

## EXPERIMENTO DIDÁTICO DE QUIMIOMETRIA EMPREGANDO IMAGENS DIGITAIS OBTIDAS POR CELULAR PARA DETERMINAR ADULTERAÇÃO DE AZEITE DE OLIVA COM ÓLEO DE SOJA: UM TUTORIAL, PARTE VI

Carlos Alberto Rios<sup>a</sup>, André Marcelo de Souza<sup>b</sup>, Ronei Jesus Poppi<sup>a,†</sup>, Márcia Cristina Breitreitz<sup>a,\*</sup>

<sup>a</sup>Instituto de Química, Universidade Estadual de Campinas, 13084-971 Campinas – SP, Brasil

<sup>b</sup>Embrapa Solos, Empresa Brasileira de Pesquisa Agropecuária, 22460-000 Rio de Janeiro – RJ, Brasil

Recebido em 15/03/2022; aceito em 20/07/2022; publicado na web em 13/09/2022

TEACHING EXPERIMENT IN CHEMOMETRICS USING DIGITAL IMAGES OBTAINED BY MOBILE PHONE TO DETERMINE ADULTERATION OF OLIVE OIL WITH SOYBEAN OIL: A TUTORIAL, PART VI. The aim of this manuscript was to show the concepts involving data treatment of RGB based image analysis by univariate and multivariate approaches – the latter using Partial Least Squares (PLS) - using as a practical example application the determination of extravirgin olive oil adulteration. Digital images were collected using a device made of LEDs, batteries and a smartphone and results are shown in a tutorial format using Matlab computing environment. The experiment can serve as an example for teaching the subject for undergraduate and graduate students.

Keywords: teaching experiment; digital image; partial least squares; extra-virgin olive oil; adulteration.

### INTRODUÇÃO

Os recentes avanços instrumentais e computacionais das últimas décadas impactaram na Química Analítica de tal maneira que o resultado foi um aumento significativo na complexidade e na quantidade de dados gerados em uma única medida. Em função da necessidade de extrair informações relevantes deste significativo volume de dados, de maneira rápida e eficiente, a Quimiometria têm sido bastante difundida na universidade e na indústria.<sup>1</sup>

A Quimiometria pode ser definida como sendo um conjunto de métodos matemáticos, estatísticos e computacionais para extrair, interpretar e prever os dados de interesse químico de maneira rápida e eficiente.<sup>1</sup> Dentre suas diversas subáreas, pode-se destacar o planejamento de experimentos, o reconhecimento de padrões e a calibração multivariada. Desde o início da década de 2010, uma sequência de tutoriais vem sendo publicados pela Química Nova com a intenção de contribuir na capacitação de profissionais dos mais variados ramos da cadeia produtiva, incluindo estudantes de graduação e pós-graduação.

No primeiro tutorial, foi usada a análise de componentes principais (*Principal Component Analysis*, PCA) para diferenciar óleos comerciais por espectroscopia no infravermelho médio<sup>1</sup> (*Mid Infrared Spectroscopy*, MIRS). No segundo, foi usada a regressão por mínimos quadrados parciais (PLS, do inglês *Partial Least Squares*) para determinar a concentração de paracetamol em comprimidos usando espectroscopia no infravermelho próximo (*Near Infrared Spectroscopy*, NIR).<sup>2</sup> Para o terceiro estudo, foi feito a determinação de Ferro (III) com *o*-fenantrolina usando espectroscopia no visível em conjunto com planejamento experimental (*design of experiment*, DoE).<sup>3</sup>

A imagem hiperespectral foi explorada no quarto tutorial por meio de um mapeamento de uma formulação farmacêutica sólida para uso oral (*pellets*) empregando espectroscopia de imagem na região do NIR, em conjunto com métodos de Resolução de Curvas (*Multivariate Curve Resolution*, MCR).<sup>4</sup> No quinto e último tutorial publicado, foi desenvolvido um método de classificação de óleos vegetais

comestíveis usando espectroscopia na região do infravermelho médio combinado com análise discriminante por mínimos quadrados parciais (*Partial Least Squares Discriminant Analysis*, PLS-DA).<sup>5</sup> Os três primeiros tutoriais mencionados têm sido utilizados em disciplinas de Química Analítica Instrumental ministradas no Instituto de Química da UNICAMP.

A proposta deste tutorial é revisitar o conceito de imagens digitais para explorar o seu uso como metodologia analítica para estimar o percentual de adulterantes em amostras de alimentos de uma maneira didática. Por definição, uma imagem digital é composta pela união de vários *pixels* (do inglês, *picture element*), sendo a quantidade dos mesmos proporcional à resolução da imagem.<sup>6</sup>

Os métodos analíticos disponíveis para quantificar espécies químicas em amostras coloridas, como o colorimétrico, fotométrico ou espectrofotométrico são onerosos se comparados com a metodologia baseada em imagens digitais.<sup>7,8</sup> Nesse sentido, o uso de dispositivos comerciais simples, como câmeras digitais, *webcams*, *scanners* e *smartphones* desempenham um papel importante como detectores genéricos visando obter o máximo de informações relevantes de maneira rápida e eficiente.<sup>9-11</sup> A grande parte dos dispositivos citados usam o sistema de cores RGB, composto pelas cores vermelho (R, *Red*), verde (G, *Green*) e azul (B, *Blue*). Além do baixo custo, o uso de imagens digitais a partir de celulares se destaca pela versatilidade, simplicidade, rapidez e pela possibilidade de medidas *in situ*.

O interesse pelo uso analítico de imagens digitais é crescente, por isso novas aplicações têm sido desenvolvidas e publicadas em diversas áreas de aplicação. Destacam-se aplicações na área de controle de qualidade de alimentos, como por exemplo, o monitoramento do processo de produção de pães,<sup>12</sup> qualidade de presunto,<sup>13</sup> análise de vinhos,<sup>14</sup> quantificação de lactobacilos em leite fermentado,<sup>15</sup> monitoramento da fermentação de chás,<sup>16</sup> detecção de doenças em plantas,<sup>17</sup> determinação do tamanho de grãos de arroz,<sup>18</sup> determinação da cor em cervejas<sup>19</sup> e diferenciação de mel e própolis por origem geográfica.<sup>20,21</sup> Além desses, é possível encontrar estudos na área biológica,<sup>22,23</sup> forense,<sup>24</sup> combustíveis,<sup>25,26</sup> documentoscopia<sup>27</sup> e arqueologia.<sup>28</sup>

No setor alimentício, produtos de relevância econômica são alvo de fraudes que podem adicionar, substituir ou adulterar um

\*e-mail: marciacb@unicamp.br

†*In memoriam*

ingrediente presente no alimento.<sup>19</sup> Em outros casos, podem ser feitas modificações na embalagem, no rótulo ou até em certificados ou declarações que atestam a qualidade do produto visando um ganho econômico e prejuízo para a vítima.<sup>29,30</sup> Além dos aspectos econômicos, essa prática pode representar riscos reais para a saúde do consumidor que não possui as ferramentas analíticas específicas para identificar e distinguir a presença de adulterantes.<sup>27,28</sup>

Um alvo frequente de falsificações é o azeite de oliva, devido ao seu alto valor econômico e nutricional.<sup>29</sup> No estudo feito por Santos e colaboradores,<sup>31</sup> foi utilizado um espectrômetro portátil na região do NIR associado a ferramentas quimiométricas (PCA, PLS). O objetivo era classificar e quantificar a presença de óleo de soja em amostras de azeite de oliva extra-virgem e de óleo composto que foram apreendidas pela polícia estadual no estado do Espírito Santo (ES), Brasil. O modelo empregando o método de regressão em mínimos quadrados parciais (PLS) foi construído a partir de 19 misturas binárias azeite-óleo, em triplicata, com um intervalo de concentração de óleo vegetal de 0 até 100%. O modelo foi gerado com 2 variáveis latentes (*latent variables*, LV) e apresentou valores para o erro quadrático médio de previsão (*root mean square error of prediction*, RMSEP) de 4,8% e um coeficiente de determinação ( $R^2$ ) de 0,975.

Já no estudo de Duraipandian e colaboradores,<sup>32</sup> foi empregada a espectroscopia Raman em casos de adulteração de azeite de oliva extra-virgem com óleos comestíveis mais baratos, como milho, soja e colza. O método de regressão adotado foi o PLS e o modelo foi construído usando 33 espectros de misturas binárias (azeite + colza), 48 espectros de misturas ternárias (azeite + colza + milho) e 133 espectros de misturas quaternárias (azeite + colza + milho + soja), totalizando 214 espectros. O modelo proposto apresentou valores para RMSEP de 5,1% e um coeficiente de determinação ( $R^2$ ) de 0,97. As misturas foram preparadas com diferentes concentrações de azeite, dentro de um intervalo de 10 até 100%.

Apesar do uso de imagens digitais já ter sido empregada em uma ampla gama de situações, ainda existe demanda por pesquisas de aplicações no monitoramento de óleos vegetais, em especial, em casos de adulteração de azeite de oliva extra-virgem. O presente estudo tem como objetivo geral contribuir para o desenvolvimento de metodologias empregando imagens digitais obtidas a partir de um *smartphone* e calibração multivariada por PLS, visando a determinação do teor de azeite de oliva extra-virgem em amostras adulteradas com óleo de soja. Com a intenção de comparar o desempenho do método multivariado baseado em PLS, também foram construídos modelos univariados empregando os canais R, G e B, que são mais simples e prontamente aplicáveis à proposta de ensino.

## PARTE EXPERIMENTAL

### Preparo de amostra

As amostras de óleo de soja e azeite de oliva extra-virgem foram adquiridas em supermercados na cidade de Campinas – SP. A Figura 1S, disponível no Material Suplementar, apresenta os frascos das amostras que foram utilizadas. No total foram adquiridas 6 marcas de azeite de oliva extra-virgem, 5 marcas de óleo de soja e 1 amostra de óleo composto (soja + azeite, 12 %).

Para construir a curva analítica, foi preparada uma única mistura com volumes iguais de todas as marcas de óleo de soja. O mesmo procedimento foi adotado para as amostras de azeite de oliva extra-virgem. A partir disso, foram preparadas misturas entre óleo de soja e azeite de oliva em diferentes proporções, de acordo com a Tabela 1.

O experimento foi realizado em triplicata e o volume total de cada mistura foi de 30 mL. Os volumes foram medidos usando uma proveta e as imagens digitais foram adquiridas, conforme será descrito a seguir.

**Tabela 1.** Proporção de óleo e azeite presente nas misturas preparadas

| Mistura | Óleo de soja (mL) | Azeite (mL) | Total (mL) | Azeite (v/v %) |
|---------|-------------------|-------------|------------|----------------|
| 1       | 30                | 0           | 30         | 0              |
| 2       | 27                | 3           | 30         | 10             |
| 3       | 24                | 6           | 30         | 20             |
| 4       | 21                | 9           | 30         | 30             |
| 5       | 18                | 12          | 30         | 40             |
| 6       | 15                | 15          | 30         | 50             |
| 7       | 12                | 18          | 30         | 60             |
| 8       | 9                 | 21          | 30         | 70             |
| 9       | 6                 | 24          | 30         | 80             |
| 10      | 3                 | 27          | 30         | 90             |
| 11      | 0                 | 30          | 30         | 100            |

### Aquisição das imagens digitais

Para a aquisição das imagens, foi construído um sistema de baixo custo a partir de uma caixa de isopor, 6 pilhas de 1,5 V e 2 LEDs (do inglês, *light emitting diode*) de alto brilho na cor branca com uma tensão de 3,5 V. A parte superior do aparato foi recortada para encaixar um *smartphone* e assim foram obtidas as imagens digitais. A Figura 2S, disponível no Material Suplementar, apresenta o dispositivo usado para a coleta das imagens.

As imagens foram obtidas empregando um *smartphone Samsung Galaxy S9* modelo SM-G9600/1DL com estabilização óptica de imagem e uma lente com duas aberturas focais, uma maior ( $f/1,5$ ) e outra menor ( $f/2,4$ ), distância focal de 26 mm, balanço de branco automático, ISO 80, tempo de exposição 1/100 s e câmera traseira de 12 MP. Todas as imagens foram obtidas com  $4032 \times 3024$  pixels (12 MB) no formato “.jpg”. A distância entre a câmera e as amostras foi de 5 cm.

Para cada mistura preparada de acordo com a Tabela 1, foram coletadas 5 imagens, totalizando 55 por replicata. No total, foram obtidas 165 imagens digitais para a construção dos modelos de calibração. Além disso, foram coletadas 5 imagens da amostra de óleo composto e essas foram usadas na etapa de previsão.

As imagens digitais usadas neste tutorial, no formato “.jpg”, estão prontamente disponíveis aos interessados por meio de contato com a autora M. C. Breitreitz (marciacb@unicamp.br).

### Softwares e algoritmos

Para realizar os cálculos descritos neste tutorial, foi usado um pacote gratuito para aplicação do algoritmo PLS que centra matrizes **X** e **y** na média de forma automática. As rotinas para o algoritmo foram disponibilizadas pelo grupo de Quimiometria da Universidade Livre de Bruxelas (Chemometrics in Analytical Chemistry, Chemo AC) disponíveis no *website*: <http://www.vub.ac.be/fabi/publiek/index.html>.

As rotinas foram executadas em ambiente Matlab 8.3. Para o tratamento das imagens e construção das curvas univariadas, foi usado o pacote *Image Processing Toolbox* do MATLAB (Mathworks, Natick, Massachusetts, EUA, v. 9.0).

### Curva analítica: metodologia multivariada

Foram coletadas 5 imagens para cada mistura preparada, sendo que para cada imagem foi gerado um histograma na forma de uma matriz com dimensões de  $1 \times 768$ . Em seguida, foi calculada a média desses histogramas para cada replicata.

A partir dos histogramas médios, o modelo PLS foi construído usando, no conjunto de calibração, as misturas com as seguintes porcentagens de azeite: 0, 10, 40, 60, 90 e 100%. Para o conjunto de validação, foram usados os histogramas das misturas com as seguintes porcentagens: 20, 50 e 80%. Na etapa de previsão, foram usados os histogramas médios da amostra de óleo composto (12%) e das misturas com o percentual de azeite de 30 e 70%. O pré-processamento adotado tanto na matriz  $X$  quanto no vetor  $y$  foi centrar os dados na média.

### Curva analítica: metodologia univariada

Abordando o problema de maneira univariada, para cada imagem, foi gerado um vetor linha com a intensidade da cor nos canais R, G e B. Em seguida, foi calculado a média desses valores para cada replicata. De posse dos valores médios, foi construída uma curva analítica para cada canal do sistema RGB. O percentual de azeite usado na calibração, validação e previsão foram os mesmos da metodologia multivariada.

## RESULTADOS E DISCUSSÃO

### Tratamento das imagens digitais

As imagens obtidas foram transferidas para o computador e passaram por etapas de pré-tratamento antes de construir os modelos de calibração univariado e multivariado. As linhas de comando usadas nestas etapas estão descritas abaixo de forma comentada e disponíveis no material suplementar no formato de uma rotina ('tratamentoimagemRGB.m', rotina 1).

#### Importar a imagem .jpg e depois abri-la no Matlab

(1) Na etapa inicial, as imagens no formato ".jpg" foram importadas para o *workspace* do Matlab no formato de matrizes de dados.

```
>> X = imread('nome da imagem.jpg');
>> imshow(X);
```

#### Fazer um recorte na imagem (Region of Interest, ROI)

(2) Nessa etapa, é feito um recorte na imagem original para selecionar uma região central que aparece somente a mistura óleo-azeite (Region of Interest, ROI). Neste artigo, a imagem foi cortada com dimensões de 501x501 pixels.

```
>> XX = X (B2:B1, A2:A1, :); %considerando a imagem formada por A x B pixels
figure (1)
>> subplot (1,2,1)
>> imshow (X)
>> subplot (1,2,2)
>> imshow (XX)
```

#### Gerar uma figura com as imagens da região de interesse (Region of Interest, ROI)

As imagens recortadas (501x501 pixels) para cada replicata estão disponíveis no material suplementar em um arquivo na extensão ".mat". ("matriz\_imagem\_replicata\_1.mat", "matriz\_imagem\_replicata\_2.mat" e "matriz\_imagem\_replicata\_3.mat").

```
>> load('matriz_imagem_replicata_1.mat'), será carregado uma série de matrizes no workspace. A0R1 = A0, percentual (0%) de azeite de oliva (A). R1, imagem 1 das 5 que foram coletadas para o percentual 0%. A30R3 = A30, percentual (30%) de azeite de oliva (A). R3, imagem 3 das 5 que foram coletadas para o percentual 30%.
```

Para gerar uma figura com as 5 imagens coletadas para o percentual 0%, é necessário usar as seguintes linhas de comando:

```
>> subplot(1,5,1)
>> imshow(A0R1)
>> subplot(1,5,2)
>> imshow(A0R2)
>> subplot(1,5,3)
>> imshow(A0R3)
>> subplot(1,5,4)
>> imshow(A0R4)
>> subplot(1,5,5)
>> imshow(A0R5)
```

Para as demais porcentagens, basta usar as mesmas linhas de comando descritas acima e substituir o "0" pelo valor desejado (10, 20, 30, etc). Para gerar imagens das outras replicatas, basta limpar o *workspace* usando a função "clear" e refazer todos os passos descritos neste tópico usando os outros dois arquivos ".mat".

#### Separar a imagem nas matrizes individuais dos canais R, G e B

(3) Em seguida, a matriz de dados gerada foi separada em três outras matrizes com as mesmas dimensões, representando os canais do sistema RGB. A Figura 3S, disponível no Material Suplementar, apresenta um esquema ilustrativo da metodologia descrita.

```
>> RGB1red = XX (:, :, 1);
>> RGB1green = XX (:, :, 2);
>> RGB1blue = XX (:, :, 3);
```

#### Obter os histogramas médios para cada imagem cortada nos canais R, G e B

Nessa etapa, foram criados dois vetores linha (1x256) para cada canal, sendo que o primeiro contém os valores crescentes de intensidade que variam numa escala de 0 a 255. O segundo, possui os valores das frequências absolutas. A Figura 4S, disponível no Material Suplementar, apresenta um esquema ilustrativo da metodologia usada.

```
>> [rgby1red,x1r] = imhist(RGB1red);
>> [rgby1green,x1g] = imhist(RGB1green);
>> [rgby1blue,x1b] = imhist(RGB1blue);
```

(4) Os vetores com as frequências absolutas foram agrupados em uma nova matriz. Em seguida, foi feito o somatório obtendo a frequência total. Por fim, foi feita a divisão da frequência absoluta pela frequência total obtendo as frequências relativas, ou seja, o histograma da imagem.

```
>> rgbimagem1 = [rgby1red,rgby1green,rgby1blue];
>> somatorio = sum(rgbimagem1);
>> rgby1red = rgby1red/somatorio(:,1);
>> rgby1green = rgby1green/somatorio(:,2);
>> rgby1blue = rgby1blue/somatorio(:,3);
>> HistR1 = [rgby1red',rgby1green',rgby1blue'];
>> plot (HistR1');
```

Após gerar os histogramas, foi empregada outra metodologia na matriz obtida na etapa (2). O intuito era gerar, para cada imagem, um único valor de R, G e B. Essa abordagem foi adotada para simplificar o problema, na tentativa de usar um método univariado.

#### Obter um valor médio para cada imagem cortada nos canais R, G e B

(5) A matriz gerada na etapa (2) foi reorganizada em uma nova matriz (XX1) com três colunas e com o número de linhas igual ao número de colunas multiplicado pelo número de linhas da

matriz original.

```
>>XX1=reshape (XX,size(XX,1)*size(XX,2), size(XX,3));
(6) Em seguida, foi feita a média das colunas da matriz "XX1" para obter os valores médios de R, G e B. Isso foi feito para cada imagem coletada.
>> valormedioRGB=mean(XX1);
```

Após coletar e realizar o pré-tratamento, foi feito um comparativo entre as imagens. O intuito foi observar se houve uma grande variação na cor e na luminosidade entre uma foto e outra. As imagens geradas após a etapa (2) estão apresentadas nas Figuras 5S e 6S, disponíveis no Material Suplementar.

A partir das figuras, foi possível observar que não houve uma grande variação de cor entre a mesma marca. No entanto, foi observada uma variação considerável entre as diferentes marcas de óleo de soja. Em relação as amostras de azeite de oliva, também foi observada uma diferença de tonalidade que pode ter relação com o grau de amadurecimento da azeitona usada na fabricação do azeite, que podem estar maduras (amarelo ouro) ou verdes (verde intenso).

Após analisar as imagens, foi observado que existe uma diferença na cor entre as diferentes marcas, mesmo se tratando de amostras que pertencem a mesma classe, seja óleo ou azeite. Isso mostra a importância do preparo das misturas adulteradas a partir de uma única mistura de óleo de soja com todas as marcas. O mesmo se aplica para as amostras de azeite de oliva extra-virgem.

As imagens digitais das misturas preparadas nas diferentes proporções de óleo-azeite estão apresentadas na Figura 7S, disponível no Material Suplementar. No aspecto visual, as imagens digitais obtidas estão de acordo com o esperado, pois à medida que a proporção de azeite aumenta, a coloração fica mais escura e intensa.

## Resultados da metodologia univariada

Após obter o valor médio de R, G e B para cada replicata, foram construídas as curvas analíticas. A seguir, serão apresentadas as etapas para a construção das curvas analíticas para calibração, validação e previsão do percentual de azeite de oliva extra-virgem presente nas misturas. As linhas de comando descritas abaixo estão disponíveis no material suplementar no formato de uma rotina ('regressaolinear.m', rotina 2).

*Carregar os vetores linha de cada replicata e obter os valores médios de R, G e B*

(7) Os vetores linha obtidos para cada uma das imagens foram carregados no *workspace*. Foram obtidos vetores linha para todas as replicatas das misturas apresentadas na Tabela 1. Em seguida, os vetores foram organizados de acordo com o percentual de azeite e o valor médio foi calculado para cada canal do sistema RGB.

```
>>load('artigovaloresRGB.mat');
>>x0=[valorRGB0R1;valorRGB0R2;valorRGB0R3];
>>x10=[valorRGB10R1;valorRGB10R2;valorRGB10R3];
>>x20=[valorRGB20R1;valorRGB20R2;valorRGB20R3];
>>x30=[valorRGB30R1;valorRGB30R2;valorRGB30R3];
>>x40=[valorRGB40R1;valorRGB40R2;valorRGB40R3];
>>x50=[valorRGB50R1;valorRGB50R2;valorRGB50R3];
>>x60=[valorRGB60R1;valorRGB60R2;valorRGB60R3];
>>x70=[valorRGB70R1;valorRGB70R2;valorRGB70R3];
>>x80=[valorRGB80R1;valorRGB80R2;valorRGB80R3];
>>x90=[valorRGB90R1;valorRGB90R2;valorRGB90R3];
>>x100=[valorRGB100R1;valorRGB100R2;valorRGB100R3];
>>x0mean=mean(x0);
>>x10mean=mean(x10);
```

```
>>x20mean=mean(x20);
>>x30mean=mean(x30);
>>x40mean=mean(x40);
>>x50mean=mean(x50);
>>x60mean=mean(x60);
>>x70mean=mean(x70);
>>x80mean=mean(x80);
>>x90mean=mean(x90);
>>x100mean=mean(x100);
>> xazeite12= valorRGBcomposto;
>> x12_mean = mean(xazeite12);
(8) Criar as matrizes do conjunto de calibração, validação e previsão. Além de criar os vetores y contendo o percentual de azeite.
>>retaR_cal=[x0mean(:,1);x10mean(:,1);x40mean(:,1);x60mean(:,1);x90mean(:,1);x100mean(:,1)];
>>retaR_val=[x20mean(:,1);x50mean(:,1);x80mean(:,1)];
>>retaRval_ext=[x12_mean(:,1);x30mean(:,1);x70mean(:,1)];
>>retaG_cal=[x0mean(:,2);x10mean(:,2);x40mean(:,2);x60mean(:,2);x90mean(:,2);x100mean(:,2)];
>>retaG_val=[x20mean(:,2);x50mean(:,2);x80mean(:,2)];
>>retaGval_ext=[x12_mean(:,2);x30mean(:,2);x70mean(:,2)];
>>retaB_cal=[x0mean(:,3);x10mean(:,3);x40mean(:,3);x60mean(:,3);x90mean(:,3);x100mean(:,3)];
>>retaB_val=[x20mean(:,3);x50mean(:,3);x80mean(:,3)];
>>retaBval_ext=[x12_mean(:,3);x30mean(:,3);x70mean(:,3)];
>>yADULTcal=[100;90;60;40;10;0];
>>yADULTval=[80;50;20];
>>yADULTval_ext=[88;70;30];
```

*Construir as curvas analíticas para cada canal do sistema de cor RGB*

(9) A seguir estão as linhas de comando para construir a curva referente ao canal R, conforme apresentado na Figura 1. As curvas para o canal G e B também estão apresentadas na Figura 1.

```
>>y = yADULTcal;
>>x = retaR_cal; (1)
>>p1 = polyfit(x,y,1)
>>figure(1)
>>plot(x,y,'ro')
>>hold on
>>plot(retaR_val,yADULTval,'bo') (2)
>>plot(retaRval_ext,yADULTval_ext,'go') (3)
>>plot(x,polyval(p1,x),'k-')
>>yfit = p1(1) * x + p1(2);
>>yresid = y - yfit;
>>SQresid = sum(yresid.^2);
>>SQtotal = (length(y)-1) * var(y);
>>R2 = 1 - SQresid/SQtotal
>>coefangular=p1(:,1);
>>coeflinear=p1(:,2);
>>formatSpec = '%s %s %d %s %s %s%d%s';
>>A1 = 'y';
>>A2 = '=';
>>A3 = '3';
>>A4 = 'x';
>>A5 = '+';
>>A6 = '(';
>>A7 = '5';
>>A8 = ')';
>>str = sprintf(formatSpec,A1,A2,A3,A4,A5,A6,A7,A8);
>>text(140,95,str,'FontSize',12,'FontName','Arial'); (4)
>>s1=sprintf('R^2 = %0.4f',R2);
```

```
>>text(140,85,s1,'FontSize',12,'FontName','Arial'); (5)
>>axis([134 196 -5 104]); (6)
>>legend('Calibração','Validação','Validação externa')
>>xlabel('Valores médios de intensidade (Canal R)', 'FontSize',12,'FontName','arial'); (7)
>>ylabel('Porcentagem de adulteração (%)', 'FontSize',12,'FontName','arial');
>>title ('Gráfico do canal Red (R)', 'FontSize',16,'FontName','arial') (8)
```

(10) Para construir as curvas referentes ao canal G (Green) e Blue (B), basta usar a mesma rotina descrita na etapa (9) e substituir as linhas de comando destacadas em negrito pelas seguintes linhas de comando:

#### Canal G

```
>>x = retaG_cal; (1)
>>plot(retaG_val,yADULTval,'bo') (2)
>>plot(retaGval_ext,yADULTval_ext,'go') (3)
>>text(112,95,str,'FontSize',12,'FontName','Arial'); (4)
>>text(112,90,s1,'FontSize',12,'FontName','Arial'); (5)
>>axis([108 178 -5 104]); (6)
>>xlabel('Valores médios de intensidade (Canal G)', 'FontSize',12,'FontName','arial'); (7)
>>title ('Gráfico do canal Green (G)', 'FontSize',16,'FontName','arial'); (8)
```

(11) Para construir as curvas referentes ao canal Blue (B), basta usar a mesma rotina descrita na **etapa (9)** e substituir as linhas

de comando destacadas em negrito pelas seguintes linhas de comando:

#### Canal B

```
>>x = retaB_cal; (1)
>>plot(retaB_val,yADULTval,'bo') (2)
>>plot(retaBval_ext,yADULTval_ext,'go') (3)
>>text(7,20,str,'FontSize',12,'FontName','Arial'); (4)
>>text(7,13,s1,'FontSize',12,'FontName','Arial'); (5)
>>axis([0 23 -2 105]); (6)
>>xlabel('Valores médios de intensidade (Canal G)', 'FontSize',12,'FontName','arial'); (7)
>>title ('Gráfico do canal Blue (B)', 'FontSize',16,'FontName','arial') (8)
```

A partir das curvas analíticas obtidas, é possível notar que os canais R e G apresentaram boa linearidade entre os valores médios de intensidade de cor do sistema RGB e o percentual de azeite presente nas misturas. Os coeficientes de determinação obtidos para o canal R e G foram 0,986 e 0,993, respectivamente. Por outro lado, o canal B não exibiu uma boa linearidade e apresentou um coeficiente de determinação de 0,395.

#### Resultados da metodologia multivariada

Após obter o conjunto de histogramas médios, partiu-se para a construção do modelo multivariado. O modelo PLS foi construído relacionando os histogramas médios com o percentual de azeite presente nas misturas. A seguir, serão apresentadas as etapas para o

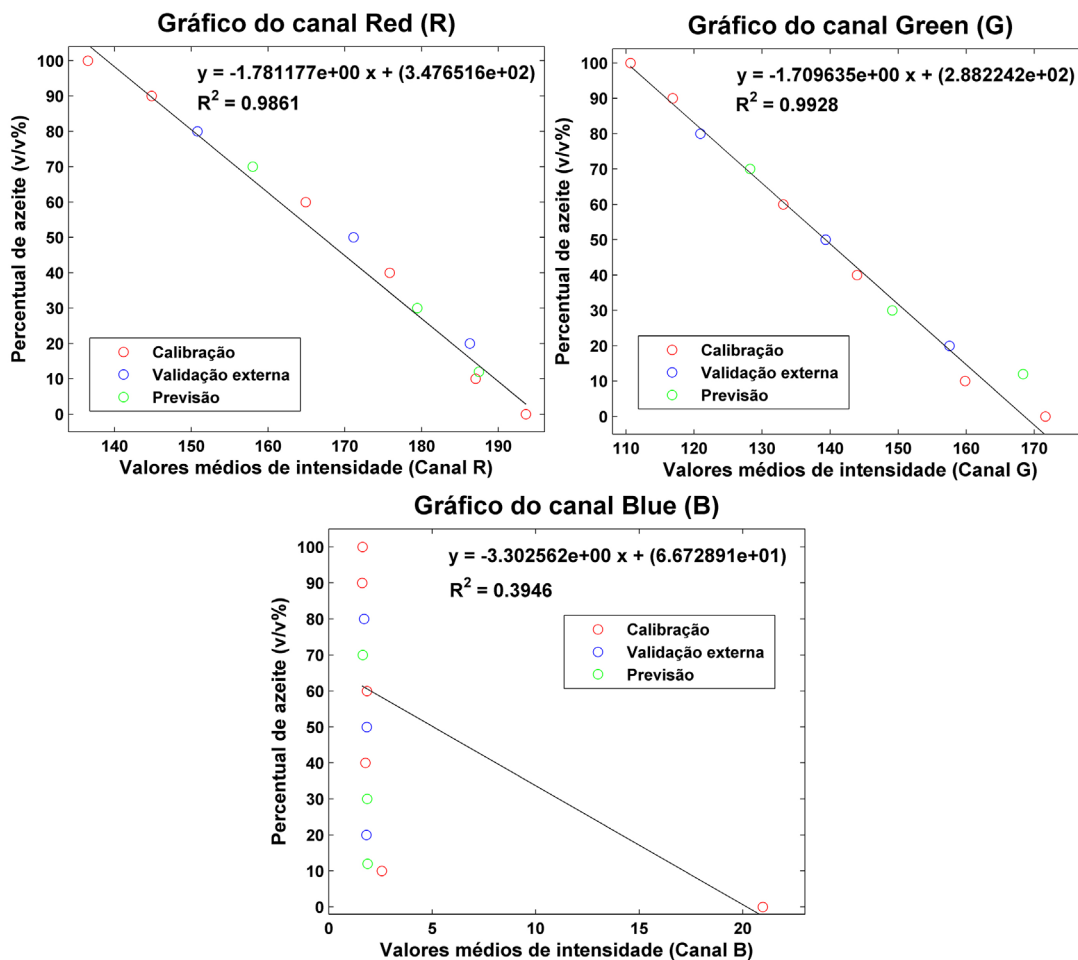


Figura 1. Gráfico das curvas analíticas usando a metodologia univariada para os canais R, G e B

desenvolvimento do modelo PLS, validação e previsão do percentual de azeite presente nas misturas adulteradas. As linhas de comando usadas nestas etapas estão disponíveis no material suplementar no formato de uma rotina ('rotinaPLS.m', rotina 3).

#### Carregar os histogramas de cada replicata e obter os histogramas médios

(12) Nesta etapa, os histogramas obtidos para cada replicata foram carregados no 'workspace'. Em seguida, os histogramas foram reorganizados de acordo com o percentual de azeite e o histograma médio foi calculado.

```
>> load('artigoHistogramasRGB.mat')
>> x0=[hist0R1;hist0R2;hist0R3];
>> x10=[hist10R1;hist10R2;hist10R3];
>> x20=[hist20R1;hist20R2;hist20R3];
>> x30=[hist30R1;hist30R2;hist30R3];
>> x40=[hist40R1;hist40R2;hist40R3];
>> x50=[hist50R1;hist50R2;hist50R3];
>> x60=[hist60R1;hist60R2;hist60R3];
>> x70=[hist70R1;hist70R2;hist70R3];
>> x80=[hist80R1;hist80R2;hist80R3];
>> x90=[hist90R1;hist90R2;hist90R3];
>> x100=[hist100R1;hist100R2;hist100R3];
>>x0mean=mean(x0);
>>x10mean=mean(x10);
>>x20mean=mean(x20);
>>x30mean=mean(x30);
>>x40mean=mean(x40);
>>x50mean=mean(x50);
>>x60mean=mean(x60);
>>x70mean=mean(x70);
>>x80mean=mean(x80);
>>x90mean=mean(x90);
>>x100mean=mean(x100);
>>xazeite12= histoleocomposto;
>>x12_mean = mean(xazeite12);
```

(13) Em seguida, as matrizes do conjunto de calibração, validação e previsão foram criadas. Além disso, foi criado os vetores y contendo o percentual de azeite.

```
>>x_cal=[x0mean;x10mean;x40mean;x60mean;x90mean;
x100mean];
>>x_val=[x20mean;x50mean;x80mean];
>>x_prev=[x12_mean;x30mean;x70mean];
>>y_cal=[0;10;40;60;90;100];
>>y_val=[20;50;80];
>>y_prev=[12;30;70];
```

#### Validação cruzada

(14) Para encontrar o número de variáveis latentes que deve ser usada no modelo, é necessário realizar o procedimento de validação cruzada, conforme descrito na linha de comando abaixo:

```
>> [rmsecv,yh,e]=PLSCV(x_cal,y_cal,0,10);
```

(15) Para criar o gráfico de RMSECV vs. Número de variáveis latentes, apresentado na Figura 2, o usuário deverá digitar o número de variáveis latentes com base no mínimo da curva gerada. De acordo com a Figura 2, o número escolhido é 2, uma vez que a partir dele, não existe variação apreciável no valor de RMSECV.

```
>>figure (1)
>>set(0,'defaultAxesFontName', 'arial')
```

```
>>axes('FontSize',24);
>>plot(rmsevcv,'k');
>>hold on
>>plot(rmsevcv,'ko')
>>ylabel('RMSECV','FontSize',30,'FontName','arial');
>>xlabel('Variáveis latentes','FontSize',30,'FontName','arial');
>>vl=input('Quantas variáveis latentes?'); Entrar com o número
de variáveis latentes e teclar o <enter>.
```

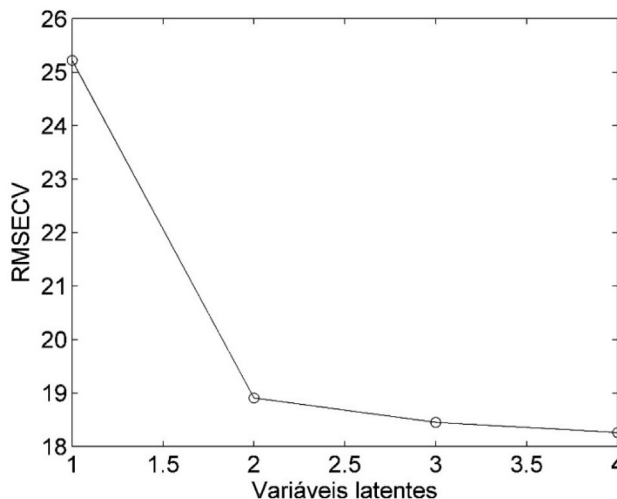


Figura 2. Gráfico do número de variáveis latentes vs o valor de RMSECV

#### Calibração

(16) Nesta etapa, o algoritmo irá fazer a previsão do percentual de azeite para as amostras de calibração e calculará os valores de RMSECV com o número de variáveis latentes indicados na etapa (15).

```
>> [ypred1multi]=PLSPRED(x_cal,y_cal,x_cal,0,vl);
>> rmsec=sqrt((sum((y_cal-ypred1).^2))/(length(y_cal)-vl-1));
```

(17) Em seguida, foi construído o gráfico de valores esperados vs valores previstos com as amostras de calibração. A Figura 8S, disponível no Material Suplementar, apresenta o gráfico gerado.

```
>>figure (2)
>>set(0,'defaultAxesFontName', 'arial')
>>axes('FontSize',24);
>>r=corrcoef(y_cal,ypred1);
>>bias=sum(y_cal-ypred1)/length(y_cal);
>>s1=sprintf('r^2c = %0.4f',r(1,2)^2);
>>s2=sprintf('rmsec = %0.4f',rmsec);
>>s3=sprintf('bias = %0.4f',bias);
>>plot(y_cal,ypred1,'ko','LineWidth',2,'MarkerSize',15,'
MarkerFaceColor',[.8 .8 .8]);lsline
>>pa1=min(y_cal);
>>pa2=max(y_cal);
>>axis([.95*pa1 1.05*pa2 .95*pa1 1.05*pa2]);
>>text(pa1-abs(pa1*0.04),1.02*pa2->>abs(pa2*0.05),s1,
'FontSize',20,'FontName','Arial');
>>text(pa1-abs(pa1*0.04),1.02*pa2->>abs(pa2*0.08),s2,
'FontSize',20,'FontName','Arial');
>>text(pa1-abs(pa1*0.04),1.02*pa2->>abs(pa2*0.11),s3,
'FontSize',20,'FontName','Arial');
>>ylabel('Valores previstos (%)','FontSize',30,'FontName',
'arial');
>>xlabel('Valores esperados (%)','FontSize',30,'FontName',
'arial');
```

**Validação Externa**

(18) O modelo de calibração construído foi validado usando um conjunto externo de amostras. O vetor “ypred2” contém os valores previstos do percentual de azeite: 20, 50 e 80%. Além disso, foi calculado o valor de RMSEV de validação.

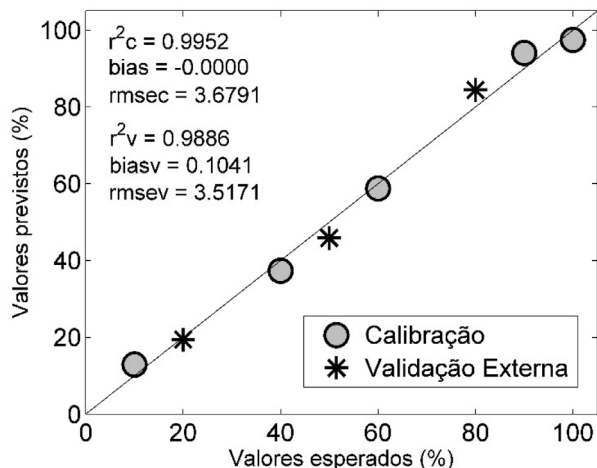
```
>>[ypred2]=plsprend(x_cal,yval,x_val,0,2);
>>rmsev=sqrt((sum((yval-ypred2).^2))/(length(yval)));
```

(19) As linhas de comando descritas abaixo foram usadas para gerar o gráfico de valores previstos vs os valores esperados para as amostras de validação. O gráfico gerado está apresentado na Figura 9S, disponível no Material Suplementar.

```
>>figure (3)
>>set(0,'defaultAxesFontName', 'arial')
>>axes('FontSize',20);
>>r=corrcoef(yval,ypred2);
>>biasv=sum(yval-ypred2)/length(yval);
>>s12=sprintf('r^2v = %0.4f',r(1,2)^2);
>>s22=sprintf('rmsev = %0.4f',rmsev);
>>s32=sprintf('biasv = %0.4f',biasv);
>>plot(yval,ypred2,'g*','LineWidth',2,'MarkerSize',14,'MarkerFaceColor',[.8 .8 .8]);lsline
>>pa3=min(yval);
>>pa4=max(yval);
>>text(pa3-abs(pa3*0.04),1.02*pa4-abs(pa4*0.05),s12,'FontSize',16,'FontName','Arial');
>>text(pa3-abs(pa3*0.04),1.02*pa4-abs(pa4*0.08),s22,'FontSize',16,'FontName','Arial');
>>text(pa3-abs(pa3*0.04),1.02*pa4-abs(pa4*0.11),s32,'FontSize',16,'FontName','Arial');
>>ylabel('Valores previstos (%)','FontSize',20,'FontName','arial');
>>xlabel('Valores esperados (%)','FontSize',20,'FontName','Z'arial');
>>axis([15 85 0 100])
```

(20) Na Figura 3, está apresentado o gráfico dos valores esperados vs. valores previstos para as amostras de calibração e de validação externa. Os comandos para gerar o gráfico estão apresentados abaixo:

```
>>figure (4)
>>set(0,'defaultAxesFontName', 'arial')
>>axes('FontSize',16);
>>plot(yval,ypred1,'ko',yval,ypred2,'k*','LineWidth',2,'MarkerSize',15,'MarkerFaceColor',[.8 .8 .8]);
>>legend('Calibração','Validação Externa');
>>axis([.95*pa1 1.05*pa2 .95*pa1 1.05*pa2]);
>>h=refline(1,0);set(h,'color','k');
>>text(pa1-abs(pa1*0.04),1.1*pa4-abs(pa4*0.05),s12,'FontSize',14,'FontName','Arial');
>>text(pa1-abs(pa1*0.04),1.1*pa4-abs(pa4*0.08),s22,'FontSize',14,'FontName','Arial');
>>text(pa1-abs(pa1*0.04),1.1*pa4-abs(pa4*0.11),s32,'FontSize',14,'FontName','Arial');
>>text(pa1-abs(pa1*0.04),1.04*pa2-abs(pa2*0.16),s1,'FontSize',14,'FontName','Arial');
>>text(pa1-abs(pa1*0.04),1.04*pa2-abs(pa2*0.19),s2,'FontSize',14,'FontName','Arial');
>>text(pa1-abs(pa1*0.04),1.04*pa2-abs(pa2*0.22),s3,'FontSize',14,'FontName','Arial');
>>ylabel('Valores previstos (%)','FontSize',14,'FontName','arial');
>>xlabel('Valores esperados (%)','FontSize',14,'FontName','arial');
```



**Figura 3.** Gráfico dos valores esperados vs valores previstos para a calibração e validação externa

De acordo com a Figura 3, os coeficientes de determinação ( $R^2$ ) para a calibração e validação foram 0,995 e 0,989, respectivamente. Isso indica que o modelo construído apresenta boa concordância entre os valores esperados e previstos e uma boa capacidade preditiva.

Os valores de erro do modelo possuem a mesma unidade da propriedade de interesse (%v/v de azeite) o que permite avaliar o erro médio do modelo PLS construído. Os valores obtidos para o RMSEC (amostras de calibração), RMSEV (validação externa) e RMSEP (previsão) foram de 3,7, 3,5 e 4,3%, respectivamente.

### Previsão do percentual de azeite (v/v %) na amostra comercial de óleo composto

#### Metodologia univariada

As linhas de comando que foram empregadas nesta etapa estão disponíveis no material suplementar no formato de uma rotina ('regressaoilinear.m', rotina 2). A partir da amostra de óleo composto, foram coletadas 5 imagens digitais e em seguida foi aplicada a mesma metodologia para obter os histogramas e os valores médios de R, G e B. As imagens e os valores de R, G e B estão apresentadas na Figura 10S, disponível no material suplementar.

Após isso, as curvas analíticas obtidas no modelo univariado foram usadas para encontrar o percentual de azeite numa etapa de previsão. Isso foi feito, substituindo o valor de y nas equações da reta obtidas para os canais R, G e B pelos valores obtidos em cada imagem. Assim, foi calculado o valor do x que corresponde ao percentual de azeite previsto para a amostra comercial (12%) e para as misturas (30 e 70%). Após obter os valores para cada imagem, foi calculado o valor médio.

*Encontrar o percentual de azeite a partir dos valores de R, G e B e construir o gráfico comparando os valores obtidos, conforme mostrado na Figura 4*

(21) A sequência de linhas de comando apresentada nesta etapa mostra como calcular o percentual de azeite usando as equações da reta obtidas para os canais R, G e B.

```
>>ypredRED1=(coefangularR*valorRGBcomposto(:,1))+(coef
linearR);
>>ypredREDmean1=mean(ypredRED1);
>>ypredGREEN1=(coefangularG*valorRGBcomposto(:,2))+
(coeflinearG);
>>ypredGREENmean1=mean(ypredGREEN1);
>>ypredBLUE1=(coefangularB*valorRGBcomposto(:,3))+(coef
linearB);
```

```

>>yprdBBLUEmean1=mean(yprdBBLUE1);
>>x30=[valorRGB30R1;valorRGB30R2;valorRGB30R3];
>>x30mean=mean(x30);
>>yprdBRED2=((coefangularR*x30mean(:,1))+(coeflinearR));
>>yprdBGREEN2=((coefangularG*x30mean(:,2))+(coeflinearG));
>>yprdBBLUE2=((coefangularB*x30mean(:,3))+(coeflinearB));
>>x70=[valorRGB70R1;valorRGB70R2;valorRGB70R3];
>>x70mean=mean(x70);
>>yprdBRED3=((coefangularR*x70mean(:,1))+(coeflinearR));
>>yprdBGREEN3=((coefangularG*x70mean(:,2))+(coeflinearG));
>>yprdBBLUE3=((coefangularB*x70mean(:,3))+(coeflinearB));
>>yprdB1linear=[yprdBREDmean1;yprdBRED2;yprdBRED3];
>>yprdB2linear=[yprdBGREENmean1;yprdBGREEN2;yprdBGREEN3];
>>yprdB3linear=[yprdBBLUEmean1;yprdBBLUE2;yprdBBLUE3];

```

(22) Nessa etapa, estão apresentadas as linhas de comando para construir o gráfico comparativo entre os valores reais e os valores previstos. Foram usadas as misturas com 30 e 70% além da amostra comercial de óleo composto. O gráfico construído está apresentado na Figura 4.

```

>>figure (4)
>>set(0,'defaultAxesFontName','arial');
>>axes('FontSize',12);
>>h1=bar([yADULTval_ext yprdB1linear yprdB2linear yprdB3linear]);colormap gray;
>>legend('Valores reais','Canal R','Canal G','Canal B');
>>axis([0.5 3.5 0 100]);
>>xlabel('Amostras de previsão: 12,30 e 70%','FontName','arial','FontSize',14);
>>ylabel('Percentual de azeite (v/v%)','FontSize',14,'FontName','arial');
>>hline(0,'k');

```

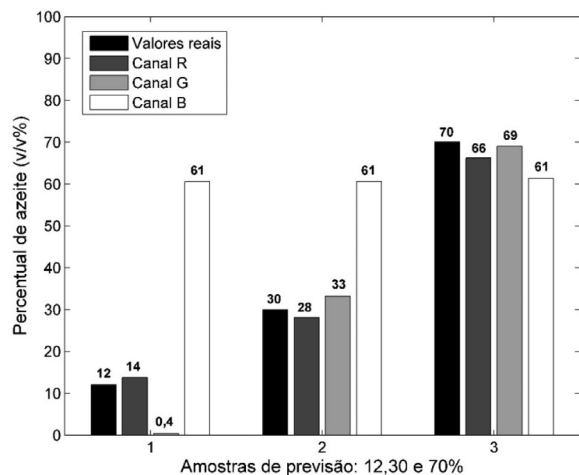


Figura 4. Valores esperados vs valores previstos para a amostra comercial de óleo composto (12%) e as misturas de 30 e 70%, nos canais R, G e B

De acordo com a Figura 4, o canal R foi o que mais se aproximou do valor real, considerando as misturas de 30, 70% e a amostra comercial de óleo composto (12%).

#### Metodologia multivariada

Nessa etapa, foi feita a previsão do teor de azeite presente nas misturas de 30%, 70% e na amostra comercial de óleo composto (12%). As linhas de comando usadas estão disponíveis no material suplementar no formato de uma rotina ("rotinaPLS", rotina 3).

Calcular o valor do RMSEP e construir o gráfico comparando os valores obtidos, conforme mostrado na Figura 5

(23) Os valores previstos do percentual de azeite na amostra de óleo composto serão mostrados no vetor 'ypred3multi'. Em seguida, foi calculado o valor do RMSEP.

```

>>[ypred3multi]=PLSPRED(xcalt_c_msc,ycal,xmedt_c_msc,0,2);
>>rmsep=sqrt((sum((ymed-ypred3multi).^2))/(length(ymed)));

```

(24) Criar um gráfico com os valores esperados vs. Valores previstos pelo modelo PLS para o conjunto de previsão, conforme mostrado na Figura 5.

```

figure (5)
>>set(0,'defaultAxesFontName','arial')
>>axes('FontSize',14);
>>h1=bar([yprev ypred3multi]);colormap gray
>>legend('Valores esperados (%)','Valores previstos (%)')
>>axis([0.5 3.6 0 80])
>>xlabel('Amostras de previsão:12,30 e 70%','FontName','arial','FontSize',14);
>>ylabel('Percentual de azeite(v/v%)','FontSize',14,'FontName','arial');
>>hline(0,'k')

```

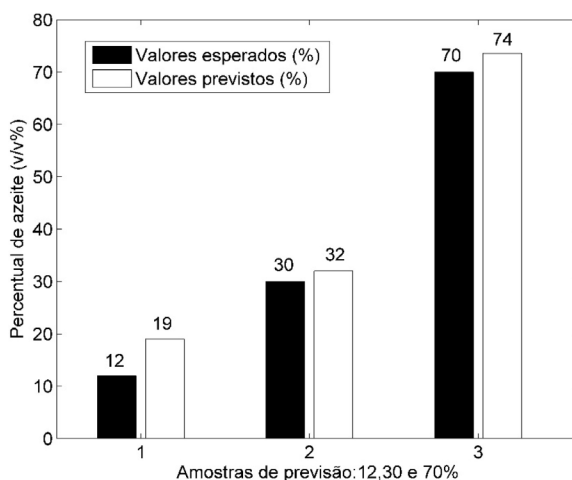


Figura 5. Valores esperados vs valores previstos para a amostra comercial de óleo composto (12%) e as misturas usadas na previsão (30 e 70%) usando o modelo multivariado

Os valores preditos pelo modelo PLS apresentaram boa proximidade com os valores de referência, uma vez que as misturas não estavam no conjunto de calibração. Considerando que o erro de previsão do modelo foi de 4,3% e o valor percentual de azeite na amostra é de 12%, a faixa de valores previstos pelo modelo construído seria de 8,3 a 16,3% (v/v %).

Por fim, foi gerado um gráfico comparativo entre os resultados obtidos usando a abordagem univariada (canal R) com a multivariada (histogramas). As linhas de comando que foram usadas estão disponíveis no material suplementar ("rotinaPLS", rotina 3). A Figura 11S, disponível no Material Suplementar, apresenta o gráfico gerado comparando as duas abordagens.

Construir o gráfico comparando os valores obtidos para a abordagem multivariada e univariada

(25) Construir o gráfico comparativo entre os resultados obtidos com o modelo multivariado e o univariado;

```

>> load('artigo histogramasRGB.mat');
>> load('artigo valoresRGB.mat');

```



```

>> rotinaartigoregressaolinear;
>> rotinaartigoPLS;
>> figure (6)
>> set(0,'defaultAxesFontName','arial');
>> axes('FontSize',14);
>> h1=bar([yprev ypred1linear ypred3multi]);colormap gray;
>> legend('Valores esperados (%)','Univariado (Canal R)','Multivariado');
>> axis([0.5 3.6 0 80]);
>> xlabel('Amostras de previsão:12,30 e 70%','FontName','arial','FontSize',14);
>> ylabel('Percentual de azeite(v/v%)','FontSize',14,'FontName','arial');
>> hline(0,'k')

```

Na abordagem univariada, os valores previstos para as replicatas da amostra de óleo composto podem ser encontrados na matriz chamada “*ypredRED1*”, gerada ao executar as linhas de comando presentes na rotina 2. (“*regressaolinear.m*”, rotina 2’). Os valores previstos para as amostras de 30 e 70 % estão nas matrizes “*ypredRED2*” e “*ypredRED3*”, respectivamente. Já para a abordagem multivariada, os valores previstos para as amostras de validação podem ser encontrados na matriz, “*ypred3multi*” gerada ao executar a rotina 3. (“*rotinaPLS.m*”, rotina 3).

De acordo com a Figura 11S, é possível concluir que ambas as metodologias apresentaram resultados parecidos e próximos ao valor percentual de azeite presente nas misturas de 30 e 70%. No entanto, para a amostra de óleo composto, a abordagem univariada (14%) obteve um resultado mais próximo ao valor esperado do que a multivariada (19%). Além disso, a metodologia univariada se destaca devido à sua simplicidade e fácil implementação, já que não é necessário usar ferramentas quimiométricas por meio de *softwares* ou *toolbox* para construir os modelos multivariados.

Outro ponto a ser destacado é que a mistura azeite-óleo não apresentou alterações de cor em toda a extensão da imagem digital, ou seja, os valores de intensidade de cor de cada *pixel* estavam muito próximos. Esse fato contribuiu para que a metodologia univariada apresentasse bons resultados, pois essa uniformidade proporcionou que apenas o valor médio de R, G e B fosse o suficiente para representar toda a variabilidade presente na imagem digital.

## CONCLUSÕES

O emprego de imagens digitais com o objetivo de quantificar a porcentagem de azeite de oliva em misturas adulteradas com óleo de soja se mostrou uma opção viável, econômica, rápida e com potencial didático para ser usada em disciplinas de graduação e pós-graduação. A estratégia de misturar as amostras comerciais em uma única amostra foi importante, pois gerou modelos mais robustos e representativos. Para quantificar o teor de azeite nas misturas, a abordagem univariada se mostrou mais adequada, em especial com o uso do canal R.

A metodologia proposta apresenta potencial para ser empregada em uma etapa de triagem de amostras apreendidas de azeite de oliva extra-virgem, possivelmente adulteradas com óleo de soja, um adulterante muito comum. Ressalta-se, no entanto, que a informação coletada é baseada em uma alteração física, ou seja, uma mudança na cor da mistura a partir da variação na porcentagem de óleo-azeite. Por isso, o emprego de imagens digitais não substitui o uso de técnicas analíticas, especialmente espectroscópicas, as quais irão fornecer informações químicas sobre a composição das amostras.

## IN MEMORIAM

Esta publicação é uma homenagem ao Prof. Dr. Ronei Jesus Poppi, que iniciou as publicações dos tutoriais de Quimiometria da Revista Química Nova e foi orientador de pós-graduação de todos os autores do trabalho.

## MATERIAL SUPLEMENTAR

As Figuras 1S até a 11S, bem como as linhas de comando para as rotinas 1, 2 e 3 estão no material suplementar disponível em <http://quimicanova.s bq.org.br>, na forma de arquivo PDF com acesso livre. Na Figura 1S são apresentados as amostras de óleo de soja, azeite de oliva e óleo composto que foram usadas no estudo. A Figura 1S-A apresenta as 5 marcas diferentes de óleo de soja (Concordia, Suavit, Vilha Velha, Soya e Liza). A Figura 1S-B apresenta as 6 marcas de azeite de oliva (Borges, Allegro, La española, gallo, O-L!ve, Andorinha) e a amostra de óleo composto (12% azeite de oliva, Maria). A Figura 2S, apresenta o dispositivo construído para coletar as imagens digitais, composto por uma caixa de isopor, LEDs de autobrilho na cor branca e pilhas. Na Figura 3S, é apresentado um esquema da metodologia usada para obter as matrizes R, G e B a partir da imagem digital original obtida pelo smartphone. Na Figura 4S é apresentado um esquema da metodologia usada para obter os histogramas médios para cada imagem digital a partir das matrizes de cada canal do sistema RGB. As Figuras 5S e 6S apresentam as imagens digitais de azeite de oliva e óleo de soja de todas as marcas adquiridas. A coleta foi feita em triplicata e evidenciou que não houve variação de uma foto para outra. A Figura 7S apresenta as imagens das misturas adulteradas preparadas entre óleo e azeite em diferentes proporções numa faixa de 0 a 100%. As Figuras 8S e 9S apresentam os gráficos dos valores esperados versus os valores previstos do modelo de calibração e o gráfico dos valores esperados versus os vs valores previstos pelo modelo para validação externa, respectivamente. Na Figura 10S estão as imagens digitais da amostra de óleo composto usada na etapa de previsão. Além disso, em cada imagem estão apresentados os valores médios de R, G e B que foram usados para obter o teor de adulteração da amostra. A Figura 11S mostra um comparativo entre as duas metodologias abordadas, univariada e multivariada. Na abordagem univariada, foi considerado o canal R, que apresentou os melhores resultados (R<sup>2</sup>).

## AGRADECIMENTOS

Os autores agradecem ao Programa estágio Docente (PED) da UNICAMP, ao apoio e incentivo do Prof. Ronei Jesus Poppi (*in memoriam*). Agradecemos também aos órgãos de fomento à pesquisa: CAPES (Finance Code 001) e ao CNPq/FAPESP por meio do INCT Bioanalítica.

## REFERÊNCIAS

- de Souza, A. M.; Poppi, R. J.; *Quim. Nova* **2012**, *35*, 223. [Crossref]
- de Souza, A. M.; Breikreitz, M. C.; Filgueiras, P. R.; Rohwedder, J. J. R.; Poppi, R. J.; *Quim. Nova* **2013**, *36*, 1057. [Crossref]
- Breikreitz, M. C.; de Souza, A. M.; Poppi, R. J.; *Quim. Nova* **2014**, *37*, 564. [Crossref]
- Da-Col, J. A.; Dantas, W. F. C.; Poppi, R. J.; *Quim. Nova* **2018**, *41*, 345. [Crossref]
- de Santana, F. B.; de Souza, A. M.; Almeida, M. R.; Breikreitz, M. C.; Filgueiras, P. R.; Sena, M. M.; Poppi, R. J.; *Quim. Nova* **2020**, *43*, 371. [Crossref]
- Gonzales, R. C.; Woods, R. E.; Eddins, S. L.; *Digital image processing using MATLAB*, 1<sup>st</sup> ed., Pearson Education: Upper Saddle River, 2004.

7. Wu, D.; Sun, D. W.; *Trends Food Sci. Technol.* **2013**, *29*, 5. [Crossref]
8. Choudhury, A. K. R.; *Principles of color and appearance measurement*, 1<sup>st</sup> ed., Woodhead Publishing: Cambridge, **2014**
9. Byrne, L.; Barker, J.; Pennarun-Thomas, G.; Diamond, D.; Edwards, S.; *TrAC – Trends Anal. Chem.* **2000**, *19*, 517. [Crossref]
10. Roda, A.; Michelini, E.; Zangheri, M.; Fusco, M.D.; Calabria, D.; Simoni, P.; *TrAC – Trends Anal. Chem.* **2016**, *79*, 317. [Crossref]
11. Diniz, P. H. G. D.; *J. Chem.* **2020**, *34*, 3242. [Crossref]
12. Paquet-Durand, O.; Solle, D.; Schirmer, M.; Becker, T.; Hitzmann, B.; *J. Food. Eng.* **2012**, 111. [Crossref]
13. Jackman, P.; Da-Wen, S.; Allen, P.; Valous, N. A.; Mendoza, F.; Ward, P.; *Meat Sci.* **2010**, *84*, 711. [Crossref]
14. Tôrres, A. R.; Lyra, W.; Andrade, S. I. E.; Andrade, R. A. N.; Silva, E. C.; Araujo, M. C. U.; Gaião, E. N.; *Talanta* **2011**, *84*, 601. [Crossref]
15. Borin, A.; Ferrão, M. F.; Mello, C.; Cordi, L.; Pataca, L. C. M.; Duran, N.; Poppi, R. J.; *Anal. Bioanal. Chem.* **2007**, 387, 1105. [Crossref]
16. Borah, S.; Bhuyan, M.; *Int. J. Food. Sci. Technol.* **2005**, *40*, 675. [Crossref]
17. Barbedo, J. G. A.; *SpringerPlus* **2013**, *2*, 660. [Crossref]
18. Van-Dalen, G.; *Food Res. Int.* **2004**, *37*, 51. [Crossref]
19. Fulgêncio, A. C. C.; Araújo, V. P. T.; Pereira, H.V.; Botelho, B. G.; Sena, M. M.; *Food Anal. Methods* **2020**, *13*, 303. [Crossref]
20. Domínguez M. A.; Diniz, P. H. G. D.; Di Nezio, M. S.; Araújo, M. C. U.; Centurión, M. E.; *Microchem. J.* **2014**, *112*, 104. [Crossref]
21. Pierini, G. D.; Fernandes, D. D. S.; Diniz, P. H. G. D.; Araújo, M. C. U.; Di Nezio, M. S.; Centurión, M. E.; *Microchem. J.* **2016**, *128*, 62. [Crossref]
22. Almeida, V. E.; Costa, G. B.; Fernandes, D. D. S.; Diniz, P. H. G. D.; Brandão, D.; Medeiros, A. C. D.; Véra, G.; *Anal. Bioanal. Chem.* **2014**, 406, 5989. [Crossref]
23. Morais, C. L. M.; Lima, K. M. G.; *Anal. Methods* **2015**, *7*, 6904. [Crossref]
24. Marcelo, M. C. A.; Mariotti, K. C.; Ortiz, R. S.; Ferrão, M. F.; Anzanello, M. J.; *Microchem. J.* **2016**, *127*, 87. [Crossref]
25. Costa, G. B.; Fernandes, D. D. S.; Almeida, V. E.; Araújo, T. S. P.; Melo, J.P.; Diniz, P. H. G. D.; Véras, G.; *Talanta* **2015**, *139*, 50. [Crossref]
26. Costa, G. B.; Fernandes, D. D. S.; Almeida, V. E.; Maia, M. S.; Araujo, M. C. U.; Veras, G.; Diniz, P. H. G. D.; *Anal. Methods* **2016**, *8*, 4949. [Crossref]
27. Valderrama, L.; Valderrama, P.; *Chemom. Intell. Lab. Syst.* **2016**, *156*, 188. [Crossref]
28. Ammari, F.; Del-Solar-Velarde, N. ; Chapoulie, R.; Bousquet, B.; *Environ. Sci. Pollut. Res.* **2017**, *24*, 2205. [Crossref]
29. Spink, J.; Moyer, D. C.; *J. Food Sci.* **2011**, *76*, 157. [Crossref]
30. Rodríguez, S. D.; Rolandelli, G.; Buera, M. P.; *Food Chem.* **2019**, *274*, 392. [Crossref]
31. Santos, P. C.; Tosato, F.; Cesconetto, M.; Corrêa, T.; Santos, F. D.; Lacerda, L.; Pires, A. A.; Ribeiro, A. V. F. N.; Filgueiras, P. R.; Romão, W.; *Quim. Nova* **2020**, *43*, 891. [Crossref]
32. Duraipandian, S.; Petersen J. C.; Lassen, M.; *Appl. Sci.* **2019**, *9*, 2433. [Crossref]