

激光与光电子学进展

基于近红外光谱的榛子蛋白质无损检测模型

张冬妍, 付聪聪*, 李丹丹, 马苗源, 黄莹

东北林业大学机电工程学院, 黑龙江 哈尔滨 150040

摘要 为了实现对榛子蛋白质的快速无损检测, 提出了一种结合近红外光谱技术和间隔随机跳蛙算法的榛子蛋白质检测模型。提取榛子的近红外光谱数据后, 对榛子光谱数据进行一阶导和标准正态变量变换预处理。针对随机跳蛙算法的初始子集以及最终波段数量阈值不确定的问题, 采用间隔随机跳蛙算法进行特征波段提取, 并对比了该算法与竞争性自适应重加权采样算法、连续投影算法和原始随机跳蛙算法的提取结果。基于提取的特征波段建立偏最小二乘回归模型。实验结果表明, 相比其他算法, 间隔随机跳蛙算法的性能最好且建立的模型稳定性更高。间隔随机跳蛙算法对交叉验证集的回归系数和均方根误差分别为 0.9082 和 0.0178, 对测试集的回归系数和均方根误差分别为 0.8999 和 0.0372。

关键词 光谱学; 近红外; 无损检测; 间隔随机跳蛙算法; 偏最小二乘法

中图分类号 O433

文献标志码 A

DOI: 10.3788/LOP213222

Nondestructive Detection Model of Hazelnut Protein Based on Near Infrared Spectroscopy

Zhang Dongyan, Fu Congcong*, Li Dandan, Ma Miaoyuan, Huang Ying

College of Mechanical and Electrical Engineering, Northeast Forestry University,
Harbin 150040, Heilongjiang, China

Abstract To achieve the quick nondestructive detection of hazelnut protein, a near infrared spectroscopy and interval random frog algorithm-based hazelnut protein detection model is proposed in this paper. After extracting the near infrared spectral data of hazelnut, first-order derivative and standard normal variable transformation preprocessing on the hazelnut spectral data is performed. Considering the uncertainty of the initial subset of the random frog algorithm and the final band number threshold's uncertainty, an interval random frog algorithm is utilized to extract the characteristic band, competitive adaptive reweighted sampling algorithm, and successive projections algorithm, and the original random frog algorithm are compared. Furthermore, a partial least squares regression model is developed based on the extracted feature bands. The experimental findings depict that the interval random frog algorithm had the best performance and the developed model is more stable when compared with other algorithms. The regression coefficient and root mean square error of the interval random frog algorithm for the cross-validation set are 0.9082 and 0.0178, respectively, and the regression coefficient and root mean square error for the test set are 0.8999 and 0.0372, respectively.

Key words spectroscopy; near infrared; nondestructive testing; interval random frog algorithm; partial least squares

1 引言

榛子是四大坚果之一,也是东北三省的主要特产之一,因其内部含有丰富的蛋白质、脂肪酸、碳水化合物、多种营养物质以及绿色环保和可加工性强的特性受到了人们的欢迎。但榛子在贮存和食品生产过程中经常出现干瘪、霉变等情况,不仅会降低蛋白质的含

量^[1],霉变的榛子还会对人体造成不良影响。蛋白质作为榛子的主要营养指标,对其含量进行检测分析有助于鉴别榛子的品质并确认其是否发生霉变,避免食品生产中出现的质量与安全问题^[2-5]。

近红外光谱技术具有无损检测以及检测精度和效率高等优点,得到了人们的广泛关注和研究^[6-9]。其原理是依据各种物质中不同官能团信息在光谱中对应不

收稿日期: 2021-11-13; 修回日期: 2021-12-20; 录用日期: 2021-12-21; 网络首发日期: 2021-01-06

基金项目: 中央高校基本科研业务费专项(2572019BF02)

通信作者: *fcc@nefu.edu.cn

同的吸收光谱,利用光谱信息进行物质含量的检测。近年来,人们关于近红外无损检测的研究取得了一定的进展。在榛子的无损检测方面:Pannico等^[10]通过结合近红外光谱检测和人工感官评价建立两个偏最小二乘法(PLS)模型,实现了对有缺陷果仁的检测和分离并根据脂质氧化水平对健康果仁进行分级;Moscetti等^[11]通过对光谱数据集的处理,测试了几种预处理方法对偏最小二乘回归(PLSR)模型的影响,并选用最佳预处理方法建立的模型对榛子仁缺陷进行检测。在蛋白质含量的检测以及特征波段的提取中,常用的特征波段提取算法有竞争性自适应重加权采样(CARS)算法^[12]、连续投影算法(SPA)^[13-14]、无信息变量消除法(UVE)^[15]以及间隔偏最小二乘法等。如:蒋大鹏等^[16]通过提取东北地区松子的近红外光谱,运用序列最小优化支持向量机(SMO-SVM)、随机次梯度下降支持向量机(Pegasos-SVM)与最小二乘支持向量机(LS-SVM)算法提取松子的特征波段后建立松子蛋白质分类相关性模型,并对相应验证集中的数据进行预测验证,然后依据近红外光谱数据实现了松子蛋白质含量等级的预测;马文强等^[17]采用间隔偏最小二乘法筛选光谱特征波段并建立偏最小二乘模型,对核桃中蛋白质的含量进行了检测,整体样本的验证集均方根误差和相关系数分别为0.021和0.913;仇逊超等^[15]将近红外光谱技术应用于东北松子蛋白质的无损检测中,采用反向间隔偏最小二乘法、无信息变量消除法筛选特征波段,该模型对带壳松子和松仁蛋白质模型校正集的相关系数分别为0.9056和0.9383,验证集的均方根误差分别为0.6670和0.5761,为带壳松子和松仁

的蛋白质无损检测提供了参考。

上述研究表明,对光谱进行特征选择可以进一步提升模型的检测精度和效率,对榛子的无损检测有一定的参考和应用价值^[18],但目前对榛子蛋白质含量检测的报道较少。因此,本文采用近红外方法提取野生榛子的光谱数据,采用多种预处理方法和改进的随机跳蛙算法提取榛子的蛋白质特征波长,并对比分析CARS算法、SPA、原始随机跳蛙算法与改进算法结合PLSR方法建立的模型估算结果,确定对榛子蛋白质无损检测的最佳波长和最优回归模型,为通过近红外无损检测榛子蛋白质含量提供了一种新方法和依据。

2 材料和方法

2.1 仪器与材料

实验使用的近红外光谱仪为Ocean optics公司制造的QUEST512,检测波长为850~1700 nm,光谱分辨率小于3 nm,使用InGaAs探测器阵列,工作电压为5 V的卤素光源。卤素光源和榛子样品通过光谱仪配套Y形光纤探头连接,另一端连接光谱仪,光谱仪通过USB数据线与计算机连接,如图1所示。光谱数据的测量在SpectraSuite软件下进行,实验中其他数据和图片均通过PyCharm Community Edition 2020.3.2 x64以及Matlab 2019b软件获得。实验使用的榛子为同一原产地的两批东北小兴安岭野生毛榛子,人工去壳后选出完整且未变质榛子共300粒,平均每粒榛子的质量为0.5 g,采用Kolmogorov-Smirnov(KS)算法按照8:2的比例划分为交叉验证集和测试集。

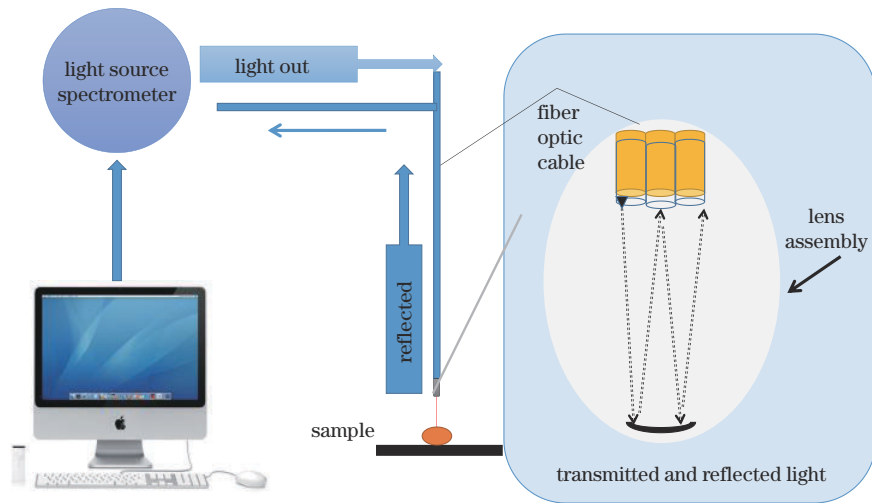


图1 硬件平台的原理图

Fig. 1 Schematic diagram of the hardware platform

2.2 数据的采集

保持室内温度20℃左右采集光谱数据,采集前需对仪器进行校准。先测量环境的光谱(暗光谱),再测量白板光谱(明光谱),将明暗光谱相减后测量榛子光

谱数据,以保证采集数据不受仪器自身和环境的影响。采集过程中,以漫反射方式测量单个榛子的近红外光谱数据10次,积分时间为50 ms,光源对准最大榛子直径的中心,测试过程中保持垂直照射,并完全阻挡光线

直接进入检测器。测量后切换不同点位,然后取三个点位的平均值作为最终光谱数据,以减少环境差异对测量结果的影响。榛子蛋白质的化学测定方法依据国家食品安全标准 GB5009.5—2016,榛子蛋白质的分布如表 1 所示。

表 1 榛子中蛋白质的含量分布
Table 1 Content distribution of protein in hazelnut

Data set	Number	Range / %	Mean / %	Standard deviation / %
Hazelnut sample set	300	20.10–26.93	23.50	1.58

2.3 预处理方法

采集近红外光谱过程中虽然避免了多种误差来源,但由于不同时间的光照强度不同以及测量角度的

误差,仍然会产生光谱散射和基线漂移,导致数据建模时的误差较大。外部环境和人工测量引起的误差可通过预处理降低,因此,采用一阶导和标准正态变量(SNV)进行预处理,以消除样本的基线偏移、散射、样本尺寸、光源以及光路导致的误差^[19-20]。

2.4 数据建模方法

实验中主要采用四种算法进行建模,对光谱数据进行预处理后,先用随机跳蛙(RF)算法提取特征波段,然后改进随机跳蛙算法的初始变量集(V_0),选取最低均方根误差(R_{MSE})对应的波段建立间隔随机跳蛙(IRF)算法并提取特征波段,再用CARS算法、SPA提取特征波段,利用四组特征波段建立PLSR模型并对比CARS、SPA、IRF三个算法建立的PLSR模型性能指标,以体现IRF在特征波段提取中的优势以及建模的精确性和稳定性。具体步骤如图 2 所示。

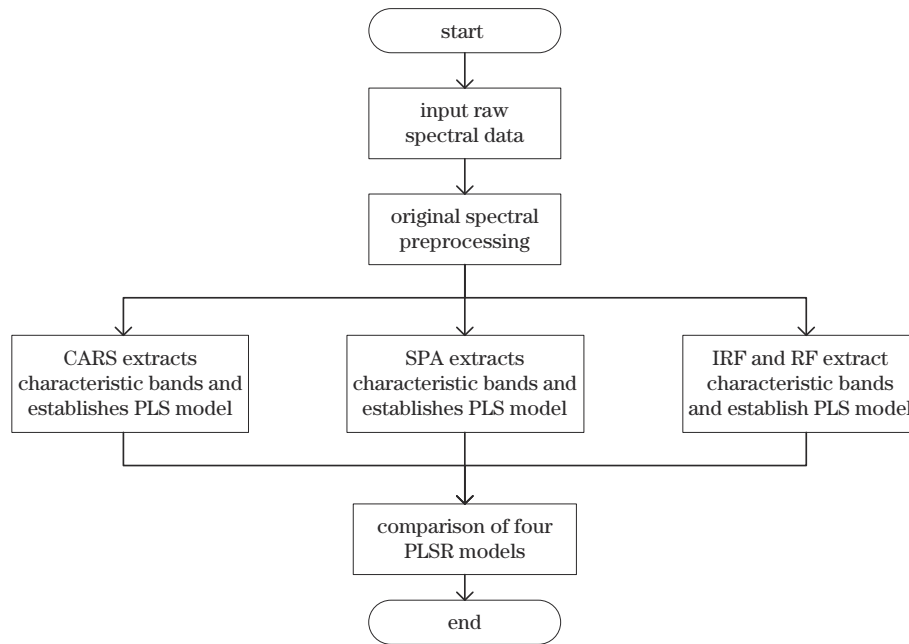


图 2 实验流程图

Fig. 2 Experimental flow chart

提取全波段后,对榛子波段采用一阶导和SNV变换进行预处理,然后提取预处理后光谱数据的特征波段。

1) CARS是通过自适应重加权采样(ARS)技术选择PLS模型中回归系数绝对值较大的特征波长,去掉权重较小的特征波长,利用交叉验证选出均方根误差最小的波长子集,获得最优波长的组合。

2) SPA是一种常用的光谱提取算法,可消除光谱变量之间的相关性,以最小的冗余和相关性获得特征波长。SPA利用矢量的投影分析将被选择波长投影到其他波长并对比投影矢量的大小,然后取投影矢量最大的波长作为最终的特征波长。

3) RF算法先设定包含 Q 个变量的原始变量子集 V_0 。然后基于原始变量子集,建立包含 Q^* 个变量的候

补变量子集 V^* ,再用 V^* 作为 V_1 代替 V_0 ,直到迭代 N 次。最后计算每个变量的选择概率作为变量选择的标准,通过人为设定阈值选择特征波段。IRF算法是基于RF算法的改进算法。RF算法虽然在特征波长选择中表现突出,但仍存在不足,如:初始变量集 V_0 是随机产生的(有效性难以保证),为保证运行过程遍历整个数据集,迭代次数 N 需要足够大,导致算法的运行时间长、收敛速度慢,效率低;RF算法在选择特征波长时,选择被选概率值大于阈值的变量为特征波长,但人为设定的阈值无理论依据。因此,IRF算法提出了两点改进:基于间隔区间建立PLS模型,将最小均方根误差对应区间的波长设为初始变量集 V_0 并开始迭代,有效消除了 V_0 的随机性,避免无信息变量及噪声的干扰,减少了迭代次数;对排序后的变量从第1个波长开

始,每次增加1个波长,建立光谱数据和待测物质含量数据之间的多元线性回归模型。计算每个模型的交叉验证均方根误差 R_{MSE}^c ,最小 R_{MSE}^c 对应的变量子集即为特征波长。

对 CARS、SPA、RF、IRF 算法提取的特征波段分别建立 PLS 模型进行回归预测,然后选取交叉验证集和预测集的回归系数 (R_c 和 R_p) 以及均方根误差 (R_{MSE}^c 和 R_{MSE}^p) 作为指标评价四个模型的性能。

3 分析与讨论

3.1 光谱测量和预处理结果

用漫反射方法采集选取的 300 粒去壳榛子的光谱信息,测量的原始光谱数据如图 3 所示。采用 KS 算法将样品按 8:2 的比例分为交叉验证集和测试集。KS 算法在空间距离均匀分布的前提下保证了数据集中的样本分布均匀。对原始光谱数据按照预处理方法中的一阶导和 SNV 进行预处理,预处理后的波形如图 4 所示。其中:一阶导可以消除基线偏移,增强光谱特性,减少背景光源基线偏移引起的误差;SNV 可以消除样品和光源变化对光谱数据的影响。对预处理前后波段建立的 PLS 模型进行预测,结果如表 2 所示。对比发现,预处理后的模型性能得到了大幅提升,显著提高了模型的预测结果。

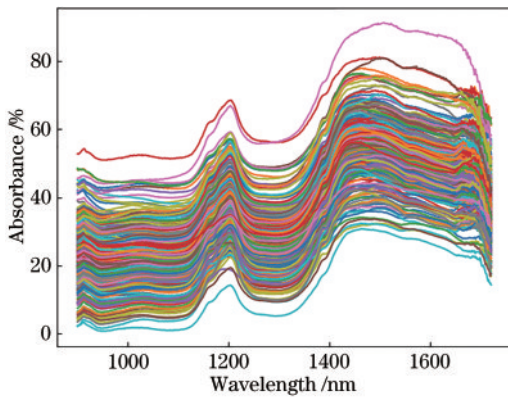


图 3 全波段的原始光谱

Fig. 3 Raw spectra of all bands

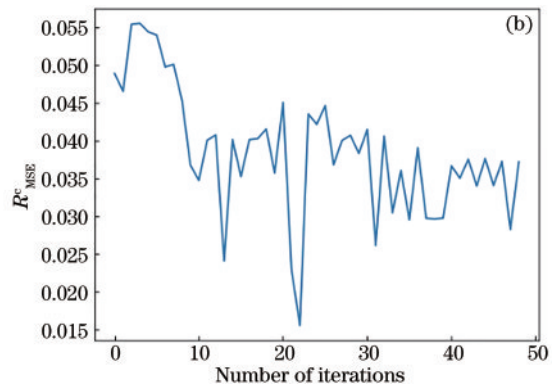
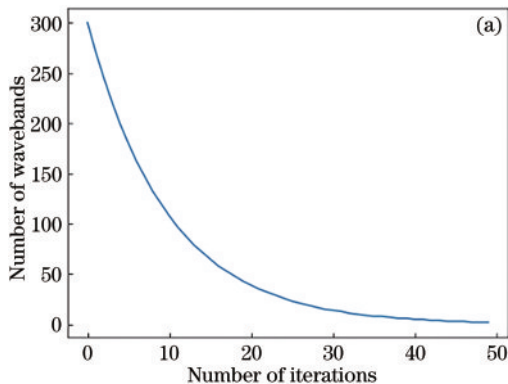


图 5 CARS 算法的最优建模结果。(a) 波段数; (b) R_{MSE}^c

Fig. 5 Optimal modeling result of the CARS algorithm. (a) Number of wavebands; (b) R_{MSE}^c

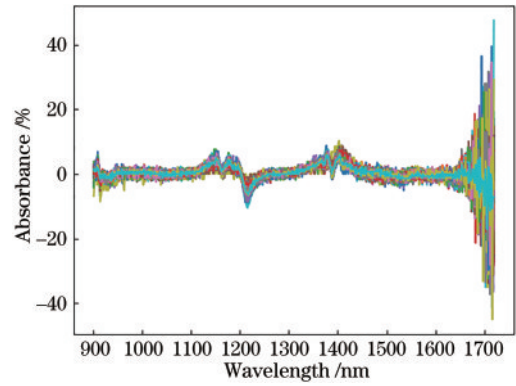


图 4 预处理后的光谱

Fig. 4 Spectra after preprocessing

表 2 预处理前后模型的预测结果
Table 2 Prediction results of the model before and after preprocessing

Method	Cross validation set		Test set	
	R_{MSE}^c	R_c	R_{MSE}^p	R_p
PLS model before pretreatment	0.0792	0.5183	0.0484	0.6089
PLS model after pretreatment	0.0550	0.7590	0.0365	0.6374

3.2 建模结果

3.2.1 CARS 算法的建模结果

CARS 算法的原理依据数据建模方法,在提取特征波段过程中将回归测试波段数量设置为 20,交叉验证集数量设置为 10。CARS 采用蒙特卡罗抽样方法选择校准集,建立 PLSR 模型,最终找到最优变量组合^[12]。CARS 每次选择的波段数量是随机的,导致实验中相应的结果也是随机的。因此,实验中 CARS 建模运行了 5 次,选择 R_{MSE}^c 最小的特征波段作为最终提取的特征波段。图 5 为 50 次采样后基于 CARS 算法的特征波段选择结果。可以发现:随着采样时间的增加,特征波段的数量逐渐减少, R_{MSE}^c 先减小后增大;前几次迭代中,特征波段数量和 R_{MSE}^c 均急剧下降,这表明粗糙选择去除了大量不相关波段; R_{MSE}^c 达到最小值之

后,随着采样次数的增加又逐渐增加,这表明与预测结果相关的一些特征波段从光谱数据中被移除,导致预测模型的效果变差。因此,将第 23 次采样后得到的波段确定为预测榛子蛋白质含量的特征波段,共 38 个变量。

3.2.2 SPA 的建模结果

SPA 是一种正向特征变量选择方法。算法中测试比例是评价算法提取特征波段是否有效的一个参数,提取过程中的测试比例选为 0.3。图 6 为 SPA-PLS 模型包含不同数量波段的 R_{MSE} 。图 7 为提取特征波段的位置。其中,横坐标表示波段序列号,纵坐标表示经 SNV 变换后的输入值。从图 6 可以发现,提取特征波段的最佳数量为 3 个, R_{MSE} 为 0.05259。从图 7 可以发现,提取的三个波段分别为第 96 个、第 254 个和第 384 个,对应波长分别为 1056.52 nm、1312.58 nm 和 1520.64 nm。

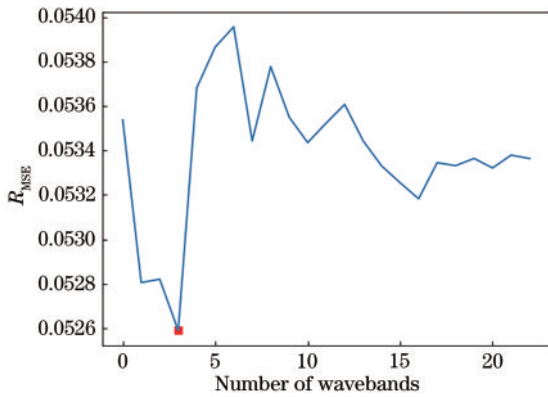


图 6 SPA 的 R_{MSE} 随波段数的变化曲线

Fig. 6 Variation curve of R_{MSE} of SPA with the number of wavebands

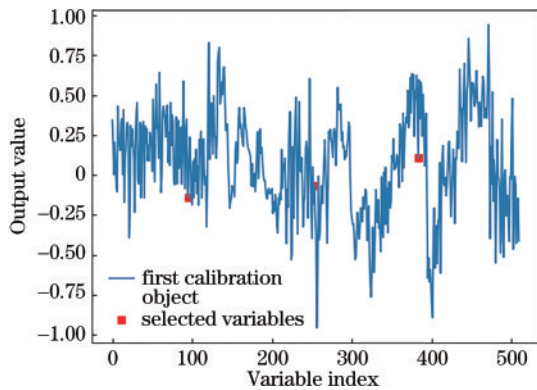


图 7 SPA 选取波段的位置

Fig. 7 Location of selected bands for SPA

3.2.3 RF 算法改进前后的建模结果

RF 算法中的迭代次数为 10000 次,变量子集的数量 Q 为 2,交叉验证波段最大数量 A 为 10,特征波段阈值为 30。IRF 算法提取特征波段过程中的迭代次数为 2000,固定窗口数量为 1,子间隔初始化数为 10,最大主分量为 10,仿真分析结果如图 8、图 9 所示。可以发

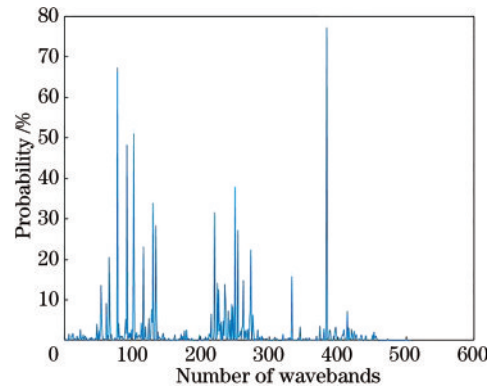


图 8 IRF 算法各波段的选取概率

Fig. 8 Selection probability of each band of the IRF algorithm

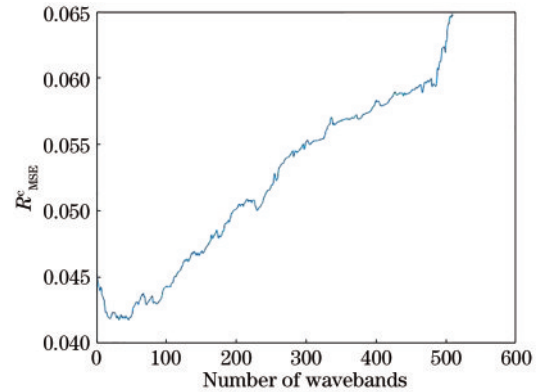


图 9 IRF 算法的 R_{MSE}^c 随特征波段数量的变化曲线

Fig. 9 Variation curve of R_{MSE}^c of the IRF algorithm with the number of characteristic bands

现,IRF 模型的 R_{MSE}^c 随着提取波段数量的增加先减少后增加,在波段数量达到 32 时 R_{MSE}^c 取得最小值,为 0.04175。因此,选取特征波段数量为 32 个,对应的波长范围为 970~1100 nm 以及 1360 nm 和 1510 nm 附近。

根据四种算法提取出的特征波段分别建立 PLSR 模型,结果如表 3 所示。可以发现,四种算法提取特征波段后建立的 PLSR 模型相比预处理后建立的模型在各个性能指标上都得到了大幅提升,且减少了无关波段带来的运算量,提升了模型的预测效率。此外,SPA、CARS、RF 算法的表现明显不如 IRF 算法,主要原因:CARS 算法在特征波段筛选过程中波段数量减

表 3 不同算法的建模结果

Table 3 Modeling results of different algorithms

Algorithm	Cross validation set		Test set	
	R_{MSE}^c	R_c	R_{MSE}^p	R_p
Pretreatment	0.0550	0.7590	0.0365	0.6374
CARS	0.0396	0.8972	0.0181	0.8840
SPA	0.0315	0.8953	0.0370	0.7963
RF	0.0186	0.8922	0.0403	0.8876
IRF	0.0178	0.9082	0.0372	0.8999

少太快,不仅没有完全去除无关波段,还将部分特征波段定义为无关波段去除;SPA 仅提取了三个特征波段,在预测过程中模型的稳定性明显低于 CARS 与 IRF 算法建立的回归模型,容易产生过拟合现象,但模型效率最高;RF 算法的特征波段由人为设定阈值,且在提取特征波段时的迭代次数较大,导致运算效率较低;IRF 算法选择了 R_{MSE}^s 最小时对应的 32 个波段,即特征波段被选择概率最高的前 32 个波段,相比其他三种算法建立的模型更稳定。此外,仇逊超等^[15]和马文强等^[17]建立的蛋白质含量预测模型提取的特征波段数量分别为 37 和 88,相比而言,IRF 算法在预测精度相当的情况下提取的波段数量更少,极大减少了模型的复杂度。

图 10、图 11 为基于 IRF 算法提取的特征波段建立 PLSR 模型的回归拟合曲线。在特征波长的选取中,IRF 选取出现概率较大的特征波长有 1520.64、1027.18、1366.14 nm。蛋白质是由 20 种不同 α -氨基酸和亚氨酸用不同肽键连接成的多肽聚合体,包含 C—H、N—H 化学键等^[21-22]。其中:N—H 键伸缩振动的一级倍频吸收带位于 1510 nm 附近;1508.95 nm 为 α -氨基酸中的氨基的吸收峰;蛋白质中 N—H 键基频

的二倍频为近红外区域 1034 nm 附近的吸收峰;甲基和亚甲基中 C—H 键伸缩振动的一级倍频和变形振动的基频吸收带位于 1360~1395 nm 范围内;1385.07 nm 位置为 C—H 键的吸收峰。在一些非蛋白质特征波段所在区域中也存在选择概率较大的波段,原因是榛子中的脂肪与脂肪酸包含相同的化学键,在后续研究中可用于对榛子脂肪以及脂肪酸等指标的修正^[23]。

4 结 论

构建了检测东北小兴安岭野生榛子蛋白质含量的回归模型。采用一阶导和 SNV 预处理后,有效消除了光谱中的基线偏移以及光源变化对光谱数据造成的影响。采用 IRF 算法对光谱波段进行优化,相比常见的 CARS 算法和 SPA,在预测精度和效率上均得到了一定提升,有效降低了波段数量以及模型的复杂度。实验结果表明,间隔随机跳蛙算法对交叉验证集的回归系数和均方根误差分别为 0.9082 和 0.0178,对测试集的回归系数和均方根误差分别为 0.8999 和 0.0372,能实现对东北小兴安岭野生榛子中蛋白质含量的快速无损检测。该研究表明波段优化在建模分析过程中的重要性,只有采取合适的特征波长提取算法才能更好地提升模型性能,为榛子蛋白质的快速无损检测提供了思路和依据。

参 考 文 献

- [1] 王进, 蒋柳庆, 马惠玲, 等. ClO_2 和 1-MCP 对青皮核桃二步贮藏的效应[J]. 中国食品学报, 2015, 15(3): 137-145.
Wang J, Jiang L Q, Ma H L, et al. Effect of ClO_2 and 1-MCP on preservation of green fruit and peeled walnut in two-step storage[J]. Journal of Chinese Institute of Food Science and Technology, 2015, 15(3): 137-145.
- [2] Ji J M, Ge Z F, Feng Y S, et al. Lipid characterization of Chinese wild hazelnuts (*Corylus mandshurica* maxim.) [J]. Journal of Oleo Science, 2019, 68(1): 13-20.
- [3] Fan L Y, Ren J, Yang Y T, et al. Comparative analysis on essential nutrient compositions of 23 wild hazelnuts (*Corylus heterophylla*) grown in northeast China[J]. Journal of Food Quality, 2020, 2020: 9475961.
- [4] 畅博奇, 季兰, 王帆, 等. 平欧杂种榛不同品种(系)坚果品质的比较研究[J]. 中国果树, 2021(11): 31-38.
Chang B Q, Ji L, Wang F, et al. Comparison of nut quality of different cultivars (lines) of interspecific hybrid F1 between *Corylus heterophylla* Fisch and *Corylus avellana* L[J]. China Fruits, 2021(11): 31-38.
- [5] 唐永康, 俞建峰, 夏晓露. 球磨联合超声波提取榛子蛋白的研究[J]. 现代化工, 2021, 41(9): 149-154.
Tang Y K, Yu J F, Xia X L. Study on extraction of hazelnut protein by ball milling and ultrasound[J]. Modern Chemical Industry, 2021, 41(9): 149-154.
- [6] Zareef M, Chen Q S, Hassan M M, et al. An overview

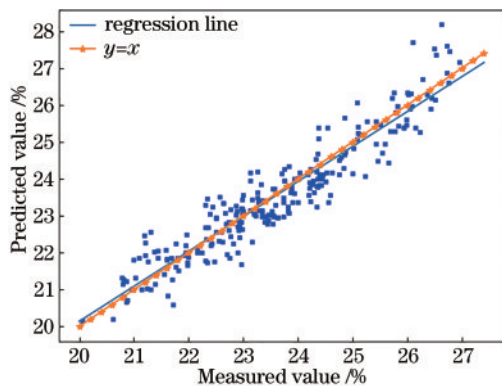


图 10 IRF 算法对交叉验证集的回归结果

Fig. 10 Regression results of the IRF algorithm on the cross-validation set

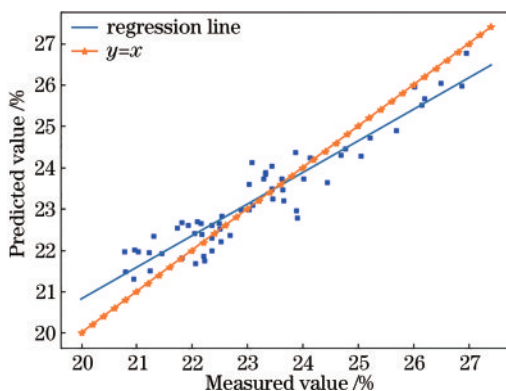


图 11 IRF 算法对测试集的回归结果

Fig. 11 Regression results of the IRF algorithm on the test set

- on the applications of typical non-linear algorithms coupled with NIR spectroscopy in food analysis[J]. *Food Engineering Reviews*, 2020, 12(2): 173-190.
- [7] Momchilova S, Nikolova-Damyanova B. Quantitative TLC and gas chromatography determination of the lipid composition of raw and microwaved roasted walnuts, hazelnuts, and almonds[J]. *Journal of Liquid Chromatography & Related Technologies*, 2007, 30(15): 2267-2285.
- [8] 韦卓, 李稳稳, 林敏, 等. 基于 Dropout 深度信念网络的棉和涤纶含量的近红外光谱检测[J]. *光学学报*, 2021, 41(17): 1730002.
Wei Z, Li W W, Lin M, et al. Near-infrared spectroscopy detection of cotton/polyester content based on Dropout deep belief network[J]. *Acta Optica Sinica*, 2021, 41(17): 1730002.
- [9] 王绪泉, 王丽丽, 方家熊. 近红外光谱传感物联网研究与应用进展[J]. *中国激光*, 2021, 48(12): 1210001.
Wang X Q, Wang L L, Fang J X. Research and application progresses of near-infrared spectral sensing Internet of Things[J]. *Chinese Journal of Lasers*, 2021, 48(12): 1210001.
- [10] Pannico A, Schouten R E, Basile B, et al. Non-destructive detection of flawed hazelnut kernels and lipid oxidation assessment using NIR spectroscopy[J]. *Journal of Food Engineering*, 2015, 160: 42-48.
- [11] Moschetti R, Haff R P, Aernouts B, et al. Feasibility of Vis/NIR spectroscopy for detection of flaws in hazelnut kernels[J]. *Journal of Food Engineering*, 2013, 118(1): 1-7.
- [12] Yu H D, Qing L W, Yan D T, et al. Hyperspectral imaging in combination with data fusion for rapid evaluation of tilapia fillet freshness[J]. *Food Chemistry*, 2021, 348: 129129.
- [13] Bai Y H, Xiong Y J, Huang J C, et al. Accurate prediction of soluble solid content of apples from multiple geographical regions by combining deep learning with spectral fingerprint features[J]. *Postharvest Biology and Technology*, 2019, 156: 110943.
- [14] 陈远哲, 王巧华, 高升, 等. 基于近红外光谱的淡水鱼贮藏期质构品质的无损检测模型[J]. *激光与光电子学进展*, 2021, 58(12): 1230001.
Chen Y Z, Wang Q H, Gao S, et al. Nondestructive testing model for textural quality of freshwater fish in storage using near-infrared spectroscopy[J]. *Laser & Optoelectronics Progress*, 2021, 58(12): 1230001.
- [15] 仇逊超, 曹军. 近红外光谱波段优化在东北松子蛋白质定量检测中的应用[J]. *现代食品科技*, 2016, 32(11): 303-309.
Qiu X C, Cao J. Application of band optimization of near-infrared spectra for quantitative detection of proteins in northeastern pine nuts[J]. *Modern Food Science and Technology*, 2016, 32(11): 303-309.
- [16] 蒋大鹏, 张冬妍, 李丹丹, 等. 基于近红外的松子蛋白质品质分类处理[J]. *计算技术与自动化*, 2018, 37(3): 180-184.
Jiang D P, Zhang D Y, Li D D, et al. Classification of pine nut protein quality based on near infrared[J]. *Computing Technology and Automation*, 2018, 37(3): 180-184.
- [17] 马文强, 张漫, 李忠新, 等. 基于近红外光谱的核桃仁蛋白质含量检测分析[J]. *农业机械学报*, 2017, 48(S1): 407-411.
Ma W Q, Zhang M, Li Z X, et al. Detection and analysis of walnut protein content based on near infrared spectroscopy[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2017, 48(S1): 407-411.
- [18] 韩亚鲁, 李绍稳, 郑文瑞, 等. 基于集成提升算法的土壤速效氮近红外光谱回归预测[J]. *激光与光电子学进展*, 2021, 58(16): 1630005.
Han Y L, Li S W, Zheng W R, et al. Regression prediction of soil available nitrogen near-infrared spectroscopy based on boosting algorithm[J]. *Laser & Optoelectronics Progress*, 2021, 58(16): 1630005.
- [19] Wang J H, Wang J, Chen Z, et al. Development of multi-cultivar models for predicting the soluble solid content and firmness of European pear (*Pyrus communis* L.) using portable Vis-NIR spectroscopy[J]. *Postharvest Biology and Technology*, 2017, 129: 143-151.
- [20] Ejeahalaka K K, On S L W. Effective detection and quantification of chemical adulterants in model fat-filled milk powders using NIRS and hierarchical modelling strategies[J]. *Food Chemistry*, 2020, 309: 125785.
- [21] 周瑞宝. 植物蛋白功能原理与工艺[M]. 北京: 化学工业出版社, 2008.
Zhou R B. Functional principle and technology of plant protein[M]. Beijing: Chemical Industry Press, 2008.
- [22] 严衍禄. 近红外光谱分析基础与应用[M]. 北京: 中国轻工业出版社, 2005.
Yan Y L. Basis and application of near infrared spectroscopy[M]. Beijing: China Light Industry Press, 2005.
- [23] 黄维, 田丰玲, 刘振尧, 等. 基于不同 PLS 算法的方竹笋中蛋白质分析的近红外光谱特征波段选择[J]. *食品科学*, 2013, 34(22): 133-137.
Huang W, Tian F L, Liu Z Y, et al. Wavelength selection for FT-NIR spectroscopic analysis of protein in chimonobambusa quadrangularis shoot based on iPLS and BiPLS models[J]. *Food Science*, 2013, 34(22): 133-137.