

Análise de águas minerais com e sem sabor com uma língua eletrónica potenciométrica

Zelda Manuel Menezes Alberto

*Dissertação apresentada à Escola Superior Agrária de Bragança
para obtenção do Grau de Mestre em Qualidade e Segurança
Alimentar*

Orientado por
António Manuel Coelho Lino Peres
Luís Avelino Guimarães Dias

**Bragança
2014**



Análise de águas minerais com e sem sabor com uma língua eletrónica potenciométrica

Zelda Manuel Menezes Alberto

*Dissertação apresentada à Escola Superior Agrária de Bragança
para obtenção do Grau de Mestre em Qualidade e Segurança
Alimentar*

Orientado por

**António Manuel Coelho Lino Peres
Luís Avelino Guimarães Dias**

**Bragança
2014**

Nome: Zelda Manuel Menezes Alberto

Orientador:

Prof. Doutor António M. Peres, Escola Superior Agrária- Instituto Politécnico de Bragança

Co-orientador:

Prof. Doutor Luís Guimarães Dias, Escola Superior Agrária- Instituto Politécnico de Bragança

DEDICATÓRIA

Em especial, aos meus queridos e amados pais
Manuel Alberto Gunza e Marcelina Menezes,
Pela dedicação, incentivo e amor inesgotável.

Às minhas irmãs, Cadete, Teresa, Otasília e
Guilhermina por serem amáveis e companheiras.

Aos meus sobrinhos Nodete, Delmara e Marcela,
pelo sorriso brilhante que me proporcionam.

Com Amor e Gratidão.

Eu dedico

Agradecimentos

É o momento de parar no tempo e relembrar o filme construído durante o caminho percorrido e refletir que alcancei apenas um degrau dentre várias outras vitórias que virão, sempre embasados na humildade, no amor e respeito ao próximo. Durante este período de formação conheci pessoas, culturas variadas, histórias marcantes, porém, tenho á agradecer;

À Deus por me conceder a graça de vivenciar este momento único e abençoado, e por trilhar sempre os meus caminhos dando-me oportunidade de crescer ao nível pessoal, científico e profissional.

À minha família, pelos ensinamentos, educação, amor incondicional, alegria, que de forma natural me encorajou a seguir os meus ideais me tornando numa pessoa segura e capaz de romper barreiras que possam surgir ao longo da vida e ser feliz o suficiente para dar significado a ela. Palavras me faltam para exprimir o quão tamanho é o sentimento que foi criado por vocês e que tornou-se num livro de apoio da vida.

Ao meu namorado Arsénio Sandando Sassussu Lucamba pela compreensão, paciência, amor incondicional, força que mesmo distante conseguiu proporcionar-me.

Ao meu orientador Professor Doutor António Peres, pelo acolhimento, encaminhamento, acompanhamento. Desde o início do trabalho que tive a perceção que estava em boas mãos e que tudo correria como o projetado, é sem dúvida um profissional excelente.

Ao meu co-orientador, Professor Doutor Luís Dias pela acreditação, acolhimento, incentivo, apoio, dedicação e sobre tudo a forma como trata com facilidade e simplicidade diversas situações sem perder o caracter, a educação que lhe são característico. É sem dúvida uma pessoa assertiva e um pedagogo por excelência.

Ao Eng. Sá Morais pelo acolhimento, simpatia, instrução, disponibilidade, ensinamentos e ajuda na realização das análises cromatográficas.

Ao Instituto Politécnico do Kuanza-Sul, Sumbe, liderado pelo Diretor Manuel Otávio Spinola, Professor Doutor Joel Ekulica, professores e funcionários das restantes áreas, pela credibilidade, oportunidade que me foi dada, para fazer parte deste projeto de formação científica.

Ao Instituto Nacional de Gestão de Bolsas de Estudo de Angola, pela autorização da formação, e apoio financeiro durante todo período concedido para a realização do curso.

Ao Instituto Politécnico de Bragança, na pessoa do Sr. Presidente João Sobrinho Teixeira, á Escola Agrária de Bragança na pessoa do Diretor Albino Bento, Professores e todos os funcionários de áreas diversas pela receção, acolhimento, formação, inserção em atividades de fórum educativo, cultural e social.

Aos estudantes Angolanos do mestrado do IPB, pela amizade, convivência, companheirismo nos bons e maus momentos, que de alguma forma simplificou a minha integração, a facilidade de vencer obstáculos, resumir a distância e simplificar saudades.

Aos colegas de curso e todos amigos que de forma direta ou indireta proporcionaram a amizade e momentos únicos.

Á todos que me apoiaram e estiveram comigo, o meu muito obrigado!

Resumo

O objetivo deste trabalho foi aplicar uma língua eletrônica potenciométrica na análise de águas minerais com e sem sabor adquiridas em várias superfícies comerciais, para averiguar se a informação obtida permitia fazer análise qualitativa e quantitativa. As águas minerais naturais e de nascente estão disponíveis comercialmente com uma grande variedade de paladares, associados à diversidade da composição físico-química. As águas minerais de sabor são conhecidas como refrigerantes por terem adição de ingredientes autorizados (ex., corantes, conservantes, aromatizantes, edulcorantes, acidulantes, antioxidantes, etc).

A língua eletrônica potenciométrica utilizada neste trabalho incluía um eletrodo de referência Ag/AgCl de dupla junção e dois sistemas de multi-sensores, conectados a um datalogger para aquisição dos sinais. Cada sistema de multi-sensores era constituído por 20 membranas lipídicas diferentes, de sensibilidade cruzada, preparadas com vários compostos lipídicos (3,0%), plastificantes (65,0%) e com o polímero PVC (32,0%). O segundo sistema era uma réplica do primeiro sistema de multi-sensores.

A caracterização físico-química das 34 amostras minerais recolhidas envolveu análises de pH e condutividade em todas as amostras e a análise por cromatografia líquida nas amostras de sabor (16 amostras), para determinar conteúdos em ácidos orgânicos (ácidos cítrico, málico e ascórbico) e açúcares (glucose, frutose e sacarose). Num primeiro estudo sobre os dados analíticos obtidos verificou-se que as amostras poderiam ser agrupadas em 4 grupos de acordo com a informação dos rótulos: águas minerais com e sem gás e águas de sabor com e sem gás. A análise de variâncias mostrou que havia diferenças significativas entre os quatros grupos considerando as variáveis da caracterização físico-química das amostras. A análise de componentes principais usando os parâmetros químicos obtidos por HPLC mostrou que os dados analíticos separavam naturalmente as amostras de águas de sabor com e sem gás.

Numa segunda fase, usaram-se os perfis de sinais obtidos com a língua eletrônica e, no âmbito da análise qualitativa, aplicou-se a análise dos componentes principais para verificar a distribuição espacial das amostras e a interdependência com os parâmetros físico-químicos determinados; posteriormente, aplicou-se análise discriminante linear acoplada ao algoritmo de arrefecimento simulado para seleção de

variáveis e estabeleceu-se o melhor modelo discriminante para os quatro grupos definidos. O desempenho do modelo obtido ao nível da precisão foi avaliado usando a validação cruzada “leave-one-out”. Verificou-se que com uma só função discriminante (explica 99,94% da variância total dos dados), contendo 24 sensores (12 sensores do primeiro sistema e 12 do segundo sistema), permitia obter 100% de classificações corretas, por validação cruzada. Neste estudo obteve-se também informação de que os sinais obtidos da língua eletrónica tinham informação sobre os valores de pH e condutividade e, por isso, estabeleceram-se modelos de regressão linear múltipla para prever resultados de pH e condutividade nas amostras. Como a regressão linear múltipla é sensível à multicolinearidade dos sinais dos sensores também, foi necessário seleccionar o melhor modelo através da seleção das variáveis independentes usando o algoritmo arrefecimento simulado. Os modelos foram testados na robustez de previsão através da validação cruzada com a técnica “leave-one-out”.

Os melhores modelos de previsão de pH e condutividade incluíam 27 sensores (13 sensores do primeiro sistema e 14 do segundo sistema) e 28 sensores (14 sensores do primeiro sistema e 14 do segundo sistema), respetivamente, mostrando que a informação estava presente nos perfis dos sinais. Verificou-se que na relação linear entre os valores previstos pelo modelo selecionado e os valores experimentais, quer para a variável pH quer para a condutividade, obtiveram-se valores de declive e ordenada na origem que, estatisticamente, podem ser considerados os teóricos (declive igual a 1 e ordenada na origem igual a zero).

Os resultados mostraram que a língua eletrónica pode ser uma ferramenta analítica prática na análise de águas minerais com e sem sabores, ao nível do controlo analítico de classificação ou discriminação das amostras, bem como, da análise de pH e condutividade.

Palavras-chave: língua eletrónica; águas minerais com e sem sabor; análise de componentes principais; análise discriminante linear; regressão linear múltipla; algoritmo de arrefecimento simulado.

Abstract

The objective of this study was to apply a potentiometric electronic tongue in the analysis of mineral waters with and without flavors acquired in various commercial areas, to ascertain whether the information obtained allowed qualitative and quantitative analysis. Mineral waters without flavors, natural or spring water (still and sparkling), are commercially available with a large variety of flavors associated with the diversity of physical and chemical composition. Mineral waters with flavors are allowed to have added ingredients (ex., dyes, preservatives, flavorings, sweeteners, acidulants, antioxidants, etc.).

The potentiometric electronic tongue used in this work included a reference electrode Ag/AgCl with a double-junction and two multi-sensor systems, connected to a datalogger for signal acquisition. Each multi-sensor system consisted of 20 different lipid membranes with cross-sensitivity prepared with various lipid compounds (3.0%), plasticizers (65.0%) and PVC polymer (32.0%). The second system was a replica of the first multi-sensor system.

The physico-chemical characterization of the 34 samples acquired involved analysis of pH and conductivity in all samples and HPLC analysis in the samples of flavors (16 samples) to determine contents in organic acids (citric, malic and ascorbic acids) and sugars (glucose, fructose and sucrose). As a first study on the data obtained it was found that the samples could be grouped into 4 groups according to information on labels: mineral waters and sparkling water and water with flavors with or without gas. The analysis of variance showed there were significant differences between the four groups for the variables of physico-chemical characterization of the samples. The principal component analysis using the chemical parameters obtained by HPLC showed that the analytical data naturally separated the samples of flavored water with and without gas.

In a second step, the signal profiles obtained from the electronic tongue were used, within the qualitative analysis, and principal component analysis was applied to verify the spatial distribution of the samples and the interdependence with certain physico-chemical parameters; subsequently, linear discriminant analysis coupled to simulated annealing algorithm for selection of variables was used to obtain the best

model to discriminate the 4 groups defined. The performance of the model at the precision level was evaluated using cross-validation "leave-one-out". It was found that with only one discriminant function (explain 99.46% of data total variance), having 21 sensors (12 sensors of the first system and 12 of the second system), allowed to get 100% of correct classification for cross-validation.

This study also allowed to verify that signals obtained from the electronic tongue had information about pH and conductivity values and, therefore, models were established using multiple linear regression to predict pH and conductivity values in the samples. As multiple linear regression is sensitive to multicollinearity of sensor signals, it was also necessary to select the best model through the selection of independent variables using the simulated annealing algorithm. The models were tested in the robustness of prediction by cross-validation with the "leave-one-out" technique. The best models for predicting pH and conductivity sensors included 27 (12 sensors of the first system) and 28 (14 sensors of the first system), respectively, showing that the information was present in the profiles of the signals sensors. It was found that the linear relation between the values provided by the selected model and the experimental values, both for the variables pH and conductivity, provided slope and intercept values that may statistically be regarded as the theoretical (slope equal to 1 and intercept equal to zero).

The results showed that the electronic tongue can be a practical analytical tool in the analysis of mineral water with or without flavors in the analytical control of samples by classification or discrimination sample, as well as the analysis of pH and conductivity of the samples.

Key-words: electronic tongue; mineral waters with and without flavors; principal component analysis; linear discriminant analysis; multiple linear regression; simulated annealing algorithm.

ÍNDICE GERAL

DEDICATÓRIA	ii
Agradecimentos	iii
Resumo	v
Abstract	vii
Índice de tabelas	xi
Índice de figuras	xii
Abreviaturas	xiv
Objetivos	xv
1. INTRODUÇÃO	1
1.1. Águas minerais naturais e de nascente	1
1.1.2. Composição das águas minerais e requisitos na sua composição.....	2
1.2. Águas minerais naturais e de nascente com sabor	3
1.2.1. Composição das águas com sabor	4
1.3. Controlo de qualidade de águas com e sem sabor	5
1.4. Língua eletrónica	6
1.4.2. Processamento de sinais	7
1.4.3. Aplicações das línguas eletrónicas.....	10
2. MATERIAL E MÉTODOS.....	11
2.1. Reagentes.....	11
2.2. Amostragem	11
2.3. Análise por HPLC	13
2.3.1. Preparação de soluções.....	14
2.3.2. Equipamento HPLC.....	15
2.3.3. Preparação das amostras	15
2.4. Medição do pH	16
2.5. Medição da condutividade	16
2.6 Língua eletrónica	18
2.8 Métodos estatísticos	19
2.8.1. Análise de componentes principais (ACP).....	19
2.8.2. Classificação com ADL	20
2.8.3 Quantificação com RLM.....	21
3. RESULTADOS E DISCUSSÃO	23
3.1. Análises físico-químicas	23
3.1.1. Medição do pH	23

3.1.2. Condutividade	24
3.2. Análise por HPLC	25
3.3. Análise exploratória dos dados químicos e físico-químicos	28
3.4 – Análise exploratória dos dados da análise com a língua eletrónica	30
3.5 – Análise qualitativa com a língua eletrónica.....	33
3.6 – Análise quantitativa com a língua eletrónica	36
4. CONCLUSÕES.....	40
BIBLIOGRAFIA	42
WEBSITES	45

Índice de tabelas

Tabela 1. Descrição das amostras e seus principais constituintes12

Tabela 2 - Concentrações médias obtidas da análise das amostras por HPLC27

Índice de figuras

Figura 1. Ultrassons Elma Transsonic 460 e sistema de filtração Phenomenex, acoplado a uma bomba de vácuo da Laboport, usando filtros de nylon (0,2 µm da Whatman).	14
Figura 2 – Milivoltímetro usado na análise do pH em águas minerais.	16
Figura 3 – Condutivímetro usado na análise da condutividade em águas minerais.	17
Figura 4 – Equipamento para a análise das águas minerais com a língua eletrónica: 1- PC para aquisição de dados; 2- DataLogger Agilent; 3- língua eletrónica; 4- Agitador magnético.	19
Figura 5 - Variação dos valores de pH e condutividade das amostras estudadas	25
Figura 6 – Análise de componentes principais (PC1 e PC2) dos dados químicos e físico-químicos obtidos da análise às amostras de sabor com gás (pontos de cor verde) ou sem gás (pontos de cor laranja). No gráfico A indica-se o sabor de cada amostras e no gráfico B, as respetivas marcas	29
Figura 7 - Perfil dos sinais obtidos da análise das amostras com a língua eletrónica.	30
Figura 8 - Extremos quartis dos sinais de potencial (mV) em função dos sensores.	31
Figura 9– Análise dos componentes principais aos dados da língua eletrónica. As cores dos pontos evidenciam 4 grupos: água mineral (pontos de cor verde), água mineral com gás (pontos de cor laranja), água de sabor com gás (pontos de cor roxo) e água de sabor sem gás (pontos de cor azul); gráficos A, B e C mostram como legendas dos pontos os sabor e os valores de pH e de condutividade, respetivamente.	32
Figura 10 - Percentagens de classificações corretas por análise discriminante linear para os dados originais (DOcorrecta) e para a validação cruzada “leave-one-out” (LOOcorrecta) de cada modelo selecionado pelo algoritmo ARRS.	35
Figura 11 - Análise discriminante linear entre 4 grupos de águas: água mineral sem gás (AM), água mineral com gás (AMgas), água de sabor sem gás (AS) e água de sabor com gás (ASgas). As elipses de cada grupo estão associadas ao intervalo de confiança a 95%, assumindo uma distribuição normal multivariada.	35
Figura 12 - Valores de RMSE e R ² de cada modelo de regressão linear múltipla usando subconjunto de sensores selecionados pelo algoritmo arrefecimento simulado para prever valores de pH.	37

Figura 13 - Valores de pH previstos pelo modelo regressão linear múltipla selecionado, quer para os dados originais quer os da validação cruzada, em função dos valores experimentais. 37

Figura 14 - Valores de RMSE e R^2 de cada modelo de regressão linear múltipla usando subconjunto de sensores selecionados pelo algoritmo arrefecimento simulado para prever valores de condutividade. 39

Figura 15 - Valores de condutividade previstos pelo modelo regressão linear múltipla selecionado, quer para os dados originais quer os da validação cruzada, em função dos valores experimentais. 39

Abreviaturas

ACP - Análises de Componentes Principais

ADL - Análise Discriminante Linear

Ag/AgCl – Prata/Cloreto de prata

AM - Água mineral natural/ nascente sem gás

AMgas - Água mineral natural/ nascente com gás

ANOVA - Análise de Variância

AS - Água mineral natural de sabor sem gás

ARRS – Arrefecimento Simulado

ASgas - Água mineral natural/nascente de sabor com gás

CP - Componentes Principais

d - detetado mas não quantificado

HPLC - Cromatografia líquida de alto desempenho (abreviatura do inglês, *high performance liquid chromatography*)

IR - Índice de Refração

LE – Língua Eletrónica

Nd - Não detetado

PVC - Cloreto de polivinílo (abreviatura do inglês “Polyvinyl Chloride”)

RL - Regressão Linear

RLM - Regressão Linear Múltipla

UV – Ultravioleta

LOO – “deixar um de fora” (abreviatura do inglês, leave-one-out)

Objetivos

Neste trabalho pretendeu-se estudar o desempenho de uma língua eletrónica potenciométrica na classificação de águas minerais comerciais (potável e mineral) com e sem sabor. Como objetivos específicos pretendeu-se verificar:

- como as amostras de água mineral com e sem sabor poderiam ser agrupadas considerando a sumária caracterização físico-química;
- como se comportam os perfis de sinais dos sensores da língua eletrónica em função da caracterização físico-química;
- se a língua eletrónica permitia discriminar as amostras (análise qualitativa);
- se era possível definir modelos de quantificação com a língua eletrónica considerando as variáveis da caracterização físico-química.

1. INTRODUÇÃO

As águas minerais e os refrigerantes são as bebidas mais consumidas no mundo e existe uma enorme variedade de composição e marcas (Steen e Ashurst, 2006).

Em Portugal o consumo de águas minerais com e sem sabor tem aumentado ao longo dos últimos anos, associado também ao aparecimento de novos produtos com diferentes sabores, conquistando um lugar de destaque no sector de bebidas. Ao nível das vendas de águas minerais naturais e de nascente (mercado nacional e exportações), a estatística da APIAM (Associação Portuguesa das Indústrias de Águas Minerais Naturais e de Nascente) mostrou uma evolução positiva de cerca de 14,5% (em volume) em 2010, com uma estagnação em 2011, uma evolução negativa acentuada em 2012 e um resultado positivo de cerca de 7% (litros) em 2013 (W1). Os refrigerantes (valores atualizados em Setembro de 2014) evidenciaram uma evolução positiva de cerca de 4,2% no último ano. No mundo, as bebidas refrescantes não alcoólicas correspondem a cerca de 49% do total do mercado de bebidas comerciais (4.541,9 milhões de litros) (W2).

1.1. Águas minerais naturais e de nascente

A variedade da composição físico-química das águas minerais naturais e de nascente, conferindo sabor único, deve-se à origem e ao percurso geológico das águas. O DL nº 156/98 de 6 de junho define e caracteriza as águas minerais em vários tipos:

- **Água mineral natural** é uma “água de circulação subterrânea, considerada bacteriologicamente própria, com características físico-químicas estáveis na origem, dentro da gama de flutuações naturais, de que podem eventualmente resultar efeitos favoráveis à saúde e que se distingue da água de beber comum, pela sua pureza original, pela sua natureza caracterizada pelo teor de substâncias minerais, oligoelementos ou outros constituintes”;

- **Água mineral natural efervescente**, a “água que liberta espontaneamente e de forma perceptível gás carbónico nas condições normais de temperatura e de pressão, quer na origem quer após engarrafamento”. Neste grupo consideram-se três categorias em função dos teores em gás carbónico: **água mineral**

natural gasosa, se o teor em gás carbónico proveniente do aquífero após decantação eventual e engarrafamento é o mesmo que à saída da captação; **água mineral natural reforçada com gás carbónico natural**, se o teor em gás carbónico proveniente do mesmo aquífero, após decantação eventual e engarrafamento, é superior ao verificado à saída da captação; e **água mineral natural gaseificada**, se a adição de gás carbónico foi de uma origem que não seja o aquífero donde a água provém.

- **Água de nascente**, a “água subterrânea, considerada bacteriologicamente própria, com características físico-químicas que a tornam adequada para o consumo humano no seu estado natural”.

1.1.2. Composição das águas minerais e requisitos na sua composição

Em geral a variação na composição dos diferentes tipos de água é ao nível dos seus componentes maioritários (bicarbonatos, sódio, cálcio, etc.) e também nos oligoelementos (como o ferro, iodo, flúor, cobre, zinco, fósforo e outros minerais, a níveis vestigiais) (W3). Os oligoelementos são elementos químicos inorgânicos que são necessários aos seres humanos em pequenas quantidades e que desempenham diversas funções metabólicas no organismo. Globalmente, estes compostos encontram-se em cada tipo de água com diferentes níveis de concentração, tornando-a diferente de todas as outras, conferindo-lhe paladares únicos.

As águas minerais naturais e de nascente são caracterizadas ao nível geológico e hidrogeológico, físico, químico, microbiológico e, se necessário, farmacológico, fisiológico e clínico (DL nº 156/98 de 6 de junho), “não podem apresentar nenhum defeito do ponto de vista organoléptico, nem qualquer tratamento de desinfeção ou adição além da separação de elementos instáveis tais como componentes de ferro e enxofre; separação do arsénio e dos compostos de ferro, manganês e de enxofre; eliminação total ou parcial do gás carbónico livre por processos exclusivamente físicos; incorporação ou reincorporação do gás carbónico legalmente”. Os tratamentos referidos não podem ter qualquer efeito de alteração da composição dessas águas de forma a manter os constituintes essenciais e, conseqüentemente, as suas propriedades.

1.2. Águas minerais naturais e de nascente com sabor

As águas com sabor são conhecidas como bebidas refrigerantes ou simplesmente refrigerante. São classificadas como um género alimentício constituído principalmente por água contendo em solução, emulsão ou suspensão ingredientes autorizados (ex., corantes, conservantes, aromatizantes, edulcorantes, acidulantes, antioxidantes, etc) e, eventualmente gaseificado com dióxido de carbono. As bebidas refrigerantes são classificadas em vários tipos (DL nº93/89 de 28 de Março) e, considerando as amostras usadas neste estudo, salientam-se os seguintes tipos de refrigerantes:

- **refrigerante de extratos vegetais** – “refrigerante turvo ou límpido resultante da diluição em água de extratos e aromatizantes podendo eventualmente incluir sumo, polme ou respetivos derivados e ainda outros ingredientes comestíveis de origem vegetal”;

- **refrigerante aromatizado** – “refrigerante límpido resultante da diluição em água de aromatizantes e, eventualmente, adicionados de açúcares. No caso de não conter açúcares, designar-se-á por refrigerante «água aromatizada»”;

- **refrigerante de sumo de frutos** – “refrigerante turvo ou límpido, resultante da diluição, em água de sumo ou polme de frutos, respetivos concentrados ou desidratados, com um teor de sumo compreendido entre os limites mínimos indicados (m/m) e a concentração mínima fixada para o néctar do mesmo fruto, podendo conter aromatizantes naturais provenientes das espécies de frutos dos quais é extraído o sumo”.

As matérias-primas usadas na produção de refrigerantes são, em geral, à base de sumo concentrado, um produto obtido do sumo por eliminação física de pelo menos, 50% da água de constituição, permitindo atribuir sabor característico às bebidas; xarope de sumo, um produto pastoso obtido de solução aquosa dos açúcares com concentração adequada adicionada de sumo, sumo concentrado, polme concentrado, néctar ou uma mistura de quaisquer destes produtos; extrato vegetal, produto com efeito aromatizante ou sávido, obtido a partir de matérias vegetais, podendo ser sujeito a processos de concentração (DL nº93/89 de 28 de Março).

1.2.1. Composição das águas com sabor.

As águas minerais naturais e de nascente para se tornarem águas de sabor, possuem na sua composição aditivos (com/sem valor nutritivo) que são substâncias incorporadas nos alimentos durante o fabrico, transformação, preparação, tratamento, acondicionamento, transporte ou armazenamento de um produto alimentar. Embora estes aditivos não sejam normalmente consumidos como alimentos, nem utilizados como ingredientes típicos dos alimentos, são adicionados para satisfazerem funções tecnológicas como melhorar a estabilidade, propriedades organoléticas ou nutritiva dos alimentos (Ashurst, 2005).

Ao nível dos constituintes das bebidas com sabor, a água potável (pureza reconhecida) é o principal constituinte, podendo ser mineral desde que satisfaça as condições legais. O dióxido de carbono é utilizado para a carbonatação das bebidas, permitindo realçar o paladar e dar ação refrescante ao produto (Steen e Ashurst, 2006). Como segundo ingrediente maioritário, os edulcorantes concedem sabor adocicado às bebidas, usando-se produtos naturais (sacarose, xarope de açúcares, xarope de derivado de milho e mel) ou produtos artificiais (aspartame, E951; acesulfame k, E950; e o ciclutamato de sódio, E952) (Ashurst, 2005; Lima e Afonso, 2009). Os corantes são usados para dar ou restaurar (durante o processamento da bebida pode ocorrer a destruição dos corantes naturais) a cor dos refrigerantes. O beta caroteno é um dos corantes naturais mais usados, possui cor amarelo/laranja (E160a) e a tartrazina (E102), corante artificial, proporciona cor amarelo/limão (Ashurst, 2005). Os conservantes permitem inibir, retardar ou deter o crescimento de microrganismos ou qualquer deterioração dos alimentos, como por exemplo o ácido benzoico, ácido málico e sórbico. Os aromatizantes proporcionam uma identidade genérica, mas também o seu carácter único às bebidas, em parte responsável por atrair a atenção do consumidor. Algumas das substâncias aromáticas são o benzaldeído, butirato de amilo, aldeído cinâmico, dentre outros (Ashurst, 2005; Steen e Ashurst, 2006).

Neste trabalho foram analisados alguns aditivos tais como: antioxidantes, acidulantes e edulcorantes. Relativamente aos acidulantes e antioxidantes analisaram-se os seguintes: o ácido cítrico, principal constituinte de vários citrinos (acidulante dos

mais usados) como o limão que contém quantidades entre 5-8% (Ashurst, 2005; Yúfera, 1998). Atualmente é obtido por ação enzimática a partir da glucose e outros açúcares, pode ser adquirido como um sólido cristalino branco na forma de pó granular no seu estado anidro ou na forma monohidratada. O ácido málico é o principal ácido das maçãs e o segundo ácido, após o cítrico, em frutas cítricas, estando presente nas mais variadas frutas. É um ácido um pouco mais forte em acidez do que o ácido cítrico e pode ser obtido como um sólido branco cristalino que é altamente solúvel em água (Ashurst, 2005; Yúfera, 1998). O ácido ascórbico (vitamina C) é um poderoso antioxidante que melhora a estabilidade dos aromas das bebidas e o tempo de vida de prateleira, sendo também utilizado como acidulante (Ashurst, 2005; Yúfera, 1998).

Os edulcorantes analisados foram a sacarose, glucose e frutose. A sacarose é o açúcar mais abundante na natureza, um dissacarídeo obtido a partir da cana-de-açúcar ou de beterraba e está disponível em forma de cristal. A glucose e a frutose são conhecidos como açúcares das frutas, estando a glucose em maior quantidade relativamente à frutose, encontram-se em frutas como: maçã, pêra, marmelo, nêspira, entre outras (Ashurst, 2005; Yúfera, 1998).

1.3. Controlo de qualidade de águas com e sem sabor

Um alimento deve ter qualidade definida e sem qualquer risco para a saúde do consumidor e, por isso, em todas as fases da sua produção devem-se adotar procedimentos corretos e ingredientes selecionados para garantir a qualidade e segurança e conformidade dos produtos (Anexo III do artigo 2º, DL nº 93/89 de 28 de Março; Steen e Ashurst, 2006).

O controlo de qualidade está associado a várias técnicas e métodos analíticos para verificar possíveis adulterações e/ou contaminações que possam surgir nos alimentos. Os métodos analíticos para o controlo da água são cada vez mais sofisticados e usados correntemente na análise de águas, para verificação da sua qualidade para o consumo humano. Em trabalhos anteriores recorreu-se ao uso da potenciometria, condutimetria, volumetria, espectrofotometria de absorção molecular, técnica da filtração em membrana, língua eletrónica e técnicas microbiológicas (Resende, 2006; Parada, 2012).

Geralmente, as adulterações nos refrigerantes de sumos tem-se baseado na diluição excessiva (Ashurst, 2005) e adição de açúcares (Ashurst, 2005; Silva et al., 1999; Vlasov et al., 2005), etc. As águas de sabor usadas neste trabalho não são à base da adição de derivados de frutas e, por isso, este tipo de adulterações não se colocam. O controlo de qualidade deste tipo de bebidas é ao nível da conformidade dos produtos.

1.4. Língua eletrónica

A língua eletrónica (LE) é um sistema de multisensores constituído por sensores químicos com resposta reprodutível, de elevada estabilidade e sensibilidade cruzada a diferentes espécies em solução (não específicos e de baixa seletividade) e que, por isso, fornecem perfis de sinais que traduzem informação global sobre as amostras (Vlasov *et al.*, 2005). Métodos estatísticos multivariados são usados no processamento dos perfis de sinais para análise qualitativa, ao nível da identificação, classificação e discriminação de amostras, e análise quantitativa de multi-espécies químicas e a avaliação do sabor das amostras (Peres et al., 2011). As LEs são sistemas que permitem adquirir informação global sobre amostras complexas que não é possível com outras metodologias analíticas, principalmente mimetizar a língua humana na apreciação de sabor (Dias *et al.*, 2011). De salientar que a capacidade analítica está dependente das características da LE usada (Vlasov *et al.*, 2005).

1.4.1. Sensores químicos

Os sensores dependem do princípio de deteção da LE que pode ser: potenciometria (Toko, 2000; Legin *et al.*, 2003), voltametria (Heyden *et al.*, 2002; Wei *et al.*, 2013), amperometria (Burattia *et al.*, 2004) e impedância eletroquímica (Ulloa *et al.*, 2013).

Neste trabalho usou-se uma LE potenciométrica que apresenta diversas vantagens: pode ser construída no laboratório (como neste trabalho), tem grande durabilidade, permite uma vasta seleção de reagentes na preparação dos sensores e é de fácil aplicação no sistema de multi-sensores. É um sistema analítico mais simples e

de fácil utilização, por exemplo, basta uma simples lavagem entre amostras para se proceder à análise seguinte.

Vários tipos de sensores têm sido usados, sendo os multisensores constituídos só com um mesmo tipo de família de sensores ou com misturas de vários tipos, como por exemplo: sensores de vidro de calcogenetos (Legin *et al.*, 2003; Rudnistskaya *et al.*, 2009), membranas poliméricas lipídicas de sensibilidade cruzada (Dias *et al.*, 2008; Eckerta *et al.*, 2011; Legin *et al.*, 2003) e membranas seletivas a iões (Gallardo *et al.*, 2005; Legin *et al.*, 2003). Alguns destes tipos de sensores são seletivos, sendo usados na LE por permitirem uma diferenciação de amostras por meio de parâmetros químicos específicos. Mas, o critério de seleção de um sensor para um sistema de multisensores que se pretende usar como LE, é ter baixa seletividade e alta sensibilidade cruzada (capacidade de um sensor responder a um número diferente de espécies em solução). Uma LE com diferentes sensores de sensibilidade cruzada vai dar sinais diferentes para a matriz da amostra em análise, que constituem um perfil de sinais tipo “espectro” da amostra, que corresponde a uma impressão digital de multi-componentes da amostra, permitindo o uso de ferramentas multivariadas de estatística para análise qualitativa e quantitativa. O número de sensores num sistema de multisensores depende do trabalho que se pretende efetuar, tendo-se usado neste trabalho 40 sensores de membrana polimérica lipídicas (20 sensores diferentes na sua composição, em duplicado), de sensibilidade cruzada.

Os sistemas com um elevado número de sensores permitem um maior conjunto de aplicações de análise qualitativa e quantitativa, usando técnicas de seleção de variáveis acoplados às metodologias de análise estatística multivariada, pois possibilitam selecionar os sensores mais adequados para cada trabalho.

1.4.2. *Processamento de sinais*

O processamento dos perfis de sinais pode ser efetuado usando métodos estatísticos multivariados não supervisionados ou supervisionados.

Os métodos não supervisionados são métodos exploratórios, sem qualquer controlo do investigador, que permitem verificar a variabilidade natural existente nas amostras analisadas com o objetivo de procurar algum padrão ou estrutura entre

amostras. A técnica mais usada é a análise de componentes principais, ACP (Dias *et al.*, 2008; Escuder-Gilaberta *et al.*, 2010).

Para se aplicar os métodos supervisionados é necessário fornecer informação inicial sobre as amostras para que seja possível estabelecer modelos de classificação (análise qualitativa) e de calibração (análise quantitativa) com os sinais da LE. Dentro desta categoria, as metodologias mais usadas e simples para classificação é a análise discriminante linear (ADL), o método dos mínimos quadrados parciais (MQP-AD) e as redes neurais artificiais (RNA); para a análise quantitativa, os mais aplicados são a regressão linear múltipla (RLM) e o método dos mínimos quadrados parciais (MQP) (Escuder-Gilaberta *et al.*, 2010; Dias *et al.*, 2011).

Este trabalho incidiu na classificação de amostras através da ADL e na análise quantitativa, em calibrações com recurso à RLM. A ADL e RLM são técnicas multivariadas supervisionadas que visam a obtenção de modelos matemáticos lineares das variáveis independentes (sinais dos sensores) que melhor separem os diferentes grupos (variável dependente) em que se pretendem classificar as amostras ou quantifiquem o parâmetro físico-químico da amostra, respetivamente.

A aplicação destas duas metodologias implicaram a execução de 2 passos: primeiro, o estudo de seleção das variáveis independentes (sensores) para obter o melhor modelo; depois, o estudo da robustez do modelo selecionado associado à classificação ou quantificação das amostras.

Na **seleção de variáveis** (redução) podem-se aplicar vários métodos que se diferenciam na forma como selecionam e testam as variáveis. As metodologias clássicas, presentes na maior parte dos programas de estatísticas comerciais, são chamados de métodos heurísticos (métodos de seleção de um subconjunto de variáveis) incluídos na classe chamada *greedy*, porque são métodos que fazem uma escolha no momento mas, mais tarde, não verificam se foi uma boa opção (Graham, 2001). Como exemplo, temos os métodos de seleção sequencial para a frente (*sequential forward selection method*), de seleção sequencial por eliminação posterior (*sequential backward elimination method*) e seleção por etapas (*stepwise method*). Estes métodos são importantes quando no estudo existem poucas variáveis e se pretende ordenar as variáveis selecionadas em termos da sua importância. Mas, no caso da LE

usada neste trabalho, utilizam-se 40 sensores que mostram multicolinearidade e, por isso, os métodos heurísticos nem sempre escolhem o melhor conjunto de variáveis independentes. Nesta situação, os métodos de seleção de variáveis meta-heurísticos permitem obter o melhor subconjunto de variáveis que correspondem a um ótimo global para um dado critério de aproximação a uma determinada função, num grande espaço de busca de outros subconjuntos de variáveis. Dois exemplos destes métodos são o algoritmo genético e o algoritmo arrefecimento simulado (ARRS) (usado neste trabalho). No algoritmo ARRS, a melhor solução é obtida num espaço selecionado aleatoriamente na vizinhança da solução corrente e testada segundo as regras de um processo de arrefecimento simulado (consiste em aquecer um metal até à sua fusão seguido, de uma lenta diminuição da temperatura para que a substância atinja a configuração mais regular possível).

O método meta-heurístico de seleção de variáveis permite ter modelos com menor número de variáveis (elimina variáveis irrelevantes ou que contribuem da mesma forma para o estudo), aumentando a exatidão de previsão. O modelo será mais simples, mais fácil de interpretar e, por se conhecerem as variáveis corretas, mais rápido nos cálculos, permitindo melhores resultados de previsão.

O **estudo da robustez** do modelo ADL é efetuado testando a sua capacidade de previsão, mesmo quando parte dos pressupostos requeridos para a aplicação desta técnica não sejam observados: normalidade nas variáveis, homogeneidade das matrizes de variância-covariância nos grupos e diferenças significativas entre os grupos. Se o modelo ADL for robusto na classificação, então um modelo pode ser usado na previsão do grupo onde se inserem novas amostras (McGarigal et al., 2000). Devido ao pequeno número de amostras presentes no trabalho, o estudo da previsão baseia-se principalmente nos resultados de classificação corretos obtidos com a validação cruzada (VC) usando a técnica “deixar um de fora” (LOO, leave-one-out). Na técnica VC-LOO remove-se uma amostra do estudo (contém n amostras), obtendo-se o modelo ADL com as restantes que depois é usado para prever a classificação da amostra removida. O processo é repetido n vezes, onde se remove sempre uma amostra diferente, o que permite obter os erros de classificação de todas as amostras e o erro global (soma dos erros de cada teste dividido por n). Na análise quantitativa usando RLM, a capacidade

de previsão do modelo é verificada usando também a VC-LOO. Os resultados da previsão quantitativa são avaliados por comparação com os dados esperados de cada amostra que são os experimentais. A comparação é visualizada numa relação linear entre os valores previstos e os valores esperados, cujo declive e ordenada na origem se esperam ter valores um e zero, respetivamente.

1.4.3. Aplicações das línguas eletrónicas

O desenvolvimento do sensor gustativo (LE) representa um grande avanço no controle de qualidade para as indústrias alimentares, podendo ainda o seu uso ser estendido a outros sectores como, por exemplo, farmacêutico e cosmético. Em comparação aos métodos convencionais de análise, os resultados obtidos com o sensor gustativo são mais económicos e rápidos.

As LEs potenciométricas têm sido usados em análise qualitativa, semi-quantitativa e quantitativa. Por exemplo, na distinção de bebidas (Dias *et al.*, 2011; Toko, 2000), alimentos (Beullens *et al.*, 2010; Peres *et al.*, 2011), classificação de vinhos (Legin *et al.*, 2003), identificação de adulterações no leite (Dias *et al.*, 2009), quantificação de sabores e alguns compostos (Toko, 2000; Legin *et al.*, 2000), classificação de cervejas (Ciosek *et al.*, 2006), discriminação do mel em diferentes grupos de méis florais (Dias *et al.*, 2013), análise semi-quantitativa de sumos e quantitativa de açúcares e ácidos orgânicos em sumos comerciais (Sequeira, 2012).

Em relação a trabalhos recentes com águas naturais e de consumo podem-se salientar alguns exemplos de aplicação da LE. Uma LE contendo elétrodos seletivos a iões no estado todo-sólido foi usada na classificação de águas comerciais, bebidas à base de laranja e amostras de chá (Gallardo *et al.*, 2005); na análise qualitativa de águas naturais usou-se uma LE preparada com fios metálicos e superfícies condutoras baseadas na tecnologia de filme espesso (Martínez-Máñez *et al.*, 2005); uma LE usando elétrodos de película de pastas de metais e óxidos de metal diferentes, permitiu a classificação das águas potáveis (Garcia-Breijoa *et al.*, 2011).

Neste trabalho aplicou-se uma LE potenciométrica de membranas poliméricas lipídicas, de sensibilidade cruzada, na análise de águas minerais com e sem gás e águas de sabor com e sem gás.

2. MATERIAL E MÉTODOS

2.1. Reagentes

As soluções utilizadas neste trabalho foram preparadas com água desionizada de qualidade do tipo II (condutividade de 10 MΩ.cm). Para análise cromatográfica foi utilizado o eluente preparado com ácido orto-fosfórico da Panreac. Para a calibração do HPLC foram usadas soluções padrão, preparadas com os seguintes reagentes: ácido cítrico mono-hidratado (Fisher), ácido málico (Merck), ácido ascórbico (panreac), sacarose (panreac), glucose (Fluka) e frutose (Fluka).

As amostras foram analisadas usando vários procedimentos, a seguir descritos; análise por HPLC, medição do pH, medição da condutividade, para relacionar com a informação obtida da análise da língua eletrónica.

2.2. Amostragem

Neste trabalho estudou-se o desempenho de uma língua eletrónica na análise de 34 amostras de águas minerais naturais e de nascente com e sem sabor, adquiridas em superfícies comerciais da cidade de Bragança (Portugal). As amostras foram mantidas à temperatura ambiente (20 °C) até serem analisadas e classificadas em quatro grupos segundo a informação presente no rótulo: água mineral natural/nascente com gás (AMgas, 12 amostras), água mineral natural/nascente sem gás (AM, 6 amostras), água mineral natural de sabor sem gás (AS, 8 amostras) e água mineral natural/nascente de sabor com gás (ASgas, 8 amostras). Três amostras não são de origem Portuguesa.

Na tabela 1 apresenta-se a indicação dos principais componentes das amostras analisadas, o tipo de bebida, bem como o número e o código de identificação atribuído.

Tabela 1. Descrição das amostras e seus principais constituintes.

Marcas	Principais Componentes	Tipo	Amostr a	Grupo
A	Água mineral natural gasocarbónica, concentrado de fruta, sumo de framboesa, maçã e pera (2,3%) à base de concentrado, ácido cítrico.	Refrigerante	1	ASgas
A	Água mineral natural gasocarbónica, concentrado de sumo de limão e maçã (2,3%) à base de concentrado, ácido cítrico.	Refrigerante	2	ASgas
B	Água mineral natural frize (98,5%), sumo de maçã à base de concentrado, ácido cítrico.	Refrigerante	3	ASgas
B	Água mineral natural frize (97,5%) sumo de maçã e groselha (1%) à base de concentrado, ácido cítrico.	Refrigerante	4	ASgas
B	Água mineral natural frize (95%), açúcar, dióxido de carbono, corante e aroma, ácido cítrico.	Refrigerante	5	ASgas
C	Água de nascente arrieiro carbonatada, aroma natural de limão e lima, sumo de maçã e limão (0,5%) à base de concentrado, ácido cítrico e ascórbico.	Refrigerante	6	ASgas
D	Água mineral natural gaseificada castello (99,9%) e extrato natural de limão (0,1%).	Refrigerante	7	ASgas
D	Água mineral natural gaseificada (99,9%) e extratos naturais de lima (0,1%).	Refrigerante	8	ASgas
E	Água mineral natural de luso (80,7%) sumo de fruta concentrado: 11,5% maçã e 0,5% limão, extrato natural de fruta, ácido cítrico.	Refrigerante	9	AS
E	Água mineral natural de luso (80%), extrato natural de fruta, concentrado de água de coco (12%), ácido cítrico.	Refrigerante	10	AS
E	Água mineral natural de luso (81,1%), sumo de fruta concentrado: (8%) maçã, (1,5%) uva, (1,5%) cenoura púrpura, (1%) morango, ácido cítrico.	Refrigerante	11	AS
E	Água mineral natural de luso 80%, sumo de maçã concentrado (15%), extrato natural de fruta, aroma natural de maçã c/outros aromas naturais, xarope de caramelo, ácido cítrico.	Refrigerante	12	AS
F	Água mineral natural vitalis (97%), extratos de frutas, dextrina de trigo (0,9%), ácido cítrico.	Refrigerante	13	AS
F	Água mineral natural vitalis, xarope de glucose, frutose, extrato de fruta, sumo concentrado de limão e maçã (0,1%), ácido málico.	Refrigerante	14	AS
F	Água mineral natural vitalis, xarope de frutose, acidificante, sumo de limão e maçã (0,1%) a base de concentrado, ácido cítrico.	Refrigerante	15	AS

Tabela 1. Descrição das amostras e seus constituintes (continuação).

Marcas	Principais Componentes	Tipo	Amostr a	Grupo
B	Água mineral natural reforçada com gás carbónico natural	Água	16	AMgas
D	Água mineral natural gaseificada.	Água	17	AMgas
D	Água mineral natural gaseificada.	Água	18	AMgas
G	Água mineral natural reforçada c/gás carbónico natural.	Água	19	AMgas
H	Água mineral natural gasocarbónica.	Água	20	AMgas
I	Água de nascente gaseificada.	Água	21	AMgas
J	Água mineral natural gasocarbónica.	Água	22	AMgas
A	Água mineral natural gaseificada.	Água	23	AMgas
K	Água da nascente gaseificada.	Água	24	AMgas
L	Água mineral natural gaseificada	Água	25	AMgas
M	Água mineral natural gaseificada	Água	26	AMgas
N	Água da nascente natural.	Água	27	AM
F	Água mineral natural.	Água	28	AM
O	Água mineral natural.	Água	29	AM
P	Água de nascente.	Água	30	AM
E	Água mineral natural.	Água	31	AM
Q	Água mineral natural.	Água	32	AM
R	Água mineral natural gaseificada	Água	33	AMgas
E	Água mineral natural de luso 83,4%, sumo de fruta concentrado 10,8% maçã e 0,3% laranja, extrato natural de fruta, ácido cítrico.	Refrigerante	34	AS

Fonte: Informação de rotulagem.

2.3. Análise por HPLC

A cromatografia líquida de alto desempenho (HPLC) foi usada neste trabalho para identificar e quantificar as concentrações dos açúcares e ácidos orgânicos nas amostras.

2.3.1. Preparação de soluções

O eluente usado para o HPLC foi o ácido orto-fosfórico (Fisher Scientific) a 0,1% mantido a 35 °C durante as análises. Todos os eluentes preparados foram previamente filtrados com um sistema de filtração da Phenomenex, acoplado a uma bomba de vácuo da Laboport, usando filtros de nylon (0,2 µm da Whatman) (Figura 1).

Os tempos de retenção dos ácidos orgânicos (ácidos cítrico, málico e ascórbico) e dos açúcares (sacarose, frutose e glucose) foram determinados após injeção de soluções independentes com concentrações de 1,0g/L. Prepararam-se ainda várias soluções padrão para proceder à calibração. As concentrações variaram entre 0,3 e 5,0 g/L para a glucose, frutose, sacarose, ácido cítrico e ácido málico. Para o ácido ascórbico, as concentrações variaram entre 0,2 e 1,0 g/L, de acordo com Sequeira (2012). Para o estudo de precisão e exatidão preparou-se uma solução de controlo de qualidade por mistura de todos compostos num balão de 200 mL: sacarose (0,5076 g), frutose (0,5268 g), glucose (0,5186 g), ácido cítrico (0,5034g), ácido málico (0,5170g) e ácido ascórbico (0,5127g).



Figura 1 – Ultrassons Elma Transsonic 460 e sistema de filtração Phenomenex, acoplado a uma bomba de vácuo da Laboport, usando filtros de nylon (0,2 µm da Whatman).

2.3.2. Equipamento HPLC

O cromatógrafo utilizado foi da marca Varian, constituído por uma bomba Prostar 220, um injetor manual Rheodyne modelo 7725i com um loop de 20 µL, um forno Jones Chromatography modelo 7981 e uma coluna Supelcogel C-610H (30 cm de comprimento e 7,8 mm de diâmetro interno). Para a análise foram usados dois detetores colocados em série; o detetor índice de refração (IR) da Varian modelo RI-4 (à temperatura de 45 °C, para medição dos açúcares: sacarose, glucose e frutose) e o detetor de UV da Varian modelo 9050 (comprimento de onda ajustado a 210 nm) para a detecção e quantificação dos ácidos orgânicos (ácido cítrico, ácido málico e ácido ascórbico).

2.3.3. Preparação das amostras

A análise de açúcares e ácidos orgânicos foi realizada em 16 amostras de águas minerais com sabor. Cada corrida cromatográfica demorou 30 minutos e foi efetuada com repetições até se obter resultados concordantes (desvio médio percentual inferior a 5%). Os dados foram recolhidos, registados e tratados através do software star Chromatography Workstation, versão 6,4.

As amostras com gás foram submetidas a um processo prévio de degaseificação utilizando um banho de ultrassons (Elma Transsonic 460/H) (Figura 1), durante cerca de 10 minutos. A quantificação dos açúcares e ácidos orgânicos foi efetuada diretamente nas amostras. As amostras cujos cromatogramas mostravam picos saturados (amostras 1-3, 5, 9, 12-15, 19 e 34) foram, posteriormente, diluídas 10 vezes (medição de 1,00 mL para um balão de 10,00 mL). As amostras e as soluções diluídas das amostras foram filtradas com o auxílio de microfiltros descartáveis (nylon 0,2 µm da Whatman) e seringas para dentro de vials de 2 mL e, posteriormente, colocados à temperatura de 5 °C até análise por HPLC.

2.4. Medição do pH

O valor de pH é medido por potenciometria através da medição da diferença de potencial nas soluções entre um eletrodo indicador e um eletrodo de referência (Harris, 1995). Para as medições usou-se um milivoltímetro de marca Hanna Instruments, pH 211 Microprocessor pHMeter, com um eletrodo de pH combinado, constituído por um só corpo com um eletrodo indicador membrana de vidro e um eletrodo de referência prata/cloreto de prata (Figura 2).

Para se obter leituras diretas de pH foi necessário efetuar a calibração prévia do eletrodo, utilizando soluções tampão de pH 4,00, 7,00 e 9,00 a 20 °C (Panreac). As amostras foram medidas diretamente exceto, as gaseificadas que sofreram uma prévia degaseificação por ultrassons (Elma Transsonic 460/H) (Figura1). As análises foram repetidas até se obterem 3 resultados concordantes (desvio padrão relativo percentual inferior a 5%).



Figura 2 – Milivoltímetro usado na análise do pH em águas minerais

2.5. Medição da condutividade

A condutimetria avalia a capacidade que uma solução tem em conduzir corrente elétrica, devido à migração dos iões sob a influência de um gradiente de potencial existente entre dois eletrodos de platina imersos numa solução (Jeffery *et al.*, 1992).

Esta relaciona-se com a concentração total de substâncias iônicas e com a temperatura a que a medição é feita. Assim quanto maior for a quantidade de iões (cálcio, magnésio, potássio, sódio, carbonatos, carbonetos, sulfatos, cloretos, etc) dissolvidos, maior será a condutividade elétrica da água.

As análises de condutividade foram efetuadas com um equipamento de análise de multiparâmetros de marca Consort C861 multi-parameter analyser com uma célula condutivimétrica (Figura 3) que foi calibrada com 3 soluções de KCl cuja relação entre concentrações e condutividades é: 0,001 mol/L equivale a 146,9 $\mu\text{S cm}^{-1}$; 0,01mol/L equivale a 1413 $\mu\text{S cm}^{-1}$; 0,1 mol/L equivale a 12,880 mS cm^{-1} . Após calibração, a condutividade de cada amostra foi medida diretamente (as soluções gaseificadas sofreram uma prévia desgaseificação em ultrassons Elma Transsonic 460/H) (figura 1-A) com repetições até se obterem resultados concordantes (desvio padrão relativo percentual inferior a 5%).



Figura 3 – Condutivímetro usado na análise da condutividade em águas minerais

2.6 Língua eletrónica

A língua eletrónica utilizada neste trabalho é constituída por dois sistemas de multi-sensores e um elétrodo de referência Ag/AgCl de dupla junção. A língua eletrónica estava ligada a um datalogger (Agilent 34970) e a aquisição do sinal de cada membrana foi efetuada através de um computador usando o software “Agilent BenchLink Data Logger” (Figura 4). As análises foram efetuadas em soluções sob agitação magnética (VELP Scientifica, modelo ARE Heating magnetic stirrer). Cada sistema de multi-sensores era constituído por 20 membranas lipídicas diferentes, de sensibilidade cruzada, preparadas com vários compostos lipídicos (3,0%), plastificantes (65,0%) e com o polímero PVC (32,0%). O segundo sistema corresponde a uma duplicação do primeiro sistema de multi-sensores.

Todas as amostras foram analisadas a 20 °C (temperatura ambiente), num copo de plástico colocado sobre um agitador magnético da marca Velp Scientifica. Os sensores da língua eletrónica estavam mergulhados diretamente nas amostras e a análise era dada por terminada após um período de estabilização de 10 minutos. As amostras que continham gás foram previamente desgaseificadas no ultrassons da Elma Transsonic 460/H (Figura 1) durante aproximadamente 5 minutos, antes de serem analisadas. Após a leitura de cada amostra era feita uma lavagem cuidadosa do sistema analítico com água desionizada e enxaguamento leve com papel absorvente, antes de ser mergulhado na amostra seguinte. O sistema permite a monitorização contínua dos sensores, pelo que foi possível verificar visualmente a variação do sinal em função do tempo de cada uma das 40 membranas poliméricas.

No final da análise todos os sinais dos sensores obtidos durante os 10 minutos da análise eram guardados em ficheiro Excel, seleccionando-se somente o último resultado de cada sensor como representativo do perfil obtido na análise da amostra. Por fim, os perfis de sinais das amostras foram recolhidos numa base de dados para tratamento posterior com técnicas de análise estatística multivariada.

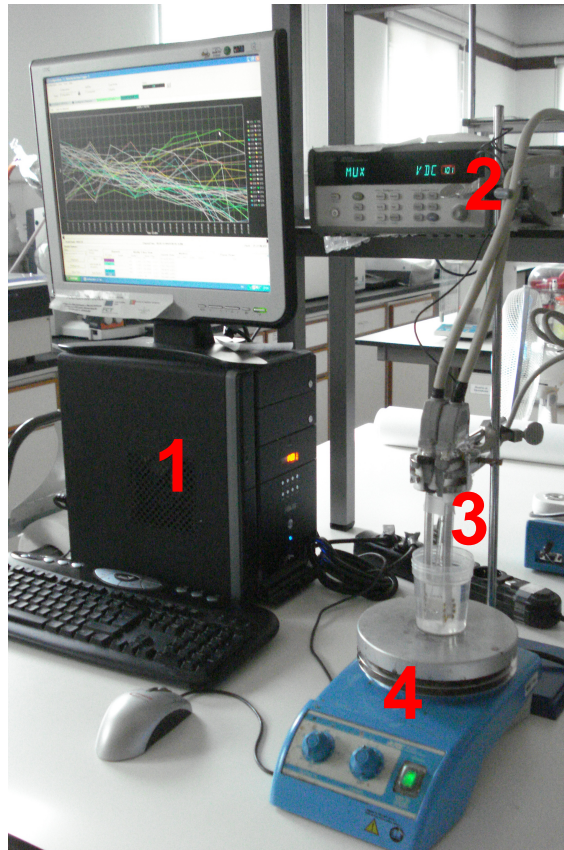


Figura 4 – Equipamento para a análise das águas minerais com a língua eletrónica: 1- PC para aquisição de dados; 2- DataLogger Agilent; 3- língua electrónica; 4- Agitador magnético.

2.8 Métodos estatísticos

Os métodos estatísticos de análise multivariada permitem a redução do número de variáveis independentes inicialmente avaliadas num dado estudo, proceder a análises qualitativas (classificação e discriminação) e quantitativas (obtenção de modelos de calibração e previsão). (Neto *et, al*, 2006). Neste trabalho, foram utilizados diferentes métodos quimiométricos.

2.8.1. Análise de componentes principais (ACP)

A análise de componentes principais (ACP) é um método não supervisionado, porque os grupos não são pré estabelecidos. É uma técnica multivariada muito utilizada pois permite a análise exploratória, transformando um conjunto de variáveis correlacionadas num conjunto de variáveis independentes entre si (ortogonais que se chamam de componentes principais, CP) que são combinações lineares das variáveis

originais. As componentes principais são calculadas por ordem decrescente, ou seja, a primeira função está orientada de forma a explicar a maior variância possível dos dados, a segunda função (ortogonal da primeira) explica parte da variância por explicar e, assim sucessivamente (Pestana e Gajairo, 2005).

Descrita desta forma, a ACP é geralmente considerada um método de redução de variáveis (Marôco, 2005). Os coeficientes (*loadings*) das combinações lineares dão uma indicação da contribuição de cada variável original para uma determinada componente e as novas coordenadas das amostras no novo sistema de eixos de CP são denominadas de *scores* (Marôco, 2005). Geralmente, a visualização da projeção das amostras com as novas funções usando os primeiros CP permitem a identificação de possíveis *outliers* (elementos que fogem da tendência geral estabelecida pelos restantes elementos do grupo) e grupos de indivíduos (*clusters*) de uma forma intuitiva, bem como, a identificação de variáveis que permitem a discriminação entre os grupos (Marôco, 2005).

2.8.2. Classificação com ADL

A análise discriminante linear (ADL) é um método de análise supervisionada pois os grupos são pré estabelecidos. É uma técnica aplicável quando a variável dependente é qualitativa (identificação dos grupos) e as variáveis independentes são quantitativas, e permite obter funções lineares de variáveis que potenciam o distanciamento entre grupos. Com o modelo ADL é possível classificar a que grupo pertencem novas amostras. Esta metodologia é sensível à multi-colinearidade entre as variáveis dependentes embora, em último caso, a robustez do modelo deve ser avaliada ao nível da previsão. Por isso, e de forma a evitar problemas de multi-colinearidade, a seleção das variáveis é um aspeto muito importante para a obtenção de modelos de previsão robustos. Também permite, em muitas situações, reduzir o ruído dos dados analíticos, por eliminar variáveis sem informação relevante para o estudo.

Os sinais dos sensores da língua eletrónica apresentam elevada colinearidade e, por isso, é fundamental usar um algoritmo de seleção de variáveis que permita a escolha dos sensores mais relevantes para cada estudo. Nesta tentativa verificou-se os resultados ao nível da percentagem de classificações corretas de modelos com

subconjuntos de sensores contendo entre 2 e 33 sensores (o número máximo de sensores deverá ser sempre inferior ao número total de amostras utilizadas, que neste estudo foram 34). Usou-se o algoritmo arrefecimento simulado, algoritmo meta-heurístico de seleção de subconjuntos de variáveis por busca local (corresponde a um ótimo global para um dado critério de aproximação a uma determinada função), para obter cada subconjunto e após 10 000 tentativas para obter esse modelo (valor que permitia chegar à mesma solução com repetição do procedimento). O critério de qualidade usado para avaliar o ajuste entre a variável dependente (grupos) usando os subconjuntos de sensores estabelecidos pelos algoritmos foi o critério ccr12, índice equivalente à maximização da primeira raiz de Roy (valor conceptualmente equivalente à razão F na análise de variância univariada) (Cadima *et al.*, 2004; Cadima *et al.*, 2012). Os modelos obtidos para cada subconjunto de sensores (2 a 33 sensores) foram testados na sua capacidade de classificação das amostras nos grupos pré-definidos, quer com dados originais quer com a metodologia validação cruzada. Neste estudo usou-se a técnica “deixar um de fora” (*leave-one-out*, LOO). Esta técnica avalia o desempenho de um modelo na previsão ao remover uma amostra do estudo contendo n amostras e prever a sua classificação com o modelo ADL obtido com as restantes (n-1). Ao repetir este processo n vezes, obtém-se os erros de classificação de todas as amostras, sendo o erro global a soma dos erros de cada teste dividido por n.

O melhor modelo de discriminação será o que permite ter modelos com menor número de variáveis, mais simples e mais fácil de interpretar, e o que permite ter melhores resultados de previsão.

2.8.3 Quantificação com RLM

A regressão linear múltipla (RLM) permite obter um modelo de previsão que representa a relação entre uma variável dependente e uma combinação de várias variáveis independentes. Esta técnica envolve uma série de pressupostos para a sua aplicação: verificação da linearidade dos dados experimentais, independência dos erros, homogeneidade das variâncias dos resíduos, normalidade de distribuição dos erros e a ausência de multi-colinearidade entre as variáveis independentes (Marôco, 2005). Embora o cumprimento dos pressupostos seja uma mais-valia para a aplicabilidade do

modelo, verifica-se que é possível ultrapassar algumas destas questões desde que se verifique a robustez do modelo ao nível da previsão. Tal como na análise discriminante linear, a multi-colinearidade presente nos sinais dos sensores é um problema. Esta situação também pode ser minimizada recorrendo ao algoritmo de arrefecimento simulado, usando como critério de qualidade o coeficiente de determinação (R^2) (Cadima *et al.*, 2004; Cadima *et al.*, 2012). Os modelos RLM obtidos para cada subconjunto de sensores (2 a 33 sensores) selecionados pelo algoritmo foram avaliados ao nível da previsão quantitativa da variável dependente quer com dados originais quer para a validação cruzada usando a técnica LOO. O melhor modelo é também aquele que possui o menor número de sensores e que permite obter um coeficiente de determinação para os dados da validação cruzada LOO mais elevado. Este modelo é testado na sua capacidade de previsão através da relação linear entre os valores previstos pela validação cruzada LOO no estabelecimento do modelo e os valores experimentais, através dos valores do declive e ordenada na origem (bem como, os respetivos intervalos de confiança), além do coeficiente de determinação ajustado.

Os dados obtidos experimentalmente neste trabalho foram tratados estatisticamente com o programa de estatística R (versão 2.15.1).

Os resultados, tratamento de dados e conclusões obtidas com base nos dados experimentais resultantes da análise por pH, condutividade, HPLC e língua eletrónica, estão apresentados no capítulo seguinte.

3. RESULTADOS E DISCUSSÃO

O trabalho experimental incidiu no estudo analítico de águas minerais com e sem sabor com uma língua eletrónica. O trabalho foi realizado na Escola Superior Agrária do Instituto Politécnico de Bragança - Laboratório de Química e Bioquímica Aplicada no âmbito do projeto de Línguas Eletrónicas para Sensorização Ambiental e Alimentar. A seguir apresenta-se os resultados e discussão da análise do pH e condutividade das 34 amostras, análise por HPLC (açúcares e ácidos orgânicos) às águas minerais de sabor e a análise com a língua eletrónica.

3.1. Análises físico-químicas

3.1.1. Medição do pH

Os valores de pH das 34 amostras de água analisadas variaram entre 2,85 e 7,42; com valor médio de 5,07 e desvio padrão de 1,5 indicativo de que a amostragem apresenta grande variabilidade. Esta variabilidade deve-se à presença de amostras com sabor, que apresentam valores de pH a variar entre 2,85 e 4,33, com exceção das amostras 7 e 8 (pH igual a 6,44 e 6,28, respetivamente). Segundo o DL nº 93/89 de 28 de Março, as águas com sabor, identificadas como refrigerantes, podem apresentar um valor de pH mínimo de 2,5. Os resultados obtidos mostram que as amostras de refrigerantes estão de acordo com os valores estabelecidos pela legislação.

As águas minerais com e sem gás apresentam um intervalo de valores de pH entre 5,45 e 7,42. Segundo o estabelecido pelo DL nº156/98 de 6 de junho e DL nº 306/2007, que define as características das águas minerais naturais e de nascente, os resultados das amostras estudadas encontram-se dentro dos limites definidos ($\text{pH} \leq 9$ e $\leq 9,5$), o que indica serem próprias para o consumo.

Num primeiro estudo, efetuou-se uma análise de variância (ANOVA) a um fator para verificar se a variabilidade nos valores de pH e condutividade poderia ser atribuída à existência de 4 grupos de amostras (águas minerais com e sem gás, e águas de sabor com e sem gás) ou se, esta variação se devia simplesmente a dois grupos de amostras (águas minerais com e sem sabor). Ao nível dos valores de pH, a ANOVA a um fator,

considerando 4 grupos de amostras, mostrou que havia diferenças significativas entre os grupos (valor-p <0,001). O teste Tukey HSD mostrou que não havia diferenças significativas (valor-p > 0.05) entre amostras de água mineral sem gás (pH médio de 6,30) e com gás (pH médio de 6,05); pelo contrário, entre as amostras de águas de sabor com gás (pH médio de 4,43) e sem gás (pH médio de 3,15), bem como entre as amostras de águas minerais e as águas com sabor, as diferenças são significativas (valores de p =0.003 e <0,001, respetivamente).

Na Figura 5 mostra um gráfico de densidade associado aos valores de pH onde parcialmente se pode averiguar a razão destes resultados, verificando-se que só as amostras de águas de sabor sem gás é que não mostram qualquer sobreposição com as águas minerais sem sabor.

3.1.2. Condutividade

Os valores analíticos obtidos globalmente mostram uma variação entre 30,3 e 2980 $\mu\text{S}/\text{cm}$, tendo-se obtido o valor médio e desvio padrão de 1137,38 e 843,96 $\mu\text{S}/\text{cm}$, respetivamente. De acordo com o DL nº 306/2007 de 27 de Agosto, as águas para o consumo devem obter valor paramétrico de condutividade correspondente a 2,5 mS/cm a 20°C. Para os refrigerantes a legislação não estabelece um valor para a condutividade, pelo que o seu valor depende da composição química de cada uma das amostras.

O estudo da ANOVA a 1 fator mostrou que havia diferenças significativas entre os 4 grupos (valor-p <0,001). Para verificar quais os grupos que tinham médias diferentes, aplicou-se o teste Tukey HSD. Entre os grupos de amostras de águas minerais sem gás (condutividade média de 0,287mS/cm) e com gás (condutividade média de 1,382mS/cm), bem como entre as águas de sabor com gás (condutividade média de 1,726mS/cm) e sem gás (condutividade média de 0,788mS/cm) há diferenças significativas (valor-p =0,002). Obtiveram-se diferenças significativas entre as médias dos grupos:

- a) águas de sabor com gás e águas minerais sem gás (valor-p =0,003);
- b) águas de sabor com gás e águas de sabor sem gás (valor-p =0,047);
- c) águas minerais com gás e águas minerais sem gás (valor-p =0,016).

Na Figura 5 é possível verificar a variabilidade dos resultados para a condutividade e perceber que também se verifica uma sobreposição de valores entre os 4 grupos, correspondendo às conclusões obtidas pela ANOVA e teste Tukey HSD, em cima referidas.

Os valores de pH e condutividade obtidos neste trabalho estão dentro dos parâmetros exigidos por lei, sendo que o conjunto de amostras analisadas apresentam grande variabilidade nos resultados devido as diferenças existente entre elas.

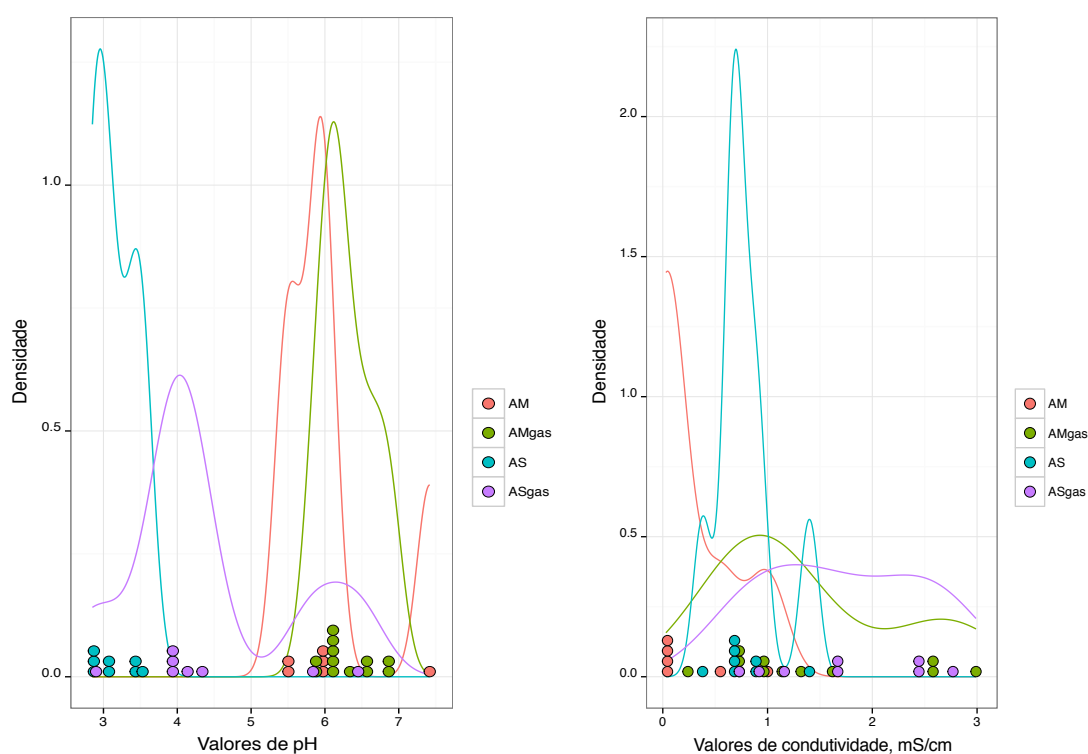


Figura 5 - Variação dos valores de pH e condutividade das amostras estudadas.

A seguir mostram-se os resultados de análise dos açúcares (glucose, frutose e sacarose) e ácidos orgânicos (ácidos cítrico, málico, ácido ascórbico) obtidos por HPLC.

3.2. Análise por HPLC

A análise por HPLC foi realizada em 16 amostras de águas de sabor com e sem gás. O método analítico usado foi testado e parcialmente validado num trabalho anterior (Sequeira, 2012) ao nível da precisão (repetibilidade e precisão intermédia) e

exatidão usando várias soluções padrão de controlo de qualidade. Neste trabalho, antes de efetuar as análises às amostras procedeu-se a uma confirmação rápida do seu desempenho analítico. As concentrações das soluções padrão de calibração variavam entre 0,3 e 5,0 g/L para a glucose, frutose, sacarose, ácido cítrico e ácido málico. O ácido ascórbico tinha concentrações a variar de 0,2 a 1,0 g/L (Sequeira, 2012). O método analítico cromatográfico foi também testado com uma solução controlo de qualidade, também preparada com mistura dos compostos estudados.

A calibração obtida mostrou resultados comparáveis ao estudo anterior e no estudo com a solução de controlo de qualidade verificou-se que a precisão e exatidão eram aceitáveis (desvio padrão relativo percentual e erro relativo percentual inferiores a 5%). Também foram feitas correções ao nível das concentrações dos compostos considerando que, no detetor (IR), a frutose é afetada pela presença do ácido málico (UV), tal como foi referido por Sequeira, 2012. Na Tabela 2 apresentam-se as concentrações médias obtidas da análise às amostras, calculadas a partir de 3 concentrações concordantes, cujo erro experimental é inferior a 5% (desvio padrão relativo percentual), o que significa que os resultados são aceitáveis.

Com base nos resultados obtidos (Tabela 2) verificou-se que nas amostras 7 e 8, nenhum dos compostos analisados foi detetado, o que indica a ausência ou presença dos compostos em quantidades muito reduzidas (inferiores ao limite de deteção da técnica utilizada). Supõe-se que nestas amostras só foram adicionadas substâncias aromatizantes. Nas restantes amostras, os açúcares analisados foram todos detetados e quantificados com exceção da sacarose nas amostras 14 e 15. Os níveis de sacarose variaram entre 1,22 e 20,94 g/L (valores das amostras 3 e 12, respetivamente); a glucose apresentou níveis de concentração a variar entre 0,08 e 21,46 g/L (valores das amostras 6 e 9, respetivamente); e os da frutose entre 0,29 e 22,29 g/L (valores das amostras 5 e 9, respetivamente).

Os resultados obtidos dos ácidos orgânicos mostram que o ácido cítrico estava presente em todas amostras com valores a variarem entre 0,04 e 2,90 g/L (valores das amostras 14 e 4, respetivamente); o ácido málico foi detetado em 14 amostras e quantificado em 9 amostras, com uma concentração a variar entre 0,05 e 0,64 g/L (valores das amostras 1 e 14, respetivamente); o ácido ascórbico foi detetado e quantificado em uma única amostra (amostra 6) com valor correspondente a 0,08 g/L.

A grande variabilidade dos resultados ao nível das concentrações mostra a variabilidade das amostras recolhidas. No entanto estas quantidades estão dentro dos parâmetros de qualidade para este tipo de bebidas. Globalmente, os açúcares apresentam as concentrações mais elevadas (em média variam entre $\approx 8,4$ - $11,3$ g/L) e com grande variabilidade nos valores para diferentes amostras (desvios padrões relativos percentuais superiores a 79%). O ácido cítrico é o ácido com concentrações mais elevadas (concentração média de 1,5 g/L e desvio padrão relativo percentual de 66%).

Tabela 2 - Concentrações médias obtidas da análise das amostras por HPLC.

Amostras	Sacarose g/L	Glucose g/L	Frutose g/L	Ácido Cítrico g/L	Ácido Málico g/L	Ácido Ascórbico g/L
1	7,87	4,88	3,35	2,66	0,05	Nd
2	20,94	12,47	7,94	2,64	0,05	Nd
3	1,22	0,72	0,93	2,44	d	Nd
4	1,57	1,24	1,45	2,90	d	Nd
5	14,15	0,62	0,29	2,79	d	Nd
6	1,53	0,08	0,31	1,09	d	0,08
7	Nd	Nd	Nd	Nd	Nd	Nd
8	Nd	Nd	Nd	Nd	Nd	Nd
9	19,74	21,46	22,29	1,11	0,28	Nd
10	19,72	17,09	11,37	1,12	0,09	Nd
11	1,36	1,26	7,88	0,61	0,23	Nd
12	21,05	13,92	15,19	0,64	0,35	Nd
13	4,90	3,09	2,19	0,42	d	Nd
14	Nd	19,26	16,23	0,04	0,64	Nd
15	Nd	17,81	14,82	1,62	0,08	Nd
34	20,94	14,89	13,96	1,01	0,18	Nd

d: detetado; Nd: não detetado

3.3. Análise exploratória dos dados químicos e físico-químicos

Para verificar como as amostras de bebidas com sabor se agrupariam atendendo aos dados químicos e físico-químicos, aplicou-se uma análise de componentes principais. A matriz de dados incluiu os valores de pH, condutividade, concentrações de açúcares (glucose, frutose e sacarose) e ácidos orgânicos (ácido cítrico e ácido málico). Nas amostras que apresentavam indicação de não detetado foi considerado o nível zero de concentração. Os valores de cada variável foram centrados e normalizados de forma a diminuir a influência das diferentes ordens de grandeza. Não se considerou neste conjunto de dados a concentração de ácido ascórbico por ter sido detetado e quantificado somente numa amostra. Na Figura 6 mostra-se a representação espacial dos dois primeiros componentes principais (PC1 e PC2) para as amostras de águas de sabor considerando se eram com ou sem gás (distinguidas pela cor dos pontos) e indicando, como legenda, o sabor de cada amostra (gráfico A) ou as respectivas marcas (gráfico B).

O método de ACP selecionou sete componentes principais que explicavam a variância total dos dados. Os dois primeiros CPs explicam cumulativamente 77,4% da variância total dos dados. A nova variável PC1 explica 51,0% e a PC2 explica 26,4% da variância total dos dados. A representação gráfica destes dois componentes principais (Figura 6) permite a visualização da distribuição espacial dos dados originais (*scores*), bem como, a influência das variáveis originais nessa distribuição (*loadings*). Na Figura 6 mostram-se 2 gráficos da mesma representação (PC2 *versus* PC1) salientando-se, em ambos os gráficos, as bebidas de sabor com gás com pontos de cor verde e as sem gás com pontos de cor laranja. Como legenda de cada ponto, no primeiro gráfico (A) indica-se o sabor de cada amostra e no segundo gráfico (B), as respectivas menções.

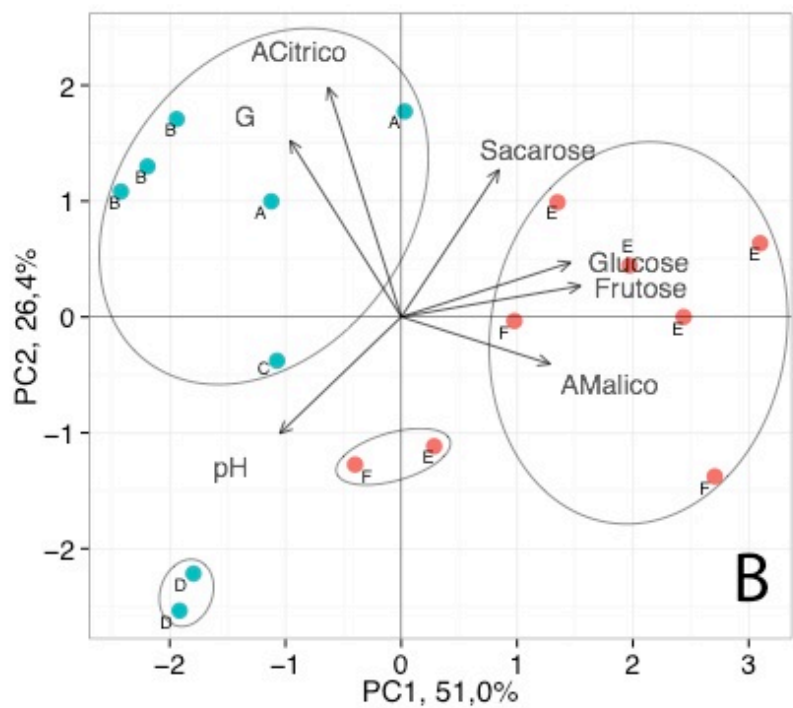
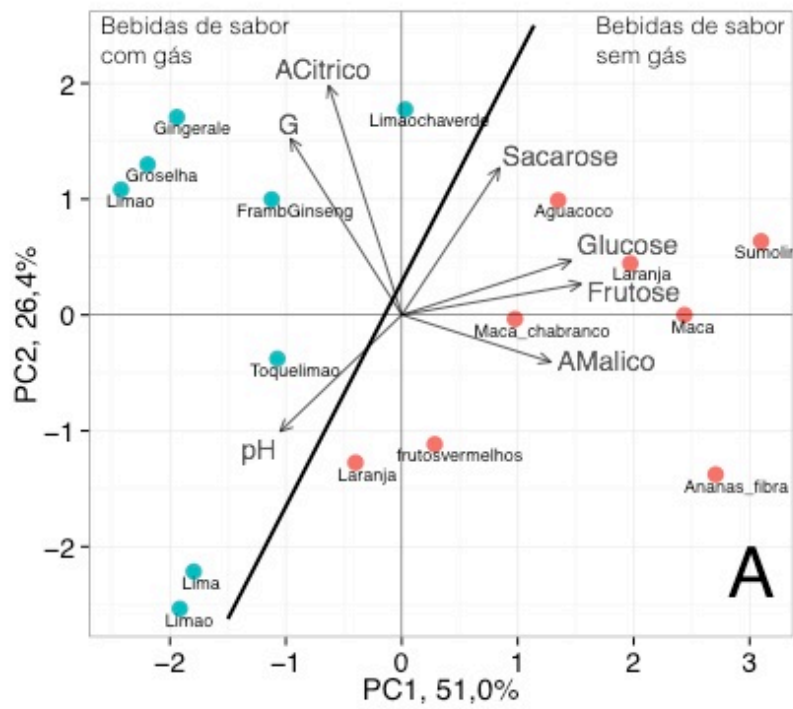


Figura 6 – Análise de componentes principais (PC1 e PC2) dos dados químicos e físico-químicos obtidos da análise às amostras de sabor com gás (pontos de cor azul) ou sem gás (pontos de cor laranja). No gráfico A indica-se o sabor de cada amostras e no gráfico B, as respetivas marcas.

3.4 – Análise exploratória dos dados da análise com a língua eletrónica

As amostras foram analisadas com a língua eletrónica e os perfis de sinais de potenciais (V) obtidos estão apresentados na Figura 7. Em geral, os perfis dos sinais obtidos mostraram ser semelhantes para todas as amostras, indicando que os sensores responderam da mesma forma.

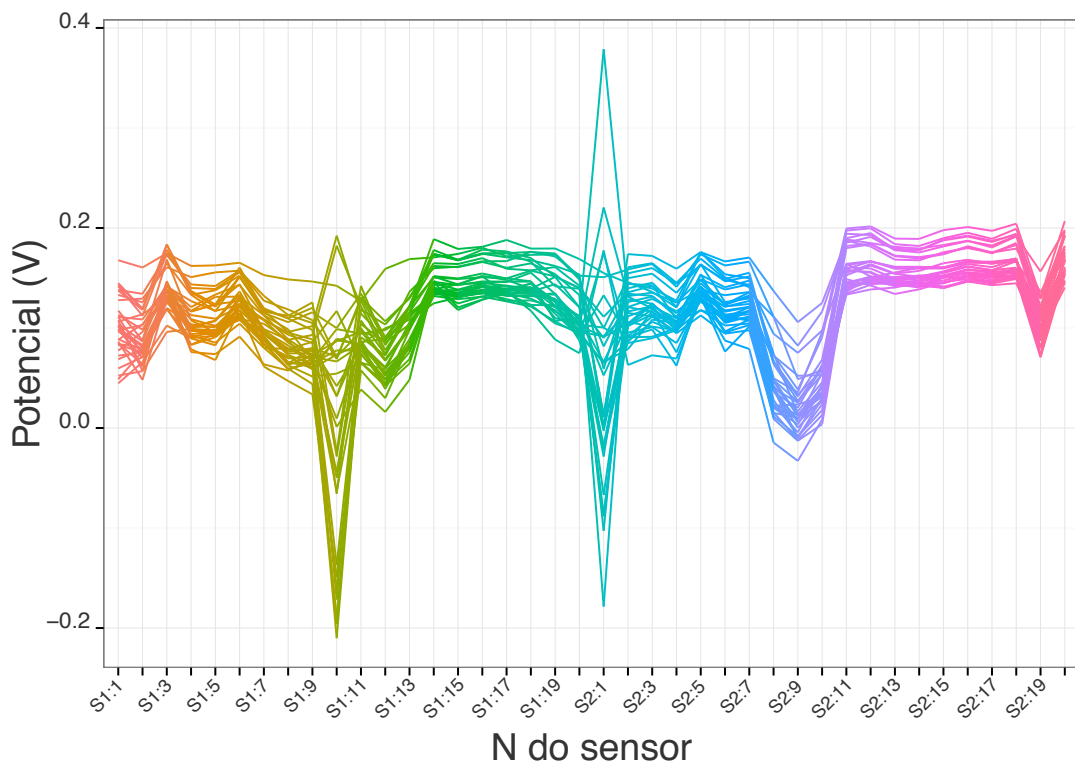


Figura 7 - Perfil dos sinais obtidos da análise das amostras com a língua eletrónica.

Na Figura 8 mostram-se os valores de potencial para cada sensor na forma de *Boxplot*, podendo-se verificar como os dados das amostras variam e a presença de extremos. Destes extremos, só a amostra 9 é que mostrou alguma prevalência por estar presente em 10 sensores, pelo que os restantes são amostras pontuais. Pode-se inferir que, os extremos são valores de sinal máximos ou mínimos para diferentes amostras e, por isso, resultam da resposta do sensor à matriz da amostra em questão. Considerando este aspeto, as amostras não foram removidas nos tratamentos a seguir apresentados.

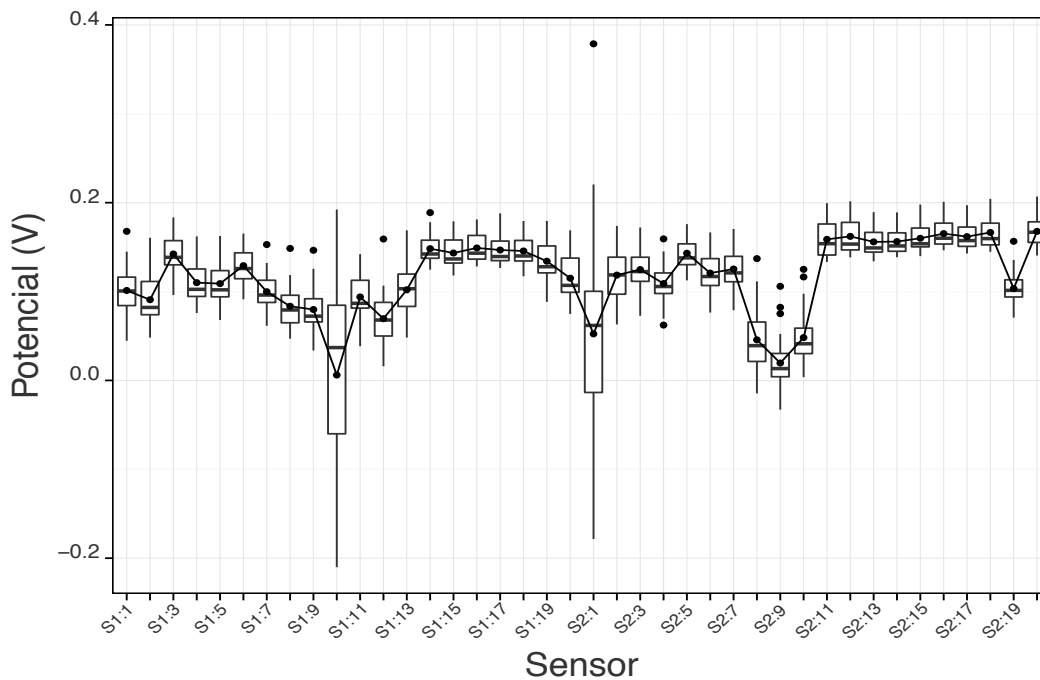


Figura 8 - Extremos quartis dos sinais de potencial (V) em função dos sensores.

Para averiguar a variabilidade dos perfis de sinais nas amostras analisadas, os valores de cada sensor da língua eletrônica foram centrados e normalizados e posteriormente aplicou-se o método não supervisionado, a análise de componentes principais. Deste tratamento multivariado obtiveram-se 28 componentes principais que explicam 100% da variabilidade dos dados. Verificou-se que os dois primeiros componentes principais permitem explicar 82,4% da variabilidade dos dados (PC1 = 70,5% e PC2=11,9%).

Na Figura 9 encontram-se representadas graficamente os dois primeiros componentes principais onde, ao nível da cor dos pontos se evidenciam 4 grupos de amostras (água mineral natural, água mineral de nascente, água de sabor com gás e água de sabor sem gás) e ao nível das legendas dos pontos, a indicação do sabor (gráfico A), o valor de pH (gráfico B) e o valor de condutividade (gráfico C) de cada amostra. Nesta figura é possível verificar que as águas de sabor com gás estão agrupadas e separadas das restantes amostras mas, não se evidenciou outra tendência de agrupamento. As amostras de sabor sem gás sobrepõem-se parcialmente às outras amostras sem sabor.

Com o objetivo de explicar a distribuição das amostras nos gráficos da Figura 10, averiguou-se a possibilidade de existir uma correlação entre os dados da primeira componente principal e os valores de pH e condutividade das amostras. Os resultados mostram que não há uma evidente tendência linear ou curvilínea entre os dados PC1 e os valores da condutividade ou de pH. No entanto verificou-se que era possível considerar dois grupos de amostras de acordo com os valores de pH e, no caso das condutividades, uma tendência de agrupamento de amostras com valores similares. Esta última situação é verificada também ao nível das marcas, o que se associa ao facto das bebidas da mesma marca terem semelhante conteúdo iónico. Em geral, no quadrante negativo do PC1 situam-se as amostras com $\text{pH} < 4$ e no quadrante positivo, as que têm $\text{pH} > 4$. No caso das águas de sabor efetuou-se também o estudo de correlação com os dados físicos-químicos e químicos tendo-se verificado que havia uma forte correlação com as concentrações de frutose ($R=0,847$ para um ajuste polinomial de 2º grau).

3.5 – Análise qualitativa com a língua eletrónica

Para verificar se os perfis de sinais da língua eletrónica permitiam classificar as amostras em quatro grupos estabeleceu-se uma metodologia de tratamento de dados em dois passos com o objetivo de selecionar o melhor modelo: menor número de sensores e 100% de classificações corretas para a validação cruzada. Para a classificação supervisionada usou-se a análise discriminante linear cujo desempenho de previsão é inferior quando as variáveis independentes são multi-colineares (em geral, com $R > 0,8$) (Vanuza e Filzmoser, 2009). No caso dos sinais da língua eletrónica existe elevada multicolinearidade pelo que, os modelos devem ser obtidos por seleção das variáveis independentes mais importantes (sensores) para evitar/reduzir este problema. Os modelos devem ser posteriormente avaliados atendendo à sua robustez ao nível das previsões, através da percentagem de classificações corretas obtidas nas amostras usando a validação cruzada. De salientar que num modelo de estimação pode ser relevante colocar todas as variáveis de forma a maximizar o ajuste para os dados originais, permitindo ter resíduos menores e coeficientes de correlação de ajuste maiores. Pelo contrário, num modelo de previsão, o interesse é maximizar o

desempenho de previsão para os dados originais que, neste trabalho, é efetuado com redução das variáveis independentes para evitar efeitos de sobre ajustamento e avaliação de desempenho de previsão através da validação cruzada usando a técnica “leave-one-out”.

Na seleção de variáveis aplicou-se o algoritmo meta-heurístico arrefecimento simulado (AS) para obter o melhor modelo para cada um dos subconjuntos de sensores definido que, neste trabalho foi de 2 a 33 sensores. O algoritmo verifica, em 10000 tentativas, qual o melhor modelo para cada subconjunto de sensores. Os modelos selecionados para cada subconjunto de sensores são posteriormente testados na sua capacidade de classificação ao nível dos dados originais e ao nível da validação cruzada usando a técnica “leave-one-out”. Na Figura 10 mostram-se os valores de % de classificações corretas para os dados originais (DOcorreta) e para a validação cruzada “leave-one-out” (LOOcorreta) obtidos por análise discriminante linear em cada modelo selecionado pelo algoritmo AS.

O modelo discriminante linear com o menor número possível de sensores (modelo mais simples) e com o maior valor de classificações corretas corresponde ao de 24 sensores que permite ter 100% de classificações corretas (Figura 10). Os sensores selecionados foram: S1:1, S1:3, S1:4, S1:6, S1:9 a S1:11, S1:15 a S1:18, S1:20, S2:2, S2:5, S2:6, 2 S2:8, S2:11, S2:13, S2:14, S2:16 a S2:20. O modelo com 34 sensores não foi possível de obter pois todas as combinações apresentavam multi-colinearidade entre variáveis. O modelo com 1 sensor não foi considerado nesta seleção devido ao fato de os usados neste trabalho serem de sensibilidade cruzada e, por isso, só terem aplicabilidade analítica quando usados em conjunto e não isolados. O modelo discriminante linear selecionado tem 3 funções que explicam 100% da variabilidade dos dados originais sendo a primeira função discriminante suficiente para discriminar as amostras dos 4 grupos estabelecidos (primeira função discriminante explica 99,94% da variância dos dados; segunda e terceira funções explicam 0,04 e 0,01%, respetivamente). Na Figura 11 mostra-se a representação gráfica das duas primeiras funções discriminantes na separação dos 4 grupos de amostras. As elipses de cada grupo estão associadas ao intervalo de confiança a 95%, assumindo uma distribuição normal multivariada.

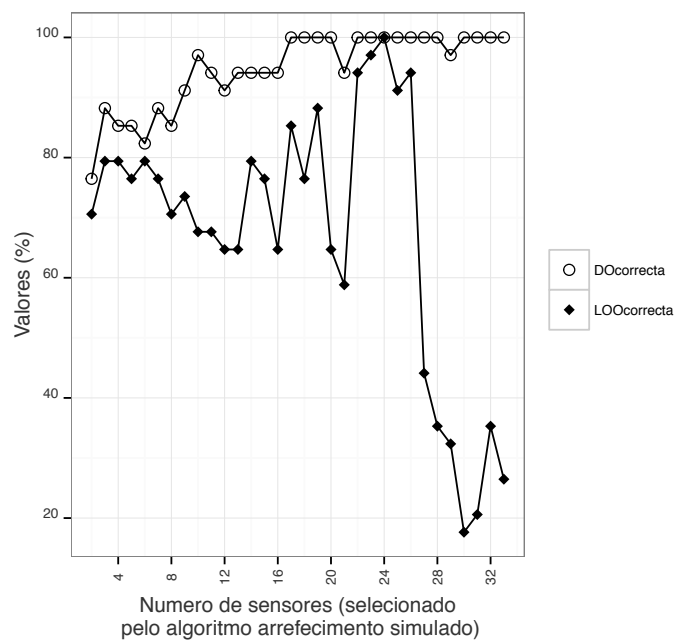


Figura 10 - Percentagens de classificações corretas por análise discriminante linear para os dados originais (DOcorrecta) e para a validação cruzada “leave-one-out” (LOOcorrecta) de cada modelo selecionado pelo algoritmo AS.

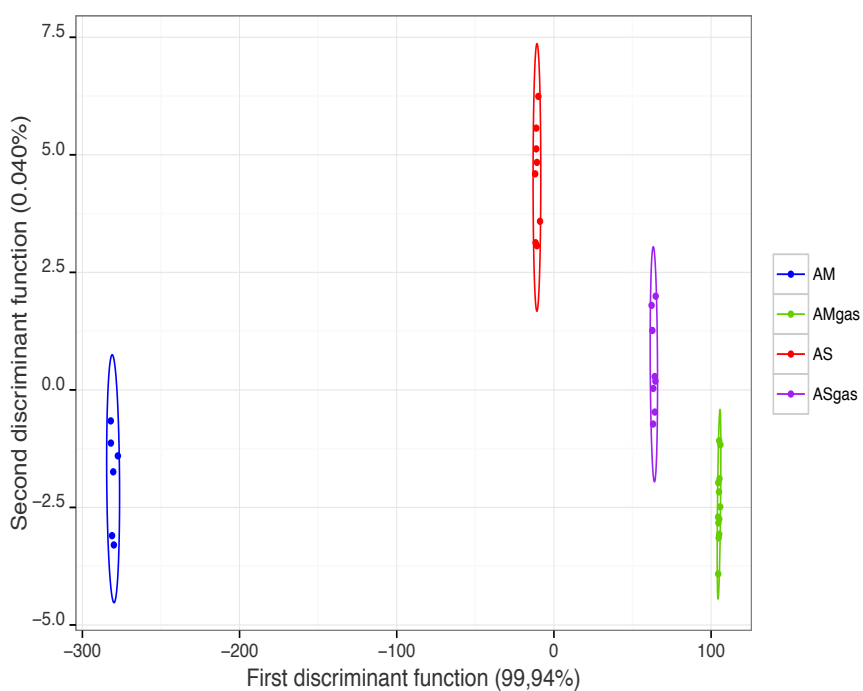


Figura 11 - Análise discriminante linear entre 4 grupos de águas: água mineral sem gás (AM), água mineral com gás (AMgas), água de sabor sem gás (AS) e água de sabor com gás (ASgas). As elipses de cada grupo estão associadas ao intervalo de confiança a 95%, assumindo uma distribuição normal multivariada.

Estes resultados mostram que a língua eletrónica usada permite extrair informação das diferentes amostras que possibilitam a análise qualitativa ao nível da classificação em quatro grupos que foram estabelecidos em função dos rótulos. Ao nível da ordem de separação de grupos, verificou-se que, da esquerda para a direita, no quadrante negativo do PC1, o grupo das amostras de águas minerais (AM) é o mais afastado, seguido do grupo de águas minerais de sabor sem gás, na zona central de PC1. No quadrante positivo do PC1, próximo da zona central, apresenta-se o grupo de águas de sabor com gás e mais afastado, o grupo de águas minerais com gás.

3.6 – Análise quantitativa com a língua eletrónica

O estudo realizado evidenciou ainda que os perfis de sinais obtidos com os sensores potenciométricos poderiam ser correlacionados com os valores de pH e condutividade das amostras. Por isso, procurou-se verificar se era possível obter um modelo de regressão linear múltipla de previsão destes parâmetros usando os sensores da língua eletrónica. O procedimento seguido é semelhante ao descrito no subcapítulo anterior, de análise qualitativa com a língua eletrónica, onde, tal como para a análise discriminante linear, a regressão linear múltipla é sensível à multi-colinearidade dos sinais dos sensores. Por isso, também foi necessário selecionar o melhor modelo através da seleção das variáveis independentes usando o algoritmo arrefecimento simulado e posteriormente, testar a sua robustez ao nível das previsões, através da validação cruzada com a técnica “leave-one-out”.

No estabelecimento do modelo de previsão do valor de pH em águas minerais com e sem sabor aplicou-se o algoritmo meta-heurístico arrefecimento simulado (AAS) para obter o melhor modelo, através de 10000 tentativas, para cada um dos subconjuntos de sensores a variar entre 2 a 33 sensores. Como critérios de seleção, o modelo deve ter o menor valor de erro RMSE (raiz do erro médio quadrático) e o maior valor de coeficiente de determinação (R^2). NA Figura 12 mostram-se os valores de RMSE e R^2 de cada modelo de regressão linear múltipla usando cada subconjunto de sensores selecionados pelo algoritmo arrefecimento simulado para prever valores de pH. O modelo com 33 sensores não foi considerado por não ter sido possível selecionar um subconjunto sem problemas de multi-colinearidade.

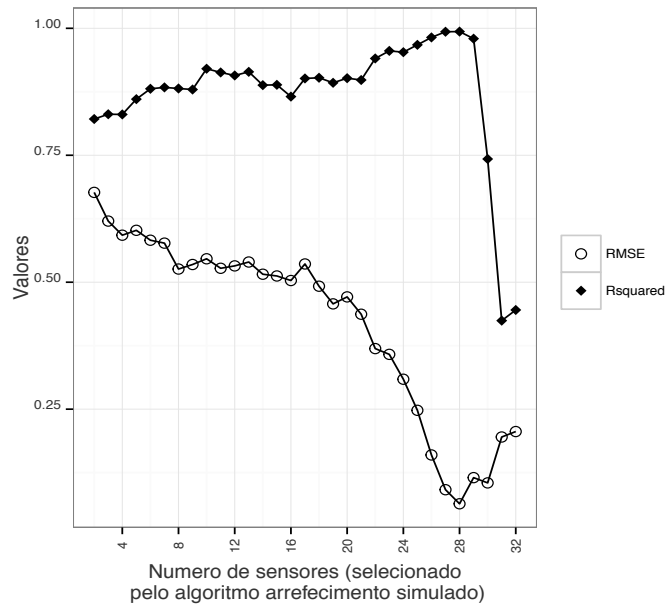


Figura 12 - Valores de RMSE e R^2 de cada modelo de regressão linear múltipla usando subconjunto de sensores selecionados pelo algoritmo arrefecimento simulado para prever valores de pH.

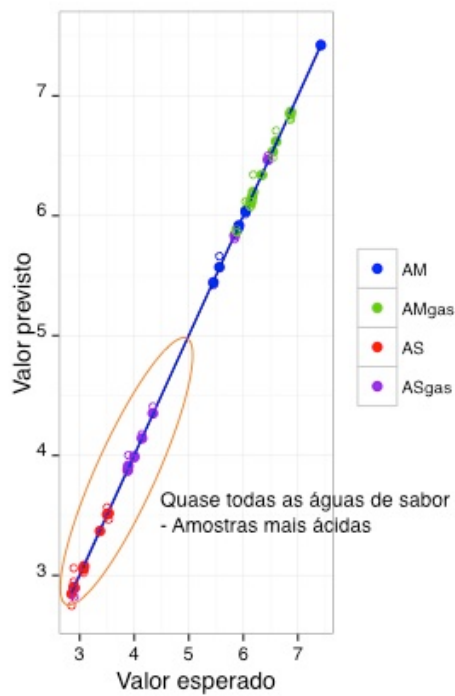


Figura 13 - Valores de pH previstos pelo modelo regressão linear múltipla selecionado, quer para os dados originais quer os da validação cruzada, em função dos valores experimentais.

O modelo selecionado contém 27 sensores (RMSE=0.06366 e $R^2=0.9936$): S1:1, S1:5, S1:6, S1:10 a S1:12, S1:14, S1:16 a S1:20, S2:1, S2:2, S2:5 a S2:9, S2:13 a S2:20.

A capacidade de previsão do modelo pode ser visualizada na Figura 13, onde se representam os valores de pH previstos pelo modelo selecionado, quer para os dados originais quer os da validação cruzada, em função dos valores experimentais.

Neste gráfico, o declive tem o valor de 0,9955 e ordenada na origem de 0,03217. Estatisticamente podem ser considerados iguais aos teóricos (declive igual a um e ordenada na origem igual a zero) pois estes valores estão incluídos nos intervalos de confiança ao nível de 95%: $0,9857 < \text{declive} < 1,005$; $-0,01951 < \text{ordenada na origem} < 0,08384$.

Tratamento semelhante foi efetuado para obter o modelo de regressão linear múltipla de previsão do valor da condutividade. O modelo selecionado entre os subconjuntos de sensores a variar entre 2 e 33 sensores selecionados pelo algoritmo arrefecimento simulado foi o de 28 sensores: S1:1 a S1:9, S1:11 a S1:14, S1:16, S2:4 a S2:6, S2:9 a S2:15, S2:17 a S2:20. O modelo apresentou valores de $RMSE=0.02568$ e $R^2=0.9990$.

Na Figura 14 apresentam-se os valores de RMSE e R^2 de cada modelo, para prever a condutividade das amostras, obtidos por regressão linear múltipla usando cada subconjunto de sensores selecionados pelo algoritmo arrefecimento simulado. Também não foi possível obter o modelo com 33 sensores por problemas de multi-colinearidade.

Na Figura 15 mostram-se os valores de condutividade previstos pelo modelo selecionado, quer para os dados originais quer os da validação cruzada, em função dos valores experimentais. A capacidade de previsão do modelo é evidente tendo-se obtido um declive com valor de 0,9905 e ordenada na origem de 0.007895. Também, estatisticamente, os valores teóricos do declive e ordenada na origem estão inseridos nos intervalos de confiança ao nível de 95%: $0.9855 < \text{declive} < 1.022$; $-0.03678 < \text{ordenada na origem} < 0.05256$.

Os resultados obtidos para os valores de pH e condutividade mostraram que a língua eletrónica usada tem sensores com sensibilidade ao nível de concentrações de sais na matriz das amostras que permitem extrair informação associada aos valores de pH e das condutividades das amostras. Considerando bom ajuste entre os valores previstos e os valores experimentais de pH e condutividade, a língua eletrónica serve também de confirmação que não foi identificado qualquer “outlier” nestes valores experimentais.

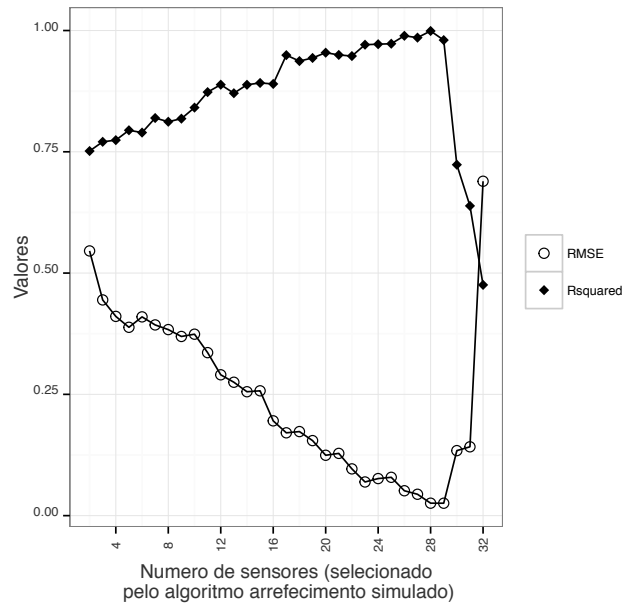


Figura 14 - Valores de RMSE e R^2 de cada modelo de regressão linear múltipla usando subconjunto de sensores selecionados pelo algoritmo arrefecimento simulado para prever valores de condutividade.

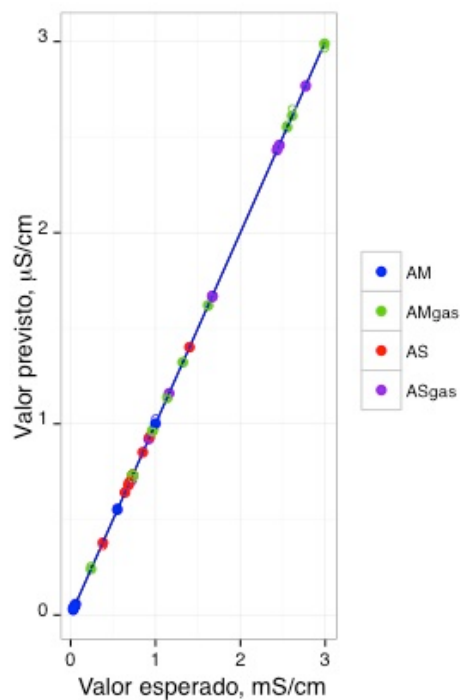


Figura 15 - Valores de condutividade previstos pelo modelo regressão linear múltipla selecionado, quer para os dados originais quer os da validação cruzada, em função dos valores experimentais.

4. CONCLUSÕES

A caracterização físico-química das 34 amostras minerais recolhidas envolveu análises de pH e condutividade de todas as amostras e a análise por cromatografia líquida das amostras de água com sabor (16 amostras) para determinar conteúdos em ácidos orgânicos (ácidos cítrico, málico e ascórbico) e açúcares (glucose, frutose e sacarose). Num primeiro estudo sobre os dados analíticos obtidos verificou-se se as amostras poderiam ser agrupadas em 4 grupos de acordo com a informação constante nos rótulos: águas minerais com e sem gás e águas de sabor com e sem gás. A análise de variâncias mostrou que havia diferenças significativas entre os quatro grupos considerando os dados físico-química das amostras. A análise de componentes principais usando os parâmetros químicos obtidos por HPLC mostrou que os dados analíticos separavam naturalmente as amostras de água de sabor com e sem gás.

Numa segunda fase, usaram-se os perfis de sinais obtidos com a língua eletrónica e, no âmbito da análise qualitativa, aplicou-se a análise de componentes principais para verificar a distribuição espacial das amostras e a interdependência com os parâmetros físico-químicos determinados; posteriormente, aplicou-se análise discriminante linear acoplada ao algoritmo de arrefecimento simulado para seleção de variáveis e estabelecer o melhor modelo discriminante para os 4 grupos definidos. O desempenho do modelo obtido ao nível da precisão foi avaliado usando a validação cruzada “leave-one-out”. Verificou-se que com uma só função discriminante, contendo 24 sensores, que permitia obter 100% de classificações corretas, por validação cruzada. Neste estudo verificou-se ainda que a informação contida nos sinais obtidos da língua eletrónica podia ser correlacionada com os valores de pH e condutividade e, por isso, estabeleceram-se modelos, por regressão linear múltipla, de previsão de resultados de pH e condutividade nas amostras. Como a regressão linear múltipla é sensível à multicolinearidade dos sinais dos sensores também, foi necessário selecionar o melhor modelo através da seleção das variáveis independentes usando o algoritmo arrefecimento simulado. Os modelos foram testados na robustez de previsão, através da validação cruzada com a técnica “leave-one-out”. Os melhores modelos de previsão de pH e condutividade incluíam 27 e 28 sensores respetivamente, mostrando que a informação estava presente nos perfis dos sinais. Verificou-se que na relação linear

entre os valores previstos pelo modelo selecionado e os valores experimentais, quer para a variável pH quer para a condutividade, obtiveram-se valores de declive e ordenada na origem que, estatisticamente, podem ser considerados os teóricos (declive igual a 1 e ordenada na origem igual a zero).

Os resultados mostraram que a Língua Eletrónica pode ser uma ferramenta analítica prática na análise de águas minerais com e sem sabores, ao nível do controlo analítico de classificação ou discriminação de amostras, bem como, da análise de pH e condutividade das amostras.

BIBLIOGRAFIA

- Ashurst, P. R. *Chemistry and Technology of Soft drinks and Fruit Juices*, Blackwell Publishing, 2^o Edition, Hereford, UK, 2005;
- Beullens, K., Mészáros, P., Vermeir, S., Kirsanov, D., Legin, A., Buysens, W., Nicolai, B.M. & Lammertyn, J. Analysis of tomato taste using two types of electronic tongues, *Sensors and Actuators: B. Chemical* 131 (2008) 10-17.
- Burattia, S., Benedettia, S., Scampicchioa, M. & Pangerod, E. C. Characterization and classification of Italian Barbera wines by using an electronic nose and an amperometric electronic tongue, *Analytica Chimica Acta* 525 (2004) 133-139.
- Cadima, J., Cerdeira, J. O. & Minhoto, M., *Computational aspects of algorithms for variable selection in the context of principal components*, *Computational Statistics & Data Analysis* 47 (2004) 225-236.
- Cadima, J., Cerdeira, J. O., Silva, P. D. & Minhoto, M., *The subselect R package*, 2012 (<http://cran.rproject.org/web/packages/subselect/vignettes/subselect.pdf>)
- Ciosek, P. & Wróblewski, W. The recognition of beer with flow-through sensor array based on miniaturized solid-state electrodes, *Talanta* 69 (2006) 1156–1161.
- Decreto-Lei nº 306/2007, Diário da república de 27 de Agosto.
- Decreto-Lei nº 93/89, Diário da república de 28 de Março.
- Decreto-Lei nº156/98, Diário da República de 6 de junho.
- Dias, L. A., Peres, A. M., Vilas-Boas, M., Rocha, M. A., Estevinho, L. & Machado, A. A. S. C. An electronic tongue for honey classification, *Microchimica Acta* 163 (2008) 97-102.
- Dias, L.G., Peres, A.M., Barcelos, T.P., Sá Morais, J., Machado, A.A.S.C. Semi-quantitative and quantitative analysis of soft drinks using an electronic tongue, *Sensors & Actuators: B. Chemical* 154 (2011) 111-118.
- Dias, M. E. B. C. de Sousa, *Língua electrónica potenciométrica: uma ferramenta para análise da qualidade do mel*, Mestrado em Química. Departamento de Química e Bioquímica da Faculdade de Ciências, Universidade do Porto, 2013.
- Eckert, C., Lutz, C., Breikreutz, J. & Woertz, K. *Quality control of oral herbal products by an electronic tongue— Case study on sage lozenges*, *Sensors and Actuators B* 156 (2011) 204-212.

- Escuder-Gilaberta, L. & Peris, M. Review: Highlights in recent applications of electronic tongues in food analysis, *Analytica Chimica Acta* 665 (2010) 15-25.
- Gallardo, J., Alegret, S. & del Valle, M. Application of a potentiometric electronic tongue as a classification tool in food analysis, *Talanta* 66 (2005) 1303-1309.
- Garcia-Breijoa, E., Atkinson, J., Gil-Sanchez, L., Masot, R., Ibanez, J., Garrigues, J., Glanc, M., Laguarda-Miro, N. & Olguin, C., A comparison study of pattern recognition algorithms implemented on a microcontroller for use in an electronic tongue for monitoring drinking waters, *Sensors and Actuators A* 172 (2011) 570–582.
- Graham, J. M., *Problems with stepwise procedures in discriminant analysis*. Paper presented at the annual meeting of the Southwest Educational Research Association, New Orleans, LA (2001).
- Harris, D., et al.; Quantitative Chemical Analysis, Fourth Edition, *W.H. Freeman and Company new york, 1995*.
- Heyden, Y., V., Popovici, S. T., & Schoenmakers, P. J. *Evaluation of size – exclusion chromatography and size exclusion electrochromatography calibration curves*, Vol. 957, 2 (2002) 127-137;
- Jeffery, G., et al.; *Análise Química Quantitativa*, 5ª Edição, LTC-Livros técnicos e científicos Editora, S.A., Rio de Janeiro, 1992.
- Legin, A., Rudnitskaya, A., Vlasov, Yu, Di Natale, C., Mazzone, E. & D’Amico, A. Application of electronic tongue for qualitative and quantitative analysis of complex liquid media, *Sensors and Actuators: B. Chemical* 65 (2000) 232-234
- Legin, A., Rudnitskaya, A., Lvova, L., Vlasov, Yu., Di Natale, C., & D’Amico, A. Evaluation of Italian wine by the electronic tongue: recognition, quantitative analysis and correlation with human sensory perception, *Analytica Chimica Acta* 484 (2003) 33-44.
- Lima, A. C. S., Afonso., J. C. A *Química do refrigerante*, vol.31, 3, Agosto 2009.
- Marôco, J *Análise Estatística, com utilização de SPSS*, 3ª Edição Sílabo, Lisboa, Portugal, 2005.
- Martínez-Máñez, R., Soto, J., Garcia-Breijoa, E., Gil, L., Ibáñez, J. & Llobet, E. An “electronic tongue” design for the qualitative analysis of natural waters, *Sensors and Actuators B* 104 (2005) 302–307.

- McGarigal, K. Cushman, S. & Stafford, S. *Multivariate Statistics for Wildlife and Ecology Research*, Springer-Verlag New York, Inc, 2000.
- Neto, B. B., Scaiminio, I. S. & Bruns, R. E. 25 Anos de quimiometria no Brasil, *Quim. Nova* (2006) 1401-1406.
- Parada, M. J. C., *Monitorização ecotoxicológica da qualidade da água da Ribeira do Portelo (NE de Portugal)*. Dissertação para obtenção do grau de mestre em tecnologia ambiental, Instituto Politécnico de Bragança, 2012.
- Peres, A. M., Dias, L. G., Veloso, A. C. A., Meirinho, S. G., Sá Morais, J. & Machado, A.A.S.C An electronic tongue for gliadins semi-quantitative detection in foodstuffs, *Talanta* 83 (2011) 857-864.
- Pestana, M. E. & Gageiro, J. N. *Análise de dados para ciências sociais "A complementaridade do SPSS"*, Edição Sílabo, Lisboa, Portugal, 2005.4ªEdição.
- Resende, A. *Análises a águas de consumo humano. Análises Microbiológica, Clássicas e Instrumentais*. Trabalho de fim de curso de Engenharia Biotecnológica, Instituto Politécnico de Bragança, 2006;
- Rudnitskaya, A., Schmidtke, M. L., Delgadillo, I., Legin, A. & Scollaryd, G. Study of the influence of micro-oxygenation and oak chip maceration on wine composition using an electronic tongue and chemical analysis, *Anal. Chim. Acta* 642 (2009) 235.
- Sequeira, C. B. *Análises de açúcares e ácidos em sumos comerciais: aplicação da língua eletrónica*. Dissertação para obtenção do grau de mestre, em qualidade e segurança alimentar, Instituto Politécnico de Bragança, 2012.
- Silva, B. M., Seabra, R. M., Andrade, P. B., Oliveira, M. B. & Ferreira, M. A. Adulteração por adição de açúcares a sumos de frutos: *uma revisão*, *Cienc. Tecnol. Aliment*, Vol. 2, 4 (1999) 184-191;
- Steen, D. P. & Ashurst, P.R. *Carbonated soft Drinks: Formulation and manufacture*, Blackwell publishing, 2006.
- Toko, K. Taste sensor, *Sensors and Actuators: B. Chemical* 64 (2000) 205-215.
- Ulloa, P. A., Guerra, R., Cavaco, A. M., R. da Costa, A. M., Figueira, A. C. & Brigas, A. F. Determination of the botanical origin of honey by sensor fusion of impedance e-tongue and optical spectroscopy, *Computers and Electronics in Agriculture* 94 (2013) 1-11.

- Vanuza, K. & Filzmoser, P., Introduction to multivariate statistical analysis in chemometrics, CRC Press (Taylor & Francis), Boca Raton, FL, USA, 2009.
- Vlasov, Yu Legin, A, Rudnitskaya, A. Natale, & D'Amico, C.A. Nonspecific sensor arrays ("electronic tongue") for chemical analysis of liquids, *IUPAC Technical Report, Pure and Applied Chemistry* 77 (2005) 1965-1983.
- Wei, Z., Wan, J. & Liao, W. Technique potential for classification of honey by electronic tongue, *Journal of Food Engineering* 94 (2009) 260-266.
- Yúfera, E. P. *Química de los alimentos*, editorial síntese, Madrid, Espanha, 1998.

WEBSITES

- W1 - <http://www.APIAM.pt> acessado em 9/10/14).
- W2 - <http://probeb.pt/conteudo/> acessado em 10/10/2014.
- W3 - www.aguadaspedras.com , acessado em 03-09-2014.