

DOI:10.3969/j.issn.1000-1298.2010.12.026

基于多元校正法的香梨糖度可见/近红外光谱检测*

徐惠荣 陈晓伟 应义斌

(浙江大学生物系统工程与食品科学学院, 杭州 310029)

【摘要】 应用可见/近红外光谱透射技术结合多元校正法探讨了样品在不同温度条件(5、10、15、20℃)下香梨糖度的快速无损检测。在波长500~900 nm范围内,用逐步多元线性回归(SMLR)、偏最小二乘法(PLS)、最小二乘支持向量机(LS-SVM1、LS-SVM2)和遗传算法-偏最小二乘法(GA-PLS)等多种多元校正法进行了建模预测比较研究。预测结果从优到差依次为LS-SVM2、LS-SVM1、GA-PLS、PLS、SMLR。

关键词: 香梨 糖度 无损检测 可见/近红外光谱 多元校正 温度

中图分类号: S123 **文献标识码:** A **文章编号:** 1000-1298(2010)12-0126-04

Multivariate Approach to Determinate Sugar Content of Fragrant Pears with Temperature Variation by Visible/NIR Spectroscopy

Xu Huirong Chen Xiaowei Ying Yibin

(School of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310029, China)

Abstract

Nondestructive determination of sugar content (SC) of fragrant pears with temperature variation (5°C, 10°C, 15°C and 20°C) was studied by visible-NIR spectroscopy coupled with multivariate calibration methods. In the region of 500~900 nm, five multivariate calibration models: stepwise multiple linear regression (SMLR), partial least squares (PLS), least squares-support vector machines (LS-SVM1, only 30 bands spectral data obtained by SMLR analysis inputted as X variable), LS-SVM2 (both 30 bands spectral data and sample temperature inputted as X variable) and genetic algorithm-partial least squares (GA-PLS) were compared. The results showed that the prediction performance was in the order of LS-SVM2, LS-SVM1, GA-PLS, PLS and SMLR.

Key words Fragrant pears, Sugar content, Nondestructive determination, Visible-NIR spectroscopy, Multivariate calibration, Temperature

引言

近红外光谱分析技术作为一种无损检测技术已被成功用于鲜果内部品质(如糖度等)的检测,然而,近红外光谱测量技术仍存在不足,主要是预测模型缺乏稳健性^[1]。主要原因有^[2]:受样品相关因素的扰动,如样品温度、大小、品种、产地和收获时间等;受仪器与环境等因素的影响,如杂散光、空气湿

度、光谱仪的温度、灯龄等。

样品温度的变化会引起近红外区域分子键振动强度和频率的变化,进而影响近红外校正模型的性能^[1,3~6]。由于在实际的生产过程中很难保证每个样品温度的一致性,因而研究建立稳健的近红外光谱校正模型实现对温度的修正或补偿是很有必要的^[1,4]。目前,研究温度修正和补偿的方法可归纳为以下两大类^[7]:①最大可能去除温度影响的方

收稿日期:2010-04-28 修回日期:2010-05-06

*“十一五”国家科技支撑计划资助项目(2006BAD11A12)

作者简介:徐惠荣,副教授,主要从事农产品品质无损检测技术与装备研究, E-mail: hrxu@zju.edu.cn

通讯作者:应义斌,教授,博士生导师,主要从事农产品品质与安全无损检测研究, E-mail: yingyb@zju.edu.cn

法,主要有选择光谱中的不变部分进行建模的方法,如无用变量剔除法(UVE)、模拟退火法(SA)等,以及光谱调整法,如分段直接校正法(PDS)、连续分段直接校正法(CPDS)和正交信号校正(OSC)等。

②纯回归方法,主要包括线性回归方法,如逐步多元线性回归(SMLR)、主成分回归(PCR)和偏最小二乘法(PLS),以及非线性回归方法,如局部加权拟合(LWR)、神经网络(NN)和最小二乘支持向量机(LS-SVM)等。

本文主要探讨利用可见/近红外光谱结合多种多元校正法对受温度变化影响的香梨糖度进行检测。

1 实验材料与方法

1.1 材料

市购新疆库尔勒香梨100个,用毛巾擦干净水果表皮,作好样品号及采样点标记后,将样品置于温控箱内(温度控制范围0~50℃,波动±0.1℃),调节温度到某一实验设计温度(5、10、15、20℃),在光谱采集前,样品在温控箱内每一设定温度条件下恒温保存24h,使样品内外温度达到一致。

1.2 仪器与方法

利用前期研究构建的检测系统^[8],分别采集样品温度为5、10、15、20℃时的可见/近红外透射光谱。光谱采集软件OOIBase32(美国海洋光学公司)参数设置如下:积分时间240 ms,平均采样5次,平滑点数为5。为了减少在实验过程中样品温度的变化,每次从温控箱中取5个样品进行光谱采集,在尽可能短的时间内完成,每个样品在赤道部位(最大横径)采集2次(标记处),其平均光谱作为样品的特征光谱。

糖度测定参照国家标准(GB/T 10650—2008),在相应的光谱采集点附近取果肉挤压取汁后,用手持糖量计(数字式折射仪PR-101型,日本ATAGO公司)进行测量,将2次测量平均值作为其糖度值,糖度分布范围为10.1~14.0 °Brix。

1.3 数据分析

100个香梨在4个温度下的光谱混合后形成400个光谱样本。有2个香梨样品存在理化分析错误,剩余98个香梨样本的糖度分布情况如图1所示,并用TQ Analyst 7.2(Thermo Electron Corp., USA)软件中的Spectrum outlier功能模块分析,剔除1个异常光谱样本,将剩余光谱样本各温度下随机选取2/3作为校正集样本(共261个),1/3作为验证集样本(共130个),在波长500~900 nm范围内(如图2所示,光谱两端信噪比不理想)进行多变量

建模预测分析。

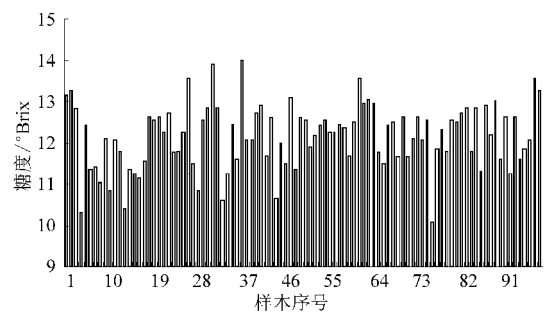


图1 样本的糖度分布情况

Fig. 1 SC distribution of samples

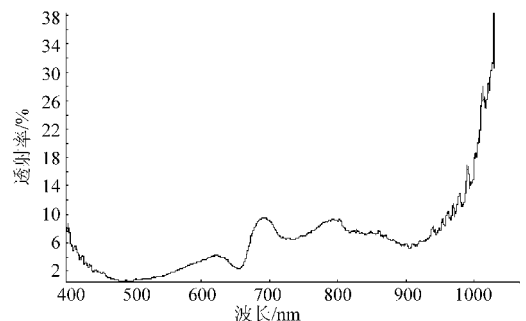


图2 样品温度为5℃时的透射光谱曲线

Fig. 2 Transmittance spectrum of pear at 5°C

分析用的多元校正法有:逐步多元线性回归(SMLR)、偏最小二乘法(PLS)、遗传算法-偏最小二乘法(GA-PLS)和最小二乘支持向量机(LS-SVM),SMLR和PLS分析用TQ Analyst 7.2软件,GA-PLS分析用PLS_Toolbox 5.2.1(Eigenvector Research, Inc., USA),LS-SVM分析用LS-SVM 1.5 toolbox(Suykens等开发, K. U. Leuven, Belgium)以及Matlab 7.0。

2 结果与讨论

2.1 逐步多元线性回归

在波长500~900 nm范围内对上述样本集(经均值中心化处理)进行SMLR回归分析,在选取30个波长时取得最好结果,校正集相关系数 $R = 0.868$,校正均方根误差 $C = 0.386$ °Brix,预测均方根误差 $P = 0.587$ °Brix,预测值与实际值间的对应关系如图3a所示。

2.2 偏最小二乘法

在波长500~900 nm范围内对上述样本集(经均值中心化处理)进行PLS回归分析,通过TQ Analyst 7.2中的PRESS分析选用16个因子建模,得到 $R = 0.860$, $C = 0.397$ °Brix, $P = 0.580$ °Brix,预测值与实际值间的对应关系如图3b所示。

可见,与前期研究^[9]在简单温度变化下(样品温度只有2个)SMLR的预测结果要优于PLS的预

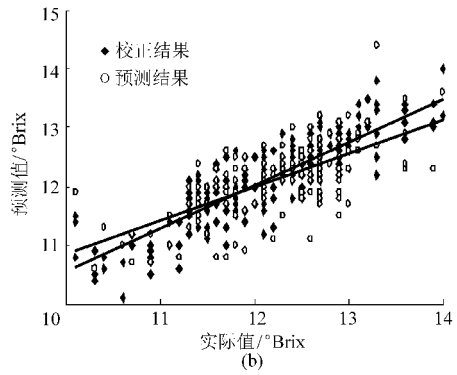
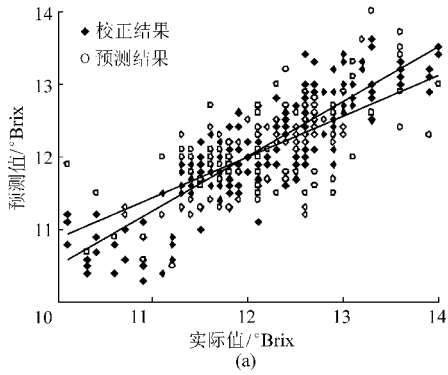


图3 预测值与实际值间的对应关系

Fig. 3 Predicted SC against the measured SC

(a) SMLR方法 (b) PLS方法

测结果相反,在温度变化范围较大条件下(样品温度有4个),SMLR方法已显得不足以进行温度的补偿,该结果与Abe等(2000)^[10]得出的结论相一致。

2.3 遗传算法-偏最小二乘法

通过PLS_Toolbox 5.2.1软件,在Matlab 7.0环境下进行GA变量选择分析。主要参数设置如下:种群大小64,窗口大小10,收敛率50%,变异率0.005,采用双点交叉重组法。对上述样本集同样进行均值中心化处理后,通过适应度(最小交互验证均方根误差)分析,比较22个独立模型后,选出410个波长。

将经GA分析选定的校正集光谱矩阵(410个波长数,261×410),调入建模预测分析程序进行PLS建模分析,同样在选用16个因子下取得最好结果, $C=0.370\text{ }^{\circ}\text{Brix}$, $P=0.555\text{ }^{\circ}\text{Brix}$,分析结果如图4所示。

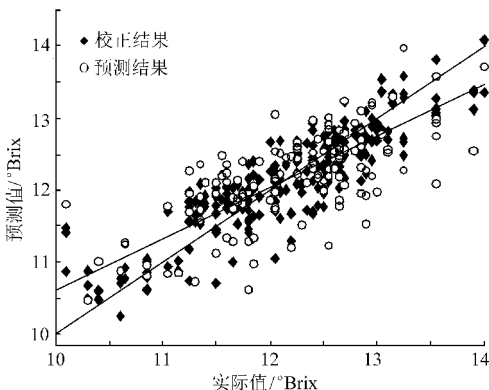


图4 GA-PLS模型预测值与实际值间的对应关系

Fig. 4 Predicted SC against the measured SC by GA-PLS

2.4 LS-SVM1

将SMLR分析中得到的30个波段的光谱值作为LS-SVM分析的输入量,选用高斯径向基函数(RBF)为核函数,其中核宽 σ^2 和正规化参数 γ 通过二步格点搜索法与L-折交互验证($L=10$)^[11]相结合的粗选和精选两个步骤进行优化,参数优化结

果如图5所示,网格点“·”反映了第一步格点搜索范围及步长,曲线反映了误差等高线,网络点“×”反映了第二步格点搜索范围及步长,得最佳 $\gamma=7\,278.439\,9$, $\sigma^2=395.942\,63$ 。利用优化后的参数进行回归分析,预测结果如图6所示, $P=0.539\text{ }^{\circ}\text{Brix}$ 。

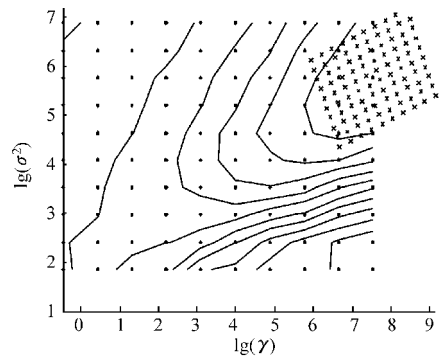
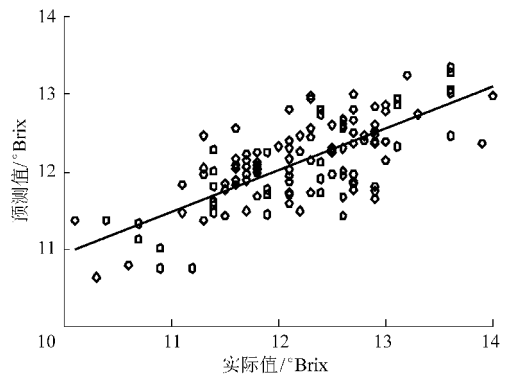
图5 LS-SVM1中参数 σ^2 和 γ 的优化过程Fig. 5 Optimization of σ^2 and γ for LS-SVM1

图6 LS-SVM1回归分析预测结果

Fig. 6 Predicted SC against the measured SC by LS-SVM1

2.5 LS-SVM2

Wülfert等(2000)^[12]认为由于温度是一个已知量,因而可追加到校正模型矩阵中,与光谱一起成为一个独立的变量来提高预测精度。本文中,对上述样本集,将SMLR分析中得到的30个波段的光谱值及其对应的样品温度值同时作为LS-SVM分析的输入量,选用RBF为核函数,其中核宽 σ^2 和正规化

参数 γ 同样通过二步格点搜索法与 L -折交互验证 ($L = 10$) 相结合的粗选和精选两个步骤进行优化, 参数优化结果如图 7 所示, 得最佳 $\gamma = 10489.6662$, $\sigma^2 = 933.567768$ 。利用优化后的参数进行回归分析, 预测结果如图 8 所示, $P = 0.523$ °Brix。

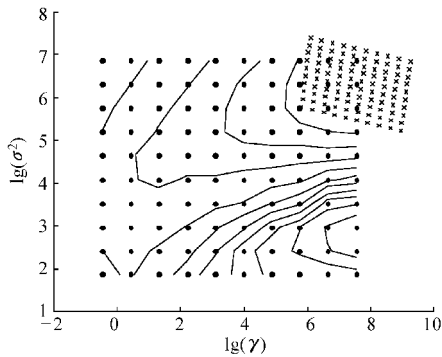


图 7 LS-SVM2 中参数 σ^2 和 γ 的优化过程

Fig. 7 Optimization of σ^2 and γ for LS-SVM2

从上述 SMLR、PLS、GA-PLS、LS-SVM1 (输入量为 30 个波段下的光谱值) 和 LS-SVM2 (输入量为 30 个波段下的光谱值和样品温度值) 等 5 种多元校正法的预测结果比较可以看出: 预测结果从优到差依次为 LS-SVM2、LS-SVM1、GA-PLS、PLS、SMLR。

3 结论

(1) 应用可见/近红外光谱技术结合多元校正

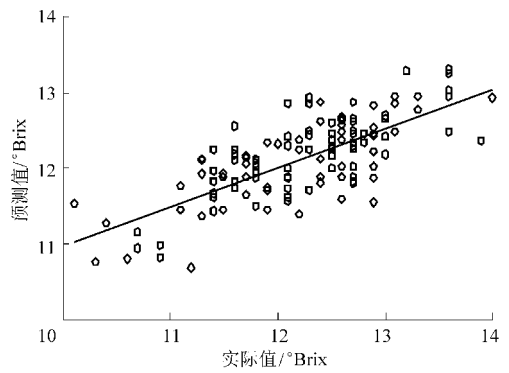


图 8 LS-SVM2 回归分析预测结果

Fig. 8 Predicted SC against the measured SC by LS-SVM2

法 (SMLR、PLS、GA-PLS、LS-SVM), 研究了复杂样品温度条件下 (5、10、15、20°C) 的香梨糖度快速检测。预测精度从优到差依次为 LS-SVM2、LS-SVM1、GA-PLS、PLS 和 SMLR。

(2) 在复杂的温度变化条件下, SMLR 已不足以进行温度的补偿。GA-PLS 通过遗传算法选取了与温度无关的波段, 使预测精度较全波段 PLS 高, 而 LS-SVM 采用了非线性模型可以有效地补偿温度的非线性影响。另外, LS-SVM2 中将温度引入了校正模型矩阵中, 在一定程度上补偿了温度的影响并提高了检测精度。

参 考 文 献

- 1 Sánchez N H, Lurol S, Roger J M, et al. Robustness of models based on NIR spectra for sugar content prediction in apples [J]. *Journal of Near Infrared Spectroscopy*, 2003, 11(2): 97 ~ 107.
- 2 Segtnan V H, Mevik B H, Isaksson T, et al. Low-cost approaches to robust temperature compensation in near-infrared calibration and prediction situations [J]. *Applied Spectroscopy*, 2005, 59(6): 816 ~ 825.
- 3 Cozzolino D, Liu L, Cynkar W U, et al. Effect of temperature variation on the visible and near infrared spectra of wine and the consequences on the partial least square calibrations developed to measure chemical composition [J]. *Analytica Chimica Acta*, 2007, 588(2): 224 ~ 230.
- 4 Kawano S, Abe H, Iwamoto M. Development of a calibration equation with temperature compensation for determining the Brix value in intact peaches [J]. *Journal of Near Infrared Spectroscopy*, 1995, 3(4): 211 ~ 218.
- 5 Arazuri S, Jare'n C, Arana J I. Selection of the temperature in the sugar content determination of kiwi fruit [J]. *International Journal of Infrared and Millimeter Waves*, 2005, 26(4): 607 ~ 616.
- 6 Blanco M, Valdés D. Suppressing the temperature effect in near infrared spectroscopy by using orthogonal signal correction [J]. *Journal of Near Infrared Spectroscopy*, 2006, 14(3): 155 ~ 160.
- 7 Hageman J A, Westerhuis J A, Smilde A K. Temperature robust multivariate calibration: an overview of methods for dealing with temperature influences on near infrared spectra [J]. *Journal of Near Infrared Spectroscopy*, 2005, 13(2): 53 ~ 62.
- 8 Lu H, Jiang H, Fu X, et al. Non-invasive measurements of the internal quality of intact 'Gannan' navel orange by VIS-NIR spectroscopy [J]. *Transactions of the ASABE*, 2008, 51(3): 1 009 ~ 1 014.
- 9 徐惠荣, 汪辉君, 黄康, 等. PLS 和 SMLR 建模方法在水蜜桃糖度无损检测中的比较研究 [J]. *光谱学与光谱分析*, 2008, 28(11): 2 523 ~ 2 526.

Xu Huirong, Wang Huijun, Huang Kang, et al. Comparison of PLS and SMLR for nondestructive determination of sugar content in honey peach using NIRS [J]. *Spectroscopy and Spectral Analysis*, 2008, 28(11): 2 523 ~ 2 526. (in Chinese)

(下转第 147 页)

- Li Baoguo, Xiao Jianjun, Zhang Yan. Research progress on corn stress cracks in drying process [J]. University of Shanghai for Science and Technology, 2001, 23(2):107~110. (in Chinese)
- 3 朱文学, 连政国. 玉米应力裂纹的生成和扩展过程模拟[J]. 农业工程学报, 2000, 16(1):113~116.
Zhu Wenxue, Lian Zhengguo. Simulation of formation and development of stress cracks in corn [J]. Transactions of the CSAE, 2000, 16(1):113~116. (in Chinese)
- 4 李心平, 高连兴, 马福丽. 玉米种子力学特性的有限元分析[J]. 农业机械学报, 2007, 38(10):64~67.
Li Xinping, Gao Lianxing, Ma Fuli. Analysis of finite element method on mechanical properties of corn seed [J]. Transactions of the Chinese Society for Agricultural Machinery, 2007, 38(10):64~67. (in Chinese)
- 5 张俊雄, 荀一, 李伟. 基于形态特征的玉米种子表面裂纹检测方法[J]. 光学精密工程, 2007, 15(6):951~956.
Zhang Junxiong, Xun Yi, Li Wei. Detection of surface cracks of corn kernel based on morphology [J]. Optics and Precision Engineering, 2007, 15(6):951~956. (in Chinese)
- 6 朱文学, 张玉先, 李栋, 等. 玉米应力裂纹扩展的分形模型及动力学分析[J]. 农业机械学报, 1999, 30(3):48~52.
Zhu Wenxue, Zhang Yuxian, Li Dong, et al. Study on fractional models and propagating kinetics of stress cracks on corn [J]. Transactions of the Chinese Society for Agricultural Machinery, 1999, 30(3):48~52. (in Chinese)
- 7 李心平, 高连兴. 种子玉米籽粒果柄断裂机理试验研究[J]. 农业工程学报, 2007, 23(11):47~51.
Li Xinping, Gao Lianxing. Experimental study on breaking mechanism of kernel stem of corn seed [J]. Transactions of the CSAE, 2007, 23(11):47~51. (in Chinese)
- 8 Bino R J, Aartse J W, Van Der Burg W J. Non destructive X-ray of *arabidopsis* embryo mutants [J]. Seed Science Research, 1993, 3(2):167~170.
- 9 Cicero S M. Evaluation of mechanical damage in seeds of maize (*Zea mays* L.) by X-ray and digital imaging [J]. Seed Science and Technology, 1998, 26(3):603~612.
- 10 徐立章, 李耀明, 李洪昌. 水稻谷粒脱粒损伤的影响因素分析[J]. 农业机械学报, 2008, 39(12):55~59.
Xu Lizhang, Li Yaoming, Li Hongchang. Analysis on factors affecting performance of rice kernel damage during threshing [J]. Transactions of the Chinese Society for Agricultural Machinery, 2008, 39(12):55~59. (in Chinese)
- 11 高连兴, 李晓峰, 接鑫, 等. 大豆内部机械损伤对发芽的影响[J]. 农业机械学报, 2010, 41(10):63~66, 102.
Gao Lianxing, Li Xiaofeng, Jie Xin, et al. Inner mechanical damage impact to germination of soybean kernels [J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(10):63~66, 102. (in Chinese)

(上接第 129 页)

- 10 Abe H, Iyo C, Kawano S. A study on the universality of a calibration with sample temperature compensation [J]. Journal of Near Infrared Spectroscopy, 2000, 8(1):209~213.
- 11 魏国, 刘剑, 孙金玮, 等. 基于 LS-SVM 的非线性多功能传感器信号重构方法研究 [J]. 自动化学报, 2008, 34(8):869~875.
Wei Guo, Liu Jian, Sun Jinwei, et al. Study on nonlinear multifunctional sensor signal reconstruction method based on LS-SVM [J]. Acta Automatica Sinica, 2008, 34(8):869~875. (in Chinese)
- 12 Wülfert F, Kok W Th, de Noord O E, et al. Linear techniques to correct for temperature induced spectra variation in multivariate calibration [J]. Chemometrics Intelligent Laboratory Systems, 2000, 51(2):189~200.