

DOI: 10.13718/j.cnki.xdzk.2020.12.007

# W·默科特果实内在品质近红外光谱检测研究

李勋兰, 尹旭敏, 杨蕾, 洪林, 韩国辉, 王武

重庆市农业科学院, 重庆 401329

**摘要:** 探索了近红外光谱技术对 W·默科特果实内在多项品质指标进行快速、无损检测的可行性。采集 1 000~2 500 nm 波段的 W·默科特果实的近红外光谱, 分别采用间隔偏最小二乘法(iPLS)、竞争性自适应重加权算法(CARS)和间隔偏最小二乘法结合连续投影算法(iPLS-SPA)3 种特征波长筛选方法对全波段进行特征波长筛选, 利用全波段及 3 种特征波长筛选方法得到的特征波段对应光谱信息建立了果实可溶性固形物(TSS)、可滴定酸(TA)和维生素 C( $V_C$ )含量的最小二乘支持向量机回归(LS-SVR)预测模型和偏最小二乘回归(PLSR)预测模型。结果显示, 采用 CARS 筛选的特征波长所对应光谱信息建立的 LS-SVR 预测模型精度最高, 模型对 TSS 和 TA 和  $V_C$  含量的预测集相关系数分别达到 0.91, 0.85 和 0.91, 且模型对应的预测集均方根误差(RMSEP)分别为 0.26, 0.03 和 0.25。说明采用近红外光谱技术结合 CARS 和 LS-SVR 可实现对 W·默科特果实 TSS 和 TA 和  $V_C$  含量的同时检测。

**关键词:** 近红外光谱; W·默科特; 内在品质; 含量检测

**中图分类号:** S666.4

**文献标志码:** A

**文章编号:** 1673-9868(2020)12-0060-07

柑橘是我国栽培面积最大的果树作物, 柑橘产业也是南方地区重要的农村经济支柱产业。全国柑橘产业标准化程度不高, 导致果实品质优良混杂, 采后果实精细分级少, 影响柑橘品牌价值和产业经济效益, 而进行柑橘果实分级对提高其商品价值和加工附加值具有积极作用<sup>[1]</sup>。目前, 我国柑橘产地主要按果实的大小、健康与病果等外表进行果实分级, 而分级指标涉及内在品质较少, 产地非损伤品质检测和监测技术发展缓慢。W·默科特果实肉质细嫩化渣且多汁, 口感酸甜适口且风味浓, 是我国南方地区主栽优良晚熟柑橘之一, 对当地农业经济发展具有重要意义。

光谱技术具有无损、快速、样品无需前处理等许多独特的优点, 利用光谱技术对果实品质进行检测的研究基本已涉及大部分水果<sup>[2-6]</sup>和果实中可溶性固形物(total soluble solids, TSS)、酸度(titratable acid, TA)、维生素 C(vitamin C,  $V_C$ )、花青素等品质指标, 研究内容主要集中于光谱获取方式、特征波段的选择、建模方法和模型评价等, 但采用光谱技术目前对水果品质的检测多为一种模型只能检测一种指标, 且模型的通用性较差。因此, 开发利用光谱技术以实现果实多指标同时检测、通用性强、精度高的柑橘果实品质检测技术具有重要意义。

本研究以 W·默科特果实为对象, 分析测定果实中可溶性固形物、可滴定酸和维生素 C 含量, 结合近

收稿日期: 2019-09-04

基金项目: 重庆市科研机构绩效激励引导专项(cstc2017jxjl80006); 重庆市基础与前沿研究计划项目(cstc2016jcyjA0046); 重庆市社会事业与民生保障科技创新专项(cstc2016shms-ztx80002); 重庆市财政基础研究项目(NKY-2017AC017)。

作者简介: 李勋兰(1988—), 女, 硕士, 助理研究员, 主要从事果树生理与信息化的研究。

通信作者: 王武, 副研究员。

红外光谱技术, 采用间隔偏最小二乘法(interval partial least squares, iPLS)、间隔偏最小二乘法结合连续投影算法(Successive projections algorithm, SPA)和竞争性自适应重加权算法(Competitive adaptive re-weighted sampling, CARS)提取特征波长, 并采用最小二乘支持向量机回归法(Least square support vector regression, LS-SVR)和偏最小二乘回归法(Partial least squares regression, PLSR)建立果实中 3 种内在品质指标含量预测模型, 以期为柑橘果实内在品质快速无损检测 and 水果智能分级技术提供参考。

## 1 材料与方法

### 1.1 实验材料

供试果实采自重庆市江津区重庆市农业科学院果树研究所现代农业示范基地, 品种为 8 年生 W·默科特(W·murcott, *Citrus reticulata* Blanco × *Citrus sinensis* Osb. cv Murcott), 枳橙砧(*Citrus siimensis* (L.) Osb. × *Poncirus trifoliata* (L.) Raf.), 于 2016—2018 年度的 1 月中旬连续 3 年对选取的 15 株 W·默科特植株进行果实采样, 从树冠四周中采集健康和大小一致的果实, 共获得样本 180 个。采用 Kennard-Stone 法按照 3 : 1 的比例将样本分为校正集和预测集, 其中校正集 135 个, 预测集 45 个。

### 1.2 光谱数据采集

将果实清洗晾干后, 利用聚光科技(杭州)股份有限公司生产的便携式近红外光谱仪(波长范围 1 000~2 500 nm, 光谱采样间隔为 1 nm, 扫描次数为 10 次, 光源采用 12V, 45W 卤钨灯), 在果面赤道线 3 个方向(夹角 120°)各进行 1 次光谱扫描, 每个果实共计 3 次, 取其平均光谱为该果实光谱信息。光谱采集前用标准白板进行参比扫描。

### 1.3 内在品质检测

对每个果实采集光谱信息以后进行单果品质检测。果实品质理化指标参照 GB/T8210-2011《柑橘鲜果检验法》中方法测定, 可溶性固形物用手持糖度折光仪(Pocket Pal-1, Atago, Japan)测定, 可滴定酸采用酸碱滴定法测定, 维生素 C 采用 2,6-二氯酚酚滴定法测定, 所有数据均进行 3 次技术重复。

### 1.4 数据处理

利用多元散射校正(Multiplicative scatter correction, MSC)法对原始光谱数据进行预处理, 以消除表面散射以及光程变化对近红外漫反射光谱的影响。

分别采用 iPLS, iPLS-SPA 和 CARS 3 种常用特征提取法对全波段光谱信息进行特征波段提取。iPLS 是一种基于移动窗口法和最小二乘法的光谱特征提取方法。SPA 是一种前向变量选择方法, 能有效寻找含有最低限度的冗余信息变量组, 使变量之间的共线性达到最小, 但是 SPA 在寻找最低限度冗余信息变量时会增加信号的信噪比, 对全波段寻找会增加计算量。因此, 本文利用 SPA 对 iPLS 得到的特征波段区间进行特征波长提取。CARS 是一种基于蒙特卡罗采样和 PLS 回归系数的特征波长选择方法, 也是常用光谱特征波长提取方法之一。

采用 LS-SVR 和 PLSR 分别对不同特征筛选方法所提取的特征光谱信息构建果实 TSS, TA 和  $V_c$  含量预测模型, 同时采用 K 折交叉验证(K-fold Cross Validation)和网格搜索法对 LS-SVR 算法中核函数参数进行优化。

## 2 结果与分析

### 2.1 果实 TSS, TA 和 $V_c$ 含量

从校正集和验证集样品果实 TSS, TA 和  $V_c$  含量统计结果(表 1)可以看出, 在所划分的校正集样品中, 果实 TSS, TA 和  $V_c$  含量最大值皆大于预测集的值, 而校正集中果实 TSS, TA,  $V_c$  含量最小值皆小于预测集的值。说明校正集样本所覆盖的范围皆大于预测集所覆盖的范围, 采用 Kennard-Stone 法所划分的校正集代表性较强, 样本划分合理, 可用于建立预测模型。

表 1 果实 TSS, TA 和  $V_c$  实测值统计

内在品质	样本集	样本数	范围	平均值	标准差	变异系数
TSS/%	校正集	135	9.90~13.20	11.28	0.80	0.07
	预测集	45	10.10~12.50	11.11	0.63	0.06
TA/g	校正集	135	0.71~1.13	0.88	0.10	0.11
	预测集	45	0.74~0.99	0.86	0.06	0.07
$V_c$ /mg	校正集	135	20.68~24.02	22.24	0.71	0.03
	预测集	45	21.01~23.49	22.35	0.60	0.02

注:表中 TA 和  $V_c$  数据以每 100 mL 果汁计。

## 2.2 果实光谱反射值曲线

从果实原始近红外光谱(图 1)可以看出,在 1 000~2 500 nm 近红外光谱区有 3 个较强的吸收峰. 1 200 nm 和 1 920 nm 左右的吸收峰,主要是 O—H 键各级倍频的变化,是由水分含量引起;1 450 nm 处吸收峰附近主要是 C—H, —CH<sub>2</sub> 键的变化,是由大分子糖类物质所引起. 说明利用近红外光谱检测柑橘果实内在品质具有可行性.

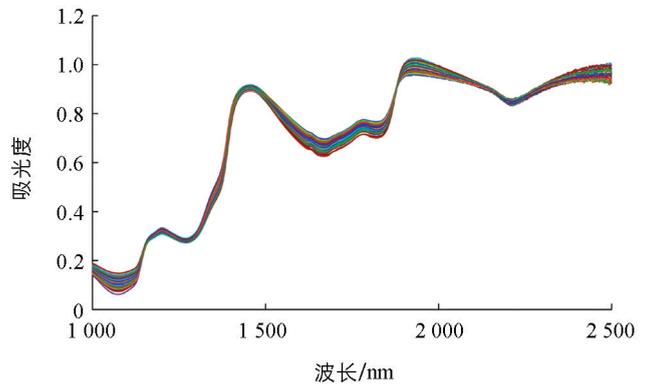


图 1 W·默科特果实光谱曲线

## 2.3 特征波长提取

对预处理后全波段光谱分别采用 CARS, iPLS 和 iPLS-SPA 提取得到特征波长(表 2,图 2).

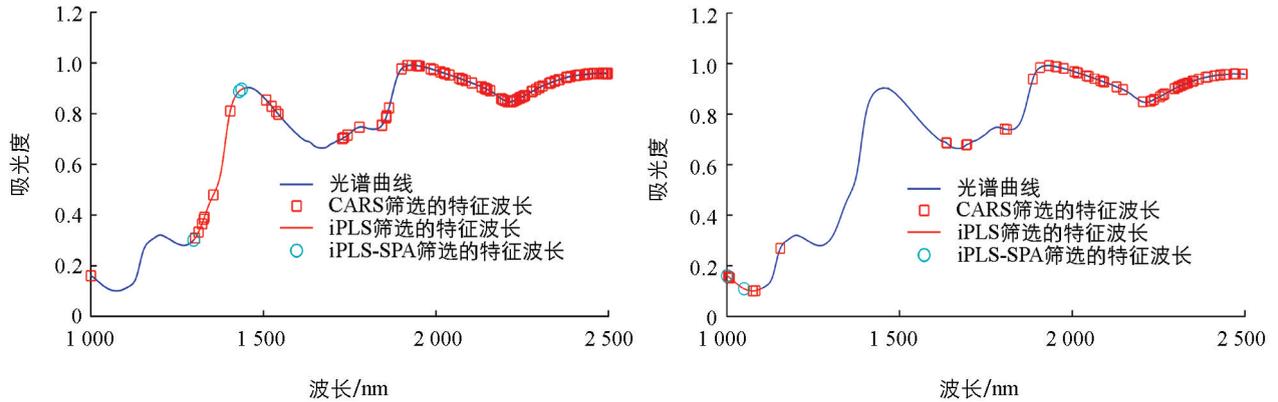
表 2 不同波长提取方法筛选特征波长结果

检测指标	波长筛选方法	波长个数	特征波长范围/nm
TSS	none	1 500	1 000~2 500
	CARS	118	图 2a
	iPLS	150	1 300~1 449
	iPLS-SPA	3	1 300,1 432,1 439
TA	none	1 500	1 000~2 500
	CARS	74	图 2b
	iPLS	100	1 000~1 099
	iPLS-SPA	2	1 000,1 050
$V_c$	none	1 500	1 000~2 500
	CARS	41	图 2c
	iPLS	167	1 000~1 166
	iPLS-SPA	5	1 066,1 149,1 155,1 136,1 165

从表 2 可以看出,采用 iPLS 和 iPLS-SPA 对 TSS 含量预测所筛选特征波长在 1 300~1 449 nm 区域,采用 CARS 对 TSS 含量预测所筛选特征波长主要为 1 303~1 544 nm,1 728~1 993 nm 和 2 012~2 498 nm,3 种方法提取的特征波长在 1 300~1 449 nm 区域具有交叉重叠;对 TA 含量预测采用 iPLS 和 iPLS-SPA 筛选的特征波长在 1 000~1 099 nm 区域,采用 CARS 筛选的特征波长区域主要集中在 1 003~1 009 nm,

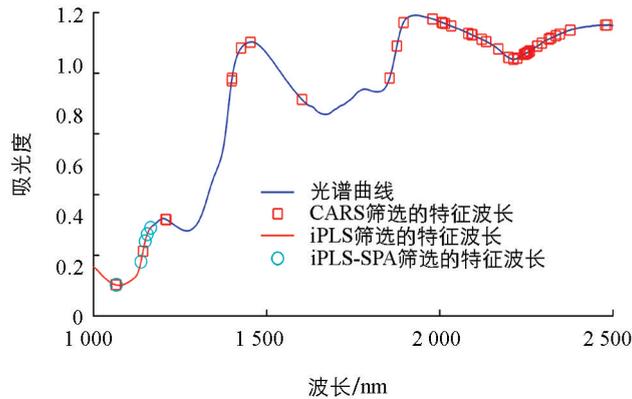
1 075~1 083 nm, 1 907~2 494 nm, 3 种方法提取的特征波长在 1 003~1 009 nm 和 1 075~1 083 nm 区域具有交叉重叠; 对  $V_c$  含量预测采用 iPLS 和 iPLS-SPA 筛选的特征波长在 1 000~1 166 nm 区域, 采用 CARS 筛选的特征波长区域主要集中在 1 064~1 066 nm, 1 208~1 602 nm, 1 855~2 483 nm, 3 种方法提取的特征波长在 1 064~1 066 nm 区域具有交叉重叠。

从图 2 可以看出, 采用不同特征提取方法所获得的特征波长区域存在差异; 3 种特征波长方法皆减少了预测模型输入的数据量, 其中采用 iPLS-SPA 所获得的特征波长个数最少; CARS 提取的波长区域范围最大, 但采用 iPLS 所提取的波长数最多。



(a) 采用3种方法对预测TSS含量所提取的特征波长

(b) 采用3种方法对预测TA含量所提取的特征波长



(c) 采用3种方法对预测 $V_c$ 含量所提取的特征波长

图 2 3 种方法所筛选特征波长

## 2.4 模型的建立

将全波长 1 500 个变量以及 CARS, iPLS 和 iPLS-SPA 提取的特征波长对应光谱信息作为输入, 建立 TSS, TA 和  $V_c$  含量 LS-SVR 和 PLSR 预测模型(表 3)。采用校正集和预测集均方根误差 ( $RMSEC$ ,  $RMSEP$ ) 和相关系数 ( $R_c$ ,  $R_p$ ) 作为模型的评价指标。从表 3 可以看出, 采用全波段, iPLS 和 iPLS-SPA 提取的特征波长对应光谱信息作为模型输入, 建立的果实 TSS 含量 LS-SVR 和 PLSR 预测模型的预测集和建模集的相关系数和  $RMSE$  接近, 但采用 CARS 结合 LS-SVR 获得的模型精度最高, 预测相关系数达到 0.91, 且  $RMSEP$  值最低; 采用全波段, iPLS 和 iPLS-SPA 结合 PLSR 建立的模型对果实 TA 含量预测精度高于 LS-SVR 模型, 采用 CARS 筛选波长对应光谱信息建立的 PLSR 和 LS-SVR 模型比采用其他 3 种波长对应光谱信息建立的模型预测精度都高, 且采用 CARS 结合 LS-SVR 建立模型对 TA 含量预测相关系数达到 0.85, 且  $RMSEP$  值最低; 采用全波段, iPLS 和 iPLS-SPA 结合 LS-SVR 和 PLSR 建立的  $V_c$  含量模型稳定性和精度皆较低, 而采用 CARS 提取的特征波长对应光谱信息建立的 LS-SVR 预测模型对  $V_c$  含量预测精度最高, 预测相关系数达到 0.91, 且  $RMSEP$  值较低。根据预测模型不但要有高的相关系数, 较小的  $RMSEP$  和  $RMSEC$ , 且  $RMSEP$ ,  $RMSEC$  之间的差异要较小的模型性能评判原则, 得出采用 CARS 提

取的特征波长建立的 LS-SVR 模型对 W·默科特果实 TSS, TA 和  $V_c$  含量预测效果最为理想.

表 3 建模预测结果

检测指标	波长筛选方法	LS-SVR				PLSR			
		RMSEC	$R_c$	RMSEP	$R_p$	RMSEC	$R_c$	RMSEP	$R_p$
TSS	none	0.30	0.93	0.34	0.84	0.35	0.90	0.34	0.84
	CARS	0.19	0.97	0.26	0.91	0.24	0.96	0.48	0.63
	iPLS	0.37	0.89	0.37	0.80	0.37	0.89	0.36	0.81
	iPLS-SPA	0.35	0.90	0.35	0.83	0.38	0.88	0.35	0.83
TA	none	0.04	0.93	0.04	0.79	0.06	0.81	0.03	0.84
	CARS	0.02	0.99	0.03	0.85	0.06	0.93	0.03	0.85
	iPLS	0.04	0.90	0.04	0.80	0.06	0.81	0.03	0.84
	iPLS-SPA	0.07	0.86	0.03	0.76	0.07	0.74	0.03	0.87
$V_c$	none	0.26	0.93	0.34	0.82	0.47	0.74	0.41	0.73
	CARS	0.19	0.96	0.25	0.91	0.33	0.88	0.29	0.87
	iPLS	0.50	0.72	0.18	0.69	0.51	0.69	0.41	0.72
	iPLS-SPA	0.40	0.68	0.35	0.65	0.53	0.66	0.34	0.83

### 3 讨 论

利用光谱技术对果实品质进行检测已有较多研究,研究者采用较多的波长范围主要集中在可见光和近红外区域.本研究采用 1 000~2 500 nm 近红外波长区域,从 W·默科特果实光谱反射曲线来看,在该区域该波形与其他柑橘品种基本一致<sup>[7-9]</sup>,与甜柿<sup>[10]</sup>、西瓜<sup>[11]</sup>、黄桃<sup>[12]</sup>等果实也具有相同的吸收峰和波形.在 1 000~1 450 nm 波段区域,该波段区域与柑橘果汁<sup>[7-8]</sup>光谱曲线具有相同波形和吸收峰,说明 1 000~1 450 nm 波段对光谱技术检测果实内部品质紧密相关.

特征提取是光谱数据挖掘中的一个核心环节,进行特征波段筛选可以缩减建模输入参数,减少模型运算时间、消除多余干扰信息,提高模型的预测精度.本研究采用 iPLS, iPLS-SPA 和 CARS 3 种方法进行特征波长提取,获得的特征波长区域存在差异,说明建模采用的特征波长与采用的提取方法相关.本研究果实 TSS 含量最佳模型采用的特征光谱波段与 Guo 等<sup>[13]</sup>建立的桃,田喜等<sup>[7]</sup>建立的脐橙, Cunha Junior 等<sup>[14]</sup>建立的棕榈果 TSS 预测最佳模型采用的特征波长相互交叉重叠,所采用的特征波长集中于 1 303~1 544 nm, 1 728~1 993 nm 和 2 300~2 399 nm 区域.果实 TA 含量最佳模型采用特征光谱波段与代芬等<sup>[15]</sup>建立的砂糖橘, Guo 等<sup>[16]</sup>建立的枣和 Louw 等<sup>[17]</sup>建立的李果实的 TA 含量预测最佳模型采用的特征波长相互交叉重叠,所采用的特征波长主要集中于 1 900~2 300 nm 区域.果实  $V_c$  含量最佳模型采用特征光谱波段与夏俊芳等<sup>[18]</sup>建立的脐橙以及高升等<sup>[19]</sup>、陈辰等<sup>[20]</sup>建立的红提果实  $V_c$  含量预测最佳模型采用的特征波长相互交叉重叠,所采用的特征波长集中于 1 064~1 066 nm, 1 208~1 602 nm 和 1 855~2 483 nm 区域.由于采用设备仪器、波长筛选方法以及涉及水果种类差异等,研究者采用光谱技术对水果品质指标检测所采用的波长不一致,即使是针对相同品种的水果不同研究者建模采用的波长范围也不一致.本研究最终采用 CARS 结合 LS-SVR 建立了 W·默科特果实 SSC, TA 和  $V_c$  含量预测的通用模型,但果实内在品质包括较多指标,是一个较为复杂的系统<sup>[21-22]</sup>,该模型对其他柑橘品种果实中 SSC, TA 和  $V_c$  含量的预测精度及适用性仍需进行验证.建立采用波数个数少、预测精度高、通用性强的果实内部品质含量模型仍是目前研究的重点.

本研究采用两种方法建立了柑橘果实内在品质预测模型,PLSR 是化学计量学分析过程中最常用的多元线性建模方法,PLSR 在计算理化指标与光谱信息之间关系是依靠寻找线性相关性来实现的;LS-SVR 是一种基于统计学习理论的机器学习算法,在计算过程中采用非线性映射的核函数,可用于处理非线性回

归问题. 总体上, 采用相同的数据建立的 LS-SVR 模型预测精度高于 PLSR 模型, 该结果与詹白勺等<sup>[23]</sup>建立的对香梨可溶性、牛晓颖等<sup>[24]</sup>建立的对草莓固酸比和可滴定酸及黄康等<sup>[25]</sup>建立的对番茄汁糖酸度预测模型结果一致. 模型预测结果说明, 果实内在品质理化指标与光谱信息的变化关系为非线性关系, 宜采用非线性算法构建果实内在品质理化指标预测模型.

## 4 结 论

本文利用近红外光谱技术对 W·默科特果实的 SSC, TA 和  $V_c$  含量进行了定量分析研究, 得出采用 CARS, iPLS 和 iPLS-SPA 3 种方法可极大地简化模型; LS-SVR 与 PLSR 相比更适用于果实内在品质定量分析模型的建立; 利用近红外光谱技术, 采用 CARS 筛选特征波长结合 LS-SVR 可用于 W·默科特果实的 SSC, TA 和  $V_c$  含量的同时检测.

## 参考文献:

- [1] 沈兆敏. 我国柑橘产销现状、发展趋势及对策建议 [J]. 果农之友, 2017(9): 1.
- [2] MENDOZA F, LU R F, CEN H Y. Grading of Apples Based on Firmness and Soluble Solids Content Using Vis/SW-NIR Spectroscopy and Spectral Scattering Techniques [J]. Journal of Food Engineering, 2014, 125: 59-68.
- [3] FAN S X, HUANG W Q, GUO Z M, et al. Prediction of Soluble Solids Content and Firmness of Pears Using Hyperspectral Reflectance Imaging [J]. Food Analytical Methods, 2015, 8(8): 1936-1946.
- [4] 孙 通, 莫欣欣, 刘木华. 果皮对脐橙可溶性固形物可见/近红外检测精度的影响 [J]. 光谱学与光谱分析, 2018, 38(5): 1406-1411.
- [5] HU M H, DONG Q L, LIU B L, et al. Prediction of Mechanical Properties of Blueberry Using Hyperspectral Interactance Imaging [J]. Postharvest Biology and Technology, 2016, 115: 122-131.
- [6] WANG A C, HU D, XIE L J. Comparison of Detection Modes in Terms of the Necessity of Visible Region (VIS) and Influence of the Peel on Soluble Solids Content (SSC) Determination of Navel Orange Using VIS-SWNIR Spectroscopy [J]. Journal of Food Engineering, 2014, 126: 126-132.
- [7] 田 喜, 何绍兰, 吕 强, 等. 脐橙可溶性固形物含量的光谱检测技术研究 [J]. 西南大学学报(自然科学版), 2014, 36(3): 122-129.
- [8] 廖秋红, 何绍兰, 谢让金, 等. 基于近红外光谱的纽荷尔脐橙产地识别研究 [J]. 中国农业科学, 2015, 48(20): 4111-4119.
- [9] 洪 涯, 洪添胜, 代 芬, 等. 连续投影算法在砂糖橘总酸无损检测中的应用 [J]. 农业工程学报, 2010, 26(S2): 380-384.
- [10] 王 丹, 鲁晓翔, 张 鹏, 等. 近红外无损检测甜柿果实质地和品质 [J]. 食品工业科技, 2013, 34(24): 53-56.
- [11] 王世芳, 韩 平, 崔广禄, 王冬, 刘珊珊, 赵跃. SPXY 算法的西瓜可溶性固形物近红外光谱检测 [J]. 光谱学与光谱分析, 2019, 39(3): 738-742.
- [12] 王铭海, 郭文川, 商 亮, 等. 基于近红外漫反射光谱的多品种桃可溶性固形物的无损检测 [J]. 西北农林科技大学学报(自然科学版), 2014, 42(2): 142-148.
- [13] GUO W C, GU J S, LIU D Y, et al. Peach Variety Identification Using Near-infrared Diffuse Reflectance Spectroscopy [J]. Computers and Electronics in Agriculture, 2016, 123: 297-303.
- [14] CUNHA JUNIOR L C, NARDINI V, KHATIWADA B P, et al. Classification of Intact Açaí (*Euterpe Oleracea* Mart.) and Juçara (*Euterpe Edulis* Mart) Fruits Based on Dry Matter Content by Means of near Infrared Spectroscopy [J]. Food Control, 2015, 50: 630-636.
- [15] 代 芬, 洪添胜, 罗 霞, 等. 基于可见-近红外光谱的砂糖橘总酸无损检测 [J]. 华中农业大学学报, 2012, 31(4): 518-523.
- [16] GUO Y, NI Y N, KOKOT S. Evaluation of Chemical Components and Properties of the Jujube Fruit Using near Infrared Spectroscopy and Chemometrics [J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2016, 153: 79-86.
- [17] LOUW E D, THERON K I. Robust Prediction Models for Quality Parameters in Japanese Plums (*Prunus salicina* L.) Using NIR Spectroscopy [J]. Postharvest Biology and Technology, 2010, 58(3): 176-184.

- [18] 夏俊芳, 李培武, 李小昱, 等. 不同预处理对近红外光谱检测脐橙  $V_C$  含量的影响 [J]. 农业机械学报, 2007, 38(6): 107-111.
- [19] 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.
- [20] 陈 辰, 鲁晓翔, 张 鹏, 等. 红提葡萄  $V_C$  含量的可见/近红外检测模型 [J]. 食品与机械, 2015, 31(5): 70-74.
- [21] 李 雪, 刘剑飞, 张 玲, 等. 伏淡季水水果实内在品质综合评价 [J]. 西南大学学报(自然科学版), 2019, 41(12): 24-32.
- [22] 游庭活, 温 露, 刘 凡, 等. 不同干燥方式及温度条件对桑椹干营养保健品质的影响 [J]. 蚕业科学, 2015, 41(5): 921-927.
- [23] 詹白勺, 倪君辉, 李 军. 高光谱技术结合 CARS 算法的库尔勒香梨可溶性固形物定量测定 [J]. 光谱学与光谱分析, 2014, 34(10): 2752-2757.
- [24] 牛晓颖, 周玉宏, 邵利敏. 基于 LS-SVM 的草莓固酸比和可滴定酸近红外光谱定量模型 [J]. 农业工程学报, 2013, 29(25): 270-274.
- [25] 黄 康, 汪辉君, 徐惠荣, 等. 基于最小二乘支持向量机的番茄汁糖酸度分析研究 [J]. 光谱学与光谱分析, 2009, 29(4): 931-934.

## Determination of the Intrinsic Quality of Citrus Variety W • Murcott by Near Infrared Spectroscopy

LI Xun-lan, YIN Xu-min, YANG Lei,  
HONG Lin, HAN Guo-hui, WANG Wu

*Chongqing Academy of Agricultural Sciences, Chongqing 401329, China*

**Abstract:** The feasibility of fast and nondestructive measurement of main internal quality indexes of the fruit of the citrus variety W • Murcott was investigated using near infrared spectroscopy in this study. Near infrared spectral data of W • murcott fruit were acquired in the spectral region of 1 000–2 500 nm. The interval partial least squares (iPLS) regression, competitive adaptive reweighting algorithm (CARS) and a hybrid variable selection algorithm based on iPLS and successive projections algorithm (iPLS-SPA) were applied for the effective spectral variable selection. Least squares support vector machine regression (LS-SVR) and partial least squares regression (PLSR) were applied to establish a prediction model of the content of soluble solids (TSS), titratable acid (TA) and vitamin C ( $V_C$ ) content in W • murcott fruit. Spectral data of full-spectrum and the effective spectral variable filtered by iPLS, CARS and iPLS-SPA were used in modeling, respectively. The results showed that the LS-SVR prediction model established with the effective spectral variable filtered by CARS best successful, the correlation coefficients of TSS and TA and  $V_C$  being 0.91, 0.85 and 0.91, respectively, and the corresponding *RMSEP* (Root-Mean-Square Error of Prediction) being 0.26, 0.03 and 0.25 respectively. The above results indicated that the simultaneous detection of TSS, TA and  $V_C$  contents in WIL-1 • murcott fruit could be achieved by using near-infrared spectroscopy combined with CARS and LS-SVR.

**Key words:** near infrared spectroscopy; W • Murcott; intrinsic quality; content detection