

大豆主要脂肪酸含量近红外模型的建立

王 铎¹,张 力²,范素杰¹,韩 笑¹,吴 晗¹,王丕武¹,姚 丹¹,张 君¹

(1. 吉林农业大学 农学院,吉林 长春 130118; 2. 双辽市农业局,吉林 双辽 136400)

摘 要:为提高大豆脂肪酸品种的育种进程,使用气相色谱法测量 289 份大豆籽粒样品的脂肪酸含量,并用近红外谷物品质分析仪 NIRFIEX N-500 采集大豆样品近红外光谱值,采用主成分回归法将化学试验测定的数值与采集的光谱曲线对应拟合整理分析建立定标模型。每种脂肪酸建立 11 个定标标准,选取决定系数定标 Q 值中最高的数值为定标模型。结果显示 5 种脂肪酸定标模型的适宜 Q 值分别为:棕榈酸0.823 5、硬脂酸0.854 1、油酸0.819 6、亚油酸0.829 1、亚麻酸0.836 3。最后验证结果表明,近红外法测量值与化学试验测定值误差均在 1.5% 误差范围之内,证明所建立模型质量较好,具有使用价值,可用于大豆育种材料脂肪酸含量的快速测量。

关键词:大豆;脂肪酸含量;近红外模型

中图分类号:S565.1 **文献标识码:**A **DOI:**10.11861/j.issn.1000-9841.2017.02.0295

Establishment of NIRS Model for the Main Fatty Acids in Soybean

WANG Duo¹, ZHANG Li², FAN Su-jie¹, HAN Xiao¹, WU Han¹, WANG Pi-wu¹, YAO Dan¹, ZHANG Jun¹
(1. Agriculture College of Jilin Agricultural University, Changchun 130118, China; 2. Agricultural Bureau of Shuangliao City, Shuangliao 136400, China)

Abstract: In order to improve the efficiency of soybean fatty acid breeding and to speed up the breeding process, the fatty acid content of 289 soybean grain samples measured by GC and NIRFIEX N-500 was used to collect the NIR spectra of soybean samples. Using the principal component regression (principle component regression, referred to as PCR), the calibration model was established by fitting and analyzing the value of the chemical experiment and the corresponding spectral curve. Eleven calibration criteria were established for each fatty acid, and the highest value of calibration coefficient Q was selected as the calibration model. The results showed the fit calibration Q value for the calibration model. The fit Q value of palmitic acid was 0.8235, stearic acid was 0.8541, oleic acid was 0.8196, linoleic acid was 0.8291, and linolenic acid was 0.8363, respectively. Validation results demonstrated that the errors between the near infrared detection and chemical test results were within 1.5%, which well explained the model and could be used to measure the fatty acid contents of soybean in breeding.

Keywords: Soybean; Fatty acid content; Near - infrared model

大豆产量占世界油料作物产量的 39%,而大豆油是世界食用植物油产量的首位,大豆是世界上重要的经济作物与油料作物之一^[1]。大豆油主要组成成分为三酰甘油,其中脂肪酸含量占其含量的 90% 左右,大豆脂肪酸主要由油酸、亚油酸、亚麻酸、硬脂酸、棕榈酸 5 种脂肪酸组成。其中硬脂酸、棕榈酸为饱和脂肪酸。油酸、亚油酸、亚麻酸为不饱和脂肪酸,不饱和脂肪酸属于人体必需脂肪酸^[2-3]。不同大豆品种中的各种脂肪酸含量均具有一定的差异性^[4]。亚油酸的含量在不同大豆品种中差异最小,相对比较稳定;而亚麻酸、油酸含量的变异较大。油酸对大豆油的稳定性有很大影响,含量越高,大豆油稳定性越好,越耐运输和贮存^[5]。硬脂酸和棕榈酸为饱和脂肪酸,饱和脂肪酸不容易被人体消化,易引起脑血管疾病与肥胖症^[1]。而不饱和脂肪酸不但对减低血脂、减少血栓等有作用,而且是健脑明目不可或缺的营养物质^[6]。所以对大豆

品种的筛选不仅要求含油量高,而且要有合理的脂肪酸含量比例,才能最大程度提高人类的身体健康。而且脂肪酸含量的比例也影响到大豆油的营养、贮藏和加工,脂肪酸的量比对决定大豆油脂品质的高低有重要的影响^[7]。但是传统方法对于大豆脂肪酸育种材料的筛选是十分繁琐的,传统方法要用索氏提取器粗提,甲脂化处理,然后进行 GC 处理分析,整个过程费时费力,消耗样品的量还大,而近红外光谱分析技术却有效的解决了这个难题^[8]。

近红外光是指波长在 780 ~ 2 526 nm 范围内的电磁波,习惯上又将其分为近红外长波(780 ~ 1 100 nm)与近红外短波(1 100 ~ 2 526 nm)。近红外光谱分析技术是 20 世纪 90 年代以来发展最快、应用最广的光谱分析技术,是光谱测量技术与化学计量学学科完美结合^[9],具有分析速度快、分析效率高、分析成本低、测试重现性好、测量样品一般无需处理、便于在线分析、无损分析等多方面优点^[10]。

收稿日期:2016-11-17
基金项目:吉林省科技发展计划项目(20150204004NY);吉林省省级粮食生产发展专项(2015001)。
第一作者简介:王铎(1991 -),男,硕士,主要从事大豆分子育种研究。E-mail:499077344@qq.com。
通讯作者:张君(1968 -),男,教授,博士,主要从事作物遗传育种研究。E-mail:zhangjun@jlau.edu.cn。

运用近红外光谱分析技术测量大豆脂肪酸的含量,每个样品的测量只需十几秒钟,分析速度非常快;无需对样品进行处理,不需要其它药品试剂或者仪器,只需要完整干净的种子即可;测量时对种子没有任何损耗和伤害;分析结果与实际化学值误差小,分析效率高;结果直接呈现在电脑上,便于处理分析。因此,基于以上特点,采用近红外光谱分析技术建立大豆籽粒样品脂肪酸含量的近红外模型是完全可行的^[11-13]。

本研究以 289 份大豆种子为材料,采用主成分回归法(principle component regression,简称 PCR),将化学试验测定的数值与采集的光谱曲线对应拟合整理分析建立定标模型,拟建立 5 种大豆主要脂肪酸含量近红外光谱分析模型,为大豆主要脂肪酸性状测定提供快速准确的分析方法,从而加速大豆品质育种进程。

1 材料与方法

1.1 材料

吉农 17 × 吉农 18 为父母本构建的重组自交系 F₂。材料 289 份,由吉林农业大学农学院育种教研室提供,于 2015 年 4 月下旬种于吉林农业大学试验田,3 行区,小区行长 4.5 m,9 月下旬对材料进行收取与保存。试验用于建立定标模型的材料除品质存在差异外,百粒重、外观颜色等数量性状也存在一定的差异与不同,因此所选材料具有较好的代表性。

1.2 方法

1.2.1 近红外光谱的测定 本试验所需仪器是瑞士进口的 NIRFIEX N-500 型近红外光谱分析仪,首先将所测样品进行处理,挑选干净、完整种子装满培养皿,然后放到近红外光谱分析仪的样品槽中,种子样品量需要盖过红外扫描感应区,用 Management Console 软件先建立光谱采集程序,然后由操作软件 Operator 来控制样品的采集,每份样品种子扫描采集 3 次光谱曲线,光谱扫描曲线结果图将自动存储在所连接的电脑中。

1.2.2 脂肪酸含量的测定 大豆种子脂肪酸含量由吉林省农业科学院农业质量标准与检测技术研究所测定。分别采用加热甲酯化提取法和气相色谱分析法(GC)^[14]。

取适量的分析纯氢氧化钾加入 100 mL 甲醇反应 2 h,加入 0.3 g 大豆样品粉末与 1 mL 正己烷,震荡后放置 5 h 离心,上清液加入 0.5 mL 甲酯化试剂,震荡反应后静置 1 h,离心处理取上清备用。

GC 分析:采用岛津 GC-2014 气相色谱仪,以氮气、氢气为载气,对制备样品进行气相色谱检测^[14]。

1.2.3 光谱预处理 所采集扫描得到的光谱除包含样品的自身信息外,还包含一些其它无关信息和

噪声,如样品背景环境、杂散光信息和电噪声等。所以,在建立定标模型过程中消除无关信息和噪音是十分必要的^[15]。

本试验采用主成分回归对光谱矩阵进行主成分分析,然后将光谱主成分得分矩阵与浓度矩阵进行回归分析。首先把原始测量数据进行主成分分析,使数目较少的主成分变为元变量的线性组合,使其最大限度的反应被测样品的组成和结构的信息,通过对主成分的筛选,剃去干扰组分和干扰因素的主成分因子,然后再用其中的几个主成分与化学成分进行多元线性回归。

定标模型要建立有效光谱,未选择属性相关联的光谱不参与建模与预测集预测。

在外部验证方法中,建模集样品应该约为总光谱的 2/3 ~ 3/4。利用这些光谱产生定标模型:计算平均值光谱和主成分。通常,C-组应包含关于物质的所有“可接受极端信息”。剩余光谱将在 V-组中使用。预测集用于验证定标模型。在样品集分集时,无需考虑样品特殊状态。C-组和 V-组使用的样品应当互相独立。

1.2.4 定标与验证 选择属性、方法、光谱(C-组和 V-组光谱)以及校准波长后,应计算化学计量学结果(主成分,分数,误差等)。对于计算,软件仅利用在与校准波长和数据预处理选择有关的 C-组中选择的光谱。不作任何预处理进行第一次计算:应使用所测量的原始光谱。V-组光谱仅用于证明和判断校准。

校准结果取决于所选择的初级和次级主成分因子数量。初级主成分因子(PCs)拥有真正的光谱信息并且它们被用于重建光谱。它们确定了光谱残差。

所使用的初级主成分因子数量越多,允许的误差越小。为避免过拟合和拟合不足现象出现,在建模过程中不要使用过多或过少的主成分因子数。

次级主成分因子用于区分不同种类物质以及簇半径计算。初级主成分因子数量限制着次级主成分因子数量。

本试验由于所用仪器的自动选择调节性,所选初级主成分因子与次级主成分因子的选择和确定均是计算机自主选择。

样品的化学数据和扫描数据经过处理后,运用定标软件 NIRCal 建立定标方程,此时,向导从最优化过程开始启动,这将花费数分钟。可在“定标 Q-值概观(calibration Q-value overview)”中直观进展。在最优化过程中,Q-值从 0(无工作定标)上升直至接近 1(=完美定标)。自动定标向导结束后,运算结果显示在一个列表中,列表顶部显示最佳建模结果(Q 值最高)。

2 结果与分析

2.1 剔除超常样品的结果分析

将超常样品曲线剔除之后,剩余的 253 份样品

光谱曲线用于构建模型。与未去除超常样品光谱曲线相比较,处理后的样品光谱曲线分布平均,趋势一致,显著改善了光谱曲线部分过于分散、集中和错误的情况,使此样品光谱曲线具有代表性。

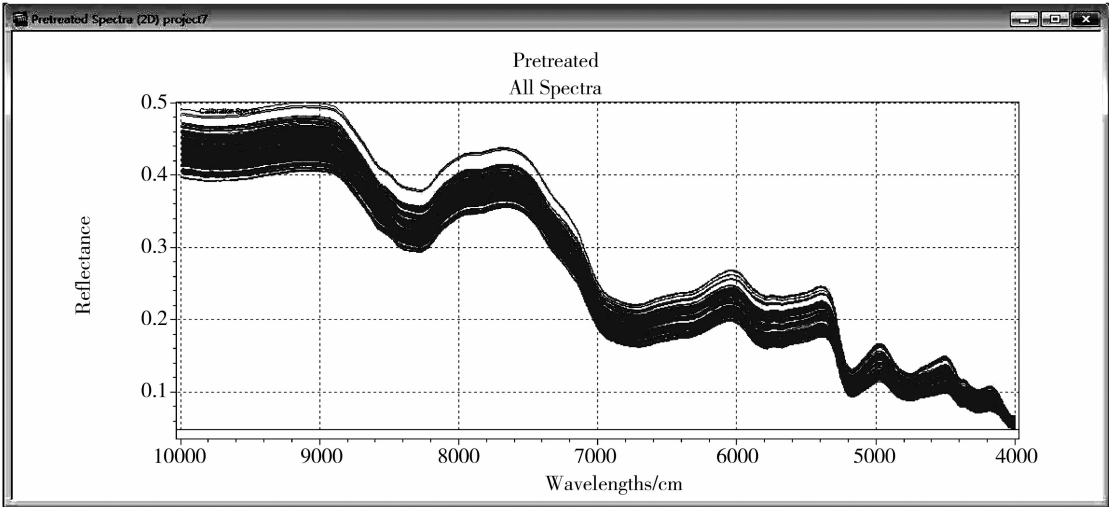


图 1 处理后剩余大豆光谱

Fig. 1 Residual soybean spectrum after treatment

2.2 脂肪酸定标集与验证集的分选

剔除超常样品后,剩余光谱曲线用于建立脂肪酸近红外模型,随机从中选取 50 份样品构成验证集,余下 203 份样品生成定标集。从表 1 中可以看出,各试验材料脂肪酸含量的验证集均包含在定标

集范围内,满足定标的要求。但是试验材料属于重组自交系,其亲缘相近,所以各脂肪酸含量变化范围比较小,相对过于集中,或许会对定标结果产生一定的局限性影响。

表 1 5 种脂肪酸含量定标集和验证集的分析数据

Table 1 Analysis data of five kinds of fatty acid content calibration set and validation set

脂肪酸	组分	最小值	最大值	均值	样本数
Fatty acid	Constituent	Minimum/%	Maximum/%	Mean/%	Sample size N
亚油酸	定标集 Calibration set	52. 16	58. 58	56. 29	203
Linoleic acid	验证集 Validation set	52. 32	58. 47	56. 25	50
硬脂酸	定标集 Calibration set	3. 39	5. 37	4. 60	203
Stearic acid	验证集 Validation set	3. 65	5. 12	4. 61	50
油酸	定标集 Calibration set	17. 41	26. 46	20. 67	203
Oleic acid	验证集 Validation set	17. 68	26. 01	20. 64	50
棕榈酸	定标集 Calibration set	9. 82	12. 31	10. 90	203
Palmitic acid	验证集 Validation set	9. 99	12. 03	11. 01	50
亚麻酸	定标集 Calibration set	6. 01	8. 83	7. 54	203
Linolenic acid	验证集 Validation set	6. 12	8. 12	7. 52	50

2.3 定标模型的建立

通过软件 NIRCal,将 203 份定标集大豆样品的 5 种脂肪酸含量化学测定值与光谱曲线数据一一对应输入程序中,然后对其进行标定处理,分别建立 5 种脂肪酸含量的定标模型。本试验采用主成分回归法将这二者拟合整理建立定标模型。每种脂肪酸建立 11 个定标标准,选取定标 Q 值最高的标准作为定标模型。结果显示:棕榈酸 Q 值为0. 823 5;硬脂酸 Q 值为0. 854 1;油酸 Q 值为0. 819 6;亚油酸 Q

值为0. 829 1;亚油酸 Q 值为0. 836 3。

2.4 模型的验证

对于建立的近红外定标模型,对其后续的验证是必不可少的过程,通过验证得到的结果能更好的评估所建立模型的优劣,检验其在实际使用中的效果。利用新建立的 NIR 定标模型,对另外 10 个大豆样品进行脂肪酸含量扫描测量,然后用化学方法对其测量,比较化学法测定值和近红外预测值,验证模型的准确性,评估其实际的应用价值。

表2 化学法与近红外测定5种脂肪酸含量比较

Table 2 Comparison of five kinds of fatty acids by chemical method and near infrared spectroscopy(%)

棕榈酸 Palmitic acid		
化学法测定值 Determination of chemical method	近红外预测值 NIR prediction	绝对误差 Absolute error
10.28	11.33	1.05
11.36	11.69	0.33
10.88	10.23	-0.65
11.53	10.99	-0.54
11.19	11.98	0.79
10.49	11.03	0.54
11.51	10.87	-0.64
11.89	11.02	-0.87
11.96	11.35	-0.61
10.20	10.88	0.68

硬脂酸 Stearic acid		
化学法测定值 Determination of chemical method	近红外预测值 NIR prediction	绝对误差 Absolute error
4.46	4.96	0.50
4.86	4.35	-0.51
4.83	4.43	-0.40
5.20	5.63	0.43
5.01	4.86	-0.15
4.66	5.43	0.77
4.57	5.65	1.08
4.42	4.01	-0.41
4.29	4.89	0.60
4.37	4.81	0.44

油酸 Oleic acid		
化学法测定值 Determination of chemical method	近红外预测值 NIR prediction	绝对误差 Absolute error
20.92	21.63	0.71
20.05	20.45	0.40
23.91	22.96	-0.95
22.20	21.86	-0.34
21.44	22.32	0.88
19.45	20.23	0.78
20.12	20.60	0.48
19.62	18.85	-0.77
19.35	19.03	-0.32
19.70	20.20	0.50

亚油酸 Linoleic acid		
化学法测定值 Determination of chemical method	近红外预测值 NIR prediction	绝对误差 Absolute error
56.82	55.96	-0.86
55.72	55.98	0.26
53.84	53.07	-0.77
53.78	54.50	0.72
55.07	55.89	0.82
58.14	59.36	1.22
56.14	55.80	-0.34
56.31	55.93	-0.38
55.46	56.55	1.09
58.00	58.31	0.31

亚麻酸 Linolenic acid		
化学法测定值 Determination of chemical method	近红外预测值 NIR prediction	绝对误差 Absolute error
7.52	7.96	0.44
8.01	7.83	-0.18
6.54	7.30	0.76
7.29	8.09	0.80
7.30	6.50	-0.80
7.26	6.84	-0.42
7.66	7.16	-0.50
7.76	7.99	0.23
7.93	8.32	0.39
7.73	7.30	-0.43

表2结果显示,近红外检测结果与标准方法测定结果误差在1.5%以内,说明近红外检测结果准确可靠,可应用于大豆脂肪酸快速测定。

3 讨论

近红外光谱分析技术具有效率高、测量速度快、不破坏样品等优点,但是其测量准确性与稳定性容易受到多种因素的影响。对这些影响因素的深入研究分析,降低负影响,是今后建立近红外模型提高效率和减少误差的关键^[16]。

试验证明,所建立的近红外模型能够准确、快速的测定大豆脂肪酸的含量,能进一步缩短育种的年限和缩小育种的规模,提高育种效率。由此可见,近红外光谱分析技术对现代育种具有非常重要的作用。但本试验所建模型并不十分理想,分析可能以下几方面的因素:1. 大豆样品籽粒的整齐度

与均一性,此原因对模型的精准度有着重要的影响;2. 光谱采集时的外部环境的变化对光谱的采集具有一定影响;3. 样品的采集光谱时间与化学测定的时间间隔较长,或许影响种子的品种;4. 气相色谱法测量值的准确性也直接影响着建模的准确性^[17]。

因为近红外光的信号比较微弱,光谱所储存的信息量大、谱区又有重叠,尤其是在测量含量低的样品时,误差往往会被放大,所以需要进一步降低这几方面的影响,提高分析结果的可靠性。所以要

对所需样品进行筛选,对光谱数据进行预处理^[18]。

由于试验材料选择对本试验的影响,虽然所用材料数量充足,但各组分的脂肪酸含量差别比较小,样品具有一定的局限性,对于所建立近红外模型的结果有一定的影响,以后应该适当补充不同大豆材料,增加样品品种和数量,使其更具有代表性。

4 结 论

本研究所建立了 5 种脂肪酸近红外模型,定标 Q 值分别为棕榈酸 0.823 5、硬脂酸 0.854 1、油酸 0.819 6、亚油酸 0.829 1、亚油酸 0.836 3;验证结果显示,近红外检测结果与标准方法测定结果误差均在 1.5% 之内,证明所建立模型较好,可用于大豆育种材料的脂肪酸含量快速测定。

参考文献

[1] 宋晓昆,张颖君,闫龙,等. 大豆脂肪酸组份相关、变异特点分析[J]. 华北农学报,2010(S2):68-73. (Song X K,Zhang Y J,Yan L,et al. A study on correlation and variability of fatty acid composition contents of soybean cultivars[J]. Acta Agriculturae Boreali-Sinica,2010(S2):68-73.)

[2] 任波,李毅. 大豆种子脂肪酸合成代谢的研究进展[J]. 分子植物育种,2005(3):301-306. (Ren B, Li Y. Research advances on fatty acid biogynthesis metabolism in soybean seed[J]. Molecular Plant Breeding,2005,3(3):301-306.)

[3] 邹筱,韩粉霞,陈明阳,等. 大豆脂肪酸主要组分含量 QTL 定位[J]. 作物学报,2014,40(9):1595-1603. (Zou X,Han F X,Chen M Y,et al. Quantitative trait loci associated with major fatty acid components in soybean[J]. Acta Agronomica Sinica,2014,40(9):1595-1603.)

[4] 王晓燕,张彩英,贾晓艳. 河北省大豆品种脂肪酸组成与含量分析[J]. 河北农业大学学报,2007(2):15-18. (Wang X Y,Zhang C Y, Jia X Y. Analysis of fatty acids composition and content in soybean varieties in Hebei province[J]. Journal of Agricultural University of Hebei,2007,30(2):15-18.)

[5] 尹田夫. 大豆油脂脂肪酸改良与生化育种策略[J]. 大豆科学,1988,7(1):75-79. (Yin T F. Improvement on fatty acid of ail lipid and strategy of biochemistry breeding in soybean[J]. Soybean Science,1988,7(1):75-79.)

[6] 李志香,沈翠平. 多不饱和脂肪酸对人体的作用[J]. 生物学通报,1998,33(1):11-12. (Li Z X,Shen C P. Role of polyunsatu-

rated fatty acids in human body[J]. Bulletin of Biology,1998,33(1):9-11.)

[7] 徐杰,胡国华,张大勇. 大豆种子脂肪酸组分的研究进展[J]. 大豆科学,2005,24(1):61-66. (Xu J, Hu G H, Zhang D Y. Study advance on fatty acid composition of soybean[J]. Soybean Science,2005,24(1):61-66.)

[8] 张颖君,高慧敏,蒋春志,等. 大豆种子脂肪酸含量的快速测定[J]. 大豆科学,2008,27(5):859-862. (Zhang Y J, Gao H M, Jiang C Z, et al. Fast analysis on fatty acids of soybean seed by gas chromatography[J]. Soybean Science, 2008, 27(5):859-862.)

[9] McClure W F. Near-infrared spectroscopy: The giant is running strong[J]. Analytical Chemistry,1994,66(1):43A-53A.

[10] 徐广通,袁洪福,陆婉珍. 现代近红外光谱技术及应用进展[J]. 光谱学与光谱分析,2000,20(2):134-142. (Xu G T, Yuan H F, Lu W Z. Development of modern near infrared spectroscopic techniques and its applications[J]. Spectroscopy and Spectral Analysis,2000,20(2):134-142.)

[11] 柴玉华,谭克竹. 基于近红外分析技术检测大豆脂肪酸含量的研究[J]. 农业工程学报,2007,23(1):238-241. (Chai Y H, Tan K Z. Measurement of soybean fatty acid content by near infrared spectroscopy[J]. Transactions of the CSAE, 2007, 23(1):238-241.)

[12] 梁丹. 应用近红外光谱分析技术定量检测植物油脂脂肪酸含量的研究[J]. 安徽农业科学,2012,40(30):14933-14936. (Li-ang D. Quantitative detection research for the fatty acid of vegetable oil using near infrared spectroscopy[J]. Journal of Anhui Agricultural Sciences,2012,40(30):14933-14936.)

[13] 闫龙,蒋春志,于向鸿,等. 大豆粗蛋白、粗脂肪含量近红外检测模型建立及可靠性分析[J]. 大豆科学,2008,27(5):833-837. (Yan L,Jiang C Z,Yu X H,et al. Development and reliability of near in frared spectroscopy (NIS) models of protein and oil content in soybean[J]. Soybean Science,2008,27(5):833-837.)

[14] 范胜栩,李斌,孙君明,等. 气相色谱方法定量检测大豆 5 种脂肪酸[J]. 中国油料作物学报,2015(4):548-553. (Fan S Y, Li B, Sun J M,et al. A quantitative gas chromatographic method for determination of soybean seed fatty acid components[J]. Chinese Journal of Oil Crop Sciences,2015(4):548-553.)

[15] 李勇,魏益民,王锋. 影响近红外光谱分析结果准确性的因素[J]. 核农学报,2005(3):236-240. (Li Y, Wei Y M, Wang F. Affecting factors on the accuracy of near-infrared spectroscopy analysis[J]. Acta Agriculturae Nucleatae Sinica,2005(3):236-240.)

[16] 刘青格. 近红外光谱的信息特征提取[D]. 镇江:江苏大学,2003. (Liu Q G. The signal features extraction of Near-infrared Spectroscopy[D]. Zhenjiang:Jiangsu University,2003.)

[17] 褚小立,袁洪福,陆婉珍. 基础数据准确性对近红外光谱分析结果的影响[J]. 光谱学与光谱分析,2005,25(6):886-889. (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,2005,25(6):886-889.)

[18] 熊英. 近红外光谱的原理及应用[J]. 中山大学研究生学刊(自然科学·医学版),2013(2):16-30. (Xiong Y. The principle and application of near infrared spectroscopy[J]. Journal of the Graduates Sun Yat-Sen University(Natural Sciences & Medicine edition),2013(2):16-30.)