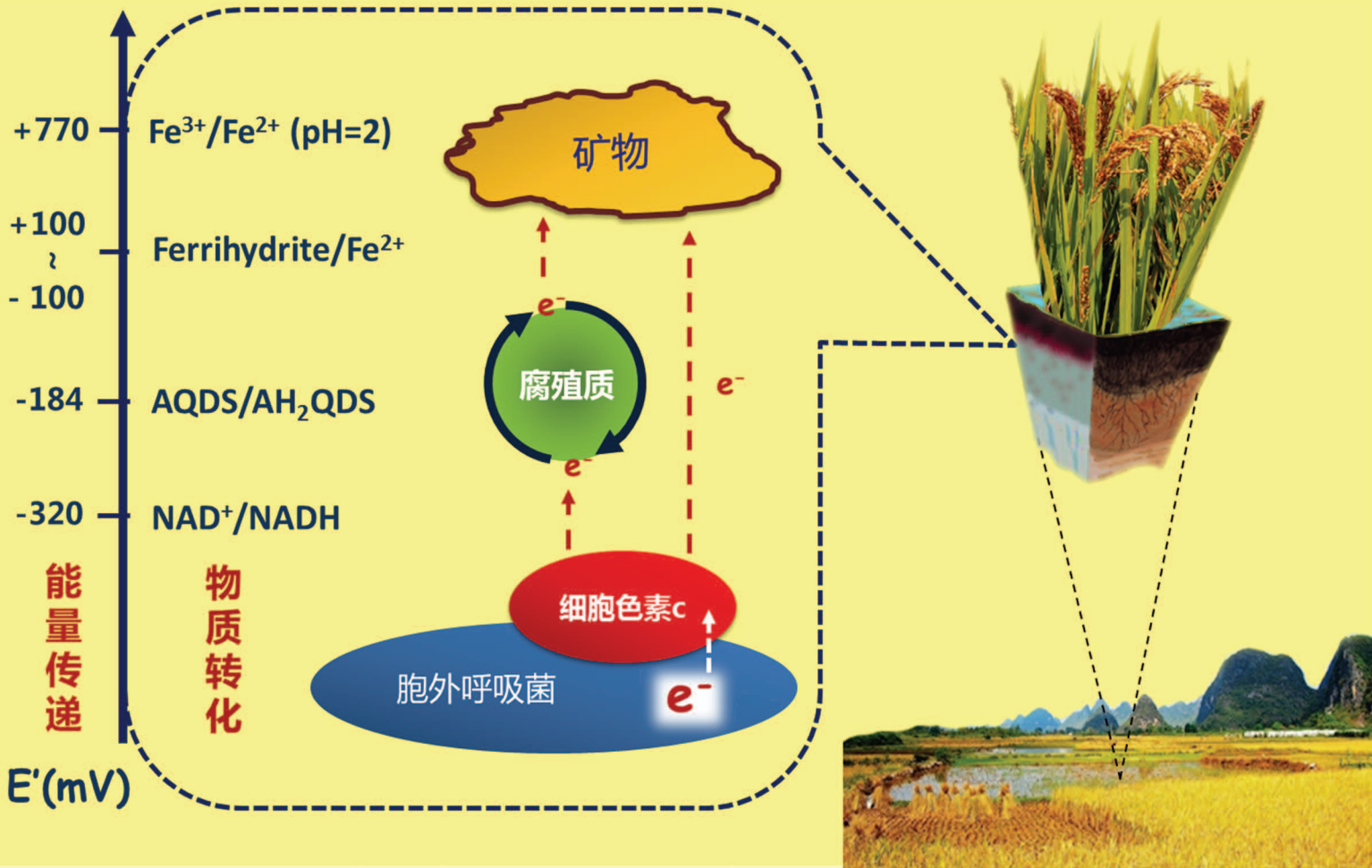


Acta Pedologica Sinica 土壤学报

Turang Xuebao



中国土壤学会 主办
科学出版社 出版

2016

第53卷 第2期

Vol.53 No.2



土壤学报

(Turang Xuebao)



第 53 卷 第 2 期 2016 年 3 月

目 次

综述与评论

- 土壤微生物—腐殖质—矿物间的胞外电子传递机制研究进展····· 吴云当 李芳柏 刘同旭 (277)
长期施肥对农田土壤氮素关键转化过程的影响····· 王 敬 程 谊 蔡祖聪等 (292)

新视角与前沿

- 土壤学不应忽视对作物土传病原微生物的研究····· 蔡祖聪 黄新琦 (305)

研究论文

- 西安少陵塬黄土—古土壤序列 S_3 剖面元素迁移及古气候意义····· 楚纯洁 赵景波 (311)
基于盲源分离的稀疏植被区土壤含盐量反演····· 刘 娅 潘贤章 石荣杰等 (322)
基于地类分层的土壤有机质光谱反演校正样本集的构建····· 刘艳芳 卢延年 郭 龙等 (332)
基于多分辨率遥感数据与随机森林算法的土壤有机质预测研究····· 王茵茵 齐雁冰 陈 洋等 (342)
鄂东南崩岗剖面土壤水分特征曲线及模拟····· 邓羽松 丁树文 蔡崇法等 (355)
放水冲刷对红壤坡面侵蚀过程及溶质迁移特征的影响····· 马美景 王军光 郭忠录等 (365)
汶川震区滑坡堆积体坡面土壤侵蚀率及水动力学参数研究····· 王仁新 何丙辉 李天阳等 (375)
咸水冻融灌溉对重度盐渍土壤水盐分布的影响····· 张 越 杨劲松 姚荣江 (388)
基于不同估算方法的贵州省土壤温度状况····· 陆晓辉 董宇博 涂成龙 (401)
拉萨灌丛草甸区土壤温度变化特征····· 巩玉玲 王兆锋 张懿锂等 (411)
砂土和黏土的颗粒差异对土壤斥水性的影响····· 杨 松 吴珺华 董红艳等 (421)
AQDS加速红壤性水稻土中DDT厌氧脱氯效应研究····· 刘翠英 王 壮 徐向华等 (427)
激发式秸秆深还对土壤养分和冬小麦产量的影响····· 赵金花 张丛志 张佳宝 (438)
臭氧污染对麦田土壤不同活性有机碳库的影响····· 寇太记 程相涵 张东亮等 (450)
黑土区水稻土有机氮组分及其对可矿化氮的贡献····· 丛耀辉 张玉玲 张玉龙等 (457)
水土保持措施对红壤缓坡地土壤活性有机碳及酶活性的影响····· 黄尚书 成艳红 钟义军等 (468)
祁连山青海云杉林叶片—枯落物—土壤的碳氮磷生态化学计量特征····· 赵维俊 刘贤德 金 铭等 (477)
基于核酸DNA/RNA同位素示踪技术的水稻土甲烷氧化微生物研究····· 郑 燕 贾仲君 (490)
适应玉米的溶磷细菌筛选及其对玉米生长的影响····· 梅新兰 闪安琪 蒋 益等 (502)
旱地红壤线虫群落对不同耕作年限的响应及指示意义····· 王明伟 刘雨迪 陈小云等 (510)
西藏“玉米田养鹅”模式下养分吸收与养分平衡特征····· 沙志鹏 张宇阳 王 超等 (523)
加工番茄连作对土壤理化性状及微生物量的影响····· 康亚龙 景 峰 孙文庆等 (533)

研究简报

- CTMAB对BS-12修饰膨润土的复配修饰模式····· 余 璐 孟昭福 李文斌等 (543)
不同机械改土方式对白浆土物理特性及酶活性的影响····· 孟庆英 张春峰 贾会彬等 (552)

信息

- 《土壤学报》2015年度审稿专家名录····· (560)

封面图片: 微生物胞外电子传递: 能量传递与物质转化 (由吴云当、李芳柏、刘同旭提供)

DOI: 10.11766/trxb201506020150

基于地类分层的土壤有机质光谱反演校正样本集的构建*

刘艳芳^{1, 2} 卢延年¹ 郭 龙¹ 肖丰涛¹ 陈奕云^{1, 2, 3†}

(1 武汉大学资源与环境科学学院, 武汉 430079)

(2 地球空间信息技术协同创新中心, 武汉大学, 武汉 430079)

(3 武汉大学苏州研究院, 江苏苏州 215123)

摘 要 以江汉平原滨湖地区不同土地利用类型的土壤样本为例, 比较了基于目标土壤理化性质的浓度梯度法、扩展的基于多种理化性质的综合法 (P-KS)、基于光谱信息的KS法、最邻近样本去除法 (reduce nearest neighbor samples, RNNS) 法和基于浓度分层并结合光谱信息的C-KS、C-RNNS法, 基于地类分层再结合上述方法, 构建具有不同层次土壤信息代表性的校正集, 采用偏最小二乘回归法, 建立土壤有机质可见光/近红外光谱反演模型。结果表明, 具有单一代表性的浓度梯度法、KS法、RNNS法难以建立适用模型; 具有光谱与理化性质二元代表性的C-KS方法模型预测精度得到了明显的提升, 相对分析误差 (ratio of performance to standard deviation, RPD) 为1.66; 考虑土地利用类型后, 浓度梯度法、RNNS法与C-KS法模型预测精度有明显的提升, RPD分别达到了1.84、1.51、1.75, 模型具有良好的适用性。说明具有多层次土壤信息代表性的校正集构建方法对提高土壤有机质可见光/近红外光谱反演模型的适用性具有较好作用。

关键词 理化性质代表性; 光谱代表性; 地类分层; 偏最小二乘回归

中图分类号 S127 **文献标识码** A

土壤有机质 (soil organic matter, SOM) 是土壤的重要组分, 是土壤肥力的重要指标, 同时也是全球碳循环重要的源和汇^[1-2]。可见光/近红外光谱分析技术因其高效、便捷和环保等特点, 已逐渐发展成包括土壤有机质含量在内的多种土壤理化参数分析的常用方法^[3]。国内外学者选取了不同研究区, 针对不同类型的土壤^[4-5], 从光谱波段与SOM的相关关系入手^[6-7], 研究了光谱预处理和建模方法^[8-10]对SOM预测精度的影响, 以期找到最适模型, 而在校正样本集选择方面的研究很少。

校正样本集的选择对土壤有机质光谱反演模型

稳定性及预测精度非常重要^[11], 样本在SOM大小和分布上要具有代表性, 能代表土壤属性的变异情况, 还要能代表土壤光谱的变异情况^[12]。目前常用的样本选择方法主要有基于土壤理化参数的浓度梯度法, 基于光谱信息的KS法、Duplex法和GN距离法 (global H and neighborhood H) 等^[13]。浓度梯度法按照待测性质排序, 按一定规则选取样本, 在待测性质上具有代表性但没有考虑光谱特征。KS法将光谱差异大的样本选入校正集, 但光谱信息组成复杂, 光谱是土壤属性的综合体现, 校正集代表性并不是基于待测性质的代表性。

* 公益性行业科研专项项目 (201412023)、卫星测绘技术与应用国家测绘地理信息局重点实验室经费项目 (Klamta-201401) 资助 Supported by the Special Fund in the Public Interest (No. 201412023), Key Laboratory of Satellite Mapping Technology and Application, National Administration of Surveying, Mapping and Geoinformation (No. Klamta-201401)

† 通讯作者 Corresponding author, E-mail: chenyy@whu.edu.cn

作者简介: 刘艳芳 (1959—), 女, 湖北武汉人, 博士, 教授, 主要从事地理信息应用等研究。E-mail: yfliu610@163.com

收稿日期: 2015-06-02; 收到修改稿日期: 2015-08-11

在回归分析过程中，局部建模样本对回归模型的贡献权重不一样，具有某一土壤特征的样本可能大量入选校正集，导致预测误差分布存在一定的空间结构^[14]，这为基于土壤特征细分和具有多元代表性的校正集样本选择方法能够提高模型预测精度提供理论基础。相对于传统的基于单一理化性质代表性或光谱代表性的样本选择方法，近期研究开始关注综合考虑理化性质和光谱信息的样本选择方法^[15]。Roberto等^[16]提出了SPXY法（sample set partitioning based on joint x-y distance），在KS方法的基础上考虑理化性质参数， p 、 q 两样本间的距离 $d_{xy}(p, q)$ 是在基于光谱信息的欧氏距离的基础上引入了理化性质，同时为了确保样本在 x 和 y 空间具有相同的权重，分别除以它们的最大值。

刘伟等^[14]提出了Rand-KS方法，以汽油中碳酸二甲酯含量和二甲亚砷溶液为研究对象，按浓度分层，再在各层使用KS选择校正集样本，对光谱反演过程中的均质化现象进行了改进，浓度和光谱上同时具有一定代表性。然而，关于土壤信息细分和多元代表性校正集样本选择对预测模型精度的影响研究尚未开展，地理信息代表性也很少在校正集样本构建中体现。

本文使用土壤信息（包括土壤理化性质信息、光谱信息和地理信息）按特征分层选取校正集，使得校正集样本具有更全面的代表性。设计了基于目标土壤理化性质的浓度梯度法、扩展的基于多种理化性质的P-KS法、基于光谱信息的KS法、RNNS法、基于浓度分层并结合光谱信息的C-KS、C-RNNS法，基于地类分层再结合上述方法的十种校正集构建方法，并比较其对模型预测精度的影响。P-KS法综合考虑对土壤光谱影响权重较大的几种参数SOM、Fe、N、P、容重，以KS算法挑选具有多种理化性质代表性的样本组成校正集；C-KS与C-RNNS法将SOM按浓度范围等分为 m 个区间，在每个区间按KS方法和RNNS方法选择三分之二样本形成校正集；基于地类分层的方法，按照旱地、水田、其他划分为三类，每一类中分别结合浓度梯度法、KS、RNNS和C-KS法选取具有SOM分布或光谱代表性的校正集，再合并形成具有地类代表性的校正集。比较了基于地类分层和具有多元代表性的校正集构建方法对模型预测精度的影响，从校正样本集构建的角度寻找建立优质模型的方法。

1 材料与方法

1.1 样品采集与制备

研究区位于湖北省江汉平原，江汉平原由长江和汉江冲积而成，分布着由河流—湖泊—沼泽—水稻田组成的湿地系统，受人类长期农业活动影响，生态稳定性差，生产力波动大，SOM水平较低，对江汉平原SOM的空间动态监测对于滨湖地区土地资源管理有重要意义^[17-18]。共采集270个样本（Dataset0， $n=270$ ），按照土地利用类型分为三类：水田（Dataset1， $n=114$ ）、旱地（Dataset2， $n=129$ ）、其他（Dataset3， $n=27$ ）；水田作物为水稻，旱地作物有玉米、棉花等，其他包含林地、园地等，分布零散，多为水杉、圃地。土壤样本经过风干、研磨、过100目筛后，每个样品分为两份，一份用于光谱测量，一份用于化学分析。

1.2 实验室分析及光谱获取

光谱测量采用ASD FieldSpec3地物光谱仪，将样品平铺在托盘内，且样品厚度不小于5mm，在暗室内用光源为50W的卤光灯照射，光照角度 45° ，光源距样本中心40cm，光谱仪探头位于土壤表面垂直上方15cm。测试之前先用白板校正，测试过程中每测十个样本校正一次，以保证测量稳定。每个样本采集10条光谱曲线，进行算术平均得到各土样的反射光谱数据。

1.3 光谱预处理

在数据分析之前，先去除噪声较大的边缘波段，保留信噪比较高的400~2350nm波段。预处理方法对模型至关重要，目前研究主要通过多次建模来挑选最优的预处理方法^[19]，缺少关于何种情况选用何种预处理方法的研究^[20]。样本光谱曲线及去除包络线光谱曲线如图1所示，固体样本光谱采集过程均会产生散射，本文采用多元散射校正（multiplicative scatter correction, MSC）方法处理；光谱在近红外部分存在一定的锯齿状噪声，去除包络线后，谱段范围有小而杂乱的吸收谷，会对反射峰位置的选取和模型造成干扰^[21]，本文选取Savitzky-Golay（SG）平滑方法对曲线平滑去噪，45点平滑数。预处理顺序为log10变换、SG平滑、MSC，最后进行均值中心化处理（mean center, MC）。

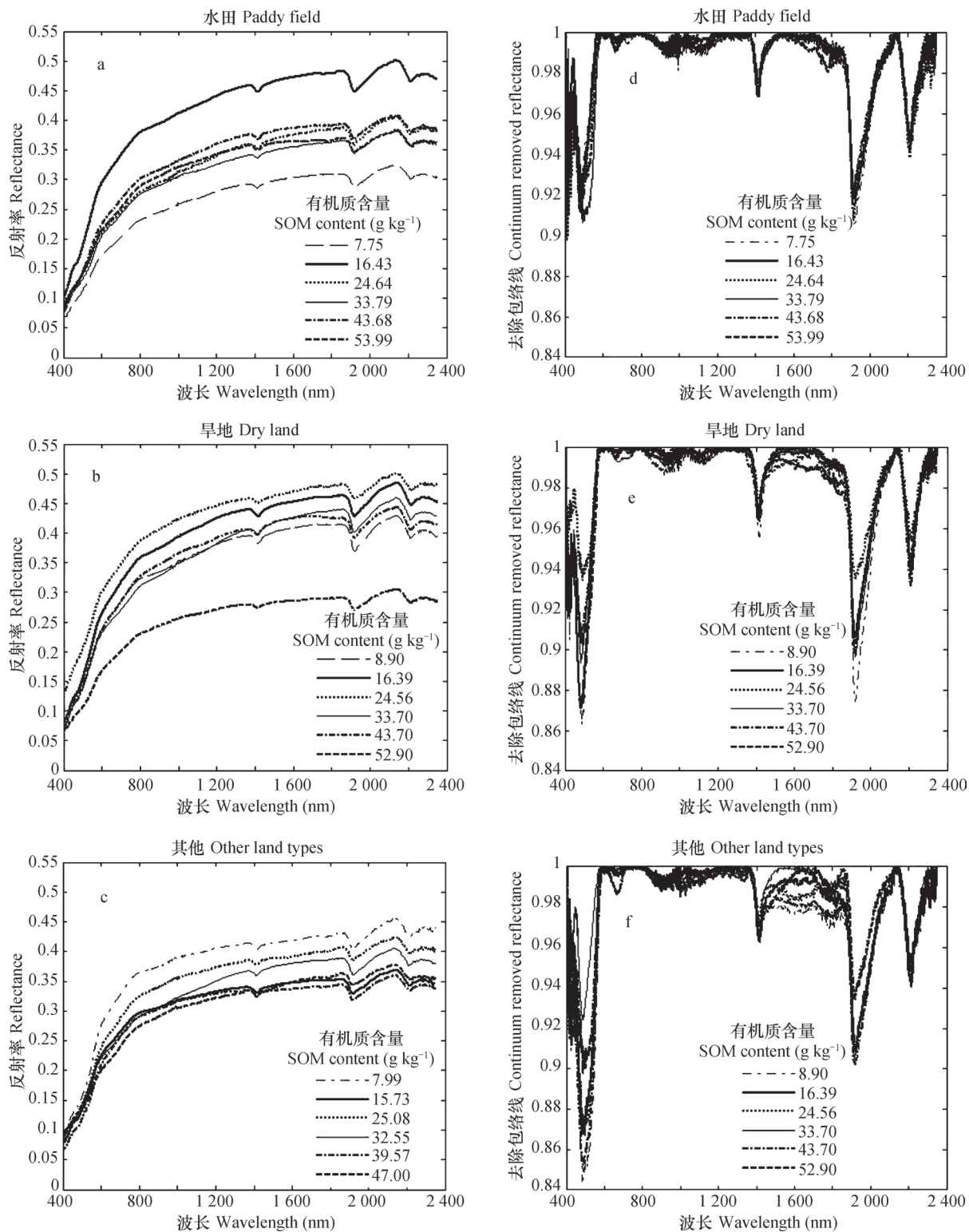


图1 三种地类的原始光谱曲线 (a、b、c) 与去除包络线光谱曲线 (d、e、f)

Fig. 1 Raw spectral curves (a, b and c) and continuum-removed spectral curves (d, e and f) of soil samples relative to land use types

1.4 实验策略设计

本文围绕土壤理化性质信息、光谱信息和地理空间信息，设计了十种具有多元土壤信息代表性的校正集构建方法，并比较了其优劣。基于土壤理化性质的样本选择方法主要是浓度梯度法，将样本按SOM大小排序，每三个样本中选择第一个和第三个组成校正集（Dataset C）。扩展浓度梯度法，综合考虑对土壤光谱影响权重较大的几种参数SOM、Fe、N、P、容重^[22]，以KS算法挑选具有多种理化性质代表性的样本组成校正集（Dataset P-KS）。基于土壤光谱信息的样本选择方法常用的是KS算法，使用欧氏距离挑选在光谱空间具有代表性的样本（Dataset KS）。相对于KS法基于最大距离选择样本，本文还采用RNNS算法，计算出两个样本的最邻近欧式距离，随机剔除最邻近样本中的一个，直至达到数量形成建模样本（Dataset RNNS）。

土壤理化性质信息和光谱信息结合的方法，将SOM按浓度范围等分为六个区间，在每个区间按KS方法和RNNS方法选择三分之二样本形成校正集（Dataset C-KS, Dataset C-RNNS），使得校正集在各个浓度区间均具有代表性。基于地类分层的样本选择方法，将样本按照土地利用类型分为水田、水浇地、其他，不同土地利用类型分别以浓度梯度法、KS法、RNNS法、C-KS法建立校正集样本，然后综合形成校正集（Dataset L-C, Dataset L-KS, Dataset L-RNNS, Dataset L-C-KS）。

1.5 模型建立与验证

偏最小二乘回归（Partial Least Squares Regression, PLSR）是土壤有机质高光谱反演建模中广泛应用并且较为成熟的建模方法^[23]，其公式为（1），式中 y 代表目标参数， x 代表土壤光谱参数， β_i （ $i=1, 2, \dots, k$ ）表示 y 关于 x_i 的回归系数。

$$\hat{y} = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (1)$$

本文采用舍一交叉验证，通过确定最小的舍一交叉验证均方根误差 $RMSE_{CV}$ 来确定主因子个数，从而确定最优模型。回归预测模型的稳定性通过模型舍一交叉验证的决定系数 R_{CV}^2 和均方根误差 $RMSE_{CV}$ 来检验；模型的预测精度基于独立验证集，通过计算 R_p^2 、预测均方根误差 $RMSE_p$ 和相对分析误差RPD来检验， $RMSE_p$ 越小， R_p^2 越大，说明模型预测能力越好。此外，当 $RPD < 1.0$ 时，认

为模型不具有预测能力，不适用于SOM的高光谱估算； $1.0 < RPD < 1.4$ 时，则认为模型较差，只能区分SOM数值的高低； $1.4 < RPD < 1.8$ 时，模型较好，可以用来进行SOM的粗略估算； $1.8 < RPD < 2.0$ 代表模型质量良好，可以用于SOM的定量预测； $2.0 < RPD < 2.5$ 代表模型质量很好； $RPD > 2.5$ 代表模型质量极好^[24-25]。样本选取和模型的建立采用Matlab2012b软件和MatlabPLStoolbox工具箱实现。

2 结 果

2.1 样本集的统计特征

图2为各数据集的直方图和箱图。由Dataset0可知，总样本SOM含量在 1.44 g kg^{-1} 至 59.86 g kg^{-1} 之间，平均值为 26.78 g kg^{-1} ，呈正偏态，土壤有机质含量整体上较低，有一个异常值点。由Dataset1、2、3可以看出总样本集的组成，其中水田SOM含量明显较高，均值为 32.53 g kg^{-1} ，接近正态分布；旱地SOM含量明显较低，均值为 22.15 g kg^{-1} ，而且分布较为集中在 10 g kg^{-1} 至 25 g kg^{-1} 之间，并出现多个异常值，呈明显的正偏态；其他地类SOM水平也较低，均值为 24.59 g kg^{-1} 。水田SOM均值要明显高于旱地，且数据分布区间较宽。不同土地利用方式SOM数据分布的差异性为基于地类分层的校正集样本选择提供基础。本文保留异常值建模，在全局范围内异常值是不存在的，异常值可能代表具有某一类特征的土壤样本。

十种样本选择方法的校正集样本分布具有相似性，SOM在含量和分布上均对原样本集具有一定的代表性。Dataset C分布与总样本集近似一致；Dataset P-KS法在SOM高值部分数量变多，平均值和标准差较大，正偏态减弱。Dataset KS与总样本集类似；Dataset RNNS中间样本分布更均匀，平均值、标准差较大，偏度较小。Dataset C-KS、Dataset C-RNNS与总样本分布类似，最小值没有入选Dataset C-KS。加入土地利用方式后，各校正集表现与不考虑地类时类似，Dataset L-RNNS与Dataset RNNS在 $10 \sim 20 \text{ g kg}^{-1}$ 之间的样本频数差异较大；Dataset L-C-KS与Dataset C-KS在高值样本选取上差异较大。

各样本集SOM统计指标类似，但实际样本组成的差异明显，样本组成在土地利用类型分布上具有

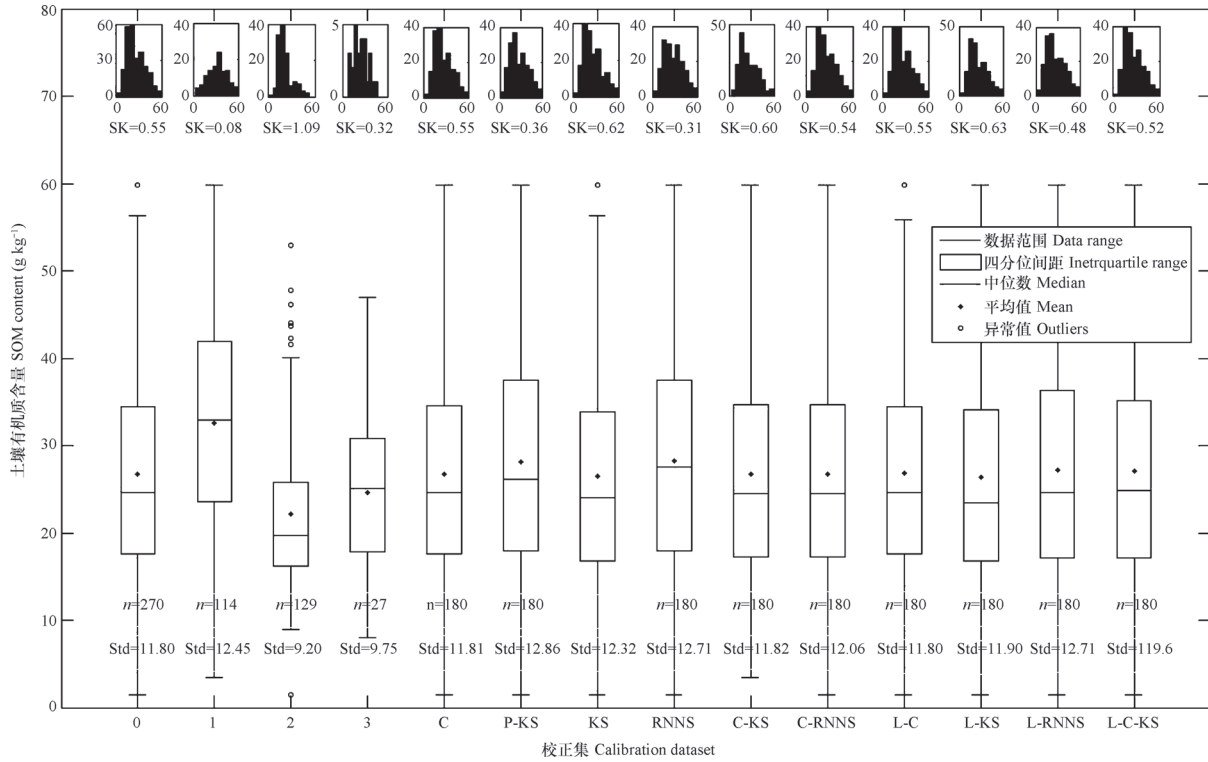


图2 各数据集SOM含量分布箱图和直方图 (“SK”为偏度, “n”为样本数, “Std”为标准差)

Fig.2 Box plots and histograms of soil organic matter content distribution in each sample set (“SK” denotes skewness, “n” denotes the number of samples, and “Std” denotes standard derivation)

差异性。基于地类分层抽样,使得样本在每个地类的SOM浓度上或者光谱上具有代表性,得到的校正集可能更具代表性。

2.2 最优模型的确定

建模和验证结果相关指标如表1。基于理化性质信息,浓度梯度法校正集建模结果 R^2_{CV} 为0.64, R^2_p 为0.50, RPD为1.38,模型稳定性较好,预测能力较差;SOM、Fe、N、P、容重五种理化参数以KS方法选择的校正集, R^2_{CV} 为0.61, R^2_p 为0.55, RPD为1.47,能够实现对江汉平原SOM定量估算,但需要改进。基于光谱信息,KS法 R^2_{CV} 为0.67, R^2_p 为0.48, RPD为1.27,模型稳定性较好,但是预测性差;RNNs法 R^2_{CV} 为0.67, R^2_p 为0.42, RPD为1.16,模型稳定性与预测能力较差。仅基于欧式距离的样本选择方法还不足以建立具有代表性的校正集。综上,常用的样本选择算法很难建立适用的预测模型。

基于土壤理化性质和光谱信息结合的方法,

C-KS法 R^2_{CV} 为0.54, R^2_p 为0.66, RPD为1.66; C-RNNs方法 R^2_{CV} 为0.70, R^2_p 为0.48, RPD为1.30。C-KS法具有较好的稳定性与预测能力,能够用于滨湖地区SOM定量预测;C-RNNs法稳定性好,预测性差,不适用于SOM定量预测,但相对于RNNs法模型预测精度有很大提升。

基地类分层选择样本,L-C法 R^2_{CV} 为0.51, R^2_p 为0.70, RPD为1.84,具有较好的稳定性和最好的预测能力,较具有单一代表性的浓度梯度法有很大提升。L-KS法 R^2_{CV} 为0.66, R^2_p 为0.44, RPD为1.27,预测精度未能改进。L-RNNs法 R^2_{CV} 为0.56, R^2_p 为0.59, RPD为1.51,对于RNNs具有很大优化,能够用于定量估测。同时考虑土地利用类型、SOM含量、光谱特征三方面代表性,L-C-KS法 R^2_{CV} 为0.54, R^2_p 为0.68, RPD为1.75,具有较好的模型稳定性和预测效果。考虑土地利用类型后,浓度梯度法、RNNs、C-KS法构建的校正样本集建模预测精度有明显的提高。

表1 不同校正集构建方法建模和预测结果

Table 1 Models using calibration sets built up with different sample selection methods and their predictions

信息类型 Information type	选择方法 Selection method	主成分数Principal component number	舍一交叉验证Leave-one out cross-validation		预测验证Prediction validation		相对分析误差Ratio of performance to standard deviation
			R^2_{CV}	RMSE _{CV}	R^2_p	RMSE _p	
理化性质 ^①	C	14	0.64	7.08	0.50	8.58	1.38
	P-KS	14	0.61	7.62	0.55	7.54	1.47
光谱信息 ^②	KS	17	0.67	7.14	0.48	8.47	1.27
	RNNS	13	0.67	7.37	0.42	7.84	1.16
理化性质和光谱信息 ^③	C-KS	14	0.54	8.06	0.64	7.13	1.66
	C-RNNS	17	0.70	6.61	0.48	8.74	1.30
理化性质、光谱信息和地理空间信息 ^④	L-C	14	0.51	8.36	0.70	6.34	1.84
	L-KS	19	0.66	7.14	0.44	8.64	1.27
	L-RNNS	14	0.56	8.47	0.59	6.47	1.51
	L-C-KS	14	0.54	8.17	0.68	6.58	1.75

①Physical and chemical properties, ②Spectral information, ③Physical and chemical properties and spectral information, ④General information

3 讨论

土壤光谱的影响因素包括直接的土壤内在理化性质属性和间接的外在环境要素，其在地理空间上均存在空间异质性，本文在校正样本集构建过程中引入土壤理化性质、光谱信息、地理信息三方面的信息，构建具有多层次代表性的校正集，目的在于改善校正集中光谱影响要素分布不均匀而导致的单一代表性校正集对总体样本变异覆盖度小的不足，使校正样本集能够代表研究区总体以及研究区局部环境变异。

浓度梯度法与KS法是可见光/近红外光谱反演中常用的样本选择方法，但在江汉平原滨湖地区的SOM反演中未取得好的效果，RNNS法也未能建立可用模型，原因可能在于土壤贫瘠、干扰因素多，有研究表明当SOM含量在 20 g kg^{-1} 以下时，其光谱特征极易被其他成分（如铁氧化物）光谱特征所掩盖^[26]。同时单一代表性校正集对总体变异性覆盖度较小，导致上述校正集构建策略下模型的预测精度偏低，滨湖地区土壤有机质光谱反演还有待进一步研究。P-KS法综合多种理化性质选取校正集，模型精度有所提升，但是P-KS法对几种入

选的参数均以欧氏距离进行量度，不同理化性质对光谱的影响程度或比例不明确，未考虑权重因素。该方法需要提前获取土壤多种参数，可以结合具有多种参数的区域土壤光谱库进行校正集的构建。

浓度梯度与光谱结合的C-KS法在浓度和光谱空间上具有双重代表性，建立了精度较好的模型，说明了总样本集细分选取具有多元代表性的校正集具有更好的预测能力。基于地类分层抽样，浓度梯度法、KS、C-KS取得了更好的预测精度，证明了地理空间信息用于更具代表性的校正集构建的可行性和潜在必要性。地理空间要素影响土壤的形成与演化，进而在一定程度上影响土壤的理化性质；土壤的反射光谱信息是具有光谱响应特征的土壤内在理化性质的外在反映，因而使用土壤光谱信息能够反演土壤理化性质信息，地理空间信息、土壤理化信息和土壤光谱信息是土壤信息的三个方面，是相互联系的整体。同种土地利用类型以及相同的农业利用和管理方法会使得土壤理化性质和光谱信息及其之间的关系逐渐趋同，这为基于地类分层或者更加细分的样本选择方法在土壤反演建模中的应用提供了可能。

4 结论

本文指出了土壤光谱校正集选取主要基于土壤理化性质或光谱信息,具有单一代表性,比较了具有多元代表性的校正集构建方法的优劣,探讨了基于地类分层的校正集构建方法对模型的影响。结果表明,浓度梯度法、KS法、RNNS法的样本选择算法分别在SOM分布或光谱上具有代表性,但是未能建立适用的模型。基于多种土壤理化性质的P-KS法较单一性质的浓度梯度法更为有效。浓度梯度法与KS和RNNS结合的方法在土壤理化性质和光谱上具有代表性,模型预测精度均有提高,达到了评估预测标准,具有多层次代表性的校正集构建方法具有较好的潜力。考虑土地利用类型后,浓度梯度法、RNNS法与C-KS法构建的校正集模型预测精度有明显的提升,其中基于地类分层结合浓度梯度法(L-C)具有最优的预测精度,是最佳的样本构建方法。基于地类分层,构建具有多元代表性的校正集方法能够有效提高江汉平原滨湖地区土壤有机质光谱反演模型的预测精度。

参考文献

- [1] Batjes N H. Total carbon and nitrogen in the soils of the world. *European Journal of Soil Science*, 1996, 47 (2): 151—163
- [2] 龙军, 张黎明, 毛艳玲, 等. 福建省不同耕地土壤和土地利用类型对“碳源/汇”的贡献差异研究. *土壤学报*, 2013, 50 (4): 664—674
Long J, Zhang L M, Mao Y L, et al. Study on the difference of contributions made by various cropland soil and land use types to carbon source/sink in Fujian Province, China (In Chinese). *Acta Pedologica Sinica*, 2013, 50 (4): 664—674
- [3] Gerbbers R, Adamchuk V I. Precision agriculture and food security. *Science*, 2010, 327: 828—831
- [4] 宋海燕, 秦刚, 韩小平, 等. 基于可见光谱的不同质地土壤有机质快速测定. *农业机械学报*, 2012, 43 (7): 69—72
Song H Y, Qin G, Han X P, et al. Rapid prediction of soil organic matter by using visible infrared spectral technology (In Chinese). *Transactions of the Chinese Society for Agricultural Machinery*, 2012, 43 (7): 69—72
- [5] 卢艳丽, 白由路, 杨俐苹, 等. 东北平原不同类型土壤有机质含量高光谱反演模型同质性研究. *植物营养与肥料学报*, 2011, 17 (2): 456—463
Lu Y L, Bai Y L, Yang L P, et al. Homogeneity of retrieval models for soil organic matter of different soil types in Northeast Plain using hyperspectral data (In Chinese). *Plant Nutrition and Fertilizer Science*, 2011, 17 (2): 456—463
- [6] 纪文君, 史舟, 周清, 等. 几种不同类型土壤的VIS-NIR光谱特性及有机质相应波段. *红外与毫米波学报*, 2012, 31 (3): 277—282
Ji W J, Shi Z, Zhou Q, et al. VIR-NIR reflectance spectroscopy of the organic matter in several types of soils (In Chinese). *Journal of Infrared and Millimeter Waves*, 2012, 31 (3): 277—282
- [7] 郭斗斗, 黄绍敏, 张水清, 等. 多种潮土有机质高光谱预测模型的对比分析. *农业工程学报*, 2014, 30 (21): 192—200
Guo D D, Huang S M, Zhang S Q, et al. Comparative analysis of various hyperspectral prediction models of fluvo-aquic soil organic matter (In Chinese). *Transactions of the Chinese Society of Agricultural Engineering*, 2014, 30 (21): 192—200
- [8] 薛利红, 周鼎浩, 李颖, 等. 不同利用方式下土壤有机质和全磷的可见近红外高光谱反演. *土壤学报*, 2014, 51 (5): 993—1002
Xue L H, Zhou D H, Li Y, et al. Prediction of soil organic matter and total phosphorus with VIS-NIR hyperspectral inversion relative to land use (In Chinese). *Acta Pedologica Sinica*, 2014, 51 (5): 993—1002
- [9] Rossel R A V, Behrens T. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 2010, 158 (1/2): 46—54
- [10] Shi T, Cui L, Wang J, et al. Comparison of multivariate methods for estimating soil total nitrogen with visible/near infrared spectroscopy. *Plant and Soil*, 2013, 336 (1/2): 363—375
- [11] Daszykowski M, Walczak B, Massart D L. Representative subset selection. *Analytica Chimica Acta*, 2002, 468: 91—103
- [12] Liu Y L, Jiang Q H, Fei T, et al. Transferability of a visible and near-infrared model for soil organic matter estimation in riparian landscapes. *Remote Sensing*, 2014, 6 (5): 4305—4322
- [13] 张小超, 吴静珠, 徐云. 近红外光谱分析技术及其在现代农业中的应用. 北京: 电子工业出版社, 2012
Zhang X C, Wu J Z, Xu Y. Near infrared spectroscopy technology and its application in modern agriculture (In Chinese). Beijing: Publishing House of Electronic Industry, 2012

- [14] 刘伟, 赵众, 袁洪福, 等. 光谱多元分析校正集和验证集样本分布优选方法研究. 光谱学与光谱分析, 2014, 34 (4) : 947—951
Liu W, Zhao Z, Yuan H F, et al. An optimal selection method of samples of calibration set and validation set for spectral multivariate analysis (In Chinese) . Spectroscopy and Spectral Analysis, 2014, 34 (4) : 947—951
- [15] 詹雪艳, 赵娜, 林兆洲, 等. 校正集选择方法对于积雪草总苷中积雪草苷NIR定量模型的影响. 光谱学与光谱分析, 2014, 34 (12) : 3267—3272
Zhan X Y, Zhao N, Lin Z Z, et al. Effect of algorithms for calibration set selection on quantitatively determining asiaticoside content in centella total glucosides by near infrared spectroscopy (In Chinese) . Spectroscopy and Spectral Analysis, 2014, 34 (12) : 3267—3272
- [16] Roberto K H G, M á rio C U A, Gledson E J, et al. A method for calibration and validation subset partitioning. Talanta, 2005, 67 (4) : 736—740
- [17] 王学雷, 吴宜进. 江汉平原湿地系统的退化与生态恢复重建. 华中农业大学学报, 2004, 23 (4) : 467—471
Wang X L, Wu Y J. Degeneration and ecological rehabilitation of wetlands in Jianghan Plain (In Chinese) . Journal of Huazhong Agricultural University, 2004, 23 (4) : 467—471
- [18] 刘刚, 沈守云, 闫文德, 等. 洪湖湿地生态系统土壤有机碳及养分含量特征. 生态学报, 2011, 31 (24) : 7625—7631
Liu G, Shen S Y, Yan W D, et al. Characteristics of organic carbon and nutrient content in five soil types in Honghu wetland ecosystems (In Chinese) . Acta Ecologica Sinica, 2011, 31 (24) : 7625—7631
- [19] 申艳, 张晓平, 杨学明, 等. 光谱预处理及不同光谱区对黑土有机碳近红外光谱模型的影响. 土壤学报, 2010, 47 (5) : 1006—1012
Shen Y, Zhang X P, Yang X M, et al. Effects of data pretreatment and spectrum bands on near infrared spectroscopy model of soil organic carbon in black soil (In Chinese) . Acta Pedologica Sinica, 2010, 47 (5) : 1006—1012
- [20] Asmund R. Pre-processing in vibrational spectroscopy—when, why and how. Analytical Methods, 2014, 6: 7124—7129
- [21] 黄明祥, 王珂, 史舟, 等. 土壤高光谱噪声过滤评价研究. 光谱学与光谱分析, 2009, 29 (3) : 722—725
Huang M X, Wang K, Shi Z, et al. Quantitative evaluation of soil hyperspectral denoising with different filters (In Chinese) . Spectroscopy and Spectral Analysis, 2009, 29 (3) : 722—725
- [22] 郑光辉. 江苏部分地区土壤属性高光谱定量估算研究. 南京: 南京大学, 2010
Zheng G H. Study on the quantitative estimating of soil properties with hyper-spectrum in some parts of Jiangsu Province (In Chinese) . Nanjing: Nanjing University, 2010
- [23] 沈掌泉, 卢必慧, 单英杰, 等. 基于变量选择的偏最小二乘回归法和田间行走式近红外光谱进行土壤含量测定研究. 光谱学与光谱分析, 2013, 33 (7) : 1775—1780
Shen Z Q, Lu B H, Shan Y J, et al. Study on soil carbon estimation by on-the-go near-infrared spectra and partial least squares regression with variable selection (In Chinese) . Spectroscopy and Spectral Analysis, 2013, 33 (7) : 1775—1780
- [24] Shi T, Cui L, Wang J, et al. Comparison of multivariate methods for estimating soil total nitrogen with visible/near infrared spectroscopy. Plant and Soil, 2013, 336 (1/2) : 363—375
- [25] Rossel R A V, McGlynn R N, McBratney A B. Determining the composition of mineral-organic mixes using UV-VIS-NIR diffuse reflectance spectroscopy. Geoderma, 137: 70—82
- [26] Krishnan P, Alexander J D, Butler B J, et al. Reflectance technique for predicting soil organic matter. Soil Science Society of America Journal, 1980, 44 (6) : 1282—1285

Construction of Calibration Set Based on the Land Use Types in Visible and Near-Infrared (VIS-NIR) Model for Soil Organic Matter Estimation

LIU Yanfang^{1, 2} LU Yannian¹ GUO Long¹ XIAO Fengtao¹ CHEN Yiyun^{1, 2, 3*}

(1 School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China)

(2 Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China)

(3 Suzhou Institute of Wuhan University, Suzhou, Jiangsu 215123, China)

Abstract Soil organic matter (SOM) is not only an important indicator of soil fertility but also an important source and sink in the global carbon cycle. Therefore, it is essential to acquire the information of SOM for soil management. The visible and near-infrared (VIS-NIR) reflectance spectroscopy technique, known as a novel, rapid, accurate, environment-friendly and efficient approach when compared with conventional laboratory analyses, is a promising one to acquisition of soil properties. Construction of a calibration set is key to use of VIS-NIR quantitative analysis in building up a quality prediction model. Conventionally, selection of samples for the calibration set is based on soil physical and chemical properties or soil spectral information, like the concentration gradient method (C) and Kennard - Stone (KS) method, which are able to select samples that may be representative of physical and chemical properties or spectra, but not of geographical space and multivariate information. Impacts of the shortages on prediction accuracy of the model have rarely been reported. The aim of this paper is to explore how sample selection methods affect accuracy of the VIS-NIR reversion model in estimation of SOM, using soil samples collected from lands under different types of land use in the riparian areas of the Jiangnan Plain. A total of 270 soil samples were collected, air dried and ground to pass a 2 mm sieve, for analysis of VIS-NIR spectra using a FieldSpec3 spectrometer. The spectral curves were preprocessed with log₁₀, Savitzky-Golay (SG), multiplicative scatter correction (MSC) and mean center (MC).

A total of four categories of ten sample selection methods based on multivariate soil information were proposed for constructing calibration sets. The first category, including the concentration gradient method and the method adopted several properties (P-KS), depends on soil physical and chemical properties; the second category, including the KS method and the Reduce on Neighbor Samples (RNNS) method, is based on spectral information; the third category, including the C-KS and C-RNNS methods, combines soil physic-chemical properties with spectral information; and the fourth category uses land use type hierarchy in combination with all the aforementioned methods. The P-KS method takes into comprehensive account parameters, like SOM, Fe, N, P and bulk density (BD), that may be quite high in weight of impacts on soil spectra and uses KS algorithm to select soil samples representative of a variety of physical-chemical properties for construction of the calibration set. The C-KS and C-RNNS methods divide SOM concentration into six levels, from each of which two-thirds of the samples were selected using the KS and RNNS methods to form the calibration sets. The methods based on land use type hierarchy divide the entire sample set into three categories, namely dry land, paddy field and the others. For each category, soil samples representative of SOM distribution or soil spectra were selected in combination with the concentration gradient method, KS, RNNS and C-KS, separately to form calibration sets, which were then merged into a calibration set representative of land use type.

On such a basis, a partial least squares regressions (PLSR) model was established, showing that in the first and second categories, the models with calibration sets formed with the C, KS and RNNS methods, representative of SOM distribution or soil spectra singularly, were not so good in prediction accuracy; and those with the P-KS method were much better, with determination coefficient for prediction (R_p^2) being 0.55, root mean squared error of prediction ($RMSE_p$) being 7.54 and ratio of performance to standard deviation (RPD) being 1.47. The models with calibration sets formed with the C-KS method, representative of both physical and chemical properties and spectra, were good in accuracy with R_p^2 being 0.64, $RMSE_p$ being 7.13 and RPD being 1.66. The inclusion of land use type in forming calibration sets, greatly improved the models using the C, RNNS and C-KS methods in prediction accuracy, bring R_p^2 up to 0.70, 0.59 and 0.68, $RMSE_p$ to 6.34, 6.47 and 6.58, and RPD to 1.84, 1.84 and 1.51, respectively. It is therefore, quite obvious that the use of calibration sets formed with soil samples representative of multi-layers of soil information can improve the models in prediction accuracy. The L-C method has turned out to be the best method for sample selection in construction of calibration sets for VIR-NIR models for prediction of soil organic matter contents in the riparian areas of the Jiangnan Plain.

Key words Representativeness of physical and chemical properties; Representativeness of spectrum; Land use type; Partial Least Squares Regression (PLSR)

(责任编辑：檀满枝)

CONTENTS

Reviews and Comments

- Mechanism of Extracellular Electron Transfer among Microbe-Humic-Mineral in Soils: A Review WU Yundang, LI Fangbai, LIU Tongxu (290)
 Effects of Long-term Fertilization on Key Processes of Soil Nitrogen Cycling in Agricultural Soil: A Review WANG Jing, CHENG Yi, CAI Zucong, et al. (303)

Insights and Perspectives

- Soil-borne Pathogens Should not Be Ignored by Soil Science CAI Zucong, HUANG Xinqi (310)

Research Articles

- Element Migration in S₃ Profile of the Shaolingyuan Loess-Paleosol Sequence in Xi'an and Its Paleoclimatic Implication CHU Chunjie, ZHAO Jingbo (320)
 Estimation of Soil Salt Content over Partially Vegetated Areas Based on Blind Source Separation LIU Ya, PAN Xianzhang, SHI Rongjie, et al. (330)
 Construction of Calibration Set based on the Land Use Types in Visible and Near-Infrared (VIS-NIR) Model for Soil Organic Matter Estimation LIU Yanfang, LU Yannian, GUO Long, et al. (340)
 Prediction of Soil Organic Matter based on Multi-resolution Remote Sensing Data and Random Forest Algorithm WANG Yinyin, QI Yanbing, CHEN Yang, et al. (353)
 Characteristic Curves and Model Analysis of Soil Moisture in Collapse Mound Profiles in Southeast Hubei DENG Yusong, DING Shuwen, CAI Chongfa, et al. (363)
 Research on Sediment and Solute Transport on Red Soil Slope under Simultaneous Influence of Scouring Flow MA Meijing, WANG Junguang, GUO Zhonglu, et al. (373)
 Research on Soil Erosion Rate and Hydrodynamic Parameters of Landslide Accumulation Slope in Wenchuan Earthquake Area WANG Renxin, HE Binghui, LI Tianyang, et al. (386)
 Effects of Saline Ice Water Irrigation on Distribution of Moisture and Salt Content in Coastal Saline Soil ZHANG Yue, YANG Jingsong, YAO Rongjiang (399)
 Soil Temperature Regime in Guizhou Province Relative to Assessment Method LU Xiaohui, DONG Yubo, TU Chenglong (409)
 Characteristics of Variation of Soil Temperature in Shrub Meadow Area of Lhasa GONG Yuling, WANG Zhaofeng, ZHANG Yili, et al. (419)
 Soil Water Repellency of Sands and Clay as Affected by Particle Size YANG Song, WU Junhua, DONG Hongyan, et al. (426)
 Effect of AQDS Accelerating Anaerobic Dechlorination of DDT in Hydragric Acrisols LIU Cuiying, WANG Zhuang, XU Xianghua, et al. (436)
 Effect of Straw Returning via Deep Burial Coupled with Application of Fertilizer as Primer on Soil Nutrients and Winter Wheat Yield ZHAO Jinhua, ZHANG Congzhi, ZHANG Jiabao (448)
 Effects of Ozone Pollution on Different Active Organic Carbon Stocks in Wheat Farmland Soil KOU Taiji, CHENG Xianghan, ZHANG Dongliang, et al. (455)
 Soil Organic Nitrogen Components and Their Contributions to Mineralizable Nitrogen in Paddy Soil of the Black Soil Region CONG Yaohui, ZHANG Yuling, ZHANG Yulong, et al. (466)
 Effects of Soil and Water Conservation Measures on Soil Labile Organic Carbon and Soil Enzyme Activity in Gentle Slope Land of Red Soil HUANG Shangshu, CHENG Yanhong, ZHONG Yijun, et al. (475)
 Ecological Stoichiometric Characteristics of Carbon, Nitrogen and Phosphorus in Leaf-Litter-Soil System of *Picea Crassifolia* Forest in the Qilian Mountains ZHAO Weijun, LIU Xiande, JIN Ming, et al. (488)
 The Application of Biomarker Genes for DNA/RNA-Stable Isotope Probing of Active Methanotrophs Responsible for Aerobic Methane Oxidation in Six Paddy Soils ZHENG Yan, JIA Zhongjun (500)
 Screening of Phosphate-solubilizing Bacteria Adaptable to Corn and Effects of the Bacteria on the Growth of Corn MEI Xinlan, SHAN Anqi, JIANG Yi, et al. (508)
 Response of Soil Nematode Community to Cultivation in Upland Red Soil Relative to Cultivation History and Its Significance as Indicator WANG Mingwei, LIU Yudi, CHEN Xiaoyun, et al. (521)
 Nutrient Absorption and Nutrient Balance in an Agro-pastoral Compound Production Pattern of "Raising Geese in Corn Fields" in Tibet SHA Zhipeng, ZHANG Yuyang, WANG Chao, et al. (531)
 Effects of Continuous Cropping of Processing Tomato on Physical-chemical Properties of and Microbial Biomass in the Soil KANG Yalong, JING Feng, SUN Wenqing, et al. (542)

Research Notes

- Mechanism of CTMAB Modifying BS-12 Modified Bentonite YU Lu, MENG Zhaofu, LI Wenbin, et al. (550)
 Effects of Mechanical Soil Amelioration Method on Physical Properties of and Enzyme Activity in Planosol MENG Qingying, ZHANG Chunfeng, JIA Huibin, et al. (559)

Cover Picture: Microbial Extracellular Electron Transfer: Energy Transfer and Substance Transformation (by WU Yundang, LI Fangbai, LIU Tongxu)

《土壤学报》编辑委员会

主 编: 史学正

执行编委: (按姓氏笔画为序)

丁维新	巨晓棠	王敬国	王朝辉	史 舟	宇万太	朱永官
李永涛	李芳柏	李保国	李 航	吴金水	沈其荣	张玉龙
张甘霖	张福锁	陈德明	邵明安	杨劲松	杨明义	杨林章
林先贵	依艳丽	周东美	周健民	金继运	逢焕成	胡 锋
施卫明	骆永明	赵小敏	贾仲君	徐国华	徐明岗	徐建明
崔中利	常志州	黄巧云	章明奎	蒋 新	彭新华	雷 梅
窦 森	廖宗文	蔡祖聪	蔡崇法	潘根兴	魏朝富	

编辑部主任: 陈德明

责任编辑: 卢 萍 檀满枝 陈荣府

土 壤 学 报

Turang Xuebao

(双月刊, 1948年创刊)

第 53 卷 第 2 期 2016 年 3 月

ACTA PEDOLOGICA SINICA

(Bimonthly, Started in 1948)

Vol. 53 No. 2 Mar., 2016

编 辑 《土壤学报》编辑委员会
地址: 南京市北京东路 71 号 邮政编码: 210008
电话: 025 - 86881237
E-mail: actapedo@issas.ac.cn

Edited by Editorial Board of Acta Pedologica Sinica
Add: 71 East Beijing Road, Nanjing 210008, China
Tel: 025 - 86881237
E-mail: actapedo@issas.ac.cn

主 编 史 学 正
主 管 中 国 科 学 院
主 办 中 国 土 壤 学 会
承 办 中国科学院南京土壤研究所

Editor-in-Chief Shi Xuezheng
Superintended by Chinese Academy of Sciences
Sponsored by Soil Science Society of China
Undertaken by Institute of Soil Science,
Chinese Academy of Sciences

出 版 科 学 出 版 社
地址: 北京东黄城根北街 16 号 邮政编码: 100717

Published by Science Press
Add: 16 Donghuangchenggen North Street,
Beijing 100717, China

印刷装订 北京中科印刷有限公司
总发行 科 学 出 版 社
地址: 北京东黄城根北街 16 号 邮政编码: 100717
电话: 010 - 64017032
E-mail: journal@mail.sciencep.com

Printed by Beijing Zhongke Printing Limited Company
Distributed by Science Press
Add: 16 Donghuangchenggen North Street,
Beijing 100717, China
Tel: 010 - 64017032
E-mail: journal@mail.sciencep.com

国外发行 中国国际图书贸易总公司
地址: 北京 399 信箱 邮政编码: 100044

Oversea distributed by China International Book Trading Corporation
Add: P. O. Box 399, Beijing 100044, China

国内统一连续出版物号: CN 32-1119/P

国内邮发代号: 2-560

国外发行代号: BM45

定价: 60.00 元

国 内 外 公 开 发 行



ISSN 0564-3929

