, 峰度为 2.83, K-S 检验显示原始数据不呈正态分布, 经对数转换后 $P = 0.146 > 0.05$, 服从正态分布。

SOM 含量差异主要是由不同土地利用方式或土壤类型引起的, 因此, 在对 SOM 插值前得消除它们引起的 SOM 差异。对不同土地利用类型、土壤类型的哑变量赋值结果使用点二列相关法分别与 SOM 含量作相关性分析 (表 2) 表明, SOM 含量和水田呈极显著正相关, 水田为还原环境, 有利于 SOM 保存和积累; 与旱地呈极显著负相关, 旱地多处于氧化环境, 扰动频繁, SOM 降解快。水稻土和 SOM 含量呈极显著正相关, 冲积土、赤红壤、红壤呈极显著负相关, 这与土地利用方式分析结果基本一致。

表 2 研究区 SOM 含量与土地利用方式和土壤类型 Pearson 相关系数

Table 2 Pearson correlation coefficient of SOM and types of land use and soil

	水田	旱地	林地	水稻土	冲积土	赤红壤	红壤	石灰土
相关系数	0.676**	-0.622**	0.090	0.526**	-0.269**	-0.153**	-0.131**	-0.047

注: **在 0.01 水平 (双侧) 上显著相关。

3.2 OLS、GWR 回归模型

3.2.1 OLS 模型

对筛选出的变量进行 OLS 模型拟合, 由拟合诊断结果 (表 3) 可知: OLS 模型带宽 AICc 为 787 m, 校正可决系数 R^2 的值为 0.56, 表明模型拟合较好, 说明

不同土地利用、土壤类型、地形因子等变量可解释因变量中大约 56% 的变化。K(BP)_Prob=0.002<0.05, 说明模型不稳定或存在空间异质性, 系数显著性通过稳健概率 (Robust_Pr) 来确定; 使用卡方统计量来确定整个模型是否显著, Wald-Prob 为 0.00, 说明整个模型是显著的; Jarque-Bera 检验 JB-Prob 为 0.00,

表 3 OLS 模型诊断结果
Table 3 Diagnostic results of OLS model

AICc	R ²	F-Stat	F-Prob	Wald-Prob	K(BP)	K(BP)_Prob	JB-Prob
787	0.56	63.9	0.00*	0.00*	25.6	0.002*	0.00*

注：**在0.01水平(双侧)上显著相关。

小于 0.05,表明研究区 SOM 含量的 OLS 模型结果是片面的。残差 Moran's I 指数为 0.39,呈空间自相关^[28]。

由不同样品类型辅助变量因子的诊断系数(表 4)可知:旱地、冲积土的系数均为负值,表明解释变量和因变量为负相关关系,水田、水稻土、赤红壤、石灰土等系数为正值,表明与因变量为正相关关系。高程、坡度、坡向、赤红壤等解释变量的系数非常接近零,表明该解释变量对模型解释程度较弱,SOM 含量受地形因子、赤红壤的因素影响较小。系数诊断表中水田、旱地、水稻土、石灰土、冲积土的 Robust_Pr 系数均显著,表示该解释变量对拟合的模型非常重要;模型方差膨胀因子(Variance Inflation Factor, VIF)值小于 7.5 不存在多重共线性问题。

3.2.2 GWR 模型

采用虚拟变量后不同土地利用类型、不同土壤类型以及 DEM 提取的地形因子作为解释变量进行 GWR 模型的参数回归。本文使用的核函数为自适

应高斯核函数,GWR 模型拟合参数见表 5。AICc 为带宽判断标准,值越小说明拟合不确定性越小,一般两个模型的 AICc 值相差大于 3,则具有较低 AICc 值的模型为更佳模型,GWR 模型中 AICc 值最佳带宽为 774 m,OLS 模型为 787 m,具有较低 AICc 值的 GWR 模型能够更好地拟合观测数据,并且 GWR 模型 R² 为 0.64,校正后的 R² 为 0.59,满足了统计模型的基本要求。SOM 含量与旱地和坡度的回归系数均为负值,呈负相关关系,而与水田、高程和坡向的回归系数均为正值,呈正相关关系,这与 OLS 的相关性检验结果基本一致。

3.3 空间变异特征分析

通过将研究区 GWR 模型残差以及其残差中值、均值中心化和 OLS 模型残差进行正态分布检验后,利用 GS+9.0 软件进行半方差函数拟合,得到参数(表 6)与 OK 插值结合,再将插值得到的结果与相应趋势项进行空间加和而得到研究区 SOM 的预测值^[4],

表 4 OLS 模型诊断系数结果
Table 4 Results of diagnostic coefficient of OLS model

变量	系数	P	Robust_SE	Robust_t	Robust_Pr	StdCoef	VIF	
截距	2.18	0.00	0.26	8.33	0.00	0.00	-	
土地利用	水田	0.91**	0.00	0.15	6.13	0.00	0.41	3.06
	旱地	-0.56**	0.00	0.08	-6.71	0.00	-0.31	2.02
土壤类型	水稻土	0.28**	0.01	0.13	2.16	0.03	0.13	2.99
	赤红壤	0.01	0.95	0.09	0.06	0.95	0.00	2.02
	石灰土	0.21*	0.03	0.09	2.33	0.02	0.10	1.98
	冲积土	-0.24*	0.02	0.09	-2.84	0.00	-0.10	1.58
地形因子	高程	0.00	0.17	0.00	1.17	0.24	0.06	2.17
	坡度	0.00	0.61	0.00	-0.52	0.60	-0.02	1.89
	坡向	0.00	0.58	0.00	-0.52	0.60	-0.02	1.09

注:**和*分别在0.01和0.05水平(双侧)上显著相关;土壤类型、土地利用类型使用哑变量处理;P为Probability 概率;Robust_SE为标准差健壮度;Robust_t为T统计量健壮度;Robust_Pr为概率健壮度;StdCoef为回归系数的标准差;VIF为方差膨胀因子。

表 5 地理加权回归模型拟合参数
Table 5 Fitting parameters of geographically weighted regression model

	平滑程度	残差平方和	标准化剩余平方和	AICc	R ²	校正R ²
GWR	3343.22	123.74	0.56	774	0.64	0.59

表 6 不同插值处理的 SOM 半方差函数模型与参数

Table 6 Semi-variogram function model for SOM and its corresponding parameters with different methods

方法	理论模型	预测系数 R^2	变程(A_0) /m	块金值(C_0)	基台值(C_0+C)	结构方差(C)	块金值/基台值(C_0/C_0+C)/%
OK	指数模型	0.692	1050	0.0155	0.139	0.1235	11.2
RK	指数模型	0.192	480	0.0565	0.371	0.3145	15.2
GWRK	指数模型	0.188	450	0.0525	0.320	0.2675	16.4
MC_OK	指数模型	0.186	450	0.0556	0.320	0.2644	17.4
MM_OK	指数模型	0.187	450	0.0522	0.320	0.2678	16.3

最后与直接用实测值数据插值比较。

利用 GS+9.0 进行地统计分析,发现拟合的理论变异函数模型均为指数模型。其块金值/基台值均小于 25%(表 6),说明均存在一定的空间自相关性, SOM 含量主要受结构性因素影响,可进行相关插值分析。经土地利用、土壤类型、地形因子等辅助变量回归后 GWRK 模型的 C_0 和 C_0/C_0+C 等参数高于 OK 的结果,说明样本空间变异是由土地利用方式、土壤类型、地形因子等随机因素造成的。

利用 5 种模型对研究区的 SOM 插值发现:高值区分布在西北部、西部和东部等石灰土分布的岩溶区和水田,低值区位于研究区北部红水河沿岸的冲

积土地带(图 3)。RK、GWRK、MM_OK 和 MC_OK 表现出一定相似性,结合土地利用类型来看,均能较好从整体和局部细节上研究区内 SOM 含量空间分布;大体上 MM_OK 的 SOM 空间分布较为连续,而其他的几种分布比较破碎,反映出不同土地利用对 SOM 的特征以及信息的丰富程度较好;MC_OK 低值区相较于其他方法少。

3.4 模型精度评价

各模型内部验证的结果显示:OK、RK、GWRK、MM_OK 和 MC_OK 的 MSE 均接近于 0, RMSS 均接近于 1(表 7),而 GWRK 的 RMS 最小, ASE 与

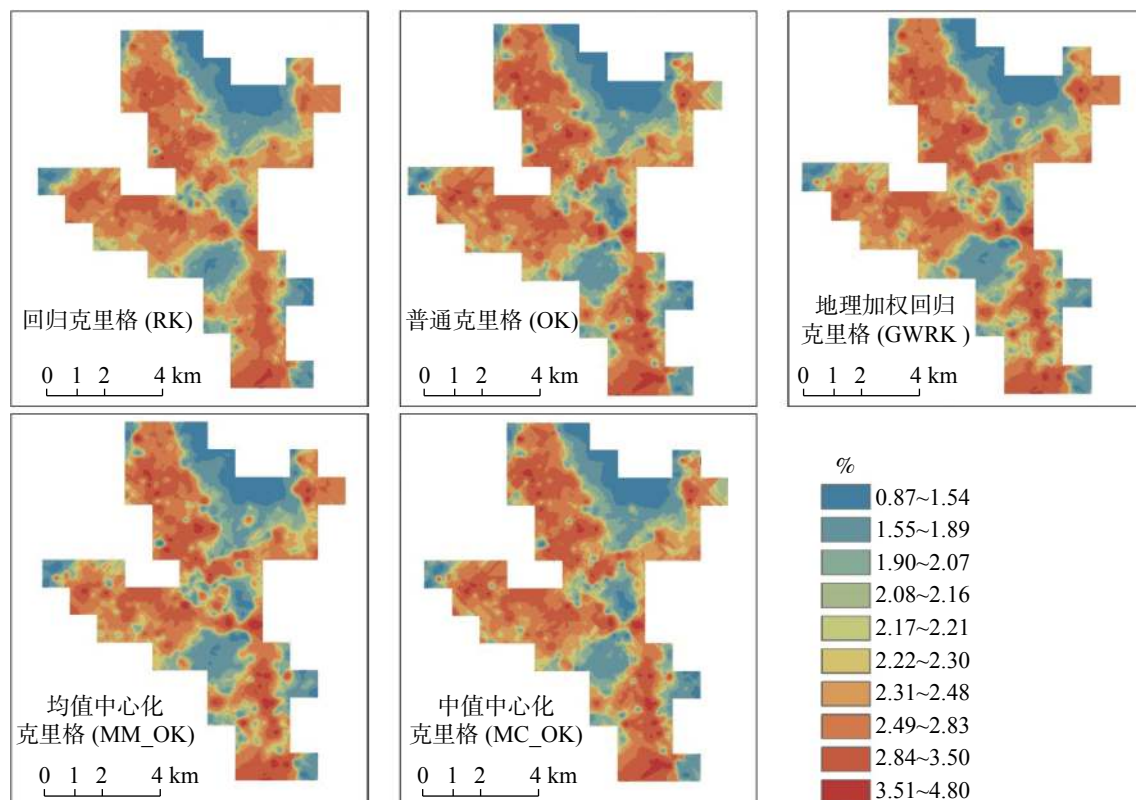


图 3 SOM 含量预测分布图

Fig. 3 Spatial distribution of SOM

表 7 不同模型预测精度评价
Table 7 Precision assessment of SOM content prediction with different methods

方法	内部验证				外部验证				
	MSE	RMSS	RMS	ASE	MAE	RMSE	AC/%	r	R ²
OK	-0.0167	1.0028	0.8166	0.8536	0.6822	0.8724	30.8	0.23	0.14
RK	0.0087	0.9836	0.5983	0.5983	0.4626	0.5844	71.9	0.71	0.48
GWRK	-0.0218	0.9797	0.5464	0.5546	0.3912	0.4960	79.4	0.79	0.63
MC_OK	0.0075	1.0764	0.5605	0.5211	0.4202	0.5215	89.3	0.77	0.58
MM_OK	0.0071	0.9890	0.5637	0.5703	0.4632	0.5775	87.0	0.72	0.49

RMS 接近, 说明 GWRK 模型拟合更好。外部验证的结果显示: GWRK 的 MAE 和 RMSE 最小, 分别指示了 GWRK 模型的精度最高和稳定性最好。GWRK 的预测集与验证集的相关系数 r 最大, 预测准确度 (AC) 由大到小依次为 MC_OK>MM_OK>GWRK>RK>OK。综上所述, 结合内部验证参数和外部验证得出 GWRK 的模型稳定性和精度较好, MC_OK 模型准确度最好。

4 讨论

马山县北部岩溶区不同土地利用方式的 SOM 存在较大差异, 表明 SOM 受地形、土壤结构性质以及土地利用方式等共同影响。SOM 具有强烈的空间自相关性, 变异范围主要在 450~480 m 之间, 超过这个范围变异影响较小。5 种克里格插值方法均能很好反映马山县北部 SOM 含量的空间分布规律。

根据 5 种方法插值图 (图 3) 可知, SOM 中高值均集中研究区的西北部、西部等地区, 南部也有零星高值出现; 结合土地利用类型和土壤类型分析可知, 中高值的土地利用方式主要为水田、林地, 土壤类型为水稻土和石灰土; SOM 含量受土壤氧化还原条件影响较大, 水田的还原条件有利于 SOM 的积累; 林地分布区受人干扰较小, 枯枝落叶、动植物残体及分解产物在表层积累, 所以该地区 SOM 也较高^[29], 钟聪等^[1] 也发现广西地区林地和水田的 SOM 高于旱地。而研究区 SOM 低值主要分布在中部、东北部、北部, 结合土地利用方式, 该区域主要为旱地, 旱地受人干扰频繁、作物轮作周期快、土壤通气状况好, SOM 分解快^[30], 累积相对少, 所以 SOM 值会相对较低。土壤类型也影响到 SOM 含量, 石灰土的

SOM 高于红壤^[1]。同时本研究发现北部靠近红水河, 成土母质为冲积土, 区域的 SOM 含量低, 不易积累, 所以会呈现出较低值, 这与前人研究结果基本一致^[13]。

不同土地利用类型对 SOM 的回归系数显示: 水田为 0.91, 而旱地 0.56, 林地为参照, 说明水田环境对有机质贡献程度较高^[4]; 回归模型分析比较发现, OLS 模型带宽 AICc 为 787 m, 校正 R² 的值为 0.56, GWR 模型中 AICc 值最佳带宽为 774 m, 校正 R² 为 0.59, 通常认为模型中带宽 AICc 小的模型为拟合度较优, 可见 GWR 模型比 OLS 模型拟合优度更高、更显著, 且 SOM 单因素方差分析 $P < 0.05$, 研究区 SOM 分布不满足方差齐性^[29]。

模型预测图中可知, 利用 GWR、OLS 拟合的残差插值的方法, 能有效降低空间变异和异常值的影响, 提升模型的预测精度, 能够有效反映出岩溶地区 SOM 空间变异的复杂性。5 种方法中 MC_OK 精度最高, 其次为 MM_OK、GWRK、RK、OK; RK、GWRK 两种模型精确度分别相对 OK 提高 2.3 倍、2.6 倍。本研究中结合辅助变量后残差中值化克里格模型的预测精度要好于均值中心化克里格, 残差数据的正态分布并不完美对称, 中值更能代表该区域 SOM 的平均水平, 这与吴子豪等^[31] 研究土壤有机碳空间异质性的结果基本一致; 张欢等^[4] 在研究青藏高原与黄土高原交界的湟水流域复杂地形区时也发现地理加权回归克里格和中值修正克里格相对于未修正克里格的准确度提升了 2% 和 4%, 本文中均值和中值中心化修正克里格相对于未修正的普通克里格的准确度提升接近 57% 和 59%, 准确度远远高于未经修正的克里格, 原因可能是研究区为岩溶地貌, 土地利用的方式和土壤类型差异相对较大, 土壤有机质有较强的空间变异性^[32]。

5 结 论

本文构建广西马山县北部岩溶地区土壤 SOM 含量的空间预测模型,分析出表层土壤有机质含量的空间分布特征,得到以下主要结论:

(1)结合土地利用方式、土壤类型和地形因子数据等辅助变量的 RK、GWRK、MC_OK 和 MM_OK4 种预测模型的稳定性和精度均远远高于 OK 模型,其中 GWRK 模型稳定性最好,MC_OK 模型预测精度最高;

(2)SOM 含量和水田呈极显著正相关,与旱地呈极显著负相关,氧化还原条件的变化影响了耕地的 SOM 含量;

(3)研究区北部岩溶地区土壤 SOM 含量在空间分布上存在异质性,土壤 SOM 含量的空间分布特征受土地利用和成土母质的影响,结合辅助变量因子预测模型能有效消除空间变异因素的影响,克服岩溶区 SOM 含量的空间非稳定性,从而提高 SOM 含量模型的稳定性和精度,服务于农业生产、环境治理和土壤碳库计量等工作。

参考文献

- [1] 钟聪,李小洁,何园燕,邱微文,李杰,张新英,胡宝清. 广西土壤有机质空间变异特征及其影响因素研究[J]. 地理科学, 2020, 40(3): 147-154.
ZHONG Cong, LI Xiaojie, HE Yuanyan, QIU Weiwen, LI Jie, ZHANG Xinying, HU Baoqing. Spatial variation of soil organic matter and its influencing factors in Guangxi, China[J]. *Scientia Geographica Sinica*, 2020, 40(3): 147-154.
- [2] 朱梓弘,杨程,谢银财,王群,朱同彬. 重度石漠化区不同土地利用方式下土壤养分特征[J]. 中国岩溶, 2018, 37(6): 842-849.
ZHU Zihong, YANG Cheng, XIE Yincai, WANG Qun, ZHU Tongbin. Characteristics of soil nutrient in karst rocky regions with heavy desertification under different land-use patterns[J]. *Carsologica Sinica*, 2018, 37(6): 842-849.
- [3] Campbell J B. Spatial Variation of Sand Content and pH Within Single Contiguous Delineations of Two Soil Mapping Units I[J]. *Soil Science Society of America Journal*, 1978, 42(3): 460-464.
- [4] 张欢,高小红. 复杂地形区土壤有机质空间变异性分析及制图[J]. 水土保持研究, 2020, 27(5): 93-100.
ZHANG Huan, GAO Xiaohong. Analysis of spatial variability and mapping of soil organic matter contents in complex terrain areas[J]. *Research of Soil and Water Conservation*, 2020, 27(5): 93-100.
- [5] 鲍丽然,周皎,李瑜,贾中民. 渝西北土壤有机质空间变异及影响因素分析[J]. 西南农业学报, 2017, 30(11): 2541-2547.
BAO Liran, ZHOU Jiao, LI Yu, JIA Zhongmin. Spatial variability of soil organic matter and its influence factors of hilly area in northwest Chongqing southwest China[J]. *Journal of Agricultural Sciences*, 2017, 30(11): 2541-2547.
- [6] Wang Y, Fu B, Lü Y, Song C, Luan Y. Local-scale spatial variability of soil organic carbon and its stock in the hilly area of the Loess Plateau, China[J]. *Quaternary Research*, 2010, 73(1): 70-76.
- [7] Zhao B, Li Z, Li P, Xu G, Gao H, Cheng Y, Chang E, Yuan S, Zhang Y, Feng Z. Spatial distribution of soil organic carbon and its influencing factors under the condition of ecological construction in a hilly-gully watershed of the Loess Plateau, China[J]. *Geoderma*, 2017, 296: 10-17.
- [8] 陆访仪,赵永存,黄标,汪景宽. 海伦市耕层土壤有机质含量空间预测方法研究[J]. 土壤通报, 2012, 43(3): 662-667.
LU Fangyi, ZHAO Yongcun, HUANG Biao, WANG Jingkuan. Comparison of predicting methods for mapping the spatial distribution of topsoil organic matter content in cropland of Hailun[J]. *Chinese Journal of Soil Science*, 2012, 43(3): 662-667.
- [9] 李启权,王昌全,岳天祥,李冰,张新,高雪松,张毅,袁大刚. 基于定性和定量辅助变量的土壤有机质空间分布预测:以四川三台县为例[J]. 地理科学进展, 2014, 33(2): 259-269.
LI Qiquan, WANG Changquan, YUE Tianxiang, LI Bing, ZHANG Xin, GAO Xuesong, ZHANG Yi, YUAN Dagang. Prediction of distribution of soil organic matter based on qualitative and quantitative auxiliary variables: A case study in Santai County in Sichuan Province[J]. *Progress in Geography*, 2014, 33(2): 259-269.
- [10] 吴才武,夏建新,段峥嵘. 土壤有机质预测性制图方法研究进展[J]. 土壤通报, 2015, 46(1): 239-247.
WU Caiwu, XIA Jianxin, DUAN Zhengrong. Technologies of predictive mapping for soil organic matter[J]. *Chinese Journal of Soil Science*, 2015, 46(1): 239-247.
- [11] Xu Y, Smith S E, Grunwald S, Abd-Elrahman A, Wani S P. Estimating soil total nitrogen in smallholder farm settings using remote sensing spectral indices and regression kriging[J]. *Catena*, 2017, 163: 111-122.
- [12] Zhao M S, Rossiter D G, Li D C, Zhao Y G, Liu F, Zhang G L. Mapping soil organic matter in low-relief areas based on land surface diurnal temperature difference and a vegetation index[J]. *Ecological Indicators*, 2014, 39: 120-133.
- [13] 李颖,刘秀明,周德全. 中国南方喀斯特地区SOC空间异质性及其对碳储量估算的指示意义[J]. 江苏农业科学, 2019, 47(8): 264-272.
LI Ying, LIU Xiuming, ZHOU Dequan. SOC spatial heterogeneity and its implications for carbon storage estimation in karst area of southern China[J]. *Jiangsu Agricultural Sciences*, 2019, 47(8): 264-272.
- [14] 谷佳慧,杨奇勇,蒋忠诚,罗为群,曾红春,覃星铭,蓝芙宁. 广

- 南县幅岩溶区与非岩溶区土壤碳氮磷生态化学计量比空间变异分析[J]. 中国岩溶, 2018, 37(5): 761-769.
- GU Jiahui, YANG Qiyong, JIANG Zhongcheng, LUO Weiqun, ZENG Hongchun, QIN Xingming, LAN Funing. Spatial variation analysis of soil carbon, nitrogen and phosphorus eco-stoichiometric ratios in karst and non-karst areas of Guangan county, Yunnan, China[J]. *Carsologica Sinica*, 2018, 37(5): 761-769.
- [15] 张双双, 靳振江, 贾远航, 李强. 岩溶区与非岩溶区3种土地利用方式下土壤细菌群落结构比较[J]. 中国岩溶, 2019, 38(2): 164-172.
- ZHANG Shuangshuang, JIN Zhenjiang, JIA Yuanhang, LI Qiang. Comparison of soil bacterial community structures from three soil land-use between karst and non-karst areas under three kinds of land use[J]. *Carsologica Sinica*, 2019, 38(2): 164-172.
- [16] Chen X B, Zheng H, Zhang W, He X Y, Li L, Wu J S, Huang D Y, Su Y R. Effects of land cover on soil organic carbon stock in a karst landscape with discontinuous soil distribution[J]. *Journal of Mountain Science*, 2014, 11: 774-781.
- [17] Liu T Z, Liu C Q, Lang Y C, Ding H. Dissolved organic carbon and its carbon isotope compositions in hill slope soils of the karst area of southwest China: Implications for carbon dynamics in limestone soil[J]. *Geochemical Journal*, 2014, 48(3): 277-285.
- [18] 景建生, 刘子琦, 罗鼎, 孙建. 喀斯特洼地土壤有机碳分布特征及影响因素[J]. 森林与环境学报, 2020, 40(2): 23-29.
- JING Jiansheng, LIU Ziqi, LUO Ding, SUN Jian. Distribution characteristics and influencing factors of soil organic carbon in karst depression[J]. *Journal of Forest and Environment*, 2020, 40(2): 23-29.
- [19] 吴海勇, 曾馥平, 宋同清, 彭晚霞, 黎星辉, 欧阳资文. 喀斯特峰丛洼地土壤有机碳和氮素空间变异特征[J]. 植物营养与肥料学报, 2009(5): 1029-1036.
- WU Haiyong, ZENG Fuping, SONG Tongqing, PENG Wanxia, LI Xinghui, OUYANG Ziwen. Spatial variations of soil organic carbon and nitrogen in peak-cluster depression areas of Karst Region[J]. *Plant Nutrition and Fertilizer Science*, 2009(5): 1029-1036.
- [20] Bai Y, Zhou Y. The main factors controlling spatial variability of soil organic carbon in a small karst watershed, Guizhou Province, China[J]. *Geoderma*, 2020, 357: 113938.
- [21] 吴敏, 刘淑娟, 叶莹莹, 张伟, 王克林, 陈洪松. 典型喀斯特高基岩出露坡地表层土壤有机碳空间异质性及其储量估算方法[J]. 中国生态农业学报, 2015, 23(6): 676-685.
- WU Min, LIU Shujuan, YE Yingying, ZHANG Wei, WANG Kelin, CHEN Hongsong. Spatial heterogeneity and storage assessment method of surface soil organic carbon in high bulk-rock ratio slopes of Karst Regions[J]. *Chinese Journal of Eco-Agriculture*, 2015, 23(6): 676-685.
- [22] Wang S, Huang M, Shao X, Mickler R A, Li K, Ji J. Vertical Distribution of Soil Organic Carbon in China[J]. *Environmental Management*, 2004, 33(1): S200-S209.
- [23] 张甘霖, 龚子同. 土壤调查实验室分析方法 [M]. 北京: 科学出版社, 2012.
- ZHANG Ganlin, GONG Zitong. Soil survey laboratory analytical methods [M]. Beijing, Science Press, 2012.
- [24] 李河, 麦劲壮, 肖敏, 杨学宁. 哑变量在Logistic回归模型中的应用[J]. 循证医学, 2008, 8(1): 42-45.
- LI He, MAI Jinzhuang, XIAO Min, YANG Xuening. Application of dummy variable in logistic regression models[J]. *The Journal of Evidence-Based Medicine*, 2008, 8(1): 42-45.
- [25] Shen H, Luo X-Q, Bi J-F. An Alternative Method for Internal Stability Prediction of Gravelly Soil[J]. *KSCE Journal of Civil Engineering*, 2018, 22(4): 1141-1149.
- [26] 陈慕松, 范晓晖, 吴寿华. 基于不同空间插值类型的耕地土壤有机质空间变异性分析[J]. 江西农业学报, 2018, 30(10): 55-59.
- CHEN Musong, FAN Xiaohui, WU Shouhua. Analysis of spatial variation of soil organic matter content in cultivated land based on different spatial interpolation styles[J]. *Acta Agriculturae Jiangxi*, 2018, 30(10): 55-59.
- [27] 文雯, 周宝同, 汪亚峰, 黄勇. 基于辅助环境变量的土壤有机碳空间插值: 以黄土丘陵区小流域为例[J]. 生态学报, 2013, 33(19): 6389-6397.
- WEN Wen, ZHOU Baotong, WANG Yafeng, HUANG Yong. Soil organic carbon interpolation based on auxiliary environmental covariates: a case study at small watershed scale in Loess Hilly region[J]. *Acta Ecologica Sinica*, 2013, 33(19): 6389-6397.
- [28] 万龙, 马芹, 张建军, 付艳玲, 张晓萍. 黄土高原降雨量空间插值精度比较: KRIGING与TPS法[J]. 中国水土保持科学, 2011, 9(3): 79-87.
- WAN Long, MA Qin, ZHANG Jianjun, FU Yanling, ZHANG Xiaoping. Precise comparison of spatial interpolation for precipitation using KRIGING and TPS(Thin plate smoothing spline) methods in Loess Plateau[J]. *Science of Soil and Water Conservation*, 2011, 9(3): 79-87.
- [29] 蒋文惠. 地形和土地利用对山区土壤养分空间变异的影响 [D]. 济南: 山东农业大学, 2014.
- JIANG Wenhui. Effects of land use and topographic factors on soil nutrients variability in mountain area [D]. Shandong: Shandong Agricultural University, 2014.
- [30] 文冬妮, 杨程, 杨霖, 秦兴华, 孟磊, 何秋香, 朱同彬, Müller C. 岩溶区农业种植对土壤有机氮矿化的影响[J]. 中国岩溶, 2020, 39(2): 189-195.
- WEN Dongni, YANG Cheng, YANG Lin, QIN Xinghua, MENG Lei, HE Qiuxiang, ZHU Tongbin, Müller C. Effects of agricultural cultivation on soil organic nitrogen mineralization in karst regions[J]. *Carsologica Sinica*, 2020, 39(2): 189-195.
- [31] 吴子豪, 刘艳芳, 陈奕云, 郭龙, 姜庆虎, 王少辰. 综合土地利用及空间异质性的土壤有机碳空间插值模型[J]. 应用生态学报, 2018, 29(1): 238-246.
- WU Zihao, LIU Yanfang, CHEN Yiyun, GUO Long, JIANG

- Qinghu, WANG Shaochen. Spatial interpolation model of soil organic carbon density considering land-use and spatial heterogeneity[J]. Chinese Journal of Applied Ecology, 2018, 29(1): 238-246.
- [32] 刘鹏, 蒋忠诚, 蓝芙宁, 李衍青, 于洋. 土地利用对溶丘洼地土壤容重、水分和有机质空间异质性的影响: 以南洞流域驻马哨洼地为例[J]. 中国岩溶, 2019, 38(1): 100-108.
- LIU Peng, JIANG Zhongcheng, LAN Funing, LI Yanqing, YU Yang. Effects of land use on spatial heterogeneity of soil bulk density, moisture and organic material in karst hilly depressions: An example of the Zhumashao depression of Nandong watershed[J]. Carsologica Sinica, 2019, 38(1): 100-108.

Spatial variation analysis of soil organic matter in karst area

ZHANG Chunlai¹, LU Laimou^{1,2}, YANG Hui¹, HUANG Fen¹

(1. Institute of Karst Geology, CAGS/ Key Laboratory of Karst Dynamics, MNR & GZAR, Guilin, Guangxi 541004, China; 2. Guangxi Huantou Jinyi Environmental Technology Co., Ltd., Nanning, Guangxi 530299, China)

Abstract As an important parameter for soil quality assessment, soil organic matter (SOM) can provide nutrients for crops, strengthen soil fertility, and improve soil physicochemical properties. Besides, SOM also plays an important role in restricting the activity of heavy metal elements and pesticide residues, and in regulating CO₂ in soil. Therefore, the study on spatial distribution of SOM, its influence factors and prediction by GIS and geostatistical analysis is important in agricultural practice, environmental management and the measurement of soil carbon storage.

The content of SOM is easily affected by the factors such as geological background, climate change, agricultural activities and land use change. Hence the complexity and non-stationarity of the spatial variability of SOM will make it difficult for the quantitative prediction. The high heterogeneity of SOM and insufficient sampling data in the karst area also leads to low accuracy of spatial prediction. Guangxi is one of the areas with the most widely distributed karst areas in China. Typical karst landforms are developed in the northern part of Mashan county, with a range of land use and parent materials. Different patterns of land use include paddy field, dry land, shrub land, forest land, orchard and grassland, accounting for 10.55%, 29.39%, 29.49%, 12.86%, 0.70% and 0.66% respectively. Paddy fields, consisting of paddy soil, are located in karst valleys; dry lands, consisting of red soil, brown lime soil, and alluvial soil, are distributed at the bottom of slopes and karst depressions on both sides of the valley; shrubland and forest land, mainly consisting of brown lime soil and red soil, spread over mountains. A total of 441 SOM data, with 8 high-density topsoil (0-20 cm) samples per km² on average, were obtained in the geochemical survey of land quality in the northern karst area of Mashan county, Guangxi in 2017. The data facilitates the spatial prediction of highly heterogeneous SOM in karst areas.

In this paper, qualitative variables such as land uses and soil types were converted into quantitative variables through the assignment method of dummy variables, and terrain factors were used as auxiliary variables to meet the requirement of geostatistical regression analysis. With the purpose of exploring the applicability of geostatistical SOM mapping in karst areas, this study respectively established five method models—ordinary Kriging (OK), regression Kriging (RK), Geographically Weighted Regression Kriging (GWRK), Ordinary Kriging of Median Centralization (MC_OK) and Mean Modified Ordinary Kriging (MM_OK)—to compare the prediction results, and select the most suitable prediction model for the karst area. Results show that the SOM content in the study area ranges from 0.81% to 5.03%, with an average of 2.36%, and the coefficient of variation is 37.3%. The spatial distribution is moderate spatial variation. Results of geostatistical analysis by GS+9.0 show that the fitted theoretical variogram models are all exponential ones. Both nugget value and abutment value are less than 25%, which indicates that there is a certain spatial autocorrelation, and the SOM content is mainly affected by structural factors; hence correlation interpolation analysis can be carried out.

In addition, the spatial variation of SOM in the karst area is jointly affected by land use, soil types and topographical factors. The areas with high content of SOM are located respectively in the karst areas and paddy fields

where lime soil is distributed in the northwest, west and east of the study area, and the areas with low content are located in the alluvial land along the Hongshui river in the north of the study area. The SOM content shows significantly positive correlation with paddy fields because the reducing environment of paddy fields is conducive to the preservation and accumulation of SOM. But SOM content shows significantly negative correlation with dry land due to the oxidative environment of this land type where frequent disturbance and rapid degradation of SOM will occur.

Finally, RK, GWRK, MM_OK and MC_OK can be used for SOM prediction and mapping in karst areas because of their good interpretability. The results of internal verification show that the root-mean-square of GWRK is the smallest, and the average standard error is close to root-mean-square, indicating the better fitness of GWRK model. The results of external verification demonstrate that the mean absolute error and root mean squares error of GWRK are the smallest, suggesting the highest precision and the best stability of GWRK model. Meanwhile, MC_OK shows its best accuracy. Combined with auxiliary variable factors such as land use, soil types, and topography, GWRK model can effectively eliminate the influence of spatial variation factors and overcome the spatial non-stationarity of SOM content in karst areas, thereby improving the stability and precision of the SOM prediction model, while MC_OK model can improve the prediction accuracy.

Key words soil organic matter, geographically weighted regression Kriging, spatial distribution, karst

(编辑 黄晨晖)