

米良1号猕猴桃贮藏过程中糖酸含量估测模型的构建

刘丽^{1,2}, 杨天意^{1#}, 董聪颖¹, 石彩云¹, 司鹏¹, 魏志峰¹, 高登涛^{1,3*}

(¹中国农业科学院郑州果树研究所, 郑州 450009; ²中国农业科学院中原研究中心, 河南新乡 453500;

³中国农业科学院西部农业研究中心, 新疆昌吉 831100)

摘要:【目的】针对传统化学方法测定猕猴桃品质存在工序复杂、费时费力、需破坏性检测等问题, 提出一种基于高光谱技术的高效无损检测方法。【方法】以110个米良1号猕猴桃(*Actinidia chinensis* var. *deliciosa* ‘Miliang-1’)样本为研究对象, 利用高光谱仪采集不同贮藏时间果实的高光谱反射光谱。利用光谱-理化值共生距离法(sample set partitioning based on joint X-Y distance sampling, SPXY)将猕猴桃样本按照8:3的数量比例划分为训练集和测试集, 统一采用支持向量机(SVM)对比分析标准正态变换(SNV)、多元散射校正(MSC)、一阶导数(1st-D)、二阶导数(2nd-D)、平滑算法(SG)对原始光谱进行预处理。使用遗传算法(genetic algorithm, GA)和随机蛙跳(random frog, RF)对猕猴桃高光谱特征波长进行筛选, 结合支持向量回归(SVR)、反向传播神经网络(BP)算法, 组合构建猕猴桃品质的回归预测模型。【结果】在组合模型中, 可溶性固形物含量的最优模型为1st-D+GA-BP, R^2 为0.903, RMSE为1.731; 可滴定酸含量的最优模型为1st-D+GA-BP, R^2 为0.857, RMSE为0.225。【结论】应用高光谱技术对米良1号猕猴桃可溶性固形物含量、可滴定酸含量进行无损检测具有可行性。为进一步研究不同品种猕猴桃可溶性固形物含量、可滴定酸含量的无损检测模型奠定了基础。

关键词:猕猴桃; 高光谱; 硬度; 可溶性固形物含量; 可滴定酸含量

中图分类号: S663.4

文献标志码: A

文章编号: 1009-9980(2025)01-0216-11

Construction of a sugar and acid content estimation model for Miliang-1 kiwifruit during storage

LIU Li^{1,2}, YANG Tianyi^{1#}, DONG Congying¹, SHI Caiyun¹, SI Peng¹, WEI Zhifeng¹, GAO Dengtao^{1,3*}

(¹Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou 450009, Henan, China; ²Zhongyuan Research Center, Chinese Academy of Agricultural Sciences, Xinxiang 453500, Henan, China; ³Western Agricultural Research Center, Chinese Academy of Agricultural Sciences, Changji 831100, Xinjiang, China)

Abstract: 【Objective】 Traditional chemical methods for assessing the storage quality of kiwifruit typically involve complex procedures and high costs, which may hinder their widespread use. These conventional approaches often require significant labor, time and expensive reagents, making them less feasible for large-scale or routine quality control applications. Additionally, these methods usually result in the destruction of the fruit samples, which is not ideal for continuous monitoring. The complexity and cost associated with traditional chemical methods create barriers for smaller producers and can lead to inconsistencies in quality control across the industry. To address these issues, we propose a non-destructive testing method based on hyperspectral technology. This study aims to develop a reliable and efficient method for evaluating the quality of kiwifruit during storage without damaging the samples, thereby providing a more practical and economical solution for the kiwifruit industry. 【Methods】 In this study, 110 Miliang-1 kiwifruit samples were used as experimental subjects. These samples were selected to represent a broad range of storage conditions and potential quality variations, ensuring that the

收稿日期: 2024-09-02 接受日期: 2024-11-25

基金项目: 河南省重点研发专项(221111111800); 国家重点研发课题(2022YFD1600703); 中央级公益性科研院所基本科研业务费专项(1610192023105)

作者简介: 刘丽, 女, 副研究员, 主要从事果树栽培生理研究。E-mail: liuli03@caas.cn。#为共同第一作者。

*通信作者 Author for correspondence. E-mail: gaodengtao@caas.cn

findings of the study would be widely applicable. During the research, a hyperspectral imaging system was used to collect hyperspectral reflectance data of these kiwifruits at different storage times. This data collection included indicators such as titratable acidity, firmness and soluble solid content, which are critical factors in determining the fruit's overall quality. Hyperspectral imaging technology can capture detailed spectral information across a wide range of wavelengths, providing rich spectral data that offers insights into the internal and external properties of the fruit. This non-invasive method enables the assessment of quality attributes without compromising the integrity of the samples, allowing for repeated measurements over time. However, to ensure the accuracy and reliability of the data, multiple preprocessing methods were employed to process the collected data. These preprocessing methods not only enhance signal quality but also effectively remove noise from the data, ensuring the precision and effectiveness of subsequent analyses. The preprocessing methods used in the study included Standard Normal Variate transformation (SNV), Multiplicative Scatter Correction (MSC), first-order derivative (1st-D), second-order derivative (2nd-D) and Savitzky-Golay smoothing (SG). These methods help correct baseline variations, scatter effects and noise, improving the quality of the spectral data. Each preprocessing technique addresses specific issues within the spectral data, such as correcting for light scattering, baseline shifts and other interferences, thereby optimizing the data for further analysis. To select the optimal hyperspectral wavelengths for predicting kiwifruit quality, Genetic Algorithm (GA) and Random Frog (RF) methods were employed. These algorithms are powerful tools for feature selection, and capable of identifying the most informative wavelengths from the hyperspectral data. By pinpointing the most relevant wavelengths, these methods reduce the dimensionality of the data and enhance the efficiency of the predictive models. The selected wavelengths were then used to construct regression prediction models for kiwifruit quality indicators, including soluble solid content (SSC), firmness and titratable acidity. The regression models utilized a combination of Support Vector Regression (SVR) and Backpropagation Neural Network (BP) algorithms to determine the optimal predictive performance for each quality indicator. These models are particularly suitable for handling the complexity and non-linearity of hyperspectral data, as they can effectively learn from the intricate relationships within the data.

【Results】 The study found that the combination of preprocessing and wavelength selection significantly impacted the accuracy of the prediction models. For soluble solid content, the best model was 1st-D + GA-BP, with a coefficient of determination (R^2) of 0.903 and a root mean square error (RMSE) of 1.731, indicating high accuracy in predicting kiwifruit SSC, and reflecting the potential relationship between spectral data and SSC. For firmness, the best prediction model was 1st-D + RF-BP, with an R^2 of 0.900 and an RMSE of 0.879, demonstrating reliable predictive capability and highlighting the robustness of the model. For titratable acidity, the best model was 1st-D + GA-BP, with an R^2 of 0.857 and an RMSE of 0.225, showing good performance in predicting acidity levels and demonstrating the model's effective generalization to new data. These results underscore the effectiveness of the developed models and the significant role of preprocessing and feature selection in enhancing model performance.

【Conclusion】 The successful application of hyperspectral technology in this study highlights its potential for non-destructive quality assessment of kiwifruit. By accurately predicting key quality attributes such as SSC, firmness and acidity, hyperspectral imaging provides a powerful alternative to traditional chemical methods. This technology not only simplifies the assessment process but also reduces costs and preserves the integrity of the fruit samples. Additionally, it allows for continuous monitoring of fruit quality during storage, enabling timely interventions to maintain optimal conditions and prevent spoilage. This advancement could revolutionize quality control in the kiwifruit industry, providing a more effi-

cient, cost-effective and sustainable approach to maintaining high standards of fruit quality. The ability to monitor quality in a non-destructive manner also opens up new possibilities for research and development in the field of agricultural sciences, potentially leading to further innovations and improvements in fruit quality assessment and management.

Key words: Kiwifruit; Hyperspectral technology; Firmness; Soluble solid content; Titratable acidity

猕猴桃原产于中国,属于猕猴桃科猕猴桃属。因其果实独特的甜酸口感和丰富的维生素C含量而在全球范围内备受欢迎,是一种具有重要经济价值的水果^[1]。在猕猴桃贮藏过程中,通过检测硬度、可溶性固形物含量(SSC)和可滴定酸含量这几个关键指标,可以有效地管理猕猴桃的存储条件和时长,以维持其最佳食用品质,延长货架期,同时减少因过度成熟或腐烂导致的损失^[2]。目前,猕猴桃的品质检测过程主要依赖于耗时而复杂的传统化学方法,这些方法不仅操作繁琐,而且具有破坏性,往往会对果实本身造成损害。因此,建立一套高效且无损的猕猴桃果实品质检测方法,不仅可以节省大量人力物力,也能有效减少猕猴桃贮藏过程中的损耗。

高光谱技术在水果非破坏性检测方面具有明显优势,国内外科研人员已广泛应用高光谱技术进行水果内部品质的深入分析和研究。例如,卢娜等^[3]利用高光谱成像系统预测草莓硬度,通过标准正态变换(SNV)、多元散射校正(MSC)和Savitzky-Golay平滑等方法预处理光谱数据,采用偏最小二乘(PLS)方法结合化学计量学建模,比较各预处理方法对模型效果的影响。结果显示,经SNV处理的PLS模型表现最佳,相关系数高达0.989,均方根误差为0.021,也间接证明了高光谱成像技术检测果实硬度的可行性。林娇娇等^[4]通过近红外高光谱成像技术分析杧果的可溶性固形物含量,以探究不同品种之间在400~1000 nm波段的光谱差异。通过采集光谱数据并运用CARS-PLS模型进行分析,结果显示该模型具有高拟合度和预测精度,相关系数高达0.880 6,均方根误差为0.636 6,也间接证明了高光谱技术快速和无损检测果实可溶性固形物含量的可行性。赵明蕊等^[5]使用近红外高光谱成像技术来检测加工番茄的品质,并通过Savitzky-Golay方法优化原始光谱数据,建立了循环神经网络(RNN)、支持向量机(SVM)、K最近邻(KNN)、随机森林(RF)和偏最小二乘法(PLS)的模型,以预测番茄的可滴定酸含量,结果显示该模型具有高拟合度和预测精

度,相关系数高达0.869,均方根误差为0.03,也间接证明了高光谱技术在快速和无损检测果实可滴定酸含量上的可行性。这些研究强调了高光谱成像技术在水果无损检测中的潜力,以及通过不同方法和模型的应用,可以实现对水果内部质量的准确和高效检测。尽管针对各类水果内部品质的研究使用高光谱技术已相当成熟^[6-11],但关于运用高光谱技术和化学计量学来分析和研究猕猴桃在贮藏时期无损检测的定量模型相对较少。

笔者在本研究中以米良1号为对象,使用高光谱成像技术获取其光谱数据,使用化学计量分析获取其硬度、可溶性固形物含量、可滴定酸含量数据,并对光谱数据特征波段进行筛选,提取其品质指标的相关特征波段,并进一步建立米良1号贮藏过程中内部品质的预测模型,以期为米良1号及其他猕猴桃品质无损检测提供参考。

1 材料和方法

1.1 试验材料

试验所用样本果实采摘于中国农业科学院郑州果树研究所猕猴桃试验园,品种为米良1号,共计采摘了110个大小均匀、表面无损伤和疤痕的样本,编号后放入冷库中储存,储藏温度为4℃。每次采集图像前,需提前12 h将猕猴桃从冷库中取出,使其与周围温度保持一致。在试验的第0、10、20、30、40 d,每次分别测量22个样本的高光谱数据、可溶性固形物含量、硬度和可滴定酸含量。

1.2 仪器和设备

FigSpec® FS1X系列-高光谱相机(线扫描),彩谱科技(浙江)有限公司生产;GY-4-J型水果硬度计,浙江托普云农科技股份有限公司生产;PAL-BX/AC-ID 8便携式数显糖酸一体机(猕猴桃),日本ATAGO爱拓公司生产。

1.3 数据获取

高光谱数据采集:高光谱成像系统由FigSpec® FS1X系列-高光谱相机(线扫描)、位移控制平台、2

个150 W的光纤卤素灯、1台戴尔数据处理机组成。硬件包括高光谱相机(线扫描)、平移台、光源、计算机和软件处理控制系统,该成像系统通过PC机进行控制,并利用HSI Analyzer高光谱图像采集软件来采集信息。成像光谱仪光谱范围为400~1000 nm,光谱分辨率为2.5 nm。为了减少误差,在采集过程中确保环境温度和光源强度的稳定性,在正常室温20~26 °C和光线阴暗条件下,采用卤素灯提供光源进行高光谱图像的拍摄。此外,将标准白板的高度调整到与猕猴桃样本在同一焦面上,设置光谱相机的曝光时间为13.5 ms,样本平台与镜头的距离为170 mm。电控移动平台的前进距离为11 cm,前进速度为0.46 cm·s⁻¹,回退速度为5 cm·s⁻¹。

硬度测定:在进行测量前,将测试状态设为峰值保持状态,并将每个样品固定在工作台上,确保果实表面与仪器保持垂直。然后,在果实的赤道部位选择3个测量点(相隔120°)。以匀速转动GY-4型硬度计升降手柄的方式,使探头下压,并读取数据。当探头达到刻线(压入10 mm)后停止施加力,读出硬度值,单位为N。并求取3个测量点的平均值作为该样品的硬度参考值。

可溶性固形物含量(SSC)、可滴定酸含量测定:在完成硬度值测定后,立即切取所需测量部位的猕猴桃。使用手动榨汁机进行榨汁,在榨汁后立刻倒入一次性杯中。摇晃混匀,然后倒入测定窗口,马上点击Start,测定可溶性固形物含量和可滴定酸含量。每一个猕猴桃取3次测量结果平均值作为最终值,每次测量完毕后,需要用蒸馏水清洗测定窗口,并擦拭干净,以免影响试验结果。

1.4 光谱数据预处理与特征波段筛选

采集的猕猴桃样本高光谱数据导入ViewS-pecPro软件中,计算各组样本的平均光谱作为该组样本的高光谱数据并导出文件,然后在MatLab软件中进行预处理。由于外界及光谱仪自身扰动的影响,获得的猕猴桃光谱在两端波段噪声较大,信息冗余,为减少计算量,故选取400~1000 nm波段作为建模的全波段。

猕猴桃样品通过联合X/Y的异常样本识别方法剔除异常样本数据后,用联合X-Y距离样本集算法划分验证集和校正集,然后对光谱进行标准正态变换(SNV)、多元散射校正(MSC)、卷积平滑滤波处理(SG)、一阶导数(1st-D)、二阶导数(2nd-D)等预

处理以消除噪声和杂散光对模型性能的影响。

为解决光谱模型建立过程中的问题,如波长数目不足、低效率及模型复杂性,提出采用特征波长的建模方法。此方法通过选择合适的算法,如随机蛙跳(RF)和遗传算法(GA),来提升模型的稳定性并减少误差。随机蛙跳算法通过随机搜索最优解,而遗传算法则模拟自然选择的过程,通过迭代选择最优个体;遗传算法则通过模拟自然选择过程,迭代地优化特征选择,有效减少模型复杂度。通过合适的算法选择,可以有效地进行特征波长的提取和建模,使得模型在实际应用中更为精确和高效。

1.5 定量建模方法

定量建模是将检测样本中得到的信息进行量化并建立数学模型的过程,该建模方法可分为线性方法(SVR)和非线性方法(BP神经网络模型)。

支持向量回归(Support Vector Regression, SVR)是一种用于预测连续值的机器学习方法,通过寻找一个最佳的线来使得大多数数据点的预测值与实际值的误差在一个容许范围内。SVR有几个显著的优点:首先,它在高维数据中表现良好,适合处理有很多特征的数据;其次,通过使用核函数,SVR可以处理复杂的非线性问题;再次,SVR对异常值有一定的鲁棒性对小误差不敏感。然而,SVR也有一些缺点。它的计算复杂度较高,尤其是在处理大规模数据时,训练过程可能会非常耗时。此外,SVR需要选择合适的参数(如核函数类型、c值和g值),这通常需要通过交叉验证等方法来进行调优。

BP神经网络(反向传播神经网络)是一种用于回归任务的人工神经网络模型,通过调整网络权重和偏置以最小化预测误差。它包含输入层、隐藏层和输出层,利用前向传播计算输出,通过误差反向传播调整权重。模型训练通过多次迭代优化误差函数,并使用早停法防止过拟合。最终,经过评估和验证的模型可应用于实际数据,实现对新数据的预测。

2 结果与分析

2.1 贮藏品质的分析

从实际测得的数据可以看出,随着贮藏时间的增加,米良1号猕猴桃硬度逐渐降低。从图1可以看出,在贮藏初期会出现硬度快速降低的现象^[12]。SSC和TA含量是评价猕猴桃风味和食用品质的关键指标,二者作为呼吸基质,也是合成ATP的主要来

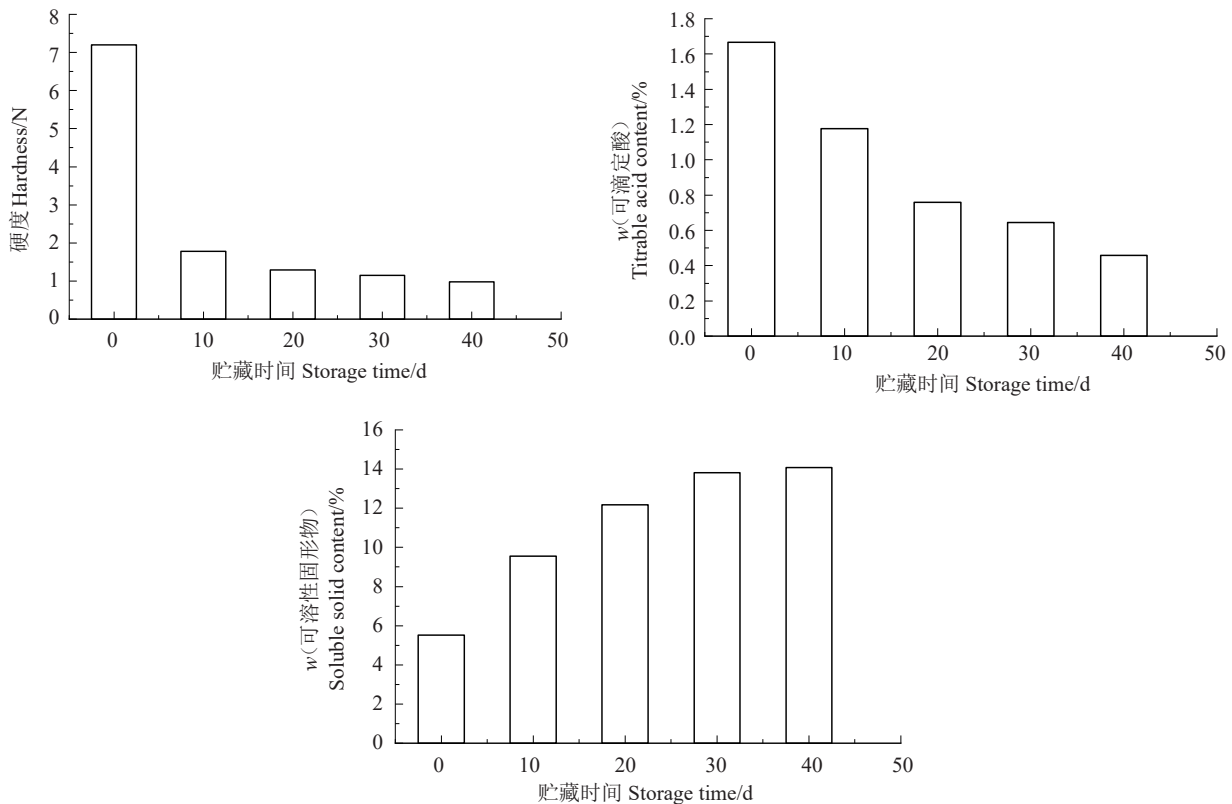


图1 米良1号贮藏过程中硬度、可滴定酸含量、可溶性固形物含量的变化

Fig. 1 Changes in hardness, titratable acid, and soluble solid content during storage of hybrid kiwifruit

源^[13]。从图1可以看出,随着猕猴桃的存放时间增加,伴随着呼吸作用,淀粉的转化使得SSC逐渐增加,TA含量逐渐下降。

2.2 样本划分

待测样本在进行光谱扫描和理化实验的过程中,由于仪器异常、操作错误和环境的影响,存在个别样本测量结果异常,笔者在本研究中采用基于XY变量联合的ODXY异常样本剔除算法进行异常样本剔除以提高模型精确度^[14]。通过SPXY算法计算样本之间的欧氏距离和标签差异,构建距离矩阵D和标签差异矩阵Dy。对距离矩阵进行归一化处理。选择距离最远的两个样本作为初始训练样本。

迭代选择距离已选训练样本最远的样本作为新的训练样本,直到达到指定的训练样本数量。将未被选为训练样本的样本作为测试样本。

对待测样本剔除异常值后,利用SPXY算法以8:3的比例划分训练集和预测集,具体的划分数据的统计结果见表1。由表1可知,训练集中包含了测试集中硬度、可滴定酸含量、可溶性固形物含量的最大值和最小值,并且其分布范围较大,也表明了SPXY划分的数据集是可靠的,这种合理的数据集划分能够确保模型的预测性能和准确性^[15]。

2.3 贮藏时期光谱分析

110个米良1号猕猴桃所获得的最初原始光谱

表1 训练集和验证集样本结果统计

Table 1 Statistics of sample results for training and validation sets

指标 Index	样本集 Sample set	数量 Number	范围 Range	平均值 Average value	标准差 Standard deviation
可溶性固形物含量 Soluble solid content	训练集 Training set	80	4.300~16.200	11.161	3.338
	测试集 Test set	29	5.100~15.733	10.863	3.960
硬度 Hardness	训练集 Training set	80	0.435~9.900	2.677	2.646
	测试集 Test set	29	0.665~7.715	2.057	2.077
可滴定酸含量 Titratable acid content	训练集 Training set	80	0.260~2.3467	0.954	0.478
	测试集 Test set	30	0.360~1.8133	0.867	0.473

数据在400~1000 nm范围的平均光谱反射率曲线如图2所示,结果表明,所有样品均表现出相似的光谱曲线趋势,110个米良1号猕猴桃所获取的最初原始光谱数据在400~1000 nm范围的平均光谱反射率曲线如图2所示,结果表明,所有样品均表现出相似的光谱曲线趋势。图3曲线是5个不同贮藏时间(0, 10, 20, 30和40 d)下,每个时期22个样本的平均光谱曲线,从曲线可以看出,400~640 nm之间反射率处于上升状态,这主要是由于猕猴桃果肉和表皮中的叶绿素和其他色素吸收引起的^[16],在640~660 nm范围内,趋于平稳但不同时期反射率而差值较大,是因为随着贮藏时间的增加,叶绿素的含量逐渐被分解,其反射率的值相对更高。670~750 nm呈急剧上升趋势,由于红边效应,可见光区转变到反射率较高

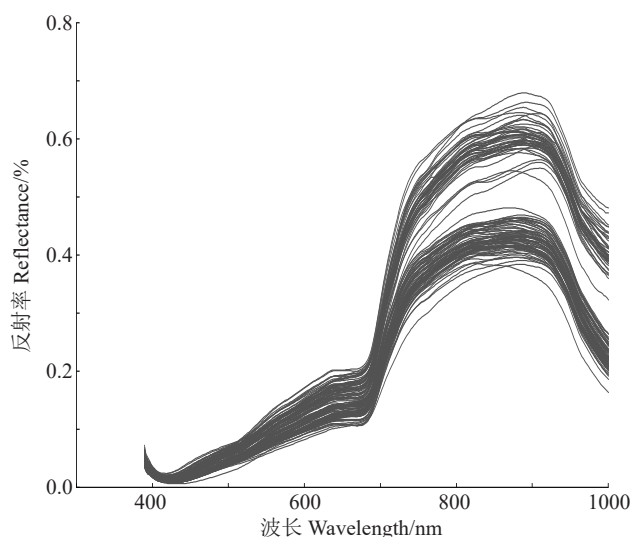


图2 原始光谱数据
Fig. 2 Original spectral data

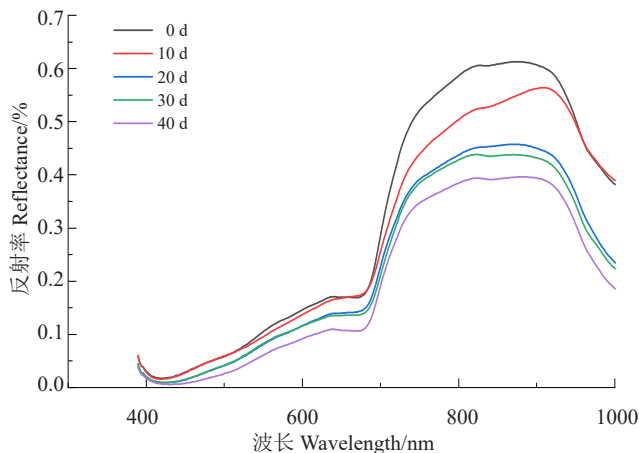


图3 不同贮藏时期的平均光谱曲线
Fig. 3 Average spectral curves at different storage periods

的近红外区的边缘^[17]。在830~900 nm呈缓慢趋势且不同时期的反射率差值较大。在810~830 nm和900~910 nm之间存在微弱波峰主要是果实中的碳水化合物和水分引起的,即体现了O-H的三级和二级倍频信息^[18]。

2.4 光谱预处理

利用SNV、MSC、SG、1st-D、2nd-D五种不同预处理方式得到米良一号猕猴桃光谱数据分别与硬度、可滴定酸含量、可溶性固形物含量进行SVR建模,得到的数据如表2。c和g是通过网格寻优的方式得到的模型最佳参数。模型评价采用RMSEP和R²指标,RMSE表示模型预测值与实际值之间差异的标准差,R²衡量模型解释能力。由表2可知,猕猴桃硬度经SNV处理得到的模型最好,R²为0.732, RMSE为1.360;可溶性固形物含量最优预处理是1st-D,R²和RMSE分别为0.805、1.523;猕猴桃可滴定酸含量最优预处理是1st-D,R²和RMSE分别为0.811、0.185。

表2 不同光谱预处理方法建模结果
Table 2 Modeling results of different spectral preprocessing methods

指标 Index	预处理方法 Preprocessing methods	R ²	RMSE	c	g
硬度 Hardness	SG	0.525	1.511	70	0.113
	1st-D	0.495	1.539	84	0.496
	2nd-D	0.694	1.362	56	0.018
	MSC	0.648	1.195	56	0.023
	SNV	0.732	1.360	9	0.064
可溶性固形物 含量 Soluble solid content	SG	0.775	1.647	61	0.034
	1st-D	0.805	1.523	88	0.184
	2nd-D	0.741	1.675	55	0.088
	MSC	0.781	1.821	52	0.004
	SNV	0.778	1.862	75	0.046
可滴定酸含量 Titrable acid content	SG	0.690	0.219	84	0.013
	1st-D	0.811	0.185	41	0.091
	2nd-D	0.665	0.298	33	0.078
	MSC	0.693	0.312	47	0.002
	SNV	0.702	0.290	40	0.029

2.5 提取特征波长

预处理后的光谱数据中含有大量的冗余信息,严重影响模型的鲁棒性和准确性,为了简化模型结果和提高预测精度,笔者使用随机蛙跳(RF)和遗传算法(GA)对猕猴桃高光谱特征波长进行筛选。由于算法具有随机性,笔者多次重复,选择选取最佳波

段。将随机蛙跳算法的参数设置迭代次数N为1000次,主成分个数10个,蛙跳初始种群数目Q为2个,以每个光谱被选择的可能性为筛选依据,运行结果为降序排列的被选择可能性,设定一阶导数预处理后数据、SNV处理后数据被选择可能性阈值分别为0.116、0.27,如图4所示随机蛙跳算法下波长选择的概率分布图,每组数据共筛选出10个特征波长,如图5所示红色边框方块为选定的变量。

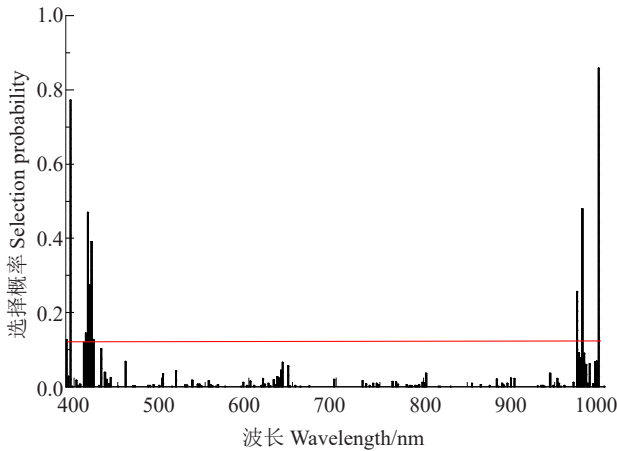


图4 RF算法波长选择概率分布

Fig. 4 Probability distribution of wavelength selection by RF algorithm

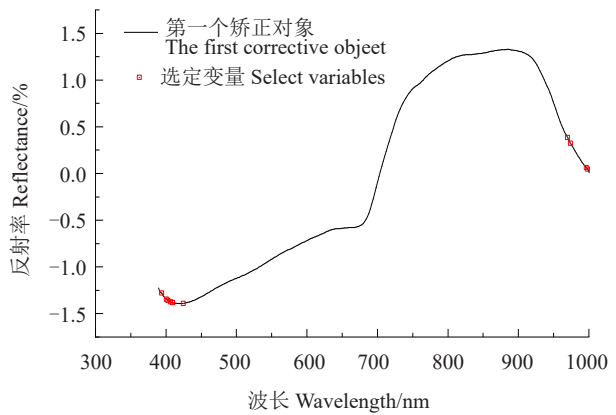
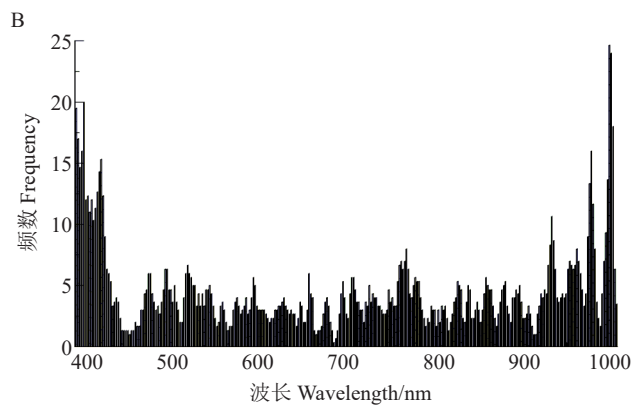
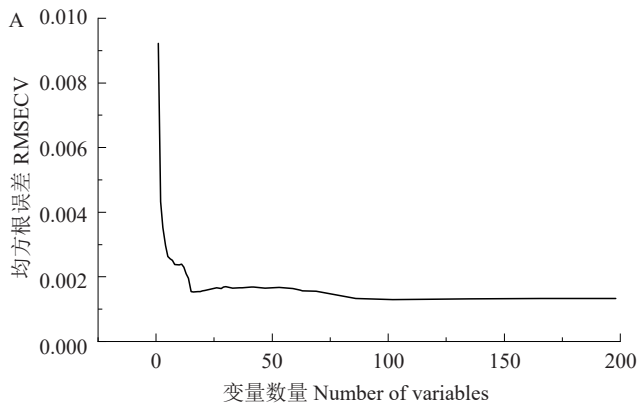


图5 RF算法选择的最优特征

Fig. 5 Optimal feature selected by wavelengths RF algorithm

遗传算法(Genetic Algorithm, GA)是一种基于自然选择和遗传机制的搜索和优化算法,用于解决复杂的优化问题。在GA运算过程中,设定初始群体为40,交叉率为0.5,变异率为0.01,迭代次数为200。以最小的RMSECV值为标准,RMSECV变化图如图6-A所示。筛选出波长点在迭代过程中出现频次较多的波长点为最优波长点,最终选定特征波长点为81个,如图6-B所示。



A. RMSECV 变化;B. GA 频率选择。

A. RMSECV variation; B. GA frequency selection.

图6 GA特征筛选的过程

Fig. 6 The process of GA feature selection

2.6 猕猴桃无损检测模型建立与分析

SVR支持向量机回归模型建立:以可溶性固形物含量、硬度和可滴定酸含量的校正集样本经RF和GA算法筛选出的特征波段作为输入,对应的含量值作为输出结合SVR算法建立回归预测模型。对可溶性固形物含量、硬度和可滴定酸含量的全谱作为输入,对应的含量值作为输出结合SVR算法建立

回归预测模型。采用网格寻优的方式寻找最佳参数c、g,建模结果如表2所示。

BP神经网络模型建立:设置BP神经网络模型参数,激活函数使用ReLU,输入层节点,学习率0.01,迭代次数1000次,以可溶性固形物含量、硬度和可滴定酸含量的校正集样本经RF和GA算法筛选出的特征波段作为输入,对应的含量值作为输出

表 2 SVR 建模结果

Table 2 SVR modeling results

指标 Index	预处理方法 Preprocessing methods	特征筛选 Feature selection	建模方法 Modeling method	R^2	RMSE	模型参数 Model parameter
可溶性固形物含量 Soluble solid content	None	GA	SVR	0.632	2.144	c=50;g=0.152
	1st-D	GA	SVR	0.834	1.405	c=52;g=0.998
硬度 Hardness	None	RF	SVR	0.776	1.850	c=94;g=0.604
		GA	SVR	0.647	1.430	c=24;g=0.027
		GA	SVR	0.831	0.916	c=73;g=0.878
可滴定酸含量 Titrable acid content	None	RF	SVR	0.871	0.849	c=96;g=0.985
		GA	SVR	0.623	0.302	c=58;g=0.008
		GA	SVR	0.751	0.254	c=45;g=0.921
		RF	SVR	0.748	0.275	c=78;g=0.763

结合BP算法建立回归预测模型。对可溶性固形物含量、硬度和可滴定酸含量的全谱作为输入,对应的含量值作为输出结合BP算法建立回归预测模型,建模结果如表3所示。

结合表2和表3可得:可溶性固形物含量的最优模型为1st-D+GA-BP, R^2 为0.903, RMSE为1.731;硬度的最优预测模型为1st-D+RF-BP, R^2 为

0.9, RMSE为0.879;可滴定酸含量的最优模型为1st-D+GA-BP, R^2 为0.857, RMSE为0.225。可溶性固形物含量、硬度和可滴定酸含量的高光谱回归预测模型预测值与真实值分析如图7、8、9所示,训练集和预测集的相关系数均在0.85以上,其均方根误差也最小,预测值也拟合在曲线周围,模型预测准确。

表 3 BP 建模结果

Table 3 Modeling results of different spectral preprocessing methods

指标 Index	预处理方法 Preprocessing methods	特征筛选 Feature selection	建模方法 Modeling method	R^2	RMSE	模型参数 Model parameter
可溶性固形物含量 Soluble solid content	None	GA	BP	0.656	1.736	5
	1st-D	GA	BP	0.903	1.163	5
		RF	BP	0.734	1.886	5
硬度 Hardness	None	GA	BP	0.664	1.261	5
	1st-D	GA	BP	0.889	0.894	5
		RF	BP	0.932	0.621	5
可滴定酸含量 Titrable acid content	None	GA	BP	0.622	0.359	5
	1st-D	GA	BP	0.857	0.225	5
		RF	BP	0.705	0.269	5

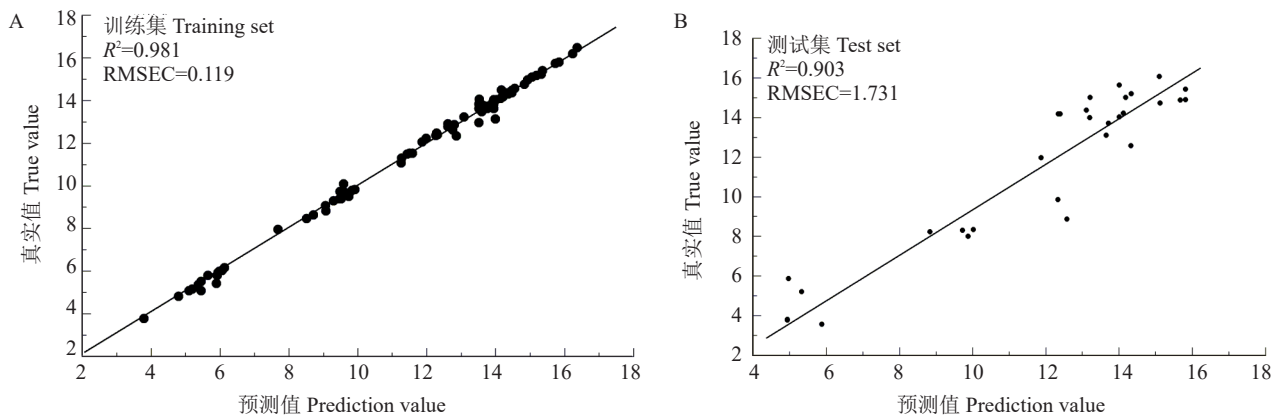


图 7 可溶性固形物含量预测效果

Fig. 7 Prediction effect of soluble solids content

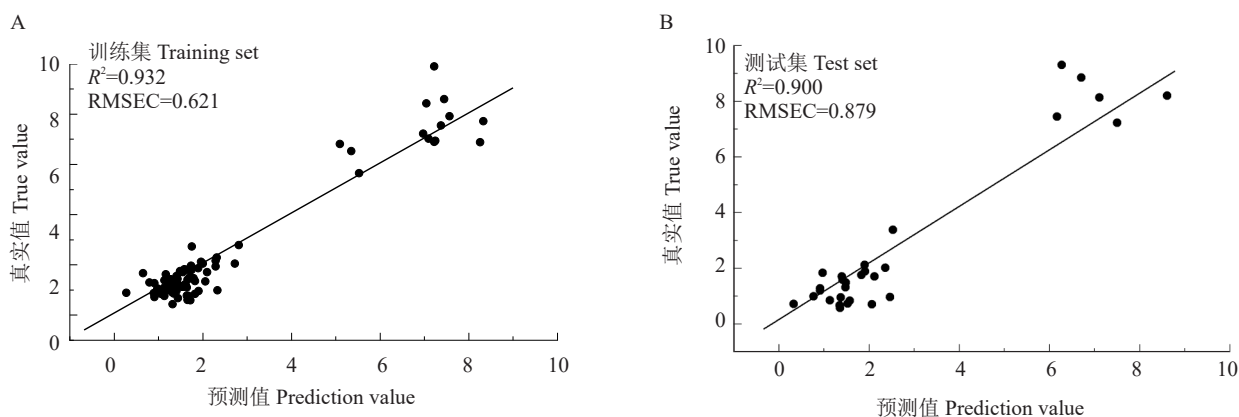


图8 硬度预测效果

Fig. 8 Hardness prediction effect

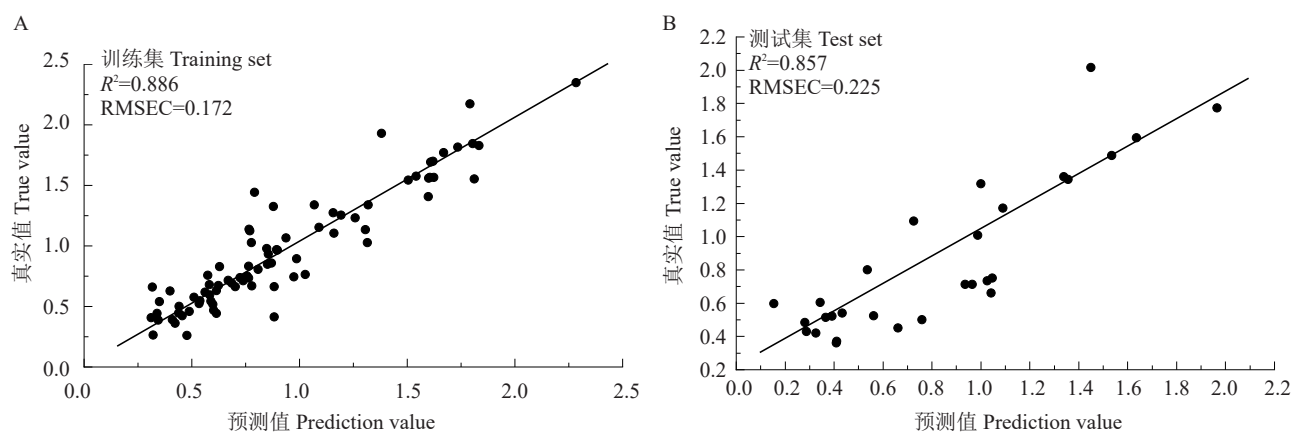


图9 可滴定酸含量预测效果

Fig. 9 Prediction effect of titratable acids content

3 讨 论

笔者在本研究中利用成像高光谱对米良1号猕猴桃可溶性固形物含量、硬度、可滴定酸含量进行无损检测,获取覆盖400~1000 nm光谱范围内300个波段的详细光谱数据。样品中可溶性固形物含量、可滴定酸含量的参考值是通过便携式数字折光仪(ATAGO,日本)经过3次单独测量获得的,以平均值作为可溶性固形物含量的参考值。硬度的参考值是通过GY-4型硬度计(浙江托普),果实的赤道部位选择3个测量点(相隔120°),以相同的速度进行按压,最后取三点的平均值作为硬度的参考值。使用ViewSpecPro软件,进行圈取猕猴桃轮廓并提取其平均光谱数据。随后采用MATLABR2023b软件进行分析,采用1st-D、2nd-D、MSC、SNV和SG等5种预处理方法来进行数据的处理与分析,采用随机蛙跳(RF)、遗传算法(GA)提取特征光谱信息,同时降

低数据的复杂度。

基于特征变量构建的1st-D+RF-BP模型,拥有较高的 $R^2=0.900$ 以及较小的 $RMSE=0.879$,说明1st-D+RF-BP模型可以很好地预测米良1号猕猴桃的硬度。相较于基于全变量构建的猕猴桃硬度无损检测模型,该模型从全部变量中筛选出较少的变量,很大程度上提升了模型的运行速率,运算时间从5.84 s降至0.84 s。可溶性固形物含量的最优模型为1st-D+GA-BP, R^2 为0.903, $RMSE$ 为1.731。相比孟庆龙等^[7]基于主成分回归的猕猴桃可溶性固形物含量无损检测和吴彦红等^[19]建立的猕猴桃多线性回归模型精度有所提升,但对比姜凤丽等^[8]研究的软枣猕猴桃SSC检测精度有所不足,可能是由于猕猴桃品种不同所致。可滴定酸含量的最优模型为1st-D+GA-BP,拥有较高的 $R^2=0.857$ 和较小的 $RMSE=0.225$,说明该模型对预测米良一号猕猴桃可滴定酸含量的可行性,但精度相较于孟庆龙等^[20]预测结果存在一定

的不足,可能是由于贮藏时间的问题。可滴定酸含量在贮藏前期变化程度较大,贮藏后期趋于平稳,调整数据采集区间可以提高模型精度。

笔者在本研究中通过提取特征波长并结合线性和非线性模型,成功预测了米良1号猕猴桃的可溶性固形物含量、硬度和可滴定酸含量,为该猕猴桃在贮藏过程中的内部生理变化提供了参考依据。未来研究重点将为选择和组合其他有效的特征波长提取方法,以解决波长变量自相关性的问题,并考虑不同品种和贮藏方式等因素的影响,从而达到提升米良1号猕猴桃预测模型准确率的目的。

4 结 论

以米良1号猕猴桃为研究对象,采集不同贮藏时间的猕猴桃高光谱信息和理化指标,通过分析1st-D、2nd-D、MSC、SNV和SG等5种预处理方法,两种特征筛选方法随机蛙跳(RF)、遗传算法(GA)再结合BP神经网络、SVR支持向量机两种算法,建立米良1号猕猴桃贮藏品质无损检测模型,最终得出如下结论:

1)通过比较5种光谱预处理方法,发现可溶性固形物含量、硬度、可滴定酸含量的最优预处理均为1st-D。

2)采用遗传算法、随机蛙跳对预处理光谱进行处理,能够有效对特征波长进行筛选,简化了模型复杂度,提高了预测精度。

3)可溶性固形物含量的最优模型为1st-D+GA-BP, R^2 为0.903, RMSE为1.731;硬度的最优预测模型为1st-D+RF-BP, R^2 为0.900, RMSE为0.879;可滴定酸含量的最优模型为1st-D+GA-BP, R^2 为0.857, RMSE为0.225。

参考文献 References:

- [1] 钟曼茜,翟舒嘉,刘伟,睢国祥,段玉权,林琮,陶鑫凉.我国即食猕猴桃产业发展现状、问题与对策[J].中国果树,2023(2):122-127.
ZHONG Manxi, ZHAI Shujia, LIU Wei, SUI Guoxiang, DUAN Yuquan, LIN Qiong, TAO Xinliang. Current situation, problems and countermeasures of the development of ready-to-eat kiwifruit industry in China[J]. China Fruits, 2023(2): 122-127.
- [2] 郭琳琳,庞荣丽,王瑞萍,乔成奎,田发军,王彩霞,李君,庞涛,成昕,谢汉忠.猕猴桃营养品质综合评价[J].果树学报,2022,39(10):1864-1872.
GUO Linlin, PANG Rongli, WANG Ruiping, QIAO Chengkui, TIAN Fajun, WANG Caixia, LI Jun, PANG Tao, CHENG Xin, XIE Hanzhong. Comprehensive trait evaluation for kiwifruit nutritional quality[J]. Journal of Fruit Science, 2022, 39(10): 1864-1872.
- [3] 卢娜,韩平,王纪华.基于高光谱成像技术的草莓硬度预测[J].软件导刊,2018,17(3):180-182.
LU Na, HAN Ping, WANG Jihua. Prediction on firmness of strawberry based on hyperspectral imaging[J]. Software Guide, 2018, 17(3): 180-182.
- [4] 林娇娇,蒙庆华,吴哲锋,常洪娟,倪淳宇,邱邹全,李华荣,黄玉清.基于近红外高光谱技术的芒果可溶性固形物含量无损检测[J].果树学报,2024,41(1):122-132.
LIN Jiaojiao, MENG Qinghua, WU Zhefeng, CHANG Hongjuan, NI Chunyu, QIU Zouquan, LI Huarong, HUANG Yuqing. Fruit soluble solids content non-destructive detection based on visible/near infrared hyperspectral imaging in mango[J]. Journal of Fruit Science, 2024, 41(1): 122-132.
- [5] 赵明蕊.基于高光谱成像技术的加工番茄果实品质无损检测研究[D].石河子:石河子大学,2023.
ZHAO Mingrui. Nondestructive testing of processing tomato fruit quality based on hyperspectral imaging[D]. Shihezi: Shihezi University, 2023.
- [6] 尚静,冯树南,谭涛,吴美芝,陈海江,孟庆龙.基于高光谱成像的贵长猕猴桃硬度快速预测[J].食品工业科技,2023,44(6):345-350.
SHANG Jing, FENG Shunan, TAN Tao, WU Meizhi, CHEN Haijiang, MENG Qinglong. Rapid prediction for the firmness of Guichang kiwifruit by hyperspectral imaging[J]. Science and Technology of Food Industry, 2023, 44(6): 345-350.
- [7] 孟庆龙,黄人帅,张艳,尚静.贮藏期内猕猴桃酸度的快速无损检测[J].农产品加工,2022(11):66-68.
MENG Qinglong, HUANG Renshuai, ZHANG Yan, SHANG Jing. Rapidly nondestructive detection for the pH of kiwifruits during storage[J]. Farm Products Processing, 2022(11): 66-68.
- [8] 姜凤利,杨磊,田有文,孙炳新,罗子旋.基于高光谱成像的软枣猕猴桃SSC检测研究[J].沈阳农业大学学报,2023,54(3):318-326.
JIANG Fengli, YANG Lei, TIAN Youwen, SUN Bingxin, LUO Zixuan. Detection of soluble solids content in *Actinidia arguta* based on hyperspectral imaging[J]. Journal of Shenyang Agricultural University, 2023, 54(3): 318-326.
- [9] GALVÃO R K H, ARAUJO M C U, JOSÉ G E, PONTES M J C, SILVA E C, SALDANHA T C B. A method for calibration and validation subset partitioning[J]. Talanta, 2005, 67(4): 736-740.
- [10] RICCIOLI C, PÉREZ-MARÍN D, GARRIDO-VARO A. Optimizing spatial data reduction in hyperspectral imaging for the prediction of quality parameters in intact oranges[J]. Postharvest Biology and Technology, 2021, 176: 111504.
- [11] AREDO V, VELÁSQUEZ L, CARRANZA-CABRERA J, SICHE R. Predicting of the quality attributes of orange fruit using

- hyperspectral images[J]. Journal of Food Quality and Hazards Control, 2019, 6: 82-92.
- [12] 陆玲鸿, 马媛媛, 古咸彬, 肖金平, 宋根华, 张慧琴. 猕猴桃果实贮藏期间细胞壁多糖物质降解特性及组织结构差异分析[J]. 核农学报, 2023, 37(3): 550-558.
LU Linghong, MA Yuanyuan, GU Xianbin, XIAO Jinping, SONG Genhua, ZHANG Huiqin. Degradation characteristics of cell wall polysaccharides and differences in tissue structure of kiwifruit fruits during storage[J]. Journal of Nuclear Agricultural Sciences, 2023, 37(3): 550-558.
- [13] 李艳杰, 孙先鹏, 郭康权, 王英. 臭氧、保鲜剂对猕猴桃贮藏保鲜效果的比较[J]. 食品科技, 2009, 34(2): 45-48.
LI Yanjie, SUN Xianpeng, GUO Kangquan, WANG Ying. Comparison of the storage effect on kiwi fruit between ozone and preservative[J]. Food Science and Technology, 2009, 34(2): 45-48.
- [14] 尹宝全, 史银雪, 孙瑞志. 近红外光谱分析中的一种基于XY变量联合的异常样本剔除算法[J]. 中国科学技术大学学报, 2016, 46(3): 208-214.
YIN Baoquan, SHI Yinxue, SUN Ruizhi. An outlier sample eliminating algorithm based on joint XY distances for near infrared spectroscopy analysis[J]. Journal of University of Science and Technology of China, 2016, 46(3): 208-214.
- [15] 张芳, 邓照龙, 田有文, 高鑫, 王开田, 徐正玉. 基于高光谱成像技术的南果梨酸度无损检测方法[J]. 沈阳农业大学学报, 2024, 55(2): 231-239.
ZHANG Fang, DENG Zhaolong, TIAN Youwen, GAO Xin, WANG Kaitian, XU Zhengyu. Non-destructive testing method for acidity of Nanguo pear based on hyperspectral imaging technology[J]. Journal of Shenyang Agricultural University, 2024, 55(2): 231-239.
- [16] 高宏盛, 郭志强, 曾云流, 丁港, 王逍遥, 李黎. 基于高光谱图像波段融合的猕猴桃软腐病早期分类检测[J]. 光谱学与光谱分析, 2024, 44(1): 241-249.
GAO Hongsheng, GUO Zhiqiang, ZENG Yunliu, DING Gang, WANG Xiaoyao, LI Li. Early classification and detection of kiwifruit soft rot based on hyperspectral image band fusion[J]. Spectroscopy and Spectral Analysis, 2024, 44(1): 241-249.
- [17] GITELSON A A, KEYDAN G P, MERZLYAK M N. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves[J]. Geophysical Research Letters, 2006, 33(11): L11402.
- [18] SPIER M R, WOICIECHOWSKI A L, DE SOUZA V L P, SOCOL C R. Production and characterization of amylases by *Aspergillus niger* under solid state fermentation using agro industrial products[J]. International Journal of Food Engineering, 2006, 2(3): 1556-3758.
- [19] 吴彦红, 严霖元, 吴瑞梅, 杨勇. 利用荧光高光谱图像技术无损检测猕猴桃糖度[J]. 江西农业大学学报, 2010, 32(6): 1297-1300.
WU Yanhong, YAN Linyuan, WU Ruimei, YANG Yong. Non-destructive determination of sugar content in kiwifruit using hyperspectral fluorescence imaging technology[J]. Acta Agriculturae Universitatis Jiangxiensis, 2010, 32(6): 1297-1300.
- [20] 孟庆龙, 黄人帅, 张艳, 尚静. 贮藏期内猕猴桃酸度的快速无损检测[J]. 农产品加工, 2022(11): 66-68.
MENG Qinglong, HUANG Renshuai, ZHANG Yan, SHANG Jing. Rapidly nondestructive detection for the pH of kiwifruits during storage[J]. Farm Products Processing, 2022(11): 66-68.

2024年度优秀审稿专家

为了感谢审稿专家对《果树学报》的支持与帮助,同时调动审稿专家的积极性,增强审稿专家的责任感和使命感,编辑部通过统计审稿数量、审稿质量和审稿时效三个指标,评选出43位优秀审稿专家,本刊对其严肃、严谨的学术作风,公平公正、认真负责的审稿态度及甘于奉献的精神表示衷心的感谢!具体名单如下(按姓氏笔画为序):

方伟超 王磊 王中堂 王少敏 王树桐 王家保 王彩霞 冉昆 冯立娟 石倩倩
 龙友华 刘辉 刘有春 吕中伟 吕德国 孙健 孙浩元 孙海悦 何业华 张蕾
 张朝红 张慧琴 李大卫 李明章 李林光 李桂荣 李燕婷 杨江山 汪松虎 陈双建
 徐小彪 钱国良 章秋平 彭福田 赖钟雄 雷新涛 谭彬 潘志勇 蒋迎春 薛晓敏
 陈宝成 赵迎丽 赵密珍

本刊编辑部