

利用可见近红外光谱多指标综合预测生鲜牛肉储存期

马世榜^{1,2}, 徐 杨^{1*}, 汤修映¹, 田潇瑜¹, 付 姓¹

1. 中国农业大学工学院, 北京 100083
2. 南阳理工学院, 河南 南阳 473004

摘 要 建立了牛肉基于 TVB-N、菌落总数、pH 值和肉色参数 L^* 多个指标的储存期预测模型, 利用可见近红外光谱(Vis/NIR)技术结合区间偏最小二乘(iPLS)和遗传算法(GA)建立了各个指标的 PLS 预测模型, 实现了多指标综合无损快速预测 4 °C 下牛肉的储存期。用 iPLS 和 iPLS-GA 提取有效波长变量建立 PLS 预测模型, 以预测相关系数和预测标准差作为模型评价标准, 结果表明用 iPLS-GA 选择变量建立的各个指标的 PLS 预测模型均优于全波段和 iPLS 组合的 PLS 模型。由多个指标的预测值和储存期的预测模型, 对校正集和预测集样品储存期进行预测, 其预测相关系数和标准差分别是 0.903, 0.897 和 1.88, 2.24。说明利用光谱技术结合得出的储存期预测模型可以实现多指标综合预测牛肉储存期, 为无损快速检测牛肉储存期或货架期提供了一种新方法。

关键词 可见近红外光谱; 牛肉储存期; 多指标检测; 变量选择; 偏最小二乘

中图分类号: S123 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2012)12-3242-05

引 言

牛肉是消费量较大的主要肉品之一, 其品质安全越来越受重视。生鲜牛肉色泽鲜艳光亮, 肉质呈中性或弱碱性。随着储存期的增长, 微生物繁殖增多, 生成的氨及胺类碱性含氮物质(简称挥发性盐基氮, TVB-N)增大, pH 值升高, 肉色变暗, 影响其品质及食用安全, 因此预测判断生鲜牛肉的储存期具有重要意义。但感官或理化实验判断生鲜牛肉的储存期较困难, 且存在人为因素影响大、效率低、周期长, 常需要对检测样品进行破坏处理, 不能实现快速无损检测。可见近红外光谱(visible and near infrared spectroscopy, Vis/NIR)分析技术因具有分析速度快、样品无需预处理、非破坏性、测试重现性好等特点^[1], 已经广泛应用于对肉品品质检测的研究^[2], Andrés^[3]等用可见近红外光谱预测了屠宰后储存期不同的牛肉的 pH 值、肉色; Prieto^[4]等用近红外光谱预测了牛肉的肉色、持水性、嫩度等; Holmer^[5]等研究了猪肉 pH 值、肉色与货架期间有较好的回归关系。但单个指标很难全面准确反映肉的储存期。用 Vis/NIR 同时预测牛肉 TVB-N、细菌总数、pH 值、肉色多个指标, 进而预测牛肉的储存期, 还未见有相关报道。本工作研究屠宰后保存在 4 °C

环境下生鲜牛肉储存期与 TVB-N、细菌总数、pH 值和肉色参数 L^* 多指标之间的关系, 通过 Vis/NIR 技术预测上述指标, 实现多指标综合无损快速预测生鲜牛肉储存期的目的。

1 实验部分

1.1 材料

实验样品从北京市御香苑畜牧有限公司取自屠宰后经 48 h 解僵、排酸的不同东北改良牛的里脊部位, 用无菌刀分割为约 8 cm×6 cm×2.5 cm 尺寸大小的 56 块, 用保鲜袋包装放置于 4~6 °C 低温保鲜箱, 运至实验室后整齐、无挤压的保存在 4 °C 冰箱中, 模拟市场上待售生鲜牛肉放置环境。实验周期为 17 d, 前 6 d 每隔 12 h 检测一块样品, 后 11 d 每隔 12 h 检测两块样品。实验时样品是随机从冰箱中取出。

1.2 仪器

实验用可见近红外光谱检测系统主要由暗箱、光源、载物台、光谱仪(Avantes, Holland)、计算机等组成。光谱波长测量范围为 200~1 750 nm, 最小采样间隔为 0.5 nm。光源为两个卤钨灯(12 V, 20 W)。此外还包括 KDY-9820 凯氏定氮仪、pH 计、精密色差仪、立式压力蒸汽灭菌器、超净工作台、电恒温培养箱、肉搅拌机、电子天平、食品样品安全蓄

收稿日期: 2012-05-17, 修订日期: 2012-09-18

基金项目: 国家自然科学基金项目(30771244)和公益性行业(农业)科研经费项目(201003008)资助

作者简介: 马世榜, 1978 年生, 中国农业大学工学院博士研究生 e-mail: mshibang@126.com

* 通讯联系人 e-mail: xuyang@cau.edu.cn

冷藏运箱、培养皿、量筒、PE 保鲜膜(100 型)等实验仪器设备和器材。

1.3 方法

1.3.1 光谱采集

采集 350~1 700 nm 波长范围的牛肉样品表面反射光谱。每个样品采集光谱前,把样品从冰箱中取出,去除保鲜袋,在空气中暴露 30 min,使样品表面水分自然挥发;调节载物台高度使样品表面距离光纤探头高度为 12 cm,每个样品采集 6 个不同点的反射光谱,取 6 个点的光谱平均值作为该样品的最终反射光谱值。

1.3.2 挥发性盐基氮的测量

按照 GB/T5009.44-2003 中的半微量定氮法,用 KDY-9820 凯氏定氮仪进行测量 TVB-N。每次测量前,做三个样品空白实验,取平均值作为样品空白实验值,每个样品做三个样液实验,取平均值作为该样品样液实验参考值。

1.3.3 菌落总数的测量

按照 GB/4789.2-2010 测量菌落总数(colony total)法,采用平板计数法,以菌落形成单位(colony-forming units, CFU·g⁻¹)表示每 g 样品中形成的微生物菌落总数。整个实验周期内,每个样品匀液做两个空白对照,按照 10 倍梯度方法稀释 3~4 个适宜稀释度,每个稀释度做两个平皿,在(36±1)℃的恒温培养箱里培养(48±2)h 后计数,取菌落总数的对数值作为实验参考值,用 logCFU·g⁻¹表示。

1.3.4 pH 值的测量

用 pH 计测量 pH 值。测量前用温度计测量样品温度和 pH 计标定缓冲液的温度,保证两者温度一致;用缓冲液对 pH 计进行三点标定。每个样品重复测量不同的点 6 次,取 6 次的平均值作为该样品的最终 pH 参考值。

1.3.5 肉色参数的测量

用精密色差仪测量 CIE(1976)颜色参数 L* (L* 代表亮度变量)。测量前用标准黑腔和白板对仪器进行校正。每个样品表面测量不同点 6 次,取 6 次的平均值作为样品颜色参数 L* 的参考值。

对四个指标参考值进行统计分析,设参考值绝对误差的阈值为样本集平均误差的 2 倍,大于阈值说明该样本参考值异常,应以剔除,TVB-N 和菌落总数参考值分别剔除 1 个和 3 个异常值样本。TVB-N、菌落总数、pH 值和 L* 指标参考值变化范围如表 1 所示。

Table 1 Statistical values of beef physical and chemistry indexes

Index(units)	Max	Min	Mean	Standard deviation
TVB-N(mg·(100 g) ⁻¹)	49.85	8.64	27.21	10.95
Colony total(log CFU·g ⁻¹)	9.98	7.23	8.68	0.54
pH	6.39	5.49	5.93	0.25
L*	49.30	31.89	38.82	5.11

1.4 数据处理及建模方法

建立 TVB-N、菌落总数、pH 值、L* 参数四个指标与储存期之间的回归模型和指标各自的光谱预测模型。为消除高

频噪声干扰、图谱偏移或漂移现象,用多元散射校正(MSC)、标准归一化(SNV)、Savitzky-Golay(SG)平滑对原始光谱进行预处理,确定最佳预处理方法。为减少无用冗余信息,提取有效光谱变量,提高光谱预测模型稳定性和预测能力,尝试采用区间偏最小二乘(iPLS)和遗传算法(GA)对全波段光谱变量进行选择,建立各指标的偏最小二乘(PLS)预测模型。具体方法实现采用 MATLAB7.6(Mathworks Inc.)编程。用所建模型的预测值与参考值的相关系数和预测值标准差作为模型评价标准,相关系数越高,标准差越小,模型的预测效果越好。

2 结果与讨论

2.1 储存期预测模型的建立

实验所测样品 TVB-N、菌落总数、pH 值随储存期的增长,呈增大趋势,L* 呈下降趋势,与样品的感官变化一致。单指标菌落总数呈指数增长变化,但多个指标间存有耦合性,Grau^[6]等研究生鲜鸡肉储存期与 TVB-N、pH 值、菌落总数之间呈现很好的线性变化关系。作储存期 d(天)与四个指标变量之间的多元线性回归,可得储存期预测模型如下式

$$d = -24.4336 + 0.2089X_{\text{TVB-N}} + 1.5068X_{\log\text{CFU}\cdot\text{g}^{-1}} + 3.0708X_{\text{pH}} - 0.0483X_{L^*} \quad (1)$$

式(1)中 d 是储存期,X 是对应的指标变量。对回归方程进行显著水平 $\alpha=0.05$ 的 F 检验,回归方程显著通过 F 检验,即储存期与 TVB-N、菌落总数、pH 值、L* 参数指标之间有显著的线性关系。用该回归模型预测样品储存期与实际储存期之间的关系如图 1 所示,预测相关系数为 0.935,标准差为 1.65。从预测结果和分布图可知在储存期为 14 d 以后的预测误差较大。分析原因可能是由于生鲜牛肉储存后期各指标达到一定极限后变化较缓慢的缘故。考虑到生鲜牛肉储存在 4℃下 14 d 以后已经腐败无法食用的实际情况,该模型可以满足实际储存期的预测应用。

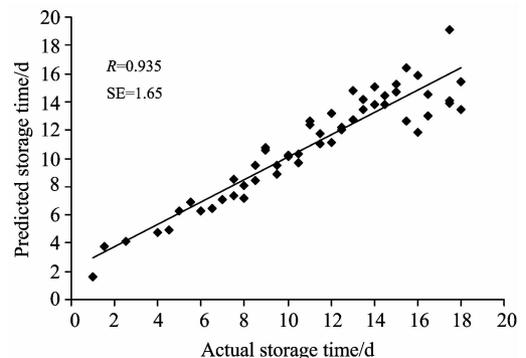


Fig. 1 Predicted vs actual values of storage time

2.2 光谱数据处理分析

对全样品光谱数据进行主成分分析,用光谱杠杆值(leverage)检验全部样品中的光谱异常值,剔除 3 个异常光谱样品,原始光谱曲线如图 2(a)所示。从图上可以看出在波长 430, 660, 925, 1 230, 1 580 nm 附近有较大的吸收峰,430 nm 附近主要是脱氧肌红蛋白的吸收峰,660 nm 附近可

能主要是 NH_3 基团的 3 级倍频吸收, 1 570 nm 附近主要源于 N—H 基团的第 1 倍频吸收, 说明光谱反映了引起 TVB-N、菌落总数、pH 值、肉色变化的信息。对四个指标多次对比经 MSC, SNV 和 SG 三种预处理光谱方法后所建立的 PLS

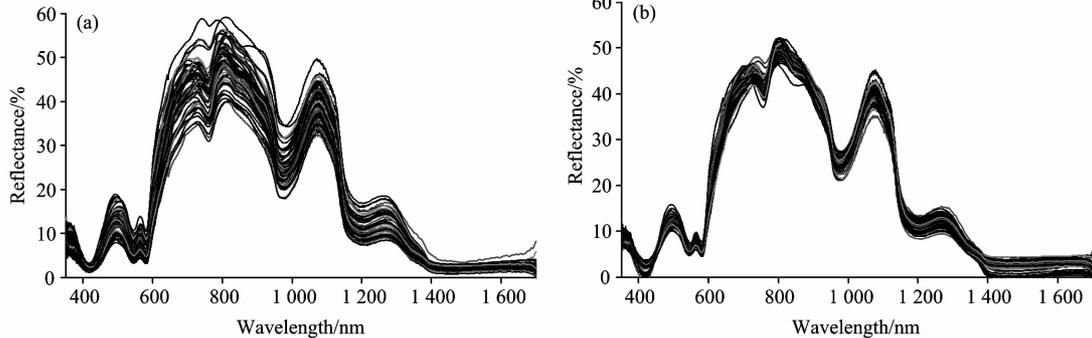


Fig. 2 (a) Raw Vis/NIR spectra of beef; (b) MSC preprocess Vis/NIR spectra of beef

2.3 指标预测模型的建立及评价

2.3.1 iPLS 选择变量及建模

全波段光谱中包含大量无用信息, 光谱变量间可能存在共线性, 降低模型稳定性和预测能力。Norgaard^[7]提出了区间偏最小二乘法(interval PLS, iPLS)对全光谱波段进行筛选, 蔡健荣^[8]等应用 iPLS 方法预测猪肉 TVB-N 得到较好效果。其基本原理是把全波段分为相等的多个子区间波段, 分别建立各个子区间波段和全波段的偏最小二乘(PLS)预测模型, 比较各子区间波段模型与全波段模型预测均方根误差(RMSE), 最小误差子区间用于最终建模。但单个子区间很难全面反映所有有用光谱信息, 我们把全波段分为 20 个区间, 用 iPLS 方法取小于或接近全波段模型 RMSE 的联合子区间分别建立 TVB-N、菌落总数、pH 值和 L^* 参数四个指

标 PLS 预测建模。建模时按 3:1 原则从总样品集选出 14 个样品作为预测集, 余下的作为校正集用于校正模型的建立, 预测集用于模型的验证。对四个指标用 iPLS 选择出来的联合子区间分别是 [1 5 7 11 12], [1 2 3 6 8 9 10 11 14~20], [1 2 4 5 6 7 11 19 20], [1 2 3 4 5 6 9 19 20], 用全波段和 iPLS 选择出的联合子区间分别建立四个指标的 PLS 模型, 预测结果如表 2 所示。其中 R_c 和 R_p 分别是校正集和预测集的相关系数, SEC (standard error of calibration) 和 SEP (standard error of prediction) 分别是校正集和预测集的标准差。从结果可以看出, 四个指标用全波段光谱变量所建 PLS 预测模型, 不同程度存在有过拟合现象, 模型稳定性差。用 iPLS 选择子区间后建模, 光谱变量个数减少, PLS 模型主成分数减少, 过拟合现象降低, 预测精度得到提高。

Table 2 PLS calibration models for prediction of 4 indexes using different methods for selection of wavelength

Selection methods	Number of variable	PLS components	TVB-N				Number of variable	PLS components	Colony total			
			R_c	SEC	R_p	SEP			R_c	SEC	R_p	SEP
PLS	1 569	12	0.939	4.107	0.762	9.167	1 569	11	0.932	0.322	0.768	0.451
iPLS	419	11	0.964	3.212	0.856	8.301	1 243	10	0.937	0.320	0.876	0.395
iPLS-GA	95	10	0.952	3.011	0.943	4.742	156	8	0.920	0.326	0.903	0.384
Selection methods	Number of variable	PLS components	pH				Number of variable	PLS components	L^*			
			R_c	SEC	R_p	SEP			R_c	SEC	R_p	SEP
PLS	1 569	13	0.994	0.026	0.873	0.087	1 569	7	0.985	0.864	0.891	3.023
iPLS	748	13	0.993	0.032	0.892	0.079	665	7	0.942	1.758	0.932	2.482
iPLS-GA	87	10	0.976	0.065	0.930	0.070	110	4	0.946	1.604	0.952	1.477

2.3.2 GA 选择变量建模

经 iPLS 选择出的各个子区间的变量间可能还存在共线性, 对联合子区间继续用遗传算法(GA)进行变量选择(简称 iPLS-GA)。GA 是模拟生物进化过程, 以目标适应度函数为判据, 通过不断对群体进行选择、交叉、变异遗传操作, 重组优化群内个体, 实现特征变量优选的方法^[9], 优选过程是以所选特征波长变量建立 PLS 模型的交叉验证标准差最小为目标函数进行计算^[10]。GA 方法是一种有效的光谱变量选择方法, 在多种光谱检测方面得到应用^[11-15]。GA 群体大小

设为 30 个个体, 每个个体基因变量设为 17 个, 变异概率和交叉概率分别设为 1% 和 50%, 遗传代数设为 100 代。由于 GA 初始群体是随机产生的, 因此对每个指标都重复 10 次筛选过程建立预测模型, 选择预测结果最好的一次作为最终建模用变量。由于所选择出的变量个数较多, 在此不一一列出。所建 PLS 模型预测结果如表 2 中示, 其中预测集 TVB-N、菌落总数、pH 值、 L^* 参数四个指标的预测相关系数分别为 0.943, 0.903, 0.930, 0.952, 预测标准差分别为 4.742 $\text{mg} \cdot (100 \text{ g})^{-1}$, 0.384 $\log\text{CFU} \cdot \text{g}^{-1}$, 0.070, 1.477, 预测

集相关性结果如图 3 所示。从表中对比结果可知用 iPLS-GA 法筛选变量后与单独用 iPLS 方法筛选变量相比, 变量个数和 PLS 模型主成分数均得到大幅减少, 模型稳定性和预测精

度得到提高。说明 iPLS-GA 方法能够提取有效光谱变量, 减少无用信息和变量间的共线性, 提高模型的预测精度和预测能力。

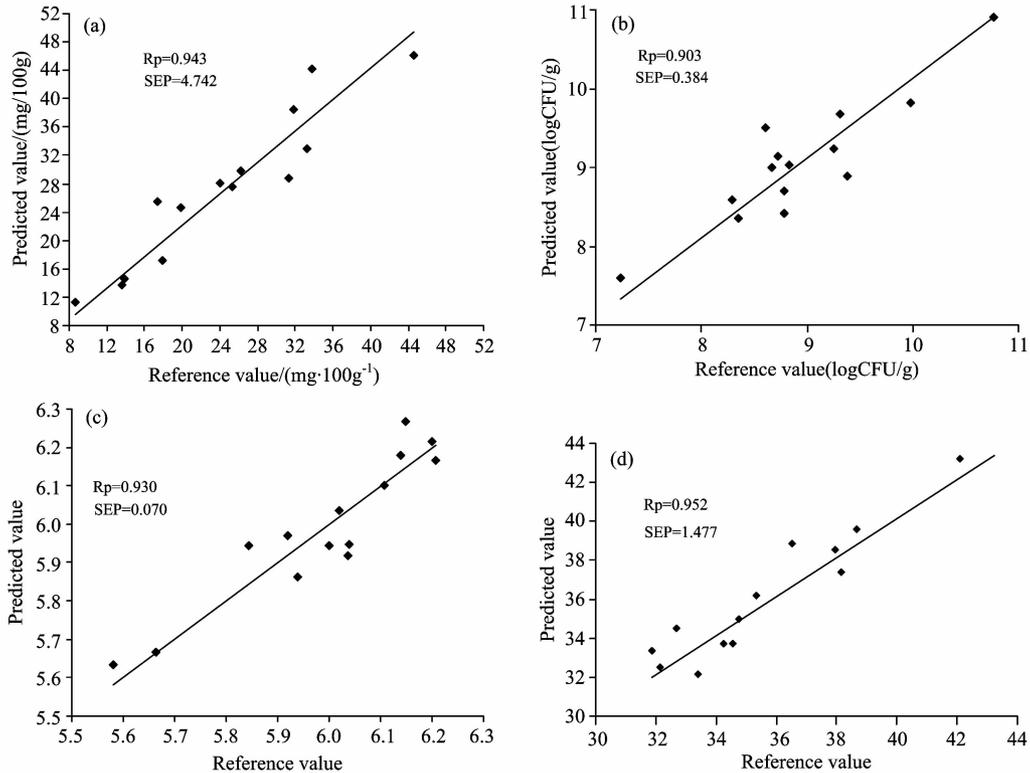


Fig. 3 Predicted vs reference values of 4 indexes in prediction set by iPLS-GA-PLS model

2.4 储存期预测及评价

根据 iPLS-GA 选择出的有效光谱变量建立的 TVB-N、菌落总数、pH、 L^* 的 PLS 预测模型分别对样品校正集和预测集进行预测, 得到四个指标的预测值, 由前述所得出的储存期预测模型式(1)分别对校正集和预测集的样品储存期进行预测, 预测结果分布图如图 4 所示。校正集和预测集样品

储存期的预测相关系数分别是 0.903 和 0.897, 预测标准差分别为 1.88 和 2.24, 达到了较好的预测效果。说明用可见近红外光谱技术能够同时快速、无损检测生鲜牛肉 TVB-N、菌落总数、pH 值、 L^* 参数多个指标, 结合所得的多指标储存期预测模型, 能够实现快速、多指标综合无损预测牛肉储存期的目的。

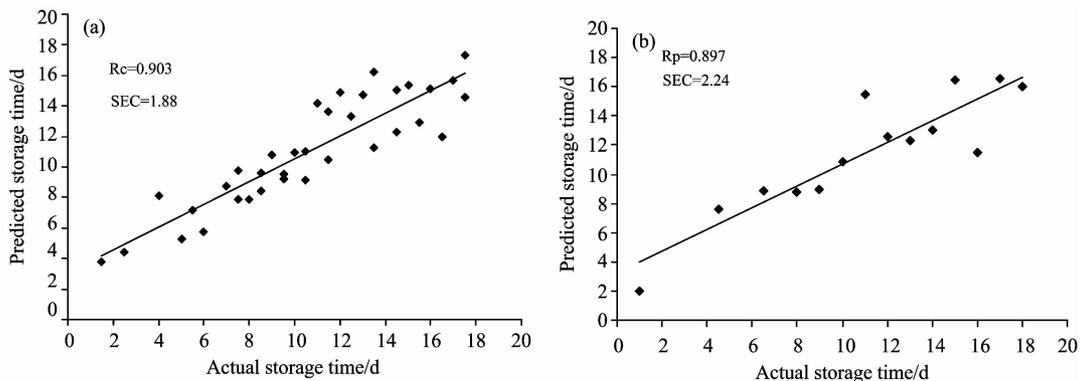


Fig. 4 Predicted vs actual value of storage time

(a): Calibration set; (b): Prediction set

3 结论

研究了储藏于 4 °C 环境下的生鲜牛肉储存期与 TVB-N、

菌落总数、pH 值、 L^* 参数多指标之间的关系并得出了基于多指标的储存期预测模型, 结合用 iPLS 和 GA 建立的 TVB-N、菌落总数、pH 值和 L^* 指标的光谱预测模型, 实现了用可见近红外光谱技术对牛肉储存期的无损快速预测。通过优

化对比, 确定 MSC 为光谱数据的最佳预处理方法。iPLS 结合 GA 提取出了有效的光谱变量, 提高了建模速度、模型的预测精度和预测能力。4 个指标预测集的预测相关系数和标准差分别是 0.943, 0.903, 0.930, 0.952 和 $4.742 \text{ mg} \cdot (100 \text{ g})^{-1}$, $0.384 \text{ logCFU} \cdot \text{g}^{-1}$, 0.070、1.477。校正集和预测集

样品储存期的预测相关系数和标准差分别是 0.903, 0.897 和 1.88, 2.24, 具有较好的预测效果。说明利用可见近红外光谱技术能够实现多指标综合无损快速预测生鲜牛肉的储存期, 为生牛肉储存期或货架期的无损快速检测提供了一种新的方法。

References

- [1] LU Wan-zhen, YUAN Hong-fu, XU Guang-tong, et al. (陆婉珍, 袁洪福, 徐广通, 等). Modern Near Infrared Spectroscopy Analytical Technology(现代近红外光谱分析技术). Beijing: China Petrochemical Press(北京: 中国石化出版社), 2006. 10.
- [2] Savenije B, Geesink G H, van der Palen J G P, et al. Meat Science, 2006, (73): 181.
- [3] Andrés S, Silva A, Soares-Pereira A L, et al. Meat Science, 2008, (78): 217.
- [4] Prieto N, Andrés S, Giraldez F J, et al. Meat Science, 2008, (79): 692.
- [5] Holmer S F, McKeith R O, Boler D D, et al. Meat Science, 2009, (82): 86.
- [6] Raúl G, Antonio J S, Joel G, et al. Food Research International, 2011, (44): 331.
- [7] Norgaard L, Saudland A, Wagneret J, et al. Applied Spectroscopy, 2000, (54): 413.
- [8] CAI Jian-rong, WAN Xin-min, CHEN Quan-sheng(蔡健荣, 万新民, 陈全胜). Acta Optica Sinica(光学学报), 2009, 2(10): 2808.
- [9] Leardi R, González A L. Chemometrics and Intelligent Laboratory Systems, 1998, (41): 195.
- [10] Leardi R. Journal of Chemometrics, 2000, (14): 643.
- [11] Ying Yibin, Liu Yande. Journal of Food Engineering, 2008, (84): 206.
- [12] Xu H R, Qi B, Sun T, et al. Journal of Food Engineering, 2012, (109): 142.
- [13] Chen Q S, Jiang P, Zhao J W. Spectrochimica Acta Part A, 2010, (76): 50.
- [14] Zou X B, Zhao J W, Huang X Y, et al. Chemometrics and Intelligent Laboratory Systems, 2007, (87): 43.
- [15] Christoffer Abrahamsson, Jonas Johansson, Anders Sparén, et al. Chemometrics and Intelligent Laboratory Systems, 2003, (69): 3.

Prediction of Storage Time of Fresh Beef with Multi-Index Using Visible and Near-Infrared Spectroscopy

MA Shi-bang^{1,2}, XU Yang^{1*}, TANG Xiu-ying¹, TIAN Xiao-yu¹, FU Xing¹

1. College of Engineering, China Agricultural University, Beijing 100083, China

2. Nanyang Institute of Technology, Nanyang 473004, China

Abstract The prediction model of beef's storage time was established based on multi indexes of fresh beef, such as TVB-N, colony total, pH value, and L^* parameter. Visible and near-infrared spectroscopy (Vis/NIR) combined with interval PLS (iPLS) and genetic algorithm (GA) was investigated for establishing PLS calibration model of above 4 indexes, respectively, and rapid and nondestructive prediction of the storage time of fresh beef stored at 4 °C was realized. PLS models of 4 indexes were built with full spectrum and effective variables selected by iPLS and iPLS-GA method, respectively. The performance of each model was evaluated according to two correlations coefficients (R) and standard error (SE) of calibration and prediction sets. Experimental results showed that the performance of all models built with effective variable selected by iPLS-GA was better than full spectrum and iPLS. The storage time of calibration and prediction sets of beef samples was predicted by storage time model with predicted values of above 4 indexes, and was achieved as follows: $R_c = 0.903$, $R_p = 0.897$, $SEC = 1.88$ and $SEP = 2.24$. The study demonstrated that the beef's storage time can be synthetically predicted with multi-index by using visible and near-infrared spectroscopy combined with the prediction model of beef's storage time. This provides a new method for rapid and non-destructive detection of beef's storage time or shelf life.

Keywords Visible and near-infrared spectroscopy; Beef storage time; Multiple determination; Variable selection; PLS

* Corresponding author