

基于光谱变换算法的不同横径沃柑的光谱与建模对比分析

Comparative analysis of spectra and modeling of different cross-diameter orah mandarin based on spectral transformation algorithm

胡峰 龚中良 文韬 代兴勇 唐子叶

HU Feng GONG Zhongliang WEN Tao DAI Xingyong TANG Ziyue

(中南林业科技大学机电工程学院, 湖南长沙 410004)

(School of Mechanical and Electrical Engineering, Central South University of Forestry and Technology, Changsha, Hunan 410004, China)

摘要:目的:减小沃柑横径对糖度预测带来的影响。方法:将132个沃柑按横径分为小等果(65~70 mm)、中等果(70~75 mm)和大等果(75~80 mm)3组。采集全体沃柑的横径光谱后,运用光谱变换算法,综合光谱信息与沃柑横径信息,将不同大小的沃柑光谱转换到同一横径基准下。将修正前的光谱与修正后的光谱分别通过预处理,目标共生距离算法(SPHY)划分,竞争性自适应权重取样法(CARS)筛选特征波长点以及偏最小二乘回归法(PLS)建立修正前后的糖度模型。结果:小等果光谱经过修正,预测集决定系数(R^2_p)由0.790提升至0.821,预测集均方根误差(RMSEP)为0.489降低至0.443;中等果光谱经过修正, R^2_p 由0.801提升为0.845, RMSEP为0.460降低至0.422;大等果光谱修正, R^2_p 由0.820提升至0.863, RMSEP为0.431降低至0.393。结论:光谱修正算法减小了沃柑横径带来的光谱差异,提升了模型的预测精度。

关键词:沃柑;横径;光谱修正;糖度

Abstract: Objective: To reduce the influence of transverse diameter of Mandarin orange on the prediction of sugar content.

Methods: 132 citrus were divided into three groups according to transverse diameter: small (65~70 mm), medium (70~75 mm) and large (75~80 mm). After collecting the transverse diameter spectrum of all citrus, the spectrum information and

transverse diameter information of citrus were synthesized by spectral transformation algorithm, and the spectrum of different sizes of citrus was converted to the same transverse diameter datum. The pre-correction and post-correction spectra were respectively preprocessed, divided by target symbiotic distance algorithm (SPHY), screened by competitive adaptive weight sampling (CARS) and selected by partial least squares regression (PLS) to establish the sugar degree model before and after correction. **Results:** The prediction set determination coefficient (R^2_p) was increased from 0.790 to 0.821, and the prediction set root mean square error (RMSEP) was decreased from 0.489 to 0.443. The middle fruit spectrum was modified, R^2_p increased from 0.801 to 0.845, RMSEP decreased from 0.460 to 0.422. R^2_p increased from 0.820 to 0.863 and RMSEP decreased from 0.431 to 0.393. **Conclusion:** The spectral correction algorithm can reduce the spectral difference caused by the transverse diameter and improve the prediction accuracy of the model.

Keywords: orah mandarin; transverse diameter; spectral correction; sugar degree

可溶性固形物是评价柑橘内部品质的重要指标之一。目前国内外学者已经将可见/近红外光谱无损检测方法应用于柑橘无损检测且取得了较好的效果,但是实际采后检测中柑橘的大小、表面颜色、检测位置、光源距离等众多因素对柑橘可溶性固形物的糖度判别造成了很大影响,其中以果实大小对柑橘可溶性固形物的透射光谱检测产生的影响最为显著^[1]。因此,消除水果尺寸差异带来的影响,对提高水果分选模型精度具有重要意义^[2]。

目前国内外众多学者在研究柑橘等其他水果的大小对可见/近红外光谱的影响中取得了一定的成就。并提

基金项目:湖南自然科学基金杰出青年基金项目(编号:2023JJ10099)

作者简介:胡峰,男,中南林业科技大学在读硕士研究生。

通信作者:龚中良(1965—),男,中南林业科技大学教授,博士。

E-mail: gzlaa@163.com

收稿日期:2023-07-21 **改回日期:**2024-03-10

出了一些修正果实大小对透射光谱影响的方法^[3]。Mc Clark 等^[4]在进行蜜橘糖度预测的研究中,发现果实赤道处测量的近红外透射光谱随着果实直径的增大而向上移动,通过将二阶导数光谱除以果实直径可以减小果实大小的影响。杜冉^[5]针对近红外透射技术在线检测苹果内部品质中果实大小的影响,安排了两组试验方案,结果表明:采用不同尺寸梯度的样品建模,可以消除水果尺寸大小对检测结果的影响,有效提高模型的稳健性。曹松涛^[6]研究了样品大小差异对梨糖度预测模型的影响及其修正方法,并提出了一种可以根据水果大小调节光源高度的检测机构设想,以确保光源光斑始终照射在水果赤道位置,提高预测精度。田世杰^[7]将直径变换法和光谱变换法应用于苹果霉心病的检测,并修正了果实大小对透射光谱的影响。雷雨等^[8-10]用果径信息参与建模,修正果实大小的方法提高了模型识别精度。河野澄夫^[11]将每个果实的光谱换算成同一直径果实的光谱,以消除果实大小带来的影响,为后续修正果实大小的研究提供了一定的思路借鉴。

研究拟采用基于消光系数的光谱修正方法,综合光谱信息与沃柑横径信息,将不同大小的沃柑光谱转换到同一果径基准下^[12]。用公式求出光在沃柑内部的消光系数,利用光在沃柑内穿透时的消光系数修正透射光谱,实现对沃柑的近红外透射光谱进行修正。建立模型并与原始光谱模型进行对比。

1 材料与方法

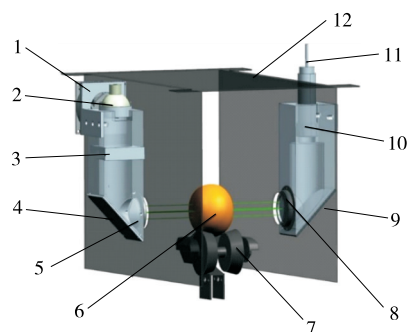
1.1 试验样本

以市售沃柑商品果为研究对象,剔除外表严重损伤的果实。用游标卡尺测量每一个沃柑的横径(各测量3次取平均)。将横径在 65~70 mm 的沃柑分为小等果;70~75 mm 分为中等果;75~80 mm 分为大等果。将不同等级果实分别从 1 开始编号。样品被分为校正组、预测组两组,其中校正组样本数 99 个,预测组样本数 32 个。

通过对全体 132 个沃柑的横径进行测量发现,小等果(65~70 mm)有 37 个;中等果(70~75 mm)有 60 个;大等果(75~80 mm)有 35 个。计算所有沃柑的横径平均值为 72.35 mm,其中位数为 72.38 mm。选取与平均值和中位数最接近的中等果第 3 号沃柑($d=72.4$ mm)作为全体沃柑修正光谱的参考样本。

1.2 试验设备与光谱采集方法

试验采用的全透射光谱照射装置如图 1 所示,其由卤素灯、照射灯筒、平面镜、凸透镜、直角准直镜、光谱仪和触发器等组成。光源选用镀金属膜的低压大功率卤素灯(JCR100W 型,12 V,日本岩崎公司)。卤素灯发出的光线经灯筒底部平面镜反射至凸透镜聚焦并准直,直射



1. 散热风扇 2. 卤素灯 3. 插拔式模块 4. 平面镜 5. 凸透镜 6. 柑橘样本 7. 双锥滚子式果杯 8. 凸透镜 9. 平面镜 10. 准直镜 11. 接收光纤 12. 黑箱箱体

图 1 光谱采集装置

Figure 1 Spectrum acquisition device

透过柑橘,透射过的光线被直角准直镜接收,通过光纤传到微型光谱仪(Ocean optics Maya 2000 Pro 型,美国 OceanInsight 公司),响应波长范围为 350~1 050 nm,分辨率 1.5 nm。

连接好光纤、数据线、光谱仪与触发件,通电光谱可移植模块。打开卤素灯,预热 30 min 后,采集聚四氟乙烯参比球作为参比光谱以及采集暗光谱。打开 Ocean View 在线采集软件,积分时间设置为 80 ms,扫描平均设为 5,滑动宽度为 12。将标好号的沃柑依次放入光谱箱中,姿态为果梗向上,使卤素灯的光线通过沃柑后被准直镜接收,采集每一个沃柑的横径静态光谱。为了获得更完整的光谱数据,每一个沃柑横径采集 3 次光谱,然后取平均光谱作为该沃柑的代表光谱。每个沃柑采集了 2 068 个波长点的光谱,由于光谱两端的信噪比很差,因此只采用 550~950 nm 波段的光谱来进行分析。

1.3 光谱修正方法

光在沃柑内部进行传输时,光程(果实直径)对光的衰减作用十分严重。如图 2 所示,光照在果实外表面的光强为 I_0 ,距离表皮距离为 d 处的光强衰减为 I 。 I 与 I_0 之间满足关系式(1)。

$$I = I_0 \exp(-u_e d), \quad (1)$$

为了进行光谱修正,将式(1)进行 Log 变换,得到式(2):

$$\ln I = \ln I_0 - u_e d, \quad (2)$$

式中:

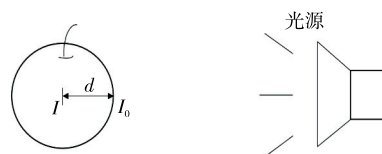


图 2 I 、 I_0 、 d 的位置

Figure 2 Position of I , I_0 , d

d ——位置距离光入射点的距离,mm;

u_e ——果实内部消光系数(或衰减系数,不同样本的消光系数不同), mm^{-1} ;

I_0 ——刚进入苹果表面时的光强值;

I ——对应深度 d 时的透射光强值。

对任意一个果实来说,如果知道了入射深度为 d_1 处的衰减光强 I_1 以及入射深度为 d_2 处的衰减光强 I_2 ,则由关系式(3)可以得到单个样本的消光系数的表达式,如式(4)所示。

$$\begin{cases} \ln I_1 = \ln I_0 - u_e d_1 \\ \ln I_2 = \ln I_0 - u_e d_2 \end{cases}, \quad (3)$$

联立可以求得:

$$-u_e = \frac{\ln I_1 - \ln I_2}{d_1 - d_2}, \quad (4)$$

式中:

d_1 ——果实内第一处的入射深度,mm;

d_2 ——果实内第二处的入射深度,mm;

I_1 ——入射深度 d_1 处的光强值;

I_2 ——入射深度 d_2 处的光强值;

u_e ——果实内部消光系数, mm^{-1} 。

如果认为 I 是光透射穿过果实后的光强,则 d 所代表的入射深度即为果实的直径。选定直径为某一数值的果实作为参考样本,其果径 d_R 对应的光强为 I_R 。

任一果实的横径为 d_i ,其对应的透射光强为 I_i 。运用式(5)可以求出所有样本的平均消光系数。

$$-u_e = \frac{\sum_{i=1}^n \left(\frac{\ln I_R - \ln I_i}{d_R - d_i} \right)}{n}, \quad (5)$$

式中:

d_R ——参考样本的横径值,mm;

I_i ——样本 i 修正前的光强值;

I_i^* ——修正后的光强值;

d_i ——样本 i 的果实大小(横径值),mm;

n ——样本总数量。

通过对式(5)进行逆运算得到式(6),可以基于消光系数来考虑果实横径大小的透射光谱修正。

$$I_i^* = \exp[-u_e(d_R - d_i) + \ln I_i], \quad (6)$$

式中:

d_R ——参考样本的横径值,mm;

I_i ——样本 i 修正前的光强值;

I_i^* ——修正后的光谱强度;

d_i ——样本 i 的果实大小(横径值),mm;

u_e ——果实内部消光系数, mm^{-1} 。

1.4 样本可溶性固形物的测定

采集光谱数据后,将样品分别剥皮并套上过滤袋,使用手动榨汁机压榨出汁。过滤袋的作用主要是减小橘络

和果肉残渣对测量结果的影响。使用的糖度计型号为 PAL-1(Atago Co,Tokyo,Japan)。使用清水标定零点后,滴取样本汁液于糖度计的检测槽内。每一个样品重复测量 3 次,取其平均值作为可溶性固形物参值。

1.5 数据处理与分析

如图 3 所示,试验中,使用 Excel 2019 来存储光谱矩阵数据,波长值与糖度值。再将波长值与糖度计导入到数学分析软件 Matlab 2016a 进行采集光谱分析。确定参考样本及其对应的参考横径,运用式(5)计算整批样本的平均消光系数,运用式(6)计算每一个沃柑的修正光强值。将原始光谱与修正光谱分别参与后续数据处理流程。

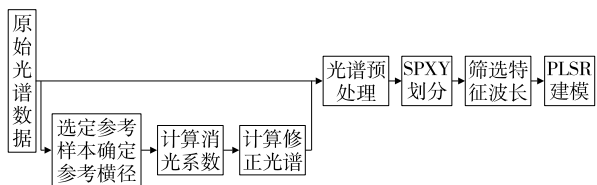


图3 数据处理分析流程图

Figure 3 Flow chart of data processing and analysis

由于光谱仪光纤探头接收到的光谱中包含各种各样的噪声,因此在数据分析之前,对光谱数据进行适当的预处理工作,以提高模型的收敛性能,对光谱分别进行标准正态变量变换、多元散射训练、移动窗口平滑、SG一阶求导等多种方式预处理,选择数据较优的预处理方法参与后续建模流程,再依据目标共生距离算法(SPXY)划分方法以 3:1 的比例进行校正集和预测集的划分,并采用竞争性自适应重加权算法筛选特征波长,最后使用建模定量分析软件 Unscrambler 中建立偏最小二乘回归法(PLS)模型,来比较各尺寸等级的沃柑修正前后所得模型的预测精度,以校正样本的相关系数(r_c)和均方根偏差(RMSEC)以及预测样本的相关系数(r_p)与均方根偏差(RMSEP)来进行评价与对比。

2 结果与讨论

2.1 样本糖度分布

图 4 为 132 个沃柑的糖度分布直方图。样本糖度范围为 $10.1 \sim 17.8$ °Brix,平均糖度为 15.36 °Brix。

2.2 修正前后的光谱对比

2.2.1 消光系数的计算 选定与整批沃柑横径平均值最相近的中等果 3 号作为参考样本。以其横径 72.4 mm 为参考横径。按式(5)求得全体沃柑的平均消光系数,其与波长的关系如图 5 所示。由图 5 可知,在 $400 \sim 600$ nm 波段,随着波长增加,平均消光系数降低;在 $600 \sim 900$ nm 波段,平均消光系数在 -0.11 左右波动;在 $900 \sim$

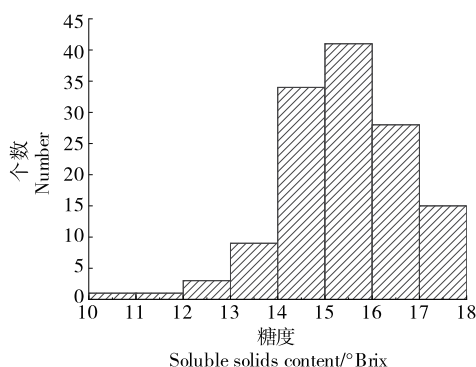


图 4 132 个沃柑的糖度分布直方图

Figure 4 Histogram of Brix distribution of 132 orah mandarins

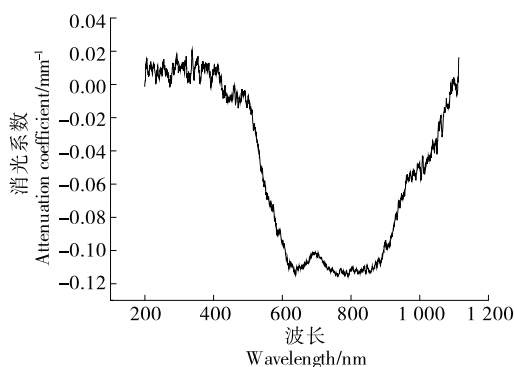


图 5 平均消光系数与波长变化趋势

Figure 5 Trend of average extinction coefficient and wavelength

1 100 nm 波段,随着波长增加,平均消光系数增大。

2.2.2 修正前后光谱 由式(6)求得全体沃柑的修正光谱。由图6可知,3个等级的沃柑透光光谱整体起伏趋势

类似,未出现明显异常情况。原始光谱与修正后的光谱都有 2 个吸收峰。其中 710 nm 和 820 nm 附近的吸收峰主要受水或 C—H 键官能团影响^[13]。

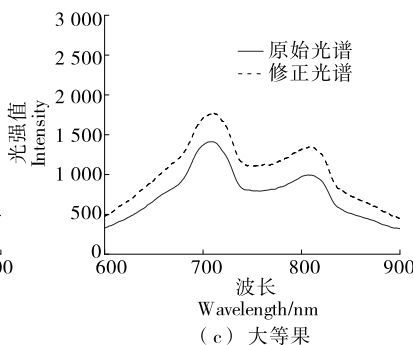
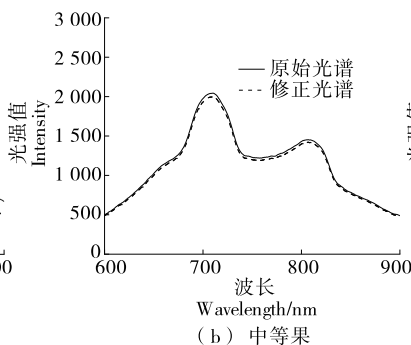
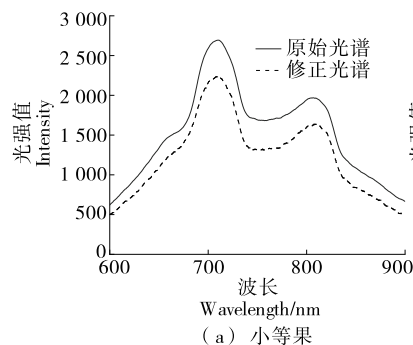


图 6 修正前后的平均光谱

Figure 6 Average spectrum before and after correction

小等果的平均原始光谱中在 710 nm 处的波峰约为 2 700,修正后约为 2 250。810 nm 处的最大波峰约为 2 000,修正后约为 1 600。修正后的光谱值均比原始光谱值低。从原理上分析,小等果的横径比参考横径要小,光透过参考样本后衰减的光强比透过小等果后衰减的要多。

中等果原始光谱与修正光谱在 710 nm 与 810 nm 处的波峰均约为 2 000。810 nm 处的最大波峰约为 1 400。中等果光谱修正前后的差异不大,因为中等果沃柑平均横径为 72.5 mm, d_R 与 d_i 相差 0.1 mm, $I_i^* \approx I_i$ 。

大等果原始光谱中在 710 nm 处的波峰大小约为 1 400,修正后约为 1 750。810 nm 处的最大波峰约为 1 000,修正后约为 1 250。修正后的光谱值比原始光谱值要高。从原理上分析,大等果的横径比参考横径要大,光透过参考样本后衰减的光强比透过大等果后衰减的要少。

对任意一个样本来说,光谱的修正是在不同的波长点处进行的。试验所采用的光谱仪采集的每一条光谱均有 2 068 个波长点。表 1~表 3 列出了不同横径的样本在

第一个波长点 $\lambda=199.583$ nm 处的原始光谱值与修正光谱值。

2.3 修正前后的模型对比

2.3.1 预处理的选择 为降低干扰信息以及散射情况对模型预测精度的影响,采用卷积平滑(SG)、多元散射校正(MSC)、标准正态变量变换(SNV)和一阶导数算法对光谱数据进行处理,其中 SG 平滑窗口宽度为 15,拟合多项式次数为 3。图 7 为小等果原始光谱的各种预处理方式的光谱图。由图 7 可知,经 MSC、SNV 预处理后的光谱相比于原始光谱更加集中,有效降低了柑橘样本由于散射、颗粒分布不均引起的光谱差异。一阶求导算法(SG-1st)放大了样本之间的差异,但经 SG-1st 处理后的光谱引入了较多的噪声。

利用处理后的数据进行偏最小二乘回归法(PLSR)建模。由表 4 可知,运用 MSC 预处理方法的预测集相关系数 R_p^2 为 0.790,略高于其他 3 种预处理方式。RMSEP 为 0.489,低于其他 3 种预处理方式。因此,选用 MSC 作为后续建模的预处理方法。

表 1 小等果在 $\lambda=199.583\text{ nm}$ 处的原始光谱与修正光谱
Table 1 The original and corrected spectra of the small isocarpous at $\lambda=199.583\text{ nm}$

序号	横径/mm	原始光谱值	修正光谱值	序号	横径/mm	原始光谱值	修正光谱值
1	67.5	139.495	145.203	20	69.3	143.650	147.341
2	67.4	140.820	146.702	21	68.9	143.895	148.076
3	69.0	142.240	146.253	22	66.2	142.235	149.639
4	68.0	139.660	144.781	23	69.0	143.230	147.271
5	69.0	142.155	146.166	24	68.4	145.060	149.887
6	68.1	140.905	145.952	25	67.4	142.155	148.093
7	68.2	141.905	146.867	26	66.3	142.565	149.863
8	67.2	141.075	147.208	27	68.5	143.320	147.968
9	69.8	141.820	144.870	28	68.2	141.320	146.262
10	69.0	139.745	143.688	29	68.6	143.650	148.188
11	67.2	143.395	149.629	30	68.5	144.645	149.336
12	68.1	141.735	146.812	31	67.5	144.895	150.824
13	69.4	141.740	145.263	32	67.4	142.320	148.265
14	69.1	142.985	146.899	33	68.6	142.980	147.496
15	69.4	141.405	144.920	34	67.0	142.900	149.357
16	68.6	142.485	146.986	35	67.2	142.565	148.763
17	67.5	139.410	145.114	36	69.2	143.650	147.462
18	67.6	142.155	147.851	37	67.7	141.905	147.470
19	68.4	140.740	145.423				

表 2 中等果在 $\lambda=199.583\text{ nm}$ 处的原始光谱与修正光谱
Table 2 The original and corrected spectra of the middle isocarpous at $\lambda=199.583\text{ nm}$

序号	横径/mm	原始光谱值	修正光谱值	序号	横径/mm	原始光谱值	修正光谱值
1	71.0	142.900	142.683	20	71.2	144.560	144.371
2	71.8	143.900	143.806	21	73.5	142.900	143.070
3	72.4	147.060	147.060	22	74.0	147.720	147.976
4	74.0	146.730	146.984	23	71.2	145.400	145.210
5	71.4	144.560	144.403	24	70.5	142.570	142.276
6	71.0	142.900	142.683	25	71.0	143.900	143.681
7	74.6	141.240	141.577	26	70.0	144.570	144.193
8	72.7	143.070	143.116	27	71.2	144.570	144.381
9	70.2	143.730	143.387	28	70.2	147.390	147.038
10	70.8	144.570	144.319	29	73.5	142.400	142.570
11	73.6	144.570	144.758	30	73.8	147.560	147.784
12	74.5	142.740	143.065	31	72.0	146.390	146.326
13	72.8	145.400	145.463	32	73.3	143.400	143.540
14	73.6	141.900	142.084	33	72.0	142.570	142.508
15	70.0	144.815	144.438	34	73.6	144.900	145.088
16	73.5	145.560	145.733	35	70.6	143.400	143.120
17	72.5	142.570	142.585	36	73.7	143.900	144.103
18	72.5	143.900	143.915	37	70.6	145.730	145.445
19	74.2	144.900	145.183	38	72.6	143.900	143.931

续表 2

序号	横径/mm	原始光谱值	修正光谱值	序号	横径/mm	原始光谱值	修正光谱值
39	74.2	143.230	143.510	50	72.6	145.060	145.091
40	72.2	144.070	144.038	51	72.5	147.720	147.736
41	71.4	143.570	143.414	52	72.8	147.230	147.293
42	71.2	145.230	145.041	53	70.5	145.900	145.599
43	70.4	143.400	143.089	54	70.5	145.400	145.100
44	71.3	147.390	147.214	55	70.2	146.230	145.881
45	73.5	144.900	145.073	56	73.8	142.070	142.285
46	72.8	146.390	146.453	57	74.3	148.220	148.525
47	72.2	145.900	145.868	58	70.2	147.060	146.709
48	71.0	148.390	148.164	59	74.0	146.060	146.313
49	74.8	143.730	144.104	60	73.3	145.400	145.542

表 3 大等果在 $\lambda=199.583\text{ nm}$ 处的原始光谱与修正光谱

Table 3 The original and corrected spectra of the big isocarpous at $\lambda=199.583\text{ nm}$

序号	横径/mm	原始光谱值	修正光谱值	序号	横径/mm	原始光谱值	修正光谱值
1	79.5	146.395	147.527	19	77.2	145.400	146.159
2	77.2	144.480	145.234	20	76.0	146.560	147.133
3	76.6	148.060	148.736	21	76.4	145.560	146.193
4	76.8	145.730	146.427	22	77.8	147.390	148.256
5	77.5	146.890	147.705	23	76.0	146.730	147.304
6	77.8	146.890	147.753	24	76.2	145.230	145.830
7	77.5	143.070	143.863	25	75.6	144.230	144.731
8	77.4	145.560	146.351	26	79.2	144.730	145.801
9	79.5	145.230	146.353	27	76.0	147.230	147.806
10	77.2	144.560	145.314	28	79.4	147.390	148.513
11	77.0	147.730	148.469	29	75.2	146.230	146.675
12	75.0	146.890	147.305	30	75.0	145.890	146.307
13	77.8	146.230	147.089	31	75.2	144.230	144.668
14	76.0	144.070	144.633	32	76.2	147.560	148.169
15	76.4	146.730	147.368	33	78.5	145.890	146.858
16	76.8	147.560	148.266	34	76.2	145.560	146.161
17	76.2	144.070	144.665	35	78.5	143.900	144.855
18	76.8	147.890	148.597				

2.3.2 SPXY 划分 图 8 为小等果原始光谱经过 SPXY 按 3 : 1 划分为校正集与预测集。校正集样本数 28, 预测集样本数 9。校正集糖度为 11.2~17.8 °Brix; 预测集糖度为 12.3~17.5 °Brix。

2.3.3 特征波长的选取 基于 MSC 与 SPXY 划分后的小等果原始光谱, 采用 CARS 算法选择出沃柑光谱的特征波长。由图 9 可知, CARS 算法随着运行次数的增加, 保留的波长变量越来越少, 在前 34 次运行中, 交叉验证均方根误差随着运行次数的增加而减少, 此时特征波长点为 13 个。由运行次数与各变量回归系数的变化关系

可知, 在运行次数为 34 时, 交互验证均方根误差出现最低点, 故保留波长为 34 次, 运行结果的 13 个波长点。

2.3.4 偏最小二乘模型的建立 由表 5、表 6 可知, 修正后各个等级的沃柑的校正集相关系数与预测集相关系数对比修正前的系数有所提高, 校正集均方根误差与预测集均方根误差均低于修正前的。小等果中, 修正前 R_p^2 为 0.790, RMSEP 为 0.489; 修正后 R_p^2 为 0.821, RMSEP 为 0.443。中等果修正前 R_p^2 为 0.801, RMSEP 为 0.460; 修正后 R_p^2 为 0.845, RMSEP 为 0.422。大等果修正前 R_p^2 为 0.820, RMSEP 为 0.431; 修正后 R_p^2 为 0.863, RMSEP 为 0.393。

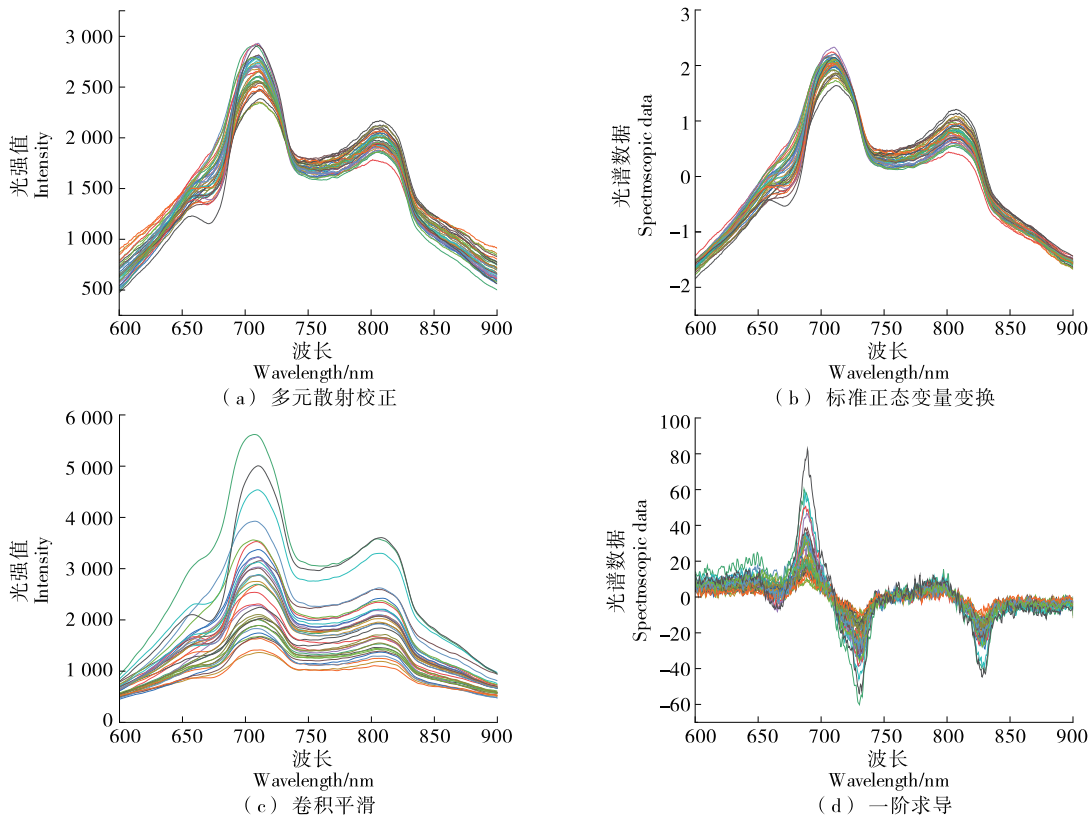


图 7 小等果原始光谱的各种预处理方法

Figure 7 Various pretreatment methods for the original spectrum of small isocarpous

表 4 预处理方法的比较

Table 4 Comparison of preprocessing methods

预处理方法	主成分数	校正集		预测集	
		R_c^2	RMSEC	R_p^2	RMSEP
MSC	13	0.867	0.423	0.790	0.489
SNV	22	0.873	0.462	0.772	0.503
SG	17	0.844	0.477	0.725	0.495
SG-1st	33	0.858	0.484	0.765	0.517

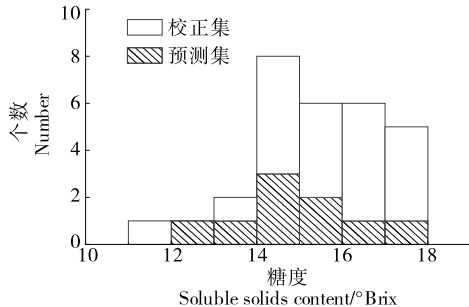


图 8 校正集与预测集糖度分布

Figure 8 Correction set and predicted glycan distribution

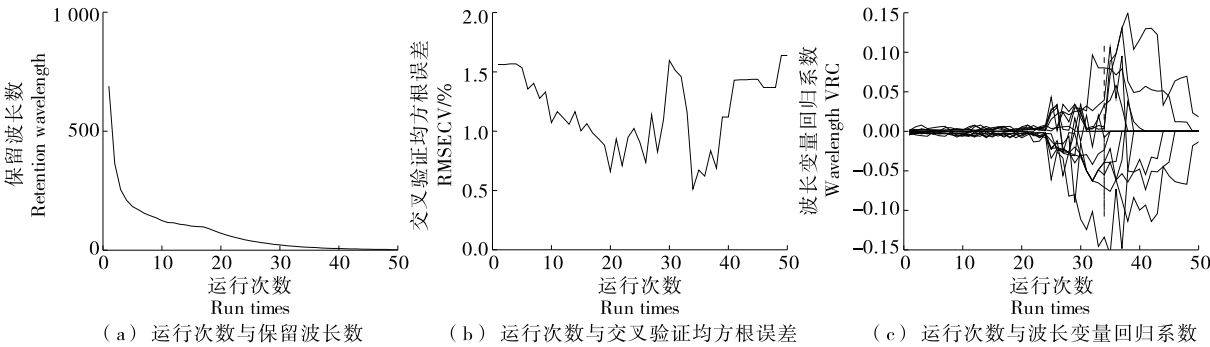


图 9 CARS 算法特征波长筛选流程

Figure 9 Characteristic wavelength selection process of CARS algorithm

表 5 原始光谱的建模数据

Table 5 Modeling data of the original spectrum

沃柑横 径等级	特征波 长数	校正集		预测集	
		R_c^2	RMSEC	R_p^2	RMSEP
小等果	13	0.867	0.423	0.790	0.489
中等果	28	0.887	0.416	0.801	0.460
大等果	11	0.892	0.395	0.820	0.431

将 Matlab 数据导入 Unscrambler 软件, 分别建立 3 个沃柑等级下的 PLS 预测模型。图 10~图 12 分别为小等果、中等果和大等果的预测值糖度与实测值糖度的

表 6 修正光谱的建模数据

Table 6 Modeling data of modified spectrum

沃柑横 径等级	特征波 长数	校正集		预测集	
		R_c^2	RMSEC	R_p^2	RMSEP
小等果	44	0.917	0.406	0.821	0.443
中等果	56	0.926	0.398	0.845	0.422
大等果	25	0.935	0.352	0.863	0.393

散点图。

2.4 未知样本的糖度预测

购入糖度未知, 软硬程度相当, 颜色相近的沃柑

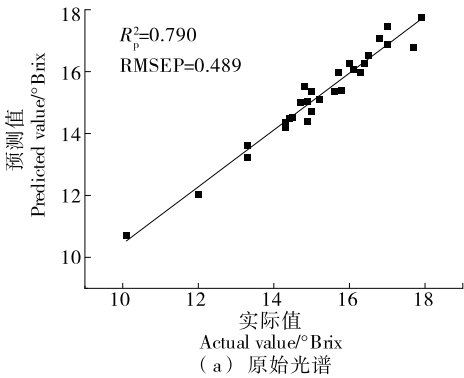


图 10 小等果的预测糖度与测量糖度的散点图

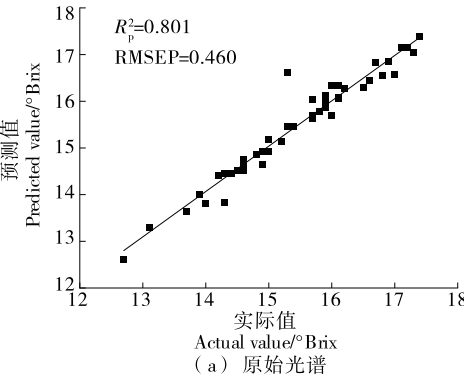
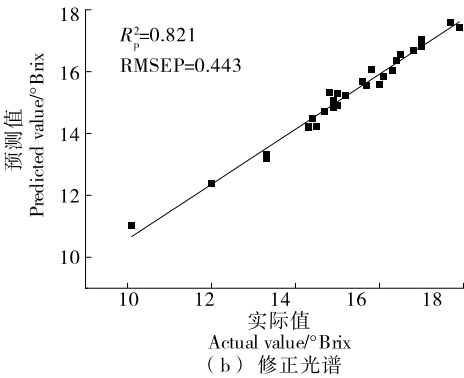


图 11 中等果的预测糖度与测量糖度的散点图

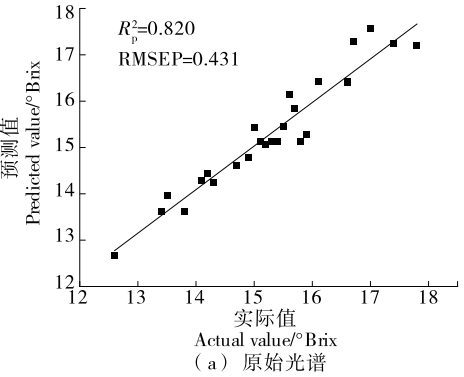
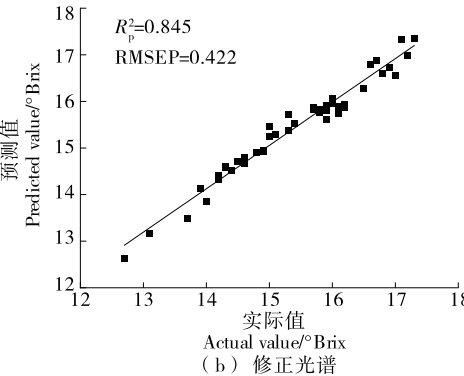


图 12 大等果的预测糖度与测量糖度的散点图

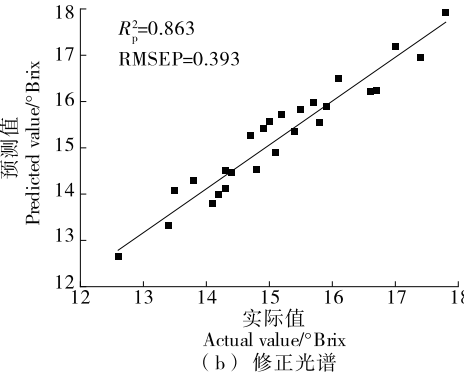


Figure 12 Scatter plot of predicted sugar content and measured sugar content of big isocarpous

15 个,对样本进行标号,用游标卡尺测量每一个沃柑的横径,并采集每一个沃柑的光谱。将光谱数据代入对应的偏最小二乘模型,分别求出修正前与修正后的糖度值。将沃柑糖度等级分为 4 级时,可溶性固形物分级区间为 $<12, 12\sim14, 14\sim16, >16$ °Brix。表 7 为未知样本的实际糖度值与修正前后预测糖度值的结果。从整体上看,经过光谱变换修正后的模型对未知样本糖度预测的误差比原始光谱的要小。原始光谱建模的 RMSE 为 0.570,修正光谱建模的 RMSE 为 0.404。

表 7 未知样本分析结果[†]

Table 7 Analysis results of unknown samples

序号	横径	糖度实测值	SSC(原)	误差(原)	SSC(修)	误差(修)
1	69.6	15.9	16.42	0.52	16.26	0.36
2	67.9	15.3	16.02	0.72	15.46	0.16
3	69.9	17.3	17.85	0.55	17.52	0.22
4	68.7	13.7	14.28	0.58	14.02	0.32
5	70.6	16.2	16.72	0.52	16.64	0.44
6	70.5	14.6	13.98	-0.62	14.14	-0.46
7	72.2	15.6	16.16	0.56	16.03	0.43
8	71.6	13.9	14.92	1.02	14.87	0.97
9	72.8	14.3	14.07	-0.23	13.92	-0.38
10	70.7	14.3	13.73	-0.57	13.95	-0.35
11	70.4	13.1	13.65	0.55	13.37	0.27
12	78.6	16.1	16.21	0.11	16.15	0.05
13	76.9	15.1	14.83	-0.27	14.92	-0.18
14	76.2	15.9	15.28	-0.62	15.55	-0.35
15	77.1	15.7	16.24	0.54	16.05	0.35

[†] SSC(原)表示用原始光谱建立的模型去预测未知样本的糖度得出的预测值;误差(原)表示 SSC(原)与 SSC(实测值)之间的误差;SSC(修)表示用修正光谱建立的模型去预测未知样本的糖度得出的预测值;误差(修)表示 SSC(修)与 SSC(实测值)之间的误差。

3 结论

研究将 132 个沃柑按横径分为 3 组,分别采集了全体沃柑的横径光谱,通过计算平均消光系数与运用光谱变换算法,综合光谱信息与沃柑横径信息,将不同大小的沃柑光谱转换到同一横径基准下。通过多元散射校正预处理,目标共生距离算法按 3 : 1 划分校正集与预测集,竞争性自适应权重取样法筛选特征波长点以及偏最小二乘回归法建立了原始光谱与修正光谱的糖度偏最小二乘模型。结果表明,小、中、大果修正后的预测集相关系数高于修正前的,预测集均方根误差均低于修正前的,说明光谱修正算法减小了沃柑横径带来的光谱差异,提升了模型的预测精度。

参考文献

- [1] 余怀鑫. 基于可见—近红外光谱技术的柑橘品质无损检测研究[D]. 武汉: 华中农业大学, 2021: 9.
YU H X. Research on non-destructive testing of citrus quality based on visible-near-infrared spectroscopy [D]. Wuhan: Huazhong Agricultural University, 2021: 9.
- [2] MILLER W M, ZUDE M. NIR-based sensing to identify soluble solids content of Florida citrus [J]. Applied Engineering in Agriculture, 2004, 20: 321-327.
- [3] HERREMANS E, VERBOVEN P, DEFRAEYE T, et al. X-ray CT for quantitative food microstructure engineering: The apple case[J]. Nuclear Instruments and Methods in Physics Research Section B-Beam Interactions with Materials and Atoms, 2014, 324: 88-94.
- [4] MC CLARK V A, FRASER D G, JORDAN R B, et al. Internal quality assessment of mandarin fruit by vis/NIR spectroscopy[J]. Journal of Near Infrared Spectroscopy, 2003, 11: 323-332.
- [5] 杜冉. 基于近红外透射光谱技术的苹果内部品质在线检测的研究[D]. 淄博: 山东理工大学, 2008: 10.
DU R. Research on on-line detection of internal quality of apple based on near-infrared transmission spectroscopy [D]. Zibo: Shandong University of Technology, 2008: 10.
- [6] 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: 12.
CAO S T. Study on the influence of sample correlation factors on visible/near-infrared spectroscopy detection of pear sugar content [D]. Hangzhou: Zhejiang University, 2017: 12.
- [7] 田世杰. 基于透射光谱的苹果霉心病判别影响因素及其修正方法研究[D]. 咸阳: 西北农林科技大学, 2020: 42.
TIAN S J. Study on the identification factors and correction method of apple mold heart disease based on transmission spectrum [D]. Xianyang: Northwest A & F University, 2020: 42.
- [8] 雷雨, 何东健, 周兆永, 等. 苹果霉心病可见/近红外透射能量光谱识别方法[J]. 农业机械学报, 2016, 47(4): 193-200.
LEI Y, HE D J, ZHOU Z Y, et al. Identification method of visible/near-infrared transmission energy spectroscopy for apple mold heart disease[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(4): 193-200.
- [9] 马敏娟, 李磊, 赵娟, 等. 融合不同成熟度的苹果可溶性固形物预测模型研究[J]. 食品安全质量检测学报, 2018, 9(11): 290-296.
MA M J, LI L, ZHAO J, et al. Study on soluble solids prediction model of apple with different maturity levels[J]. Journal of Food Safety and Quality Inspection, 2018, 9(11): 290-296.
- [10] 吴辰星. 融合多源信息的苹果霉心病在线检测方法研究[D]. 咸阳: 西北农林科技大学, 2017: 15.
WU C X. Research on online detection method of apple mold heart disease with multi-source information[D]. Xianyang: Northwest A & F University, 2017: 15.

(下转第 218 页)

- heat-treatment on the physico-chemical properties of egg white proteins: A kinetic study[J]. *Journal of Food Engineering*, 2006, 75(3): 316-326.
- [77] MA Y Q, SHAN A S, WANG R H, et al. Characterization of egg white powder gel structure and its relationship with gel properties influenced by pretreatment with dry heat[J]. *Food Hydrocolloids*, 2021, 110: 106149.
- [78] CHENG Y, WANG J T, CHI Y, et al. Effect of dry heating on egg white powder influencing water mobility and intermolecular interactions of its gels[J]. *Journal of the Science of Food and Agriculture*, 2021, 101(2): 433-440.
- [79] 熊舟翼, 马美湖, 蔡朝霞, 等. 卵白蛋白修饰改性对功能活性影响的研究进展[J]. *湖北农业科学*, 2013, 52(15): 3 473-3 477.
- XIONG Z Y, MA M H, CAI C X, et al. Research advance on the influence of modification on functional activity of ovalbumin[J]. *Hubei Agricultural Sciences*, 2013, 52(15): 3 473-3 477.
- [80] SUN J, MU Y Y, JING H, et al. Effects of single- and dual-frequency ultrasound on the functionality of egg white protein[J]. *Journal of Food Engineering*, 2020, 277: 109902.
- [81] XIONG W F, WANG Y T, ZHANG C L, et al. High intensity ultrasound modified ovalbumin: Structure, interface and gelation properties[J]. *Ultrasonics Sonochemistry*, 2016, 31: 302-309.
- [82] WANG Y C, ZHANG M, ADHIKARI B, et al. The application of ultrasound pretreatment and pulse-spouted bed microwave freeze drying to produce desalted duck egg white powders[J]. *Drying Technology*, 2013, 31(15): 1 826-1 836.
- [83] 刘松, 张东旭, 堵国成, 等. 微生物谷氨酰胺转氨酶的表达及分子改造研究进展[J]. *生物工程学报*, 2011, 27(12): 1 681-1 689.
- LIU S, ZHANG D X, DU G C, et al. Progress in expression and molecular modification of microbial transglutaminase[J]. *Chinese Journal of Biotechnology*, 2011, 27(12): 1 681-1 689.
- [84] HUYST A M R, DELEU L J, LUYCKX T, et al. Impact of heat and enzymatic treatment on ovalbumin amyloid-like fibril formation and enzyme-induced gelation [J]. *Food Hydrocolloids*, 2022, 131: 107784.
- [85] LU X, HUANG X, MA B, et al. Modification methods and applications of egg protein gel properties: A review [J]. *Comprehensive Reviews in Food Science and Food Safety*, 2022, 21(3): 2 233-2 252.

(上接第 87 页)

- [16] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[J/OL]. *arXiv*. (2017-12-05) [2024-03-19]. <https://doi.org/10.48550/arXiv.1706.05587>.
- [17] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 40(4): 834-848.
- [18] 马冬梅, 李鹏辉, 黄欣悦, 等. 改进 DeepLabV3+ 的高效语义分割[J]. *计算机工程与科学*, 2022, 44(4): 737-745.
- MA D M, LI P H, HUANG X Y, et al. Improved efficient semantic segmentation of DeepLab V3+ [J]. *Computer Engineering and Science*, 2022, 44(4): 737-745.
- [19] 王淑青, 黄剑锋, 张鹏飞, 等. 基于 YOLOv4 神经网络的小龙虾质量检测方法[J]. *食品与机械*, 2021, 37(3): 120-124, 194.
- WANG S Q, HUANG J F, ZHANG P F, et al. Crayfish quality detection method based on YOLOv4[J]. *Food & Machinery*, 2021, 37(3): 120-124, 194.
- [20] 郑帅. 语义分割技术在船舶卫星图像识别中的应用[J]. *舰船科学技术*, 2022, 44(14): 155-158.
- ZHENG S. Application of semantic segmentation technology in ship satellite image recognition[J]. *Ship Science and Technology*, 2022, 44(14): 155-158.
- [21] 赵世达, 王树才, 李振强, 等. 基于 U 型卷积神经网络的羊肋排图像分割[J]. *食品与机械*, 2020, 36(9): 116-121, 154.
- ZHAO S D, WANG S C, LI Z Q, et al. Image segmentation of sheep ribs based on U-shaped convolutional neural network[J]. *Food & Machinery*, 2020, 36(9): 116-121, 154.
- [22] 马冬梅, 黄欣悦, 李煜. 基于特征融合和注意力机制的图像语义分割[J]. *计算机工程与科学*, 2023, 45(3): 495-503.
- MA D M, HUANG X Y, LI Y. Image semantic segmentation based on feature fusion and attention mechanism [J]. *Computer Engineering and Science*, 2023, 45(3): 495-503.

(上接第 121 页)

- [11] 河野澄夫. 近赤外分光法による果実糖度の測定[J]. *食糧その科学と技術*, 2005, 43(3): 69-86.
- HEYE C F. Determination of fruit sugar content by near-infrared spectroscopy[J]. *Food and Science and Technology*, 2005, 43(3): 69-86.
- [12] 常汉. 水心病苹果水心程度与可溶性固形物含量在线无损检测方法与分级装备研究[D]. 杭州: 浙江大学, 2021: 22.
- CHANG H. Research on online nondestructive testing method and grading equipment for water-centered degree and soluble solids content of apple with water heart disease[D]. Hangzhou: Zhejiang University, 2021: 22.
- [13] WORKMAN J R J, WEYER L. Practical guide to interpretive near-infrared spectroscopy[M]. Boca Raton: CRC Press, 2007: 27.