

激光与光电子学进展

基于模拟退火的土壤钾含量近红外光谱波段优选

李春婷¹, 张优优¹, 陈华舟^{1*}, 辜洁², 莫丽娜³, 李小玉¹

¹桂林理工大学理学院, 广西 桂林 541004;

²重庆人文科技学院机电与信息工程学院, 重庆 401524;

³桂林旅游学院旅游数据学院, 广西 桂林 541006

摘要 土壤钾含量对作物的品质和产量有重要影响,因此,用近红外(NIR)光谱技术结合特征波段筛选方法对土壤钾含量进行快速定量检测。首先,结合模拟退火算法和区间偏最小二乘法(SA-iPLS)建立土壤钾含量的特征波段优化 NIR 模型。然后,通过调试子区间数量得到 SA-iPLS 的最优特征波段。最后,根据模型的评价指标将 SA-iPLS 模型与偏最小二乘法(PLS)、区间 PLS(iPLS)、协同 iPLS(SiPLS)模型进行对比。结果表明,当子区间数为 90 时,SA-iPLS 模型在训练集上的表现最优,对测试集的预测均方根误差和相关系数分别为 0.0117 和 0.8884。全谱 PLS 模型对测试集样品的预测均方根误差和相关系数分别为 0.0140 和 0.8506,iPLS、SiPLS 模型的最优子区间数分别为 80、70,对测试集样品的预测均方根误差分别为 0.0155、0.0145,相关系数分别为 0.7786、0.8420。相比常规的 iPLS、SiPLS 模型,SA-iPLS 模型能保留更多有用的光谱信息,提高对土壤钾含量的预测精度。

关键词 光谱学; 近红外光谱; 土壤钾含量; 波段优选; 模拟退火算法; 区间偏最小二乘法

中图分类号 O657.33

文献标志码 A

DOI: 10.3788/LOP202259.1330002

Near-Infrared Spectral Waveband Selection for Soil Potassium Content Based on Simulated Annealing

Li Chunting¹, Zhang Youyou¹, Chen Huazhou^{1*}, Gu Jie², Mo Lina³, Li Xiaoyu¹

¹College of Science, Guilin University of Technology, Guilin 541004, Guangxi, China;

²School of Electrical and Information Engineering, Chongqing College of Humanities, Science & Technology, Chongqing 401524, China;

³School of Tourism Data, Guilin Tourism University, Guilin 541006, Guangxi, China

Abstract Soil potassium content affects the quality and yield of crops. In this study, near-infrared (NIR) spectroscopy is combined with a method of selecting feature wavebands for the rapid quantitative detection of soil potassium content. First, the NIR model for the feature waveband optimization of the soil potassium content is established by combining the simulated annealing algorithm and interval partial least squares (SA-iPLS). Then, the optimal feature waveband of SA-iPLS is obtained by adjusting the number of subintervals. Finally, the SA-iPLS model is compared with the partial least squares (PLS), interval PLS (iPLS), and synergy iPLS (SiPLS) models according to the evaluation indicators of the model. The results show that the SA-iPLS model exhibited the best performance on the training set when the number of subintervals is 90, and the prediction root mean square error and correlation coefficient of the test set are 0.0117 and 0.8884, respectively. The prediction root mean square error and correlation coefficient of the full-spectrum PLS model for the test set samples are 0.0140 and 0.8506, respectively. The optimal number of subintervals for the iPLS and SiPLS models are 80 and 70, respectively; the prediction root mean square errors for the test set samples are 0.0155 and 0.0145, respectively; the correlation coefficients are 0.7786 and 0.8420, respectively. Compared to the conventional iPLS and SiPLS models, the SA-iPLS model can retain more useful spectral information and improve the prediction accuracy of the soil potassium content.

Key words spectroscopy; near-infrared spectroscopy; soil potassium content; waveband optimization; simulated annealing algorithm; interval partial least square

收稿日期: 2021-07-08; 修回日期: 2021-08-01; 录用日期: 2021-08-18

基金项目: 国家自然科学基金(61505037)、广西自然科学基金(2018GXNSFAA050045)

通信作者: hzchengut@foxmail.com

1 引言

钾是作物生长代谢所需的重要常量元素,测定土壤中的钾含量可以了解土壤的供钾潜力、制定施肥决策,以保证作物品质和产量^[1-3]。传统化学实验方法检测土壤中钾含量的成本高、过程繁琐,对实验室条件有一定的要求,且使用化学试剂,容易造成环境污染。近红外(NIR)光谱分析技术利用待测样品的光谱矩阵进行建模,对待测成分进行定性或定量分析,具有快速、非破坏性、无污染以及增加实验结果对比性等优点,已被广泛应用到农业、食品、药物、炼油与化工等领域^[4-6]。

土壤光谱中包含大量土壤属性信息,复杂的土壤光谱特性和人为实验误差会使光谱数据中包含很多冗余信息和随机噪声,影响后续的建模效果。不同的土壤待测物质会在某些特定光谱区域表现出较高的灵敏度^[7-8],因此建模前常常需要进行波段优选。偏最小二乘法(PLS)是一种在全光谱区域进行建模的定量分析方法^[9]。近年来,人们将波段优选和PLS相结合,先后提出了区间PLS(iPLS)和协同iPLS(SiPLS),并分析了模型的优化程度^[10-12]。这两种方法均是根据一定的原则和运行过程筛选出灵敏度较高的波段,然后对模型进行优化。iPLS对全光谱检测区域进行等宽划分,在各个区间构建PLS模型,通过对比各子区间模型的预测精度选择最优特征波段,但该波段优选方法存在建模区间单一的问题^[13]。SiPLS模型将其中若干个子区间联合起来,通过对比各联合区间的模型性能选择最优特征波段,但该模型需要遍历整个解空间,计算量和计算时长会随划分子区间数和联合子区间数的增加而快速增长^[14]。目前,已有研究将遗传算法(GA)、蚁群算法(ACO)等智能算法与iPLS模型进行结合,以优化光谱波段,从而有效剔除冗余信息,使模型的预测能力和精度得到一定程度的提高^[15-16]。GA的全局搜索能力极强,ACO具有正反馈机制、分布式计算的优点,但这两种算法都易陷于局部最优解。模拟退火(SA)算法是模拟固体升温再冷却过程的一种随机组

合优化算法,具有操作简单、运行效率高、能获得全局近似最优解等优点^[17],降温过程在解空间搜索新解,Metropolis准则以一定概率接受恶化解,随着温度的下降,对恶化解的接受能力逐渐降低,使算法逐渐收敛,从而跳出局部最优解,并在较短的时间内达到全局近似最优解^[18]。因此,该算法常被应用到NIR光谱分析的参数组合优化,以解决典型的局部优化问题^[19-20]。

本文结合模拟退火算法与区间偏最小二乘法(SA-iPLS)进行波段优选,对广西某地土壤钾含量建立SA-iPLS模型。首先,以训练集样品的均方根误差为目标函数进行迭代寻优,设置冷却表和终止迭代条件,筛选出能反映样品信息的最优特征波段。然后,对比SA-iPLS模型与iPLS、SiPLS两种波段优选模型的性能,以在建模时减少土壤钾含量NIR光谱定量分析模型计算量和计算时长的同时增加更多有用的光谱信息,提高模型的预测精度。

2 实验部分

2.1 实验数据

土壤样品来自广西桂南、桂中、桂东北地区的几个农田场地,每个场地选择15~20个采样点,在每个采样点使用5点采样法,采样深度为20 cm,得到具有代表性的135份样品。样品经过风干、碾磨并过2 mm筛。采用火焰光度法测定每个土壤样品的钾含量,作为NIR模型分析的化学参考值。经统计分析,所有样品的化学参考值范围为0.053%~0.192%,标准偏差和平均值分别为0.031%和0.107%。利用Perkin-Elmer公司的Spectrum One NTS傅里叶变换NIR光谱仪检测样品光谱,每份土壤样品经系统自动扫描64次后输出平均光谱,光谱测量范围为10000~4000 cm^{-1} ,光谱数据采集间隔约为4 cm^{-1} ,全谱段共包括1512个波长点。采用Savitzky-Golay卷积平滑(SG平滑)^[21]、标准正态变换(SNV)和小波包变换(WPT,小波函数为bior3.6,尺度参数为7)^[22]对广西土壤光谱数据进行预处理,135份广西土壤的原始NIR光谱和预处理后的光谱如图1所示。

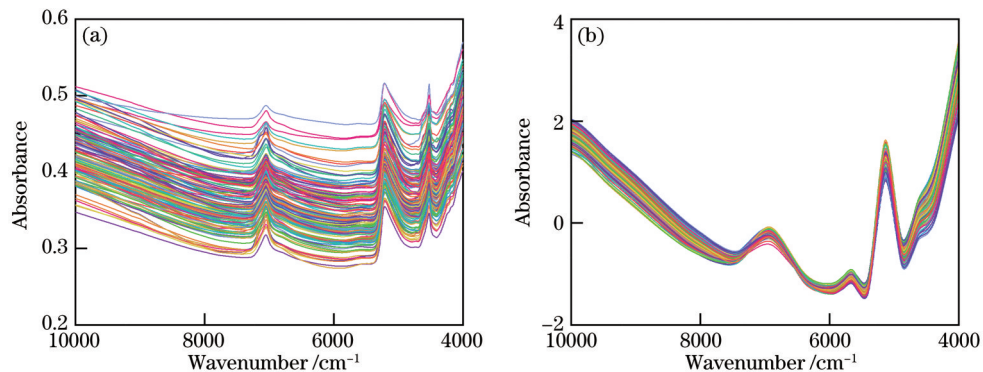


图1 135份土壤样品的NIR光谱图。(a)原始光谱;(b)预处理后的光谱

Fig. 1 NIR spectra of 135 soil samples. (a) Raw spectra; (b) spectra obtained after preprocessing

2.2 样品集划分与模型评价指标

建模过程中需要将样品划分为训练集和测试集,采用基于联合 X - Y 距离的样品集划分 (SPXY) 方法按 6:1 的比例进行划分,训练集包含 115 个样品,测试集包含 20 个样品。各样品集的统计数据如表 1 所示。此外,需要对训练集样品和测试集样品进行评价,采用的评价指标为均方根误差 (R_{MSE}) 和相关系数 (R),可分别表示为

$$R_{MSE} = \sqrt{\frac{\sum_{j=1}^n (y_j - y'_j)^2}{n-1}}, \quad (1)$$

表 1 样品集中钾元素的质量分数

Table 1 Mass fraction of potassium element in the sample set

Sample set	Number of samples	Maximum	Minimum / %	Average / %	Standard deviation / %
Calibration set	115	0.192	0.053	0.108	0.032
Test set	20	0.162	0.065	0.101	0.023

2.3 近红外光谱波段优选算法

iPLS 模型是将全光谱检测区域划分为等宽的 N 个区间,分别在每个子区间构建 PLS 模型,并选取 R_{MSE}^V 最小的子区间作为 iPLS 模型的优选波段。SiPLS 在 iPLS 的基础上,对其中 m 个子区间进行穷尽排列组合,组合成一个联合区间,共建立 C_N^m 个 PLS 模型,以 R_{MSE}^V 最小为原则选择区间,作为 SiPLS 模型的最优特征波段, m 一般取为 2~4^[14]。SA-iPLS 模型是一种将 SA 算法以一定概率接受新解的思想运用到 iPLS 组合子区间过程,以得到最优联合区间的特征波段,主要参数包括等宽子区间数 N 、初始温度 T_0 、外循环最大迭代次数 N_c 、Metropolis 链长 L 、退温速率 q 。首先,将全光谱检测区域划分为 N 个等宽子区间,随机抽取若干个子区间组合成联合区间作为初始解,在联合区间上构建 PLS 模型,将训练集的 R_{MSE}^V 作为 SA 的适应度函数。然后,通过变异操作产生新解即新的子区间组合,计算新解对应的适应度函数,并根据 Metropolis 准则判断是否用该新解更替当前解;在相同温度下进行 L 次寻优,得到当前温度的最优解后利用退温函数 $T_{i+1} = q \times T_i$ 对温度进行退温;继续迭代寻优直至满足迭代终止条件,输出的最优解为 SA-iPLS 模型选择的最优特征波段。SA-iPLS 模型的迭代终止条件为外循环达到最大迭代次数 N_c ,此时该模型的交互检验次数稳定在 $N_c \times L$ 次。

3 结果与讨论

3.1 波段优选

在土壤样品的 NIR 全光谱区域构建 PLS 模型,交互检验得到的 R_{MSE}^V 为 0.0272。为对子区间划分数 N 做优化选择,设置 $N=60, 70, 80, 90, 100$,利用训练集样品对每一个 N 构建 iPLS、SiPLS 和 SA-iPLS 模型,

$$R = \frac{\sum_{j=1}^n (y_j - y_m)(y'_j - y'_m)}{\sqrt{\sum_{j=1}^n (y_j - y_m)^2 \sum_{j=1}^n (y'_j - y'_m)^2}}, \quad (2)$$

式中, y_j 为指定样品集(训练集或测试集)中第 j 个样品的参考化学值, y'_j 为指定样品集中第 j 个样品的 NIR 光谱预测值, y_m 为指定样品集中所有样品化学参考值的平均值, y'_m 为指定样品集中所有样品 NIR 光谱预测值的平均值, n 为指定样品集的样品个数。在建模过程中,训练集的两个模型评价指标分别记为 R_{MSE}^V 和 R_V ,测试集的两个模型评价指标分别记为 R_{MSE}^T 和 R_{T0} 。

各个模型在不同 N 下得到的最优特征波段 R_{MSE}^V 如图 2 所示。其中:iPLS 模型在划分子区间数为 80 时,交互检验得到的 R_{MSE}^V 最小,为 0.0269,对应的最优特征波段范围为 8036~7980 cm^{-1} ,包含 19 个波长点;SiPLS 模型在每一个划分子区间数 N 下,对 m 个子区间进行穷尽排列组合,组成 C_N^m 个待优化的特征波段,并对每个波段进行交互检验,选出 R_{MSE}^V 最小值的特征波段作为该 N, m 下 SiPLS 优选特征波段。一般情况下,选取 2~4 个子区间进行组合。在划分子区间数 N 为 60,组合区间数 m 为 2 时,模型交互检验次数为 1770 次, m 增加为 4 时,模型交互检验次数达到 487635 次;当划分子区间数 N 为 100,组合区间数 m 为 2 时,模型交互检验次数为 4950 次, m 为 3、4 时,交互检验次数增至 161700、3921225 次。这表明 SiPLS 模型的交互检验次数和计算时间会随划分子区间数 N 和联合子区间数 m 的增加呈指数增长。实验中将组合区间数 m 设置为 2,当划分 70 个子区间数时, SiPLS 模型对广西土

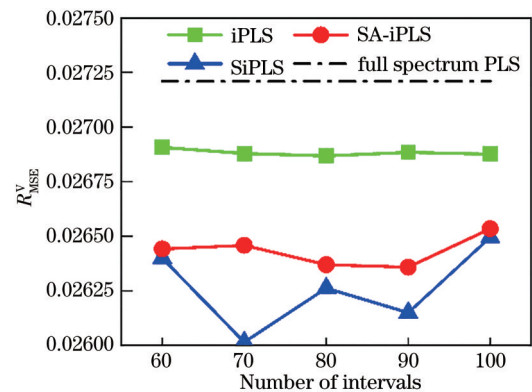


图 2 不同模型最优特征波段对应的 R_{MSE}^V
Fig. 2 R_{MSE}^V corresponding to the optimal feature waveband of different models

壤的钾含量 NIR 分析预测效果达到最优, R_{MSE}^V 为 0.0260, 对应的子区间组合为 SiPLS 优选特征波段, 波段范围为 $9168\sim 9089\text{ cm}^{-1}$ 、 $5395\sim 5311\text{ cm}^{-1}$, 波长点数为 43。

SA-iPLS 模型是将模拟退火算法运用到 iPLS 模型的子区间组合过程中, 使模型在趋向全局最优解的同时增加筛选的组合子区间个数以及可利用的信息变量, 且计算量和计算时长稳定、可调。算法参数 T_0 、 N_c 、 L 、 q 分别设置为 1000、500、10、0.95, 模型交互检验次数稳定在 5000 次。迭代的目标函数为每个子区间组合构建的 PLS 模型的 R_{MSE}^V , 500 次迭代优化的过程

如图 3 所示。可以发现: 算法迭代初期温度较大, 接受恶化的概率大, R_{MSE}^V 波动较大; 随着温度逐渐下降, 接受恶化的能力下降, R_{MSE}^V 逐渐趋于平稳。根据 R_{MSE}^V 最小原则获得各个子区间数 N 下最优的特征波段。从图 2 可以发现, 当子区间数为 90 时, SA-iPLS 模型在训练集上的性能最优, 对应的 R_{MSE}^V 为 0.0264。图 4 为 SA-iPLS 在划分子区间数为 90 时筛选的最优特征波段, 由 27 个子区间联合而成, 共 453 个波长点。综上所述, iPLS、SiPLS 和 SA-iPLS 模型在不同子区间划分情况下在训练集上的表现差别不大, 相对于全谱建模的 PLS 模型, 性能仅略有提升。

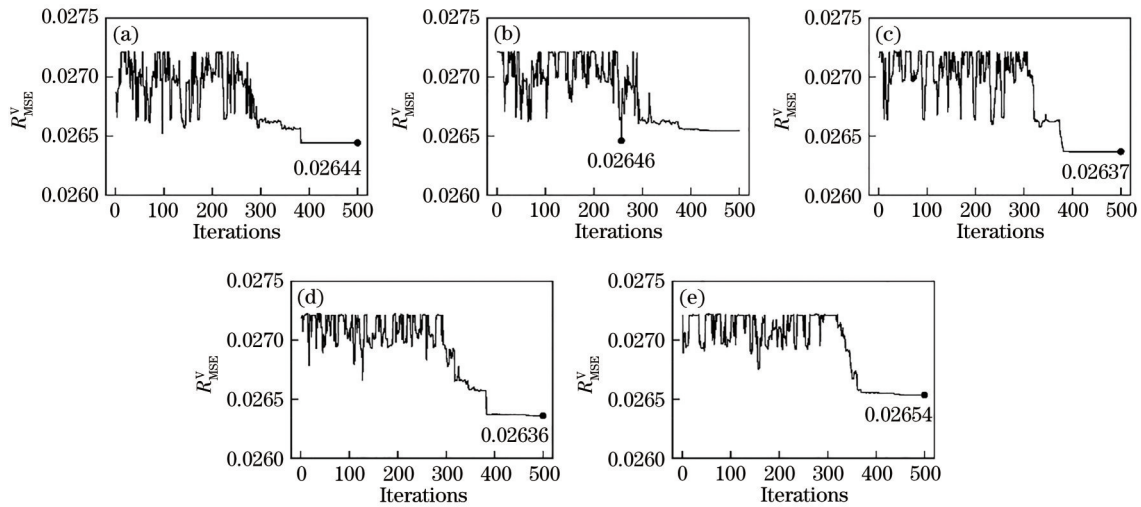


图 3 SA-iPLS 模型的迭代优化过程。(a) $N=60$; (b) $N=70$; (c) $N=80$; (d) $N=90$; (e) $N=100$

Fig. 3 Iterative optimization process of the SA-iPLS model. (a) $N=60$; (b) $N=70$; (c) $N=80$; (d) $N=90$; (e) $N=100$

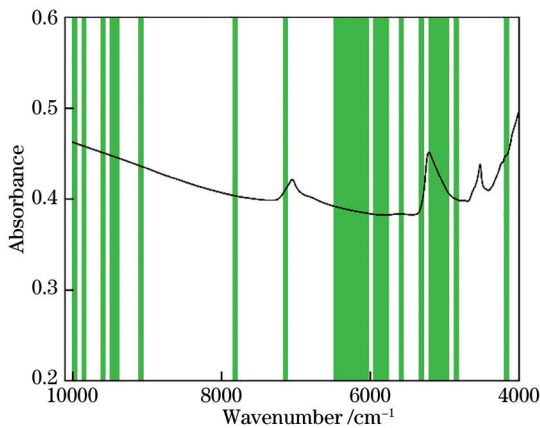


图 4 SA-iPLS 模型 ($N=90$) 筛选的最优特征波段

Fig. 4 Optimal feature waveband selected by SA-iPLS model ($N=90$)

用 iPLS、SiPLS 和 SA-iPLS 模型对采集的土壤光谱数据筛选最优特征波段, 结果如表 2 所示。可以发现, 不同的筛选波段方法得到最优特征波段子区间数 N 不同, iPLS 模型为 80, SiPLS 模型为 70, SA-iPLS 模型为 90。iPLS、SiPLS、SA-iPLS 的最优特征波段分别包括 1、2、27 个子区间, 16、37、453 个波长点, SA-iPLS

表 2 不同模型筛选的最优特征波段

Table 2 Optimal characteristic bands screened by different models

Model	Number of subintervals	Number of selected subintervals	Selected wavebands / cm^{-1}
iPLS	80	1	8036-7980
SiPLS	70	2	9168-9089, 5395-5311
SA-iPLS	90	27	10000-9942, 9875-9815, 9621-9561, 9494-9434, 9430-9371, 9112-9053, 7845-7782, 7170-7106, 6495-6431, 6427-6364, 6360-6296, 6292-6229, 6225-6161, 6157-6094, 6090-6026, 6022-5959, 5955-5891, 5887-5824, 5820-5756, 5617-5553, 5347-5283, 5212-5148, 5144-5081, 5077-5013, 5009-4946, 4874-4811, 4199-4135

建模利用的波长变量多于 iPLS、SiPLS 模型,以保留更多有用的光谱信息,且 SA-iPLS 模型的计算量和计算时长稳定、可调,不会出现 SiPLS 模型计算量和验证时间随划分子区间数和联合子区间数增加呈指数增长的问题。

3.2 模型预测

利用训练好的 PLS、iPLS、SiPLS、SA-iPLS 模型分别对测试集样品中钾含量进行预测,四种模型的建模效果如表 3 所示。可以发现:iPLS、SiPLS、SA-iPLS 波段优选模型的 R_{MSE}^V 分别为 0.0269、0.0260、0.0264,均低于全谱 PLS 模型的 R_{MSE}^V (0.0272);在测试集上,SA-iPLS 优选模型的表现最好,主因子数为 17, R_{MSE}^T 为 0.0117, R_T 为 0.8884,PLS、iPLS、SiPLS 模型的主因子数分别为 19、2、17, R_{MSE}^T 分别为 0.0140、0.0155、0.0145, R_T 分别为 0.8506、0.7786、0.8420。结果表明,SA-iPLS 方法能有效保留更多的有用光谱区间,筛选出灵敏度更高的特征波段,提高了模型的精度和预测能力,且 NIR 光谱分析模型交互检验次数更稳定可控。

表 3 不同模型的建模结果
Table 3 Modeling results of different models

Model	PLS factors	R_{MSE}^V	R_{MSE}^T	R_T
PLS	19	0.0272	0.0140	0.8506
iPLS	2	0.0269	0.0155	0.7786
SiPLS	17	0.0260	0.0145	0.8420
SA-iPLS	17	0.0264	0.0117	0.8884

采用上述子区间数和波段筛选过程得到的最优子区间数和最优特征波段的 SA-iPLS 模型预测 20 个广西土壤测试集样品的钾含量,其预测相关性如图 5 所示。可以发现,利用 SA-iPLS 模型对样品钾含量的预测结果与使用生化实验方法测定的测量值的相关度较高,这表明该模型的预测精度较高。

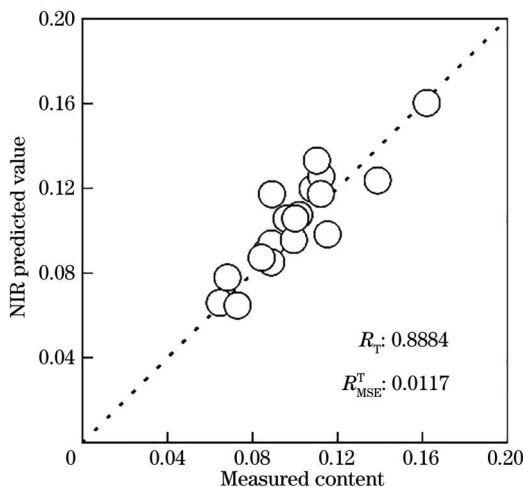


图 5 SA-iPLS 对测试集样品的预测结果

Fig. 5 Prediction results of SA-iPLS for the test samples

4 结 论

研究了 SA-iPLS 特征波段优选算法在土壤钾含量 NIR 光谱测定中的应用。设定不同的划分子区间数并分别构建 SA-iPLS 模型,获得的最优划分子区间数为 90,并得到由 27 个子区间组合成的最优特征波段。结果表明,该模型对测试集样品的 R_{MSE}^T 为 0.0117, R_T 为 0.8884。相比 iPLS、SiPLS 波段优选模型,SA-iPLS 增加了区间选择的灵活性,能有效保留更多的有用光谱区间,且 SA-iPLS 可以通过调节模型参数使模型计算量和计算时长稳定、可控,在测试集上的表现更优,能在土壤钾含量的定量快速检测上取得良好的预测效果,有利于光谱快速检测技术在现代农业中的应用。

参 考 文 献

- [1] Amtmann A, Hammond J P, Armengaud P, et al. Nutrient sensing and signalling in plants: potassium and phosphorus[J]. *Advances in Botanical Research*, 2005, 43: 209-257.
- [2] Amtmann A, Troufflard S, Armengaud P. The effect of potassium nutrition on pest and disease resistance in plants[J]. *Physiologia Plantarum*, 2008, 133(4): 682-691.
- [3] 董大明, 郑文刚, 赵春江, 等. 农田土壤中钾元素含量的激光诱导击穿光谱测量方法[J]. *光谱学与光谱分析*, 2013, 33(3): 785-789.
Dong D M, Zheng W G, Zhao C J, et al. Determination of potassium in farmland soil using laser-induced breakdown spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2013, 33(3): 785-789.
- [4] 郝勇, 吴文辉, 商庆园, 等. 山茶油中油酸和亚油酸近红外光谱分析模型[J]. *光学学报*, 2019, 39(9): 0930004.
Hao Y, Wu W H, Shang Q Y, et al. Analysis model of oleic and linoleic acids in camellia oil via near-infrared spectroscopy[J]. *Acta Optica Sinica*, 2019, 39(9): 0930004.
- [5] Chen H Z, Liu Z Y, Cai K, et al. Grid search parametric optimization for FT-NIR quantitative analysis of solid soluble content in strawberry samples[J]. *Vibrational Spectroscopy*, 2018, 94: 7-15.
- [6] 褚小立, 陈瀑, 李敬岩, 等. 近红外光谱分析技术的最新进展与展望[J]. *分析测试学报*, 2020, 39(10): 1181-1188.
Chu X L, Chen P, Li J Y, et al. Progresses and perspectives of near infrared spectroscopy analytical technology[J]. *Journal of Instrumental Analysis*, 2020, 39(10): 1181-1188.
- [7] 林卡, 李德成, 刘峰, 等. 基于可见-近红外反射光谱的土壤碳酸钙含量与反演效果关系研究[J]. *土壤学报*, 2018, 55(2): 304-312.
Lin K, Li D C, Liu F, et al. Study on relationship between soil calcium carbonate content and inversion effect based on visible near-infrared reflectance spectra[J]. *Acta Pedologica Sinica*, 2018, 55(2): 304-312.

- [8] 冉思, 丁建丽, 葛翔宇, 等. 基于稀疏网络的可见光/近红外反射光谱土壤有机质含量估算[J]. 激光与光电子学进展, 2020, 57(24): 242803.
Ran S, Ding J L, Ge X Y, et al. Estimation method of VIS-NIR spectroscopy for soil organic matter based on sparse networks[J]. Laser & Optoelectronics Progress, 2020, 57(24): 242803.
- [9] Nørgaard L, Saudland A, Wagner J, et al. Interval partial least-squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy[J]. Applied Spectroscopy, 2000, 54(3): 413-419.
- [10] Chen Q S, Zhao J W, Liu M H, et al. Determination of total polyphenols content in green tea using FT-NIR spectroscopy and different PLS algorithms[J]. Journal of Pharmaceutical and Biomedical Analysis, 2008, 46(3): 568-573.
- [11] Mo L N, Chen H Z, Chen W H, et al. Study on evolution methods for the optimization of machine learning models based on FT-NIR spectroscopy[J]. Infrared Physics & Technology, 2020, 108: 103366.
- [12] Barra I, Haefele S M, Sakrabani R, et al. Soil spectroscopy with the use of chemometrics, machine learning and pre-processing techniques in soil diagnosis: recent advances-a review[J]. TrAC Trends in Analytical Chemistry, 2021, 135: 116166.
- [13] Sampaio P S, Soares A, Castanho A, et al. Optimization of rice amylose determination by NIR-spectroscopy using PLS chemometrics algorithms[J]. Food Chemistry, 2018, 242: 196-204.
- [14] 孔庆明, 谷俊涛, 高睿, 等. 基于组合偏最小二乘的特征波段优选方法在氨、碱化处理玉米秸秆粗蛋白检测中的研究[J]. 分析测试学报, 2020, 39(11): 1334-1343.
Kong Q M, Gu J T, Gao R, et al. Study on detection of crude protein in ammonified and alkalinized corn straw by spectrum characteristic band selection method based on synergy interval partial least squares[J]. Journal of Instrumental Analysis, 2020, 39(11): 1334-1343.
- [15] Huang X W, Zou X B, Zhao J W, et al. Measurement of total anthocyanins content in flowering tea using near infrared spectroscopy combined with ant colony optimization models[J]. Food Chemistry, 2014, 164: 536-543.
- [16] Song X Z, Huang Y, Yan H, et al. A novel algorithm for spectral interval combination optimization[J]. Analytica Chimica Acta, 2016, 948: 19-29.
- [17] Sodeifian G, Sajadian S A, Ardestani N S. Experimental optimization and mathematical modeling of the supercritical fluid extraction of essential oil from *Eryngium billardieri*: application of simulated annealing (SA) algorithm[J]. The Journal of Supercritical Fluids, 2017, 127: 146-157.
- [18] 石吉勇, 邹小波, 赵杰文, 等. BiPLS 结合模拟退火算法的近红外光谱特征波长选择研究[J]. 红外与毫米波学报, 2011, 30(5): 458-462.
Shi J Y, Zou X B, Zhao J W, et al. Selection of wavelength for strawberry NIR spectroscopy based on BiPLS combined with SAA[J]. Journal of Infrared and Millimeter Waves, 2011, 30(5): 458-462.
- [19] Guo Z M, Barimah A O, Yin L M, et al. Intelligent evaluation of taste constituents and polyphenols-to-amino acids ratio in matcha tea powder using near infrared spectroscopy[J]. Food Chemistry, 2021, 353: 129372.
- [20] Liu J M, Jin S, Bao C H, et al. Rapid determination of lignocellulose in corn stover based on near-infrared reflectance spectroscopy and chemometrics methods[J]. Bioresource Technology, 2021, 321: 124449.
- [21] Steinier J, Termonia Y, Deltour J. Smoothing and differentiation of data by simplified least square procedure [J]. Analytical Chemistry, 1972, 44(11): 1906-1909.
- [22] Walczak B, van den Bogaert B, Massart D L. Application of wavelet packet transform in pattern recognition of near-IR data[J]. Analytical Chemistry, 1996, 68(10): 1742-1747.