

## 西瓜光透射规律与品质属性的内在联系

李 雄, 刘燕德\*, 孙旭东, 欧阳爱国, 姜小刚, 王观田, 欧阳玉平

华东交通大学机电与车辆工程学院, 江西 南昌 330013

**摘 要** 果品内部化学基础信息与光谱信息较好对应是提高模型的关键。大量的皮厚、体积大的水果果肉在可见-近红外区域透光性差、光折射角度难以确定、化学基础信息获取不准, 导致品质预测效果差。以西瓜为研究对象, 对西瓜不同区域的可溶性固形物与品质属性的内在联系进行探讨。水果市场购买 360 个西瓜样品, 在线检测装置采集光谱时参数用两种设置: 积分时间 100 ms、电流 8.0 A 和积分时间为 150 ms, 电流 8.15 A, 后者的光谱吸收峰值强度更高。西瓜可溶性固形物含量测定时, 将西瓜分为 8 份, 分别测量心糖、中糖、外围糖、底边糖(SSC)和混合糖(SSC)的平均值, 西瓜内部不同区域可溶性固形物有较大的差异, 果中心的心糖值最高, 而越靠近瓜皮区域的糖度值越低。以西瓜不同区域可溶性固形物为因变量, 卷积平滑(S-G)降解光谱噪声后的光谱为自变量建立可溶性固形物偏最小二乘预测模型, 建模集 270 个, 预测集 90 个。对比模型发现, 提高分选装置的积分时间和卤钨灯电流, 可以增加可溶性固形物模型预测精度; 局部区域的可溶性固形物作为模型的因变量预测效果也高于混合糖为因变量建立的模型。由于可见-近红外入射后在瓜果内部发生一定角度的折射、光停留在浅层区透射等原因, 靠近瓜皮的底边糖区域表征了较多的西瓜果肉信息, 建模效果最佳, 预测集相关系数为 0.89, 均方根误差为 0.24, 建模集相关系数为 0.96, 均方根误差为 0.18。而中糖、外围糖等具有一定深度且在光的直线区域表征的西瓜果实内部信息较少, 建模效果较差。因此西瓜底边糖为最佳的可溶性固形物采集区域。研究结果揭示了水果光散射规律特征及其与品质属性的内在联系, 可供实现光谱数据库和分析模型的在线更新参数。

**关键词** 西瓜; 可见-近红外; 装置参数; 可溶性固形物; 预测模型

**中图分类号:** S24 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)10-3265-06

### 引 言

中国是西瓜生产和消费最大的国家, 而西瓜内部含有人体所需的各种营养成分, 具有良好的保健功效及很高经济价值, 随着人们现代生活水平的提高, 人们对西瓜的内在品质提出了更高的要求。现有西瓜品质选择主要依靠经验判断, 如瓜的外观纹路、颜色等, 缺乏有效的西瓜无损分选方法。近红外光谱检测技术在苹果、梨等水果品质方面已经得到广泛的应用, 因此, 基于近红外检测技术在线检测西瓜的内部品质探索研究具有可行性<sup>[1-3]</sup>。

国内外研究人员应用近红外光谱检测技术对水果内部品质进行了大量的动态在线检测研究, Zhang 等<sup>[4]</sup>联合近红外和高光谱技术, 采用 CARS-SPA 模型提取 1 000~2 500 nm

的有效波长, 建立不同产地苹果可溶性固形物分级的通用数学模型, 相关系数达 0.917。Guo 等<sup>[5]</sup>采用 SPA 算法有效筛选近红外光谱, 建立猕猴桃可溶性固形物无损检测模型, 获得较好的测定效果。欧阳爱国等<sup>[6]</sup>研究不同变量筛选方法对近红外光谱检测不知火杂柑可溶性固形物模型影响, 结果表明光谱预处理方法 MSC 结合偏最小二乘方法缩短了建模时间, 改善了模型精度。

以上的研究, 都是对皮薄、体积小形状规则的水果, 以整个果品混合糖为因变量建立相关检测模型, 研究的重点是模型算法的优化。然而化学基础信息与光谱信息的对应才是提高模型的关键。针对水果市场上大量的皮厚、体积大的水果, 果肉区域可见-近红外透光性差、光线折射区域难以确定、化学基础信息采集区域不当等问题, 导致的品质信息预测效果差, 以西瓜为研究对象, 探讨西瓜不同区域的可溶性

收稿日期: 2019-08-23, 修订日期: 2019-12-19

基金项目: 国家自然科学基金项目(31760344), 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), 江西省优势科技创新团队项目(20153BCB24002)资助

作者简介: 李 雄, 1992年生, 华东交通大学机电与车辆工程学院博士研究生 e-mail: 1598657020@qq.com

\* 通讯联系人 e-mail: jxliuyd@163.com

固形物对西瓜品质检测的影响,揭示西瓜光透射规律特征与品质属性的内在联系。

## 1 实验部分

### 1.1 试验材料

试验所用西瓜购自江西南昌市,共 360 个大小相近品种为麒麟瓜的实验样品,用蒸馏水将外部擦洗干净,编号后在室内温度 25 °C 存放 1 d 后采集光谱,麒麟瓜的重量、可溶性固形物的理化指标如表 1 所示,其中 270 个西瓜样品为建模集,90 个样品为预测集。

表 1 建模集与预测集西瓜样品可溶性固形物统计结果

Table 1 Statistical results of soluble solids in modeling set and prediction set

组别	建模集	预测集
数量 N	270	90
重量/kg	2.48	5.09
可溶性固形物范围 RS/%	12.7~4.8	11.9~4.8
可溶性固形物平均值 AVG/%	9.45	9.74
可溶性固形物标准差 SD	0.85	0.84

### 1.2 光谱采集

西瓜内部的光谱采集是通过在线检测装置来完成的,在线检测装置结构如图 1 所示,包含了检测系统、控制系统和传输系统,具体零件有触发装置、计步盘、齿轮、链条、光纤、暗室、卤素灯、光谱仪、PLC(programmable logic controller)、样品杯、触发器、分级出口、电动机等。本装置通过各个机械联合运作完成西瓜的光谱采集。图 1 中的 7 所示为两个卤钨灯,为西瓜的漫透射方式采集西瓜光谱信息提供光源,透过西瓜内部的光被 8 所示的光纤接收传输至电脑。西瓜以 5 个/s 的传输速度通过检测系统,卤素灯照亮经过的西瓜,西瓜每经过一次触发 Ocean 光谱仪配置的 SpectraSuit 软

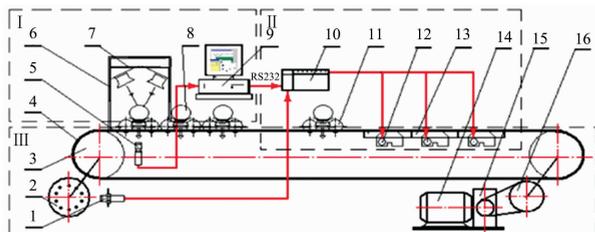


图 1 西瓜可溶性固形物分选装置

I: 检测系统; II: 控制系统; III: 传送系统

1: 触发装置; 2: 计步盘; 3: 齿轮; 4: 链条; 5: 光纤; 6: 暗室;

7: 卤素灯; 8: 样品; 9: 光谱仪; 10: PLC; 11: 样品杯;

12: 触发器; 13: 分级出口; 14: 电动机

Fig. 1 The selection device of SSC for fruit

I: Detection; II: Control; III: Transfer;

1: Trigger; 2: Step counter; 3: Gear; 4: Chain;

5: Fiber; 6: Darkroom; 7: Halogen lamp; 8: Sample;

9: Spectrometer; 10: PLC; 11: Sample cup;

12: Trigger; 13: Graded export; 14: Motor

件采集样品的光谱能量值,并传输至计算机保存。光谱采集前置装置预热 30 min,卤钨灯光照强度为 100 W,在线检测装置积分时间为 100 ms(额定积分时间)和 150 ms 两种设置,西瓜果柄朝上,底部果蒂在正下方,每个西瓜划分成 8 个面,采集 8 条光谱,取其中 4 条光谱的平均值为区域光谱,另外 4 条光谱的平均值为混合糖度光谱。

### 1.3 可溶性固形物含量的测定

切下西瓜相应区域的果肉,挤出适量的西瓜汁滴入到折射式数字糖度计(PR-101a)测量可溶性固形物,每次测量后用纯净水标定糖度 0% 后继续测量下一个点,每个西瓜切下区域相应位置的平均数作为该区域的可溶性固形物值<sup>[7]</sup>。

### 1.4 数据分析

获取的西瓜光谱和对应区域的可溶性固形物输入到 MATLAB 和 Unscrambler 软件,处理后建立相应的预测模型,选取最佳的模型输入到在线检测装置可实现西瓜的糖度在线分级。模型的优劣由相关系数和均方根误差来评价。由于不同区域的可溶性固形物差异较大,所以利用不同区域的可溶性值建立预测模型,寻找较高的相关系数和较小的均方根误差模型<sup>[8]</sup>。

### 1.5 可见-近红外光谱的采集及预处理

#### 1.5.1 西瓜平均光谱特性分析

图 2 所示是西瓜在分选装置两种参数下采集的光谱,分选装置旧参数为积分时间 100 ms,电流 8.0 A,新参数的积分时间为 150 ms,电流 8.15 A,电流是指卤钨灯的电流强度,实验装置的卤钨灯最大承受电流 8.3 A,长期的实验表明当卤钨灯通常在 8.0 A 的电流下得到稳定,如果电流过高容易烧毁灯泡和损伤果品,电流过低采集的光谱强度不够;大胆尝试了 8.15 A 的卤钨灯电流作为新参数进行试验。积分时间是表示光谱在水果内部的照射时间,长期的实验表明,采集果品的积分时间设置为 100 ms(旧参数)能够满足光谱采集的信息,本实验以 50 ms 为一个梯度,增加一个梯度的积分时间(新参数)进行新旧参数的对比。图中可以看出新参数在波长相同的范围内,吸光度的强度更高。在卤钨灯的额定电流环境下,新参数增加了 0.15 A 的电流,从而增加卤钨灯的光照强度,光纤获取的吸光度则更强,此外,新参数增加了光谱在水果内部品质的光照时间,获取更全面的水果内部信息。

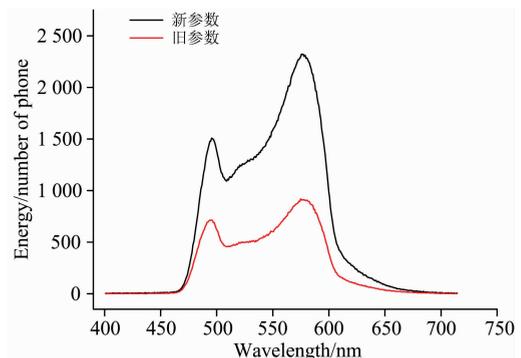


图 2 西瓜近红外漫透射原始光谱

Fig. 2 Near-infrared diffuse transmission original spectra of watermelon

西瓜内部含有各种复杂的化合物, 这些化合物基本是由 C—H, N—H 和 O—H 等官能团组成。图中显示, 在 460~620 nm 波段范围内出现了不同层次的波峰和波谷, 在波长为 496 nm 的吸收峰可能是因为电子跃迁导致, 577 nm 吸收峰可能是因为 N—H 官能团伸缩振动和变形振动一级倍频导致<sup>[9]</sup>。

### 1.5.2 近红外光谱预处理

为了提高 PLS 预测模型的精准度, 减少非实验因素导致的误差, 需要消除噪声、识别奇异性样本、增强光谱信号。对吸光度光谱值进行卷积平滑(S-G), 卷积平滑通过最小二乘拟合的多项式移动窗口数据加权平均来降解光谱噪声, 从而提高信噪比, 保留西瓜吸光度光谱有用的信息, 如式(1)所示

$$\alpha_{\lambda, \text{smooth}} = \bar{\alpha}_{\lambda} = \frac{1}{C} \sum_{\delta=-\omega}^{\omega} \alpha_{\lambda+\delta} \quad (1)$$

其中,  $c_{\delta}$  是最小二乘原理多项式拟合系数,  $C = \sum_{\delta=-\omega}^{\omega} c_{\delta}$  为归一化因子数,  $\alpha$  是西瓜样品反射率,  $c_{\delta}$  和与每个实际光谱值相乘的结果, 降低了噪声的干扰。

## 2 结果与讨论

### 2.1 可溶性固形物区域选择

与其他皮薄、体积小的水果不同, 西瓜内部结构复杂、体积大、可溶性固形物内部不均匀, 不同区域的可溶性固形物有较大的差异。图 3 所示是西瓜光谱采集和糖度采集区域划分。西瓜分为 8 份, 为了探究可见-近红外光在西瓜内部透射路径, 寻找与光谱相关性最高的区域糖度, 西瓜底部做圆形标记, 光纤头放置在此区域接收光谱, 其中 A, B, C 和 D 区域切开后将分别测量并求出心糖、中糖、外围糖、底边糖平均值, 而 1, 2, 3 和 4 等区域切开后将整块西瓜挤汁, 求综合糖度值。近红外光透过果肉传输给西瓜底部光纤, 采集西瓜光谱信息。

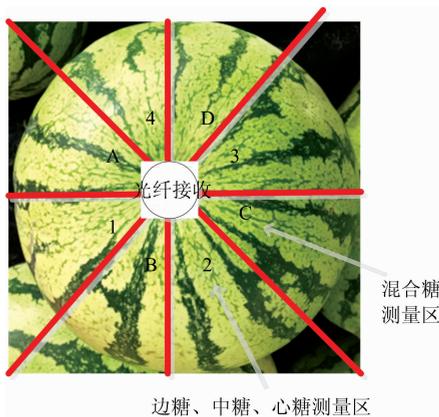


图 3 西瓜光谱采集和糖度采集区域划分

Fig. 3 Watermelon is divided into several regions for spectral collection and sugar content collection

图 4 所示为西瓜由果中心向外延伸可溶性固形物的划分, 第 1 个 2 cm 区域的糖称之为心糖, 第 2 个 2 cm 区域的糖为中糖、第 3 个 2 cm 区域的糖为外围糖, 最后约 2 cm 区域的糖为底边糖, 距离西瓜中心最远的区域为白色瓜皮, 本实验不研究瓜皮可溶性固形物。如图 5 所示为心糖、中糖、外围糖和底边糖的变化趋势, 其中心糖、中糖、外围糖、底边糖区域的可溶性固形物的平均值分别为 11.27, 11.05, 10.66 和 6.07 brix, 整块西瓜一起榨汁计算出的混合糖(SSC)平均值为 9.9 brix。西瓜不同区域可溶性固形物(SSC)有明显的差异, 距离果中心最近的心糖值最高、其次是中糖、外围糖、底边糖, 越靠近瓜皮区域的糖度值越低。由此推测, 西瓜糖度值选择不当, 对模型精度的影响将较大。

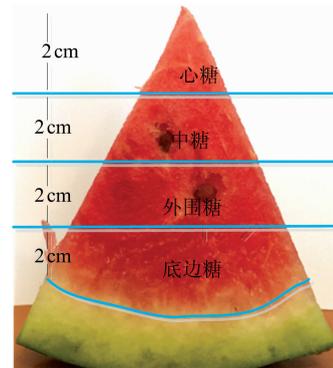


图 4 可溶性固形物位置划分

Fig. 4 Location of soluble solids

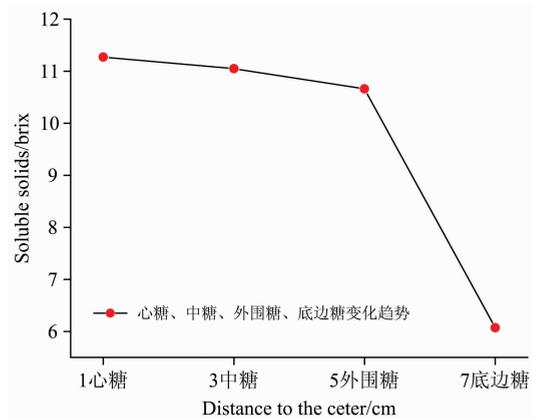


图 5 西瓜可溶性固形物含量随深度变化趋势

Fig. 5 The content of soluble solids in watermelon varies with depth

### 2.2 西瓜糖度区域对模型精度的影响

分别将西瓜按照图 5 的方式切开后, 测量西瓜内部的心糖、中糖、外围糖和底边糖(SSC)等化学值矩阵, 数据导入到 Matlab 数据处理软件中, 分为建模集 270 个, 预测集 90 个样品建立偏最小二乘(PLS)数学模型, 建模前通过程序实现光谱的预处理。如表 2 所示是以不同区域可溶性固形物为因变量, 接收的光谱值为自变量建立的 PLS 模型结果。

表 2 不同区域可溶性固形物建模结果对比

Table 2 Comparison of modeling results of soluble solids in different regions

装置参数	可溶性固形物区域	主成分因子数	$R_p$	RMSEP	$R_c$	RMSEC	斜率
旧参数	心糖	6	0.84	0.28	0.95	0.12	9.87
	中糖	7	0.74	0.56	0.87	0.34	9.30
	外围糖	8	0.74	0.75	0.92	0.34	7.83
	底边糖	7	0.87	0.57	0.96	0.27	5.6
	混合糖	7	0.86	0.44	0.92	0.89	9.09
新参数	心糖	8	0.80	0.25	0.87	0.21	10.25
	中糖	9	0.80	0.31	0.92	0.22	10.20
	外围糖	7	0.70	0.56	0.84	0.42	8.75
	底边糖	8	0.89	0.24	0.96	0.18	3.4
	混合糖	8	0.82	0.34	0.95	0.24	4.23

通过对比不同区域可溶性固形物(SSC)的建模结果发现,在额定功率的前提下,适当提高分选装置的积分时间和卤钨灯电流,整体会上提高可溶性固形物的无损检测模型的预测精度。可能是因为增加了光强,提高了水果内部的通透性,也增加了光在水果内部的停留时间,水果内部获取了较多的信息。而对比心糖、中糖、外围糖和底边糖(SSC)等区域的化学值建模结果,发现如果以常规的方式切下的西瓜挤汁得到的混合可溶性固形物作为模型输入自变量得到的效果并不是最佳,由于西瓜果肉多,取局部区域的可溶性固形物作为模型的因变量可适当的提高模型精度。在西瓜可溶性固形物区域选择方面,区域的选择不同导致的建模相关系数会有较大的差异。选取底边糖(SSC),最靠近西瓜皮的 2 cm 果肉的可溶性固形物作为模型的因变量,建模效果最好。其次是最靠近果肉区域的可溶性固形物也就是心糖(SSC)作为模型的因变量模型效果较好,而外围糖和中糖的效果最差。

如图 6 所示是分选装置新参数采集光谱后,以西瓜底边糖(SSC)为因变量建立的可溶性固形物偏最小二乘(PLS)预测模型散点图,360 个西瓜样品划分好 90 个预测集和 270 个建模集。预测集相关系数为 0.89,预测集均方根误差为 0.24,建模集相关系数为 0.96,建模集均方根误差 0.18。图 7 为西瓜 PLS 可溶性固形物定量模型的主成分因子数图,建模时选择较低的主成分因子(PC)会覆盖西瓜果肉光谱中的有效信息,降低定量模型精度,例如选择 PC 为 5 或者 6 时出现“欠拟合”,而 PC 为 9 或者 10 会出现“过拟合”,最佳的

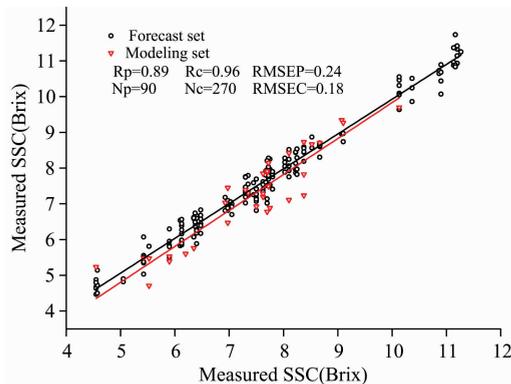


图 6 PLS 模型建模散点图

Fig. 6 PLS model modeling scatter plot

PC 为 8,预测精度达到最优,所以本研究西瓜可溶性固形物定量模型的最佳主成分数为 8。

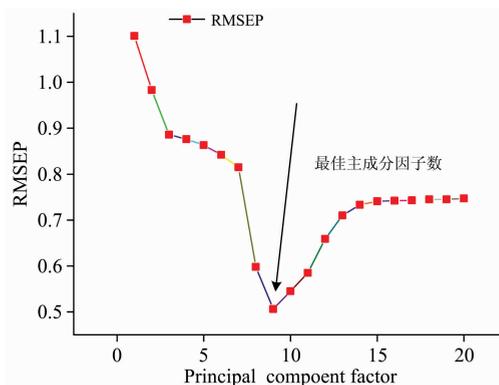


图 7 主成分因子数和均方根误差

Fig. 7 Number of principal component factors and root mean square error

图 8 是定量模型的回归系数图,全谱中各个变量均在西瓜可溶性固形物预测中发挥了作用,而系数较大的波长点对西瓜可溶性固形物变量起的贡献率最大,例如,在 525, 535, 552, 564 和 578 nm 处的波长点对因变量的预测的作用最大,将采集到西瓜能量谱输入到模型中,经过计算可以得到西瓜的可溶性固形物。

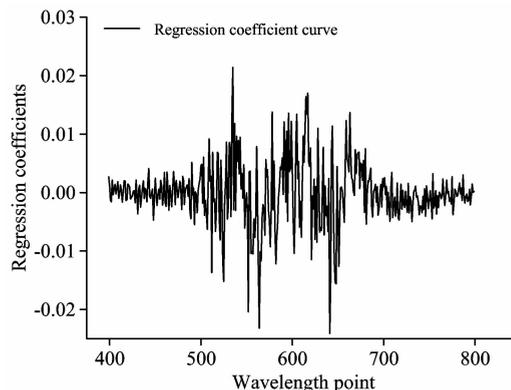


图 8 西瓜可溶性固形物预测模型回归系数图

Fig. 8 Regression coefficient diagram of watermelon soluble solids prediction model

### 2.3 光学特性分析

通过比较发现,在西瓜底部采集光谱,针对果肉内部不同区域化学信息指标检测西瓜果品可溶性固形物的效果有较大的差异。底边区域为可溶性固形物的最佳采集区域,中糖、外围糖区域效果最差,因此西瓜可溶性固形物预测建模有着一定的规律性:由于利用漫透射方式采集西瓜光谱,西瓜有 2 cm 左右的瓜皮,阻挡了部分光透射到水果内部,将可

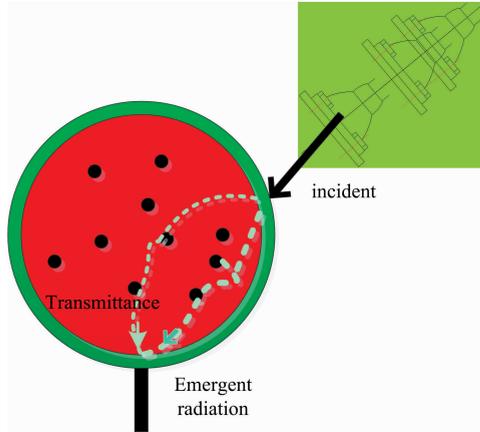


图 9 光在西瓜内部输运示意图

Fig. 9 Schematic diagram of light transmitting inside watermelon

见-近红外光停留在底边层的信息最多,底边糖的建模效果最佳。其次由于光学折射原理,入射光经过西瓜果皮后会有一定角度的折射,直线传播经过的中糖、外围糖区域表征的西瓜果实内部信息较少。如图 9 所示,是可见-近红外光谱在西瓜内部输运的示意图,可见-近红外透过细胞紧密的瓜皮后,在果肉的机械组织区域发生折射,由于西瓜光强的限制以及西瓜体积大、内部结构复杂的特点,光经过靠近果皮底边糖表征了最多的果肉信息。

### 3 结 论

果品不同区域果肉的可溶性含量有明显差异,心糖区域的可溶性固形物含量最高,靠近果皮的底边糖区域可溶性固形物最低。而可溶性固形物等化学基础信息与光谱信息较好的对应可提高预测模型精度。通过偏最小二乘算法,将西瓜内部组织分层处理,研究漫透射方式下西瓜心糖、中糖、外围糖、底边糖区域的可溶性固形物对模型精度的影响,其中底边糖区域可溶性固形物建立模型最佳,预测集相关系数达到 0.89,均方根误差为 0.24,高于其他区域模型精度。说明西瓜底边层包含了可见-近红外光谱的最大量信息,为最佳的可溶性固形物采集区域,探讨了水果光散射规律特征及其与品质属性的内在联系,为光谱数据库和分析模型的在线更新提供参数。

### References

- [1] Martí Raúl, Sánchez Guadalupe, Valcárcel Mercedes, et al. Food Chemistry, 2019, 300: 125227.
- [2] 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.
- [3] HAN Dong-hai, CHANG Dong, SONG Shu-hui, et al(韩东海,常冬,宋曙辉,等). Transactions of the Chinese Society of Agricultural Machinery(农业机械学报), 2013, 44(7): 174.
- [4] Zhang D Y, Xu Y F, Huang W Q, et al. Infrared Physics & Technology, 2019, 98: 297.
- [5] Guo W, Fan Z, Dong J. Food Analytical Methods, 2016, 9(1): 38.
- [6] OUYANG Ai-guo, WU Ming-ming, WANG Hai-yang, et al(欧阳爱国,吴明明,王海阳,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2017, 37(5): 1497.
- [7] Hu Rong, Zhang Lixin, Yu Zhiyuan, et al. Infrared Physics and Technology, 2019, 11(26): 102.
- [8] Fan Shuxiang, Huang Wenqian, Guo Zhiming, et al. Food Analytical Methods, 2015, 8(8): 1936.
- [9] Song Jie, Li Guanglin, Yang Xiaodong. Journal of the Science of Food and Agriculture, 2019, 99(11): 4898.

# Interent Relation Between the Light Transmission Law and Quality Attributes of Watermelon

LI Xiong, LIU Yan-de\*, SUN Xu-dong, OUYANG Ai-guo, JIANG Xiao-gang, WANG Guan-tian, OUYANG Yu-ping  
School of Mechatronics Engineering, East China Jiaotong University, Nanchang 330013, China

**Abstract** Corresponding to the internal chemical basic information of the fruit and the spectral information is the key to improve the model. For the current large thickness and a large volume of fruit, the visible/near infrared is poor in light transmission in the flesh region, the light refraction angle is difficult to determine, and the basic chemical information is not Accurate, resulting in poor quality prediction. In this paper, watermelon is the research object, and the intrinsic relationship between soluble solids and quality attributes in different regions of watermelon is discussed. Three hundred sixty samples are purchased in the fruit market. The parameters of the online detection device are divided into old parameters: integration time 100 ms, current 8.0 A and new parameters. The integration time is 150 ms, the current is 8.15 A to collect the watermelon spectrum, and the spectral absorption peak intensity is higher under the new parameters of the device. When the content of soluble solids in watermelon was determined, the watermelon was divided into 8 parts, and the heart sugar, medium sugar, peripheral sugar, base sugar (SSC) and mixed sugar (SSC) were determined respectively. The soluble solids in different parts of the pulp were large. The difference is that the most recent heart sugar value is the closest to the fruit center, and the lower the sugar value is, the closer it is to the melon skin area. Taking soluble solids in different regions of watermelon as the dependent variable, the new and old parameters of convolution smoothing (S-G) degraded spectral noise was used to establish a soluble solids partial least squares prediction model with independent variables, with 270 modeling sets and 90 prediction sets. The comparison model found that the new parameters improved the integration time of the sorting device and the tungsten halogen lamp current, which increased the prediction accuracy of the soluble solid model; while the soluble solids in the local area as the model's dependent variable prediction effect was also higher than the mixed sugar. The model established by the variable, the bottom sugar region of the region close to the melon skin, due to visible/near-infrared incidence, and then a certain angle of refraction inside the melon, which indicates more watermelon fruit information, the modeling effect is the best, the prediction set the correlation coefficient is 0.89, and the root mean square error is 0.24. However, the internal information of watermelon fruit characterized by medium sugar and peripheral sugar region is less, and the modeling effect is poor. Therefore, the bottom edge of the watermelon is the best soluble solids collection area. This study reveals the characteristics of fruit light scattering and its intrinsic relationship with quality attributes, and realizes online updating of spectral database and analytical model.

**Keywords** Watermelon; Visible/near infrared; Device parameters; Soluble solids; Prediction model

(Received Aug. 23, 2019; accepted Dec. 19, 2019)

\* Corresponding author