

刘燕德, 胡宣, 朱明旺, 等. 近红外在线检测装置参数对苹果糖度模型适用性的影响 [J]. 华南农业大学学报, 2022, 43(5): 108-114.
LIU Yande, HU Xuan, ZHU Mingwang, et al. Influence of near-infrared on-line detection device parameters on the applicability of apple soluble solid content model[J]. Journal of South China Agricultural University, 2022, 43(5): 108-114.

近红外在线检测装置参数对苹果糖度模型适用性的影响

刘燕德^{1,2,3}, 胡 宣^{1,2,3}, 朱明旺^{1,3}, 姚金良¹, 李 斌^{1,3}, 廖 军^{1,3}, 张剑一¹, 景寒松¹

(1 华东交通大学 智能机电装备创新研究院, 江西 南昌 330013; 2 江西省光电检测工程技术中心, 江西 南昌 330013; 3 华东交通大学 机电与车辆工程学院, 江西 南昌 330013)

摘要:【目的】近红外(Near-infrared)光谱在线检测装置的检测速度和积分时间等因素会影响所建立苹果糖度模型的性能, 本文旨在分析检测速度和积分时间对模型适用性的影响, 以提高在线检测的精度。【方法】近红外光谱在线检测装置检测速度和积分时间分别设置为 0.3 m/s+100 ms、0.5 m/s+70 ms、0.5 m/s+100 ms、0.5 m/s+120 ms、0.5 m/s+150 ms, 共 5 个实验组, 试验所用苹果样本共 180 个, 在 350~1 150 nm 波长下采集 5 个试验组苹果的近红外光谱, 应用偏最小二乘法(Partial least squares, PLS)建立苹果可溶性固形物含量(Soluble solid content, SSC)的预测模型。【结果】近红外在线检测装置积分时间对苹果糖度的检测存在阈值, 当积分时间低于 70 ms 时, 模型预测性能较差。预测模型中建模集与预测集的检测速度与积分时间相同时的预测效果优于二者不同时的。检测速度和积分时间会影响在线检测的精度, 不同检测速度和积分时间下, 光线在苹果内部的传输路线不同, 会导致光纤探头获得的内部信息有所差异, 使预测性能变差。在 0.3 m/s+100 ms、0.5 m/s+100 ms、0.5 m/s+120 ms 和 0.5 m/s+150 ms 4 个试验组中使用 Kennard-Stone 算法挑选出 135 个具有代表性的样本光谱, 建立了混合检测速度和积分时间的预测模型, 其预测集相关系数(R_p)均在 0.85 以上, 预测集均方根误差(RMSEP)均低于 0.65。【结论】本研究建立的混合检测速度和积分时间的预测模型可对苹果的糖度达到更好的预测, 满足不同检测装置参数下苹果糖度在线检测的要求。

关键词: 近红外光谱; 检测速度; 积分时间; 阈值; 混合模型; 苹果糖度

中图分类号: S661.1; O657.3

文献标志码: A

文章编号: 1001-411X(2022)05-0108-07

Influence of near-infrared on-line detection device parameters on the applicability of apple soluble solid content model

LIU Yande^{1,2,3}, HU Xuan^{1,2,3}, ZHU Mingwang^{1,3}, YAO Jinliang¹, LI Bin^{1,3}, LIAO Jun^{1,3}, ZHANG Jianyi¹, JING Hansong¹

(1 Intelligent Electromechanical Equipment Innovation Research Institute, East China Jiaotong University, Nanchang 330013, China; 2 Jiangxi Province Optoelectronic Detection Engineering Technology Center, Nanchang 330013, China; 3 School of Mechatronics and Vehicle Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: 【Objective】The performance of the apple soluble solid content model was influenced by the detection speed and integration time of the near-infrared (NIR) spectroscopy on-line detection device. The aim of this study was to analyze the influence of detection speed and integration time on the applicability of the

收稿日期: 2021-12-15 网络首发时间: 2022-06-06 15:08:36

网络首发地址: <https://kns.cnki.net/kcms/detail/44.1110.S.20220602.2325.006.html>

作者简介: 刘燕德, 教授, 博士, 主要从事近红外无损检测研究, E-mail: jxliuyd@163.com

基金项目: 江西省教育厅研究生创新资金项目 (YC 2020-s336); 国家自然科学基金 (31760344)

model, and improve the accuracy of on-line detection. 【Method】 The detection speed and integration time of the on-line detection device of NIR spectroscopy were set as 0.3 m/s and 100 ms, 0.5 m/s and 70 ms, 0.5 m/s and 100 ms, 0.5 m/s and 120 ms, 0.5 m/s and 150 ms respectively. A total of 180 apple samples were used for the experiment, the NIR spectra of five experimental groups of apples were collected at 350~1 150 nm, and the partial least squares (PLS) method was applied to establish the prediction model of apple soluble solid content (SSC). 【Result】 There was a threshold for the integration time. When the integration time was shorter than 70 ms, the prediction performance of the model was poor. The prediction performance of the prediction model with the same detection speed and integration time for modeling set and prediction set was superior to that of the prediction model with different detection speed and integration time. Detection speed and integration time would affect the accuracy of on-line detection. The different transmission routes of light inside the apple at different detection speeds and integration time could lead to differences in the internal information obtained by the fiber optic probe, making the prediction performance worse. The 135 representative sample spectra were selected from four groups of 0.3 m/s and 100 ms, 0.5 m/s and 100 ms, 0.5 m/s and 120 ms, 0.5 m/s and 150 ms using the Kennard-Stone algorithm, the prediction models of the mixed detection speed and integration time were established, with the correlation coefficients (R_p) of the prediction set all above 0.85, the root mean square errors (RMSEP) all below 0.65. 【Conclusion】 The established prediction model of mixed detection speed and integration time can better predict apple sugar content and meet the requirements of apple sugar content on-line detection under different parameters of detection devices.

Key words: Near-infrared spectra; Detection speed; Integration time; Threshold; Mixed model; Apple soluble solid content

我国苹果产量及种植面积在全球占比超 50%，但是由于采后自动检测和分级技术不足，导致我国优质果率不足，苹果出口量仅占总产量的 10%，亟需水果品质在线检测装置提升我国苹果在国际市场的竞争力。近红外光谱在线检测装置作为一种快速、无损和绿色的检测装备，已应用于检测苹果、草莓、柑橘、梨、西瓜等水果的内部品质^[1-4]。近红外光谱在线检测装置的检测速度和积分时间等因素会影响所建立的糖度模型的性能，因此，建立检测速度和积分时间的通用模型非常必要。

学者应用近红外光谱技术对水果内部品质做了很多研究：郭志明等^[5]利用近红外透射法建立了苹果腐心病的在线检测系统，其建立的预测模型的相关系数为 0.92；Liu 等^[6]建立了脐橙可溶性固形物含量 (Soluble solid content, SSC) 的近红外漫反射在线检测模型，其预测相关系数为 0.90；李龙等^[7]利用近红外光谱技术搭建了苹果在线无损检测装备，并对苹果 SSC 建立了预测模型，其相关系数达到 0.949，预测集均方根误差为 0.449；韩东海等^[8]和刘新鑫等^[9]使用近红外透射光谱结合波段筛选方法对苹果的 2 种病害进行判别，其判别模型准确率达到 95.7%；Xu 等^[10]研究比较了单点和双点检测对苹果 SSC 在线检测精度的影响，

使用双分支光纤的系统证明了卓越的鲁棒性，而使用单分支光纤的系统证明了卓越的准确性，其预测集决定系数达到 0.63。以上学者进行的研究都未考虑在线检测速度和积分时间对模型的影响，所建立的模型的性能较低。崔丰娟等^[11]研究了近红外光谱在线采集装置的移动速度对所建立的模型性能的影响，基于 2 种运动速度建立了混合运动速度模型，其预测集相关系数为 0.94，均方根误差为 0.289；Sun 等^[12]利用近红外动态在线检测装备，研究了装置不同运动速度对建立的翠冠梨糖度预测模型的影响，结果发现，当运动速度为 0.5 m/s 时，所建立的糖度模型的性能较好。以上学者探究了检测速度的因素，但是没有涉及对积分时间的研究。本文应用近红外光谱在线检测系统采集不同检测速度和积分时间的苹果光谱数据，建立不同检测速度和积分时间的苹果糖度通用模型。

1 材料与方法

1.1 试验材料

试验使用的苹果样品品种为‘红富士’，共计 180 个，订购于山东烟台某果园，苹果直径为 75~85 mm，到货后，用湿巾将苹果表面灰尘擦除，置于环

境温度为 25 ℃ 的室内保存 24 h 后采集苹果的试验光谱, 经 Kennard-Stone(K-S) 算法分类后, 其中, 建模集样本光谱 135 个, 用于建立 SSC 模型, 预测集样本光谱 45 个, 用于对建立的模型性能进行评判。

1.2 试验装置与光谱采集

本文采用的近红外光谱采集装置是由本研究前期自主研发的动态在线漫透射检测装置^[13], 如图 1 所示, 光源为 2 排卤素灯, 1 排 5 个, 共计 10 个, 卤素灯的参数为 12 V、100 W, 为漫透射方式采集光谱信息提供光源。苹果放在果杯上, 经链条传输至暗箱中, 由卤素灯照亮经过的苹果, 透过苹果内部的光被光纤接收并通过光谱仪传输至电脑, 光谱仪的波长为 350~1 150 nm, 通过配套的光谱采集软件调节样品的积分时间。光谱采集前装置预热 30 min, 通过转动检测速度旋钮调节果杯的移动速度, 从而调节苹果的检测速度。试验分为 5 组, 5 组的检测速度和积分时间分别为: 第 1 组 0.3 m/s、100 ms, 第 2 组 0.5 m/s、70 ms, 第 3 组 0.5 m/s、100 ms, 第 4 组 0.5 m/s、120 ms, 第 5 组 0.5 m/s、150 ms。每个样品在赤道部位采集 4 次光谱, 分布在 4 个面上, 取其平均光谱作为该样品的试验光谱。

1.3 SSC 测定

苹果样品的 SSC 采用折射式数字糖度仪 (PR-101a, 日本) 测量, 用水果刀切下 4 个面光谱采集部位的部分果肉, 将果肉挤出果汁, 滴在糖度仪的测量位置, 测量苹果此面的糖度值, 取 4 个面的平均

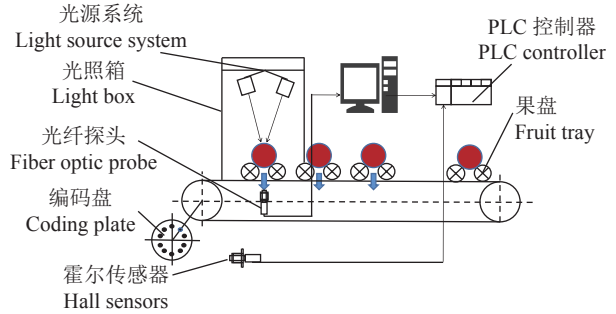


图 1 近红外漫透射在线检测装置

Fig. 1 Device of near-infrared diffuse transmittance on-line detection

糖度值作为该苹果样品的糖度值。

1.4 数据处理

首先应用 K-S 算法将采集到的苹果样本进行分类, 使用 Unscrambler 软件导入采集到的光谱数据, 建立苹果的 SSC 模型, 模型的性能采用预测集相关系数 (R_p) 和均方根误差值 ($RMSEP$) 进行评判。

2 结果与分析

2.1 苹果 SSC 测量结果分析

对 180 个苹果样品分别使用 K-S 算法进行建模集与预测集的分类, 其中建模集 135 个, 预测集 45 个, 苹果的 SSC 测量值如表 1 所示。建模集的 SSC 含量范围较广, 为 9.05~16.40°Brix, 大于预测集的 SSC 含量范围, 9.65~14.85°Brix, 可以对苹果糖度模型取得较好的预测效果。

表 1 苹果 SSC 统计结果
Table 1 Statistical results of apple SSC

数据集 Data set	样品数量 Sample quantity	最小值/°Brix Minimum	最大值/°Brix Maximum	平均值/°Brix Mean	标准偏差/°Brix Standard deviation
建模集 Modeling set	135	9.05	16.40	12.78	1.24
预测集 Prediction set	45	9.65	14.85	12.73	1.15
全部 All	180	9.05	16.40	12.77	1.22

2.2 不同检测速度和积分时间下光谱特性分析

2.2.1 不同检测速度和积分时间的光谱对比 取 180 个样品中具有代表性的苹果样本, 比较 2 种检测速度的光谱, 如图 2 所示。0.3 m/s 采集到的光谱能量较强; 2 种检测速度采集到的光谱波形基本相同, 波峰位于 640、710、和 800 nm 处, 波谷位于 675、755 nm 处。在 710 nm 处的波峰主要与 C—H 和 O—H 键的倍频伸缩振动有关^[14-15], 在 800 nm 处的波峰主要与 C—H 键及 N—H 键的二级倍频吸收有关^[16-17]。0.3 m/s 检测速度下样品的能量光谱高于 0.5 m/s, 造成此现象的原因是 0.3 m/s 的检测速度下, 样品的曝光时间长, 透过样品的光

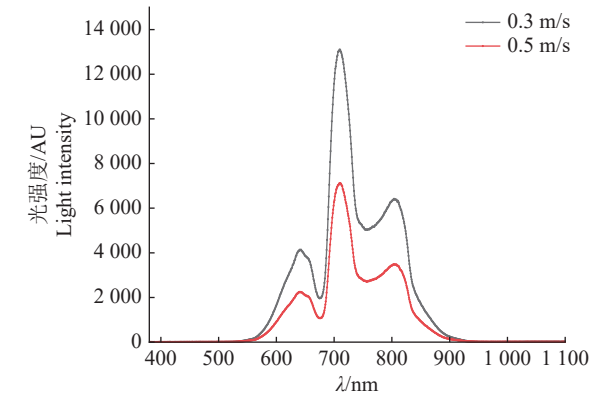


图 2 不同检测速度下的苹果近红外光谱

Fig. 2 Near-infrared spectra of apples at different detection speeds

能量较多, 探测器接受的能量光谱值较高。

取与图 2 相同的苹果样品, 在 4 种积分时间下的光谱如图 3 所示。整体光谱能量强弱关系是 $150\text{ ms} > 120\text{ ms} > 100\text{ ms} > 70\text{ ms}$, 4 种积分时间所采集到的光谱波形基本相同, 在 640 、 710 、 800 nm 附近存在波峰, 波谷位于 675 、 755 nm 附近, 积分时间越长, 光谱的能量值越高。

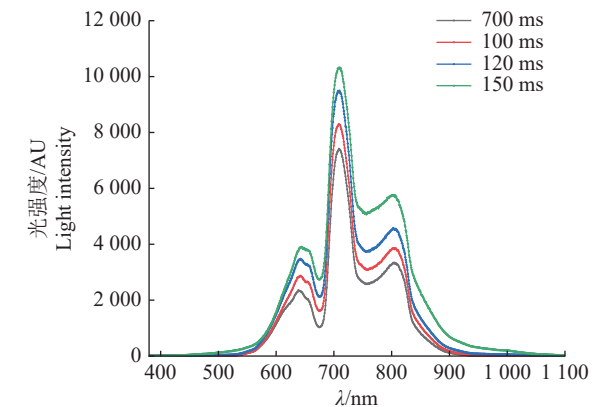


图 3 不同积分时间下的苹果近红外光谱

Fig. 3 Near-infrared spectra of apples with different integration time

2.2.2 不同检测速度和积分时间的光路分析 光线在苹果内部传输路径如图 4 所示。在 0.3 m/s 检测速度下, 样品在光源下的曝光时间长, 获得充足光线的同时也带入许多杂散光, 导致下方光纤接收到的信息含有许多无用信息; 在 0.5 m/s 检测速度下, 传输速度加大导致携带苹果内部品质信息的光线传输到果杯壁上, 没有被果杯下的光纤探头接收。对于不同的积分时间, 短积分时间导致较少的光线进入苹果内部, 光纤探头获取的光谱信息量不足, 从而导致获取的苹果内部品质信息不足; 长积分时间时进入苹果内部的光线较多, 同时光纤探头接收的苹果内部品质信息已经饱和, 更多地获取了

杂散光的信息^[18]。以上因素都会导致所建立的苹果 SSC 预测模型性能变差。

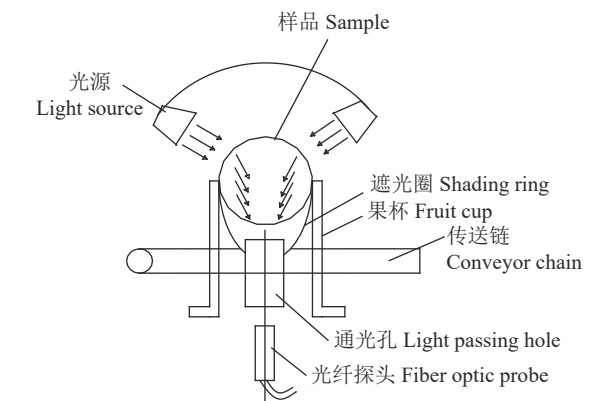


图 4 光线在苹果内部传输路径图

Fig. 4 Light transmission path in apple

2.3 单一检测速度和积分时间建立模型的适用性

2.3.1 检测速度与积分时间相同时建模集与预测集 使用偏最小二乘法 (Partial least squares, PLS) 建立苹果糖度预测模型, 为防止模型过拟合或欠拟合, 潜变量的数量 (Latent variables, LVs) 设定为 $1\sim 20$ 。5 组检测速度和积分时间建立的苹果糖度 PLS 模型结果如表 2 所示, 建模集与预测集检测速度与积分时间相同时苹果 SSC 散点图如图 5 所示。结果表明: 积分时间在 $100\sim 150\text{ ms}$ 时, 建模集和预测集模型的相关系数 (R_C 、 R_P) 都在 0.85 以上, 且预测集均方根误差 ($RMSEP$) 较小, 模型精度较高; 积分时间为 70 ms 时, 建立的模型精度较低, 积分时间对苹果糖度模型存在阈值, 下面只研究 $100\sim 150\text{ ms}$ 积分时间提高模型精度的方法。

2.3.2 检测速度与积分时间不同时建模集与预测集 由表 2 可以看出, 当检测速度为 0.5 m/s 、积分时间为 100 ms 时, 建立的苹果 SSC 模型性能最

表 2 5 组检测速度和积分时间建模结果¹⁾

Table 2 Modeling results of five groups of detection speed and integration time							
组别 Group	检测速度/($\text{m}\cdot\text{s}^{-1}$) Detection speed	积分时间/ms Integration time	LVs	R_C	$RMSE_C$	R_P	$RMSE_P$
1	0.3	100	10	0.945	0.405	0.886	0.549
2	0.5	70	8	0.760	0.806	0.749	0.772
3	0.5	100	9	0.921	0.482	0.921	0.451
4	0.5	120	9	0.876	0.598	0.898	0.506
5	0.5	150	9	0.925	0.473	0.894	0.516

1) LVs: 潜变量的数量; R_C 、 R_P : 建模集与预测集的相关系数; $RMSE_C$ 、 $RMSE_P$: 建模集与预测集的均方根误差

1) LVs: Latent variables; R_C 、 R_P : Correlation coefficient of modeling set and prediction set; $RMSE_C$ 、 $RMSE_P$: Root mean square error of modeling set and prediction set respectively

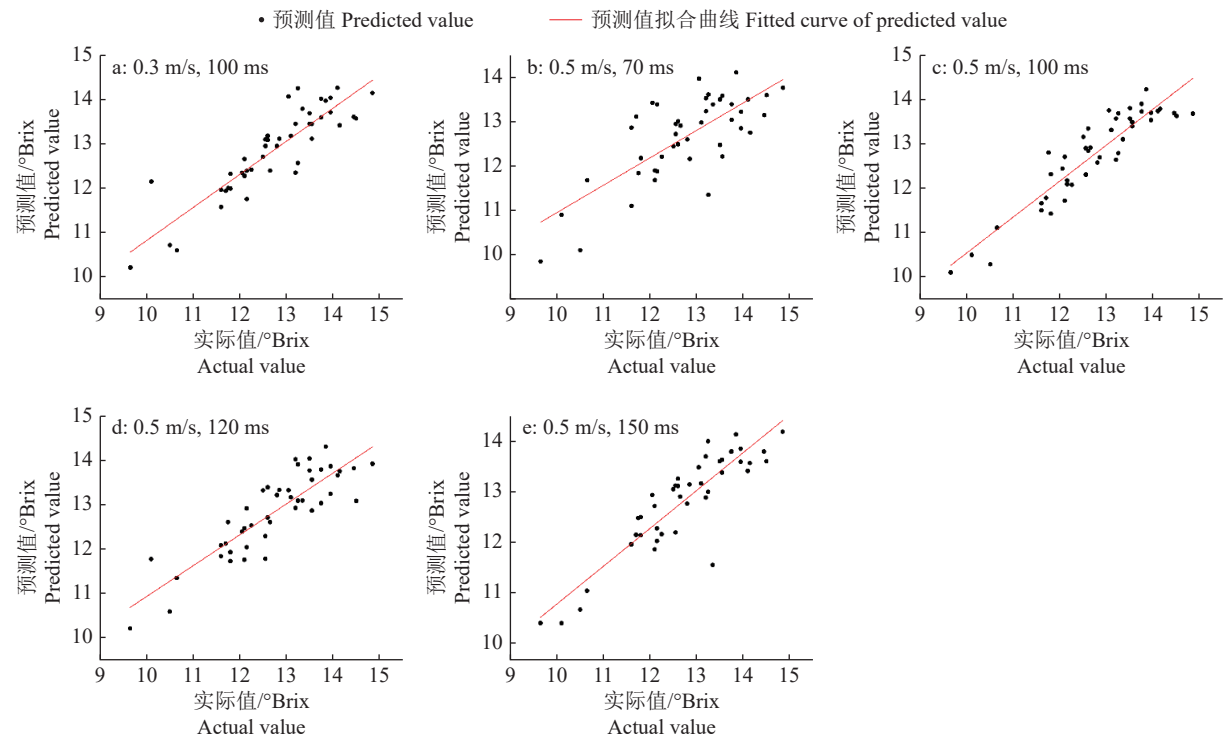


图 5 建模集与预测集的检测速度与积分时间相同时苹果 SSC 散点图

Fig. 5 Scatter diagram of apple SSC under the same detection speed and integration time of modeling set and prediction set

好, R_p 为 0.921, $RMSEP$ 为 0.451。选取检测速度为 0.5 m/s、积分时间为 100 ms 组中的建模集作为建模集与预测集检测速度与积分时间不同时模型的建模集, 分别对其他组别的预测集进行预测, 其结果如表 3 所示, 其预测模型散点图如图 6 所示。结果表明: 当建模集与预测集检测速度与积分时间不同时, $RMSEP$ 较建模集与预测集检测速度与积分时间相同时大, 预测效果较差。在实际的水果在线检测线上, 需建立在一定检测速度和积分时间范围内均适用的模型, 因此下面将研究如何提高苹果糖度模型在不同检测速度和积分时间下的适用性。

表 3 检测速度 0.5 m/s、积分时间 100 ms 时预测其他检测速度和积分时间模型的结果¹⁾

Table 3 Prediction results of other detection speeds and integration time models with detection speed of 0.5 m/s and integration time of 100 ms

检测速度/(m·s ⁻¹)	积分时间/ms	LVs	R_p	$RMSEP$
Detection speed	Integration time			
0.3	100	13	0.683	0.839
0.5	120	10	0.877	0.570
0.5	150	9	0.772	0.730

1) LVs: 潜变量的数量; R_p : 预测集的相关系数; $RMSEP$: 预测集的均方根误差

1) LVs: Latent variables; R_p : Correlation coefficient of prediction set; $RMSEP$: Root mean square error of prediction set

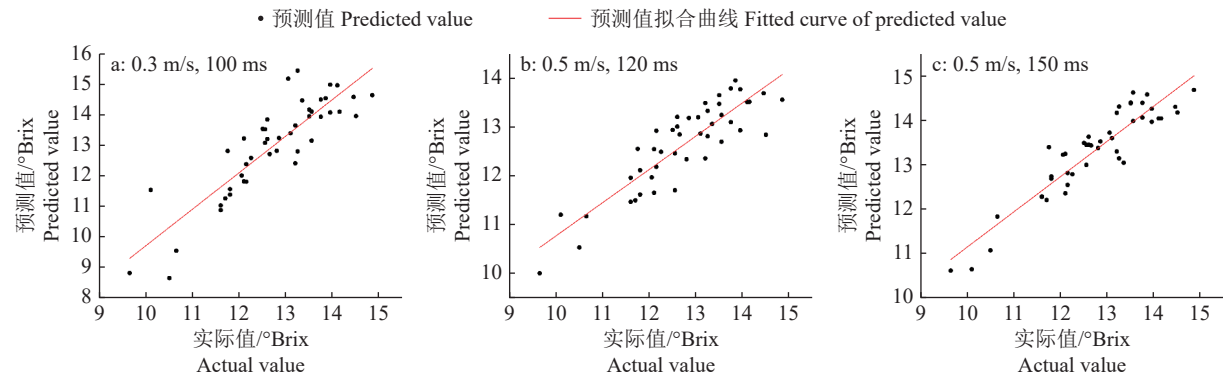


图 6 检测速度 0.5 m/s、积分时间 100 ms 预测其他试验组苹果 SSC 预测值与实际值散点图

Fig. 6 Scatter diagram of predicted value and actual value of apple SSC for other experimental groups predicted by detection speed of 0.5m/s and integration time of 100 ms

2.4 混合检测速度和积分时间建立模型的适用性

在“2.3”4 组试验组中的建模集中, 使用 K-S 算法共选取 135 个具有代表性的样本光谱作为混合检测速度与积分时间预测模型的建模集, 分别对各试验组进行预测, 预测结果如表 4 所示, 其预测模型散点图如图 7 所示。结果表明: 混合检测速度和积分时间建立的预测模型拟合效果较好, 模型性能较优, 相对于单一速度和积分时间作为建模集建立的预测模型, R_p 明显提高, $RMSEP$ 明显减小, 模型稳定性明显提高, 能在一定检测速度和积分时间范围内, 对苹果 SSC 达到更好的预测, 满足不同环境下苹果糖度在线检测的要求。

表 4 混合检测速度和积分时间预测模型预测结果¹⁾
Table 4 Prediction results of prediction model with mixed detection speed and integration time

检测速度/(m·s ⁻¹)	积分时间/ms	LVs	R _p	RMSEP _p
Detection speed	Integration time			
0.3	100	9	0.850	0.633
0.5	100	12	0.919	0.477
0.5	120	12	0.894	0.488
0.5	150	9	0.893	0.501

1) LVs: 潜变量的数量; R_p : 预测集的相关系数; $RMSEP_p$: 预测集的均方根误差

1) LVs: Latent variables; R_p : Correlation coefficient of prediction set; $RMSEP_p$: Root mean square error of prediction set

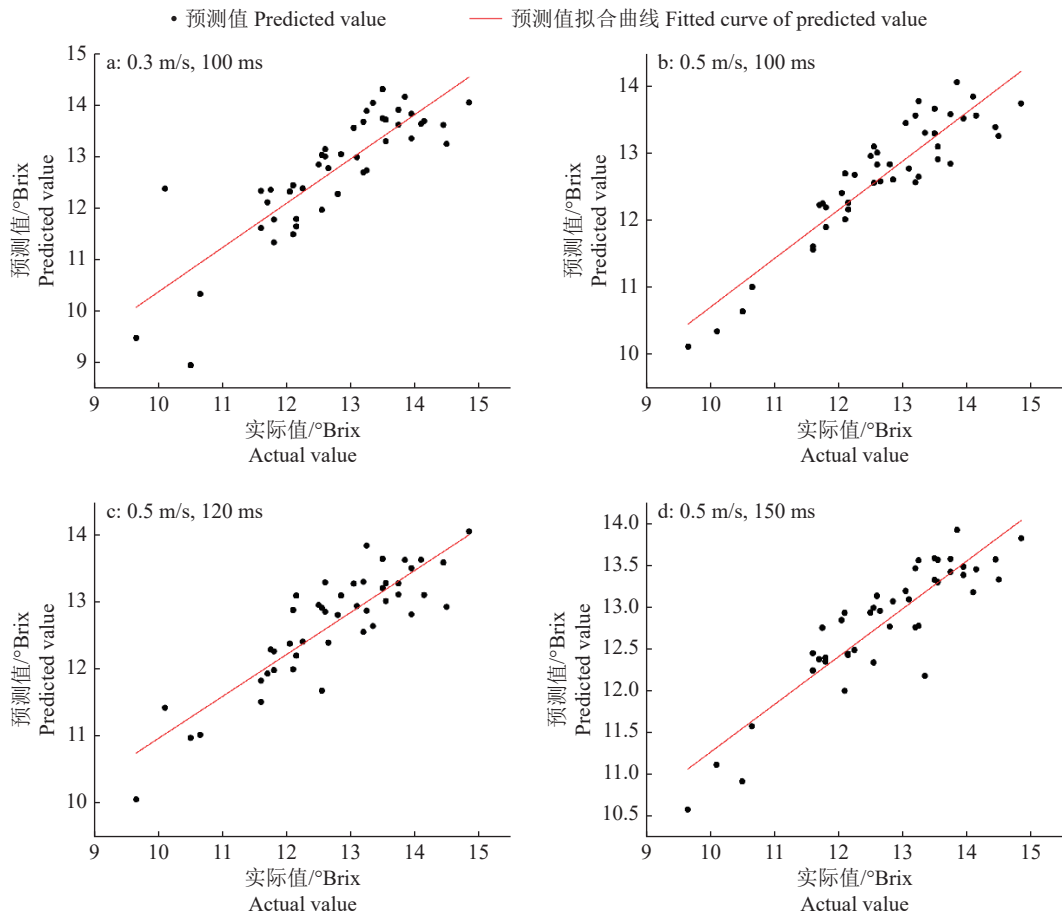


图 7 混合检测速度和积分时间预测模型苹果 SSC 预测值与实际值散点图

Fig. 7 Scatter diagram of predicted value and actual value of apple SSC of prediction model with mixed detection speed and integration time

3 结论

本文研究了不同检测速度和积分时间对苹果 SSC 预测模型的影响, 预测模型中建模集和预测集的检测速度和积分时间相同时的预测效果优于建模集与预测集的检测速度和积分时间不同时。为满足在线检测的需要, 预测模型应在一定的检测速度和积分时间范围内, 保持良好的预测性能, 为此建

立了混合检测速度和积分时间模型; 相对于单一检测速度和积分时间作为建模集建立的预测模型, R_p 由 0.877 提高至 0.919, $RMSEP$ 由 0.570 减小至 0.477, 对苹果 SSC 达到更好的预测效果, 满足了不同环境下苹果糖度在线检测的要求。

参考文献:

[1] XU X, XU H, XIE L, et al. Effect of measurement posi-

tion on prediction of apple soluble solids content (SSC) by an on-line near-infrared (NIR) system[J]. *Journal of Food Measurement and Characterization*, 2019, 13(1): 506-512.

[2] CHEN H, LIU Z, 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.

[3] LEE H, KIM M S, LIM H S. et al. Detection of cucumber green mottle mosaic virus-infected watermelon seeds using a near-infrared (NIR) hyperspectral imaging system: Application to seeds of the “Sambok Honey” cultivar[J]. *Biosystems Engineering*, 2016, 148: 138-147.

[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] 郭志明, 黄文倩, 陈全胜, 等. 苹果腐心病的透射光谱在线检测系统设计及试验[J]. *农业工程学报*, 2016, 32(6): 283-288.

[6] LIU Y, SUN X, OUYANG A. Nondestructive measurement of soluble solid content of navel orange fruit by visible-NIR spectrometric technique with PLSR and PCA-BPNN[J]. *LWT: Food Science and Technology*, 2010, 43(4): 602-607.

[7] 李龙, 彭彦昆, 李永玉, 等. 苹果内外品质在线无损检测分级系统设计与试验[J]. *农业工程学报*, 2018, 34(9): 267-275.

[8] 韩东海, 刘新鑫, 鲁超, 等. 苹果内部褐变的无损检测研究[J]. *农业机械学报*, 2006(6): 86-88.

[9] 刘新鑫, 韩东海, 涂润林, 等. 苹果水心病在贮藏期变化的无损检测[J]. *农业工程学报*, 2004(1): 211-214.

[10] XU X, MO J, XIE L, et al. Influences of detection position and double detection regions on determining soluble solids content (SSC) for apples using on-line visible/near-infrared (Vis/NIR) spectroscopy[J]. *Food Analytical Methods*, 2019, 12(9): 2078-2085.

[11] 崔丰娟, 阚建文. 近红外透射苹果运动速度模型适用性的研究[J]. *农机化研究*, 2010, 32(11): 170-173.

[12] SUN T, LIN H, XU H, et al. Effect of fruit moving speed on predicting soluble solids content of ‘Cuiguan’ pears (*Pomaceae pyrifolia* Nakai cv. Cuiguan) using PLS and LS-SVM regression[J]. *Postharvest Biology and Technology*, 2009, 51(1): 86-90.

[13] 刘燕德, 张宇, 姜小刚, 等. 不同贮藏期水蜜桃硬度及糖度的检测研究[J]. *光谱学与光谱分析*, 2021, 41(1): 243-249.

[14] LEROY B, LAMBOTTE S, DOTREPPE O, et al. Prediction of technological and organoleptic properties of beef *Longissimus thoracis* from near-infrared reflectance and transmission spectra[J]. *Meat Science*, 2004, 66(1): 45-54.

[15] MCDEVITT R M, GAVIN A J, ANDRÉS S, et al. The ability of visible and near-infrared reflectance spectroscopy to predict the chemical composition of ground chicken carcasses and to discriminate between carcasses from different enotypes[J]. *Journal of Near Infrared Spectroscopy*, 2005, 13(3): 109-117.

[16] ZHOU L J, WU H, LI J T, et al. Determination of fatty acids in broiler breast meat by near-infrared reflectance spectroscopy[J]. *Meat Science*, 2012, 90(3): 658-664.

[17] PRIETO N, ROSS D W, NAVAJAS E A, et al. Online prediction of fatty acid profiles in crossbred Limousin and Aberdeen Angus beef cattle using near infrared reflectance spectroscopy[J]. *Animal*, 2011, 5(1): 155-165.

[18] 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.

【责任编辑 李庆玲】