

• 工业生物技术 •

基于傅里叶变换近红外光谱实时分析 1,3-丙二醇发酵过程生物量的在线监测方法

王路¹, 刘涛¹, 陈洋², 孙亚琴², 修志龙²

1 大连理工大学 控制科学与工程学院, 辽宁 大连 116024

2 大连理工大学 生命科学与技术学院, 辽宁 大连 116024

王路, 刘涛, 陈洋, 等. 基于傅里叶变换近红外光谱实时分析 1,3-丙二醇发酵过程生物量的在线监测方法. 生物工程学报, 2017, 33(1): 68–78.

Wang L, Liu T, Chen Y, et al. On-line monitoring of biomass in 1,3-propanediol fermentation by Fourier-transformed near-infrared spectra analysis. Chin J Biotech, 2017, 33(1): 68–78.

摘要: 生物量是反映生物发酵过程进展的重要参数, 对生物量进行实时监测可用于对发酵过程的调控优化。为克服目前主要采用的离线方法检测生物量时间滞后和人工测量误差较大等缺点, 本研究针对 1,3-丙二醇发酵过程设计了一个基于傅里叶变换近红外光谱实时分析技术的生物量在线监测实验平台, 通过对实时采集光谱预处理以及敏感光谱段分析, 应用偏最小二乘算法, 建立了 1,3-丙二醇发酵过程生物量变化的动态预测模型。以底物甘油浓度为 60 g/L 和 40 g/L 的发酵过程作为外部验证实验, 分析得到模型的预测均方根误差分别为 0.341 6 和 0.274 3, 结果表明所建立的模型具有较好的实时预测能力, 能够实现对 1,3-丙二醇发酵过程中生物量的有效在线监测。

关键词: 近红外光谱分析技术, 在线监测, 1,3-丙二醇发酵过程, 生物量, 偏最小二乘算法

On-line monitoring of biomass in 1,3-propanediol fermentation by Fourier-transformed near-infrared spectra analysis

Lu Wang¹, Tao Liu¹, Yang Chen², Yaqin Sun², and Zhilong Xiu²

1 School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, Liaoning, China

2 School of Life Science and Biotechnology, Dalian University of Technology, Dalian 116024, Liaoning, China

Abstract: Biomass is an important parameter reflecting the fermentation dynamics. Real-time monitoring of biomass can

Received: June 29, 2016; **Accepted:** September 26, 2016

Supported by: National Natural Science Foundation of China (Nos. 61473054, 61633006, 21306021), National Thousand Talents Program of China, Fundamental Research Funds for the Central Universities of China (No. DUT15ZD108).

Corresponding author: Tao Liu. Tel: +86-411-84706465; Fax: +86-411-84706706; E-mail: tliu@dlut.edu.cn

国家自然科学基金 (Nos. 61473054, 61633006, 21306021) 第三批国家青年千人计划, 大连理工大学重点培育基金项目 (No. DUT15ZD108) 资助。

be used to control and optimize a fermentation process. To overcome the deficiencies of measurement delay and manual errors from offline measurement, we designed an experimental platform for online monitoring the biomass during a 1,3-propanediol fermentation process, based on using the fourier-transformed near-infrared (FT-NIR) spectra analysis. By pre-processing the real-time sampled spectra and analyzing the sensitive spectra bands, a partial least-squares algorithm was proposed to establish a dynamic prediction model for the biomass change during a 1,3-propanediol fermentation process. The fermentation processes with substrate glycerol concentrations of 60 g/L and 40 g/L were used as the external validation experiments. The root mean square error of prediction (RMSEP) obtained by analyzing experimental data was 0.341 6 and 0.274 3, respectively. These results showed that the established model gave good prediction and could be effectively used for on-line monitoring the biomass during a 1,3-propanediol fermentation process.

Keywords: near-infrared spectra analysis technology, on-line monitoring, 1,3-propanediol fermentation process, biomass, partial least-squares (PLS) algorithm

发酵技术现已被应用于很多生产领域,关于发酵过程监控与优化方面的研究受到越来越多的关注和探讨^[1-3]。相对于很多其他生化工程,发酵过程大多数成分复杂且变化较快^[4],因此要对发酵过程实施在线调控,首先要解决的问题就是如何实现对过程重要参数的实时监测^[5]。生物量是发酵过程的一个重要参数^[6-7],目前采用的细胞干重、分光光度计测量吸光光度值(OD)等监测方法比较费时,并且不能实现在线监测,因而造成目前发酵过程控制明显落后于现今工业过程控制的发展水平。

近红外光谱技术拥有无损、快速以及无需样品预处理等优点^[8-9],通过结合先进的化学计量学方法,可以实现对工业过程的实时在线监控。目前已经越来越受到人们的关注,被广泛应用于食品、农业、医药以及化工等领域^[10-13]。将近红外光谱分析技术应用于生物发酵过程参数的监测也已经有一些报道,但是大都应用离线方法。桂勇利等^[14]通过建立近红外定量校正模型对谷氨酸发酵过程中乳酸的浓度进行快速离线检测。Guo等^[15]通过近红外光谱技术离线监测乳酸链球菌发酵过程的乳酸链球菌效价、还原糖浓度、细胞浓度和pH这4个关键参数。Cruz等^[16]应用近红外光谱技术对废弃食用油发酵过程的胞内物质P(3HB)等重要参数进行了快速离线检测。

本文针对克雷伯氏杆菌发酵生成1,3-丙二醇过程中的重要参数——生物量,开展在线监测研究,实时采集发酵过程的近红外光谱,对光谱进行敏感谱段选择,通过比较多种光谱预处理方法,对发酵过程生物量应用偏最小二乘算法建立较为准确的在线监测模型。以测量吸光光度值(OD)的方法作为检测生物量的离线参考方法,进行实验验证,结果表明本文所建立的模型能够实现1,3-丙二醇发酵过程中生物量的有效实时在线监测。

1 材料与方 法

1.1 主要仪器与1,3-丙二醇发酵过程在线监测实验平台

本文使用的仪器和软件主要包括:傅立叶在线近红外分析仪(型号:TALYS-ASP531,配以波数范围为4 000-12 500 cm^{-1} 的浸没式透反射探头,ABB公司),近红外光谱分析与处理软件HorizonMB(ABB公司),5 L发酵罐(型号:5BG,上海保兴生物设备有限公司),722 s分光光度计(波长范围:340-1 000 nm,上海精密科学仪器有限公司)等。通过对仪器的调整,本实验室设计并搭建了基于近红外光谱分析技术的1,3-丙二醇发酵过程在线监测实验平台,如图1所示。

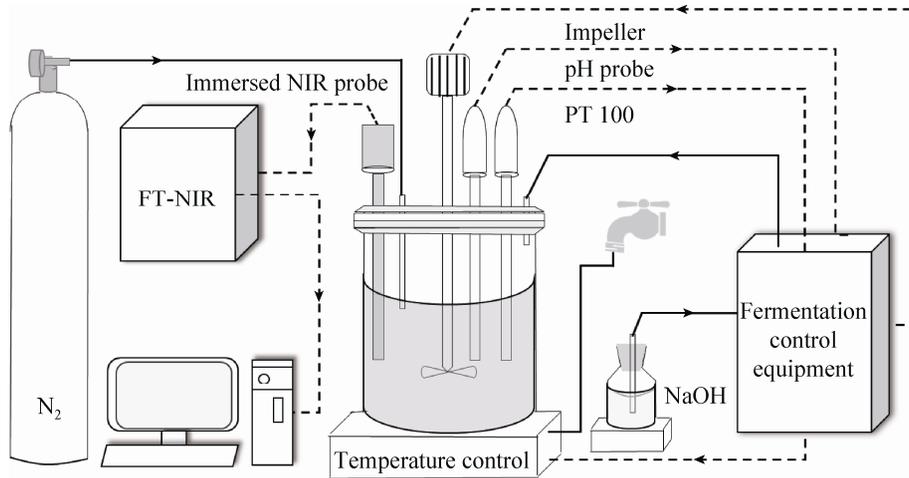


图 1 1,3-丙二醇发酵过程在线监测实验平台

Fig. 1 Experimental platform for on-line monitoring a 1,3-propanediol fermentation process.

本实验平台主要用于研究基于近红外光谱分析技术在线监测 1,3-丙二醇发酵过程的方法。其中发酵反应在 5 L 发酵罐内进行,发酵过程中使用生物发酵控制器保持发酵罐内的温度(使用 PT100 探头采集温度信息,并使用加热装置和冷水进行调节)和 pH(使用 pH 计采集酸碱度信息,并使用 NaOH 进行调节)稳定于设定值,并控制搅拌桨对发酵液进行稳定的搅拌。通过通入氮气的方式使发酵罐内保持厌氧状态,以满足菌体生长的需要。在发酵反应进行的同时,使用配以浸没式透反射探头的傅里叶在线近红外光谱仪(探头和光谱仪之间使用光纤连接)对整个发酵过程进行近红外光谱的在线采集,并可通过计算机进行实时的谱图分析与在线监测工作。

1.2 菌种、培养基组成与实验方法

菌种: 本文所进行的实验中使用的菌种均为克雷伯氏肺炎菌 *Klebsiella pneumoniae* CGMCC 2028, 是大连理工大学生命科学与技术学院课题组筛选诱变、保存于中国微生物菌种保藏中心的菌种。

种子培养基组成如表 1 所示。

其中微量元素 A 的组成如表 2 所示。

其中, Fe^{2+} 溶液由 12 mol/L HCl (4 mL/L) 和 $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ (5 g/L) 组成, Ca^{2+} 溶液由 CaCl_2 (20 g/L) 组成。

发酵培养基组成如表 3 所示。

其中微量元素 B 的组成如表 4 所示。

表 1 种子培养基组成

Table 1 Compositions of the seed culture medium

Substances	Concentration	Substances	Concentration
Glycerol	20.0 g/L	$(\text{NH}_4)_2\text{SO}_4$	2.0 g/L
KH_2PO_4	1.3 g/L	$\text{K}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$	4.454 g/L
CaCO_3	2.0 g/L	$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0.2 g/L
Ca^{2+}	1.0 mL/L	Trace elements A	2.0 ml/L
Fe^{2+}	1.0 mL/L	Yeast extract powder	1.0 g/L

表 2 微量元素 A 组成

Table 2 Compositions of the trace elements A

Substances	Concentration	Substances	Concentration
12M HCl	0.9 mL/L	$\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$	20 mg/L
H_3BO_3	60.0 mg/L	$\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$	100 mg/L
ZnCl_2	70.0 mg/L	$\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$	200 mg/L
$\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$	25.0 mg/L	$\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$	35 mg/L

以上培养基均经过 121 °C、20 min 高温灭菌处理。

种子培养方法：按 1% (V/V) 的接种量将菌种接种到两瓶 100 mL 种子培养基内，于 37 °C、150 r/min 摇床培养 12 h。

发酵实验方法：将培养好的种子接种到装有 2 L 发酵培养基的 5 L 发酵罐中，整个发酵过程中通过温度控制设备将发酵液温度保持在 (37±0.5) °C 范围内，并使用 5 mol/L 的 NaOH 溶液保持发酵液的 pH 稳定在 (7±0.5) 范围内。整个发酵过程以 0.1 vvm 速率通入氮气以保持发酵过程的厌氧状态，并保持 250 r/min 的搅拌速率，使发酵液保持均匀状态，全程发酵时间 13 h。

1.3 发酵过程取样、光谱数据采集与吸光光度值 (OD) 数据采集

在发酵过程中每隔 1 h 对发酵液取样 1 次，

表 3 发酵培养基组成

Substances	Concentration	Substances	Concentration
Glycerol	60.00 g/L	MgCl ₂ ·6H ₂ O	0.26 g/L
KH ₂ PO ₄	1.36 g/L	(NH ₄) ₂ SO ₄	6.61 g/L
CaCl ₂	0.29 g/L	Trace elements B	5.00 mL/L
Citric acid	0.42 g/L	Yeast extract powder	2.00 g/L

表 4 微量元素 B 组成

Table 4 Compositions of the trace elements B

Substances	Concentration	Substances	Concentration
12 mol/L HCl	10.00 mL/L	CuCl ₂ ·2H ₂ O	0.47 g/L
ZnCl ₂	0.68 g/L	FeCl ₂ ·4H ₂ O	3.97 g/L
H ₃ BO ₃	0.06 g/L	MnCl ₂ ·4H ₂ O	0.17 g/L
CoCl ₂ ·6H ₂ O	0.47 g/L	Na ₂ MoO ₄ ·2H ₂ O	0.005 g/L

表 5 校正集与验证集数据分析

Table 5 Analysis on the calibration data set and validation data set

	Numbers of samples	Ratio (%)	OD	Mean	Variance
Calibration set	46	67.65	0.046 0–8.261 0	3.196 5	7.262 2
Validation set	22	32.35	0.061 2–8.048 3	3.148 9	6.988 6

为拓展模型的适用范围并提高建模的准确性，对未接菌时的培养基也进行取样和后续的分析处理。使用测量吸光光度值 (OD) 的方法获取发酵过程生物量的参考数据 (下文称为实测值)。吸光光度值 (OD) 用光密度法以蒸馏水为背景，适当稀释发酵液后在 650 nm 波长下测定。在取样的同时进行近红外光谱的在线采集。以分辨率 16，扫描次数 128 次 (约 50 s)，检测器增益 237.84，空气背景，吸收光谱的模式在 4 000–12 000 cm⁻¹ 的光谱范围上连续采集近红外光谱 3 次作为与该样品对应的近红外光谱数据。为对发酵过程进行全面的分析，对取出的样品进行离心处理得到上清液，用水浴锅保持上清液与发酵过程相同的温度 (37 °C) 并以相同的光谱采集条件进行近红外光谱的离线采集。

2 1,3-丙二醇发酵过程生物量在线动态监测模型建立

2.1 吸光光度值 (OD) 数据分析与处理

在上面介绍的实验条件下进行 5 个批次的发酵实验，得到 68 个样本用于在线动态监测模型的建立。随机的把样本点分为 3 份，取其中 1 份作为验证集，剩余的作为校正集。校正集与验证集的数据特性如表 5 所示。从表中可以看到在吸光光度值 (OD) 范围上校正集略大于验证集，在平均数、方差方面都比较相似，根据文献[17]可知这样的数据特性对建立准确可靠的多元校正模型具有重要意义。

2.2 近红外光谱数据分析与处理

在上文介绍的光谱采集条件下对 5 个批次的发酵实验进行光谱在线采集,得到原始光谱如图 2 所示。

该图为一个批次发酵实验并对含有大量噪声的 $4\ 000\text{--}4\ 800\ \text{cm}^{-1}$ 谱段进行删除处理后所得到的原始光谱图。结合发酵过程信息可以明显看出随着发酵时间的增长,光谱基线不断升高,这与随着发酵时间的增长生物量不断增长(吸光度值 (OD) 不断升高)是相关的。对比图 3 发酵过程上清液的光谱图可以看到,排除细胞影响后发酵过程物质浓度变化给光谱带来的影响较小,这说明在线采集光谱的基线变化主要由生物量变化导致(发酵过程中 pH , 温度等条件保持恒定,可以排除这些干扰的影响),由底物甘油和主产物 1,3-丙二醇等物质吸收带来的影响相对较小。这种现象产生的主要原因是生物量浓度的提高使发酵液对近红外光的散射能力增强,导致光谱的吸光度变高。放大分析光谱在不同波段上的敏感性可以发现,在谱段 $7\ 200\text{--}12\ 000\ \text{cm}^{-1}$ (下文称为谱段 A) 和谱段 $5\ 400\text{--}6\ 440\ \text{cm}^{-1}$ (下文称为谱段 B) 上光谱对生物量变化的敏感性更强。

同时 A、B 这两个谱段也基本涵盖了 C-H 等基团的吸收区域^[18],由于细胞本身也含有很多近红外光谱可以分析的 C-H、O-H 等基团,会在发酵过程中对光谱的变化产生一定的影响,因此对光谱作以下光谱预处理,以便减少由散射、仪器噪声等带来的干扰。本文主要使用的光谱预处理方法包括多元散射校正(MSC)^[19]、标准正态变量变换(SNV)^[19]、SG 平滑^[19]及一阶导数和二阶导数处理。经一阶导数和二阶导数处理后的光谱如图 4 和图 5 所示。从图中可以看到导数处

理把光谱的基线拉平,基本上消除了光谱的噪声。

2.3 建模方法

在对吸光度值 (OD) 数据和近红外光谱数据进行上文的分析与处理的基础上应用偏最小二乘算法^[20-22]建模。在偏最小二乘算法的建模过程中因子数的选择是一个关键的问题,本文使用了交叉有效性原则对模型的因子数进行选择。此原则主要考察因子的边际贡献率^[23],可以减

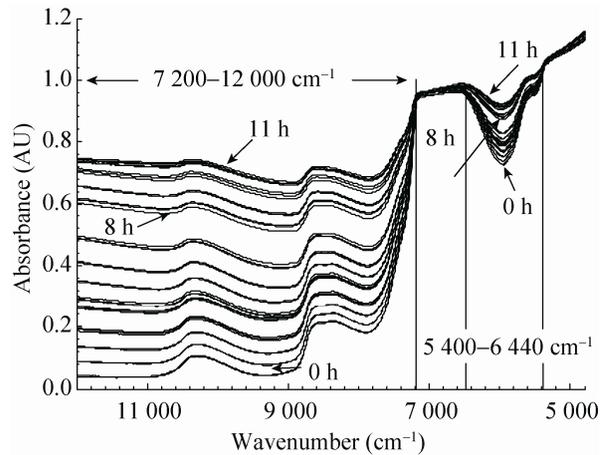


图 2 1,3-丙二醇发酵过程的原始光谱图

Fig. 2 Original spectra of a 1,3-propanediol fermentation process.

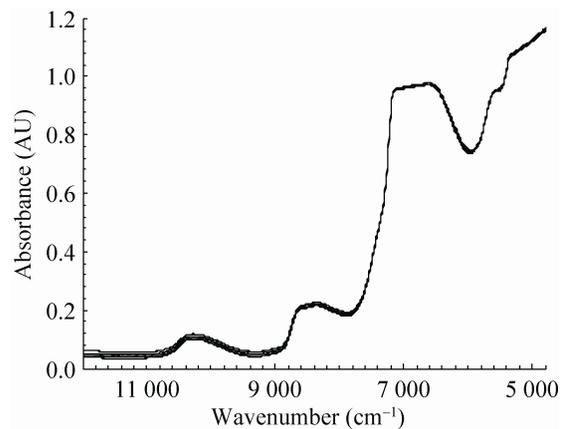


图 3 发酵过程离心上清液光谱图

Fig. 3 The original spectra of supernatant.

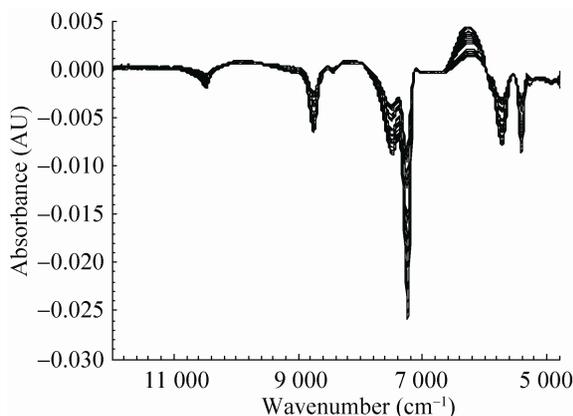


图 4 经一阶导数处理后的光谱图
Fig. 4 The first derivative spectra.

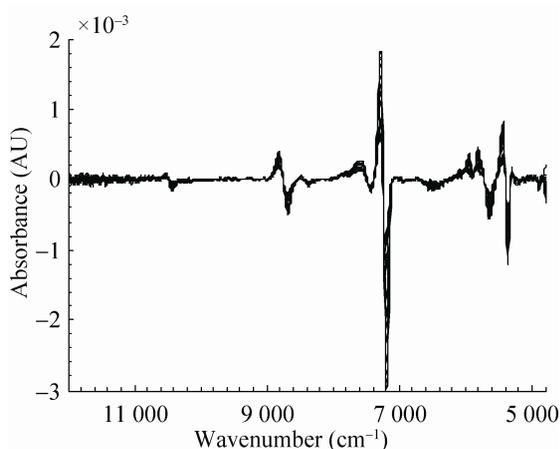


图 5 经二阶导数处理后的光谱图
Fig. 5 The second derivative spectra.

少选取因子数的数量,降低模型过拟合的概率,使模型更具可靠性。1,3-丙二醇发酵过程通常伴有噪声与干扰,如通气带来的气泡与强力搅拌带来的干扰等^[24]。因而需要可靠性好的在线动态监测模型来实现对该发酵过程的有效在线监测。

交叉有效性原则主要是通过考察取 h 个因子数时, Q_h^2 是否大于等于 0.097 5 这个阈值来决定此因子的取舍,当满足条件时考虑引进此因子,反之则不引进此因子。其公式如式(1)所示。

$$Q_h^2 = 1 - \frac{PRESS_h}{SS_{h-1}} \quad (1)$$

其中, h 为因子数, $PRESS_h$ 为取 h 个因子数时,利用除去样本点 i 的数据集建模后再对该样本点进行预测,并对每个样本点重复上述操作后得到的预测残差平方和,其公式如式(2)所示。 SS_h 为利用所有样本点,取 h 个因子数建模时的误差平方和,其公式如式(3)所示。

$$PRESS_h = \sum_{i=1}^n (y_i - \hat{y}_{h(-i)})^2 \quad (2)$$

$$SS_h = \sum_{i=1}^n (y_i - \hat{y}_{hi})^2 \quad (3)$$

其中, n 为样本数量, h 为因子数, y_i 为实测数据, $\hat{y}_{h(-i)}$ 为建模时删去样本点 i 后,取 h 个因子数建模后,再用所建立的模型计算的 y_i 的预测值, \hat{y}_{hi} 为用全部样本点并取 h 个因子数建模后第 i 个样本的预测值。

本文使用交叉验证均方根误差 ($RMSECV$)、校正均方根误差 ($RMSEC$) 和预测均方根误差 ($RMSEP$) 来对模型进行评价,其公式如式(4)、(5)和(6)所示。

$$RMSECV = \sqrt{\frac{PRESS_h}{n_c}} \quad (4)$$

其中, n_c 为校正集样本数量, $PRESS_h$ 为由公式(2)计算得到的校正集数据的预测残差平方和。

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n_c} (\hat{y}_{ci} - y_{ci})^2}{n_c - h - 1}} \quad (5)$$

其中, n_c 为校正集样本数量, \hat{y}_{ci} 为校正集样本的预测值, y_{ci} 为校正集样本的实测值, h 为模型所选择的因子数。

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n_p} (\hat{y}_{pi} - y_{pi})^2}{n_p}} \quad (6)$$

其中, n_p 为验证集样本数量, \hat{y}_{pi} 为验证集样本的预测值, y_{pi} 为验证集样本的实测值。

3 结果与分析

由对近红外光谱的敏感性分析可知,谱段 A 和谱段 B 与发酵过程中生物量变化有明显的相关关系,且可以看到相比谱段 B,谱段 A 的光谱吸光度变化范围更大,因而与发酵过程中生物量变化的相关关系更加明显,所以考虑以谱段 A 为核心谱段来进行谱段选择的建模比较。分别采用了谱段 A、谱段 A+B、全谱段(4 800–12 000 cm^{-1}) 这 3 种谱段选择方式,并在无预处理,多元散射校正 (MSC),标准正态变量变换 (SNV),SG 平

滑条件下一阶导数和二阶导数的光谱预处理条件下进行了建模比较。根据不同的谱段选择得到的建模结果分别列在表 6、7、8 中。表中使用 $RMSECV$ 作为评价模型预测偏差的标准 (该值越小说明模型的预测误差越小^[25]),使用 $RMSEP$ 与 $RMSEC$ 的比值作为评价模型过拟合和欠拟合程度的标准 (该值过大易过拟合,过小易欠拟合,通常认为该值在 0.8–1.2 范围内时,所建立的模型是可以接受的^[25]),使用测定系数 (R^2) 作为评价校正集数据实测值与预测值相关性的标准 (该值越接近 1 说明相关性越好^[25])。

表 6 采用谱段 A 和不同光谱预处理方法的建模结果

Table 6 Modeling results by using the spectra band A and different spectral pretreatment methods

Model	Pretreatment method	$RMSECV$	$RMSEP$	$RMSEP/RMSEC$	Factors	R^2
1	None	0.561 5	0.449 9	0.807 8	1	0.957 0
2	MSC	0.377 7	0.322 0	0.896 8	3	0.982 4
3	SNV	0.363 0	0.327 3	0.945 4	3	0.983 6
4	The 1st	0.310 8	0.293 8	1.086 2	4	0.990 1
5	The 2nd	0.352 2	0.379 6	1.199 5	4	0.986 4

表 7 采用谱段 A+B 和不同光谱预处理方法的建模结果

Table 7 Modeling results by using the spectra band A+B and different spectral pretreatment methods

Model	Pretreatment method	$RMSECV$	$RMSEP$	$RMSEP/RMSEC$	Factors	R^2
6	None	0.560 8	0.449 4	0.808 1	1	0.957 1
7	MSC	0.384 9	0.349 7	0.955 9	2	0.981 6
8	SNV	0.369 2	0.342 6	0.968 1	2	0.982 7
9	The 1st	0.337 7	0.341 5	1.039 8	2	0.985 1
10	The 2nd	0.299 8	0.283 6	1.048 4	4	0.990 1

表 8 采用全谱段和不同光谱预处理方法的建模结果

Table 8 Modeling results by using all the wave numbers and different spectral pretreatment methods

Model	Pretreatment method	$RMSECV$	$RMSEP$	$RMSEP/RMSEC$	Factors	R^2
11	None	0.288 4	0.267 2	0.961 9	4	0.988 4
12	MSC	0.349 0	0.328 4	0.955 2	1	0.983 6
13	SNV	0.352 5	0.329 8	0.949 8	1	0.983 3
14	The 1st	0.362 4	0.383 1	1.091 7	2	0.983 0
15	The 2nd	0.329 3	0.324 7	1.076 9	4	0.987 7

分析表 6-8 中的数据可以看到,表中所得到的 15 个模型的参数 $RMSEP/RMSEC$ 均在 0.8-1.2 范围内,说明这些模型都没有产生过拟合或欠拟合,但模型 1、6、7、8、9、12、13 和 14 通过采用交叉有效性原则选择出的因子数较少,可能会严重影响模型的预测性能。为此,对因子数少于 3 的模型进行因子数的扩展,以验证是否该原则选择出较少的因子数会明显影响模型预测性能,扩展因子数后的建模结果如表 9、10、11 所示。

综合分析表 6-11 可以看到,进行因子数扩展后,除对模型 1 和 6 稍有一些精度提高外,对其他模型的性能提高均较小,而且对所建立的 27 个模型进行比较可以看到,未经因子数扩展的模型

11 所得的交叉验证均方根误差最小且 R^2 和 $RMSEP/RMSEC$ 均接近 1。可知该模型具有较小的预测误差且具有较小的过拟合与欠拟合风险。

因此,在排除了选择较少因子数导致模型性能严重降低的风险后,本文最终确定模型 11 为最佳模型。模型 11 的校正集和验证集的预测值(此处的预测值为 1,3-丙二醇发酵过程生物量在线监测模型给出的结果,后文同)与实测值(此处的实测值为使用传统检测方法得到的吸光光度值 (OD) ,下文同)比较如图 6 和图 7 所示。

从图 6 和图 7 中可以看到,所选择的最佳模型校正集与验证集的预测值与实测值之间具有较小误差,且相关性较好。

表 9 在谱段 A 条件下利用扩展因子数的建模结果

Table 9 Modeling results by using the spectra band A and the expansion factors

The former model	Model	Pretreatment method	$RMSECV$	$RMSEP$	$RMSEP/RMSEC$	Factors	R^2
1	16	None	0.540 4	0.453 0	0.845 0	2	0.960 4
	17		0.385 4	0.339 6	0.904 3	3	0.980 7

表 10 在谱段 A+B 条件下利用扩展因子数的建模结果

Table 10 Modeling results by using the spectra band A+B and the expansion factors

The former model	Model	Pretreatment method	$RMSECV$	$RMSEP$	$RMSEP/RMSEC$	Factors	R^2
6	18	None	0.529 3	0.451 3	0.864 3	2	0.962 5
	19		0.433 7	0.374 7	0.878 6	3	0.975 1
7	20	MSC	0.343 9	0.304 5	0.931 8	3	0.985 4
8	21	SNV	0.339 1	0.311 7	0.964 9	3	0.985 7
9	22	The 1st	0.324 3	0.291 3	0.926 4	3	0.986 5

表 11 在全谱段条件下利用扩展因子数的建模结果

Table 11 Modeling results by using all the wave numbers and the expansion factors

The former model	Model	Pretreatment method	$RMSECV$	$RMSEP$	$RMSEP/RMSEC$	Factors	R^2
12	23	MSC	0.335 2	0.312 0	0.953 5	2	0.985 3
	24		0.321 2	0.299 4	1.000 2	3	0.987 8
13	25	SNV	0.336 6	0.312 5	0.951 2	2	0.985 1
	26		0.321 6	0.301 8	1.002 1	3	0.987 6
14	27	The 1st	0.330 1	0.291 7	0.922 0	3	0.986 3

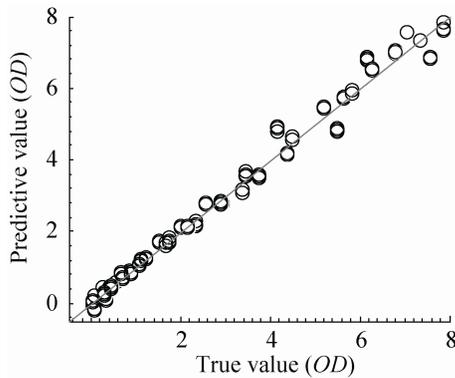


图 6 模型 11 校正集的预测值与实测值比较
Fig. 6 Predictive and true values of calibration set of model 11.

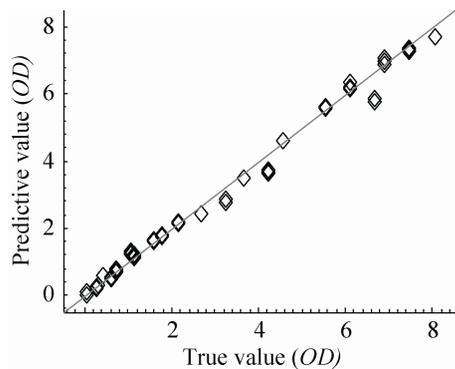


图 7 模型 11 验证集的预测值与实测值比较
Fig. 7 Predictive and true values of validation set of model 11.

选择的最佳模型是利用无光谱预处理的全谱段 ($4\ 800\text{--}12\ 000\ \text{cm}^{-1}$) 近红外光谱得到的。采用全谱段建模说明,在光谱变化不明显的谱段也包含了建模所需要的必要信息。采用含有更多细胞散射信息的原始光谱进行建模说明,细胞散射信息相比于细胞含氢基团的光谱吸收信息,对于准确和可靠的建模具有不可忽视的作用。

4 实验验证

通过两个批次的 1,3-丙二醇发酵实验 (初始甘油浓度为 $60\ \text{g/L}$ 和 $40\ \text{g/L}$), 对本文建立的 1,3-丙

二醇发酵过程生物量在线监测模型的监测效果和模型的适用性进行验证。在实验中,使用本文建立的模型可以实现对生物量的实时在线监测 (可 $1\ \text{min}$ 检测 1 次数据), 相比测量吸光光度值 (OD) 的传统检测方法 ($1\ \text{h}$ 检测 1 个数据点) 具有明显的实时性优势。从图 8 和图 9 中可以看到,应用本文建立的模型可以得到与传统方法实测值相比趋势一致且误差较小的实时监测值,并且可以看到本文所建立的模型在培养基物质浓度变化的情况下,也可得到稳定且误差较小的实时监测结果。

模型的在线监测数据及与发酵实验实测数据的比较如表 12 所示。对于甘油浓度 $60\ \text{g/L}$ 和 $40\ \text{g/L}$ 的验证实验,模型的预测值与实测值之间的测定系数 (R^2) 分别为 0.9976 和 0.9928 ,说明预测值与实测值之间具有很好的相关性。预测均方根误差 (RMSEP) 分别为 0.3416 和 0.2743 ,说明所建立的模型能够较准确的实现对 1,3-丙二醇发酵过程生物量的实时在线监测。

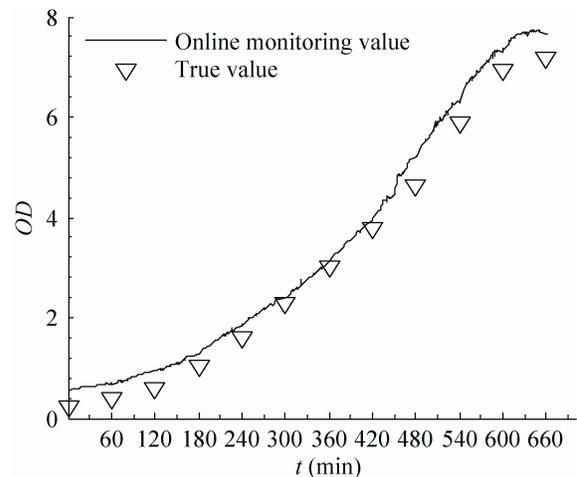


图 8 光谱模型预测值与实测值对比 (甘油 $60\ \text{g/L}$)
Fig. 8 Comparison between spectra model prediction and the real measurement (Glycerol $60\ \text{g/L}$).

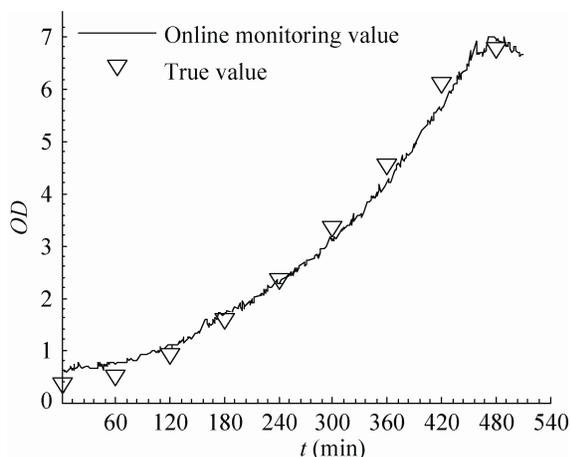


图 9 光谱模型预测值与实测值对比 (甘油 40 g/L)
Fig. 9 Comparison between spectra model prediction and the real measurement (Glycerol 40 g/L).

表 12 1,3-丙二醇发酵过程生物量在线监测模型预测结果

Table 12 Model predictive results for on-line monitoring the biomass of 1, 3-propanediol fermentation

Glycerol 60 g/L	True value	Monitoring value
OD	0.229 3–7.170 2	0.529 8–7.675 7
R^2		0.997 6
RMSEP		0.341 6
Glycerol 40 g/L	True value	Monitoring value
OD	0.338 6–6.941 0	0.532 3–6.959 0
R^2		0.992 8
RMSEP		0.274 3

5 结论

对发酵过程生物量进行在线监测能有效地改善发酵过程的调控优化。本文设计了一个基于近红外光谱分析技术的 1,3-丙二醇发酵过程在线监测实验平台,通过对采集的光谱数据进行预处理以及不同谱段的敏感性分析,应用偏最小二乘算法对发酵过程中生物量建立了在线动态监测模型。根据选定的最佳模型,发现细胞散射信息相对于细胞含氢基团的光谱吸收信息,对于准确建模具有重要作用。通过实验测试与离线方法

的检测结果对比,验证了本文建立的模型能对该发酵过程生物量进行有效的实时在线监测,相比传统的监测方法具有明显的实时性优势。

REFERENCES

- [1] Zhuang YP, Chen HZ, Xia JY, et al. Progress in industrial bioprocess engineering in China. *Chin J Biotech*, 2015, 31(6): 778–796 (in Chinese). 庄英萍, 陈洪章, 夏建业, 等. 我国工业生物过程工程研究进展. *生物工程学报*, 2015, 31(6): 778–796.
- [2] Lourenço ND, Lopes JA, Almeida CF, et al. Bioreactor monitoring with spectroscopy and chemometrics: a review. *Anal Bioanal Chem*, 2012, 404(4): 1211–1237.
- [3] Landgrebe D, Haake C, Höpfner T, et al. On-line infrared spectroscopy for bioprocess monitoring. *Appl Microbiol Biotechnol*, 2010, 88(1): 11–22.
- [4] Chu XL, Yuan HF, Lu WZ. The on-line near infrared spectroscopy process analytical technique and its applications. *Mod Sci Instrum*, 2004, (2): 3–21 (in Chinese). 褚小立, 袁洪福, 陆婉珍. 在线近红外光谱过程分析技术及其应用. *现代科学仪器*, 2004, (2): 3–21.
- [5] Arnold SA, Gaensakoo R, Harvey LM, et al. Use of at-line and *in-situ* near-infrared spectroscopy to monitor biomass in an industrial fed-batch *Escherichia coli* process. *Biotechnol Bioeng*, 2002, 80(4): 405–413.
- [6] Hall JW, McNeil B, Rollins MJ, et al. Near-infrared spectroscopic determination of acetate, ammonium, biomass, and glycerol in an industrial *Escherichia coli* fermentation. *Appl Spectrosc*, 1996, 50(1): 102–108.
- [7] Crowley J, Arnold SA, Wood N, et al. Monitoring a high cell density recombinant *Pichia pastoris*, fed-batch bioprocess using transmission and reflectance near infrared spectroscopy. *Enzyme Microb Technol*, 2005, 36(5/6): 621–628.
- [8] Cervera AE, Petersen N, Lantz AE, et al. Application of near-infrared spectroscopy for

- monitoring and control of cell culture and fermentation. *Biotechnol Progr*, 2009, 25(6): 1561–1581.
- [9] Lu WZ. *Modern Near Infrared Spectroscopy Analytical Technology*. 2nd ed. Beijing: China Petrochemical Press, 2006: 1–11 (in Chinese).
陆婉珍. 现代近红外光谱分析技术. 2 版. 北京: 中国石化出版社, 2006: 1–11.
- [10] Huang HB, Yu HY, Xu HR, et al. Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: a review. *J Food Eng*, 2008, 87(3): 303–313.
- [11] Roggo Y, Chalou P, Maurer L, et al. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. *J Pharm Biomed Anal*, 2007, 44(3): 683–700.
- [12] Chu XL, Lu WZ. Research and application progress of near infrared spectroscopy analytical technology in China in the past five years. *Spectrosc Spect Anal*, 2014, 34(10): 2595–2605 (in Chinese).
褚小立, 陆婉珍. 近五年我国近红外光谱分析技术研究与应用进展. *光谱学与光谱分析*, 2014, 34(10): 2595–2605.
- [13] Watson DG. *Pharmaceutical analysis : a textbook for pharmacy students and pharmaceutical chemists*. 2nd ed. London: Elsevier Press, 2005: 128–134.
- [14] Gui YL, Liang JB, Ma L, et al. Model construction for lactate concentration prediction in glutamate fermentation process relying on near-infrared spectroscopy technology. *Food Ferment Ind*, 2014, 40(8): 1–6 (in Chinese).
桂勇利, 梁静波, 马雷, 等. 基于近红外技术谷氨酸发酵过程中乳酸浓度预测模型的建立. *食品与发酵工业*, 2014, 40(8): 1–6.
- [15] Guo WL, Du YP, Zhou YC, et al. At-line monitoring of key parameters of nisin fermentation by near infrared spectroscopy, chemometric modeling and model improvement. *World J Microbiol Biotechnol*, 2012, 28(3): 993–1002.
- [16] Cruz MV, Sarraguça MC, Freitas F, et al. Online monitoring of P(3HB) produced from used cooking oil with near-infrared spectroscopy. *J Biotechnol*, 2015, 194: 1–9.
- [17] Lu WZ, Yuan HF, Chu XL. *Near Infrared Spectroscopy Instrument*. Beijing: Chemical Industry Press, 2010: 17–61 (in Chinese).
陆婉珍, 袁洪福, 褚小立. 近红外光谱仪器. 北京: 化学工业出版社, 2010: 17–61.
- [18] Workman J, Weyer L. *Practical Guide and Spectral Atlas for Interpretive Near-Infrared Spectroscopy*. 2nd ed. New York: CRC Press, 2012: 195–226.
- [19] Rinnan Å, Van Den Berg F, Engelsen SB. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends Anal Chem*, 2009, 28(10): 1201–1222.
- [20] Haaland DM, Thomas EV. Partial least-squares methods for spectral analyses: 2. application to simulated and glass spectral data. *Anal Chem*, 1988, 60(11): 1202–1208.
- [21] Gemperline P. *Practical guide to chemometrics*. 2nd ed. New York: Crc Press, 2006: 167–216.
- [22] Wold S, Sjöström M, Eriksson L. PLS- regression: a basic tool of chemometrics. *Chemom Intell Lab Syst*, 2001, 58(2): 109–130.
- [23] Wang HW. *Partial least-Squares Regression- Method and Applications*. Beijing: National Defense Industry Press, 1999: 155–157 (in Chinese).
王惠文. 偏最小二乘回归方法及其应用. 北京: 国防工业出版社, 1999: 155–157.
- [24] Menzel K, Zeng AP, Biebl H, et al. Kinetic, dynamic, and pathway studies of glycerol metabolism by *Klebsiella pneumoniae*, in anaerobic continuous culture: I. The phenomena and characterization of oscillation and hysteresis. *Biotechnol Bioeng*, 1996, 52(5): 549–60.
- [25] Yan YL, Chen B, Zhu DZ, et al. *Near Infrared Spectroscopy-Principles, Technologies and Applications*. Beijing: China Light Industry Press, 2013: 162–175 (in Chinese).
严衍禄, 陈斌, 朱大洲, 等. 近红外光谱分析的原理、技术与应用. 北京: 中国轻工业出版社, 2013: 162–175.

(本文责编 郝丽芳)