

基于机载高光谱影像的农田尺度土壤有机碳密度制图

刘潜¹, 王梦迪¹, 郭龙², 王冉³, 贾中甫³, 胡献君⁴,
唐乾坤⁵, 石铁柱¹

1. 深圳大学 自然资源部大湾区地理环境监测重点实验室&广东省城市空间信息工程重点实验室&深圳市空间信息智能感知与服务重点实验室, 深圳 518060;
2. 华中农业大学 资源与环境学院, 武汉 430070;
3. 内蒙古自治区测绘地理信息中心, 呼和浩特 010020;
4. 海军工程大学, 武汉 430032;
5. 自然资源部第三航测遥感院&四川测绘地理信息局, 成都 610100

摘要: 准确监测土壤有机碳密度SOCD (Soil Organic Carbon Density) 对调控土壤碳汇、合理利用土壤资源具有重要意义。机载高光谱影像为精细化SOCD制图提供了重要数据源。由于机载高光谱在数据收集过程中易受到外部因素的影响, 光谱中存在噪声影响SOCD的估算精度。因此, 本研究旨在探究基于机载高光谱影像估算SOCD的技术流程。对原始光谱进行预处理, 包括一阶微分FD (First Derivative) 和包络线去除CR (Continuum Removal) 变换。采用遗传算法GA (Genetic Algorithm) 选择特征波段, 并结合不同回归方法, 如偏最小二乘回归PLSR (Partial Least Square Regression)、多元线性回归MLR (Multiple Linear Regression)、支持向量机SVM (Support Vector Machine) 和人工神经网络ANN (Artificial Neural Network) 估算SOCD。结果表明, 在经过GA特征波段选择后, 原始光谱、FD光谱和CR光谱预测SOCD的精度均有所提高。使用原始光谱特征波段, PLSR、MLR、SVM和ANN共4种模型预测SOCD的决定系数 R^2 分别为0.672、0.621、0.551和0.678。使用FD与CR光谱特征波段的 R^2 范围分别在0.452—0.593和0.332—0.602, 具有较大的误差。利用原始光谱的特征波段进行SOCD数字制图, 不同回归模型预测的SOCD在空间上具有较为相似的变化趋势, 与SOCD测量值较为相近, 绝对误差较大的点多出现在采样点边缘附近。

关键词: 土壤有机碳密度, 机载高光谱, 遗传算法, 数字土壤制图

中图分类号: S159.9/P2

引用格式: 刘潜, 王梦迪, 郭龙, 王冉, 贾中甫, 胡献君, 唐乾坤, 石铁柱. 2024. 基于机载高光谱影像的农田尺度土壤有机碳密度制图. 遥感学报, 28(1): 293–305

Liu Q, Wang M D, Guo L, Wang R, Jia Z F, Hu X J, Tang Q K and Shi T Z. 2024. Mapping of soil organic carbon density at farmland scale based on airborne hyperspectral images. National Remote Sensing Bulletin, 28(1): 293–305 [DOI: 10.11834/jrs.20221805]

1 引言

目前, 全球土壤有机碳SOC (Soil Organic Carbon) 储量约为15500亿吨, 约为大气碳储量的两倍, 化石燃烧年排放量的150倍 (Guo等, 2020; 罗梅等, 2020)。SOC显著影响土壤肥力和作物产量 (于雷

等, 2016), 同时, 在调节全球气候变化方面也发挥着重要作用 (Lal, 2004; 陈朝等, 2011)。在全球土壤碳库研究中, 土壤有机碳密度SOCD (Soil Organic Carbon Density) 是一种表征土壤有机碳储量的常用指标 (刘艳芳等, 2017)。准确监测SOCD对调控土壤碳汇以及合理利用土壤资源具有重要意义。

收稿日期: 2021-12-07; 预印本: 2022-06-21

基金项目: 国家自然科学基金 (编号: 41890854)

第一作者简介: 刘潜, 研究方向为土壤遥感。E-mail: 2060325016@email.szu.edu.cn

通信作者简介: 石铁柱, 研究方向为农业定量遥感。E-mail: tiezhushi@szu.edu.cn

随着遥感技术的发展, 遥感影像为精确的数字土壤制图提供了有效的数据支撑 (史舟等, 2018; 文凤平等, 2021)。Castaldi等 (2016) 利用4种卫星数据 (Landsat 8 OLI、Sentinel-2 MSI、EO-1 ALI 和 Hyperion) 和3种拟发射的卫星数据 (EnMAP、PRISMA 和 HypIRI) 预测了SOC含量, 展示了不同卫星数据在数字土壤制图中的作用。然而, 卫星遥感影像的空间分辨率较低, 难以满足小尺度农田土壤的精细化监测。相比较卫星平台, 机载平台具有更高的空间分辨率和更好的灵活性。机载遥感监测平台在小规模农田土壤制图中具有更大的优势 (Guo等, 2021)。机载平台通过搭载RGB、多光谱和高光谱等传感器, 采集地表空间连续的光谱信息。高光谱影像可以提供更为丰富的光谱信息, 以确定与土壤属性关系密切的特征波段 (Castaldi等, 2016)。因此, 机载高光谱影像是估算土壤属性的重要数据源。然而, 机载高光谱影像在采集过程中, 易受天气条件、植被覆盖等外部因素的干扰, 进而影响模型的预测精度 (郑光辉等, 2013; Nouri等, 2017)。这对机载高光谱数据的处理提出了更高的要求。

遗传算法GA (Genetic Algorithm) 常被用来选择敏感的光谱波段, 以减小数据冗余, 提高模型的预测能力, 被广泛应用于近端高光谱分析中 (邹小波和赵杰文, 2007; 褚小立等, 2001)。机器学习方法常被用于构建SOC估算模型。如Viscarra Rossel和Behrens (2010) 利用可见光近红外高光谱数据比较了支持向量机SVM (Support Vector Machine)、随机森林、人工神经网络ANN (Artificial Neural Network) 在预测SOC上的能力。有必要基于机载高光谱影像数据, 探讨GA和不同回归方法在预测SOCD中的效果。

综上所述, 本研究的主要目的是基于机载高光谱影像, 比较不同数据处理方法和回归方法预测SOCD的效果。旨在 (1) 建立一套适用于机载高光谱影像的处理流程和方法; (2) 探索机载高光谱数据在SOCD制图方面的能力; (3) 评估GA在敏感波段选择上的效果; (4) 分析使用不同回归方法预测SOCD的空间分布特征。

2 材料与方法

2.1 研究区与土壤采样数据

研究区位于美国Iowa州Iowa县的一个农场

(41°44'0"N, 91°56'30"W), 面积约为2 km², 其中85%以上的土地为农田, 其余土地覆盖为植被、水体以及建筑。该研究区地形较为平坦, 属于温带大陆性气候, 夏热冬凉, 年降水量通常能达到880 mm。受益于密西西比河, 该农场的土壤富饶, 适合耕作, 主要作物类型为玉米和大豆。研究区主要土壤类型为暗沃弱发育湿润淋溶土 (Mollic Hapludalfs)、典型粘淀湿润软土 (Typic Argiudolls)、氧化粘淀湿润软土 (Oxyaquic Argiudolls) 和潮湿弱发育湿润软土 (Aquic Hapludolls) (Hong等, 2020)。

2015年10月, 收集了50份0—15 cm的表层土壤样本, 采样点之间的距离不少于130 m (图1)。将土壤样本置于20℃—25℃的室温中风干14 d, 然后将土壤样本碾碎, 加入稀盐酸除去碳酸盐, 最后采用元素分析仪 (Vario EL Cube) 测量SOC。采用重量法测量土壤水分。土壤容重SBD (Soil Bulk Density) 由环刀中湿土的体积和重量计算。筛分和称量砾石得到砾石浓度。SOC_D表示为 (刘艳芳等, 2017):

$$\text{SOCD} = \sum_{i=1}^n (1 - \theta_i \%) \times \text{SBD}_i \times \text{SOC}_i \times H_i / 100 \quad (1)$$

$$\text{SBD} = \frac{(m_1 - m_2)}{V} \quad (2)$$

式中, i 表示为土壤层数 (1, 2, 3, ..., n), 在本研究中 i 设为1, θ_i 为砾石浓度, SBD_i 为土壤容重, SOC_i 为土壤有机碳含量, H_i 为土壤厚度 (15 cm), m_1 为烘干后环刀与干土的总重量, m_2 为环刀重量, V 为环刀体积。

2.2 机载高光谱影像的获取

本研究将 Headwall Micro-Hyperspec 高光谱传感器搭载在直升机上以获取高光谱影像。对耕地进行翻耕, 土壤表面无秸秆覆盖, 以确保传感器对裸土进行直接观测。获取时间为2015年11月19日上午10:00至下午14:00, 该时间段光照充足, 天空无云。该传感器主要包括A系列 (可见光—近红外) 和X系列 (短波红外) 数据。许多学者认为SOC对可见光区域的光谱反射率影响最为显著 (Ben-Dor, 2002; 史舟等, 2011; 彭杰等, 2013), 因此, 本研究使用A系列成像光谱仪采集的影像数据。该影像数据包含332个光谱波段, 光谱波长介于398—1000 nm。光谱分辨率为1.9 nm, 空间分辨率约为1 m。近红外范围内900—1000 nm

的光谱由于受水汽的严重影响, 故予以剔除。对影像进行几何校正和镶嵌处理, 使用 FLAASH (Fast Line-of-sight-Atmospheric Analysis of Spectral Hypercubes) 模型进行大气校正。对原始高光谱影像进行 Savitzky-Golay 滤波处理, 减少噪声的同时, 尽可能保留原始光谱曲线的形态特征, 本研究将

滤波窗口设为 10, 多项式次数设为 2 (Gomez 等, 2018)。一阶微分 FD (First Derivative) 和包络线去除 CR (Continuum Removal) 处理能够增强原始光谱中的细微吸收峰, 这些微弱吸收峰可能包含适合 SOCD 估算的重要光谱信息。因此, 对光谱数据进行了一阶微分 FD 和包络线去除 CR 处理。

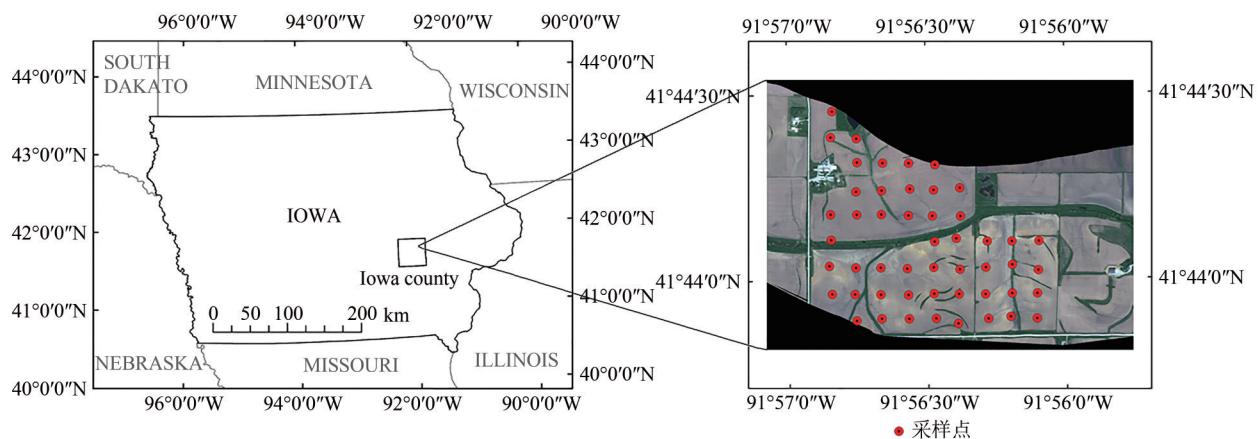


图1 研究区地理位置及采样点分布

Fig. 1 Location of study area and spatial distribution of sample points

2.3 基于 GA-PLS 算法的波段筛选

对于高光谱数据而言, 光谱变量中通常包含噪声或者干扰信号, 这可能会降低回归模型的精度。将受噪声影响严重或对响应变量不敏感的光谱变量剔除, 往往能够提高模型的精度 (詹白勺等, 2014)。高光谱变量筛选的方法主要有小波分析、连续投影法和主成分分析等。相比于这几种算法, 遗传算法提供了一种更为高效的光谱特征选取方法, 且保留了原始光谱变量的物理意义 (乔天等, 2018)。本研究采用遗传算法与偏最小二乘法 (GA-PLS) 结合的方法进行光谱变量的筛选。

GA 通过模拟生物界自然选择与进化机制, 不断地遗传迭代, 最终保留下目标函数值最优的变量组合 (Leardi 等, 1992; 蔡建楠等, 2020)。遗传算法由 5 个基本步骤组成: (1) 变量的编码; (2) 种群初始化; (3) 适应度评价; (4) 繁殖; (5) 突变。不断重复步骤 (3)–(5), 当适应度函数不再增加或达到给定的迭代次数时终止程序, 此时可筛选出最优的变量组合。为确保模型的可靠稳定, 本研究在蔡建楠等 (2020) 的研究基础上, 通过反复调试, 设置种群大小为 64, 窗口宽度为 1, 初始变量集为 10%, 突变率为 0.5%。为了防止过拟合, 采用随机交叉验证并设置迭代次数为 100。

2.4 SOCD 预测模型的建立

将经过 GA-PLS 筛选后的波段组合用于建立 SOCD 预测模型。本研究建立了偏最小二乘 PLSR (Partial Least Square Regression)、补充完整 MLR、SVM 以及 ANN 等 4 种预测模型。同时将全光谱波段的反射率数据也用于建模, 以比较 GA 模型与全光谱波段模型的预测效果。

2.4.1 PLSR

PLSR 算法是由 Wold (1966) 提出的一种多元回归方法, 其被广泛地应用于光谱分析。该算法能够对给定的自变量进行压缩, 并将其投影到几个潜在变量 LVs (Latent Variables) 中去。这些少量的 LVs 既能够很好地表达自变量, 又与因变量高度相关。PLSR 的特点是即使自变量存在严重的多重相关性也能够进行回归建模。在 PLSR 模型中, 选取 LVs 的数量对结果的输出具有重要影响。在防止过拟合的前提下, 本研究通过分析交叉验证的均方根误差 RMSECV (Root Mean Square Error of Cross Validation) 来选择最优的 LVs 数量。

2.4.2 MLR

MLR 是一种传统的回归方法, 其建立多个自变量与因变量之间的线性关系, 并通过该线性关

系去预测因变量的值。它要求自变量与因变量之间存在线性关系, 以及自变量彼此之间的相关性要低。MLR的一般方程表示为 (Xie 等, 2021):

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_ix_i$$

式中, y 为因变量, b_0 为常数, x_i 为第 i 个自变量、 b_i 为 x_i 对应的回归系数。

2.4.3 SVM

SVM是由 Vapnik (1998) 提出的一种基于核的机器学习方法, 该模型由多个支持向量和非线性模型系数组成, 它们定义了自变量的非线性映射, 可用于预测连续的因变量。其中心思想就是尽可能地将更多的训练数据映射到超平面中, 同时限制决策边界的宽度。本研究使用高斯径向基函数作为 SVM 的核函数, 需要对 epsilon、cost 和 gamma 参数进行优化。其中, epsilon 被用于控制决策边界的宽度, 并设为 1.0、0.1、0.01; gamma 是用于定义高斯函数宽度的调优参数, 设为以对数均匀分布在 10^{-6} 到 10 之间的 15 个值; cost 则为每个支持向量的影响, 设为以对数均匀分布在 10^{-3} —100 的 11 个值。SVM 通过使总损失最小化同时使间隔最大化而得到最终的优化模型。

2.4.4 ANN

ANN 由相互连接的非线性处理单元 (神经元) 组成, 这些神经元协同工作, 解决各类问题。多层感知器 MLP (Multilayer Perceptron) 是土壤科学中最常用的神经网络结构 (Besalatpour 等, 2013), 其由输入层、隐藏层和输出层组成。隐藏层中的神经元数量对最终的输出具有重要影响。权重和偏差是神经网络的两个参数, 通过调整参数优化算法和学习能力以减少下一次迭代的误差。本研究使用了基于反向传播 BP (Backpropagation) 的前馈 MLP, BP 可以连续不断地更新或训练权值, 从而减小预测误差。训练由 “learn cycles” 和 “learn rate” 两个参数控制, learn rate 决定了第一个 learn cycle 的训练持续时间, 后续每个 learn cycle 的持续时间是前一个 learn cycle 持续时间的两倍。本研究中, learn rate 设为 0.125, learn cycle 设为 20。

2.5 模型精度评价

本研究评估模型精度采用的是留一法交叉验证。SOCD 预测模型的精度通过交叉验证决定

系数 (R^2)、RMSECV、平均绝对误差 MAE (Mean Absolute Error) 和相对均方根误差 rRMSE (relative Root Mean Square Error) 进行评估。具体计算如下。

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

$$\text{RMSECV} = \sqrt{\sum_{i=1}^n (\hat{y}_i - y_i)^2 / n} \quad (4)$$

$$\text{MAE} = \sum_{i=1}^n |\hat{y}_i - y_i| / n \quad (5)$$

$$\text{rRMSE} = \text{RMSECV} / \bar{y} \quad (6)$$

式中, $\hat{y}_i$ 表示预测值, $\bar{y}$ 为平均测量值, y_i 为测量值, n 为土壤样本的数量。

3 结果

3.1 SOCD 统计描述

图 2 显示了实验室测量的 SOCD 基本信息。

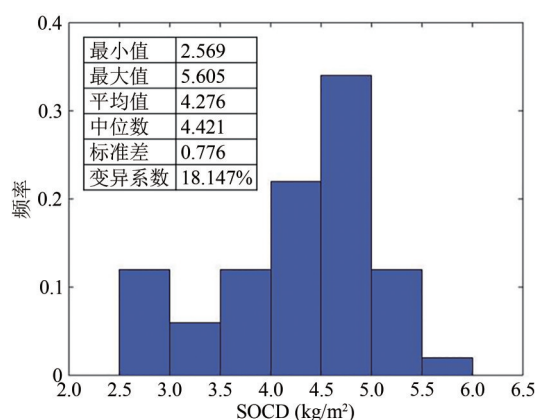


图2 50个土壤样本的SOCD频率分布直方图

Fig. 2 SOCD histograms of 50 soil samples

50 个土壤样本的 SOCD 值分布在 2.569—5.605 kg/m², 平均值为 4.276 kg/m², 中位数为 4.421 kg/m²。标准差 (0.776 kg/m²) 和变异系数 (18.147%) 表明 SOCD 具有中等的变化程度。对 SOCD 值的正态分布进行 Jarque-Bear 检验, 结果显示其在 0.01 的水平上显著。

3.2 土壤反射光谱特征

图 3(a) 展示了 4 种不同 SOCD 梯度条件下的原始光谱的平均反射率曲线。原始土壤光谱反射率总体上呈现出随波长增大而增大的趋势, 并在 900 nm 处达到最大值。原始光谱反射率随着 SOCD 的增加具有明显的降低趋势, 表明较高的 SOCD 将增加土

壤对 400—900 nm 区间电磁波的吸收能力。这主要是受到土壤发色团与有机质本身黑色的影响, 在视觉上表现为暗黑色的土壤比亮色的土壤有机碳含量更高 (史舟 等, 2015)。图 3(c) 展示了在不同 SOCD 梯度下的四条 FD 光谱反射率曲线, 这些曲线都具有相似的形状, 其值主要分布在 -0.004—0.006 之间。与原始光谱相比, FD 光谱凸显了更多的

细节, 但不具有平滑性。图 3(e) 显示了不同 SOCD 梯度条件下的平均 CR 光谱曲线。4 条 CR 光谱反射率曲线具有相似的变化趋势, 在 400—900 nm 波长范围内, 4 条曲线先由最高点 (CR 值=1) 急剧下降, 在 425 nm 附近下降至最低点后, 又开始呈现上升趋势, 随后逐渐趋于平缓, 并在 900 nm 时, 四条曲线重新回到最高点 (CR 值=1)。

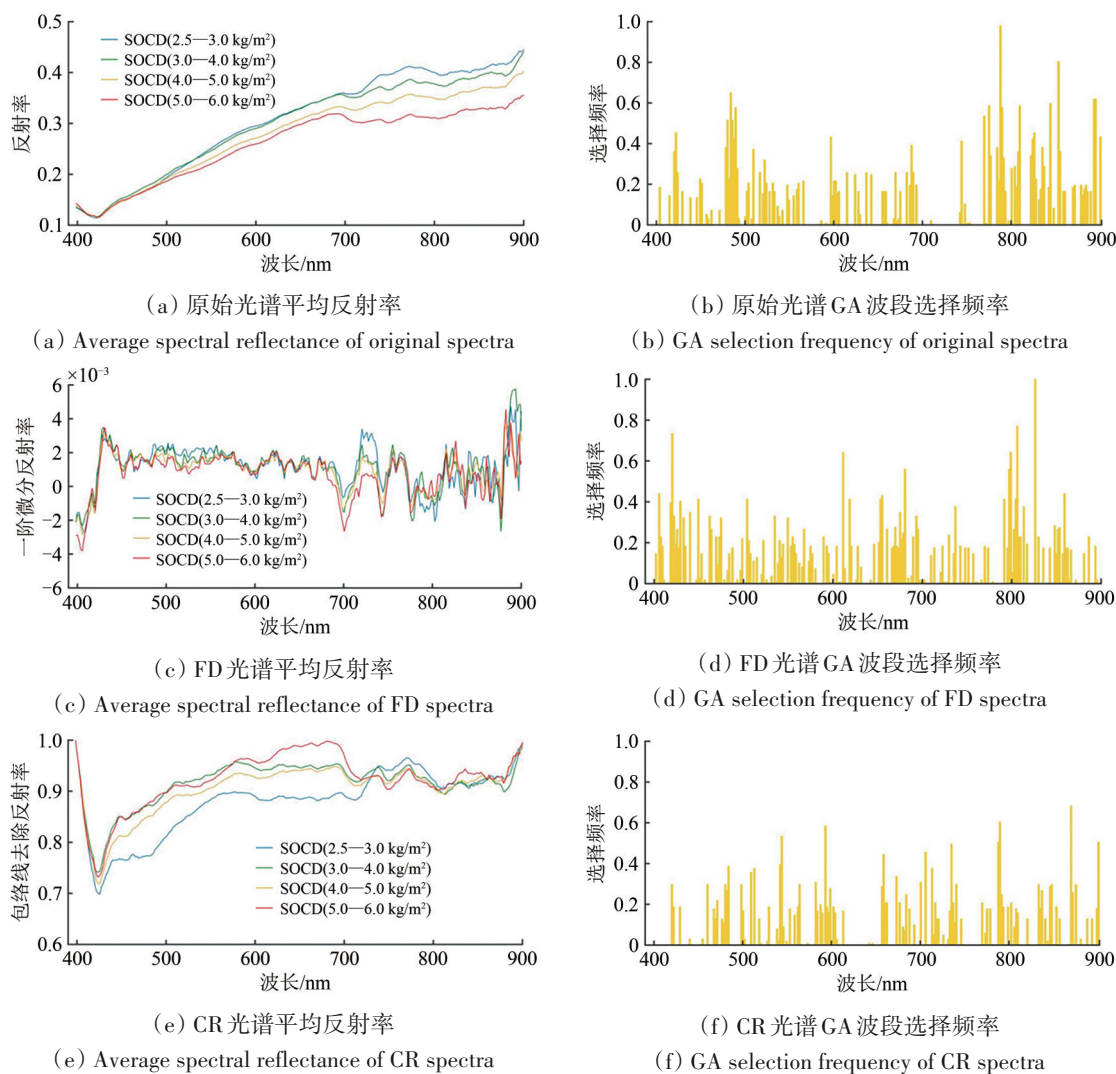


图 3 不同预处理光谱的平均反射率及 GA 波段选择频率

Fig. 3 Average spectral reflectance and GA selection frequency of different pre-processed spectra

3.3 光谱变量选择

图 3(b) 显示了原始光谱变量在 GA 模型中的使用频率。FD、CR 光谱波段的使用频率分别对应图 3(d) 和图 3(f)。若一个变量被使用的频率越高表明它更有可能是特征波段。可以发现 3 种光谱对于各个波段的使用频率具有较大差异, 说明不同

的处理方法将改变光谱对 SOCD 的敏感性。被选择的特征波段位置如图 4 所示。原始光谱的特征波段主要集中在 450—530 nm、600—700 nm 和 770—860 nm 的光谱区域, 而在 730 nm 附近的光谱区域缺乏特征波段。FD 光谱的特征波段主要分布在 400—460 nm、550—700 nm 和 760—840 nm 光谱范围。

CR 光谱的特征波段集中在 430—550 nm、650—750 nm 及 780—890 nm 光谱范围。3 种光谱特征波段的波长范围各有差异, 表明不同预处理方法会影响 SOCD 估算的特征波段。3 种光谱都较为频繁地选取了位于蓝光 (450 nm) 和近红外区域 (760—900 nm) 的波段。对于可见光区域, 450 nm 蓝波段的土壤反射率与 SOC 含量息息相关 (Hong 等, 2020)。对于近红外区域, SOC 的光谱特征通常与某些官能团的吸收特征有关, 如 C—H、C—O 和 N—H 键等 (Shi 等, 2014; Knadel 等, 2013)。经过 GA 变量选择, 最终选取了 29 个原始特征波段、36 个 FD 特征波段以及 24 个 CR 特征波段。

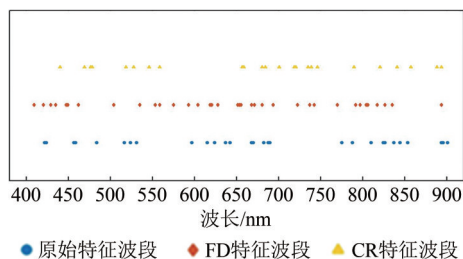


图4 原始光谱、FD光谱和CR光谱的GA选择波段

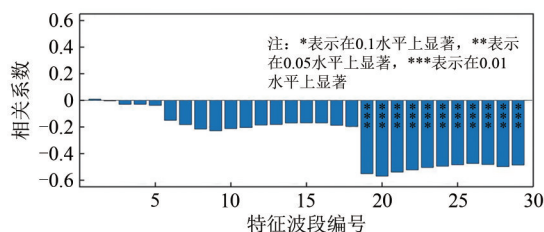
Fig. 4 GA selected bands from original, FD and CR spectra

本研究对原始光谱、FD光谱和CR光谱的特征波段进行分析。3种光谱的特征波段与SOCD间的相关系数如图5所示。对于原始光谱, 超过一半的特征波段与SOCD相关系数的绝对值大于0.2, 其中, 部分波段的相关系数的绝对值超过了0.4, 且均在0.01水平上显著。但也存在相关性很低的波段, 大部分原始光谱的特征波段与SOCD呈现负相关。FD光谱的特征波段与SOCD的相关性则普遍较低, 有少量波段的相关系数绝对值超过了0.2, 其中有大部分在0.05水平上显著。CR光谱的特征波段与SOCD的相关性也较低, 仅有7个特征波段的相关系数高于0.2, 大部分特征波段的相关系数都较低, 与SOCD呈现正相关的CR特征波段多于负相关的特征波段。整体来看, 原始光谱、FD光谱和CR光谱与SOCD相关性较高的波段都集中在近红外范围 (780—900 nm), 但FD和CR光谱处理方法降低了光谱对SOCD的敏感性。

3.4 SOCD预测效果及精度评价

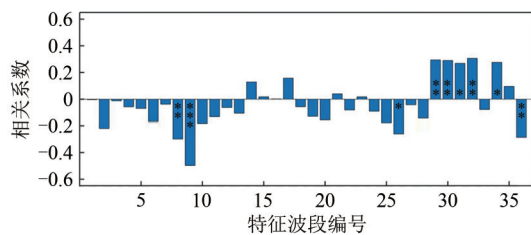
由表1所示, 对于原始光谱, GA-ANN模型 ($R^2=0.678$) 的预测精度最高, GA-PLSR模型 ($R^2=0.672$) 与 GA-MLR模型 ($R^2=0.621$) 也具有较好

的预测效果, GA-PLSR ($R^2=0.526$) 和 GA-ANN ($R^2=0.505$), 而 GA-SVM模型 ($R^2=0.452$) 的精度最低。对于CR光谱, GA-MLR ($R^2=0.602$) 要优于 GA-ANN ($R^2=0.464$)、GA-PLSR ($R^2=0.372$) 和 GA-SVM ($R^2=0.332$) 模型。原始光谱全波段与特征波段估算SOCD的散点图如图6所示。与原始光谱的全波段相比, 使用其特征波段的散点图更接近1:1线。上述结果表明GA波段选择方法能够提高模型预测的精度。



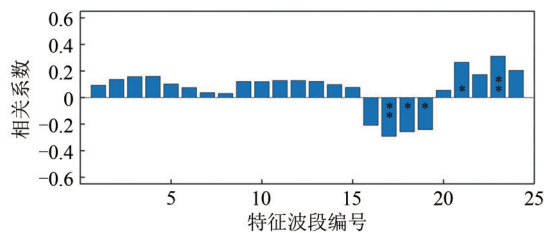
(a) 原始光谱特征波段与SOCD的相关系数

(a) Correlation coefficients between SOCD and the bands selected from original spectra



(b) FD光谱特征波段与SOCD的相关系数

(b) Correlation coefficients between SOCD and the bands selected from FD spectra



(c) FD光谱特征波段与SOCD的相关系数

(c) Correlation coefficients between SOCD and the bands selected from CR spectra

图5 不同预处理光谱的特征波段与SOCD的相关系数
Fig. 5 Correlation coefficients between SOCD and the bands selected from different pre-processed spectra

对于原始光谱、FD光谱和CR光谱, 均通过GA波段选择方法提高了SOCD的估算精度, 表现为更高的 R^2 与更低的 RMSECV、MAE 和 rRMSE (表1)。使用原始光谱特征波段预测SOCD的 R^2 范围为 0.551—0.678, 而使用FD光谱特征波段的 R^2

范围为0.452—0.593, CR光谱特征波段的 R^2 则在0.332—0.602。只有原始光谱的特征波段具有较好的预测效果。而使用FD光谱与CR光谱的特征波段存在较大的误差,无法满足SOC_D制图的精度要求。基于上述结果,本研究利用原始光谱的特征波段绘制了研究区的SOC_D。

表1 基于不同光谱、不同回归模型的SOC_D预测效果

Table 1 Prediction results of SOC _D by using multivariate regression models and different pre-processed spectra									
光谱类型	指标	GA特征波段				全光谱波段			
		PLSR	MLR	SVM	ANN	PLSR	MLR	SVM	ANN
原始光谱	R^2	0.672	0.621	0.551	0.678	0.388	0.255	0.315	0.343
	RMSECV	0.458	0.514	0.525	0.449	0.665	0.867	0.644	0.630
	MAE	0.380	0.385	0.420	0.358	0.493	0.616	0.484	0.476
	rRMSE/%	10.7	12.0	12.3	10.5	15.6	20.3	15.1	14.7
FD光谱	R^2	0.526	0.593	0.452	0.505	0.284	0.094	0.193	0.201
	RMSECV	0.549	0.586	0.578	0.552	0.659	0.907	0.691	0.693
	MAE	0.462	0.471	0.436	0.449	0.515	0.734	0.535	0.530
	rRMSE/%	12.8	13.7	13.5	12.9	15.4	21.2	16.2	16.2
CR光谱	R^2	0.372	0.602	0.332	0.464	0.231	0.162	0.212	0.274
	RMSECV	0.628	0.538	0.629	0.571	0.733	0.973	0.689	0.656
	MAE	0.504	0.521	0.467	0.443	0.578	0.745	0.520	0.593
	rRMSE/%	14.7	12.6	14.7	13.4	17.1	22.8	16.1	15.3

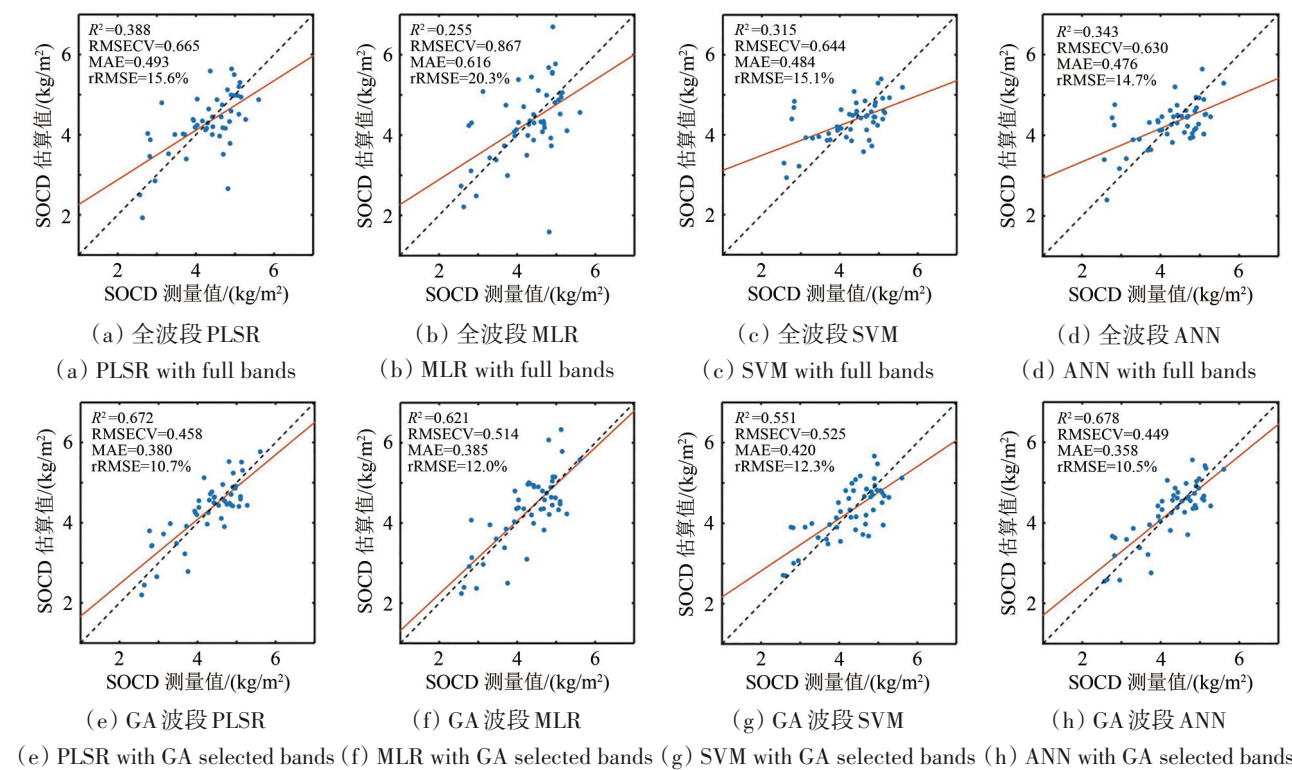


图6 原始光谱全波段与GA特征波段估算SOC_D的散点图

Fig. 6 Scatter plots of SOC_D estimated by full bands and GA selected bands of original spectra

3.5 SOC_D数字土壤制图

基于原始光谱特征波段估算整幅影像的SOC_D,其空间分布如图7所示。PLSR、MLR与ANN等

3个模型的SOC_D空间分布具有较高的相似性,但ANN相较于PLSR与MLR的SOC_D变化更为平缓。其中,PLSR模型的SOC_D估算值变化范围分别为

0—6.66 kg/m², MLR 与 ANN 模型估算的 SOCD 变化范围分别为 0—7.33 kg/m² 及 2.20—5.97 kg/m²。在 SOCD 的空间分布上, SVM 模型与 PLSR、MLR、ANN 模型具有较大的差异, 表现为 SOCD 总体上要高于其他 3 个预测模型, 其研究区内的 SOCD 范围为 0—6.92 kg/m²。总体而言, 在研究区的西部

和东南处, 4 种模型的 SOCD 变化趋势较为相似。然而, 在研究区的东北角却有不同, 表现为 PLSR 与 SVM 模型预测的 SOCD 比 MLR 和 ANN 模型的要高。这种不确定性是由于研究区的东北部缺乏样本导致的。

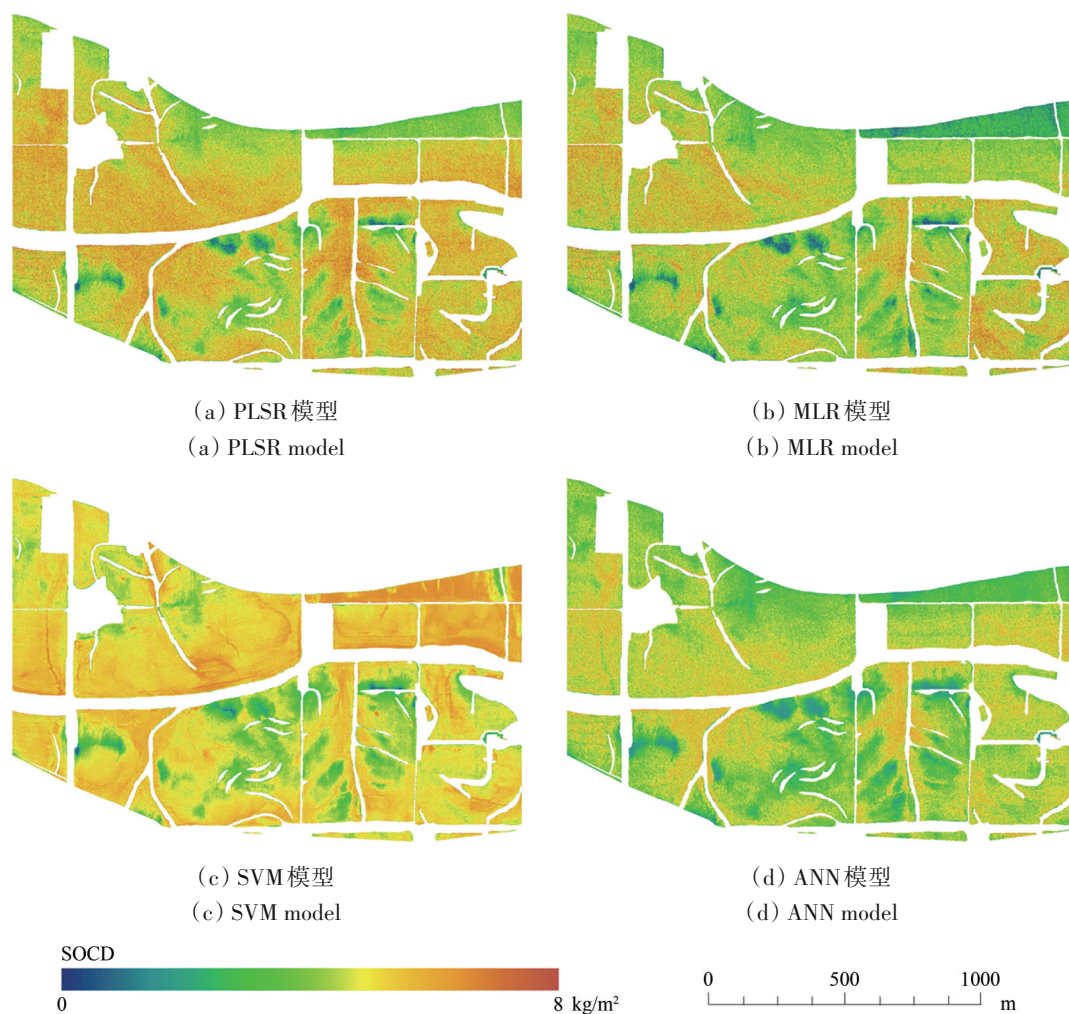


图7 4种回归模型的SOCD制图结果
Fig. 7 SOCD mapping by four regression models

3.6 不同SOCD预测模型之间的比较

不同预测模型估算的SOCD空间变化趋势总体来说相似。线性模型(PLSR、MLR)与非线性模型(SVM、ANN)预测的SOCD在内部存在明显差异。表2比较了SOCD测量值与4种模型预测值的基本统计信息。其中, 土壤样本SOCD测量值与PLSR、MLR、SVM和ANN预测值的平均值相近, 分别为4.276、4.307、4.302、4.347和4.297。标准差分别为0.776、0.762、0.921、0.598和0.745, 其

中, PLSR、ANN预测值的标准差与测量值的标准差较为相近, MLR预测值的标准差要高于测量值的标准差, 而SVM预测值的标准差则低于测量值的标准差, 这表明MLR模型加剧了SOCD的离散程度, SVM模型则会减弱SOCD的离散程度。为了进一步探究SOCD残差的空间结构特征, 采用全局Moran's I来分析各预测模型SOCD残差的空间特征。PLSR、MLR、SVM和ANN模型的残差全局Moran's I分别为-0.020、-0.329、-0.059和-0.064。

4种模型的全局 Moran's I 均为负数表明SOCD残差呈现出一定的空间离散性,然而,只有MLR模型残差的离散性较为明显,其余3种模型残差的空间分布则较为随机。

表2 SOCD测量值与4种模型预测值的基本统计信息
Table 2 Basic statistics of SOCD measured values and predicted values of four models

统计指标	测量值	模型			
		PLSR	MLR	SVM	ANN
最小值/(kg/m ²)	2.569	2.205	2.234	2.715	2.546
最大值/(kg/m ²)	5.605	5.771	6.337	5.456	5.566
平均值/(kg/m ²)	4.276	4.307	4.306	4.347	4.297
标准差/(kg/m ²)	0.776	0.762	0.921	0.598	0.745
残差全局 Moran's I		-0.020	-0.329	-0.059	-0.064

4 讨 论

尽管本研究没有使用短波红外光谱信息,但机载 Headwall Micro-Hyperspec 的 A 系列影像仍然能实现中等精度的SOCD制图。在可见光近红外区域具有与SOCD相关的光谱范围,这是SOCD能被成功预测的关键(刘艳芳等,2017)。除此之外,该光谱区域也经常用于SOC和SOM的估算。如Castaldi等(2016)使用400—600 nm的可见光波段估算了SOC,取得了良好的效果。Liu等(2009)利用620—810 nm波段的吸收特征对SOM进行了成功预测。本研究发现SOCD预测的关键波段主要集中在450—530 nm以及770—860 nm波长范围,与SOC和SOM预测的重要波长较为吻合,这些结果表明SOCD、SOC和SOM之间可能具有相似的光谱吸收特征。Shi等(2014)认为在400—900 nm光谱范围内,用于SOC估算的可见光波段主要由土壤发色团和腐殖酸的暗黑程度决定,SOC在近红外的光谱特征可能与官能团(如C—H、C—O和N—H键)基频振动的泛音和合谱带有关。

聂哲等(2019)的研究表明FD和CR方法提高了500—920 nm的光谱对SOC的相关性,并认为光谱处理方法(一阶导数、包络线去除等)往往能得到更好的预测结果。但本研究发现,对于机载高光谱数据,FD和CR处理方法均降低了对SOCD的相关性,使用平滑的原始光谱的预测效果要比FD和CR光谱更好。Vaudour等(2016)的研究也发现了这种现象。其原因可能是机载高光谱影像与近端光谱相比存在较高的噪声,一阶导数和包络线去除处理会放大这种噪声,并会改变光谱对SOCD的敏感性,从而影响预测精度。

基于原始特征波段,采用PLSR、MLR、SVM和ANN预测SOCD的精度各异, R^2 都介于0.55—0.68,估算精度均属于中等。估算精度从高到低分别为:ANN($R^2=0.678$)>PLSR($R^2=0.672$)>MLR($R^2=0.621$)>SVM($R^2=0.551$)。这些结果与以往的研究差异较大,如Reda等(2021)人使用PLSR与ANN模型估算SOC的 R^2 分别为0.90与0.88,取得了较高的精度。卢延年等(2014)的研究结果表明PLSR和SVM的预测精度要优于MLR。造成这种不同的原因有很多,包括不同的数据源、不同的预处理方法等。另外,不同的试验条件和土壤类型会导致光谱与土壤属性之间线性和非线性关系发生改变。

样本量较少往往会降低模型的整体精度,本研究采用的4种预测模型表现出了中等精度,受到了样本量不足的困扰。然而,高光谱数据具有丰富的光谱信息,而GA特征选择算法又可从中筛选出判别力强的特征子集,因此仍可满足空间制图的精度要求(Li等,2007)。同时,本研究所使用的SOCD样本趋于正态分布,小样本对于SOCD这种普遍呈正态分布的土壤属性影响较小(陈奕云等,2017)。PLSR与MLR线性模型的估算精度高于SVM非线性模型,这也体现了样本量少的影响。然而,与SVM不同,非线性模型ANN甚至要优于PLSR、MLR两种线性模型,这可能与ANN对非线性交互数据强大的处理能力有关(纪文君等,2012)。

图8显示了PLSR、MLR、SVM和ANN预测SOCD的绝对误差的空间分布特征,绝对误差是由预测值与实测值之差的绝对值计算而来。总体来说,四种模型预测SOCD的绝对误差在空间上具有

相似的分布, 绝对误差大多在 0.5 kg/m^2 以下。但也存在绝对误差较大的点, 这些点大多集中分布,

且多出现在采样点的边缘附近。

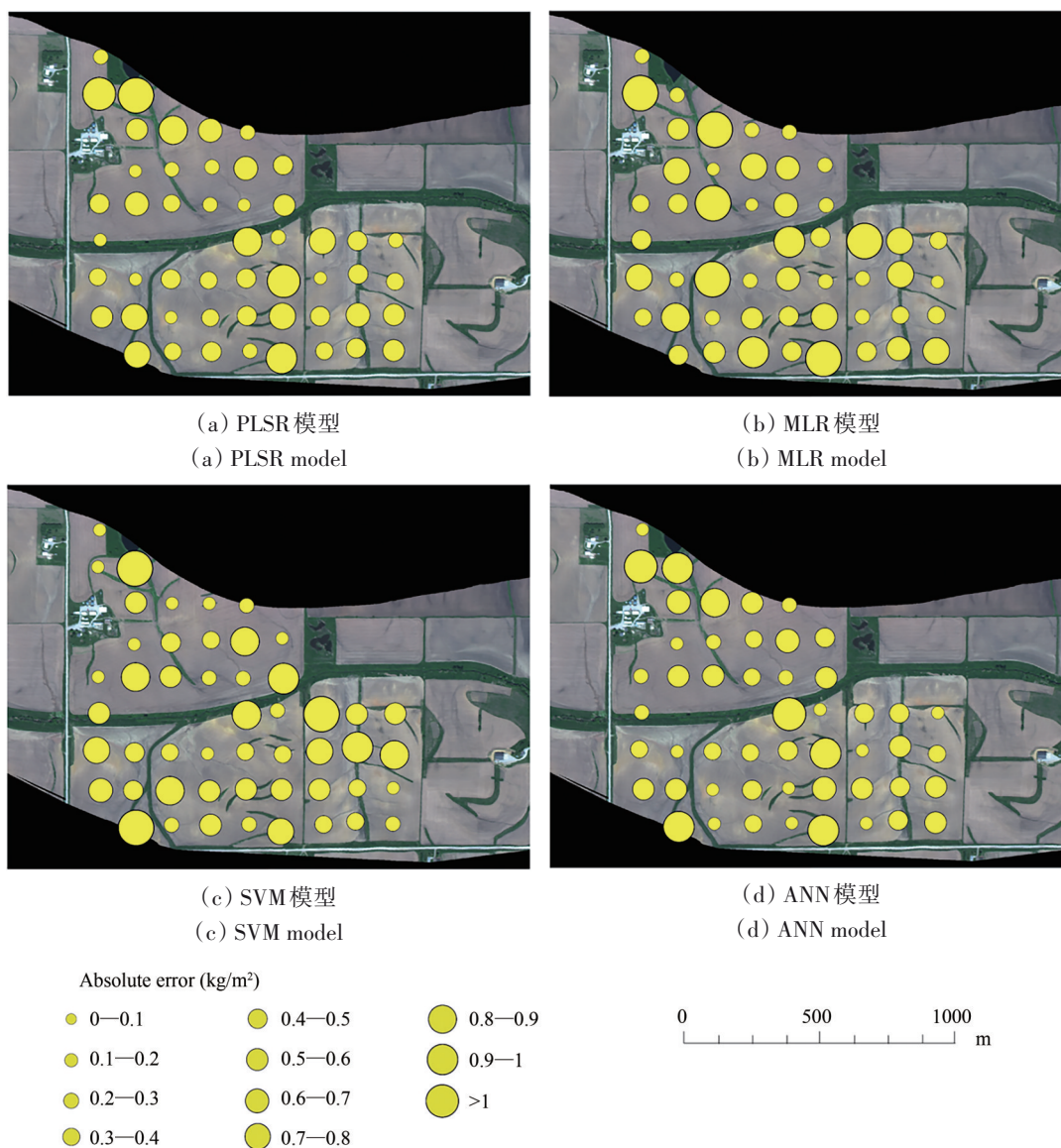


图8 4种模型预测SOCD的绝对误差的空间分布情况

Fig. 8 The spatial distribution of the absolute error of SOCD predict by four models

5 结 论

本研究使用GA方法对机载高光谱影像进行了SOCD预测。本研究考虑了不同预处理方法(一阶导数变换、连续统去除变换)对预测结果的影响。此外, 还比较了不同回归模型(PLSR、MLR、SVM和ANN)在SOCD预测上的效果。结果表明, GA可以选择与SOCD相关的敏感波段, 提高机载高光谱预测SOCD的能力。与全光谱波段相比, 使用原始光谱、FD和CR特征波段预测SOCD的精度均有所提高。采用四种回归模型对原始光谱的特

征波段进行SOCD数字制图, ANN是最佳的预测模型($R^2=0.678$), SVM的预测精度最差($R^2=0.551$), PLSR和MLR预测的 R^2 分别为0.672和0.621。本研究证实了机载高光谱影像在SOCD制图方面的能力。在未来的研究中, 还需要在不同数据集、不同土壤类型及不同地理尺度上进行测试和验证。

参考文献(References)

Ben-Dor E. 2002. Quantitative remote sensing of soil properties. *Advances in Agronomy*, 75: 173-243 [DOI: 10.1016/S0065-2113(02)75005-0]

- Besalatpour A A, Ayoubi S, Hajabbasi M A, Mosaddeghi M R and Schulin R. 2013. Estimating wet soil aggregate stability from easily available properties in a highly mountainous watershed. *CATE-NA*, 111: 72-79 [DOI: 10.1016/j.catena.2013.07.001]
- Cai J N, Liu H L, Jiang B, He T H, Chen W J, Feng Z W, Li Z L and Xing Q G. 2020. Using hyperspectral imagery and GA-PLS algorithm to estimate chemical oxygen demand concentration of water in river network. *Journal of Irrigation and Drainage*, 39(9): 126-131 (蔡建楠, 刘海龙, 姜波, 何甜辉, 陈文杰, 冯志伟, 黎倬琳, 邢前国. 2020. 基于 GA-PLS 算法的河网水体化学需氧量高光谱反演. *灌溉排水学报*, 39(9): 126-131) [DOI: 10.13522/j.cnki.gggs.2020063]
- Castaldi F, Palombo A, Santini F, Pascucci S, Pignatti S and Casa R. 2016. Evaluation of the potential of the current and forthcoming multispectral and hyperspectral imagers to estimate soil texture and organic carbon. *Remote Sensing of Environment*, 179: 54-65 [DOI: 10.1016/j.rse.2016.03.025]
- Chen Y Y, Qi T C, Huang Y J, Wan Y, Zhao R Y, Qi L, Zhang C and Fei T. 2017. Optimization method of calibration dataset for VIS-NIR spectral inversion model of soil organic matter content. *Transactions of the Chinese Society of Agricultural Engineering*, 33(6): 107-114 (陈奕云, 齐天赐, 黄颖菁, 万远, 赵瑞瑛, 仝林, 张超, 费腾. 2017. 土壤有机质含量可见-近红外光谱反演模型校正集优选方法. *农业工程学报*, 33(6): 107-114) [DOI: 10.11975/j.issn.1002-6819.2017.06.014]
- Chen Z, Lü C H, Fan L and Wu H. 2011. Effects of land use change on soil organic carbon: a review. *Acta Ecologica Sinica*, 31(18): 5358-5371 (陈朝, 吕昌河, 范兰, 武红. 2011. 土地利用变化对土壤有机碳的影响研究进展. *生态学报*, 31(18): 5358-5371)
- Chu X L, Yuan H F, Wang Y B and Lu W Z. 2001. Variable selection for partial least squares modeling by genetic algorithms. *Chinese Journal of Analytical Chemistry*, 29(4): 437-442 (褚小立, 袁洪福, 王艳斌, 陆婉珍. 2001. 遗传算法用于偏最小二乘方法建模中的变量筛选. *分析化学*, 29(4): 437-442) [DOI: 10.3321/j.issn:0253-3820.2001.04.018]
- Gomez C, Adeline K, Bacha S, Driessen B, Gorretta N, Lagacherie P, Roger J M and Briottet X. 2018. Sensitivity of clay content prediction to spectral configuration of VNIR/SWIR imaging data, from multispectral to hyperspectral scenarios. *Remote Sensing of Environment*, 204: 18-30 [DOI: 10.1016/j.rse.2017.10.047]
- Guo L, Fu P, Shi T Z, Chen Y Y, Zhang H T, Meng R and Wang S Q. 2020. Mapping field-scale soil organic carbon with unmanned aircraft system-acquired time series multispectral images. *Soil and Tillage Research*, 196: 104477 [DOI: 10.1016/j.still.2019.104477]
- Guo L, Sun X R, Fu P, Shi T Z, Dang L N, Chen Y Y, Linderman M, Zhang G L, Zhang Y, Jiang Q H, Zhang H T and Zeng C. 2021. Mapping soil organic carbon stock by hyperspectral and time-series multispectral remote sensing images in low-relief agricultural areas. *Geoderma*, 398: 115118 [DOI: 10.1016/J.GEODERMA.2021.115118]
- Hong Y S, Chen S C, Chen Y Y, Linderman M, Mouazen A M, Liu Y L, Guo L, Yu L, Liu Y F, Cheng H and Liu Y. 2020. Comparing laboratory and airborne hyperspectral data for the estimation and mapping of topsoil organic carbon: feature selection coupled with random forest. *Soil and Tillage Research*, 199: 104589 [DOI: 10.1016/j.still.2020.104589]
- Ji W J, Li X, Li C X, Zhou Y and Shi Z. 2012. Using different data mining algorithms to predict soil organic matter based on visible-near infrared spectroscopy. *Spectroscopy and Spectral Analysis*, 32(9): 2393-2398, 2408 (纪文君, 李曦, 李成学, 周银, 史舟. 2012. 基于全谱数据挖掘技术的土壤有机质高光谱预测建模研究. *光谱学与光谱分析*, 32(9): 2393-2398, 2408) [DOI: 10.3964/j.issn.1000-0593(2012)09-2393-06]
- Knadel M, Viscarra Rossel R A, Deng F, Thomsen A and Greve M H. 2013. Visible-near infrared spectra as a proxy for topsoil texture and glacial boundaries. *Soil Science Society of America Journal*, 77(2): 568-579 [DOI: 10.2136/sssaj2012.0093]
- Lal R. 2004. Soil carbon sequestration impacts on global climate change and food security. *Science*, 304(5677): 1623-1627 [DOI: 10.1126/science.1097396]
- Leardi R, Boggia R and Terrile M. 1992. Genetic algorithms as a strategy for feature selection. *Journal of Chemometrics*, 6(5): 267-281 [DOI: 10.1002/cem.1180060506]
- Li L, Ustin S L and Riano D. 2007. Retrieval of fresh leaf fuel moisture content using genetic algorithm partial least squares (GA-PLS) modeling. *IEEE Geoscience and Remote Sensing Letters*, 4(2): 216-220 [DOI: 10.1109/LGRS.2006.888847]
- Liu H J, Zhang Y Z and Zhang B. 2009. Novel hyperspectral reflectance models for estimating black-soil organic matter in Northeast China. *Environmental Monitoring and Assessment*, 154(1/4): 147-154 [DOI: 10.1007/s10661-008-0385-4]
- Liu Y F, Song Y L, Guo L, Chen Y Y, Lu Y N and Liu Y. 2017. Geostatistical models of soil organic carbon density prediction based on soil hyperspectral reflectance. *Transactions of the Chinese Society of Agricultural Engineering*, 33(2): 183-191 (刘艳芳, 宋玉玲, 郭龙, 陈奕云, 卢延年, 刘以. 2017. 结合高光谱信息的土壤有机碳密度地统计模型. *农业工程学报*, 33(2): 183-191) [DOI: 10.11975/j.issn.1002-6819.2017.02.025]
- Lu Y N, Liu Y F, Chen Y Y and Jiang Q H. 2014. Optimization of the hyperspectral prediction model of soil organic carbon contents of Jiangnan plain. *Chinese Agricultural Science Bulletin*, 30(26): 127-133 (卢延年, 刘艳芳, 陈奕云, 姜庆虎. 2014. 江汉平原土壤有机碳含量高光谱预测模型优选. *中国农学通报*, 30(26): 127-133) [DOI: 10.11924/j.issn.1000-6850.2014-0439]
- Luo M, Guo L, Zhang H T, Wang S Q and Liang P. 2020. Characterization of spatial distribution of soil organic carbon in China based on environmental variables. *Acta Pedologica Sinica*, 57(1): 48-59 (罗梅, 郭龙, 张海涛, 汪善勤, 梁攀. 2020. 基于环境变量的中国土壤有机碳空间分布特征. *土壤学报*, 57(1): 48-59) [DOI: 10.11766/trxb201812110454]
- Nie Z, Li X F, Lv J X and Zheng X. 2019. Hyperspectral retrieval of surface soil organic matter content in a typical black soil region of

- northeast China. *Chinese Journal of Soil Science*, 50(6): 1285-1293 (聂哲, 李秀芬, 吕家欣, 郑晓. 2019. 东北典型黑土区表层土壤有机质含量高光谱反演研究. *土壤通报*, 50(6): 1285-1293) [DOI: 10.19336/j.cnki.trtb.2019.06.04]
- Nouri M, Gomez C, Gorretta N and Roger J M. 2017. Clay content mapping from airborne hyperspectral Vis-NIR data by transferring a laboratory regression model. *Geoderma*, 298: 54-66 [DOI: 10.1016/j.geoderma.2017.03.011]
- Peng J, Li X, Zhou Q, Shi Z, Ji W J and Wang J Q. 2013. Influence of iron oxide on the spectral characteristics of organic matter. *Journal of Remote Sensing*, 17(6): 1396-1412 (彭杰, 李曦, 周清, 史舟, 纪文君, 王家强. 2013. 氧化铁对有机质光谱特性的影响分析. *遥感学报*, 17(6): 1396-1412) [DOI: 10.11834/jrs.20132273]
- Qiao T, Lv C W, Xiao W P, Lv K and Shui H W. 2018. Hyperspectral prediction modeling of soil texture based on genetic algorithm. *Chinese Journal of Soil Science*, 49(4): 773-778 (乔天, 吕成文, 肖文凭, 吕凯, 水宏伟. 2018. 基于遗传算法的土壤质地高光谱预测模型研究. *土壤通报*, 49(4): 773-778) [DOI: 10.19336/j.cnki.trtb.2018.04.03]
- Reda R, Saffaj T, Derrouz H, Itiq S E, Bouzida I, Saidi O, Lakssir B and El Hadrami E M. 2021. Comparing CalReg performance with other multivariate methods for estimating selected soil properties from Moroccan agricultural regions using NIR spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, 211: 104277 [DOI: 10.1016/j.chemolab.2021.104277]
- Shi T Z, Chen Y Y, Liu H Z, Wang J J and Wu G F. 2014. Soil organic carbon content estimation with laboratory-based visible-near-infrared reflectance spectroscopy: feature selection. *Applied Spectroscopy*, 68(8): 831-837 [DOI: 10.1366/13-07294]
- Shi Z, Guo Y, Jin X and Wu H X. 2011. Advancement in study on proximal soil sensing. *Acta Pedologica Sinica*, 48(6): 1274-1281 (史舟, 郭燕, 金希, 吴豪翔. 2011. 土壤近地传感器研究进展. *土壤学报*, 48(6): 1274-1281) [DOI: 10.11766/trxb201012070517]
- Shi Z, Liang Z Z, Yang Y Y and Guo Y. 2015. Status and prospect of agricultural remote sensing. *Transactions of the Chinese Society for Agricultural Machinery*, 46(2): 247-260 (史舟, 梁宗正, 杨媛媛, 郭燕. 2015. 农业遥感研究现状与展望. *农业机械学报*, 46(2): 247-260) [DOI: 10.6041/j.issn.1000-1298.2015.02.037]
- Shi Z, Xu D Y, Teng H F, Hu Y M, Pan X Z and Zhang G L. 2018. Soil information acquisition based on remote sensing and proximal soil sensing: current status and prospect. *Progress in Geography*, 37(1): 79-92 (史舟, 徐冬云, 滕洪芬, 胡月明, 潘贤章, 张甘霖. 2018. 土壤星地传感技术现状与发展趋势. *地理科学进展*, 37(1): 79-92) [DOI: 10.18306/dlkxjz.2018.01.009]
- Vapnik V N. 1998. *Statistical Learning Theory*. New York: John Wiley & Sons
- Vaudour E, Gilliot J M, Bel L, Lefevre J and Chehdi K. 2016. Regional prediction of soil organic carbon content over temperate croplands using visible near-infrared airborne hyperspectral imagery and synchronous field spectra. *International Journal of Applied Earth Observation and Geoinformation*, 49: 24-38 [DOI: 10.1016/j.jag.2016.01.005]
- Viscarra Rossel R A and Behrens T. 2010. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 158(1/2): 46-54 [DOI: 10.1016/j.geoderma.2009.12.025]
- Wen F P, Zhao W, Hu L, Xu H X and Cui Q. 2021. SMAP passive microwave soil moisture spatial downscaling based on optical remote sensing data: a case study in Shandian river basin. *National Remote Sensing Bulletin*, 25(4): 962-973 (文凤平, 赵伟, 胡路, 徐红新, 崔倩. 2021. 耦合MODIS数据的SMAP被动微波土壤水分空间降尺度研究——以闪电河流域为例. *遥感学报*, 25(4): 962-973) [DOI: 10.11834/jrs.20219393]
- Wold H. 1966. Nonlinear estimation by iterative least squares procedures//David F N and Neyman J, eds. *Research Papers in Statistics*, Festschrift for J. Neyman. New York: Wiley: 441-444
- Xie X F, Wu T, Zhu M, Jiang G J, Xu Y, Wang X H and Pu L J. 2021. Comparison of random forest and multiple linear regression models for estimation of soil extracellular enzyme activities in agricultural reclaimed coastal saline land. *Ecological Indicators*, 120: 106925 [DOI: 10.1016/j.ecolind.2020.106925]
- Yu L, Hong Y S, Zhou Y, Zhu Q, Xu L, Li J Y and Nie Y. 2016. Wavelength variable selection methods for estimation of soil organic matter content using hyperspectral technique. *Transactions of the Chinese Society of Agricultural Engineering*, 32(13): 95-102 (于雷, 洪永胜, 周勇, 朱强, 徐良, 李冀云, 聂艳. 2016. 高光谱估算土壤有机质含量的波长变量筛选方法. *农业工程学报*, 32(13): 95-102) [DOI: 10.11975/j.issn.1002-6819.2016.13.014]
- Zhan B S, Ni J H and Li J. 2014. Hyperspectral technology combined with CARS algorithm to quantitatively determine the SSC in Korla fragrant pear. *Spectroscopy and Spectral Analysis*, 34(10): 2752-2757 (詹白勺, 倪君辉, 李军. 2014. 高光谱技术结合CARS算法的库尔勒香梨可溶性固形物定量测定. *光谱学与光谱分析*, 34(10): 2752-2757) [DOI: 10.3964/j.issn.1000-0593(2014)10-2752-06]
- Zheng G H, Wang M J, Jiao C X and Sun D M. 2013. Review on prediction of soil organic matter with reflectance spectroscopy. *Journal of Nanjing University of Information Science and Technology (Natural Science Edition)*, 5(6): 481-486 (郑光辉, 王明江, 焦彩霞, 孙东敏. 2013. 土壤有机质高光谱估算研究进展. *南京信息工程大学学报(自然科学版)*, 5(6): 481-486) [DOI: 10.13878/j.cnki.jnuist.2013.06.001]
- Zou X B and Zhao J J. 2007. Methods of characteristic wavelength region and wavelength selection based on genetic algorithm. *Acta Optica Sinica*, 27(7): 1316-1321 (邹小波, 赵杰文. 2007. 用遗传算法快速提取近红外光谱特征区域和特征波长. *光学学报*, 27(7): 1316-1321) [DOI: 10.3321/j.issn:0253-2239.2007.07.032]

Mapping of soil organic carbon density at farmland scale based on airborne hyperspectral images

LIU Qian¹, WANG Mengdi¹, GUO Long², WANG Ran³, JIA Zhongfu³, HU Xianjun⁴,
TANG Qiankun⁵, SHI Tiezhu¹

1. MNR Key Laboratory for Geo-Environmental Monitoring of Great Bay Area & Guangdong Key Laboratory of Urban Informatics & Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China;

2. College of Resources and Environment, Huazhong Agricultural University, Wuhan 430070, China;

3. Inner Mongolia Autonomous Region Surveying, Mapping and Geoinformation Center, Hohhot 010020, China;

4. Naval University of Engineering, Wuhan 430032, China;

5. MNR The Third Institute of Photogrammetry and Remote Sensing & Sichuan Bureau of Surveying, Mapping and Geoinformation, Chengdu 610100, China

Abstract: Accurate monitoring of Soil Organic Carbon Density (SOCD) is important for regulating soil carbon sinks and rationally using soil resources. Airborne hyperspectral images provide important data sources for SOCD mapping. The noise in the spectrum affects the accuracy of SOCD estimation because airborne hyperspectral images are easily affected by external factors during data collection. A set of technical processes that are suitable for airborne hyperspectral data processing is still lacking. Therefore, this study aims to investigate the technical process of SOCD estimation based on airborne hyperspectral images. The original spectra are preprocessed by First Derivative (FD) and Continuum Removal (CR) transform. Genetic Algorithm (GA) was used to select the feature bands. Different regression methods, such as Partial Least-Squares Regression (PLSR), Multiple Linear Regression (MLR), Support Vector Machine (SVM), and Artificial Neural Network (ANN), were used to estimate SOCD. Results showed that the accuracy of SOCD prediction for original, FD, and CR spectra was improved after feature band selection by GA. With the feature bands of original spectra, the R^2 of SOCD predicted by PLSR, MLR, SVM, and ANN are 0.672, 0.621, 0.551, and 0.678, respectively. The range of R^2 are 0.452—0.593 and 0.332—0.602 with FD and CR feature bands, respectively, which demonstrate large errors. The feature bands of the original spectrum were used in this study for SOCD mapping. The SOCD predicted by four regression models has a highly similar trend in space and is similar to the SOCD measured value. The points with large absolute errors mostly occur near the edges of the sampling points.

Key words: soil organic carbon density, airborne hyperspectral images, genetic algorithm, digital soil mapping

Supported by National Natural Science Foundation of China (No. 41890854)