

## ANÁLISE DE COMPONENTES PRINCIPAIS POR INTERVALOS (IPCA) COMO MÉTODO DE SELEÇÃO DE REGIÃO ESPECTRAL NO INFRAVERMELHO MÉDIO E PRÓXIMO PARA DISCRIMINAÇÃO DE ÓLEOS VEGETAIS

Gilson Augusto Helfer<sup>1</sup>, Lilian de Fátima Ferreira da Silva<sup>2</sup>, Fernanda Carla Bock<sup>2</sup>, Luciano Marder<sup>1,2\*</sup>

<sup>1</sup> Programa de Pós-graduação em Sistemas e Processos Industriais, Universidade de Santa Cruz do Sul (UNISC), 96815-900, Santa Cruz do Sul – Brasil

<sup>2</sup> Departamento de Química e Física, Universidade de Santa Cruz do Sul (UNISC), 96815-900, Santa Cruz do Sul – Brasil

\*E-mail: [lucianomarder@unisc.br](mailto:lucianomarder@unisc.br)

Recebido em 05/07/2013  
Aceito em 09/11/2013

### RESUMO

Neste trabalho foi empregada a análise de componentes principais por intervalo (iPCA) como método de seleção de variáveis, associada à espectroscopia no infravermelho próximo e médio, visando à discriminação de óleos vegetais. Os seguintes óleos vegetais foram analisados: soja, canola, girassol, arroz, milho, algodão e amendoim. Os espectros foram adquiridos em quintuplicata, utilizando um espectrômetro Perkin Elmer Spectrum 400 no intervalo entre 650-4.000  $\text{cm}^{-1}$  (infravermelho médio) e 4.000-10.000  $\text{cm}^{-1}$  (infravermelho próximo), com resolução de 4  $\text{cm}^{-1}$  e 32 varreduras. Após, os dados espectrais foram normalizados e posteriormente corrigidos com as transformadas SNV (variação normal padrão) e primeira derivada com uma janela de 5 pontos. Os modelos foram desenvolvidos através do aplicativo Matlab empregando análise de componentes principais por intervalos (iPCA) disponível no pacote iToolbox. Foram realizadas análises subdividindo os espectros em 8, 16 e 32 intervalos, com os dados centrados na média. O presente trabalho indica que a análise de componentes principais por intervalos (iPCA), como ferramenta de seleção de variáveis, possibilita encontrar regiões representativas.

**Palavras-chave:** Análise multivariada, óleos vegetais, quimiometria, iPCA

### 1 Introdução

Autenticidade é um critério de qualidade muito importante para óleos e gorduras comestíveis, devido a uma grande diferença de preços entre os diferentes tipos destes produtos. No entanto, a determinação dessa autenticidade é tradicionalmente um processo demorado e trabalhoso, geralmente usando métodos cromatográficos. Combinada com métodos quimiométricos, a espectroscopia vibracional na região do infravermelho é uma técnica analítica que vem sendo utilizada para verificar a autenticidade dos óleos e gorduras comestíveis, devido a sua simplicidade, rapidez e facilidade de preparação da amostra [1].

Ao se utilizar esta técnica, associada à quimiometria, depara-se com um grande número de variáveis, incluindo aquelas de potencial interferente e correlacionadas a não linearidades, que implicam desde problemas no desempenho computacional até nos modelos gerados. Para resolver este impasse são utilizados métodos de seleção de variáveis. Estes métodos são baseados no princípio da escolha de um pequeno número de variáveis, segregadas a partir do espectro original, onde a remoção de variáveis não informativas vai produzir modelos mais robustos [2]. Os métodos de seleção incluem

abordagens clássicas, como a manual (baseada no conhecimento), de seleção “sequencial” e “univariada”; e métodos sofisticados, como o algoritmo de sucessivas projeções - SPA (*Successive Projection Algorithm*) e técnicas heurísticas, como redes neurais artificiais -ANN (*Artificial Neural Networks*), algoritmos genéticos - GA (*Genetic Algorithm*), entre outras [3].

A complexidade de um modelo pode ser reduzida através de uma pré-seleção criteriosa de comprimentos de onda. O simples fato de se dividir o espectro em regiões proporcionais em tamanho e de se analisar cada um separadamente, já caracteriza um método de seleção [3]. Neste sentido, o presente trabalho sugere o emprego da análise de componentes principais por intervalos (iPCA) como um método de seleção de variáveis na discriminação de óleos vegetais.

#### 1.1 Óleos vegetais

Os óleos e gorduras, também chamados de lipídios, são substâncias insolúveis em água (hidrofóbicas), de origem animal, vegetal ou até mesmo microbiana. São formadas a partir da condensação entre glicerol e ácidos graxos, chamados triglicerídeos, e ácidos graxos livres, que chegam a representar até 96% do peso total dessas moléculas. São eles os principais

combustíveis da maioria dos organismos e constituem, na verdade, uma das mais importantes formas de armazenamento de energia química [4]. As unidades fundamentais da maioria dos lipídios são os ácidos graxos, que podem ser saturados (ligações simples) e insaturados (duplas ligações). Os óleos vegetais têm como principal componente em proporção cadeias insaturadas [5]. Zambiasi [6], via cromatografia gasosa (GC-FID), identificou a composição dos ácidos graxos em diversos óleos vegetais comercializados no Brasil, sendo os resultados apresentados na Tabela 1.

Tabela 1. Composição das amostras de óleos vegetais utilizadas no estudo.

Tipo de óleo	Saturado (%)	Monoinsaturado (%)	Poli-insaturado (%)
Canola	6,98	64,42	28,60
Girassol	12,36	15,93	71,71
Milho	13,87	24,76	61,37
Soja	15,10	21,73	63,17
Amendoim	18,38	50,33	31,29
Arroz	20,68	41,41	37,91
Algodão	25,73	17,49	56,78

### 1.2 Espectroscopia no infravermelho

É uma técnica que permite identificar uma amostra através da radiação infravermelha. Essa radiação ao ser absorvida causa alteração nos modos rotacionais e vibracionais das moléculas. A diferença entre a radiação emitida pela fonte e a radiação absorvida pela amostra é registrada por um detector, gerando um espectro de absorção no infravermelho [7].

Assim como ocorre em outros tipos de absorção de energia, que caracterizam um processo quantizado, as moléculas são excitadas para atingir um estado maior de energia quando absorvem radiação no infravermelho. Uma molécula absorve apenas frequências (energias) selecionadas de radiação, ocorrendo sua vibração por deformações axiais e/ou angulares [8].

A região do infravermelho é absorvida pela maioria dos compostos orgânicos e inorgânicos que possuem ligações covalentes. A região do infravermelho próximo (4.000 a 14.290  $\text{cm}^{-1}$ ) tem como características picos largos e de baixa intensidade, enquanto no infravermelho médio (200 a 4.000  $\text{cm}^{-1}$ ) aparecem picos muito intensos e geralmente estreitos [9,10].

### 1.3 Espectroscopia por reflexão

As técnicas de reflexão vêm cada vez mais sendo empregadas na caracterização de alimentos, bem como no monitoramento dos processos produtivos dos mesmos. Os três métodos de reflexão no infravermelho que apresentam aplicações práticas na aquisição de espectros são: método de reflexão especular (ou externa), método por refletância difusa e o método por reflexão interna. A técnica de reflexão difusa vem sendo amplamente aplicada, associada aos equipamentos que operam no infravermelho próximo, com os quais é geralmente denominada de espectroscopia de reflexão no infravermelho próximo - NIR (*Near Infrared Reflectance Spectroscopy*). A reflexão total atenuada - ATR (*Attenuated Total Reflectance*) é uma técnica de amostragem rápida que requer uma mínima preparação da amostra e, principalmente, que apresenta fácil obtenção de espectros [11].

A espectroscopia no infravermelho próximo com esfera de integração - NIRA (*Near Infrared Reflectance Spectroscopy Accessory*) apresenta refletância difusa e especular, detecção uniforme mesmo em amostras heterogêneas e redução dos efeitos de polarização oriundos do feixe de iluminação e da amostra. É uma excelente ferramenta para análises qualitativas e quantitativas, possui detector Índio-Gálio-Arsênio (InGaAs) proporcionando uma alta relação sinal/ruído. O acessório de refletância atenuada universal - UATR (*Universal Attenuated Total Reflectance*), utilizado no infravermelho médio, oferece uma análise de fácil execução e limpeza para a maioria das amostras, como os óleos em geral, com um mínimo de preparação e alta reprodutibilidade [12].

### 1.4 Quimiometria

Quimiometria é a aplicação de técnicas matemáticas e estatísticas à análise de dados relativos aos processos químicos. Sua importância é caracterizada pela complexidade dos dados que podem ser obtidos via análise de rotina. A necessidade de se analisar um conjunto extenso de variáveis demonstra um aspecto multivariado [13].

Os dados multivariados, geralmente correspondem a uma matriz X de valores, correspondendo a m variáveis para n amostras, conforme ilustra a Figura 1.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & \cdots & x_{nm} \end{bmatrix}$$

Figura 1. Representação de uma matriz de dados. Fonte: [14].

### 1.5 Análise por Componentes Principais – PCA

A base fundamental da maioria dos métodos modernos para tratamento de dados multivariados é o PCA (*Principal Component Analysis*), que consiste numa manipulação da matriz de dados com objetivo de representar as variações presentes em muitas variáveis, através de um número menor de "fatores". Constrói-se um novo sistema de eixos (denominados rotineiramente de fatores, componentes principais, variáveis latentes ou ainda autovetores) para representar as amostras, no qual a natureza multivariada dos dados pode ser visualizada em poucas dimensões [15].

A análise de fatores é realizada sobre uma matriz de dados que relaciona um conjunto de variáveis a diversos experimentos (amostras). Esta matriz de dados  $A$  pode ser centrada na média ou escalonada, sendo decomposta no produto de três matrizes através do algoritmo de decomposição de valor singular - SVD (*Singular Value Decomposition*), resultando na equação  $X=U*S*V^T$ .

A matriz  $V$  (transposta) é denominada de matriz dos *loadings*, as colunas da matriz  $V$  correspondem aos autovetores e  $S$  é uma matriz diagonal. As matrizes  $U$  e  $V$  são ortogonais entre si e o produto entre matrizes  $U$  e  $S$  é denominado de *escores*. Como resultado da análise de componentes principais, o conjunto de dados originais é agrupado em função da correlação existente entre as variáveis, gerando um novo conjunto de eixos (componentes principais) ortogonais entre si e de mais simples manipulação matemática [16].

### 1.6 Análise por Componentes Principais em intervalos – iPCA

O algoritmo da iPCA divide o conjunto de dados (variáveis) em um número de intervalos equidistantes. Para cada intervalo realiza-se uma PCA, que descreve a variação (ou dispersão) de um determinado conjunto de dados em intervalos pré-estabelecidos, apresentando os resultados em gráfico de *escores*. Este método objetiva dar uma visão geral dos dados e pode ser útil em estudos exploratórios e interpretativos em relação aos agrupamentos entre as amostras [17].

## 2 Metodologia

### 2.1 Amostragem

Para o desenvolvimento desse trabalho foram utilizados sete diferentes óleos vegetais apresentados na Tabela 2.

Tabela 2. Amostras de óleos vegetais utilizadas no estudo

Óleo	Código	Marca	Lote	Fabricação	Validade
Soja	SOYB	Bunge	L0912	15/09/12	06/2013
Soja	SOYL	Leve	C5112	16/06/12	06/2013
Canola	CANB	Bunge	L0312	28/03/12	12/2012
Canola	CANL	Liza	L05 C	11/05/12	05/2013
Girassol	SFWB	Bunge	L0412	29/04/12	01/2013
Girassol	SFWL	Liza	L06 C	08/06/12	06/2013
Arroz	RISE	Carreteiro	L180	06/12	06/2013
Milho	CRNB	Bunge	L0312	26/03/12	12/2012
Milho	CRNL	Liza	L08 C	27/08/12	08/2013
Algodão	COTT	Triângulo	PD-071	12/06/12	-
Amendoim	PEAT	Triângulo	PD-069	12/06/12	-

### 2.2 Aquisição dos espectros no infravermelho

Os espectros foram obtidos em um espectrômetro Perkin-Elmer 400 Series empregando-se um acessório contendo esfera de integração, equipado com um detector Índio-Gálio-Arsênio (InGaAs), para as análises no infravermelho próximo (FT-NIR), enquanto que, no infravermelho médio (FT-MIR), empregou-se um acessório de refletância total atenuada universal (UATR). Foram utilizados os intervalos entre 650 e 4.000  $\text{cm}^{-1}$  (FT-MIR) e entre 4.000 e 10.000  $\text{cm}^{-1}$  (FT-NIR), com resolução de 4  $\text{cm}^{-1}$  e 32 varreduras. Para cada amostra de óleo foram adquiridos cinco espectros, e na sequência, os dados obtidos foram normalizados. A aquisição dos espectros foi realizada no período entre outubro e novembro de 2012. O conjunto de espectros médios das amostras encontra-se na Figura 2 e na Figura 3.

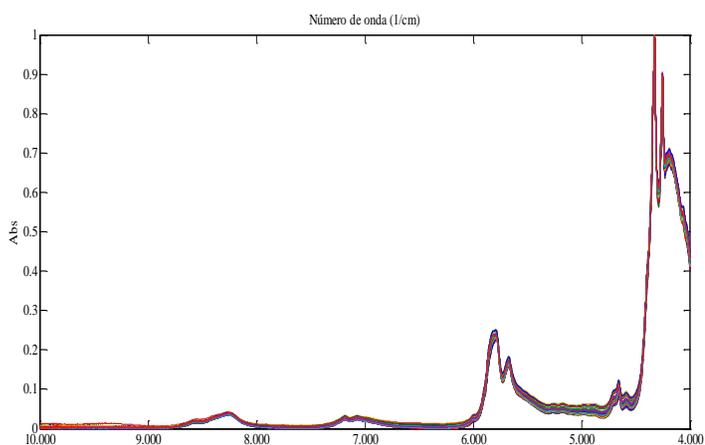


Figura 2. Espectro normalizado das amostras de óleo na faixa de 10.000 a 4.000  $\text{cm}^{-1}$  (FT-NIR).

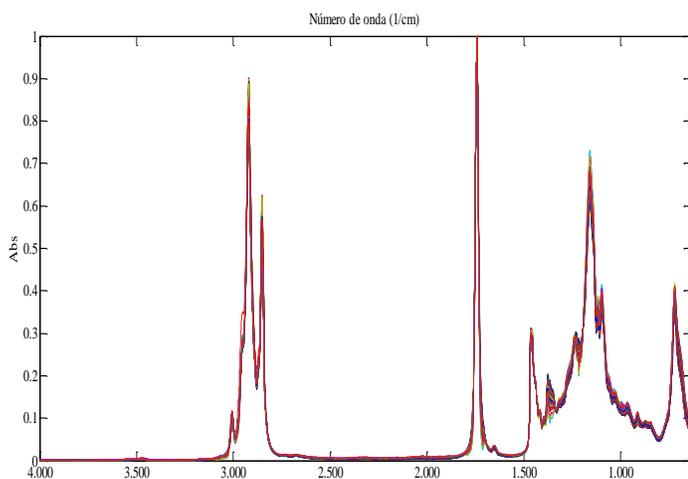


Figura 3. Espectro normalizado das amostras de óleo na faixa de 4.000 a 650  $\text{cm}^{-1}$  (FT-MIR).

### 2.3 Análise dos dados

A matriz de espectros foi analisada por *iPCA*, através do *software* Matlab 7.0, empregando-se o pacote *iToolbox* desenvolvido por Norgaard [18] com o objetivo de identificar a região espectral que caracteriza a separação dos diferentes tipos de óleo vegetal. Para esta análise, os espectros foram divididos em 8, 16 e 32 regiões equidistantes. Além disso, foram aplicados métodos de correção de sinais, como a SNV (variação normal padrão) e a primeira derivada (janela de 5 pontos). Finalmente, visando melhorar o desempenho quimiométrico, os dados foram centrados na média. Nas Figuras 4 e 5 podem ser observados os espectros das amostras de óleos com a aplicação da correção por SNV e primeira derivada, para os dados FT-NIR e FT-MIR, respectivamente.

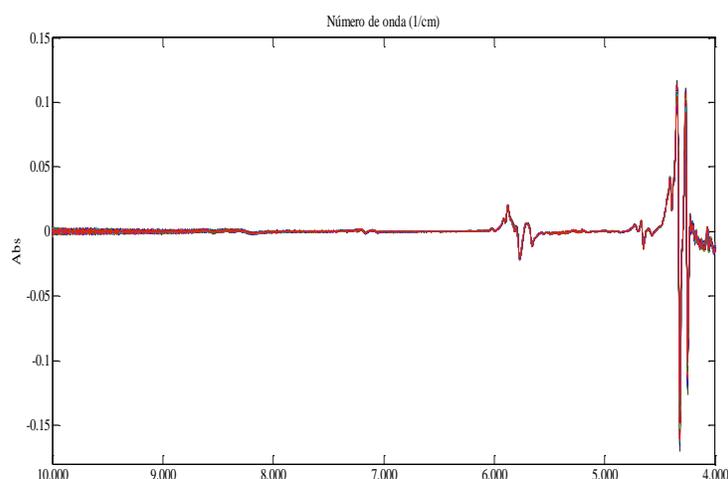


Figura 4. Espectros das amostras de óleos com aplicação da correção por SNV e primeira derivada (FT-NIR).

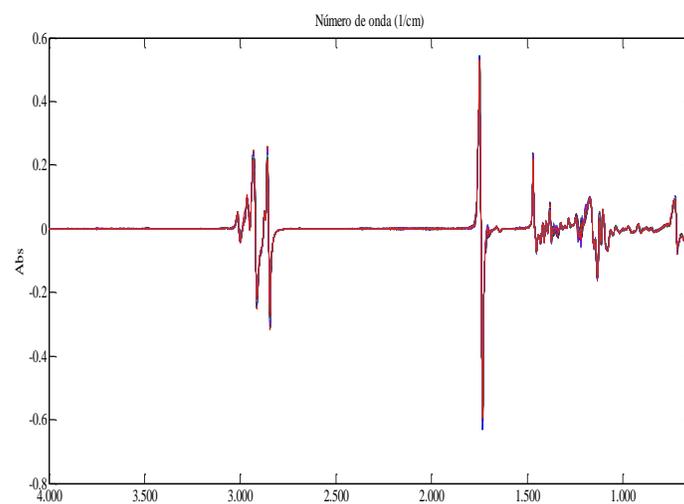


Figura 5. Espectros das amostras de óleos com aplicação da correção por SNV e primeira derivada (FT-MIR).

## 3 Resultados e discussões

### 3.1 Infravermelho Próximo (FT-NIR)

Primeiramente foi realizada uma PCA com toda a faixa espectral, ou seja, de 4.000 a 10.000  $\text{cm}^{-1}$ . Como pode ser observado na Figura 6, utilizando toda essa faixa, não se obteve uma boa separação dos óleos vegetais.

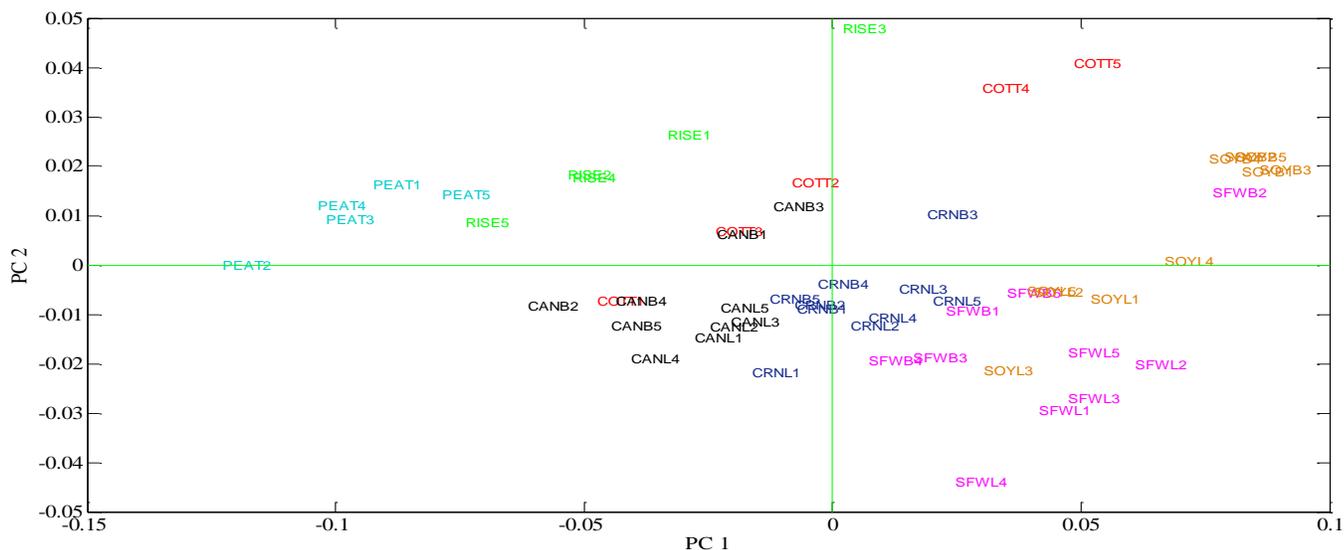


Figura 6. Gráfico de escores PC1 x PC2 - modelo utilizando toda faixa espectral (FT-NIR).

A seguir, com os dados FT-NIR, foram avaliadas com tPCA as subdivisões em 8, 16 e 32 intervalos. Os melhores resultados foram obtidos dividindo o espectro em 32 intervalos. A Figura 7 apresenta o gráfico da variância percentual para cada região estudada. Neste gráfico, para cada intervalo, ou

seja, para cada região do espectro, foi calculada a variância, em percentual, em cada componente principal. O intervalo 23 acumula 94,70% das informações nas duas primeiras componentes principais.

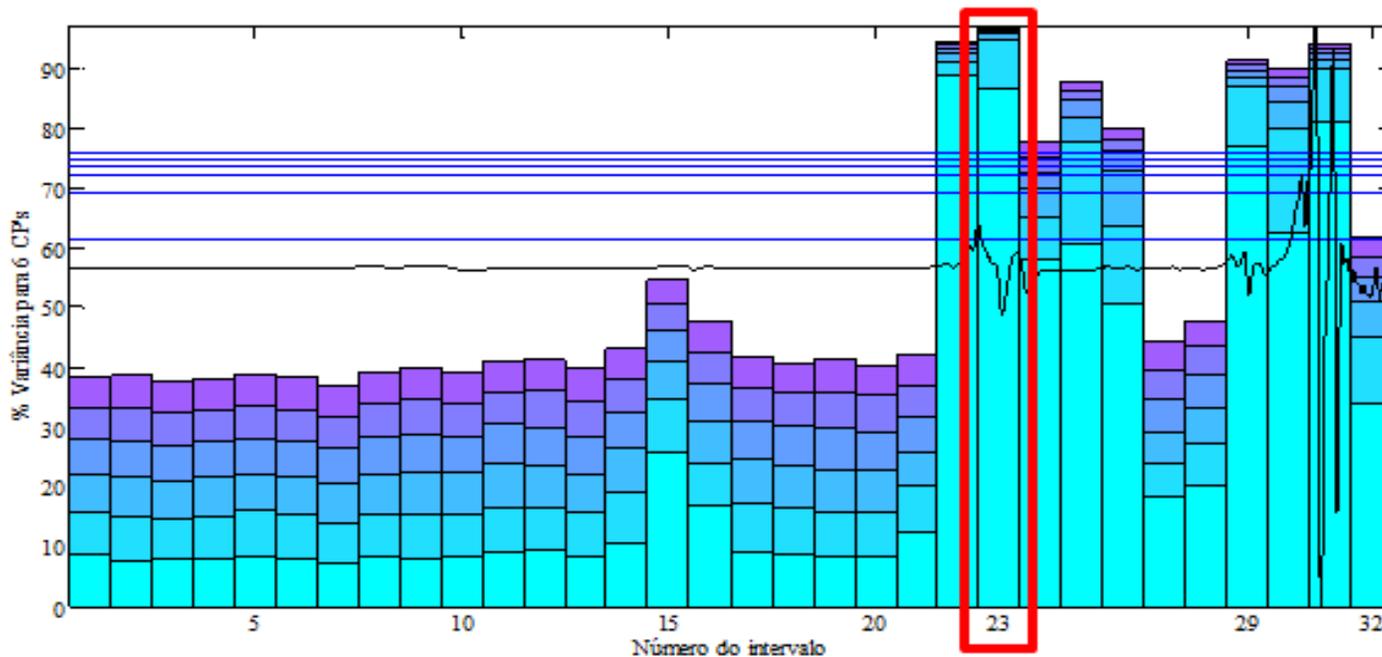


Figura 7. Variação percentual das componentes principais divididas em 32 intervalos (FT-NIR).

As alturas das barras representam, em forma percentual, a variância contida em cada componente principal para cada intervalo. Cabe ressaltar que a linha horizontalmente demarcada representa a variância de cada uma das

componentes principais para a análise de PCA para toda a informação do espectro. Os valores das variâncias referentes ao intervalo 23 são apresentados na Tabela 3.

Tabela 3. Valores para variância e % acumulado da variância (PCA).

Componente principal	% Variância	% Acumulado variância
1	86,75	86,75
2	7,94	94,70
3	1,27	95,97
4	0,44	96,42
5	0,33	96,75
6	0,30	97,05

A região espectral onde se obteve uma melhor discriminação dos óleos vegetais correspondeu à faixa entre 5.683 e 5.869  $\text{cm}^{-1}$  (intervalo 23). Segundo Yang [1] e Xiaobo [3], a região entre 5.500 e 6.000  $\text{cm}^{-1}$  representa a vibração

axial (estiramento) da ligação C–H dos grupos funcionais –  $\text{CH}_2$ , – $\text{CH}_3$  e –  $\text{CH}=\text{CH}$  nos óleos e gorduras comestíveis. A Figura 8 apresenta o gráfico dos escores para PC1 x PC2, obtidos para o intervalo 23.

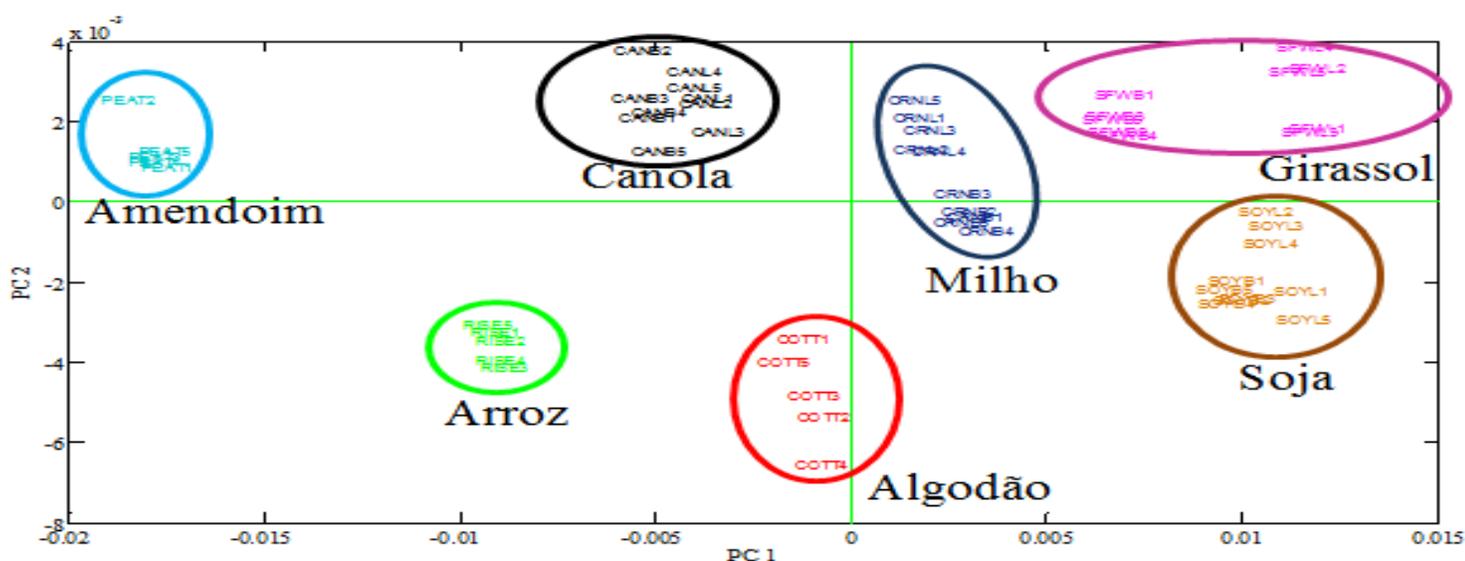


Figura 8. Gráfico de escores PC1 x PC2 – intervalo 23 (FT-NIR).

Analisando o gráfico dos escores da PC1 x PC2 (Figura 8), é possível observar que a PC1 separa os óleos de milho, soja e girassol, em valores positivos e os óleos de algodão, amendoim, arroz e canola, em valores negativos. As amostras de milho, soja e girassol se agrupam por possuírem maiores quantidades de ácidos graxos poli-insaturados em sua composição (Tabela 1). Em relação às demais, a PC2 separa os óleos com maior concentração de ácidos graxos monoinsaturados (canola e amendoim) em valores positivos e

dos óleos com maior concentração de ácidos graxos saturados (algodão e arroz) em valores negativos.

### 3.2 Infravermelho Médio (FT-MIR)

Assim como no infravermelho próximo, foi realizada uma PCA para os dados no infravermelho médio com toda a faixa espectral, ou seja, de 650 a 4.000  $\text{cm}^{-1}$ . Como pode ser verificado na Figura 9, utilizando toda essa faixa, não se obteve uma satisfatória separação dos óleos vegetais.

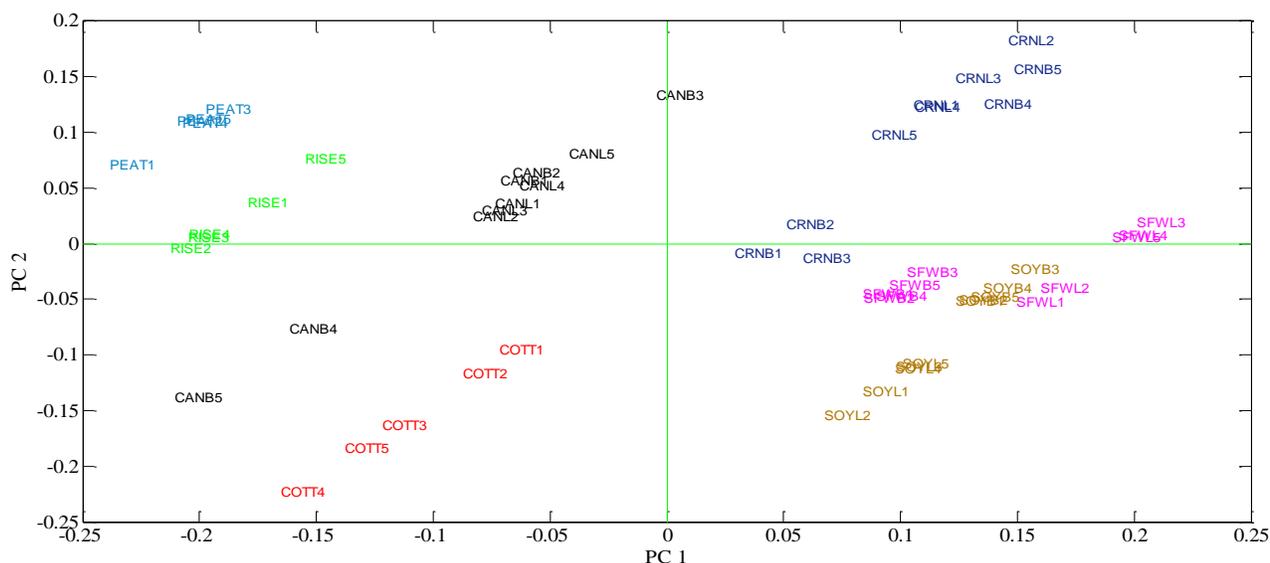


Figura 9. Gráfico de escores PC1 x PC2 - modelo utilizando toda faixa espectral (FT-MIR).

Posteriormente, com os dados FT-MIR, foram avaliadas com *i*PCA as subdivisões em 8, 16 e 32 intervalos. O melhor desempenho dos resultados foi alcançado dividindo o espectro em 32 intervalos. A Figura 10 apresenta o gráfico da variância percentual para cada região proposta. Neste gráfico, para cada

intervalo, ou seja, para cada região do espectro, foi calculada a variância, em percentual, em cada componente principal. O intervalo 28 acumula 95,49% das informações nas duas primeiras componentes principais.

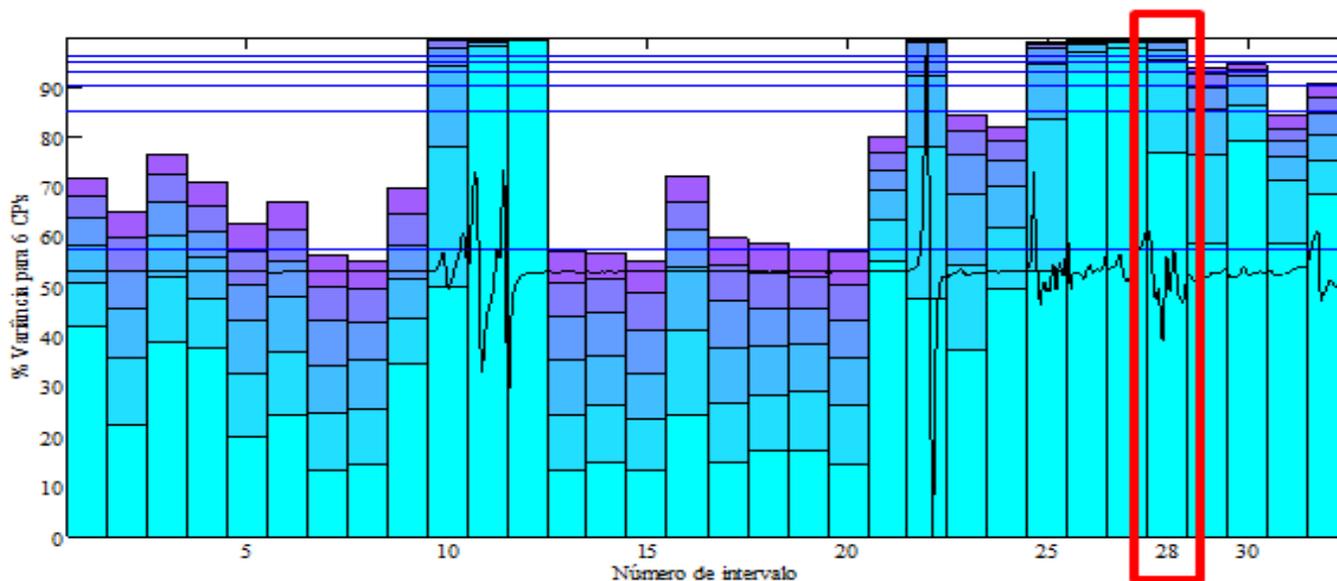


Figura 10. Variação percentual das componentes principais divididas em 32 intervalos (FT-MIR).

As alturas das barras representam, em forma percentual, a variância contida em cada componente principal para cada intervalo. A linha traçada horizontalmente representa a variância de cada uma das componentes principais para a

análise de PCA para toda a informação do espectro. Os valores das variâncias referentes ao intervalo 28 são apresentados na Tabela 4.

Tabela 4. Valores para variância e % acumulado da variância (PCA).

Componente principal	% Variância	% Acumulado variância
1	76,68	76,68
2	18,81	95,49
3	2,14	97,63
4	1,30	98,93
5	0,60	99,54
6	0,12	99,66

A região espectral onde se obteve uma melhor discriminação dos óleos vegetais correspondeu à faixa entre 1.169 e 1.066  $\text{cm}^{-1}$  (intervalo 28). Segundo Yang [1] e Guillén [19], a região entre 1.100 e 1.200  $\text{cm}^{-1}$  representa a vibração axial (estiramento) da ligação C-O dos ésteres (C-C(=O)-O) e a vibração angular (flexão) da ligação C-H dos grupos funcionais -CH<sub>2</sub> nos óleos e gorduras comestíveis. A Figura 11 apresenta o gráfico dos escores para PC1 x PC2, obtido para o intervalo 28.

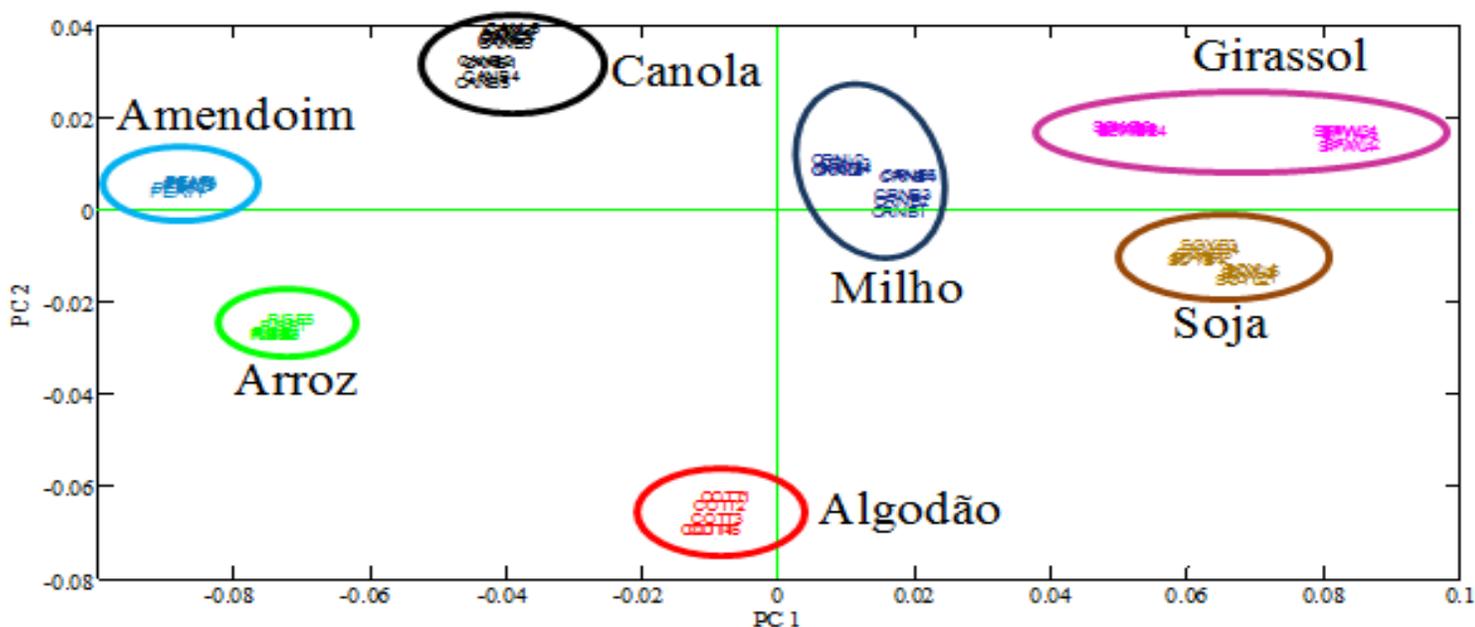


Figura 11. Gráfico de escores PC1 x PC2 – intervalo 23 (FT-MIR).

Da mesma forma que o observado no infravermelho próximo (FT-NIR) é possível verificar no gráfico dos escores da PC1 x PC2 (Figura 11) que a PC1 separa os óleos de milho, soja e girassol em valores positivos e os óleos de algodão, amendoim, arroz e canola, em valores negativos. Assim, a PC2 separa os óleos de canola e amendoim, em valores positivos e os óleos de algodão e arroz, em valores negativos. A separação dos óleos vegetais está associada ao fato de que cada óleo/gordura difere na composição, comprimento e grau de insaturação de ácidos graxos nas cadeias dos triglicerídeos [4].

#### 4 Conclusões

O presente trabalho indica que a análise de componentes principais por intervalos (*i*PCA), como ferramenta de seleção de variáveis, possibilita encontrar regiões representativas visando à discriminação de diferentes óleos vegetais. À medida

que os espectros foram sendo divididos em diferentes intervalos, foram se obtendo regiões espectrais mais seletivas e que possibilitaram uma melhor distinção dos óleos de acordo com sua composição (cadeias saturadas e insaturadas). A região espectral onde se obteve uma melhor discriminação dos óleos vegetais correspondeu à faixa entre 5.683 e 5.869  $\text{cm}^{-1}$  na região do infravermelho próximo e entre 1.169 e 1.066  $\text{cm}^{-1}$  na região do infravermelho médio.

#### Agradecimentos

Os autores agradecem à CAPES, pela bolsa de mestrado, à FAPERGS e à UNISC, pelas bolsas de iniciação científica, e à Triângulo Alimentos Ltda. pela doação dos óleos de algodão e amendoim utilizados no estudo.

**INTERVAL PRINCIPAL COMPONENTS ANALYSIS (IPCA) AS A METHOD OF SPECTRAL RANGE SELECTION IN MID INFRARED AND NEAR INFRARED TO DISCRIMINATION OF VEGETABLE OILS**

**ABSTRACT:** Interval Principal Component Analysis (*i*PCA) was used in this study as variable selection methods associated with mid and near-infrared spectroscopy, aiming discrimination of vegetable oils. The following oils were analyzed: soybean, canola, sunflower, rice, corn, cotton and peanuts. The spectra were acquired in quintuplicate using a Perkin Elmer Spectrum 400 spectrometer in the range of 650-4000  $\text{cm}^{-1}$  (mid-infrared), and 4000-10000  $\text{cm}^{-1}$  (near-infrared) with a resolution of 4  $\text{cm}^{-1}$  and 32 scans. So the spectral data were normalized and later corrected with the transformed SNV (standard normal variation) and first derivative with a 5 point window. The models were developed using the Matlab application employing Interval Principal Component Analysis (*i*PCA) available in the *i*Toolbox package. Analyses subdividing the spectra in 8, 16 and 32 intervals were made, with mean centering data. This study indicates that Interval Principal Component Analysis (*i*PCA) as a tool for variable selection allows to find representative regions aiming discrimination of different vegetable oils.

**Keywords:** Multivariate analysis, vegetable oils, chemometrics, *i*PCA

**Referências**

[1] YANG, H.; IRUDAYARAJ, J.; PARADKAR, M. M. . Food Chemistry, n. 93, p. 25-32, 2005.

[2] COSTA FILHO, P. A.; POPPI, R. J. Aplicação de algoritmos genéticos na seleção de variáveis em espectroscopia no infravermelho médio. Determinação simultânea de glicose, maltose e frutose. Química Nova, Vol. 25, n. 1, 2002.

[3] XIAOBO, Z. et al. Analytica Chimica Acta, n. 667, p. 14-32, 2010.

[4] MORETTO, E.; FETT, R. Tecnologia de óleos e gorduras vegetais na indústria de alimentos. São Paulo: Varela, 1998.

[5] VISENTAINER, J. V.; FRANCO, M. R. B. Ácidos graxos em óleos e gorduras: identificação e quantificação. São Paulo: Varela, 2006

[6] ZAMBIAZI, R. C. et al Boletim do Centro de Pesquisa de Processamento de Alimentos, Curitiba, Vol. 25, n. 1, p. 111-120, jan.- jun. 2007.

[7] BARBOSA, L. C. Espectroscopia no Infravermelho na Caracterização de Compostos Orgânicos. Ed. da UFV: Viçosa. 189 p., 2007.

[8] PAVIA, Donald L. et al. São Paulo: Cengage Learning, 2010.

[9] PASQUINI, C. Near Infrared Spectroscopy: fundamentals, practical aspects and analytical applications. J. Braz. Chem. Soc., vol.14. 198-219 p. 2003.

[10] SILVERSTEIN, Robert M.; WEBSTER, Francis X.; KIEMLE, David J. Identificação espectrométrica de compostos orgânicos. 7. ed. Rio de Janeiro: LTC, 2007.

[11] FERRÃO, M.F. . Tecno-lógica., Santa Cruz do Sul, Vol. 5, n.1, p.63-85, 2001.

[12] PERKIN ELMER, User Guide. Inc. Massachusetts: 2010.

[13] TRINDADE, M. M. et al.. Tecno-Lógica, Vol. 9, n. 1, p. 59-74, jan./jun. 2005.

[14] HELFER, G. A.; FERRÃO, M. F.; FERREIRA, C. V; HERMES, N. Publicação de métodos de análise multivariada no controle qualitativo de essências alimentícias empregando espectroscopia no infravermelho médio. Ciência e Tecnologia dos Alimentos, vol. 26, n. 4, dez. 2006.

[15] FERREIRA, M. M. C. et al.. Química Nova, vol. 22, n. 5, set. 1999.

[16] FERRÃO, M. F. Aplicação de técnicas espectroscópicas de reflexão no infravermelho no controle de qualidade de farinha de trigo. 2000. Tese (Doutorado) - Universidade Estadual de Campinas, 2000.

[17] LEARDI, R.; NORGAARD, L. Journal of Chemometrics, Vol. 18, n. 11, p. 486-497, 2004.

[18] NORGAARD, Lars. *i*Toolbox for MATLAB. University of Copenhagen. Disponível em: <<http://www.models.life.ku.dk/itoolbox>>. Acesso em: 07 mar. 2012.

[19] GUILLÉN, M. D.; CABO, N.. Journal of the Science of Food and Agriculture. N. 75, p. 1-11, 1997.