

食用油脂酸值近红外光谱特征波长优选

王立琦¹, 刘亚楠¹, 张青², 崔月¹, 葛慧芳¹, 于殿宇^{2,*}

(1.哈尔滨商业大学计算机与信息工程学院, 黑龙江 哈尔滨 150028; 2.东北农业大学食品学院, 黑龙江 哈尔滨 150030)

摘要:以近红外光谱快速检测大豆油脂酸值为目标, 研究间隔偏最小二乘(interval partial least square, iPLS)结合遗传算法(genetic algorithm, GA)及连续投影算法(successive projection algorithm, SPA)的特征波长变量优选方法。制备不同酸值的大豆油脂样品100个, 并在4 000~12 000 cm⁻¹范围内采集了油样的近红外透射光谱。首先用iPLS法从原始光谱中初步筛选出4 540~5 346 cm⁻¹和6 807~7 004 cm⁻¹组合特征波段, R^2 和预测均方根误差(root mean square error of prediction, RMSEP)分别为0.978 9和0.064 3; 然后分别用GA和SPA从特征光谱区域中筛选出与油脂酸值密切相关的特征波长变量, 从GA和SPA 2种选择结果中各选取前6个波长点, 以12个特征波长变量建立PLS校正模型, 其 R^2 和RMSEP分别为0.985 9和0.045 1。研究表明, 在油脂酸值近红外光谱分析中, 采用iPLS-GA-SPA相结合的方法进行特征波长选择能有效去除冗余信息, 降低模型复杂度, 可为快速无损检测油脂酸值提供重要理论依据。

关键词:油脂酸值; 近红外光谱; 特征波长变量; 间隔偏最小二乘; 遗传算法; 连续投影算法

Optimization of Characteristic Wavelength Variables of Near Infrared Spectroscopy for Detecting Edible Oil Acid Value

WANG Liqi¹, LIU Yanan¹, ZHANG Qing², CUI Yue¹, GE Huifang¹, YU Dianyu^{2,*}

(1. School of Computer and Information Engineering, Harbin University of Commerce, Harbin 150028, China;

2. School of Food Science and Technology, Northeast Agricultural University, Harbin 150030, China)

Abstract: With the goal of achieving rapid detection of soybean oil acid value by using near-infrared (NIR) spectroscopy, this study optimized the selection of the characteristic wavelength variables by combined use of interval partial least square (iPLS), genetic algorithm (GA) and successive projection algorithm (SPA). A total of 100 soybean oil samples with different acid values were collected, their NIR transmittance spectra in the range of 4 000–12 000 cm⁻¹ were acquired. Firstly, the characteristic wavebands of 4 540–5 346 cm⁻¹ and 6 807–7 004 cm⁻¹ were extracted from the original spectra by iPLS, with a determination coefficient (R^2) and root mean square error of prediction (RMSEP) of 0.978 9 and 0.064 3, respectively. Then, the characteristic wavelength variables closely related to oil acid value were selected by GA and SPA from the previously selected bands, respectively. It was shown that the PLS calibration model established using 12 variables consisting of the top 6 characteristic wavelengths from optimal selection results of each of the two algorithms was optimum, with R^2 and RMSEP of 0.985 9 and 0.045 1, respectively. The research indicated that selection of the characteristic wavelength variables by iPLS-GA-SPA in NIR analysis for oil acid value could effectively remove redundant information, and decrease the complexity of the model. This paper can offer important reference for rapid and non-destructive detection of oil acid value.

Key words: oil acid value; near infrared spectroscopy (NIR); characteristic wavelength variables; interval partial least square (iPLS); genetic algorithm (GA); successive projection algorithm (SPA)

DOI:10.7506/spkx1002-6630-201616033

中图分类号: TS227

文献标志码: A

文章编号: 1002-6630 (2016) 16-0205-06

引文格式:

王立琦, 刘亚楠, 张青, 等. 食用油脂酸值近红外光谱特征波长优选[J]. 食品科学, 2016, 37(16): 205-210. DOI:10.7506/spkx1002-6630-201616033. <http://www.spkx.net.cn>

WANG Liqi, LIU Yanan, ZHANG Qing, et al. Optimization of characteristic wavelength variables of near infrared spectroscopy for detecting edible oil acid value[J]. Food Science, 2016, 37(16): 205-210. (in Chinese with English abstract)

DOI:10.7506/spkx1002-6630-201616033. <http://www.spkx.net.cn>

收稿日期: 2016-03-16

基金项目: 国家自然科学基金面上项目 (31271886)

作者简介: 王立琦 (1966—), 女, 教授, 博士, 主要从事食品安全快速检测研究。E-mail: hsdwlq@163.com

*通信作者: 于殿宇 (1964—), 男, 教授, 博士, 主要从事粮油加工技术研究。E-mail: dyu2000@126.com

酸值是油脂中游离脂肪酸含量的标志,是评价食用油脂质量的重要指标之一,也是油脂碱炼脱酸时计算用碱量的主要依据,因此在油脂加工过程中需要定时检测油脂的酸值^[1]。传统的检测方法是通过扦样,然后在实验室进行化学分析后反馈给生产一线,检测过程受人为因素影响大,使用化学滴定造成环境污染,检测时间长,不利于生产和调控^[2]。

近红外光谱分析技术具有检测速度快、绿色环保、无损等优点,近年来被广泛应用于食品检测、品质分析和掺伪鉴别^[3-5]。Aderval等^[6]通过多元分类与近红外光谱技术结合的方法准确鉴别出大豆油样是否为转基因产品;Armata等^[7]用近红外光谱技术对不同种类和地区的葵花籽油、橄榄油及玉米油的过氧化值等指标参数进行测定,并比较了不同预处理方法对模型回归预测的影响;Siemens等^[8]将中红外与近红外光谱相结合,快速测定出了菜籽油、亚麻油中的脂肪酸组成,并对油酸、亚油酸、亚麻酸的含量及碘值的测量方法进行了评价。

随着近红外光谱仪器的发展,人们可以采集到包含上千个波长点的光谱数据,这些光谱信息之间存在多重相关性,全谱波长信噪比低,模型计算量大,计算过程繁琐,导致预测精度降低,而且不同基团的近红外吸收特性不同,某些波长点处没有目标组分的吸收,波长冗余现象严重^[9-11]。考虑到近红外光谱技术应用时,必须保证所分析样品的光谱维数与建模中所采用的光谱数据维数一致,随着波长变量的增加,不仅会提高模型传递的难度,而且会增加专用分析仪器的硬件成本。通过特定方法从得到的谱图中筛选出对建模影响较大的特征波长变量,在简化模型的同时可以提高模型的稳健性,有利于性价比高的专用近红外光谱仪器的开发,对近红外光谱分析技术的应用和推广具有很高的实用价值。张曼等^[12]以66个小麦样品为实验材料,研究了岭回归方法在近红外光谱定量分析中的应用,最后从1 297个波长点中优选出4个波长点并建立预测模型,取得了很好的效果;Wu Di等^[13]应用近红外光谱检测食用植物油中的 α -亚麻酸含量和亚油酸含量,并进一步用小波变换结合无信息变量消除的波长优选方法对偏最小二乘(partial least square, PLS)回归模型进行简化;Alessandra等^[14]研究了4种不同种类的食用植物油中酸值、折射率和黏度在3 068~6 184 cm^{-1} 光谱区域内的吸收特性,通过间隔PLS(interval PLS, iPLS)方法分别筛选出这3个参数在该光谱区域中的特征吸收范围,实验结果显示酸值在中红外区域的吸收波段为3 108~3 356 cm^{-1} ,折射率和黏度在近红外区域的吸收波段为5 636~5 888 cm^{-1} ,然后结合连续投影算法(successive projection algorithm, SPA)在特征波段

中提取各参数的特征波长,达到简化多元线性回归模型(multiple linear regression, MLR)的效果。

本实验研究油脂酸值在近红外光谱区域的特征吸收,通过iPLS、遗传算法(genetic algorithm, GA)、SPA等多种波长变量选择方法相结合,优选出油脂酸值在近红外光谱区域的特征吸收波长变量,为快速无损检测油脂酸值提供重要理论依据。

1 材料与方法

1.1 材料

一级大豆油 市购。

1.2 仪器与设备

Antaris傅里叶变换近红外光谱仪 美国Thermo Nicolet公司;MATLAB R2010b(MathWorks)、Unscrambler X10.3 挪威CAMO公司。

1.3 方法

1.3.1 样品处理

利用合格的一级大豆油为基料,通过添加一定量的亚油酸以及等物质的量的甘油二酯来调节酸值,调制出不同酸值的大豆油脂样品100个,并按GB/T 5530—2005《动植物油脂:酸值和酸度测定》测得准确的酸值化学值。

1.3.2 近红外光谱采集

利用傅里叶变换近红外光谱仪对大豆油脂样品进行透射光谱扫描,扫描范围为4 000~12 000 cm^{-1} ,分辨率为4 cm^{-1} ,样品池采用直径为8 mm透明玻璃管,以空气作参比,扫描次数64次,为减小光谱数据的测量误差,每个样品测量3次,再以3次的平均值作为最终的光谱值。

1.3.3 特征波长变量的优选

根据大豆油脂样品对近红外光谱的吸收特征,首先采用多种PLS方法,从原始光谱中提取油脂酸值特征波段,再分别利用GA和SPA进行特征波长点的筛选,最后将GA和SPA的选择结果结合以确定最终入选的大豆油脂酸值的特征波长变量。

2 结果与分析

2.1 大豆油脂样品的分集

采用基于马氏距离的Kennard-Stone样品划分算法,从100个油脂样品中选取80个作为校正集用于建模,剩余20个样品作为预测集用于验证模型的适用性,如表1所示,预测集样品的化学值范围很好地包含于校正集样品中,样品分布比较均匀,校正集与预测集偏差较小。酸值的大小代表油脂中游离脂肪酸的百分含量(以油酸计)。

表1 样品酸值统计

Table 1 Statistical data of acid values of samples

样品	数目	最小值/%	最大值/%	平均值/%	标准偏差/%
校正集	80	0.110	1.612	0.930	0.386
预测集	20	0.273	1.482	0.945	0.382

2.2 油脂酸值特征波段选择

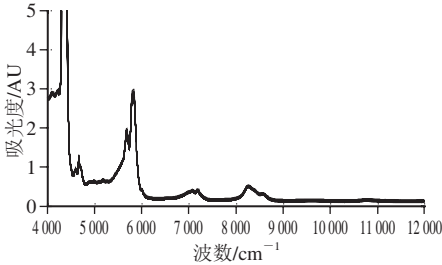
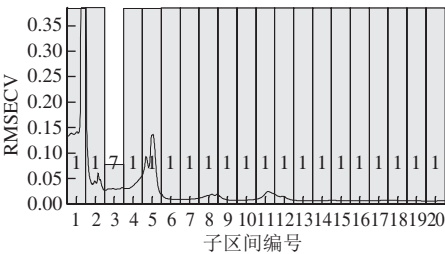


图1 大豆油脂样品近红外光谱
Fig. 1 NIR spectra of soybean oil samples

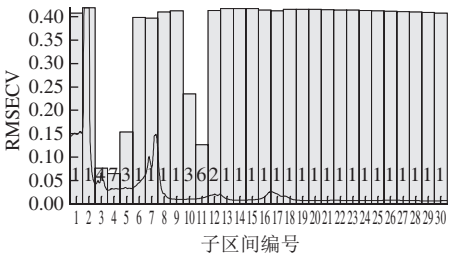
扫描不同酸值的100 个大豆油脂样品, 获得其近红外光谱如图1所示。本实验主要目的是筛选出油脂酸值特征波长, 过多的预处理有时会过滤掉光谱中的有用信息, 所以只采用平滑方法对原始光谱进行了去噪, 平滑窗口大小为11 点。根据已有的研究^[15]成果, 发现在油脂酸值近红外光谱分析中采用经典的PLS模型可以达到理想的预测效果。因此, 为了提高选择结果的准确性, 分别采用iPLS、组合区间PLS (synergy interval PLS, SiPLS) 和后向区间PLS (backward interval PLS, BiPLS) 进行对比分析^[16], 以期优选出与大豆油脂酸值密切相关的近红外光谱特征波段。

2.2.1 基于iPLS的特征波段选择

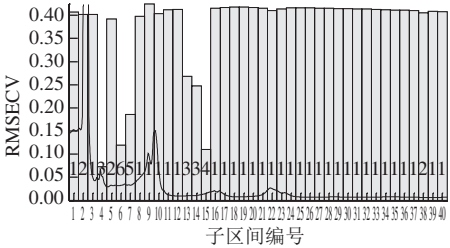
在4 000~12 000 cm⁻¹范围内, 以4 cm⁻¹分辨率共采集了2 075 个波长点数据, 波长变量编号为1~2 075。将原始光谱分别平均划分为20、30、40、50 个子区间, 应用iPLS在各子区间分别建立模型, 如图2所示。横坐标代表子区间的编号, 纵坐标指示出不同子区间的交互验证均方根误差 (root mean square error of cross validation, RMSECV) 值, 每一个子区间对应模型的最佳主因子数在柱状图底部给出。



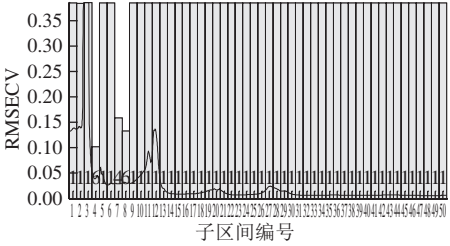
a. 20 个子区间建模结果



b. 30 个子区间建模结果



c. 40 个子区间建模结果



d. 50 个子区间建模结果

图2 不同子区间数的iPLS建模结果

Fig. 2 iPLS modeling results for different subintervals

从图2可以看出, 在整个波段选择过程中, 划分的子区间数大小对于模型的建立产生较大影响。图2a中的20 个子区间模型, 由于区间数过少, 易遗漏细节信息; 图2d中的50 个子区间模型的RMSECV值明显较高, 对整体模型产生一定破坏性。而图2b、c中的各个子区间对应的RMSECV值相对较小, 因此选择30 个及40 个子区间模型, 再从图2b中选取模型效果较好的区间3、4、5及区间11, 分别对应序号为141~210、211~280、281~350及696~764的波长变量, 从图2c中选取模型效果较好的区间4、6及区间15, 分别对应序号为157~208、261~312及729~780的波长变量。综合图2b及图2c的筛选结果, 确定2 个特征波段, 波长变量序号分别为141~350和729~780。

2.2.2 基于SiPLS的特征波段选择

与iPLS一样, 分别将原始光谱平均分为20、30、40 个和50 个子区间, 然后再分别联合2、3 个和4 个子区间建立模型以考察不同子区间数量对模型和选择结果的影响, 结果见表2。

表 2 SiPLS方法选择结果

Table 2 Selection results by SiPLS

子区间数	组合区间	R^2	RMSECV	相对标准偏差/%
20	[3, 13]	0.944 4	0.090 2	9.70
	[3, 13, 16]	0.952 6	0.098 9	10.63
	[3, 6, 10, 12]	0.947 6	0.091 6	9.85
30	[3, 8]	0.957 3	0.079 4	8.54
	[3, 18, 23]	0.967 5	0.077 4	8.32
	[3, 11, 12, 30]	0.972 9	0.063 0	6.77
40	[4, 40]	0.970 0	0.066 5	7.15
	[4, 17, 40]	0.969 5	0.067 8	7.29
	[4, 17, 38, 40]	0.970 3	0.065 4	7.03
50	[4, 7]	0.968 3	0.070 1	7.54
	[4, 6, 21]	0.968 7	0.075 8	8.15

从表2可以看出,在子区间数为30,联合4个子区间时获得的SiPLS模型精度最高,其 R^2 和RMSECV分别为0.972 9和0.063 0,此时入选组合区间为[3, 11, 12, 30],分别对应序号为141~210、696~833及2 007~2 075的波长变量。

2.2.3 基于BiPLS的特征波段选择

将原始光谱区间分别划分为20、30、40个和50个子区间,在不同子区间数的情况下,BiPLS模型如表3所示。

表 3 BiPLS方法选择结果

Table 3 Selection results by BiPLS

子区间数	入选区间	R^2	RMSECV	相对标准偏差/%
20	[3]	0.925 3	0.081 2	8.73
30	[3, 10, 13, 15, 25, 26, 29]	0.971 2	0.065 6	7.05
40	[4, 6, 7, 17, 30, 33, 38, 39]	0.973 8	0.061 8	6.65
50	[4, 5, 10, 18, 34, 37, 45, 48]	0.970 1	0.071 8	7.72

由表3可以看出,当原始光谱被平均划分为40个子区间时,能获得最佳的BiPLS模型,最优建模区间组合为[4, 6, 7, 17, 30, 33, 38, 39],分别对应序号为157~208、261~364、833~884、1 509~1 560及1 923~2 024的波长变量,该模型的 R^2 和RMSECV分别为0.973 8和0.061 8。

2.2.4 预测模型验证

综合上述结果,分别以iPLS、SiPLS和BiPLS选择结果中的最佳光谱波段建立校正模型,并与全谱模型对比,对验证集的20个未知样品进行预测,如表4所示。

表 4 不同波段优选方法的建模结果

Table 4 Modeling results of different methods for wavebands selection

建模方法	波长变量范围	波长变量数	R^2	RMSEP	相对标准偏差/%
PLS	1~2 075	2 075	0.840 6	0.104 2	11.03
iPLS	141~350、729~780	262	0.978 9	0.064 3	6.80
SiPLS	141~210、696~833、2 007~2 075	277	0.973 1	0.072 8	7.703
BiPLS	157~208、261~364、833~884、1 509~1 560、1 923~2 024	414	0.975 7	0.082 3	8.71

综合分析表4可以看出,运用3种方法优选特征波段后,预测结果均优于全谱模型,其中iPLS方法预测

精度最高,且实际参与建模的波长变量数最少,共计262个波长变量,对应光谱范围为4 540~5 346 cm^{-1} 和6 807~7 004 cm^{-1} 。根据相关文献[17-18]报道,4 500~4 800 cm^{-1} 正是油脂—COOR伸缩振动组合频吸收谱带,而6 080~7 014 cm^{-1} 则是一 CH_2 —的吸收谱带,为了避免特征波长的漏选,将4 540~5 346 cm^{-1} 和6 807~7 004 cm^{-1} 组合波段作为油脂酸值近红外光谱的特征波段以进一步进行特征波长变量选择。

2.3 油脂酸值特征波长变量选择

2.3.1 基于GA的特征波长变量选择

在选择出油脂酸值特征吸收波段的基础上,采用GA进一步优选特征波长变量。算法在Riccardo Leardi编写的GA-PLS工具箱中实现。该工具箱首先通过GA随机初始化染色体,最优染色体被作为PLS回归模型的输入进行留一法交叉验证得到一个交互验证解释方差(cross-validated explained variance, C.V%)值,在每一次计算过程中得到的最高C.V%值所对应的染色体就是所寻找的最优变量子集。通过PLS回归模型来验证GA选择的变量,可以避免单独使用GA而引起的过拟合^[19]。GA的参数设置为最大繁殖代数100、种群大小30、交叉率50%、变异率1%,选择10次迭代结果,针对各个波长点进行平均选择次数计算,根据选择次数进行排序而选取出相应的波长点,用各波段选择出的特征波长变量建立PLS模型,并对未知样品进行预测分析,预测结果见表5。表5中所选波长变量序号是按该波长点被选频数由高到低排序的,可以看出从4 540~5 346、6 807~7 004 cm^{-1} 组合波段内选择出的28个波长变量建模效果最好, R^2 和预测均方根误差(root mean square error of prediction, RMSEP)分别为0.980 3和0.049 8,用少数的特征波长变量进一步优化了校正模型的预测能力。

表 5 GA选择的特征波长变量建模结果

Table 5 Modeling results from characteristic wavelength variables selected by GA

波长范围/ cm^{-1}	所选波长变量序号	所选波长变量数	R^2	RMSEP	相对标准偏差/%
4 540~5 346	167、142、147、181、150、165、179、185、190、192、196、162、193、149、246、302、307、154、175、187、200、222、274、143、184、265、306、312、327、152、149、186	32	0.975 1	0.056 9	6.02
	755、768、743、742、766、777、757、760、780、773	10	0.975 4	0.058 9	6.23
4 540~5 346、6 807~7 004	167、161、184、190、195、181、186、187、163、196、149、170、142、166、198、229、239、293、759、154、155、174、160、178、199、204、226、270	28	0.980 3	0.049 8	5.27

但是,作为一种随机优化方法,GA在遗传迭代中会有权值、阈值的随机调整,且选择、交叉、变异过程也都带有很强的随机性,使得每次迭代结果出现不同^[20-21]。

为了提高波长变量选择的可靠性和准确性,另选取一种确定性的波长选择方法SPA,并与GA方法进行对比分析。

2.3.2 基于SPA特征波长变量选择

SPA是一种前向循环的变量选择方法,它从一个波长开始,每次计算它在未选入波长上的投影,将投影向量最大的波长选入到波长组合中,直到循环选出指定波长数目,每一个新选入的波长与前一个波长的线性关系都是最小的^[22-23]。SPA的运行步骤与文献[24]报道相同。

应用SPA在2个特征波段及其组合波段中提取特征波长点,并根据选择结果建立PLS模型,建模结果如表6所示,表中所选波长变量序号按该波长的重要性由高到低排序。可以看出,在4 540~5 346 cm^{-1} 和6 807~7 004 cm^{-1} 的组合波段建立校正模型的预测能力最好, R^2 值最高,为0.982 6, RMSEP值最低,为0.057 4,共计12个波长变量。

表6 SPA选择的特征波长变量建模结果
Table 6 Modeling results from characteristic wavelength variables selected by SPA

波长范围/ cm^{-1}	所选波长变量序号	所选波长变量数	R^2	RMSEP	相对标准偏差/%
4 540~5 346	279、188、297、142、197、150、149、170、153、213、176、159、171、177、339、328	16	0.963 2	0.069 8	7.39
6 807~7 004	755、768、743、774、750、762	6	0.937 4	0.090 8	9.61
4 540~5 346、6 807~7 004	266、298、148、142、170、212、181、190、186、171、257、223	12	0.982 6	0.057 4	6.07

2.3.3 GA与SPA选择结果融合

从上述2种方法的最优结果中选取2组数据,一组来源于GA的28个波长变量,一组来源于SPA的12个波长变量。为了提高选择结果的可靠性,将GA和SPA的选择结果进行融合,按被选波长变量的重要性排序进行组合来建立PLS模型,组合方法及建模结果如表7所示。

表7 GA与SPA选择结果融合
Table 7 Fusion results of GA and SPA

用于建模的波长变量	波长变量数	包含共同变量数	R^2	RMSEP	相对标准偏差/%
GA和SPA各取前5个波长变量	10	0	0.976 1	0.057 8	6.12
GA和SPA各取前6个波长变量	12	0	0.985 9	0.045 1	4.77
GA和SPA各取前7个波长变量	13	1	0.977 8	0.051 7	5.47
GA和SPA各取前8个波长变量	14	2	0.980 6	0.049 4	5.23
GA和SPA各取前9个波长变量	15	3	0.981 7	0.045 9	4.86
GA和SPA各取前10个波长变量	17	3	0.981 0	0.048 3	5.11
GA和SPA各取前11个波长变量	19	3	0.986 3	0.044 7	4.73
GA和SPA各取前12个波长变量	20	4	0.984 2	0.047 5	5.03
GA和SPA所有入选波长变量	35	5	0.982 9	0.048 4	5.12

从表7可以看出,各取GA和SPA选择结果的前11个波长变量,其中包含3个共同变量,共计19个波长变量,基于这19个波长变量的模型预测精度最高,其 R^2 和

RMSEP分别为0.986 3和0.044 7;各取GA和SPA选择结果的前6个波长变量,共12个波长变量,基于这12个波长变量的模型次之,其 R^2 和RMSEP分别为0.985 9和0.045 1,与基于19点的模型预测能力相当,但是后者参与建模的变量数与前者相比减少很多,模型的计算量更小,开发专用仪器时硬件成本也更低,因此最终确定12个油脂酸值特征波长变量,它们在光谱图中的位置如图3所示。

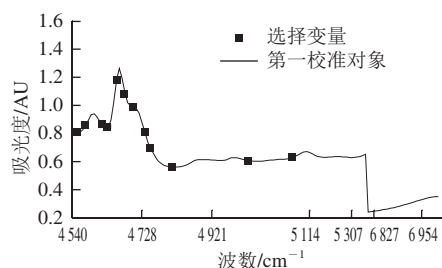


图3 12个油脂酸值近红外光谱特征波长变量位置

Fig. 3 Positions of 12 NIR characteristic wavelength variables for oil acid value

2.3.4 预测模型验证

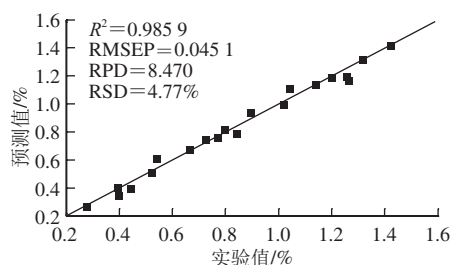


图4 20个样本预测结果

Fig. 4 Predicted results of 20 samples

用这12个特征波长变量建立模型对20个未知样品进行预测,预测效果如图4所示, R^2 和RMSEP分别为0.985 9和0.045 1,相对分析误差值为8.470,相对标准偏差为4.77%,残差值范围为0.015 6~0.109 4,说明模型的预测效果很好,证明了以上特征波长选择结果的准确性。

3 结论

通过采集不同酸值的大豆油脂样品的近红外光谱,应用iPLS、BiPLS和SiPLS方法结合GA和SPA进行油脂酸值特征波长变量优选。利用原始光谱建立的PLS校正模型,其 R^2 和RMSEP分别为0.840 6和0.104 2,参与建模的波长变量为2 075个;通过iPLS方法选择出的特征波段建立PLS校正模型,其 R^2 和RMSEP分别为0.978 9和0.064 3,参与建模的波长变量减少到262个;利用GA与SPA相结合优选出来的特征波长变量建立PLS校正模型,

R^2 和RMSEP分别为0.985 9和0.045 1, 而参与建模的波长变量仅为12个, 证明在油脂酸值近红外光谱分析中采用iPLS-GA-SPA相结合的方法进行特征波长优选能有效去除冗余信息, 降低模型复杂度, 提高模型的预测精度, 为开发专用油脂近红外光谱分析仪器及实现在线监测提供数据支持。该方法也可以推广到其他种类油脂的特征波长优选中, 为油脂各指标的快速检测提供理论依据。

参考文献:

- [1] 周石洋, 陈玲. 食用油脂中酸值测定的研究[J]. 粮食科技与经济, 2014, 39(4): 37-38. DOI:10.3969/j.issn.1007-1458.2014.04.012.
- [2] 刘娅, 赵国华, 陈宗道, 等. 近红外光谱技术及其在油脂检测中应用[J]. 粮食与油脂, 2002(11): 35-37. DOI:10.3969/j.issn.1008-9578.2002.11.015.
- [3] 刘桦, 叶晓岚, 杨光, 等. 近红外光谱技术在线监测积雪草药材活性成分的大孔树脂分离纯化过程[J]. 光谱学与光谱分析, 2013, 33(1): 98-101. DOI:10.3964/j.issn.1000-0593(2013)01-0098-04.
- [4] 徐霞, 成芳, 应义斌. 近红外光谱技术在肉品检测中的应用和研究进展[J]. 光谱学与光谱分析, 2009, 29(7): 1876-1880. DOI:10.3964/j.issn.1000-0593(2009)07-1876-05.
- [5] 朱伟兴, 江辉, 陈全胜. 特征波长筛选在近红外光谱测定梨硬度中的应用[J]. 农业工程学报, 2010, 26(8): 368-372. DOI:10.3969/j.issn.1002-6819.2010.08.062.
- [6] ADERVAL S L, ARNALDO P D, JESSICA S A P. Rapid characterization of transgenic and non-transgenic soybean oils by chemometric methods using NIR spectroscopy[J]. Spectrochimica Acta Part A, 2013, 100: 115-119. DOI:10.1016/j.saa.2012.02.085.
- [7] ARMENTA S, GARRIGUES S, de la GUARDIA M. Determination of edible oil parameters by near infrared spectrometry[J]. Analytica Chimica Acta, 2007, 596: 330-337. DOI:10.1016/j.aca.2007.06.028.
- [8] SIEMENS B J, DAUN K J. Determination of the fatty acid composition of canola, flax and solin by near-infrared spectroscopy[J]. Journal of the American Oil Chemists' Society, 2005, 82: 153-157. DOI:10.1007/s11746-005-5165-5.
- [9] ZOU Xiaobo, ZHAO Jiwen, MALCOLM J W P, et al. Variable selection methods in near infrared spectroscopy[J]. Analytica Chimica Acta, 2010, 667: 14-22. DOI:10.1016/j.aca.2010.03.048.
- [10] FRANCO A, ALEJANDRO C O. A new and efficient variable selection algorithm based on ant colony optimization. Applications to near infrared spectroscopy/partial least-squares analysis[J]. Analytica Chimica Acta, 2011, 6/9(1): 18-25. DOI:10.1016/j.aca.2011.04.061.
- [11] 陈鑫, 刘飞. MIV方法在苹果糖度近红外分析中的应用[J]. 计算机与应用化学, 2012, 29(7): 812-816. DOI:10.3969/j.issn.1001-4160.2012.07.010.
- [12] 张曼, 刘旭华, 何雄奎, 等. 岭回归在近红外光谱定量分析及最优波长选择中的应用研究[J]. 光谱学与光谱分析, 2010, 30(5): 1214-1217. DOI:10.3964/j.issn.1000-0593(2010)05-1214-04.
- [13] WU Di, CHEN Xiaojing, SHIA P, et al. Determination of α -linolenic acid and linoleic acid in edible oils using near-infrared spectroscopy improved by wavelet transform and uninformative variable elimination[J]. Analytica Chimica Acta, 2009, 634: 166-171. DOI:10.1016/j.aca.2008.12.024.
- [14] ALESSANDRA F C P, MÁRCIO J C P, FRANCISCO F G N, et al. NIR spectrometric determination of quality parameters in vegetable oils using iPLS and variable selection[J]. Food Research International, 2008, 41: 341-348. DOI:10.1016/j.foodres.2007.12.013.
- [15] 王立琦. 食用油质量近红外光谱分析技术研究[D]. 哈尔滨: 哈尔滨理工大学, 2011.
- [16] 石吉勇, 皱小波, 赵杰文, 等. BiPLS结合模拟退火算法的近红外光谱特征波长选择研究[J]. 红外与毫米波学报, 2011, 30(5): 458-462.
- [17] ALFRED A C, SUMAPORN K, DU Yiping, et al. The detection and quantification of adulteration in olive oil by near-infrared spectroscopy and chemometrics[J]. Analytical Sciences, 2004, 20: 935-940.
- [18] CASALEA M, SINELLI N, OLIVERI P, et al. Chemometrical strategies for feature selection and data compression applied to NIR and MIR spectra of extra virgin olive oils for cultivar identification[J]. Talanta, 2010, 80(5): 1832-1837. DOI:10.1016/j.talanta.2009.10.030.
- [19] RICCARDO L. Application of genetic algorithm-PLS for feature selection in spectral data sets[J]. Journal of Chemometric, 2000, 14: 643-655.
- [20] YING Yibin, LIU Yande. Non-destructive measurement of internal quality in pear using genetic algorithms and FT-NIR spectroscopy[J]. Journal of Food Engineering, 2008, 84: 206-213.
- [21] 乔立岩, 彭喜元, 马云彤. 基于遗传算法和支持向量机的特征子集选择方法[J]. 电子测量与仪器学报, 2006, 20(1): 1-5. DOI:10.13382/j.jemi.2006.01.015.
- [22] 成忠, 张立庆, 刘赫扬, 等. 连续投影算法及其在小麦近红外光谱波长选择中的应用[J]. 光谱学与光谱分析, 2010, 30(4): 949-952. DOI:10.3964/j.issn.1000-0593(2010)04-0949-04.
- [23] 钱海波, 孙来军, 王乐凯, 等. 基于连续投影算法的小麦湿面筋近红外校正模型优化[J]. 中国农学通报, 2011, 27(18): 51-56.
- [24] 孙旭东, 郝勇, 蔡丽君. 基于抽取和连续投影算法的可见近红外光谱变量筛选[J]. 光谱学与光谱分析, 2011, 31(9): 2399-2402. DOI:10.3964/j.issn.1000-0593(2011)09-2399-04.