

赵爽, 宋京燕, 陈国兴, 等. 基于水稻种粒高光谱的品质性状预测方法[J]. 华中农业大学学报, 2023, 42(3): 211-219.
DOI: 10.13300/j.cnki.hnlkxb.2023.03.025

基于水稻种粒高光谱的品质性状预测方法

赵爽¹, 宋京燕¹, 陈国兴¹, 宋鹏¹, 冯慧², 杨万能¹

1. 华中农业大学作物遗传改良全国重点实验室, 武汉 430070;
2. 华中农业大学信息学院/农业生物信息湖北省重点实验室, 武汉 430070

摘要 为探究作物种子品质性状无损检测方法, 以100份水稻核心种质资源为试验材料, 基于近红外高光谱相机采集水稻种粒的透射、反射光谱数据并提取光谱参数, 水稻种粒脱壳后使用近红外谷物分析仪测定米粒品质; 以水稻种粒光谱参数为自变量、米粒品质指标为因变量, 建立米粒品质预测模型。结果显示, 使用单一光谱建模时, 透射光谱建模效果优于反射光谱建模效果; 结合透射-反射特征光谱集合建模可使粗蛋白预测模型 R^2 从0.74提高至0.91, 可使直链淀粉预测模型 R^2 从0.40提高至0.69, 可使水分预测模型 R^2 从0.53提高至0.68。结果表明, 使用水稻种粒光谱参数可无损预测稻米品质, 同时利用透射、反射光谱可提升建模效果。

关键词 水稻; 稻米品质; 无损检测; 高光谱; 预测模型

中图分类号 TP391.41; TS210.7 **文献标识码** A **文章编号** 1000-2421(2023)03-0211-09

水稻作为我国主要的粮食作物之一, 在保障粮食安全方面具有不可替代的作用, 在国民经济发展和社会稳定中具有重要意义。随着生活品质的提升, 国民对稻米的品质要求也逐渐提高。直链淀粉含量和蛋白质含量是衡量稻米品质的2个关键指标^[1]。直链淀粉含量丰富的米饭蓬松、偏硬、黏性小; 直链淀粉含量少的米饭软且黏性大^[2]。稻米中的蛋白质在人体中容易被吸收、利用率高, 摄入稻米是人体从外界获取蛋白质的途径之一^[3]。稻米与贮藏环境直接接触, 水分含量过高或过低不但影响稻米的食用品质, 甚至可能导致稻米霉变、虫变^[4]。因此, 在育种、加工生产中, 有必要对稻米直链淀粉、蛋白质和水分含量进行测定^[5]。

传统的稻米直链淀粉和蛋白质含量测定采用生化方法。常见的蛋白质含量测定方法有Folin-酚法、凯氏定氮法、双缩脲法、考马斯亮蓝法等^[6]。常见的直链淀粉含量测定方法有碘比色法、电流滴定法、电压滴定法等^[7-8]。常见的水分含量测定方法为烘干法^[9]。这些测定方法低效、有损且样本不能被再次利用^[10]。高光谱因其波段众多, 涵盖光谱和空间信息, 可以获得高分辨率的图像和光谱参数^[11]。不同物质的结构以及化学成分不同, 其光谱反射率、透射率有

所差异, 使用高光谱相机无损采集水稻种粒的光谱数据, 提取出光谱参数结合米粒品质指标可建立相应的预测模型。传统的稻米品质测定方法均需要对种粒脱壳再进行测定, 利用高光谱技术可避免对种粒脱壳、达到无损检测的目的。

目前, 高光谱无损检测农产品品质性状技术已经得到广泛应用。Ma等^[12]基于高光谱技术建立了水稻籽粒蛋白质含量的偏最小二乘回归模型, 其 R^2 达0.801, 均方根误差为0.52。戴国新等^[13]采用逐步线性回归建立了基于高光谱的水稻精米直链淀粉和蛋白质含量预测模型, R^2 分别可达0.823和0.837。Sun等^[14]基于光谱性状对水稻种子蛋白质含量做全基因组关联分析, 光谱性状能够鉴定到与传统方法相同的SNP位点, 表明通过光谱可建立有效的水稻种子蛋白质含量预测模型。张晶等^[15]基于近红外高光谱采用偏最小二乘法建立了荞麦营养成分快速检测模型, 其中蛋白质、淀粉含量预测模型的决定系数分别为99.97%、99.86%, 且验证结果表明预测值和真实值差异不显著。梁秀英等^[16]基于近红外高光谱成像技术采用偏最小二乘回归法、小生境蚁群算法、支持向量机等方法建立了玉米籽粒蛋白质含量预测模型, 其预测效果均优于原始光谱数据预测效果。

收稿日期: 2022-10-06

基金项目: 国家自然科学基金联合基金项目(U21A20205); 中央高校基本科研业务费专项(2662021JC008)

赵爽, E-mail: 1006852680@qq.com

通信作者: 冯慧, E-mail: fenghui@mail.hzau.edu.cn

Heman等^[17]利用近红外高光谱技术结合偏最小二乘法建立了稻米水分含量无损预测模型,其 R^2 大于0.9。于宏威等^[18]利用回归系数法筛选特征波段建立花生蛋白质含量的光谱预测模型,结果表明基于高光谱技术测定花生中蛋白质含量具有可行性。李栓明等^[19]、孟庆琰等^[20]、曹志敏等^[21]、曹晓宁等^[22]也基于高光谱参数建立了农产品品质性状无损预测模型。前人的研究多是采用传统的光谱处理方法,预处理方法主要包括平滑^[23]、导数法^[24]、多元散射校正^[25-26]及特征筛选算法建立多种回归预测模型。这些研究表明高光谱技术在无损检测领域具有应用潜力。

本研究基于近红外波段高光谱相机采集100份水稻核心种质资源材料的谷粒光谱数据,分别以反射光谱、透射光谱以及结合反射-透射光谱作为模型自变量,以米粒品质指标作为因变量,通过相关性分

析、逐步线性回归等方法筛选重要特征波段,建立米粒直链淀粉、粗蛋白和水分含量预测模型,旨在为作物种子品质性状无损检测提供科学依据。

1 材料与方法

1.1 试验材料

从533份水稻核心种质资源材料中随机选取100份作为试验材料。每份材料随机取100粒水稻种粒进行试验,种粒光谱采集、米粒品质测定时间为2021年9月,试验地点为华中农业大学作物遗传改良国家重点实验室作物表型中心。

1.2 试验设计与流程

技术路线如图1所示,主要包括光谱参数提取、米粒品质指标获取以及模型构建三部分。本研究基于高光谱表型平台采集水稻种粒光谱数据,经图像分割、图像掩膜提取光谱参数。米粒品质指标由近

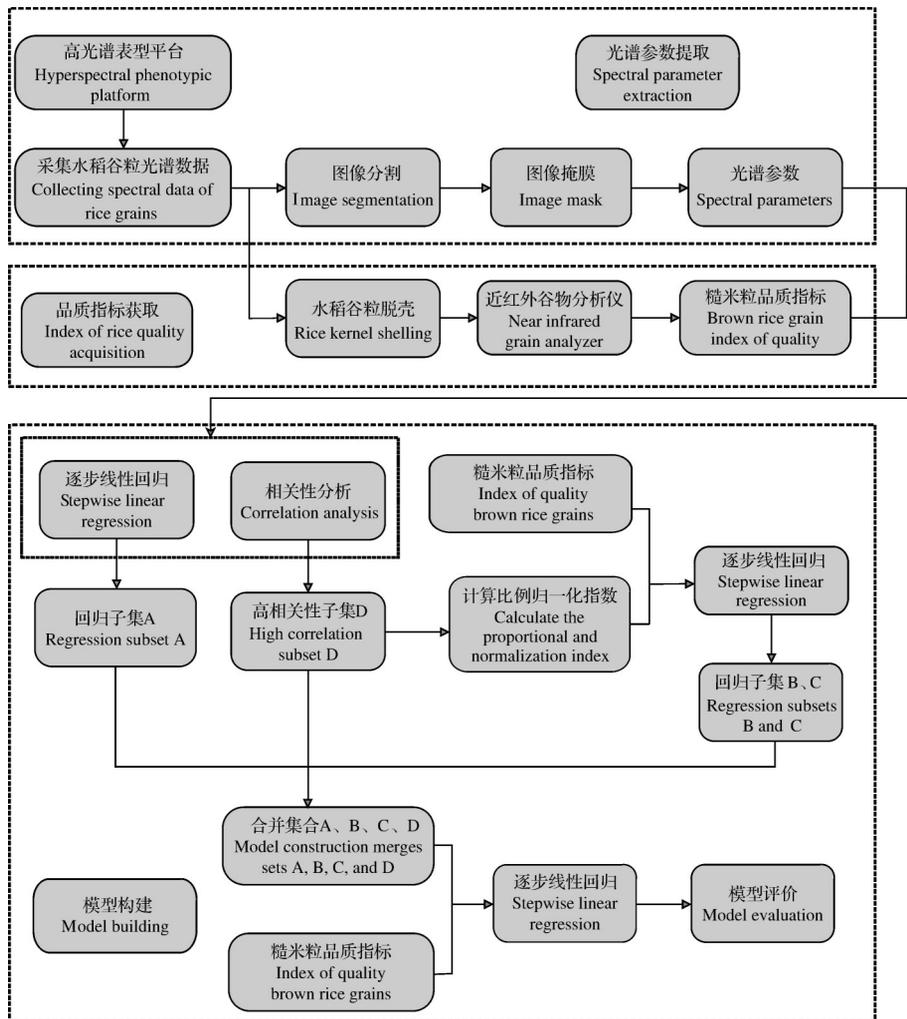


图1 技术路线

Fig.1 Technical diagram

红外谷物分析仪测得。以水稻种粒光谱参数为自变量、米粒品质为因变量构建预测模型。

本研究光谱参数提取基于Labview2015(National Instruments, USA)实现,逐步线性回归分析基于SPSS.22 (IBM, USA)软件实现,其他计算在Python3.8语言环境下通过编程实现。

1.3 试验方法

1)光谱参数提取。本研究采用近红外-高光谱成像系统,该系统硬件主要包括近红外波段高光谱相机(Headwall Photonics, USA)、卤素灯光源以及平移台。高光谱相机成像范围为900~1700 nm,光谱波段间隔为4.65 nm,总共172波段。卤素灯为系统提供稳定光源,水稻种粒放置在平移台上进行数据采集,平移台移动速度为2 mm/s。整个系统置于暗室中,工作时仅有卤素灯光源。

采集光谱数据前,需对系统进行黑白板校正。盖上镜头盖,采集1帧暗电流光谱数据。将标准白板置于高光谱相机镜头正下方,打开镜头盖,采集一帧白板光谱数据,根据公式 $R_0 = \frac{R - R_d}{R_w - R_d}$ 对采集的图像进行校正,其中, R_0 为校正后的灰度值, R 为原始光谱图像, R_d 为暗电流校正图像, R_w 为白板校正图像。试验过程中每隔1 h需对系统校正1次,以保证光谱数据的稳定性。

高光谱相机在光谱范围内共采集到172张图像,每张图像的灰度范围均不同。采用大津阈值分割算法(OTSU)进行图像分割,其核心思路是寻找一个阈值 X 把所有的像素点分成2类:一类的像素值 $\leq X$,则被置为背景区域;另一类的像素值 $> X$,则被置为前景区域。当这2类的类间方差取得最大值时,则 X 为最佳分割阈值。设背景区域的像素点数为 N_0 ,其占图像总像素的比例为 P_0 ,平均像素值为 M_0 ;前景区域所有像素点数为 N_1 ,其占图像总像素的比例为 P_1 ,平均像素值为 M_1 ;图像所有像素点数为 N ,平均像素值为 M 。则满足以下关系:

$$M = \frac{M_0 \times P_0 + M_1 \times P_1}{N} \quad (1)$$

$$M = M_0 \times P_0 + M_1 \times P_1 \quad (2)$$

$$P_0 + P_1 = 1 \quad (3)$$

则类间方差 X 满足:

$$X = P_0 \times P_1 \times (M_0 - M_1)^2 \quad (4)$$

提取全波段光谱图像,挑选1365 nm波段下的图像,经OTSU算法分割得到水稻种粒的二值图。水稻种粒二值图与全波段光谱图像掩膜得到水稻种粒的全波段高光谱图像,进而计算光谱参数,包括总反射率 T 、平均反射率 A 、平均反射率的一阶导 dA 、平均反射率的二阶导 ddA 以及 $\lg A$ 。反射光谱参数获取流程如图2所示。透射光谱参数由近红外谷物分析仪测得。

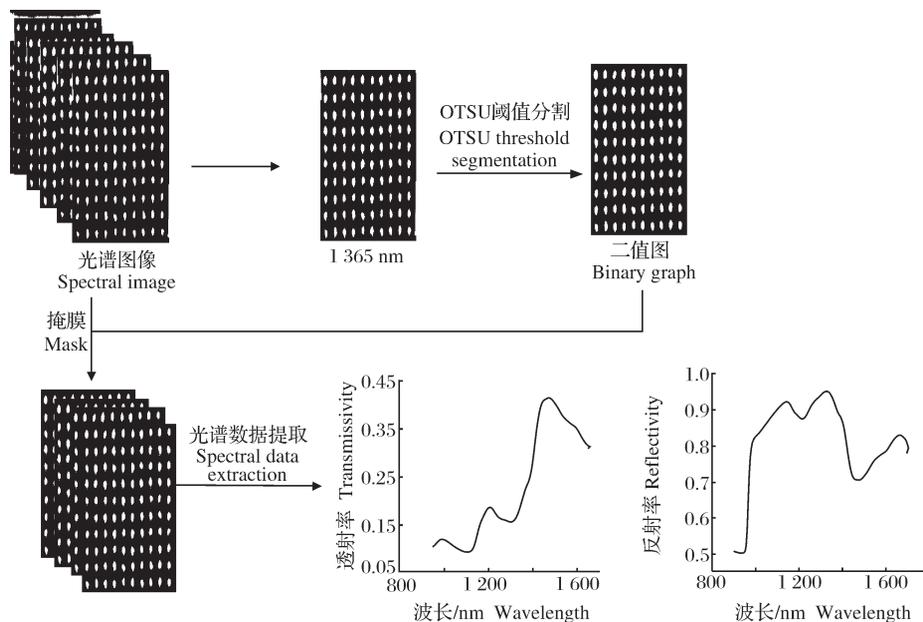


图2 光谱参数提取过程

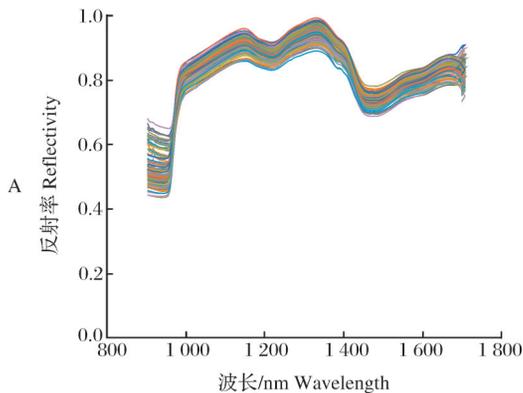
Fig.2 Spectral parameter extraction process

2)米粒品质指标测定。采用砻谷机(新丰牌JL-GJ4.5)对水稻种粒脱壳得到糙米粒,采用近红外谷

物分析仪(Perten DA7250型)测得米粒粗蛋白、水分和直链淀粉含量相对值。

3)逐步线性回归及相关性分析。采用逐步线性回归法及相关性分析法去除冗余光谱参数,筛选出对米粒品质指标有预测效果的光谱特征。将米粒粗蛋白、水分以及直链淀粉含量作为因变量分别与光谱参数进行逐步线性回归分析,得到回归子集A。计算光谱参数的比例和归一化指数,由于光谱参数众多,比例指数 $\frac{R_i}{R_j}$ 和归一化指数 $\frac{R_i - R_j}{R_i + R_j}$ 两两间计算量大,因此,在计算前,对光谱参数和因变量做相关性分析,筛选出相关性绝对值大于0.3的光谱特征,得到高相关性光谱特征集合D。对集合D的光谱特征计算比例和归一化指数,比例和归一化指数分别与因变量进行逐步线性回归分析得到回归子集B和回归子集C。整合集合A、B、C、D与因变量再次进行逐步线性回归,最终筛选得到有预测效果的特征集合。本研究中的透射光谱数据与反射光谱数据处理方式相同。

4)模型评价。模型评价用来评估模型的准确性和稳定性。本研究采用5折交叉验证对模型进行评估。5折交叉验证即将数据集随机划分为5组,其中4组用于模型训练,1组用于模型测试,每组数据分别



充当1次测试集,得到5个模型的结果,取平均值则为1次交叉验证结果。根据预测值和真实值可分别计算训练集、测试集的均方根误差(RMSE)、平均绝对百分比误差(MAPE)和决定系数(R^2)。其中, y 为真实值, \hat{y} 为预测值。

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y})^2} \quad (5)$$

$$MAPE = \frac{1}{m} \sum_{i=1}^m \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (6)$$

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\hat{y}_i - \bar{y})^2} \quad (7)$$

2 结果与分析

2.1 水稻种粒的透射、反射光谱曲线

图3展示了水稻种粒的光谱曲线。样本间反射率、透射率曲线波峰、波谷位置基本一致,但由于不同品种品质指标及种粒结构存在差异,导致在相同的波段下反射率和透射率大小存在差异。其中,透射光谱曲线的峰、谷分布在1450 nm和1100 nm附近,反射光谱曲线的峰、谷分布在1300 nm和1400 nm附近。

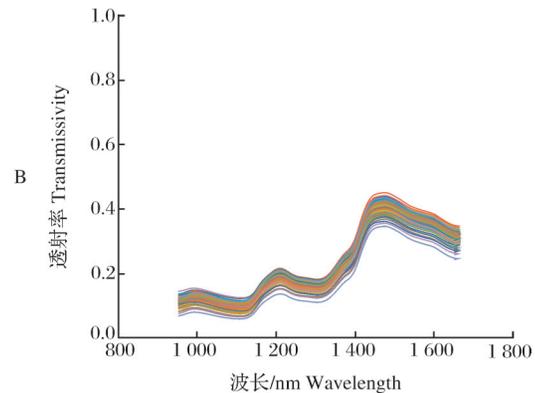


图3 水稻种粒反射(A)、透射(B)光谱特征曲线

Fig.3 Characteristic curve of reflection(A) and transmission(B) spectrum of rice seed

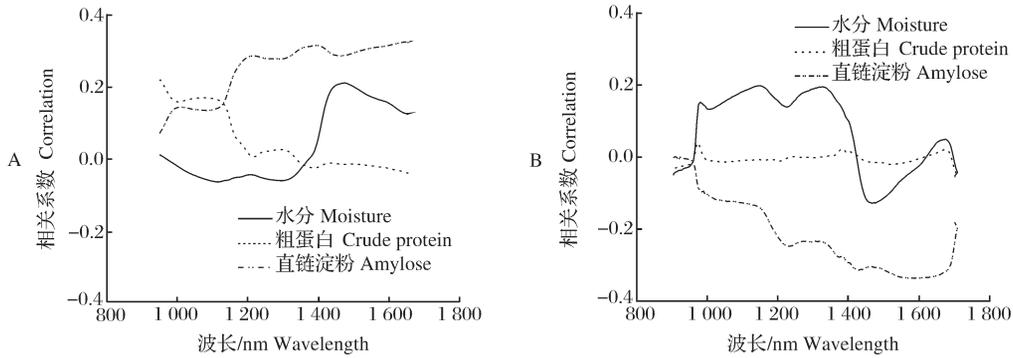
2.2 米粒品质指标与反射率、透射率的相关性

粗蛋白、直链淀粉和水分含量与反射光谱、透射光谱的相关性关系见图4,其中,水分和直链淀粉含量分别在透射、反射光谱上的相关性具有相反的趋势。水分含量在反射光谱相关性的波谷位置,恰好对应透射光谱的波峰位置1500 nm附近。粗蛋白含量与反射光谱无较大相关性,但在透射光谱波段上存在一定的相关性。

2.3 全波段透射、反射光谱回归建模

水分、粗蛋白和直链淀粉含量作为因变量分别

与全波段反射、透射光谱参数进行逐步线性回归分析,得到回归子集A。图5展示了自变量个数与 R^2 的关系,随着自变量个数的增加, R^2 有所提升,为保证模型的泛化能力,引入模型的自变量个数为样本总数的10%。其中,粗蛋白含量在反射、透射光谱建立的回归模型表现最好,基于透射光谱建立的回归模型 R^2 为0.85,基于反射光谱建立的回归模型 R^2 为0.74。基于透射光谱建立的直链淀粉含量回归模型 R^2 为0.64,基于反射光谱建立的直链淀粉含量回归模型 R^2 为0.46。基于透射光谱建立的水分含量回归



A:透射光谱 Transmission spectrum; B:反射光谱 Reflection spectrum.

图4 基于光谱的米粒品质相关性分析

Fig.4 Correlation analysis of rice grain composition based on spectrum

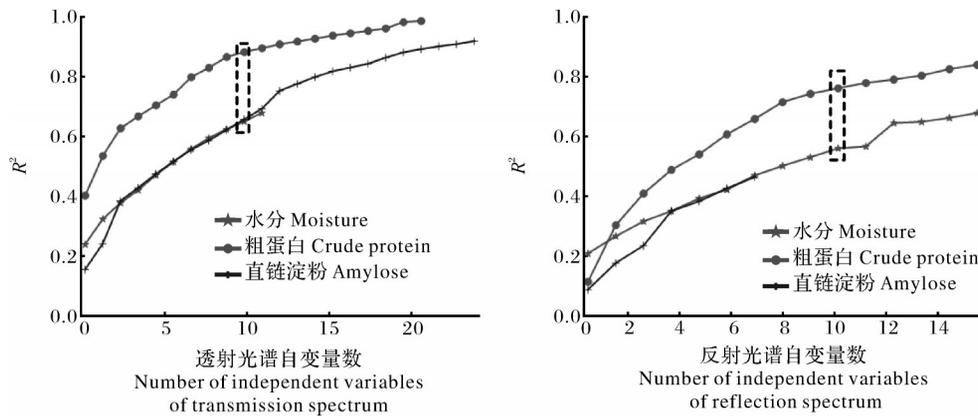


图5 基于全波段光谱的米粒品质指标逐步线性回归

Fig.5 Stepwise linear regression of rice grain composition content based on all-band spectrum

模型 R^2 为 0.63, 基于反射光谱建立的水分含量回归模型 R^2 为 0.55。结果表明, 基于透射光谱建立的回归模型较反射光谱建立的回归模型预测效果要好。

为了更好地检验模型泛化能力, 采用5折交叉验证对模型进行评估, 结果如表1所示。采用全波段透

射光谱建立的米粒品质预测模型优于采用全波段反射光谱建立的预测模型的效果。透射光谱模型评价指标 RMSE、MAPE 值均小于反射光谱, R^2 值均大于反射光谱, 说明基于透射光谱数据建立的预测模型较基于反射光谱建立的预测模型更稳定。

表1 基于全波段光谱的米粒品质指标逐步线性回归5折交叉验证结果

Table 1 Results of 5-fold cross verification of indexes of rice quality by stepwise linear regression based on all-band spectrum

方法 Methods	因变量 Dependent variable	训练集 Training sets			测试集 Testing sets		
		R^2	RMSE	MAPE/%	R^2	RMSE	MAPE/%
反射光谱 Reflection spectrum	直链淀粉含量 Content of amylose	0.431	1.525	8.00	0.400	1.655	8.80
	粗蛋白含量 Content of crude protein	0.714	0.673	10.40	0.683	0.741	11.70
	水分含量 Content of moisture	0.538	0.223	1.40	0.476	0.244	1.60
透射光谱 Transmission spectrum	直链淀粉含量 Content of amylose	0.609	0.037	2.50	0.532	0.043	2.90
	粗蛋白含量 Content of crude protein	0.848	0.041	4.90	0.794	0.050	6.00
	水分含量 Content of moisture	0.546	0.007	0.57	0.506	0.008	0.60

2.4 透射、反射特征光谱回归建模

直链淀粉、粗蛋白和水分含量作为因变量分别与透射、反射光谱数据进行相关性分析,筛选出相关性绝对值大于0.3的自变量集合 D ,计算集合 D 光谱参数间的比例和归一化指数。以比例和归一化指数集合为自变量进行逐步线性回归分析,筛选出对因变量有预测效果的回归子集 B 、 C 。整合回归子集 A 、高相关性光谱参数集合 D ,比例和归一化回归子集 B 、 C 作为最终的特征光谱集合与因变量进行逐步线性回归,结果如表2所示,基于透射光谱取特征光谱集合进行逐步线性回归能得到较好的拟合效果,其中粗蛋白的拟合 R^2 为0.89,较全波段回归提高了

表2 基于水稻种粒特征光谱集合的米粒品质逐步线性回归结果

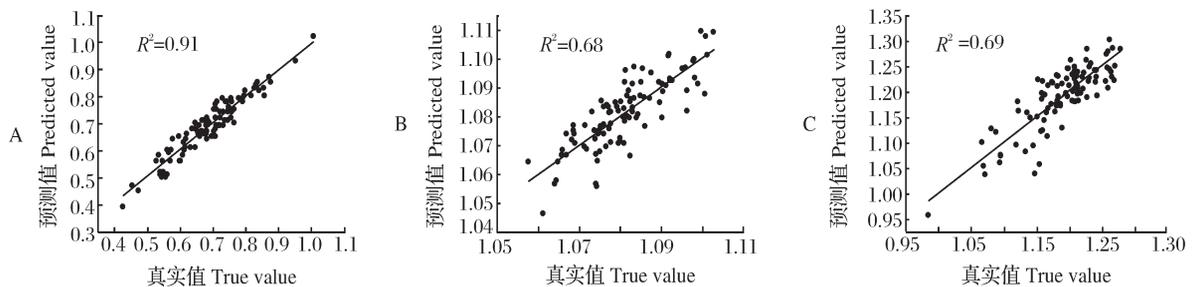
Table 2 Stepwise linear regression results of indexes of rice quality based on characteristic spectral set of rice grains

反射光谱 Reflection spectrum		透射光谱 Transmission spectrum	
因变量 Dependent variable	R^2	因变量 Dependent variable	R^2
直链淀粉含量 Amylose	0.40	直链淀粉含量 Amylose	0.67
粗蛋白含量 Crude protein	0.78	粗蛋白含量 Crude protein	0.89
水分含量 Content of moisture	0.58	水分含量 Content of moisture	0.53

0.04。直链淀粉含量的拟合 R^2 为0.67,较全波段回归提高了0.03。

2.5 结合透射-反射特征光谱回归建模

同时使用透射、反射光谱进行建模分析,以直链淀粉、粗蛋白以及水分含量分别作为因变量,本文“2.4”计算的特征光谱集合作为自变量进行逐步线性回归。图6为模型真实值和预测值的散点图,同时使用透射和反射光谱参数对米粒品质进行预测,较单一光谱取特征集合建模效果有所提升。仅使用反射光谱建立的粗蛋白含量预测模型 R^2 为0.78,仅使用透射光谱建立的粗蛋白含量预测模型 R^2 为0.89,同时使用透射、反射光谱进行建模可使模型 R^2 提升至0.91。仅使用反射光谱建立的水分含量预测模型 R^2 为0.58,仅使用透射光谱建立的水分含量预测模型 R^2 为0.53,同时使用透射、反射光谱进行建模可使 R^2 提升至0.68。仅使用反射光谱建立的直链淀粉含量预测模型 R^2 为0.40,仅使用透射光谱建立的直链淀粉含量预测模型 R^2 为0.67,同时使用透射、反射光谱进行建模可使模型 R^2 升至0.69。模型的真实值和预测值方差分析结果(表3)显示, P 值大于0.01,说明模型预测值和真实值差异不显著,模型具有一定预测效果。为进一步验证模型稳定性和健壮性,对模型进行了五折交叉验证,结果(表4)显示,训练集、测试集的RMSE、MAPE均在统计学接受范围内,模型稳定性、准确性较高。



A:粗蛋白 Crude protein; B:水分 Moisture; C:直链淀粉 Amylose.

图6 米粒品质真实值与预测值散点图

Fig.6 Scatter plot of true and predicted of indexes of rice quality

3 讨论

本研究基于水稻种粒透射光谱、反射光谱以及结合透射-反射光谱分别对米粒品质指标进行建模预测。其中直链淀粉含量的 R^2 为0.67,与路辉等^[27]的研究结果一致。但水分含量的预测模型 R^2 稍低,可能是由于水稻种粒光谱采集和米粒品质指标测定的时间差导致。同时使用水稻种粒透射、反射光谱

数据对粗蛋白、直链淀粉含量进行建模预测,其 R^2 高于唐延林等^[28]利用稻谷粉末进行预测的 R^2 ,水稻种粒较粉末达到了无损效果,也便于光谱数据的采集,同时使用透射、反射光谱数据包含的信息更丰富,能取得更好的预测效果。戴国新等^[13]采集水稻精米的反射光谱数据基于逐步线性回归对精米蛋白质含量进行预测, R^2 为0.837,本研究基于水稻种粒的透射-

表3 米粒品质真实值与预测值方差分析结果

Table 3 Results of variance analysis of real and predicted indexes of rice quality

品质指标 Indexes of rice quality		自由度 Degree of freedom	平方和 Quadratic sum	均方 Mean square	F值 F-value	P
直链淀粉含量 Amylose	组间 Groups	1	9.3421E-11	9.3421E-11	2.6896E-8	0.999 869
	组内 Interclass	198	0.687 743	0.003 473		
	总计 Total	199	0.687 743			
水分含量 Content of moisture	组间 Groups	1	1.5014E-12	1.5014E-12	1.2114E-8	0.999 912
	组内 Interclass	198	0.024 539	0.000 124		
	总计 Total	199	0.024 539			
粗蛋白含量 Crude protein	组间 Groups	1	1.9354E-10	1.9354E-10	1.662E-8	0.999 897
	组内 Interclass	198	2.305 620	0.011 645		
	总计 Total	199	2.305 620			

表4 基于特征光谱集合的米粒品质回归模型5折交叉验证结果

Table 4 Five-fold cross-validation results of regression model of indexes of rice quality based on characteristic spectral set

因变量 Dependent variable	训练集 Training sets			测试集 Testing sets		
	R ²	RMSE	MAPE/%	R ²	RMSE	MAPE/%
直链淀粉含量 Amylose	0.670	0.034	2.36	0.623	0.039	2.71
粗蛋白含量 Crude protein	0.904	0.036	4.47	0.868	0.040	5.03
水分含量 Content of moisture	0.650	0.007	0.51	0.595	0.007	0.59

反射光谱信息对米粒品质指标进行建模预测,粗蛋白含量的预测模型 R^2 为0.91,优于戴国新等^[13]仅使用反射光谱建模的效果,采集水稻种粒的光谱数据较采集糙米或者精米的光谱数据更便捷。同时使用透射和反射光谱进行预测较单一光谱所包含的信息更丰富,故能取得更好的预测效果。使用全波段光谱建模效果稍优于使用特征波段建模效果,但是特征波段建模不仅精简了模型,更在于使用相对较少的特征取得较好的预测效果。前人对米粒的性状研究多是基于米粒的光谱信息,本研究提出了1种基于水稻种粒的光谱信息预测米粒品质的建模方法,该方法可为今后其他作物的更多性状含量检测提供思路。本研究样本量有限,今后可开展群体研究,进一步验证所建模型的稳定性。

本研究使用全波段反射光谱进行建模,水分、粗蛋白和直链淀粉含量的 R^2 分别为0.55、0.74、0.46;使用全波段透射光谱进行建模,水分、粗蛋白、直链淀粉含量的 R^2 分别为0.63、0.85、0.64,结果表明仅使用透射光谱进行建模预测效果优于仅使用反射光谱的预测效果。使用逐步线性回归方法对反射光谱取特征波段回归建模,水分、粗蛋白、直链淀粉含量的 R^2

分别为0.58、0.78、0.40。对透射光谱取特征波段集合回归建模,水分、粗蛋白、直链淀粉含量的 R^2 分别为0.53、0.89、0.67。结合反射光谱和透射光谱进行特征波段建模可使水分含量预测模型的 R^2 从0.53提高至0.68,可使粗蛋白含量预测模型 R^2 从0.74提高至0.91,可使直链淀粉含量预测模型决定系数从0.40提高至0.69,5折交叉验证结果显示模型具有稳定性,预测模型具有统计学意义。通过建立水稻种粒高光谱参数预测米粒品质指标的模型可为其他作物种子品质性状检测提供参考。

参考文献 References

- [1] 余四斌,熊银,肖景华,等.杂交稻与绿色超级稻[J].科学通报,2016,61(35):3797-3803.YU S B, XIONG Y, XIAO J H, et al. Hybrid rice and green super rice [J]. Chinese science bulletin, 2016, 61(35):3797-3803 (in Chinese with English abstract).
- [2] 卢林,孙成效,朱智伟,等.我国稻米品质标准及检测技术创新概述[J].中国稻米,2022,28(1):1-6.LU L, SUN C X, ZHU Z W, et al. Overview of rice quality standard and innovative detection technology in China [J]. China rice, 2022, 28(1): 1-6 (in Chinese with English abstract).
- [3] 夏朵,周浩,何予卿.稻米品质的遗传研究及分子育种进展[J].

- 华中农业大学学报, 2022, 41(1): 48-61. XIA D, ZHOU H, HE Y Q. Progress of genetics and molecular breeding of rice quality [J]. Journal of Huazhong Agricultural University, 2022, 41(1): 48-61 (in Chinese with English abstract).
- [4] 孙俊, 唐凯, 毛罕平, 等. 基于 MEA-BP 神经网络的大米水分含量高光谱技术检测 [J]. 食品科学, 2017, 38(10): 272-276. SUN J, TANG K, MAO H P, et al. Hyperspectral detection of moisture content in rice based on MEA-BP neural network [J]. Food science, 2017, 38(10): 272-276 (in Chinese with English abstract).
- [5] 赵飞, 刘建, 曹高巍, 等. 水稻品质育种与优质化栽培问题的探讨 [J]. 杂交水稻, 2022, 37(2): 1-6. ZHAO F, LIU J, CAO G Y, et al. Discussion on grain quality in breeding and cultivation of rice [J]. Hybrid rice, 2022, 37(2): 1-6 (in Chinese with English abstract).
- [6] 彭波, 孙艳芳, 韩秋, 等. 水稻种子蛋白质含量检测方法的比较分析 [J]. 江苏农业科学, 2018, 46(4): 22-28. PENG B, SUN Y F, HAN Q, et al. Comparative analysis of detection methods of protein content in rice seeds [J]. Jiangsu agricultural sciences, 2018, 46(4): 22-28 (in Chinese).
- [7] 王仪春, 张小明, 石春海. 稻米直链淀粉含量测定方法的探讨 [J]. 中国农学通报, 2001, 17(5): 30-32, 51. WANG Y C, ZHANG X M, SHI C H. Studies on the analyzing method of amylose contents in rice [J]. Chinese agricultural science bulletin, 2001, 17(5): 30-32, 51 (in Chinese with English abstract).
- [8] 张小明, 石春海, 鲍根良, 等. 稻米直链淀粉含量测定方法研究 [J]. 植物遗传资源科学, 2002(1): 31-35. ZHANG X M, SHI C H, BAO G L, et al. Studies on the analyzing method of amylose contents in rice grain [J]. Journal of plant genetic resources, 2002(1): 31-35 (in Chinese with English abstract).
- [9] 盛宏达, 阎金生, 张克朋. 用烘干法测定大米水分时须注意的几个问题 [J]. 中国稻米, 2005, 11(1): 40. SHENG H D, YAN J S, ZHANG K P. Several problems needing attention in determining rice moisture by drying method [J]. China rice, 2005, 11(1): 40 (in Chinese).
- [10] 王冬, 王坤, 吴静珠, 等. 基于光谱及成像技术的种子品质无损快速检测研究进展 [J]. 光谱学与光谱分析, 2021, 41(1): 52-59. WANG D, WANG K, WU J Z, et al. Progress in research on rapid and non-destructive detection of seed quality based on spectroscopy and imaging technology [J]. Spectroscopy and spectral analysis, 2021, 41(1): 52-59 (in Chinese with English abstract).
- [11] ZHU M, HUANG D, HU X J, et al. Application of hyperspectral technology in detection of agricultural products and food: a review [J]. Food science & nutrition, 2020, 8(10): 5206-5214.
- [12] MA C Y, REN Z S, ZHANG Z H, et al. Development of simplified models for nondestructive testing of rice (with husk) protein content using hyperspectral imaging technology [J/OL]. Vibrational spectroscopy, 2021, 114: 103230 [2022-10-06]. <https://doi.org/10.1016/j.vibspec.2021.103230>.
- [13] 戴国新, 陈国兴, 杨万能, 等. 基于高光谱的水稻精米品质参数测量技术研究 [J]. 农业大数据学报, 2019, 1(2): 51-63. DAI G X, CHEN G X, YANG W N, et al. Measurement technology of quality parameters of rice grain based on hyperspectral imaging on the visible-near infrared [J]. Journal of agricultural big data, 2019, 1(2): 51-63 (in Chinese with English abstract).
- [14] SUN D W, CEN H Y, WENG H Y, et al. Using hyperspectral analysis as a potential high throughput phenotyping tool in GWAS for protein content of rice quality [J/OL]. Plant methods, 2019, 15: 54 [2022-10-06]. <https://doi.org/10.1186/s13007-019-0432-x>.
- [15] 张晶, 郭军, 张美莉, 等. 基于近红外建立荞麦营养成分快速检测模型 [J]. 中国粮油学报, 2020, 35(6): 151-158. ZHANG J, GUO J, ZHANG M L, et al. Establishment of rapid detection model of buckwheat nutritional components based on near infrared spectroscopy [J]. Journal of the Chinese cereals and oils association, 2020, 35(6): 151-158 (in Chinese with English abstract).
- [16] 梁秀英, 李小昱, 杨万能. 奇异数据筛选法在玉米籽粒蛋白质近红外光谱检测中的应用 [J]. 激光生物学报, 2015, 24(1): 38-45. LIANG X Y, LI X Y, YANG W N. Outlier detection for measurement of protein content in maize kernels based on near-infrared reflectance spectroscopy [J]. Acta laser biology sinica, 2015, 24(1): 38-45 (in Chinese with English abstract).
- [17] HEMAN A, HSIEH C L. Measurement of moisture content for rough rice by visible and near-infrared (NIR) spectroscopy [J]. Engineering in agriculture, environment and food, 2016, 9(3): 280-290.
- [18] 于宏威, 王强, 石爱民, 等. 高光谱成像技术结合化学计量学可视化花生中蛋白质含量分布 [J]. 光谱学与光谱分析, 2017, 37(3): 853-858. YU H W, WANG Q, SHI A M, et al. Visualization of protein in peanut using hyperspectral image with chemometrics [J]. Spectroscopy and spectral analysis, 2017, 37(3): 853-858 (in Chinese with English abstract).
- [19] 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 with English abstract).
- [20] 孟庆琰, 何建国, 刘贵珊, 等. 基于近红外光谱技术的马铃薯全粉蛋白质无损检测 [J]. 食品科技, 2015, 40(3): 287-291. MENG Q Y, HE J G, LIU G S, et al. Potato granules protein nondestructive testing based on NIR spectroscopy analysis technology [J]. Food science and technology, 2015, 40(3): 287-291 (in Chinese with English abstract).
- [21] 曹志敏, 张志肖, 王彦, 等. 近红外光谱法非破坏性测定绿豆籽粒粗蛋白含量的研究 [J]. 河北农业科学, 2016, 20(5): 104-108. CAO Z M, ZHANG Z X, WANG Y, et al. Non-destructive test of crude protein in mungbean grain by near infrared reflectance spectroscopy [J]. Journal of Hebei agricultural sciences, 2016, 20(5): 104-108 (in Chinese with English abstract).
- [22] 曹晓宁, 田翔, 赵小娟, 等. 基于近红外光谱法快速检测藜麦淀粉含量 [J]. 江苏农业科学, 2017, 45(4): 147-149. CAO X N, TIAN X, ZHAO X J, et al. Rapid detection of starch content in quinoa based on near infrared spectroscopy [J]. Jiangsu agricultural sciences, 2017, 45(4): 147-149 (in Chinese).

- [23] 刘红芸, 吴雪梅, 李德仑, 等. 基于高光谱技术的采摘期烟叶水分含量研究[J]. 中国农机化学报, 2021, 42(9): 157-163. LIU H Y, WU X M, LI D L, et al. Study on the moisture content of tobacco leaves during the picking period based on hyperspectral technology [J]. Journal of Chinese agricultural mechanization, 2021, 42(9): 157-163 (in Chinese with English abstract).
- [24] 许建东, 张淑娟, 郑小南, 等. 高光谱技术结合变量选择方法的甘薯冻害检测研究[J]. 食品与发酵工业, 2021, 47(8): 197-203. XU J D, ZHANG S J, ZHENG X N, et al. Study on the detection of sweet potato freezing damage based on hyperspectral technology and variable selection method [J]. Food and fermentation industries, 2021, 47(8): 197-203 (in Chinese with English abstract).
- [25] TYBURCZY C, MOSSOBA M M, FARDIN-KIA A R, et al. Evaluation of low transfat edible oils by attenuated total reflection-Fourier transform infrared spectroscopy and gas chromatography: a comparison of analytical approaches [J]. Analytical and bioanalytical chemistry, 2012, 404(3): 809-819.
- [26] 罗曦, 吴方喜, 谢鸿光, 等. 近红外光谱的水稻抗性淀粉含量测定研究[J]. 光谱学与光谱分析, 2016, 36(3): 697-701. LUO X, WU F X, XIE H G, et al. Research on resistant starch content of rice grain based on NIR spectroscopy model [J]. Spectroscopy and spectral analysis, 2016, 36(3): 697-701 (in Chinese with English abstract).
- [27] 路辉, 彭彬倩, 冯晓宇, 等. 大米直链淀粉、蛋白质、脂肪、水分含量的近红外光谱检测模型优化[J]. 中国稻米, 2020, 26(6): 55-59, 63. LU H, PENG B Q, FENG X Y, et al. Model optimization for determination of amylose, protein, fat and moisture content in rice by near-infrared spectroscopy [J]. China rice, 2020, 26(6): 55-59, 63 (in Chinese with English abstract).
- [28] 唐延林, 黄敬峰, 王人潮. 利用高光谱法估测稻穗稻谷的粗蛋白质和粗淀粉含量[J]. 中国农业科学, 2004, 37(9): 1282-1287. TANG Y L, HUANG J F, WANG R C. Study on estimating the contents of crude protein and crude starch in rice panicle and paddy by hyperspectra [J]. Scientia agricultura sinica, 2004, 37(9): 1282-1287 (in Chinese with English abstract).

A method for predicting rice quality based on hyperspectral analysis of rice seed

ZHAO Shuang¹, SONG Jingyan¹, CHEN Guoxing¹, SONG Peng¹, FENG Hui², YANG Wanneng¹

1. National Key Laboratory of Corp Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China;

2. College of Informatics, Hubei Key Laboratory of Agricultural Bioinformatics, Huazhong Agricultural University, Wuhan 430070, China

Abstract The content of amylose, crude protein and water is an important index to measure the quality of rice grain. The transmittance and reflectance spectral data of rice grains from 100 rice core germplasm resources were collected using a near-infrared hyperspectral camera, and the spectral parameters were extracted to study the method for the non-destructive testing of quality traits in rice seed. Index of rice quality components was measured using a near infrared grain analyzer after the rice kernel was shelled. A model for predicting index of rice quality was established using the spectral parameters of rice grains as independent variables and index of rice quality as dependent variables. The results showed that the modeling effect of transmission spectrum was better than that of reflection spectrum when a single spectral model was used. Combined with characteristic spectral sets of transmission and reflection, the R^2 of the model for predicting crude protein, amylose and water was increased from 0.74 to 0.91, from 0.40 to 0.69, and from 0.53 to 0.68, respectively. It is indicated that the modeling effect can be improved by using both spectrum of transmission and reflection, and index of rice quality can be predicted nondestructively by using the spectral parameters of rice grains.

Keywords rice; rice quality; nondestructive examination; hyperspectral; prediction model

(责任编辑: 张志钰)