

网络出版时间:2013-05-02 10:22

网络出版地址: <http://www.cnki.net/kcms/detail/61.1390.S.20130502.1022.003.html>

食醋总酸和还原糖含量近红外光谱检测模型的优化研究

郝 勇¹, 余永建², 朱胜虎²

(1 华东交通大学 机电工程学院, 江西 南昌 330013; 2 江苏恒顺醋业股份有限公司, 江苏 镇江 212143)

[摘 要] **【目的】**采用近红外光谱(Near-infrared spectroscopy, NIRS)分析技术对食醋中的总酸和还原糖进行快速定量分析,为食醋品质的检测提供技术支持。**【方法】**用近红外透射光谱法采集食醋的光谱信息,用 Savitzky-Golay 一阶导数和小波导数(Wavelet derivative, WD)对采集到的食醋光谱进行预处理,采用连续投影算法(Successive projections algorithm, SPA)的波长筛选方法对食醋的总酸以及还原糖定量分析模型进行优化。**【结果】**食醋光谱经小波导数预处理和基于 SPA 的波长优选后,总酸和还原糖含量分析精度得以提高,模型得到简化。总酸的预测均方根误差(Root mean square error of prediction, RMSEP)由 0.282 降低为 0.227,预测相关系数(Correlation coefficient of prediction, R_p)由 0.965 提高到 0.977,建模变量由 680 个精简为 51 个;还原糖的 RMSEP 由 0.627 降低为 0.606,预测相关系数 R_p 由 0.996 提高至 0.997,建模变量精简为 50 个。**【结论】**利用近红外光谱分析技术结合小波导数和 SPA 分析方法,可用于食醋品质的定量分析,为食醋品质的快速检测提供了一种可行的方法。

[关键词] 食醋;品质检测;总酸;还原糖;小波导数;连续投影算法;近红外光谱

[中图分类号] O657.33;TS264.2⁺2 **[文献标志码]** A **[文章编号]** 1671-9387(2013)05-0094-05

Model optimization of near-infrared spectroscopy for determination of total acid and reducing sugar in vinegar

HAO Yong¹, YU Yong-jian², ZHU Sheng-hu²

(1 College of Mechanical and Electronic Engineering, East China Jiaotong University, Nanchang, Jiangxi 330013, China;

2 Jiangsu Hengshun Vinegar Industry CO. LTD., Zhenjiang, Jiangsu 212143, China)

Abstract: **【Objective】** Near-infrared spectroscopy (NIRS) was used for the determination of total acid and reducing sugar in vinegar samples. It can provide technical support for vinegar quality detection. **【Method】** Near-infrared transmittance spectroscopy was adopted for spectral acquisition of vinegar samples. Savitzky-Golay 1st derivative and wavelet derivative were used for spectral pretreatment. Wavelength selection method of successive projections algorithm (SPA) was used for model optimization of total acid and reducing sugar in vinegar samples. **【Result】** The results showed that the wavelet derivative (WD) combined with SPA could not only improve the analysis precision, but also simplify the modeling variables. For the total acid model, the root mean square error of prediction (RMSEP) and modeling variables were reduced from 0.282 and 680 to 0.227 and 51, respectively, and the correlation coefficient of prediction (R_p) was improved from 0.965 to 0.977. For the reducing sugar model, the RMSEP and modeling variables were

[收稿日期] 2012-08-02

[基金项目] 国家自然科学基金项目(21265006);江西省青年科学基金项目(20114BAB2011010);华东交通大学博士启动基金项目(01309021);江苏省企业博士集聚计划项目

[作者简介] 郝 勇(1978—),男,内蒙古呼和浩特人,讲师,博士,主要从事光谱分析及化学计量学方法研究。

E-mail: haonm@163.com

reduced form 0.627 and 680 to 0.606 and 50, respectively, and the R_p was improved from 0.996 to 0.997. **【Conclusion】** NIRS combined with WD-SPA method is a feasible and fast method for the determination of vinegar quality.

Key words: vinegar; quality determination; total acid; reducing sugar; wavelet derivative; successive projections algorithm (SPA); near-infrared spectroscopy (NIRS)

随着食醋酿造工艺和口感风味的多样化,食醋作为一种调味佳品或饮品已逐步走进了广大消费者的厨房。在全国大多数地区都有酿醋作坊和地方醋产品品牌,如恒顺香醋、山西陈醋、天津独流老醋以及四川保宁醋等。依据生产工艺的不同,可将食醋分为酿造食醋和配制食醋两大类。其中酿造食醋是指单独或混合使用各种含有淀粉、糖的原料或酒精,进行微生物发酵酿制而成的液体调味品;配制食醋是以酿造食醋为主体,将50%以上的酿造食醋与食用冰醋酸以及食品添加剂等混合配制而成的调味食醋。对于各种食醋,国家标准 GB 18187—2000 对其中的总酸、还原糖、不挥发酸、总脂以及可溶性无盐固形物等含量都有严格的要求,同时也是食醋分级的重要依据。然而,由于食醋行业中允许勾兑,使得一些不法分子采用硫酸、尿素、工业盐及碱等工农业用品制造食品色素,这些色素流入食醋等酿造企业,通过和醋精、防腐剂进行勾兑来生产“陈醋”;还有一些酿造企业采用勾兑食醋假冒酿造食醋,严重影响了食醋行业的良性发展和消费者的身体健康。

食醋中的总酸和还原糖含量是衡量食醋品质的两项重要指标。目前,国家标准规定的测量方法主要为化学滴定法,需要配制多种试剂,并且在测量前需进行标定和样品预处理,整个检测过程复杂耗时(大约需要3 h),对检测人员的技术水平要求较高,且分析误差较大^[1-2]。近红外光谱(Near-infrared spectroscopy, NIRS)由于其分析效率高、速度快、成本低、非破坏性和易于在线分析等特点被广泛应用于复杂样品的定性和定量分析^[3-6]。然而 NIRS 不仅可以反映物质的化学组成和含量,同时也包含了由被测物的温度、表面纹理、密度以及内部组分分布不均匀性等因素引起的光谱响应。因此,在进行光谱分析时,有效信息的提取、分析模型的建立和优化均非常重要^[7-8]。目前,将近红外透射光谱分析法用于食醋总酸和还原糖的研究已有报道^[9-12]。陈斌等^[13]采用红外光谱对18份恒顺香醋的总酸和还原糖分别采用多元线性回归方法进行定量分析,表明采用三元线性回归建立的校正模型的相关系数均大于0.990;李华北等^[14]采用近红外光谱对66份恒顺

香醋的还原糖进行了定量分析,表明小波变换结合三元线性回归可以实现还原糖的定量分析,相关系数为0.991。上述研究分别对总酸和还原糖进行了定量分析,然而由于样品数量较少且原产地来源较为单一,所建分析模型的适应性和稳健性尚需进一步验证。

为了将近红外光谱分析方法应用于食醋总酸和还原糖的定量分析,本研究分别收集了江苏、天津和山西等省市不同食醋酿造企业的食醋样品,采用近红外透射光谱分析技术,结合不同的光谱预处理方法,建立食醋总酸和还原糖定量分析的偏最小二乘(Partial least squares, PLS)模型,并引入连续投影算法(Successive projections algorithm, SPA)对分析模型进行简化,以期对食醋品质的近红外光谱快速分析提供试验基础和方法参考。

1 材料与方法

1.1 食醋样品的收集

样品来源于江苏恒顺、天津独流、山西东湖、山西水塔以及山西紫林公司提供的瓶装食醋,共计80份样品,生产年份的跨度为2006—2011年。

1.2 总酸和还原糖含量的测定

食醋总酸和还原糖的质量分数测定按照 GB/T 12456—2008 和 GB/T 5009.7—2008 的方法进行。每个样品平行测定2次,最终结果取2次测定的平均值。根据食醋样品总酸和还原糖质量分数的平均分布原则,将80份样品按照近似2:1的比例划分为校正集和验证集,以保证分组后样品集的平均值和标准偏差较为接近。样品的统计信息如表1所示。

1.3 近红外光谱的采集

用 Tensor37(德国 Bruker 公司)傅立叶变换近红外光谱仪(RT-Pbs 检测器)进行食醋样品光谱的采集。以空气为参比,将食醋样品注入2 mm 比色皿,光谱采用64次扫描,分辨率为 8 cm^{-1} ,光谱采集范围为 $5\,850\sim 11\,290\text{ cm}^{-1}$,每条光谱包含680个数据点,每个样品测试3次,以平均光谱作为最终建模光谱。食醋的近红外透射光谱如图1所示。

表 1 80 份食醋样品总酸和还原糖质量分数的测定与统计

Table 1 Statistical information of total acid and reducing sugar contents in vinegar samples

品质指标 Quality indexes	样品集 Sample dataset	样品数 Sample number	质量分数/% Mass fraction			
			最小值 Minimum	最大值 Maximum	平均值 Average	标准偏差 Standard deviation
总酸 Total acid	校正集 Calibration	53	2.51	7.03	4.94	1.01
	验证集 Validation	27	2.51	6.61	5.01	1.04
还原糖 Reducing sugar	校正集 Calibration	53	0.98	30.65	5.00	6.04
	验证集 Validation	27	1.11	30.65	5.61	7.37

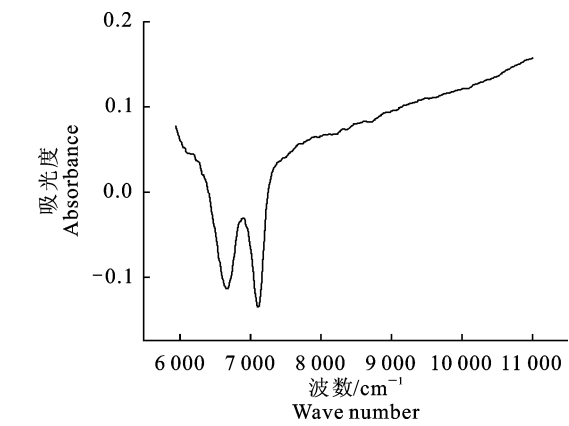


图 1 食醋样品的近红外透射光谱图

Fig. 1 NIR transmittance spectra of a vinegar sample

1.4 模型的评价指标

采用 PLS 建立光谱模型,PLS 因子数采用蒙特卡罗交互验证(Monte carlo cross-validation, MC-CV)法确定,模型的评价指标包括交叉验证均方根误差(Root mean square error of cross validation, RMSECV)、预测均方根误差(Root mean square error of prediction, RMSEP)和预测相关系数(Correlation coefficient of prediction, R_p)。采用 Matlab R2010a 进行数据处理和模型构建。其中 RMSECV 和 RMSEP 值越小, R_p 值越大;且模型的 RMSECV 和 RMSEP 值越接近,则模型的稳健性和预测能力越好。

表 2 不同光谱预处理方法处理后模型校正和预测结果比较

Table 2 Calibration and validation results of models with different preprocessing methods

品质指标 Quality indexes	预处理方法 Preprocessing methods	因子数 Factor	RMSECV	RMSEP	R_p
总酸 Total acid	原始光谱 Raw spectra	8	0.223	0.282	0.965
	SG	15	0.363	0.428	0.918
	WD	7	0.267	0.233	0.976
还原糖 Reducing sugar	原始光谱 Raw spectra	8	0.786	0.627	0.996
	SG	5	1.127	0.943	0.993
	WD	7	0.777	0.617	0.996

比较表 2 中 SG 导数和 WD 导数 2 种方法的处理结果可知,虽然 2 种导数方法都可以消除光谱的

2 结果与分析

2.1 光谱预处理方法的优化

食醋为液体样品,在采集近红外光谱时不存在散射效应,因此能够对其光谱产生影响的仅是由于温度变化引起的基线漂移。导数运算是一种比较理想的消除光谱基线漂移的方法。对光谱信号进行求导运算可以增加光谱的分辨率,使重叠光谱得到进一步分离,从而提高光谱的信噪比;但是导数运算对高频噪声特别敏感,因此在对光谱进行求导运算前,需要对光谱进行适度的平滑,而且求导的阶数不能太高。

光谱求导主要包括 Savitzky-Golay(SG)导数和小波导数^[16](Wavelet derivative, WD)2 种方法,表 2 所示为食醋总酸和还原糖质量分数经不同导数处理后的建模结果,其中 SG 导数采用 5 点 3 次拟合进行一阶导数运算,WD 采用“Haar”小波基,小波尺度为 100。由表 2 可知,对于食醋总酸的近红外模型,经 SG 导数预处理后,模型的预测结果比原始光谱差;光谱经 WD 处理后,与原始光谱相比,模型的因子数由 8 变为 7,模型的 R_p 由 0.965 提高至 0.976,而且模型的 RMSECV 和 RMSEP 更加接近,模型更加稳健。对于还原糖光谱模型,光谱经 WD 预处理后,模型的因子数、RMSECV 和 RMSEP 都得到了减小,模型的精度得到了提高。

基线漂移,然而 WD 处理后的光谱比原始光谱和 SG 导数光谱更加平滑(图 2),有效地抑制了高频噪声。

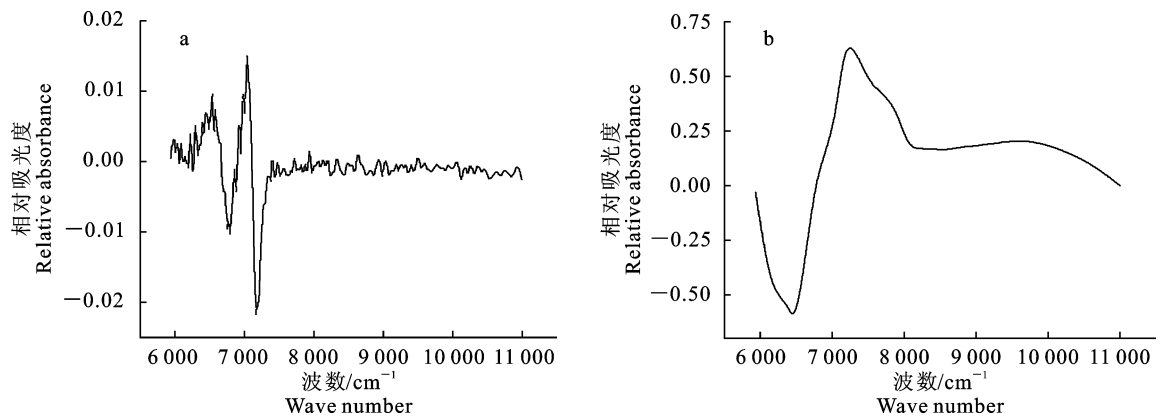


图 2 食醋近红外光谱经 SG(a)、WD(b)导数处理后的光谱图

Fig. 2 NIR transmittance spectra for vinegar by SG(a) and WD(b) derivative

2.2 光谱建模变量的选择

由于近红外光谱的波长点大量冗余,利用光谱的全谱信息建模往往会使模型的稳健性和预测能力下降。对于复杂样品的近红外光谱,尽管 PLS 等多元校正方法可以有效地提取与分析组分的相关信息,但对于大部分样品而言,完全消除其他组分对目标组分的干扰仍非常困难。因此,采用变量筛选方法进行建模波长优选,选择与待测目标组分相关的光谱变量进行建模,可以提高模型的稳定性和预测能力。本研究采用的样品分别来自江苏、山西和天津 3 个省市,食醋酿造工艺和所采用的水质决定了不同食醋样品的光谱间存在背景差异,因此采用变量筛选方法提取与目标组分相关的特征波长点。

SPA 方法能够有效消除光谱中众多变量间的共线性影响,使向量间的共线性达到最小,从而降低模型的复杂度,因此在光谱变量选取中得到了广泛应用^[10]。表 3 所示为食醋总酸和还原糖定量分析模型的参数经 SPA 方法优选后,模型的保留变量数、PLS 因子数以及模型的校正和预测结果。与表 2 的计算结果相比,光谱经 WD-SPA 方法处理后,食醋总酸和还原糖定量分析模型的精度均得到明显提高,建模变量大大简化,由原始光谱的 680 个分别减少为 51 和 50 个。选择的变量分布如图 3 所示,这些变量主要集中于低波数区(6 000~8 000 cm⁻¹),该区域主要反映了 C—H 基团的倍频吸收^[15]。

表 3 食醋总酸和还原糖定量分析模型经 SPA 处理后的变量、因子数及校正和预测结果

Table 3 Calibration and validation results, variables, and number of factors for models of total acid and reducing sugar processed by SPA

品质指标 Quality indexes	保留变量数 Retained variable number	因子数 Factor	RMSECV	RMSEP	R _p
总酸 Total acid	51	7	0.263	0.227	0.977
还原糖 Reducing sugar	50	7	0.763	0.606	0.997

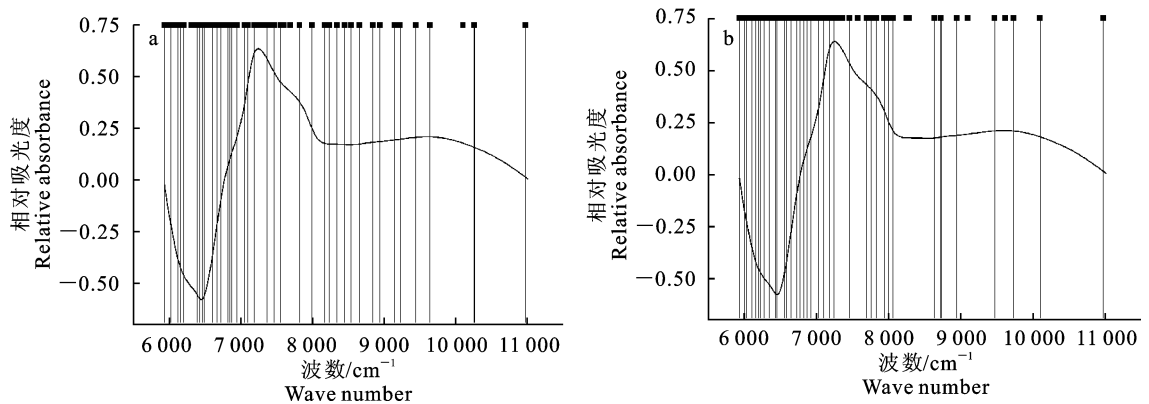


图 3 食醋总酸(a)和还原糖(b)定量分析模型选择的建模变量示意图

Fig. 3 Selected modeling variables for total acid (a) and reducing sugar (b)

图 4 所示为验证集样品总酸和还原糖的实测值与模型预测值的相关关系。由图 4 可知,对于总酸定量分析模型,总酸质量分数集中于 3.5%~6.5%,且分布比较均匀,无法从酸度的分布情况判断食醋的品牌;对于还原糖定量分析模型,大部分样

品的还原糖质量分数集中于 2%~7%,只有 3 个果醋样品的还原糖质量分数大于 10%,虽然质量分数分布较为分散,但模型预测集的相关系数仍然较高,说明模型具有很好的外延性。

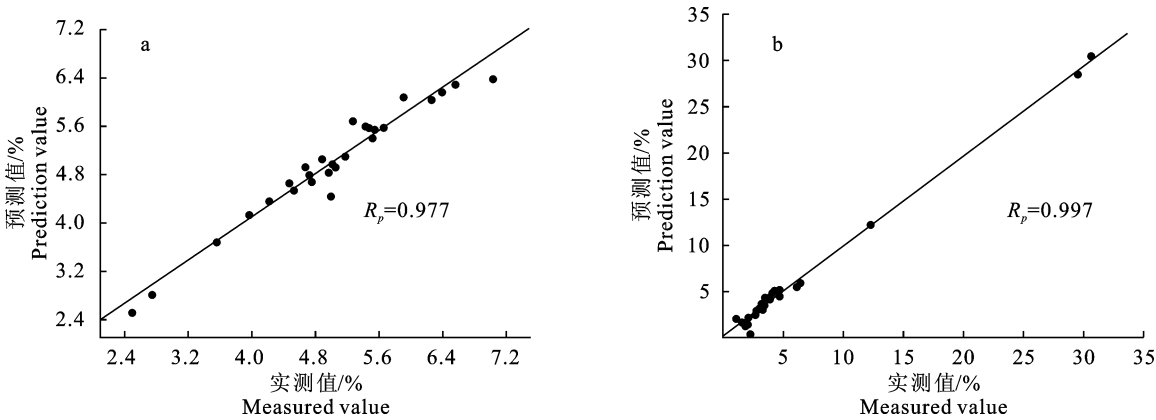


图 4 食醋总酸(a)和还原糖(b)质量分数实测值和预测值的相关分析

Fig. 4 Relationship between predicted values with the reference values for total acid(a) and reducing sugar(b)

3 结 论

本研究通过收集多家食醋生产企业的食醋样品,以实现食醋中总酸和还原糖的近红外光谱快速定量分析为目标,采用 PLS 方法对食醋的近红外透射光谱进行建模,结合小波导数(WD)预处理方法和 SPA 变量筛选方法,对分析模型进行了优化精简。研究表明,连续小波导数运算可以实现光谱的求导和平滑;小波导数结合 SPA 方法不仅可以提高食醋总酸和还原糖的分析精度,而且可以极大地简化建模变量数。近红外光谱结合小波导数和 SPA 方法可用于食醋总酸和还原糖的定量分析,为食醋品质检测提供了技术支持。

[参考文献]

[1] Saiz-Abajo M J, Gonzalez-Saiz J M, Pizarro C. Prediction of organic acids and other quality parameters of wine vinegar by near-infrared spectroscopy: A feasibility study [J]. Food Chemistry, 2006, 99(3): 615-621.

[2] Pizarro C, Esteban-Diez I, Saenz-Gonzalez C, et al. Vinegar classification based on feature extraction and selection from headspace solid-phase microextraction/gas chromatography volatile analyses: A feasibility study [J]. Analytica Chimica Acta, 2008, 608(1): 38-47.

[3] 刘 卉, 郭文川, 岳 绒. 猕猴桃硬度近红外漫反射光谱无损检测 [J]. 农业机械学报, 2011, 42(3): 145-149.

Liu H, Guo W C, Yue R. Non-destructive detection of kiwifruit firmness based on near-infrared diffused spectroscopy [J].

Transactions of the Chinese Society for Agricultural Machinery, 2011, 42(3): 145-149. (in Chinese)

[4] 石吉勇, 邹小波, 赵杰文, 等. 黄瓜叶片叶绿素含量近红外光谱无损检测 [J]. 农业机械学报, 2011, 42(5): 178-182.

Shi J Y, Zou X B, Zhao J W, et al. NIR spectra in non-invasive measurement of cucumber leaf chlorophylls content [J]. Transactions of the Chinese Society for Agricultural Machinery, 2011, 42(5): 178-182. (in Chinese)

[5] Gaydou V, Kister J, Dupuy N. Evaluation of multiblock NIR/MIR PLS predictive models to detect adulteration of diesel/biodiesel blends by vegetal oil [J]. Chemometrics and Intelligent Laboratory Systems, 2011, 106(2): 190-197.

[6] Yan H, Han B X, Wu Q Y, et al. Rapid detection of *Rosa laevigata* polysaccharide content by near-infrared spectroscopy [J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2011, 79(1): 179-184.

[7] Shen F, Ying Y B, Li B B, et al. Prediction of sugars and acids in Chinese rice wine by mid-infrared spectroscopy [J]. Food Research International, 2011, 44(5): 1521-1527.

[8] Sinelli N, Casiraghi E, Barzaghi S, et al. Near infrared (NIR) spectroscopy as a tool for monitoring blueberry osmo-air dehydration process [J]. Food Research International, 2011, 44(5): 1427-1433.

[9] Smyth H E, Cazzolino D, Cynkar W U, et al. Near infrared spectroscopy as a rapid tool to measure volatile aroma compounds in riesling wine: Possibilities and limits [J]. Analytical and Bioanalytical Chemistry, 2008, 390(7): 1911-1916.

[10] Liu F, He Y. Application of successive projections algorithm for variable selection to determine organic acids of plum vinegar [J]. Food Chemistry, 2009, 115: 1430-1436.