

基于变量选择的小麦粗蛋白含量近红外光谱检测

吕程序¹ 姜训鹏² 张银桥¹ 张小超¹ 毛文华¹

(1. 中国农业机械化科学研究院, 北京 100083; 2. 中粮营养健康研究院有限公司, 北京 102209)

摘要: 小麦粗蛋白含量是其品质评价的重要指标,为探讨基于选择的短波近红外光谱变量定量判别小麦籽粒粗蛋白的可能性,采集了52份小麦籽粒样本,用湿化学方法分析其粗蛋白含量,获取其900~1700 nm波段的光谱,进而利用该光谱进行预处理方法的优化研究及小麦籽粒蛋白敏感变量的优选研究,以偏最小二乘的方法建立了基于短波近红外光谱的小麦籽粒蛋白定量模型。结果表明:多元散射校正和小波变换结合是短波近红外光谱定量判别小麦籽粒粗蛋白含量较优的预处理方法;利用200次竞争性自适应重加权变量优选的统计结果,优选出1 028、1 158、1 199、1 367、1 407、1 445、1 478、1 494、1 550、1 584、1 661、1 686 nm 12个变量为小麦籽粒蛋白敏感变量,占全谱的2%,该方法可稳定、高效地优选光谱变量,降低水分对模型的影响;结合预处理优化及变量优选建立偏最小二乘模型,模型预测决定系数和预测均方差分别为0.961和0.369。可见优选的短波近红外光谱变量可用于定量判别小麦籽粒粗蛋白含量。

关键词: 小麦;粗蛋白;近红外光谱;变量选择

中图分类号: S24 **文献标识码:** A **文章编号:** 1000-1298(2016)S0-0340-07

Variable Selection Based Near Infrared Spectroscopic Quantitative Analysis on Wheat Crude Protein Content

Lü Chengxu¹ Jiang Xunpeng² Zhang Yinqiao¹ Zhang Xiaochao¹ Mao Wenhua¹

(1. Chinese Academy of Agricultural Mechanization Sciences, Beijing 100083, China

2. COFCO Nutrition and Health Research Institute, Beijing 102209, China)

Abstract: Wheat is one of the main cereal grain which was produced not only in China but also at abroad. The aim of this paper is to study the feasibility of several selected variables from short wavelength near infrared spectroscopy to quantify crude protein of the whole wheat grain. In total, 52 whole wheat grain samples were collected, including 39 samples for calibration and 13 samples for validation. On the one hand, the crude protein of these samples were analyzed by using the Chinese standard of Kjeldahl method; on the other hand, those were scanned to obtain near infrared spectra with the wavelength range of 900~1700 nm by using a wheat analysis system developed by Chinese Academy of Agricultural Mechanization Sciences (CAAMS). Both of spectroscopic pretreatment method and sensitive variables were optimized, then the model was built by partial least squares regression method. The results showed the combination of multiple scattering correction and wavelet transform performed better. Competitive adaptive reweighted sampling method showed an efficient variable selection, which picking 12 variables and taking 2% of the full range spectral variables, including 1 028, 1 158, 1 199, 1 367, 1 407, 1 445, 1 478, 1 494, 1 550, 1 584, 1 661, 1 686 nm. Based on the optimized pretreatment method and selected variables, the model showed that the prediction determination coefficient and prediction root mean square error were 0.961 and 0.369, respectively. Competitive adaptive reweighted sampling variable selection

收稿日期: 2016-07-20 修回日期: 2016-08-12

基金项目: 国家国际科技合作专项(2014DFA31660)

作者简介: 吕程序(1985—),女,工程师,博士,主要从事农机机电技术研究,E-mail: lvchengxu@caams.org.cn

通信作者: 毛文华(1974—),女,研究员,博士生导师,主要从事农机机电技术研究,E-mail: mwh-924@163.com

based short wavelength near infrared spectroscopic technique showed a potential for crude protein quantitative analysis on whole wheat grain.

Key words: wheat; crude protein; near infrared spectroscopy; variable selection

引言

小麦是全球范围内的主要粮食作物,其产量仅次于玉米和水稻。我国是小麦生产大国,2014年小麦产量为1.26亿t,占全球小麦产量的17%^[1]。小麦籽粒通常经过研磨成粉后进一步加工使用。根据用途及后期加工特性,小麦粉可分为特高筋粉、高筋粉、中筋粉、低筋粉等,分类指标之一即为粗蛋白含量(质量分数)。因此,小麦籽粒粗蛋白含量的检测在小麦生产(种植、收获)、市场交易、食品加工等多个环节均有应用。

常用的小麦籽粒粗蛋白含量检测方法有凯氏定氮法、杜马氏燃烧法、近红外光谱法等。凯氏定氮法利用腐蚀性酸,把被测物中有机氮转化为氨,进而定量测量。杜马氏燃烧法通过燃烧把被测物中的氮元素转化成分子氮,进而利用热导检测器定量检测。凯氏定氮法、杜马氏燃烧法均为湿化学分析方法,测量精度高,用于仲裁检验,但有操作复杂、耗时长、污染环境等缺点。近红外光谱法具有快速、无损、无污染等优点,且基于近红外光谱的小麦籽粒蛋白质检测技术已日渐成熟。在中国及欧盟,基于小麦籽粒近红外光谱的蛋白分析方法已被列为标准方法^[2-3],但这些标准方法中均未明确细化谱区范围光谱等测量参数。谱区范围是光谱检测的重要参数,其直接决定模型的判别精度,也关系到测量的效率与成本。近红外光谱源于分子键的合频与倍频振动吸收,该区域光谱谱峰宽且高度重叠,部分共线性变量不仅不能为模型提供有效信息,反而由于引入噪声而影响模型精度,选择适合的谱区或敏感变量,可以提高模型精度及稳定性^[4]。选择的谱区或光谱敏感变量较全谱区更易获取。靠近可见光或短波近红外区域的光谱变量没有长波近红外区域光谱信息丰富,但该波段信号更易由廉价的检测器采集获得。因此研究基于敏感变量优选的短波近红外波段小麦籽粒的蛋白检测研究有重要实践意义与应用价值。

孙来军等^[5]、FOX等^[6]、POJIC等^[7]学者对小麦籽粒570~1100nm波段内的短波近红外光谱进行了研究,结果表明该波段可用于小麦籽粒粗蛋白含量的定量判别^[5-7]。李栓明等^[8]对小麦籽粒350~2500nm波段近红外光谱的蛋白敏感变量进行了筛选,结合无信息变量剔除、连续投影算法、逐步多元线性回归的3种算法,筛选出了15个敏感变

量,该优选的变量用于小麦籽粒蛋白定量判别,预测均方根误差为0.59。宦克为等^[9]对小麦籽粒950~1690nm波段近红外光谱的蛋白敏感变量进行了筛选,结合蒙特卡罗采样技术与特征投影图方法,优选出12个敏感变量,该优选的变量用于小麦籽粒蛋白定量判别较全波段模型预测精度高,预测均方根误差为0.25。李勇等^[10]对基于833~2500nm波段光谱的小麦籽粒蛋白质模型进行了研究,结果表明选择有效区间可避开水分特征区间,降低水分对模型的影响。

LI等^[11]于2009年提出竞争性自适应重加权(Competitive adaptive reweighted sampling, CARS)的变量选择理论,该方法综合无信息变量有效去除与共线性变量删减2种手段,可高效地优选关键敏感变量。该方法模拟达尔文生物进化论的自然选择和遗传学机理的生物进化过程,每次优选结果均不相同。为获得较稳定、一致的CARS变量选择结果,本文利用多次CARS运算的统计结果来优选敏感变量,并对优选的短波近红外光谱变量定量判别小麦籽粒粗蛋白含量进行研究。

1 材料与方法

1.1 材料制备

收集不同产地、品种的成熟小麦籽粒样本52个。剔除虫蚀、破损、霉变、病斑、生芽等不完善籽粒和杂质后,每个样本取1kg自然晾晒风干。风干的样本通过四分法分成2份:其中一份按照GB/T 5511—2008的凯氏方法测定粗蛋白含量,并按照GB/T 21305—2007的干燥法测定含水率;另一份抽真空避光于4℃样品冷藏箱中保存,用于扫描光谱。

为了探讨预处理方法及变量选择的优化,针对52个小麦籽粒样本,根据其粗蛋白含量,按隔三取一的原则分集,其中39个样本被划分为定标集,13个样本被划分为验证集。为探讨蛋白质敏感变量选择与消除水分对模型影响的关系,根据样本的含水率,分别设计2个独立含水率的定标集和2个独立含水率的验证集,研究特定含水率范围的蛋白质模型预测含水率范围外小麦籽粒样本粗蛋白含量的能力。具体方法为:根据含水率排序定标集的39个样本,取低含水率的13个样本,定义为低含水率定标集,后文简称C1;取高含水率的13个样本,定义为高含水率定标集,后文简称C2。这就保证了

C1、C2 两个定标集样本的含水率独立、不交叉。在验证集 13 个样本中,选取含水率在 C2 含水率范围内,粗蛋白含量在 C1 粗蛋白含量范围内的样本,作为高含水率验证集,后文简称 P1;选取含水率在 C1 含水率范围内,粗蛋白含量在 C2 粗蛋白含量范围内的样本,作为低含水率验证集,后文简称 P2。

1.2 光谱采集

使用中国农业机械化科学研究院研制的 CA-06 型光栅光谱分析仪^[12],包括:固定光栅光学系统、CCD 驱动控制系统、CCD 信号采集系统、探测器温度控制单元、高稳定度光源控制系统、旋转样品池及其控制单元和计算机信号处理系统 7 部分。

实验开始前,将冷藏的小麦籽粒样品在实验室放置 24 h 至温度达到室温(20~22℃)。小麦籽粒光谱扫描参数为:旋转样本杯模式采集漫反射光谱,光谱范围:900~1 700 nm,分辨率:1.56 nm,采集数据点数:512,相同样本重复装样次数:3,每条光谱重复扫描次数:10,采集光谱以 ASCII 格式输出。

1.3 数据处理

所有的数据处理均在 Matlab 7.0 软件(The MathWorks, 美国)配合 PLS_toolbox 6.7(Eigenvector research, 美国)和 CARS 2.0 工具包实现。

模型的建立使用偏最小二乘(Partial least squares, PLS)的方法,使用全交互验证的方法,根据交互验证均方差最小(RMSECV)的方法确定建模潜变量数(Latent variable, LVs)。模型的评价指标有:定标集、全交互验证集、独立验证集的决定系数和均方差,分别用 R_C^2 、 R_{CV}^2 、 R_p^2 、RMSEC、RMSECV、RMSEP 表示。

为减小随机噪声对光谱信号的影响、提高光谱

信号的信噪比,采用 SG9 点平滑和小波变换(Wavelet transform, WT)的方法预处理光谱,其中小波基选用多贝西 5 阶小波,分解层为 5 层。为消除光谱信号的基线漂移、放大光谱信号差异,选用 SG9 点一阶和二阶导数的方法预处理光谱。为消除样本粒度不均引起的散射影响,使用多元散射校正(Multiple scattering correction, MSC)和标准正态变量(Standard normal variate, SNV)的方法预处理光谱。基于单一预处理方法优选结果,探讨适宜小麦籽粒粗蛋白含量建模的最优预处理组合方法。

利用 2 种基于 CARS 的变量优选方法。一种是通过 200 次 CARS 变量优选,综合适度函数值与选择变量数目,选择较优的一次优选变量结果。另一种是通过统计 200 次 CARS 变量优选的结果,根据优选变量的分布,选择所有优选变量区间内的高频被选变量。为验证多次统计 CARS 变量优选结果的稳定性,共计 3 组 CARS 运算结果被统计,每组包括 200 次 CARS 运算。CARS 运算使用定标集光谱数据,选用优选的预处理方法,提取的最大主成分数为优选预处理模型中的 LVs,蒙特卡罗取样次数为 200,内部验证选用 7 层交叉检验。

2 结果与讨论

2.1 小麦籽粒与光谱

小麦籽粒粗蛋白含量和含水率如表 1 所示,可见小麦籽粒的粗蛋白含量在 11.07%~16.93% 之间,分布较为广泛。水分在近红外区域吸收显著,为影响检测精度的关键因素。本研究中小麦籽粒经过风干的预处理,含水率在 9.66%~12.28% 之间,含水率的差异较小。

表 1 小麦籽粒样本粗蛋白含量及含水率

Tab.1 Protein and moisture content of wheat samples

样本集	样本数	粗蛋白含量/%				含水率/%			
		最小值	最大值	均值	方差	最小值	最大值	均值	方差
定标集	39	11.07	16.37	14.16	1.33	9.81	12.28	11.03	0.49
验证集	13	11.75	16.93	14.25	1.39	9.66	12.28	11.01	0.73

52 个小麦籽粒样本的短波近红外光谱图如图 1 所示。在 900~1 700 nm 波段,小麦籽粒存在明显的吸收峰,该区域光谱主要为 C—H、N—H、O—H 等键的倍频及合频吸收。其中,波长 970 nm 和 1 450 nm 为水分子 O—H 键对称和反对称伸缩振动的组合频吸收谱带^[13-14],由图 1 可见,1 408~1 602 nm 波段的水分吸收对光谱的影响显著;1 150~1 210 nm 波段为 C—H 键振动的二级倍频谱带;1 300~1 500 nm 波段为 C—H 键的组合频吸收谱带;1 470~1 590 nm 波段

为 N—H 键的倍频吸收区域。

900~1 700 nm 波段小麦籽粒光谱复杂。不同来源的吸收峰交错重叠,显示为以波长 1 185 nm 和 1 430 nm 为中心的 2 个宽峰。谱图曲线存在一定的噪声。受基质及外界环境因素的影响,光谱存在谱峰漂移,且因小麦籽粒粒度影响,还存在一定的散射。

2.2 全波段模型

利用 700~1 900 nm 全波段光谱建立小麦籽粒

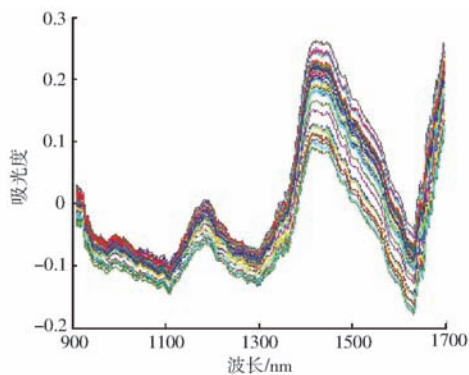


图1 小麦籽粒近红外光谱图

Fig. 1 Near infrared spectra of wheat

粗蛋白定量模型,其验证结果如表2所示。

由于图1所示的原始光谱存在大量的噪声,平滑和小波变换2种降噪预处理方式较无预处理的模型, R_{CV}^2 增加2%,RMSECV减少5%,模型质量略微提高。一阶导数和二阶导数模型较无预处理模型, R_{CV}^2 减少2%~13%,RMSECV增加5%~33%,导数处理使模型质量较原始光谱模型降低,因为导数的预处理方法在放大光谱信号的同时也放大了噪声,在本文光谱噪声较大的情况下,导数处理并不能改善模型质量,反而导致模型质量下降,且高阶导数模型质量下降更加显著。由图1可见原始光谱存在散

表2 全波段模型结果

Tab. 2 Model built with whole spectral range

预处理	LVs	R_C^2	R_{CV}^2	R_p^2	RMSEC	RMSECV	RMSEP
无	7	0.939	0.853	0.963	0.325	0.504	0.269
平滑	7	0.945	0.866	0.969	0.307	0.480	0.268
小波变换	7	0.944	0.867	0.972	0.310	0.480	0.266
一阶导数	6	0.930	0.839	0.946	0.348	0.529	0.355
二阶导数	6	0.917	0.739	0.927	0.379	0.671	0.416
MSC	6	0.948	0.857	0.967	0.298	0.499	0.271
SNV	6	0.947	0.851	0.972	0.303	0.511	0.245
MSC + 平滑	6	0.950	0.868	0.962	0.292	0.479	0.274
MSC + 小波变换	6	0.948	0.872	0.959	0.299	0.473	0.290

射,使用MSC和SNV2种数据标准化的预处理方式,其中MSC提高了模型质量,较无预处理模型的 R_{CV}^2 提高5%,RMSECV减少1%;但SNV预处理方法并未提高模型质量。

以 R_{CV}^2 和RMSECV为评价指标、以MSC结合小波变换为小麦籽粒粗蛋白含量短波近红外光谱的最优预处理方式,模型结果如图2所示, R_p^2 为0.959, RMSEP为0.290。

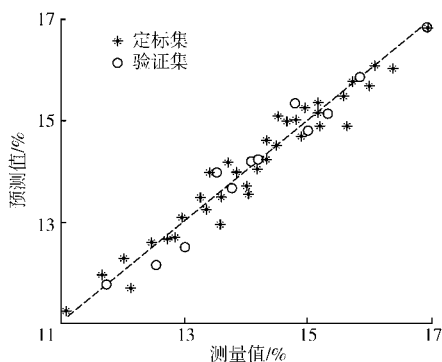


图2 全波段小麦籽粒粗蛋白定量模型

Fig. 2 Wheat protein model with whole range spectra

2.3 单次CARS变量优选模型

以MSC结合小波变换的定标集光谱数据为基础,选取RMSECV为适度函数,进行200次CARS变量优选的结果如图3所示。可见CARS算法优选后的光谱变量为6~40个,仅占全谱512个光谱

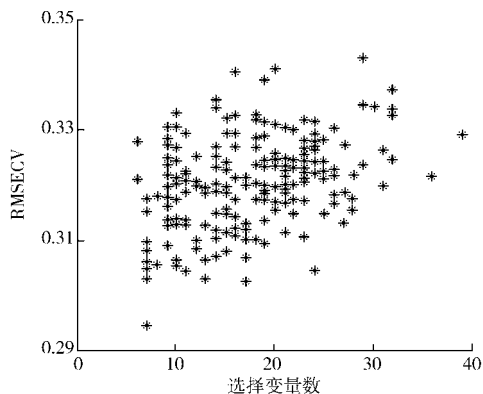


图3 CARS运算结果

Fig. 3 CARS results

变量的1%~8%,显著减少了92%以上的光谱变量。

兼顾选择变量数少和RMSECV小的原则,当选择波长1444、1480、1481、1552、1658、1661、1689 nm处的7个变量时,适度函数值最小。利用该7个变量建立小麦籽粒粗蛋白定量模型,结果如图4所示,定标集、交互验证集、独立验证集的 R^2 分别为0.940、0.924、0.912;均方差分别为0.323、0.362、0.429。该单次CARS优选变量的模型较全谱模型减少了99%的输入数据量,虽然定标集和独立验证集的评价指标(决定系数和均方差)较全波段模型略差,但优选变量模型明显改善了全交互验

证集的预测精度,表明基于 CARS 的变量优选方法有改善小麦籽粒短波近红外光谱模型稳定性的潜力。

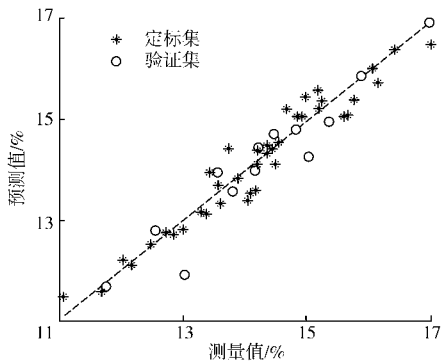


图4 CARS 优选7变量的小麦籽粒粗蛋白定量模型

Fig.4 Wheat protein model with seven variables selected by CARS

2.4 多次统计 CARS 变量优选模型

2.4.1 统计的变量优选结果

图3所示的200次CARS优选变量统计结果如图5所示。可见200次CARS运算所选敏感变量分别集中在1 025~1 032 nm、1 141~1 164 nm、1 197~1 204 nm、1 360~1 380 nm、1 402~1 422 nm、1 437~1 456 nm、1 470~1 483 nm、1 487~1 497 nm、1 540~1 572 nm、1 577~1 593 nm、1 653~1 668 nm、1 682~1 697 nm 12个变量区间内。该12个优选变量区间内的高频被选变量分别为1 028、1 158、1 199、1 367、1 407、1 445、1 478、1 494、1 550、1 584、1 661、1 686 nm。

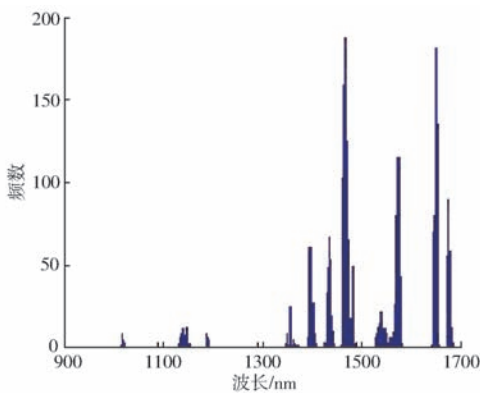


图5 CARS 选择波长统计结果

Fig.5 Statistical result of variables selected by CARS

2.4.2 优选变量建模与验证

利用该12个变量建立模型结果如图6所示,定标集、交互验证集、独立验证集的 R^2 分别为0.949、0.922、0.961;RMSEC、RMSECV和RMSEP分别为0.296、0.368、0.369。该模型较单次CARS变量优选模型, R^2_{CV} 和RMSECV结果相似,但 R^2_p 、RMSEC和RMSEP有了明显的改善,其中RMSEP减少了14%。

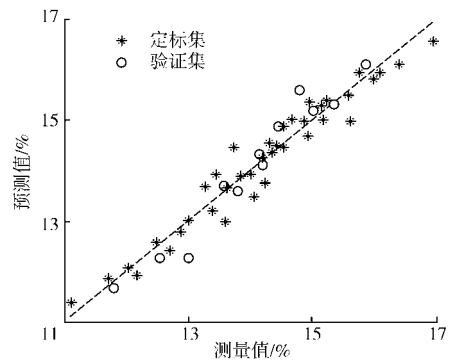


图6 CARS 优选12变量的小麦籽粒粗蛋白定量模型

Fig.6 Wheat protein model with 12 variables selected by CARS

该模型较全波段模型 R^2_{CV} 和RMSECV明显改善,但RMSEP增加。

2.4.3 多次统计 CARS 优选变量稳定性

为验证多次统计CARS优选变量结果的稳定性,图7所示为分别重新计算2组200次CARS优选变量统计结果。可见图7重新运算2次的统计结果与图5所示一致,均存在12个变量集中分布区间。其中图7a所示的12个优选变量区间内的高频被选变量分别为1 030、1 156、1 198、1 367、1 406、1 445、1 476、1 494、1 551、1 584、1 661、1 689 nm;图7b所示的12个优选变量区间内的高频被选变量分别为1 030、1 149、1 198、1 367、1 406、1 445、1 476、

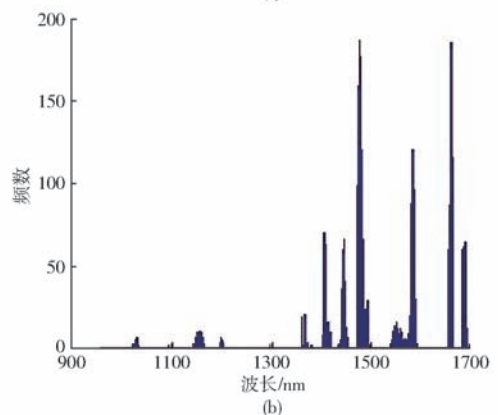
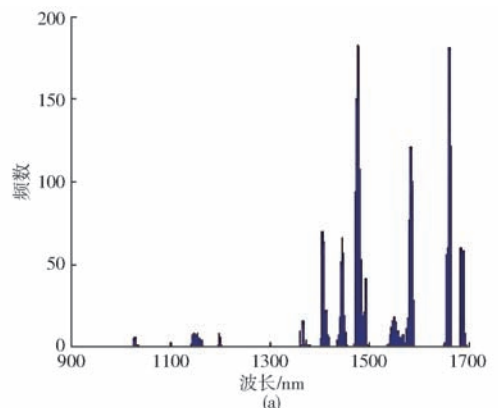


图7 2组CARS选择波长统计结果

Fig.7 Statistical result of variables selected by CARS

1 494、1 550、1 584、1 661、1 684 nm。图 5 和图 7 所示 3 组多次统计 CARS 优选变量的标准偏差为 0 ~ 4.7 nm, 3 组优选的 12 个变量的平均标准偏差为 1 nm。由于临近的近红外光谱变量存在高度的共线性, 3 组多次统计 CARS 优选变量结果表明该方法可以稳定地优选小麦籽粒蛋白近红外敏感变量。

2.4.4 优选的变量与已有研究比较

已有研究对蛋白的敏感谱带进行了解析。973 ~ 1 020 nm附近为蛋白质 N—H 键伸缩振动的二倍频吸收^[8], 波长 1 010 nm 为蛋白酰胺 N—H 键伸缩振动二倍频吸收^[15], 1 125 ~ 1 152 nm 为小麦粗蛋白的特征吸收谱带^[16], 1 435 ~ 1 446 nm 为小麦粗蛋白的特征吸收区域^[16], 1 506 nm 为蛋白质的特征吸收区域^[17], 1 500 ~ 1 530 nm 源于蛋白质 N—H 伸缩振动一级倍频吸收^[8]。这与本研究变量优选的结果一致。

2.4.5 蛋白质敏感变量选择与消除水分影响的关系探讨

独立含水率定标集和验证集的样本信息如表 3 所示。可见 C1 定标-P1 验证的组合中, 验证集的粗蛋白含量 13.00% ~ 14.78% 在定标集粗蛋白含

量 12.01% ~ 15.21% 之内, 验证集的含水率 11.85% ~ 12.09% 完全不在定标集的含水率 9.81% ~ 10.91% 之内; C2 定标-P2 验证的组合中, 验证集的粗蛋白含量 12.54% ~ 15.02% 在定标集粗蛋白含量 11.07% ~ 16.37% 之内, 验证集的含水率 10.28% ~ 10.71% 完全不在定标集的含水率 11.37% ~ 12.28% 之内。为比较考察样本水分对全谱段模型和优选变量模型二者的影响, 使用上述 2 组定标-验证组合方式, 分别构建并评价全谱段 MSC + 小波变换的小麦籽粒蛋白质模型和优选的 12 个光谱变量 MSC + 小波变换的小麦籽粒蛋白质模型, 结果如表 4 所示。表 4 可见两组预测方式优选变量模型较全谱段模型均显著降低了 RMSEP, 降幅为 47% ~ 56%, 优选变量模型的效果较优。这表明, 基于变量优选的小麦籽粒蛋白质模型预测能力受定标及验证样本水分的干扰较全谱段模型小, 这与已有文献报道研究结果一致^[10]。由于本文小麦籽粒样本量较少, 从中挑选特定蛋白质及含水率的样本的能力有限, 后期研究团队将考虑扩充样本量, 进一步讨论优选变量与消除水分影响之间的关系。

表 3 独立含水率定标集和验证集样本信息

Tab. 3 Information of calibration set and verification set for independent moisture content

样本集	符号	样本数	粗蛋白含量/%				含水率/%			
			最小值	最大值	均值	方差	最小值	最大值	均值	方差
低含水率定标集	C1	13	12.01	15.21	13.79	1.08	9.81	10.91	10.17	0.36
高含水率定标集	C2	13	11.07	16.37	14.26	1.61	11.37	12.28	11.69	0.33
高含水率验证集	P1	3	13.00	14.78	13.97	0.90	11.85	12.09	11.94	0.13
低含水率验证集	P2	5	12.54	15.02	13.94	0.95	10.28	10.71	10.59	0.18

表 4 固定含水率模型蛋白预测结果

Tab. 4 Predicted results of protein for fixed moisture content

组合方式	模型类型	定标集		验证集	
		R_c^2	RMSEC	R_p^2	RMSEP
C1 定标-P1 验证	全谱段	0.896	0.335	0.342	0.827
	优选变量	0.976	0.161	0.992	0.438
C2 定标-P2 验证	全谱段	0.895	0.503	0.774	0.762
	优选变量	0.956	0.325	0.974	0.338

谱定量判别小麦籽粒粗蛋白含量较优的预处理方法。利用竞争性自适应重加权的方法可以高效、稳定地优选光谱敏感变量, 减小水分对模型判别的影响, 优选结果为 1 028、1 158、1 199、1 367、1 407、1 445、1 478、1 494、1 550、1 584、1 661、1 686 nm 的 12 个变量, 占全谱数据点个数的 2%。结合预处理优化及变量优选建立偏最小二乘模型, 模型预测决定系数和预测均方差分别为 0.961 和 0.369。优选的短波近红外光谱变量可用于定量判别小麦籽粒粗蛋白含量。

3 结束语

多元散射校正和小波变换结合是短波近红外光

参 考 文 献

- 1 国家统计局. 2014 年国民经济和社会发展统计公报[R]. 中华人民共和国国家统计局, 2015-02-26.
- 2 DIN EN 15948—2012 Cereals-determination of moisture and protein-method using near infrared spectroscopy in whole kernels [S]. 2015.
- 3 GB/T 24899—2010 粮油检验-小麦粗蛋白含量测定-近红外法[S]. 2010.
- 4 ZOU X, ZHAO J, POVEY M J W, et al. Variables selection methods in near-infrared spectroscopy[J]. Analytica Chimica Acta,

- 2010, 667(1-2): 14-32.
- 5 孙来军,王乐凯,钱海波,等. 基于近红外透射光谱分析技术的小麦蛋白含量测定[J]. 中国农学通报,2011,27(3):35-39.
SUN L J, WANG L K, QIAN H B, et al. Determination of protein content in wheat based on near infrared transmittance spectroscopy[J]. Chinese Agricultural Science Bulletin, 2011,27(3):35-39. (in Chinese)
 - 6 FOX G P, BLOUSTEIN G, SHEPPARD J. On-the-go NIT technology to assess protein and moisture during harvest of wheat breeding trials[J]. Journal of Cereal Science, 2010,51(1):171-173.
 - 7 POJIĆ M, RAKIĆ D, LAZIĆ Ž. Chemometric optimization of the robustness of the near infrared spectroscopic method in wheat quality control[J]. Talanta,2015,131:236-242.
 - 8 李栓明,郭银巧,王克如,等. 小麦籽粒蛋白质光谱特征变量筛选方法研究[J]. 中国农业科学,2015,48(12):2317-2326.
LI S M, GUO Y Q, WANG K R, et al. Research on variable selection of wheat kernel protein content with near infrared spectroscopy[J]. Scientia Agricultura Sinica,2015,48(12):2317-2326. (in Chinese)
 - 9 宦克为,刘小溪,郑峰,等. 基于蒙特卡罗特征投影法的小麦蛋白质近红外光谱测量变量选择[J]. 农业工程学报,2013,29(4):266-271.
HUAN K W, LIU X X, ZHENG F, et al. Variable selection of near-infrared spectroscopy for measuring wheat protein based on MC-LPG[J]. Transactions of the CSAE,2013,29(4):266-271. (in Chinese)
 - 10 李勇,魏益民,张波,等. 近红外水分稳健分析模型研究[J]. 光谱学与光谱分析,2005,25(12):1963-1967.
LI Y, WEI Y M, ZHANG B, et al. Study on the robust NIR calibration models for moisture[J]. Spectroscopy and Spectral Analysis,2005,25(12):1963-1967. (in Chinese)
 - 11 LI H, LIANG Y, XU Q, et al. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration[J]. Analytica Chimica Acta,2009,648:77-84.
 - 12 张银桥,赵博,王辉,等. 基于固定光栅和 CCD 技术的近红外光谱分析系统设计[J]. 农业工程学报,2009,25(9):176-181.
ZHANG Y Q, ZHAO B, WANG H, et al. Technology of near infrared spectroscopy analysis system design based on fixed grating and charge-coupled device[J]. Transactions of the CSAE, 2009,25(9):176-181. (in Chinese)
 - 13 WORKMAN J, WEYER L. Practical guide to interpretive near infrared spectroscopy [M]. London: Taylor & Francis, 2007: 239-263.
 - 14 BURNS D A. Handbook of near infrared analysis third edition [M]. London: Taylor & Francis,2007:263.
 - 15 PEDERSEN D K, MARTENS H, NIELSEN J P, et al. Near infrared absorption and scattering separated by extended inverted signal correction: analysis of near infrared transmittance spectra of single wheat seeds[J]. Applied Spectroscopy,2002,56(9): 1206-1214.
 - 16 ZHU D, WANG K, ZHANG D, et al. Quality assessment of crop seeds by near infrared hyperspectral imaging[J]. Sensor Letters,2011,9(3):1144-1150.
 - 17 BUTKUTE B, CESEVICIENE J. Development of equations for wheat grain quality assessment by a NIRS-6500 spectrometer I Grain quality and optical data base characterization[J]. Zemdirbyste Agriculture,2009,96(4):62-77.