戴炼, 宋光均, 钟武杰, 等. 基于近红外光谱的类胡萝卜素发酵过程总糖与总氮快速定量监测 [J]. 食品工业科技, 2026, 47(15): 337−344. doi: 10.13386/j.issn1002-0306.2025080068

DAI Lian, SONG Guangjun, ZHONG Wujie, et al. Rapidly Quantifying Total Sugar and Nitrogen Contents during Carotenoid Fermentation Processes Using Near-infrared Spectroscopy[J]. Science and Technology of Food Industry, 2026, 47(15): 337−344. (in Chinese with English abstract). doi: 10.13386/j.issn1002-0306.2025080068

· 分析检测 ·

基于近红外光谱的类胡萝卜素发酵过程总糖与总氮快速定量监测

戴　炼[1], 宋光均[2], 钟武杰[1], 丁　敏[1], 万语悦[1], 谢吉红[1], 蹇华丽[1,*]

（1.华南农业大学食品学院, 广东广州 510642;
2.广州标旗光电科技发展股份有限公司, 广东广州 510663）

摘　要: 为建立胶红酵母（*Rhodotorula mucilaginosa*，*R. mucilaginosa*）类胡萝卜素发酵过程中总糖和总氮的快速定量分析方法，本研究以新会陈皮柑果肉为原料，采用近红外（Near-infrared spectroscopy，NIR）光谱技术结合偏最小二乘回归（Partial least squares regression，PLSR），构建了发酵液中总糖（Total soluble sugar，TSS）和总氮（Total nitrogen，TN）的 NIR 定量模型。以预测集决定系数（Coefficient of determination for prediction set，R^2p）、预测均方根误差（Root mean square error of prediction，RMSEP）和相对分析误差（Relative percent deviation，RPD）等作为模型评价指标，系统地比较了不同光谱预处理方法及特征波数筛选对模型性能的影响，同时通过独立样本对模型进行外部验证，结果表明：采用多元散射校正（Multiplicative scatter correction，MSC）进行光谱预处理，并结合竞争性自适应重加权采样法（Competitive adaptive reweighted sampling，CARS）筛选特征波数，所建立的总糖及总氮定量模型效果最佳。其中 TSS 定量模型 R^2p 为 0.9670，RMSEP 为 2.6587，外部验证 RPD 值为 4.8386，可用于高精度定量分析。TN 定量模型 R^2p 和 RMSEP 分别为 0.9502 和 0.0624，外部验证 RPD 值为 2.7423，适用于常规定量分析。TSS 及 TN 模型的 t 检验结果均显示预测值与实测值之间无显著差异（P>0.05），具有较好的准确性和稳定性，可用于快速监测胶红酵母类胡萝卜素发酵液样品的 TSS 和 TN 含量。

关键词: 近红外光谱技术, 化学计量学, 胶红酵母, 总糖, 总氮, 定量分析模型

中图分类号: TS207.3　　**文献标识码:** A　　**文章编号:** 1002−0306(2026)15−0337−08

DOI: 10.13386/j.issn1002-0306.2025080068

本文网刊:

Rapidly Quantifying Total Sugar and Nitrogen Contents during Carotenoid Fermentation Processes Using Near-infrared Spectroscopy

DAI Lian[1], SONG Guangjun[2], ZHONG Wujie[1], DING Min[1], WAN Yuyue[1], XIE Jihong[1], JIAN Huali[1,*]

（1.College of Food Science, South China Agricultural University, Guangzhou 510642, China;
2.Guangzhou Biaoqi Optoelectronic Technology Development Co., Ltd., Guangzhou 510663, China）

Abstract: Xinhui dried tangerine peel fruit pulp was used as a raw material to establish a rapid and quantitative analytical method for monitoring the total soluble sugar (TSS) and nitrogen (TN) contents during carotenoid fermentation by *Rhodotorula mucilaginosa* (*R. mucilaginosa*). Near-infrared spectroscopy (NIR) was integrated with partial least squares regression (PLSR) to construct models for quantifying the TSS and TN in fermentation broth. The coefficient of determination for the prediction set (R^2p), root mean square error of prediction (RMSEP), and relative percent deviation (RPD) were used as evaluation metrics, and the effects of different spectral preprocessing methods and feature wavenumbers on model

收稿日期: 2025−08−11

基金项目: 广东省科技创新战略专项资金（重点领域研发计划项目）“发酵食品精准制造关键技术及产业化示范”（2020B020226008）。

作者简介: 戴炼（2001−）, 女, 硕士研究生, 研究方向: 食品发酵与酿造, E-mail: 1321201971@qq.com。

*** 通信作者:** 蹇华丽（1977−）, 女, 博士, 副教授, 研究方向: 食品发酵与酿造, E-mail: jianhualij@scau.edu.cn。

performance were systematically compared. The results were externally validated using independent samples. The optimal models for quantifying the TSS and TN contents were obtained using multiplicative scatter correction (MSC) for preprocessing the spectrum and competitive adaptive reweighting sampling (CARS) for selecting the feature wavenumber. The model for quantifying TSS achieved a R^2p of 0.9670, a RMSEP of 2.6587, and an external validation RPD of 4.8386, indicating high-precision quantitative analysis. The model for quantifying TN yielded R^2p and RMSEP values of 0.9502 and 0.0624, respectively, with an external validation RPD of 2.7423, which indicated suitability for routine quantitative analyses. The *t*-test results for the TSS and TN models indicated that the predicted and measured values did not significantly differ (P>0.05), indicating high accuracy and stability. These models can be used for rapidly monitoring the TSS and TN contents in *R. mucilaginosa* carotenoid fermentation broth samples.

Key words: near-infrared spectroscopy; chemometrics; *Rhodotorula mucilaginosa*; total soluble sugar; total nitrogen; quantitative analysis model

类胡萝卜素(Carotenoids)是一种脂溶性萜类化合物[1]。因其具有多种生物学功能[2−4],近年来应用越来越广泛且市场需求增长迅速。相较于传统来源,微生物发酵法生产类胡萝卜素具有产率高、周期短、安全环保等显著优势[5−6]。以新会陈皮柑加工副产物—果肉为原料,通过胶红酵母(*Rhodotorula mucilaginosa*, *R. mucilaginosa*)发酵可实现类胡萝卜素的高效转化,兼具废弃物资源化与低成本生产的双重价值[7−8]。然而,对于发酵过程中的关键参数—总糖(Total soluble sugar, TSS)和总氮(Total nitrogen, TN),目前普遍分别采用滴定法或分光光度法和凯氏定氮法作为检测方法[9−10],这些传统检测方法存在操作繁琐、检测速度慢、数据滞后等问题,难以满足基于碳、氮源含量变化进行高精度实时反馈控制补料的工艺需求。因此,亟需开发更高效、便捷、准确的发酵过程关键参数(TSS 和 TN)监测方法。

近红外光谱技术(Near-infrared spectroscopy, NIR)是一种基于分子振动能级跃迁产生的特征吸收光谱(来源于含氢基团 C-H、N-H、O-H 等内部振动的倍频与合频吸收)[11−12]。该光谱被广泛用于食品成分及发酵关键参数的分析[13−14]。但由于 NIRS 原始光谱的可解释性不强,通常需要借助化学计量学方法,采用统计学及计算机技术,设计和选择最优光谱测量程序和实验方案,并通过解释测量数据而获得最大限度的信息,建立校正模型,从而实现过程关键参数的定性或定量预测[15]。近年来,近红外光谱结合化学计量学方法的分析技术因其快速、无损等优势在多元混合体系物质测定中应用越来越广泛。王春媛等[16] 利用近红外光谱技术结合偏最小二乘回归,建立了发酵麦麸中还原糖和可溶性蛋白质含量的快速检测方法。郝超等[17] 利用概率偏最小二乘的多元统计过程监控方法与近红外光谱技术相结合来对柠檬酸发酵液化清液生产过程的总氮和总糖含量进行监控。Zhao 等[18] 利用近红外光谱法与偏最小二乘法、反向传播神经网络及其组合监测康普茶的总糖含量。Kozma 等[19] 利用 NIRS 研究了 CHO 细胞在不同发酵罐发酵体系下葡萄糖含量的变化规律,为流加式操作过程中葡萄糖的浓度提供快速监测手段。Li 等[20] 采用近红外光谱结合深度迁移学习(DTL),建立了用于实时监控柑橘果醋发酵过程中总糖和总氮含量的快速检测方法等。

鉴于类胡萝卜素发酵生产规模放大过程中补料工艺对发酵液中碳、氮源含量变化高精度实时监测的迫切需求,本文拟通过近红外光谱结合相应的化学计量学方法对胶红酵母-陈皮柑果肉发酵体系的 TSS 与 TN 进行监测,建立 TSS 与 TN 的快速定量模型,替代传统离线检测方法,为优化实时补料策略提供可靠数据支撑,从而达到提高类胡萝卜素产率、提升资源利用效率的目的,这对后续实现发酵过程的精确控制,类胡萝卜素工业化量产具有重要意义。

1 材料与方法

1.1 材料与仪器

新会陈皮柑果肉　江门市新会区锦胜果园提供;硫酸　分析纯,广州市梓行化玻仪器有限公司;DNS 试剂　分析纯,上海科丰化学试剂有限公司;盐酸、硫酸铵、乙酸锌、亚铁氰化钾、硼酸、甲基红、溴甲酚绿、氢氧化钠　均为分析纯,国药集团化学试剂有限公司;胶红酵母野生型菌株 GDMCC 2.30　购自广东省菌种保藏中心;本研究使用的突变株 JH-R23 由 GDMCC 2.30 经“新一代载人飞船试验船”空间搭载诱变、初筛、复筛及纯化所得,由华南农业大学食品学院发酵工程实验室保藏;种子液　挑取 JH-R23 单菌落接种于马铃薯葡萄糖液体培养基(PDB)中,于 28 ℃、180 r/min 条件下振荡培养 36 h 制得,培养基购自广东环凯微生物有限公司。

Antaris Ⅱ型傅里叶变换近红外光谱仪　配有投射分析模块及光程为 1 mm 玻璃比色皿,美国赛默飞公司;赛诺 KT8400 全自动凯氏定氮仪　福斯华科贸有限公司;VersaMax 光栅型酶标仪　美国 Molecular Devices 公司;HWS-24 型电热恒温水浴锅　上海恒一科学仪器有限公司;AL204 型电子天平　梅勒特-托利多上海有限公司;JYL-C19V 榨汁机　九阳股份有限公司;TGL-16gR 台式高速冷冻离心机　上海一恒科学仪器有限公司。

1.2 实验方法

1.2.1 样品的制备与处理　称取适量新鲜新会柑果

肉，使用榨汁机榨汁。所得果汁经双层滤布过滤后，于 5000 r/min 离心 10 min。收集上层澄清液，调节其 pH 至中性，随后在 121 ℃ 下高压灭菌 20 min 备用。将种子液以 4% 的接种量接入灭菌后的果汁，于 28 ℃ 发酵 5 d，每天取 8 个样品，将取得的发酵液经 10000 r/min 离心 10 min 后滤去菌体。共进行 5 批独立发酵，最终得到 205 个样品（含每批初始发酵液 1 个）。取其中的 185 个样品用于建模分析、20 个样品用于外部验证。

1.2.2　近红外光谱的采集　将仪器开机预热 30 min 后自检，将样品混匀并排出底部气泡后，转置于 1 mm 光程的比色皿中采集样品光谱数据。仪器工作条件：在室温环境（22±2 ℃）下，以水为参比，扫描范围为 4000~10000 cm^{-1}，扫描次数 32 次，分辨率为 8 cm^{-1}。原始光谱经仪器配套的 TQ Analyst 软件按 4 cm^{-1} 间隔插值处理后，每个样品的光谱包含 1557 个变量（波数点）。每个样品采集 3 次，取平均光谱用于数据分析。

1.2.3　理化数据的测定　总糖含量测定参考 NY/T 2742-2015《水果及制品可溶性糖的测定 3,5-二硝基水杨酸比色法》中方法。总氮含量测定参照 GB 5009.5-2016《食品中蛋白质的测定》第一法中的自动凯氏定氮法。

1.2.4　样品集划分　在建立定量分析模型前，采用马氏距离法剔除 185 个建模样本中的异常值[21]，再按照 3:1 的比例将剩余样本划分为训练集和预测集。训练集用于建立测定样品中 TSS 和 TN 的定量校正模型，预测集则用于验证所建立校正模型的预测性能。本研究采用基于 XY 联合距离的样本集划分方法（Sample set partitioning based on joint X-Y distance, SPXY）。该方法同时考虑了光谱变量（X）和理化值变量（Y）的信息[22−23]。

1.2.5　光谱预处理　在光谱采集过程中，仪器自身的响应漂移、环境温度与湿度的波动、样品光散射效应以及杂散光干扰等因素，均会引入噪声和基线漂移等非有效信息，导致降低所建定量校正模型的稳定性与准确性[24]。因此，在利用光谱数据建立可靠的定量模型之前，需对原始光谱进行适当的预处理。本研究考察了四种常见的预处理方法，包括卷积平滑（Savitzky-Golay smoothing, SG）、多元散射校正（Multiplicative scatter correction, MSC）、标准正态变量变换（Standard normal variate, SNV）、一阶导数（First derivative, FD）。其中，SG 用于平滑处理光谱，能够增加样本信号的信噪比；MSC 和 SNV 可以消除颗粒大小、分布不均匀及表面散射对光谱信号的影响；FD 旨在扣除背景漂移，提高光谱分辨率[25−26]。

1.2.6　特征波数的筛选　近红外光谱的全波段包含大量变量数，其中存在较多冗余信息。直接利用全波段建模会降低模型精度[27]。为此，本研究通过比较竞争性自适应全加重法（Competitive adaptive reweighted sampling, CARS）、无信息变量筛除法（Uninformative variable elimination, UVE）和随机蛙跳法（Random frog, RF）3 种方法特征筛选后的建模效果来确定最佳的波数筛选方法，从而提高模型的精确度并降低模型的复杂度。

1.2.7　模型的建立与评价　本研究采用偏最小二乘回归（Partial least squares regression, PLSR）建立胶红酵母类胡萝卜素发酵过程中关键参数（总糖与总氮）的定量预测模型。模型质量通过训练集决定系数（R^2c）、训练集均方根误差（RMSEC）参数评估：R^2c 越接近 1，表明近红外分析结果与化学测定值的吻合度越高，模型可靠性越强。模型预测性能则通过预测集决定系数（R^2p）和预测集均方根误差（RMSEP）评价：R^2p 越大且 RMSEP 越小，表明模型预测精度越高。同时，引入相对分析误差（Relative percent deviation, RPD）评估模型的实际预测能力：RPD<2.5 时，表明该模型预测能力不足，无法满足定量分析需求；2.5≤RPD<3.0 表明模型可用于常规定量分析；RPD>3.0 则表明模型具备高精度预测能力[28]。为进一步验证模型的可靠性，采用 20 个独立样本进行外部验证，并引入 t 检验对近红外预测值与理化实测值进行统计检验[29]。

1.3　数据处理

每个样品的理化指标均采集三次数据后取平均。采用 Excel 2019 进行数据处理与汇总，使用 OriginPro 2021 绘图，并利用 MATLAB R2021a 进行光谱预处理、特征波数筛选及建模分析。

2　结果与分析

2.1　样品集划分结果

类胡萝卜素发酵液的建模样本采用马氏距离法将距离阈值大于 4 的异常值剔除，结果如图 1 所示。本研究采用 SPXY 法将 178 个有效样本划分样品集。建立一个准确的模型关键在于训练集需广泛覆盖预测集的理化值范围[30]。如表 1 所示，总糖和总氮定量模型的预测集理化值均落在其对应训练集的取值范围内，因此该样品集划分方案合理，能够较好地满足模型构建的需求。

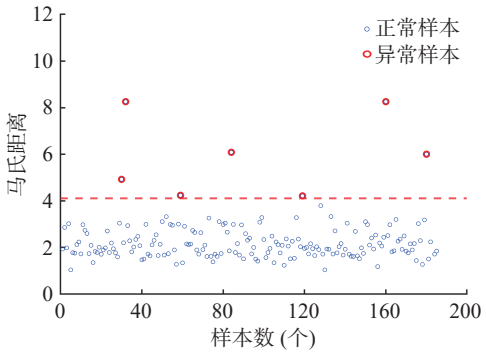


图 1　所有样本马氏距离

Fig.1　Mahalanobis distance for all samples

表 1 SPXY 法划分类胡萝卜素发酵液样品集结果

Table 1 Results of the partitioning of carotenoid fermentation broth samples by SPXY

指标	TSS		TN	
	训练集	预测集	训练集	预测集
样本量	133	45	133	45
最大值(mg/mL)	88.26	78.91	2.54	2.19
最小值(mg/mL)	13.09	16.79	0.75	1.03
平均值(mg/mL)	48.16	55.29	1.59	1.53
标准偏差	16.83	15.03	0.38	0.35

2.2 原始光谱和预处理结果

类胡萝卜素发酵液的原始光谱数据如图 2(a)所示。由于样品中总糖(TSS)和总氮(TN)含量存在差异,其吸光度值存在明显差异。样品的近红外光谱吸收主要集中在三个波段:7500~10000、5500~7500 和 4000~5500 cm^{-1}。其中,7500~10000 cm^{-1} 波段主要对应含氢基团伸缩振动的二级倍频及其合频吸收,但吸收强度相对较弱。不同样品间的光谱差异主要体现在 5500~7500 cm^{-1} 和 4000~5500 cm^{-1} 波段,这两个区域反映了含氢基团的一级倍频以及 N-H 键的一级倍频及其合频吸收[31−33]。5200 cm^{-1} 附近区域是 O-H 键伸缩振动[34],因水的强吸收而处于饱和状态,故避免使用该区间的波数建模。

对原始光谱数据采用不同预处理方法后的建模结果如表 2 所示:多元散射校正 MSC 是类胡萝卜素发酵液总糖和总氮定量分析模型的最佳预处理方法。经该方法处理后,其预测集表现出较高的预测决定系数 R^2p 和较低的预测均方根误差 RMSEP,且相对分析偏差 RPD 值高于其他预处理方法。与原始光谱相比,经 MSC 预处理后的光谱基线漂移现象得到了显著改善。由于果汁发酵液经过滤所得,其中仍含有少量细小颗粒,MSC 能有效消除因颗粒大小及分布不均匀引起的表面散射对光谱的干扰。此外,两个模型的 R^2p 经 10 次计算的标准偏差(SD)分别为 0.0295 和 0.0902,表明模型具有较好的稳定性和可靠性。但经预处理后的预测集 R^2p 值仍低于 0.95,这说明模型的预测精度虽可接受,但仍存在一定的提升空间,需要进行特征波数的筛选以提高预测的准确性。

2.3 特征波数的筛选

本研究的每个原始光谱包含 1557 个变量,其中存在大量冗余信息和噪声干扰。为优化模型性能,系统比较了 UVE、RF 和 CARS 三种特征波数筛选方法。如表 3 所示,三种方法均能在一定程度上提升模型预测精度,但提升效果存在显著差异。总糖和总氮模型经 UVE 法筛选后分别保留了 245 和 382 个变量,但由于未能有效剔除非信息变量,其预测精度相比全光谱模型的提升幅度有限(图 3(a)和(b))。

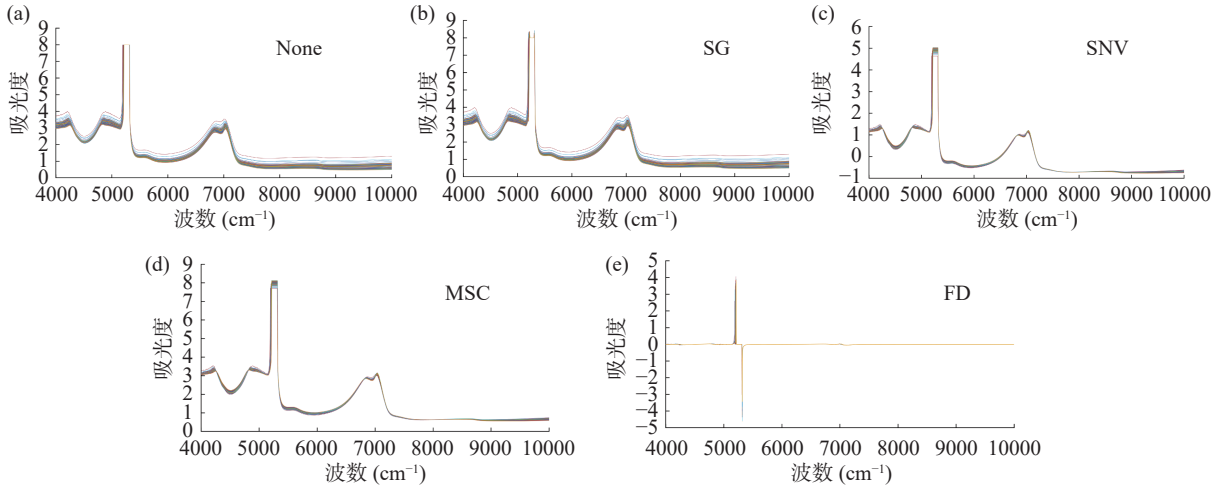


图 2 类胡萝卜素发酵液 NIR 原始光谱图(a)和预处理光谱图(b~e)

Fig.2 NIR raw spectrum (a) and preprocessed spectra (b~e) of carotenoid fermentation broth

表 2 光谱预处理方法对 PLRS 模型结果的影响

Table 2 Effect of spectral preprocessing methods on PLSR model results

预处理方法	TSS				TN			
	R^2p	RMSEP	SD	RPD	R^2p	RMSEP	SD	RPD
None	0.8067	6.5002	0.0572	2.5114	0.8553	0.3674	0.0581	2.5627
SG	0.8478	5.4509	0.0435	2.9406	0.8824	0.1812	0.0380	2.8370
SNV	0.8635	4.5917	0.0423	3.2213	0.9023	0.1257	0.0758	3.1604
MSC	0.9013	3.8071	0.0295	3.4197	0.9064	0.1266	0.0902	3.1999
FD	0.8697	4.4056	0.0439	3.1048	0.8894	0.2581	0.0844	2.8906

注:None,原始光谱;SG,卷积平滑;SNV,标准正态变量变换;MSC,多元散射校正;FD,一阶导数;R^2p,预测决定系数;RMSEP,预测均方根误差;SD,标准偏差,分别10次计算 R^2p的统计结果。

表 3　MSC 结合特征筛选的 TSS 和 TN 模型预测结果

Table 3　Prediction effect of TSS and TN models developed using MSC combined with different feature selection methods

指标	特征筛选方法	变量数	主因子数	R^2p	RMSEP	RPD
TSS	None	1557	10	0.9013	3.8071	3.4197
	UVE	245	8	0.9417	3.4433	4.4233
	RF	80	6	0.9337	3.5949	4.1500
	CARS	68	9	0.9670	2.6587	5.6627
TN	None	1557	9	0.9064	0.1266	3.1999
	UVE	382	8	0.9344	0.0718	3.9267
	RF	80	7	0.9301	0.0751	3.8687
	CARS	46	9	0.9502	0.0624	4.6604

注：None，无特征变量筛选；UVE，无信息变量筛除法；RF，随机蛙跳法；CARS，竞争性自适应全加重法。

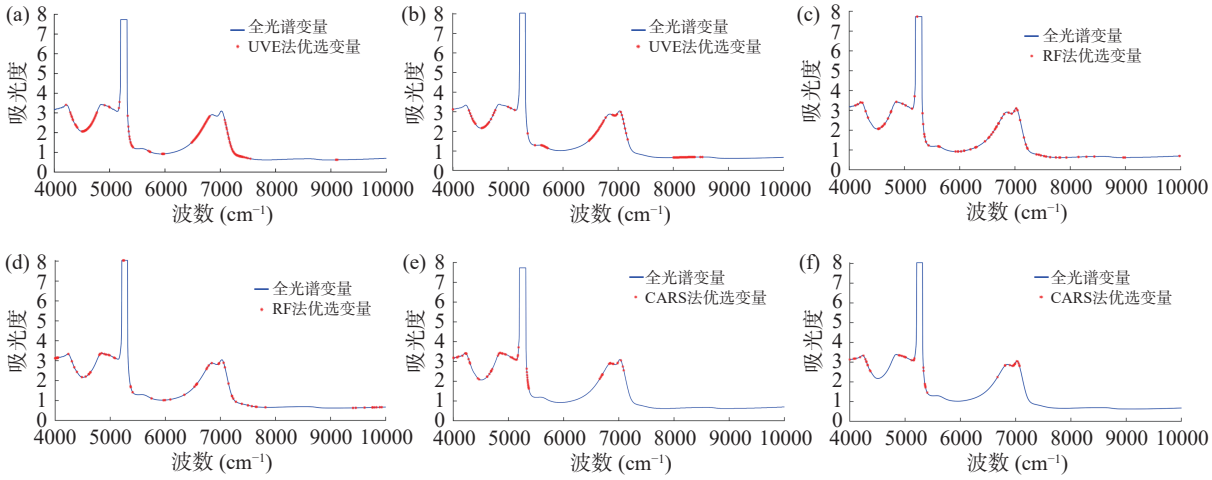


图 3　不同特征筛选方法对类胡萝卜素发酵液 TSS 与 TN 模型的特征波数选择结果

Fig.3　Characteristic wavenumbers selected by different feature selection methods for TSS models and TN models in carotenoid fermentation broth

注：TSS 模型：(a)UVE，(c)RF，(e)CARS；TN 模型：(b)UVE，(d)RF，(f)CARS。

RF 法则因其算法固有的随机性，易误选 5200 cm^{-1} 附近水的 O-H 伸缩振动平顶峰区域波数，导致模型过拟合，预测精度提升最小(图 3(c)和(d))。

相比之下，CARS 法通过蒙特卡洛采样与指数衰减机制，实现了基于化学结构导向的特征波数提取[35]。该方法显著降低了特征变量数量，其中 TSS 和 TN 模型特征变量数分别降至 68 和 46。如图 3(e)所示，CARS 法在总糖模型提取的特征波数主要集中在 4000~5100、5300~5500 和 6500~7200 cm^{-1} 范围，对应含氢基团的一级倍频吸收。该模型预测精度显著提升，R^2p 达到 0.9670，RMSEP 为 2.6587。对于总氮模型(图 3(f))，经筛选出的特征波数主要集中在 6500~7200 cm^{-1}(N-H 键一级倍频)和 4000~5500 cm^{-1}(N-H 键合频吸收)区域，其 R^2p 提升至 0.9502，RMSEP 为 0.0624。

2.4　定量模型的建立与评价

由 2.3 可知，采用 MSC 进行光谱预处理，并结合 CARS 筛选特征光谱，所建立的总糖及总氮定量模型效果最佳。其真实值和预测值的相关性如图 4 所示。两个模型的训练集和预测集的数据点分布均匀，且均紧密环绕在代表理想预测线(Y=X)的黑色实线附近。TSS 定量模型训练集 R^2c 为 0.9747、RMSEC

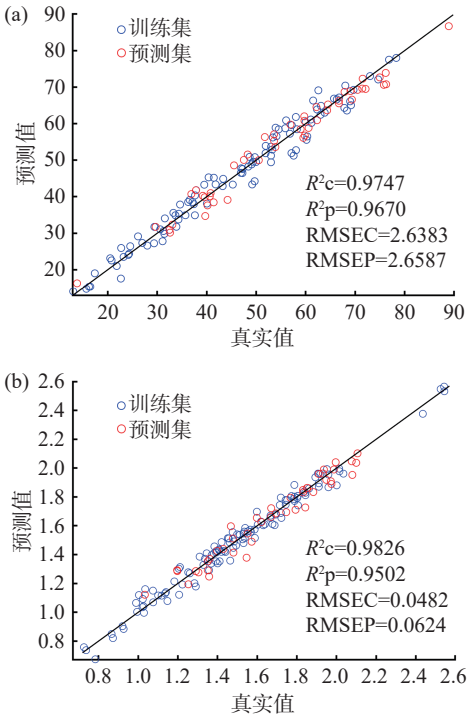


图 4　TSS(a)和 TN(b)定量模型的真实值与预测值的相关性

Fig.4　Correlation plots between measured and predicted values for TSS (a) and TN (b) quantitative models

为 2.6383；预测集 R^2p 为 0.9670、RMSEP 为 2.6587；TN 模型训练集 R^2c、RMSEC 分别为 0.9826 和 0.0482；其预测集 R^2p、RMSEP 分别为 0.9502 和 0.0624。上述模型的训练集和预测集 R^2 值均大于 0.95，表明所建立的定量模型具有较高的预测准确性。

2.5 模型的外部验证

外部验证是评估模型可靠性的关键步骤。本研究采用 20 个独立样本进行验证，并对其预测值与真实值进行 t 检验，以判定两组结果是否存在显著性差异。结果如表 4 所示，总糖定量模型的外部验证 RPD 值为 4.8386（RPD>3.0），RMSEP 为 2.7064，符合高精度定量分析的要求；而总氮模型由于其建模样本的浓度范围较窄（0.75~2.54 mg/mL），样本间理化性质差异较小，其外部验证 RPD 值为 2.7423（2.5≤RPD<3.0），RMSEP 为 0.1628，适用于常规定量分析。近红外预测值与化学测定值的 t 检验结果显示，两组数据的 P 值均大于 0.05，表明近红外光谱法的检测准确性与传统化学方法相当。在此基础上，本研究所建立的近红外定量模型还具备显著优势：其一，样品前处理更简便，近红外光谱法仅需离心处理，而传统方法检测发酵液总糖时需经过蛋白沉淀、过滤等步骤，检测总氮更是需要将样品进行高温消解；其二，检测效率更高，传统方法单次检测发酵液总糖含量需 40 min，总氮检测更需长达 2 h 以上；而基于 MSC+CARS+PLSR 建立的近红外分析方法，仅需几分钟就能完成单一样本的检测。相较于传统检测方法，近红外分析法实现了快速、无损检测，检测效率得到大幅提升。

表 4 外部验证结果

Table 4 Results of external validation

组分	样品量	模型	RMSEP	RPD	P值	适用性
总糖	20	MSC+CARS+PLSR	2.7064	4.8386	0.2300	高精度定量分析
总氮	20	MSC+CARS+PLSR	0.1628	2.7423	0.9972	常规定量分析

注：P值为近红外模型预测值与化学测定值的配对t检验结果。

3 结论

本研究利用 NIR 技术结合化学计量学方法，成功构建了用于快速、准确监测胶红酵母发酵陈皮柑果肉生产类胡萝卜素过程中关键参数（TSS 和 TN）含量的定量分析模型。探讨了 SG、MSC、SNV 等多种光谱预处理方法，以及 CARS、UVE 等多种波数筛选方法对预测模型性能的影响，并对最佳预测模型进行了外部验证。结果表明：采用 MSC 进行光谱预处理，并结合 CARS 筛选特征波数，所建立的总糖及总氮定量模型效果最佳。其中，TSS 最佳定量模型性能优异，其预测集 R^2p 和 RMSEP 分别为 0.9670 和 2.6587，外部验证 RPD 值为 4.8386（RPD>3.0），表明该模型具有较高的预测精度，满足精确监测发酵过程中 TSS 含量的要求。TN 最佳定量模型 R^2p 和 RMSEP 分别为 0.9502 和 0.0624，外部验证 RPD 值为 2.7423（2.5≤RPD<3.0），表明该模型具有良好的预测能力，适用于发酵液中 TN 含量的常规定量分析。外部验证 t 检验结果显示 TSS 和 TN 定量模型 P 值均大于 0.05，表明样品的近红外模型预测值与理化实测值在 α=0.05 水平上无显著性差异，具有较好的稳定性和可靠性。

综上，本研究建立的 NIR 定量模型，能有效克服传统检测 TSS 和 TN 方法操作繁琐、耗时长的局限，为基于 TSS 和 TN 动态变化的实时反馈补料策略优化奠定了坚实基础。该模型的应用对提高类胡萝卜素产率、提升陈皮柑果肉废弃物资源化利用效率具有重要实践意义。

© The Author(s) 2026. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

参考文献

[1] NEAGU S, COJOC R, ENACHE M, et al. Biotransformation of waste glycerol from biodiesel industry in carotenoids compounds by halophilic microorganisms[J]. Waste and Biomass Valorization, 2019, 10: 45−52.

[2] MA W, CHEN X, WANG B O, et al. Characterization, antioxidativity, and anti-carcinoma activity of exopolysaccharide extract from *Rhodotorula mucilaginosa* CICC 33013[J]. Carbohydrate Polymers, 2018, 181: 768−777.

[3] NIRANJANA R, GAYATHRI R, MOL S N, et al. Carotenoids modulate the hallmarks of cancer cells[J]. Journal of Functional Foods, 2015, 18: 968−985.

[4] 周琳, 梁轩铭, 赵磊. 天然类胡萝卜素的生物合成研究进展[J]. 生物技术通报, 2022, 30(7): 119−127. [ZHOU L, LIANG X M, ZHAO L. Biosynthesis of natural carotenoids: Progress and perspective[J]. Biotechnology Bulletin, 2022, 30(7): 119−127.]

[5] FONSECA K R S, LOTAS M K, CORTEZ A C A, et al. Exploration of carotenoid-producing *Rhodotorula* yeasts from Amazonian substrates for sustainable biotechnology applications[J]. Current Research in Microbial Sciences, 2025, 8: 100373.

[6] 汪杰, 孙向阳, 姚红梅, 等. 一株产类胡萝卜素菌株的鉴定和培养条件优化及抗氧化特性分析[J]. 西华大学学报(自然科学版), 2022, 41(4): 82−89. [WANG J, SUN X Y, YAO H M, et al. Identification of a carotenoid-producing strain and optimization of its culture conditions and analysis of its antioxidant activities[J]. Journal of Xihua University (Natural Science Edition), 2022, 41(4): 82−89.]

[7] SHARMA R, GHOSHAL G. Optimization of carotenoids pro-

duction by *Rhodotorula mucilaginosa*（MTCC-1403）using agro-industrial waste in bioreactor: A statistical approach[J]. Biotechnology Reports, 2020, 25: e00407.
[8] 李义勇, 陈柏忠, 沈许源, 等. 利用新会柑果浆培养海洋假丝酵母的发酵工艺[J]. 食品工业, 2022, 43(6): 55−58. [LI Y Y, CHEN B Z, SHEN X Y, et al. The fermentation process of the Xinhui citrus pulp for *Candida orthopsilosis* culture[J]. Food Industry, 2022, 43(6): 55−58.]
[9] 苏骏敏, 沈昌莹, 张树权. 食品中糖的检测方法研究进展[J]. 现代食品, 2022, 28(6): 43−47. [SU J M, SHEN C Y, ZHANG S Q. Research progress on the detection methods of sugar content in food[J]. Modern Food, 2022, 28(6): 43−47.]
[10] 王芳, 包怡红, 于震. 食品中蛋白质快速检测技术的研究[J]. 食品工业科技, 2014, 35(11): 372−375. [WANG F, BAO Y H, YU Z. Study on rapid determination of protein in food[J]. Science and Technology of Food Industry, 2014, 35(11): 372−375.]
[11] 蓝希华, 王志国, 栾小丽, 等. 基于近红外光谱技术的木糖含量快速检测[J]. 光谱学与光谱分析, 2025, 45(7): 1916−1923. [LAN X H, WANG Z G, LUAN X L, et al. Rapid detection for xylose content using near-infrared spectroscopy technology[J]. Spectroscopy and Spectral Analysis, 2025, 45(7): 1916−1923.]
[12] 余松柏, 黄张君, 吴奇霄, 等. 基于近红外光谱构建酒用高粱主要理化指标的快速无损分析模型[J]. 食品工业科技, 2023, 44(10): 311−319. [YU S B, HUANG Z J, WU Q X, et al. Constructing rapid and undamaged detection models for main physicochemical indexes of brewing sorghum based on near infrared spectrum[J]. Science and Technology of Food Industry, 2023, 44(10): 311−319.]
[13] KASEMSUMRAN S, BOONDAENG A, NGOWSUWAN K, et al. Simultaneous monitoring of the evolution of chemical parameters in the fermentation process of pineapple fruit wine using the liquid probe for near-infrared coupled with chemometrics[J]. Foods, 2022, 11(3): 377.
[14] BANERJEE S, MANDAL S, JESUBALAN G N, et al. NIR spectroscopy-CNN-enabled chemometrics for multianalyte monitoring in microbial fermentation[J]. Biotechnology and Bioengineering, 2024, 121(6): 1803−1819.
[15] 李浩源, 马晓博, 李连, 等. 近红外光谱分析技术在发酵工程中的应用[J]. 生命的化学, 2024, 44(7): 1178−1183. [LI H Y, MA X B, LI L, et al. The application of near infrared spectroscopy in fermentation engineering[J]. Chemistry of Life, 2024, 44(7): 1178−1183.]
[16] 王春媛, 齐景伟, 安晓萍, 等. 基于 NIR 定量模型快速预测发酵麸皮中还原糖和可溶性蛋白成分的含量[J]. 中国农业大学学报, 2024, 29(8): 190−203. [WANG C Y, QI J W, AN X P, et al. Rapid prediction of reducing sugar and soluble protein content in fermented bran based on NIR quantitative analysis model[J]. Journal of China Agricultural University, 2024, 29(8): 190−203.]
[17] 郝超, 赵忠盖, 刘飞. 基于近红外光谱的柠檬酸发酵液化清液概率偏最小二乘法监控[J]. 食品与发酵工业, 2020, 46(20): 214−220. [HAO C, ZHAO Z G, LIU F. Statistical process monitoring of clear liquefied solution for citric acid fermentation based on near infrared spectroscopy[J]. Food and Fermentation Industries, 2020, 46(20): 214−220.]
[18] ZHAO S, ADADE S Y S S, WANG Z, et al. On-line monitoring of total sugar during kombucha fermentation process by near-infrared spectroscopy: Comparison of linear and non-linear multiple calibration methods[J]. Food Chemistry, 2023, 423: 136208.
[19] KOZMA B, SALGÓ A, GERGELY S. On-line glucose monitoring by near infrared spectroscopy during the scale up steps of mammalian cell cultivation process development[J]. Bioprocess and Biosystems Engineering, 2019, 42(6): 921−932.
[20] LI C, JIN C, ZHAI Y N, et al. Simultaneous detection of citrus internal quality attributes using near-infrared spectroscopy and hyperspectral imaging with multi-task deep learning and instrumental transfer learning[J]. Food Chemistry, 2025, 481: 143996.
[21] CHE W K, SUN L J, ZHANG Q, et al. Application of visible near-infrared spectroscopy in the prediction of azodicarbonamide in wheat flour[J]. Journal of Food Science, 2017, 82(10): 2516−2525.
[22] 黄华丹, 邓健豪, 曹庸, 等. 基于近红外光谱技术快速检测广式酱油发酵过程中主要理化指标的含量[J]. 现代食品科技, 2024, 40(12): 386−401. [HUANG H D, DENG J H, CAO Y, et al. Rapid determination of main physicochemical indexes in the fermentation process of Cantonese soy sauce based on near infrared spectroscopy[J]. Modern Food Science and Technology, 2024, 40(12): 386−401.]
[23] DU C S, SUN L J, BAI H Y, et al. Quantitative detection of talcum powder in wheat flour based on near infrared spectroscopy and hybrid feature selection[J]. Infrared Physics & Technology, 2022, 123: 104185.
[24] 左旭丽, 潘秀珍, 耿泽宇, 等. 经 19 种光谱预处理筛选建立 3 种连翘叶茶近红外定性定量模型[J]. 食品安全质量检测学报, 2022, 13(19): 6431−6440. [ZUO X L, PAN X Z, GENG Z Y, et al. Establishment of the near infrared qualitative and quantitative models of 3 kinds of *Forsythia suspense* leaves tea by 19 kinds of spectral pretreatment and screening[J]. Journal of Food Safety and Quality, 2022, 13(19): 6431−6440.]
[25] 王甜甜, 冯国红, 朱玉杰. 近红外光谱结合化学计量学方法快速检测蓝莓可溶性固形物和维生素 C 含量[J]. 食品工业科技, 2023, 44(16): 297−305. [WANG T T, FENG G H, ZHU Y J. Rapid determination of soluble solids and vitamin C in blueberry by near infrared spectroscopy combined with chemometrics[J]. Science and Technology of Food Industry, 2023, 44(16): 297−305.]
[26] 梁子兆, 李欣, 刘朴, 等. 基于可见-近红外光谱的多品种猕猴桃贮藏品质的多指标综合预测模型研究[J]. 食品工业科技, 2025, 46(13): 282−291. [LIANG Z Z, LI X, LIU P, et al. Comprehensive multi-indicator prediction model for storage quality of multi-cultivar kiwifruit based on visible-near infrared spectroscopy[J]. Science and Technology of Food Industry, 2025, 46(13): 282−291.]
[27] 陈蓓, 郑恩让, 郭拓. 多种光谱变量筛选算法在红参提取近红外建模中的应用[J]. 光谱学与光谱分析, 2021, 41(8): 2443−2449. [CHEN B, ZHENG E R, GUO T. Application of various algorithms for spectral variable selection in NIRS modeling of red ginseng extraction[J]. Spectroscopy and Spectral Analysis, 2021, 41(8): 2443−2449.]
[28] COZZOLINO D, MURRAY I. Near-infrared spectroscopy in food analysis[M]. Cham: Springer, 2021: 1−20.
[29] SUN H, MA J, XU C, et al. Application research of near-infrared spectroscopy（NIRS）in online monitoring of the mixing uniformity of intermediate granules[J]. International Journal of Clinical and Experimental Medicine Research, 2025, 9(3): 258−267.

［30］郭歌, 张梦玲, 巩志杰, 等. 基于近红外光谱和复杂样品划分集合的生物质灰分含量模型构建[J]. 光谱学与光谱分析, 2023, 43(10): 3143−3149. ［GUO G, ZHANG M L, GONG Z J, et al. Construction of biomass ash content model based on near-infrared spectroscopy and complex sample set partitioning[J]. Spectroscopy and Spectral Analysis, 2023, 43(10): 3143−3149.］

［31］陈毅, 胡芸, 何军, 等. 基于近红外光谱法建立的模型预测烟用料液中水溶性糖的含量[J]. 理化检验-化学分册, 2022, 58(4): 381−386. ［CHEN Y, HU Y, HE J, et al. Prediction of water-soluble sugars in tobacco casings with models based on near infrared spectrometry[J]. Physical and Chemical Testing-Chemistry Section, 2022, 58(4): 381−386.］

［32］JENEPHA M S J, PRADHAN S, JAMES C. Molecular structure, NBO analysis of the hydrogen-bonded interactions, spectroscopic (FTIR, FTRaman), drug likeness and molecular docking of the novel anti COVID-2 molecule (2E)-N-methyl-2-[(4-oxo-4H-chromen-3-yl) methylidene]hydrazinecarbothioamide (dimer)-quantum chemical approach[J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2021, 251: 119388.

［33］李琳, 孙慧慧, 曹荣, 等. 基于近红外光谱技术的南极磷虾品质快速评定[J]. 食品工业科技, 2026, 47(1): 318−325. ［LI L, SUN H H, CAO R, et al. Rapid quality assessment of Antarctic krill using near-infrared spectroscopy technology[J]. Science and Technology of Food Industry, 2026, 47(1): 318−325.］

［34］邓孺孺, 何颖清, 秦雁, 等. 近红外波段(900-2500 nm)水吸收系数测量[J]. 遥感学报, 2012, 16(1): 192−206. ［DENG R R, HE Y Q, QIN Y, et al. Measuring pure water absorption coefficient in the near-infrared spectrum (900-2500 nm)[J]. Journal of Remote Sensing, 2012, 16(1): 192−206.］

［35］母雯竹, 张贵宇, 张维, 等. 基于 CARS-SPA 特征提取的黄水淀粉近红外光谱定量模型优化[J]. 食品科学, 2024, 45(19): 8−14. ［MU W Z, ZHANG G Y, ZHANG W, et al. Optimization of quantitative modeling of starch in Huangshui based on near-infrared spectral feature extraction using competitive adaptive reweighted sampling combined with successive projections algorithm[J]. Food Science, 2024, 45(19): 8−14.］