

近红外光谱法检测 婴幼儿营养米粉蛋白质含量

于清丽^{1,2}, 石磊^{1,2}, 耿响^{1,2,*}, 刘秀红^{1,2}

(1. 江西出入境检验检疫局综合技术中心, 江西南昌 330038;

2. 江西省红外光谱应用工程技术研究中心, 江西南昌 330038)

摘要:以 160 个婴幼儿营养米粉为样品, 应用近红外光谱技术和偏最小二乘法建立了婴幼儿营养米粉中蛋白质含量的近红外模型。比较了 2、3、4、5、6、7、8 mm 不同装样厚度(即光程)下样品近红外光谱的稳定性, 通过样品光谱在不同波数下的标准差选定了 5 mm 为最佳的装样厚度; 比较了变量标准化、导数、多元散射校正和 9 点平滑预处理方法对模型建立的影响, 根据模型的相关系数和均方根误差确定了最佳的预处理方法, 为变量标准化、一阶导数和 9 点平滑, 预处理后的光谱数据结合国标方法测定的蛋白质含量建立了婴幼儿营养米粉中蛋白质含量的近红外定量分析模型, 定标集和验证集相关系数分别为 0.9856、0.9841, 均方根误差分别为 0.2431%、0.2456%, 对已建立的近红外光谱定量分析模型进行了外部样品验证, 预测准确度达到 96%, 验证结果证明了近红外检测技术与国标方法一致性较好, 速度快且不需要样品前处理, 在婴幼儿营养米粉蛋白质含量分析中具有广泛的应用价值。

关键词:近红外, 婴幼儿营养米粉, 蛋白质

Determination of Protein Contents in Infant Nutritious Rice Flour by Near Infrared Spectroscopy

YU Qing-li^{1,2}, SHI Lei^{1,2}, GENG Xiang^{1,2,*}, LIU Xiu-hong^{1,2}

(1. Technical Center of Inspection and Quarantine Jiangxi Entry-Exit Inspection
and Quarantine Bureau of P.R.China, Nanchang 330038, China;

2. Jiangxi Province Engineering Research Center of Infrared Spectroscopy Application, Nanchang 330038, China)

Abstract: A near infrared analysis model of protein contents in nutritious infant rice flour was established by using near infrared spectroscopy and partial least square method. The spectral stability of the sample with different sample thickness 2, 3, 4, 5, 6, 7, 8 mm were compared. The standard deviation of the sample spectrum under different wave number determined that 5 mm was the best loading thickness. The effects of variable standardization, first order derivative, multivariate scattering correction and 9-point smoothing preprocessing method on the model establishment were compared. According to the correlation coefficient and root mean square error of the model, the best pretreatment methods were variable standardization, first derivative and 9-point smoothing. The near infrared quantitative analysis model of protein contents in infant nutritious rice flour was established. The calibration correlation coefficients of calibration set and verification set were 0.9856, 0.9841, and the root mean square error were 0.2431%, 0.2456%, respectively. The established NIR quantitative analysis model was verified by external samples. The accuracy of prediction was 96%. The results showed that the near infrared detection method was consistent with the national standard method, and it was fast and didn't need sample pretreatment. It would have wide application value in the analysis of protein content of infant nutritious rice flour.

Key words: near infrared spectroscopy; infant nutrition rice flour; protein contents

中图分类号: TS207.3

文献标识码: A

文章编号: 1002-0306(2019)06-0237-05

doi: 10.13386/j.issn1002-0306.2019.06.039

引文格式: 于清丽, 石磊, 耿响, 等. 近红外光谱法检测婴幼儿营养米粉蛋白质含量[J]. 食品工业科技, 2019, 40(6): 237-240, 251.

婴幼儿营养米粉指的是以大米为主要原料, 同时加入维生素、营养强化剂、矿物质等其他辅料, 经

收稿日期: 2017-11-30

作者简介: 于清丽(1987-), 女, 硕士, 工程师, 研究方向: 食品快速检测, E-mail: leeyu0104@126.com。

* 通讯作者: 耿响(1981-), 女, 博士, 高级工程师, 研究方向: 分子光谱快速检测技术, E-mail: gengxiang2005@sina.com。

基金项目: 江西省重点研发计划项目(20161BBE50101)。

加工制成的粉状产品,适用于婴幼儿食用,营养成分能满足婴幼儿的部分营养需要^[1-2],是母乳或婴幼儿配方食品不能满足婴幼儿营养以及婴幼儿断奶期间,为婴幼儿补充营养的辅助食品。近年来,市场上屡见不鲜的食品安全事件从侧面揭示了食品工业中的薄弱环节和漏洞,引起了社会各界对婴幼儿营养米粉的关注。

目前国内外学者对婴幼儿营养米粉的研究,主要集中在其对婴幼儿营养特性方面,很少关注营养成分检测方面^[3-5],蛋白质对正在成长发育的婴幼儿来说尤为重要,年龄越小,成长发育速度也就越快,所需蛋白质也就相应的越多,能够现场快速检测蛋白质含量对婴幼儿辅助食品的生产加工非常有意义。近红外光谱技术已被应用于液体奶等食品的检测中^[6-12],但迄今尚未有对婴幼儿营养米粉中蛋白质快速检测的研究报道。

针对已有研究对婴幼儿米粉的营养成分研究较少的情况,本文采用近红外光谱技术这一快速、高效、环保,无需前处理的分析方法,结合化学计量学方法及常规化学检测数据,建立蛋白质含量快速检测分析定量模型,为快速定量检测婴幼儿营养米粉中蛋白质含量提供科学依据。

1 材料与方法

1.1 材料与仪器

婴幼儿营养米粉 共4个品牌,分别是亨氏43个、雅士利39个、贝因美45个、美庐33个,共160个,来源于2015年至2016年食品药品监督管理部门在江西省境内开展质量监管过程中抽检的婴幼儿营养米粉样品。

NIRFlex N-500 傅立叶近红外光谱仪 瑞士步琦(BUCHI)公司。

1.2 实验方法

1.2.1 蛋白质含量的测定 采用国标方法^[14]中的凯氏定氮法测定婴幼儿米粉样品中的蛋白质含量。

1.2.2 光程的选择 从160个实验样品中随机选取3个米粉样品,在相同的实验环境下,将样品置于样品池中,用反射板压实样品至下底面平整无裂缝,分别在装填厚度2、3、4、5、6、7、8 mm,波数范围4000~10000 cm⁻¹、扫描次数32次条件下进行近红外光谱采集,每种装填厚度进行4次平行采集实验,计算4次平行实验各波数下反射率的标准差,以确定最佳光程。

1.2.3 样品原始光谱的采集 由于米粉属于小颗粒样品,放置在采样盘中,加盖反射板后较为紧密,所以直接测定样品,不进行粉碎过筛处理。光谱采集时,室温10~30℃,相对湿度20%~80%,波数范围4000~10000 cm⁻¹,扫描次数32次,装样厚度5 mm。

1.2.4 光谱预处理 样品的近红外光谱所包含的信息中有一部分与待测样品性质无关,且为后面的近红外数据处理带来干扰,会导致近红外光谱噪声较大、稳定性较差、所建立模型的准确性降低,因此,有必要对样品的原始近红外光谱进行预处理。对于固体样品,其光谱的预处理方法主要包括导数、变量标

准化、平滑和多元散射校正等。

本文采用近红外光谱仪自带的数据处理软件NIRCAL,从中选择一阶导数、二阶导数、变量标准化、9点平滑、多元散射校正预处理方法对样品的近红外光谱数据进行处理,然后利用偏最小二乘法建立近红外定量模型,通过模型的相关系数和均方根误差比较不同预处理方法的效果,选择相关系数最大,均方根最小的预处理方法作为最佳预处理方法。

1.3 模型评价

本论文中选择最佳的预处理方法是根据不同预处理后的样品数据所建立模型的相关系数和均方根误差来判定,其中均方根误差按公式(1)计算:

$$\text{均方根误差} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n-1}} \quad \text{式(1)}$$

式中: $\hat{y}_i$ —校正集(或预测集)中第*i*个样品蛋白质含量的近红外预测值; y_i —校正集(或预测集)中第*i*个样品蛋白质含量的标准方法测定值; n —校正集(或预测集)中样品个数。

2 结果与分析

2.1 光程的选择

光程由样品装填厚度直接影响,样品装填厚度应保证光谱采集时大部分光不会透过样品。装填厚度过薄,无法保证样品完全覆盖样品池的底面,近红外光穿过样品缝隙,造成能量损失,光谱所含信息不准确;装填过厚则造成浪费。为了确定最佳的光谱采集光程,随机选择了3个样品,7种装填厚度下,每种装填厚度分别重新装样4次,得到该样品在该光程下的4条光谱,计算4条光谱的标准差,比较不同装填厚度的稳定性。图1为3个样品在不同装填厚度时不同波数下的标准差。

从图1看出:当样品装填厚度为2、3、4 mm时,光谱标准差均较大,说明此时光谱的稳定性较差,当装填厚度达到5 mm及以上时,标准差明显减小,标准差大小与装填厚度为6、7、8 mm时相差不大,3个样品在不同装填厚度的标准差比较均有此规律,所以在满足实验要求的情况下,为了减少样品的实验用量,所以本实验选择最佳光程为5 mm。

2.2 光谱采集

采集160个婴幼儿米粉样品的近红外光谱,每个样品重新装样三次采集光谱,将三次的近红外光谱取平均作为该样品的近红外光谱。图2给出的是160个样品的原始近红外光谱,从图2中可以看出所有样品的近红外光谱谱型一致,不存在异常样品,米粉中的蛋白质信息主要分布在4500~4700 cm⁻¹以及5760 cm⁻¹附近,其他出峰主要为米粉中的脂肪、水分、纤维素等物质的信息。

2.3 定标集、验证集划分

目前,定标集和验证集的划分方法主要包括随机法和kenard-stone(KS)法,本研究中选用随机法将160个米粉样品随机划分为定标集和验证集,但样品中蛋白质含量的最大值和最小值包含在定标集中,以保证定标集涵盖所有样品。表1给出了划分后的

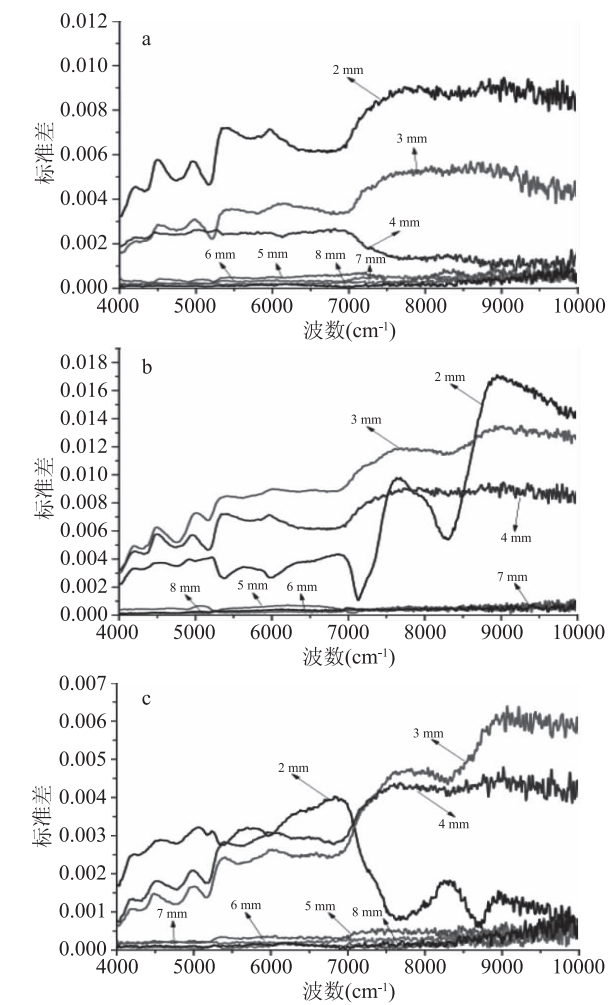


图1 3个样品在不同装填厚度下各波数的反射率标准差
Fig.1 Reflectivity standard deviation of three samples in different wave numbers with different loading thickness

定标集和验证集的蛋白质含量。从表1可知:定标集样品的蛋白质含量为5.43%~16.8%,最小值为5.43%,验证集样品蛋白质含量为5.76%~12.4%,国标^[1]中规定婴幼儿谷类辅助食品中蛋白质含量不小于1.94%,本试验收集的婴幼儿米粉的蛋白质含量均满足国标要求。

2.4 模型建立

表2给出的是将160个样品的近红外光谱采用

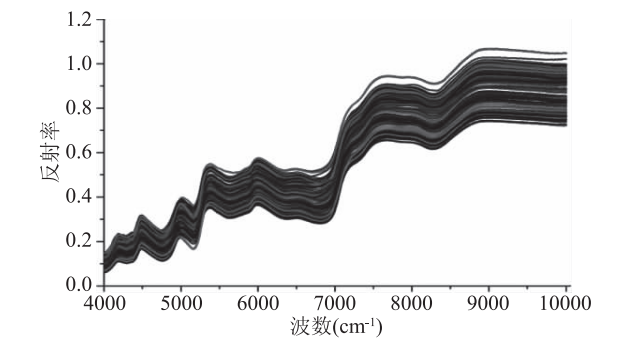


图2 160个样品的近红外光谱图
Fig.2 Near infrared spectra of 160 infant nutrition rice flour

表1 定标集和验证集蛋白质含量
Table 1 Contents of protein of calibration sets and validation sets

项目	样品数	最大值 (%)	最小值 (%)	平均值 (%)	标准偏差 (%)
总样本	160	16.80	5.43	8.28	1.88
定标集	110	16.80	5.43	8.11	1.93
验证集	50	12.40	5.76	8.65	1.77

不同的方法预处理后,利用偏最小二乘方法建立的蛋白质定量分析模型。变量标准化可以消除固体颗粒大小、光程变化等对NIR漫透射光谱的影响;多元散射校正主要用来消除颗粒分布不均匀及颗粒大小产生的散射影响;导数是基线校正的方法,对光谱进行一阶导数处理能够明显消除光谱的基线漂移,从而减弱谱带重叠,突出谱带特征;对光谱进行二阶导数处理则可以消除光谱的旋转误差,但进行二阶导数会使得随机噪声变大,降低信噪比;平滑可以降低随机误差和高频噪声。从表2中可以看出,变量标准化、一阶导数和9点平滑的效果最好,可能是由于米粉样品颗粒度较为均匀,颗粒度不大,多元散射校正效果不明显,而变量标准化可以进一步消除颗粒度及光程对样品光谱的影响,使其一致性和稳定性更好,二阶导数会使得有效信息强度增强,但随机噪声也随之增大了,导致模型效果不佳。从表2可以看出,经过变量标准化、一阶导数和9点平滑预处理后,所建立的定量分析模型的 R_c 、 R_p 为0.9856、

表2 不同预处理方法及所建立近红外模型的结果

Table 2 Results of calibration sets spectrum with different pretreatment methods

预处理方法	R_c	RMSEC (%)	R_p	RMSEP (%)
变量标准化	0.9354	0.6686	0.9037	0.8178
变量标准化、一阶导数和9点平滑	0.9856	0.2431	0.9841	0.2456
变量标准化、二阶导数和9点平滑	0.9591	0.4840	0.9305	0.6914
多元散射校正	0.9406	0.6142	0.9121	0.7624
多元散射校正、一阶导数和9点平滑	0.9243	0.7217	0.8987	0.8568
多元散射校正、二阶导数和9点平滑	0.8859	0.8699	0.8820	0.9002

注: R_c :校正集相关系数,即校正集样品的常规方法测定值与模型预测值间的相关系数; R_p :验证集相关系数,即验证集样品的蛋白质常规方法测定值和模型预测值间的相关系数;RMSEC:root mean square error of calibration,校正均方根误差;RMSEP:root mean square error of validation,预测均方根误差,RMSEC和RMSEP均指近红外光谱法测定值与标准方法测定值之间残差的标准差。

0.9841, 均方根误差 RMSEC 和 RMSEP 分别为 0.2431、0.2456, 较其他预处理方法效果更好, 相关系数也说明近红外方法与经典方法测定的样品蛋白质含量相关性较好, 且 RMSEC 与 RMSEP 均小于 0.3%, 均小于国标方法测定的样品蛋白质含量的 10%, 满足国标方法^[14]中对于蛋白质含量测定方法精度的要求。

为了更好地对所建立的蛋白质近红外定量分析模型的预测精确度和可靠性进行评价, 进一步计算定标集和验证集的相对分析误差 RPD, 以验证模型的预测能力, 蛋白质定量分析模型的定标与验证集的 RPD 结果见表 3。

表3 蛋白质近红外模型的定标与验证结果

Table 3 Results of calibration and verification of NIRS protein content analysis model

项目	相关系数 r	RPD
定标集	0.9856	3.59
验证集	0.9841	3.61

注: RPD(相对分析误差) = 标准差/均方根误差。

相关系数 R 越接近 1, 则表示建立的预测模型越精确, RMSEP/RMSEC 越小, RPD 越大 (>3), 则表示模型的预测效果越好。从表 3 可看出, 对于定标集和验证集, 样品蛋白质含量的国标方法分析测定值和预测值的相关系数均在 0.98 以上, RPD 均大于 3, 说明所建立的近红外定量分析模型可以准确地预测婴幼儿营养米粉中蛋白质的含量, 说明采用该预处理方法后所建立的近红外定量分析模型具有很好的精确度和稳定性。

2.5 外部验证结果

为了进一步验证模型的适用性, 用建立的蛋白质近红外定量模型预测另外收集的未参加近红外模型建立的 50 个婴幼儿营养米粉样品的蛋白质含量, 与国标方法^[14]所得结果进行比较, 计算两组结果的绝对差值与算术平均值的比值, 其比值小于 10% 即认为预测准确。图 3 给出了对 50 个样品预测结果的准确度。

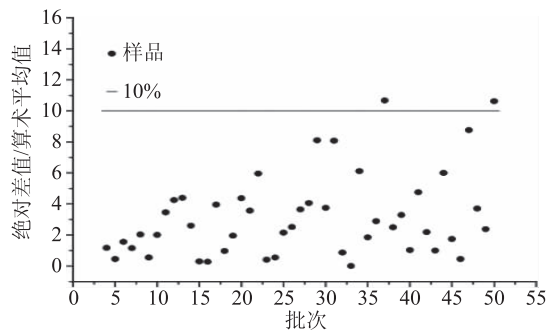


图3 蛋白质含量预测准确度

Fig.3 Forecast accuracy of protein contents

从图 3 可看出, 用建立的近红外定量模型预测 50 个婴幼儿营养米粉样品的蛋白质含量, 与国标方法相比, 仅有 2 个样品的绝对差值与算术平均值的比值大于 10%, 即预测结果误差较大, 该外部验证证

明所建立的近红外婴幼儿米粉的蛋白质定量模型的预测准确度达到 96%。

3 结论

对于 160 个婴幼儿营养米粉样品, 首先比较了不同装填厚度 2、3、4、5、6、7、8 mm 对样品近红外光谱的影响, 确定了最佳的装填厚度为 5 mm; 在波数 4000~10000 cm^{-1} 、扫描次数 32 次, 样品装填厚度 5 mm 条件下, 采集了所有样品的近红外光谱。为了确定最佳的预处理方法, 比较了导数、变量标准化、多元散射校正和平滑预处理方法对所建立模型的影响, 确定了最佳的预处理方法为变量标准化、一阶导数和 9 点平滑, 结合国标方法测定的蛋白质含量, 利用偏最小二乘方法建立了婴幼儿营养米粉中蛋白质含量的近红外分析模型, 模型的定标集和预测集的相关系数均大于 0.98, 均方根误差均小于 0.3%, 且 RPD 均小于 3, 又进一步利用未参与建模的 50 个样品进行外部验证, 模型的预测准确度达到 96%, 该结果说明了近红外光谱法快速检测婴幼儿营养米粉中蛋白质含量的可行性, 为近红外光谱技术在婴幼儿食品营养指标快速检测上的应用提供了依据。

参考文献

- [1] 中华人民共和国卫生部. GB 10769-2010 食品安全国家标准婴幼儿谷类辅助食品[S]. 北京: 中国标准出版社, 2010.
- [2] 中华人民共和国卫生部. GB 10767-2010 食品安全国家标准较大婴儿和幼儿配方食品[S]. 北京: 中国标准出版社, 2010.
- [3] 潘菁. 婴幼儿营养米粉配方优化及加工关键技术研究[D]. 无锡: 江南大学, 2012.
- [4] 潘菁, 汪何雅, 钱和. 一种婴幼儿低敏营养米粉的配方及关键工艺的研究[J]. 食品工业科技, 2012, 33(9): 268-270.
- [5] 魏红, 徐宏伟, 钟红舰, 等. 膳食纤维的应用及检测方法[J]. 海峡预防医学杂志, 2004(2): 25-27.
- [6] 姚宁, 程顺国. 近红外光谱分析技术及发展前景[J]. 现代制造技术与装备, 2012(5): 71-74.
- [7] 朱向荣, 单杨, 李高阳, 等. 近红外光谱法快速测定液态奶中蛋白质和脂肪含量[J]. 食品科学, 2011, 32(12): 191-195.
- [8] 谷月, 赵武奇, 李晓丹, 等. 近红外光谱技术对液态奶制品中乳糖、脂肪、蛋白质的定量检测[J]. 宁夏大学学报, 2012, 33(4): 376-380.
- [9] 李德海, 孙常雁, 孙莉洁, 等. 婴幼儿断乳期营养米粉的研究进展[J]. 食品工业科技, 2010(3): 401-404.
- [10] Ouissan Abbas, Bernard Lecler, Pierre Dardenne, et al. Detection of melamine and cyanuric acid in feed ingredients by near infrared spectroscopy and chemometrics[J]. Journal of Near Infrared Spectroscopy, 2013, 21(3): 183.
- [11] Andrei Kalinin, Viktor Krashennnikov, Sergei Sadovskiy, et al. Determining the composition of proteins in milk using a portable near infrared spectrometer[J]. Journal of Near Infrared Spectroscopy, 2013, 21(5): 409.
- [12] Cédric Camps, Marianne Gérard, Mélanie Quennoz, et al. Prediction of essential oil content of oregano by hand-held and

(下转第 251 页)

- [4] 黄韬睿,王鑫,孟甜,贾腌菜中亚硝酸盐检测方法的研究进展[J].食品研究与开发,2016,37(6):204-206.
- [5] 王君,刘蓉.近红外光谱技术在液态食品掺假检测中的应用[J].食品工业科技,2016,37(7):374-386.
- [6] 杰克·沃克曼,洛伊斯·文依 编著,诸小立,等译.近红外光谱解析实用指南[M].北京:化学工业出版社,2009:12-13.
- [7] Hirschfeld T. Salinity determination using NIRA[J]. Applied spectroscopy, 1985, 39(4):740-741.
- [8] Hans Buning-Pfaue. Analysis of water in food by near infrared spectroscopy[J]. Food Chemistry, 2003, 82:107-115.
- [9] Lin M, Cavinato A G, Huang Y, et al. Predicting sodium chloride content in commercial king (*Oncorhynchus tshawytscha*) and chum (*O. keta*) hot smoked salmon fillet portions by short-wavelength near-infrared (SW-NIR) spectroscopy [J]. Food Research International, 2003, 36(8):761-766.
- [10] De M M, Manuelian C L, Ton S, et al. Prediction of sodium content in commercial processed meat products using near infrared spectroscopy[J]. Meat Science, 2017, 125:61-65.
- [11] 中华人民共和国卫生部. GB 5009.33-2016 食品安全国家标准食品中亚硝酸盐与硝酸盐的测定[S]. 北京:中国标准出版社, 2016:149-153.
- [12] 刘翠玲,吴景珠,孙晓荣.近红外光谱技术在食品品质检测方法中的研究[M]. 北京:机械工业出版社, 2015:24-47.
- [13] Ferreira D S, Galão O F, Pallone J A L, et al. Comparison and application of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for determination of quality parameters in soybean samples[J]. Food Control, 2014, 35(1):227-232.

(上接第 240 页)

- Fourier transform NIR spectroscopy[J].Journal of the Science of Food and Agriculture,2014,94(7):1397-1402.
- [13] 刘晓春,陈磊,丁勇宝,等.常量法检测蛋白质的方法改进[J].轻工科技,2013(5):126-127.
- [14] 中华人民共和国卫生部.GB 5009.5-2010,食品安全国家标准食品中蛋白质的测定[S].北京:中国标准出版社,2010.
- [15] 陆婉珍.现在近红外光谱分析技术第二版[M].北京:中国石化出版社,2006.
- [16] 王伟,张玉,王楠,等.基于傅里叶变换近红外光谱的奶粉

(上接第 244 页)

- 2000,21(8):46-48.
- [15]王丽玲,白艳玲.水产品组胺测定中盐酸量对结果影响的探讨[J].现代科学仪器,2001(5):55.
- [16]胡红美.青条鱼中组胺检测方法探讨[J].安徽预防医学杂志,2011,17(6):468-469.
- [17]金安宝.水产品中组胺测定方法的改进[J].中国卫生检验杂志,2001,17(1):149-150.
- [18]冯美西,陈伟珍,侯丽琼,等.湛江海域鱼类组胺含量的检测分析[J].保鲜与加工,2016,16(3):97-101.
- [19]张继红,马俊华,刘平.分光光度法测定鱼类产品中组胺[J].医学动物防制,2013,29(9):1060-1061.
- [20]丁海燕,王联珠,盛晓风,等.非挥发性生物胺对组胺的可见分光光度法检测结果的影响[J].中国渔业质量与标准,2015,5(4):17-22.
- [21]王新惠,白婷,李俊霞,等.发酵肉制品中组胺含量测定方法的比较[J].食品与机械,2015,31(3):37-39.

- [14] 严衍禄,赵龙莲,韩东海.近红外光谱分析基础与应用[M].北京:中国轻工业出版社,2005:29-81.
- [15] 鲁超,韩东海,温朝晖,等.基于近红外光谱检测腐乳盐坯中盐分含量的研究[J].中国酿造,2006(4):9-13.
- [16] 刘冰.近红外光谱法在涪陵榨菜品质检测方面应用研究[D].重庆:西南大学,2011:53-59.
- [17] 朱向荣,李高阳.基于近红外光谱与组合间隔偏最小二乘法的稻米镉含量快速检测[J].食品与机械,2015,3(4):43-50.
- [18] Chen J, Zhu S P. Rapid determination of total protein and wet gluten in commercial wheat flour using siSVR-NIR[J]. Food Chemistry, 2017, 221: 1939-1946.
- [19] 陈大伟,闫昭,刘昊岩. SVD 系列算法在评分预测中的过拟合现象[J]. 山东大学学报:工学版, 2014, 44(3): 15-21.
- [20] 公丽艳,孟宪军. 基于主成分与聚类分析的苹果加工品质评价[J]. 农业工程学报, 2004, 30(13): 276-283.
- [21] 褚小立,邓勇,杜一平. 近红外光谱分析技术实用手册[M]. 北京:机械工业出版社 2016:1-4, 114-160.
- [22] 蒋霞,张晓,白铁成,等. 近红外光谱技术结合 PLS 和 SPA 检测鲜冬枣表面农药残留量的方法[J]. 江苏农业科学, 2018, 46(2): 146-149.
- [23] 靳庭良,张宝青. 回归分析中 t 检验与 F 检验关系的进一步探讨[J]. 统计与决策, 2009, 297(21): 7-9.
- [24] 李云雁,胡传荣. 试验设计与数据处理[M]. 北京:化学工业出版社. 2008, 23-41.
- [25] 刘润碧. 用 Excel 做 t 检验和 F 检验[J]. 海峡预防医学杂志, 2002, 8(5): 67-68.

- [illegible]