〈红外应用〉

发酵液中甘油和丁醇含量的近红外快速检测方法

赵 昇, 王晓荣, 张进明

(南京工业大学 电气工程与控制科学学院, 江苏 南京 211800)

摘要: 为了提高生物发酵过程的控制与优化水平,试对发酵底物甘油和目标产物丁醇的含量进行快速无损检测进行研究。首先对 80 份甘油与丁醇的单体系样本进行近红外扫描并进行光谱分析。为提高模型的预测能力,分别采用了偏最小二乘法、间隔偏最小二乘法、向前间隔偏最小二乘法、向后间隔偏最小二乘法和窗口移动最小二乘法5种定量校正方法建立甘油和丁醇含量的近红外检测模型并对模型进行分析与比较。结果表明,向后间隔偏最小二乘法建立的模型效果较好,甘油和丁醇的单组份溶液的检测模型相关系数分别达到 0.99932 (甘油)和 0.98843 (丁醇)。为测量真实的甘油发酵液中甘油和丁醇含量,搭建并建立了甘油和丁醇浓度监测平台,验证得相关系数分别达到 0.99074 (甘油)和 0.99261 (丁醇)。结果表明建立的近红外快速检测模型在检测的准确性和快速性上均有优异的性能,为发酵行业快速检测提供了新的检测手段。

关键词: 近红外光谱技术; 向后间隔偏最小二乘; 甘油发酵; 丁醇

中图分类号: O657.33 文献标识码: A 文章编号: 1001-8891(2020)05-0488-06

Rapid Detection of Glycerol and Butanol in Fermentation Broth Using Near-Infrared Spectroscopy

ZHAO Sheng, WANG Xiaorong, ZHANG Jinming

(College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing 211800, China)

Abstract: In this paper, to improve the control and optimization of the bio-fermentation process, rapid and non-destructive detection of glycerol and butanol was studied. To this end, 80 samples of glycerol and butanol were scanned and analyzed by means of near-infrared spectroscopy. Furthermore, in order to improve the predictive ability of the model, five quantitative calibration methods, namely, the partial least squares method, interval partial least squares method, forward interval partial least squares method, backward interval partial least squares method, and minimum forward interval method were used to establish the near-infrared detection model for glycerol and butanol content. The models were then analyzed and compared. The results showed that the model established by means of the backward interval partial least squares method was effective, and the correlation coefficients of the single component solution of glycerol and butanol were 0.99932 (glycerol) and 0.98843 (butanol), respectively. To measure the content of glycerol and butanol in the glycerol fermentation broth, a monitoring platform for glycerol and butanol concentrations was established. The correlation coefficients were found to be 0.99074 (glycerol) and 0.99261 (butanol), respectively. The results show that the NIR rapid detection model demonstrates excellent performance with regards to accuracy and rapidity and provides a new detection method for the rapid detection of glycerol and butanol in the fermentation industry.

Key words: near-infrared spectroscopy, backward interval partial least squares, glycerol fermentation, butanol

收稿日期: 2018-10-16; 修订日期: 2020-03-30.

作者简介: 赵昇(1995-),男,江苏常州人,硕士研究生,主要从事嵌入式系统开发研究。E-mail: techmanx@163.com。

通信作者: 王晓荣(1972-), 男,副教授,硕士生导师,主要研究方向为分析仪器和嵌入式系统设计。E-mail: wang@njut.edu.cn。

0 引言

生物发酵行业是我国的支柱行业之一,但由于生 产技术比较落后,发酵产物中附加产物较多,而且新 一代的发酵制品如生物酶、干扰素等前沿产物主要被 欧美国家控制,我国发酵工业的技术水平与国外仍有 较大差距。由于发酵过程比较复杂, 传统的生物传感 器在自动控制系统中不能满足生产需求, 生产过程最 直接的控制参数生物(生化)参数(包括底物、重要 中间代谢物和目标产物等) 快速检测或在线检测技术 的缺乏,是整个行业面临的瓶颈技术。可见,快速监 测发酵液底物与产物的变化对于实现工业发酵从"粗 放化、间歇式"向"精密化、连续化"的生产方式的 转型具有重要的现实意义[1]。而丁醇作为一种有机化 工原料,其热值与辛烷值和乙醇十分相似,且具有能 量含量高、挥发性低、减少温室气体排放等优点被认 为是比乙醇更为理想的燃料, 丁醇还可以用于塑料制 品的制造及食用香味的提取等。但工业生产丁醇主要 以石油为原料,而随着能源的枯竭,发酵法的发展前 景越来越受重视, 大有取代化学法的趋势。因此, 十 分有必要研究并提供一种快速检测发酵液中甘油与 丁醇含量的方法[2]。

近红外光谱技术在检测含氢、碳、氮等有机物的 成分上其优势非常明显,因其速度快、无损性、易操 作等检测的优点,被广泛应用[3]。国内外很多专家学 者应用近红外技术在发酵检测领域进行学术研究。 Cavinato 等采用短波近红外光谱技术对发酵过程中乙 醇的含量进行了测定^[4]。Macalony 等人在工业上高密 度大肠杆菌发酵过程中采用了近红外光谱测定了甘 油、铵盐和乙酸的浓度^[5]。Arnold 等人采用近红外光 谱技术测定 CHO 细胞分批发酵过程中氨酰胺、氨基 酸、葡萄糖与乳酸的含量[6]。但甘油和丁醇这些中工业 发酵过程起关键作用而光谱相似的状态变量却是鲜有 报道。而且基于 MEMS 的近红外光谱仪体积小、分辨 率高、价格低,解决了目前由于光谱仪体积大,价格 高等原因造成其在发酵工业中不能普及的问题^[7]。因 此,本文开展了基于 MEMS 近红外分析技术在甘油 转丁醇的发酵过程对参数进行快速检测的应用与研 究,为甘油发酵过程中检测甘油与丁醇含量提供了一 种快速、低价格的检测手段。

1 检测模型的比较

1.1 试剂与仪器

DLP NIRscan Nano EVM 近红外光谱仪,电子天平(精度为 0.0001 g)、恒温箱、去离子蒸馏水、量

筒、培养皿、滴管、甘油、丁醇等。

1.2 仪器参数设定与光谱采集

配置不同浓度的甘油与丁醇溶液,采用去离子蒸馏水精确配置容量为 $4 \, \text{mL}$,浓度为 $1 \sim 80 \, \text{g/L}$ 的甘油与丁醇的样本溶液 $80 \, \text{份}$,浓度间隔为 1%。采用美国 TI 公司生产的 NIRscan Nano 近红外光谱仪,扫描 32 次,分辨率为 $7 \, \text{cm}^{-1}$,波数间隔为 $2 \, \text{cm}^{-1}$ 。以空气为参比,室内温度保持在 $25 \, \text{℃左右}$,湿度基本保持不变,每个样本平行采集 $3 \, \text{次}$,取其平均光谱作为该样本的原始光谱。

1.3 样本集的划分

若选择的校正样本比较具有代表性,则在建模的时候可以减少很多的工作量,而且可以提高模型的精确度^[8-9]。浓度梯度法可以有效地涵盖组分含量,避免了样本集划分的不均匀,而且保证了后续处理和建模方法的适用范围。根据浓度梯度法本文将正常的74个甘油与丁醇样本按照浓度梯度法2:1的比例进行排序。表1为划分校正集与预测集的有机质含量统计表,校正集中甘油含量的范围为:1~80g/L,预测集中甘油含量为:4~79g/L;丁醇校正集含量范围为:1~80g/L,预测集含量为:3~79g/L。

1.4 光谱预处理

在实验过程中仪器的噪音总会避免不了,因此在模型建立之前要先对光谱进行预处理^[10]。本文采用 S-G、多元散射校正、导数校正和标准正态校正对甘油和丁醇样本的光谱进行预处理,结果如表 2 所示。

从表 2 中我们可以看出,不同的预处理方法,对建立的模型精度有不同的影响。模型评价指标有很多,本文选用了相关系数 (R_c) 和交互验证均方根误差 (Root Mean Square Error For Validation,RMSECV) 作为评价模型的可信度。一般来说,相关系数越大,交互验证均方根误差越小,说明近红外模型的测定值与实际值越吻合,模型的准确度越高,也即所建立的模型效果越好。明显可以看出采用标准正态变量(Standard Normal Variables,SNV)的方法在 R_c 和 (Root Mean Square Error For Prediction,RMSECP) 两个指标上性能优异,是最优的预处理方法。

1.5 模型的建立与最优模型的建立

本实验分别在甘油和丁醇的原始光谱图采用 SNV 预处理之后,建立了偏最小二乘法(Partial Least Squares,PLS)、间隔偏最小二乘(Interval Partial Least Squares,IPLS)、向前间隔偏最小二乘法(Forward Interval Partial Least Squares,FIPLS)、向后间隔偏最小二乘法(backward Interval Partial Least Squares,BIPLS)、移动窗口偏最小二乘法(Moving Window Partial Least

表 1 校正集与预测集生物量的测量结果 Table 1 Modeling and forecasting sets biomass measurement results

Process parameter	Sample size	Sample size	Concentration range(g/L)	Average value(g/L)	Standarddeviation
Glycerol	Calibration set	50	1-80	40.7143	18.9819
	Prediction set	24	4-78	41.5806	18.7139
Butanol	Calibration set	50	1-80	40.6032	18.89449
	Prediction set	24	3-78	41.3871	18.6818

表 2 不同预处理结果甘油和丁醇的预测模型结果

Table 2 Prediction results of different pretreatment results for Glycerol and Butanol

	Glycerol				Butanol			
Method	Method Calibratio		ation set Validation set		Calibration set		Validation set	
	$R_{\rm c}$	RMSECV	$R_{ m v}$	RMSECP	$R_{\rm c}$	RMSECV	$R_{ m v}$	RMSECP
Original	0.88355	2.79	0.86703	2.37	0.89974	3.07	0.88767	3.11
S-G	0.90137	2.28	0.80077	2.08	0.90251	2.87	0.89912	2.79
MSC	0.90546	2.27	0.80024	2.07	0.90792	2.27	0.90471	2.56
1st	0.91738	1.65	0.83166	1.87	0.91377	1.91	0.91896	2.14
2st	0.92859	1.54	0.85919	1.69	0.93055	1.80	0.92034	1.91
SNV	0.96979	1.24	0.89245	1.31	0.95443	1.29	0.95331	1.48

Squares,MWPLS)5 种校正集样本模型^[11-13]。建模结果采用 RMSECV 作为模型的最终评价指标,表 3 为PLS 模型预测结果。

表 3 PLS 模型的预测结果

Table 3 Predictive results of the PLS model

Method	Glycerol (RMSECV)	Butanol (RMSECV)
PLS	1.50	1.56
IPLS	1.31	1.52
FIPLS	1.25	1.47
BIPLS	1.21	1.23
MWPLS	1.25	1.51

从甘油与丁醇的预测结果来看,5种 PLS 回归模型中,iPLS、FiPLS、BiPLS、MWPLS 都是对偏最小二乘法的改进,模型的稳定性与准确度有一定的提高。通过数据比较,可以得出 BiPLS 模型可以减少无效波段对模型的影响,同时减少了计算机的计算量,

建立的模型效果最好。但真实的发酵过程,发酵液的 组成成分与标准溶液有很大的区别,且丁醇的含量是 逐渐上升的,模型的效果还需经过真实发酵液检测进 行验证。

2 模型的验证

2.1 发酵液的制备与测定

仪器:生物传感仪、DLP NIRscan 近红外光谱仪 (美国德州仪器)、电子天平(精度 0.0001 g)、5 L 发酵罐、TQ Analyst 计量学分析系统。

试剂: 甘油、酵母粉、二硫苏糖醇、L-半胱氨酸盐酸盐、九水硫化钠、氯化钠、六水氯化镁、氯化铵等。菌体培养基参数如表 4 所示。

在实验过程中,要保持发酵罐中的 pH 值为 7.0~7.2; 溶氧控制维持: 0~10%; 温度控制: 初始温度为 34%。

表 4 培养基参数列表

Table 4 Media parameter list

	90.7110	Table 4 Wedia parameter list	
Formula	Description	Contents/(g/L)	Total/g
Trimethylol methylaminoethyl		4.85	24.25
	PH buffer		
$\mathrm{KH_{2}PO_{4}}$		0.75	3.75
K_2HPO_4		0.75	3.75
Yeast powder		3	15
L-cysteine hydrochloride		0.0242	0.121
$Na_2S.9H_2O$	Reducing agent	0.048	0.24
Dithiothreotol (DTT)		0.0771	0.3855
NaCl		1	5
MgCl ₂ .6H ₂ O		0.5	2.5
$\mathrm{KH_{2}PO_{4}}$		0.2	1
$\mathrm{NH_4Cl}$	Salt	0.3	1.5
KCl		0.3	1.5
CaCl ₂ .2H ₂ O		0.015	0.075

含量测定:实时取出的发酵液中甘油和丁醇的含量采用生物传感仪进行测量,测量的数据用来建立甘油发酵液模型。然后再从发酵罐中取出 10 组未知浓度样品检测模型的准确性如表 5 所示(F_1 已知样本, F_2 未知样本)。

表 5 发酵液中甘油与丁醇含量

Table 5 The content of glycerol and butanol in fermentation broth

				g/L	
Formantation time/h	F	1	F_2		
Fermentation time/h	Glycerol	Butanol	Glycerol	Butanol	
4	98.27	0.64			
5			97.89	0.74	
8	97.01	1.19			
12	95.89	1.32			
15			94.05	1.87	
16	93.62	2.07			
20	90.38	3.19			
24	86.79	4.11			
25			85.01	4.86	
28	81.88	6.07			
32	74.33	9.17			
35			65.99	13.34	
36	66.91	14.26			
40	57.11	18.11			
44	49.47	22.33			
45			45.23	24.47	
48	37.29	29.35			
52	33.44	30.70			
55			27.89	32.92	
56	29.85	31.67			
60	26.25	32.30			
64	22.89	32.97			
65			20.55	33.72	
68	17.11	34.58			
72	14.09	35.27			
75			12.11	35.33	
76	12.89	35.49			
80	10.71	35.86			
84	8.65	36.26			
85			7.71	36.35	
88	6.69	36.50			
92	5.03	36.70			
95			3.44	36.86	
96	3.77	36.95			
100	2.85	37.05			

2.2 近红外光谱采集

采用近红外漫反射法对上述发酵液进行采集,以空气为参比。仪器参数设置为:分辨率为 10 nm,扫描次数 30 次。采集的光谱图如图 1 所示。

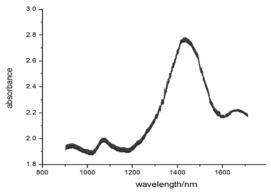


图 1 甘油发酵液的近红外光谱图

Fig.1 Near-infrared spectra of glycerol fermentation broth

2.3 样本集的划分

按照浓度梯度法对样本集按照校正集:预测集 =2:1 的比例进行划分。样本划分如表 6 所示。

2.4 模型的建立

本节直接将样本光谱经过 SNV 预处理之后,再结合 BiPLS 建立模型,其结果如图 2 和图 3 所示。

由图 2 和图 3 可以看出,甘油发酵液的甘油校正模型相关系数为 0.99074,交互验证均方根误差为 1.11,而甘油发酵液的丁醇模型相关系数为 0.99261,交互验证均方根误差为 1.09,相关系数均接近 1, RMSEC 数值均较小,表明模型的预测能力较好,精确度较高。

2.5 传统的生物测量法与近红外检测法的比较

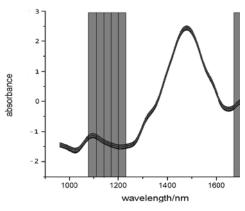
使用生物传感仪对未知样本进行浓度测定,然后 采用近红外光谱仪对未知样本进行光谱采集,使用建 立的甘油发酵模型预测甘油与丁醇的浓度如表 7,结 果比较如图 4 所示。

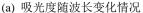
生物发酵中,基本上采用生物传感仪测量发酵液的组分含量,目前市场上高精度的生物传感仪的测量误差基本也保持在1%~2%以内。从表7可以看出,当以生物传感仪测量值为标准值时,采用近红外光谱技术测量值甘油的平均测量误差为3.29%,丁醇的平均测量误差为4.58%。发酵过程需要控制测定范围,

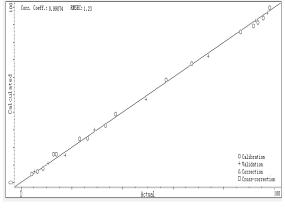
	表 6	样本集的划分	Table 6 Division	Table 6 Division of sample sets	
Process parameter	Sample size	Sample size	Concentration range/(g/L)	Average value/(g/L)	Standard deviation
Glycerol	Calibration set	17	2.85-98.27	44.9128	35.49578
	Prediction set	8	5.03-97.01	46.6925	36.65657
Butanol	Calibration set	17	0.64-37.05	22.5640	14.59323
	Prediction set	8	1.32-36.25	21.8371	14.98168

测量误差为 5%,显然精度是满足要求的。从图 4 可以看出模型对未知样本的预测与传统方法测定的真实值时极为接近的,模型的预测能力较好。近红外光谱方法相比传统的生物方法有着明显的优势,尤其在测量速度上,通常可以在几秒内就可以完成,可应用

于过程分析,而且不需对发酵液进行预处理,降低了人为因素误差。因此,近红外光谱分析技术测定甘油 发酵过程中甘油与丁醇的含量是一种切实可行的快速检测方法。







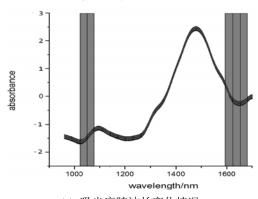
(b) 校正模型预测图

(a) Absorbance varies with wavelength

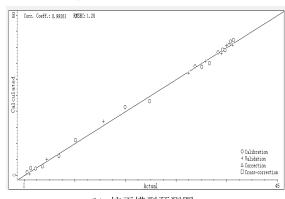
(b) Calibration Model Prediction Diagram

图 2 甘油发酵液中甘油校正模型预测图

Fig.2 Glycerin Correction Model Prediction Chart in Glycerin Fermentation Broth



(a) 吸光度随波长变化情况
(a) Absorbance varies with wavelength



(b) 校正模型预测图

(b) Calibration Model Prediction Diagram

图 3 甘油发酵液中丁醇校正模型预测图

Fig.3 Butanol Correction Model Prediction Chart in Glycerin Fermentation Broth

表 7 传统生物检测与近红外检测结果

Table 7 Traditional Biological Detection and Near-infrared Detection Results

	Glycer	rol/(g/L)	Butanol/(g/L)		
Fermentation time/h	Biometric	Near-infrared	Biometric	Near-infrared	
	method	detection	method	detection	
5	97.89	97.21	0.74	0.87	
15	94.05	94.77	1.87	2.02	
25	85.01	83.04	4.86	5.16	
35	65.99	67.01	13.34	14.08	
45	45.23	43.88	24.77	24.15	
55	27.89	27.11	32.92	33.53	
65	20.55	20.32	33.72	34.17	
75	12.11	13.09	35.33	35.69	
85	7.71	8.15	36.35	36.66	
95	3.44	4.27	36.86	37.22	

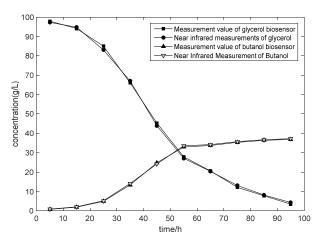


图 4 传统生物测量方法与近红外检测结果示意图

Fig.4 Comparison of traditional detection methods and near-infrared detection methods

3 总结

本文研究了近红外光谱技术在发酵制备丁醇过程中底物甘油和丁醇含量的快速检测方法。首先根据甘油与丁醇的近红外微观特性,确定了甘油和丁醇的吸收峰分别为(1149~1265 nm 和 1666~1801 nm),然后采用 PLS、iPLS、FiPLS、BiPLS、MWPLS 对甘油与丁醇的单体系溶液,建立了校正模型。通过对比分析,BiPLS 可以减少无效波段对模型的影响,同时减少了计算机的运算量,建立的模型效果最好,为真实甘油发酵的模型建立提供了数据基础。针对甘油发酵液的特性,采用了 BiPLS 建立了甘油发酵液的校正模型,并采用传统生物检测方法和甘油发酵模型对未知样本进行预测。实验表明,在甘油转丁醇的发酵过程,NIRS 可以作为一种快速、低价格的检测手段,为发酵行业实现自动化控制提供了理论基础与数据支持。

参考文献:

- 1] 欧阳平凯. **发酵工程关键技术及其应用[M]**. 北京: 化学工业出版社, 2005: 15-20.
 - OUYANG Pingkai. Key technology and application of fermentation engineering[M]. Beijing: Chemical Industry Press, 2005: 15-20.
- [2] 宋钢,郑璞,倪晔,等.木薯发酵产丁醇的研究[J]. 生物加工过程, 2012, **10**(2): 6-10.
 - SONG Gang, ZHEN Pu, NI Ye, et al. Study on the production of butanol from cassava[J]. Chinese Journal of Bioprocess Engineering, 2012,

10(2):6-10.

- [3] 任东,瞿芳芳,陆安详,等. 近红外光谱分析技术与应用[M]. 北京: 科学出版社, 2017.
 - REN Dong, QU Fangfang, LU Anxiang, et al. *Near Infrared Spectroscopy Technology and Application*[M]. Beijing: Science Press, 2017: 18-28.
- [4] Arnold M A, Small G W. Determination of physiological levels of glucose in an aqueous matrix with digitally filtered Fourier transform near-infrared spectra[J]. *Analytical chemistry*, 1990, 62(14): 1457-1464.
- [5] Macaloney G, Hall J W, Rollins M J, et al. The utility and performance of near-infra red spectroscopy in simultaneous monitoring of multiple components in a high cell density recombinant Escherichiacoli production process[J]. *Bioprocess Engineering*, 1997, 17(3): 157-167.
- [6] Arnold M Opiyo, Ying Tiejin. Effect of 1-Methylcy- clopropene Post-harvest Treatment on Ripening Process in Cherry Tomato Fruit[J]. Journal of Plant Physiology and Molecular Biology, 2005(1): 27-33.
- [7] 于新洋, 卢启鹏, 高洪智, 等. 便携式近红外光谱仪器现状及展望[J]. 光谱学与光谱分析, 2013(11): 2.
 - YU Xinyang, LU Qipeng, GAO Hongzhi, et al. Present situation and Prospect of portable near infrared spectrometer[J]. *Spectroscopy and Spectral Analysis*, 2013(11): 2.
- [8] 宋相中. 近红外光谱定量分析中三种新型波长选择方法研究[D]. 北京: 中国农业大学, 2017.
 - SONG Xiangzhong. Research of three new wavelength selection methods in near infrared spectroscopy quantitative analysis area[D]. Beijing: China Agricultural University, 2017.
- [9] 于燕波. 近红外光谱分析技术在转基因水稻识别和高油棉籽筛选中的应用研究[D]. 北京: 中国农业大学, 2014.
 - YU Yanbo. Application of Near Infrared Spectroscopy Analysis Technology in Transgenic Rice Identification and High Oil Cotton Seed Screening[D]. Beijing: China Agricultural University, 2014.
- [10] Davies T. Book review. Near infrared technology in the agricultural and food industries[J]. *Nir News*, 2002, **13**(5): 7-7.
- [11] WANG Z, Felmy A R, Thompson C J, et al. Near-infrared spectroscopic investigation of water in supercritical CO₂ and the effect of CaCl₂[J]. Fluid Phase Equilibria, 2013, 338(4): 155-163.
- [12] M Khalique Ahmed, Jacob Levenson. Application of Near-infrared Spectroscopy to the Quality Assurance of Ethanol and Butanol Blended Gasoline[J]. Petroleum Science and Technology, 2012, 30(2): 2-5.
- [13] 邹孝恒, 郝中骐, 易荣兴, 等. 基于遗传算法和偏最小二乘法的土壤 激光诱导击穿光谱定量分析研究[J]. 分析化学研究报告, 2015(2): 3. ZOU Xiaoheng, HAO Zhongqi, YI Rongxing, et al. Quantitative Analysis of Soil by Laser-induced Breakdown Spectroscopy Using Genetic Algorithm-Partial Least Squares[J]. Chinese Journal of Analytical Chemistry, 2015(2): 3.