

近红外光谱法快速检测甜菜糖度的模型优化

李毛毛^{1,4}, 郑喜群¹, 任健¹, 赵丽影³, 杨勇^{1,2*}

- (1. 齐齐哈尔大学食品与生物工程学院, 农产品加工黑龙江省普通高校重点实验室, 齐齐哈尔 161006;
2. 东北农业大学食品学院, 哈尔滨 150030; 3. 博天糖业有限公司, 北京 100029;
4. 聚光科技(杭州)股份有限公司, 杭州 310052)

摘要: **目的** 建立起近红外光谱技术关于甜菜糖度的最佳预测模型。**方法** 研究了 Savitzky-Golay 平滑处理、Savitzky-Golay 导数、均值中心化、差分求导、净分析信号、去趋势校正、标准正态变量变换和多元散射校正等 8 种预处理方法的多方法联用处理进行光谱数据的预处理, 结合光谱波段优选, 建立甜菜糖度与近红外光谱的预测模型。**结果** 在进行模型的评价时, 以误差均方根(SEP)、校正标准误差(SEC)与交叉检验误差(SECV)作为评价指标。**结论** 发现经过光谱波段优选之后, 结合 Savitzky-Golay 平滑、Savitzky-Golay 导数、去趋势校正及均值中心化进行光谱数据的预处理得到的模型效果最佳。

关键词: 近红外光谱法; 甜菜; 糖度; 偏最小二乘法

Model optimization on rapid detection of beet sugar content by near infrared spectroscopy

LI Mao-Mao^{1,4}, ZHENG Xi-Qun¹, REN Jian¹, ZHAO Li-Ying³, YANG Yong^{1,2*}

- (1. Key Laboratory of Processing Agricultural Products of Heilongjiang Province, College of Food and Bioengineering, Qiqihar University, Qiqihar 161006, China; 2. College of Food Science, Northeast Agricultural University, Harbin 150030, China; 3. Bo-Tian Sugar Limited Company, Beijing 100029, China;
4. Focused Photonics (Hangzhou), Inc. Hangzhou 310052, China)

ABSTRACT: Objective To establish the optimal forecast model of beet sugar content based on near-infrared spectroscopy. **Methods** Eight kinds of methods as Savitzky-Golay smoothing, Savitzky-Golay derivative, mean centering, differential derivative, net analytic signal, to the tendency calibration, standard normal variable transformations and multiplicative scatter correction were associated and combined with preferable spectral bands choice to build the beet sugar content and the near-infrared spectrum prediction model. **Results** SEP, SEC and SECV were set as the evaluation index when evaluating the models. **Conclusion** The results show after the preferred spectrum bands choice, the preprocessing that combines Savitzky-Golay smoothing, Savitzky-Golay derivative, to the trend correction and mean centering can obtain the optimal model.

KEY WORDS: near infrared spectroscopy; sugar beet; sugar content; partial least squares

基金项目: 黑龙江省自然科学基金项目(C201331)、齐齐哈尔市科技局农业攻关项目(NYGG-201206-3)、黑龙江省大学生创新创业训练计划项目(201410221040)

Fund: Supported by the Natural Science Foundation of Heilongjiang Province (C201331), Qiqihar Science and Technology Bureau of Agricultural Research Projects (NYGG-201206-3), Innovation and Entrepreneurship Training Program for College Students in Heilongjiang Province (201410221040)

*通讯作者: 杨勇, 副教授, 主要研究方向为农产品快速检测技术。E-mail: yangyong7904@163.com

*Corresponding author: YANG Yong, Associate Professor, College of Food and Bioengineering, Qiqihar University, Qiqihar 161006. E-mail: yangyong7904@163.com

1 引言

近红外光谱分析技术以其快速、安全、无损、无污染等优势逐步向各个领域渗透。相关报道显示, 20 世纪 80 年代以来, 近红外在制糖方面的研究主要集中在美国、澳大利亚等一些国家^[1-4], 但是这些研究主要集中在蔗糖方面。甜菜作为制糖工业的一个重要的原料作物, 其相关的报道却是少之又少^[4-9]。考察期间重点关注了甜菜的收购方式。现在的甜菜在收购初期, 完全是沿袭传统的按重量进行收购的方式, 这样的收购方式会使得一部分的甜菜种植者以利益为驱动, 造成在甜菜种植时仅仅注重甜菜的质量而不追求甜菜的品质^[10]。低质量甜菜的收购, 会使得甜菜加工企业在收购及加工期间产生多余的支出, 给甜菜加工企业带来负担。本研究着手于甜菜收购环节, 利用近红外光谱技术建立快速甜菜检测机制^[11-14], 为甜菜的按质论价提供技术支持。

2 材料与方方法

2.1 实验材料

甜菜样品共 28 个品种, 734 个样品, 由博天糖业有限公司提供。

2.2 仪器

SupNIR-2720 近红外光谱仪(钢镓砷检测器, 波长范围 1000~1800 nm, 聚光科技(杭州)股份有限公司); ADS220 型自动数字糖量计(英国 Bellingham Stanley 公司); LOS-M-60 型可移动式甜菜样品处理机(波兰 Wikipol 公司)。

2.3 试验方法

2.3.1 甜菜样品的选取及糖度的测量

本实验选取具有代表性的 734 个特征样品, 选取其中的 587 个作为校正集样品, 剩余的 147 个样品作为预测集样品。将甜菜样品经 LOS-M-60 型可移动式甜菜样品处理机(波兰 Wikipol 公司)处理成糊状后, 采用旋光法进行甜菜中糖度的测量。

2.3.2 甜菜近红外光谱采集

将甜菜样品均匀平铺在样品盘中, 使用近红外分析仪以 1 nm 的采样间隔进行光谱数据的采集。设置光谱平均次数为 30 次, 样品在采集时保持样品及环境温度均为(25±1) °C, 光谱扫描范围为 1000~1800 nm, 得出扫描光谱图。每个样品重复扫描 3 次, 每次扫描时样品要求重新装样, 保持装样的均一性, 求得平均光谱曲线。

2.4 光谱数据处理与分析

利用 RIMP V1.4 软件对平滑处理(savitzky-golay smoothing, SGM)、导数(savitzky-golay derivative, SGD)、均值中心化(mean centering, MC)、差分求导(difference drivative, DD)、净分析信号(net analyte signal, NAS)、去趋势校正(dislodge tendency, DT)、标准正态变量变换(standard normal variate, SNV)和多元散射校正(multiplicative signal correction, MSC)等 8 种光谱预处理方法的消噪效果进行比较, 应用偏最小二乘法(partial least square, PLS)校正方法建立数学模型。在进行光谱波段优选的时间根据甜菜的主要的吸收峰的位置, 选择 1000~1650nm 为优选之后的波段。最后以校正集的交叉验证均方根(RMSECV)和预测集的预测均方根误差(RMSEP)作为评价指标, 计算公式为:

$$RMSECV = \sqrt{\frac{\sum (y_i' - y_i)^2}{n-1}} \quad RMSEP = \sqrt{\frac{\sum (y_i' - y_i)^2}{n}}$$

式中: y_i' 为近红外预测值;

y_i 为标准方法得到的值;

n 为样品数。

3 结果与分析

3.1 甜菜样品糖度化学实测值的分析

校正集和预测集甜菜样品糖度的实测结果见表 1, 587 个校正集样品的糖度值范围为 13%~17%, 涵盖了大多数现有的甜菜糖度值分布范围, 说明本实验选用的样品具有很强的代表性。

表 1 甜菜糖度测定结果

Table 1 The sugar content of beet samples used for calibration and validation

分组	样品数量/个	糖度范围/%	平均值/%	标准偏差/%
校正集	587	13.00 ~ 17.00	14.75	0.9342
预测集	147	13.04 ~ 17.00	13.87	1.0135

3.2 最佳光谱预处理方法的确定

甜菜样品的大小和均匀程度会影响光谱的信噪比,因此,采用 Savitzky-Golay 平滑、Savitzky-Golay 导数、多元散射校正(MSC)、均值中心化、差分求导、净分析信号(NAS)和去趋势校正(DT)共 7 种光谱预处理方法对甜菜的原始光谱进行预处理,然后分别采用 PLS 法建立相关模型。实验结果如表 2、3 所示。实验通过对近红外光谱进行波长优选,以模型的 r^2 、RMSECV 和 RMSEP 为模型精度评价指标, r^2 越大、RMSECV 和 RMSEP 越小,说明模型的精度越高,所建模型的预测能力和稳健性越高。587 样本去除了重复性样本和异常值样本,最终确定由 734 样本建立了该糖度模型。

表 2 表明采用 SGM、SGD、DT、MC 综合处理光谱数据,以 PLS 法建立回归模型的预测能力优于另外 2 种预处理方法。最优模型的糖度模型的交互验证得到的预测值和实际值的散点图以及残余方差图,如图 1~图 2 所示,线性方程为 $Y=0.86564X+2.0075$ 。

3.3 甜菜糖度模型的可靠性验证

选取主成分数 8,用 PLS 法处理光谱建立甜菜糖度模型,将外部验证所选取的 70 个样品代入模型中,进行甜菜样品的糖度测定,预测值和实测值的分布见图 1 所示。结果表明,甜菜糖度预测值和实测值的 r^2 为 0.9214,说明近红外光谱预测值与糖度实测值具有极显著的相关性,经单因素方差分析得到 $P=0.9549 > 0.05$,即在 95%的置信区间内,甜菜糖度实测值与预测值无显著性差异。 $Bias=0.023$ 说明预测模型的准确性良好, $SEP=0.439$ 表示预测结果的精密程度良好。

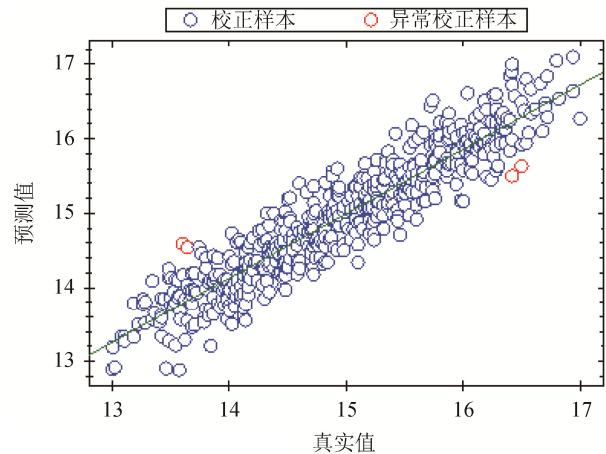


图 1 PLS 校正模型内部交互验证的预测值和实际值的散点图

Fig. 1 Scatter plot of predicted and measured values of calibration set using PLS model

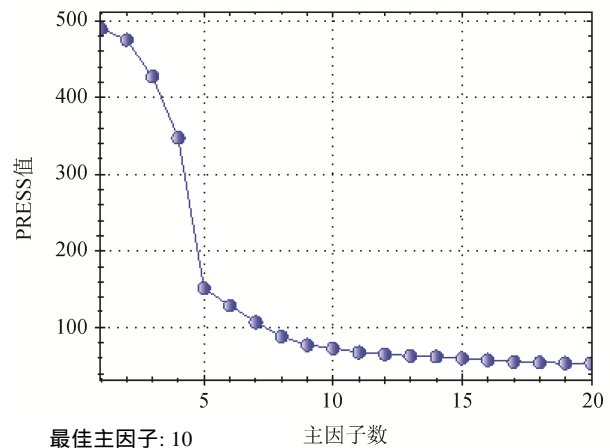


图 2 PLS 糖度模型的残余方差与主成分因子数的关系
Fig. 2 Relationship between number of principal component and residual validation variance of sugar content using PLS model

表 2 波长优选前的不同光谱预处理结果

Table 2 Pretreating results of differentspectrum before optimization of wave length

预处理方式	主因子数	SEC	SEP	Bias	SECV
SGM、SGD、MSC、MC	8	0.4733	0.4686	0.0622	0.4659
SGM、DD、NAS、MC	11	0.4959	0.6678	0.0034	0.5987
SGM、SGD、DT、MC	10	0.3536	0.4120	0.0362	0.3659

表 3 波长优选后的不同光谱预处理结果

Table 3 Pretreating results of different spectrum after optimization of wave length

预处理方式	主因子数	SEC	SEP	Bias	SECV
SGM、SGD、MSC、MC	5	0.4190	0.4190	0.0431	0.4524
SGM、DD、NAS、MC	15	0.4878	0.6434	-0.0031	0.5455
SGM、SGD、DT、MC	10	0.3424	0.4086	0.0318	0.3562

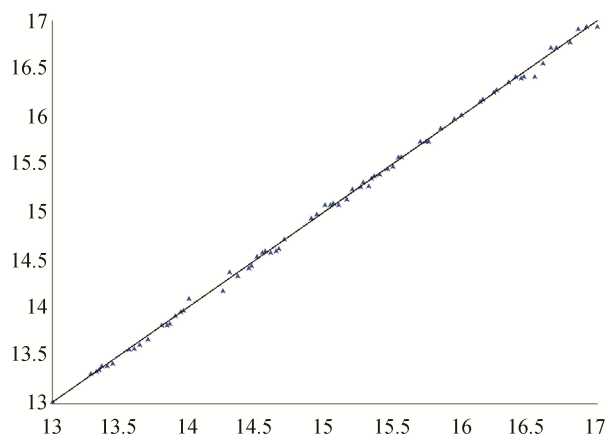


图 3 甜菜糖度值 PLS 模型的外部验证结果

Fig. 3 External verification results of beet sugar content using PLS model

4 结 论

结果发现经过光谱波段优选之后, 结合 Savitzky-Golay 平滑、Savitzky-Golay 导数、去趋势校正及均值中心化进行光谱数据的预处理得到的模型效果最佳。

近红外光谱法作为一种简单、快速、无损、环保的检测方法, 能够良好的评价甜菜的糖度, 可以运用该方法来实现甜菜收购环节的定等分级和按质论价。

参考文献

- [1] Kuczejda M, Ramirez S, Yilmaz S. NIR polarimetry/sample preparation for sucrose measurement and assessment of an automated system for color, turbidity and ash determination [J]. *Int Sugar J*, 2005, 107(1277): 302–307.
- [2] Haelele D, Sevenich D, Jones D, *et al.* NIR-based analytical systems enable development of corn seed products and corn grain markets that are specific for dry-grind ethanol production [J]. *Int Sugar J*, 2007, 109(1299): 154–158.
- [3] Simpson R, Oxley J. Routine analysis of molasses and mixed juice by NIR spectroscopy [J]. *Int Sugar J*, 2009, 111(1326): 387–402.
- [4] Magaña CN, Núñez-Sánchez VM, Fernández-Cabanás, *et al.* Direct prediction of bioethanol yield in sugar beet pulp using near infrared spectroscopy [J]. *Bioresource Technol*, 2011, 102(20): 9542–9549.
- [5] Maalouly J, Eveleigh L, Douglas N, *et al.* Application of 2D correlation spectroscopy and outer product analysis to infrared

spectra of sugar beets [J]. *Vib Spectrosc*, 2004, 36(2): 279–285.

- [6] MroczykWB, Michalski KM. Quantitative and qualitative analyses in near infrared analysis of basic compounds in sugar beet leaf [J]. *Comput Chem*, 1995, 19(3): 299–301.
- [7] Pan LQ, Zhu QB, Lu RF, *et al.* Determination of sucrose content in sugar beet by portable visible and near-infrared spectroscopy [J]. *Food Chem*, 2015, 167(15): 264–271.
- [8] Pan LQ, Zhu QB, Lu RF, *et al.* Measurement of moisture, soluble solids, sucrose content and mechanical properties in sugar beet using portable visible andnear-infrared spectroscopy [J]. *Postharvest Biol Tec*, 2015, 102: 42–50.
- [9] Ferdinand P, Thomas S. Direct prediction of bioethanol yield in sugar beet pulp using Near infrared reflectance spectroscopy [J]. *Bioresource Technol*, 2011, 102(3): 2834–2841.
- [10] 刘凤华, 李洵微. 用近红外光谱法测定甜菜粕中总糖分和水分的方法研究[J]. *光谱实验室*, 1995, 12(4): 31–34 .
Liu FH, Li XW. Determination of total sugar and water in sugar beet by near infrared spectroscopy [J]. *Chin J Spectrosc Lab*, 1995, 12(4): 31–34.
- [11] Ainara L, Arazuri S, Jarén C, *et al.* Crude protein content determination of potatoes by NIRS technology [J]. *Procedia Technol*, 2013, 8: 488–492.
- [12] Prieto N, Roehle R, LavínP, *et al.* Application of near infrared reflectance spectroscopy to predict meat and meat products quality: A review [J]. *Meat Sci*, 2009, 83(2): 175–186.
- [13] Woo K, Chung MK, Kang N, *et al.* Effect of resistant starch from corn or rice on glucose control, colonic events, and blood lipid concentrations instreptozotocin induced diabeticroats [J]. *J Nutr Biochem*, 2003, (14): 166–172.
- [14] Blanco M, Villarroya I. NIR spectroscopy: a rapid-response analytical tool [J]. *Trend Anal Chem*, 2002, 21(4): 240–250.

(责任编辑: 李振飞)

作者简介



李毛毛, 本科, 主要研究方向为近红外光谱研究。

E-mail: maomao_li@fpi-inc.com



杨 勇, 博士, 副教授, 主要研究方向为农产品快速检测技术。

E-mail: yangyong7904@163.com