



基于近红外漫反射光谱法的大豆粗蛋白和粗脂肪含量的快速检测

王丽萍, 陈文杰, 赵兴忠, 张 新

(陕西省杂交油菜研究中心/国家油料作物改良中心陕西分中心, 陕西 杨陵 712100)

摘要:为满足大豆品质育种快速筛选的需求, 本文详细探讨了利用近红外漫反射光谱法对大豆粗蛋白和粗脂肪含量实现快速测定的可行性。采用凯氏定氮法和索氏抽提法测定了 120 份大豆粗蛋白和粗脂肪的含量, 分别采集大豆整粒和粉末两种状态的近红外光谱, 然后运用化学计量学方法 PLS 建立近红外光谱与化学值之间的关系模型。其中粉末大豆样品建立的粗蛋白校正模型的决定系数 R^2 为 0.978 7, 校正标准误差 $RMSECV$ 为 0.003 8, 该模型对 24 份待测样品进行测定的预测标准误差 $RMSEP$ 为 0.002 84; 粗脂肪校正模型的 R^2 为 0.934 1, $RMSECV$ 为 0.003 69, $RMSEP$ 为 0.003 53。整粒大豆建立的粗蛋白校正模型的 R^2 为 0.872 4, $RMSECV$ 为 0.009 07, $RMSEP$ 为 0.007 49; 粗脂肪校正模型的 R^2 为 0.876 5, $RMSECV$ 为 0.005 08, $RMSEP$ 为 0.004 66。对比发现, 建模样品的状态对近红外模型的预测性能有重要影响, 样品在粉末状态下建立的粗蛋白和粗脂肪近红外模型的预测效果更好。另一方面, 由于整粒样品建立的近红外模型的 R^2 均在 0.87 以上, 因此当样品量较少没有足够样品可用于粉碎时, 该模型可以满足对整粒大豆品质进行粗测的需求。该结果对大豆育种早期筛选工作具有重要意义。

关键词: 大豆; 近红外模型; 快速检测; 粗蛋白; 粗脂肪

Rapid Determination of Crude Protein and Crude Oil Content of Soybean Based on Near Infrared Diffuse Reflectance Spectroscopy

WANG Li-ping, CHEN Wen-jie, ZHAO Xing-zhong, ZHANG Xin

(Hybrid Rapeseed Research Center of Shaanxi Province / Shaanxi Branch of National Oil Crop Improvement Center, Yangling 712100, China)

Abstract: In order to meet the need of quick determination of soybean quality, the feasibility of rapid determination of crude protein and crude fat content in soybean by near infrared diffuse reflectance spectroscopy was discussed in detail. The chemical values of crude protein and crude fat content of 120 soybeans were assayed by Kjeldahl and Soxhlet methods separately. Meanwhile, near-infrared spectra of samples in the two states of whole seed and powder were collected. Finally, the correlation models between soybean spectra and chemical values were built by the partial least square (PLS) method in chemometrics. The determination coefficient (R^2) of calibration model of crude protein content was 0.978 7 and the root mean square error of cross validation ($RMSECV$) was 0.003 8 with soybean samples in the state of powder. When 24 test samples were predicted, the root meansquare error of prediction ($RMSEP$) was 0.002 84. The R^2 and $RMSECV$ of crude oil content model was 0.934 1 and 0.003 69 respectively, and $RMSEP$ was 0.003 53. The R^2 and $RMSECV$ of protein model set up by soybean whole seed samples was 0.872 4 and 0.009 07 respectively, and $RMSEP$ was 0.007 49. The R^2 and $RMSECV$ of oil content model was 0.876 5 and 0.005 08 respectively, and $RMSEP$ was 0.004 66. It was found that the state of calibration samples had a significant effect on the prediction ability of the model. The results indicated that the performance of crude protein and crude oil content models built by powder samples was better. On the other hand, the R^2 of all models that were built by soybean whole seed samples were more than 0.87, so the models could be used to measure soybean quality roughly when the sample was not enough to grind. The results were of great importance in early screening of soybean breeding.

Keywords: Soybean; Near infrared model; Rapid determination; Protein; Crude oil

大豆含有丰富的营养物质, 是植物蛋白和植物油的主要来源之一。为满足人们的多种需求, 育种家在加强大豆高产育种的同时, 也注重品质育种的研究^[1]。在大豆品种审定中, 粗蛋白和粗脂肪含量被作为衡量大豆品质优劣的两个基本指标, 其常规

测定方法分别是凯氏定氮法和索氏抽提法, 这两种方法虽然可靠性高, 但存在操作复杂和耗时多等缺点, 为提高大豆育种工作的效率, 急需开发一种新型快速测定方法。

近红外光谱分析在 20 世纪中叶首先从农业领

收稿日期: 2018-10-09

基金项目: 陕西省现代农作物种业项目。

第一作者简介: 王丽萍(1979 -), 女, 硕士, 助理研究员, 主要从事品质分析检测工作。E-mail: 421536061@qq.com。

域开始,之后迅速发展成为近代一项高新技术^[2]。它利用有机物在近红外光谱区的振动吸收而快速测定样品中多种化学成分含量。蛋白质、脂肪糖类等含有的各种含氢基团的倍频与合频谱带恰好落在近红外区,通过得到这些有机分子含氢基团的特征振动信息,从而测定其化学成分的含量^[3-4]。近红外漫反射光谱分析技术(NIRS)作为一种新型的物理测试技术,具有无损、快速、高通量、低成本等优点^[5],已被广泛应用于大豆^[6-7]、小麦^[8-9]、油菜^[10-11]、玉米^[12-13]、水稻^[14-15]等种子的质量检测。虽然近红外漫反射光谱法已较早应用于大豆的品质检测,但现有研究中都提到需对样品进行粉碎,而在实际育种过程中很多样品由于珍贵难得,需要进行无损伤检测。因此,本试验利用 MATRX-I 近红外光谱仪采集大豆颗粒和粉末两种状态的近红外光谱,并采用凯氏定氮法和索氏抽提法测定样品的粗脂肪和粗蛋白含量,然后利用 OPUS 软件自动优化建模谱区和选择数据预处理方法,构建近红外光谱与含量之间的偏最小二乘法(PLS)关系模型,最后依据软件给出的模型参数确定最优模型并通过预测集对定标模型进行验证。通过分析比较不同样品状态下建立的近红外模型预测性能的差异,探讨利用近红外漫反射光谱法测定大豆籽粒品质的可行性。

1 材料与方法

1.1 材料

1.1.1 供试样品 试验样品由陕西省杂交油菜研究中心豆麦研究室提供,主要包括大豆品种资源 15 份、大豆育种材料 75 份和国家农作物品种区域试验大豆 30 份。

1.1.2 仪器 光谱采集使用德国 Bruker 公司的 MPA 型傅里叶变换近红外光谱仪(配有镀金积分球,样品旋转器和 4.2 cm 石英样品杯,单粒样品台

和 2.0 cm 安培瓶,镀金背景和 PbS 检测器)。脂肪测定仪:海能 SOX606;凯氏定氮仪:BUCHI K370;粉碎机:天津泰斯特 FW80。

1.2 试验设计

从1 000份大豆种质资源中选取具有代表性的大豆样品 120 份,粗蛋白和粗脂肪含量呈均匀分布,粗蛋白含量的分布范围为 34.53%~48.13%,粗脂肪含量的分布范围为 14.98%~22.53%。将入选样品依据成份含量排序后按照 4:1 将样品分为校正集和预测集两部分,使两组样品的成份含量均呈一定梯度,校正集用于建立粗蛋白和粗脂肪的近红外模型,预测集用于对模型的预测能力进行评价。

1.3 方法

1.3.1 索氏抽提法 采用 GB 5009.6-2016 国家标准方法测定大豆粗脂肪含量。

1.3.2 凯氏定氮法 采用 GB/T 14489.2-2008 国家标准方法测定大豆粗蛋白含量。

1.3.3 近红外光谱法 光谱采集条件:谱区范围 12 000~4 000 cm⁻¹,分辨率 8 cm⁻¹,扫描次数 64 次,室温 25~26℃。为了获得丰富的样品信息,每个样品 3 次重复。考虑到大豆籽粒及其间隙均较大,选用不同的样品杯采集样品的近红外光谱,大豆籽粒选用大尺寸的旋转杯而粉末样品则选用小尺寸的安培瓶。

1.4 数据分析

利用 OPUS 7.0 软件建立近红外光谱与成份含量值之间的数学模型,并通过预测集对近红外模型的准确度做出评价,同时采用 Origin 8.6 绘图。

2 结果与分析

2.1 样品粗蛋白、粗脂肪含量

分别采用凯氏定氮法和索氏抽提法测定 120 份大豆样品的粗蛋白质和粗脂肪含量的分布情况见表 1。

表 1 大豆粗蛋白、粗脂肪含量化学法测定结果

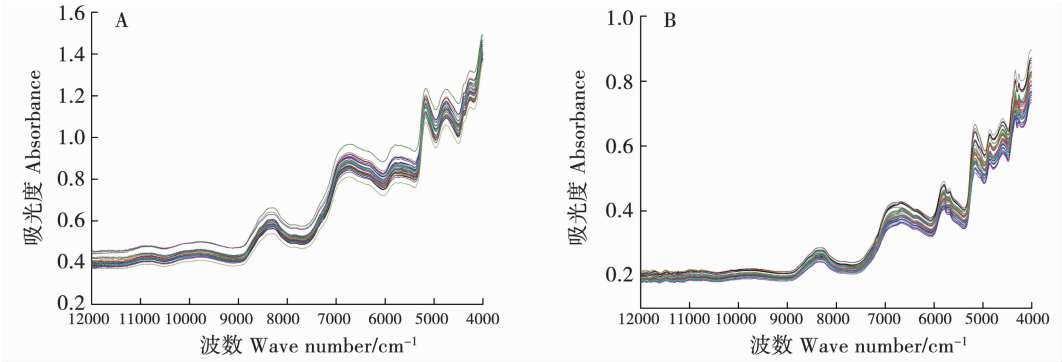
Table 1 Chemical determination results of crude protein and crude oil content

品质指标 Quality index		样品数 Sample number	最小值 Minimum /%	最大值 Maximum /%	平均值 Average /%
粗蛋白 Crude protein	总计 Total	120	34.53	48.13	42.64
	校正集 Calibration set	96	34.53	48.13	42.74
	预测集 Prediction set	24	36.63	47.32	42.50
粗脂肪 Crude oil	总计 Total	120	14.98	22.53	18.79
	校正集 Calibration set	96	14.98	22.53	18.82
	预测集 Prediction set	24	15.63	22.21	18.60

2.2 光谱数据采集

大豆颗粒及粉末样品的近红外光谱图如图 1 所示。通过谱图比对发现:吸收光谱在4 000 ~ 6 000

cm⁻¹这个区间段存在较大的差异,粉末样品的光谱图中有更多的吸收峰,其携带了更丰富的样品内部信息。



A:大豆整粒; B:大豆粉末。下同。
A: Soybean whole seed; B: Soybean powder. The same below.

图1 大豆样品的近红外光谱图

Fig. 1 Near-infrared spectra of soybean

2.3 模型的构建优化

为了消除背景干扰,采用消除常数偏移量、减去一条直线、矢量归一化、最大-最小归一化、多元散射校正、内部标准、一阶导数、二阶导数、一阶导数+减去一条直线、一阶导数+矢量归一化和一阶导数+MSC共11种光谱预处理方法对大豆的原始光谱进行处理后作为输入变量,通过特征向量的提取与内部交叉验证建立化学值与近红外光谱之间的偏最小二乘法(PLS)回归预测模型。根据OPUS 7.0分析软件得出的决定系数(R^2)、校正标准误差(RMSECV)作为评判指标,反复剔除异常样品,最终

确定最优模型,其模型的相关参数见表2及图2和3。从图2可以直观的看出由粉末样品建立的大豆蛋白质校正曲线的预测值与实测值的吻合程度更高,样品在中心线的分布更集中,同时由表2可知粉末样品模型 R^2 高达0.9787,而颗粒样品模型的 R^2 却仅为0.8724。从表2和图3可以看出,不同样品状态下建立的大豆粗脂肪模型的校正曲线差异显著,粉末样品模型的预测值和实测值之间的线性相关性更高,对应的 R^2 及RMSECV均较优。因此,对样品进行粉碎可以显著提高大豆近红外模型的预测准确度,从而改善校正模型的预测效果。

表2 在不同状态样品大豆粗蛋白和粗脂肪含量的近红外模型参数

Table 2 Parameters of NIR model of crude protein and crude oil content of soybean in different states							
校正模型		校正集 Calibration set			预测集 Prediction set		
Calibration model		含量范围 Content range/%	R^2	RMSECV	含量范围 Content range/%	$RMSEP$	RPD
大豆整粒粗蛋白模型 Crude protein model of soybean whole seed		34.53 ~ 48.13	0.8724	0.00907	36.63 ~ 47.32	0.00749	3.56
	大豆粉末粗蛋白模型 Crude protein model of soybean powder	34.53 ~ 48.13	0.9787	0.0038	36.63 ~ 47.32	0.00284	9.88
大豆整粒粗脂肪模型 Crude oil model of soybean whole seed		14.98 ~ 22.53	0.8765	0.00508	15.63 ~ 22.21	0.00466	3.13
	大豆粉末粗脂肪模型 Crude oil model of soybean powder	14.98 ~ 22.53	0.9341	0.00369	15.63 ~ 22.21	0.00353	4.19

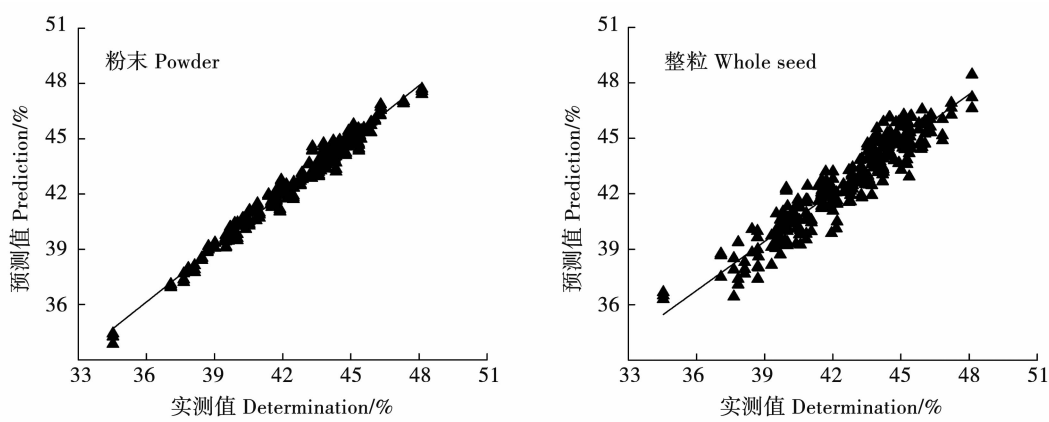


图 2 两种样品状态下大豆粗蛋白校正集的实测值与近红外预测值的相关图
Fig. 2 Correlation between NIR prediction and determination of soybean protein for calibration sets in the two states

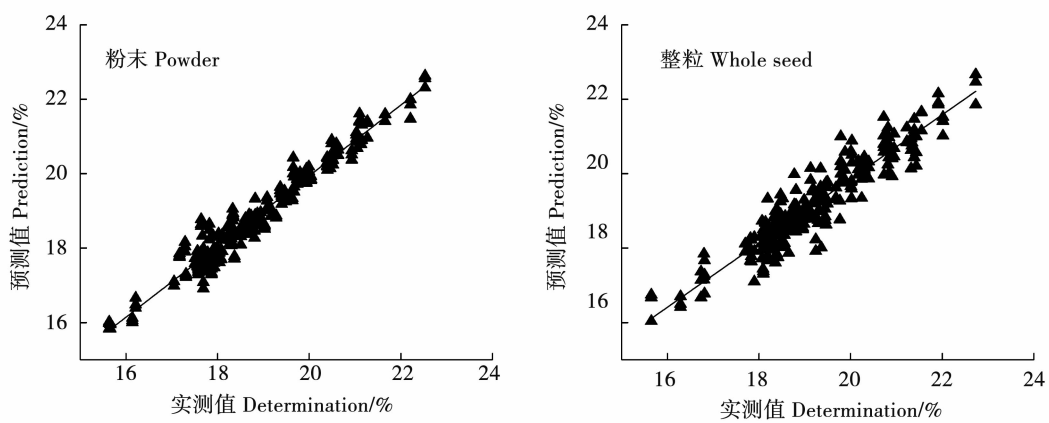


图 3 两种样品状态大豆粗脂肪校正集的实测值与近红外预测值的相关图
Fig. 3 Correlation between measured value and NIR predicted value of soybean crude fat content calibration sets in the two states

2.4 稳定性和准确性评价

利用 24 个预测集样品对相应的近红外模型预测能力进行验证,以预测标准误差($RMSEP$)和残留预测偏差(RPD)作为评价指标,其结果见表 2 及图

4、5。由表 2 可知,对样品进行粉碎后,粗蛋白模型预测集的 $RMSEP$ 由 0.007 49 降低为 0.002 84, RPD 由 3.56 提高到了 9.88,粗脂肪模型预测集的 $RMSEP$ 由 0.004 66 降低为 0.003 53, RPD 由 3.13 提高到 4.19。

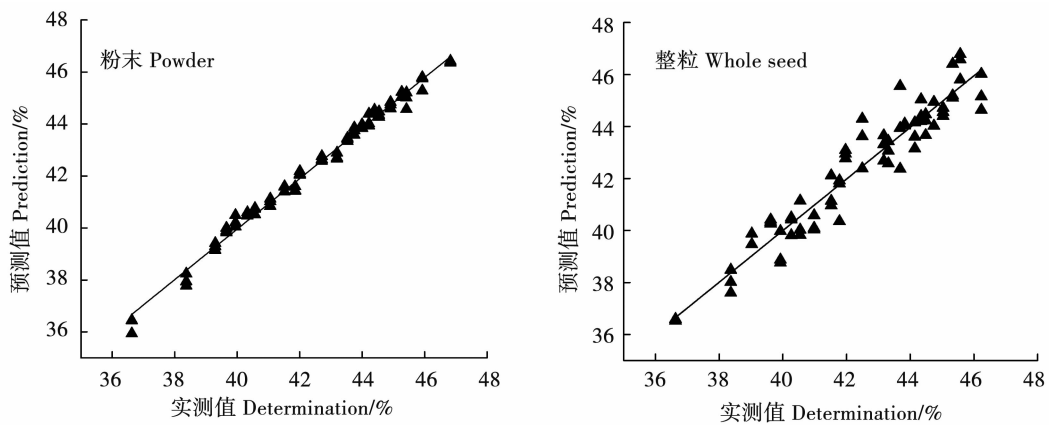


图 4 两种样品状态下大豆粗蛋白含量检验集的实测值与近红外预测值相关图
Fig. 4 Correlation between measured value and NIR predicted value of soybean crude protein content prediction sets in the two states

由图4、5可以看出,样品在粉末状态下建立的大豆粗蛋白和粗脂肪近红外模型对预测集进行拟合时有更小的预测误差,准确度更高一些。以上试验结果进一步证明建模样品的状态对近红外模型的预测性能有重要影响,粉末样品建立的校正模型

有更高的准确度和稳定性。同时,整粒大豆建立的近红外模型的 $RMSEP$ 均较小,且 RPD 大于3,说明由大豆籽粒建立的模型准确度也较好,能够满足在大豆高世代育种中的初筛需求。

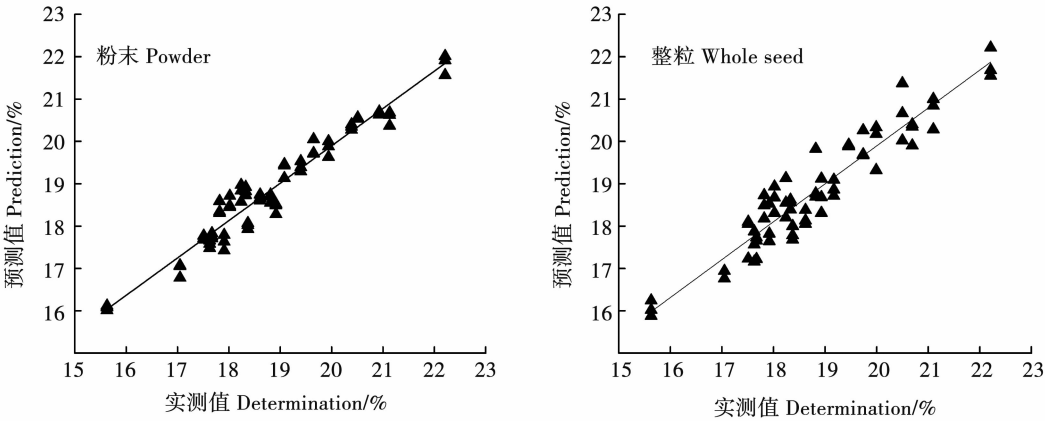


图5 两种样品状态下大豆粗脂肪含量检验集的实测值与近红外预测值相关图
Fig. 5 Correlation between measured value and NIR predicted value of soybean crude oil content content prediction sets in the two states

3 讨论

目前大豆粗蛋白和粗脂肪含量的快速测定方法主要是近红外光谱法,依据光谱采集方式可分为漫反射光谱法和透射光谱法。很多文献中提到需对大豆样品进行粉碎,而对于那些量少的样品一经粉碎就无法延续后代,使其在大豆品种改良后代的选育过程中应用受到限制。为此,本文探讨利用近红外漫反射光谱法测定整粒大豆品质的可行性。分别采集整粒和粉末样品的近红外光谱并建立各自的近红外模型,试验发现,粉末样品建立的近红外校正模型的性能远远好于整粒大豆建立的模型,这可能是由于样品颗粒间的空隙较大,漫反射光波长较长不易深入大豆颗粒内部,漫反射取得的光信号主要来自于浅层样品;而粉末样品由于颗粒较小,其表层样品与内部样品的信息基本一致。因此,粉末样品的近红外光谱携带更多的样品信息,由其建立的近红外模型的准确度也更高一些。另一方面虽然整粒样品建立的近红外模型的预测效果略差一些,但其相关系数均在0.87以上且 RPD 大于3,可以满足大豆品种选育过程中的初筛需求。因此,在样品的测定过程中,可以根据实际的需求灵活选择不同的模型,对于量少的种子只能选择整粒模型对样品进行粗测,而量大的样品既可以粗测又能够精测。

4 结论

本文建立了利用近红外漫反射光谱法对大豆粗蛋白和粗脂肪含量快速测定的方法。其中大豆粉末建立的粗蛋白近红外模型的 R^2 为0.978 7, $RMSECV$ 为0.003 8, $RMSEP$ 为0.002 84, RPD 为9.88;粗脂肪近红外模型的 R^2 为0.934 1, $RMSECV$ 为0.003 69, $RMSEP$ 为0.003 53, RPD 为4.19。大豆颗粒建立的粗蛋白校正模型的 R^2 为0.872 4, $RMSECV$ 为0.009 07, $RMSEP$ 为0.007 49, RPD 为3.56;粗脂肪校正模型的 R^2 为0.876 5, $RMSECV$ 为0.005 08, $RMSEP$ 为0.004 66, RPD 为3.13。结果表明,样品的状态对近红外模型的预测性能有重要影响,对样品进行粉碎可以提高模型的准确度。虽然,整粒样品近红外模型的预测效果略差一些,但对保留量少的种子具有重要的现实意义。同时,大豆粉末建立的粗蛋白和粗脂肪的近红外校正模型具有较高的稳定性和准确度,可以实现对大豆粗蛋白和粗脂肪含量进行准确的分析,提高了检测效率,满足大豆育种快速筛选的需求。

参考文献

[1] 王海杰,林家贵,王敏芬. 大豆品质育种研究进展[J]. 生物灾害科学,2015, 38(2): 102-105. (Wang H J, Lin J G, Wang M F. Research advance of soybean quality breeding[J]. Biological Disaster Science, 2015, 38(2): 102-105.)

- [2] 严衍禄, 陈斌, 朱大洲, 等. 近红外光谱分析的原理、技术与应用[M]. 北京: 中国轻工业出版社, 2013: 178. (Yan Y L, Chen B, Zhu D Z, et al. Near infrared spectroscopy principles, technologies and applications[M]. Beijing: China Light Industry Press, 2013: 178.)
- [3] 刘晓庚. 近红外光谱技术在粮油储藏及其品质分析中的应用[J]. 粮食储藏, 2007, 36(1): 29-35. (Liu X G. Application of near infrared spectroscopy in grain and oil storage and quality analysis[J]. Grain Storage, 2007, 36(1): 29-35.)
- [4] 姚鑫森, 张瑞英, 李霞辉, 等. 近红外透射光谱法(NITS)分析大豆品质的研究[J]. 大豆科学, 2006, 25(4): 417-420. (Yao X M, Zhang R Y, Li X H, et al. The study of soybean quality analysis by near infrared transmission spectroscopy (NIRS) [J]. Soybean Science, 2006, 25(4): 417-420.)
- [5] 朱丽伟, 马文广, 胡晋, 等. 近红外光谱技术检测种子质量的应用研究进展[J]. 光谱学与光谱分析, 2015, 35(2): 346-349. (Zhu L W, Ma W G, Hu J, et al. Advances of NIR spectroscopy technology applied in seed quality detection[J]. Spectroscopy and Spectral Analysis, 2015, 35(2): 346-349.)
- [6] 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 infrared spectroscopy(NIRS) models of protein and oil content in soybean[J]. Soybean Science, 2008, 27(5): 833-837.)
- [7] 王燕, 鞠涛, 刘晓兰, 等. 近红外光谱法预测大豆营养成分含量模型的建立和应用[J]. 营养饲料, 2014, 50(7): 62-65. (Wang Y, Ju T, Liu X L, et al. Building and application of soybean nutrients content model predicted by near Infrared spectral method[J]. Nutrition and Feedstuffs, 2014, 50(7): 62-65.)
- [8] 李小龙, 秦丰, 赵龙莲, 等. 近红外光谱技术的小麦条锈病严重程度分级识别[J]. 光谱学与光谱分析, 2015, 35(2): 367-371. (Li X L, Qin F, Zhao L L, et al. Identification and classification of disease severity of wheat stripe rust using near infrared spectroscopy technology[J]. Spectroscopy and Spectral Analysis, 2015, 35(2): 367-371.)
- [9] 时伟芳, 谢宗铭, 杨丽明, 等. 基于近红外光谱技术的春小麦单粒种子活力鉴定[J]. 麦类作物学报, 2016, 36(2): 200-205. (Shi W F, Xie Z M, Yang L M, et al. Identification of single seed vigor of spring wheat based on near-infrared spectroscopy [J]. Journal of Triticeae Crops, 2016, 36(2): 200-205.)
- [10] 王耐红, 张敏, 王瑞, 等. 甘蓝型油菜硫苷组分近红外检测模型的建立[J]. 西南大学学报(自然科学版), 2015, 37(3): 35-41. (Wang N H, Zhang M, Wang R, et al. The construction of a near infrared inspection model for glucosinolate constituent contents in brassica napus L[J]. Journal of Southwest University (Natural Science Edition), 2015, 37(3): 35-41.)
- [11] 芮玉奎, 黄昆伦, 王为民, 等. 近红外光谱技术在检测转基因油菜籽中芥酸和硫甙上的应用研究[J]. 光谱学与光谱分析, 2006, 26(12): 2190-2192. (Riu Y K, Huang K L, Wang W M, et al. Detection of erucic acid and glucosinolate in intact rapeseed by near-infrared diffuse reflectance spectroscopy[J]. Spectroscopy and Spectral Analysis, 2006, 26(12): 2190-2192.)
- [12] 张新玉, 王颖杰, 刘若西, 等. 近红外光谱技术应用于玉米单籽粒蛋白质含量检测分析的初步研究[J]. 中国农业大学学报, 2017, 22(5): 25-31. (Zhang X Y, Wang Y J, Liu R X, et al. Application of near-infrared spectroscopy technology to analyze protein content in single kernel maize seed[J]. Journal of China Agricultural University, 2017, 22(5): 25-31.)
- [13] 姚鑫森, 卢淑雯, 解铁民, 等. 玉米子粒颜色对近红外模型预测其淀粉含量的影响[J]. 玉米科学, 2013, 21(4): 153-156. (Yao X M, Lu S W, Xie T M, et al. Effect of kernel color on corn starch content by near-infrared transmittance spectroscopy [J]. Journal of Maize Sciences, 2013, 21(4): 153-156.)
- [14] 宋乐, 王琦, 王纯阳, 等. 基于近红外光谱的单粒水稻种子活力快速无损检测[J]. 粮食储藏, 2015(1): 20-23. (Song L, Wang Q, Wang C Y, et al. Qualitative analysis of single rice seed vigor using near infrared reflection spectroscopy[J]. Grain Storage, 2015(1): 20-23.)
- [15] 宋雪健, 钱丽丽, 周义, 等. 近红外漫反射光谱技术对水稻产地溯源的研究[J]. 农产品加工, 2017(5): 13-15. (Song X J, Qian L L, Zhou Y, et al. Study on the origin of rice by near infrared diffuse reflectance spectroscopy [J]. Farm Products Processing, 2017(5): 13-15.)