

基于可见-近红外光谱技术的苹果糖度光照位置优化研究

刘燕德, 崔惠桢, 李 斌, 王观田, 徐 振, 李茂鹏

华东交通大学机电与车辆工程学院, 江西 南昌 330013

摘 要 在利用可见-近红外漫透射光谱技术对苹果的可溶性固形物(SSC)检测时,由于卤素灯光照射在苹果上的位置不同,采集到的苹果光谱中所包含的可溶性固形物信息不同,导致模型得出的结果不同;找到一个最好的苹果光照位置有利于得到最佳的可溶性固形物评价模型。利用多模式可调节的光学结构在相同的实验环境和实验条件下采集了购买于同一水果批发商的尺寸相近但照射位置不同的两批苹果的近红外漫透射光谱,探索苹果可溶性固形物模型建立过程中最佳的照射位置从而得到最佳位置的可溶性固形物评价模型。通过对样品进行光谱采集、糖度真值采集并结合化学计量学方法得出最佳的建模位置,照射位置为上部且光谱没有预处理时的偏最小二乘回归(PLS)模型性能为 RMSEC 为 0.288 2, RMSEP 为 0.343 6, R_c 为 0.960 6, R_p 为 0.934 9;照射位置为斜上部且光谱没有预处理的 PLS 模型性能为 RMSEC 为 0.340 7, RMSEP 为 0.513 3, R_c 为 0.931 1, R_p 为 0.863 6;照射位置为上部且光谱没有预处理的主成分分析回归(PCR)模型性能为 RMSEC 为 0.573 6, RMSEP 为 0.601 4, R_c 为 0.842 4, R_p 为 0.800 7;照射位置为斜上部且光谱没有预处理的 PCR 模型性能为 RMSEC 为 0.709 2, RMSEP 为 0.797 4, R_c 为 0.701 4, R_p 为 0.670 7,最佳照射位置为苹果上部;进一步地采用多种预处理方法对照射位置为上部的 PLS 模型进行对比,得到最优模型为 MSC-PLS 模型,其 RMSEC 为 0.2264 4, RMSEP 为 0.301 5, R_c 为 0.966 9, R_p 为 0.949 9。最后再对相同的 46 个苹果进行相同的实验操作得到光谱、真值后,代入到建立的 MSC-PLS 模型中进行外部验证,结果显示外部验证的相关系数为 0.930 58,验证均方根误差为 0.843 59,验证了建立的 MSC-PLS 模型的稳定性和可靠性,进一步表明光谱采集位置为苹果上部时的近红外漫透射模型有很好的预测能力,该研究为预测苹果可溶性固形物的检测提供了技术支持。

关键词 近红外;光照位置;苹果;位置优化;可溶性固形物

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)11-3373-07

引 言

随着社会的发展和进步,人们对食品的要求也在不断地提高,苹果作为一种富含维生素和矿物质的优质水果,在广大消费者中很受欢迎,苹果的内外品质也一直是人们在购买时所关注的重点。可见-近红外无损检测技术作为一种高效率、低成本且对检测对象不会造成破坏的检测手段,在苹果、西瓜、脐橙、柑橘、柚子、草莓、哈密瓜等水果的内外品质检测方面被广泛使用^[1]。

在对水果进行糖度检测时,漫反射、漫透射等检测方式建立的模型会有不同的预测结果,采用同一个检测方式的同时,检测位置的不同也会导致建立模型的预测能力不同。刘

旭文等在建立梨的可溶性固形物的近红外光谱检测模型稳定性及优化研究中,通过对翠冠梨径向三个不同检测位置(果梗、赤道、花萼)采集光谱和测量糖度建立不同位置的可溶性固形物评价模型,用 CARS 筛选变量,筛选出 36 个特征变量后用偏最小二乘回归(partial least square, PLS)建立模型,得到用赤道区附近光谱及糖度建立的模型预测结果更准。李雄^[2]等通过研究西瓜光投射规律与品质属性的内在联系时提出,不同的积分时间和电流强度会导出不同的建模结果;在其他条件相同情况下,根据西瓜瓜皮到果心的距离将西瓜分为心糖、中糖、外糖、底边糖和混合糖,通过卷积平滑降解光谱噪声并用偏最小二乘建立可溶性固形物的糖度模型后,得到的结果表明用底边糖区域的光谱信息建模效果最佳。目前还没有见到针对苹果漫透射采集光谱的不同光照位

收稿日期: 2021-08-20, 修订日期: 2022-02-23

基金项目: 国家“十二五”(863 计划)项目(SS2012AA101906), 南方山地果园智能化管理技术与装备协同创新中心项目(赣教高字[2014]60 号), 国家自然科学基金项目(31760344), 江西省教育厅科学技术研究青年项目(GJJ190348)资助

作者简介: 刘燕德,女,1967年生,华东交通大学机电与车辆工程学院教授 e-mail: jxliuyd@163.com

置对建立模型影响的研究报道。

将果茎为 70 mm 苹果果实放置在近红外静态光谱采集装置上；该近红外静态光谱采集装置的卤素灯可以调整其照射位置及角度，通过调整卤素灯光对苹果不同部位的照射区域，把照射位置分为上部(与水平方向夹角为 80°)和斜上部(与水平位置夹角为 50°)，探究苹果果实不同的光照位置对建立基于可见-近红外的苹果可溶性固形物的预测模型的影响，找到在检测苹果可溶性固形物时，照射在苹果果实上的最佳光照位置。

1 实验部分

1.1 材料

实验所需的苹果样品购买于江西省南昌市某水果市场，按照相同尺寸购买苹果样品，苹果尺寸为 (73 ± 2) mm。买来的苹果用纯净水和消毒纸巾将表面清洗擦拭干净后，放在室温 20°C 环境下静置 24 h，因为购置的苹果的温度不一样，静置 24 h 目的是为了消除在建立可溶性固形物模型时温度造成的影响。静置 24 h 的苹果剔除在裂果，碰伤果和品质相差较大的果实后随机分成两组，一组照射果实的上部，另一组照射果实的斜上部；照射苹果上部的果实根据 K-S 算法分成训练集 188 个，预测集 62 个。所有果实的果茎范围为 (73 ± 2) mm，用来建模的苹果的可溶性固形物的糖度区间为 $8.2 \sim 16.8^\circ\text{Brix}$ ，用来预测的模型性能的苹果可溶性糖度区间为 $8.8 \sim 16.2^\circ\text{Brix}$ 。照射苹果斜上部的果实训练集为 188 个，预测集为 62 个，所有果实的果茎范围为 $70 \sim 75$ mm，训练集糖度范围为 $9.1 \sim 16.6^\circ\text{Brix}$ ，预测集糖度范围为 $10.5 \sim 16.6^\circ\text{Brix}$ 。

表 1 苹果的理化指标范围统计结果

Table 1 Statistical results of physical and chemical indices of apple

照射位置	数据集	样品数 / 个	果径范围 / mm	糖度范围 / ($^\circ\text{Brix}$)
上部	训练集	188	73 ± 2	$8.2 \sim 16.8$
	预测集	62	73 ± 2	$8.8 \sim 16.2$
斜上部	训练集	188	73 ± 2	$9.1 \sim 16.6$
	预测集	62	73 ± 2	$10.5 \sim 16.6$

1.2 装置和光谱采集

可见近红外漫透射光谱采集装置如图 1 所示，该装置是由可调灯架、支架、卤素灯、光纤固定支架、光纤、光谱仪、计算机以及果杯组成。可调灯架用于调整卤素灯照射角度以及卤素灯与样品的距离；可调灯架与支架用螺丝螺母固定，两个可调灯架分别安装两盏 12 V, 100 W 的卤素灯，采用稳压电源供电。光纤通过光纤固定装置固定在支架上；果杯安装在光纤固定装置上部；光纤另一端与光谱仪连接，光谱仪另一端与计算机连接。采集光谱时，将卤素灯调整至与水平夹角 45° 处，此时采集的位置为苹果的斜上部，采集完斜上部的样品后，将卤素灯调整至与水平夹角 75° 处，此时的采集位置为上部，具体如图 1 所示。整个装置放在一个暗箱中，

光谱采集的全过程在暗箱中进行。光源按照漫透射的方式照射到苹果，透射光谱被传送到光谱仪，然后再传输至计算机。

光谱采集时，要提前打开电源使卤素灯和光谱采集装置预热 20 min，预热后首先对聚四氟乙烯球进行光谱采集，之后分别对照射位置为上部的苹果和照射位置为斜上部的苹果采集光谱后取平均光谱，将平均光谱作为苹果样品的漫透射光谱。实验参数为：积分时间 100 ms，光强 100 W，卤素灯 6 盏。

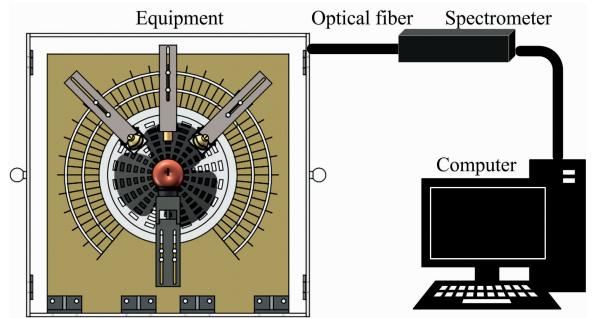


图 1 光谱检测系统示意图

Fig. 1 Schematic diagram of the spectrum detection system

1.3 糖度判定

对采集光谱后的苹果，用刀沿着光谱采集位置切下 4 mm 果肉薄片榨取相应的苹果汁，用滤网过滤掉榨汁残留的果肉后用一次性吸管吸取榨取的果汁，滴到 ATAGO PAL-1 便携式数显折射计镜面上。在吸取果汁测定之前，首先用蒸馏水清洗测定仪器的测定部分，完成仪器测量之前的调零，一个检测面的糖度测定三次，取平均值为其真实糖度。在这个过程中，需要注意的是，榨取完第一个位置的苹果片、测完相应糖度后，在进行下一个位置的苹果片榨取测定时，要将水果刀及榨汁的设备用蒸馏水洗刷干净，更换一次性吸管，并且对 ATAGO PAL-1 设备用清水重新调零，避免因果汁残留导致糖度测量出现偏差进而导致实验数据有误。

1.4 数据处理

用 MATLAB2016A 提取光谱及波段数据，并将数据导出到 EXCEL 表格后在相同的 EXCEL 中输入测得的真实糖度值，将真实糖度和对应的光谱一起输入到 The Unscrambler，首先剔除苹果光谱和糖度数据中的异常样本，对照射位置分别为上部和斜上部的苹果光谱和可溶性固形物数据分别采用相同的回归方法进行分析，选择训练集、检验集和外部测试集后再选择合适的线性或者非线性方法进行建模。

在统计学中，回归分析指的是两种或两种以上变量之间的相关性关系的定量分析。按照一个变量涉及的另几个变量的数目，可以分为一元线性回归和多元线性回归^[3]，一元线性回归是一个变量涉及另外一个变量，探究两个变量之间的相关线性关系；多元线性回归则是一个变量涉及两个或两个以上变量，探究至少三个变量之间的相关线性关系；按照自变量和因变量之间的关系，又可以把回归分为线性回归和非线性回归^[4-7]。

1.5 模型评价

建立的定量模型需要有性能评价标准，会用到一些统计参数，例如校正标准偏差 RMSEC，预测标准偏差 RMSEP 和决定系数 R^2 ，通过比较这三个指标，选取最优的建模参数和模型^[8]。一般来说，校正标准偏差 RMSEC 的数值越低，表明所建立预测模型的相关性越好，预测标准偏差 RMSEP 的数值越低，则说明所建立模型的预测准确性更准确^[9-10]。

三个评价指标的计算公式如下。

① 训练集的均方根误差

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n_c} (y_c - y)^2}{n_c}} \quad (1)$$

式(1)， y 是样品指标的真实值； y_c 是模型根据建模数据得出的预测值； n_c 是校正集建模所需的样品数量。

② 预测集的均方根误差

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n_p} (y_p - y)^2}{n_p}} \quad (2)$$

式(2)， y 是样品指标的真实值； y_p 是模型根据建模数据得出的预测值； n_p 是预测集建立所需的样品数量。

通常情况下，RMSEC 和 RMSEP 两者的值越小并且两者间的差值越接近零时，可以做出判断，建立的模型效果最好，否则模型效果不佳。

③ 决定系数

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{i, \text{actual}} - y_{i, \text{predicted}})^2}{\sum_{i=1}^n (y_{i, \text{actual}} - \bar{y}_{\text{actual}})^2} \quad (3)$$

式(3)中， $y_{i, \text{actual}}$ 为第 i 样本参考方法的测定值； \bar{y}_{actual} 为校正集或测试集所有样本参考方法测定值的平均值； $y_{i, \text{predicted}}$ 为校正集或测试集预测过程中第 i 样本的预测值， n 为校正集或测试集的样本数^[11]。

2 结果与讨论

2.1 光谱分析

2.1.1 不同光照位置的差异影响分析

图 2 是近红外漫透射苹果不同光照位置光路示意图。采用漫透射的方式研究苹果不同光照位置对苹果可溶性固形物检测模型的影响时，如图 2 灯 B 所示，灯 B 的光路示意线是照射位置为苹果斜上部位置，光源光路主要携带信息为浅层的果肉信息，光穿过果核的部分少，不会采集到果核的信息，避免了果核信息对整个可溶性固形物模型建立带来的干扰；如图 2 灯 A 所示，其光路示意线是照射苹果上部位置的光线示意图，在此位置，光线可以穿过苹果的大部分区域，也包含果核的大部分信息，所包含的苹果的信息内容更多更丰富。

2.1.2 不同光照位置的光谱对比分析

如图 3 所示，为基于近红外漫透射的苹果不同照射位置的光谱光强示意图，图示为平均原始光谱。由图可得，照射位

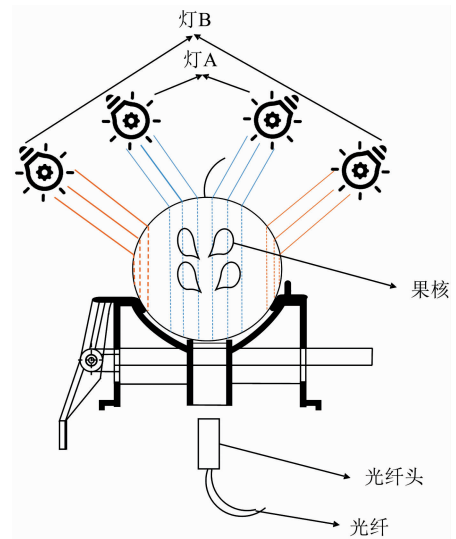


图 2 近红外漫透射苹果不同光照位置光路示意图

Fig. 2 Schematic diagram of light rays go through the apple along two paths designed for near infrared diffuse transmission experiment

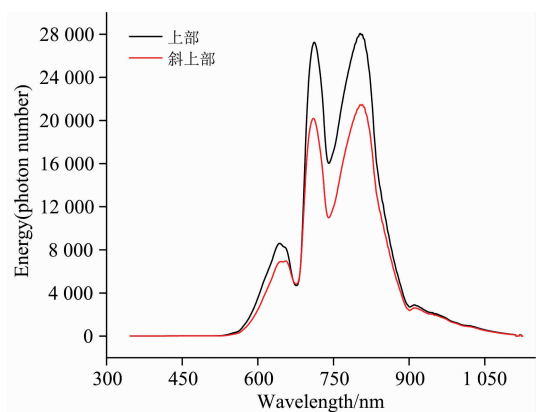


图 3 基于近红外漫透射的苹果不同光照位置的光谱光强示意图

Fig. 3 Schematic diagram of intensity distributions of light go through the apple at different illumination positions based on near infrared diffuse transmission

置为上部的光谱漫透射能量比照射位置为斜上部的苹果漫透射能量强；两个不同照射位置的近红外漫透射光谱的吸收峰和波谷所在的波段位置基本相同，在 645, 710 和 820 nm 波段处出现了光谱波峰，且 710 nm 处和 820 nm 处的能量峰值相差较大，在 645 nm 处的能量相差较小；在 675 和 740 nm 处存在波谷，740 nm 处的波谷差值比 675 nm 处的波谷差值大。由于采集到的光谱在 500 nm 之前和 1 100 nm 之后含有的有效信息较少，且含有大量的噪声信息和无关信息，所以在建立可溶性固形物模型时，选取 500~1 100 nm 的光谱建模。在波段为 675 nm 附近处，可能是由于样品中叶绿素含量吸收变化的强度不同导致了光谱吸收峰的出现^[12]。在 740 nm 处的光谱吸收峰与羟基的第三倍频和 C—H 键的第四倍

频相关。上述的光谱波段所携带的信息都与苹果中可溶性固形物(soluble solids content, SSC)的含量有一定的联系。

2.1.3 可溶性固形物统计

表 2 为基于近红外光谱不同光照位置对苹果可溶性固形物建模影响的样品数据统计,其中照射位置为上部的样品训练集有 188 个苹果样品,苹果的平均糖度值为 12.969 4° Brix,糖度范围为 8.2~16.8° Brix,照射位置为上部的样品预测集有 62 个苹果样品,苹果的平均糖度值为 13.180 7° Brix,糖度范围为 8.8~16.2° Brix;照射位置为斜上部的训练集有 188 个苹果样品,苹果的平均糖度值为 13.259 7° Brix,糖度范围为 9.1~16.6° Brix;照射位置为斜上部的预测集有 62 个苹果样品,苹果的平均糖度值为 13.373 9° Brix,糖度范围为 10.5~16.6° Brix,预测集的糖度范围被包含在训练集的糖度范围内,数据选择合理。

表 2 校正集和预测集中可溶性固形物 SSC 含量的范围和平均值

Table 2 Ranges and average values of SSC in calibration set and prediction set

位置	Parameter	Data Set	Samples	Mean	Range
上部	SSC/(°Brix)	Calibration	188	12.969 4	8.2~16.8
		Prediction	62	13.180 7	8.8~16.2
斜上部	SSC/(°Brix)	Calibration	188	13.259 7	9.1~16.6
		Prediction	62	13.373 9	10.5~16.6

2.2 建模与分析

2.2.1 不同预处理方法的 PLS 模型

表 3 所示为基于近红外光谱不同光照位置对苹果可溶性固形物建模影响的不同预处理方法的 PLS 模型,从表 3 中可以得出,当光源照射位置为苹果上部时,没有对光谱进行预处理的原始光谱 RMSEC 为 0.288 2, R_c 为 0.960 6, RMSEP 为 0.343 6, R_p 为 0.934 9,而照射位置为斜上部的 RMSEC 为 0.340 7, R_c 为 0.931 1, RMSEP 为 0.513 3, R_p 为 0.863 6,从照射位置为上部原始光谱的数据来看, RMSEC 和 RMSEP 的差值为 0.055 4,而照射位置为斜上部的原始光谱的 RMSEC 和 RMSEP 的差值为 0.172 6,照射位置为上部原始光谱的 R_p 为 0.934 9,照射位置为斜上部原始光谱的 R_p 为 0.863 6,照射位置为上部的原始光谱的 R_c 为 0.906 0,照射位置为斜上部的原始光谱的 R_c 为 0.931 1,采用不同的光谱预处理方法所得的结果中,不同光谱预处理得到的所有的 RMSEP, RMSEC, R_c 和 R_p 均显示,基于近红外的苹果漫透射可溶性固形物模型照射位置为上部时的模型比斜上部的模型效果、精度更好。

2.2.2 不同预处理方法的 PCR 模型

表 4 为基于近红外光谱不同光照位置对苹果可溶性固形物建模影响的不同预处理方法的主成分分析回归(principal component regression, PCR)模型,从表 4 中可以得出,当光源照射位置为苹果上部时,没有对光谱进行预处理的原始光谱 RMSEC 为 0.576 3, R_c 为 0.842 4, RMSEP 为 0.601 4, R_p 为 0.800 7,而照射位置为斜上部的 RMSEC 为 0.709 2,

R_c 为 0.701 4, RMSEP 为 0.797 4, R_p 为 0.670 7,从照射位置为上部原始光谱的数据来看, RMSEC 和 RMSEP 的差值为 0.025 1,而照射位置为斜上部的原始光谱的 RMSEC 和 RMSEP 的差值为 0.088 2,照射位置为上部原始光谱的 R_p 为 0.800 7,照射位置为斜上部原始光谱的 R_p 为 0.670 7,照射位置为上部的原始光谱的 R_c 为 0.842 4,照射位置为斜上部的原始光谱的 R_c 为 0.701 4,采用不同的光谱预处理方法所得的结果中,不同光谱预处理得到的所有的 RMSEP, RMSEC, R_c 和 R_p 均显示,基于近红外的苹果漫透射可溶性固形物模型照射位置为上部时的模型比斜上部的模型效果、精度更好。

表 3 不同预处理方法的苹果上部、斜上部漫透射 PLS 模型处理结果

Table 3 Modeling results of PLS with different pretreatment methods based on diffuse transmission data of apple samples using illuminations of upper and inclined upper

检测方式	预处理方法	Calibration		Prediction	
		RMSEC	R_c	RMSEP	R_p
上部	Raw	0.288 2	0.960 6	0.343 6	0.934 9
	SG-PLS	0.408 1	0.921 0	0.400 0	0.911 9
	SG-Derivative	0.245 4	0.971 4	0.432 7	0.896 8
	SNV	0.264 9	0.966 7	0.303 7	0.949 2
斜上部	MSC	0.264 4	0.966 9	0.301 5	0.949 9
	Raw	0.340 7	0.931 1	0.513 3	0.863 6
	SG-PLS	0.474 9	0.866 1	0.579 3	0.826 1
	SG-Derivative	0.295 8	0.948 1	0.633 1	0.792 4
斜上部	SNV	0.317 7	0.940 1	0.451 7	0.894 3
	MSC	0.322 9	0.938 1	0.454 6	0.893 0

表 4 不同预处理方法的苹果上部、斜上部漫透射 PCR 模型处理结果

Table 4 Modeling results of PCR with different pretreatment methods based on diffuse transmission data of apple samples using illuminations of upper and inclined upper

检测方式	预处理方法	Calibration		Prediction	
		RMSEC	R_c	RMSEP	R_p
上部	Raw	0.576 3	0.842 4	0.601 4	0.800 7
	SG	0.577 0	0.842 1	0.599 6	0.801 9
	SG-Derivative	0.859 3	0.949 6	0.982 6	0.468 1
	SNV	0.538 2	0.862 5	0.579 5	0.815 0
斜上部	MSC	0.540 0	0.861 8	0.585 2	0.811 4
	Raw	0.709 2	0.701 4	0.797 4	0.670 7
	SG	0.708 5	0.702 0	0.794 8	0.672 8
	SG-Derivative	0.903 6	0.515 3	0.968 9	0.513 7
斜上部	SNV	0.687 8	0.719 1	0.781 5	0.683 6
	MSC	0.695 1	0.713 2	0.770 7	0.692 4

2.3 预测模型建立与分析

近红外光谱的苹果糖度光照位置优化研究中,从两个不

同照射位置的光谱差异来看，照射位置为上部的平均光谱光强强于照射位置为斜上部的平均光谱光强，从近红外漫反射的光路分析来看，照射位置为上部时，光谱包含了苹果本身更多的可溶性固形物信息，虽然照射位置为苹果斜上部时，包含的信息里大部分为苹果外部浅层果肉和部分的中部果肉信息，很少包含了果核信息，但是无论从得到的光谱的光强角度来看，还是光路分析来看，照射位置为苹果上部时，建立的苹果可溶性固形物模型建模和预测的效果都明显高于照射位置为苹果斜上部时，所以可以得出结论，基于近红外光谱的漫透射的苹果可溶性固形物模型建立时最好的光照位置为苹果的上部，此位置尽可能全面的包含了苹果可溶性固形物建立所需要的信息。

2.3.1 预测模型建立

综合以上结论可以得出在探究基于近红外漫透射苹果可溶性固形物模型建立最佳位置的实验中，最佳的照射位置为苹果的上部，在建立照射位置为上部的模型中，如表 5 所示，不论是原始光谱还是对光谱进行预处理之后，光谱的 PLS 模型建模效果比 PCR 效果好很多，在 PLS 模型中，SG-PLS 模型的性能略低于没有光谱预处理的光谱的模型性能，但 SG-一阶导数-PLS、SNV-PLS 和 MSC-PLS 的模型的性能比没有光谱预处理的模型性能提高了很多，其中 SNV-PLS 和 MSC-PLS 的模型是所有模型性能中评价指标最好的两个模型，SNV-PLS 的 RMSEC 比原始未光谱预处理模型的 RMSEC 降低了 0.023 3，RMSEP 降低了 0.039 9， R_c 相比较未光谱预处理的 R_c 提高了 0.006 1， R_p 提高了 0.014 3；而 MSC-PLS 的 RMSEC 比原始未光谱预处理模型的 RMSEC 降低了 0.023 8，RMSEP 降低了 0.042 1， R_c 相比较未光谱预处理的 R_c 提高了 0.006 3， R_p 则提高了 0.015，综上所述，基于近红外漫透射苹果可溶性固形物模型建立的最佳照射为苹果的上部，且照射位置为上部的最佳模型是 MSC-PLS 模型，最佳模型的 RMSEC 为 0.264 4，RMSEP 为 0.301 5， R_c 为 0.966 9， R_p 为 0.949 9。图 4 为 MSC-PLS 建模数据真实糖度和预测糖度值的线性拟合示意图，可以看到预测值较为平均且接近的拟合在拟合线附近，散点越靠近拟合线，说明模型的预测性能越好。

表 5 上部照射位置不同模型评价指标
Table 5 Different model evaluation indexes of upper irradiation position

建模方式	预处理方法	Calibration		Prediction	
		RMSEC	R_c	RMSEP	R_p
PLS	Raw	0.288 2	0.960 6	0.343 6	0.934 9
	SG-PLS	0.408 1	0.921 0	0.400 0	0.911 9
	SG- Derivative	0.245 4	0.971 4	0.432 7	0.896 8
	SNV	0.264 9	0.966 7	0.303 7	0.949 2
	MSC	0.264 4	0.966 9	0.301 5	0.949 9
PCR	Raw	0.576 3	0.842 4	0.601 4	0.800 7
	SG-PLS	0.577 0	0.842 1	0.599 6	0.801 9
	SG- Derivative	0.859 3	0.949 6	0.982 6	0.468 1
	SNV	0.538 2	0.862 5	0.579 5	0.815 0
	MSC	0.540 0	0.861 8	0.585 2	0.811 4

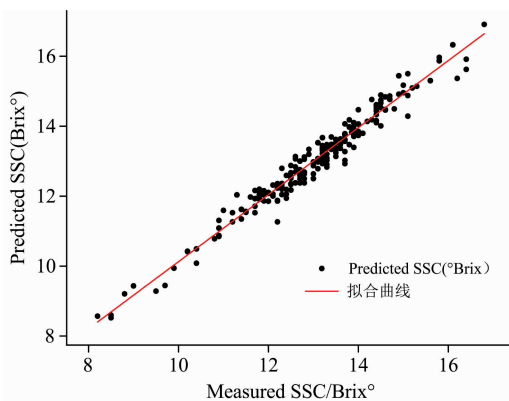


图 4 MSC-PLS 的糖度真实值和预测值的线性拟合
Fig. 4 Linear fitting between real and predicted sugar contents of msc-pls

2.3.2 模型外部验证

在建立了最佳光照位置为上部的 MSC-PLS 模型后，为了对模型进行一个稳定性和可靠性检验，又从购买苹果的水果批发市场购买了相同品种，尺寸接近的 46 个同批次的苹果，在对新购买的苹果进行同样的样品处理后，在同样的环境、温度、湿度下采集光照位置为上部的近红外漫透射光谱，进行外部验证^[13]。结果如图 5 所示，外部验证的 46 个苹果样品的糖度预测和标准理化值得相关系数为 0.9305 8，验证均方根误差为 0.8435 9，所得分析数据进一步说明了照射位置为上部的基于近红外漫透射的苹果可溶性固形物建立的 MSC-PLS 模型是具有良好稳定性，可靠性的苹果可溶性固形物预测模型。

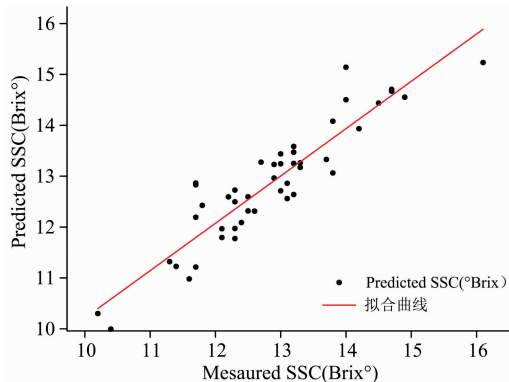


图 5 外部验证糖度真值和预测值线性拟合示意图
Fig. 5 Schematic diagram of linear fitting between true value and predicted value of sugar content for external verification

3 结 论

利用自行研发且已证实过可行性的近红外光谱采集仪器，通过在相同的实验环境和实验条件的前提下，对从同一个水果批发市场采购的尺寸接近的同批次苹果采用不同的光照位置进行近红外漫透射光谱的采集，并对其糖度进行测量，分别对照射位置为上部 and 照射位置为斜上部的光谱进行

建模, 照射位置为上部且光谱没有预处理时的 PLS 模型性能为 RMSEC 为 0.288 2, RMSEP 为 0.343 6, R_c 为 0.960 6, R_p 为 0.934 9; 照射位置为斜上部且光谱没有预处理的 PLS 模型性能为 RMSEC 为 0.340 7, RMSEP 为 0.513 3, R_c 为 0.931 1, R_p 为 0.863 6; 照射位置为上部且光谱没有预处理的主成分回归模型性能为 RMSEC 为 0.573 6, RMSEP 为 0.601 4, R_c 为 0.842 4, R_p 为 0.800 7; 卤素灯照射位置为苹果的斜上部且光谱没有预处理的 PCR 模型性能为 RMSEC 为 0.709 2, RMSEP 为 0.797 4, R_c 为 0.701 4, R_p 为 0.670 7, 所以无论是 PLS 模型还是 PCR 模型, 都说明了建立可溶性固形物糖度模型时, 最好的光照位置是上部, 上部

包含了更多的可溶性固形物信息, 且 PLS 模型的模型性能好于 PCR 模型。进一步地, 对光谱进行预处理后, 得到了最优的模型 MSC-PLS 模型, 该模型的模型性能为 RMSEC 为 0.226 44, RMSEP 为 0.301 5, R_c 为 0.966 9, R_p 为 0.949 9。进一步地, 又在相同的实验条件和实验环境下, 另采购 46 个苹果进行同样操作后进行外部验证, 得出 MSC-PLS 模型是稳定可靠的。本文探究了基于近红外漫透射的苹果可溶性固形物模型建立的最佳光照位置为上部, 且最优的模型为 MSC-PLS, 为研究苹果的可溶性固形物和苹果的品质检测提供了新的技术支撑。

References

- [1] LIU Yan-de, CHENG Meng-jie, HAO Yong(刘燕德, 程梦杰, 郝勇). Journal of East China Jiaotong University(华东交通大学学报), 2018, 35(4): 1.
- [2] LI Xiong, LIU Yan-de, SUN Xu-dong, et al(李雄, 刘燕德, 孙旭东, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2020, 40(10): 3265.
- [3] GUO Tao, HUANG You-qin, DAI Lu-ming, et al(郭涛, 黄右琴, 代露茗, 等). Chinese Journal of Animal Nutrition(动物营养学报), 2021, 33(5): 2939.
- [4] WU Yan-yong, SHI Yan-cai, LÜ Xue-juan, et al(吴延勇, 史艳财, 吕雪娟, 等). Hubei Agricultural Sciences(湖北农业科学), 2021, 60(12): 128.
- [5] LIU Yan-de, ZHU Dan-ning, WU Ming-ming, et al(刘燕德, 朱丹宁, 吴明明, 等). Food and Machinery(食品与机械), 2016, 32(10): 115, 163.
- [6] Gabriel A Leiva-Valenzuela, Lu Renfu, José Miguel Aguilera. Journal of Food Engineering, 2013, 115(1): 91.
- [7] CAI De-ling, TANG Chun-hua, LIANG Yu-ying, et al(蔡德玲, 唐春华, 梁玉英, 等). Food and Fermentation Industries(食品与发酵工业), 2020, 46(7): 218.
- [8] MENG Qing-long, SHANG Jing, HUANG Ren-shuai, et al(孟庆龙, 尚静, 黄人帅, 等). Food and Fermentation Industries(食品与发酵工业), 2020, 46(19): 205.
- [9] LEI Ying, LIU Cui-ling, ZHOU Zi-yan(雷鹰, 刘翠玲, 周子彦). Journal of Food Science and Technology(食品科学技术学报), 2018, 36(6): 95.
- [10] ZHOU Jing, DING Li, ZHANG Jun-lin, et al(周静, 丁黎, 张俊林, 等). Acta Armamentarii(兵工学报), 2018, 39(3): 463.
- [11] ZHANG Jin, CAO Xiao-ning, TIAN Xiang, et al(张晋, 曹晓宁, 田翔, 等). Journal of Anhui Agricultural Sciences(安徽农业科学), 2021, 49(9): 175.
- [12] WU Shuang, LI Guo-jian, JIE Deng-fei(吴爽, 李国建, 介邓飞). Food and Machinery(食品与机械), 2020, 36(12): 132.
- [13] SHEN Guang-hui, CAO Yao-yao, LIU Xin, et al(沈广辉, 曹瑶瑶, 刘馨, 等). Jiangsu Journal of Agricultural Sciences(江苏农业学报), 2021, 37(2): 509.

Study on Optimization of Apple Sugar Degree and Illumination Position Based on Near-Infrared Technology

LIU Yan-de, CUI Hui-zhen, LI Bin, WANG Guan-tian, XU Zhen, LI Mao-peng

School of Mechanical, Electrical and Vehicle Engineering, East China Jiaotong University, Nanchang 330013, China

Abstract When using near-infrared diffuse transmission technology to detect SSC of apple, the information of soluble solids in the collected apple spectrum is different due to the different positions of halogen lamp on apple, which will lead to the different performance of the model. Finding the best illumination position of apple is conducive to obtaining the best evaluation model of soluble solids. Using the multi-mode adjustable optical structure, the near-infrared diffuse transmission spectra of two batches of apples purchased from the same fruit wholesaler with similar size but different irradiation positions were collected under the same experimental environment and conditions. The best irradiation position in the process of establishing the apple soluble solid model was studied, and the evaluation model of the best position of soluble solid was obtained. The best modeling position is obtained by spectrum collection, true sugar degree value collection and chemometrics method. When the irradiation position is the upper part and the spectrum is not pretreated, the PLS (Partial Least Square) model performance is RMSEC 0.288 2, RMSEP 0.343 6, R_c 0.960 6 and R_p 0.934 9. The performance of the PLS model with oblique upper irradiation position and no spectral pretreatment is RMSEC 0.340 7, RMSEP 0.513 3, R_c 0.931 1 and R_p 0.863 6. The performance of the PCR (Principle Component Regression) model with upper irradiation position and no spectral pretreatment was RMSEC 0.573 6, RMSEP 0.601 4, R_c 0.842 4 and R_p 0.800 7. The performance of the PCR model with oblique upper irradiation position and no spectral pretreatment was RMSEC 0.709 2, RMSEP 0.797 4, R_c 0.701 4, R_p 0.670 7. The best irradiation position is the upper part of the apple; Further, a variety of pretreatment methods are used to compare the PLS model with the upper irradiation position. The optimal model is the MSC-PLS model. Its model performance is RMSEC 0.226 44, RMSEP 0.301 5, R_c 0.966 9 and R_p 0.949 9. Finally, after the same experimental operation is carried out on the same 46 apples, the spectra and true values are obtained and substituted into the established MSC-PLS (Multiplicative Scatter Correction-Partial Least Square) model for external verification. The results show that the correlation coefficient of external verification is 0.930 58, and the root mean square error of verification is 0.843 59, which verifies the stability and reliability of the established MSC-PLS model. It further shows that the near-infrared diffuse projection model has good prediction ability when the spectral acquisition position is the upper part of the apple. This paper provides technical support for predicting the detection of soluble solids in the apple.

Keywords Near-infrared; Lighting position; Apple; Location optimization; Soluble solids

(Received Aug. 20, 2021; accepted Feb. 23, 2022)