

鸭肉中谷氨酸含量的可见-近红外光谱测定研究

赵进辉 刘木华 吁芳 沈洁 涂东成

(江西农业大学工学院, 江西南昌 330045)

摘要:为实现鸭肉中谷氨酸含量的快速测定,提出了利用可见-近红外光谱结合 PCA、BP 神经网络来建立鸭肉中谷氨酸含量测定的定量分析模型。采集试验首先采集光谱范围在 350~1800nm 的鸭肉可见-近红外反射光谱,并在 430~1000、1001~1400 和 430~1400nm 3 个光谱范围内分别用一阶导数(FD)、二阶导数(SD)、多元散射校正(MSC)、标准正交变量变换(SNV)4 种方法对原始光谱进行光谱预处理;然后以前 8 个主成分得分和鸭肉中谷氨酸含量分别作为 BP 神经网络的输入变量和输出变量来建立鸭肉中谷氨酸含量预测模型。试验结果表明,在 430~1000nm 光谱范围内,采用 SNV 光谱预处理建立的 BP 神经网络模型为最优,其验证集的相关系数为 0.9564,预测样本均方根误差为 0.058572,能够满足鸭肉中谷氨酸含量快速测定的实际要求。

关键词:鸭肉;谷氨酸;可见-近红外光谱;BP 神经网络;主成分分析

QUANTITATIVE ANALYSIS OF GLUTAMATE CONTENT IN DUCK MEAT USING VISIBLE-NEAR INFRARED SPECTROSCOPY

ZHAO Jin-hui LIU Mu-hua YU Fang SHEN Jie TU Dong-cheng

(College of Engineering, Jiangxi Agricultural University, Nanchang, Jiangxi 330045)

Abstract: Quantitative analysis models of glutamate contents in duck meat were established using visible-near infrared spectroscopy combining with principal component analysis (PCA) and BP neural network to achieve the rapid prediction of glutamate contents in duck meat. Visible-near infrared spectroscopies of duck meat in the range of 350~1800nm were obtained, and first derivative (FD), second derivative (SD), multiplicative scatter correction (MSC) and standard orthogonal variable transformation (SNV) were used to complete the pretreatment of original spectroscopy in the spectral region of 430~1000, 1001~1400 and 430~1400nm, respectively. The scores of 8 principal components (PCs) and glutamate contents in duck meat were selected as the inputs and outputs of BP neural network to build the prediction model of glutamate content in duck meat, respectively. The results showed that the model of BP neural network established using the pretreatment spectroscopy by SNV in the spectral region of 430~1000nm was optimal. The correlation coefficient and the root mean squared error of prediction samples were 0.9564 and 0.058572 in the validation set respectively, and the model could meet the demands of rapid detection of duck meet glutamate content.

Key words: duck meat; glutamate; visible-near infrared spectroscopy; BP neural network; principal component analysis

鸭肉是人们食用较多的一种禽肉产品,大量研究表明,鸭肉中富含的谷氨酸不仅可参与脑蛋白和碳水

化合物的代谢,增强大脑功能,缓解疲劳,加速溃疡愈合,同时也是一种重要的肉类滋味呈味物质。谷氨酸

收稿日期:2011-22-22 接受日期:2011-05-20

基金项目:国家高技术研究与发展技术(863 计划)资助项目(2008AA10Z209)

作者简介:赵进辉(1978-),男,湖南华容人,博士,讲师,主要从事图像处理与农畜产品无损检测研究。E-mail:zjhxiaocao@sina.com

通讯作者:刘木华(1969-),男,江西赣州人,博士,教授,主要从事农畜产品无损检测研究。E-mail:suikelmh@sina.com

含量是评价鸭肉品质的重要指标之一。因此研究鸭肉中谷氨酸含量的快速、无损测定方法具有重要的社会和经济意义。传统的谷氨酸含量测定方法,如离子交换色谱法(IEC)或高效液相色谱法(HPLC)虽然具有较高的准确度和良好的重复性,但检测费时,费用高,且需使用有毒化学试剂^[1-3]。

可见-近红外光谱分析技术具有操作简单、无毒害、速度快、稳定性好,能实现在线无损分析的特点^[4-6]。国内外已有学者利用近红外光谱来测定鱼粉、饲料、大豆、稻米及中国黄酒中的谷氨酸含量^[7-12],但目前还没有利用近红外光谱测定鸭肉中谷氨酸含量的报道。

本研究采用可见-近红外反射光谱技术,结合主成分分析和BP神经网络技术来建立鸭肉中谷氨酸含量快速测定的定量分析模型,旨在满足鸭肉中谷氨酸含量快速测定的实际需要。

1 材料与方 法

1.1 试验材料

试验对象麻鸭购于江西农业大学菜市场。将鸭去毛后悬挂在铁架上,自然状态下冷却1h后进行取样。由鸭的生理结构可知,鸭的整体谷氨酸含量高低可以在特殊部位得到体现,谷氨酸含量高的鸭,其特殊部位谷氨酸含量也高,这也是许多研究者取特殊部位进行品质分析的依据^[13]。本研究取鸭身上的特殊部位鸭胸脯肉(整片)作为试验样本,共取70只,鸭胸肉分别编号为01~70号。

1.2 仪器设备

试验用光谱仪为QualitySpec Pro光谱仪(Analytical Spectral Devices, Inc., USA),光谱测量范围为350~1800nm,其中在350~1000nm采样间隔为1.4nm,分辨率为3nm,在1000~1800nm采样间隔为2nm,分辨率为10nm;日立L-8800型氨基酸自动分析仪。

1.3 可见-近红外反射光谱采集与化学测定

为防止干扰,光谱采集在密闭的暗箱中进行。采集样本光谱前,对系统配置优化标准白板进行标定。采集样本光谱时,将样本置于圆形石英器皿中,用光谱仪自带附件高密度光源检测探头采集样本的可见-近红外反射光谱。每个样本在3个不同部位进行扫描,并取平均值作为该样本的原始光谱。应用Indico4.0(Analytical Spectral Devices, Inc., USA)软件和与之配套的标准白板采集归一化后的光谱数据并转换格式,

以便后续数据处理和分析。

样本在完成光谱采集后,对鸭肉中的谷氨酸含量按照GB/T 5009.124-2003食品中氨基酸的测定^[14]进行。

1.4 数据分析

用数据处理软件Unscramble V9.7和DPS V7.05分析鸭胸脯肉的可见-近红外反射光谱与鸭胸脯肉中谷氨酸含量之间的关系。

2 结果与讨论

2.1 光谱波段范围的选择

对获取的70个样本按照谷氨酸含量顺序,每间隔4个样本选择1个样本,作为验证集样本,共13个,即05、10、15、20、25、30、35、40、45、50、55、60、65号;其余的57个作为校正集样本。

图1为校正集样本的原始光谱。由于光谱仪在测量范围的两端存在较多噪音,这一点从350~430nm范围内光谱曲线可以看出;同时,在1400~1800nm范围内光谱曲线比较平直,相对反射率变化不大;此外,在350~1100nm光谱范围内,为低噪声512阵元PDA探测器,在1000~1800nm光谱范围内,为InGaAs探测器。综上所述,本研究选择430~1000、1001~1400和430~1400nm 3个波段进行对比分析,以选择最佳波段范围。

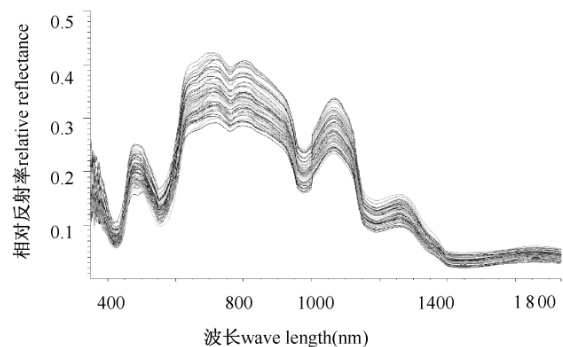


图1 鸭肉校正集样本的可见-近红外光谱图

Fig. 1 Visible-near infrared spectra of calibration set samples of duck meat

2.2 光谱预处理

由于在采集的可见-近红外光谱中往往包含一些与待测样品性质无关的噪声,导致可见-近红外光谱的基线漂移、旋转和光谱的不重复^[15],因此,为了得到更加稳定和有效的预测模型,有必要对原始光谱进行预处理。本研究采用一阶导数(first derivative, FD)、

二阶导数 (second derivative, SD)、多元散射校正 (multiplicative scattering correction, MSC) 和标准正交变量变换 (standard normal variate, SNV) 4 种方法对原始光谱进行预处理,并从中选择一种较优的光谱预处理方法。

2.3 主成分分析

若将全波段的波长变量作为 BP 神经网络的输入,会大大地增加计算量,同时,输入变量越多,建模训练样本也越多,一些重复和次要的变量也会使模型的预测能力下降,因此,有必要减少输入变量数目。本研

究采用 PCA 分析,以得到合适的 BP 神经网络的输入变量。本研究以波长点的相对反射率作为 PCA 分析的输入变量,通过交互验证 (cross validation) 方法得到相应的主成分的特征值及累计可信度。表 1 给出了在 3 个光谱范围内采用不同光谱预处理方法得到的前 10 个主成分的累计可信度。由表 1 可知,前 8 个主成分的累积可信度在 94% 以上,能够反映光谱数据绝大部分的特征信息。因此,本研究以前 8 个主成分得分作为 BP 神经网络的输入变量,这样就把原来较多的波长变量压缩成了彼此正交的 8 个新变量。

表 1 不同光谱范围和预处理方法下的前 10 个主成分的累计可信度

Table 1 Accumulative reliabilities of the first 10 PCs of different preprocessing methods and spectral ranges

光谱范围 spectral range (nm)	预处理方法 preprocessing method	主成分 principal components	主成分									
			PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
430 ~ 1000	原始光谱 original spectroscopy	累计可信度	90.834	98.748	99.414	99.669	99.813	99.905	99.959	99.976	99.983	99.99
	一阶导数 FD	accumulative	66.777	79.132	88.029	90.871	93.401	94.823	95.627	96.234	96.742	97.113
	二阶导数 SD	reliability	83.871	89.556	90.791	91.868	92.877	93.629	94.317	94.881	95.413	95.912
	多元散射校正 MSC	(%)	84.824	93.169	95.97	97.271	98.484	99.353	99.758	99.848	99.893	99.926
	标准正交变量变换 SNV		83.946	92.884	95.968	97.263	98.454	99.325	99.735	99.844	99.891	99.923
1001 ~ 1400	原始光谱 original spectroscopy	累计可信度	92.423	99.137	99.813	99.898	99.958	99.98	99.987	99.99	99.993	99.995
	一阶导数 FD	accumulative	75.45	88.041	91.314	92.258	93.122	93.789	94.44	95.035	95.532	95.998
	二阶导数 SD	reliability	88.145	89.299	90.361	91.289	92.14	92.867	93.516	94.112	94.655	95.169
	多元散射校正 MSC	(%)	85.387	99.045	99.492	99.722	99.883	99.939	99.962	99.973	99.98	99.984
	标准正交变量变换 SNV		85.261	99.039	99.487	99.718	99.877	99.939	99.962	99.973	99.98	99.983
430 ~ 1400	原始光谱 original spectroscopy	累计可信度	84.658	95.249	98.729	99.281	99.621	99.75	99.854	99.925	99.96	99.972
	一阶导数 FD	accumulative	67.576	80.585	87.728	90.25	92.272	93.718	94.634	95.268	95.786	96.252
	二阶导数 SD	reliability	84.058	89.237	90.357	91.255	92.107	92.833	93.498	94.059	94.547	95.02
	多元散射校正 MSC	(%)	79.081	92.288	96.534	97.849	98.719	99.336	99.641	99.778	99.835	99.882
	标准正交变量变换 SNV		78.685	92.099	96.409	97.723	98.594	99.255	99.574	99.755	99.819	99.871

2.4 预测模型的建立

本研究采用只含有 1 个隐含层的 3 层的 BP 神经网络结构来建立 BP 神经网络模型。以前 8 个主成分得分作为 BP 神经网络的输入,鸭肉中谷氨酸含量作为 BP 神经网络的输出,因此,输入层和输出层节点数分别确定为 8 和 1。目前还没有一个很好的解析式来表达隐含层节点数,本研究根据如下经验公式^[16,17]来初步确定隐含层节点数。

$$j = \sqrt{n + m} + \alpha \quad (1)$$

式中: j 为隐含层节点数, n 为输入层节点数, m 为输出层节点数, α 为 1~10 的常数。

由于输入层节点数为 8,输出层节点数为 1,根据式(1)初步确定隐含层节点数为 4、5、6、7、8、9、10、11、

12 和 13 中的 1 个。为从中选择 1 个作为隐含层节点数,通过多次改变隐含层节点数进行分析对比,以确定最佳隐含层节点数。网络训练的其他的参数为:网络的初始权值为 -0.15 ~ +0.15 之间的随机数,最大迭代次数为 1000 次,最小训练速率取 0.1,动态参数为 0.6,允许误差为 0.0001。表 2 给出了在不同光谱范围和 FD、SD、MSC、SNV 光谱预处理方法下的最佳隐含层节点数和校正模型结果。

2.5 模型预测结果与分析

相关系数 (R^2) 越大,校正均方根误差 (RMSEC)、预测均方根误差 (RMSEP) 越小,说明模型的预测能力越强。由表 2 可知,在 430~1000、1001~1400 和 430~1400nm 光谱范围内,用 SNV 光谱预处理方法建立

的 BP 模型的 RMSEC 比其他几种光谱处理方法建立的 BP 模型的明显降低,而 R^2 则明显增加,说明在鸭肉中谷氨酸含量的可见-近红外光谱测定中用 SNV 光谱预处理方法建立的 BP 模型的预测能力比其他几种光谱预处理方法建立的 BP 模型更强。因此,本研究确定选用 SNV 光谱预处理方法作为鸭肉中谷氨酸含量测定的可见-近红外光谱预处理方法。由表 2 可

知,在 430 ~ 1000nm 光谱范围内,用 SNV 光谱预处理方法建立的 BP 模型的 R^2 和 RMSEC 分别为 0.9987 和 0.00972;在 1001 ~ 1400nm 光谱范围内,用 SNV 光谱预处理方法建立的 BP 模型的 R^2 和 RMSEC 分别为 0.808 和 0.110854;在 430 ~ 1400nm 光谱范围内,用 SNV 光谱预处理方法建立的 BP 模型的 R^2 和 RMSEC 分别为 0.9963 和 0.013972。

表 2 不同光谱范围和预处理方法下的校正模型结果

Table 2 Calibration model results of different preprocessing methods and spectral ranges

光谱范围 spectral range (nm)	预处理方法 preprocessing method	最佳隐含层节点数 No. of appropriate hidden node	校正均方根误差 RMSEC	相关系数 (R^2) correlation coefficient
430 ~ 1000	原始光谱 original spectroscopy	6	0.061921	0.9321
	一阶导数 FD	7	0.228691	0.6123
	二阶导数 SD	5	0.229558	0.4731
	多元散射校正 MSC	6	0.135765	0.6557
	标准正交变量变换 SNV	6	0.00972	0.9987
1001 ~ 1400	原始光谱 original spectroscopy	6	0.181276	0.3835
	一阶导数 FD	7	0.229186	0.4508
	二阶导数 SD	6	0.229583	0.0562
	多元散射校正 MSC	7	0.155741	0.5492
	标准正交变量变换 SNV	6	0.110854	0.808
430 ~ 1400	原始光谱 original spectroscopy	4	0.084126	0.8674
	一阶导数 FD	4	0.227037	0.5389
	二阶导数 SD	7	0.229032	0.5263
	多元散射校正 MSC	6	0.112374	0.7666
	标准正交变量变换 SNV	6	0.013972	0.9963

表 3 给出了不同光谱范围条件下通过 SNV 光谱预处理的预测模型结果。由表 3 可知,在 430 ~ 1000 和 430 ~ 1400nm 光谱范围内,用 SNV 光谱预处理方法建立的 BP 模型的 R^2 在 0.95 以上, RMSEP 在 0.06 以下,说明在 430 ~ 1000 和 430 ~ 1400nm 这 2 个光谱范围内都可作为用 SNV 光谱预处理方法建立的 BP 模型的光谱波段范围;在 1001 ~ 1400nm 光谱范围内,用 SNV 光谱预处理方法建立的 BP 模型的 R^2 在 0.8 以下,说明在这个光谱范围内用 SNV 光谱预处理方法建立的 BP 模型的预测能力较差,不适合作为用 SNV 光谱预处理方法建立的 BP 模型的光谱波段范围。由上面分析可知,虽然在 430 ~ 1000 和 430 ~ 1400nm 光谱范围内用 SNV 光谱预处理方法建立的 BP 模型的 R^2 和 RMSEP 分别相差不大,都可以作为用 SNV 光谱预处理方法建立的 BP 模型的光谱波段范围,但由于 430 ~ 1400nm 光谱范围的波长点数比 430 ~ 1000nm 光谱范围的波长点数要多,模型建立的计算量会更大。因此,在鸭肉中谷氨酸含量的快速测定中可以优先选择 430 ~ 1000nm 作为建立模型的光谱波段范围。图 2 给出了在 430 ~ 1000nm 谱范围内用 SNV 光谱预处理方法建立的 BP 模型预测未知样本的预测值和实际值之

间的关系图。

表 3 不同光谱范围条件下的预测模型结果

Table 3 Predictive model results of different spectral ranges

光谱范围 spectral range (nm)	预处理方法 preprocessing method	预测均方 根误差 RMSEP	相关系数 (R^2) correlation coefficient
430 ~ 1000	标准正交变量变换 SNV	0.058572	0.9564
1001 ~ 1400	标准正交变量变换 SNV	0.122372	0.7887
430 ~ 1400	标准正交变量变换 SNV	0.049823	0.9626

3 结论

在 430 ~ 1000nm 光谱范围内,采用 SNV 光谱预处理建立的 BP 神经网络模型为最优,其验证集的相关系数为 0.9564,预测样本均方根误差为 0.058572,能够满足鸭肉中谷氨酸含量快速测定的实际应用要求。

参考文献:

[1] 万丹晶, 陈妙芬, 翟咏虹. 用 HPLC 法和氨基酸分析仪测定多维

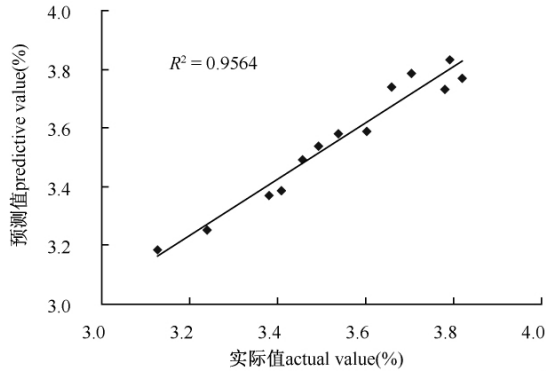


图2 鸭肉中谷氨酸含量测定中预测值和实际值之间的关系
Fig.2 Relationship between actual value and predictive value of glutamate contents in duck meat

氨基酸片中18种氨基酸[J]. 药学服务与研究, 2006, 6(3): 212-214

[2] 李 歆, 李小玲. 高效液相色谱技术在饲料安全生产中的应用[J]. 畜牧与饲料科学, 2009, 30(1): 92-94

[3] 梁丽绒, 吕利华, 赵良启. 饮品与调味品中香味成分分析方法[J]. 山西化工, 2006, 26(2): 28-32

[4] 吴 军, 白琪林, 苏胜宝, 陈绍江, 孟庆翔, 严衍禄. 近红外反射光谱法分析玉米秸秆纤维素含量的研究[J]. 分析化学, 2005, 33(10): 1421-1423

[5] 申 艳, 张晓平, 梁爱珍, 时秀焕, 范如芹, 杨学明. 近红外光谱法在土壤有机质研究中的应用[J]. 核农学报, 2010, 24(1): 199-207

[6] 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, 19(3): 236-240

[7] Shen F, Niu X, Yang D, Ying Y, Li B, Zhu G, Wu J. Determination of amino acids in Chinese rice wine by fourier transform near-infrared spectroscopy[J]. J Agric Food Chem, 2010, 58(17): 9809-9816

[8] González-Martín I, Alvarez-García N, González-Cabrera JM. Near-infrared spectroscopy (NIRS) with a fibre-optic probe for the prediction of the amino acid composition in animal feeds [J]. Talanta 2006, 69(3): 706-710

[9] 牛智有, 韩鲁佳. 鱼粉中氨基酸近红外光谱定量分析[J]. 农业机械学报, 2007, 38(5): 114-117

[10] 刘波平, 秦华俊, 罗 香, 曹树稳, 王俊德. 偏最小二乘-反向传播-近红外光谱法同时测定饲料中4种氨基酸[J]. 分析化学, 2007, (4): 525-528

[11] 陆艳婷, 张小明, 叶胜海, 祁永斌, 金庆生. 稻米品质近红外光谱法分析中定标样品选择与模型优化[J]. 核农学报, 2008, 22(6): 823-828

[12] 王文真, 付翠真. 利用近红外反射光谱快速测定大豆籽粒蛋白质、脂肪和部分氨基酸含量[J]. 中国种业, 1994, (1): 31-32

[13] 刘海霞, 章建浩, 王永丽, 靳国锋. 木瓜蛋白酶嫩化处理对淘汰蛋鸡肌肉蛋白质水解和嫩度等品质的影响[J]. 食品工业科技, 2009, 30(1): 174-177

[14] GB/T5009.124-2003, 食品中氨基酸的测定[S]

[15] 黎 静, 薛龙, 刘木华, 王 晓, 罗春生. 基于可见-近红外光谱识别氧乐果污染的脐橙[J]. 农业工程学报, 2010, 26(2): 366-169

[16] 肖钢铭, 钟 璐, 陈立耀. 神经网络结构与训练参数选取[J]. 武汉工业大学学报, 1997, 19(2): 108-110

[17] 张德丰. MATLAB神经网络应用设计[M]. 北京: 机械工业出版社, 2009

(责任编辑 高美须 裴 颖)

(上接第501页)

[7] 徐建龙, 林贻滋, 奚永安, 张小明, 鲍根良, 张炳林, 姚建龙, 朱卫东, 朱茂丹, 庐王印, 严培鸿, 陈 斌. 高产优质早熟晚糯航育1号的选育、特征特性及其栽培技术[J]. 浙江农业科学, 1999, (2): 74-76

[8] 中国农业科学院. 中国稻作学[M]. 北京: 农业出版社, 1986: 379-418

[9] 肖武名, 杨祁云, 陈志强, 王 慧, 郭 涛, 刘永柱, 朱小源. 空间诱变水稻品系抗稻瘟病遗传及微卫星多态性分析[J]. 中国农业科学, 2008, 41(12): 3952-3958

[10] 石春海, 申宗坦. 早粳粒型的遗传和改良[J]. 中国水稻科学, 1995, 9(1): 27-32

[11] 李源祥, 李金国, 蒋兴村, 华育坚. 丰产高抗“赣早粳47号”的航天育种研究[A]. 刘纪原. 中国航天诱变育种[M]. 北京: 中国宇航出版社, 2007: 169-176

[12] 谢华安. 超级杂交稻恢复系“航1号”的选育与应用[A]. 刘纪原. 中国航天诱变育种[M]. 北京: 中国宇航出版社, 2007: 103-110

(责任编辑 王媛媛)