

范会平,杜朝炜,李真,等. 基于近红外光谱技术的小麦专用粉品质特性快速检测研究[J]. 轻工学报, 2025,40(2):51-60.

FAN H P, DU Z W, LI Z, et al. Rapid detection of wheat special flour quality characteristics based on near-infrared spectroscopy technology[J]. Journal of Light Industry, 2025, 40(2):51-60. DOI:10.12187/2025.02.006

基于近红外光谱技术的小麦专用粉品质特性 快速检测研究

范会平1,2,杜朝炜1,2,李真1,2,杨勇1,2,任广跃3,张德榜4,艾志录1,2

1. 河南农业大学 食品科学技术学院,河南 郑州 450002;

2. 农业农村部 大宗粮食加工重点实验室,河南 郑州 450002;

3. 河南科技大学 食品与生物工程学院,河南 洛阳 471023;

4. 郑州万谷机械股份有限公司,河南 荥阳 450041

摘要:基于近红外光谱技术,结合不同预处理和特征波长筛选方法,构建小麦专用粉的破损淀粉含量、降落数值、吸水率、稳定时间、拉伸面积、延伸度和最大拉伸阻力的偏最小二乘(Partial Least Squares, PLS)预测模型和总体预测模型,并对模型的预测能力进行评估。结果表明:去线性趋势(Detrend, DT)是破损淀粉含量和吸水率预测模型的最佳预处理方法,Savitzky-Gloay(SG)卷积平滑是降落数值和拉伸面积预测模型的最佳预处理方法,标准正态变量变换(Standard Normal Variable Transformation, SNV)是延伸度和最大拉伸阻力预测模型的最佳预处理方法。竞争性自适应重加权法(Competitive Adaptive Reweighted Sampling, CARS)可有效提高破损淀粉含量、降落数值、吸水率、拉伸面积和最大拉伸阻力预测模型的预测精度,预测决定系数分别为0.9641、0.7140、0.9755、0.9434和0.8283;连续投影算法(Successive Projections Algorithm, SPA)可有效提高稳定时间和延伸度预测模型的效果,预测决定系数分别为0.7135和0.9530。总体预测模型对稳定时间、拉伸面积和最大拉伸阻力的预测效果均有所提升,剩余预测偏差(Residual Predictive Deviation, RPD)分别从1.86、4.27和2.51提升到2.43、5.26和3.11。综上可知,近红外光谱技术对小麦专用粉品质特性的无损快速检测是有效的、可行的。

关键词:小麦专用粉;近红外光谱;品质特性;偏最小二乘;快速检测 中图分类号:TS211.4 文献标识码:A 文章编号:2096-1553(2025)02-0051-10

0 引言

随着生活水平的不断提高,人们对面食种类和 品质提出了更高要求,对各类小麦专用粉的生产需 求也随之显著增加^[1]。配粉是小麦专用粉生产过 程中的关键环节,有利于保持产品品质的均衡稳 定,且能配制出多种用途的小麦专用粉。粉流在线 配粉法是制粉企业常用的一种配粉方法,该方法在

收稿日期:2024-07-19;修回日期:2024-11-06;出版日期:2025-04-15

基金项目:河南省重大科技专项项目(221100110800)

作者简介:范会平(1972—),女,河南省新密市人,河南农业大学教授,博士,主要研究方向为粮油精深加工、薯类主食化。 E-mail:fanhuiping1972@ hotmail.com

通信作者:艾志录(1965—),男,河南省新乡市人,河南农业大学教授,博士,主要研究方向为农产品加工与品质控制。E-mail: zhilafood@ sina. com

小麦专用粉生产过程中,依据各粉流的品质差异, 将不同粉路的小麦粉进行混合,以生产出一种或多 种小麦专用粉^[2]。因此,快速、准确地掌握不同粉 路小麦粉的品质对于实现精准配粉至关重要。降 落数值和吸水率、稳定时间、拉伸面积、延伸度、最 大拉伸阻力等粉质拉伸特性指标是小麦专用粉的 重要评价指标,此外,破损淀粉含量对面团的流变 特性和发酵特性也有一定影响,会影响焙烤、蒸煮 食品的品质,故在小麦专用粉生产过程中需不断调 整各粉路小麦粉的流量以确保终产品品质的稳定 性^[3-4]。然而,这些指标的测定都较为繁琐,尤其是 粉质拉伸特性,不仅需要不断地尝试并调整加水 量,还需要很长的醒面时间,耗时费力,严重制约了 小麦专用粉在线配粉的实时调控。

近红外光谱技术不仅可以反映 O—H、C—H 等 化学键的分子振动信息和吸收状态,还具有较强的 穿透性,适用于液体、固体、粉末等多种形态样品的 定量和定性检测分析^[5-6]。利用该技术可以快速检 测苹果酒中的挥发性化合物、速溶绿茶粉中的咖啡 $[D^{[7-9]}$ 及烟叶中的叶黄素和 β -胡萝卜素。在小麦粉 品质特性的检测方面,近红外光谱技术也取得了较 好的效果。孙晓荣等[10] 通过近红外光谱技术构建 了小麦粉水分、灰分和面筋含量的快速检测模型, 预测效果较理想。肖润君^[11]采用近红外光谱技术 对小麦全麦粉中的蛋白组分进行了定量预测,其中 醇溶蛋白和谷蛋白的预测模型效果较好,可用于实 际定量分析。W. Z. Zhou 等^[12]将近红外光谱技术 与化学计量学相结合,建立了小麦粉中湿面筋、沉 降值等指标的定量测定方法,准确率均在90%以 上。然而,这些研究主要集中在水分、灰分、粗蛋白 和面筋等组分含量的检测,对降落数值、吸水率、稳 定时间、拉伸面积、延伸度、最大拉伸阻力等反映小 麦专用粉品质特性指标的快速检测研究仍相对较 少。因此,在小麦专用粉的生产过程中,仍依赖于 传统的离线分析法监测其品质,难以实现真正的在 线实时监测。

鉴于此,本文拟基于近红外光谱技术,结合不 同预处理和特征波长筛选方法,建立小麦专用粉的 破损淀粉含量、降落数值、吸水率、稳定时间、拉伸 面积、延伸度和最大拉伸阻力等品质特性的偏最小 二乘(Partial Least Squares, PLS)预测模型和总体预 测模型,并评估其预测能力,以期为小麦专用粉生产 过程中加工参数和配粉方案的实时调控提供参考。

1 材料与方法

1.1 主要材料与试剂

小麦粉样品,分别取自中粮面业(漯河)有限公 司、中粮(郑州)粮油工业有限公司、想念食品股份 有限公司等小麦粉生产企业不同粉路的小麦粉在 制品和终产品;硼酸、氯化钠,国药集团化学试剂有 限公司;碘化钾,上海麦克林生化科技有限公司;硫 代硫酸钠,郑州派尼化学试剂厂。所有化学试剂均 为分析纯。

1.2 主要仪器与设备

2600XT-3 型近红外谷物分析仪,法国 Chopin 公司;Y41 型破损淀粉分析仪,土耳其 Yucebas 公 司;FN-II 型降落值测定仪,杭州大成科学仪器有限 公司;Farinograph-E 型粉质仪、Extensograph-E 型电 子拉伸仪,德国 Brabender 公司。

1.3 实验方法

1.3.1 小麦粉指标测定 破损淀粉含量参考《粮 油检验 小麦粉损伤淀粉测定 安培计法》(GB/T 31577—2015)的方法^[13],使用破损淀粉分析仪进行 测定;降落数值参考《小麦、黑麦及其面粉,杜伦麦 及其粗粒粉降落数值的测定 Hagberg-Perten 法》 (GB/T 10361—2008)的方法^[14],使用降落值测定仪 进行测定;粉质特性和拉伸特性分别参考《粮油检 验 小麦粉面团流变学特性测试 粉质仪法》(GB/T 14614—2019)和《粮油检验 小麦粉面团流变学特性 测试 拉伸仪法》(GB/T 14615—2019)的方法^[15-16], 使用粉质仪和电子拉伸仪进行测定。

1.3.2 近红外光谱采集 将小麦粉样品充分混匀, 使用近红外谷物分析仪采集近红外光谱。采用漫 反射模式,将小麦粉样品放在采集窗口上,利用近 红外光源照射样品,反射光被检测器收集,生成光 谱数据。每份样品采集3次,取平均光谱作为实验 光谱。采集的光谱波长范围为680~2600 nm,分辨 率为1 nm。

1.4 数据处理

1)异常样本剔除。使用主成分分析(Principal Component Analysis, PCA)结合马氏距离(Mahalanobis Distance, MD)的方法^[17]对采集的小麦粉光谱数据 进行异常值测定,以 5% 为阈值, 剔除距离过大的 样本。

2)样本集划分。分别采用随机划分(Random Selection, RS)法、Kennard Stone 划分(KS)法和光 谱-理化值共生距离划分(Sample Set Partitioning Based on Joint X-Y Distance, SPXY)法对样本集进行 划分,分析不同划分方法对模型的影响。

3)光谱数据预处理。为减少无用信息的干扰, 突出有效信息,采用去线性趋势(Detrend,DT)、多 元散射校正(Multiplicative Scatter Correction,MSC)、 标准正态变量变换(Standard Normal Variable Transformation,SNV)、归一化(Normalize,NOR)和 Savitzky-Golay(SG)卷积平滑5种预处理方法对光 谱数据进行处理。通过对比各预处理方法的建模 结果,选出最佳预处理方法。

4)特征波长筛选。所采集的小麦粉近红外光 谱共有 1921 个波长,存在大量的冗余信息和多重共 线性信息。通过筛选特征波长代替全光谱波长进 行建模,可以有效地简化模型,提高模型的稳定性 和计算效率^[18]。因此,在对光谱数据预处理后,采 用连续投影(Successive Projection Algorithm, SPA) 法、竞争性自适应重加权(Competitive Adaptive reweighted Sampling, CARS)法和无信息变量消除(Uninformative Variable Elimination, UVE)法筛选特征 波长。

5) 预测模型构建。利用 PLS 对处理后的光谱 数据,建立预测模型。

6) 预测模型评价。通过校正决定系数 (Determination Coefficient of Calibration, R_e^2)、预测决 定系数(Determination Coefficient of Prediction, R_p^2)、 校正均方根误差(Root Mean Squared Error of Calibration, *RMSEC*)和预测均方根误差(Root Mean Squared Error of Prediction, *RMSEP*)评估预测模型 性能。决定系数取值范围 0~1,数值越接近 1 表示 预测模型对数据的拟合程度越好;均方根误差表示 预测值与真实值之间的平均偏差程度,数值越小表 示模型的预测结果越准确^[19]。剩余预测偏差 (Residual Predictive Deviation, *RPD*)是预测集数据 的标准差与*RMSEP*的比值,也是用来评价预测模 型性能的常用指标。通常*RPD*在1.6~2.0之间表 示模型具有一定的预测能力,但需进一步优化;*RPD* 在2.0~2.5之间表示模型可以达到粗估的效果; *RPD*超过2.5时表示模型的预测性能良好^[20]。

7) 总体预测模型建立。对小麦粉样品的破损 淀粉含量、降落数值、吸水率等 7 项指标进行 PCA 分析,选择累计方差解释程度超过 90% 的主成分作 为输出建立 PLS 预测模型。

2 结果与分析

2.1 小麦粉品质指标测定结果分析

对小麦粉样品的破损淀粉含量、降落数值、吸 水率等品质指标进行测定,结果见表1。由表1可 知,各指标取值范围较广,基本可以涵盖各种小麦 专用粉的品质特性。

2.2 异常样本剔除分析

在采集近红外光谱时,由于外界环境的不稳 定,以及仪器和样品批次差异等原因,可能导致光 谱数据偏离正常的分布。MD 是一种有效的度量方 法,用于确定一个点相对于一个分布的距离。它考 虑了数据中各个变量的协方差,能够在空间中更好 地反映数据点之间的真实距离^[21]。但直接在原始 特征空间中使用 MD,可能对不具备明显线性关系 或较为复杂的高维数据不够敏感。先利用 PCA 对 数据进行降维,在降维后的空间中计算每个点的 MD,可以更准确地反映异常信息^[22]。图 1 为小麦

表1 小麦粉样品品质指标测定结果 Table 1 Results of quality indicators

for wheat flour samples

		_		
指标	最大值	最小值	平均值	标准差
破损淀粉含量/UCD _c	32.17	13.01	23.55	3.81
降落数值/s	896	333	539	72
吸水率/%	84.2	54.5	63.9	6.2
稳定时间/min	49.80	2.45	15.90	10.24
拉伸面积/cm ²	236	22	84	42
延伸度/mm	263	45	128	39
最大拉伸阻力/BU	1060	226	468	123



图 1 小麦粉样品的 MD 分布图 Fig. 1 MD distribution of wheat flour samples

粉样品的 MD 分布图。由图 1 可知,从用于破损淀 粉含量、降落数值、吸水率和稳定时间建模的 211 份 小麦样品中剔除了 11 个异常样本,从用于拉伸面 积、延伸度和最大拉伸阻力建模的 161 份小麦样品 中剔除了 8 个异常样本。

2.3 样本集划分结果分析

由于数据分布的差异,不同样本集划分方法对 建模效果具有一定的影响。选择合适的样本集划 分方法能够提升模型稳定性和预测精度。RS 法将 样本集随机划分为校正集和预测集;KS 法基于欧几 里得距离选择样本,以确保所选样本可以覆盖整个 特征空间;SPXY 法则进一步考虑化学值,计算样本 在 X 和 Y 空间的联合距离^[23]。本研究分别使用上 述3种划分方法,按照4:1的比例将样本集划分为 校正集和预测集,并对每个指标构建 PLS 回归模 型,以比较不同划分方法的建模效果,结果见表 2。 由表2可知,对于破损淀粉含量、稳定时间、延伸度 和最大拉伸阻力这 4 项指标,采用 SPXY 法划分的 效果最好;破损淀粉和稳定时间经 SPXY 划分后建 立的模型 $R_{\rm p}^2$ 和 RMSEP 均优于其他两种划分方法; 延伸度和最大拉伸阻力经3种方法划分后的R_n²相 差不大,但 SPXY 法划分后的 RMSEP 最小;吸水率 和拉伸面积使用 RS 法划分后建立的模型 R_{0}^{2} 最大, RMSEP 最小,模型效果最好;而降落数值则更适合 使用 KS 法进行划分。

2.4 光谱数据预处理对模型预测能力的影响 分析

为了减少噪声和冗余信息的干扰,本研究采用 DT、SNV、NOR、MSC和SG卷积平滑这5种预处理 表 2 不同样本集划分方法的建模结果

Table 2 Modeling results of different sate	nple
--	------

set partitioning methods

指标	划分方法	$R_{ m c}^2$	RMSEC	$R_{ m p}^2$	RMSEP
	RS	0.908 6	0.988 2	0.8557	1.3493
破顶淀粉 含量/UCD。	KS	0.9606	0.8434	0.8663	1.134 3
ц <u>т</u> , е о (SPXY	0.9599	0.8545	0.8947	0.9809
	RS	0.5674	47.8801	0.276 3	57.9949
降洛致值/	KS	0.6963	38.7601	0.527 8	49.5691
~	SPXY	0.6206	46.6606	0.178 0	47.9254
	RS	0.9877	0.6597	0.9667	1.122 4
吸水率/ %	KS	0.989 2	0.635 1	0.9306	1.414 6
10	SPXY	0.988 6	0.6619	0.884 9	1.1966
<i>(</i> (), (), (), (), (), (), (), (), (), (),	RS	0.616 2	6.3868	0.471 3	7.318 6
稳定时间/ min	KS	0.661 0	5.608 4	0.558 0	7.7184
	SPXY	0.688 9	5.6195	0.601 0	6.954 5
护伸云和了	RS	0.934 9	10.366 9	0.9301	12.5369
拉伸面积∕ cm ²	KS	0.937 0	12.159 5	0.9264	13.327 8
em	SPXY	0.938 6	11.602 2	0.901 0	12.9934
	RS	0.953 8	8. 1835	0. 9139	12.265 8
₩仲度/ cm	KS	0.984 8	4.6608	0. 9285	11.3614
CIII	SPXY	0.960 5	8.0121	0.9160	9.822 0
	RS	0.7317	59.025 5	0.6190	62.777 6
菆 大拉伸 阻力/Β□	KS	0.792 1	59.842 9	0.6017	62.831 6
,	SPXY	0.768 6	50.369 2	0.6098	53.687 6

方法对光谱数据进行处理,图 2 为小麦粉原始近红 外光谱图及经不同预处理后的近红外光谱图。DT 可以有效消除原始近红外光谱数据在采集时产生 的偏移,SNV 和 MSC 主要用于消除粉末样品采集时 因颗粒分布不均匀产生的影响^[24]。由图 2 可知,经 过 DT、SNV 和 MSC 预处理后的近红外光谱在 1400 nm 和 1800 nm 处的特征峰更加明显(见图 2b)、c)和 e));NOR 可以将近红外光谱数据缩放至 一个标准范围内,消除不同特征之间由于量纲不同 而产生的影响(见图 2d));SG 卷积平滑可以使近红 外光谱更加平滑,减少噪声的影响,但同时也会削 弱近红外光谱的整体特征(见图 2f))^[24]。 分别对原始近红外光谱及经 5 种预处理后的近 红外光谱建立全波段的 PLS 回归模型,建模结果见 表 3。



图 2 小麦粉原始近红外光谱图及经不同预处理后的近红外光谱图

Fig. 2 Raw near-infrared spectras of wheat flour and near-infrared spectras after different pretreatments

	表3	不同预处理方法的 PLS 模型结果
--	----	-------------------

Table 3	Results	of PLS	models	with	different	pretreatment	methods
						1	

							-				
指标	预处理方法	$R_{ m c}^2$	RMSEC	$R_{ m p}^2$	RMSEP	指标	预处理方法	$R_{ m c}^2$	RMSEC	$R_{ m p}^2$	RMSEP
	未处理	0.9599	0.854 5	0.8947	0.9809		NOR	0.626 8	6.154 6	0.5134	7.6796
	DT	0.957 1	0.8097	0.9409	0.816 2		MSC	0.6611	5.865 2	0.5717	7.2106
7世中 2011年1月	SNV	0.9508	0.8671	0.902 8	1.047 3		SG 卷积	0 672 2	5 968 1	0 598 8	6 973 2
版顶徙初 含量/UCDc	NOR	0.925 1	1.0703	0.8924	1.1019		半滑	0.072 2	5. 700 1	0. 570 0	0. 715 2
	MSC	0.982 6	0.8156	0.9257	0.9156		未处理	0.9349	10.366 9	0.930 1	12.5369
	SG 卷积 平滑	0.962 5	0.8574	0.914 0	0.984 8	的估子在了	DT SNV	0. 923 0 0. 922 4	11. 076 9 11. 123 5	0. 926 9 0. 910 3	13. 293 7 14. 729 0
	未处理	0.6963	38.7601	0.527 8	49.5691	拉伸面积∕ cm ²	NOR	0.906 2	12.225 1	0.9327	12.753 1
降落数值/ s	DT	0.6948	50. 521 4	0.5628	48.5900	Cim	MSC	0.922 5	11.112 2	0.9113	14.6439
	SNV	0.6687	51.808 1	0. 545 9	48.607 1	延伸度/ cm	SG 卷积	0.9351	10.868 0	0.9401	11.972 5
	NOR	0.641 1	53. 139 9	0.558 6	47.924 6		<u></u> 未か理	0.960.5	8 012 1	0.916.0	9 822 0
	MSC	0.666 2	53.931 3	0.544 6	48.677 2		不定理 DT	0.942.2	8 724 6	0.908.4	9 694 9
	SG 卷积 亚漫	0. 710 9	37.007 1	0.5636	47.6533		SNV	0. 946 8	8. 298 4	0. 925 2	8. 763 4
吸水率/ %		0 987 7	0 659 7	0 966 7	1 122 4		NOR	0.947 0	9. 279 4	0.923 1	8.8832
	DT	0.990 5	0.580.3	0.972.8	1. 014 3		MSC	0.9469	8. 595 9	0.925 1	8.9676
	SNV	0.978 0	0.882 6	0. 959 4	1. 240 2		SG 卷积 平滑	0.954 2	8.6218	0.9077	9.7314
	NOR	0.974 9	0.9427	0.958 0	1.2604		未处理	0.768 6	50.369 2	0.609 8	53.6876
	MSC	0.9859	0.781 9	0.970 0	1.064 6		DT	0.784 2	44.4662	0.6401	50.643 5
	SG 卷积 平滑	0.980 5	0.8311	0.9691	1.0808	最大拉伸	SNV	0.763 3	38.017 3	0.7191	49.245 9
	未处理	0.688.9	5, 619 5	0.601.0	6,954.5	阻力/BU	NOR	0.6974	51.741 4	0.6857	55.9807
稳定时间/	DT	0.648 1	5, 976 9	0.5697	7. 221 9		MSC	0.762 9	42.156 9	0.714 3	52.808 6
min	SNV	0.661 3	5.863 5	0. 572 3	7. 199 8		SG 卷积 平滑	0.7393	48.3192	0.7217	52.408 2

由表3可知,对于破损淀粉含量和吸水率,采用 DT预处理后建立的模型效果最佳,*RMSEP*分别减 小了16.76%和9.63%;对于降落数值和拉伸面 积,经SG卷积平滑预处理后建立的预测模型的 R_p^2 最高,*RMSEP*最小,预测效果最好;对于延伸 度和最大拉伸阻力,采用SNV预处理后建立的模 型效果最佳,虽然延伸度在校正集上的表现略有 降低,但在预测集上的表现较好,说明模型泛化 能力有所提高;稳定时间则是使用原始近红外光 谱建模的效果最好,这可能是因为稳定时间所对 应的特征数据在原始近红外光谱中已足够突出, 预处理反而会使这些特征的对比度降低,影响建 模效果。

2.5 特征波长筛选及预测模型构建结果

对预处理后的光谱进行特征波长筛选,可以有 效减少冗余信息对模型性能和稳定性的影响^[25]。 SPA 是通过比较选择不同数量波长时对应的 *RMSEC* 大小来确定特征波长的数量^[25]。CARS 根 据每个波长对预测结果的贡献程度来调整波长权 重,通过迭代逐步淘汰权重较小的波长,保留对模 型贡献最大的波长^[26]。UVE 通过添加一组统计特 性与实际特征相似但与响应变量无关的噪声,再以 噪声变量为性能基线,评估每一个特征的重要性, 进而移除重要性较低的变量^[26]。

在对光谱进行预处理和特征波长筛选后,对 各指标分别建立 PLS 回归预测模型。建模结果和 预测结果分别见表 4 和图 3。由表 4 可知,与全光 谱建模相比,各指标经过筛选特征波长后获得的 最优模型的预测效果均有所提高。其中,破损淀 粉含量、降落数值、吸水率、拉伸面积和最大拉伸 阻力使用 CARS 筛选特征波长后建立的模型效果 最好;与全光谱建模相比,上述指标经特征波长筛 选后所建立模型的 R²_p均有所提高,*RMSEP* 分别降 低了 22.1%、21.2%、5.1%、2.27%和16.4%,*RPD* 分别为 5.34、1.95、6.47、4.27 和 2.51。对于稳定 时间和延伸度,经 SPA 筛选波长后建立的模型效果 较好;延伸度校正集和预测集的 R² 均大于 0.95, *RPD* 为 4.69;稳定时间校正集和预测集的 R² 均大 于 0.70,*RPD* 为 1.86。由图 3 可知,破损淀粉、吸 水率、拉伸面积和延伸度(分别见图 3a)、c)、e)和 f))的预测模型对数据的拟合程度均较好,预测偏 差较小。

本研究构建的吸水率 CARS-PLS 预测模型的 RMSEP为 0.963 0, 低于陈嘉伟等^[27]的研究结果 (1.876),预测效果更好,但稳定时间 SPA-PLS 预测 模型的 RMSEP 为 6.001 1,高于其报道的 2.459,这 可能是由其研究中使用高斯回归对样本进行模糊 分类后建立的 GPR-PLSR 模型对线性程度较低指标 有更好的预测效果所致。J. C. Chen 等^[28]使用高光 谱对小麦籽粒建立了降落数值的 PLSR 预测模型, 相关系数 R 为 0.72,平均绝对误差为 56,本研究结 果与之相近。N. Gerisch 等^[29]利用近红外光谱结合 PLS 构建了大麦降落数值的预测模型. RMSEP 为 59.95,高于本研究结果的 37.540 8,这可能是因为 报道中所用近红外光谱波长区间为950~1650 nm, 每2 nm 采集一个波长点,波长区间和波长点数量 较少,忽略了一些与降落数值相关的有效信息,导 致 RMSEP 较高。

降落数值的特征波长主要分布于 700 nm、 1100 nm 和 1800 nm 附近。对于小麦粉破损淀粉 含量,筛选的特征波长主要分布在 900 nm、1800 nm 和 2100~2500 nm 附近,这些波长包含了 C-H 键、C---O键和 C---C键伸缩振动的一级倍频及 0—H键伸缩振动的一级和二级倍频,与小麦粉中 淀粉的特征波长相对应^[29]。吸水率的特征波长主 要集中在 2300 nm 附近,包含了 N-H 键、O-H 键、C==O键和 C---C键的伸缩振动及 C----H键的 变形振动,对应氨基酸、淀粉等物质[30],且小麦 粉的吸水率与其破损淀粉和蛋白质含量均有显 著的正相关关系^[31]。稳定时间、拉伸面积、延伸 度和最大拉伸阻力的特征波长主要集中在 700~ 1200 nm、2000~2300 nm 和 2550 nm 附近,这些波长 包含了 N—H 键、C ==O 键的伸缩振动及 N—H 键 伸缩振动的三级倍频,对应氨基酸、蛋白质等物 质^[30]。小麦粉中蛋白质的含量和质量决定了面筋 的弹性、抗延伸性等特性,蛋白质含量较高的小麦 粉通常表现出更好的拉伸特性和更长的稳定 时间[32]。

7

Table 4 Company	on of the regression	prediction	mouels after	leature wavere	enguis select	eu by unieren	t memous
指标	特征波长选择方法	变量数	$R_{ m c}^2$	RMSEC	$R_{ m p}^2$	RMSEP	RPD
	全光谱	1921	0.957 1	0.8097	0.9409	0.816 2	4.17
□ 世 定 小 小 CD	SPA	14	0.903 3	1.2159	0.943 0	0.802 0	4.24
风坝徙初百里/UUDC	CARS	79	0.9657	0.664 1	0.964 1	0.6365	5.34
	UVE	221	0.933 3	1.010 3	0.934 0	0.8628	3.94
	全光谱	1921	0.7109	37.007 1	0.563 6	47.6533	1.53
陈菇粉店 /。	SPA	12	0.6402	48.1947	0.6428	43.081 5	1.70
降落数值/s	CARS	49	0.7207	36.403 8	0.714 0	37.5408	1.95
	UVE	334	0. 599 9	50.267 0	0.582 6	56.5992	1.29
	全光谱	1921	0.9905	0.5803	0.972 8	1.014 3	6.14
吸水率/%	SPA	27	0.9676	1.071 3	0.966 2	1.1304	5.51
	CARS	32	0.978 1	0.8808	0.975 5	0.963 0	6.47
	UVE	777	0.978 0	0.8819	0.9607	1.2191	5.11
	全光谱	1921	0.688 9	5.6195	0.601 0	6.954 5	1.60
<u> </u>	SPA	27	0.7202	5.3296	0.713 5	6.001 1	1.86
稳定时间/min	CARS	23	0.6753	5.741 4	0.682 4	6.2047	1.80
	UVE	530	0.582 2	6.5122	0.5871	6.884 8	1.62
	全光谱	1921	0.935 1	10.868 0	0.9401	11.972 5	4.18
治体云和 (2	SPA	15	0.911 2	11.8984	0.9337	12.661 3	3.95
拉伸曲你/ cm	CARS	16	0.9364	10.8290	0.943 4	11.7003	4.27
	UVE	606	0.9369	10.0300	0.929 5	13.053 9	3.83
	全光谱	1921	0.946 8	8.2984	0.925 2	8.763 4	3.72
延伸度/am	SPA	15	0.974 0	6.472 9	0.953 0	6.944 4	4.69
延仲度/ cm	CARS	28	0.964 5	7.5980	0.936 0	8.1063	4.02
	UVE	156	0.944 9	9.463 0	0.928 3	8.5787	3.80
	全光谱	1921	0.763 3	38.017 3	0.7191	49.245 9	2.10
是十拉伸阻力/DU	SPA	19	0.8347	49.732 6	0.805 5	46.8132	2.21
取八121甲阻力/ BU	CARS	32	0.843 2	43.3347	0.828 3	41.163 6	2.51
	UVE	204	0.8819	39.5002	0.6869	55.5810	1.86



Table 4 Comparison of PLS regression prediction models after feature wavelengths selected by different methods



Fig. 3 Prediction results of various indicator models

2.6 总体预测模型建立结果

由于原始指标数量较多、相互之间存在一定的 相关性,直接对这些指标进行预测可能会导致模型 复杂程度较高。通过 PCA 降维可以提取具有代表 性的少数几个主成分,进而简化模型^[33]。对 7 项指 标进行 PCA 分析,结果见图 4。选择累计方差解释 度达到 90%以上的前 4 个主成分作为输出,结合光 谱数据建立 PLS 预测模型。经光谱预处理和特征 波长筛选后,确定经 SNV 预处理和 CARS 筛选波长 后建立的 PLS 模型效果最佳,总体预测模型预测效 果见表 5。由表 5 可知,对于破损淀粉含量、降落数 值、吸水率和延伸度,总体预测模型的预测效果不 如独立的预测模型;但总体预测模型对稳定时间、拉 伸面积和最大拉伸阻力的预测效果均有所提升,*RPD* 分别由 1.86、4.27 和 2.51 提升至 2.43、5.26 和 3.11。







表5 总体预测模型预测效果

 Table 5
 Overall prediction model

 prediction performance

_								
	评价 指标	破损淀粉 含量	降落 数值	吸水率	稳定 时间	拉伸 面积	延伸度	最大拉 伸阻力
	$R_{ m p}^2$	0.842 1	0.6819	0.883 9	0.825 5	0.9627	0.931 1	0.893 0
ŀ	RMSEP	1.132 7	46.484 6	1.9789	4.9901	9.4939	11. 151 9	38.8034
	RPD	2.56	1.80	2.98	2.43	5.26	3.87	3.11

3 结论

本文利用近红外光谱技术,结合多种预处理和 特征波长筛选方法,构建了小麦专用粉的破损淀粉 含量、降落数值、吸水率、稳定时间、拉伸面积、延伸 度和最大拉伸阻力的 PLS 回归模型及总体预测模 型,并对其预测能力进行了评估。结果表明,破损 淀粉含量和吸水率预测模型的最佳预处理方法是 DT,延伸度和最大拉伸阻力更适合采用 SNV 进行预 处理,降落数值和拉伸面积更适合采用 SG 卷积平 滑进行预处理,而稳定时间的建模效果在不进行预 处理时更佳。对光谱进行预处理、特征波长筛选 后,各指标预测模型效果均有所提升。破损淀粉含 量、吸水率和拉伸面积的 CARS-PLS 预测模型和延 伸度的 SPA-PLS 预测模型的 R^2 均在 0.9 以上,最 大拉伸阻力的 CARS-PLS 预测模型的 R² 达到 0.8 以上,RPD均大于2.5,预测效果较理想。降落数值 的 CARS-PLS 预测模型和稳定时间的 SPA-PLS 预 测模型的 R² 均大于 0.7, RPD 分别为 1.95 和 1.86, 对指标有一定的预测能力。对这些指标建立总体 预测模型,其对稳定时间、拉伸面积和最大拉伸阻 力的预测能力均得到提升, RPD 分别由 1.86、4.27 和 2.51 提高至 2.43、5.26 和 3.11。本研究实现了 对小麦专用粉生产过程中各粉路小麦粉品质特性 的快速预测,在配粉环节可以快速掌握各粉路小麦 粉品质特性,提高小麦专用粉生产效率,但降落数 值和稳定时间预测模型的预测精度仍需进一步提 高,在后续研究中将对这两个指标的预测模型进行 优化,以提高模型的稳定性和准确性。

参考文献:

- [1] 刘鑫,王晓曦,史建芳. 配麦和配粉在专用粉生产中的应用[J]. 现代面粉工业,2010,24(1):1-4.
- 李林轩,王晓芳,李硕,等.专用小麦粉生产的基本方法、必备条件和关键环节[J].现代面粉工业,2021,35
 (5):5-8.
- [3] 张蓓,张剑,李梦琴,等.小麦粉破损淀粉含量对面团发 酵性能的影响[J].食品与机械,2019,35(12):70-74.
- [4] LI E P, DHITAL S, HASJIM J. Effects of grain milling on starch structures and flour/starch properties [J]. Starch-Stärke, 2014, 66(1/2):15-27.
- [5] 朱丽伟,严金欣,黄娟,等.近红外光谱和人工神经网络的金苦荞氨基酸快速测定[J].光谱学与光谱分析, 2022,42(1):49-55.
- [6] 纪红昌,邱晓臣,柳文浩,等.花生籽仁含油量近红外 模型的构建及其应用[J].中国油料作物学报,2022, 44(5):1089-1097.
- [7] YE M Q, GAO Z P, LI Z, et al. Rapid detection of volatile

• 59 •

compounds in apple wines using FT-NIR spectroscopy [J]. Food Chemistry, 2016, 190:701-708.

- [8] SINIJA V R, MISHRA H N. FT-NIR spectroscopy for caffeine estimation in instant green tea powder and granules [J]. LWT-Food Science and Technology, 2009, 42(5): 998-1002.
- [9] 何屹,杨本刚,尹嵩,等.基于随机蛙跳筛选的初烤烟
 叶中β-胡萝卜素和叶黄素含量近红外模型的建立
 [J].轻工学报,2024,39(2):100-106.
- [10] 孙晓荣,张晨光,刘翠玲,等.近红外光谱技术对小麦粉品质定量快速检测[J].食品科技,2023,48(11):246-252.
- [11] 肖润君.基于近红外光谱的小麦主要品质指标速测技术研究[D].保定:河北农业大学,2022.
- [12] ZHOU W Z, LEI Y Q, ZHOU Q D, et al. A rapid determination of wheat flours components based on near infrared spectroscopy and chemometrics [J]. Vibrational Spectroscopy, 2024, 130:103650.
- [13] 国家质量监督检验检疫总局,中国国家标准化管理委员会.粮油检验 小麦粉损伤淀粉测定 安培计法: GB/T 31577—2015[S].北京:中国标准出版社,2015.
- [14] 国家质量监督检验检疫总局.小麦、黑麦及其面粉,杜 伦麦及其粗粒粉降落数值的测定 Hagberg-Perten 法: GB/T 10361—2008[S].北京:中国标准出版社,2008.
- [15] 国家市场监督管理总局,国家标准化管理委员会.粮 油检验 小麦粉面团流变学特性测试 粉质仪法: GB/T 14614—2019[S].北京:中国标准出版社,2019.
- [16] 国家市场监督管理总局,国家标准化管理委员会.粮 油检验 小麦粉面团流变学特性测试 拉伸仪法: GB/T 14615—2019 [S]. 北京:中国标准出版 社,2019.
- [17] 陈斌, 邹贤勇, 朱文静. PCA 结合马氏距离法剔除近红 外异常样品[J]. 江苏大学学报(自然科学版), 2008, 29(4):277-279, 292.
- [18] 张旭,白雪冰,汪学沛,等.近红外特征光谱的羊肉 TVB-N浓度预测模型[J].光谱学与光谱分析,2021, 41(11):3377-3384.
- [19] TANG C C, JIANG B Z, EJAZ I, et al. High-throughput phenotyping of nutritional quality components in sweet potato roots by near-infrared spectroscopy and chemometrics methods[J]. Food Chemistry(X), 2023, 20:100916.
- [20] KOVALENKO I V, RIPPKE G R, HURBURGH C R. Determination of amino acid composition of soybeans

(*Glycine* max) by near-infrared spectroscopy [J]. Journal of Agricultural and Food Chemistry, 2006, 54(10): 3485–3491.

- [21] 刘翠玲, 胡玉君, 吴胜男, 等. 近红外光谱奇异样本剔除 方法研究[J]. 食品科学技术学报, 2014, 32(5): 74-79.
- [22] 宋可可. 基于近红外光谱的稳健建模技术研究[D]. 青岛:青岛科技大学,2023.
- [23] 张乐,吴静珠,李江波,等.基于随机森林的单粒玉米 种子水分近红外快速定量检测[J].中国粮油学报, 2021,36(12):114-119.
- [24] 袁媛.基于近红外光谱技术的面粉品质无损检测研究
 [D].长沙:湖南农业大学,2023.
- [25] 李江波,郭志明,黄文倩,等.应用 CARS 和 SPA 算法 对草莓 SSC 含量 NIR 光谱预测模型中变量及样本筛 选[J].光谱学与光谱分析,2015,35(2):372-378.
- [26] LIN H J, YING Y B. Theory and application of near infrared spectroscopy in assessment of fruit quality: A review[J]. Sensing and Instrumentation for Food Quality and Safety, 2009, 3(2):130-141.
- [27] 陈嘉伟,周德强,崔晨昊,等.近红外光谱的小麦粉粉 质特性预测模型研究[J].光谱学与光谱分析,2023,43(10):3089-3097.
- [28] CHEN J C, HU Y, LI X, et al. An independent validation reveals the potential to predict Hagberg-Perten falling number using spectrometers [J]. The Plant Phenome Journal, 2023, 6(1):e20070.
- [29] GERISCH N, GUERREIRO R, WESPEL F, et al. Development of a near-infrared spectroscopy calibration for Hagberg falling number assessment of barley (*Hordeum vulgare*): A comparison of methods [J]. Plant Breeding, 2022,141(3):355-365.
- [30] 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,2005:31-39.
- [31] SEDLÁČEK T, ŠVEC I. Water absorption of food wheat flour-comparison of direct and indirect methods [J]. Croatian Journal of Food Science and Technology, 2023, 15(2):163-172.
- [32] 涂金金,高淑娟,苏煜娴,等. 面筋蛋白在发酵面团加 工中的作用机制及研究进展[J]. 中国粮油学报, 2024,39(4):210-222.
- [33] RUIZ C, ALAÍZ C M, DORRONSORO J R. A survey on kernel-based multi-task learning [J]. Neurocomputing, 2024,577:127255.

Rapid detection of wheat special flour quality characteristics based on near-infrared spectroscopy technology

FAN Huiping^{1,2}, DU Zhaowei^{1,2}, LI Zhen^{1,2}, YANG Yong^{1,2}, REN Guangyue³, ZHANG Debang⁴, AI Zhilu^{1,2}

1. College of Food Science and Technology, Henan Agricultural University, Zhengzhou 450002, China;

2. Key Laboratory of Staple Grain Processing, Ministry of Agriculture and Rural Affairs, Zhengzhou 450002, China;

detection

College of Food and Biological Engineering, Henan University of Science and Technology, Luoyang 471023, China; A. Zhengzhou Wangu Machinery Co., Ltd., Xingyang 450041, China

Abstract: Based on near-infrared spectroscopy technology, combined with different preprocessing and characteristic wavelength screening methods, partial least squares (PLS) prediction models and overall prediction model were established for indicators such as damaged starch content, falling number, water absorption rate, stability time, stretching area, extensibility and maximum resistance. The results showed that detrend (DT) was the best preprocessing method for the prediction model of damaged starch content and water absorption rate, savitzky-gloay (SG) convolutional smoothing was the best preprocessing method for the prediction model of falling number and stretching area, and standard normal variable transformation (SNV) was the best preprocessing method for the prediction model of extensibility and maximum resistance. Competitive adaptive reweighted sampling (CARS) could effectively improve the prediction accuracy of models for damaged starch content, falling number, water absorption rate, stretching area and maximum resistance, with prediction determination coefficients of 0.964 1, 0.714 0, 0.975 5, 0.943 4 and 0.828 3, respectively, successive projections algorithm (SPA) had improved the performance of stability time and extensibility prediction models, with prediction determination coefficients of 0.713 5 and 0.953 0, respectively. The overall prediction model had improved its predictive performance for stability time, stretching area and maximum resistance, their residual predictive deviation increased from 1.86, 4. 27 and 2. 51 to 2. 43, 5. 26 and 3. 11, respectively. In summary, near-infrared spectroscopy technology was effective and feasible for a non-destructive and rapid detection of the quality characteristics of wheat flour. Key words; wheat special flour; near-infrared spectroscopy; quality characteristics partial least squares; rapid

[责任编辑:王晓波]

本刊数字网络传播声明

本刊已许可中国知网,万方数据资源系统,维普网,国家科技学术期刊开放平台,博看网, 超星,中国科技论文在线,中教数据库,长江文库,CABI,CA,AJ. VINITI,EBSCOhost,Ulrichs, FSTA等在其系列数据库产品中以数字化方式复制、汇编、发行、信息网络传播本刊全文。其相 关著作权使用费与本刊稿酬一并支付。作者向本刊提交文章发表的行为即视为同意我刊上述 声明。