

## Química Nova



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution Non-Commercial, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que sem fins comerciais e que o trabalho original seja corretamente citado. Fonte: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-40422016000700876&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-40422016000700876&lng=en&nrm=iso). Acesso em: 12 mar. 2018.

## REFERÊNCIA

SILVA NETO, Gerson F.; FONSECA, Alexandre; BRAGA, Jez W. B. Classificação de águas minerais baseada em imagens digitais obtidas por smartphones. **Química Nova**, São Paulo, v. 39, n. 7, p. 876-881, ago. 2016. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-40422016000700876&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-40422016000700876&lng=en&nrm=iso)>. Acesso em: 12 mar. 2018. doi: <http://dx.doi.org/10.5935/0100-4042.20160088>.

## CLASSIFICAÇÃO DE ÁGUAS MINERAIS BASEADA EM IMAGENS DIGITAIS OBTIDAS POR SMARTPHONES

Gerson F. Silva Neto, Alexandre Fonseca\* e Jez W. B. Braga

Instituto de Química, Universidade de Brasília, 70910-900, Brasília – DF, Brasil

Recebido em 27/01/2016; aceito em 03/03/2016; publicado na web em 25/05/2016

CLASSIFICATION OF MINERAL WATERS BASED ON DIGITAL IMAGES ACQUIRED BY SMARTPHONES. This work describes a new procedure for classification of mineral waters based on digital images acquired by smartphones. Commercial waters from eight mineral springs plus distilled water and tap water were combined with eriochrome T black or murexide and transferred to a cuvette, which was positioned into a light controlled chamber. RGB (*Red, Blue and Green*) measurements of cuvette images were acquired in real time, using a free smartphone app, and employed as variables for the exploratory analysis. 2D data dispersion along component B for murexide (x axis) and component R for eriochrome T black (y axis) provides the clear visualization of clusters using the raw variables. Hierarchical cluster analysis (HCA) applied to this data confirmed the efficient discrimination of samples providing the characterization of nine clusters for the ten classes of water investigated. The classification of samples based on a k-nearest neighbors (k-NN) modelled to the efficiency rate of 100% for 8 classes and of 94.4% and 50% for the remaining classes, respectively, indicating the adequate performance of the proposed strategy. Considering the facilities to acquire the data, such as low cost instrumentation and reagents, and the rapidity of the procedures, this alternative may be applied for verification of commercial water adulteration.

Keywords: smartphone; mineral waters; colorimetry; exploratory analysis.

## INTRODUÇÃO

O uso de imagens digitais como ferramenta para análises químicas tem sido amplamente explorado nos últimos anos.<sup>1</sup> Empregando-se imagens obtidas por câmeras fotográficas, *scanners*, *webcams* ou telefones celulares, conhecidos como *smartphones*, é possível realizar determinações qualitativas e quantitativas com desempenho adequado, sem a necessidade de instrumentos como espectrofluorímetros ou espectrofotômetros.<sup>2-4</sup> De fato, a literatura relata que, utilizando informações simples das imagens, como os valores de cores usualmente baseadas no sistema RGB (do inglês *Red, Blue e Green*) consegue-se obter resultados analíticos similares àqueles adquiridos com instrumentação já bem estabelecida, caracterizando esta estratégia em uma alternativa de menor custo para a realização de determinadas medidas instrumentais.

Ao se considerar as aplicações analíticas que podem ser conduzidas por meio de imagens digitais, merece destaque aquelas que envolvem a identificação de substâncias ou a discriminação/classificação de amostras baseadas em medidas colorimétricas.<sup>5</sup> Neste sentido, arranjos com sensores contendo diferentes reagentes colorimétricos imobilizados sobre filmes poliméricos ou de papel têm sido aplicados com sucesso na resolução de muitos desafios analíticos, como a discriminação de explosivos ou de aminas e a identificação de substâncias de interesse variado como compostos orgânicos voláteis, produtos químicos tóxicos ou aromatizantes do café.<sup>6-11</sup> Na maioria destas aplicações, imagens do arranjo de sensores, obtidas após a interação deste com o sistema em estudo, proporcionam uma espécie de impressão digital da amostra, o que permite classificá-los, muitas vezes, como narizes ou línguas eletrônicas.<sup>5</sup>

Embora este tipo de estratégia forneça um conjunto de dados amplo, pois considera as informações do sistema RGB para cada reagente colorimétrico imobilizado, e torne o procedimento de medidas simples, necessitando apenas da interação da amostra com o arranjo de sensores, há sempre a necessidade do desenvolvimento dos

dispositivos, o que pode ser considerado um limitador para alguns grupos de pesquisa. Além disso, quando as respostas para apenas um reagente colorimétrico ou para um pequeno número deles (no máximo três) é suficiente para a resolução do problema o desenvolvimento dos sensores pode ser desnecessário. De fato, muitos trabalhos demonstram que a adição dos reagentes colorimétricos diretamente à amostra, com posterior realização das medidas por imagens podem fornecer resultados qualitativos ou quantitativos satisfatórios, permitindo avaliar preliminarmente o potencial de determinados reagentes para o futuro desenvolvimento de sensores imobilizados. Oliveira *et al.* aplicaram a análise de componentes principais à diversos descritores de cores de imagens digitais, entre eles o sistema RGB, para a classificação de solos contaminados com diferentes teores de cromo.<sup>12</sup> Os autores relataram que o método proposto é bastante simples e permite verificar se os solos apresentam teores de cromo superiores ou inferiores ao estabelecido por legislações de diferentes países. Damasceno e colaboradores utilizaram a análise multivariada de imagens para estimar o pH de amostras de água potável a partir da mistura das amostras com o reagente azul de bromotimol e aplicação das informações do sistema RGB em modelo de calibração por mínimos quadrados parciais.<sup>13</sup> Lima *et al.* introduziram a combinação de imagens digitais, obtidas por uma *webcam*, com um mini-sistema fluxo-batelada. O sistema foi aplicado para a determinação de ácido tânico em chá-verde, sendo obtidos resultados que demonstraram um desempenho analítico satisfatório, além de permitir uma redução significativa no consumo de reagentes e amostras, assim como a diminuição na geração de resíduos.<sup>14</sup>

Segundo o Departamento Nacional de Produção Mineral o Brasil apresenta-se como o quarto mercado consumidor de água engarrafada do mundo com um consumo *per capita* da ordem de 90 litros/ano, o que corresponde a cerca de 17 bilhões de litros comercializados todo ano.<sup>15</sup> Por se tratar de um produto para consumo humano, as águas minerais devem ser submetidas à caracterização química para aferição de sua composição mineral, o que permite diferenciá-las da água potável da rede pública assim como das chamadas águas mineralizadas. Para este propósito, técnicas analíticas como a potenciometria,

\*e-mail: afonseca@unb.br

espectrofotometria, espectrometria atômica e a volumetria são quase sempre utilizadas para a determinação dos diferentes constituintes, o que requer pessoal treinado e um tempo considerável para obtenção dos resultados.<sup>16,17</sup> Neste sentido, o desenvolvimento de métodos de reconhecimento de padrão baseado em medidas qualitativas ou semi-quantitativas por imagens digitais podem surgir como uma alternativa para a classificação de águas minerais de forma mais simples e rápida, constituindo-se em uma ferramenta que permita a verificação preliminar de possíveis adulterações sem a necessidade das análises quantitativas mais laboriosas. Em um trabalho desenvolvido por Fonseca et al. foi demonstrado que a combinação das águas minerais com complexantes metálicos e medidas fotométricas realizadas com um instrumento *lab-made* à base de diodos emissores de luz (LED) é capaz de proporcionar a discriminação adequada de diferentes águas minerais. No referido trabalho, as medidas realizadas com dois agentes quelantes e seleção de apenas quatro comprimentos de onda foram suficientes para a classificação correta das amostras. Isto mostra que o procedimento não exige um grande número de variáveis, de modo que a aplicação de medidas colorimétricas realizadas diretamente sobre a amostra complexada com o sistema RGB a partir de um smartphone pode proporcionar resultados semelhantes, facilitando e tornando ainda menos oneroso o procedimento de medidas.

Considerando estes aspectos, o presente manuscrito descreve um método rápido, simples e de baixo custo para a classificação de águas minerais comerciais baseada na reação dos íons metálicos constituintes com reagentes colorimétricos, aquisição de parâmetros RGB de imagens das soluções resultantes com um *smartphone* e aplicação do método de reconhecimento de padrões pela regra do vizinho mais próximo (k-NN) aos dados gerados. O desempenho apresentado demonstra a possibilidade de uso da estratégia proposta para facilitar a classificação de águas minerais além de contribuir para o futuro desenvolvimento de arranjos de sensores colorimétricos específicos para a discriminação/classificação de águas minerais.

## PARTE EXPERIMENTAL

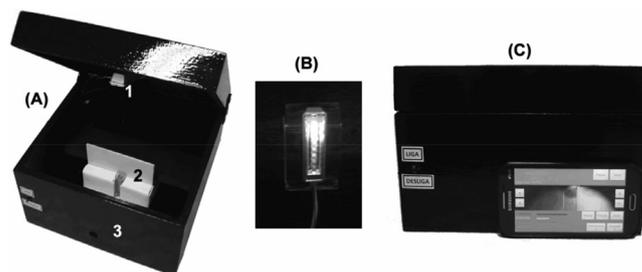
### Aparatos

Medidas de pH das soluções empregadas foram realizadas com um pHmetro HANNA pH 21. Cubetas de poliestireno com caminho óptico de 10 mm e capacidade de 4,5 mL foram utilizadas em todos os experimentos. Um suporte para cubeta (Figura 1) foi construído em poli-metilmetacrilato (PMMA) e revestido com espuma vinílica acetinada (E.V.A.) de cor branca nas regiões de contato com as faces laterais e traseira da cubeta, deixando a face frontal aberta para a aquisição das imagens.

Todas as medidas foram realizadas em um ambiente de luz controlada em uma caixa de madeira (Figura 1 A) com dimensões de 22 cm x 22 cm x 16 cm (largura x comprimento x altura), a qual foi utilizada para acomodar o suporte para cubeta e um arranjo com três diodos emissores de luz branca (LED), conforme mostrado na Figura 1 B. Um orifício com diâmetro de 1,0 cm foi confeccionado na parte frontal da caixa e utilizado como janela de observação para as medidas. Com este arranjo, o celular era posicionado manualmente a 8,0 cm da face frontal da cubeta na parte externa da caixa (Figura 1 C). Vale destacar que existem trabalhos na literatura mostrando a possibilidade de correção do sinal RGB para diferentes condições de exposição à luz, o que pode eliminar a necessidade de ambientes de luz controlada como o utilizado aqui.<sup>18</sup> Contudo, tendo em vista que o controle de exposição à luz pode ser facilmente obtido pelo sistema proposto, optou-se por não empregar esse tipo de correção de sinal.

As leituras das componentes RGB de uma pequena área da imagem (aproximadamente 0,25 cm<sup>2</sup>) foram efetuadas em tempo real a

partir do aplicativo para smartphones *color-picker* (Real time) e de um celular equipado com câmera fotográfica com resolução de 5.0 Mp.<sup>19</sup>



**Figura 1.** Fotografias da caixa de madeira utilizada como ambiente de luz controlada para a realização das medidas. Visão geral (A) com indicações do arranjo de LED (1), suporte para cubeta (2) e janela de observação (3), detalhe do arranjo de iluminação com três LED (B) e posicionamento do celular à frente da caixa (C)

### Reagentes e Soluções

Todas as soluções foram preparadas com água purificada por osmose reversa e com reagentes apresentando, ao menos, 99% de pureza. Uma solução tampão amônia/cloreto de amônio 0,3 mol L<sup>-1</sup> e pH 10 foi preparada dissolvendo 8,025 g de NH<sub>4</sub>Cl (Synth) em 450 mL de água, ajustando-se o pH com uma solução de hidróxido de sódio (5,0 mol L<sup>-1</sup>) e completando-se o volume à 500 mL com água. Uma solução de murexida (MUR) 0,3 mmol L<sup>-1</sup> foi preparada dissolvendo-se 0,022 gramas do indicador (Merck) em 25 mL de etanol com posterior diluição a 250,00 mL com água. Uma solução de negro de eriocromo T (EBT) 0,2 mmol L<sup>-1</sup> foi preparada dissolvendo 0,01 g do indicador (Synth) em 100 mL de uma solução 0,16% (m V<sup>-1</sup>) de tetraborato de sódio decahidratado (Synth), a qual foi preparada em metanol. Ambas as soluções de indicadores foram preparadas no mesmo dia em que foram utilizadas e armazenadas em frascos de vidro âmbar.

Águas minerais comerciais provenientes de oito fontes distintas foram adquiridas no mercado local de Brasília e analisadas sem nenhum pré-tratamento. Nos referidos estudos, foram utilizadas cinco amostras de cada fonte, as quais pertenciam a lotes diferentes. Além destas águas, foram utilizadas também amostras de água destilada e de torneira como amostras de controle, sendo todas as águas avaliadas listadas na Tabela 1.

Para as medidas, 15,00 mL da amostra foram transferidos para um balão volumétrico de 25,00 mL, seguido da adição de 5,00 mL de solução tampão e 3,00 mL de solução de MUR ou 75 µL de solução de EBT, completando-se o volume remanescente com água purificada.

**Tabela 1.** Amostras de águas analisadas durante os estudos

| Águas | Fontes ou tipo | Datas dos envases       |
|-------|----------------|-------------------------|
| 1     | Mina da Lua    | 21/11/2014 a 06/01/2015 |
| 2     | Oriente        | 30/03/2014 a 22/11/2014 |
| 3     | Quartzo        | 09/09/2014 a 29/01/2015 |
| 4     | Primavera      | 26/09/2014 a 21/01/2015 |
| 5     | Alexânia       | 22/05/2014 a 23/12/2014 |
| 6     | Água Santa     | 29/09/2014 a 28/12/2014 |
| 7     | Marisa         | 23/09/2014 a 15/11/2014 |
| 8     | Primavera      | 24/09/2014 a 27/10/2014 |
| 9     | Destilada      | -                       |
| 10    | Torneira       | -                       |

## Procedimentos

A fim de obter um controle mais apurado do tempo de reação entre os reagentes colorimétricos (EBT ou MUR) com os íons metálicos presentes nas amostras, realizaram-se as leituras dos sinais RGB um minuto após a adição do indicador e ajuste do volume do balão. Posteriormente, as soluções foram transferidas para a cubeta, a qual foi acoplada ao suporte no interior da caixa com luz controlada. Concluída esta etapa, a câmera do celular foi posicionada na janela de leitura e a aquisição dos sinais foi realizada a partir do aplicativo, sendo os dados anotados para posterior avaliação. Vale destacar que foram realizadas medidas em triplicata para as cinco amostras de águas minerais de cada fonte, totalizando um conjunto de 15 medidas instrumentais por fonte.

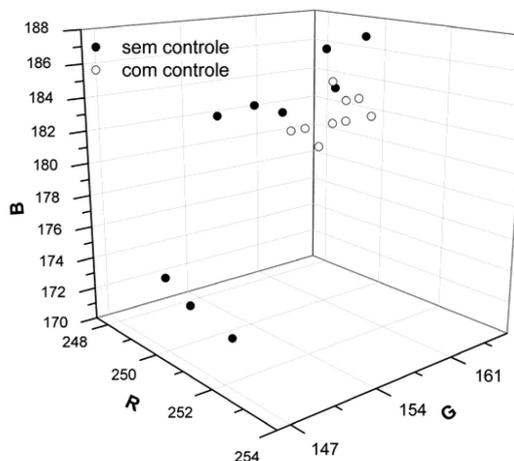
## Análise quimiométrica

Para verificar a capacidade de discriminar entre as diferentes águas minerais estudadas através dos dados colorimétricos, uma análise exploratória inicial dos dados foi realizada empregando análise hierárquica de agrupamentos (HCA, do inglês *Hierarchical Cluster Analysis*), executada com base na distância euclidiana e por algoritmo de ligação completa. A análise de componentes principais (PCA, do inglês *Principal Component Analysis*) também foi empregada utilizando as variáveis RGB obtidas com os dois reagentes colorimétricos para visualizar a separação dos grupos e determinar quais as variáveis mais importantes para a separação observada. Por fim, as amostras foram divididas em dois conjuntos de dados independentes na proporção de 2/3 para o desenvolvimento e 1/3 para a validação de um método de classificação baseado na regra do vizinho mais próximo (k-NN, do inglês *k-Nearest Neighbor*).<sup>20</sup> Todos os cálculos foram executados empregando o programa *Matlab* (versão 2011b) e *PLS-Toolbox* (versão 7.03).

## RESULTADOS E DISCUSSÃO

Os reagentes colorimétricos MUR e EBT são utilizados como indicadores para íons metálicos em titulações complexométricas e produzem soluções com cores ligeiramente diferentes a depender da concentração e da identidade dos íons no meio.<sup>21</sup> Baseado nestas diferenças foi demonstrado que a análise exploratória de águas minerais pode ser realizada com sucesso a partir de medidas fotométricas em sistemas de análise em fluxo e do tratamento quimiométrico dos dados obtidos.<sup>22</sup> Considerando que no presente trabalho a análise em fluxo não foi utilizada, foi preciso avaliar a influência do controle de tempo de reação na repetibilidade/dispersão dos dados obtidos.

Estudos preliminares com o reagente murexida mostraram que a leitura do sinal RGB deve ser realizada com o controle do tempo de reação entre os íons metálicos da amostra e o reagente colorimétrico. De fato, conforme apresentado na Figura 2, que mostra a projeção dos dados em função das variáveis RGB, os resultados em triplicata obtidos para três amostras de uma mesma água, sem o controle do tempo reacional, ou seja, com cada triplicata sendo medida com tempos diferentes após a mistura com os reagentes (30 s, 1 min. e 2 min.), tendem a apresentar uma maior dispersão do que aqueles obtidos quando todas as leituras são realizadas exatamente um minuto após a adição dos reagentes colorimétricos. Deste modo, todos os estudos posteriores foram realizados com o controle de tempo entre a adição do reagente colorimétrico e a leitura dos sinais RGB com o objetivo de minimizar a dispersão dos dados não relacionada à composição das amostras e a qual poderia limitar a análise exploratória das águas estudadas.



**Figura 2.** Dispersão dos dados obtidos sem o controle de tempo reacional (medidas das triplicatas realizadas a 30s, 1 min. e 2 min. após a mistura dos reagentes) e com o controle do tempo reacional (triplicatas medidas exatamente 1 minuto após a adição dos reagentes e homogeneização da solução) para três amostras de água de torneira em relação às componentes RGB das imagens obtidas com o reagente murexida

Conforme apresentado na Figura 3(A), a distribuição dos dados obtidos a partir das medidas com EBT permitiram a discriminação em quatro grupos de amostras referentes a diferentes águas (3, 4, 9 e 10) sendo que as demais amostras (1, 2, 5, 6, 7 e 8) não formaram agrupamentos individuais. Vale destacar que as águas 9 e 10 referem-se à água destilada e água de torneira, respectivamente, para as quais já era esperada uma boa discriminação dada a diferença de concentração de íons metálicos nestas águas em relação às águas minerais.

Utilizando a projeção dos dados com relação aos valores de R, G e B obtidos apenas com o reagente MUR foi possível observar uma melhor discriminação das águas avaliadas. A Figura 3(B) mostra que, neste caso, ocorre a separação individual das amostras 1, 2, 4, 8 e 10, sendo a amostra 10 referente à água de torneira. Além destes agrupamentos de amostras, observou-se também grupos formados pela combinação das amostras pertencentes aos grupos 3 e 9 e dos grupos 5, 6 e 7. A justificativa para o agrupamento destas amostras pode ser obtida das concentrações de íons cálcio fornecida nos rótulos e apresentadas na Tabela 2 e pelo fato do reagente MUR apresentar maior seletividade aos íons cálcio do que aos íons magnésio.<sup>21</sup> Como pode ser observado, a água 3 apresenta a menor concentração de cálcio, o que pode ter resultado no agrupamento desta água com a água destilada (praticamente isenta de cálcio). Da mesma forma, para as águas 5, 6 e 7, apresentado no detalhe da Figura 3 (B) observa-se uma grande semelhança nas concentrações de íons cálcio o que certamente acarretou em valores semelhantes para as componentes RGB destas amostras.

Outra constatação importante destes estudos iniciais é que, embora as amostras 3 e 9 não tenham formado grupos individuais, ao se utilizar o indicador MUR, grupos bem definidos destas amostras foram observados para as medidas com EBT, possivelmente pelo fato do EBT responder a outros íons presentes nas amostras de água 3. Desta observação propôs-se a combinação das variáveis obtidas com os dois reagentes a fim de se otimizar a discriminação das águas. Para avaliar a contribuição de cada variável para a variância total dos dados, determinou-se a variância das medidas realizadas com cada uma das componentes R, G e B com ambos os reagentes, sendo os resultados apresentados na Tabela 3. Como pode ser observado, os maiores coeficientes de variação foram determinados para as variáveis B da MUR e R do EBT, as quais foram utilizadas para compor o gráfico de dispersão mostrado na Figura 4. Neste gráfico, observa-se a

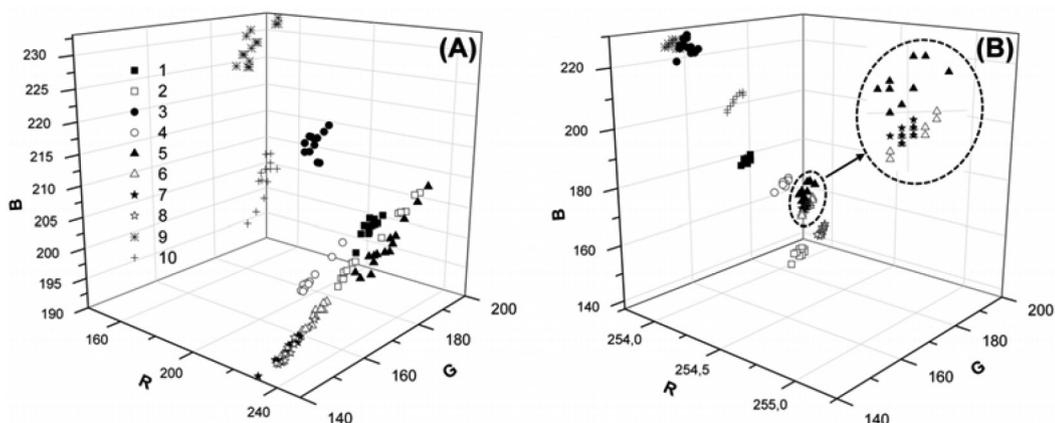


Figura 3. Dispersão dos dados obtidos para as medidas em triplicata das amostras de águas avaliadas utilizando-se os indicadores EBT (A) e MUR (B)

Tabela 2. Concentrações em cálcio e magnésio apresentadas nos rótulos das águas minerais estudadas

| Águas minerais | [Cálcio] mg L <sup>-1</sup> | [Magnésio] mg L <sup>-1</sup> |
|----------------|-----------------------------|-------------------------------|
| 1              | 4,701                       | 3,452                         |
| 2              | 26,490                      | 11,210                        |
| 3              | 0,119                       | 0,266                         |
| 4              | 12,598                      | 1,598                         |
| 5              | 14,291                      | 4,556                         |
| 6              | 16,107                      | 9,076                         |
| 7              | 16,780                      | 9,606                         |
| 8              | 24,200                      | 14,220                        |

formação de oito agrupamentos, sendo sete deles correspondentes às amostras individuais 1, 2, 3, 4, 8, 9 e 10 e um grupo composto pelas amostras 5, 6 e 7, o que mostra que apenas estas duas variáveis são capazes de produzir uma separação mais eficiente das águas estudadas em relação ao uso das variáveis RGB obtidas com cada reagente cromogênico de forma individual (Figura 3). É importante destacar que a dispersão dos dados em três dimensões utilizando as variáveis G e B para MUR e R para EBT ou R e G para EBT e B para MUR não proporcionaram separações visualmente mais eficientes do que aquela conseguida com apenas duas variáveis (B para MUR e R para EBT). A Análise PCA também foi empregada utilizando todas as seis variáveis e dados centrados na média ou auto-escalados, sendo obtidos resultados muito similares aos apresentados na Figura 4 utilizando as duas primeiras componentes principais, que explicaram, ao menos, 95,50% da variância total dos dados. Com este estudo (Figura 1S) foi possível discriminar satisfatoriamente 9 classes de amostras, sendo que as amostras 6 e 7 formaram um único agrupamento, de forma muito semelhante ao demonstrado na Figura 4. A análise PCA também confirmou que as variáveis B para MUR e R para EBT contribuem de forma significativa para discriminação, uma vez que estas variáveis apresentaram os maiores pesos (*loadings*) para o cálculo das duas primeiras componentes principais nos estudos com PCA (Figura 1S).

Considerando estes resultados, pode-se afirmar que a estratégia proposta permitiu a análise exploratória dos dados sem a necessidade de uma ferramenta quimiométrica como a PCA, o que pode facilitar a interpretação dos resultados por usuários pouco ou não familiarizados com quimiometria.

Observando-se o gráfico na Figura 4 e os dados na Tabela 2 é possível associar a variável B do reagente MUR à concentração de íons cálcio nas amostras de água. Nota-se que para maiores valores

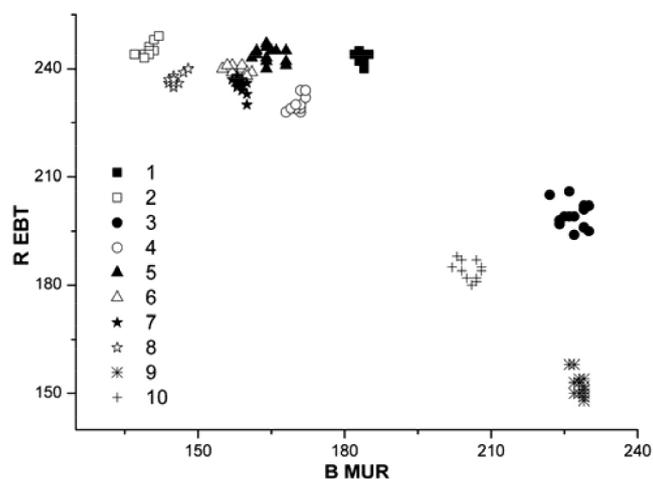
desta componente estão as águas com menores concentrações de íons cálcio, como a água mineral 3 e a água 9 (destilada), enquanto no outro extremo (menores valores na abscissa) encontram-se as amostras mais concentradas em cálcio. De fato, amostras que apresentam praticamente o mesmo teor de cálcio, como as amostras 6 e 7, tendem a exibir coordenadas em “x” muito semelhantes. Além disso, a literatura relata que o reagente MUR é um indicador mais seletivo a íons cálcio do que a íons magnésio, reforçando a relação desta variável com o teor de cálcio.<sup>21</sup>

Com relação a variável R do EBT (ordenada do gráfico de dispersão), observa-se que a mesma proporciona a separação de grupos de amostras que contém baixas concentrações de íons cálcio e magnésio (3,9 e 10) em relação àqueles grupos onde a concentração destes íons é relativamente maior (demais grupos). Uma hipótese plausível para tal constatação, é que, devido ao reagente EBT ser um bom indicador tanto para íons cálcio quanto para íons magnésio, os sinais obtidos nesta componente para as amostras mais concentradas estão muito próximas do nível de saturação (valor máximo igual a 255), o que possivelmente não favoreceu a discriminação das amostras com relação a este eixo.<sup>21</sup>

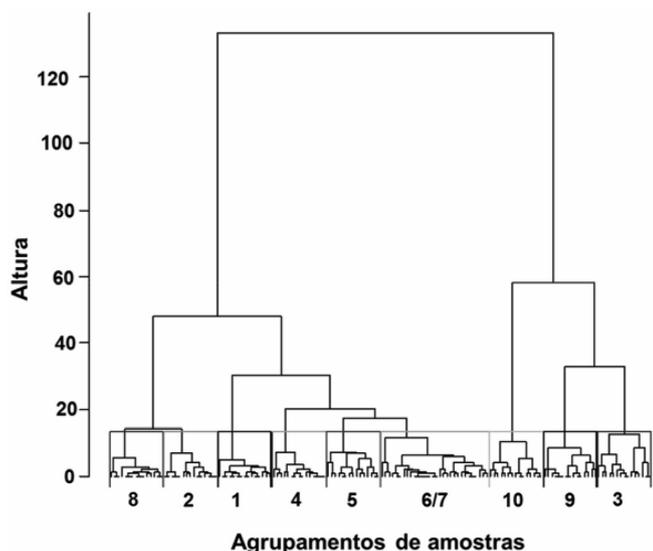
Tabela 3. Variância dos dados para as diferentes variáveis avaliadas

| Variáveis | Média  | Desvio Padrão | Coefficiente de variação (%) |
|-----------|--------|---------------|------------------------------|
| R – MUR   | 254,00 | 0,00          | 0,0                          |
| G – MUR   | 176,80 | 16,25         | 9,2                          |
| B – MUR   | 178,18 | 30,43         | 17,1                         |
| R – EBT   | 221,11 | 30,12         | 13,6                         |
| G – EBT   | 163,96 | 13,91         | 8,5                          |
| B – EBT   | 205,48 | 10,50         | 5,1                          |

A avaliação exploratória foi também realizada por meio da análise de agrupamentos hierárquicos (HCA) utilizando-se as duas variáveis que apresentaram maior variância (B para MUR e R para EBT) sem nenhum pré-tratamento dos dados.<sup>20</sup> O dendrograma apresentado na Figura 5 mostra a formação de nove agrupamentos, sendo oito destes correspondentes a amostras individuais e um agrupamento formado pelas amostras 6 e 7. Como pode ser observado, os dados referentes às amostras da classe 5 foram agrupados em uma classe diferente da classe constituída pelos dados das classes 6 e 7, o que não poderia ser afirmado com segurança ao se observar apenas a dispersão apresentada na Figura 4. Neste ponto, é importante ressaltar que os resultados obtidos através da HCA indicam a discriminação adequada das águas



**Figura 4.** Dispersão dos dados obtidos para as medidas em triplicata das amostras de águas avaliadas utilizando-se os dados para a componente R com EBT e para a componente B com MUR



**Figura 5.** Dendrograma obtido com as variáveis R para o EBT e B para MUR

estudadas, o que foi conseguido a partir de uma instrumentação alternativa de baixo custo e de um procedimento bastante simples.

Com o objetivo de verificar a qualidade das duas variáveis selecionadas na construção de um modelo para a classificação das águas estudadas, propôs-se o uso do método k-NN utilizando três amostras de cada fonte para a construção do modelo e duas amostras para a sua validação.<sup>20</sup> Vale destacar que as amostras utilizadas para a construção ou validação foram selecionadas de forma aleatória, por sorteio, sendo as replicatas autênticas das medidas para estas amostras (9 replicatas para calibração e 6 replicatas para validação) consideradas como amostras diferentes e independentes no estudo realizado. Adicionalmente, nenhum pré-processamento dos dados foi empregado para a construção do modelo e o número de vizinhos para verificação das distâncias foi otimizado a cinco, empregando-se para isto a estratégia de validação cruzada.<sup>20</sup> De maneira geral, o método k-NN prevê se uma determinada amostra pertence a uma classe específica através da distância desta amostra em relação a outras amostras (vizinhos mais próximos à amostra em questão) em um espaço n-dimensional definido, por exemplo, pelas variáveis experimentais consideradas na análise. Para o caso estudado, as distâncias de certa amostra em relação aos seus cinco vizinhos mais próximos são determinadas com relação ao espaço bidimensional

formado pelas variáveis B MUR (eixo x) e R EBT (eixo y). As distâncias determinadas para os cinco vizinhos mais próximos são então elencadas em ordem crescente e cada vizinho mais próximo fornece um voto para certa classe de amostras baseado nestas distâncias. No presente estudo, quando três ou mais vizinhos mais próximos atribuem o mesmo voto (classe) para a amostra esta é então classificada como pertencente àquela classe. Caso ocorra um empate entre duas classes (dois votos para duas classes distintas e um voto para uma terceira classe), é utilizado como critério de desempate soma das distâncias dos vizinhos nessas duas classes, sendo que a classe que apresentar a menor soma será a classe atribuída à amostra em questão.<sup>20</sup>

A literatura mostra que uma forma adequada de avaliação do modelo de calibração obtido por k-NN consiste em determinar a taxa de eficiência (TE) deste modelo para cada uma das classes avaliadas. Para isso, determinou-se o percentual de falsos positivos (FP) e falsos negativos (FN) para cada classe, sendo TE estimada a partir da equação 1.<sup>23</sup>

$$TE = 100 - (FP + FN)\% \quad (1)$$

Com base nesta equação, o modelo de calibração desenvolvido apresentou taxas de eficiência abaixo de 100% apenas para as fontes 6 (86,4%) e 7 (76,6%), o que era previsto se considerada a análise exploratória desenvolvida anteriormente. Além disso, as eficiências acima de 75% para estas classes e o fato de todas as outras oito classes apresentarem 100% de eficiência demonstraram uma modelagem adequada para a calibração.

A Tabela 4 lista os resultados obtidos para a validação do método k-NN. Novamente, foram observadas taxas de eficiência de 100% para oito das dez classes estudadas, sendo que a classe 7, uma das classes com eficiência menor, apresentou um índice de 94,4% para este quesito, o qual pode ser considerado elevado. Deste modo, pode-se afirmar que o método apresentou um desempenho adequado para a classificação das águas minerais avaliadas, uma vez que apenas a classe de amostras 6 apresentou uma baixa taxa de eficiência (50%). Deve-se salientar ainda que as amostras não classificadas corretamente na classe 6, foram classificadas como amostras da classe 7 e que estas classes apresentam concentrações dos íons metálicos majoritários (Tabela 2) bastante semelhantes, o que dificultou a distinção destas amostras pelo modelo. Considerando este aspecto, é possível esperar que a inclusão de novos reagentes colorimétricos seletivos a outros íons metálicos (minoritários) pode melhorar a classificação inclusive para estas classes (6 e 7). Entretanto, deve-se ter em mente que, a depender da concentração destes elementos na amostra, a resposta

**Tabela 4.** Resultados obtidos para a validação do modelo k-NN

| Classes de amostras | Verdadeiros Positivos (%) | Falsos Positivos (%) | Verdadeiros Negativos (%) | Falsos Negativos (%) | Taxa de Eficiência (%) |
|---------------------|---------------------------|----------------------|---------------------------|----------------------|------------------------|
| 1                   | 100                       | 0                    | 100                       | 0                    | 100                    |
| 2                   | 100                       | 0                    | 100                       | 0                    | 100                    |
| 3                   | 100                       | 0                    | 100                       | 0                    | 100                    |
| 4                   | 100                       | 0                    | 100                       | 0                    | 100                    |
| 5                   | 100                       | 0                    | 100                       | 0                    | 100                    |
| 6                   | 50                        | 0                    | 100                       | 50                   | 50                     |
| 7                   | 100                       | 5,6                  | 94,4                      | 0                    | 94,4                   |
| 8                   | 100                       | 0                    | 100                       | 0                    | 100                    |
| 9                   | 100                       | 0                    | 100                       | 0                    | 100                    |
| 10                  | 100                       | 0                    | 100                       | 0                    | 100                    |

colorimétrica obtida pode não resultar na otimização dos resultados, o que deve ser avaliado em trabalhos futuros.

## CONCLUSÃO

Os resultados obtidos permitem concluir que a classificação de águas minerais pode ser realizada de forma simples e eficiente empregando-se medidas colorimétricas obtidas a partir de um aplicativo para *smartphone* sem a necessidade de determinações quantitativas mais laboriosas.<sup>16,17</sup> A estratégia proposta neste trabalho poderá ser utilizada no futuro como ferramenta para a verificação preliminar de possíveis adulterações de águas minerais comerciais utilizando um modelo de classificação desenvolvido em um aplicativo para *smartphones*, facilitando ainda mais a sua aplicação. Adicionalmente, o presente trabalho poderá servir de base para o desenvolvimento de sensores colorimétricos específicos para a classificação de águas minerais.

## MATERIAL SUPLEMENTAR

Os resultados obtidos com a análise de componentes principais aplicada aos dados centrados na média são apresentados na Figura 1S. O gráfico de *scores* (Figura 1S (A)) indica a separação das amostras de forma similar ao obtido pela dispersão dos dados com relação às variáveis originais R-EBT e B-MUR. Com o gráfico de pesos (Figura 1S (B)) é possível notar que as variáveis R-EBT e B-MUR contribuem significativamente para o cálculo das duas primeiras componentes principais, o que corrobora para a seleção destas variáveis no estudo realizado. O leitor pode acessar livremente o conteúdo deste material no website (<http://quimicanova.sbq.org.br/>).

## AGRADECIMENTOS

Os autores agradecem ao Instituto Nacional de Ciências e Tecnologias Analíticas Avançadas (INCTAA) do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq – processo 573894/2008-6). O estudante Gerson F. Silva Neto agradece ao CNPq/PIBIC-UnB pela bolsa concedida.

## REFERÊNCIAS

1. Byrne, L.; Barker, J.; Pennarun-Thomas, G.; Diamond, D.; *Trends in Anal. Chem.* **2000**, *19*, 517.
2. Choodum, A.; Parabun, K.; Klawach, N.; Daeid, N. N.; Kanatharana, P.; Wongniramaikul, W.; *Talanta* **2013**, *115*, 143.
3. Gomes, M. S.; Trevizan, L. C.; Nóbrega, J. A.; Kamogawa, M. Y.; *Quim. Nova* **2008**, *31*, 1577.
4. Gaiao, E. N.; Martins, V. L. M.; Lyra, W. S.; Almeida, L. F.; Silva, E. C.; Araujo, M. C. U.; *Anal. Chim. Acta* **2006**, *570*, 283.
5. Askim, J. R.; Mahmoud, M.; Suslick, K. S.; *Chem. Soc. Rev* **2013**, *42*, 8649.
6. Salles, M. O.; Meloni, G. N.; Araujo, W. R.; Paixão, T. R. L. C.; *Anal. Methods* **2014**, *6*, 2047.
7. Bueno, L.; Meloni, G. N.; Reddy, S. M.; Paixão, T. R. L. C.; *RSC Adv.* **2015**, *5*, 20148.
8. Feng, L.; Musto, C. J.; Kemling, J. W.; Lim, S. H.; Suslick, K. S.; *Chem. Commun.* **2010**, *46*, 2037.
9. Feng, L.; Musto, C. J.; Kemling, J. W.; Lim, S. H.; Zhong, W.; Suslick, K. S.; *Anal. Chem.* **2010**, *82*, 9433.
10. Suslick, K. S.; Bailey, D. P.; Ingison, C. K.; Jazen, M.; Kosal, M. E.; McNamaraIII, W. B.; Rakow, N. A.; Sen, A.; Weaver, J. J.; Wilson, J. B.; Zhang, C.; Nakagaki, S.; *Quim. Nova* **2007**, *30*, 677.
11. Suslick, B. A.; Feng, L.; Suslick, K. S.; *Anal. Chem.* **2010**, *82*, 2067.
12. Oliveira, L. F.; Canevari, N. T.; Jesus, A. M. D.; Pereira-Filho, E. R.; *Revista Virtual de Química* **2013**, *5*, 169.
13. Damasceno, D.; Toledo, T. G.; Godinho, M. S.; Silva, C. P.; Oliveira, S. B.; Oliveira, A. E.; *Quim. Nova* **2015**, *38*, 836.
14. Lima, M. B.; Andrade, S. L. E.; Barreto, I. S.; Almeida, L. F.; Araújo, M. C. U.; *Microchem. J.* **2013**, *106*, 238.
15. [https://sistemas.dnpm.gov.br/publicacao/mostra\\_imagem.asp?IDBancoArquivoArquivo=8963](https://sistemas.dnpm.gov.br/publicacao/mostra_imagem.asp?IDBancoArquivoArquivo=8963), acessado Maio 2016.
16. Morgano, M. A.; Schatti, A. C.; Enriques, H. A.; Mantovani, D. M. B.; *Ciênc. Tecnol. Aliment.* **2002**, *22*, 239.
17. Correia, P. M. R.; Ferreira, M. M. C.; *Quim. Nova* **2007**, *30*, 481.
18. Jia, M. Y.; Wu, Q. S.; Li, H.; Zhang, Y.; Gan, Y. F.; Feng, L.; *Biosens. Bioelectron.* **2015**, *74*, 1029.
19. [https://play.google.com/store/apps/details?id=ecolor.com&hl=pt\\_BR](https://play.google.com/store/apps/details?id=ecolor.com&hl=pt_BR), acessado Maio 2016.
20. Ferreira, M. M. C.; *Quimiometria: Conceitos, Métodos e Aplicações*, Editora da Unicamp: Campinas, 2015.
21. Vogel; Mendham, J.; Barnes, J. D.; Denney, R. C.; Thomas, M. J. K.; *Análise Química Quantitativa*, 6ª edição, LTC: Rio de Janeiro, 2002.
22. Fonseca, A.; Raimundo Jr, I. M.; *Anal. Chim. Acta* **2007**, *596*, 66.
23. Botelho, B. R.; Reis, N.; Oliveira, L. S.; Sena, M. M.; *Food Chem.* **2015**, *181*, 31.