

基于 GF-1 遥感数据预测区域森林土壤有机质含量^①

李莹莹¹, 赵正勇^{1*}, 杨旗¹, 丁晓纲², 孙冬晓², 韦孙玮¹

(1 广西大学林学院广西森林生态与保育重点实验室, 南宁 530004; 2 广东省林业科学研究院, 广州 510520)

摘要: 为探索国产卫星 GF-1 预测土壤有机质(SOM)的能力, 本研究以广东省云浮市的罗定市为研究区, 以 GF-1 多光谱遥感影像衍生的 9 个遥感变量和 DEM 提取的 9 个地形水文变量为预测因子, 建立 2 种人工神经网络模型(A 模型: 地形水文; B 模型: 地形水文+遥感), 对 5 个土壤深度(L1: 0~20 cm, L2: 20~40 cm, L3: 40~60 cm, L4: 60~80 cm, L5: 80~100 cm) 的 SOM 进行预测。结果表明: 5 个深度的 B 模型全都比 A 模型的精度高, 尤其是 L1、L2 土层, 精度提升明显, 其 R^2 分别提高了 13% 和 10%; 而深层土壤(L3、L4、L5)的精度提升较小, 仅提高了 4%、5% 和 4%。另外, 两个评价指标 RMSE 和 ROA $\pm 10\%$ 也表现出相似的趋势。总体而言, GF-1 遥感数据显著改善了上层(0~40 cm)森林土壤人工神经网络模型的预测精度, 对下层(40~100 cm)森林土壤模型改善尺度较低, 是预测森林土壤 SOM 含量可观的新遥感数据源。

关键词: 土壤预测; 人工神经网络模型; GF-1; 遥感数据; 多层土壤

中图分类号: S159-3; S714 **文献标志码:** A

Prediction of Soil Organic Matter Content Based on Artificial Neural Network Model and GF-1 Remote Sensing Data

LI Yingying¹, ZHAO Zhengyong^{1*}, YANG Qi¹, DING Xiaogang², SUN Dongxiao², WEI Sunwei¹

(1 Guangxi Key Laboratory of Forest Ecology and Conservation, College of Forestry, Guangxi University, Nanning 530004, China; 2 Guangdong Academy of Forestry, Guangzhou 510520, China)

Abstract: To explore the capability of GF-1 satellite to predict soil organic matter (SOM), Luoding City of Yunfu City, Guangdong Province was taken as the study area, and 9 multi-spectral remote sensing variables retrieved from GF-1 and 9 terrain variables derived from DEM were used as predictors to establish two kinds of artificial neural network models (Model A: terrain; Model B: terrain & remote sensing) for predicting soil organic matter (SOM) at five soil depths (L1: 0–20 cm, L2: 20–40 cm, L3: 40–60 cm, L4: 60–80 cm, and L5: 80–100 cm). The results showed that the accuracies of SOM full-variable B model at five depths was higher than those of A model with topographic variables only. Especially for the L1 and L2 layers of soil, the accuracy was obviously improved. The R^2 of the L1 and L2 layers of SOM were increased by 13% and 10% respectively. However, the accuracies of deep soils (L3, L4, L5) were only improved by 4%, 5% and 4%, respectively, and RMSE and ROA $\pm 10\%$ also showed a similar trend. The results show that GF-1 remote sensing image can be used as a new data source to predict SOM.

Key words: Soil prediction; Artificial neural network model; GF-1; Remote sensing data; Multi-layer soil

土壤有机质(SOM)的预测方法有空间插值法、线性回归模型和回归克里格法等, 这些方法常需要大量的样点并假设空间自相关。而在多样化的环境中进行 SOM 预测时, 样点数据匮乏, SOM 与环境之间的关系更多是非线性过程^[1]。因此, 学者开始用人工智能技术来解决难题。例如, Mcbratney 等^[2]使用人工神经网络(artificial neural network, ANN)模型预测土壤

物理性质。Henderso 等^[3]使用决策树生成各种土壤性质图。这些模型大多是基于数字高程模型(DEM)衍生的地形和水文变量, 模型输入数据源较单一。随着卫星遥感技术迅速发展, 记录着更多土壤信息的遥感影像丰富了土壤制图的数据集^[4]。以往研究表明, Landsat 系列卫星的遥感影像最为广泛运用。例如, 张法升等^[5]利用 Landsat 5 影像预测了辽宁省阜新镇

①基金项目: 广西自然科学基金项目(2018GXNSFBA138035; 2018GXNSFAA050135)和广东省林业科技计划项目(2019-07)资助。

* 通讯作者(z.zhao@unb.ca)

作者简介: 李莹莹(1996—), 女, 广西南宁人, 硕士研究生, 主要从事森林土壤制图研究。E-mail: 228757496@qq.com

土壤表层的 SOM, Wang 等^[6]使用 Landsat 8 影像预测表层土壤有机碳(SOC)的含量。他们发现,在森林地区进行土壤成分预测的关键环境变量是遥感数据。但是, Landsat 5 遥感卫星早已退役, Landsat 7 则由于内部故障使遥感影像出现条带,最新的 Landsat 8 的 30 m 分辨率影像还无法满足高精度的预测模型。在小区域中, Spot、Quick Bird 等高分辨率遥感影像也有应用,但波段少、幅宽小,并且价格昂贵^[7],对大尺度土壤预测难普及。高分一号(GF-1)是中国首颗高分辨率对地观测遥感卫星,对我国的遥感技术的发展具有重要的战略意义。目前 GF-1 影像大多用于森林检测、地物识别和灾害监测等,用于估测土壤成分的研究非常少,尚处于探索阶段。

本文将构建 2 个 ANN 模型,即以粗分辨率土壤图为基础, A 模型的候选输入变量为 DEM 衍生出来的 9 个地形变量, B 模型继承了 A 模型筛选出来的预测能力最强的地形变量,又增加从 GF-1 影像衍生的 9 个遥感变量,建立罗定市 0 ~ 100 cm 5 个土层 SOM 的 ANN 反演模型,探索 GF-1 号遥感影像的 4 个波段以及衍生的 5 个植被变量在预测 SOM 上的反演能力,为今后利用 GF-1 号遥感影像预测其他土壤成分的研究提供参考。

1 材料与方法

1.1 研究区概况

研究区在广东省云浮市的罗定市(22°25'11" ~ 22°57'34"N, 111°03'08" ~ 111°52'44"E),全市面积 2 327.5 km²,其中森林面积约 1 426.9 km²,森林覆盖率达 61.3%。研究区内雨热同期,具有高温高湿的特点,有利于植物生长。其中,热带季风森林类型占主导地位,主要包括天然次生常绿阔叶林、针叶林和混交林。在当地气候、地形和植被类型的共同作用下,罗定市主要以红壤和赤红壤为主。

1.2 数据来源及处理

1.2.1 土壤样点数据 研究区的 260 个土壤样点数据由广东省林业科学研究院提供,具体采样时间在 2015 年 11—12 月(图 1)。样点数据由随机布点、专题布点和林分布点 3 种采样方案所得,充分代表了不同地形、植被和土壤类型等条件下的 SOM 特征。依据来自中科院南京土壤研究所全国土壤调查数据的 SOM 图(<http://www.soilinfo.cn/map/index.aspx>),将 0 ~ 100 cm 土壤剖面分为 5 层,即 L1: 0 ~ 20 cm, L2: 20 ~ 40 cm, L3: 40 ~ 60 cm, L4: 60 ~ 80 cm, L5: 80 ~ 100 cm,采用重铬酸氧化-外加热法测定 SOM 含量。

1.2.2 粗分辨率的 SOM 图 在 ArcGIS 中将中科院南京土壤研究所 1 : 100 万比例尺的 SOM 矢量图转化成 1 000 m 分辨率栅格图,即本研究用的粗分辨率 SOM(coarse-resolution soil organic matter, CSOM)图,再采用传统的反距离加权插值法重采样至 10 m 分辨率。本文以 CSOM 代表区域内 SOM 的平均值,空间上的 SOM 变化主要由地形水文和植被环境影响。

1.2.3 DEM 衍生的 9 个地形水文变量 本研究的 DEM 数据来自广东省林业科学研究院从 Cartosat-1 (IRS-P5)获得的立体图像,采用反距离加权插值法重采样至 10 m 分辨率。根据 Zhao 等^[8-9]对土壤性质的研究,坡度(Slope)、坡向(Aspect)、地形位置指数(topographical position index, TPI)、潜在太阳辐射(potential solar radiation, PSR)、潜在地下水位深(depth to water, DTW)、泥沙输移比(sediment delivery ratio, SDR)、水流长度(flow length, FL)、水流流向(flow direction, FD)、土壤地形因素(soil terrain factor, STF)是预测模型较为优越的变量。9 个地形水文变量主要参照 Meng 等^[10]和 Zhao 等^[8]的提取方法,利用 ArcGIS 中的空间分析工具和基于 ArcGIS 的脚本语言 Avenue 开发的森林水文工具从 DEM 获取。

1.2.4 GF-1 号衍生的 9 个遥感变量 综合考虑影像的能见度和采样时间,从地理空间数据云网站(<http://www.gscloud.cn>)选择覆盖研究区的 7 景 8 m 分辨率的 GF-1 多光谱遥感影像,波段范围为 450 ~ 890 nm,成像时间为 2014 年 10 月 8 日 3 景,2014 年 1 月 23 日 3 景,2013 年 11 月 15 日 1 景。虽然,选取影像的成像时间不一致,某些影像也不在生长季内,但是罗定市属于亚热带季风气候,森林植被终年常绿,季节变化不明显,不同季节的影像相差不大。首先,利用 ENVI5.3 软件对遥感影像进行预处理,主要包括在“Radiometric Calibration”模块中完成辐射定标;再对影像进行 FLASS 大气校正,以消除图像中由大气散射引起的辐射误差;再选择“RCP Orthorectification Workflow”模块进行正射校正;最后进行镶嵌、裁剪和重采样(反距离加权插值法)至 10 m 分辨率。模型采用的遥感变量有蓝波段(B)、绿波段(G)、红波段(R)、近红外波段(NIR),旨在探索 GF-1 影像 4 波段在预测 SOM 上的能力。另外,利用 ENVI5.3 的 Band math 模块进行 4 种植被指数 NDVI、DVI、RVI、RDVI 的计算,用于反映植物生长状况。在遥感科学领域,这些植被指数对土壤背景的变化非常敏感,能较好地反映植被覆盖度和生长

状况的差异,是植被生长的良好指标。

为了提高森林分类的准确性,在易康(eCognition)软件中采用分层分类方法对处理好的 8 m 分辨率的影像进行森林分类。分类样本数据从广东省林业科学研究院提供的森林小班矢量图的各个森林类型较大斑块的几何中心选取 900 个尽可能均匀分布的森林类型样点。将 900 个数据样点随机分成两份,600 个用于分类训练,300 个用于精度检验。根据 eCognition 的“Feature Space Optimization”选择可分离性最优的特征,先区分植被和非植被,在植被基础上再分林地和耕地,最后在林地上区分针叶林、阔叶林和混交林。分类精度评价使用 eCognition 中的基于样本的误差矩阵(Error Mateix based on Sample),针叶林、阔叶林和针叶林的总体精度为 0.81, KIA 为 0.77。分类结果见图 1。由于森林分类技术的局限,无法对研究区内面积很小的灌木、竹林和其他类植被实现自动分类,则依据小班图进行人工分类^[11],最后采用反距离加权插值法重采样至 10 m 分辨率。

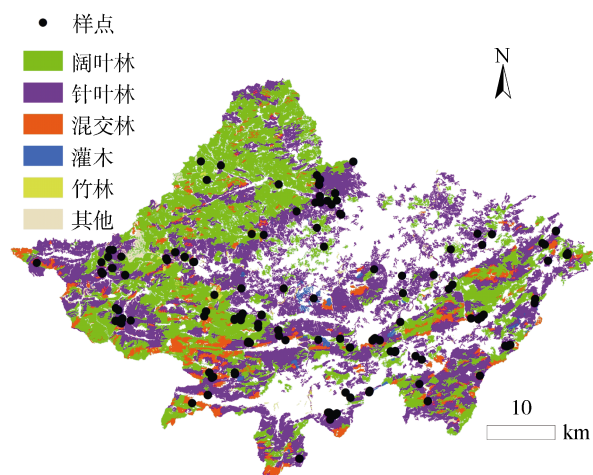


图 1 研究区采样点和森林类型图

Fig. 1 Sampling points and forest types in the study area

1.3 研究方法

1.3.1 ANN 模型结构 本研究采用的 ANN 模型是基于 Zhao 等^[9]构建的后向反馈人工神经网络模型。ANN 模型具有 3 层结构:输入层、输出层和隐含层。输入层为模型预测自变量,本研究的输入自变量分为两种类型:必选变量和候选变量。必选变量是模型必定会纳入的数据,候选变量则是从一个至多个逐渐增加的组合输入模型。本研究构建的 2 种 ANN 模型主要区别在于模型的输入变量上。A 模型(地形)输入层的必选变量为 CSOM,候选变量为 DEM 衍生的 9 个地形水文变量,候选参数组合为 C_9^1 、 C_9^2 、 C_9^4 、 C_9^5 、

C_9^6 、 C_9^7 、 C_9^8 、 C_9^9 共 511 种,最优 ANN 模型将从 511 个模型中筛选。B 模型(地形+遥感)输入层的必选变量除了 CSOM 之外,还包括了 A 模型筛选出的最优地形水文变量,候选变量则是 GF-1 获取的 9 个遥感数据,同样以组合方式输入模型。模型输出层为预测因变量,即各层土壤 SOM 的实测数据。输入层和输出层之间的为隐含层。隐含层每个节点都与输入层和输出层连接,输入层与隐含层之间的所有链接构成输入权矩阵,隐含层与输出层之间的所有链接构成输出权矩阵,使用 Sigmoid 激励函数实现非线性的映射。另外,模型采用 Levenberg-Marquardt 算法进行训练,确定隐含层节点数量,当均方误差(MSE)小于 0.01 时,停止训练。隐含层节点最终固定在 40。

1.3.2 评价指标 为了比较模型的性能,ANN 模型采用 10-fold Cross Validation(10 重交叉验证)评估模型精度。具体采用 3 个模型精度评价指标,包括均方根误差(RMSE)、决定系数(R^2)、相对整体精度($ROA \pm 10\%$)。具体公式如下:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - X_i)^2}{n}} \quad (1)$$

$$R^2 = \frac{\sum_{i=1}^n (X_i - Y_i)^2}{\sum_{i=1}^n (X_i - \bar{Y})^2} \quad (2)$$

$$ROA = \frac{\sum_{i=1}^n \left\{ \begin{array}{l} 1 \text{ if } \text{abs}\left(\frac{Y_i - X_i}{X_i}\right) \times 100 < T \\ 0 \text{ else} \end{array} \right\}}{n} \times 100 \quad (3)$$

式中: Y_i 为预测值; X_i 为模型的测量值; n 为总样本数; $\bar{Y}$ 为模型预测值的平均值; T 是精度阈值,本研究中 T 为 10。

2 结果与分析

2.1 研究区土壤有机质描述性统计

研究区采样点 SOM 含量的统计结果如表 1 所示。SOM 的变化范围在 0.05 ~ 103.07 g/kg,其中 L1 土层的变化最大,在 0.07 ~ 103.70 g/kg。SOM 的平均含量和标准差随土壤深度增加而下降,深层土壤(L4、L5)的 SOM 约是表层(L1)的 1/2。此外,每一层 SOM 的变异系数都比较高,都在 50% 以上,研究区 SOM 的空间分布属于中等变异。

2.2 研究区土壤有机质与预测变量的关系

SOM 与森林类型(Forest)等 19 个变量的 Pearson

表 1 各土层 SOM 的描述性统计(g/kg)
Table 1 Statistical information of SOM of each soil layer

土层	最小值	最大值	平均值	标准差	变异系数(%)
L1	0.07	103.07	23.19	12.33	53
L2	2.86	55.19	16.41	8.42	51
L3	0.44	47.51	13.42	7.09	53
L4	0.88	55.02	11.98	7.47	62
L5	0.05	49.81	10.16	6.45	63

相关关系如表 2 所示。每层土壤的 SOM 均与 CSOM 显著正相关。在遥感变量中, 每层土壤的 SOM 与多个遥感变量之间的相关性都显著, 特别是与 NDVI、RDVI。另外, 每层土壤的 SOM 仅与少数地形变量的相关性显著。统计表明, 本研究区的 SOM 与地形变量的线性关系较弱, 与遥感变量线性关系较强。因此, 本研究采用经典的机器学习算法——ANN 模型来解决非线性的问题。

2.3 研究区土壤有机质预测模型比较与评估

2.3.1 模型输入变量最优组合筛选 最优模型应在继续添加变量后精度不会显著提高, 则该输入组合才是模型最优的变量组合。因此, 最佳模型应具有较高的 R^2 和 $ROA \pm 10\%$, 较低的 RMSE。本研究 2 种模

型的预测结果如表 3 所示。在 L1 土层中, A 模型(地形)逐渐添加地形水文变量后显著提高了模型的性能, $ROA \pm 10\%$ 的值从 35% 逐渐增加到 47%、50%、56%、63%、72%, 但进一步添加变量, $ROA \pm 10\%$ 不再继续增加, 甚至在增加到 8 个变量时 $ROA \pm 10\%$ 降到了 63%。模型的 RMSE 和 R^2 也都反映出相同的先改善后降低的趋势。因此, A 模型预测 L1 土层 SOM 的最佳输入变量为 6 个, 基于此选择方案, 预测 L2、L3、L4、L5 土层 SOM 的最佳输入变量为 6、7、5、6 个。而对于 B 模型, 预测 5 个土层 SOM 的最佳输入变量个数分别为 6、4、6、6、7 个。

2.3.2 模型预测精度比较 从表 3 可知, B 模型与 A 模型相比普遍具有最低的 RMSE, 以及最高的 R^2 和 $ROA \pm 10\%$, 因此, B 模型的预测结果较佳。并且, 可以观察到 B 模型对表层土壤(L1、L2)比对深层土壤(L3、L4、L5)的预测精度提高尺度更大。相对于 A 模型, B 模型对 L1、L2 土层预测的 R^2 提高了 13% 和 10%, 而 L3、L4、L5 土层仅提高了 4%、5%、4%。同样地, RMSE 和 $ROA \pm 10\%$ 也表现出相似的趋势。这表明遥感影像在改善研究区 SOM 的预测能力方面是有一定的局限性, 对于深层土壤改善不够显著。

表 2 各层土壤 SOM 与输入变量的相关性
Table 2 Correlation between SOM and input variables of each soil layer

变量	SOM_L1	SOM_L2	SOM_L3	SOM_L4	SOM_L5
CSOM	0.144*	0.236**	0.170**	0.177**	0.155*
Forest	0.162*	0.156*	0.006	0.108	0.130*
NDVI	0.146*	0.215**	0.127*	0.193**	-0.021
RVI	0.131*	0.212**	0.120	0.197**	0.013
DVI	0.008	0.087	-0.014	0.120	-0.010
RDVI	0.349**	0.365**	0.605**	0.323**	0.251**
B	-0.117	-0.206**	-0.187**	-0.142*	-0.064
G	-0.137*	-0.156*	-0.155*	-0.089	-0.016
R	-0.152*	-0.182**	-0.150*	-0.127*	0.006
NIR	-0.035	0.029	-0.053	0.074	-0.008
TPT	0.083	0.039	0.025	-0.004	0.010
Slope	0.178**	0.283**	0.218**	0.215**	0.241**
Aspect	0.015	-0.026	0.038	0.161*	-0.009
STF	-0.126	-0.213**	-0.104	-0.110	-0.213**
SDR	0.134*	0.065	0.105	0.092	0.074
DTW	0.138*	0.149*	0.057	0.097	0.081
FD	-0.090	0.014	-0.069	-0.073	-0.018
FL	-0.006	-0.006	0.120	-0.061	-0.058
PSR	-0.174**	-0.086	-0.152*	-0.184**	-0.064

注: **表示相关性在 $P < 0.01$ 水平显著; *表示相关性在 $P < 0.05$ 水平显著。

表 3 SOM 的 2 种 ANN 模型最优变量组合及预测精度
Table 3 Optimal variable combination and prediction accuracies of two ANN models of SOM

模型	深度	变量数	RMSE	R ²	ROA±10%(%)	最优模型变量组合
A	L1	1	165.48	0.21	35	TPI
		2	123.71	0.35	47	TPI、Slope
		3	97.78	0.47	50	TPI、Slope、STF
		4	77.15	0.59	56	TPI、Slope、STF、SDR
		5	69.03	0.64	63	TPI、Slope、STF、SDR、DTW
		6	46.66	0.75	72	TPI、Slope、STF、SDR、DTW、FL
		7	57.52	0.72	64	TPI、Slope、STF、SDR、DTW、FL、PSR
		8	56.76	0.71	63	TPI、Slope、STF、SDR、DTW、FL、PSR、Aspect
		9	55.74	0.71	68	TPI、Slope、STF、SDR、DTW、FL、PSR、Aspect、FD
	L2	6	16.11	0.76	73	TPI、Slope、Aspect、SDR、DTW、PSR
	L3	7	11.02	0.79	76	TPI、Slope、Aspect、STF、SDR、FD、PSR
	L4	5	15.10	0.74	71	TPI、Slope、Aspect、SDR、FL
	L5	6	12.05	0.76	76	TPI、Slope、Aspect、STF、DTW、FD
B	L1	6	23.03	0.88	87	A1+ Forest、NDVI、RDVI、G、B、R
	L2	4	10.98	0.86	86	A2+ Forest、NDVI、RVI、DVI
	L3	6	9.76	0.83	82	A3+Forest、RDVI、G、B、RVI
	L4	6	11.39	0.79	80	A4+Forest、NDVI、RDVI、DVI、G、NIR
	L5	7	9.68	0.80	82	A5+Forest、NDVI、RDVI、RVI、DVI、R、NIR

注：表中 A1 代表 TPI、Slope、STF、SDR、DTW、FL；A2 代表 TPI、Slope、Aspect、SDR、DTW、PSR；A3 代表 TPI、Slope、Aspect、STF、SDR、FD、PSR；A4 代表 TPI、Slope、Aspect、SDR、FL；A5 代表 TPI、Slope、Aspect、STF、DTW、FD。

2.4 研究区土壤有机质空间格局

采用 B 模型筛选出的最优 ANN 模型，以 10 m×10 m 的分辨率绘制罗定市森林土壤 SOM 的空间分布(图 2)。SOM 的含量随着土层深度的加深而减少，其中，L1 土层 SOM 含量最高，其空间分布含量普遍高

于 30 g/kg；L2 土层 SOM 含量主要在 20~30 g/kg；L3 土层 SOM 含量则大多处在 10~20 g/kg 和 20~30 g/kg 等级；L4 和 L5 土层 SOM 含量减少得更明显，主要在 10~20 g/kg 等级范围内。从水平结构的整体来看，每层土壤较高的 SOM 主要出现在阔叶林地区。

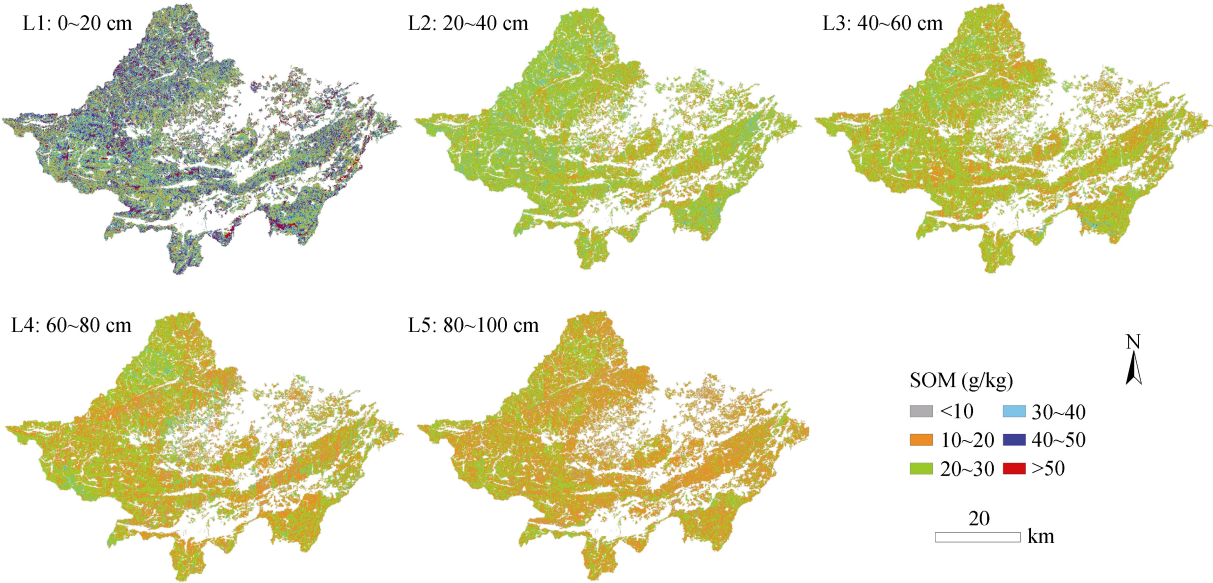


图 2 研究区各土层的 SOM 空间分布
Fig. 2 Spatial distribution of SOM in 5 soil layers in study area

3 讨论

本研究主要探索 GF-1 的 4 个多光谱波段与衍生的 5 个植被变量对传统的地形水文 ANN 模型的影响,结果表明,GF-1 衍生数据提高了地形水文土壤模型的预测精度。这一研究结果与 Wang 等^[12]和 Zhou 等^[13]的报告一致,他们认为全变量模型(地形水文+遥感)比地形水文模型预测更好。在改善地形水文模型的精度上,遥感数据对上层(0~40 cm)土壤的预测提升幅度较大,对深层(40~100 cm)土壤提升幅度较小。这与 Kempen 等^[14]和 Minasny 等^[15]的结果一致,他们都报道了土壤预测模型的性能随深度而下降。这可能是由于辅助环境变量的映射能力随着土壤深度加深而降低,所使用的大部分辅助数据不能有效地捕获这些土层(40~100 cm)的 SOM 变化,其不确定性随着土壤深度的增加而增加^[9]。

通过最佳模型(B 模型)的变量组合来看,所有遥感变量中森林类型(Forest)和 NDVI 对 SOM 含量表现出较强的预测能力。这与 Winowiecki 等^[16]的研究一致,其发现有机碳的空间分布格局与植被类型及密度密切相关。在森林生态系统中,SOM 的主要来源为植被凋落物,植被越密,凋落物越多。植被凋落物的特征也是影响 SOM 积累的重要原因,针叶类凋落叶分解速率较慢,不利于 SOM 的积累,而阔叶树凋落叶容易分解,促进 SOM 的积累^[17]。一些落叶阔叶林还发生季节性落叶,增加了 SOM 的来源。植被指数广泛用于评估植被覆盖、生物量、作物产量和植物健康等植物状况。最优模型变量组合显示,NDVI、RDVI、DVI 和 RVI 均有出现在不同土层的最优模型中。这些指标对植被覆盖、植被色素含量和叶片水分的变化最为敏感,已成功用于评估植被覆盖状况、预测土壤养分和土壤盐分^[18-19]。考虑到土壤性质与植被覆盖状况的密切关系,植被指数可以反映土壤性质的变化。因此,可以引入光谱植被指数作为间接指标来了解 SOM 的状况。

遥感数据是数字土壤制图研究中应用广泛的辅助变量。但是,以往研究中极少探索 GF-1 数据在土壤制图中的应用。本研究发现,GF-1 的 4 个波段作为 ANN 模型输入变量进行评估 SOM 是有效的。这与马驰^[7]的结果一致,他应用 GF-1 的 4 个波段较好地预测了 SOM 的空间分布。另外,以往研究大多数都是采用了 Moidis、Landsat 和 Sentinel 等卫星的绿波段、红波段和近红外波段等,都未探索蓝波段与土壤性质的关系。而最优模型的变量组合显示 GF-1 的

蓝波段对 L1、L3 土层的 SOM 预测具有一定的重要性。该结论与郭燕等^[20]研究一致,她们发现 GF-1 的 4 个波段中,蓝波段与 SOM 的相关性最强。Streck 等^[21]认为,土壤电磁光谱可见光区和近红外区等的光谱反射率揭示了大量关于土壤条件的信息。事实上,土壤光谱数据受到土壤性质的影响,包括土壤 SOM、植被、土壤质地、土壤水分、土壤盐分和氧化铁含量等。因此,光谱数据是研究和测绘土壤性质的一个很好的工具。

Pouladi 等^[22]使用遥感数据预测表层土壤有机碳含量,发现遥感变量解释了 89% 的有机碳含量的空间变化,并认为当采样点足够密集时,无需在模型中引入其他环境变量。但本研究表明,地形水文变量在预测土壤性质时起到了很重要的作用,仅地形水文变量就可以解释 70% 左右的 SOM 含量分布状况。并且,遥感变量对表土层的预测效果较好,对深层土壤预测效果较差。因此,地形水文环境变量是必要的,尤其是在地势起伏大的森林地区^[23]。从本研究 A 模型筛选出来的最优变量组合来看,预测 SOM 含量的代表性地形水文变量是 TPI 和 Slope。TPI 反映的是研究点与其邻域内其他点的相对高度差^[24]。TPI 为下坡位和山谷地区,坡度较小,海拔较低,人类活动干扰多,土壤侵蚀严重,容易造成 SOM 流失。相反,在坡度大的山脊和上坡位地区,海拔较高,人类活动较少,土壤动物和微生物活跃,有利于 SOM 形成。海拔还通过影响小气候的形成,影响土壤动物和微生物的活性,进而影响 SOM 的积累。

从 SOM 的垂直分布来看,随着土壤深度的增加,SOM 含量呈逐渐下降的趋势,这与邓小军等^[25]对亚热带广西山地红壤区域的研究结果一致。主要原因是凋落物、土壤动物和微生物主要分布在土壤表层,且表层土壤通风、湿热条件优于深层土壤,使得 SOM 在表层积累较高^[26]。从 SOM 的水平结构来看,SOM 含量高的主要分布在阔叶林。这与最优模型的变量组合一致,森林类型(Forest)是 5 个土层 SOM 预测的最重要的变量。SOM 的积累主要受到高等植物凋落物的性质,以及土壤动物和微生物的影响。阔叶林的凋落物比针叶林凋落物易分解。另外,罗定市主要的针叶树种是马尾松、杉木和桉树林等人工林,受到人类林业活动的影响较多,不利于 SOM 的积累。

本研究中,加上遥感变量的 ANN 模型在预测性能上得到了提高,但是,研究中存在不确定性。首先,数据采集和实验室分析中会存在一些采样误差和实验误差。其次,由于地形和云层的影响,高海拔地区

在图像分割过程中易产生阴影,从而导致卫星图像数据中出现较大的反射率误差^[13]。另外,陆地表面特性随时间的变化和遥感数据采集的时间也会影响预测结果。最后,本研究只从 GF-1 提取了 9 个遥感数据,可能未充分挖掘更好的遥感潜在变量。

4 结论

本研究利用 2 种 ANN 模型预测了广东省云浮市罗定市森林土壤 5 个土层的 SOM 含量,结果表明,国产卫星 GF-1 号衍生的 9 个遥感变量对研究区 SOM 的预测是有效的,其对上层(L1 和 L2)森林土壤 SOM 的预测效果较好,遥感数据的加入使得基于地形水文数据的 ANN 土壤模型的精度(R^2)分别提高了 13% 和 10%,而对 40 cm 以下深层土壤的预测性能提高较小,仅提高了 4% 或 5%,需要引入更多的遥感变量,如多光谱波段和全色波段的纹理特征等辅助变量来改善深层土壤的预测。因此,在未来的森林土壤制图中应注重选择和开发与植被相关的变量,进一步探索国产遥感卫星与土壤制图技术的结合,推动其在全国范围内林业、农业等方面的进一步应用与发展。

参考文献:

- [1] McBratney A B, Odeh I O A, Bishop T F A, et al. An overview of pedometric techniques for use in soil survey[J]. *Geoderma*, 2000, 97(3/4): 293–327.
- [2] McBratney A B, Mendonça Santos M L, Minasny B. On digital soil mapping[J]. *Geoderma*, 2003, 117(1/2): 3–52.
- [3] Henderson B L, Bui E N, Moran C J, et al. Australia-wide predictions of soil properties using decision trees[J]. *Geoderma*, 2005, 124(3/4): 383–398.
- [4] 周萍, 王润生, 阎柏琨, 等. 高光谱遥感土壤有机质信息提取研究[J]. *地理科学进展*, 2008, 27(5): 27–34.
- [5] 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.
- [6] Wang S, Gao J H, Zhuang Q L, et al. Multispectral remote sensing data are effective and robust in mapping regional forest soil organic carbon stocks in a northeast forest region in China[J]. *Remote Sensing*, 2020, 12(3): 393.
- [7] 马驰. 基于 GF-1 土壤有机质含量估测的研究[J]. *西南农业学报*, 2018, 31(1): 126–130.
- [8] Zhao Z, Yang Q, Benoy G, et al. Using artificial neural network models to produce soil organic carbon content distribution maps across landscapes[J]. *Canadian Journal of Soil Science*, 2010, 90(1): 75–87.
- [9] Zhao Z Y, Yang Q, Sun D X, et al. Extended model prediction of high-resolution soil organic matter over a large area using limited number of field samples[J]. *Computers and Electronics in Agriculture*, 2020, 169: 105172.
- [10] Meng F R, Castonguay M, Ogilvie J, et al. Developing a GIS-based flow-channel and wet areas mapping framework for precision forestry planning[EB/OL]. Stellenbosch: IUFRO Precision Forestry Symposium, 2006.
- [11] 孙庆林. 基于 GF-2 的八达岭林场主要森林类型分类方法研究[D]. 北京: 北京林业大学, 2017.
- [12] Wang S, Zhuang Q L, Jin X X, et al. Predicting soil organic carbon and soil nitrogen stocks in topsoil of forest ecosystems in northeastern China using remote sensing data[J]. *Remote Sensing*, 2020, 12(7): 1115.
- [13] Zhou T, Geng Y J, Chen J, et al. Mapping soil organic carbon content using multi-source remote sensing variables in the Heihe River Basin in China[J]. *Ecological Indicators*, 2020, 114: 106288.
- [14] Kempen B, Brus D J, Stoorvogel J J. Three-dimensional mapping of soil organic matter content using soil type-specific depth functions[J]. *Geoderma*, 2011, 162(1/2): 107–123.
- [15] Minasny B, McBratney A B, Malone B P, et al. Digital mapping of soil carbon[M]//*Advances in Agronomy*. Amsterdam: Elsevier, 2013: 1–47.
- [16] Winowiecki L, Vågen T G, Huising J. Effects of land cover on ecosystem services in *Tanzania*: A spatial assessment of soil organic carbon[J]. *Geoderma*, 2016, 263: 274–283.
- [17] 王薪琪, 王传宽. 东北 5 种温带人工林表层土壤碳氮含量的分异[J]. *应用生态学报*, 2019, 30(6): 1911–1918.
- [18] Rivero R G, Grunwald S, Binford M W, et al. Integrating spectral indices into prediction models of soil phosphorus in a subtropical wetland[J]. *Remote Sensing of Environment*, 2009, 113(11): 2389–2402.
- [19] Taghizadeh-Mehrjardi R, Minasny B, Sarmadian F, et al. Digital mapping of soil salinity in Ardakan region, central Iran[J]. *Geoderma*, 2014, 213: 15–28.
- [20] 郭燕, 程永政, 王来刚, 等. 利用高光谱和 GF-1 模拟多光谱进行土壤有机质预测和制图研究[J]. *土壤通报*, 2016, 47(3): 537–542.
- [21] Streck N A, Rundquist D, Connot J. Spectral signature of selected soils[J]. *Agrometeorol*, 2015, 11(1): 184.
- [22] Pouladi N, Möller A B, Tabatabai S, et al. Mapping soil organic matter contents at field level with Cubist, Random Forest and kriging[J]. *Geoderma*, 2019, 342: 85–92.
- [23] Li X H, Ding J L, Liu J, et al. Digital mapping of soil organic carbon using sentinel series data: A case study of the ebinur lake watershed in Xinjiang[J]. *Remote Sensing*, 2021, 13(4): 769.
- [24] 李小川, 丁晓纲. 广东省云浮市森林土壤养分调查与评价[M]. 北京: 中国林业出版社, 2018.
- [25] 邓小军, 曹继钊, 宋贤冲, 等. 猫儿山自然保护区 3 种森林类型土壤养分垂直分布特征[J]. *生态科学*, 2014, 33(6): 1129–1134.
- [26] 何林. 汝城县森林土壤养分含量的垂直分布规律[J]. *湖南林业科技*, 1988, 15(1): 59–63.