

利用近红外光谱技术检测掺假豆浆

李东华, 潘园园, 张 卉^{*}

(沈阳化工大学制药与生物工程学院, 沈阳 110142)

摘 要: 为了对豆乳内在营养指标及掺假豆乳进行快速检测, 试验运用近红外光谱技术, 利用偏最小二乘法进行回归分析, 分别建立 83 个真伪豆浆样品的蛋白质和总固形物含量定标模型, 并对模型的预测性能进行分析。结果表明: 选取主成分数为 12 和 14, 蛋白质和总固形物含量的近红外光谱预测值与化学实测值之间的相关系数 R 分别为 0.9756 和 0.9489, 校正均方根误差分别为 0.186 和 0.175, 预测集样品的预测值和实测值之间的残差值均较小、接近零, 残差之和分别为 -0.074 和 -1.191, 说明建立的定标模型可以准确预测豆浆中蛋白质和总固形物含量, 且预测性能较好; 通过对预测集样品的预测值与豆浆行业标准规定值相比较, 确定预测集样品中掺假豆浆的正确判别率为 100%, 说明建立的蛋白质和总固形物含量定标模型可以应用于掺假豆浆的判别检测, 且判别结果准确率高。本试验表明利用近红外光谱技术可实现对豆浆主要品质指标的快速无损检测, 也可准确进行真伪豆浆的快速判别, 本检测方法可为豆乳行业健康持续发展提供一定的技术支撑。

关键词: 近红外光谱; 蛋白质; 无损检测; 豆浆; 掺假检测; 总固形物

doi: 10.3969/j.issn.1002-6819.2014.03.031

中图分类号: TS207.3

文献标志码: A

文章编号: 1002-6819(2014)-03-0238-05

李东华, 潘园园, 张 卉. 利用近红外光谱技术检测掺假豆浆[J]. 农业工程学报, 2014, 30(3): 238—242.

Li Donghua, Pan Yuanyuan, Zhang Hui. Adulteration detection of soymilk based on near-infrared spectroscopy[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2014, 30(3): 238—242. (in Chinese with English abstract)

0 引 言

豆浆是一种营养丰富植物蛋白饮品, 以豆浆的方式补充蛋白可以使人的抗病能力增强, 延缓衰老^[1-3]。近几年豆浆掺假现象日益严重, 这不仅有碍于豆浆行业的发展, 而且掺入的非豆乳成分会危及到了消费者的人身健康, 因此建立一种快速检测豆浆内在营养成分含量的方法是十分必要的。

近红外光谱分析技术主要是通过物质内部成分对近红外光的特征吸收实现定性和定量分析, 是近年来发展起来的一种高效快速的现代分析技术, 在牛乳、果酒、蜂蜜、豆油等液体食品的内在品质分析以及掺假成分判别中已有广泛的应用^[4-8]; 目前利用近红外技术对牛乳类的研究比较广泛, 而对豆乳等植物蛋白饮料未见报道, 对牛乳类的研究主要是营养指标的检测以及掺假成分的定性和定量

判别, 例如: Melfsen A 等^[9]利用近红外光谱技术建立了 1 610 个原料乳的脂肪、蛋白、乳糖含量的定标模型, 相关系数分别为 0.998、0.94 和 0.73, 但对原料乳中尿素含量检测效果差于前 3 个指标; 褚莹等^[10-11]分别对纯山羊奶和掺有奶油、还原奶的掺假山羊奶的近红外光谱进行了定性和定量研究, 采用 MLP 神经网络法建立校正模型的正判率达到 99.4%, 验证集的正判率达到 100%, 证明采用近红外光谱技术结合适当的化学计量学方法可以实现羊奶掺假检测的快速无损鉴别。

蛋白质和总固形物含量是评价豆浆品质的重要指标^[12-14], 本试验考虑市场上常见的豆浆掺假手段, 制备了 31 个掺假豆浆和 62 个合格豆浆样品, 并随机从市场购买 14 个样品, 利用近红外光谱分析技术, 进行真伪豆浆内在成分的定量检测, 建立豆浆的蛋白质和总固形物含量的定标模型, 并对模型的性能进行预测。

1 材料与方法

1.1 材料与仪器

制备豆浆的大豆原料购买于沈阳富友种子有限公司, 收集的豆浆样品购买于沈阳市内多个农贸市场, 豆浆香精、增稠剂、色素等均购买于沈阳南二添加剂市场。试验一共制备 107 个样品, 83 个样

收稿日期: 2013-09-10 修订日期: 2013-12-10

基金项目: 辽宁省教育厅项目(2009A582); 辽宁省自然科学基金项目(2013020060)

作者简介: 李东华(1981—)女, 博士, 讲师, 从事食品质量控制和无损检测方向的教学与科研工作。沈阳 沈阳化工大学制药与生物工程学院, 110142。Email: lidonghua2004168@163.com

※通信作者: 张 卉(1968—), 女, 博士, 副教授, 主要从事食品生物技术方面的教学与科研工作。沈阳 沈阳化工大学制药与生物工程学院, 110142。Email: zhuhui-99@163.com

品作为训练集样品, 24 个样品作为预测集样品, 其中训练集样品包括 57 个纯豆浆和 26 个掺假豆浆, 预测集样品包括纯豆浆和掺假豆浆各 5 个, 随机购买的 14 个豆浆样品; 将样品按照顺序依次编号, 供光谱和化学值的测定。

Purespect 型近红外透射光谱仪(日本杂贺技术研究所); 紫外可见双光束扫描分光光度计(日本日立公司); 电热恒温鼓风干燥箱(上海森信实验仪器有限公司); 电子天平(北京赛多利斯仪器系统有限公司)。

1.2 方法

1.2.1 豆浆的制作

通过市场调查和有关掺假豆浆的报道, 掺假豆浆主要是通过向纯豆浆中加入大量的水、增稠剂、豆浆香精、色素以及甜味剂等勾兑而成。本试验先按照豆浆的典型制作工艺^[15]制作出蛋白质含量符合标准^[13-14](蛋白质质量分数 $\geq 2\%$)的纯豆浆, 在纯豆浆的基础上根据现有的掺假手段, 制备掺假豆浆。最终确定用于制备掺假豆浆的主辅料分别为: 纯豆浆(豆水质量比为 1:10)、豆浆香精、豆浆增稠剂、甜蜜素、日落黄。

1.2.2 蛋白质和总固形物含量的测定

采用考马斯亮蓝 G-250 法进行蛋白质含量的测定^[16-17], 对豆浆中总固形物含量的测定采用常压干燥法, 先测定豆浆中水分再折算成总固形物的含量^[13-14]。

1.2.3 近红外透射光谱的测定

采用日本杂贺技术研究所提供的 Purespect 型近红外光谱透射仪获取豆浆的近红外光谱, 准确量取 50 mL 的豆浆样品倒入烧杯中, 然后将烧杯放入样品杯内, 随操作台转动进行光谱采集, 在完全遮光的实验室内进行样品光谱的采集; 仪器扫描波长范围为 643.26~954.15 nm, 采点间隔为 1.29 nm, 进行全光谱采集, 每个样品扫描 3 次。

1.2.4 光谱数据的处理

采集的光谱数据利用 ChangeCR 软件中心化处理后存入 Excel 中, 根据比尔定律将光谱数据转化为光密度值, 再导入 Unscrambler 6.1 软件对光谱数据进行处理。利用 Unscrambler 6.1 软件的 Savitzky-Golay 对光谱进行平滑处理, 采用二阶导数法进行光谱的背景扣除及信号的分离, 再利用偏最小二乘法(partial least squares, PLS)对光谱数据进行统计处理并建立模型^[18-20]。

2 结果与分析

2.1 蛋白质含量和总固形物含量实测值的分析

按照豆浆的典型制备工艺进行制备纯豆浆, 通过添加水、增稠剂、香精等配制掺假豆浆, 共收集 83 个训练集样品和 24 个预测集样品。对训练集和预测集样品的蛋白质和总固形物含量进行测定, 然后进行统计, 具体结果见表 1。

表 1 训练集和预测集样品的蛋白质和总固形物含量的统计结果

Table 1 Statistics results of protein and total solids content for calibration and prediction samples

项目 Item	训练集样品 Calibration samples		预测集样品 Prediction samples	
	蛋白质质量分数 Protein content/%	总固形物质量分数 Total solids content/%	蛋白质质量分数 Protein content/%	总固形物质量分数 Total solids content/%
变化范围 Range of variation	0.9694~3.9054	7.551~5.0716	1.106~3.4602	7.399~5.0716
平均值 Mean	2.4723	6.3438	1.9148	5.8293
标准差 Standard deviation	0.8526	0.5583	0.6642	0.6438
变异系数 Coefficient of variation	0.3449	0.0880	0.3469	0.1104

注: 训练集样品 $n=83$, 预测集样品 $n=24$ 。

Note: Calibration samples, $n=83$, prediction samples, $n=24$.

由表 1 可知, 训练集和预测集样品的蛋白质和总固形物含量基本覆盖了高、中、低水平, 并且预测集样品的蛋白质和总固形物含量范围都在训练集范围内, 说明样品具有一定的代表性, 可准确地进行模型的建立。

2.2 模型的建立及预测性能的分析

近红外光谱信号的处理方法有变换、平滑、相关、卷积、微分和积分等。导数光谱既可以消除基线漂移或平缓背景干扰的影响, 又可以提供比原来光谱更高的分辨率和更清晰的光谱轮廓变化^[21-22]。本试验利用 Unscrambler 6.1 软件对豆浆的近红外原始吸收光谱进行平滑和二阶导数处理后得到了信

号清晰的光谱图, 结果见图 1, 随后选用偏最小二乘法进行回归分析, 选取主成分数为 12 和 14, 分别建立豆浆的蛋白质和总固形物含量的校正模型, 图 2 为蛋白质和总固形物含量的校正结果。

由图 2 可知, 蛋白质含量和总固形物含量的校正结果都较好, 相关系数均达到 0.9 以上, 校正均方根误差也较低, 但一个优秀的近红外模型除了模型自身相关性要较高之外, 还要具备较好的预测性能, 在实际应用中后者往往尤为重要。因此利用已建立的数学模型分别对 24 个预测集样品进行预测, 根据预测结果的可靠性来分析模型的预测性能, 表 2 为 24 个预测集样品的预测结果。

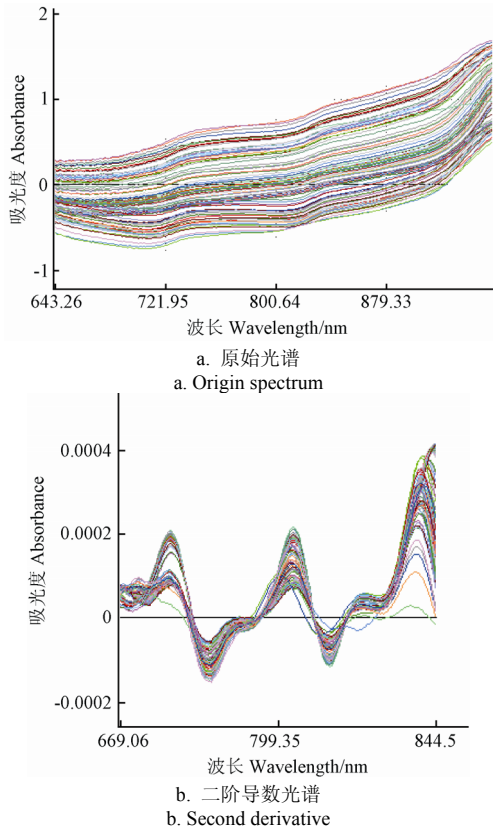


图 1 所有训练集样品的吸收光谱图

Fig.1 Absorption spectrum for all calibration soymilk samples

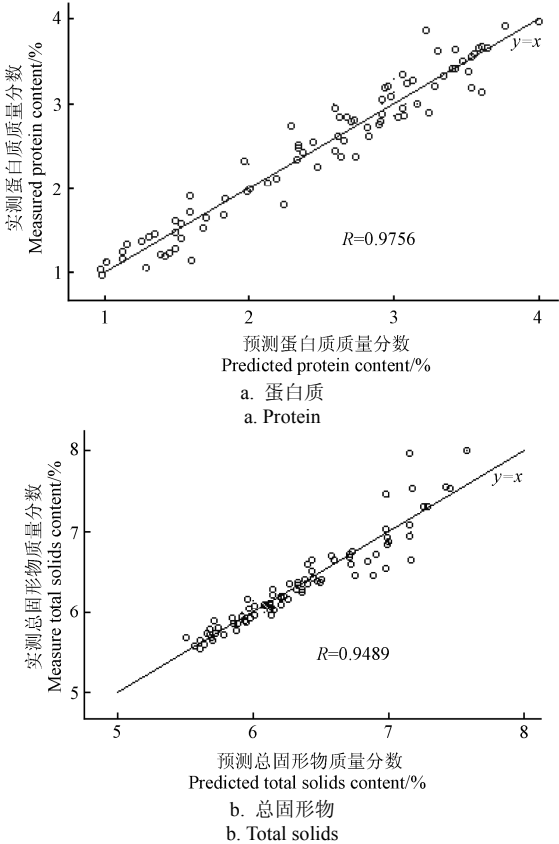


图 2 训练集样品的蛋白质、总固形物预测值与实测值对应的散点图

Fig.2 Scatter plots of predicted protein, total solids content and measured value for calibration samples

表 2 预测集样品的预测结果

Table 2 Prediction results of prediction samples

编号 No.	蛋白质质量分数 Protein content/%			总固形物质量分数 Total solids content/%		
	真实值 Actual value	预测值 Predicted value	残差 Residual error	真实值 Actual value	预测值 Predicted value	残差 Residual error
1	3.9	4.184	-0.284	7.355	8.062	-0.707
2	2.957	3.042	-0.085	7.231	7.197	0.034
3	3.223	3.121	0.102	7.259	7.359	-0.1
4	3.302	3.401	-0.099	7.247	7.858	-0.611
5	2.664	2.412	0.252	6.358	6.225	0.133
6	1.609	1.592	0.017	5.857	5.69	0.167
7	1.535	1.84	-0.305	5.904	5.91	-0.006
8	1.413	1.519	-0.106	5.749	5.955	-0.206
9	1.502	1.656	-0.154	5.793	5.908	-0.115
10	1.911	1.619	0.292	5.832	5.884	-0.052
11	1.387	1.465	-0.078	5.798	5.831	-0.033
12	1.568	1.449	0.119	5.867	5.9	-0.033
13	1.664	1.81	-0.146	5.899	6.117	-0.218
14	1.658	1.689	-0.031	5.732	6.003	-0.271
15	1.623	1.41	0.213	6.129	5.901	0.228
16	1.708	1.678	0.03	6.026	5.985	0.041
17	1.833	1.726	0.107	5.89	5.806	0.084
18	1.774	1.774	0	5.845	5.941	-0.096
19	1.547	1.368	0.179	5.784	5.694	0.09
20	1.297	1.549	-0.252	5.89	5.687	0.203
21	1.139	1.251	-0.112	5.655	5.564	0.091
22	1.106	1.095	0.011	5.59	5.472	0.118
23	3.46	3.541	-0.081	7.312	7.26	0.052
24	2.716	2.379	0.337	6.169	6.153	0.016
求和 Sum			-0.074			-1.191

注：1~5 号为自制的合格样品；6~10 号为自制的掺假样品；11~24 为购买的样品。
Note: No1~ No5 were normal samples confecting in lab; No6~ No10 were adulteration samples confecting in lab No11~ No24 were samples collecting from market.

由表 2 可看出，2 个模型的预测残差均较小，分析偏差和，蛋白质含量实测和预测值的残差之和（-0.074）接近于 0，说明建立的蛋白质含量定标模型预测性能较好，而总固形物含量实测和预测值的残差之和为-1.191，误差较大，造成这一现象的原因可能是由于水分蒸发时，部分可溶性成分也随着挥发，导致结果存在偏大误差。

根据豆浆行业标准^[13-14]确定，当样品中的蛋白质质量分数小于 2%，总固形物质量分数小于 6%时，样品不达标，可判定为不合格，本试验对预测集样品的预测值与标准规定值相比较，发现 6~22 号的蛋白质质量百分含量不达标，这些样品中除了 13、14 号外，总固形物质量含量也不达标，试验最终以蛋白质质量百分含量为主要判别指标，判定 6~22 号均为不合格产品，与样品采集实际情况和化学测定真实值对照，掺假豆浆的正确判别率为 100%。

3 结论与讨论

本试验建立了豆浆中蛋白质和总固形物含量的近红外透射光谱的数学模型, 并对其预测性能进行了分析, 结果显示所建立的定标模型相关系数均达到 0.9 以上, 校正均方根误差也较低; 进而对预测集样品进行了预测, 通过对预测相关系数、预测均方根误差以及残差值的分析表明: 模型能够较准确预测豆浆预测集样品的蛋白质和总固形物的含量, 预测性能较好, 且对掺假豆浆地正确判别率达到 100%。

由于市场上的掺假豆浆层出不穷, 本试验配制的掺假豆浆种类较少, 因此建立的定标模型的性能有一定的局限性, 在实际应用中应不断扩充掺假豆浆的种类, 来提高模型的预测精度和适应范围; 另外近红外检测模型的建立还受其他因素的影响, 例如样品的表面状态(均一性、稳定性), 仪器的状态、环境条件等, 因此应该对各影响因素进行综合分析, 从而提高模型的适用性。

[参 考 文 献]

- [1] 李里特. 大豆加工与利用[M]. 化学工业出版社, 2003.
- [2] 贺丰霞, 芮汉明. 豆浆对糖液清净效果的研究[J]. 食品工业科技, 2009, 30(4): 255—257.
He Fengxia, Yin Hanming. Research on lustration effect of soybean milk on syrup[J]. Science and Technology of Food Industry, 2009, 30(4): 255—257. (in Chinese with English abstract)
- [3] 梁晓丽, 许钰麒, 范志红. 豆浆对慢性病的预防与控制作用研究进展[J]. 中国食物与营养, 2010(10): 73—76.
Liang Xiaoli, Xu Yuqi, Fan Zhihong. Research advancement of soy milk's prevention and treatment effect on chronic diseases[J]. Food and Nutrition in China, 2010(10): 73—76. (In Chinese with English abstract)
- [4] 刘燕德, 万常斓. 芝麻油掺伪的近红外透射光谱检测技术[J]. 农业机械学报, 2012, 43(7): 136—139.
Liu Yande, Wan Changlan. Analysis of sesame oil adulteration using near infrared transmission spectroscopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(7): 136—139. (in Chinese with English abstract)
- [5] 屠振华, 朱大洲, 籍保平, 等. 基于近红外光谱技术的蜂蜜掺假识别[J]. 农业工程学报, 2011, 27(11): 382—387.
Tu Zhenhua, Zhu Daozhou, Ji Baoping, et al. Adulteration detection of honey based on near-infrared spectroscopy[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2011, 27(11): 382—387. (in Chinese with English abstract)
- [6] 钟珍珍, 张立国, 张鑫, 等. 近红外技术识别掺假原奶的可行性研究[J]. 计算机与应用化学, 2010, 27(12): 1691—1693.
Zhong Zhenzhen, Zhang Liguang, Zhang Xin, et al. The feasibility study of discriminating adulteration raw milk by near infrared technology[J]. Computers and Applied Chemistry, 2010, 27(12): 1691—1693. (in Chinese with English abstract)
- [7] Lopez-calleja, Gonzalez, Fajardov. Application of an Indirect Elisa and a PCR technique for detection of cows, milk in sheeps and goat, milk cheeses[J]. International Dairy Journal, 2007, 17(1): 87—93.
- [8] Daniel Kelly, Cristina Petisco, Gerard Downey. Potential of near infrared transreflectance spectroscopy to detect adulteration of Irish honey by beet invert syrup and high fructose corn syrup[J]. Journal of Near Infrared Spectroscopy, 2006, 14(2): 139—146.
- [9] Melfsen A, Hartung E, Haeussermann A. Accuracy of milk composition analysis with near infrared spectroscopy in diffuse reflection mode[J]. Biosystems Engineering, 2012, 112(3): 210—217.
- [10] 褚莹, 丁武, 丁丁. 近红外光谱法用于掺假羊奶的快速无损鉴别[J]. 中国乳品工业, 2012, 40(2): 51—53.
Chu Ying, Ding Wu, Ding Ding. Rapid and nondestructive identification of adulteration goat's milk by near infrared spectroscopy[J]. China Dairy Industry, 2012, 40(2): 51—53. (in Chinese with English abstract)
- [11] 褚莹. 基于近红外光谱技术实现羊奶新鲜度检测及掺假检测的研究[D]. 陕西: 西北农林科技大学, 2012.
Zhu Ying. Study on Discrimination of Freshness of Goat's Milk and Adulteration Goat's Milk by Near Infrared Spectroscopy Technology[D]. Shanxi, Northwest A&F University, 2012. (in Chinese with English abstract)
- [12] 九阳股份有限公司. LS/T 3241-2012 豆浆用大豆[S]. 北京: 中国标准出版社, 2012.
- [13] 中国轻工业联合会. QB/T 2132-2008 植物蛋白饮料豆奶(豆浆)和豆奶饮料[S]. 北京: 中国轻工业出版社, 2008.
- [14] 中华人民共和国商务部. SB/T 10633-2011 豆浆类[S]. 北京: 中国标准出版社, 2012.
- [15] 马儒珍, 刘金福. 食品工艺学实验技术[M]. 北京: 化学工业出版社, 2011.
- [16] 南亚, 李宏高. 考马斯亮蓝 G-250 法快速测定牛乳中的蛋白质[J]. 饮料工业, 2007, 12(10): 41—44.
Nan Ya, Li Honggao. Quick determination of protein in fresh milk with Coomassie Brilliant Blue G-250[J]. Detection and Analysis, 2007, 12(10): 41—44. (in Chinese with English abstract)
- [17] 孙士青, 王少杰, 李秋顺, 等. 考马斯亮蓝法快速测定乳品中蛋白质含量[J]. 山东科学, 2011, 24(6): 53—55.
Sun Shiqing, Wang Shaojie, Li Qiushun, et al. Coomassie brilliant blue method based protein content determination in milk[J]. Shandong Science, 2011, 24(6): 53—55. (in Chinese with English abstract)
- [18] 严衍禄, 赵龙莲, 韩东海, 等. 近红外光谱分析基础与应用[M]. 北京: 中国轻工业出版社, 2005: 142—145.
- [19] 李水芳, 单杨, 朱向荣, 等. 近红外光谱结合化学计量学方法检测蜂蜜产地[J]. 农业工程学报, 2011, 27(8): 350—354.
Li Shuifang, Dan Yang, Zhu Xiangrong, et al. Detection of geographical origin of honey using near-infrared spectroscopy and chemometrics[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions

- of the CSAE), 2011, 27(8): 350—354. (in Chinese with English abstract)
- [20] Chen Lanzhen, Wang Jiahua, Ye Zhihua, et al. Classification of Chinese honeys according to their floral origin by near infrared spectroscopy[J]. Food Chemistry, 2012, 135(2): 338—342.
- [21] Carlos E B, Olivieri A C. Net analytic preprocessing: a new and versatile multivariate calibration technique analysis of mixture of rubber antioxidants by NIR spectroscopy[J]. Near Infrared Spectroscopy, 2001(9): 245—249.
- [22] 杨燕, 聂鹏程, 杨海清, 等. 基于可见-近红外光谱技术的蜜源快速识别方法[J]. 农业工程学报, 2010, 26(3): 238—242.
- Yang Yan, Nie Pengcheng, Yang Haiqing, et al. Rapid recognition method of nectar plant based on visible-near infrared spectroscopy[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2010, 26(3): 238—242. (in Chinese with English abstract)

Adulteration detection of soymilk based on near-infrared spectroscopy

Li Donghua, Pan Yuanyuan, Zhang Hui^{*}

(Department of Pharmaceutical Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China)

Abstract: In order to rapidly detect the internal nutritive index and discriminate adulteration soymilk, the near infrared transmission spectrometer such as Purespect was used to obtain spectrums for 83 unadulterated and adulterated soymilk samples. The spectral scanning procedure was conducted in dark room, 643.26-954.15 nm wavelength range was chosen, scanning wavelength interval was 1.29 nm. Each soymilk sample was scanned three times. Pure soymilks were made according to the regulations in soymilk products industry standard SB/T 10633-2011. A lot of water, essence, thickening agent, food sunset yellow pigment were added to unadulterated soymilk samples in order to obtain adulterated soymilk samples. In this study, 31 adulteration samples and 62 unadulterated samples were prepared in the processing laboratories. 14 soymilk samples were gathered from the market. All samples were used to scan the spectrum and determining chemical composition. The experimental results indicated that smooth lines and clear spectrogram were obtained using Savitzky-Golay and the second derivative method. Chemometrics method of partial least squares (PLS) was used to the model calibration for protein and total solid content in samples. The correlation coefficient of predicted value and measured value of protein and total solid content for soymilk calibration samples were 0.9756 and 0.9489 respectively. The correction of root mean square error for soymilk calibration samples were 0.186 and 0.175 respectively. 12 and 14 was selected for principal component number respectively. 24 prediction samples were prepared for analyzing predictive capability. The results indicated that the residual values of predicted value and measured value for soymilk prediction samples were small and close to zero. The distribution of residual was uniform for both sides of zero line. The residual sums between predicted and measured protein and total solid content values were -0.074 and -1.191 respectively. The results verified that the calibration models could accurately predict the protein and total solid content for soymilk samples. According to the standard of soymilk industry, the internal nutrition of sample was satisfied and the sample was disqualified when the protein mass fraction in soymilk samples was less than 2%, the total solids mass fraction in soymilk samples was less than 6%. Through comparing the predicted values of prediction samples with the regulation values by soymilk industry standard, No. 6-22 samples were disqualified. The resolution of individually adulterated soymilk from all prediction samples was 100% based on the practical sample collection conditions and measured value by chemical methods. The results verified that the protein and total solid content calibration models were capable of discriminating the adulteration soymilk. This experiment indicated that rapidly detect the major quality index and discriminate adulteration soymilk were achieved based on NIR spectra. This detection method can be used to support for the healthy and abidingly development of soymilk industry.

Key words: near infrared spectroscopy; proteins; nondestructive examination; soymilk; adulteration detection; total solids content

(责任编辑: 刘丽英)