

## 发酵液中乙醇含量的近红外光谱 NIRS 定量分析与验证

李燕萍, 许宏贤, 钱莹, 段钢

(丹尼斯克能科, 江苏 无锡, 214082)

**摘 要** 应用近红外光谱技术(NIRS), 采用偏最小二乘法(PLS)建立发酵液中乙醇含量定量分析预测模型, 并比较发酵液原液和发酵液上清液 NIRS 预测模型对预测乙醇含量的效果差异。结果表明, 发酵液原液所建立的乙醇含量和发酵上清液所建立的乙醇含量预测模型校正误差(SEC)、交叉检验标准误差(SECV)、校正相关系数(RSQ)、交叉验证相关系数(1-VR)皆有很好的关联性。由此可见, 通过内部交叉检验和外部验证, 用近红外光谱法测定发酵液中的乙醇含量具有很高的准确度, 为近红外光谱法快速监测乙醇发酵提供了新的方法。

**关键词** 近红外光谱, 发酵, 乙醇含量, 预测模型

目前测定发酵液中乙醇含量的方法有填充柱气相色谱法、密度瓶法、酒精计法、高效液相色谱法等。密度瓶法和酒精计法需先将样品溶液蒸馏后再测定, 较为繁琐。填充柱法由于在运行中存在严重的涡流扩散, 传质阻力也较大, 影响了柱效的提高, 因此逐渐被淘汰<sup>[1,2]</sup>。高效液相色谱法则需要先将样品进行前期处理, 且测定一个样品时间较长, 测量的结果由于前处理可能造成一定的误差。

近红外光谱分析(near infrared reflectance spectroscopy, NIRS)是利用有机化学物质在其近红外光谱区内的光学特性快速估测样品中的一项或多项化学成分含量的新技术<sup>[3]</sup>。近红外光是指波长在780~2 500 nm的红外光。其原理主要是由分子的振动能级变化产生的, 它是由与H有关的基团, 如O-H、N-H、C-H键的伸缩振动产生<sup>[4]</sup>。基于以上原理, 近红外光谱分析具有分析过程中样品基本不需处理、不消耗和破坏样品、仪器自身无污染等特点和优势, 堪称是绿色分析技术的代表。因此, 近红外光谱在分析测试领域, 特别是作为过程分析仪器, 在在线分析和工业控制领域正发挥着越来越重要的作用。

利用近红外光谱法测定发酵液中乙醇含量的研究在国内、外报道甚少, 一般都是利用NIRS测定白酒、啤酒等中的乙醇含量, 而直接测定谷物发酵液中乙醇含量的研究较少。此外, 在直接测定的同时比较发酵液原液与上清液中乙醇含量关系并建立其测定模型的研究尚未见报道。

本研究利用NIRS可快速测定发酵液中乙醇的

含量, 并比较发酵液原液和发酵液上清液 NIRS 定标模型对预测乙醇含量的效果差异, 获得满意的结果, 与传统的比重法、气相色谱法及高效液相色谱法相比, 更快捷、准确、经济和环保<sup>[5]</sup>。

## 1 材料与方法

### 1.1 仪器

FOSS 公司 InfraXact™近红外光谱仪。

### 1.2 波段指标吸收峰

波长范围 570~1 880 nm, 发酵液中乙醇在此波段都有较强的吸收和稳定的信号。

### 1.3 测量方式

反射、漫反射。

### 1.4 进样系统

采用 FOSS 公司的专用样品杯, 该杯表面有特殊金镀层处理, 能增强反射光谱。

### 1.5 软件系统

WinISI III 定量分析方式为偏最小二乘法(PLS)。

### 1.6 样品来源

实验室自制, 选择 120 个发酵样品。

### 1.7 样品成分含量液相色谱法测定。

样品分析采用 HPLC 法(Agilent 1100); 色谱柱 Rezex ROA-Organic Acid; 示差检测器; 流动相为 0.005 mol/L H<sub>2</sub>SO<sub>4</sub>, 流速 0.6 mL/min, 柱温 6℃, 进样量 20 μL。

### 1.8 定标样品的收集

从光谱特征出发, 对 120 个样品的图谱进行聚类分析, 类别数为 60, 每 1 类取 1 个图谱作为标准样品, 标准样品集共 100 个样品, 标准样品集的乙醇含

第一作者: 学士, 工程师(段钢博士为通讯作者)。

收稿日期: 2007-12-29

量在 3.07%~8.64%，原 120 个样品的乙醇含量在 2.95%~8.69%，基本能覆盖全部样品的乙醇含量范围，这样可以保证在预测时避免预测方程外推而产生较大的误差。

1.9 样品的处理

(a)直接取 3 mL 发酵液原液；(b)取 3 mL 发酵液原液，离心取其上清液。

1.10 光谱采集及数据录入

分别将 a 和 b 样品倒于专用杯中，加上特殊金反射盖。在波长 570~1 780 nm，每个样品图谱采用扫描 3 次，平均得到 1 条近红外光谱，并转化成  $\log(1/R)$  储存于计算机内，如图 1 和图 2 所示。

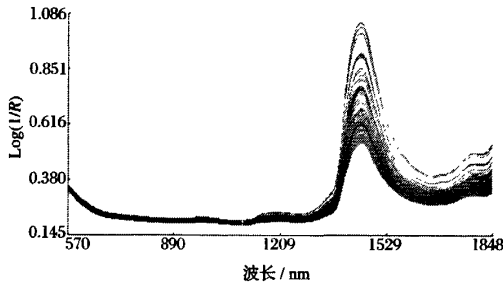


图 1 发酵液原液的近红外光谱图

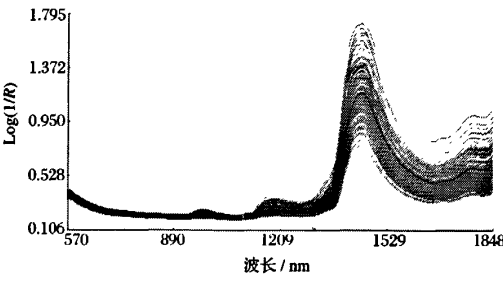


图 2 发酵液上清液的近红外光谱图

1.11 光谱的数学处理

光谱分析采用 FOSS 公司提供的 WinISS III 软件系统，用“标准正常化处理和散射处理”数学处理方法，对所收集的光谱进行处理和比较，了解光谱的特性并剔除异常光谱，并使用软件的自动条件筛选功能（软件优化条件：马氏距离  $GH=3.0$ ，邻居距离  $NH=0.6$ ），去除离散性大的样品和性质相似的样品，对样品群体的代表性进行优化。参数设置：导数处理

(Derivative)=1，光谱点间隔(GAP)=4，一次平滑(Smooth1)=4，二次平滑(Smooth2)=1。

1.12 定标建模

在 WinISI 软件系统的化学计量方法中，采用 PLS 最小二乘法回归技术和定标样品全光谱建立模型，将样品的光谱数据和化学值进行关联，确定二者间的定性或定量关系。样品集 a 所建立的模型为 A，样品集 b 所建立的模型为 B。以近红外分析结果为 X 轴，色谱分析结果为 Y 轴，通过线性回归对预测集进行分析。

1.13 定标方程的验证

收集适当的未知样品，用近红外定标方程分析出数据，作为 NIR 数据，同时用色谱法进行实验，所得的数据为 LAB 数据，用数理统计方法对这些数据进行处理，以验证方程的准确度与精密度。样品分为定标集和验证集，先用定标样品集建立定标模型并作内部交叉验证，再用验证样品集进行外部验证，最后根据定标相关系数(RSQ)，交叉验证相关系数(1-VR)，定标标准误差(SEC)和交叉检验标准误差(SECV)等指标确定最优校正模型。

2 结果与分析

2.1 预测模型参数结果

预测模型的优劣取决于所建模型相关系数的高低及误差的大小。较高的相关系数和较低的标准误差对一个可靠的预测模型是十分必要的，而定标样品的数量及含量范围则决定了预测模型的适用范围。

通常用定标相关系数(RSQ)、交叉验证相关系数(1-VR)、定标标准误差(SEC)和交叉检验标准误差(SECV)来衡量所建模型的好坏。定标集标准误差(SEC)的大小反映了预测方程对自身的预测能力。定标集交互验证的标准偏差(SECV)的大小则反映了预测方程的预测能力。(RSQ)和(1-VR)越大，(SEC)和(SECV)越小，则所建模型越好。

在建立预测模型时，需对光谱进行预处理，这样可消除各方面因素对光谱信息的影响。对光谱进行不同的数学处理和散射处理后对定标效果的影响如表 1~表 4 所示。

表 1 不同的数学处理对发酵液原液乙醇含量定标集 A 的影响

成 分	数学处理	模型 A			
		定标标准误差 SEC	定标相关系数 RSQ	交叉检验标准误差 SECV	交叉验证相关系数 1-VR
乙 醇	1,4,4,1	0.127 2	0.993 0	0.154 1	0.989 8
	2,4,4,1	0.138 6	0.991 6	0.173 7	0.986 9
	1,7,7,1	0.141 5	0.991 3	0.164 5	0.988 4

注：回归分析采用 PLS，光谱校正采用 SNV+Detrend。

表 2 不同的数学处理对发酵液上清液乙醇含量定标集 B 的影响

成 分	数学处理	模型 B			
		定标标准误差 SEC	定标相关系数 RSQ	交叉检验标准误差 SECV	交叉验证相关系数 1-VR
乙 醇	1,4,4,1	0.113 6	0.994 7	0.131 7	0.992 9
	2,4,4,1	0.118 6	0.994 1	0.132 9	0.992 6
	1,7,7,1	0.123 6	0.993 5	0.133 4	0.992 3

注：回归分析采用 PLS，光谱校正采用 SNV+Detrend。

表 3 不同的散射校正方法对发酵液原液乙醇含量定标集 A 的影响

成 分	散射处理	模型 A			
		定标标准误差 SEC	定标相关系数 RSQ	交叉检验标准误差 SECV	交叉验证相关系数 1-VR
乙 醇	SNV+Detrend	0.127 2	0.993 0	0.154 1	0.989 8
	SNV only	0.136 3	0.991 9	0.172 4	0.987 2
	Standard Msc	0.136 5	0.991 4	0.171 7	0.981 5

注：回归分析采用 PLS，数学处理采用 1,4,4,1。

表 4 不同的散射校正方法对发酵液上清液乙醇含量定标集 B 的影响

成 分	散射处理	模型 B			
		定标标准误差 SEC	定标相关系数 RSQ	交叉检验标准误差 SECV	交叉验证相关系数 1-VR
乙 醇	SNV+Detrend	0.113 6	0.994 7	0.131 7	0.992 9
	SNV only	0.113 6	0.994 1	0.131 9	0.992 6
	Standard Msc	0.131 4	0.992 6	0.140 0	0.991 6

注：回归分析采用 PLS，数学处理采用 1,4,4,1。

如表 1 和表 2 中所示，对发酵液原液和发酵液上清液中乙醇含量，采用 3 种数学处理方法，显示出不同的预处理效果，从表 1 和表 2 的数据中看出，采用 1,4,4,1 组合，其标准误差 (SEC、SECV) 最小，定标相关系数 (RSQ) 和 (1-VR) 也最大。

如表 3 和表 4 中所示，采用不同的散射处理会产生不同的效果，其“SNV+Detrend” (标准正态变量转换法结合趋势变换法) 处理方法，稍好于另外 2 个处理方法。

2.2 不同形态的样品集的比较

对上述 2 个定标样品集 a(发酵液原液)、b(发酵液上清液) 分别建立预测模型 A 和 B，采用最佳的优化组合，“SNV+Detrend” 散射处理和 1,4,4,1 数学处理组合，其预测模型参数如表 5、表 6 所示。从表 5 和表 6 的定量分析数据可见，此 2 个乙醇预测方程有很高的相关系数 (RSQ) 0.9930 和 0.9947；2 个乙醇的预测模型都有较低的预测标准误差 (SEC) 为 0.1272 和 0.1136。

表 5 发酵液原液乙醇含量定标集 A 模型统计参数

成 分	模型 A					
	最大值	最小值	定标标准误差 SEC	定标相关系数 RSQ	交叉检验标准误差 SECV	交叉验证相关系数 1-VR
乙 醇	8.641 3	3.071 0	0.127 2	0.993 0	0.154 1	0.989 8

注：回归分析采用 PLS，光谱校正“SNV+Detrend” 数学处理采用 (1,4,4,1)。

表 6 发酵上清液乙醇含量定标集 B 模型统计参数

成 分	模型 B					
	最大值	最小值	定标标准误差 SEC	定标相关系数 RSQ	交叉检验标准误差 SECV	交叉验证相关系数 1-VR
乙 醇	8.641 3	3.071 0	0.113 6	0.994 7	0.131 7	0.992 9

注：回归分析采用 PLS，光谱校正“SNV+Detrend” 数学处理采用 (1,4,4,1)。

从图 3 和图 4 可以看出，2 个预测模型对定标集样品 a、b，近红外光谱预测的值与其相应的真值的吻合程度很好，样品比较集中地分布在拟和线附近。从表 5、表 6 看出，A、B 模型的 (SECV) 值皆较小，说明这 2 个乙醇预测模型的预测效果较好。

2.3 模型预测效果的验证

通常采用外部验证的方法，对所建模型的质量或实际预测效果进行评价。所谓外部验证是指用另一批完全独立的、化学成分已知的验证集样品对所建模型的实际预测能力进行验证。选取 20 份样品，分别对创建的 2 个定标模型测试，并将 NIRS 的测定值与色谱法实测值作比较，如表 7、表 8 所示。

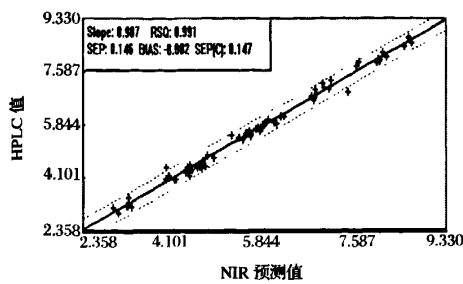


图3 发酵液中乙醇含量定标模型 A  
内部交叉证实相关图

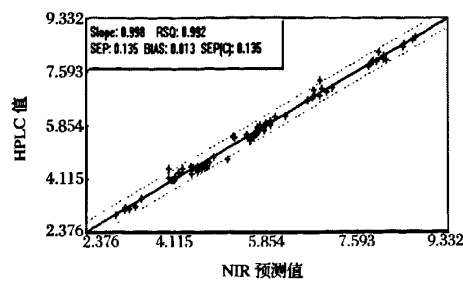


图4 发酵液中乙醇含量定标模型 B  
内部交叉证实相关图

表 7 发酵液原液乙醇含量验证样品集 A 模型统计参数

模型 A						
成 分	最大值	最小值	定标标准误差 SEC	定标相关系数 RSQ	交叉检验标准误差 SECV	交叉验证相关系数 1-VR
乙 醇	8.550 1	3.108 0	0.159 6	0.990 5	0.160 1	0.979 0

表 8 发酵液上清液乙醇含量验证样品集 B 模型统计参数

模型 B						
成 分	最大值	最小值	定标标准误差 SEC	定标相关系数 RSQ	交叉检验标准误差 SECV	交叉验证相关系数 1-VR
乙 醇	8.550 1	3.108 0	0.131 0	0.992 1	0.142 1	0.980 9

表 7 和表 8 的数据表明,A 模型的相关系数(RSQ)为 0.990 5,标准误差(SEC)值为 0.159 6;B 模型的相关系数(RSQ)为 0.992 1,标准误差(SEC)值为 0.131 0。

且从表 9 可以看出,A 模型的近红外预测值与色谱分析值误差在-0.201~0.220,B 模型的近红外预测值与色谱分析值误差在-0.086~0.100,说明 2 个定标模型可以较好地进行盲样测定。

表 9 验证集样品乙醇含量的真值,预测值,绝对误差

样品名	HPLC 值	模型 A		模型 B	
		NIR 预测值	绝对误差	NIR 预测值	绝对误差
1	3.157	3.278	-0.121	3.218	-0.061
2	3.954	4.029	-0.075	3.951	0.003
3	4.082	3.943	0.139	4.088	-0.006
4	4.165	4.004	0.161	4.093	0.072
5	3.184	3.172	0.012	3.138	0.046
6	5.553	5.487	0.066	5.464	0.089
7	4.079	4.107	-0.028	4.046	0.033
8	4.073	4.137	-0.064	4.089	-0.016
9	5.868	6.069	-0.201	5.922	-0.054
10	6.139	6.208	-0.069	6.204	-0.065
11	5.916	5.870	0.046	5.833	0.083
12	6.106	6.145	-0.039	6.014	0.092
13	6.003	5.902	0.101	5.924	0.079
14	7.026	7.083	-0.057	6.926	0.100
15	6.800	6.734	0.066	6.886	-0.086
16	8.087	8.200	-0.113	8.101	-0.014
17	7.893	7.673	0.220	7.869	0.024
18	8.406	8.529	-0.123	8.465	-0.059
19	7.944	8.062	-0.118	7.970	-0.026
20	8.546	8.543	0.003	8.552	-0.006

上述内部交叉验证和外部验证结果证明,利用发酵液原液和发酵液上清液所建预测模型的相关系数均较高,预测标准误差均较小,因此它们均可应用于乙醇含量分析;相对而言,发酵液上清液所建模型的预测准确性比发酵液模型的预测准确性稍高。

### 3 结 论

(1)应用近红外光谱分析技术和偏最小二乘法(PLS),选择最佳的光谱处理方法建立了定量测定发酵液中的乙醇含量的数学模型。内部交叉验证预测值和真值之间的相关系数(RSQ)分别为 0.993 0(发酵液原液)和 0.994 7(发酵液上清液),相应的标准误差(SEC)分别为 0.127 2 和 0.113 6;外部验证预测值和真值之间的 RSQ 值分别为 0.990 5(发酵液原液)和 0.992 1(发酵液上清液),其相对应的标准误差 SEC 分别为 0.159 6 和 0.131 0。说明发酵液原液和发酵液上清液 NIRS 数学模型均具有较高的预测准确性,可应用于酒精发酵在线控制。同时对不同样品形态所建立的数学模型的效果进行了比较。在定标样品集相同的情况下,发酵上清液模型的预测准确性比发酵液原液的稍高。所以在实际应用近红外技术时,可以根据不同的研究目的,选择利用不同的预测模型。

(2)近红外分析属于“二手分析方法”,其可靠程度依赖于实验室的分析方法,实验室分析测试数据的

准确性将直接决定建立模型的可靠性与适用性;且参与定标建模的样品光谱应具有相似性又要能代表光谱间最大差异,应覆盖实际应用所分析样品的所有变异性。稳定性好,适应性强的预测模型,才能得到可靠的预测分析结果。

(3)近红外光谱技术用于检测发酵液中乙醇的含量,具有速度快、准确度高的优点,表明该方法测量发酵液中的乙醇含量是可行的。这种方法可以及时地根据测定数据在生产过程中实时监控,对指导生产过程有重要的意义。

## 参 考 文 献

1 肖新亮. 实用分析化学(修订版)[M]. 天津:天津大学出版社,2000. 307

- 2 刘虎威. 气相色谱方法及应用[M]. 北京:化学工业出版社,2000. 19,60
- 3 张子仪,陈雪秀,任 鹏,等. 近红外光谱分析技术[M]. 北京:中国农业科技出版社,1992
- 4 丁丽敏,计 成,戒 易,等. 近红外光谱分析技术及其在评定饲料营养价值中的应用[J]. 中国饲料,1997,(1~2): 11~13
- 5 史月华. 傅立叶变换近红外光谱法同时定量分析水溶液中的乙醇果糖和葡萄糖[J]. 分析化学,2001,(10): 1 213~1 215

## New Models for NIRS Analysis of Alcohol in Fermentation Broth

Li Yanping, Xu Hongxian, Qian Ying, Duan Gang

(Genencor(Wuxi) Bioproducts Co Ltd. A Danisco Division, Wuxi 214082, China)

**ABSTRACT** With near infrared reflectance spectroscope (NIRS), models for quantitative analysis of alcohol content in different types of fermentation broth were established by using partial least square (PLS) method. The results indicated good correlations among the models by the standard errors (SEC) for calibration, SECV for cross validation and determination coefficients RSQ for calibration and 1-VR for calibration. The method of NIRS has high accuracy in predicting tests of cross validation and independent validation, thus offering a new and efficient way for quick analysis/on-line monitoring of fermentation alcohol production.

**Key words** near infrared reflectance spectroscopy, fermentation, alcohol, calibration model

(上接第 135 页)

## Determination of Acrylamide in Heat Treatment Food by Ultra High Performance Liquid Chromatography Tandem Mass Spectrometry

GeYu<sup>1</sup>, Zhao Xuli<sup>2</sup>, Chao Qiangguo<sup>1</sup>

1(National Food Quality Supervision and Inspection Center(Shanghai)/ Shanghai Institute of Quality Inspection and Technical Research, Shanghai 200233, China)

2 (East China University of Science & Technology Bioengineering Department, Shanghai 200237, China)

**ABSTRACT** This research built up a method of ultra high performance liquid chromatography tandem mass spectrometry for the determination of acrylamide in heated food. The pre-preparation steps include defat using hexane, distill with sodium chloride, extraction with ethyl acetate and solid phase extraction with Oasis HLB. What the separation of sample uses is ACQUITY UPLC HSS T<sub>3</sub> (1.8μm, 150mm×2.1mm), with 0.1% formic acid and methanol(the volume ratio is 90:10) as mobile phase. This method has high sensitivity and good repeatability through the validation of precision, repeatability and recovery, and it can satisfy trace analytical request. More than 121 kinds of heated food sold in the city are investigated after this method are established.

**Key words** UPLC-MS/MS, heated food, acrylamide, distribution