基于近红外光谱分析贡梨可溶性固形物全局模型的鲁棒性

刘燕德,廖 军,李 斌,姜小刚,朱明旺,姚金良,王 秋

华东交通大学机电与车辆工程学院,智能机电装备创新研究院,江西南昌 330013

摘 要 贡梨是大众喜爱的水果,为研究不同检测方向对近红外在线检测贡梨可溶性固形物 SSC 的影响, 提出全局模型并分析其鲁棒性。在贡梨六个方向上收集光谱:茎-花萼轴垂直,茎向上(A1)和茎向下(A5), 茎-花萼轴和水平之间 45°, 茎向上倾斜(A2)和茎向下倾斜(A4), 茎-花萼轴水平, 茎朝向右侧光(A3), 茎花 萼轴水平, 茎朝向带移动方向(A6)。SSC 范围为 9.53~14.70 的 150 个样品分为 115 个标准偏差为 1.05 的 校准集和 35 个标准偏差为 0.93 的预测集。采用偏最小二乘回归 PLSR 分别建立六个局部模型和一个全局 模型,局部模型由各方向的115个校正集数据经过 Savitzky-Golay 卷积平滑、多元散射校正 MSC、高斯滤波 平滑 GFS 三种不同的预处理方法处理后使用偏最小二乘回归 PLSR 建立而来;用本方向校正集数据建立的 局部模型验证本方向的 35 个预测集数据,比较这三种预处理方法后所建立的 PLSR 模型,结果表明经过 GFS 处理后建立的模型验证效果最好,因此六个局部模型和全局模型均采用 GFS 处理后建立的 PLSR 模 型。全局模型是由 A1, A2, A3, A4, A5 和 A6 六个方向的 690 个校正集光谱数据经过 GFS 预处理后采用 PLSR 建立的贡梨 SSC 模型。各方向的预测集分别对七个模型进行验证,验证结果表明,局部模型虽然在本 方向的预测效果强于全局模型,但无法验证其他方向,鲁棒性差,由此可知检测方向的不同对预测效果的影 响很大;全局模型能够准确预测各个检测方向的贡梨 SSC,全局模型的校正集相关系数 R。为 0.828,校正集 均方根误差 RMSEC 为 0.424; A1, A2, A3, A4, A5 和 A6 方向的预测集相关系数 R。分别为 0.818, 0.765, 0.799, 0.821, 0.794 和 0.824, 预测集均方根误差 RMSEP 分别为 0.446, 0.525, 0.478, 0.538, 0.486 和 0.619; 六个方向的 R。与 R。比较接近且均在 0.800 左右, RMSEC 与 RMSEP 均在 0.500 左右, 结 果表明全局模型在检测不同方向的贡梨 SSC 上有着极好的鲁棒性。

关键词 近红外; 贡梨; 可溶性固形物 SSC; 全局模型; 鲁棒性 中图分类号: O657.3 文献标识码: A DOI: 10.3964/j.issn.1000-0593(2022)09-2781-07

引 言

贡梨是一种受人喜爱的水果,以皮薄多汁、味浓甘甜而 闻名。贡梨中的可溶性固形物含量是影响贡梨品质的重要内 部属性之一,也是决定水果成熟度和收获时间的关键参数之 一。因此,开发无损、快速的贡梨可溶性固形物含量的检测 方法来满足消费者对稳定、优质水果日益增长的需求具有重 要的价值。近红外光谱技术能够实现食品的无损检测分析, 而且几乎不需要对样品进行处理即可进行检测,不会产生废 物,并允许同时评估几种成分^[1]。近红外在线检测已被应用 于测定梨^[2]、橙子^[3]、橄榄^[4]、苹果^[5]等水果的 SSC。

水果的物理性质如颜色[6]、温度[7]、静止位置[8]、移动

速度^[9]等在检测过程中可能会对光谱信息采集有很大影响, 除这些物理性质以外,测量方向也是影响光谱信息采集的重 要因素之一。Yu等^[10]研究了基于近红外光谱的红富士苹果 果实方位变化对可溶性固形物在线预测的影响;韩东海 等^[11]比较了3种果实方向即茎-花萼轴水平、茎-花萼轴垂 直、茎向上和茎向下在线检测苹果中的SSC,基于茎-花萼轴 垂直、茎向上的方向得到了较好的结果;在透射模式下,基 于茎萼轴水平方向检测梨褐心得到了较好的结果^[12],茎萼 轴垂直取向也获得了测量苹果的SSC和硬度的良好结果。如 今近红外在线检测水果内部品质的设备多采用局部模型对水 果进行预测,由于局部模型是利用单个局部方向获得的光谱 数据单独开发的,在预测其他水果方向上的SSC时,其精度 会大打折扣。目前解决该问题主要依靠人工摆正果位或增加

收稿日期: 2021-07-17,修订日期: 2021-10-10

基金项目:国家自然科学基金项目(31760344),国家科技奖后备项目培育计划项目(20192AEI91007),江西省教育厅科学技术研究项目 (GJJ200615),江西省教育厅科学技术研究项目(GJJ190306)资助

作者简介:刘燕德,女,1967年生,华东交通大学智能机电装备创新研究院教授 e-mail:jxliuyd@163.com

摆正机构,费钱费力。建立一个基于全方向的全局模型也许 能够解决上面的问题,也还没有人将全局模型应用于近红外 在线漫透射检测装备来预测不同方向运送下的贡梨 SSC。

本研究利用课题组自主研发的在线漫透射检测装备采集 六个不同检测方向下的贡梨光谱,使用 Savitzky-Golay 卷积 平滑、多元散射校正 MSC、高斯滤波平滑 GFS 对平均后的 150 贡梨光谱进行预处理,对光谱进行修正和优化,最后通 过偏最小二乘回归 PLSR 建立了六个局部模型和一个全局模 型。六个方向的预测集分别对七个模型进行预测,对预测结 果进行比较并评判各局部模型与全局模型的鲁棒性。为全局 模型是否能够应用于近红外在线漫透射检测装备提供参考。

1 实验部分

1.1 样本

于 2021 年四月下旬在中国南昌某水果超市购买 150 个 贡梨,150 个贡梨均没有表面缺陷或污染物。贡梨的赤道直 径范围为 85~95 mm,所有样品都单独清洗并编号;实验前 把样品放在实验室储存 24 h,使贡梨温度与实验室温度一 致,以减少温度对预测精度的影响。 将 150 个样本随机分配到校准集和预测集。选择 115 个 样本作为用于开发校准模型的校准集,其余 35 个样本用作 预测集来验证校准模型的预测性能。

1.2 光谱收集

样品的漫透射光谱由本课题组自主研发的水果内部品质 近红外漫透射光谱在线检测装备(图1)测得^[13]。

使用反射率超过 95%的聚四氟乙烯制成的直径 $\phi = 90$ mm 的白色球体测量参考光谱。参考测量三次,计算平均光 谱作为最终的 R_{white} 。当所有的灯都关了,灯的入口打开时, 收集暗光谱 R_{dark} 。

如图 2 所示^[14],茎-花萼轴垂直,茎向上(A1)和茎向下(A5),茎-花萼轴和水平之间 45°,茎向上倾斜(A2)和茎向下倾斜(A4),茎-花萼轴水平,茎朝向右侧光(A3),茎花萼轴水平,茎朝向带移动方向(A6)。实验前随机选取了 60 个贡梨,让它们从上一传送通道上随机传过来,贡梨在果杯上的位置都在这六个方向附近,这可能是因为贡梨本身的结构造成的。且两排光源以果杯为基准为对称光源,这六个方向相当于一个大区间,从上一传送通道上传过来的贡梨在果杯上的位置都在这一区间内。



图 1 水果内部品质近红外漫透射光谱在线检测装备示意图

1: 触发开关; 2: 编码盘; 3: 链轮; 4: 链条; 5: 光纤; 6: 暗箱; 7: 光源; 8: 样品; 9: 终端 PC 机;

10: PLC 控制柜; 11: 果杯; 12: 弹跳器; 13: 分级出口; 14: 电动机; 15: 减速器; 16: 驱动链轮

Fig. 1 Schematic diagram of on-line detection equipment for fruit internal quality by near infrared diffuse transmission spectroscopy

- 1: Porximity switch; 2: Coding mask; 3: Chain wheel; 4: Link chain; 5: Optica fiber; 6: Camera bellows; 7: Optical source;
 - 8: Specimen; 9: Terminal PC; 10: PLC control cabinet; 11: Fruit cup; 12: Bionic boot; 13: Grading export;
 - 14: Electromotor; 15: Reducer; 16: Drive sprocket



Fig. 2 Diffuse transmission testing mechanism (a), Gongli placement position (b)

分别在六个方向测量光谱数据,光谱数据按 A1, A2, A3, A4, A5 和 A6 的顺序采集。用手将每个样品稳定地放在 水果盘上。本课题组自主研发的水果内部品质近红外漫透射 光谱在线检测装备(图 1)能够保证运行平稳,果杯上的贡梨 震感极弱,在传送贡梨的过程中可以保证贡梨的平稳传送。 仅保留 346~1 125 nm 的光谱数据,由 1 044 个数据点即光 谱变量组成,用于进一步分析。然后,原始光谱 R_{nw}根据式 (1)转换为校准光谱 R:

$$R = \frac{R_{\rm raw} - R_{\rm dark}}{R_{\rm white} - R_{\rm dark}} \tag{1}$$

考虑到贡梨形状的不规则性,为了减少操作人员和仪器 引入的误差,提高信噪比,在每个水果方位采集三个光谱, 样品绕茎萼轴旋转 0°,120°和 240°,对三个光谱进行平均得 到平均光谱。样品以 30 cm • s⁻¹的速度由传送带传送,同时 所有光谱数据存储在工业计算机中。

在光谱采集之后,立即通过传统的破坏性试验确定 SSC 值。手动挤压每个完整样品的汁液,并将其滴到日本东京 ATAGO有限公司生产的数字折光仪 PAL-1上(测量精度为 Brix±0.2%,且具有温度补偿功能,适用温度范围 10~40 ℃),以测量 SSC 值,绕茎萼轴旋转 0°,120°和 240°,为每个 样品收集三个 SSC 值,并根据这三个值来获取平均值。

1.3 光谱数据预处理

由于采集的光谱可能包含无用的背景信息或噪声,而光 谱预处理往往可以提高模型的性能。采用不同的预处理方法 对原始光谱数据进行预处理,包括 Savitzky-Golay 卷积平滑 SGS、多元散射校正 MSC 和高斯滤波平滑 GFS 等来优化原 始光谱数据。SGS 可以有效减少光谱的噪声并提高信噪比, 多元散射校正 MSC 可以消除水果内部不均导致的散射效 应;高斯滤波 GFS 可以消除光谱中的随机噪声。

1.4 偏最小二乘回归分析

偏最小二乘回归 PLSR 是一种无监督的多变量统计分析 方法^[15],被广泛应用于 NIRS 分析研究,具有较强的抗干扰 性。PLSR 的目的是通过将光谱矩阵与浓度向量投影到新的 潜在空间,使其中的潜在变量之间的协方差最大化,并最终 执行潜变量线性回归模型。

1.5 模型的建立

为了全面检验测量方向对贡梨 SSC 光谱分析的影响,使 用 PLSR 分别为每个方向和所有方向数据集建立 SSC 模型, 即局部模型与全局模型。每个局部模型都是由该方向上的 115 个校正集光谱数据建立的。因此,分别建立了六个局部 模型,如图 3,校正集用长方形表示,预测集用圆形表示;在 局部模型中,实线所指为本方向的预测集对该模型进行验证;虚线为其他方向的预测集对该模型进行验证,这是为了 评估局部模型的鲁棒性。一些物理性质的全局模型已经成功 地用于减少色差^[8]、温度波动^[9]和生物可变性对近红外模型 性能的影响。在本研究中,全局模型中的校正集由六个方向 的校正集数据组成,总共包括 690 个光谱。为了验证全局模 型的预测能力,分别用六个预测集对全局模型进行验证。比 较七个模型的性能,选择出鲁棒性最好的模型。



1.6 模型性能分析

根据校准和预测的相关系数、校准的均方根误差 RM-SEC、预测的均方根误差 RMSEP 来评估校准模型的性能。 R_c , R_p , RMSEC 和 RMSEP 的计算定义如式(2)一式(5)

$$R_{c} = \sqrt{\sum_{i=1}^{n_{c}} (y_{pi} - y_{mi})^{2}} / \sqrt{\sum_{i=1}^{n_{c}} (y_{pi} - y_{meanc})^{2}}$$
(2)

$$R_{p} = \sqrt{\sum_{i=1}^{n_{p}} (y_{pi} - y_{mi})^{2}} / \sqrt{\sum_{i=1}^{n_{p}} (y_{pi} - y_{meanp})^{2}}$$
(3)

RMSEC =
$$\sqrt{\frac{1}{n_c} \sum_{i=1}^{n_c} (y_{pi} - y_{mi})^2}$$
 (4)

RMSEP =
$$\sqrt{\frac{1}{n_{\rho}} \sum_{i=1}^{n_{\rho}} (y_{\rho i} - y_{m i})^2}$$
 (5)

其中, y_{pi} 为第 i 个水果的 SSC 的预测值, y_{mi} 为第 i 个贡梨 SSC 的实测值, y_{meanc} 和 y_{meanp} 分别为校准集和预测集中 SSC 的平均值, n_c 和 n_p 分别为校准集和预测集中水果个数。— 般来说,选择一个好的模型,不仅要根据较高的 R_c 和 R_p , 较低的 RMSEC 和 RMSEP,还要根据 RMSEC 和 RMSEP 之 间的微小差异来选择。

2 结果与讨论

2.1 光谱特征与光谱分析

在150个贡梨中随机选取第11号果,如图4所示为11 号果在A1—A6六个方向上获得的346~1125 nm 的漫透射 平均原始光谱。由图4可知,A1方向的漫透射光谱的能量明 显强于其他方向,处于其他方向的漫透射光谱上方;其他方 向的漫透射光谱的能量由强到弱依次是A5,A6,A2,A3和 A4,其中A6,A2,A3和A4的漫透射光谱的能量相差不大; 各方向的漫透射光谱的能量不同可能与贡梨所受的有效光源 照射面积有关,其中A1与A5所受的有效光源照射面积最 大。从六个方向所采集的光谱的吸收峰和波谷所处的波段大 致相同,只是光强大小不同。其中波谷在750 nm 附近,在 710与800 nm 附近各有一个波峰。750 nm 附近的波谷可能 由O—H 三倍频伸缩振动导致。





2.2 可溶性固形物含量分析

表1为六个方向的校正集和预测集可溶性固形物的范围,标准偏差和平均值。从表1可看出含115个样本的校正 集的可溶性固形物含量的范围大于含35个样本的预测集的 可溶性固形物含量的范围。

表 1 校正集和预测集中可溶性固形物 SSC 含量的范围、标准偏差和平均值

 Table 1
 Range, standard deviation and average value of SSC content in calibration set and prediction set

Parameter	Data set	Samples	Mean	S. D	Range
SSC/	Calibration	115	12.06	1.05	9.53~14.70
(°Brix)	Prediction	35	12.09	0.93	9.60~13.37

2.3 局部模型建立及验证

以 A1 方向为例, 用在 A1 方向采集的 115 个校正集光 谱数据建立模型,再用另外35个预测集数据对其进行局部 验证。其他五个方向同理建立模型并进行验证。每种预处理 方法侧重的方面都有所不同,因此对原始光谱用不同的预处 理方法会有不同的效果,为了有效减少光谱的噪声并提高信 噪比,使用 SGS 对六个局部模型进行预处理;考虑到水果内 部不均易导致散射效应,采用 MSC 对六个局部模型进行预 处理;由于光谱中的随机噪声,选用 GFS 对六个局部模型进 行预处理。表 2 记录了使用 SGS, MSC 和 GFS 预处理后结 合 PLSR 建立的模型以及预测结果。可以发现,六个方向的 原始数据使用 SGS 和 MSC 预处理后建立的 PLSR 模型, 它 们的 R。和 R。都会有所减小,校正集和预测集的 RMSE 大 多都会变大;当使用 GFS 分别对六个方向原始数据预处理 后,A1,A2,A4,A5和A6的 R_c 会变小, R_p 会变大,这就 使得校正集和预测集的相关系数更接近, 拟合性更好。因 此,结果表明通过 GFS 预处理后建立的六个局部模型比用 其他两种预处理后建立的结果更好。

将经过 GFS 预处理后的光谱建立的六个局部模型及验证结果进行比较,可以发现在 A3 方向的预测效果最好,它 的 R_e 和 R_p 分别是 0.913 和 0.872, RMSEC 和 RMSEP 分别 为 0.397 和 0.342。A6 方向的局部模型的局部预测效果最

差,它的 R。和 R。分别是 0.831 和 0.794, RMSEC 和 RM-SEP 分别为 0.534 和 0.508。结果表明,贡梨 SSC 局部模型 在相同方向上的预测效果非常好。但是在线检测时,如果没 有机构或人工摆正水果的位置,水果在经过检测装置时不可 能只是单一方向的。因此,要想证明局部模型该情况下的效 果,还需要看其他五个方向的预测集对其进行验证的效果。

表 2 六个方向的局部模型与局部预测效果 Table 2 Local model and local prediction effect in six directions

0	Destassion	Calib	oration	Pred	iction
Orientation	Fretreatment	$R_{ m c}$	RMSEC	$R_{ m p}$	RMSEP
A1	Raw	0.916	0.406	0.870	0.456
	SGS	0.870	0.497	0.873	0.453
	MSC	0.886	0.468	0.839	0.503
	GFS	0.898	0.443	0.873	0.452
A2	Raw	0.987	0.151	0.898	0.327
	SGS	0.955	0.280	0.869	0.374
	MSC	0.938	0.327	0.879	0.398
	GFS	0.982	0.175	0.897	0.331
A3	Raw	0.922	0.377	0.919	0.363
	SGS	0.900	0.424	0.856	0.339
	MSC	0.913	0.397	0.885	0.315
	GFS	0.913	0.397	0.872	0.342
A4	Raw	0.961	0.278	0.871	0.399
	SGS	0.922	0.387	0.883	0.375
	MSC	0.965	0.262	0.865	0.402
	GFS	0.950	0.311	0.880	0.386
A5	Raw	0.945	0.301	0.863	0.384
	SGS	0.892	0.414	0.846	0.407
	MSC	0.932	0.333	0.896	0.344
	GFS	0.930	0.337	0.859	0.389
A6	Raw	0.835	0.527	0.794	0.508
	SGS	0.825	0.542	0.794	0.509
	MSC	0.910	0.402	0.762	0.542
	GFS	0.831	0.534	0.794	0.508

2.4 全集模型的建立及验证

使用 A1—A6 六个方向上的 690 个校正集数据建立全局 模型,各方向的预测集分别对全局模型进行验证。然而从表 3 和表 4 中可以看到,经过 GFS 预处理后建立的 PLSR 模 型,校正集与各个方向的预测集的拟合性变好;校正集的 R_c 为 0. 828, A1, A2, A3, A4, A5 和 A6 的 R_p 分别为 0. 818, 0. 765, 0. 799, 0. 821, 0. 794 和 0. 824,校正集的 RMSEC 为 0. 424, RMSEP 分别为 0. 446, 0. 525, 0. 478, 0. 538, 0. 486 和 0. 619,可以看出全局模型对各个方向都有较好的预测 效果。

为了进一步评估全局模型对全方向模型的预测效果,把 六个方向的预测集合成一个预测集,再对全集模型进行验证,该预测集的 R_p和 RMSEP 分别为 0.748 和 0.548,结果 表明依旧有较好的预测效果。图 3 为贡梨 SSC 测量真值与预 测值的散点图。

	Tuble e	Giobai	mouel une	i prediction	critect or		The unice	.10115					
Orientation	Pretreatment		Calib	Coliberation		Prediction							
		LVs	Cambration		A1		A2		A3				
			$R_{ m c}$	RMSEC	$R_{ m p}$	RMSEP	$R_{ m p}$	RMSEP	$R_{ m p}$	RMSEP			
Ommidianational	Raw	16	0.844	0.406	0.810	0.497	0.766	0.523	0.792	0.479			
Omnidirectional	GFS	16	0.828	0.424	0.818	0.446	0.765	0.525	0.799	0.478			

表 3 全局模型以及 A1, A2 和 A3 方向预测效果 Table 3 Global model and prediction effect of A3, A4 and A5 directions

表 4 全局模型以及 A4, A5 和 A6 方向预测效果 Table 4 Global model and prediction effect of A4, A5 and A6 directions

Orientation			Calib	ration	Prediction							
	Pretreatment	LVs	Cambration		A4		A5		A6			
			$R_{ m c}$	RMSEC	$R_{ m p}$	RMSEP	$R_{ m p}$	RMSEP	$R_{ m p}$	RMSEP		
Omnidirectional	Raw	16	0.844	0.406	0.801	0.538	0.785	0.492	0.821	0.612		
	GFS	16	0.828	0.424	0.821	0.538	0.794	0.486	0.824	0.619		



2.5 局部模型与全集模型的对比

用六个方向的预测集数据分别对 A1, A2, A3, A4, A5 和 A6 的局部模型和全局模型进行验证。如表 5 和表 6, 局部 模型虽然对本方向的预测效果很好,但是对于其他方向的预

测效果却很差,例如局部模型 A1 预测 A2, A3, A4 和 A5 的 RMSEP 分别为 1.431, 0.816, 0.767 和 1.015; 局部模型 A2 预测 A1, A4, A5 和 A6 的 R_p 均在 0.1 以下, 不可信赖; 其 他局部模型均有类似情况。在线检测时如果没有人工对水果 进行方向调整或者没有设计制造调整方向的摆正机构,那么 局部模型是不具备在线检测能力的,局部模型在验证其他方 向时所表现出来的鲁棒性太差。六个方向的预测集对全局模 型进行验证的效果虽不如局部模型预测本方向时的验证效 果,但是效果都还不错,全局模型的校正集 R。为 0.828, RMSEC为 0.424; A1, A2, A3, A4, A5 和 A6 方向的预测 集分别为 0.818, 0.765, 0.799, 0.821, 0.794 和 0.824, RMSEP 分别为 0.446, 0.525, 0.478, 0.538, 0.486 和 0.619;结果表明六个方向的 R_p 与 R_c 均比较接近且均在 0.800 左右, RMSEC 与 RMSEP 均在 0.500 左右, 因此全局 模型在验证六个方向时是鲁棒的。但是总体来看 R 都只有 0.8 左右,这可能是因为贡梨六个方向上的 SSC 含量分布不 均匀,而本实验所测得的 SSC 含量真值为贡梨整果的平均 SSC含量真值。

表 5 局部模型和全局模型对 A1, A2, A3 方向预测集的预测效果 Table 5 Prediction effects of local model and global model on A1, A2 and A3 direction validation sets

		Vs Calibration		Prediction							
Orientation	LVs			Al		A	42	A3			
		$R_{ m c}$	RMSEC	$R_{ m p}$	RMSEP	$R_{ m p}$	RMSEP	$R_{ m p}$	RMSEP		
Omnidirectional	16	0.828	0.424	0.818	0.446	0.765	0.525	0.799	0.478		
A1	13	0.898	0.443	0.873	0.452	0.182	1.431	0.731	0.816		
A2	14	0.982	0.175	NA	0.336	0.897	0.331	0.765	0.914		
A3	9	0.913	0.397	NA	13.972	NA	3.493	0.872	0.342		
A4	12	0.950	0.311	0.328	0.410	0.532	1.735	0.814	0.545		
A5	14	0.930	0.337	NA	0.134	0.552	1.058	0.797	0.633		
A6	7	0.831	0.534	0.348	0.330	NA	1.133	0.794	0.056		

注: NA 为"NOT Available"(下同)

Note: NA represents "NOT Available" (the same below)

表 6 局部模型和全局模型对 A4, A5 和 A6 预测集的预测效果

6	Prediction	effects o	of local	model	and	global	model	on A4	I, A5	and	A6	direction	validation	sets
---	------------	-----------	----------	-------	-----	--------	-------	-------	-------	-----	----	-----------	------------	------

		Calibration		Prediction							
Orientation	LVs			1	44	1	A 5	A6			
		$R_{ m c}$	RMSEC	$R_{ m p}$	RMSEP	$R_{ m p}$	RMSEP	$R_{ m p}$	RMSEP		
Omnidirectional	16	0.828	0.424	0.821	0.538	0.794	0.486	0.824	0.619		
A1	13	0.898	0.443	0.684	0.767	0.490	1.015	0.372	0.089		
A2	14	0.982	0.175	NA	1.196	NA	2.024	NA	1.618		
A3	9	0.913	0.397	0.714	0.791	0.471	1.564	0.709	0.820		
A4	12	0.950	0.311	0.880	0.386	0.153	2.144	0.190	1.946		
A5	14	0.930	0.337	0.752	0.650	0.859	0.389	0.722	0.659		
A6	7	0.831	0.534	0.728	0.632	0.827	0.537	0.794	0.508		

3 结 论

采用本课题组自主研发的水果内部品质近红外漫透射光 谱在线检测装备采集 150 个贡梨六个方向的光谱,研究了光 谱测量方向的变化对贡梨 SSC 在线测定的影响。使用 GFS 预处理后通过 PLSR 建立的模型效果都比较好,局部模型能 够很好的预测本方向的贡梨 SSC,全局模型可以较好的预测 全方向的贡梨 SSC。通过建立局部模型和全局模型并分别用 六个方向的预测集对它们进行验证,结果表明样品的检测方 向极大地影响了贡梨 SSC 的预测精度,局部模型的鲁棒性 差;全局模型与局部模型相比,基于所有方向的全局模型获 得了更准确的预测结果,六个方向的 R_p 与 R_e 均比较接近且 均在 0.800 左右, RMSEC 与 RMSEP 均在 0.500 左右,鲁棒 性更好。因此,全局模型将有助于消除光谱测量中测量方向 的影响,实现贡梨 SSC 的整体方向在线检测,这对近红外在 线检测工业化有实际意义。

References

- [1] Gabrieli Alves de Oliveira, Sylvie Bureau, Catherine Marie-Geneviève Claire Renard, et al. Food Chemistry, 2014, 143: 223.
- [2] QIN Shan-zhi, CHEN Bin, LU Dao-li, et al(秦善知,陈 斌,陆道礼,等). Jiangsu Agricultural Sciences(江苏农业科学), 2014, 42 (8); 284.
- [3] JIANG Shui-quan, SUN Tong(江水泉,孙 通). Food and Machinery(食品与机械), 2020, 36(2): 89.
- [4] Lourdes Salguero-Chaparro, Vincent Baeten, Ouissam Abbas, et al. Journal of Food Engineering, 2012, 112(3): 152.
- [5] MENG Qing-long, SHANG Jing, HUANG Ren-shuai, et al(孟庆龙,尚 静,黄人帅,等). Food and Fermentation Industry(食品与发酵工业), 2020, 46(19): 205.
- [6] Guo Zhiming, Huang Wenqian, Peng Yankun, et al. Postharvest Biology and Technology, 2016, 115: 81.
- [7] MO Bao-qing, ZHU Hui-juan, XIAO Li(莫宝庆,朱慧娟,肖 黎). Chinese Journal of Health Inspection(中国卫生检验杂志), 2012, 22 (1): 125.
- [8] Fan Shuxiang, Zhang Baohua, Li Jiangbo, et al. Biosystems Engineering, 2016, 143:9.
- [9] Xia Yu, Huang Wenqian, Fan Shuxiang, et al. Journal of Food Process Engineering, 2018, 41(8): e12915.
- [10] Xia Yu, Huang Wenqian, Fan Shuxiang, et al. Infrared Physics and Technology, 2019, 97: 467.
- [11] HAN Dong-hai, ZHOU En-yang, QI Shu-ye(韩东海,周恩洋,威淑叶). Transactions of the Chinese Society for Agricultural Machinery (农业机械学报), 2015, 46(11): 260.
- [12] SUN Xu-dong, LIU Yan-de, LI Yi-fan, et al(孙旭东,刘燕德,李铁凡,等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2016, 47(1): 227.
- [13] 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.
- [14] Liu Yande, Zhang Yu, Jiang Xiaogang, et al. Vibrational Spectroscopy, 2020, 111: 103152.
- [15] HUANG Jian-jun, LI Xue-mei, TENG Hong-quan(黄建军,李雪梅,滕宏泉). Disaster Science(灾害学), 2021, 36(2): 60.

Table

Robustness of Global Model of Soluble Solids in Gongli Pear Based on Near-Infrared Spectroscopy

LIU Yan-de, LIAO Jun, LI Bin, JIANG Xiao-gang, ZHU Ming-wang, YAO Jin-liang, WANG Qiu

School of Electromechanical and Vehicle Engineering, East China Jiaotong University, Institute of Intelligent Electromechanical Equipment Innovation, Nanchang 330013, China

Abstract Gongpear is a popular fruit. In order to study the influence of different detection directions on the online detection of soluble solid SSC in Gongpear by NIR, a global model was proposed, and its robustness was analyzed. The spectra were collected from Gongpears in six directions: stem-calyx axis vertical, stem-upward (A1) and stem-downward (A5), between the stem-calyx axis and horizontal 45°, stem-upward sloping (A2) and stem-downward sloping (A4), stem-calyx axis horizontal, stem-right light oriented (A3), stem-calyx axis horizontal, stem-band moving direction (A6). The 150 samples with SSC ranging from 9.53 to 14. 70 were divided into 115 calibration sets with a standard deviation of 1.05 and 35 prediction sets with a standard deviation of 0.93. Six local models and one global model were established by partial least-squares regression (PLSR). The local models were established by partial least-squares regression (PLSR) after 115 calibration sets of data in each direction were preprocessed by Savitzky-Golay convolution smoothing, Multiple Scattering Correction (MSC) and Gaussian Filtering Smoothing (GFS). The local model established by the local correction set was used to verify the data of 35 prediction sets in the local direction. Compared with the PLSR model established by the three pretreatment methods, the results showed that the model established by GFS processing had the best validation effect. Therefore, the PLSR model established by GFS processing was used for all the six local and global models. The global model is a Gongpears SSC model established by PLSR after GFS pretreatment from 690 calibration sets of spectral data in A1, A2, A3, A4, A5 and A6. The prediction sets in each direction verified the seven models. The verification results showed that although the prediction effect of the local model was stronger than that of the global model in the local direction, it could not be verified in other directions and the robustness was poor. Therefore, different detection directions had a great influence on the prediction effect. The global model can accurately predict the SSC of Gongpears pear in each detection direction. The global model's correlation coefficient (R_c) is 0.828, and the root mean square error RMSEC is 0.424. The correlation coefficients (R_p) of A1, A2, A3, A4, A5 and A6 prediction sets were 0.818, 0.765, 0.799, 0.821, 0.794 and 0.824, and the root mean square errors RMSEP were 0.446, 0.525, 0.478, 0.538, 0.486 and 0.619, respectively. The $R_{\rm p}$ and $R_{\rm c}$ in six directions are close to each other and are around 0.800, the RMSEC and RMSEP are around 0. 500. The results show that the global model has excellent robustness in detecting gongpear SSC in different directions.

Keywords Near-infrared; Gongpear; Soluble solid SSC; Global model; Robustness

(Received Jul. 17, 2021; accepted Oct. 10, 2021)