

MONITORAMENTO DA PRODUÇÃO DE ETANOL CELULÓSICO POR ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO

A.S. S. PINTO^{1,2}, S. C. PEREIRA², M. P. A. RIBEIRO¹ e C. S. FARINAS^{1,2}

¹Universidade Federal de São Carlos, Departamento de Engenharia Química.

²Empresa Brasileira de Pesquisa Agropecuária, Embrapa Instrumentação.

E-mail para contato: ariane_sbrice@hotmail.com

RESUMO - O monitoramento do processo de fermentação alcoólica é essencial para o controle de sua eficiência. Contudo, ainda não há métodos bem definidos que sejam rápidos e práticos. O presente estudo utilizou espectroscopia no infravermelho próximo (NIR) associada à análise multivariada para monitorar a produção de etanol celulósico a partir de bagaço e palha de cana-de-açúcar. Analisaram-se os espectros obtidos a partir de amostras sintéticas e reais. Os espectros foram modelados utilizando-se a Regressão por Mínimos Quadrados Parciais (PLS) através do *software* Matlab. A quantidade de variáveis latentes foi definida com base nos resultados da Validação Cruzada (*leave-one-out*). Diversos conjuntos de amostras foram analisados a fim de garantir o melhor modelo de previsão, sendo que este foi modelado com dados sintéticos e reais. Assim, obtiveram-se erros de previsão na validação externa de 1,68% para etanol e de 3,68% para glicose, considerando-se 13 variáveis latentes para ambos compostos.

1. INTRODUÇÃO

Com a Crise do Petróleo na década de 70, o governo brasileiro passou a utilizar o etanol como combustível alternativo aos combustíveis fósseis. Em 1973, a partir da implementação do Programa Nacional do Álcool (PROÁLCOOL), o Brasil tornou-se o pioneiro na produção de bioenergia em larga escala e, desta forma, o álcool etílico tornou-se um produto industrial importante para a economia do país.

Segundo o balanço energético brasileiro de 2011, o país produziu cerca de 22,9 bilhões de litros de etanol e pesquisas preveem uma demanda potencial de 47,5 bilhões de litros para 2015 (Milanez, 2012). Neste cenário, avanços tecnológicos que levem à otimização do processo produtivo do etanol são de suma importância para as indústrias do setor.

A fermentação alcoólica é realizada pela levedura *Saccharomyces cerevisiae*, que utiliza glicose como substrato na geração de etanol. Este processo ocorre a partir do consumo de açúcares simples, que são encontrados no caldo de cana-de-açúcar e podem ser obtidos a partir da hidrólise de açúcares superiores presentes no material celulósico das plantas. A produção de etanol celulósico, ou etanol de segunda geração, é uma alternativa para aumentar a produção deste biocombustível sem a necessidade de expandir áreas de cultivo da cana-de-açúcar.

Mensurar a diferença da concentração do substrato e do produto ao longo do tempo é

crucial para o controle do balanço de massa da reação bioquímica e, conseqüentemente, determinante da economicidade do processo em questão. Os componentes químicos da fermentação podem ser quantificados com precisão a partir da Cromatografia Líquida de Alta Eficiência (do inglês, *High Performance Liquid Chromatography*, HPLC), embora este método seja demorado e haja perda das amostras analisadas.

A Espectroscopia no Infravermelho Próximo (NIR, *Near Infrared Spectroscopy*) associada às técnicas de calibração multivariada representam uma alternativa à quimiometria convencional. A regressão por mínimos quadrados parciais (PLS, *Partial Least Squares*) é uma das principais técnicas de análise multivariada encontradas na literatura. Segundo Morita et al. (2011), o etanol e a glicose podem ser quantificados em meio fermentativo a partir do PLS-NIR com sucesso. A técnica propostamostrou-se eficiente no monitoramento de diversos produtos em misturas de compostos, destacando-se por quantificar multi-constituintes com o mínimo, ou nenhuma, preparação da amostra de forma rápida e não destrutiva (Chongqing e Hong hang, 2008).

No presente trabalho, avaliou-se a utilização de PLS-NIR para a análise multivariada dos principais componentes da fermentação alcoólica (glicose e etanol), comparando-se diferentes formas de montar a matriz de calibração, sendo que esta foi formada tanto por amostras sintéticas quanto por amostras obtidas do meio fermentativo.

2. MATERIAIS E MÉTODOS

2.1 Amostras Reais: Processo de Produção do Etanol Celulósico

Preparação de biomassa: Bagaço e palha de cana pré-tratados foram lavados e os materiais lignocelulósicos foram secos a 45 °C durante 48 horas. Em seguida foram processados em um moinho de faca com granulometria inferior a 2 mm.

A hidrólise enzimática: Bagaço e palha de cana explodidos a vapor foram submetidos à hidrólise enzimática em 0,1 M de tampão de citrato (pH 5) à razão sólido/ líquido de 1/10. Utilizaram-se 30 FPU de extrato enzimático por grama residual de celulose no material pré-tratado. Os ensaios foram feitos em agitação a 50°C e 200 rpm durante 24 h.

Fermentação alcoólica: A fração líquida do hidrolisado foi fermentada inoculando-se o meio com 25 g/L de levedura e dada concentração inicial de glicose padronizada em 100 g/L. Os experimentos foram conduzidos sob agitação a 31°C e 100 rpm.

O monitoramento da glicose foi realizado por kit enzimático (Doles[®], Brasil) e a produção de etanol foi monitorada através do HPLC (Pereira *et al.* 2013).

2.2 Amostras Sintéticas

Amostras puras e misturas: As amostras sintéticas foram preparadas a partir da diluição de uma solução tampão de Citrato de Sódio (5 mM e pH= 4,5) com concentração inicial de etanol e glicose igual a 200 e 180g/L, respectivamente. A partir destas soluções aferiram-se diferentes diluições para cada componente estudado (P). Partindo-se da mesma solução tampão, obtiveram-se misturas ideais (M) de produto e substrato em diversas proporções, as

quais foram escolhidas aleatoriamente.

2.3 Varredura no Infravermelho Próximo

Um espectrômetro FT-NIR da Perkin Elmer (*Spectrum 100N*) – propriedade da Embrapa Instrumentação – foi utilizado na leitura de todas as amostras consideradas neste estudo. A resposta do equipamento consistiu na média de 32 varreduras de absorbância numa região de comprimentos de onda.

2.4 Tratamento e Análise Matemática dos Espectros

Os procedimentos de refinamento dos dados, bem como as análises estatísticas necessárias neste desenvolvimento foram implementados em MATLAB (*Matlab R2010b*).

Pré-tratamento dos espectros: Os espectros originais foram pré-tratados, sendo que para este estudo, considerou-se o alisamento pela média com e sem a redução do número de pontos, bem como as derivadas de primeira e segunda ordem dos espectros alisados.

Calibração: Os dados tratados foram modelados a partir da multicalibração. A fim de verificar qual modelo prevê a concentração dos analitos com mais acurácia, avaliaram-se 5 conjuntos de calibração distintos formados a partir de amostras sintéticas puras, misturas e reais. O modelo de calibração utilizado para este estudo foi o da Regressão por Mínimos Quadrados Parciais (*Partial Least Squares*, PLS) (Haaland, 1998).

Validação cruzada e validação externa: A fim de definir a quantidade de variáveis latentes do modelo, aplicou-se a validação cruzada (*leave-one-out*, LOO) do sistema e calculou-se a soma do quadrado dos erros de previsão (PRESS), ou desvio padrão de previsão, para cada modelo gerado. O cálculo do *i*-ésimo PRESS foi realizado a partir da Equação 1.

$$PRESS^i = \sqrt{\frac{1}{m} \sum_{k=1}^m (c_k^{(i)} - \hat{c}_k^{(i)})^2} \quad (1)$$

Onde $c_k^{(i)}$ e $\hat{c}_k^{(i)}$ são a referência e o valor predito das concentrações usadas na validação cruzada. Assim, o número de fatores mais adequado para descrever o sistema em questão, ou seja, o número de regressores (NR) ótimo foi aquele cujo valor de PRESS do componente foi mínimo.

A validade do modelo obtido foi verificada a partir do erro de validação externa (EVE), na qual se utilizaram 3 amostras do conjunto de dados da fermentação para avaliar a previsão das concentrações de etanol e glicose. Os erros da validação externa foram calculados analogamente ao PRESS, dado que $\hat{c}_k^{(i)}$ admite o valor predito pela PLS (Breton, 2007).

3. RESULTADOS E DISCUSSÃO

Neste trabalho, consideraram-se 22 amostras reais de fermentações distintas. Obtiveram-se concentrações variadas no meio fermentativo para os compostos químicos estudados, dentre as quais as máximas foram de 28,33g/L para o etanol e de 34,72g/L para a glicose, bem como a concentração mínima foi de 17,41 e de 21,31 g/L de etanol e glicose, respectivamente.

Já para as amostras sintéticas, analisaram-se 32 amostras, sendo 10 para o conjunto de soluções de apenas um componente (denominadas aqui de amostras puras) e 22 para as soluções de misturas de etanol e glicose em tampão citrato (denominadas aqui de misturas ideais). A Tabela 1 contém as concentrações puras.

Tabela 1-Concentrações de etanol (C_E) e glicose (C_G) puros em diferentes diluições.

C_E (g/L)	200	160	120	80	40
C_G (g/L)	180	144	108	72	36

A Tabela 2 apresenta as concentrações referentes às misturas ideais de glicose e etanol consideradas.

Tabela 2- Concentrações mássicas de etanol (C_G) e glicose (C_E) nas amostras sintéticas.

Amostras	Concentração mássica (% m/m)	
	C_G	C_E
1	100,0	0,0
2	94,4	5,6
3	88,9	11,1
4	81,8	18,2
5	74,7	25,3
6	66,4	33,6
7	57,0	43,0
8	45,7	54,3
9	32,9	67,1
10	18,1	81,9
11	0,0	100,0

Após obter os espectros no NIR, realizaram-se os pré-tratamentos, que são exemplificados pelos gráficos dispostos nas Figuras 1 e 2. Neste estudo, consideraram-se o alisamento dos espectros, bem como suas derivadas primeira e segunda.

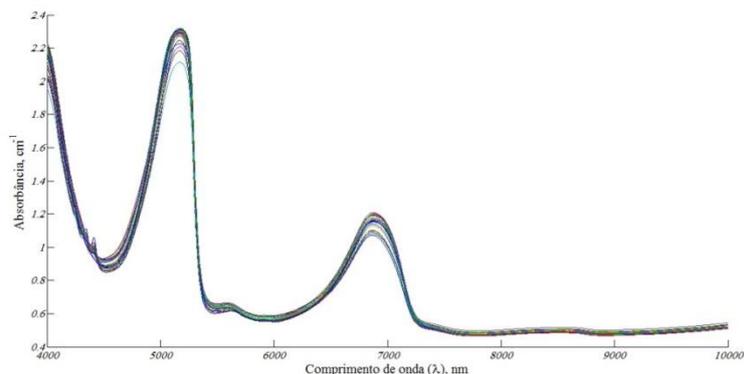
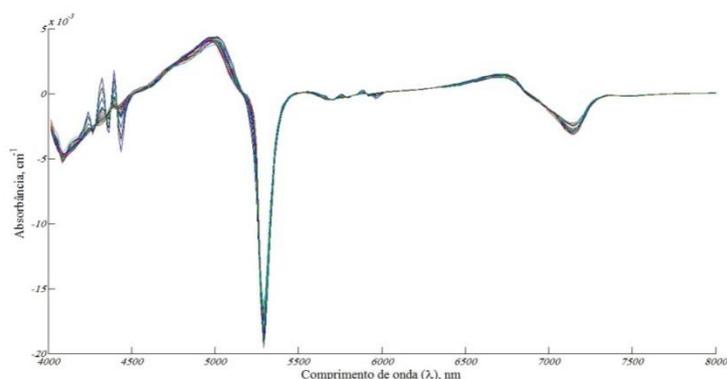
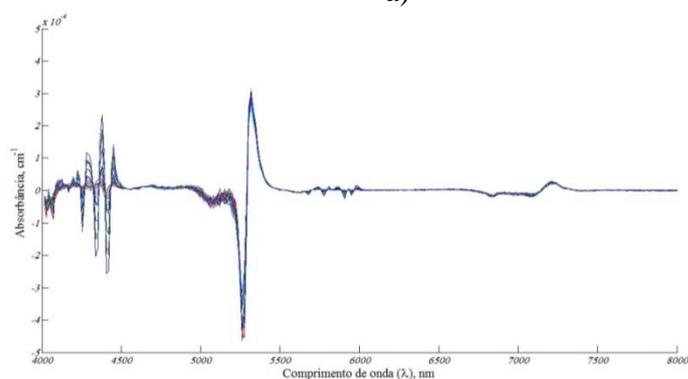


Figura 1- Espectros do conjunto de amostras sintéticas (puras e misturas ideais) alisados pela média móvel.



a)



b)

Figura 2- Espectros do conjunto de amostras sintéticas alisadas com derivada de a) primeira e b) segunda ordem.

A priori, avaliou-se o conjunto de calibração formado apenas por amostras puras e sintéticas. Para este conjunto de amostras, analisaram-se os erros da validação cruzada para diferentes combinações de pré-tratamentos. Nesta análise, 5 variáveis latentes foram consideradas em todas as modelagens. Os dados estão dispostos na Tabela 3.

O procedimento de alisamento não levou a grandes erros da validação para ambos componentes, sendo que estes foram inferiores a 5%. Este resultado apresentou erros equivalentes àqueles obtidos a partir de espectros brutos, sendo assim, consideraram-se apenas os PRESSs dos dados alisados.

A utilização das derivadas primeira e segunda dos espectros (alisados ou não) não melhorou o desempenho do modelo. Como pode ser visto nas Figuras 1 e 2 não há translações acentuadas nas linhas de base dos espectros de absorvância e de sua primeira derivada, as quais justificariamos uso da primeira e segunda derivada, respectivamente, como pré-tratamento. Assim, a utilização de derivadas numéricas dos dados ruidosos levou a um pior desempenho desses modelos, principalmente para os dados de etanol.

Tabela 3-Resumo da análise dos pré-tratamentos estudados.

	<i>Modelos</i>					
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	
<i>Alisamento pela Média(perda de resolução)</i>	X	-	-	-	X	
<i>Alisamento pela Média Móvel</i>	-	X	-	-	-	
<i>Derivada Primeira</i>	-	-	X	-	X	
<i>Derivada Segunda</i>	-	-	-	X	-	
PRESS (%)	<i>Glucose</i>	1,0	1,2	1,1	1,3	1,5
	<i>Etanol</i>	5,0	5,0	10,0	13,0	9,5

A partir destes resultados, considerou-se apenas o alisamento pela média com perda de dados para o pré-tratamento dos demais espectros, visto que esta técnica forneceu o menor erro de validação cruzada para ambos componentes.

Além de amostras da fermentação etanólica real, analisaram-se também amostras sintéticas, visto que o processo fermentativo para a produção de etanol de segunda geração ocorre em um meio complexo no qual coexistem diferentes compostos químicos além daqueles que se deseja mensurar e que podem prejudicar a previsão da concentração dos compostos químicos. Portanto, para que as bandas do infravermelho próximo da glicose e do etanol sejam facilmente identificados na solução real, obtiveram-se, a priori, espectros de meios ideais.

Determinaram-se os regressores dos demais conjuntos de dados utilizados para a calibração, bem como os respectivos erros referentes à validação cruzada (PRESS) e à validação externa. As Tabelas 4 e 5 apresentam as variáveis latentes e os erros obtidos nas etapas mencionadas, bem como as faixas de concentração referentes a cada modelo gerado. Salienta-se que os fatores foram escolhidos cautelosamente a fim de evitar modelos sobreajustados (*over-fitting*).

Como descrito, inicialmente, consideraram-se somente os dados sintéticos no conjunto de calibração, ou seja, de amostras puras e de misturas ideais (PM). Em seguida, consideraram-se apenas os dados da fermentação etanólica real (F). Avaliaram-se também os erros na modelagem baseada no conjunto de calibração formado por espectros de amostras puras, misturas ideais e caldo de fermentação real (PMF). A Tabela 4 contém os resultados referentes a estas modelagens.

Tabela 4- Erros de validação cruzada e externa dos modelos obtidos com espectros alisados para conjuntos de calibração formados por: PM; F e PMF.

	Conjuntos de Calibração					
	PM		F		PMF	
	Etanol	Glicose	Etanol	Glicose	Etanol	Glicose
Faixa de Validade do Modelo (g/L)	0-160	0-180	17,41-28,33	21,31-34,72	0-160	0-180
NR	6	6	2	6	13	13
PRESS (%)	0,98	5,44	9,86	9,98	1,55	4,6
EVE (%)	4,51	8,44	22,96	8,43	1,68	3,68

O conjunto de dados formados pelas amostras reais foi satisfatório somente para a previsão da glicose, visto que o erro foi inferior a 10%. Enquanto que para o etanol, este conjunto apresentou um erro de previsão superior a 20%, assim o ajuste do modelo para o etanol partindo destas amostras não apresentou resultados promissores. A possível explicação para este resultado pode estar atrelada à presença de diversos componentes químicos no meio fermentativo que podem ter absorvido no mesmo comprimento de onda que os componentes estudados alterando, assim, o padrão espectral destas substâncias. A análise do conjunto de calibração, neste caso, poderia conter faixas de concentração mais amplas para os componentes analisados, assim os erros seriam reduzidos, visto que a influência dos compostos interferentes presentes no caldo seria menor. Os demais conjuntos de espectros estudados geraram bons modelos de previsão, apresentando erros aceitáveis que variaram entre 3,5 e 8,5%.

A união dos espectros das amostras sintéticas e reais para formar um conjunto de calibração (PMF) forneceu os melhores resultados. Os erros de previsão dos modelos foram inferiores a 4% tanto para o etanol quanto para a glicose. Além de menores erros de previsão, estes modelos podem ser utilizados para prever uma faixa de concentração elevada e ainda incorporam a presença dos possíveis interferentes do sistema, pois contêm amostras reais no conjunto de calibração. Modelos com erros de validação externa ainda menores foram obtidos, mas, nestes casos, os modelos, provavelmente, apresentaram *over-fitting*, visto que estes modelos de previsão apresentavam grande número de variáveis latentes.

Avaliou-se ainda a influência das amostras puras no modelo de previsão, ou seja, a calibração do modelo foi obtida sem a referência de espectros das substâncias puras. Assim, consideraram-se dois conjuntos de calibração, o primeiro, formado somente por espectros das misturas ideais (M) e o segundo formado por estes dados mais aqueles referentes à fermentação real (MF). A Tabela 5 apresenta os resultados obtidos.

Tabela 5- Erros de validação cruzada e externa dos modelos obtido com espectros alisados para os seguintes conjuntos de calibração: M e MF.

	Conjuntos de Calibração			
	M		MF	
	Etanol	Glicose	Etanol	Glicose
Faixa de Validade do Modelo (g/L)	10,14-90,42	5,37-45,92	10,14-90,42	5,37-45,92
NR	3	3	10	14
PRESS (%)	12,88	13,40	5,25	5,42
EVE (%)	16,43	46,67	2,36	6,11

Ao analisar o conjunto de calibração M, verificaram-se erros de validação externa altos, chegando a 47% para a glicose o que inviabilizaria o monitoramento. A união dos dados de misturas ideais e reais (MF) para formar o conjunto de calibração reduziu os erros de previsão do modelo (~ 6%), mas não teve um desempenho melhor que o conjunto PMF.

4. CONCLUSÃO

Neste trabalho foram avaliados diferentes conjuntos de calibração multivariada, a fim de obter aquele que apresentasse o melhor monitoramento dos principais componentes da fermentação alcoólica de segunda geração (substrato e produto). Verificou-se que o melhor conjunto de calibração foi aquele composto por espectros de amostras sintéticas e reais (PMF), visto que, para este caso, obtiveram-se erros de validação externa de até 4,0% para ambos analitos. Além deste resultado, constatou-se que o conjunto de calibração composto apenas por amostras puras e misturas também pode ser utilizado na previsão do produto e substrato deste processo, contudo com um erro maior na predição da glicose (~8,5 %).

5. REFERÊNCIAS BIBLIOGRÁFICAS

BRERETON, R. G.. Applied Chemometrics for Scientists. Inglaterra, John Wiley & Sons Ltd, 2007.

CHONGQING L., HONG HANG W.. Near-infrared spectroscopy with a fiber-optic probe for state variables determination in solid-state fermentation. *Process Biochem.* 43(5): 511-516, 2008.

HAALAND, D. M., THOMAS, E. V. Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Anal. Chem.*, v. 60, p. 1193-1202, 1988.

MILANEZ, A. Y.; NYKO D.; GARCIA, J. L. F.; REIS, B. L. S. F. S.. O déficit de produção de etanol no Brasil entre 2012 e 2015: determinantes, consequências e sugestões políticas. BNDES Setorial, março de 2012.

MORITA H., HASUNUMA T., VASSILEVA M., TSENKOVA R., KONDO A. Near Infrared Spectroscopy as High-Throughput Technology for Screening of Xylose-Fermenting Recombinant *Saccharomyces cerevisiae* Strains. *Anal Chem.*, v. 83, p. 4023–4029, 2011.

PEREIRA, S. C.; RODRÍGUEZ-ZÚÑIGA; GOMES, B. L.; MARTELLI, F. H.; BARROS, G. O. F.; SILVA, W. T. L.; FARINAS, C.S.. Bioethanol Production from sugarcane bagasse pretreated by steam explosion: chemical and Morphological insights. Publicado em XIX Simpósio Nacional de Bioprocessos, 2013.