文章编号：1005-0906(2008)03-0057-03

近红外反射光谱测定玉米完整子粒蛋白质和淀粉含量的校正模型

王铁固[1,2],刘新香[1],库丽霞[1],魏良明[3],吴连成[1],陈彦惠[1]

(1.河南农业大学农学院,郑州 450002；2.河南科技学院农学系,河南 新乡 453003；3.河南省农科院粮食作物研究所,郑州 450002)

摘　要：以玉米完整子粒为实验材料,采用偏最小二乘回归法建立近红外反射光谱测定其蛋白质和淀粉含量的数学模型,光谱预处理结果表明采用矢量归一化和一阶导数+多元散射校正分别建立蛋白质含量和淀粉含量的校正模型效果最佳，外部验证结果证明校正后模型预测结果与化学值之间的相关系数分别达到0.946 9 和 0.924 0。

关键词：玉米;近红外反射光谱;品质分析;校正模型

中图分类号：S513　　　　**文献标识码：**A

Calibration Model of Determining Protein and Starch Contents of Whole Maize Kernel by Near Infrared Reflectance Spectroscopy

WANG Tie-gu[1,2], LIU Xin-xiang[1], KU Li-xia[1], WEI Liang-ming[3], WU Lian-cheng[1], CHEN Yan-hui[1]

(1. *Agronomy College of Henan Agricultural University, Zhengzhou* 450002;
2. *Department of Agronomy, Henan Institute of Science and Technology, Xinxiang* 453003;
3. *Food Crops Institute, Henan Academy of Agricultural Sciences, Zhengzhou* 450002, China)

Abstract: Protein and starch contents of whole maize kernel were determined by near infrared reflectance spectroscopy(NIRS). Math model was established by partial least square regression. Result of spectrum pretreatment indicated that the calibration models developed by vector normalization and first order derivative with multivariate scattering correction were the best for protein and starch, respectively. Exterior validation suggested that the correlations between the results of predicted NIRS by calibration and those obtained from chemical methods of protein and starch content were improved. The correlation coefficient is 0.9469 and 0.9240 respectively.

Key words: Maize; NIRS; Attribute Analysis; Calibration model

近红外反射光谱(Near Infrared Reflectance Spectroscopy,NIRS)分析是近年来迅速发展起来的一项测试技术，是利用有机化学物质在其近红外光谱区内的光学特性快速估测样品中的一项或多项化学成分含量，在国外已成为分析农作物品质的重要手段。利用近红外反射光谱进行玉米完整子粒的蛋白质和淀粉测定,可以克服化学方法对玉米子粒破坏的缺点,对玉米品质育种工作有重要的实践意义。本实验以玉米完整子粒为实验材料,采用偏最小二乘(partial least square,PLS)回归法建立近红外反射光谱测定其蛋白质和淀粉含量的数学模型,并进行校正和维护。

1　材料与方法

1.1　实验材料

按照建立 NIRS 校正模型的要求，设计标样。实验材料包括 73 个自交系、35 个杂交种(F_1)、50 个 F_2 材料(F_1 自交)。子粒样品分别来自 2003 年海南加代、2004 年河南农大试验田及河南省区试的材料。其中 108 份样品作校正集,50 份样品作验证集。

1.2　仪器与样品光谱测定

实验仪器为 MATRIX-I 型傅立叶变换近红外

收稿日期：2008-01-22
基金项目：国家“863”计划项目(2002AA207008)
作者简介：王铁固(1971-),男,河南伊川人,博士,从事玉米遗传育种研究。Tel:0371-63558032
E-mail:chy9890@163.com
陈彦惠为本文通讯作者。

光谱仪(德国BRUKER公司)。将30～50 g子粒样品盛于直径5 cm的旋转样品池中，在4 000～12 000 cm(波长800～2 400 nm)谱区范围内扫描64次，采集样品的吸收光谱。每个样品重复取样5次，计算其平均光谱贮存于计算机中。

1.3 子粒品质测定

蛋白质含量按照GB/T14771-12.191993进行测定；淀粉含量按照NY/T11-1985进行测定。所有样品均重复测定3次以上，结果以干基(%)表示。

1.4 NIRS数学模型的建立

对复杂有机物的光谱分析，要利用其全波长信息，运用化学计量学方法进行研究。本实验建立模型的化学计量学方法为偏最小二乘(PLS)回归法。在近红外漫反射光谱中，各种基团存在特定吸收峰，由于不同基团产生的光谱在吸收峰位、强度上不同，随着样品组成的变化，光谱特征也将发生变化。因此需要根据测定样品的近红外光谱与它们的组成或性质数据，通过合理的光谱处理及校正方法，建立样品光谱与组成或性质间的校正模型。外部验证是指用另一批完全独立的、化学成分已知的样品对模型预测效果进行检验。实验利用交叉验证均方差RMSECV、交叉验证决定系数R^2来衡量不同光谱预处理方法对模型的校正效果，利用预测值与化学测定值间的相关系数衡量校正后模型的优劣。采用BRUKER公司OPUS/QUANT 4.0商用光谱定量分析软件(Windows XP下运行)，进行上述光谱数据预处理、谱区选择与回归统计分析。

2 结果与分析

2.1 校正集和外部验证集样品中蛋白质和淀粉的含量分布

本研究所用的材料包括自交系、区试杂交种和F_2材料，其子粒类型包括马齿型、硬粒型及中间型，杂交种材料来自多种环境。从表1可以看出，子粒蛋白质、淀粉的含量范围变化分别为8.29%～15.52%和62.38%～74.55%，基本上覆盖了玉米生产和育种材料中蛋白质和淀粉含量的常态分布范围。

表1 校正集和验证集样品中蛋白质和淀粉含量分布

Table 1 The distributing of protein and starch content in emendation-gather specimen and validation-gather specimen %

统计参数 Statistical parameter	蛋白质含量 Protein content	淀粉含量 Starch content
校正集		
平均值	10.56	70.56
最小值	8.29	62.38
最大值	15.52	74.55
标准偏差	0.72	2.01
外部验证集		
平均值	10.77	72.16
最小值	8.50	66.88
最大值	13.26	73.99
标准偏差	1.25	2.12

2.2 光谱预处理对模型的校正效果

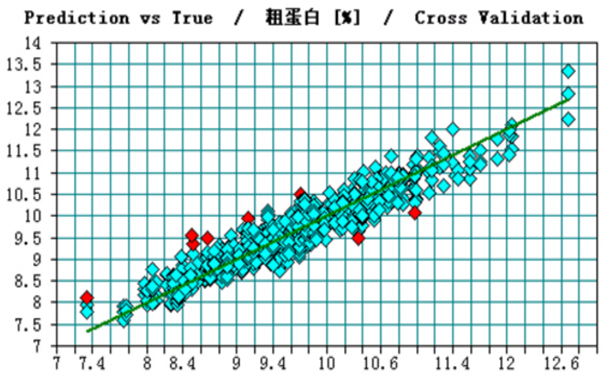
表2 不同光谱预处理方法对模型校正效果的比较

Table 2 The compare of model emendation effect with methods of different spectral pretreatment

预处理方法 Method of pretreatment	蛋白质含量 Protein content		淀粉含量 Starch content	
	R^2(%)	RMSECV	R^2(%)	RMSECV
光谱未处理	85.33	0.352	60.30	1.17
COE	86.95	0.332	61.64	1.15
SLS	87.85	0.320	63.48	1.12
多元散射校正	87.91	0.319	61.97	1.14
二阶导数	80.81	0.402	59.27	1.18
矢量归一化	89.02	0.303	61.86	1.14
一阶导数	86.34	0.339	62.42	1.14
一阶导数+SLS	87.01	0.331	63.04	1.13
一阶导数+多元散射校正	87.26	0.328	63.55	1.12
一阶导数+矢量归一化	87.13	0.329	62.07	1.14
最大最小归一化	88.39	0.312	62.53	1.13

各种光谱预处理方法对校正集实验材料玉米子粒蛋白质和淀粉含量预测模型的校正效果见表2。除二阶导数预处理的效果不理想外，其它光谱预处理方法都对提高蛋白质含量和淀粉含量的预测效果具有一定校正作用，但是相互之间差异不明显。相比较而言，对于蛋白质含量以矢量归一化

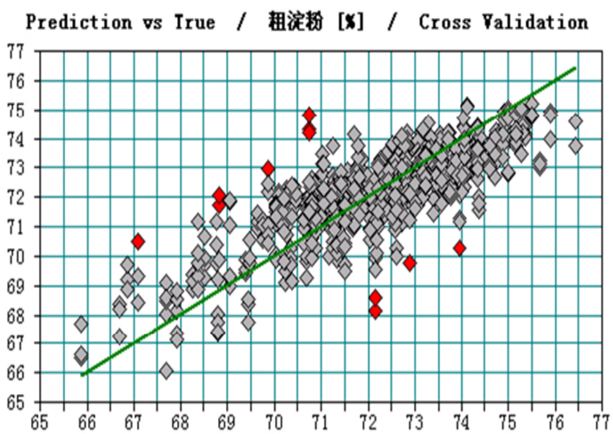
处理时的校正效果最好(图 1),其交叉验证的决定系数为 89.02%,交叉验证均方差为 0.303。对于淀粉含量来说，各种光谱预处理方法的交叉验证决定系数和交叉验证均方差分别在 60%和 1.15%左右,相比蛋白质含量来说其总体效果要差一些。比较而言，以一阶导数 + 多元散射校正光谱预处理方法的校正效果最好(图 2)。其交叉验证决定系数和交叉验证均方差分别为 63.55%和 1.12%。根据不同光谱预处理方法对模型校正结果影响的比较分析，分别采用矢量归一化和一阶导数 + 多元散射校正光谱预处理方法校正维护蛋白质含量和淀粉含量的数学模型。



Rank:10　　R^2=89.02　　RMSECV=0.303

图 1　蛋白质矢量归一化的校正效果

Fig.1　Emendation effect of vector unity of protein content



Rank:9　　R^2=63.55　　RMSECV=1.12

图 2　淀粉一阶导数 + 多元散射校正的校正效果

Fig.2　Emendation effect of uni-differential coefficien+ Mul-dispersion emendation of starch content

2.3　校正模型实际预测效果评价

为了进一步评价模型的质量和实际预测效果，随机选取与上述校正集不同的 50 份 F_2 材料对所建模型进行预测。按照国家和部颁标准方法测定它们的蛋白质和淀粉的化学值，并利用预测值与化学测定值间的相关系数和平均相对误差(ARE)对其预测效果进行评价。将玉米子粒两个品质性状模型校正前后预测结果与化学测定值间的相关系数和平均相对误差列入表 3。

表 3　预测值与化学值间的相关系数及 ARE

Table 3　Related coefficients of predicted mean and chemic mean and ARE

模型 Model	蛋白质含量 Protein content	淀粉含量 Starch content
校正前	0.9385**	0.8863**
校正后	0.9469**	0.9240**
ARE	-0.01	-0.01

从表 3 可以看出，校正后蛋白质含量和淀粉含量预测值与化学值的相关系数分别为 0.9469 和 0.9240,都达到极显著水平。校正后模型预测值与化学值之间的平均相对误差 ARE 也比较小,均为负值。上述外部验证结果证明校正后模型的准确性与化学方法相当。

3　讨　论

由于近红外光谱分析技术具有分析速度快、多组分同时测定、样品不需预处理、非破坏性分析、低分析成本和操作简单等优点,其应用已经受到越来越多人的关注，尤其是农作物领域中的应用进展迅速。

近红外光谱分析依靠样品间光谱信息的细微差别来对样品进行定量或定性分析，是一种物理的分析方法,分析结果受多种因素影响。大量的研究表明,粒度大小、样品均匀程度、水分含量和温度等的不同都会影响测试的准确性。试验中采用旋转样品池、重复装样后多次测量取平均值的方法,来减少玉米子粒样品因粒度大小、均匀程度等因素对光谱的影响。在校正维护模型的过程中,为减少水分对近红外预测值的影响，在定标时尽量使样品保持在稳定的、自然状态下的水分含量,并保证待分析样品水分含量与建立模型时条件相一致。对于温度对光谱的影响,采用在不同的温度下扫描样品光谱的方法，所建立的模型就具有温度的自动矫正功能。光谱预处理可有效地去除噪声,去除干扰因素的影响,使有用信息凸现出来,提高建模质量。由于近红外分析方法属于二次分析方法，所以其预测效果的好坏直接取决于化学方法测试的准确性。在试验中,有关化学成分的分析,采用国家和部颁标准方法，并进行 3 次以上的重复平行测定,减少实验误差。近红外光谱分析是从复杂、重叠、变动的背景中提取弱信息的一种分析技术,样品的代表性最重要,研究所选(下转第 63 页)

(上接第 59 页)用玉米子粒类型包括马齿型、硬粒型及中间型,区试材料来自多种环境;蛋白质和淀粉的含量范围变化基本上覆盖了玉米生产和育种材料中两种成分含量的常态分布范围，具有很好的代表性，所得到的校正模型的外部验证结果表明模型的预测值与化学值的相关系数分别达到 0.946 9 和 0.924 0，可以利用该模型无损预测玉米完整子粒蛋白质和淀粉的含量，指导玉米育种的品质遗传改良。

本实验所得到的校正模型仍不完善，还需要收集更多种类和数量的样品，同时注意细节进行校正和维护。

参考文献：

[1] 戴景瑞.我国玉米生产发展的前景及对策[J]. 作物杂志,1998(5):6-11.

[2] Norris K H, Barnes R F, Moore J E, et al. Prediction forage quality by NIRS[J]. Animal Sci., 1976,43(4): 899-897.

[3] William s P C, Mackemzie S L. Determination of methionine in peas by NIRS[J]. J. Agric, Food Chem., 1985, 33: 811-815.

[4] 徐广通,袁洪福,陆婉珍.现代近红外光谱技术及应用进展[J]. 光谱学与光谱分析,2000,20(2):134-142.

[5] Philip Williams, Karl Norris. Near infrared technology in the agriculture and food industries[J]. AACC, inc. St, MN USA, 2001(3): 6-11.

[6] 史永刚,冯新泸,李子存,等.近红外光谱分析中应考虑的几个问题[J]. 光谱实验室,2001(4):435-437.

[7] 陆婉珍,袁洪福,徐广通,等.现代近红外光谱分析技术[M]. 北京:中国石化出版社, 2000.

[8] 卢利军,庄树华,李爱军.应用近红外技术测定黄豆粕中水分、蛋白、粗脂肪[J]. 分子科学学报,2001,17(2):115-120.

[9] 魏良明,姜海鹰,李军会,等.玉米杂交种品质性状的近红外光谱分析技术研究[J]. 光谱学与光谱分析,2005,25(9):1404-1407.

（责任编辑：李万良）