doi:10.16104/j.issn.1673-1891.2021.03.005

# 饲用高粱 3 个品质性状近红外模型的构建研究

王丽华,蔡孟禹,陈 蒙,李杰勤\*

(安徽科技学院农学院,安徽凤阳 233100)

摘 要:通过化学方法测定了 245 个高粱材料整株的蛋白质、中性洗涤纤维(NDF)、酸性洗涤纤维(ADF)的含量,并采集其整 株粉碎样品的近红外光谱,构建了高粱 3 个饲用品质性状的近红外模型。结果表明:这 245 份高粱材料整株的蛋白质均值为 5.97%;变幅为 2.38%~13.48%,NDF 均值为 64.67%;变幅为 40.34%~82.45%,ADF 均值为 32.37%;变幅为 13.25%~52.86%, 说明实验材料具有良好的代表性。相关性分析表明蛋白质与 NDF 和 ADF 之间都呈显著负相关,而 NDF 和 ADF 之间呈显著 正相关。利用近红外光谱模型软件构建这 3 个饲用品质性状的近红外模型,利用偏最小二乘法和多元信号修正方法建立原 始模型。进一步分析表明,优化 6 即二阶导数光谱、未滤嗓、未优化波长且手动去除异常数据点的优化可获得 3 个品质性状的 最优近红外模型。

关键词:高粱;蛋白质;NDF;ADF;近红外模型 中图分类号:S548 文献标志码:A 文章编号:1673-1891(2021)03-0022-06

# Development of Near Infrared Model for Three Quality Traits of Feed Sorghum

WANG Lihua, CAI Mengyu, CHEN Meng, LI Jieqin\*

(College of Agriculture, Anhui Science and Technology University, Fengyang, Anhui 233100, China)

Abstract: The content of protein, neutral detergent fiber (NDF) and acid detergent fiber (ADF) in 245 strains of sorghum were determined by chemical method. The near infrared spectra of whole plant were collected and the near infrared model was developed by the partial least square method (PLS). The results showed that the average protein content was 5.97%, with a variation range of  $2.38\% \sim 13.48\%$ , the average NDF content was 64.67%, varying between 40.34% and 82.45%, the average ADF content was 32.37% and the amplitude was  $13.25\% \sim 52.86\%$ . The correlation analysis showed that there was a significant negative correlation between protein and NDF and ADF, while NDF and ADF showed a significant positive correlation. The results showed that the near infrared model had a better predictive performance, and the optimization 6 was the best model for all three quality traits through the optimization study.

Keywords: sorghum; protein; neutral detergent fiber (NDF); acid detergent fiber (ADF); Near Infrared Model

# 0 引言

高粱是禾本科(Gramineae)高粱属(Sorghum bicolor L.)作物。根据用途不同,可以将高粱分为食 用高粱、饲用高粱、能源用高粱、酒用高粱和帚用高 粱。随着近年来人们生活水平的提高,饲用高粱在 我国的种植面积越来越大,在畜牧业发展中发挥着 越来越重要的作用<sup>[1]</sup>。蛋白质含量和纤维素含量 是评价饲用高粱品质性状的重要指标。传统化学 法测定饲用高粱蛋白质含量和纤维素含量的测定 时间长、样品用量大、成本较高。因此,当育种中需 要对大量的材料进行筛选时,传统的化学法则成了 提高筛选效率的瓶颈。

近红外反射光谱技术(NIRS, near-infrared reflectance spectroscopy)是一项新型物理测定技术,具 有样品前处理简单、快速,可同时测定一种或多种 化学成分含量的特点<sup>[2]</sup>,在农业、食品、化工、医药 等多个领域得到了广泛应用<sup>[3-7]</sup>。NIRS 技术还具 有快速、无损特点,使其在测定完成后,样品无损且 保持活性,所以在作物育种、种子质量检测等领域 具有良好的应用前景。目前,在高粱的籽粒性状和 饲用品质性状等方面也有一些近红外光谱应用的

收稿日期:2021-06-01

基金项目:安徽省高校自然科学研究项目(KJ2019A0811)。

作者简介:王丽华(1980—),女,河北定州人,实验师,硕士,研究方向:作物遗传育种。\*通信作者:李杰勤(1980—),四川宜 宾人,教授,博士,研究方向:作物遗传育种。

相关研究。刘敏轩等<sup>[8]</sup>建立了 60 份高粱及其近缘 种籽粒中多酚类物质的近红外光谱模型; 邵春莆 等<sup>[9]</sup>建立了 63 份酒用高粱籽粒中总淀粉、直链和支 链淀粉含量的近红外光谱模型; 王勇生等<sup>[10]</sup>利用 110 份高粱构建了粗脂肪、粗纤维和灰分的近红外 光谱模型。但是,这些研究主要集中在高粱的籽粒 性状, 饲用品质性状的研究较少, 同时使用的材料 数量也都较少。因此, 建立一个具有广泛适应性的 高粱饲用品质性状的近红外模型具有非常重要的 意义。

本文利用化学法测定了 245 个高粱品种(系) 全株的蛋白质、中性洗涤纤维(NDF)和酸性洗涤纤 维(ADF)含量,并对这些样品进行近红外光谱扫 描,构建了高粱3个饲用品质性状的近红外光谱分 析模型,简化了品质性状测定流程,提高了品质性 状测定的工作效率,也降低了测定性状所需的费 用,为饲用高粱的品质育种奠定了良好的基础。

## 1 材料与方法

#### 1.1 试验材料和样品处理

本试验选用国际半干旱热带作物研究所(ICRI-SAT)提供的高粱微核心种质材料,245个高粱品种 (系)名称及来源地如表1所示。在245个高粱品 种(系)抽穗前刈割地上部分,每个品系取3株,每 个单株分别置于120℃杀青1h,105℃烘干72h, 直至质量不发生改变。将烘干后的样品单株粉碎 过孔径0.425 mm 筛。过筛样品用于后续的化学分 析和近红外光谱扫描。

序号 序号 来源地 序号 来源地 序号 序号 来源地 来源地 来源地 1 USA 99 India South Africa 197 Lesotho 50 Japan 148 2 USA Pakistan 149 South Africa 198 Lesotho 51 100 Ethiopia 3 South Africa USA 52 South Africa 101 Sudan 150 199 Lesotho 4 USA 53 Uganda 102 Sudan 151 Zambia 200 Lesotho 5 USA 54 103 Bangladesh 152 Zambia 201 Lesotho Uganda 6 India 55 Kenya 104 Botswana 153 Ghana 202 Lesotho 7 105 Botswana 154 203 South Africa India 56 Kenva Ethiopia 8 China 57 Zimbabwe 155 Ethiopia 204 South Africa Kenva 106 9 China 58 Sudan 107 India 156 Rwanda 205 China 10 59 Thailand 206 Zimbabwe China 108 Senegal 157 Mali 11 India 60 Chad 109 Niger 158 Mali 207 Zimbabwe 12 Chad 159 208 Zimbabwe South Africa 110 Niger Mali 61 13 South Africa 62 USA 111 USA 160 Mali 209 Zimbabwe 14 South Africa 210 Zimbabwe 63 Ethiopia 112 USA 161 Mali 15 South Africa 64 Ethiopia 113 USA 162 Mali 211 Zimbabwe 16 Iran 65 Ethiopia 114 USA 163 Togo 212 Zimbabwe 17 Afghanistan Ethiopia USA 164 Benin 213 Zimbabwe 66 115 18 South Africa Zimbabwe USA 165 Madagascar 214 Zimbabwe 67 116 19 South Africa 68 Sudan USA 166 215 China Egypt 117 20 South Africa Nigeria 69 Australia 118 USA 167 216 China 21 Kenva 70 USA 119 USA 168 South Africa 217 China 22 South Africa 71 Yemen 120 Indonesia 169 South Africa 218 China 23 170 Sudan 219 India 72 Turkey 121 Kenya China 24 Burkina Faso 220 India 73 India 122 Malawi 171 China 74 Sierra Leone 25 India Ethiopia 123 Malawi 172 221 China 26 75 173 222 India Nicaragua 124 Syrian Arab Morocco China 27 Cuba 174 South Africa 223 India 76 125 Botswana Korea South Africa 28 India 77 Venezuela 126 Botswana 175 224 Korea 176 29 India 78 Mexico 127 Sri Lanka Yemen 225 Korea 30 India 79 South Africa MYA 177 Yemen 226 128 Korea 31 India 80 South Africa 129 Somalia 178 Yemen 227 Korea 32 South Africa 179 228 India 81 130 Somalia Yemen Cameroon 33 India 82 South Africa 131 Sudan 180 Yemen 229 Cameroon

表 1 245 个高粱品种(系)的编号和来源

第35卷

									天水
序号	来源地	序号	来源地	序号	来源地	序号	来源地	序号	来源地
34	India	83	South Africa	132	Zambia	181	Yemen	230	Uganda
35	India	84	Argentins	133	Ethiopia	182	Yemen	231	Uganda
36	India	85	Botswana	134	Ethiopia	183	Yemen	232	Uganda
37	India	86	Cameroon	135	Ethiopia	184	Yemen	233	Burundi
38	India	87	Cameroon	136	Ethiopia	185	Yemen	234	Zaire
39	India	88	Cameroon	137	Ethiopia	186	Swaziland	235	Algeria
40	India	89	Cameroon	138	Gambia	187	Swaziland	236	Yemen
41	India	90	Cameroon	139	Mozambique	188	Swaziland	237	Yemen
42	India	91	Cameroon	140	Yemen	189	Swaziland	238	Yemen
43	Uganda	92	Cameroon	141	Yemen	190	Swaziland	239	India
44	Nigeria	93	Cameroon	142	Tanzania	191	Swaziland	240	India
45	Nigeria	94	Cameroon	143	Tanzania	192	Swaziland	241	Somalia
46	Nigeria	95	Cameroon	144	Tanzania	193	Swaziland	242	Tanzania
47	Nigeria	96	Cameroon	145	India	194	Swaziland	243	Honduras
48	Nigeria	97	India	146	South Africa	195	Lesotho	244	Kenya
49	Nigeria	98	India	147	South Africa	196	Lesotho	245	India

#### 1.2 化学分析方法

#### 1.2.1 样品处理

称取1g左右样品粉末,记录其质量,放入消化 管中。先加少量水润湿,再加入5mL浓硫酸,在消 化炉上先进行预热,再将消化炉温度设定在380℃ 进行消煮,时间为20min;消煮液冷却后,加入过氧 化氢20滴,继续预热并消煮;重复3~4次,直至液 体变为透明,继续加热,将其中的过氧化氢完全受 热分解后,冷却并移至100mL容量瓶中,用蒸馏水 定容至100ml后装瓶、编号。

## 1.2.2 蛋白质、NDF 和 ADF 质量分数测定

蛋白质含量测定采用全自动连续流动分析仪 (Seal AutoAnalyzer AA3); NDF 和 ADF 测定采用半 自动纤维分析仪(美国 ANKOMA200i)。

## 1.3 近红外光谱数据采集及模型建立

#### 1.3.1 粉末样品的光谱数据采集

试验采用 Antaris II (Thermo Fisher) 近红外光谱 分析仪。将样品装满旋转样品杯,在3 999.63~ 10 000.00 cm<sup>-1</sup>谱区范围扫描 64 次,分辨率设置为 8.0/cm,采集样品的吸收光谱。为了消除空气影响, 在扫描样品前扫描背景光谱,扫描次数和样品扫描 次数相同。

### 1.3.2 近红外模型的建立和优化

将扫描的光谱数据导入 TQ Analyst 软件中,并 在软件中输入化学分析方法所测得蛋白质、NDF 和 ADF 的数据。通过软件设置通用选项为偏最小二 乘(PLS)回归,并选择光程类型选项为多元信号修 正(MSC)。这是因为粉末样品光谱采集时,由于样 品颗粒尺寸、均匀性等的影响,光程无法保持恒定, 故采用 MSC 消除影响。波长设置为原始波长 3 999.63~10 000.00 nm。运行软件建立模型,记录 校正标准偏差(RMSEC)及其相关系数、预测标准偏 差(RMSEP)及其相关系数和使用因子数。此时建 立的模型为原始模型,通过以下几步分别设置不同 的优化,记录 RMSEC 及其相关系数、RMSEP 及其相 关系数和使用因子数。优化方法如下:

优化1:选择光谱数据格式为一阶导数光谱。

优化2:选择光谱数据格式为二阶导数光谱。

优化3:在二阶导数光谱基础上,选择滤噪方法 为 Savitzky - Golay,其 Data points 参数为7, Polynomial order 参数为3。

优化4:在二阶导数光谱基础上,选择滤噪方法为 Norris derivative,其 Segment length 参数为5,Gap between segments 参数为5。

优化5:在二阶导数光谱基础上,不选择滤噪, 选择自动优化波长。

优化6:在模型中观察,手动去除异常数据点, 其他同优化2。

## 1.4 数据的处理

化学分析方法测得的数据导入 SPSS 数据处理 软件中,计算数据的平均值,变幅、变异系数及相关 系数。

#### 2 结果与分析

# 2.1 245 份高粱全株蛋白质、NDF、ADF 质量分数及 其相关性

从表 2 可知,245 个高粱全株的这 3 个性状的 变异系数较大,分别为 30.47%,11.11%,16.68%,说

住。

表 2 高刻	<b>梁的蛋白质、</b> NDF	、ADF 的化学法	测定值统计
品质性状	均值±标准差	变幅	变异系数
蛋白质	$5.97 \pm 1.82$	2.38~13.48	30.47
NDF	64.67±7.18	40.34~82.45	11.11
ADF	$32.37 \pm 5.40$	13.25~52.86	16.68

进一步对这 3 个性状进行相关性分析,结果表明:蛋白质与 NDF 和 ADF 极显著负相关, NDF 和 ADF 极显著正相关(表 3)。相关性的分析,对于近 红外模型的构建有一定的参考价值,可以用于构建 模型的结果检验。

表 3 高粱全株蛋白质含量、NDF 和 ADF 间的相关性分析

	蛋白质	NDF	ADF
蛋白质	1		
NDF	-0.262 * *	1	
ADF	-0.261 * *	0.875 * *	1

注:"\*\*"表示在 0.01 水平显著相关。

# 2.2 245 份高粱秸秆品质性状模型的建立、优化及 分析

### 2.2.1 近红外光谱优化分析

有应用价值的近红外光谱模型应具有较高的 光谱分辨率和灵敏度的特点,并且其相关系数高、 预测误差小<sup>[3]</sup>。在未进行任何优化时,生成的光谱 图较粗糙(图1)。从图1可以看出,在不同波段有 几个较为明显的峰。进一步对原始光谱进行一阶 导数和二阶导数优化,对光谱进行一阶导数优化 后,相比原始模型,峰的数量和幅度明显增加(图 2a)。对光谱进行二阶导数处理后,峰的数量和幅 度进一步增加,但噪声峰也明显增加(图 2b)。进一 步滤噪处理后(图 2c 和 2d),峰的平滑度增加,噪声 明显较小。这说明一阶导数和二阶导数处理后,滤 噪处理是必不可少的步骤。



## 2.2.2 蛋白质含量模型建立、优化及分析

将全株蛋白质含量的数据与不同优化条件的 光谱进行模型构建(表 4)。相比于未优化模型, 优化2,4和6的RMSEC都小于未优化模型,而优化



1,3,5则大于未优化模型。优化6的RMSEC最小,为0.990,相关系数最大,为0.8215;其RMSEP值也较小。

## 2.2.3 NDF 含量模型建立、优化及分析

将全株 NDF 的数据与不同优化条件的光谱进 行模型构建(表5)。由表5可知,这些优化条件下, RMSEC 和 RMSEP 相差不大。说明用全株 NDF 含 量数据进行模型构建具有很好的拟合性。在这些 优化条件中,优化 6 的 RMSEC 和 RMSEP 值都最小,而且 RMSEC 相关系数的值也最大,达到 0.9057; RMSEP 相关系数也较大,为 0.8634, 排第 2 名。这说明,优化 6 是最优的模型。另外,相比于未优化模型,6 个优化模型中只有优化 5 的 RMSEC 略大于未优化模型,这 6 个优化模型的 RMSEP 值都小于未优化模型,而且 RMSEC(除优化 5)和 RMSEP 对应的相关系数都大于未优化模型。

表 4 蛋白质含量模型优化结果对比

农** 虽口灰百里铁主优化和木利比								
优化	波长/cm <sup>-1</sup>	RMSEC	相关系数	RMSEP	相关系数	因子数		
未优化	3 999.63~10 000.00	1.09	0.795 8	1.24	0.814 0	8		
优化 1	3 999.63~10 000.00	1.11	0.771 6	1.27	0.824 6	6		
优化 2	3 999.63~10 000.00	0.995	0.820 2	1.44	0.777 6	6		
优化 3	3 999.63~10 000.00	1.18	0.735 3	1.34	0.809 9	4		
优化 4	3 999.63~10 000.00	1.06	0.786 4	1.27	0.823 7	7		
优化 5	4277.34~4 300.48	1.43	0.571 7	1.74	0.650 8	6		
优化 6	3 999.63~10 000.00	0.990	0.821 5	1.44	0.778 9	6		

表 5 NDF 含量模型优化结果对比

优化	波长/cm <sup>-1</sup>	RMSEC	相关系数	RMSEP	相关系数	因子数
未优化	3 999.63~10 000.00	0.044 5	0.814 7	0.045 1	0.789 8	5
优化 1	3 999.63~10 000.00	0.039 6	0.856 1	0.042 0	0.837 3	7
优化 2	3 999.63~10 000.00	0.039 5	0.857 0	0.037 0	0.859 7	5
优化 3	3 999.63~10 000.00	0.039 9	0.853 8	0.039 7	0.842 7	5
优化 4	3 999.63~10 000.00	0.039 8	0.854 9	0.037 1	0.873 8	9
优化 5	4 254.20~4 354.48	0.044 7	0.812 9	0.040 5	0.825 9	8
优化 6	3 999.63~10 000.00	0.030 5	0.905 7	0.036 2	0.863 4	6

#### 2.2.4 ADF 含量模型建立、优化及分析

将全株 ADF 的数据与不同优化条件的光谱进 行模型构建。从表 6 可知, RMSEC 在 6 种优化条件 下相差不大。说明用全株 ADF 含量数据进行模型 构建具有很好的拟合性。在这些优化条件中, 优化 1 和 6 的 RMSEC 小于未优化模型,所有的 RMSEP 都小于未优化模型。相比于优化 1,优化 6 模型的 RMSEP 值更小,而其对应的相关系数也大于优化 1。因此,综合考虑 RMSEC 和 RMSEP,优化 6 是最 优模型。

表 6 ADF 含量模型优化结果对比

优化	波长/cm <sup>-1</sup>	RMSEC	相关系数	RMSEP	相关系数	因子数	
未优化	3 999.63~10 000.00	0.030 1	0.815 5	0.064 2	0.388 0	6	
优化 1	3 999.63~10 000.00	0.026 4	0.861 8	0.063 6	0.445 3	8	
优化 2	3 999.63~10 000.00	0.030 3	0.818 1	0.045 0	0.695 2	5	
优化 3	3 999.63~10 000.00	0.030 9	0.804 2	0.057 3	0.503 9	6	
优化 4	3 999.63~10 000.00	0.031 5	0.795 6	0.059 2	0.461 6	5	
	4 254.20~4 269.63						
优化 5	4 304.34~4 412.33	0.031 1	0.801 8	0.056 0	0.536 3	5	
	5 122.01~5 133.58						
优化 6	3 999.63~10 000.00	0.027 0	0.847 8	0.055 0	0.552 1	5	

## 3 讨论与结论

近红外光谱分析是复杂、重叠光谱分析,测定

样品不经预处理,故样品的状态、测定方式以及测定的条件都会影响测定结果<sup>[11]</sup>。因此得到的光谱 模型是由样品待测成分、样品背景成分、测量条件、 测定环境等信息变量的复杂参数构成<sup>[12]</sup>。为了构 建较为理想的近红外模型,本文选择了 245 份不同 来源的高粱品种(系),其均值、变幅及变异系数说 明这些试验材料来源丰富且具有代表性。祝诗平 等<sup>[13]</sup>研究发现过孔径 0.425 mm 筛的样品 r<sup>2</sup>值最高 (0.936 4),建模效果最好,本试验样品也选择过 40 目筛,旨在尽可能保留高粱整株成分,有助于模型 构建。

在模型的优化上,光程选择 MSC,MSC 主要是 消除颗粒分布不均匀及颗粒大小产生的散射影 响<sup>[14]</sup>,说明样品的物理状态对光谱的扫描影响不 同。方彦等<sup>[15]</sup>认为采用二阶导数处理效果最好,其 次为一阶导数处理,而三阶导数处理效果最差,本 研究中也发现二阶导数光谱可以很好地优化近红 外模型,降低 RMESC 值,在以后的建模过程中,可 以着重使用二阶导数光谱优化;平滑处理是滤除噪 声的常见方法,方彦等<sup>[15]</sup>将平滑波长点设定为 6, 本研究滤噪优化在对比原始模型数据后发现,滤噪 有一定的优化效果,可以为优化模型提供参考,滤 噪是为了减弱以至消除各种噪声对光谱信息的影 响,但在减弱的过程中也可能会影响到正常的光谱 信息;在光波范围优化后,其 RMSEC 值会有不同程 度的改变,但对结果的影响不大;异常样品是近红 外预测模型的构建过程中经常遇见的一个问题,对 其合理的剔除,有利于提高预测模型的预测精 度<sup>[16]</sup>,本试验中3个饲用品质性状的近红外模型构 建都选择了优化6即二阶光谱下,未滤噪优化,光谱 范围未优化,手动去除异常样品的模型为最优模 型,说明异常样品的剔除对不同品质性状模型的优 化具有很好的效果。

综合来看,本研究构建了高粱蛋白质、NDF和ADF的近红外模型,其RMSEC值分别为0.990,0.0305,0.0270,其RMSEC相关系数分别为0.8215,0.9057,0.8478。其值在不同优化中数值最优或相对较优,因此,构建的模型都有一定的应用价值。

#### 参考文献:

- [1] 王晓娜,徐春城,温定英,等.不同测定方法对青贮饲料中 NDF 和 ADF 含量的影响[J].草业科学,2012,29(1):144-149.
- [2] OSBORNE B G, FEARN T, HINDLE P H. Practical near infrared spectroscopy with applications in food and beverage analysis
  [M].UK:Longman Scientific and Technical, 1993:227.
- [3] 李佳佳,洪慧龙,万明月,等.基于近红外光谱的大豆茎秆化学组分含量检测模型构建与应用[J].中国农业科学,2021,54 (5):887-900.
- [4] 吴鹏,宋海燕,杨威,等[J].食品工业科技,2020,41(22):227-231.
- [5] 耿锐.近红外光谱分析技术在油品分析中的应用[J].化工设计通讯,2020,46(9):34+69.
- [6] 杨桂玲,范文昌,郭成栓.不同来源陈皮药材的近红外光谱法分析[J].时珍国医国药,2020,31(4):870-871.
- [7] 岑忠用, 雷顺新, 雷蕾, 等. 近红外光谱法鉴别 6 种根茎类中药材 [J]. 华中农业大学学报, 2021, 6:1-7.
- [8] 刘敏轩,王赟文,韩建国.高粱籽粒中多酚类物质的傅立叶变换近红外光谱分析[J].分析化学,2009,37(9):1275-1280.
- [9] 邵春甫,李长文,王珊,等.近红外光谱应用于高粱中总淀粉、直链淀粉与支链淀粉的定量分析[J].粮油食品科技,2016, 24(2):60-64.
- [10] 王勇生,李洁,王博,等.基于近红外光谱扫描技术对高粱中粗脂肪、粗纤维、粗灰分含量的测定方法研究[J].中国粮油 学报,2020,35(3):181-185.
- [11] 严衍禄.近红外光谱分析基础与应用[M].北京:中国轻工业出版社,2005.
- [12] 甘莉,孙秀丽,金良王,等.NIRS 定量分析油菜种子含油量、蛋白质含量数学模型的创建[J].中国农业科学,2003,36 (12):1609-1613.
- [13] 祝诗平,王刚,杨飞,等.粉末样品颗粒大小对花椒挥发油近红外光谱定量预测的影响研究[J].光谱学与光谱分析, 2008,28(4):775-779.
- [14] Users' manual in OPUS spectroscopic software 6 version [Z]. Germany:Bruker Optik GmbH, 2008:18-19.
- [15] 方彦.近红外光谱测定玉米蛋白质含量校正模型的建立[J].光谱实验室,2011,28(3):1050-1053.
- [16] 褚小立,王艳斌,陆婉珍.近红外光谱定量校正模型的建立及应用[J].理化检验(化学分册),2008(8):796-800.