

基于可见-近红外光谱的水稻土全磷反演研究^①

周鼎浩¹, 薛利红^{2*}, 李颖¹, 杨林章²

(1 中国科学院南京土壤研究所, 南京 210008; 2 江苏省农业科学院, 南京 210014)

摘要:采用 PLSR 偏最小二乘法回归结合留一法交叉验证, 利用长期定位试验田以及直湖港小流域面上的水稻土土壤样本建立最优模型, 研究了不同光谱预处理方式对水稻土全磷可见-近红外光谱反演精度的影响, 探索水稻土全磷光谱反演的可行性。并结合简单相关系数法以及 PLSR 模型回归系数法分析了水稻土全磷光谱反演的重要波段。结果表明, 光谱预处理方法对土壤全磷反演精度的影响不大; 基于 PLSR 建立的水稻土全磷光谱反演模型的校正决定系数达 0.85, 交叉验证决定系数为 0.70, RPD 为 1.8, 有较好的模型精度; 440 ~ 740 nm 为土壤全磷光谱反演的重要波段。利用 PLSR 对水稻土全磷进行光谱预测是可行的。

关键词:土壤全磷; 可见-近红外光谱; 偏最小二乘回归; 光谱预处理; 敏感波段; 水稻土
中图分类号: S153.2

获取土壤磷素养分的空间分布信息是控制农田磷素面源污染、精确施用磷肥的关键。而传统的化学分析方法需要破坏性采集样品, 劳动强度大, 费时费力, 成本相对较高。而光谱分析是一种快速、低廉且无损的分析方法。自从 1965 年光谱技术首次被用来分析土壤质地和水分含量以来, 很多学者利用土壤的可见-近红外光谱对土壤特性如水分、有机质、碳、氮、阳离子交换量等进行了调查研究, 结果均表明利用光谱技术尤其是光谱分辨率更高的高光谱技术能够成功地预测这些土壤特性^[1-4]。我国是广种水稻的国家, 稻谷生产在粮食上占有很大的比重。所以, 水稻土的研究在科学上和生产上都有极其重要的意义^[5]。

利用光谱技术进行土壤磷的测试诊断研究是近年来的一个新研究热点。由于土壤总磷主要以无机形态存在, 用近红外分析时存在着一定的困难。因此, 基于总磷的近红外校正模型结果较差, R^2 在 0.4 ~ 0.5 之间^[6-8]。然而, Bogrecki 和 Lee^[9]在美国佛罗里达奥基乔比湖区的细沙土壤中发现了光谱吸收率与磷浓度之间的高度相关性。杜昌文和周健民^[10]利用红外光声光谱对潮土的土壤有效磷进行了预测, 发现建立的偏最小二乘模型和人工神经网络模型均可以用于土壤有效磷的预测。当选用的土壤质地/类型相对

比较单一时, 模型精度较高, 但当利用多种土壤类型进行预测时, 预测精度往往受到影响^[11-13]。徐永明等^[14]研究了可见/近红外反射光谱与土壤氮、磷、钾的关系, 基于反射率、一阶导数、倒数之对数及波段深度的 PLSR 模型对磷的检验相关系数在 0.68 ~ 0.73 ($n = 20$) 之间, 预测均方根差在 0.66 ~ 0.70 mg/kg 之间。李希灿等^[15]利用模糊识别理论建立土壤质量指标光谱反演模型, 其中全磷和速效磷的平均预测误差分别为 8.44% 和 13.01% ($n = 14$)。

由此可见, 虽然利用可见-近红外光谱测量土壤全磷含量被证明是一种可行的方法, 但实验结果和精度仍受土壤类型、实验处理和数据分析方法等因素影响。前人虽对高光谱的数据处理方法、土壤类型的影响或者土壤磷素的光谱反演各有所研究, 但针对我国广泛存在且在科学和生产上具有重要意义的水稻土的全磷光谱反演的研究较少, 尤其是甚少对光谱预处理方法和敏感波段进行系统研究。

因此, 本文在前人的研究基础上有针对性地以单一土壤类型即水稻土为研究对象, 系统地探索了十余种预处理方法对光谱建模精度的影响, 利用苏州农科院长期定位试验田和太湖流域直湖港小流域稻田土壤的混合样本结合交叉验证思想建立 PLSR 偏最小二乘回归模型并验证, 从而探讨利用可见-近红外光

①基金项目: 国家自然科学基金项目(40901104、41171235)和中国科学院知识创新工程重要方向项目(KZCX2-YW-QN406)资助。

* 通讯作者(lhxue@issas.ac.cn)

作者简介: 周鼎浩(1988—), 男, 江苏南京人, 硕士研究生, 研究方向为土壤磷的高光谱反演及其在磷流失高风险区识别中的应用。E-mail: 429270977@qq.com

谱来反演水稻土土壤全磷的可行性 ;并对土壤全磷光谱反演的敏感波段或有用波段进行了讨论分析 ,以期 为今后水稻土全磷的光谱反演研究提供借鉴。

1 材料与方法

1.1 样本采集与制备

土壤样品分别采集于苏州市农业科学院的长期定位试验田(31°32'45"N, 120°41'57"E)以及太湖流域直湖港小流域。苏州市农业科学院的长期定位试验田供试土壤类型为黄泥土, 试验设 14 个处理 :无肥对照(CK)、单施氮肥(N)、氮磷配施(NP)、氮钾配施(NK)、磷钾配施(PK)、氮磷钾配施(NPK)、氮肥加秸秆(RN)、单施粪肥(M)、粪肥加氮肥(MN)、粪肥加氮磷化肥(MNP)、粪肥加氮钾化肥(MNK)、粪肥加磷钾化肥(MPK)、粪肥加氮磷钾(MNPK)、粪肥加氮肥加秸秆(MRN)。试验从 1980 年开始, 稻麦轮作, 经过 24 年的施肥, 土壤氮、磷、钾的含量差异明显。分别采集了 2008、2009 和 2010 三年水稻成熟期土壤表层样品(0~20 cm), 共 74 个样本。此外, 2011 年水稻收获后, 在太湖流域直湖港小流域稻田随机采集稻田土壤样本 34 个, 与农科院 74 个样本混合用于最优模型的建立及检验。每个样本分为两份, 一份用作全磷含量的化学分析, 采用 NaOH 碱熔钼锑抗比色法(国家标准法)^[16], 一份用作室内光谱分析。所有的样本均经过风干, 拣出根系以及其他杂质, 研磨过 20 目筛处理。

1.2 光谱测定

采用 ASD 公司产的 FieldSpec 3 Hi-Res (350 ~ 2 500 nm) 对水稻土风干土样进行了室内高光谱测定, 采样间隔分别为 1.4 nm(350~1 050 nm)和 2 nm (1 000~2 500 nm)。光谱分辨率为 3 nm(350~700 nm)、8.5 nm(700~1 400 nm)和 6.5 nm(1 400~2 100 nm), 测定范围为 350~2 500 nm, 探头视场角为 25°, 室内照明光源采用与光谱仪配套的 2 个 14.5 v 卤素灯。

将制备好的土壤样品均匀平铺于直径约 6.7 cm, 深度约 0.8 cm 的铝盒中。探头距土壤样品表面约 13 cm。探头入射角 90°, 室内光源入射角为 45°。每个样本扫描 5 次, 取其平均值作为最终光谱反射率数据。

1.3 光谱数据分析

用光谱仪自身携带的软件 Viewspec 中的 splice correction 功能对光谱曲线进行断点拼接, 并结合前人研究经验去除 350~380 nm 和 2 400~2 500 nm 噪声波段^[17-18]。在 10 nm 重采样间隔下, 采用不同预处理方法, 利用苏州农科院长期定位试验田的水稻土样品作为数据集, 结合留一法交叉验证(LOOCV)建立土壤全磷偏最小二乘回归(PLSR)模型, 以模型的校正决定系数 R^2_{cal} 、交叉验证决定系数 R^2_{val} 、校正均方根误差 $RMSEC$ 、交叉验证均方根误差 $RMSEP$ 、 RPD ^[19]为评价指标, 筛选土壤全磷预测的最佳光谱处理方法。并将苏州农科院与直湖港小流域面上的样本混合建立最优模型, 检验是否可以建立一个最优模型来进行水稻土全磷光谱反演, 验证水稻土全磷光谱反演的可行性及其精度。最后通过简单相关分析和 PLSR 模型各波段的回归系数分析, 讨论了土壤全磷的光谱敏感响应波段。

所有数据分析均在 Unscrambler9.7、SPSS13.0 和 EXCEL2007 软件中进行。

2 结果与分析

2.1 样本全磷含量状况

数据集主要来源于两部分。苏州农科院长期定位试验田, 共 74 个样本, 土壤全磷含量变化在 0.38 ~ 1.37 g/kg 之间, 有较好的正态分布特征, 偏度 0.697, 峰度-0.623。直湖港小流域, 共 34 个样本, 全磷含量变化在 0.03 ~ 0.68 g/kg, 亦有较好的正态分布特征, 偏度 -0.175, 峰度 -0.451。综合数据集为二者之和(表 1)。

表 1 土壤样品基本信息
Table 1 Status of soil sample properties

数据集	样本数	最大值 (g/kg)	最小值 (g/kg)	平均值 (g/kg)	标准差 (g/kg)
农科院数据集	74	1.37	0.38	0.73	0.31
直湖港数据集	34	0.68	0.03	0.40	0.20
综合数据集	108	1.37	0.03	0.63	0.32

2.2 基于全波段的 PLSR 建模分析

2.2.1 预处理方法对建模精度的影响 光谱数据的预处理可以消除或者减少各种因素对光谱产生的干扰, 减小数据的共线性、维数、背景影响、噪

声水平和光谱的基线漂移偏移等对光谱建模的影响^[20]。偏最小二乘法 PLSR 有机结合了多元线性回归分析、主成分分析和单相关分析, 在自变量多、变量间存在严重多重相关性方面具有独特的优

势^[21-23]。本研究以苏州农科院 74 样本为基础，剔除 350 ~ 380 nm 和 2 400 ~ 2 500 nm 噪声波段后，在 10 nm 重采样间隔下对不同的预处理方法进行了探讨，以 R^2_{cal} 、 $RMSEC$ 、 R^2_{val} 、 $RMSEP$ 和 RPD 作为 PLSR 模型精度评价指标，结果见表 2。结果表明，

本次试验中各种光谱预处理方法对土壤全磷光谱 PLSR 回归模型结果影响不大， R^2_{cal} 变化在 0.90 ~ 0.97 之间。其中窗口移动平均法平滑对模型精度的提高有一定帮助， R^2_{val} 达到 0.80， $RMSEP$ 为 0.14 g/kg， RPD 为 2.2。

表 2 预处理方法对建模精度的影响
Table 2 Influence of different pretreatments

数据处理方式	主成分数	R^2_{cal}	$RMSEC(g/kg)$	R^2_{val}	$RMSEP(g/kg)$	RPD
原始光谱	12	0.96	0.06	0.74	0.16	1.9
S-G 平滑(2 阶, 3)	11	0.95	0.07	0.75	0.16	1.9
窗口移动平均法平滑	10	0.92	0.09	0.80	0.14	2.2
高斯滤波平滑	10	0.94	0.07	0.75	0.16	1.9
中值滤波	10	0.92	0.08	0.79	0.14	2.2
一阶微分	12	0.97	0.05	0.65	0.19	1.6
基线校正	11	0.96	0.06	0.74	0.16	1.9
多元散射校正	11	0.96	0.06	0.70	0.17	1.8
标准正态变量变换 SNV	11	0.96	0.06	0.71	0.17	1.8
平均归一化	11	0.96	0.06	0.73	0.16	1.9
最大值归一化	12	0.96	0.06	0.71	0.17	1.8
范围归一化	12	0.96	0.06	0.69	0.17	1.8
面积归一化	11	0.96	0.06	0.72	0.16	1.9
窗口移动平均法平滑 + 一阶微分	8	0.93	0.08	0.79	0.14	2.2
窗口移动平均法平滑 + SNV	9	0.90	0.10	0.76	0.15	2.0
窗口移动平均法平滑 + 面积归一化	9	0.90	0.10	0.78	0.15	2.1
窗口移动平均法平滑 + 范围归一化	12	0.96	0.06	0.69	0.17	1.8
高斯滤波平滑 + 面积归一化	9	0.92	0.09	0.73	0.16	1.9

2.2.2 最优光谱反演模型的建立 采用窗口移动平均平滑法对光谱进行预处理，利用 PLSR 偏最小二乘回归法对苏州农科院水稻土 74 个样本及直湖港 34 样本共 108 样本综合数据集进行建模，得到土壤全磷的最优光谱反演模型；采用留一法交叉验证(LOOCV)对建立的模型进行验证。可以看出，模型有很好的拟合效果(图 1)，实测值和预测值较好地分布在 1 : 1 线左右，模型校正决定系数 R^2_{cal} 达 0.85，交叉验证的决定系数 R^2_{val} 为 0.70， $RMSEC$ 为 0.12 g/kg， $RMSEP$ 为 0.18 g/kg， RPD 为 1.8。PLSR 模型提取出 10 个主成分，方差累计贡献率占校正模型的 85%，占验证模型的 70%。结果表明，该模型能够用于水稻土的光谱磷预测。

2.3 土壤光谱对全磷含量的响应关系分析

高光谱波段多，信息量大，如何有效地从海量数据中精简信息，提取有用波段是可见-近红外光谱的研究热点。本文分别采取简单线性相关分析和基于 PLSR 回归系数分析两种方法尝试探讨光谱全磷响应重要波段。

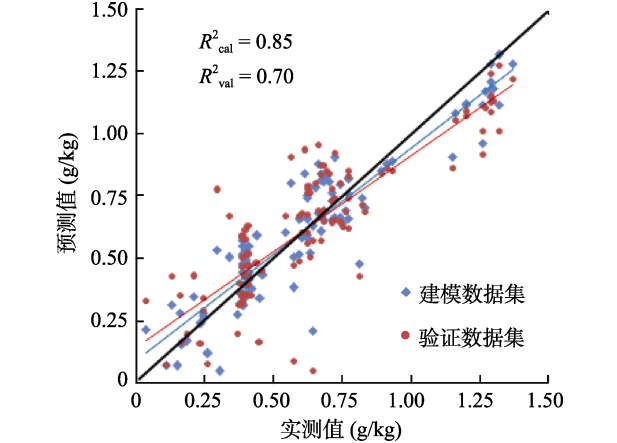


图 1 PLSR 模型拟合效果
Fig. 1 Fit effect of PLSR model

2.3.1 基于简单线性相关分析的土壤全磷光谱响应

重要波段筛选 以苏州农科院 74 样本为研究对象，利用 SPSS 软件对 380 ~ 2 400 nm 原始光谱各波段与全磷含量作简单相关性分析，得到如图 2 的简单相关系数图。可以看出，所有波段光谱反射率均与土壤全磷含量呈正相关，但相关性总体不高。其

中, 641 ~ 1 440 nm 通过了 $P < 0.05$ 水平下的显著性检验。利用这些通过显著检验的波段进行 PLSR 建模, 模型 R^2_{cal} 为 0.79, $RMSEC$ 为 0.14 g/kg, R^2_{val} 为 0.66, $RMSEP$ 为 0.18 g/kg, 与全波段 PLSR 模型相比, 精度并没有提高。表明, 简单相关性分析并不能有效地筛选出土壤全磷响应波段。

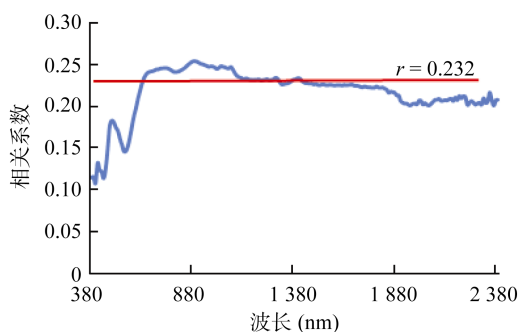


图 2 380 ~ 2 400 nm 各波段全磷含量相关系数
Fig. 2 Correlation of reflectance and TP in 380~2 400 nm

2.3.2 基于 PLSR 模型回归系数分析的土壤全磷光谱响应重要波段筛选 多元线性回归模型中, 因变量回归系数的大小在一定程度上反映了该因变量对自变量的影响程度, 回归系数越大, 其对自变量的影响程度愈大。以苏州农科院 74 样本为研究对象, 利用 Unscrambler9.7 中 PLSR 下的 regression coefficients 以及 important variables 功能分别对 10 nm 重采样间隔下的原始土壤光谱进行了波段筛选。研究发现, 对全磷含量影响较大的光谱波段主要为 381 ~ 410、440 ~ 450、470 ~ 610、630 ~ 740、780 ~ 860、1 150 ~ 1 190 nm。利用筛选出来的重要波段重新进行 PLSR 建模, 结果如表 3。由表 3 可见, 基于 381 ~ 410、440 ~ 450、470 ~ 610、630 ~ 740、780 ~ 860、1 150 ~ 1 190 nm 波段的 PLSR 模型验证决定系数达到 0.79, 验证均方根误差为 0.14 g/kg。而基于 440 ~ 610 和 630 ~ 740 nm 波段的 PLSR 模型验证决定系数达到 0.83, 验证均方根误差为 0.13 g/kg, 精度比全波段的 PLSR 模型略有提高。表明, 对于本试验供试土壤, 440 ~ 740 nm 波段对土壤磷的预测非常重要。

表 3 10 nm 采样间隔下基于重要波段的土壤全磷 PLSR 模型与全波段的土壤全磷 PLSR 模型的比较

Table 3 Comparison between PLSR models with important bands and full range in 10 nm sampling interval

波段	R^2_{cal}	$RMSEC$ (g/kg)	R^2_{val}	$RMSEP$ (g/kg)
波段 a	0.85	0.12	0.81	0.13
波段 b	0.59	0.20	0.42	0.24
波段 c	0.26	0.26	0.13	0.29
波段 a+b+c	0.83	0.12	0.79	0.14
波段 a+b	0.86	0.11	0.83	0.13
380 ~ 2 400 nm	0.96	0.06	0.74	0.16

注: 波段 a 代表 440 ~ 610 nm, b 代表 630 ~ 740 nm, c 代表 381 ~ 410、780 ~ 860、1 150 ~ 1 190 nm。

3 讨论与小结

本文利用苏州农科院长期定位田的水稻土 74 个样本, 初步探讨了光谱预处理方法对土壤总磷光谱预测 PLSR 模型精度的影响。在此基础上探讨了是否可以针对相同土壤类型但数据来源不同的样本建立统一的模型进行光谱全磷反演以及土壤全磷的光谱响应重要波段。

研究表明, 不同光谱预处理方法对 PLSR 建模精度影响不大, 这可能是由于室内光谱测量时较好地控制了实验环境, 减小了噪声、基线漂移等造成的误差, 与野外测量难以控制测量环境有所不同。此外去除掉 350 ~ 380、2 400 ~ 2 500 nm 噪声波段以及波段融合后, 信噪比已经得到较大提高, 因此平滑、求导等预处理方法帮助不大。利用偏最小二乘法结合交叉验证法建立的最优模型校正相关系数达 0.85, 预测相关系数 0.70, $RMSEC$ 为 0.12 g/kg, $RMSEP$ 为 0.18 g/kg, RPD 为 1.8。1.75 < RPD < 2.25 时, 表明该模型能较有效地进行土壤全磷的预测^[19]。对单一水稻土而言, 可以将不同数据来源的样本混合建立统一的模型, 利用土壤反射光谱来反演水稻土土壤全磷含量是可行的。

此外, 研究还发现简单相关性分析并不能有效地提取出土壤全磷的敏感光谱波段, 而基于 PLSR 模型回归系数分析则能提取出对土壤全磷预测有用的重要波段, 为 440 ~ 740 nm 波段。这主要是因为磷在土壤中并没有明确的吸收峰, 对其进行光谱反演必须依赖综合的光谱信息, 传统的简单相关分析并没有考虑到高光谱数据各波段之间的相互关系, 而 PLSR 法则把所有光谱波段作为一个综合体进行考虑, 充分考虑了波段之间的相互作用关系。基于 440 ~ 740 nm 可见光波段建立的 PLSR 模型, 比全波段的 PLSR 模型精度略有提高, 表明该波段范围是土壤全磷预测十分有用的光谱波段。这与前人得出的近红外波段最优、可见光次之、紫外波段较差的研究结果不太吻合。这可能是由于土壤类型不同导致^[24]。本文研究结果表明, 利用 PLSR 模型回归系数法可以用来寻求预期属性的光谱响应重要波段, 尤其是那些没有显著吸收峰的属性, 这在以往的研究中较少提及。

然而, 当利用提取的这些重要波段对直湖港流域水稻土全磷进行预测时精度一般。可能原因是因为直湖港土壤样本与农科院不完全一致, 土壤全磷的分布区间明显小于农科院。这也表明, 提取的 440 ~ 740 nm 波段虽然对土壤全磷预测比较重要, 但删去的其余波

段仍然包含了一定的光谱信息,对建模精度提高没有帮助,但对提高模型的稳定性等仍然具有一定的作用。因此,在今后的研究工作中,可以借鉴植被光谱中光谱指数的构建思路,对能提高模型精度的敏感波段和不能提高模型精度但有可能包含一定光谱信息的不敏感波段进行组合,寻找一种可较广泛适用的波段组合方法和新的参数来进行土壤全磷预测。

参考文献：

- [1] Bowers SA, Hanks RJ. Reflection of radiant energy from soil[J]. *Soil Science*, 1965, 100(3): 130-138
- [2] Stoner ER, Baumgardner MF. Characteristic variations in reflectance of surface soils[J]. *Soil Science Society of America Journal*, 1981, 45(6): 1 161-1 165
- [3] Sudduth KA, Hummel JW. Evaluation of reflectance methods for soil organic matter sensing[J]. *Transactions of the ASAE*, 1991, 34(4): 1 900-1 909
- [4] Dalal RC, Henry RJ. Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry[J]. *Soil Science Society of America Journal*, 1986, 50(1): 120-123
- [5] 徐琪, 陆彦椿, 刘元昌, 朱洪官. 中国太湖地区水稻土[M]. 上海: 上海科学技术出版社, 1980
- [6] Chang CW, Laird DA, Mausbach MJ, Hurburgh CR. Near-infrared reflectance spectroscopy—principal components regression analyses of soil properties[J]. *Soil Science Society of America Journal*, 2001, 65(2): 480-490
- [7] Thomasson JA, Sui R, Cox MS, Al-Rajehy A. Soil reflectance sensing for determining soil properties in precision agriculture[J]. *Transactions of the ASAE*, 2001, 44(6): 1 445-1 453
- [8] 宋海燕, 何勇. 近红外光谱法分析土壤中磷、钾含量及 pH 值的研究[J]. *山西农业大学学报*, 2008, 28(3): 275-278
- [9] Bogrekci I, Lee WS. Spectral soil signatures and sensing phosphorus[J]. *Biosystems Engineering*, 2005, 92(4): 527-533
- [10] 杜昌文, 周健民. 傅里叶变换红外光声光谱法测定土壤中有效磷[J]. *分析化学*, 2007, 35(1): 119-122
- [11] 陈鹏飞, 刘良云, 王纪华, 沈涛, 陆安祥, 赵春江. 近红外光谱技术实时测定土壤中总氮及磷含量的初步研究[J]. *光谱与光谱学分析*, 2008, 28(2): 295-298
- [12] Hartmann HP, Appel T. Calibration of near infrared spectra for measuring decomposing cellulose and green manure in soils[J]. *Soil Biology and Biochemistry*, 2006, 38(5): 887-897
- [13] 田永超, 张娟娟, 朱艳, 姚霞, 曹卫星. 不同类型土壤的光谱特征及其有机质含量预测[J]. *中国农业科学*, 2009, 42(9): 3 154-3 163
- [14] 徐永明, 蔺启忠, 王璐, 黄秀华. 基于高分辨率反射光谱的土壤营养元素估算模型[J]. *土壤学报*, 2006, 43(5): 709-716
- [15] 李希灿, 王静, 李玉环, 王芳, 赵杰. 基于模糊集分析的土壤质量指标高光谱反演[J]. *地理与地理信息科学*, 2008, 24(4): 25-28
- [16] 鲍士旦. 土壤农化分析[M]. 北京: 中国农业出版社, 1999: 76-78
- [17] 王淼, 解宪丽, 周睿, 王宝良, 王昌昆, 刘娅, 潘剑君, 沈润平, 潘贤章. 基于可见光-近红外漫反射光谱的红壤有机质预测及其最优波段选择[J]. *土壤学报*, 2011, 48(5): 1 083-1 089
- [18] 王淼, 潘贤章, 解宪丽, 王昌昆, 刘娅, 李燕丽, 潘剑君. 土壤含水量对反射光谱法预测红壤土壤有机质的影响研究[J]. *土壤*, 2012, 44(4): 645-651
- [19] Maleki MR, van Holm L, Ramon H, Merckx R, de Baerdemaeker J, Mouazen AM. Phosphorus sensing for fresh soils using visible and near infrared spectroscopy[J]. *Biosystems Engineering*, 2006, 95(3): 425-436
- [20] 陆婉珍. 现代近红外光谱分析技术[M]. 2 版. 北京: 中国石化出版社, 2010
- [21] 王纪华, 黄文江, 劳彩莲, 张录达, 罗长兵, 王韬, 刘良云, 宋晓宇, 马智宏. 运用 PLSR 算法由小麦冠层反射光谱反演氮素垂直分布[J]. *光谱学与光谱分析*, 2007, 27(7): 1 319-1 322
- [22] 翁永玲, 戚浩平, 方洪宾, 赵福岳, 路云阁. 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): 1 255-1 263
- [23] Mouazen AM, de Baerdemaeker J, Ramon H. Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer[J]. *Soil and Tillage Research*, 2005, 80(1/2): 171-183
- [24] Bogrekci I, Lee WS. Comparison of ultraviolet, visible, and near infrared sensing for soil phosphorus[J]. *Biosystems Engineering*, 2007, 96(2): 293-299

Visible-Near Infrared Reflectance Spectroscopy for Prediction of Total Phosphorus Content in Paddy Soil

ZHOU Ding-hao¹, XUE Li-hong^{2*}, LI Ying¹, YANG Lin-zhang²

(1 *Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China;*

2 *Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China*)

Abstract: PLSR regression with cross validation(LOOCV) was used to establish the optimal model using the dataset of a long-term plot experiment in Suzhou and dataset of Zhihugang river watershed aimed to discuss the influence of different pretreatment methods on prediction of total phosphorus(TP) in paddy soil and explor the feasibility of total phosphorus prediction with hyperspectrum. Important spectrum bands for total phosphorus prediction were also explored with correlation analysis and PLSR regression coefficient analysis. Results showed that effect of pretreatment methods of hyperspectral data on the precision of PLSR model was little. The precision of optimum PLSR model was good with the coefficient of determination (R^2) of 0.85 for calibration and 0.70 for validation, and RPD of 1.8. 440–740 nm was important for TP prediction with hyperspectrum. It is feasible to predict soil TP using visible-near infrared reflectance spectroscopy with PLSR regression method.

Key words: Soil total phosphorus, Visible-near infrared reflectance spectroscopy, Partial least square regression, Hyperspectrum pretreatment, Sensitive band, Paddy soil