

CHEMOSTAT, UM SOFTWARE GRATUITO PARA ANÁLISE EXPLORATÓRIA DE DADOS MULTIVARIADOS

Gilson A. Helfer^{a,*,#}, Fernanda Bock^{a,#}, Luciano Marder^{a,#}, João C. Furtado^{b,#}, Adilson B. da Costa^{c,#} e Marco F. Ferrão^d

^aDepartamento de Química e Física, Universidade de Santa Cruz do Sul, 96815-900 Santa Cruz do Sul – RS, Brasil

^bDepartamento de Informática, Universidade de Santa Cruz do Sul, 96815-900 Santa Cruz do Sul – RS, Brasil

^cDepartamento de Biologia e Farmácia, Universidade de Santa Cruz do Sul, 96815-900 Santa Cruz do Sul – RS, Brasil

^dInstituto de Química, Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre – RS, Brasil

Resultados do Matlab®

Para realizar a validação da HCA com ligação completa (*complete linkage*), utilizaram-se as mesmas amostras e as mesmas regiões analisadas pelo HCA do ChemoStat (Figura 4 do artigo). A Figura 1S ilustra o resultado dessa análise no Matlab®.

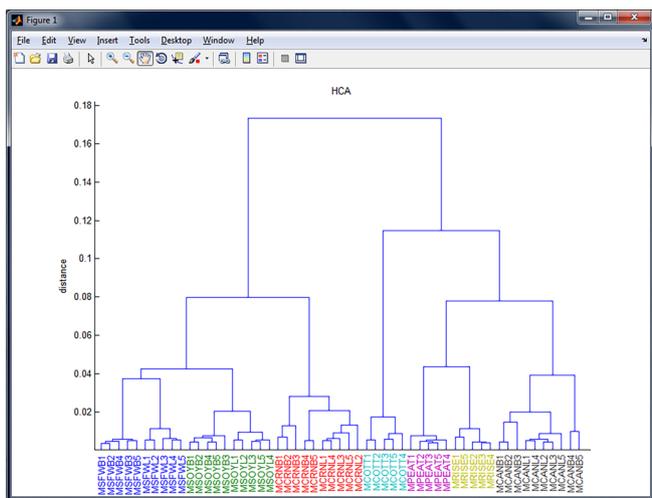


Figura 1S. Gráfico do tipo dendrograma HCA – ligação completa do Matlab®

A Figura 2S exibe o gráfico de *scores* a partir da análise dos componentes principais (PCA) aplicada ao conjunto de espectros dos óleos

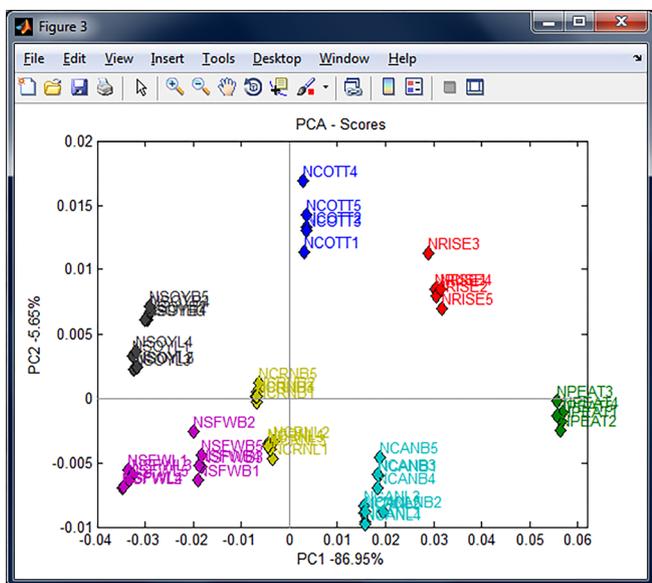


Figura 2S. Gráfico de scores PCA do Matlab®

vegetais na faixa entre 5500 e 6000 cm^{-1} , normalizados entre os limites zero e um, corrigidos pelo método do valor normal padrão (“SNV”) e, posteriormente, centrados na média, em comparação à Figura 5 do artigo.

A Figura 3S exibe o gráficos de loadings para a PC1 da Figura 2S, empregando o Matlab® em comparação à Figura 6 do artigo.

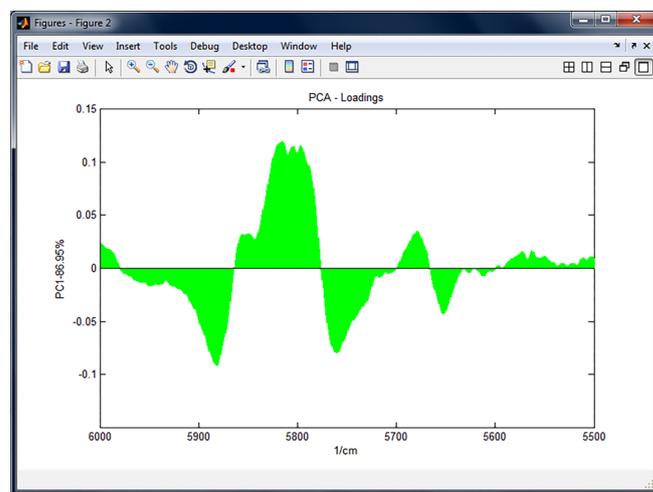


Figura 3S. Gráfico de loadings PCA do Matlab®

A Figura 4S apresenta as projeções dos *scores* e *loadings* utilizando a opção “Biplot”, para os óleos vegetais estudados na faixa espectral entre 5500 e 6000 cm^{-1} , empregando o Matlab® em comparação à Figura 7 do artigo.

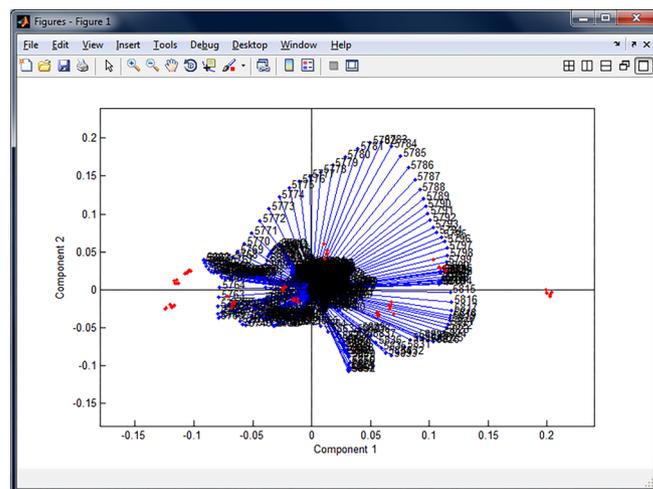


Figura 4S. Gráfico Biplot (scores x loadings) do Matlab®

*e-mail: ghelfer@gmail.com

#Programa de Pós-Graduação em Sistemas e Processos Industriais

Para detecção de amostras anômalas foi utilizado o cálculo da decomposição do T2 de Hotelling para dados multivariados, com

um nível de confiança de 95%, sendo atribuído à variável alpha da distribuição de Fisher-Snedecor, e ilustrado pela Figura 5S em comparação à Figura 8 do artigo.

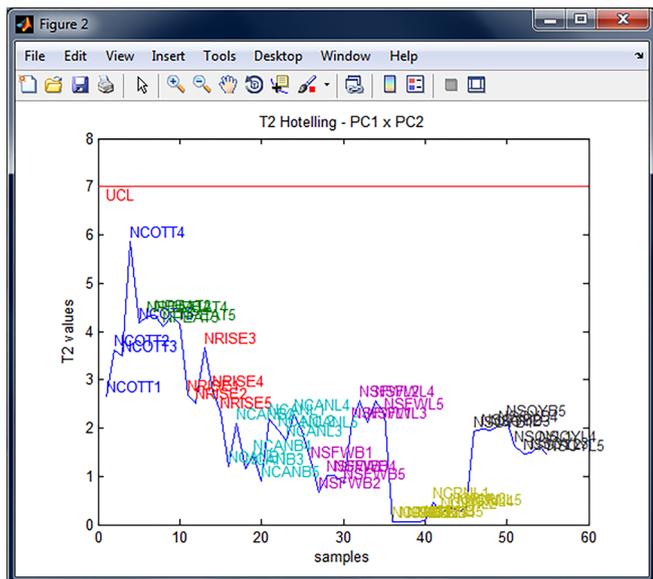


Figura 5S. Gráficos de Outliers a partir do T2 de Hotelling do Matlab®

Na comparação dos resultados do iPCA (Figura 9 do artigo), são apresentadas a Figura 6S e a Tabela 1S.

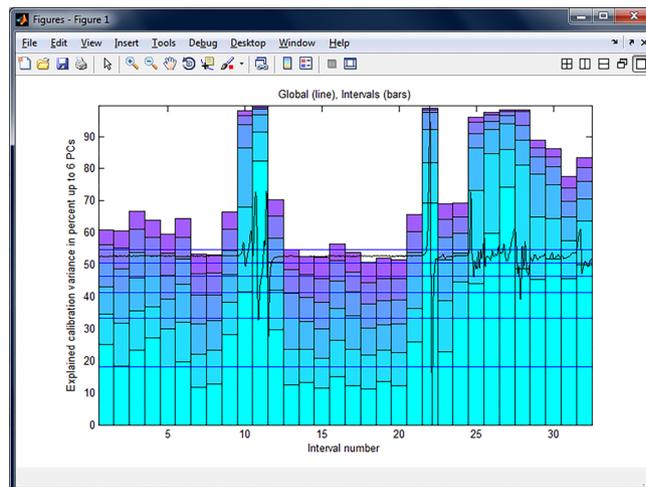


Figura 6S. Variação percentual das componentes principais divididas em 32 intervalos aplicados nos espectros de óleos vegetais (FT-MIR) - Matlab®

Tabela 1S. Valores para variância e variância acumulada das seis primeiras componentes principais dos dados de óleos vegetais - ChemoStat e Matlab® para o intervalo número 28

CP	ChemoStat var (%)	ChemoStat var acum (%)	Matlab® var (%)	Matlab® var acum (%)
1	48,71	48,71	48,71	48,71
2	32,40	81,12	32,40	81,12
3	7,69	88,81	7,69	88,81
4	4,72	93,52	4,72	93,52
5	4,33	97,86	4,33	97,86
6	0,54	98,39	0,54	98,39