



基于混合样品集的大豆种子蔗糖含量近红外光谱定标模型建立

丛媛媛¹, 刘浩², 衣志刚², 王博², 刘念析², 陈亮²

(1. 吉林省农业科学院 农业经济与信息研究所, 吉林 长春 130033; 2. 吉林省农业科学院 大豆研究所/大豆国家工程研究中心, 吉林 长春 130033)

摘要:为快速测定原料大豆的蔗糖指标, 及时为调整豆制品生产工艺参数或原料配比提供数据支撑, 以满足快速准确鉴定高蔗糖大豆育种新品系的需要, 本研究建立基于混合样品集的大豆种子蔗糖含量近红外光谱定标模型。将天然和人工添加蔗糖的大豆样品结合起来作为混合参试样品集, 先利用近红外光谱分析仪采集光谱数据, 再通过酶比色法测定化学值, 进行光谱数据预处理, 采用偏最小二乘回归统计方法建立大豆蔗糖含量的近红外光谱定量分析模型, 并利用未知蔗糖含量的大豆种子样品对模型的预测性能进行验证。以混合参试样品集建立的近红外光谱定标模型能够在较宽的蔗糖含量范围(37.83 ~ 139.35 g·kg⁻¹)内获得较为精确的预测结果, 并将二阶导数联合多元散射校正作为光谱数据的最佳预处理方法。经外部验证, 样品的化学测定值与定标模型预测值极显著正相关, 且决定系数达0.943, 预测均方差为3.17, 相对误差为5.923%。所建立近红外光谱定标模型的预测能力较强、适用性较好, 能够有效应用于大豆种子蔗糖含量的快速准确测量。

关键词:大豆; 品质; 蔗糖含量; 近红外光谱定标模型

Establishment of NIRS Calibration Models for Sucrose Content in Soybean Based on the Combined Sample Set

CONG Yuan-yuan¹, LIU Hao², YI Zhi-gang², WANG Bo², LIU Nian-xi², CHEN Liang²

(1. Institute of Agricultural Economics and Information, Jilin Academy of Agricultural Sciences, Changchun 130033, China; 2. Soybean Research Institute, Jilin Academy of Agricultural Sciences / National Engineering Research Center for Soybean, Changchun 130033, China)

Abstract: This study established a near-infrared spectroscopy calibration model for sucrose content in soybean. It can provide data support for the rapid determination of sucrose in the raw materials of soybean, and then adjust the processing parameters or optimum material proportion of soy products. In addition, it can meet the needs of breeding high-sucrose soybean lines. In this study, natural and sucrose-added soybean samples were combined as the sample set. The spectral data was firstly collected using a near-infrared spectrometer. Then the near-infrared spectroscopy quantitative analysis model of soybean sucrose content was established by the determination of chemical value of enzyme colorimetry, spectral data preprocessing, and partial least square regression statistical method. Finally the prediction performance of the model was verified by using soybean seed samples with unknown sucrose content. In this study, the near-infrared spectroscopy calibration model based on the combined sample set could obtain more accurate prediction results in a wide range of sucrose content (37.83 – 139.35 g·kg⁻¹). The second derivative combined with multivariate scattering correction was used as the best preprocessing method for spectral data. Through external verification, the chemical determination value of the sample showed a significant positive correlation with the prediction value of the calibration model, the determination coefficient was 0.943, the prediction mean square deviation was 3.17, and the relative error was 5.923%. The established near-infrared spectroscopy calibration model has strong predictive ability and good applicability, and could be effectively applied to the rapid and accurate measurement of sucrose content of soybean seeds.

Keywords: Soybean; Quality; Sucrose content; NIRS calibration model

随着生活水平的日益提高, 人们对自身健康越来越关注, 并愈发重视增加膳食中豆制品的摄入量^[1]。低聚糖是大豆中所含可溶性糖类的总称, 主要由蔗糖、水苏糖、棉子糖、果糖、葡萄糖等构成, 其

中的蔗糖含量最高, 且其它的低聚糖组分大多属于蔗糖的衍生产物^[2-3]。已有研究证实, 大豆中蔗糖含量会显著影响豆制品风味、口感等食味特性。近年高蔗糖已被日本、韩国的育种者列为优质食用大

收稿日期: 2020-04-24

基金项目: 吉林省科技发展规划项目(20200201136JC); 中国博士后科学基金(2019M661234)。

第一作者简介: 丛媛媛(1982-), 女, 硕士, 助理研究员, 主要从事计算机数理统计分析与大豆田间试验设计研究。E-mail: dayuan_cc@163.com。

通讯作者: 陈亮(1981-), 男, 博士, 副研究员, 主要从事大豆优质遗传育种研究。E-mail: liangchen119@126.com。

豆的重要育种目标之一^[4-5]。有研究报道,大豆蔗糖含量受遗传及环境两方面作用的影响^[6-9]。美国研究者发现 241 份大豆种质的蔗糖含量变化范围为 1.6 ~ 95.4 mg·g⁻¹,且变幅及变异系数均较大^[10]。肌醇-3-磷酸合成酶基因(*MIPS1*)的突变会引起大豆蔗糖组分的大量积累,突变体株系种子中的蔗糖含量高达 75 ~ 119 mg·g⁻¹,而非突变体大豆种子中的蔗糖含量仅为 44 ~ 69 mg·g⁻¹^[11-12]。

高效液相色谱、酸解比色法和酶解比色法由于具有灵敏度高、检出限量小、干扰少等优点,在当前食品工业中的蔗糖含量测定领域获得了较大的关注^[13-14]。然而,以上方法存在操作复杂、速度慢、检测成本高等问题,不适合优质食用大豆选育过程中对种子蔗糖含量的大批量鉴定,以及豆制品加工企业对原料大豆蔗糖指标的快速测定。随着光学、计算机数据处理及化学计量等技术的发展,近红外光谱分析技术日趋完善,现已被广泛应用于大豆水份、蛋白质、氨基酸、脂肪和脂肪酸等含量的测定工作^[15-18]。然而,近红外光谱分析方法在大豆蔗糖含量检测方面的应用研究较少,原因主要有二:一是缺乏高蔗糖大豆参考样品,导致近红外预测模型的适用范围较狭窄;二是蔗糖由果糖和葡萄糖通过糖苷键相连而成,大豆中果糖和葡萄糖的存在会导致近红外预测模型的抗干扰能力较差。

如何拓宽和提高模型的预测范围和测量精准度是当前近红外光谱分析仪测定大豆糖类组分亟待解决的问题。本研究拟建立大豆蔗糖含量的近红外光谱定标模型,并在目前尚无高蔗糖参考样品的情况下,通过将天然与人工添加蔗糖的大豆样品相结合,探讨扩展定量模型应用范围、提高检测精度的可行性,为大豆制品原料豆蔗糖指标的快速测量、高蔗糖大豆优异材料的准确鉴定提供依据,加速优质食用大豆的育种及加工应用进程。

1 材料与方法

1.1 材料

供试材料为 236 份东北大豆种质资源,由吉林省农业科学院大豆研究所保存,其中包括育成品种 188 份,地方品种 33 份,国外引种材料 15 份。

2018 年春季,将参试大豆材料种植于吉林省农业科学院公主岭试验基地。秋季收获后,将 200 g 种子样品在室温下干燥 15 ~ 20 d,使其含水量达到约 8%,随后用于常规化学和近红外光谱分析。

1.2 方法

1.2.1 近红外光谱的收集 使用 NIRFIEX N-500 型近红外光谱分析仪对研磨过筛后的 236 个大豆样品进行光谱扫描,3 次生物学重复,扫描后的干涉图被记录为反射光谱。

称量纯度为 99.8% 的蔗糖粉,并添加到初始蔗糖含量为 34.75 ~ 78.28 g·kg⁻¹ 的大豆样品中,将大豆和蔗糖粉混合物在研磨机中均质化。以 236 个大豆样品为天然样品集,以添加蔗糖的样品为人工样品集,以天然样品和人工样品的组合为混合样品集,构建 3 个近红外光谱预测模型。根据 3 个模型的决定系数、均方差及所覆盖的蔗糖含量范围确定研究使用的大豆种子蔗糖含量近红外光谱定标建模的样品集。

1.2.2 蔗糖含量化学值的测定 使用酶学试剂盒(R-biopharm,德国拜发),在 β -D-果糖苷酶(β -FS)的催化下,大豆参试样品中的蔗糖被专一性地酶解为葡萄糖和果糖。葡萄糖氧化酶(GOD)在有氧条件下,催化 β -D-葡萄糖水溶液氧化,生成 D-葡萄糖酸- δ -内酯和过氧化氢。受过氧化物酶(POD)催化,过氧化氢与 4-氨基安替吡啉和苯酚生成红色醌亚胺。在波长 505 nm 处测定醌亚胺的吸光度,换算成为天然大豆参试样品的蔗糖含量。

1.2.3 光谱数据集预处理 以天然和人工添加蔗糖的大豆样品组合作为定标样品集,将以上定标样品集的近红外反射光谱和试剂盒测定获得的蔗糖含量参考数据相关联,建立带参考数据的光谱文件。数据预处理采用 NIRX 公司开发的 nirsLAB 近红外数据分析软件。首先对采集到的原始光谱数据进行归一化处理,随后采取卷积平滑法(CS)、二阶导数(2D)和多元散射校正(MSC)等方法进行光谱数据的预处理。

1.2.4 定标模型建立 采用偏最小二乘法建立大豆蔗糖含量的定标模型,选取 R_c^2 (定标模型决定系数)、 $RMSEC$ (定标模型均方差)、 R_{cv}^2 (交互验证决定系数)、 $RMSECV$ (交互验证均方差)作为评价模型预测效果优劣的参数。决定系数越大,均方差越小,模型的预测效果越好。

1.2.5 建模效果验证与评价 选取 40 份未知蔗糖含量的大豆种子样品,分别用酶比色法和已建立的近红外模型测定其种子蔗糖量进行比较验证。估算两种方法测定值的 R^2 (决定系数)、 $RMSEP$ (预测均方差)和 RE (相对误差),对模型的可靠性进行检

验。*RMSEP* 值越小, *R*² 值越高, 模型预测效果越好; 当 *RE* < 20% 时, 认为所建模型具有良好的适用性。

2 结果与分析

2.1 定标样品集的确定

236 个天然大豆样品(天然样品集)中蔗糖含量范围为 34.75 ~ 78.28 g·kg⁻¹, 均值为 55.42 g·kg⁻¹。初始蔗糖含量为 37.83 ~ 76.93 g·kg⁻¹ 的大豆样品中有 174 份, 剔除 22 个建模异常样品, 在 152 份样品中人工添加蔗糖构成人工样品集, 其蔗糖含量范围为 49.99 ~ 136.76 g·kg⁻¹, 均值为 89.25 g·kg⁻¹。

对采集到的原始光谱数据进行归一化处理后用多种方法进行光谱数据的预处理, 结果如表 1 所示, 与人工样品集相比, 混合样品集在定标和交叉验证两个方面, 均具有较高的决定系数以及较低的均方差。此外, 混合样品集所覆盖的蔗糖含量范围为 37.83 ~ 139.35 g·kg⁻¹, 显著高于天然样品集和人工样品集(表 1)。综合考量以上因素, 将混合样

品集确定为大豆种子蔗糖含量近红外光谱定标建模的样品。

2.2 光谱数据预处理方法的确定

在光谱数据解析前, 应先对光谱信号进行一些变换处理, 以减弱甚至消除非目标因素对检测信号信息的影响。由表 1 可知, 天然样品集的光谱数据由 CS 处理后的效果最好, 与其它光谱处理方法相比, 其 *RMSECV* 值最小(4.466)、*R*_{CV}² 值最大(0.936)。因此, 选择卷积平滑作为天然样品集的光谱数据预处理方法。人工样品集的光谱数据在 MSC 处理后, *RMSECV* 值最小(4.093)、*R*_{CV}² 值最大(0.986), 故以多元散射校正作为人工样品集的光谱数据预处理方法。混合样品集的光谱数据预处理方面, “2D + CS” 和 “2D + MSC” 两种光谱数据预处理方法的效果都较好, 但 “2D + MSC” 删除的数据点数明显少于 “2D + CS”(表 1)。因此, 将二阶导数联合多元散射校正作为大豆种子蔗糖含量近红外光谱定标模型的最佳预处理方法。

表 1 不同大豆蔗糖含量预测模型的建模参数比较

Table 1 Parameter comparison of sucrose content in soybean seed in different calibration models							
样品集 Sample set	蔗糖含量范围 Sucrose content range/(g·kg ⁻¹)	删除点数 Outlier number	光谱数据预处理 Spectroscopic data Pretreatment	预测模型决定 系数 <i>R</i> _C ²	预测模型均方差 <i>RMSEC</i>	交互验证决定 系数 <i>R</i> _{CV} ²	交互验证均方差 <i>RMSECV</i>
天然样品集 Original sample set	34.75 ~ 78.28	8	CS	0.901	3.725	0.936	4.466
		13	2D	0.844	5.815	0.880	5.505
		17	MSC	0.862	4.857	0.829	6.294
		6	2D + CS	0.869	4.405	0.884	5.388
		13	CS + MSC	0.879	4.019	0.867	5.903
		20	2D + MSC	0.883	3.942	0.899	4.904
人工样品集 Artificial sample set	49.99 ~ 136.76	9	CS	0.944	5.128	0.973	4.214
		11	2D	0.959	5.077	0.977	4.101
		10	MSC	0.966	3.639	0.986	4.093
		9	2D + CS	0.923	5.897	0.909	5.339
		4	CS + MSC	0.940	5.482	0.948	4.291
		5	2D + MSC	0.962	3.985	0.921	4.549
混合样品集 Combined sample set	37.83 ~ 139.35	25	CS	0.919	4.261	0.921	6.077
		22	2D	0.915	4.489	0.959	5.235
		31	MSC	0.917	4.399	0.950	5.989
		19	2D + CS	0.958	3.629	0.974	4.798
		28	CS + MSC	0.924	4.044	0.954	5.692
		11	2D + MSC	0.955	4.022	0.969	5.034

2.3 近红外定标模型的建立

将处理后的光谱数据和化学测定值数据进行偏最小二乘回归运算,建立并优化近红外光谱定标模型,其交互验证的相关图如图 1 所示,混合样品集定标模型的蔗糖含量预测范围可达 37.83 ~ 139.35 g·kg⁻¹,且具有较高的 R_{CV}^2 值(0.969),较低的 $RMSEC$ 值(5.034)(表 1),说明定标模型预测效果的精度较高且稳定性较好。

由图 2 可知,基于混合样品集的近红外预测模型为 9 000 ~ 4 000 cm⁻¹(波长 1 111 ~ 2 500 nm)波段范围内有较强的吸收峰,基本涵盖了光谱信号信息丰富的长波近红外光谱的整个区域,因此能够准确预测大豆种子的蔗糖含量,故选取此部分波段范围内的光谱作为后续构建定标模型的基础数据。

比较天然样品集与人工样品集的光谱图发现,后者在 3 个不同频率的红外线波长处会显示出额外的红外吸收信号(数据未提供),因此将 3 个异常频率的波长区统一标识在了混合样品集二阶导数换算处理过的近红外吸收光谱图中(图 2),分别位于 6 215 ~ 6 111 cm⁻¹(波长 1 609 ~ 1 636 nm)、4 851 ~ 4 739 cm⁻¹(波长 2 061 ~ 2 110 nm)和 4 281 ~ 4 185 cm⁻¹(波长 2 336 ~ 2 389 nm)3 个波段区域内,这 3 个异常谱峰处的特征值曲线形态较为松散、数据抖动也较大,说明这 3 处的光谱数据相关性稍低。

进一步对混合样品集整个近红外谱区中每个光谱点的吸光度与化学测定参比值进行相关性分析,得到二者的回归系数。为表征特定光谱区域与预测蔗糖含量的相关性,在 10 000 ~ 3 500 cm⁻¹(波长 1 000 ~ 2 857 nm)波段近红外光谱范围内,分析

基于混合样品集的每个波长吸光度值对其蔗糖含量的回归系数趋势结果,如图 3 所示,其中 7 545, 7 109,6 250,6 099,5 946和 5 304 cm⁻¹(波长分别为 1 325,1 407,1 600,1 640,1 682和 1 885 nm)波数处的回归系数峰值对于定标模型建立时谱区的选定十分重要。由于在采用偏最小二乘法筛选近红外光谱特征波长时,已剔除了以上异常谱峰所对应的波长,由人工添加蔗糖所引起 3 个异常谱峰区并未影响到定标模型建立时谱区的选择(图 2 和 3)。此外,鉴于以上异常谱峰并不存在于天然大豆样品中,以人工样品集建立的近红外光谱模型的预测性能将会较差。

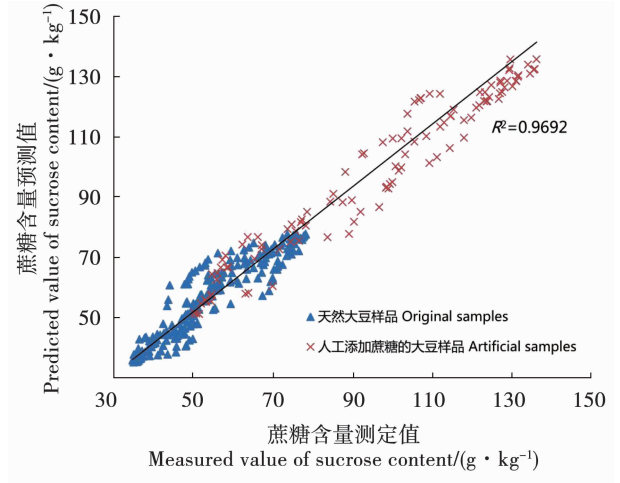


图 1 基于混合样品集的大豆蔗糖含量近红外预测值与化学实测值相关性

Fig. 1 Correlation between NIRS prediction and determination of sucrose content in soybean seed for calibration set

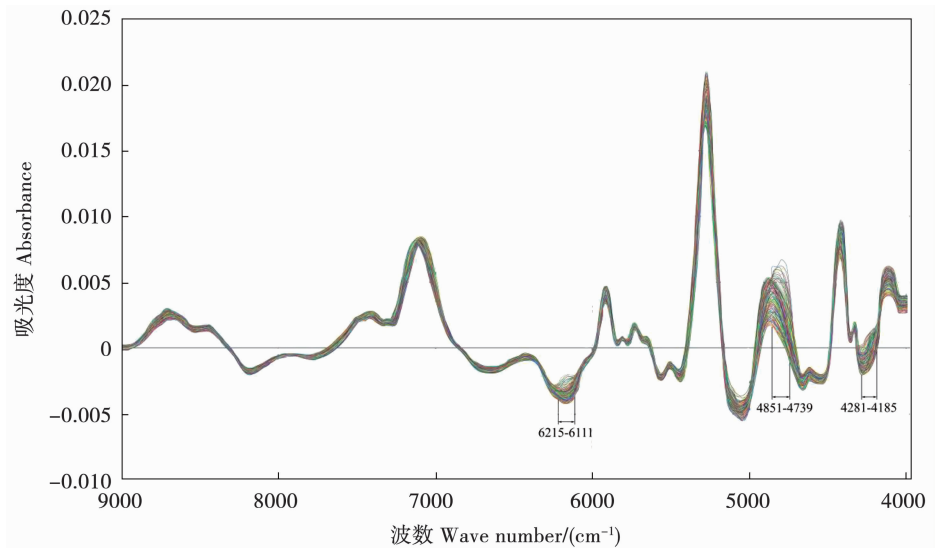


图 2 二阶导数联合多元散射校正处理后混合样品集的近红外光谱图

Fig. 2 NIRS spectra of sucrose content in soybean seed based on combined sample set after 2nd-derivative + MSC pre-treatment

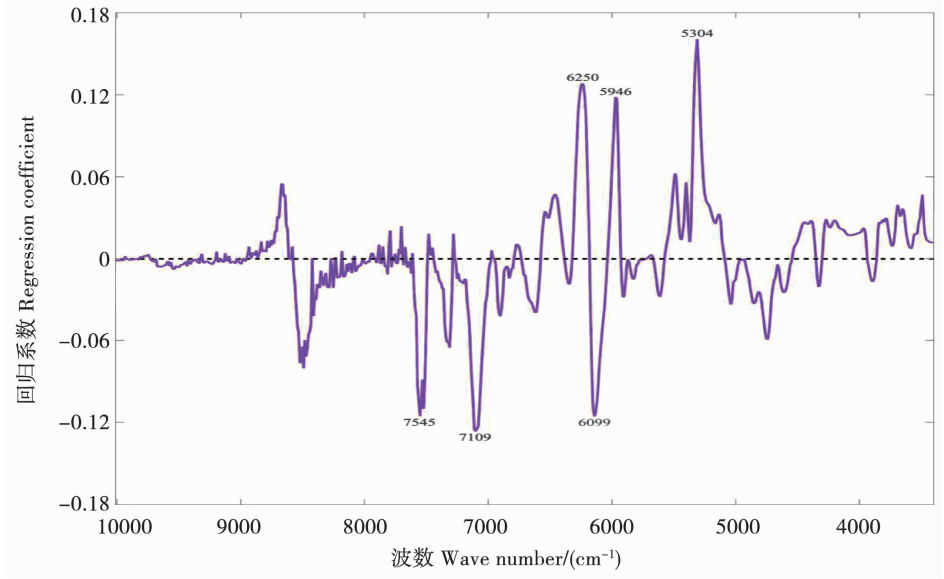


图3 基于混合样品集近红外光谱定标模型的回归系数

Fig. 3 Regression coefficient of sucrose content in soybean seed based on combined sample set

2.4 近红外定标模型检验

为验证定标模型的准确性,选取 40 份未知蔗糖含量的大豆种子样品,分别用酶比色法和建立的近红外模型测定样品的蔗糖含量值,对实测值与预测值作线性相关分析。其中,基于混合样品集的近红外定标模型的决定系数(R^2)值最大(0.943)、预测均方差(RMSEP)值最小(3.127)、相对误差(RE)值最小(5.923%);而以人工样品集建立的近红外模型的蔗糖含量预测值与实测值的决定系数仅为0.429(表2)。由此可以推测,基于混合样品集的近红外光谱定标模型对于筛选较大蔗糖含量范围的大豆样品是可行的,该模型基本可替代常规化学测定方法;而基于人工样品集的预测模型无法应用于未添加蔗糖的天然大豆样品蔗糖含量的测定。

表2 基于不同样品集建模的外部检验
Table 2 External-validation of sucrose content in soybean seed in different calibration models based on different sample sets

模型 Model	蔗糖含量 Sucrose content /(g·kg ⁻¹)	决定系数 R^2	预测均方差 RMSEP	相对误差 RE/%
天然样品集模型 Original sample set	46.36 ~ 78.39	0.776	4.292	27.541
人工样品集模型 Artificial sample set	59.67 ~ 90.86	0.429	5.856	31.388
混合样品集模型 Combined sample set	47.32 ~ 89.91	0.943	3.127	5.923
酶比色法实测 Measured value of enzyme colorimetry	48.26 ~ 90.45	-	-	-

3 讨论

大豆含有丰富的营养成分,其中的蔗糖是除蛋白质、脂肪以外的重要品质指标^[2,4]。如何在育种后代材料的筛选过程中,简捷准确地测定大豆籽粒中的蔗糖含量,一直是育种家长期探索和期待解决的问题。近年来,光学、计算机数据处理及计量化学等学科不断发展,使得近红外光谱检测技术的稳定性、实用性和准确性都得到了极大的改善与提高,特别是在大豆优质育种领域得到了广泛的认可与应用^[19-20]。为满足高蔗糖大豆新品系的快速筛选和豆制品加工原料大豆蔗糖指标高效测定的需要,本研究建立了一种基于近红外反射光谱的蔗糖含量测定方法。

近红外光谱定量分析技术作为一种间接的测量方法,其预测模型的准确性除了受到建模定标样品化学值准确度的影响,还受到样品自身代表性和变化范围的影响^[14]。一般来说,样品变化范围越大,模型的适用性越宽,但是分析结果的精度可能变差;样品变化范围较小,模型适用范围小,分析结果的精度相对较高,但适用性会变窄^[17, 21]。一直以来,高蔗糖大豆参考样品的缺乏极大地限制了近红外定量分析模型的开发。虽然本研究的参试大豆材料已具有较好的代表性,但样品中蔗糖含量的变化范围仍不够宽泛。为了扩大模型中蔗糖含量的预测范围,本研究将蔗糖粉与粉碎过筛后的大豆样品加以均质混合,获得人工添加蔗糖的大豆样品,再进一步将天然与人工添加蔗糖的大豆样品加以组合成为定标样品集,一方面可以增加建模参考样品的代表性,另一方面可以避免片面追求样品的特

殊性,更好地覆盖今后高蔗糖大豆育种材料中蔗糖成分的变化范围。

与天然大豆样品相比,人工样品中添加的蔗糖组分并未嵌入天然大豆基质中,导致在不同频率的红外线波长处,人工样品会显示出额外的近红外吸收信号。据报道,糖类化合物的近红外信号主要出现在第一倍频和合频吸收区^[21-22],与本研究观察到的特殊谱峰的位置基本一致。若将这些额外的近红外吸收信号用于定量分析模型的建立,势必会导致定量模型的预测性能较差。本研究在使用天然与人工的混合样品集构建近红外预测模型时,仅使用了与天然大豆基质样品相关的波长,而排除了包含额外近红外吸收信号的波长范围,使所建立定标模型的预测性能得到了极大地提高。此外,与天然样品集的预测模型相比,在基于混合样品集的定标模型中,其蔗糖含量的预测范围被拓宽,从而有助于模型自身精度的提高。其原因是当近红外模型待测组分的变化范围较大时,由光谱噪声、杂散光、基线漂移和平缓背景等引起的仪器系统误差被放大的程度会随之变小^[23]。

本研究将二阶导数联合多元散射校正确定为基于混合样品集的、大豆种子蔗糖含量近红外光谱定标模型的最佳预处理方法,这与许多前人的研究结果基本一致^[24-25]。已有研究证实,使用多元散射校正、二阶导数等光谱数据预处理方法可以消除固体颗粒大小、表面散射及光程变化对近红外漫反射光的影响,降低光谱漂移现象的发生概率,更重要的是能够使近红外光谱图上出现更明显的吸收峰,显著提高光谱间的差异^[13, 26]。

类似这种借助混合样品集构建近红外定量分析模型的方法已被广泛应用于鉴别原料肉注水、饮料色素超标、掺杂或伪劣白酒及其它领域^[27-29]。本研究的结果再次证实添加蔗糖的人工参考样品对于近红外光谱预测模型适用范围的拓宽以及分析精度的提高都是有帮助的。然而,鉴于被测物质的近红外光谱取决于样品的组成和结构,如何借助化学计量学方法区分多组分的近红外光谱重叠区域是今后以混合样品集建立近红外模型需要重点改进的方面。此外,很多研究已证实,当检测条件发生改变时,分析模型可能会失效^[30-31]。为了保证模型的稳健性和适应性,今后会考虑借助一些有效的波长选择方法,找到最适合模型转移的光谱波段,进而改善近红外模型在仪器间的转移效果,使模型在新的条件下同样适用,为大豆蔗糖含量近红外光谱测定技术的推广应用打下坚实基础。

4 结 论

本研究将天然和人工添加蔗糖的大豆样品结

合起来作为参试样品,先利用近红外光谱分析仪采集混合样品集的光谱数据,再通过样品化学值测定、光谱数据预处理以及偏最小二乘法回归统计方法建立了大豆蔗糖含量的近红外光谱预测模型,利用 40 份未知蔗糖含量的大豆样品对模型准确性进行的检验结果表明,确立大豆种子蔗糖含量的近红外定标模型与酶比色法化学测定值的决定系数达 0.943、预测均方差为 3.127。该方法一方面可满足大豆育种单位快速准确鉴定高蔗糖大豆育种新品系的需要;另一方面可快速测定原料大豆的蔗糖含量品质指标,为豆制品加工企业及时调整生产工艺参数或原料配比提供数据支撑。

致谢: 特别感谢吉林省农业科学院大豆研究所王曙明研究员,提供了大部分的大豆参试品种(系),并在试验设计中给予指导!感谢吉林大学赵熹博士,在试验过程中给予的帮助!

参考文献

[1]

Anderson E J, Ali M L, Beavis W D, et al. Soybean [*Glycine max* (L.) Merr.] breeding: History, improvement, production and future opportunities [M]// *Advances in Plant Breeding Strategies: Legumes*. Springer, 2019: 431-516.

[2]

邱红梅,王曙明,于妍,等. 大豆低聚糖种质筛选及遗传研究进展[J]. 大豆科学, 2014, 33 (5): 768-772. (Qiu H M, Wang S M, Yu Y, et al. Research progression germplasm screening and genetic with oligosaccharides of soybean [J]. Soybean Science, 2014, 33 (5): 768-772.)

[3]

李菁华,张明聪,金喜军,等. 高油型和高蛋白型大豆鼓粒期的糖分积累规律[J]. 大豆科学, 2017, 36 (1): 68-73. (Li J H, Zhang M C, Jin X J, et al. Sugar accumulation rule of high oil and high protein soybean during the seed-filling period [J]. Soybean Science, 2017, 36 (1): 68-73.)

[4]

孙明明,王萍,李智媛,等. 大豆活性成分研究进展[J]. 大豆科学, 2018, 37 (6): 975-983. (Sun M M, Wang P, Li Z Y, et al. Research progress of soybean active ingredients [J]. Soybean Science, 2018, 37 (6): 975-983.)

[5]

Hagely K B, Jo H, Kim J H, et al. Molecular-assisted breeding for improved carbohydrate profiles in soybean seed[J]. *Theoretical and Applied Genetics*, 2020, 133 (4): 1189-1200.

[6]

Lee J S, Kim S M, Kang S. Fine mapping of quantitative trait loci for sucrose and oligosaccharide contents in soybean [*Glycine max* (L.) Merr.] using 180 K Axiom® SoyaSNP genotyping platform [J]. *Euphytica*, 2016, 208 (1): 195-203.

[7]

Dobbels A A, Michno J M, Campbell B W, et al. An induced chromosomal translocation in soybean disrupts a *KASI* ortholog and is associated with a high-sucrose and low-oil seed phenotype[J]. *Agronomy and Plant Genetics*, 2017, 7 (4): 1215-1223.

[8]

Matei G, Woyann L G, Meneguzzi C, et al. Profiling and genotype × environment interactions of seed sugar contents in Brazilian soybean genotypes[J]. *Euphytica*, 2017, 213 (9): 203.

[9]

Patil G, Vuong T D, Kale S, et al. Dissecting genomic hotspots underlying seed protein, oil, and sucrose content in an

- interspecific mapping population of soybean using high-density linkage mapping [J]. *Plant Biotechnology Journal*, 2018, 16 (11): 1939-1953.
- [10] Medic J, Atkinson C, Hurburgh C R. Current knowledge in soybean composition [J]. *Journal of the American Oil Chemists' Society*, 2014, 91(3): 363-384.
- [11] Maupin L M, Rosso M L, Rainey K M. Environmental effects on soybean with modified phosphorus and sugar composition [J]. *Crop Science*, 2011, 51(2): 642-650.
- [12] Rosso M L, Bursleson S A, Maupin L M, et al. Development of breeder-friendly markers for selection of *MIPSI* mutations in soybean [J]. *Molecular Breeding*, 2011, 28(1): 127-132.
- [13] 王潇潇, 李军涛, 孙祥丽, 等. 近红外反射光谱快速测定四种大豆制品中寡糖含量的研究 [J]. *光谱学与光谱分析*, 2018, 38(1): 58-61. (Wang X X, Li J T, Sun X L, et al. Study on rapid determination of oligosaccharides in soybean products by near-infrared reflectance spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2018, 38(1): 58-61.)
- [14] 杨泉女, 周权驹, 王蕴波, 等. 甜玉米葡萄糖、果糖和蔗糖含量近红外反射光谱模型构建 [J]. *中国农业科技导报*, 2018, 20(1): 137-146. (Yang Q N, Zhou Q J, Wang Y B, et al. Establishment of NIRS models for the content of glucose, fructose and sucrose in sweet corn [J]. *Journal of Agricultural Science and Technology*, 2018, 20(1): 137-146.)
- [15] 薛忠财, 高辉远. 反射光谱技术在监测大豆叶片光化学效率和水分状况中的应用 [J]. *大豆科学*, 2016, 35(3): 436-441. (Xue Z C, Gao H Y. Application of spectral reflectance technique on monitoring the photochemical efficiency and water condition in soybean leaves [J]. *Soybean Science*, 2016, 35(3): 436-441.)
- [16] 王铎, 张力, 范素杰, 等. 大豆主要脂肪酸含量近红外模型的建立 [J]. *大豆科学*, 2017, 36(2): 295-299. (Wang D, Zhang L, Fan S J, et al. Establishment of NIRS model for the main fatty acids in soybean [J]. *Soybean Science*, 2017, 36(2): 295-299.)
- [17] Cortés V, Blasco J, Aleixos N, et al. Monitoring strategies for quality control of agricultural products using visible and near-infrared spectroscopy: A review [J]. *Trends in Food Science and Technology*, 2019, 85: 138-148.
- [18] 王丽萍, 陈文杰, 赵兴忠, 等. 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 [J]. *大豆科学*, 2019, 38(2): 280-285. (Wang L P, Chen W J, Zhao X Z, et al. Rapid determination of crude protein and crude oil content of soybean based on near infrared diffuse reflectance spectroscopy [J]. *Soybean Science*, 2019, 38(2): 280-285.)
- [19] 王丽萍, 赵兴忠, 陈文杰, 等. 油菜籽含油量傅里叶变换近红外模型的修正 [J]. *激光生物学报*, 2013, 22(1): 44-47. (Wang L P, Zhao X Z, Chen W J, et al. Optimization of quantitative analysis model for oil content in rapeseed with Fourier transform near infrared spectroscopy [J]. *Acta Laser Biology Sinica*, 2013, 22(1): 44-47.)
- [20] 王翠秀, 曹见飞, 顾振飞, 等. 基于近红外光谱大豆蛋白质、脂肪快速无损检测模型的优化构建 [J]. *大豆科学*, 2019, 38(6): 968-976. (Wang C X, Cao J F, Gu Z F, et al. Rapid nondestructive test of soybean protein and fat by near infrared spectroscopy combined with different model methods [J]. *Soybean Science*, 2019, 38(6): 968-976.)
- [21] Wee C D, Hashiguchi M, Ishigaki G, et al. Evaluation of seed components of wild soybean (*Glycine soja*) collected in Japan using near-infrared reflectance spectroscopy [J]. *Plant Genetic Resources*, 2018, 16(2): 94-102.
- [22] Choung M G. Determination of sucrose content in soybean using near-infrared reflectance spectroscopy [J]. *Journal of the Korean Society for Applied Biological Chemistry*, 2010, 53(4): 478-484.
- [23] 惠文凯, 王益, 陈晓阳. 麻疯树种子含油量近红外光谱定标模型的建立 [J]. *北京林业大学学报*, 2018, 40(1): 1-7. (Hui W K, Wang Y, Chen X Y. Calibration model building for investigating seed oil content of *Jatropha curcas* using near infrared spectroscopy [J]. *Journal of Beijing Forestry University*, 2018, 40(1): 1-7.)
- [24] 张娟娟, 熊淑萍, 时雷, 等. 基于近红外光谱分析的土壤全氮含量估测研究 [J]. *土壤*, 2015, 47(4): 653-657. (Zhang J J, Xiong S P, Shi L, et al. Soil nitrogen content prediction with near infrared spectroscopy [J]. *Soils*, 2015, 47(4): 653-657.)
- [25] 刘燕德, 肖怀春, 邓清, 等. 柑桔黄龙病近红外光谱无损检测 [J]. *农业工程学报*, 2016, 32(14): 202-208. (Liu Y D, Xiao H C, Deng Q, et al. Nondestructive detection of citrus greening by near infrared spectroscopy [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2016, 32(14): 202-208.)
- [26] 芦兵, 孙俊, 杨宁, 等. 基于 SAGA-SVR 预测模型的水稻种子水分含量高光谱检测 [J]. *南方农业学报*, 2018, 49(11): 2342-2348. (Lu B, Sun J, Yang N, et al. Hyperspectral detection for moisture in rice seeds by SAGA-SVR prediction model [J]. *Journal of Southern Agriculture*, 2018, 49(11): 2342-2348.)
- [27] 杨志敏, 丁武, 张瑶. 应用近红外技术快速鉴别原料肉注水的研究 [J]. *食品研究与开发*, 2012, 33(5): 118-120, 128. (Yang Z M, Ding W, Zhang Y. Study on discrimination of raw meat and water-injected meat based on near-infrared spectroscopy and artificial neural network model [J]. *Food Research and Development*, 2012, 33(5): 118-120, 128.)
- [28] 杨国强, 张淑娟, 赵艳茹. 基于近红外透射光谱的汾阳王酒快速鉴别 [J]. *农业机械学报*, 2013, 44(S1): 189-193. (Yang G Q, Zhang S J, Zhao Y R. Fast discrimination of adulterated Fenyangwang wine based on near infrared spectroscopy [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2013, 44(S1): 189-193.)
- [29] 田晶, 李巧玲. 近红外光谱法快速检测饮料中的食用合成色素 [J]. *食品科学*, 2018, 39(2): 293-297. (Tian J, Li Q L. Rapid determination of synthetic pigments in fruit-flavored beverage by near infrared spectroscopy [J]. *Food Science*, 2018, 39(2): 293-297.)
- [30] 曹晓兰, 邓梦洁, 汪佩佩. 基于 PLSR 的苕麻叶片含水量估测模型建立及优化 [J]. *激光生物学报*, 2018, 27(5): 467-472. (Cao X L, Deng M J, Wang P P. Building and optimizing of the PLSR based estimation model ramie leaf's water content [J]. *Acta Laser Biology Sinica*, 2018, 27(5): 467-472.)
- [31] 张津源, 张德贤, 张苗. 基于连续投影算法的小麦蛋白质含量近红外光谱预测分析 [J]. *江苏农业学报*, 2019, 35(4): 960-964. (Zhang J Y, Zhang D X, Zhang M. Prediction and analysis of wheat protein content by near-infrared spectroscopy based on successive projections algorithm [J]. *Jiangsu Journal of Agricultural Sciences*, 2019, 35(4): 960-964.)