

# 基于不同光谱变换的土壤盐含量光谱特征分析

朱 贇<sup>1,3</sup>, 申广荣<sup>1,2,3\*</sup>, 项巧巧<sup>1,3</sup>, 吴 裕<sup>1,3</sup>

(1.上海交通大学 农业与生物学院和低碳农业研究中心, 上海 200240; 2.都市农业(南方)重点实验室, 上海 200240;  
3.国家林业局上海城市森林生态系统国家定位观测研究站, 上海 200240)

**摘 要:**跟踪初生盐渍土壤的微生物修复实验, 采用同步实测得土壤盐含量和光谱数据, 详细分析了基于34种光谱变换, 修复过程中盐渍土的光谱特征。对于选取的6种光谱变换, 采用全波段(400~1650 nm)和分析获得的最佳敏感波段分别建立了土壤盐含量的光谱反演PLSR(partial least squares regression)模型。研究表明, 光谱变换处理使土壤盐含量与平滑后的光谱反射数据的相关性明显增强, 且最佳敏感波段范围进一步聚焦。本研究得到最佳光谱变换为导数变换, 基于全波段的土壤盐含量预测模型以SGSD变换效果最好, 与原始光谱相比, 模型的 $r$ 、RMSEP分别从0.537和1.928改善到0.823和1.256。而SGSD(LogR)是基于最佳波段所建立的盐含量预测模型的有效光谱变换方法, 该研究为进一步实现盐渍土中盐含量快速定量分析提供了方法和数据参考。

**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:**S15      **文献标识码:**A      **文章编号:**0564-3945(2017)03-0560-09

DOI: 10.19336/j.cnki.trtb.2017.03.08

朱 贇, 申广荣, 项巧巧, 吴 裕. 基于不同光谱变换的土壤盐含量光谱特征分析[J]. 土壤通报, 2017, 48(3): 560-568

ZHU Yun, SHEN Guang-rong, XIANG Qiao-qiao, WU Yu. Spectral Characteristics of Soil Salinity Based on Different Pre-processing Methods [J]. Chinese Journal of Soil Science, 2017, 48(3): 560-568

土壤盐分及其光谱特征是土壤盐渍化高光谱遥感定量监测的基础。土壤的光谱反射特征是土壤母质、盐分、水分、有机质含量、氧化物、质地等各种理化性状互相作用综合影响的结果。具有精细光谱分辨率、纳米级连续光谱的高光谱信息反映了土壤光谱的细微特征, 使得深入分析土壤盐分含量的高光谱特征成为了可能。因而分析研究盐渍土光谱吸收、反射特征随土壤盐分含量的变化规律, 进而建立土壤盐渍化高光谱遥感定量监测模型成为了当前的研究热点<sup>[1,2]</sup>。

土壤盐含量(soil salinity content, SSC)高光谱定量反演模型的精度是其有效性与适用性的关键因素<sup>[3]</sup>。光谱数据的变换分析作为提高模型反演精度的有效途径之一, 受到了广泛的关注。研究表明光谱数据的变换对数据挖掘非常有益<sup>[4-6]</sup>, 合适的光谱变换方法可减弱乃至消除光谱数据的无关信息, 突出微妙的光谱特征, 有助于提高土壤盐含量模型的预测精度与稳健性<sup>[7,8]</sup>。选择合适的光谱变换方法是模型建立的重要一步, 一阶导数、二阶导数和连续统去除等变换在目前研究中已有广泛应用<sup>[9]</sup>。导数法可以消除基线漂移或平缓背景干扰的影响, 提供比原光谱更清晰的光谱轮廓变化, 而连续

统去除则可以更好地突出土壤光谱曲线的吸收和反射特征, 从而提取出光谱特征波段。陶兰花和陈东强<sup>[10,11]</sup>等发现土壤反射率的变换形式中, 二阶导数是构建预测模型的最佳光谱方式; Said Nawar 和曹文涛等<sup>[11,12]</sup>则认为一阶导数形式是盐含量的最佳估算的变换形式; 而刘娅<sup>[13]</sup>等用原始光谱反射率建立的盐含量预测模型取得了最佳效果; 也有研究者对原始光谱反射率进行连续统去除变换后, 建立了效果较好的土壤盐含量的PLSR预测模型<sup>[14,15]</sup>。盐渍化程度的不同其光谱最佳变换方式也不相同, 研究表明, 非盐渍化、中度盐渍化土壤的二阶导数变换形式效果更佳, 轻度、重度盐渍化土壤的一阶导数变换形式效果更佳<sup>[16]</sup>。由于土壤具有高度异质性等特点, 有关土壤光谱变换分析研究所获结果各有不同, 有的因研究区不同而不同<sup>[10,12]</sup>, 有的因土壤类型或盐渍化程度不同而不同<sup>[11,13]</sup>, 即使同一研究区针对土壤不同属性的最佳变换方法也不相同<sup>[17]</sup>。因此, 深入分析探究不同光谱变换的盐渍土盐分含量高光谱响应特征对高效、高精度的盐渍土高光谱遥感定量监测具有重要意义。

基于以上分析, 本研究基于盐渍化土壤修复过程

收稿日期: 2016-11-23; 修订日期: 2016-11-23

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(41471120)、社科重大项目(14ZDB139)及上海交大农工交叉项目(Agri-X2015004)资助

作者简介: 朱 贇(1991-), 女, 上海人, 硕士生, 主要从事盐渍土的光谱特征研究。E-mail: zy\_1125@sjtu.edu.cn

\* 通讯作者: E-mail: sgrong@sjtu.edu.cn

中不同盐分含量的光谱测试,对土壤原始光谱进行光谱平滑,以及平滑后导数变换、连续统去除、光谱增强、归一化、均值中心化、基线校正等 34 种光谱变换,进而分析土壤盐含量的原始光谱、平滑光谱及平滑后的 34 种变换的光谱特征,力求从冗余的高光谱数据中选择对土壤盐分含量变化敏感的最佳波段,进而利用 PLSR 法以全波段和最佳响应波段分别进行土壤盐含量的定量反演模型研究,为提高土壤盐渍化高光谱遥感反演的准确性和实用性提供科学依据。

## 1 材料与方法

### 1.1 实验设计

本研究土壤样品的采集区域为山东省北部的东营市(北纬 36°55′ ~ 38°10′,东经 118°07′ ~ 119°10′),土壤类型为初生盐渍土,其盐渍化土壤为砂质壤土,黏粒含量为 10 ~ 15%,pH 为 7.850,有机质含量为 2.210 g kg<sup>-1</sup>,土壤总氮含量为 0.258 g kg<sup>-1</sup>。土壤盐分中,阴离子以 Cl<sup>-</sup> 为主,阳离子以 Na<sup>+</sup> 为主,盐分含量主要以氯化钠为主(表 1)。

表 1 盐分离子含量统计

Table 1 Statistics of salinity ions

盐分离子含量 Salinity ion content	最小值 Minimum (g kg <sup>-1</sup> )	最大值 Maximum (g kg <sup>-1</sup> )	均值 Mean (g kg <sup>-1</sup> )	分别占阴(阳)离子比例 Proportion (%)
Cl <sup>-</sup>	0.800	2.477	1.766	57.889
SO <sub>4</sub> <sup>2-</sup>	0.075	3.312	0.985	32.284
HCO <sub>3</sub> <sup>-</sup>	0.195	0.557	0.300	9.827
CO <sub>3</sub> <sup>2-</sup>	0	0	0	0
Na <sup>+</sup>	0.227	2.256	0.465	28.823
Ca <sup>2+</sup>	0.166	1.056	0.480	28.377
Mg <sup>2+</sup>	0.284	0.896	0.474	28.045
K <sup>+</sup>	0.038	0.627	0.249	14.755
SSC	6.248	13.804	8.211	-

为了获取连续的、渐变土壤盐分含量对应的光谱数据,本研究跟踪土壤修复实验并同步进行光谱测试。修复实验采用已获专利(CN102676492A)<sup>[18]</sup>并在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。

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,未经任何其它处理,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,各个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。温室为玻璃温室,室内设有智能环境控制设备,温度维持在 40 °C 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。

### 1.2 数据获取

**1.2.1 土样的采集与理化分析** 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,采集的深度为 0 ~ 20 cm,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,先清除土样中的一些杂草和小石块等非土壤成分物质,于实验室自然风干后,经手工磨碎、过 100 目筛(0.15 mm),装入编号的密封袋中,分为两份,一份供农化分析使用,另一份供光谱测定使用。盐分含量分析时按 m(土):m(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<sup>+</sup>、K<sup>+</sup>、Mg<sup>2+</sup>、Ca<sup>2+</sup>、Cl<sup>-</sup>、HCO<sub>3</sub><sup>-</sup>、CO<sub>3</sub><sup>2-</sup>、SO<sub>4</sub><sup>2-</sup>)含量,土壤盐含量通过电导率获得<sup>[19-21]</sup>。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<sup>[12, 22]</sup>,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 土样的光谱数据测定** 土样光谱数据测定采用 AvaField-2 便携式高光谱地物波谱仪(荷兰 Avantes 公司),该光谱仪波长范围为 350 ~ 1700 nm,采样间隔为 0.59 nm(350 ~ 1040 nm)和 3.8 nm(1040 ~ 1700 nm)。实验室光谱测定在暗室内进行,将处理好的土样置于培养皿(直径 10 cm、深 2 cm)内,装满后,采用直尺将其表面刮平。以光谱仪配套的功率为 1000 W 的卤素灯为光源,调节光照角度使得灯光最亮处位于样品中央,传感器探头的视场角为 5°,到土壤表面距离为 15 cm,使探头接收土壤光谱的区域远小于培养皿面积,以保证探头所接收的均为土样的反射光谱。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,取其算术平均值为该土壤样品的实际反射率。

### 1.3 研究方法

**1.3.1 光谱数据平滑去噪** 在获取土样的光谱数据时,由于样品自身、环境或人为操作的影响,使土样光谱曲线中不可避免地包含噪声与干扰信息,光谱数据

平滑可以有效地消除光谱数据无关信息和噪声, 突出光谱有用的信息。应用 OriginPro 9.1 软件对光谱曲线进行平滑去噪处理, 平滑方法选用 FFT 滤波, 去除噪声较大的边缘波段 350 ~ 400 nm 和 1650 ~ 1700 nm。

### 1.3.2 光谱数据变换分析 光谱变换是盐渍土壤光谱

特征分析中重要内容, 理想的光谱变换可以有效提高模型的预测精度与稳健性。本研究针对平滑后的光谱数据尝试了光谱增强、归一化、均值中心化、基线校正、导数光谱法和简单的数学变换等 34 种光谱变换(表 2)。

表 2 光谱变换

Table 2 Spectral transformation

类型 Type	光谱变换方式 Pre-processing method	英文全称 English name	简写 Abbreviation
原始	原始光谱反射率	raw reflectance spectra	R
平滑去噪	FFT 滤波器平滑	FFT Filters	FFT
归一化	面积归一化	area normalization	AN
	单位向量归一化	unit vector normalization	UVN
	平均值归一化	mean normalization	MN
	最大值归一化	maximum normalization	MXN
	范围归一化	range normalization	RN
光谱增强	多元散射校正	multiplicative scatter correction	MSC
	标准正态变量变换	standard normal variate	SNV
基线校正	线性基线校正	linear baseline correction	LBC
均值中心化	均值中心化	mean center	MC
导数	一阶微分(差分)	first derivative	FD
	二阶微分(差分)	second derivative	SD
	一阶微分(S.Golay)	Savitzky-Golay first derivative	SGFD
	二阶微分(S.Golay)	Savitzky-Golay second derivative	SGSD
	倒数的一阶微分(差分)	inverse of first derivative	FD(1/R)
	倒数的一阶微分(S.Golay)	inverse of Savitzky-Golay first derivative	SGFD(1/R)
	倒数的二阶微分(S.Golay)	inverse of Savitzky-Golay second derivative	SGSD(1/R)
	对数的一阶微分(差分)	logarithm of first derivative	FD(LogR)
	对数的一阶微分(S.Golay)	logarithm of Savitzky-Golay first derivative	SGFD(LogR)
	对数的二阶微分(S.Golay)	logarithm of Savitzky-Golay second derivative	SGSD(LogR)
	对数倒数的一阶微分(差分)	inverse of LogR of first derivative	FD(1/LogR)
	对数倒数的一阶微分(S.Golay)	inverse of LogR of Savitzky-Golay first derivative	SGFD(1/LogR)
	对数倒数的二阶微分(S.Golay)	inverse of LogR of Savitzky-Golay second derivative	SGSD(1/LogR)
	倒数对数的一阶微分(差分)	logarithm of 1/R of first derivative	FD(Log(1/R))
	倒数对数的一阶微分(S.Golay)	logarithm of 1/R of Savitzky-Golay first derivative	SGFD(Log(1/R))
	倒数对数的二阶微分(S.Golay)	logarithm of 1/R of Savitzky-Golay second derivative	SGSD(Log(1/R))
	均方根的一阶微分(差分)	square root of R of first derivative	FD(√R)
	均方根的一阶微分(S.Golay)	square root of R of Savitzky-Golay first derivative	SGFD(√R)
	均方根的二阶微分(S.Golay)	square root of R of Savitzky-Golay second derivative	SGSD(Log(√R))
简单数学变换	倒数	inverse of R	1/R
	对数	logarithm of R	LogR
	对数的倒数	inverse of LogR	1/LogR
	倒数的对数	logarithm of 1/R	Log(1/R)
	均方根	square root of R	√R
	连续统去除	continuum removal	CR

光谱增强可以凸显在原始光谱中不明显的光谱特征, 包括多元散射校正(multiplicative scatter correction, MSC)和标准正态变量变换(standard normal variate, SNV), 两者的目的基本相同, 即消除各土样之间因颗粒度、填满密度、分布均匀等差异所产生的影响, 增强与盐含量相关的光谱信息。MSC 是由 Geladi<sup>[23]</sup>等人提出的, 它是通过测量光谱之间的回归分析, 使用回归拟合得到的截距和斜率来校正测得的每个光谱数据, 截距大小反映的是样品独特反射作用, 斜率大小则反映土样的均匀性。SNV 与 MSC 有所不同, 它是通过对每

一个光谱中的各波长点的反射率应满足一定分布(如正态分布)的假设, 对每一条光谱曲线进行校正。

归一化(normalization)的算法较多, 有面积归一化法、最大归一化法、单位向量归一化、平均值归一化、最大值归一化和范围归一化等, 光谱归一化的目的是消除光程或样品稀释等变化对光谱响应造成的影响。它是将土样光谱数据的范围集中在 0 ~ 1 之间, 使得各土样光谱特征控制在相同范围内。

均值中心化(mean centering, MC)往往用来增加土样光谱之间的差异, 从而提高模型的稳健性与预测能

力,它是将样本光谱减去样本集的平均光谱,以此来消除光谱数据的绝对吸收值。

基线校正(baseline correction)又称为背景校正,是消除由于光谱测量环境的变化等导致光谱平移或旋转且基线不能保证完全不变问题的有效措施之一,通过去除基线,可以使测得的土样光谱在同一标准上进行比较<sup>[7]</sup>。

导数应用于光谱分析中主要用于校正光谱基线,消除其他背景干扰,提供比原光谱更高的分辨率和更清晰的光谱轮廓变化,但它同时会放大噪声,导致信噪比降低。因此,对光谱进行导数运算前,应首先进行平滑处理。主要包括一阶(first derivative, FD)和二阶导数(second derivative, SD)处理,一般采用直接差分法(也称为 Norris 求导法<sup>[24, 25]</sup>)和 Savitzky-Golay 卷积求导法<sup>[26]</sup>,直接差分法对于分辨率高、波长采样点多的光谱具有较好的适用性,它是一种最简单的离散波谱求导方法。但对于稀疏波长采样点的光谱,则采用 Savitzky-Golay 卷积求导法更合适,它是通过最小二乘法可以计算得到平滑系数相似的导数系数。

连续统去除法(continuum removal, CR)可以有效地突出光谱曲线的吸收和反射特征<sup>[27]</sup>。从直观上来看,连续统是一个直线逐点连接光谱上局部光谱反射极值点的凸壳(convex hull),相当于光谱曲线的“外壳”,本研究也对平滑后的光谱曲线进行了连续统去除的变换,用于识别土样光谱曲线特征吸收谷的位置,以便更好地反映土壤光谱曲线的总体形状特征。

对反射率进行光谱倒数的对数变换能够在一定程度上提高可见光波段的光谱差异<sup>[28]</sup>。本文中采用的简单数学变换分析包括倒数、对数、对数的倒数、倒数的对数和均方根。

**1.3.3 土壤盐含量最佳敏感波段及其反演模型构建** 本研究分别基于全波段和敏感波段建立土壤盐含量反演模型,并分析其中不同光谱变换对相关模型的影响和作用。在最佳敏感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(extremums of correlation coefficients, ECE)<sup>[29]</sup>,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(different ranges of correlation coefficients, DRCC)<sup>[29]</sup>,分别对不同相关系数范围( $|\text{r}| \geq 0.5$ 、 $\geq 0.6$ 、 $\geq 0.7$ )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,最低的标准差的模型,选取其对应的相关系数范围下的波段。两种方法选取的最佳敏感波段均需在

0.05 置信水平下,通过  $P$  值小于 0.01 的显著性检验来确定最佳敏感波段范围。

在进行土壤盐含量反演模型的构建时,首先对土壤样本数据进行预处理,对采集来的 216 个(6 个处理,每个处理 3 个梯度,每个梯度 3 个重复,采样 4 次)土样的测试数据,每个修复处理下各个梯度的 3 次重复的盐含量分别求取平均值,这样,每个修复时间段内共得到 18 个土样数据,4 次采样共得到 72 个土壤样本数据,剔除样本集中浓度标准值或光谱数据偏差较大的异常样本(outliers)。本研究将去除异常样本后的 63 个样本分为建模样本 42 个(2/3)和校正样本 21 个(1/3)<sup>[20]</sup>。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 42 个样本来作建模样本,剩余的 21 个样本作为校正样本。

本文采用 PLSR<sup>[30]</sup>的建模方法对盐渍土中的盐含量进行预测,PLSR 方法能够在减少光谱维数的同时,从光谱数据中揭示最大盐含量变化的主控因子,使建立的模型具有更好的鲁棒性<sup>[22, 31]</sup>。分别以全波段和 2 种方法选取的最佳敏感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,建模的过程中的验证方法采用交叉验证(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,选取相关系数( $r$ )、建模样本的标准差(RMSEC)和校正样本的标准差(RMSEP)来衡量评价并判断选择最优模型<sup>[26, 32]</sup>。

## 2 结果与分析

### 2.1 修复过程中盐渍土光谱特征分析

在整个土壤修复实验过程中,土壤盐含量呈逐步下降趋势。土壤盐含量从修复前的  $15.453 \text{ g kg}^{-1}$  降至修复 10 天的  $11.340 \text{ g kg}^{-1}$ ,修复 30 天的  $8.887 \text{ g kg}^{-1}$  及修复 50 天的  $8.489 \text{ g kg}^{-1}$ ,修复前 50 天的修复效果较为理想,而“返盐”<sup>[33]</sup>现象使得修复 70 天的土壤盐含量上升为  $13.432 \text{ g kg}^{-1}$ 。相应地,以每个修复时间段内所有土壤样品 SSC 含量的均值和土壤光谱反射率的关系看(图 1(a)),修复土壤的光谱反射率在 0 ~ 50 天随着盐含量的降低也随之下降,在修复 70 天盐含量升高后,光谱反射率也随之上升,说明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。修复过程中,各盐含量光谱曲线在形态上几乎一致,在  $1411.90 \text{ nm}$  左右有明显的吸收谷;在  $400.11 \sim 1376.24 \text{ nm}$  的范围内,土壤反射率随着波长的增加呈

现明显的上升趋势;在 1376.24 ~ 1411.90 nm 范围内的光谱反射率随着波长增加而递减;在 1411.90 ~ 1649.15 nm 的范围内,土壤反射率又整体缓慢增大并逐渐趋于平缓,这些特征与相关研究结果基本一致<sup>[34]</sup>。

不同的光谱变换对修复过程中盐渍土光谱特征的影响不同。连续统去除使得盐渍土光谱反射率明显增强,相比原始光谱曲线,吸收特征值得到了明显的放大,如 485.83 nm、907.92 nm 和 1411.9 nm 的微

弱吸收带在连续统去除曲线中明显可见,其中波谷最深的为反映水汽吸收的 1411.9 nm,通常认为 1400 nm 附近随着水分含量增加,水汽吸收增强,相对应的光谱吸收谷也增强<sup>[35]</sup>(图 1(b))。一阶微分变换处理后,土壤光谱特征值明显增多,可清晰的看出,盐含量特征波段集中在 442.48 nm、472.96 nm、558.59 nm、1059.47 nm、1397.67 nm、1419 nm 等波段附近(图 1(c))。

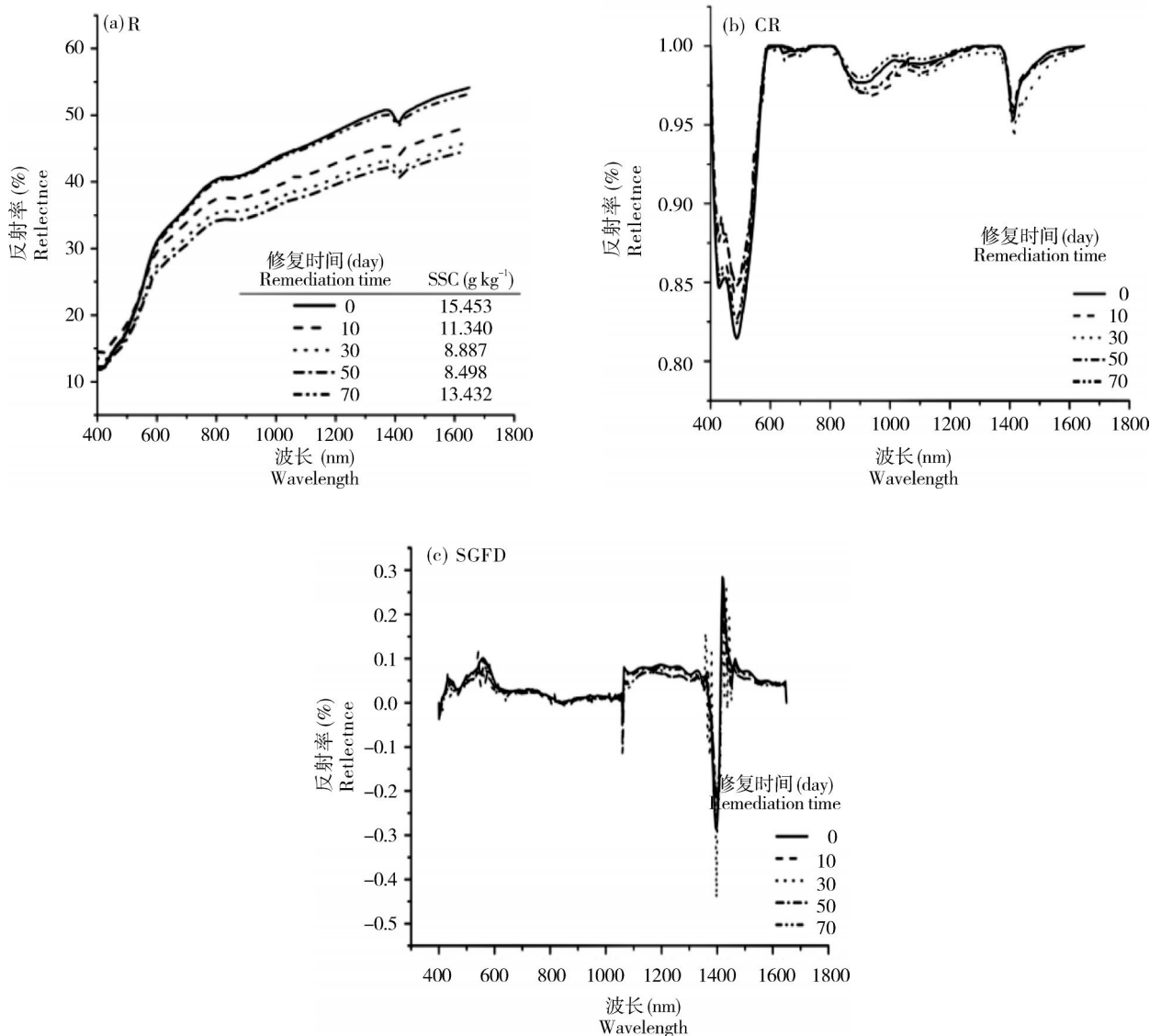


图 1 a.各组土样光谱反射率曲线,b.各组土样光谱连续统去除后反射率曲线,c.各组土样光谱一阶微分后反射率曲线

Fig. 1 a. Reflectance spectra of soil samples in each group; b. Continuum removal reflectance spectra of soil samples in each group; c. Savitzky-golay first derivative reflectance spectra of soil samples in each group.

## 2.2 不同光谱变换对盐渍土壤最佳敏感波段的影响

对原始光谱、平滑光谱及平滑后的 34 种不同变换光谱数据即共 36 种光谱数据集(表 2)分别进行全波段的 PLSR 土壤 SSC 反演建模,并将每一种光谱变换分析对应的模型精度与平滑光谱数据对应的 PLSR 模

型精度进行比较,以模型的相关系数显著提高幅度,标准差变小程度为评判基准<sup>[36]</sup>,选择得到有效提高模型精度的 FD、SGFD、SGSD、SGSD(LogR)、FD(1/LogR)和 SGSD(Log(1/R))6 种光谱变换数据集,用于进一步分析盐含量光谱特征。

表 3 土壤盐含量与不同变换形式光谱的相关性分析

Table 3 Correlation analysis between different spectral transformation and soil salinity content

光谱变换方式 Pre-processing method	最大相关系数所在波段 (nm) Maximum correlated spectra	相关系数 (r) Correlation coefficient	相关系数提升比率 (%) Promotion of correlation coefficient
R	1182.3	0.632**	-
FD	1411.9	0.721**	14
SGFD	1411.9	0.721**	14
SGSD	1394.11	0.643**	2
SGSD(LogR)	1419	0.755**	20
FD(1/LogR)	1457.81	0.770**	22
SGSD(Log(1/R))	1419	0.755**	20

注:表中\*\*表示极显著相关( $P < 0.01$ )。

原始光谱与土壤盐含量相关系数大多在  $-0.172 \sim 0.610$  之间,相关系数最大达到 0.632(1182.3 nm),而经平滑后的光谱数据经 FD、SGFD、SGSD、SGSD(LogR)、FD(1/LogR)和 SGSD(Log(1/R))变换后,光谱反射率与盐含量的相关性呈显著增强,相关系数最大值相对于原始光谱均有提升(表 3),其中,SGSD(LogR)、FD(1/LogR)和 SGSD(Log(1/R))分别提升了 20%、22% 和 20%,且这 6 种变换的相关系数最大值的波长位置大多集中于 1394.11 ~ 1457.82 nm 邻近。

基于相关系数极值和不同相关系数范围<sup>[29]</sup>两种方法提取的不同光谱变换数据集的最佳敏感波段见表 4。在相关系数极值方法下,相比原始光谱所提取的 3 个最佳敏感波段,基于 SGSD(LogR)变换的最佳敏感波段

范围集中于 1394.11 nm 和 1419 nm 波段处。而根据不同相关系数范围方法,基于原始光谱提取的最佳敏感波段数多达 13 个,除了基于 SGSD(Log(1/R))变换没有减少最佳波段个数外,其余 5 种变换的最佳敏感波段数均有减少,尤其 FD 变换,仅为 3 个波段。因此,对平滑后的光谱数据进行 6 种变换分析都不同程度地使最佳敏感波段更为集中,使进一步构建的预测模型得到简化。综合分析,所有变换的最佳敏感波段大多集中于近红外区域,这些波段的交集也即本研究所确定的土壤盐含量最佳光谱响应波段为 947.11 ~ 949.31 nm、1340.27 nm、1394.11 nm、1419 nm、1457.81 ~ 1461.31 nm、1537.68 ~ 1551.39 nm 和 1602.32 nm,这些与已有研究结果一致<sup>[12, 37, 38]</sup>。

表 4 不同光谱变换下土壤盐含量的最佳波段 PLSR 模型预测结果

Table 4 Prediction results of soil salinity content with PLSR models based on optimal sensitive bands pretreated by different spectral transformation

光谱变换 Pre-processing transformations of spectrum	方法 Method	波段范围 (nm) Spectral range	r	RMSEP
R	ECE	1163.6,1415.45 ,1422.54	0.699	1.635
	DRCC	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 ~ 1649.15	0.551	1.848
FFTS	ECE	1415.45 ,1422.54	0.764	1.437
	DRCC	629.6 ~ 643.38,659.44 ~ 660.58,689.74 ~ 697.15,698.87 ~ 721.63,795.71 ~ 804.15,843.94 ~ 1523.92	0.540	1.990
FD	ECE	947.11, 948.21 ,948.76, 1411.9	0.734	1.535
	DRCC	520.24 ~ 524.32, 946.56 ~ 949.31, 1411.9	0.732	1.525
SGFD	ECE	947.11, 947.66, 948.76, 949.31, 1411.9	0.725	1.604
	DRCC	520.24 ~ 524.32, 585.79 ~ 586.95, 946.56 ~ 949.31, 1059.47, 1038.6, 1411.9	0.774	1.205
SGSD	ECE	1050.07, 1376.24, 1379.82, 1394.11	0.631	1.782
	DRCC	1057.85 ,1042.42 ~ 1050.07, 1325.79, 1358.29, 1376.24, 1379.82, 1394.11, 1401.24	0.672	1.658
SGSD(LogR)	ECE	1394.11, 1419	0.797	1.345
	DRCC	658.29 ~ 660.58, 941.61 ~ 944.91, 951.51 ~ 954.81, 1340.27, 1343.88, 1358.29,1401.24, 1419	0.812	1.284
FD(1/LogR)	ECE	1329.42, 1333.04, 1336.65, 1340.27,1408.35, 1411.9 ,1419 , 1457.81, 1461.31, 1534.25, 1537.68, 1541.11, 1544.54, 1547.97, 1551.39,1554.81, 1558.23, 1602.32, 1605.69	0.787	1.380
	DRCC	1340.27, 1408.35 ,1411.9 ,1419, 1457.81, 1461.31 ,1537.68 ~ 1551.39, 1602.32	0.786	1.384
SGSD(Log(1/R))	ECE	1394.11, 1419,1457.81	0.790	1.369
	DRCC	658.29 ~ 660.58,941.61 ~ 944.91, 951.51 ~ 954.81 ,1340.27, 1343.88,1358.29,1386.97 ~ 1394.11 , 1401.24, 1419,1422.54,1457.81,1461.31,1468.32	0.794	1.349

注:表中,ECE 为相关系数极值法的简称,DRCC 为不同相关系数范围法的简称。

### 2.3 基于不同光谱变换土壤盐含量光谱反演模型的验证与比较

针对选定的 6 种光谱数据集, 分别以全波段和最佳敏感波段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, 建立土壤盐含量的全波段和最佳敏感波段光谱反射率的 PLSR 模型。相比原始光谱数据, 平滑后的光谱变换数据集的土壤盐含量反演模型精度大多都有所提高, 但不同光谱变换下土壤盐含量反演模型精度不同。以原始光谱反射率建立的盐含量预测模型中, 全波段的模型相关系数为 0.537; 基于 6 种变换光谱的模型相关系数均高于原始光谱, FD、SGFD、SGSD、(SGSD(LogR))、FD(1/LogR)和 SGSD(Log(1/R))的预测模型相关系数分别为 0.807、0.800、0.823、0.796、0.711 和 0.796, 相比原始光谱, 其模型精度分别提高了 50%、49%、53%、48%、32%和 48%, 综合来看, SGSD 是全波段下土壤盐含量预测模型的最佳变换。

同样地, SGSD(LogR)是 ECE 和 DRCC 的最佳敏感波段提取方法下的最理想的光谱变换形式。以原始光谱建立的最佳敏感波段(ECE、DRCC)模型相关系数分

别为 0.699 和 0.551, 而光谱变换 SGSD(LogR)在基于 ECE 和 DRCC 选取的最佳敏感波段, 分别建立的土壤盐含量预测模型, 其相关系数分别达到 0.797 和 0.812, 相比原始光谱模型精度提高了 14%和 47%, 并且 SGSD (LogR) 变换下的最佳敏感波段 PLSR 模型的 RMSEP 均较小, 说明其预测效果好。

校正集的预测值和实测值相关图是判断验证模型结果最为直观的方法, 首先, 选出全波段模型中最佳的光谱变换 SGSD 与最佳敏感波段模型中最佳光谱变换 SGSD (LogR) 的 PLSR 预测模型, 然后以校正集实测 SSC 值为 X 轴, 预测 SSC 值为 Y 轴, 做出散点图, 并添加趋势线。

土壤盐含量实测值与反演模型预测值的散点图如图 2 所示, 图 2(a)为全波段 SGSD 实测值和模型预测值的散点图, 图 2(b)为相关系数极值提取的最佳敏感波段 SGSD(LogR)实测值和模型预测值的散点图, 图 2(c)为不同相关系数提取的最佳敏感波段 SGSD (LogR)实测值和模型预测值的散点图。以校正集的预测值和测量值拟合趋势线接近 1 : 1 为标准, 以全波段 SGSD 变换下的预测模型最佳, 全波段的 PLSR 模型反演效果

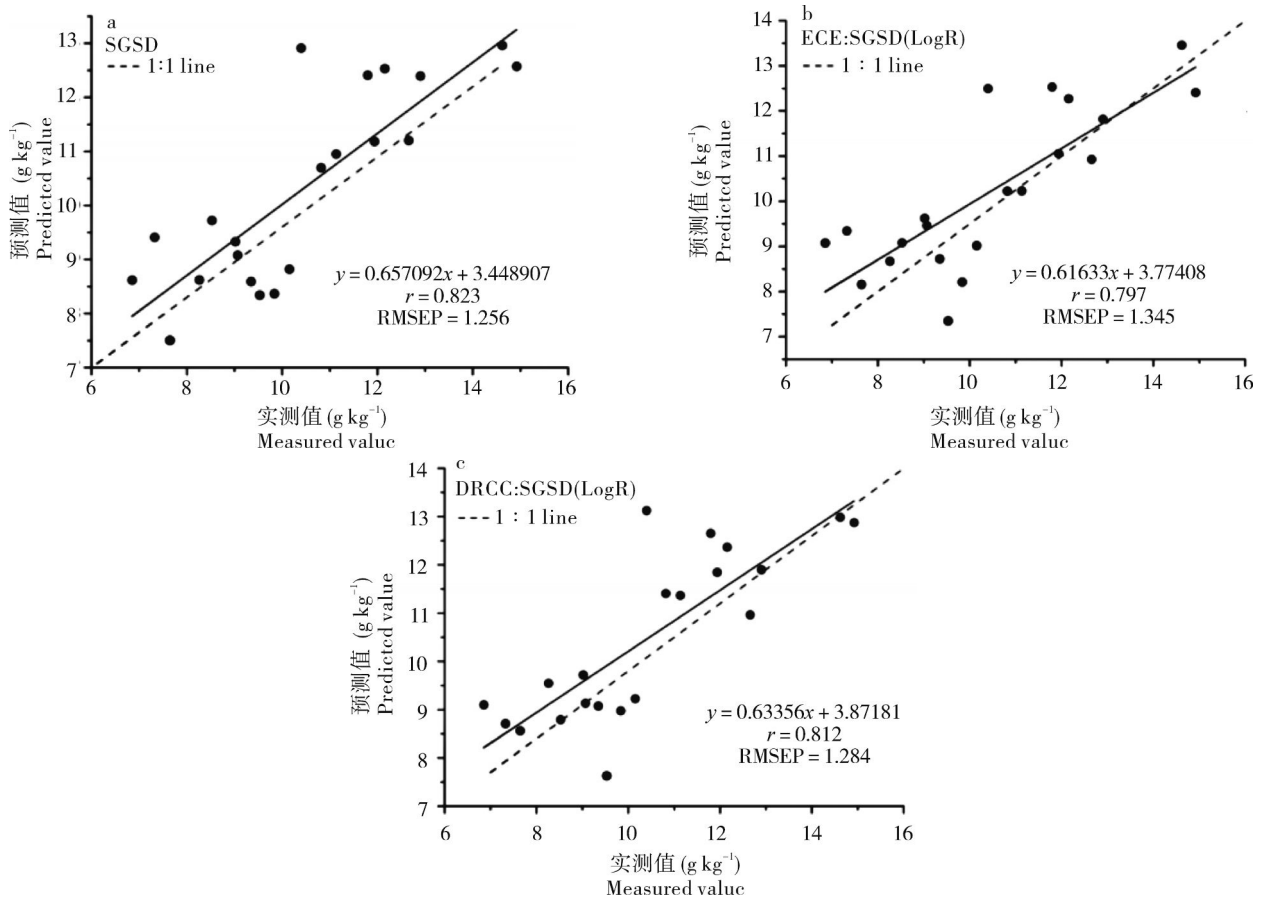


图 2 土壤盐含量实测值和模型预测值的散点图

Fig. 2 Scatterplot between measured values and predicted values in the model of soil salinity contents

优于最佳敏感波段的 PLSR 模型,这是由于最佳敏感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,但最佳敏感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<sup>[9]</sup>的研究结果相似。

### 3 结论与讨论

为定量分析探究盐渍土的光谱特征,本文基于盐渍土修复过程中,渐变的土壤盐含量及其光谱反射数据的连续测试,详细分析基于 34 种不同光谱变换方法,修复过程中盐渍土的光谱特征;不同光谱变换下盐分含量最佳光谱响应波段的变化及其对进一步构建的盐含量光谱反演模型精度的影响。

(1)修复过程中盐渍土,对应不同修复时间段即不同土壤盐含量,土壤光谱反射率差异性明显,随着土壤盐含量的逐渐减少,盐渍化土壤光谱反射率在 560.33 ~ 1649.15 nm 波长范围内均明显降低,不同盐含量土壤光谱曲线均在 1411.90 nm 附近有明显的水分吸收带。连续统去除和一阶导数变换,使得原始光谱反射率的光谱特征放大并突出,在原有基础上,提供了更多的特征波段,有利于更好地研究修复过程中盐渍土的光谱特征。

(2)原始光谱反射率与盐含量相关性不高,平滑后的光谱数据基于选定的 6 种形式的光谱变换能够有效提高光谱反射率与盐含量的相关性,并且进一步聚焦了最佳敏感波段范围。在这 6 种光谱变换中,导数光谱法或基于简单数学变换后再进行的导数变换对增强光谱反射率与盐含量的相关系数曲线的变化趋势,提高二者的相关性效果最佳。

(3)研究分析表明,基于不同光谱变换分析方法土壤盐含量光谱反演模型的模型精度,以二阶导数的效果最为突出。利用 SGSD 变换建立的全波段 PLSR 模型校正集相关系数  $r$  为 0.825, RMSEP 为 1.256,模型精度相比原始光谱建立的全波段 PLSR 模型提高了 53%,优于同类其他光谱变换方法;而 SGSD(logR)均为两种最佳敏感波段提取方法下建立预测模型的最佳变换。

在光谱特征增强的数学变换方法中,本研究选用了应用较多的 34 种进行分析研究。进一步的研究中可以针对光谱形态和土壤盐分含量的吸收反射特性,对各种增强方法进行更加细致地调整、改进,或者加以综合运用,并选取恰当的吸收特征描述方法,以提取稳定性更强、敏感程度更好的特征参数,为采用可见光/近红外反射光谱分析技术快速检测土壤盐含量,提供更

实用、有效的途径。

#### 参考文献:

- [1] 曹文涛,吴泉源,王菲,等.基于野外实测光谱的淮北地区土壤全盐量监测研究[J].土壤通报,2016,47(2):265-271
- [2] MA C, SHEN G R, ZHI Y E, et al. Spectral Characteristics of Salinized Soils during Microbial Remediation Processes [J]. Spectroscopy and Spectral Analysis, 2015, 35(9): 2602 - 2608.
- [3] 向红英,柳维扬,彭杰,等.基于连续统去除法的南疆水稻土有机质含量预测土壤[J].土壤,2016,48(2):389-394.
- [4] SAWUT M, GHULAM A, TIYIP T, et al. Estimating soil sand content using thermal infrared spectra in arid lands [J]. International Journal of Applied Earth Observation and Geoinformation, 2014, 33: 203 - 210.
- [5] ABDEL-RAHMAN E M, MUTANGA O, ODINDI J, et al. A comparison of partial least squares (PLS) and sparse PLS regressions for predicting yield of Swiss chard grown under different irrigation water sources using hyperspectral data [J]. Comput Electron Agr, 2014, 106: 11 - 19.
- [6] RADY A M, GUYER D E, KIRK W, et al. The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict processing-related constituents of potatoes [J]. Journal of Food Engineering, 2014, 135: 11 - 25.
- [7] SEASHOLTZ M B, KOWALSKI B. The parsimony principle applied to multivariate calibration [J]. Analytica Chimica Acta, 1993, 277: 165 - 177.
- [8] XIAO Z Z, LI Y, FENG H. Hyperspectral Models and Forecasting of Physico-Chemical Properties for Salinized Soils in Northwest China[J]. Spectroscopy and Spectral Analysis, 2016, 36(5): 1615 - 1622.
- [9] XU LU, WANG Zhi-Chun, NYONGESAH M J, et al. Determining soil water content of salt-affected soil using far-infrared spectra: laboratory experiment [J]. Journal of Applied Remote Sensing, 2015, 9: 095983-1-095983-15.
- [10] 陶兰花,塔西甫拉提·特依拜,买买提·沙吾提,等.土壤氯化钠含量高光谱估算模型研究[J].地理与地理信息科学,2014,30(3):46-50.
- [11] 陈东强,王让会.准噶尔盆地人工林地土壤全盐的高光谱反演[J].干旱区研究,2013,30:444-448.
- [12] NAWAR S, BUDDENBAUM H, HILL J. Estimation of soil salinity using three quantitative methods based on visible and near-infrared reflectance spectroscopy: a case study from Egypt [J]. Arabian Journal of Geosciences, 2015, 8: 5127 - 5140.
- [13] 刘娅,潘贤章,王昌昆,等.基于可见-近红外光谱的滨海盐土土壤盐分预测方法[J].土壤学报,2012,49:824-829.
- [14] 李晓明,韩霖昌,李娟.典型半干旱区土壤盐分高光谱特征反演[J].光谱学与光谱分析,2014,34(4):1081-1084.
- [15] 李微,李媛媛,田彦,等.基于包络线法的滨海滩涂 PLSR 盐分模型研究[J].海洋科学进展,2014,32(4):501-507.
- [16] 赵振亮,塔西甫拉提·特依拜,孙倩,等.土壤光谱特征分析及盐渍化信息提取——以新疆渭干河/库车河绿洲为例[J].地理科学进展,2014,33(2):280-288.
- [17] BUDDENBAUM H, STEFFENS M. The effects of spectral pretreatments on chemometric analyses of soil profiles using laboratory imaging spectroscopy[J]. Applied and Environmental Soil Science, 2012, 1 - 12.



- [18] 支月娥, 周培, 冯海玮, 等. 一种高效秸秆降解复合菌剂及其有机肥制备方法[P]. CN 102676492 A, 2012 - 05 - 30.
- [19] SONMEZ S, BUYUKTAS D, OKTUREN F, et al. Assessment of different soil to water ratios (1: 1, 1: 2.5, 1: 5) in soil salinity studies[J]. Geoderma, 2008, 144: 361 - 369.
- [20] AMEZKETA E. An integrated methodology for assessing soil salinization, a pre-condition for land desertification [J]. J Arid Environ, 2006, 67: 594 - 606.
- [21] 刘广明, 杨劲松, 姜艳. 江苏典型滩涂地区地下水及土壤的盐分特征研究[J]. 土壤, 2005, 37(2), 163 - 168.
- [22] 翁永玲, 戚浩平, 方洪宾, 等. 基于 PLSR 方法的青海茶卡 - 共和盆地土壤盐分高光谱遥感反演 [J]. 土壤学报, 2010, 47(6): 1255 - 1263.
- [23] GOLDSHLEGER N, CHUDNOVSKY A, BEN-BINYAMIN R. Predicting salinity in tomato using soil reflectance spectra [J]. International journal of remote sensing, 2013, 34: 6079 - 6093.
- [24] 褚小立, 袁洪福, 陆婉珍. 近红外分析中光谱预处理及波长选择方法进展与应用[J]. 化学进展, 2004, 16: 528 - 542.
- [25] 李民赞, 韩东海, 王秀. 光谱分析技术及其在应用 [M]. 北京: 科学出版社, 2006.
- [26] 褚小立. 化学计量学方法与分子光谱分析技术[M]. 北京: 化学工业出版社, 2011.
- [27] 江泽慧, 费本华, 杨忠. 光谱预处理对近红外光谱预测木材纤维素结晶度的影响[J]. 光谱学与光谱分析, 2007, 27: 435 - 438.
- [28] 何挺, 王静, 林宗坚, 等. 土壤有机质光谱特征研究[J]. 武汉大学学报 (信息科学版), 2006, 31: 975 - 979.
- [29] GAO X H, YANG Y, ZHANG W, et al. Visible-near infrared reflectance spectroscopy for estimating soil total nitrogen contents in the Sanjiang Yuan Regions, China: a case study of Yushu County and Maduo County, Qinghai province. SPIE Asia Pacific Remote Sensing: International Society for Optics and Photonics [C], 2014, p. 926310-0-12.
- [30] 于雷, 洪永胜, 耿雷, 等. 基于偏最小二乘回归的土壤有机质含量高光谱估算[J]. 农业工程学报, 2015, 31(4): 103 - 109.
- [31] WOLD S, SJÖSTRÖM M, ERILSSON L. PLS-regression: a basic tool of chemometrics [J]. Chemometr Intell Lab Syst, 2001, 58: 109 - 130.
- [32] 杨福芹, 冯海宽, 李振海, 等. 基于赤池信息量准则的冬小麦叶面积指数估算[J]. 农业机械学报, 2015, 46: 112 - 120.
- [33] 张辉, 李维炯, 倪永珍. 生物有机无机复合肥对土壤微生物活性的影响[J]. 农村生态环境, 2004, 20: 37 - 40.
- [34] MOREIRA L C J, DOS SANTOS TEIXEIRA A, GALVÃO L S. Laboratory salinization of Brazilian alluvial soils and the spectral effects of Gypsum[J]. Remote Sensing, 2014, 6: 2647 - 2663.
- [35] 任少亭, 浩折霞, 李蒙, 等. 新疆玛纳斯河流域不同水分含量盐渍化土壤光谱信息特征研究 [J]. 土壤通报, 2013, 44 (5): 1091 - 1095.
- [36] 彭杰, 刘焕军, 史舟, 等. 盐渍化土壤光谱特征的区域异质性及盐分反演[J]. 农业工程学报, 2014, 30(17): 167 - 174.
- [37] 张飞, 塔西甫拉提·特依拜, 丁建丽, 等. 塔里木河中游典型绿洲盐渍化土壤的反射光谱特征 [J]. 地理科学进展, 2012, 31(7): 921 - 932.
- [38] HOWARI F M, GOODELL P C, MIYAMOTO S. Spectral properties of salt crusts formed on saline soils [J]. Journal Of Environmental Quality, 2002, 31:1453 - 1461.

## Spectral Characteristics of Soil Salinity Based on Different Pre-processing Methods

ZHU Yun<sup>1,3</sup>, SHEN Guang-rong<sup>1,2,3\*</sup>, XIANG Qiao-qiao<sup>1,3</sup>, WU Yu<sup>1,3</sup>

(1. School of Agriculture and Biology, Research Center for Low-Carbon Agriculture, Shanghai Jiao Tong University, Shanghai 200240, China; 2. Key Laboratory of Urban Agriculture (South), Ministry of Agriculture, Shanghai 200240, China; 3. Shanghai Urban Forest Ecosystem Research Station of National Positioning and Observation, State Forestry Administration, Shanghai 200240, China)

**Abstract:** In this paper, soil salinity content (SSC) and its spectral reflectance were measured during the microbial remediation process of saline soil. The aim of the paper was to analyze and compare the effects of 34 pre-processing methods on spectral characteristics of saline soil during the remediation process. Partial least squared regression (PLSR) analysis was then used to predict SSC based on reflectance spectra by using full bands (400 ~ 1650 nm) and the optimal sensitive bands for 6 selected pre-processing methods. The results showed that spectral pre-processing methods had the advantage of enhancing the correlation of SSC and smoothed reflectance spectra, and the range of optimal sensitive bands was further focused. The derivative turned out to be the best pre-processing methods in this study, and the prediction accuracy of SGSD was the best in full bands. Compared to the raw reflectance spectra (R), the corresponding  $r$  and RMSEP of the predicted model were improved, respectively, from 0.537 and 1.928 to 0.823 and 1.256. Based on optimal sensitive bands of PLSR predicting models of SSC, SGSD(LogR) obtained more robust calibration and prediction accuracies than other pre-processing inversion models. The results obtained in this study provided a method and data reference for further quantitative analysis of SSC in saline soil quickly.

**Key words:** Saline soil; Microbial remediation; Pre-processing methods of spectra; PLSR [责任编辑: 孙福军]