

土壤修复过程中盐含量及其光谱特征分析研究

朱 赟¹, 申广荣^{1, 2*}, 王紫君¹, 陆邵明³, 支月娥², 项巧巧¹

1. 上海交通大学农业与生物学院, 低碳农业研究中心, 上海 200240
2. 农业部都市农业(南方)重点实验室, 上海 200240
3. 上海交通大学船舶海洋与建筑工程学院, 上海 200240

摘 要 基于盐渍土修复过程中盐分含量和同步实测光谱数据, 通过对原始光谱数据、平滑光谱数据及平滑后的不同变换光谱数据等八种光谱数据集, 分别以相关系数的极值和不同相关系数范围两种方法分析其最佳敏感波段范围, 深入分析了不同变换下土壤的光谱响应特征。在此基础上, 运用偏最小二乘回归方法, 以全波段(400~1 650 nm)和分析获得的最佳敏感波段建立了基于修复过程的土壤盐含量和光谱反射率的关系模型。结果表明: 针对八种光谱数据集, 采用两种方法提取的土壤最佳敏感波段, 均集中在 947.11~949.31, 1 340.27, 1 394.11, 1 419, 1 457.81~1 461.31, 1 537.68~1 551.39 和 1 602.32 nm; 且最佳波段的土壤盐含量反演模型, 以模型评价参数的决定系数(R^2)和均方根误差(RMSE), 以及赤池信息量准则(akaike's information criterion, AIC)作为选择最佳模型的标准, 均以 SGSD(Log R)模型的建模和预测结果比其他光谱变换的模型更为显著。基于全波段的 PLSR 建模效果总体上稍优于最佳波段的模型, 其中以 SGSD 的预测精度最为突出, 其模型的决定系数 R^2 与标准差 RMSEP 分别为 0.673 和 1.256; 基于两种方法获得的最佳波段的 PLSR 模型与全波段对比在模型精度方面虽有一定差距, 但从模型的复杂程度比较, 具有模型简单、变量更少及运算量小的特点。该研究可在土壤盐含量及其光谱特征的研究中, 为实现土壤盐渍化定量、快速、便捷的监测和检测提供参考。

关键词 盐渍化土壤; 微生物修复; 光谱变换; 偏最小二乘法

中图分类号: TP7 文献标识码: A DOI: 10.3964/j.issn.1000-0593(2017)05-1507-07

引 言

土壤盐渍化是一定的气候、地形、水文地质等自然条件共同对水盐运动产生影响的结果。土壤光谱响应模式与土壤物理性质、土壤水分、盐分、有机质含量及肥力等有着密切关系, 即土壤的化学成分是决定其反射光谱特征的重要因素。研究土壤盐渍化这一复杂的动力学过程中其盐分含量及其光谱特征的变化对于应用高光谱技术快速、便捷有效地诊断、监测土壤盐渍化具有重要的意义。

近年来随着遥感技术的发展, 土壤盐渍化的高光谱特征及其定量反演研究成为人们关注的热点并取得了重要的成果。不同化学成分的盐, 如含水氯化钠、硫酸钠、硫酸钾、硫酸钙、硫酸镁等会引起吸收峰位置微妙且较稳定的变化^[1]。

进一步分析表明, 以氯化钠为主导性盐分的盐渍土的敏感波段范围为 1 370~1 445 和 1 447~1 608 nm^[2]; 盐渍化程度的不同其光谱反射吸收特征也不相同, 与正常土壤相比, 重度盐渍土在 680, 1 180 及 1 780 nm 处具有明显的特征, 而在 2 200 nm 处, 土壤的盐度越高吸收就越弱, 并且光谱曲线在 1 400 和 1 900 nm 左右具有较宽且左右两边不对称的吸收谷^[3]。非盐渍化土壤的特征波段为 617 nm, 中度盐渍化土壤的特征波段为 751 nm^[4]。在数据分析和建模方法方面, 目前常用的如逐步线性回归(simple Linear Regression, SMLR)、主成分分析(principal component analysis, PCA)和偏最小二乘回归(partial least squares regression, PLSR)等方法。不同的研究其侧重点和方法手段各异, 所建的土壤盐渍化光谱反演模型也都不同。获得满足实际需要, 且排除水分、有机质、土壤类型等干扰、具有较高精度的高光谱盐分反演模型仍是

收稿日期: 2015-12-28, 修订日期: 2016-04-08

基金项目: 高分国土资源遥感应用示范系统(一期)项目(04-Y30B01-9001-12/15), 国家自然科学基金项目(41471120), 社科重大项目(14ZDB139)和上海交大农工交叉项目(Agri-X2015004)资助

作者简介: 朱 赟, 女, 1991 年生, 上海交通大学农业与生物学院硕士研究生 e-mail: zy_1125@sjtu.edu.cn

* 通讯联系人 e-mail: sgrong@sjtu.edu.cn

人们探索的目标^[5]。为此,关注土壤盐渍化的整个过程,研究分析盐分含量的渐变及其光谱特征的变化是一个有效途径^[2,6],而最佳光谱变换形式的选择及光谱特征与敏感波段分析是建立盐渍化土壤光谱反演模型的基础和重要内容。

基于以上分析,本研究跟踪盐渍化土壤修复的整个过程,通过对原始光谱数据、平滑光谱数据及平滑后的不同变换光谱数据等八种光谱数据集,分别以两种不同方法分析其最佳敏感波段范围,分析土壤盐分含量渐降及其同步的光谱响应特征。在此基础上,利用偏最小二乘回归法以全波段与最佳响应波段两种建模方式建立土壤盐含量的光谱反演模型。

1 实验部分

1.1 设计及材料

以山东东营盐渍土为样品,采用已获专利(CN102676492A)并在次生盐渍化土壤的微生物修复中有较好应用效果的微生物菌种(TNC 冈贮;复合菌)进行修复实验。供试土壤采自山东省北部的东营市(北纬 36°55′-38°10′,东经 118°07′-119°10′),其盐渍化土壤为砂质壤土,黏粒含量为 10%~15%,土壤容重为 1.40 g·cm⁻³,在该研究区的土

壤盐分中,阴离子以 Cl⁻ 为主,占阴离子总量的 57.889%,SO₄²⁻ 次之,其他的阴离子含量很少;阳离子以 Na⁺ 为主,占阳离子总量的 28.823%,余下的依次为 Ca²⁺,Mg²⁺,K⁺,因此,盐分含量主要以氯化钠为主(表 1)。

修复实验在上海交通大学浦江基地 7 号温室内进行。供试土壤从研究区域采样后,未经任何处理,与肥料混合均匀后盛装在 PVC 框(61 cm×39.5 cm×35 cm)中置于温室内,各处理的填土方式相同。PVC 框底部留有出口,与管道连通并装有控制阀门,用于收集实验中的淋盐(或淋溶水)。混合时肥料按照每 100 kg 土壤施用 5,15,30 kg 的用量加入。玻璃温室内设有智能环境控制设备,温度维持在 40℃左右,可为微生物的生理活动提供合适的温度及相关环境条件。其中每 10 天浇一次水,每次采样时先采样再浇水,每次浇水浇透即可,通过 PVC 框底部管道和阀门确保 PVC 框中无积水,无水流出,以保证实验中水量仅供微生物新陈代谢。实验设 6 个处理(对照、有机肥、有机肥+EM 菌剂、有机肥+JFB 菌剂、菇科菌肥、菇科菌肥+EM 菌剂和菇科菌肥+JFB 菌剂),每个处理包括 3 个肥料施用量(5,15,30 kg),每个梯度 3 个重复。

表 1 供试土壤的基本理化性质/(g·kg⁻¹)

Table 1 The physico-chemical characteristics of the tested soil/(g·kg⁻¹)

指标	pH	有机质	总氮 TN	Ca ²⁺	Mg ²⁺	Na ⁺	K ⁺	Cl ⁻	HCO ₃ ⁻	SO ₄ ²⁻
实测值	7.850	2.210	0.258	0.373	0.478	0.656	0.049	3.118	0.335	0.576

1.2 数据获取

1.2.1 土壤样本的采集与理化分析

土壤修复实验历经 70 天,分别在修复的第 10 天(2012 年 12 月 8 日)、30 天(2012 年 12 月 28 日)、50 天(2013 年 1 月 17 日)及 70 天(2013 年 2 月 6 日)进行土壤样品采样分析。采用棋盘式土壤采样法,采集各个样点 0~20 cm 的表层土壤,每次采集 54 个土壤样本,四次采集共 216 个土样。带回实验室后自然风干,经简单的处理后(除去一些杂草和石头,手工磨碎),过 100 目筛(0.15 mm),装入编号的密封袋中,分为两份,一份用来测定土壤的理化指标,另一份用来测定土壤的光谱。盐分含量分析时按 $m(\text{土}):m(\text{水})=1:5$ 制备待测液,浸提后,按照《土壤农化分析》中分析方法分别测定 pH、电导率和可溶性盐离子(Na⁺,K⁺,Mg²⁺,Ca²⁺,Cl⁻和 SO₄²⁻,CO₃²⁻)含量。为尽可能排除水分因素的影响,以保证盐分含量对盐渍土光谱反射率影响的单一性及可靠性^[5,7],土壤样本在经风干处理后(含水量均控制在 10%以下)进行光谱测试。

1.2.2 土壤样本的光谱数据测量

实验室光谱测量在暗室内进行,使用荷兰 Avantes 公司生产的 AvaField-2 便携式高光谱地物波谱仪(野外光谱辐射仪),该光谱仪光谱响应范围 350~1700 nm,采样间隔在 350~1040 nm 范围为 0.59 nm,在 1040~1700 nm 范围为 3.8 nm。将研磨后的土样置于直径 10 cm、深 2 cm 的培养皿

中,装满后,采用直尺将其表面刮平。采用光谱仪配套的功率为 1000 W 的卤素灯作为光谱测量的光源,调节光照角度使得灯光最亮处位于样品中央,探头视角角为 5°,探头到土壤表面距离为 15 cm。每个土壤样品测量前均进行白板校正,各土样分别采集 5 条光谱曲线,取其平均值为该土壤样品的反射率。

1.3 方法

1.3.1 光谱数据的平滑去噪

消除光谱数据无关信息和噪声的预处理是分析光谱数据的基础。本研究首先对原始光谱曲线进行 FFT 滤波器平滑、Savitzky-Golay 卷积平滑法、移动平均法、中值滤波器平滑、高斯滤波器平滑^[8]等五种不同的平滑去噪处理。对比原始光谱曲线,分析比较每种平滑处理对消除土样高光谱数据噪声的影响,以能保持光谱曲线总体趋势且光谱信息丢失最少为原则进行平滑方法的筛选。具体通过对每一种平滑变换下全波段模型与原始光谱全波段模型的均方根误差(RMSE)进行比较,选取 RMSE 最小的平滑方法。

1.3.2 光谱变换分析

光谱变换是盐渍土壤光谱特征分析的重要内容。本研究尝试了 31 种光谱变换对光谱平滑后的数据进行分析,选取了各个变换的全波段建模结果较理想的前 6 种变换用于进一步的分析。首先对原光谱数据及每一种变换分析后的光谱数据建立全波段 PLSR 预测模型,然后,将每一种变换对应的

PLSR 模型精度与原光谱数据对应的模型精度进行比较,以模型的相关系数与决定系数显著提高,标准偏差与标准差变为评判基准^[8],对各变换后所建模型精度的改善效果从高到低排列,挑选前 20% 的变换进行下一步处理。本文选择的变换包括一阶微分的差分求导(FD)、一阶微分的 SG 卷积求导(SGFD)、二阶微分的 SG 卷积求导(SGSD)、对数的二阶的 SG 卷积求导[SGSD(LogR)]、对数倒数的一阶差分求导[FD(1/Log R)]和倒数对数的二阶 SG 卷积求导[SGSD(Log(1/R))]^[9]。

1.3.3 最佳响应波段分析

用偏最小二乘法建立盐渍土含盐量的预测模型时,全波段参加建模将有 1 336 个波段的光谱反射率作为自变量,最佳响应波段的分析是建立有效土壤盐含量光谱反演模型的基础。

针对原始光谱数据、平滑光谱数据及平滑后的六种变换光谱数据即共八种不同形式的光谱数据,基于土壤盐含量与光谱反射的相关系数,在最佳波段的分析上尝试了两种不同的方法。第一种方法为相关系数极值提取法^[10],选定盐含量与光谱反射率之间相关系数曲线的波峰点(系数最高)的多个波段为最佳波段;第二种方法是根据盐含量与光谱反射率之间相关系数的范围^[10],分别对不同相关系数范围($|r| \geq 0.5$, ≥ 0.6 , ≥ 0.7)所对应的波段建立偏最小二乘模型,根据模型精度比较,即三个模型中拥有最高的相关系数与决定系数,最低的标准偏差与标准差的模型,选取其对应的相关系数范围下的波段。进一步,对以上两种方法挑选出的敏感波段,再分别利用偏最小二乘回归处理,由回归系数大小确定每种方法分别对应的土壤盐含量光谱敏感波段^[5],用于参与最终的盐含量光谱反演模型的构建。偏最小二乘回归分析在软件 Unscrambler 9.7 中完成。

1.3.4 关系模型的建立

首先对土壤样本进行预处理,即建模集中异常样本的剔除和建模集样本的选取。所谓异常样本,是指浓度标准值或光谱数据存在较大误差的样品。本研究将去除异常样本后的 62 个样本分为建模样本 41 个(2/3)和校正样本 21 个(1/3)^[8]。建模样本的选取采用含量梯度法,即将样本集中的样品按盐分含量值范围分为 7 个子范围,然后统计落入各个子范围内的样品数,最后在这些子范围内根据落入样品数的多少随机选取 41 个本来作建模样本,剩余的 21 个样本作为校正样本。

偏最小二乘回归(PLSR)是将因子分析和回归分析相结合的方法,它不仅仅考虑因变量与自变量集合的回归建模,而是采用成分提取的方法,在变量系统中提取对系统有最佳解释能力的新综合变量,再对它们进行回归建模^[11]。本工作采用 PLSR 的建模方法对盐渍土中的盐含量进行预测,分别以全波段的反射率和两种方法选取的最佳波段的反射率为自变量建立土壤盐渍化光谱反演模型,建模在 Unscrambler9.7 软件中完成,建模的过程中采用交叉验证(Full Cross Validation),模型因子数直接采用软件建议的个数。

对模型精度和质量的分析,通过计算比较模型的相关系数(r)、决定系数(R^2)、建模样本的标准偏差(SEC)、标准差

(RMSEC)和校正样本的标准偏差(SEP)、标准差(RMSEP)以及赤池信息量准则(akaike's information criterion, AIC)来评价并判断选择最优模型^[8, 12]。

2 结果与讨论

2.1 修复过程中土壤盐分含量及其光谱反射率的变化

本研究采用的盐渍土属于极重度盐化土,其盐含量为 1.55%。盐渍土 70 天的修复过程中,盐分含量及其光谱反射率的变化如图 1 所示。其中,SSC 盐含量为每个修复时间段内所有土壤样品的均值,反射率为全波段光谱反射率的均值。盐含量的变化呈先大幅降低然后回升的趋势。相应地,同步测量的土壤光谱反射率呈基本相似的变化趋势,特别是在修复的 30 天到 50 天。

整个修复过程中,土壤盐含量从修复开始的 1.55% 降到修复 10 天的 1.34%,30 天的 0.89% 和 50 天的 0.85%。由于“返盐”现象,在修复的第 70 天,土壤盐含量上升为 1.34%。土壤修复过程中,土壤盐含量变化曲线的斜率在第一阶段(修复开始到 10 天),第二阶段(修复 10 天到 30 天)和第三阶段(修复 30 到 50 天)以及第四阶段(修复 50 天到 70 天)分别是 -0.041, -0.012, -0.002 和 0.024,相应的光谱曲线斜率分别是 -0.005, -0.003, -0.002 和 0.007。在修复 50 天后,由于土壤中微生物菌剂活性减弱和微生物数量及种类的减少,不能产生大量的多糖和粘胶来改善土壤团粒结构,土壤中水分通过毛细管上升,盐分离子向上运动并在地表累积导致土壤修复 70 天时“返盐”现象,相应的光谱反射率的变化表明,光谱反射率对于 SSC 的变化是敏感的,在土壤修复过程中其光谱反射率对盐渍土盐含量变化具有较好的指示作用。

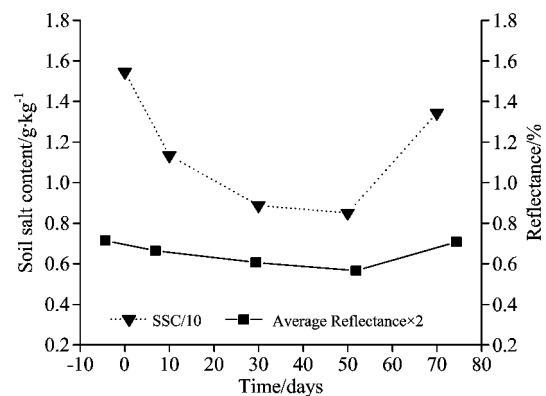


图 1 土壤盐含量与光谱反射率的发展趋势

Fig 1 The trends of soil salt content and soil reflectance indicators

2.2 土壤光谱曲线的平滑去噪和特征分析

对于修复过程中 10 天,30 天,50 天,70 天的土壤样品光谱数据分别进行的平滑处理表明,FFT 滤波器平滑效果较好,其平滑处理后的全波段模型的 RMSE 在所有平滑方法中最小,为 1.392。

分析盐渍土光谱特征(见图 2),所有土样的光谱曲线在

形态上几乎一致, 每条光谱曲线基本平行。在 400~760 nm 的可见光波段范围内, 土壤反射率呈现明显的上升趋势, 在 760~1 650 nm 的近红外波段范围内, 土壤光谱反射率增长

趋于平缓, 在 1 400 nm 附近有明显的水分吸收谷, 通常认为 1 400 nm 附近为羟基(-OH)带谱^[13]。

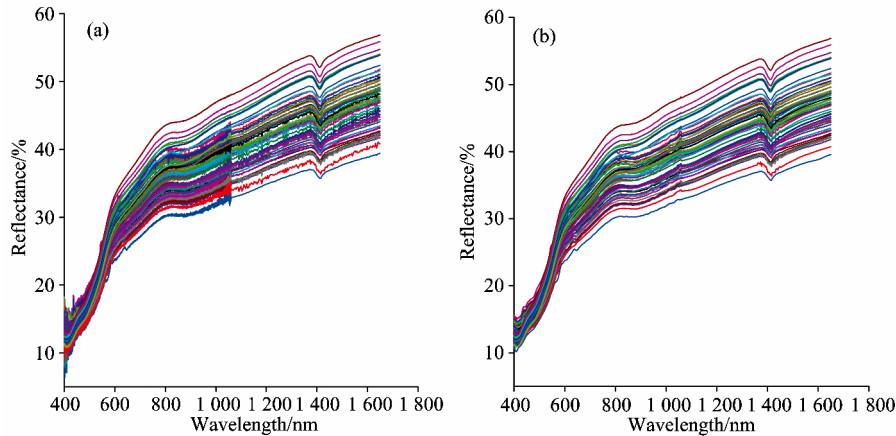


图 2 (a)原始光谱曲线, (b)平滑后的光谱曲线

Fig 2 (a) the original spectral curve, (b) the spectral curve after smoothing

对比分析土壤修复过程中不同时间段即不同盐分含量的光谱反射率变化情况。图 3 是土壤修复过程中不同时间段土壤样本的光谱反射率平均值在全波段的光谱响应曲线。在土壤修复的过程中, 土样反射率随着盐分的减少而降低, 由于土壤修复 70 天的“返盐”现象, 其土样的光谱反射率也相应上升。另外, 修复过程中不同盐含量的盐渍土光谱曲线形状很相似, 曲线斜率变化比较平缓且基本平行; 盐渍土除

1 411. 90 nm 处的吸收谷在反射率光谱曲线上能非常清晰识别外, 其他吸收特征宽且弱, 吸收谷不明显, 难以在原始光谱曲线上找出吸收谷的具体位置。对反射率光谱的连续统去除可以增加细小而较弱的吸收特征, 由图 3(b)可见, 盐渍土光谱具有 3 个不同深度的吸收谷, 分别为 485. 83、907. 92 和 1 411. 90 nm。

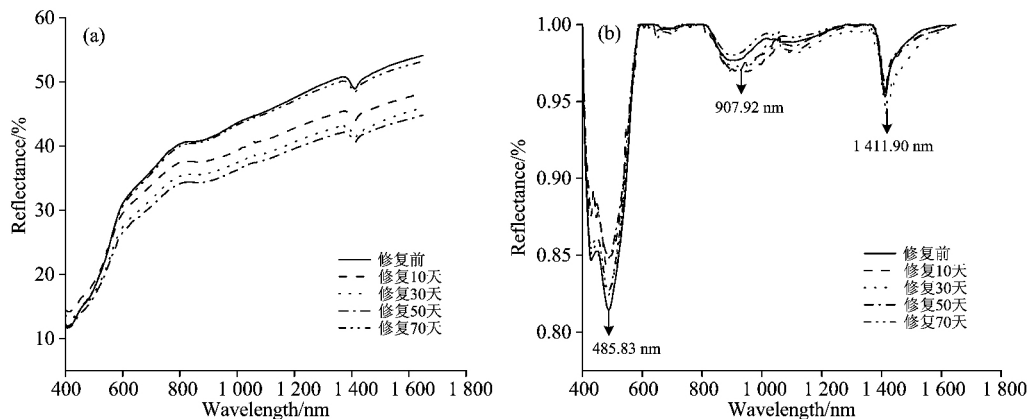


图 3 (a)各组土样光谱反射率曲线, (b)各组土样光谱连续统去除后反射率曲线

Fig 3 (a) the reflectance spectra of soil samples in each group, (b) the continuum removal reflectance spectra of soil samples in each group

2.3 最佳敏感波段选取

2.3.1 极值提取最佳波段范围

基于相关系数极值^[10]提取的最佳波段见表 2, 除 FD(1/Log R)外, 大部分光谱数据集的敏感波段数为 2~5 个。

2.3.2 不同相关系数范围提取最佳波段

采用不同相关系数范围选取土壤盐分最佳敏感波段分三个步骤: 首先, 对选定的六种光谱变换及原光谱和平滑后光谱共 8 个数据集, 分别通过建模选择确定各数据集最佳敏感

波段及其对应的相关系数范围($|r| \geq 0.5$, ≥ 0.6 和 ≥ 0.7 中的一个)。以光谱数据的 SGSD[Log(1/R)]为例, 在建模集中, R^2 值在 $|r| \geq 0.5$, ≥ 0.6 和 ≥ 0.7 时, 分别为 0.759, 0.460, 0.597, 相对应的 SEC 和 RMSEC 分别为 0.946 和 0.935, 1.416 和 1.399, 1.224 和 1.209; 在校正集中, R^2 值相应为 0.681, 0.484, 0.625, SEP 和 RMSEP 分别为 1.272 和 1.241, 1.601 和 1.577, 1.359 和 1.345, 综合各模型精度参数, SGSD(LogR)的最佳波段在 $|r| \geq 0.5$ 的波段范围内选

取。其次，鉴于不同光谱变换后的光谱数据集第一步选取的初级敏感光谱波段的个数均较多，因而进行进一步处理，即将相邻的波段合并为一个波段，其反射率值为以此波段范围内光谱反射率的平均值。第三，对合并处理后的敏感波段的反射率及其盐分含量进行 PLSR 回归处理，选取对因变量盐分含量影响显著的波段为最终敏感波段。表 3 是基于不同相关系数范围方法获得的不同光谱数据集最终敏感波段。除了

R 与 $SGSD[\text{Log}(1/R)]$ 数据集，其余的光谱变换形式下所选取的最佳波段个数均在 3~8 个之间。另外，每个光谱数据集变换形式的最佳敏感波段的选取都是在不同相关系数水平下，如 R 为 0.5, FFT 为 0.6 等(表 3)，这也说明不同光谱变换对分析研究土壤盐分含量的最佳响应波段的重要意义和必要性。

表 2 极值提取最佳波段范围

Table 2 SSC optimal bands by extremums of the correlation coefficient

光谱变换形式	最佳波段/nm	自变量个数
R	1 163. 6, 1 415. 45, 1 422. 54	3
FFT	1 415. 45, 1 422. 54	2
FD	947. 11, 948. 21, 948. 76, 1 411. 9	4
SGFD	947. 11, 947. 66, 948. 76, 949. 31, 1 411. 9	5
SGSD	1 050. 07, 1376. 24, 1379. 82, 1 394. 11	4
SGSD(Log R)	1 394. 11, 1 419	2
FD(1/Log R)	1 329. 42, 1 333. 04, 1 336. 65, 1 340. 27, 1 408. 35, 1 411. 9, 1 419, 1 457. 81, 1 461. 31, 1 534. 25, 1 537. 68, 1 541. 11, 1 544. 54, 1 547. 97, 1 551. 39, 1 554. 81, 1 558. 23, 1 602. 32, 1 605. 69	19
SGSD[Log(1/ R)]	1 394. 11, 1 419, 1 457. 81	3

表 3 不同相关系数提取最佳波段范围

Table 3 SSC optimal bands by different r value of the correlation coefficient

光谱变换形式	最佳波段/nm	自变量个数
$R(0.5)$	505. 09, 507. 43, 511. 51, 512. 09, 513. 84, 514. 42, 516. 17, 517. 97, 518. 5, 520. 24~521. 41, 522. 57~523. 74, 524. 9~534. 79, 535. 95~1 649. 15	13
FFTS(0.6)	629. 6~643. 38, 659. 44~660. 58, 689. 74~697. 15, 698. 87~721. 63, 795. 71~804. 15, 843. 94~1 523. 92	6
FD(0.6)	520. 24~524. 32, 946. 56~949. 31, 1 411. 9	3
SGFD(0.6)	520. 24~524. 32, 585. 79~586. 95, 946. 56~949. 31, 1059. 47, 1 038. 6, 1 411. 9	6
SGSD(0.6)	1 057. 85, 1 042. 42~1 050. 07, 1 325. 79, 1 358. 29, 1 376. 24, 1 379. 82, 1 394. 11, 1 401. 24	8
SGSD(Log R)(0.5)	658. 29~660. 58, 941. 61~944. 91, 951. 51~954. 81, 1 340. 27, 1 343. 88, 1 358. 29, 1 401. 24, 1 419	8
FD(1/Log R)(0.7)	1 340. 27, 1 408. 35, 1 411. 9, 141 9, 1 457. 81, 1 461. 31, 1 537. 68~1 551. 39, 1 602. 32	8
SGSD[Log(1/ R)](0.5)	658. 29~660. 58, 941. 61~944. 91, 951. 51~954. 81, 1 340. 27, 1 343. 88, 1 358. 29, 1 386. 97~1 394. 11, 1 401. 24, 1 419, 1 422. 54, 1 457. 81, 1 461. 31, 1 468. 32	13

分析比较，两种最佳波段的选取方法下，八种光谱数据集所确定的土壤盐分含量最佳光谱响应波段集中在 947. 11~949. 31, 1 340. 27, 1 394. 11, 1 419, 1 457. 81~1 461. 31, 1 537. 68~1 551. 39 和 1 602. 32 nm 波段范围，这些结果与已有研究结果基本一致^[2, 7, 14]。

2.4 土壤盐含量和光谱反射率的关系模型

分别以最佳波段和全波段(400~1 650 nm)的光谱数据为自变量，盐含量为因变量，采用 leave-one-out 交叉验证法确定回归模型中的最佳因子数建立土壤盐含量的 PLSR 模型。

2.4.1 全波段模型的建立

以原始光谱数据、平滑光谱数据反射率及其六种光谱变换数据分别与土壤盐含量建立的 PLSR 全波段模型表明，相比于原始和平滑反射率来看，SGSD, SGFD 和 FD 变换对全波段反演模型的精度具有明显的改善作用，SGSD(Log R)和

SGSD(Log(1/ R))处理后模型精度有所上升，而 FD(1/Log R)改善模型精度作用较其他变换稍弱。其中，SGSD 的校正集的相关系数和决定系数最高，分别为 0.823 和 0.673，并且建模集和校正集的标准偏差较接近，分别为 0.980 和 1.256。因此，在本研究中，采用该光谱变换形式建立的全波段盐渍土盐含量 PLSR 预测模型最为理想。

2.4.2 基于最佳波段土壤盐分含量光谱模型的建立

基于表 2 与表 3 中两种最佳波段提取方法所列出的敏感波段光谱数据作为自变量，土壤盐含量作为因变量，利用交叉验证的方法分别建立偏最小二乘模型。综合比较，SGSD(Log R)是极值提取最佳波段方法下最理想的光谱变换形式，根据 AIC 的原则^[12]，SGSD(Log R)与其他光谱变换方式相比，AIC 值最小，为 18.460。另外，其校正集模型的相关系数为 0.797，校正集样本的 SEP 为 1.359, RMSEP 为 1.345，与建模样本的标准偏差和标准差均很接近。基于

SGSD(Log R) 变换的偏最小二乘回归模型的表达式为式(1)

$$y = 20.947\ 975 + 9\ 752 \times R_{1\ 394.11} - 17\ 230 \times R_{1\ 419} \quad (1)$$

式(1)中, y 为土壤盐含量, $R_{1\ 394.11}$ 和 $R_{1\ 419}$ 分别为 1 394.11 和 1 419 nm 处的光谱反射率。

采用不同相关系数范围提取最佳波段的方法, 分析表明, SGSD(Log R) 也是最佳的光谱变换形式, 其 AIC 值最小, 为 28.516; 其校正样本的相关系数为 0.812, 校正集样本的标准偏差(SEP)为 1.315, 校正标准差(RMSEP)为 1.284, 与建模样本的标准偏差和标准差均很接近, 说明其预测效果好。其偏最小二乘回归模型的表达式为式(2)

$$y = 16.200\ 312 - 75\ 220 \times R_{658.29 \sim 660.58} + 52\ 120 \times R_{941.61 \sim 944.91} - 19\ 740 \times R_{951.51 \sim 954.81} + 6\ 608 \times R_{1\ 340.27 \sim 9263} \times R_{1\ 343.88} - 13\ 950 \times R_{1\ 358.29} + 240.870 \times R_{1401.24} - 11\ 130 \times R_{1\ 419} \quad (2)$$

式(2)中, y 为土壤盐含量, $R_{658.29 \sim 660.58}$ 为 658.29~660.58 nm 波段的光谱平均反射率, 其他的以此类推。

以上两种不同的最佳波段提取方法所建立的建模集和校正集模型有相同的特点。综合考虑相关系数、决定系数、建模集与校正集中标准差与标准偏差大小, 最理想的光谱变换形式为 SGSD(Log R), 而原始光谱与平滑光谱的处理下的建模效果均不是很理想。

比较全波段和最佳波段分别建立的 PLSR 模型精度, 平滑光谱数据集(FFT)和四种光谱变换 FD, SGFD, SGSD 和 SGSD[Log(1/ R)]数据集的光谱反射率所建立的全波段模型精度优于最佳波段模型精度; 而原光谱数据集 R 和两种光谱变换[FD(1/Log R) 和 SGSD(Log R)]所建立的最佳波段模型精度优于全波段模型精度。

3 结 论

(1)在跟踪土壤修复过程的 70 天中, 光谱反射率与土壤盐含量的变化总体一致。在土壤修复的 30~50 天, 土壤盐含量和光谱反射率下降趋势尤为接近, 在 50~70 天之间, 随着

“返盐”现象导致 SSC 的含量升高, 平均反射率也呈不同程度的增强。光谱反射率对于土壤盐含量渐变过程的敏感, 说明土壤光谱反射率对盐渍土盐含量变化具有较好的指示作用。

(2)在土壤修复的过程中, FFT 滤波器平滑效果较好, 光谱曲线的噪声大大降低, 并且完整地保存了光谱曲线的变化趋势。不同盐含量的盐渍土光谱的连续统去除法处理显现了盐渍土细小而较弱的吸收特征, 3 个不同深度的吸收谷, 分别在 485.83, 907.92 和 1411.90 nm 处, 与已有同类研究结果^[5,7]相似。

针对八种光谱数据集, 基于相关系数的极值和不同相关系数范围两种方法提取的土壤最佳敏感波段, 大致集中在相同范围, 且与领域内已有相关研究结果基本一致^[2,7,14]。在相关系数极值提取法下的自变量波段个数大多为 2~5 个, 而不同相关系数范围提取法下的自变量波段数为 3~8 个。极值提取下的最佳波段数较少, 简化了模型的运算, 但所建模型精度略低于基于不同相关系数范围方法所建模型。

(3)应用 PLSR 方法, 基于全波段和最佳波段光谱反射率建立的不同变换下土壤盐含量预测模型, 平滑光谱数据集(FFT)和四种光谱变换(FD, SGFD, SGSD 和 SGSD[Log(1/ R)]数据集的光谱反射率所建立的全波段模型精度优于最佳波段模型精度, SGSD 模型定量反演效果最为理想; 而原光谱数据集 R 和两种光谱变换[FD(1/Log R) 和 SGSD(Log R)]所建立的最佳波段模型精度优于全波段模型精度, 且后者效果更为理想。总之, 全波段的 PLSR 模型反演效果总体上优于最佳波段的 PLSR 模型, 这是由于最佳波段的建模模型只运用了全波段的部分光谱, 容易造成一些信息的丢失, 但最佳波段的选取后模型具有变量少、运算量小的特点, 可适用于对模型精度要求不高的场合。这一结果与于雷^[15]的研究结果相似。

更为完善的结果比如本研究运用的两种波段提取方法所得到的最佳波段范围是否能够进一步缩小并得到更高的模型精度有待进一步的探究。另外, 更多土壤类型数据的支持和验证也是本研究进一步深化研究的重要内容。

References

- [1] Howari F M. Journal of Applied Spectroscopy, 2003, 70(5): 782.
- [2] MA Chuang, SHEN Guang-rong, ZHI Yue-e, et al. Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2015, 35(9): 2602.
- [3] Dehaan R L, Taylor G R. Remote sensing of Environment, 2002, 80(3): 406.
- [4] ZHAO Zhen-liang, Taxipulati Teyibai, SUN Qian, et al(赵振亮, 塔西甫拉提·特依拜, 孙倩, 等). Progress in Geography(地理科学进展), 2014, 33(2): 280.
- [5] WENG Yong-ling, QI Hao-ping, FANG Hong-bin, et al(翁永玲, 戚浩平, 方洪宾, 等). Acta Pedologica Sinica(土壤学报), 2010, (6): 1255.
- [6] MA Chuang, SHEN Guang-rong, WANG Zi-jun, et al(马创, 申广荣, 王紫君, 等). Chinese Journal of Soil Science(土壤通报), 2015, (2): 292.
- [7] Nawar S, Buddenbaum H, Hill J. Arabian Journal of Geosciences, 2015, 8: 5217.
- [8] CHU Xiao-li(褚小立). Molecular Spectroscopy Analytical Technology Combined with Chemometrics and Its Applications(化学计量学方法与分子光谱分析技术). Beijing: Chemical Industry Press(北京: 化学工业出版社), 2011. 75.
- [9] PENG Jie, LIU Huan-jun, SHI Zhou, et al(彭杰, 刘焕军, 史舟, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2014, 30(17): 167.
- [10] Gao Xiaohong, Yang Yang, Zhang Wei, et al. SPIE Asia Pacific Remote Sensing, 2014, 9623: 962310-1.

- [11] YIN Li, LIU Qiang, WANG Hui-wen(尹力, 刘强, 王惠文). Journal of System Simulation(系统仿真学报), 2003, 15(1): 135.
- [12] YANG Fu-qin, FENG Hai-kuan, LI Zhen-hai, et al(杨福芹, 冯海宽, 李振海, 等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2015, 46(11): 112.
- [13] PENG Jie, ZHOU Qing, ZHANG Yang-zhu, et al(彭杰, 周清, 张杨珠, 等). Acta Pedologica Sinica(土壤学报), 2013, 50(3): 517.
- [14] Howari F M, Goodell P C, Miyamoto S. Journal of Environmental Quality, 2002, 31(5): 1453.
- [15] YU Lei, HONG Yong-sheng, GENG Lei, et al(于雷, 洪永胜, 耿雷, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2015, 31: 103.

Soil Salt Content and Its Spectral Characteristics During Microbial Remediation Processes

ZHU Yun¹, SHEN Guang-rong^{1, 2*}, WANG Zi-jun¹, LU Shao-ming³, ZHI Yue-e², XIANG Qiao-qiao¹

1. Research Center for Low-Carbon Agriculture, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, China
2. Key Laboratory of Urban Agriculture (South), Ministry of Agriculture, Shanghai 200240, China
3. School of Naval Architecture, Ocean & Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract In this paper, the soil salt content (SSC) and the associated spectral reflectance were measured and analyzed during the microbial remediation process of saline soil. The two methods including extremums of correlation coefficients and the different ranges of correlation coefficients were used to find the optimal sensitive bands of SSC for eight spectral data sets covering the raw spectral reflectance, the smoothed spectral reflectance and six different pre-processing transformations of spectral data of saline soil. With this basis, partial least squares regression (PLSR) was used to build relational models between SSC and spectral reflectance based on full bands (400~1 650 nm) and optimal bands, respectively. The results showed that the optimal spectral bands for eight spectral data sets, concentrated on 947.11~949.31, 1 340.27, 1 394.11, 1 457.81~1 461.31, 1 537.68~1 551.39 and 1 602.32 nm. Taking the coefficient of determination (R^2), root mean squared error (RMSE) and akaike's information criterion(AIC) as criteria to select the best model. For the PLSR predicting models of SSC based on optimal bands from two different ways, the SGSD (LogR) obtained more robust calibration and prediction accuracies than other pre-processing inversion models. Compared with optimal bands, the full bands using PLSR method could obtain better prediction accuracies on the whole. Among all of the eight spectral data sets in full bands, the prediction accuracy of SGSD was the best, the corresponding R^2 and RMSEP of the predicted model were 0.673 and 1.256. For the inversion models based on optimal bands, although there was a slight gap in the prediction accuracy with that based on full bands, they also had their own merits; these models were much simpler and thus the reducing model computation and modelling speed were more important than improving prediction accuracy. The results of this study showed that the method had a great potential for diagnosing and monitoring soil salinization quickly and conveniently in researching the relation between SSC and soil reflectance spectra.

Keywords Saline soil; Microbial remediation; Spectral transformations; PLSR

(Received Dec. 28, 2015; accepted Apr. 8, 2016)

* Corresponding author