

基于 DWT-GA-PLS 的土壤碱解氮含量高光谱估测方法 *

陈红艳¹ 赵庚星^{1**} 李希灿² 王向锋³ 李玉玲⁴

(¹土肥资源高效利用国家工程实验室/山东农业大学资源与环境学院, 山东泰安 271018; ²山东农业大学信息科学与工程学院, 山东泰安 271018; ³垦利县国土资源局, 山东垦利 257500; ⁴齐河县农业局, 山东齐河 251100)

摘要 以山东齐河县为研究区, 实地采集土壤样本, 在土样高光谱测试并进行一阶导数变换的基础上, 先运用离散小波变换(DWT)对土壤光谱去噪降维, 然后采用遗传算法(GA)筛选土壤碱解氮定量估测模型的参与变量, 最后应用偏最小二乘(PLS)回归构建土壤碱解氮含量的估测模型。结果表明: 离散小波变换结合遗传算法和偏最小二乘法(DWT-GA-PLS)用于土壤碱解氮含量定量估测, 不仅可压缩光谱变量、减少模型参与变量, 而且可改善模型估测准确度; 较之于采用土壤全谱, 小波离散分解1~2层低频系数构建的模型在参与变量大幅减少的情况下, 取得更准确或与之相当的预测结果, 其中, 基于第2层小波低频系数采用GA筛选变量构建的PLS模型的预测效果表现最好, 预测 R^2 达到0.85, RMSE为8.11 mg·kg⁻¹, RPD为2.53。说明DWT-GA-PLS用于土壤碱解氮含量高光谱定量估测的有效性。

关键词 碱解氮 高光谱 离散小波变换 遗传算法

文章编号 1001-9332(2013)11-3185-07 **中图分类号** S127, S153.2 **文献标识码** A

Hyper spectral estimation method for soil alkali hydrolysable nitrogen content based on discrete wavelet transform and genetic algorithm in combining with partial least squares (DWT-GA-PLS). CHEN Hong-yan¹, ZHAO Geng-xing¹, LI Xi-can², WANG Xiang-feng³, LI Yu-ling⁴ (¹National Engineering Laboratory for Efficient Utilization of Soil and Fertilizer Resources, College of Resources and Environment, Shandong Agricultural University, Tai'an 271018, Shandong, China; ²College of Information Science and Engineering, Shandong Agricultural University, Tai'an 271018, Shandong, China; ³Kenli County Bureau of Land and Resources, Kenli 257500, Shandong, China; ⁴Qihe Bureau of Agriculture, Qihe 251100, Shandong, China). -Chin. J. Appl. Ecol., 2013, 24(11): 3185–3191.

Abstract: Taking the Qihe County in Shandong Province of East China as the study area, soil samples were collected from the field, and based on the hyperspectral reflectance measurement of the soil samples and the transformation with the first deviation, the spectra were denoised and compressed by discrete wavelet transform (DWT), the variables for the soil alkali hydrolysable nitrogen quantitative estimation models were selected by genetic algorithms (GA), and the estimation models for the soil alkali hydrolysable nitrogen content were built by using partial least squares (PLS) regression. The discrete wavelet transform and genetic algorithm in combining with partial least squares (DWT-GA-PLS) could not only compress the spectrum variables and reduce the model variables, but also improve the quantitative estimation accuracy of soil alkali hydrolysable nitrogen content. Based on the 1–2 levels low frequency coefficients of discrete wavelet transform, and under the condition of large scale decrement of spectrum variables, the calibration models could achieve the higher or the same prediction accuracy as the soil full spectra. The model based on the second level low frequency coefficients had the highest precision, with the model predicting R^2 being 0.85,

* 国家自然科学基金项目(41271235)、国家科技计划项目(2013BAD05B06, 2011BAD21B0601)、高校博士点基金项目(20103702110010)和山东省自主创新专项(2012CX90202)资助。

** 通讯作者. E-mail: zhaogx@sdaau.edu.cn

2013-03-19 收稿, 2013-09-02 接受.

the RMSE being $8.11 \text{ mg} \cdot \text{kg}^{-1}$, and RPD being 2.53, indicating the effectiveness of DWT-GA-PLS method in estimating soil alkali hydrolysable nitrogen content.

Key words: alkali hydrolysable nitrogen; hyper spectra; discrete wavelet transform; genetic algorithm.

近年来,随着精准农业的发展,应用光谱技术快速准确获取土壤参数信息已经成为研究热点。土壤氮素是植物生长发育的基础,自20世纪80年代光谱数据应用于土壤氮素含量估测开始,土壤氮素的定量遥感估测已成为最近几年研究较多的方向之一^[1-2]。

采用适当的光谱预处理和波长选择方法能够简化和提取特征光谱信息,提高土壤参数定量模型的预测能力^[3-4]。在现有光谱预处理方法中,小波变换(wavelet transform, WT)在信号去噪和数据压缩方面具有独特的优势^[5-7],如刘炜等^[6]对土壤原始光谱及其A值的一阶导数进行小波去噪处理,改善了它们与土壤有机质的相关性;Viscarra-Rosset 和 Lark^[7]结合小波变换和回归方法预测土壤有机碳含量,不仅提高了模型决定系数,而且压缩了数据、简化了模型。但单纯依赖小波系数建模,参与变量仍较多,土壤参数模型的实用性受到影响^[8],仍需探索一定的优化选择方法筛选建模变量。遗传算法(genetic algorithm, GA)作为一种通过模拟自然进化过程并行搜索最优解的方法^[9],在光谱分析中,结合偏最小二乘(partial least squares, PLS)可用于多组分分析的波长选择或谱区选择,其研究和应用日益成为热点^[10-12]。采用高分辨率的土壤全谱作为遗传算法的搜索空间,其波长数目及优化搜索空间十分巨大,从而影响寻优结果,因此现有研究多采用谱区选择^[13-14],如温健婷等^[15]采用遗传算法进行土壤重金属的谱区选择,并进行偏最小二乘建模,获得了较高的土壤重金属高光谱估测精度;Li等^[16]基于遗传算法结合偏最小二乘在植物水分近红外光谱分析模型中进行谱区选择,简化了分析模型,增强了模型的稳健性。其中,适宜谱区区间大小和数目的选择是关键。

本文采用离散小波变换对土壤光谱进行降维去噪,将小波低频系数作为遗传算法的注入变量,进而筛选土壤碱解氮估测模型的参与变量,构建其含量的偏最小二乘回归估测模型,旨在探索土壤碱解氮含量的高光谱估测方法,以简化模型,提高估测的准确度和稳定性。

1 材料与方法

1.1 土壤样本及碱解氮化学分析

本文选择的研究区为山东齐河县,境内为黄河冲积扇平原,土壤类型为潮土,肥力水平较高,其粘土矿物以水云母为主。在研究区首先根据土壤质地、肥力等级划分采样单元,每个采样单元的平均面积为 26.67 hm^2 ,采用网格布点方式,每个采样单元采用10点混合的方式采集1个混合土样,采样深度为0~20 cm,取土1 kg左右,共采集土样71个。将土样自然风干、敲碎,剔除其他侵入体,过筛(2 mm),并混合均匀,用四分法取样200 g,分成两份,一份装于盛样皿(直径10 cm,深2 cm)中,稍稍压平,使其表面平整,用于土壤高光谱数据的采集;另一份再过1 mm筛用于土壤碱解氮含量的碱解扩散法测定,先用碱液(NaOH)处理过筛土样,易水解的有机氮及铵态氮转化为氨,硝态氮则先经硫酸亚铁转化为铵,以硼酸吸收氨,再用标准酸(H₂SO₄)滴定,并做空白试验,然后计算水解性氮含量。

1.2 光谱测定及预处理

土壤高光谱数据采用美国ASD Fieldspec 3光谱仪暗室采集,并对其进行噪音波段去除、断点校正等处理,光谱测定条件、操作及预处理见文献[8, 17]。将400~2450 nm的反射率作为土壤原始光谱。

1.3 研究方法

1.3.1 光谱变换 本文对光谱反射率进行一阶导数变换,以消除低频噪声对目标光谱的影响,导数变换采用的计算公式见文献[8]。

1.3.2 基于离散小波变换(discrete wavelet transform, DWT)的光谱预处理方法 对土壤导数光谱进行离散小波变换,舍去小波高频系数,提取小波低频系数,能够一定程度上剔除由光谱仪精度、测试条件等不确定因素影响的高频噪音,得到光谱基本信息^[18-19]。目前,常用的小波系有Haar、Daubechies(DB n)、Biorthogonal、Symlet等,笔者曾在文献[8]中探讨了小波适宜分解尺度并分析了上述常用的4种小波函数的表现,研究表明,1~5层的小波低频系数与原始光谱的相关性较高,各小波母函数对土壤养分建模精度影响不大,其中,Bior 1.3建模精度

略高。因此,本文对土壤反射率的一阶导数光谱采用 Bior 1.3 小波母函数进行 1~5 层离散分解,提取不同分解层的小波低频系数,以用于遗传算法筛选变量,相关计算采用 Matlab 7.0 软件编程实现。

1.3.3 基于遗传算法和偏最小二乘 (GA-PLS) 的变量筛选 本文采用遗传算法和偏最小二乘法结合 (GA-PLS) 筛选模型变量。首先将土壤全谱和上述提取的小波低频系数作为基因,并用 0/1 字符进行二进制编码,将土壤碱解氮含量作为优化目标,随机生成初始群体;然后采用 PLS 交叉校验均方根误差 (RMSECV) 作为适应度函数,当 RMSECV 达到最小或没有显著变化时,所筛选的变量和确定的因子数用于建模预测,经过 5 次重复筛选后,确定选中频率最高的变量参与建模。在遗传算法中,种群大小、交叉、变异算子等是主要影响因素,种群越大,收敛效果越好,但计算量也会增大;对于交叉概率,一般染色体都会参与交叉,但要考虑交叉后解的可行性和运算速度;变异是为了避免把个体局限在较优的局部解空间内,因此变异概率值通常设置较低^[20]。经过多次反复计算,考虑到运算量和运算精确度,确定相应参数为:种群大小 30,最大繁殖代数 100,交叉概率 0.5,变异概率 0.01。采用 Matlab 7.0 软件编程实现变量筛选。

1.3.4 基于偏最小二乘 (PLS) 回归的估测模型建立与验证 采用每次剔除样本组中的一个样品进行交叉检验建模验证。首先将土样分组,按照土壤碱解氮含量进行排序,每间隔 1 个样本选择 3 个样本作为建模样本集,其余样本为验证样本集,并把土壤碱解氮含量相同的 2 个建模样本再分开,有 1 个划入了验证集。71 个样本分为两组,一组 51 个用于建模,另一组 20 个用于验证。然后利用建模样本,运用 SAS 9.0 软件,采用 PLS 回归,分别对土壤全谱、1~5 层小波低频系数和遗传算法筛选的各层变量建立土壤碱解氮含量高光谱估测模型。PLS 分析采用舍一交叉验证法 (leave one out cross validation) 确定主成分数。并利用验证样本数据检验模型,模型精度用 R^2 、RMSE 和 RPD 进行评价。

2 结果与分析

2.1 样本土壤碱解氮含量状况

利用统计方法分析 71 个样本数据,样本碱解氮含量最小值是 $53.0 \text{ mg} \cdot \text{kg}^{-1}$,最大值是 $169.0 \text{ mg} \cdot \text{kg}^{-1}$,平均值为 $111.24 \text{ mg} \cdot \text{kg}^{-1}$,标准差为 $25.51 \text{ mg} \cdot \text{kg}^{-1}$ 。

2.2 基于 DWT 提取的小波低频系数

利用 71 个土壤样本的数据,对土壤反射率的一阶导数光谱采用 Bior 1.3 小波母函数进行 1~5 层离散分解,提取小波低频系数,将小波系数进行光谱重构并与土壤全谱进行相关性分析。由表 1 可以看出,小波系数的数目随着分解层数的增加而逐渐减少,下层数目约为上层数目的 $1/2$,数据逐渐得到压缩;与土壤全谱的相关性也逐渐减小,即小波系数反映输入光谱信息的能力逐渐减弱。其中,1~3 层低频系数和输入光谱数据的相关系数大于 0.90,足以表示光谱的整体特征,可用以代表输入土壤光谱;4~5 层的相关系数大于 0.5,具有一定的相关性。

2.3 基于 GA-PLS 筛选的变量

基于土壤全谱和各分解层的小波低频系数,采用遗传算法结合偏最小二乘法筛选土壤碱解氮定量模型的入选变量。采用遗传算法注入变量数目应少于 200 个,以避免过度拟合现象的发生,对于光谱数据,可通过求取若干个连续波长的平均值,减少注入变量^[21]。因此,对于变量数目超过 200 个的土壤全谱和 1~3 层小波低频系数,本文先采用一定的间隔划分为若干变量区间。

光谱区间大小直接影响遗传算法的准确度。为确定最适宜的区间划分大小,对于土壤全谱,本文基于全部土壤样本的平均光谱,按照波长范围 2 倍递增并且就近取奇数值的原则,分别采用波长 3、7、15、31 和 71 nm 为区间大小均分原始光谱,然后求取不同大小的区间光谱平均和最大标准偏差(表 2)。标准偏差越小,区间数据越集中,反之,数据越分散。相对于其他区间大小变化,区间大小由 7 nm 减少到 3 nm,平均和最大标准偏差变化均较小,但大于 7 nm 的光谱区间标准差的差异显著。为了既保障区间数据相近,又尽量减少变量数目,本文选择 7 nm 作为划分土壤全谱区间的适宜大小。

表 1 各层小波低频系数及其重构光谱与原始光谱的相关性

Table 1 Low frequency wavelet coefficients and the correlation between its restructuring spectrum and original spectrum

分解层数 Decomposing level	小波系数数目 Wavelet coefficient number	压缩率 Compression rate (%)	相关系数 Correlation coefficient (R)
0	2051	100	1.00
1	1028	50.1	1.00
2	516	25.2	0.96
3	260	12.7	0.93
4	132	6.4	0.80
5	68	3.3	0.56

小波低频系数下层数目约为上层数目的 $1/2$,参考土壤全谱7 nm的划分间隔和遗传算法注入变量低于200的要求,对1~3层小波低频系数分别采用4、3、2的间隔划分区间。

对划分的区间求取平均值并作为遗传算法的注入变量,将土壤碱解氮含量作为优化目标,运用GA重复筛选5次,然后将选中区间内原波长点或系数作为最终的建模变量。

由表3可以看出,采用遗传算法筛选变量,选中变量数目相对小波系数大幅减少,最少只有22个,最多为147个。可见,在小波系数的基础上,采用遗

表2 区间大小分析

Table 2 Analysis on region size

区间大小 Region size (nm)	平均标准偏差 Average SD	最大标准偏差 Max SD	区间数目 Region number
3	1.65×10^{-6}	8.77×10^{-4}	684
7	1.98×10^{-6}	9.86×10^{-4}	293
15	5.49×10^{-5}	1.53×10^{-3}	137
31	1.05×10^{-4}	1.80×10^{-3}	66
71	1.75×10^{-4}	1.32×10^{-3}	29

传算法可减少模型参与变量。

对比光谱与各层小波系数曲线的相似性可见,曲线整体趋势和走向保持一致,但随着小波分解层数的增多,反映光谱噪声和微小吸收特征的信息逐渐被剔除,光谱的基本信息被渐进反映;第1~2层小波低频系数与原始光谱较相似,在反映土壤光谱基本特征的同时保留了许多细节,诸如400~470、

表3 小波系数及GA筛选的变量

Table 3 Wavelet coefficients and selected variables based on GA

分解层数 Decompo- sing level	小波系数 数目 Wavelet coefficient number	区间 间隔 Interval	GA筛选 变量数目 Selected variable number based on GA	变量占小波系数 的百分比 Percentage of selected variable to wavelet coefficient
0	2051	7	147	7.2
1	1028	4	111	10.8
2	516	3	43	8.3
3	260	2	25	9.6
4	132	1	37	28.0
5	68	1	22	32.4

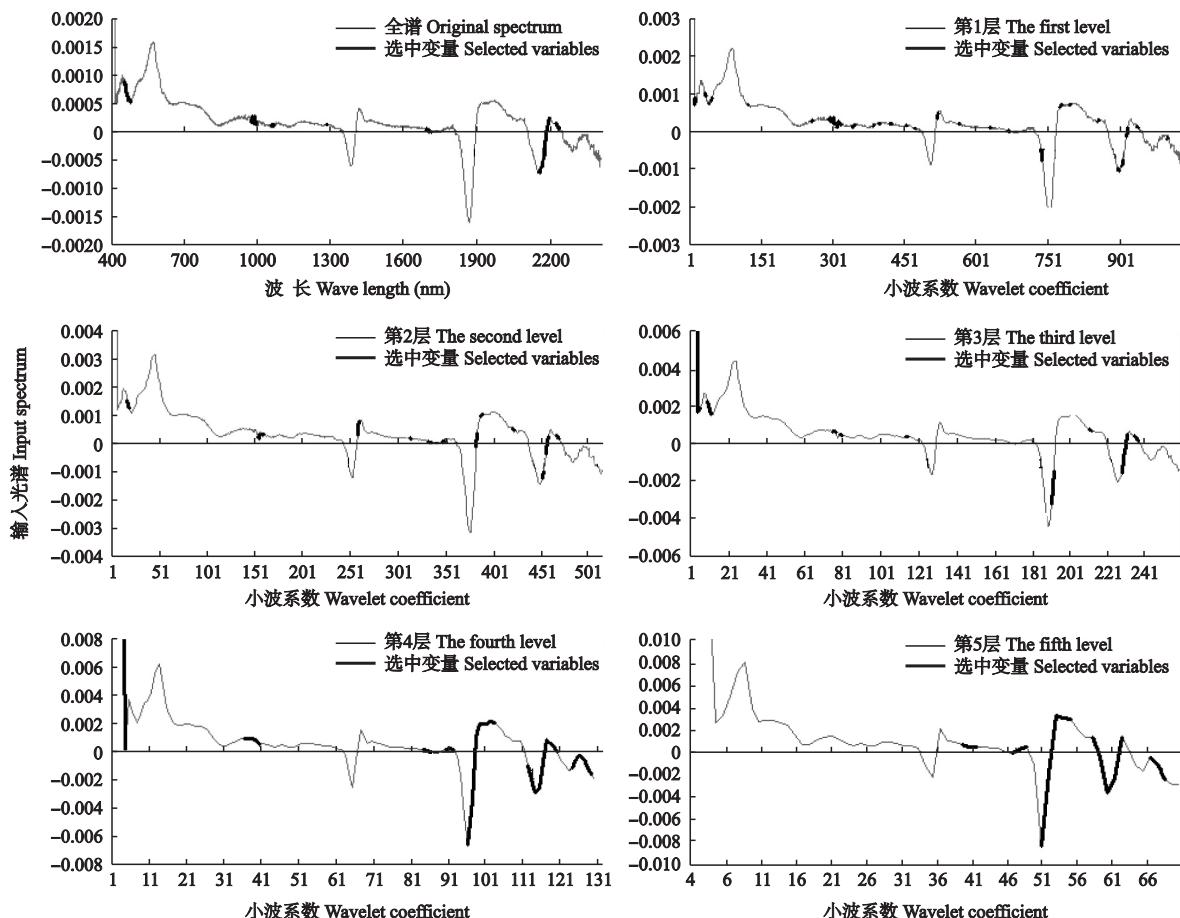


图1 全谱、1~5层小波低频系数及GA筛选的变量

Fig. 1 Original spectrum, wavelet coefficients of levels 1~5 and selected variables based on GA.

900~1100、1700~1800 nm 及相应的小波系数中仍有明显的毛刺,而 3~5 层小波低频系数曲线较光滑,显然剔除了较多的细节特征(图 1)。比较选中变量位置的相近性可见,相对原始光谱,第 1~2 层小波低频系数筛选的变量大多位置与之较一致,而且表现为零散地分布在整个光谱或小波系数区间;3~5 层小波低频系数筛选的变量位置与之差异较大,而且大多集中连片分布,尤其是 4~5 层小波低频系数筛选的变量。

2.4 基于 PLS 的土壤碱解氮含量高光谱估测模型

分别基于土壤全谱、小波系数和遗传算法筛选的变量,采用 PLS 回归构建土壤碱解氮含量估测模型,并进行模型验证和比较。

由表 4 可以看出,基于离散小波变换结合偏最小二乘(DWT-PLS)构建的土壤碱解氮含量估测模型中,土壤全谱和第 1~2 层低频系数建模精度较高,第 3~5 层低频系数建模精度较低;而且基于第 1~2 层低频系数建模精度与土壤全谱基本相当,建模集决定系数(R^2)在 0.67~0.71,均方根差(RMSE)在 11.88~12.26 mg·kg⁻¹,验证集 R^2 在 0.77~0.78, RMSE 在 10.59~10.64 mg·kg⁻¹,相对分析误差(RPD)大于 2.00,可用于土壤碱解氮含量的估测;土壤全谱建模变量达 2051 个,第 1 层和第

2 层小波低频系数分别约是全谱的 1/2 和 1/4。可见,基于小波变换 1~2 层低频系数既可压缩模型变量,又可用于土壤碱解氮含量的高光谱估测。

由表 5 可以看出,基于离散小波变换结合遗传算法和偏最小二乘(DWT-GA-PLS)的土壤碱解氮含量估测模型,与基于离散小波变换结合偏最小二乘(DWT-PLS)的估测模型(表 4)在精度变化趋势上趋于一致,即土壤全谱和第 1~2 层低频系数模型精度较高,第 3~5 层低频系数模型精度较低,而且基于第 1~2 层低频系数的模型较之土壤全谱,在参与变量大幅减少的前提下,模型精度略高或基本相当。不同的是,采用 DWT-GA-PLS 得到的模型精度(表 5)普遍优于 DWT-PLS(表 4);基于土壤全谱和第 1~2 层低频系数,DWT-GA-PLS 对土壤碱解氮含量具有较准确的预测性,相对 DWT-PLS 的模型决定系数 R^2 提高了 4%~14%,均方根差降低 0.46~2.53 mg·kg⁻¹, RPD 提高 0.3~0.5;且 DWT-GA-PLS 构建的模型参与变量也明显少于 DWT-PLS,约是其对应系数的 8% 左右。由此可见,相对 DWT-PLS,采用 DWT-GA-PLS 可在模型参与变量大幅减少的情况下,精度得到一定幅度的提高。根据预测精度确定最佳模型为:基于第 2 层小波低频系数采用 GA 筛选变量构建的 PLS 模型,该模型参与变量 43 个,建模

表 4 基于离散小波变换结合偏最小二乘(DWT-PLS)的土壤碱解氮估测模型

Table 4 Estimation models of soil alkali hydrolysable nitrogen content based on discrete wavelet transform and partial least squares

小波系数 Wavelet coefficient	建模集 Calibration set		预测集 Validation set			主成分数量 Component number	变量数量 Variable number
	R^2	RMSE (mg·kg ⁻¹)	R^2	RMSE (mg·kg ⁻¹)	RPD		
ca0	0.71	11.88	0.78	10.63	2.09	3	2051
ca1	0.70	11.97	0.78	10.59	2.10	3	1028
ca2	0.67	12.26	0.77	10.64	2.05	3	516
ca3	0.37	12.60	0.61	11.76	1.38	2	260
ca4	0.35	12.39	0.59	11.80	1.26	2	132
ca5	0.30	11.98	0.49	12.10	1.04	2	68

表 5 基于离散小波变换结合遗传算法和偏最小二乘(DWT-GA-PLS)的土壤碱解氮估测模型

Table 5 Estimation models of soil alkali hydrolysable nitrogen content based on discrete wavelet transform in combination with genetic algorithm and partial least squares

小波系数 Wavelet coefficient	建模集 Calibration set		预测集 Validation set			主成分数量 Component number	变量数量 Variable number
	R^2	RMSE (mg·kg ⁻¹)	R^2	RMSE (mg·kg ⁻¹)	RPD		
ca0	0.82	9.97	0.85	8.68	2.48	6	147
ca1	0.80	10.49	0.82	9.58	2.31	5	111
ca2	0.71	11.80	0.85	8.11	2.53	7	43
ca3	0.66	12.34	0.69	11.48	1.76	6	25
ca4	0.54	12.99	0.64	11.66	1.63	4	37
ca5	0.36	12.53	0.50	12.02	1.42	3	22

表 6 3 种方法的模型结果与比较

Table 6 Model results and comparison of three methods

方法 Method	建模集 Calibration set		验证集 Validation set			变量数量 Variable number
	R^2	RMSE ($\text{mg} \cdot \text{kg}^{-1}$)	R^2	RMSE ($\text{mg} \cdot \text{kg}^{-1}$)	RPD	
相关分析 Correlation analysis	0.39	21.03	0.35	22.05	0.73	4
多元逐步回归 Stepwise multiple linear regression	0.78	13.27	0.72	14.27	1.38	8
DWT-GA-PLS	0.71	11.80	0.85	8.11	2.53	43

R^2 为 0.71, 建模 RMSE 为 $11.80 \text{ mg} \cdot \text{kg}^{-1}$, 预测 R^2 达到 0.85, 预测 RMSE 为 $8.11 \text{ mg} \cdot \text{kg}^{-1}$ (为样本碱解氮平均含量的 7.4%), 预测 RPD 为 2.53, 具有较好的预测准确性.

2.5 与传统特征波段筛选方法的比较

基于试验数据, 分别利用相关分析、多元逐步回归分析筛选的特征波段构建土壤碱解氮含量估测的回归模型, 其中, 多元逐步回归设定变量的贡献显著水平为 0.05, 方差膨胀因子小于 10, 以消除变量间多重共线性的影响. 与 DWT-GA-PLS 方法的最佳结果进行比较可见, DWT-GA-PLS 方法与传统的相关分析和多元逐步回归法相比, 模型参与变量相对较多, 但模型的预测精度和稳定性有大幅提高, 可更好地用于土壤碱解氮估测(表 6).

3 讨 论

本文采用 DWT-GA-PLS 估测土壤碱解氮含量, 有效降低了模型复杂度, 改善了预测准确度. 现有研究中, 也有学者采用小波变换与遗传算法结合用于土壤参数的光谱分析, 并且与本文的研究结论相似^[22-24], 如方利民等^[23]基于可见/近红外光谱将离散小波变换、快速独立分量分析、遗传算法和 BP 神经网络结合建立了土壤有机碳和阳离子交换量的 BP 分析模型, 并获得了较好的预测准确度. 原因在于小波离散变换可消除光谱噪声, 压缩光谱变量, 注入遗传算法可简化搜索空间; 遗传算法可进一步筛选待测参数的特征变量从而优化模型; 加之, 本文采用偏最小二乘回归可消除变量间的多重共线性, 这些可有效简化模型, 减少模型随机性, 提高模型精度, 因此 DWT-GA-PLS 用于土壤碱解氮含量的高光谱估测具有较好的效果.

遗传算法可从全局最优角度搜索使目标函数值较优的变量, 是一种较理想的特征选择工具^[25], 但该方法收敛速度慢、结构参数选取存在随机性. 为提高 GA 分析的有效性、准确性和稳定性, 也有学者对 GA 注入变量进行了研究, 如陈斌等^[26]将相关系数法生成的初始化波长作为启发性信息注入 GA 算法

中, 简化搜索空间, 引导搜索方向, 从而达到快速收敛的目的; 李鹏飞等^[27]利用反向区间偏最小二乘法(BiPLS)对原始光谱波长进行初步定位, 再用遗传算法进行二次选择, 减少了运算时间, 提高了模型的预测能力和精度. 在光谱消噪降维中, 小波变换是最常用和有效的工具, 本文将离散小波变换的低频系数作为遗传算法的注入变量, 将小波变换的压缩消噪和遗传算法的特征选择相结合, 同以往单纯将小波系数作为模型变量相比^[8], 在进一步减少建模变量的基础上, 提高了预测精度和稳定性, 取得了更好的土壤碱解氮含量高光谱估测效果, 对从复杂、重叠、变动的光谱信号中提取有用信息并建立校正模型具有积极的参考意义.

无论对小波低频系数及 GA 变量筛选结果的分析, 还是对所构建估测模型精度的分析, 基于 Bior 1.3 函数 1~2 层小波低频系数构建的估测模型精度均相当或略优于土壤全谱, 而入选变量数目却大幅减少, 尤以第 2 层小波分解效果最佳. 本文仅针对山东省齐河县的潮土类型和土壤碱解氮进行估测, 对于其他差异较大的土壤类型或其他土壤参数的适用性还有待研究核实.

在土壤光谱划分区间大小选择时, 按照“波长范围 2 倍递增并且就近取奇数值的原则”, 本文确定 7 nm 为最适宜的光谱区间划分大小, 对于其他的土壤全谱或 1~3 层小波系数区间大小划分方案, 研究结果是否相同, 还有待进一步研究探索.

本文基于室内光谱, 采用 DWT-GA-PLS 构建的最佳模型有 43 个参与变量, 在室内可采用计算机技术或程序语言快速、自动实现碱解氮含量估测, 对田间原状土碱解氮估测具有参考意义, 但推广到田间原位实时测量, 仍需进一步探索. 今后将采用田间原位光谱进行土壤碱解氮估测研究.

参考文献

- [1] Couillard A, Turgeon AJ, Shenk JS, et al. Near infrared reflectance spectrometry for analysis of turf soil profiles. *Crop Science*, 1997, 37: 1554-1559
- [2] Zhang J-J (张娟娟), Tian Y-C (田永超), Yao X (姚

- 霞), et al. Application of wavelet packet analysis in estimating soil parameters based on NIR spectra: Spectroscopy and spectral analysis. *Transactions of the Chinese Society of Agricultural Engineering* (农业工程学报), 2012, **28**(12): 183–188 (in Chinese)
- [3] Ehsani MR, Upadhyaya SK, Fawcett WR. Feasibility of detecting soil nitrate content using a mid-infrared technique. *Transactions of the American Society of Agricultural Engineers*, 2001, **44**: 1931–1940
- [4] Shen Y (申 艳), Zhang X-P (张晓平), Liang A-Z (梁爱珍), et al. Near infrared spectroscopy in determining organic carbon and total nitrogen in black soil of Northeast China. *Chinese Journal of Applied Ecology* (应用生态学报), 2010, **21**(1): 109–114 (in Chinese)
- [5] Cheng T, Rivard B, Sánchez-Azofeifa A. Spectroscopic determination of leaf water content using continuous wavelet analysis. *Remote Sensing of Environment*, 2011, **115**: 659–670
- [6] Liu W (刘 炜), Chang Q-R (常庆瑞), Guo M (郭曼), et al. Application of wavelet transformation in detection of organic matter content based on visible/near infrared reflectance spectroscopy. *Agricultural Research in the Arid Areas* (干旱地区农业研究), 2010, **28**(5): 241–246 (in Chinese)
- [7] Viscarra-Rossel RA, Lark RM. Improved analysis and modelling of soil diffuse reflectance spectra using wavelets. *European Journal of Soil Science*, 2009, **60**: 453–464
- [8] Chen H-Y (陈红艳), Zhao G-X (赵庚星), Li X-C (李希灿), et al. Application of wavelet analysis for estimation of soil available potassium content with hyperspectral reflectance. *Scientia Agricultura Sinica* (中国农业科学), 2012, **45**(7): 1425–1431 (in Chinese)
- [9] Chu X-L (褚小立), Yuan H-F (袁洪福), Lu W-Z (陆婉珍). Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique. *Progress in Chemistry* (化学进展), 2004, **16**(4): 528–542 (in Chinese)
- [10] Michael F, Julio C, Leyden F, et al. Genetic algorithm optimization in drug design QSAR: Bayesian-regularized genetic neural networks (BRGNN) and genetic algorithm-optimized support vectors machines (GA-SVM). *Molecular Diversity*, 2011, **15**: 269–289
- [11] Elliott GN, Worgan H, Broadhurst D, et al. Soil differentiation using fingerprint fourier transform infrared spectroscopy, chemometrics and genetic algorithm-based feature selection. *Soil Biology & Biochemistry*, 2007, **39**: 2888–2896
- [12] Zou X-B (邹小波), Zhao J-W (赵杰文). Methods of characteristic wavelength region and wavelength selection based on genetic algorithm. *Acta Optica Sinica* (光学学报), 2007, **27**(7): 1316–1321 (in Chinese)
- [13] Tu Z-H (屠振华), Ji B-P (籍保平), Meng C-Y (孟超英), et al. Analysis of NIR characteristic wavelengths for apple flesh firmness based on GA and iPLS. *Spectroscopy and Spectral Analysis* (光谱学与光谱分析), 2009, **29**(10): 2760–2764 (in Chinese)
- [14] Wang J-H (王加华), Han D-H (韩东海). Analysis of near infrared spectra of apple SSC by genetic algorithm optimization. *Spectroscopy and Spectral Analysis* (光谱学与光谱分析), 2008, **28**(10): 2308–2311 (in Chinese)
- [15] Weng J-T (温健婷), Zhang X (张 霞), Zhang B (张 兵), et al. A study of band selection method for retrieving soil lead content with hyperspectral remote sensing data. *Advances in Earth Science* (地球科学进展), 2010, **25**(6): 625–629 (in Chinese)
- [16] Li L, Cheng Y, Ustin S, et al. Retrieval of vegetation equivalent water thickness from reflectance using genetic algorithm(GA)-partial least squares (PLS) regression. *Advances in Space Research*, 2008, **41**: 1755–1763
- [17] Chen H-Y (陈红艳), Zhao G-X (赵庚星), Li X-C (李希灿), et al. Hyper-spectral estimation of soil organic matter content based on wavelet transformation. *Chinese Journal of Applied Ecology* (应用生态学报), 2011, **22**(11): 2935–2942 (in Chinese)
- [18] Si BC. Spatial scaling analysis of soil physical properties: A review of spectral and wavelet methods. *Vadose Zone Journal*, 2008, **7**: 547–562
- [19] Zheng L-H (郑丽华), Li M-Z (李民赞), Pan L (潘凌), et al. Application of wavelet packet analysis in estimating soil parameters based on NIR spectra. *Spectroscopy and Spectral Analysis* (光谱学与光谱分析), 2009, **29**(6): 1549–1552 (in Chinese)
- [20] Wang X-P (王晓平), Cao L-M (曹立明). Genetic Algorithm: Theory, Application, Soft Implementation. Xi'an: Xi'an Jiaotong University Press, 2002 (in Chinese)
- [21] Millán CP, Forina M, Casolino C, et al. Extraction of representative subsets by potential functions method and genetic algorithms. *Chemometrics and Intelligent Laboratory Systems*, 1998, **40**: 33–52
- [22] Wang G-Q (王国庆), Shao X-G (邵学广). A discrete wavelet transform-genetic algorithm-cross validation approach for high ratio compression and variable selection of near-infrared spectral data. *Chinese Journal of Analytical Chemistry* (分析化学), 2005, **33**(2): 191–194 (in Chinese)
- [23] Fang L-M (方利民), Feng A-M (冯爱明), Lin M (林 敏). Rapid prediction of total organic carbon content and CEC in soil using visible/near infrared spectroscopy. *Spectroscopy and Spectral Analysis* (光谱学与光谱分析), 2010, **30**(2): 327–330 (in Chinese)
- [24] Wang MH, Tseng YF. A novel analytic method of power quality using extension genetic algorithm and wavelet transform. *Expert Systems with Application*, 2011, **38**: 12491–12496
- [25] Cho HW, Kim SB, Jeong MY, et al. Genetic algorithm-based feature selection in high-resolution NMR spectra. *Expert Systems with Applications*, 2008, **35**: 967–975
- [26] Chen B (陈 斌), Wang H (王 豪), Lin S (林 松), et al. Analysis of near infrared spectroscopy of beer alcohol content by correlation coefficients and genetic algorithms. *Transactions of the Chinese Society of Agricultural Engineering* (农业工程学报), 2005, **21**(7): 99–102 (in Chinese)
- [27] Li P-F (李鹏飞), Wang J-H (王加华), Cao N-N (曹楠宁), et al. Selection of variables for MLR in Vis/NIR spectroscopy based on BiPLS combined with GA. *Spectroscopy and Spectral Analysis* (光谱学与光谱分析), 2009, **29**(10): 2637–2641 (in Chinese)

作者简介 陈红艳,女,1979年生,博士,讲师。主要从事农业遥感与信息技术研究,发表论文13篇。E-mail: chenhy@sdu.edu.cn

责任编辑 杨 弘