

# Determination of Reducing Sugar and Alcoholic Strength in Jujube Vinegar Fermented Liquid by Hyperspectral Image Technology

Huixia Jiang, LiuJun Jia, Haihong Zhang\*, Baoting Wu, Dongdong Li, Ziwen Li

Agricultural College, Ningxia University, Yinchuan Ningxia  
Email: \*1042132475@qq.com

Received: Apr. 6<sup>th</sup>, 2018; accepted: Apr. 20<sup>th</sup>, 2018; published: Apr. 27<sup>th</sup>, 2018

## Abstract

Reducing sugar and alcoholic strength content in jujube vinegar fermented processing was carried out by hyperspectral image technology, and quantitative analysis model was established by partial least squares method, while the spectral characteristic wavelength of the spectral region was screened by uninformative variables elimination (UVE), competitive adaptive reconvolutions (CARS) and genetic algorithm (GA). Determination coefficient  $R^2$ , root-mean-square error of prediction (RMSEP), relative percent deviation (RPD) and the best principle factors were employed as evaluation indexes. The optimization effect of CARS was the best,  $R^2$  of optimization mode were 0.9045 and 0.8993, RMSEP were 1.3635 and 1.2878, RPD were 3.24 and 3.58, and best principle factors were 7 and 8. The results showed that: variable selection can improve accuracy and stability of the model of sugar and alcoholic strength content quantitative analysis, reduce analytical difficulty, and optimize the model. It is feasible that reducing sugar and alcoholic content of jujube vinegar fermented liquid can be analyzed rapidly, quantitatively and accurately by hyperspectral image technology.

## Keywords

Jujube Vinegar, Hyperspectral Image Technology, Quantitative Analysis, Reducing Sugar, Alcoholic Strength

# 枣醋发酵液中还原糖含量和酒精度的可见 - 近红外高光谱快速定量分析

蒋慧霞, 贾柳君, 张海红\*, 吴宝婷, 李冬冬, 李子文

\*通讯作者。

宁夏大学农学院, 宁夏 银川  
Email: 1042132475@qq.com

收稿日期: 2018年4月6日; 录用日期: 2018年4月20日; 发布日期: 2018年4月27日

## 摘要

本文利用可见-近红外高光谱检测技术对枣醋发酵过程中还原糖含量和酒精度进行了定量分析, 并通过偏最小二乘法(PLS)建立定量分析模型, 同时采用无信息变量消除法(UVE)、竞争性自适应加权算法(CARS)和遗传算法(GA)对整个谱区进行光谱特征波长变量筛选, 以决定系数( $R^2$ )、预测标准偏差(RMSEP)、相对分析误差(RPD)以及最佳主因子数作为模型质量的评价指标。其中用CARS法挑选波长后对模型的优化效果最佳, 还原糖含量和酒精度CARS-PLS模型的 $R^2$ 分别达到0.9045和0.8993, RMSEP为1.3635和1.2878, RPD为3.24和3.58, 最佳主因子数为7和8。结果表明: 进行变量筛选可提高枣醋酿造过程中还原糖含量和酒精度模型准确度和稳定性, 降低解析难度, 达到优化模型的作用, 可见-近红外高光谱检测技术可实现枣醋发酵液中还原糖含量和酒精度的快速、定量、精确分析。

## 关键词

枣醋, 高光谱检测技术, 定量分析, 还原糖, 酒精度

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

红枣富含糖、酸、维生素、矿物质、蛋白质及多酚等生理活性物质, 营养价值极高。以其为原料, 经酒精发酵、醋酸发酵酿制而成的枣醋是一种集营养与保健于一体的新型保健产品, 具有降低血压、抗衰老、抗氧化、促进新陈代谢, 维持体内酸碱平衡等多种保健功能[1][2]。枣醋的品质与酿造过程中各项参数的变化密切相关, 其中还原糖含量、酒精度直接影响着发酵进程及转化率, 适时快速监测与调控这两项变量参数, 是确保枣醋酿造质量的关键环节。枣醋酿造过程中还原糖为酵母的生长提供营养物质, 然而其含量过高会抑制酵母的生长繁殖, 含量过低又会使得酒精发酵不完全进而影响酒精量的生成。酒精度的变化则会直接影响到醋酸量的形成, 酒精度过高会使酒精转酸率降低进而影响枣醋的品质[3]。目前, 这两项指标的常规检测方法前处理复杂, 监测信息滞后, 很难进行大量样品的快速检测, 因此, 迫切需要寻找一种快速、准确的检测方法, 以指导酿造生产, 提升枣醋品质。

高光谱检测技术是一种新的基于超多波段的影像数据技术, 在可见-近红外(400~1100 nm)范围内能够获得波段数多达上百个的光谱连续、分辨率高达纳米数量级的图像数据, 同时可以得到图像信息和光谱信息, 其光谱信息能够反映样本化学成分和分子信息等内部信息, 图像信息能够反映样本大小、形状、颜色和纹理等外部属性, 可对样品进行定性与定量分析。由于高光谱检测技术具有无需样品前处理, 分析速度快, 效率高, 操作简单等优点, 近年来已在食品、药品等行业的检测方面得到了广泛的关注[4][5][6][7][8]。邹小波等[9]利用高光谱检测技术和化学计量学方法实现翻醅均匀性快速判别, 以翻醅前后醋醅为研究对象, 镇江香醋固态发酵过程中总酸质量分数和 pH 值为表征指标, 采集其高光谱信息, 并检

测对应的总酸质量分数和 pH 值, 采用 SiPLS 和 GA 优选特征变量, PLS 和 LS-SVM 方法建立快速预测模型, 结果表明利用高光谱技术快速预测醋醅均匀性是可行的。申婷婷等[10]研究采用高光谱检测技术检测醋酸发酵中醋醅样品的总酸、pH 值、水分和不挥发酸含量的可视化分布图, 并提取了理化指标分布的定量特征, 实现了醋醅理化指标的可视化分布检测及其分布特征的数字化描述。

本文拟对枣醋液态发酵过程中还原糖含量和酒精度进行快速分析, 采用偏最小二乘法(PLS)建立定量分析模型, 同时采用无信息变量消除法(UVE)和竞争性自适应加权算法(CARS)和遗传算法(GA)对整个谱区进行光谱特征波长变量筛选。比较分析这三种光谱筛选方法对枣醋酿造过程中还原糖含量和酒精度预测模型的影响, 降低模型的复杂程度, 提高模型精度及运算速度。

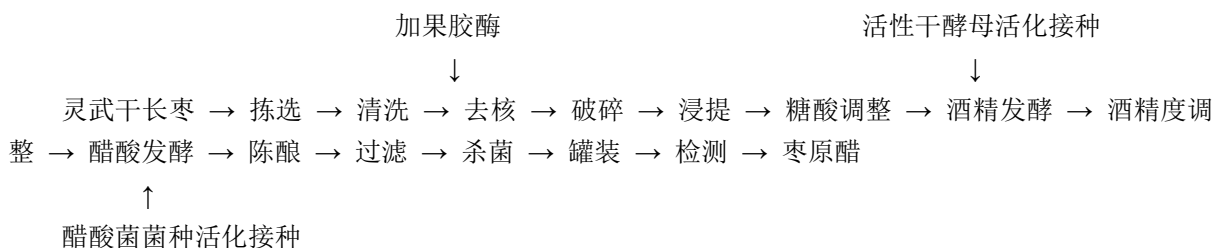
## 2. 材料与方法

### 2.1. 材料与试剂

灵武干长枣(市场自购), 果胶酶(安琪果胶酶), 酵母菌(安琪牌), 醋酸菌(河南雅大股份有限公司), 葡萄糖, 果葡糖浆, 硫酸铜, 酒石酸钾钠等均为国产分析纯。

### 2.2. 样本制备

枣醋酿造工艺流程:



采用二次发酵法进行枣醋的酿造, 先将干枣破碎、浸提、发酵成高纯度枣酒, 再进行醋酸发酵将枣酒发酵为枣原醋。其中还原糖含量、酒精度直接影响着发酵进程及转化率, 适时快速监测与调控这两项变量参数, 是确保枣醋酿造质量的关键环节。在酒精发酵、醋酸发酵、陈酿的 30 天中, 每天取样, 先用高光谱检测系统扫描, 采集数据后, 进行还原糖和酒精度的测定。

### 2.3. 理化指标测试仪器与设备

PL203 电子天平(梅特勒-托利多仪器(上海)有限公司); SW-CJ-IFD 无菌超净操作台(苏净集团安泰公司制造); 电热恒温水浴锅(北京长源实验设备厂); DSX-280B 高压灭菌锅(上海申安医疗器械厂); 酒精计(河北省武强县双明仪表厂); SPX-25°C 恒温培养箱(上海博讯实业有限公司); 722-型可见分光光度计(上海棱光技术有限公司); BS-2F 振荡培养箱

### 2.4. 理化指标常规测定方法

- 1) 总糖的测定: 菲林间接法[11]。
- 2) 酒精度的测定: 蒸馏法[12]。

### 2.5. 高光谱数据采集设备

实验使用光谱范围为 400~1000 nm 的 HyperSpec VNIR 可见近红外高光谱成像系统, 该系统生产于美国 Headwall Photonics 公司, 由高光谱成像光谱仪(Inspector N 系列, Golden Way Scientific Co., Ltd., US)光谱分辨率 2.5 nm; CCD 摄像机(G4-232, Golden Way Scientific Co. Ltd., US)像素尺寸 8.0 μm; 1 个

线光源(90-254VAC, 47~63 Hz, Golden Way Scientific Co., Lab., EQUIP)与 1 个卤钨灯(250 V, Headwall Photonics Instruments Co., Ltd., Beijing, China)构成的光源系统; 电控位移平台(VT-80, Headwall Photonics Instruments Co., Ltd., Beijing, China); 计算机(ThinkPadX220 Inter(R)), RAM3.41G 和数据采集软件(Hyperspec-N for AndorLuca Rev A.3.1.4.vi, Headwall Photonics Instruments Co., Ltd., Beijing, China)等 6 个部件组成。其中, 高光谱成像仪的波段范围为 400~1100 nm, 光谱分辨率是 2.5 nm, 因此在该范围内共有 125 个波长。从样品反射的光线被光谱仪中的分光系统分离成特征波长再进入物镜; 摄像头中形成的二维图像在电脑上保存。通过使用一个电动平台使样品在物镜面前移动, 并且过程重复; 二维图像在样品的邻点上被堆叠以形成三维超立方体, 存储在计算机中进行进一步的分析[13]。

## 2.6. 实验方法

### 2.6.1. 光谱数据采集

从枣醋酿造的第一天开始到发酵结束至陈酿阶段的 30 天中, 每天分别取 50 mL 的发酵液样 7 个。将发酵液样置于测试盒中, 编号。按采集的顺序放置于电控位移平台, 经黑白校正后进行光谱信息采集。系统参数设置如下: 光谱范围 400~1000 nm, 曝光时间为 15 ms, 步进速度为 160  $\mu\text{m/s}$ , 扫描起始位置为 90 mm, 扫描宽度为 80 mm。

### 2.6.2. 校正集与验证集的划分

将采集到的光谱数据分为校正集和验证集, 其中, 校正集用于建立相应的定量检测模型, 验证集用来检验所建立的模型的可靠性。在随机保留 30 个枣醋酿造过程样本作为独立测试集的基础上, 采用 Kennard-Stone (K-S)法将剩余 180 个样品进行样本集和验证集的划分。选择校正集样本 120 个, 验证集样本 60 个。校正集与验证集的值统计如表 1 所示。

### 2.6.3. 光谱预处理

高光谱在采集过程中非常容易受到基线漂移, 附带噪声和不重复性等干扰因素对模型的影响, 为了消除高光谱成像中干扰因素对模型的影响, 提高模型稳定性与准确度, 本试验采用多元散射校正(MSC)对光谱进行预处理。

### 2.6.4. 光谱波长变量选择

本实验分别采用 UVE 法、CARS 法、GA 法对全光谱 125 个波长变量进行优化选择, 同时采用 PLS 法建立模型。选取决定系数( $R^2$ )、预测标准偏差(RMSEP)、相对分析误差(RPD)以及最佳主因子数来评价模型稳定性与预测能力[14]。

### 2.6.5. 数据处理与分析

UVE、CARS、GA 等程序通过运行 MATLAB 完成, 偏最小二乘计算应用 Unscrambler X10.3 光谱分析软件(挪威 CAMO 公司)实现。

**Table 1.** Statistical results of calibration set and validation set

**表 1.** 校正集与验证集统计结果

	样品数	样本性质	平均值/( $\text{g}\cdot\text{L}^{-1}$ )	最大值/( $\text{g}\cdot\text{L}^{-1}$ )	最小值/( $\text{g}\cdot\text{L}^{-1}$ )	标准差
校正集	120	还原糖	6.69	16.5	2.2	4.48
		酒精度	4.98	11.9	0.0	4.63
验证集	60	还原糖	6.65	16.1	2.6	4.42
		酒精度	4.87	11.6	0.1	4.61

### 3. 结果与分析

#### 3.1. 光谱波段优选

##### 3.1.1. 无信息变量消除法(UVE)优选特征波长变量

在全光谱的基础上,对枣醋酿造过程中还原糖含量和酒精度指标利用无信息变量消除法(UVE)进行处理,UVE 优选特征波长变量结果如图 1 所示,竖直实线左边是枣醋的波长变量,右边为引入的系统噪音变量。虚线表示变量稳定性的上下阈值,处于两阈值之间的变量可认为是无关的信息量,超出阈值的部分为有用信息波长变量[15]。

##### 3.1.2. 竞争性自适应重加权算法(CARS)优选特征波长变量

CARS 在进行波长变量筛选时,保留回归系数绝对值大的变量,剔除回归系数绝对值小的变量,重复运行筛选出最佳变量子集。如图 2(A)中(a)~(c)各分图所示,(a)表示随着运行次数增加,变量数的下降趋势由前段快速减少到后段逐渐平缓,反映了 CARS 的先“粗选”和后“精选”过程;(b)表示交互验证均方

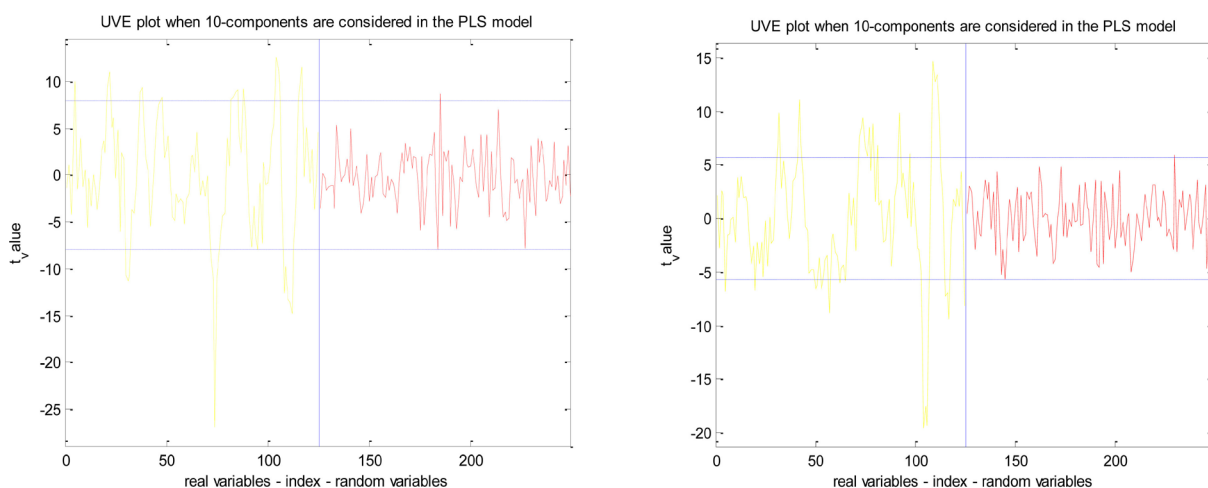


Figure 1. UVE variable stability analysis results

图 1. UVE 变量稳定性分析结果

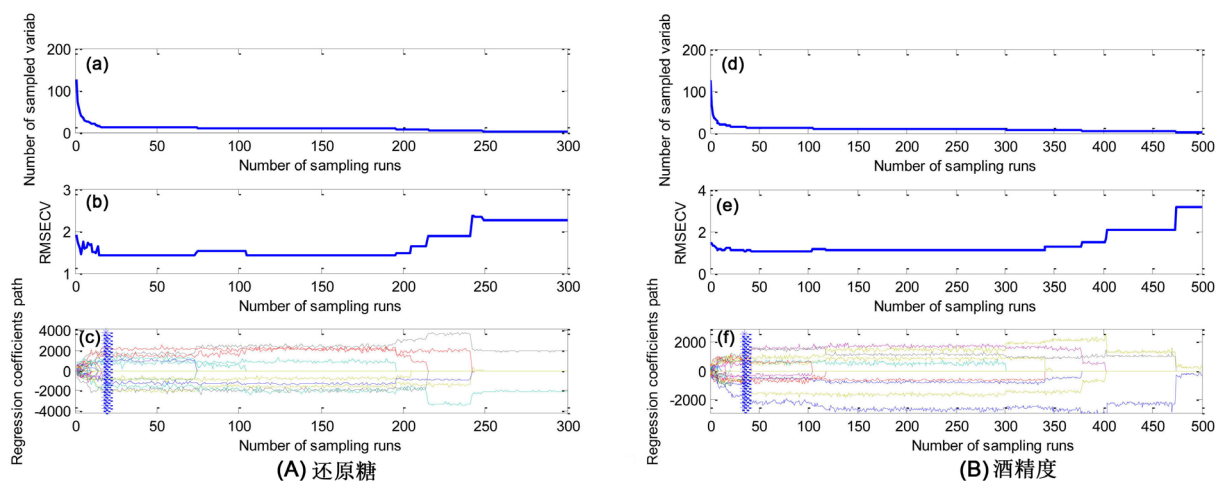


Figure 2. Key variable selection results by CARS

图 2. CARS 法关键变量选择结果



差的变化趋势,随着运行次数的增加, RMSECV 值先减少后增大,先减少表明筛选过程剔除与样本性质无关的变量,后又增大则可能剔除了关键变量,从而导致残差增大;(c) 表示回归系数的变化趋势,与“\*” 相对的点为 RMSECV 值的最低点。经 CARS 法筛选所得变量数分别为 19 和 35 个,仅占全光谱的 15.2%、28%,表明通过波长变量筛选,可有效降低模型的复杂程度,提高预测效率。

### 3.1.3. 遗传偏最小二乘波段选择法优选特征波长变量

GA 是基于生物进化论,模拟自然界进化机制的一种优化算法,通过选择频率最高的波长建模来挑选特征波长变量。设定 GA 优化参量:初始群体为 30,交叉概率为 0.5,变异概率为 0.01,遗传迭代次数为 100。图 3 中(A)、(B)分别为还原糖、酒精度含量两个指标通过 GA 法筛选后的各变量被选用的频次图,分别筛选出 26、37 个波长变量,占全光谱的 20.8%、29.6%。

## 3.2. 模型建立与评价

经过上述三种方法的筛选,分别建立枣醋酿造过程中还原糖、酒精度的全谱-PLS、UVE-PLS、CARS-PLS、GA-PLS 定量模型,并对决定系数( $R^2$ )、预测标准偏差(RMSEP)、相对分析误差(RPD)以及最佳主因子数进行比较,来评价模型效果,模型优化结果如表 2 所示。

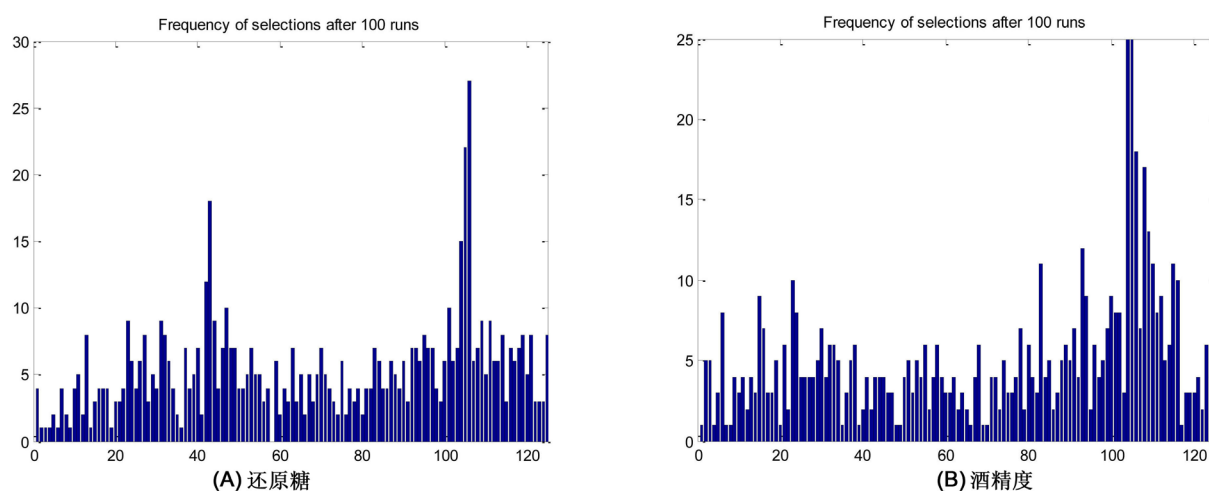


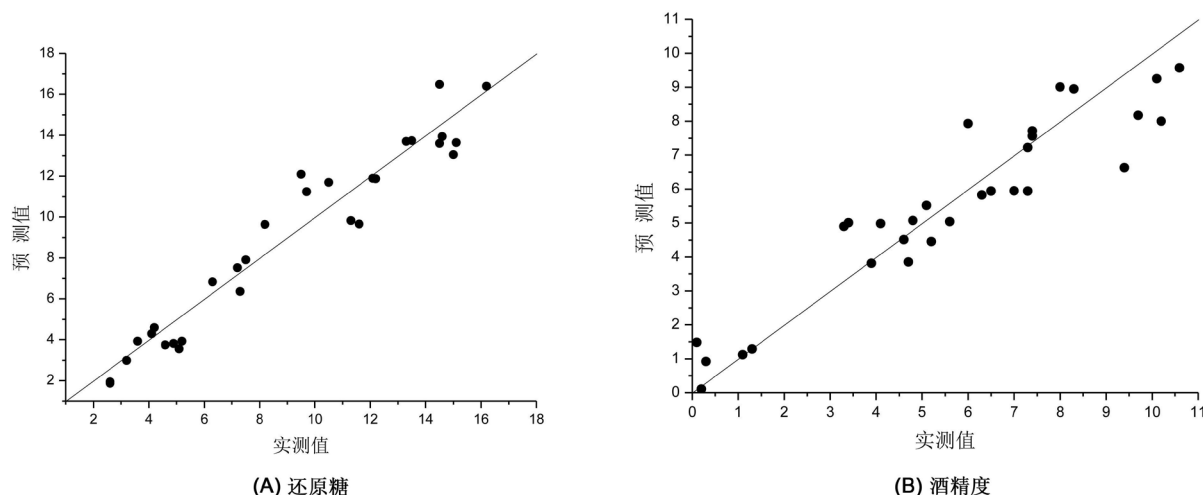
Figure 3. The frequency of each variable by chosen

图 3. 各变量被选用的频次图

Table 2. Different models and performance evaluation results of reducing sugar and alcoholic strength

表 2. 还原糖和酒精度的不同模型及性能评价结果

样本性质	模型	变量数	$R^2$	RMSEP	RPD	主成分数
还原糖	全谱-PLS	125	0.8422	1.6278	2.71	12
	UVE-PLS	31	0.8566	1.4898	2.97	12
	CARS-PLS	19	0.9045	1.3635	3.24	7
	GA-PLS	26	0.8878	1.4356	3.07	10
酒精度	PLS-全谱	125	0.8449	1.5309	3.01	12
	UVE-PLS	34	0.8856	1.3675	3.37	9
	CARS-PLS	35	0.8993	1.2878	3.58	8
	GA-PLS	37	0.8843	1.3705	3.36	9



**Figure 4.** Predicted vs. reference values of reducing sugar and alcoholic strength in CARS model

**图 4.** 还原糖和酒精度的 CARS 模型实测值与预测值分布

从表 2 可看出,与全光谱建模相比,采用上述 3 种方法进行变量筛选后,建模所用变量数显著减少, $R^2$  普遍比全光谱增大, RMSEP 和最佳主因子数相对减少,模型分析效果均优于全谱模型。UVE、CARS 和 GA 淘汰了光谱中共线性变量及受外界因素影响较大的波长变量,优选出关键性变量,减少了变量数,有效地降低了模型复杂程度,提高了模型精确度。相比之下,GA 和 UVE 特征变量筛选结果不如 CARS 理想,CARS 方法更适用于高光谱数据特征波长变量的筛选,可选出高光谱数据中信息更强且对其他外界影响因素不敏感的变量,模型的预测效率得到极大提升的同时不失参数代表性,模型效果略优。

其中,采用 CARS 筛选特征波长变量后所建立的枣醋酿造过程中,还原糖含量和酒精度 CARS-PLS 定量模型的  $R^2$  分别达到 0.9045 和 0.8993, RMSEP 为 1.3635 和 1.2878, RPD 为 3.24 和 3.58,建模所用变量数最少,优化效果最佳,模型稳定性和预测能力最佳。

### 3.3. 模型验证

将独立样本测试集中 30 个样品的光谱通过 CARS-PLS 模型进行验证,如图 4 中(A)、(B)所示,还原糖含量和酒精度的实测值与预测值点呈现对角线分布。且经过成对 t 检验,还原糖含量和酒精度的实测值与预测值无显著差异,说明模型的预测精度较高。

## 4. 结论

本研究采用可见-近红外高光谱检测技术对枣醋酿造过程中还原糖含量和酒精度指标进行了定量分析。分别采用无信息变量消除法(UVE)、竞争性自适应重加权算法(CARS)、遗传算法(GA)对光谱特征波长进行变量筛选,并结合偏最小二乘法(PLS)建立模型。结果表明,波长变量筛选在保留还原糖和酒精度含量特征波长的同时可剔除大量冗余无效信息,简化了模型复杂程度,提高了模型准确度和稳定性。采用 CARS 法进行波段筛选后所建模型的效果优于 GA 法和 UVE 法。枣醋酿造过程中,还原糖含量和酒精度 CARS-PLS 定量模型的  $R^2$  分别达到 0.9045 和 0.8993, RMSEP 为 1.3635 和 1.2878, RPD 为 3.24 和 3.58。因此,基于可见-近红外高光谱检测技术,利用 CARS-PLS 法建立模型,可满足对枣醋酿造过程中还原糖含量和酒精度进行快速、实时、准确检测的需求,为以上指标的实时快速检测提供了理论依据。

## 基金项目

宁夏高校科学研究项目(No.NGY2016019);宁夏十三五重点专业建设项目。

## 参考文献

- [1] 李宏高, 吴忠会, 白文涛, 等. 红枣醋饮料工艺研究[J]. 食品科学, 2006, 27(10): 645.
- [2] 李红蕊, 李志西, 赵晓野, 等. 红枣醋和枳椇减肥降血脂作用研究[J]. 西北农业学报, 2009, 18(2): 257-260.
- [3] 李秀萍, 郑平, 吴幼茹, 等. 甘蔗果酒酿造酵母筛选及其发酵特性研究[J]. 中国酿造, 2014, 33(6): 63-66.
- [4] 吴龙国, 王松磊, 康宁波, 等. 基于可见-近红外高光谱成像技术的灵武长枣缺陷识别[J]. 农业工程学报, 2015(20): 281-286.
- [5] 吴龙国, 何建国, 贺晓光, 等. 高光谱图像技术在水果无损检测中的研究进展[J]. 激光与红外, 2013, 43(9): 990-996.
- [6] 石吉勇. 基于高光谱图像技术的设施栽培作物营养元素亏缺诊断研究[D]: [博士学位论文]. 镇江: 江苏大学, 2012.
- [7] 王雷, 乔晓艳, 董有尔, 等. 高光谱图像技术在农产品检测中的应用进展[J]. 应用光学, 2009, 30(4): 639-645.
- [8] 田有文, 王晓娟. 基于高光谱图像技术的农产品品质无损检测[J]. 农机化研究, 2009, 31(10): 220-222.
- [9] 邹小波, 申婷婷, 朱瑶迪, 等. 基于高光谱图像的香醋醋醅均匀性快速判别[J]. 农业机械学报, 2016, 47(1): 215-220.
- [10] 申婷婷. 高光谱图像技术检测镇江香醋醋醅理化指标分布研究[D]: [硕士学位论文]. 镇江: 江苏大学, 2016.
- [11] 张意静. 食品分析技术[M]. 北京: 中国轻工业出版社, 2001: 149-151.
- [12] 中国标准出版社. 中华人民共和国国家标准: 食品卫生检验方法(理化部分)[M]. 北京: 中国标准出版社, 2012: 205-206.
- [13] Gowen, A., O'Donnell, C., Cullen, P., et al. (2007) Hyperspectral Imaging—An Emerging Process Analytical Tool for Food Quality and Safety Control. *Trends in Food Science & Technology*, **18**, 590-598.  
<https://doi.org/10.1016/j.tifs.2007.06.001>
- [14] 张初, 刘飞, 孔汝汝. 利用近红外高光谱图像技术快速鉴别西瓜种子品种[J]. 农业工程学报, 2013, 29(20): 270-277.
- [15] 熊雅婷, 李宗朋, 王健, 等. 近红外光谱波段优化在白酒酒醋成分分析中的应用[J]. 光谱学与光谱分析, 2016, 36(1): 84-90.

### 知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2164-5507, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [hjas@hanspub.org](mailto:hjas@hanspub.org)