

姜小刚, 朱明旺, 姚金良, 等. 基于近红外光谱技术苹果尺寸差异对糖度模型适用性的影响[J]. 华中农业大学学报, 2024, 43(1): 242-248.
DOI: 10.13300/j.cnki.hnlkxb.2024.01.028

基于近红外光谱技术苹果尺寸差异 对糖度模型适用性的影响

姜小刚^{1,2}, 朱明旺¹, 姚金良¹, 李斌^{1,2}, 廖军¹, 张宇翔³, 刘燕德^{1,2}

1. 华东交通大学智能机电装备创新研究院/机电与车辆工程学院, 南昌 330013;
2. 江西省光电检测工程技术中心, 南昌 330013; 3. 淮阴工学院自动化学院, 淮安 223003

摘要 为消除水果自身尺寸差异对其糖度预测模型的不利影响, 进一步提高水果分选模型精度, 应用近红外光谱在线检测装置采集不同果径苹果的近红外光谱, 对光谱进行多种预处理后, 分别建立苹果可溶性固形物的偏最小二乘法模型, 再用苹果果径75~85 mm组中的建模集预测苹果果径分别为65~75、85~95 mm组中的预测集样品, 最后用果径组65~75、75~85、85~95 mm中的建模集和预测集, 分别作为混合苹果尺寸糖度预测模型的建模集和预测集, 并利用特征光谱选择算法对模型进行简化, 建立苹果糖度通用预测模型。结果显示: 与建模集和预测集果径不同时所建立的苹果糖度预测模型最优组相比, 其相关系数 R_p 由0.805提高至0.943, 预测集均方根误差值RMSEP由0.778减小至0.480, RPD由0.96增加至3.05, 再对建立的通用模型进行简化, 可以降低苹果尺寸对苹果糖度模型的影响, 提高模型预测性能。

关键词 苹果; 近红外光谱; 混合尺寸模型; 尺寸差异; 水果分选; 无损检测

中图分类号 O657.3 **文献标识码** A **文章编号** 1000-2421(2024)01-0242-07

苹果富含膳食纤维及微量元素, 是人们最常食用的水果之一, 其糖度的高低主要取决于可溶性糖的含量, 而可溶性糖含量与可溶性固形物密切相关。苹果果径大小存在的差异会影响所建立的糖度模型的性能, 糖度模型的准确性直接影响水果分选模型的精度, 因此研究苹果果径对其糖度模型的影响是非常有必要的。近年来, 近红外光谱检测技术已大量运用在水果的采后分选^[1-4]。崔丰娟等^[5]研究了近红外在线检测装置运动速度对苹果糖度模型的影响, 结果表明传输速度为0.5 m/s时, 所建立的苹果糖度模型预测性能较好。Xu等^[6]比较了单点和双点检测对苹果可溶性固形物含量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。上述研究均未考虑苹果果径对模型的影响, 建立的模型性能较低。曹松涛等^[7]针对梨的尺寸差异、产地差异因素对其内部品质模

型的影响进行研究, 采取对光谱进行二阶导数处理可以提高模型的预测性能。刘燕德等^[8]建立了不同尺寸脐橙的近红外光谱检测模型, 使用MSC和SNV预处理可以解决脐橙尺寸差异的影响, 提高预测模型的精度。Tian等^[9]建立了近红外光谱不同尺寸苹果核霉变的判别模型, 提出近红外光谱强度与其光程呈指数关系, 并修正了近红外光谱, 修正后的模型预测集判别精度达到90.2%, 该方法可以修正果实大小对透射光谱的影响, 以提高对染病苹果的识别。刘燕德等^[10]使用模型更新加变量筛选方法对5个品种的苹果可溶性固形物含量进行预测, 建立的模型对新品种苹果的预测性能有明显改善, RMSEP为0.587%, R_p 为0.928。Mcglone等^[11]开发了2个原型在线近红外系统, 一种是基于时延积分光谱仪, 另一种是基于大口径光谱仪, 后一种系统具有较高的精度, 经过偏最小二乘校正后, 预测的均方根误差为4.1%, 但实验仅选取赤道平均直径为76 mm (SD =

收稿日期: 2023-04-20

基金项目: 国家自然科学基金项目(31760344); 江西省自然科学基金项目(20171BAB212021); 江西省教育厅科学技术研究项目(GJJ200652; GJJ200615)

姜小刚, E-mail: jxg_ecjtu@163.com

通信作者: 刘燕德, E-mail: jxliuyd@163.com

2.8 mm)的苹果,未研究果实大小对褐变组织检测的影响。研究表明,果径差异会对近红外在线检测装置产生影响。

本研究采集不同尺寸苹果的近红外光谱,建立各种光谱预处理后糖度含量预测模型,结合波段筛选算法,降低苹果尺寸差异对苹果糖度模型的影响,提高不同尺寸苹果糖度预测模型的性能,为进一步提高水果分选模型精度提供参考。

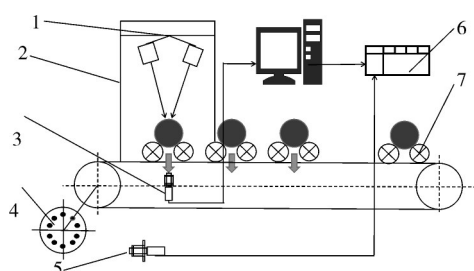
1 材料与方法

1.1 试验材料

本研究采用的苹果品种为红富士,订购于山东某果园,苹果无外部机械损伤和缺陷,共计240个,其中果径为65~75 mm(即 $65\text{ mm} < \text{果径} \leq 75\text{ mm}$)的苹果80个,75~85 mm(即 $75\text{ mm} < \text{果径} \leq 85\text{ mm}$)的苹果80个,85~95 mm(即 $85\text{ mm} < \text{果径} \leq 95\text{ mm}$)的苹果80个,到货后,放置于室温为24℃的密闭房间内保存24 h,并用湿巾将苹果外部的灰尘擦尽,风干后采集近红外光谱。

1.2 试验装置与光谱采集

采用笔者所在实验室自主研发的近红外光谱在线检测装置^[12]。装置如图1所示,光源系统采用2排卤素灯,1排5个,提供近红外光。采集过程为:将苹果水平放置于果盘中,经链条传动将苹果传至光照箱,光照箱为一密闭暗室,可以防止自然光的干扰,当苹果赤道部位正对果盘下的光纤探头时,PLC触发,采集苹果的近红外光谱。光谱仪的波长范围为350~1 150 nm,采集苹果光谱前,装置预热30 min,传动链的运动速度为0.5 m/s,光谱采集的积分时间为100 ms,每个苹果样品均在其赤道位置采集4次光谱,每次将苹果旋转90°,取其平均光谱作为该苹果样本近红外光谱。



1. 光源系统 Near-infrared light; 2. 光照箱 Testing room; 3. 光纤探头 Fiber optic probe; 4. 编码盘 Counters; 5. 霍尔传感器 Hall sensor; 6. PLC控制器 PLC controller; 7. 果盘 Pallets.

图1 近红外漫透射在线检测装置

Fig.1 The device of NIR diffuse transmittance detection

1.3 可溶性固形物含量测定

用水果刀在每个苹果样本赤道处的测量位置切下一块厚度约为20 mm的苹果块,再将切下的苹果块榨汁,用定量滤纸(颗粒截留水平低至2.5 μm)过滤后,用折射式数字糖度仪(PR-101a,日本)测量苹果的可溶性固形物含量(soluble solids content, SSC),取苹果4个赤道面测量位置SSC的平均值,作为该苹果样本的SSC。使用数显游标卡尺(Mitutoyo-500,日本)测量苹果赤道位置的果径,每个苹果在赤道位置测量4次,取平均果径作为该果的果径。

1.4 数据处理

在建立苹果糖度预测模型前,需对苹果样本进行分类,使用Kennard-Stone(K-S)算法,将不同果径组苹果分为建模集与预测集,建立苹果糖度PLS预测模型,采用预测集相关系数 R_p 、预测集均方根误差值(root mean square error of prediction, RMSEP)和相对分析误差(residual predictive deviation, RPD)对模型的预测性能进行评判,其 R_p 越大, RMSEP越小, RPD越大,其模型性能越优^[10]。

1.5 光谱预处理

由于光谱存在噪声,对光谱进行预处理可以提高模型性能,其中多元散射校正(multiple scattering correction, MSC)、标准正态变量变换(standard normal variable transformation, SNV)和卷积平滑(Savitzky-Golay, S-G)是解决苹果果径差异的重要预处理方法^[12]。本研究采用上述3种预处理方法对苹果光谱进行预处理,比较3种预处理方法对解决苹果果径差异造成其糖度模型预测性能差的能力。

2 结果与分析

2.1 苹果SSC测量结果

240个苹果试验样品采用K-S算法分成3种不同果径,并将不同果径苹果样本以3:1的比例划分为建模集和预测集,其中,每个果径组下80个样本,其中建模集60个、预测集20个,苹果的SSC测量值如表1所示。由表1可知,该方法保证了训练集中样本按空间距离均匀分布,使所建立的模型最能反映样本信息^[13-14]。

2.2 不同果径苹果近红外光谱特性分析

取3种果径组苹果的平均光谱,比较3种不同果径的光谱(图2)。随着果径的增大,所对应的苹果光谱能量越小,苹果果径在65~75 mm时采集到的光谱能量最强,苹果果径在85~95 mm时采集到的光

表1 苹果试验样品SSC统计结果

Table 1 The statistical results of soluble solids content of apple

样品 Samples	数据集 Data	最小值/% Minimum	最大值/% Maximum	平均值/% Average	标准偏差/% Standard deviation	
全部 All	80	9.85	16.70	13.84	1.54	
65~75 mm	建模集 Modeling	60	9.85	16.70	13.73	1.60
	预测集 Prediction	20	11.55	16.70	14.15	1.32
	全部 All	80	8.75	16.85	13.44	1.74
75~85 mm	建模集 Modeling	60	8.75	16.85	13.41	1.70
	预测集 Prediction	20	8.85	16.50	13.53	1.84
	全部 All	80	11.50	15.45	13.28	0.83
85~95 mm	建模集 Modeling	60	11.50	15.45	13.32	0.85
	预测集 Prediction	20	11.60	14.80	13.16	0.75

谱能量最弱,3种不同果径苹果所采集到光谱波形基本相同,波峰位于640、710、800 nm处,波谷位于675、755 nm,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805 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]。

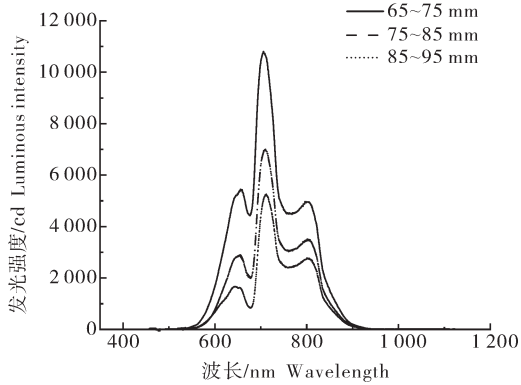


图2 3种果径苹果近红外光谱图

Fig.2 Near-infrared spectra of three fruit sizes of apples

果径为65~75 mm苹果的能量光谱高于果径为75~85 mm和85~95 mm苹果的能量光谱,原因是近红外光在苹果内部随着光程的增加,其携带的能量会发生衰减,在给定的波长下,进入苹果内部的光消光率近似为对数函数^[18],可拟合为: $I = I_0 \exp(-u_c d)$,其中: I 为进入苹果内部的光强, I_0 为苹果下方光纤探头所接收到的光强, u_c 为消光系数, d 为光线进入苹果点到光线出苹果点的距离。在图1的近红外在线检测装置中,上式中的 d 为苹果的果径,随着苹果果径的增大,光线的能量在苹果内部的损失越来越大,导致其采集到的近红外光谱能量值较低,由此猜测,苹果不同的果径会造成其近红外光谱有所差异,从而对由近红外光谱建立的苹果糖度预测模型的性能产生影响。

2.3 3组果径苹果糖度预测模型的建立

使用偏最小二乘回归(PLS)建立苹果糖度预测模型,为防止模型过拟合或欠拟合,LVs的数量设定为1~20。分别采用MSC、SNV和S-G平滑对光谱进行预处理,3组果径苹果建立的糖度PLS模型结果如表2所示。

果径为65~75 mm的苹果近红外光谱经SNV处理后,其糖度模型预测效果最好,其 R_p 为0.922, RMSEP为0.542, RPD值为2.44,其散点图如图3A所示。果径为75~85 mm的苹果近红外光谱经SNV处理后,其糖度模型预测效果最好,其 R_p 为0.986, RMSEP为0.323, RPD值为5.70,其散点图如图3B所示。果径为85~95 mm的苹果近红外光谱经SNV处理后,其糖度模型预测效果最好,其 R_p 为0.923, RMSEP为0.312, RPD值为2.40,其散点图如图3C所示。3种果径下原始光谱建模和各种预处理后建模结果可知:SNV在一定程度上可以解决由于苹果果径差异所造成其糖度预测模型性能差的问题,因为SNV作为一种可以消除样品颗粒大小、表面散射以及光程变化的预处理方法^[19],可以解决样品尺寸不均造成的光谱散射的影响。

2.4 果径75~85 mm苹果预测其余果径组苹果糖度预测模型

由表2知,果径为75~85 mm苹果糖度预测模型的建模集较其他果径组的模型性能好,故选择果径75~85 mm苹果的建模集作为混合预测模型的建模集,探究当建模集中苹果果径与预测集中苹果果径不同时,其建立的苹果糖度预测模型的性能。用果径75~85 mm组别中的建模集,分别预测果径65~75、85~95 mm组别中的预测集。当建模集苹果果径与预测集苹果果径差别较大时,所建立的苹果糖度

表 2 3 组果径分别建立的苹果糖度模型建模结果

Table 2 Modeling results of apple brix modeled for each of the three fruit size groups

果径/mm Fruit diameter	预处理方法 Preprocessing method	LVs	R_c	RMSEC	R_p	RMSEP	RPD
65~75	Raw	11	0.968	0.398	0.866	0.698	1.89
	MSC	10	0.970	0.385	0.896	0.646	2.04
	SNV	11	0.977	0.364	0.922	0.542	2.44
	S-G平滑 Savitzky-Golag smoothing algorithm	12	0.969	0.389	0.871	0.652	2.02
75~85	Raw	10	0.980	0.337	0.979	0.389	4.73
	MSC	11	0.986	0.277	0.983	0.346	3.82
	SNV	13	0.994	0.186	0.986	0.323	5.70
	S-G平滑 Savitzky-Golag smoothing algorithm	11	0.988	0.298	0.983	0.377	4.88
85~95	Raw	9	0.910	0.351	0.904	0.337	2.23
	MSC	10	0.932	0.308	0.918	0.328	2.29
	SNV	10	0.939	0.293	0.923	0.312	2.40
	S-G平滑 Savitzky-Golag smoothing algorithm	9	0.909	0.369	0.901	0.345	2.17

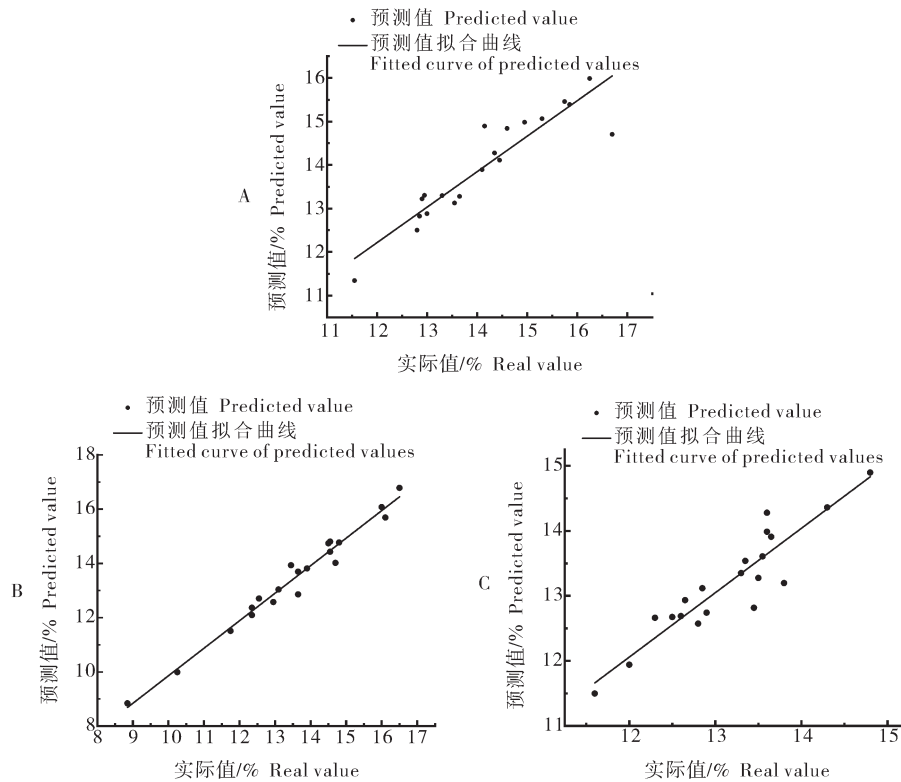


图 3 果实直径 65~75 mm (A)、75~85 mm(B)和 85~95 mm(C)的糖度预测散点图

Fig.3 Predicted scatter diagram of apple brix for fruit diameter 65-75 mm(A), fruit diameter 75-85 mm(B), and fruit diameter 85-95 mm(C)

预测模型预测性能较差,与表 2 单独组果径原始光谱建模结果相比,果径为 65~75 mm 的苹果糖度模型预测集 R_p 由 0.866 下降至 0.768, RMSEP 由 0.698 增加到 0.853, RPD 值由 1.89 下降至 1.55, 其预测模型散点图如图 4A 所示。果径为 85~95 mm 的苹果糖度模型预测集 R_p 由 0.904 下降至 0.805, RMSEP 由 0.337 增加到 0.778, RPD 值由 2.23 下降至 0.96, 其预

测模型散点图如图 4B 所示。对于一个稳定的模型,其 RPD 值要大于 2, 故建模集与预测集苹果果径差别较大时建立的糖度预测模型不可靠, 在实际的水果分选线中, 苹果果径会有很大的差异, 会导致水果分选模型的分选性能差, 故需要建立不同果径苹果糖度预测模型的通用模型, 以增加不同果径苹果糖度的分选性能。

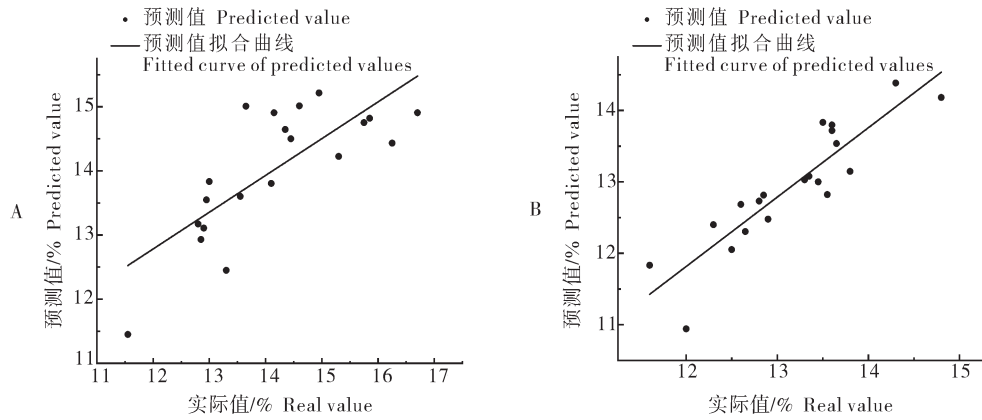


图4 75~85 mm果径预测65~75 mm(A)和85~95 mm(B)果径糖度散点图

Fig.4 Scatter diagram of fruit diameter 75-85 mm predicting fruit diameter 65-75 mm(A) and 85-95 mm(B)

2.5 不同果径糖度通用预测模型

由图4可知,苹果果径尺寸差异会对其近红外光谱产生影响,造成所建立的苹果糖度预测模型的预测性能差,故需建立不同果径苹果糖度通用预测模型,解决苹果果径差异对模型的影响,采用果径组65~75、75~85、85~95 mm中的建模集和预测集,作为混合苹果尺寸糖度预测模型的建模集和预测集,建立不同苹果果径糖度预测模型,其预测模型散点图如图5所示。采用果径75~85 mm组中的建模集,预测果径65~75、85~95 mm组别中的预测集模型,其相关系数 R_p 由0.805提高至0.943, RMSEP由0.778减小至0.480, RPD由0.96增加至3.05,模型稳定性明显提高,可以降低苹果果径对苹果糖度模型的影响。

2.6 通用模型的简化

为满足在线检测中对稳定和快速的需求,需对模型进行简化,使用连续投影算法(successive projection algorithm, SPA)和竞争性自适应重加权算法(competitive adaptive reweighting algorithm, CARS)对不同尺寸糖度通用预测模型光谱变量进行特征提取,2种算法挑选出的特征波长如图6所示,再将挑选出的光谱变量建立PLS模型,模型效果如表3所示。

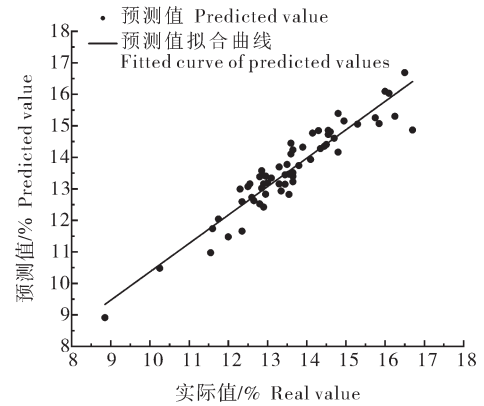


图5 混合果径糖度预测模型散点图

Fig.5 Scatter diagram of brix mixed fruit diameter prediction model

表3 2种波长选择算法的PLS模型效果
Table 3 PLS model of two wavelength selection algorithms

样本选择算法 Algorithm	变量数 Number of variables	R_c	RM-SEC	R_p	RM-SEP	RPD
SPA	23	0.939	0.495	0.915	0.581	2.58
CARS	65	0.964	0.384	0.963	0.391	3.84

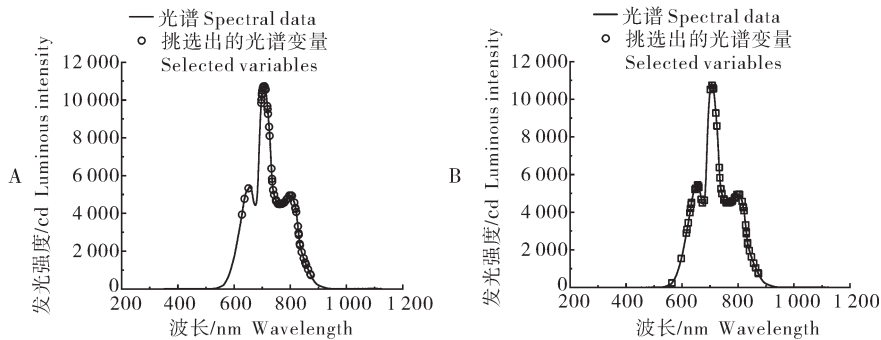


图6 SPA(A)和CARS(B)挑选出的光谱变量

Fig.6 Spectra variables selected by SPA(A) and CARS(B)

由图6可知,SPA和CARS挑选出的特征光谱大多数位于波峰波谷处,说明此处含有大量关于苹果糖度的信息,SPA挑选出23个特征光谱变量,其建立的糖度预测模型预测集相关系数 R_p 为0.915, RMSEP为0.581,相比于采用各果径组中建模集和预测集建立的混合苹果尺寸糖度预测模型,其预测模型性能有所下降,原因是SPA去除了许多与糖度有关的信息,导致其模型预测性能差。CARS挑选出65个特征光谱点,其建立的糖度预测模型预测集相关系数 R_p 为0.963, RMSEP为0.391,相比于采用各果径组中建模集和预测集建立的混合苹果果径糖度预测模型,其糖度预测模型性能有所提升,且模型所用光谱变量由1044个下降至65个,提高了模型运算速度。

3 讨论

为消除水果自身尺寸差异对其糖度预测模型不利影响,进一步提高水果分选模型精度,本研究应用近红外光谱在线检测装置采集不同果径苹果的近红外光谱。对比各组果径苹果光谱发现,苹果尺寸差异会对其光谱产生影响,最终会对其苹果糖度模型的预测性能产生影响,为此,研究了不同尺寸差异的解决方法,分别建立了相同果径组不同预处理糖度预测模型,其光谱经SNV预处理后,建立的糖度模型预测效果最好,其 R_p 为0.922, RMSEP为0.542, RPD值为2.44。随后建立不同果径糖度预测模型,用果径组75~85 mm的建模集分别预测65~75 mm和85~95 mm预测集,其建立的糖度预测模型性能较差。

为解决苹果尺寸差异造成的苹果糖度预测模型预测性能差的问题,建立了混合果径模型,其相关系数 R_p 由0.805提高至0.943, RMSEP由0.778减小至0.480, RPD由0.96增加至3.05。为了满足在线检测对稳定和快速的需求,对模型进行简化,采用CARS挑选出65个特征光谱点,建立的糖度预测模型。预测集相关系数 R_p 为0.963, RMSEP为0.391,预测模型性能有所提升,模型所用光谱变量由1044个下降至65个,模型稳定性明显提高,可以减小苹果尺寸差异对苹果糖度模型的影响,提高模型的预测性能。

本研究结果表明,不同果径苹果糖度通用预测模型与波段筛选算法结合的策略能够有效应用于无损检测中不同苹果果径糖度通用模型的建立,具有一定的可行性。后续研究可以将该策略应用于其他水果或其他内部品质预测通用模型的建立,并通过改进预处理、波段筛选和建模方法提升模型性能。

参考文献 References

- [1] TIAN X, LI J B, WANG Q Y, et al. A bi-layer model for nondestructive prediction of soluble solids content in apple based on reflectance spectra and peel pigments [J]. *Food chemistry*, 2018, 239: 1055-1063.
- [2] CUBERO S, ALEIXOS N, MOLTÓ E, et al. Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables [J]. *Food and bioprocess technology*, 2011, 4(4): 487-504.
- [3] NICOLAÏ B M, BEULLENS K, BOBELYN E, et al. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review [J]. *Postharvest biology and technology*, 2007, 46(2): 99-118.
- [4] 黎丽莎, 刘燕德, 胡军, 等. 近红外无损检测技术在水果成熟度判别中的应用研究 [J]. *华东交通大学学报*, 2021, 38(6): 95-105. LI L S, LIU Y D, HU J, et al. Application of near infrared nondestructive testing technology in fruit maturity discrimination [J]. *Journal of East China Jiaotong University*, 2021, 38(6): 95-105 (in Chinese with English abstract).
- [5] 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. CUI F J, ZHA J W. Study of the applicability of the apple speed model by near-infrared transmission [J]. *Journal of agricultural mechanization research*, 2010, 32(11): 170-173 (in Chinese with English abstract).
- [6] XU X, MO J C, XIE L J, 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.
- [7] 曹松涛. 样品相关因素对梨糖度可见/近红外光谱检测影响的研究 [D]. 杭州: 浙江大学, 2017. CAO S T. Effects of sample related factors on detection of soluble solids content in pear by vis/NIR spectroscopy [D]. Hangzhou: Zhejiang University, 2017 (in Chinese with English abstract).
- [8] 刘燕德, 饶宇, 孙旭东, 等. 尺寸差异对脐橙糖度可见近红外光谱检测模型影响 [J]. *光谱学与光谱分析*, 2020, 40(10): 3241-3246. LIU Y D, RAO Y, SUN X D, et al. Size effect on the near-infrared spectroscopy detection model of navel orange [J]. *Spectroscopy and spectral analysis*, 2020, 40(10): 3241-3246 (in Chinese with English abstract).
- [9] TIAN S J, ZHANG J H, ZHANG Z X, et al. Effective modification through transmission vis/NIR spectra affected by fruit size to improve the prediction of moldy apple core [J]. *Infrared physics & technology*, 2019, 100: 117-124.
- [10] 刘燕德, 黎丽莎, 李斌, 等. 多品种苹果可溶性固形物近红外无损检测通用模型研究 [J]. *华中农业大学学报*, 2022, 41(2): 237-244. LIU Y D, LI L S, LI B, et al. General near-infrared model of soluble solids content in multi-variety apples [J]. *Journal of Huazhong Agricultural University*, 2022, 41(2): 237-244 (in Chinese with English abstract).
- [11] MCGLONE V A, MARTINSEN P J, CLARK C J, et al. Online detection of brownheart in Braeburn apples using near infrared transmission measurements [J]. *Postharvest biology and tech-*

- nology, 2005, 37(2): 142-151.
- [12] 刘燕德, 张雨, 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. LIU Y D, ZHANG Y, JIANG X G, et al. Detection on firmness and soluble solid content of peach during different storage days[J]. Spectroscopy and spectral analysis, 2021, 41(1): 243-249 (in Chinese with English abstract).
- [13] WU W, WALCZAK B, MASSART D L, et al. Artificial neural networks in classification of NIR spectral data: design of the training set [J]. Chemometrics and intelligent laboratory systems, 1996, 33(1): 35-46.
- [14] WILLIAMS P. Near-infrared technology in the agricultural and food industries [J]. American association of cereal chemists, 1987, 10(2): 15-28.
- [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 genotypes [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] FRASER D G. Near infra-red (NIR) light penetration into an apple [J]. Postharvest biology & technology, 2001, 2(3): 8-19.
- [19] 李杰, 李尚科, 蒋立文, 等. 基于近红外光谱技术与化学计量学的绿茶无损鉴别方法研究[J]. 分析测试学报, 2020, 39(11): 1344-1350. LI J, LI S K, JIANG L W, et al. A nondestructive method identifying varieties of green tea based on near infrared spectroscopy and chemometrics [J]. Journal of instrumental analysis, 2020, 39(11): 1344-1350 (in Chinese with English abstract).

Effects of apple size on applicability of model for predicting content of sugar based on near infrared spectroscopy

JIANG Xiaogang^{1,2}, ZHU Mingwang¹, YAO Jinliang¹, LI Bin^{1,2},
LIAO Jun¹, ZHANG Yuxiang³, LIU Yande^{1,2}

1. Institute of Intelligent Mechanical and Electrical Equipment Innovation/School of Mechanical and Electrical and Vehicle Engineering, East China Jiaotong University, Nanchang 330013, China;
2. Jiangxi Photoelectric Detection Engineering Technology Center, Nanchang 330013, China;
3. Automation College of Huaiyin Institute of Technology, Huaian 223003, China

Abstract The size difference of the fruit itself results in poor robustness and low accuracy of the model for predicting the content of sugar. Eliminating the influence of fruit size differences is of great significance to improve the accuracy of fruit sorting models. The NIR spectra of apples with different fruit diameters were collected by an online NIR spectroscopy detection device, and the partial least squares (PLS) models of the content of sugar in apple (SSC) were established after various pre-processing of the spectra. The modeling sets in the diameter of apple fruit group with 75-85 mm were used to predict the prediction set samples in the diameter of apple fruit group with 65-75 mm and 85-95 mm, respectively. The modeling and prediction sets in the diameter of apple fruit group with 65-75 mm, 75-85 mm, and 85-95 mm were used as the modeling and prediction sets of the mixed size model for predicting the content of sugar in apple, respectively. The model was simplified by using the feature spectral selection algorithm. The correlation coefficient R_p was increased from 0.805 to 0.943, the root mean square error value RMSEP was reduced from 0.778 to 0.480, and the RPD was increased from 0.96 to 3.05 compared with the optimal set of model for predicting the content of sugar in apple established when the modeling set and the prediction set had different fruit diameters. It is indicated that simplifying the general model established can reduce the effects of apple size on the model for predicting the content of sugar in apple and improve the prediction performance of the model.

Keywords apple; near-infrared spectroscopy; mixed-size model; size difference; fruit grading; non-destructive testing

(责任编辑:陆文昌)