

基于高光谱的土壤全氮含量估测

张娟娟^{1,2}, 田永超¹, 姚霞¹, 曹卫星¹, 马新明², 朱艳^{1*}

(1. 南京农业大学 国家信息农业工程技术中心, 江苏省信息农业高新技术研究重点实验室, 南京 210095;
2. 河南农业大学 信息与管理科学学院, 郑州 450002)

摘要: 基于高光谱(350~2 500 nm)数据,研究了我国中、东部地区5种主要类型土壤全氮含量与高光谱反射率之间的定量关系,构建了基于偏最小二乘法(PLS)、BP神经网络(BPNN)和特征光谱指数的土壤全氮含量估算模型。结果表明,以500~900 nm、1 350~1 490 nm区域波段反射率经Norris滤波平滑后的一阶导数光谱为基础,构建的基于PLS和BPNN的土壤全氮含量估算模型精度较高,建模决定系数分别为0.81和0.98;独立观测资料检验结果显示,模型预测决定系数分别为0.81和0.93,均方根误差RMSE为0.219 g·kg⁻¹和0.149 g·kg⁻¹,相对分析误差RPD为2.28和3.36,说明PLS和BPNN模型对土壤全氮含量具有较高的预测精度。在光谱指数的分析中,基于近红外872 nm和1 482 nm两个波段的差值光谱指数DI(NDR₈₇₂, NDR₁₄₈₂)对土壤全氮含量最敏感,建模决定系数、预测决定系数、RMSE和RPD分别为0.66、0.53、0.31 g·kg⁻¹和1.60。比较而言,三种方法估算土壤氮含量的精度顺序为BPNN模型>PLS>DI(NDR₈₇₂, NDR₁₄₈₂),基于PLS和BPNN两种方法建立的土壤全氮含量高光谱估测模型具有较高的精度,可以用来精确估算土壤全氮含量;基于两波段构建的DI(NDR₈₇₂, NDR₁₄₈₂)预测效果低于前两者,但也可以用来粗略估测土壤中的全氮含量。

关键词: 土壤; 全氮; 高光谱; 偏最小二乘法; BP神经网络; 光谱指数

中图分类号: S153.6 **文献标志码:** A **文章编号:** 1000-3037(2011)05-0881-10

土壤氮是植物生长发育的基础,其丰缺状况直接影响作物生长及产量和品质的形成。土壤氮的传统化学测定方法如Kjeldahl或Bremner法通常耗时、费工、有害,且实效性弱。近年来,光谱分析技术被证明是一项有潜力的化学组分快速测定技术^[1-3]。Dalal等^[4]研究发现土壤光谱在1 700~2 100 nm与氮素关系密切。徐永明等^[5]利用室内土壤高光谱反射率,运用一阶导数、倒数、反射率倒数的对数以及波段深度与总氮含量进行逐步回归分析,确定了土壤氮敏感波谱带为2 160~2 280 nm和2 448~2 476 nm。卢艳丽等^[6]利用光谱指数建立了东北黑土的全氮含量估测模型,结果表明基于550 nm和450 nm构建的归一化光谱指数能较好地估测土壤全氮含量。Lee等^[7]认为土壤氮的吸收波段在510 nm左右,但不同的土壤类型有较大差异。Reeves等^[8]发现基于NIRS(近红外光谱)的土壤总氮预测精度受土壤类型的影响较大。Chang等^[9]利用偏最小二乘法(PLS)建立了基于光谱分析技术的土壤全氮含量估测模型,其预测值与化学测定值之间的相关系数达到0.90以上。孙建英等^[10]

收稿日期: 2010-09-14; 修订日期: 2010-12-24。

基金项目: 教育部新世纪优秀人才支持计划(NCET-08-0797); 国家自然科学基金项目(30871448); 江苏省创新学者攀登计划(BK2008037); 江苏省自然科学基金项目(BK2008330)。

第一作者简介: 张娟娟(1979-),女,讲师,博士,主要从事农田土壤监测方面的研究。

* 通信作者, E-mail: yanzhu@njau.edu.cn

利用近红外光谱结合偏最小二乘法建立了土壤全氮含量的光谱模型,且筛选出估测全氮的最佳波长为 2 262 nm 和 1 888 nm。赵锁劳等^[11]利用逐步多元回归建立了我国黄土样品的总氮量估测模型,定标相关系数为 0.94。于飞健等^[12]利用偏最小二乘法预测了土壤总氮含量,对于 2 mm 和 0.15 mm 风干土样实测值和预测值间的决定系数分别为 0.88 和 0.90。张雪莲等^[13]利用偏最小二乘法分别建立了 5 种不同类型土壤不同深度的全氮含量光谱估测模型,认为模型受土壤类型的影响较大,某地区建立的模型用于另一地区土壤样品预测时存在较大的局限性。郑立华等^[14]利用神经网络方法建立了土壤全氮含量估测模型,模型预测精度达到 0.81。以上研究表明,利用光谱分析技术预测土壤全氮含量是可行的,然而,不同类型土壤光谱差异较大,针对多种类型土壤的全氮含量进行系统预测的研究还较少;另外,在光谱分析中,利用不同方法建立模型预测精度存在较大差别,尚需系统比较和评价。

因此,本文拟采用高光谱分析技术,在采集我国中、东部地区 5 种主要类型土壤样本的光谱反射率和全氮含量基础上,系统研究土壤全氮的特征敏感光谱区域,并运用偏最小二乘法、BP 神经网络和特征光谱指数 3 种方法构建并比较土壤全氮含量的光谱估测模型,以期对土壤全氮含量信息的快速获取提供技术依据。

1 材料与方法

1.1 土样采集和制备

试验土壤样品为我国中、东部地区 5 种主要土壤类型:水稻土、潮土、砂浆黑土、盐潮土和盐土。所有土壤样品均为基于梅花法 5 点采集的混合样本,经自然风干后,用木棒压磨后粗略去除沙砾及植物残体,然后过 20 目筛子混合均匀,供光谱测试和化学分析用。土壤样本基本信息参见文献[15]。

1.2 土壤全氮含量测定

采用凯氏定氮法分析土壤全氮含量。表 1 为不同采样地点获取的土壤样品全氮含量统计特征。不同地区农田土壤全氮含量差异较大,样本中最大值出现在杭州($3.04 \text{ g} \cdot \text{kg}^{-1}$),其次是宿州,最小值出现在郑州($0.35 \text{ g} \cdot \text{kg}^{-1}$)。样本全氮含量的标准偏差为 0.51,表明样本覆盖土壤类型广、肥力变异度较大,用于建模研究具有较好的代表性。

表 1 土壤全氮含量统计特征

Table 1 Statistic characteristics of soil total nitrogen content

取样地点	样本个数	最大值/ $(\text{g} \cdot \text{kg}^{-1})$	最小值/ $(\text{g} \cdot \text{kg}^{-1})$	平均值/ $(\text{g} \cdot \text{kg}^{-1})$	标准差
郑州	121	1.14	0.35	0.71	0.20
滑县	45	1.14	0.58	0.85	0.15
宿州	76	2.02	0.75	1.25	0.23
南京	128	1.69	0.62	1.11	0.29
杭州	48	3.04	0.91	2.02	0.78
大丰	24	1.20	0.68	0.91	0.18
上海	18	2.19	0.96	1.53	0.41
所有样本	460	3.04	0.35	1.10	0.51
建模样本	331	3.04	0.35	1.11	0.52
检验样本	129	2.99	0.42	1.07	0.50

1.3 光谱测量

运用 ASD FieldSpec Pro 地物高光谱仪^[15] 测量土壤样品的反射光谱数据。仪器光源为功率 1 000 W 的卤素灯,探头视场角为 5°。测量在黑暗环境的实验室进行,光源照射方向与垂直方向夹角 30°,探头距离土样表面 15 cm,把适量经过处理的土壤样品倒入盛样皿中,用玻璃稍稍压实,保持表面平整,厚 1.5 mm。每个样品旋转 3 次,每个角度扫描 12 次,共 36 次,取平均值作为该样本光谱。

1.4 土壤光谱的预处理

建立模型前,采用 Savitzky-Golay 平滑(SG 平滑)^[16] 及 Norris 平滑滤波^[17] 两种方式对一阶导数光谱进行预处理,以降低光谱背景噪音的影响。

1.5 数据分析方法

1.5.1 光谱指数法

将 460 个混合样本随机分成两组,一组用于土壤全氮含量估测模型的构建(建模样本,共 331 个),另一组用于模型的验证(测试样本,共 129 个)。利用 350~2 500 nm 波段的光谱反射率,系统构造了原始光谱和一阶导数光谱任意两波段组合而成的比值光谱指数(RI)、差值光谱指数(DI)和归一化光谱指数(ND):

$$RI(R_{\lambda_1}, R_{\lambda_2}) = R_{\lambda_1} / R_{\lambda_2} \quad (1)$$

$$DI(R_{\lambda_1}, R_{\lambda_2}) = R_{\lambda_1} - R_{\lambda_2} \quad (2)$$

$$ND(R_{\lambda_1}, R_{\lambda_2}) = (R_{\lambda_1} - R_{\lambda_2}) / (R_{\lambda_1} + R_{\lambda_2}) \quad (3)$$

其中, λ_2 和 λ_1 为 350~2 500 nm 波段范围内任意两波段, R_{λ_1} 和 R_{λ_2} 为 350~2 500 nm 波段范围内任意两波段的原始反射率, SDR_{λ_1} 和 SDR_{λ_2} 为经 SG 平滑后的任意两波段反射率的一阶导数值, NDR_{λ_1} 和 NDR_{λ_2} 为经 Norris 滤波平滑后的任意两波段反射率的一阶导数值。通过分析上述光谱参数与土壤全氮含量间的定量关系,以筛选估测土壤全氮含量的最佳光谱参数。光谱参数的计算利用 Matlab 7.0^[21] 自行编写程序实现。

1.5.2 偏最小二乘法

进一步构建基于偏最小二乘法的土壤全氮含量估测模型。PLS 定标采用 Thermo 公司提供的 TQ Analyst 7.1 分析软件实现,利用“剔一法(Leave-one-out)”计算内部交互验证均方根误差(RMSECV)及决定系数(R_{cv}^2)来优化建模参数。

1.5.3 神经网络法

将 PLS 建模分析所提取的光谱特征变量,作为 BP 神经网络法(BPNN)的输入值,以减少 BPNN 模型的计算量。隐含层和输出层传递函数采用 Tansig 函数和 Purelin 函数,网络训练函数采用 Trainlm 函数。隐含层结点数选择通过多次反复试验确定,目标误差、学习速率及训练迭代次数根据目标要求进行设定。BP 神经网络的学习利用 Matlab 7.0^[18] 编程实现。

土壤全氮含量估测模型的预测精度采用预测值和观测值间的决定系数 R^2 、均方根误差 RMSE 和相对分析误差 RPD 来衡量。 R^2 越大, RMSE 越小,说明模型的精度越高。另外,当 $RPD > 2$ 时表明模型具有极好的预测能力,当 $1.4 < RPD < 2$ 表明模型可对样品作粗略估测,而 $RPD < 1.4$ 则表明模型无法对样品进行准确预测^[9]。

2 结果与分析

2.1 土壤全氮含量与光谱反射率及其一阶导数的相关性

图 1(a)为土壤全氮含量与原始光谱反射率和一阶导数光谱的相关系数图。可以看出,

土壤全氮含量与原始反射率在 350~560 nm 和 1 913~1 932 nm 区域正相关,其余波段区域呈负相关,整体相关系数均较低(绝对值低于 0.40)。一阶导数光谱在 556 nm 附近与土壤全氮含量呈较好的负相关,相关系数达到 0.67,而在 850 nm、1 400 nm 附近呈较好的正相关,相关系数达到 0.5~0.6,可能是一阶导数变换能较好地消除背景噪音的原因。

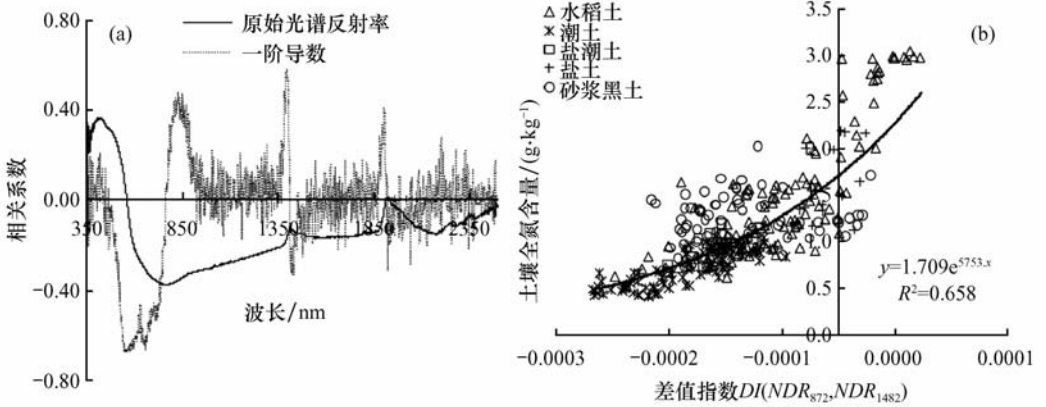


图 1 土壤全氮含量与原始光谱反射率及一阶导数的相关性(a)和土壤全氮含量与差值指数 $DI(NDR_{872}, NDR_{1482})$ 的关系(b)

Fig. 1 Correlations of soil total nitrogen content to original spectral reflectance and its first derivative value (a) and relationships between soil total nitrogen content and $DI(NDR_{872}, NDR_{1482})$ (b)

2.2 基于光谱指数的土壤全氮含量估测

研究了常见的两波段原始光谱及一阶导数光谱组合而成的比值、差值和归一化指数与土壤全氮含量的关系。结果表明,原始光谱反射率与全氮含量相关性较好的波段组合主要集中在两个区域,一个区域位于可见光和近红外波段,且基本位于 1 000 nm 以内;另外一个区域位于 1 800~2 200 nm 范围内,且比值、归一化和差值指数所涉及的波段区域基本一致。而类似的一阶导数波段组合主要集中在 530~680 nm 与 1 360~1 400 nm、800~930 nm 与 1 420~1 500 nm 区域。图 2 展示了两波段原始反射率和经 Norris 滤波平滑后的一阶导数光谱组合而成的差值指数与全氮含量的相关等势图。比值和归一化指数与全氮含量的相关性

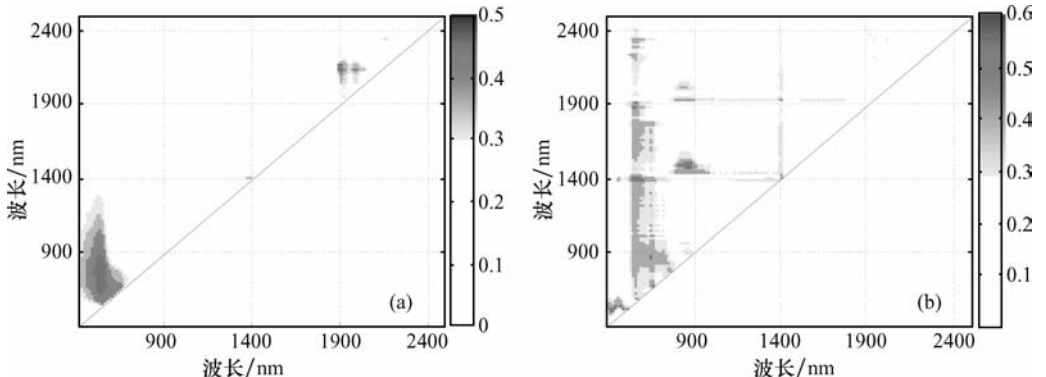


图 2 土壤全氮含量与两波段原始光谱差值指数(a)及两波段导数差值指数(b)的相关性

Fig. 2 Correlations of soil total nitrogen content to two-band difference index based on spectral reflectance (a) and corrected first derivative spectra with Norris smoothing filter (b)

弱于差值指数,一阶导数光谱的组合与土壤全氮含量的相关性整体高于原始光谱指数。其中,以近红外区 872 nm 和 1 482 nm 两个波段的一阶导数构成的差值指数即 $DI(NDR_{872}, NDR_{1482})$ 表现最好,建模样本决定系数为 0.658[图 1(b)]。

2.3 基于 PLS 的土壤全氮含量估测

分别以全谱波段(350~2 500 nm)、前文研究提出的与土壤全氮含量相关性较高的波段区域(称为敏感波段)为光谱建模样本,建立了基于 PLS 的土壤全氮含量估测模型。表 2 显示了基于不同光谱预处理方法和建模区域相结合的 PLS 模型参数,选取敏感波段作为建模区域所建立的全氮含量估测模型优于基于全谱波段的模型。其中以 500~900 nm、1 350~1 490 nm 经 Norris 平滑的一阶导数光谱为建模区域,当主因子数为 10 时,内部交互均方根偏差最小[图 3(a)],所建模型精度最高,建模决定系数 Rcv^2 、内部交互验证均方根误差 $RMSECV$ 和相对分析误差 RPD 分别为 0.81、0.23 $g \cdot kg^{-1}$ 和 2.26,图 3(b)为该模型内部交互验证结果。

表 2 基于不同光谱预处理方法的 PLS 模型内部交互验证结果

Table 2 Cross-validation results of the PLS models based on different pre-processing methods

预处理方法	建模区域/nm	主因子个数	决定系数 R^2	内部交互验证均方根误差 $RMSECV$	相对分析误差 RPD
R	350~2 500	13	0.78	0.24	2.14
SG + D	350~2 500	5	0.49	0.37	1.39
Norris + D	350~2 500	14	0.78	0.25	2.12
R	400~1 000	12	0.79	0.24	2.14
SG + D	500~900, 1 350~1 490	8	0.66	0.31	1.69
Norris + D	500~900, 1 350~1 400	10	0.81	0.23	2.26

注:R 表示反射率;SG 表示 Savitky-Golay 平滑;Norris 表示 Norris 平滑;D 表示一阶导数。

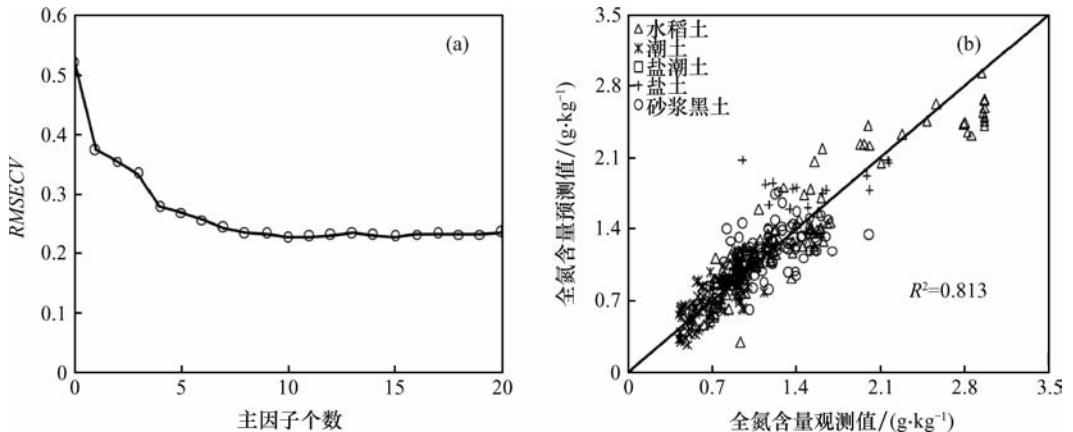


图 3 主因子数对 PLS 定量分析模型 $RMSECV$ 的影响(a)和基于 PLS 的土壤全氮含量监测模型内部交互验证结果(b)

Fig. 3 The cross validated root mean square error ($RMSECV$) of prediction against the number of factor (a) and cross-validation between measured and estimated values of total nitrogen content based on PLS model (b)

2.4 基于 BP 神经网络的土壤全氮含量估测

与主成分分析相比,PLS 同时考虑了光谱自变量和因变量的影响,提取的主成分更具有

代表性^[19]。因此,在前文基础上,进一步选取经 Norris 平滑的一阶导数光谱(500~900 nm、1 350~1 490 nm),基于 PLS 进行主成分提取。结果显示,前 10 个主成分能够解释原始光谱变量的 99.68% [图 4(a)],说明前 10 个主成分基本代表了原始的光谱信息。将这 10 个主成分作为 BP 网络的输入节点,输出节点为土壤全氮含量,选取 Tansig 函数作为隐层传递函数, Purelin 为输出层函数, Trainlm 为训练函数,学习速率设为 0.01,最大训练次数为 3 000,模型期望误差为 0.001。当隐含层节点为 28 时,模型预测效果达到最好,当训练次数达到 1 300 时,网络训练达到目标要求,模型训练图如图 4(b),网络的拟合精度达到了 0.98 [图 5(a)],满足精确估测土壤全氮含量的要求。

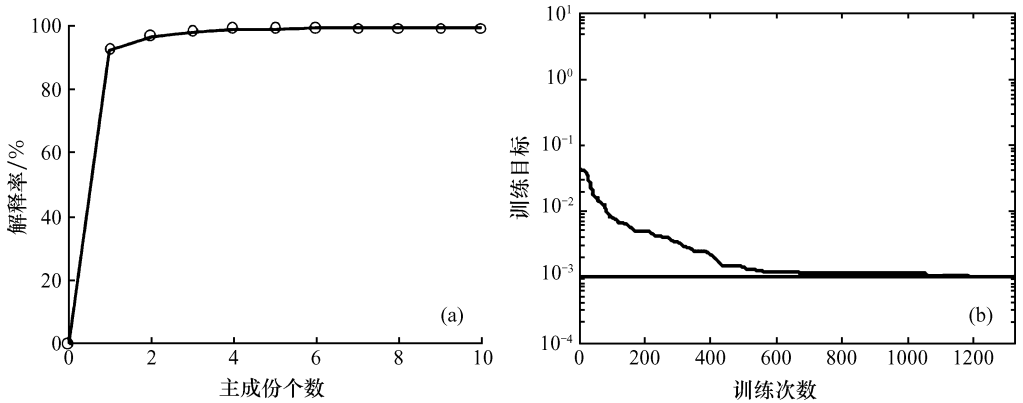


图 4 压缩主成分解释图(a)和 BP 神经网络训练曲线(b)

Fig. 4 Accumulative reliabilities plot of the first 10 principal components by PLS (a) and training curve of BP neural network (b)

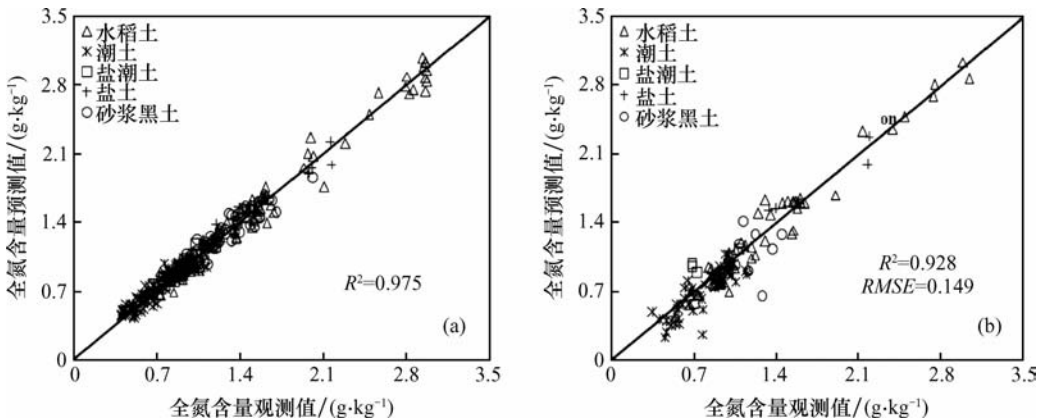


图 5 利用 BP 神经网络对土壤建模样本的拟合(a)和基于 BP 神经网络模型的预测(b)

Fig. 5 Calibration result based on BP neural network (a) and validation result based on BP neural network (b)

2.5 土壤全氮含量估测模型的检验与比较

为了比较 3 种不同方法(偏最小二乘法、BP 神经网络及特征光谱指数)对土壤全氮含量的预测能力,利用独立于建模样本以外的试验资料对上述估测模型进行了测试和检验。结果显示,基于 PLS、BPNN 及 $DI(NDR_{872}, NDR_{1482})$ 三种方法的模型预测决定系数分别为 0.81、0.93 和 0.54, $RMSE$ 分别为 $0.22 \text{ g} \cdot \text{kg}^{-1}$ 、 $0.15 \text{ g} \cdot \text{kg}^{-1}$ 和 $0.31 \text{ g} \cdot \text{kg}^{-1}$, RPD 分别为

2.28、3.36 和 1.60 [图 6(a)、5(b) 和 6(b)]。表明采用 BP 神经网络建立的土壤全氮估测模型具有较高的预测精度,其次是 PLS,而采用 $DI(NDR_{872}, NDR_{1482})$ 的预测结果稍差, $DI(NDR_{872}, NDR_{1482})$ 用于土壤全氮预测时还有待于进一步测试和完善。

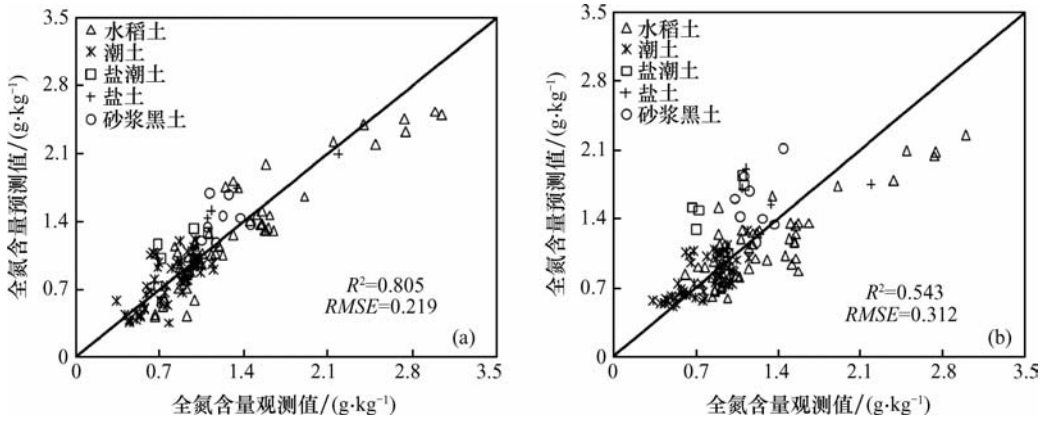


图 6 基于 PLS(a) 和 $DI(NDR_{872}, NDR_{1482})$ (b) 的土壤全氮含量观测值和预测值的比较

Fig. 6 Comparisons between observed and predicted total nitrogen content based on different methods

3 讨论与结论

光谱预处理能够较好地消除土壤类型(化学形态、物理形态如颗粒度)及所处环境(温度、湿度、pH 值等)等因素的影响,从而突出光谱反射率与养分含量之间的相关性^[20]。徐永明等^[5]运用一阶导数、倒数、反射率倒数的对数以及波段深度等光谱形式研究了与总氮含量的关系,表明利用一阶导数和反射率倒数的对数预测土壤总氮精度最高;Zornoza 等^[21]利用多元散射校正和一阶导数对近红外光谱进行预处理,提高了对土壤全氮含量的预测精度;Brunet 等^[22]比较了一阶导数、二阶导数及归一化处理等预处理方法,显示一阶导数光谱数据对土壤全氮含量的估测结果较好。本研究结果证实了这一点,不同类型土壤高光谱数据,经 Norris 滤波平滑后的一阶导数与全氮含量具有良好的相关性。

借鉴植被光谱分析方法^[23],系统分析了高光谱 350~2 500 nm 波段范围内任意两波段的比值、差值和归一化指数与土壤全氮含量的关系,原始光谱和一阶导数与全氮含量的定量关系均以差值光谱指数建立的预测模型结果最好。其中,872 nm 和 1 482 nm 2 个波段的一阶导数光谱经 Norris 滤波平滑后所构建的导数差值指数 $DI(NDR_{872}, NDR_{1482})$ 表现最好。然而由于 N-H 键多以二阶以上倍频或合频存在,相比较 O-H、C-H 键在土壤光谱上反应弱^[24],从而影响了土壤全氮信息的提取。因此,基于 $DI(NDR_{872}, NDR_{1482})$ 的土壤全氮含量预测模型还有待于进一步测试和完善。

以往较多研究利用 PLS 采用全波段光谱建模,如徐永明等^[25]利用 350~2 500 nm 范围高光谱建立了土壤全氮的估测模型,独立样本检验的决定系数为 0.8。于飞健等^[12]利用近红外光谱 800~2 500 nm 波段光谱信息估测了土壤全氮含量。本文分析了不同类型土壤全氮含量与高光谱的相关关系,表明在 350~2 500 nm 范围内的部分高光谱波段存在较大噪音,从而有必要剔除出噪音较大的波段,进而选取敏感波段进行 PLS 建模。结果显示,相比全波段(350~2 500 nm)而言,采用 500~900 nm、1 350~1 490 nm 波段范围经 Norris 滤波平滑后的一阶导数光谱为建模区域,能够显著提高对土壤全氮含量的预测精度。此外,相比于以

往研究^[14],本文建立的 BPNN 模型是以 PLS 提取的主成分光谱作为输入变量,该方法同时考虑了光谱自变量和因变量的信息,提取的有效特征成分具有良好的代表性,模型对不同类型土壤的全氮含量具有较好的预测性。比较而言,BP 神经网络模型预测精度较高,其次是偏最小二乘法,可能是因为土壤全氮含量与反射光谱之间是一种非线性关系,PLS 更适合线性建模,而 BPNN 能进行良好的非线性关联^[26]。基于 $DI(NDR_{872}, NDR_{1482})$ 的模型预测精度低于前两种方法,但它需要的波段信息少,所建立的模型简单,是一种较有潜力的土壤全氮快速估测方法。基于土壤和高光谱采样的局限性,本文所建模型还需利用更为广泛的土壤类型资料进行测试和完善。

参考文献 (References):

- [1] Bendor E, Banin A. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties[J]. *Soil Science Society of America Journal*, 1995, 59: 364-372.
- [2] 屈晓晖, 庄大方, 彭望碌, 等. 基于 ANN 分类的农田遥感动态监测模型研究[J]. 自然资源学报, 2007, 22(2): 193-197. [QU Xiao-hui, ZHUANG Da-fang, PENG Wang-lu, et al. Studies on remote sensing dynamic detection model of cropland based on the classification of artificial neural network. *Journal of Natural Resources*, 2007, 22(2): 193-197.]
- [3] 张春桂, 张星, 陈敏敏, 等. 福建近岸海域悬浮泥沙浓度遥感定量监测研究[J]. 自然资源学报, 2008, 23(1): 150-160. [ZHANG Chun-gui, ZHANG Xing, CHEN Min-yan, et al. Study on remote sensing quantitative model of suspended sediments in the coastal waters of Fujian. *Journal of Natural Resources*, 2008, 23(1): 150-160.]
- [4] Dalal R C, Henry R J. Simultaneous determination of moisture, organic carbon, and total nitrogen by near-infrared reflectance spectrophotometry[J]. *Soil Science Society of America Journal*, 1986, 50: 120-123.
- [5] 徐永明, 蔺启忠, 黄秀华, 等. 利用可见光/近红外反射光谱估算土壤总氮含量的实验研究[J]. 地理与地理信息科学, 2005, 21(1): 19-22. [XU Yong-ming, LIN Qi-zhong, HUANG Xiu-hua, et al. Experimental study on total nitrogen concentration in soil by VNIR reflectance spectrum. *Geography and Geo-Information Science*, 2005, 21(1): 19-22.]
- [6] 卢艳丽, 白由路, 王磊, 等. 黑土土壤中全氮含量的高光谱预测分析[J]. 农业工程学报, 2010, 26(1): 256-261. [LU Yan-li, BAI You-lu, WANG Lei, et al. Determination for total nitrogen content in black soil using hyperspectral data. *Transactions of the CSAE*, 2010, 26(1): 256-261.]
- [7] Lee W S, Mylavarapu R S, Choe J S, et al. Study on soil properties and spectral characteristics in Florida[R]. American Society for Aerospace Education Paper, 2001.
- [8] Reeves J, McCarty G, Mesinger J. Near infrared reflectance spectroscopy for the analysis of agricultural soils[J]. *Journal of Near Infrared Spectroscopy*, 1999, 7: 179-193.
- [9] Chang C, David Laird A. Near-infrared reflectance spectroscopic analysis of soil C and N[J]. *Soil Science*, 2002, 167: 110-116.
- [10] 孙建英, 李民赞, 郑立华, 等. 基于近红外光谱的北方潮土土壤参数实时分析[J]. 光谱学与光谱分析, 2006, 26(3): 426-429. [SUN Jian-ying, LI Min-zan, ZHENG Li-hua, et al. Real-time analysis of soil moisture, soil organic matter, and soil total nitrogen with NIR spectra. *Spectroscopy and Spectral Analysis*, 2006, 26(3): 426-429.]
- [11] 赵锁芳, 彭玉魁. 我国黄土区土壤水分、有机质和总氮的近红外光谱分析[J]. 分析化学, 2002, 30(8): 978-980. [ZHAO Suo-lao, PENG Yu-kui. Analysis of soil moisture, organic matter and total nitrogen content in loess in China with near infrared spectroscopy. *Chinese Journal of Analytical Chemistry*, 2002, 30(8): 978-980.]
- [12] 于飞健, 闵顺耕, 巨晓棠, 等. 近红外光谱法分析土壤中的有机质和氮素[J]. 分析实验室, 2002, 1(3): 49-51. [YU Fei-jian, MIN Shun-geng, JU Xiao-tang, et al. Determination the content of nitrogen and organic substance in dry soil by using near infrared diffusion reflectance spectroscopy. *China Journal of Analysis Laboratory*, 2002, 1(3): 49-51.]
- [13] 张雪莲, 李晓娜, 武菊英, 等. 不同类型土壤总氮的近红外光谱技术测定研究[J]. 光谱学与光谱分析, 2010, 30(4): 906-910. [ZHANG Xue-lian, LI Xiao-na, WU Ju-ying, et al. Study on the determination of total nitrogen (TN) in different types of soil by near-infrared spectroscopy (NIS). *Spectroscopy and Spectral Analysis*, 2010, 30(4): 906-910.]
- [14] 郑立华, 李民赞, 潘奕, 等. 基于近红外光谱技术的土壤参数 BP 神经网络预测[J]. 光谱学与光谱分析, 2008,

- 28(5): 1160-1164. [ZHENG Li-hua, LI Min-zan, PAN Yi, *et al.* Estimation of soil organic matter and soil total nitrogen based on NIR spectroscopy and BP neural network. *Spectroscopy and Spectral Analysis*, 2008, 28(5): 1160-1164.]
- [15] 张娟娟, 田永超, 朱艳, 等. 不同类型土壤的光谱特征及其有机质含量预测研究[J]. 中国农业科学, 2009, 42(9): 3154-3163. [ZHANG Juan-juan, TIAN Yong-chao, ZHU Yan, *et al.* Spectral characteristics and estimating of organic matter contents of different soil types. *Scientia Agricultura Sinica*, 2009, 42(9): 3154-3163.]
- [16] Savitzky A, Golay M J E. Smoothing and differentiation of data by simplified least squares procedures[J]. *Analytical Chemistry*, 1964, 36(8): 1627-1639.
- [17] Norris K H, Williams P C. Optimization of mathematical treatments of raw near-infrared signal in the measurement of protein in hard Red Spring wheat. I. Influence of particle size[J]. *Cereal Chemistry*, 1984, 62(2): 158-165.
- [18] 李强, 赵伟. MATLAB 数据处理与应用[M]. 北京: 国防工业出版社, 2001: 1-60. [LI Qiang, ZHAO Wei. *Data Processing and Application in MATLAB*. Beijing: National Defense Industry Press, 2001: 1-60.]
- [19] 刘波平, 秦华俊, 罗香, 等. PLS-BP 法近红外光谱同时检测饲料组分的研究[J]. 光谱学与光谱分析, 2007, 27(10): 2005-2009. [LIU Bo-ping, QIN Hua-jun, LUO Xiang, *et al.* Determination of four contents of feedstuff powder using near infrared spectroscopy by PLS-BP model. *Spectroscopy and Spectral Analysis*, 2007, 27(10): 2005-2009.]
- [20] 丁海泉, 卢启鹏, 朴仁官, 等. 土壤有机质近红外光谱分析组合波长的优选[J]. 光学精密工程, 2007, 12(15): 1946-1957. [DING Hai-quan, LU Qi-peng, PU Ren-gong, *et al.* Optimum choice of combination wavelengths in near infrared analysis for soil organic matter. *Optics and Precision Engineering*, 2007, 15(12): 1946-1951.]
- [21] Zornoza R, Guerrero C, Mataix-Solera J. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils[J]. *Soil Biology & Biochemistry*, 2008, 40: 1923-1930.
- [22] Brunet D, Barthes B G, Chotte J L, *et al.* Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity[J]. *Geoderma*, 2007, 139: 106-117.
- [23] 田永超, 杨杰, 姚霞, 等. 高光谱植被指数与水稻叶面积指数的定量关系[J]. 应用生态学报, 2009, 20(7): 1685-1690. [TIAN Yong-chao, YANG Jie, YAO Xia, *et al.* Quantitative relations between hyperspectral vegetation index and rice leaf area index. *Chinese Journal of Applied ecology*, 2009, 20(7): 1685-1690.]
- [24] 李民赞. 光谱分析技术及其应用[M]. 北京: 科学出版社, 2006. [LI Min-zan. *Spectral Analysis Technology and Application*. Beijing: Science Press, 2006.]
- [25] 徐永明, 蔺启忠, 王璐, 等. 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(9): 709-716. [XU Yong-ming, LIN Qi-zhong, WANG Lu, *et al.* Model for estimating soil nutrient elements based on high resolution reflectance spectra. *Acta Pedologica Sinica*, 2006, 43(9): 709-716.]
- [26] 陆婉珍. 现代近红外光谱分析技术[M]. 北京: 中国石化出版社, 2000. [LU Wan-zhen. *Modern Near Infrared Spectral Analysis Technology*. Beijing: China Petrochemical Press, 2000.]

Estimating Soil Total Nitrogen Content Based on Hyperspectral Analysis Technology

ZHANG Juan-juan^{1,2}, TIAN Yong-chao¹, YAO Xia¹, CAO Wei-xing¹, MA Xin-ming², ZHU Yan¹

(1. National Engineering and Technology Center for Information Agriculture, Jiangsu Key Laboratory for Information Agriculture, Nanjing Agricultural University, Nanjing 210095, China; 2. College of Information and Management, Henan Agricultural University, Zhengzhou 450002, China)

Abstract: Quantitative relationships between soil total nitrogen content (TN) and hyperspectra in visible and near-infrared region (VIS-NIR) (350 – 2500 nm) were studied for five soil types (paddy soil, fluvo-aquic soil, salinized fluvo-aquic soil, saline soil, dark soil with lime concretion) collected from central and East China. Based on three different methods of spectral index, partial least square (PLS) and back propagation neural network (BPNN), the models were developed for estimating TN content in soil. The results showed that the newly developed PLS and BPNN models for estimating TN content based on the corrected first derivative spectra of 500 – 900 nm and 1350 – 1490 nm regions with Norris smoothing filter performed well, with R^2 of calibration as 0.81 and 0.98, respectively. The R^2 , $RMSE$ and RPD of validation were 0.81, 0.219 $g \cdot kg^{-1}$ and 2.28 for the method of PLS, and were 0.93, 0.149 $g \cdot kg^{-1}$ and 3.36 for the method of BPNN, respectively. In addition, $DI(NDR_{872}, NDR_{1482})$ composed of the corrected first derivative spectra of 872 nm and 1482 nm with Norris smoothing algorithm also had a good correlation with soil TN content. Testing of the estimating model based on $DI(NDR_{872}, NDR_{1482})$ with independent datasets from different types of soil samples resulted in R^2 , $RMSE$ and RPD as 0.66, 0.53 $g \cdot kg^{-1}$ and 1.60, respectively. Comparison of the above three methods, the sequence of prediction accuracy was PLS-BPNN model > PLS > $DI(NDR_{872}, NDR_{1482})$, which indicated that the newly developed BPNN and PLS models were reliable for estimating soil TN content with high prediction accuracy, and $DI(NDR_{872}, NDR_{1482})$ maybe a good indicator of soil TN content.

Key words: soil; total nitrogen; hyperspectra; partial least square; BP neural network; spectral index