

樊泳灼, 李新国. 湖滨绿洲棕漠土有机碳含量高光谱估算[J]. 江苏农业学报, 2023, 39(6): 1341-1348.

doi: 10.3969/j.issn.1000-4440.2023.06.009

湖滨绿洲棕漠土有机碳含量高光谱估算

樊泳灼^{1,2}, 李新国^{1,2}

(1.新疆师范大学地理科学与旅游学院, 新疆 乌鲁木齐 830054; 2.新疆干旱区湖泊环境与资源实验室, 新疆 乌鲁木齐 830054)

摘要: 以博斯腾湖湖滨绿洲为研究区, 利用实测棕漠土有机碳含量与高光谱(350~2 500 nm)数据, 应用竞争性自适应重加权采样算法(CARS)、连续投影算法(SPA)、竞争性自适应重加权采样-连续投影算法(CARS-SPA)筛选棕漠土有机碳含量响应的高光谱特征波段, 分别采用全波段和特征波段结合随机森林(RF)模型构建棕漠土有机碳含量估算模型。结果表明: 博斯腾湖湖滨绿洲棕漠土0~50.0 cm 土层有机碳含量为1.40~40.92 g/kg, 平均值为14.20 g/kg, 变异系数为55.54%, 呈中等变异水平。CARS、SPA、CARS-SPA等算法筛选出的棕漠土有机碳含量响应特征波段分别为122个、11个和10个。基于CARS-SPA算法筛选出的特征波段数据输入RF模型估算效果最好, 验证集检验的决定系数(R^2)、相对分析误差(RPD)、均方根误差(RMSE)分别为0.85、2.59和2.72 g/kg, 该方法能有效减少光谱数据冗余, 提高模型估算精度和运行效率。本研究结果为研究区棕漠土有机碳含量的估算提供参考。

关键词: 土壤有机碳含量; 棕漠土; 高光谱; 竞争性自适应重加权采样-连续投影算法(CARS-SPA); 随机森林

中图分类号: S127

文献标识码: A

文章编号: 1000-4440(2023)06-1341-08

Hyperspectral prediction of organic carbon content of brown desert soil in the lakeside oasis

FAN Yong-zhuo^{1,2}, LI Xin-guo^{1,2}

(1. College of Geographical Sciences and Tourism, Xinjiang Normal University, Urumqi 830054, China; 2. Xinjiang Key Laboratory of Lake Environment and Resources in Arid Regions, Urumqi 830054, China)

Abstract: Using the measured organic carbon content of brown desert soil and hyperspectral (350–2 500 nm) data acquired from the lakeside oasis of Bosten Lake, competitive adaptive reweighted sampling (CARS), successive projection algorithm (SPA) and competitive adaptive reweighted sampling-successive projection algorithm (CARS-SPA) were used to screen the characteristic bands of organic carbon content response in brown desert soil. The prediction model of organic carbon content in brown desert soil was constructed by using full band and characteristic band combined with random forest (RF) model. The results showed that the organic carbon content in the 0–50.0 cm soil layer of the brown desert soil in the lakeside oasis of Bosten Lake was 1.40–40.92 g/kg, with an average of 14.20 g/kg, and the coefficient of variation was 55.54%, showing a

moderate variation level. The response characteristic bands of brown desert soil organic carbon content screened by CARS, SPA and CARS-SPA were 122, 11 and 10, respectively. The best prediction effect was obtained when the characteristic band data selected by the CARS-SPA algorithm were input into the RF model. The determination coefficient (R^2), relative percentage difference (RPD) and root mean square error (RMSE) of the validation set test

收稿日期: 2023-01-18

基金项目: 新疆维吾尔自治区自然科学基金项目(2022D01A214); 国家自然科学基金项目(41661047)

作者简介: 樊泳灼(1999-), 女, 新疆塔城人, 硕士研究生, 主要从事干旱区土壤资源变化及遥感应用研究。(E-mail) yong-zhuofan@sina.com

通讯作者: 李新国, (E-mail) onlinelxg@sina.com

were 0.85, 2.59 and 2.72 g/kg, respectively. This method could effectively reduce the redundancy of spectral data and improve the prediction accuracy and operation efficiency of the model. The results of this study provided a reference for the prediction of organic carbon content in brown desert soil in the lakeside oasis of Bosten Lake.

Key words: soil organic carbon content; brown desert soil; hyperspectral; competitive adaptive reweighted sampling-successive projection algorithm (CARS-SPA); random forest

土壤有机碳含量是耕地质量、土地生产力水平的重要指标^[1-2]。湖滨绿洲是干旱区发展农业的重要区域,干旱区绿洲土壤有机碳在全球生态环境及碳循环方面有重要作用,对绿洲农业的发展有重要意义^[3-4]。传统的土壤有机碳含量测定方法是人工取样后进行实验室分析,费时费力且效率低下。近年来,随着高光谱技术的发展,利用波长为 350~2 500 nm 的高光谱数据进行土壤有机碳含量准确高效监测已成为农业遥感的新热点^[5-7]。

虽然高光谱数据丰富,但同时亦带来了大量的冗余信息。因此,利用高光谱数据进行土壤属性监测的一个重要环节是特征光谱的筛选^[8],即通过筛选对土壤属性敏感的特征波段可有效减少光谱数据冗余,提高估算模型精度和运行效率。吴俊等^[6]通过竞争性自适应重加权采样(CARS)算法筛选特征波段,再利用偏最小二乘回归(PLSR)、支持向量机(SVM)、随机森林(RF)、反向传播神经网络(BPNN)等模型对江西省土壤有机碳含量进行了估算。唐海涛等^[9]利用 CARS 算法筛选特征波段并结合随机森林(RF)算法建立不同类型土壤有机质含量估算模型。章海亮等^[10]利用连续投影(SPA)算法筛选特征波段,并结合遗传算法实现了土壤有机质的监测。Wang 等^[11]利用 SPA 算法提取特征波段并结合偏最小二乘回归模型(PLSR)实现了黄灌区盐碱地土壤含盐量的估算,有效提高了估算值的决定系数(R^2)和相对分析误差(RPD)。Liu 等^[12]利用 CARS 方法筛选特征波段结合非线性模型 RF 估算靖边县土壤有机质含量,估算精度较全波段建模取得了一定的提升。上述研究结果表明利用不同方法筛选特征波段输入估算模型可以较好地减少数据冗余,优化模型结构,提高模型的运行效率。另外,模型的选择也会影响土壤有机碳含量的估算精度。吴俊等^[6]利用高光谱估算江西省土壤有机碳含量的研究表明非线性模型 RF 相较于线性模型 PLSR 建模精度更高。Cheng 等^[13]基于 CARS 算法筛选的特征波段比较了极限学习机(ELM)、BP

神经网络(BPNN)、支持向量机回归(SVMR)和偏最小二乘回归(PLSR)等模型估算土壤全氮含量的差异,结果表明利用 ELM 模型能获得更好的精度。综合上述研究结果,特征光谱的筛选和估算模型的选择均会影响估算模型的精度和效率,且不同的土壤类型、监测的土壤属性指标也会影响到特征光谱的筛选方案和估算模型的选择。目前基于 CARS、SPA、CARS-SPA 算法进行特征光谱的筛选及利用非线性模型 RF 进行干旱区湖滨绿洲棕漠土有机碳含量的估算研究还鲜有报道。

本研究以新疆博斯腾湖湖滨绿洲棕漠土为研究对象,利用 ASD FieldSpec3 地物光谱仪采集棕漠土 350~2 500 nm 波长范围的光谱反射率,以 CARS、SPA 和 CARS-SPA 3 种算法筛选土壤有机碳含量的特征波段,并利用随机森林模型(RF)进行模型构建,实现棕漠土有机碳含量的估算,为区域棕漠土有机碳含量的高效监测提供参考。

1 材料与方法

1.1 研究区概况

博斯腾湖湖滨绿洲地处新疆焉耆盆地,隶属于新疆博湖县^[14-15],位于 $41^{\circ}45' \sim 42^{\circ}10'N$, $86^{\circ}20' \sim 86^{\circ}50'E$ (图 1)。博斯腾湖湖滨绿洲是典型的人工绿洲和自然绿洲混合的湖滨绿洲,面积约为 1 360.0 km²,年平均降水量约 83.6 mm,年平均气温约 8.0~8.6℃,光照充足,属于大陆性荒漠气候^[14-15]。棕漠土是博斯腾湖湖滨绿洲的重要耕作土壤,占研究区土壤总面积的 90%以上^[14-15]。

1.2 棕漠土样品采集与处理

根据研究区的地形、植被及土地利用类型等因素,结合研究区的土壤调查现状,2019 年 7 月在常年耕作的棕漠土地块上选取 10 个代表性样地按照“S”型布设样点并采集 0~10.0 cm、10.1~20.0 cm、20.1~30.0 cm、30.1~40.0 cm、40.1~50.0 cm 土层的土壤样品,采样时去除地表的植物根系及石块等杂质,利用 GPS 记录样点经纬度并记录采样点周围环境,每个样

点运用四分法选取约 200 g 土样装袋标号密封带回实验室,共采集 50 份棕漠土样品。将棕漠土样品进行自然风干、研磨并过 2 mm 筛后封装为 2 部分,一部分利用重铬酸钾容量法测定土壤有机碳含量,另一部

分用于高光谱测定^[9,14]。采用浓度梯度法对土壤样本按照 3:1 比例划分训练集与验证集^[16],50 份土壤样品,将 37 份土壤样品作为训练集用于建模,13 份土壤样品作为验证集用于模型验证。

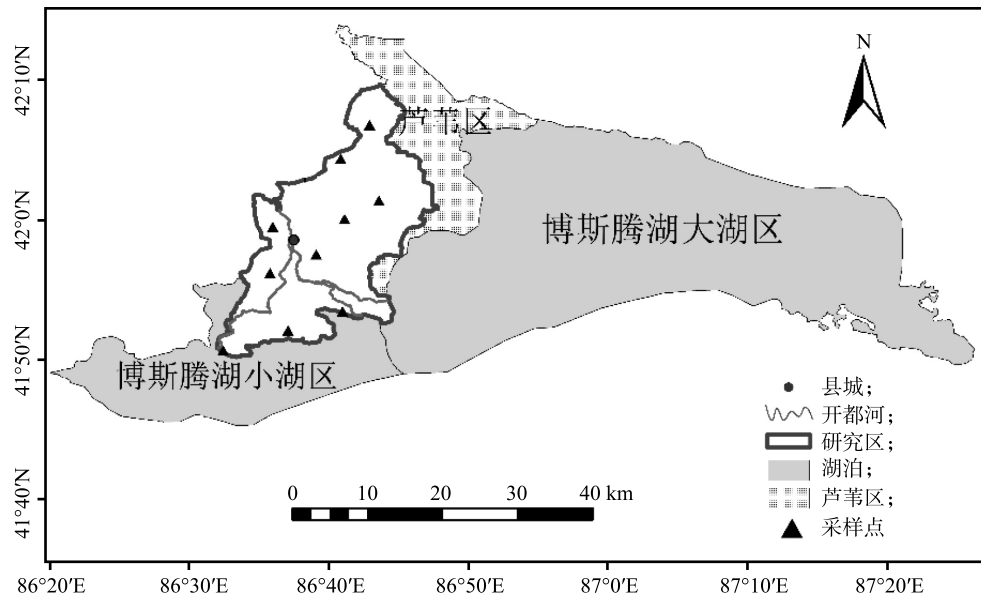


图1 研究区位置及采样地分布

Fig.1 Location of the study area and distribution of sampling sites

1.3 光谱数据采集及预处理

使用 ASD FieldSpec3 地物光谱仪(美国 ASD 公司产品)选择晴朗无风天气于室外测定棕漠土样品的高光谱数据,测定时间为北京时间 12:00–14:00,每个样品采集 10 条光谱数据以平均值为该样品的最终光谱数据。剔除信噪比较低的 2 430~2 500 nm 波段及受水汽影响较大的 1 300~1 450 nm 波段和 1 800~1 950 nm 波段^[14]。为提升光谱曲线信噪比,利用 The Unscrambler 软件对光谱数据进行 Savitzky-Golay(SG)滤波处理^[6]。考虑到低阶微分处理光谱数据可有效减少噪声的干扰^[17],因此再对 SG 滤波处理后数据进行一阶微分变换,进而进行特征波段的筛选。

1.4 棕漠土有机碳含量特征光谱筛选

本研究采用竞争性自适应重加权采样算法(CARS)、连续投影算法(SPA)、竞争性自适应重加权采样-连续投影算法(CARS-SPA)3 种算法筛选棕漠土有机碳含量特征光谱。研究中基于预处理后光谱数据及有机碳含量数据,采用 MATLAB 软件结合 CARS、SPA、CARS-SPA 等算法进行特征波段筛选^[6,9]。

1.4.1 竞争性自适应重加权采样算法(CARS)

CARS 是一种结合蒙特卡洛采样与偏最小二乘回归(PLSR)的特征变量选择方法,其关键在于利用指数衰减函数(Exponentially decreasing function)和自适应重加权采样法(Adaptive reweighted sampling)两步骤对关键变量进行选择。研究中,设置蒙特卡洛采样次数为 100,通过对不同波段组合的反复迭代,利用蒙特卡洛交叉验证法筛选出交叉验证均方根误差($RMSE_{cv}$)最小的波段组合,即为最优波段组合^[6]。

1.4.2 连续投影算法(SPA) SPA 是一种矢量空间共线性最小化的前向变量选择算法,使用 SPA 算法可以压缩光谱数据的数量,通过不断的计算筛选出均方根误差($RMSE$)最小的最优波段组合作为最后结果输出,其在一定程度上可以消除光谱多重共线性和重叠干扰,从而提高模型的验证效率与精度^[14]。

1.4.3 CARS-SPA 算法 CARS-SPA 算法是利用 SPA 算法对 CARS 算法筛选后的特征波段进行二次筛选优化,可以有效减少模型输入的特征波段数目,减少数据冗余,提高模型的运行效率和精度。

1.5 模型构建及精度验证

随机森林是一种基于决策树的集成算法,其中每一棵决策树都是从建模集中有放回的随机取样构建的,最终估算结果由多数表决原则来决定,在进行拟合前,需要调节的参数为决策树数量^[18-19]。

分别以原始全波段、预处理后全波段及 CARS、SPA、CARS-SPA 3 种算法筛选的特征波段为自变量,棕漠土土壤有机碳含量为因变量,输入随机森林模型中进行估算,棕漠土有机碳含量估算模型构建流程如图 2 所示。研究中用决定系数(R^2)、均方根误差(RMSE)及相对分析误差(RPD)3 个指标进行模拟精度评价。其中,决定系数(R^2)和相对分析误差(RPD)越大,均方根误差(RMSE)越小^[19],说明模型较为稳定,估算性能较好。当 $RPD < 1.40$ 时,说明模型估算性能较差;当 $1.40 \leq RPD < 2.00$ 时,模型估算性能一般;当 $RPD \geq 2.00$ 时,模型模拟性能较好^[5]。

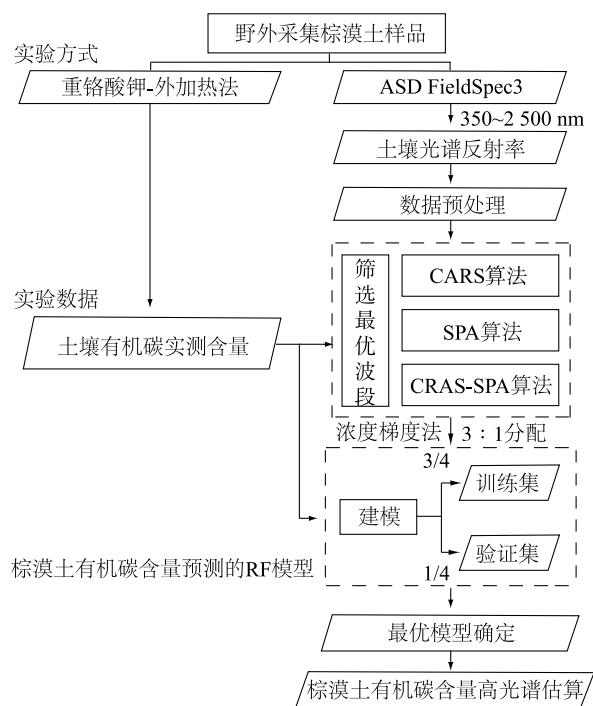


图2 棕漠土有机碳含量高光谱估算流程图

Fig.2 Hyperspectral estimation flow chart of organic carbon content in brown desert soil

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

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

$$RPD = \frac{SD}{RMSE} \quad (3)$$

式中: y_i 和 y_p 分别表示土壤有机碳含量的实测值和估算值, \bar{y} 表示土壤有机碳含量实测值的平均值, n 是样本数量, SD 是验证集实测值的标准偏差。

2 结果与分析

2.1 棕漠土有机碳含量描述统计特征

棕漠土不同样本集有机碳含量的统计特征如表 1 所示。从表 1 可以看出,研究区棕漠土 0~50.0 cm 土层的有机碳含量总体变化范围在 1.40~40.92 g/kg,平均值为 14.20 g/kg,标准差为 7.89 g/kg。建模集和验证集的平均值分别为 15.16 g/kg、11.49 g/kg,总样本集、建模集及验证集土壤有机碳含量变异系数分别为 55.54%、53.32%、59.89%,均处于中等变异水平,这表明研究区内棕漠土有机碳含量具有一定的空间变异性,数据比较离散,这一特征有利于监测技术的开展^[20-21]。

2.2 棕漠土有机碳含量估算特征波段筛选

SG 滤波及其一阶微分变换处理后的光谱曲线如图 3 所示。从图 3A 可知,SG 滤波后的棕漠土光谱曲线整体变化趋势基本一致,在 350~1 950 nm 波段,光谱反射率随波长升高而增强,在 2 200 nm 波段之后反射率逐渐减弱,在可见光波段光谱曲线上上升速度较快,在短波近红外及部分长波近红外波段上升相对缓慢。图 3B 是对 SG 滤波后的光谱曲线进行一阶微分处理,可有效减少噪声干扰。

全波段经过 CARS、SPA、CARS-SPA 3 种算法选择特征波段过程及结果如图 4 所示,从图 4 中可以看出不同算法筛选特征波段结果存在差异。在 CARS 算法筛选特征波段的过程中,随着运行次数的增加,筛选出的波段数量呈指数衰减趋势,即早期快速衰减,后期变化平缓,而 $RMSE_{cv}$ 波动较大,当运行次数为 40 时, $RMSE_{cv}$ 达到最小值 (3.47 g/kg) (图 4A)。由图 4B 可知,在 $RMSE_{cv}$ 最小时, CARS 算法筛选的特征波段为 122 个。利用 SPA 算法进行特征波段筛选时,随着变量(波段)数量的增加,交叉检验均方根误差总体呈减少趋势。当筛选出 11 个特征波段时, $RMSE_{cv}$ 值达到最小值 (2.63 g/kg) (图 4C),筛选出的特征波段分布如图 4D 所示。与 CARS 算法相比, SPA 算法更能有效地减少筛选出的特征波段数量,即 CARS 算法虽然能有效减少特征波段数量,但其筛选出的特征波段数量仍较多,这势必会影响下一步的估算效率。因此,在

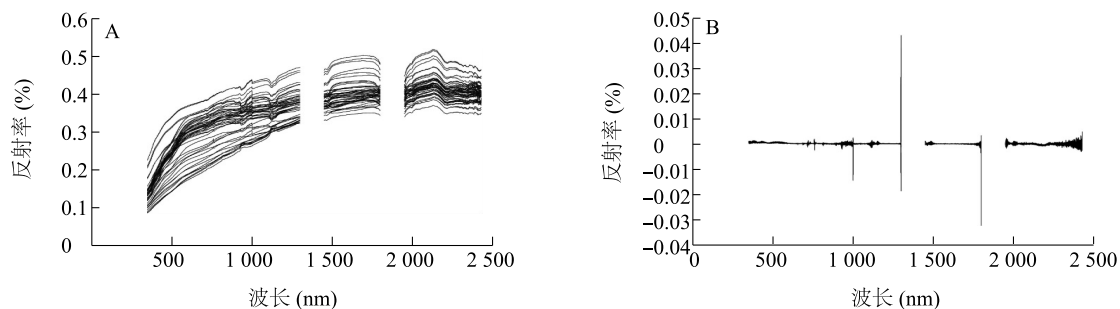
CARS 算法筛选出特征波段基础上,进一步利用 SPA 算法进行再次筛选。与 SPA 算法相似,随着特征波段数量的增加,CARS-SPA 算法的交叉检验均方根误差总体呈减少趋势,当筛选出 10 个特征波段时, $RMSE_{CV}$ 值达到最小值(2.96 g/kg)(图 4E)。筛

选出的特征波段分布如图 4F 所示。对比图 4D 和图 4F,可以看出,SPA 算法与 CARS-SPA 算法得到的特征波段数量虽然比较接近,但其对应的波段却有较大差异。

表 1 不同数据集棕漠土有机碳含量统计特征

Table 1 Statistical characteristics of organic carbon content of brown desert soil in different data sets

样本集	样本数	最小值 (g/kg)	最大值 (g/kg)	平均值 (g/kg)	标准差 (g/kg)	变异系数 (%)
总样本	50	1.40	40.92	14.20	7.89	55.54
建模集	37	1.40	40.92	15.16	8.08	53.32
验证集	13	3.81	23.29	11.49	6.88	59.89



A:SG 滤波后的棕漠土光谱曲线;B:一阶微分变换处理 SG 滤波后的棕漠土光谱曲线。

图 3 SG 平滑及其一阶微分变换处理后的棕漠土光谱反射率曲线

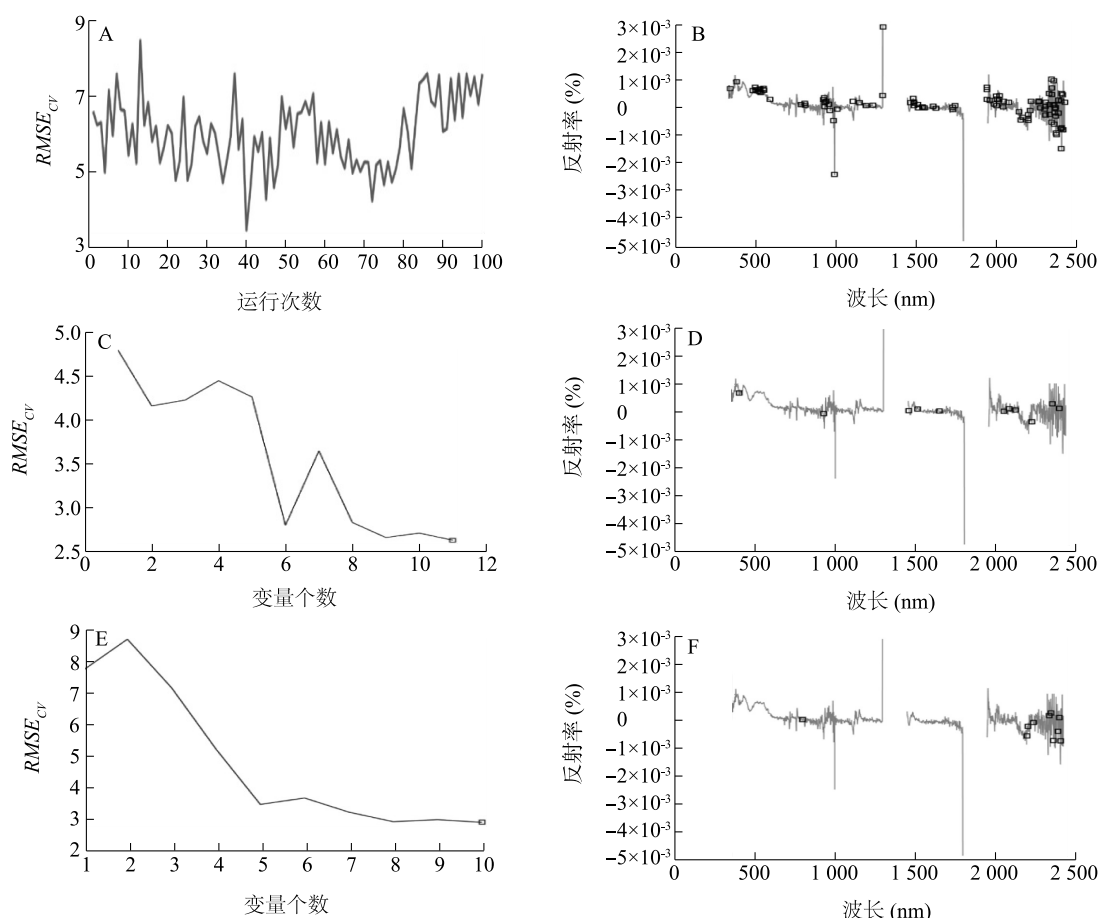
Fig.3 Brown desert soil spectral reflectance curves after SG smoothing and SG-first-order derivative pretreatment

2.3 模型建立与检验

基于不同数据集输入 RF 模型进行建模,得到的训练集回代检验结果和验证集检验结果分别如图 5 和图 6 所示。利用原始全波段数据建模,训练集回代检验的 R^2 和 $RMSE$ 分别为 0.59 和 5.06 g/kg(图 5A),而利用一阶微分变换预处理后的全波段光谱数据进行建模后,训练集回代检验的 R^2 和 $RMSE$ 分别为 0.79 和 3.50 g/kg(图 5B),模型精度有所提升。利用 CARS、SPA、CARS-SPA 3 种算法筛选的特征波段后,输入 RF 模型进行建模,训练集的回代检验结果如图 5C、图 5D、图 5E 所示。相比于预处理后的全波段数据建模, R^2 分别提高了 0.11、0.10、0.11, $RMSE$ 分别降低了 1.01 g/kg、0.87 g/kg、1.01 g/kg。

利用原始全波段光谱数据建模,验证集检验的 R^2 、 $RMSE$ 和 RPD 分别为 0.45、5.51 g/kg 和 1.35(图 6A)。由于 RPD 小于 1.4,故构建的模型对棕漠土有机碳含量估算效果较差,棕漠土有机碳含量

的实测值和估算值在 1:1 线附近较为分散。利用一阶微分变换后的光谱数据进行建模后,验证集检验的 R^2 、 $RMSE$ 和 RPD 分别为 0.63、4.84 g/kg 和 1.65(图 6B)。这说明光谱经过低阶微分处理后不仅能降低噪声干扰更能凸显光谱特征,从而提升模型估算精度^[17]。利用 CARS 算法筛选得到的特征波段数据进行建模,验证集检验的 R^2 、 $RMSE$ 和 RPD 分别为 0.75、3.53 g/kg 和 1.99(图 6C)。由于 RPD 小于 2.00,说明模型对研究区棕漠土有机碳含量的估算效果一般。而基于 SPA 算法和 CARS-SPA 算法筛选得到的特征光谱数据建模,验证集检验的 R^2 分别为 0.82 和 0.85, $RMSE$ 分别为 3.07 g/kg 和 2.72 g/kg, RPD 分别为 2.34 和 2.59(图 6D,图 6E)。由于通过这 2 种特征波段筛选算法, RPD 均大于 2,表明这 2 种算法筛选的波段输入 RF 模型均能较好地估算棕漠土有机碳含量。对比两者的检验指标,利用 CARS-SPA 算法筛选的特征波段构建 RF 模型的效果最优。



A: CARS 算法筛选特征波段; B: CARS 算法筛选的特征波段分布; C: SPA 算法筛选特征波段; D: SPA 算法筛选的特征波段分布; E: CARS-SPA 算法筛选特征波段; F: CARS-SPA 算法筛选的特征波段分布。

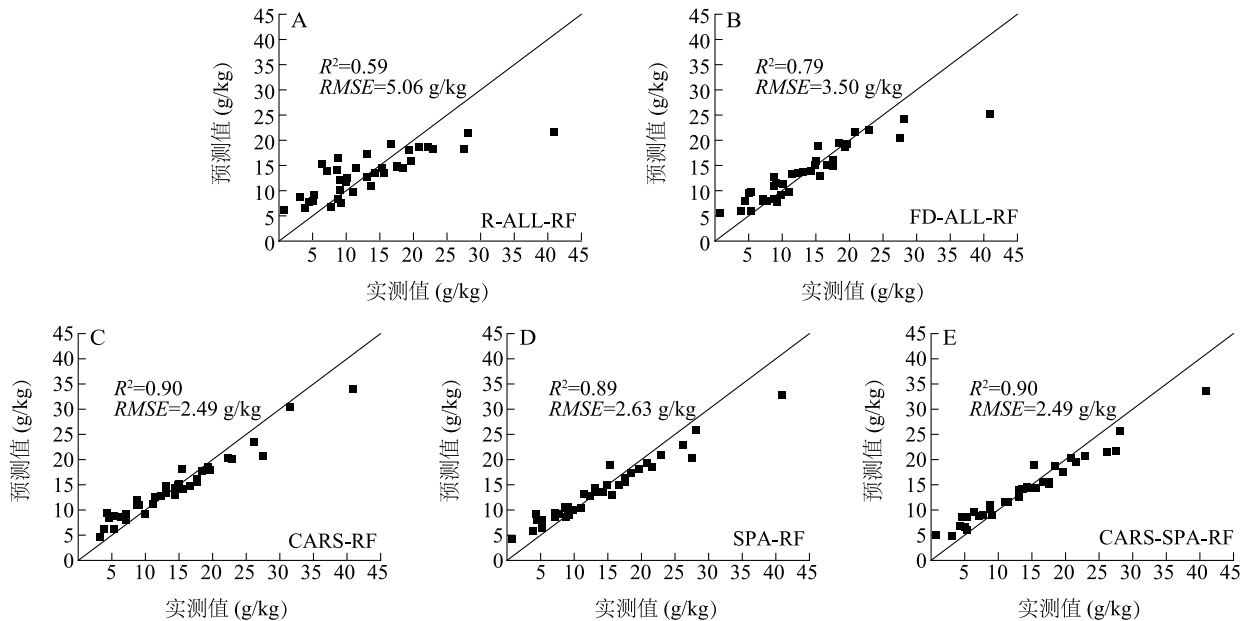
图 4 不同算法筛选特征波段过程及结果

Fig.4 The process and results of different algorithms for selecting feature bands

3 讨论

目前已有研究表明进行室内土壤光谱测定可以降低周围环境对光谱数据影响,但不能模拟采样的室外自然环境,存在一定局限性^[22]。本研究选择在室外进行棕漠土光谱测定,可以更大程度模拟野外采样的自然环境,实现对光谱数据的有效测定^[23]。光谱数据经过低阶微分处理以后,可有效降低噪声的干扰^[17]。本研究利用全波段原始光谱数据建模后,验证集检验的决定系数 R^2 和 RPD 分别为 0.45 和 1.35,而利用一阶微分预处理后的全波段光谱数据后,分别提升为 0.63 和 1.65,模型的估算精度有了很大提升,这与张娟娟等^[20]、牛芳鹏等^[14]研究结果基本一致。棕漠土的光谱曲线除水分吸收峰及噪声影响外,整体趋势较为平缓,这与砂姜黑土^[20]及砂壤潮土^[7]的光谱曲线趋势大致相同。

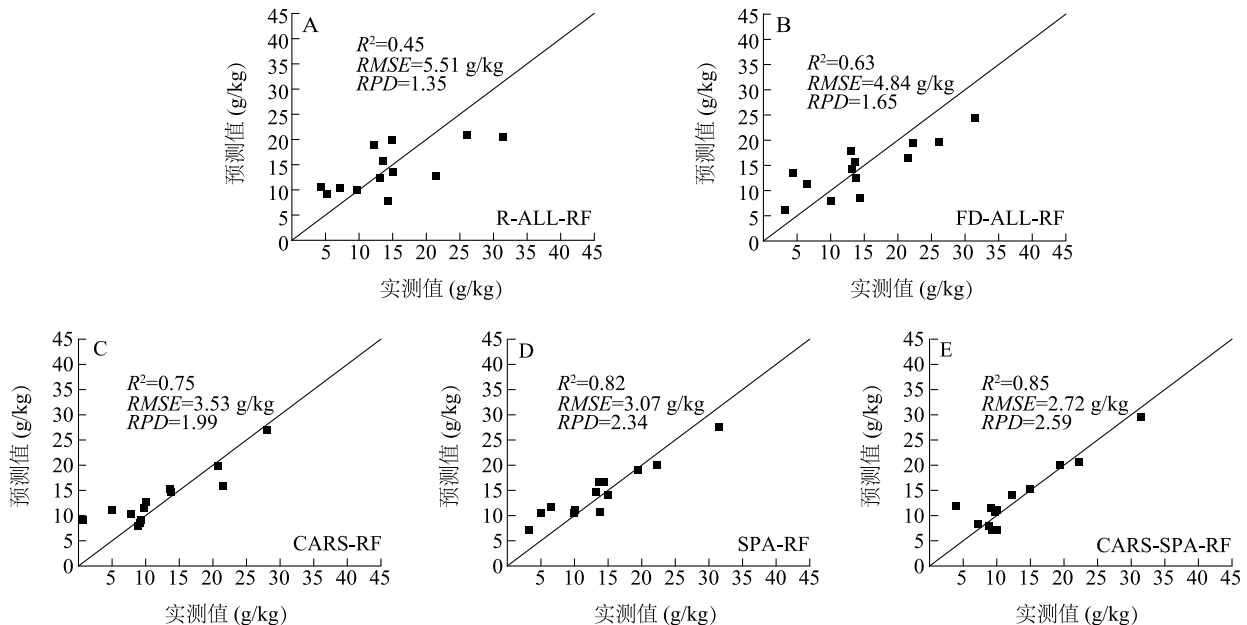
CARS、SPA、CARS-SPA 3 种算法筛选出的棕漠土特征波段数分别是 122 个、11 个、10 个,即上述算法均能有效地降低有机碳估算模型的输入波段数目,提高模型的运行速率,这一结论与钟翔君等^[7]、牛芳鹏等^[14]、唐海涛等^[9]研究结果基本一致。通过不同筛选算法筛选出的特征波段存在一定的差异,即不同算法筛选出的特征波段具有一定的不确定性。基于 CARS 算法筛选出的特征波段分布较为分散,在 1 100~2 500 nm,主要是受到羰基、酰胺和羟基等基团分子振动的倍频与合频吸收影响^[9];1 000 nm 以下存在少量的特征波段,这可能是因为 CARS 采用交叉验证筛选出的特征波段为 $RMSE_{cv}$ 较小的最优波段集而非相关性高的波段集。利用 SPA 对 CARS 筛选出的特征波段再次筛选后,筛选出的特征波段主要集中在近红外波段 760~2 500 nm,这与钟翔君等^[7]为预测砂壤潮土有机质含量,利用 CARS-SPA



R-ALL-RF:原始全波段数据输入随机森林模型;FD-ALL-RF:原始全波段数据经一阶微分变换预处理后输入随机森林模型;CARS-RF:竞争性自适应重加权采样算法筛选后的特征波段数据输入随机森林模型;SPA-RF:连续投影算法筛选后的特征波段数据输入随机森林模型;CARS-SPA-RF:竞争性自适应重加权采样算法筛选后的特征波段数据再经连续投影算法筛选后输入随机森林模型。

图5 基于RF模型的训练集回代检验结果

Fig.5 The test results of training set based on RF model



R-ALL-RF:原始全波段数据输入随机森林模型;FD-ALL-RF:原始全波段数据经一阶微分变换预处理后输入随机森林模型;CARS-RF:竞争性自适应重加权采样算法筛选后的特征波段数据输入随机森林模型;SPA-RF:连续投影算法筛选后的特征波段数据输入随机森林模型;CARS-SPA-RF:竞争性自适应重加权采样算法筛选后的特征波段数据再经连续投影算法筛选后输入随机森林模型。

图6 基于RF模型的验证集检验结果

Fig.6 Verification results of validation set based on RF model

算法筛选出的特征波段分布结果有所差异,这可能是土壤类型不同导致的。现有研究表明,使用

非线性模型能取得更好的估算效果^[6,9],本研究利用CARS-SPA算法筛选的特征波段结合非线性的随

机森林模型,验证集检验的 R^2 、 RPD 分别达到 0.85 和 2.59,取得了较好的估算效果,即本研究筛选的特征波段及模型构建方法适用于基于高光谱的湖滨绿洲棕漠土有机碳含量估算。

4 结 论

(1) 博斯腾湖湖滨绿洲棕漠土 0~50.0 cm 土层有机碳含量变化范围为 1.40~40.92 g/kg,平均值为 14.20 g/kg,标准差 7.89 g/kg,变异系数为 55.54%,整体呈现中等变异水平。

(2) CARS、SPA、CARS-SPA 3 种算法筛选出的特征波段数及分布有较大差异,基于 CARS-SPA 算法筛选出的特征波段主要集中在近红外波段 760~2 500 nm,且数目更少。

(3) 基于 CARS-SPA 算法筛选出的特征波段构建的 RF 模型在估算棕漠土有机碳含量时效果最好,验证集 R^2 为 0.85, $RMSE$ 为 2.72 g/kg, RPD 为 2.59。即利用该算法筛选特征变量可以减少冗余波段数据的干扰,提高模型估算精度和运行效率,实现研究区棕漠土有机碳含量的有效估算。

参考文献:

- [1] 胡莹洁,孔祥斌,姚静韬.北京市平原区土壤有机碳垂直分布特征[J].生态学报,2019,39(2):561-570.
- [2] BURAKA T, ELIAS E, LELAGO A. Soil organic carbon and its' stock potential in different land-use types along slope position in Coka watershed, Southern Ethiopia[J]. Heliyon, 2022, 8(8): e10261.
- [3] 周倩倩,丁建丽,黄 帅. 干旱区绿洲土壤有机碳高光谱估测及其影响因子分析[J].干旱地区农业研究,2018,36(5):200-206,214.
- [4] 贡 璐,朱美玲,刘曾媛,等. 塔里木盆地南缘典型绿洲土壤有机碳、无机碳与环境因子的相关性[J].环境科学,2016,37(4):1516-1522.
- [5] 赵明松,谢 毅,陆龙妹,等. 基于高光谱特征指数的土壤有机质含量建模[J].土壤学报,2021,58(1):42-54.
- [6] 吴 俊,郭大千,李 果,等. 基于 CARS-BPNN 的江西省土壤有机碳含量高光谱预测[J].中国农业科学,2022,55(19):3738-3750.
- [7] 钟翔君,杨 丽,张东兴,等. 砂壤潮土有机质含量可见-近红外光谱预测[J].光谱学与光谱分析,2022,42(9):2924-2930.
- [8] LAAMRANI A, BERG A A, VORONEY P, et al. Ensemble identification of spectral bands related to soil organic carbon levels over an agricultural field in southern Ontario, Canada[J]. Remote Sensing, 2019, 11(11):1298.
- [9] 唐海涛,孟祥添,苏循新,等. 基于 CARS 算法的不同类型土壤有机质高光谱预测[J].农业工程学报,2021,37(2):105-113.
- [10] 章海亮,罗 微,刘雪梅,等. 应用遗传算法结合连续投影算法近红外光谱检测土壤有机质研究[J].光谱学与光谱分析,2017,37(2):584-587.
- [11] WANG S, CHEN Y, WANG M, et al. SPA-based methods for the quantitative estimation of the soil salt content in saline-alkali land from field spectroscopy data: A case study from the Yellow River irrigation regions[J]. Remote Sensing, 2019,11(8):967.
- [12] LIU J B, DONG Z Y, XIA J S, et al. Estimation of soil organic matter content based on CARS algorithm coupled with random forest[J]. Spectrochimica Acta, Part A: Molecular and Biomolecular Spectroscopy,2021,258: 119823.
- [13] CHENG H, WANG J, DU Y. Combining multivariate method and spectral variable selection for soil total nitrogen estimation by Vis-NIR spectroscopy[J]. Archives of Agronomy and Soil Science, 2021, 67(12): 1665-1678.
- [14] 牛芳鹏,李新国,麦麦提吐尔逊·艾则孜,等. 基于连续投影算法的博斯腾湖西岸湖滨绿洲土壤有机碳含量的高光谱估算[J].浙江大学学报(农业与生命科学版),2021,47(5):673-682.
- [15] 赵 慧,李新国,靳万贵,等. 基于地理加权回归模型的博斯腾湖湖滨绿洲土壤盐分离子含量高光谱估算[J].土壤,2021,53(3):646-653.
- [16] 肖云飞,高小红,李冠稳. 土壤有机质可见光-近红外光谱预测样本优化选择[J].土壤,2020,52(2):404-413.
- [17] CLOUTIS E A. Hyperspectral geological remote sensing: evaluation of analytical techniques[J]. International Journal of Remote Sensing,1996,17(12): 2215-2242.
- [18] DING J, YANG A, WANG J, et al. Machine-learning-based quantitative estimation of soil organic carbon content by VIS/NIR spectroscopy[J]. Peer J, 2018, 6: e5714.
- [19] 包青岭,丁建丽,王敬哲,等. 基于随机森林算法的土壤有机质含量高光谱检测[J].干旱区地理,2019,42(6):1404-1414.
- [20] 张娟娟,席 磊,杨向阳,等. 砂姜黑土有机质含量高光谱估测模型构建[J].农业工程学报,2020,36(17):135-141.
- [21] XU S, ZHAO Y, WANG M, et al. Comparison of multivariate methods for estimating selected soil properties from intact soil cores of paddy fields by VIS-NIR spectroscopy[J]. Geoderma, 2018, 310: 29-43.
- [22] 刘焕军,张新乐,郑树峰,等. 黑土有机质含量野外高光谱预测模型[J].光谱学与光谱分析,2010,30(12):3355-3358.
- [23] 马利芳,熊黑钢,张 芳. 基于野外 VIS-NIR 光谱的土壤盐分主要离子预测[J].土壤,2020,52(1):188-194.

(责任编辑:石春林)