

# 三级大豆油酸价的近红外光谱检测

王铭义<sup>1</sup>, 郭建英<sup>1,\*</sup>, 张佳宁<sup>2</sup>, 李越<sup>2</sup>, 于殿宇<sup>2</sup>

(1. 哈尔滨理工大学测控技术与通信工程学院, 黑龙江 哈尔滨 150086;

2. 东北农业大学食品学院, 黑龙江 哈尔滨 150030)

**摘要:** 为改进现有的油脂检测方法, 以三级大豆油脂酸价检测为研究对象, 利用近红外透射光谱技术, 结合化学计量学方法, 选择  $4500\sim 6000\text{cm}^{-1}$  为特征波段建立油脂酸价校正模型, 首先研究基于小波变换的光谱预处理方法, 通过详细比较不同小波分解层数对建模的影响。结果确定 db4 小波的四层分解去噪效果最佳。结合滤波后重构的光谱信号建立油脂酸价 BP 神经网络校正模型, 利用预测集样本对模型进行验证, 决定系数  $R^2$  和预测均方根误差分别为 0.9743 和 0.1036。证明利用近红外光谱分析技术快速检测油脂酸价是完全可行的。

**关键词:** 油脂酸价检测; 近红外光谱分析; 小波预处理; 偏最小二乘回归; BP 神经网络

## Determination of Acid Value in Third Grade Soybean Oil by Near Infrared Spectroscopy

WANG Ming-yi<sup>1</sup>, GUO Jian-ying<sup>1,\*</sup>, ZHANG Jia-ning<sup>2</sup>, LI Yue<sup>2</sup>, YU Dian-yu<sup>2</sup>

(1. School of Measurement-control Technology and Communication Engineering, Harbin University of Science and Technology, Harbin 150086, China; 2. School of Food, Northeast Agricultural University, Harbin 150030, China)

**Abstract:** Near-infrared spectroscopy combined with chemometric methods was employed to determine the acid value of third grade soybean oil as an improvement over the current methods. The characteristic waveband of  $4500\sim 6000\text{cm}^{-1}$  was selected to establish an acid value calibration model. Spectral pre-processing methods based on wavelet transform were investigated and the effects of different wavelet decomposition levels on prediction results of acid value were compared in detail. The results showed that Daubechies4 (db4) wavelets with four levels of decomposition revealed the best noise removal. A back propagation (BP) neural network calibration model was built based on reconstructed spectral signals with wave filtering and was validated using the prediction set with a determination coefficient of 0.9743 and a RMSEP of 0.1036. This study supports a high feasibility to apply near-infrared spectroscopy for the rapid determination of acid value of oils.

**Key words:** acid value detection in oil; near infrared spectroscopy analysis; wavelet pre-processing; partial least square regression; BP neural network

中图分类号: TS207.3

文献标识码: A

文章编号: 1002-6630(2011)22-0171-04

油脂作为人类膳食的必需营养素之一、人体能量的来源, 在人类生活中具有极其重要的作用。随着人民生活水平的提高, 对油脂质量的要求也越来越高。酸价是衡量油脂质量的主要指标之一, 反映了油脂加工的质量以及储藏过程中品质的变化, 是油脂质量的必检项目。在油脂加工过程中, 酸价的大小也是油脂碱炼脱酸时计算用碱量的依据, 需要经常进行检测、调控, 以求得油脂加工工艺的最佳脱酸条件、提高产品质量, 从而获得最大的利润。因此油脂酸价测定在实际生产中

有着十分重要的作用<sup>[1]</sup>。

在我国, 按照国标要求, 油脂酸价的测定采用滴定法, 该法为一种主观的测定方法, 需要有经验的检验人员才能得到较好的可靠性和重现性<sup>[2]</sup>。对相同的样品, 不同操作者因视觉判断差异会得到不一致的测定结果, 即存在个人偏差, 特别是在判断滴定终点时, 不容易掌握准确, 尤其是操作不甚熟练的人员测定时, 其误差较大。虽然人们采取了各种方法来解决这些问题<sup>[3]</sup>, 但终究还是属于传统的实验室检测化验方法。最重要的

收稿日期: 2011-05-31

基金项目: 国家“863”计划项目(2010AA101503); 黑龙江省教育厅科学技术研究项目(11551109)

作者简介: 王铭义(1973—), 男, 博士研究生, 研究方向为测试计量技术及仪器。E-mail: wang\_mingyi163@163.com

\* 通信作者: 于殿宇(1964—), 男, 教授, 硕士, 研究方向为油脂加工。E-mail: dyu2000@yahoo.com.cn

是这种方法只能进行间断操作,不能实现连续化和自动化。因此,本研究提出将近年来发展迅速并广泛应用的近红外光谱分析技术引入油脂酸价检测中,该方法的优点在于分析成本低、绿色环保、操作简便、不必对样本预处理、易于实现在线检测和调控<sup>[4-6]</sup>。利用近红外光谱分析进行油脂酸价在线检测,可以进一步提高成品油的合格率,优化各项工艺指标,降低生产消耗和产品成本,提升我国油脂加工行业的自动化水平。

## 1 材料与方法

### 1.1 样本化学值的测定

目前我国对油脂酸价化学值的测定是采用国标法(GB 5530—85《植物油脂检验:酸价测定法》):先利用中性乙醚-乙醇(2:1, V/V)混合溶液溶解样本油脂及其所含的脂肪酸,再用碱标准溶液(通常为0.1mol/L氢氧化钾溶液)利用酸碱中和原理进行滴定,最后根据选取样本试样的质量以及所消耗碱溶液的体积计算出酸价值。计算公式如下:

$$\text{酸价}/(\text{mg KOH/g}) = \frac{V \times c \times 56.1}{m}$$

式中:  $V$  为滴定消耗的氢氧化钾溶液体积/mL;  $c$  为 KOH 溶液浓度/(mol/L);  $m$  为试样质量/g; 56.1 为 KOH 的摩尔质量/(g/mol)。

两次平行试验结果允许差不超过 0.2mg KOH/g, 求其平均数, 即为测定结果。

本研究从黑龙江省油脂加工企业生产线上收集了非转基因大豆生产的三级油脂样本 60 个, 采用国标法测定其酸价值, 经过筛选, 最后确定了 49 个典型样本用于油脂酸价近红外光谱分析, 样本酸价化学值范围为 0.473~3.005。将 49 个样品按酸价大小排序, 从中按一定间隔选出 9 个作为验证集, 其余 40 个作为校正集建立模型。

### 1.2 样本光谱数据的采集

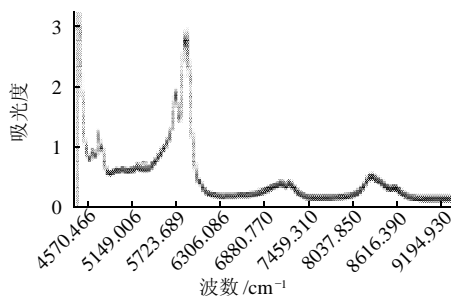


图1 大豆油脂近红外光谱图

Fig.1 Near-infrared spectra of soybean oil samples

利用美国 Thermo Nicolet 公司 Antaris 傅里叶变换近红外光谱仪对大豆油脂样本进行透射光谱扫描, 波长范

围 12000~4000cm<sup>-1</sup>, 分辨率 8cm<sup>-1</sup>。样品池采用直径为 8mm 透明玻璃管, 以空气作参比, 扫描次数 64 次。为减小测量误差, 每个样本测量 3 次, 取平均光谱, 其谱图如图 1 所示。

### 1.3 数据处理

小波变换、BP 神经网络在 Matlab 7.0 中编程实现, 偏最小二乘回归算法采用 CAMO 的 Unscrambler 9.7 数据分析软件实现。

## 2 结果与分析

### 2.1 基于小波变换的光谱预处理

噪声永远伴随信号存在, 传统的去噪方法通常无法刻画信号的非平稳性、提高信号的信噪比。小波变换具有多分辨率性、选基灵活性、低熵性和去相关性等优点, 在正交小波中, 正交基的选择更能接近真实信号本身, 通过正交小波变换能更容易地分离出噪声或其他不被需要的信息, 因此, 在信号去噪处理中小波变换具有传统方法无可比拟的优势<sup>[7-9]</sup>。

小波阈值去噪是 1992 年斯坦福大学的 Johnstone 和 Donoho 教授提出来的<sup>[10]</sup>, 它是一种非线性去噪法, 区别于以往的小波去噪, 在最小均方误差意义下能近似达到最优, 是一种实用简单、效果较好的去噪方法。它的基本思想是根据正交小波变换的去数据相关性, 将信号的能量集中在小波域中一些大而有限的系数中、将噪声能量分布在小波域的整个区间, 先对信号进行小波分解, 然后选择合适的门限阈值, 将各层分解系数的模值与阈值比较, 并将大于和小于该阈值的系数进行分别处理, 再对处理后的小波系数进行重构, 即可完成信号去噪。

本实验选取 db4 小波基对油脂近红外光谱信号预处理。dbN 系列小波具有很好的正交性, 虽无解析的表达式, 但其扩展性和光滑性好, 可灵活权衡增加支集长度带来的边界问题, 提高信号能量的集中程度, 对信号去噪有非常好的效果<sup>[11]</sup>。

任意信号的小波分解都是根据信号长度确定最大分解层数的, 在 MATLAB 的小波工具箱中, 通过语句  $N = \text{wmaxlev}(n, 'wname')$  来实现<sup>[12]</sup>。研究中选取 4500~6000cm<sup>-1</sup> 为特征波段建立油脂酸价校正模型, 采集的光谱信号波长点数为 188, 计算得到最大分解层数是 4, 也就是说分解层数可以是 2、3 或 4。为了达到最佳去噪效果, 本实验分别进行不同层数的小波分解, 选用缺省阈值处理方式, 对原始光谱和小波去噪后重构的光谱建立 PLS 回归模型, 利用 9 个预测集样本验证模型, 以最佳主因子数、决定系数  $R^2$  和预测误差均方根 (root-mean-square error of prediction, RMSEP) 的大小评价模型质量, 预测结果如表 1 所示。

表1 不同小波分解层数预测结果

Table 1 Prediction results with different levels of wavelet decomposition

预测集	化学值	原始光谱	2层分解	3层分解	4层分解
样本1	0.665	0.5739	0.4985	0.44432	0.46925
样本2	0.913	0.92904	0.94978	0.96748	0.94575
样本3	1.181	1.18387	1.26481	1.25387	1.22679
样本4	1.398	1.62994	1.55339	1.51228	1.55508
样本5	1.629	1.85729	1.8552	1.81972	1.81944
样本6	1.874	1.85173	1.83652	1.8366	1.86139
样本7	2.115	2.1609	2.16207	2.1579	2.17821
样本8	2.414	2.35719	2.45037	2.451	2.39161
样本	92.704	2.55441	2.60042	2.61291	2.61765
最佳主因子数		6	6	6	4
$R^2$		0.962	0.9662	0.9683	0.9695
RMSEP		0.1259	0.1188	0.1151	0.1128

由表1可以看出,经过不同分解层次的小波去噪后,模型质量较原始光谱均有所改善,其中4层分解效果最好,不仅提高了模型预测精度,而且最佳主因子数也从6变为4,减少了模型复杂度。

## 2.2 BP神经网络校正模型的建立

### 2.2.1 BP神经网络概述

人工神经网络(ANN)是近年发展起来的数学处理方法,在近红外光谱分析领域已被广泛用于建立非线性体系较强的近红外分析模型。在诸多算法中,目前应用最多的是BP人工神经网络<sup>[13]</sup>。

BP神经网络通常采用基于BP神经元的多层前向神经网络的结构形式,一般由3个神经元层次组成,即输入层、输出层和隐含层。数据由输入层输入,经标准化处理,并施以权重传输到第二层,即隐含层,隐含层经过权值、阈值和激励函数运算后,传输到输出层,输出层给出神经网络的预测值,并与期望值进行比较,若存在误差,则从输出开始反向传播该误差,进行权值、阈值调整,使网络输出值逐渐与希望输出值一致。

### 2.2.2 油脂酸价BP神经网络设计

根据Kolmogorov映射网络存在定理,三层BP神经网络(即一个隐含层)可以以任意精度拟合任意连续函数<sup>[14]</sup>,所以本实验选取三层结构的BP神经网络建立油脂酸价近红外光谱校正模型。

近红外光谱往往具有上千个数据点,若直接将这些数据点作为网络的输入,则网络规模会非常庞大。同时这些数据点可能存在严重的共线性,不仅使网络运算速度慢,更易陷入局部最小。因此,有必要对光谱数据进行降维处理。首先将40个校正集样本的近红外光谱数据进行db4四层小波分解去噪,进行标准化处理,再利用PLS交互验证选择最佳主因子数4,将主因子得分作为神经网络的输入变量,即输入层节点为4。输出层采用单节点即输出层的变量数为1,为油脂酸价化学测定值。

在BP网络的建立过程中,隐含层节点数的选择是关键问题。一般来说,较多的隐含层节点可存储较多的信息,但随着隐含层节点数的增加,权数呈平方增加,由此带来训练时间变长的问题,而且对于较多的隐含层节点而言,通常需要更多的训练样本,否则所得数学模型不稳定。较少的隐含层节点则存储信息较少,不能充分反映输入和输出变量之间复杂的函数关系,而且在训练过程中很容易陷入局部极小<sup>[15]</sup>。

通过建立不同隐含层节点的神经网络模型,利用预测集样本验证,根据预测误差均方根RMSEP的大小做评价指标,选择最佳隐含层节点数,隐含层节点数对网络的影响如图2所示。

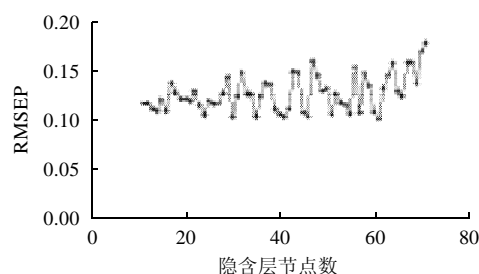


图2 隐含层节点数对BP神经网络的影响

Fig.2 Influence of hidden layer node number on BP neural network.

由图2可以看出,当隐含层节点数为60时, RMSEP最小。设定隐含层传递函数为logsig(对数S型传递函数),输出层传递函数为purelin(线性传递函数),网络训练迭代次数为3500次,目标误差设定为0.005,建立BP神经网络。最后对9个预测集样本进行验证,结果如图3所示。

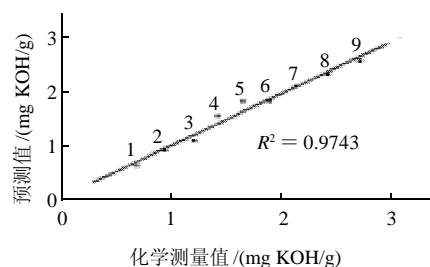


图3 油脂酸价化学测量值与预测值相关图

Fig.3 Correlation between chemically measured and predicted acid values

## 3 结 论

传统的油脂酸价检测都是利用化学方法在实验室进行的,费时费力耗资,不能实现实时有效的在线检测,常导致油脂质量不稳定以及油脂质量等级定标模糊等问

题。本课题以三级大豆油脂酸价检测为研究对象,利用现代近红外光谱分析技术,结合小波变换及BP神经网络建立校正模型,预测结果决定系数 $R^2$ 和预测均方根误差RMSEP分别为0.9743和0.1036。以上研究可为近红外光谱分析技术在食用油脂酸价定量检测中的应用提供理论依据,也为进一步实现油脂酸价的实时在线检测和调控提供参考。

#### 参考文献:

- [1] 薛雅琳, 马榕. 油脂质量标准体系与质量检验体系的建立[J]. 中国油脂, 2006, 31(1): 53-56.
- [2] 韩菊, 董建芳, 李芳. 油脂中脂肪酸的分析测定[J]. 河北科技大学学报, 2007, 28(3): 209-212.
- [3] 刘瑞兴, 吴苏喜. 深色油脂酸值测定新方法的研究[J]. 粮油加工与食品机械, 2006(5): 51-52.
- [4] 褚小立, 袁洪福, 陆婉珍. 近年来我国近红外光谱分析技术的研究与发展[J]. 分析仪器, 2006(2): 1-10.
- [5] 陆婉珍. 现代近红外光谱分析技术[M]. 北京: 中国石化出版社, 2007.
- [6] HOLMAN R T, EDMONDSON P R. Near-infrared spectra of fatty acids and some related substances[J]. Anal Chem, 1956, 28(10): 1533-1538.
- [7] PASTI L, WALCZAK B, MASSART D L. Optimization of signal denoising discrete wavelet transform[J]. Chemometrics and Intelligent Laboratory Systems, 1999, 48(1): 21-34.
- [8] KARIM G O, DAVID J A. Noise reduction in multichannel neural recordings using a new array wavelet denoising algorithm[J]. Neurocomputing, 2001, 38(4): 1687-1693.
- [9] 马兰, 夏俊芳, 张战锋, 等. 基于小波变换的番茄总糖近红外无损检测[J]. 农业工程学报, 2009, 25(10): 350-354.
- [10] DONOHO D L, JOHNSTONE I M. Wavelet shrinkage: asymptotia[J]. Journal of the Royal Statistical Society: Series (B), 1995, 7(2): 301-369.
- [11] 刘娟花, 李福德. 基于小波变换的信号去噪研究[J]. 西安理工大学学报, 2004, 20(3): 289-292.
- [12] 董长虹. Matlab 小波分析工具箱原理与应用[M]. 北京: 国防工业出版社, 2005.
- [13] 翁欣欣, 陆峰, 王传现, 等. 近红外光谱-BP神经网络-PLS法用于橄榄油掺杂分析[J]. 光谱学与光谱分析, 2009, 29(12): 3283-3287.
- [14] 杨南林, 程翼宇, 瞿海斌. 用人工神经网络近红外光谱法测定冬虫夏草中的甘露醇[J]. 分析化学, 2003, 31(6): 664-668.
- [15] 单杨, 朱向荣, 许青松, 等. 近红外光谱结合小波变换-径向基神经网络用于奶粉蛋白质与脂肪含量的测定[J]. 红外与毫米波学报, 2010, 29(2): 128-131.