

doi:10.3969/j.issn.1002-2481.2024.03.01

施用羊粪下小米粗脂肪、粗蛋白质含量高光谱检测

王国梁, 张艾英, 王丽霞, 赵培悦, 刘鑫, 成锴, 郭二虎

(山西农业大学 谷子研究所/省部共建有机旱作农业国家重点实验室(筹), 山西 长治 046011)

摘要:运用先进技术实现有机肥高效利用,因地制宜发展区域农业是有机旱作农业发展的重要手段。羊粪施用量的选择与谷子种植用养结合能够有效提高小米品质质量,区别于传统检测手段,高光谱成像对待测样本可以做到检测快速、无损,因而在实现小米粗脂肪、粗蛋白质快速检测中具有巨大潜力。为了为谷子栽培措施的改进及利用高光谱成像对小米营养成分快速检测提供参考和理论支持,试验采集358份小米样本高光谱图像数据并利用传统检测手段进行粗脂肪、粗蛋白质含量测定,同时运用化学计量学方法进行2种成分含量预测。结果表明,连续投影算法结合迭代保留信息变量算法建立的偏最小二乘回归模型在2种成分含量检测上精度最高,其中模型预测粗蛋白质含量相关系数、均方根误差、相对分析误差分别为0.88、0.59、1.99,利用该回归模型进行可视化表达可直观看出不同羊粪施用量下粗脂肪、粗蛋白质的积累规律,并得出每公顷施用90 m³肥料为最佳施用量。

关键词:小米粗脂肪;粗蛋白质;高光谱;化学计量学方法;可视化

中图分类号:O657.3 文献标识码:A 文章编号:1002-2481(2024)03-0001-08

Hyperspectral Detection of Crude Fat and Protein Content of Millet under Sheep Manure Application

WANG Guoliang, ZHANG Aiyong, WANG Lixia, ZHAO Peiyue,

LIU Xin, CHENG Kai, GUO Erhu

(Institute of Millet, Shanxi Agricultural University/State Key Laboratory of Sustainable Dryland Agriculture(in preparation), Changzhi 046011, China)

Abstract: Efficient utilization of organic fertilizer through advanced technology and development of regional agriculture according to local conditions are a crucial method for the development of organic dry farming. Combining sheep manure with millet cultivation can significantly enhance millet quality. Unlike traditional detection methods, hyperspectral imaging offers a rapid, non-destructive means of analyzing samples, therefore, it has great potential in achieving rapid detection of millet crude fat and protein. In order to provide reference and theoretical support for improvement of millet cultivation practices and rapid assessment of nutrient content in millet through hyperspectral imaging, in this study, hyperspectral image data of 358 millet samples were collected, traditional detection methods were used to determine the content of crude fat and crude protein. Simultaneously, chemometrics methods were utilized to predict the content of the two components. The results indicated that the partial least squares regression(PLSR) model established by employing the successive projections algorithm(SPA) and iteratively retaining informative variables(IRIV) had the highest accuracy. The correlation coefficient(R^2) predicted by the model for crude protein content was 0.88, with a root mean squared error(RMSE) of 0.59 and a relative percent deviation(RPD) of 1.99. The visualization representations by the regression model vividly illustrated the accumulation patterns of crude fat and protein at various application rates of sheep manure, and it was determined that applying 90 m³ of fertilizer per hectare was the optimal application amount.

Key words: crude fat and protein of millet; hyperspectral; chemometrics methods; visualization

小米中粗脂肪、粗蛋白质含量分别占比在5% 于小麦和大米,且涵盖氨基酸种类齐全,比例适宜,和10%左右^[1-2]。研究表明,小米中蛋白质含量高 熬制小米粥后,上浮“米油”对特殊人群极具一定药

收稿日期:2023-07-03

基金项目:山西农业大学省部共建有机旱作农业国家重点实验室自主研发项目(202105D121008-2-5);山西省现代农业产业技术体系谷子体系(2023CYJSTX04-04);杂粮种质资源创新与分子育种国家实验室项目(202204010910001-13);国家现代农业产业技术体系建设专项(CARS-06-14.5-A21)

作者简介:王国梁(1985-),男,山西大同人,助理研究员,博士,主要从事农产品品质检测研究工作。

通信作者:郭二虎(1965-),男,山西长治人,研究员,主要从事谷子新品种选育及栽培研究工作。

用保健功能,小米以其丰富的营养价值现已被大众公认为优质谷物,对于小米粗脂肪和粗蛋白质的快速检测也变得日益重要。

高光谱检测具有无损、快速、无接触及化学试剂零使用等技术特点,目前,运用光谱技术针对农作物营养成分检测的研究也渐渐被农业科研人员所关注^[3-5]。田翔等^[6]对小米淀粉、直链淀粉、蛋白质和脂肪应用近红外光谱法进行快速检测,结果表明,采用化学法和近红外法检测结果无显著差异;王浩等^[7]运用近红外结合 PLSR 模型实现小米硒含量的快速预测。此外,高光谱数据具有变量维度高的特点,为了降低变量维度、提升预测模型预测精度,一般采用特征波段提取算法对目标变量数据组进行特征波段选择。其中,变量组合集群分析(Variable Combination Population Analysis, VCPA)算法能够将所有存在相互影响的原始变量自由组合,并通过预设终端输出变量波段数阈值,选取极具代表性的变量波段,赵环等^[8]运用该算法预测小麦蛋白质含量并取得较理想预测效果;SPA 算法通过选择初始向量,计算在未选变量上的投影,选出最佳波段组合,李颖^[9]运用 SPA 算法对木材可见-近红外光谱数据进行特征波段提取并建立 PLS 模型,在木材分类过程中取得较为理想效果;而 IRIV 算法可以在剔除冗余波段的同时将变量强、弱信息波段尽可能多地保留,于雷等^[10]运用 IRIV 算法优选了大豆叶片高光谱波长并对叶绿素含量进行估测,为高光谱技术估测叶片生理参数提供新思路。目前,特征波段提取算法联用在特征波段选择上应用渐广,余海东^[11]提出近红外光谱波长选择采用 3 步联用策略,提高模型预测精度同时,既突出各算法优势,也克服了算法局限性。因此,通过运用高光谱成像结合化学计量学相关知识实现小米粗脂肪、粗蛋白质快速检测具有重要的现实意义。

本试验主要对不同羊粪施用量处理下的小米高光谱数据,运用传统化学方法测定粗脂肪和粗蛋白质含量,并结合化学计量方法建立快速检测小米粗脂肪和粗蛋白质含量的预测模型,通过特征变量提取联用建立的预测模型进行小米粗脂肪和粗蛋白质含量积累规律的可视化表达,旨在探寻快速、无毒、零接触、低成本和可视化预测小米粗脂肪和粗蛋白质含量的方法。通过研究高光谱检测技术结合化学计量学在小米粗脂肪和粗蛋白质检测方面应用的可行性,为进一步在谷子农艺性状选择、育种栽培方式优化、产品质量分类定级以及提升小

米终端商品价值提供参考依据。

1 材料和方法

1.1 试验地概况

试验地位于山西农业大学谷子研究所试验基地,该区域属暖温带大陆性气候,年平均气温 9.3℃,无霜期 150 d。年平均降雨量 580 mm。土壤属于红壤,pH 值为 7.4。前茬作物是玉米。基本肥力为:土壤含水量 15.9%,有效氮含量 58.67 mg/kg,有效磷含量 15.54 mg/kg,有效钾含量 95.20 mg/kg,有机质含量 10.75 g/kg,土壤肥力处于中等水平。

1.2 试验材料

供试长生 13(品种登记号:晋审谷(认)2015008)由山西农业大学谷子研究所提供。羊粪购买自农户。羊粪有机质含量为 42.7%,氮含量为 1.86%,磷含量为 1.47%,钾含量为 2.02%。

1.3 试剂与仪器设备

石油醚、浓硫酸、脱脂滤纸、过氧化氢、溴甲酚绿、甲基红、无水乙醇、盐酸(国药集团试剂有限公司)。

LTJM2099 散热型小米磨谷机(郑州中谷机械设备有限公司),WF-20B 型粉碎机(南京科益机械设备有限公司),HWS24 型电热恒温水浴锅(上海恒科技有限公司),8400 型全自动定氮仪(丹麦 FOSS 公司),Starter Kit 型高光谱成像仪(近红外光程 900~1 700 nm,数码镜头焦距 25 mm,光圈 F/1.3,光源型号 PHILIPS 6995 P,工作色温 3 200 K,额定电压/功率为 230 V/1 000 W,美国 Headwall Photonics 公司),铝盒(中山市天美五金制品有限公司)。

1.4 试验方法

羊粪施用量设置 4 个处理:S1. 60 m³/hm²,S2. 75 m³/hm²,S3. 90 m³/hm²,S4. 105 m³/hm²,同时设 S5 为空白对照。羊粪都是底施,试验一重复按 1—5 正序依次排列,二重复按 5—1 倒序排列。取样时,采用棋盘采样法,每小区采样 180 个点,每点面积 2 m²,相邻 5 点混装,编号记录,经晾晒、砻碾、精选等工序后,共采集待测小米样本 358 份。

1.4.1 小米高光谱数据采集 小米高光谱数据采集过程中,被测样装入直径 5 cm、深 1 cm 的器皿内,保证被测样品表面平整、紧实,载物台平移速度 16 mm/s,相机曝光时间 0.9 ms,样品表面距离镜头 25 mm,每份被测样品采集光谱数据 3 次,数据按顺序编号、保存。数据采集开始前进行黑白校正,后根据式(1)运算得到样本相对光强 X,每采集 3 张高

光谱图像,重复进行上述校正过程^[3]。采集设备如图1所示。

$$X = \frac{I - B}{W - B} \quad (1)$$

式中, B 为暗背景光强(cd); W 为白背景光强(cd); I 为被测样品采集光强(cd)。

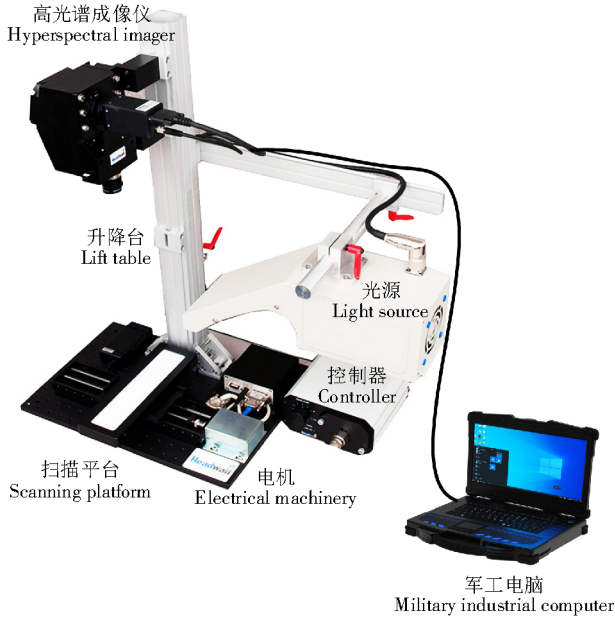


图1 高光谱成像仪结构
Fig.1 Structure of hyperspectral imager

1.4.2 粗脂肪含量测定 采用索氏提取法对小米粗脂肪含量(T)进行测定^[12]。

$$T = \frac{W_2 - W_3}{W_2 - W_1} \times 100 \quad (2)$$

式中, W_1 为脱脂滤纸质量(g); W_2 为脱脂滤纸与样本总质量(g); W_3 为样本完成脱脂过程与脱脂滤纸总质量(g)。

1.4.3 粗蛋白质含量测定 脱脂后样本采用凯氏定氮法对小米粗蛋白质含量(P)进行测定^[13]。

$$P = \frac{(A - B) \times C \times 14 \times V_1}{M \times V_2} \times 6.25 \quad (3)$$

式中, A 为样品滴定时消耗盐酸量(mL), B 为对照滴定时消耗盐酸量(mL), C 为标准盐酸摩尔浓度(mol/L),14为盐酸标准滴定溶液相当的氮质量(mg), V_1 为消化液总量(mL), M 为样品质量(mg), V_2 为蒸馏时消化液用量(mL),6.25为总粗蛋白质含量转换系数。

1.5 高光谱数据分析

对小米进行高光谱数据处理,采用 Kennard-Stone 算法将待测样本光谱数据按照 2:1 分为训练集和预测集^[14],采用特征波段提取联用算法进行光

谱数据特征波段提取并建立 PLSR 模型,进行小米粗脂肪、粗蛋白质含量的预测,为更直观描述不同羊粪施用量下粗脂肪、粗蛋白质的积累规律,利用最佳回归方程对这 2 种成分进行可视化表达。

1.5.1 高光谱数据分析软件 样本高光谱数据采集及可视化处理使用软件 ENVI 5.1,数据分析使用 MATLAB 2020a。

1.5.2 特征变量提取 特征波段提取是从高光谱高维数据中提取、挖掘隐匿信息,消除无关信息干扰并提高回归模型预测能力,提升运算速度的方法。本研究中,提出 VCPA-IRIV、SPA-IRIV 特征变量提取联用算法,对初次提取特征变量进行“精选”,最大程度上剔除冗余波段,提高模型精度,其中,VCPA 变量选择基于指数衰减函数(Uninformative variable elimination strategy of index decreasing function, EDF)、二进制矩阵采样法(Binary matrix sampling, BMS)和模型总体分析(Model population analysis, MPA)3 个重要步骤,EDF 使变量空间不断收缩,剩余变量占比(r_i)计算如式(4)。

$$r_i = e^{-\theta i} \quad (4)$$

式中, θ 为控制 EDF 曲线的参数,与下降曲线速度呈正相关,当 i 为 0 时, $r_0=1$;经过 N 次迭代后,剩余变量数为预设阈值 ω ,即最终特征波段数($r_N = \omega/p$, p 为变量数),利用式(5)求得 θ 值。

$$\theta = \frac{\ln\left(\frac{p}{\omega}\right)}{N} \quad (5)$$

每次执行 EDF 时,BMS 和 MPA 也被执行,BMS 生成与原数据矩阵等尺寸的二进制矩阵,变量被选与否根据 BMS 中“1”“0”分布决定,“1”为选择,“0”为不选,MPA 利用偏最小二乘法计算子集交叉验证均方根误差,并保留该值取最小的变量子集,重复此运算过程,即得到所需波段组合,该算法能够随执行进度剔除低贡献变量,变量空间随之变小,提高高贡献变量组合出现概率,较为有效避免无效信息和干扰信息出现^[15-16];SPA 是一种前向特征变量选择算法,利用向量投影分析,将变量投影于其他变量上,以投影向量最大波长为选择波长,通过预设变量选择数,最终获得含最少冗余信息和最小共性的变量组合^[17-18];IRIV 依据模型集群分析思想,并充分考量变量间联合效应,将变量随机组合生成二进制矩阵,该矩阵行为变量组合、列为变量数,基于各行分别建立 PLS 模型,通过交叉验证均方根误差评估这些组合模型效果。最终,将贡献

分为强信息变量、弱信息变量、无信息变量以及干扰信息变量 4 种,保留变量由 *DMEAN*(Difference of mean values)及 *P* 值决定,前者为波长变量计算包含和不包含时的均方根误差均值,后者为曼-惠特尼 *U* 检验 *P* 值。执行算法多次迭代后排除无效及干扰变量,并经逆向消除从强、弱信息变量中提取获得最佳特征波段^[19]。

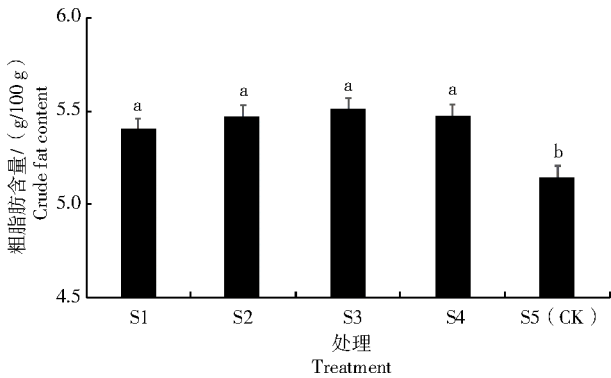
1.5.3 预测模型建立及评价 预测模型采用 PLSR 模型,该模型是反演样品理化特性中最为常用的一种多元统计数据分析法^[20],运用该算法建模时,能兼顾样本光谱信息及其理化特性,探究二者线性关系。利用 *R*、*RMSE*、*RPD* 指标评价预测模型,依据文献[21]的 *RPD* 值将预测模型预测精度分为 A、B、C 等 3 类,其中, A 类的 *RPD* > 2,代表模型预测精度较高; B 类的 $1.4 \leq RPD \leq 2$,代表模型具

有一般预测精度,仅适用于粗定量估计; C 类的 *RPD* < 1.4,代表模型精度较差,此模型不适用于该数据集的预测。

2 结果与分析

2.1 粗脂肪、粗蛋白质含量差异分析

羊粪施用量的不同会反映在样本粗脂肪、粗蛋白质含量差异上。从图 2 可以看出, S3 处理 (90 m³/hm² 羊粪) 的小米粗脂肪、粗蛋白质含量均达到最高,平均粗脂肪和粗蛋白质含量分别为 5.52、11.05 g/100 g。S1~S4 处理间的粗脂肪含量差异不显著,但均显著高于 S5(CK) (*P* < 0.05); 而 S2 处理与 S3 处理间的粗蛋白含量差异不显著,但与其他处理之间均差异显著 (*P* < 0.05)。

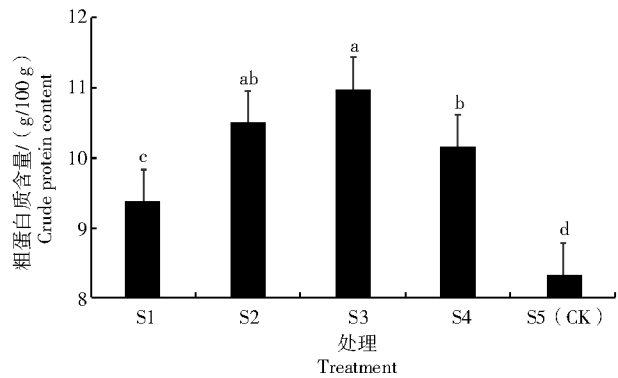


不同小写字母表示不同处理在 0.05 水平上差异显著

Different lowercase letters indicated significant differences in different treatment at the 0.05 level

图 2 不同羊粪施用量下小米的粗脂肪、粗蛋白质含量

Fig.2 The content of crude fat and protein of millet with the different application rates of sheep manure



2.2 不同羊粪施用量下小米光谱特性

图 3 为 5 个处理下小米样本采集绘制而成的平均光谱曲线,图中首尾两端反射噪声大,数据稳定性差,选取各样本 950~1 650 nm 光谱数据为基础光谱数据,在该光谱区间内待测物光谱数据含有丰富的化学官能团信息。平均光谱曲线出现波谷处代表该波段区域内样本分子对辐射光具有吸附作用,产生了化学键的振动。

由图 3 可知,5 条曲线变化趋势基本一致,羊粪施用量处理不同平均光谱曲线在各个波段下呈现不同反射率数值, S3 处理相较于其他处理反射率普遍偏低,反映出 S3 处理在籽粒营养物质的积累上相较于其他 4 个处理偏高,光谱反射率 950~980、1 100~1 250、1 300~1 450 nm 区域内呈现下降趋势,其中 S3 处理反射率整体偏低, S2 处理相较于 S3 处理偏高, S1、S4 处理在 950~1 200 nm 基本重

合, S5 处理反射率最高,在 1 200~1 350 nm 波段范围内梯度明显,不同处理反射率从小到大表现为 S3 < S2 < S4 < S1 < S5(CK)。反射率较低者,其籽粒粗脂肪、粗蛋白质含量较高,而 1 450 nm 之后的波段 5 个处理曲线变化平滑且梯度不明显,这一波段范围内,光谱反射率会受到样本内水分的影响较大^[22]。

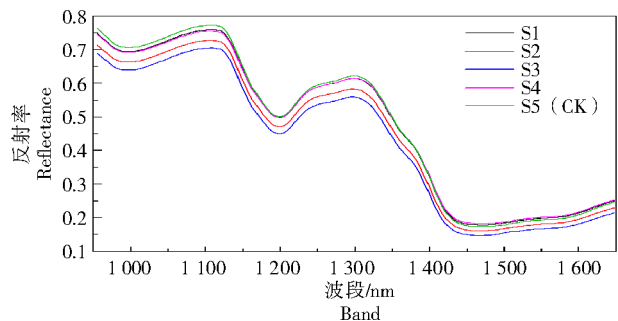


图 3 不同羊粪施用量下小米平均光谱曲线

Fig.3 The average spectral curves of millet with the different application rates of sheep manure

而小米的粗脂肪、粗蛋白结构中存在CH、OH、NH等化学官能团。由图3可知,在近红外光谱带中,CH二级倍频吸收谱带在1 150~1 250 nm,组合频在1 360~1 390、1 410~1 450 nm;游离OH二级倍频吸收谱带为960~980 nm和1 400~1 450 nm,结合OH二级倍频吸收谱带为1 000~1 130 nm;NH一级倍频吸收谱带为1 490~1 600 nm^[23]。

2.3 小米粗脂肪、粗蛋白质特征波段提取PLSR预测

为了降低冗余信息对预测模型精度的影响,对小米光谱进行特征波段提取,其中SPA基于逐步选择思想,所选波段具有直观物理统计意义,VCPA通过进化算法搜索最优的变量组合,而IRIV通过迭代的方式逐步选择保留最具信息量的变量,从而有效地提高了特征选择的准确性和效率。因而,本研

究采用VCPA、SPA、VCPA-IRIV、SPA-IRIV算法进行特征波段提取,选择出的波段如表1所示。

运用VCPA算法时,参考相关文献^[4]及试验调试后,设定采样次数1 000,即通过BMS对训练集进行1 000次采样,并取得采用概率相同的1 000个组合变量,经偏最小二乘法计算得1 000个交叉验证均方根误差值,设置变量子集中剩余子集占比10%,选择最小交叉验证均方根误差值100组,运用EDF剔除出现频次少的变量,后组成新变量集,最大迭代次数50,并重复上述执行过程,从而选择出相应的特征变量;SPA算法是目前较为常用的前向迭代搜索算法,能够较为准确剔除冗余变量来解决变量共线问题。

表1 特征波段提取结果
Tab.1 Results of key band extraction

成分指标 Component index	特征波段选择方法 Method to select key bands	被选波段数量/个 Number of bands	被选波段/nm Bands selected
粗脂肪 Crude fat	VCPA	10	1 044、1 049、1 120、1 304、1 308、1 313、1 318、1 337、1 360、1 577
	VCPA-IRIV	7	1 049、1 120、1 304、1 318、1 337、1 360、1 577
	SPA	10	1 120、1 134、1 181、1 205、1 266、1 360、1 398、1 426、1 454、1 582
	SPA-IRIV	8	1 120、1 134、1 181、1 266、1 360、1 398、1 426、1 454
粗蛋白质 Crude protein	VCPA	12	1 011、1 035、1 087、1 091、1 110、1 271、1 275、1 365、1 370、1 374、1 440、1 577
	VCPA-IRIV	9	1 011、1 087、1 091、1 271、1 275、1 365、1 370、1 440、1 577
	SPA	10	955、974、978、1 360、1 374、1 398、1 421、1 516、1 520、1 601
	SPA-IRIV	8	974、978、1 360、1 398、1 421、1 516、1 520、1 601

为减少波段冗余,VCPA、SPA分别与IRIV联用的特征波段提取方法对已进行特征波段提取后的变量进行二次提取,IRIV算法优势在于利用 $DMEAN < 0, P > 0.05$ 这2条判断准则将变量中无效信息和干扰信息剔除或保留强、弱信息,最终通过逆向运算最大限度“精选”变量,试验中根据文献^[23]及试验调试设置IRIV算法交叉验证次数5次,最大主成分10,运算程序随迭代数增加,变量维数随之降低,为了提升后续预测模型精度,本研究中保留了容易被剔除的弱信息变量。由表1可知,经过特征波段算法选取的特征波段,几近全部被选波段均落在CH、OH、NH吸收谱带内,特征波段联用算法在降低光谱数据维度、剔除冗长波段同时,尽可能多地完整保留粗脂肪、粗蛋白光谱信息,并为后续提升回归模型预测精度提供数据基础。

运用IRIV算法对光谱数据进行“精选”后,利用P值将粗脂肪和粗蛋白质光谱变量划分为强、弱信息,通过这一方法可更加直观反映样本光谱信

息与成分含量间的映射关系。特征变量联用算法在粗脂肪变量信息划分中,2种算法均划分4条强信息变量和6条弱信息变量,而粗蛋白质中采用VCPA-IRIV划分7条强信息变量5条弱信息变量,SPA-IRIV划分7条强信息变量3条弱信息变量。从数据结果看,粗蛋白质特征变量提取强信息保留较多,经过逆向运算后在粗脂肪VCPA-IRIV保留7条特征波段,SPA-IRIV保留8条特征波段,粗蛋白质VCPA-IRIV保留9条特征波段,SPA-IRIV保留8条特征波段。

对粗脂肪和粗蛋白质提取后的特征波段分别建立PLSR预测模型,预测结果如表2所示,粗蛋白质预测结果明显好于粗脂肪预测结果,其中采用SPA进行特征波段提取效果高于VCPA,这也验证了SPA具有能够从光谱数据矩阵中选择最小共线波段组合这一优点,而采用BMS为抽样方式的VCPA算法在特征波段选择上具有一定随机性,波段选择结果相较于SPA算法精度稍低,以粗蛋白质为例,

运用 SPA-IRIV 算法建立的 PLSR 预测模型 R_p 、RMSEP 分别为 0.88、0.59, 且 RPD 值为 1.99, 评价模型能达到 B 级, 精度最高。

利用精度最高的预测模型 SPA-IRIV-PLSR 获得粗脂肪和粗蛋白质回归方程分别为式(6)和(7)。

$$Y_{fat} = 604.1\lambda_{1120} - 675.4\lambda_{1134} - 39.9\lambda_{1181} - 535.9\lambda_{1266} + 1151.3\lambda_{1360} - 223.8\lambda_{1398} - 902.4\lambda_{1426} + 600.1\lambda_{1454} \quad (6)$$

$$Y_{protein} = 222.17\lambda_{974} - 261.91\lambda_{978} + 232.35\lambda_{1360} - 232.92\lambda_{1398} - 6.29\lambda_{1421} + 247.58\lambda_{1516} - 235.94\lambda_{1520} + 60.24\lambda_{1601} \quad (7)$$

式中, Y_{fat} 为粗脂肪的预测值, $Y_{protein}$ 为粗蛋白质的预测值, λ_i 为选出的特征波段 (i 为波长)。

2.4 小米粗脂肪、粗蛋白质含量可视化表达

利用 2.3 求得回归方程进行粗脂肪、粗蛋白质含量的可视化表达, 图 4 中 5 组图像为在不同羊粪施用量处理下粗脂肪和粗蛋白质含量积累变化可视化表达图像, 每幅图片下标出了对应各处理下 2 种成分平均含量, 右侧为成分变化渐变色带。从图 4 可清晰看出, S3 处理下, 粗脂肪与粗蛋白质含量较高, 且粗蛋白质含量在不同处理下区别较大。因而, 利用可视化可直观对小米粗脂肪、粗蛋白质含量进行表达, 同时也能反映出不同肥力下小米粗脂肪、粗蛋白质含量积累情况。该研究结论能够为实现不同肥力下小米营养成分的可视化快速检测提供理论依据。

表 2 小米粗脂肪和粗蛋白质不同特征波段提取建立 PLSR 预测
Tab.2 PLSR prediction established by different key band extractions from crude fat and protein of millet

成分指标 Component index	处理方法 Treatment	训练集 Training			预测集 Predicting		
		R_c	RMSEC	RPD_c	R_p	RMSEP	RPD_p
粗脂肪 Crude fat	RAW	0.54	1.14	1.05	0.53	1.25	0.78
	VCPA	0.39	1.25	0.31	0.37	1.41	0.37
	VCPA-IRIV	0.38	1.32	0.30	0.36	1.52	0.25
	SPA	0.40	1.33	0.50	0.39	1.31	0.47
	SPA-IRIV	0.46	1.23	0.52	0.4	1.10	0.50
粗蛋白质 Crude protein	RAW	0.84	0.64	1.51	0.84	0.72	1.43
	VCPA	0.80	0.78	1.51	0.78	0.80	1.45
	VCPA-IRIV	0.61	1.02	1.35	0.59	1.13	1.27
	SPA	0.88	0.60	2.09	0.87	0.63	1.95
	SPA-IRIV	0.90	0.57	2.16	0.88	0.59	1.99

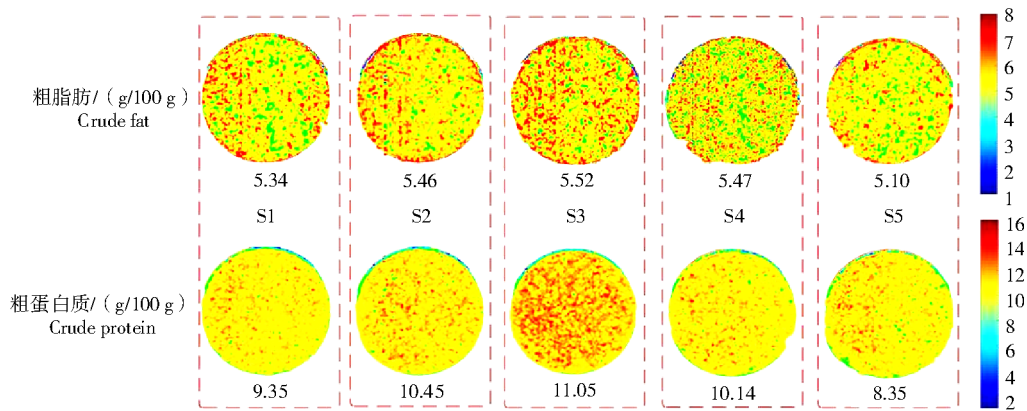


图 4 不同羊粪施用量下小米粗脂肪和粗蛋白质含量可视化表达对比
Fig.4 Comparison of visualization representations of crude fat and protein content of millet with the different application rates of sheep manure

3 结论与讨论

随着羊养殖业规模化发展, 羊场废弃物羊粪处理成为影响全舍饲羊的制约因素, 而羊粪经过处理用于种植业不但可以实现废弃物高效利用、减少不

当处理带来的污染, 也能有效改善农作物种植环境、提升土壤肥力、提高农产品品质^[23-25]。研究表明^[26], 施用腐熟羊粪可以提高作物根部土壤微生物种群多样性和种群丰度, 稳定群落结构, 促进有机质分解、营养物质循环利用, 达到改善作物养分吸

收效果的目的。本研究将高光谱成像与化学计量学相结合,设定不同的羊粪施用量,并收集了358份小米粉样品,用于检测粗脂肪和粗蛋白质含量。采用SPA算法能够选择光谱信息最少冗余的波长以解决共线性问题,简化关键变量提取过程。张恒等^[27]利用SPA提取特征变量后建立预测模型,对土壤全氮和碱解氮含量进行预测,并实现了凉山州植烟土壤理化指标快速预测。IRIV算法利用Mann-Whitney U检验来确定变量的特征,并且最优变量集不包含无信息变量和干扰信息变量,通过对变量集进行逆消除,可得到变量的最优组合。杨欢等^[28]利用IRIV提取特征变量后建立玉米种子品种纯度分类模型,模型识别效果较好。为了得到最优变量组合,ZHAO等^[29]通过特征变量提取联用算法对番茄色素含量进行预测,梁琨等^[30]采用ARS-IRIV结合提取特征变量后建立的预测模型对库尔勒香梨可溶性固形物含量进行检测。以上研究表明,弱信息变量在组合中起着重要作用,有效地提高了模型的预测性能,证明了本研究针对小米粗脂肪和粗蛋白质含量高光谱检测提出的“从粗到细”特征变量提取策略的可靠性。根据可视化结果显示,不同羊粪施用量下小米中粗脂肪和粗蛋白质的积累情况,所选试验品种具有良好的适应性,粗脂肪含量在小米营养成分中占比较小,含量的差异取决于基因的稳定表达。检测模型对粗蛋白质含量取得了良好的预测结果,积累的可视化显示了不同羊粪处理下的不同变化,这与物理和化学测量结果一致。粗脂肪含量的预测性能并不理想,在可视化积累方面存在一定的偏差,但通过可视化仍可以对小米的粗脂肪含量的积累做出定性判断。从试验结果来看,公顷施90 m³羊粪最适合籽粒的粗脂肪、粗蛋白质的合成,可得出该肥力对籽粒粗脂肪、粗蛋白质的积累具有积极作用。而利用小米粗脂肪和粗蛋白质含量可视化表达能够更加直观反映在不同施肥处理下粗蛋白质的图像区别度高于粗脂肪,这与小米内部粗蛋白质含量高于粗脂肪,且受羊粪施用量的影响大于粗脂肪有关,也与粗脂肪预测模型精度较低有关;对比粗脂肪和粗蛋白质含量,公顷施90 m³羊粪最高,对照最低,这与理化测定结果相符,因而通过采用特征波段联用算法结合可视化表达等数据处理方式直观描述小米粗脂肪、粗蛋白质含量具有一定优势,同时亦可以图像化表达最适宜羊粪施用量在小米粗脂肪和粗蛋白质含量积累上的优势。

本研究运用特征波段联用算法对小米高光谱数据进行特征波段提取后建立了小米粗脂肪、粗蛋

白质高光谱PLSR预测模型,并对小米粗脂肪和粗蛋白质含量进行可视化表达并进行对比分析。数据分析中分别对小米粗脂肪和粗蛋白质的全波段光谱数据和经特征波段联用算法提取特征波段后建立PLSR模型,经过模型对比选择SPA-IRIV-PLSR作为最佳预测模型,为更加直观反映不同羊粪施用量对含量积累规律的影响,利用粗脂肪和粗蛋白质回归方程进行2种成分含量不同羊粪施用量下积累规律的可视化表达,从所得图像可以清晰看到公顷施90 m³羊粪处理在小米粗脂肪和粗蛋白质含量的积累优势。综上,该数据分析方法可应用于多变量降维、成分含量可视化表达的研究中,研究结论为高光谱成像应用于小米其他营养成分的检测中提供了理论基础。

参考文献:

- [1] 李君霞,马小倩,代书桃,等. 谷子品质性状研究进展[J]. 河南农业科学,2023,52(9):14-23.
LI J X, MA X Q, DAI S T, et al. Research progress on quality traits of foxtail millet[J]. Journal of Henan Agricultural Sciences, 2023, 52(9): 14-23.
- [2] 刘宇杰,陈银焕,杨修仕,等. 小米营养及功能成分研究进展[J]. 粮食与油脂,2020,33(5):1-3.
LIU Y J, CHEN Y H, YANG X S, et al. Research progress on nutrition and functional components of millet[J]. Cereals & Oils, 2020, 33(5): 1-3.
- [3] WU D, SUN D W. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: a review: Part II: applications[J]. Innovative Food Science & Emerging Technologies, 2013, 19: 15-28.
- [4] XIA Y, XU Y F, LI J B, et al. Recent advances in emerging techniques for non-destructive detection of seed viability: A review[J]. Artificial Intelligence in Agriculture, 2019, 5(1): 35-47.
- [5] SAHA D, MANICKAVASAGAN A. Machine learning techniques for analysis of hyperspectral images to determine quality of food products: a review[J]. Current Research in Food Science, 2021, 4: 28-44.
- [6] 田翔,刘思辰,王海岗,等. 近红外漫反射光谱法快速检测谷子蛋白质和淀粉含量[J]. 食品科学,2017,38(16):140-144.
TIAN X, LIU S C, WANG H G, et al. Application of near infrared diffuse reflectance spectroscopy in rapid detection of crude protein and starch in foxtail millet[J]. Food Science, 2017, 38(16): 140-144.
- [7] 王浩,于港华,侯颖,等. 小米硒含量近红外预测模型的建立与评价[J]. 中国农业大学学报,2021,26(10):157-163.
WANG H, YU G H, HOU Y, et al. Establishment and evaluation of near-infrared prediction model for selenium content in millet[J]. Journal of China Agricultural University, 2021, 26(10): 157-163.
- [8] 赵环,宦克为,郑峰,等. 基于变量组合集群分析法的小麦蛋白质近红外光谱变量选择方法研究[J]. 长春理工大学学报(自然科学版),2016,39(5):51-54.
ZHAO H, HUAN K W, ZHENG F, et al. Research on variable selection of protein in wheat near infrared spectroscopy based on

- latent projective graph[J]. Journal of Changchun University of Science and Technology(Natural Science Edition), 2016, 39(5): 51-54.
- [9] 李颖. 基于 Vis: NIR 光谱的木材树种鉴别及密度模型研究[D]. 哈尔滨: 东北林业大学, 2019.
- LI Y. Study on identification of wood species and density model based on Vis: NIR spectrum[D]. Harbin: Northeast Forestry University, 2019.
- [10] 于雷, 章涛, 朱亚星, 等. 基于 IRIV 算法优选大豆叶片高光谱特征波长变量估测 SPAD 值[J]. 农业工程学报, 2018, 34(16): 148-154.
- YU L, ZHANG T, ZHU Y X, et al. Determination of soybean leaf SPAD value using characteristic wavelength variables preferably selected by IRIV algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(16): 148-154.
- [11] 余海东. 近红外光谱波长选择方法及其在罗非鱼片新鲜度检测中的应用[D]. 海口: 海南大学, 2021.
- YU H D. Research on wavelength selection method in near-infrared spectra and its application for the rapid detection of freshness in tilapia fillets[D]. Haikou: Hainan University, 2021.
- [12] 国家质量监督检验检疫总局, 中国国家标准化管理委员会. 植物油料 含油量测定: GB/T 14488.1—2008[S]. 北京: 中国标准出版社, 2009.
- General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Oilseeds-determination of oil content: GB/T 14488.1-2008[S]. Beijing: Standards Press of China, 2009.
- [13] 国家卫生和计划生育委员会, 国家食品药品监督管理总局. 食品安全国家标准 食品中蛋白质的测定: GB 5009.5—2016[S]. 北京: 中国标准出版社, 2017.
- National Health and Family Planning Commission, State Food and Drug Administration. National food safety standards-determination of protein for food safety: GB 5009.5-2016[S]. Beijing: Standards Press of China, 2017.
- [14] WU Z S, OUYANG G Q, SHI X Y, et al. Absorption and quantitative characteristics of C-H bond and O-H bond of NIR[J]. Optics and Spectroscopy, 2014, 117(5): 703-709.
- [15] FAN N Y, MA X, LIU G S, et al. Rapid determination of TBARS content by hyperspectral imaging for evaluating lipid oxidation in mutton[J]. Journal of Food Composition and Analysis, 2021(103): 1-10.
- [16] WAN G L, LIU G S, HE J G, et al. Feature wavelength selection and model development for rapid determination of myoglobin content in nitrite-cured mutton using hyperspectral imaging[J]. Journal of Food Engineering, 2020(287): 1-9.
- [17] MARTÍN J F G. Optical path length and wavelength selection using Vis/NIR spectroscopy for olive oil's free acidity determination[J]. International Journal of Food Science and Technology, 2015, 50(6): 1461-1467.
- [18] LI Y. Rapid quality discrimination of grape seed oil using an extreme machine learning approach with Near-Infrared(NIR) spectroscopy[J]. Spectroscopy, 2022, 36: 14-20.
- [19] YUN Y H, BIN J, LIU D L, et al. A hybrid variable selection strategy based on continuous shrinkage of variable space in multivariate calibration[J]. Analytica Chimica Acta, 2019, 1058: 58-69.
- [20] 褚小立. 化学计量学方法与分子光谱分析技术[M]. 北京: 化学工业出版社, 2011.
- CHU X L. Molecular spectroscopy analytical technology combined with chemometrics and its applications[M]. Beijing: Chemical Industry Press, 2011.
- [21] CHANG C W, LAIRD D A, MAUSBACH M J, et al. Near-infrared reflectance spectroscopy - principal components regression analyses of soil properties[J]. Soil Science Society of America Journal, 2001, 65(2): 480-490.
- [22] ZHAN Y M, GUO W C. Moisture content detection of maize seed based on visible/near-infrared and near-infrared hyperspectral imaging technology[J]. International Journal of Food Science and Technology, 2020, 55(2): 631-640.
- [23] SUN H X, ZHANG S J, REN R, et al. Detection of soluble solids content in different cultivated fresh jujubes based on variable optimization and model update[J]. Foods, 2022, 11(16): 2522.
- [24] 芦伟, 高月锋, 简路洋, 等. 羊粪作为有机肥对植物和土壤的影响[J]. 家畜生态学报, 2019, 40(9): 86-90.
- LU W, GAO Y F, JIAN L Y, et al. Application effect of sheep manure on plant and soil[J]. Journal of Domestic Animal Ecology, 2019, 40(9): 86-90.
- [25] 李亚飞, 张翔, 常栋, 等. 施用量及施用方式对连作烟田土壤理化性状和经济效益的影响[J]. 山西农业科学, 2020, 48(12): 1964-1968, 1983.
- LI Y F, ZHANG X, CHANG D, et al. Effects of fertilization amount and patterns of sheep dung on soil properties and tobacco economic performance in continuous tobacco planting fields[J]. Journal of Shanxi Agricultural Sciences, 2020, 48(12): 1964-1968, 1983.
- [26] 朱云芬, 李蓉, 向极钎, 等. 羊粪资源化利用的研究进展[J]. 湖北农业科学, 2021, 60(11): 12-15.
- ZHU Y F, LI R, XIANG J Q, et al. Research advances in resources utilization of sheep manure[J]. Hubei Agricultural Sciences, 2021, 60(11): 12-15.
- [27] 张恒, 梁太波, 冯文强, 等. 基于连续投影算法的土壤全氮和碱解氮含量高光谱估测[J]. 中国烟草科学, 2023, 44(5): 103-113.
- ZHANG H, LIANG T B, FENG W Q, et al. Hyperspectral estimation of total nitrogen and alkali hydrolysable nitrogen contents in tobacco growing soil based on successive projection algorithm[J]. Chinese Tobacco Science, 2023, 44(5): 103-113.
- [28] 杨欢, 罗斌, 张晗, 等. 基于高光谱成像技术和 IRIV 算法的玉米种子品种纯度识别[J]. 江苏大学学报(自然科学版), 2023, 44(2): 159-165.
- YANG H, LUO B, ZHANG H, et al. Recognition of maize seed variety purity based on hyperspectral imaging technology and IRIV algorithm[J]. Journal of Jiangsu University (Natural Science Edition), 2023, 44(2): 159-165.
- [29] ZHAO J G, CHEN N, ZHU T Y, et al. Simultaneous quantification and visualization of photosynthetic pigments in *Lycopersicon esculentum* mill. under different levels of nitrogen application with visible-near infrared hyperspectral imaging technology[J]. Plants, 2023, 12(16): 2956.
- [30] 梁琨, 刘全祥, 潘磊庆, 等. 基于高光谱和 CARS-IRIV 算法的‘库尔勒香梨’可溶性固形物含量检测[J]. 南京农业大学学报, 2018, 41(4): 760-766.
- LIANG K, LIU Q X, PAN L Q, et al. Detection of soluble solids content in 'Korla fragrant pear' based on hyperspectral imaging and CARS-IRIV algorithm[J]. Journal of Nanjing Agricultural University, 2018, 41(4): 760-766.