

京郊鲜食杏白利糖度的便携式光谱快速无损检测方法研究

刘翠玲^{1,2}, 闻世震^{1,2*}, 孙晓荣^{1,2}, 张善哲^{1,2}, 姜传智^{1,2}, 殷莺倩^{1,2}

(1. 北京工商大学人工智能学院, 北京 100048; 2. 北京工商大学食品安全大数据技术北京市重点实验室, 北京 100048)

摘要: 目的 建立京郊鲜食杏白利糖度的定量分析预测模型, 实现对京郊鲜食杏品质的快速无损检测。**方法** 使用便携式近红外光谱仪采集 900~1700 nm 下鲜食杏的漫反射光谱信息, 使用多元散射校正(multiplicative scatter correction, MSC)、标准正态变量变换(standard normal variable transformation, SNV)和 Savitzky-Golay 卷积平滑(Savitzky-Golay smooth, S-G)对原始光谱数据进行预处理, 使用 Kennard-Stone (K-S) 算法以 3:1 比例将样本集划分成校正集和预测集, 利用竞争自适应重加权采样(competitive adaptive reweighted sampling, CARS)算法和连续投影算法(successive projections algorithm, SPA)对光谱进行特征波长筛选, 使用偏最小二乘回归(partial least squares regression, PLSR)算法建立京郊鲜食杏白利糖度的预测模型。**结果** 以 MSC+S-G+CARS+PLSR 算法建立的北京鲜食杏的白利糖度预测模型取得较好的预测精度, 模型的校正集均方根误差、校正集相关系数、预测集均方根误差、预测集相关系数分别为 0.3502、0.9747、0.4698、0.9616。**结论** 基于便携式近红外光谱技术所采集数据构建的京郊鲜食杏白利糖度预测模型准确性较高, 可以快速准确检测鲜食杏白利糖度, 从而实现对鲜食杏品质的快速无损检测, 为鲜食杏的品质检测提供了理论依据和方法指导。

关键词: 鲜食杏; 白利糖度; 便携式; 无损检测; 偏最小二乘回归

Research on the Brix content of fresh *Armeniaca* in suburbs of Beijing based on portable spectroscopic rapid non-destructive detection method

LIU Cui-Ling^{1,2}, WEN Shi-Zhen^{1,2*}, SUN Xiao-Rong^{1,2}, ZHANG Shan-Zhe^{1,2},
JIANG Chuan-Zhi^{1,2}, YIN Ying-Qian^{1,2}

(1. School of Artificial Intelligence, Beijing Technology and Business University, Beijing 100048, China; 2. Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, Beijing 100048, China)

ABSTRACT: Objective To establish a quantitative analysis and prediction model for the Brix content of fresh *Armeniaca* in the suburbs of Beijing, realize rapid non-destructive testing of the quality of fresh *Armeniaca* in the suburbs of Beijing. **Methods** Diffuse reflectance spectral information of fresh *Armeniaca* in the suburbs of Beijing was collected by a portable near-infrared spectrometer. The raw spectral data were preprocessed using multiplicative

基金项目: 北京市自然科学基金项目(4222043)

Fund: Supported by the Beijing Natural Science Foundation of China (4222043)

*通信作者: 闻世震, 硕士研究生, 主要研究方向为光谱检测技术与系统研制。E-mail: wsz412801@163.com

*Corresponding author: WEN Shi-Zhen, Master Degree, School of Artificial Intelligence, Beijing Technology and Business University, No.11, Fucheng Road, Haidian District, Beijing 100048, China. E-mail: wsz412801@163.com

scatter correction (MSC), standard normal variable transformation (SNV), and Savitzky-Golay smooth (S-G). The sample set was divided into calibration set and prediction set according to the ratio of 3:1 using Kennard-Stone algorithm. The characteristic wavelengths of the spectrum were selected by the competitive adaptive reweighted sampling (CARS) algorithm and the successive projections algorithm (SPA). A prediction model of Beijing fresh *Armeniaca* Brix was established using the partial least squares regression (PLSR) algorithm. **Results** The prediction model for the Brix content of fresh *Armeniaca* in Beijing suburbs established by the MSC+S-G+CARS+PLSR algorithm had better prediction accuracy, and the root mean square of calibration, correlation coefficient of calibration, root mean square of prediction, and correlation coefficient of prediction of the model was respectively 0.3502, 0.9747, 0.4698, and 0.9616. **Conclusion** The prediction model of the Brix content of fresh *Armeniaca* in suburban Beijing constructed based on the data of the portable spectrometer has high accuracy, which can quickly and accurately detect the Brix content of fresh *Armeniaca*, and can realize rapid and non-destructive testing of the quality of fresh *Armeniaca*. The theoretical basis and method guidance are provided for the quality detection of fresh *Armeniaca*.

KEY WORDS: fresh *Armeniaca*; Brix; portable; non-destructive detection; partial least squares regression

0 引言

杏为蔷薇科(Rosaceae)杏属(*Armeniaca*)植物,原产于我国,资源丰富,地理分布广泛^[1]。北京位于华北平原西北部,鲜食杏(*Armeniaca vulgaris* Lam.)产业发展历史悠久,根据 2018 年的北京市杏和李产业调查报告,北京市鲜食杏的种植面积 9.63 万亩,年产 2720.69 万公斤^[2]。目前北京的鲜食杏存在成熟期集中、采摘周期短和品质一般等问题。果实品质在很大程度上取决于果实内部所含糖的种类和数量^[3]。白利糖度代表在 20℃下每 100 g 水溶液中溶解的蔗糖克数,在水果产业中代表果实内部的糖分含量^[4]。因此,检测鲜食杏的白利糖度对于鲜食杏的品质至关重要,对精准把控鲜食杏采摘时间和选育高品质鲜食杏具有重要意义。

目前已有学者对杏子果实的品质进行了相关研究。HUANG 等^[5]分析了 11 个杏品种果实在成熟过程中糖含量与花青素的关系。NSIBI 等^[6]采用基因组选择方法,使用贝叶斯 LASSO 模型和基因组选择模型 RR-BLUP 预测杏的葡萄糖含量和乙烯释放量,RR-BLUP 模型预测效果好于贝叶斯 LASSO 模型,但模型的精度不高。LIU 等^[7]比较了树上干杏在多种低温条件下糖、酸和香气挥发物的变化,发现杏子的风味主要与香气挥发物有关,近冰点储存方式能够有效保持杏的风味。上述研究基于化学计量学方法,需要对杏果实进行破损,再与化学试剂反应,返回结果时间较长,难以实现对杏子品质的快速准确检测。

随着光谱检测技术的发展,便携式光谱仪因其体积小、快速、无损检测等优点,逐渐在食品快速检测中普及和应用^[8-11]。徐金阳等^[12]采用 AMBERII 手持光谱仪采集小麦的近红外光谱,基于遗传算法和偏最小二乘法建立小麦粉的水分、灰分和面筋预测模型,实现对小麦粉品质的快速检测。李庆等^[13]利用便携式近红外技术对名贵药材西

红花和常见假冒品进行快速真伪鉴别和掺假定量预测。袁雷明等^[14]基于可见/近红外光谱技术自主研发便携式光谱仪,对红富士苹果的可溶性固形物含量具有较高预测精度。孙阳等^[15]使用便携式近红外光谱仪实现对樱桃番茄糖分的现场快速无损检测。光谱仪利用其内部的光源照射到水果表面时,部分光透过水果表皮,在水果内部发生漫反射,通过分析漫反射光谱与水果内部理化指标的关系,可实现对水果内部品质的无损检测^[16]。基于以上研究,便携式光谱技术在农业和食品检测领域具有广阔的应用前景和研究潜力。

目前鲜食杏的品质检测依然采用人工判断和随机采样破坏法,这种判别方法误差较大且效率低。因此,本研究以北京郊区的鲜食杏作为研究对象,利用便携式光谱技术采集样本光谱数据,分析鲜食杏的光谱特征和白利糖度变化,然后对鲜食杏的原始光谱进行预处理和特征波长筛选,采用偏最小二乘回归(partial least squares regression, PLSR)算法建立京郊鲜食杏白利糖度的最优预测模型,从而实现对京郊鲜食杏品质的快速无损检测,为鲜食杏的品质检测提供理论依据和技术支撑。

1 材料与方法

1.1 材料与试剂

在 5 月份鲜食杏成熟期,分别在北京海淀区、平谷区、延庆区的杏子采摘园,采集个体大小均匀、表面无破损、软硬适中的鲜食杏共计 109 颗作为实验样本,其中海淀玉巴达杏(*Armeniaca* 'Yubada') 35 颗,平谷北寨红杏(*Armeniaca* 'Beizhaihongxing') 37 颗,延庆半香白杏(*Armeniaca* 'Banxiangbaixing') 37 颗。将采集的鲜食杏样本流水洗净表面浮沉,按照种类分别在载物托盘(30 cm×20 cm)上间隔 5 cm 依次放置样本,然后在实验室干燥阴凉处自然晾

干, 同时为排除温度和湿度对实验结果的干扰, 保持实验室恒温恒湿。

1.2 仪器与设备

NIR-M-R2 便携式近红外光谱仪(深圳市谱研互联科技有限公司); PAL-1 数显折射仪(日本 ATAGO 公司)。

1.3 方法

1.3.1 样本光谱数据采集

实验在恒温恒湿的环境条件下进行, 室温 26℃, 湿度 36.25%, 使用便携式近红外光谱仪采集鲜食杏的光谱信息。光谱仪参数: 狭缝尺寸 25 μm , 光谱分辨率 10 nm, 波长范围 900~1700 nm, 平滑宽度 2.34 nm, 曝光时间 0.635 ms, 重复次数 6 次。光谱仪在进行扫描前应使用白板校正, 校正后开始采集样本的光谱数据。采集样品的光谱数据时, 光谱仪镜头紧贴果实表面, 在果实横径最大且果面较为平整处进行光谱采集。每个果实样本采集 5 次, 取平均值作为样本最终光谱数据, 最终得到 109 组吸光度数据。

1.3.2 白利糖度的测定

使用数显折射仪测定 109 个样本的白利糖度值。对每个样本, 使用手动榨汁器获取果肉浆液, 使用一次性滴管将果肉浆液滴入数显测糖仪, 测定样本白利糖度, 重复测定 5 次, 取平均值作为样本白利糖度的真实值。

1.3.3 光谱预处理

采集的光谱波段首尾会存在较大噪声, 降低数据的准确性, 故截取 950~1650 nm 光谱波段。为了消除光谱数据中包含的大量噪声和干扰信息, 使用多元散射校正(multiplicative scatter correction, MSC)算法^[17]、标准正态变量变换(standard normal variable transformation, SNV)^[18]和 Savitzky-Golay 卷积平滑(Savitzky-Golay smooth, S-G)^[19]对光谱数据进行预处理。

1.3.4 样本集划分

样本集中会存在差异性较小的样本, 包含这些样本建立校正模型会降低模型的准确性和稳定性。使用 Kennard-Stone (K-S)算法^[20]对样本集进行划分, 利用空间变量之间的相对欧氏距离, 选择距离大的样本为校正集, 剩余样本作为预测集。

1.3.5 光谱的特征波长筛选

在建立近红外分析模型时, 光谱仪器噪声或外界环境因素变化使某些波段下的样本光谱信噪比较低, 影响模型稳定性^[21]。利用特征波长筛选选择光谱数据的最优波长进行建模, 可提高模型的预测能力和稳定性。使用竞争性自适应重加权采样法(competitive adaptive reweighted sampling, CARS)^[22]和连续投影算法(successive projections algorithm, SPA)^[23]对光谱数据进行特征波长筛选。

1.4 模型建立及评价指标

PLSR 算法通过最小化偏差平方和对曲线进行线性拟合,

结合了多元线性回归、相关分析和主成分分析的优点^[24]。模型常见的评价指标有校正集相关系数(correlation coefficient of calibration, R_c)、校正集均方根误差(root mean square of calibration, RMSEC)、预测集相关系数(correlation coefficient of prediction, R_p)和预测集均方根误差(root mean square of prediction, RMSEP)。相关系数越接近 1, 均方根误差越接近 0, 模型的准确性和稳定性越好。

2 结果与分析

2.1 鲜食杏的光谱特征和白利糖度变化分析

图 1 显示了 109 个鲜食杏实验样本的光谱。在 980、1180 和 1450 nm 处有较为明显的吸收峰。其中, 980 nm 吸收峰对应 O—H 键的 3 倍频吸收峰, 1180 nm 对应 C—H 键的 3 倍频吸收带, 1450 nm 的吸收峰对应 H₂O 的 1 倍频吸收带。

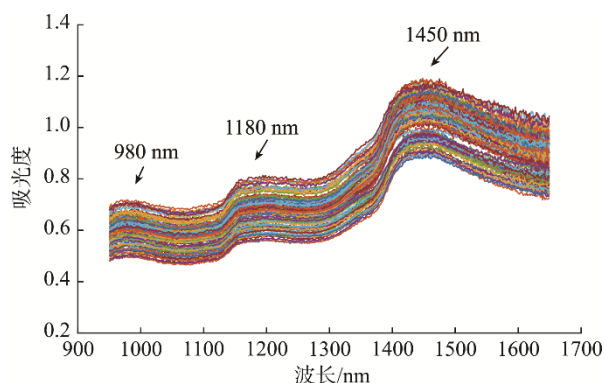


图 1 北京鲜食杏的漫反射光谱

Fig.1 Diffuse reflectance spectra of fresh *Armeniaca* in Beijing

未成熟的杏中含有大量的淀粉、有机酸和鞣质^[25], 在成熟过程中, 淀粉等糖类物质分解转化为可溶性碳水化合物, 有机酸和鞣质被氧化分解, 同时细胞壁中层的果胶质转变为可溶性果胶, 果肉细胞中的淀粉粒转化为可溶性糖^[26], 提高了杏的白利糖度。在味觉上表现为酸度下降、涩味消失、甜度上升。

2.2 光谱数据预处理及样品集划分

由于采集光谱数据时环境、样本个体差异、表面光度及表皮厚度等^[27]噪声因素影响, 致使原始光谱重叠且不平滑, 有效的光谱信息掺杂较多的白噪声, 从而影响预测模型的准确性。为了保证模型的准确性, 本研究采用 MSC、SNV 和 S-G 算法对光谱数据进行预处理。从图 2 可以看出, MSC 和 SNV 对原始光谱预处理后, 虽然大幅减少了光谱重叠现象, 但光谱曲线不平滑的问题依然存在, 而 MSC+S-G 和 SNV+S-G 预处理后, 解决了光谱曲线不平滑的问题。由于无法直接通过预处理后的光谱表征来判断预处理算法的优劣, 需要对光谱数据建立相应模型来评价。

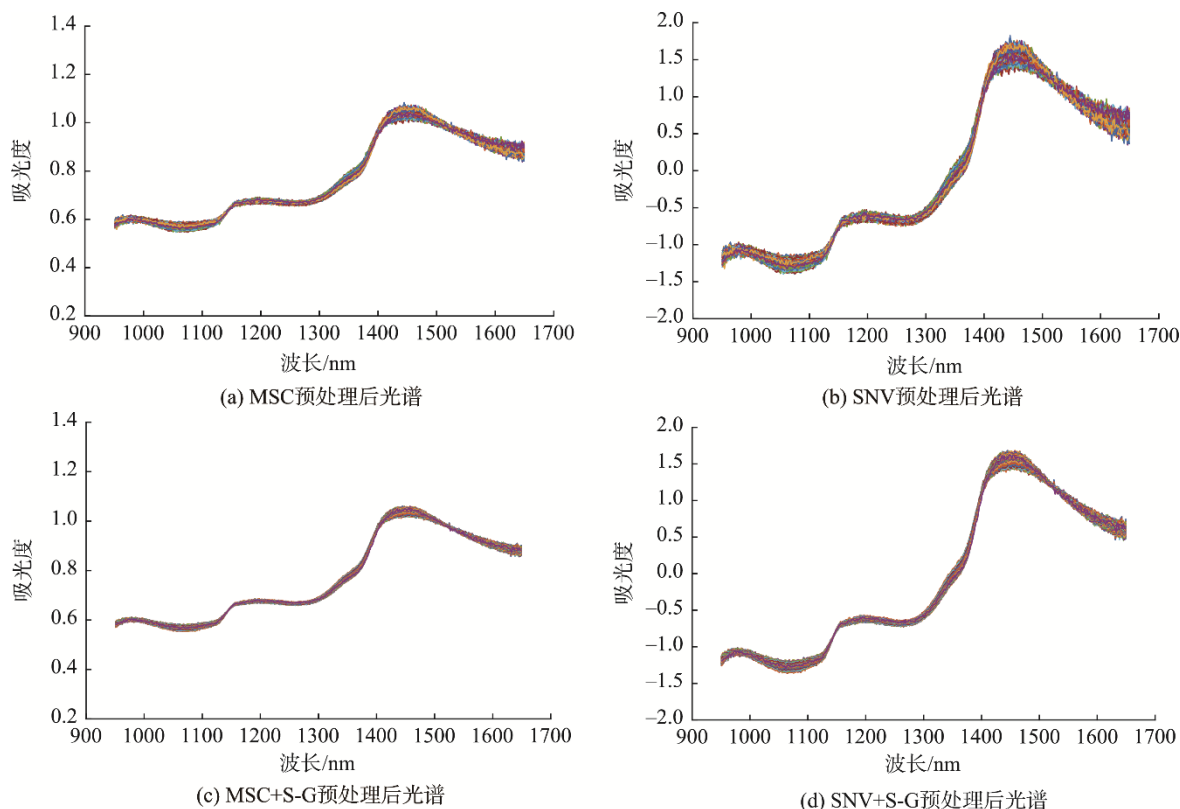


图 2 不同预处理方法后的光谱

Fig.2 Spectra after different preprocessing methods

对于 109 个北京鲜食杏样本, 利用 K-S 算法按照 3:1 的比例划分出校正集和预测集。划分结果如表 1 所示, 玉巴达杏的白利糖度范围在 12.8~15.7°Brix, 北寨红杏的白利糖度范围在 11.9~15.2°Brix, 半香白杏的白利糖度范围在 16.2~19.7°Brix, 半香白杏与玉巴达杏和北寨红杏的白利糖度差异较大; 校正集样本 80 个, 最大值 19.7°Brix, 最小值 11.9°Brix; 预测集样本 29 个, 最大值 19.0°Brix, 最小值 12.7°Brix。校正集和预测集中, 玉巴达杏、北寨红杏和半香白杏的数量分布均匀, 表明数据分布较为合理, 有利于提高预测模型的稳定性。

表 1 K-S 划分样本集结果统计
Table 1 K-S division sample set result statistics

样本指标	品种	数量	最小值 /°Brix	最大值 /°Brix	平均值 /°Brix
校正集	玉巴达杏	27	12.8	15.7	14.0
	北寨红杏	30	11.9	15.2	13.2
	半香白杏	23	16.2	19.7	18.3
预测集	玉巴达杏	8	12.7	14.7	13.6
	北寨红杏	7	13.2	14.3	13.8
	半香白杏	14	17.9	19.0	18.4

采用不同的预处理方法建立全波长的北京鲜食杏白利糖度的 PLSR 预测模型, 结果如表 2 所示。原始光谱经过 MSC+S-G 预处理后所建的预测模型效果最好, RMSEP 和 R_p 分别为 0.4750 和 0.9599, 与未进行预处理建立的模型相比, R_p 提高了 29.66%, RMSEP 减少了 61.25%, 大幅度提高了预测模型精度。

表 2 不同的预处理方法建立 PLSR 预测模型结果统计
Table 2 Statistics of the results of the PLSR prediction models established using different preprocessing methods

预处理方法	校正集		预测集	
	RMSEC	R_c	RMSEP	R_p
未处理	0.9059	0.8351	1.2261	0.7403
MSC	1.0281	0.7729	1.1279	0.7715
SNV	1.0365	0.7691	1.1450	0.7646
MSC+S-G	0.3969	0.9679	0.4750	0.9599
SNV+S-G	0.4112	0.9605	0.4958	0.9498

2.3 特征波长筛选

2.3.1 CARS 算法特征波长筛选

CARS 算法每次蒙特卡洛(Monte Carlo, MC)采样选择 PLSR 模型中回归系数绝对值大的变量点, 去除权重较小

的变量点, 利用 K 折交叉验证选出 PLSR 模型子集中交叉验证均方根误差(cross-validation root mean squared error, RMSECV)最小的 PLSR 模型, 该模型所包含的量子集即为选择的最优特征波长^[28]。本研究设置 MC 采样次数为 50 次, 采用 5 折交叉验证法计算。图 3 表示 CARS 算法特征波长筛选的结果。在图 3(a)中, 前 5 次 MC 采样中变量数大量减少, 后续采样中变量数减少速度变缓, 表明了 CARS 算法筛选特征波长的趋势。图 3(b)表明随着取样运行次数的增加, RMSECV 先减小后增大; 当 RMSECV 逐渐减小时, 与鲜食杏白利糖度预测无关的信息或部分共线信息被剔除; 当 RMSECV 逐渐增大时, 表明与鲜食杏白利糖度预测有关的信息被剔除。因此, RMSECV 最小时的采样次数为最优结果的采样次数, 即图 3(c)中星型垂线对应的第 29 次采样次数。最终共选出 19 个特征波长, 对应波长为 1002.66、1022.87、1102.51、1194.57、1212.42、1214.79、1372.79、1385.09、1390.66、1396.22、1399.56、1475.01、1477.16、1501.79、1503.92、1535.65、1541.94、1634.32、1640.34 nm。

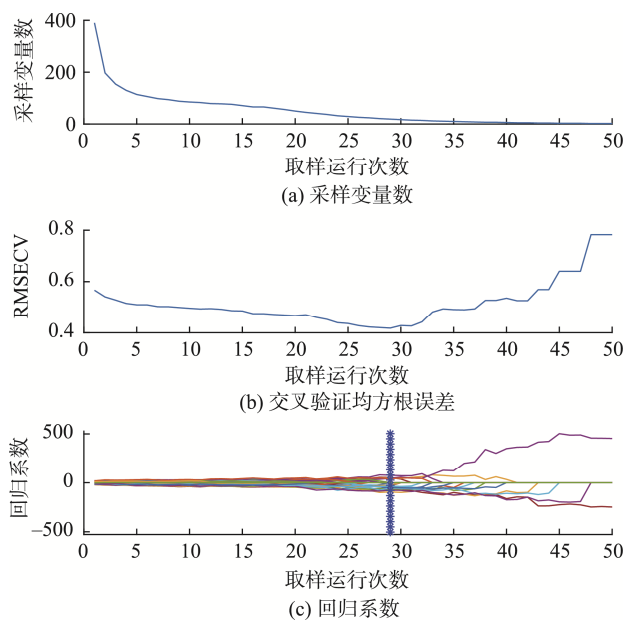
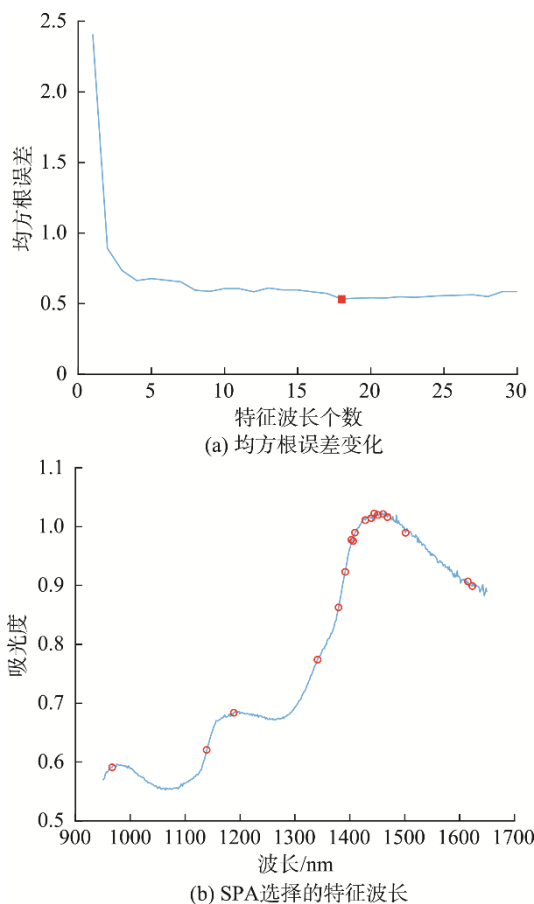


图 3 基于 CARS 算法的北京鲜食杏白利糖度特征波长筛选图
Fig.3 Characteristic wavelength selection maps of Beijing fresh *Armeniaca* Brix based on CARS algorithm

2.3.2 SPA 特征波长筛选

SPA 是一种向前循环选择方法^[29], 在使用 SPA 进行特征波长筛选时, 采用不同数量的波长分别建立 PLSR 模型, 计算这些模型的均方根误差(root mean square, RMSE), 选择最小的 RMSE 值对应的特征波长数量作为最优特征波长数量^[30]。对校正集样本进行 SPA 特征波长筛选, 设定模型选择的波长个数最小值为 1, 最大值为 30。图 4(a)为不同特征波长个数对应模型的 RMSE 变化趋势图, 图中

RMSE 随着特征波长个数的变化不断波动, 在特征波长个数为 18 时(图中实心方框), RMSE 达到最小值, 因此, 选取 18 为最优波长个数, 选取的波长在原始光谱中的位置如图 4(b)所示, 所选波长占原始光谱信息的 4.61%。



注: (a)图中实心方框为特征波长个数为 18 时的均方根误差值; (b)图的中国为各特征波长的位置。

图 4 基于 SPA 的北京鲜食杏白利糖度特征波长筛选图
Fig.4 Characteristic wavelength selection maps of Beijing fresh *Armeniaca* Brix based on SPA

2.3.3 基于 CARS 和 SPA 特征波长筛选的 PLSR 建模效果

为了研究 CARS 和 SPA 两种特征波长算法对模型精度的影响, 在不同预处理方法下, 使用 CARS 和 SPA 特征波长筛选算法分别建立北京鲜食杏白利糖度的 PLSR 预测模型, 结果如表 3 所示。结果表明, 在使用 CARS 波长筛选算法建立的 PLSR 模型中, MSC+S-G+CARS 所建的模型效果较好, R_c 和 R_p 分别为 0.9747、0.9616, 都大于其他模型, RMSEC 和 RMSEP 分别为 0.3502、0.4698, 都小于其他模型; 在使用 SPA 波长筛选算法建立的 PLSR 模型中, SNV+S-G+SPA 所建的模型效果较好, R_c 和 R_p 分别为 0.9588、0.9398, 都大于其他模型, RMSEC 和 RMSEP 分别为 0.4735、0.5137, 都小于其他模型。MSC+S-G+CARS 所建 PLSR 模型与 SNV+S-G+SPA 所建 PLSR 模型相

比, MSC+S-G+CARS 所建 PLSR 模型的 R_c 和 R_p 比 SNV+S-G+SPA 所建 PLSR 模型分别高出 0.0159、0.0218, SNV+S-G+SPA 所建 PLSR 模型的 RMSEC 和 RMSEP 比 MSC+S-G+CARS 所建 PLSR 模型分别高出 0.1233、0.0439, 因此本研究中 MSC+S-G+CARS+PLSR 是北京鲜食杏白利糖度的最优建模算法, 此外与 MSC+S-G+PLSR 算法的建模结果相比, 模型精度有略微提升, 但特征波长个数从 390 个减少到 19 个, 数据维度大幅降低。由上述分析可知, 特征波长筛选不但能够保持模型预测精度, 而且能够有效

降低数据维度。

2.4 基于最优特征波长筛选方法的 PLSR 建模

使用最优波长筛选算法 MSC+S-G+CARS 建立北京鲜食杏白利糖度的 PLSR 预测模型, 校正集建立的模型效果如图 5 所示, 模型的 RMSEC 和 R_c 分别为 0.3502、0.9747, 表明模型具有较好的预测能力。预测集模型效果如图 6 所示, 预测值与真实值之间差异不显著, 具有高度相关性, 模型预测效果较好。

表 3 基于不同特征波长提取算法建立的 PLSR 模型结果统计

Table 3 Statistical results of establishing PLSR model based on different feature wavelength extraction algorithms

特征波长提取算法	特征波长个数	预处理算法	RMSEC	R_c	RMSEP	R_p
未处理	390	未处理	0.9059	0.8351	1.2261	0.7403
	390	MSC	1.0281	0.7729	1.1279	0.7715
	390	SNV	1.0365	0.7691	1.1450	0.7646
	390	MSC+S-G	0.3969	0.9679	0.4750	0.9599
	390	SNV+S-G	0.4112	0.9605	0.4958	0.9498
CARS	56	未处理	0.7927	0.8743	1.0266	0.8137
	35	MSC	0.6316	0.9157	1.0294	0.8183
	38	SNV	0.5646	0.9364	0.8509	0.8714
	19	MSC+S-G	0.3502	0.9747	0.4698	0.9616
	31	SNV+S-G	0.3915	0.9651	0.4795	0.956
SPA	25	未处理	0.8316	0.8675	1.1136	0.7963
	20	MSC	0.7526	0.8972	1.1159	0.8135
	21	SNV	0.5946	0.9137	0.8841	0.8542
	18	MSC+S-G	0.5372	0.9513	0.5646	0.9317
	17	SNV+S-G	0.4735	0.9588	0.5137	0.9398

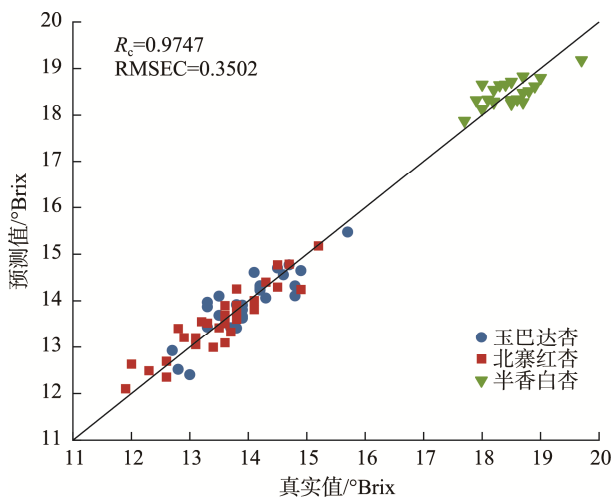


图 5 MSC+S-G+CARS+PLSR 模型的校正集效果
Fig.5 Effect of calibration set model based on MSC+S-G+CARS+PLSR algorithm

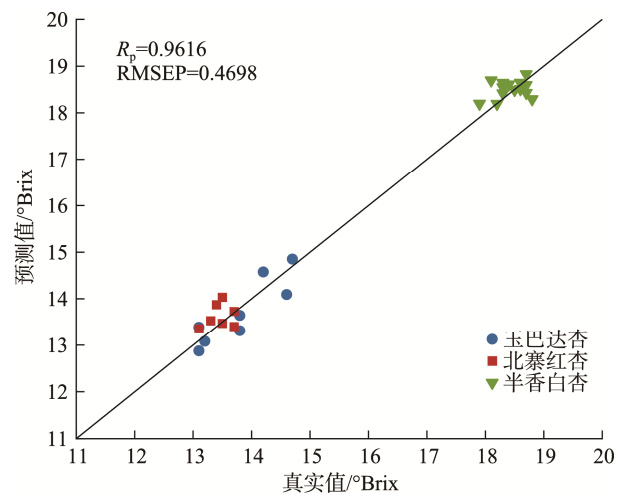


图 6 MSC+S-G+CARS+PLSR 模型的预测集效果
Fig.6 Effects of prediction set model based on MSC+S-G+CARS+PLSR algorithm

2.5 京郊鲜食杏白利糖度最优模型验证

为了验证基于 MSC+S-G+CARS+PLSR 算法建立的京郊鲜食杏白利糖度最优模型的准确性, 选取与样本无关的 15 个鲜食杏作为测试样本, 其中玉巴达杏 5 个、北寨红杏 5 个、半香白杏 5 个。鲜食杏测试样本的光谱数据和白利糖度理化值使用与建模样本同样的方法获得。测试结果如表 4 所示, 最大误差为 0.75, 最小误差 0.05, 结果表明模型有较好的检测精度, 可以实现对鲜食杏白利糖度的较为准确的预测。

表 4 15 个鲜食杏测试样本白利糖度的预测结果
Table 4 Prediction results of Brix of 15 fresh *Armeniaca* test samples

样本编号	品种	真实值 /°Brix	预测值 /°Brix	误差绝对值 /°Brix
1	玉巴达杏	13.20	13.62	0.42
2	玉巴达杏	14.60	14.16	0.44
3	玉巴达杏	14.30	13.55	0.75
4	玉巴达杏	13.90	14.34	0.44
5	玉巴达杏	13.50	14.14	0.64
6	北寨红杏	13.40	14.02	0.62
7	北寨红杏	13.50	13.43	0.07
8	北寨红杏	13.50	14.01	0.51
9	北寨红杏	13.50	13.77	0.27
10	北寨红杏	14.30	13.70	0.60
11	半香白杏	18.50	18.92	0.42
12	半香白杏	18.20	18.59	0.39
13	半香白杏	18.60	18.65	0.05
14	半香白杏	17.90	18.03	0.13
15	半香白杏	18.50	18.30	0.20

3 结 论

本研究基于便携式近红外光谱技术, 以北京郊区鲜食杏为研究对象, 使用便携式近红外光谱仪采集鲜食杏的光谱数据, 同时使用数显折射仪测定鲜食杏的白利糖度, 对光谱数据使用 MSC、SNV、S-G 算法进行预处理, 再结合 CARS 和 SPA 算法选择特征波长, 建立京郊鲜食杏白利糖度的最佳 PLSR 模型。结果表明, 最佳预处理方法为 MSC+S-G, 此时模型的 RMSEP 和 R_p 分别为 0.4750、0.9599, 在经过特征波长筛选后, MSC+S-G+CARS 是北京鲜食杏白利糖度的最优 PLSR 建模算法, 模型的 RMSEP 和 R_p 分别为 0.4698、0.9616, 相比较 MSC+S-G 建立的 PLSR 模型算法, RMSEP 减少, R_p 提高, 波长变量个数由 390 个减少到 19 个, 模型大小大幅降低且精度提高。基于 MSC+S-G+CARS+PLSR 算法建立鲜食杏白利糖度最优模型对 15 个鲜食杏样本的预测结果表明模型较为准确预测京郊鲜食杏白利糖度, 结合便携式近红外光谱技术可实现对京郊鲜食杏

品质的快速无损检测。在后续改进中, 可通过增加样本数量等方式提高预测模型的准确性和稳定性。本研究基于便携式光谱技术构建了京郊鲜食杏白利糖度的最佳预测模型, 满足了鲜食杏品质的快速无损检测需求, 为现场快速无损检测鲜食杏的品质提供了重要的技术支持。

参考文献

- [1] 徐铭, 刘威生, 王爱德, 等. 杏主要经济性状遗传分析[J]. 果树学报, 2020, 37(1): 1–10.
XU M, LIU WS, WANG AID, *et al.* Genetic analysis of main economic traits of apricot [J]. J Fruit Sci, 2020, 37(1): 1–10.
- [2] 王玉柱. 中国杏和李产业调查报告[M]. 北京: 中国农业出版社, 2017.
WANG YZ. China apricot and plum industry survey report [M]. Beijing: China Agric Press, 2017.
- [3] 孙肖瑞, 谢兆森, 尤佳玲, 等. 葡萄果实中糖转运、代谢与调控研究进展[J]. 植物生理学报, 2021, 57(3): 542–550.
SUN XR, XIE ZS, YOU JL, *et al.* Research progress on sugar transport, metabolism, and regulation in grape fruit [J]. Plant Physiol J, 2021, 57(3): 542–550.
- [4] 安涛, 张海亮, 陈少侃, 等. 荷斯坦牛初乳白利糖度值和初乳产量特征及影响因素分析[J]. 中国畜牧杂志, 2019, 55(9): 53–57.
AN T, ZHANG HL, CHEN SK, *et al.* Characteristics and influencing factors of colostrum brix and colostrum yield of Holstein cattle [J]. Chin J Anim Sci, 2019, 55(9): 53–57.
- [5] HUANG Z, WANG Q, XIA L, *et al.* Preliminary exploring of the association between sugars and anthocyanin accumulation in apricot fruit during ripening [J]. Sci Hort, 2019, 248: 112–117.
- [6] NSIBI M, GOUBLE B, BUREAU S, *et al.* Adoption and optimization of genomic selection to sustain breeding for apricot fruit quality [J]. G3: Genes, Genomes, Genet, 2020, 10(12): 4513–4529.
- [7] LIU B, JIAO W, WANG B, *et al.* Near freezing point storage compared with conventional low temperature storage on apricot fruit flavor quality (volatile, sugar, organic acid) promotion during storage and related shelf life [J]. Sci Hort, 2019, 249: 100–109.
- [8] 郭志明, 郭闯, 王明明, 等. 果蔬品质安全近红外光谱无损检测研究进展[J]. 食品安全质量检测学报, 2019, 10(24): 8280–8288.
GUO ZM, GUO C, WANG MM, *et al.* Research progress on near-infrared spectroscopy non-destructive testing of fruit and vegetable quality and safety [J]. J Food Saf Qual, 2019, 10(24): 8280–8288.
- [9] 雷鹰, 刘翠玲, 周子彦. 应用便携式近红外光谱仪研究苹果糖度的快速分析模型[J]. 食品科学技术学报, 2018, 36(6): 95–100.
LEI Y, LIU CL, ZHOU ZY. A fast analytical model for apple sugar content study using portable near-infrared spectrometer [J]. J Food Sci Technol, 2018, 36(6): 95–100.
- [10] MOGOLLON MR, CONTRERAS C, FREITAS ST, *et al.* NIR spectral models for early detection of bitter pit in asymptomatic ‘Fuji’ apples [J]. Sci Hort, 2021, 280: 109945.
- [11] 王冬, 张鹤冬, 朱业伟, 等. 基于 LVF-NIR 的樱桃可溶性固形物含量无损检测模型的建立与优化[J]. 食品安全质量检测学报, 2020, 11(3): 854–859.
WANG D, ZHANG HD, ZHU YW, *et al.* Establishment and optimization of a nondestructive rapid measurement model for soluble solids content in cherry based on LVF-NIR [J]. J Food Saf Qual, 2020, 11(3): 854–859.

- [12] 徐金阳, 刘翠玲, 周子彦, 等. 基于便携式仪器的小麦面粉品质检测方法研究[J]. 食品安全质量检测学报, 2019, 10(6): 1734-1739.
XU JY, LIU CL, ZHOU ZY, *et al.* Research on quality detection method of wheat flour based on portable instrument [J]. *J Food Saf Qual*, 2019, 10(6): 1734-1739.
- [13] 李庆, 闫晓剑, 赵魁, 等. 基于云端-互联便携式近红外技术现场快检西红花真伪[J]. 光谱学与光谱分析, 2020, 40(10): 3029-3037.
LI Q, YAN XJ, ZHAO K, *et al.* On-site quick inspection of the authenticity of saffron based on cloud-connected portable near-infrared technology [J]. *Spectrosc Spec Anal*, 2020, 40(10): 3029-3037.
- [14] 袁雷明, 郭珍珠, 陈孝敬, 等. 基于可见/近红外光谱技术的便携分析仪的应用[J]. 食品安全质量检测学报, 2017, 8(9): 3455-3460.
YUAN LM, GUO ZZ, CHEN XJ, *et al.* Application of portable analyzer based on visible/near-infrared spectroscopy [J]. *J Food Saf Qual*, 2017, 8(9): 3455-3460.
- [15] 孙阳, 刘翠玲, 孙晓荣, 等. 基于便携式近红外光谱仪的樱桃番茄糖分快速分析模型[J]. 食品与发酵工业, 2021, 47(23): 214-220.
SUN Y, LIU CL, SUN XR, *et al.* Rapid analysis model of cherry tomato sugar based on portable near-infrared spectrometer [J]. *Food Ferment Ind*, 2021, 47(23): 214-220.
- [16] 王燕, 付琪, 李颖, 等. 近红外光谱分析技术快速检测藕粉品质[J]. 食品安全质量检测学报, 2022, 13(15): 5026-5034.
WANG Y, FU Q, LI Y, *et al.* Rapid detection of lotus root starch quality by near-infrared spectroscopy [J]. *J Food Saf Qual*, 2022, 13(15): 5026-5034.
- [17] 黄燕, 王璐, 关海鹏, 等. 基于优选 NIR 光谱波数的绿豆产地无损检测方法[J]. 光谱学与光谱分析, 2021, 41(4): 1188-1193.
HUANG Y, WANG L, GUAN HOU, *et al.* Nondestructive detection method of mung bean origin based on optimized NIR spectral wavenumber [J]. *Spectrosc Spec Anal*, 2021, 41(4): 1188-1193.
- [18] 王冬, 王世芳, 罗娜, 等. 基于数字光处理技术的梨可溶性固形物含量的无损速测研究[J]. 食品安全质量检测学报, 2018, 9(11): 2722-2727.
WANG D, WANG SF, LUO N, *et al.* Nondestructive and rapid measurement of pear soluble solids content based on digital light processing technology [J]. *J Food Saf Qual*, 2018, 9(11): 2722-2727.
- [19] AIGAWWAM S, BENAÏSSA M. Robust eye blink detection based on eye landmarks and Savitzky-Golay filtering [J]. *Information*, 2018, 9(4): 93.
- [20] PHUPHAPHUD A, SAENGPRACHATANARUG K, POSOM J, *et al.* Non-destructive and rapid measurement of sugar content in growing cane stalks for breeding programmers using visible-near infrared spectroscopy [J]. *Biol Eng*, 2020, 197: 76-90.
- [21] 王转卫, 迟茜, 郭文川, 等. 基于近红外光谱技术的发育后期苹果内部品质检测[J]. 农业机械学报, 2018, 49(5): 348-354.
WANG ZW, CHI Q, GUO WC, *et al.* Detection of apple internal quality in late development stage based on near-infrared spectroscopy [J]. *Trans Chin Soc Agric Mach*, 2018, 49(5): 348-354.
- [22] 张静, 徐阳, 姜彦武, 等. 近红外光谱技术在葡萄及其制品品质检测中的应用研究进展[J]. 光谱学与光谱分析, 2021, 41(12): 3653-3659.
ZHANG J, XU Y, JIANG YW, *et al.* Research progress on the application of near-infrared spectroscopy in the quality inspection of grapes and their products [J]. *Spectrosc Spec Anal*, 2021, 41(12): 3653-3659.
- [23] LIANG K, HUANG J, HE R, *et al.* Comparison of Vis-NIR and SWIR hyperspectral imaging for the non-destructive detection of DON levels in *Fusarium* head blight wheat kernels and wheat flour [J]. *Infrared Phys Technol*, 2020, 106: 103281.
- [24] ZHANG M, SHEN M, PU Y, *et al.* Rapid identification of apple maturity based on multispectral sensor combined with spectral shape features [J]. *Sci Hortic*, 2022, 8(5): 361.
- [25] 高升, 王巧华, 李庆旭, 等. 基于近红外光谱的红提维生素 C 含量、糖度及总酸含量无损检测方法[J]. 分析化学, 2019, 47(6): 941-949.
GAO S, WANG QH, LI QX, *et al.* Non-destructive testing method for vitamin C content, sugar content and total acid content in red extract based on near-infrared spectroscopy [J]. *Chin J Anal Chem*, 2019, 47(6): 941-949.
- [26] 章世奎, 罗晓琴, 王亚铜, 等. 调环酸钙对轮台白杏营养生长和果实品质的影响[J]. 新疆农业科学, 2021, 58(5): 846-853.
ZHANG SK, LUO XQ, WANG YT, *et al.* Effects of prohexadione calcium on vegetative growth and fruit quality of Luntai white apricot [J]. *Xinjiang Agric Sci*, 2021, 58(5): 846-853.
- [27] 于丹, 刘娟娟, 毕泉鑫, 等. 低温胁迫及恢复对杏树品种雌蕊生长发育的影响[J]. 东北林业大学学报, 2021, 49(1): 1-5.
YU D, LIU XJ, BI QX, *et al.* Effects of low temperature stress and restoration on the growth and development of pistils of apricot cultivars [J]. *J Northeast For Univ*, 2021, 49(1): 1-5.
- [28] 彭海根, 金楹, 詹菡国, 等. 近红外光谱技术结合竞争自适应重加权采样变量选择算法快速测定土壤水解性氮含量[J]. 分析测试学报, 2020, 39(10): 1305-1310.
PENG HG, JIN Y, ZHAN YG, *et al.* Rapid determination of soil hydrolyzable nitrogen content by near infrared spectroscopy combined with competitive adaptive weighted sampling variable selection algorithm [J]. *J Instrum Anal*, 2020, 39(10): 1305-1310.
- [29] 高升, 王巧华, 施行, 等. 便携式红提葡萄多品质可见/近红外检测仪设计与试验[J]. 农业机械学报, 2021, 52(2): 308-319.
GAO S, WANG QH, SHI X, *et al.* Design and experiment of a portable multi-quality visible/near-infrared detector for red grapes [J]. *Trans Chin Soc Agric Mach*, 2021, 52(2): 308-319.
- [30] 张立欣, 杨翠芳, 陈杰, 等. 基于变量优选的苹果糖分含量近红外光谱检测[J]. 食品与机械, 2021, 37(10): 112-118.
ZHANG LX, YANG CF, CHEN J, *et al.* Detection of sugar content in apples by near-infrared spectroscopy based on variable optimization [J]. *Food Mach*, 2021, 37(10): 112-118.

(责任编辑: 韩晓红 于梦娇)

作者简介



刘翠玲, 教授, 博士生导师, 主要研究方向为智能测量技术与数据处理研究。
E-mail: liucl@btbu.edu.cn



闻世震, 硕士研究生, 主要研究方向为光谱检测技术与系统研制。
E-mail: wsz412801@163.com