

## EXPERIMENTO DIDÁTICO DE QUIMIOMETRIA PARA O MAPEAMENTO DE PELLETS FARMACÊUTICOS UTILIZANDO ESPECTROSCOPIA DE IMAGEM NA REGIÃO DO INFRAVERMELHO PRÓXIMO E RESOLUÇÃO MULTIVARIADA DE CURVAS COM MÍNIMOS QUADRADOS ALTERNANTES: UM TUTORIAL, PARTE IV

José Augusto Da-Col, Willian Francisco Cordeiro Dantas e Ronei Jesus Poppi\*

Departamento de Química Analítica, Instituto de Química, Universidade Estadual de Campinas, CP 6154, 13084-971 Campinas – SP, Brasil

Recebido em 24/05/2017; aceito em 11/09/2017; publicado na web em 08/11/2017

DIDACTIC EXPERIMENT OF CHEMOMETRICS FOR PHARMACEUTICAL PELLETS MAPPING USING NEAR-INFRARED SPECTROSCOPY AND MULTIVARIATE CURVE RESOLUTION WITH ALTERNATING LEAST SQUARES: A TUTORIAL, PART IV. The aim of this manuscript is to present a didactic experiment based on the combination of Near Infrared chemical imaging and Multivariate Curve Resolution for beginners, undergraduate and graduate students. In this tutorial, using MATLAB environment, the instrumental signals will be separated into two matrices, of spectral profiles and relative concentrations in the samples. Then, chemical images were generated to identify and verify the distribution of the components in the pellet samples. Chemical imaging was obtained from the mapping of predetermined area and the results for relative concentrations obtained in the ALS optimization for the analyzed area. This approach was successfully tested in chemistry undergraduate course at the Chemistry Institute of Campinas State University.

Keywords: NIR; chemical imaging; multivariate curve resolution; pharmaceutical pellets; teaching experiment.

### INTRODUÇÃO

A Quimiometria é uma área da Química extremamente difundida nos dias atuais e muito útil na extração de informações dos mais variados sistemas químicos. Além da pesquisa básica, a indústria química utiliza a análise multivariada para otimizar seus processos e, nesse contexto, formar profissionais com o conhecimento necessário é muito importante para o desenvolvimento do setor no país.<sup>1</sup> Assim, esse artigo pretende dar sequência à série de tutoriais já publicados pela Química Nova para o ensino de Quimiometria.<sup>2-4</sup> No primeiro desses estudos, a discriminação de óleos comerciais por espectroscopia no infravermelho médio foi utilizada para o ensino da Análise de Componentes Principais (*Principal Component Analysis*, PCA).<sup>2</sup> O segundo trabalho empregou a determinação da concentração do princípio ativo paracetamol em comprimidos com espectroscopia no infravermelho próximo e regressão por mínimos quadrados parciais (PLS, do inglês *Partial Least Squares*) para o ensino de calibração multivariada.<sup>3</sup> Por fim, no terceiro tutorial, a tradicional determinação de Ferro (III) com *o*-fenantrolina através da espectroscopia na região do visível foi usada para apresentar os conceitos do planejamento experimental (*design of experiments*, DOE).<sup>4</sup> Todos esses trabalhos foram aplicados em disciplinas de Química Analítica Instrumental lecionadas no Instituto de Química da Unicamp, com muito boa aceitação pelos alunos.

A proposta desse atual tutorial é apresentar uma visão diferenciada das análises por espectroscopia no infravermelho próximo (*Near Infrared Spectroscopy*, NIR), em que análises convencionais são substituídas por mapeamentos de regiões bem definidas e mais representativas da totalidade da amostra. Essa iniciativa pode ser mais eficiente quando se empregam ferramentas quimiométricas para o tratamento desses dados, como a Resolução Multivariada de Curvas (*Multivariate Curve Resolution*, MCR). Com isso, não só informações químicas das substâncias que constituem as amostras, mas também

informações relacionadas à distribuição desses constituintes no material podem ser obtidas.<sup>5</sup> Ao se aplicar essa estratégia, são geradas imagens hiperespectrais, que são, resumidamente, um conjunto de espectros de uma região delimitada da amostra, obtidos com uma técnica analítica com capacidade de responder à determinada propriedade físico-química do material. Para que seja gerada a imagem, a área da amostra deve ser subdividida em unidades menores e, para cada um desses pixels, deve ser obtido um espectro de absorção, emissão ou reflectância, por exemplo.<sup>6,7</sup> Na literatura, são encontrados muitos exemplos dessa abordagem, seja na discriminação de nanopartículas magnéticas dispersas em polímero,<sup>8</sup> na identificação de bactérias<sup>9</sup> ou na análise de produtos alimentícios como carne e presunto cru.<sup>10,11</sup>

Para a geração das imagens hiperespectrais, dentre várias técnicas analíticas, podem ser empregadas a espectroscopia Raman<sup>12</sup> ou a espectroscopia no infravermelho médio,<sup>13</sup> por exemplo. Porém, uma técnica muito útil é a espectroscopia no infravermelho, principalmente na região do infravermelho próximo (NIR), de 750 a 2.500 nm ou 13.000 a 4.000 cm<sup>-1</sup>, envolvendo fótons com energia compreendida entre 2,65 x 10<sup>-19</sup> a 7,96 x 10<sup>-20</sup> J.<sup>14</sup> Nesses valores de energia podemos observar vibrações de sobretons (primeiro, segundo e terceiro), que são aquelas que ocorrem do nível vibracional fundamental (n = 0) para os níveis vibracionais superiores (n = 2, 3, 4...). Pela regra de seleção para o infravermelho, as transições de sobretons não seriam permitidas, mas acabam acontecendo devido ao fenômeno da anarmonicidade mecânica, ou seja, pelas diferenças de massas dos átomos envolvidos na ligação química. Além dos sobretons, bandas de combinação também são observadas pela soma ou subtração dos números de onda, que acabam gerando uma banda única. Como na região do NIR muitos sobretons e bandas de combinação estão sobrepostos, não há picos específicos a serem atribuídos aos componentes químicos e as diferenças e semelhanças entre os espectros podem ser mais bem compreendidas com o auxílio da Quimiometria.<sup>14-16</sup> Quando se utiliza a espectroscopia NIR na geração de imagens hiperespectrais, temos então a espectroscopia vibracional de imagem (*Near Infrared-Chemical Imaging*, NIR-CI), que possui como uma das principais características poder operar com

\*e-mail: ronei@iqm.unicamp.br

reflectância difusa, diretamente na superfície da amostra.<sup>17</sup> Por ser uma técnica não destrutiva e por requerer pouco ou nenhum preparo de amostra,<sup>14</sup> NIR-CI tem sido recomendada pela agência norte-americana *Food and Drug Administration* (FDA) para uso no controle de processos farmacêuticos.<sup>18</sup> Sua contribuição consiste em verificar se um medicamento está adequado para a venda e consumo ou se o princípio ativo (*active pharmaceutical ingredient*, API) está disponível no produto. Essas informações são obtidas de forma rápida e limpa e, muitas vezes, sem consumo de reagentes e geração de resíduos tóxicos. Várias são as contribuições listadas na literatura para a área farmacêutica ao se empregar NIR-CI.<sup>19-27</sup>

Uma aplicação citada é o estudo e desenvolvimento de *pellets*, um dos vários tipos disponíveis de preparações sólidas para uso oral. *Pellets* são pequenas estruturas esféricas, com tamanho de partícula controlado e sua produção normalmente é feita pela aglomeração das substâncias na forma de pós finos ou por revestimento por filme, proporcionando a incorporação de drogas que sejam incompatíveis entre si. Esse revestimento por filme pode contribuir também para características específicas dos *pellets*, desde estéticas até a disponibilização uniforme e seletiva do API, evitando sinais de concentração no plasma ou a agressão da mucosa intestinal.<sup>28</sup> Além disso, a forma com que o API está disperso no medicamento pode afetar sua solubilização, o que se relaciona diretamente com a biodisponibilidade e, por consequência, com a eficácia terapêutica do fármaco.<sup>29</sup> Normalmente, essa heterogeneidade da amostra não pode ser avaliada por técnicas convencionais como HPLC ou UV, mas somente os valores globais de seus teores. Já a morfologia das dispersões sólidas pode ser avaliada por microscopia óptica ou eletrônica de varredura (MEV), sem fornecer informações químicas das amostras.<sup>30</sup> Portanto, conhecer a distribuição dos seus componentes é de grande importância para a qualidade do produto final.<sup>31,32</sup>

### Resolução multivariada de curvas com mínimos quadrados alternantes

As imagens geradas por NIR-CI contêm uma enorme quantidade de informação e esses conjuntos de espectros e sua interpretação pode ser realizada com métodos de análise multivariada. Tais métodos podem ser tanto de reconhecimento de padrões (como PCA), de classificação (como *Linear Discriminant Analysis*, LDA) ou de calibração multivariada (como PLS).<sup>33</sup> A resolução multivariada de curvas por mínimos quadrados alternantes (MCR-ALS) resolve misturas de sinais de forma iterativa, pela otimização dos perfis de concentração e dos espectros (matrizes **C** e **S<sup>T</sup>**) com a aplicação de restrições prévias. Essas restrições são baseadas no conhecimento químico ou matemático sobre as características dos dados.<sup>34,35</sup>

O método MCR pode ser classificado como uma técnica de análise de fatores,<sup>36</sup> na qual se busca isolar as fontes de variação de um conjunto de dados, além de resolver e quantificar tais variações. Para isso, a matriz dos dados instrumentais (**D**) é decomposta pelo algoritmo de quadrados mínimos alternados (ALS) em uma matriz referente à concentração relativa de cada substância pura (**C**) e aos espectros dessas substâncias (**S**). A expressão para o MCR é descrita na Equação 1, onde **E** é uma matriz de resíduos.<sup>37</sup>

$$\mathbf{D} = \mathbf{CS}^T + \mathbf{E} \quad (1)$$

Para a resolução da equação anterior pelo algoritmo ALS é necessário construir uma estimativa inicial, para que a otimização possa ser começada. Se optarmos por uma estimativa para a matriz **S**, ela pode ser conseguida através de uma informação já conhecida, como espectros das substâncias constituintes do sistema ou um sinal simulado por uma ferramenta matemática que consiga extrair as informações menos aleatórias do sinal original. Além dessa informação, restrições

podem ser aplicadas para limitar as possibilidades de soluções para o sistema. Um exemplo são restrições de não-negatividade, em que soluções que contemplem valores negativos para o sinal (**S**) ou para as concentrações relativas (**C**) são desconsideradas. Assim, as matrizes **C** e **S<sup>T</sup>** podem ser construídas em ciclos iterativos de duas etapas, conforme mostrado nas Equações 2 e 3:<sup>38</sup>

$$\mathbf{C} = \mathbf{DS}(\mathbf{S}^T\mathbf{S})^{-1} \quad (2)$$

$$\mathbf{S}^T = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{D} \quad (3)$$

Uma vez estimadas as matrizes **C** e **S** nas etapas indicadas pelas Equações 2 e 3, seu produto pode ser calculado (**CS<sup>T</sup>**) e uma nova matriz, **D\***, é reconstruída a cada iteração. Essas etapas são repetidas até que algum critério de convergência estabelecido previamente seja alcançado, como um número máximo de iterações a ser realizado ou caso a variação entre o resultado da iteração atual e aquele da iteração anterior seja menor que um valor pré-estabelecido.<sup>37,38</sup> É importante ressaltar que, ainda que a convergência seja alcançada nessa etapa iterativa, a solução obtida não é necessariamente a melhor resposta para o conjunto de dados estudado, pois há sempre o risco de se alcançar um mínimo local apesar das restrições pré-definidas.

Dessa comparação da matriz de dados originais **D** com essa nova matriz **D\*** podem ser calculados dois importantes parâmetros, que fornecem informações sobre a qualidade do modelo construído: a porcentagem de falta de ajuste (*Lack of Fit*, %LOF) e a porcentagem de variância explicada ( $R^2$ ). O primeiro desses parâmetros é calculado conforme descrito na Equação 4 e compara cada termo da nova matriz com o termo correspondente da matriz original. Como o desejado é que essas matrizes sejam as mais semelhantes, os valores dos erros ( $e$ ) devem ser pequenos e o valor de %LOF deverá estar mais próximo a zero.<sup>37,38</sup>

$$\%LOF = 100 \sqrt{\frac{\sum e_{ij}^2}{\sum d_{ij}^2}} \quad (4)$$

O segundo parâmetro segue o mesmo raciocínio, comparando as duas matrizes, mas dessa vez descrevendo o quanto da informação original está sendo modelada pelo MCR, ou seja, o quanto da matriz original **D** está contida na nova matriz **D\***. A forma de calcular a porcentagem da variância explicada é descrita na Equação 5 e o ideal é que 100% da informação original seja expressa pelo modelo, sendo portanto desejado um valor para  $R^2$  próximo a esse.<sup>37,38</sup>

$$R^2 = 100 \left( 1 - \frac{\sum e_{ij}^2}{\sum d_{ij}^2} \right) \quad (5)$$

Com a intenção de difundir e popularizar o uso do MCR como ferramenta analítica, algumas boas iniciativas foram desenvolvidas, como por Jaumot *et al.*,<sup>35</sup> em 2005. Os autores disponibilizaram gratuitamente uma interface gráfica para ambiente MATLAB® que permite, de forma intuitiva, a escolha do número de componentes, da estimativa inicial e também permite a aplicação de restrições. Isso facilita o uso principalmente para usuários menos familiarizados com o software. Essa ferramenta foi sendo aperfeiçoada e hoje conta com uma atualização, que oferece opções adicionais de restrições, como a restrição de correlação e a inclusão de modelos pré-definidos (*hard-modeling*), além de formas mais eficientes de apresentação dos resultados.<sup>39</sup> Em 2014, Março *et al.*<sup>37</sup> publicaram uma completa revisão, em língua portuguesa, descrevendo os aspectos teóricos e os avanços do método até então. São discutidos pontos como o uso de estimativas iniciais, a otimização pelo ALS e o uso de restrições para minimizar o efeito de ambiguidade rotacional e sua leitura é recomendada para aqueles que desejarem se aprofundar no tema.

Pelo exposto, observa-se que o método de MCR tem grande utilidade na previsão da concentração e na distribuição dos constituintes das amostras farmacêuticas, sendo viável seu uso em processos industriais e no desenvolvimento de medicamentos.<sup>25,40-45</sup> O que se propõe então, nesse tutorial, é apresentar uma opção para o ensino desses tópicos em disciplinas experimentais de Química Analítica, baseando-se em experiências bem-sucedidas no Instituto de Química da Unicamp.

### Pré-tratamento

Uma realidade encontrada pelo analista é que, junto da informação desejada, informações não relevantes são obtidas no momento da análise e uma boa parte dela pode ser removida em uma etapa criteriosa de pré-tratamento dos dados. Um exemplo são variações presentes em espectros obtidos por reflectância difusa em NIR, que ocorrem basicamente por fenômenos de espalhamento aditivo e multiplicativo da radiação na superfície das partículas. O fator mais importante para o espalhamento da radiação está relacionado com o tamanho das partículas do material, sendo um fenômeno essencialmente físico e que normalmente se manifesta como uma variação na inclinação do espectro. Dentre as possibilidades para a correção desses efeitos,<sup>46</sup> uma delas é o uso da Padronização Normal de Sinal (*Standard Normal Variate*, SNV), que corresponde a autoescalar cada linha da matriz original dos dados.<sup>47</sup> São calculados o espectro médio ( $\bar{x}$ ) e o desvio padrão ( $\sigma_n$ ) das amostras, onde  $n$  é o número de amostras e, por consequência, de espectros (Equação 6).

$$SNV_n = (x_n - \bar{x})/\sigma_n = (x_n - \bar{x})/\sqrt{\frac{\sum(x_n - \bar{x})^2}{n-1}} \quad (6)$$

Outra situação em que a correção se faz necessária é na presença de ruído aleatório ao sinal instrumental, composto de erros experimentais, erros de medida e de outras fontes de variabilidade, como oscilações de corrente elétrica ou variações de temperatura durante a medida. Esse ruído pode ser reduzido, mas não completamente eliminado e essa etapa deve ser muito bem conduzida para também não remover sinal verdadeiro. Uma forma de remover ruído instrumental é por alisamento, ou seja, pela aplicação de um filtro que possibilitará a melhora da razão sinal/ruído (razão entre a intensidade do sinal analítico e a raiz da variância associada ao sinal). Dentre as técnicas de alisamento mais empregadas, estão aquelas que utilizam uma janela móvel, que percorrerá todo o espectro.<sup>1</sup> Definido o tamanho dessa janela, todos os pontos nesse intervalo serão utilizados para gerar uma única resposta que será atribuída ao centro da janela. Um método muito comum para esse fim é o método Savitzky-Golay,<sup>48</sup> que gera uma média ponderada com pesos obtidos por um ajuste polinomial como a resposta do ponto central. A janela é movida e o polinômio é ajustado novamente, num processo que só termina quando todas as janelas tiverem sido alisadas. A escolha do grau do polinômio afeta grandemente o resultado do alisamento, com polinômios de graus mais baixos removendo mais ruído, mas com risco maior de distorção do sinal. Da mesma forma, o uso de janelas maiores causa diminuição do ruído, mas também apresenta risco maior de distorcer o sinal. A escolha desses parâmetros deve, portanto, ser cuidadosamente avaliada.<sup>1</sup>

## PARTE EXPERIMENTAL

### Materiais e reagentes

Para o estudo, foi utilizado um *pellet* comercial cujo princípio ativo é o cetoprofeno, um fármaco da classe dos anti-inflamatórios

não-esteroides e que apresenta ação anti-inflamatória e analgésica.<sup>49</sup> Esses *pellets* são constituídos de um núcleo inerte e de camadas que recobrem a região central, possuindo diâmetro aproximado de 900  $\mu\text{m}$ . Para a análise, as amostras foram cortadas ao meio com auxílio de lâmina de barbear em uma lupa (Carl Zeiss, modelo Stemi DV4, 8 a 32x de magnificação). Após isso, foram afixadas em uma placa de vidro com fita dupla face transparente, para que fossem obtidos os dados que irão corresponder às imagens químicas dos componentes da amostra.

Para a comparação dos sinais instrumentais originais com os recuperados pelo modelo MCR, foram usados os reagentes de grau analítico cetoprofeno (Sigma-Aldrich), amido (Synth), sacarose (Synth) e dióxido de silício (Sigma-Aldrich), que foram acondicionados em um porta-amostras de alumínio, cuja cavidade foi completada com o pó do composto químico. Para minimizar o espalhamento da radiação, a superfície foi nivelada para que ficasse o mais uniforme possível.

### Condições operacionais

A aquisição dos espectros das amostras de *pellets* e das substâncias puras foi realizada no espectrômetro Spectrum 100 N FT-NIR, acoplado a um sistema de imagem Spotlight 400 N, ambos da Perkin-Elmer. As condições operacionais utilizadas são informadas na Tabela 1.

**Tabela 1.** Condições operacionais do espectrômetro Spotlight 400N FT-NIR Imaging System nas análises dos pellets e das substâncias puras

Condições operacionais	Imagem do pellet	Substâncias puras
Tamanho de pixel ( $\mu\text{m}$ )		25
Resolução ( $\text{cm}^{-1}$ )		16
Número de <i>scans</i>		32
Faixa espectral ( $\text{cm}^{-1}$ )		7800 – 4000
Área ( $\mu\text{m}$ )	1300 x 1300	300 x 300
Área (pixel)	64 x 64	12 x 12

### Tratamento dos dados experimentais

Os espectros foram obtidos em modo de reflectância (R) e convertidos para pseudo-absorbância ( $\log 1/R$ ) e, como pré-tratamento, foi aplicada a padronização normal de sinal,<sup>47</sup> para minimizar possíveis efeitos de espalhamento da radiação e alterações causadas pelo tamanho de partículas. Os dados também foram alisados com janela de 19 pontos para remoção de ruído experimental pelo método Savitzky-Golay.<sup>48</sup>

Foi realizada uma etapa de seleção de espectros para a remoção dos pixels referentes à placa de vidro e à fita adesiva através da escolha de valores de escores de uma PCA previamente realizada com os dados centrados na média. Assim, foi selecionada uma faixa de valores de escores que corresponde apenas aos espectros dos *pellets*. Esse processo pôde ser realizado de forma gráfica com o pacote HYPER-Tools, que foi criado especificamente para o tratamento de imagens químicas.<sup>50</sup> Por essa razão e também por estar livremente disponível, foi empregado no estudo.

Com esse novo conjunto de dados, um modelo MCR foi construído com 4 componentes. A escolha do número de componentes levou em consideração o resultado da PCA com os dados centrados na média, que explicaram mais de 99% da variância.

Como estimativa inicial para o modelo MCR, foi construída uma matriz com o espectro médio de cada uma das substâncias puras, ou seja, a média dos 144 espectros obtidos experimentalmente para cada uma dessas 4 substâncias. Como condições para o modelo

MCR, foram impostas as restrições de máximo de 500 iterações, erro inferior a 0,1% entre as iterações, além da não negatividade para a concentração.

Todos os cálculos foram efetuados em ambiente MATLAB R2016b (MathWorks, Natick, MA, USA), utilizando os pacotes gratuitos HYPER-Tools<sup>50</sup> e MCR-ALS.<sup>51</sup> Como alternativa gratuita ao software MATLAB, pode ser encontrada uma versão em ambiente R no site *Comprehensive R Archive Network* (CRAN).<sup>52</sup> Já para modelos MCR-ALS, existe um pacote, também gratuito, disponível no mesmo site, na seção *Chemometrics and Computational Physics*.<sup>53</sup>

## RESULTADOS E DISCUSSÃO

### Tratamento dos dados

(1) Uma vez finalizadas as análises das amostras e das substâncias puras, conforme indicado na seção Condições operacionais, são gerados pelo equipamento arquivos com a extensão .fsm (extensão específica para esse tipo de dados para os equipamentos da Perkin-Elmer), que deverão ser copiados para o diretório de trabalho do MATLAB. O primeiro passo é construir uma estrutura de dados para conter cada um dos sinais (*Pellet*, Sacarose, Amido, dióxido de silício e Cetoprofeno).

```
>>DADOS(1).nomes = 'Pellet';
>>DADOS(2).nomes = 'Sacarose';
>>DADOS(3).nomes = 'Amido';
>>DADOS(4).nomes = 'SiO_2';
>>DADOS(5).nomes = 'Cetoprofeno';
```

(2) Como arquivos com extensão .fsm não são reconhecidos pelo MATLAB, é necessário que esses dados sejam extraídos com o emprego da rotina fsmload.m, fornecida pelo fabricante do equipamento, a Perkin-Elmer. Para que essa etapa seja completada, os arquivos devem ter os nomes *pellet.fsm*, *sacarose.fsm*, *amido.fsm*, *sio2.fsm* e *cetoprofeno.fsm* ou a referência a eles nas linhas de comando.

```
>>[DADOS(1).cubo,~,~,SIZE.zAxis,~]=fsmload('pellet.fsm');
>>DADOS(2).cubo=fsmload('sacarose.fsm');
>>DADOS(3).cubo=fsmload('amido.fsm');
>>DADOS(4).cubo=fsmload('sio2.fsm');
>>DADOS(5).cubo=fsmload('cetoprofeno.fsm');
```

(3) As matrizes de dados criadas possuem três dimensões, sendo duas delas relacionadas com as dimensões espaciais e a terceira, com a dimensão espectral. Para o *pellet*, tem-se a dimensão 52x52x476 e para os demais compostos puros, 12x12x476. Porém, para poder ser resolvido pelo MCR, os cubos de dados devem ser desdobrados em matrizes de duas dimensões, o que é feito com a rotina reshape.

```
>>[SIZE.nX,SIZE.nY,SIZE.nZ]=size(DADOS(1).cubo);
>>DADOS(1).brutos = reshape(DADOS(1).cubo,SIZE.nX*SIZE.nY,SIZE.nZ);
>>[SIZE.nx,SIZE.ny,SIZE.nz]=size(DADOS(2).cubo);
>>DADOS(2).brutos = reshape(DADOS(2).cubo,SIZE.nx*SIZE.ny,SIZE.nz);
>>DADOS(3).brutos = reshape(DADOS(3).cubo,SIZE.nx*SIZE.ny,SIZE.nz);
>>DADOS(4).brutos = reshape(DADOS(4).cubo,SIZE.nx*SIZE.ny,SIZE.nz);
>>DADOS(5).brutos = reshape(DADOS(5).cubo,SIZE.nx*SIZE.ny,SIZE.nz);
```

(4) Os dados brutos podem ser graficados em figuras, conforme

descrito abaixo. O mesmo tipo de rotina é empregado para a geração das demais figuras no artigo, com adequações nas matrizes e dimensões de cada caso.

```
>>[~,SIZE.k]=size(DADOS);
>>for i=1:SIZE.k;
>> subplot(1,ceil(SIZE.k),i);plot(SIZE.zAxis,DADOS(i).brutos');
>> xlabel('Número de onda (cm^{-1})','FontSize',12.5,'FontWeight','bold');
>> ylabel('Reflectância','FontSize',12.5,'FontWeight','bold');
>> if SIZE.zAxis(1) > SIZE.zAxis(2)
>> set(gca,'XDir','reverse')
>> end
>> axis square
>> axis tight
>> h = title([(DADOS(i).nomes) ' Brutos']);
>> set(h,'Color','black','FontSize',12.5,'FontWeight','bold');
>> set(gca,'FontWeight','bold','FontSize',12)
>>end
>>clear i h
```

Os espectros para os dados brutos do *pellet*, em valores de reflectância, podem ser vistos na Figura 1. Os demais espectros, para as substâncias puras, são apresentados na seção Material Suplementar (Figura 1S).

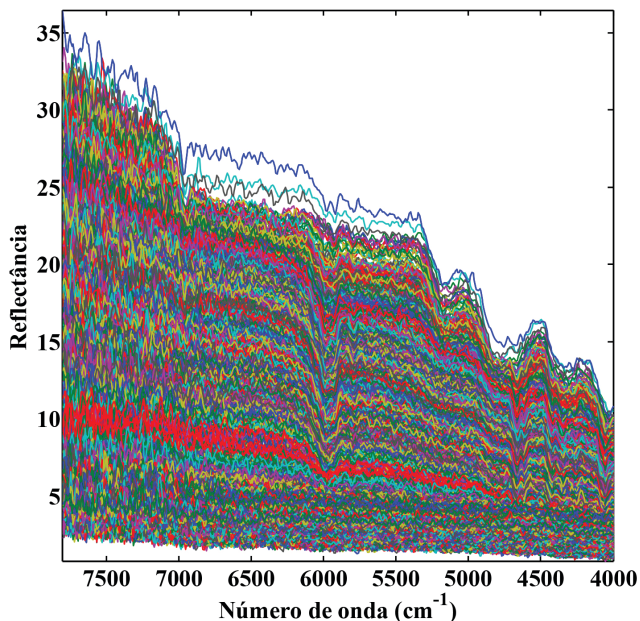


Figura 1. Espectros NIR sobrepostos em reflectância para os dados brutos da amostra do pellet

(5) Para o tratamento dos dados espectrais de NIR, normalmente realiza-se uma etapa de conversão dos valores de reflectância (R) para  $\log 1/R$ , que é conhecida como pseudoabsorbância. Isso é necessário para que se possa relacionar a concentração diretamente com a intensidade do sinal medido, numa forma análoga à lei de Beer-Lambert. Porém, como essa lei é válida somente para materiais homogêneos e transparentes à radiação, Kubelka e Munk,<sup>16,54</sup> desenvolveram um tratamento matemático que consegue relacionar valores de reflectância difusa com as concentrações das espécies presentes em uma determinada amostra. O tratamento matemático é bastante rigoroso, mas pode ser simplificado uma vez que na maioria das aplicações há um comportamento linear do sinal com a concentração. Nesses casos,



passa a ser válida a aproximação  $\log 1/R$ .

```
>>DADOS(1).absorbancia = (log10(1./(DADOS(1).brutos./100)));
>>DADOS(2).absorbancia = (log10(1./(DADOS(2).brutos./100)));
>>DADOS(3).absorbancia = (log10(1./(DADOS(3).brutos./100)));
>>DADOS(4).absorbancia = (log10(1./(DADOS(4).brutos./100)));
>>DADOS(5).absorbancia = (log10(1./(DADOS(5).brutos./100)));
```

Dessa forma, são obtidos os seguintes perfis dos espectros de pseudoabsorbância sobrepostos para a amostra do *pellet* (Figura 2). Na seção Material Suplementar (Figura 2S), são apresentadas as imagens dos espectros em pseudoabsorbância para as substâncias puras.

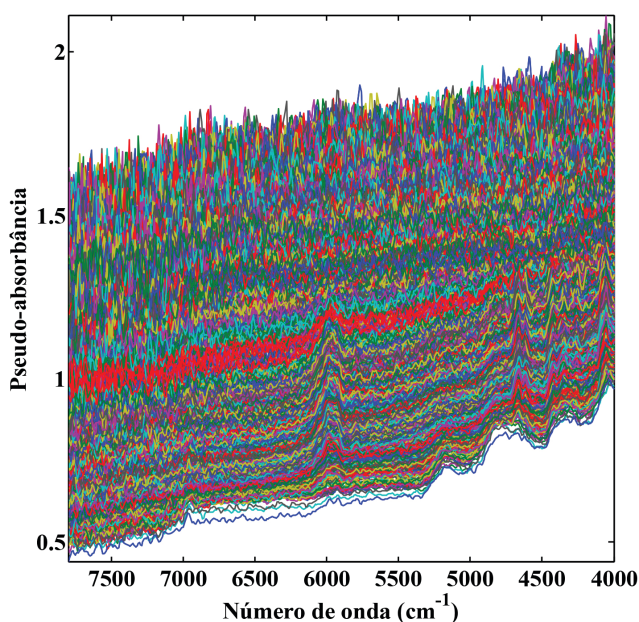


Figura 2. Espectros NIR sobrepostos em pseudo-absorbância para os dados brutos da amostra do *pellet*

(6) Uma etapa importante para o bom desenvolvimento do modelo MCR é o pré-tratamento dos dados, com ferramentas que resolvem problemas intrínsecos das análises instrumentais. Em análises com NIR, possíveis efeitos de espalhamento da radiação e alterações causadas pelo tamanho de partículas podem proporcionar variações nos sinais, como aumentos na linha de base. Para tanto, aplica-se a padronização normal de sinal (SNV), que minimiza esses efeitos. Outra consideração é sobre o ruído instrumental causado, por exemplo, por variações de temperatura no detector e oscilações eletrônicas, que podem ser corrigidas através de alisamento. Nesse estudo, aplicou-se o filtro Savitzky-Golay, com janela de 19 pontos. As funções para o SNV e alisamento estão disponíveis no pacote HYPER-Tools e são, respectivamente, *htsnv* e *filpoly*. As instruções para uso podem ser consultadas com o comando `open htsnv` ou `open filpoly`.

```
>>DADOS(1).snv = htsnv(DADOS(1).absorbancia);
>>DADOS(1).alisamento= filpoly(DADOS(1).snv,[],[-9 9],2,0,[0 0 0]);
>>DADOS(2).snv = htsnv(DADOS(2).absorbancia);
>>DADOS(2).alisamento= filpoly(DADOS(2).snv,[],[-9 9],2,0,[0 0 0]);
```

```
>>DADOS(3).snv = htsnv(DADOS(3).absorbancia);
>>DADOS(3).alisamento= filpoly(DADOS(3).snv,[],[-9 9],2,0,[0 0 0]);
>>DADOS(4).snv = htsnv(DADOS(4).absorbancia);
>>DADOS(4).alisamento= filpoly(DADOS(4).snv,[],[-9 9],2,0,[0 0 0]);
>>DADOS(5).snv = htsnv(DADOS(5).absorbancia);
>>DADOS(5).alisamento= filpoly(DADOS(5).snv,[],[-9 9],2,0,[0 0 0]);
```

Com o pré-tratamento, os espectros do *pellet* passaram a apresentar o perfil visualizado na Figura 3 e os das demais substâncias podem ser vistos na seção Material Suplementar, na Figura 3S.

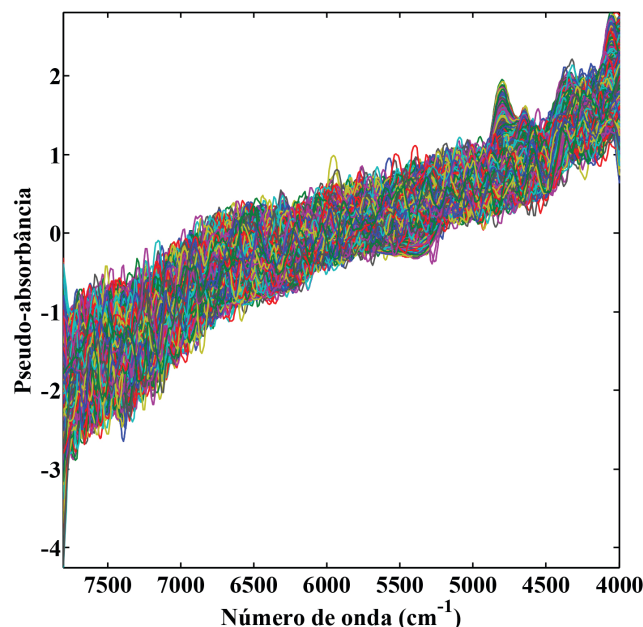


Figura 3. Espectros normalizados e alisados para os dados da amostra do *pellet*

(7) Após os dados serem processados, o passo seguinte é criar a matriz de estimativas que será usada para iniciar o modelo MCR. Para tanto, deve-se fazer a média dos 144 espectros obtidos instrumentalmente para cada uma das 4 substâncias puras, gerando 4 espectros médios (Figura 4), cada um deles em uma linha dessa matriz de estimativas.

```
>>DADOS(1).estimativa=[mean(DADOS(2).alisamento);mean(-DADOS(3).alisamento);mean(DADOS(4).alisamento);mean(-DADOS(5).alisamento)];
```

(8) Uma vez criada a matriz de estimativas para iniciar o modelo, é conveniente realizar uma última etapa de tratamento dos dados experimentais. Quando a região escolhida é mapeada, além da superfície do *pellet*, há uma área vizinha à amostra que gera sinal da lâmina de vidro e da fita adesiva. Se o modelo fosse criado com todo o sinal, ao menos uma componente adicional seria necessária para descrever essa região, o que aumentaria a complexidade do sistema e até mesmo poderia comprometer a capacidade de descrever a informação de interesse. Nesse momento, realiza-se uma PCA com os dados brutos centrados na média e, através dos valores dos escores, seleciona-se uma faixa que represente apenas o sinal da amostra, removendo-se o sinal que não é de interesse. Esse procedimento pode ser realizado com o uso da rotina *htmaskingtool*, disponibilizada gratuitamente no pacote HYPER-Tools.<sup>50</sup> Cabe ressaltar que, para o perfeito funcionamen-

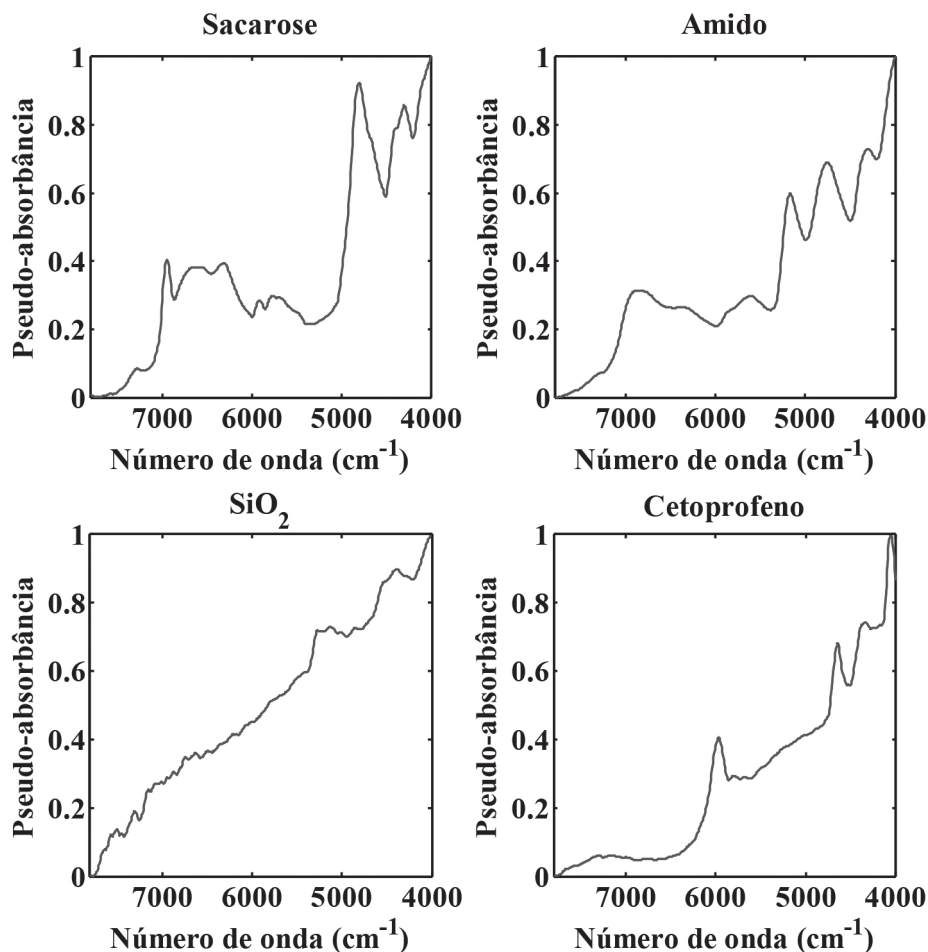


Figura 4. Espectros médios para as substâncias puras, que serão utilizados como estimativa inicial do modelo MCR-ALS

to do pacote HYPER-Tools, o pacote *Image Processing Toolbox*, que é comercializado pela Mathworks, deve estar instalado no MATLAB®. Caso isso não seja possível, a única alternativa é a de se escolher a região pelos valores dos escores da PCA, como indicado a seguir (as opções que devem ser utilizadas estão indicadas em negrito). De qualquer forma, essa opção é capaz de gerar uma seleção dos pixels muito eficiente.

```
>>MCR(1).mascara=htmaskingtool(reshape(DADOS(1).brutos,SIZE.nX,SIZE.nY,SIZE.nZ),[])
>>Masking using PCA and histograms --> 1
>>Masking using ROI in one area --> 2
>>Masking using ROI in more than one area --> 3
>>Please, select the type of masking and press enter 1
>>Based on PCA --> 1
>>Based on average image --> 2
>>Please, select the method and press enter 1
>>Mean center, press --> 1
>>NO mean center, press --> 0
>>Do you want to mean center the spectra? 1
>><SPACE>
>>Select component for masking 1
>><ENTER>
```

Seleciona-se a região de interesse observando-se os valores correspondentes ao *pellet* na barra de cores da imagem superior. Para o caso específico, essa região deve começar na região inicial, à esquerda, e ir até próximo a 30 na escala do gráfico inferior. Se a região escolhida estiver correta, finaliza-se o processo clicando na barra de espaço e respondendo à pergunta abaixo com o número 1.

Caso seja necessário repetir a seleção, deve-se responder à questão com 0 e escolher nova região, até que a escolha seja adequada.

```
>><SPACE>
>>Are you happy with the threshold? yes (1) / no(0) 1
```

A região é, portanto, selecionada graficamente e as etapas descritas acima podem ser visualizadas nas Figuras 4S e 5S, disponíveis no Material Suplementar.

(9) Escolhidos os pixels referentes à amostra, o passo a seguir é construir uma matriz com os espectros selecionados. Isso causa uma diminuição da dimensão da matriz dos dados, que originalmente era composta de 2704 espectros e, após a etapa de seleção, passa a ter 1158 linhas, cerca de 40%. Se por um lado a quantidade total de informação diminui, por outro lado essa informação é mais seletiva, já que corresponde apenas ao *pellet*. Isso reduz a complexidade da informação a ser modelada e facilita a sua interpretação ao final do processo. Os espectros selecionados podem ser vistos na Figura 5.

(10) Criada a estimativa inicial com os sinais instrumentais das substâncias puras (passo 7) e selecionados os espectros referentes ao *pellet* (passo 8 e 9), o modelo MCR pode ser executado através da rotina disponibilizada por Tauler em <http://www.mcrals.info/>. Essa é uma etapa onde devem ser informadas algumas condições, como a matriz de dados originais (**MCR(1).pixelintotal**), a matriz com a estimativa inicial (**DADOS(1).estimativa**), o número de iterações do modelo (100) e o valor máximo permitido para a

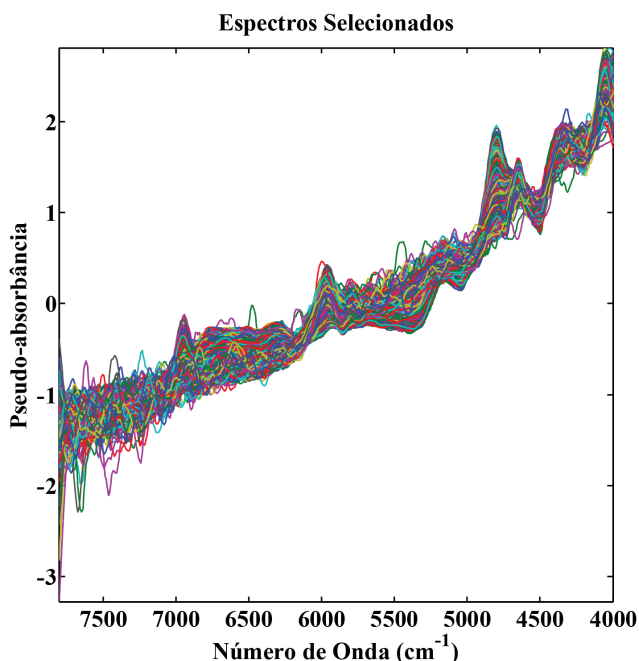


Figura 5. Espectros do pellet após a seleção dos pixels pela ferramenta de criação de máscara *htmaskingtool*, disponível no pacote *HYPER-Tools*

falta de ajuste (0,1%)

```
>>[MCR(1).copt,MCR(1).sopt]=als(MCR(1).
pixelinttotal,DADOS(1).estimativa,1,100,0.1);
```

Ao se executar a linha de comando acima, a rotina pedirá que sejam informadas as opções e restrições do modelo MCR, indicadas em negrito nas linhas abaixo:

```
>>INPUT TYPE OF CONSTRAINTS TO BE APPLIED
>>None (0)
>>Non-negativity (1)
>>Unimodality (2)
>>Closure (3)
>>Equality (known) / Lower than (selectivity) constraints in
conc profiles (4)
>>Equality (known) / Lower than (selectivity) constraints in
spectra profiles (5)
>>Shape constraints (6)
>>Three-way data structure constraints (7)
>>Enter a vector with the selected constraints, e.g. [1,3,5]
>>Enter the constraints to be applied in the optimization: 1
>>Enter the non-negativity constraints (1=conc / 2=conc and
spectra / 3=spectra): 1
>>Enter the selected algorithm for conc (0 = forced to zero / 1
= nnls / 2 = fnls): 2
>>normalization 0, 1 or 2?: 0
>>Do you want graphic output during the ALS optimization
(y/n)? y
>>Is all the input information correct (y/n)? y
```

Nesse ponto, a rotina irá fornecer algumas informações, que deverão ser copiadas manualmente, já que não são salvas de forma automática. As principais figuras de mérito são a porcentagem de variância explicada (%R<sup>2</sup>) e a porcentagem de falta de ajuste (%LOF) e são elas que indicam a qualidade do modelo criado. A porcentagem de variância explicada está relacionada com o quanto da informação original, na forma do sinal instrumental, está sendo descrita pelo modelo MCR. Quanto maior for a variância explicada, ou seja, quanto mais próximo de 100% for o valor

calculado, mais fiel será o modelo criado em relação aos sinais originais. No caso calculado, o valor foi de 99,91% e indica que apenas uma pequena parcela da informação original não pode ser expressa e está contida na matriz de resíduos (**E**). A mesma tendência é verificada pela porcentagem de falta de ajuste, que demonstra a diferença entre a matriz de dados originais (**D**) e a matriz reproduzida pelo modelo MCR (matriz **D**<sup>\*</sup> ou **CS**<sup>T</sup>).<sup>35</sup> Como esses valores devem ser inferiores a um e os menores possíveis, o valor de 0,00726 garante que a informação recuperada pelo modelo é similar ao sinal original.

- (11) Agora podemos comparar os perfis espectrais das substâncias puras com aqueles recuperados pelo MCR, que nos permitirá a possibilidade de identificar os componentes do sistema. Um parâmetro que auxilia nessa verificação é o coeficiente de correlação (r<sup>2</sup>), que é relacionado com a matriz de covariância de cada espectro, ou seja:

$$r(i,j) = \frac{C(i,j)}{\sqrt{C(i,i)C(j,j)}} \quad (7)$$

```
>>a=corrcoef(MCR(1).sopt(1,:),DADOS(1).estimativa(1,:));
>>b=corrcoef(MCR(1).sopt(2,:),DADOS(1).estimativa(2,:));
>>c=corrcoef(MCR(1).sopt(3,:),DADOS(1).estimativa(3,:));
>>d=corrcoef(MCR(1).sopt(4,:),DADOS(1).estimativa(4,:));
>>MCR(1).coeficientes=[a(1,2) b(1,2) c(1,2) d(1,2)];
>>clear a b c d
```

Pelos resultados, podemos afirmar que há uma excelente correlação entre os perfis para os componentes sacarose (r<sup>2</sup> = 0,9867), amido (r<sup>2</sup> = 0,9861) e cetoprofeno (r<sup>2</sup> = 0,9970), com bandas coincidentes e altos valores de coeficiente de correlação, conforme visto na Figura 6.

Para o dióxido de silício, não existem tantas bandas coincidentes, mas sim uma tendência de similaridade entre os perfis. Esse comportamento pode ser atribuído a uma baixa concentração do componente no *pellet* e também à sua natureza química. O dióxido de silício coloidal apresenta baixa densidade, de 0,029 a 0,042 g cm<sup>-3</sup>,<sup>41</sup> o que faz com que grande parte da radiação seja difundida em seu interior, não retornando ao detector e dificultando sua identificação. Isso se reflete no menor valor obtido para o coeficiente de correlação, de 0,9522.

- (12) Como a matriz de dados foi desdobrada para a realização do tratamento quimiométrico (passo 3), uma vez terminado o modelo MCR devemos retomar a dimensão original dos dados, recolocando os espectros selecionados pela máscara (passo 8) em sua posição original.

```
>>[SIZE.nx,SIZE.ny]=size(MCR(1).copt);
>>MCR(1).novo_copt=nan(SIZE.nx*SIZE.ny,SIZE.ny);
>>for i=1:MCR(1).comprimento;
>>MCR(1).novo_copt(MCR(1).mascara.
pixelinttotal(i,1),:)=MCR(1).copt(i,:);
>>end
>>clear i
```

- (13) Dessa forma, imagens com 52 x 52 pixels podem ser construídas para cada um dos componentes resolvidos no modelo MCR-ALS, como indicado abaixo (Figura 7).

```
>>[SIZE.k,~]=size(DADOS(1).estimativa);
>>for i=1:SIZE.k;
>>subplot(1,ceil(SIZE.k),i);imagesc(reshape(MCR(1).novo_
copt(:,i),SIZE.nx,SIZE.ny,1));
```

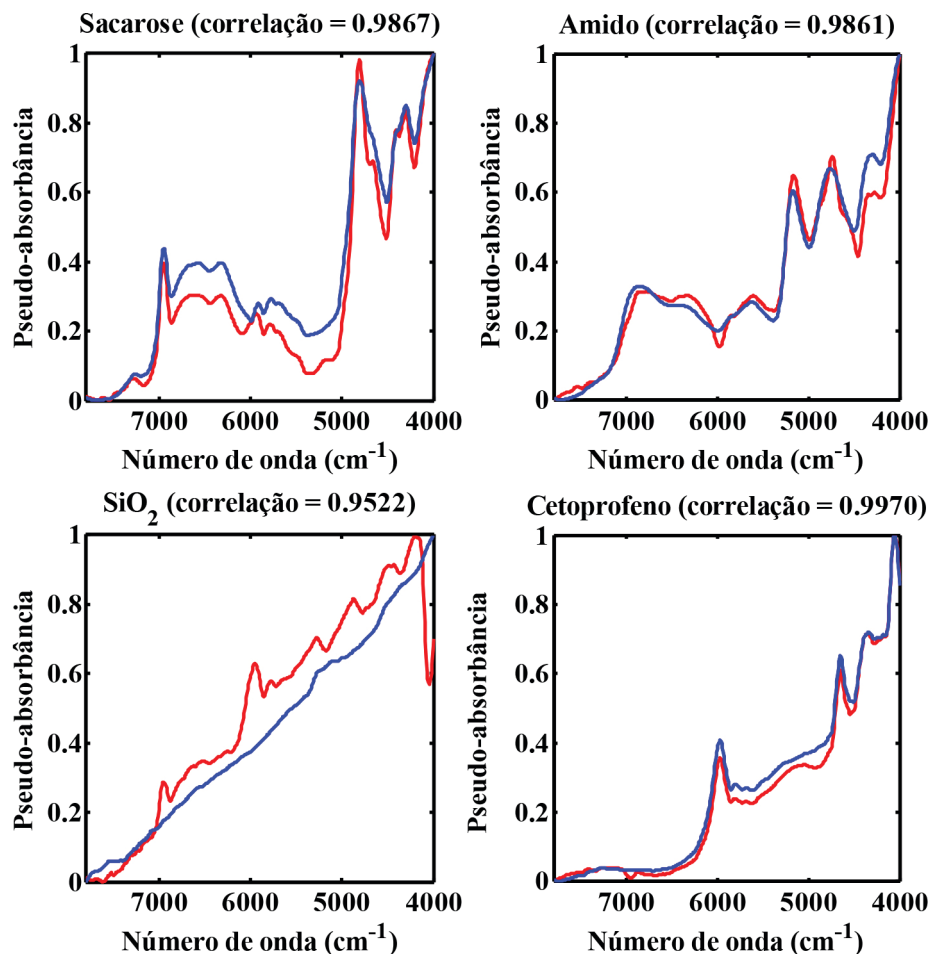


Figura 6. Comparação dos espectros das substâncias puras obtidos experimentalmente (azul) e os espectros recuperados pelo modelo MCR (em vermelho), além dos valores dos coeficientes de correlação ( $r^2$ ) para cada um dos componentes

```
>>axis square
>>h = title(DADOS(i+1).nomes);
>>set(h,'Color','black','FontSize',12.5,'FontWeight','bold')
>>set(gca,'FontWeight','bold','FontSize',12)
>>end
>>clear h k i
```

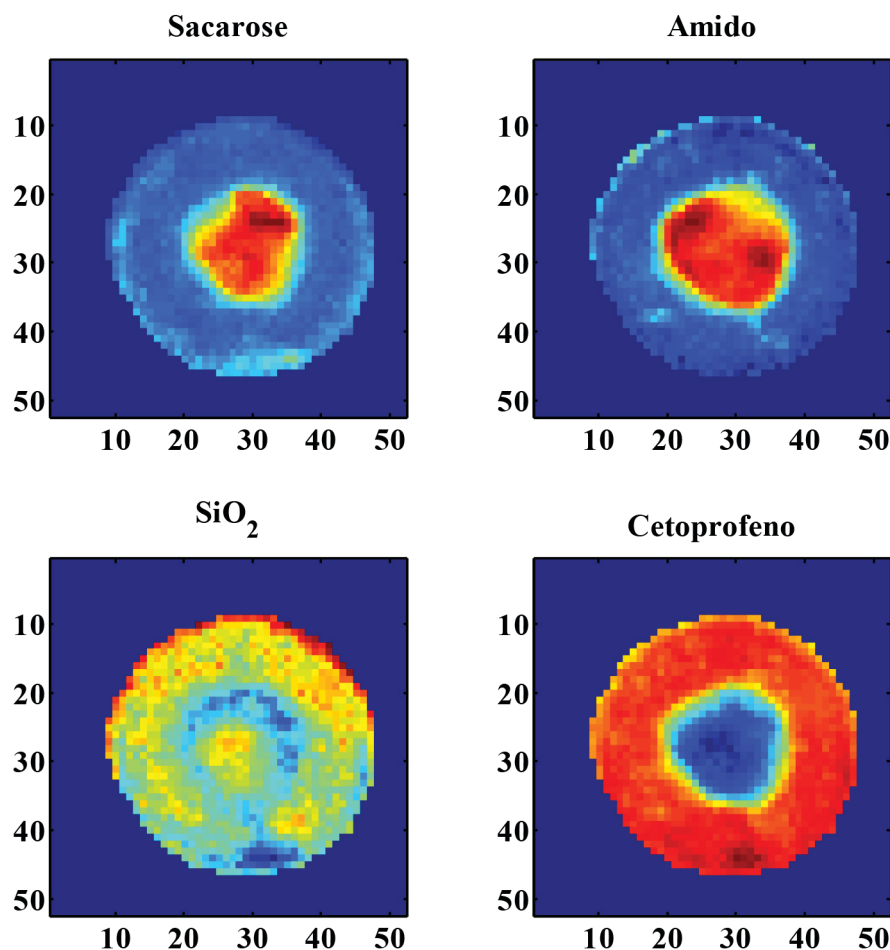
A avaliação das imagens geradas para a distribuição das espécies recuperadas pelo modelo deixa evidente que existem regiões bastante distintas no *pellet*. Na área central da amostra, há um núcleo inerte contendo sacarose e amido. O uso de um *pellet* inerte como partida para a deposição dos compostos de interesse é uma prática usual na indústria farmacêutica.<sup>41,55,56</sup> Imediatamente após esse núcleo inerte, há uma camada composta basicamente pelo API cetoprofeno, o que está indicado na imagem como uma região com coloração vermelha intensa. Por fim, a última espécie estudada é o dióxido de silício, que aparentemente está distribuído mais externamente na amostra, talvez como uma camada de revestimento ou proteção. Porém, como as condições de obtenção de sinal instrumental para esse composto não é tão ideal quanto para os demais, há de se ter certo cuidado com os resultados para esse componente. Essas informações não poderiam ser obtidas em análises por NIR de forma convencional, em que apenas as informações químicas estariam disponíveis, mas não sua distribuição espacial. Técnicas como cromatografia líquida poderiam identificar e quantificar esses compostos, mas de forma global na amostra. Isso indica o quão útil o emprego da técnica de NIR-CI pode ser aos pesquisadores para casos semelhantes.

## CONCLUSÃO

No estudo apresentado, foi proposto um experimento didático que permite o ensino da quimiometria para além dos métodos mais usualmente abordados, como PCA e PLS. O uso de espectroscopia NIR e imagem hiperespectral com MCR-ALS, pelas suas características, pode ser aplicado em disciplinas de graduação, pós-graduação e também para profissionais que não possuam experiência com a técnica, expandindo o uso da espectroscopia no infravermelho próximo. Como as etapas de preparo de amostras e de aquisição dos dados instrumentais são bastante simples e rápidas, isso proporciona ao docente a oportunidade de se dedicar à discussão dos conceitos envolvidos em todo o experimento. Nesse momento, pode-se ressaltar as vantagens do uso da espectroscopia de imagem em relação às análises convencionais por NIR, bem como descrever o passo-a-passo da técnica de MCR-ALS, para que os alunos entendam perfeitamente seu funcionamento. A presença de todos os comandos e procedimentos descritos de forma detalhada e com o uso de rotinas de acesso livre na internet facilitam a aplicação do experimento. Para a avaliação do aprendizado dos alunos, além do relatório que é normalmente exigido, indica-se a eventual aplicação de um formulário online para inferir o real nível de entendimento dos alunos sobre o tema proposto.

Uma vez compreendido o funcionamento do experimento proposto, a técnica pode ser aplicada a comprimidos que possuam camadas distintas ou preparações farmacêuticas em cremes ou suspensões, entre tantas outras classes distintas de amostras. A única condição





**Figura 7.** Imagens da distribuição química das espécies na amostra de *pellet* para os 4 componentes estudados (as cores quentes, vermelho, indicam a presença do componente e as frias, azul, a ausência)

é que essas amostras possuam regiões com características distintas, que possam ser discriminadas pelo modelo.

#### MATERIAL SUPLEMENTAR

No material suplementar a esse artigo, que está disponível para download livre, na forma de arquivo PDF, em <http://quimicanova.sbq.org.br>, são apresentadas figuras dos sinais dos excipientes amido, sacarose e dióxido de silício, além do princípio ativo cetoprofeno. A Figura 1S mostra os espectros como obtidos pelo equipamento, em reflectância e sem nenhum tratamento prévio. Esses espectros convertidos em valores de pseudoabsorbância são vistos na Figura 2S e, após serem normalizados e alisados, na Figura 3S. A etapa de seleção dos pixels referentes ao *pellet* é detalhada nas Figuras 4S e 5S. A Figura 4S apresenta como é feita a escolha dos valores dos escores da PCA, para a seleção da região de interesse, enquanto na Figura 5S, a região efetivamente selecionada em comparação com a imagem original. Além disso, todos os passos realizados no estudo estão reunidos no Quadro 1S, o que permite que seja copiado e colado no editor do programa MATLAB® para que possa ser facilmente utilizado.

#### AGRADECIMENTOS

Os autores agradecem às agências de fomento Capes e CNPq pelo auxílio financeiro e ao técnico do Laboratório de Quimiometria em Química Analítica (LAQQA), H. D. Machado, pelo apoio às atividades do projeto.

#### REFERÊNCIAS

1. Ferreira, M. M. C.; *Quimiometria - Conceitos, Métodos e Aplicações*, 1ª ed., Editora da Unicamp: Campinas, 2015.
2. de Souza, A. M.; Poppi, R. J.; *Quim. Nova* **2012**, *35*, 223.
3. de Souza, A. M.; Breikreitz, M. C.; Filgueiras, P. R.; Rohwedder, J. J. R.; Poppi, R. J.; *Quim. Nova* **2013**, *36*, 1057.
4. Breikreitz, M. C.; de Souza, A. M.; Poppi, R. J.; *Quim. Nova* **2014**, *37*, 564.
5. Fernandez Pierna, J. A.; Baeten, V.; Dardenne, P.; Dubois, J.; Lewis, E. N.; Burger, J. Em *Comprehensive Chemometrics* Elsevier: Amsterdam, 2009.
6. Geladi, P.; Burger, J.; Lestander, T.; *Chemom. Intell. Lab. Syst.* **2004**, *72*, 209.
7. Amigo, J. M.; Babamoradi, H.; Elcoroaristizabal, S.; *Anal. Chim. Acta* **2015**, *896*, 34.
8. Chiriac, A. P.; Nita, L. E.; Tartau, L.; Neamtu, I.; Tudorachi, N.; Diaconu, A.; *Rev. Roum. Chim.* **2016**, *61*, 345.
9. Dubois, J.; Neil Lewis, E.; Fry, F. S.; Calvey, E. M.; *Food Microbiol.* **2005**, *22*, 577.
10. Kamruzzaman, M.; ElMasry, G.; Sun, D.-W.; Allen, P.; *Innov. Food Sci. Emerg. Technol.* **2012**, *16*, 218.
11. Garrido-Novell, C.; Garrido-Varo, A.; Pérez-Marín, D.; Guerrero-Ginel, J. E.; Kim, M.; *J. Food Eng.* **2015**, *153*, 117.
12. Smith, J. P.; Smith, F. C.; Krull-Davatzes, A. E.; Simonson, B. M.; Glass, B. P.; Booksh, K. S.; *Chem. Data Collect.* **2017**, *9–10*, 35.
13. Araya, J. A.; Carneiro, R. L.; Arévalo, C.; Freer, J.; Castillo, R. del P.; *Microchem. J.* **2017**, *134*, 164.

14. Treado, P. J.; Priore, R. J.; Nelson, M. P. *Handbook of Vibrational Spectroscopy* John Wiley & Sons, Ltd: Chichester, UK, 2006.
15. Workman, J. J.; Weyer, L.; *Practical Guide to Interpretative Near-Infrared Spectroscopy* Taylor & Francis: Boca Raton, 2008.
16. Pasquini, C.; *J. Braz. Chem. Soc.* **2003**, *14*, 198.
17. Lopes, M. B.; Wolff, J.-C.; Bioucas-Dias, J. M.; Figueiredo, M. A. T.; *Anal. Chim. Acta* **2009**, *641*, 46.
18. Ravn, C.; Skibsted, E.; Bro, R.; *J. Pharm. Biomed. Anal.* **2008**, *48*, 554.
19. Amigo, J. M.; *Anal. Bioanal. Chem.* **2010**, *398*, 93.
20. Zhou, L.; Xu, M.; Wu, Z.; Shi, X.; Qiao, Y.; *Drug Test. Anal.* **2016**, *8*, 72.
21. Khorasani, M.; Amigo, J. M.; Sun, C. C.; Bertelsen, P.; Rantanen, J.; *Eur. J. Pharm. Biopharm.* **2015**, *93*, 293.
22. Yekpe, K.; Abatzoglou, N.; Bataille, B.; Gosselin, R.; Sharkawi, T.; Simard, J.; Cournoyer, A.; *Int. J. Pharm.* **2015**, *486*, 242.
23. Koide, T.; Yamamoto, Y.; Fukami, T.; Katori, N.; Okuda, H.; Hiyama, Y.; *Chem. Pharm. Bull. (Tokyo)* **2015**, *63*, 663.
24. Khorasani, M.; Amigo, J. M.; Bertelsen, P.; Sun, C. C.; Rantanen, J.; *Powder Technol.* **2015**, *300*, 120.
25. Franch-Lage, F.; Amigo, J. M.; Skibsted, E.; MasPOCH, S.; Coello, J.; *Int. J. Pharm.* **2011**, *411*, 27.
26. Sabin, G. P.; de Carvalho Rocha, W. F.; Poppi, R. J.; *Microchem. J.* **2011**, *99*, 542.
27. Cruz, J.; Blanco, M.; *J. Pharm. Biomed. Anal.* **2011**, *56*, 408.
28. Dukic-Ott, A.; Thommes, M.; Remon, J. P.; Kleinebudde, P.; Vervaet, C.; *Eur. J. Pharm. Biopharm.* **2009**, *71*, 38.
29. Leuner, C.; *Eur. J. Pharm. Biopharm.* **2000**, *50*, 47.
30. Juan, A. De; Tauler, R.; Dyson, R.; Marcolli, C.; Rault, M.; Maeder, M.; *TrAC -- Trends Anal. Chem.* **2004**, *23*, 70.
31. Santos, H. M. M.; Veiga, F. J. B.; de Pina, M. E. T.; de Sousa, J. J. M. S.; *Rev. Bras. Ciências Farm.* **2004**, *40*, 455.
32. Santos, H. M. M.; Veiga, F. J. B.; de Pina, E. M. S. T.; de Sousa, J. J. M. S.; *Rev. Bras. Ciências Farm.* **2006**, *42*, 309.
33. Geladi, P.; Sethson, B.; Nyström, J.; Lillhonga, T.; Lestander, T.; Burger, J.; *Spectrochim. Acta, Part B* **2004**, *59*, 1347.
34. Tauler, R.; *Chemom. Intell. Lab. Syst.* **1995**, *30*, 133.
35. Jaumot, J.; Gargallo, R.; de Juan, A.; Tauler, R.; *Chemom. Intell. Lab. Syst.* **2005**, *76*, 101.
36. Xu, C. J.; Gourvéneq, S.; Liang, Y. Z.; Massart, D. L.; *Anal. Chim. Acta* **2006**, *575*, 1.
37. Março, P. H.; Valderrama, P.; Alexandrino, G. L.; Poppi, R. J.; Tauler, R.; *Quim. Nova* **2014**, *37*, 1525.
38. de Juan, A.; Jaumot, J.; Tauler, R.; *Anal. Methods* **2014**, *6*, 4964.
39. Jaumot, J.; de Juan, A.; Tauler, R.; *Chemom. Intell. Lab. Syst.* **2015**, *140*, 1.
40. de Juan, A.; Tauler, R.; *Anal. Chim. Acta* **2003**, *500*, 195.
41. Rowe, R. C.; Sheskey, P. J.; Quinn, M. E. *Handbook of Pharmaceutical Excipients*, 6<sup>a</sup> ed., Pharmaceutical Press: London, 2009.
42. Piqueras, S.; Burger, J.; Tauler, R.; de Juan, A.; *Chemom. Intell. Lab. Syst.* **2012**, *117*, 169.
43. Terra, L. A.; Poppi, R. J.; *Chemom. Intell. Lab. Syst.* **2014**, *130*, 91.
44. Monteyne, T.; Adriaensens, P.; Brouckaert, D.; Remon, J.-P.; Vervaet, C.; De Beer, T.; *Int. J. Pharm.* **2016**, *512*, 158.
45. Grangeiro Junior, S.; de Moura França, L.; Pimentel, M. F.; Albuquerque, M. M.; de Santana, D. P.; Santana, A. K. M.; Souza, J. A. L.; Simões, S. S.; *Microchem. J.* **2015**, *118*, 252.
46. Rinnan, Å.; Berg, F. van den; Engelsen, S. B.; *TrAC -- Trends Anal. Chem.* **2009**, *28*, 1201.
47. Barnes, R. J.; Dhanoa, M. S.; Lister, S. J.; *Appl. Spectrosc.* **1989**, *43*, 772.
48. Savitzky, A.; Golay, M. J. E.; *Anal. Chem.* **1964**, *36*, 1627.
49. <https://chem.nlm.nih.gov/chemidplus/sid/0022071154>, acessada em Julho 2017.
50. <https://www.hypertools.org/>, acessada em Julho 2017.
51. <http://www.mcrals.info/>, acessada em Julho 2017.
52. <https://cran.r-project.org/index.html>, acessada em Julho 2017.
53. <https://cran.r-project.org/web/views/ChemPhys.html>, acessada em Julho 2017.
54. Kubelka, P.; Munk, F.; *Z. Tech. Phys.* **1931**, *12*, 593.
55. Avalle, P.; Pollitt, M. J.; Bradley, K.; Cooper, B.; Pearce, G.; Djemai, A.; Fitzpatrick, S.; Capece, M.; Barrows, J.; Davé, R. N.; *Eur. J. Pharm. Biopharm.* **2014**, *87*, 244.
56. Sinchaipanid, N.; Chitropas, P.; Mitrejev, A.; *Pharm. Dev. Technol.* **2004**, *9*, 163.