

# 山茶油中油酸和亚油酸近红外光谱分析模型

郝勇\*, 吴文辉, 商庆园, 耿佩

华东交通大学机电与车辆工程学院, 江西 南昌 330013

**摘要** 将近红外光谱分析技术结合化学计量学方法用于山茶油混合油品中油酸和亚油酸含量的快速检测。配制了 76 种山茶油混合油样本用于近红外光谱的采集, 将不同的光谱预处理方法用于光谱有效信息的提取; 将蒙特卡罗无信息变量消除(MCUVE)和变量组合集群分析(VCPA)方法用于建模变量的选择; 将偏最小二乘回归(PLSR)用于脂肪酸含量定量分析模型的构建。结果表明: 经 NWD1<sup>st</sup>-MSC 预处理后, 两种脂肪酸的近红外光谱的校正均得到最好的结果; 采用基于 VCPA 的变量优选方法极大地改善了模型精度, 实现了建模变量数量的有效压缩。对于油酸模型, 建模变量数量由 1501 减少为 7, 交叉验证均方根误差和校正相关系数分别为 1.107 和 0.984, 预测均方根误差和测试集的预测相关系数分别为 1.178 和 0.981; 对于亚油酸模型, 建模变量数量由 1501 减少为 8, 交叉验证均方根误差和校正相关系数分别为 0.089 和 0.987, 预测均方根误差和测试集的预测相关系数分别为 0.105 和 0.982。近红外光谱分析技术结合 NWD1<sup>st</sup>-MSC-VCPA-PLSR 的方法为山茶油混合油品中脂肪酸含量的测定提供了一种快速简单的分析方法。

**关键词** 光谱学; 近红外光谱; 脂肪酸; 变量筛选; 蒙特卡罗无信息变量消除; 变量组合集群分析

中图分类号 S816.2

文献标识码 A

doi: 10.3788/AOS201939.0930004

## Analysis Model of Oleic and Linoleic Acids in Camellia Oil via Near-Infrared Spectroscopy

Hao Yong\*, Wu Wenhui, Shang Qingyuan, Geng Pei

School of Mechatronics & Vehicle Engineering, East China Jiaotong University, Nanchang, Jiangxi 330013, China

**Abstract** Near-infrared spectroscopy (NIRS), combined with chemometrics methods, is applied to rapid quantitative determination of oleic acid and linolenic acid in camellia oil blends. 76 camellia oil samples are prepared and used for near-infrared spectral collection. Different spectral preprocessing methods are applied to effective information extraction. Two variable selection methods, Monte Carlo uninformative variable elimination (MCUVE) and variable combination population analysis (VCPA), are applied to select characteristic NIRS variables for the two fatty acids in camellia oil blends. Quantitative analysis models of the fatty acids are built using partial least-square regression. The results show that NWD1<sup>st</sup>-MSC preprocessing can be used for optimization of near-infrared spectra of the two fatty acids in camellia oil blends. It is found that the VCPA method can greatly improve the precision of the model and compress the modeling variables. For the oleic acid model, the modeling variables decrease from 1501 to 7, the root-mean-square error of cross-validation and correlation coefficient of calibration are 1.107 and 0.984, respectively, and the root-mean-square error and correlation coefficient of prediction are 1.178 and 0.981, respectively. For the linoleic acid model, the modeling variables decrease from 1501 to 8, the root-mean-square error of cross-validation and correlation coefficient of calibration are 0.089 and 0.987, respectively, and the root-mean-square error and correlation coefficient of prediction are 0.105 and 0.982, respectively. NIRS combined with NWD1<sup>st</sup>-MSC-VCPA-PLSR provides a quick and easy analysis method for measuring fatty acids in camellia oil blends.

**Key words** spectroscopy; near-infrared spectroscopy; fatty acid; variable selection; Monte Carlo uninformative variable elimination; variable combination population analysis

**OCIS codes** 300.1030; 300.6340; 120.4290

收稿日期: 2019-03-06; 修回日期: 2019-04-04; 录用日期: 2019-05-14

基金项目: 国家自然科学基金(21265006)、国家质量监督检验检疫总局科技计划项目(2015IK160)、江西省自然科学基金(2015ZBAB201003)

\* E-mail: haonm@163.com

# 1 引 言

山茶油是一种营养价值非常高且易被人体消化吸收的绿色食用油,含有丰富的油酸和亚油酸等脂肪酸。孕妇在怀孕期间食用山茶油既可以增加母乳,又有利于胎儿健康成长;老年人食用可以去火、养颜,并且可以降低心脑血管疾病的发病率<sup>[1-3]</sup>。

近红外光谱技术作为一种绿色、高效、无损的现代分析技术已被应用于众多领域,如:郑瑞娜等<sup>[4]</sup>利用近红外技术对海藻糖浆组分进行了检测;孔清清等<sup>[5]</sup>利用近红外技术对烟叶产地进行了识别;王丽杰等<sup>[6]</sup>利用近红外技术对牛奶中脂肪、蛋白质和乳糖的浓度进行了检测;Cascañt 等<sup>[7]</sup>利用近红外技术测定了鲑鱼油中脂肪酸的含量,测量结果的相对均方根预测误差(RRMSEP)不大于 1.8%;Li 等<sup>[8]</sup>利用近红外技术测定了沙棘籽油中脂肪酸的含量,其中油酸和亚油酸的相对误差分别为 1.20%和 1.02%;Yuan 等<sup>[9]</sup>利用近红外技术测定了山茶油中的成分,其中油酸和亚油酸的预测相关系数( $R^2$ )分别为 0.9424 和 0.9682。以上研究结果表明,近红外技术在脂肪酸定量分析中具有一定的可行性。采用近红外光谱分析技术对油料作物中的脂肪酸进行定

量分析时,模型精度易受油品组分差异(产地和制取方式不同)和建模变量的影响,模型的稳定性和泛化能力较差。基于此,通过收集不同来源和不同制取方式的植物油结合人工配制的方式构成分析样品集,采用近红外光谱分析技术结合不同的光谱预处理方法以及蒙特卡罗无信息变量消除(MCUVE)、变量组合集群分析(VCPA)的变量选择方法,构建山茶油混合油品中油酸和亚油酸含量的偏最小二乘回归(PLSR)模型,期望能为山茶油混合油品中脂肪酸含量的测定提供稳定可靠的参考方法,为后续便携式近红外脂肪酸含量快速测定仪光源发射波长和检测器探测范围的选择提供理论依据。

## 2 实验材料与方 法

### 2.1 样品的制备

山茶油(压榨和浸出两种制取方式)、棕榈油、菜籽油由江西省 A 市和 B 市两家茶油厂提供,其中山茶油为 100%的纯山茶油(经江西省出入境检验检疫局测定)。以山茶油为基础油,分别向其中加入不同体积分数的棕榈油和菜籽油<sup>[10]</sup>,样品的组成信息如表 1 所示。共配制 76 个样品。

表 1 样品配制信息

Table 1 Sample configuration information

Sample composition	Number	Volume fraction/%
Camellia oil mixed with rapeseed oil (pressed)	1-19	3, 5, 7, 10, 13, 15, 18, 20, 25, 28, 30, 32, 35, 38, 40, 43, 45, 48, 50
Camellia oil mixed with rapeseed oil (leaching)	20-38	3, 5, 7, 10, 13, 15, 18, 20, 25, 28, 30, 32, 35, 38, 40, 43, 45, 48, 50
Camellia oil mixed with palm oil (leaching)	39-57	3, 5, 7, 10, 13, 15, 18, 20, 25, 28, 30, 32, 35, 38, 40, 43, 45, 48, 50
Camellia oil mixed with palm oil (pressed)	58-76	3, 5, 7, 10, 13, 15, 18, 20, 25, 28, 30, 32, 35, 38, 40, 43, 45, 48, 50

### 2.2 脂肪酸含量的测定

安捷伦-6890 气相色谱仪用于样品中油酸和亚油酸含量的测定。气相色谱仪的测定条件如表 2 所示。

### 2.3 光谱采集及样品集划分

将步琦 NIRFlex N-500 近红外光谱仪用于样品光谱的采集,利用培养皿附件采集样品的近红外光谱,采集时保证油液厚度一致,波数范围为 4000~10000  $\text{cm}^{-1}$ ,分辨率为 8  $\text{cm}^{-1}$ ,扫描次数为 32,温度

为 10~30  $^{\circ}\text{C}$ ,相对湿度为 30%~70%。

采用浓度均匀分布的方式将样品集划分为建模集和测试集,保证建模集样品的浓度范围涵盖整个测试集样品的浓度范围,并且要求样品均匀分布。为了保证小样本模型的稳健性,一般要求建模集样品和测试集样品数量之比为 2:1。因此,将 76 个样品进行划分,其中的 51 个样品用于模型的建立,剩下的 25 个样品用于模型的测试。

表 2 气相色谱检测条件

Table 2 Detection conditions of gas chromatograph

Content	Parameter
Instrument model	Agilent-6890
Detector	FID
Sampler	Split
Flow rate of H <sub>2</sub> / (mL·min <sup>-1</sup> )	1
Initial temperature / °C	140
Import sample temperature / °C	260
Termination temperature / °C	240
Heating rate / (°C·min <sup>-1</sup> )	4
Carrier gas	N <sub>2</sub>
Split ratio	40:1
Sample size / μL	1
Measure time / min	40

## 2.4 光谱数据的预处理方法

Norris-Williams (NW) 求导是由 Norris 在 1983 提出的一种用于近红外光谱求导的算法,该算法是消除光谱基线漂移的常用方法。NW 求导主要包括两个步骤,即首先对光谱进行平滑处理,然后根据给定的参数(窗口宽度和拟合次数)对平滑后的光谱求一阶导数或二阶导数<sup>[11]</sup>。

标准正态变量变换(SNV)主要用来消除表面散射和光程变化对近红外光谱的影响<sup>[12]</sup>,对于光谱  $\mathbf{X}_{i,k}$  进行 SNV 变换的计算公式为

$$\mathbf{X}_{i,SNV} = (\mathbf{X}_{i,k} - \bar{X}_i \mathbf{I}) / \sqrt{\frac{\sum_{k=1}^m (\mathbf{X}_{i,k} - \bar{X}_i \mathbf{I})^2}{m-1}}, \quad (1)$$

式中: $\mathbf{X}_{i,SNV}$  为 SNV 预处理后的光谱; $\bar{X}_i$  为第  $i$  条光谱的平均值(标量); $\mathbf{X}_{i,k}$  为需要进行 SNV 变换的光谱; $k=1,2,\dots,m$ ,  $m$  为波长点数; $i=1,2,\dots,n$ ,  $n$  为校正集的样本数; $\mathbf{I}$  为  $1 \times m$  的单位向量。

多元散射校正(MSC)可以有效剔除一些物理因素(相对湿度、样品不均匀、装填密度等)导致的散射的影响,提高光谱的信噪比<sup>[12]</sup>,具体算法如下。

- 1) 计算校正集样品的平均光谱  $\bar{\mathbf{A}}$ ;
- 2) 进行一元线性回归,即

$$\mathbf{A}_i = a_i \bar{\mathbf{A}} + \mathbf{I} b_i, \quad (2)$$

- 3) 计算 MSC 预处理后的光谱,即

$$\mathbf{A}_{i,MSC} = (\mathbf{A}_i - \mathbf{I} b_i) / a_i, \quad (3)$$

式中: $\mathbf{A}_i$  为第  $i$  条光谱; $a_i$ 、 $b_i$  分别为相对偏移系数

和偏移量; $\mathbf{A}_{i,MSC}$  为经 MSC 预处理后的光谱。

## 2.5 变量优选方法

### 2.5.1 VCPA

VCPA 是基于模型集群分析思想的一种新颖的变量选择方法,包括两个关键步骤。首先,利用达尔文自然进化理论中“适者生存”的简单有效的指数递减函数(EDF)来确定变量空间保持和连续收缩的变量数;其次,在每个 EDF 运行中,采用二进制矩阵采样(BMS)策略,给每个变量相同的机会去选择,并生成不同的变量组合,用于产生子集,以构造子模型的种群,第  $\lambda$  次 EDF 运行时计算要保留的变化比率  $r_\lambda$ ;然后计算变量之间所有组合的交叉验证均方根误差(RMSECV);最后将 RMSECV 值由小到大排列,选择前 10%子集作为最后运算结果<sup>[13]</sup>。变化比率的表达式为

$$r_\lambda = \exp(-\theta \lambda), \quad (4)$$

其中,

$$\theta = \frac{\ln(p/14)}{N}, \quad (5)$$

式中: $\theta$  为常数,用来控制 EDF 的曲线; $p$  为变量的总数; $N$  为 EDF 预定的运行步数。

### 2.5.2 MCUVE

UVE 是基于 PLSR 系数向量  $\mathbf{s}$  的变量选择方法,用来消除那些不提供信息的变量。基于此,MCUVE 在 UVE 方法的可变空间中采用蒙特卡罗进行采样。回归系数矩阵在第  $\psi$  次运行后计算得到的回归系数矩阵  $\mathbf{S}_o = [s_1, \dots, s_\omega]$ ,  $o=1,2,\dots,\psi$ ,稳定性的定义为<sup>[14]</sup>

$$h_j = \text{mean}(s_j) / \text{std}(s_j), j = 1, 2, \dots, \omega, \quad (6)$$

式中: $\text{mean}(s_j)$  和  $\text{std}(s_j)$  为每个变量绝对回归系数的均值和标准差; $s_j$  为回归系数矩阵  $\mathbf{S}$  的第  $j$  列向量; $\omega$  为回归系数矩阵列向量的总数。

## 2.6 模型的建立和评价

经过 VCPA 和 MCUVE 变量优选后,采用 PLSR 方法建立油酸和亚油酸含量的定量分析模型。采用模型的校正相关系数( $R_c$ )和 RMSECV 作为评价指标对光谱的预处理方法和变量选择方法进行优选,采用测试集的预测相关系数( $R_p$ )和预测均方根误差(RMSEP)用于最终优化模型精度的评价。MATLAB R2014a 用于光谱数据的预处理、变量优选以及偏最小二乘模型的建立。其中  $R_c$  越大, RMSECV 越小,模型的精度越高; $R_p$  越大, RMSEP 越小,模型的预测能力越好。

### 3 结果与分析

#### 3.1 光谱预处理方法的选择

图 1 所示为样本的原始近红外光谱图。从图 1 可知,样本的原始光谱图中不仅包含其组分信息,还包括光谱仪组成器件引入的噪声、杂散光和样品背景等干扰信息引起的光谱反射率的变化。因此,需要对原始光谱进行预处理,对干扰信息进行抑制或滤除,以提高后续模型的分析精度。

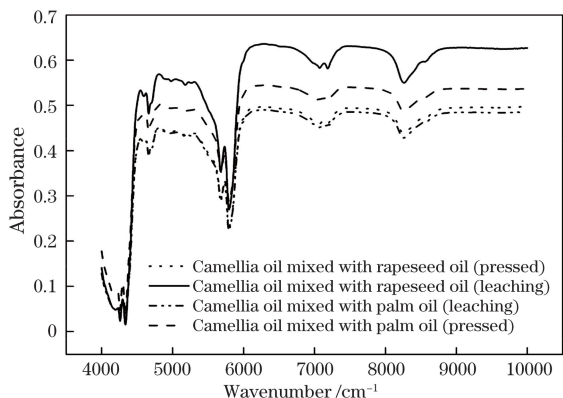


图 1 样本的原始近红外光谱图

Fig. 1 Original near-infrared spectra of samples

分别采用  $NWD1^{st}$ 、 $NWD2^{nd}$ 、MSC、SNV、 $NWD1^{st}$ -MSC 这 5 种预处理方法对原始光谱进行信息变换和提取,结果如表 3 所示。由表 3 可知,光谱预处理后,山茶油混合油品中两种脂肪酸的 PLSR 模型结果相差较大,其中光谱经  $NWD1^{st}$ -MSC 方法预处理后,两种脂肪酸的  $R_c$  和 RMSECV 都得到了明显改善。图 2 所示为  $NWD1^{st}$ -MSC 处理后的光谱图,可见,与原始光谱相比,预处理后的光谱变得更加集中和光滑。

表 3 不同光谱预处理方法的 PLSR 模型结果比较

Table 3 Comparison of PLSR modeling results with different spectral preprocessing methods

Method	Oleic acid		Linoleic acid	
	$R_c$	RMSECV	$R_c$	RMSECV
None	0.726	5.197	0.354	0.657
$NWD1^{st}$	0.733	5.105	0.669	0.622
$NWD2^{nd}$	0.701	5.444	0.546	0.740
MSC	0.985	1.086	0.950	0.171
SNV	0.985	1.096	0.949	0.173
$NWD1^{st}$ -MSC	<b>0.986</b>	<b>1.021</b>	<b>0.968</b>	<b>0.138</b>

#### 3.2 PLSR 模型建模变量的优选

随着仪器科学的发展,光谱图中包含的信息越来越丰富,冗余信息的存在不仅会增加计算开销,还

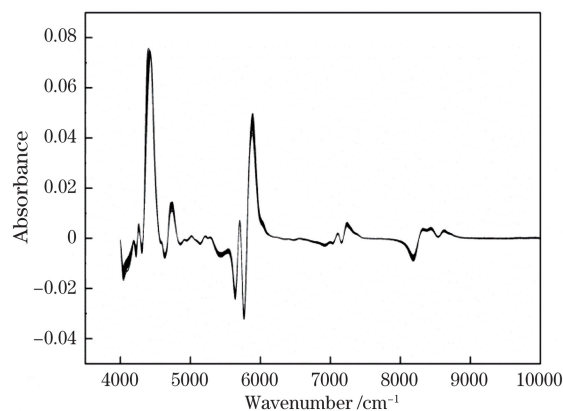


图 2  $NWD1^{st}$ -MSC 处理后的光谱图

Fig. 2 Spectrum after processing with  $NWD1^{st}$ -MSC

会影响待测对象的分析精度。变量优选能够剔除冗余信息,降低模型的复杂度,提高程序的运行速度。分别采用 MCUVE 和 VCPA 方法对光谱的建模变量进行优选。光谱经  $NWD1^{st}$ -MSC 预处理后的两种变量优选方法的 PLSR 模型分析结果如表 4 所示。由表 4 可知,对于油酸模型,光谱经  $NWD1^{st}$ -MSC 预处理,并经 MCUVE 优选变量后,PLSR 模型的校正结果得到进一步提高,而采用 VCPA 优选变量后,模型的精度稍有降低,但模型得到了极大精简,建模变量由 1501 减小为 7;对于亚油酸模型,光谱经  $NWD1^{st}$ -MSC 预处理,并经 MCUVE 和 VCPA 优选变量后,精度均得到提高,建模变量均减少,分别由 1501 减小为 500 和 8。对于山茶油混合油品中两种脂肪酸模型,采用 VCPA 方法在不影响定量分析精度的前提下得到了最精简的模型。采用 VCPA 方法选择的变量分布如图 3 所示,可知,油酸模型选择的变量主要集中在  $4200\text{ cm}^{-1}$  附近和  $5700\text{ cm}^{-1}$  附近,亚油酸模型选择的变量主要在  $4200\text{ cm}^{-1}$  附近和  $4400\sim 4900\text{ cm}^{-1}$  区间。根据文献[15-16]可知,在  $4200\text{ cm}^{-1}$  波数附近选择的变量

表 4 两种变量优选方法的 PLSR 模型结果比较

Table 4 Comparison of PLSR modeling results with two kinds of variable selection methods

Analysis object	Variable selection method	Number of variables	$R_c$	RMSECV
Oleic acid	None	1501	0.986	1.021
	VCPA	7	0.984	1.107
	MCUVE	180	0.994	0.688
Linoleic acid	None	1501	0.968	0.138
	VCPA	8	0.987	0.089
	MCUVE	500	0.989	0.080



对应于 $-\text{CH}_2$ 和 $-\text{CH}_3$ 基团的弯曲振动和 $-\text{CH}$ 的拉伸,在 $4400\sim 4900\text{ cm}^{-1}$ 波数范围内选择的变量对应于 $\text{C}=\text{O}$ 伸缩振动合频吸收、顺式双键的 $-\text{C}-\text{H}$ 伸缩振动组合频,在 $5700\text{ cm}^{-1}$ 波数附近选

择的变量是 $-\text{CH}_2$ 中 $\text{C}-\text{H}$ 伸缩振动的第一级泛音。脂肪酸中的分子结构由 $\text{C}-\text{H}$ 、 $\text{C}-\text{O}$ 、 $-\text{CH}_3$ 、 $-\text{CH}_2$ 、 $-\text{HC}=\text{CH}$ 等键组成,表明建模选择的点符合样本中各物质的分子结构。

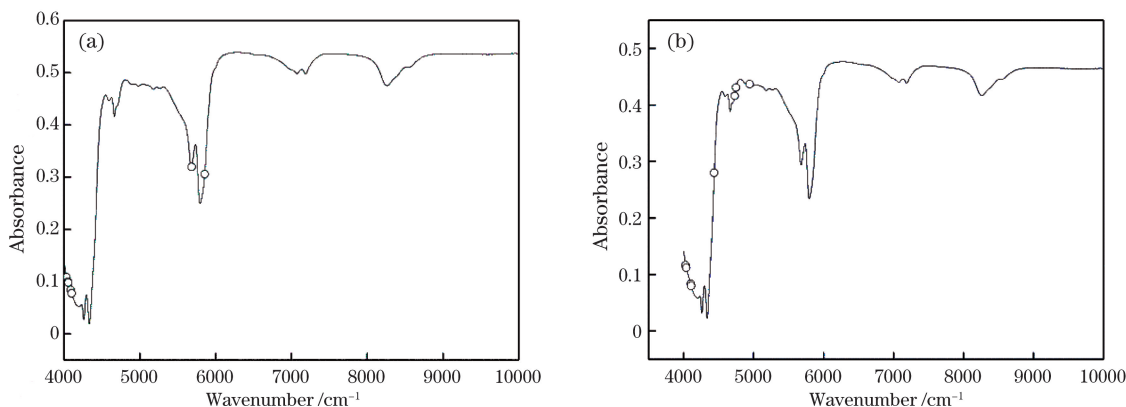


图3 两种脂肪酸PLSR模型建模变量样本点分布图。(a)油酸PLSR模型;(b)亚油酸PLSR模型

Fig. 3 Sample point distributions of PLSR modelling variables for two kinds of fatty acids.

(a) Oleic acid PLSR model; (b) linoleic acid PLSR model

### 3.3 PLSR模型的建立

采用PLSR建模方法从自变量和因变量矩阵中可以提取偏最小二乘的成分,可有效地实现降维,消除可能存在的自变量间的复共线性关系,从而显著提高数据结果的可靠性和准确性<sup>[17]</sup>。通过对多种预处理方法和变量优选方法进行对比分析可知,对于亚油酸和油酸两种脂肪酸光谱,采用 $\text{NWD1}^{\text{st}}\text{-MSC-VCPA}$ 方法处理后,其PLSR分析模型的精度

和复杂度均最优。分别对测试集样本中的两种脂肪酸含量进行预测分析,山茶油混合油品中两种脂肪酸模型的预测含量与实测含量间的相关性如图4所示。由图4可知:油酸和亚油酸质量分数的参考值和预测值之间存在较好的相关关系;油酸预测样品集的 $R_p$ 和RMSEP分别为0.980和1.178,标准误差(SD)为2.565;亚油酸预测样品集的 $R_p$ 和RMSEP分别为0.981和0.105,标准误差为0.447。

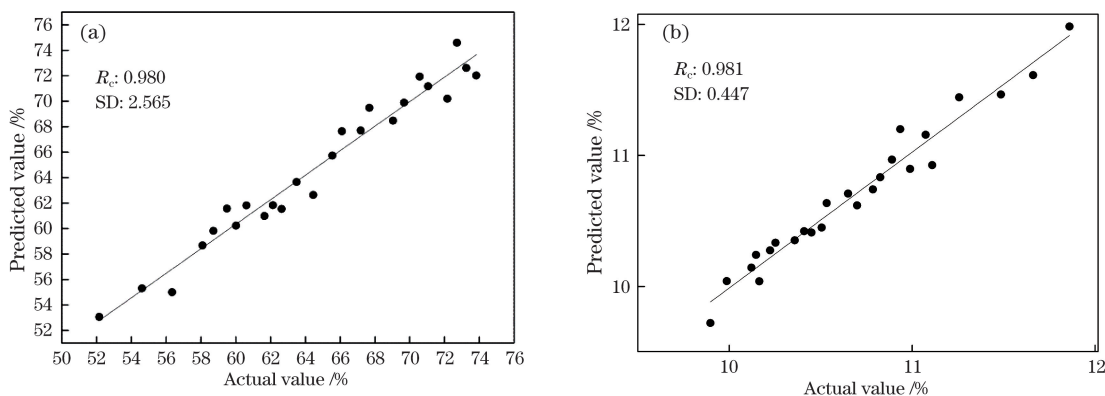


图4 两种脂肪酸含量实测值与模型预测值的相关关系图。(a)油酸;(b)亚油酸

Fig. 4 Correlation relationships of measured and predicted values of models for two kinds of fatty acids.

(a) Oleic acid samples; (b) linoleic acid samples

## 4 结 论

将近红外光谱分析技术化学计量学方法用于山茶油混合油品中油酸和亚油酸含量的快速分析,将5种光谱预处理方法和两种变量优选的方法用于光谱信息的提取和模型优化。研究表明:近红外

光谱分析技术结合 $\text{NWD1}^{\text{st}}\text{-MSC-VCPA-PLSR}$ 方法可以用于山茶油混合油品中油酸和亚油酸的快速准确分析;油酸预测样品集的 $R_p$ 和RMSEP分别为0.980和1.178,亚油酸预测样品集的 $R_p$ 和RMSEP分别为0.981和0.105。该方法有望成为山茶油混合油品中脂肪酸含量快速分析的方法之一,同时,该

方法可为便携式近红外脂肪酸含量快速测定仪光源和检测器的选用提供选择依据。

## 参 考 文 献

- [1] Zeb A. Triacylglycerols composition, oxidation and oxidation compounds in camellia oil using liquid chromatography-mass spectrometry[J]. *Chemistry and Physics of Lipids*, 2012, 165(5): 608-614.
- [2] Li X, Tan Y S, Ma G G, *et al.* Advance in research and application of camellia oil [J]. *Journal of the Chinese Cereals and Oils Association*, 2017, 32(11): 191-196.  
李雪, 谭运寿, 马贵刚, 等. 山茶籽油研究应用进展 [J]. *中国粮油学报*, 2017, 32(11): 191-196.
- [3] Sun T, Wu Y Q, Li X Z, *et al.* Discrimination of camellia oil adulteration by NIR spectra and subwindow permutation analysis [J]. *Acta Optica Sinica*, 2015, 35(6): 0630005.  
孙通, 吴宜青, 李晓珍, 等. 基于近红外光谱和子窗口重排分析的山茶油掺假检测 [J]. *光学学报*, 2015, 35(6): 0630005.
- [4] Zheng R N, Xie D, Yang Q Y. Rapid quantitative detection of trehalose syrup component based on near infrared spectroscopy[J]. *Food & Machinery*, 2017, 33(10): 60-63, 134.  
郑瑞娜, 谢定, 杨倩圆. 基于近红外光谱的海藻糖浆组分快速定量检测 [J]. *食品与机械*, 2017, 33(10): 60-63, 134.
- [5] Kong Q Q, Ding X Q, Gong H L. Application of improved random forest pruning algorithm in tobacco origin identification of near infrared spectrum [J]. *Laser & Optoelectronics Progress*, 2018, 55(1): 013006.  
孔清清, 丁香乾, 宫会丽. 改进的修剪随机森林算法在烟叶近红外光谱产地识别中的应用研究 [J]. *激光与光电子学进展*, 2018, 55(1): 013006.
- [6] Wang L J, Yang Y Y. Purification and noise elimination of near infrared spectrum in rapid detection of milk components concentration by using principal component weight resetting[J]. *Acta Optica Sinica*, 2017, 37(10): 1030003.  
王丽杰, 杨羽翼. 利用主成分权重重置实现牛奶成分浓度快速检测中近红外光谱的净化去噪 [J]. *光学学报*, 2017, 37(10): 1030003.
- [7] Cascant M M, Breil C, Fabiano-Tixier A S, *et al.* Determination of fatty acids and lipid classes in salmon oil by near infrared spectroscopy [J]. *Food Chemistry*, 2018, 239: 865-871.
- [8] Li Z P, Wang J, Xiong Y T, *et al.* The determination of the fatty acid content of sea buckthorn seed oil using near infrared spectroscopy and variable selection methods for multivariate calibration [J]. *Vibrational Spectroscopy*, 2016, 84: 24-29.
- [9] Yuan J J, Wang C Z, Chen H X, *et al.* Prediction of fatty acid composition in camellia oleifera oil by near infrared transmittance spectroscopy (NITS) [J]. *Food Chemistry*, 2013, 138(2/3): 1657-1662.
- [10] Wang L, Lee F S C, Wang X R, *et al.* Feasibility study of quantifying and discriminating soybean oil adulteration in camellia oils by attenuated total reflectance MIR and fiber optic diffuse reflectance NIR [J]. *Food Chemistry*, 2006, 95(3): 529-536.
- [11] Asachi M, Hassanpour A, Ghadiri M, *et al.* Assessment of near-infrared (NIR) spectroscopy for segregation measurement of low content level ingredients [J]. *Powder Technology*, 2017, 320: 143-154.
- [12] Chu X L, Yuan H F, Lu W Z. Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique [J]. *Progress in Chemistry*, 2004, 16(4): 528-542.  
褚小立, 袁洪福, 陆婉珍. 近红外分析中光谱预处理及波长选择方法进展与应用 [J]. *化学进展*, 2004, 16(4): 528-542.
- [13] Yun Y H, Wang W T, Deng B C, *et al.* Using variable combination population analysis for variable selection in multivariate calibration [J]. *Analytica Chimica Acta*, 2015, 862: 14-23.
- [14] Yan H, Song X Z, Tian K D, *et al.* Quantitative determination of additive chlorantraniliprole in abamectin preparation: investigation of bootstrapping soft shrinkage approach by mid-infrared spectroscopy [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2018, 191: 296-302.
- [15] Grabska J, Beć K B, Ishigaki M, *et al.* Spectra-structure correlations of saturated and unsaturated medium-chain fatty acids. Near-infrared and anharmonic DFT study of hexanoic acid and sorbic acid [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2017, 185: 35-44.
- [16] Armenta S, Garrigues S, de la Guardia M. Determination of edible oil parameters by near infrared spectrometry [J]. *Analytica Chimica Acta*, 2007, 596(2): 330-337.
- [17] Hao Y, Xue L. Study on modeling methods for determination of the linolenic acid by using infrared spectroscopy [J]. *Journal of Chinese Agricultural Mechanization*, 2015, 36(3): 164-168.  
郝勇, 薛龙. 亚麻酸红外光谱定量分析模型构建方法研究 [J]. *中国农机化学报*, 2015, 36(3): 164-168.