



QUANTIFICAÇÃO DE GLICOSE E ETANOL NA FERMENTAÇÃO ALCÓOLICA POR ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO (NIR)

A.S.S. Pinto¹, S.C. Pereira², M.P.A. Ribeiro¹, C.S. Farinas²

- (1) Universidade Federal de São Carlos, DEQ/UFSCar, Rod. Washington Luís, SP 310, Km 235, 13565-905, São Carlos, SP, ariane_sbrice@hotmail.com, marceloribeiro@ufscar.br
(2) Embrapa Instrumentação, R. Quinze de Novembro, 1452, 13560-970 São Carlos, SP, sandracerqueirapereira@gmail.com, cristiane.farinas@embrapa.br

Resumo: O monitoramento do processo de fermentação alcoólica é essencial para o controle de sua eficiência. Contudo, ainda não há métodos rápidos e práticos bem definidos. O presente estudo utilizou a técnica de espectroscopia na região do infravermelho próximo (NIR) associada à análise multivariada para monitorar a produção de etanol celulósico. A biomassa foi proveniente do processamento de quatro variedades de cana-de-açúcar. O conjunto de dados espectrais foi baseado na análise de amostras sintéticas e reais. Os espectros foram modelados utilizando-se a Regressão por Mínimos Quadrados Parciais (PLS). O número de variáveis latentes foi definido com base nos resultados da Validação Cruzada (*leave-one-out*). Diversas combinações empregando dados espectrais de amostras sintéticas e reais foram analisadas a fim de garantir o melhor modelo de previsão. Consideraram-se oito variáveis latentes para ambos compostos. Como resultado, obtiveram-se erros de previsão na validação externa (RMSEP) de $\pm 0,80$ g/L para etanol e de $\pm 1,20$ g/L para glicose. Portanto, a metodologia PLS-NIR foi capaz de quantificar as concentrações analisadas com acurácia, mostrando-se, assim, eficaz na previsão da composição do caldo fermentativo da produção de etanol celulósico.

Palavras-chave: etanol celulósico, espectroscopia no infravermelho próximo, análise multivariada.

MEASUREMENT OF GLUCOSE AND ETHANOL IN ALCOHOLIC FERMENTATION BY NEAR-INFRARED SPECTROSCOPY (NIR)

Abstract: The monitoring of the fermentation process is essential for the control of its efficiency. However, there are no well-defined methods that are fast and practical for such application. The present study used the technique of near-infrared spectroscopy (NIR) associated with multivariate analysis to monitor the production of cellulosic ethanol. The biomass came from the processing of four varieties of sugarcane. The set of spectral data was based on the analysis of real and synthetic samples. The spectra were modeled using the Partial Least Squares Regression (PLS). The number of latent variables was defined based on the results of Cross Validation (*leave-one-out*). The calibration matrices were constructed with different sets of data to ensure the best model. Eight latent variables were considered for both compounds. As a result, there were obtained prediction errors in the external validation (RMSEP) around ± 0.80 g/L for ethanol and ± 1.20 g/L for glucose. Therefore, the methodology PLS-NIR was able to accurately measure the concentrations in the fermentation broth of the of cellulosic ethanol production process.

Keywords: cellulosic ethanol, near-infrared spectroscopy, multivariate analysis.

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 do Programa Nacional do Alcool, 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 2014, a demanda energética de etanol teve um crescimento de 20,0% em 2013 com relação ao ano anterior. Pesquisas preveem uma demanda potencial de 47,5 bilhões de litros deste biocombustível 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 na maioria dos processos industriais pela levedura *Saccharomyces cerevisiae*, que utiliza açúcares simples como substrato na geração de etanol. Este processo ocorre a partir do consumo de açúcares, que são encontrados no caldo de cana-de-açúcar e podem ser alternativamente 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 potencial estratégia para aumentar a produção deste biocombustível sem a

necessidade de expandir áreas de cultivo da cana-de-açúcar e competir com as áreas cultiváveis utilizadas para a produção de alimentos.

Mensurar a concentração do substrato e do produto ao longo do tempo de fermentação 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 (do inglês, *Near Infrared Spectroscopy*, NIR) associada às técnicas de calibração multivariada (PLS-NIR) representam uma alternativa à quimiometria convencional. A regressão por mínimos quadrados parciais (do inglês, *Partial Least Squares*, PLS) é umas das principais técnicas de análise multivariada encontradas na literatura. O PLS-NIR mostrou-se eficiente no monitoramento de diversos produtos em misturas de compostos e destaca-se por quantificar multi-constituintes com o mínimo, ou nenhuma, preparação da amostra, de forma rápida e não destrutiva (HONGQIANG e HONGZHANG, 2008).

No presente estudo, avaliou-se a utilização de PLS-NIR para a análise multivariada dos principais componentes da fermentação alcoólica para obtenção de etanol celulósico, comparando-se diferentes formas de montar a matriz 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

Bagaço e palha de cana-de-açúcar foram secos e processados de acordo com o modelo de produção de etanol celulósico desenvolvido pela Embrapa Agroenergia. O monitoramento da glicose foi realizado por kit enzimático (Doles[®], Brasil) e a produção de etanol foi monitorada através da Cromatografia Líquida de Alta Eficiência (*High Performance Liquid Chromatography*, HPLC) (PEREIRA et al. 2013).

2.2. Amostras sintéticas

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 200 g/L de etanol e de 180 g/L de glicose. Aferiram-se diferentes diluições destas soluções para cada componente estudado puro (P). Partindo-se do mesmo tampão, obtiveram-se misturas ideais (M) de produto e substrato em diversas proporções, as quais foram escolhidas aleatoriamente. O conjunto de amostras composto por ambas soluções, P e M, foi definido como conjunto de amostras sintéticas (PM).

2.3. Varredura no infravermelho próximo, tratamento e análise matemática dos espectros

Um espectrômetro FT-NIR da Perkin Elmer (*Spectrum 100N*) foi utilizado na leitura de todas as amostras deste estudo. Os espectros foram obtidos na temperatura de, aproximadamente, 25 °C, consistindo na média de 32 varreduras numa região de comprimentos de onda variando entre 4.000 a 10.000 cm⁻¹. Os espectros foram obtidos na temperatura de, aproximadamente, 25 °C, consistindo na média de 32 varreduras numa região de comprimentos de onda variando entre 4.000 a 10.000 cm⁻¹. Os espectros originais foram pré-tratados, considerando-se o alisamento pela média com e sem a redução do número de pontos, bem como as derivadas primeira e segunda. Os espectros tratados foram modelados a partir da multicalibração. A fim de verificar qual modelo prevê a concentração dos analitos com maior acurácia, avaliaram-se três conjuntos de calibração distintos formados a partir de amostras puras (P), misturas (M) e reais (F). O modelo de calibração utilizado para este estudo foi o do *Partial Least Squares* (PLS) (HAALAND, 1998).

A fim de definir a quantidade de variáveis latentes (NR), aplicou-se a validação cruzada (*leave-one-out*) do sistema e calculou-se desvio padrão de previsão (*Root Mean Square Error Cross Validation*, RMSECV), para cada modelo gerado. O cálculo do *i*-ésimo RMSECV foi realizado a partir da Equação 1.

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

Em que, $c_k^{(i)}$ e $\hat{c}_k^{(i)}$ são a referência e o valor predito das concentrações usadas na validação cruzada. Assim, o número de variáveis latentes (NR) mais adequado para descrever o sistema em questão, foi aquele cujo valor de RMSECV do analito foi mínimo. A validade do modelo obtido foi verificada a partir do erro de validação externa (*Root Mean Square Error Prediction*, RMSEP). Os erros da validação externa foram calculados analogamente ao RMSECV, dado que $\hat{c}_k^{(i)}$ admite o valor predito pelo PLS (BRERETON, 2007).

3. Resultados e Discussão

Neste trabalho, consideraram-se 37 amostras reais de fermentações. Avaliaram-se conjuntos de calibração contendo 33 amostras reais (F), sendo que os quatro dados remanescentes foram destinados à validação externa. Já para as amostras sintéticas, analisaram-se 32 amostras, sendo 10 para o conjunto de soluções de apenas um componente (P) e 22 para as soluções de misturas (M).

Após obter os espectros no NIR, realizaram-se os pré-tratamentos, exemplificados pelos gráficos das Figuras 1 e 2. A priori, avaliou-se o conjunto de calibração formado apenas por amostras sintéticas – conjunto formado por amostras puras e misturas (PM). Para este conjunto de amostras, analisaram-se os erros da validação cruzada para diferentes combinações de pré-tratamentos. Nesta análise, cinco variáveis latentes foram consideradas em todas as modelagens. Os dados estão dispostos na Tabela 1.

O procedimento de alisamento espectral não levou a grandes erros de RMSECV para ambos componentes. Este resultado apresentou erros equivalentes àqueles obtidos a partir de espectros brutos. Já utilização das derivadas primeira e segunda dos espectros (alisados ou não) não melhorou o desempenho do modelo.

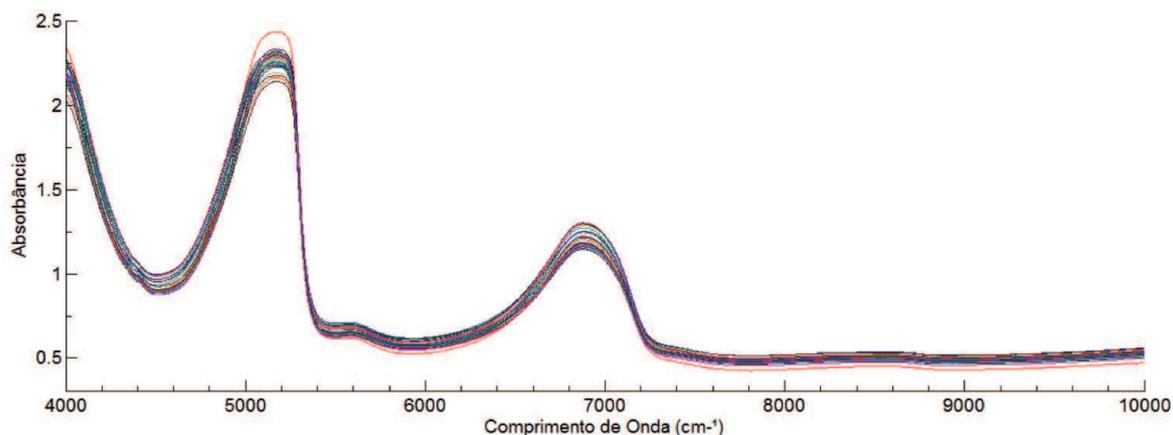


Figura 1. Espectros do conjunto de amostras sintéticas (PM) alisados pela média móvel (técnica utilizada: perda de pontos).

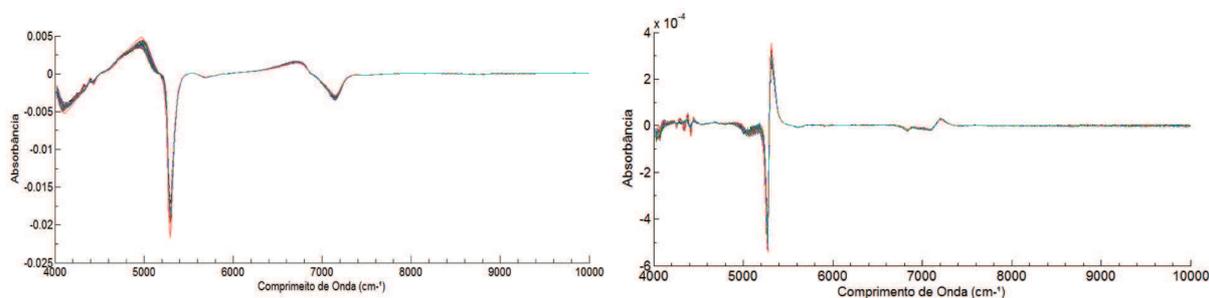


Figura 2. Espectros da primeira e segunda derivada dos Espectros mostrados na Figura 1.

A partir dos resultados, considerou-se apenas o alisamento pela média móvel com perda de dados como o pré-tratamento dos demais espectros.

Tabela 1. Resumo da análise dos pré-tratamentos estudados, em que X sinaliza o pré-tratamento que foi estudado para cada modelagem avaliada (Mi).

Pré-tratamentos		M ₁	M ₂	M ₃	M ₄	M ₅
Alisamento pela Média (perda de resolução)		X	-	-	-	X
Alisamento pela Média Móvel		-	X	-	-	-
Derivada Primeira		-	-	X	-	X
Derivada Segunda		-	-	-	X	-
RMSECV (g/L)	Glicose	5,0	5,0	10,0	13,0	9,5
	Etanol	1,0	1,2	1,1	1,3	1,5

Salienta-se que, após a análise inicial, amostras anômalas foram removidas e as variáveis latentes (NR) foram escolhidas cautelosamente a fim de evitar modelos com RMSEP's elevados ou sobreajustados (*over-fitting*). Avaliaram-se também os erros de previsão na modelagem baseada no conjunto de calibração formado pelos es-

pectros das amostras de fermentação real (F), soluções dos anlitos puros e em mistura (PM), bem como o conjunto formado por amostras puras, misturas e fermentação (PMF). A Tabela 2 contém estes resultados.

Tabela 2. Variáveis latentes (NR), erros de validação cruzada (RMSECV) e externa (RMSEP) dos modelos obtidos com espectros alisados para conjuntos de calibração formados por: PM, F e PMF.

Matriz Calibração	PM		F		PMF	
	Etanol	Glicose	Etanol	Glicose	Etanol	Glicose
Faixa de Validade do Modelo (g/L)	0-160,00	0-180,00	2,32-28,33	21,31-92,85	0-160,00	0-180,00
NR	6	6	6	5	8	8
RMSECV (g/L)	0,98	5,44	4,30	2,10	0,30	0,90
RMSEP (g/L)	4,51	8,44	1,80	1,80	0,80	1,20

Constatou-se que o conjunto de calibração composto apenas por amostras puras e misturas (PM) pode ser utilizado na previsão do produto e substrato deste no processo seja admissível erros de validação externa de, aproximadamente, $\pm 5,0$ g/L para o etanol e de $\pm 8,0$ g/L para a glicose.

A união dos espectros das amostras sintéticas e reais para formar um conjunto de calibração (PMF) forneceu os melhores resultados. Destaca-se que essa matriz calibração forneceu um modelo PLS-NIR cujos RMSEP's foram de, aproximadamente, $\pm 1,0$ g/L para ambos compostos.

Morita et al. (2011) realizou um estudo no qual utilizaram-se na matriz calibração 266 amostras provenientes da fermentação alcoólica realizada com leveduras geneticamente modificadas. Os erros de previsão da modelagem, deste estudo foram estatisticamente equivalentes aos obtidos no presente trabalho para os dados PMF, ou seja, os RMSEP's foram de, aproximadamente, $\pm 1,0$ g/L. Embora os erros sejam significativamente iguais, a matriz calibração proposta no atual estudo contém apenas 65 amostras, dentre as quais 32 amostras correspondem à sintéticas e os demais referem-se à fermentação real (PMF).

4. Conclusões

Neste trabalho, avaliaram-se diferentes matrizes calibração a fim de obter aquela cujo modelo PLS-NIR apresentasse o menor erro de previsão (RMSEP) dos componentes da fermentação alcoólica de segunda geração (substrato - glicose e produto - etanol). O conjunto de calibração PMF forneceu o modelo PLS-NR mais eficaz, embora as matrizes calibração PM e F também forneçam modelos de previsão satisfatórios. Portanto, foi possível mensurar etanol e glicose no caldo fermentativo advindo da produção de etanol de segunda geração com acurácia de, aproximadamente, $\pm 0,80$ e $\pm 1,20$ g/L (RMSEP_{PMF}), respectivamente.

Agradecimentos

À Universidade Federal de São Carlos, à Embrapa Instrumentação e ao financiamento do CNPq (PIBIQ/CNPq -Processo nº 122979/2013-3).

Referências

- BRERETON, R. G.. Applied Chemometrics for Scientists. Inglaterra, John Wiley & Sons Ltd, 2007.
- 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.
- HONGQIANG, L., HONGZHANG, C. Near-infrared spectroscopy with a fiber-optic probe for state variables determination in solid-state fermentation. *Process Biochem.*, v. 43, p. 511-516, 2008.
- 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.