

甘蔗蔗汁锤度和还原糖分傅立叶变换近 红外定量分析数学模型

曹 干^{1,2}, 谭中文², 梁计南², 龙永惠¹, 周学秋³

(1 广东省农业科学院 作物研究所, 广东 广州 510640; 2 华南农业大学 农学院, 广东 广州 510642;

3 布鲁克光谱仪器公司, 北京 100081)

摘要:以傅立叶变换近红外光谱技术为基础, 采用透射测定方式和 PLS 算法, 建立蔗汁锤度和还原糖分定量分析数学模型。蔗汁锤度和还原糖分预测数学模型的决定系数(R^2)分别为 0.999 0 和 0.995 3, 均方根差分别为 0.115 0% 和 0.072 2%, 预测误差接近常规分析方法的误差。

关键词:傅立叶变换近红外光谱; 定量分析; 甘蔗; 锤度; 还原糖分

中图分类号: S566.1; O174.2

文献标识码: A

文章编号: 1001-411X(2004)02-0001-04

Mathematic models for quantitative analysis of Brix and reducing sugars in sugarcane juice with Fourier transform near infrared spectroscopy

CAO Gan^{1,2}, TAN Zhong-wen², LIANG Ji-nan², LONG Yong-hui¹, ZHOU Xue-qiu³

(1 Crop Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China;

2 College of Agriculture, South China Agric. Univ., Guangzhou 510642, China; 3 Bruker Instruments Ltd, Beijing 100081, China)

Abstract: Basic on the technique of Fourier transform near infrared (FT-NIR) spectroscopy, the calibration models for quantitative analysis of Brix and reducing sugars in sugarcane juice were developed by using transmission mode and calibrating with partial least square (PLS) algorithm. The determination coefficients (R^2) of the predicted models for Brix and reducing sugars in juice were 0.999 0 and 0.995 3, respectively; the root mean square errors of cross validation were 0.115 0% and 0.072 2% for Brix and reducing sugars in juice, respectively. The predictive errors measured by FT-NIR methods were close to those by routine laboratory methods.

Key words: Fourier transform near infrared spectroscopy; quantitative analysis; sugarcane; Brix; reducing sugars

在甘蔗品质分析中, 需要测定多项品质指标。除蔗糖分和转光度 2 个重要的品质指标外, 锤度和还原糖分同样是不可或缺的指标, 其在制糖、甘蔗选育种和栽培中对甘蔗品质的综合评价有重要作用。

蔗汁锤度和还原糖分的常规测定方法繁琐、费力和耗时, 在制糖过程中更无法实现在线分析, 是长期以来制糖、甘蔗科研和生产中需要解决的问题。Brotherton 等^[1]应用近红外技术测定原料蔗、原糖、蔗

渣的锤度、还原糖分、转光度等多个品质成分, 由偏小二乘法建立的预测数学模型获得较好的效果。Schaffler 等^[2,3]用近红外光谱方法测定混合蔗汁和糖蜜的蔗糖分、果糖、葡萄糖、锤度、转光度、灰分等项目, 结果表明, 除灰分的预测效果差外, 其余项目的准确度均可以接受。Berding 等^[4]用近红外方法测定甘蔗的多个品质性状, 其中蔗汁锤度近红外预测值与化学值的相关系数为 0.97, 预测标准差为 2.36

g/kg,认为准确度能满足甘蔗选育种的要求.国内鲜见有关应用近红外方法测定甘蔗蔗汁锤度和还原糖分的报道.

笔者曾以傅立叶变换近红外光谱技术为基础,探索了在不同测定光程下用透射方式快速测定蔗汁品质成分的可行性,已建立的蔗汁蔗糖分和转光度的近红外定量分析数学模型有很高的准确性^[5,6].在研究的基础上,通过建立蔗汁锤度和还原糖分的定量分析数学模型,期望通过一次测定就可同时获得多个品质参数,解决甘蔗品质常规分析程序繁琐、耗时费力的问题,为甘蔗品质管理和选育种提供一种快速准确的品质分析方法.

1 材料和方法

1.1 材料

在1999~2003年共4个榨季中,选择甘蔗不同基因型50个以上,基因型类型包括早、中、晚熟品种和高、中、低糖品种.在甘蔗不同成熟期(10月至次年3月)按标准方法取6~8条生长正常蔗茎,共获得蔗茎样品150个.

1.2 方法

1.2.1 蔗汁锤度和还原糖分的常规分析 蔗茎样品在小型压榨机上榨汁.蔗汁锤度采用比重法测定,还原糖分采用兰-艾农法测定^[7].由常规分析获得的数据称为化学值,由近红外分析获得的数据则称为预测值.

1.2.2 蔗汁近红外光谱的测定 用快速滤纸过滤除去蔗汁中的蜡粉、纤维碎粒等杂质后,取适量过滤蔗汁置于光程为5.0 mm的玻璃比色杯中,用透射方式扫描获得蔗汁近红外吸收光谱图.近红外光谱的测定使用德国布鲁克公司制造的VECTOR 22/N型傅立叶变换近红外光谱仪,扫描谱区为4 000~10 000 cm^{-1} ,扫描次数为20次,分辨率为8 cm^{-1} .

1.2.3 化学计量学方法 采用偏最小二乘法(PLS)^[8]建立数学模型,在布鲁克公司提供的OPUS定量分析软件上完成.

1.2.4 模型质量的判别 用内部交叉证实和外部检验对数学模型进行验证.通过比较模型的决定系数(R^2)和均方根差(RMSECV)来衡量模型的质量.均方根差和 R^2 由以下公式计算:

$$\text{均方根差} = \sqrt{\frac{1}{M} \sum (D_i)^2},$$

$$R^2 = 1 - \frac{\sum (D_i)^2}{\sum (y_i - y_m)^2},$$

公式中, D_i 表示第*i*个样品的化学值和预测值之差, M 为样品数, y_i 为第*i*个样品的化学值, y_m 为*M*个样品预测值的平均值.

2 结果与分析

2.1 常规分析基本参数

表1为150个样品的蔗汁锤度和还原糖分常规分析结果.本试验中样品蔗汁锤度(w)的变幅为5.20%~23.40%.在生产实践中,以全蔗茎样本为基础的蔗汁锤度出现高于24%的例子是很罕见的.甘蔗幼嫩未成熟节间的蔗汁锤度较低,变幅在5%~8%左右.因此,样品蔗汁锤度的变化范围具有足够大的覆盖面,样品有较好的代表性.蔗茎还原糖分的高低与品种、气候和栽培措施等因子有关.成熟蔗茎的还原糖分很低,未成熟蔗茎一般较高,成熟期的变幅一般在0.5%~3.0%.本试验中,样品的还原糖分的变幅为0.094 7%~3.958 8%.常规方法测定蔗汁锤度的随机误差较小,而测定还原糖分的随机误差则较大.由于兰-艾农法检出还原糖的痕量约为0.1%(w),在低浓度情况下有较大的测定误差.在本试验条件下,蔗汁锤度和还原糖分常规分析的误差控制在合理范围内.

表1 蔗汁样品常规分析基本参数

Tab. 1 Basic parameters of routine lab analysis in juice samples of sugarcane

项目 items	样品数 samples	极差 extremum	平均值 mean value	标准差 SD
蔗汁锤度 Brix in juice(w)/%	150	5.200~23.40	18.219	0.05
蔗汁还原糖分 reducing sugars in juice (w)/%	150	0.094 7~3.958 8	1.279 7	0.055 8

2.2 蔗汁的近红外光谱

图1为蔗汁近红外吸收光谱图.蔗汁中有大量的水分,还含有蔗糖、还原糖、有机酸、氨基酸等富含碳氢基团的有机物质.水分子在近红外谱区有强烈的吸收,在使用5.0 mm光程时,水分子在4 000~

5 300 cm^{-1} 和6 600~7 100 cm^{-1} 处吸收超界,在建立模型时这些谱区不能使用.在5 300~6 600 cm^{-1} 和7 100~10 000 cm^{-1} 分别是与碳氢基团有关的一倍频和二倍频吸收谱区,这些谱区为蔗汁锤度和还原糖分的定量分析提供了丰富的信息.

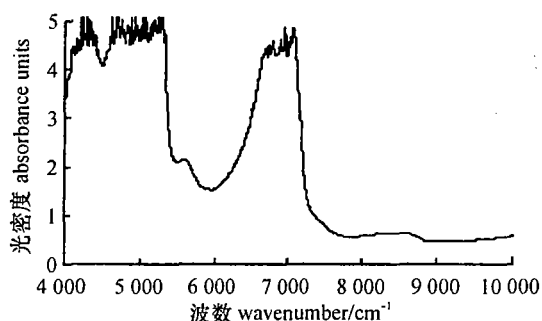


图1 光程为5.0 mm的甘蔗蔗汁近红外光谱图

Fig. 1 NIR spectra of sugarcane juice with optic path of 5.0 mm

2.3 数学模型的建立与优化

在利用 PLS 算法来建立预测数学模型时,需要选择合适的光谱数据预处理方法、谱区范围和主成分维数,作为建模的最佳条件. 利用在 OPUS 软件中设置的自动寻找和优化功能寻找建立模型的最佳条件. 通过比较各种可能组合下预测模型的决定系数 (R^2) 和均方根差,选取 R^2 尽可能大而均方根差尽可能小的组合,获得建模的最佳光谱数据预处理方法、最佳谱区范围和最佳主成分维数(表2).

表2 建立蔗汁锤度和还原糖分近红外预测模型的最佳条件

Tab. 2 The optimal conditions for establishing NIR predicted models of Brix and reducing sugars in juice of sugarcane

建模条件 conditions for models	蔗汁锤度 Brix in juice	蔗汁还原糖分 reducing sugars in juice
光谱预处理方法 spectral data preprocessing	最小最大归一化	一阶导数 + 矢量归一化
谱区范围 range of wavenumber / cm^{-1}	8 751 ~ 7 498, 5 778 ~ 5 450	10 000 ~ 7 498, 6 102 ~ 5 774
主成分维数 ranks	8	7

在常规分析和近红外光谱测定中误差的出现是不可避免的,因此,在建模时需要对样品光谱进行统计学上的显著性检验. 在进行 F 检验时,设置检验异常阈值(F 概率值)为 0.99,将每一个建模样品的化学值与预测值之间的误差与平均误差进行比较,得到相应的 F 值和 F 概率值. 经 F 值检验,剔除误差显著的异常样品.

另外,样品集的个数并非越多越好. 根据“少而精”的原则,对建模样品进行选择和优化. 在保证样

品集含量变幅大、含量梯度分布均匀的前提下,剔除在空间分布过度密集样品,同时尽量避免空间分布出现空洞.

由此而建立的数学模型获得良好效果. 蔗汁锤度和还原糖分数学模型的 R^2 分别为 0.999 0 和 0.995 3,均方根差分别为 0.115% 和 0.072 6%,预测误差接近常规方法的分析误差(表3). 预测值和化学值之间存在高度的相关性(图2).

表3 蔗汁锤度和还原糖分近红外预测模型的效果

Tab. 3 Calibration results of predicted models for brix and reducing sugars in juice

项目 items	样品数 samples	决定系数 R^2	均方根差 the root mean square errors of cross validation
蔗汁锤度 Brix in juice (w)/%	79	0.999 0	0.115 0
蔗汁还原糖分 reducing sugars in juice (w)/%	106	0.995 3	0.072 2

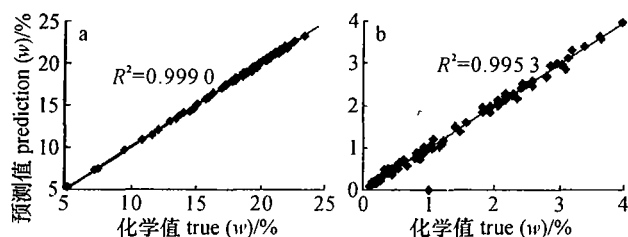


图2 蔗汁锤度(a)和还原糖分(b)化学值与预测值之间的相关性

Fig. 2 The correlations between trues and predictions of brix value (a) and polarization(b) in juice of sugarcane

2.4 预测效果验证

内部交叉证实是依次剔除建模样品集中一个(或多个)样品,用剩余样品来建模预测被剔除样品

的含量,比较被剔除样品预测值与化学值的差异,由此判断模型的预测准确性. 外部验证是选择一批与建模样品集独立无关的样品,通过比较独立样品预测值与化学值的差异来判断模型的预测准确性.

蔗汁锤度和还原糖分的内部交叉证实和外部验证结果列于表4. 无论是内部交叉证实还是外部验证,蔗汁锤度的绝对误差和相对误差都较小,表明有很高的预测准确性. 蔗汁还原糖分的绝对误差较小,相对误差则较大,但并不表明还原糖分模型的预测准确性很差. 因为还原糖分的测定值较小,相对于较小的测定值而言,蔗汁还原糖分预测结果有较大的相对误差是可以理解的,特别是在还原糖分很小的

情况下,由于还原糖的常规分析方法本身就存在着较大的随机误差,基于常规方法而建立起来的近红外预测模型也不可避免存在较大的误差,而还原糖分近红外预测数学模型的均方根差与常规方法的标

准差接近,可见,数学模型对还原糖分的预测准确性接近常规分析方法.因此,在实践上其预测结果是可以接受的.

表4 预测模型的内部交叉证实和外部验证结果

Tab. 4 The results of cross validation and independent validation for predicted models

验证方式 validation mode	项目 items	样品数 samples	绝对误差 absolute error			相对误差 relative error ¹⁾		
			最大	最小	平均	最大	最小	平均
			max.	min.	mean	max.	min.	mean
内部交叉证实 cross validation	蔗汁锤度 Brix in juice(w)/%	79	0.270	0.004 6	0.092 0	2.21	0.03	0.55
	蔗汁还原糖分 reducing sugars in juice(w)/%	93	0.221	0.000 1	0.055 1	42.59	0.02	6.88
外部验证 independent validation	蔗汁锤度 Brix in juice(w)/%	50	0.470	0.023 6	0.239 0	3.86	0.11	1.40
	蔗汁还原糖分 reducing sugars in juice(w)/%	25	0.383	0.003 0	0.141	44.34	1.39	23.92

1) 相对误差 = (|绝对误差| / 化学值) × 100%

3 讨论

蔗汁锤度的近红外分析数学模型有很高的预测准确性,利用近红外方法代替蔗汁锤度常规测定方法是完全可行的。蔗汁还原糖分的近红外分析数学模型的预测效果不如蔗汁锤度,主要与测定还原糖的常规方法误差较大有关。由于近红外法测定还原糖分的误差与常规方法的误差接近,其预测结果是可接受的。如果要进一步提高还原糖分近红外分析数学模型的准确性,必须要提高还原糖分常规分析的准确性,以及适当增加建模样品集的样品数量。

在先前的研究中,已经成功建立了蔗汁蔗糖分和转光度的近红外分析数学模型^[5]。蔗汁锤度和还原糖分分析数学模型的建立,使分析人员能够通过时间不超过20 s的一次扫描,就可同时得到蔗汁蔗糖分、转光度、锤度和还原糖分4个品质指标。如果再编制简单的计算程序,就可十分方便地由蔗汁蔗糖分、转光度和锤度计算出蔗汁重力纯度和简纯度。近红外光谱技术在甘蔗品质分析中的应用,为甘蔗品种选育和大田生产管理带来了极大的方便,也为糖厂原料蔗品质的管理、在线分析的实现和按质论价体系的建立提供重要的技术支持,展示了广阔的应用前景。

参考文献:

- [1] BROTHERTON G A, BERDING N, EGAN B T. Near infrared spectroscopic applications for milling: prospects and

implications[J]. Proceedings of the Australian Society of Sugar Cane Technologists, 1995, 17: 21 - 29.

- [2] SCHAFFLER K J, MEYER J H. Near infrared analysis of shredded cane: a potential replacement for the direct analysis of cane[J]. Proceedings of the Annual Congress South African Sugar Technologists' Association, 1996, 70: 131 - 139.
- [3] SCHAFFLER K J, de GAYE M T. Rapid near infrared estimation of multi-components in mixed juice and final molasses: The possibility of day to day control of raw sugar factories using NIR[J]. Proceedings of the Annual Congress South African Sugar Technologists' Association, 1997, 71: 153 - 160.
- [4] BERDING N; BROTHERTON G A. Near infrared reflectance spectroscopy for analysis of sugarcane from clonal evaluation trials: II. Expressed juice[J]. Crop Science, 1991, 31(4): 1 024 - 1 028.
- [5] 曹 干, 周学秋. 近红外光谱法快速测定甘蔗蔗汁蔗糖分的研究[J]. 甘蔗, 2000, 8(3): 20 - 25.
- [6] CAO G, TAN Z W, LIANG J N, et al. Mathematic models for analysis of quality components in sugarcane juice with fourier transform near infrared spectroscopy[J]. Agricultural Sciences of China, 2003, 2(2): 190 - 194.
- [7] 华南工学院. 制糖工业分析[M]. 北京: 轻工业出版社, 1984. 24 - 32, 114 - 123.
- [8] 吉海彦, 严衍禄. 用主成分回归和偏最小二乘法定量测定谷物成分[J]. 北京农业大学学报, 1994, 20(1): 59 - 63.

【责任编辑 周志红】