

不同品种猪肉 pH 值高光谱检测的模型传递修正算法*

李小昱¹ 钟雄斌¹ 刘善梅² 黄涛¹ 武振中¹

(1. 华中农业大学工学院, 武汉 430070; 2. 华中农业大学理学院, 武汉 430070)

摘要: 针对猪肉 pH 值高光谱检测模型受品种差异影响存在适用性差的问题, 比较了不同算法, 提出了一种基于光谱值校正的模型传递算法, 以用于品种之间的模型传递。以山黑猪为主品种建立 PLS 模型, 用该模型直接预测零号土猪样本时, 预测相关系数仅为 0.415, 预测均方根误差为 0.180 4, 预测精度较差。分别用斜率/截距 (S/B) 算法、模型更新算法以及光谱值校正传递算法对山黑猪模型进行修正或传递并进行了比较。采用 S/B 法时, 山黑猪模型对零号土猪的预测相关系数仍为 0.415, 预测均方根误差由 0.180 4 降至 0.134 3, 下降了 25.54%。采用模型更新算法时, 把 14 个零号土猪样本添加到山黑猪校正集, 修正后的山黑猪模型对零号土猪样本的预测性能较优, R_p 由 0.415 提高至 0.797, 提高 92.05%, 预测均方根误差由 0.180 4 降低为 0.112 1, 下降了 37.86%。采用光谱-理化值共生距离法结合 DS 算法的光谱值校正传递算法时, 山黑猪模型对零号土猪样本的预测相关系数由 0.415 提高至 0.837, 提高了 101.69%, 预测均方根误差由 0.180 4 降低至 0.085 6, 下降了 52.55%。结果表明, 光谱值校正的传递算法能够有效消除品种之间光谱差异, 提高了山黑猪模型的适用性, 且传递修正效果优于 S/B 算法和模型更新算法。

关键词: 猪肉 pH 值 高光谱检测 模型传递修正 光谱值校正传递算法

中图分类号: S123; O657.3

文献标识码: A

文章编号: 1000-1298(2014)09-0216-07

引言

利用光谱数据建立的多元校正模型, 由于其样本本身化学组成或物理性质的变化 (如品种的差异), 会使光谱差异变大, 进而可能导致原模型不再适用于未知样本特征值的预测, 存在适用性差的问题^[1]。因此, 模型对其他品种样本进行预测前, 需要对其适用性进行检验, 如果模型出现较大的预测误差, 则需要对原模型修正和传递^[2-3]。目前, 大多数研究针对使用不同仪器而产生的模型适用性问题, 仅有少数学者对模型在品种之间的适用性问题开展相关研究^[4-7]。

现阶段的研究大多是通过模型更新算法或建立全局校正模型来实现品种之间模型的修正。模型更新算法是向原模型的校正集添加新样本, 扩充建模样本集的覆盖范围, 使模型得到修正从而可以预测待测新样本^[8-10]。虽然该方法简单有效, 在模型修正的研究中得到广泛的应用, 但是缺点是向原模型的校正集添加了新样本后, 需重新建立分析模型, 故修正后的模型需要对原模型的预测集样本重新预测, 即存在着模型回检的问题^[11], 而且修正后的模

型对校正集样本预测效果可能面临变差的风险。全局校正模型在建模时需要把可能预测到的因素尽可能包括在校正集中, 每个品种都需要选择较多的校正样本, 且所建立的混合模型预测性能往往不如单个品种所建立的模型, 对提高模型的适用性作用有限^[12]。因此, 借鉴仪器之间模型的传递思路, 本文提出一种基于光谱值校正的品种之间的模型传递算法。

1 材料与方法

1.1 试验材料

冷鲜猪肉样本为不同猪胴体上的里脊肉, 分别购于武汉市沃尔玛与中百仓储超市, 采集时间为 2012 年 12 月份。最终共采集了 82 个山黑猪样本和 79 个零号土猪样本。样本采集后, 用保鲜袋包装, 运至实验室, 上下表面修整, 切成尺寸 (长 × 宽 × 厚) 约为 8 cm × 5 cm × 2.5 cm 的肉块, 装入保鲜袋并逐个编号后, 依次进行高光谱数据采集和 pH 值测定。

1.2 高光谱数据采集

试验采用 HyperSIS 型农产品高光谱成像系统,

收稿日期: 2013-10-09 修回日期: 2013-11-27

* 公益性行业 (农业) 科研专项经费资助项目 (2010033008)

作者简介: 李小昱, 教授, 博士生导师, 主要从事智能化检测与控制技术研究, E-mail: lixiaoyu@mail.hzau.edu.cn

该系统主要由光源、图像光谱仪 (SPECIM, V10E, Finland)、CCD 摄像头 (Andor, Clara, Britain)、暗箱、电控位移台等部件组成。采集猪肉样本的高光谱图像时,相关参数设置为:曝光时间 0.1 s,电控位移台运行速度 2.1 mm/s,扫描长度 120 mm,有效波长范围 391.21 ~ 1 043.14 nm,波长间隔 1.25 nm,波段个数 520。在所有样本图像采集之前,室温条件下将光谱仪预热 30 min 待照明强度稳定后再进行光谱数据采集。为了减少光谱仪暗电流和室内光线对图像的影响,采用黑物遮盖镜头和采用标准白板来获取黑、白板图像,对所有样本图像进行黑白校正^[13]。最后采用 ENVI 4.7 软件,选择感兴趣区域 (ROI),提取每个样本的反射光谱。

1.3 样本 pH 值理化分析

将猪肉样本绞碎,置于烧杯中,使用 pH 计测量 (METTLER TOLEDO, FE20, Switzerland) pH 值。测量前用缓冲液对 pH 计进行校准,测量时将电极深入样品中。每个样本测量 4 次,取平均值作为最终的 pH 值。

1.4 模型传递修正方法

模型的修正方法主要有斜率/截距 (S/B) 算法和模型更新算法。本文提出了光谱值校正传递算法,以实现模型在品种之间的传递。模型传递修正的评价指标主要是预测相关系数 R_p 和预测均方根误差。

S/B 算法首先将主品种模型对目标品种标准化样本的预测值与实测值用一条直线进行拟合得到直线的斜率和截距,作为模型的修正系数。然后利用该修正系数对目标品种待测样本预测值进行校正后作为真正的预测值^[14]。该算法只对预测结果进行校正,不需要改变模型和 R_p ,仅改变预测均方根误差。模型更新算法要添加新样本重新建立新模型,使原模型增加与新样本相适配的新信息,以形成新的模型。

品种之间模型的传递不同于仪器之间模型的传递,标准化样本组的选取不仅要考虑主、目标品种样本的光谱有差异,还要考虑主、目标品种样本的理化特性存在差异。光谱值校正传递算法首先需要选取一组标准化样本。标准化样本组的选取可采用 2 个品种样本之间的光谱-理化值共生距离进行选取,该距离能同时考虑到光谱差异与理化值差异。选取某个品种作为主品种,计算得到目标品种样本与主品种样本之间的光谱-理化值共生距离矩阵,由距离最短原则配对地选出标准化样本组。标准化样本组选取后对它们的理化值再进行方差分析以验证,若结果

表明标准样本组的理化值差异不显著,则表明标准样本组的理化值是一致的,差异主要来自于光谱之间的差异,故设法消除光谱的差异。本文利用直接校正法 (DS) 算法来消除它们之间的光谱差异^[15],以得到相关的转换系数矩阵,保存后对目标品种待测样本的光谱值进行修正。该传递算法具体步骤如下:

(1) 计算主品种和目标品种的光谱欧氏距离,计算公式为

$$d_x(a,b) = \sqrt{\sum_{i=1}^n [x_a(i) - x_b(i)]^2} \tag{1}$$

式中 $d_x(a,b)$ ——主品种第 a 个样本与目标品种第 b 个样本的光谱欧氏距离

$x_a(i)$ ——主品种第 a 个样本第 i 个波长变量的光谱值

$x_b(i)$ ——目标品种第 b 个样本第 i 个波长变量的光谱值

n ——波长变量的个数

(2) 计算主品种与目标品种样本的理化值距离。计算公式为

$$d_y(a,b) = \sqrt{(y_a - y_b)^2} \tag{2}$$

式中 $d_y(a,b)$ ——主品种第 a 个样本与目标品种第 b 个样本的理化值距离

y_a ——主品种第 a 个样本的理化值

y_b ——目标品种第 b 个样本的理化值

(3) 计算主品种样本与目标品种样本的光谱-理化值共生距离。计算公式为

$$d_{xy}(a,b) = \frac{d_x(a,b)}{\max_{a \in [1,N] \atop j \in [1,M]} d_x(a,j)} + \frac{d_y(a,b)}{\max_{a \in [1,N] \atop j \in [1,M]} d_y(a,j)} \tag{3}$$

式中 $d_{xy}(a,b)$ ——主品种第 a 个样本与目标品种第 b 个样本的光谱-理化值共生距离

N ——主品种样本数量

M ——目标品种样本数量

(4) 转换系数矩阵的计算:利用光谱-理化值共生距离最短原则从距离矩阵中选取与主品种样本光谱和理化值性质都相似的目标品种样本,一一对应组合作为标准化样本组,利用直接校正方法计算光谱转换系数矩阵 F 。

(5) 将待测目标品种的样本光谱修正为与主品种样本光谱和理化值相一致的光谱,修正公式为

$$X'_{\text{target}} = X_{\text{target}} F \tag{4}$$

式中 X_{target} ——待测目标品种样本的光谱

X'_{target} ——修正后的待测目标品种样本光谱

2 结果与分析

2.1 猪肉 pH 值定量检测模型的建立

采用蒙特卡罗交叉验证算法^[16]剔除异常样本,最终山黑猪样本和零号土猪样本各有 4 个和 3 个被剔除,表 1 为剔除异常样本后样本的 pH 值统计结果。

表 1 猪肉的 pH 值统计结果

品种	样本数量	平均值	最大值	最小值	标准差
山黑猪	78	5.581 0	6.410 0	5.240 0	0.214 5
零号土猪	76	5.588 8	6.352 5	5.242 5	0.221 4

选取山黑猪作为主品种,零号土猪作为目标待测品种,采用 SPXY 算法按 2:1 比例选取 53 个山黑猪样本作为校正集,25 个样本作为验证集,用于建立山黑猪 pH 值定量分析模型。同时采用 SPXY 算法选取 40 个零号土猪样本用于检测山黑猪模型的适用性。剩余的 36 个零号土猪样本用于选取 S/B 算法的标准化样本,或在模型更新算法中添加到山黑猪模型校正集进行修正。山黑猪样本光谱经过多种预处理方法处理后,采用 PLS 建模方法建立检测模型。经过比较确定变量标准化(Normalize)为最优预处理方法,所建立的模型交叉验证相关系数 R_c 及预测相关系数 R_p 分别为 0.886 和 0.864,交叉验证均方根误差和预测均方根误差分别为 0.112 9 和 0.105 9。由此可知,所建立的山黑猪模型对本品种验证样本的预测效果良好。

2.2 山黑猪模型的适用性检测

每个校正模型都有一定的适用范围,当待测样本的物化信息超出了模型所适应的范围,则原模型将不能很好地预测新样本的品质信息,需对模型进行修正和传递。采用山黑猪样本与零号土猪样本的主成分得分空间分布图可以直观地考察校正模型的适用性^[17-18],如图 1 所示。由图 1 可知,零号土猪样本的主成分覆盖范围超出了山黑猪样本主成分所能涵盖的范围,说明山黑猪样本模型的适用范围较

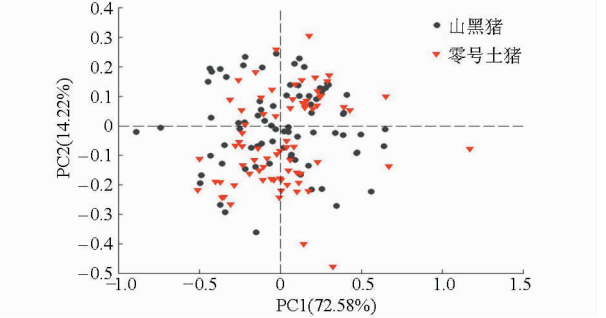


图 1 山黑猪与零号土猪样本主成分得分
Fig. 1 PCA scores of Shanhei and Linghao pig samples

小,而零号土猪样本需要较大的适用范围,故建立的山黑猪模型不能较好地预测零号土猪样本。

为了进一步量化地描述模型的适用性,采用 Fisher 值^[19]对不同品种样本之间的光谱差异进行分析。Fisher 值用 $F(i)$ 表示,其计算公式为

$$F(i) = \frac{\sum_{i=1}^k n_i (\bar{x}_{ij} - \bar{x}_{.j})^2}{\sum_{i=1}^k ((n_i - 1) S_{ij}^2)} \quad (5)$$

式中 k ——品种数
 n_i ——第 i 个品种的样本数
 $\bar{x}_{ij}$ ——第 i 个品种中所有样本在波长 j 中的反射率平均值
 $\bar{x}_{.j}$ ——所有样本(包括所有品种)在波长 j 中的反射率平均值
 S_{ij} ——第 i 个品种样本在波长 j 处的标准偏差

Fisher 值表示品种间方差与品种内方差之比,当 $F(i) < 1$,表明在该波长点品种内方差要大于品种间方差。 $F(i) > 1$,表明在该波长点品种间方差大于品种内方差,这些波长点的存在不利于模型在品种之间的适用性,且 $F(i)$ 值越大,产生的不利影响越大。

图 2 和图 3 反映了山黑猪与零号土猪样本原始光谱的 Fisher 值图和 Normalize 预处理光谱的 Fisher 值图。由图 2 可知,原始光谱 Fisher 值大于 1 的区域主要在光谱的 390 ~ 430 nm、610 ~ 910 nm、1 020 ~ 1 040 nm 范围,说明在这些波长区域内,品种间方差要大于品种内方差,两品种光谱差异比较大。对于

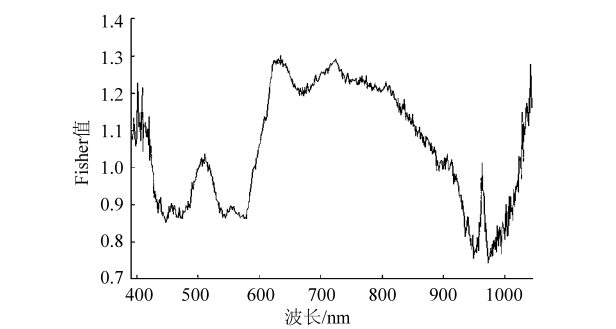


图 2 原始光谱的 Fisher 值
Fig. 2 Fisher value of original spectra

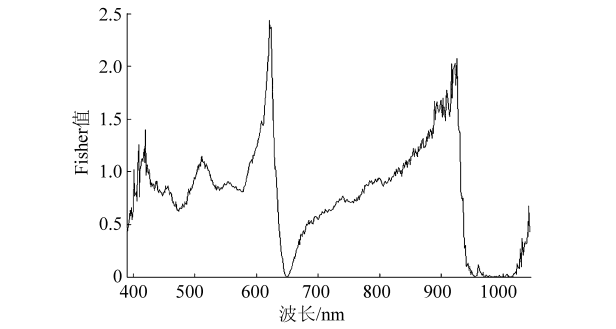


图 3 Normalize 预处理光谱的 Fisher 值
Fig. 3 Fisher value of the data obtained from Normalize

经过 Normalize 预处理后的光谱,在 400 ~ 430 nm、505 ~ 525 nm、580 ~ 630 nm、830 ~ 930 nm 区间内,共有 146 个波长点的 Fisher 值都大于 1,表明在这些光谱区域内,品种间的差异要大于同品种内的差异,故这些波长点在一定程度上影响了山黑猪模型的适用性,降低了山黑猪模型对零号土猪样本的预测性能。

利用 2.1 节中所建立的山黑猪模型对 40 个零号土猪样本进行预测,预测相关系数 R_p 为 0.415,预测均方误差达到 0.180 4,预测误差较大,需对山黑猪模型进行传递修正,分析结果与根据主成分得分分布图、Fisher 值图分析得出的结果相一致。

2.3 基于斜率/截距算法的模型修正

从 2.1 节中的 36 个零号土猪样本中,依次随机选取标准化样本个数 5、10、15、20、25、30、35。山黑猪模型分别对其 pH 值进行预测后,采用 S/B 算法得到斜率和截距,作为修正系数。由于 S/B 算法不改变山黑猪模型对零号土猪样本的预测相关系数 $R_p^{[20]}$,故 R_p 仍为 0.415,只需利用修正系数对 40 个零号土猪样本的 pH 值预测值进行修正后,计算修正后的预测值与实测值之间的预测均方根误差。不同标准化样本数量的修正结果如图 4 所示。

从图 4 可知,经过 S/B 算法修正后,预测均方根误差有一定程度上降低,当标准化样本个数为 10 时,预测均方根误差达到最小值,为 0.134 3。与模型修正之前的 0.180 4 相比,预测均方根误差仅下降了 25.54%,由此可见 S/B 算法对山黑猪模型的修正效果有限。

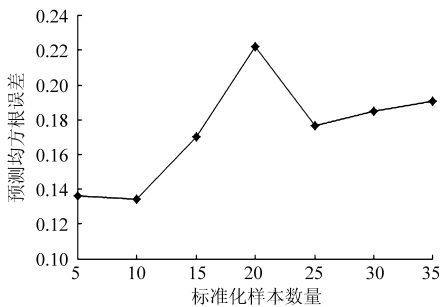


图 4 不同标准化样本数目的预测均方根误差
Fig. 4 RMSEP with different number of standardization samples

2.4 基于模型更新算法的模型修正

采用 SPXY 算法从 36 个零号土猪样本中,选取不同数量的样本添加到山黑猪校正集对山黑猪模型进行修正,并对 40 个零号土猪样本进行预测,结果如图 5 所示。从图 5 可知,直接用山黑猪原模型来预测零号土猪样本时,即没有向山黑猪校正集添加零号土猪样本,预测相关系数 R_p 为 0.415,预测均方根误差为 0.180 4,预测出现较大的偏差。随着向原

模型校正集添加新样本数量的增加, R_p 逐渐在增加,预测均方根误差逐渐减低。当零号土猪样本的添加数目为 14 时,修正后的模型对零号土猪样本的预测性能得到了较大的改善, R_p 由 0.415 提高至 0.797,提高了 92.05%,预测均方根误差由 0.180 4 降低为 0.112 1,下降 37.86%。此后,随着添加数量的增加,修正后的模型对零号土猪样本的预测性能趋向稳定。由此可知,向山黑猪原模型的校正集添加零号土猪样本修正后,能改善山黑猪模型对零号土猪样本的预测性能,提高模型的适用性。

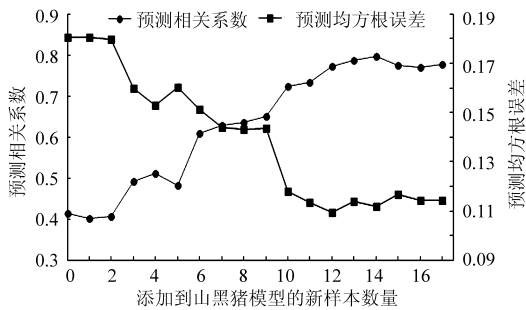


图 5 添加新样本修正后的山黑猪模型预测效果
Fig. 5 Prediction results of Shanhei pig model based on a variable number of new samples

为了检验修正后的山黑猪模型对山黑猪样本的预测性能是否因为加入新样本后变差,故用添加新样本数目为 14 个的山黑猪修正模型重新预测山黑猪验证集中的 25 个样本。预测相关系数 R_p 有所下降,由 0.864 降低为 0.798,同时预测均方根误差略有增加,预测均方根误差由 0.105 9 增加为 0.107 1。由此表明,添加新样本的修正模型对主品种验证集的预测性能在一定程度上可变量。

2.5 基于光谱值校正传递算法的模型传递

利用式(3)计算得到山黑猪样本与零号土猪样本的光谱-理化值共生距离矩阵。首先采用 SPXY 算法选出不同数量的山黑猪标准化样本,山黑猪标准化样本数目取值范围为 10 ~ 50 个,间隔为 5,结果如表 2 所示。然后从距离矩阵中挑选出与每一个山黑猪标准化样本距离为最短的零号土猪样本,一一对应组成标准化样本组。选择出标准化样本组后,需要对标准化样本组中山黑猪与零号土猪样本的理化值进行方差分析,分析结果见表 2,每个标准化样本组的方差分析概率 P 值都大于 0.05,表明标准化样本组的理化值差异不显著,故验证了由光谱值-理化值共生距离法选择出来的两品种样本的理化值是一致的。通过光谱值校正传递算法得到转换系数,对 40 个零号土猪样本光谱进行转换后,用 2.1 节中的山黑猪模型对其重新进行预测。

由表 2 可知,采用 DS 法能有效消除光谱的差

异。当山黑猪标准化样本数目过少时,有效信息较少,而山黑猪标准化样本数目过多时,引进了冗余信息,故传递效果均不佳。当标准化样本数目为 35 时,预测均方根误差最小,为 0.085 6,预测相关系数 R_p 为 0.837。采用光谱值校正传递算法校正后,山黑猪模型对零号土猪样本的预测相关系数 R_p 由

0.415 提高至 0.837,提高了 101.69%,预测均方根误差由 0.180 4 降低至 0.085 6,下降了 52.55%,模型的适用性得到了较大改善和提高。由此表明,通过光谱-理化值共生距离法结合 DS 算法的光谱值校正传递算法能有效提高山黑猪模型的适用性,实现了模型在品种之间的传递。

表 2 不同标准化样本组的 R_p 、预测均方根误差以及方差分析 P 值
Tab.2 R_p , RMSEP and P value of variance analysis of different number of standardization samples

参数	标准化样本数量								
	10	15	20	25	30	35	40	45	50
R_p	0.150	0.300	0.666	0.738	0.822	0.837	0.827	0.802	0.783
预测均方根误差	0.344 5	0.274 7	0.172 7	0.121 5	0.095 5	0.085 6	0.091 1	0.098 7	0.104 1
理化值方差分析 P 值	0.836 6	0.920 4	0.929 3	0.893 4	0.937 1	0.923 5	0.918 5	0.757 5	0.745 0

2.6 算法的传递修正效果比较

表 3 列出了 3 种算法对山黑猪模型传递修正结果的比较。结果表明:首先,3 种算法在一定程度可改善山黑猪模型对零号土猪样本的预测性能,但光谱值校正传递算法的改善效果最优。其次,光谱值校正传递算法和 S/B 算法均不需要改变山黑猪原模型,不同的是光谱值校正传递算法是通过消除品

种之间的光谱值差异而达到模型传递的效果,S/B 算法则是对预测结果的校正,而前者更为有效。再者,虽然模型更新法使山黑猪修正模型对零号土猪样本的预测性能得到了改善,但是修正后的山黑猪模型与原模型相比,其 R_c 降低,交互验证均方根误差增加,对山黑猪样本预测相关系数 R_p 降低,预测均方根误差增加,模型的预测性能变差。

表 3 不同算法对山黑猪模型传递修正效果的比较
Tab.3 Comparison of the transfer and correction effect of different algorithms to the Shanhei pig model

模型	模型校正集		山黑猪预测集		零号土猪预测集	
	R_c	交互验证均方根误差	R_p	预测均方根误差	R_p	预测均方根误差
山黑猪原模型	0.886	0.112 9	0.864	0.105 9	0.415	0.180 4
S/B 算法修正后的模型	0.886	0.112 9	0.864	0.105 9	0.415	0.134 3
模型更新法修正后的模型	0.819	0.154 5	0.798	0.107 1	0.797	0.112 1
光谱值校正传递算法传递后的模型	0.886	0.112 9	0.864	0.105 9	0.837	0.085 6

2.7 讨论

为了进一步验证光谱值校正传递算法的有效性,采用 2013 年 10 月份采集的不同批次的 20 个零号土猪样本来再次验证。直接用 2.1 节中的山黑猪原模型预测零号土猪样本时, R_p 为 0.187,预测均方根误差达到 0.408 9。当采用 S/B 算法,取 5 个 2013 年 10 月份的零号土猪样本作为标准化样本时, R_p 仍为 0.187,预测均方根误差为 0.197 1;采用模型更新方法,添加 6 个 2013 年 10 月份的零号土猪样本时, R_p 为 0.230,预测均方根误差为 0.351 1;当采用光谱值校正传递算法,标准化样本数目为 15 时, R_p 提高至 0.700,预测均方根误差降低为 0.149 3。结果表明,虽然光谱值校正传递算法应用于 2013 年 10 月份的样本数据时,相关指标 R_p 与预测均方根误差要比表 3 中应用于 2012 年 12 月份样本时略低,这是由于不同批次样本的系统误差(如仪器调节、化学测量值等不是同一人完成)所造成的,但是传

递效果仍然明显优于 S/B 算法、模型更新算法两者的修正效果,表明光谱值校正传递算法用于模型在不同批次不同品种之间的传递是有效的,进一步作深入研究可以提高模型的适用性。

3 结论

(1)利用高光谱数据建立了山黑猪 pH 值的 PLS 定量分析模型,其 R_c = 0.886,交叉验证均方根误差为 0.112 9, R_p = 0.864,预测均方根误差为 0.105 9,结果表明高光谱成像技术可用于预测猪肉的 pH 值。然而,用该模型去预测另一个品种零号土猪样本时,模型的预测性能出现了较大的偏差, R_p 仅为 0.415,预测均方根误差达到 0.180 4,模型的适用性较差。

(2)针对山黑猪模型对零号土猪样本适用性差的问题,提出了光谱-理化值共生距离法结合 DS 算法的光谱值校正传递算法。该算法不改变山黑猪原

模型,而且由光谱-理化值共生距离最短原则选取的标准化样本组能同时兼顾光谱和理化值性质的差异,利用了DS算法的优势,改善了山黑猪模型对零号土猪待测样本的预测性能,实现了模型在品种之间的传递。当采用该传递算法时,山黑猪模型对零号土猪样本的预测相关系数 R_p 由0.415提高至0.837,提高了101.69%,预测均方根误差由0.1804

降低为0.0856,降低了52.55%。
(3)将光谱值校正传递算法与S/B算法、模型更新法进行了比较,结果表明,光谱值校正传递算法能有效消除品种对高光谱模型的影响,比S/B算法和模型更新算法有更好的模型传递修正能力,能较好地提高pH值高光谱预测模型的适用性。

参 考 文 献

- 1 张学博,冯艳春,胡昌勤.近红外多元校正模型传递的进展[J].药物分析杂志,2009,29(8):1390-1399.
Zhang Xuebo, Feng Yanchun, Hu Changqin. Progress in calibration transfer of near-infrared multivariate model[J]. Journal of Pharmaceutical Analysis, 2009, 29(8): 1390-1399. (in Chinese)
- 2 褚小立,袁洪福,陆婉珍.光谱多元校正中的模型传递[J].光谱学与光谱分析,2001,21(6):881-885.
Chu Xiaoli, Yuan Hongfu, Lu Wanzhen. Model transfer in multivariate calibration [J]. Spectroscopy and Spectral Analysis, 2001, 21(6):881-885. (in Chinese)
- 3 严衍祿.近红外光谱分析基础与应用[M].北京:中国轻工业出版社,2005.
- 4 Schimleck L R, Kube P D, Raymond C A, et al. Extending near infrared reflectance (NIR) pulp yield calibration to new sites and species [J]. Journal of Wood Chemistry and Technology, 2006, 26(4):299-311.
- 5 Camps C, Christen D. Non-destructive assessment of apricot fruit quality by portable visible-near infrared spectroscopy[J]. LWT-Food Science and Technology,2009,42(6):1125-1131.
- 6 潘璐,王加华,李鹏飞,等.砂梨糖度近红外光谱波段遗传算法优化[J].光谱学与光谱分析,2009,29(5):1246-1250.
Pan Lu, Wang Jiahua, Li Pengfei, et al. Region optimization of SSC model for pyrus pyrifolia by genetic algorithm [J]. Spectroscopy and Spectral Analysis,2009, 29(5):1246-1250. (in Chinese)
- 7 文东东,李小昱,赵政,等.不同品种牛肉新鲜度光谱检测模型的维护方法[J].食品安全质量检测学报,2012,3(6):621-626.
Wen Dongdong, Li Xiaoyu, Zhao Zheng, et al. Maintenance methods for beef freshness detection model of different species based on spectrum[J]. Journal of Food Safety and Quality, 2012, 3(6): 621-626. (in Chinese)
- 8 褚小立,袁洪福,王艳斌,等.近红外稳健分析校正模型的建立(I)-样品温度的影响[J].光谱学与光谱分析,2004,24(6):666-671.
Chu Xiaoli, Yuan Hongfu, Wang Yanbin, et al. Developing robust near infrared calibration models[J]. Spectroscopy and Spectral Analysis, 2004, 24(6):666-671. (in Chinese)
- 9 谢丽娟,应义斌.转基因番茄鉴别模型维护方法[J].江苏大学学报:自然科学版,2012,33(5):538-542.
Xie Lijuan, Ying Yibin. Priliminary research on calibration model maintenance method for transgenic tomato discriminant [J]. Journal of Jiangsu University, 2012, 33(5): 538-542. (in Chinese)
- 10 Candolfi A, Massart D L. Model updating for the identification of NIR spectra from a pharmaceutical excipient[J]. Applied Spectroscopy, 2000, 54(1): 48-53.
- 11 姚胜,武国峰,姜亦飞,等.相思树聚戊糖含量近红外光谱分析模型的建立及修正[J].光谱学与光谱分析,2010,30(5):1206-1209.
Yao Sheng, Wu Guofeng, Jiang Yifei, et al. Extending hemicelluloses content calibration of *Acacia* spp using NIR to new sites [J]. Spectroscopy and Spectral Analysis, 2010, 30(5): 1206-1209. (in Chinese)
- 12 付庆波,索辉,贺馨平,等.温度影响下短波近红外酒精度检测的传递校正[J].光谱学与光谱分析,2012,32(8):2080-2084.
Fu Qingbo, Suo Hui, He Xinping, et al. Transfer calibration for alcohol determination using temperature-induced shortwave near infrared spectra[J]. Spectroscopy and Spectral Analysis, 2012, 32(8):2080-2084. (in Chinese)
- 13 王伟,彭彦昆,张晓莉.基于高光谱成像的生鲜猪肉细菌总数预测建模方法研究[J].光谱学与光谱分析,2010,30(2):411-415.
Wang Wei, Peng Yankun, Zhang Xiaoli. Study on modeling method of total viable count of fresh pork meat based on hyperspectral imaging system[J]. Spectroscopy and Spectral Analysis, 2010, 30(2):411-415. (in Chinese)
- 14 肖武,李小昱,李培武,等.基于近红外光谱土壤水分检测模型的适应性[J].农业工程学报,2009,25(3):33-36.
Xiao Wu, Li Xiaoyu, Li Peiwu, et al. Adaptability of the model for soil moisture measurement based on near-infrared spectroscopy[J]. Transactions of the CSAE, 2009, 25(3): 33-36. (in Chinese)
- 15 王菊香,李华,邢志娜,等.波长筛选结合直接校正法用于近红外光谱模型传递研究[J].分析测试学报,2011,30(1):43-47.
Wang Juxiang, Li Hua, Xing Zhina, et al. Application of wavelength selection- direct standarization algorithm to calibration

- transfer of near infrared analysis[J]. Journal of Instrumental Analysis, 2011,30(1):43–47. (in Chinese)
- 16 Cao Dongsheng, Liang Yizeng, Xu Qingsong, et al. A new strategy of outlier detection for QSAR/QSPR[J]. Journal of Computational Chemistry, 2010,31(3):592–602.
- 17 徐广通, 袁洪福, 陆婉珍. 近红外光谱定量校正模型适用性研究[J]. 光谱学与光谱分析, 2001, 21(4):459–463.
Xu Guangtong, Yuan Hongfu, Lu Wanzhen. Study of quantitative calibration model suitability in near-infrared spectroscopy analysis[J]. Spectroscopy and Spectral Analysis, 2001,21(4):459–463. (in Chinese)
- 18 刘泽春, 张峰, 谢卫. 烟草近红外光谱模型的适配性研究[J]. 烟草化学, 2008(5):34–37.
Liu Zechun, Zhang Feng, Xie Wei. Study on suitability of near-infrared spectroscopy model for tobacco[J]. Tobacco Science & Technology, 2008(5):34–37. (in Chinese)
- 19 Candolfi A, Wu W, Massart D L, et al. Comparison of classification approaches applied to NIR-spectra of clinical study lots[J]. Journal of Pharmaceutical and Biomedical Analysis, 1998,16(8):1329–1347.
- 20 胡润文, 夏俊芳. 脐橙总糖近红外光谱模型传递研究[J]. 食品科学, 2012, 33(3):28–32.
Hu Runwen, Xia Junfang. Transfer of NIRS calibration model for determining total sugar content in navel orange[J]. Food Science, 2012, 33(3):28–32. (in Chinese)

Comparison of Transfer and Correctional Methods for Pork pH Value Detection of Different Varieties by Hyperspectral Imaging Technique

Li Xiaoyu¹ Zhong Xiongbai¹ Liu Shanmei² Huang Tao¹ Wu Zhenzhong¹

(1. College of Engineering, Huazhong Agricultural University, Wuhan 430070, China

2. College of Sciences, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: The calibration model of the pork quality based on hyperspectral data was easily affected by pork varieties and had poor applicability. So different algorithms were compared and a calibration transfer algorithm based on spectral value correction was proposed to improve the model applicability between different varieties. The PLS model for Shanhei pig was established, then was used to predict Linghao pig samples. Prediction accuracy was very poor with only R_p of 0.415, and RMSEP of 0.180 4. To improve model applicability, S/B algorithm, model updating and transfer algorithm based on spectral value correction were respectively adopted and compared. For S/B algorithm, R_p of the model prediction for Linghao pig samples was still 0.415, and RMSEP decreased from 0.180 4 to 0.134 3, only down by 25.54%. For the model updating method, when 14 Linahao pig samples were added to the calibration dataset, the model prediction performance for Linghao pig samples achieved optimal. R_p increased to 0.797, improved by 92.05%, and RMSEP reduced to 0.112 1, dropped by 37.86%. For transfer algorithm based on spectral value correction which combined spectral value-physical value coexist distance with DS algorithm, R_p of the model prediction for Linghao pig samples increased to 0.837, rose by 101.69%, and RMSEP reduced to 0.0856, fell by 52.55%. The results showed that the transfer algorithm based on spectral value correction could eliminate the difference of spectral value of different varieties efficiently and improve the model applicability. It gave the best transfer and correction result than the two other algorithms.

Key words: Pork pH value Hyperspectral detection Model transfer and correction Spectral correction